

ZERO-SUM GAMES FOR CONTINUOUS-TIME JUMP MARKOV PROCESSES IN POLISH SPACES: DISCOUNTED PAYOFFS

XIANPING GUO,* *Zhongshan University*

ONÉSIMO HERNÁNDEZ-LERMA,** *CINVESTAV-IPN*

Abstract

This paper is devoted to the study of two-person zero-sum games for continuous-time jump Markov processes with a discounted payoff criterion. The state and action spaces are all Polish spaces, the transition rates are allowed to be *unbounded*, and the payoff rates may have *neither upper nor lower bounds*. We give conditions on the game's *primitive data* under which the existence of a solution to the Shapley equation is ensured. Then, from the Shapley equation, we obtain the existence of the value of the game and of a pair of optimal stationary strategies using the *extended infinitesimal operator* associated with the transition function of a possibly nonhomogeneous continuous-time jump Markov process. We also provide a *recursive* way of computing (or at least approximating) the value of the game. Moreover, we present a 'martingale characterization' of a pair of optimal stationary strategies. Finally, we apply our results to a controlled birth and death system and a Schlögl first model, and then we use controlled Potlach processes to illustrate our conditions.

Keywords: Zero-sum game; jump Markov process; pair of optimal stationary strategies; martingale characterization

2000 Mathematics Subject Classification: Primary 91A15; 91A25; 60J27

1. Introduction

In this paper we are concerned with a class of *zero-sum stochastic dynamic games*. These games have been widely studied in the literature. The existing works can be roughly classified into three main groups. The first group deals with *discrete-time games* (see, for instance, [6], [15], [23], [26], [28], and the references therein), in which the evolution of the states of the game is described by a discrete-time Markov chain. The second group deals with *stochastic differential games* (see, for instance, [2], [7], [17], and the references therein), in which the states evolve according to a stochastic differential equation. The third group deals with *semi-Markov games* in which the state of the game is described by a semi-Markov process and the players can choose their actions only at *certain* (random) epochs; see, for instance, [16], [19], [22], [29], and the references therein. Here we study a *fourth* class of stochastic games, namely, continuous-time Markov games in which the state process evolves according to a continuous-time jump Markov process. The latter class has been considered in [12], [13], and [18]. However, the

Received 7 March 2006; revision received 26 July 2007.

* Postal address: School of Mathematics and Computational Science, Zhongshan University, Guangzhou, 510275, P. R. China. Email address: mcsgxp@mail.sysu.edu.cn

Research supported by the NSFC and RFDP.

** Postal address: Department of Mathematics, CINVESTAV-IPN, A. Postal 14-740, México D.F. 07000, Mexico.

Email address: ohernand@math.cinvestav.mx

Research partially supported by CONACyT Grant 45693-F.

treatments in [12] and [13] are restricted to the case of a denumerable state space; and although the state space in [18] is general, the concluding remarks in [22] raise questions as to whether the main results are correct.

In this paper we study the more general case in which the state and action spaces are all *Polish spaces*, the transition rates are allowed to be *unbounded*, the payoff rates may have *neither upper nor lower bounds*, and, moreover, the players can select their actions *continuously* in time. The main motivation for studying this general case stems from the zero-sum nature of many real-world situations, such as inventory/production processes, telecommunication systems, water reservoir systems, controlled queueing networks, and some models of chemical reactions. In such systems the payoff rates may be unbounded, and the state or the action space may be nondenumerable.

More precisely, we deal with stochastic games specified by four *primitive data*: a state space S ; control constraint sets $A(x)$ for player 1 and $B(x)$ for player 2, which may depend on the current state $x \in S$; transition rates $q(\cdot \mid x, a, b)$; and payoff rates $r(x, a, b)$. The state space S is a Polish space, the transition rates $q(\cdot \mid x, a, b)$ may be *unbounded*, and the payoff rates $r(x, a, b)$ may have *neither upper nor lower bounds*. We impose suitable conditions on these *primitive data*. Under these conditions, we first construct a probability space on which we introduce the game's state and action processes as well as some of their fundamental properties. We then prove that the Shapley (or dynamic programming or Hamilton–Jacobi–Bellman) equation has a solution. This solution is used to show that the game has a value, and that a pair of optimal stationary strategies exists. In addition, we provide a *recursive* way of computing (or at least approximating) the value of the game, and also present a ‘martingale characterization’ of pairs of optimal stationary strategies. Moreover, our conditions are illustrated by generalized Potlach processes with controlled parameters [3], and our results are applied to controlled birth and death systems and to a Schlögl first model. A key feature of this paper is that our assumptions are imposed on the game's *primitive data*.

The rest of the paper is organized as follows. In Section 2 we introduce the game model we are interested in. In Section 3 we define the family of admissible strategies, and present some properties of the state and the action processes. The optimality criterion we are concerned with is presented in Section 4. Our main optimality results are stated in Section 5, and illustrated with examples in Section 6; their proofs are postponed to Section 7.

2. The game model

We begin this section by introducing some notation. If X is a Polish space (that is, a complete and separable metric space), we denote by $\mathcal{B}(X)$ its Borel σ -algebra and by $\mathcal{P}(X)$ the set of probability measures on $\mathcal{B}(X)$, endowed with the topology of weak convergence.

In this section we introduce the following (continuous-time, time-homogeneous) two-person zero-sum stochastic game model:

$$\{S, (A(x) \subseteq A, B(x) \subset B, x \in S), q(\cdot \mid x, a, b), r(x, a, b)\}, \quad (2.1)$$

where S is the *state space*, a Polish space, $A(x)$ and $B(x)$ are the (nonempty) Borel *sets of admissible actions* for players 1 and 2 in state x , and A and B are Polish spaces representing the action spaces for players 1 and 2, respectively. Let

$$K := \{(x, a, b) \mid x \in S, a \in A(x), b \in B(x)\},$$

which is a Borel subset of $S \times A \times B$ [21].

The function $q(\cdot \mid x, a, b)$ in (2.1) is the function of *transition rates* and satisfies the following properties.

(T₁) For each fixed $(x, a, b) \in K$, $q(\cdot \mid x, a, b)$ is a signed measure on $\mathcal{B}(S)$, whereas for each fixed $D \in \mathcal{B}(S)$, $q(D \mid \cdot)$ is a real-valued Borel-measurable function on K ;

(T₂) $0 \leq q(D \mid x, a, b) < \infty$ for all $(x, a, b) \in K$ and $x \notin D \in \mathcal{B}(S)$; and

(T₃) $q(S \mid x, a, b) = 0$ and $0 \leq -q(\{x\} \mid x, a, b) < \infty$ for all $(x, a, b) \in K$.

