MATHEMATICS OF OPERATIONS RESEARCH
Vol. 31, No. 3, August 2006, pp. 637–648
ISSN 0364-765X | EISSN 1526-5471 | 06 | 3103 | 0637

DOI 10.1287/moor.1060.0211

Excludability and Bounded Computational Capacity

Ehud Lehrer, Eilon Solan
School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel
{lehrer@post.tau.ac.il, eilons@post.tau.ac.il}

We study the notion of excludability in repeated games with vector payoffs, when one of the players is restricted to strategies with bounded computational capacity. We show that a closed set is excludable by Player 2 when Player 1 is restricted to using only bounded-recall strategies if and only if it does not contain a convex approachable set. We provide partial results when Player 1 is restricted to using strategies that can be implemented by automata.

Key words: repeated games; vector payoffs; Blackwell's approachability; excludability; bounded-recall strategies; finite automata
MSC2000 subject classification: Primary: 91A20, 91A05
OR/MS subject classification: Primary: games/group decisions; secondary: noncooperative
History: Received September 9, 2003; revised March 15, 2005, and November 2, 2005.

1. Introduction. In a seminal paper, Blackwell [2] introduced and studied the notions of approachability and excludability in repeated games with vector payoffs. These notions are analogues to the max-min and the min-max levels in repeated games with scalar payoffs.

In repeated games with vector payoffs the outcome of players' actions is not a scalar utility for each player, but rather a vector in some Euclidean space. A set of vectors is *approachable* by Player 1 if he has a strategy that guarantees, with arbitrarily high probability, that the average (vector) payoff is asymptotically close to the set, regardless of the strategy employed by Player 2. A set is *excludable* by Player 2 if she has a strategy that guarantees, with arbitrarily high probability, that the average payoff is asymptotically bounded away from the set, regardless of the strategy used by Player 1.

Blackwell [2] provided a geometric condition that guarantees that a set is approachable, and proved that any convex set is either approachable by Player 1, or excludable by Player 2. Hou [5] and Spinat [14] fully characterized the family of approachable sets. Both proved that if a closed set is *minimally* (w.r.t. set inclusion) approachable, then it satisfies Blackwell's geometric condition. Vieille [15] studied the notions of weak approachability and weak excludability, which were also introduced by Blackwell [2], and proved that any set is either weakly approachable or weakly excludable. For partial results on weak approachability for two-dimensional games where both players have two actions, see Hou [4].

Our objective is to study repeated games with vector payoffs when one of the players is restricted to strategies with bounded computational capacity. Two classes of such strategies were extensively studied in the literature: those that can be implemented by finite automata (see, e.g., Neyman [9], Rubinstein [11], Ichiishi et al. [6]), and those with bounded recall—that is, strategies that may condition only on actions that were played in the last t stages, for some fixed integer t (see, e.g., Lehrer [7], Aumann and Sorin [1]).

In a companion paper (Lehrer and Solan [8]), we study the notions of approachability with automata and with bounded-recall strategies. We characterized the sets that are approachable by Player 1 if he may only use strategies that can be implemented by automata, or bounded-recall strategies (and Player 2 is not restricted). We proved that the following three statements are equivalent for closed sets.

- F is approachable with automata by Player 1.
- F is approachable with bounded-recall strategies by Player 1.
- F contains a convex set that is approachable by Player 1.

In the present paper we concentrate on the notions of excludability against automata and against bounded-recall strategies. A set is *excludable against automata* (respectively, *against bounded-recall strategies*) by Player 2 if she has a strategy that ensures that when Player 1 plays a finite automaton (respectively, a bounded-recall strategy), the long-run average payoff is bounded away from the set.

Our first result is a complete characterization of the sets that are excludable against bounded-recall strategies: A closed set is excludable against bounded-recall strategies by a player if and only if it is not approachable with bounded-recall strategies by that player—namely, if and only if it does not contain any convex set that is approachable by the player.

We then provide partial results concerning sets that are excludable against automata. We do not know whether any set that is not approachable with automata by Player 1 is excludable against automata by Player 2. Solving this problem will reveal whether the two notions related to strategies with bounded computational capacity are equivalent.

2. The model and the main results. In this section we define the model of repeated games with vector payoffs, and two types of bounded computational capacity strategies. We then define the notion of excludability against such strategies, and finally, we state our results.

2.1. Repeated games with vector payoffs. A two-player *repeated game with vector payoffs* is a triplet (I, J, U), where I and J are finite sets of actions of the players, and $U = (u_{i,j})_{i\in I,\, j\in J}$ is a d-dimensional vector-payoff matrix: $u_{i,j} \in \mathbb{R}^d$ for every $i \in I$ and $j \in J$. We assume w.l.o.g. that all the payoffs are in the unit ball of $\mathbb{R}^d$. We denote by $\Delta(I)$ and $\Delta(J)$ the players' mixed action sets (that is, the sets of probability distributions over I and J, respectively). By $\mathcal{S}$ and $\mathcal{T}$ we denote the sets of mixed strategies in the repeated game of the Players 1 and 2, respectively.

The pair (i_k, j_k) denotes the joint action played by Players 1 and 2 at stage k. The average (vector) payoff up to stage n is $\bar{x}_n = \sum_{k=1}^n u_{i_k, j_k}/n$.

2.2. On bounded computational capacity strategies. Let t be a natural number. A *t-bounded-recall strategy* of Player 1 is a pair (m, σ), where $m \in (I \times J)^t$ and σ: $(I \times J)^t \to \Delta(I)$. When playing a t-bounded-recall strategy (m, σ), at any stage Player 1 plays $\sigma(w)$, where w is the string of the last t joint actions. He starts the game with the (virtual) memory m. Thus, at the first stage he plays the mixed action $\sigma(m)$, at the second stage he plays $\sigma(m', i_1, j_1)$, where m' are the last $t-1$ coordinates of m and (i_1, j_1) is the realized pair of actions of the two players at the first stage, and so on.[1]

A *bounded-recall* strategy is a strategy that is t-bounded recall for some $t \in \mathbb{N}$. We denote by $\mathcal{S}_{BR}$ the set of all bounded-recall strategies of Player 1.

A (nondeterministic) *automaton A* is given by (i) a finite set of states; (ii) a probability distribution over the set of states, according to which the initial state is chosen; (iii) a finite set of inputs; (iv) a finite set of outputs; (v) a function that assigns to every state a probability distribution over outputs; and finally (vi) a transition rule that assigns to every state and every input a probability distribution over states.

When the set of outputs coincides with the set I of actions of Player 1, and the set of inputs coincides with the set $I \times J$ of action pairs, an automaton implements a strategy of Player 1 as follows. The initial state of the automaton is chosen according to the initial distribution given in (ii). At every stage, as a function of the current state an action of Player 1 is chosen by the probability distribution given in (v), and a new state is chosen as a function of the pair of actions played (by both players), according to the probability distribution given in (vi). $\mathcal{S}_A$ denotes the set of Player 1's strategies that can be implemented by an automaton.

Observe that every t-bounded-recall strategy can be implemented by an automaton with $|I \times J|^t$ states.

