

# Zero-Sum Discounted Reward Criterion Games for Piecewise Deterministic Markov Processes

O. L. V. Costa<sup>1</sup> · F. Dufour<sup>2</sup>

Published online: 25 April 2017  
© Springer Science+Business Media New York 2017

**Abstract** This paper deals with the zero-sum game with a discounted reward criterion for piecewise deterministic Markov process (PDMPs) in general Borel spaces. The two players can act on the jump rate and transition measure of the process, with the decisions being taken just after a jump of the process. The goal of this paper is to derive conditions for the existence of min–max strategies for the infinite horizon total expected discounted reward function, which is composed of running and boundary parts. The basic idea is, by using the special features of the PDMPs, to re-write the problem via an embedded discrete-time Markov chain associated to the PDMP and re-formulate the problem as a discrete-stage zero sum game problem.

**Keywords** Zero sum games · Continuous-time · Discounted reward criterion · General borel spaces

**Mathematics Subject Classification** Primary 90C40 · Secondary 91A05 · 91A15 · 93E03

---

✉ O. L. V. Costa  
oswaldo@lac.usp.br

F. Dufour  
dufour@math.u-bordeaux1.fr

<sup>1</sup> Departamento de Engenharia de Telecomunicações e Controle, Escola Politécnica da Universidade de São Paulo, CEP: 05508 900 São Paulo, Brazil

<sup>2</sup> Institut Polytechnique de Bordeaux, Institut de Mathématiques de Bordeaux, INRIA Bordeaux Sud Ouest, 200 Avenue de la Vieille Tour, 33405 Talence Cedex, France

## 1 Introduction

Piecewise deterministic Markov processes (PDMPs) were introduced in [2] and [3] as a general family of continuous-time non-diffusion stochastic models, suitable for formulating many optimization problems in queuing and inventory systems, maintenance-replacement models, and many other areas of engineering and operations research. PDMPs are determined by three local characteristics: the flow  $\phi$ , the jump rate  $\lambda$ , and the transition measure  $Q$ . Starting from  $x$ , the motion of the process follows the flow  $\phi(x, t)$  until the first jump time  $T_1$ , which occurs either spontaneously in a Poisson-like fashion with rate  $\lambda$  or when the flow  $\phi(x, t)$  hits the boundary of the state space. In either case the location of the process at the jump time  $T_1$  is selected by the transition measure  $Q(\phi(x, T_1), \cdot)$  and the motion restarts from this new point as before. As shown in [3], a suitable choice of the state space and the local characteristics  $\phi$ ,  $\lambda$ , and  $Q$  provides stochastic models covering a great number of problems of engineering and operations research (see, for instance, [3,4]).

Zero-sum stochastic dynamic games have been recently widely studied in the literature, in discrete as well as continuous time. Regarding the first case the evolution of the process follows a discrete-time Markov process, and we can mention the book [6] and the papers [7, 10, 15–17] as a sample of works dealing with the discrete-time case. For these problems the admissible strategy of the players can be past-depend on the previous states and actions, and the optimal equilibrium solution is usually obtained from stochastic measurable selectors (thus depending only on the present state value) that satisfy a min–max optimality equation. For the continuous-time case the definition of the admissible strategies depends on the model that it is considered. In the so-called semi-Markov case (see, for instance, [13, 14, 22]) the controlled process is defined in terms of a sequence of random decision (jump) epochs and post-jump locations, and the decisions are taken immediately after a jump. The admissible strategies for the players are transition probabilities that may depend on the whole past history of the process and actions up to the present value of the state value. Notice that in this semi-Markov case there is no motion of the process between the jumps. Another approach for the problem is to consider that the state process evolves according to a continuous-time jump Markov process (see, for instance, [8, 9]). In this formulation both players 1 and 2 observe continuously the current state of the system and whenever it is at some state  $x(t)$  they choose independently their actions  $a(t)$  and  $b(t)$  according to stochastic kernels  $\pi_t^1(\cdot|x(t))$  and  $\pi_t^2(\cdot|x(t))$ . Notice that in this case there is no dependence on previous actions and state values, and the strategies depend only on the present state value  $x(t)$ .

In this paper we consider the zero-sum games with an infinite horizon discounted reward criterion for PDMP in general Borel spaces. The two players can act on the jump rate and transition measure of the process, with the decisions being taken just after a jump of the process. We assume that the players' decisions may depend on the previous actions and the post-jump locations up to the present location. When compared with [13, 14, 22], the process considered in this paper is more general since it includes a possible flow motion between jumps and also jumps whenever the process touches the frontier. Note that a semi-Markov process can be written as a PDMP when it is “markovianized” as shown, for instance, in [21]. Indeed, as presented in [21],

pages 71–72, to markovianize a semi-Markov model taking values in a state space  $\tilde{\mathbf{X}}$  and with probability distribution function  $\tilde{F}(x, t)$  for the sojourn time at the state  $x$ , the state space has to be enlarged to  $\mathbf{X} = \tilde{\mathbf{X}} \times [0, \infty)$  so that for  $(x, t) \in \mathbf{X}$  we have that  $x$  represents the location of the process and  $t$  the elapsed sojourn time in state  $x$ . Writing this model as a PDMP, the flow would be  $\phi((x, t), s) = (x, t + s)$ , and  $\lambda(x, t)$  the failure rate of  $\tilde{F}(x, t)$  so that, in this sense, the PDMP can be seen as more general than the semi-Markov case. On the other hand in order to get our closed expressions for the min–max optimality equation we exclude from the admissible strategies the dependence on the previous inter-arrival jump times. Indeed, the basic idea of our approach is to re-write the min–max continuous-time problem in a discrete-time way, and derive the optimality equations by iterations of a kernel  $G(\cdot|x, a, b)$ , to be defined in (17), where  $x$  will represent the post-jump location and  $a$  and  $b$  the post-jump actions from players 1 and 2 respectively. Thus to get our iterative procedure through kernel  $G$  we had to exclude the dependence on the sojourn times. When compared with [8, 9] our approach has the advantage of allowing the dependence on the previous actions and post-jump location of the process, being more within the context of a game.

This paper is organized as follows. In Sect. 2 we present the notation and problem formulation. Sections 3 presents the main operators that will be required in the paper. In Sect. 4 we present some auxiliary results. In Sect. 5 we derive conditions for the existence and characterization of min–max strategies for the infinite horizon total expected discounted payoff function, which is the main result of the paper. In the Appendix we present the proof of an auxiliary result.

## 2 Notation and Problem Formulation

In this section we start by introducing in Sect. 2.1 the main notation that will be used along the paper. Section 2.2 aims at presenting the spaces and parameters related to the problem. In Sect. 2.3 we introduce the construction of the process while in Sect. 2.4 we define the set of admissible strategies and the associated conditional distribution of the controlled process.

### 2.1 Notation

The following notation will be used in this paper:  $\mathbb{N}$  is the set of natural numbers including 0,  $\mathbb{N}^* = \mathbb{N} - \{0\}$ ,  $\mathbb{R}$  denotes the set of real numbers,  $\mathbb{R}_+$  the set of non-negative real numbers,  $\mathbb{R}_+^* = \mathbb{R}_+ - \{0\}$ ,  $\mathbb{R}_+^* \cup \{+\infty\}$  and  $\overline{\mathbb{R}}_+^* = \mathbb{R}_+^* \cup \{+\infty\}$ . For  $X$  a Borel space (i.e. a Borel-measurable subset of a complete and separable metric space), we denote by  $\mathcal{B}(X)$  its associated Borel  $\sigma$ -algebra. For  $X, Y$  Borel spaces, we write  $\mathbb{M}(X, Y)$  for the space of Borel-measurable functions from  $X$  to  $Y$ . The set of Borel-measurable and Borel-measurable and bounded real-valued functions defined on the Borel space  $X$  is denoted respectively by  $\mathbb{M}(X)$  and  $\mathbb{B}(X)$ . By  $\mathbb{M}(X)_+$  we mean the set of non-negative Borel-measurable real-valued functions, and similarly for  $\mathbb{B}(X)_+$ . For  $g \in \mathbb{M}(X)$  with  $g(x) > 0$  for all  $x \in X$ ,  $\mathbb{B}_g(X)$  is the set of functions  $v \in \mathbb{M}(X)$  such that  $\sup_{x \in X} \frac{|v(x)|}{g(x)} < +\infty$ . For any set  $A$ ,  $I_A$  denotes the indicator

function of the set  $A$ .  $\mathcal{P}(X)$  is the set of probability measures defined on  $(X, \mathcal{B}(X))$ , and  $\mathcal{P}(X|Y)$  is the set of stochastic kernels on  $X$  given  $Y$  where  $Y$  denotes a Borel space. For any point  $x \in X$ ,  $\delta_x$  denotes the Dirac measure defined by  $\delta_x(\Gamma) = I_\Gamma(x)$  for any  $\Gamma \in \mathcal{B}(X)$ . If  $R$  is a kernel on  $Y$  given  $X$  and  $f \in \mathbb{M}(Y)$ , then for any  $x \in X$ ,  $Rf(x)$  denotes  $\int_Y f(y)R(dy|x)$  provided the integral exists. Finally, the infimum over an empty set is understood to be equal to  $+\infty$ , and we set  $e^{-\infty} = 0$ .

### 2.2 Preliminaries

For the definition of the state space of the PDMP we will consider for notational simplicity that  $\mathbf{X}$  is an open subset of  $\mathbb{R}^n$  ( $n \in \mathbb{N}^*$ ) with  $\partial\mathbf{X}$  denoting the boundary of  $\mathbf{X}$ , and  $\bar{\mathbf{X}}$  its closure. This definition could be easily generalized to include some boundary points and countable union of sets as in [3, Sect. 24]. In what follows the sets  $\mathbf{A}$  and  $\mathbf{B}$  are the action spaces for players 1 and 2, respectively, and assumed to be Borel spaces. For each  $x \in \mathbf{X}$ , we define the subsets  $\mathbf{A}(x)$  of  $\mathbf{A}$  and  $\mathbf{B}(x)$  of  $\mathbf{B}$  as the set of feasible control actions for players 1 and 2, respectively, that can be taken when the state process is in  $x \in \mathbf{X}$ . Let  $\mathbf{U}$  be another Borel space associated to the control process.

We introduce next some data that will be used to define the controlled PDMP.

- The flow  $\phi(x, t)$  is a function  $\phi : \mathbb{R}^n \times \mathbb{R}_+ \rightarrow \mathbb{R}^n$  continuous in  $(x, t)$  and such that  $\phi(x, t + s) = \phi(\phi(x, t), s)$ .
- For each  $x \in \mathbf{X}$ , the time the flow takes to reach the boundary starting from  $x$  is defined as

$$t_*(x) \doteq \inf\{t > 0 : \phi(x, t) \in \partial\mathbf{X}\}.$$

It is assumed that  $t_* \in \mathbb{M}(\mathbf{X}, \bar{\mathbb{R}}_+)$  (see [3, Lemma 27.1] for conditions that assure that  $t_*$  is Borel measurable). For  $x \in \mathbf{X}$  such that  $t_*(x) = \infty$  (that is, the flow starting from  $x$  never touches the boundary), we set  $\phi(x, t_*(x)) = \Delta$ , where  $\Delta$  is a fixed point in  $\partial\mathbf{X}$ .

- The jump rate  $\lambda \in \mathbb{M}(\mathbf{X} \times \mathbf{U})_+$ .
- The transition measure  $Q$  which is a stochastic kernel in  $\mathcal{P}(\mathbf{X}|\bar{\mathbf{X}} \times \mathbf{U})$ . To avoid jumps to the same point, we assume that  $Q(\{x\}|x, u) = 0$  for any  $x \in \mathbf{X}, u \in \mathbf{U}$ .
- The pre-defined control function  $\ell \in \mathbb{M}(\mathbf{X} \times \mathbf{A} \times \mathbf{B} \times \mathbb{R}_+, \mathbf{U})$ .

*Remark 2.1* The idea behind the definition above is that after a jump from a point  $x \in \mathbf{X}$  an action  $a \in \mathbf{A}(x)$  will be chosen for player 1, and similarly an action  $b \in \mathbf{B}(x)$  will be chosen for player 2. Actions  $a$  and  $b$  will parametrize the function  $\ell(x, a, b, t)$ , with  $0 \leq t \leq t_*(x)$ , which will regulate the jump rate and transition measure of the PDMP until the next jump time. Therefore in the model considered in this paper the decisions for players 1 and 2 are taken only after a jump time, and the behavior of  $\lambda$  and  $Q$  will depend on the pre-defined function  $\ell(x, a, b, t)$  for  $0 \leq t \leq t_*(x)$ .

