CrossMark

On discounted AR–AT semi-Markov games and its complementarity formulations

P. Mondal · S. Sinha · S. K. Neogy · A. K. Das

Abstract In this paper, we introduce a class of two-person finite discounted AR–AT (Additive Reward–Additive Transition) semi-Markov games (SMGs). We provide counterexamples to show that AR–AT and AR–AT–PT (Additive Reward–Additive Transition Probability and Time) SMGs do not satisfy the ordered field property. Some results on AR–AT–AITT (Additive Reward–Additive Transition and Action Independent Transition Time) and AR–AIT–ATT (Additive Reward–Action Independent Transition and Additive Transition Time) games are obtained in this paper. For the zero-sum games, we prove the ordered field property and the existence of pure stationary optimals for the players. Moreover, such games are formulated as a vertical linear complementarity problem (VLCP) and have been solved by Cottle-Dantzig's algorithm under a mild assumption. We illustrate that the nonzero-sum case of such games do not necessarily have pure stationary equilibria. However, there exists a stationary equilibria which has at most two pure actions in each state for each player.

Keywords Semi-Markov games · AR–AT · Optimal/nash equilibrium stationary strategies, VLCP · Cottle-Dantzig's algorithm

P. Mondal (✉) · S. Sinha
Mathematics Department, Jadavpur University, Kolkata 700032, India
e-mail: prasenjit1044@yahoo.com

S. Sinha
e-mail: sagnik62@yahoo.co.in

S. K. Neogy
Indian Statistical Institute, Delhi Centre, New Delhi 110016, India
e-mail: skn@isid.ac.in

A. K. Das
Indian Statistical Institute, Kolkata Centre, Kolkata 700108, India
e-mail: akdas@isical.ac.in

1 Introduction

Stochastic (Markov) games (SGs) with additive rewards and additive transitions (AR–AT) have been widely studied in the literature (e.g., Himmelberg et al. 1976; Raghavan et al. 1985). Raghavan et al. (1985) proved that for zero-sum AR–AT stochastic games, the ordered field property holds and both players have pure optimal stationary strategies for the discounted as well as the undiscounted payoff criteria. However, for nonzero-sum games in this class, a pure stationary equilibria may not exist. Raghavan et al. (1985) suggested a finite step method which involves solving a finite number of Markov decision problems to compute a pair of pure stationary optimal strategies and the value of the zero-sum AR–AT stochastic games. Mohan et al. (1999) formulated such games as a single vertical linear complementarity problem (VLCP) which can be solved by Cottle-Dantzig's algorithm under a mild assumption.

In this paper, we introduce a class of two-person finite AR–AT semi-Markov games (SMGs) under the discounted payoff criterion. In Sect. 2, we define two-person finite SMGs and supply relevant material on VLCP. In Sect. 3, counterexamples are given to show that zero-sum AR–AT and AR–AT–PT (Additive Reward–Additive Transition Probability and Time) games do not satisfy the ordered field property. We prove that AR–AT–AITT (Additive Reward–Additive Transition and Action Independent Transition Time) and AR–AIT–ATT (Additive Reward–Action Independent Transition and Additive Transition Time) SMGs satisfy the ordered field property and possess pure optimal stationary strategies. We show by illustrations that equilibrium point of pure stationary strategies may not exist in the nonzero-sum case of such games. However, the existence of a stationary equilibrium point which has at most two pure actions for each player in each state is proved. In Sect. 4, we formulate the zero-sum AR–AT–AITT and AR–AIT–ATT games as VLCP and show that the Cottle-Dantzig's algorithm can process these problems under some mild assumptions.

2 Background and preliminaries

2.1 Two-person finite semi-Markov games

A finite two-person semi-Markov game is defined by a collection of objects $<$ S, $\{A(s) : s \in S\}, \{B(s) : s \in S\}, q, P, r_1, r_2 >$ where S $= \{1, 2, \ldots, k\}$ is the finite nonempty state space and for each $s \in$ S, $A(s)$ and $B(s)$ are the finite sets of admissible actions in state s for Player 1 (P1) and Player 2 (P2) respectively. Let $|A(s)| = m_s$ and $|B(s)| = n_s$ which denote the cardinality of the sets $A(s)$ and $B(s)$ respectively. The set of admissible triplets is defined by $\mathscr{K} = \{(s, i, j)| s \in S, i \in A(s), j \in B(s)\}$. Let $\mathbf{P}(D)$ be the family of probability distributions on a finite set D. Then the map $q : \mathscr{K} \to \mathbf{P}(S)$ is called the transition function of the game. Therefore, for each $(s, i, j) \in \mathscr{K}$, a transition probability vector $q(s, i, j) = (q(1|s, i, j), q(2|s, i, j), \ldots, q(k|s, i, j)) \in [0, 1]^k$ such that $\sum_{s'=1}^{k} q(s'|s, i, j) = 1$. Given $(s, i, j) \in \mathscr{K}$ and $s' \in$ S, let $\tau_{ss'}^{ij}$ be the transition (sojourn) time random variable which denotes the time for a transition from a state s to a state s' by a pair of actions $(i, j) \in A(s) \times B(s)$. For a finite

SMG, we assume that each $\tau_{ss'}^{ij}$ has a finite support $\{1, 2, \ldots, T\}$. Then, the map $P : \mathcal{K} \times S \rightarrow \mathbf{P}(\{1, 2, \ldots, T\})$ is called the conditional transition time distribution. Symbolically, for each $(s, i, j) \in \mathcal{K}$, $s' \in S$, a transition time distribution vector $P_{ss'}^{ij} = (P_{ss'}^{ij}(1), \ldots, P_{ss'}^{ij}(T)) \in [0, 1]^T$ such that $\sum_{t=1}^{T} P_{ss'}^{ij}(t) = 1$. Finally, $r_1, r_2 : \mathcal{K} \rightarrow \mathbf{R}$ are called the payoff functions and they represent the immediate (expected) rewards of P1 and P2 respectively.

If $r_1(s, i, j) + r_2(s, i, j) = 0$ for all $(s, i, j) \in \mathcal{K}$, then we have a two-person zero-sum game. Otherwise, the game is nonzero-sum. For the zero-sum case, we define $r = r_1 = -r_2$ and consider P1 as the maximizer and P2 as the minimizer.

A finite SMG is played over the infinite future as follows: At the 0th decision epoch, the game starts at a state $s_0 \in S$ and the players choose simultaneously (and hence independently of each other) a pure action pair $(i_0, j_0) \in A(s_0) \times B(s_0)$, then the following happens: P1 and P2 receive immediate rewards $r_1(s_0, i_0, j_0)$ and $r_2(s_0, i_0, j_0)$ respectively and the system moves to a new state s_1 with probability $q(s_1|s_0, i_0, j_0)$. The time until the transition occurs is determined by a discrete random variable $\tau_{s_0 s_1}^{i_0 j_0}$ having the probability mass function $P_{s_0 s_1}^{i_0 j_0}(\cdot)$. Once the transition to s_1 occurs on the next decision epoch, the entire process is repeated over and over again.

Clearly, when all the transition times are identical, the game is just a finite stochastic game (Shapley 1953).

A behavioral strategy π_1 for P1 is a sequence $\{\pi_{1n}\}_{n=0}^{\infty}$, where π_{1n} is a probability distribution on $A(s_n)$ depending on the history $h_n = (s_0, i_0, j_0, \ldots, s_{n-1}, i_{n-1}, j_{n-1}, s_n)$ of the system up to the n-th decision epoch, where $(s_k, i_k, j_k) \in \mathcal{K}$ are respectively the state and actions of the players at the k-th decision epoch ($k = 0, 1, 2, \ldots, n$).

