

On pure stationary almost Markov Nash equilibria in nonzero-sum ARAT stochastic games

Anna Jaśkiewicz · Andrzej S. Nowak

Received: 30 April 2014 / Accepted: 22 December 2014 / Published online: 4 January 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract We deal with discounted ARAT stochastic games on a Borel state space with finite action spaces and nonatomic transition probabilities. We prove the existence of pure Nash equilibria in stationary almost Markov strategies that depend only on the current and previous state of the game. Our proof is based on an existence theorem for correlated equilibria in stochastic games and some results on the integrals of set-valued mappings with respect to a probability measure depending on a parameter.

Keywords Nonzero-sum stochastic game · Borel state space · Additive reward · Additive nonatomic transition probability · Pure stationary almost Markov equilibrium

1 Introduction

An integral analysis of set-valued mappings plays an important role in the study of various equilibria in nonzero-sum stochastic games with general state spaces; see [Mertens and Parthasarathy \(1991, 2003\)](#), [Nowak and Raghavan \(1992\)](#), [Jaśkiewicz and Nowak \(2005\)](#) and [Barelli and Duggan \(2014\)](#). In particular, certain results on

The authors gratefully acknowledge the financial support of the National Science Center under grant DEC-2011/03/B/ST1/00325.

A. Jaśkiewicz (✉)
Department of Mathematics, Wrocław University of Technology,
50-370 Wrocław, Poland
e-mail: anna.jaskiewicz@pwr.edu.pl

A. S. Nowak
Faculty of Mathematics, Computer Science and Econometrics,
University of Zielona Góra, Podgórna 50, 65-246 Zielona Góra, Poland
e-mail: a.nowak@wmie.uz.zgora.pl

the integrals with respect to a parametrized measure, proved by Artstein (1989) or Mertens (2003), are significant in showing the existence of Nash equilibria in general models of games in the class of randomised history dependent or semi-Markov strategies; see Mertens and Parthasarathy (1991, 2003), Barelli and Duggan (2014). In this paper, we focus on some special class of nonzero-sum stochastic games with Borel state space, finite action spaces and additive reward and transition structure, called in the sequel ARAT games. We are interested in Nash equilibria among pure stationary almost Markov strategies that depend, in each period of the game, on the current and previous state of the game. This class of strategies is a subclass of stationary semi-Markov strategies considered in Barelli and Duggan (2014). More precisely, they showed the existence of randomised Nash equilibria in the class of strategies, where dependence on the current, previous states and the previous actions is assumed. A stationary Markov Nash equilibrium consists of strategies that depend only on the current state. For convenience, they will be also called stationary. This term is common and has been used in a number of papers on stochastic games.

ARAT games with Borel state and finite action spaces were first studied by Himmelberg et al. (1976), who showed the existence of stationary Nash equilibria for p -almost all initial states. Their result was strengthened by Parthasarathy (1982), who obtained stationary Nash equilibria for all initial states. Pure stationary Markov Nash equilibria may not exist in ARAT stochastic games; see Example 3.1 (a game with 4 states) in Raghavan et al. (1985) or Example 4 (a game with 2 states) in Nowak (2006). The existence of pure stationary ϵ -equilibria in ARAT games was proved by Nowak (1987) under the assumption that the transition probabilities are nonatomic and the discounted payoffs are “averaged” with respect to a nonatomic distribution of the initial state. Markovian ϵ -equilibria in pure strategies for ARAT games can be shown to exist by using the backward induction in the finite horizon game; see Rieder (1979). Thuijsman and Raghavan (1997) (for finite state space) and Künle (1999) (for Borel state space and compact metric action spaces) established the existence of nonstationary history dependent pure Nash equilibria in ARAT games. Their proofs are based on the well-known idea of threats used frequently in repeated games. Stochastic games with average payoffs, additive transitions and finite state and action spaces were studied by Flesch et al. (2007), where some results on ϵ -equilibria were given in the class of randomised strategies.