2.3. Excludability against bounded-capacity strategies. Let $d(x, y)$ denote the Euclidean distance between the points x and y in $\mathbb{R}^d$. For every set F in $\mathbb{R}^d$ and every $x \in \mathbb{R}^d$, let $d(x, F) = \inf_{y\in F} d(x, y)$ be the distance between x and F. For every $\delta > 0$, let $B(F, \delta) = \{x \in \mathbb{R}^d\colon d(x, F) \le \delta\}$ be the closed set of all points that are δ-close to F.

Blackwell [2] defined the notion of excludability in repeated games with vector payoffs. A set F is excludable by Player 2 if she can guarantee with arbitrarily high probability that the average payoff is asymptotically bounded away from F.

Definition 2.1 (Blackwell 1956). A set F is *excludable* by Player 2 if there exists a strategy $\tau \in \mathcal{T}$ such that

$$\exists\, \varepsilon > 0, \forall\, \eta > 0, \exists\, N \in \mathbb{N}, \forall\, \sigma \in \mathcal{S}, \quad \mathbf{P}_{\sigma,\tau}\Big(\inf_{n\ge N} d(\bar{x}_n, F) < \varepsilon\Big) < \eta,$$

where $\mathbf{P}_{\sigma,\tau}$ is the probability distribution over the space of infinite plays induced by the pair (σ, τ).

We are interested in studying when a given set is excludable by Player 2, provided that Player 1 is restricted to using bounded-recall strategies, or strategies that can be implemented by automata.

Definition 2.2. (1) A set F is *excludable against bounded-recall strategies* by Player 2 if there exists a strategy $\tau \in \mathcal{T}$ such that

$$\exists\, \varepsilon > 0, \forall\, \eta > 0, \forall\, \sigma \in \mathcal{S}_{BR}, \exists\, N \in \mathbb{N}, \quad \mathbf{P}_{\sigma,\tau}\left(\inf_{n\ge N} d(\bar{x}_n, F) < \varepsilon\right) < \eta.$$

(2) The set is *excludable against automata* if a similar condition holds when $\mathcal{S}_{BR}$ is replaced by $\mathcal{S}_A$.

[1] One could replace the virtual memory used in the first t stages by t functions σ_r: $(I \times J)^{r-1} \to \Delta(I)$ for $r = 1, \ldots, t$; the function σ_r is used to choose the action at stage r. Our results hold with this more general definition of a bounded-recall strategy.

Observe that in this definition N may depend on the strategy used by Player 1, whereas in Definition 2.1 it does not. If N is required to be independent of the strategy employed by Player 1, excludability against bounded-recall strategies (or against automata) turns out to be equivalent to excludability. Nevertheless, it is desirable that N depends only on the *size* of the memory of the strategy (or on the size of the automaton), and not on the strategy itself. Studying excludability against bounded computational capacity strategies under this stronger definition is left for future research.

2.4. Results and examples. To present our result concerning excludability against bounded-recall strategies, we need the notion of approachability. A set F is approachable if Player 1 can guarantee with arbitrarily high probability that the long-run average payoff will be arbitrarily close to F. Formally,

Definition 2.3 (Blackwell 1956). A set F is *approachable* by Player 1 if there exists a strategy $\sigma \in \mathcal{S}$ such that

$$\forall \varepsilon > 0, \forall \eta > 0, \exists N \in \mathbb{N}, \forall \tau \in \mathcal{T}, \quad \mathbf{P}_{\sigma,\tau}\left(\sup_{n\geq N} d(\bar{x}_n, F) \geq \varepsilon\right) < \eta.$$

In this case we say that σ *approaches* F.

Blackwell [2] provided a geometric condition that guarantees that a set is approachable.

2.4.1. Excludability against bounded-recall strategies. Our main result concerning excludability against bounded-recall strategies, which is proved in §3.1, is the following.

Theorem 2.1. *A closed set that does not contain a convex set that is approachable by Player* 1 *is excludable against bounded-recall strategies by Player* 2.

The dual notion to excludability against bounded-recall strategies is approachability with bounded-recall strategies.

Definition 2.4 (Lehrer and Solan 2003). A set F is *approachable with bounded-recall strategies* by Player 1 if for every $\delta > 0$ there exists a bounded-recall strategy σ that approaches $B(F, \delta)$.

We note that a set cannot be both approachable with bounded-recall strategies and excludable against bounded-recall strategies.

Together with Lehrer and Solan [8, Proposition 2], Theorem 2.1 implies the following characterization of sets that are excludable against bounded-recall strategies.

Corollary 2.1. *The following three statements are equivalent for a closed set* $F \subseteq \mathbb{R}^d$:
- F *is excludable against bounded-recall strategies by Player* 2.
- F *is not approachable with bounded-recall strategies by Player* 1.
- F *does not contain a convex set which is approachable by Player* 1.

As Hou [5] and Spinat [14] provided a geometric characterization to approachable sets, Corollary 2.1 provides a geometric characterization for sets that are excludable against bounded-recall strategies.

If in addition F is convex, then we have the following result.

Corollary 2.2. *A convex set is excludable by Player* 2 *if and only if it is excludable against bounded-recall strategies by Player* 2.

Indeed, by Blackwell [2] a convex set F is either approachable by Player 1 or excludable by Player 2. In the former case, because F is convex, it contains a convex set that is approachable by Player 1, so that by Corollary 2.1 it is not excludable against bounded-recall strategies by Player 2. In the latter case, because it is excludable by Player 2, it is in particular excludable against bounded-recall strategies by Player 2.

2.4.2. Excludability against automata. We now present our results concerning excludability against automata. We start with some definitions.

For every pair (p, q) of mixed actions, denote by $u_{p,q} = \sum_{i,j} p_i u_{i,j} q_j$ the expected vector payoff. This is the expected stage-payoff when Player 1 plays the mixed action p and Player 2 plays the mixed action q.

For any mixed action q of Player 2, denote

$$H(q) = \{u_{p,q}\colon p \text{ is a mixed action of Player 1}\}.$$

When Player 2 plays the mixed action q, the expected stage-payoff is always in $H(q)$, regardless of the mixed action chosen by Player 1.

Definition 2.5. Let x be a point in a set F. x is *sharp* in F if for every $\varepsilon > 0$ there is some mixed action q of Player 2 such that $H(q) \cap F \subseteq B(x, \varepsilon)$.

Because the family $\{H(q); q$ is a mixed action of Player 2$\}$ depends on the game, the property of being sharp is relative to the game.

A point x is sharp in F if for every $\varepsilon > 0$ Player 2 has a stationary strategy that guarantees that the long-run average payoff is either ε-close to x, or outside F.

Example 2.1. Consider the following game, where both players have two actions, payoffs are two dimensional, and the payoff matrix is

	L	R
T	$-1, 1$	$1, -1$
B	$2, 2$	$-1, -1$

For every mixed action $q = (q_L, 1 - q_L)$ of Player 2, where q_L is the probability that Player 2 plays the left column,

$$H(q) = \text{conv}\{(1 - 2q_L, -1 + 2q_L), (-1 + 3q_L, -1 + 3q_L)\}.$$

The following figure (Figure 1) depicts a few representatives of the family $\{H(q), 0 \le q \le 1\}$ (in light lines), together with two sets (in gray line).

