האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

CORRELATION THROUGH BOUNDED RECALL STRATEGIES

By

RON PERETZ

Discussion Paper # 579 July 2011

המרכז לחקר הרציונליות

THE CENTER FOR THE STUDY
OF RATIONALITY

Feldman Building, Givat-Ram, 91904 Jerusalem, Israel
PHONE: [972]-2-6584135 FAX: [972]-2-6513681
E-MAIL: ratio@math.huji.ac.il
URL: http://www.ratio.huji.ac.il/

Correlation through Bounded Recall Strategies

Ron Peretz*

June 21, 2011

Abstract

Two agents independently choose mixed m-recall strategies that take actions in finite action spaces A_1 and A_2. The strategies induce a random play, $a_1, a_2, \ldots$, where a_t assumes values in $A_1 \times A_2$. An M-recall observer observes the play. The goal of the agents is to make the observer believe that the play is similar to a sequence of i.i.d. random actions whose distribution is $Q \in \Delta(A_1 \times A_2)$. For nearly every t, the following event should occur with probability close to one: "the distribution of a_{t+M} given $a_t, \ldots, a_{t+M-1}$ is close to Q." We provide a sufficient and necessary condition on m, M, and Q under which this goal can be achieved (for large m).

This work is a step in the direction of establishing a folk theorem for repeated games with bounded recall. It tries to tackle the difficulty in computing the individually rational levels (IRL) in the bounded recall setting. Our result implies, for example, that in some games the IRL in the bounded recall game is bounded away below the IRL in the stage game, even when all the players have the same recall capacity.

1 Introduction

The motivation behind this work comes from the study of the set of equilibrium payoffs in repeated games played through strategies of bounded recall (bounded-recall games). It turns out that the main problem is to estimate the individually rational level (IRL) of the bounded-recall games. In repeated games of perfect recall, the IRL is equal to that of the one-step game and therefore the folk theorem allows us to express the set of equilibrium payoffs in terms of the IRL of the one-step game. In bounded-recall games, the IRL may differ from that of the one-step game depending on the recall capacities of the players. So, if we are to find a "folk theorem" for bounded-recall games it should be expressed in terms of the IRL of the bounded-recall games.

*School of mathematical sciences, Tel-Aviv University.
This work is based on the authors Ph.D. thesis done under the supervision of Professor Abraham Neyman in the Center for the Study of Rationality, Hebrew University of Jerusalem. It is supported in part by grant #212/09 of the Israel Science Foundation and by the Google Inter-university center for Electronic Markets and Auctions.
Email: ronprtz@gmail.com.

In the special case of two-person games, the IRL is the value of a two-person zero-sum game. Two pioneering works on the value of two-person zero-sum bounded-complexity games are by Ben-Porath [BP93] (bounded-memory) and Lehrer [Leh88] (bounded-recall). Recently there has been some progress in the study of the value of two-person zero-sum bounded-complexity games due to the introduction of information theoretic methods by Neyman and Okada [NO00].[1] Among the works that employ these methods are [NO09, Ney08, NS10, GT07] and the present work. In [Per08] we characterize the asymptotic value of zero-sum games, and hence[2], also, the set of equilibrium payoffs in non-zero-sum two-person games.

In three-person games, the picture is less clear. Consider for example the case in which all three players have the same recall capacity. Lehrer's result on two-person games implies that each player can simulate a mixed action such that the punished player cannot correlate with it. By each player simulating a mixed action independently of the others, the IRL in the bounded-recall game is shown to be (asymptotically) no greater than the IRL in the one-step game. Bavly and Neyman [BN03] show that in the presence of a deterministic public signal (that depends on an exponentially long recall) the team can simulate a correlated punishment (they term it "concealed correlation"); thus the IRL in the bounded-recall game with public signaling converges, as the recall capacity grows, to the value that the punished player can secure against the (correlated) team in the one-step game.

Is it possible to obtain concealed correlation without a public signal? This question motivates the present work. We analyze the class of random processes that can be implemented through bounded-recall strategies. We refer, in particular, to cyclic approximations of i.i.d. processes. For this class of processes, we obtain an asymptotic characterization of what can be implemented through bounded-recall strategies. It should be noted that a tight estimation of the IRL, and hence an asymptotic characterization of the set of equilibrium payoffs, remains an open problem.

Our possibility result, Theorem 2.1, is proved by constructing strategies for the punishing team. This construction relies on an explicit version of the construction needed in the zero-sum case.[3] The present construction relies on a cycle of actions in which a large majority of the actions, in a fixed set of positions, are chosen uniformly at random, whereas the rest of the actions are chosen such that the entire cycle satisfies some local constraints. The local constraints are of the form that can be dealt with by Lovász's local lemma [AS00, p. 65] obtaining an implicit construction. The fact that most of the actions are chosen uniformly at random guarantees that the entire cycle preserves properties of a true random cycle. This feature is hard to achieve through an implicit construction.

Our impossibility result, Theorem 2.3, is obtained by proving information constraints on the limiting average of the play induced by mixed m-recall strategies. The stationarity of the limiting average play makes it suitable for information theoretic analysis.

In Section 4, we relate the analysis of implementable processes to the computation of the IRL in bounded-recall games. We establish a few results and present several problems for future research.

[1] Methods of information theory appear also in [Leh88]

[2] See Theorem 4 in [Leh88]

[3] See Proposition 4.9 in [Per08].

2 Results

Let $A = A_1 \times A_2$ be a finite alphabet. A *pure m-recall strategy*[4] for agent i ($i = 1, 2$) is a function $\sigma : \bigcup_{n=0}^{\infty} A^n \to A_i$, satisfying $\forall n > m$ $\forall a \in A^n$

$$\sigma(a) = \sigma(a_{n-m+1}, \dots, a_n).$$

The set of pure m-recall strategies for agent i is denoted $\Sigma^i(m)$. A pair of pure (m-recall) strategies σ, τ induces a play $(a_t)_{t\in\mathbb{N}}$. The *induced play* is defined recursively by

- $a_1 = (\sigma(\emptyset), \tau(\emptyset))$,
- $a_{n+1} = (\sigma(a_1, \dots, a_n), \tau(a_1, \dots, a_n))$.

A mixed m-recall strategy for agent i is a probability measure on (the discrete space) $\Sigma^i(m)$. A pair of mixed (m-recall) strategies induces a measure on the set of plays.

The *entropy* of a discrete random variable X is defined by

$$H(X) = -\sum_{\mathbf{x}} \Pr(X = \mathbf{x}) \log(\Pr(X = \mathbf{x})).$$

The entropy is a function of the distribution of X; therefore if Q is a probability measure over a finite set, we can define

$$H(Q) = -\sum_{\mathbf{x}} Q(\mathbf{x}) \log(Q(\mathbf{x})).$$

We denote by Q^n the n-fold product of Q. The entropy of Q^n is exactly $nH(Q)$. The distribution of a sequence of random variables $x_1, \dots, x_n$ is called an *ϵ-approximation of Q^n*, if

(i) $\frac{1}{n} H(x_1, \dots, x_n) \geq H(Q) - \epsilon$, and

(ii) $\sum_{\mathbf{x}} \left| \frac{1}{n} \sum_{t=1}^{n} \Pr(X_t = \mathbf{x}) - Q(\mathbf{x}) \right| \leq \epsilon$.

To illustrate the meaning of an ϵ-approximation it can be shown that for every $\epsilon > 0$ there exists $\delta > 0$ such that for every n, if the distribution of $x_1, \dots, x_n$ is a δ-approximation of Q^n then

$$\frac{1}{n} \sum_{t=1}^{n} \mathop{\mathbf{E}}_{x_1,\dots,x_{t-1}} \left[\sum_{\mathbf{x}} |\Pr(X_t = \mathbf{x} | x_1, \dots, x_{t-1}) - Q(\mathbf{x})| \right] < \epsilon.$$

The left-hand side of the above inequality is the expected L_1-distance between Q and the distribution of x_t given $x_1, \dots, x_{t-1}$; where t is chosen at random from $\{1, \dots, n\}$. Moreover, by Markov's equality, the following event occurs in probability greater than $1 - \sqrt{\epsilon}$: the L_1-distance between Q and the distribution of x_t given $x_1, \dots, x_{t-1}$ is less than $\sqrt{\epsilon}$.

The present paper characterizes the asymptotic values of m and M, for which a team of two m-recall agents can produce an ϵ-approximation of Q^M. The significant parameters turn out to be the ratio between M and m and the entropy of Q, Q_1 and Q_2, where Q_i is the marginal of Q on the ith coordinate.