In this paper, we study pure strategies in ARAT games, which are special cases of stationary semi-Markov ones introduced recently by Barelli and Duggan (2014). Assuming that the action spaces are finite and the transition probabilities are dominated by a nonatomic probability measure on a Borel state space, we prove the existence of pure Nash equilibria in stationary almost Markov strategies that depend only on the current and previous state of the game. Similarly as in Barelli and Duggan (2014), our proof is a combination of some arguments used by Nowak and Raghavan (1992) to study correlated equilibria and a measurable selection theorem for parametrized set-valued integrals due to Mertens (2003). Our main result on pure Nash equilibria in the class of above-mentioned strategies contributes to the literature on ARAT games. This result, however, cannot be extended to ARAT stochastic games with transition probabilities involving atoms. We give an example with two states where pure station-

ary almost Markov Nash equilibria do not exist. This fact shows that our assumption on nonatomic transitions is indeed essential.

2 The model and main result

We consider a two-person nonzero-sum discounted stochastic game G with additive rewards and additive transitions (an ARAT game for short) for which:

- (i) (S, \mathcal{B}) is a nonempty Borel state space with its Borel σ -algebra \mathcal{B} .
- (ii) $A = \{1, 2, \dots, n_1\}$ and $B = \{1, 2, \dots, n_2\}$ are action spaces for players 1 and 2, respectively.
- (iii) $A(s) \subset A, B(s) \subset B$ are nonempty sets of actions available to player 1 and 2 in state $s \in S$. Assume that the set-valued mappings $s \mapsto A(s)$ and $s \mapsto B(s)$ are lower measurable. Define

$$D = \{(s, a, b) : s \in S, a \in A(s), b \in B(s)\}.$$

Then D is a Borel subset of $S \times A \times B$.

- (iv) Let $u_n : S \times A \mapsto \mathbb{R}$ and $w_n : S \times B \mapsto \mathbb{R}$ be bounded Borel measurable functions for $n = 1, 2$. The reward (or payoff) function for player $n = 1, 2$ is given by

$$r_n(s, a, b) = u_n(s, a) + w_n(s, b), \text{ where } (s, a, b) \in D.$$

- (v) $q : D \times \mathcal{B} \mapsto [0, 1]$ is a transition probability such that

$$q(\cdot | s, a, b) = q_1(\cdot | s, a) + q_2(\cdot | s, b) \text{ for each } (s, a, b) \in D$$

and for some Borel measurable subtransition probabilities q_1 and q_2 . We assume that there exists a nonatomic probability measure μ on (S, \mathcal{B}) such that $q(\cdot | s, a, b) \ll \mu$ for all $(s, a, b) \in D$.

- (vi) $\beta \in (0, 1)$ is a discount factor.

These above components describe a discrete-time dynamic game in which each period $t \in \mathbb{N}$ begins with a state $s_t \in S$, and after observing s_t , the players simultaneously choose their actions $a_t \in A(s_t), b_t \in B(s_t)$ and obtain rewards $r_1(s_t, a_t, b_t)$ and $r_2(s_t, a_t, b_t)$. A new state s_{t+1} , is realised from the distribution $q(\cdot | s_t, a_t, b_t)$ and new period begins with rewards discounted by β . The game is played with past history $h_t = (s_1, a_1, b_1, \dots, a_{t-1}, b_{t-1}, s_t)$ as common knowledge for both players, where s_k is the state in the k -th period of the game, $a_k \in A(s_k)$ and $b_k \in B(s_k)$ are the actions taken by the players at period $k = 1, \dots, t, t \in \mathbb{N}$. In this paper, we are interested in Nash equilibria in pure strategies. Therefore, we do not define randomised strategies. A *pure strategy* for player 1 (2) is a sequence $\pi = (\pi_t)$ ($\sigma = (\sigma_t)$) of Borel measurable mappings, where each π_t (σ_t) associates with each given history h_t an action $a_t \in A(s_t)$ ($b_t \in B(s_t)$). Let F_1 (F_2) be the set of all Borel measurable functions $f : S \times S \mapsto A$ ($g : S \times S \mapsto B$) such that $f(s, s') \in A(s')$ ($g(s, s') \in B(s')$) for