A stationary strategy f_1 for P1 consists of a k-tuple $f_1 = (f_1(1), f_1(2), \ldots, f_1(k))$, where $f_1(s) \in \mathbf{P}(A(s))$ for each $s \in S$. The probability of choosing an action i in state s under f_1 will be denoted by $f_1(s, i)$. A stationary strategy f_1 is called pure if $f_1(s, i) \in \{0, 1\}$ for all $s \in S$ and $i \in A(s)$.

For P2, a behavioral strategy π_2 and a stationary strategy f_2 are similarly defined. Let Π_1, Π_2 and $\mathcal{F}_1, \mathcal{F}_2$ be respectively the classes of all behavioral and stationary strategies for P1 and P2 in a semi-Markov game.

Let $(X_0, A_0, B_0, X_1, A_1, B_1, \ldots)$ be the coordinate sequence on the infinite product space $S \times (\{A(s) : s \in S\} \times \{B(s) : s \in S\} \times S)^{\infty}$. Let $r(X_m, A_m, B_m)$ be the immediate reward on the m-th decision epoch ($m = 0, 1, 2, \ldots$), where X_m, A_m and B_m are respectively the state and actions chosen in that epoch. Once the initial state s and a pair of strategies (π_1, π_2) are specified, then so is the probability distribution of $r(X_m, A_m, B_m)$ for every m. The expectation of r is well defined and will be denoted by $\mathbf{E}_{\pi_1, \pi_2}[r(X_m, A_m, B_m) \mid X_0 = s]$. Let $\mathcal{B}(S)$ be the set of all bounded functions on S.

Definition 1 (*Discounted payoffs*) For $\beta \in (0, 1)$ and $(\pi_1, \pi_2) \in \Pi_1 \times \Pi_2$, the expected total β-discounted payoff to Player l, $l = 1, 2$, is defined by

$$V_{\pi_1, \pi_2}^{l\beta}(s) = \mathbf{E}_{\pi_1, \pi_2}\left[\sum_{n=0}^{\infty} \beta^{(\tau_0 + \ldots + \tau_n)} r_l(X_n, A_n, B_n)|X_0 = s\right] \quad \text{for all } s \in S,$$

where $\tau_0 = 0$ and τ_n is the time between the $(n-1)$-th and n-th decision epochs. Clearly, τ_{n+1} is a random variable depending on X_n, A_n, B_n and X_{n+1} for each n and $A_n \in A(X_n)$ and $B_n \in B(X_n)$ are pure actions of the players at the n-th decision epoch. For the zero-sum case, $V_{\pi_1,\pi_2}^{\beta}(s) = V_{\pi_1,\pi_2}^{1\beta}(s) = -V_{\pi_1,\pi_2}^{2\beta}(s)$ for all $s \in$ S.

Definition 2 (*Value and optimal strategies*) A zero-sum two-person β-discounted semi-Markov game is said to have a value vector $v_\beta = \left[v_\beta(s)\right]_{s \in S}$ if

$$\inf_{\pi_2 \in \Pi_2} \sup_{\pi_1 \in \Pi_1} V_{\pi_1,\pi_2}^{\beta}(s) = \sup_{\pi_1 \in \Pi_1} \inf_{\pi_2 \in \Pi_2} V_{\pi_1,\pi_2}^{\beta}(s) = v_\beta(s) \quad \text{for all } s \in \text{S}.$$

A pair (π_1^*, π_2^*) is said to be an optimal strategy pair if for all $s \in$ S,

$$V_{\pi_1^*,\pi_2}^{\beta}(s) \geq v_\beta(s) \geq V_{\pi_1,\pi_2^*}^{\beta}(s) \quad \text{for all } (\pi_1, \pi_2) \in \Pi_1 \times \Pi_2.$$

To develop a standard SMG model, we require a regularity condition (that ensures that an infinite number of transitions do not occur in a finite time interval) and various boundedness and continuity assumptions (see Sect. 4 in Luque-Vasquez 2002). If we endow $\mathcal{K}$ with the discrete topology in a finite SMG, all these assumptions are trivially satisfied. Luque-Vasquez (2002) proved that a β-discounted semi-Markov game has value and a pair of stationary optimal strategies for countable state space, compact action spaces and continuous sojourn times. He obtained a Shapley equation for such games. For a finite semi-Markov game, we observe the following result:

Theorem 1 *The value vector v_β of a zero-sum two-person finite β-discounted semi-Markov game satisfies the Shapley equation*

$$v_\beta(s) = val\left[r(s,i,j) + \sum_{s' \in S} q(s'|s,i,j)v_\beta(s') \sum_{t=1}^{T} \beta^t P_{ss'}^{ij}(t)\right] \quad \textit{for all } s \in \text{S.(1)}$$

Further, $(f_1^*, f_2^*) \in \mathscr{F}_1 \times \mathscr{F}_2$ is an optimal strategy pair if and only if $(f_1^*(s), f_2^*(s))$ is a pair of optimal strategies of the matrix game in (1) above for all $s \in$ S.

For each $u \in \mathscr{B}(\text{S})$, $(s,i,j) \in \mathscr{K}$ and $(f_1, f_2) \in \mathscr{F}_1 \times \mathscr{F}_2$, we define:

$$H(u,s,i,j) = r(s,i,j) + \sum_{s' \in S} q(s'|s,i,j)u(s') \sum_{t=1}^{T} \beta^t P_{ss'}^{ij}(t)$$

$$T_\beta u(s) = \min_{j \in B(s)} \max_{i \in A(s)} H(u,s,i,j)$$

$$T_\beta(f_1,f_2)u(s) = H(u,s,f_1,f_2) = \sum_{i \in A(s)} \sum_{j \in B(s)} H(u,s,i,j)f_1(s,i)f_2(s,j).$$

We call T_β the Shapley operator. Since the reward is bounded, $T_\beta u$ and $T_\beta(f_1, f_2)u \in \mathscr{B}(\text{S})$ for each $u \in \mathscr{B}(\text{S})$. Note that Theorem 1 can be written as $v_\beta(s) = T_\beta v_\beta(s)$ and $T_\beta(f_1^*, f_2^*)v_\beta(s) = v_\beta(s)$ for all $s \in$ S.

We need the following lemmas to prove Theorem 1:

Lemma 1 *(a) For each $u \in \mathcal{B}(S)$, there exists a pair $(f_1^*, f_2^*) \in \mathcal{F}_1 \times \mathcal{F}_2$ such that*

$$T_\beta u(s) = H(u, s, f_1^*, f_2^*) = \max_{f_1 \in \mathcal{F}_1} H(u, s, f_1, f_2^*)$$

$$= \min_{f_2 \in \mathcal{F}_2} H(u, s, f_1^*, f_2) \ \forall s \in S.$$

(b) Both T_β and $T_\beta(f_1, f_2)$ are contraction operators with modulus less than 1.

Proof The proof follows on similar lines as Lemmas 5.1 and 5.2 in Luque-Vasquez (2002) and the required assumptions are trivially true for our finite SMGs. □

Since both T_β and $T_\beta(f_1, f_2)$ are contraction operators, by Banach's Fixed Point Theorem there exist unique functions v^* and v_{f_1, f_2}^* in $\mathcal{B}(S)$ such that $T_\beta v^*(s) = v^*(s)$ and $T_\beta(f_1, f_2) v_{f_1, f_2}^*(s) = v_{f_1, f_2}^*(s)$ for all $s \in S$.

