

Stochastic games on a product state space: the periodic case

János Flesch · Gijs Schoenmakers · Koos Vrieze

Accepted: 29 January 2009 / Published online: 4 March 2009

© The Author(s) 2009. This article is published with open access at Springerlink.com

Abstract We examine so-called product-games. These are n -player stochastic games played on a product state space $S^1 \times \dots \times S^n$, 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. (Math Oper Res 33, 403–420, 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. (Math Oper Res 33, 403–420, 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

J. Flesch
Department of Quantitative Economics, Maastricht University,
P.O. Box 616, 6200 MD Maastricht, The Netherlands
e-mail: j.flesch@ke.unimaas.nl

G. Schoenmakers (✉) · K. Vrieze
Department of Mathematics, Maastricht University,
P.O. Box 616, 6200 MD Maastricht, The Netherlands
e-mail: gm.schoenmakers@micc.unimaas.nl

K. Vrieze
e-mail: koos@ictvrieze.nl

1 Introduction

Stochastic games An n -player stochastic game is given by (1) a set of players $N = \{1, \dots, n\}$, (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 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 starts in an initial state $s \in S$ and is to be played at stages in \mathbb{N} in the following way. At any stage m , in the current state $s_m \in S$, every player $i \in N$ has to choose an action a_m^i from his action set $A_{s_m}^i$, independently of the other players. The chosen joint action $a_m = (a_m^1, \dots, a_m^n)$ induces a payoff $r_{s_m}^i(a_m)$ to each player i , and a transition to a new state according to the transition probability distribution $p_{s_m a_m}$, where play will continue at stage $m + 1$. We assume complete information (i.e. the players know all the data of the stochastic game), full monitoring (i.e. the players observe the current state and the actions chosen by all the players), and perfect recall (i.e. the players remember all previous states and actions).

Strategies A mixed action x_s^i for player i in state $s \in S$ is a probability distribution on A_s^i . The set of mixed actions for player i in state s is denoted by X_s^i . 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 $\pi_s^i(h) \in X_s^i$ in the current state s as a function of the 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 action to be played with probability 1.

If the mixed actions prescribed by a strategy depend only on the current state then the strategy is called stationary. Thus, the stationary strategy space for player i is $X^i = \times_{s \in S} X_s^i$. We use the notation x^i for stationary strategies for player i , while x_s^i 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 $A^i = \times_{s \in S} A_s^i$.

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 for player i is

$$\gamma_s^i(\pi) := \liminf_{M \rightarrow \infty} \mathbb{E}_{s\pi} \left(\frac{1}{M} \sum_{m=1}^M R_m^i \right) = \liminf_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M \mathbb{E}_{s\pi} \left(R_m^i \right),$$

where R_m^i 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 .

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

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

where $\pi^{-i} = (\pi^j)_{j \in N - \{i\}}$, 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.

The game called the Big Match, introduced by Gillette (1957) and solved by Blackwell and Ferguson (1968), and the game in Sorin (1986) demonstrate that 0-equilibria need not exist with respect to the average reward. In fact, history dependent strategies are indispensable for establishing ε -equilibria, for $\varepsilon > 0$.

For two-player stochastic games, Vieille (2000a,b) managed to establish the existence of ε -equilibria, for all $\varepsilon > 0$. However, little is known about n -player stochastic games. It is unknown whether these games always possess ε -equilibria, for all $\varepsilon > 0$, and this question is the most challenging open problem in the field of stochastic games. The main difficulties are caused by the discontinuity of the average reward on the spaces of stationary strategies. Given these difficulties, interest arises in finding equilibria in special classes.

The class of product-games 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 $A^i_{s^i}$ for each state $s^i \in S^i$; (3) a transition probability distribution $p^i_{s^i a^i_{s^i}}$ over the state space S^i for each state $s^i \in S^i$ and for each action $a^i_{s^i} \in A^i_{s^i}$. 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 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 $\Gamma^1, \Gamma^2, \dots, \Gamma^n$, is an n -player stochastic game for which (1) the set of players is $N = \{1, \dots, n\}$; (2) the state space is $S = S^1 \times \dots \times S^n$; (3) the action set for each player $i \in N$ in each state $s = (s^1, \dots, s^n) \in S$ is $A^i_s = A^i_{s^i}$; (4) the transition probability distribution p_{sa_s} , for each state $s = (s^1, \dots, s^n) \in S$ and for each joint action $a_s = (a^1_s, \dots, a^n_s) \in \times_{i \in N} A^i_s$, is

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

for state $\bar{s} = (\bar{s}^1, \dots, \bar{s}^n) \in S$. Note that no condition is imposed on the payoff structure. As a consequence, play of the product game G can be viewed as simultaneous play of the n Markov transition structures $\Gamma^1, \dots, \Gamma^n$, which are linked by payoff functions r^1, \dots, r^n that may depend on all n current states as well as on all n actions chosen by the players.

This kind of transition structure is natural, and has already been considered in applications. For applications in wireless networks, Altman et al. (2005) (see also Altman et al. 2007a,b) examined two-player product-games, although in a somewhat different fashion. They assumed that the sum of the payoffs is always equal to zero (zero-sum games), and dropped the assumption of full monitoring by letting each player observe only his own coordinate of the current state and the action chosen by himself. As a result, both players have to make choices without knowledge of 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). Other models related to product-games are certain types of multi-agent systems. These models use the same kind of transition structure as product-games, but here the players (or agents) have a common goal. The main difficulty is caused by the assumption that the players cannot fully monitor the other players' behavior. We refer to [Becker et al. \(2003\)](#), and all the references therein.

Note that stochastic games with a single controller (cf. [Parthasarathy and Raghavan 1981](#) or [Filar and Vrieze 1996](#)), i.e. when one player controls the transitions, 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.

For the class of n -player product-games, we prove the existence of 0-equilibria (cf. [Theorem 1](#)). This extends [Flesch et al. \(2008a\)](#), where aperiodicity was assumed on the transition structure of the product-game (cf. [Sect. 2](#), for a precise definition of aperiodicity). The construction makes use of the fact that, in a product-game, player i controls the i -th component of the transitions. This property will enable us to analyse the Markov transition structure of each player separately.

Zero-sum games and optimality In the study of stochastic games, the class of zero-sum stochastic games play a special role. These are two-player stochastic games for which $r_s^2(a_s) = -r_s^1(a_s)$, for each state s and for each joint action a_s . In these games the two players have completely opposite interests. Thus, in a zero-sum game, player 1 wants to maximize his own reward, while player 2 tries to minimize player 1's reward. For simplicity, let $\gamma = \gamma^1$. [Mertens and Neyman \(1981\)](#) showed that, for all $s \in S$,

$$\sup_{\pi^1} \inf_{\pi^2} \gamma_s(\pi^1, \pi^2) = \inf_{\pi^2} \sup_{\pi^1} \gamma_s(\pi^1, \pi^2) := v_s;$$

here v_s is called the value for initial state s . 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. [Mertens and Neyman \(1981\)](#) proved that both players have ε -optimal strategies for any $\varepsilon > 0$, though history dependent strategies may be necessary for ε -optimality.

For the class of zero-sum product-games, we show that both players have stationary 0-optimal strategies (cf. [Theorem 2](#)). In addition, we analyse the structure of the value of these games.

The structure of the article In [Sect. 2](#), we discuss preliminary concepts. In [Sect. 3](#), we present our main results and discuss the main difficulties which we encounter when facing periodic product-games. The proofs are given in [Sect. 4](#).