each $s, s' \in S$. A *pure stationary almost Markov strategy* π for player 1 is such that for some $f \in F_1$ we have $\pi_1(s_1) = f(s_1, s_1)$ for all $s_1 \in S$ and $\pi_t(h_t) = f(s_{t-1}, s_t)$ for every h_t and $t \geq 2$. In other words, a pure stationary almost Markov strategy for player 1 may only depend on the previous and current state for any $t \geq 2$. We will identify any pure stationary almost Markov strategy for player 1 with $f \in F_1$. Similarly, we define pure stationary almost Markov strategies for player 2 and identify them with Borel measurable mappings $g \in F_2$. We would like to indicate that F_1 and F_2 are special cases of the classes of stationary semi-Markov strategies considered in [Barelli and Duggan \(2014\)](#), where dependence of π_t (or σ_t) on the current state, previous state and previous actions is assumed. Furthermore, a pure stationary almost Markov strategy is called *Markov*, if it is independent of the previous state. The set of all strategies is denoted by Π for player 1 and by Σ for player 2.

For any strategies $\pi \in \Pi$ and $\sigma \in \Sigma$ we define the expected discounted payoff or reward function for player n :

$$J_n(s, \pi, \sigma) = E_s^{\pi\sigma} \left(\sum_{k=1}^{\infty} \beta^{k-1} r_n(s_k, a_k, b_k) \right),$$

where $E_s^{\pi\sigma}$ is the expectation operator corresponding to the unique probability measure $P_s^{\pi\sigma}$ defined on the space of all feasible infinite histories of the process starting in state $s = s_1 \in S$ and induced by the transition probability q and strategies π and σ .

A pair of strategies $(\pi^*, \sigma^*) \in \Pi \times \Sigma$ is called a *Nash equilibrium*, if

$$J_1(s, \pi^*, \sigma^*) \geq J_1(s, \pi, \sigma^*) \text{ for all } \pi \in \Pi$$

and

$$J_2(s, \pi^*, \sigma^*) \geq J_2(s, \pi^*, \sigma) \text{ for all } \sigma \in \Sigma.$$

In the sequel, we shall refer to the game \tilde{G} with the state space $S \times S$ and action spaces $\tilde{A}(s, s') = A(s')$, $\tilde{B}(s, s') = B(s')$ for all $(s, s') \in S \times S$. The reward functions in game \tilde{G} , denoted by \tilde{r}_n , are defined as follows

$$\tilde{r}_n((s, s'), a, b) := r_n(s', a, b) \text{ for all } (s, s') \in S \times S, a \in A(s'), b \in B(s').$$

The transition probability \tilde{q} in game \tilde{G} is defined as follows

$$\tilde{q}(C_1 \times C_2 | (s, s'), a, b) := \delta_{s'}(C_1)q(C_2 | s', a, b)$$

for all $(s, s') \in S \times S, a \in A(s'), b \in B(s')$ and $C_1, C_2 \in \mathcal{B}$. Here $\delta_{s'}$ denotes the Dirac measure concentrated at s' . Hence, $\tilde{q}(\{s'\} \times C_2 | (s, s'), a, b) := q(C_2 | s', a, b)$.

Pure strategies are defined in game \tilde{G} in an obvious manner. Note that each strategy $f \in F_1$ or $g \in F_2$ in game G is stationary Markov in game \tilde{G} . The discounted payoffs for player $n = 1, 2$ is denoted by $\tilde{J}_n((s, s'), \pi, \sigma)$. Note that if $s = s'$, then

$$J_n(s', f, g) = \tilde{J}_n((s, s'), f, g) \text{ for all } f \in F_1, g \in F_2, s' \in S, n = 1, 2. \tag{1}$$

We can now formulate our main result.

Theorem *Every ARAT game G satisfying assumptions (i)–(vi) has a Nash equilibrium in pure stationary almost Markov strategies.*

Proof Let $B(S \times S)$ be the space of all bounded Borel measurable real-valued functions on $S \times S$. For any $s \in S$ and $v = (v_1, v_2)$, where $v_1, v_2 \in B(S \times S)$, we consider a static game $\Gamma_v(s)$ where the payoff to player $n = 1, 2$ is given by

$$U_n^v(s, a, b) := r_n(s, a, b) + \beta \int_S v_n(s, s')q(ds'|s, a, b), \tag{2}$$

where $a \in A(s)$, $b \in B(s)$. We shall also consider the payoff functions of the form (2) where v_1 and v_2 depend only on $s' \in S$. Let $N_v(s)$ be the set of all pure Nash equilibria in the game $\Gamma_v(s)$. Under the ARAT assumption $N_v(s) \neq \emptyset$. Indeed, $(a_0, b_0) \in N_v(s)$ if and only if

$$a_0 \in \arg \max_{a \in A(s)} \left[u_1(s, a) + \beta \int_S v_1(s, s')q_1(ds'|s, a) \right]$$

and

$$b_0 \in \arg \max_{b \in B(s)} \left[w_2(s, b) + \beta \int_S v_2(s, s')q_2(ds'|s, b) \right].$$