Our first theorem is a possibility result. For a finite or countable set A we denote the set of probability measures on A by $\Delta(A)$.

[4]We use the word "agent" and not "player" since no game is being played. There are no incentives. We only study the class of random sequences that can be implemented through mixed m-recall strategies. This has implications to the study of the equilibria, as shown in Section 4.

Theorem 2.1. *Let A_1 and A_2 be finite alphabets. Let $Q \in \Delta(A_1 \times A_2) \setminus (\Delta(A_1) \times \Delta(A_2))$. For every $\epsilon > 0$ there exists $m_0 \in \mathbb{N}$ such that for every $M, m \geq m_0$, if*

$$\frac{m}{M} \geq 1 - \frac{H(Q)}{H(Q_1) + H(Q_2)} \tag{2.1}$$

then there exist mixed m-recall strategies for agents one and two that take actions in A_1 and A_2, respectively, whose induced play, $a_1, a_2, \ldots$, satisfies that for every $t \geq \exp\left((H(Q_1) + H(Q_2))m\right)$, the distribution of $a_{t+1}, \ldots, a_{t+M}$ is an ϵ-approximation of Q^M.

The next theorem is a converse version of Theorem 2.1.

Theorem 2.2. *Let $A = A_1 \times A_2$ be a finite alphabet. Let $Q \in \Delta(A) \setminus (\Delta(A_1) \times \Delta(A_2))$. For every $\delta > 0$ there exists $\epsilon > 0$ such that for every m, M and every pair of mixed m-recall strategies, if the induced play, $a_1, a_2, \ldots$, is such that there exists t_0 such that for every $t \geq t_0$, $a_{t+1}, \ldots, a_{t+M}$ is an ϵ-approximation of Q^M, then*

$$\frac{m}{M} + \delta \geq 1 - \frac{H(Q)}{H(Q_1) + H(Q_2)}.$$

Let us now examine condition (2.1). It can be rearranged as

$$mH(Q_1) + mH(Q2) \geq M(H(Q_1) + H(Q_2) - H(Q)).$$

The expression $H(Q_1) + H(Q_2) - H(Q)$ is the mutual information of the random variables $(a_1, a_2) \mapsto a_i$ defined on the probability space $\langle A_1 \times A_2, Q \rangle$. Formally, the *mutual information* of two discrete random variables X and Y is defined by

$$I(X;Y) = H(X) + H(Y) - H(X,Y).$$

Another way of defining the mutual information is in terms of *conditional entropy*. The *entropy of X given Y* is defined by

$$H(X|Y) = -\sum_{\mathbf{y}} \Pr(Y = \mathbf{y}) \sum_{\mathbf{x}} \Pr(X = \mathbf{x}|Y = \mathbf{y}) \log(\Pr(X = \mathbf{x}|Y = \mathbf{y})).$$

One can verify [CT06, p. 16] the *chain rule of entropy* which is

$$H(X|Y) = H(X,Y) - H(Y).$$

The chain rule implies that

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X).$$

Thus, condition (2.1) can be written as

$$mH(Q_1) + mH(Q2) \geq MI_{Q(x,y)}(x;y) \tag{2.1a}$$

The concavity of the entropy function and the definition of conditional entropy imply the *conditional entropy inequality*

$$H(X|Y) \leq H(X),$$

and hence

$$0 \leq I(X;Y) \leq \min\{H(X), H(Y)\}. \tag{2.2}$$

For $M \leq 2m$, inequality (2.2) implies that the condition (2.1a), and thus the condition of Theorem 2.1, holds for every Q.

Theorem 2.2 is, in fact, a consequence of a more general impossibility result, Theorem 2.3. To obtain an impossibility result, one has to consider every play that can be implemented through mixed m-recall strategies. The set of these plays is not a "nice" mathematical object. To make it more tractable we consider the limiting average play (defined subsequently) instead of the play itself. The limiting average serves as a smoothing operator mapping eventually periodic plays (defined subsequently) to stationary plays. Although it is not injective, we don't lose valuable information by applying this operator since the payoff in the repeated game is a function of the limiting average play.

Note that the play induced by bounded-recall strategies enters a cycle at some point. In other words, it is eventually periodic. Formally, the group $\mathbb{Z}$ acts on $A^{\mathbb{Z}}$ by $(n.a)_t = a_{t+n}$. For a positive integer n and a sequence $a \in A^{\mathbb{N}}$, $n.a$ is also well defined. We say that a play $a \in A^{\mathbb{N}}$ is *eventually periodic*[5] if there exists $T > 0$ such that $T.a = (2T).a$.

The action of $\mathbb{Z}$ extends to measures on $A^{\mathbb{Z}}$ by $(n.\mu)(F) := \mu(\{a : n.a \in F\})$. Let μ be a probability measure on the set of (eventually periodic) plays. The *limiting average* of μ (if it exists) is the following weak limit:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} n.\mu.$$

Note that the limiting average is stationary (i.e., $\mathbb{Z}$-invariant), if it exists. If μ is supported in the set of eventually periodic sequences, then its limiting average exists. In particular, the limiting average exists for the play induced by mixed m-recall strategies.

The following theorem provides constraints on the (random) plays that can be implemented through mixed m-recall strategies.

Theorem 2.3. *Let σ^1, σ^2 be mixed m-recall strategies of agents one and two, respectively. The strategies induce a random play. Let $a_1, a_2, \ldots$ be random variables that distribute according to the limiting average of the induced play. For every integer $M \geq m$. There exist random variables f, g that are functions of $a_1, \ldots, a_M$ such that*

$$\frac{m}{M+1} \geq \frac{\sum_{i=1,2} H(a^i_{M+1}|f)}{\sum_{i=1,2} H(a^i_{M+1})} - \frac{H(a_{M+1}|f)}{\sum_{i=1,2} H(a^i_{M+1})} \tag{2.3}$$

$$\frac{m}{M+1} \geq 1 - \frac{H(a_{M+1}|g)}{\sum_{i=1,2} H(a^i_{M+1}|g)} - \frac{I(a_1, \ldots, a_m; a_{m+1}, \ldots, a_{M+1})}{(M+1)\sum_{i=1,2} H(a^i_{M+1}|g)}. \tag{2.4}$$

Let Q_f be the distribution of a_{M+1} given f. We can rewrite (2.3) as

$$m(H(a^1_{M+1}) + m(H(a^2_{M+1}) \geq (M+1)\, \mathbf{E}\, I_{Q_f(x,y)}(x;y).$$

The above can be expressed in terms of *conditional mutual information.* The *mutual information of X and Y given Z* is defined by

$$I(X;Y|Z) = H(X|Z) + H(Y|Z) - H(X,Y|Z),$$

[5]Equivalently, we say that $a \in A^{\mathbb{N}}$ is *periodic* if there exist $T > 0$ such that $a = T.a$, and $a \in A^{\mathbb{N}}$ is *eventually periodic* if there exists $S > 0$ such that $S.a$ is periodic.

where X, Y, and Z are discrete random variables. In analogy to (2.1a), inequality (2.3) can be rewritten as

$$mH(a^1_{M+1}) + mH(a^2_{M+1}) \geq (M+1)I(a^1_{M+1}; a^2_{M+1}|f). \tag{2.3a}$$

3 Proofs

We begin with some preliminaries. Let A be a finite set. The cardinality of A is denoted by $|A|$. Let l be a positive integer. The set $\{1, \ldots, l\}$ is denoted by $[l]$.

Let $a_1, \ldots, a_l \in A$. The empirical distribution of the sequence $(a_1, \ldots, a_l)$ is the probability measure on A given by

$$\text{emp}(a_1, \ldots, a_l)(\mathbf{a}) = \frac{1}{l} \left|\{t \in [l] : a_t = \mathbf{a}\}\right|.$$

For $Q \in \Delta(A)$, the set of all sequences of length l whose empirical distribution is Q is denoted $T^Q(l)$. Namely,

$$T^Q(l) = \left\{a \in A^l : \text{emp}(a) = Q\right\}.$$

Proposition 3.1. *If $T^Q(l) \neq \emptyset$, then*

$$0 \leq H(Q) - \frac{\log\left|T^Q(l)\right|}{l} \leq \frac{\log|A| + 1}{l}.$$

Proof. See [CT06]. □

3.1 A Proof for Theorem 2.1

We would like to reduce Theorem 2.1 to an analogous combinatorial statement. The combinatorial statement will use the following notion:

Definition 3.2. Let A and B be (finite) sets. A relation $\Gamma \subset A \times B$ is called *homogeneous* if for every $(a, b), (a', b') \in \Gamma$ there exist two permutations $\alpha : A \to A$ and $\beta : B \to B$ such that

- $\alpha(a) = a'$,
- $\beta(b) = b'$,
- $(\alpha, \beta)(\Gamma) \subset \Gamma$.