Lemma 2 *For each $(f_1, f_2) \in \mathcal{F}_1 \times \mathcal{F}_2$, $V_{f_1, f_2}^\beta(\cdot)$ is the unique fixed point of $T_\beta(f_1, f_2)$.*

Proof In what follows $\mathbf{E}_{f_1, f_2}(\cdot \mid X_0 = s)$ is denoted by $\mathbf{E}$. Now, by definition

$$V_{f_1, f_2}^\beta(s) = \mathbf{E}\left[\sum_{n=0}^\infty \beta^{(\tau_0 + \ldots + \tau_n)} r(X_n, A_n, B_n)\right]$$

$$= \mathbf{E}[r(X_0, A_0, B_0)] + \mathbf{E}\left[\mathbf{E}_{f_1, f_2}\left\{\sum_{n=1}^\infty \beta^{(\tau_1 + \ldots + \tau_n)} r(X_n, A_n, B_n) \mid h_1, \tau_1\right\}\right]$$

$$= \sum_{i \in A(s)} \sum_{j \in B(s)} r(s, i, j) f_1(s, i) f_2(s, j) + \mathbf{E}[\beta^{\tau_1} V_{f_1, f_2}^\beta(X_1)]$$

$$= r(s, f_1, f_2) + \sum_{i \in A(s)} \sum_{j \in B(s)} \left[\sum_{s' \in S} q(s'|s, i, j) V_{f_1, f_2}^\beta(s') \sum_t \beta^t P_{ss'}^{ij}(t)\right]$$

$$f_1(s, i) f_2(s, j) = T_\beta(f_1, f_2) V_{f_1, f_2}^\beta(s) \quad \text{for all} \quad s \in S.$$

□

Lemma 3 *A pair of strategies $(f_1^*, f_2^*) \in \mathcal{F}_1 \times \mathcal{F}_2$ is optimal if and only if its expected payoff satisfies Shapley equation $T_\beta V_{f_1^*, f_2^*}^\beta(s) = V_{f_1^*, f_2^*}^\beta(s)$ for all $s \in S$.*

Proof We use Lemmas 1 and 2 and the rest of the proof follows on similar lines as in Luque-Vasquez (2002). For details, see proof of Theorem 4.3, pp. 10–12. □

Proof of Theorem 1 The result now follows from Lemmas 2 and 3. □

Let $p(s'|s, i, j) = q(s'|s, i, j) \sum_{t=1}^T \beta^t P_{ss'}^{ij}(t)$ be the probability for a transition from state s to s' in the next day by the pair of actions $(i, j) \in A(s) \times B(s)$. Then the system will stop in state s with probability $p'(s, i, j) = 1 - \sum_{s' \in S} p(s'|s, i, j)$ for

$(i, j) \in A(s) \times B(s)$. Since $\beta \in (0, 1)$, $0 < \sum_{s' \in S} p(s'|s, i, j) < 1$ and hence $p'(s, i, j) \in (0, 1)$ for all $(s, i, j) \in \mathcal{K}$ and the game ends with probability 1 after a finite number of steps. Thus a finite discounted SMG reduces to a stopping stochastic game (Shapley 1953), where the optimality equation becomes

$$v_\beta(s) = val \left[r(s, i, j) + \sum_{s' \in S} p(s'|s, i, j) v_\beta(s') \right] \quad \text{for all} \quad s \in S. \tag{2}$$

Note that Shapley Eqs. (1) and (2) are identical. However, (1) is more convenient for studying a structured (may be with AR–AT) β-discounted SMG.

Remark 1 One can rewrite a finite semi-Markov game as a stochastic game on a larger state space by replacing the transition time by staying in states, with certain probabilities, in which the payoff is zero. However, the transformed SG obtained in this way may not work for the original SMG. See Schweitzer (1971) and Jianyong and Xiaobo (2004) (Example 2.1, p. 342–343) for details.

Definition 3 (*Nash equilibrium*) A pair of strategies $(f_1^*, f_2^*) \in \mathcal{F}_1 \times \mathcal{F}_2$ constitutes a stationary Nash equilibrium pair in the discounted SMG, if for all $s \in S$,

$$V_{f_1^*, f_2^*}^{1\beta}(s) \geq V_{f_1, f_2^*}^{1\beta}(s) \quad \text{for all} \quad f_1 \in \mathcal{F}_1$$

$$\text{and} \quad V_{f_1^*, f_2^*}^{2\beta}(s) \geq V_{f_1^*, f_2}^{2\beta}(s) \quad \text{for all} \quad f_2 \in \mathcal{F}_2.$$

Suppose that for all $s \in S$, $\bar{V}_{f_2}^{1\beta}(s) = \max_{f_1 \in \mathcal{F}_1} V_{f_1, f_2}^{1\beta}(s)$ and $\bar{V}_{f_1}^{2\beta}(s) = \max_{f_2 \in \mathcal{F}_2} V_{f_1, f_2}^{2\beta}(s)$. If one of the players, say P2, chooses a fixed stationary strategy f_2, then the semi-Markov game problem reduces to a discounted semi-Markov decision problem for P1. In this problem, $\bar{V}_{f_2}^{1\beta}(s)$ is the maximal (optimal) β-discounted payoff for P1 in state s and there exists a β-discounted stationary maximal (optimal) policy f_1^* for P1. The existence of such an f_1^* (or f_2^* for P2) is ensured by the finite assumptions of state and action spaces (Ross 1970). We now state the following useful result. For a proof we refer to Ross (1970).

Lemma 4 *For* $(f_1, f_2) \in \mathcal{F}_1 \times \mathcal{F}_2$, $\left(\bar{V}_{f_2}^{1\beta}(\cdot), \bar{V}_{f_1}^{2\beta}(\cdot) \right)$ *is the unique solution in* $\mathscr{B}(S) \times \mathscr{B}(S)$ *of the following system of equations in* $\left(v^1(\cdot), v^2(\cdot) \right)$:

$$v^1(s) = \max_{f_1 \in \mathbf{P}(A(s))} \left[r_1(s, f_1, f_2) + \sum_{i \in A(s)} \sum_{j \in B(s)} \left\{ \sum_{s' \in S} q(s'|s, i, j) v^1(s') \sum_t \beta^t P_{ss'}^{ij}(t) \right\} \right.$$
$$\left. f_1(s, i) f_2(s, j) \right], \quad s \in S \tag{3}$$

$$v^2(s) = \max_{f_2 \in \mathbf{P}(B(s))} \left[r_2(s, f_1, f_2) + \sum_{i \in A(s)} \sum_{j \in B(s)} \left\{ \sum_{s' \in S} q(s'|s, i, j) v^2(s') \sum_t \beta^t P_{ss'}^{ij}(t) \right\} \right.$$
$$\left. f_1(s, i) f_2(s, j) \right], \quad s \in S. \tag{4}$$

Furthermore, $(f_1^*, f_2^*) \in \mathscr{F}_1 \times \mathscr{F}_2$ is a pair of optimal response for the two players iff $f_1^*(s)$ and $f_2^*(s)$ maximize (3) and (4) respectively for each $s \in$ S.