By $P_v(s)$ [$coP_v(s)$] we denote the set of all payoff vectors [convex combinations of payoff vectors] corresponding to equilibria in $N_v(s)$. Let $B(S)$ be the space of all bounded Borel measurable real-valued functions on S . A simple adaptation of the arguments given in Nowak and Raghavan (1992) (see also page 35 in Jaskiewicz and Nowak (2005)) yields the existence of some $w^* = (w_1^*, w_2^*)$ with $w_n^* \in B(S)$ for $n = 1, 2$ such that

$$w^*(s) \in coP_{w^*}(s) \text{ for all } s \in S.$$

Assume that $A \times B$ is endowed with the lexicographic order and every $A(s) \times B(s)$ is given the induced order. Write $A \times B = \{p_1, p_2, \dots, p_d\}$ with $d = n_1n_2$. Define

$$q_0(C|s, p) := \begin{cases} q(C|s, p), & \text{if } p \in A(s) \times B(s) \\ \mu(C), & \text{if } p \notin A(s) \times B(s), \end{cases}$$

where $C \in \mathcal{B}$. Define the d -dimensional stochastic kernel

$$K(\cdot|s) := (q_0(\cdot|s, p_1), q_0(\cdot|s, p_2), \dots, q_0(\cdot|s, p_d)).$$

Let $w_0^*(s, p) := \int_S w^*(s')q_0(ds'|s, p)$, $p \in A \times B$. Put

$$\bar{w}_0^*(s) := \int_S w^*(s')K(ds'|s) = (w_0^*(s, p_1), w_0^*(s, p_2), \dots, w_0^*(s, p_d)).$$

The set-valued mappings $s \mapsto P_{w^*}(s)$ and $s \mapsto coP_{w^*}(s)$ are lower measurable (see Lemma 6 in Nowak and Raghavan (1992)). Let Ψ be the graph of the mapping $s \mapsto \int_S P_{w^*}(s')K(ds'|s)$, i.e.,

$$\Psi := \left\{ (s, y) \in S \times \mathbb{R}^{2d} : s \in S, y \in \int_S P_{w^*}(s')K(ds'|s) \right\}.$$

By part 3 in Theorem 2 of Mertens (2003), there exists a Borel measurable mapping $\phi : \Psi \times S \mapsto \mathbb{R}^2$ such that $\phi(s, y, s') \in P_{w^*}(s')$ for all $(s, y) \in \Psi, s \in S$ and

$$y = \int_S \phi(s, y, s')K(ds'|s), \quad s \in S. \tag{3}$$

From Lyapunov’s theorem (see Corollary 18.1.10 in Klein and Thompson (1984)), it follows that $\int_S P_{w^*}(s')K(ds'|s) = \int_S coP_{w^*}(s')K(ds'|s)$. Hence, $\bar{w}_0^*(s) \in \int_S P_{w^*}(s')K(ds'|s)$ for each $s \in S$. Put

$$v^*(s, s') := \phi(s, \bar{w}_0^*(s), s'), \quad s, s' \in S.$$

Clearly $v^* \in B(S \times S)$. By (3), we have that

$$\bar{w}_0^*(s) = \int_S v^*(s, s')K(ds'|s) = \int_S w^*(s')K(ds'|s) \text{ for every } s \in S, s' \in S.$$

This fact implies that

$$\int_S v^*(s, s')q(ds'|s, a, b) = \int_S w^*(s')q(ds'|s, a, b) \text{ for every } (s, a, b) \in D. \tag{4}$$