Lemma 3.3. *Let A and B be finite alphabets. Let $\Gamma \subsetneq A \times B$ be a non-empty homogeneous relation. For every $\epsilon > 0$, there exists $m_0 \in \mathbb{N}$ such that for every M, $m \geq m_0$, if*

$$\frac{m}{(1+\epsilon)M} \geq 1 - \frac{\log|\Gamma|}{\log|A \times B|}$$

then there exist mixed m-recall strategies σ and τ that take actions in the disjoint union $A\dot\cup\{\#\}$ and $B\dot\cup\{\#\}$, respectively, whose induced play, $(a_1, b_1), (a_2, b_2), \ldots$, satisfies that for every $t \geq |A \times B|^m$

$$\frac{1}{M} H((a_{t+1}, b_{t+1}), \ldots, (a_{t+M}, b_{t+M})) \geq \log|\Gamma| - \epsilon,$$

$$\frac{1}{M} \sum_{j=1}^{M} \Pr((a_{t+j}, b_{t+j}) \in \Gamma) \geq 1 - \epsilon.$$

In addition, for every $t \geq 0$ there exists $1 \leq j \leq \frac{m}{2}$ such that $a_{t+j} = b_{t+j} = \#$.

We first argue that Lemma 3.3 implies Theorem 2.1.

Proof. By continuity of the entropy function, we may assume w.l.o.g. that $\frac{m}{(1+\epsilon)M} \geq 1 - \frac{H(Q)}{H(Q_1)+H(Q_2)}$. To see this, consider convex combinations of Q and $Q_1 \otimes Q_2$. For a positive integer w let $\hat{Q}$ be a closest element to Q in $\Delta(A_1 \times A_2) \setminus \Delta(A_1) \times \Delta(A_2)$ with respect to the property $T^{\hat{Q}}(w) \neq \emptyset$. Consider the auxiliary alphabets $T^{\hat{Q}_1}(w)$ and $T^{\hat{Q}_2}(w)$ and the relation $T^{\hat{Q}}(w) \subsetneq T^{\hat{Q}_1}(w) \times T^{\hat{Q}_2}(w)$. The symmetry group S_w acts on $T^{\hat{Q}_1}(w) \times T^{\hat{Q}_2}(w)$ by $\pi.(q_1, \ldots, q_w) = (q_{\pi(1)}, \ldots, q_{\pi(w)})$. The relation $T^{\hat{Q}}(w)$ is an orbit; therefore it is homogeneous.

Since $\lim_{w\to\infty} \frac{1}{w} \log \left| T^{\hat{P}}(w) \right| = H(P)$, for $P = Q, Q_1, Q_2$, we can apply Lemma 3.3 (with a large w) to obtain strategies σ^1 and σ^2 that induce a play with the desired properties.

Two concerns remain: (i) How to interpret the # action; (ii) Since the auxiliary alphabet is composed of sequences of true actions, one has to be able to tell where an auxiliary action begins given a history of the past m true actions.

Both concerns, (i) and (ii), are dealt with together. Consider agent i ($i \in \{1, 2\}$). Choose $\alpha \in A_i$ such that $\hat{Q}_i(\alpha) \neq 1$. Let $\beta \neq \alpha$ be another element of A_i. The interpretation of # will be the word $(\alpha, \ldots, \alpha)$ (i.e., w times α). The desired strategy (of agent i) will be the interpretation of σ^i where any two successive auxiliary actions are separated by a β. The separator β is not a sub-string in the interpretation of the # action; therefore agent i can always identify the # actions. Since it is provided that at least one # action is taken in any time interval of length m, agent i is able to tell exactly where each auxiliary action begins. □

We turn now to proving Lemma 3.3. First we make the following assumptions without loss of generality:

- The sets A and B are minimal with respect to the property that $\Gamma \subset A \times B$.
- $M \geq m$.

We shall describe the strategy of the first agent, σ. The strategy of the second agent, τ, is defined analogously (by replacing A with B as necessary).

Let us write $M = k_1 + \ldots + k_l$, where $\left|k_i - \frac{m}{4}\right| \leq \frac{m}{16}$, $\forall i \in [l]$. For large m, it can be done by taking $l = \lfloor M/\lfloor \frac{m}{4} \rfloor \rfloor$ and $|k_i - k_j| \leq 1$, $\forall i, j \in [l]$. Let $K_i = \sum_{j<i} k_i$, $i \in [l+1]$. At time $t = K_i + 1 \mod M$ the agent takes the action #. At time $t \neq K_i + 1 \mod M$ the agent takes actions in A. The time intervals between the # actions are called *cells*. That is, $sM + [K_{i+1}] \setminus [K_i] = \{sM + K_i + 1, \ldots, sM + K_{i+1}\}$ is a cell , for $s \in \mathbb{N}$ and $i \in [l]$. A time interval of the form $sM + [M] = \{sM + 1, \ldots, (s+1)M\}$ is called an *arch*. The actions in each cell are organized in a structure that will be described later.

The strategy acts in two stages. In the first stage the agent draws a (partially) random cycle that consists of $L_1 \cdot M$ actions. The cycle of the second agent consists of $L_2 \cdot M$ actions, where $\gcd(L_1, L_2) = 1$. Each agent repeats her cycle until they encounter a *matching* arch. The term "matching" will be defined later. Roughly speaking, it is a situation in which a great majority of the actions in the arch are in Γ. In the second stage, the agents repeat the actions of the matching arch forever.

The cell's structure

Consider some cell, $sM+[K_{i+1}]\setminus[K_i]$. The first part of the cell is called the *computed region*. It consists of the time interval $sM+[K_i+r]\setminus[K_i]$, where $r=\mathrm{O}(\log m)$. The purpose of the computed region is to make sure that the strategy is of m-recall. The rest of the cell is called the *random region*. The actions in the random region are drawn randomly, uniformly, and independently of the rest of the actions in the random regions of the entire cycle.

The computed region is composed of three components:

- the cell's number, $i \in [l]$,
- a state, $s \in \{\mathrm{SEARCHING}, \mathrm{SEARCH\_FAILED}, \mathrm{LOOPING}\}$,
- a color (described later).

The information in each component is encoded in the agent's alphabet, A. Since $l \le 6M/m \le \frac{6\log|A\times B|}{\log|A\times B|-\log|\Gamma|}$, the size (number of letters) of the first two components is fixed (it does not depend on m, only on $|\Gamma|$ and $|A\times B|$). The size of the third component is $\mathrm{O}(\log m)$. The precise size of the color component is discussed later.

The cell number

An encoding of the number $i \in [l]$ in the alphabet A.

The state

The state enables the search for a matching arch despite the fact that M can be greater than m, the recall capacity of the agents. We wish to describe the current state as a function of the actions in the previous cell. We begin with a formal definition of a matching arch.

Definition 3.4. Let $a=\left\{(a_t^1,a_t^2)\right\}_{t\in\mathbb{Z}}$ be a sequences of joint actions of the agents. We say that $sM+[M]$ is a *matching arch* of a, if $(a_t^1,a_t^2)\in\Gamma$, for every $t \in sM+\bigcup_{i\in[l]}[K_{i+1}]\setminus[K_i+r])$.

With the above definition in mind, finding a satisfying search scheme is straightforward. Here is the alternative we've chosen. Suppose the previous cell number was i and the previous state was s. The actions taken in the previous cell were $(a_t,b_t)_{t\in[sM+K_{i+1}]\setminus[sM+K_i]}$. The current cell number i' is given by $i'=i+1 \mod l$. The current state s' is given by the following rules:

- if $s=\mathrm{LOOPING}$, then $s'=\mathrm{LOOPING}$.
- if $s=\mathrm{SEARCHING}$,
  - if $\forall t\in[sM+K_{i+1}]\setminus[sM+K_i+r]$, $(a_t,b_t)\in\Gamma$,
    - if $i\ne l$, then $s'=\mathrm{SEARCHING}$.
    - if $i=l$, then $s'=\mathrm{LOOPING}$.
  - otherwise,
    - if $i\ne l$, then $s'=\mathrm{SEARCH\_FAILED}$.
    - if $i=l$, then $s'=\mathrm{SEARCHING}$.
- if $s=\mathrm{SEARCH\_FAILED}$,
  - if $i\ne l$, then $s'=\mathrm{SEARCH\_FAILED}$.
  - if $i=l$, then $s'=\mathrm{SEARCHING}$.

