Bounded Memory with Finite Action Spaces*

Mehmet Barlo   Guilherme Carmona
Sabancı University  Universidade Nova de Lisboa

Hamid Sabourian
University of Cambridge
January, 2008

Abstract

This study establishes that the Folk Theorem holds for any 2 player repeated game with (time-independent) limited-memory pure strategies, and with discounting. That is, we prove that any strictly individually rational payoff can be approximated with a limited-memory subgame perfect strategy profile (involving only pure actions) when players are sufficiently patient.

Journal of Economic Literature Classification Numbers: C72; C73; C79

Keywords: Repeated Games; Memory; Bounded Rationality; Folk Theorem

*Correspondent: Mehmet Barlo, FASS, Sabancı University Orhanlı; 34956 Tuzla Istanbul Turkey; Email: barlo@sabanciuniv.edu; telephone: +(90) 216 483 9284; fax: +(90) 216 483 9250.
1 Introduction

Repeated games provide a framework in which long-term relationships can be analyzed. In particular, the repeated prisoners’ dilemma is used to study whether or not it is possible for two individuals to cooperate even when they have a short-term incentive for not doing so. The well-known answer displays that if players are sufficiently patient, then cooperation in every period will be a (subgame perfect) equilibrium outcome. But, because players do not possess a short-term incentive to cooperate, in any equilibrium strategy that results in cooperation, current play must depend on the past. Similar considerations apply to any (strictly) individually rational payoff vector of any repeated game: They can be approximated with subgame perfection due to the Folk Theorems of Fudenberg and Maskin (1986) and Aumann and Shapley (1994); and in any equilibrium strategy supporting them, players must remember and condition their play on the past. Therefore, it is reasonable to expect, as suggested by Aumann (1981), that the extensive multiplicity of equilibria described by the Folk Theorem may be reduced by restricting players to use limited memory strategies.

In contrast, this study shows that the Folk Theorem continues to hold with limited memory strategies in every full-dimensional 2 player discounted repeated game with pure strategies. Specifically, Theorem 1 establishes that any strictly individually rational payoff of any two person repeated game can be approximated with a time-independent limited memory subgame perfect strategy profile (involving only pure actions) when players are sufficiently patient. It needs to be mentioned that the size of the memory and the complexity of the strategy used in the proof depends on how fine the approximation is, but not on the discount
rate. In general, as the approximation gets finer, the size of the time-independent memory explodes to infinity.

Even though our result holds for any 2 player repeated game, we find it helpful to discuss our construction first in the context of the prisoners’ dilemma with discounting to provide an easier reading. It is well known that in the context of the prisoners’ dilemma without any restrictions on memory the Folk Theorem could be easily reached: Knowing that defection in every period is the most severe credible punishment for both of the players, we could use an equilibrium simple strategy inducing a path with the desired payoff vector\(^1\)

However, with limited memory such strategies may fail to be subgame perfect, and a player’s deviation may fail to trigger perpetual defection by the other. To see this, suppose that we want to implement a cycle consisting of \(((C, D), (D, C))\), which yields an average payoff strictly higher than the minmax return to each player\(^2\). The simple strategy inducing this cycle, denoted by \(\pi = \{\pi_t\}_{t=1}^\infty\), and involving the play of \((D, D)\) forever for any history inconsistent with the equilibrium path, is subgame perfect with unbounded memory and with sufficiently high discount factors. However, this strategy is not subgame perfect with limited memory. This is because, if players can remember at most \(M\) periods, one player prefers to deviate at a history with its last \(M\) entries equal to \((a_1, \pi_2, \ldots, \pi_M)\) with \(a_1 \neq \pi_1\) instead of playing the punishment: If \(\pi_M = (D, C)\), then player 1 can play \(C\) instead of \(D\), which will make the play return to the equilibrium outcome in the next period\(^3\). Hence,

\(^1\)For a definition of simple strategies, see Abreu (1988).
\(^2\)We are thankful to an anonymous referee for providing us with a similar example.
\(^3\)Notice that if \(\pi_M = (C, D)\) by a similar argument, player 2 would deviate.
player 1’s continuation payoff in that history strictly exceeds the payoff he would receive by not deviating.

The way we overcome this difficulty is by allowing the play to continue along the equilibrium path even in some histories that are inconsistent with the equilibrium path. However, this alone is not sufficient as the above example illustrates. In fact, we could modify the above strategy so that it recommends \( \pi_{M+1} \) at a history whose \( M \)-tail equals \((a_1, \pi_2, \ldots, \pi_M)\) with \( a_1 \neq \pi_1 \); and, other than this modification the strategy is unchanged. Then, player \( i \) for whom \( \pi_i^M = C \), will find it profitable to deviate from \( D \) to \( C \) at a history with its \( M \)-tail equal to \((a_1, a_2, \pi_2, \ldots, \pi_{M-1})\), for any \( a_2 \neq \pi_1 \) and any \( a_1 \). By doing so, he produces a history with its \( M \)-tail equal to \((a_2, \pi_2, \ldots, \pi_M)\) and brings the play back to the equilibrium path. Therefore, if we continue to change the strategy by allowing the play to return to the equilibrium path at these problematic histories, an inductive argument would imply that the play must be the equilibrium path after any possible history, a requirement clearly incompatible with subgame perfection. Therefore, this examples shows that the approach of changing the strategy as suggested is useful only when the equilibrium path satisfies certain properties.

This is why, to obtain the limited memory Folk Theorem in the context of the prisoners’ dilemma, attention needs to be restricted to the following kind of outcomes: Cycles that have at least two consecutive \((C,C)\)s followed by at least one \((D,D)\); and ordered so that all \((C,C)\)s come first, followed by all the \((D,D)\)s, and the rest (without loss of generality) composed of either only \((C,D)\)s or only \((D,C)\)s. Letting \( M \) equal to the size of the cycle,
the time-independent $M$-memory strategy we specify recommends that at a history $h$ with its $M$-tail either consistent with the equilibrium outcome, or given by $(a_1, \ldots, a_{M-m}, \pi_1, \ldots, \pi_m)$ for some $a_1, \ldots, a_{M-m}$, $m$ greater or equal to the number of consecutive $(C, C)$s in the cycle, players should play the next equilibrium action. Moreover, for any other history, the strategy recommends to play $D$. As a result, if a player deviates singly, then the other player will play $D$, thus, players will never observe two consecutive $(C, C)$s. The reason why we need at least two consecutive $(C, C)$s followed by at least one $(D, D)$ is illustrated in the following examples: Suppose first that the cycle we analyze has two consecutive $(C, C)$s, but not followed by a $(D, D)$. Then, if the $M$-tail of a history is given by $(a_1, \ldots, a_{M-2}, (C, C), (C, C))$ for some $a_1, \ldots, a_{M-2}$, and $(D, C)$ is the next action profile in the cycle, then player 1 can deviate: Play $C$ and this deviation does not trigger player 2 to play $D$ forever, because next period the $M$-tail of the resulting history would be $(a_2, \ldots, a_{M-1}, (C, C), (C, C))$. Second, suppose that the cycle consists of $((C, C), (D, D), (D, C))$. Then, if the $M$-tail of an history equals $(a_1, \ldots, a_{M-2}, (C, C), (D, D))$ for some $a_1, \ldots, a_{M-2}$, then the strategy recommends $(D, C)$. However, player 1 can deviate and play $C$ and this deviation does not trigger player 2 to play $D$ forever; in contrast, the play continues along the equilibrium path. Note that even though these deviations in both of these examples are clearly not profitable, these examples illustrate that a single agent deviation does not necessarily trigger the other to play $D$ forever.

of only $(C, C)$s and $(D, D)$s suffices. On the other hand, if the payoff of the second agent is strictly higher than that of the first, the cycle constructed will involve only $(C, D)$s following $(C, C)$s and $(D, D)$s. It should be noted that the reverse situation can then be handled with only $(D, C)$s.
When generalizing these observations to any 2 player repeated game, we use the standard construction of Fudenberg and Maskin (1986) involving $R$ period common minmaxing, $R$ a natural number, with the following important modification discussed below. Let $S = (s^{(1)}, \ldots, s^{(r)})$ be the joint action space, and $\bar{m}$ be the common minmax with the convention that $s^{(2)} = \bar{m}$ for all $i = 1, 2$, and $s^{(3)} = \bar{m}$. Let $u$ be any strictly individually rational payoff, and let $\pi$ be an outcome that approximates $u$, and consists of the repetition of the cycle $((s^{(1)}; p_1), (s^{(2)}; p_2), \ldots, (s^{(r)}; p_r))$ with $p_2 \geq 2, p_j > 0$ for all $j = 2, \ldots, r$, and $\sum_{j=1}^{r} p_j = K$. Moreover, let $\hat{\pi}$ be the outcome path that the play will follow after a possible $R$ period common minmax phase, and it is given by the repetition of the cycle $((s^{(2)}; \hat{p}_2), (s^{(3)}; \hat{p}_3), (s^{(1)}; \hat{p}_1), (s^{(4)}; \hat{p}_4), \ldots, (s^{(r)}; \hat{p}_r))$ with $\hat{p}_3 = 2p_3 + 1$ and $\hat{p}_j = 2p_j$ for all $j \neq 3$.