2 Preliminary concepts

From now on, we use superscripts to index the player and subscripts for the state. Whenever one of them is omitted, we 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 $\times_{s \in S} A_s^i$. 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 $\times_{j \in N - \{i\}} S^j$.

Minmax-levels Take an arbitrary n -player stochastic game. For any player $i \in N$ and initial state $s \in S$, let

$$v_s^i := \inf_{\pi^{-i} \in \Pi^{-i}} \sup_{\pi^i \in \Pi^i} \gamma_s^i(\pi^i, \pi^{-i}). \tag{1}$$

Here v_s^i 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 secure 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 for any player i

$$v_s^i = \min_{x_s^{-i} \in X_s^{-i}} \max_{x_s^i \in X_s^i} \sum_{t \in S} p_{s, (x_s^i, x_s^{-i})}(t) v_t^i. \tag{2}$$

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 $s \in S$. Moreover, for zero-sum games, the value satisfies $v = v^1 = -v^2$, shown by [Mertens and Neyman \(1981\)](#).

Markov chains 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 the $(s, \bar{s})^{\text{th}}$ entry of $P(x)$ gives the transition probability $p_{s, \bar{s}}(x)$ for moving from state s to state \bar{s} when the joint mixed action x_s is played in state s . Associated with this Markov-chain, there are transient and recurrent states. We can group the recurrent states into ergodic sets. 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 $W^1, \dots, W^{\lambda(W)}$, i.e. when starting in any $s \in W^l$, the process will move through the cyclic sets in the order $W^l, W^{l+1}, \dots, W^{\lambda(W)}, W^1, \dots, W^{l-1}, W^l, \dots$. 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 and Snell \(1960\)](#) for a more detailed discussion on Markov chains.

Some of the contents of the remainder of this section are 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. \(2008a\)](#). For an illustration of the concepts below, we refer to [Example 1](#).

Classification of states First, we analyse the Markov transition structure Γ^i of each player i separately. (Note that a separate analysis of the transition structure is possible only because each 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 $s^i \in S^i$ is of type 1 if it is transient for each stationary strategy x^i of player i . Otherwise, s^i is of type 2, in which case player i has a stationary strategy for which s^i is recurrent.

Maximal communicating sets Two states s_1^i and s_2^i of type 2 are said to communicate with each other, if there exists a stationary strategy x^i of player i such that s_1^i and s_2^i belong to the same ergodic set. We note that communication between states have been used extensively in the literature of stochastic games (cf. [Vieille 2000a,b](#); [Solan and Vohra 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, called maximal communicating sets. Every maximal communicating set E^i has the following properties:

(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 steps without leaving E^i .

(2) If player i decides to leave E^i , the probability that he returns to E^i is strictly less than 1, regardless of 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 of 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_ρ with probability at least $1 - \rho$).

(4) There is at least one maximal communicating set which player i is unable to leave, i.e. no transitions to states outside are possible.

(5) For any strategy of player i , regardless of 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 K^i denote the index set for player i 's maximal communicating sets. We use the notation $E_{k^i}^i$, where $k^i \in K^i$, for a maximal communicating set for player i . In every state s^i of $E_{k^i}^i$, for every $k^i \in K^i$, let $\bar{A}_{s^i}^i$ denote the set of those actions $a_{s^i}^i \in A_{s^i}^i$ which keep play in $E_{k^i}^i$ with probability 1. The sets $\bar{A}_{s^i}^i$ are clearly nonempty. For every state $s = (s^1, \dots, s^n) \in S$, we also let $\bar{A}_s^i := \bar{A}_{s^i}^i$. Notice that, if x^i is a stationary strategy for player i which, in every state $s^i \in E_{k^i}^i$, places a positive probability on all actions in $\bar{A}_{s^i}^i$ and places probability zero on all actions in $A_{s^i}^i - \bar{A}_{s^i}^i$, then $E_{k^i}^i$ is an ergodic set for x^i .

Periodicity and segments The period of $E_{k^i}^i$, denoted by $\lambda_{k^i}^i$, is defined as the period of the Markov chain on $E_{k^i}^i$ associated with a stationary strategy x^i of player i that uses only completely mixed actions on $\bar{A}_{s^i}^i$ for all $s^i \in E_{k^i}^i$. (Obviously, the period is independent of the particular choice of x^i .) The cyclic sets of $E_{k^i}^i$ are denoted by

$T_{k^i}^i(1), \dots, T_{k^i}^i(\lambda_{k^i}^i)$. For convenience, let $T_{k^i}^i(u \cdot \lambda_{k^i}^i + w) := T_{k^i}^i(w)$ for all $u \in \mathbb{N}$ and $w \in \{1, \dots, \lambda_{k^i}^i\}$.

Let $K := \times_{i=1}^n K^i$. Consider the product $E_k := \times_{i=1}^n E_{k^i}^i$ for some $k = (k^1, \dots, k^n) \in K$. The period of E_k , denoted by λ_k , is defined as the period of the Markov chain on E_k associated with a joint stationary strategy x that uses only joint completely mixed actions on \bar{A}_s for all $s \in E_k$. Clearly, λ_k equals the least common multiple of $\lambda_{k^1}^1, \dots, \lambda_{k^n}^n$. 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 $s^i \in T_{k^i}^i(l_i)$ for some $l_i \in \{1, \dots, \lambda_{k^i}^i\}$, for all $i \in N$, then the segment F containing state s has cyclic sets of the form

$$T_F(m) := T_{k^1}^1(l_1 + m - 1) \times \dots \times T_{k^n}^n(l_n + m - 1), \quad m = 1, \dots, \lambda_k. \tag{3}$$

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

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 $E_k = \times_{i=1}^n E_{k^i}^i$. 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}_s^i , 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, due to the following observation. Recall that, for any initial state and strategies of the players, each player i eventually settles in one of his maximal communicating sets $E_{k^i}^i$, with probability 1. Hence, with probability 1, play will eventually *settle* in a segment $F \subset E_k$ 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 $\bar{v}_{F,s}^i$ denote the minmax-level of player i in \bar{G}_F for initial state $s \in F$. If, for some player i , the inequality $\bar{v}_{F,s}^i \geq v_s^i$ holds for all initial states $s \in F$, then we call \bar{G}_F *satisfactory* for player i . Otherwise, \bar{G}_F is called *unsatisfactory* for player i . In words, if \bar{G}_F is satisfactory for 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 for all players. Further, let $\mathcal{F}[i]$ denote the set of segments F such that \bar{G}_F is unsatisfactory for player i but \bar{G}_F is satisfactory for all players $j \in \{1, \dots, i - 1\}$. Obviously, $\mathcal{F}^*, \mathcal{F}[1], \dots, \mathcal{F}[n]$ forms a partition of all segments.