The color

Recall that in the first stage the agent draws a (partially) random cycle $a_1, a_2, \ldots, a_{ML}, a_{ML+1} = a_1, \ldots$ (whose state and color components are as yet undefined). In order to make the cycle implementable through an m-recall strategy, it is sufficient to have

$$\forall 1 \leq s < t \leq M \cdot L \qquad (a_{s+1}, \ldots, a_{s+m}) \neq (a_{t+1}, \ldots, a_{t+m}) \tag{3.1}$$

In the second stage, once a matching arch (say $\Omega = \{sM+1, \ldots, (s+1)M\}$) is found, the agent repeats the actions of that arch forever. In order to implement this behavior with m-recall it is sufficient to have

$$\forall 1 \leq s < t \leq M \cdot L \qquad (a_{s-s\%M+(s+j)\%M})_{j\in[m]} \neq (a_{t-t\%M+(t+j)\%M})_{j\in[m]} \tag{3.2}$$

where $x\%y$ is defined by

$$x\%y \in [y],$$
$$x\%y = x \mod y.$$

Our goal is to set the content of color components such that (3.1) and (3.2) will hold. To do so, we define a graph $G = \langle E, V\rangle$ whose vertices correspond to the cells, i.e., $V = [L] \times [l]$. The edges, E, correspond to places where there is a possibility (by an improper coloring) that either (3.1) or (3.2) will not hold.

Note that the placement of # actions and numbering of cells guarantee that (3.1) and (3.2) can only fail if $s = t \mod M$. We divide (3.1) and (3.2) to $\lfloor \frac{m}{2} \rfloor$ cases (not necessarily disjoint). The kth case is that of $s = t = K_i - k \mod M$ for some $i \in [l]$. Accordingly, let us define sets of edges $E^{(1)}_1, \ldots, E^{(1)}_{\lfloor \frac{m}{2} \rfloor}, E^{(2)}_1, \ldots, E^{(2)}_{\lfloor \frac{m}{2} \rfloor}$, and set

$$E = \bigcup_{\substack{*=1,2\\ 1\leq k\leq \lfloor \frac{m}{2} \rfloor}} E^{(*)}_k.$$

For every $k = 1, \ldots, \lfloor \frac{m}{2} \rfloor$ define:

$$\begin{aligned}
E^{(1)}_k = \Big\{ \{(s,i),(t,i)\} : {}& \forall j \in [m]; \\
& \text{if } a_{s'+j} \text{ and } a_{t'+j} \text{ are defined, then they are equal,} \\
& \text{where } x' = (x-1)M + K_i + 1 - k, \text{ for } x = s,t. \Big\} \\
E^{(2)}_k = \Big\{ \{(s,i),(t,i)\} : {}& \forall j \in [m]; \\
& \text{if } a_{s'-s'\%M+(s'+j)\%M} \text{ and } a_{t'-t'\%M+(t'+j)\%M} \text{ are defined,} \\
& \text{then they are equal,} \\
& \text{where } x' \ (x = s,t) \text{ is the integer that satisfies} \\
& (x-1)M < x' \leq xM, \text{ and } x' = K_i + 1 - k \mod M \Big\}
\end{aligned}$$

Note that if we manage to encode a vertices coloring of G within the color component of the corresponding cells, then (3.1) and (3.2) will be satisfied. To ensure (with high probability) that the chromatic number of G is small enough to fit in the color component, we restrict L from above. Let $\delta > 0$. We restrict L by

$$L \leq \left(\frac{|A|}{1+\delta}\right)^m \tag{3.3}$$

Let d be an integer such that

$$(1+\delta)^d > |A| \tag{3.4}$$

We would like to color G with md colors. For our purposes the size of the color component can be $\lceil \log_A(md) \rceil$. The size of the rest of the computed region is constant (as m grows); therefore the size of the entire computed region, r, is less than $2\log_A(md)$.

The set of vertexes, V, can be partitioned into disconnected sets $V_i = [L] \times \{i\}$, $i \in [l]$. Since l is bounded (by a constant independent of m) we only have to consider the induced graph on V_i for every $i \in [l]$. Let $i \in [l]$. Consider the induced graphs $G_k^{(*)} = \langle V_i, E_k^{(*)}|_{V_i}\rangle$, $k \in [\lfloor \frac{m}{2} \rfloor]$, $* = 1, 2$. In these graphs every connected component is a clique. We estimate the probability that $G_k^{(*)}$ contains a clique of size $d+1$.

$$\begin{aligned}\Pr(G_k^{(*)} \text{ contains a } \mathbf{K}_{d+1}) &\leq \binom{L}{d+1} |A|^{-m(1-\frac{lr}{m})d} \leq L(L^d\,|A|^{-md})\,|A|^{lrd} \leq \\ &\leq L(1+\delta)^{-md}(md)^{2dl} \leq \left[\frac{|A|}{(1+\delta)^d}\right]^m (md)^{2dl}\end{aligned}$$

Thus,

$$\begin{aligned}\Pr(\chi(G) > md) &\leq \Pr(\deg(G) \geq md) \leq \Pr(\exists *, k \ \deg(G_k^{(*)}) \geq d) = \\ &= \Pr(\exists *, k \ G_k^{(*)} \text{ contains a } \mathbf{K}_{d+1}) \leq \left[\frac{|A|}{(1+\delta)^d}\right]^m (md)^{2dl} m \xrightarrow[m\to\infty]{} 0\end{aligned}$$

Let $\mathcal{E}$ be the event that the graph G is md colorable for both agents. We have, so far, proved that $\Pr(\mathcal{E}) \to 1$ (as $m \to \infty$) and that, in $\mathcal{E}$, σ and τ are m-recall strategies. Also, the induced play contains the # action as frequently as required.

We turn, now, to the stopping time in which a matching arch is encountered. Call this stopping time T.

$$T = \min\{sM : [sM] \setminus [(s-1)M] \text{ is a matching arch}\}$$

Clearly, a matching arch exists iff $T < \infty$ iff $T \leq ML_1L_2$. By the restriction we have already posed on L, $ML_1L_2 \leq |A \times B|^m$, for every m large enough. The lemma may be concluded by the following claims:

Claim 3.5. *Conditioned on the event $\{T < \infty\} \cap \mathcal{E}$ the actions in the random region of the matching arch are i.i.d. random variables that distribute uniformly on Γ.*

Claim 3.6.

There exists $\delta > 0$ such that if $L_1 > \left(\frac{|A|}{(1+\delta)}\right)^m$ and $L_2 > \left(\frac{|B|}{(1+\delta)}\right)^m$, then $Pr(T < \infty) \to 1$, as $m \to \infty$.

Proof of Claim 3.5. Denote the random region by $R = \bigcup_{i\in[l]} [K_{i+1}] \setminus [K_i + r]$. Let $(\mathbf{a}, \mathbf{b}), (\mathbf{a}', \mathbf{b}') \in \Gamma^R$. Let $(\alpha, \beta) = (\alpha_j, \beta_j)_{j\in R}$ be permutations that preserve Γ, provided by the homogeneity of Γ. The permutation (α, β) defines a transformation, θ, on $A^{R\times[L_1]} \times B^{R\times[L_2]}$, by operating on every arch at once. The transformation θ is a permutation. The distribution of the random actions is uniform on $A^{R\times[L_1]} \times B^{R\times[L_2]}$; therefore it is θ invariant.

For a random variable X defined on $A^{R\times[L_1]} \times B^{R\times[L_2]}$, denote $\theta.X(\omega) = X(\theta(\omega))$. Since $(\alpha, \beta)(\Gamma^R) = \Gamma^R$, the stopping time T is θ invariant, i.e., $\theta.T = T$. The graph G, described above, is also θ invariant; therefore $\theta(\mathcal{E}) = \mathcal{E}$. Consider the play $(a_t, b_t)_{t\in\mathbb{N}}$ conditioned on $\{T < \infty\} \cap \mathcal{E}$. Apply θ and get a new process $\theta.(a_t, b_t)_{t\in\mathbb{N}}$ that has the same distribution. Since the stopping time is the same for both processes we have

$$\begin{aligned}
&\Pr\left\{(a_t, b_t)_{t\in T-M+R} = (\mathbf{a}', \mathbf{b}')\right\} = \\
&\Pr\left\{\theta.((a_t, b_t)_{t\in T-M+R}) = (\mathbf{a}', \mathbf{b}')\right\} = \\
&\Pr\left\{(\alpha, \beta)((a_t, b_t)_{t\in \theta.T-M+R}) = (\mathbf{a}', \mathbf{b}')\right\} = \\
&\Pr\left\{(\alpha, \beta)((a_t, b_t)_{t\in T-M+R}) = (\mathbf{a}', \mathbf{b}')\right\} = \\
&\Pr\left\{(a_t, b_t)_{t\in T-M+R} = (\mathbf{a}, \mathbf{b})\right\}
\end{aligned}$$

□

Proof of Claim 3.6. Denote the random region by $R = \bigcup_{i\in[l]} [K_{i+1}] \setminus [K_i + r]$. Denote the cycle that agent one (resp. two) draws a (resp. b). For $s \in [L1]$ and $t \in [L2]$, let $E_{s,t}$ be the event that $(a_{sM+j}, b_{tM+j}) \in \Gamma$, $\forall j \in R$. Note that $\bigcup E_{s,t} = \{T < \infty\}$.