In words, the $M$-memory strategy $f$ that we will employ in the proof of the Folk Theorem, $M > 4K + 1 > R$, will allow the play go back to $\pi$ in some histories that are not consistent with $\pi$ (with $M$ observations), and involve $R$ periods of common minmaxing followed by $\hat{\pi}$ (not $\pi$). Specifically, it is given as follows: Let $h \in H$. If length of $M$ tail of history $h$ is equal to $M$, then $f$ prescribes the following behavior: If the $M$ tail of $h$ is given by $(a_1, \ldots, a_{M-k}, \pi_1, \ldots, \pi_k)$ for some $a_1, \ldots, a_{M-k} \in S$ and for some $p_1 + p_2 \leq k \leq M$, then $f(h) = \pi_{k+1}$. If the $M$ tail of $h$ is $(a_1, \ldots, a_{M-k}, \hat{\pi}_1, \ldots, \hat{\pi}_k)$ for some $a_1, \ldots, a_{M-k} \in S$, and for some $\hat{p}_2 \leq k \leq M$, then $f(h) = \hat{\pi}_{k+1}$. If $M$ tail of $h$ is given by $((\bar{m}; M - k), (s^{(2)}, k))$ for some $k$, $0 \leq k < \hat{p}_2$, then $f(h) = s^{(2)}$. For other histories with the length of $M$ tail of history $h$ is equal to $M$, $f(h) = \bar{m}$. The definition of $f$ is standard for histories with the length of
$M$ tail less than $M$, because then we have full memory.

At the hearth of the above difficulties lies the restriction to pure strategies and the use of time-independent memory. In fact, when players’ action spaces are connected and payoff functions continuous, players can encode the relevant information about the history at an arbitrarily low cost by using the richness of the action spaces. As a consequence, Barlo, Carmona, and Sabourian (2006) established that the Folk Theorem continues to hold even with time-independent one-memory strategies. However, this richness is lost with finite action spaces, and such encoding becomes difficult, if not impossible. On the other hand, Barlo and Carmona (2007) proves that in the context of prisoners’ dilemma (with the cooperative payoff being Pareto optimal) the order of memory in the limited-memory Folk Theorem does not depend on the fineness of the approximation with time-dependent limited memory pure strategies. Sabourian (1998) obtained a perfect Folk Theorem with bounded memory strategies for the case of repeated games with no discounting and finite number of pure actions. Furthermore, Hörner and Olszewski (2007) also present a perfect Folk Theorem with bounded memory strategies for games with (public or private but almost

\footnote{Formally, that study displays that as long as the cooperative payoff is Pareto optimal, for any $\varepsilon > 0$ there exists a natural number $M_\varepsilon$ such that for any strictly individually rational payoff bounded away by $\varepsilon$ from the minmax return, there is a sufficiently high discount factor with which this payoff can be arbitrarily approximated employing a time-dependent $M_\varepsilon$-memory subgame perfect strategy, which involves only pure actions.}

imperfect monitoring and finite action and outcome spaces. Their result, however, requires a rich set of public signal and displays a trade-off between the discount factor and the length of the memory.

2 Notation and Definitions

The stage game:

A normal form game $G$ is defined by $G = (N, (S_i)_{i \in N}, (u_i)_{i \in N})$, where $N = \{1, \ldots, n\}$ is a finite set of players, $S_i$ is the set of player $i$’s actions and $u_i : \prod_{i \in N} S_i \rightarrow \mathbb{R}$ is player $i$’s payoff function. We assume that $S_i$ is finite and $|S_i| \geq 2$ for all $i \in N$. Hence, $|\prod_{i \in N} S_i| \geq 2^n$.

Let $S = \prod_{i \in N} S_i$ and $S_{-i} = \prod_{j \neq i} S_j$. Also, for any $i \in N$ denote respectively the minmax payoff and a minmax profile for player $i$ by $v_i = \min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} u_i(s_i, s_{-i})$ and $m^i \in \arg\min_{s_{-i} \in S_{-i}} \max_{s_i \in S_i} u_i(s_i, s_{-i})$. If $G$ is a 2-player game, a mutual minmax profile is $\bar{m} = (m^1, m^2)$.

For all $\epsilon \geq 0$, let $U^\epsilon = \{u \in \text{co}(u(A)) : u_i > v_i + \epsilon$ for all $i \in N\}$. We refer to $U^\epsilon$ as the set of $\epsilon$-strictly individually rational payoffs. In particular, $U^0$ denotes the set of strictly individually rational payoffs. Finally, let $U = \{u \in \text{co}(u(A)) : u_i \geq v_i$ for all $i \in N\}$.

The repeated game:

The supergame $G^\infty(\delta)$ of $G$ consists, for all $\delta \in (0, 1)$, of an infinite sequence of repetitions of $G$. We denote the action of any player $i$ in $G^\infty(\delta)$ at any date $t = 1, 2, 3, \ldots$ by $s^t_i \in S_i$. Also, let $s^t = (s^t_1, \ldots, s^t_n)$ be the profile of choices at $t$.

For $t \geq 1$, a $t$-stage history is a sequence $h_t = (s^1, \ldots, s^t)$. The set of all $t$-stage histories
is denoted by $H_t = S^t$ (the $t$-fold Cartesian product of $S$). We use $H_0$ to represent the initial (0-stage) history. The set of all histories is defined by $H = \bigcup_{t=0}^{\infty} H_t$.

For every $h \in H$, define $h^r \in A$ to be the projection of $h$ onto its $r$th coordinate. For every $h \in H$ we let $\ell(h)$ denote the length of $h$. For two positive length histories $h$ and $\bar{h}$ in $H$ we define the concatenation of $h$ and $\bar{h}$, in that order, to be the history $(h \cdot \bar{h})$ of length $\ell(h) + \ell(\bar{h})$: $(h \cdot \bar{h}) = (h^1, h^2, \ldots, h^{\ell(h)}, \bar{h}^1, \bar{h}^2, \ldots, \bar{h}^{\ell(\bar{h})})$. We follow the convention that $e \cdot h = h \cdot e = h$ for every $h \in H$. For a history $h \in H$ and an integer $0 \leq m \leq \ell(h) - 1$, the $m$-stage tail of $h$ is denoted by $T^m(h) \in H : T^0(h) = e$ and $(T^m(h))^j = h^{\ell(h)-(m+1)+j}$ for $j = 1, 2, \ldots, m$ and $1 \leq m \leq \ell(h) - 1$. We also follow the convention that $T^m(h) = h$, for all $m \geq \ell(h)$.

For all $i \in \mathbb{N}$, a strategy for player $i$ is a function $f_i : H \rightarrow S_i$ mapping histories into actions. The set of player $i$’s strategies is denoted by $F_i$, and $F = \prod_{i \in \mathbb{N}} F_i$ is the joint strategy space. Finally, a strategy vector is $f = (f_1, \ldots, f_n)$. Given an individual strategy $f_i \in F_i$ and a history $h \in H$ we denote the individual strategy induced at $h$ by $f_i|_h$. This strategy is defined pointwise on $H$: $(f_i|_h)(\bar{h}) = f_i(h \cdot \bar{h})$, for every $\bar{h} \in H$. We will use $(f|_h)$ to denote $(f_1|_h, \ldots, f_n|_h)$ for every $f \in S$ and $h \in H$. We let $F_i(f_i) = \{f_i|h : h \in H\}$ and $F(f) = \{f|h : h \in H\}$.

Given a strategy of player $i$, $f_i \in F_i$, we say that $f_i$ is a $m$-memory strategy if $m$ is the smallest integer satisfying $f_i(h) = f_i(\bar{h})$ for all $h, \bar{h} \in H$ with $T^m(h) = T^m(\bar{h})$. If $f_i$ is a strategy with $m$-memory, we write $\text{rec}(f_i) = m$.