The set F_1 has no sharp points, while all the points in F_2 are sharp.

Our main result concerning excludability against automata, which is proved in §3.2, is the following.

Theorem 2.2. *Let F be a closed set. If the convex hull of the sharp points in F is not contained in F, then F is excludable against automata.*[2]

By Theorem 2.2 the set F_2 in Figure 1(B) is excludable against automata by Player 2. To explain the main ideas in the proof of Theorem 2.2, we elaborate on Example 2.1.

Example 2.1—continued. Let $F_2 = [(0, 0), (2, 2)] \cup [(0, 0), (1, -1)]$ be the set that appears in Figure 1(B). One can verify that F_2 satisfies Blackwell's [2] geometric condition, and therefore it is approachable by Player 1. Theorem 2.2 F_2 is excludable against automata by Player 2. We now explain how to construct a strategy of Player 2 that ensures the long-run average payoff remains far from F_2, provided that Player 1 uses strategies that can be implemented by automata.

Let R^* (respectively, L^*) be the stationary strategy of Player 2 in which she always plays the right column (respectively, left column). If Player 2 uses one of these strategies, the evolution of the state of the automaton that implements Player 1's strategy follows a Markov chain. Thus, the average payoff converges to a limit. Because $H(R) \cap F_2$ consists only of the point $(1, -1)$, when Player 2 uses R^*, this limit is either out of F_2, or equal to $(1, -1)$. Similarly, because $H(L) \cap F_2 = (2, 2)$, when Player 2 uses L^* the limit is either out of F_2, or equal to $(2, 2)$.

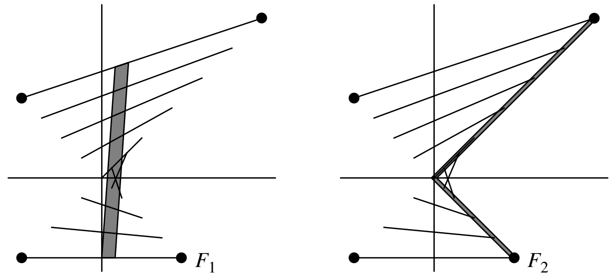
Player 2 is going to play in blocks whose length is polynomial in k. In each block Player 2 follows either R^* or L^*. If the average payoff within block k is far from F_2, Player 2 plays in block $k + 1$ the same strategy he played along block k. If the average payoff within block k is close to F_2, Player 2 switches in block $k + 1$ to the other strategy.

Let us now verify that the average payoff is asymptotically far from F_2. If the limit payoff is far from F_2, with high probability the average payoff within the block is far from F_2. The definition of the strategy implies that Player 2 will keep on playing the same stationary strategy, so that the long-run average payoff converges to the limit, which is outside F_2. Otherwise, Player 2 alternates between playing L^* (within such blocks the average payoff is close to $(2, 2)$) and playing R^* (within such blocks the average payoff is close to $(1, -1)$). Therefore, the long-run average payoff converges to $\frac{1}{2}(1, -1) + \frac{1}{2}(2, 2) = (\frac{3}{2}, \frac{1}{2})$, which is outside F_2.

In general, the length of block k slowly increases with k. It increases, so that whichever automaton Player 1 uses, for k is sufficiently large, the average payoff within block k is close to the limit payoff. It increases slowly, so that the ratio between the length of block k and the total length of all preceding blocks goes to zero. This property is needed so that the average payoff will indeed be close to the desired convex combination.

To make the discussion complete, we briefly mention the notion of excludability with a bounded computational capacity strategy. A set is *excludable with a bounded-recall strategy* (respectively, *with automata*) by Player 1 if Player 1 has a bounded-recall strategy (respectively, a strategy that can be implemented by an automaton) that ensures the long-run average payoff remains far from the set. Theorem 1 in Lehrer and Solan [8] implies the following proposition, which makes the connection between excludability with bounded capacity and approachability.

[2] The convex hull of an empty set is an empty set.



Part A Part B

FIGURE 1.

PROPOSITION 2.1. *A set with no sharp points* (F_1) *and a set in which all points are sharp* (F_2).

3. Proofs. The following estimate on Markov chains with finite state space will be useful. Chebyshev's inequality (see Shiryayev [13, p. 121]) and the decomposition theorem of irreducible chains (see Feller [3, p. 405]) imply that for every finite irreducible Markov chain[3] there is a constant $C > 0$ such that for every two states s, s', every $\lambda > 0$, and every $n \in \mathbb{N}$,

$$\mathbf{P}\left(\left| \frac{\mathbb{1}_{s_1=s} + \cdots + \mathbb{1}_{s_n=s}}{n} - \nu(s) \right| > \lambda \;\middle|\; s_0 = s' \right) < \frac{C}{n\lambda^2}, \tag{1}$$

where $\mathbb{1}_{s_t=s} = 1$ if the state at stage t is s and 0 otherwise, and $\nu = \big(\nu(s)\big)_s$ is the (unique) invariant distribution.

3.1. Excludability against bounded-recall strategies. Here we prove Theorem 2.1, which states that a closed set that does not contain a convex set that is approachable by Player 1 is excludable against bounded-recall strategies by Player 2.

The structure of the proof is as follows. We first prove in Proposition 3.1 that for every given bounded-recall strategy (m, σ) of Player 1 there is a strategy $\tau = \tau(m, \sigma)$ of Player 2 such that under $((m, \sigma), \tau)$ the long-run average payoff remains far from F with probability at least $1/2$. We then show in Lemma 3.2 that the collection $\{\tau(m, \sigma)\colon (m, \sigma) \in \mathcal{S}_{BR}\}$ can be taken to be countable. Finally, we construct a strategy τ^* that ensures that the long-run average payoff remains far from F. The strategy τ^* plays in blocks; each might be finite or infinite. A block ends once the average payoff along the block is close to F. If the average payoff is never close to F, the block never ends (i.e., it is an infinite block). In each block τ^* follows one of the countably many strategies in the collection. τ^* is designed in such a way that if it happens that there are infinitely many finite blocks, then each strategy in the collection is followed infinitely often (that is, in infinitely many blocks).

When Player 1 follows the bounded-recall strategy (m, σ), in each block where $\tau(m, \sigma)$ is played there is a probability of at least $1/2$ that the average payoff will never be close to F (so that the block is infinite). The probability that there will be infinitely many finite blocks is zero. Indeed, because $\tau(m, \sigma)$ is played in infinitely many blocks, and in each the probability that the block is finite is at most $1/2$, therefore, with probability one, one of the blocks is infinite, which means that the average payoff is far from F all along that block. This ensures that F is excludable.

We will need the following simple observation, which holds because under the Hausdorff metric, the collection of compact sets is compact, and the limit of approachable sets is approachable.

