Stochastic games on a product state space: the periodic case

János Flesch, Gijs Schoenmakers, Koos Vrieze

Stochastic Games on a Product State Space: The Periodic Case

RM/08/016

JEL code: C73

Maastricht research school of **E**conomics
of **TE**chnology and **OR**ganizations

Universiteit Maastricht
Faculty of Economics and Business Administration
P.O. Box 616
NL - 6200 MD Maastricht

phone : ++31 43 388 3830
fax : ++31 43 388 4873

Stochastic Games on a Product State Space: The Periodic Case

János Flesch, Gijs Schoenmakers, Koos Vrieze*

June 10, 2008

Abstract

We examine so-called product-games. These are n-player stochatic games played on a product state space $S^1 \times \cdots \times S^n$, in which player i controls the transitions on S^i. For the general n-player case, we establish the existence of 0-equilibria. In addition, for the case of two-player zero-sum games of this type, we show that both players have stationary 0-optimal strategies.

In the analysis of product-games, interestingly, a central role is played by the periodic features of the transition structure. Flesch et al. [2008] showed the existence of 0-equilibria under the assumption that, for every player i, the transition structure on S^i is aperiodic. In this article, we examine product-games with periodic transition structures. Even though a large part of the approach in Flesch et al. [2008] remains applicable, we encounter a number of tricky problems that we have to address. We provide illustrative examples to clarify the essence of the difference between the aperiodic and periodic cases.

Keywords: Noncooperative Games, Stochastic Games, Periodic Markov Decision Problems, Equilibria.

1 Introduction

Stochastic games and product-games. An n-player *stochastic game* is given by (1) a set of players $N = \{1, \ldots, n\}$, (2) a nonempty and finite set of states S, (3) for each state $s \in S$, a nonempty and finite set of actions A_s^i for each player i, (4) for each

*Addresses: János Flesch: Department of Quantitative Economics/ Gijs Schoenmakers & Koos Vrieze: Department of Mathematics. University of Maastricht, P.O.Box 616, 6200 MD Maastricht, The Netherlands.

state $s \in S$ and each joint action $a_s \in \times_{i \in N} A_s^i$, a payoff $r_s^i(a_s) \in \mathbb{R}$ to each player i, (5) for each state $s \in S$ and each joint action $a_s \in \times_{i \in N} A_s^i$, a transition probability distribution $p_{sa_s} = (p_{sa_s}(t))_{t \in S}$.

The game is to be played at stages in $\mathbb{N}$ in the following way. Play starts at stage 1 in an initial state, say in state $s_1 \in S$. In s_1, each player $i \in N$ has to choose an action a_1^i from his action set $A_{s_1}^i$. These choices have to be made independently. The chosen joint action $a_1 = (a_1^1, \ldots, a_1^n)$ induces an immediate payoff $r_{s_1}^i(a_1)$ to each player i. Next, play moves to a new state according to the transition probability distribution $p_{s_1 a_1}$, say to state $s_2 \in S$. At stage 2, a new action $a_2^i \in A_{s_2}^i$ has to be chosen by each player i in state s_2. Then, given action combination $a_2 = (a_2^1, \ldots, a_2^n)$, player i receives payoff $r_{s_2}^i(a_2)$ and the play moves to some state s_3 according to the transition probability distribution $p_{s_2 a_2}$, and so on. We assume complete information (i.e. the players know all the data of the stochastic game), full monitoring (i.e. the players observe the present state and the actions chosen by all the players), and perfect recall (i.e. the players remember all previous states and actions).

A *Markov transition structure* Γ^i for player $i \in N$ is given by (1) a nonempty and finite state space S^i; (2) a nonempty and finite action set $A_{s^i}^i$ for each state $s^i \in S^i$; (3) a transition probability distribution $p_{s^i a_{s^i}^i}^i$ over the state space S^i for each state $s^i \in S^i$ and for each action $a_{s^i}^i \in A_{s^i}^i$. Note that, if we also assigned a payoff in every state to every action, then we would obtain the well-known model of a Markov decision problem for player i.

We will now consider a special type of n-player stochastic games, called product-games, in which the transition structure is derived by taking the product of n Markov transition structures. A *product-game* G, associated to the Markov transition structures $\Gamma^1, \Gamma^2, \ldots, \Gamma^n$, is an n-player stochastic game for which (1) the set of players is $N = \{1, \ldots, n\}$; (2) the state space is $S = S^1 \times \cdots \times S^n$; (3) the action set for each player $i \in N$ in each state $s = (s^1, \ldots, s^n) \in S$ is $A_s^i = A_{s^i}^i$; (4) the transition probability distribution p_{sa_s}, for each state $s = (s^1, \ldots, s^n) \in S$ and for each joint action $a_s = (a_s^1, \ldots, a_s^n) \in \times_{i \in N} A_s^i$, is

$$p_{sa_s}(\bar{s}) = \prod_{i \in N} p_{s^i a_s^i}^i(\bar{s}^i)$$

for state $\bar{s} = (\bar{s}^1, \ldots, \bar{s}^n) \in S$. Note that there is no condition imposed on the payoff structure.

Observe that (1) the action space of player i only depends on the i-th coordinate of the state, (2) the i-th coordinate of the transitions from any state s only depend on the i-th coordinate s^i of the state and on the action a_s^i chosen by player i. Therefore, as far as the actions and the transitions are concerned, player i can play on the i-th coordinate of the game G without the interference of the other players. As a consequence, play of the product game G can be viewed as simultaneous play of the n Markov transition structures $\Gamma^1, \ldots, \Gamma^n$, which are linked by payoff functions $r^1, \ldots, r^n$ that may depend on all n current states as well as on all n actions chosen by the players.

Product-games have been introduced in Altman et al. [2005], although in a somewhat different fashion. They only examined two-player games in which the sum of the payoffs is always equal to zero (zero-sum games), and dropped the assumption of full monitoring by letting each player only observe his own coordinate of the present state and only the action chosen by himself. As a result, both players have to make choices without noticing anything about the other player's behavior. They showed that a linear programming formulation is sufficient to solve these games, i.e. to find the value and stationary optimal strategies (cf. the definitions below). Moreover, they displayed possible applications of product-games in wireless networks (see also Altman et al. [2007,a] and Altman et al. [2007,b]).

Note that the class of product-games, as defined in our paper, differs essentially from other known classes of n-player stochastic games. Stochastic games with a single controller (cf. Parthasarathy & Raghavan [1981] or Filar & Vrieze [1996]), i.e. when one player controls the transitions, however, fall into the class of product-games. Indeed, a stochastic game which is controlled by player i can be seen as a product-game in which S^j is a singleton for all players $j \neq i$. Finally, we wish to mention the class of stochastic games with additive transitions (AT-games, cf. Flesch et al. [2007]), i.e. when the transitions are additively decomposable into player-dependent components, in contrast with a product decomposition. Not surprisingly, the structure of product-games and AT-games differ essentially, and product-games require new ideas and an entirely different approach.

From now on, we will consequently use the upper-index for the player and the lower-index for the state. Whenever one of them is omitted, we will then mean a vector in the case of quantities and a product in the case of sets, for all possible players or states respectively. For example, A^i denotes $\times_{s \in S} A^i_s$. Finally, we denote the set of opponents of any player i by $-i := N - \{i\}$. Then, $-i$ in the upper-index will mean a vector or product for all players $j \neq i$. For example, S^{-i} denotes $\times_{j \in N - \{i\}} S^j$.

Strategies. A mixed action x^i_s for player i in state $s \in S$ is a probability distribution on A^i_s. The set of mixed actions for player i in state s is denoted by X^i_s. A mixed action is called completely mixed, if it assigns a positive probability to each available action. A (history dependent) strategy π^i for player i is a decision rule that prescribes a mixed action $\pi^i_s(h) \in X^i_s$ in the present state s depending on the past history h of play (i.e. the sequence of all past states and all past actions chosen by all the players). We use the notation Π^i for the set of strategies for player i. A strategy π^i for player i is called pure if π^i prescribes, for every state and every possible history, one specific action to be played with probability 1. Given a strategy π^i for player i and a history h, the strategy π^i conditional on h, denoted by $\pi^i[h]$, is the strategy which prescribes a mixed action $\pi^i_s[h](h')$ in any present state s for any history h' as if h had happened before h', i.e. $\pi^i_s[h](h') = \pi^i_s(h \oplus h')$, where $h \oplus h'$ is the history consisting of h concatenated by h'. In fact, $\pi^i[h]$ is just the continuation strategy of π^i after history h.

If the mixed actions prescribed by a strategy only depend on the present state then the strategy is called stationary. Thus, the stationary strategy space for player i is $X^i = \times_{s\in S} X^i_s$. We use the notation x^i for stationary strategies for player i, while x^i_s refers to the corresponding mixed action for player i in state s. Note that the set of pure stationary strategies for player i is simply $A^i = \times_{s\in S} A^i_s$.

A joint stationary strategy $x = (x^i)_{i\in N}$ induces a Markov-chain on the state space S with transition matrix $P(x)$, where entry $(s, \bar{s})$ of $P(x)$ gives the transition probability $p_{sx_s}(\bar{s})$ for moving from state s to state $\bar{s}$ when the joint mixed action x_s is played in state s. With respect to this Markov-chain, we can speak of transient and recurrent states. A state is called recurrent if, when starting there, play will eventually return with probability 1; otherwise the state is called transient. If play is in a recurrent state, then this state will be visited infinitely often with probability 1, while transient states can only be visited finitely many times, with probability 1. We can group the recurrent states into minimal closed sets, into so-called ergodic sets. An ergodic set is a collection W of recurrent states with the property that, when starting in any of the states in W, all states in W will be visited infinitely often and play will remain in W forever with probability 1. The period of a state $s \in W$ is defined as the greatest common divisor of all numbers $m \in \mathbb{N}$ such that returning to state s in m steps has a positive probability. It is known that every state in an ergodic set W has the same period, which we denote by $\lambda(W)$. When $\lambda(W) = 1$, the set W is called aperiodic. Moreover, W can be uniquely divided into $\lambda(W)$ pairwise-disjoint cyclic sets $W^1, \ldots, W^{\lambda(W)}$, i.e. when starting in any $s \in W^l$, the process will move through the cyclic sets in the order $W^l, W^{l+1}, \ldots, W^{\lambda(W)}, W^1, \ldots, W^{l-1}, W^l, \ldots$. It is known that there exists a $\mu > 0$ and a stage M such that at any stage $m \geq M$, the process can be, with probability at least μ, in any state of the cyclic set appropriate for the moment. We refer to Kemeny & Snell [1960] for a more detailed discussion on Markov chains.

Let

$$Q(x) := \lim_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} P^m(x); \tag{1}$$

the limit is known to exist (cf. Doob [1953], theorem 2.1, page 175). Entry $(s, \bar{s})$ of the stochastic matrix $Q(x)$, denoted by $q_{sx}(\bar{s})$, is the expected frequency of stages for which the process is in state $\bar{s}$ when starting in s. The matrix $Q(x)$ has the well known properties (cf. Doob [1953]) that

$$Q(x) = Q(x)\,P(x) = P(x)\,Q(x) = Q^2(x). \tag{2}$$

Rewards. For a joint strategy $\pi = (\pi^i)_{i\in N}$ and initial state $s \in S$, the sequences of payoffs are evaluated by the expected average reward, simply reward, which is given for player i by

$$\gamma^i_s(\pi) := \liminf_{M\to\infty} \mathbb{E}_{s\pi} \left(\frac{1}{M} \sum_{m=1}^{M} R^i_m \right) = \liminf_{M\to\infty} \frac{1}{M} \sum_{m=1}^{M} \mathbb{E}_{s\pi} \left(R^i_m \right),$$

where R^i_m is the random variable for the payoff for player i at stage m, and where $\mathbb{E}_{s\pi}$ stands for expectation with respect to the initial state s and the joint strategy π.

With regard to a joint stationary strategy $x = (x^i)_{i\in N}$, we obtain more explicit formulas for the induced reward. Let $r^i_s(x_s)$ denote the expected immediate payoff for player i in state s if the joint mixed action x_s is played. By definition, for every player i's reward we have

$$\gamma^i(x) = Q(x)\,r^i(x), \tag{3}$$

hence by (2) we also obtain

$$\gamma^i(x) = P(x)\,\gamma^i(x) \tag{4}$$

$$\gamma^i(x) = Q(x)\,r^i(x) = Q^2(x)\,r^i(x) = Q(x)\,\gamma^i(x). \tag{5}$$

For any player $i \in N$ and initial state $s \in S$, let

$$v^i_s := \inf_{\pi^{-i}\in\Pi^{-i}} \sup_{\pi^i\in\Pi^i} \gamma^i_s(\pi^i, \pi^{-i}). \tag{6}$$

Here v^i_s is called the minmax-level for player i in state s. Intuitively, this is the highest reward that player i can defend against any strategies of the other players if the initial state is s. Note that, in order to defend his minmax-level, (1) against different joint strategies of players $-i$, player i may have to use different strategies and (2) for the choice of the mixed action at stage m, player i does not need to know player $-i$'s joint strategy for stages beyond m (cf. Neyman [2003]). It is known that the minmax-level of any player i satisfies

$$v_s^i = \min_{x_s^{-i} \in X_s^{-i}} \max_{x_s^i \in X_s^i} \sum_{t \in S} p_{s,(x_s^i, x_s^{-i})}(t)\, v_t^i, \tag{7}$$

which is an easy consequence of the definition of v_s^i and equality (4). Furthermore, by Thuijsman & Vrieze [1991] (their proof is given for only two players but directly extends to the n-player case in combination with Neyman [2003], who showed that the minmax-levels equal the limit of the discounted minmax-levels in n-player stochastic games), there always exists an initial state s in the set $\{t \in S \,|\, v_t^i = \min_{t' \in S} v_{t'}^i\}$ for which players $-i$ have a joint stationary strategy x^{-i} such that $\gamma_s^i(\pi^i, x^{-i}) \leq v_s^i$ for all strategies π^i for player i. In other words, the infimum in expression (6) is attained for such a state s at stationary strategies.

Equilibria. A joint strategy $\pi = (\pi^i)_{i \in N}$ is called a (Nash) ε-equilibrium for initial state $s \in S$, for some $\varepsilon \geq 0$, if

$$\gamma_s^i\left(\sigma^i, \pi^{-i}\right) \leq \gamma_s^i\left(\pi\right) + \varepsilon \qquad \forall \sigma^i \in \Pi^i,\ \forall i \in N,$$

which means that no player can gain more than ε by a unilateral deviation. If π is an ε-equilibrium for all initial states, then we call π an ε-equilibrium. It is clear from the definition of the minmax-level v that if π is an ε-equilibrium then $\gamma_s^i(\pi) \geq v_s^i - \varepsilon$ for each player i and each initial state $s \in S$.