Furthermore, the model in (2.1) is assumed to be *stable*, which means that

$$q(x) := \sup_{(a,b) \in A(x) \times B(x)} (-q(\{x\} \mid x, a, b)) < \infty \quad \text{for each } x \in S. \tag{2.2}$$

Finally, $r(x, a, b)$ is a real-valued function on K that stands for the reward rate function for player 1 (or the cost rate function for player 2).

The game is played as follows. Players 1 and 2 *continuously* observe the current state of the system. Whenever the system is at state $x(t) \in S$ at time $t \geq 0$, players 1 and 2 independently choose actions $a_t \in A(x(t))$ and $b_t \in B(x(t))$ according to some strategies, respectively. As a consequence of this, the following happens:

1. player 1 receives an immediate reward at rate $r(x(t), a_t, b_t)$ and player 2 incurs a cost at rate $r(x(t), a_t, b_t)$; and
2. the system moves to a new state set governed by a possibly nonhomogeneous transition probability function, which is determined by the transition rates $q(\cdot \mid x(t), a_t, b_t)$.

Thus, the goal of player 1 is to maximize his/her rewards, whereas that of player 2 is to minimize his/her costs with respect to some performance criterion $J(\cdot, \cdot, \cdot)$, which in our present case is defined by (4.1), below.

To illustrate the game model, (2.1), we consider the following two examples.

Example 2.1. (*A controlled birth and death system.*) Consider a birth and death system in which the state variable denotes the population size at time $t \geq 0$. There are ‘natural’ birth and death rates λ and μ , respectively, which are nonnegative constants, emigration parameters h_1 controlled by a player 1, and immigration parameters h_2 controlled by a player 2. When the state of the system is $x \in S := \{0, 1, \dots\}$, player 1 takes an action a from a given set $A(x)$, which gives a reward at rate $r_1(x, a)$ and which may increase ($h_1(x, a) \geq 0$) or decrease ($h_1(x, a) \leq 0$) the emigration parameter. Conversely, player 2 takes an action b from a set $B(x)$, which incurs a cost with rate $c(x, b)$, and which may increase ($h_2(x, b) \geq 0$) or decrease ($h_2(x, b) \leq 0$) the immigration. In addition, there is a reward rate p_0x for each unit of time during which the system remains in state x , where $p_0 > 0$.

We now formulate this system as the game model (2.1) with S , $A(x)$ and $B(x)$ as above, $A := \bigcup_{x \in S} A(x)$, $B := \bigcup_{x \in S} B(x)$, and the corresponding transition rates $q(y \mid x, a, b)$ and reward rates $r(x, a, b)$ are given as follows. For $x = 0$ and each $a \in A(0)$ and $b \in B(0)$,

$$q(1 \mid 0, a, b) := -q(0 \mid 0, a, b) := h_2(0, b),$$

and, for each $x \geq 1$, $a \in A(x)$, and $b \in B(x)$,

$$q(y \mid x, a, b) := \begin{cases} \mu x + h_1(x, a) & \text{if } y = x - 1, \\ -(\mu + \lambda)x - h_1(x, a) - h_2(x, b) & \text{if } y = x, \\ \lambda x + h_2(x, b) & \text{if } y = x + 1, \\ 0 & \text{otherwise,} \end{cases} \tag{2.3}$$

$$r(x, a, b) := p_0 x + r_1(x, a) + c(x, b). \tag{2.4}$$

Example 2.2. (*A controlled Schlögl first model.*) The Schlögl first model [27] is a model of a chemical reaction with diffusion in a container, and it is a typical model of nonequilibrium systems. Here we formulate it as the game problem.

Let $E := \{1, \dots, i, \dots, N\}$ with $N < \infty$. We view each $i \in E$ as a small vessel in which there is a reaction. The states of the model are vectors $x = (x_i, i \in E)$, where $x_i \geq 0$ is the number of particles in vessel i . Thus, the state space is $S := \mathbb{Z}_+^N$ with $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$. In each vessel $i \in E$, the rate function of the reaction is of a birth–death type as

$$q_r(y \mid x, a, b) := \begin{cases} \beta_0 + \beta_1 x_i & \text{if } y = x + e_i, \\ \delta_1 x_i + \delta_2 x_i(x_i - 1) & \text{if } y = x - e_i, \\ 0 & \text{otherwise,} \end{cases}$$

with $a := (\beta_0, \beta_1)$ and $b := (\delta_1, \delta_2)$, where e_i is the element in S whose value corresponding to i is 1, and all other values are 0. Here we interpret the parameters (β_0, β_1) and (δ_1, δ_2) as the actions controlled by players 1 and 2, which take values in compact sets A and B in \mathbb{R}_+^2 , respectively, where $\mathbb{R}_+ := (0, \infty)$. Moreover, when using actions a and b in state $x \in S$, player 1 receives a reward $r_1(x, a)$ and player 2 obtains a reward $r_2(x, b)$, so that the payoff for this model turns out to be $r(x, a, b) := r_1(x, a) - r_2(x, b)$.

The other part of the model consists of diffusion between any two vessels, which are described by a transition probability matrix $(p(i, j) : i, j \in E)$. This means that, if there are x_i particles in vessel i , then the rate function of the diffusion from vessel i to j is $x_i p(i, j)$. Thus, the rate function of the diffusion becomes

$$q_d(y \mid x) := \begin{cases} x_i p(i, j) & \text{if } y = x - e_i + e_j, \\ 0 & \text{otherwise.} \end{cases}$$

Then the Schlögl first model can be expressed as the game model (2.1) with $S, A(x), B(x)$, and $r(x, a, b)$ as above, $A := \bigcup_{x \in S} A(x)$, $B := \bigcup_{x \in S} B(x)$, and the transition rates $q(y \mid x, a, b) := q_r(y \mid x, a, b) + q_d(y \mid x)$. Explicitly, for each $a := (\beta_0, \beta_1) \in A$ and $b := (\delta_1, \delta_2) \in B$,

$$q(y \mid x, a, b) := \begin{cases} \beta_0 + \beta_1 x_i & \text{if } y = x + e_i, \\ \delta_1 x_i + \delta_2 x_i(x_i - 1) & \text{if } y = x - e_i, \\ x_i p(i, j) & \text{if } y = x - e_i + e_j, \\ 0 & \text{otherwise,} \end{cases} \tag{2.5}$$

and $q(x \mid x, a, b) = -\sum_{y \neq x} q(y \mid x, a, b)$.

The two examples will be further studied in Section 6, below.