Example 1 As an illustration, consider the product-game in Fig. 1 with two players. 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 payoffs to the respective players are given in the upper-left corner, while the next state is indicated in the bottom-right

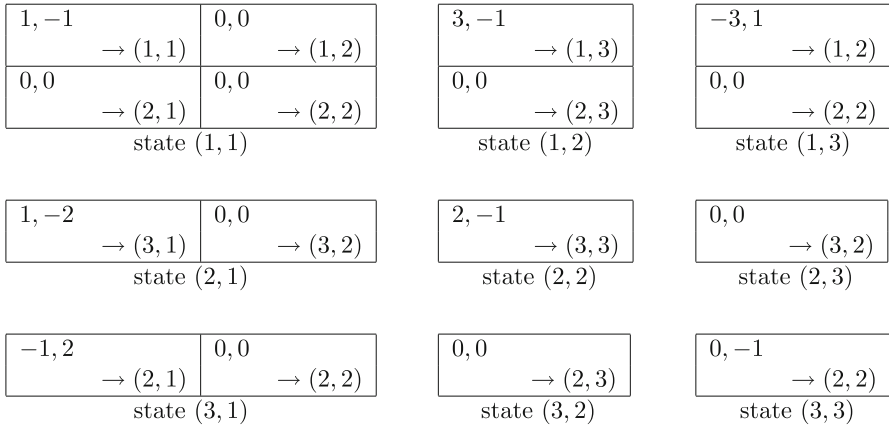


Fig. 1 Game of Example 1

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 $A_1^1 = \{1, 2\}$, $A_2^1 = A_3^1 = \{1\}$, and transition probabilities

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

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 $E_I^1 := \{1\}$ and $E_{II}^1 := \{2, 3\}$ are maximal communicating sets, with index-set $K^1 = \{I, II\}$. Moreover, E_I^1 is aperiodic (i.e. has periodicity 1) whereas E_{II}^1 has periodicity 2. As for the actions which keep play in these maximal communicating sets, we obtain $\bar{A}_1^1 = \bar{A}_2^1 = \bar{A}_3^1 = \{1\}$.

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

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

Further, $E_I^2 := \{1\}$ and $E_{II}^2 := \{2, 3\}$ are maximal communicating sets, with index-set $K^2 = \{I, II\}$. The maximal communicating 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.

3 The main results

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

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 derive the existence of stationary solutions.

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

We focus on the Theorem 1 and its proof, as Theorem 2 will follow directly (cf. end of Sect. 4.3).

Existence of stationary 0-equilibria Our construction for 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 case 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. Corollary 8 below (through corollary 12) establishes for such games that all minmax-levels are constant on the whole state space S . It is unclear how to exploit this fact.

The difference between the periodic and aperiodic cases Interestingly, the period of the product-game plays a central role in the analysis. In Flesch et al. (2008a), the special case of aperiodic product-games has been extensively studied, and the validity of Theorem 1 has been shown for all aperiodic product-games. The approach presented there is applicable to periodic product-games as well, but the periodic case poses a number of additional problems that we address. In the periodic case, as discussed in Sect. 2, if E_{ki}^i denotes a maximal communicating set for every player i , then the product $E_k = \times_{i=1}^n E_{ki}^i$ may split 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. (2008a) 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 segment of E_k separately, cf. Corollary 8 together with Corollary 12).

B. The central lemma of the aperiodic case is invalid for periodic product-games and has to be modified. We refer to Lemma 5 and the remark after that.

C. In the aperiodic case, moving to a set E_k can be achieved by letting each player i move to E_{ki}^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 $s^i \in E_{ki}^i$ is the state and m^i is the stage at which player i enters E_{ki}^i . Thus, it is crucial to arrive at the right segment within E_k . See the proof of Lemma 5 and the example after that.

An attempt to transform periodic product-games into aperiodic ones In order to show 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 $E_k = \times_{i=1}^n E_{ki}^i$, where E_{ki}^i denotes a maximal communicating set for every player i . As discussed in Sect. 2, if the product-game is aperiodic, then E_k is one segment and all states in E_k communicate. If the product-game is periodic, then E_k may split 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 .

4 The formal proofs of main Theorems 1 and 2

In this section, we provide proofs for Theorems 1 and 2. We will focus on Theorem 1, as Theorem 2 will follow (cf. the end of Sect. 4.3). In Sect. 4.1, we examine restricted games. In Sect. 4.2, we analyze the minmax-levels of the players in so-called simple product-games, and then in Sect. 4.3 we extend this to the general case. By combining these results, we prove Theorem 1 in Sect. 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 \tilde{G}_F , corresponding to segment F . First, we analyse the minmax-levels of the players in \tilde{G}_F , and then we discuss possible equilibrium rewards within \tilde{G}_F .

First, it turns out that each player i 's minmax-level in \tilde{G}_F is constant. In addition, players $-i$ can make sure in \tilde{G}_F that player i 's reward is at most his minmax-level (i.e. the infimum is attained in (1) for the game \tilde{G}_F).

Lemma 3 *Let G be a product-game. Consider the restricted game \tilde{G}_F , corresponding to some segment F , and an arbitrary player i . Then, the minmax-level \bar{v}_F^i of player i in \tilde{G}_F is constant, i.e. $\bar{v}_{F,s}^i = \bar{v}_{F,t}^i (=:\bar{v}_F^i)$ for all states $s, t \in F$. Moreover, in \tilde{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}_F^i .*

The proof is analogous to the proof of lemma 3.1 for aperiodic product-games in Flesch et al. (2008a). We also refer to Flesch et al. (2008b). Here, we provide a brief outline. Consider a restricted game \tilde{G}_F and a player i . Let $\alpha^i := \min_{t \in F} \bar{v}_{F,t}^i$, which is the lowest minmax-level in the restricted game \tilde{G}_F for player i . The idea of the proof

is to find a set $U \subset \{t \in F \mid \bar{v}_{F,t}^i = \alpha^i\}$ and a joint stationary strategy x^{-i} (in \tilde{G}_F) such that, irrespective of the strategy of player i , the following hold: (1) once play reaches U , it remains in U forever, and player i 's reward within U is at most α^i , (2) play reaches U with probability 1. Clearly, these properties will then imply that the minmax-level \bar{v}_F^i of player i in \tilde{G}_F equals the constant α^i , and that x^{-i} satisfies the second part of the lemma. We remark that the existence of such U and x^{-i} is based on the work of [Thuijsman and Vrieze \(1991\)](#) on stationary strategies for particular initial states, in combination with [Neyman \(2003\)](#).

As an illustration, we now revisit the game in [Example 1](#). We find in accordance with [Lemma 3](#) that the minmax-levels of the players are constant in the restricted games. Indeed, for player 1 we have that

$$\bar{v}_{F(I,I)}^1 = 1, \quad \bar{v}_{F(I,II)}^1 = 0, \quad \bar{v}_{F(II,I)}^1 = 0, \quad \bar{v}_{F(II,II),1}^1 = 1, \quad \bar{v}_{F(II,II),2}^1 = 0, \quad (4)$$

while for player 2 that

$$\bar{v}_F^2 = -\bar{v}_F^1 \quad (5)$$

for any segment F .

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 $\pi_s^i[h](h')$ in any current state s for any history h' as if h had happened before h' , i.e. $\pi_s^i[h](h') = \pi_s^i(h \oplus h')$, 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 .

Next, we present a possible equilibrium for the restricted game \tilde{G}_F . We show that there exists a 0-equilibrium in \tilde{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 three players) and [Simon \(2003\)](#) (with only two players) constructed examples proving that such a result does not hold for all stochastic games.

Lemma 4 *Let G be a product-game. Consider the restricted game \tilde{G}_F , corresponding to some segment F . Then, there exists a 0-equilibrium π in \tilde{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 $\gamma_s^i(\pi[h])$ is independent of the initial state $s \in F$ and the history h , given h occurs with a positive probability with respect to π .*

The proof is the same as for lemma 3.7 in [Flesch et al. \(2008a\)](#). We also refer to [Flesch et al. \(2008b\)](#). Here, we provide only a brief outline. Note:

- (i) The minmax-levels of the players in \tilde{G}_F are constant, by [Lemma 3](#).
- (ii) The set of feasible rewards in \tilde{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. \(2008a\)](#), based on [Dutta \(1995\)](#).

