A Minority Game with Bounded Recall

Jérôme Renault
CEREMADE
Université Paris Dauphine
Pl. du Marechal de Lattre de Tassigny
F–75775 Paris Cedex 16, France
renault@ceremade.dauphine.fr

Marco Scarsini
Dipartimento di Statistica e Matematica Applicata
Università di Torino
Piazza Arbarello 8
I–10122 Torino, Italy
marco.scarsini@unito.it

Tristan Tomala
CEREMADE
Université Paris Dauphine
Pl. du Marechal de Lattre de Tassigny
F–75775 Paris Cedex 16, France
tomala@ceremade.dauphine.fr

July 2006

Abstract

This paper studies a repeated *minority game* with public signals, symmetric bounded recall and pure strategies. We investigate both *public* and *private* equilibria of the game with fixed recall size. We first show how public equilibria in such repeated games can be represented as colored sub-graphs of a de Bruijn graphs. Then we prove that the set of public equilibrium payoffs with bounded recall converges to the set of uniform equilibrium payoffs as the size of the recall increases. We also show that private equilibria behave badly: a private equilibrium payoff with bounded recall need not be a uniform equilibrium payoff.

Key words: folk theorem, de Bruijn sequence, imperfect monitoring, uniform equilibrium, public equilibrium, private equilibrium.

MSC 2000 Subject Classification: Primary 91A20. Secondary 91A18, 91A43.

OR/MS subject classification: Primary: Games/group decisions, noncooperative.

1 Introduction

Repeated games with complete information are known to have multiple equilibria. The prominent result in this direction is the folk theorem which asserts that in games with perfect monitoring and perfectly rational players, every feasible and individually rational payoff can be sustained by an equilibrium of the repeated game. A more realistic model to study involves games with imperfect monitoring, where players observe imperfectly other players' actions, and bounded rationality, where players have limited information processing abilities. Typically these two problems have been studied separately in the literature. A notable exception is a recent paper by Cole and Kocherlakota (2005).

The literature on games with imperfect monitoring seeks to characterize the set of equilibrium payoffs (see e.g., Lehrer (1992a,b), Abreu et al. (1990), Fudenberg and Levine (1994), Tomala (1998), Renault and Tomala (2004)), and the literature on games with bounded rationality examines whether equilibrium payoffs of the unrestricted repeated game can be approximated by equilibrium payoffs of the repeated game with bounded rationality (see e.g., Rubinstein (1986), Abreu and Rubinstein (1988), Kalai and Stanford (1988), Lehrer (1988, 1994) Ben-Porath (1990, 1993), Sabourian (1998), Neyman (1998), Bavly and Neyman (2005)).

The present paper aims at blending these two approaches in the analysis of a *minority game*. In this class of games each player has two actions and aims at choosing the action that is less popular among all players. The game is repeated and after each stage the most popular (or equivalently the less popular) action is publicly announced.

In many real life situations it is preferable to be in the minority. Think for instance of a residential suburban area that is linked to downtown by two main roads. Commuters have to decide every morning which road to take and, for obvious reasons, they all want to avoid traffic. Since the commuters typically do not recognize their fellow commuters on the road, but only perceive the existence of traffic, this phenomenon can be modelled as a game with imperfect public monitoring. Attention to these phenomena originates in some papers by Arthur (1994, 1999). In general this class of games is interesting when several agents must take decentralized decision on whether to access a scarce resource, knowing that at most a fixed number of them will be able to enjoy its benefits. Similar situations have been analyzed by the empirical economic literature on market entry games (see e.g., Selten and Güth (1982), Ochs (1990, 1995), Rapoport et al. (2002), Erev and Rapoport (1998)). In these models players must decide independently whether to enter a market (and incur an entry cost). Since capacity is limited, the entrants will reap a reward only if their number is smaller than a fixed threshold. This clearly creates a problem of coordination. Similar ideas have been used to analyze speculative behavior in financial markets, and minority games have been used as a formalization of concepts, like the *contrarian investment strategy*, previously considered in empirical studies (see e.g., De Bondt and Thaler (1985), Chan (1988)). Most of the literature on the topic can be found in theoretical physics journals and has a non-strategic approach. The reader is referred to the recent books by Challet et al. (2005) and Coolen (2005) for a history of the problem, its statistical-mechanics analysis, and some applications to financial markets. An analysis of minority games from the learning viewpoint can be found in Bottazzi et al. (2003, 2002), Bottazzi and Devetag (2004). A minority game is basically a repeated coordination game, the specific feature being that players want to coordinate negatively and be where their fellow players are not. The stage game used in a repeated minority game is strongly related to congestion games introduced by Rosenthal (1973) and to crowding games studied by Milchtaich (1998, 2000).

Renault et al. (2005) are the first to consider a minority game from a traditional strategic viewpoint. In their model an odd number of players have to choose simultaneously one of two rooms. The players who choose the less crowded room receive a reward of one euro. The others receive nothing. The game is repeated over time. At each step, after the players' action choices, only a public signal (the majority room) is announced to everybody, so players do not observe the actions or the payoffs of the other players. Renault et al. (2005) prove an undiscounted folk theorem for this game, and characterize the set of uniform equilibrium payoffs, i.e. they show that any feasible payoff is an equilibrium payoff. In particular, they construct a uniform equilibrium where the payoff of each player is zero. It is interesting to notice that a folk theorem exists for this game even if no identifiability condition à la Fudenberg et al. (1994) holds.

The paper by Renault et al. (2005) is in the tradition of repeated games with imperfect public monitoring. Examples of such games go back to Rubinstein (1979), Rubinstein and Yaari (1983), and Radner (1985) with reference to principal-agent models and by Green and Porter (1984) with reference to oligopoly. More systematic analyses of games with imperfect public monitoring have been provided by Lehrer (1989) and Tomala (1998) in the undiscounted case and Abreu et al. (1990) Fudenberg et al. (1994), Fudenberg and Levine (1994) in the discounted case. Most papers focus on public equilibria, namely, equilibria in which each player uses only strategies that depend on the public signal and not on her private history. More recently a considerable attention has been devoted to games with imperfect private monitoring. See for instance the whole issue of *Journal of Economic Theory* dedicated to this topic, with the introduction by Kandori (2002).

Discounted and finitely repeated versions of the minority game are studied by Renault et al. (2006). Their model deals with an intermediate situation where the signal is public, but strategies are private, namely, they depend on the public signal *and* on the private history of each player. Games with public monitoring and private strategies have been studied by Mailath et al. (2002) and Kandori and Obara (2006). Mailath et al. (2002) deal with finitely repeated games and show three examples of substantially different behavior of private versus public strategies in games with imperfect public monitoring. Kandori and Obara (2006) consider infinitely repeated games and provide a method to construct private strategies that are more efficient than public strategies.

In the present paper a three-player minority game with imperfect public monitoring and bounded recall is studied, and only pure strategies are considered. Bounded recall and public signal is a typical assumption for minority games in the physics literature. We first analyze public equilibria. Public strategy profile in those games can be represented as the choice of a subgraph in a de Bruijn graph, together with a coloring of the vertices, i.e. a rule that assigns each vertex to a player. Using these tools we compute some equilibria.

We look then at the asymptotic behavior of the set of bounded recall equilibrium payoffs. For any game with bounded recall and imperfect public monitoring, the set of public equilibria with bounded recall is a subset of the set of public equilibria with unbounded recall and the set of public-equilibrium payoffs increases with the size of the recall. But, for some games, it may not converge to the set of unbounded recall public equilibrium payoffs. For instance, consider a repeated game with a public blank signal. Since player have no information, the set of unbounded recall equilibrium payoffs is the convex hull of stage-Nash payoffs. In a game with bounded recall and public strategies, the public memory is always empty, so players always choose the same action and bounded recall public equilibria are nothing but stage-Nash equilibria.

For the minority game, we show that the set of public equilibrium payoffs does converge to the set of unbounded recall public equilibrium payoffs, as the length of recall increases.

The set of private equilibria lacks the nice properties of public equilibria and we exhibit a private equilibrium with recall 3 whose payoff does not lie in the set of unbounded-recall-private-equilibrium payoffs. These results are somehow connected to Mailath et al. (2002) and Kandori and Obara (2006), who also compare public and private equilibria, but, to the best of our knowledge, this paper is the first that considers such a comparison in a bounded recall framework.

The paper is organized as follows. Section 2 describes a model of repeated games with imperfect public monitoring and bounded recall. Section 3 deals with a minority game and gives the main results. In Section 4 examples of other repeated games are considered. Finally Section 5 contains the proofs of the results.

2 Repeated games with public signals

2.1 Description of the model

Consider a stage game

$$G = \langle N, (A^i)_{i \in N}, (g^i)_{i \in N} \rangle. \tag{2.1}$$

In this setting N is a set of players, for each $i \in N$, A^i is the set of actions available to player i, $A := \times_{i \in N} A^i$ is the set of action profiles, and the map $g^i : A \to \mathbb{R}$ is the payoff function for player i. Denote by $g : A \to \mathbb{R}^N$ the vector payoff function $(g^i)_{i \in N}$. For every $i \in N$, put $A^{-i} = \times_{j \in N, j \neq i} A^j$, therefore $a^{-i} \in A^{-i}$ will be a shortcut for $(a^j : j \neq i) \in \times_{j \in N, j \neq i} A^j$. Consider then a set of signals U and a mapping $\ell : A \to U$. In the whole paper the sets N, A^i, U are assumed nonempty and finite.

This game is repeated over time. At each round $t = 1, 2, \ldots$, players choose actions and if $a_t \in A$ is the action profile at stage t, they observe a public signal $u_t = \ell(a_t)$ before proceeding to the next stage. The set of histories of length $t \geq 0$ for player i is $\mathcal{H}^i_t := (A^i \times U)^t$, $\mathcal{H}^i_0$ being a singleton, and $\mathcal{H}^i = \cup_{t \geq 0} \mathcal{H}^i_t$ is the set of all histories for player i.

When $U = A$ and ℓ is the identity mapping on A, each player fully observes the action profile. When the function ℓ is constant, no player receives information on the action profile. These two cases will be referred to as *perfect monitoring* and *trivial monitoring*, respectively.

A strategy for player i is a mapping $\sigma^i : \mathcal{H}^i \to A^i$. The set of strategies for player i is denoted by Σ^i, and similar conventions are adopted as for actions: $\Sigma = \times_{j \in N} \Sigma^j$, $\Sigma^{-i} = \times_{j \in N, j \neq i} \Sigma^j$. A profile of strategies $\sigma = (\sigma^i)_{i \in N}$ generates a unique history $(a_t(\sigma), u_t(\sigma))_{t \geq 1} \in (A \times U)^\infty$, where, for each t, $u_t(\sigma) = \ell(a_t(\sigma))$. In the whole paper only pure strategies are considered.

Given a strategy profile σ, the average payoff for player i up to time T is $\gamma^i_T(\sigma) = \frac{1}{T} \sum_{t=1}^T g^i(a_t(\sigma))$, and $\gamma^i(\sigma) = \lim_{T \to \infty} \gamma^i_T(\sigma)$, when the limit exists.

Let Γ_∞ be the infinitely repeated game. The next definition recalls the concept of uniform equilibrium.

Definition 2.1. A strategy profile σ is a *uniform equilibrium* of Γ_∞ if

(a) for all $i \in N$, $\gamma^i(\sigma)$ exists.