We define  $\Xi = \{x \in \partial\mathbf{X} : x = \phi(y, t) \text{ for some } y \in \mathbf{X} \text{ and } t \in \mathbb{R}_+^*\}$ , the so called active boundary. As usual we will assume that the set  $\mathbf{K} = \{(x, a, b) : x \in \mathbf{X}, a \in \mathbf{A}(x), b \in \mathbf{B}(x)\}$  is a Borel subset of  $\bar{\mathbf{X}} \times \mathbf{A} \times \mathbf{B}$ .

### 2.3 Construction of the Process

Let  $\mathbf{X}_\infty = \mathbf{X} \cup \{x_\infty\}$ , where  $x_\infty$  is an isolated artificial point corresponding to the case when no jumps occur in the future. Similarly  $\mathbf{A}_\infty = \mathbf{A} \cup \{a_\infty\}$ ,  $\mathbf{B}_\infty = \mathbf{B} \cup \{b_\infty\}$ ,  $\mathbf{A}(x_\infty) = \{a_\infty\}$ , and  $\mathbf{B}(x_\infty) = \{b_\infty\}$  where  $a_\infty, b_\infty$  are isolated artificial actions for players 1 and 2 corresponding to the case when no jumps occur in the future. For notational convenience, we introduce  $\mathbf{K}_\infty = \{(x, a, b) : x \in \mathbf{X}_\infty, a \in \mathbf{A}(x), b \in \mathbf{B}(x)\}$ .

We put  $\Omega_n = \mathbf{X} \times (\mathbf{A} \times \mathbf{B} \times \mathbb{R}_+^* \times \mathbf{X})^n \times (\{a_\infty\} \times \{b_\infty\} \times \{\infty\} \times \{x_\infty\})^\infty$ . The canonical space denoted by  $\Omega$  is defined as  $\Omega = \bigcup_{n=1}^\infty \Omega_n \cup ((\mathbf{X} \times \mathbf{A} \times \mathbf{B} \times \mathbb{R}_+^*)^\infty)$  and is endowed with its Borel  $\sigma$ -algebra denoted by  $\mathcal{F}$ . For notational convenience,  $\omega \in \Omega$  will be represented as

$$\omega = (x_0, a_0, b_0, \theta_1, x_1, a_1, b_1, \theta_2, x_2, a_2, b_2, \theta_3, x_3, \dots).$$

Here,  $x_0 \in \mathbf{X}$  is the initial state of the controlled point process  $\xi$  with values in  $\mathbf{X}$ , defined below. For  $n \in \mathbb{N}^*$ , the components  $\theta_n > 0$  and  $x_n$  correspond to the intervals between two consecutive jumps and the values of the process immediately after jumps, and  $a_n, b_n$  the actions taken by players 1 and 2 respectively, also immediately after jumps. In case  $\theta_n < \infty$  and  $\theta_{n+1} = \infty$ , the trajectory has only  $n$  jumps, and we put  $\theta_m = \infty$  and  $x_m = x_\infty, a_m = a_\infty, b_m = b_\infty$  (artificial points) for all  $m \geq n + 1$ . Between jumps, the state of the process  $\xi$  moves according to the flow  $\phi$ .

The path up to  $n \in \mathbb{N}$  is denoted by  $h_n = (x_0, a_0, b_0, \theta_1, x_1, a_1, b_1, \theta_2, \dots, x_{n-1}, a_{n-1}, b_{n-1}, \theta_n, x_n)$  (thus excluding the decisions at  $n$ ), and the collection of all such paths is denoted by  $\mathbf{H}_n$ . For  $n \in \mathbb{N}$ , introduce the mappings  $X_n : \Omega \rightarrow \mathbf{X}_\infty$  by  $X_n(\omega) = x_n$ ,  $A_n : \Omega \rightarrow \mathbf{A}_\infty$  by  $A_n(\omega) = a_n$ ,  $B_n : \Omega \rightarrow \mathbf{B}_\infty$  by  $B_n(\omega) = b_n$  and, for  $n \geq 1$ , the mappings  $\Theta_n : \Omega \rightarrow \mathbb{R}_+^*$  by  $\Theta_n(\omega) = \theta_n$ ;  $\Theta_0(\omega) = 0$ . The sequence  $(T_n)_{n \in \mathbb{N}^*}$  of  $\mathbb{R}_+^*$ -valued mappings is defined on  $\Omega$  by  $T_n(\omega) = \sum_{i=1}^n \Theta_i(\omega) = \sum_{i=1}^n \theta_i$  and  $T_\infty(\omega) = \lim_{n \rightarrow \infty} T_n(\omega)$ . We denote by  $H_n = (X_0, A_0, B_0, \Theta_1, X_1, A_1, B_1, \dots, A_{n-1}, B_{n-1}, \Theta_n, X_n)$  the  $n$ -term random history process taking values in  $\mathbf{H}_n$  for  $n \in \mathbb{N}$ .

The random measure  $\mu$  associated with  $(\Theta_n, X_n, A_n, B_n)_{n \in \mathbb{N}}$  is a measure defined on  $\mathbb{R}_+^* \times \mathbf{X} \times \mathbf{A} \times \mathbf{B}$  by

$$\mu(\omega; dt, dx, da, db) = \sum_{n \geq 1} I_{\{T_n(\omega) < \infty\}} \delta_{(T_n(\omega), X_n(\omega), A_n(\omega), B_n(\omega))}(dt, dx, da, db).$$

Roughly speaking, for any  $\Gamma \in \mathcal{B}(\mathbb{R}_+^* \times \mathbf{X} \times \mathbf{A} \times \mathbf{B})$ ,  $\mu(\Gamma)$  gives the number of elements of the sequence  $(T_n, X_n, A_n, B_n)_{n \in \mathbb{N}}$  that are in  $\Gamma$ . For notational convenience the dependence on  $\omega$  will be suppressed and, instead of  $\mu(\omega; dt, dx, da, db)$ , it will be written  $\mu(dt, dx, da, db)$ . Moreover, we will denote the marginal of the measure  $\mu$  on  $\mathbb{R}_+^*$  by  $\mu(dt, \mathbf{X} \times \mathbf{A} \times \mathbf{B})$ , that is

$$\mu(dt, \mathbf{X} \times \mathbf{A} \times \mathbf{B}) = \sum_{n \geq 1} I_{\{T_n(\omega) < \infty\}} \delta_{T_n(\omega)}(dt). \tag{1}$$

Define  $\mathcal{F}_t = \sigma\{H_0\} \vee \sigma\{\mu([0, s] \times B) : s \leq t, B \in \mathcal{B}(\mathbf{X} \times \mathbf{A} \times \mathbf{B})\}$  for  $t \in \mathbb{R}_+$ . Finally, we define the controlled process  $\{\xi(t)\}_{t \in \mathbb{R}_+}$  and the action processes  $\{a(t)\}_{t \in \mathbb{R}_+}, \{b(t)\}_{t \in \mathbb{R}_+}$  as follows:

$$\xi(t) = \begin{cases} \phi(X_n, t - T_n) & \text{if } T_n \leq t < T_{n+1} \text{ for } n \in \mathbb{N}; \\ x_\infty, & \text{if } T_\infty \leq t, \end{cases} \tag{2}$$

$$a(t) = \begin{cases} A_n, & \text{if } T_n \leq t < T_{n+1} \text{ for } n \in \mathbb{N}; \\ a_\infty, & \text{if } T_\infty \leq t, \end{cases} \tag{3}$$

$$b(t) = \begin{cases} B_n, & \text{if } T_n \leq t < T_{n+1} \text{ for } n \in \mathbb{N}; \\ b_\infty, & \text{if } T_\infty \leq t. \end{cases} \tag{4}$$

Obviously, the process  $(\xi(t), a(t), b(t))_{t \in \mathbb{R}_+}$  can be equivalently described by the sequence of random variables  $(\Theta_n, X_n, A_n, B_n)_{n \in \mathbb{N}}$ . We define the random process  $\{u(t)\}_{t \in \mathbb{R}_+}$  taking values in  $\mathbf{U}$  as follows:

$$u(t) = \sum_{n \in \mathbb{N}} I_{\{T_n < t \leq T_{n+1}\}} \ell(X_n, A_n, B_n, t - T_n),$$

for  $t \in \mathbb{R}_+^*$ . The process  $\{u(t)\}_{t \in \mathbb{R}_+}$  is  $\{\mathcal{F}_t\}_{t \in \mathbb{R}_+}$ -predictable with values in  $\mathbf{U}$ .

### 2.4 Admissible Strategies and Conditional Distribution of the Controlled Process

In what follows we will consider strategies that depend, just after the  $n^{th}$  jump, on the past values of the post-jump location  $X_k, k = 0, \dots, X_n$ , and the previous actions  $A_k, B_k, k = 0, \dots, n - 1$ . For this we define, following the definition presented before,  $\tilde{h}_n = (x_0, a_0, b_0, x_1, a_1, b_1, \dots, x_{n-1}, a_{n-1}, b_{n-1}, x_n)$  (thus excluding the decisions at  $n$ , and all the inter-jump times  $\theta_k$ ) and by  $\tilde{\mathbf{H}}_n$  the collection of all such paths. An admissible strategy for players 1 and 2 respectively is a sequence  $\pi = (\pi_n)_{n \in \mathbb{N}}$  and  $\gamma = (\gamma_n)_{n \in \mathbb{N}}$  such that, for any  $n \in \mathbb{N}$ ,

- $\pi_n$  is a stochastic kernel on  $\mathbf{A}$  given  $\tilde{\mathbf{H}}_n$ . For  $\tilde{h}_n = (x_0, a_0, b_0, x_1, \dots, x_n) \in \tilde{\mathbf{H}}_n$ , it satisfies  $\pi_n(\mathbf{A}(x_n) | \tilde{h}_n) = 1$ .
- $\gamma_n$  is a stochastic kernel on  $\mathbf{B}$  given  $\tilde{\mathbf{H}}_n$ . For  $\tilde{h}_n = (x_0, a_0, b_0, x_1, \dots, x_n) \in \tilde{\mathbf{H}}_n$ , it satisfies  $\gamma_n(\mathbf{B}(x_n) | \tilde{h}_n) = 1$ .

For simplicity we denote the set of admissible strategies for player 1 by  $\Pi$  and the set of admissible strategies for player 2 by  $\Gamma$ .

**Definition 2.2** A randomized Markov strategy for player 1 is of the form  $p = (p_0, p_1, \dots)$  where  $p_k$  is a stochastic kernel on  $\mathbf{A}$  given  $\mathbf{X}$  satisfying  $p_k(\mathbf{A}(x) | x) = 1$  and similarly for a randomized Markov strategy for player 2, denoted by  $q = (q_0, q_1, \dots)$ . We denote the set of Markov randomized strategies for player 1 by  $\Pi^M$  and the set of randomized Markov strategies for player 2 by  $\Gamma^M$ . The case in which  $p_k$  and  $q_k$  are measurable functions is referred to as the set of deterministic Markov strategies for players 1 and 2, and denoted respectively by  $\Pi^D$  and  $\Gamma^D$ . The stationary case

corresponds to  $p_k = p$  and  $q_k = q$  for all  $k$ , where  $p$  and  $q$  are stochastic kernels on  $\mathbf{A}$  given  $\mathbf{X}$  and  $\mathbf{B}$  given  $\mathbf{X}$  respectively, satisfying  $p(\mathbf{A}(x)|x) = 1$  and  $q(\mathbf{B}(x)|x) = 1$ . This case will be denoted by  $\mathbf{P}$  and  $\mathbf{Q}$  respectively.