Regarding general stochastic games, the famous game called the Big Match, which was introduced by Gillette [1957] and solved by Blackwell & Ferguson [1968], and the game in Sorin [1986] demonstrated that 0-equilibria do not necessarily exist with respect to the average reward. They made it clear, moreover, that history dependent strategies are indispensable for establishing ε-equilibria, for $\varepsilon > 0$.

For two-player stochastic games, Vieille [2000-a,b] managed to establish the existence of ε-equilibria, for all $\varepsilon > 0$. However, only little is known about n-player stochastic games, and it is unresolved whether they always possess ε-equilibria, for all $\varepsilon > 0$. This is probably the most challenging open problem in the field of stochastic games these days.

For the class of n-player product-games, we will answer this question in the affirmative by proving the existence of 0-equilibria (cf. **Main Theorem 1**). This extends Flesch et al. [2008], where a certain type of aperiodicity was assumed on the transition structure of the product-game (cf. section 2 below, for a precise definition of aperiodicity).

Zero-sum games and optimality. In the development of stochastic games, a special role has been played by the class of zero-sum stochastic games, which are two-player stochastic games for which $r_s^2(a_s) = -r_s^1(a_s)$ (meaning that the sum of the payoffs is zero), for each state s and for each joint action a_s. In these games the two players have completely opposite interests. Mertens & Neyman [1981] showed that for such games $v^2 = -v^1$. Here $v := v^1$ is called the value of the game. They also showed that, if instead of using liminf one uses limsup in the definition of the reward, one would find precisely the same value v. Thus, in a zero-sum game, player 1 wants to maximize his own reward, while at the same time player 2 tries to minimize player 1's reward. For simplicity, let $\gamma = \gamma^1$. A strategy π^1 for player 1 is called ε-optimal for initial state $s \in S$, for some $\varepsilon \geq 0$, if $\gamma_s(\pi^1, \pi^2) \geq v_s - \varepsilon$ for any strategy π^2 of player 2, while a strategy π^2 for player 2 is called ε-optimal for initial state $s \in S$ if $\gamma_s(\pi^1, \pi^2) \leq v_s + \varepsilon$ for any strategy π^1 of player 1. If π^1 or π^2 is ε-optimal for all initial states, then we call π^1 or π^2 an ε-optimal strategy. For simplicity, 0-optimal strategies are briefly called optimal. Mertens and Neyman [1981] proved (even in a stronger form) that both players have ε -optimal strategies for any $\varepsilon > 0$, even though history dependent strategies may be necessary for ε-optimality.

For the class of zero-sum product-games, we will provide a proof that both players have stationary 0-optimal strategies (cf. **Main Theorem 2**). In addition, we analyse the structure of the value of these games.

The structure of the article. In section 2, we discuss preliminary concepts. Then, in section 3, we present our main results, discuss the main difficulties which we encounter when facing periodic product-games and provide a general idea of the proof. The formal proofs are given in sections 4, 5, 6 and 7.

2 Preliminary concepts

Some of the contents of this section is very similar to the decomposition presented in Ross and Varadarajan [1991] for Markov decision problems (i.e. stochastic games with only one player). We also refer to Flesch et al. [2008].

Classification of states. First, we analyse the Markov transition structure Γ^i of each player i separately. (Note that such a separate analysis of the transition structure is only possible due to the fact that player i controls the transitions on his own coordinate.) We distinguish between two basic types of states in the state space S^i of Γ^i. A state $s^i \in S^i$ belongs to type 1 if it is transient for each stationary strategy x^i of player

i. Otherwise, s^i belongs to type 2, in which case player i has a stationary strategy for which s^i is recurrent.

Maximal communicating sets. Two states s^i_1 and s^i_2 of type 2 are said to communicate with each other, if there exists a stationary strategy x^i of player i such that s^i_1 and s^i_2 belong to the same ergodic set. We note that communication between states have been used extensively in the literature of stochastic games (cf. Vieille [2000-a,b], Solan & Vieille [2002], Solan [2003]).

This relationship of communication is an equivalence relation on the set of states of type 2. As such, it induces equivalence classes, which for obvious reasons are called maximal communicating sets. Therefore, every maximal communicating set E^i has the properties that (1) player i can go from any state in E^i to any other state in E^i, with probability 1, possibly in a number of moves without leaving E^i, and (2) if player i decides to leave E^i, the probability that he ever comes back to E^i is strictly less than 1, regardless his strategy (and since the state and action spaces are finite, these probabilities have an upper-bound strictly smaller than 1). The latter observation further implies that (3) the total number of times during the whole play that player i switches from a maximal communicating set to another one is finite with probability 1, regardless the initial state and player i's strategy; (in fact, for every $\rho > 0$ there exists an $L_\rho \in \mathbb{N}$ such that the number of times that play moves from one maximal communicating set to another is at most L^1_ρ with probability at least $1-\rho$); (4) there is always at least one amongst the maximal communicating sets which player i is unable to leave, i.e. there are no transitions to states outside; (5) for any strategy of player i, regardless the initial state, player i eventually *settles*, with probability 1, in one of his maximal communicating sets E^i, i.e. after finitely many stages, player i remains forever in E^i (it is possible that player i would be able to leave E^i with a different strategy).

Let $E^i_{k^i}$, where $k^i \in K^i$, denote the maximal communicating sets for player i. In every state s^i of the communicating set $E^i_{k^i}$, for every $k^i \in K^i$, let $\bar{A}^i_{s^i}$ denote the set of those actions $a^i_{s^i} \in A^i_{s^i}$ which keep play in $E^i_{k^i}$ with probability 1. The sets $\bar{A}^i_{s^i}$ are clearly nonempty. For every state $s = (s^1, \ldots, s^n) \in S$, we also let $\bar{A}^i_s := \bar{A}^i_{s^i}$.

Periodicity and Segments. The period of $E^i_{k^i}$, denoted by $\lambda^i_{k^i}$, is defined as the period of the Markov chain on $E^i_{k^i}$ associated to a stationary strategy x^i of player i that only uses completely mixed actions on $\bar{A}^i_{s^i}$ for all $s^i \in E^i_{k^i}$. (Obviously, the period is independent of the particular choice of x^i.) The cyclic sets of $E^i_{k^i}$ are denoted by $T^i_{k^i}(1), \ldots, T^i_{k^i}(\lambda^i_{k^i})$. For convenience, let $T^i_{k^i}(u \cdot \lambda^i_{k^i} + w) := T^i_{k^i}(w)$ for all $u \in \mathbb{N}$ and

$w \in \{1, \ldots, \lambda^i_{k^i}\}$.

Let $K := \times_{i=1}^n K^i$. Consider the product $E_k := \times_{i=1}^n E^i_{k^i}$ for some $k = (k^1, \ldots, k^n) \in K$. The period of E_k, denoted by λ_k, is defined as the period of the Markov chain on E_k associated to a joint stationary strategy x that only uses joint completely mixed actions on $\bar{A}_s$ for all $s \in E_k$. Clearly, λ_k equals the least common multiple of $\lambda^1_{k^1}, \ldots, \lambda^n_{k^n}$. Notice that this Markov chain has no transient states and consists of a number of ergodic sets, which we call segments. Each segment F has period λ_k and is determined by the starting state. If $s \in E_k$ is the starting state with $s^i \in T^i_{k^i}(l_i)$ for some $l_i \in \{1, \ldots, \lambda^i_{k^i}\}$, for all $i \in N$, then the segment F containing state s has cyclic sets of the form

$$T_F(m) := T^1_{k^1}(l_1 + m - 1) \times \cdots \times T^n_{k^n}(l_n + m - 1), \quad m = 1, \ldots, \lambda_k. \tag{8}$$

We remark that the number of segments within E_k equals the greatest common divisor of $\lambda^1_{k^1}, \ldots, \lambda^n_{k^n}$. For convenience, let $T_F(u \cdot \lambda_F + w) := T_F(w)$ for all $u \in \mathbb{N}$ and $w \in \{1, \ldots, \lambda_F\}$.

Finally, the period of the whole product-game is defined as the least common multiple of the periods of all its segments. In aperiodic product-games (i.e. which have period 1), each set E_k is just one segment.

Restricted games. Take an arbitrary segment F, within some $E_k = \times_{i=1}^n E^i_{k^i}$. By restricting the state space to $F \subset S$, and the action set of every player i in any state $s \in F$ to $\bar{A}^i_s$, we obtain a restricted game $\bar{G}_F$. Note that $\bar{G}_F$ is a stochastic game, but not necessarily a product-game (the state space F of $\bar{G}_F$ is only a product if $F = E_k$).

These restricted games play a key role in the analysis of product-games, which is due to the following observation. As pointed out above, for any initial state and strategies of the players, each player i eventually settles in one of his maximal communicating sets $E^i_{k^i}$, with probability 1. Hence, with probability 1, play will eventually *settle* in a segment $F \subset E_k$ and in the corresponding restricted game $\bar{G}_F$. The study of these restricted games is therefore of great importance.

For a restricted game $\bar{G}_F$, let $\bar{v}^i_{F,s}$ denote the minmax-level of player i in $\bar{G}_F$ for initial state $s \in F$. If, for some player i, the inequality $\bar{v}^i_{F,s} \geq v^i_s$ holds for all initial states $s \in F$, then we call $\bar{G}_F$ satisfactory to player i. Otherwise, $\bar{G}_F$ is called unsatisfactory to player i. In words, if $\bar{G}_F$ is satisfactory to player i, then player i weakly prefers $\bar{G}_F$ to G, as far as his minmax-level is concerned on F. Let $\mathcal{F}^*$ denote the set of segments F such that $\bar{G}_F$ is satisfactory to all players. Further, let $\mathcal{F}[i]$ denote the set of segments F such that $\bar{G}_F$ is unsatisfactory to player i but $\bar{G}_F$ is satisfactory to all players $j \in \{1, \ldots, i-1\}$. Obviously, $\mathcal{F}^*, \mathcal{F}[1], \ldots, \mathcal{F}[n]$ forms a partition of all segments.

$1,-1$ $\to (1,1)$	$0,0$ $\to (1,2)$
$0,0$ $\to (2,1)$	$0,0$ $\to (2,2)$

state $(1,1)$

$3,-1$ $\to (1,3)$
$0,0$ $\to (2,3)$

state $(1,2)$

$-3,1$ $\to (1,2)$
$0,0$ $\to (2,2)$

state $(1,3)$

$1,-2$ $\to (3,1)$	$0,0$ $\to (3,2)$

state $(2,1)$

$2,-1$ $\to (3,3)$

state $(2,2)$

$0,0$ $\to (3,2)$

state $(2,3)$

$-1,2$ $\to (2,1)$	$0,0$ $\to (2,2)$

state $(3,1)$

$0,0$ $\to (2,3)$

state $(3,2)$

$0,-1$ $\to (2,2)$

state $(3,3)$

Figure 1: Game of Example 1

$\star$ *Example 1.* As an illustration, consider the product-game with two players given in figure 1. This is a game with nine states. In each state, the actions of player 1 are represented by the rows, and the actions of player 2 by the columns. So each cell of each state corresponds to a pair of actions. In each cell, the two payoffs to the respective players are given in the upper-left corner, while the next state is indicated in the bottom-right corner. In this game all the transitions are pure, i.e. each transition probability distribution assigns probability 1 to a certain state.

The underlying Markov transition structure for player 1 is given by state space $S^1 = \{1,2,3\}$, action sets $A^1_1 = \{1,2\}$, $A^1_2 = A^1_3 = \{1\}$, and transitions

$$p^1_{11} = (1,0,0),\ \ p^1_{12} = (0,1,0),\ \ p^1_{21} = (0,0,1),\ \ p^1_{31} = (0,1,0).$$

So in state 1, player 1 can either stay or leave for state 2, while he moves between state 2 and 3 back and forth. Regarding the classification of the states in S^1, both $E^1_I := \{1\}$ and $E^1_{II} := \{2,3\}$ are maximal communicating sets, with index-set $K^1 = \{I, II\}$. Moreover, E^1_I is aperiodic (i.e. has periodicity 1) whereas E^1_{II} has periodicity 2. As for the actions which keep play in these maximal communicating sets, we obtain $\bar{A}^1_1 = \bar{A}^1_2 = \bar{A}^1_3 = \{1\}$.

The underlying Markov transition structure for player 2 is identical. So, the state space is $S^2 = \{1,2,3\}$, the action sets are $A_1^2 = \{1,2\}$, $A_2^2 = A_3^2 = \{1\}$, and the transitions are

$$p_{11}^2 = (1,0,0),\ \ p_{12}^2 = (0,1,0),\ \ p_{21}^2 = (0,0,1),\ \ p_{31}^2 = (0,1,0).$$

Further, $E_I^2 := \{1\}$ and $E_{II}^2 := \{2,3\}$ are maximal communicating sets, with index-set $K^2 = \{I, II\}$. The maximal commincating set E_I^2 is aperiodic, whereas E_{II}^2 has periodicity 2, and $\bar{A}_1^2 = \bar{A}_2^2 = \bar{A}_3^2 = \{1\}$.

Note that $E_{(I,I)} = E_I \times E_I = \{(1,1)\}$, $E_{(I,II)} = \{(1,2),(1,3)\}$ and $E_{(II,I)} = \{(2,1),(3,1)\}$ all consist of one segment, which we denote by $F_{(I,I)}$, $F_{(I,II)}$ and $F_{(II,I)}$ respectively, while $E_{(II,II)} = \{2,3\}^2$ falls apart into two segments, i.e. segment $F_{(II,II),1} = \{(2,2),(3,3)\}$ and segment $F_{(II,II),2} = \{(2,3),(3,2)\}$.

There are five restricted games corresponding to these five different segments. For instance, the restricted game $\bar{G}_{F_{(I,I)}}$ consists of the top-left cell in state $(1,1)$, while $\bar{G}_{F_{(II,I)}}$ consists of the left cells of states $(2,1)$ and $(3,1)$. Note that, in every restricted game, the reward is unique to every player. $\star$

3 The main results

For the class of product-games, we present the following result concerning existence of equilibria.

Main Theorem 1. *There exists a 0-equilibrium in every n-player product-game.*

In addition, for the special case of two-player zero-sum product-games, we show the existence of stationary solutions.