(b) for all $\epsilon > 0$ there exists T_0 such that for all $T \geq T_0$, for all $i \in N$, for all $\tau^i \in \Sigma^i$, $\gamma^i_T(\tau^i, \sigma^{-i}) \leq \gamma^i_T(\sigma) + \epsilon$.

Denote by E_∞ the set of uniform equilibrium payoffs of Γ_∞, i.e., the set of vectors $(\gamma^i(\sigma))_{i \in N}$, where σ is a uniform equilibrium of Γ_∞.

2.2 Public strategies

Definition 2.2. Let $i \in N$. The strategy $\sigma^i \in \Sigma^i$ is called *public* if for all $t \geq 1$, and for all histories of length t, $h = (a^i_1, u_1, \ldots, a^i_t, u_t)$ and $h' = (b^i_1, v_1, \ldots, b^i_t, v_t)$,

$$(\forall s \in \{1, \ldots, t\},\ u^i_s = v^i_s) \implies \sigma^i(h) = \sigma^i(h').$$

In words a public strategy depends only on public signals. The set of public strategies of player i is denoted by $\widehat{\Sigma}^i$. A strategy profile σ is a *public equilibrium* if it is a uniform equilibrium and each player's strategy is public. The corresponding set of equilibrium payoffs is denoted by $\widehat{E}_\infty$. In the case of perfect monitoring, any strategy is public, since the public history contains all the past.

In repeated games with unbounded recall every pure strategy is equivalent to a public strategy. Knowing her own strategy and the history of public signals, a player can deduce the actions she played in the past (see e.g. Tomala (1998)). More precisely the following lemma holds.

Lemma 2.3. *For every $\sigma^i \in \Sigma^i$, there exists $\widehat{\sigma}^i \in \widehat{\Sigma}^i$ such that for all $\tau^{-i} \in \Sigma^{-i}$ and for each stage t*

$$a_t(\sigma^i, \tau^{-i}) = a_t(\widehat{\sigma}^i, \tau^{-i}).$$

The proof is straightforward: the action played by σ at the first stage depends on σ^i only, therefore the action played at the second stage depends only on σ^i and on the first public signal and so on, by induction.

Corollary 2.4. $\widehat{E}_\infty = E_\infty$.

To emphasize the dependence on the player's own past actions, a strategy that is not public will be called *private*. As it will be seen in the sequel, in games with bounded recall, considering public or private strategies makes a big difference.

2.3 Bounded recall

Consider now players who recall only recent observations. Informally, a strategy has recall k, if the player who uses it remembers only what happened on the k previous stages, and plays in a stationary way, i.e., this player has no clock and relies on her recall, but not on time. The formal definition is the following.

Definition 2.5. Given an integer $k \in \mathbb{N}$, the strategy $\sigma^i \in \Sigma^i$ has *recall* k if there exists a mapping $f : (A^i \times U)^k \to A^i$ such that for all $t \geq k$ and for all histories $h = (a_1^i, u_1, \dots, a_t^i, u_t) \in \mathcal{H}_t^i$

$$\sigma^i(h) = f(a_{t-k+1}^i, u_{t-k+1}, \dots, a_t^i, u_t).$$

By convention, a strategy that has recall 0 is a constant mapping on $\mathcal{H}^i$.

Lehrer (1988, 1992a,b) and Bavly and Neyman (2005) use a somewhat different definition: in those papers, a bounded recall strategy is the choice of an initial recall plus the mapping f. This implies that whenever the initial recall re-appears during the course of the game, the player will play in the same way as at early stages. In the definition given here, a player plays as she wishes before stage k and then uses the stationary rule f. We believe that asymptotic results are unlikely to differ using one or another definition, however for small values of k, the initialization phase might be critical. Also note that Sabourian (1998) uses the same definition as the one given above.

The set of strategies for player i that have recall k is denoted by Σ_k^i and $\Sigma_k := \times_{i \in N} \Sigma_k^i$. Since the game is finite, for each $\sigma \in \Sigma_k$, the sequence $a_t(\sigma)$ is eventually periodic, i.e. periodic from some stage on, which implies the existence of $\gamma^i(\sigma)$. The normal form game $\Gamma_k = \langle N, (\Sigma^i_k), (\gamma^i) \rangle$ is thus well defined and the set of Nash equilibrium payoffs of Γ_k in pure strategies is denoted by E_k.

Let $\widehat{\Sigma}^i_k = \widehat{\Sigma}^i \cap \Sigma^i_k$ be the set of public strategies with recall k, $\widehat{\Gamma}_k = \langle N, (\widehat{\Sigma}^i_k), (\gamma^i) \rangle$ be the public-strategy game with recall k, and $\widehat{E}_k$ be the set of its (pure) Nash equilibrium payoffs.

Remark 2.6. In games with bounded recall, considering public strategies is a true restriction. From Lemma 2.3, every pure strategy σ^i is equivalent to a public strategy $\widehat{\sigma}^i$ but the bounded recall property is not preserved. It might be that σ^i has recall k but $\widehat{\sigma}^i$ does not. For example, consider trivial monitoring (the mapping ℓ is constant). Given any recall k, there is only one history of public signals, thus a public strategy with bounded recall is a constant strategy. By contrast, a private strategy (of recall 1) can simply alternate between two actions. The equivalent public strategy alternates between the two actions according to time and thus is not a public strategy with bounded recall according to Definition 2.5.

Remark 2.7. In the game with recall 0, strategies are constant and thus $\widehat{E}_0 = E_0$. This set further coincides with the set of pure Nash equilibrium payoffs of the stage game.

Remark 2.8. All the equilibrium notions defined in this section might well be empty since we are dealing with pure strategies. However, when the stage game has a pure Nash equilibrium, playing this equilibrium at each stage regardless of history is an equilibrium of the repeated game in any sense defined above: uniform, k-recall public, k-recall private. The rest of the paper deals mainly with the minority game which has pure Nash equilibria.

2.4 The repeated minority game

In the *minority game* (MG) three players have to choose simultaneously one of two rooms: L (left) or R (right). For each profile of action $a = (a^1, a^2, a^3) \in \{L, R\}^3$, call *minority room* the less crowded room and *majority room* the most crowed room. Player i's payoff is then 1 if she chooses the minority room and 0 otherwise. Hence the payoff matrix of the MG is as follows, where player 1 chooses the row, player 2 the column, and player 3 the matrix.

	L	R
L	$0,0,0$	$0,1,0$
R	$1,0,0$	$0,0,1$

L

	L	R
L	$0,0,1$	$1,0,0$
R	$0,1,0$	$0,0,0$

R

The profile where one player chooses L and the two other players choose R is a Nash equilibrium. All pure Nash equilibria of this game are obtained by permutation of players and rooms. Denote by C be the convex hull of payoff vectors generated by these equilibria. If $e(i) \in \mathbb{R}^3$ is the vector whose i-th component is 1 and the other components are 0, then

$$C = \operatorname{conv}\{e(i) : i \in \{1,2,3\}\} = \left\{ x \in [0,1]^3 : \sum_{i=1}^{3} x^i = 1 \right\}.$$

It is worth noticing that this is also the set of Pareto-efficient payoffs in the game.

Consider now the repeated game where the majority room is publicly observed. At each stage $t = 1, 2, \ldots$, players choose their room and before stage $t+1$, the majority room is publicly announced: $U = \{L, R\}$, and

$$\ell(a) = \begin{cases} L & \text{if } \#\{i : a^i = L\} \geq 2, \\ R & \text{if } \#\{i : a^i = R\} \geq 2. \end{cases}$$

The rest of the paper deals with the repeated minority game with these public signals. The following Folk-theorem-like result holds.

Proposition 2.9. *In the minority game $E_\infty = C$.*

3 Main results

3.1 Public equilibria and de Bruijn graphs

We give here a combinatorial representation of k-recall strategies using de Bruijn graphs. We consider a directed graph T_k, where each of the 2^k nodes is labeled by a k-letter word written with the alphabet $\{L, R\}$. For $i \in \{1, \ldots, k\}$ let $x_i \in \{L, R\}$. The word $x = (x_1, \ldots, x_k)$ precedes the word $y = (y_1, \ldots, y_k)$ if $(x_2, \ldots, x_k) = (y_1, \ldots, y_{k-1})$. The word y succeeds x whenever x precedes y. Hence each node (i. e. the word associated to it) precedes only two nodes. Such a graph is called de Bruijn graph (see e.g. de Bruijn (1946) and Yoeli (1962) for some properties of these graphs). The following figure shows a de Bruijn graph T_3 based on sequences written with the alphabet $\{L, R\}$.

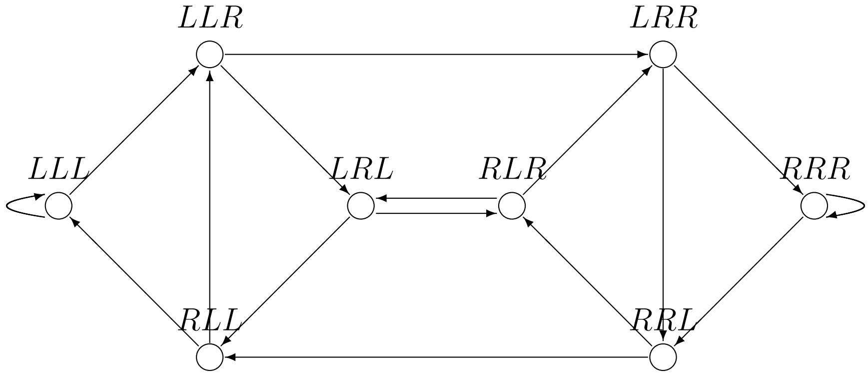

Figure 1. de Bruijn graph T_3

A proof of the following result can be found in Yoeli (1962)(see Lempel (1971) for a generalization to any finite alphabet).

Proposition 3.1. *For every p in $\{1,\ldots,2^k\}$, there exists in the de Bruijn graph T_k a cycle with length p.*

The link with public strategies is the following. Let $\sigma = (\sigma^1, \sigma^2, \sigma^3)$ be a k-recall strategy profile, $(a_t(\sigma))_t$ the induced sequence of action profiles and $(u_t(\sigma))_t$ the induced sequence of public signals. We denote by $x_{t-1} = (u_{t-k}(\sigma), \ldots, u_{t-1}(\sigma))$ the *public memory* before stage t. Let $f^i : \{L,R\}^k \to \{L,R\}$ be the mapping associated to σ^i and set $f = (f^1, f^2, f^3)$. The mapping f associates to every public memory $x \in \{L,R\}^k$ the next action profile. From stage k on, the play of the game is perfectly determined by f, that is $f(x_{t-1}) = a_t(\sigma)$ for each $t > k$.

The sequence $(x_t)_t$ is eventually periodic: there exist two integers t_0 and p such that $x_{t+p} = x_t$, $\forall t \geq t_0$. The payoff associated to σ is thus the average payoff over a period: $\gamma^i(\sigma) = \frac{1}{p}\sum_{t=t_0+1}^{t_0+p} g^i(f(x_t))$. Let us call a *cycle* of σ a tuple $(x_{t+1}, \ldots, x_{t+p})$ with $t \geq t_0$.