The cumulative jump rate  $\Lambda^{a,b}(x, t)$  is given by

$$\Lambda^{a,b}(x, t) \doteq \int_{[0,t[} \lambda(\phi(x, s), \ell(x, a, b, s)) ds, \tag{5}$$

for  $(x, a, b) \in \mathbf{K}$ , and  $t \in [0, t_*(x)]$ . With a slight abuse of notation, we denote

$$\lambda Q(A|\phi(x, t), \ell(x, a, b, t)) \doteq \lambda(\phi(x, t), \ell(x, a, b, t)) Q(A|\phi(x, t), \ell(x, a, b, t)) \tag{6}$$

for  $(x, a, b) \in \mathbf{K}$ ,  $t \in [0, t_*(x)]$ , and  $A \in \mathcal{B}(\mathbf{X})$ . Now, let us introduce the stochastic kernel  $D$  on  $\overline{\mathbb{R}}_+^* \times \mathbf{X}_\infty$  given  $\mathbf{K}_\infty$  describing the joint distribution of the next sojourn time and state of the process:

$$\begin{aligned} D(\Gamma \times S|x, a, b) &= \left[ I_{\{x=x_\infty\}} + e^{-\Lambda^{a,b}(x,+\infty)} I_{\{x \in \mathbf{X}\}} I_{\{t_*(x)=+\infty\}} \right] \delta_{(+\infty, x_\infty)}(\Gamma \times S) \\ &\quad + I_{\{x \in \mathbf{X}\}} \left[ I_{\{t_*(x) < +\infty\}} \delta_{t_*(x)}(\Gamma) Q(S|\phi(x, t_*(x)), \right. \\ &\quad \left. \ell(x, a, b, t_*(x))) e^{-\Lambda^{a,b}(x, t_*(x))} \right. \\ &\quad \left. + \int_{\Gamma \cap ]0, t_*(x)[} \lambda Q(S|\phi(x, t), \ell(x, a, b, t)) e^{-\Lambda^{a,b}(x,t)} dt \right], \tag{7} \end{aligned}$$

for  $\Gamma \in \mathcal{B}(\overline{\mathbb{R}}_+^*)$ ,  $S \in \mathcal{B}(\mathbf{X}_\infty)$ .

Roughly speaking, given  $x$  the last post-jump location of the process,  $a$  the action for player 1 and  $b$  the action for player 2, the first line in the previous equation gives the probability of the next sojourn time and the state of the process to be equal to  $(+\infty, x_\infty)$ , that is,

$$D(\{\infty \times x_\infty\}|x, a, b) = \begin{cases} e^{-\Lambda^{a,b}(x,+\infty)} & \text{for } x \in \mathbf{X}, \quad t_*(x) = +\infty, \\ 1 & \text{for } x = x_\infty. \end{cases} \tag{8}$$

The second line gives the probability of the next sojourn time to be equal to  $t_*(x)$  (corresponding to a jump at the boundary) and the state of the process to be in  $S$ , that is, for  $x \in \mathbf{X}$  such that  $t_*(x) < \infty$ , we have that

$$D(\{t_*(x)\} \times S|x, a, b) = Q(S|\phi(x, t_*(x)), \ell(x, a, b, t_*(x))) e^{-\Lambda^{a,b}(x, t_*(x))}. \tag{9}$$

The third line gives the probability of the next sojourn time to be less than  $t_*(x)$  (corresponding to a natural jump) and the state of the process to be in  $S$ , that is, for  $x \in \mathbf{X}$  and  $\tau < t_*(x)$ , we have that

$$D([0, \tau] \times S|x, a, b) = \int_0^\tau \lambda Q(S|\phi(x, t), \ell(x, a, b, t))e^{-\Lambda^{a,b}(x,t)} dt. \tag{10}$$

Consider the strategies  $(\pi, \gamma)$  for players 1 and 2 and an initial state  $x_0 \in \mathbf{X}$ . From Remark 3.43 in [12], there exists a probability  $\mathbb{P}_{x_0}^{\pi, \gamma}$  on  $(\Omega, \mathcal{F})$  and a sequence of random variables  $(\Theta_n, X_n, A_n, B_n)_{n \in \mathbb{N}}$  (or equivalently a stochastic process  $(\xi(t), a(t), b(t))_{t \in \mathbb{R}_+}$ , see Eqs. (2), (3), (4)) such that the conditional distribution of  $(\Theta_{n+1}, X_{n+1}, A_{n+1}, B_{n+1})$  given  $\mathcal{F}_{T_n}$  under  $\mathbb{P}_{x_0}^{\pi, \gamma}$  is determined by the stochastic kernel  $D_n^{\pi, \gamma}$  on  $\overline{\mathbb{R}}_+^* \times \mathbf{K}_\infty$  given  $\mathbf{H}_n$  given by

$$D_n^{\pi, \gamma}(ds, dx, da, db|h_n) := D(ds, dx|X_n, a, b)\pi_n(da|\tilde{h}_n)\gamma_n(db|\tilde{h}_n).$$

We write  $\mathbb{E}_{x_0}^{\pi, \gamma}(\cdot)$  to denote the expectation under the probability  $\mathbb{P}_{x_0}^{\pi, \gamma}$ . We can summarize the dynamics of the stochastic process  $(\xi(t), a(t), b(t))_{t \in \mathbb{R}_+}$  as follows. At time  $t = 0$  the first actions for players 1 and 2, denoted by  $A_0$  and  $B_0$ , are obtained randomly from the probability measures  $\pi_0(\cdot|x_0)$  and  $\gamma_0(\cdot|x_0)$  respectively. The first jump time  $T_1$  is a random random variable with distribution given by

$$\mathbb{P}_{x_0}^{\pi, \gamma}(T_1 > t|H_0) = \begin{cases} e^{-\Lambda^{A_0, B_0}(x, t)} & \text{for } t < t_*(x_0), \\ 0 & \text{for } t \geq t_*(x_0). \end{cases} \tag{11}$$

If  $T_1$  is equal to infinity, then  $\xi(t) = \phi(x_0, t)$ ,  $a(t) = A_0$ ,  $b(t) = B_0$  for  $t \in \mathbb{R}_+$ , and  $X_n = x_\infty$ ,  $A_n = a_\infty$ ,  $B_n = b_\infty$  for  $n \in \mathbb{N}^*$ . Otherwise select independently an  $\mathbf{X}$ -valued random variable  $X_1$  having distribution, for  $S \in \mathcal{B}(\mathbf{X})$ , given by

$$\mathbb{P}_{x_0}^{\pi, \gamma}(X_1 \in S|H_0) = Q(S|\phi(x_0, T_1), \ell(x, A_0, B_0, T_1)). \tag{12}$$

The trajectory of  $\{\xi(t)\}$  starting from  $x_0$  and for  $t \leq T_1$  is given by (2), and for  $\{a(t)\}$  and  $\{b(t)\}$  for  $t < T_1$  is as given in (3) and (4). In general, at time  $T_n$  and starting from  $X_n$ , we select the actions for players 1 and 2, denoted by  $A_n$  and  $B_n$ , randomly from the probability measures  $\pi_1(\cdot|\tilde{H}_n)$  and  $\gamma_1(\cdot|\tilde{H}_n)$  respectively, and the next inter-jump time  $T_{n+1} - T_n$  and post-jump location  $X_{n+1}$  as in (11) and (12) respectively.

The value function for the min–max problem will contain two terms, a running reward function  $f$  associated to the gradual actions of players 1 and 2, and a boundary reward function  $r$ , associated with the impulsive actions on the boundary  $\Xi$  of players 1 and 2. We assume that  $f \in \mathbb{M}(\mathbf{X}_\infty)$  and  $r \in \mathbb{M}(\Xi)$ .

The associated  $T_n$ -horizon and infinite-horizon discounted payoff criterion corresponding to strategies  $(\pi, \gamma)$  for players 1 and 2 are defined by

$$\begin{aligned} \mathcal{D}(n, \pi, \gamma, x_0) = & \mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_{]0, T_n[} e^{-\alpha t} f(\xi(t), u(t)) dt \right. \\ & \left. + \int_{]0, T_n[} e^{-\alpha t} I_{\{\xi(t-) \in \Xi\}} r(\xi(t-), u(t-)) \mu(dt, \mathbf{X} \times \mathbf{A} \times \mathbf{B}) \right], \end{aligned} \tag{13}$$



and

$$\mathcal{D}(\pi, \gamma, x_0) = \mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_{]0, \infty[} e^{-\alpha t} f(\xi(t), u(t)) dt + \int_{]0, \infty[} e^{-\alpha t} I_{\{\xi(t-) \in \Xi\}} r(\xi(t-), u(t-)) \mu(dt, \mathbf{X} \times \mathbf{A} \times \mathbf{B}) \right] \tag{14}$$

where the measure  $\mu(dt, \mathbf{X} \times \mathbf{A} \times \mathbf{B})$  has been defined in (1). In the previous expression,  $\alpha > 0$  is the discount factor,  $\mathcal{D}(n, \pi, \gamma, x_0)$  and  $\mathcal{D}(\pi, \gamma, x_0)$  are understood to be equal to  $+\infty$  if the integrals of both the positive and negative parts of the integrand are infinite. Note that, for any strategy  $\pi \in \Pi, \gamma \in \Gamma$ , the functions  $\mathcal{D}(n, \pi, \gamma, \cdot)$  and  $\mathcal{D}(\pi, \gamma, \cdot)$  are measurable. The  $T_n$ -horizon and infinite horizon lower value (denoted by the superscript  $l$ ) and upper value (denoted by the superscript  $u$ ) problems for the discounted payoff games are defined respectively as:

$$\mathcal{J}^l(n, x_0) = \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} \mathcal{D}(n, \pi, \gamma, x_0), \quad \mathcal{J}^l(x_0) = \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} \mathcal{D}(\pi, \gamma, x_0), \tag{15}$$

$$\mathcal{J}^u(n, x_0) = \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} \mathcal{D}(n, \pi, \gamma, x_0), \quad \mathcal{J}^u(x_0) = \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} \mathcal{D}(\pi, \gamma, x_0). \tag{16}$$

Clearly we have that  $\mathcal{J}^l(n, x_0) \leq \mathcal{J}^u(n, x_0)$  and  $\mathcal{J}^l(x_0) \leq \mathcal{J}^u(x_0)$ . If  $\mathcal{J}^l(n, x_0) = \mathcal{J}^u(n, x_0)$  ( $\mathcal{J}^l(x_0) = \mathcal{J}^u(x_0)$ ) then the common value is called the value of the game and denoted by  $\mathcal{V}(n, x_0)$  ( $\mathcal{V}(x_0)$  respectively). If the infinite horizon game has a value  $\mathcal{V}$  then a strategy  $\pi^* \in \Pi$  is said to be optimal for player 1 if  $\inf_{\gamma \in \Gamma} \mathcal{D}(\pi^*, \gamma, x_0) = \mathcal{V}(x_0)$  and similarly  $\gamma^* \in \Gamma$  is said to be optimal for player 2 if  $\sup_{\pi \in \Pi} \mathcal{D}(\pi, \gamma^*, x_0) = \mathcal{V}(x_0)$ . The pair  $(\pi^*, \gamma^*)$  is said to be a pair of optimal strategies if  $\pi^*$  is optimal for player 1 and  $\gamma^*$  is optimal for player 2. Similar definitions hold the finite horizon case.

### 3 Main Operators

In this section we present some important operators associated to the  $T_n$ -horizon and infinite horizon min-max problems problem posed in (15) and (16) respectively. Let us introduce the kernel  $G$  on  $\mathbf{X}_\infty$  given  $\mathbf{K}_\infty$  as follows:

$$\begin{aligned} G(A|x, a, b) &\doteq I_{\{x \in \mathbf{X}\}} \int_{\mathbb{R}_+} e^{-\alpha s} D(ds, A|x, a, b) \\ &= I_{\{x \in \mathbf{X}\}} \left[ \int_{]0, t_*(x)[} e^{-\alpha s - \Lambda^{a,b}(x,s)} \lambda Q(A \cap \mathbf{X} | \phi(x, s), \ell(x, a, b, s)) ds \right. \\ &\quad \left. + e^{-\alpha t_*(x) - \Lambda^{a,b}(x, t_*(x))} Q(A \cap \mathbf{X} | \phi(x, t_*(x)), \ell(x, a, b, t_*(x))) \right] \end{aligned} \tag{17}$$

for  $(x, a, b) \in \mathbf{K}_\infty$ ,  $A \in \mathcal{B}(\mathbf{X}_\infty)$  and the kernel  $L$  (respectively,  $H$ ) defined on  $\mathbf{X} \times \mathbf{U}$  (respectively,  $\Xi \times \mathbf{U}$ ) given  $\mathbf{K}_\infty$  as follows:

$$L(A|x, a, b) \doteq I_{\{x \in \mathbf{X}\}} \int_{]0, t_*(x)[} e^{-\alpha s - \Lambda^{a,b}(x,s)} \delta_{(\phi(x,s), \ell(x,a,b,s))}(A) ds, \tag{18}$$

$$H(B|x, a, b) \doteq I_{\{x \in \mathbf{X}\}} e^{-\alpha t_*(x) - \Lambda^{a,b}(x,t_*(x))} \delta_{(\phi(x,t_*(x)), \ell(x,a,b,t_*(x)))}(B), \tag{19}$$

for  $(x, a, b) \in \mathbf{K}_\infty$ ,  $A \in \mathcal{B}(\mathbf{X} \times \mathbf{U})$ ,  $B \in \mathcal{B}(\Xi \times \mathbf{U})$ .