3. Strategies

A randomized Markov strategy for player 1 is a family $\pi^1 = (\pi_t^1) (t \geq 0)$ that satisfies the following conditions.

- (i) For each $t \geq 0$, π_t^1 is a stochastic kernel on A given S , that is, for each $D \in \mathcal{B}(A)$, $\pi_t^1(D | \cdot)$ is a Borel function on S , and for each $x \in S$, $\pi_t^1(\cdot | x)$ is a probability measure on A , and, furthermore, $\pi_t^1(\cdot | x)$ is concentrated on $A(x)$, i.e. $\pi_t^1(A(x) | x) = 1$; and
- (ii) for every $D \in \mathcal{B}(A)$ and $x \in S$, $\pi_t^1(D | x)$ is a Borel measurable function in $t \geq 0$.

For each $t \geq 0$ and $x \in S$, by (i), we shall regard $\pi_t^1(\cdot | x)$ as a probability measure on $A(x)$. We denote by Π_1^m the family of all randomized Markov strategies for player 1.

A strategy $\pi^1 = (\pi_t^1) \in \Pi_1^m$ is called *stationary* if it does not depend on t ; that is, if there is a stochastic kernel π^1 on A given S such that

$$\pi_t^1(\cdot | x) \equiv \pi^1(\cdot | x) \quad \text{for all } x \in S \text{ and } t \geq 0.$$

We denote this stationary strategy by $\pi^1(\cdot | x)$. The set of all stationary strategies for player 1 is denoted by Π_1^s .

The sets of all randomized Markov strategies Π_2^m and all stationary strategies Π_2^s for player 2 are defined similarly, with $B(x)$ in lieu of $A(x)$.

For each pair of strategies $(\pi^1, \pi^2) := (\pi_t^1, \pi_t^2) \in \Pi_1^m \times \Pi_2^m$, we define the associated transition rates $q(\cdot | x, \pi_t^1, \pi_t^2)$ as follows. For each $C \in \mathcal{B}(S)$, $x \in S$, and $t \geq 0$,

$$q(C | x, \pi_t^1, \pi_t^2) := \int_{B(x)} \int_{A(x)} q(C | x, a, b) \pi_t^1(da | x) \pi_t^2(db | x). \tag{3.1}$$

In particular, when both π^1 and π^2 are stationary, we write the left-hand side of (3.1) as $q(C | x, \pi^1, \pi^2)$.

It follows that, for each fixed pair of strategies $(\pi^1, \pi^2) := (\pi_t^1, \pi_t^2) \in \Pi_1^m \times \Pi_2^m$, the transition rates function $q(\cdot | x, \pi_t^1, \pi_t^2)$ is an infinitesimal generator [5]. (For a more precise statement of this fact, see Lemma 7.2(b) with $u = \mathbf{1}_C$, the indicator function of a set C , and Remark 7.1, below.) As is well known, any (possibly substochastic and nonhomogeneous) transition (probability) function $\tilde{p}(s, x, t, C; \pi^1, \pi^2)$ depending on (π^1, π^2) and satisfying

$$\lim_{\Delta t \rightarrow 0^+} \frac{\tilde{p}(t, x, t + \Delta t, C; \pi^1, \pi^2) - \mathbf{1}_C(x)}{\Delta t} = q(C | x, \pi_t^1, \pi_t^2)$$

for all $x \in S$, $C \in \mathcal{B}(S)$, and $t \geq 0$, is called a *Q-process* with transition rates $q(C | x, \pi_t^1, \pi_t^2)$.

To guarantee the existence of such a Q-process we restrict ourselves to admissible policies in the classes Π_1 and Π_2 defined as follows.

Definition 3.1. The classes Π_1 and Π_2 denote the subsets of randomized Markov strategies $\pi^k = (\pi_t^k)$ in Π_k^m for player $k (k = 1, 2)$ such that $\pi_t^1(D_1 | x)$ and $\pi_t^2(D_2 | x)$ are continuous in $t \geq 0$, for each fixed $x \in S$, $D_1 \in \mathcal{B}(A)$, and $D_2 \in \mathcal{B}(B)$.

Remark 3.1. (a) Observe that $\Pi_1 \times \Pi_2$ is nonempty because it contains $\Pi_1^s \times \Pi_2^s \neq \emptyset$. Moreover, we will provide an example to show that Π_1 and Π_2 do not contain the stationary strategies only; see Example 6.4, below.

(b) By (3.1), $q(C \mid x, \pi_t^1, \pi_t^2)$ is continuous in $t \geq 0$ for each $x \in S$, $C \in \mathcal{B}(S)$, and $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$.

By (T₁)–(T₃), (2.2), Definition 3.1, and (3.1), we have the following facts.

Lemma 3.1. *Let (π^1, π^2) be in $\Pi_1 \times \Pi_2$. Then*

- (a) *for each $x \in S$, $t \geq 0$, and $C \in \mathcal{B}(S)$,*
 - (i) $q(\cdot \mid x, \pi_t^1, \pi_t^2)$ *is a signed measure on $\mathcal{B}(S)$;*
 - (ii) $0 \leq q(C \mid x, \pi_t^1, \pi_t^2) < \infty$ *when $x \notin C$;*
 - (iii) $q(S \mid x, \pi_t^1, \pi_t^2) = 0$ *and* $0 \leq -q(\{x\} \mid x, \pi_t^1, \pi_t^2) < \infty$;
 - (iv) $q(C \mid x, \pi_t^1, \pi_t^2)$ *is continuous in $t \geq 0$ and measurable in $x \in S$,*
- (b) *there exists a Q-process $\tilde{p}(s, x, t, C; \pi^1, \pi^2)$ with transition rates $q(C \mid x, \pi_t^1, \pi_t^2)$.*

Proof. Parts (i)–(iii) in (a) follow from (3.1) and the definition of the model, (2.1). Part (iv) follows from (2.2) and Definition 3.1. Finally, part (b) follows from part (a) together with Theorem 1 of [5].

Lemma 3.1(b) guarantees the existence of a Q-process, such as the *minimum* Q-process denoted by $p^{\min}(s, x, t, C; \pi^1, \pi^2)$, which it is so-named because $p^{\min}(s, x, t, C; \pi^1, \pi^2) \leq \tilde{p}(s, x, t, C; \pi^1, \pi^2)$ for any Q-process $\tilde{p}(s, x, t, C; \pi^1, \pi^2)$. The minimum Q-process can be directly constructed from the given transition rates $q(\cdot \mid x, \pi_t^1, \pi_t^2)$; see (3.3), below. As is well known [1], [5], however, such a Q-process might not be regular, that is, we might have $p^{\min}(s, x, t, S; \pi^1, \pi^2) < 1$ for some $x \in S$ and $t \geq s \geq 0$. To ensure the regularity of a Q-process we propose the following ‘drift’ conditions.