Lemma 3.2. *If σ is a public equilibrium of Γ_k, then for each x in the cycle of σ, $f(x)$ is a Nash equilibrium of the one-shot game.*

Proof. Otherwise there is an x in the cycle of σ such that $f(x) = (L,L,L)$ (or $= (R,R,R)$). Then player 1 deviates and plays R (or L) whenever the public memory is x and plays like σ^1 otherwise. This deviation does not affect the sequence of public signals and thus does not affect the behavior of other players. It is profitable since at least once every p stages, player collects a payoff of 1 instead of 0. □

Thanks to this lemma, we can restrict our attention to mappings f that map public memories (i.e. $\{L,R\}^k$) to Nash equilibria of the one-shot game. Notice now that a Nash equilibrium of the minority game is fully described by

(i) the player who gets 1 and

(ii) the majority room.

That is, to specify the mapping f, we must attach to each public memory

(i) a winning player and

(ii) the next public signal.

We can thus describe a strategy profile in the de Bruijn graphs by selecting one outgoing edge for each node and by coloring the nodes: each node is assigned to a player, or to nobody if the players are all in the same room. As we said before, in equilibrium every node is assigned to a player.

Note that a node is assigned to player i when she is the winning player. So if she changes action at this node, first she gets a bad payoff, and second, she does not change the public signal. A deviation of player i can thus be regarded as an alternative choice of an outgoing edge at each node that is not assigned to her.

To sum up, a public equilibrium in the k-recall game can be described as follows:

- for each node of T_k, one outgoing edge and one player ar chosen in such a way that
- no player i can induce a more profitable cycle in the graph by changing outgoing edges at nodes not assigned to her.

3.2 Some public equilibria

We describe now some public equilibrium payoffs.

Lemma 3.3. (a) *For any* $k \geq 0$, $\left(\frac{k}{k+1}, \frac{1}{k+1}, 0\right) \in \widehat{E}_k$.

(b) *For any* $k \geq 2$, $\left(\frac{k-2}{k}, \frac{1}{k}, \frac{1}{k}\right) \in \widehat{E}_k$.

(c) *For any* $k \geq 2$, $\left(\frac{k-2}{k}, \frac{2}{k}, 0\right) \in \widehat{E}_k$.

As shown by the proofs of these results, the representation of strategies in de Bruijn graphs is fundamental to determine bounded-recall equilibria in the minority game. The constructions will identify in the graphs cycles that represent the equilibrium play, and will correctly assign a player to each node.

Remark 3.4. Define the *effective recall* of a strategy as the smallest k for which this strategy has recall k. In our equilibrium constructions, the effective recall of the three players are different. For instance when $k = 3$, in the equilibrium of Lemma 3.3(a) the effective recalls of the three players are 0, 0, and 3, respectively. In fact player 1 always plays L, player 2 always plays R. In (b) the recalls are 1, 3, and 3, and in (c) they are 1, 2, and 3. The following question, raised by an anonymous referee, remains open: does there exist an equilibrium payoff in $\widehat{E}_k$, such that in every equilibrium yelding this payoff with recall at most k and public strategies, all the strategies of the players have effective recall k?

The partial results of Lemma 3.3 enable us to completely describe the set of public equilibrium payoffs for small values of k.

Proposition 3.5. (a) $\widehat{E}_0 = \{(1,0,0),(0,1,0),(0,0,1)\}$.

(b) $\widehat{E}_1 = \widehat{E}_0 \cup \left\{\left(\frac{1}{2},\frac{1}{2},0\right),\left(\frac{1}{2},0,\frac{1}{2}\right),\left(0,\frac{1}{2},\frac{1}{2}\right)\right\}$.

(c) $\widehat{E}_2 = \widehat{E}_1 \cup \left\{\left(\frac{1}{3},\frac{2}{3},0\right),\left(\frac{1}{3},0,\frac{2}{3}\right),\left(0,\frac{1}{3},\frac{2}{3}\right),\left(\frac{2}{3},\frac{1}{3},0\right),\left(\frac{2}{3},0,\frac{1}{3}\right),\left(0,\frac{2}{3},\frac{1}{3}\right)\right\}$.

Note that for $k \leq 2$, all public equilibrium payoffs are on the boundary of the triangle C. A direct consequence of Lemma 3.3 is that $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3}\right) \in \widehat{E}_3$, so when the recall is $k \geq 3$ there exists a public equilibrium payoff in the interior of C.

3.3 Convergence of $\widehat{E}_k$

We use the standard notion of Hausdorff convergence of closed sets and get the following convergence result.

Theorem 3.6. *In the minority game* $\lim_{k\to+\infty} \widehat{E}_k = E_\infty = C$, *that is for every* $\varepsilon > 0$, *there exists* k_0 *such that for each* x *in* C *and each* $k \geq k_0$, *there exists* $y \in E_k$ *such that* $\|x-y\| \leq \varepsilon$.

The construction, like standard Folk-theorems, uses a main path and punishments. Players agree on a cycle over the set of stage-Nash equilibria leading approximately to the target payoff. Since only stage-equilibria are played, a deviation that does not modify the signals is not profitable. When players see unexpected signals, they punish the deviator by staying for a long time in the same room where they were at the deviation stage. The punishment is effective since only a player who gets a zero payoff (i.e., is not alone in a room) can modify the signal. Before the deviation signal leaves the public recall, players re-write it in the recall. Two players can do so by playing the same action thus controlling the public signal. The detailed construction is given in Section 5.

Remark 3.7. Theorem 3.6 easily extends to a $(2n+1)$-player minority game (each player has to choose between L and R and receives a payoff of 1 if she is in the minority room and zero otherwise). However, the proof heavily relies on the specific properties of the game and signal function. Since convergence of $\widehat{E}_k$ to E_∞ is not always guaranteed, a challenging and open problem is to characterize $\lim_k \widehat{E}_k$.

3.4 Private equilibria

The following proposition shows that under bounded recall, the set of private equilibrium payoffs is strictly larger than the set of public equilibrium payoffs. Furthermore, private equilibria may not be equilibria of the unbounded-recall, i.e. uniform equilibria, and we find a private equilibrium payoff for $k = 3$ which lies outside E_∞.

Proposition 3.8. (a) *E_2 differs from $\widehat{E}_2$ since* $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \in E_2 \setminus \widehat{E}_2$.

(b) *In the minority game* $(3/7, 3/7, 0) \in E_3$ *and thus* $E_3 \not\subset E_\infty$.

This last point is proved by constructing explicitly an equilibrium $\sigma = (\sigma^1, \sigma^2, \sigma^3)$ of Γ_3 with payoff $(3/7, 3/7, 0)$. The proof is quite lengthy and involved and seems to indicate that more general results in this direction are quite hard to obtain.

As mentioned in the introduction, Mailath et al. (2002) and Kandori and Obara (2006) compare the behavior of public and private strategies in games with public signals and unbounded recall. Proposition 3.8 does something of that sort in a bounded recall framework.

4 Beyond the minority game

The main message of the paper is that, for the game under study, equilibria are quite hard to analyze and even asymptotic results requires a lot of structure and involved constructions, compared to regular Folk theorems. We conclude with two examples of other games showing that such phenomena are bound to appear in many cases.

Let us first consider a game with perfect monitoring. In this case, public and private equilibria coincide. The following example shows that even under perfect monitoring, convergence of $(\widehat{E}_k)_k$ to E_∞ may fail.

There are three players, $N = \{1, 2, 3\}$. Player 3 only has one action, player 1 chooses the line, player 2 chooses the column, and as usual the first (resp. second, resp. third) coordinate of the vector payoff corresponds to the first (resp. second, resp. third) player.

	L	R
T	$\sqrt{2}, -\sqrt{2}, 1$	$-1, 1, 1$
B	$0, 0, 0$	$0, 0, 0$

Players with unbounded recall get $(0, 0, 1)$ in (uniform) equilibrium: player 1 plays T and player 2 alternates between L and R with the correct frequencies. Since $\sqrt{2}$ is irrational, these frequencies must also be irrational but players with bounded recall eventually enter a cycle on a finite number of actions, generating rational frequencies.

Proposition 4.1. (a) *The payoff* $(0, 0, 1) \in E_\infty$.

(b) *The set* $\widehat{E}_k = E_k = \{(0, 0, 0)\}$.

We know from the study of the minority game that private bounded recall equilibria may not be equilibria of the unrestricted game. We strengthen this result by giving an example where there is a payoff in $\cap_{k \geq 1} E_k \setminus E_\infty$, i.e. there exists a bounded recall equilibrium payoff bounded away from E_∞ as k grows.

Consider the following two-player game, with $A^1 = \{T, M, B_1, B_2\}$, $A^2 = \{L, R\}$ and $U = \{u, v\}$. The payoffs and the public signals are indicated below.

	L	R
T	$0, 0,$	$0, 0$
M	$2, 2$	$0, 3$
B_1	$2, 1$	$0, 0$
B_2	$2, 1$	$0, 0$

payoffs

	L	R
T	u	u
M	u	u
B_1	u	v
B_2	u	v

signals

Proposition 4.2. *In the above game, for each $k \geq 1$, the payoff $(2, 4/3) \in E_k$, but $(2, 4/3) \notin E_\infty$.*

5 Proofs

The following general lemma will be used in the sequel.

Lemma 5.1. (a) *If a strategy σ is an equilibrium of $\widehat{\Gamma}_k$, then σ is a uniform equilibrium of $\widehat{\Gamma}_\infty$. Thus, $\widehat{E}_k \subset \widehat{E}_\infty$.*

(b) *If a strategy σ is an equilibrium of $\widehat{\Gamma}_k$, then σ is an equilibrium of Γ_k. Thus, $\widehat{E}_k \subset E_k$.*

(c) *If a strategy σ is an equilibrium of $\widehat{\Gamma}_k$, then σ is an equilibrium of $\widehat{\Gamma}_{k+1}$. Thus, $\widehat{E}_k \subset \widehat{E}_{k+1}$.*

Proof. (a) This kind of result is common in the literature on games with bounded complexity (see e.g., Neyman (1998), Ben-Porath (1993), Lehrer (1988, 1994)) and relies on a usual dynamic programming argument. Let σ be an equilibrium of $\widehat{\Gamma}_k$. For each player i, finding a best reply in Σ^i to σ^{-i} amounts to solving a dynamic programming problem, where the state space is U^k, the set of public histories of length k, the action space is A^i, the payoff in state $h = (u_1, \ldots, u_k)$, given action a^i is $g^i(a^i, \sigma^{-i}(h))$, and the new state is $(u_2, \ldots, u_k, \ell(a^i, \sigma^{-i}(h)))$. It is well known (see Blackwell (1962)) that there exists a stationary optimal strategy. Thus, the best reply of player i to a profile of public strategies with recall k is a public strategy with recall k (see Abreu and Rubinstein (1988, Lemma 1)). Therefore σ is a uniform equilibrium of Γ_∞.

(b) This follows directly from the previous point. The game $\widehat{\Gamma}_k$ is a *subgame* of Γ_k in the sense that the set of strategies of each player in $\widehat{\Gamma}_k$ is a subset of the set of strategies of this player in Γ_k. Let then σ be a strategy profile in $\widehat{\Gamma}_k$, if σ is not an equilibrium of Γ_k, then a player i has a profitable deviation in $\Sigma^i_k \subset \Sigma^i$, thus σ is not a uniform equilibrium contradicting the previous point.