*Remark 3.1* When  $t_*(x) = \infty$  for  $x \in \mathbf{X}$  we have that  $e^{-\alpha t_*(x)} = 0$  and thus the kernels  $G$  and  $H$  have a special form. Indeed in this case  $G(A|x, a, b) = \int_0^{t_*(x)} e^{-\alpha s - \Lambda^{a,b}(x,s)} \lambda Q(A|\phi(x, s), \ell(x, a, b, s)) ds$  (see the notation in (6)) and  $H(B|x, a, b) = 0$ , for  $(x, a, b) \in \mathbf{K}$ ,  $A \in \mathcal{B}(\mathbf{X})$ ,  $B \in \mathcal{B}(\Xi \times \mathbf{U})$ .

We conclude this section introducing the following notation. For  $\varrho \in \mathcal{P}(\mathbf{A}(x))$ ,  $\chi \in \mathcal{P}(\mathbf{B}(x))$  and a function  $h \in \mathbb{M}(\mathbf{K})$ , we write

$$h(x, \varrho, \chi) = \int_{\mathbf{A}(x)} \int_{\mathbf{B}(x)} h(x, a, b) \varrho(da) \chi(db). \tag{20}$$

For  $\bar{\pi} \in \mathcal{P}(\mathbf{A}_\infty|\mathbf{X}_\infty)$  and  $\bar{\gamma} \in \mathcal{P}(\mathbf{B}_\infty|\mathbf{X}_\infty)$  respectively, satisfying  $\bar{\pi}(\mathbf{A}(x)|x) = 1$  and  $\bar{\gamma}(\mathbf{B}(x)|x) = 1$ , we write

$$h(x, \bar{\pi}, \bar{\gamma}) = \int_{\mathbf{A}(x)} \int_{\mathbf{B}(x)} h(x, a, b) \bar{\pi}(da|x) \bar{\gamma}(db|x), \tag{21}$$

and for admissible strategies  $\pi = (\pi_n)_{n \in \mathbb{N}} \in \Pi$  and  $\gamma = (\gamma_n)_{n \in \mathbb{N}} \in \Gamma$  for players 1 and 2 respectively, we write

$$h(x_k, \pi_k(\cdot|\tilde{h}_k), \gamma_k(\cdot|\tilde{h}_k)) = \int_{\mathbf{A}(x)} \int_{\mathbf{B}(x)} h(x_k, a, b) \pi_k(da|\tilde{h}_k) \gamma_k(db|\tilde{h}_k), \tag{22}$$

where we recall that  $\tilde{h}_k = (x_0, a_0, b_0, x_1, a_1, b_1, \dots, x_{k-1}, a_{k-1}, b_{k-1}, x_k)$ . Following (21) and (22) we set, for  $\bar{\pi} \in \mathcal{P}(\mathbf{A}_\infty|\mathbf{X}_\infty)$  and  $\gamma = (\gamma_n)_{n \in \mathbb{N}} \in \Gamma$  an admissible strategy for player 2,

$$h(x_k, \bar{\pi}, \gamma_k(\cdot|\tilde{h}_k)) = \int_{\mathbf{A}(x)} \int_{\mathbf{B}(x)} h(x_k, a, b) \bar{\pi}(da|x_k) \gamma_k(db|\tilde{h}_k), \tag{23}$$

and similarly for  $h(x_k, \pi(\cdot|\tilde{h}_k), \bar{\gamma})$ , for the case in which  $\pi = (\pi_n)_{n \in \mathbb{N}} \in \Pi$  is an admissible strategy for player 1 and  $\bar{\gamma} \in \mathcal{P}(\mathbf{B}_\infty|\mathbf{X}_\infty)$ .

### 4 Assumptions and Auxiliary Results

The purpose of this section is to introduce the main assumptions and present some auxiliary results that will be needed for deriving our main results. The first assumption is related to an upper bound for the jump rate  $\lambda$ .

**Assumption A** There exists  $\bar{\lambda} \in \mathbb{M}(\mathbf{X})$  satisfying  $\int_0^t \bar{\lambda}(\phi(x, s)) ds < \infty$  for  $t \in [0, t_*(x))$  such that, for any  $(x, r) \in \mathbf{X} \times \mathbf{U}$ ,  $\lambda(x, r) \leq \bar{\lambda}(x)$ .

The next proposition will be used in the sequel to establish an iterative procedure to get upper and lower bounds for the payoff functions (15) and (16), using the operator  $G$  defined in (17).

**Proposition 4.1** *Suppose that Assumption A holds and that there exist Borel-measurable functions  $\mathcal{W} : \mathbf{X}_\infty \mapsto \mathbb{R}_+$ ,  $\mathcal{S} : \mathbf{X}_\infty \mapsto \mathbb{R}$  and  $\mathcal{C} : \mathbf{K}_\infty \mapsto \mathbb{R}$  and a constant  $M$  satisfying  $G\mathcal{W}(x, a, b) \leq M\mathcal{W}(x)$ ,  $|\mathcal{S}(x)| \leq M\mathcal{W}(x)$  and  $|\mathcal{C}(x, a, b)| \leq M\mathcal{W}(x)$  for any  $(x, a, b) \in \mathbf{K}_\infty$ . Consider  $x_0 \in \mathbf{X}$ ,  $\pi = (\pi_k)_{k \in \mathbb{N}} \in \Pi$  and  $\gamma = (\gamma_k)_{k \in \mathbb{N}} \in \Gamma$  such that*

$$\int_{\mathbf{A}} \int_{\mathbf{B}} \left[ G\mathcal{S}(x_k, a_k, b_k) + \mathcal{C}(x_k, a_k, b_k) \right] \pi_k(da_k | \tilde{h}_k) \gamma_k(db_k | \tilde{h}_k) \leq \mathcal{S}(x_k) \tag{24}$$

for any  $k \in \mathbb{N}$  and  $\tilde{h}_k \in \tilde{\mathbf{H}}_k$ . We have that

$$\mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_{n+1}} \mathcal{S}(X_{n+1}) \right] + \sum_{k=0}^n \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} \mathcal{C}(X_k, A_k, B_k) \right] \leq \mathcal{S}(x_0). \tag{25}$$

*Proof* See Appendix. □

Condition (24) is usually hard to be checked since it is written in terms of the operator  $G$ , which involves an integral with respect to the primitive data  $Q$  and  $\lambda$  of the process, as well as some boundary conditions. The next assumption presents an infinitesimal condition written directly in terms of the primitive data  $Q$  and  $\lambda$ , and the boundary conditions, that will be used to verify (24). But first we need to introduce the set  $\mathbb{M}^{ac}(\mathbf{X}_\infty)$  of real valued measurable functions defined on  $\mathbf{X}_\infty$  which are absolutely continuous with respect to the flow  $\phi$ , that is, the set of functions  $g \in \mathbb{M}(\mathbf{X}_\infty)$  such that for any  $x \in \mathbf{X}$ , the function  $g(\phi(x, \cdot))$  is absolutely continuous on  $[0, t_*(x)[$  and  $\lim_{t \rightarrow t_*(x)} g(\phi(x, t))$  exists whenever  $t_*(x) < \infty$ . From Lemma 2.2 in [1], if  $g \in \mathbb{M}^{ac}(\mathbf{X}_\infty)$  then there exists a real-valued measurable function  $\mathcal{X}g$  defined on  $\mathbf{X}$  satisfying

$$g(\phi(x, t)) = g(x) + \int_{[0, t]} \mathcal{X}g(\phi(x, s)) ds, \tag{26}$$

for any  $x \in \mathbf{X}$ , and  $t \in [0, t_*(x)[$ . Notice that the domain of definition of the mapping  $g \in \mathbb{M}^{ac}(\mathbf{X})$  can be extended to  $\mathbf{X}_\infty \cup \Xi$  by setting  $g(z) = \lim_{t \rightarrow t_*(x)} g(\phi(x, t))$  where  $z = \phi(x, t_*(x)) \in \Xi$ .

**Assumption B** There exist constants  $0 < c_1 < \alpha$ ,  $d_1 \geq 0$  and a function  $W \in \mathbb{M}^{ac}(\mathbf{X}_\infty)$  satisfying  $W \geq 1$  such that for all  $(x, a, b) \in \mathbf{K}$ , and  $0 \leq t < t_*(x)$ ,

$$\begin{aligned} & \mathcal{X}W(\phi(x, t)) \\ & \quad - \lambda(\phi(x, t), \ell(x, a, b, t)) \left[ W(\phi(x, t)) - QW(\phi(x, t), \ell(x, a, b, t)) \right] \\ & \leq c_1 W(\phi(x, t)) + d_1, \end{aligned} \tag{27}$$

and

$$QW(\phi(x, t_*(x)), \ell(x, a, b, t_*(x))) \leq (1 - (\alpha - c_1))W(\phi(x, t_*(x))), \tag{28}$$

whenever  $t_*(x) < \infty$ .

*Remark 4.2* Similarly as in Remark 3.2 of [8], Assumption B can be seen as an extension of the “drift condition” presented in (2.4) of [20] for PDMPs, and it is also known as a Lyapunov or Foster-Lyapunov condition. This condition is usually used to obtain growth conditions as in Proposition 4.3 below, and also for some forms of ergodicity, see, for instance, [13]. In Remark 4.5 we show that, for the continuous-time jump Markov process in Polish spaces case as considered in [8], condition (27) becomes condition (a) in Assumption 3.1 of [8].

Combining Assumptions A and B and using the operators  $L$  and  $H$  as defined in (18) and (19) we obtain a condition similar to (24) for fixed actions  $a$  and  $b$ .

**Proposition 4.3** *Suppose that Assumptions A and B hold. For any  $\bar{M}_1 > 0$  define*

$$\bar{C}(x, a, b) = LW(x, a, b) + HW(x, a, b), \tag{29}$$

and

$$S(x) = \frac{\bar{M}_1}{\alpha - c_1} W(x) + \frac{d_1 \bar{M}_1}{\alpha(\alpha - c_1)}, \tag{30}$$

where  $W$ ,  $c_1$  and  $d_1$  are as in Assumption B. Then, for any  $(x, a, b) \in \mathbf{K}_\infty$ ,

$$GS(x, a, b) + \bar{M}_1 \bar{C}(x, a, b) \leq S(x). \tag{31}$$

*Proof* From the definition of the kernels  $L$  and  $H$  in (18) and (19) we have that  $LW(x_\infty, a_\infty, b_\infty) = 0$  and  $HW(x_\infty, a_\infty, b_\infty) = 0$  since  $x_\infty \notin \mathbf{X}$ . Thus from (29) we get that  $GS(x_\infty, a_\infty, b_\infty) = \bar{C}(x_\infty, a_\infty, b_\infty) = 0$  and so (31) is trivially satisfied. Now, consider  $(x, a, b) \in \mathbf{K}$ . After some algebraic manipulations we get from (27) and (28) in Assumption B and  $S$  as defined in (30) that

$$\begin{aligned} & \mathcal{X}S(\phi(x, t)) - \alpha S(\phi(x, t)) + \bar{M}_1 W(\phi(x, t)) \\ & \quad - \lambda(\phi(x, t), \ell(x, a, b, t)) \left[ S(\phi(x, t)) - QS(\phi(x, t), \ell(x, a, b, t)) \right] \leq 0, \end{aligned} \tag{32}$$

$$\begin{aligned} & \bar{M}_1 W(\phi(x, t_*(x))) + QS(\phi(x, t_*(x)), \ell(x, a, b, t_*(x))) \\ & \leq S(\phi(x, t_*(x))) \text{ for } t_*(x) < \infty. \end{aligned} \tag{33}$$

Now, multiplying the inequality (32) by  $e^{-\alpha t - \Lambda^{a,b}(x,t)}$  and integrating over  $[0, s]$  for  $s < t_*(x)$ , we get

$$\begin{aligned} & e^{-\alpha s - \Lambda^{a,b}(x,s)} S(\phi(x, s)) - S(x) + \int_0^s e^{-\alpha t - \Lambda^{a,b}(x,t)} \lambda QS(\phi(x, t), \ell(x, a, b, t)) dt \\ & + \bar{M}_1 \int_0^s e^{-\alpha t - \Lambda^{a,b}(x,t)} W(\phi(x, t)) dt \leq 0 \end{aligned} \tag{34}$$

where we have used Assumption A to claim that

$$\begin{aligned} & e^{-\alpha s - \Lambda^{a,b}(x,s)} S(\phi(x, s)) - S(x) \\ & = \int_0^s e^{-\alpha t - \Lambda^{a,b}(x,t)} \left( \mathcal{X}S(\phi(x, t)) - [\alpha + \lambda(\phi(x, t), \ell(x, a, b, t))] S(\phi(x, t)) \right) dt. \end{aligned}$$

Consider first the case where  $t_*(x) < \infty$ . Recalling that  $S(\phi(x, \cdot))$  is absolutely continuous on  $[0, t_*(x)]$  we obtain

$$\begin{aligned} & e^{-\alpha t_*(x) - \Lambda^{a,b}(x,t_*(x))} S(\phi(x, t_*(x))) - S(x) \\ & + \int_0^{t_*(x)} e^{-\alpha t - \Lambda^{a,b}(x,t)} \lambda QS(\phi(x, t), \ell(x, a, b, t)) dt \\ & + \bar{M}_1 \int_0^{t_*(x)} e^{-\alpha t - \Lambda^{a,b}(x,t)} W(\phi(x, t)) dt \leq 0 \end{aligned}$$

by taking the limit in (34) as  $s$  tends to  $t_*(x)$ . Now, by using (33) we easily get the result.