Assumption 3.1. *There exists a measurable function $w_1 \geq 1$ on S , and constants $c_1 \neq 0$, $b_1 \geq 0$, and $M_q > 0$ such that*

- (a) $\int_S w_1(y)q(dy \mid x, a, b) \leq c_1 w_1(x) + b_1$ *for all $(x, a, b) \in K$; and*
- (b) $q(x) \leq M_q w_1(x)$ *for all $x \in S$, with $q(x)$ as defined in (2.2).*

Remark 3.2. (a) For the case of uniformly bounded transition rates (i.e. $\sup_{x \in S} q(x) < \infty$), Assumption 3.1(b) is not required.

(b) Assumption 3.1(a) is an extension of the ‘drift condition’ (2.4) of [20], for a *homogeneous* Q-process, and it is also known as a Lyapunov or Foster–Lyapunov condition. It is typically used to obtain growth conditions, as in Lemma 3.2(a), below, or some forms of ergodicity, as in [11], [16], and [29], for instance.

Assumption 3.1 is supposed to hold throughout the following, unless explicitly stated otherwise.

Lemma 3.2. *Suppose that Assumption 3.1 holds. Then, for each $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$, and all $t \geq s \geq 0$ and $x \in S$, we have*

- (a) $\int_S w_1(y)p^{\min}(s, x, t, dy; \pi^1, \pi^2) \leq e^{c_1(t-s)} w_1(x) + (b_1/c_1)(e^{c_1(t-s)} - 1)$, *and*
- (b) $p^{\min}(s, x, t, S; \pi^1, \pi^2) = 1$.

Proof. (a) Part (a) can be obtained as in the proof of Lemma 3.2 of [11].

(b) Let $\gamma(C \mid x, \pi_z^1, \pi_z^2) := q(C \setminus \{x\} \mid x, \pi_z^1, \pi_z^2)$. For each $x \in S, C \in \mathcal{B}(S), n \geq 0$, and $t \geq s \geq 0$, let

$$\begin{aligned}
 & p_{n+1}(s, x, t, C, \pi^1, \pi^2) \\
 & := \int_s^t \exp\left(\int_s^z q(\{x\} \mid x, \pi_v^1, \pi_v^2) dv\right) \int_S p_n(z, y, t, C, \pi^1, \pi^2) \gamma(dy \mid x, \pi_z^1, \pi_z^2) dz,
 \end{aligned}
 \tag{3.2}$$

with $p_0(s, x, t, C, \pi^1, \pi^2) := \mathbf{1}_C(x) \exp(\int_s^t q(\{x\} \mid x, \pi_v^1, \pi_v^2) dv)$. Then, by Theorem 2 of [5], we have

$$p^{\min}(s, x, t, C; \pi^1, \pi^2) = \sum_{n=0}^{\infty} p_n(s, x, t, C, \pi^1, \pi^2).
 \tag{3.3}$$

Moreover, let $L_n(s, x, t, \pi^1, \pi^2) := -\int_s^t \int_S q(\{y\} \mid y, \pi_v^1, \pi_v^2) p_n(s, x, v, dy, \pi^1, \pi^2) dv$ for all $x \in S$ and $t \geq s \geq 0$. Then, by Assumption 3.1(b) and (2.2) we have

$$L_n(s, x, t, \pi^1, \pi^2) \leq \int_s^t \int_S M_q w_1(y) p_n(s, x, v, dy, \pi^1, \pi^2) dv,$$

which together with (3.3) and part (a) gives $\lim_{n \rightarrow \infty} L_n(s, x, t, \pi^1, \pi^2) = 0$. Hence, part (b) follows from the corollary after Theorem 1 of [5].

Lemma 3.2(b) shows that a Q-process with transition rates $q(C \mid x, \pi_t^1, \pi_t^2)$ is *regular*. Thus, under Assumption 3.1 we shall write the regular transition function $p^{\min}(s, x, t, C; \pi^1, \pi^2)$ simply as $p(s, x, t, C; \pi^1, \pi^2)$.

To define our game’s payoff criterion, we need to introduce the state and action processes for players 1 and 2. Thus, we first construct a probability space for these processes.

For each $s \geq 0$, let $G^s := \{(t_0, t_1, \dots, t_n) \mid s \leq t_0 < t_1 < \dots < t_n < \infty, n \geq 0\}$ directed by inclusion, i.e. $\tau := (t_0, t_1, \dots, t_n) < \tau' \in G^s$ means that $\tau' = (t_0, t_1, \dots, t_n, \dots, t_m)$ for some positive numbers $t_m > t_{m-1} > \dots > t_n$ with $m \geq n + 1$.

For each $(\pi^1, \pi^2) := (\pi_t^1, \pi_t^2) \in \Pi_1 \times \Pi_2$, let $p(s, x, t, C; \pi^1, \pi^2)$ be the corresponding regular Q-process. Let $E := S \times A \times B$ with the usual product σ -algebra.

For each $s \geq 0$, let μ^s be an arbitrary probability distribution on S . Then, for each vector $\tau := (t_0, t_1, \dots, t_n)$ in G^s , we can define the product probability measure $P_\tau^{\mu^s, \pi^1, \pi^2}$ on $\Omega_\tau := E^{n+1}$ such that, for each $F := C_{t_0} \times D_{t_0}^1 \times D_{t_0}^2 \times C_{t_1} \times D_{t_1}^1 \times D_{t_1}^2 \times \dots \times C_{t_n} \times D_{t_n}^1 \times D_{t_n}^2$ with $C_{t_k} \in \mathcal{B}(S), D_{t_k}^1 \in \mathcal{B}(A)$, and $D_{t_k}^2 \in \mathcal{B}(B)$,

$$\begin{aligned}
 P_\tau^{\mu^s, \pi^1, \pi^2}(F) &= \int_{C_{t_0}} \mu^{t_0}(dx_{t_0}) \pi_{t_0}^1(D_{t_0}^1 \mid x_{t_0}) \pi_{t_0}^2(D_{t_0}^2 \mid x_{t_0}) \int_{C_{t_1}} p(t_0, x_{t_0}, t_1, dx_{t_1}; \pi^1, \pi^2) \\
 &\quad \times \pi_{t_1}^1(D_{t_1}^1 \mid x_{t_1}) \pi_{t_1}^2(D_{t_1}^2 \mid x_{t_1}) \cdots \int_{C_{t_n}} p(t_{n-1}, x_{t_{n-1}}, t_n, dx_{t_n}; \pi^1, \pi^2) \\
 &\quad \times \pi_{t_n}^1(D_{t_n}^1 \mid x_{t_n}) \pi_{t_n}^2(D_{t_n}^2 \mid x_{t_n}).
 \end{aligned}
 \tag{3.4}$$

Let $T := [0, \infty)$, and let (Ω, \mathcal{F}) be the (canonical) product measurable space with $\Omega := E^T$, and \mathcal{F} the product σ -algebra. Moreover, let g_τ be the coordinate projection, i.e.

$$g_\tau(e) := e_\tau \quad \text{for each } e = (e(t), t \geq 0) \in \Omega,
 \tag{3.5}$$

where $e_\tau := (e(t_0), e(t_1), \dots, e(t_n))$ and $\tau = (t_0, t_1, t_2, \dots, t_n)$. We then have the following result.