Main Theorem 2. *In two-player zero-sum product-games, both players have a stationary 0-optimal strategy.*

Existence of stationary 0-equilibria. Our construction for Main Theorem 1 will only provide 0-equilibria in history-dependent strategies. It remains unclear whether 0-equilibria always exist within the class of stationary strategies. This question is already challenging in the situation where each player i's state space S^i is just one aperiodic maximal communicating set. In this case, the whole state space S is just one segment. Even though, corollary 6 below (through corollary 10) will yield for such a game that all minmax-levels are constant on the whole state space S, it is still not evident how one should get a grip on the problem.

The difference between the periodic and aperiodic cases. Interestingly, it turns out that the period of the product-game plays a central role in the analysis. In Flesch et al. [2008], the special case of aperiodic product-games has been extensively studied, and the validity of Main Theorem 1 has been shown for all aperiodic product-games. The approach presented there is applicable for a large part to periodic product-games as well, but the periodic case poses a number of additional problems that we have to address. The main cause of these problems is that, in the periodic case, as we discussed in section 2, if $E^i_{k^i}$ denotes a maximal communicating set for every player i, then the product $E_k = \times_{i=1}^n E^i_{k^i}$ may fall apart into a number of segments, which do not communicate with each other. The main problems that we encounter are the following:

A. Several properties which Flesch et al. [2008] derived for the sets E_k in the aperiodic case do not hold for the periodic case. Luckily, however, we are able to derive similar properties for each segment of the sets E_k. (For example, the minmax-levels of the players are no longer constants on the whole set E_k, just on each segement of E_k separately, cf. corollary 6 together with corollary 10.)

B. The central lemma of the aperiodic case loses its validity for periodic product-games and has to be modified. We refer to lemma 3 and the remark after that.

C. In the aperiodic case, we often used that, moving to a set E_k can be achieved by letting each player i move to $E^i_{k^i}$. This is insufficient for the periodic case, as we need to move to certain segments within E_k. Note that the segment which the players enter in E_k will be determined by the collection $(s^i, m^i)_{i=1}^n$, where $s^i \in E^i_{k^i}$ is the state and m^i is the stage at which player i enters $E^i_{k^i}$. Thus, it becomes a real precision work to arrive at the right segment and not at another one within E_k. In particular, we refer to the proof of lemma 3 and the example after that.

An attempt to transform periodic product-games into aperiodic ones. In order to show Main Theorem 1, one could try to transform every periodic product-game G into an aperiodic one G' and hope that the 0-equilibrium in G' reveals a 0-equilibrium for the original product-game G. For example, in the context of Markov chains, it is known that if P is the transition matrix of a Markov chain on finitely many states, then for any $\mu \in (0, 1]$, the transition matrix $\mu \cdot P + (1 - \mu) \cdot I$, where I is the identity matrix, induces the same ergodic structure and the same set of invariant distributions. This is particularly interesting when P is periodic, as $\mu \cdot P + (1 - \mu) \cdot I$ is aperiodic for all $\mu \in (0,1)$.

In periodic product-games, such transformations are bound to fail, for the following reason. Consider a set $E_k = \times_{i=1}^{n} E^i_{k^i}$, where $E^i_{k^i}$ denotes a maximal communicating set for every player i. As we discussed in section 2, if the product-game is aperiodic, then E_k is one segment and all states in E_k communicate. On the other hand, if the product-game is periodic, then E_k may fall apart into several segments, which do not communicate with each other. Hence, any transformation which would unite these segments into one segment, would change the structure of the game so radically that the 0-equilibrium that one finds in G' will not generally correspond to a 0-equilibrium in the original product-game G.

The general idea of the construction of a 0-equilibrium. The general idea of the construction of an equilibrium η, for a product-game G, is as follows. The equilibrium η will prescribe to follow a joint strategy π, unless some player i deviates from π^i by playing an action outside the support of π^i (i.e. an action on which π^i puts probability zero). If player i deviates in such a way, then from the next state, say state s, players $-i$ switch to a joint stationary strategy y^{-i} and push down player i's reward to his minmax-level v^i_s. In fact, y^{-i} acts as a threat strategy, whose task is to force player i to follow the prescriptions of π^i. Punishment with y^{-i} will be shown to be severe enough. Finally, our construction will guarantee that no deviation inside the support of π^i (such deviations are hard to detect) is profitable for any player i.

Now let us briefly describe the construction of π, which shows a number of similarities with the construction in Vieille [2000-a,b] and Flesch et al. [2008]. The joint strategy π prescribes to play in the following way:

(1) When entering some segment F, with $F \in \mathcal{F}^*$ (i.e. the corresponding restricted game $\bar{G}_F$ is satisfactory to all players): In this case, π will prescribe to stay on F and play a certain equilibrium in $\bar{G}_F$. Here, the players collect "high" payoffs. (Cf. solvable sets in Vieille [2000-a,b].)

(2) When entering some segment F, with $F \in \mathcal{F}[i]$ (i.e. the corresponding restricted game $\bar{G}_F$ is unsatisfactory to player i): In this case, π will prescribe player i to exit $E^i_{k^i}$ (and thereby to leave F), while all other players wait for player i's exit patiently. It will be taken care of that no player's minmax-level drops in expectation by this exit. Payoffs in F are disregarded. (Cf. controlled sets in Vieille [2000-a,b].)

(3) Outside all joint maximal communicating sets: In this case, π will let the players play for their future perspectives. Payoffs in these states are disregarded.

Note that, according to π, play will surely settle in a restricted game belonging to case (1).

4 The formal proofs of Main Theorems 1 and 2

In this section, we provide a proof for Main Theorems 1 and 2. We will focus on Main Theorem 1, as Main Theorem 2 will follow (cf. the end of section 4.3) along the way without major additional difficulties. In section 4.1, we examine restricted games. In section 4.2 we analyze the minmax-levels of the players in so-called simple product-games, and then in section 4.3 we extend this to the general case. By combining these results, we prove Main Theorem 1 in section 4.4.

4.1 Analysis of the restricted games

We know that, for any strategies of the players and for any initial state, play will eventually settle in some restricted game, with probability 1. Therefore, it is essential to know what perspectives each restricted game can offer to the players in terms of minmax-levels and equilibrium rewards. Consider an arbitrary restricted game $\bar{G}_F$, corresponding to segment F. First, we analyse the minmax-levels of the players in $\bar{G}_F$, and then we discuss possible equilibrium rewards within $\bar{G}_F$.

First we will show that each player i's minmax-level in $\bar{G}_F$ is constant. Moreover, players $-i$ can make sure in $\bar{G}_F$ that player i's reward is at most his minmax-level (i.e. the infimum is attained in (6) for the game $\bar{G}_F$).

Lemma 1 *Let G be a product-game. Consider the restricted game $\bar{G}_F$, corresponding to some segment F, and an arbitrary player i. Then, the minmax-level $\bar{v}^i_F$ of player i in $\bar{G}_F$ is constant, i.e. $\bar{v}^i_{F,s} = \bar{v}^i_{F,t} (=: \bar{v}^i_F)$ for all states $s, t \in F$. Moreover, in $\bar{G}_F$, players $-i$ have a joint stationary strategy x^{-i} which guarantees that player i's reward from any initial state $s \in F$ is at most his minmax-level $\bar{v}^i_F$, i.e. for all strategies π^i for player i in $\bar{G}_k$ we have*

$$\bar{\gamma}^i_s(\pi^i, x^{-i}) \leq \bar{v}^i_F,$$

where $\bar{\gamma}$ denotes the reward for the game $\bar{G}_F$.

We will prove this lemma in section 5.

$\star$ As an illustration, we now revisit the game in example 1. We find in accordance with lemma 1 that the minmax-levels of the players are constant in the restricted games. Indeed, for player 1 we have that

$$\bar{v}^1_{F_{(I,I)}} = 1,\ \bar{v}^1_{F_{(I,II)}} = 0,\ \bar{v}^1_{F_{(II,I)}} = 0,\ \bar{v}^1_{F_{(II,II),1}} = 1,\ \bar{v}^1_{F_{(II,II),2}} = 0, \tag{9}$$

while for player 2 that

$$\bar{v}_F^2 = -\bar{v}_F^1 \tag{10}$$

for any segment F. $\star$

Next, we present a possible equilibrium for the restricted game $\bar{G}_F$. We show that there exists a 0-equilibrium in $\bar{G}_F$ in which, if no player deviates, the players' future expectations remain unchanged during the whole play. Note that Flesch et al. [1997] (with 3 players) and Simon [2003] (with only 2 players) constructed examples proving that such a result would not hold for all stochastic games in general.

Lemma 2 *Let G be a product-game. Consider the restricted game $\bar{G}_F$, corresponding to some segment F. Then, there exists a 0-equilibrium π in $\bar{G}_F$ such that the corresponding rewards are independent of the initial state and all the continuation rewards remain unchanged with probability 1 during the whole play. More precisely, the reward $\bar{\gamma}_s^i(\pi[h])$ is independent of the initial state $s \in F$ and the history h, given h occurs with a positive probability with respect to π. Here $\bar{\gamma}$ denotes the reward for the restricted game $\bar{G}_F$.*

The proof is the same as for lemma 3.7 in Flesch et al. [2008]. Here, we only provide a brief outline. Note:

(i) The minmax-levels of the players in $\bar{G}_F$ are constant, by lemma 1.

(ii) The set of feasible rewards in $\bar{G}_F$ (i.e. the rewards that can be obtained by some joint strategy) is the same from any initial state in F. This is an immediate consequence of the fact that the players can move from any state in F to any other one in F, in a finite number of steps.

(iii) The extreme points of the set of feasible rewards are induced by pure stationary strategies. This is shown in Flesch et al. [2008], based on Dutta [1995].

Given these three observations, this game situation is almost identical to a repeated game. The proof further is direct and uses ideas and arguments that are standard in various kinds of Folk-theorems.

4.2 The minmax-levels in simple product-games

A product-game G is called <u>simple</u> if, within every restricted game $\bar{G}_F$, every player i has a unique payoff, i.e. $r_s^i(a_s) = r_t^i(b_t)$ for all states $s, t \in F$ and for all joint actions

$1,-1$ $\rightarrow (1,1)$	$0,0$ $\rightarrow (1,2)$
$0,0$ $\rightarrow (2,1)$	$0,0$ $\rightarrow (2,2)$

state $(1,1)$

$0,0$ $\rightarrow (1,3)$
$0,0$ $\rightarrow (2,3)$

state $(1,2)$

$0,0$ $\rightarrow (1,2)$
$0,0$ $\rightarrow (2,2)$

state $(1,3)$

$0,0$ $\rightarrow (3,1)$	$0,0$ $\rightarrow (3,2)$

state $(2,1)$

$1,-1$ $\rightarrow (3,3)$

state $(2,2)$

$0,0$ $\rightarrow (3,2)$

state $(2,3)$

$0,0$ $\rightarrow (2,1)$	$0,0$ $\rightarrow (2,2)$

state $(3,1)$

$0,0$ $\rightarrow (2,3)$

state $(3,2)$

$1,-1$ $\rightarrow (2,2)$

state $(3,3)$

Figure 2: Game of Example 2

$a_s \in \bar{A}_s$, $b_t \in \bar{A}_t$. Let z_F^i denote this unique payoff for player i in the restricted game $\bar{G}_F$. Thus, in simple product-games, when play settles in one of the restricted games $\bar{G}_F$, the rewards of the players will equal z_F.

$\star$ *Example 2:* Consider the simple product-game G with two players given in figure 2. This game is obtained from the game in example 1 by replacing all payoffs by 0 in the restricted games corresponding to segments $F_{(I,II)} = \{(1,2),(1,3)\}$ and $F_{(II,I)} = \{(2,1),(3,1)\}$, and by replacing all payoffs for player 1 by 1 in the restricted game corresponding to segment $F_{(II,II),1} = \{(2,2),(3,3)\}$. This game is simple according to the definition above.

Let us examine the players' minmax-levels in G. For player 1, we will argue that

$$v_s^1 = \begin{cases} 0 & \text{if } s \in F_{(II,I)} \cup F_{(II,II),2} = \{(2,1),(3,1),(2,3),(3,2)\} \\ 1 & \text{if } s \in F_{(I,I)} \cup F_{(I,II)} \cup F_{(II,II),1} = \{(1,1),(1,2),(1,3),(2,2),(3,3)\}. \end{cases} \tag{11}$$

Obviously, $v_s^1 = 0$ for $s \in \{(2,3),(3,2)\}$, while player 1's minmax-level is also 0 for initial states $(2,1)$ and $(3,1)$ in view of player 2's first action. Now consider an arbitrary other initial state $s \in \{(1,1),(1,2),(1,3),(2,2),(3,3)\}$. Obviously, $v_s^1 \leq 1$. On the other hand, player 1 can guarantee reward 1 for state s by the pure stationary strategy x^1 which plays action 1 in states $(1,1)$ and $(1,2)$, while action 2 in state $(1,3)$. Hence, $v_s^1 = 1$ indeed.

We similarly find that

$$v_s^2 = -v_s^1 \tag{12}$$

for all $s \in S$. $\star$

Consider a state $s \in S$ within a simple product-game. The following lemma considers the situation where s^i is of type 2 for player i. Then, as we know, s^i belongs to a cyclic set $T^i_{k^i}(m^i)$ of some maximal communicating set $E^i_{k^i}$. Suppose player i can choose between one of the following two options: (1) player i can choose any state t^i in the next cyclic set $T^i_{k^i}(m^i+1)$, and the new state of the game becomes (t^i, s^{-i}), or (2) player i can choose any action a^i_s in state s^i, and the new state of the game becomes (t^i, s^{-i}), $t^i \in S^i$, with probability $p^i_{s^i a^i_s}(t^i)$. Note that players $-i$ remain in state s^{-i} in either case. In the following lemma we show that option 1 is always at least as good as option 2, as far as player i's minmax-level is concerned. A similar statement is valid for players $-i$.

