MATHEMATICS OF OPERATIONS RESEARCH
Vol. 23, No. 3, August 1998
Printed in U.S.A.

FINITELY REPEATED GAMES WITH FINITE AUTOMATA

ABRAHAM NEYMAN

In honor of R. J. Aumann's 65th birthday

The paper studies the implications of bounding the complexity of the strategies players may select, on the set of equilibrium payoffs in repeated games. The complexity of a strategy is measured by the size of the minimal automation that can implement it.

A finite automation has a finite number of states and an initial state. It prescribes the action to be taken as a function of the current state and a transition function changing the state of the automaton as a function of its current state and the present actions of the other players. The size of an automaton is its number of states.

The main results imply in particular that in two person repeated games, the set of equilibrium payoffs of a sequence of such games, $G(n)$, $n = 1, 2, \ldots$, converges as n goes to infinity to the individual rational and feasible payoffs of the one shot game, whenever the bound on one of the two automata sizes is polynomial or subexponential in n and both, the length of the game and the bounds of the automata sizes are at least n.

A special case of such result justifies cooperation in the finitely repeated prisoner's dilemma, without departure from strict utility maximization or complete information, but under the assumption that there are bounds (possibly very large) to the complexity of the strategies that the players may use.

1. Introduction. A fundamental message of the theory of repeated games is that the cooperative outcomes of multi person games, provided those games are repeated over and over, are consistent with the usual ''selfish'' utility-maximizing behavior assumed in economic theory. For example, in the prisoner's dilemma of Figure 1, the only rational outcome in noncooperative play of the one shot game is (1, 1). But in infinitely repeated play, the players can achieve the cooperative outcome (3, 3) in equilibrium.

Indeed, the Folk theorem and several other results (Aumann 1959, 1960, 1981, Aumann and Shapley 1994, Fudenberg and Maskin 1986, Rubinstein 1994, and Sorin 1986, 1990, 1992) assert that cooperative outcomes of the one-shot game are equilibria (and also perfect equilibria) of the infinite repetition of that game. Cooperation is also rationalized by Nash equilibria or even perfect equilibria in some classes of finitely repeated games (Benoit and Krishna 1985, 1987 and Gossner 1995). However, there are games, including the prisoner's dilemma, that are not in this class; indeed, in any finite repetition of the prisoner's dilemma, all equilibria (and all correlated equilibria and all communication equilibria) lead to the noncooperative outcome at each stage. This contrasts with the common observation in the experiments involving finite repetitions of the prisoner's dilemma, that players do not always choose the single-period dominant actions, but instead achieve some mode of cooperation.

The present paper justifies cooperation in the finitely repeated prisoner's dilemma, as well as in other finitely repeated games, without departure from the strict utility maximization, but under the assumption that there are bounds (possibly very large) to the complexity of the strategies that players may use.

2. The model.

2.1. Strategic games and equilibria. Let G be an n-person game, $G = (N; (A^i)_{i\in N}; (r^i)_{i\in N})$, where $N = \{1, 2, \ldots, n\}$ is the set of players, A^i is a finite set of actions for

0364-765X/98/2303/0513/$05.00

	Unfriendly	Friendly
Unfriendly	1, 1	4, 0
Friendly	0, 4	3, 3

FIGURE 1.

Player i, $i = 1, \dots, n$, and $r^i : A^1 \times \cdots \times A^n \to \mathbb{R}$ is the payoff function of Player i. The set A^i is called also the set of pure strategies of Player i. The Cartesian product $\times_{i \in N} A^i = A^1 \times A^2 \times \cdots \times A^n$ is denoted by A and $r : A \to \mathbb{R}^N$ denotes the vector valued function whose ith component is r^i, i.e., $r(a) = (r^1(a), \dots, r^n(a))$. We write (N, A, r) for short for $(N; (A^i)_{i \in N}; (r^i)_{i \in N})$. For any finite set B we denote by $\Delta(B)$ the set of all probability distributions on B. For any player i and any n-person game G, we denote by $v^i(G)$ his individual rational payoff in the mixed extension of the game G, i.e., $v^i(G) = \min \max r^i(a^i, \sigma^{-i})$ where the max ranges over all pure strategies of Player i, and the min ranges over all $N\setminus\{i\}$-tuples of mixed strategies of the other players, and r^i denotes also the payoff to Player i in the mixed extension of the game. We denote by $u^i(G)$ the individual rational payoff of Player i in pure strategies, i.e., $u^i(G) = \min \max r^i(a^i, a^{-i})$ where the max ranges over all pure strategies of Player i, and the min ranges over all $N\setminus\{i\}$-tuples of pure strategies of the other players. Obviously $u^i(G) \geq v^i(G)$. An equilibrium of a strategic game (N, A, r) is an N-tuple of (mixed) strategies $\sigma = (\sigma^i)_{i \in N}$, $\sigma^i \in \Delta(A^i)$, such that for every $i \in N$ and any strategy of Player i, $\tau^i \in A^i$, $r^i(\tau^i, \sigma^{-i}) \leq r^i(\sigma^i, \sigma^{-i})$. If σ is an equilibrium, the vector payoff $r(\sigma)$ is called an equilibrium payoff. For any game G in strategic form we denote by $E(G)$ the set of all equilibrium payoffs in the game G.

2.2. The finitely repeated game G^T. Given an n-person game, $G = (N; (A^i)_{i \in N}; (r^i)_{i \in N})$, we define a new game in strategic form $G^T = (N; (\Sigma^i(T))_{i \in N}; (r^i_T)_{i \in N})$ which models a sequence of T plays of G, called *stages*. After each stage, each player is informed of what the others did at the previous stage, and he remembers what he himself did and what he knew at previous stages. Thus, the information available to each player before choosing his action at stage t is all past actions of the players in previous stages of the game. Formally, let H_t, $t = 1, \dots, T$, be the Cartesian product of A by itself $t-1$ times, i.e., $H_t = A^{t-1}$, with the common set theoretic identification $A^0 = \{\emptyset\}$, and let $H = \bigcup_{t=1}^T H_t$. A pure strategy σ^i of Player i in G^T is a function $\sigma^i : H \to A^i$. Obviously, H is a disjoint union of H_t, $t = 1, \dots, T$ and therefore one often defines $\sigma^i_t : H_t \to A^i$ as the restriction of σ^i to H_t. We denote the set of all pure strategies of Player i in G^T by $\Sigma^i(T)$. The set of pure strategies of Player i in the infinitely repeated game G^* is denoted Σ^i.

Any N-tuple $\sigma = (\sigma^1, \dots, \sigma^n) \in \times_{i \in N} \Sigma^i(T)$ of pure strategies in G^T induces a play $\omega(\sigma) = (\omega_1(\sigma), \dots, \omega_T(\sigma))$ defined by induction: $\omega_1(\sigma) = (\sigma^1(\emptyset), \dots, \sigma^n(\emptyset)) = \sigma(\emptyset)$ and $\omega_t(\sigma) = \sigma(\omega_1(\sigma), \dots, \omega_{t-1}(\sigma))$ or in other words $\omega^i_1(\sigma) = \sigma^i(\emptyset)$, and $\omega^i_t(\sigma) = \sigma^i(\omega_1(\sigma), \dots, \omega_{t-1}(\sigma)) = \sigma^i_t(\omega_1(\sigma), \dots, \omega_{t-1}(\sigma))$.

Set

$$r_T(\sigma) = \frac{r(\omega_1(\sigma)) + \cdots + r(\omega_T(\sigma))}{T}.$$

Two strategies σ^i and τ^i of Player i in G^T are *equivalent* if for every $N\setminus\{i\}$-tuple of pure strategies $\sigma^{-i} = (\sigma^j)_{j \in N\setminus\{i\}}$, $\omega_t(\sigma^i, \sigma^{-i}) = \omega_t(\tau^i, \sigma^{-i})$ for every $1 \leq t \leq T$. The equivalence classes of pure strategies are called *reduced strategies*.

2.3. Finite automata. A *finite automaton* for Player i is a four-tuple $\langle M^i, q^i, f^i, g^i\rangle$, where M^i is a finite set, $q^i \in M^i$, $f^i : M^i \to A^i$, and $g^i : M^i \times A^{-i} \to M^i$. The set M^i is the set of possible states of the automaton, q^i is the initial state, $f^i(q)$ is the action taken by the automaton when in state q, and g^i describes the transition of the automaton from state to state; if at state q the other players choose the action tuple a^{-i}, then the automaton's next state is $g^i(q, a^{-i})$. The *size* of the finite automaton is the number of states.

A finite automaton for Player i can be viewed as a prescription for Player i to choose his actions in the various stages of the repeated game. The action to be taken at stage 1 is $f^i(q^i)$. The action in stage 2 is $f^i(g^i(q^i, a_1^{-i}))$ where a_1^{-i} is the actions taken by the other players in stage 1. More generally, if we define inductively,

$$g^i(q, b_1, \ldots, b_t) = g^i(g^i(q, b_1, \ldots, b_{t-1}), b_t),$$

where $b_j \in A^{-i}$, the action prescribed by the automaton for Player i at stage t is $f^i(g^i(q^i, a_1^{-1}, \ldots, a_{t-1}^{-i}))$ where a_j^{-i}, $1 \le j < t$, is the $N\setminus\{i\}$-tuple of actions at stage j. Therefore, any automaton $\alpha = \langle M^i, q^i, f^i, g^i\rangle$ of Player i induces a strategy σ_α^i in G^T that is given by

$$\sigma_\alpha^i(a_1, \ldots, a_{t-1}) = f^i(g^i(q^i, a_1^{-i}, \ldots, a_{t-1}^{-i})).$$

Note also that an automaton α of Player i induces also a strategy σ_α^i of Player i in the infinitely repeated game G^*. A strategy σ^i of Player i in G^* (in G^T) is *implemented* by the automaton α of Player i if σ^i is equivalent to σ_α^i, i.e., if for every $\sigma^{-i} \in \times_{j\neq i}\Sigma^j$ ($\sigma^{-i} \in \times_{j\neq i}\Sigma^j(T)$), $\omega(\sigma^i, \sigma^{-i}) = \omega(\sigma_\alpha^i, \sigma^{-i})$.

A finite sequence of actions $(a_1, \ldots, a_t)$ and a pure strategy σ^i of Player i in G^* are *compatible*, if for every $1 \le s \le t$, $\sigma^i(a_1, \ldots, a_{s-1}) = a_s^i$. The set of all sequences of actions of length n that are compatible with σ^i is denoted $A^n(\sigma^i)$. Given a strategy σ^i of Player i in G^*, any sequence of actions $(a_1, \ldots, a_t)$, *induces* a strategy $(\sigma^i | a_1, \ldots, a_t)$ in G^*, by

$$(\sigma^i | a_1, \ldots, a_t)(b_1, \ldots, b_s) = \sigma^i(a_1, \ldots, a_t, b_1, \ldots, b_s).$$

Section 3 shows that the number of different reduced strategies that are induced by a given pure strategy σ^i of Player i in G^* and all σ^i-compatible sequences of actions equals the size of the smallest automaton that implements σ.

2.4. Finitely repeated games with finite automata. Given a game G in strategic form and positive integers $m_1, \ldots, m_n$, we define $\Sigma^i(T, m_i)$ to be all pure strategies in $\Sigma^i(T)$ that are induced by an automaton of size m_i. Note that if a strategy is induced by an automaton of size m_i and $m_i' \ge m_i$ then it is also induced by an automaton of size m_i'. The game $G^T(m_1, \ldots, m_n)$ is the strategic game $(N; (\Sigma^i(T, m_i))_{i\in N}; r_T)$ where r_T here is the restriction of our earlier payoff function r_T to $\times_{i\in N}\Sigma^i(T, m_i)$.

3. Automata and strategic complexity. We define in this section two measures of complexity of strategies in the repeated game. One complexity measure is the size of the smallest automaton that implements σ, and the other one is the number of different reduced strategies that are induced by a given pure strategy σ^i of Player i in G^* and all σ^i-compatible sequences of actions. It is proved that the two complexity measures coincide. For an analog result for "exact automata" see Kalai (1990).

More precisely we define the first measure of complexity of a pure strategy $\sigma \in \Sigma^i$ (or $\sigma \in \Sigma^i(T)$), $\text{comp}_1(\sigma)$, as the smallest size of an automaton that implements σ^i. The second measure of complexity (of a pure strategy), $\text{comp}_2(\sigma^i)$, is defined as the cardinality of the set of equivalence classes of $\{(\sigma^i|a_1, \ldots, a_n)|n \in \mathbb{N}_0$ and $(a_1, \ldots, a_n) \in A^n(\sigma^i)\}$, where $\mathbb{N}_0$ is the set of nonnegative integers.

The measure of complexity $\text{comp}_2(\sigma)$ is defined on Σ^i, the strategies in the infinitely repeated game. It has a natural extension to a measure of complexity for strategies in the finitely repeated game G^T; let $\sigma = (\sigma_t)_{t=1}^T \in \Sigma^i(T)$, and define

$$\text{comp}_2(\sigma) = \min\{\text{comp}_2(\tau) : \tau \in \Sigma^i \text{ and } \forall t, 1 \leq t \leq T, \sigma_t = \tau_t\}.$$

PROPOSITION 1. *For every pure strategy $\sigma^i \in \Sigma^i$, or in $\Sigma^i(T)$,*

$$\text{comp}_1(\sigma^i) = \text{comp}_2(\sigma^i).$$

PROOF. Let $\sigma^i \in \Sigma^i$ be implemented by the automaton $\langle M^i, q^i, f^i, g^i\rangle$. We will show first that $\text{comp}_2(\sigma^i) \leq |M^i|$, where $|M^i|$ is the number of elements in the set M^i. Let $a = (a_1, \ldots, a_n)$ and $b = (b_1, \ldots, b_k)$ be two plays that are compatible with σ^i, and such that $(\sigma^i|a)$ is not equivalent to $(\sigma^i|b)$. Then there is a play $c = (c_1, \ldots, c_p)$ that is compatible with both $(\sigma^i|a)$ and $(\sigma^i|b)$ and such that $(\sigma^i|a)(c_1, \ldots, c_p) \neq (\sigma^i|b)(c_1, \ldots, c_p)$. Therefore,

$$g^i(q^i, a_1^{-i}, \ldots, a_n^{-i}, c_1^{-i}, \ldots, c_p^{-i}) \neq g^i(q^i, b_1^{-i}, \ldots, b_k^{-i}, c_1^{-i}, \ldots, c_p^{-i}).$$

As

$$g^i(q^i, a_1^{-i}, \ldots, a_n^{-i}, c_1^{-i}, \ldots, c_p^{-i}) = g^i(g^i(q^i, a_1^{-i}, \ldots, a_n^{-i}), c_1^{-i}, \ldots, c_p^{-i})$$

and

$$g^i(q^i, b_1^{-i}, \ldots, b_k^{-i}, c_1^{-i}, \ldots, c_p^{-i}) = g^i(g^i(q^i, b_1^{-i}, \ldots, b_k^{-i}), c_1^{-i}, \ldots, c_p^{-i}),$$

we deduce that

$$g^i(q^i, a_1^{-i}, \ldots, a_n^{-i}) \neq g^i(q^i, b_1^{-i}, \ldots, b_k^{-i}).$$

Thus, $\text{comp}_2(\sigma^i) \leq |M^i|$ which implies that

$$\text{comp}_1(\sigma^i) \geq \text{comp}_2(\sigma^i).$$

Let M be the set of equivalence classes of $\{(\sigma^i|a) : n \in \mathbb{N}_0$ and $a \in A^n(\sigma^i)\}$, and we identify here a strategy with its equivalence class. Let $q^i = \sigma^i$, and for $n \in \mathbb{N}_0$ and $a = (a_1, \ldots, a_n) \in A^n(\sigma^i)$,

$$f^i(\sigma^i|a_1, \ldots, a_n) = (\sigma^i|a_1, \ldots, a_n)_1,$$

and

$$g^i((\sigma^i|a_1, \ldots, a_n), a^{-i}) = (\sigma^i|a_1, \ldots, a_n, ((\sigma^i|a_1, \ldots, a_n)_1, a^{-i})).$$

Then, σ^i is implemented by $\langle M, q^i, f^i, g^i\rangle$ and thus we conclude that

$$\text{comp}_1(\sigma^i) \leq \text{comp}_2(\sigma^i),$$

which completes the proof of the proposition. □

4. Automata and play complexity. We define in this section a complexity measure for each player on plays (and on sets of plays) of the repeated game, and derive various inequalities regarding the complexity of a play. The definitions and results in this section are of independent interest and do serve as a good introduction to the more involved complexity counting in the proof of our main result. Let $G = (N, A, r)$ be a strategic game. A play of the repeated game is an element of the set $\bigcup_{t=1}^{\infty} A^t \cup A^{\infty}$. A finite (infinite) play is an element of $\bigcup_{t=1}^{\infty} A^t$ (A^{∞}). A strategy σ^i of Player i in G^* is *compatible* with the infinite play $(a_1, \cdots)$, if for every positive integer t, the finite play $(a_1, \ldots, a_t)$ is compatible with σ^i. Let ω be a play. We define the ith player complexity of the play ω, $\operatorname{comp}^i(\omega)$, as the smallest complexity of a strategy σ^i of Player i which is compatible with ω, i.e.,

$$\operatorname{comp}^i(\omega) = \inf\{\operatorname{comp}(\sigma) \colon \sigma \in \Sigma^i \text{ is compatible with } \omega\}.$$

LEMMA 1. *Let $a = (a_1, \ldots, a_t) \in A^t$. Then*

$$\operatorname{comp}^i(a) \le t.$$

PROOF. The strategy σ^i of Player i which is implemented by the automaton $\langle M^i, q^i, f^i, g^i \rangle$ where $M^i = \{1, \ldots, t\}$, $q^i = 1$, $f^i : M^i \to A^i$ is given by $f^i(s) = a_s^i$, and $g^i(s, *) = \min(s + 1, t)$ is compatible with $a = (a_1, \ldots, a_t)$. □

If n and m are two positive integers, $a = (a_1, \ldots, a_n) \in A^n$ and $b = (b_1, \ldots, b_m) \in A^m$, then we denote by $a + b$ the element of A^{n+m} that is defined by

$$a + b = (a_1, \ldots, a_n, b_1, \ldots, b_m).$$

LEMMA 2. *Let $a = (a_1, \ldots, a_t) \in A^t$ and $b = (b_1, \ldots, b_s) \in A^s$. Then*

$$\operatorname{comp}^i(a + b) \ge \max(\operatorname{comp}^i(a), \operatorname{comp}^i(b)).$$

PROOF. Let σ^i be a strategy of Player i which is compatible with the play $(a_1, \ldots, a_t, b_1, \ldots, b_s)$. Then σ^i is compatible with a and therefore $\operatorname{comp}^i(a + b) \ge \operatorname{comp}^i(a)$. On the other hand, $(\sigma^i | a_1, \ldots, a_t)$ is compatible with b and as $\operatorname{comp}(\sigma^i) \ge \operatorname{comp}(\sigma^i | a_1, \ldots, a_t)$, $\operatorname{comp}^i(a + b) \ge \operatorname{comp}^i(b)$. □

For $a \in A^n$ and a positive integer d we define $d * a$ by induction on d: $1 * a = a$ and

$$(d + 1) * a = d * a + a.$$

LEMMA 3. *Let $a = (a_1, \ldots, a_t) \in A^t$ with $a_1 = a_2 = \cdots = a_{t-1}$ and $a_{t-1}^i \ne a_t^i$. Then*

$$\operatorname{comp}^i(a) = t.$$

PROOF. By Lemma 1 $\operatorname{comp}^i(a) \le t$. Let σ^i be a strategy which is compatible with a. Then the strategies $\sigma, (\sigma | a_1), \ldots, (\sigma | a_1, \ldots, a_{t-1})$ are t different strategies: for $0 \le s < q \le t - 1$,

$$(\sigma^i | s * a_1)((t - q - 1) * a_1) = a_1^i \ne a_t^i = (\sigma^i | q * a_1)((t - q - 1) * a_1),$$

where $\tau(0 * a_1) = \tau_1$ and $(\tau | 0 * a_1) = \tau$. Thus the strategies $(\sigma^i | s * a_1)$, $0 \le s \le t - 1$, are t different strategies, and therefore $\operatorname{comp}^i(a) \ge t$. □

For $a \in A^n$, $b \in A^k$ and a positive integer s with $\min(n, k) \ge s - 1$, we define $a =_s b$ if $a_t = b_t$ for every $t < s$.