LEMMA 3.1. *If a closed set F does not contain a convex approachable set, then there is $\varepsilon > 0$ such that $B(F, \varepsilon)$ does not contain a convex approachable set.*

An element of $(I \times J)^t$ will be called a *memory*. When Player 1 plays a t-bounded-recall strategy, we denote by m_n his memory at stage n, that is, the t-history composed of the t pairs of actions played in stages $n-t, n-t+1, \ldots, n-1$. As Player 2 observes past play, she knows m_n.

The next proposition asserts that as long as the closed set F does not contain any convex approachable set, for every fixed bounded-recall strategy of Player 1 there is a reply of Player 2 that ensures the long-run average payoff remains far from F with nonnegligible probability.

[3] A finite Markov chain is *irreducible* if any state can be reached from any other state.

Proposition 3.1. *Let F be a closed set that does not contain any convex approachable set. There exists $\varepsilon > 0$ such that for every bounded-recall strategy σ of Player* 1 (*without an initial memory*) *there is a strategy τ of Player* 2 *and an integer N_* that satisfies*

$$\mathbf{P}_{(m,\sigma),\tau}\Big(\inf_{n\geq N_*} d(\bar{x}_n, F) \geq 2\varepsilon\Big) \geq \tfrac{1}{2}, \quad \forall\, m. \tag{2}$$

Observe that in Equation (2) the quantity ε is independent of the strategy of Player 1, while N_* does depend on it. When the length of the memory is fixed, the number of initial memories is finite, so that N_* need not depend on the initial memory.

Proof of Proposition 3.1. Let F be a closed set that does not contain any convex approachable set. By Lemma 3.1 there is $\varepsilon > 0$ such that $B(F, 6\varepsilon)$ does not contain any convex approachable set. By Lehrer and Solan ([8, Proposition 3]), $B(F, 6\varepsilon)$ is not approachable with bounded-recall strategies by Player 1.

We fix throughout a t-bounded-recall strategy σ of Player 1 (without the initial memory) and proceed in five steps.

Step 1. *Irreducible sets in the space of memories.* For every two memories m, m', we say that *m may lead to m'* if and only if Player 2 has an action that makes the game move in a single stage from memory m to memory m' with positive probability. Formally, this happens if and only if $m = (i_1, j_1, \ldots, i_t, j_t)$, $m' = (i_2, j_2, \ldots, i_t, j_t, i', j')$, and $\sigma(i' \mid m) > 0$.

A subset M of $(I \times J)^t$ is *irreducible* (w.r.t. the "may lead to" relation) if every memory $m \in M$ may lead only to memories in M, and for every $m, m' \in M$ there is a sequence of memories $m = \hat{m}_1, \hat{m}_2, \ldots, \hat{m}_L = m'$ such that $\hat{m}_l$ may lead to $\hat{m}_{l+1}$ for each $l = 1, 2, \ldots, L-1$. We denote by $\mathcal{M}$ the collection of all irreducible sets. When we say that $m \in \mathcal{M}$, we mean that $m \in M$, for some $M \in \mathcal{M}$.

The definition implies that if $m_n \in \mathcal{M}$, all subsequent memories belong to the same irreducible set.

Let $\hat{\tau}$ be the strategy of Player 2 in which she plays all actions with equal probability at every stage, irrespective of past play. By following $\hat{\tau}$, Player 2 ensures that the play moves in a finite number of stages to a memory that belongs to an irreducible set. Formally,

$$\mathbf{P}_{(m,\sigma),\hat{\tau}}(m_n \in \mathcal{M} \text{ for some } n \in \mathbb{N}) = 1, \quad \forall\, m \in (I \times J)^t. \tag{3}$$

Moreover, by following $\hat{\tau}$ Player 2 guarantees that the play reaches any memory in an irreducible set, provided it starts at that set. Formally,

$$\mathbf{P}_{(m,\sigma),\hat{\tau}}(m_n = m' \text{ for some } n \in \mathbb{N}) = 1, \quad \forall\, M \in \mathcal{M}, \quad \forall\, m, m' \in M. \tag{4}$$

Step 2. *Finding Good Play Paths.* An infinite sequence $h = (i_1, j_1, i_2, j_2, \ldots) \in (I \times J)^{\mathbb{N}}$ is called a *play path*.

Fix $m \in \mathcal{M}$. Because $B(F, 6\varepsilon)$ is not approachable by bounded-recall strategies, it is in particular not approachable by (m, σ). Therefore, there is a strategy τ_m of Player 2 such that on a set of play paths with positive probability the average payoff is infinitely often far from F by more than 6ε. Denoting

$$H_* = \{h\colon d(\bar{x}_n, F) > 6\varepsilon \text{ infinitely often}\}$$

we have

$$\mathbf{P}_{(m,\sigma),\tau_m}(H_*) > 0, \quad \forall\, m \in \mathcal{M}.$$

As the payoff space is bounded, for each play path in H_* there is a vector $c \notin B(F, 6\varepsilon)$ such that $d(\bar{x}_n, c) < \varepsilon$ infinitely often. In particular, for every initial memory $m \in \mathcal{M}$ there is a vector $c_m \notin B(F, 6\varepsilon)$ such that the set of the play paths along which the average payoff is infinitely often ε-close to c_m has a positive probability. That is,

$$\mathbf{P}_{(m,\sigma),\tau_m}\Big(\liminf_{n\to\infty} d(\bar{x}_n, c_m) < \varepsilon\Big) > 0, \quad \forall\, m \in \mathcal{M}. \tag{5}$$

Denote by H_m the set of play paths that satisfy $\liminf_{n\to\infty} d(\bar{x}_n, c_m) < \varepsilon$.

Step 3. *Empirical Distribution of Actions.* For every play path h, every memory $m \in (I \times J)^t$, every action $j \in J$, and every two positive integers k and l that satisfy $k < l$, denote[4]

$$\alpha^{m,j}_{k\to l}(h)(i) = \frac{\#\{k \leq n \leq l\colon m_n = m, i_n = i, j_n = j\}}{\#\{k \leq n \leq l\colon m_n = m, j_n = j\}}.$$

[4] Recall that m_n, i_n, and j_n depend on h.

$(\alpha^{m,j}_{k\to l}(h)(i))_{i\in I}$ is the empirical distribution of actions of Player 1 between stages k and l played after the memory m when Player 2 plays j. This quantity is defined whenever the denominator does not vanish.

