REPEATED GAMES WITH PARTIAL MONITORING: THE STOCHASTIC SIGNALING CASE

JOHN HILLAS AND MIN LIU

ABSTRACT. In this work we extend a result of Lehrer characterizing the correlated equilibrium payoffs in undiscounted two player repeated games with partial monitoring to the case in which the signals are permitted to be stochastic. In particular we develop appropriate versions of Lehrer's concepts of "indistinguishable" and "more informative." We also show that any payoff associated with a (correlated) distribution on strategy vectors in the stage game such that neither player can profitably deviate from one of his strategies to another that is indistinguishable and more informative is the payoff of a correlated equilibrium of the supergame.

CONTENTS

Date: December, 1995.

We are most grateful to Jean-François Mertens who initially suggested the question examined in this paper to Min Liu and who told us the ideas behind much of the analysis. We are also grateful to Abraham Neyman and Sylvain Sorin for comments during various stages of this work.

1. INTRODUCTION

The folk theorem tells us that in a repeated game with perfect monitoring any feasible and individually rational payoff is an equilibrium payoff as long as the players are sufficiently patient. The original and cleanest version of this result deals with undiscounted games.

Without perfect monitoring the result is not true. One may think of the case in which no player ever learns anything about how the others played as an extreme example. In a series of papers Lehrer (1990, 1991, 1992a, 1992b, 1992c) gave some characterization of the equilibrium payoffs in undiscounted games with partial monitoring. Like the folk theorem (most of) Lehrer's results characterize the equilibrium payoffs in the repeated game in terms of easily calculated aspects of the stage game. Lehrer's work deals with both two player games and n-player games. In this paper we will deal with only the two player case.

The cleanest of Lehrer's results deals with correlated equilibria. There is a clear intuition as to why this should be the case. With partial monitoring each player receives at the end of each stage some signal that depends on the actions taken in that stage. In general different players will receive different signals. Thus from the next stage on an equilibrium of the original game will generate not a Nash equilibrium but rather a correlated equilibrium.

Lehrer considers the set of (correlated) distributions on vectors of moves such that no player could gain by deviating in an undetectable way. He shows that any payoff associated with such a distribution is achievable as a correlated equilibrium of the supergame.

There are two aspects of the notion of a deviation being undetectable. The first is that if player 1 changing from action s to action t would lead to a different signal to player 2 then that deviation would be detectable. And even if this was only true if player 2 was taking some particular action one could think of player 2 as occasionally taking that action to "check" on player 1. Lehrer calls two actions of player 1 that generate the same signal for player 2 whatever player 2 does "indistinguishable." One might think that a deviation that involved deviating from some action to an action that was indistinguishable from it could not be detected. This however is not true.

The problem is that while two actions may generate the same signal for the other player they may not lead to the same information for the player taking the action. Thus player 1 might get perfect information about the action taken by player 2 if he takes action s and no information if he takes action t. Then player 2 could check if player 1 was playing s when he was expected to by requiring player 1 to take actions leading to different signals to player 2 depending on what player 1 had learned about what player 2 had done. If player 1 had deviated to t he would be unable to respond appropriately.

Thus Lehrer defines also a notion of one strategy being "more informative" than another. A deviation from action s to t is undetectable if t is both indistinguishable from s and more informative than s. Of course to make the argument in the previous paragraph required player 2 to be able to tell if player 1 was responding appropriately. Thus there must be actions by player 1 that lead to different signals to player 2. Lehrer called a game such that each player had some possibility to communicate with the other through his choice of action one with a "nontrivial signaling structure," and his results deal with games with such a signaling structure.

In this paper we extend one of Lehrer's results. In Lehrer's models the signals are deterministic. That is, for given actions of the players the signals to the players are determined. We extend the analysis to allow the signals to be stochastic. This introduces some complications and requires some statistical computations but the idea is in its essential features no different than that developed by Lehrer. Much of the statistical computation is done using Blackwell's approachability theorem.

2. PRELIMINARIES

We start by recalling the definition of a repeated game. The notation and definitions follow Sorin (1990) and Mertens, Sorin, and Zamir (1994).

We let $I = \{1,2\}$ denote the set of players, S^i for $i = 1,2$ the finite set of actions for player i in the stage game, and $g : S = S^1 \times S^2 \to \mathbb{R}^2$ the payoff function. We assume that g is normalized to take values between 0 and 1. We denote by G the normal form game defined by I, S^i $(i = 1,2)$ and g; by X^i $(i = 1,2)$ the set of mixed strategies of player i, that is, probabilities on S^i, with $X = X^1 \times X^2$; and by g also the multilinear extension of the payoff function to X.

The supergame Γ with stage game G is played as follows: At stage $n+1$, each player observes a signal a^i_n from some finite set A^i, jointly generated (perhaps stochasticly) by s_n the vector of actions of the players in the previous stage. We assume perfect recall, that is, that each player remembers the signals he received in the past and the actions he took. For notational convenience we assume that a^i_n reveals, at least, s^i_n. Thus player i makes his decision at stage $n+1$ based on the sequence of signals $\{a^i_1, a^i_2, \ldots, a^i_n\}$. Each player i, at stage $n+1$, chooses a move s^i_{n+1} and the signals a_{n+1} are generated, and each player i is informed of a^i_{n+1}. The game continues to stage $n+2$. Finally the above description is commonly known.

A play in the game Γ is an infinite sequence $(s_1, g_1, a_1, s_2, \ldots, a_{n-1}, s_n, g_n, a_n, \ldots)$ where $g_n = g(s_n)$, and a_n is the realization of the signal generated by the action s_n. The set of all plays is denoted by H_∞. The initial part of a play ending in stage n (that is, a finite sequence $(s_1, g_1, a_1, s_2, \ldots, a_{n-1}, s_n, g_n, a_n)$) is called an n-history and the set of all n-histories is H_n. The set of all histories is $H = \cup_n H_n$.

We let $\mathcal{H}_n$ be the σ-algebra generated by H_n, $\mathcal{H}_\infty$ the product σ-algebra $\vee_n \mathcal{H}_n$, and $\mathcal{H}$ the induced σ-algebra on H. Similarly we let H^i_n be the set of sequences $(a^i_1, \ldots, a^i_n)$—recall that we assume that at each stage a^i reveals s^i. And we let $\mathcal{H}^i_n$ be the σ-algebra on H_∞ generated by H^i_n and $\mathcal{H}^i$ be player i's information partition on H generated by the restriction of each $\mathcal{H}^i_n$ to H_n.

A pure strategy for player i is a $\mathcal{H}^i$-measurable function from H to S^i. A behavior strategy for player i is a $\mathcal{H}^i$-measurable function from H to X^i. And a mixed strategy is a probability distribution over pure strategies.

