האוניברסיטה העברית בירושלים
THE HEBREW UNIVERSITY OF JERUSALEM

UNCOUPLED AUTOMATA AND PURE NASH EQUILIBRIA

by

YAKOV BABICHENKO

Discussion Paper # 459 August 2007

מרכז לחקר הרציונליות

CENTER FOR THE STUDY OF RATIONALITY

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

Uncoupled Automata and Pure Nash Equilibria*

Yakov Babichenko†

August 2, 2007

Abstract

We study the problem of reaching Nash equilibria in multi-person games that are repeatedly played, under the assumption of uncoupledness: every player knows only his own payoff function. We consider strategies that can be implemented by finite-state automata, and characterize the minimal number of states needed in order to guarantee that a pure Nash equilibrium is reached in every game where such an equilibrium exists.

1 Introduction

We study the problem of reaching Nash equilibria in multi-person games, where the players play the same game repeatedly. The main assumption, called *uncoupledness* (see [1]), is that every player knows only his own utility function. The resulting play of the game yields an *uncoupled dynamic.*

Hart and Mas-Colell show in [1] that if the game is played in continuous time, and the moves of every player are deterministic, then uncoupled dynamics cannot always lead

*This is part of author's M.Sc. thesis. I wish to thank Sergiu Hart for his support and guidance, and Noam Nisan for helpful discussions.

†Center for Study of Rationality and The Institute of Mathematics, The Hebrew University of Jerusalem, 91904 Jerusalem, Israel.

to Nash equilibria. In [2] they show that the situation is different when stochastic moves are allowed and the game is played in discrete time: if the players know the history of play,[1] then there are uncoupled strategies that lead to a Nash equilibrium. The question is whether it is necessary to know the whole history in order to reach a Nash equilibrium. The answer is no. It was proved in [2], Theorems 4 and 5, that under the assumption of uncoupledness, convergence of the long-run empirical distribution of play to a (pure or mixed) Nash equilibrium can be guaranteed by using only the history of the last R periods of play, for some finite R. This is called a *finite-recall strategy*. Although finite-recall uncoupled strategies can guarantee convergence of the distribution of play to a Nash equilibrium, it is shown in [2], Theorem 6, that this *cannot* hold for the period-by-period behavior probabilities. If however, instead of finite recall one uses *finite memory* (e.g., finitely many periods of history but not necessarily the last ones), then the convergence of the behavior can be guaranteed as well ([2], Theorem 7).

This leads us to the study of uncoupled strategies with finite memory, i.e., *finite-state automata*. In this paper we deal with convergence to pure Nash equilibria in games which have such equilibria. In [2], Theorem 3, it is shown that in order to guarantee convergence to pure Nash equilibria one needs recall of size $R = 2$. Since finite recall is a special case of finite automata, the question we address here concerns the *minimum number of states required* for uncoupled finite automata to reach a pure Nash equilibrium. There are four classes of finite-state automata: the actions in every state can be deterministic (pure) or stochastic (mixed), and the transitions between states can be deterministic or stochastic. We will analyze each of the four classes in turn.

Section 2 presents the model, defines the relevant concepts and present the total results of the paper. Since the results are different for two-player games than for games with more than two players, we consider two-player games in Section 3 and n-player games for $n \geq 3$ in Section 4. In Sections 3 and 4 we discuss each of the four automata classes separately. Appendix A and Appendix B containing the proofs of Theorems 6 and 7.

[1] I.e., the past actions of all the players.

2 The Model

2.1 The Game

A basic static (one-shot) game Γ is given in strategic (or normal) form as follows. There are $n \geq 2$ players, denoted $i = 1, 2, ..., n$. Each player i has a finite set of pure actions $A^i = \{a^i_1, ..., a^i_{m^i}\}$; let $A := A^1 \times A^2 \times ... \times A^n$ be the set of action combinations. The payoff function (or utility function) of player i is a real-valued function $u^i : A \rightarrow \mathbb{R}$. The set of mixed (or randomized) actions of player i is the probability simplex over A^i, i.e., $\Delta(A^i) = \{x^i = (x^i(a^i_j))_{j=1,...,m^i} : \Sigma^{m^i}_{j=1} x^i(a^i_j) = 1$ and $x^i(a^i_j) \geq 0$ for $j = 1, ..., m^i\}$; payoff functions u^i are multilinearly extended, and so $u^i : \Delta(A^1) \times \Delta(A^2) \times ... \times \Delta(A^n) \rightarrow \mathbb{R}$.

We fix the set of players n and the action sets A^i, and identify a game by its n-tuple of payoff functions $U = (u^1, u^2, ..., u^n)$. Let $\mathcal{U}^i$ be the set of payoff functions of player i, and $\mathcal{U} := \mathcal{U}^1 \times ... \times \mathcal{U}^n$.

Denote the actions of all the players except player i by a^{-i}, i.e., $a^{-i} = (a^1_{j_1}, ..., a^{i-1}_{j_{i-1}}, a^{i+1}_{j_{i+1}}, ..., a^n_{j_n})$, and denote the set of actions of all the players except player i by $A^{-i} = A^1 \times ... \times A^{i-1} \times A^{i+1} \times ... \times A^n$. An action $a^i_j \in A^i$ will be called a best reply to a^{-i} if $u^i(a^i_j, a^{-i}) \geq u^i(a^i_k, a^{-i})$ for every $a^i_k \in A^i$. A pure Nash equilibrium is an action combination $a = (a^1_{j_1}, a^2_{j_2}, ..., a^n_{j_n}) \in A$, such that $a^i_{j_i}$ is a best reply to a^{-i} for all i.

For every game U, let $\widetilde{U} = (\widetilde{u}^1, \widetilde{u}^2, ..., \widetilde{u}^n)$ denote the resulting *best-reply game*, which is defined by

$$\widetilde{u}^i(a) = \begin{cases} 1 \text{ , if } a^i \text{ is a best reply to } a^{-i} \\ 0 \text{ , otherwise} \end{cases}$$

Note that a is a pure Nash equilibrium of U if and only if it is a pure Nash equilibrium of $\widetilde{U}$.

2.2 The Dynamic Setup

The dynamic setup consists of the repeated play, at discrete-time periods $t = 1, 2, ...,$ of the static game U. Let $a^i(t) \in A^i$ denote the action of player i at time t, and put

$a(t) = (a^1(t), a^2(t), ..., a^n(t)) \in A$ for the combination of actions at t. We assume that there is standard monitoring: at the end of period t each player i observes everyone's action, i.e., $a(t)$; when the choices are random, the players observe only the realized actions $a(t)$.

2.3 Automata

An *automaton*[2] *for player* i is a 4-tuple $\Lambda^i := < \Psi^i, \mathfrak{s}_0^i, f^i, g^i >$. Ψ^i is the set of *states*; $\mathfrak{s}_0^i \in \Psi^i$ is the *starting state*; $f^i : \Psi^i \to \Delta(A^i)$ is the *action function*; and $g^i : A \times \Psi^i \to \Delta(\Psi^i)$ is the *transition function*. Let $\mathcal{A}^i$ denote the set of all automata of player i. An automaton $\Lambda^i \in \mathcal{A}^i$ will be called *a pure-action automaton* if the actions in all states are pure, i.e.,[3] $\mathrm{Im}(f^i) \subset A^i$. Otherwise it will be called a *mixed-action automaton*. An automaton $\Lambda^i \in \mathcal{A}^i$ will be called a *deterministic-transition automaton* if all the transitions are deterministic, i.e., $\mathrm{Im}(g^i) \subset \Psi^i$. Otherwise it will be called a *stochastic-transition automaton*. An automaton $\Lambda^i \in \mathcal{A}^i$ will be called a k^i*-automaton* if it has k^i states, i.e., $|\Psi^i| = k^i$.

Let $(\Lambda^1, \Lambda^2, ..., \Lambda^n)$ be n automata, where Λ^i is a k^i-automaton for player i. The play proceeds as follows. At time $t = 1$ every player i is at his starting state $\mathfrak{s}_0^i$, and plays an action $a^i(1)$ according to the probability distribution $f^i(\mathfrak{s}_0^i)$. Let the realized actions of all the players be $a(1) := (a^1(1), ..., a^n(1))$. Then every player i moves to a new state according to the transition probabilities $g^i(a(1), \mathfrak{s}_0^i)$. Now assume that at time t player i is in state $s^i \in \Psi^i$, and hence at time $t+1$ player i plays an action $a^i(t)$ according to the probability distribution $f^i(s^i)$. The actions of all the players are $a(t+1)$, and every player i then moves to a new state according to the transition probabilities $g^i(a(t+1), s^i)$.