Proposition 3.1. *Suppose that Assumption 3.1 holds. Then, for each $(\pi^1, \pi^2) = (\pi_t^1, \pi_t^2) \in \Pi_1 \times \Pi_2$ and any given (initial) distribution μ^s on S at time $s \geq 0$, there exists a unique probability measure $\tilde{P}_{\pi^1, \pi^2}^{\mu^s}$ on (Ω, \mathcal{F}) such that, for each $t \geq s$,*

$$\tilde{P}_{\pi^1, \pi^2}^{\mu^s}(W_t) = \int_{C_s} \mu^s(dx_s) \int_C p(s, x_s, t, dx_t; \pi^1, \pi^2) \pi_t^1(D_1 | x_t) \pi_t^2(D_2 | x_t), \tag{3.6}$$

$$\tilde{P}_{\pi^1, \pi^2}^{\mu^s}(W_s) = \int_{C_s} \mu^s(dx_s) \pi_s^1(D_1 | x_s) \pi_s^2(D_2 | x_s), \tag{3.7}$$

where

$$W_t := \{e \in \Omega : (e(s), e(t)) \in C_s \times A \times B \times C \times D_1 \times D_2\},$$

$$W_s := \{e \in \Omega : e(s) \in C_s \times D_1 \times D_2\},$$

with $C_s, C \in \mathcal{B}(S)$, $D_1 \in \mathcal{B}(A)$, and $D_2 \in \mathcal{B}(B)$.

Proof. Because any probability measure on a Polish space is ‘inner regular’ [25], by (3.4) we conclude that the probability measure $P_\tau^{\mu^s, \pi^1, \pi^2}$ is also inner regular for each $\tau \in G^s$. Then the existence of the unique probability measure $\tilde{P}_{\pi^1, \pi^2}^{\mu^s}$ follows from Theorem 4 of [25, p. 20], for instance. Moreover, the results in (3.6) and (3.7) follow from (3.4) and (3.5), together with Proposition 1 of [25, p. 18].

Definition 3.2. (State and action processes.) For each $e = (e(t), t \geq 0) \in \Omega$ and $t \geq 0$, let $e(t) := (e^0(t), e^1(t), e^2(t)) \in E = S \times A \times B$. Then the coordinate process $\xi(t)(e) := e^0(t)$ defines the system’s state process, and $\eta_1(t)(e) := e^1(t)$ and $\eta_2(t)(e) := e^2(t)$ are the action processes for players 1 and 2, respectively.

Let $\tilde{E}_{\pi^1, \pi^2}^{\mu^s}$ denote the expectation operator associated with $\tilde{P}_{\pi^1, \pi^2}^{\mu^s}$. We write $\tilde{P}_{\pi^1, \pi^2}^{s,x}$ for $\tilde{P}_{\pi^1, \pi^2}^{\mu^s}$ and $\tilde{E}_{\pi^1, \pi^2}^{s,x}$ for $\tilde{E}_{\pi^1, \pi^2}^{\mu^s}$ when μ^s is the Dirac measure at $x \in S$. We then obtain the following result.

Proposition 3.2. *Suppose that Assumption 3.1 holds. Then, for each $x, y \in S$, $(\pi^1, \pi^2) = (\pi_t^1, \pi_t^2) \in \Pi_1 \times \Pi_2$, $C \in \mathcal{B}(S)$, $D_1 \in \mathcal{B}(A)$, $D_2 \in \mathcal{B}(B)$, and $t \geq s$,*

- (a) $\tilde{P}_{\pi^1, \pi^2}^{s,x}(\xi(t) \in C) = p(s, x, t, C; \pi^1, \pi^2)$,
- (b) $\tilde{P}_{\pi^1, \pi^2}^{s,x}(\eta_1(t) \in D_1, \eta_2(t) \in D_2 | \xi(t) = y) = \pi_t^1(D_1 | y) \pi_t^2(D_2 | y)$,
- (c) $\tilde{E}_{\pi^1, \pi^2}^{s,x} r(\xi(t), \eta_1(t), \eta_2(t))$ is Borel measurable in $t \geq s$.

Proof. Parts (a) and (b) follow from (3.6) and Definition 3.2. To prove part (c) we define reward rates $r(x, \pi_t^1, \pi_t^2)$ as follows:

$$r(x, \pi_t^1, \pi_t^2) := \int_{A(x)} \int_{B(x)} r(x, a, b) \pi_t^2(db | x) \pi_t^1(da | x) \quad \text{for each } x \in S \text{ and } t \geq 0. \tag{3.8}$$

By Definition 3.1, we see that $r(x, \pi_t^1, \pi_t^2)$ is Borel measurable in $(t, x) \in [0, \infty) \times S$. Thus, by (3.6) and (3.8) we have

$$\tilde{E}_{\pi^1, \pi^2}^{s,x} r(\xi(t), \eta_1(t), \eta_2(t)) = \int_S p(s, x, t, dy; \pi^1, \pi^2) r(y, \pi_t^1, \pi_t^2), \tag{3.9}$$

which is Borel measurable in (t, x) , because $p(s, x, t, C; \pi^1, \pi^2)$ is continuous in $t \geq 0$; see [5]. Hence, (c) follows.

We will write the left-hand side of (3.8) as $r(x, \pi^1, \pi^2)$ when both π^1 and π^2 are stationary.