Any strategy $f \in F$ induces an outcome at any date as follows: $\pi^1(f) = f(H_0)$ and
$\pi^t(f) = f(\pi^1(f), \ldots, \pi^{t-1}(f))$ for any $t > 1$. Denote the set of outcome paths by $\Pi = S \times S \times \cdots$ and define the outcome path induced by any strategy profile $f \in F$ by $\pi(f) = \{\pi^1(f), \pi^2(f), \ldots\} \in \Pi$.

We assume that all agents discount the future returns by a common discount factor $\delta \in (0, 1)$. Thus the payoff in the supergame $G^\infty(\delta)$ of $G$ is given by

$$U_i(f) = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i(\pi^t(f)).$$

Also, for any $\pi \in \Pi$, $t \in \mathbb{N}$, and $i \in N$, let $V_i^t(\pi) = (1-\delta) \sum_{r=t}^{\infty} \delta^{r-t} u^i(\pi^r)$ be the continuation payoff of player $i$ at date $t$ if the outcome path $\pi$ is played. For simplicity, we write $V_i(\pi)$ instead of $V_i^1(\pi)$.

A strategy vector $f \in F$ is a Nash equilibrium of $G^\infty(\delta)$ if $U_i(f) \geq U_i(\hat{f}_i, f_{-i})$ for all $i \in N$ and all $\hat{f}_i \in F_i$. A strategy vector $f \in F$ is a SPE of $G^\infty(\delta)$ if $f|h$ is a Nash equilibrium for all $h \in H$. An outcome path $\pi$ is a subgame perfect outcome path if there exists a SPE $f$ such that $\pi = \pi(f)$.

We also define a $M$ -- memory SPE as a SPE with the additional property that it has $M$ -- memory.

### 3 Two-player games

Our main result for 2-player game is the following folk theorem.

**Theorem 1** Let $G$ be a 2-player game with $U^0 \neq \emptyset$. Then, for all $\varepsilon > 0$, there exists $M \in \mathbb{N}$ and $\delta^* \in (0, 1)$ such that for all $u \in U$ and $\delta \geq \delta^*$, there exists a time-independent, subgame
perfect equilibrium strategy $f$ of $G(\delta)$ with $\text{rec}(f) \leq M$ and $\|U(f) - u\| < \varepsilon$.

**Proof.** Let $\varepsilon > 0$ be given and consider $\bar{u} \in \mathcal{U}^0$. Then, there is $\zeta > 0$ such that $\bar{u} \in \mathcal{U}^\zeta$. Let $B > \max_{w \in \text{co}(u(A)) \setminus \mathcal{U}^\zeta} \|w - \bar{u}\| > 0$ (note that $\text{co}(u(A)) \setminus \mathcal{U}^\zeta$ is compact). Then, given $u \in \mathcal{U}$, define $x = \lambda \bar{u} + (1 - \lambda)u$ with $\lambda = \varepsilon/B$. Then, it is easy to check that $\|x - u\| < \varepsilon$ and $x \in \mathcal{U}^{c \varepsilon}$ with $c = \zeta/B$. Let $M$ be given by Lemma 1 corresponding to $\eta = c \varepsilon$. Thus, the conclusion follows from Lemma 1.

As shown above, Theorem 1 follows from the following Lemma.

**Lemma 1** Let $G$ be a 2-player game with $\mathcal{U}^0 \neq \emptyset$. Then, for all $\varepsilon, \eta > 0$, there exists $M \in \mathbb{N}$ and $\delta^* \in (0, 1)$ such that for all $u \in \mathcal{U}^0$ and $\delta \geq \delta^*$, there exists a $M$-memory SPE $f$ of $G(\delta)$ with $\|U(f) - u\| < \varepsilon$.

**Proof.** Clearly, $v_i \geq u_i(\bar{m})$. For convenience, we normalize payoffs so that $u_i(\bar{m}) = 0$ for both $i = 1, 2$. Let $B = \max_i \max_s |u_i(s)|$ and $s^* \in S$ be such that $s^*_i \neq \bar{m}_i$ for all $i$. Order $S$ and write $S = \{s_1, \ldots, s_r\}$ with $s_2 = s^*$ and $s_3 = \bar{m}$.

Define $0 < \gamma < \min\{\eta/3, \varepsilon/2\}$. Let $\mathcal{U}_k$ be the set of those vectors $w \in \mathbb{R}^2$ satisfying $w > v$ and

$$w = \sum_{s \in S} \frac{p_s u(s)}{k}$$

for some $(p_s)_{s \in S} \in \mathbb{N}_{++}$ such that $\sum_{s \in S} p_s = k$ and $p_2 \geq 2$. Let $K \in \mathbb{N}$ be such that

$$D \subseteq \bigcup_{x \in \mathcal{U}_K} B_\gamma(x) \text{ and } \frac{2K}{2K + 1} (\max_i v_i + \eta - \gamma) - \gamma > \max_i v_i + \eta - 3\gamma,$$

(1)
(this is possible since \( \mathcal{U}_k \) converges to \( \mathcal{U} \) in the Hausdorff distance and since \( \lim_{K \to \infty} 2K(\max_i v_i + \eta - \gamma)/(2K + 1) - \gamma = \max_i v_i + \eta - 2\gamma > \max_i v_i + \eta - 3\gamma \)),

\[
\hat{K} = 2K + 1,
\]

\( \beta > 0 \) such that

\[
\beta < \min \left\{ \gamma, \frac{\eta - \gamma}{2K} \right\},
\]

\( \xi > 0 \) such that

\[
2\xi < \min \left\{ \eta - 3\gamma, \frac{\eta - \gamma}{K} - 2\beta \right\},
\]

\( T' \in \mathbb{N} \) be such that

\[
T' > \frac{B}{\xi},
\]

\( R \in \mathbb{N} \) be such that

\[
R > \max \left\{ \frac{B\hat{K}}{\xi}, 2K \right\},
\]

\( T \in \mathbb{N} \) be such that

\[
T = T' + R + \hat{K},
\]

\( M \in \mathbb{N} \) be such that

\[
M = T,
\]
and \( \tilde{\delta} \in (0, 1) \) be such that for all \( \delta \in [\tilde{\delta}, 1) \)

\[
\frac{1 - \delta^R}{1 - \delta^K} > \frac{B}{\tilde{K} \xi},
\]

\[
\frac{1 - \delta^{T'+1}}{1 - \delta} > T',
\]

\[
\delta^T > \frac{B}{B + \xi},
\]

\[
\delta^R > \frac{B + \xi}{B + 2 \xi},
\]

\[
\sup_{x \in [-B,B]} \left| \frac{1 - \delta}{1 - \delta^j} \sum_{k=1}^J \delta^{k-1} x_k - \frac{1}{J} \sum_{k=1}^J x_k \right| < \beta \quad \text{for all } J \in \{K, \hat{K}\}
\]

and for all \( i = 1, 2 \), \( 1 \leq k \leq \hat{K} \) and all \( \hat{K} \)-length sequences \( \{a^1, \ldots, a^\hat{K}\} \) of actions with

\[
\sum_{k=1}^\hat{K} u_i(a^k) / \hat{K} > \xi
\]

\[
\sum_{t=k}^{\hat{K}} \delta^{t-1} u_i(a_t) > -\frac{(\hat{K} - k + 1)B}{\hat{K} \xi}
\]

and

\[
\sum_{t=1}^{\hat{K}} \delta^{t-1} u_i(a_t) > -(k - 1)B
\]

Note that such \( \tilde{\delta} \in (0, 1) \) exists since

\[
\lim_{\delta \to 1} \frac{1 - \delta^R}{1 - \delta^K} = \frac{R}{K} > \frac{B}{\tilde{K} \xi},
\]

\[
\lim_{\delta \to 1} \frac{1 - \delta^{T'+1}}{1 - \delta} = T' + 1 > T',
\]

\[
\lim_{\delta \to 1} \frac{\sum_{t=k}^{\hat{K}} \delta^{t-1} u_i(a_t)}{\sum_{t=1}^{\hat{K}} \delta^{t-1} u_i(a_t)} = \frac{\sum_{t=k}^{\hat{K}} u_i(a_t)}{\sum_{t=1}^{\hat{K}} u_i(a_t)} > -\frac{(\hat{K} - k + 1)B}{\hat{K} \xi}
\]

and

\[
\lim_{\delta \to 1} \frac{-\sum_{t=k}^{\hat{K}} \delta^{t-1} u_i(a_t)}{\sum_{t=1}^{\hat{K}} \delta^{t-1} u_i(a_t)} = \frac{-\sum_{t=k}^{\hat{K}} u_i(a_t)}{\sum_{t=1}^{\hat{K}} u_i(a_t)} > \frac{(k - 1)B}{\hat{K} \xi}
\]

for all \( i = 1, 2 \) and \( 1 \leq k \leq \hat{K} \).
Now fix any $\delta \geq \bar{\delta}$ and consider any $u \in U^\eta$. We will show that there is a $M$–memory SPE $f$ with $||U(f, \delta) - u|| < \varepsilon$.