Put $v^* = (v_1^*, v_2^*)$ and let s^- denote the previous state. Since $v^*(s, s') \in P_{w^*}(s')$ for every $s, s' \in S$ or equivalently, $v^*(s^-, s) \in P_{w^*}(s)$ for every $s^-, s \in S$, by Filippov’s implicit function theorem (see Theorem 18.17 in Aliprantis and Border (2006) or Lemma 4 in Nowak and Raghavan (1992)), there exists a pair $(f^*, g^*) \in F_1 \times F_2$ such that

$$v_1^*(s^-, s) = U_1^{w^*}(s, f^*(s^-, s), g^*(s^-, s)) = \max_{a \in \tilde{A}(s^-, s)} U_1^{w^*}(s, a, g^*(s^-, s)) \tag{5}$$

and

$$v_2^*(s^-, s) = U_2^{w^*}(s, f^*(s^-, s), g^*(s^-, s)) = \max_{b \in \tilde{B}(s^-, s)} U_2^{w^*}(s, f^*(s^-, s), b) \tag{6}$$

for all $s^-, s \in S$. By (4), we have $U_n^{w^*}(s, a, b) = U_n^{v^*}(s, a, b)$ for each $(s, a, b) \in D$ and $n = 1, 2$. Thus, from (5) and (6), we conclude that

$$v_1^*(s^-, s) = U_1^{v^*}(s, f^*(s^-, s), g^*(s^-, s)) = \max_{a \in \tilde{A}(s^-, s)} U_1^{v^*}(s, a, g^*(s^-, s))$$

and

$$v_2^*(s^-, s) = U_2^{v^*}(s, f^*(s^-, s), g^*(s^-, s)) = \max_{b \in \tilde{B}(s^-, s)} U_2^{v^*}(s, f^*(s^-, s), b)$$

for every $s^-, s' \in S$. We have obtained two Bellman equations (for players 1 and 2) in the game \tilde{G} . By standard dynamic programming arguments (see Blackwell (1965)), these equations imply that

$$\tilde{J}_1((s^-, s), f^*, g^*) = \max_{f \in F_1} \tilde{J}_1((s^-, s), f, g^*) \tag{7}$$

and

$$\tilde{J}_2((s^-, s), f^*, g^*) = \max_{g \in F_2} \tilde{J}_2((s^-, s), f^*, g) \tag{8}$$

for all $s^-, s \in S$. Putting $s^- = s$ in (7), (8) and using (1), we obtain

$$J_1(s, f^*, g^*) = \max_{f \in F_1} J_1(s, f, g^*), \quad J_2(s, f^*, g^*) = \max_{g \in F_2} J_2(s, f^*, g)$$

for each $s \in S$. These equations and standard dynamic programming arguments (see Blackwell (1965)) imply that (f^*, g^*) is a Nash equilibrium in the class of all strategies of the players. □

Remark 1 From the proof, it follows that (f^*, g^*) is subgame perfect in the sense of Selten (1975). Pure stationary Markov Nash equilibria can be shown to exist in the same manner as in Theorem 2 in Nowak (2006), if the transition probability is a convex combination of finitely many nonatomic measures on S . The existence of pure stationary Markov Nash equilibria under assumptions made in this paper (i.e. with non-atomic additive transitions and additive rewards) is an open problem. Studying stochastic ARAT games K uenle (1999) did not assume that the transition probability is dominated by any probability measure. He obtained a pure Nash equilibrium (π^*, σ^*) where each π_t^* and σ_t^* depends on the entire history h_t . We assume the dominance of q with respect to an nonatomic measure μ and obtain an equilibrium in the simplest possible class of strategies. Our result cannot be extended to the case where μ has some atoms, which is illustrated in the next section. The survey of the existing results on randomised Nash equilibria in stochastic games without ARAT structure the reader may find in Nowak (2003), Ja skiewicz and Nowak (2005) and Barelli and Duggan (2014). Related result to our theorem on randomised equilibria in games with finite state independent action sets is mentioned on page 147 in Mertens and Parthasarathy (1991).

Remark 2 The Nash equilibrium strategy for each player considered in this paper is called ‘‘stationary’’, since it is determined by a single function independent of calendar time. The term ‘‘almost Markov’’, on the other hand, refers to the property that a strategy does not depend only on the current state (at any stage $t \geq 2$), but also on the previous state. The fact that our strategies in Nash equilibrium depend on the current and previous state follows from applying a parametrised version of Lyapunov’s theorem given by Mertens (2003).