(c) The argument is similar to the one used for point (b), $\widehat{\Gamma}_k$ is a subgame of $\widehat{\Gamma}_{k+1}$: any strategy with recall k can be played in the game with recall $k+1$. So, if a strategy profile σ in $\widehat{\Gamma}_k$ is not an equilibrium of $\widehat{\Gamma}_{k+1}$, then some player i has a profitable deviation in $\widehat{\Sigma}^i_{k+1} \subset \Sigma^i$, thus σ is not a uniform equilibrium contradicting point (a). □

Proof of Proposition 2.9. This follows directly from the characterization given in Tomala (1998, Theorem 5.1, page 104), but we provide a simple direct proof. First note that, since C is the convex hull of Nash equilibrium payoffs of the one-shot game, then $C \subset E_\infty$. Given any point x in C, one can find a sequence of Nash equilibria $(a_t)_t$ of the minority game, such that the average payoff vector along this sequence converges to x. Then, the strategy profile such that for each player i and stage t, player i plays a^i_t at stage t, irrespective of the history, is clearly a uniform equilibrium with payoff x.

To get the converse, note that there are two types of action profiles: either two players are in the same room and the profile is an equilibrium of the MG, or the three players are in the same room. In the latter case, each player has a profitable deviation (she prefers to switch room) and further this deviation does not change the majority room, i.e., the public signal. If at a strategy profile the three players are in the same room on a non-negligible set of stages, then player 1 can switch rooms at these stages. This increases her payoff at these stages without affecting public signals, hence without affecting the behavior of the other players. Such a strategy profile cannot be a uniform equilibrium and therefore $E_\infty \subset C$. □

The following notation and terminology will be used in the sequel.

$$L^p = \underbrace{L \cdots L}_{p \text{ times}}, \quad R^q = \underbrace{R \cdots R}_{q \text{ times}}.$$

Call *word* any finite sequence of signals. Given two words $u = (u_1, \ldots, u_p)$ and $v = (v_1, \ldots, v_q)$, denote by uv the concatenated word $uv = (u_1, \ldots, u_p, v_1, \ldots, v_q)$.

Consider the minority game with recall k, and its associated de Bruijn graph T_k. Call m-cycle a cycle of length m, and call *stable* the cycles where all the nodes have the same number of L's. Among the stable cycles having s L's, say, call *main* all the cycles containing the nodes $L^s R^{k-s}$ or $R^{k-s} L^s$. In T_k there are $k-1$ main k-cycles and 2 main 1-cycles.

Proof of Lemma 3.3. (a) Consider the $(k+1)$-cycle that contains R^k and all the nodes whose label contains just one L. In equilibrium, players cycle on this $(k+1)$-cycle and elsewhere they go to this cycle as fast as possible.

Assign node R^k to player 2 and all the other nodes in the graph to player 1.

Player 1 can deviate only on R^k, and she has no incentive to do it, because that would induce a cycle on the node R^k, that is assigned to player 2. Player 2 can deviate anywhere else, but she has no incentive to do it, since she cannot find a cycle that contains R^k and is shorter than the equilibrium cycle.

Player 3 can always deviate, but, since she would get a zero payoff anyway, she has no incentive to deviate.

The following figure shows the above equilibrium for $k = 3$.

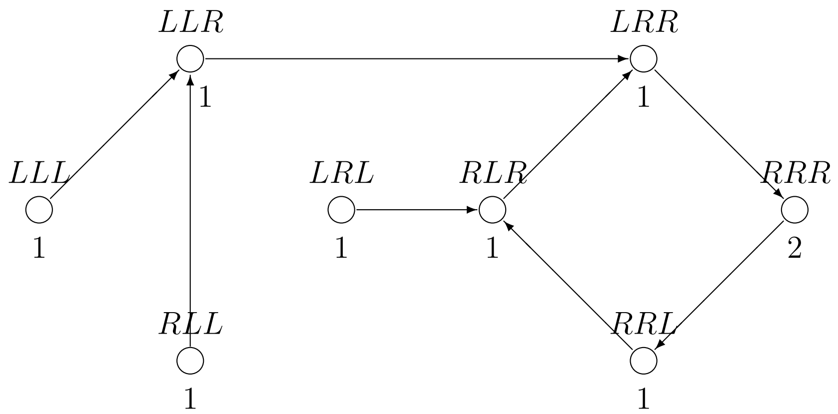

Figure 2. Equilibrium with 3-recall and payoff $\left(\frac{3}{4}, \frac{1}{4}, 0\right)$

(b) For every $s \in \{1, \ldots, k\}$ assign nodes $R^{k-s}L^s$ to player 2, nodes $L^{k-s}R^s$ to player 3, and the other nodes to player 1. In equilibrium players cycle on the main k-cycles and elsewhere they move to the closest main cycle.

Assume for instance that we start with the memory L^k. The closest main k-cycle is the cycle of nodes that contain only one R, i.e. the cycle

$$L^{k-1}R, L^{k-2}RL, L^{k-3}RL^2, \ldots, LRL^{k-2}, RL^{k-1}, L^{k-1}R, \ldots$$

Remark that all nodes are assigned to player 1 except $L^{k-1}R$ and RL^{k-1}. Assume that player 1 deviates at node $L^{k-1}R$. The next node is $L^{k-2}R^2$ which is not assigned to her. If she deviates again, the next node is still not assigned to her, and so on. Thus, her only possibility to collect payoffs is to stop deviating and follow the equilibrium. Indeed, when player 1 is at a node not assigned to her, under the equilibrium strategy she will be winning at the next $k-2$ nodes whereas if she deviates she will spend more time in nodes where she gets 0.

More generally, consider a node where player 1 could possibly deviate, namely the nodes assigned either to player 2 or to player 3. One can check that

1. any deviation in a node not assigned to player 1 leads to another node not assigned to player 1,
2. the shortest path from that node to the closest node assigned to player 1 is via an equilibrium path, the shortest path from that node to the second closest node assigned to player 1 is via an equilibrium path, and so on.

Therefore any non-equilibrium cycle that is forced by player 1 with a finite sequence of deviations is longer than k and the proportion of nodes in this cycle assigned to player 1 cannot be larger than $(k-2)/k$. Thus there is no finite sequence of deviations that would make player 1 better off.

For instance, if $k = 3$, deviating in LLL (resp. RRR) would force the 1-cycle $LLL\cdots$ (resp.$RRR\cdots$). Deviating in LLR (resp. RRL) would increase the distance to the next 1-node from 1 to at least 3, hence the deviation would be profitable only if it induced a 5-cycle with two nodes assigned to player 1, but this is not possible since player 1 cannot deviate on her own nodes. Deviating in RLL (resp. LRR) would increase the distance to the next 1-node from 2 to at least 3. Using the same argument as before, we can see that this deviation is not profitable.

The argument is similar for the other players. Consider now a node where player 2 could possibly deviate, namely the nodes assigned either to player 1 or to player 3. It is not difficult to verify that

1. any deviation in a node not assigned to player 2 leads to another node not assigned to player 2,

2. the shortest path from that node to the closest node assigned to player 2 is via an equilibrium path, the shortest path from that node to the second closest node assigned to player 2 is via an equilibrium path, and so on.

Therefore any non-equilibrium cycle that is forced by player 2 with a finite sequence of deviations is longer than k and the proportion of nodes in this cycle assigned to player 2 cannot be larger than $1/k$.

By symmetry the argument for player 3 is the same.

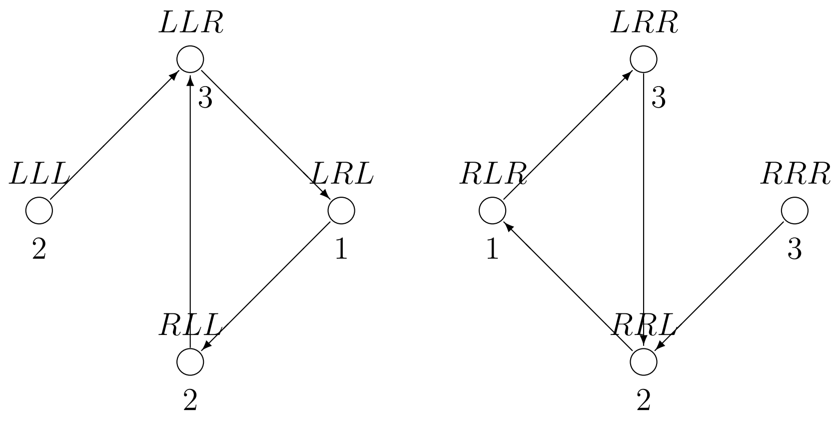

Figure 3. Equilibrium with 3-recall and payoff $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$

(c) As above. Just assign to player 2 the nodes that were assigned to 3, and repeat the argument.

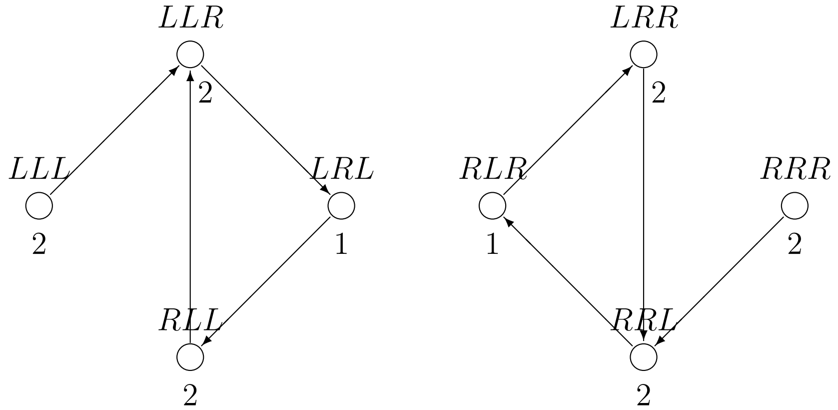

Figure 4. Equilibrium with 3-recall and payoff $\left(\frac{1}{3}, \frac{2}{3}, 0\right)$

$\square$

Proof of Proposition 3.5. (a) By Remark 2.7, the only possible Nash equilibria of $\widehat{E}_0$ are repetitions of the same Nash equilibrium of the stage game.

(b) By Lemma 5.1(c), $\widehat{E}_0 \subset \widehat{E}_1$. Furthermore by Lemma 3.3(a) the payoffs $\left(\frac{1}{2}, \frac{1}{2}, 0\right)$, $\left(\frac{1}{2}, 0, \frac{1}{2}\right)$, $\left(0, \frac{1}{2}, \frac{1}{2}\right) \in \widehat{E}_1$. No other equilibrium payoff can be obtained with recall 1, since the maximal length of a cycle in the de Bruijn graph T_1 is 2.

(c) By Lemma 5.1(c), $\widehat{E}_1 \subset \widehat{E}_2$. Furthermore by Lemma 3.3(a) the payoffs $\left(\frac{1}{3}, \frac{2}{3}, 0\right)$, $\left(\frac{1}{3}, 0, \frac{2}{3}\right)$, $\left(0, \frac{1}{3}, \frac{2}{3}\right)$, $\left(\frac{2}{3}, \frac{1}{3}, 0\right)$, $\left(\frac{2}{3}, 0, \frac{1}{3}\right)$, $\left(0, \frac{2}{3}, \frac{1}{3}\right) \in \widehat{E}_2$.

No other equilibrium payoff can be obtained with recall 2.

First we prove that $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) \notin \widehat{E}_2$. In fact the maximal length of a cycle in the de Bruijn graph T_2 is 4. Hence, in order to obtain such a payoff in equilibrium, the players would have to cycle on a 3-cycle of T_2, and each node should be assigned to a different player. There are only two such cycles. Take for instance the cycle $LL \to LR \to RL$, and assume that these nodes are assigned to players 1, 2, and 3, respectively. Then player 2 deviating in RL induces the cycle $LR \to RL \to LR \dots$ and gets a payoff of $\frac{1}{2}$. An analogous argument can be used for the cycle $LR \to RR \to RL$.