Given these three observations, this game situation is almost identical to a repeated game. The proof 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 *simple* if, within every restricted game \tilde{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 $a_s \in \bar{A}_s, b_t \in \bar{A}_t$. Let z_F^i denote this unique payoff for player i in the restricted game \tilde{G}_F . Thus, in simple product-games, when play settles in one of the restricted games \tilde{G}_F , the rewards of the players will equal z_F .

Example 2 Consider the simple product-game G with two players given in Fig. 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 argue that

$$v_s^1 = \begin{cases} 0 & \text{if } s \in F_{(II,I)} \cup F_{(II,II),2} = \{(2, 1), (3, 1), (2, 3), (3, 2)\} \\ 1 & \text{if } s \in F_{(I,I)} \cup F_{(I,II)} \cup F_{(II,II),1} = \{(1, 1), (1, 2), (1, 3), (2, 2), (3, 3)\}. \end{cases} \quad (6)$$

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.

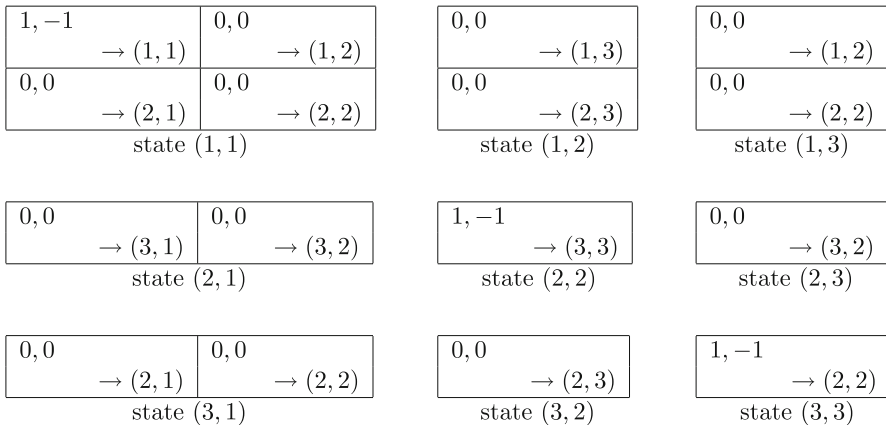


Fig. 2 Game of Example 2

We similarly find that

$$v_s^2 = -v_s^1 \tag{7}$$

for all $s \in S$.

Consider a state $s \in S$ within a simple product-game, such that state s^i is of type 2 for player i . (Recall that a state is of type 2, if it belongs to a maximal communicating set). Suppose that s^i belongs to a cyclic set $T_{k^i}^i(m^i)$ of a maximal communicating set $E_{k^i}^i$. Consider the situation in which 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 $T_{k^i}^i(m^i + 1)$, and the new state of the game becomes (t^i, s^{-i}) , or (2) player i can choose any action a_s^i in state s^i , and the new state of the game becomes (t^i, s^{-i}) , $t^i \in S^i$, with probability $p_{s^i a_s^i}^i(t^i)$. Note that players $-i$ remain in state s^{-i} in either case. In the lemma below we show that option 1 is always at least as good as option 2, as far as player i 's minmax-level is concerned. We also show the validity of a similar statement for players $-i$.

The proof is far from straightforward. We provide an intuition for the lemma. Consider player i and his two options described above. By taking option 1, player i is certain to remain in the same maximal communicating set $E_{k^i}^i$. On the other hand, by playing an action $a_{s^i}^i$ in option 2, player i possibly leaves $E_{k^i}^i$ and strategically restricts himself, as he will not be able to return to $E_{k^i}^i$ with probability 1 (cf. Property 2 of maximal communicating sets in Sect. 2). In this sense, waiting in $E_{k^i}^i$ provides no worse future prospects.

As an illustration, we revisit the simple product-game in Example 2. Consider player $i = 1$ and state $s = (1, 1)$ belonging to the maximal communicating set $E_1^1 = \{1\}$ for player 1. As we know, E_1^1 is aperiodic, thus E_1^1 has only one cyclic set, i.e. E_1^1 . For option 1 of player 1, consider $t^1=1$ yielding state $(1, 1)$; while for option 2, consider action 2 yielding state $(2, 1)$ with probability 1. As we know, $v_{(1,1)}^1 = 1 \geq 0 = v_{(2,1)}^1$, which means that option 1 is at least as good as option 2 indeed.

Lemma 5 *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 $T_{k^i}^i(m^i)$ of some maximal communicating set $E_{k^i}^i$. Consider any action $a_{s^i}^i \in A_{s^i}^i$ in state s^i for player i . Then, for any state $t^i \in T_{k^i}^i(m^i + 1)$, we have*

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

- (2) *Suppose that state s^j is of type 2 for every player $j \neq i$, and belongs to cyclic set $T_{k^j}^j(m^j)$ of some maximal communicating set $E_{k^j}^j$. Thus, $s^{-i} \in T_{k^{-i}}^{-i}(m^{-i})$. Consider any joint action $a_s^{-i} \in A_s^{-i}$ for players $-i$. Then, for any joint state $t^{-i} \in T_{k^{-i}}^{-i}(m^{-i} + 1)$, we have*

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

Proof 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 $T_{k^i}^i(1), \dots, T_{k^i}^i(\lambda_{k^i}^i)$ 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

$$\sum_{u^i \in S^i} p_{s^i a_{s^i}^i}^i(u^i) v_{(u^i, s^{-i})}^i \leq v_{(t^i, s^{-i})}^i. \tag{8}$$

□

The idea of the proof Let $\varepsilon > 0$. We compare two specific games Ω and $\tilde{\Omega}$. The game Ω is the original game G starting in state (t^i, s^{-i}) , whereas $\tilde{\Omega}$ is the game G which starts in initial state (u^i, s^{-i}) with probability $p_{s^i a_{s^i}^i}^i(u^i)$. We define two joint strategies

π for Ω and $\tilde{\pi}$ for $\tilde{\Omega}$ in such a way that the following properties hold:

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

Property (B) for $\tilde{\pi}$ in $\tilde{\Omega}$: $\gamma_{(u^i, s^{-i})}^i(\tilde{\pi}) \geq v_{(u^i, s^{-i})}^i - \varepsilon$, for every $u^i \in S^i$.

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

$$\gamma_{(t^i, s^{-i})}^i(\pi) = \sum_{u^i \in S^i} p_{s^i a_{s^i}^i}^i(u^i) \gamma_{(u^i, s^{-i})}^i(\tilde{\pi}).$$

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

$$\begin{aligned} \sum_{u^i \in S^i} p_{s^i a_{s^i}^i}^i(u^i) v_{(u^i, s^{-i})}^i &\leq \sum_{u^i \in S^i} p_{s^i a_{s^i}^i}^i(u^i) (\gamma_{(u^i, s^{-i})}^i(\tilde{\pi}) + \varepsilon) \\ &= \sum_{u^i \in S^i} p_{s^i a_{s^i}^i}^i(u^i) \gamma_{(u^i, s^{-i})}^i(\tilde{\pi}) + \varepsilon \\ &= \gamma_{(t^i, s^{-i})}^i(\pi) + \varepsilon \\ &\leq v_{(t^i, s^{-i})}^i + 2\varepsilon. \end{aligned}$$

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 $\gamma_{(t^i, s^{-i})}^i(\pi^{-i}, \sigma^i) \leq v_{(t^i, s^{-i})}^i + \varepsilon$ for any strategy σ^i for player i . Such a joint strategy π^{-i} exists by the definition of the minmax-level $v_{(t^i, s^{-i})}^i$. Thus, irrespective of the choice of π^i , property A will be satisfied.