A vector of behavior strategies σ defines a probability P_σ on $(H_\infty, \mathcal{H}_\infty)$. We let $\overline{\gamma}_n(\sigma) = E_\sigma(\frac{1}{n}\sum_{m=1}^n g_m)$ denote the expected average payoff for the first n stages under σ. For some such strategies these averages may not converge in the usual sense. Thus, in order to be able to define equilibria in the usual sense we use the notion of a Banach limit. For any Banach limit $\mathcal{L}$ and for any pair of strategies this defines the payoff to that pair of strategies and with the payoffs well defined the concept of equilibrium is defined in the usual way. We call such equilibria $\mathcal{L}$-equilibria.

Definition 1. Let ℓ^∞ denote the space of all bounded sequences of real numbers. A linear functional $F : \ell^\infty \to \mathbb{R}$ is called a *Banach limit* if $F(\{\xi_n\}) \leq \limsup(\{\xi_n\})$ for all $\{\xi_n\}$ in ℓ^∞ and $\eta_n = \xi_{n+1}$ for all n implies that $F(\{\eta_n\}) = F(\{\xi_n\})$.

We also define a somewhat stronger notion of equilibrium than $\mathcal{L}$-equilibrium, that of uniform equilibria.

Definition 2. The strategy σ is a *uniform equilibrium* if $\overline{\gamma}^i_n(\sigma)$ converges to some $\overline{\gamma}^i(\sigma)$ and for all $\varepsilon > 0$, there exists N such for all $n > N$, for each i, and for all τ^i, $\overline{\gamma}^i_n(\sigma^{-i}, \tau^i) \leq \overline{\gamma}^i(\sigma) + \varepsilon$.

And we define the notion of correlated equilibria as the equilibria of the game obtained from the original game by augmenting it with a correlation device.

Definition 3. A *correlation device c* (for the player set I) is a probability space $(E, \mathcal{E}, P)$ together with sub σ-fields $(\mathcal{E}^i)_{i\in I}$ of $\mathcal{E}$. The extension Γ_c of a game Γ by c is the game where first nature selects e from E according to P, next each player i in I is informed of the events in $\mathcal{E}^i$ which contain e, then Γ is played. A *correlated equilibrium* of Γ is a pair (c, equilibrium of Γ_c).

The previous two definitions are from Mertens, Sorin, and Zamir (1994).

3. The Model with Non-stochastic Signals

Here we develop formally the model of Lehrer of games with nonstochastic signals. After each move s, each player i observes $a^i = Q^i(s)$, where Q^i is a mapping from S to some finite set of signals A^i. Recall that we assume the game has perfect recall and that Q^i reveals i's move, that is, for any vectors of actions s and t if $s^i \neq t^i$ then $Q^i(s) \neq Q^i(t)$.

We now define what it means for a game to have a nontrivial signaling structure. We shall be concerned only with games satisfying this condition. Games for which this condition is not satisfied are quite easy to deal with, since at least one of the players never observes anything about what the other player is doing.

Definition 4 (Nontrivial signaling structure)**.** A game is said to have a *nontrivial signaling structure* if for both players $i = 1,2$ there exists s^i in S^i and s^j, t^j in S^j with $j \neq i$ such that $Q^i(s^i,s^j) \neq Q^i(s^i,t^j)$.

Next we define the notion of actions being indistinguishable. This means that the other player cannot tell them apart on the basis of the signal he receives.

Definition 5 (Indistinguishable)**.** Two strategies of player i, s^i and t^i are said to be *indistinguishable* if $Q^j(s^i,s^j) = Q^j(t^i,s^j)$ for all s^j in S^j.

And finally one action is more informative than another if taking the first action gives at least as much information about what the other has done as taking the second action. Since we shall only be concerned with situations in which actions are indistinguishable and one is more informative we shall include the requirement that the actions be indistinguishable part of the definition of more informative.

Definition 6 (More informative)**.** One strategy of player i, s^i, is said to be *more informative* than another, t^i, if s^i and t^i are indistinguishable and $Q^i(t^i,s^j) \neq Q^i(t^i,t^j)$ implies that $Q^i(s^i,s^j) \neq Q^i(s^i,t^j)$ for all s^j and t^j in S^j.

This definition means that player i always gets no more information on j's move by playing t^i than by playing s^i.

Comment 1. The indistinguishable relation is an equivalence relation and the more informative relation is a partial order. In particular, both are transitive.

Let P be the set of probabilities on S (correlated moves). The payoff function is extended to P by integration. The sets of equilibrium payoffs are characterized in terms of the sets

$$A^i = \{P \in P \mid \sum_{s^j} P(s^i,s^j) g^i(s^i,s^j) \geq \sum_{s^j} P(s^i,s^j) g^i(t^i,s^j)$$
$$\text{for all } s^i, t^i \text{ in } S^i \text{ with } t^i \text{ more informative than } s^i\}$$

We denote the set of correlated equilibrium payoffs by C_∞, and the set of individually rational payoffs by IR. Lehrer (1992a) proves the following theorem.

Theorem 1. *The set of payoffs of correlated equilibria equals* $g(A^1 \cap A^2) \cap IR$.

4. The Model with Stochastic Signals

In this section we describe a model in which the signals are stochastic. That is, the actions of the players determine the distribution of the generated signals, not the actual realization. Given the actions (s^i,s^j) of the players we let $\theta((\cdot,\cdot) \mid s^i,s^j)$ be the joint distribution on the space of signals $A = A^i \times A^j$. We also denote the derived distribution $\theta^j(\cdot \mid s^i,s^j)$ the marginal distribution on A^j and extend this to mixed strategies of player i in the obvious way as

$$\theta^j(\cdot \mid x^i,s^j) = \sum_{s^i} x^i(s^i)\theta^j(\cdot \mid s^i,s^j).$$

Definition 7 (Nontrivial signaling structure)**.** A game is said to have a *nontrivial signaling structure* if for both players $i = 1,2$ there exists s^i in S^i and s^j, t^j in S^j with $j \neq i$ such that $\theta^i(\cdot \mid s^i,s^j) \neq \theta^i(\cdot \mid s^i,t^j)$.

Logically, perhaps, we should use mixed strategies in the previous definition but it is trivially equivalent to use pure strategies. In the definitions that follow we do need to explicitly consider mixed strategies. In the situation in which the signaling is deterministic if there are no pure actions that are indistinguishable from a given action then there can be no mixed actions that are indistinguishable either. However the same is not true in the situation with stochastic signaling. Here the relevant notion of indistinguishable is that the actions produce the same distribution on the signals of the other player so that the other player could with some statistical test tell, with some degree of certainty, which of the actions was being taken. And it could well be that while the distributions on the signals of the others from two pure actions are both different from the distribution generated by a third action that some mixture of the first two actions might generate exactly the distribution generated by the third action. And since we need to consider mixed actions when defining the set of actions indistinguishable from a given action it is also necessary to define the notion of more informative for mixed actions.