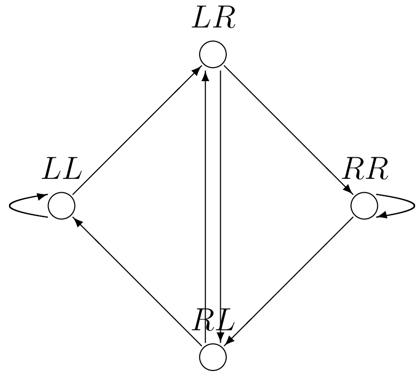

Figure 5. de Bruijn graph T_2

We claim now that the payoff $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$ (or its permutations) cannot be obtained at equilibrium. Consider the only 4-cycle in the graph T_2, namely, $LL \to LR \to RR \to RL$. This cycle cannot give a payoff $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$ or its permutations. In fact if (i_1, i_2, i_3, i_4) denotes the strategy profile that assigns LL to player i_1, LR to player i_2, etc., then none of the configurations that give a payoff $\left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right)$ is an equilibrium:

- $(1, 1, 2, 3)$ is not an equilibrium, since player 3 would deviate in LR,
- $(1, 1, 3, 2)$ is not an equilibrium, since player 2 would deviate in LR,
- $(1, 2, 1, 3)$ is not an equilibrium, since player 3 would deviate in LR, and player 2 would deviate in RL,
- $(1, 3, 1, 2)$ is not an equilibrium, since player 2 would deviate in LR, and player 3 would deviate in RL,
- $(1, 2, 3, 1)$ is not an equilibrium, since player 1 would deviate in LR,
- $(1, 3, 2, 1)$ is not an equilibrium, since player 1 would deviate in LR,
- $(2, 1, 1, 3)$ is not an equilibrium, since player 1 would deviate in RL,
- $(3, 1, 1, 2)$ is not an equilibrium, since player 1 would deviate in RL,
- $(2, 1, 3, 1)$ is not an equilibrium, since player 2 would deviate in LR, and player 3 would deviate in RL,
- $(3, 1, 2, 1)$ is not an equilibrium, since player 3 would deviate in LR, and player 2 would deviate in RL,
- $(3, 2, 1, 1)$ is not an equilibrium, since player 2 would deviate in RL,
- $(2, 3, 1, 1)$ is not an equilibrium, since player 3 would deviate in RL.

The above cycle cannot give the payoff $\left(\frac{3}{4}, \frac{1}{4}, 0\right)$ or its permutations, either. Using the same notation as before

- $(1,1,1,2)$ is not an equilibrium, since player 2 would deviate in LR,
- $(1,1,2,1)$ is not an equilibrium, since player 2 would deviate in RL,
- $(1,2,1,1)$ is not an equilibrium, since player 2 would deviate in RL,
- $(2,1,1,1)$ is not an equilibrium, since player 2 would deviate in LR.

□

Proof of Theorem 3.6. Let $m \geq 2$ be an integer and let C_m be the set of vectors of $x \in C$ with rational components of the form $x^i = m^i/m$ with $m^i \geq 2$ integers. Then C_m converges to C as m goes to infinity i.e. $\sup_{x \in C} \inf_{y \in C_m} \|x - y\|$ goes to 0 as m goes to infinity. Therefore Theorem 3.6 follows from Lemma 5.1(c) and from Lemma 5.2 below. □

Lemma 5.2. *For every integers $m \geq 2$ and $K \geq 2m$, $C_m \subset \widehat{E}_k$ for $k = Km$.*

The following terminology will be used in the proof of Lemma 5.2. A word of length k is called a *public recall.* Given a public recall M, a word u of length $l \leq k$ is called a *sub-word* of M if there exist two words v, w (possibly of length 0) such that $M = vuw$. The word consisting of $L \dots L$, q times is denoted L^q. If u is a sub-word of M, define the *position* of u in M as the rank of the first letter of u. For instance, if M begins with u, then u has position 1; if M ends with u, then u has position $k - l + 1$.

Proof of Lemma 5.2. Let $m \geq 2$ be an integer and $x \in C_m$. The aim is to construct a strategy profile σ with payoff x which is an equilibrium of $\widehat{\Gamma}_k$ for $k = Km$, with $K \geq 2m$. The strategy construction is in a folk-theorem spirit. First the right payoff is obtained by playing an adapted main path. In case of a detected deviation, punishments have to be performed. Because of finite recall, the evidence that a deviation occurred may disappear from the recall. To get a deviating player to be punished forever, players are asked to rewrite periodically a word in the public recall, indicating that a deviation has occurred and which actions should be used to punish. This construction relies heavily on properties of the minority game and the majority room as a signal. The following properties will be used extensively.

- A player who is in the minority room at some stage cannot change the signal at that stage. This implies that a player who gets a payoff of 1 at a given stage has no incentive to deviate at that stage since it can only decrease the stage-payoff and has no impact whatsoever on the future.

- The main path is constructed so that at each stage a Nash equilibrium of the one-shot game is played. Thus at each stage there is one player in the minority room and the other two players are in the majority room, both receiving a payoff of zero. If the signal changes, that means that one of the two players in the majority room deviated, but the public signal does not tell who did. A simple way to punish the deviating player without knowing her identity is to apply the following policy: "If I see a wrong signal at stage t, then I remain in the room where I was at stage t." This insures that the deviating player, who was in the majority room when the deviation was detected, remains in the majority room as long as the punishment phase lasts.

- Any payoff vector can be obtained by two action profiles giving different public signals (just exchange L and R).

- Two players can write any word in the public recall, whatever the behavior of the third player is.

Pick now a point $x = (x^i)_i \in C_m$. Then $x = \sum_i x^i e(i)$, where for each $i \in N$, $x^i = m^i/m$, with $m^i \geq 2$, so $x^i \geq 2/m$. Let $H = (a_1^*, \ldots, a_m^*) \in A^m$ be a sequence of action profiles of length m such that

1. the average payoff along H is
$$x = \frac{1}{m} \sum_{t=1}^{m} g(a_t^*),$$

2. the public history $(\ell(a_1^*), \ldots, \ell(a_m^*))$ associated to H is L^m.

Such a sequence exists: it suffices to play a sequence of Nash equilibria of the MG such that player i gains 1 exactly m^i times and the majority room is always L. For each room $r \in \{L, R\}$, let $\bar{r}$ be the other room, and, if a is an action profile, let $\bar{a}$ be the action profile where every player has switched room. Let $\bar{H} \in A^m$ be the sequence obtained from H by switching rooms: $\bar{H} = (\bar{a}_1^*, \ldots, \bar{a}_m^*)$. The main path will be the periodic repetition of the sequence $H\bar{H}$. Here is how to construct a profile of strategies of recall k that generates this periodic sequence of action profiles.

Let $W := L^m$ be the word induced by H. A word w is a sub-word of W if $w = L^q$ with $0 \leq q \leq m$. If a periodic repetition of $H\bar{H}$ is played, at each stage the public recall ends by a word of the type $\bar{W}w$ or $W\bar{w}$ with w sub-word of W (possibly of length 0). Call such words *end-words*. An end-word writes either $L^m R^q$ or $R^m L^q$, $0 \leq q < m$. The aim is to play a periodic repetition of $H\bar{H}$. In order to do that, at each stage knowledge of the end-word is sufficient to know what action profile should be played at the next stage. Thus, letting E be the set of end-words, there exists a mapping f which maps E to pure Nash equilibria of the MG and such that for each end-word e, $f(e) = (f^i(e))_{i\in N}$ is the action profile that follows e in the periodic repetition of $H\bar{H}$.

Consider now deviations. After each end-word e, $f(e)$ should be played. On the main path $f(e)$ induces a winning player $i(e)$ and a signal $r(e)$. If $\bar{r}(e)$ is observed, then some player $j \neq i(e)$ has deviated. Let us call *deviation-word*, a word of the type $e\bar{r}(e)$: a deviation word writes either $L^m R^q L$ or $R^m L^q R$, $0 \leq q < m$. If a deviation-word $e\bar{r}(e)$ appears in the recall, the strategy prescribes to keep playing $f(e)$ as long as the position of $e\bar{r}(e)$ is greater than $2m$. During this punishing phase the signal is completely controlled by the punished player, hence this player could write in the recall *another* deviation-word $e'\bar{r}(e')$. To prevent other end-words to appear in the recall, if L^{m-1} (resp. R^{m-1}) appears, all players must play R (resp. L). Finally, when the position of $e\bar{r}(e)$ becomes less than or equal to $2m$, the players must rewrite this word in the recall by all playing the same actions for an appropriate number of times.

The exact definition of the strategy profile σ is given now.

- **Initialization**. At the first m stages each players plays L. For the next m stages, each player plays R, for the next m stages each player plays L, and so on until stage k.
- **Main path**. If the recall contains no deviation-word and ends by the end-word e, each player i plays $f^i(e)$.
- **Early punishments**.
  - If the recall contains a deviation-word $e\bar{r}(e)$ whose position is greater than $2m$, and if the recall does not end by L^{m-1} or by R^{m-1}, then each player i plays $f^i(e)$.
  - If the recall contains a deviation-word $e\bar{r}(e)$ whose position is greater than $2m$, and if the recall ends by L^{m-1}, then each player i plays R.
  - If the recall contains a deviation-word $e\bar{r}(e)$ whose position is greater than $2m$, and if the recall ends by R^{m-1}, then each player i plays L.
- **Late punishments**. If the recall contains a deviation word $e\bar{r}(e) = L^m R^q L$ with $0 \leq q < m$, let p be its position.
  - If $m < p \leq 2m$, then each player i plays L.
  - If $m - q < p \leq m$, then each player i plays R.
  - If $p = m - q$, then each player i plays L.

  And similarly for $e\bar{r}(e) = R^m L^q R$.
- **Other memories**. For all other memories, each player plays L.

It remains to prove that the above-defined strategy profile σ has payoff x and is an equilibrium of $\widehat{\Gamma}_k$.

If all players play this strategy, the public recall after stage k is either $L^m R^m \dots L^m R^m$ or $L^m R^m \dots L^m R^m L^m$ depending on the parity of K. It ends by an end-word e and contains no deviation word. The next action profile is then $f(e)$ and the public recall still ends by an end-word so the strategy uses f again. By construction of f, this strategy profile generates the periodic repetition of $H\bar{H}$ and the payoff is indeed x.

Suppose that player i deviates. First, player i cannot modify the signals in the initialization phase, and, since this phase is transient, it is irrelevant for payoffs. We consider thus deviations at later stages.

If the deviation never changes the signals, then player i changes action only at stages where she was in the minority room. Therefore she loses payoff at these stages and does not affect the behavior of other players. Such a deviation is thus not profitable.