By (1), let $p_1, \ldots, p_r$ be such that $p_k > 0$ for all $1 \leq k \leq r$, $p_2 \geq 2$, $\sum_{k=1}^{r} p_k = K$ and

$$\left| \sum_{k=1}^{r} \frac{p_k u(s_k)}{K} - u \right| < \gamma. \tag{16}$$

Let $u' = \sum_{k=1}^{r} \frac{p_k u(s_k)}{K}$ and $\pi$ consist of repetitions of the cycle

$$((s_1; p_1), (s_2; p_2), (s_3; p_3), (s_4; p_4), \ldots, (s_r; p_r)).$$

Then,

$$||V(\pi, \delta) - u|| \leq ||V(\pi, \delta) - u'|| + ||u' - u|| < 2\gamma < \varepsilon.$$

Thus, it is enough to show that there is a $M$–memory SPE $f$ with $\pi(f) = \pi$.

Let $\hat{\pi}$ consist of repetitions of

$$((s_2; 2p_2), (s_3; 2p_3 + 1), (s_1; 2p_1), (s_4; 2p_4), \ldots, (s_r; 2p_r))$$

and let

$$\hat{u} = \sum_{k=1}^{2K+1} \frac{u(\hat{\pi}^k)}{2K + 1} = \frac{2K}{2K + 1} u'.$$

We have that

$$V_i^t(\hat{\pi}) > \frac{2K u'_i}{2K + 1} - \beta > \frac{2K}{2K + 1} (u_i - \gamma) - \beta > \frac{2K}{2K + 1} (v_i + \eta - \gamma) - \gamma > v_i + \eta - 3\gamma > v_i + 2\xi \tag{17}$$

for all $i = 1, 2$ and $t \in \mathbb{N}$ (the first inequality follows from (16), the second since $u \in U^\eta$, the third since $\beta < \gamma$ by (3), the forth by (1) and the fifth by (4)). Note that

$$u'_i - \hat{u}_i = u'_i - \frac{2K}{2K + 1} u'_i = \frac{1}{K} u'_i > \frac{u_i - \gamma}{K} > \frac{\eta - \gamma}{K}. \tag{18}$$
for all $i = 1, 2$ (the inequalities follows by (16) and since $u \in \mathcal{U}$ and $v_i \geq 0$ for all $i$). Thus,

$$V_i^t(\pi) > u'_i - \beta > \hat{u}_i + \frac{\eta - \gamma}{K} - \beta > V_i^k(\hat{\pi}) + \frac{\eta - \gamma}{K} - 2\beta > V_i^k(\hat{\pi}) + 2\xi$$  \hspace{1cm} (19)

for all $i = 1, 2$ and $t, k \in \mathbb{N}$ (the first inequality follows from (13), the second from (18), the third from (13) and the forth by (4)).

We seek to define a common punishment path that consists of $\bar{m}$ for the first $T - R = T' + \hat{K}$ periods followed by a path $\hat{\pi}$ yielding a payoff bounded away from minmax for each player. Furthermore, $\hat{\pi}$ must be such that its initial payoff is the lowest possible (i.e., $V_i^t(\hat{\pi}, \delta) \geq V_i(\hat{\pi}, \delta)$ for all $i$ and $t$) to prevent players to deviate at latter stages.

We define $\hat{\pi}$ as follows: it consists of $\bar{m}$ for the first $R$ periods and then it consists of $\hat{\pi}$, i.e.,

$$\hat{\pi}^t = \begin{cases} 
\bar{m} & \text{if } t \leq R, \\
\hat{\pi}^{t-R} & \text{if } t > R
\end{cases}$$

Before proceeding with the construction of equilibrium strategy, we shall next establish the outcome $\hat{\pi}$ has the following “stick and carrot” property.

**Claim 1** For all $i = 1, 2$ and $t \in \mathbb{N}$, $V_i^t(\hat{\pi}, \delta) \geq V_i(\hat{\pi}, \delta)$.

**Proof of Claim 1** Recall that $\hat{u}_i > v_i + 2\xi > \xi$ and $u_i(\bar{m}) = 0$ for all $i$. Hence, for all $2 \leq t \leq R$,

$$V_i^t(\hat{\pi}, \delta) = \delta^{R+1-t} \frac{1 - \delta}{1 - \delta^K} \sum_{l=1}^{\hat{K}} \delta^{l-1} u_i(\hat{\pi}^l) > \delta^{R} \frac{1 - \delta}{1 - \delta^K} \sum_{l=1}^{\hat{K}} \delta^{l-1} u_i(\hat{\pi}^l) = V_i(\hat{\pi}, \delta).$$

Consider $t > R$ and let $1 \leq k \leq \hat{K}$ be such that $\pi^t = \hat{\pi}^k$. Then,

$$V_i^t(\hat{\pi}, \delta) = \frac{1 - \delta}{1 - \delta^K} \left( \sum_{l=k}^{\hat{K}} \delta^{l-1} u_i(\hat{\pi}^l) + \delta^{\hat{K}-k+1} \sum_{l=1}^{k-1} \delta^{l-1} u_i(\hat{\pi}^l) \right).$$
Hence, in order to show that \( V^l_i(\tilde{\pi}, \delta) \geq V^l_i(\tilde{\pi}, \delta) \), it suffices to show that

\[
\sum_{l=k}^{K} \delta^{l-1} u_i(\tilde{\pi}^l) + \delta^{K-k+1} \sum_{l=1}^{k-1} \delta^{l-1} u_i(\tilde{\pi}^l) > \delta^R \sum_{l=1}^{K} \delta^{l-1} u_i(\tilde{\pi}^l),
\]

or equivalently,

\[
\delta^R < 1 + \frac{\sum_{l=k}^{K} \delta^{l-1} u_i(\tilde{\pi}^l) + \delta^{K-k+1} \sum_{l=1}^{k-1} \delta^{l-1} u_i(\tilde{\pi}^l) - \sum_{l=1}^{K} \delta^{l-1} u_i(\tilde{\pi}^l)}{\sum_{l=1}^{K} \delta^{l-1} u_i(\tilde{\pi}^l)}.
\]

Clearly, (20) holds if \( k = 1 \). If \( k \neq 1 \), it follows that

\[
\sum_{l=k}^{K} \delta^{l-1} u_i(\tilde{\pi}^l) + \delta^{K-k+1} \sum_{l=1}^{k-1} \delta^{l-1} u_i(\tilde{\pi}^l) - \sum_{l=1}^{K} \delta^{l-1} u_i(\tilde{\pi}^l) =
\]

\[
(1 - \delta^{k-1}) \sum_{l=k}^{K} \delta^{l-1} u_i(\tilde{\pi}^l) - (1 - \delta^{K-k+1}) \sum_{l=1}^{k-1} \delta^{l-1} u_i(\tilde{\pi}^l).
\]

Hence, (20) becomes

\[
\delta^R < 1 + (1 - \delta^{k-1}) \frac{\sum_{l=k}^{K} \delta^{l-1} u_i(\tilde{\pi}^l)}{\sum_{l=1}^{K} \delta^{l-1} u_i(\tilde{\pi}^l)} + (1 - \delta^{K-k+1}) \frac{\sum_{l=1}^{k-1} \delta^{l-1} u_i(\tilde{\pi}^l)}{\sum_{l=1}^{K} \delta^{l-1} u_i(\tilde{\pi}^l)}.
\]

Since \( \delta \geq \bar{\delta} \), it follows from (14) and (15) that

\[
(1 - \delta^{k-1}) \frac{\sum_{l=k}^{K} \delta^{l-1} u_i(\tilde{\pi}^l)}{\sum_{l=1}^{K} \delta^{l-1} u_i(\tilde{\pi}^l)} + (1 - \delta^{K-k+1}) \frac{\sum_{l=1}^{k-1} \delta^{l-1} u_i(\tilde{\pi}^l)}{\sum_{l=1}^{K} \delta^{l-1} u_i(\tilde{\pi}^l)} >
\]

\[
(1 - \delta^{k-1}) \frac{(\hat{K} - k + 1)B}{\hat{K} \xi} + (1 - \delta^{K-k+1}) \frac{-(k - 1)B}{\hat{K} \xi}.
\]