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

(i) $\tilde{\pi}^{-i}$ for $\tilde{\Omega}$ is obtained by copying π^{-i} in the sense of Lemma 14. Here, by copying we mean that the joint strategy $\tilde{\pi}^{-i}$ is going to “mimic” π^{-i} by using the same mixed actions. At this point, it is crucial that players $-i$ start in s^{-i} in both games Ω and $\tilde{\Omega}$. This is discussed more precisely below. For the technical details, we refer to Lemma 14 in Appendix.

(ii) $\tilde{\pi}^i$ for $\tilde{\Omega}$ is then obtained by taking a strategy which defends player i 's minmax-level v , up to ε , against $\tilde{\pi}^{-i}$ (cf. the discussion 1). Note that $\tilde{\pi}$ will satisfy property B.

(iii) π^i for Ω is obtained by copying $\tilde{\pi}^i$ in the sense of Lemma 14. 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 $\tilde{\Omega}$. We now describe this initial phase more precisely. During this phase, the strategy π^i prescribes for player i to move from his initial state t^i to state s^i , within the maximal communicating set $E^i_{k^i}$. This can be done by 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 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 $\tilde{\Omega}$). 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 $\tilde{\pi}^{-i}$ is defined up to some stage m , so is $\tilde{\pi}^i$.

At stage 1: The history is empty. For all $u^i \in S$, let

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

which means that, in any initial state (u^i, s^{-i}) of $\tilde{\Omega}$, the joint strategy $\tilde{\pi}^{-i}$ prescribes the same joint mixed action as π^{-i} in the initial state (u^i, s^{-i}) of Ω . Note that in both cases, players $-i$ are in joint state s^{-i} . Moreover, Lemma 14 would prescribe the same mixed action for $\tilde{\pi}^{-i}$, when copying π^{-i} .

Given $\tilde{\pi}^{-i}$ at stage 1, we know $\tilde{\pi}^i$ 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 $\tilde{\pi}^{-i}$ copies π^{-i} at stage m in the sense of Lemma 14. Given $\tilde{\pi}^{-i}$ up to stage m , we know $\tilde{\pi}^i$ 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 $\tilde{\Omega}$.

By regarding this as the initial state and stage, π^i copies $\tilde{\pi}^i$ in the sense of Lemma 14,

based on $(\tilde{\pi}^i, \tilde{\pi}^{-i})$ up to stage m . (So, π^i for stage $m = l \cdot \lambda + m'$ is a copy of $\tilde{\pi}^i$ at stage m').

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

It remains to verify property C. Consider the joint strategies $\pi = (\pi^i, \pi^{-i})$ in Ω and $\tilde{\pi} = (\tilde{\pi}^i, \tilde{\pi}^{-i})$ in $\tilde{\Omega}$. We know that, 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 $\tilde{\Omega}$ too) is simple, by definition, the payoffs are constant in each restricted game \tilde{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 \tilde{G}_F) is the same with respect to $\pi = (\pi^i, \pi^{-i})$ in Ω and $\tilde{\pi} = (\tilde{\pi}^i, \tilde{\pi}^{-i})$ in $\tilde{\Omega}$.

Take an arbitrary segment F within some $E_k = \times_{j=1}^n E_{kj}^j$, where, as usual, E_{kj}^j denotes a maximal communicating set for player j . For any player j , state $u^j \in E_{kj}^j$ and stage m^j , let $\omega^j(u^j, m^j)$ denote the event that player j settles in E_{kj}^j in state u^j at stage m^j . Since $\tilde{\pi}^{-i}$ is a copy of π^{-i} , event $\omega^j(u^j, m^j)$, for any $j \neq i$ and any $u^j \in E_{kj}^j$ and stage m^j , has the same probability with respect to $\pi = (\pi^i, \pi^{-i})$ in Ω and with respect to $\tilde{\pi} = (\tilde{\pi}^i, \tilde{\pi}^{-i})$ in $\tilde{\Omega}$. By construction, π^i is also a copy of $\tilde{\pi}^i$, with a “delay” of $l \cdot \lambda$ stages. Therefore, event $\omega^i(u^i, m^i)$, for any $u^i \in E_{ki}^i$ and stage m^i , has the same probability with respect to $\tilde{\pi} = (\tilde{\pi}^i, \tilde{\pi}^{-i})$ in $\tilde{\Omega}$ 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 λ_{ki}^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 $\tilde{\pi} = (\tilde{\pi}^i, \tilde{\pi}^{-i})$ in $\tilde{\Omega}$. This completes the proof.

As an illustration of the proof of Lemma 5, consider the game with two players given in Fig. 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 5, and take player $i = 2$ and state $s = (1, 1)$. Note that $\{1, 2, 3\}$ 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 5, we should have $v_{(1,3)}^2 \leq v_{(1,2)}^2$.

Suppose, as in the proof of the lemma, that the game Ω starts in state $(1, 2)$, while the game $\tilde{\Omega}$ starts in state $(1, 3)$. In the picture, the x -path will show how play develops in $\tilde{\Omega}$, while the y -path will show how play develops in Ω 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 $\tilde{\pi}^1 = \tilde{\pi}^{-i}$ is a copy of π^1 , the strategy $\tilde{\pi}^1$ prescribes in

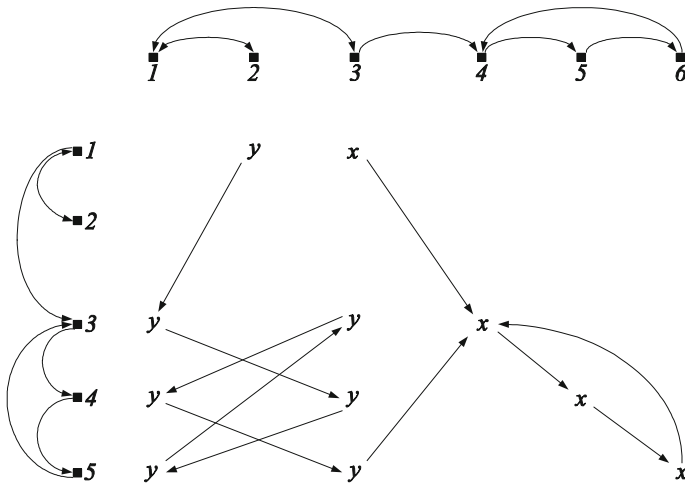


Fig. 3 Illustration for Lemma 5

$\tilde{\Omega}$ 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 $\tilde{\pi}^2 = \tilde{\pi}^i$ in $\tilde{\Omega}$ 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 $\tilde{\pi}^2$ 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 $\tilde{\Omega}$ started. Recall that $\lambda = 6$. In Fig. 3, in the y-path, player 2 arrives at state 3 at stage $7 = 6 + 1$. After this, player 2 will copy strategy $\tilde{\pi}^2$, 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 $\tilde{\Omega}$ (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 $\tilde{\pi}^2$, 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)}.