We show that $E_{s,t}$ are pairwise independent. Consider w.l.o.g. the events $E_{s,t}$ and $E_{s,t'}$, $t \neq t'$. Fix any $(a) \in A^R$. Obviously, $E_{s,t}$ and $E_{s,t'}$ are independent given $(a)_{s+R}$; therefore it is sufficient to show that $\Pr(E_{s,t} | (a)_{s+R} = \mathbf{a})$ does not depend on $\mathbf{a}$. For $a \in A$ let Γ_a be the section $\{(a, b) \in \Gamma : b \in B\}$. By the assumption that A and B are minimal w.r.t. the property $\Gamma \subset A \times B$, and by the homogeneity of Γ, we have that the cardinality of Γ_a is the same for every $a \in A$. Clearly, $\Pr(E_{s,t} | (a)_{s+R} = \mathbf{a}) = \prod_{j\in R} \frac{|\Gamma_{\mathbf{a}_{s+j}}|}{|B|}$.

Define a random variable $X = \sum_{s,t} \mathbf{1}_{E_{s,t}}$. Write $\alpha = \frac{|\Gamma|}{|A\times B|}$. By Markov's inequality we have

$$\begin{aligned}
\Pr(X = 0) \leq \Pr((X - \mathbf{E}\,X)^2 \geq (\mathbf{E}\,X)^2) \leq \frac{\operatorname{Var} X}{(\mathbf{E}\,X)^2} \leq \frac{1}{\mathbf{E}\,X} = \\
= \frac{1}{L_1 L_2 \Pr(E_{1,1})} \leq \left[\frac{(1+\delta)^2}{|A \times B|}\right]^m \left[\frac{|A \times B|}{|\Gamma|}\right]^M
\end{aligned} \tag{3.5}$$

The lemma assumes that $|A \times B|^m \geq \left[\frac{|A\times B|}{|\Gamma|}\right]^M$. This assumption can be strengthened w.l.o.g. to $|A \times B|^m \geq \left[\frac{|A\times B|}{|\Gamma|}\right]^{M(1+\xi)}$, for some $\xi > 0$. Let δ be so small that $\left[\frac{1}{(1+\delta)^2}\right]^{1-\frac{|\Gamma|}{|A\times B|}} \geq \left[\frac{|A\times B|}{|\Gamma|}\right]^{\frac{\xi}{2}}$, hence

$$\left[\frac{|A \times B|}{(1+\delta)^2}\right]^m \geq \left[\frac{|A \times B|}{|\Gamma|}\right]^{M(1+\frac{\xi}{2})} \tag{3.6}$$

Combining (3.5) and (3.6) we have

$$\Pr(T = \infty) \leq \left[\frac{|\Gamma|}{|A \times B|}\right]^{M(\frac{\xi}{2})} \xrightarrow[M\to\infty]{} 0$$

□

This concludes the proof of Lemma 3.3.

3.2 A proof for Theorem 2.3

For any discrete random variables X, Y and Z, the mutual information of X and Y given Z satisfies

$$I(X;Y|Z) = H(X,Y|Z) - H(X|Y,Z) - H(Y|X,Z). \tag{3.7}$$

In addition, if A and B are independent, then

$$H(X|Y) \leq H(Y) + H(X|Y,A) + H(X|Y,B). \tag{3.8}$$

Consider an arbitrary pair of mixed m-recall strategies for agents one and two. The strategies induce a play. Let $a_1, a_2 \ldots$ be random variables that distribute according to the limiting average of that play. By (3.7), the stationarity of the limiting average, the chain rule of entropy, (3.8), and the conditional entropy inequality (C.E.I.) we get

$$\sum_{t=1}^{M+1-m} I(a^1_{M+1}; a^2_{M+1}|a_t, \ldots, a_M) \underset{\substack{\uparrow\\(3.7)}}{=}$$

$$\sum_{t=1}^{M+1-m}\left[H(a_{M+1}|a_t,\ldots,a_M) - \sum_{i=1,2} H(a_{M+1}|a_t,\ldots,a_M,a^i_{M+1})\right] \underset{\substack{\uparrow\\\text{stationarity}}}{=}$$

$$\sum_{t=m+1}^{M+1}\left[H(a_t|a_1,\ldots,a_{t-1}) - \sum_{i=1,2} H(a_t|a_1,\ldots,a_{t-1},a^i_t)\right] \underset{\substack{\uparrow\\\text{chain rule}}}{=}$$

$$H(a_{m+1},\ldots,a_{M+1}|a_1,\ldots,a_m) - \sum_{t=m+1}^{M+1}\sum_{i=1,2} H(a_t|a_1,\ldots,a_{t-1},a^i_t) \underset{\substack{\uparrow\\(3.8)}}{\leq}$$

$$H(a_1,\ldots,a_m) +$$

$$\sum_{i=1,2}\left[H(a_{m+1},\ldots,a_{M+1}|a_1,\ldots,a_m,\sigma^i) - \sum_{t=m+1}^{M+1} H(a_t|a_1,\ldots,a_{t-1},a^i_t)\right] =$$

$$\underset{\substack{\uparrow\\\text{chain rule}}}{=} H(a_1,\ldots,a_m) +$$

$$\sum_{i=1,2}\sum_{t=m+1}^{M+1}\left[H(a_t|a_1,\ldots,a_{t-1},\sigma^i) - H(a_t|a_1,\ldots,a_{t-1},a^i_t)\right] \leq$$

$$\underset{\substack{\uparrow\\\text{C.E.I: } a^i_t \text{ is a function of } a_{t-m},\ldots,a_{t-1} \text{ and } \sigma^i}}{\leq} H(a_1,\ldots,a_m) \tag{3.9}$$

First, we refer to (2.3), which is equivalent to (2.3a) and to

$$I(a^1_{M+1}; a^2_{M+1}|f) \leq \frac{m}{M+1} \sum_{i=1,2} H(a^i_1).$$

Denote $\alpha = \min_{f:A^M \to A^M} I(a^1_{M+1}; a^2_{M+1}|f(a_1, \ldots, a_M))$. By (3.9) we have

$$(M+1-m)\alpha \leq \sum_{t=1}^{M+1-m} I(a^1_{M+1}; a^2_{M+1}|a_t, \ldots, a_M) \leq$$
$$H(a_1, \ldots, a_m) \leq$$
$$mH(a_1) = m\left[H(a^1_1) + H(a^2_1) - I(a^1_1; a^2_1)\right] \leq m\left[H(a^1_1) + H(a^2_1) - \alpha\right].$$

This concludes (2.3).