Definition 8 (Indistinguishable)**.** Two mixed strategies of player i, x^i and y^i are said to be *indistinguishable* if $\theta^j(\cdot \mid x^i, s^j) = \theta^j(\cdot \mid y^i, s^j)$ for all s^j in S^j.

As in the model with non-stochastic signaling player j can't tell, on the basis of the signal he received whether player i played x^i or y^i.

We use Blackwell's (1951) idea of one experiment being more informative than another to define a notion of one strategy being more informative than another in the setting in which signals are stochastic. We shall first recall Blackwell's definition in a more general setting.

Let M be the space of parameters in which we are interested. An experiment consists of Ω a set of possible observations and $Y_m(\cdot)$ a set of probability distributions over Ω for every m in M.

Definition 9 (Blackwell, 1951)**.** The experiment Y^1 is more informative than Y^2 if there exists $L_{\omega_1}(\cdot)$ (that is, for all ω_1 in Ω_1, a probability distribution over Ω_2) such that for all m in M, for all ω in Ω_2, $Y^2_m(\omega) = \sum_{\omega_1 \in \Omega_1} Y^1_m(\omega_1) L_{\omega_1}(\omega)$.

In the definition that follows we define one action to be more informative than another if it is more informative in the sense of Blackwell about both the action the other takes and the signal he receives. (That is, the space of parameters specifies both the action taken and the signal observed by the other.) Again this differs a little from the situation with deterministic signaling. There the signal of the other was completely determined by the actions and so there was no need for an independent concern with the observation of the other.

Definition 10 (More informative)**.** One mixed strategy x^i of player i is said to be *more informative* than another y^i if x^i and y^i are indistinguishable and there exist $L_{a^i}(\cdot)$ in ΔA^i (that is, for all a^i in A^i, a probability distribution over A^i) such that for all a^j in A^j and for all s^j in S^j,

$$\theta^i(\cdot \mid y^i, a^j, s^j) = \sum_{a^i \in A^i} \theta^i(a^i \mid x^i, a^j, s^j) L_{a^i}(\cdot)$$

where $\theta^i(a^i \mid y^i, a^j, s^j) = \sum_{s^i} y^i(s^i) \theta((a^i, a^j) \mid s^i, s^j) / \sum_{s^i} y^i(s^i) \theta^j(a^j \mid s^i, s^j)$.

Comment 2. Again, the indistinguishable relation is an equivalence relation and the more informative relation is a partial order.

We now define the sets we will use to characterize the equilibrium payoffs. Let P be the set of probabilities on S (correlated moves) and extend the payoff to P by integration. Let

$$\overline{A}^i = \{P \in P \mid \sum_{s^j} P(s^i,s^j) g^i(s^i,s^j) \geq \sum_{s^j} P(s^i,s^j) g^i(y^i,s^j)$$
$$\text{for all } s^i \text{ in } S^i \text{ and } y^i \text{ in } \Delta S^i \text{ with } y^i \text{ more informative than } s^i\}$$
$$= \{P \in P \mid \sum_{s^j} P(s^i,s^j) g^i(s^i,s^j) \geq \sum_{s^j} \sum_{s^{i'}} P(s^i,s^j) g^i(s^{i'},s^j) y^i(s^{i'} \mid s^i)$$
$$\text{for all } s^i, s^{i'} \text{ in } S^i \text{ and } y^i \text{ in } \Delta S^i \text{ with } y^i \text{ more informative than } s^i\},$$

that is, the set of correlated strategies such that, when advised to play s^i, player i cannot gain by deviating to y^i in ΔS^i for any y^i more informative than s^i.

5. Preliminary Results

In this section we prove two preliminary results. Both results and both proofs are very minor modifications of results in Mertens, Sorin, and Zamir (1994). The first result is Proposition 4.5 from their Chapter IV and the proof given here is a somewhat more detailed version of their proof with some minor modifications to make clear that it does not depend on the assumption of nonstochastic signaling.

Proposition 1. *The payoff vector d is a uniform equilibrium payoff if and only if there exists a decreasing sequence ε_m converging to 0, and sequences N_m and σ_m such that σ_m is an ε_m-equilibrium in Γ_{N_m} leading to a payoff within ε_m of d.*

Proof. Necessity: Suppose that σ is a uniform equilibrium with $\overline{\gamma}^i(\sigma) = d^i$. Take any decreasing sequence ε_m converging to 0 and let σ_m=σ. Now, let N_m be such that $|\overline{\gamma}^i_{N_m}(\sigma) - \overline{\gamma}^i(\sigma)| < \varepsilon_m/2$ for $i = 1,2$ and $\overline{\gamma}^i_{N_m}(\sigma^{-i},\tau^i) \leq \overline{\gamma}^i(\sigma) + (\varepsilon_m/2)$ for $i = 1,2$ and for all τ^i. Thus $\overline{\gamma}^i(\sigma) < \overline{\gamma}^i_{N_m}(\sigma) + (\varepsilon_m/2)$. And

$$\overline{\gamma}^i_{N_m}(\sigma^{-i},\tau^i) \leq \overline{\gamma}^i(\sigma) + (\varepsilon_m/2) < \overline{\gamma}^i_{N_m}(\sigma) + (\varepsilon_m/2) + (\varepsilon_m/2) = \overline{\gamma}^i_{N_m}(\sigma) + \varepsilon_m$$

for $i = 1,2$ and for all τ^i. That is, σ is an ε_m-equilibrium in Γ_{N_m} leading to a payoff within ε_m of d.

Sufficiency: We construct σ such that σ is a uniform equilibrium of Γ. That is, $\overline{\gamma}^i_n(\sigma)$ converges to $\overline{\gamma}^i(\sigma)$, and for any ε we can find $N(\varepsilon)$ such that if $n > N(\varepsilon)$, then $\overline{\gamma}^i_n(\sigma^{-i},\tau^1) \leq \overline{\gamma}^i(\sigma) + \varepsilon$, for both i and all τ^i. This is equivalent to showing that for any ε we can find $N(\varepsilon)$ such that if $n > N(\varepsilon)$ then σ is an ε-equilibrium of Γ.

Define superblocks M_m as a sequence of l_m blocks of size N_m and let σ be play σ_m on each N_m block of M_m, that is, starting with an empty history after each cycle of N_m moves. Choose l_m such that $N_{m+1}/(l_m N_m) \leq \varepsilon_{m+1}$. It follows that if n is in N_m, σ is an ε-equilibrium of Γ_n.