Suppose now that player i changes the signal at some stage, therefore i is in the majority room at this stage. This generates a deviation-word $e\bar{r}(e)$. As long as the position of $e\bar{r}(e)$ is greater than $2m$, the other players play $f(e)$ so player i receives a payoff of zero, except if she generates words of the type L^{m-1} or R^{m-1}. In such cases, the other players will play both R or both L. Such situations appear at most every m stages. So, the only opportunities to player i to gain a payoff of 1 are when other players rewrite the deviations word (at most $2m$ stages), and once every m stages for $k-2m$ stages. The average payoff for player i is thus no more than

$$\begin{aligned}\frac{2m+\dfrac{k-2m}{m}}{k} &= \frac{2m+K-2}{Km} \\ &\leq \frac{2}{K}+\frac{1}{m} \\ &\leq x^i,\end{aligned}$$

since $x^i \geq 2/m$, and $K \geq 2m$. □

Proof of Proposition 3.8. (a) First remark that with private strategies player i can play a periodic sequence of actions with cycle RLL by using a strategy that relies on her own actions only, and does not regard public signals whatsoever. Therefore consider the strategy profile obtained by cycling:

$$\begin{matrix} R & L & L \\ L & R & L \\ L & L & R \end{matrix}$$

where the i-th row indicates the strategy of the i-th player. This is clearly an equilibrium of Γ_2: it is a repetition of one-stage Nash equilibria so no player can increase her stage payoff by deviating and further, since player do not regard public signals, no player can change the future behavior of her opponents. The associated payoff is then $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$. By Proposition 3.5(c) the payoff $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ is not in $\widehat{E}_2$.

(b) We construct an equilibrium $\sigma = (\sigma^1, \sigma^2, \sigma^3)$ of Γ_3 with payoff $(3/7, 3/7, 0)$. Given strategies of recall 3, the action played by a player at some stage depends only on her last 3 actions and on the last 3 public signals. The last $3 \wedge t$ actions or signals at time t will be called *available*.

The profile σ is defined as follows:

(a) If at least one available public signal is R, then σ recommends to each player to switch room, i.e., to play L if she played R at the previous stage, and vice-versa.

(b) Assume now that all available public signals are L.

(b1) Regarding the first three stages, as long as the public signal is L, σ recommends to play as follows:

stage $\rightarrow$	1	2	3
P1	L	L	R
P2	R	L	L
P3	L	R	L

For example, the symbol R in line P3 means that at stage 2, σ^3 asks player 3 to play R if the public signal of stage 1 was L.

(b2) At every stage $t \geq 3$, if the last 3 public signals are L, then each player $i \in \{1, 2, 3\}$ plays the action $f^i(a^i_{t-3}, a^i_{t-2}, a^i_{t-1}) \in \{L, R\}$ where $a^i_{t'}$ denotes the action played by player i at stage t' and the functions f^1, f^2, f^3 are described below.

last own actions	P1	P2	P3
LLL	R	R	L
LLR	R	L	L
LRL	L	R	L
LRR	L	L	L
RLL	L	L	L
RLR	L	R	L
RRL	R	L	L
RRR	L	L	L

Figure 6

At the intersection of column P2 and line RLL, the symbol L means that $f^2(RLL) = L$, i.e., at any stage $t \geq 3$, if the last 3 public signals were L, and the last actions played by player 2 were R (at stage $t-3$), L (at stage $t-2$), and L (at stage $t-1$), then player 2 following σ^2 should play L. This ends the definition of σ.

The proof is complete once Lemma 5.3 below is proved. □

Lemma 5.3. (a) *The payoff induced by* σ *is* $(3/7, 3/7, 0)$.

(b) *The strategy* σ *is an equilibrium of* Γ_3.

Proof. (a) Assume that σ is played. The induced play can be represented as follows.

stage →	1	2	3	4	5	6	7	8	9	10	11	12	13	
action P1	L	L	Ⓡ	Ⓡ	L	Ⓡ	L	L	L	Ⓡ	Ⓡ	L	Ⓡ	...
action P2	Ⓡ	L	L	L	Ⓡ	L	Ⓡ	Ⓡ	L	L	L	Ⓡ	L	...
action P3	L	Ⓡ	L	L	L	L	L	L	L	L	L	L	L	...
public signal	L	L	L	L	L	L	L	L	L	L	L	L	L	...

Figure 7

The action of a player in the minority room, if any, is emphasized with a circle. The public signal is L at every stage, the induced play eventually has period 7 (one can see a period from stage 3 to stage 9), and the induced payoff is (3/7,3/7,0).

(b) This part is a direct consequence of the next three lemmata, where the best reply condition is checked for every player. □

Lemma 5.4. *In* Γ_3, σ^3 *is a best reply against* σ^{-3}.

Proof. Let τ^3 be any strategy of player 3 in Σ_3^3. It is necessary to prove that $\gamma^3(\tau^3, \sigma^{-3}) \leq \gamma^3(\sigma) = 0$. Assume in the sequel that (τ^3, σ^{-3}) is played, and distinguish two cases.

Case 1. Assume that the sequence of public signals never contains the symbol R. Then the sequence of actions played by players 1 and 2 is the same as in Figure 7. So at stages $3, 4, 5, 6$ player 3 is playing L (otherwise the public signal will be R at some stage). Since τ^3 has recall 3, it implies that player 3 will play L at every stage $t \geq 3$. Since L is at each stage the majority room, $\gamma^3(\tau^3, \sigma^{-3}) = 0$.

Case 2. Assume that at some stage the public signal is R. Consider the first stage $\bar{t}$ where this happens. Up to stage $\bar{t}$, the actions played by player 1 and 2 correspond to Figure 7, so at stage $\bar{t}$ it is not possible that both players 1 and 2 play R. Consequently, at stage $\bar{t}$: either (players 1 and 3 play R and player 2 plays L), or (players 2 and 3 play R and player 1 plays L). Recall now that σ^1 and σ^2 ask players 1 and 2 to change rooms whenever one of the available signals is R.

As long as one of the available public signals is R, players 1 and 2 will exchange rooms at each stage and, since players 1 and 2 are not in the same room, the payoff for player 3 will be zero. So to get out of this punishment phase, player 3 has to play three consecutive times L in order to induce three consecutive signals L. So it is possible to assume w.l.o.g. that there exists a stage t where the situation is as follows:

stage $\rightarrow$	t	$t+1$	$t+2$	$t+3$
action P1	L	R	L	$L^{(a)}$
action P2	R	L	R	$R^{(b)}$
action P3	L	L	L	
public signal	L	L	L	

or

stage $\rightarrow$	t	$t+1$	$t+2$	$t+3$
action P1	R	L	R	$L^{(c)}$
action P2	L	R	L	$R^{(d)}$
action P3	L	L	L	
public signal	L	L	L	

Figure 8

$^{(a)}$ because $f^1(L, R, L) = L$ (see Figure 6),
$^{(b)}$ because $f^2(R, L, R) = R$,
$^{(c)}$ because $f^1(R, L, R) = L$,
$^{(d)}$ because $f^2(L, R, L) = R$.

If player 3 plays R at stage $t+3$, then at this stage (players 1 and 3 play R and player 2 plays L) or (players 2 and 3 play R and player 1 plays L), and player 3 does not get out of the punishment phase where players 1 and 2 exchange rooms at each stage, and player 3's payoff is zero at each stage.

So let us assume that player 3 plays L at stage $t+3$. But since τ^3 has recall 3, player 3 will continue to play L as long as the public signal is L. The situation at the end of stage $t+2$ is similar to the situation at the end of stage 7 (left table) or stage 6 (right table) of Figure 7, and from this stage on player 3 will be in the majority room (the L room) hence will also have payoff zero. So $\gamma^3(\tau^3, \sigma^{-3}) = 0$. $\square$

Lemma 5.5. *In Γ_3, σ^1 is a best reply against σ^{-1}.*

Proof. Let τ^1 be a strategy profile of player 1 in Σ_3^1. It is necessary to prove that $\gamma^1(\tau^1, \sigma^{-1}) \leq \gamma^1(\sigma) = 3/7$. Assume that (τ^1, σ^{-1}) is played. Two cases are possible.

Case 1. Assume that at each stage the public signal is L. Then the situation is as follows:

stage $\rightarrow$	1	2	3	4	5	6	7	8	9	10	11	
action P1	L	L	X	Y	L	Z	L	L				
action P2	Ⓡ	L	L	L	Ⓡ	L	Ⓡ	Ⓡ	L	L	L	
action P3	L	Ⓡ	L	L	L	L	L	L	L	L	L	...
public signal	L	L	L	L	L	L	L	L	L	L	L	...

with X, Y, Z in $\{L, R\}$.

If $(X, Y) = (L, L)$, then player 1 only plays L since σ^1 has recall 3. And $\gamma^1(\tau^1, \sigma^{-1}) = 0 \leq 3/7$. So it is possible to assume w.l.o.g. that $(X, Y) \neq (L, L)$. The same argument shows that $Z = R$.

If $(X, Y) = (L, R)$, then the actions played by player 1 are $LLLRLRLL$. Since signals are assumed to be L at each stage, the next action of player 1 depends on her available actions only and one sees that is this word, the first appearance of LRL

is followed by R and the second il followed by L. This sequence of actions is thus unachievable with recall 3. If $(X, Y) = (R, L)$, then player 1 plays $LLR\ LLR\ LLR\ LLR\ldots$. But then at some stage the public signal will be R, yielding a contradiction.

The last case to consider is $(X, Y) = (R, R)$. In such a case:

stage →	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	
action P1	L	L	$\textcircled{R}$	$\textcircled{R}$	L	$\textcircled{R}$	L	L	T	U							
action P2	$\textcircled{R}$	L	L	L	$\textcircled{R}$	L	$\textcircled{R}$	$\textcircled{R}$	L	L	L	$\textcircled{R}$	L	$\textcircled{R}$	$\textcircled{R}$	L	
action P3	L	$\textcircled{R}$	L	L	L	L	L	L	L	L	L	L	L	L	L	L	...
public signal	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	...

If $T = R$, then player 1 plays the following sequence with period 6: $LLRRLR\ LLRRLR\ LLRRLR\ldots$. Since player 2 plays a sequence with period 7 and $\gcd(6, 7) = 1$, at some stage the signal will be R, yielding a contradiction. So $T = L$.

Now if $U = L$, the memory of player 1 at stage 10 is the same than at stage 9. She will thus plays always L and get a payoff of zero. If $U = R$, this is exactly in the case of Figure 7, and $\gamma^1(\tau^1, \sigma^{-1}) = 3/7$.

Case 2. Assume that there exists some stage where the public signal is R. It is possible to proceed as in the proof of Lemma 5.4 (Case 2). Since $f^2(L, R, L) = f^2(R, L, R) = R$ and $f^3(L, R, L) = f^3(R, L, R) = L$, also in this case $\gamma^1(\tau^1, \sigma^{-1}) = 0 \leq 3/7$. □

Lemma 5.6. *In Γ_3, σ^2 is a best reply against σ^{-2}.*

Proof. Let τ^2 in Σ_3^2 be a strategy of player 2. It is necessary to show that $\gamma^2(\tau^2, \sigma^{-2}) \leq 3/7 = \gamma^2(\sigma)$. Assume for the sake of contradiction that $\gamma^2(\tau^2, \sigma^{-2}) > 3/7$.

Claim. It cannot happen that at some stage, both players 1 and 3 play R.

Assume on the contrary that there exists a first stage $\bar{t}$ where both player 1 and player 3 play R. Necessarily $\bar{t} \geq 3$ and since player 3 plays R at $\bar{t}$, $\bar{t}$ cannot be the first stage where the signal is R. So there exists some stage $\widehat{t} < \bar{t}$ such that the signal at stage $\widehat{t}$ is R, and the signal at every stage t, $\widehat{t} < t < \bar{t}$ is L.

Since player 3 plays R at $\bar{t}$, then $\bar{t} \leq \widehat{t} + 3$. By definition of $\bar{t}$, at stage $\widehat{t}$: the signal is R, either player 1 or player 3 play L, and player 2 plays R. So after stage $\widehat{t}$, players 1 and 3 start to exchange rooms and this contradicts the fact that both player 1 and player 3 play R at $\bar{t}$.