Therefore, it is enough to show that

\[
1 - \delta^R > (1 - \delta^{k-1}) \frac{(\hat{K} - k + 1)B}{\hat{K} \xi} + (1 - \delta^{K-k+1}) \frac{(k - 1)B}{\hat{K} \xi}.
\]

Clearly, it suffices to establish that

\[
1 - \delta^R > (1 - \delta^{\hat{K}})(\hat{K} - k + 1)B / \hat{K} \xi + (1 - \delta^{\hat{K}})(k - 1)B / \hat{K} \xi =
\]

\[
(1 - \delta^{\hat{K}})B / \xi,
\]

which follows from (9).
For convenience, define $\theta_i = V_i(\hat{\pi}, \delta)$ for all $i$. Then, $V_i(\hat{\pi}) > \theta_i = \delta^R V_i(\hat{\pi}) > \delta^R (v_i + 2\xi) > v_i + \xi$ for all $i$, where the second inequality follows from (17) and the third from (12). Hence,

$$V_i(\hat{\pi}) > \theta_i > v_i + \xi \text{ for all } i. \tag{24}$$

Since $M = T \geq R + \hat{K} > 2K + 2K + 1 = 4K + 1$, it follows by Lemma 2 that there is a $M$–memory strategy $f$ that implements $\pi$ and satisfies the following properties: (a) it has the one-shot deviation property, (b) for all $h \in H$, either

$$\pi(f|h) = (\pi_t, \pi_{t+1}, \ldots) \text{ for some } t \geq 1, \tag{25}$$

or

$$\pi(f|h) = (\hat{\pi}_t, \hat{\pi}_{t+1}, \ldots) \text{ for some } t \geq 1, \tag{26}$$

or

$$\pi(f|h) = ((\bar{m}; M - t), \hat{\pi}_1, \hat{\pi}_2, \ldots) \text{ for some } 0 \leq t < M \tag{27}$$

and (c) for all $h \in H$, a single-player, one-shot deviation $(g_i, f_{-i}|h)$ from $f|h$ leads to $\pi(g_i, f_{-i}|h) = ((\bar{m}; k), \hat{\pi}_1, \hat{\pi}_2, \ldots)$ with $k \geq M - K = T' + R$ when $\pi(f|h)$ satisfies either (25) or (26) and with $k = M = T$ if it satisfies (27).

We next show that $f$ is SPE. Since $f$ has the one-shot deviation property, it is enough to show that one-shot deviations are not profitable.

Let $h \in H$ be such that $\pi(f|h) = (\pi_t, \pi_{t+1}, \ldots)$ for some $t \in \{1, \ldots, K\}$ and assume that player $i$ deviates from $\pi_i^t$ and plays $s_i$ instead. Since $\bar{m}$ will be played at least $M - \hat{K} = T' + R$,
we have that

\[ V_t^i(\pi) - ((1 - \delta)u_i(s_i, \pi_{-i}^t) + \delta V_i(\pi(f|h \cdot (s_i, \pi_{-i}^t)))) \geq \]

(28)

\[ V_t^i(\pi) - ((1 - \delta)u_i(s_i, \pi_{-i}^t) + \delta^{T'+1} \theta_i) \geq \]

(29)

\[ V_t^i(\pi) - (1 - \delta)B + \delta^{T'+1} V_t^i(\pi) \] .

(30)

(The last inequality in the above follows from \([19]\) and \([24]\)). Put \(\alpha_i = V_t^i(\pi) > v_i + 4\xi > \xi\); since \(\alpha_i > (1 - \delta)B + \delta^{T'+1}\alpha_i\) is equivalent to \(\frac{1 - \delta^{T'+1}}{1 - \delta} > \frac{B}{\alpha_i}\), and

\[
\frac{1 - \delta^{T'+1}}{1 - \delta} > T' > \frac{B}{\xi} \geq \frac{B}{\alpha_i}
\]

(which holds by \([5]\) and \([10]\)), it follows that

\[ V_t^i(\pi) > (1 - \delta)u_i(s_i, \pi_{-i}^t) + \delta V_i(\pi(f|h \cdot (s_i, \pi_{-i}^t))). \]

Second, consider \(h \in H\) such that \(\pi(f|h) = (\hat{\pi}_t, \hat{\pi}_{t+1}, \ldots)\) for some \(t \in \{1, \ldots, K\}\) and assume that player \(i\) deviates from \(\hat{\pi}_t^i\) and plays \(s_i\) instead. Since \(\bar{m}\) will be played at least \(M - \tilde{K} = T' + R\), we have that

\[ V_t^i(\hat{\pi}) - ((1 - \delta)u_i(s_i, \pi_{-i}^t) + \delta V_i(\pi(f|h \cdot (s_i, \pi_{-i}^t)))) \geq \]

(31)

\[ V_t^i(\hat{\pi}) - ((1 - \delta)u_i(s_i, \pi_{-i}^t) + \delta^{T'+1} \theta_i) \geq \]

(32)

\[ V_t^i(\hat{\pi}) - (1 - \delta)B + \delta^{T'+1} V_t^i(\hat{\pi}) \] .

(33)

(The last inequality in the above follows from the fact that \(V_t^i(\hat{\pi}) = V_t^{i+R}(\tilde{\pi}) \geq V_i(\tilde{\pi}) = \theta_i\)). Since \(V_t^i(\hat{\pi}) > v_i + 2\xi > \xi\), we can use the same argument used above to conclude that

\[ V_t^i(\hat{\pi}) > (1 - \delta)u_i(s_i, \pi_{-i}^t) + \delta V_i(\pi(f|h \cdot (s_i, \pi_{-i}^t))). \]
Finally, consider \( h \in H \) such that \( \pi(h) = ((\bar{m}; M-t), \hat{\pi}_1, \hat{\pi}_2, \ldots) \) for some \( 0 \leq t < M \).

Recall that \( M = T \). If player \( i \) deviates from \( \bar{m} \) by playing \( s_i \), it follows that

\[
V_i(\pi(f|h)) - ((1 - \delta)v_i(s_i, \bar{m}_{-i}) + \delta V_i(\pi(f|h \cdot (s_i, \bar{m}_{-i})))) \geq (34)
\]

\[
\delta^{T-t}V_i(\hat{\pi}) - ((1 - \delta)v_i + \delta^{T+1}V_i(\hat{\pi})) \geq (35)
\]

\[
\delta^T V_i(\hat{\pi}) - ((1 - \delta)v_i + \delta^{T+1}V_i(\hat{\pi})) \geq (36)
\]

Put \( \alpha_i = V_i(\hat{\pi}) > v_i + 2\xi \) by (17); since \( \delta^T \alpha_i > (1 - \delta)v_i + \delta^{T+1}\alpha_i \) is equivalent to \( \delta^T > \frac{v_i}{\alpha_i} \), and

\[
\delta^T > \frac{B}{B + \xi} \geq \frac{v_i}{v_i + \xi} > \frac{v_i}{\alpha_i}
\]

(which holds by (11)), it follows that

\[
V_i(\pi(f|h)) > (1 - \delta)v_i(s_i, \bar{m}_{-i}) + \delta V_i(\pi(f|h \cdot (s_i, \bar{m}_{-i}))).
\]

Let \( f \) be a strategy and \( i \in N \). Then \( g_i \in F_i \) is a one-shot deviation by player \( i \) from \( f \) if there exists \( h \in H \) such that \( g_i(h) \neq f_i(h) \) and \( g_i(h') = f_i(h') \) for all \( h' \in H \setminus \{h\} \). Let \( G_i(f) \) denote the set of all one-shot deviations by player \( i \) from \( f \). A strategy \( f \) has the one-shot deviation property if for all \( i \in N \),

\[
U_i(f_i, f_{-i}) \geq U_i(g_i, f_{-i}) \quad \text{for all } g_i \in G_i(f)
\]

implies that

\[
U_i(f_i, f_{-i}) \geq U_i(g_i, f_{-i}) \quad \text{for all } g_i \in F_i.
\]

Let \( s^* \in S \) be such that \( s^*_i \neq \bar{m}_i \) for all \( i = 1, 2 \) and order \( S \) so that \( S = \{s_1, \ldots, s_r\} \) with \( s_2 = s^* \) and \( s_3 = \bar{m} \).
Lemma 2 Let $\pi$ consist on the repetition of the cycle $((s_1;p_1),(s_2;p_2),\ldots,(s_r;p_r))$ with $p_2 \geq 2$, $p_j > 0$ for all $j = 2,\ldots,r$ and $\sum_{j=1}^r p_j = K$ and $\hat{\pi}$ consist on the repetition of the cycle $((s_2;\hat{p}_2),(s_3;\hat{p}_3),(s_1;\hat{p}_1),(s_4;\hat{p}_4),\ldots,(s_r;\hat{p}_r))$ with $\hat{p}_3 = 2p_3 + 1$ and $\hat{p}_j = 2p_j$ for all $j \neq 3$.