Let $N^{m,j}_{k\to l}(h) = \#\{k \le n \le l\colon m_n = m, j_n = j\}$ be the number of times the memory m is followed by the action j of Player 2. The average payoff along h between stages k and l is given by

$$Z = \frac{1}{l-k+1}\sum_{n=k}^{l} x_n = \sum_{m,j}\frac{N^{m,j}_{k\to l}(h)}{l-k+1} + \sum_{i}\alpha^{m,j}_{k\to l}(h)(i)u(i,j). \tag{6}$$

If along h the action j appears after the memory m only finitely many times, we can choose k sufficiently large such that along h the memory m is never followed by the action j after stage k. Therefore, for every play path h there is a stage k such that

$$\lim_{l\to\infty} N^{m,j}_{k\to l}(h) = \infty \quad \text{or} \quad \lim_{l\to\infty} N^{m,j}_{k\to l}(h) = 0. \qquad \forall\, m, \forall\, j \in J. \tag{7}$$

Set

$$H_1 = \left\{h\colon \lim_{l\to\infty}\alpha^{m,j}_{1\to l}(h) = \sigma(m), \forall\,(m,j) \text{ s.t. } \lim_{l\to\infty} N^{m,j}_{1\to l}(h) = \infty\right\}.$$

Because the choices of the two players are conditionally independent,

$$\mathbf{P}_{(m,\sigma),\tau}(H_1) = 1, \quad \forall\,\tau, \forall\, m. \tag{8}$$

Step 4. *Periodic Partial Histories.* Fix $m_* \in (I\times J)^t$, and fix $h \in H_{m_*} \cap H_1$. By (5) and (8) such a play path exists. Because $h \in H_{m^*}$, infinitely often along h the payoff is in $d(c_{m_*}, \varepsilon)$. Let $\delta > 0$ be arbitrary. Because the number of memories is finite, and by (7), there are positive integers k and l that satisfy $k < l$ such that the following conditions are simultaneously satisfied:

(A1) For every $(m,j) \in (I\times J)^t \times J$, either $\lim_{l\to\infty} N^{m,j}_{k\to l}(h) = 0$ or $\lim_{l\to\infty} N^{m,j}_{k\to l}(h) = \infty$.
(A2) $\|\alpha^{m,j}_{1\to l}(h) - \sigma(m)\| < \delta$ for every (m,j) with $\lim_{l\to\infty} N^{m,j}_{1\to l} = \infty$.
(A3) The average payoff until stage l is in $B(c_{m_*}, \varepsilon)$.
(A4) $m_k = m_l$.
(A5) $k \le \frac{1}{2}\varepsilon l$.

By (A3) and (A5) it follows that the average payoff along h between stages k and l is in $B(c_m, 2\varepsilon)$.

The set of memories that occur along h between stages k and l is

$$M_{k\to l}(h) = \{m \in (I\times J)^t\colon \exists k \le n \le l \text{ s.t. } m_n = m\}.$$

The partial history h between stages k and l induces a Markov chain over the space $M_{k\to l}(h)$: if the current memory is $m = (i_1, j_1, \ldots, i_t, j_t)$, the new memory is $m' = (i_2, j_2, \ldots, i_t, j_t, i', j')$ with probability

$$\frac{N^{m,j'}_{k\to l}(h)}{l-k+1} \times \alpha^{m,j'}_{k\to l}(h)(i').$$

This is the empirical frequency of the pair (i', j') after the memory m along h between stages k and l. (A4) implies that this Markov chain is irreducible, and its invariant distribution ν_α is actually given by

$$\nu_\alpha(m) = \frac{\#\{k \le n < l\colon m_n = m\}}{l-k+1}, \quad \forall\, m \in M_{k\to l}(h). \tag{9}$$

A second Markov chain over the space $M_{k\to l}(h)$ is the following one. If the current state is $m = (i_1, j_1, \ldots, i_t, j_t)$, the new memory is $m' = (i_2, j_2, \ldots, i_t, j_t, i', j')$ with probability

$$\frac{N^{m,j'}_{k\to l}(h)}{l-k+1} \times \sigma(m)(i').$$

Provided δ is sufficiently small, (A2) implies that this Markov chain is irreducible. Denote its invariant distribution by ν_β. Observe that this is the Markov chain that is induced over $M_{k\to l}(h)$ when Player 1 follows σ and Player 2 follows the t-bounded-recall strategy $\hat\tau$ that is defined by

$$\hat\tau(m)(j) = \frac{N^{m,j}_{k\to l}(h)}{l-k+1}, \quad \forall\, j \in J. \tag{10}$$

If δ is sufficiently small, by virtue of (A1), (A2), and O'Cinneide ([10, Theorem 1]) we have

$$\|\nu_\alpha - \nu_\beta\| \leq \frac{\varepsilon}{|(I \times J)^t|}. \tag{11}$$

Provided the initial memory is in $M_{k\to l}(h)$, the long-run average payoff under $(\sigma, \hat{\tau})$ is

$$Y = \sum_{m,i,j} \nu_\beta(m)\hat{\tau}(m(j))\sigma(m)(i)u(i,j). \tag{12}$$

By (A2), (A5), (11), and (9), the difference between Y and Z (see Equation (6)) is at most $\varepsilon + \delta$. Because $Z \in B(c_{m^*}, \varepsilon)$, and because δ is sufficiently small, it follows that $Y \in B(c_{m^*}, 3\varepsilon)$, as desired.

Step 5. *Constructing the Desired Strategy.* We here construct a strategy τ that ensures the average payoff remains far from F when Player 1 follows σ.

- Phase 1: τ follows $\hat{\tau}$ until a memory in $\mathcal{M}$ is reached. Denote by $m_* \in \mathcal{M}$ that memory. Fix $h \in H_{m_*} \cap H_1$, let $\delta > 0$ be sufficiently small, and let k and l be two positive integers such that (11) and (A1)–(A5) hold.
- Phase 2: τ keeps following $\hat{\tau}$ until a memory in $M_{k\to l}(h)$ is reached.
- Phase 3: From that stage on, τ follows $\hat{\tau}$.

Let us verify that (2) holds for τ. Indeed, by (3) there is $N_*^1 \in \mathbb{N}$ such that with probability greater than $5/6$ Phase 1 ends before stage N_*^1. By (4) there is $N_*^2 > N_*^1$ such that with probability greater than $4/6$ Phase 2 ends before stage N_2^*.

Equation (1) implies that there is $N_* > N_*^2$ such that with probability at least $\frac{1}{2}$, the average payoff up to stage n is within $d(c_{m_*}, 4\varepsilon)$ for every $n \geq N_*$. Because $c_{m_*} \notin B(F, 6\varepsilon)$, the desired result follows. □

Remark 3.1. Note that the strategy τ defined in the proof of Proposition 3.1 is asymptotically t-bounded recall. That is, with probability 1 there is a stage such that after that stage, Player 2 follows a t-bounded-recall strategy.

Our next step is to prove that there is a countable collection of strategies of Player 2 such that for every bounded-recall strategy of Player 1, one of the strategies in this collection ensures that the long-run average payoff will be far from F.