The next lemma provides a lower bound for the complexity of a play which repeats d times a play (of the form $t * a + b$) and in which the actions of Player i initiate a deviation from the periodic play.

LEMMA 4. *Let* $a = (a_1, \ldots, a_k) \in A^k$ *and* $b = (b_1, \ldots, b_n) \in A^n$ *with* $a_1^i \neq b_1^i$, $t \geq 0$ *and* $d \geq 1$. *Assume that* $\omega = (\omega_1, \ldots, \omega_s) \in A^s$ *with* $(d-1)(tk+n) + tk + 1 < s \leq (d+1)(tk+n)$ *and*

$$d * (t * a + b) =_s \omega \quad \text{and} \quad ((d+1) * (t * a + b))_s^i \neq \omega_s^i.$$

Then

$$\text{comp}^i(\omega) \geq d(t+1).$$

PROOF. Let σ^i be a strategy which is compatible with ω. Consider the set of strategies $\{(\sigma^i | j * (t * a + b) + l * a) : 0 \leq j < d \text{ and } 0 \leq l \leq t\}$, i.e., the set of strategies $\{(\sigma^i | \omega_1, \ldots, \omega_{j(tk+n)+lk})\}$ (and where for $j = l = 0$ we mean by $(\sigma^i | j * (a + b) + l * a)$ or by $(\sigma^i | \omega_1, \ldots, \omega_{j(tk+n)+lk})$ the strategy σ^i. If $(j, l) \neq (j', l')$ then either $l \neq l'$ and without loss of generality $l < l'$, and then $(\sigma^i | j * (t * a + b) + l * a)((t - l') * a) = a_1^i \neq b_1^i = (\sigma | j' * (t * a + b) + l' * a)((t - l') * a)$ or $l = l'$ and without loss of generality $j < j'$, and then

$$(\sigma | j * (t * a + b) + l * a)(\omega_{j'(tk+n)+lk+1}, \ldots, \omega_{s-1}) = (d * (t * a + b))_s^i$$

and

$$(\sigma | j' * (t * a + b) + l * a)(\omega_{j'(tk+n)+lk+1}, \ldots, \omega_{s-1}) = \omega_s^i.$$

As $\omega_s^i \neq (d * (t * a + b))_s^i$, the cardinality of this set of strategies is $d(t+1)$. Therefore $\text{comp}^i(\omega) \geq d(t+1)$. □

Lemma 4 is of interest also for the special case when $k = 1$. Elaboration on such complexity counting (where $k = 1$) appears later in the proof of the main result.

REMARK 1. *The conclusion of Lemma* 4 *does not hold when replacing the assumption* $a_1^i \neq b_1^i$ *with the assumption* $a_1 \neq b_1$.

Indeed, assume that there are two players, $i = 1, 2$, and for each Player i, $\{0, 1\} \subset A^i$. Consider the action pairs $a = (0, 0)$ and $b = (0, 1)$. For any $d \in \mathbb{N}$, the play $\omega = (\omega_1, \ldots, \omega_s)$ with $s = d(t+1) + 1$ and $\omega = d * (t * a + b) + (1, 0)$ satisfies $d * (t * a + b) =_s \omega$ and $(d+1) * (t * a + b)_s^i \neq \omega_s$ while

$$\text{comp}^1(d * (t * a + b) + (1, 0)) = d + 1.$$

The complexity of the strategy of Player 1, which plays 0 as long as the number of past action 1 is at most d and plays 1 otherwise, equals $d + 1$.

The next lemma provides a complexity lower bound for a play which departs from a fully coordinated periodic play after completing a fixed given number of cycles. If $f : A^1 \to A^2$ is a $1 - 1$ function and $a = (a_1, \ldots, a_n)$ is a play with $a_t^2 = f(a_t^1)$ for every $1 \leq t \leq n$ we call the play a a *coordinated* play.

LEMMA 5. *Let* $a = (a_1, \ldots, a_n)$ *be a coordinated play*, $b \in A$ *with* $b^1 \neq a_1^1$, *and* $d \in \mathbb{N}$. *Then*

$$\text{comp}^1(d * a + b) \geq (d-1)n + 1.$$

PROOF. Assume that $m \leq n$ is the period of $(d+1) * a$. If $m < n$, there is a play $c = (c_1, \ldots, c_m)$, a positive integer $\bar{d} \in \mathbb{N}$ and a play e, such that $d * a = e + \bar{d} * c$, $b^1 \neq c_1^1$, and $(\bar{d} - 1)m \geq (d-1)n$. It follows from Lemma 2 that we can assume without loss of generality that $m = n$ is the period of $(d + 1) * a$. Let $\omega = (\omega_1, \dots, \omega_{dn+1}) = d * a + b$. Let σ be a strategy of Player 1 which is compatible with ω. If $\text{comp}^i(\sigma) < (d - 1)n + 1$, there are two positive integers s and t with $1 \leq s < t \leq (d - 1)n + 1$ such that

$$(\sigma | \omega_1, \dots, \omega_{s-1}) = (\sigma | \omega_1, \dots, \omega_{t-1})$$

where for $s = 1$ the left-hand side of the above equality is σ. As $d * a$ is a coordinated play, it follows by induction on k that for every $0 \leq k \leq dn + 1 - t$, $\omega^1_{s+k} = \omega^1_{t+k}$. In particular, setting $k = dn + 1 - t$, $\omega^1_{s+dn+1-t} = \omega^1_{dn+1} = b^1$. As $b^1 \neq a^1_1$ it follows that $t - s$ is not a multiple of n. On the other hand, it implies that

$$(1) \qquad (\omega^1_s, \dots, \omega^1_{s+n-1}) = (\omega^1_t, \dots, \omega^1_{t+n-1})$$

As $(\omega^1_1, \dots, \omega^1_{dn})$ is n-periodic we may assume without loss of generality that (1) holds for $s \leq n$ and $t < s + n$. As $d * a$ is n-periodic, if $s > 1$, $\omega^1_{s-1} = \omega^i_{s+n-1} = \omega^i_{t+n-1} = \omega^i_{t-1}$ and therefore (1) holds for $s - 1$ and $t - 1$ and thus without loss of generality we may assume that $s = 1$ in (1), and thus $d * a$ has a period of size $t - s < n$. □

The lower bound in the previous lemma can be replaced with $(d - 1) * n + 2$ and this is the best one possible. Indeed, if $a = (0, 0) + (n - 1) * (1, 1)$ then $\text{comp}^i(d * a + (1, 1) = (d - 1)n + 2$.

The next lemma provides a complexity lower bound for a play which departs from a periodic play a after completing d cycles.

LEMMA 6. *Let $a = (a_1, \dots, a_t)$ be a play and d a positive integer. Let $B^i \subset A^i$ be a nonempty subset of the actions of Player i. Assume that $k : B^i \to \mathbb{N}$ is such that for every $b^i \in B^i$ there is $s = s(b^i) < t - k(b^i)$ with $a_{s+1} = \cdots = a_{s+k(b^i)}$ and $b^i = a^i_{s+1} \neq a^i_{s+k(b^i)+1}$. Then $\text{comp}^i(a) \geq \sum_{a^i \in B^i} k(a^i)$, and if $\omega = (\omega_1, \dots, \omega_s)$ is a play with $td < s \leq t(d + 1)$, $(d + 1) * a =_s \omega$ and $((d + 1) * a)^i_s \neq \omega^i_s$ then*

$$\text{comp}^i(\omega) \geq d \sum_{a^i \in B^i} k(a^i).$$

PROOF. Assume that $\sigma \in \Sigma^i$ is compatible with ω. For any $b^i, c^i \in B^i$, any two positive integers m, n with $s(b^i) \leq m < s(b^i) + k(b^i)$, $s(c^i) \leq n < s(c^i) + k(c^i)$, and $0 \leq q \leq p < d$, if $(b^i, m, p) \neq (c^i, n, q)$, then

$$(2) \qquad (\sigma | \omega_1, \dots, \omega_{m+pt}) \neq (\sigma | \omega_1, \dots, \omega_{n+qt}).$$

Indeed, if $b^i \neq c^i$ then ($m \neq n$ and)

$$(\sigma | \omega_1, \dots, \omega_{m+pt})_1 = b^i \neq c^i = (\sigma | \omega_1, \dots, \omega_{n+qt})_1.$$

If $b^i = c^i$ and $m < n$,

$$(\sigma | \omega_1, \dots, \omega_{m+pt})((s(b^i) + k(b^i) - n) * a_m) = b^i$$

and

$$(\sigma | \omega_1, \dots, \omega_{n+qt})((s(b^i) + k(b^i) - n) * a_m) = a^i_{s(b^i)+k(b^i)+1}.$$

As $a^i_{s(b^i)+k(b^i)+1} \neq b^i$, (2) follows.

If $b^i = c^i$, and $m = n$ and $0 \leq p < q < d$,

$$(\sigma | \omega_1, \ldots, \omega_{m+pt})(\omega_{m+pt+1} \cdots, \omega_{s-t(q-p)-1}) = ((d+1) * a)^i_s$$

and

$$(\sigma | \omega_1, \ldots, \omega_{n+qt})(\omega_{m+qt+1} \cdots, \omega_{s-1}) = \omega^i_s.$$

As $(\omega_{m+pt+1} \cdots, \omega_{s-t(q-p)-1}) = (\omega_{m+qt+1} \cdots, \omega_{s-1})$ and $((d+1) * a)^i_s \neq \omega^i_s$, (2) follows. Finally, observe that the number of triples (b^i, m, p) with $b^i \in B^i$, $s(b^i) \leq m < s(b^i) + k(b^i)$ and $0 \leq q < d$ equals $d \sum_{a^i \in B^i} k(a^i)$. □

A set of plays Q is *conforming* for Player i if for all $(a_1, \ldots, a_t, b_1, \cdots)$ and $(a_1, \ldots, a_t, c_1, \cdots)$ in Q,

$$b^i_1 = c^i_1.$$

A pure strategy σ^i of Player i is *conformable* to Q if it is compatible with any $\omega \in Q$. Notice that a set of plays Q is conforming for Player i if and only if there is a pure strategy of Player i that is conformable to Q. The ith player complexity of a set of plays Q that is conforming for Player i is defined as the smallest complexity of a strategy σ^i of Player i that is conformable to Q, i.e.,

$$\text{comp}^i(Q) = \inf\{\text{comp}^i(\sigma) : \sigma \in \Sigma^i \text{ is conformable to } Q\}.$$

In what follows, we discuss the complexity of some special class of sets of plays. We hope that these remarks will help the reader in following the corresponding parts in the proofs.

Consider an arbitrary 2-player game G with two actions for each player labeled 0 and 1. Let E be a set of sequences of zeroes and ones of length k and such that for every $\epsilon = (\epsilon_1, \ldots, \epsilon_k) \in E$ and $1 \leq i < k$, $(\epsilon_1, \ldots, \epsilon_i, 0, \ldots, 0) \in E$.

Assume that for every $\epsilon \in E$, $\gamma(\epsilon)$ is a play of length $k_1 > k$ with $\gamma^1_t(\epsilon) = 0$ for every $1 \leq t \leq k_1$ and $\gamma^2_t(\epsilon) = \epsilon_t$ for every $1 \leq t \leq k$. We associate with every injective (1 − 1) function β from E to the set of coordinated plays of length k_2 a set Q of plays of the repeated game, and we comment on its complexity.

Fix positive integers $l > k_1 + k_2$ and $d \geq 2$. What is the complexity of Player 1 of the set

$$Q = \{\gamma(\epsilon) + d * ((l - k_2) * (0, 0) + \beta(\epsilon)) | \epsilon \in E\}.$$

A simple upper bound for $\text{comp}^1(Q)$ is $|E|(k_1 + l)$ where $|E|$ is the number of elements of E. If $\epsilon^2_{k_1} = 1$ for every $\epsilon \in E$, then, $\text{comp}^1(Q) \leq |E| l$. Indeed, let $M = E \times \{1, 2, \ldots, l\}$ be the set of states of an automaton of Player 1, and the initial state of the automaton is $(\epsilon^*, p))$, where $p < l - k_1 - k_2$ and $\epsilon^* = (0, \ldots, 0) \in E$. The action function $f : M \to A^1$ is defined by $f(\epsilon, i) = 0$ if $1 \leq i \leq l - k_2$ and $f(\epsilon, i) = \beta^1_{i+k_2-l}(\epsilon)$ if $l - k_2 < i \leq l$. Finally the transition function $g : M \times A^2 \to M$ obeys: $g((\epsilon, i), f(\epsilon, i))$ equals $(\epsilon, i+1)$ if $1 \leq i < l$ and it equals $(\epsilon, 1)$ if $i = l$. This ensures that once the automaton is at state $(\epsilon, 1)$ it follows the loop $(l - k_2) * (0, 0) + \beta(\epsilon)$ as long as Player 2 follows the loop. The strategy of Player 1 that is induced by the automaton $\langle M, (\epsilon^*, p), f, g\rangle$ is compatible with any play in Q whenever $g((\epsilon^*, p), \gamma^2_1(\epsilon), \ldots, \gamma^2_{k_1}(\epsilon)) = (\epsilon, 1)$ for every $\epsilon \in E$. There are many transition functions with the above property. If, in addition, $\beta_{k_2}(\epsilon) = (1, 1)$ for every $\epsilon \in E$, the complexity of Player 1 of the set of plays Q equals $l|E|$.

5. Statements of the main results. The main results of the present paper address the asymptotic behavior of the sets of equilibrium payoffs, $E(G^T(m_1, m_2))$, of the two player games $G^T(m_1, m_2)$ as T, m_1 and m_2 go to ∞. All convergence of sets is with respect to the Hausdorff topology. The main result will follow from theorems which provide conditions on a list of variables:

a feasible payoff $x \in \text{co}(r(A))$,
a positive constant $\epsilon > 0$,
the number of repetitions T, and
the bounds of the automata sizes, m_1 and m_2,

that guarantee the existence of an equilibrium payoff y of the game $G^T(m_1, m_2)$ that is ϵ-close to x. One of the conditions will ensure that a payoff in a sufficiently small neighborhood of x is generated by strategies that are implemented by automata of sizes which are less than m_1 and m_2. This condition is stated by means of the inequalities $m_i \geq m_0$ where m_0 is sufficiently large. Another condition requires the bounds of one or both automata sizes to be subexponential in the number of repetitions, i.e., a condition that asserts that $(\log m_i)/T$ is sufficiently small. Theorems 1 and 2 require a subexponential (as a function of T) bound of $\min(m_1, m_2)$ while Theorems 3 and the main theorem require a subexponential bound of $\max(m_1, m_2)$ (as a function of $\min(T, \min(m_1, m_2))$).

THEOREM 1. *Let $G = (\{1, 2\}, A, r)$ be a two person game in strategic form. Then for every $\varepsilon > 0$ sufficiently small, there are positive integers T_0 and m_0, such that if $T \geq T_0$, and $x \in \text{co}(r(A))$ with $x^1 > u^1(G)$, and $x^2 > u^2(G)$, and*

$$m_0 \leq \min(m_1, m_2) \leq \exp(\varepsilon^3 T),$$

then there is $y \in E(G^T(m_1, m_2))$ with

$$|y^i - x^i| < \varepsilon.$$

Special cases of the above theorem have been stated in previous publications. Neyman (1985) states that in the case of the finitely repeated prisoner's dilemma G, for any positive integer k, there is T_0 such that if $T \geq T_0$ and $T^{1/k} \leq \min(m_1, m_2) \leq \max(m_1, m_2) \leq T^k$, then there is a mixed strategy equilibrium of $G^T(m_1, m_2)$ in which the payoff is $1/k$-close to the ''cooperative'' payoff of G. Papadimitriou and Yannakakis (1994) state the special case of Theorem 1 for games with rational payoffs and payoff vectors x in $r(A)$. They also state a result for a subset of $\text{co}(r(A))$ with the additional assumption that the bounds on both automata are subexponential in the number of repetitions. The sketched proof in Papadimitriou and Yannakakis (1994) is, however, incomplete; it is tailored for the specific prisoner's dilemma depicted in the introduction, and assumes that $m_1 = m_2$. The detailed proofs appearing in Papadimitriou and Yannakakis (1995, 1996) also have gaps, and we do not see how this line of proof can be validated for general integer payoff matrices that are ordinally equivalent to the prisoner's dilemma. For details, see §6.3. Other related results include Megiddo and Wigderson (1986) and Zemel (1989).

The conclusion of Theorem 1 fails if we replace in the assumptions of the theorem the strict inequality $x^1 > u^1(G)$ by the weak inequality $x^1 \geq u^1(G)$. For example, in the game

0, 4	1, 3
1, 1	1, 0

the only equilibrium payoff in $G^T(m_1, m_2)$ with $m_2 \geq 2^T - 1$ is $(1, 1)$. Indeed, if (σ, τ) is an equilibrium of $G^T(m_1, m_2)$, $r^1_T(\sigma, \tau) \geq 1$, and therefore for every $1 \leq t \leq T$, $r^1_t(\sigma, \tau) = 1$ with probability (with respect to the probability on plays induced by the mixed strategy pair (σ, τ)) 1. If $r^2_T(\sigma, \tau) > 1$, let S, $1 \leq S \leq T$, be the largest positive integer $t \leq T$ such that with positive probability $r(\omega_t(\sigma, \tau)) = (1, 3)$. It follows that for $S < t \leq T$, $r^2(\omega_t(\sigma, \tau)) \leq 1$. Therefore, if τ' is the strategy of Player 2 which coincides with the strategy τ up to stage $S - 1$, and for $t \geq S$ plays Left, $r^2_T(\sigma, \tau') > r^2_T(\sigma, \tau)$. As $\Sigma^2(T, 2^T - 1) = \Sigma^2(T, \infty)$, the result follows. However, the theorem remains intact if we replace the assumption $x^1 > u^1(G)$ with the weak inequality $x^1 \geq u^1(G)$, and in addition we assume that there is a vector payoff $y \in \mathrm{co}(r(A))$ with $y^i > u^i(G)$, $i = 1, 2$.

The next theorem is obviously a generalization of Theorem 1. We do state both theorems because we believe that Theorem 1 is of independent interest and its proof avoids a few complications that arise in the proof of Theorem 2. Also a first reading of the proof of Theorem 1 will help in the reading of the proof of Theorem 2. For simplicity, the statements of the next two theorems are nonsymmetric with respect to the two players.