Theorem 2 *There exists a Nash equilibrium pair $(f_1^*, f_2^*) \in \mathscr{F}_1 \times \mathscr{F}_2$ in any finite two-person nonzero-sum β-discounted semi-Markov game.*

Proof Polowczuk (2000) proved that there exists a β-discounted stationary Nash equilibrium pair for any N-person nonzero-sum semi-Markov game with a countable state space, compact metric action spaces and continuous sojourn times. For the finite case, the result follows on similar lines using the optimality Eqs. (3) and (4).   □

Definition 4 (*Semi-Markov games with ordered field property*) A two-person finite semi-Markov game with rational inputs, that is, with rational payoffs, rational transition probabilities and transition time distributions, and rational discount factor is said to possess the (Archimedean) ordered field property if it has a pair of optimal/Nash equilibrium strategies whose coordinates are rational.

It follows that the value/Nash equilibrium payoffs of the game are rational too.

It is not required for the ordered field property that the game can be solved with stationary strategies. Thuijsman and Raghavan (1997) proved the ordered field property for (nonzero-sum) AR–AT stochastic games by allowing non-stationary strategies. Mondal and Sinha (2013, 2015) proved the ordered field property in a subclass of SeR-SIT and for one player control finite semi-Markov games.

2.2 Vertical linear complementarity problem

Vertical linear complementarity problem (VLCP) is a generalization of the linear complementarity problem (LCP). Let A be a matrix of order $m \times k$ with $m \geq k$ and let q be an m-vector. Suppose that A and q are partitioned in the following form

$$A = \begin{bmatrix} A^1, A^2, \ldots, A^k \end{bmatrix}^t, \quad q = \begin{bmatrix} q^1, q^2, \ldots, q^k \end{bmatrix}^t$$

where $A^j \in \mathbf{R}^{m_j \times k}$ and $q^j \in \mathbf{R}^{m_j}$, $1 \leq j \leq k$ with $\sum_{j=1}^{k} m_j = m$. We call A as a vertical block matrix of type $(m_1, m_2, \ldots, m_k)$. Now, the generalized LCP (also known as vertical generalization of the LCP) is to find $w \in R^m$ and $z \in R^k$ such that

$$w - Az = q, \quad w \geq \mathbf{0}_m, \quad z \geq \mathbf{0}_k$$
$$z_j \prod_{i=1}^{m_j} w_i^j = 0, \quad j = 1, 2, \ldots, k.$$

This is denoted by VLCP(q, A). When $m_j = 1$ for each j, this reduces to the standard LCP(q, A). For a detail study on LCP, we refer the book of Cottle et al. (1992). The VLCP arises in solving some classes of stochastic game problems (see Mohan et al. 1999, 2001). Cottle and Dantzig (1970) described a procedure for solving a VLCP, which is an extension of Lemke's algorithm (1965) for the LCP. Under the standard non-degeneracy assumption, the algorithm either terminates in a solution to the VLCP(q, A) or in a secondary proper ray.

Definition 5 *A* is said to be a vertical block $E(d)$-matrix (denoted by $A \in VBE(d)$) for some $d > 0$ if VLCP(d, A) has a unique solution $w = d, z = 0$.

Definition 6 *A* is said to be a vertical block R_0-matrix (denoted by $A \in VBR_0$) if VLCP(0, A) has a unique solution $w = 0, z = 0$.

If the vertical block matrix $A \in VBE(d) \cap VBR_0$, then VLCP(q, A) is processable by Cottle-Dantzig's algorithm (Mohan et al. (1996)).

3 Main results

We first describe the different conditions used throughout this paper:

AR (Additive Rewards): The rewards can be written as the sum of two functions, one depending on P1 and other on P2. So, for the zero-sum case,

$$r(s, i, j) = r_1(s, i) + r_2(s, j) \ \forall \ (s, i, j) \in \mathcal{K},$$

and for the nonzero-sum case,

$$r_1(s, i, j) = r_{11}(s, i) + r_{12}(s, j) \quad \text{and} \quad r_2(s, i, j) = r_{21}(s, i) + r_{22}(s, j) \ \forall \ (s, i, j) \in \mathcal{K}.$$

AT (Additive Transitions): The transition probabilities can be written as the sum of two functions, one depending on P1 and other] on P2, i.e.,

$$q(s'|s, i, j) = q_1(s'|s, i) + q_2(s'|s, j) \ \forall \ (s, i, j) \in \mathcal{K}, \quad s' \in S, \ni q_1, q_2 \geq 0.$$

ATT (Additive Transition Times): The transition times are additive, i.e.,

$$P_{ss'}^{ij}(t) = (P_1)_{ss'}^{i}(t) + (P_2)_{ss'}^{j}(t),$$
$$t = 1, 2, \ldots, T; \quad \forall \ (s, i, j) \in \mathcal{K}, \quad s' \in S, \ni P_1, P_2 \geq 0.$$

AIT (Action Independent Transitions) Transition probabilities are independent of the actions of the players. That means

$$q(s'|s, i, j) = q(s'|s, i', j') \equiv q(s'|s) \ \forall \ (s, i, j), (s, i', j') \in \mathcal{K}, \quad s' \in S.$$

AITT (Action Independent Transition Times): Transition times are independent of the actions of the players. Thus,

$$P_{ss'}^{ij}(t) = P_{ss'}^{i'j'}(t) \equiv P_{ss'}(t), \ t = 1, 2, \ldots, T; \forall \ (s, i, j), (s, i', j') \in \mathcal{K}, \quad s' \in S.$$

Definition 7 (*AR–AT games*) A finite semi-Markov game with AR and AT conditions is called an AR–AT (Additive Rewards Additive Transitions) game.

Definition 8 (*AR–AT–PT games*) A finite semi-Markov game with AR, AT and ATT conditions is called an AR–AT–PT (Additive Rewards and Additive Transition Probabilities and Times) game.

Definition 9 (*AR–AT–AITT games*) A finite semi-Markov game is called an AR–AT–AITT (Additive Rewards Additive Transitions and Action Independent Transition Times) game if AR, AT and AITT conditions hold.

Definition 10 (*AR–AIT–ATT games*) A finite semi-Markov game is called an AR–AIT–ATT (Additive Rewards Action Independent Transitions and Additive Transition Times) game if AR, AIT and ATT conditions hold.

Fishery games (Filar and Vrieze 1997) were proposed as an application of AR–AT stochastic games. One can construct a Petroleum game model (where the state space represents the reservation of petroleum in a country and the players are different petroleum companies) for AR–AT–AITT semi-Markov games in a similar manner.

Zero-sum discounted AR–AT semi-Markov games do not possess the ordered field property as the following example shows:

Example 1 Let $p^1 = q^1 = (1/2, 0)$, $p^2 = q^2 = (0, 1/2)$ and $a = 2, b = 1, c = 1$, $d = 0$.