The total gain from deviating in the stages up to the end of superblock M_{m-1} is at most $\sum_{k=1}^{m-1} \varepsilon_k l_k N_k$; the total gain from deviating from the end of superblock M_{m-1} to the end of block N_m, the block which is just before the one containing n, is less than or equal to $(n - \sum_{k=1}^{m-1} l_k N_k)\varepsilon_m$; and $N_m \times 1$ is an upper bound on the gain from deviating in the N_m block containing n. Thus the most that either player can gain by deviating is

$$\begin{aligned}
\frac{1}{n}[\sum_{k=1}^{m-1} \varepsilon_k l_k N_k + (n - \sum_{k=1}^{m-1} l_k N_k)\varepsilon_m + N_m \times 1]& \\
< \frac{1}{n}[\sum_{k=1}^{m-1} \varepsilon_k l_k N_k + n\varepsilon_m + N_m]& \\
< \frac{\sum_{k=1}^{m-1} \varepsilon_k l_k N_k}{\sum_{k=1}^{m-1} l_k N_k} + \varepsilon_m + \frac{N_m}{l_{m-1} N_{m-1}}& \\
\leq \frac{\sum_{k=1}^{m-1} \varepsilon_k l_k N_k}{\sum_{k=1}^{m-1} l_k N_k} + \varepsilon_m + \varepsilon_m.&
\end{aligned}$$

Thus σ is a $((\sum_{k=1}^{m-1} \varepsilon_k l_k N_k / \sum_{k=1}^{m-1} l_k N_k) + 2\varepsilon_m)$ equilibrium of Γ_n. Now

$$\begin{aligned}
\frac{\sum_{k=1}^{m-1} \varepsilon_k l_k N_k}{\sum_{k=1}^{m-1} l_k N_k} &= \frac{\sum_{k=1}^{m'} \varepsilon_k l_k N_k}{\sum_{k=1}^{m-1} l_k N_k} + \frac{\sum_{k=m'+1}^{m-1} \varepsilon_k l_k N_k}{\sum_{k=1}^{m-1} l_k N_k} \\
&\leq \frac{\sum_{k=1}^{m'} \varepsilon_k l_k N_k}{\sum_{k=1}^{m-1} l_k N_k} + \frac{\sum_{k=m'+1}^{m-1} \varepsilon_k l_k N_k}{\sum_{k=m'+1}^{m-1} l_k N_k}. \qquad (1)
\end{aligned}$$

Now let m' be such that $\varepsilon_{m'} < \varepsilon/8$ and let $N_1(\varepsilon)$ be such that if $n > N_1(\varepsilon)$ and n is in M_m, then $\sum_{k=1}^{m'} \varepsilon_k l_k N_k / \sum_{k=1}^{m-1} l_k N_k < \varepsilon/8$. Thus the first term in the inequality (1) is less than $\varepsilon/8$. Also the second term is no more than $\varepsilon_{m'+1} < \varepsilon/8$. (Recall that ε_m is a decreasing sequence.) Thus, for any $\varepsilon > 0$ if n is in N_m and $n > N_1(\varepsilon)$,

$$\frac{\sum_{k=1}^{m-1} \varepsilon_k l_k N_k}{\sum_{k=1}^{m-1} l_k N_k} + 2\varepsilon_m < 4(\varepsilon/8) = \varepsilon/2$$

and so $\vec{\gamma}_n^i(\sigma^{-i}, \tau^1) \leq \vec{\gamma}_n^i(\sigma) + (\varepsilon/2)$. And, by construction, $\vec{\gamma}_n^i(\sigma)$ converges to $\vec{\gamma}^i(\sigma)$. That is, we can find $N_2(\varepsilon)$ such that if $n > N_2(\varepsilon)$, then $|\vec{\gamma}_n^i(\sigma) - \vec{\gamma}^i(\sigma)| < \varepsilon/2$. Letting $N(\varepsilon) = \max\{N_1(\varepsilon), N_2(\varepsilon)\}$ we obtain that if $n > N(\varepsilon)$

then

$$\begin{aligned}\vec{\gamma}^i_n(\sigma^{-i},\tau^1) &\le \vec{\gamma}^i_n(\sigma)+\frac{\varepsilon}{2}\\ &< \vec{\gamma}^i(\sigma)+\frac{\varepsilon}{2}+\frac{\varepsilon}{2}\\ &= \vec{\gamma}^i(\sigma)+\varepsilon\end{aligned}$$

and so σ is a uniform equilibrium of Γ_n. Thus the proposition is proved. □

Comment 3. The result remains true if we consider equilibria of the game augmented by a correlation device. The proof of necessity is unchanged. In the proof of sufficiency the strategy we constructed behaved independently on each of the N_m blocks. Thus the needed correlation device for the M_m superblock would be the independent product l_m copies of the correlation device for the game Γ_{N_m} and the correlation device for the whole game would be the product of the devices for the superblocks. Because of the way that the equilibrium strategy was constructed it would not matter if the signals of the device were announced at the beginning of the game or at the beginning of each block.

The following lemma is a (rather trivial) extension of Lemma 4.6 of Chapter 4 of Mertens, Sorin, and Zamir (1994) to the stochastic signaling case.

Lemma 1. *In the two-person stochastic signaling supergame, given a pure strategy* σ^1*, at each history h player* 1 *can use any (mixed) action* y^1 *rather than* $\sigma^1(h)=s^1$ *with* y^1 *more informative than* s^1*, while still inducing the same probability distribution on* H^2*.*

Proof. The signal distributions to player 2 will be the same if player 1 plays y^1 instead of s^1 because the two strategies are indistinguishable. Now at the next stage since y^1 is more informative than s^1, player 1 can generate the same probability distribution over his space of signals if he had played s^1 from the signals that he actually observed when playing y^1 and play accordingly in the future. □

6. The Main Result

In this section we prove our main result, extending Lehrer's result to the model with stochastic signaling. Many of the details of the proof follow quite closely the proof of Theorem 1 given in Mertens, Sorin, and Zamir (1994).

Theorem 2. *The set of payoffs of correlated equilibria equals* $g(\bar{A}^1\cap\bar{A}^2)\cap IR$*.*

Proof. We shall show that $C_\infty \subset g(\bar{A}^1 \cap \bar{A}^2) \cap IR \subset C_\infty$. To show the first inclusion let $d = (d^1, d^2)$ be an L-equilibrium payoff not in $g(\bar{A}^1 \cap \bar{A}^2)$. (The inclusion in IR is clear, because if not, the players will always be better off by deviating to individual rational strategies contradicting the supposition that d was an equilibrium payoff.)

Now we show that any strategy leading to a payoff outside $g(\bar{A}^1 \cap \bar{A}^2)$ cannot be a correlated equilibrium. The idea is simple. From such a strategy a player can gain by replacing any action recommended to him by his most preferred (mixed) action among those that are more informative than the recommended action.

The equilibrium strategies σ together with the correlation device give, for each stage n, an induced distribution on S, which we denote P_n. We let $\bar{P}_n = \frac{1}{n}\sum_{m=1}^{n} P_m$ and $\tilde{P} = L(\bar{P}_n)$. Thus $d = g(\tilde{P})$.