Now let us assume that  $t_*(x) = \infty$ . Recalling that  $S$  is positive we get from (34) that

$$\begin{aligned} & \int_0^{t_*(x)} e^{-\alpha t - \Lambda^{a,b}(x,t)} \lambda QS(\phi(x, t), \ell(x, a, b, t)) dt \\ & + \bar{M}_1 \int_0^{t_*(x)} e^{-\alpha t - \Lambda^{a,b}(x,t)} W(\phi(x, t)) dt \leq S(x) \end{aligned}$$

which is precisely the claim since  $t_*(x) = \infty$  (see Remark 3.1). □

In the next assumption we consider an upper bound function  $W(x)$  for  $|f(x, u)|$  and  $|r(z, u)|$ . This function  $W(x)$  will be the same function as in Assumption B.

**Assumption C** There exists  $M_1 > 0$  such that  $|f(x, u)| \leq M_1 W(x)$  for all  $x \in \mathbf{X}$ ,  $u \in \mathbf{U}$ , and  $|r(z, u)| \leq M_1 W(z)$  for all  $z \in \mathbf{Z}$ ,  $u \in \mathbf{U}$ .

**Lemma 4.4** *Suppose that Assumptions A, B, and C hold, and consider  $\bar{C}$  as in (29). Then the real-valued function  $C$  defined on  $\mathbf{K}_\infty$  by*

$$C(x, a, b) = Lf(x, a, b) + Hr(x, a, b) \tag{35}$$

satisfies  $\sup_{(x,a,b) \in \mathbf{K}_\infty} \frac{|C(x, a, b)|}{W(x)} < \infty$ . Moreover,  $\sup_{(x,a,b) \in \mathbf{K}_\infty} \frac{\bar{C}(x, a, b) + GW(x, a, b)}{W(x)} < \infty$ .

*Proof* It is a straightforward application of Proposition 4.3 and Assumption C.  $\square$

The next assumption is needed to guarantee the convergence of the sum of some discounted payoff functions related to the infinite horizon problem (14).

**Assumption D** There exist constants  $0 < c_2 < \alpha$ ,  $d_2 \geq 0$  and a function  $W_2 \in \mathbb{M}^{ac}(\mathbf{X}_\infty)$  such that for all  $(x, a, b) \in \mathbf{K}$  with  $x \in \mathbf{X}$ , and  $0 \leq t < t_*(x)$ ,

$$\begin{aligned} & \mathcal{X}W_2(\phi(x, t)) \\ & - \lambda(\phi(x, t), \ell(x, a, b, t)) \left[ W_2(\phi(x, t)) - QW_2(\phi(x, t), \ell(x, a, b, t)) \right] \\ & \leq c_2 W_2(\phi(x, t)) + d_2, \end{aligned} \tag{36}$$

and

$$QW_2(\phi(x, t_*(x)), \ell(x, a, b, t_*(x))) \leq (1 - (\alpha - c_2))W_2(\phi(x, t_*(x))), \tag{37}$$

whenever  $t_*(x) < \infty$ . Moreover,

$$W(x) \leq M_2(LW_2(x, a, b) + HW_2(x, a, b)), \quad \text{for all } (x, a, b) \in \mathbf{K}, \tag{38}$$

for some  $M_2 > 0$  and for the function  $W$  introduced in Assumption B.

*Remark 4.5* We show in this remark that for the case in which there is no flow, that is,  $\phi(x, t) = x$  for all  $t$  (and thus there is no boundary), Assumptions B and D are similar to Assumptions 3.1(a) and Assumption 5.2(d) in [8] obtained for a continuous-time jump Markov process in Polish spaces. Indeed, for the case in which there is no motion we would have  $\phi(x, t) = x$ ,  $\ell(x, a, b, t) = (a, b)$ ,  $t_*(x) = \infty$  for all  $x \in \mathbf{X}$ ,  $t \in \mathbb{R}_+$ . For each  $x \in \mathbf{X}$ ,  $a \in \mathbf{A}(x)$ ,  $b \in \mathbf{A}(x)$ , define the signed measure  $q(\cdot|x, a, b)$  on  $\mathbb{B}(\mathbf{X})$  as

$$q(S|x, a, b) = \lambda(x, a, b)(Q(S|x, a, b) - \delta_x(S)), \quad S \in \mathbb{B}(\mathbf{X}). \tag{39}$$

Then function  $q(\cdot|x, a, b)$ , referred to as the function of transition rates in [8], satisfies the conditions  $(T_1)$ ,  $(T_2)$ ,  $(T_3)$  in [8], and  $-q(\{x\}|x, a, b) = \lambda(x, a, b)$ , so that

$$q(x) \doteq \sup_{a \in \mathbf{A}(x), b \in \mathbf{B}(x)} (-q(\{x\}|x, a, b)) = \sup_{a \in \mathbf{A}(x), b \in \mathbf{B}(x)} \lambda(x, a, b). \tag{40}$$

Note that  $W(\phi(x, t)) = W(x)$ ,  $W_2(\phi(x, t)) = W_2(x)$  for all  $t \in \mathbb{R}_+$ , so that the derivative with respect to  $t$  is zero, that is,  $\mathcal{X}W(\phi(x, t)) = 0$  and  $\mathcal{X}W_2(\phi(x, t)) = 0$ . Using the notation in (39) we have that (27) and (36) can be written respectively as  $qW(x, a, b) \leq c_1W(x) + d_1$  and  $qW_2(x, a, b) \leq c_2W_2(x) + d_2$ , which corresponds to Assumption 3.1(a) and the second part of Assumption 5.2(d) in [8]. Notice now that  $HW_2(x, a, b) = 0$  and

$$LW_2(x, a, b) = \int_0^\infty e^{-(\alpha+\lambda(x,a,b))s} ds W_2(x) = \frac{1}{\alpha + \lambda(x, a, b)} W_2(x)$$

so that (38) can be re-written as  $(\alpha + \lambda(x, a, b))W(x) \leq M_2W_2(x)$  which is equivalent, using the notation in (40), to  $(\alpha + q(x))W(x) \leq M_2W_2(x)$  for some  $M_2 > 0$ . We show next that this is equivalent to the first part of the Assumption 5.2(d) in [8], that is,  $q(x)W(x) \leq \tilde{M}_2W_2(x)$ , for some  $\tilde{M}_2 > 0$ , assuming that  $0 < q_{min} \leq q(x)$  for all  $x \in \mathbf{X}$ . In fact from  $(\alpha + q(x))W(x) \leq M_2W_2(x)$  it is immediate that  $q(x)W(x) \leq (\alpha + q(x))W(x) \leq M_2W_2(x)$ . On the other hand, if  $q(x)W(x) \leq \tilde{M}_2W_2(x)$  then  $W(x) \leq \frac{\tilde{M}_2}{q_{min}}W_2(x)$  and  $(\alpha + q(x))W(x) \leq \tilde{M}_2(1 + \frac{\alpha}{q_{min}})W_2(x) = M_2W_2(x)$  with  $M_2 = (1 + \frac{\alpha}{q_{min}})\tilde{M}_2$ , showing the equivalence.

The next result shows the convergence of the expected discounted sum of the function  $W$ , that will be used for the infinite horizon problem (14).

**Proposition 4.6** Consider Assumptions A, B and D. For  $x_0 \in \mathbf{X}$ ,  $\pi = (\pi_k)_{k \in \mathbb{N}} \in \Pi$  and  $\gamma = (\gamma_k)_{k \in \mathbb{N}} \in \Gamma$ ,  $n \in \mathbb{N}$  we have that

$$\sum_{k=0}^n \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} W(X_k) \right] \leq S_2(x_0) \tag{41}$$

where  $S_2(x) = \frac{M_2}{\alpha - c_2} W_2(x) + \frac{d_2 M_2}{\alpha(\alpha - c_2)}$ . In particular,  $\mathbb{P}_{x_0}^{\pi, \gamma} (\{T_\infty < \infty\}) = 0$  for any  $\pi = (\pi_k)_{k \in \mathbb{N}} \in \Pi$  and  $\gamma = (\gamma_k)_{k \in \mathbb{N}} \in \Gamma$ .

*Proof* Using the same arguments as in the Proposition 4.3 we get that for any  $(x, a, b) \in \mathbf{K}_\infty$   $GS_2(x, a, b) + M_2\bar{C}_2(x, a, b) \leq S_2(x)$  with  $\bar{C}_2(x, a, b) = LW_2(x, a, b) + HW_2(x, a, b)$ . Now, we can apply Proposition 4.1 to the functions  $\mathcal{W}(x) = \mathcal{S}(x) = S_2(x)$  and  $\mathcal{C}(x, a, b) = M_2\bar{C}_2(x, a, b)$  to get

$$\mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_{n+1}} S_2(X_{n+1}) \right] + M_2 \sum_{k=0}^n \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} \bar{C}_2(X_k, A_k, B_k) \right] \leq S_2(x_0).$$

Recalling that  $S_2$  is positive we obtain that

$$\sum_{k=0}^n \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} I_{\{X_k \in \mathbf{X}\}} W(X_k) \right] \leq S_2(x_0) \tag{42}$$

by using inequality (38). Since  $\mathbb{P}_{x_0}^{\pi, \gamma}(\{X_k = x_\infty\} \cap \{T_k = \infty\}) = 1$  we get that

$$\mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} I_{\{X_k = x_\infty\}} W(X_k) \right] = 0,$$

and thus we have that

$$\begin{aligned} \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} W(X_k) \right] &= \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} I_{\{X_k \in \mathbf{X}\}} W(X_k) \right] + \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} I_{\{X_k = x_\infty\}} W(X_k) \right] \\ &= \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} I_{\{X_k \in \mathbf{X}\}} W(X_k) \right]. \end{aligned} \tag{43}$$

From (42) and (43) we get (41). Now, since  $W \geq 1$ , we get from (41) and the Monotone Convergence Theorem that  $\mathbb{E}_{x_0}^{\pi, \gamma} \left[ \sum_{k=0}^\infty e^{-\alpha T_k} \right] \leq S_2(x_0)$ , implying the last part of the result.  $\square$

The following auxiliary results will be useful in the sequel, in order to re-write our min–max continuous-time problem in a discrete-time framework, in which the stages are defined by the jump times  $T_k$  of the PDMP. The first result gives an interpretation of (18), (19), in terms of the jump time  $T_1$ .