2.4 Strategy Mappings

Let $\varphi : \mathcal{U} \to \mathcal{A}^1 \times ... \times \mathcal{A}^n$ be a mapping that associates to every game $U = (u^1, ..., u^n) \in \mathcal{U}$ an n-tuple of automaton strategies $\varphi(U) = (\varphi^1(U), ..., \varphi^n(U))$ (with $\varphi^i(U)$ the automaton of player i). We will call the mapping φ *uncoupled* if, for each player i, the ith coordinate

[2] This is short for "a strategy implemented by an automaton."

[3] We identify A^i with the unit vectors in $\Delta(A^i)$.

φ^i of φ depends only on u^i, i.e., $\varphi^i : \mathcal{U}^i \to \mathcal{A}^i$ (rather than $\varphi^i : \mathcal{U} \to \mathcal{A}^i$). That is, φ^i associates to each payoff function $u^i \in \mathcal{U}^i$ of player i an automaton $\varphi^i(u^i) \in \mathcal{A}^i$, and

$$\varphi(U) = \varphi(u^1, u^2, ..., u^n) = (\varphi^1(u^1), \varphi^2(u^2), ..., \varphi^n(u^n)).$$

We will refer to $\varphi^i : \mathcal{U}^i \to \mathcal{A}^i$ as an *uncoupled strategy mapping to automata for player* i; thus φ^i "constructs" an automaton for player i by considering u^i only.[4] If $\varphi^i(u^i) \in \mathcal{A}^i$ is an automaton of size (at most) k^i for every payoff function $u^i \in \mathcal{U}^i$, we will say that φ^i is an *uncoupled strategy mapping to* k^i*-automata.*

Finally, we will say that the mapping φ is a *Pure Nash mapping*, or *PN-mapping* for short, if the strategies $\varphi(U)$ yield almost sure convergence of play to a pure Nash equilibrium in every game $U \in \mathcal{U}$ where such an equilibrium exists.

2.5 The Results

Clearly, every finite-recall strategy is in particular a finite-automaton strategy. Indeed, a strategy with recall of size R can be implemented by an automaton of size $|A|^R = \left(\prod_{i=1}^n m^i\right)^R$ (i.e. one state for each possible recall). Therefore, by Theorem 3 in [2], there is uncoupled PN-mapping to automata of size $\left(\prod_{i=1}^n m^i\right)^2$. The question we address here is whether there is uncoupled PN-mapping to automata with fewer states.

Our purpose is thus to characterize minimal numbers $k^1, ..., k^n$ such that there exists uncoupled PN-mapping where, for each i, the range is k^i-automata. We will analyze each of the four cases (pure or mixed-action automata, and deterministic or stochastic-transition automata) separately.

The results are the following:

For two-player games ($n = 2$):

[4] We assume that every player i knows his index i.

- There exists uncoupled PN-mapping to automata of sizes:

	mixed actions	pure actions
stochastic transitions	$\begin{cases} m^1 \\ m^2+1 \end{cases}$ or $\begin{cases} m^1+1 \\ m^2 \end{cases}$	$\begin{cases} m^1 \\ m^2+1 \end{cases}$ or $\begin{cases} m^1+1 \\ m^2 \end{cases}$
deterministic transitions	$\begin{cases} m^1+2 \\ m^2+2 \end{cases}$	$\begin{cases} 4m^1+O(1) \\ 4m^2+O(1) \end{cases}$

- There is no uncoupled PN-mapping to automata of sizes m^1, m^2.

For n-player games ($n \geq 3$):

- There exists uncoupled PN-mapping to automata of sizes:

	mixed actions	pure actions
stochastic transitions	$2m^i$	$2m^i$
deterministic transitions	$2m^i+3$	$O(m^i + n\log n)$

- Let $k^1, k^2, ..., k^n$ such that $\forall i = 1, ..., n : k^i < 2m^i$. Then there is no uncoupled PN-mapping to automata of sizes $k^1, k^2, ..., k^n$.

3 Two-Player Games

3.1 Stochastic transitions and mixed actions

We will show that there exists an uncoupled PN-mapping where the range for player 1 is (m^1+1)-automata and the range for player 2 is m^2-automata or, symmetrically, the range for player 1 is m^1-automata and the range for player 2 is (m^2+1)-automata. On the other hand, we will show that there is no PN-mapping where the ranges of the players are smaller.

Theorem 1 *Let $k^1 \geq m^1$ and $k^2 \geq m^2 + 1$. Then, for each player $i = 1, 2$, there exists an uncoupled strategy mapping to k^i-automata with stochastic transitions and mixed actions that guarantees almost sure convergence of play to a pure Nash equilibrium of the stage game in every game where such an equilibrium exists.*

Proof. We define the mapping φ as follows:

Given a game $U = (u^1, u^2)$, the automaton $\varphi^1(u^1) = \Lambda^1 \in \mathcal{A}^1$ is constructed as follows:

Denote $\varphi^1(u^1) = \Lambda^1 =< \Psi^1, \mathfrak{s}_0^1, f^1, g^1 >$ when Λ^1 is a m^1-automaton. We denote the states of Λ^1 by $\Psi^1 = \{s_1^1, ..., s_{m^1}^1\}$.

$\mathfrak{s}_0^1 := s_1^1$.

$f^1(s_i^1) := a_i^1 \equiv (0, ..., 0, \overset{i}{1}, 0, ..., 0)$.

$$g^1(a, s_i^1) = g^1((a_i^1, a^2), s_i^1) := \begin{cases} s_i^1 \equiv (0, ..., 0, \overset{i}{1}, 0, ..., 0) & \text{if } a_i^1 \text{ is a best reply to } a^2 \\ (\frac{1}{m^1}, ..., \frac{1}{m^1}) & \text{otherwise} \end{cases}$$

In state s_i^1 player 1 plays action a_i^1. He stays in this state if a_i^1 is a best reply to the action of player 2; otherwise he moves randomly to any one of the m^1 states with equal probability $\frac{1}{m^1}$. Note that whether an action of player 1 is a best reply or not depends only on his payoff function; therefore, Λ^1 depends on u^1 only.

Now we construct the automaton $\varphi^2(u^2) = \Lambda^2 \in \mathcal{A}^2$ as follows:

Denote $\Lambda^2 =< \Psi^2, \mathfrak{s}_0^2, f^2, g^2 >$ when Λ^2 is a $(m^2 + 1)$-automaton. We denote the states of Λ^2 by $\Psi^2 = \{s_0^2, s_1^2, ..., s_{m^1}^2\}$.

$\mathfrak{s}_0^2 := s_0^2$.

$$f^2(s_j^2) := \begin{cases} (\frac{1}{m^2}, ..., \frac{1}{m^2}) & j = 0 \\ a_j^2 & j \geq 1 \end{cases}$$

$$g^2(a, s_j^2) = g^2((a^1, a_j^2), s_j^2) := \begin{cases} (\frac{1}{m^2+1}, ..., \frac{1}{m^2+1}) & j = 0 \\ s_j^2 & j \geq 1 \text{ and } a_j^2 \text{ is a best reply to } a^1 \\ (\frac{1}{m^2+1}, ..., \frac{1}{m^2+1}) & j \geq 1 \text{ and } a_j^2 \text{ is not a best reply to } a^1 \end{cases}$$

In the state s_0^2, player 2 plays the mixed action $(\frac{1}{m^2}, ..., \frac{1}{m^2})$, and moves to any of the $m^2 + 1$ states with probability $\frac{1}{m^2+1}$.

In the states s_i^2, $i \geq 1$, player 2 plays action a_i^2. He stays in this state if a_i^2 is a best reply to the action of player 1; otherwise he moves to the state s_0^2.

Now we will prove that (φ^1, φ^2) is a PN-mapping.

We partition the space $\Psi^1 \times \Psi^2$ of the automata states into four regions:

$P_1 := \{(s_i^1, s_j^2), 1 \le i \le m^1, 1 \le j \le m^2 : \widetilde{u}^1(a_i^1, a_j^2) = 1, \widetilde{u}^2(a_i^1, a_j^2) = 1\}$; i.e., in this case (a_i^1, a_j^2) is a pure Nash equilibrium.

$P_2 := \{(s_i^1, s_0^2), 1 \le i \le m^1\}$.

$P_3 := \{(s_i^1, s_j^2), 1 \le i \le m^1, 1 \le j \le m^2 : \widetilde{u}^2(a_i^1, a_j^2) = 0\}$.