Now, for any P in $\mathcal{P}$ we define P^1 in $\bar{A}^1$ that results when player 1 replaces any action by one that is most preferred among those that are more informative. First let the map $\varphi^1 : S^1 \to X^1$ be such that $\varphi^1(s^1)$ maximizes $\sum_{s^2} P(s^1, s^2) g^1(y^1, s^2)$ on the set $U = \{y^1 \in X^1 \mid y^1$ is more informative than $s^1\}$. Now let $P^1(s^1, s^2) = \sum_{t^1} \varphi^1(t^1)(s^1) P(t^1, s^2)$. (The term $\varphi^1(t^1)(s^1)$ is the weight that the mixed action $\varphi^1(t^1)$ puts on the action s^1.)

Suppose that $\tilde{P}$ is not in $\bar{A}^1$. Let $\tilde{\varphi}^1$ be the map described in the previous paragraph when $P = \tilde{P}$. Let τ^1 be obtained from σ^1, the equilibrium strategy of player 1, by using, at each stage, $\tilde{\varphi}^1(s^1)$ rather than s^1 and generating for the following stages (by adding some random element, if necessary) a signal having the same distribution as the distribution on the signals that would have resulted from using s^1. Thus from Lemma 1

$$L(\bar{\gamma}_n^1(\tau^1, \sigma^2)) = g^1(\tilde{P^1}) > g^1(\tilde{P})$$

contradicting the supposition that σ was an equilibrium.

To show the inclusion of $g(\bar{A}^1 \cap \bar{A}^2) \cap IR$ in C_∞ consider P in $\bar{A}^1 \cap \bar{A}^2$ with $g(P)$ in IR. By Proposition 1 it is enough to construct, for any $\varepsilon > 0$, an ε-equilibrium in a finite game with payoff within ε of $g(P)$. Let $\varepsilon = 5\varepsilon_1$.

We construct the finite game in the following way. First we define a block consisting of a large number of "normal" stages followed by a phase in which each randomly chooses for the other the stage about which he is to report, and then each reports the signal he observed at that stage. We can accomplish this in such a way that

1. the fraction of "normal" stages is high
2. during the "normal" stages the players play in a way that leads to the desired payoff,
3. each player communicates to the other with high accuracy the stage that the other is to report, and

4. neither player can profitable deviate without substantially changing the distribution on the signals observed by the other player.

We next consider a larger block, consisting of a large number of the previously defined blocks. Since in the smaller blocks neither player could profitably deviate without changing the distribution on the other player's signals, in the larger block each player can partition his observations into two sets so that

1. if the other player has not deviated the observation will lie, with high probability, in the first set, while
2. if the other player has deviated in a manner giving him a significant gain the observation will, with high probability, lie in the second set.

Finally we put a large number of these larger blocks together so that for almost all of those blocks there remains time to punish deviations.

We first observe that the hypothesis of nontrivial signaling means that the players can communicate with each other. Whatever information they wish to communicate can first be encoded as a binary number. Then player 2 has at least two strategies s^2 and t^2 and a strategy s^1 of player 1 with $\theta^1(\cdot \mid s^1, s^2) \neq \theta^1(\cdot \mid s^1, t^2)$. And similarly for player 1. Now if we want player 2 to communicate a binary number we let s^2 denote "0" and t^2 denote "1". For example, if we want to have player 2 communicate the signal that he observed we would label each signal in A^2 with a binary number and translate this to a sequence of s^2 and t^2. If we wanted the information communicated with high probability we would have to repeat each "1" or "0" a large number of times. (Below we see that we do wish to communicate the randomly chosen time accurately, but do not need to communicate the observation accurately.) For future use we let W be the number of stages, that is, digits, the players need to report their observed signal according to some binary code.

The strategies are defined on blocks of stages of size $N_1 = 2^n + 2nK + 2W$. For the first 2^n stages of the block the players receive a recommendation from some correlation device $\bar{R}$. The correlation device $\bar{R}$ is the independent product of 2^n copies of the probability R on $\Omega^1 \times \Omega^2 = (S^1 \cup (S^1 \times S^2)) \times (S^2 \cup (S^2 \times S^1))$. The distribution R is obtained by taking the convex combination of the uniform distribution on S (with probability η) and P (with probability $(1-\eta)$) and independently announcing with probability η to one of the players the move of his opponent. That is,

$$\begin{aligned} R(s) &= (\eta/(\#S) + (1-\eta)P(s))/(1+2\eta) \\ R(s^1, \{s^2, s^1\}) &= \eta R(s) \\ R(\{s^1, s^2\}, s^2) &= \eta R(s). \end{aligned}$$

The device operates as follows. An outcome in $(\Omega^1 \times \Omega^2)^{2^n}$ is randomly selected according to $\overline{R}$. Player i is informed of its projection on $(\Omega^i)^{2^n}$ and is asked to follow the projection of this on $(S^k)^{2^n}$. Now, at every one of the first 2^n stages each move of each player is announced with positive probability and for each pair of recommended moves there is a positive probability that player 1 will also be told player 2's recommendation. And similarly there is a positive probability that player 1 will also be told player 2's recommendation.

During the next n subsequences of K stages, during each subsequence, player 1 plays an independent mixture $(1/2, 1/2)$ on the moves (s^1, t^1) at the first stage. Once this move (s^1 or t^1) is realized, he plays the same action for the following $K-1$ stages. At the same time, player 2 plays s^2. We choose s^1, t^1 and s^2 so that $\theta^2(\cdot \mid s^1, s^2) \neq \theta^2(\cdot \mid t^1, s^2)$. For the next n subsequences we reverse the roles of players 1 and 2. These random moves are used to generate random times $\mathbf{m}^2$ and $\mathbf{m}^1$ independently and uniformly distributed on the previous 2^n stages and communicate them with a precision depending on K.

Given ε_1 let $\eta < \varepsilon_1/3$. The following claim says that we can let $K = K(\varepsilon_1, n)$ be such that each player i hears correctly $\mathbf{m}^i$ with probability at least $1-\varepsilon_1$ and let n be such that $(2nK(\varepsilon_1, n) + 2W))/2^n \leq \eta$. We leave its proof until after we have completed the main part of the proof of the theorem.