3 A counterexample

Below we give an example of a stochastic ARAT game with finite state and action spaces having no pure stationary almost Markov Nash equilibrium.

Let $S = \{1, 2\}$, $A(1) = B(2) = \{1, 2\}$, $A(2) = B(1) = \{1\}$. Assume that $r_1(1, a, 1) = 0$ for $a \in A(1)$ and $r_1(2, 1, b) = 6$ if $b = 1$, and $r_1(2, 1, b) = -6$ if $b = 2$. Let

$$r_2(1, a, 1) = u_2(1, a) = \begin{cases} 7 & \text{for } a = 1 \\ 0 & \text{for } a = 2 \end{cases}$$

and

$$r_2(2, 1, b) = w_2(1, b) = \begin{cases} 0 & \text{for } b = 1 \\ 1 & \text{for } b = 2. \end{cases}$$

We assume that the transition probability in state $s = 1$ is controlled by player 1 and in state $s = 2$ is controlled by player 2. We define

$$q(1|1, a, 1) = \begin{cases} \frac{2}{3} & \text{if } a = 1 \\ \frac{1}{3} & \text{if } a = 2 \end{cases} \quad \text{and} \quad q(1|2, 1, b) = \begin{cases} \frac{2}{3} & \text{if } b = 1 \\ \frac{1}{3} & \text{if } b = 2 \end{cases}$$

and $q(2|s, a, b) = 1 - q(1|s, a, b)$ for each $s \in S, a \in A(s), b \in B(s)$.

Let $\tilde{S} = \{s_1, s_2, s_3, s_4\}$ where $s_1 = (1, 1), s_2 = (1, 2), s_3 = (2, 1), s_4 = (2, 2)$. Note that any pure stationary almost Markov strategy for player 1 can be defined as $f_{ij}(s_1) = i, f_{ij}(s_2) = j$, and $f_{ij}(s_3) = f_{ij}(s_4) = 1$. Thus, player 1 has four pure stationary almost Markov strategies. A pure stationary almost Markov strategy for player 2 is denoted by g_{ij} , where $g_{ij}(s_3) = i, g_{ij}(s_4) = j$, and $g_{ij}(s_1) = g_{ij}(s_2) = 1$. In order to compute the discounted expected rewards to the players for any pair (f_{ij}, g_{kl}) of strategies we consider an auxiliary game \tilde{G} (defined in Sect. 2), in which f_{ij} and g_{kl} are pure stationary strategies (or pure stationary Markov strategies). For computational purposes, it is convenient to use the standard matrix notation that is common in the finite state space case. By $Q(f_{ij}, g_{kl})$ we denote the transition probability matrix induced by \tilde{q} and strategies f_{ij}, g_{kl} . We assume that the rows and columns of $Q(f_{ij}, g_{kl})$ are labeled by s_1, s_2, s_3 and s_4 . Let

$$\tilde{R}_n(f_{ij}, g_{kl}) := [\tilde{r}_n(s_1, f_{ij}, g_{kl}) \tilde{r}_n(s_2, f_{ij}, g_{kl}) \tilde{r}_n(s_3, f_{ij}, g_{kl}) \tilde{r}_n(s_4, f_{ij}, g_{kl})]^T$$

be the vector of rewards of player $n = 1, 2$ in the auxiliary game, induced by strategies f_{ij} and g_{kl} . By $\tilde{J}_n(s_m, f_{ij}, g_{kl})$, we denote the discounted expected payoff to player n in game Γ . Note that

$$\tilde{J}_n(s_1, f_{ij}, g_{kl}) = J_n(1, f_{ij}, g_{kl}) \quad \text{and} \quad \tilde{J}_n(s_4, f_{ij}, g_{kl}) = J_n(2, f_{ij}, g_{kl})$$

for each strategy pair (f_{ij}, g_{kl}) . For any player n , define

$$\tilde{J}_n(f_{ij}, g_{kl}) := [\tilde{J}_n(s_1, f_{ij}, g_{kl}) \tilde{J}_n(s_2, f_{ij}, g_{kl}) \tilde{J}_n(s_3, f_{ij}, g_{kl}) \tilde{J}_n(s_4, f_{ij}, g_{kl})]^T.$$