If $M > 4K + 1$, then there is a $M$-memory strategy $f$ that implements $\pi$ and satisfies the following properties: (a) it has the one-shot deviation property, (b) for all $h \in H$, either

$$\pi(f|h) = (\pi_t,\pi_{t+1},\ldots) \text{ for some } t \geq 1,$$

(37)

or

$$\pi(f|h) = (\hat{\pi}_t,\hat{\pi}_{t+1},\ldots) \text{ for some } t \geq 1,$$

(38)

or

$$\pi(f|h) = ((\bar{m};M-t),\hat{\pi}_1,\hat{\pi}_2,\ldots) \text{ for some } 0 \leq t < M$$

(39)

and (c) for all $h \in H$, a single-player, one-shot deviation $(g_i,f_{-i}|h)$ from $f|h$ leads to $\pi(g_i,f_{-i}|h) = ((\bar{m};k),\hat{\pi}_1,\hat{\pi}_2,\ldots)$ with $k \geq M - (2K + 1)$ when $\pi(f|h)$ satisfies either (37) or (38) and with $k = M$ if it satisfies (39).

Proof. Define the strategy $f$ as follows. In the definition of $f$, and in what follows, the condition

$$T^M(h) = (a_1,\ldots,a_{M-q-k},(\bar{m};q),\pi_1,\ldots,\pi_k)$$

means that there exist $a_1,\ldots,a_{M-q-k} \in S$ such that the equality holds. The same applies to other conditions similar to that. Let $h \in H$. Consider first the case where the length of
\(T^M(h)\) is equal to \(M\). If

\[
T^M(h) = (a_1, \ldots, a_{M-k}, \pi_1, \ldots, \pi_k) \text{ for some } p_1 + p_2 \leq k \leq M,
\]

then \(f(h) = \pi_{k+1}\). If

\[
T^M(h) = (a_1, \ldots, a_{M-k}, \hat{\pi}_1, \ldots, \hat{\pi}_k) \text{ for some } \hat{p}_2 \leq k \leq M,
\]

then \(f(h) = \hat{\pi}_{k+1}\). If

\[
T^M(h) = ((\bar{m}; M - k), (s^*, k)) \text{ for some } 0 \leq k < \hat{p}_2,
\]

then \(f(h) = s^*\). Otherwise, \(f(h) = \bar{m}\).

Finally, consider the case where the length of \(T^M(h)\) is less than \(M\) (and so \(T^M(h) = h\)). If

\[
T^k(h) = (\pi_1, \ldots, \pi_k) \text{ for some } k \geq p_1 + p_2,
\]

then \(f(h) = \pi_{k+1}\); if

\[
T^k(h) = (\hat{\pi}_1, \ldots, \hat{\pi}_k) \text{ for some } k \geq \hat{p}_2,
\]

then \(f(h) = \hat{\pi}_{k+1}\); otherwise, \(f(h) = \bar{m}\).

Clearly, \(f\) is a \(M\) – memory strategy since, for all \(h \in H\), \(f(h)\) depends only on \(T^M(h)\).

Define \(H^1\) as the set of histories \(h \in H\) satisfying (40) or (44), \(H^2\) those satisfying (41) or (44) and \(H^3\) those satisfying (42). Also, let \(H_E = H^1 \cup H^2 \cup H^3\) and \(H_P = H \setminus H_E\).

We first note that \(f\) is well defined. If \(h\) is such that

\[
T^M(h) = (a_1, \ldots, a_{M-k}, \pi_1, \ldots, \pi_k) = (a_1, \ldots, a_{M-k'}, \pi_1, \ldots, \pi_{k'})
\]
for some $k, k' \geq p_1 + p_2$, then either $k = k'$ or $k = k' + \alpha K$ for some $\alpha \in \mathbb{N}$, implying that $\pi_{k+1} = \pi_{k'+1}$. A similar argument shows that if $T^M(h) = (a_1, \ldots, a_{M-k'}, \hat{\pi}_1, \ldots, \hat{\pi}_k) = (a_1, \ldots, a_{M-k'}, \hat{\pi}_1, \ldots, \hat{\pi}_k')$ for some $k, k' \geq \hat{p}_2$, then either $k = k'$ or $k = k' + \alpha \hat{K}$ for some $\alpha \in \mathbb{N}$, and so $\hat{\pi}_{k+1} = \hat{\pi}_{k'+1}$.

It is clear that $H^1 \cap H^3 = \emptyset$ and $H^1 \cap H^3 = \emptyset$. Furthermore,

Claim 2 $H^1 \cap H^2 = \emptyset$.

Proof. We will show that if $h \in H^2$, then $h \notin H^1$. Let $k \geq \hat{p}_2$ be such that $T^k(h) = (\hat{p}_1, \ldots, \hat{p}_k)$. Thus, there are $\hat{p}_2$ consecutive occurrences of $s_2$ in $T^k(h)$. Thus, if there exists $j \geq p_1 + p_2$ such that $T^j(h) = (\pi_1, \ldots, \pi_j)$, it must be that $k - \hat{p}_1 \geq p_1 + p_2$ and $(\hat{\pi}_{\hat{p}_1+1}, \ldots, \hat{\pi}_k) = (a'_1, \ldots, a_{k-\hat{p}_1-j}, \pi_1, \ldots, \pi_j)$. But this implies that $s_2$ follows from $s_1$ in $\hat{\pi}$, a contradiction. ■

The following claim states that if players play according to $f$ starting from an history $h \in H^1$, then they play the equilibrium path.

Claim 3 If $h \in H^1$, then $\pi(f|h) = (\pi_k, \pi_{k+1}, \ldots)$ for some $1 \leq k \leq K$.

Proof. This follows from the fact that if $T^k(h) = (\pi_1, \ldots, \pi_k)$ for some $M > k \geq p_1 + p_2$, then $T^{k+1}(h \cdot f(h)) = (\pi_1, \ldots, \pi_{k+1})$ and, clearly, $k + 1 \geq p_1 + p_2$. Moreover, if $k = M$, then $T^M(h \cdot f(h)) = (\pi_2, \ldots, \pi_K, \pi_1, \ldots)$ and so $T^{M-K+1}(h \cdot f(h)) = (\pi_1, \ldots, \pi_{M-K+1})$ with $M - K + 1 > p_1 + p_2$. ■

The following claim states that if players play according to $f$ starting from an history $h \in H^2 \cup H^2$, then they play the path $\hat{\pi}$.
Claim 4 If $h \in H^2 \cup H^3$, then $\pi(f|h) = (\hat{\pi}_k, \hat{\pi}_{k+1}, \ldots)$ for some $1 \leq k \leq 2K + 1$.

Proof. Note first that if $h \in H^3$, then it cannot satisfy neither (40) nor (41). Hence, players will play $s_2 = s^*$ for $p_2 - k$ periods, after which the history is such that $T^M(h \cdot (s^*; p_2 - k)) = ((\bar{m}; M - p_2), (s^*, p_2)) \in H^2$.

Hence, it suffices to show that if $h \in H^2$, then $h \cdot f(h) \in H^2$. It is clear that $h \cdot f(h)$ satisfies (41): If $T^k(h) = (\hat{\pi}_1, \ldots, \hat{\pi}_k)$ for some $M > k \geq \hat{p}_2$, then $T^{k+1}(h \cdot f(h)) = (\hat{\pi}_1, \ldots, \hat{\pi}_{k+1})$ and, clearly, $k + 1 \geq \hat{p}_2$. Moreover, if $k = M$, then $T^M(h \cdot f(h)) = (\hat{\pi}_2, \ldots, \hat{\pi}_{2K+1}, \hat{\pi}_1, \ldots)$ and so $T^{M-2K}(h \cdot f(h)) = (\hat{\pi}_1, \ldots, \hat{\pi}_{M-2K})$ with $M - 2K > 2K + 1 > \hat{p}_2$. ■

The next claim shows that if players play according to $f$ starting from an history $h \in H_P$, then they play $\bar{m}$ until players observe $(\bar{m}; M)$, after which they play $\hat{\pi}$.