Lemma 3 *Let G be a simple product-game. Take an arbitrary player i and a state $s = (s^i, s^{-i}) \in S$.*

(1) Suppose that state s^i is of type 2 for player i, and belongs to cyclic set $T^i_{k^i}(m^i)$ of some maximal communicating set $E^i_{k^i}$. Consider any action $a^i_{s^i} \in A^i_{s^i}$ in state s^i for player i. Then, for any state $t^i \in T^i_{k^i}(m^i+1)$, we have

$$\sum_{u^i \in S^i} p^i_{s^i a^i_{s^i}}(u^i)\, v^i_{(u^i, s^{-i})} \le v^i_{(t^i, s^{-i})}.$$

(2) Suppose that state s^j is of type 2 for every player $j \neq i$, and belongs to cyclic set $T^j_{k^j}(m^j)$ of some maximal communicating set $E^j_{k^j}$. Thus, $s^{-i} \in T^{-i}_{k^{-i}}(m^{-i})$. Consider any joint action $a^{-i}_s \in A^{-i}_s$ for players $-i$. Then, for any joint state $t^{-i} \in T^{-i}_{k^{-i}}(m^{-i}+1)$, we have

$$\sum_{u^{-i} \in S^{-i}} p^{-i}_{s^{-i} a^{-i}_{s^{-i}}}(u^{-i})\, v^i_{(s^i, u^{-i})} \ge v^i_{(s^i, t^{-i})}.$$

We will prove this lemma in Section 6. The proof is far from straightforward. Nevertheless, we will provide now an intuition behind the lemma. Consider part 1 of the lemma, and the two options we described before the lemma. By taking option 1, player i is certain to remain in the same maximal communicating set $E^i_{k^i}$. On the other hand, by playing an action $a^i_{s^i}$ in option 2, player i possibly leaves $E^i_{k^i}$ and strategically restricts himself, as he will not be able to return to $E^i_{k^i}$ with probability 1 (cf. property 2 of maximal communicating sets in section 2). In this sense, waiting in $E^i_{k^i}$ provides no worse future prospects. Hence the inequality for player i's expected minmax-levels after the transition. A similar reasoning supports part 2 of the lemma.

$\star$ As an illustration for part 1 of the lemma, we revisit the simple product-game in example 2. Consider player $i = 1$, state $s = (1, 1)$, and action 2 and state $t^1 = 1$ for player 1. Then, $v^1_{(2,1)} = 0 \leq 1 = v^1_{(1,1)}$, in accordance with part 1 of the lemma. $\star$

Remark. A special case arises when the product-game is aperiodic. Consider part 1 of the lemma. Then, as $T^i_{k^i}(m^i) = T^i_{k^i}(m^i + 1) = E^i_{k^i}$, we obtain for every action $a^i_{s^i}$ of player i that

$$\sum_{u^i \in S^i} p^i_{s^i a^i_{s^i}}(u^i)\, v^i_{(u^i,s^{-i})} \leq v^i_{(s^i,s^{-i})}. \tag{13}$$

This means that even if player i had a solitary move, i.e. he could play an arbitrary action $a^i_{s^i}$ in state (s^i, s^{-i}), while every other player j remains in the same state s^j, he would not be able to improve on his minmax-level in expectation. This was in fact the central result for the aperiodic case in Flesch et al. [2008] (cf. lemma 3.2).

This is, however, no longer valid for periodic product-games. In the game in example 2, for instance, a solitary move for player 1 in state $(3, 2)$ would lead to state $(2, 2)$, improving player 1's minmax-level. Hence, inequality (13) would not hold.

Suppose player i is in a state of type 2 within a maximal communicating set $E^i_{k^i}$. We show that, irrespectively of the joint action chosen by players $-i$, the actions which keep him in $E^i_{k^i}$ with probability 1 provide the best expected minmax-level after the transition.

Lemma 4 *Let G be a simple product-game. Take an arbitrary player i.*

(1) Let $s \in S$ be such that s^i is of type 2. Consider any actions $a^i_s \in \bar{A}^i_s$ and $b^i_s \in A^i_s$ for player i and any joint action $a^{-i}_s \in A^{-i}_s$ for players $-i$. Then,

$$\sum_{t \in S} p_{s,(a^i_s,a^{-i}_s)}(t)\, v^i_t \geq \sum_{t \in S} p_{s,(b^i_s,a^{-i}_s)}(t)\, v^i_t.$$

(2) Let $s \in S$ be such that s^j is of type 2 for all players $j \neq i$. Consider any joint actions $a^{-i}_s \in \bar{A}^{-i}_s$ and $b^{-i}_s \in A^{-i}_s$ for players $-i$ and any action $a^i_s \in A^i_s$ for player i. Then,

$$\sum_{t \in S} p_{s,(a^i_s,a^{-i}_s)}(t)\, v^i_t \leq \sum_{t \in S} p_{s,(a^i_s,b^{-i}_s)}(t)\, v^i_t.$$

Proof. We will prove part 1; the proof of part 2 is similar.

So, consider part 1. Since s^i is of type 2, s^i belongs to a cyclic set $T^i_{k^i}(m^i)$ of some maximal communicating set $E^i_{k^i}$. Then, by playing action a^i_s, player i actually moves to the next cyclic set $T^i_{k^i}(m^i+1)$. Hence, by part 1 of lemma 3, for any $t^{-i} \in S^{-i}$ we have

$$\sum_{t^i \in S^i} p^i_{s^i a^i_s}(t^i)\, v^i_{(t^i,t^{-i})} \geq \sum_{t^i \in S^i} p^i_{s^i b^i_s}(t^i)\, v^i_{(t^i,t^{-i})}.$$

Therefore,

$$\begin{aligned}
\sum_{t \in S} p_{s,(a^i_s,a^{-i}_s)}(t)\, v^i_t &= \sum_{t^{-i} \in S^{-i}} p^{-i}_{s^{-i}a^{-i}_s}(t^{-i}) \left[\sum_{t^i \in S^i} p^i_{s^i a^i_s}(t^i)\, v^i_{(t^i,t^{-i})}\right] \\
&\geq \sum_{t^{-i} \in S^{-i}} p^{-i}_{s^{-i}a^{-i}_s}(t^{-i}) \left[\sum_{t^i \in S^i} p^i_{s^i b^i_s}(t^i)\, v^i_{(t^i,t^{-i})}\right] \\
&= \sum_{t \in S} p_{s,(b^i_s,a^{-i}_s)}(t)\, v^i_t,
\end{aligned}$$

completing the proof. ∎

In view of (7), the lemma above has the following implication.

Corollary 5 *Let G be a simple product-game. Take an arbitrary player i.*

(1) Let $s \in S$ be such that s^i is of type 2. Consider any action $a^i_s \in \bar{A}^i_s$ for player i and any joint action $a^{-i}_s \in A^{-i}_s$ for players $-i$. Then,

$$\sum_{t \in S} p_{s,(a^i_s,a^{-i}_s)}(t)\, v^i_t \geq v^i_s.$$

(2) Let $s \in S$ be such that s^j is of type 2 for all players $j \neq i$. Consider any joint action $a^{-i}_s \in \bar{A}^{-i}_s$ for players $-i$ and any action $a^i_s \in A^i_s$ for player i. Then,

$$\sum_{t \in S} p_{s,(a^i_s,a^{-i}_s)}(t)\, v^i_t \leq v^i_s.$$

As we know, all states within a segment F communicate through joint actions $a_s \in \bar{A}_s$, $s \in F$. Since, due to the above corollary, these actions do not change the minmax-levels of the players in expectation, we may conlude the following result.

Corollary 6 *Let G be a simple product-game, and F a segment. Then, the minmax-level v^i of every player i is constant on F, i.e. $v^i_s = v^i_t (=: v^i_F)$ for all $s,t \in F$.*

$\star$ As an illustration, consider the simple product-game in example 2. In view of (11) and (12), we find, in accordance with the corollary above, that both players minmax-levels are constant on each segment. $\star$

The next lemma is more specific on constant minmax-levels.

Lemma 7 *Let G be a simple product-game. Take an arbitrary player i.*

(1) Consider a cyclic set $T^i_{k^i}(m)$ of some maximal communicating set $E^i_{k^i}$ for player i. Then, for any two states $s^i, t^i \in T^i_{k^i}(m)$ and any joint state $s^{-i} \in S^{-i}$ of players $-i$, the minmax-level of player i satisfies $v^i_{(s^i,s^{-i})} = v^i_{(t^i,s^{-i})}$.

(2) Consider a cyclic set $T^j_{k^j}(m^j)$ of some maximal communicating set $E^j_{k^j}$ for all players $j \neq i$. Then, for any two joint states $s^{-i}, t^{-i} \in \times_{j \neq i} T^j_{k^j}(m^j)$ and any state $s^i \in S^i$ of player i, the minmax-level of player i satisfies $v^i_{(s^i,s^{-i})} = v^i_{(s^i,t^{-i})}$.

Proof. We will prove part 1; the proof of part 2 is similar.

Take any joint state $s^{-i} \in S^{-i}$ for players $-i$. Let $s^i \in T^i_{k^i}(m)$ be such that $v^i_{(s^i,s^{-i})} \leq v^i_{(w^i,s^{-i})}$ for all $w^i \in T^i_{k^i}(m)$, and let $t^i \in T^i_{k^i}(m)$. It suffices to show that $v^i_{(s^i,s^{-i})} = v^i_{(t^i,s^{-i})}$.

Take any state $u^i \in T^i_{k^i}(m-1)$, with $T^i_{k^i}(0) := T^i_{k^i}(\lambda^i_{k^i})$, and an action $a^i_{u^i} \in \bar{A}^i_{u^i}$ such that $a^i_{u^i}$ in state u^i induces transition to state t^i with a positive probability (obviously, such a state and action exist, due to the definitions of cyclic sets). Note that part 1 of lemma 3 yields

$$\sum_{w^i \in S^i} p^i_{u^i a^i_{u^i}}(w^i)\, v^i_{(w^i,s^{-i})} \leq v^i_{(s^i,s^{-i})}.$$

As action $a^i_{u^i}$ in state u^i only induces transition to states in $T^i_{k^i}(m)$, by the choice of s^i, we have $v^i_{(s^i,s^{-i})} = v^i_{(w^i,s^{-i})}$ for all $w^i \in S^i$ for which $p^i_{u^i a^i_{u^i}}(w^i) > 0$. In particular, $v^i_{(s^i,s^{-i})} = v^i_{(t^i,s^{-i})}$, which completes the proof. ∎

Recall that, in simple product-games, z^i_F denotes the unique payoff for player i in the restricted game $\bar{G}_F$. When $z^i_F \geq v^i_F$ for every player i, i.e. when the restricted game $\bar{G}_F$ is satisfactory to all players, then we can let the players stay in F and collect the individually rational rewards z_F. However, we still have to examine what happens in the situation where $z^i_F < v^i_F$ for some player i, i.e. when the restricted game $\bar{G}_F$ is unsatisfactory to player i. The next lemma proposes a way for player i to exit $\bar{G}_F$, by playing a certain action a^i_s in one of the cyclic sets of F. A similar result holds for players $-i$.

Lemma 8 *Let G be a simple product-game, and F a segment within some $E_k = \times_{i=1}^n E^i_{k^i}$ for some $k = (k^1, \ldots, k^n) \in K$. Suppose that F has cyclic sets of the form (cf. (8))*

$$T_F(m) := T^1_{k^1}(m) \times \cdots \times T^n_{k^n}(m), \;\; m = 1, \ldots, \lambda_k.$$

Consider player i. Let z^i_F denote player i's unique payoff in the restricted game $\bar{G}_F$, and v^i_F be player i's minmax-level on F in the game G (a constant, cf. corollary 6).

(1) Suppose $z^i_F < v^i_F$. Then, for player i, there exists an $m \in \{1, \ldots, \lambda_k\}$, a state $s^i \in T^i_{k^i}(m)$, and an action $a^i_{s^i} \in A^i_{s^i} - \bar{A}^i_{s^i}$ in state s^i such that if player i plays action $a^i_{s^i}$ in any state $s = (s^i, s^{-i}) \in T_F(m)$, then player i's minmax-level cannot decrease in expectation from state s, regardless the actions played by players $-i$. More precisely, for any $a^{-i}_s \in A^{-i}_s$ we have

$$\sum_{t \in S} p_{s,(a^i_{s^i}, a^{-i}_s)}(t)\, v^i_t \geq v^i_F.$$

(2) Suppose $z^i_F > v^i_F$. Then, for players $-i$, there exists an $m \in \{1, \ldots, \lambda_k\}$, a joint state $s^{-i} \in T^{-i}_{k^{-i}}(m)$, and a joint action $a^{-i}_{s^{-i}} \in A^{-i}_{s^{-i}} - \bar{A}^{-i}_{s^{-i}}$ (i.e. at least one player $j \neq i$ plays outside $\bar{A}^j_{s^j}$) in joint state s^{-i} such that if players $-i$ play joint action $a^{-i}_{s^{-i}}$ in any state $s = (s^i, s^{-i}) \in T_F(m)$, then player i's minmax-level cannot increase in expectation from state s, regardless the action played by player i. More precisely, for any $a^i_s \in A^i_s$ we have

$$\sum_{t \in S} p_{s,(a^i_s, a^{-i}_{s^{-i}})}(t)\, v^i_t \leq v^i_F.$$

We will prove this lemma in Section 7.

$\star$ As an illustration for part 1 of the lemma, we revisit the simple product-game in example 2. Consider the segment $F_{(I,II)} = \{(1,2), (1,3)\}$, where $v^1_{F_{(I,II)}} = 1 > 0 = z^1_{F_{(I,II)}}$, by (11). Segment $F_{(I,II)}$ has period 2 and two cyclic sets, i.e. $T_{F_{(I,II)}}(1) = \{(1,2)\}$ and $T_{F_{(I,II)}}(2) = \{(1,3)\}$. Notice that player 1 can exit $F_{(I,II)}$ through action 2 in state $(1,3)$, and by doing so, play moves to state $(2,2)$, where his minmax-level is $v^1_{(2,2)} = 1$. Note that $v^1_{(2,2)} \geq v^1_{(1,3)}$, in accordance with part 1 of the lemma. Thus, one can choose $m = 2$, $s^1 = 1$ and $a^1_{s^1} = 2$. $\star$

4.3 The minmax-levels in general product-games

Take an arbitrary product-game G. The next lemma presents a natural way of transforming G into a simple product-game $\widetilde{G}$, by replacing payoffs by minmax-levels, and claims that the minmax-levels of the players remain unchanged under this transformation. The idea to replace payoffs by minmax-levels, in the context of stochastic games, also appeared in Solan [1999] and in a more sophisticated way in Solan & Vohra [2002].

Lemma 9 *Take an arbitrary product-game G, with v_s^i denoting the minmax-level for every player i and for every state $s \in S$. Let $\bar{v}_F^i$ denote player i's minmax-level in every restricted game $\bar{G}_F$ (which is constant, cf. lemma 1). Let $\widetilde{G}$ denote the simple product-game which is derived from G by replacing every player i's payoffs in every restricted game $\bar{G}_F$ by his minmax-level $\bar{v}_F^i$. Further, let w_s^i denote every player i's minmax-level in $\widetilde{G}$ in state s.*

Then, the minmax-levels of the product-games G and $\widetilde{G}$ are equal, i.e. $v_s^i = w_s^i$ for all players i and for all states $s \in S$.