THEOREM 2. *Let $G = (\{1, 2\}, A, r)$ be a two person game in strategic form. Then for every $\varepsilon > 0$ sufficiently small, there are positive integers T_0 and m_0, such that if $T \geq T_0$ and $x \in \mathrm{co}(r(A))$ with $x^1 > v^1(G)$, and $x^2 > u^2(G)$, and*

$$m_0 \leq m_1 \leq \min(m_2, \exp(\varepsilon^3 T)),$$

there is $y \in E(G^T(m_1, m_2))$ with

$$|y^i - x^i| < \varepsilon.$$

THEOREM 3. *Let $G = (\{1, 2\}, A, r)$ be a two person game in strategic form. Then for every $\varepsilon > 0$ sufficiently small, there are positive integers m_0 and T_0 such that if $T \geq T_0$ and x is a point in $\mathrm{co}(r(A))$ with $x^i \geq v^i(G)$, and*

$$m_0 \leq m_1 \leq m_2 \leq \exp(\varepsilon^3 \min(T, m_1)),$$

then there is $y \in E(G^T(m_1, m_2))$ with

$$|y^i - x^i| < \epsilon.$$

The equilibrium strategies in Theorems 2 and 3 are robust in the following sense. Assume that, in addition to the assumptions (on G, ϵ, T, x, m_1 and m_2) in each theorem, there are action pairs $a = (a^1, a^2) \in A$ and $(a^1, b^2) \in A$, such that $r^2(a^1, b^2) > r^2(a^1, c^2)$ whenever $b^2 \neq c^2$, and $r^2(a) > v^2(G)$. Then there is a strategy pair (σ, τ) in $G^T(m_1, m_2)$, with

$$|r^i_T(\sigma, \tau) - x^i| < \varepsilon$$

which is an equilibrium of $H^T(m_1, m_2)$ for every two person game H with payoffs that are within ε of the payoffs in G.

We are ready now to state our main theorem, which relates the equilibrium payoffs of $G^T(m_1, m_2)$ to the equilibrium payoffs of the undiscounted infinitely repeated game G^*_∞. Recall that the Folk Theorem asserts that

$$E(G^*_\infty) = \{x \in \mathrm{co}(r(A)) | x^1 \geq v^1(G) \text{ and } x^2 \geq v^2(G)\}.$$

MAIN THEOREM. *Let* $G = (\{1, 2\}, A, r)$ *be a two person game in strategic form, and let* $(T, m_1(T), m_2(T))_{T=1}^{\infty}$ *be a sequence of triples of positive integers with* $\min_{i=1,2} m_i(T) \to \infty$ *as* $T \to \infty$, *and*

$$\lim_{T\to\infty} \frac{\log \max_{i=1,2} m_i(T)}{\min(m_1(T), m_2(T), T)} = 0.$$

Then,

$$\lim_{T\to\infty} E(G^T(m_1(T), m_2(T))) = E(G^*_\infty).$$

The inequality $m_2 \leq \exp(\epsilon^3 m_1)$ in Theorem 3, could probably be replaced with an alternative lower bound, as a function of T, on m_1, provided that we also replace the weak inequality $x^1 \geq v^1(G)$ with the strict inequality $x^1 > v^1(G)$.

CONJECTURE 1. *Let* $G = (\{1, 2\}, A, r)$ *be a two person game in strategic form. Then for every* $\epsilon > 0$ *sufficiently small, there is a positive integer* T_0 *such that if* $T \geq T_0$ *and* x *is a point in* $\text{co}(r(A))$ *with* $x^i > v^i(G)$, *and*

$$\varepsilon T \leq m_1 \leq \min(m_2, \exp(\varepsilon^3 T)),$$

then there is $y \in E(G^T(m_1, m_2))$ *with*

$$|y^i - x^i| < \varepsilon.$$

The next theorem is straightforward and very easy. We state it as a contrast to the previous results. It shows that the subexponential bounds on the sizes of the automata as a function of the number of repetitions is essential to obtain equilibrium payoffs that differ from those of the finitely repeated game G^T.

THEOREM 4. *For every game* G *in strategic form there exists a constant* c *such that if* $m_i \geq \exp(cT)\ \forall i$ *then*

$$E(G^T(m_1, \ldots, m_n)) = E(G^T).$$

6. Preliminaries.

6.1. Notation. Let $G = (N; A; r)$ be an N-player game in strategic form. We denote by $K(G)$ or K, for short, twice the largest absolute value of a payoff in the game G. Thus $r^i(a) - r^i(b) \leq K(G)$ for every a and b in A. We denote by R the vector valued function defined on all finite plays as the average payoff, i.e., for $a = (a_1, \ldots, a_n) \in A^n$,

$$R(a) = R((a_1, \ldots, a_n)) = \frac{r(a_1) + \cdots + r(a_n)}{n}.$$

The integer part of a real number x is denoted $[x]$, i.e., $[x]$ is the largest integer that is less than or equal to x. The length n of a play $c = (c_1, \ldots, c_n) \in A^n$ is denoted $|c|$. The number of elements in a set X is denoted $|X|$. Given two sets of real numbers X and Y, $X + Y = \{x + y \mid x \in X, y \in Y\}$.

There are several constructions and functions that are used repeatedly in our proofs. Therefore, we introduce here several of these as general notations. Given two positive integers m_1 and l, we define the nonnegative integer $k(m_1, l)$ as the smallest nonnegative integer k such that $2^k l > m_1 - l$. The set $Q(m_1, l)$ is defined as a subset Q of $\{0, 1\}^{k(m_1,l)}$ satisfying:

$$Q \supset \{(\epsilon_1, \ldots, \epsilon_{k-1}, 0) : \epsilon_i \in \{0, 1\}\} \qquad (1, \ldots, 1) \notin Q$$

$$|Q| = [(m_1 - l)/l].$$

Note that the definition of $k = k(m_1, l)$ implies that $2^{k-1} l \leq (m_1 - l) < 2^k l$, and thus $2^{k-1} \leq [(m_1 - l)/l] < 2^k$. Therefore, such a subset Q of $\{0, 1\}^k$ exists. Note also that for every $(\epsilon_1, \ldots, \epsilon_k) \in Q$, $\sum_{i=1}^{k} \epsilon_i < k$. In the constructed equilibrium, $|Q|$ is the number of possible proposed plays, $2k$ is the duration of the communication phase, and l is the length of the cycle following the communication phase. The role of the first two properties of Q is to simplify the description of the proposed play.

In each of the constructed equilibria two actions of each player are labeled 0 and 1. To each $\epsilon \in Q$ we will associate two plays of length $2k$: the communication play $\overline{\theta}(\epsilon)$ and the verification play $\theta^*(\epsilon)$.

6.2 The idea of the proofs of the main results.

(a) for the Prisoner's Dilemma and $m_1 < T$.

We outline the proof in a few special cases where G is the Prisoner's Dilemma given in the introduction. We consider first two instances in which the payoff vector is $x = (3, 3)$. We exhibit a pure strategy equilibrium (σ, τ) of $G^T(m_1, m_2)$ where $2 \leq m_1 \leq m_2 < T$, resulting in a payoff of 3 for each player in each stage. Label the Friendly action by 0, and the Unfriendly action by 1. For any play $\omega = (\omega_1, \ldots, \omega_T)$ let σ^ω and τ^ω be the strategies that follow the play ω as long as the other player follows it, and switch to punishing forever as soon as a deviation from the proposed play is observed. Notice that ω is the outcome of the strategy pair $(\sigma^\omega, \tau^\omega)$. Now let $\omega = T * (0, 0)$. The only play with average payoff greater than 3 to Player 2 that is compatible with the strategy σ^ω is $(T - 1) * (0, 0) + (0, 1)$. By Lemma 3 such a play requires a strategy with complexity at least T. Obviously, $\sigma^\omega \in \Sigma^1(2) \subset \Sigma^1(m_1)$ and $\tau^\omega \in \Sigma^2(m_2)$ and thus $(\sigma^\omega, \tau^\omega)$ is an equilibrium of $G^T(m_1, m_2)$ with outcome ω.

Now fix δ with $0 < \delta < 1$. Let T be sufficiently large, $m_1 \leq m_2$, and $m_1 < \delta T$. We now construct a play ω with average payoff within ϵ of the friendly payoff $(3, 3)$ such that the pure strategy pair $(\sigma^\omega, \tau^\omega)$ is an equilibrium of $G^T(m_1, m_2)$. Choose an integer d sufficiently large that $R(d * (0, 0) + (1, 0) + (0, 1))$ is within $\varepsilon/2$ of $(3, 3) = r(0, 0)$ and that $[(T - 1)/(d + 2)](d + 1) > \delta T$. Let ω be the $d + 2$ periodic play of G^T where the last string of $d + 2$ actions is $d * (0, 0) + (1, 0) + (0, 1)$. For T sufficiently large $R(\omega)$ is within ϵ of x. The strategy pair, σ^ω of Player 1 and τ^ω of Player 2, is an equilibrium of $G^T(m_1, m_2)$. Indeed, the complexity of each of these strategies is no more than $d + 3 \leq m_1 \leq m_2$, implying that $\sigma^\omega \in \Sigma^1(T, m_1)$ and $\tau^\omega \in \Sigma^2(T, m_1)$. No other play which is compatible with σ^ω results in a higher payoff to Player 2. And the only play θ with $R^1(\theta) > R^1(\omega)$ that is compatible with τ^ω is $(\omega_1, \ldots, \omega_{T-1}, (1, 1))$, which by Lemma 4 has complexity for Player 1 at least $[(T - 1)/(d + 2)](d + 1) > \delta T$. Thus any strategy of Player 1 that does better than σ^ω is not in $\Sigma^1(T, m_1)$, and so the strategy pair $(\sigma^\omega, \tau^\omega)$ is an equilibrium of $G^T(m_1, m_2)$.

The role of the action pair $(0, 1)$ at the end of the cycle together with the synchronization of the play so that $\omega_T = (0, 1)$ is to ensure that Player 2 has no desire to deviate. The role of the action pair $(1, 0)$ following each d plays of the action pair $(0, 0)$ is to ensure that any strategy of Player 1 that follows the play ω up to close to the end and then deviates is sufficiently complex, as illustrated in Lemma 4 and Remark 1. For example, the play $\omega = 1000 * (4 * (0, 0) + (0, 1))$ is not an equilibrium play of the 5,000 repeated prisoners' dilemma for $m_1 = 1003$. Consider the strategy of Player 1 implemented by an automaton with 1003 states: one absorbing state in which the Unfriendly action is played and 1002 states in which the Friendly action is taken. The automaton starts in state one and until state 999 moves to the next state whenever the other player plays the unfriendly action and stays in its present state if the other player plays the friendly action. For states 1000 to 1002 the automaton moves to the next state whatever the other player does. Against any strategy of Player 2 consistent with ω this strategy of Player 1 follows the play ω up to stage 4998 and plays the unfriendly action in the last 2 stages of the game, which is better for Player 1 than ω.

(b) for general games and $m_1 < T/4$.

Now consider an arbitrary game G and suppose that $m_1 < T/4$. For any vector payoff $x \in \operatorname{co} r(A)$ with $x^i > u^i(G)$ we will construct a d-periodic play ω with a corresponding payoff of approximately x. Further, the play ω is such that the strategy τ^ω is a best reply to σ^ω among all strategies of Player 2, and σ^ω is a best reply to τ^ω among all strategies in $\Sigma^1(T, m_1)$.

We start by constructing the cycle $c = d_1 * a_1 + b_1 + d_2 * a_2 + b_2 + d_3 * a_3 + b_3 + d_4 * a_4$ with $r^2(a_1) \leq r^2(a_2) \leq r^2(a_3) \leq r^2(a_4) = \max_{a \in A} r^2(a)$, d_4 sufficiently large, $R(c)$ is approximately x, and $\max_{i=1}^{3} d_i/|c| > \frac{1}{4}$. Two of the terms b_i are the empty string. The b_i following the longest string $d_i * a_i$, $1 \leq i \leq 3$, is an action pair with $b_i^1 \neq a_i^1$. The play ω is the periodic play of c such that the play of the last $|c|$ stages is c. The cycle, and thus also the play ends with a string $d_4 * a_4$. Since $r^2(a_1) \leq r^2(a_2) \leq r^2(a_3)$ this implies that τ^ω is a best reply to σ^ω. Indeed, except for one stage in the cycle the payoffs of Player 2 increase through the cycle. The value of d_4 is sufficiently large that the contribution of this stage is negligible and that any one stage gain from deviation before the last d_4 stages is offset by the loss in the remaining stages. The role of the action pairs b_i, $1 \leq i \leq 3$, is to ensure that the complexity of a strategy of Player 1 that deviates from the proposed play after completing k cycles c, is at least $k \max_{1 \leq i \leq 3} d_i$. As any play θ with $R^1(\theta) > R^1(\omega)$ that is compatible with τ^ω, coincides with ω in all but the few very last stages we conclude that indeed $r^1(\sigma^\omega, \tau^\omega) \geq r^1(\sigma, \tau^\omega)$ for any $\sigma \in \Sigma^1(T, m_1)$.

So far we have described instances of cases in which we were able to describe a pure strategy equilibrium of $G^T(m_1, m_2)$. For sufficiently large m_i, $\Sigma^i(T, m_i) = \Sigma^i(T)$, and so all equilibrium payoffs of $G^T(m_1, m_2)$ are equilibrium payoffs of G^T. For pure strategy equilibria significantly smaller bounds on m_i suffice: if m_1 and m_2 are both at least T, any pure strategy equilibrium payoff of the repeated prisoner's dilemma results in repeated play of the unfriendly actions.

(c) for $m_i \geq T/4$.

When $m_i \geq T/4$ we will construct mixed strategy equilibria. We now describe the outline of the proof in case that $m_1 > T/4$ and $m_2 > T$.

For any subset of plays $Q \subset A^T$, let τ^Q be the mixed strategy of Player 2 that is a mixture of the pure strategies τ^ω, $\omega \in Q$, each equally likely. Recall that a set of plays Q is conforming for Player 1 if the actions of Player 1 at each stage are functions of past actions, i.e., for any two plays $\omega, \theta \in Q$ and any $1 \leq t \leq T$, $\theta_t^1 = \omega_t^1$ whenever $(\omega_1, \ldots, \omega_{t-1}) = (\theta_1, \ldots, \theta_{t-1})$. A pure strategy σ of Player 1 is conformable to Q if and only if for every $\omega \in Q$ the outcome of the strategy pair (σ, τ^ω) is ω. For any set of plays Q that is conforming to Player 1, there is a pure strategy of Player 1 that is conformable to Q.

The equilibrium strategy of Player 2 is a mixed strategy of the form τ^Q where Q is a set of plays that is conforming for Player 1. The equilibrium strategy of Player 1, σ^*, is a mixture of strategies that are conformable to Q. The play ω that is selected by τ^Q via the realization τ^ω is called the proposed play. We will construct the set Q so that the number of plays in Q is at least 2^{k-1} and less than 2^k (where the value k depends on the parameters m_1 and T). Therefore we rename the set of the possible proposed plays and identify each play with a sequence $\epsilon = (\epsilon_1, \ldots, \epsilon_k)$ of zeros and ones. We use the notation $\omega(\epsilon) = (\omega_1(\epsilon), \ldots, \omega_T(\epsilon))$ for the proposed play associated with ϵ, and τ^ϵ for $\tau^{\omega(\epsilon)}$.

The action of Player 1 at stage $1 \leq t \leq k$ in the proposed plays $\omega(\epsilon) \in Q$ is independent of ϵ and the sequence of actions of Player 2 in the first k stages of the game identifies the proposed play $\omega(\epsilon)$. This implies that Q is conforming for Player 1.

We interpret the actions of Player 2 in the first k stages of the game as a signal that Player 2 sends to Player 1. The signal specifies one of finitely many plays of the repeated game. Each player follows the proposed play as long as the other player follows it, and switches to punishing the other player as soon as he detects a deviation from the proposed play. The strategy of Player 2 detects immediately any deviation by Player 1 from the proposed play.

Each one of the proposed plays enters a cycle of action pairs with associated payoff approximately x. Thus, in any one of the proposed plays, Player 1 has no incentive to deviate prior to the very last stages of the finitely repeated game. The set of possible proposed plays is such that Player 1 is unable to follow each one of the proposed plays and deviate in even one of them at the very late stages of the finitely repeated game.

There are several properties that need to be satisfied in order to construct such an equilibrium. The resulting expected payoff needs to be approximately the fixed payoff vector x. This will be achieved by each one of the proposed plays separately, by playing in most stages a cycle of action pairs where the average payoff over the cycle is approximately x. (When $x = r(a)$, $a \in A$, most of the action pairs in the cycle will be a.)

As Player 2 specifies the proposed play, the payoff to Player 2 needs to be independent of the proposed play. This will be accomplished by Player 2 sending his signal during the first k stages of the repeated game, using two actions, say 0 and 1. The next k stages are used to balance the number of times that each one of these two actions of Player 2 appears in the first $2k$ stages, i.e., ensuring that each one of them appears exactly k times. Player 1 plays a fixed action, say 0 during the first $2k$ stages. We refer to the first $2k$ stages as the communication phase. Following the communication phase the play will enter one of finitely many possible cycles where the number of times each action pair appears in the different cycles is constant and moreover, the leftovers needed to complete the full play are independent of the proposed play.

Many simplifications of the ideas result by assuming $x = r(a) \in r(A)$ rather than $x \in \mathrm{co}\, r(A)$. Label the action pair $a = (a^1, a^2)$ by $(0, 0)$ and the action pair of the punishing strategies $(1, 1)$. Following the communication phase (of length $2k$) in which Player 2 transmits a message $\epsilon \in Q$ to Player 1 (during the first k stages), the play enters a cycle $c(\epsilon)$, which is a coordinated play. The length of the coordinated play, $|c(\epsilon)|$, is independent of ϵ and $|c(\epsilon)| \gg 2k$. Most of the action pairs in the coordinated play $c(\epsilon)$ are $(0, 0)$. However, the last $2k$ stages of the coordinated play $c(\epsilon)$ depend on ϵ. This part is called the verification play (and is denoted $\theta^*(\epsilon)$).

The set of messages Q and the corresponding verification play are such that an automaton of Player 1 which wishes to follow each one of the proposed cycles for at least two rounds can not use the same states in different cycles (corresponding to different messages) and moreover, it needs $|c|$ different states for each cycle. We thus fill up his capacity by generating enough messages so that $m_1 - |Q||c|$ is sufficiently small (no more than $|c|$) preventing him from selecting even just one proposed play and being able to deviate in the last stage in that play while repeating the cycle in all other proposed plays. One needs further to choose the set of proposed plays so that Player 1 is unable to increase his own payoff by neglecting a subset of messages and using the freed upon states for sufficient gain in the late stages of the game. We label the automaton states used in playing the coordinated cycle $c(\epsilon)$ by $(\epsilon, 1), \ldots, (\epsilon, |c|)$.

Player 1 also has to process the message sent during the communication phase, and for that it might seem that an additional number of about $|Q|$ states is needed. However, the constructed equilibrium is such that he uses the very same states used to follow the different cycles.

An additional problem arises. If a state of the automaton of Player 1 is used both in one of the cycles as well as in the communication phase, it may have to tolerate the two actions 0 and 1 of Player 2. Therefore, there are deviations by Player 2 from the set of plays Q that are left unpunished. Such deviations can start during the communication phase, e.g., Player 2 can choose the action 0 at a stage during the communication phase that dictates to Player 2 the action 1. Or, they can start after completing the communication phase during a play of a cycle, by taking the action 1 instead of 0 at a stage in which the state of the automaton of Player 1 tolerates both actions 0 and 1 of Player 2. Therefore, if Player 2 knows an exact stage in one of the cycles in which the state of the automaton of Player 1 tolerates the two actions 0 and 1, he may take advantage over it in future stages of the game. Moreover, the availability of undetectable deviations by Player 2 may alter his indifference among the various plays in Q. Thus the mixed strategy of Player 1 conceals the exact states that tolerate both actions, 0 and 1. But how can the play enter now the coordinated play; after all, Player 2 has to be informed of the time Player 1 enters the start of the verification play. This is accomplished by ensuring that the communication phase ends with an action pair $(0, 1)$ in a state of the automaton which is reused in the coordinated cycle for a $(0, 0)$ action pair, and the transition function of the automaton of Player 1 will change the automaton state to $(\epsilon, 1)$. The above discussion indicates one role of a mixed strategy of Player 1 in our constructed equilibrium.