Now we turn to (2.4). We assume w.l.o.g. that $M \geq 2m$. Otherwise the theorem holds trivially. Write $\mathcal{I} = I(a_1, \ldots, a_m; a_{m+1}, \ldots, a_{M+1})$. By the chain rule and stationarity, we have

$$H(a_1, \ldots, a_m) = \mathcal{I} + H(a_1, \ldots, a_m|a_{m+1}, \ldots, a_{M+1}) =$$
$$\mathcal{I} + H(a_{M+2-m}, \ldots, a_{M+1}|a_1, \ldots, a_{M+1-m}) =$$
$$\mathcal{I} + \sum_{t=1}^{m} H(a_{M+1}|a_t, \ldots, a_M) \leq$$

by C.E.I and assuming $M \geq 2m$

$$\leq \mathcal{I} + \frac{m}{M+1-m} \sum_{t=1}^{M+1-m} H(a_{M+1}|a_t, \ldots, a_M) =$$
$$\mathcal{I} + \frac{m}{M+1-m} \sum_{t=1}^{M+1-m} \left[\sum_{i=1,2} H(a^i_{M+1}|a_t, \ldots, a_M) - I(a^1_{M+1}; a^2_{M+1}|a_t, \ldots, a_M)\right]. \quad (3.10)$$

Combining m times (3.9) with $M+1-m$ times (3.10) we obtain

$$(M+1-m)\mathcal{I} \geq \sum_{t=1}^{M+1-m} \left[(M+1)I(a^1_{M+1}; a^2_{M+1}|a_t, \ldots, a_M) - m \sum_{i=1,2} H(a^i_{M+1}|a_t, \ldots, a_M)\right].$$

Take a minimal summand, indexed t', from the right-hand side of the above inequality. Let $g = (a_{t'}, \ldots, a_M)$. Thus,

$$\mathcal{I} \geq (M+1)I(a^1_{M+1}; a^2_{M+1}|g) - m \sum_{i=1,2} H(a^i_{M+1}|g),$$

which is equivalent to (2.4). □

3.3 A proof for Theorem 2.2

The assumption implies that the limiting average play $b_1, b_2, \ldots$ satisfies

$$\frac{1}{M} H(b_1, \ldots, b_M) \geq H(b_1) - r(\epsilon) \qquad (3.11)$$
$$\|\mathrm{P}(b_1) - Q\| \leq \epsilon$$

where $r(\epsilon) \to 0$, as $\epsilon \to 0$.

Let $\delta > 0$. Choose $0 < \alpha < 1$ such that

$$\frac{\delta}{2} > (1-\alpha)\left(1 - \frac{H(Q)}{H(Q_1) + H(Q_2)}\right). \qquad (3.12)$$

Let $M' = \lfloor \alpha M \rfloor$. By (3.11) and the stationarity of the limiting average play,

$$H(b_1) - r(\epsilon) \leq \frac{1}{M} H(b_1, \ldots, b_M) \leq$$
$$\frac{1}{M} H(b_1, \ldots, b_{M'}) + \frac{M - M'}{M} H(b_{M'+1} | b_1, \ldots, b_{M'}) \leq$$
$$\frac{M'}{M} H(b_1) + \frac{M - M'}{M} H(b_{M'+1} | b_1, \ldots, b_{M'}) \leq H(b_1);$$

hence

$$H(b_1) - \frac{r(\epsilon)}{1-\alpha} \leq H(b_{M'+1} | b_1, \ldots, b_{M'}) \leq H(b_1);$$

hence, for every random variable[6] $f = f(b_1, \ldots, b_{M'})$,

$$H(b_1) - \frac{r(\epsilon)}{1-\alpha} \leq H(b_{M'+1} | f) \leq H(b_1).$$

For $i = 1, 2$, the above implies

$$H(b_1^i) - \frac{r(\epsilon)}{1-\alpha} \leq H(b_{M'+1}^i | f) \leq H(b_1^i), \qquad (3.13)$$

since

$$0 \leq H(b_{M'+1}^i) - H(b_{M'+1}^i | f) = H(f) - H(f | b_{M'+1}^i) \leq$$
$$H(f) - H(f | b_{M'+1}) = H(b_{M'+1}) - H(b_{M'+1} | f) \leq \frac{r(\epsilon)}{1-\alpha}.$$

Let $f = f(b_1, \ldots, b_{M'})$ be the random variable provided by Theorem 2.3. Theorem 2.3, (3.12), and (3.13) imply that, if ϵ is chosen small enough, then

$$\frac{m}{M} + \delta \geq 1 - \frac{H(Q)}{H(Q_1) + H(Q_2)}.$$

[6] I.e., f is measurable with respect to the σ-algebra generated by $b_1, \ldots, b_{M'}$.

4 The individually rational level

In this section we obtain non-trivial results on the individually rational level in repeated games with bounded recall. We begin by introducing the notation, then presenting previously known results, and conclude by presenting new results and a few conjectures.

Let $G = \langle [N], A = A_1 \times \cdots \times A_N, g : A \to \mathbb{R}^N \rangle$ be a finite N-person game in normal form. The *individually rational level* (also called the *min-max level*) of player i is the level below which the team of all players but player i cannot force the payoff of player i. Formally, it is defined by

$$\mathrm{IRL}_i\, G = \min_{\sigma^{-i} \in \prod_{j \neq i} \Delta(A_j)} \max_{\sigma^i \in A_i} g^i(\sigma).$$

The *security level* (also called the *max-min level*) of player i is the value that player i can guarantee against all other players. It is the value of the two-person zero-sum game defined by player i's payoff function. We denote this game G^i. The security level of player i is the following quantity:

$$\begin{aligned}\operatorname{val} G^i &= \max_{\sigma^i \in \Delta(A_i)} \min_{\sigma^{-i} \in \prod_{j \neq i} A_j} g^i(\sigma) \\ &= \min_{\sigma^{-i} \in \Delta(\prod_{j \neq i} A_j)} \max_{\sigma^i \in A_i} g^i(\sigma).\end{aligned}$$

We denote by $G[m_1, \ldots, m_N]$ the infinitely repeated version of G, where each player i is restricted to m_i-recall strategies.

Consider a sequence of games $G[m_1(n), \ldots, m_N(n)]$, $n \in \mathbb{N}$. By considering only uncorrelated actions, Theorem 1 in [Leh88] implies that if $\lim_{n\to\infty} \frac{\log m_i(n)}{m_j(n)} = 0$, for all $i, j \in [N]$, then

$$\begin{aligned}\operatorname{val} G^i = \lim_{n\to\infty} \operatorname{val}\, (G[m_1(n), \ldots, m_N(n)])^i \leq \\ \liminf_{n\to\infty} \mathrm{IRL}_i\, G[m_1(n), \ldots, m_N(n)] \leq \\ \limsup_{n\to\infty} \mathrm{IRL}_i\, G[m_1(n), \ldots, m_N(n)] \leq \mathrm{IRL}_i\, G. \quad (4.1)\end{aligned}$$

In the case of two-person games (i.e. $N = 2$), the individually rational and security levels are the same; therefore inequality (4.1) provides a tight estimation for the bounded-recall individually rational level.

In the case of $N > 2$ the picture is less clear. Even the special case where all the players have the same recall capacity is not resolved. Currently, we are not even able to show the existence of the limit of $\mathrm{IRL}_i\, G[m, \ldots, m]$, as m grows. One may erroneously presume that if a player's recall capacity is at least as large as that of the other players, then the other players cannot device a correlated punishment against that player. The following corollary of Theorem 2.1 shows that $\limsup_{m\to\infty} \mathrm{IRL}_i G[m_1, \ldots, m_N]$ can be arbitrarily close to $\operatorname{val} G^i$ even if player i's recall capacity is greater than that of the other players:

Corollary 4.1. *For every $\epsilon > 0$ and every positive integer C, there exists a finite three-player game G such that*

- $\mathrm{IRL}_3\, G = 1$,

- $\operatorname{val} G^3 = 0$,
- $\limsup_{m\to\infty} \mathrm{IRL}_3\, G[m,m,Cm] < \epsilon$.

Proof. Let $G = \langle [3], A = A_1 \times A_2 \times A_3, g : A \to \mathbb{R}^3 \rangle$ be a three-person game. We say that two actions of player one, a and b, are strategically equivalent if $g(a,x,y) = g(b,x,y)$, for every pair of actions x and y of player two and three, respectively. Let $\{G_n\}$ be games obtained from G by appending player one's action space with n new actions $\{x_1,\ldots,x_n\}$, which are strategically equivalent to some action $a \in A_1$.

Claim 4.2. *For every positive integer C,*

$$\lim_{n\to\infty} \limsup_{m\to\infty} \mathrm{IRL}_3\, G_n[m,m,Cm] = \operatorname{val} G^3.$$

Corollary 4.1 stems from Claim 4.2 by choosing G such that $\mathrm{IRL}_3\, G = 1$ and $\operatorname{val} G^3 = 0$. Such a game exists since any game in which the individually rational level and security level differ can be transformed to a game with the desired properties through an appropriate affine transformation of the payoff function of player three.