Claim 5 If $h \in H_P$ and $T^M(h) = (a_1, \ldots, a_{M-k}, (\bar{m}; k))$ for some $0 \leq k < M$ and $a_{M-k} \neq \bar{m}$, then

$$\pi(f|h) = ((\bar{m}; M - k), \hat{\pi}_1, \hat{\pi}_2, \ldots).$$

Proof. Note that it is enough to show that $h \cdot (\bar{m}; l) \in H_P$ for all $l = 0, \ldots, M - k - 1$ since then (42) applies. Also, note that the conclusion holds for $l = 0$ by assumption.

Fix $l \in \{1, \ldots, M - k - 1\}$ and assume that $h \cdot (\bar{m}; l-1) \in H_P$. If $h \cdot (\bar{m}; l) \in H^1$, it would follow that $T^k(h \cdot (\bar{m}; l)) = (\pi_1, \ldots, \pi_k)$ for some $k > p_1 + p_2$ since $\pi_k = T^1(h \cdot (\bar{m}; l)) = \bar{m}$. Hence, $T^{k-1}(h \cdot (\bar{m}; l-1)) = (\pi_1, \ldots, \pi_{k-1})$ with $k-1 \geq p_1 + p_2$ and so $h \in H^1$, a contradiction. An analogous argument shows that $h \cdot (\bar{m}; l) \in H^2$, then $h \cdot (\bar{m}; l-1) \in H^2$, a contradiction.

If $h \cdot (\bar{m}; l) \in H^3$, then it must be that the length of $T^M(h)$ is $M$ and $T^M(h \cdot (\bar{m}; l)) = (\bar{m}; M)$ since $l \geq 1$ (i.e., $T^1(h \cdot (\bar{m}; l)) = \bar{m}$). Since $T^M(h \cdot (\bar{m}; l)) = (a_{l+1}, \ldots, a_{M-k}, (\bar{m}; k+l))$
and $k + l \leq M - 1$, this implies $a_{M-k} = \bar{m}$, a contradiction. ■

It follows from Claims 3, 4 and 5 that property (b) holds. Furthermore, it follows from Claim 3 that $f$ implements $\pi$.

The following claim shows the consequence of a single player deviation from $f$ at an history $h \in H_E$.

**Claim 6** If $h \in H_E$, $a_i \neq f_i(h)$ and $a_{-i} = f_{-i}(h)$ for some $i \in \{1, 2\}$, then $h \cdot a \in H_P$.

**Proof.** Suppose, in order to reach a contradiction, that $h \cdot a \in H^1$. Then, $T^k(h \cdot a) = (\pi_1, \ldots, \pi_k)$ for some $k \geq p_1 + p_2$. Hence, $a = \pi_k$ and $T^{k-1}(h) = (\pi_1, \ldots, \pi_k)$, $k \geq p_1 + p_2$. If $k > p_1 + p_2$, then $h \in H^1$ and $f(h) = \pi_k = a$, a contradiction. Thus, $k = p_1 + p_2$, $a = s^*$ and $T^{p_1+p_2-1}(h) = ((s_1; p_1), (s^*; p_2 - 1))$. Since $h \in H_E$ and $T^1(h) = s^*$ (because $p_2 \geq 2$), it follows that $f(h) = s^*$ or $f(h) = \bar{m}$. Thus, either $f(h) = a$ or $f_j(h) \neq a_j$ for all $j = 1, 2$, and we reach a contradiction in both cases.

Suppose next that $h \cdot a \in H^2$. Then, $T^k(h \cdot a) = (\hat{\pi}_1, \ldots, \hat{\pi}_k)$ for some $k \geq \hat{p}_2$. Hence, $a = \hat{\pi}_k$ and $T^{k-1}(h) = (\hat{\pi}_1, \ldots, \hat{\pi}_k)$. If $k > \hat{p}_2$, then $h \in H^2$ and so $f(h) = \hat{\pi}_k = a$, a contradiction. If $k = \hat{p}_2$, then $a = s^*$ and

$$T^{\hat{p}_2-1}(h) = (s^*; \hat{p}_2 - 1).$$

Since $h \in H_E$ and $T^1(h) = s^*$ (because $\hat{p}_2 \geq 2$) it follows that $f(h) = s^*$ or $f(h) = \bar{m}$. Thus, either $f(h) = a$ or $f_j(h) \neq a_j$ for all $j = 1, 2$, and we reach a contradiction in both cases.

Finally, suppose that $h \cdot a \in H^3$ and so the length of $T^M(h)$ equals $M$ and $T^M(h \cdot a) = ((\bar{m}; M - k), (s^*; k))$ for some $0 \leq k < \hat{p}_2$. If $k = 0$, then $T^M(h \cdot a) = (\bar{m}; M)$, $a = \bar{m}$ and
$T^M(h) = (a_1, (\bar{m}; M-1))$. Hence, we obtain $a_1 = \bar{m}$ since otherwise $h \in H_P$. Thus, $T^M(h) = (\bar{m}; M)$ and $f(h) = s^*$, which implies that $a_{-i} = \bar{m}_{-i} \neq s^*_{-i} = f_{-i}(h)$, a contradiction. Hence, it must be that $k > 0$. Then, $a = s^*$ and $T^M(h) = (a_1, (\bar{m}; M-k), (s^*; k-1))$. Since $k-1 < \hat{p}_2$, $h \in H_E$ implies that $a_1 = \bar{m}$. We then obtain $T^M(h) = ((\bar{m}; M - (k-1)), (s^*; k-1))$ and so $f(h) = s^* = a$, a contradiction. 

**Corollary 1** If $h \in H_E$, $a_i \neq f_i(h)$ and $a_{-i} = f_{-i}(h)$ for some $i \in \{1, 2\}$, then

$$
\pi(f|h \cdot a) = \begin{cases} 
(\bar{m}; M), \hat{\pi}_1, \hat{\pi}_2, \ldots) & \text{if } a \neq \bar{m}, \\
(\bar{m}; M-1), \hat{\pi}_1, \hat{\pi}_2, \ldots) & \text{if } a = \bar{m} \text{ and } T^1(h) \neq \bar{m}, \\
(\bar{m}; M - p_3 - 1), \hat{\pi}_1, \hat{\pi}_2, \ldots) & \text{if } a = T^1(h) = \bar{m} \text{ and } h \in H^1, \\
(\bar{m}; M - \hat{p}_3 - 1), \hat{\pi}_1, \hat{\pi}_2, \ldots) & \text{if } a = T^1(h) = \bar{m} \text{ and } h \in H^2.
\end{cases}
$$

**Proof.** Note first that $h \cdot a \in H_P$ by Claim 6 and so if $L$ denotes the length of $T^M(h)$ and $0 \leq k \leq L$ is the largest $0 \leq k' \leq L$ such that $T^{k'}(h \cdot a) = (\bar{m}; k')$, and so $T^L(h) = (a_1, \ldots, a_{L-k}, (\bar{m}; k))$ with $a_{L-k} \neq \bar{m}$, then $\pi(f|h \cdot a) = ((\bar{m}; M - k), \hat{\pi}_1, \hat{\pi}_2, \ldots)$ by Claim 5. So, if $a \neq \bar{m}$, then $k = 0$ and so $\pi(f|h \cdot a) = (\bar{m}; M), \hat{\pi}_1, \hat{\pi}_2, \ldots)$. If $a = \bar{m}$ and $T^1(h) \neq \bar{m}$, then $k = 1$ and so $\pi(f|h \cdot a) = ((\bar{m}; M - 1), \hat{\pi}_1, \hat{\pi}_2, \ldots)$.

Finally, consider the case $a = T^1(h) = \bar{m}$. Clearly, $f(h) \neq s^*$ (since $f_{-i}(h) = a_{-i} = \bar{m}_{-i} \neq s^*_{-i}$) and $f(h) \neq \bar{m}$ (since $f(h) \neq a = \bar{m}$). This rules out the possibility that $h \in H^3$. Since $h \in H_E$, then either $h \in H^1$ or $h \in H^2$. Let $L'$ denote the length of $T^M(h)$.

Consider first the case $h \in H^1$. Hence, we can write $T^k(h) = (\pi_1, \ldots, \pi_k)$ for some $k \geq p_1 + p_2$; we may assume that $k$ is maximal, i.e., there is no $p_1 + p_2 \leq k' \leq L'$ such that $k' > k$ and $T^{k'}(h) = (\pi_1, \ldots, \pi_{k'})$. Since $\pi_k = T^1(h) = \bar{m}$ and $\pi_{k+1} = f(h) \neq \bar{m}$, then it
must be that \( k \geq p_1 + p_2 + p_3; \pi_{k-p_3+1} = \cdots = \pi_k = \bar{m} \) and \( \pi_{k-p_2-p_3+1} = \cdots = \pi_{k-p_3} = s^* \).