Lemma 3.2. *There is $\varepsilon' > 0$ and a sequence $(\tau_l, N_l)_{l\in\mathbb{N}}$, where for every $l \in \mathbb{N}$, τ_l is a strategy of Player 2 and $N_l \in \mathbb{N}$, such that for every bounded-recall strategy σ (without an initial memory) there is an index $l \in \mathbb{N}$ such that Equation (2) is satisfied for ε', σ, τ_l, and N_l.*

Proof. Let ε be given by Proposition 3.1. Fix $t \in \mathbb{N}$ and a t-bounded-recall strategy σ (without an initial memory). Let τ be the strategy we constructed in the proof of Proposition 3.1 for σ, and let N_* be the corresponding finite horizon that appears in (2). Let $D(\sigma)$ be the set of t-bounded recall strategies σ' such that the support of $\sigma(\cdot \mid m)$ and $\sigma'(\cdot \mid m)$ coincide for every $m \in (I \times J)^t$. For every $\delta > 0$ define $B_\delta(\sigma)$ to be the set of all the t-bounded recall strategies $\sigma' \in D$ that satisfy $|1 - (\sigma(i \mid m)/\sigma'(i \mid m))| < \delta$ for every $i \in I$ and $m \in (I \times J)^t$.[5]

We claim that if δ is small enough, then for each $\sigma' \in B_\delta(\sigma)$, Equation (2) is satisfied with $\varepsilon' = \varepsilon/2, \sigma', \tau$, and N_*. Indeed, because $\sigma' \in D(\sigma)$ the partitions into irreducible sets induced by σ and by σ' coincide. If δ is sufficiently small, then

(C1) the distribution over histories of length N_* under $((m, \sigma), \tau)$ and the corresponding distribution under $((m, \sigma'), \tau)$ differ (in the L_1-norm) by at most $\varepsilon/2$, and

(C2) by O'Cinneide ([10, Theorem 1]), for each irreducible set M the L_1-distance between the invariant distribution ν_β that was defined in the proof of Proposition 3.1 for σ and that defined for σ' (for the same history) is at most $\varepsilon/2$.

It follows that under $((m, \sigma'), \tau)$, with probability at least $\frac{1}{2}$ the average payoff remains out of $B(F, \varepsilon)$ from stage N_* and on. We let $\delta = \delta(\sigma) > 0$ sufficiently small that both (C1) and (C2) hold.

For any two t-bounded-recall strategies σ and σ', $D(\sigma)$ and $D(\sigma')$ are either identical or disjoint. Thus, $\{D(\sigma)\}$ is a finite partition of the set of t-bounded-recall strategies.

Moreover, for every σ the set $D(\sigma)$ is σ-compact. The sets $(B_{\delta(\sigma)}(\sigma'))_{\sigma' \in D(\sigma)}$ are open in $D(\sigma)$ and cover it. Thus, there is a countable set of strategies $(\sigma_l)_{l\in\mathbb{N}} \subset D$ such that $\bigcup_{l\in\mathbb{N}} B_{\delta(\sigma_l)}(\sigma_l) = D(\sigma)$.

We conclude that a countable union of sets of the kind $B_\delta(\sigma)$ covers all the t-bounded-recall strategies, and therefore there is a countable collection of these sets whose union covers the set of *all* bounded-recall strategies. The result follows. □

[5] By convention $0/0 = 1$.

We are now ready to prove Theorem 2.1. We will use Lehrer and Solan [8, Theorem 1], which states that a closed set is approachable by bounded-recall strategies if and only if it contains a convex approachable set.

PROOF OF THEOREM 2.1. Let F be a closed set that does not contain the convex hull of any approachable set. Let ε' and $(\tau_l, N_l)_{l\in\mathbb{N}}$ be given by Lemma 3.2. Then, for every bounded-recall strategy (m,σ) of Player 1 there is τ_l that excludes (m,σ) from $B(F,\varepsilon')$ with probability greater than $\frac{1}{2}$.

We now construct the strategy τ^* of Player 2 that was described above, and which excludes F against *every* bounded-recall strategy with probability 1.

Let $r=(r_1,r_2)\colon \mathbb{N}\to\mathbb{N}^2$ be a 1-1 and onto function. τ^* plays in blocks of varying (possibly infinite) size. In block k, τ^* follows $\tau^*_{r_1(k)}$. Block k terminates once the following two conditions are simultaneously satisfied: (i) the length of the block is at least $N_{r_1(k)}$, and (ii) the average payoff within this block is in $B(F,\varepsilon')$. Once block k terminates, block $k+1$ begins. The second coordinate r_2 ensures that for each l there are infinitely many ks that satisfy $r_1(k)=l$.

We now verify that τ^* excludes F, provided Player 1 uses a bounded-recall strategy, say (m,σ). Suppose that $\sigma\in B_{\delta(\sigma_l)}(\sigma_l)$. We argue that with probability one there are finitely many blocks. Indeed, consider the event $\mathscr{E}_1$ that all blocks are finite. On this event there are infinitely many blocks in which τ_l is followed. By (2) the probability that such a block is finite is at most $\frac{1}{2}$. Because play in different blocks is conditionally independent, $P(\mathscr{E}_1)=0$, as claimed.

This implies that

$$\mathbf{P}_{(m,\sigma),\tau^*}\left(\liminf_{n\to\infty} d(\bar{x}_n,F)<\varepsilon'\right)=0.$$

It follows that F is excludable against bounded-recall strategies, as desired. □

3.2. Excludability against automata. Here we prove Theorem 2.2, which states that any closed set that does not contain the convex hull of its sharp points is excludable against automata.

PROOF OF THEOREM 2.2. Let F be a closed set that does not contain the convex hull of its sharp points. Thus, there are sharp points $x_1,\ldots,x_L$ in F and nonnegative numbers $\lambda_1,\ldots,\lambda_L$ such that $\sum_{l=1}^L\lambda_l=1$ and $z:=\sum_{l=1}^L\lambda_l x_l\notin F$. Choose $\delta>0$ such that $d(z,F)>2\delta$, so that $d(\sum_{l=1}^L\lambda_l B(x_l,\delta),F)>\delta$.[6]

Because $(x_l)_{l=1}^L$ are sharp in F, there are mixed actions $q_1,\ldots,q_L\in\Delta(J)$ such that $H(q_l)\cap F\subseteq B(x_l,\delta/2)$, for $l=1,\ldots,L$.

Step 1. *A Definition of a Strategy.* We now define a strategy τ^* that plays in blocks with increasing length. Below we prove that this strategy guarantees that the average payoff remains far from F.

- Before start of play, choose a number β uniformly from the interval $[(3/4)\delta,\delta]$.
- There are L types of blocks, one for each $l=1,\ldots,L$. The length of block k is $k+k^4$, and its type is denoted by l_k. A block of type l is referred to as an l-block.
- For every l, at every stage of an l-block τ^* plays the mixed action q_l.
- The type of the first block is $l=1$.
- The type of block $k+1$ is determined as follows.

— If the average payoff within block k is not in $B(x_{l_k},\beta)$, the type of block $k+1$ is the same as the type of block k.

— Otherwise, denote by n the last stage of block k, and by m_l the overall number of stages up to stage n spent in l-blocks (in particular, $\sum_{l=1}^L m_l=n$). Block $k+1$ is a j-block if j is the minimal index in $\arg\max_{l=1,\ldots,L}(\lambda_l-m_l/n)$.

When Player 1 follows an automaton and Player 2 follows the stationary strategy q_l the long-run average payoff exists. Provided a block is sufficiently long, if the limit average payoff is in $B(x_l,\beta)$, then with high probability the average payoff along an l-block is in $B(x_l,\beta)$, while if the limit average payoff is outside $B(x_l,\beta)$, then with high probability the average payoff along the l-block is outside $B(x_l,\beta)$. When the average payoff is on the boundary of $B(x_l,\beta)$, the average payoff along the l-block may be in or outside $B(x_l,\beta)$. This case creates a technical difficulty, which is avoided by choosing β at random, thereby ensuring that the probability that the average payoff lies on the boundary of $B(x_l,\beta)$ is zero.