$P_4 := \{(s_i^1, s_j^2), 1 \le i \le m^1, 1 \le j \le m^2 : \widetilde{u}^1(a_i^1, a_j^2) = 0, \widetilde{u}^2(a_i^1, a_j^2) = 1\}$.

These four regions clearly cover the space $\Psi^1 \times \Psi^2$. In fact, player 2 can be in the state s_0^2 (P_2) or in any other state ($P_1 \cup P_3 \cup P_4$). If player 2 is not in the state s_0^2, then the action of player 2 can be a best reply ($P_1 \cup P_4$) or not (P_3). If it is a best reply, then the action of player 1 can be a best reply (P_1) or not (P_4)).

Players 1 and 2 stay at the same state if their action is a best reply; i.e., each state in P_1 is absorbing. We will prove that there is a positive probability of reaching a state from P_1, in finitely many periods, from any other state $s \in \Psi^1 \times \Psi^2$.

Definition 2 *An action $a_j^i \in A^i$ of player i will be called **dominant** if for every $a^{-i} \in A^{-i}$ a_j^i is a best reply to a^{-i}.*

$s = (s_i^1, s_0^2) \in P_2$: The actions are $(a_i^1, (\frac{1}{m^2}, ..., \frac{1}{m^2})) = (f^1(s_i^1), f^2(s_0^2))$. If a_i^1 is a dominant action, then denote by a_l^2 an action that is a best reply to a_i^1. Player 2 moves to s_l^2 with probability $\frac{1}{m^2+1}$. Then $(s_i^1, s_l^2) \in P_1$. If a_i^1 is not a dominant action, then denote by a_k^2 an action such that a_i^1 is not a best reply to it, and then with probability $\frac{1}{m^2}$ player 2 plays action a_k^2. Now both players move randomly over all their states and with positive probability they will get to P_1.

$s = (s_i^1, s_j^2) \in P_3$: The actions are (a_i^1, a_j^2); a_j^2 is not a best reply. Therefore, player 2 moves to s_0^2. Denote the state to which player 1 moves by s_k^1. Then $(s_k^1, s_0^2) \in P_2$.

$s = (s_i^1, s_j^2) \in P_4$: The actions are (a_i^1, a_j^2); a_j^2 is a best reply, and a_i^1 is not. Therefore, player 2 stays in s_j^2, whereas player 1 move to s_k^1 with probability $\frac{1}{m^1}$, where a_k^1 is a best reply of player 1 to a_j^2. Now either $(s_k^1, s_j^2) \in P_1$ or $(s_k^1, s_j^2) \in P_3$, depending on whether a_j^2 is a best reply to a_k^1.

In each of the above cases there is positive probability of reaching an absorbing state in P_1 in at most 3 steps. ■

Definition 3 *A game U will be called a **full game** if, for every action $a_j^i \in A^i$ of every player i, there exists $a^{-i} \in A^{-i}$ such that a_j^i is a best reply to a^{-i}.*

We prove a general result about n-player full games which will be useful in the sequel.

Lemma 4 *Let $\varphi = (\varphi^1, ..., \varphi^n)$ be an uncoupled strategy mapping that guarantees almost sure convergence of play to a pure Nash equilibrium of the stage game in every game where such an equilibrium exists. Then for every full game $U = (u_1, ..., u_n)$ and for every player i, there exist m^i nonempty sets of states $B_1^i, ..., B_{m^i}^i$ in Ψ^i (the set of states of the automaton $\varphi^i(u^i) = \Lambda^i$) such that in every state $s_k^i \in B_j^i$ player i plays a_j^i (with probability 1), and stays in B_j^i (with probability 1) if his action is a best reply to the actions of the other players.*

Proof. (By contradiction.) Assume that there exists a full game U s.t. for player 1 the set B_j^1 does not exist (or is empty). U is a full game, and so there exists $a^{-1} \in A^{-1}$ such that a_j^1 is a best reply to it. Consider the game $\overline{U} = (\overline{u^1}, \overline{u^2}, ..., \overline{u^n})$ when $\overline{u^1} := u^1$, and $\overline{u^i}(a) := \begin{cases} 1 \text{ if } a = (a_j^1, a^{-1}) \\ 0 \text{ otherwise} \end{cases}$. The only Nash equilibrium of $\overline{U}$ is (a_j^1, a^{-1}). By uncoupledness we get $\Lambda^1 = \varphi^1(u^1) = \varphi^1(\overline{u^1}) = \overline{\Lambda^1}$. If (a_j^1, a^{-1}) has been played, the next period player 1 will not play a_k^1 with probability 1 (otherwise the set B_j^1 could not be empty), and the equilibrium in the game $\overline{U}$ will never be reached with probability 1 (in contradiction to the assumption). ■

Theorem 5 *Let $k^1 = m^1$ and $k^2 = m^2$. Then there are no uncoupled strategy mappings to k^i-automata with stochastic transitions and mixed actions, that guarantee almost sure convergence of play to a pure Nash equilibrium of the stage game in every game where such an equilibrium exists.*

Proof. (By contradiction). Let U be a full game. Consider the sets $B_1^i, ..., B_{m^i}^i$ in Λ^i (see Lemma 3). By assumption $|\Lambda^i| \leq m^i$. On the one hand, $\left|B_j^i\right| \geq 1$, and, on the other hand,

$\sum_j |B_j^i| \leq |\Lambda^i| = m^i$, and so $|B_j^i| = 1$ and $\underset{j}{\cup} B_j^i = \Psi^i$. In other words, every B_j^i includes exactly one state in which player i plays a_j^i and stay there if a_j^i is a best reply, and there are no other states. Therefore, the strategy of player i is such that if his action is a best reply to the action of the other player, then in the next step he plays the same action. In [2], Proof of Theorem 1, Hart and Mas-Colell show that such a strategy cannot always lead to a pure Nash equilibrium, contradicting our assumption. ■

3.2 Stochastic transitions and pure actions

We will show the result of Theorem 1 continues to hold when the automata are restricted to be pure-actions automata. As was shown in Theorem 5, however, there is no PN-mapping where the ranges of the players are smaller.

Theorem 6 *Let $k^1 \geq m^1$ and $k^2 \geq m^2 + 1$. Then, for each player $i = 1, 2$, there exists an uncoupled strategy mapping to k^i-automata with stochastic transitions and pure actions, that guarantees almost sure convergence of play to a pure Nash equilibrium of the stage game in every game where such an equilibrium exists.*

The proof is relegated to Appendix A.

3.3 Deterministic transitions and mixed actions

We will show that there exists an uncoupled PN-mapping where the range for player i is $(m^i + 2)$-automata. Clearly, every deterministic-transition automaton is a particular case of stochastic-transition automaton, and so Theorem 5 holds here as well.

Theorem 7 *Let $k^i \geq m^i + 2$. Then for each player i there exists an uncoupled strategy mapping to k^i-automata with deterministic transitions and mixed actions, that guarantee almost sure convergence of play to a pure Nash equilibrium of the stage game in every game where such an equilibrium exists.*

The proof is relegated to Appendix B.

3.4 Deterministic transitions and pure actions

In Theorem 13, we will show for general n-player games that there exists an uncoupled PN-mapping where the range for player i is $(O(m^i + n \log n))$-automata, where m^i is the number of actions of player i and n is the number of players. In the case of 2 players, the construction of the automata in the proof of Theorem 13 proves the existence of uncoupled PN-mapping where the range for player i is $(4m^i + O(1))$-automata.

Other uncoupled PN-mapping, specifically for 2 players, has a range of $(5m^1 + m^2 - 5)$-automata for player 1, and $(5m^2 + 2m^1 - 9)$-automata for player 2. We will not show this construction here but the idea is to go through all the possible actions (a_i^1, a_j^2) in some "economical" way.

4 n-Player Games ($n \geq 3$)

4.1 Stochastic transitions, and pure or mixed actions

We will show that there exists an uncoupled PN-mapping where the range for player i is $2m^i$-automata. On the other hand, we will show that there is no uncoupled PN-mapping whose range is smaller for all players.

Theorem 8 *Let $k^i \geq 2m^i$. Then for each player i there exists an uncoupled strategy mapping to k^i-automata with stochastic transitions and pure actions that guarantee almost sure convergence of play to a pure Nash equilibrium of the stage game in every game where such an equilibrium exists.*

Proof. Let us introduce the mappings $\varphi^i(u^i) = \Lambda^i$ given a payoff function $U = (u^1, ..., u^n)$.