We turn, now, to proving Claim 4.2. Let $Q \in \Delta(A_1 \times A_2)$ be an optimal strategy for the team of players one and two in the game G^3. Choose $P_n \in \Delta(\{x_1,\ldots,x_n\} \times A_2)$, a strategy for the team in the game G_n, such that $H(P_n) \geq \log n$. Let $Q_n = (1 - \frac{1}{\sqrt{\log n}})Q + \frac{1}{\sqrt{\log n}} P_n$. Let $M = M_{m,n}$ be the largest integer that satisfies

$$\frac{m}{M_{m,n}} \geq 1 - \frac{H(Q_n)}{H((Q_n)_1) + H((Q_n)_2)} \tag{4.2}$$

The right-hand side of (4.2) converges to zero, as $n \to \infty$. Let $1 \geq \epsilon > 0$ and $n \in \mathbb{N}$ such that

$$\frac{1}{\sqrt{\log n}} \leq \frac{\epsilon}{3},$$
$$\limsup_m \frac{Cm}{M_{m,n} - Cm} \log |A_3| \leq \frac{\epsilon}{3}.$$

By Theorem 2.1, there exists m_0 such that for every $m \geq m_0$ there exist m-recall strategies for players one and two that ignore player three's actions and the limiting average of the induced play $(a_t^{1,2})_{t=1}^\infty$ satisfies

$$\frac{Cm}{M - Cm} \log |A_3| \leq \frac{\epsilon}{3},$$
$$\frac{1}{M} H(a_1^{1,2},\ldots,a_M^{1,2}) \geq H(a_1^{1,2}) - \frac{\epsilon}{3},$$
$$\sum_{\mathbf{a}} \left| \Pr(a_1^{1,2} = \mathbf{a}) - Q_n(\mathbf{a}) \right| \leq \frac{\epsilon}{3}.$$

We use an argument called Neyman-Okada's criterion (see [Per08, pp. 7-8]). Let τ be a pure Cm-recall strategy for player three. The team's mixed strategies and τ induce a play $(\hat{a}_t)_{t=1}^\infty$, $\hat{a}_t = (\hat{a}_t^1, \hat{a}_t^2, \hat{a}_t^3)$. Let $\hat{\tau}_t = \tau_{|\hat{a}_1,\ldots,\hat{a}_t}$ be player three's strategy given the history up to time t. Denoting by $(a_t, \tau_t)_{t\in\mathbb{Z}}$ the limiting average of the process $(\hat{a}_t, \hat{\tau}_t)_{t=1}^{\infty}$, we have

$$\Pr(a_1 \in A) I(\ a_1^{1,2};\ a_1^3 | a_1 \in A) \le I(\ a_1^{1,2};\ a_1^3) = H(\ a_1^{1,2}) - H(\ a_1^{1,2}|\ a_1^3) \le$$
$$H(\ a_1^{1,2}) - \frac{1}{M-Cm} H(\ a_{Cm+1}^{1,2}, \ldots,\ a_M^{1,2}|\ \tau_{Cm}) \le$$
$$H(\ a_1^{1,2}) - \frac{1}{M-Cm} H(\ a_{Cm+1}^{1,2}, \ldots,\ a_M^{1,2} | a_1^{1,2}, \ldots,\ a_{Cm}^{1,2},\ a_1^3, \ldots,\ a_{Cm}^3) \le$$
$$H(\ a_1^{1,2}) - \frac{1}{M-Cm} \Big[H(\ a_1^{1,2}, \ldots,\ a_M^{1,2}) - H(a_1^{1,2}, \ldots,\ a_{Cm}^{1,2}) -$$
$$H(\ a_1^3, \ldots,\ a_{Cm}^3) \Big] \le H(\ a_1^{1,2}) - \frac{M}{M-Cm} \left[H(\ a_1^{1,2}) - \frac{\epsilon}{3} \right] +$$
$$\frac{Cm}{M-Cm} \left[H(\ a_1^{1,2}) \right] + \frac{Cm}{M-Cm} \log |A_3| \le 2\frac{\epsilon}{3} + \frac{\epsilon}{3}.$$

Consequently, since $\Pr(a_1 \in A) \ge 1 - \frac{2}{3}\epsilon \ge \frac{1}{3}$,

$$\mathbf{E}[g^3(a_1)|a_1 \in A] \le \max\{ \mathop{\mathbf{E}}_{P(a)} g^3(a) :$$
$$P \in \Delta(A),\ I_{P(a)}(a^{1,2}; a^3) \le 3\epsilon,\ \|P_{A_1\times A_2} - Q\| \le \epsilon\} =: V_\epsilon;$$

hence

$$\mathbf{E}[g^3(a_1)] \le (1-\epsilon)V_\epsilon + \epsilon \max_{a\in A} g^3(a) \xrightarrow[\epsilon\to 0]{} \operatorname{val} G^3.$$

We have just established the following upper bound:

$$\forall C \in \mathbb{N}\ \limsup_n \limsup_m \operatorname{IRL}_3 G_n[m, m, Cm] \le \operatorname{val} G^3.$$

The proof is concluded with an equal lower bound, given in (4.1). □

Theorems 2.1 and 2.3 suggest that the ratio between the recall capacities of the players is the relevant quantity in estimating the individually rational level. Theorem 2.1 and Neyman-Okada's criterion [Per08, pp. 7-8] resolve the extreme case in which the recall capacity of the maximizing player is much smaller than that of the other players. Let G be a finite three-person game.

Proposition 4.3.

$$\lim_{C\to\infty} \lim_{m\to\infty} \operatorname{IRL}_1 G[m, Cm, Cm] = \operatorname{val} G^1.$$

The other extreme case, where the recall capacity of the maximizing player is much greater than that of the other players is left as a conjecture.

Conjecture 4.4.

$$\lim_{C\to\infty} \lim_{m\to\infty} \operatorname{IRL}_1 G[Cm, m, m] = \operatorname{IRL}_1 G.$$

Proposition 4.3 can be strengthened. The number of players can be greater than tree. Moreover, the minimizing team may be restricted to pure Cm-recall strategies, for some constant C that depends on the one-step game.

Proposition 4.5. *Let G be an N-person game. There exists a constant $C = C(G)$, such that*

$$\lim_{m\to\infty} \min_{\substack{\sigma^j \in \Sigma_j(Cm)\\ j=2,\dots,N}} \max_{\sigma^1\in\Sigma_1(m)} g^1(\sigma) = \operatorname{val} G^1,$$

where $g^1(\sigma) := \mathbf{E}\, g^1(a_1)$, and $a_1, a_2, \ldots$ is the limiting average play induced by $\sigma = (\sigma^1, \ldots, \sigma^N)$.

Proof. Restricting the team to pure strategies turns the game into a two-person zero-sum game, the game $G^1[m, Cm]$; therefore we may assume without loss of generality that $G = \langle A = A_1 \times A_2, g : A \to \mathbb{R}\rangle$ is in deed a two-person zero-sum game. Let $q \in \Delta(A_2)$ be an optimal strategy for player two. If $H(q) = 0$, then the result holds trivially. Let $C > \frac{\log|A|}{H(q)}$.

Let $\delta = 1/2(H(q)C - \log|A|)$, and $n = n_m = \lfloor \exp((H(q)C - \delta)m)\rfloor$. The first stage in the construction of the strategies in the proof of Lemma 3.3 (alternatively, Proposition 4.9 in [Per08]) show that there exists an oblivious mixed strategy $\mu_m \in \Delta(\Sigma_2(Cm))$ that induces a play $y_1, y_2, \ldots$ such that for every $T \in \mathbb{N}$

- $\|\mathrm{emp}(y_{T+1}, \ldots, y_{T+n}) - q\| = \mathrm{o}(1)$ μ_m almost surely,
- $\frac{1}{n} H(y_{T+1}, \ldots, y_{T+n}) \geq H(q) - \mathrm{o}(1)$.

We would like to show that there exist strategies $\tau_m \in \mathrm{support}(\mu_m)$ that guarantee a payoff of at most $\operatorname{val} G + \mathrm{o}(1)$ against any m-recall strategy of player one. Assume by negation that there exist $\epsilon > 0$ such that for every $m_0 \in \mathbb{N}$ there exists $m > m_0$ and a mapping $\sigma : \Sigma_2(Cm) \to \Sigma_1(m)$ such that $g(\sigma(\tau), \tau) \geq \operatorname{val} G + \epsilon$, $\mu_m(\tau)$ almost surely. It follows that $\mathbf{E}_{\mu_m} g(\sigma(\tau), \tau) \geq \operatorname{val} G + \epsilon$; and hence there exists $T \in \mathbb{N}$ such that

$$\mathop{\mathbf{E}}_{\mu_m}\left[\frac{1}{n}\sum_{l=T+1}^{T+n} g(x_l, y_l)\right] \geq \operatorname{val} G + \epsilon,$$

where $x_1, y_1, x_2, y_2, \ldots$ is the play induced by $\sigma(\tau)$ and τ.