**Lemma 4.7** *Suppose that Assumptions A, B, and C hold. For  $x_0 \in \mathbf{X}$ ,  $\pi = (\pi_n)_{n \in \mathbb{N}} \in \Pi$  and  $\gamma = (\gamma_n)_{n \in \mathbb{N}} \in \Gamma$ ,  $k \in \mathbb{N}$*

$$\begin{aligned} \mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_{]T_k, T_{k+1}[} e^{-\alpha t} f(\xi(t), u(t)) dt | H_k \right] &= e^{-\alpha T_k} Lf(X_k, \pi_k(\cdot | \tilde{H}_k), \gamma_k(\cdot | \tilde{H}_k)), \\ \mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_{]T_k, T_{k+1}[} e^{-\alpha t} I_{\{\xi(t-) \in \Xi\}} r(\xi(t-), u(t-)) \mu(dt, \mathbf{X}) | H_k \right] \\ &= e^{-\alpha T_k} Hr(X_k, \pi_k(\cdot | \tilde{H}_k), \gamma_k(\cdot | \tilde{H}_k)). \end{aligned}$$

*Proof* From(8) to (10) we have that

$$\mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_0^{T_1} e^{-\alpha s} f(\phi(x_0, s), \ell(x_0, A_0, B_0, s)) ds | H_0 \right] = Lf(x_0, A_0, B_0), \tag{44}$$

$$\mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_1} I_{\{T_1 = t_*(x_0)\}} r(\phi(x_0, t_*(x_0)), \ell(x_0, A_0, B_0, t_*(x_0))) | H_0 \right] = Hr(x_0, A_0, B_0). \tag{45}$$

Moreover, denoting  $S_{k+1} = T_{k+1} - T_k$ , we have from (44) and similar reasoning as in (43) that

$$\begin{aligned} \mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_{]T_k, T_{k+1}[} e^{-\alpha t} f(\xi(t), u(t)) dt | H_k \right] \\ = \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} \int_0^{S_{k+1}} e^{-\alpha s} f(\phi(X_k, s), \ell(X_k, A_k, B_k, s)) ds | H_k \right] \end{aligned}$$



$$\begin{aligned}
 &= e^{-\alpha T_k} \mathbb{E}_{x_0}^{\pi, \gamma} \left[ Lf(X_k, A_k, B_k, s) dS | H_k \right] \\
 &= e^{-\alpha T_k} Lf(X_k, \pi_k(\cdot | \tilde{H}_k), \gamma_k(\cdot | \tilde{H}_k)).
 \end{aligned}$$

Similarly, from (45) we get that

$$\begin{aligned}
 &\mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_{]T_k, T_{k+1}[} e^{-\alpha t} I_{\{\xi(t-) \in \Xi\}} r(\xi(t-), u(t-)) \mu(dt, \mathbf{X}) | H_k \right] \\
 &= \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} \left( e^{-\alpha S_{k+1}} r(\phi(X_k, t_*(X_k)), \ell(X_k, A_k, B_k, t_*(X_k))) \right) I_{\{S_{k+1} = t_*(X_k)\}} | H_k \right] \\
 &= e^{-\alpha T_k} \mathbb{E}_{x_0}^{\pi, \gamma} \left[ Hr(X_k, A_k, B_k) | H_k \right] = e^{-\alpha T_k} Hr(X_k, \pi_k(\cdot | \tilde{H}_k), \gamma_k(\cdot | \tilde{H}_k)),
 \end{aligned}$$

completing the proof. □

The next result re-writes the payoff functions in a discrete-time fashion, using the operators  $L$  and  $H$  defined in (18) and (19) respectively.

**Proposition 4.8** Consider Assumptions A, B, C and D. For  $x_0 \in \mathbf{X}$ ,  $\pi = (\pi_k)_{k \in \mathbb{N}} \in \Pi$  and  $\gamma = (\gamma_k)_{k \in \mathbb{N}} \in \Gamma$ ,  $n \in \mathbb{N}$ , we have that

$$\begin{aligned}
 \mathcal{D}(n + 1, \pi, \gamma, x_0) &= \sum_{k=0}^n \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} C(X_k, A_k, B_k) \right] \\
 &= \sum_{k=0}^n \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} C(X_k, \pi_k(\cdot | \tilde{H}_k), \gamma_k(\cdot | \tilde{H}_k)) \right] \tag{46}
 \end{aligned}$$

and  $\mathcal{D}(\pi, \gamma, x_0) = \lim_{n \rightarrow \infty} \mathcal{D}(n, \pi, \gamma, x_0)$ .

*Proof* Recalling the definition of  $\mathcal{D}(n + 1, \pi, \gamma, x_0)$  (see Eq. (13)) and  $C$  (see Eq. (35)), we easily obtain the first part of the claim by using Lemma 4.7. Moreover, we have that

$$\begin{aligned}
 &\mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_{]0, \infty[} e^{-\alpha t} W(\xi(t)) dt + \int_{]0, \infty[} e^{-\alpha t} I_{\{\xi(t-) \in \Xi\}} W(\xi(t-)) \mu(dt, \mathbf{X}) \right], \\
 &= \lim_{n \rightarrow \infty} \mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_{]0, T_n[} e^{-\alpha t} W(\xi(t)) dt + \int_{]0, T_n[} e^{-\alpha t} I_{\{\xi(t-) \in \Xi\}} W(\xi(t-)) \mu(dt, \mathbf{X}) \right]
 \end{aligned}$$

from the Monotone Convergence Theorem and since  $\lim_{k \rightarrow \infty} T_k = \infty$ ,  $\mathbb{P}_{x_0}^{\pi, \gamma}$ -a.s. (see Lemma 4.7). Recalling the definition of  $\bar{C}$  in equation (29) we have that

$$\begin{aligned}
 &\lim_{n \rightarrow \infty} \mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_{]0, T_n[} e^{-\alpha t} W(\xi(t)) dt + \int_{]0, T_n[} e^{-\alpha t} I_{\{\xi(t-) \in \Xi\}} W(\xi(t-)) \mu(dt, \mathbf{X}) \right] \\
 &= \lim_{n \rightarrow \infty} \sum_{k=0}^n \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} \bar{C}(X_k, \pi_k(\cdot | \tilde{H}_k), \gamma_k(\cdot | \tilde{H}_k)) \right]
 \end{aligned}$$

$$\leq M \lim_{n \rightarrow \infty} \sum_{k=0}^n \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} W(X_k) \right]$$

for some positive constant  $M$  where we have used the fact that  $\bar{C}(x, a, b) \leq MW(x)$  (see Eqs. (30) and (31)) to get the last inequality. Proposition 4.6 gives that

$$\mathbb{E}_{x_0}^{\pi, \gamma} \left[ \int_{]0, \infty[} e^{-\alpha t} W(\xi(t)) dt + \int_{]0, \infty[} e^{-\alpha t} I_{\{\xi(t-) \in \Xi\}} W(\xi(t-)) \mu(dt, \mathbf{X}) \right] < \infty.$$

Consequently, by using Assumption C and the Bounded Convergence Theorem, we get the last part of the result. □

### 5 Main Results

In this section we present the main results of this paper. We start by introducing some continuity and compactness conditions on the parameters of the problem. Proposition 5.1 establishes the existence of optimal Markov strategies for the finite horizon problem (13). The infinite horizon case (14) is considered in Propositions 5.2 and 5.3. It is shown first in Proposition 5.2 the existence of a solution for the optimality equation associated to the zero-sum game problem. Proposition 5.3 establishes uniqueness of the optimality equation and the existence of optimal stationary Markov strategies.

Consider the following assumptions:

- Assumption E** (E1) The set-valued mappings defined by  $x \mapsto A(x)$  and  $x \mapsto B(x)$  defined on  $\mathbf{X}_\infty$  are Borel-measurable and compact valued.
- (E2) For each  $x \in \mathbf{X}_\infty$ ,  $C(x, a, b)$  in continuous in  $(a, b) \in \mathbf{A}(x) \times \mathbf{B}(x)$ .
- (E3) For each  $x \in \mathbf{X}_\infty$  and  $u \in \mathbb{B}(\mathbf{X}_\infty)$ ,  $Gu(x, a, b)$  is continuous in  $(a, b) \in \mathbf{A}(x) \times \mathbf{B}(x)$ .
- (E4) For each  $x \in \mathbf{X}_\infty$ ,  $GW(x, a, b)$  is continuous in  $(a, b) \in \mathbf{A}(x) \times \mathbf{B}(x)$ .

From Lemma 4.4, it follows that  $C(x, \varrho, \chi) + Gh(x, \varrho, \chi)$  is well defined for any  $x \in \mathbf{X}_\infty$ ,  $\varrho \in \mathcal{P}(\mathbf{A}(x))$  and  $\chi \in \mathcal{P}(\mathbf{B}(x))$  whenever  $h \in \mathbb{B}_W(\mathbf{X}_\infty)$ . Consequently, proceeding as in the proof of Theorem 5.1 (c) in [8], it easily follows from Assumption E that the functions  $T$  and  $R$  defined on  $\mathbb{B}_W(\mathbf{X}_\infty)$  by

$$Th(x) = \inf_{\chi \in \mathcal{P}(\mathbf{B}(x))} \sup_{\varrho \in \mathcal{P}(\mathbf{A}(x))} \left[ C(x, \varrho, \chi) + Gh(x, \varrho, \chi) \right], \tag{47}$$

$$Rh(x) = \sup_{\varrho \in \mathcal{P}(\mathbf{A}(x))} \inf_{\chi \in \mathcal{P}(\mathbf{B}(x))} \left[ C(x, \varrho, \chi) + Gh(x, \varrho, \chi) \right], \tag{48}$$

are well defined and maps  $\mathbb{B}_W(\mathbf{X}_\infty)$  into  $\mathbb{B}_W(\mathbf{X}_\infty)$ . Moreover, from Fan’s min–max theorem in [5], and the min–max measurable selection theorem in [18] and [19], there exist  $p \in \mathbf{P}$  and  $q \in \mathbf{Q}$  such that

$$Th(x) = Rh(x) = \min_{\chi \in \mathcal{P}(\mathbf{B}(x))} \max_{\varrho \in \mathcal{P}(\mathbf{A}(x))} \left[ C(x, \varrho, \chi) + Gh(x, \varrho, \chi) \right]$$

$$\begin{aligned}
 &= \max_{\varrho \in \mathcal{P}(\mathbf{A}(x))} \min_{\chi \in \mathcal{P}(\mathbf{B}(x))} \left[ C(x, \varrho, \chi) + Gh(x, \varrho, \chi) \right] \\
 &= \max_{\varrho \in \mathcal{P}(\mathbf{A}(x))} \left[ C(x, \varrho, q) + Gh(x, \varrho, q) \right] \\
 &= \min_{\chi \in \mathcal{P}(\mathbf{B}(x))} \left[ C(x, p, \chi) + Gh(x, p, \chi) \right] \\
 &= C(x, p, q) + Gh(x, p, q).
 \end{aligned} \tag{49}$$

Set recursively

$$V_k(x) = TV_{k-1}(x), \quad V_{-1}(x) = 0, \quad k = 0, 1, \dots \tag{50}$$

**Proposition 5.1** Consider Assumptions *A, B, C* and *D* and *E*. Then there exist  $p = (p_n)_{n \in \mathbb{N}} \in \Pi^M$  and  $q = (q_n)_{n \in \mathbb{N}} \in \Gamma^M$  such that

$$V_{k+1}(x) = TV_k(x) = \left[ C(x, p_k, q_k) + GV(x, p_k, q_k) \right]$$

for  $k \in \mathbb{N}$ . Moreover, the finite horizon game has a value  $\mathcal{V}(n, x)$  satisfying

$$\mathcal{V}(n, x_0) = \mathcal{D}(n + 1, p, q, x_0) = V_n(x_0) \tag{51}$$

and  $|\mathcal{V}(n, x)| \leq S(x)$  for any  $x \in \mathbf{X}$ .

*Proof* The first statements can be easily obtained by using similar arguments as in Theorem 4.1 in [15] and Eq. (49). To show that  $|\mathcal{V}(n, x)| \leq S(x)$ , we first notice that  $GS(x, a, b) + C(x, a, b) \leq S(x)$  for any  $(x, a, b) \in \mathbf{K}_\infty$  by combining (31) and Assumption C. From Proposition 4.1 considering the functions  $\mathcal{W}(x) = W(x)$ ,  $S(x) = S(x)$  and  $\mathcal{C}(x, a, b) = C(x, a, b)$  we get the desired result.  $\square$

Define now the sequence  $U_{k+1}(x) = TU_k(x)$ , with  $U_0(x) = -S(x)$ .

**Proposition 5.2** Suppose that Assumptions *A, B, C* and *D* and *E* hold. Then there exists a function  $U^* \in \mathbb{B}_W(\mathbf{X}_\infty)$  such that  $U^* = TU^* = RU^*$  and  $\lim_{k \rightarrow \infty} U_k(x) = U^*(x)$  for each  $x \in \mathbf{X}_\infty$ .