4. The discounted payoff criterion

Fix a discount factor $\alpha > 0$. Then, for each pair of strategies $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$ and $x \in S$, the *discounted payoff* criterion $J(\cdot, \cdot, \cdot)$ is defined as

$$J(x, \pi^1, \pi^2) := \int_0^\infty (\tilde{E}_{\pi^1, \pi^2}^{0,x} e^{-\alpha t} r(\xi(t), \eta_1(t), \eta_2(t))) dt. \tag{4.1}$$

We also need the following concepts. The functions on S defined as

$$L(x) := \sup_{\pi^1 \in \Pi_1} \inf_{\pi^2 \in \Pi_2} J(x, \pi^1, \pi^2) \quad \text{and}$$

$$U(x) := \inf_{\pi^2 \in \Pi_2} \sup_{\pi^1 \in \Pi_1} J(x, \pi^1, \pi^2),$$

are called the *lower value* and the *upper value*, respectively, of the discounted payoff game. It is clear that

$$L(x) \leq U(x) \quad \text{for all } x \in S. \tag{4.2}$$

When the equality holds in (4.2) we obtain the following definitions from the theory of dynamic games [2], [6], [7], [12], [13], [15], [16], [19], [22], [26], [28].

Definition 4.1. If $L(x) = U(x)$ for all $x \in S$, then the *common* function is called the *value* of the game and denoted by V .

Definition 4.2. Suppose that the game has a value V . Then a strategy π_*^1 in Π_1 is said to be optimal for player 1 if

$$\inf_{\pi^2 \in \Pi_2} J(x, \pi_*^1, \pi^2) = V(x) \quad \text{for all } x \in S.$$

Similarly, $\pi_*^2 \in \Pi_2$ is optimal for player 2 if

$$\sup_{\pi^1 \in \Pi_1} J(x, \pi^1, \pi_*^2) = V(x) \quad \text{for all } x \in S.$$

If $\pi_*^k \in \Pi_k$ is optimal for player k ($k = 1, 2$), then (π_*^1, π_*^2) is called a pair of optimal strategies.

The aim of this paper is to give conditions for the existence of pairs of optimal stationary strategies, and to present a ‘martingale characterization’ of such pairs. To this end we first introduce some notation.

For each $s \geq 0$, $x \in X$, and each pair of strategies $(\pi^1, \pi^2) := (\pi_t^1, \pi_t^2) \in \Pi_1 \times \Pi_2$, we denote by $E_{s,x}^{\pi^1, \pi^2}$ the expectation operator associated with the probability measure $P_{s,x}^{\pi^1, \pi^2}$

which, as in Proposition 3.1, is completely determined by $p(s, x, t, C; \pi^1, \pi^2)$. Then, from [8, pp. 107–109], there exists a right-continuous, with finite left-hand limits, Markov process $x(t)$ with values in S and $p(s, x, t, C; \pi^1, \pi^2)$ as its transition function. It is well known that such a process $x(t)$ is Borel measurable as a function of $t \geq 0$.

Thus, by (3.9) we have, for each $t \geq 0$ and $x \in S$,

$$\tilde{E}_{\pi^1, \pi^2}^{0,x} r(\xi(t), \eta_1(t), \eta_2(t)) = E_{0,x}^{\pi^1, \pi^2} r(x(t), \pi_t^1, \pi_t^2), \tag{4.3}$$

and so, by (4.3) and (4.1),

$$J(x, \pi^1, \pi^2) = \int_0^\infty (E_{0,x}^{\pi^1, \pi^2} e^{-\alpha t} r(x(t), \pi_t^1, \pi_t^2)) dt. \tag{4.4}$$

Remark 4.1. By (4.4), without loss of generality we will restrict ourselves to using $x(t)$ and the corresponding probability $P_x^{\pi^1, \pi^2} := P_{0,x}^{\pi^1, \pi^2}$ and expectation $E_x^{\pi^1, \pi^2} := E_{0,x}^{\pi^1, \pi^2}$ throughout the following.

5. Main results

In this section we state our main results. The proofs are postponed to Section 7.

To guarantee the finiteness of the discounted criterion $J(x, \pi^1, \pi^2)$, by Lemma 3.2(a) and (4.4) it is natural to propose the following conditions, in which w_1 is the function in Assumption 3.1.

Assumption 5.1. (a) $|r(x, a, b)| \leq M_1 w_1(x)$ for all $(x, a, b) \in K$, for some constant $M_1 > 0$.

(b) $\alpha > \max\{0, c_1\}$, with c_1 as in Assumption 3.1.

Conversely, to ensure the existence of a pair of optimal strategies, in addition to Assumptions 3.1 and 5.1, we use the following hypotheses.

Assumption 5.2. (a) For each $x \in S$, $A(x)$ and $B(x)$ are compact.

(b) For each fixed $x \in S$, $r(x, a, b)$ is continuous in $(a, b) \in A(x) \times B(x)$.

(c) For each $x \in S$, the function $\int_S u(y)q(dy \mid x, a, b)$ is continuous in $(a, b) \in A(x) \times B(x)$, for $u = w_1$ and also for each bounded measurable function u on S .

(d) There exists a nonnegative measurable function w_2 on S , constants $c_2 > 0$, $b_2 \geq 0$, and $M_2 > 0$ such that (with $q(x)$ as defined in (2.2))

$$q(x)w_1(x) \leq M_2 w_2(x) \quad \text{and} \quad \int_S w_2(y)q(dy \mid x, a, b) \leq c_2 w_2(x) + b_2,$$

for all $(x, a, b) \in K$.

Remark 5.1. Assumptions 5.2(a)–(c) are similar to the standard continuity-compactness hypotheses for discrete-time Markov control processes; see, for instance, [14], [24], and the references therein. Assumption 5.2(d) allows us to use Dynkin’s formula, but it can be removed when $q(x)$ is bounded on S .

Given a measurable function w on S with $w(x) \geq 1$ for all $x \in S$, a function u on S is called w -bounded if the w -norm $\|u\|_w$ defined as

$$\|u\|_w := \sup_{x \in S} \frac{|u(x)|}{w(x)}$$

is finite. Such a function w will be referred to as a *weight function*. We denote by $B_w(S)$ the Banach space of all w -bounded measurable functions on S .