Remark A special case of Lemma 5 arises when the product-game is aperiodic. Consider part 1 of the Lemma 5. Then, as $T_{ki}^i(m^i) = T_{ki}^i(m^i + 1) = E_{ki}^i$, we obtain for every action $a_{s^i}^i$ of player i that

$$\sum_{u^i \in S^i} p_{s^i a_{s^i}^i}^i(u^i) v_{(u^i, s^{-i})}^i \leq v_{(s^i, s^{-i})}^i. \tag{9}$$

This means that even if player i had a solitary move, i.e. he could play an arbitrary action $a_{s^i}^i$ 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. (2008b) (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 (9) would not hold.

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

Lemma 6 *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 $a_s^i \in \bar{A}_s^i$ and $b_s^i \in A_s^i$ for player i and any joint action $a_s^{-i} \in A_s^{-i}$ for players $-i$. Then,*

$$\sum_{t \in S} p_{s,(a_s^i, a_s^{-i})}(t) v_t^i \geq \sum_{t \in S} p_{s,(b_s^i, a_s^{-i})}(t) v_t^i.$$

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

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

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

Since s^i is of type 2, s^i belongs to a cyclic set $T_{ki}^i(m^i)$ of some maximal communicating set E_{ki}^i . Then, by playing action a_s^i , player i actually moves to the next cyclic set $T_{ki}^i(m^i + 1)$. Hence, by part 1 of Lemma 5, for any $t^{-i} \in S^{-i}$ we have

$$\sum_{t^i \in S^i} p_{s^i a_s^i}^i(t^i) v_{(t^i, t^{-i})}^i \geq \sum_{t^i \in S^i} p_{s^i b_s^i}^i(t^i) v_{(t^i, t^{-i})}^i.$$

Therefore,

$$\begin{aligned} \sum_{t \in S} p_{s,(a_s^i, a_s^{-i})}(t) v_t^i &= \sum_{t^{-i} \in S^{-i}} p_{s^{-i} a_s^{-i}}^{-i}(t^{-i}) \left[\sum_{t^i \in S^i} p_{s^i a_s^i}^i(t^i) v_{(t^i, t^{-i})}^i \right] \\ &\geq \sum_{t^{-i} \in S^{-i}} p_{s^{-i} a_s^{-i}}^{-i}(t^{-i}) \left[\sum_{t^i \in S^i} p_{s^i b_s^i}^i(t^i) v_{(t^i, t^{-i})}^i \right] \\ &= \sum_{t \in S} p_{s,(b_s^i, a_s^{-i})}(t) v_t^i, \end{aligned}$$

completing the proof. □

Lemma 6 has useful implications. Suppose that player i is in a state of type 2 within a maximal communicating set E_{ki}^i . By using Lemma 6, we can deduce that

any action of player i , which keeps him in $E_{k_i}^i$ with probability 1, guarantees that his minmax-level cannot decrease in expectation after a transition. A similar result holds for players $-i$.

Corollary 7 *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 $a_s^i \in \bar{A}_s^i$ for player i and any joint action $a_s^{-i} \in A_s^{-i}$ for players $-i$. Then,*

$$\sum_{t \in S} p_{s, (a_s^i, a_s^{-i})}(t) v_t^i \geq v_s^i.$$

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

$$\sum_{t \in S} p_{s, (a_s^i, a_s^{-i})}(t) v_t^i \leq v_s^i.$$

Proof We prove part 1; the proof of part 2 is similar. Take a state s , action a_s^i and joint action a_s^{-i} as in part 1. From part 1 of Lemma 6, we have for any joint action $b_s^{-i} \in A_s^{-i}$ that

$$\sum_{t \in S} p_{s, (a_s^i, b_s^{-i})}(t) v_t^i = \max_{b_s^{-i} \in A_s^{-i}} \sum_{t \in S} p_{s, (b_s^{-i}, a_s^i)}(t) v_t^i.$$

Hence, by linearity, we may conclude for mixed actions $x_s^{-i} \in X_s^{-i}$ for players $-i$ in state s that

$$\sum_{t \in S} p_{s, (a_s^i, x_s^{-i})}(t) v_t^i = \max_{b_s^{-i} \in A_s^{-i}} \sum_{t \in S} p_{s, (b_s^{-i}, x_s^{-i})}(t) v_t^i = \max_{x_s^{-i} \in X_s^{-i}} \sum_{t \in S} p_{s, (x_s^{-i}, a_s^i)}(t) v_t^i.$$

Therefore,

$$\begin{aligned} \sum_{t \in S} p_{s, (a_s^i, a_s^{-i})}(t) v_t^i &\geq \min_{x_s^{-i} \in X_s^{-i}} \sum_{t \in S} p_{s, (a_s^i, x_s^{-i})}(t) v_t^i \\ &= \min_{x_s^{-i} \in X_s^{-i}} \left[\max_{x_s^i \in X_s^i} \sum_{t \in S} p_{s, (x_s^i, x_s^{-i})}(t) v_t^i \right] \\ &= v_s^i, \end{aligned}$$

where the last equality is due to (2). □

The next corollary derives an important property of the minmax-levels on segments.

Corollary 8 *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. $v_s^i = v_t^i (= v_F^i)$ for all $s, t \in F$.*

Proof Take a player i and let $\alpha^i := \max_{s \in F} v_s^i$, which is the highest minmax-level on F for player i . Let $U = \{s \in F | v_s^i = \alpha^i\}$. We need to show that $U = F$. Suppose by way of contradiction that $U \subsetneq F$. Since all states within segment F communicate through joint actions $a_s \in \bar{A}_s$, $s \in F$, there must exist a state $s \in U$ and a joint action $a_s \in \bar{A}_s$ such that a_s induces a transition to $F - U$ with a positive probability. But then

$$\sum_{t \in S} p_{sa_s}(t) v_t^i < v_s^i,$$

which contradicts part 1 of Corollary 7. Hence, $U = F$. □

As an illustration, consider the simple product-game in Example 2. In view of (6) and (7), we find, in accordance with the corollary above, that both players minmax-levels are constant on each segment.

Lemma 9 *Let G be a simple product-game. Take an arbitrary player i .*

- (1) *Consider a cyclic set $T_{ki}^i(m)$ of some maximal communicating set E_{ki}^i for player i . Then, for any two states $s^i, t^i \in T_{ki}^i(m)$ and any joint state $s^{-i} \in S^{-i}$ of players $-i$, the minmax-level of player i satisfies $v_{(s^i, s^{-i})}^i = v_{(t^i, s^{-i})}^i$.*
- (2) *Consider a cyclic set $T_{kj}^j(m^j)$ of some maximal communicating set E_{kj}^j for all players $j \neq i$. Then, for any two joint states $s^{-i}, t^{-i} \in \times_{j \neq i} T_{kj}^j(m^j)$ and any state $s^i \in S^i$ of player i , the minmax-level of player i satisfies $v_{(s^i, s^{-i})}^i = v_{(s^i, t^{-i})}^i$.*

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

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

Take any state $u^i \in T_{ki}^i(m - 1)$, with $T_{ki}^i(0) := T_{ki}^i(\lambda_{ki}^i)$, and an action $a_{u^i}^i \in \bar{A}_{u^i}^i$ such that $a_{u^i}^i$ 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 5 yields

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

As action $a_{u^i}^i$ in state u^i only induces transition to states in $T_{ki}^i(m)$, by the choice of s^i , we have $v_{(s^i, s^{-i})}^i = v_{(w^i, s^{-i})}^i$ for all $w^i \in S^i$ for which $p_{u^i a_{u^i}^i}^i(w^i) > 0$. In particular, $v_{(s^i, s^{-i})}^i = v_{(t^i, s^{-i})}^i$, which completes the proof. □

Recall that, in simple product-games, z_F^i denotes the unique payoff for player i in the restricted game \bar{G}_F . When $z_F^i \geq v_F^i$ for every player i , i.e. when the restricted game \bar{G}_F is satisfactory for 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 $z_F^i < v_F^i$ for some player i , i.e. when the restricted game \bar{G}_F is unsatisfactory for player i . The next lemma proposes a way for player i to exit \bar{G}_F , by playing a certain action a_s^i in one of the cyclic sets of F . A similar result holds for players $-i$.