Additional caution is needed here. Player 1 needs also to verify that Player 2 does indeed balance his two actions 0 and 1 during the first $2k$ stages of the repeated game. Otherwise, as Player 2 may prefer the action pair $(0, 0)$ to the action pair $(0, 1)$ he may choose to play the action 0 rather than 1 in some stages t with $k + 1 \leq t \leq 2k$. Such a deviation by Player 2 has a negligible effect on the total payoff but may not leave Player 2 indifferent to the different proposed plays. The states of the automaton of Player 1 that are used in stages $t = k + 1, \ldots, 2k$ are also used at later stages of the repeated game (during the cycle play) for the action pair $(0, 0)$. How can Player 1 verify that Player 2 does indeed balance his two actions in the first $2k$ stages of the repeated game? The automaton transition function is such that if Player 2 takes action 0 when in a reused state (ϵ, j), the next state of the automaton is $(\epsilon, j + 1)$ as expected in the course of the cycle play. The different pure strategies in the support of the equilibrium (mixed) strategy of Player 1 react differently in the reused states to the action 1 of Player 2. Player 2 is uncertain about the pure strategy realization, and therefore if Player 2 deviates from the proposed play at one of the stages $t = k + 1, \ldots, 2k$ he will not know when to start with the verification play. The above discussion indicates a second role of a mixed strategy of Player 1 in our constructed equilibrium.

Player 1 may have an incentive to accept only a subset of the proposed plays. This can happen if the number of the proposed plays is large and by neglecting $o(T)$ proposed plays he is able to follow most other proposed plays up to stage T and deviate at stage T. This is excluded by our construction, by requiring Player 1 to report periodically the proposed play.

6.3. An alternative idea of a proof. Papadimitriou and Yannakakis (1995, 1996) proposed a different line of proof for games with integer payoffs and vector payoffs x in $r(A)$.

Consider the prisoner's dilemma depicted in the following figure:

	C	D
C	$7, 7$	$0, 8$
D	$8, 0$	$2, 2$

Denote by $g = (g^1, g^2)$ the vector payoff function of this game. We describe in what follows the ''equilibrium'' strategies and equilibrium play (proposed in Papadimitriou and Yannakakis 1995, 1996) of the T-stage repetition of this game where Player 1 is restricted to automata of size m_1 ($>T^2$) and Player 2 is restricted to automata of size $m_2 \geq m_1$ with a payoff vector that is close to $g(C, C) = (7, 7)$.

Our discussed communication phase is replaced in Papadimitriou and Yannakakis with two phases: the business card exchange phase and the fixup phase. The first d stages are called (by Papadimitriou and Yannakakis) the business card exchange phase. Player 2 transmits his business card, $y \in \{C, D\}^d$, (by playing at stage $1 \leq t \leq d$ the action y_t,) to Player 1, who plays the (same pure) action C through this phase. All business cards of Player 2 are equally likely.

The quotes in the following paragraph are from Papadimitriou and Yannakakis (1996). The second phase, called (by Papadimitriou and Yannakakis) the fixup phase, ''has the purpose of equalizing the value of the different business cards; it consists of one random step by both players, followed by a sequence of deterministic coordinated moves by the two players that balance the payoffs.'' The coordinated moves in their proof are (C, C) and (D, D). (''In the first step the two players choose randomly, but possibly with different probabilities, among two (distinct) strategies, which we shall still call C and D. . . Then they go on to play for L more steps using only two strategy pairs, which we call (A, B), (A', B'); . . . The only requirement is that the quantities $\Delta = g^2(A, B) - g^2(A', B')$ and $\Delta' = g^1(A, B) - g^1(A', B')$ are both nonzero.'' Thus we may assume that $A = A' = C$ and $B = B' = D$.) ''The number of AB (i.e., (C, C)) steps they play, is determined by the outcome of the random step: If Player 1 played C and Player 2 also C, then it is x_1; if the combination was C, D then x_2; if D, C, then x_3; and if D, D, then x_4.''

Thus, the fixup phase last for $L + 1$ stages and is parametrized by a vector $x = (x_1, \ldots, x_4)$ (which depends on past actions) with $\max x_i \leq L$.
Let

$$\Delta = g^2(C, C) - g^2(D, D).$$

In the first stage of the fixup phase, Player 1 (Player 2) plays C with probability $p = p(x)$ ($q = q(x)$). In the next L stages of the fixup phase the players play a deterministic sequence of coordinated action pairs. If the random action pair in the first stage of the fixup phase is (C, C) they play $x_1 * (C, C) + (L - x_1) * (D, D)$; if the action pair was (C, D) then $x_2 * (C, C) + (L - x_2) * (D, D)$; if the action pair was (D, C) then $x_3 * (C, C) + (L - x_3) * (D, D)$; and if the action pair was (D, D) then $x_4 * (C, C) + (L - x_4) * (D, D)$. Recall that x_i are integers.

Player 2 is indifferent about his two possible actions in the first stage of the fixup phase only if

$$px_1\Delta + pg^2(C, C) + (1 - p)x_3\Delta + (1 - p)g^2(D, C)$$

equals

$$px_2\Delta + pg^2(C, D) + (1 - p)x_4\Delta + (1 - p)g^2(D, D).$$

Setting

$$\alpha(x) = (x_4 - x_3)\Delta + g^2(D, D) - g^2(D, C) \quad \text{and}$$

$$\beta(x) = (x_1 - x_2)\Delta + g^2(C, C) - g^2(C, D),$$

Player 2 is indifferent only if

$$p(x) = \frac{\alpha(x)}{\alpha(x) + \beta(x)}.$$

Let $F(x)$ denote the sum of the payoffs to player 2 in the fixup phase, and $G(x) = F(x) - Lg^2(D, D)$. We write next the expression for $G(x)$ and $G(y)$ for two possible values of the vectors $x = (x_1, \ldots, x_4)$ and $y = (y_1, \ldots, y_4)$.

$$G(x) = \frac{\alpha(x)x_1\Delta}{\alpha(x) + \beta(x)} + \frac{\beta(x)x_3\Delta}{\alpha(x) + \beta(x)} + \frac{\alpha(x)g^2(C, C)}{\alpha(x) + \beta(x)} + \frac{\beta(x)g^2(D, C)}{\alpha(x) + \beta(x)}$$

and

$$G(y) = \frac{\alpha(y)y_1\Delta}{\alpha(y) + \beta(y)} + \frac{\beta(y)y_3\Delta}{\alpha(y) + \beta(y)} + \frac{\alpha(y)g^2(C, C)}{\alpha(y) + \beta(y)} + \frac{\beta(y)g^2(D, C)}{\alpha(y) + \beta(y)}.$$

Let $B = g^2(C, C) + g^2(D, D) - g^2(C, D) - g^2(D, C)$. Notice that for every integer vector x, $\alpha(x) + \beta(x) = B(\text{mod } \Delta)$ and that $\alpha(x) = \alpha(y)(\text{mod } \Delta)$ and $\beta(x) = \beta(y)(\text{mod } \Delta)$. Therefore

$$F(x) - F(y) = G(x) - G(y) = \frac{N\Delta}{M\Delta + B^2}$$

where N and M are integers. Therefore if the imbalance in the first phase of the game is s and sB^2 is not a multiple of Δ, there are no integer values of the vectors x and y with $F(x) - F(y) = s$.

6.4. Zero-sum games with finite automata. In this section we present results about the value of 2-person 0-sum repeated games with finite automata. The first result follows from the proof of the result of Ben-Porath (1993), and is used in our proof of Theorems 2 and 3. In all that follows in this section we denote by G a fixed 2-person 0-sum game, and for a 2-person 0-sum game, H, we denote by $\text{Val}(H)$ its minimax value.

THEOREM 5. *Let $\delta > 0$. For every $\varepsilon > 0$ sufficiently small, if*

$$\exp(\varepsilon^{2+\delta} m_1) \geq m_2 > 1,$$

then for every positive integer T,

$$\text{Val}(G^T(m_1, m_2)) \geq \text{Val}(G) - \varepsilon.$$

The next result asserts that if the bound on the sizes of the automata of Player 2 is larger than an exponential of the size of the automata of Player 1, then Player 2 could hold Player 1 down to his maxmin in pure strategies.

THEOREM 6. *For every $K > |A^1|$, m_1 and $\epsilon > 0$, there exist a positive integer T_0 and a strategy $\tau^* \in \Delta(\Sigma^2(m_2))$ where $m_2 \geq K^{m_1}$, such that for every $T \geq T_0$, and any strategy $\sigma \in \Sigma^2(m_1)$,*

$$\text{Val}(G^T(m_1, m_2)) \leq r_T^1(\sigma, \tau^*) \leq \max_{a^1 \in A^1} \min_{a^2 \in A^2} r^1(a^1, a^2) + \epsilon.$$

PROOF. The proof is given here for completeness; it is given also in Neyman (1997). Without loss of generality we assume that $G = (\{1, 2\}, A, r)$ is a 2-player 0-sum game.

The idea of the proof is as follows: there is a subset $\mathcal{T} \subset \Sigma^2(m_1)$ with $|\mathcal{T}| = m_1|A^1|^{m_1}$ such that for every strategy $\sigma \in \Sigma^1(m_1)$ there is a strategy $\tau \in \mathcal{T}$ such that for every positive integer t, $\omega_t^2(\sigma, \tau)$ is a best reply of Player 2 to $\omega_t^1(\sigma, \tau)$ and therefore $r_T^1(\sigma, \tau) \leq \max_{a^1 \in A^1} \min_{a^2 \in A^2} r^1(a^1, a^2)$. Player 2 chooses at random a strategy from $\mathcal{T}$ and switches to another randomly chosen strategy from $\mathcal{T}$ if it does not fit. The conditional probability of success at each round is at least $1/(m_1|A^1|^{m_1})$ and therefore the probability of success in one of the first $Cm_1|A^1|^{m_1}$ rounds approaches 1 as $C \to \infty$, and the result follows. Formally, let $b: A^1 \to A^2$ be a selection from the best reply correspondence of Player 2. Construct the following mixed strategy of Player 2, τ^*, which is implemented with an automation with state space

$$M^2 = \{1, \ldots, m_1\} \times \{1, \ldots, l\}$$

where $l = [K^{m_1}/m_1]$. The initial state of the automation of Player 2 is $(1, 1)$. Let $a: M^2 \to A^1$ be a random function, each such function equally likely, i.e., for every $1 \leq i \leq m_1$, and every $1 \leq j \leq l$, $a(i, j)$ is a random element of A^1 each one equally likely, and the various $a(i, j)$ are independent. We define now the random action function of the automation

$$f^2(i, j) = b(a(i, j)).$$

The transition function of the automation, depends on a random sequence $k = k_1, \ldots, k_l$ each such sequence equally likely and the sequence is independent of the function a. We are ready now to define the transition function which depends on the functions b and a and the random sequence k.

$$g^2((i, j), c) = \begin{cases} (i + 1, j) & \text{if } i < m_1 \text{ and } c = a(i, j), \\ (k_j, j) & \text{if } i = m_1 \text{ and } c = a(i, j), \\ (1, j + 1) & \text{if } j < l \text{ and } c \neq a(i, j), \\ (1, 1) & \text{otherwise.} \end{cases}$$

Let σ be a pure strategy of Player 1 that is implemented by an automaton of size m_1. Let $a_1, a_2, \ldots$ where $a_t = (a_t^1, a_t^2)$ be the random play induced by the strategy pair σ and τ^*, and let $q_1^i, q_2^i, \ldots$ be the random sequence of states of the automation of Player i. Fix $1 \leq j \leq l$ and let $t = t_j$ be the random time of the first stage t with $q_t^2 = (1, j)$. Note that

$$\text{Prob}(a_{t+s}^1 = a(s + 1, j) \; \forall \; 0 \leq s < m_1) = \frac{1}{|A^1|^{m_1}},$$

and if $a_{t+s}^1 = a(s + 1, j) \; \forall \; 0 \leq s < m_1$ (and thus also $a_{t+s}^2 = b(a(s + 1, j)) \; \forall \; 0 \leq s < m_1$) then either there are $0 \leq s < s' < m_1$ with $q_{t+s}^1 = q_{t+s'}^1$ and then $q_{t+m_1}^1 = q_{t+m_1+s-s'}^1$ and therefore $q_{t+m_1}^1 \in \{q_t^1, \ldots, q_{t+m_1-1}^1\}$ or $|\{q_t^1, \ldots, q_{t+m_1-1}^1\}| = m_1$ and then $q_{t+m_1}^1 \in \{q_t^1, \ldots, q_{t+m_1-1}^1\}$. In either case there exists $0 \leq s < m_1$ such that the state of the automation of Player 1 at stage $t + m_1$, $q_{t+m_1}^1$ coincides with its state in stage $t + s$. Therefore if $k_j = s + 1$, the play will enter a cycle in which the payoff to Player 1 is at most $\max_{a^1 \in A^1} \min_{a^2 \in A^2} r^1(a^1, a^2)$. Therefore the conditional probability, given the history of play up to stage t_j that the payoff to Player 1 in any future stage is at most $\max_{a^1 \in A^1} \min_{a^2 \in A^2} r^1(a^1, a^2)$, and that $t_{j+1} = \infty$ is at least $1/(|A^1|^{m_1} m_1)$. Otherwise, if $t_{j+1} < \infty$, for every $t_j \leq t < t' < t_{j+1}$ $(q_t^1, q_t^2) \neq (q_{t'}^1, q_{t'}^2)$ and therefore $t_{j+1} \leq t_j + m_1^2$.

Therefore, if $t_l = \infty$, for every stage $t > lm_1^2$, the payoff to Player 1 at stage t is at most $\max_{a^1 \in A^1} \min_{a^2 \in A^2} r^1(a^1, a^2)$. The definition of l and the previous inequalities imply that

$$\text{Prob}(t_l = \infty) \geq 1 - (1 - 1/(m_1 |A^1|^{m_1}))^{l-1} \to 1 \quad \text{as } m_1 \to \infty,$$

which completes the proof of the theorem. □

A positive resolution of the next conjecture would provide a positive answer to conjecture 1.

CONJECTURE 2. *For every* $\varepsilon > 0$,

$$\text{Val}(G^T(m_1, \infty)) \to \text{Val}(G) \quad \text{as } m_1 \to \infty \text{ and } m_1 \geq \varepsilon T.$$

The truth of the above conjecture implies that there is a function $h: N \to N$ with $\lim_{T\to\infty} h(T)/T = 0$, and such that

$$\lim_{T\to\infty} \text{Val}(G^T(h(T), \infty)) = \text{Val}(G). \tag{3}$$

An interesting open problem is to find the order of magnitude of the smallest function h obeying (3). Neyman and Okada (1996) show that if $h: N \to N$ is such that $\lim_{T\to\infty} h(T)/(T/\log T) = 0$, then

$$\lim_{T\to\infty} \text{Val}(G^T(h(T), \infty)) = \max_{a^1 \in A^1} \min_{a^2 \in A^2} r^1(a^1, a^2).$$

CONJECTURE 3. *If* $h: N \to N$ *is such that* $\lim_{T\to\infty} (T/\log T)/h(T) = 0$, *then*

$$\lim_{T\to\infty} \text{Val}(G^T(h(T), \infty)) = \text{Val}(G).$$

7. The proofs of the main theorems. Theorem 4 is straightforward, and is a result of the following observation.

PROPOSITION 2. *Let* $G = (N; (A^i)_{i\in N}; (r^i)_{i\in N})$, *be an* N*-person game. Then any pure strategy of Player* i *in* G^T *is implemented by an automation of Player* i *of size* $\sum_{t=1}^{t=T} |A^{-i}|^{t-1}$.

PROOF. Let $M_1 = \{\emptyset\}$, $M_t = (A^{-i})^{t-1}$. Let σ be a pure strategy of Player i in G^T, and let $\alpha = \langle M^i, q^i, f^i, g^i \rangle$ be the automaton of Player i given by:

$$M^i = \bigcup_{t=1}^{t=T} M_t, \qquad q^i = \emptyset,$$

and for every $h \in M_t$, $1 \leq t \leq T$,

$$f^i(h) = \sigma(h'), \qquad g^i(h, a^{-i}) = (h, a^{-i}),$$

where h' is the unique history of length t that is consistent with both h and σ. Then, σ is implemented by the automaton α. □

PROOF OF THEOREM 1. Let $G = (\{1, 2\}, A, r)$ be a 2-person game, $x \in \text{co}(r(A))$ with $x^1 > u^1(G)$, and $x^2 > u^2(G)$, and $\varepsilon > 0$ sufficiently small. Without loss of generality we assume that $m_1 \leq m_2$. Assume that T_0 and m_0 are sufficiently large and that the triple (T, m_1, m_2) satisfies the following inequalities:

$$T \geq T_0,$$

and

$$m_0 \leq m_1 \leq \exp(\varepsilon^3 T).$$

W.l.o.g. we assume that $\varepsilon < 1$. We will construct an equilibrium (σ, τ) of $G^T(m_1, m_2)$ with associated equilibrium payoffs (y^1, y^2) satisfying $|y^i - x^i| < \varepsilon$.

(A) The case $m_1 < T/4$.

For small values of m_1 (relative to the value of T), we will construct a pure strategy equilibrium, and for larger values of m_1 the constructed equilibrium will be a mixed strategy one.

For $m_1 < T/4$ we will define a play $\omega = (\omega_1, \ldots, \omega_T)$ of G^T which is a periodic sequence of elements of A with period d, and such that the pair of strategies, σ of Player 1 and τ of Player 2, that follow the play ω as long as the other player follows the play ω and triggers to punishing forever with a pure strategy as soon as a deviation was observed, is an equilibrium of $G^T(m_1, m_2)$.

We start with the construction of the play ω. Recall that K is twice the largest absolute value of a payoff in the game G, and assume that $\epsilon > 0$ is sufficiently small with $\varepsilon < \min(1, K/4)$, and let $x \in \mathrm{co}(r(A))$ with $x^i > u^i(G)$.

Without loss of generality we assume that $x^i > u^i(G) + 2\varepsilon$. Otherwise, let $y \in \mathrm{co}(r(A))$ with $y^i > u^i(G)$ and without loss of generality assume that $y^1 - u^1(G) = y^2 - u^2(G)$ and ε is sufficiently small so that $\varepsilon < y^i - u^i(G)$, set $\alpha = 2\epsilon/3(y^i - u^i(G))$, $\bar{x} = \alpha y + (1 - \alpha)x$, and $\varepsilon' = \varepsilon/4$. It follows that $\bar{x}^i > u^i(G) + 2\varepsilon'$ and $|x^i - \bar{x}^i| < 2\varepsilon/3$ and therefore any point z with $|z^i - \bar{x}^i| < \varepsilon'$ satisfies $|z^i - x^i| < \varepsilon$. Let d be a fixed positive integer with $d > 3(K/\varepsilon)^2$.