State 1:

$a + c$ $p^1 + q^1$	$a + d$ $p^1 + q^2$	3 $(1, 0)$	2 $(1/2, 1/2)$
$b + c$ $p^2 + q^1$	$b + d$ $p^2 + q^2$	2 $(1/2, 1/2)$	1 $(0, 1)$

State 2:

$a + b$ $p^2 + q^2$	3 $(0, 1)$

where a cell $\begin{array}{c} r \\ (q_1, q_2) \end{array}$ corresponds to an immediate reward r and a transition with probability q_1 to state 1 and probability q_2 to state 2 if this cell is chosen in the present state. Let P1 and P2 have pure strategies x_1, x_2 and y_1, y_2 respectively in state 1, x_3 and y_3 in state 2. Transition times of this game are defined by $P_{11}^{x_1 y_1} = P_{12}^{x_2 y_2} = \{1(1)\}$, $P_{11}^{x_1 y_2} = P_{12}^{x_1 y_2} = P_{11}^{x_2 y_1} = P_{12}^{x_2 y_1} = \{2(1/2), 3(1/2)\}$ and $P_{22}^{x_3 y_3} = \{1(1)\}$, where the entries inside the round brackets are the probabilities of the respective transition times. Let $\beta = 1/2$. By Shapley Eq. (1) we get,

$$v_2 = val \left[3 + \frac{v_2}{2} \right] = 3 + \frac{v_2}{2} \Rightarrow v_2 = 6$$

and $v_1 = val \begin{bmatrix} 3 + \frac{v_1}{2} & 2 + \frac{3v_1}{32} + \frac{3v_2}{32} \\ 2 + \frac{3v_1}{32} + \frac{3v_2}{32} & 1 + \frac{v_2}{2} \end{bmatrix} = val \begin{bmatrix} 3 + \frac{v_1}{2} & \frac{41}{16} + \frac{3v_1}{32} \\ \frac{41}{16} + \frac{3v_1}{32} & 4 \end{bmatrix}$.

Solving the above matrix game by Kaplansky (1945), we get $v_1 = \frac{64\sqrt{455} - 182}{329}$ which is an irrational number. This implies that the ordered field property does not hold.

The question arises whether the zero-sum AR–AT–PT games (a natural extension of AR–AT stochastic games) possess the ordered field property. The answer is negative which is shown in the following example:

Example 2 In the Example 1, we modify transition times to additive transition times but the rest of the data for this game is same as before. Let

$$(P_1)_{11}^{x_1} = \{1(1/2), 3(0)\}, \quad (P_2)_{11}^{y_1} = \{1(1/2), 3(0)\}$$
$$(P_1)_{11}^{x_2} = \{1(0), 3(1/2)\}, \quad (P_2)_{11}^{y_2} = \{1(0), 3(1/2)\}$$
$$(P_1)_{12}^{x_1} = \{1(0), 4(1/2)\}, \quad (P_2)_{12}^{y_1} = \{1(0), 4(1/2)\}$$
$$(P_1)_{12}^{x_2} = \{1(1/2), 4(0)\}, \quad (P_2)_{12}^{y_2} = \{1(1/2), 4(0)\}$$

and $P_{ss'}^{x_i y_j} = (P_1)_{ss'}^{x_i} + (P_2)_{ss'}^{y_j}, \quad s, s' = 1, 2, i = 1, 2; j = 1, 2$

$$(P_1)_{22}^{x_3} = \{1(1/2)\}, \quad (P_2)_{22}^{y_3} = \{1(1/2)\} \quad \text{and} \quad P_{22}^{x_3 y_3} = (P_1)_{22}^{x_3} + (P_2)_{22}^{y_3}.$$

Choose $\beta = 1/2$. Now by Shapley equation (1) we get, $v_2 = 6$,

and $v_1 = val \begin{bmatrix} 3 + \frac{v_1}{2} & 2 + \frac{5v_1}{32} + \frac{9v_2}{64} \\ 2 + \frac{5v_1}{32} + \frac{9v_2}{64} & 1 + \frac{v_2}{2} \end{bmatrix} = val \begin{bmatrix} 3 + \frac{v_1}{2} & \frac{91}{32} + \frac{5v_1}{32} \\ \frac{91}{32} + \frac{5v_1}{32} & 4 \end{bmatrix}.$

Using Kaplansky (1945) and solving, we get $v_1 = \frac{48\sqrt{382}-103}{217}$ which is an irrational number. Thus in this case, the ordered field property does not hold.

We observe the following result:

Theorem 3 *A zero-sum two-person β-discounted AR–AT–AITT semi-Markov game possesses the ordered field property and each player has an optimal stationary strategy that is pure.*
Similar result holds for an AR–AIT–ATT semi-Markov game.

Proof Consider the Shapley Eq. (1) of an AR–AT–AITT semi-Markov game:

$$v_\beta(s) = val\Big[r_1(s, i) + r_2(s, j) + \sum_{s' \in S}(q_1(s'|s, i)$$
$$+ q_2(s'|s, j))v_\beta(s') \sum_{t=1}^{T} \beta^t P_{ss'}(t)\Big] \forall s \in S.$$

The above equation can be decomposed in the following way:

$$v_\beta(s) = \min_{\mu \in \mathbf{P}(B(s))} \max_{\lambda \in \mathbf{P}(A(s))} \Bigg[\Bigg\{ r_1(s, \lambda) + \sum_{s' \in S} q_1(s'|s, \lambda)v_\beta(s') \sum_{t=1}^{T} \beta^t P_{ss'}(t) \Bigg\}$$
$$+ \Bigg\{ r_2(s, \mu) + \sum_{s' \in S} q_2(s'|s, \mu)v_\beta(s') \sum_{t=1}^{T} \beta^t P_{ss'}(t) \Bigg\} \Bigg]$$
$$= \max_{\lambda \in \mathbf{P}(A(s))} \Bigg[r_1(s, \lambda) + \sum_{s' \in S} q_1(s'|s, \lambda)v_\beta(s') \sum_{t=1}^{T} \beta^t P_{ss'}(t) \Bigg]$$

$$+ \min_{\mu \in \mathbf{P}(B(s))} \left[r_2(s, \mu) + \sum_{s' \in S} q_2(s'|s, \mu) v_\beta(s') \sum_{t=1}^{T} \beta^t P_{ss'}(t) \right]$$

$$= \max_{\lambda \in \mathbf{P}(A(s))} G_{1s}(v_\beta) + \min_{\mu \in \mathbf{P}(B(s))} G_{2s}(v_\beta)s \quad \text{for all } s \in S.$$

Thus, when solving $v_\beta(s)$, P1 needs to consider only $G_{1s}(v_\beta)$ and P2 only needs to look at $G_{2s}(v_\beta)$. From the above observation, it is easy to see that both players have optimal pure actions in each state $s \in S$. Therefore, the ordered field property holds.

For an AR–AIT–ATT game, the proof follows on similar lines. □

Now we ask the following questions. Is there exist any equilibrium point of pure stationary strategy pair for nonzero-sum AR–AT–AITT and AR–AIT–ATT games? Examples 3 and 4 below answer this question in the negative.

Example 3 (AR–AT–AITT game): Let $p^1 = (0, 0, 1/2)$, $p^2 = (1/4, 1/4, 0)$, $q^1 = (0, 1/4, 1/4)$, $q^2 = (1/4, 0, 1/4)$; $a = (1, 1)$, $b = (2, -4)$, $c = (0, 0)$, $d = (-8, 0)$.