*Proof* It follows similar arguments as in the proof of Theorem 5.1 in [8]. First we show by induction on  $k$  that  $|U_k(x)| \leq S(x)$ . For  $k = 0$  it is immediate by definition. If it holds for  $k$  then, from Proposition 4.3,  $|C(x, a, b) + GU_k(x, a, b)| \leq \overline{C}(x, a, b) + GS(x) \leq S(x)$  showing that  $|U_{k+1}(x)| \leq S(x)$ . We also have from Proposition 4.3 that  $(U_k(x))_{k \in \mathbb{N}}$  is a pointwise non-decreasing sequence of functions since  $C(x, a, b) + GU_0(x, a, b) \geq -\overline{C}(x, a, b) - GS(x) \geq -S(x)$ , and thus  $U_1(x) \geq U_0(x)$ , and the operator  $T$  is monotone. From this we have that there exists  $U^*(x) = \lim_{k \rightarrow \infty} U_k(x) \leq S(x)$ , and so  $U^* \in \mathbb{B}_W(\mathbf{X}_\infty)$ . Since  $T$  is monotone and  $U^* \geq U_k$  it follows that  $TU^* \geq TU_k = U_{k+1}$ , which shows that  $TU^* \geq U^*$ . From (49) there exists  $q_n \in \mathbf{Q}$  such that for any  $\varrho' \in \mathcal{P}(\mathbf{A}(x))$ ,

$$U^*(x) \geq U_{n+1}(x) = TU_n(x) = \min_{\chi \in \mathcal{P}(\mathbf{B}(x))} \max_{\varrho \in \mathcal{P}(\mathbf{A}(x))} \left[ C(x, \varrho, \chi) + GU_n(x, \varrho, \chi) \right]$$

$$= \max_{\varrho \in \mathcal{P}(\mathbf{A}(x))} \left[ C(x, \varrho, q_n) + GU_n(x, \varrho, q_n) \right] \geq \left[ C(x, \varrho', q_n) + GU_n(x, \varrho', q_n) \right]. \tag{52}$$

Since  $\mathcal{P}(\mathbf{B}(x))$  is compact it can be assumed without loss of generality that  $q_n(\cdot|x) \rightarrow \chi'$  as  $n \rightarrow \infty$  for some  $\chi' \in \mathcal{P}(\mathbf{B}(x))$ . From the extended Fatou’s lemma (see Lemma 8.3.7 in [11]) and the continuity assumptions made we get from (52) that

$$\begin{aligned} U^*(x) &\geq \overline{\lim}_{n \rightarrow \infty} \left[ C(x, \varrho', q_n) + GU_n(x, \varrho', q_n) \right] \geq C(x, \varrho', \chi') + GU^*(x, \varrho', \chi') \\ &\geq \min_{\chi \in \mathcal{P}(\mathbf{B}(x))} \left[ C(x, \varrho', \chi) + GU^*(x, \varrho', \chi) \right]. \end{aligned} \tag{53}$$

From (53) it follows that  $U^*(x) \geq RU^*(x) = TU^*(x)$  completing the proof.  $\square$

**Proposition 5.3** *Suppose that Assumptions A, B, C and D and E hold and consider  $U^*$  as in Proposition 5.2. Then  $U^*$  is the unique solution for  $V = TV$  with  $V \in \mathbb{B}_W(\mathbf{X}_\infty)$ . Moreover the discounted infinite horizon game has a value  $\mathcal{V}$  satisfying*

$$\mathcal{V}(x_0) = U^*(x_0) = \mathcal{D}(p^*, q^*, x_0)$$

where the pair of optimal strategies  $(p^*, q^*) \in \mathbf{P} \times \mathbf{Q}$  is such that  $U^*(x) = C(x, p^*, q^*) + GU^*(x, p^*, q^*)$ .

*Proof* Let  $V \in \mathbb{B}_W(\mathbf{X}_\infty)$  satisfy  $V = TV$ . The idea of the proof is to show that the discounted infinite horizon game has a value  $\mathcal{V}$ , and that  $\mathcal{V} = V$ , so that  $\mathcal{V}$  is the unique solution of  $V = TV$  with  $V \in \mathbb{B}_W(\mathbf{X}_\infty)$ , and from Proposition 5.2 we have that  $\mathcal{V} = U^*$ . Consider any strategy  $\pi \in \Pi$  for player 1 and set  $q^* \in \mathbf{Q}$  such that (see (49))

$$\begin{aligned} V(x) = TV(x) &= \min_{\chi \in \mathcal{P}(\mathbf{B}(x))} \max_{\varrho \in \mathcal{P}(\mathbf{A}(x))} \left[ C(x, \varrho, \chi) + GV(x, \varrho, \chi) \right] \\ &= \max_{\varrho \in \mathcal{P}(\mathbf{A}(x))} \left[ C(x, \varrho, q^*) + GV(x, \varrho, q^*) \right] \geq C(x, \varrho, q^*) + GV(x, \varrho, q^*) \end{aligned} \tag{54}$$

for any  $\varrho \in \mathcal{P}(\mathbf{A}(x))$ . From Proposition 4.1 considering the functions  $\mathcal{W}(x) = W(x)$ ,  $\mathcal{S}(x) = V(x)$ ,  $\mathcal{C}(x, a, b) = C(x, a, b)$ , and using the inequality  $V(x) \geq C(x, \varrho, q^*) + GV(x, \varrho, q^*)$  for any  $\varrho \in \mathcal{P}(\mathbf{A}(x))$  obtained from (54), we get that

$$\mathbb{E}_{x_0}^{\pi, q^*} \left[ e^{-\alpha T_{n+1}} V(X_{n+1}) \right] + \sum_{k=0}^n \mathbb{E}_{x_0}^{\pi, q^*} \left[ e^{-\alpha T_k} C(X_k, \pi_k(\cdot|\tilde{H}_k), q^*) \right] \leq V(x_0). \tag{55}$$

Taking the limit in (55) as  $n \rightarrow \infty$  we obtain from Proposition 4.6 that

$$\begin{aligned} \mathcal{D}(\pi, q^*, x_0) \leq V(x_0) &\Rightarrow \mathcal{J}^u(x_0) = \inf_{\gamma \in \Gamma} \sup_{\pi \in \Pi} \mathcal{D}(\pi, \gamma, x_0) \\ &\leq \sup_{\pi \in \Pi} \mathcal{D}(\pi, q^*, x_0) \leq V(x_0). \end{aligned} \tag{56}$$

Similarly consider any strategy  $\gamma \in \Gamma$  for player 2 and set  $p^* \in \mathbf{P}$  such that (see (49))

$$\begin{aligned} V(x) &= TV(x) = RV(x) = \max_{\varrho \in \mathcal{P}(\mathbf{A}(x))} \min_{\chi \in \mathcal{P}(\mathbf{B}(x))} \left[ C(x, \varrho, \chi) + GV(x, \varrho, \chi) \right] \\ &= \min_{\chi \in \mathcal{P}(\mathbf{B}(x))} \left[ C(x, p^*, \chi) + GV(x, p^*, \chi) \right] \leq C(x, p^*, \chi) + GV(x, p^*, \chi) \end{aligned} \tag{57}$$

for any  $\chi \in \mathcal{P}(\mathbf{B}(x))$ . From Proposition 4.1 considering the functions  $\mathcal{W}(x) = W(x)$ ,  $\mathcal{S}(x) = -V(x)$ ,  $\mathcal{C}(x, a, b) = -C(x, a, b)$ , and using the inequality  $V(x) \leq C(x, p^*, \chi) + GV(x, p^*, \chi)$  for any  $\chi \in \mathcal{P}(\mathbf{B}(x))$  obtained from (57), we get that

$$\mathbb{E}_{x_0}^{p^*, \gamma} \left[ e^{-\alpha T_{n+1}} V(X_{n+1}) \right] + \sum_{k=0}^n \mathbb{E}_{x_0}^{p^*, \gamma} \left[ e^{-\alpha T_k} C(X_k, p^*, \gamma_k(\cdot | \tilde{H}_k)) \right] \geq V(x_0). \tag{58}$$

Taking the limit in (58) as  $n \rightarrow \infty$  and from Proposition 4.6 we get that

$$\begin{aligned} \mathcal{D}(p^*, \gamma, x_0) \geq V(x_0) &\Rightarrow \mathcal{J}^l(x_0) = \sup_{\pi \in \Pi} \inf_{\gamma \in \Gamma} \mathcal{D}(\pi, \gamma, x_0) \\ &\geq \inf_{\gamma \in \Gamma} \mathcal{D}(p^*, \gamma, x_0) \geq V(x_0). \end{aligned} \tag{59}$$

From (56) and (59) we get that

$$V(x_0) \leq \mathcal{J}^l(x_0) \leq \mathcal{J}^u(x_0) \leq V(x_0)$$

and thus the discounted infinite horizon game has a value  $\mathcal{V}$ , and  $\mathcal{V} = V$ . Therefore for any  $V \in \mathbb{B}_W(\mathbf{X}_\infty)$  satisfying  $V = TV$  we have that  $V = \mathcal{V}$ , which shows that  $\mathcal{V}$  is the unique fixed point solution of  $V = TV$  with  $V \in \mathbb{B}_W(\mathbf{X}_\infty)$ . From Proposition 5.2 we get that  $\mathcal{V}(x) = U^*(x)$ . Moreover by taking  $\pi = p^*$  in (59) and  $\gamma = q^*$  in (56) we get that  $\mathcal{V}(x_0) = \mathcal{D}(p^*, q^*, x_0)$  completing the proof.  $\square$

**Acknowledgements** The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve the paper. This work was partially supported by FAPESP (Research Council of the State of São Paulo) Grant 2013/50759-3. O.L.V. Costa received financial support from CNPq (Brazilian National Research Council), Grant 304091/2014-6, project INCT under the Grant CNPq 465755/2014-3, FAPESP 2014/50851-0, and FAPESP/BG Brasil through the Research Centre for Gas Innovation, FAPESP Grant 2014/50279-4.

### Appendix: Proof of Proposition 4.1

For the proof of this proposition, we need first to derive some auxiliary technical results. In what follows we write for notational convenience  $\tilde{h}_k = (\tilde{h}_{k-1}, a_{k-1}, b_{k-1}, x_k)$  and we introduce

$$v_n^n(\tilde{h}_n) = C(x_n, \pi_n(\cdot | \tilde{h}_n), \gamma_n(\cdot | \tilde{h}_n)), \tag{60}$$

$$v_{k-1}^n(\tilde{h}_{k-1}) = \int_{\mathbf{A}} \int_{\mathbf{B}} \int_{\mathbf{X}} v_k^n((\tilde{h}_{k-1}, a_{k-1}, b_{k-1}, x_k)) G(dx_k | x_{k-1}, a_{k-1}, b_{k-1}) \pi_{k-1}(da_{k-1} | \tilde{h}_{k-1}) \gamma_{k-1}(db_{k-1} | \tilde{h}_{k-1}) \tag{61}$$

and

$$s_n^n(\tilde{h}_n) = G\mathcal{S}(x_n, \pi_n(\cdot | \tilde{h}_n), \gamma_n(\cdot | \tilde{h}_n)), \tag{62}$$

$$s_{k-1}^n(\tilde{h}_{k-1}) = \int_{\mathbf{A}} \int_{\mathbf{B}} \int_{\mathbf{X}} s_k^n((\tilde{h}_{k-1}, a_{k-1}, b_{k-1}, x_k)) G(dx_k | x_{k-1}, a_{k-1}, b_{k-1}) \pi_{k-1}(da_{k-1} | \tilde{h}_{k-1}) \gamma_{k-1}(db_{k-1} | \tilde{h}_{k-1}) \tag{63}$$

for  $\pi = (\pi_k)_{k \in \mathbb{N}} \in \Pi$ ,  $\gamma = (\gamma_k)_{k \in \mathbb{N}} \in \Gamma$  and  $n \in \mathbb{N}$ ,  $k \in \mathbb{N}_n^*$  (for simplicity we omit the dependence on  $(\pi, \gamma)$  for the functions  $v_k^n, s_k^n$  defined below). Observe that for any  $n \in \mathbb{N}$ ,  $k \in \mathbb{N}_n$  and  $\tilde{h}_k \in \tilde{\mathbf{H}}_k$ ,  $v_k^n(\tilde{h}_k)$  and  $s_k^n(\tilde{h}_k)$  are well defined by using the hypotheses on  $\mathcal{W}, \mathcal{S}$  and  $\mathcal{C}$ .