Claim 1. *For any ε_1 and n one can choose $K = K(\varepsilon_1, n)$ so that player 2 may with probability at least $1-\varepsilon_1$ choose correctly the stage that player 1 is trying to tell him. And similarly with the roles reversed. Moreover the $K(\varepsilon_1, n)$'s can be chosen so that the fraction of time spent outside the "normal" part of the block converges to zero as n goes to infinity.*

Finally during the last $2W$ stages the previously defined code is alternatively used by each player i $(i = 1, 2)$ to report the signal he got at stage $\mathbf{m}^i$ $(i = 1, 2)$. Note that neither player may know very well exactly what the report of the other player was. But since in any case the signal that the other player saw is not precisely known it does not make the proof any easier to repeat the message until it is very likely to be accurately heard. All that really matters here is that there is some distribution on the signals that player 1 will observe when player 2 is reporting and that player 2 cannot generate a similar distribution if he deviates in a way that gives him significant gains.

We now show that if a large number M of these N_1 blocks put together then each player may make a statistical test to check if the other player is deviating—or at least to check if he is deviating very often. These statistical calculations could be done directly. However a corollary of the approachability theorem gives us the statistics in almost exactly the form we need.

(See Appendix A for a brief discussion of games with vector payoffs and the approachability theorem.)

Consider an artificial game with vector payoffs with the "stages" being N_1 blocks from the original game. Let $\tilde{S}^1 = \{\vec{s}^1\}$ be the action set of player 1 where $\vec{s}^1$ means that player 1 follows the recommendation at each stage in the original game. Player 2's pure strategy set $\tilde{S}^2$ in the artificial game is the set of his pure strategies in the N_1 block in the true game (including the correlation device). We denote the equilibrium strategy of player 2 in the original game by $\vec{s}^2$.

Let $v = \#S^1 \times \#S^2 \times \#A^1 \times 2^n + \#S^1 \times \#S^2 \times \#A^1 \times (\#A^1)^W$. We now define the vector payoff of dimension $v+1$ for this game. Each choice of action by player 2 leads to a distribution over the histories in the N_1 block. (Recall that player 1 has only one action.) For the first v dimensions we describe the map from such histories to payoffs. For the final dimension we define the payoff directly on the strategies.

The first large block has $\#S^1 \times \#S^2 \times \#A^1 \times 2^n$ dimensions. In this block, there are 2^n "periods." For each "period," there are $\#S^1 \times \#S^2 \times \#A^1$ dimensions. Thus, each dimension of each "period" is indexed by a triple (s^1, s^2, a^1) with s^1 in S^1, s^2 in S^2, and a^1 in A^1. For each period $k = 1, \dots, 2^n$, if in the original game, in stage k, player 1 was recommended to play s^1 and was told that player 2's recommendation was s^2 and player 1 observed signal a^1, then we put 1 in the (s^1, s^2, a^1) dimension of period k. Otherwise we put 0.

The second large block of the vector payoff consists of $\#S^1 \times \#S^2 \times \#A^1$ parts. Each part is indexed by a particular (s^1, s^2, a^1), meaning that in the stage that player 1 randomly chose for player 2 to report in the original game, player 1 had been told both players' recommendations (s^1, s^2) and had observed a^1. Each part consists of $(\#A^1)^W$ dimensions, each indexed by a particular sequence of observed signals of player 1. If in the stage he choose for player 2 to report, player 1 received the recommendation (s^1, s^2), observed the signal a^1 and then observed a particular sequence of signals during player 2's reporting stage we put 1 in the dimension with this index. Otherwise we put 0.

The third block is the simplest. It has only one dimension, and we denote it by c. We let c be 0 if player 2 does not deviate with positive probability to an action which gives him $\varepsilon/3$ more than following the recommendation and 1 otherwise. Let V^2 denote the set of strategies of player 2 that involve such a deviation. Note that we are dealing here only with pure strategies. We use the term"with positive probability" only because of the stochastic signals and the randomizations of the correlation device.

Thus we have a map from the action space to probability distributions over $v+1$ dimensional vectors of zeros and ones, a finite subset of $\mathbb{R}^{v+1}$, as required for the version of the approachability theorem given in Appendix A. We denote this map by $\tilde{\varphi} : \tilde{S}^2 \to \Delta(\{0,1\}^{v+1})$. Notice that both the first and second moments of all elements of $\tilde{\varphi}(s^2)$ are bounded by 1 for every s^2. Moreover, since player 1 has only one strategy we can, without loss of generality, assume that he observes the realized payoff.

Let $\bar{\nu}$ be the expected value of the first v dimensions of the payoff vector when player 2 plays $\bar{s^2}$, that is, $\bar{\nu} = \text{proj}E_{\tilde{\varphi}(\bar{s}^2)}\nu$, where proj denotes the projection onto the first v dimensions.

Let $\tilde{f}(s^2) = E_{\tilde{\varphi}(\bar{s}^2)}\nu$ and $Z = \text{Co}\{\tilde{f}(s^2) \mid s^2 \in \tilde{S}^2\}$. Now $(\bar{\nu}, 0) = \tilde{f}(\bar{s^2})$ is in Z. And $(\bar{\nu}, \delta)$ is not in Z for any $\delta > 0$. For suppose that $(\bar{\nu}, \delta)$ was in Z, then $(\bar{\nu}, \delta) = \sum_{s^2 \in \tilde{S}^2} \alpha_{s^2} \tilde{f}(s^2)$ with $0 \le \alpha_{s^2} \le 1$ and $\sum_{s^2} \alpha_{s^2} = 1$. The vector (α_{s^2}) is essentially a mixed action putting positive weight on those strategies involving positive probability of a deviation to an $\bar{\varepsilon}$-gaining strategy (since $\delta > 0$). This means, the mixed action at one of the 2^n stages of the original game must involve distribution over player 2's actions that gains him at least $\bar{\varepsilon}$ compared to his recommended action. By the construction of $\bar{A}^2$, this mixed action cannot be more informative than his recommended action. "Not more informative" involves two possibilities. One possibility is that the mixed action is not indistinguishable from the recommended action, which gives different distributions over player 1's signals and then leads to a different distribution on one of the first blocks of the vector payoff. The other possibility is that the distribution over player 2's signals gives him less information about player 1's action and signal distribution. This means, there is no function from his observation to his actions in the reporting stage that will give the same distribution over his action as if he had played as recommended, thus not the same distribution over player 1's signals in the reporting phase. This will make the second block of the vector payoff different from $(\bar{\nu}, \delta)$.

Now let $Z_\delta = \{\nu \in Z \mid \nu_{v+1} \ge \delta\}$. and let Z_0 be the projection on the first v dimensions of $Z_{(\varepsilon_1/4)}$. Clearly $\bar{\nu}$ is not in Z_0 and Z_0 is a closed and convex set. Let $d(Z_0, \bar{\nu}) = 3q$ and let O_0 be the open q-ball around Z_0. (That is, $O_0 = \cup_{z \in Z_0} B_q(z)$.)

Let $O_1 = \{\nu \in [0,1]^{v+1} \mid \text{either } \nu_{v+1} < (\varepsilon_1/3) \text{ or } (\nu_1, \ldots, \nu_v) \in O_0\}$. Notice that Z is contained in O_1. Also let $\delta_{n'} = d(Z, \bar{\nu}_{n'})$, where $\bar{\nu}_{n'}$ is the average payoff for the first n' stages of the artificial game.