Moreover, for any $C \in \mathcal{B}(S)$, and probability measures $\phi \in P(A(x))$ and $\psi \in P(B(x))$, let

$$q(C \mid x, \phi, \psi) := \int_{B(x)} \int_{A(x)} q(C \mid x, a, b) \phi(da) \psi(db), \tag{5.1}$$

$$r(x, \phi, \psi) := \int_{B(x)} \int_{A(x)} r(x, a, b) \phi(da) \psi(db). \tag{5.2}$$

For each $\pi^1 = \pi^1(\cdot \mid x) \in \Pi_1^s$ and $\pi^2 = \pi^2(\cdot \mid x) \in \Pi_2^s$, we write

$$\begin{aligned} q(C \mid x, \phi, \pi^2) &:= q(C \mid x, \phi, \pi^2(\cdot \mid x)), & q(C \mid x, \pi^1, \psi) &:= q(C \mid x, \pi^1(\cdot \mid x), \psi), \\ r(x, \phi, \pi^2) &:= r(x, \phi, \pi^2(\cdot \mid x)), & r(x, \pi^1, \psi) &:= r(x, \pi^1(\cdot \mid x), \psi). \end{aligned}$$

Let $q(x)$ be as defined in (2.2), and take an arbitrary real-valued measurable function $m(x)$ on S such that

$$m(x) > q(x) \geq 0 \quad \text{for all } x \in S.$$

Moreover, let w_1 be as in Assumption 3.1, and define an operator T on $B_{w_1}(S)$ as follows: for $u \in B_{w_1}(S)$ and $x \in S$,

$$Tu(x) := \max_{\phi \in P(A(x))} \min_{\psi \in P(B(x))} \left\{ \frac{r(x, \phi, \psi)}{\alpha + m(x)} + \frac{m(x)}{\alpha + m(x)} \int_S u(y) P(dy \mid x, \phi, \psi) \right\}, \tag{5.3}$$

where the *probability measure* $P(\cdot \mid x, \phi, \psi)$ is defined as

$$P(C \mid x, \phi, \psi) := \frac{q(C \mid x, \phi, \psi)}{m(x)} + \mathbf{1}_C(x) \quad \text{for all } C \in \mathcal{B}(S). \tag{5.4}$$

Moreover, for each $n \geq 1$ and $x \in S$, let

$$u_n(x) := Tu_{n-1}(x) \quad \text{with } u_0(x) := -\left(\frac{b_1 M_1}{\alpha(\alpha - c_1)} + \frac{M_1 w_1(x)}{\alpha - c_1} \right), \tag{5.5}$$

where the constants c_1, b_1 , and M_1 are as in Assumptions 3.1 and 5.1. By Assumption 5.1(b), $u_0(\cdot) \leq 0$.

Theorem 5.1. *Suppose that Assumptions 3.1, 5.1, and 5.2 hold. Then we have the following results.*

- (a) *We have $|J(x, \pi^1, \pi^2)| \leq -u_0(x)$ for all $x \in S$ and $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$.*
- (b) *The limit $\lim_{n \rightarrow \infty} u_n(x) := u^*(x)$ exists and belongs to $B_{w_1}(S)$ and, moreover, the function u^* satisfies Shapley’s equation, i.e.*

$$\alpha u^*(x) = \max_{\phi \in P(A(x))} \min_{\psi \in P(B(x))} \left\{ r(x, \phi, \psi) + \int_S u^*(y) q(dy \mid x, \phi, \psi) \right\} \quad \text{for all } x \in S. \tag{5.6}$$

(c) There exists a pair of stationary strategies $(\pi_*^1, \pi_*^2) \in \Pi_1^s \times \Pi_2^s$ such that, for all $x \in S$,

$$\alpha u^*(x) = r(x, \pi_*^1, \pi_*^2) + \int_S u^*(y)q(dy \mid x, \pi_*^1, \pi_*^2) \tag{5.7}$$

$$= \max_{\phi \in P(A(x))} \left\{ r(x, \phi, \pi_*^2) + \int_S u^*(y)q(dy \mid x, \phi, \pi_*^2) \right\} \tag{5.8}$$

$$= \min_{\psi \in P(B(x))} \left\{ r(x, \pi_*^1, \psi) + \int_S u^*(y)q(dy \mid x, \pi_*^1, \psi) \right\}. \tag{5.9}$$

(d) We have $u^*(x) = L(x) = U(x)$ for all $x \in S$, which means that the value V of the game exists and equals the function $u^*(x)$, and, furthermore, $u^*(x) = J(x, \pi_*^1, \pi_*^2)$ for all $x \in S$.

(e) In part (c), (π_*^1, π_*^2) is a pair of optimal stationary strategies.

Proof. See Section 7.

Theorem 5.1 is a key result: part (d) gives the existence of the value of the game, whereas (b) provides a recursive way of computing (or at least approximating) the value of the game. Moreover, Theorem 5.1(e) shows the existence of a pair of optimal stationary strategies. Next, we give an interesting characterization of a pair of optimal stationary strategies.

Under the hypotheses of Theorem 5.1, for each $x \in S$, $\phi \in P(A(x))$, and $\psi \in P(B(x))$ let

$$\Delta(x, \phi, \psi) := r(x, \phi, \psi) + \int_S V(y)q(dy \mid x, \phi, \psi) - \alpha V(x),$$

where $V(x)$ is the value of the game; see Theorem 5.1(d). Moreover, for each $\pi^1 := \pi^1(\cdot \mid x) \in \Pi_1^s$, $\pi^2 := \pi^2(\cdot \mid x) \in \Pi_2^s$, and $t \geq 0$, let

$$\begin{aligned} \Delta(x, \pi^1, \pi^2) &:= \Delta(x, \pi^1(\cdot \mid x), \pi^2(\cdot \mid x)), \\ \mathcal{F}_t &:= \sigma\{x(s), 0 \leq s \leq t\} \quad (\text{the } \sigma\text{-algebra generated by } x(\cdot)), \end{aligned} \tag{5.10}$$

and define the (continuous-time) stochastic process by

$$M(t, \pi^1, \pi^2) := \int_0^t e^{-\alpha\tau} r(x(\tau), \pi^1, \pi^2) d\tau + e^{-\alpha t} V(x(t)) \quad \text{for all } t \geq 0. \tag{5.11}$$

Theorem 5.2. *Suppose that Assumptions 3.1, 5.1, and 5.2 hold. Then, for a given pair of strategies (π_*^1, π_*^2) in $\Pi_1^s \times \Pi_2^s$, the following statements are equivalent.*

(a) The pair $(\pi_*^1, \pi_*^2) \in \Pi_1^s \times \Pi_2^s$ is optimal.

(b) For each $x \in S$,

$$\Delta(x, \pi_*^1, \pi_*^2) = \min_{\psi \in P(B(x))} \Delta(x, \pi_*^1, \psi) = \max_{\phi \in P(A(x))} \Delta(x, \phi, \pi_*^2) = 0.$$

(c) For each $x \in S$, $\pi^1 \in \Pi_1^s$ and $\pi^2 \in \Pi_2^s$,

(i) $\{M(t, \pi_*^1, \pi_*^2), \mathcal{F}_t\}$ is a $P_x^{\pi_*^1, \pi_*^2}$ martingale,

(ii) $\{M(t, \pi_*^1, \pi^2), \mathcal{F}_t\}$ is a $P_x^{\pi_*^1, \pi^2}$ submartingale, and

(iii) $\{M(t, \pi^1, \pi_*^2), \mathcal{F}_t\}$ is a $P_x^{\pi^1, \pi_*^2}$ supermartingale.