State 1:

$\begin{array}{c} a+c \\ p^1+q^1 \end{array}$	$\begin{array}{c} a+d \\ p^1+q^2 \end{array}$		(1, 1)	(−7, 1)
			(0, 1/4, 3/4)	(1/4, 0, 3/4)
$\begin{array}{c} b+c \\ p^2+q^1 \end{array}$	$\begin{array}{c} b+d \\ p^2+q^2 \end{array}$	=	(2, −4)	(−6, −4)
			(1/4, 1/2, 1/4)	(1/2, 1/4, 1/4)

State 2:
(0, 0)
(0, 1, 0)

State 3:
(2, 2)
(0, 0, 1)

The action independent transition times are defined by $P_{11} = \{1(1/2), 2(1/2)\}$, $P_{12} = \{1(3/4), 2(1/4)\}$, $P_{13} = \{2(1/2), 3(1/2)\}$, $P_{22} = \{1(1/4), 2(3/4)\}$, $P_{33} = \{1(1)\}$. Let the players have pure strategies x_1, x_2 and y_1, y_2 respectively in state 1. Take $\beta = 1/2$. It is obvious to show that the Nash equilibrium payoffs corresponding to the absorbing states 2 and 3 are $(0, 0)$ and $(4, 4)$ respectively. Computing $V_{x_i y_j}^{l\beta}(1)$ for $l = 1, 2, i = 1, 2$ and $j = 1, 2$, we obtain the bimatrix game for state 1 as

$$\begin{bmatrix} \left(\frac{25}{16}, \frac{25}{16} \right) & \left(\frac{-206}{29}, \frac{50}{29} \right) \\ \left(\frac{70}{29}, \frac{-122}{29} \right) & \left(\frac{-93}{13}, \frac{-61}{13} \right) \end{bmatrix}. \tag{5}$$

For example, $V_{x_1 y_2}^{1\beta}(1)$ can be computed as the unique solution v of

$$v = -7 + \frac{1}{4} \cdot v \cdot \left(\frac{1}{2} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} \right) + \frac{3}{4} \cdot 4 \cdot \left(\frac{1}{4} \cdot \frac{1}{2} + \frac{1}{8} \cdot \frac{1}{2} \right) \Rightarrow v = \frac{-206}{29}.$$

Similarly, other $V_{x_i y_j}^{k\beta}(1)$ can be verified easily.

From the bimatrix game (5), we can show that there exists no equilibrium point in pure stationary strategies and that for this example, the unique Nash equilibrium point in stationary strategies is completely mixed for state 1.

Example 4 (*AR–AIT–ATT game*): Let $a = (0,0), b = (1,-3), c = (0,0), d = (-15,0)$; $q(1) = (q(1|1), q(2|1)) = (1/2, 1/2), q(2) = (q(1|2), q(2|2)) = (0,1)$.

State 1:

$\begin{array}{c} a+c \\ q(1) \end{array}$	$\begin{array}{c} a+d \\ q(1) \end{array}$
$\begin{array}{c} b+c \\ q(1) \end{array}$	$\begin{array}{c} b+d \\ q(1) \end{array}$

$=$

$\begin{array}{c} (0,1) \\ (1/2,1/2) \end{array}$	$\begin{array}{c} (-15,0) \\ (1/2,1/2) \end{array}$
$\begin{array}{c} (1,-3) \\ (1/2,1/2) \end{array}$	$\begin{array}{c} (-14,-3) \\ (1/2,1/2) \end{array}$

State 2:

$(1,1)$
$(0,1)$

.

Let the players have pure strategies x_1, x_2 and y_1, y_2 respectively in state 1, x_3 and y_3 respectively in state 2. Take $\beta = 1/2$. The additive transition times are defined by

$$(P_1)_{11}^{x_1} = \{1(0), 2(1/2)\}, \quad (P_1)_{12}^{x_1} = \{1(1/4), 2(1/4)\}$$
$$(P_1)_{11}^{x_2} = \{1(1/2), 2(0)\}, \quad (P_1)_{12}^{x_2} = \{1(1/4), 2(1/4)\},$$
$$(P_2)_{11}^{y_1} = \{1(1/2), 2(0)\}, \quad (P_2)_{12}^{y_1} = \{1(1/8), 2(3/8)\}$$
$$(P_2)_{11}^{y_2} = \{1(1/4), 2(1/4)\}, \quad (P_2)_{12}^{y_2} = \{1(0), 2(1/2)\},$$

where $P_{ss'}^{x_i y_j} = (P_1)_{ss'}^{x_i} + (P_2)_{ss'}^{y_j}$ for $s = 1$; $s' = 1, 2$; $i, j = 1, 2$,

$$(P_1)_{22}^{x_3} = \{1(1/2)\} (P_2)_{22}^{y_3} = \{1(1/2)\} \quad \text{and} \quad P_{22}^{x_3 y_3} = (P_1)_{22}^{x_3} + (P_2)_{22}^{y_3}.$$

The Nash equilibrium payoffs corresponding to the absorbing state 2 are $(2,2)$. The bimatrix game corresponding to state 1 can be computed as (similar to Example 3):

$$\begin{bmatrix} \left(\frac{11}{26}, \frac{11}{26}\right) & \left(\frac{-420}{27}, \frac{10}{27}\right) \\ \left(\frac{43}{24}, \frac{-85}{24}\right) & \left(\frac{-438}{25}, \frac{-86}{25}\right) \end{bmatrix}.$$

It is easy to show that there exists a unique Nash equilibrium point in stationary strategies which is completely mixed in state 1. This implies that pure stationary Nash equilibria does not exist.

For $x = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$, we define $car(x) = \{k : x_k \neq 0\}$. Now for the discounted nonzero-sum AR–AT–AITT and AR–AIT–ATT semi-Markov games, we have the following remarkable result:

Theorem 4 *A discounted nonzero-sum AR–AT–AITT semi-Markov game has a Nash equilibrium point* $(\tilde{f}_1, \tilde{f}_2) \in \mathscr{F}_1 \times \mathscr{F}_2$ *such that*

$$|car(\tilde{f}_1(s))| \leq 2 \quad and \quad |car(\tilde{f}_2(s))| \leq 2 \quad for \ each \ s \in S.$$

Similar result holds for an AR–AIT–ATT semi-Markov game.

Proof We use (3) and (4) and the rest of the proof follows on similar lines as in Raghavan et al. (1985) (Theorem 3.1, p. 462) for AR–AT stochastic games. □

4 Computing optimal pure strategies of the subclasses of AR–AT games

We show that the zero-sum discounted AR–AT–AITT and AR–AIT–ATT semi-Markov games can be formulated as a VLCP which can be solved by a pivot scheme under some mild assumption. We follow the similar approach used in Mohan et al. (1999).

4.1 Discounted zero-sum AR–AT–AITT semi-Markov games

4.1.1 VLCP formulation

The Shapley Eq. (1) gives us the following for state $s, s \in S$:

$$r(s, i, j) + \sum_{s' \in S} q(s'|s, i, j)v_\beta(s') \sum_{t=1}^{T} \beta^t P_{ss'}^{ij}(t) \le v_\beta(s) \ \forall i \in A(s) \quad \text{and a fixed } j \in B(s).$$