The standard formula yields that

$$\tilde{J}_n(f_{ij}, g_{kl}) = [I - \beta Q(f_{ij}, g_{kl})]^{-1} \tilde{R}_n(f_{ij}, g_{kl}).$$

If $[I - \beta Q(f_{ij}, g_{kl})]_m^{-1}$ is the m -th row of the matrix $[I - \beta Q(f_{ij}, g_{kl})]^{-1}$, then we have

$$J_n(1, f_{ij}, g_{kl}) = [I - \beta Q(f_{ij}, g_{kl})]_1^{-1} \tilde{R}_n(f_{ij}, g_{kl}) \tag{9}$$

and

$$J_n(2, f_{ij}, g_{kl}) = [I - \beta Q(f_{ij}, g_{kl})]_4^{-1} \tilde{R}_n(f_{ij}, g_{kl}). \tag{10}$$

For an illustration assume that $\beta = 3/4$ and consider the pair (f_{21}, g_{12}) . Then

$$\begin{aligned} \tilde{R}_1(f_{21}, g_{12}) &= [0 \ 0 \ 6 \ -6]^T, & \tilde{R}_2(f_{21}, g_{12}) &= [0 \ 7 \ 0 \ 1]^T, \\ Q(f_{21}, g_{12}) &= \frac{1}{3} \begin{bmatrix} 1 & 0 & 2 & 0 \\ 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 1 \\ 0 & 1 & 0 & 2 \end{bmatrix}, & I - \beta Q(f_{21}, g_{12}) &= \frac{1}{4} \begin{bmatrix} 3 & 0 & -2 & 0 \\ -2 & 4 & -1 & 0 \\ 0 & -2 & 4 & -1 \\ 0 & -1 & 0 & 2 \end{bmatrix}, \end{aligned}$$

and

$$[I - \beta Q(f_{21}, g_{12})]^{-1} = \frac{4}{61} \begin{bmatrix} 27 & 10 & 16 & 8 \\ 16 & 24 & 14 & 7 \\ 10 & 15 & 24 & 12 \\ 8 & 12 & 7 & 34 \end{bmatrix}.$$

Using these data and (9)–(10), we can easily obtain that

$$J_1(1, f_{21}, g_{12}) = 192/61, \quad J_1(2, f_{21}, g_{12}) = -648/61$$

and

$$J_2(1, f_{21}, g_{12}) = 312/61, \quad J_2(2, f_{21}, g_{12}) = 472/61.$$

We can similarly compute the discounted expected payoffs to the players for each pair (f_{ij}, g_{kj}) of strategies and consider two bimatrix games corresponding to states 1 and 2, respectively. The rows (columns) of the matrices given below are labeled by $f_{11}, f_{12}, f_{21}, f_{22}$ ($g_{11}, g_{12}, g_{21}, g_{22}$). The payoff matrices in state $s = 1$ are:

$$M_1(1) = \begin{bmatrix} 6 & \frac{24}{11} & -\frac{24}{17} & -8 \\ \frac{48}{7} & \frac{32}{13} & -\frac{96}{61} & -\frac{96}{11} \\ \frac{96}{11} & \frac{192}{61} & -\frac{192}{95} & -\frac{192}{17} \\ \frac{48}{5} & \frac{24}{7} & -\frac{24}{11} & -12 \end{bmatrix} \quad \text{and} \quad M_2(1) = \begin{bmatrix} 21 & \frac{228}{11} & \frac{348}{17} & 20 \\ 16 & \frac{632}{39} & \frac{1000}{61} & \frac{184}{11} \\ \frac{56}{11} & \frac{312}{61} & \frac{488}{95} & \frac{88}{17} \\ 0 & \frac{4}{7} & \frac{12}{11} & 2 \end{bmatrix}.$$

Observe that this bimatrix game has no pure Nash equilibrium. In state $s = 2$, the payoff matrices are:

$$M_1(2) = \begin{bmatrix} 12 & -\frac{120}{11} & \frac{120}{17} & -16 \\ \frac{96}{7} & -\frac{136}{13} & \frac{408}{61} & -\frac{192}{11} \\ \frac{144}{11} & -\frac{648}{61} & \frac{648}{95} & -\frac{288}{17} \\ \frac{72}{5} & -\frac{72}{7} & \frac{72}{11} & -18 \end{bmatrix} \quad \text{and} \quad M_2(2) = \begin{bmatrix} 14 & \frac{136}{11} & \frac{232}{17} & 12 \\ 4 & \frac{200}{39} & \frac{264}{61} & \frac{60}{11} \\ \frac{84}{11} & \frac{472}{61} & \frac{728}{95} & \frac{132}{17} \\ 0 & \frac{16}{7} & \frac{8}{11} & 3 \end{bmatrix}.$$

This bimatrix game has no pure Nash equilibrium either.