There are three strategy pairs a_1, a_2 and a_3 in A and three nonnegative numbers α_1, α_2 and α_3 with $\sum_{j=1}^{3} \alpha_j = 1$ such that $x = \sum_{j=1}^{3} \alpha_j r(a_j)$. Let $a_4 \in A$ be a point that maximizes the payoff to Player 2, i.e., $r^2(a_4) \geq r^2(a)$ for every $a \in A$. Let d_4 be the smallest positive integer so that $d_4(r^2(a_4) - u^2(G)) > K$. In particular $(d_4 - 1)2\varepsilon < K$, i.e., $d_4 < 1 + K/2\varepsilon$ and therefore $d_4 K/d < \varepsilon/6 + K/d$. Setting $d_1 = [\alpha_1(d - d_4 - 1)]$, $d_2 = [\alpha_2(d - d_4 - 1)]$ and $d_3 = d - d_1 - d_2 - d_4 - 1$, we deduce that $|\sum_{j=1}^{3} d_j(r^i(a_j) - x^i)| \leq 2K$ and therefore

$$\left| \frac{d_1 r^i(a_1) + d_2 r^i(a_2) + d_3 r^i(a_3) + d_4 r^i(a_4)}{d} - x^i \right| < \varepsilon/2 - K/d.$$

Without loss of generality we assume that $r^2(a_1) \leq r^2(a_2) \leq r^2(a_3) \leq r^2(a_4)$, and let $1 \leq k \leq 3$ be such that $d_k = \max_{j=1}^{3} d_j$. It follows that $d_k > 2d/7$. Let $b_k \in A$ be such that $b_k^1 \neq a_k^1$. Define the periodic play $\omega = (\omega_1, \ldots, \omega_T)$ with period d as follows. The play in the last d stages consists of 5 strings. Four of the strings consist of d_j plays of the strategy pair a_j. The fifth string consist of one play of b_k, and it follows the string $d_k * a_k$. Symbolically we could write the play of the last d stages as

$$(\omega_{T-d+1}, \ldots, \omega_T) = d_1 * a_1 + \cdots + d_k * a_k + b_k + \cdots + d_4 * a_4$$

and therefore $(\omega_1, \ldots, \omega_d) = (\omega_{T-i+1}, \ldots, \omega_T, \omega_{T-d+1}, \ldots, \omega_{T-i})$ where $i = T (\mathrm{mod}\ d)$. Therefore,

$$\left| \frac{\sum_{j=1}^{4} d_j r^i(a_j) + r^i(b_k)}{d} - x^i \right| < \varepsilon/2,$$

and therefore for sufficiently large values of T, e.g., $T \geq 2dK/\varepsilon$,

$$(4) \qquad \left| \frac{\sum_{t=1}^{T} r^i(\omega_t)}{T} - x^i \right| < \varepsilon.$$

Note that the average payoff to Player 1 in each cycle is at least $u^1(G) + \varepsilon$ and that for any $a \in A$ and any $1 \le t \le T$ we have that $r^1(a) - r^1(\omega_t) \le K$. Therefore, for every $t < T - dK/\varepsilon$, and every $a^1 \in A^1$,

$$(5) \qquad r^1(a^1, \omega_t^2) + (T - t)u^1(G) < \sum_{s=t}^{T} r^1(\omega_s).$$

Furthermore, the periodic play ω is such that for every $a^2 \in A^2$ and every $T - d_4 < t \le T$, we have that $r^2(\omega_t^1, a^2) + \sum_{s>t}^{T} u^2(G) \le \sum_{s=t}^{T} r^2(\omega_s)$. As the sequence $r^2(a_i)$ is nondecreasing, i.e., $r^2(a_1) \le r^2(a_2) \le r^2(a_3) \le r^2(a_4)$ we deduce that for every $t \le T - d_4$ we have that $r^2(\omega_t^1, a^2) + \sum_{s>t}^{T} u^2(G) \le r^2(\omega_t^1, a^2) + \sum_{s>t}^{T-d_4} u^2(G) + d_4 u^2(G) \le \sum_{s=t}^{T} r^2(\omega_s)$. Altogether we conclude that for every $a^2 \in A^2$ and every $t \le T$,

$$(6) \qquad r^2(\omega_t^1, a^2) + \sum_{s>t}^{T} u^2(G) \le \sum_{s=t}^{T} r^2(\omega_s).$$

We will show now that the pair of strategies in $G^T(m_1, m_2)$, σ of Player 1 and τ of Player 2, that follow the play ω as long as the other player follows the play ω and trigger to punishing forever with a pure strategy as soon as a deviation is observed, is an equilibrium of $G^T(m_1, m_2)$. Note that such a strategy is in $\Sigma^i(m_i)$ whenever $d < m_i$. Moreover, it follows from (5) that if $\sigma' \in \Sigma^1$ is such that $r_T^1(\sigma, \tau) < r_T^1(\sigma', \tau)$, then $\omega_t(\sigma', \tau) = \omega_t$ for every $t < T - dK/\varepsilon$. Thus, the strategy σ' must deviate from the periodic play ω, after following the d-cycle for at least $[(T - dK/\varepsilon - 1)/d]$ rounds. Therefore, the complexity of σ' is at least $d_k(T - 2dK/\varepsilon)/d > 2(T - 2dK/\varepsilon)/7$, which for a sufficiently large value of T_0 is larger than $T/4$, and thus such a strategy σ' is not in $\Sigma^1(m_1)$. Therefore, for every $\sigma' \in \Sigma^1(m_1)$,

$$r_T^1(\sigma, \tau) \ge r_T^1(\sigma', \tau).$$

It follows from (6) that for every τ' in Σ^2,

$$r_T^2(\sigma, \tau) \ge r_T^2(\sigma, \tau').$$

Altogether we conclude that for sufficiently large values of m_0 and T_0, (σ, τ) is an equilibrium of $G^T(m_1, m_2)$ with payoff vector y that satisfies $|y^i - x^i| < \varepsilon$.

(B) The case $m_1 \ge T/4$.

Assume that $m_1 \ge T/4$. Let $x \in \text{co}(r(A))$. Then x is a convex combination of at most three elements of $r(A)$. We consider the following three cases, according to the minimal number of elements of $r(A)$ that contain in their convex hull the point x. (1) There is $a \in A$ with $x = r(a)$. (2) There are two different elements a_1, and a_2 and positive numbers $\lambda_1 > 0$ and $\lambda_2 > 0$, such that $\lambda_1 + \lambda_2 = 1$ and $\sum_{i=1}^{2} \lambda_i r(a_i) = x$. (3) There are three different elements $a_1, a_2, a_3 \in A$ and positive numbers $\lambda_1 > 0$, $\lambda_2 > 0$ and $\lambda_3 > 0$, with $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and $\sum_{i=1}^{3} \lambda_i r(a_i) = x$. Case 3 is partitioned into three subcases according to the relative position of the entries a_1, a_2, and a_3 in the matrix.

(3.1) $a_1^1 = a_2^1 = a_3^1$,

(3.2) $a_1^1 = a_2^1 \ne a_3^1$,

(3.3) $|\{a_1^1, a_2^1, a_3^1\}| = 3$.

Proving the result for Case 3 is obviously sufficient to establish the theorem; any vector payoff $x \in \mathrm{co}(r(A))$ could be approximated as the convex combination $\Sigma_{i=1}^{3} \lambda_i r(a_i) = x$, with $\lambda_1 > 0$, $\lambda_2 > 0$, $\lambda_3 > 0$, and $\lambda_1 + \lambda_2 + \lambda_3 = 1$, and a_1, a_2, a_3 are three different elements in A. However, the result is easier in case 1, so to help the reader grasp the ideas of the general proof, we first prove the results in case 1.

(B.1) subcase 1 with $m_2 > T$.

The following is a construction of an equilibrium for Case 1, i.e., $x = r(a)$, and under the assumptions $m_1 \geq m_0$ and $m_2 > T$. We will construct an equilibrium point (σ^*, τ^*) of $G^T(m_1, m_2)$ with associated equilibrium vector payoff (y^1, y^2) with $|y^i - r^i(a)| < \epsilon$.

We use the following notations here. The strategy b^2 of Player 2 in the (one shot) game G is a best reply of Player 2 to the strategy a^1 of Player 1. The pure strategy a^i is denoted by 0. Let D^i be the punishing strategy of Player i, i.e., Player i's strategy that holds Player $3 - i$ down to $u^{3-i}(G)$. Note that $D^i \neq a^i$, and denote the pure strategy D^i by 1.

The mixed equilibrium strategy of Player 2, τ^*, chooses randomly a pure strategy τ^ϵ where ϵ is an element of a message space Q. The message space Q is a set of sequences of zeros and ones of length k, where k depends on the parameters of the game, T and m_1. The mixed strategy σ^* of Player 1 and the pure strategy τ^ϵ of Player 2 induce a play $\omega(\sigma^*, \tau^\epsilon) = (\omega_1(\sigma^*, \tau^\epsilon), \ldots, \omega_T(\sigma^*, \tau^\epsilon))$ that depends on ϵ, and therefore we may denote it as $\omega(\epsilon) = (\omega_1(\epsilon), \ldots, \omega_T(\epsilon))$ and call it the proposed play. Player 2 communicates his choice of $\epsilon = (\epsilon_1, \ldots, \epsilon_k)$ in Q to Player 1 during the first k stages of the repeated game by playing a^2 in all stages t with $\epsilon_t = 0$ and playing D^2 in all stages t with $\epsilon_t = 1$. The next k stages of the proposed play are used to balance the number of stages $1 \leq t \leq 2k$ with $\omega_t^2(\epsilon) = 0$ and the number of stages $1 \leq t \leq 2k$ with $\omega_t^2(\epsilon) = 1$. The proposed play satisfies also for every $1 \leq t \leq 2k$, $\omega_t^1(\epsilon) = 0$. The sequence of actions of Player 2 in stages $1 \leq t \leq 2k$ of the proposed play depends thus on ϵ and is denoted by $\theta(\epsilon) = (\theta_1(\epsilon), \ldots, \theta_{2k}(\epsilon))$. We refer to the first $2k$ stages of the repeated game as the communication phase. Following the communication phase $\omega_1(\epsilon), \ldots, \omega_{2k}(\epsilon)$ with $\omega_t(\epsilon) = (0, \theta_t(\epsilon))$, and excluding the last stage of the game, the proposed play enters a cycle of length l, where l depends on T, e.g., l will be the integer part of $2T/9$, and l is much larger than k. Both players cooperate during the first $l - 2k$ stages of the cycle, i.e., $\omega_t(\epsilon) = (0, 0)$ whenever $t < T$ and $t(\mathrm{mod}\ l) \geq 2k + 1$ or $0 = t(\mathrm{mod}\ l)$. Following these initial $l - 2k$ stages in the cycle, the players play the string $((\theta_1(\epsilon), \theta_1(\epsilon)) \ldots, (\theta_{2k}(\epsilon)\theta_{2k}(\epsilon)))$, i.e., for $2k < t$ with $1 \leq t(\mathrm{mod}\ l) \leq 2k$, $\omega_t(\epsilon) = (\theta_{t(\mathrm{mod}\ l)}(\epsilon), \theta_{t(\mathrm{mod}\ l)}(\epsilon))$.

The strategy of Player 1 will detect with positive probability any deviation of Player 2. Some deviations of Player 2 will be detected with positive probability immediately, and others will lead to a detection with positive probability in a future stage. The strategy of Player 1 triggers to punishing (playing D^1) forever once it detects a deviation by Player 2. We turn now to the formal definition of the proposed play and the construction of the equilibrium strategies.

We start with the construction of the set Q, and the integers k and l. Let L be a sufficiently large number, e.g., $L = 4$ will do. Let $l = [T/(L + 1/2)]$, and let $k = k(m_1, l)$, i.e., k is the smallest integer such that $2^k l > m_1 - l$. It follows that

$$2^{k-1} \leq [(m_1 - l)/l] < 2^k \tag{7}$$

and therefore there is a subset $Q = Q(m_1, l)$ of $\{0, 1\}^k$ such that $|Q| = [(m_1 - l)/l]$. It follows that $(m_1 - l)/l - 1 < |Q| \leq (m_1 - l)/l$ and thus

$$m_1 - l < |Q|l + l \leq m_1. \tag{8}$$

In addition we require the subset Q of $\{0, 1\}^k$ to satisfy the following two conditions:

$$Q \supset \{(\epsilon_1, \ldots, \epsilon_{k-1}, 0) : \epsilon_i \in \{0, 1\}\} \qquad (1, \ldots, 1) \notin Q.$$

These additional two requirements are feasible by (7) and they are used to simplify the explicit description of the proposed play as a function of $\epsilon \in Q$ as well as the description of the equilibrium strategy of Player 1.

Given $\epsilon = (\epsilon_1, \ldots, \epsilon_k)$ in Q we denote by $\theta(\epsilon) = (\theta_1(\epsilon), \ldots, \theta_{2k}(\epsilon))$ the element of $\{0, 1\}^{2k}$ given by: $\theta_i(\epsilon) = \epsilon_i$ if $1 \le i \le k$, $\theta_i(\epsilon) = 0$ if $k < i \le k + \Sigma\epsilon_i$ and $\theta_i(\epsilon) = 1$ if $k + \Sigma\epsilon_i < i \le 2k$. Note that $\theta(\epsilon)$ is a sequence of length $2k$ of zeros and ones having exactly k zeros and k ones. The first k coordinates coincide with $\epsilon_1, \ldots, \epsilon_k$ and the remaining coordinates start with a string of zeros followed by a string of ones so that the total number of ones equals k. Note that for every $(\epsilon_1, \ldots, \epsilon_k) \in Q$, $\Sigma_{i=1}^{k} \epsilon_i < \mathrm{k}$, and therefore $\theta_{2k}(\epsilon) = 1$ for every ϵ in Q. For every ϵ in Q we associate a play $\omega(\epsilon)$ of G^T, i.e., a sequence $\omega(\epsilon) = (\omega_1(\epsilon), \ldots, \omega_T(\epsilon))$ with $\omega_t(\epsilon) = (\omega_t^1(\epsilon), \omega_t^2(\epsilon))$ in A as follows: $\omega_T(\epsilon) = (0, b^2)$ and for $t < T$,

$$\omega_t(\epsilon) = \begin{cases} (0, \theta_t(\epsilon)) & \text{if } 1 \le t \le 2k, \\ (0, 0) & \text{if } 0 < (t - 2k)(mod\ l) \le l - 2k, \\ (\theta_i(\epsilon), \theta_i(\epsilon)) & \text{if } (t - 2k)(mod\ l) = l - 2k + i < l, \\ (\theta_{2k}(\epsilon), \theta_{2k}(\epsilon)) & \text{if } (t - 2k)(mod\ l) = 0. \end{cases}$$

Setting

$$\overline{\theta}(\epsilon) = ((0, \theta_1(\epsilon)), \ldots, (0, \theta_{2k}(\epsilon))),$$

$$\theta^*(\epsilon) = ((\theta_1(\epsilon), \theta_1(\epsilon)), \ldots, (\theta_{2k}(\epsilon), \theta_{2k}(\epsilon))),$$

and $d = T - 2k - Ll$,

$$\omega(\epsilon) = \overline{\theta}(\epsilon) + L * ((l - 2k) * (0, 0) + \theta^*(\epsilon)) + (d - 1) * (0, 0) + (0, b^2).$$

We derive in the following two lemmas two important properties of the proposed plays.

LEMMA 7. *The vector payoff* $\Sigma_{t=1}^{T} r(\omega_t(\epsilon))$ *is independent of* ϵ, *and for sufficiently large values of* T,

$$\|R(\omega(\epsilon)) - r(0, 0)\| < \varepsilon.$$

PROOF. For every $\epsilon \in Q$, $\Sigma_{i=1}^{2k} \theta_i(\epsilon) = k$. Therefore, $\Sigma_{i=1}^{2k} r(\overline{\theta}_i(\epsilon) = kr(0, 1) + kr(0, 0)$ and $\Sigma_{i=1}^{2k} r(\theta_i^*(\epsilon)) = kr(1, 1) + kr(0, 0)$ for every $\epsilon \in Q$. Therefore,

$$\sum_{t=1}^{T} r(\omega_t(\epsilon)) = kr(0, 1) + Lkr(1, 1) + r(0, b^2) + (T - Lk - k - 1)r(0, 0).$$

As the right hand side of the above equality is independent of ϵ, the first conclusion of the lemma follows. As for T sufficiently large, $(Lk + k + 1)/T$ is sufficiently small, the second conclusion follows. □

Notice that for any $\epsilon \in Q$ and any $2k < t \le 2k + l$, the play $(\omega_t(\epsilon), \ldots, \omega_{t+l-1}(\epsilon))$ is a coordinated play. The next lemma asserts that all these coordinated plays are $|Q|l$ different ones.

LEMMA 8. *For every*

$$(\epsilon, t), (\epsilon', t') \in Q \times \{2k+1, 2k+2, \ldots, 2k+l\}$$

with $(\epsilon, t) \neq (\epsilon', t')$,

$$(\omega_t(\epsilon), \ldots, \omega_{t+l-1}(\epsilon)) \neq (\omega_{t'}(\epsilon'), \ldots, \omega_{t'+l-1}(\epsilon')),$$

and thus there exists $0 \leq s < l$, *with*

$$\omega^1_{t+s}(\epsilon) \neq \omega^1_{t'+s}(\epsilon')$$

and such that for every $0 \leq s' < s$,

$$(\omega^2_t(\epsilon), \ldots, \omega^2_{t+s'}(\epsilon)) = (\omega^2_{t'}(\epsilon'), \ldots, \omega^2_{t'+s'}(\epsilon')).$$

PROOF. Note that in each one of the proposed plays $\omega(\epsilon)$, $\epsilon \in Q$, both players play in stages $t = 2k+1, \ldots, T-1$ in concert, i.e., $\omega^1_t(\epsilon) = \omega^2_t(\epsilon)$. Therefore, it is enough to prove that for any such pair (ϵ, t) and (ϵ', t'), there is $0 \leq s < l$ with

$$(\omega_t(\epsilon), \ldots, \omega_{t+s}(\epsilon)) \neq (\omega_{t'}(\epsilon'), \ldots, \omega_{t'+s}(\epsilon')).$$

or equivalently that there is $0 \leq s < l$ with

$$\omega_{t+s}(\epsilon) \neq \omega_{t'+s}(\epsilon').$$

Assume first that $t = t'$ and thus $\epsilon \neq \epsilon'$. As the map $\epsilon \to \theta^*(\epsilon)$ is $1-1$, $\theta^*(\epsilon) \neq \theta^*(\epsilon')$ and therefore there is $0 \leq s < l$ with

$$(\omega_t(\epsilon), \ldots, \omega_{t+s}(\epsilon)) \neq (\omega_{t'}(\epsilon'), \ldots, \omega_{t'+s}(\epsilon')).$$

Next assume that $t < t'$. If $t' - t > l - 2k$, setting $s = 2k + l - t'$, $t' + s = l + 2k$ and thus $\omega_{t'+s}(\epsilon') = (1, 1)$ while $2k < t + s < 4k$ and thus $\omega_{t+s}(\epsilon) = (0, 0)$; and if $t' - t \leq l - 2k$, setting $s = 2k + l - t$, $t + s = l + 2k$ and thus $\omega_{t+s}(\epsilon) = (1, 1)$ while $2k + l < t' + s \leq 2l$ and thus $\omega_{t'+s}(\epsilon) = (0, 0)$. □

We describe now the equilibrium strategy of Player 2. The strategy of Player 2 calls for playing according to the proposed play as long as Player 1 follows the proposed play, and it triggers to punishing (playing D^2) forever as soon as it observes a deviation by Player 1 from the proposed play $\omega(\epsilon)$. Thus, for any ϵ in Q, $\tau^\epsilon = (\tau^\epsilon_t)^T_{t=1}$ is the pure strategy of Player 2 defined by,

$$\tau^\epsilon_t(s_1, \ldots, s_{t-1}) = \begin{cases} \omega^2_t(\epsilon) & \text{if } (s_1, \ldots, s_{t-1}) = (\omega_1(\epsilon), \ldots, \omega_{t-1}(\epsilon)), \\ 1 & \text{otherwise.} \end{cases}$$

Note that the strategy τ^* communicates its choice of ϵ in Q to Player 1 during the first $2k$ stages of the repeated game by defecting (playing D^2) in all stages t for which $\theta_t(\epsilon) = 1$.