Denote the states of Λ^i by $\Psi^i = \{s^i_{1,0}, s^i_{1,1}, s^i_{2,0}, s^i_{2,1} ..., s^i_{m^i,0}, s^i_{m^i,1}\}$. The states $s^i_{j,0}$ will be called **0-states**; the states $s^i_{j,1}$, **1-states**.

Definition 9 *Given a state $s = (s^1, ..., s^n) \in \Psi^1 \times ... \times \Psi^n$ we will say that player i is* ***fit*** *at s if*

- player i is at a 0-state and player $i+1 (\operatorname{mod} n)$ is at a state $s^i_{j,k}$ for $k \in \{0,1\}$ and $j \neq 1$,

or

- player i is at a 1-state and player $i+1(\mathrm{mod}\, n)$ is at a state $s^i_{1,k}$ for $k \in \{0,1\}$.

In every state $s^i_{j,l}$ player i plays action a^i_j. If a^i_j is a best reply to what the other players played, and player i fits player $i+1(\mathrm{mod}\, n)$, player i stays in $s^i_{j,l}$. Otherwise he moves to any one of the $2m^i$ states with equal probability $\frac{1}{2m^i}$.

Let the starting states be $\mathfrak{s}^i_0 := s^i_{1,0}$.

To prove that these automata reach a pure Nash equilibrium we partition the space $\Psi^1 \times ... \times \Psi^n$ of the automata states into $n+2$ regions:

$P_1 := \{(s^1_{k_1,l_1}, ..., s^n_{k_n,l_n}), 1 \leq k_i \leq m^i, l_i = 0, 1 : (a^1_{k_1}, ..., a^n_{k_n})$ is a pure Nash equilibrium and all the players are fit$\}$.

Note that for every pure Nash equilibrium $(a^1_{k_1}, ..., a^n_{k_n})$ there is a state $s \in P_1$ where the players play $(a^1_{k_1}, ..., a^n_{k_n})$. Take $s^i_{k_i,l_i}$ with $l_i = 1$ when $k_{i+1} = 1$ and $l_i = 0$ otherwise.

For $0 \leq r \leq n-1$ $\;P_{2,r} := \{(s^1_{k_1,l_1}, ..., s^n_{k_n,l_n})$: there exist exactly r players that are fit$\}$.

$P_3 := \{((s^1_{k_1,l_1}, ..., s^n_{k_n,l_n})$: all the players are fit, but $(a^1_{k_1}, ..., a^n_{k_n})$ is not a pure Nash equilibrium$\}$.

Clearly each state in P_1 is absorbing. Next we claim that a state in P_1 is reached with positive probability, in finitely many periods, from any other state $s \in \Psi^1 \times ... \times \Psi^n$.

$s \in P_{2,0}$: all the players are not fit, and so all the players move randomly over all their states, and there is a positive probability of reaching P_1.

For $1 \leq r \leq n-1$: $s = (s^1_{k_1,l_1}, ..., s^n_{k_n,l_n}) \in P_{2,r}$. Assume player i is fit, but player[5] $i+1$ is not. Such i exist, because we have a circle of players of which some are fit, and some are not. There is a positive probability that all the players except $i+1$ will stay at their states, and player $i+1$ (who moves randomly because he is not fit) will move in the following way: if $k_{i+1} = 1$ then he moves to $s^{i+1}_{2,l_{i+1}}$, and if $k_{i+1} \geq 2$, then he moves to $s^{i+1}_{1,l_{i+1}}$. Now all the players except i and $i+1$ remain fit/not fit, as they were before, because neither they nor the next player change their state. Player $i+1$ does not change his l_{i+1}, player $i+2$ does not change his state, and so player $i+1$ stays not fit, as he was before. Player i was fit but after the move of player $i+1$, he is not fit. The only player that changes his fitness

[5] From here till the end of the proof, we will write $i+1$ instead of $i+1(\mathrm{mod}\, n)$.

is player i, and they get to $P_{2,r-1}$. By induction, with positive probability they get to $P_{2,0}$ in r steps.

$s = (s^1_{k_1,l_1}, ..., s^n_{k_n,l_n}) \in P_3$: The action is $a = (a^1_{k_1}, ..., a^n_{k_n})$ and it is not a pure Nash equilibrium; therefore, there exists player i s.t. $a^i_{k_i}$ is not a best reply to a^{-i}, and player i moves randomly over all the states. Hence, there is a positive probability that all the players except i will stay at their states and player i will move to $s^i_{k_i,1-l_i}$. Now all the players except player i stay fitted, but player i do not fit. And so $(s^1_{k_1,l_1}, ...s^i_{k_i,1-l_i}, ..., s^n_{k_n,l_n}) \in P_{2,n-1}$.

From any state s there is a positive probability of reaching an absorbing state in P_1 in at most $n + 2$ steps. ■

Theorem 10 *Let $n \geq 4$, and let $k^1, ..., k^n$ satisfy $k^i < 2m^i$ for all $i = 1, ..., n$, (except, perhaps, for one of them). Then there is no uncoupled strategy mapping to k^i-automata with stochastic transitions and mixed actions, that guarantees almost sure convergence of play to a pure Nash equilibrium of the stage game in every game where such an equilibrium exists.*

Proof. Assume on the contrary that such a strategy mapping exists, and that $k^i < 2m^i$ for all $i = 2, ..., n$. By Lemma (2) in Λ^i there exist m^i nonempty sets of states $B^i_1, ..., B^i_{m^1}$ s.t. in every state $s^i_j \in B^i_k$ player i plays a^i_k. $|\Lambda^i| < 2m^i$, and by the pigeon hole principle there exists $k(i)$ s.t. $\left|B^i_{k(i)}\right| = 1$. Therefore, every player i has a state $s^i_{k(i)}$ where he plays $a^i_{k(i)}$, and he stays there if it is a best reply. Consider a four-players game where every player has 2 actions.

Consider the following utility function of players 2, 3, 4:

u^2 :

		a_1^4			a_2^4	
		a_1^2	a_2^2		a_1^2	a_2^2
a_1^3	a_1^1	1	0	a_1^1	1	1
	a_2^1	1	0	a_2^1	1	1
		a_1^2	a_2^2		a_1^2	a_2^2
a_2^3	a_1^1	1	1	a_1^2	0	1
	a_2^1	1	1	a_2^2	0	1

u^3 :

		a_1^4			a_2^4	
		a_1^2	a_2^2		a_1^2	a_2^2
a_1^3	a_1^1	1	1	a_1^1	1	0
	a_2^1	1	1	a_2^1	1	0
		a_1^2	a_2^2		a_1^2	a_2^2
a_2^3	a_1^1	0	1	a_1^2	1	1
	a_2^1	0	1	a_2^2	1	1

u^4 :

		a_1^4			a_2^4	
		a_1^2	a_2^2		a_1^2	a_2^2
a_1^3	a_1^1	1	1	a_1^1	0	1
	a_2^1	1	1	a_2^1	0	1
		a_1^2	a_2^2		a_1^2	a_2^2
a_2^3	a_1^1	1	0	a_1^2	1	1
	a_2^1	1	0	a_2^2	1	1

Player $i = 2, 3, 4$ gets 1 if he plays the same action as one of the players 2, 3, 4 (except himself). Otherwise he gets 0.

The strategy mapping $\varphi^i : u^i \to \mathcal{A}^i$ constructs an automaton. As mentioned, there exists a state $s^i_{k(i)}$ where player i plays $a^i_{k(i)}$, and he stays there if it is a best reply. There are 2 actions for every player, and 3 actions $a^i_{k(i)}$ $i = 2, 3, 4$. So there exist 2 players i, j who have the same action $a^i_{k(i)}, a^j_{k(j)}$, where $k(i) = k(j)$. Because of the symmetry of the functions u^2, u^3, u^4, assume $k(i) = k(j) = 1$, and assume that the two players are players 3 and 4.

Let us now consider the following game:

Γ_1 :

		a_1^4			a_2^4	
		a_1^2	a_2^2		a_1^2	a_2^2
a_1^3	a_1^1	$1, 0, 1, 1$	$0, 1, 1, 1$	a_1^1	$0, 0, 1, 0$	$0, 0, 0, 1$
	a_2^1	$0, 1, 1, 1$	$1, 0, 1, 1$	a_2^1	$0, 0, 1, 0$	$0, 0, 0, 1$
		a_1^2	a_2^2		a_1^2	a_2^2
a_2^3	a_1^1	$0, 0, 0, 1$	$0, 0, 1, 0$	a_1^2	$1, 1, 1, 1$	$1, 1, 1, 1$
	a_2^1	$0, 0, 0, 1$	$0, 0, 1, 0$	a_2^2	$1, 1, 1, 1$	$1, 1, 1, 1$

Players 3 and player 4 have the utility functions u^3 and u^4 respectively. Therefore, the automaton that their strategy mapping constructs include states s_1^3, s_1^4, where they play action a_1^3, a_1^4 (respectively) and stay there if it is a best reply. If players 3 and 4 get to the states $s^3_{k(3)}$ and $s^4_{k(4)}$, then the pure Nash equilibrium will never be reached.