Now, by Corollary 1 of Appendix A, $\Pr(\sup_{n' \ge M_a} \delta_{n'} \ge q) \le 8/(q^2 M_a)$ for any $\varepsilon_1 > 0$ and any integer M_a. In order to have $8/(q^2 M_a) \le \varepsilon_1$, we need to choose $M_a \ge 8/(\varepsilon_1 q^2)$.

We also want to make sure that if player 2 does follow the recommendation there will be a large probability that $\bar{\nu}_{n'}$ will be very close to the point

$(\bar{\nu},0)$. For $k=1,\ldots,v$, by the Chebyshev Inequality,

$$\Pr(|\,\bar{\nu}_{n',k}-E(\bar{\nu}_{n',k})\,|\geq \frac{q}{\sqrt{v}})\leq \frac{\mathrm{Var}(\bar{\nu}_{n',k})}{(\frac{q}{\sqrt{v}})^2}\leq \frac{v}{n'q^2}$$

since $\mathrm{Var}(\bar{\nu}_{n',k})\leq(1/n')$. Also $\Pr(|\bar{\nu}_{n',v+1}-E(\bar{\nu}_{n',v+1})|=0)=1$. Thus

$$\begin{aligned}\Pr(|\,\bar{\nu}_{n'}-E(\bar{\nu}_{n'})\,|\geq q) &= \Pr(\sqrt{(\bar{\nu}_{n',1}-E(\bar{\nu}_{n',1}))^2+\cdots+(\bar{\nu}_{n',v}-E(\bar{\nu}_{n',v}))^2}\geq q)\\ &= \Pr(((\bar{\nu}_{n',1}-E(\bar{\nu}_{n',1}))^2+\cdots+(\bar{\nu}_{n',v}-E(\bar{\nu}_{n',v}))^2)\geq q^2)\\ &\leq \sum_{k=1}^{v}\Pr((\bar{\nu}_{n',k}-E(\bar{\nu}_{n',k}))^2\geq \frac{q^2}{v})\\ &= \sum_{k=1}^{v}\Pr(|\bar{\nu}_{n',k}-E(\bar{\nu}_{n',k})|\geq \frac{q}{\sqrt{v}})\\ &\leq \frac{v^2}{n'q^2}.\end{aligned}$$

If we want this probability to be very small (at most ${\varepsilon_1}^2$) the number of stages has to be at least $M_b=v^2/({\varepsilon_1}^2q^2)$.

Now the most that player 2 can gain in the final $2K(\varepsilon_1,n)+2W$ stages is $\varepsilon_1/3$. And the most he can gain without playing a strategy in V^2 is $\varepsilon_1/3$. Thus if he is to gain ε_1 he must deviate to a strategy in V^2 at least $\varepsilon_1/3$ fraction of the time. If he does this and is within q of Z then $(\nu_1,\ldots,\nu_v)$ must lie in O_0. Also if $\bar{\nu}_{n'}$ is within q of $\bar{\nu}$ then it is not in O_0.

Let $M\geq\max\{M_a,M_b\}$ and let $N_2=MN_1$ Thus in an N_2 superblock, any deviation from the recommended strategy which gives player 2 at least ε_1 gain will be detected by player 1 with probability at least $1-\varepsilon_1$. If player 1 detects a deviation by player 2 he will hold player 2 to his individually rational level for the rest of the game.

Let M' be the smallest integer greater than $1/\varepsilon_1$ so that the relative size of a block N_2 in games of length $N=M'N_2$ is (almost exactly) ε_1. Then for any strategy τ^2 and for the strategy σ^1 we have described for player 1

$$\begin{aligned}\bar{\gamma}^2_N(\sigma^1,\tau^2) &\leq \varepsilon_1+(1-\varepsilon_1)(\gamma^2(P)+\varepsilon_1+\frac{1}{M})+2\varepsilon_1\\ &= \varepsilon_1+\gamma^2(P)+\varepsilon_1+\frac{1}{M}-\varepsilon_1\gamma^2(P)-\varepsilon_1^2-\frac{\varepsilon_1}{M}+2\varepsilon_1\\ &\leq \gamma^1(P)+5\varepsilon_1\\ &= \gamma^1(P)+\varepsilon_0.\end{aligned}$$

The reasoning is as follows. The first N_2 block in which player 2 deviates to a strategy in V^2 he will be undetected with probability at most ε_1 in which case he could obtain at most 1 (in every period). If he is detected in the first

N_2 block in which he deviated to a strategy in V^2 he could have gained at most ε_1 by deviating to strategies not in V^2 and could gain at most 1 for the $1/M$ of time of that block. Also, by construction, the equilibrium strategies lead to a payoff that is within $2\varepsilon_1$ of $\gamma^2(P)$. (When following the described strategies in an N_1 block the equilibrium payoff differs by at most ε_1 from $\gamma^2(P)$ and the probability of player 1 observing a signal in O_0 in some N_1 block, and so punishing player 2, is at most ε_1—or more accurately $M{\varepsilon_1}^2$ which is almost exactly the same thing.) □

We now prove the claim that we made, but did not prove, above. This will complete the proof of the result.

Proof of Claim 1. We divide player 1's signal space A^1 into two categories, A^1_0 and A^1_1, according to the ratio of $\Pr(a^1 \mid s^1, s^2)/\Pr(a^1 \mid s^1, t^2)$. If the ratio is between 0 and 1 then a^1 is in A^1_1; if the ratio is larger than 1 then a^1 is in A^1_0. Thus $\Pr(A^1_0 \mid s^1, s^2) > \Pr(A^1_0 \mid s^1, t^2)$ and $\Pr(A^1_1 \mid s^1, s^2) < \Pr(A^1_1 \mid s^1, t^2)$. Let Y be the number of times in the sequence of K trials that the observed signal is in A^1_0. Let $P_s = \Pr(A^1_0 \mid s^1, s^2), P_t = \Pr(A^1_0 \mid s^1, t^2)$. If player 2 plays s^2, $KP_s = E(Y \mid s^1, s^2)$, and if player 2 plays t^2, $KP_t = E(Y \mid s^1, t^2)$.

We will next select K so that $\Pr(Y \le K(P_s + P_t)/2 \mid s^1, s^2) \le \varepsilon_1/n$ and $\Pr(Y \ge K(P_s + P_t)/2 \mid s^1, t^2) \le \varepsilon_1/n$. Then the probability that in any of the n K-blocks that Y_n will be in the wrong set is less than ε_1.