The proof is almost the same as for lemma 3.6 in Flesch et al. [2008]. Here, we only provide a brief outline. We will argue that $v_s^i \leq w_s^i$ for all states $s \in S$. Since $v_s^i \geq w_s^i$ for all $s \in S$ follows in a similar fashion, the proof will then be complete.

In order to show that $v_s^i \leq w_s^i$ for all $s \in S$, we will prove for the game G that players $-i$ have a joint stationary strategy x^{-i} which guarantees that player i's reward from any initial state $s \in S$ is at most w_s^i, i.e. for all strategies π^i for player i we have

$$\gamma_s^i(\pi^i, x^{-i}) \leq w_s^i. \tag{14}$$

Note first that the minmax-level w_s^i equals some constant w_F^i on every segment F, by applying corollary 6 to the game $\widetilde{G}$. We construct the joint stationary strategy x^{-i} by distinguishing the following three mutually exclusive cases.

Case 1: States $s \in S$ which do not belong to any segment: In this case, let $x_s^{-i} \in X_s^{-i}$ be a joint mixed action for players $-i$ such that for any mixed action $x_s^i \in X_s^i$ of player i we have

$$\sum_{t \in S} p_{s,(x_s^{-i}, x_s^i)}(t)\, w_t^i \leq w_s^i.$$

By expression (7) for player i's minmax-level w^i in $\widetilde{G}$, such a joint mixed action exists.

Case 2: In a segment F with $\bar{v}_F^i \leq w_F^i$: In this case, players $-i$ play a joint stationary strategy in the corresponding restricted game $\bar{G}_F$ (which is a part of the original game G) as in lemma 1.

Case 3: In a segment F with $\bar{v}_F^i > w_F^i$: In this case, part 2 of lemma 8 (for the game $\widetilde{G}$ with minmax-level w^i for player i) provides a joint "exit" state and a joint "exit" action for players $-i$. So, in this state players $-i$ play this "exit" action, while in all other states $s \in F$ they play an arbitrary joint completely mixed on $\bar{A}_s^{-i}$.

Take an arbitrary strategy π^i for player i, and consider (π^i, x^{-i}) with an arbitrary initial state $s \in S$. As we know, play will eventually settle in a restricted game $\bar{G}_F$. Observe that (1) the minmax-level w^i cannot increase in expectation until settling in $\bar{G}_F$ (in case 1 by the definition of x^{-i}, while in cases 2 and 3 by part 2 of corollary 5 and part 2 of lemma 8); and (2) the segment F can only belong to case 2 (due to the exits in case 3), offering player i a reward of at most $\bar{v}^i_F \leq w^i_F$. Combining these two observations, it follows easily that player i's reward is at most w^i_s in expectation, proving (14).

$\star$ For an illustration of the above lemma, we refer to the games in examples 1 and 2. Indeed, the product-game in example 1 (which is now game G with minmax-levels v) leads to the simple product-game in example 2 (which is now game $\widetilde{G}$ with minmax-levels w). Just as in the proof of the above lemma, we can construct a stationary strategy y^1 for player 1 (y^1 being x^{-2} for players $-i$ with $i = 2$) which guarantees in G that player 2's reward is not more than w^2_s for all initial states $s \in S$. Recall from (12) and (11) that

$$\bar{w}^2_{F_{(I,I)}} = -1,\ \bar{w}^2_{F_{(I,II)}} = -1,\ \bar{w}^2_{F_{(II,I)}} = 0,\ \bar{w}^2_{F_{(II,II),1}} = -1,\ \bar{w}^2_{F_{(II,II),2}} = 0.$$

and from (10) and (9) that player 2's minmax-levels within the restricted games are

$$\bar{v}^2_{F_{(I,I)}} = -1,\ \bar{v}^2_{F_{(I,II)}} = 0,\ \bar{v}^2_{F_{(II,I)}} = 0,\ \bar{v}^2_{F_{(II,II),1}} = -1,\ \bar{v}^2_{F_{(II,II),2}} = 0.$$

Following the proof, as the segments $F_{(I,I)}$, $F_{(II,I)}$ and $F_{(II,II),1}$ and $F_{(II,II),2}$ all belong to case 2 (i.e. $\bar{v}^2_F \leq w^2_F$ when F equals any of these four segments F), the strategy y^1 has to guarantee in the corresponding restricted games that player 2's reward in G is not more than $\bar{v}^2_F$. Also, y^1 has to leave $F_{(I,II)}$, belonging to case 3 (i.e. $\bar{v}^2_{F_{(I,II)}} > w^2_{F_{(I,II)}}$). It is easy to see that the pure stationary strategy y^1 which plays action 1 in states $(1,1)$ and $(1,2)$, while action 2 in state $(1,3)$ satisfies all these requirements. $\star$

The previous lemma (and its proof) has useful consequences.

Corollary 10 *The results of lemmas 3 and 4, corollaries 5 and 6, and lemma 7 are also valid for any general product-game G. Lemma 8 extends as well if one interprets z^i_F as the minmax-level $\bar{v}^i_F$ of player i in the restricted game $\bar{G}_F$.*

Also, the infimum in expression (6) of the minmax-levels is attained at stationary strategies, for all product-games. This is stated next.

Corollary 11 *(of the proof of lemma 9) Take a product-game G and an arbitrary player i. Then, players $-i$ have a joint stationary strategy x^{-i} which guarantees that player i's reward from any initial state $s \in S$ is at most his minmax-level v_s^i, i.e. for all strategies π^i for player i we have*

$$\gamma_s^i(\pi^i, x^{-i}) \leq v_s^i.$$

With the help of this corollary, we are now ready to prove Main Theorem 2, which claimed that, in every two-player zero-sum product-game, both players have a stationary 0-optimal strategy.

Proof of Main Theorem 2. Take an arbitrary two-player zero-sum product-game, and take player $i = 1$. By corollary 11, there exists a stationary strategy x^{-1} for player 2 (as players -1 is simply player 2) which guarantees that player 1's reward is not more than v_s^1 for any initial state $s \in S$. Hence, x^{-1} is 0-optimal for player 2. One finds similarly a stationary 0-optimal strategy for player 1, which completes the proof.

$\star$ Thus, in our illustrative game in example 1, the pure stationary strategy y^1 for player 1 which plays action 1 in states $(1,1)$ and $(1,2)$, while action 2 in state $(1,3)$ is 0-optimal. $\star$

4.4 The construction of 0-equilibria in product-games

In this section, we will prove Main Theorem 1, which claimed that, in any product-game G, there exists a 0-equilibrium. Given the results above, the proof resembles the proof given in Flesch et al. [2008] for the aperiodic case. In our case, the proof is more complicated, as the sets E_k may split up into several segments, in contrast with the aperiodic case.

Proof of Main Theorem 1. For the general idea of the construction, we refer to section 3.

By corollary 6 and lemma 1, we know that, on each segment F, the minmax-values in G and in $\bar{G}_F$ are both a constant v_F and $\bar{v}_F$, respectively. Recall, from section 2, that $\mathcal{F}^*$ denotes the set of segments F such that $\bar{G}_F$ is satisfactory to all players (i.e. $\bar{v}_F^i \geq v_F^i$ for all players i), while $\mathcal{F}[i]$ denotes the set of segments F such that $\bar{G}_F$ is unsatisfactory to player i but $\bar{G}_F$ is satisfactory to all players $j \in \{1, \ldots, i-1\}$ (i.e. $\bar{v}_F^i < v_F^i$ and $\bar{v}_F^j \geq v_F^j$ for all $j \in \{1, \ldots, i-1\}$).

In every restricted game $\bar{G}_F$ with $F \in \mathcal{F}^*$, take a 0-equilibrium σ_F as in lemma 2. Let z^i_F denote the corresponding reward for any player i, which is independent of the initial state on F. Then, as $F \in \mathcal{F}^*$, we have for all players i that

$$z^i_F \geq \bar{v}^i_F \geq v^i_F. \tag{15}$$

In every restricted game $\bar{G}_F$ with $F \in \mathcal{F}[i]$, take a state s^i_F and an "exit" action a^i_F for player i in state s^i_F as in part 1 of lemma 8 (with corollary 10).

The proof of Main Theorem 1 consists of the following steps. In step 1, we construct a joint stationary strategy x^* which will be used to reach segments $F \in \mathcal{F}^*$ (this step is only needed if there are states outside the segments). Then, in step 2 we focus on states within the segments, and "extend" x^* to the joint strategy π according to which the players also receive rewards z_F in every segment $F \in \mathcal{F}^*$. Finally, in step 3, we will complete the proof by showing that π supplemented with appropriately chosen joint stationary strategies y^{-i}, for all i, forms a 0-equilibrium.

Step 1: *The construction of the joint stationary strategy x^* satisfying a number of properties*. First, for every player i and state $s \in S$ for which s^i is of type 2, fix an action $\bar{d}^i_s \in \bar{A}^i_s$. The only requirement we have is that in each segment $F \in \mathcal{F}[i]$, player i will eventually reach his "exit" state s^i_F through these actions (given the other players do not leave F).

Further, consider a player i and a state s^i of type 1, and some cyclic set $T^j_{k^j}(m^j)$ of a maximal communicating set $E^j_{k^j}$, for every player $j \neq i$. We will argue that there exists an action $a^i_{s^i}[k^{-i}, m^{-i}]$, or $a^i_{s^i}$ for short, for player i in state s^i with the following property: in every state $s = (s^i, s^{-i})$, with $s^{-i} \in \times_{j \neq i} T^j_{k^j}(m^j)$, action $a^i_{s^i}$ is providing the best transitions with respect to player i's minmax-level, against the joint action $\bar{d}^{-i}_s$, i.e. for all $b^i_s \in A^i_s$ we have

$$\sum_{t \in S} p_{s,(a^i_{s^i}, \bar{d}^{-i}_s)}(t)\, v^i_t \geq \sum_{t \in S} p_{s,(b^i_s, \bar{d}^{-i}_s)}(t)\, v^i_t. \tag{16}$$

One can verify this claim as follows. Since every player $j \neq i$ plays $\bar{d}^j_s \in \bar{A}^j_s$, player j moves next to cyclic set $T^j_{k^j}(m^j+1)$. Thus, players $-i$ move to $\times_{j \neq i} T^j_{k^j}(m^j+1)$. In view of part 2 of lemma 7, player i's minmax-level is some constant $v^i[t^i]$ on the set of states $t = (t^i, t^{-i})$ with $t^{-i} \in \times_{j \neq i} T^j_{k^j}(m^j+1)$. Hence, the left-hand-side of (16) can be written as

$$\sum_{t \in S} p_{s,(a^i_{s^i}, \bar{d}^{-i}_s)}(t)\, v^i_t = \sum_{t^i \in S^i} p^i_{s^i a^i_{s^i}}(t^i)\, v^i[t^i]. \tag{17}$$

Consequently, any action $a^i_{s^i}$ for player i maximizing the right-hand-side of (17) will prove our claim in (16).

Now, we define a subset $X^{*i}_s \subset X^i_s$ of mixed actions for every player i in every state $s = (s^1, \ldots, s^n) \in S$ as follows:

Case (1): s lies in segment F with $F \in \mathcal{F}^*$: In this case, we let $X^{*i}_s = \{\bar{d}^i_s\}$ for all players i.

Case (2): s lies in segment F with $F \in \mathcal{F}^*[i]$: In this case, for players $j \neq i$ we let $X^{*j}_s = \{\bar{d}^j_s\}$. As for player i, we let $X^{*i}_s = \{\bar{d}^i_s\}$ if $s^i \neq s^i_F$ while $X^{*i}_s = \{a^i_F\}$ if $s^i = s^i_F$ (recall that a^i_F is player i's "exit" action in state s^i_F; cf. the beginning of the proof).

Case (3): s^i is of type 1, but s^j is of type 2 for every player $j \neq i$: In this case, for every player j, state s^j belongs to some cyclic set $T^j_{k^j}(m^j)$ of a maximal communicating set $E^j_{k^j}$. For every player $j \neq i$, we let $X^{*j}_s = \{\bar{d}^j_s\}$. As for player i, we let $X^{*i}_s = \{a^i_{s^i}[k^{-i}, m^{-i}]\}$ (as discussed in (16)).

Case (4): s^i is of type 1 for at least two players i: In this case, for every player i for whom s^i is of type 1, we let $X^{*i}_s = X^i_s$. For every player j for whom s^j is of type 2, we let $X^{*j}_s = \{\bar{d}^j_s\}$.

Note that $X^{*i}_s = X^i_s$ or X^{*i}_s is a singleton consisting of one action, for any player i in any state $s \in S$. In fact player i is unrestricted exactly when he and at least one more player are in a state of type 1. Moreover, the ergodic sets for all $x \in X^*$ are all within the segments $F \in \mathcal{F}^*$ (cf. case 1), due to player i's "exit" action a^i_F in each segment $F \in \mathcal{F}^*[i]$ (cf. case 2). Observe also that the ergodic sets are the same for all $x \in X^*$.

Let G^* denote the stochastic game (not necessarily a product-game) which is derived from G by replacing every player i's payoffs by z^i_F in every restricted game $\bar{G}_F$ with $F \in \mathcal{F}^*$, and subsequently by restricting each player i in each state $s \in S$ to the space X^{*i}_s of mixed actions. Recall that z^i_F denotes player i's reward in restricted game $\bar{G}_F$ with respect to the 0-equilibrium σ_F (cf. beginning of the proof).

As the ergodic sets are the same for all $x \in X^*$, there exists a stationary 0-equilibrium $x^* \in X^*$ for the game G^* (cf. Flesch et al. [2008] for a detailed proof). Obviously, x^* is also a joint stationary strategy in the original game G, but not necessarily a 0-equilibrium.

We wish to point out three properties of x^*.

Property (1): If player i is unrestricted in state s in the game G^*, i.e. $X^{*i}_s = X^i_s$, then player i cannot go to better states regarding his reward in G^* by unilaterally deviating from x_s^{*i}, i.e. for every action $b_s^i \in A_s^i$ we have

$$\sum_{t \in S} p_{s,(b_s^i, x_s^{*-i})}(t)\, \gamma_t^{*i}(x^*) \leq \sum_{t \in S} p_{s x_s^*}(t)\, \gamma_t^{*i}(x^*),$$

where γ^{*i} denotes the reward to player i in the game G^*. This property follows from the fact that x^* is a 0-equilibrium in G^*.