Observe that $\tau^\epsilon \in \Sigma^2(T, T+1)$, i.e., τ^ϵ is implemented by an automaton of size $T + 1$; indeed, let $\langle \{1, \ldots, T, T+1\}, 1, f^2_\epsilon, g^2_\epsilon \rangle$ be the automation with action function f^2_ϵ defined by $f^2_\epsilon(t) = \omega^2_t(\epsilon)$ if $t \leq T$, $f^2_\epsilon(T+1) = D^2$, and $g^2_\epsilon(t, a) = t + 1$ if $a = \omega^1_t(\epsilon)$ and $t \leq T$, and $g^2_\epsilon(t, a) = T + 1$ otherwise, i.e., if $a \neq \omega^1_t(\epsilon)$, or if $t = T + 1$.

The equilibrium strategy τ^* in $\Delta(\Sigma^2(T, m_2))$ chooses an element ϵ from Q, each element equally likely, i.e., with probability $1/|Q|$, and given its choice ϵ, plays τ^ϵ. In other words, τ^* is the mixed strategy that selects the pure strategy τ^ϵ with probability $1/|Q|$, or in other words, τ^* is the probability distribution on $\Sigma^2(T, m_2)$ which assigns probability $1/|Q|$ to each pure strategy τ^ϵ in $\Sigma(T, m_2)$.

Let σ be a strategy of Player 1, and $\epsilon \in Q$, with $r_T^1(\sigma, \tau^\epsilon) \geq \Sigma_{t=1}^T r^1(\omega_t(\epsilon))/T$. Then $\omega_t(\sigma, \tau^\epsilon) = \omega_t(\epsilon)$ for any $t \leq T - n$ where n is a fixed number that depends on a and G alone. Therefore, for any strategy σ of Player 1,

$$r_T^1(\sigma, \tau^\epsilon) \leq \sum_{t=1}^T r^1(\omega_t(\epsilon))/T + C/T, \tag{9}$$

where C is a constant that depends on G alone.

Let σ be a pure strategy for Player 1 with

$$r_T^1(\sigma, \tau^*) \geq \sum_{t=1}^T r^1(\omega_t(\epsilon))/T,$$

and such that σ is implemented by an automation of size m_1.

Set

$$Q(1, \sigma) = \left\{ \epsilon \in Q : r_T^1(\sigma, \tau^\epsilon) > \sum_{t=1}^T \frac{r^1(\omega_t(\epsilon))}{T} \right\}, \tag{10}$$

$$Q(2, \sigma) = \left\{ \epsilon \in Q \setminus Q(1, \sigma) : r_T^1(\sigma, \tau^\epsilon) \geq \sum_{t=1}^T \frac{r^1(\omega_t(\epsilon))}{T} - \frac{x^1 - u^1(G)}{3} \right\}, \tag{11}$$

and

$$Q(3, \sigma) = \left\{ \epsilon \in Q : r_T^1(\sigma, \tau^\epsilon) < \sum_{t=1}^T \frac{r^1(\omega_t(\epsilon))}{T} - \frac{x^1 - u^1(G)}{3} \right\}. \tag{12}$$

LEMMA 9. *For every pair (ϵ, t) and (ϵ', t'), with $(\epsilon, t) \neq (\epsilon', t')$ and $t \geq t'$, in the union of the two sets*

$$Q(1, \sigma) \times \{2k + 1, 2k + 2, \ldots, 3l + 2k\}$$

and

$$Q(2, \sigma) \times \{2k + 1, \ldots, 2k + l\}$$

there exists $s < T - t$ such that

$$(\omega_t^2(\epsilon), \ldots, \omega_{t+s}^2(\epsilon)) = (\omega_{t'}^2(\epsilon'), \ldots, \omega_{t'+s}^2(\epsilon'))$$

and

$$\sigma(\omega_1(\epsilon), \ldots, \omega_{t+s}(\epsilon)) \neq \sigma\,(\omega_1(\epsilon'), \ldots, \omega_{t'+s}(\epsilon')).$$

PROOF. For every $\epsilon \in Q(1, \sigma)$, $r^1(\sigma, \tau^\epsilon) > R^1(\omega(\epsilon))$ and therefore there is a deviation from the proposed play but no deviation prior to stage $2k + 4l$, i.e., for every $\epsilon \in Q(1, \sigma)$ and every $t \leq 2k + 4l$, $\omega_t(\sigma, \tau^\epsilon) = \omega_t(\epsilon)$. For every $\epsilon \in Q(2, \sigma)$, $r^1(\sigma, \tau^\epsilon)$

$> R^1(\omega(\epsilon)) - (x^1 - u^1(G))/3$ and therefore no deviation from the proposed play prior to stage $2k + 2l$, i.e., for every $\epsilon \in Q(2, \sigma)$ and every $t \leq 2k + 2l$, $\omega_t(\sigma, \tau^\epsilon) = \omega_t(\epsilon)$.

Therefore, if either $t \neq t'(\text{mod } l)$ or $t = t'(\text{mod } l)$ and $\epsilon \neq \epsilon'$ apply Lemma 7. If $t > t'$, $t = t'(\text{mod } l)$ and $\epsilon \in Q(1, \sigma)$, let s be the largest positive integer such that

$$(\omega_t(\epsilon), \ldots, \omega_{t+s}(\epsilon)) = (\omega_{t'}(\sigma, \tau^\epsilon), \ldots, \omega_{t'+s}(\sigma, \tau^\epsilon)).$$

As $r^1(\sigma, \tau^\epsilon) \neq R^1(\omega(\epsilon))$, $s < T - t$ and

$$\sigma(\omega_1(\epsilon), \ldots, \omega_{t+s}(\epsilon)) \neq \sigma(\omega_1(\epsilon'), \ldots, \omega_{t'+s}(\epsilon')). \quad \square$$

Lemma 8 implies that

$$\text{comp}(\sigma) \geq 3l|Q(1, \sigma)| + l|Q(2, \sigma)|. \tag{13}$$

LEMMA 10. *For any strategy* $\sigma \in \Sigma^1(m_1)$,

$$r_T^1(\sigma, \tau^*) \leq \frac{1}{|Q|} \sum_Q R^1(\omega(\epsilon)).$$

PROOF. If $r_T^1(\sigma, \tau^*) \geq \Sigma_{t=1}^T r^1(\omega_t(\epsilon))/T$, it follows from (9) and the definition of $Q(3, \sigma)$ that

$$\frac{C}{T} |Q(1, \sigma)| \geq \frac{r^1(a) - u^1(G)}{3} |Q(3, \sigma)|,$$

i.e.,

$$|Q(1, \sigma)| \geq \frac{T(r^1(a) - u^1(G))}{3C} |Q(3, \sigma)|.$$

Either $Q(3, \sigma) = \emptyset$ and then $|Q(1, \sigma)| + |Q(2, \sigma)| = |Q|$, thus using (13) and (8),

$$\begin{aligned} m_1 &\geq |Q(1, \sigma)|3l + |Q(2, \sigma)|l \\ &= |Q|l + |Q(1, \sigma)|2l \\ &> m_1 - 2l + |Q(1, \sigma)|2l \end{aligned}$$

which is possible only when $Q(1, \sigma) = \emptyset$ and then $r_T^1(\sigma, \tau^*) \leq \Sigma_{t=1}^T r^1(\omega_t(\epsilon))/T$. Or, $Q(3, \sigma) \neq \emptyset$, and $T > 6C/(r^1(a) - u^1(G))$. Then,

$$|Q(1, \sigma)| > 2|Q(3, \sigma)| \geq 2.$$

In that case,

$$\begin{aligned} m_1 &\geq |Q(1,\sigma)|3l + |Q(2,l)|l \\ &= |Q(1,\sigma)|\frac{l}{2} + |Q(1,\sigma)|\frac{5l}{2} + |Q(2,l)|l \\ &> |Q|l + |Q(1,\sigma)|\frac{3l}{2} \\ &> m_1 - 2l + |Q(1,\sigma)|l \\ &> m_1, \end{aligned}$$

which is impossible. □

We construct now two pure strategies of Player 1, σ^p, $p = [l/3]$ and $p = 2[l/3]$ in $\Sigma^1(m_1)$ that satisfy the following two properties:

$$\omega(\sigma^p, \tau^\epsilon) = \omega(\epsilon)$$

and if σ^* is the mixture of the two strategies σ^p, each with probability $1/2$, then for every pure strategy τ in Σ^2 of Player 2,

$$r_T^2(\sigma^*, \tau) \leq \sum_{t=1}^{T} r^2(\omega_t(\epsilon))/T.$$

It will follow that (σ^*, τ^*) is indeed an equilibrium of $G^T(m_1, m_2)$.

Each of the pure strategies, σ^p, is implemented by an automaton with state space $Q \times \{1, \ldots, l\} \cup \{\emptyset\}$. The state $\emptyset$ of the automaton is interpreted as the punishing state. Once the automaton moves into that state, it stays there and ‘‘punishes’’ forever, i.e., plays 1 repeatedly. The initial state of the automaton is $(0, p)$ where 0 is the sequence of zeros in Q. The action function of the automaton implementing σ^p is independent of the value of p and is given by:

$$f^1(\emptyset) = 1$$

and for ϵ in Q and $1 \leq j \leq l$,

$$f^1(\epsilon, j) = \begin{cases} 0 & \text{if } 1 \leq j \leq l - 2k, \\ \theta_i(\epsilon) & \text{if } j = l - 2k + i. \end{cases}$$

We may visualize the states of the automaton of the form (ϵ, j) as if they are arranged in a rectangular array with $|Q|$ rows and l columns. The rows are indexed by the different elements ϵ in Q and the columns are indexed $1, \ldots, l$. Thus the action function assigns to each state in the first $(l - 2k)$ columns the action a^1, and in all other columns an action that depends on the row.

Figure 2 is a partial illustration of the automaton of Player 1, in the case that $k = 3$, $|Q| = 7$ and $p = 2[l/3]$. The only states depicted in Figure 2 are the elements of $Q \times \{p, \ldots, p + 6, l - 6, \ldots, l - 3\}$. The disks (circles with the center filled in) represent states of the automaton that take the action 1. The other circles represent states of the automaton that take the action 0 (a^1). For every $1 \leq t \leq 3$, the action of Player 1 in state $((\epsilon_1, \epsilon_2, \epsilon_3), l - 6 + t)$ is ϵ_t. Therefore, the sequence of the last three circles and disks in each row identifies the element $\epsilon \in Q$. The initial state, $((0, 0, 0), p)$ is marked with an $*$.

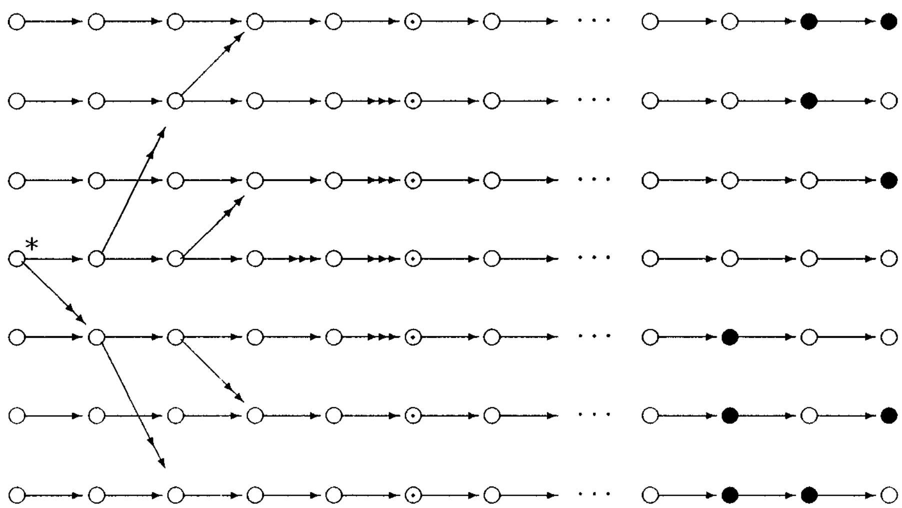

FIGURE 2.

The horizontal arrows indicate the transitions of the automaton given a coordinated play, as expected in the cycles. The double arrows indicate the transitions of the automaton given action 1 by Player 2. The triple arrows indicate the transitions of the automaton when the action of Player 2 is either 0 or 1. The circles with a dot at their center, indicate the states $(\epsilon, p+5)$. If the automaton of Player 1 is state $(\epsilon, p+5)$ and Player 2 takes the action 1, the next state of the automaton of Player 1 is $(\epsilon, 1)$.

The transition function of the automaton of Player 1 that implements σ^p, is described by the function $g^1 = g^1_p$ *which depends on the integer* p. There is one state of the automaton, the punishing state $\emptyset$, which is an absorbing state, i.e.,

$$g^1_p(\emptyset, *) = \emptyset.$$

During the first k stages of the game, the automaton is moving between the different rows in such a way that in stage $k+1$ the state of the automaton is the row that corresponds to the sequence of actions of Player 2 in the first k stages of the game. This is achieved by the following partial definition of the transition function. If $\epsilon = (\epsilon_1, \ldots, \epsilon_k) = (\epsilon_1, \ldots, \epsilon_{j-p}, 0, \ldots, 0)$ and $p \le j < p+k$, then,

$$g^1((\epsilon, j), 1) = ((\epsilon_1, \ldots, \epsilon_{j-p}, 1, 0, \ldots, 0), j+1).$$

One property of the transition function of the automaton implementing σ^p, that will follow from our definition of g^1_p, is:

$$g^1_p((0, p), (\theta_1(\epsilon), \cdots, \theta_{2k}(\epsilon))) = (\epsilon, 1).$$

The next step in our construction of the transition function g^1_p, ensures that the strategy σ^p follows the proposed play $\omega(\epsilon)$ as long as the other player follows the proposed play.

$$g^1((\epsilon, j), 0) = (\epsilon, j+1) \quad \text{if } j \le l - 2k,$$
$$g^1((\epsilon, j), c) = (\epsilon, j+1) \quad \text{if } j = l - 2k + i \text{ and } \theta_i(\epsilon) = c.$$

In most other cases, the automaton will trigger to punishing forever, i.e., will move to the state $\emptyset$. However, there are few states which tolerate a play by Player 2 that differ from 0. As follows from our construction, the states of the form (ϵ, j), with $p \leq j < p + k$, are used by the automaton of Player 1 to differentiate among the possible messages $\epsilon \in Q$ that Player 2 may transmit in the first k stages of the game, and therefore they do tolerate the actions 0 and 1 of Player 2. The states of the automaton that are of the form (ϵ, j), with $p + k \leq j < p + k + \Sigma_{i=1}^{k} \epsilon_i$ do not tolerate any deviation, i.e.,

$$g^1((\epsilon, j), 1) = \emptyset \quad \text{if } p + k \leq j < p + k + \sum_{i=1}^{k} \epsilon_i .$$

Several states of the form (ϵ, j) with $j \geq p + k + \Sigma_{i=1}^{k} \epsilon_i$ tolerate also the action 1 of Player 2. These states and transitions do depend on the value of p and are described below. The uncertainty of Player 2 about the value of p disables him from exploitation of this toleration in future stages of game without risking detection with probability at least $1/2$. If $p = [l/3]$, then

$$g^1((\epsilon, j), 1) = (\epsilon, j + 2) \quad \text{if } p + k + \sum_{i=1}^{k} \epsilon_i \leq j < p + 3k - 2 - \sum_{i=1}^{k} \epsilon_i \text{ and}$$

$$j - p - k - \sum_{i=1}^{k} \epsilon_i \text{ is even,}$$

$$g^1((\epsilon, j), 1) = (\epsilon, 1) \quad \text{if } j = p + 3k - 2 - \sum_{i=1}^{k} \epsilon_i .$$

If $p = 2[l/3]$, then

$$g^1((\epsilon, j), 1) = (\epsilon, j + 1) \quad \text{if } p + k + \sum_{i=1}^{k} \epsilon_i \leq j < p + 2k - 1,$$

$$g^1((\epsilon, j), 1) = (\epsilon, 1) \quad \text{if } j = p + 2k - 1.$$

In all other cases the value of g_p^1 equals $\emptyset$. In order to prove that (σ^*, τ^*) is an equilibrium of $G^T(m_1, m_2)$ it suffices to show that τ^ϵ is indeed a best reply of Player 2 to σ^*.

LEMMA 11. *For every strategy $\tau \in \Sigma^2$ and every $\epsilon \in Q$,*

$$\sum_{t=1}^{T} r^2(\sigma^*, \tau) \leq \sum_{t=1}^{T} r^2(\sigma^*, \tau^\epsilon).$$

PROOF. Assume first that τ is a pure strategy of Player 2 such that for some $\epsilon \in Q$, $\omega_t(\sigma^*, \tau) = \omega_t(\epsilon)$ for every $1 \leq t \leq 2k$, and $r^2(\sigma^*, \tau) \geq r^2(\sigma^*, \tau^*) = r^2(\sigma^*, \tau^\epsilon)$. Either $\omega_t(\sigma^*, \tau) = \omega_t(\epsilon)$ for every $t \leq T$, and then $r^2(\sigma^*, \tau) = r^2(\sigma^*, \tau^*)$, or there is $2k < s \leq T$ with $\omega_s(\sigma^*, \tau) \neq \omega_s(\epsilon)$ and without loss of generality assume that for every $1 \leq t < s$, $\omega_t(\sigma^*, \tau) = \omega_t(\epsilon)$. Observe that $T - 2k - Ll$ is of the order of $l/2$ and therefore for sufficiently large values of T, $[l/3] + 3k < T - 2k - Ll < 2[l/3]$, and thus in stages $T > t > 2k + Ll + [l/3] + 3k$ the strategy σ^* does not tolerate any deviation from the proposed play. Also in these stages $\omega_t(\epsilon) = (0, 0)$ and $r^2(0, b^2) \geq r^2(0, 0) > u^2(G)$. Therefore, if $T > s > 2k + Ll + [l/3] + 3k$, $r^2(\sigma^*, \tau) < r^2(\sigma^*, \tau^\epsilon)$, and if $s = T$, $r^2(\sigma^*, \tau) \leq r^2(\sigma^*, \tau^\epsilon)$. If $2k < s \leq 5k + Ll + [l/3]$, then at least one of the two strategies σ^p, $p = [l/3]$ or $p = 2\,[l/3]$ detects immediately the deviation of Player

2 from the proposed play $\omega(\epsilon)$, and thus with probability at least 1/2 Player 1 detects immediately the deviation of Player 2. Observe that there is a constant C such that any play ω of G^T which is compatible with σ^p satisfies $R^2(\omega) \le R^2(\omega(\epsilon)) + Ck/T$, and on the other hand, if in addition a deviation by Player 2 is detected prior to stage $2k + Ll + [l/3] + 3k$, $r^2(\omega) < r_T^2(\omega(\epsilon)) - 1/C$. Thus for any strategy τ of Player 2 which is compatible with $\bar{\theta}(\epsilon)$ and deviates from the proposed play prior to stage $2k + Ll + [l/3] + 3k$, $r^2(\sigma^*, \tau) < r^2(\sigma^*, \tau^*)$. If $(\omega_1^2(\sigma^*, \tau), \ldots, \omega_{2k}^2(\sigma^*, \tau))$ is not in $\theta(Q)$, then with probability at least 1/2 Player 1 detects the deviation of Player 2 prior to stage l, and therefore his loss is approximately at least $(T - l)(x^2 - u^2(G))/T$, which is of the order of a positive constant. His possible gain is at most of an order of k/T. Therefore τ^* is indeed a best reply against σ^*. □

This completes the proof of case 1 under the assumption that $m_2 > T$.