Lemma 10 *Let G be a simple product-game, and F a segment within some $E_k = \times_{i=1}^n E_{ki}^i$ for some $k = (k^1, \dots, k^n) \in K$. Suppose that F has cyclic sets of the form (cf. 3)*

$$T_F(m) := T_{k^1}^1(m) \times \dots \times T_{k^n}^n(m), \quad m = 1, \dots, \lambda_k.$$

Consider player i . Let z_F^i denote player i 's unique payoff in the restricted game \bar{G}_F , and v_F^i be player i 's minmax-level on F in the game G (a constant, cf. Corollary 8).

(1) *Suppose $z_F^i < v_F^i$. Then, for player i , there exists an $m \in \{1, \dots, \lambda_k\}$, a state $s^i \in T_{k^i}^i(m)$, and an action $a_{s^i}^i \in A_{s^i}^i - \bar{A}_{s^i}^i$ in state s^i such that if player i plays action $a_{s^i}^i$ 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 of the actions played by players $-i$. More precisely, for any $a_{s^{-i}}^{-i} \in A_{s^{-i}}^{-i}$ we have*

$$\sum_{t \in S} p_{s, (a_{s^i}^i, a_{s^{-i}}^{-i})}(t) v_t^i \geq v_F^i.$$

(2) *Suppose $z_F^i > v_F^i$. Then, for players $-i$, there exists an $m \in \{1, \dots, \lambda_k\}$, a joint state $s^{-i} \in T_{k^{-i}}^{-i}(m)$, and a joint action $a_{s^{-i}}^{-i} \in A_{s^{-i}}^{-i} - \bar{A}_{s^{-i}}^{-i}$ (i.e. at least one player $j \neq i$ plays outside $\bar{A}_{s^{-i}}^j$) in joint state s^{-i} such that if players $-i$ play joint action $a_{s^{-i}}^{-i}$ 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 of the action played by player i . More precisely, for any $a_s^i \in A_s^i$ we have*

$$\sum_{t \in S} p_{s, (a_s^i, a_{s^{-i}}^{-i})}(t) v_t^i \leq v_F^i.$$

The proof is analogous to the proof of lemma 3.6 for aperiodic product-games in Flesch et al. (2008a). We also refer to Flesch et al. (2008b). Here, we provide only a brief outline for part 1; the proof of part 2 is similar. In part 1, due to $z_F^i < v_F^i$, when starting in segment F , player i can only defend his minmax-level v_F^i if he leaves F . Therefore, there must be at least one state $s^* \in F$, joint action $a_{s^*}^{-i} \in \bar{A}_{s^*}^{-i}$ for players $-i$ and action $a_{s^*}^i \in A_{s^*}^i - \bar{A}_{s^*}^i$ for player i such that

$$\sum_{t \in S} p_{s^*, (a_{s^*}^i, a_{s^*}^{-i})}(t) v_t^i \geq v_F^i.$$

Take the unique $m \in \{1, \dots, \lambda_F\}$ for which $s^* \in T_F(m)$. One can show that this m , state s^{*i} and action $a_{s^*}^i$ satisfy part 1 of the lemma. Notice that, as $a_{s^*}^i$ guarantees for

player i in state s^* a minmax-level of at least v_F^i against $a_{s^*}^{-i} \in \bar{A}_{s^*}^{-i}$, based on part 2 of corollary 6, it cannot do worse against other joint actions of players $-i$ in state s^* . It is not very difficult to verify this for all states of the form $s = (s^{*i}, s^{-i}) \in T_F(m)$.

As an illustration for part 1 of Lemma 10, we revisit the simple product-game in Example 2. Consider the segment $F_{(I,II)} = \{(1, 2), (1, 3)\}$, where $v_{F_{(I,II)}}^1 = 1 > 0 = z_{F_{(I,II)}}^1$, by (6). 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 $v_{(2,2)}^1 = 1$. Note that $v_{(2,2)}^1 \geq v_{(1,3)}^1$, in accordance with part 1 of the lemma. Thus, one can choose $m = 2$, $s^1 = 1$ and $a_{s^1}^1 = 2$.

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 \tilde{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 and Vohra (2002).

Lemma 11 *Take an arbitrary product-game G , with v_s^i denoting the minmax-level for every player i in every state $s \in S$. Let \tilde{v}_F^i denote player i 's minmax-level in every restricted game \tilde{G}_F (which is constant, cf. lemma 3). Let \tilde{G} denote the simple product-game which is derived from G by replacing every player i 's payoffs in every restricted game \tilde{G}_F by his minmax-level \tilde{v}_F^i . Further, let w_s^i denote every player i 's minmax-level in \tilde{G} in state s .*

Then, the minmax-levels of the product-games G and \tilde{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 analogous to the proof of lemma 3.6 in Flesch et al. (2008a). We also refer to Flesch et al. (2008b). Here, we provide only an outline of the proof. 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

$$\gamma_s^i(\pi^i, x^{-i}) \leq w_s^i. \tag{10}$$

By applying Corollary 8 to the game \tilde{G} , we deduce that the minmax-level w_s^i equals some constant w_F^i on every segment F . We construct the joint stationary strategy x^{-i} by distinguishing 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 (2) for player i 's minmax-level w^i in \tilde{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 \tilde{G}_F (which is a part of the original game G) as in lemma 3.

Case 3: In a segment F with $\bar{v}_F^i > w_F^i$: In this case, part 2 of Lemma 10 (for the game \tilde{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 action 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 \tilde{G}_F . Observe that (1) the minmax-level w^i cannot increase in expectation until settling in \tilde{G}_F (in case 1 by the definition of x^{-i} , while in cases 2 and 3 by part 2 of Corollary 7 and part 2 of Lemma 10); 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 $\bar{v}_F^i \leq w_F^i$. Combining these two observations, it follows easily that player i 's reward is at most w_s^i in expectation, proving (10).

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 \tilde{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_s^2 for all initial states $s \in S$. Recall from (7) and (6) that

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

and from (5) and (4) that player 2's minmax-levels within the restricted games are

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

Following the proof, as the segments $F(I,I)$, $F(I,II)$ and $F(II,II),1$ and $F(II,II),2$ all belong to case 2 (i.e. $\bar{v}_F^2 \leq w_F^2$ 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}_F^2 . Also, y^1 has to leave $F(I,II)$, belonging to case 3 (i.e. $\bar{v}_{F(I,II)}^2 > w_{F(I,II)}^2$). 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.

Lemma 11 (and its proof) has useful consequences.

Corollary 12 *The results of Lemmas 5 and 6, Corollaries 7 and 8, and Lemma 9 are valid for any general product-game G . Lemma 10 extends as well if one interprets z_F^i as the minmax-level \bar{v}_F^i of player i in the restricted game \tilde{G}_F .*

Also, in view of the joint stationary strategy x^{-i} in the proof of Lemma 11, the infimum in expression (1) of the minmax-levels is attained at stationary strategies, for all product-games. This is stated next.