For a larger number of actions the same proof works, if we take all the actions $a_2^i, ..., a^i_{m^i}$ to be identical to the action a_2^i in this proof.

For a larger number of players we take the utility functions of players 2, 3, and 4 to be the same as in the case of 4 players and independent of the actions of the other player $(1, 5, ..., n)$. And in the game Γ_1 the utility functions of players $5, ..., n$ will be 1 if players 3 and 4 played the same action, and 0 otherwise. ■

4.2 Deterministic transitions and mixed actions

We will show that there exists an uncoupled PN-mapping such that the range for player i is $(2m^i + 3)$-automata. Clearly, every deterministic-transition automaton is a particular case of stochastic-transition automaton, and so Theorem 10 holds here as well.

Theorem 11 *Let $k^i \geq 2m^i + 3$. Then for each player i there exists an uncoupled strategy mapping to k^i-automata with deterministic transitions and mixed actions, that guarantees almost sure convergence of play to a pure Nash equilibrium of the stage game in every game where such an equilibrium exists.*

Proof. We introduce the mappings $\varphi^i(u^i) = \Lambda^i$ given a payoff function $U = (u^1, ..., u^n)$:

As in Theorem 8, we will use the states $\{s^i_{1,0}, s^i_{1,1}, s^i_{2.0}, s^i_{2,1}..., s^i_{m^i,0}, s^i_{m^i,1}\}$. These are the same states, exapt that their transitions are deterministic. Denote the states of Λ^i by $\Psi^i = \{s^i_1, s^i_2, s^i_3\} \cup \{s^i_{1,0}, s^i_{1,1}, s^i_{2.0}, s^i_{2,1}..., s^i_{m^i,0}, s^i_{m^i,1}\}$.

The states $s^i_{j,l}$ are similar to the states $s^i_{j,l}$ in Theorem 8. In every state $s^i_{j,l}$ player i plays action a^i_j. If a^i_j is a best reply to what the other players played, and player i fits, then player i stays on it (exactly as before). Otherwise he moves to s^i_1.

In the state s^i_1 player i plays $\left(\frac{1}{m^i}, ..., \frac{1}{m^i}\right)$. If he played a^i_1 he stays in s^i_1. If he played a^i_2 he moves to s^i_2. If he played a^i_j $j \neq 1, 2$ he moves to $s^i_{j,0}$.

In the state s^i_2 player i plays $\left(\frac{1}{m^i}, ..., \frac{1}{m^i}\right)$. If he played a^i_3 he stays in s^i_1. If he played a^i_4 he moves to s^i_3. If he played a^i_j $j \neq 3, 4$ he moves to $s^i_{j,1}$.

In the state s^i_3 player i plays $\left(\frac{1}{m^i}, ..., \frac{1}{m^i}\right)$. If he played a^i_1 he moves to $s^i_{1,0}$. If he played a^i_2 he moves to $s^i_{2,0}$. If he played a^i_3 he moves to $s^i_{3,1}$. If he played a^i_4 he moves to $s^i_{4,1}$. If he played a^i_5 he stays in s^i_3. If he played a^i_j $j \geq 6$, he moves to s^i_1.

Let the starting states be $\mathfrak{s}^i_0 := s^i_1$.

The proof of the claim that this mapping is a PN-mapping is proven similarly to Theorem 7 with 2 players. ∎

4.3 Deterministic transitions and pure actions

We will show that there exists an uncoupled PN-mapping such that the range for player i is $O(m^i + n \log n)$-automata. Clearly, every deterministic-transition automaton is a particular case of a stochastic-transition automaton, and so Theorem 10 holds here as well.

Lemma 12 *For every $m, n \in \mathbb{N}$ there exist n different prime numbers $p_1, ..., p_n$, $p_i \geq m$ for every i, such that $p_i = O(m + n \log n)$ for every i.*

Proof. We know that in $\{1,2,...,m\}$ there exist a maximum of $\alpha\frac{m}{\log m}$ prime numbers (α is a constant).

We also know that in $\{1,2,...,\beta(n+\alpha\frac{m}{\log m})\log(n+\alpha\frac{m}{\log m})\}$ there exist a minimum of $n+\alpha\frac{m}{\log m}$ prime numbers (β is a constant).

Therefore, in $\{m+1,m+2,...,\beta(n+\alpha\frac{m}{\log m})\log(n+\alpha\frac{m}{\log m})\}$ there exist a minimum of n prime numbers, and we can take different prime numbers $p_1,...,p_n$ s.t. $\forall i: m<p_i<\beta(n+\alpha\frac{m}{\log m})\log(n+\alpha\frac{m}{\log m})$.

To complete the proof we will show that $\beta(n+\alpha\frac{m}{\log m})\log(n+\alpha\frac{m}{\log m})=O(m+n\log n)$.

–If $m\leq n$ then

$$\beta(n+\alpha\frac{m}{\log m})\log(n+\alpha\frac{m}{\log m})=O((1+\alpha)n\log((1+\alpha)n))=O(n\log n)=O(m+n\log n).$$

–If $n<m\leq n\log n$, then

$$\frac{m}{\log m}\leq 2n\frac{\log n}{2\log m}=2n\frac{\log n}{\log m^2}\leq 2n\Longrightarrow \beta(n+\alpha\frac{m}{\log m})\log(n+\alpha\frac{m}{\log m})=$$
$$=O((1+2\alpha)n\log((1+2\alpha)n))=O(n\log n)=O(m+n\log n).$$

–If $n\log n<m\leq n^2$, then

$$n\leq\frac{m}{\log m}\frac{\log m}{\log n}\leq\frac{m}{\log m}\frac{\log n^2}{\log n}\leq 2\frac{m}{\log m}\Longrightarrow$$
$$\beta(n+\alpha\frac{m}{\log m})\log(n+\alpha\frac{m}{\log m})=O((2+\alpha)\frac{m}{\log m}\log((2+\alpha)\frac{m}{\log m}))=$$
$$=O(\frac{m}{\log m}(\log m-\log\log m)=O(m)=O(m+n\log n).$$

–If $n^2<m$, then

$$\beta(n+\alpha\frac{m}{\log m})\log(n+\alpha\frac{m}{\log m})=O((1+\alpha)\frac{m}{\log m}\log((1+\alpha)\frac{m}{\log m}))=$$
$$=O(\frac{m}{\log m}(\log m-\log\log m)=O(m)=O(m+n\log n).$$

In any case $\beta(n+\alpha\frac{m}{\log m})\log(n+\alpha\frac{m}{\log m})=O(m+n\log n)$. ■

Theorem 13 *Let $k^i\geq O(m+n\log n)$, where $m=\max\{m^i\}$. Then for each player i there exists an uncoupled strategy mapping to k^i-automata with deterministic transitions and pure actions, that guarantees almost sure convergence of play to a pure Nash equilibrium of the stage game in every game where such an equilibrium exists.*

Proof. Let $p_1, ..., p_n$ be different prime numbers s.t. $\forall i \ p_i > m^i$.[6] By Lemma 12 we can take $p_1, ..., p_n$ s.t. $p_i = O(m + n \log n)$.

We will show that there exists a PN-mapping $(\varphi^1, ..., \varphi^n)$ such that the range of φ^i is $(4p_i + 3)$-automata, and in doing so we will have concluded our proof.

We introduce the mappings $\varphi^i(u^i) = \Lambda^i$, given a payoff function $U = (u^1, ..., u^n)$:

The automaton Λ^i consists of, starting state, a state after that and p_i regions. The first region has five states. The other regions have the same structure of four states. Denote the p_i regions of Λ^i by $Q_1^i = \{s_{j,1}^i, s_{j,2}^i, ..., s_{j,5}^i\}$, and $Q_j^i = \{s_{j,1}^i, ..., s_{j,4}^i\}$ for $j = 2, ..., p_i$. Denote the starting state by $\mathfrak{s}_0^i := s_1^i$ and the state after it by s_2^i.

s_1^i : player i plays a_1^i. If $(a_1^1, ..., a_1^n)$ was played he moves to $s_{1,1}^i$ (the starting state of Q_1^i). Otherwise, if a_1^i is a best reply, then he stays at s_1^i; if a_1^i is not a best reply, then he moves to s_2^i.

s_2^i : player i plays a_2^i and in any case moves to $s_{1,1}^i$ (the starting state of Q_1^i).