If player 1 plays s^1 and player 2 plays s^2 then $E(Y \mid s^1, s^2) = KP_s$ and $\mathrm{Var}(Y \mid s^1, s^2) = KP_s(1 - P_s)$. Thus, from the Chebyshev Inequality,

$$\begin{aligned}\Pr(Y \le K(P_s + P_t)/2 \mid s^1, s^2) &= \Pr(Y - KP_s \le K(P_t - P_s)/2 \mid s^1, s^2) \\ &\le \Pr(\mid Y - KP_s \mid \ge K(P_s - P_t)/2 \mid s^1, s^2) \\ &\le \frac{KP_s(1 - P_s)}{(K(P_s - P_t)/2)^2} = \frac{4P_s(1 - P_s)}{K(P_s - P_t)^2}.\end{aligned}$$

Similarly

$$\Pr(Y \ge \frac{1}{2}K(P_s + P_t) \mid s^1, t^2) \le \frac{4P_t(1 - P_t)}{K(P_s - P_t)^2}.$$

And so if we choose K to be (the smallest integer greater than) $4n/(\varepsilon_1(P_s - P_t)^2))$ then both probabilities will be less than ε_1/n.

Also, $K < (4n/(\varepsilon_1(P_s - P_t)^2))) + 1$ and so

$$\begin{aligned}\lim_{n\to\infty} \frac{nK}{2^n} &\le \lim_{n\to\infty} \left(\frac{4n^2}{2^n \varepsilon_1 (P_s - P_t)^2} + \frac{n}{2^n} \right) \\ &= 0.\end{aligned}$$

So $(2nK(\varepsilon_1, n) + 2W)/2^n$ converges to 0 as n goes to infinity and we can choose n so that this is less than η. □

Comment 4. In fact the set of payoffs to communication equilibria (see Forges (1986)) is also the same set. Since the notion of communication equilibrium is more general than that of correlated equilibrium one needs only to check the inclusion in $g(\bar{A}^1 \cap \bar{A}^2) \cap IR$. This proof is essentially no different than the proof given for the inclusion of the set of payoffs to correlated equilibria.

Comment 5. In the proof we used the notion of L-equilibrium in proving the first inclusion and the notion of uniform equilibrium in proving the second. In both cases this was the stronger result. Thus the result is proved for both notions of equilibrium.

Appendix A. The Approachability Theorem

In this appendix we give some of the basic results concerning matrix games with vector payoffs. The results are due to Blackwell (1956) and our treatment follows (in a less general setting) that of Mertens, Sorin, and Zamir (1994).

As before we consider a finite stage game with pure action sets S^1 and S^2 and mixed action sets X^1 and X^2. Rather than having a payoff associated with each pair of strategies we assume that there is a function φ from $S = S^1 \times S^2$ to the set of probability distributions over some finite subset of $\mathbb{R}^k$. The game is played more or less as before. At each stage n player i chooses an action in S^i and then a point g_n is chosen at random according to $\varphi(s^1_n, s^2_n)$. Both players then obtain some signal that reveals for player 1, at least, g_n. We let $\bar{g}_n = \frac{1}{n}\sum_{t=1}^{n} g_t$.

Definition 11 (Approachable)**.** A set C in $\mathbb{R}^k$ is said to be *approachable* by player 1 if there is a strategy for player 1 in the infinitely repeated game for which $d(\bar{g}_n, C)$ converges to zero almost surely.

Let $f(s^1, s^2)$ be the expected value of $\varphi(s^1, s^2)$ and for any x^1 in X^1 let $Z(x^1)$ be the convex hull of the points in $\{\sum_{s^1 \in S^1} x^1(s^1) f(s^1, s^2) \mid s^2 \in S^2\}$.

Theorem 3 (The approachability theorem)**.** *Let C be any closed set in $\mathbb{R}^k$. Suppose that for any g not in C there is x^1 $(= x^1(g))$ in X^1 such that the hyperplane through $h(g)$ a closest point in C to g perpendicular to the line segment between g and $h(g)$ separates g from $Z(x^1(g))$. Then C is approachable by player 1 using strategy $\sigma^1(\cdot)$, a strategy depending on the history only through $\bar{g}_n$ where*

$$\sigma^1(\bar{g}_n) = \begin{cases} x^1(\bar{g}_n) & \text{if } n > 0 \text{ and } \bar{g}_n \notin C, \\ \text{arbitrary} & \text{otherwise.} \end{cases}$$

With that strategy,

$$E(d(\bar{g}_n, C)^2) \leq 4K/n \tag{2}$$

and

$$P(\sup_{n \geq N} d(\bar{g}_n, C) \geq \varepsilon) \leq 8K/(\varepsilon^2 N). \tag{3}$$

where K is a bound on the second order moments of $\varphi(s^1, s^2)$ for all s^1 and s^2.

We need, in fact, only one relatively simple implication of the approachability theorem.

Corollary 1 (Mertens, Sorin, and Zamir, 1994, Corollary II.4.4)**.** *For any x^1 in X^1 the set $Z(x^1)$ is approachable by player 1 using the constant strategy $\sigma^1(\cdot) = x^1$. And, again, with this strategy inequalities (2) and (3) hold.*

REFERENCES

ROBERT J. AUMANN (1989): *Lectures on Game Theory,* Westview Press, Bolder.

DAVID BLACKWELL (1951): "Comparison of Experiments," in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability,* 93–102, University of California Press, Berkeley and Los Angeles.

DAVID BLACKWELL (1956): "An Analog of the Minimax Theorem For Vector Payoffs," *Pacific Journal of Mathematics,* 6, 1–8.

FRANÇOISE FORGES (1986): "An Approach to Communication Equilibria," *Econometrica*, 54(6), 1375–1385.

EHUD LEHRER (1990): "Nash Equilibria of *n*-Player Repeated Games with Semi-Standard Information," *International Journal of Game Theory,* 19, 191–217.

EHUD LEHRER (1991): "Internal Correlation in Repeated Games," *International Journal of Game Theory,* 19, 431–456.

EHUD LEHRER (1992a): "Correlated Equilibria In Two-Player Repeated Games with Nonobservable Actions," *Mathematics of Operations Research*, 17(1), 175–199.

EHUD LEHRER (1992b): "On the Equilibrium Payoff Set of Two-Player Repeated Games with Imperfect Monitoring," *International Journal of Game Theory*, 20, 211–226.

EHUD LEHRER (1992c): "Two-Player Repeated Games with Nonobservable Actions and Observable Payoffs," *Mathematics of Operations Research*, 17(1), 200–224.

JEAN-FRANÇOIS MERTENS, SYLVAIN SORIN, AND SHMUEL ZAMIR (1994): "Repeated Games," CORE Discussion Papers 9420, 9421, and 9422.

SYLVAIN SORIN (1990): "Supergames," in *Game Theory and Applications*, edited by Tatsuro Ichiishi, Abraham Neyman, and Yair Tauman, 46–63, Academic Press, San Diego.