Property (2): If player i is restricted in state s in the game G^*, i.e. when X_s^{*i} is a singleton, then player i cannot improve on his expected minmax-level in G by unilaterally deviating from x_s^{*i}, i.e. for every action $b_s^i \in A_s^i$ we have

$$\sum_{t \in S} p_{s,(b_s^i, x_s^{*-i})}(t)\, v_t^i \leq \sum_{t \in S} p_{s x_s^*}(t)\, v_t^i. \tag{18}$$

We remark that this property, in view of (7), implies

$$v_s^i \leq \sum_{t \in S} p_{s x_s^*}(t)\, v_t^i. \tag{19}$$

To show (18), we distinguish three possibilities: (i) Whenever player i is restricted to action $\bar{d}_s^i$, then (18) follows from part 1 of lemma 4 (with corollary 10). (ii) Whenever player i is restricted to action $a^i_{s^i}[k^{-i}, m^{-i}]$ in case 3, then (18) follows from (16). (iii) Whenever player i is restricted to his "exit" action a_F^i in case 2, then by part 2 of corollary 5 and part 1 of lemma 8 (with corollary 10) we have respectively

$$\sum_{t \in S} p_{s,(b_s^i, x_s^{*-i})}(t)\, v_t^i \leq v_F^i; \quad v_F^i \leq \sum_{t \in S} p_{s x_s^*}(t)\, v_t^i,$$

implying (18).

Property (3): x^* yields rewards in G^* that are at least the minmax-levels in G, i.e. $\gamma_s^{*i}(x^*) \geq v_s^i$ for all players i and for all initial states $s \in S$. For a proof, we refer to Flesch et al. [2008].

Step 2. The construction of the joint strategy π for the original game G. Given x^* from step 1, the definition of π is easy. Let π be the joint strategy which prescribes to play as follows:

Case (1): when play enters a segment F with $F \in \mathcal{F}^*$. In this case, the players switch to the joint strategy σ_F (cf. begin of the proof).

Case (2): when play enters a segment F with $F \in \mathcal{F}^*[i]$. In this case, players $-i$ switch to a joint stationary strategy as in lemma 1, while player i follows x^*, i.e. plays the mixed action x_s^{*i} in state $s \in F$. This means that player i exits F through moving to state s_F^i and then playing action a_F^i (repeatedly until exit ocurs).

Case (3): in any state $s \in S$ outside the segments. In this case, each player i follows x^*, i.e. plays the mixed action x_s^{*i}.

We remark the following properties:

(A) Play according to π keeps play in every segment $F \in \mathcal{F}^*$ (cf. case 1) and leaves every segment $F \in \mathcal{F}^*[i]$ (cf. case 2) with the guidance of x^{*i}. Consequently, π will eventually settle in a segment $F \in \mathcal{F}^*$, with probability 1. Moreover, in a segment $F \in \mathcal{F}^*$, by switching to σ_F, each player i receives reward z_F^i. So in some sense, x^* is used to reach a segment $F \in \mathcal{F}^*$, and then σ_F is used to induce rewards z_F to the players.

(B) Whenever a player was unrestricted in some state s in the game G^*, which only happened outside the segments, then π prescribes x_s^* (cf. case 3). Whenever a player i was restricted in state s to action $\bar{d}_s^i$ in the game G^*, then π^i prescribes a mixed action in $\bar{X}_s^i$. Whenever a player i was restricted in state s to action $a_{s^i}^i[k^{-i}, m^{-i}]$ in the game G^*, then π^{-i} prescribes for players $-i$ the joint action $\bar{d}_s^{-i}$. Finally, whenever a player i was restricted in state s to an "exit" action (cf. action a_F^i in case 2) in the game G^*, then π^{-i} prescribes for players $-i$ a joint mixed action in $\bar{X}_s^{-i}$.

(C) Whenever player i leaves segment $F \in \mathcal{F}^*[i]$ from some state $s = (s_F^i, s^{-i}) \in F$ (cf. case 2), then the probability that the next segment play visits is some F', is independent of s^{-i}. This is due to the following reasons. Suppose F lies within some $E_k = \times_{j=1}^n E_{k^j}^j$, and that s belongs to a joint cyclic set $T_F(m)$ (cf. part 1 of lemma 8). Suppose also that player i exits F and moves to a state t^i outside $E_{k^i}^i$. At the same time, independently of the particular choice of s^{-i} in $T_F^{-i}(m)$, players $-i$ move to the next joint cyclic set $T_F^{-i}(m+1)$. Consequently, if t^i is of type 2, then play entered a new segment, and this segment is independent of s^{-i}. On the other hand, if t^i is of type 1, then case 3 of step 1 takes care that player i's behavior is independent of players $-i$'s state within $T_F^{-i}(m+1)$, and this remains so until a new segment is entered.

(D) By the construction of π, the joint strategies π and x^* induce the same probability that play eventually settles in any segment F, from any initial state $s \in S$. Thus, by property 3 in step 1, we obtain

$$\gamma_s^i(\pi) = \gamma_s^{*i}(x^*) \geq v_s^i \tag{20}$$

for all initial states $s \in S$ and for all players i, which means that π induces individually rational rewards in G.

Step 3. *Proving that π supplemented with the joint stationary strategies y^{-i}, for all players i, is a 0-equilibrium.* For any player i, in view of corollary 11, we may take a joint stationary strategy y^{-i} for players $-i$ such that for all initial states $s \in S$ and for all strategies τ^i for player i we have

$$\gamma_s^i(\tau^i, y^{-i}) \leq v_s^i.$$

Let the joint strategy η be defined as in section 3. Note that the expected rewards are equal with respect to η and with respect to π, hence by (20)

$$\gamma_s^i(\eta) = \gamma_s^i(\pi) = \gamma_s^{*i}(x^*) \geq v_s^i$$

for all initial states $s \in S$ and for all players i. Notice also that if h denotes a history and $s \in S$ a state such that, with a positive probability, h can occur and s can be the present state after h with respect to η (or equivalently with respect to π), then

$$\gamma_s^i(\eta[h]) = \gamma_s^i(\pi[h]) = \gamma_s^i(\pi) = \gamma_s^{*i}(x^*) \geq v_s^i, \tag{21}$$

where for the second equality we used property C from step 2 and that for σ_k the "continuation rewards" remain the same due to lemma 2. We may thus conclude that η yields individually rational rewards in G, for all players i, and for such histories h and states $s \in S$.

It remains to show that η is a 0-equilibrium in G.

Deviations inside the support of η : Here, we only consider deviations by playing actions with a positive probability according to η. We show that such deviations by a player cannot improve his expected reward. Indeed: (i) within a segment $F \in \mathcal{F}^*$ (cf. case 1 in step 2), the players play the 0-equilibrium σ_F in $\bar{G}_F$; (ii) within a segment $F \in \mathcal{F}^*[i]$ (cf. case 2 in step 2), such a deviation by players $-i$ would not change the probability of eventually moving to another segment F' (cf. property C in step 2); (iii) within a segment $F \in \mathcal{F}^*[i]$ (cf. case 2 in step 2), player i has an incentive to "exit", since within $\bar{G}_F$ he can get at most $\bar{v}_F^i$, while $\bar{v}_F^i < v_F^i$; (iv) in states $s \in S$ outside the segments (cf. case 3 in step 2), whenever for some player i, the strategy π^i (or equivalently, x^{*i}) uses at least two actions, then player i was not restricted in the game G^* (i.e. $X_s^{*i} = X_s^i$) and hence, property 1 in step 1 together with the equalities in (21) guarantee that player i cannot go to better states regarding his reward.

Deviations outside the support of η: Consider now a deviation when, for the first time, say after history h in state s, while the players should play a joint mixed action x'_s according to η, some player i deviates and plays an action b_s^i which has probability zero according to η^i, i.e. $x_s'^i(b_s^i) = 0$. This deviation is immediately noticed by players $-i$ and, according to η, they switch to the joint stationary strategy y^{-i} from the next state, say state t. Consequently, player i's reward will be at most v_t^i in expectation. Obviously, without deviation player i would receive reward $\gamma_s^i(\eta[h]) = \gamma_s^{*i}(x^*)$, in view of (21). Now, observe the following.

(A) Suppose player i is unrestricted in state s in the game G^*, i.e. $X_s^{*i} = X_s^i$. Then, $x_s' = x_s^*$ (cf. property B in step 2), and player i's expected reward after this deviation is at most

$$\sum_{t\in S} p_{s,(b_s^i,x_s^{*-i})}(t)\, v_t^i \le \sum_{t\in S} p_{s,(b_s^i,x_s^{*-i})}(t)\, \gamma_t^{*i}(x^*) \le \sum_{t\in S} p_{sx_s^*}(t)\, \gamma_t^{*i}(x^*) = \gamma_s^{*i}(x^*) = \gamma_s^i(\eta[h]),$$

where the inequalities follow from properties 3 and 1 in step 1, respectively, while the equalities from (4) and (21). Hence, the deviation is not profitable.

(B) Suppose player i is restricted in state s in the game G^*. First we show

$$\sum_{t\in S} p_{s,(b_s^i,x_s'^{-i})}(t)\, v_t^i \le \sum_{t\in S} p_{sx_s'}(t)\, v_t^i. \tag{22}$$

We distinguish three possibilities: (i) Whenever player i is restricted to action $\bar{d}_s^i$, then $x_s'^i \in \bar{X}_s^i$ by property B in step 2, hence (22) follows from part 1 of lemma 4 (with corollary 10). (ii) Whenever player i is restricted to action $a_{s^i}^i[k^{-i}, m^{-i}]$ in case 3 of step 1, then $x_s'^{-i} = \bar{d}_s^{-i}$ by property B in step 2, hence (22) follows from inequality (16). (iii) Whenever player i plays his "exit" action a_F^i in case 2 of step 1, then $x_s'^{-i} \in \bar{X}_s^{-i}$ by property B in step 2, hence by part 2 of corollary 5 and part 1 of lemma 8 (with corollary 10) we have respectively

$$\sum_{t\in S} p_{s,(b_s^i,x_s'^{-i})}(t)\, v_t^i \le v_F^i; \quad v_F^i \le \sum_{t\in S} p_{sx_s'}(t)\, v_t^i,$$

implying (22).

Hence, by (21), we obtain that player i's expected reward after this deviation is at most

$$\sum_{t\in S} p_{s,(b_s^i,x_s'^{-i})}(t)\, v_t^i \le \sum_{t\in S} p_{sx_s'}(t)\, v_t^i \le \gamma_s^i(\eta[h]),$$

which means that the deviation is not profitable again.

In conclusion, no deviation is profitable, and η is a 0-equilibrium in G. This completes the proof of Main Theorem 1. ■

$\star$ Finally, let us revisit the game in example 1. As we know, the minmax-levels of this game coincide with the minmax-levels of the game in example 2, hence by (11) and (12), we have that

$$v^1_{F_{(I,I)}} = 1,\ v^1_{F_{(I,II)}} = 1,\ v^1_{F_{(II,I)}} = 0,\ v^1_{F_{(II,II),1}} = 1,\ v^1_{F_{(II,II),2}} = 0$$

and $v^2_F = -v^1_F$ for all segments F. Recall from (9) and (10) that the players minmax-levels within the restricted games are

$$\bar{v}^1_{F_{(I,I)}} = 1,\ \bar{v}^1_{F_{(I,II)}} = 0,\ \bar{v}^1_{F_{(II,I)}} = 0,\ \bar{v}^1_{F_{(II,II),1}} = 1,\ \bar{v}^1_{F_{(II,II),2}} = 0$$

and $\bar{v}^2_F = -\bar{v}^1_F$ for all segments F. Hence, regarding which segments are satisfactory, we obtain that $F_{(I,I)}$, $F_{(II,I)}$ and $F_{(II,II),1}$ and $F_{(II,II),2}$ all belong to $\mathcal{F}^*$ (i.e. $\bar{v}_F \leq v_F$ when F equals any of these four segments), whereas $F_{(I,II)}$ belongs to $\mathcal{F}^*[1]$.

Consider the stationary strategy x^1 for player 1 which plays action 1 in states $(1,1)$ and $(1,2)$, while action 2 in state $(1,3)$, and the stationary strategy x^2 for player 2 which plays action 1 in all states. This pair (x^1, x^2) actually could play the role of π in this example. Indeed, in each restricted game $\bar{G}_F$ corresponding to segments $F \in \mathcal{F}^*$, the pair (x^1, x^2) lets the players play a 0-equilibrium, while x^1 leaves segment $F_{(I,II)}$. Notice that no threat strategies are needed here, so (x^1, x^2) is a 0-equilibrium. $\star$

5 Proof of Lemma 1

In this section, we prove lemma 1. Consider a restricted game $\bar{G}_F$ and a player i. Let $\alpha^i := \min_{t\in F} \bar{v}^i_{F,t}$. The idea of the proof is to find a set $W \subset \{t \in F | \bar{v}^i_{F,t} = \alpha^i\}$ and a joint stationary strategy x^{-i} such that, irrespective of the strategy of player i, the following hold: (1) once play reaches W, it remains in W forever, and player i's reward within W is at most α^i, (2) play reaches W with probability 1. Clearly, these properties will then imply that the minmax-level $\bar{v}^i_F$ of player i in $\bar{G}_F$ equals the constant α^i, and that x^{-i} satisfies the second part of the lemma.

Step 1: finding the set W and the joint stationary strategy x^{-i}.

As is mentioned in the introduction, by applying Thuijsman & Vrieze [1991] together with Neyman [2003] for the game $\bar{G}_F$, there exists a state $s' \in \{t \in F | \bar{v}^i_{F,t} = \alpha^i\}$ for which players $-i$ have a joint stationary strategy y^{-i} such that for all strategies π^i for player i in $\bar{G}_F$ we have

$$\bar{\gamma}^i_{s'}(\pi^i, y^{-i}) \leq \bar{v}^i_{F,s'} = \alpha^i.$$

Let Z denote the set of all those states $s \in \{t \in F | \bar{v}^i_{F,t} = \alpha^i\}$ for which this y^{-i} satisfies for all strategies π^i for player i in $\bar{G}_F$ that

$$\bar{\gamma}^i_s(\pi^i, y^{-i}) \leq \alpha^i. \tag{23}$$

In particular, $s' \in Z$.

Let y^i be a completely mixed stationary strategy in $\bar{G}_F$ for player i. By the definitions of Z and y^{-i}, the set Z is closed for the joint stationary strategy (y^i, y^{-i}) (i.e. play does not leave Z). Hence, there must exist an ergodic set $W \subset Z$ for (y^i, y^{-i}). Define a joint stationary strategy x^{-i} for players $-i$ in $\bar{G}_F$ as follows: let $x^{-i}_t = y^{-i}_t$ for all $t \in W$ and let x^{-i}_t be an arbitrary completely mixed action on $\bar{A}^{-i}_t$ for all $t \in (F - W)$.