These two starting states guarantee us two things: First, if the equilibrium is $(a_1^1, ..., a_1^n)$ then the players will stay there. Second: if it is not, then all the players get to $s_{1,1}^i$ simultaneously. We check the actions $(a_1^1, ..., a_1^n)$ separately, because in the continuation of the construction of the automata we will use the fact that $(a_1^1, ..., a_1^n)$ is not an equilibrium.

The regions $Q_1^i, Q_2^i, ..., Q_{p_i}^i$ are arranged in a circle when the players move to the region $Q_{j+1(\mathrm{mod}\, p_i)}^i$ from the previous region Q_j^i.

For $j \neq 1$ the construction of Q_j^i is the following:

---

[6] Every player i has to choose his number p_i to construct his automaton. But he doesn't know the numbers of the other players: $p_1, ..., p_{i-1}, p_{i+1}, ..., p_n$. Yet we have to ensure that every player will choose a different number. Therefore, we have to define a choice function: $\chi : \mathbb{N}^n \to (PRIME)^n$, known to the players. The choice function chooses for every $(m^1, ..., m^n)$ n prime numbers: $\chi(m^1, ..., m^n)$, and then every player i will choose the number $p_i := (\chi(m^1, ..., m^n))_i$.

As an example of choice function, let $\{w_k\}_{k=1}^{\infty}$ be the sequence of all the prime numbers in increasing order. For every i let $k(i)$ be the minimal number such that $\begin{cases} w_{k(i)} \geq m^i \\ k(i) = i(\mathrm{mod}\, n) \end{cases}$, and define $p_i := w_{k(i)}$.

The starting state $s^i_{j,1}$: player i plays a^i_j. If player $i+1(\operatorname{mod} n)$ played[7] a^{i+1}_1, player i moves to $s^i_{j,2}$; otherwise he moves to $s^i_{j,3}$. If $(a^1_1, ..., a^n_1)$ was played, he moves to[8] $s^i_{j+1,1}$.

$s^i_{j,2}$: player i plays a^i_j. If it is a best reply, and also player $i+1$ played a^{i+1}_1 (the same action as before) then player i stays there. Otherwise he moves to $s^i_{j,4}$. If $(a^1_1, ..., a^n_1)$ was played, he moves to $s^i_{j+1,1}$.

$s^i_{j,3}$: player i plays a^i_j. If it is a best reply, and also player $i+1$ played a^{i+1}_k $k \geq 2$ (the same action as before), player i stays there. Otherwise he moves to $s^i_{j,4}$. If $(a^1_1, ..., a^n_1)$ was played, he moves to $s^i_{j+1,1}$.

$s^i_{j,4}$: player i plays a^i_1. If $(a^1_1, ..., a^n_1)$ was played, he moves to $s^i_{j+1,1}$; otherwise he stays at $s^i_{j,4}$.

For $j = 1$ the construction of Q^i_1 is quite similar. The only difference is that the state $s^i_{j,4}$, changed by two $s^i_{1,4}, s^i_{1,5}$:

$s^i_{1,4}$: player i plays a^i_2. In any case, he moves to $s^i_{1,5}$.

$s^i_{1,5}$: player i plays a^i_1. If $(a^1_1, ..., a^n_1)$ was played, he moves to $s^i_{j+1,1}$. Otherwise he stays in $s^i_{1,5}$.

For $m^i < j \leq p_i$ let $Q^i_j := Q^i_{m^i}$ denoted that the act in every $s^i_{j,k}$ $m^i < j \leq p_i, k = 1, .., 4$, is identical to the act in $s^i_{m^i,k}$.

In the state $s^i_{j,1}$ player i informs player $i-1$ what he played.

If the players located at $Q^1_{k_1}, ..., Q^n_{k_n}$, and $(a^1_{k_1}, ..., a^n_{k_n})$ is an equilibrium, then all the players will stay at $s^i_{j,2}, s^i_{j,3}$.

In the state $s^i_{j,4}$ player i plays the "opposite" action than he played before and informs player $i-1$ that it is not an equilibrium.

In the state $s^i_{j,4}$ for $j \neq 1$, and $s^i_{1,5}$ for $j = 1$ player i waits until all the are informed that it is not an equilibrium. Then all the players can move to their the next region simultaneously.

To summarize, if the players are located at $Q^1_{k_1}, ..., Q^n_{k_n}$, and $(a^1_{k_1}, ..., a^n_{k_n})$ is an equilibrium, then all the players stay at the equilibrium all the time. Otherwise they all will move to $Q^1_{k_1+1}, ..., Q^n_{k_n+1}$ simultaneously.

[7] From here till the end of the proof, we will write $i \pm 1$ instead of $i \pm 1(\operatorname{mod} n)$.

[8] From here till the end of the proof, we will write $j+1$ instead of $j+1(\operatorname{mod} p_i)$.

Since the players move from Q^i_j to Q^i_{j+1} simultaneously, by the Chinese Remainder theorem it follows that they will visit every $(Q^1_{k_1},...,Q^n_{k_n}) : 1 \leq k_i \leq m^i$, until they become stuck in some $(Q^1_{k_1},...,Q^n_{k_n})$ for which $(a^1_{k_1},...,a^n_{k_n})$ is an equilibrium. Therefore, if a pure equilibrium exists, the automata will eventually reach it. ■

5 Appendix A. Proof of Theorem 6.

Proof. We define the mapping φ as follows:

Given a game $U = (u^1, u^2)$, the automaton $\varphi^1(u^1) = \Lambda^1 \in \mathcal{A}^1$ is constructed as follows:

Denote $\varphi^1(u^1) = \Lambda^1 =< \Psi^1, \mathfrak{s}^1_0, f^1, g^1 >$ when Λ^1 is a m^1-automaton. We denote the states of Λ^1 by $\Psi^1 = \{s^1_1, ..., s^1_{m^1}\}$.

$\mathfrak{s}^1_0 := s^1_1$.

In state s^1_i player 1 plays action a^1_i. He stays in this state if a^1_i is a best reply to the action of player 2; otherwise he moves randomly to any one of the m^1 states with equal probability $\frac{1}{m^1}$.

In order to define the mapping φ^2, we start by considering the following action b^2 of player 2: for every action a^2_j of player 2, let $\#BR(a^2_j)$ be the number of 1s in the column a^2_j in the table $\widetilde{u}^2$; i.e., $\#BR(a^2_j) := |\{a^1_i|\widetilde{u}^2(a^1_i, a^2_j) = 1\}| = |\{a^1_i|a^2_j$ is a best reply to $a^1_i\}|$. Consider an action a^2_k with a maximal number of 1s in its column: $\#BR(a^2_k) = \max\{\#BR(a^2_j)|a^2_j \in A^2\}$. Without loss of generality assume $k = 1$; i.e., the first column of $\widetilde{u}^2$ has no fewer 1s than any other column. Denote this action by $b^2 := a^2_1$.

Now we construct the automaton $\varphi^2(u^2) = \Lambda^2 \in \mathcal{A}^2$ as follows:

Denote $\Lambda^2 =< \Psi^2, \mathfrak{s}^2_0, f^2, g^2 >$ when Λ^2 is a $(m^2 + 1)$-automaton. We denote the states of Λ^2 by $\Psi^2 = \{s^2_0, s^2_1, ..., s^2_{m^1}\}$.

$\mathfrak{s}^2_0 := s^2_0$.

In the state s^2_0, player 2 plays the action b^2, and moves to any of the $m^2 + 1$ states with probability $\frac{1}{m^2+1}$.

In the states s^2_i, $i \geq 1$, player 2 plays action a^2_i. He stays in this state if a^2_i is a best reply to the action of player 1; otherwise he moves randomly to any one of the $m^2 + 1$ states with equal probability $\frac{1}{m^2+1}$.

Now we will prove that (φ^1, φ^2) is PN-mapping. Let us consider two cases:

Case 1: For every $i = 1, ..., m^1 : \widetilde{u}^2(a_i^1, b^2) = 1$; i.e., b^2 is a dominant action.

We partition the space $\Psi^1 \times \Psi^2$ of the automata states into four regions:

$P_1 := \{(s_i^1, s_j^2), 1 \leq i \leq m^1, 1 \leq j \leq m^2 : \widetilde{u}^1(a_i^1, a_j^2) = 1, \widetilde{u}^2(a_i^1, a_j^2) = 1\}$; i.e., in this case (a_i^1, a_j^2) is a pure Nash equilibrium.