From now on we fix an automaton A that is used by Player 1, and $\eta>0$. Our goal is to prove that there is $N_{A,\eta}\in\mathbb{N}$ such that

$$\mathbf{P}_{\sigma,\tau^*}\left(\inf_{n\geq N_{A,\eta}} d(\bar{x}_n,F)<\frac{\delta}{2}\right)<\eta.$$

[6] For every two sets $A,B\in\mathbb{R}^d$ and every $\lambda_1,\lambda_2>0$, $\lambda_1A+\lambda_2B:=\{\lambda_1x+\lambda_2y\colon x\in A, y\in B\}$, and $d(A,B)=\inf\{d(x,y)\colon x\in A, y\in B\}$.

Step 2. *An auxiliary Markov Chain.* Denote by S the set of states of the automaton, and by s_k the state of the automaton at stage k. Any mixed action q of Player 2 induces a Markov chain $\mathcal{M}(q)$ over S, which reflects the evolution of the automaton when Player 2 plays at every stage (regardless of past play) the mixed action q. We denote by $\mathbf{P}_{(A,s),q}$ the probability over play paths induced by this Markov chain, when $s \in S$ is the initial state of the automaton.

We will consider the Markov chain $\mathcal{M}(q_l)$ for every $l = 1, \ldots, L$. The type of block $k+1$ is determined according to whether the distance between the average payoff in block k and x_{l_k} is higher or lower than β. Provided the block is sufficiently long, the average payoff along the block is close to the long-run average payoff in the Markov chain $\mathcal{M}(q_l)$.

Step 3. *Estimating Empirical Distribution Along a Block.* By Seneta ([12, Theorem 4.7]) there are constants $C_1 > 0$ and $\rho \in (0,1)$ (which depend on the automaton) such that for every $l = 1, \ldots, L$, the probability that by stage k no irreducible subset[7] of $\mathcal{M}(q_l)$ is reached is at most $C_1 \rho^k$.

$$\mathbf{P}_{(A,s),q_l}\big(s_k \text{ is in some irreducible subset of } \mathcal{M}(q_l)\big) \geq 1 - C_1\rho^k, \quad \forall\, s \in S, k \in \mathbb{N}. \tag{13}$$

Note that the family of irreducible sets depends on q_l.

Denote by $\mathscr{E}_1(k')$ the event that an irreducible subset (in the respective Markov chain) is reached by stage k of block k, for every $k \geq k'$. By (13), and because $\sum_{k\in\mathbb{N}} C_1\rho^k < \infty$, there is $K_1 \in \mathbb{N}$ such that $\mathbf{P}_{(A,s),q_l}(\mathscr{E}_1(K_1)) \geq 1 - \eta/4$. On $\mathscr{E}_1(K_1)$, in each block $k \geq K_1$ the process spends at least k^4 stages in an irreducible subset.

By Equation (1) (with $\lambda = 1/\sqrt{k}$), there is a constant $C_2 > 0$ such that for every $l = 1, \ldots, L$, and every pair of states s, s' in the same irreducible set E of $\mathcal{M}(q_l)$,

$$\mathbf{P}_{(A,s'),q_l}\left(\left|\frac{\mathbb{1}_{s_1=s} + \cdots + \mathbb{1}_{s_n=s}}{n} - \nu_l^E(s)\right| > \frac{1}{\sqrt{k}}\right) < \frac{C_2}{k^3}, \quad \forall\, k \in \mathbb{N}, \quad \forall\, n \geq k^4, \tag{14}$$

where $\nu_l^E(s)$ is the invariant distribution of $\mathcal{M}(q_l)$ within E.

Step 4. *Estimating Average Payoff Along a Block.* For every irreducible set E of $\mathcal{M}(q_l)$, denote by y_l^E the long-run average payoff when the initial state of the automaton A is in E, and Player 2 plays the mixed action q_l at every stage. It is given by $y_l^E = \sum_{s\in E} \nu_l^E(s) u_{p_s, q_l} \in H(q_l)$, where p_s is the mixed action played by the automaton at state s.

Denote by E_k the irreducible set to which the Markov chain $\mathcal{M}(q_{l_k})$ is absorbed. On $\mathscr{E}_1(K_1)$ the set E_k is defined for every $k \geq K_1$.

By Equation (14), and because payoffs are bounded by one, with high probability the average payoff in the last k^4 stages of each block is within $1/\sqrt{k}$ of $y_{l_k}^{E_k}$. In this case, the average payoff along the block is within

$$\frac{k}{k+k^4} + \frac{k^4}{k+k^4} \times \frac{1}{\sqrt{k}} \leq \frac{2}{\sqrt{k}}$$

of $y_{l_k}^{E_k}$.

Denote by X_k the average payoff in block k. For every $k' \geq K_1$ let

$$\mathscr{E}_2(k') = \left\{ \big|X_k - y_{l_k}^{E_k}\big| \leq \frac{2}{\sqrt{k}}, \quad \forall\, k \geq k' \right\} \cap \mathscr{E}_1(K_1). \tag{15}$$

By (14) there is $K_2 \geq K_1$ such that $\mathbf{P}_{A,\tau^*}(\mathscr{E}_2(K_2)) \geq 1 - \eta/2$. From now on we restrict ourselves to the event $\mathscr{E}_2(K_2)$.

Let

$$\zeta = \min\big\{\big|\beta - d(x_l, y_l^E)\big| \colon l = 1, \ldots, L, E \text{ is irreducible set of } \mathcal{M}(q_l)\big\}.$$

Because there are finitely many irreducible sets, and because β was chosen uniformly in the range $[\frac{3}{4}\delta, \delta]$, one has, with probability one, $\zeta > 0$. Set $K_0 := \max\{K_2, 4/\zeta^2, 4/(\beta-\delta)^2\}$.

In both Steps 5 and 6 we will use the fact that the ratio between the length of block k and the total length of the first $k-1$ blocks goes to zero:

$$\frac{k^4 + k}{\sum_{j<k}(j^4 + j)} < \frac{10k^4}{k^5} = \frac{10}{k}. \tag{16}$$

[7] A subset of states E is *irreducible* if any state in E can be reached from any other state in E, and no state outside E can be reached from any state in E.

Step 5. *The Case $d(x_{l_k}, y_{l_k}^{E_k}) > \beta$ for Some $k \geq K_0$.* Suppose there is $k \geq K_0$ such that $d(x_{l_k}, y_{l_k}^{E_k}) > \beta$. By the definition of ζ one has $d(x_{l_k}, y_{l_k}^{E_{l_k}}) > \beta + \zeta$. Because $k \geq 4/\zeta^2$, and because we restrict ourselves to the event $\mathscr{E}_2(K_2)$, we have $d(x_{l_k}, X_k) > \beta$. By the definition of τ^*, the type of block $k+1$ is $l_{k+1} = l_k$. In particular, $q_{l_{k+1}} = q_{l_k}$: The stationary strategy that was played during block k is played during block $k+1$ as well. Because in block k the state of the automaton has reached the irreducible set E_k of $\mathscr{M}(q_{l_k})$, it will remain in E_k all through block $k+1$, so that $y_{l_{k+1}}^{E_{k+1}} = y_{l_k}^{E_k}$.