We conclude this section by pointing out that the above game has a randomised stationary Markov Nash equilibrium (f^*, g^*) where $f^*(1) = (\frac{1}{2}, \frac{1}{2})$ is a mixed strategy for player 1 in $s = 1$ and $g^*(2) = (\frac{5}{8}, \frac{3}{8})$ is a mixed strategy for player 2 in state $s = 2$.

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

References

- Aliprantis C, Border K (2006) *Infinite dimensional analysis: a hitchhiker's guide*. Springer, New York
- Artstein Z (1989) Parametrized integration of multifunctions with applications to control and optimization. *SIAM J Control Optim* 27:1369–1380
- Barelli P, Duggan J (2014) A note on semi-Markov perfect equilibria in discounted stochastic games. *J Econ Theory* 151:596–604
- Blackwell D (1965) Discounted dynamic programming. *Ann Math Stat* 36:226–235
- Flesch J, Thuijsman F, Vrieze OJ (2007) Stochastic games with additive transitions. *Eur J Oper Res* 179:483–497
- Himmelberg CJ, Parthasarathy T, Raghavan TES, Van Vleck FS (1976) Existence of p-equilibrium and optimal stationary strategies in stochastic games. *Proc Am Math Soc* 60:245–251
- Jaśkiewicz A, Nowak AS (2005) Nonzero-sum semi-Markov games with the expected average payoffs. *Math Methods Oper Res* 62:23–40
- Klein E, Thompson AC (1984) *Theory of correspondences*. Wiley, New York
- Küenle HU (1999) Equilibrium strategies in stochastic games with additive cost and transition structure and Borel state and action spaces. *Int Game Theory Rev* 1:131–147
- Mertens JF (2003) A measurable measurable choice theorem. In: Neyman A, Sorin S (eds) *Stochastic games and applications*. Kluwer, Dordrecht, pp 107–130
- Mertens JF, Parthasarathy T (1991) Nonzero-sum stochastic games. In: Raghavan et al (eds) *Stochastic games and related topics*. Kluwer, Dordrecht, pp 145–148
- Mertens JF, Parthasarathy T (2003) Equilibria for discounted stochastic games. In: Neyman A, Sorin S (eds) *Stochastic games and applications*. Kluwer, Dordrecht, pp 131–172
- Nowak AS (1987) Nonrandomized strategy equilibria in noncooperative stochastic games with additive transition and reward structure. *J Optim Theory Appl* 52:429–441
- Nowak AS (2003) N-person stochastic games: extensions of the finite state space case and correlation. In: Neyman A, Sorin S (eds) *Stochastic games and applications*. Kluwer, Dordrecht, pp 93–106
- Nowak AS (2006) Remarks on sensitive equilibria in stochastic games with additive reward and transition structure. *Math Methods Oper Res* 64:481–494
- Nowak AS, Raghavan TES (1992) Existence of stationary correlated equilibria with symmetric information for discounted stochastic games. *Math Oper Res* 17:519–526

- Parthasarathy T (1982) Existence of equilibrium stationary strategies in discounted stochastic games. *Sankhya Ser A* 44:114–127
- Raghavan TES, Tijs SH, Vrieze OJ (1985) On stochastic games with additive reward and transition structure. *J Optim Theory Appl* 47:451–464
- Rieder U (1979) Equilibrium plans for non-zero sum Markov games. In: Moeschlin O, Pallaschke D (eds) *Game theory and related topics*. North-Holland, Amsterdam, pp 91–102
- Selten R (1975) Re-examination of the perfectness concept for equilibrium points in extensive games. *Int J Game Theory* 4:25–55
- Thuijsman F, Raghavan TES (1997) Perfect information stochastic games and related classes. *Int J Game Theory* 26:403–408