$P_2 := \{(s_i^1, s_0^2), 1 \leq i \leq m^1\}$

$P_3 := \{(s_i^1, s_j^2), 1 \leq i \leq m^1, 1 \leq j \leq m^2 : \widetilde{u}^2(a_i^1, a_j^2) = 0\}$

$P_4 := \{(s_i^1, s_j^2), 1 \leq i \leq m^1, 1 \leq j \leq m^2 : \widetilde{u}^1(a_i^1, a_j^2) = 0, \widetilde{u}^2(a_i^1, a_j^2) = 1\}$

These four regions clearly cover the space $\Psi^1 \times \Psi^2$ (because player 2 can be in the state s_0^2 (P_2) or in any other state ($P_1 \cup P_3 \cup P_4$). The action of player 2 can be a best reply ($P_1 \cup P_4$) or not (P_3). If it is a best reply, then the action of player 1 can be a best reply (P_1) or not (P_4).)

Players 1 and 2 stay at the same state if their action is a best reply; i.e., each state in P_1 is absorbing. We will prove that there is a positive probability of reaching a state from P_1, in finitely many periods, from any other state $s \in \Psi^1 \times \Psi^2$.

$s = (s_i^1, s_0^2) \in P_2$: The actions are $(a_i^1, b^2) = (f^1(s_i^1), f^2(s_0^2))$. Wether a_i^1 is a best reply or not, player 1 has a positive probability (1 or $\frac{1}{m^1}$ correspondingly) to move to s_k^1, where a_k^1 is a best reply of player 1 to b^2. Player 2 will move to the state s_1^2 with probability $\frac{1}{m^2+1}$, where b^2 is a best reply to every action of player 1, in particular to a_k^1. Now $(s_k^1, s_1^2) \in P_1$.

$s = (s_i^1, s_j^2) \in P_3$: The actions are (a_i^1, a_j^2); a_j^2 is not a best reply. Therefore, player 2 moves to s_0^2 with probability $\frac{1}{m^2+1}$. Denote the state to which player 1 moves by s_k^1. Then $(s_k^1, s_0^2) \in P_2$.

$s = (s_i^1, s_j^2) \in P_4$: The actions are (a_i^1, a_j^2); a_j^2 is a best reply, a_i^1 is not. Therefore, player 2 stays in s_j^2, and player 1 move to s_k^1 with probability $\frac{1}{m^1}$, where a_k^1 is a best reply of player 1 to a_j^2. Now either $(s_k^1, s_j^2) \in P_1$ or $(s_k^1, s_j^2) \in P_3$, depending on whether a_j^2 is a best reply to a_k^1.

In all the cases there is a positive probability of reaching an absorbing state in P_1 in at most 3 steps.

Case 2: There exists an action of player 1, say a_l^1, such that b^2 is not a best reply to it; i.e., there exists $l = 1, ..., m^1$ such that $\widetilde{u}^2(a_l^1, b^2) = 0$.

Before proving the Theorem, we will prove the following simple claim about the configuration of 0s and 1s in $\widetilde{u}^2$. This claim will be useful later.

Claim 1: If there exists $l = 1, ..., m^1$, such that $\widetilde{u}^2(a_l^1, b^2) = 0$, then:

(a) Let $a_j^2 \in A^2$. Then there exists $a_i^1 \in A^1$ such that $\widetilde{u}^2(a_i^1, a_j^2) = 0$.

(b) Let $a_i^1 \in A^1$. Then there exists $a_j^2 \in A^2$ such that $\widetilde{u}^2(a_i^1, a_j^2) = 1$.

(c) Let $a_i^1 \in A^1, a_j^2 \in A^2$, such that $\begin{cases} \widetilde{u}^2(a_i^1, a_j^2) = 1 \\ \widetilde{u}^2(a_i^1, b^2) = 0 \end{cases}$. Then there exists $a_k^1 \in A^1$ such that $\begin{cases} \widetilde{u}^2(a_k^1, a_j^2) = 0 \\ \widetilde{u}^2(a_k^1, b^2) = 1 \end{cases}$.

Proof: (a) Otherwise there would be a column in $\widetilde{u}^2$ that includes 1s only. By the assumption in this case there exist $i = 1, ..., m^1$, such that $\widetilde{u}^2(a_i^1, b^2) = 0$, which contradicts the fact that b^2 is the column with the maximal number of 1s.

(b) There is some action $a_j^2 \in A^2$ that is a best reply to the action $a_i^1 \in A^1$. This action satisfies $\widetilde{u}^2(a_i^1, a_j^2) = 1$.

(c) Otherwise the number of 1s in the j-th column would be bigger than in the first column.$\square$

We partition the space $\Psi^1 \times \Psi^2$ of the automata states into regions:

For every $x, y, z \in \{0, 1\}$ put:

$P_{xyz} := \{(s_i^1, s_j^2), 1 \leq i \leq m^1, 1 \leq j \leq m^2 : \widetilde{u}^1(a_i^1, a_j^2) = x, \widetilde{u}^2(a_i^1, a_j^2) = y, \widetilde{u}^2(a_i^1, b^2) = z\}$

and

$Q_{xy} := \{(s_i^1, s_0^2), 1 \leq i \leq m^1 : \widetilde{u}^1(a_i^1, b^2) = x, \widetilde{u}^2(a_i^1, b^2) = y\}$.

In words Q for a region we mean that player 2 is in the state s_0^2, and by P we mean that he is not. There are three indices for region P, and two Boolean indices for region Q. The first index in P, Q corresponds to player 1 and indicates whether his action is a best reply (1) or not (0). The second index in P, Q corresponds similarly to player 2. The third index in P indicates whether action b^2 is a best reply to the action that player 1 played (1) or not (0).

Clearly $(\bigcup_{x,y,z\in\{0,1\}} P_{xyz}) = \Psi^1 \times (\Psi^2 \backslash \{s_0^2\})$ and $(\bigcup_{x,y\in\{0,1\}} Q_{xy}) = \Psi^1 \times \{s_0^2\}$, since we are considering all the possibilities. Therefore, $(\bigcup_{x,y,z\in\{0,1\}} P_{xyz}) \cup (\bigcup_{x,y\in\{0,1\}} Q_{xy}) = \Psi^1 \times \Psi^2$.

Consider the region $P_{11\bullet}$. By the definition of P_{xyz} we can see that $P_{11\bullet} = \{(s_i^1, s_j^2), 1 \leq i \leq m^1, 1 \leq j \leq m^2 : \widetilde{u}^1(a_i^1, a_j^2) = \widetilde{u}^2(a_i^1, a_j^2) = 1\} = \{(s_i^1, s_j^2), 1 \leq i \leq m^1, 1 \leq j \leq m^2 : (a_i^1, a_j^2) \text{ is a pure Nash equilibrium}\}$.

Players 1 and 2 stay at the same state if their action is a best reply; i.e., each state in $P_{11\bullet}$ is absorbing. If we can show that a state from $P_{11\bullet}$ is reached with positive probability, in finitely many periods, from all the other regions that we have already defined, then we will have concluded our proof.

$P_{00\bullet}$: This is the region of all the states where the actions of both players are not a best reply, and so both players move randomly over all their states, therefore, they reach a pure Nash equilibrium with probability $\frac{1}{m^1} \cdot \frac{1}{m^2+1}$; i.e., they reach $P_{11\bullet}$.

$Q_{0\bullet}$: Player 1's action is not a best reply, and player 2 is in state s_0^2. Therefore, as before both players move randomly over all their states, and they reach $P_{11\bullet}$ with probability $\frac{1}{m^1} \cdot \frac{1}{m^2+1}$, as before.

Q_{11} : Player 1's action is a best reply and so he stays in the same state. Player 2 is in state s_0^2 and so he moves randomly over all his states. Player 2 will move to s_1^2 with probability $\frac{1}{m^2+1}$, and they get to $P_{11\bullet}$.

P_{101} : Player 1 stays, and player 2 randomizes. Since $\widetilde{u}^2(a_i^1, b^2) = 1$ player 2 will move to s_0^2 with probability $\frac{1}{m^2+1}$, and they get to Q_{01} or Q_{11}.

P_{010} : Player 1 randomizes and player 2 stays. By claim 1(c) there exists $a_k^1 \in A^1$ such that $\begin{cases} \widetilde{u}^2(a_k^1, a_j^2) = 0 \\ \widetilde{u}^2(a_k^1, b^2) = 1 \end{cases}$. Player 1 will move to s_k^1 with probability $\frac{1}{m^1}$, and they get to P_{001} or P_{101}.