Step 2: proving properties (1) and (2) for W and x^{-i}.

Regarding property (1): Consider (π^i, x^{-i}) with an arbitrary strategy π^i for player i in $\bar{G}_F$. As W is closed with respect to (y^i, y^{-i}) and x^{-i} equals y^{-i} on W, once play with respect to (π^i, x^{-i}) reaches W, it will never leave it. Moreover, in view of (23), player i's reward will be at most α^i within W. Hence property (1).

Regarding property (2): We make the following observations.

(a) For any joint completely mixed stationary strategy z^{-i} of player $-i$, there exists a $\mu > 0$ and a stage M satisfying the following property: if players $-i$ use z^{-i} and play starts in any cyclic set $T_F(l)$ of segment F, then at all stages $m \geq M$, players $-i$ will be in any state $u^{-i} \in T^{-i}_F(l + m - 1)$ with probability at least μ. This follows from the discussion in section 1 on Markov chains.

(b) For any u^i belonging to a cyclic set $T^i_F(l)$, there exists a joint state $u^{-i} \in T^{-i}_F(l)$ such that $(u^i, u^{-i}) \in W$. We argue as follows. Consider (y^i, y^{-i}). Due to the choice of y^i, if play starts in any cyclic set $T_F(l')$, then at sufficiently large stages m, player i will be in every state in $T^i_F(l' + m - 1)$ with a positive probability. Since W is closed for (y^i, y^{-i}), property (b) follows.

It is clear, based on the choice of x^{-i} outside W, that properties (a) and (b) imply property (2). Hence, the proof of lemma 1 is complete.

6 Proof of Lemma 3

In this section, we prove lemma 3. To this end, however, we first need to show an auxiliary result.

Consider an arbitrary an arbitrary (non-empty) collection of players $N' \subset N$. Then, by a history for players N' we mean the sequence of past joint states of players N' and the past joint actions played by players N'. Formally, if the history is

$h = (u_1, a_1, \ldots, u_m, a_m)$, where u_l and a_l denote the state and joint action for any stage $l = 1, \ldots, m$, then the history of players N' is simply $h^{N'} = (u_1^{N'}, a_1^{N'}, \ldots, u_m^{N'}, a_m^{N'})$.

Take a joint strategy π with some initial state $s \in S$. Then, π from state s generates a probability distribution on all possible histories for players N'. The following lemma claims that, given players N' start in state $s^{N'}$, they can generate this probability distribution on their histories, even against other strategies of players outside N'. It is hardly surprising as, in a product-game, each player controls play on his own coordinate.

Lemma 12 *Let π be an arbitrary joint strategy, and $s \in S$ be some initial state. Consider an arbitrary (non-empty) collection of players $N' \subset N$. Then, there exists a joint strategy $\sigma^{N'} = (\sigma^i)_{i \in N'}$ for players in N' such that for any state $t^{N-N'} \in S^{N-N'}$ and any joint strategy $\sigma^{N-N'} = (\sigma^i)_{i \in N-N'}$ for players outside N', we have*

$$\mathbb{P}_{(s^{N'}, t^{N-N'}), \sigma}(h_m^{N'}) = \mathbb{P}_{s\pi}(h_m^{N'}) \tag{24}$$

for any joint history $h_m^{N'}$ of players in N', up to any stage $m \in \mathbb{N}$. Moreover, the joint mixed actions prescribed by $\sigma^{N'}$ in any state $u \in S$ at any stage m only depend on $u^{N'}$, on the history of players N', and on the joint mixed actions prescribed by π at stages $1, \ldots, m-1$ and by $\pi^{N'}$ at stage m.

Proof. The construction of σ^i for each player $i \in N'$ is simple. Consider some present state $u_m \in S$ at stage m and some history $h_m = (s_1, a_1, \ldots, s_{m-1}, a_{m-1})$. If the probability that $h_m^{N'}$ occurs and $u_m^{N'}$ becomes the present state for players N' is zero, with respect to π and initial state s, then the mixed action $\sigma_{u_m}^i(h_m)$ is arbitrary. Otherwise, let

$$\sigma_{u_m}^i(h_m)(a_m^i) := \sum_{\widetilde{h}_m, \widetilde{u}_m} \mathbb{P}_{s\pi}(\widetilde{h}_m, \widetilde{u}_m | \, h_m^{N'}, u^{N'}) \cdot \pi_{\widetilde{u}_m}^i(\widetilde{h}_m)(a_m^i) = \mathbb{P}_{s\pi}(a_m^i | \, h_m^{N'}, u_m^{N'}); \tag{25}$$

that is, player i should play action a_m^i with the same probability as according to the joint strategy π conditionally on joint history $h_m^{N'}$ and present state $u_m^{N'}$ for players N'.

Given $\sigma^{N'} = (\sigma^i)_{i \in N'}$, one can show (24) by using induction on m. ■

Proof of Lemma 3. We only show part 1 of the lemma; part 2 can be proven similarly. Recall that $\lambda_{k^i}^i$ denotes the period of the maximal communicating set $E_{k^i}^i$. Let $T_{k^i}^i(1), \ldots, T_{k^i}^i(\lambda_{k^i}^i)$ denote the cyclic sets of $E_{k^i}^i$, and suppose for simplicity that $s^i \in T_{k^i}^i(1)$. Take an action $a_{s^i}^i \in A_{s^i}^i$ and a state $t^i \in T_{k^i}^i(2)$. We need to prove that

$$\sum_{u^i \in S^i} p_{s^i a_{s^i}^i}^i(u^i) \, v_{(u^i, s^{-i})}^i \leq v_{(t^i, s^{-i})}^i. \tag{26}$$

The idea of the proof: Let $\varepsilon > 0$. We will compare two specific games Ω and $\widetilde{\Omega}$. The game Ω is the original game G starting in state (t^i, s^{-i}), whereas $\widetilde{\Omega}$ is the game G which starts in initial state (u^i, s^{-i}) with probability $p^i_{s^i a^i_{s^i}}(u^i)$. We will define two joint strategies π for Ω and $\widetilde{\pi}$ for $\widetilde{\Omega}$ in such a way that the following properties hold:

Property (A) for π in Ω: $\gamma^i_{(t^i,s^{-i})}(\pi) \leq v^i_{(t^i,s^{-i})} + \varepsilon$.

Property (B) for $\widetilde{\pi}$ in $\widetilde{\Omega}$: $\gamma^i_{(u^i,s^{-i})}(\widetilde{\pi}) \geq v^i_{(u^i,s^{-i})} - \varepsilon$, for every $u^i \in S^i$.

Property (C) for the rewards: π and $\widetilde{\pi}$ yield the same expected rewards in respectively Ω and $\widetilde{\Omega}$, i.e.

$$\gamma^i_{(t^i,s^{-i})}(\pi) = \sum_{u^i \in S^i} p^i_{s^i a^i_{s^i}}(u^i)\, \gamma^i_{(u^i,s^{-i})}(\widetilde{\pi}).$$

It will then follow from properties (A), (B) and (C) that

$$\begin{aligned}
\sum_{u^i \in S^i} p^i_{s^i a^i_{s^i}}(u^i)\; v^i_{(u^i,s^{-i})} &\leq \sum_{u^i \in S^i} p^i_{s^i a^i_{s^i}}(u^i)\,(\gamma^i_{(u^i,s^{-i})}(\widetilde{\pi}) + \varepsilon) \\
&= \sum_{u^i \in S^i} p^i_{s^i a^i_{s^i}}(u^i)\, \gamma^i_{(u^i,s^{-i})}(\widetilde{\pi}) + \varepsilon \\
&= \gamma^i_{(t^i,s^{-i})}(\pi) + \varepsilon \\
&\leq v^i_{(t^i,s^{-i})} + 2\varepsilon.
\end{aligned}$$

As $\varepsilon > 0$ was arbitrary, the proof will then be complete.

Construction of π^{-i} in Ω: Let π^{-i} be a joint strategy for players $-i$ in Ω such that $\gamma^i_{(t^i,s^{-i})}(\pi^{-i}, \sigma^i) \leq v^i_{(t^i,s^{-i})} + \varepsilon$ for any strategy σ^i for player i. Such a joint strategy π^{-i} exists by the definition of the minmax-level $v^i_{(t^i,s^{-i})}$. Thus, irrespectively of the choice of π^i, property A will be satisfied.

Construction of π^i for Ω, and $\widetilde{\pi}^{-i}$ and $\widetilde{\pi}^i$ for $\widetilde{\Omega}$: These strategies are defined step by step. Roughly speaking:

(i) $\widetilde{\pi}^{-i}$ for $\widetilde{\Omega}$ is obtained by copying π^{-i} in the sense of lemma 12 (players $-i$ start in s^{-i} in both games Ω and $\widetilde{\Omega}$).

(ii) $\widetilde{\pi}^i$ for $\widetilde{\Omega}$ is then obtained by taking a strategy which defends player i's minmax-level v, up to ε, against $\widetilde{\pi}^{-i}$ (cf. the discussion below (6)). Note that $\widetilde{\pi}$ will satisfy property B.

(iii) π^i for Ω is obtained by copying $\widetilde{\pi}^i$ in the sense of lemma 12. Before starting copying, though, π^i is in an initial phase in which player i moves from state t^i to s^i and subsequently plays action $a^i_{s^i}$, so that player i is in state u^i with probability $p^i_{s^i a^i_{s^i}}(u^i)$, just as in game $\widetilde{\Omega}$. We describe now this initial phase more precisely. The strategy π^i prescribes to first move from player i's initial state t^i to state s^i, within the maximal communicating set $E^i_{k^i}$. This can be done by always choosing, according to the uniform distribution, an action from the set $\bar{A}^i_{w^i}$ in every state $w^i \in E^i_{k^i}$. Note that, as the game Ω starts in $t^i \in T^i_{k^i}(2)$, player i is then in cyclic set $T^i_{k^i}(1)$ at stages of the form $l \cdot \lambda^i_{k^i}$, for all $l \in \mathbb{N}$. So, s^i can be reached at stages $l \cdot \lambda^i_{k^i}$, for large $l \in \mathbb{N}$. Recall that λ denotes the period of the whole game G (so λ is also the period of Ω and $\widetilde{\Omega}$). Thus, λ is a multiple of $\lambda^i_{k^i}$. Hence, s^i can also be reached at stages $l \cdot \lambda$, for large $l \in \mathbb{N}$. Now, let π^i prescribe to move to s^i this way, and when player i is in s^i at a stage $l \cdot \lambda$, then to play action $a^i_{s^i}$. (It will be important for property C that player i uses a stage of the form $l \cdot \lambda$, and not only $l \cdot \lambda^i_{k^i}$. See the example after the proof.)

Now, we define these strategies more precisely. Note that, as soon as $\widetilde{\pi}^{-i}$ is defined up to some stage m, so is $\widetilde{\pi}^i$.

At stage 1: The history is thus empty. Let for all $u^i \in S$

$$\widetilde{\pi}^{-i}_{(u^i,s^{-i})}(\emptyset) := \pi^{-i}_{(t^i,s^{-i})}(\emptyset),$$

which means that, in any initial state (u^i, s^{-i}) of $\widetilde{\Omega}$, the joint strategy $\widetilde{\pi}^{-i}$ prescribes the same joint mixed action as π^{-i} in the initial state (t^i, s^{-i}) of Ω. Note that in both cases, players $-i$ are in joint state s^{-i}. Moreover, lemma 12 would prescribe the same mixed action for $\widetilde{\pi}^{-i}$, when copying π^{-i}.

Given $\widetilde{\pi}^{-i}$ at stage 1, we know $\widetilde{\pi}^i$ at stage 1 as well.

At stage 1, π^i is still in the initial phase described in (iii). Hence, π^i is defined at stage 1.

At an arbitrary stage m: Given (π^i, π^{-i}) for stages up to $m-1$, and π^{-i} for stage m, the joint strategy $\widetilde{\pi}^{-i}$ copies π^{-i} at stage m in the sense of lemma 12.

Given $\widetilde{\pi}^{-i}$ up to stage m, we know $\widetilde{\pi}^i$ up to stage m as well.

As for π^i, there are three cases. (1) If player i has not reached s^i at a stage $l \cdot \lambda$, then he continues trying to get to s^i as prescribed in (iii) above. (2) If he has just arrived at s^i at stage $m = l \cdot \lambda$, then he plays action $a^i_{s^i}$, in accordance with (iii) above. (3) Suppose that player i has reached s^i at a stage $l \cdot \lambda$ and played action $a^i_{s^i}$. Then, player i was in state u^i at stage $l \cdot \lambda + 1$ with probability $p^i_{s^i a^i_{s^i}}(u^i)$, just as in the game

$\widetilde{\Omega}$. By regarding this as the initial state and stage, π^i copies $\widetilde{\pi}^i$ in the sense of lemma 12, based on $(\widetilde{\pi}^i, \widetilde{\pi}^{-i})$ up to stage m. (So, π^i for stage $m = l \cdot \lambda + m'$ is a copy of $\widetilde{\pi}^i$ at stage m'.)

The joint strategies satisfy properties A, B and C: Now we will verify properties A, B and C, which will complete the proof of the lemma. As we mentioned above, properties A and B are satisfied due to the definitions of π^{-i} and $\widetilde{\pi}^i$, respectively.

It remains to verify property C. Consider the joint strategies $\pi = (\pi^i, \pi^{-i})$ in Ω and $\widetilde{\pi} = (\widetilde{\pi}^i, \widetilde{\pi}^{-i})$ in $\widetilde{\Omega}$. As we know, with respect to any joint strategy and initial state, play eventually settles in a segment (or equivalently, in a restricted game), with probability 1. Since the game G (and therefore Ω and $\widetilde{\Omega}$ too) is simple, by definition, the payoffs are constant in each restricted game $\bar{G}_F$. Hence, in order to show C, it suffices to show that the probability that play settles in a segment F (or in the corresponding restricted game $\bar{G}_F$) is the same with respect to $\pi = (\pi^i, \pi^{-i})$ in Ω and $\widetilde{\pi} = (\widetilde{\pi}^i, \widetilde{\pi}^{-i})$ in $\widetilde{\Omega}$.