Corollary 13 (of the proof of lemma 11) *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 Theorem 2, which claimed that, in every two-player zero-sum product-game, both players have a stationary 0-optimal strategy.

Proof of Theorem 2 Take an arbitrary two-player zero-sum product-game, and take player $i = 1$. By corollary 13, 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.

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. \square

4.4 The construction of 0-equilibria in product-games

This section is devoted to the proof of theorem 1. The proof is analogous to the proof for the aperiodic case in Flesch et al. (2008a). We also refer to Flesch et al. (2008b).

We provide a general idea of the construction of a 0-equilibrium η . The equilibrium η will prescribe that all players follow a joint strategy π , unless some player i deviates 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} as in Corollary 13 and push down player i 's reward to his minmax-level v_s^i . In fact, y^{-i} acts as a threat strategy, whose task is to force player i to follow the prescriptions of π^i . 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 also shows a number of similarities with the construction in Vieille (2000a,b). The joint strategy π prescribes play in the following way:

(1) When entering some segment F , with $F \in \mathcal{F}^*$ (i.e. the corresponding restricted game \tilde{G}_F is satisfactory for all players): In this case, π will prescribe to stay on F and play an equilibrium in \tilde{G}_F as in Lemma 4. Here, the players have no reason to leave and collect "high" payoffs. (Cf. solvable sets in Vieille 2000a,b).

(2) When entering some segment F , with $F \in \mathcal{F}[i]$ (i.e. the corresponding restricted game \tilde{G}_F is unsatisfactory for player i): In this case, π will prescribe for players $-i$

to play a joint stationary strategy as in Lemma 3, whereas for player i to leave E_{ki}^i (and thereby to leave F) through moving to his “exit” state and subsequently to play his “exit” action provided by part 1 of 10. Note that player i is happy to leave F as he receives at most \bar{v}_F^i if he stays in F due to players $-i$ strategies, while $\bar{v}_F^i < v_F^i$ as $F \in \mathcal{F}[i]$. (Cf. controlled sets in Vieille 2000a,b).

(3) Outside all joint maximal communicating sets: In these states, which are transient, π will let the players play for their future perspectives, given their behavior in the segments according to (1) and (2) above.

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

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 (6) and (7), we have that

$$v_{F(I,I)}^1 = 1, \quad v_{F(I,II)}^1 = 1, \quad v_{F(II,I)}^1 = 0, \quad v_{F(II,II),1}^1 = 1, \quad v_{F(II,II),2}^1 = 0$$

and $v_F^2 = -v_F^1$ for all segments F . Recall from (4) and (5) that the players minmax-levels within the restricted games are

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

and $\bar{v}_F^2 = -\bar{v}_F^1$ 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 \tilde{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.

Appendix

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

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' . This lemma is hardly surprising as, in a product-game, each player controls play on his own coordinate only.

Lemma 14 *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*

$$\mathbb{P}_{(s^{N'}, t^{N-N'}, \sigma)}(h_m^{N'}) = \mathbb{P}_{s\pi}(h_m^{N'}) \tag{11}$$

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 $1, \dots, m - 1$ and by $\pi^{N'}$ at stage m .

Proof The construction of σ^i for each player $i \in N'$ is simple. Consider some current state $u_m \in S$ at stage m and some history $h_m = (s_1, a_1, \dots, s_{m-1}, a_{m-1})$. If the probability that $h_m^{N'}$ occurs and $u_m^{N'}$ becomes the current 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

$$\sigma_{u_m}^i(h_m)(a_m^i) := \sum_{\tilde{h}_m, \tilde{u}_m} \mathbb{P}_{s\pi}(\tilde{h}_m, \tilde{u}_m | h_m^{N'}, u^{N'}) \cdot \pi_{\tilde{u}_m}^i(\tilde{h}_m)(a_m^i) = \mathbb{P}_{s\pi}(a_m^i | h_m^{N'}, u_m^{N'}); \tag{12}$$

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 current state $u_m^{N'}$ for players N' .

Given $\sigma^{N'} = (\sigma^i)_{i \in N'}$, one can show (11) by using induction on m . □

Acknowledgments We would like to thank Rakesh V. Vohra and two anonymous referees for their suggestions, which considerably improved the presentation of the article.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Altman E, Avrachenkov K, Marquez R, Miller G (2005) Zero-sum constrained stochastic games with independent state processes. *Math Methods Oper Res* 62:375–386

Altman E, Avrachenkov K, Bonneau N, Debbah M, El-Azouzi R, Sadoc Menasche D (2007a) Constrained cost-coupled stochastic games with independent state processes (preprint)

Altman E, Avrachenkov K, Garnaeu A (2007b) A jamming game in wireless networks with transmission cost. *Lect Notes Comp Sci* 4465:1–12

Becker R, Zilberstein S, Lesser V, Goldman CV (2003) Transition-independent decentralized Markov decision processes. In: *Proceedings of the second international joint conference on autonomous agents and multi agent systems*, ACM Press, Melbourne, pp 41–48

Blackwell D, Ferguson TS (1968) The big match. *Ann Math Stat* 39:159–163

Dutta PK (1995) A Folk theorem for stochastic games. *J Econ Theory* 66:1–32

Filar JA, Vrieze OJ (1996) *Competitive Markov decision processes—theory, algorithms, and applications*. Springer, New York

Fink AM (1964) Equilibrium in a stochastic n -person game. *J Sci Hiroshima Univ Ser A–I* 28:89–93

- Flesch J, Thuijsman F, Vrieze OJ (1997) Cyclic Markov equilibria in a cubic game. *Int J Game Theory* 26:303–314
- Flesch J, Thuijsman F, Vrieze OJ (2007) Stochastic games with additive transitions. *Eur J Oper Res* 179:483–497
- Flesch J, Schoenmakers G, Vrieze OJ (2008a) Stochastic games on a product state space. *Math Oper Res* 33:403–420
- Flesch J, Schoenmakers G, Vrieze OJ (2008b) Stochastic games on a product state space: the periodic case. Research memorandum RM/08/016, University of Maastricht, The Netherlands
- 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, Princeton, pp 179–187
- Mertens JF, Neyman A (1981) Stochastic games. *Int J 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, Dordrecht, pp 173–193
- Parthasarathy T, Raghavan TES (1981) An orderfield property for stochastic games when one player controls transition probabilities. *J Optim Theory Appl* 33:375–392
- Ross KW, Varadarajan R (1991) Multichain Markov decision processes with a sample path constraint: a decomposition approach. *Math Oper Res* 16:195–207
- Simon R (2003) Value and perfection in stochastic games. Working paper, Goettingen University, Germany. *Israel J Math* (to appear)
- Solan E (1999) Three-Player absorbing games. *Math Oper Res* 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, Dordrecht, pp 265–280
- Solan E, Vieille N (2002) Correlated equilibrium in stochastic games. *Games Econ Behav* 38:362–399
- Solan E, Vohra R (2002) Correlated equilibrium and public signalling in absorbing games. *Int J Game Theory* 31:91–121
- Sorin S (1986) Asymptotic properties of a non-zero-sum game. *Int J Game Theory* 15:101–107
- Takahashi M (1964) Equilibrium points of stochastic noncooperative n -person games. *J Sci Hiroshima Univ Ser 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, Dordrecht, pp 85–100
- Vieille N (2000a) Equilibrium in 2-person stochastic games I: a reduction. *Israel J Math* 119:55–91
- Vieille N (2000b) Equilibrium in 2-person stochastic games II: the case of recursive games. *Israel J Math* 119:93–126