P_{100} : Player 1 stays and player 2 randomizes. By claim 1(b) there exists $a_j^2 \in A^2$ such that $\widetilde{u}^2(a_i^1, a_j^2) = 1$. Player 2 will move to s_j^2 with probability $\frac{1}{m^2+1}$. Note that a_i^1 does not change and so still $\widetilde{u}^2(a_i^1, b^2) = 0$ and they get to $P_{0,1,0}$ or P_{110} (i.e., $P_{\bullet 10}$).

P_{011} : Player 1 randomizes and player 2 stays. By claim 1(a) there exists $a_k^1 \in A^1$ such that $\widetilde{u}^2(a_k^1, a_j^2) = 0$. Player 1 will move to s_k^1 with probability $\frac{1}{m^1}$, and they get to $P_{00\bullet} \cup P_{100} \cup P_{101} = P_{\bullet 0 \bullet}$.

Q_{10} : Player 1 stays and player 2 randomizes. Player 2 moves to s_1^2 with probability

$\frac{1}{m^2+1}$, and they get to P_{100}.

Thus we have covered all the regions and shown that in at most 5 periods there is a positive probability of reaching the absorbing state $P_{11\bullet}$. The regions cover all the space $\Psi^1 \times \Psi^2$, and so the automata will reach a pure Nash equilibrium when such an equilibrium exists with probability 1. ■

6 Appendix B. Proof of Theorem 7.

Proof. We define the mapping φ as follows:

Given a game $U = (u^1, u^2)$, the automaton $\varphi^1(u^1) = \Lambda^1 \in \mathcal{A}^1$ is constructed as follows:

Denote the states of Λ^i by $\Psi^i = \{s^i_{00}, s^i_{01}, s^i_1, ..., s^i_{m^1}\}$.

In the state s^i_j player i plays action a^i_j. He stays in this state if a^i_j is the best reply to the action of player 2; otherwise he moves to s^i_{00}.

In the state s^i_{00} player i plays $\left(\frac{1}{m^i}, ..., \frac{1}{m^i}\right)$. If he played a^i_1, he stays in s^i_{00}. If he played a^i_2, he moves to s^i_{01}. If he played a^i_j, $j \geq 3$, he moves to s^i_j.

In the state s^i_{01} player i plays $\left(\frac{1}{m^i}, ..., \frac{1}{m^i}\right)$. If he played a^i_1, he moves to s^i_1. If he played a^i_2, he moves to s^i_2. If he played a^i_3, he stays in s^i_{01}. If he played a^i_j $j \geq 4$, he moves to s^i_{00}.

Let the starting state be $\mathfrak{s}^i_0 := s^i_{00}$.

The proof that these mapping is PN-mapping requires consideration of four cases separate: if a pure Nash equilibrium is (a^1_k, a^2_l) then we will consider the cases $\{k \leq 2, l \leq 2\}$, $\{k \leq 2, l > 2\}$, $\{k > 2, l \leq 2\}$, and $\{k > 2, l > 2\}$. We will prove only ona case, says $\{k \leq 2, l > 2\}$, since the proofs for the other cases is similar.

We partition the space $\Psi^1 \times \Psi^2$ of the automata states into regions:

$P_1 := \{(s^1_i, s^2_j), 1 \leq i \leq m^1, 1 \leq j \leq m^2 : (a^1_i, a^2_j) \text{ is a pure Nash equilibrium}\}$

$P_2 := \{(s^1_{01}, s^2_{00})\}$

$P_3 := \{(s^1_{00}, s^2_{00}) \cup (s^1_{00}, s^2_{01}) \cup (s^1_{01}, s^2_{01})\}$

For every $x \in \{00, 01\}, y \in \{\leq, >\}$ put:

$Q_{x,y} := \{(s^1_x, s^2_j), 1 \leq j \leq m^2 : \text{there exist} \begin{cases} i \leq 2 \text{ if } y \text{ is } \leq \\ i > 2 \text{ if } y \text{ is } > \end{cases} \text{such that } a^2_j \text{ is not a best reply to } a^1_i\}$

For every $x \in \{\leq, >\}, y \in \{00, 01\}$ put:

$Q_{x,y} := \{(s_i^1, s_y^2), 1 \leq i \leq m^1 : \text{there exist} \begin{cases} j \leq 2 \text{ if } x \text{ is } \leq \\ j > 2 \text{ if } x \text{ is } > \end{cases} \text{such that } a_i^1 \text{ is not a best}$ reply to $a_j^2\}$

$P_4 := \{(s_i^1, s_j^2), 1 \leq i \leq m^1, 1 \leq j \leq m^2 : (a_i^1, a_j^2) \text{ is not a pure Nash equilibrium}\}$

Clearly, $P_2 \cup P_3 = \{s_{00}^1, s_{01}^1\} \times \{s_{00}^2, s_{01}^2\}$, $P_1 \cup P_4 = \{s_1^1, ..., s_{m^1}^1\} \times \{s_1^2, ..., s_{m^2}^2\}$, $\underset{x\in\{00,01\},y\in\{\leq,>\}}{\cup} Q_{x,y} = \{s_{00}^1, s_{01}^1\} \times \{s_1^2, ..., s_{m^2}^2\}$, $\underset{x\in\{\leq,>\},y\in\{00,01\}}{\cup} Q_{x,y} = \{s_1^1, ..., s_{m^1}^1\} \times \{s_{00}^2, s_{01}^2\}$. Therefore, the union of all the regions is $\Psi^1 \times \Psi^2$.

Players 1 and 2 stay at the same state if their action is a best reply; i.e., each state in P_1 is absorbing. If we can show that a state from P_1 is reached with positive probability, in finitely many periods, from all the other regions that we have already defined, then we will have concluded our proof.

P_2 : Both players randomize. They play (a_k^1, a_l^2) with positive probability, and then they get to P_1.

P_3 : Both players randomize. From every one of the states there is a positive probability to get to P_2. For example, if they are in the states (s_{00}^1, s_{00}^2), then if they play (a_2^1, a_1^2) they move to $(s_{01}^1, s_{00}^2) \in P_2$.

$Q_{00,\leq}$: Player 1 randomizes. With positive probability the actions are (a_i^1, a_j^2) such that a_j^2 is not a best reply to a_i^1, so player 1 stays at $\{s_{00}^1, s_{01}^1\}$ and the action of player 2 is not a best reply. Therefore, player 2 moves to s_{00}^2, and so $(s_{00}^1, s_{00}^2), (s_{01}^1, s_{00}^2) \in P_2 \cup P_3$.

$Q_{01,\leq}$: Player 1 randomizes. With positive probability the actions are (a_4^1, a_j^2). If a_j^2 is a best reply to a_4^1, they move to $(s_{00}^1, s_j^2) \in Q_{00,\leq}$. Otherwise they move to (s_{00}^1, s_{00}^2).

$Q_{01,>}, Q_{00,>}$: Similar to $Q_{00,\leq}, Q_{01,\leq}$

$Q_{\leq,00}, Q_{\leq,01}, Q_{>,01}, Q_{>,00}$: Symmetric to $Q_{00,\leq}, Q_{01,\leq}, Q_{01,>}, Q_{00,>}$

P_4 : The action of one of the players is not a best reply to the action of the other, and so one of them will move to the state s_{00} (i.e., s_{00}^1 or s_{00}^2). Hence, they get to one of the previous regions.

We have covered all the regions and shown that in at most 5 periods there is positive probability of reaching the absorbing state $P_{11\bullet}$. The regions cover all the space $\Psi^1 \times \Psi^2$; therefore, the mapping is PN-mapping.

In the other cases where the pure Nash equilibrium (a^1_k, a^2_l) satisfies $\{k \leq 2, l \leq 2\}$, $\{k > 2, l \leq 2\}$, or $\{k > 2, l > 2\}$ the only difference is in how the regions P_2 and P_3 are defined. For example, for the case $\{k \leq 2, l \leq 2\}$, P_2 and P_3 will be defined by $P_2 := \{(s^1_{01}, s^2_{01})\}$ and $P_3 := \{(s^1_{00}, s^2_{00}) \cup (s^1_{00}, s^2_{01}) \cup (s^1_{01}, s^2_{00})\}$. $\blacksquare$

References

[1] S. Hart, A. Mas-Colell, "Uncoupled Dynamics do not Lead to Nash Equilibrium", *American Economic Review,* **93** (2003), 1830-1836.

[2] S. Hart, A. Mas-Colell, "Stochastic Uncoupled Dynamics and Nash Equilibrium", *Games and Economic Behavior,* **57** (2006) 286.