Thus, \( T^L(h \cdot a) = (a'_1, \ldots, a'_{L-p_3-1}, (\bar{m}; p_3 + 1)) \) with \( a'_{L-p_3-1} \neq \bar{m} \). Hence, \( k = p_3 + 1 \) and so \( \pi(f|h \cdot a) = ((\bar{m}; M-p_3-1), \hat{\pi}_1, \hat{\pi}_2, \ldots) \).

The proof of the case \( h \in H^2 \) is analogous. If \( h \in H^2 \), then we can write \( T^k(h) = (\hat{\pi}_1, \ldots, \hat{\pi}_k) \) for some \( \hat{p}_2 \leq k \leq L' \); we may assume that \( k \) is maximal, i.e., there is no \( \hat{p}_2 \leq k' \leq L' \) such that \( k' > k \) and \( T^{k'}(h) = (\hat{\pi}_1, \ldots, \hat{\pi}_{k'}) \). Since \( \hat{\pi}_k = T^1(h) = \bar{m} \) and \( \hat{\pi}_{k+1} = f(h) \neq \bar{m} \), then it must be that \( k \geq \hat{p}_2+\hat{p}_3, \pi_{k-\hat{p}_3+1} = \cdots = \pi_k = \bar{m} \) and \( \pi_{k-\hat{p}_2-\hat{p}_3+1} = \cdots = \pi_{k-\hat{p}_3} = s^* \). Thus, \( T^k(h \cdot a) = (a'_1, \ldots, a'_{L-\hat{p}_3-1}, (\bar{m}; \hat{p}_3 + 1)) \) with \( a'_{L-\hat{p}_3-1} \neq \bar{m} \). Hence, \( k = \hat{p}_3 + 1 \) and so \( \pi(f|h \cdot a) = ((\bar{m}; M-\hat{p}_3-1), \hat{\pi}_1, \hat{\pi}_2, \ldots) \).

The following corollary is a consequence of Claims 5 and 7.

**Claim 7** If \( h \in H_P, a_i \neq f_i(h) \) and \( a_{i-1} = f_{i-1}(h) \) for some \( i \in \{1, 2\} \), then \( h \cdot a \in H_P \).

**Proof.** Since \( h \in H_P \), then \( f(h) = \bar{m} \). Thus, \( a \neq \bar{m} \) and \( a \neq s^* \). This implies that \( h \cdot a \) cannot belong to \( H^3 \). In fact, if \( T^M(h \cdot a) = ((\bar{m}; M-k), (s^*, k)) \) for some \( 0 \leq k < \hat{p}_2 \), then \( a = \bar{m} \) (if \( k = 0 \)) or \( a = s^* \) (if \( k > 0 \)), a contradiction in both cases.

Thus, if \( h \cdot a \in H^1 \), then \( T^k(h \cdot a) = (\pi_1, \ldots, \pi_k) \) with \( k \geq p_1 + p_2 \) and we can choose \( k \) to be maximal, i.e., there is no \( k' \geq p_1 + p_2 \) such that \( T^{k'}(h) = (\pi_1, \ldots, \pi_k') \) and \( k' > k \). Since \( \pi_k = a \neq \bar{m} \) and \( \pi_k = a \neq s^* \), it follows that \( k > p_1 + p_2 + p_3 \). Thus, \( T^{k-1}(h) = (\pi_1, \ldots, \pi_{k-1}) \) with \( k-1 \geq p_1 + p_2 \). Thus, \( h \in H_E \), a contradiction. The case when \( h \cdot a \in H^2 \) is analogous.

The following corollary is a consequence of Claims 5 and 7.
Corollary 2 If \( h \in H_P, \ a_i \neq f_i(h) \) and \( a_{-i} = f_{-i}(h) \) for some \( i \in \{1, 2\} \), then \( \pi(f|h \cdot a) = ((\bar{m}; M), \hat{\pi}_1, \hat{\pi}_2, \ldots) \).

Proof. Note first that \( h \cdot a \in H_P \) by Claim 6. Let \( k \geq 0 \) be the largest \( k' \geq 0 \) such that \( T^{k'}(h \cdot a) = (\bar{m}; k') \). Since \( h \in H_P \), then \( f(h) = \bar{m} \) and so \( a \neq \bar{m} \). Thus, \( k = 0 \), i.e., \( a_M \neq \bar{m} \), and so \( \pi(f|h \cdot a) = ((\bar{m}; M), \hat{\pi}_1, \hat{\pi}_2, \ldots) \) by Claim 5.

It follows from Corollaries 1 and 2 that property (c) holds.

For every player \( i \), player \( i \)'s maximization problem given \( f_{-i} \) is a stationary discounted Markov decision problem. Indeed, this follows from Claims 3, 4 and 5 and Corollaries 1 and 2 by letting the set of states be

\[ \{\pi_1, \ldots, \pi_K, \hat{\pi}_1, \ldots, \hat{\pi}_K, 1, 2, \ldots, M\} \]

and defining the following transition probabilities, \( \text{prob}(q'|q, s) \), where \( s \) is the current action, \( q \) is the current state and \( q' \) is the future state: for all \( 1 \leq k \leq K \), \( \text{prob}(\pi_{k+1 \mod K}|\pi_k, \pi_k) = 1 \) and if \( s_i \neq \pi_i^k \),

\[ \text{prob}(M|\pi_k, (s_i, \pi_{-i}^k)) = 1 \quad \text{if} \quad (s_i, \pi_{-i}^k) \neq \bar{m}, \]
\[ \text{prob}(M - 1|\pi_k, (s_i, \pi_{-i}^k)) = 1 \quad \text{if} \quad (s_i, \pi_{-i}^k) = \bar{m} \text{ and } \pi_k \neq \bar{m}, \]
\[ \text{prob}(M - p_3 - 1|\pi_k, (s_i, \pi_{-i}^k)) = 1 \quad \text{if} \quad (s_i, \pi_{-i}^k) = \bar{m} \quad \text{and} \quad \pi_k = \bar{m}; \]

for all \( 1 \leq k \leq \hat{K} \), \( \text{prob}(\hat{\pi}_{k+1 \mod \hat{K}}|\hat{\pi}_k, \hat{\pi}_k) = 1 \) and if \( s_i \neq \hat{\pi}_i^k \),

\[ \text{prob}(M|\hat{\pi}_k, (s_i, \hat{\pi}_{-i}^k)) = 1 \quad \text{if} \quad (s_i, \hat{\pi}_{-i}^k) \neq \bar{m}, \]
\[ \text{prob}(M - 1|\hat{\pi}_k, (s_i, \hat{\pi}_{-i}^k)) = 1 \quad \text{if} \quad (s_i, \hat{\pi}_{-i}^k) = \bar{m} \text{ and } \hat{\pi}_k \neq \bar{m}, \]
\[ \text{prob}(M - \hat{p}_3 - 1|\hat{\pi}_k, (s_i, \hat{\pi}_{-i}^k)) = 1 \quad \text{if} \quad (s_i, \hat{\pi}_{-i}^k) = \bar{m} \quad \text{and} \quad \hat{\pi}_k = \bar{m}; \]
and for all $1 \leq k \leq M$,

$$\text{prob}(k - 1|k, \bar{m}) = 1 \quad \text{if} \quad k > 1,$$

$$\text{prob}(\hat{\pi}_1|k, \bar{m}) = 1 \quad \text{if} \quad k = 1,$$

and if $s_i \neq \bar{m}_i$, $\text{prob}(M|k, (s_i, \bar{m}_{-i})) = 1$. Thus, analogously as in Abreu (1988, Proposition 1), $f$ has the one-shot deviation property. This completes the proof. ■

4 Concluding Remarks

This study shows that the Folk Theorem continues to hold with limited memory strategies and finite actions in every 2 player discounted repeated game satisfying the full dimensionality assumption.

We are currently working in extending this result to $n > 2$ player games. The techniques presented in this paper cannot be used in such an extension, because with the $R$ period common minmaxing construction it might be impossible to reward the punishers. Instead, we have to employ $R_i$ periods of minmaxing an agent $i$ who has singly deviated, and that phase must be followed by a path satisfying the required incentive conditions and which can be distinguished from not only the equilibrium outcome path, but also from the paths following the minmaxing phase of other players.
References