(B.2) subcase 1 with $m_2 \le T$.

The handling of Case 1 under the assumption that $T/4 \le m_2 \le T$ could be done in various ways. For example, replace in the proposed play $(0, b)$ with $(0, 0)$, and Player 1 sends a message at the first stages of the game, before Player 2 sends his message, and the proposed play depends also on the message sent by Player 1, so that Player 2 is unable to count to stage T and deviate then, or by relying on our handling of the other cases in which Player 2 does not wish to deviate in the last few stages.

(B.3) subcase 3.

We turn now to the proof in Case 3. Recall that a proof of Case 3 provides also a proof for the other cases. There are several features present in the proof of Case 3 that were not present in the above proof in Case 1.

We start handling subcase (3.2). Assume that $a_1^1 = a_2^1 \neq a_3^1$ and assume further that $a_1^2 = a_3^2$. Again, Player 2 communicates its choice of $\epsilon \in Q \subset \{0, 1\}^k$ during the communication phase which lasts for $2k$ stages. The resulting play in the communication phase is denoted $\bar{\theta}(\epsilon)$. Following the communication phase the play enters a cycle of length l, part of which is a verification phase in which Player 1 communicates back the chosen ϵ.

Assume that $x \in \text{co}(r(A))$ with $x^i > u^i(G)$. As in the proof of the case $m_1 < T/4$ (where we replaced ε by $\varepsilon/4$ and x by a vector payoff $y > (u^1(G) + 2\varepsilon/3, u^2(G) + 2\varepsilon/3)$ which is $2\varepsilon/3$ apart from x), we may assume without loss of generality that $x^i > u^i(G) + 2\varepsilon$, and that $\varepsilon > 0$ is sufficiently small and the inequality $m_1 \le \exp(\varepsilon^3 T)$ is replaced by $m_1 \le \exp(64\varepsilon^3 T)$. Without loss of generality we denote a_1^1 and a_1^2 by 0 and a_3^1, a_2^2 and a_3^2 by 1, and thus we assume that

$$x = \lambda_0 r(0, 0) + \lambda_1 r(1, 1) + \lambda_2 r(0, 1),$$

with $\lambda_0 > 0$, $\lambda_1 > 0$, $\lambda_2 > 0$, and $\Sigma_{i=0}^2 \lambda_i = 1$. Let $L = [3K/\varepsilon]$. Either $\lambda_0 r^2(0, 0) + \lambda_2 r^2(0, 1) > (u^2(G) + 2\varepsilon)(\lambda_0 + \lambda_2)$, or $r^2(1, 1) > u^2(G) + 2\varepsilon$. Assume first that $r^2(1, 1) > u^2(G) + 2\varepsilon$. Set

$$\begin{aligned} l &= [T/(L + 1 - \delta)], \\ d_1 &= [\lambda_1 l], \\ d &= L^4, \\ d_2 &= [l\lambda_2/d], \\ d_3 &= [l\lambda_0(1 - 1/L)/d], \quad \text{and} \\ d_0 &= l - d_1 - dd_2 - d(d + 1)/2 - dd_3. \end{aligned}$$

The number l is the number of stages in each cycle following the communication phase.

The action pairs in each cycle are $(0, 0)$, $(0, 1)$ and $(1, 1)$. The number of plays of the action pairs $(0, 0)$, $(0, 1)$ and $(1, 1)$ in each cycle will turn out to be approximately $l\lambda_0$, $l\lambda_2$ and $l\lambda_1$, and thus the average payoff in the cycle is approximately x. The number of complete cycles in the proposed play will turn out to be equal to L. The $(L+1)$-th cycle will almost reach its end. The variable δ, which appears in the definition of the length l of the cycle, is sufficiently small, so that

$$(L+1)l - T \ll l.$$

On the other end, it is not too small. It is sufficiently large so that $(L+1)l \geq T$. E.g., $\delta = \varepsilon^2$. The last two inequalities will enable us to start playing the proposed cycle immediately after the communication phase and define the cycles $c(\epsilon)$ in such a way that: (1) the last part of the cycle is independent of ϵ, (2) the last string of action pairs in the proposed play is of the form $\beta*(a^1, a^2) + (a^1, b^2)$ where β is a sufficiently large positive integer, $r^2(a^1, a^2) > u^2(G) + \varepsilon$, b^2 is a best reply of Player 2 to the action a^1 of Player 1.

Let $l_1 = d_0 + d_1$. Player 1's complexity of the repeated play of each one of the proposed cycles will turn out to equal l_1. Note that d_0 is approximately $l\lambda_0/L$. Let $k = k(m_1, l_1)$, i.e., k is the smallest integer such that $2^k l_1 > m_1 - l_1$. As $m_1 \leq \exp(64\varepsilon^3 T)$,

$$k < 90\varepsilon^3 T. \tag{14}$$

Let $Q = Q(m_1, l_1)$, as in case 1. Recall that

$$m_1 - l_1 < |Q| l_1 + l_1 \leq m_1. \tag{15}$$

Recall that for every $\epsilon \in Q$ we associate two plays; the communication play $\overline{\theta}(\epsilon)$, and the verification play $\theta^*(\epsilon)$.

$$\overline{\theta}(\epsilon) = (0, \epsilon_1) + \cdots + (0, \epsilon_k) + \sum_{i=1}^{k} \epsilon_i * (0, 0) + \left(k - \sum_{i=1}^{k} \epsilon_i\right) * (0, 1),$$

and

$$\theta^*(\epsilon) = (\epsilon_1, \epsilon_1) + \cdots + (\epsilon_k, \epsilon_k) + \sum_{i=1}^{k} \epsilon_i * (0, 0) + \left(k - \sum_{i=1}^{k} \epsilon_i\right) * (1, 1).$$

Recall that for every $\epsilon \in Q$, $\Sigma_{i=1}^{k} \epsilon_i < \mathrm{k}$, and thus the communication play ends with $(0, 1)$ and the verification play ends with $(1, 1)$. Define the play c^* by,

$$c^* = \sum_{i=1}^{d} (d_3 * (0, 0) + d_2 * (0, 1) + (i-1) * (0, 0) + (0, 1)).$$

Note that for every $1 \leq i \leq d$,

$$|d_3 * (0, 0) + d_2 * (0, 1) + (i-1) * (0, 0) + (0, 1)| = d_3 + d_2 + i$$

and

$$\left\|R(d_3 * (0,0) + d_2 * (0,1) + (i-1) * (0,0) + (0,1)) - \frac{\lambda_0 r(0,0) + \lambda_2 r(0,1)}{\lambda_0 + \lambda_2}\right\|$$
$$< O(1/L).$$

Define the play $c = c(\epsilon)$ of length l by,

$$c = \theta^*(\epsilon) + c^* + (d_0 - 2k) * (0,0) + d_1 * (1,1).$$

The following lemma asserts that the average payoff per stage in the play c is approximately x.

LEMMA 12. *The vector payoff $R(c(\epsilon))$ is independent of ϵ, and for sufficiently large values of T,*

$$|R^i(c) - x^i| < \varepsilon/2.$$

PROOF. The number of times that each one of the action pairs, $(0,0)$, $(0,1)$ and $(1,1)$, appears in the play c, equals $k + dd_3 + d(d-1)/2 + d_0 - 2k = l - dd_2 - d - d_1 - k$, $dd_2 + d$, and $d_1 + k$, respectively. The inequality (14) implies that for sufficiently large values of T,

$$|l\lambda_1 - (d_1 + k)| < 91\varepsilon^3 T,$$

and by the definition of d_2,

$$|l\lambda_2 - (dd_2 + d)| \le d,$$

and thus using the above two inequalities,

$$|l\lambda_0 - (k + dd_3 + d(d-1)/2 + d_0 - 2k)| < 91\varepsilon^3 T + d.$$

Therefore, for sufficiently large values of T and ε sufficiently small,

$$|R^i(c) - x^i| < \varepsilon/2. \quad \square$$

Define the proposed play $\omega(\epsilon)$ by, $\omega_T(\epsilon) = (1, b)$ where b is a best reply of Player 2 to the action 1 of Player 1, and

$$\omega(\epsilon) = {}_T\overline{\theta}(\epsilon) + (L+1) * c.$$

Recall that the last $d_1 = [\lambda_1 l]$ action pairs of the play c are $(1,1)$, and that as $l = [T/(L + 1 - \delta)]$,

$$\delta l + 2k - L - 1 < (L+1)l + 2k - T \le \delta l + 2k \ll d_1,$$

and thus the proposed play ends with a long string of $(1,1)$ followed by the action pair $(1, b^2)$.

The following lemma asserts that the average payoff per stage in the proposed play $\omega(\epsilon)$ is independent of ϵ and is approximately x.

LEMMA 13. *The payoff per stage in the proposed play $\omega(\epsilon)$, $R(\omega(\epsilon))$, is independent of ϵ, and for sufficiently large values of T,*

$$|R^i(\omega(\epsilon)) - x^i| < \varepsilon.$$

PROOF. Note that $R^i(\omega_1(\epsilon), \ldots, \omega_{2k}(\epsilon))$ is independent of ϵ, and so is $R^i(c(\epsilon))$. Therefore, $R^i(\omega_1(\epsilon), \ldots, \omega_{2k+Ll}(\epsilon))$ is independent of ϵ. Observe also that

$$(\omega_{2k+Ll+1}, \ldots, \omega_{T-1}) + (2k + (L+1)l + 1 - T) * (1, 1) = c(\epsilon)$$

and $\omega_T(\epsilon) = (1, b^2)$. Therefore, as the vector payoff $R(c(\epsilon))$ is independent of ϵ, so is $R(\omega_{2k+Ll+1}, \ldots, \omega_T)$. Thus we deduce on the one hand that $R(\omega(\epsilon))$ is independent of ϵ, and on the other hand, using the inequality $2k \le 2k + (L+1)l - T \le l$, $|R^i(\omega(\epsilon) - R^i(c)| < K/L \le \varepsilon/3$. As $|R^i(c) - x^i| < \varepsilon/2$, the result follows. □

As usual, for every $\epsilon \in Q$, τ^ϵ is the pure strategy of Player 2 that follows the proposed play $\omega(\epsilon)$ as long as Player 1 follows it, and it triggers to punishing forever as soon as a deviation from the proposed play is observed. The mixed equilibrium strategy of Player 2, $\tau^* \in \Delta(\Sigma^2(T, m_2))$, chooses an element $\epsilon \in Q$, each element with probability $1/|Q|$, and given its choice ϵ, plays the pure strategy τ^ϵ. The mixed equilibrium strategy of Player 1, $\sigma^* \in \Delta(\Sigma^1(T, m_1))$, is a mixture of pure strategies, each being implemented by an automation with state space

$$M^1 = \{\emptyset\} \cup Q \times \{1, \ldots, l_1\}.$$

The action function of the automaton is given by,

$$f^1(\emptyset) = D^1,$$

and

$$f^1(\epsilon, j) = \begin{cases} \theta_j(\epsilon) & \text{if } 1 \le j \le 2k, \\ 0 & \text{if } 2k < j \le d_0, \\ 1 & \text{if } d_0 < j \le d_1. \end{cases}$$

The transitions of the automaton will be defined so that for each fixed $\epsilon \in Q$, if Player 2's strategy is τ^ϵ, the state of the automaton at stage $t = l + 2k + j$ with $1 \le j \le 2k$, or at stage $t = l + 2k - l_1 + j$ with $2k + 1 \le j \le l_1$, is (ϵ, j). This leads to the following defined transitions:

$$g^1((\epsilon, j), 0) = \begin{cases} (\epsilon, j+1) & \text{if } 1 \le j < 2k \text{ and } \theta_j(\epsilon) = 0, \\ (\epsilon, j+1) & \text{if } 2k < j \le d_0, \end{cases}$$

and

$$g^1((\epsilon, j), 1) = \begin{cases} (\epsilon, j+1) & \text{if } 1 \le j < 2k \text{ and } \theta_j(\epsilon) = 1, \\ (\epsilon, j+1) & \text{if } d_0 < j < l_1, \\ (\epsilon, 1) & \text{if } j = l_1. \end{cases}$$

The states of the automaton of the form (ϵ, j) with $1 \le j \le 2k$ or $j > d_0$ expect a coordinated play. Any deviation from a coordinated play (either $(0, 0)$ or $(1, 1)$) at these states results in punishing forever. This is accomplished by defining the following transitions:

$$g^1((\epsilon, j), 0) = \begin{cases} \emptyset & \text{if } 1 \le j \le 2k \text{ and } \theta_j(\epsilon) = 1, \\ \emptyset & \text{if } d_0 < j \le l_1, \end{cases}$$

$$g^1((\epsilon, j), 1) = \emptyset \quad \text{if } 1 \le j < 2k \text{ and } \theta_j(\epsilon) = 0,$$

and

$$g^1(\emptyset, *) = \emptyset.$$

So far we have defined parts of the transition function of the automaton of Player 1 which are independent of the pure strategies in the support of σ^*. Other parts of the transition function, and the initial state, are random and do depend on the following independent random elements. A random integer p, $d_0 - 2\bar{L}k \le p < d_0 - 3k$, where $\bar{L}$ is a sufficiently large positive integer, random integers q, $1 \le q \le 2$, and z, $1 \le z \le L$, a random increasing function j, $j : \{1, \ldots, L\} \to \{2k + d_3 + 1, \ldots, d_0 - 3Lk\}$ with $j(i + 1) > j(i) + 2d_2 + d_3 + d$, and a random sequence $i_1, \ldots, i_d$ of elements of $\{1, \ldots, L\}$. The initial state of the automaton is $(0, p)$ where 0 is the sequence of zeros in Q. We define now those transitions that enable Player 1 to record the chosen $\epsilon \in Q$. If $\epsilon = (\epsilon_1, \ldots, \epsilon_k) = (\epsilon_1, \ldots, \epsilon_{j-p}, 0, \ldots, 0)$ and $p \le j < p + k$, and $(\epsilon_1, \ldots, \epsilon_{j-p}, 1, 0, \ldots, 0) \in Q$, then,

$$g^1((\epsilon, j), 1) = ((\epsilon_1, \ldots, \epsilon_{j-p}, 1, 0, \ldots, 0), j + 1).$$

The states of the automaton that are of the form (ϵ, j), with $p + k \le j < p + k + \sum_{i=1}^k \epsilon_i$ do not tolerate any deviation, i.e.,

$$g^1((\epsilon, j), 1) = \emptyset \quad \text{if } p + k \le j < p + k + \sum_{i=1}^{k} \epsilon_i .$$

Several states of the form (ϵ, j) with $p + k + \sum_{i=1}^k \epsilon_i \le \mathrm{j} \le \mathrm{p} + 3k$ tolerate also the action 1 of Player 2. These states do depend on the values of p and q. The uncertainty of Player 2 about the values of p and q disables him from exploiting this toleration in future stages of the game without risking detection with high probability. If $q = 1$,

$$g^1((\epsilon, j), 1) = \begin{cases} (\epsilon, j + 1) & \text{if } p + k + \sum_{i=1}^{k} \epsilon_i \le j < p + 2k, \\ (\epsilon, 1) & \text{if } j = p + 2k - 1, \end{cases}$$

and if $q = 2$,

$$g^1((\epsilon, j), 1) = \begin{cases} (\epsilon, j + 2) & \text{if } p + k + \sum_{i=1}^{k} \epsilon_i \le j < p + 3k - \sum_{i=1}^{k} \epsilon_i , \\ & \text{and } j - p - k - \sum_{i=1}^{k} \epsilon_i \text{ is even}, \\ (\epsilon, 1) & \text{if } j = p + 3k - \sum_{i=1}^{k} \epsilon_i - 2, \end{cases}$$

and

$$g^1((\epsilon, 2k), 1) = (\epsilon, j(i_1) - d_3).$$

For every $1 \leq t \leq d$, we define the following transitions.

$$g^1((\epsilon, j(i_t) + zs), 1) = (\epsilon, j(i_t) + zs + z) \quad \text{if } 0 \leq s < d_2,$$

and

$$g^1((\epsilon, j(i_t) + s), 1) = \begin{cases} (\epsilon, j(i_{t+1}) - d_3) & \text{if } s = j(i_t) + zd_2 + t \text{ and } t < d, \\ (\epsilon, 2k + 1) & \text{if } s = j(i_t) + zd_2 + t \text{ and } t = d. \end{cases}$$

Note that our assumptions on the random sequence $i_1, \ldots, i_d$ and the random function j, imply that for $1 \leq t < \bar{t}$, $j(i_t) + t \neq j(i_{\bar{t}}) + \bar{t}$, and thus the above transitions are well defined. In all other cases the automaton moves to the punishing state $\emptyset$.

LEMMA 14. *For every strategy $\sigma \in \Sigma^1$,*

$$\text{comp}(\sigma) \geq 3l_1|Q(1, \sigma)| + l_1|Q(2, \sigma)|,$$

where $Q(1, \sigma)$ and $Q(2, \sigma)$ are defined as in (10) *and* (11), *respectively*.

Let $|c^*|$ denote the length of the play c^*, i.e., $|c^*| = \Sigma_{i=1}^{d}(d_3 + d_2 + i)$. Consider the two sets

$$X = Q(1, \sigma) \times (\{4k + |c^*|, \ldots, 4k + l - 1\} + \{0, l, 2l\})$$

and

$$Y = Q(2, \sigma) \times \{4k + |c^*|, \ldots, 4k + l - 1\}.$$

By the definition of the complexity of a strategy, it suffices to show that for every pair $(\epsilon, t) \neq (\epsilon', t')$ with $t \geq t'$ in the union of the two sets, $X \cup Y$, $(\omega_1(\epsilon), \ldots, \omega_t(\epsilon))$ is compatible with σ, and

$$(16) \qquad (\sigma|\omega_1(\epsilon), \ldots, \omega_t(\epsilon)) \neq (\sigma|\omega_1(\epsilon'), \ldots, \omega_{t'}(\epsilon')).$$

For every $\epsilon \in Q(1, \sigma)$, $r^1(\sigma, \tau^\epsilon) > R^1(\omega(\epsilon))$ and therefore there is a deviation from the proposed play but no deviation prior to stage $4k + 4l$, i.e., for every $\epsilon \in Q(1, \sigma)$ and every $t \leq 4k + 4l$, $\omega_t(\sigma, \tau^\epsilon) = \omega_t(\epsilon)$. In particular, $(\omega_1(\epsilon), \ldots, \omega_{4k+4l}(\epsilon))$ is compatible with σ. For every $\epsilon \in Q(2, \sigma)$, $r^1(\sigma, \tau^\epsilon) > R^1(\omega(\epsilon)) - (x^1 - u^1(G))/3$ and therefore no deviation from the proposed play prior to stage $4k + 2l$, i.e., for every $\epsilon \in Q(2, \sigma)$ and every $t \leq 4k + 2l$, $\omega_t(\sigma, \tau^\epsilon) = \omega_t(\epsilon)$. In particular $(\omega_1(\epsilon), \ldots, \omega_{4k+2l}(\epsilon))$ is compatible with σ.