Let Q_l be the distribution of (x_l, y_l) and $Q = \frac{1}{n}\sum_{l=T+1}^{T+n} Q_l$. The negation assumption yields $g(Q) \geq \operatorname{val} G + \epsilon$.

Lemma 4.2 in [Per08] ensures that

$$I_{Q(x,y)}(x; y) \leq \left[H_{Q(x,y)}(y) - \frac{1}{n} H(y_{T+1}, \ldots, y_{T+n})\right] + \frac{1}{n} H(\sigma(\tau)) \leq$$
$$\mathrm{o}(1) + \frac{\log|\Sigma_1(m)|}{n} \leq \mathrm{o}(1) + \mathrm{O}\left(\left[\frac{|A|}{|A| + \delta}\right]^m\right) \xrightarrow[m\to\infty]{} 0,$$

contradicting the optimality of q. □

The intuition for Conjecture 4.4 is as follows: Assume that neither of the players observed player three's actions. Under this signaling structure, Theorem 2.3 ensures that there exists a random variable f, such that f is observable by player three and the limiting average play of the team (players one and two) satisfies

$$I(a^1_{Cm+1}; a^2_{Cm+1} | f) < \frac{1}{C}(H(a^1_{Cm+1}) + H(a^2_{Cm+1})) \leq \frac{1}{C}\log|A| \xrightarrow[C\to\infty]{} 0 \qquad (4.3)$$

In other words, as C grows, the actions taken by the team tend to be independent conditional on the last Cm periods of history. Alas, since the team *can* respond to the actions of player three our argument fails to prove Conjecture 4.4. It does, however, prove the weaker statement of Proposition 4.7, as will be shown subsequently.

Denote by $\mathring{G}[m_1, m_2, m_3]$ the game with the signaling structure that allows everybody to see only the actions of players one and two. For a positive real number C, define a set

$$F_G(C) = \left\{ Q \in \Delta(A_1 \times A_2) : \frac{1}{C} \geq 1 - \frac{H(Q)}{H(Q_1) + H(Q_2)} \right\}.$$

Theorem 2.1 ensures that

$$\limsup_m \mathrm{IRL}_3\, \mathring{G}[m, m, \lfloor Cm \rfloor] \leq \min_{\sigma^{-3} \in F_G(C)} \max_{\sigma^3 \in A_3} g^3(\sigma). \tag{4.4}$$

Since $F(2) = \Delta(A_1 \times A_2)$, we conclude that

$$\lim_m \mathrm{IRL}_3\, \mathring{G}[m, m, 2m] = \mathrm{val}\, G^3.$$

The same argument can be pushed a little further. Define

$$\bar{F}_G(C) = \left\{ Q \in \Delta(A_1 \times A_2 \times \mathbb{N}) : \frac{1}{C} \geq 1 - \frac{H(x, y|n)}{H(x|n) + H(y|n)} \right\}$$

where the distribution of (x, y, n) is Q.

Proposition 4.6.

$$\limsup_m \mathrm{IRL}_3\, \mathring{G}[m, m, \lfloor Cm \rfloor] \leq \inf_{Q \in \bar{F}_G(C)} \sup_{a:\mathbb{N} \to A_3} \mathop{\mathbf{E}}_{Q(x,y,n)} g^3(x, y, a(n))$$

Remark. It can be shown that the infimum is obtained by a finitely supported measure; therefore the infimum can be replaced by a minimum and the supremum by a maximum.

Proof. Let G_n be the n-fold repeated version of G with average per-stage payoffs. Note that, for every $n \in \mathbb{N}$,

$$\mathrm{IRL}_3\, \mathring{G}[m, m, \lfloor Cm \rfloor] \leq \mathrm{IRL}_3\, \mathring{G}_n[\lfloor m/n \rfloor - 1, \lfloor m/n \rfloor - 1, \lceil Cm/n \rceil]. \tag{4.5}$$

Define

$$F_G^n(C) = \left\{ Q \in \bar{F}_G(C) : \forall k \in \mathbb{N}\ Q(A_1 \times A_2 \times \{k\})n \in \mathbb{Z} \right\}.$$

Note that for every $Q \in F_G^n(C)$, $|\text{support}\, Q_{\mathbb{N}}| \leq n$. Note, also, that $\bar{F}_G(C)$ is the Hausdorff limit of $F_G^n(C)$ with respect to the total variation norm. By (4.5) and (4.4), we have

$$\begin{aligned}\limsup_m \mathrm{IRL}_3\, \mathring{G}[m, m, \lfloor Cm \rfloor] &\leq \min_{\sigma^{-3} \in F_{G_n}(C)} \max_{\sigma^3 : (A_1 \times A_2)^{<n} \to A_3} g_n^3(\sigma) \leq \\ &\min_{Q \in F_G^n(C)} \max_{a:\mathbb{N} \to A_3} \mathop{\mathbf{E}}_{Q(x,y,n)} g^3(x, y, a(n)) \xrightarrow[n \to \infty]{} \\ &\min_{Q \in \bar{F}_G(C)} \max_{a:\mathbb{N} \to A_3} \mathop{\mathbf{E}}_{Q(x,y,n)} g^3(x, y, a(n)).\end{aligned} \tag{4.6}$$

The second inequality follows from restricting σ^{-3} to oblivious strategies whose induced play is a sequence of independent random actions. The best response of player three is to play a fixed sequence of actions. □

Define

$$\bar{\bar{F}}_G(C) = \left\{ Q \in \Delta(A_1 \times A_2 \times \mathbb{N}) : \frac{1}{C} \geq \frac{I(x;y|n)}{H(x)+H(y)} \right\},$$

where the distribution of (x,y,n) is Q. Note that $\bar{\bar{F}}_G(C) \supset \bar{F}_G(C)$ since

$$\bar{F}_G(C) = \left\{ Q \in \Delta(A_1 \times A_2 \times \mathbb{N}) : \frac{1}{C} \geq \frac{I(x;y|n)}{H(x|n)+H(y|n)} \right\}.$$

The inequality (4.3) proves the following lower bound:

Proposition 4.7.

$$\liminf_m \mathrm{IRL}_3\, \mathring{G}[m,m,\lfloor Cm \rfloor] \geq \inf_{Q \in \bar{\bar{F}}_G(C)} \sup_{a:\mathbb{N}\to A_3} \mathop{\mathbf{E}}_{Q(x,y,n)} g^3(x,y,a(n))$$

Remark. A different lower bound can be deduced similarly from (2.4). Nevertheless, it will not close the gap between the lower and upper bounds.

Acknowledgments

The author would like to thank his Ph.D. advisor Professor Abraham Neyman for his guidance. The author is also grateful to his hosts at Tel-Aviv University, Eilon Solan and Ehud Lehrer, for their comments and advise.

References

[AS00] Noga Alon and Joel H. Spencer. *The Probabilistic Method.* John Wiley & Sons, New York, second edition, 2000.

[BN03] Gilad Bavly and Abraham Neyman. Online concealed correlation by boundedly rational players. Discussion Paper 336, Center for the Study of Rationality, Hebrew University, Jerusalem, September 2003.

[BP93] Elchanan Ben-Porath. Repeated games with finite automata. *Journal of Economic Theory*, 59(1):17–32, February 1993.

[CT06] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory.* Wiley Interscience, New York, 2006.

[GT07] Olivier Gossner and Tristan Tomala. Secret correlation in repeated games with imperfect monitoring. *Mathematics of Operations Research*, 32:413–424, May 2007.

[Leh88] Ehud Lehrer. Repeated games with stationary bounded recall strategies. *Journal of Economic Theory*, 46(1):130–144, October 1988.

[Leh94] Ehud Lehrer. Finitely many players with bounded recall in infinitely repeated games. *Games and Economic Behavior*, 7(3):390–405, November 1994.

[Ney08] Abraham Neyman. Learning effectiveness and memory size. Discussion Paper 476, Center for the Study of Rationality, Hebrew University, Jerusalem, February 2008.

[NO00] Abraham Neyman and Daijiro Okada. Repeated games with bounded entropy. *Games and Economic Behavior*, 30(2):228–247, February 2000.

[NO09] Abraham Neyman and Daijiro Okada. Growth of strategy sets, entropy, and nonstationary bounded recall. *Games and Economic Behavior*, 66(1):404 – 425, 2009.

[NS10] Abraham Neyman and Joel Spencer. Complexity and effective prediction. *Games and Economic Behavior*, 69(1):165–168, May 2010.

[Per08] Ron Peretz. The strategic value of recall. Submitted to *Games and Economic Behavior*, 2008.