In particular, suppose that an optimal pure action in state s is i_0 for P1 and j_0 for P2 (by Theorem 3). Then, for an AR–AT–AITT semi-Markov game, we have

$$r_1(s, i) + r_2(s, j_0) + \sum_{s' \in S} q_1(s'|s, i)v_\beta(s') \sum_t \beta^t P_{ss'}(t) + \sum_{s' \in S} q_2(s'|s, j_0)v_\beta(s')$$

$$\times \sum_t \beta^t P_{ss'}(t) \le v_\beta(s) \ \text{ for all } i \in A(s)$$

$$\Rightarrow r_1(s, i) + \sum_{s' \in S} q_1(s'|s, i)v_\beta(s') \sum_t \beta^t P_{ss'}(t) \le v_\beta(s) - \eta_\beta(s) = \xi_\beta(s) \ \forall i,$$

where

$$\eta_\beta(s) = r_2(s, j_0) + \sum_{s' \in S} q_2(s'|s, j_0)v_\beta(s') \sum_t \beta^t P_{ss'}(t)$$

$$\text{and} \quad \xi_\beta(s) = r_1(s, i_0) + \sum_{s' \in S} q_1(s'|s, i_0)v_\beta(s') \sum_t \beta^t P_{ss'}(t).$$

$$\text{Let} \quad w_1(s, i) = -r_1(s, i) - \sum_{s' \in S} q_1(s'|s, i)\xi_\beta(s') \sum_t \beta^t P_{ss'}(t) + \xi_\beta(s)$$

$$- \sum_{s' \in S} q_1(s'|s, i)\eta_\beta(s') \sum_t \beta^t P_{ss'}(t) \ge 0 \ \text{ for all } s \in S, \quad i \in A(s) \tag{6}$$

$$\text{and} \quad w_2(s, j) = r_2(s, j) + \sum_{s' \in S} q_2(s'|s, j)\eta_\beta(s') \sum_t \beta^t P_{ss'}(t) - \eta_\beta(s)$$

$$+ \sum_{s' \in S} q_2(s'|s, j)\xi_\beta(s') \sum_t \beta^t P_{ss'}(t) \ge 0 \ \text{ for all } s \in S, \quad j \in B(s). \tag{7}$$

We may assume without loss of generality that $\eta_\beta(s)$, $\xi_\beta(s)$ are all positive. Since, there is at least one inequality in each of (6) and (7) for each $s \in$ S that holds as an equality, the following complementarity conditions will hold:

$$\eta_\beta(s) \prod_{i \in A(s)} w_1(s, i) = 0 \quad \text{for } s \in \{1, 2, \ldots, k\} \tag{8}$$

$$\text{and} \quad \xi_\beta(s) \prod_{j \in B(s)} w_2(s, j) = 0 \quad \text{for } s \in \{1, 2, \ldots, k\}. \tag{9}$$

The inequalities (6) and (7) along with the complementarity conditions (8) and (9) lead to the VLCP (q, A), where the matrix A is of the form

$$A = \begin{bmatrix} -\bar{M}_1 & E - \bar{M}_1 \\ -F + \bar{M}_2 & \bar{M}_2 \end{bmatrix} \quad \text{and} \quad q = \begin{bmatrix} -r_1(\cdot, \cdot) \\ r_2(\cdot, \cdot) \end{bmatrix}.$$

In the above VLCP, $\bar{M}_1 = \left[q_1(s'|s, i) \sum_t \beta^t P_{ss'}(t) \right]$ of order $(\sum m_s) \times k$ and $\bar{M}_2 = \left[q_2(s'|s, j) \sum_t \beta^t P_{ss'}(t) \right]$ of order $(\sum n_s) \times k$. For each $s \in$ S, $\bar{M}_1(s)$ and $\bar{M}_2(s)$ denote the matrices of order $m_s \times k$ and $n_s \times k$ respectively. E and F are vertical block identity matrices, corresponding to P1 and P2 respectively, of the form

$$\begin{bmatrix} e^1 & 0 & \ldots & 0 \\ 0 & e^2 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & e^k \end{bmatrix} \tag{10}$$

where e^s, $s \in \{1, 2, \ldots, k\}$ is a column vector of all 1's of appropriate order.

4.1.2 Convergence of Cottle-Dantzig algorithm for AR–AT–AITT games

We show that the vertical block matrix arising from a zero-sum discounted AR–AT–AITT game belongs to a processable class under a mild assumption.

Theorem 5 *Consider the vertical block matrix A arising from the zero-sum AR–AT–AITT game. Then $A \in VBE(l)$ where l is the vector each of whose entries is 1.*

Proof Let $d = \begin{bmatrix} d^1 \\ d^2 \end{bmatrix}$ where $d^1 > 0$ and $d^2 > 0$. We shall show by contradiction that VLCP(d, A) has only the trivial solution $w = d$, $z = 0$, when $d = l$.
Let $w = \begin{bmatrix} w^1 \\ w^2 \end{bmatrix}$ and $z = \begin{bmatrix} \eta_\beta \\ \xi_\beta \end{bmatrix}$ be a solution to VLCP(d, A). Then $\begin{bmatrix} w^1 \\ w^2 \end{bmatrix} = \begin{bmatrix} d^1 \\ d^2 \end{bmatrix} + A \begin{bmatrix} \eta_\beta \\ \xi_\beta \end{bmatrix}$.
Assume $\begin{bmatrix} \eta_\beta \\ \xi_\beta \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Let $v_\beta(s) = \xi_\beta(s) + \eta_\beta(s)$ and $v_\beta(s^*) = \max_{s \in S} v_\beta(s) > 0$.

Case 1: Let $\eta_\beta(s^*) > 0$. Then, there exists an $i \in A(s^*)$ such that

$$\xi_\beta(s^*) = -d^1(i) + \sum_{s' \in S} q_1(s'|s^*, i)v_\beta(s') \sum_t \beta^t P_{s^*s'}(t). \tag{11}$$

We also have from the feasibility condition

$$d^2(j) + \sum_{s' \in S} q_2(s'|s^*, j)v_\beta(s') \sum_t \beta^t P_{s^*s'}(t) \geq \eta_\beta(s^*). \tag{12}$$

From (11) and (12), we have

$$d^2(j) - d^1(i) + \sum_{s' \in S} q(s'|s^*, i, j)v_\beta(s') \sum_t \beta^t P_{s^*s'}(t) \geq v_\beta(s^*). \tag{13}$$

Since, $d = l$ implies $d^2(j) = d^1(i)$ for all i and j, so we have a contradiction unless $v_\beta(s^*) = 0$ or $v_\beta(s') = 0$ for all $s' \in S$ or $\xi_\beta(s') = \eta_\beta(s') = 0$ for all $s' \in S$.

Case 2: Let $\xi_\beta(s^*) > 0$. Then by complementarity there exists a $j \in B(s^*)$ such that $\eta_\beta(s^*) = d^2(j) + \sum_{s' \in S} q_2(s'|s^*, j)v_\beta(s') \sum_t \beta^t P_{s^*s'}(t)$. Since $d^2(j) > 0$, it follows that $\eta_\beta(s^*) > 0$. Hence, the theorem follows. □

We denote $\sum_{s' \in S} q(s'|s, i, j) \sum_t \beta^t P_{ss'}(t) = \mathbf{E}(\beta^{\bar{\tau}_s^{ij}})$, where $\bar{\tau}_s^{ij}$ is the transition time random variable when the game is AITT. We have the following theorem.

Theorem 6 *Consider the vertical block matrix A arising from the zero-sum AR–AT–AITT game. Then $A \in VBR_0$ if either the condition (a) or the set of conditions (b) stated below is satisfied:*

(a) For each $s \in S$ and each $j \in B(s)$, $q_2(s|s, j) > 0$.
(b) (i) For each $s \in S$, the matrix $\bar{M}_1(s)$ does not contain any zero column and
(ii) the matrix $\bar{M}_2(s)$ is a non-null matrix and $\mathbf{E}(\beta^{\bar{\tau}_s^{ij}})$ is independent of i and j.

Proof We shall show by contradiction that VLCP$(0, A)$ has only the trivial solution $w = 0, z = 0$. Let $w = \begin{bmatrix} w^1 \\ w^2 \end{bmatrix}$ and $z = \begin{bmatrix} \eta_\beta \\ \xi_\beta \end{bmatrix}$ be a solution to VLCP$(0, A)$. Suppose $\begin{bmatrix} \eta_\beta \\ \xi_\beta \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. Let $v_\beta(s) = \xi_\beta(s) + \eta_\beta(s)$ and $v_\beta(s^*) = \max_{s \in S} v_\beta(s) > 0$.