**Proposition 6.1** For  $x_0 \in \mathbf{X}$ ,  $\pi = (\pi_k)_{k \in \mathbb{N}} \in \Pi$  and  $\gamma = (\gamma_k)_{k \in \mathbb{N}} \in \Gamma$ ,  $m \in \mathbb{N}$ , we have that

$$\begin{aligned} \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha \Theta_{m-k}} v_{m-k}^m(\tilde{H}_{m-k}) | H_{m-(k+1)} \right] &= v_{m-(k+1)}^m(\tilde{H}_{m-(k+1)}), \text{ fork } \in \mathbb{N}_{m-1}, \\ \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha \Theta_{m+1-k}} s_{m+1-k}^m(\tilde{H}_{m+1-k}) | H_{m-k} \right] &= s_{m-k}^m(\tilde{H}_{m-k}), \text{ fork } \in \mathbb{N}_m. \end{aligned}$$

*Proof* It is an immediate application of the construction of the process. □

As a consequence of the previous proposition, we have the following result.

**Proposition 6.2** For  $x_0 \in \mathbf{X}$ ,  $\pi = (\pi_k)_{k \in \mathbb{N}} \in \Pi$  and  $\gamma = (\gamma_k)_{k \in \mathbb{N}} \in \Gamma$ ,  $m \in \mathbb{N}$  we have that

$$\mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_m} \mathcal{C}(X_m, \pi_m(\cdot | \tilde{H}_m), \gamma_m(\cdot | \tilde{H}_m)) \right] = v_0^m(x_0, \pi_0, \gamma_0), \tag{64}$$

$$\mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_{m+1}} \mathcal{S}(X_{m+1}) \right] = s_0^m(x_0, \pi_0, \gamma_0). \tag{65}$$

*Proof* By definition, we have

$$\begin{aligned} &\mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_m} \mathcal{C}(X_m, \pi_m(\cdot | \tilde{H}_m), \gamma_m(\cdot | \tilde{H}_m)) \right] \\ &= \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_{m-1}} \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha \Theta_m} v_m^m(\tilde{H}_m) | H_{m-1} \right] \right], \end{aligned}$$

and so from Proposition 6.1,

$$\begin{aligned} &\mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_m} \mathcal{C}(X_m, \pi_m(\cdot | \tilde{H}_m), \gamma_m(\cdot | \tilde{H}_m)) \right] \\ &= \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_{m-1}} v_{m-1}^m(\tilde{H}_{m-1}) \right]. \end{aligned}$$

Repeating this procedure we get (64), and similarly we get (65). □

For notational convenience, let us define

$$g_k^n(\tilde{h}_k) = \sum_{i=k}^n v_k^i(\tilde{h}_k). \tag{66}$$

**Proposition 6.3** *For  $k \in \mathbb{N}_n$ , we have that*

$$s_k^n(\tilde{h}_k) + g_k^n(\tilde{h}_k) \leq \mathcal{S}(x_k). \tag{67}$$

*Proof* Let us prove (67) by induction on  $k$ . For  $k = n$  we have from (60), (62) and (66) that  $s_n^n(\tilde{h}_n) = G\mathcal{S}(x_n, \pi(\cdot|\tilde{h}_n), \gamma_n(\cdot|\tilde{h}_n))$ ,  $g_n^n(\tilde{h}_n) = v_n^n(\tilde{h}_n) = \mathcal{C}(x_n, \pi_n(\cdot|\tilde{h}_n), \gamma_n(\cdot|\tilde{h}_n))$  and thus from (24),

$$\begin{aligned} & s_n^n(\tilde{h}_n) + g_n^n(\tilde{h}_n) \\ &= \int_{\mathbf{A}} \int_{\mathbf{B}} \left( G\mathcal{S}(x_n, a_n, b_n) + \mathcal{C}(x_n, a_n, b_n) \right) \pi_n(da_n|\tilde{h}_n), \gamma_n(db_n|\tilde{h}_n) \leq \mathcal{S}(x_n). \end{aligned}$$

This proves the result for  $n$ . Suppose (67) holds for  $k$ . Let us show that it also holds for  $k - 1$ . We have from (60), (61), (63), (66), the induction hypothesis (67), and (24), that

$$\begin{aligned} & \int_{\mathbf{A}} \int_{\mathbf{B}} \int_{\mathbf{X}} \left( s_k^n((\tilde{h}_{k-1}, a_{k-1}, b_{k-1}, x_k)) \right. \\ & \quad \left. + g_k^n((\tilde{h}_{k-1}, a_{k-1}, b_{k-1}, x_k)) \right) G(dx_k|x_{k-1}, a_{k-1}, b_{k-1}) \\ & \quad \pi_{k-1}(da_{k-1}|\tilde{h}_{k-1}) \gamma_{k-1}(db_{k-1}|\tilde{h}_{k-1}) \\ &= \int_{\mathbf{A}} \int_{\mathbf{B}} \int_{\mathbf{X}} s_k^n((\tilde{h}_{k-1}, a_{k-1}, b_{k-1}, x_k)) \\ & \quad G(dx_k|x_{k-1}, a_{k-1}, b_{k-1}) \pi_{k-1}(da_{k-1}|\tilde{h}_{k-1}) \gamma_{k-1}(db_{k-1}|\tilde{h}_{k-1}) \\ & \quad + \sum_{i=k}^n \int_{\mathbf{A}} \int_{\mathbf{B}} \int_{\mathbf{X}} v_k^i((\tilde{h}_{k-1}, a_{k-1}, b_{k-1}, x_k)) \\ & \quad G(dx_k|x_{k-1}, a_{k-1}, b_{k-1}) \pi_{k-1}(da_{k-1}|\tilde{h}_{k-1}) \gamma_{k-1}(db_{k-1}|\tilde{h}_{k-1}) \\ &= s_{k-1}^n(\tilde{h}_{k-1}) + \sum_{i=k}^n v_{k-1}^i(\tilde{h}_{k-1}) \\ & \leq \int_{\mathbf{A}} \int_{\mathbf{B}} \int_{\mathbf{X}} \mathcal{S}(x_k) \\ & \quad G(dx_k|x_{k-1}, a_{k-1}, b_{k-1}) \pi_{k-1}(da_{k-1}|\tilde{h}_{k-1}) \gamma_{k-1}(db_{k-1}|\tilde{h}_{k-1}) \\ &= \int_{\mathbf{A}} \int_{\mathbf{B}} \\ & \quad G\mathcal{S}(x_{k-1}, a_{k-1}, b_{k-1}) \pi_{k-1}(da_{k-1}|\tilde{h}_{k-1}) \gamma_{k-1}(db_{k-1}|\tilde{h}_{k-1}) \\ & \leq \int_{\mathbf{A}} \int_{\mathbf{B}} \left( \mathcal{S}(x_{k-1}) - \mathcal{C}(x_{k-1}, a_{k-1}, b_{k-1}) \right) \end{aligned}$$

$$\begin{aligned}
& \pi_{k-1}(da_{k-1}|\tilde{h}_{k-1})\gamma_{k-1}(db_{k-1}|\tilde{h}_{k-1}) \\
&= \mathcal{S}(x_{k-1}) - \mathcal{C}(x_{k-1}, \pi(\cdot|\tilde{h}_{k-1}), \gamma(\cdot|\tilde{h}_{k-1})) \\
&= \mathcal{S}(x_{k-1}) - v_{k-1}^{k-1}(\tilde{h}_{k-1}),
\end{aligned}$$

where the last inequality follows from (24). Thus re-arranging the terms we get that  $s_{k-1}^n(\tilde{h}_{k-1}) + \sum_{i=k-1}^n v_{k-1}^i(\tilde{h}_{k-1}) \leq \mathcal{S}(x_{k-1})$  showing (67) for  $k-1$ , completing the proof.  $\square$

Now the proof of Proposition 4.1 is a straightforward consequence of Propositions 6.2 and 6.3. From (67), we have  $s_0^n(\tilde{h}_0) + g_0^n(\tilde{h}_0) \leq \mathcal{S}(x_0)$ . Moreover, combining (64) and (66) we get  $g_0^n(\tilde{h}_0) = \sum_{k=0}^n \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_k} \mathcal{C}(X_k, \pi_k(\cdot|\tilde{H}_k), \gamma_k(\cdot|\tilde{H}_k)) \right]$  and from (65),  $s_0^n(x_0, \pi_0, \gamma_0) = \mathbb{E}_{x_0}^{\pi, \gamma} \left[ e^{-\alpha T_{n+1}} \mathcal{S}(X_{n+1}) \right]$  giving the result.

## References

- Costa, O.L.V., Dufour, F.: Continuous average control of piecewise deterministic Markov processes. Springer, New York (2013)
- Davis, M.H.A.: Piecewise-deterministic Markov processes: a general class of non-diffusion stochastic models. *J. R. Stat. Soc. (B)* **46**(3), 353–388 (1984)
- Davis, M.H.A.: Markov Models and Optimization. Chapman and Hall, London (1993)
- Davis, M.H.A., Dempster, M.A.H., Sethi, S.P., Vermes, D.: Optimal capacity expansion under uncertainty. *Adv. Appl. Probab.* **19**(1), 156–176 (1987)
- Fan, K.: Minimax theorems. *Proc. Natl. Acad. Sci. USA* **39**, 42–47 (1953)
- Filar, J.A., Vrieze, K.: Competitive Markov decision processes. Springer, New York (1997)
- González-Trejo, J.I., Hernández-Lerma, O., Hoyos-Reyes, L.F.: Minimax control of discrete-time stochastic systems. *SIAM J. Control Optim.* **41**, 1626–1659 (2003)
- Guo, X., Hernandez-Lerma, O.: Zero-sum games for continuous-time jump Markov processes in Polish spaces: discounted payoffs. *Adv. Appl. Probab.* **39**, 646–668 (2007)
- Guo, X.P., Hernández-Lerma, O.: New optimality conditions for average-payoff continuous-time Markov games in Polish spaces. *Sci. China Math.* **54**, 793–816 (2011)
- Hernández-Lerma, O., Lasserre, J.B.: Zero-sum stochastic games in Borel spaces: average payoff criterion. *SIAM J. Control Optim.* **39**, 1520–1539 (2001)
- Hernández-Lerma, O., Lasserre, J.B.: Further Topics on Discrete-Time Markov Control Processes. Applications of Mathematics, vol. 42. Springer, New York (1999)
- Jacod, J.: Calcul stochastique et problèmes de martingales. Lecture Notes in Mathematics, vol. 714. Springer, Berlin (1979)
- Jaśkiewicz, A.: Zero-sum semi-Markov games. *SIAM J. Control Optim.* **41**, 723–739 (2002)
- Jaśkiewicz, A.: Zero-sum ergodic semi-Markov games with weakly continuous transition probabilities. *J. Optim. Theory Appl.* **141**, 321–347 (2009)
- Jaśkiewicz, A., Nowak, A.S.: Zero-sum ergodic stochastic games with Feller transition probabilities. *SIAM J. Control Optim.* **45**(3), 773–789 (2006)
- Jaśkiewicz, A., Nowak, A.S.: Stochastic games with unbounded payoffs: applications to robust control in economics. *Dyn. Games Appl.* **1**, 253–279 (2011)
- Kuenle, Heinz-Uwe: On Markov games with average reward criterion and weakly continuous transition probabilities. *SIAM J. Control Optim.* **45**, 2156–2168 (2007)
- Nowak, A.S.: Measurable selection theorems for minimax stochastic optimization problems. *SIAM J. Control Optim.* **23**, 466–476 (1985)
- Rieder, U.: On semi-continuous dynamic games. Technical report, University of Karlsruhe, Karlsruhe, Germany, 1978
- Tweedie, Richard L., Lund, Robert B., Meyn, Sean P.: Computable exponential convergence rates for stochastically ordered Markov processes. *Ann. Appl. Probab.* **6**(1), 218–237 (1996)



21. Van der Duyn Schouten, F.A.: Markov decision drift processes. In: Janssen, J. (ed.) *Semi-Markov Models: Theory and Applications*, Chapter 2, pp. 63–78. Springer, New York (1984)
22. Vega-Amaya, O.: Zero-sum average semi-Markov games: fixed-point solutions of the Shapley equation. *SIAM J. Control Optim.* **42**, 1876–1894 (2003)