Given this claim, two cases, and several sub-cases are possible.

Case 1. Assume that eventually the sequence of signals only contains L. There exists $\bar{t}$ with $u_t(\tau^2, \sigma^{-2}) = L$ for all $t \geq \bar{t}$.

Then for each stage $t \geq \bar{t} + 3$, player 3 will play L (see Figure 6), and given the definition of f^1, player 1 will eventually play the following sequence with period 7: $LLLRRLR\ LLLRRLR\ LLLRRLR\ldots$

Since it was assumed that $\gamma^2(\tau^2, \sigma^{-2}) > 3/7$, there must exist 7 consecutive stages among which player 2 is in the minority room for at least 4 stages. Since the majority room should be L from some stage on, the sequence played by player 2 depends on her own actions only and therefore has period at most $2^3 = 8$. One can then check that the only possibility for player 2 to win at least 4 times out of seven is to play the periodic sequence $RRRLLRL\ RRRLLRL\ RRRLLRL\dots$, so there must exist $t \geq \bar{t}$ such that the play is:

stage →	t	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11	+12	+13
action P1	L	L	L	Ⓡ	Ⓡ	L	Ⓡ	L	L	L	Ⓡ	Ⓡ	L	Ⓡ
action P2	Ⓡ	Ⓡ	Ⓡ	L	L	Ⓡ	L	Ⓡ	Ⓡ	Ⓡ	L	L	Ⓡ	L
action P3	L	L	L	L	L	L	L	L	L	L	L	L	L	L
public signal	L	L	L	L	L	L	L	L	L	L	L	L	L	L

This sequence of actions of player 2 will be denoted by ω in the sequel.

Subcase 1.a. Assume that all signals are L. Then the situation is as follows.

stage →	1	2	3	4
action P1	L	L	R	R
action P2	X	L	L	L
action P3	L	R	L	L
public signal	L	L	L	L

It must be $X = R$ otherwise player 2 only plays L and $\gamma^2(\tau^2, \sigma^{-2}) = 0$. So player 2, at stage 4, plays L after RLL. This is not compatible with the sequence ω.

Subcase 1.b. Assume that there exists a last stage $\bar{t}$ where the public signal is R. Since player 1 and player 3 never play R at the same time, two possibilities can occur at stage $\bar{t}$.

Subsubcase 1.b.1. If player 1 plays R at stage $\bar{t}$, then

stage →	$\bar{t}$	+1	+2	+3	+4	+5	+6	+7
action P1	R	$L^{(a)}$	$R^{(a)}$	$L^{(a)}$	$L^{(b)}$	$L^{(c)}$	$R^{(d)}$	
action P2	R	$L^{(e)}$	$L^{(e)}$	$L^{(e)}$	X	Y	$L^{(e)}$	
action P3	L	$R^{(a)}$	$L^{(a)}$	$R^{(a)}$	L	L	L	L
public signal	R	L	L	L	L	L	L	L

$^{(a)}$ player 1 and player 3 change rooms after a public signal R,
$^{(b)}$ because $f^1(L, R, L) = L$,
$^{(c)}$ because $f^1(R, L, L) = L$,
$^{(d)}$ because $f^1(L, L, L) = R$,
$^{(e)}$ by assumption, the signal has to be L at every stage $\geq \bar{t} + 1$.

If $X = L$, then player 2 will always play L and have a payoff of zero. So $X = R$. Then $Y = L$ because of the periodic sequence ω. But using ω again, at stage $\bar{t} + 6$ player 2 should play R, yielding a contradiction.

Subsubcase 1.b.2. If player 3 plays R at stage $\bar{t}$, then

stage →	$\bar{t}$	+1	+2	+3	+4	+5	+6	+7	+8	+9
action P1	L	R	L	R	L	L	L	R	R	L
action P2	R	L	L	L	X	Y	Z	L	L	
action P3	R	L	R	L	L	L	L	L	L	
public signal	R	L	L	L	L	L	L	L	L	L

It must be that $X = R$, otherwise player 2 will always play L after $\bar{t}$. The sequence ω then gives $Y = L$, and $Z = R$. But by ω again at stage $\bar{t} + 7$, player 2 should play R, yielding a contradiction.

Case 2. It remains to consider the case with an infinite number of stages where the public signal is R.

Take any interval of stages $\{t_1, \ldots, t_2\}$, where $t_1 < t_2$, $u_{t_1}(\tau^2, \sigma^{-2}) = u_{t_2}(\tau^2, \sigma^{-2}) = R$, and for every $t \in \{t_1 + 1, \ldots, t_2 - 1\}$, $u_t(\tau^2, \sigma^{-2}) = L$. To conclude the proof, it is sufficient to show that the average payoff of player 2 at stages $t_1, \ldots, t_2 - 1$ is at most $3/7$.

Assume by contradiction that it is not the case, i.e., assume that the average payoff of player 2 at stages $t_1, \ldots, t_2 - 1$ is greater than $3/7$. Since player 1 and player 3 never play R at the same stage, at stage t_1, either (players 1 and 2 play R, player 3 plays L) or (players 3 and 2 play R, player 1 plays L). In each case, players 1 and 3 are going to exchange rooms at stages $t_1 + 1, t_1 + 2, t_1 + 3$, so the payoff of player 2 is zero at each stage t in $\{t_1, t_1 + 1, t_1 + 2, t_1 + 3\}$. It was assumed that the average payoff of player 2 between stage t_1 and stage $t_2 - 1$ is greater than $3/7$. This implies that $t_2 \geq t_1 + 8$. So the signal at stages $t_1 + 1, \ldots, t_1 + 7$ is L. Two cases are possible.

Subcase 2.a. At stage t_1, player 3 plays L.

stage →	t_1	+1	+2	+3	+4	+5	+6	+7	+8	...	t_2
action P1	R	L	Ⓡ	L	L	L	Ⓡ	Ⓡ	L	...	R
action P2	R	L	L	L	X	Y	L	L	Z	...	R
action P3	Ⓛ	Ⓡ	L	Ⓡ	L	L	L	L	L	...	Ⓛ
public signal	R	L	L	L	L	L	L	L	L	...	R

By a standard argument $X = R$ (otherwise player 2 plays only L and gets 0). If $Y = L$, then, since player 2 has recall 3

	t_1	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10	+11	+12	+13	+14	+15	t_2
P1	R	L	Ⓡ	L	L	L	Ⓡ	Ⓡ	L	Ⓡ	L	L	L	Ⓡ	R	L	R
P2	R	L	L	L	Ⓡ	L	L	L	Ⓡ	L	L	L	Ⓡ	L	L	L	R
P3	Ⓛ	Ⓡ	L	Ⓡ	L	L	L	L	L	L	L	L	L	L	L	L	Ⓛ
signal	R	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	R

Then $t_2 = t_1 + 16$, and the average payoff of player 2 is 3/16. So to conclude subcase 2.a., it remains to consider the case when $Y = R$.

stage $\rightarrow$	t_1	+1	+2	+3	+4	+5	+6	+7	+8	+9
action P1	R	L	$\textcircled{R}$	L	L	L	$\textcircled{R}$	$\textcircled{R}$	L	R
action P2	R	L	L	L	$\textcircled{R}$	$\textcircled{R}$	L	L	Z	T
action P3	$\textcircled{L}$	$\textcircled{R}$	L	$\textcircled{R}$	L	L	L	L	L	L
public signal	R	L	L	L	L	L	L	L	L	

Since player 2 has recall 3 and the memory RLL was met at stage $t_1 + 3$, $Z = L$, and since LLL was met at stage $t_1 + 4$, we have $T = R$. Thus, $t_2 = t_1 + 9$ and the average payoff of player 2 is at most 3/9.

Subcase 2.b. At stage t_1, player 1 plays L.

stage $\rightarrow$	t_1	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10
action P1	$\textcircled{L}$	$\textcircled{R}$	L	$\textcircled{R}$	L	L	L	$\textcircled{R}$	$\textcircled{R}$	$L^{(d)}$	$R^{(f)}$
action P2	R	L	L	L	$\textcircled{R}^{(a)}$	Y	Z	L	$L^{(b)}$		
action P3	R	L	$\textcircled{R}$	L	L	L	L	L	L	$L^{(d)}$	$L^{(f)}$
public signal	R	L	L	L	L	L	L	L	$L^{(c)}$	$L^{(e)}$	

$^{(a)}$ standard argument because player 2 has recall 3,
$^{(b)}$ the only possibility is L otherwise there is no chance for the average payoff of player 2 to be greater than 3/7. Furthermore $^{(b)}$ implies $^{(c)}$, $^{(c)}$ implies $^{(d)}$, $^{(d)}$ implies $^{(e)}$, and $^{(e)}$ implies $^{(f)}$.

Now, $(Y, Z) = (L, L)$ is not possible because player 2 would play $LLLL$ at stages $t_1 + 5, t_1 + 6, t_1 + 7, t_1 + 8$. The case $(Y, Z) = (L, R)$ also is not possible, because player 2 would have to play the same action at both stages $t_1 + 6$ and $t_1 + 8$.

Assume that $(Y, Z) = (R, L)$. Then

stage $\rightarrow$	t_1	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10
action P1	$\textcircled{L}$	$\textcircled{R}$	L	$\textcircled{R}$	L	L	L	$\textcircled{R}$	$\textcircled{R}$	L	R
action P2	R	L	L	L	$\textcircled{R}$	$\textcircled{R}$	L	L	L	$\textcircled{R}$	R
action P3	R	L	$\textcircled{R}$	L	L	L	L	L	L	L	$\textcircled{L}$
public signal	R	L	L	L	L	L	L	L	L	L	R

Here $t_2 = t_1 + 10$. The average payoff for player 2 at stages $t_1, t_1 + 1, \ldots, t_2 - 1$ is only 3/10. The last case to consider is $(Y, Z) = (R, R)$.

stage $\rightarrow$	t_1	+1	+2	+3	+4	+5	+6	+7	+8	+9	+10
action P1	$\textcircled{L}$	$\textcircled{R}$	L	$\textcircled{R}$	L	L	L	$\textcircled{R}$	$\textcircled{R}$	L	R
action P2	R	L	L	L	$\textcircled{R}$	$\textcircled{R}$	$\textcircled{R}$	L	L	X'	Y'
action P3	R	L	$\textcircled{R}$	L	L	L	L	L	L	L	L
public signal	R	L	L	L	L	L	L	L	L	L	

Necessarily $Y' = R$, and $t_2 = t_1 + 10$. The average payoff for player 2 is then at most 4/10 ($< 3/7$). $\square$

Proof of Proposition 4.1. (a) Note that $(0,0,1) \in E_\infty$ by the classical folk theorem. This payoff is feasible by a sequence of action profiles where player 1 plays T at each stage and player 2 alternates between L and R with respective frequencies $\left(\frac{1}{1+\sqrt{2}}, \frac{\sqrt{2}}{1+\sqrt{2}}\right)$. Player 1 punishes deviations from player 2 by playing B and player 2 punishes by playing R.
(b) Consider now bounded recall strategies. Take k in $\mathbb{N}$ and (x,y,z) in E_k $(=\widehat{E}_k)$. Since the play induced by a pure strategy profile with bounded recall is periodic, the average frequencies λ_1, λ_2, λ_3, λ_4 of the pure action profiles (respectively of (T,L), (T,R), (B,L), (B,R)) are non-negative rational numbers, summing up to one. We have:

$$\begin{aligned} x &= \lambda_1\sqrt{2} - \lambda_2, \\ y &= -\lambda_1\sqrt{2} + \lambda_2, \\ z &= \lambda_1 + \lambda_2. \end{aligned}$$

By individual rationality $x \geq 0$ and $y \geq 0$, so $\lambda_1\sqrt{2} = \lambda_2$. Since $\sqrt{2}$ is irrational, this implies $\lambda_1 = \lambda_2 = 0$. So $(x,y,z) = (0,0,0)$. Since $(0,0,0)$ is an equilibrium payoff, one obtains $E_k = \{(0,0,0)\}$ for each k. □

Proof of Proposition 4.2. Consider a pure equilibrium of the repeated game with unbounded recall. If at some stage (M,L) is played, then player 2 may play R at this stage and get a payoff of 3 instead of 2, without any further consequence because the signal induced by (M,R) is the same as the signal induced by (M,L). Thus, in equilibrium, (M,L) cannot be played with positive frequency, and therefore $E_\infty \subset \{(x,y) \in \mathbb{R}^2, x+y \leq 3\}$.