Case 1: Let $\eta_\beta(s^*) > 0$. Since $d^2(j) = d^1(i) = 0$ for all i and j, we arrive a contradiction in (13), unless $\xi_\beta(s') = \eta_\beta(s') = 0$ for all $s' \in S$.

Case 2: Next suppose $\eta_\beta(s^*) = 0$. This implies $\xi_\beta(s^*) > 0$. Therefore, by complementarity, there exists a $j \in B(s^*)$ such that

$$\sum_{s' \in S} q_2(s'|s^*, j)v_\beta(s') \sum_t \beta^t P_{s^*s'}(t) = \eta_\beta(s^*).$$

Suppose now condition (a) holds. Then, $q_2(s^*|s^*, j) > 0$. Also, we have $v_\beta(s^*) > 0$ and $\sum_t \beta^t P_{s^*s^*}(t) > 0$. It follows that $\eta_\beta(s^*) > 0$. Hence Case 2 does not arise if

condition (a) holds. Now suppose the set of conditions (b) holds. Since for each $s \in S$, the matrix $\bar{M}_1(s)$ does not have a zero column, we have by the feasibility condition

$$\xi_\beta(s)\, e^s - \bar{M}_1(s)\,(\eta_\beta + \xi_\beta) \geq 0$$

where e^s denotes the vector of 1's of order m_s . Then it follows that $\xi_\beta(s) > 0$ for each $s \in S$. Using condition (b)(ii), it is easy to see (similar to Lemma 1 in Mohan et al. 1999) that for $s = s^*$, $q_2(s'|s^*, j) > 0$ for some $s' \in S$ and thus $\eta_\beta(s^*) > 0$.

Hence again Case 2 does not arise. $\qquad\square$

4.2 Discounted zero-sum AR–AIT–ATT games

4.2.1 VLCP formulation

We use Shapley Eq. (1) and proceed on similar lines as in Sect. 4.1.1 and obtaining the following VLCP(q, A), where the matrix A is of the form

$$A = \begin{bmatrix} -\tilde{M}_1 & E - \tilde{M}_1 \\ -F + \tilde{M}_2 & \tilde{M}_2 \end{bmatrix} \quad \text{and} \quad q = \begin{bmatrix} -r_1(\cdot, \cdot) \\ r_2(\cdot, \cdot) \end{bmatrix}.$$

In the above VLCP, $\tilde{M}_1 = \left[q(s'|s) \sum_t \beta^t (P_1)^i_{ss'}(t) \right]$ of order $(\sum m_s) \times k$ and $\tilde{M}_2 = \left[q(s'|s) \sum_t \beta^t (P_2)^j_{ss'}(t) \right]$ of order $(\sum n_s) \times k$. For each $s \in S$, $\tilde{M}_1(s)$ and $\tilde{M}_2(s)$ denote the matrices of order $m_s \times k$ and $n_s \times k$ respectively. E and F are vertical block identity matrices, corresponding to P1 and P2 respectively, of the form (10).

4.2.2 Convergence of Cottle-Dantzig algorithm for AR–AIT–ATT games

We write $\sum_{s' \in S} q(s'|s) \sum_t \beta^t P^{ij}_{ss'}(t) = \mathbf{E}(\beta^{\tilde{\tau}^{ij}_s})$, where $\tilde{\tau}^{ij}_s$ is the transition time random variable when the game is AIT.

Following the same techniques as in Sect. 4.1.2, we have the following results:

Theorem 7 *Consider the vertical block matrix A arising from the zero-sum AR–AIT–ATT game. Then $A \in VBE(l)$ where l is the vector each of whose entries is 1.*

Theorem 8 *Consider the vertical block matrix A arising from the zero-sum AR–AIT–ATT game. Then $A \in VBR_0$ if either the condition (a) or the set of conditions (b) stated below is satisfied:*

(a) For each $s \in S$ and each $j \in B(s)$, $q(s|s) \sum_t \beta^t (P_2)^j_{ss}(t) > 0$.

(b) (i) For each $s \in S$, the matrix $\tilde{M}_1(s)$ does not contain any zero column and

(ii) the matrix $\tilde{M}_2(s)$ is a non-null matrix and $\mathbf{E}(\beta^{\tilde{\tau}^{ij}_s})$ is independent of i and j.

5 Further areas of research

In this paper, we deal with pivotal algorithm for some subclasses of AR–AT semi-Markov games. It is interesting to know how neural network algorithm (see Neogy

et al. 2008) perform for this class of games. For the undiscounted case of such games, the existence of value vector and Nash equilibrium pair is still an open problem.

References

Cottle RW, Dantzig GB (1970) A generalization of the linear complementarity problem. J Comb Theory 8:79–90

Cottle RW, Pang JS, Stone RE (1992) The linear complementarity problem. Academic Press, New York

Filar JA, Vrieze K (1997) Competitive Markov decision processes. Springer, New York

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

Jianyong L, Xiaobo Z (2004) On average reward semi-Markov decision processes with a general multichain structure. Math Oper Res 29(2):339–352

Kaplansky I (1945) A contribution to von Neumanns theory of games. Ann Math 46(2):474–479

Lemke C (1965) Bimatrix Equilibrium Points and Mathematical Programming. Management Science 11:681–689

Luque-Vasquez F (2002) Zero-sum semi-Markov games in Borel spaces: discounted and average payoff. Bol. Soc. Mat. Mexicana 8:227–241

Mohan SR, Neogy SK, Parthasarathy T, Sinha S (1999) Vertical linear complementarity and discounted zero-sum stochastic games with ARAT structure. Math Prog Ser A 86:637–648

Mohan SR, Neogy SK, Parthasarathy T (2001) Pivoting algorithms for some classes of stochastic games: a survey. Int Game Theory Rev 3(2 & 3):253–281

Mohan SR, Neogy SK, Sridhar R (1996) The generalized linear complementarity problem revisited. Math Program 74:197–218

Mondal P, Sinha S (2013) Ordered field property in a subclass of finite SeR-SIT semi-Markov games. Int Game Theory Rev 15(4):1340026-1–1340026-20. doi:10.1142/S0219198913400264

Mondal P, Sinha S (2015) Ordered field property for semi-Markov games when one player controls transition probabilities and transition times. Int Game Theory Rev. 17(1):1540022-1–1540022-26. doi:10.1142/S0219198915400228

Nash JF (1951) Non cooperative games. Ann Math 54:286–295

Neogy SK, Das AK, Das P (2008) Complementarity problem involving a Vertical Block Matrix and its Solution using Neural Network Model. Mathematical Programming and Game Theory for decision making. ISI platinum jubilee series on statistical science and interdisciplinary research 1:113–130

Polowczuk W (2000) Nonzero-sum semi-Markov games with countable state spaces. Appl Math 27(4):395–402

Raghavan TES, Tijs SH, Vrieze OJ (1985) On stochastic games with additive reward and transition structure. J Optim Theory Appl 47:375–392

Ross SM (1970) Applied probability models with optimization applications. Holden-Day, San Francisco

Shapley L (1953) Stochastic games. Proc Natl Acad Sci USA 39:1095–1100

Schweitzer PJ (1971) Iterative solution of the functional equations of undiscounted Markov renewal programming. J Math Anal Appl 34(3):495–501

Thuijsman F, Raghavan TES (1997) Perfect information stochastic games and related classes. Int J Game Theory 26:403–408