Applying the same argument to block $k+1$, we conclude that on $\mathscr{E}_2(K_2)$ the average payoff in each block $k' > k$ is in $B(y_{l_{k'}}^{E_{k'}}, 2/\sqrt{k'})$.

It follows that on $\mathscr{E}_2(K_2)$ the long-run average payoff converges to $y_{l_k}^{E_k}$:

$$\mathbf{P}_{(m,\sigma),\tau^*}\left(\lim_{n\to\infty} d(\bar{x}_n, y_{l_k}^{E_k}) = 0 \,\middle|\, \mathscr{E}_2(K_2), d(x_{l_k}, y_{l_k}^{E_k}) > \beta\right) = 1.$$

Recall that $y_{l_k}^{E_k} \in H(q_{l_k})$. Because $H(q_{l_k}) \cap F \subseteq B(x_{l_k}, \delta/2)$, but $d(y_{l_k}^{E_k}, x_{l_k}) > \beta \geq (3/4)\delta$, we deduce that $d(y_{l_k}^{E_k}, F) \geq \delta/4$. It follows that

$$\mathbf{P}_{(m,\sigma),\tau^*}\left(d\left(\lim_{n\to\infty} \bar{x}_n, F\right) \geq \frac{\delta}{4} \,\middle|\, \mathscr{E}_2(K_2), d(x_{l_k}, y_{l_k}^{E_k}) > \beta\right) = 1. \tag{17}$$

Step 6. *The case $d(x_{l_k}, y_{l_k}^{E_k}) < \beta$ for all $k \geq K_0$.* By the definition of ζ, $d(x_{l_k}, y_{l_k}^{E_k}) < \beta - \zeta$ for every $k \geq K_0$. Because $K_0 \geq 4/\zeta^2$, on $\mathscr{E}_2(K_2)$ we have $d(x_{l_k}, X_k) < \beta$ for every $k \geq K_0$. The definition of τ^* implies that the type of block $k+1$ is chosen so as to maximize the difference between λ_l and the fraction of stages spent so far in l-blocks.

We will now show that the fraction of stages spent in l-blocks converges to λ_l. Denote by $(m_{l,n})_{n\in\mathbb{N}}$ the number of stages spent in l-blocks up to stage n.

Fix $l \in \{1, \ldots, L\}$, and consider the function $n \mapsto m_{l,n}/n$. If the type of the current block is not l, along the current block this function decreases. If the type of the current block is l, then at the beginning of the current block $\lambda_l - m_{l,n}/n$ was nonnegative, so that $m_{l,n}/n \leq \lambda_l$. In particular, at the end of the current block $m_{l,n}/n \leq \lambda_l + (k^4 + k)/n$, where k is the length of the block. We conclude that for every block k and every stage n along the block

$$\frac{m_{l,n}}{n} \leq \lambda_l + \frac{k^4 + k}{n}.$$

Because $\sum_l \lambda_l = 1 = \sum_l m_{l,n}/n$,

$$\frac{m_{l,n}}{n} \geq \lambda_l - L\frac{k^4 + k}{n}.$$

By (16) the quantity $(k^4 + k)/n$ goes to zero as n goes to infinity, so that indeed $m_{l,n}/n$ converges to λ_l. This implies that the average payoff is asymptotically in $B(z, \beta) \subseteq B(z, \delta)$. By the choice of δ, this set is disjoint of $B(F, \delta)$:

$$\mathbf{P}_{(m,\sigma),\tau^*}\left(\liminf_{n\to\infty} d(\bar{x}_n, F) \geq \delta \,\middle|\, \mathscr{E}_2(K_2), d(x_{l_k}, y_{l_k}^{E_k}) < \beta \quad \forall\, k \geq K_2\right) = 1. \tag{18}$$

Step 7. *Conclusion.* Because β is chosen uniformly in the interval $[(3/4)\delta, \delta]$, the probability that for some $1 \leq l \leq L$ and some irreducible set E in $\mathscr{M}(q_l)$ the limit payoff y_l^E lies on the boundary of $B(x_l, \beta)$ is zero.

Therefore, Equations (17) and (18) imply that

$$\mathbf{P}_{(m,\sigma),\tau^*}\left(\liminf_{n\to\infty} d(\bar{x}_n, F) \geq \frac{\delta}{4} \,\middle|\, \mathscr{E}_2(K_2)\right) = 1. \tag{19}$$

The desired result follows.

Acknowledgments. The research of the first author was supported by the Israel Science Foundation (Grant 762/04). The research of the second author was supported by the Israel Science Foundation (Grant 69/01). The authors thank two anonymous referees and the associate editor for their comments, which substantially improved the paper.

References

[1] Aumann, R. J., S. Sorin. 1989. Cooperation and bounded recall. *Games Econom. Behavior* **1** 5–39.
[2] Blackwell, D. 1956. An analog of the minimax theorem for vector payoffs. *Pacific J. Math.* **6** 1–8.
[3] Feller, W. 1968. *An Introduction to Probability Theory and Its Applications*, 3rd ed. John Wiley and Sons, New York.
[4] Hou, T. F. 1969. Weak approachability in a two-person game. *Ann. Math. Statist.* **40** 789–813.
[5] Hou, T. F. 1971. Approachability in a two-person game. *Ann. Math. Statist.* **42** 735–744.
[6] Kalai, E. 1990. Bounded rationality and strategic complexity in repeated games. T. Ichiishi, A. Newman, Y. Tauman, eds. *Game Theory and Applications.* Academic Press, New York, 131–157.
[7] Lehrer, E. 1988. Repeated games with stationary bounded recall strategies. *J. Econom. Theory* **46** 130–144.
[8] Lehrer, E., E. Solan. 2003. No-regret with bounded computational capacity. CMS-EMS Discussion Paper 1373, Kellogg School of Management, Northwestern University, Evanston, IL.
[9] Neyman, A. 1985. Bounded complexity justifies cooperation in finitely repeated prisoner's dilemma. *Econom. Lett.* **9** 227–229.
[10] O'Cinneide, C. A. 1993. Entrywise perturbation theory and error analysis for Markov chains. *Numerische Mathematik* **65** 109–120.
[11] Rubinstein, A. 1986. Finite automata play the repeated prisoner's dilemma. *J. Econom. Theory* **39** 83–96.
[12] Seneta, E. 1981. *Non-Negative Matrices and Markov Chains*. Springer-Verlag, New York.
[13] Shiryayev, A. N. 1984. *Probability*. Springer-Verlag, New York.
[14] Spinat, X. 2002. A necessary and sufficient condition for approachability. *Math. Oper. Res.* **27** 31–44.
[15] Vieille, N. 1992. Weak approachability. *Math. Oper. Res.* **17** 781–791.