Fix a positive integer k and define $\sigma = (\sigma^1, \sigma^2) \in \Sigma_k$ as follows.

- σ^2 plays L at each stage whatever happens.
- σ^1 plays B_1 at stage 1, and is defined via a main phase and a transition phase. After stage 1, player 1 using σ^1 says that she is in the main phase if and only if (all public signals in her recall equal u, and the last action played by player 1 is not T). If this condition is not satisfied, then player 1 says that she is in the transition phase.
  - In the transition phase, player 1 plays B_1 if her last k actions all equal T, and plays T otherwise.
  - In the main phase, player 1 induces the following periodic sequence of actions

    $$\underbrace{B_1B_1\ldots B_1}_{k \text{ times}}\underbrace{B_2B_2\ldots B_2}_{k \text{ times}}\underbrace{MM\ldots M}_{k \text{ times}}\underbrace{B_1B_1\ldots B_1}_{k \text{ times}}\underbrace{B_2B_2\ldots B_2}_{k \text{ times}}\underbrace{MM\ldots M}_{k \text{ times}}\ldots$$

    That is, player 1 plays B_2 (resp. M, resp. B_1) if her last k actions are all B_1 (resp. B_2, resp. M), and repeats her last action otherwise.

This ends the definition of σ.

Under σ, the play remains forever in the main phase, inducing the payoff

$$\frac{1}{3}\,g(B_1,L)+\frac{1}{3}\,g(B_2,L)+\frac{1}{3}\,g(M,L)=\frac{1}{3}(2,1)+\frac{1}{3}(2,1)+\frac{1}{3}(2,2)=\left(2,\frac{4}{3}\right).$$

We check now that σ is an equilibrium of Γ_k.

The strategy σ^1 is obviously a best response against σ^2 because the maximal payoff for player 1 is 2. Let now τ^2 be any strategy of player 2 with recall k and assume that $\gamma^2(\sigma^1,\tau^2)>1$. There must exist some first stage $\bar{t}$ where player 1 plays M. Then necessarily the following happened:

stage $\rightarrow$	$\bar{t}-2k$	$\dots$	$\bar{t}-(k+1)$	$\bar{t}-k$	$\dots$	$\bar{t}-1$	$\bar{t}$
action P1	B_1	$\dots$	B_1	B_2	$\dots$	B_2	M
public signal	u	$\dots$	u	u	$\dots$	u	

This implies that player 2 has played L and the signal was u at every stage $\bar{t}-2k,\dots,\bar{t}-1$. Since $2k>k$ and τ^2 has recall k, player 2 using τ^2 will play L at every stage $t\geq\bar{t}-2k$. So $\gamma^2(\sigma^1,\tau^2)=4/3=\gamma^2(\sigma)$. Consequently σ^2 is a best response against σ^1 in the game with private strategies and recall 3, and $(2,4/3)\in E_k$. Hence $(2,4/3)\in\cap_{k\geq1}E_k\backslash E_\infty$. □

Acknowledgment

The work of Marco Scarsini was partially supported by MIUR-COFIN. This paper was written while he was visiting CEREMADE, whose nice hospitality is gratefully acknowledged.

The authors wish to thank two referees, an associate editor, and the area editor Eilon Solan for their careful reading and insightful comments and criticisms, that lead to a re-writing of the paper.

References

Abreu, D., Pearce, D., and Stacchetti, E. (1990) Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica* **58**, 1041–1063.

Abreu, D. and Rubinstein, A. (1988) The structure of Nash equilibrium in repeated games with finite automata. *Econometrica* **56**, 1259–1281.

Arthur, B. W. (1994) Inductive reasoning and bounded rationality. *Amer. Econom. Rev. Papers Proc.* **84**, 406–411.

Arthur, B. W. (1999) Complexity and the economy. *Science* **284**, 107–109.

Bavly, G. and Neyman, A. (2005) Online concealed correlation by bounded rationality. Mimeo.

Ben-Porath, E. (1990) The complexity of computing a best response automaton in repeated games with mixed strategies. *Games Econom. Behav.* **2**, 1–12.

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

Blackwell, D. (1962) Discrete dynamic programming. *Ann. Math. Statist.* **33**, 719–726.

Bottazzi, G. and Devetag, G. (2004) Coordination and self-organization in minority games: experimental evidence. In *The complex dynamics of economic interaction*, volume 531 of *Lecture Notes in Econom. and Math. Systems*, 283–300. Springer, Berlin.

Bottazzi, G., Devetag, G., and Dosi, G. (2002) Adaptive learning and emergent coordination in minority games. *Simul. Model. Practice Theory* **10**, 321–347.

Bottazzi, G., Devetag, G., and Dosi, G. (2003) Adaptive coordination and aggregate efficiency in minority games. In *Heterogenous agents, interactions and economic performance (Maastricht, 2001)*, volume 521 of *Lecture Notes in Econom. and Math. Systems*, 19–42. Springer, Berlin.

de Bruijn, N. G. (1946) A combinatorial problem. *Nederl. Akad. Wetensch., Proc.* **49**, 758–764 = *Indagationes Math.* **8**, 461–467 (1946).

Challet, D., Marsili, M., and Zhang, Y.-C. (2005) *Minority Games.* Oxford University Press, Oxford.

Chan, K. C. (1988) On the contrarian investment strategy. *J. Bus.* **61**, 147–163.

Cole, H. L. and Kocherlakota, N. R. (2005) Finite memory and imperfect monitoring. *Games Econom. Behav.* **53**, 59–72.

Coolen, A. C. C. (2005) *The Mathematical Theory of Minority Games.* Oxford University Press, Oxford.

De Bondt, W. F. M. and Thaler, R. (1985) Does the stock market overreact? *J. Finance* **40**, 793–805.

Erev, I. and Rapoport, A. (1998) Magic, reinforcement learning and coordination in a market entry game. *Games Econom. Behav.* **23**, 146–175.

Fudenberg, D., Levine, D., and Maskin, E. (1994) The folk theorem with imperfect public information. *Econometrica* **62**, 997–1039.

Fudenberg, D. and Levine, D. K. (1994) Efficiency and observability with long-run and short-run players. *J. Econom. Theory* **62**, 103–135.

Green, E. J. and Porter, R. H. (1984) Noncooperative collusion under imperfect price information. *Econometrica* **52**, 87–100.

Kalai, E. and Stanford, W. (1988) Finite rationality and interpersonal complexity in repeated games. *Econometrica* **56**, 397–410.

Kandori, M. (2002) Introduction to repeated games with private monitoring. *J. Econom. Theory* **102**, 1–15.

Kandori, M. and Obara, I. (2006) Efficiency in repeated games revisited: The role of private strategies. *Econometrica* **74**, 499–519.

Lehrer, E. (1988) Repeated games with stationary bounded recall strategies. *J. Econom. Theory* **46**, 130–144.

Lehrer, E. (1989) Lower equilibrium payoffs in two-player repeated games with nonobservable actions. *Internat. J. Game Theory* **18**, 57–89.

Lehrer, E. (1992a) Correlated equilibria in two-player repeated games with nonobservable actions. *Math. Oper. Res.* **17**, 175–199.

Lehrer, E. (1992b) Two-player repeated games with nonobservable actions and observable payoffs. *Math. Oper. Res.* **17**, 200–224.

Lehrer, E. (1994) Finitely many players with bounded recall in infinitely repeated games. *Games Econom. Behav.* **7**, 390–405.

Lempel, A. (1971) m-ary closed sequences. *J. Combinatorial Theory Ser. A* **10**, 253–258.

Mailath, G. J., Matthews, S. A., and Sekiguchi, T. (2002) Private strategies in finitely repeated games with imperfect public monitoring. *Contrib. Theor. Econ.* **2**, Art. 2, 23 pp. (electronic).

Milchtaich, I. (1998) Crowding games are sequentially solvable. *Internat. J. Game Theory* **27**, 501–509.

Milchtaich, I. (2000) Generic uniqueness of equilibrium in large crowding games. *Math. Oper. Res.* **25**, 349–364.

Neyman, A. (1998) Finitely repeated games with finite automata. *Math. Oper. Res.* **23**, 513–552.

Ochs, J. (1990) The coordination problem in decentralized markets: an experiment. *Quart. J. Econom.* **105**, 545–559.

Ochs, J. (1995) Coordination problems. In Kagel, J. and Roth, A. (eds.), *Handbook of Experimental Economics*. Princeton University Press, Princeton.

Radner, R. (1985) Repeated principal-agent games with discounting. *Econometrica* **53**, 1173–1198.

Rapoport, A., Seale, D., Erev, I., and Sundali, J. A. (2002) Equilibrium play in large group market entry games. *Management Sci.* **44**, 119–141.

Renault, J., Scarlatti, S., and Scarsini, M. (2005) A folk theorem for minority games. *Games Econom. Behav.* **53**, 208–230.

Renault, J., Scarlatti, S., and Scarsini, M. (2006) Discounted and finitely repeated minority games. Technical report, CEREMADE, Université de Paris Dauphine.

Renault, J. and Tomala, T. (2004) Communication equilibrium payoffs in repeated games with imperfect monitoring. *Games Econom. Behav.* **49**, 313–344.

Rosenthal, R. W. (1973) A class of games possessing pure-strategy Nash equilibria. *Internat. J. Game Theory* **2**, 65–67.

Rubinstein, A. (1979) Equilibrium in supergames with the overtaking criterion. *J. Econom. Theory* **21**, 1–9.

Rubinstein, A. (1986) Finite automata play the repeated prisoner's dilemma. *J. Econom. Theory* **39**, 83–96.

Rubinstein, A. and Yaari, M. E. (1983) Repeated insurance contracts and moral hazard. *J. Econom. Theory* **30**, 74–97.

Sabourian, H. (1998) Repeated games with M-period bounded memory (pure strategies). *J. Math. Econom.* **30**, 1–35.

Selten, R. and Güth, W. (1982) Equilibrium point selection in a class of market entry games. In Diestler, M., Fürst, E., and Schwadiauer, G. (eds.), *Games, Economics, and Time Series Analysis*. Physica-Verlag, Wien-Würzburg.

Tomala, T. (1998) Pure equilibria of repeated games with public observation. *Internat. J. Game Theory* **27**, 93–109.

Yoeli, M. (1962) Binary ring sequences. *Amer. Math. Monthly* **69**, 852–855.

July 15, 2006