The play $(\omega_{4k+1+|c^*|}(\epsilon), \ldots, \omega_{4k+l+d_3}(\epsilon))$ is a coordinated play with the first $d_0 - 2k$ and the last d_3 action pairs being $(0, 0)$ and $\omega_{4k+l}(\epsilon) = (1, 1)$. As $d_3 > 2k$, the string $(1, 1) + d_3 * (0, 0)$ appears only at the end of the play, and therefore, if $4k + |c^*| \leq t' < t < 4k + l$,

$$(\omega_{t+1}(\epsilon), \ldots, \omega_{4k+l}(\epsilon)) \neq (\omega_{t'+1}(\epsilon'), \ldots, \omega_{4k+l+t'-t}(\epsilon)).$$

As each one of these two plays is a coordinated play, (16) follows. The same argument applies for all pairs $(\epsilon, t) \neq (\epsilon', t')$ with $t \neq t' (\text{mod } l)$.

We consider next the case $t = t' \pmod{l}$ and $\epsilon \neq \epsilon'$. Note that the play c^* is independent of ϵ and therefore so is the play

$$(\omega_{4k+l+1}(\epsilon), \ldots, \omega_{4k+l+|c_*|}(\epsilon)) = c^*.$$

On the other hand,

$$(\omega_{4k+l+|c_*|+1}(\epsilon), \ldots, \omega_{4k+2l}(\epsilon)) \quad \text{is a coordinated play.}$$

Therefore, if $t = t' \pmod{l}$ and $\epsilon \neq \epsilon'$,

$$(\omega_{t+1}(\epsilon), \ldots, \omega_{t+l}(\epsilon)) \neq (\omega_{t'}(\epsilon'), \ldots, \omega_{t'+l}(\epsilon')).$$

Let s be the smallest nonnegative integer with

$$(\omega_{t+1}(\epsilon), \ldots, \omega_{t+s}(\epsilon)) \neq (\omega_{t'+1}(\epsilon'), \ldots, \omega_{t'+s}(\epsilon')).$$

It follows that $\omega^1_{t+s}(\epsilon) \neq \omega^1_{t'+s}(\epsilon')$ and thus (16) holds.

Assume next that $t > t'$, $t - t' = 0 \pmod{l}$ and $\epsilon = \epsilon' \in Q(1, \sigma)$. Let s be the smallest positive integer such that

$$\omega^1_{t+s}(\epsilon) \neq \omega^1_{t'+s}(\sigma, \tau^{\epsilon}).$$

As $r^1(\sigma, \tau^{\epsilon}) \neq R^1(\omega(\epsilon))$, $s \leq T - t$ and

$$(\omega_{t+1}(\epsilon), \ldots, \omega_{t+s-1}(\epsilon)) = (\omega_{t'+1}(\epsilon'), \ldots, \omega_{t'+s-1}(\epsilon')),$$

and thus (16) follows. □

LEMMA 15 *For every* $\sigma \in \Sigma^1(m_1)$,

$$r^1_T(\sigma, \tau^*) \leq \sum_{t=1}^{T} \frac{r^1(\omega_t(\epsilon))}{T}.$$

PROOF. Let σ be a pure strategy for Player 1 with

$$r^1_T(\sigma, \tau^*) \geq \sum_{t=1}^{T} \frac{r^1(\omega_t(\epsilon))}{T},$$

and such that σ is implemented by an automaton of size m_1. Recall that σ satisfies the inequality (9). Define $Q(1, \sigma)$, $Q(2, \sigma)$, and $Q(3, \sigma)$ as in (10), (11), and (12), respectively. By Lemma 14,

$$m_1 \geq \text{comp}(\sigma) \geq 3l_1|Q(1, \sigma)| + l_1|Q(2, \sigma)|. \tag{17}$$

Either $Q(3, \sigma) = \emptyset$, and then $|Q(1, \sigma)| + |Q(2, \sigma)| = |Q|$. However, $\text{comp}(\sigma) \leq m_1$ which is compatible with (15) and (17) only if $Q(1, \sigma) = \emptyset$ and then $r^1_T(\sigma, \tau^*) \leq r^1_T(\sigma^*, \tau^*)$. Or, $Q(3, \sigma) \neq \emptyset$. Then, it follows from (17) and (15), that $|Q(3, \sigma)| l_1 + l_1 \geq 2l_1|Q(1, \sigma)|$, and therefore for sufficiently large T, $r^1_T(\sigma, \tau^*) \leq r^1_T(\sigma^*, \tau^*)$. □

Next we will prove that τ^* is a best reply of Player 2 to σ^*.

LEMMA 16. *For any* $\tau \in \Sigma^2$,

$$r^2(\sigma^*, \tau) \le r^2(\sigma^*, \tau^*) = R^2(\omega(\epsilon)).$$

PROOF. Let τ be a pure strategy of Player 2. Note that the induced play $\omega(\sigma^*, \tau)$ is a random sequence. We will prove that

$$E(R^2(\omega(\sigma^*, \tau))) = r^2(\sigma^*, \tau) \le r^2(\sigma^*, \tau^*).$$

Assume first that there is $\epsilon \in Q$ with

$$\omega_t(\sigma^*, \tau) = \omega_t(\epsilon) \quad \forall t \le 2k.$$

Note that for any $t \le T$ and any $a \in A^2$,

$$r^2(\omega_t^1(\epsilon), a) + (T - t)u^2(G) \le \sum_{s=t}^{T} r^2(\omega_s(\epsilon)).$$

Let t be the smallest integer with $\omega_t(\sigma^*, \tau) \ne \omega_t(\epsilon)$. Then, if $\omega_t(\epsilon) = (1, 1)$ Player 1 will trigger to punishing forever, and then $r_T^2(\sigma^*, \tau) \le r_T^2(\sigma^*, \tau^*)$. If $\omega_t(\epsilon) = (0, 0)$ then $t \le T - \lambda_1 l/3$ and with probability close to one, Player 1 triggers to punishing forever resulting in future losses to Player 2 of $2\epsilon(\lambda_1 - \epsilon^2)l$. The one time deviation can generate a gain of the order of a constant, and no sufficient certainty to generate additional gains in the future. Therefore $r_T^2(\sigma^*, \tau) \le r_T^2(\sigma^*, \tau^*)$. If $\omega_t(\epsilon) = (0, 1)$ then $t \le T - d_1$ and with probability close to one, Player 1 triggers to punishing forever in the next $d_0 - 2k - 1$ stages. As L is sufficiently large, and d_0 and d_1 are approximately $l\lambda_0/L$ and $l\lambda_1$, respectively, it follows that $r_T^2(\sigma^*, \tau) \le r_T^2(\sigma^*, \tau^*)$. Finally, observe that if $(\omega_1^2(\sigma^*, \tau), \dots, \omega_{2k}^2(\sigma^*, \tau) \notin \theta(Q)$, then with probability at least 1/2 Player 1 will trigger to punishing forever in one of the next $2Lk$ stages, and the possible gains from such a deviation are offset by the loss in case Player 1 realizes a deviation. Altogether, we conclude that (σ^*, τ^*) is an equilibrium of $G^T(m_1, m_2)$. If $r^2(1, 1) \le u^2(G) + 2\epsilon$ then one modifies the proposed play so that the game ends with the last string $(d_0 - 2k) * (0, 0)$ (if $r^2(0, 0) \ge r^2(0, 1)$), or the last string $d_2 * (0, 1)$. For example, by adding the play $d_4 * (1, 1)$ following the communication phase where $d_4 < d_1$ and if needed changing δ. The proof in the subcases 3.2 when $|\{a_1^2, a_2^2, a_3^2\}| = 3$, follows the same lines as our present proof: Without loss of generality $a_1 = (0, 0)$, $a_2 = (0, 1)$ and $a_3 = (1, 2)$. The action pairs $(0, 1)$ in the communication play are replaced by $(0, 2)$ and the action pairs $(1, 1)$ in the proposed play are replaced by $(1, 2)$. In subcase 3.1, one adds an action pair a_4 with $a_4^1 \ne a_1^1$ and approximate the vector payoff x by a convex combination of $r(a_i)$, $1 \le i \le 4$, and the rest of the proof is very similar to our handling of subcase 3.2. In case 3.3, assume first that $a_1^2 = a_2^2 = a_3^2$. W.l.o.g. $r^2(a_1) \ge r^2(a_2) \ge r^2(a_3)$, $a_1 = (0, 1)$, $a_2 = (1, 1)$, $a_3 = (2, 1)$, $a_4 = (0, 0)$ and 0 is a best reply of Player 2 to the action 0 of Player 1. One approximates x as a convex combination of $r(a_i)$, $1 \le i \le 4$, and designs a similar proposed play as in our case 3.2, making sure that the synchronization is such that the game ends with a long string of $(0, 0)$. This completes the proof of Theorem 1 under the condition $m_2 > T$. If $m_2 \le T$ one either makes the modification indicated at the close of the proof of case 1), or refers to the next comment.

Any payoff $x \in \operatorname{co} r(A)$ is an average $\sum_{a \in A} \lambda_a r(a)$ where for every $a \in A$ $\lambda_a > 0$ and $\sum_{a \in A} \lambda_a = 1$. It should be clear from our proof of subcase 3.2 that one can actually generate an equilibrium of $G^T(m_1, m_2)$ that consists of a communication phase and the play in the cycle runs over all action pairs $a \in A$ with frequencies which are approximately λ_a. Moreover, one can synchronize the play so that the game terminates with a long string of best replies of Player 2 yielding him a payoff above his individual rational one. This completes the proof of Theorem 1. □

PROOF OF THEOREMS 2 AND 3. Theorems 2 and 3 (assuming that there is $x \in \operatorname{co}(r(A))$ with $x^i > v^i(G)$) follow directly from our proof of Theorem 1 together with Theorem 5, by replacing the single punishing state in the automaton of each player with a set of punishing states of size l_1.

Assume next that there is no $x \in \operatorname{co}(r(A))$ with $x^i > v^i(G)$. Either $x = (v^1(G), v^2(G))$ for every $x \in \operatorname{co}(r(A))$ with $(x^1, x^2) \geq (v^1(G), v^1(G))$, and then the theorem follows directly from Theorem 5, or we may assume without loss of generality that $x^2 \leq v^2(G)$ for every $x \in r(A)$. Therefore it must be the case that there is a strategy, say 1, of Player 2, such that for every strategy c of Player 1, $r^2(c, 1) = v^2(G)$. Assume that $b \in A^1$ maximizes $r^1(c, 1)$, and $a \in A^1$ minimizes $r^1(c, 1)$ with $a \neq b$. Note that it follows that $v^1(G) \leq r^1(b, 1)$. The play $T * (b, 1)$ is an equilibrium play of $G^T(m_1, m_2)$. If $x = \lambda_a r^1(a, 1) + \lambda_b r^1(b, 1)$ with $x^1 > v^1(G)$, $\lambda_a, \lambda_b > 0$ and $\lambda_a + \lambda_b = 1$. Then a play of the form $d_0 * (b, 1) + (L * (d_a * (a, 1) + d_b * (b, 1))$ is an equilibrium play of $G^T(m_1, m_2)$ for d_b sufficiently large, $x^1 < R^1(d_a * (a, 1) + d_b * (a, 1)) < x^1 + \epsilon/2$, L sufficiently large and $d_0 + L(d_a + d_b) = T$. This illustrates the result for payoff vectors in the interval $[r(a, 1), r(b, 1)]$. Assume next that there is an action pair $(0, 0)$ with $r^1(0, 0) > r^1(b, 1)$ and $r^2(0, 0) = r^2(b, 1)$. Then the action pair $(0, 0)$ differs from either $(b, 1)$ or $(a, 1)$ in each coordinate. Any individual rational payoff is either in the interval $[r(a, 1), r(b, 1)]$ or a convex combination of the payoffs to two strategy pairs $(0, 0)$ and $(b, 1)$. Label the action b by 1 if $b \neq 0$, or label the action a by 1 if $b = 0$, and construct an equilibrium of $G^T(m_1, m_2)$ in which the communication and verification phases are the same as in our proof of Theorem 1 and the cycle consists of one string of the strategy pair $(0, 0)$ and another string of the strategy pair $(1, 1)$. □

PROOF OF THE MAIN THEOREM. Let $G = (\{1, 2\}, A, r)$ be a two person game in strategic form, and let $(T, m_1(T), m_2(T))_{T=1}^{\infty}$ be a sequence of triples of positive integers with $\min_{i=1,2} m_i(T) \to \infty$ as $T \to \infty$, and

$$\lim_{T\to\infty} \frac{\log \max_{i=1,2} m_i(T)}{\min(m_1(T), m_2(T), T)} = 0.$$

By Theorem 3 it follows that

$$\liminf_{T\to\infty} E(G^T(m_1(T), m_2(T))) \supset E(G^*_\infty).$$

On the other hand, any $x \in E(G^T(m_1(T), m_2(T)))$ is obviously in $\operatorname{co}(r(A))$, and by Theorem 5, for every $\epsilon > 0$, $x^i > v^i(G) - \epsilon$ for sufficiently large values of T. Therefore,

$$\limsup_{T\to\infty} E(G^T(m_1(T), m_2(T))) \subset E(G^*_\infty)$$

which together with the previous inclusion proves the Main Theorem. □

8. Remarks. It is of interest to complete the study of the asymptotics of the equilibrium payoff sets $E(G^T(m_1, \ldots, m_n))$ of finitely repeated n-person games. Several extensions need only minor modifications. For example, the generalization of Theorem 1 for the case that $m_1 = \cdots = m_{n-1} \leq m_n$ and $x = \Sigma_{i=1}^k \lambda_i r(a_i)$ with $\lambda_i \geq 0$, $\Sigma_{i=1}^k \lambda_i = 1$ and $|\{a_i^j \mid 1 \leq i \leq k\}| = k$ for every $1 \leq j \leq n$. Other extensions are more intricate. The conclusion of Theorem 1 continues to hold for n-person finitely repeated games under the assumptions that $m_1 \leq m_2 \leq \cdots \leq m_{n-1} \leq m_n$ and $m_{n-1} \leq \exp(\epsilon^3 T)$, and that there are points $y, z \in \operatorname{co}(r(A))$ with $y^i < z^i$, $i = h - 1, n$. There are, however, difficulties in

0, 0, 0	0, 0, 8
0, 0, 8	0, 0, 8

0, 0, 8	0, 0, 8
0, 0, 8	0, 0, 0

3, 3, 5	3, 3, 5
3, 3, 5	3, 3, 5

FIGURE 3.

extending our main theorem to n-person games. The naive generalization is not correct, however. Consider the 3-player game G of Figure 3.

Player 1 chooses the row, Player 2 the column, and Player 3 chooses the matrix. Note that $v^1(G) = 0 = v^2(G)$ and $v^3(G) = 6$. Thus $E(G_\infty^*)$ is the set co$\{(0, 0, 8), (0, 0, 6), (2, 2, 6)\}$. In particular, $(3, 3, 5) \notin E(G_\infty^*)$. Denote by $w^i(G)$ the max min of Player i where he maximizes over his mixed strategies and the min is over the pure strategies of the other players. In the game above, $w^3(G) = \max_{x \in \Delta(A^3)} \min_{(a^1,a^2) \in A^{1} \times A^{2}} r^1(a^1, a^2, x) = 4$ and thus by using either Proposition 2 or Proposition 3 of Neyman (1997) one may construct sequences $m_1(T)$, $m_2(T)$ and $m_3(T)$, with $m_1(T) \le m_3(T)$ and $\lim_{T\to\infty} \min\{m_1(T), m_2(T)\} = \infty$ such that $(3, 3, 5) \in \lim\sup E(G^T(m_1(T), m_2(T), m_3(T))$. For example, if $(m_1(T) \log m_1(T))/\min(T, m_2(T)) \to 0$ as $T \to \infty$, repeated play of the right matrix is the outcome of a pure strategy equilibrium of $G^T(m_1(T), m_2(T), m_3(T))$ and if $m_2(T) \le m_1(T) = o(T)$ with $\log m_1(T)/\min(m_2(T), \log m_3(T)) \to 0$ as $T \to \infty$, repeated play of the right matrix is the outcome of a mixed strategy equilibrium of $G^T(m_1(T), m_2(T), m_3(T))$. We hope to provide details of our findings for n-person finitely repeated games with finite automata in the future.

Acknowledgment. The author wishes to thank J. Hillas for his many valuable suggestions and his help in preparing the revision of this paper.

References

Aumann, R. J. (1959). Acceptable points in general cooperative n-person games. *Contributions to the Theory of Games, Vol. IV, Ann. of Math. Stud.*, Vol. 40, Princeton University Press, Princeton, New Jersey, 287–324.

——— (1960). Acceptable points in games of perfect information. *Pacific J. Math.* **10** 381–387.

——— (1981). Survey of Repeated Games, in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern* 11–42, Wissenschaftsverlag, Bibliographisches Institut, Mannheim, Wien, Zurich.

———, L. S. Shapley (1994). Long term competition—A game theoretic approach. N. Meggido, ed., *Essays in Game Theory in Honor of Michael Maschler*, Springer-Verlag, Berlin, 1–15.

Benoit, J.-P., V. Krishna (1985). Finitely repeated games. *Econometrica* **53** 905–922.

———, and ——— (1987). Nash equilibria of finitely repeated games. *Internat. J. Game Theory* **16** 197–204.

Ben-Porath, E. (1993). Repeated games with finite automata. *J. Econom. Theory* **59** 17–32.

Fudenberg, D., E. Maskin (1986). The folk theorem in repeated games with discounting and with incomplete information. *Econometrica* **54** 533–554.

Gossner, O. (1995). The folk theorem for finitely repeated games with mixed strategies, *Internat. J. Game Theory* **24** 95–107.

Kalai, E. (1990). Bounded rationality and strategic complexity in repeated games. T. Ichiishi, A. Neyman, Y. Tauman, eds., *Game Theory and Applications*, Academic Press, New York, 131–157.

Megiddo, N., A. Wigderson (1986). On play by means of computing machines. J. Halpern, ed. *Proc. 1st Cong. on Theoretical Aspects of Reasoning and Knowledge*, Morgan Kaufman, Los Altos, California, 259–274.

Neyman, A. (1985). Bounded complexity justifies cooperation in the finitely repeated prisoner's dilemma. *Economics Letters* **19** 227–229.

——— (1997). Cooperation, repetition, and automata. S. Hart, A. Mas Colell, eds., *Cooperation: Game-Theoretic Approaches*, NATO ASI Series F, **155**. Springer-Verlag, 233–255.

———, D. Okada (1996). *Strategic entropy and complexity in repeated games*, DP 104, Center for Rationality and Interactive Decision Theory, Hebrew University, Jerusalem. Forthcoming in *Games and Economic Behavior*.

Papadimitriou, C. H., M. Yannakakis (1994). On complexity as bounded rationality, (extended abstract). STOC-94, 726–733.

———, ——— (1995, 1996). On bounded rationality and complexity, manuscript (1995, revised 1996).
Rubinstein, A. (1994). Equilibrium in supergames. N. Meggido, ed., *Essays in Game Theory in Honor of Michael Maschler*, Springer-Verlag, Berlin, 17–27.
Sorin, S. (1986). On repeated games with complete information. *Math. Oper. Res.* **11** 147–160.
——— (1990). Supergames. T. Ichiishi, A. Neyman, Y. Tauman, eds., *Game Theory and Applications*, Academic Press, New York, 46–63.
——— (1992). Repeated games with complete information, Chapter 4. R. J. Aumann, S. Hart, eds., *Handbook of Game Theory*, North-Holland, Amsterdam, 71–107.
Zemel, E. (1989). Small talk and cooperation: A note on bounded rationality, *J. Economic Theory* **49**(1) 1–9.

A. Neyman: Institute of Mathematics, Hebrew University, Givat Ram, 91904 Jerusalem, Israel