Take an arbitrary segment F within some $E_k = \times_{j=1}^n E^j_{k^j}$, where, as usual, $E^j_{k^j}$ denotes a maximal communicating set for player j. For any player j, state $u^j \in E^j_{k^j}$ and stage m^j, let $\omega^j(u^j, m^j)$ denote the event that player j settles in $E^j_{k^j}$ in state u^j at stage m^j. Since $\widetilde{\pi}^{-i}$ is a copy of π^{-i}, event $\omega^j(u^j, m^j)$, for any $j \neq i$ and any $u^j \in E^j_{k^j}$ and stage m^j, has the same probability with respect to $\pi = (\pi^i, \pi^{-i})$ in Ω and with respect to $\widetilde{\pi} = (\widetilde{\pi}^i, \widetilde{\pi}^{-i})$ in $\widetilde{\Omega}$. By construction, π^i is also a copy of $\widetilde{\pi}^i$, with a "delay" of $l \cdot \lambda$ stages. Therefore, event $\omega^i(u^i, m^i)$, for any $u^i \in E^i_{k^i}$ and stage m^i, has the same probability with respect to $\widetilde{\pi} = (\widetilde{\pi}^i, \widetilde{\pi}^{-i})$ in $\widetilde{\Omega}$ as event $\omega^i(u^i, m^i + l \cdot \lambda)$ with respect to $\pi = (\pi^i, \pi^{-i})$ in Ω. Since λ, the period of the whole game, is a multiple of the period $\lambda^i_{k^i}$, so is the "delay" $l \cdot \lambda$. Hence, the probability that play settles in F is the same with respect to $\pi = (\pi^i, \pi^{-i})$ in Ω and $\widetilde{\pi} = (\widetilde{\pi}^i, \widetilde{\pi}^{-i})$ in $\widetilde{\Omega}$. This completes the proof of lemma 3. ■

$\star$ As an illustration of the proof of lemma 3, consider the game with two players given in Figure 3.The underlying Markov transition structure for player 1 is as follows. Player 1 has 5 states, corresponding to the rows. He can move along the cycles on states $\{1, 2\}$ or on states $\{3, 4, 5\}$. Additionally, he can move from state 1 to state 3. As for player 2, he has 6 states, corresponding to the columns. Player 2 can move from state 1 to states 2 and 3, from state 2 to state 1, and from state 3 to states 1 and 4. Further, player 2 has a cycle on states $\{4, 5, 6\}$. As there are cycles of lengths 2 and 3, the game has period $\lambda = 6$. We want to focus on the transitions, so the payoffs are omitted.

Consider part 1 of lemma 3, and take player $i = 2$ and state $s = (1, 1)$. Note that

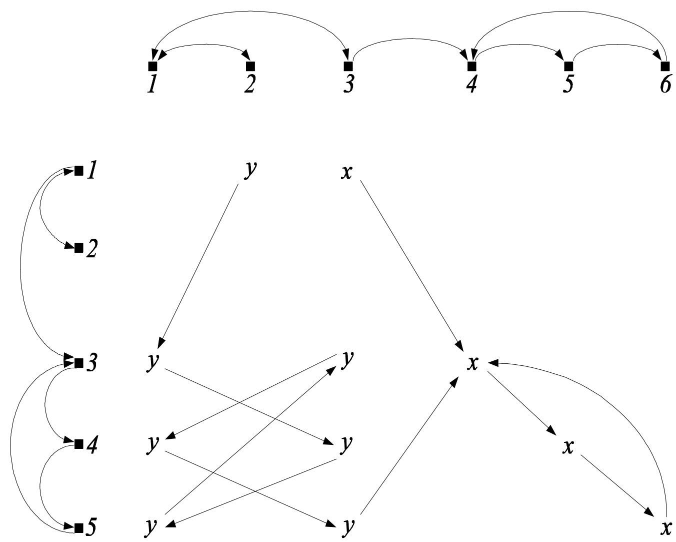

Figure 1: Figure 3: Illustration for lemma 3

$\{1,2,3\}$ is a maximal communicating set of player 2, with cyclic sets $\{1\}$ and $\{2,3\}$. Consider the action for player 2 in state 1, say action a_1^2, which moves him to state 3. Then, according to part 1 of lemma 3, we should have $v^2_{(1,3)} \leq v^2_{(1,2)}$.

Suppose, as in the proof of the lemma, that the game Ω starts in state $(1,2)$, while the game $\widetilde{\Omega}$ starts in state $(1,3)$. In the picture, the x-path will show how play develops in $\widetilde{\Omega}$, while the y-path will show how play develops in the beginning before joining the x-path. Suppose that $\pi^1 = \pi^{-i}$ in Ω prescribes for player 1 in state $(1,2)$ to move to state 3. As, afterwards, player 1 can only move along the cycle $\{3,4,5\}$, the strategy π^1 is unique for the rest of play. As $\widetilde{\pi}^1 = \widetilde{\pi}^{-i}$ is a copy of π^1, the strategy $\widetilde{\pi}^1$ prescribes in $\widetilde{\Omega}$ for player 1 in state $(1,3)$ to also move to state 3, and subsequently to move along the cycle $\{3,4,5\}$. Assume that $\widetilde{\pi}^2 = \widetilde{\pi}^i$ in $\widetilde{\Omega}$ prescribes for player 2 in state $(1,3)$ to move to state 4. Afterwards, player 2 can only move along the cycle $\{4,5,6\}$ for the rest of play. Now, as described in the proof, the strategy $\pi^2 = \pi^i$ in Ω will copy $\widetilde{\pi}^2$ after an initial phase. In this initial phase, player 2 has to reach state 3 at a stage of the form $l \cdot \lambda + 1$, where $\widetilde{\Omega}$ started. Recall that $\lambda = 6$. In figure 3, in the y-path, player 2 arrives at state 3 at stage $7 = 6 + 1$. After this, player 2 will copy strategy $\widetilde{\pi}^2$, and accordingly, he moves to state 4 and then follows the cycle $\{4,5,6\}$. As we can see, play in Ω (the y-path moving onto the x-path) and play in $\widetilde{\Omega}$ (the x-path) come together in the same segment, i.e. segment $\{(3,4),(4,5),(5,6)\}$.

It is essential that player 2 waits for a stage $l \cdot \lambda + 1$ before closing the initial phase of π^2. In our case, player 2's first visit to state 3 is at stage 3, when play in Ω is in state $(4,3)$. If player 2 decided to start copying $\widetilde{\pi}^2$, then, after state $(4,3)$, the y-path would continue $(5,4),(3,5)$, and so on, yielding a different segment, i.e. segment $\{(3,5),(4,6),(5,4)\}$. $\star$

7 Proof of Lemma 8

In this section, we prove lemma 8. We will only prove part 1; the proof of part 2 is similar.

Due to $z^i_F < v^i_F$, when starting in segment F, player i can only defend his minmax-level v^i_F if he leaves F. Therefore, there must be at least one state $s^* \in F$, joint action $a^{-i}_{s^*} \in \bar{A}^{-i}_{s^*}$ for players $-i$ and action $a^i_{s^*} \in A^i_{s^*} - \bar{A}^i_{s^*}$ for player i such that

$$\sum_{t \in S} p_{s^*,(a^i_{s^*},a^{-i}_{s^*})}(t)\, v^i_t \geq v^i_F. \tag{27}$$

Take the unique $m \in \{1, \ldots, \lambda_F\}$ for which $s^* \in T_F(m)$.

Now we will show that this m, state s^{*i} and action $a^i_{s^*}$ satisfy part 1 of the lemma. To this end, consider any state $(s^{*i}, s^{-i}) \in T_F(m)$ and a joint action $a^{-i}_{s^{-i}}$ for players $-i$ in joint state s^{-i}. We need to show

$$\sum_{t \in S} p_{(s^{*i},s^{-i}),(a^i_{s^*},a^{-i}_{s^{-i}})}(t)\, v^i_t \geq v^i_F. \tag{28}$$

Take a joint action $b^{-i}_{s^{-i}} \in \bar{A}^{-i}_{s^{-i}}$ in state s^{-i} for players $-i$. As $b^{-i}_{s^{-i}} \in \bar{A}^{-i}_{s^{-i}}$ and $a^{-i}_{s^*} \in \bar{A}^{-i}_{s^*}$, and both joint states s^{-i} and s^{*-i} belong to $T^{-i}_{k^{-i}}(m)$, we conclude that when players $-i$ either use joint action $b^{-i}_{s^{-i}}$ in state s^{-i} or joint action $a^{-i}_{s^*}$ in state s^{*-i}, they are certain to move to states in $T^{-i}_{k^{-i}}(m+1)$. Hence, by part 1 of lemma 7, we have for any $t^i \in S^i$ that

$$\sum_{t^{-i} \in S^{-i}} p^{-i}_{s^{-i} b^{-i}_{s^{-i}}}(t^{-i})\, v^i_{(t^i,t^{-i})} = \sum_{t^{-i} \in S^{-i}} p^{-i}_{s^{*-i} a^{-i}_{s^*}}(t^{-i})\, v^i_{(t^i,t^{-i})},$$

which yields

$$\begin{aligned}
\sum_{t\in S} p_{(s^{*i},s^{-i}),(a^i_{s^*},b^{-i}_{s^{-i}})}(t)\, v^i_t &= \sum_{t^i\in S^i} p^i_{s^{*i}a^i_{s^*}}(t^i)\left[\sum_{t^{-i}\in S^{-i}} p^{-i}_{s^{-i}b^{-i}_{s^{-i}}}(t^{-i})\, v^i_{(t^i,t^{-i})}\right] \\
&= \sum_{t^i\in S^i} p^i_{s^{*i}a^i_{s^*}}(t^i)\left[\sum_{t^{-i}\in S^{-i}} p^{-i}_{s^{*-i}a^{-i}_{s^*}}(t^{-i})\, v^i_{(t^i,t^{-i})}\right] \\
&= \sum_{t\in S} p_{s^*,(a^i_{s^*},a^{-i}_{s^*})}(t)\, v^i_t.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{t\in S} p_{(s^{*i},s^{-i}),(a^i_{s^*},a^{-i}_{s^{-i}})}(t)\, v^i_t &\geq \sum_{t\in S} p_{(s^{*i},s^{-i}),(a^i_{s^*},b^{-i}_{s^{-i}})}(t)\, v^i_t \\
&= \sum_{t\in S} p_{s^*,(a^i_{s^*},a^{-i}_{s^*})}(t)\, v^i_t \\
&\geq v^i_F,
\end{aligned}$$

where the first inequality follows from part 2 of corollary 4 (as $b^{-i}_{s^{-i}} \in \bar{A}^{-i}_{s^{-i}}$), whereas the last one from (27). This completes the proof of (28) and part 1 of lemma 8. ■

8 References

Altman E, Avrachenkov K, Bonneau N, Debbah M, El-Azouzi R & Sadoc Menasche D [2007,a]: Constrained Cost-Coupled Stochastic Games with Independent State Processes. Preprint.

Altman E, Avrachenkov K & Garnaev A [2007,b]: A Jamming Game in Wireless Networks with Transmission Cost. *Lecture Notes in Computer Science* 4465, 1-12.

Altman E, Avrachenkov K, Marquez R & Miller G [2005]: Zero-sum constrained stochastic games with independent state processes. *Mathematical Methods of Operations Research* 62, 375-386.

Blackwell D & Ferguson TS [1968]: The big match. *Annals of Mathematical Statistics* 39, 159-163.

Doob JL [1953]: *Stochastic processes.* Wiley, New York.

Dutta PK [1995]: A Folk theorem for stochastic games. *Journal of Economic Theory* 66, 1-32.

Filar JA & Vrieze OJ [1996]: *Competitive Markov Decision Processes - Theory, Algorithms, and Applications.* Springer-Verlag, New York.

Fink AM [1964]: Equilibrium in a stochastic n-person game. *Journal of Science of Hiroshima University,* Series A-I 28, 89-93.

Flesch J, Thuijsman F & Vrieze OJ [1997]: Cyclic Markov equilibria in a cubic game. *International Journal of Game Theory* 26, 303-314.

Flesch J, Thuijsman F & Vrieze OJ [2007]: Stochastic games with additive transitions. *European Journal of Operational Research* 179, 483-497.

Flesch J, Schoenmakers G & Vrieze OJ [2008]: Stochastic games on a product state space. *Mathematics of Operations Research* 33, 403-420.

Gillette D [1957]: Stochastic games with zero stop probabilities. In: Dresher M, Tucker AW & Wolfe P (eds.), Contributions to the theory of games III, *Annals of Mathematical Studies* 39, Princeton University Press, 179-187.

Mertens JF & Neyman A [1981]: Stochastic games. *International Journal of Game Theory* 10, 53-66.

Neyman A [2003]: *Existence of the minmax.* In: Neyman A & Sorin S (eds.), *Stochastic Games and Applications,* NATO Science series, Vol. 570, Kluwer Academic Press, 173-193.

Parthasarathy T & Raghavan TES [1981]: An orderfield property for stochastic games when one player controls transition probabilities. *Journal of Optimization Theory and Applications* 33, 375-392.

Ross KW & Varadarajan R [1991]: Multichain Markov decision processes with a sample path constraint: A decomposition approach. *Mathematics of Operations Research* 16, 195-207.

Simon R [2003]: Value and perfection in stochastic games. Working paper, Goettingen University, Germany. To appear in *Israel Journal of Mathematics*.

Solan E [1999]: Three-Player Absorbing Games. *Mathematics of Operations Research* 24, 669-698.

Solan E [2003]: *Perturbations of Markov chains with applications to stochastic games.* In: Neyman A & Sorin S (eds.), *Stochastic Games and Applications,* NATO Science series, Vol. 570, Kluwer Academic Press, 265-280.

Solan E & Vieille N [2002]: Correlated Equilibrium in Stochastic Games. *Games and*

Economic Behavior 38, 362-399.

Solan E & Vohra R [2002]: Correlated Equilibrium and Public Signalling in Absorbing Games. *International Journal of Game Theory* 31, 91-121.

Sorin S [1986]: Asymptotic properties of a non-zerosum game. *International Journal of Game Theory,* 15, 101-107.

Takahashi M [1964]: Equilibrium points of stochastic noncooperative n-person games. *Journal of Science of Hiroshima University,* Series A-I 28, 95-99.

Thuijsman F & Vrieze OJ [1991]: Easy initial states in stochastic games. In: Raghavan TES, Ferguson TS, Vrieze OJ & Parthasarathy T (eds.), *Stochastic Games and Related Topics*, Kluwer Academic Press, 85-100.

Vieille N [2000-a]: Equilibrium in 2-person stochastic games I: A reduction. *Israel Journal of Mathematics,* 119, 55-91.

Vieille N [2000-b]: Equilibrium in 2-person stochastic games II: The case of recursive games. *Israel Journal of Mathematics,* 119, 93-126.