Proof. See Section 7.

Theorem 5.2 gives, in particular, a ‘martingale characterization’ of a pair of optimal stationary strategies. Theorems 5.1 and 5.2 are extensions of the results in [9] and [10] for control problems (or one-player games).

6. Examples

In this section we first apply our results to the controlled queueing system in Example 2.1 and to the controlled Schögl model in Example 2.2, and then introduce Examples 6.3 and 6.4 to illustrate our assumptions.

Example 6.1. (*Example 2.1 continued.*) To find conditions ensuring the existence of a pair of optimal stationary strategies for Example 2.1, we consider the following assumptions.

(E₁) We have $\alpha + \mu - \lambda > 0$; $\mu x + h_1(x, a) \geq 0$ and $\lambda x + h_2(x, b) \geq 0$ for all $a \in A(x)$, $b \in B(x)$, and $x \geq 1$; in addition $h_1(0, a) = 0$ and $h_2(0, b) \geq 0$ for all $a \in A(0)$ and $b \in B(0)$.

(E₂) For each $x \in S$, $A(x)$ and $B(x)$ are compact subsets of metric spaces A and B , respectively.

(E₃) For each $x \in S$, the functions $r_1(x, a)$ and $h_1(x, a)$ are continuous in $a \in A(x)$, while the functions $c(x, b)$ and $h_2(x, b)$ are continuous in $b \in B(x)$. Moreover, $h_1(x, a)$ and $h_2(x, b)$ are assumed to be bounded in the supremum norm, that is

$$\|h_1\| := \sup_{x \in S, a \in A(x)} |h_1(x, a)| < \infty, \quad \|h_2\| := \sup_{x \in S, b \in B(x)} |h_2(x, b)| < \infty.$$

Also, $|r_1(x, a)| \leq L_1(x + 1)$ for all $a \in A(x)$, and $|c(x, b)| \leq L_2(x + 1)$ for all $b \in B(x)$, for some positive constants L_k ($k = 1, 2$).

Under these conditions we obtain the following result.

Proposition 6.1. *Under assumptions E₁, E₂, and E₃, the controlled birth and death system in Example 6.1 satisfies Assumptions 3.1, 5.1, and 5.2. Therefore, by Theorem 5.1, there exists a pair of optimal stationary strategies.*

Proof. We shall first verify Assumption 3.1. Let $w_1(x) := x + 1$ for all $x \in S$, and $M_q := \mu + \lambda + \|h_1\| + \|h_2\|$. Then Assumption 3.1(b) follows from (2.3). Also, for each $x \in S$, $a \in A(x)$, and $b \in B(x)$, from E₁ and (2.3) we have, for $x \geq 1$,

$$\begin{aligned} \sum_{y \in S} q(y | x, a, b)w_1(y) &= (\lambda - \mu)(x + 1) + \mu - \lambda - h_1(x, a) + h_2(x, b) \\ &\leq (\lambda - \mu)w_1(x) + M_q. \end{aligned} \tag{6.1}$$

For $x = 0$, we obtain

$$\sum_{y \in S} q(y | 0, a, b)w_1(y) = h_2(0, b) \leq (\lambda - \mu)w_1(0) + M_q. \tag{6.2}$$

By inequalities (6.1) and (6.2) we see that Assumption 3.1(a) holds with $c_1 := \lambda - \mu$ and $b_1 := M_q$, and so Assumption 3.1 follows.

Furthermore, by (2.4) and E₃, we have

$$|r(x, a, b)| \leq p_0x + L_1(x + 1) + L_2(x + 1) \leq (p_0 + L_1 + L_2)w_1(x)$$

for all $x \in S$, which together with E₁ gives Assumption 5.1.

Finally, to verify Assumption 5.2 let

$$w_2(x) := (x + 1)^2 \quad \text{for all } x \in S.$$

Then

$$q(x)w_1(x) \leq M_2w_2(x) \quad \text{for all } x \in S,$$

with $M_2 := \mu + \lambda + \|h_1\| + \|h_2\|$.

Then as in (6.1) and (6.2), we can see that Assumption 5.2(d) is true. Conversely, by (2.3) together with E_2 and E_3 , Assumptions 5.2(a), 5.2(b), and 5.2(c) are all satisfied, and so Assumption 5.2 follows.

Example 6.2. (*Example 2.2 continued.*) To find conditions that ensure the existence of a pair of optimal stationary strategies for Example 2.2, we assume that $r_1(x, a)$ and $r_2(x, b)$ are continuous in a and b , respectively, and satisfy $|r_1(x, a)| \leq L_1 \sum_{i=1}^N x_i^2$ and $|r_2(x, b)| \leq L_2 \sum_{i=1}^N x_i^2$ for some positive constants L_k ($k = 1, 2$). Thus,

$$|r(x, a, b)| \leq (L_1 + L_2) \sum_{i=1}^N x_i^2 \quad \text{for all } a \in A, b \in B, \text{ and } x \in S. \tag{6.3}$$

Then the Schögl first model satisfies Assumptions 3.1, 5.1, and 5.2. Therefore, by Theorem 5.1, there exists a pair of optimal stationary strategies.

Indeed to verify Assumption 3.1, we let $w_1(x) := \sum_{i=1}^N x_i^2$. Since A and B are compact sets in \mathbb{R}_+^2 , there exist positive constants ρ_1 and ρ_2 such that $\rho_1 \leq \max\{\beta_0, \beta_1, \delta_1, \delta_2\} \leq \rho_2$ for all $(\beta_0, \beta_1) = a \in A$ and $(\delta_1, \delta_2) = b \in B$. Thus, for any $x \in S$ such that $x_i \geq 1$ with $i \in E$, $a \in A$, and $b \in B$,

$$\begin{aligned} \sum_{y \in S} q(y | x, a, b)w_1(y) &= \sum_{i=1}^N (1 + 2x_i)(\beta_0 + \beta_1 x_i) + \sum_{i=1}^N (1 - 2x_i)(\delta_1 x_i + \delta_2 x_i(x_i - 1)) \\ &\quad + \sum_{i \neq j} x_i p(i, j)((1 - 2x_i) + (1 + 2x_j)) \\ &=: \sum_{i=1}^N (-2\delta_2 x_i^3 + c_i^1 x_i^2 + c_i^2 x_i + c_i^3) \\ &\leq \sum_{i=1}^N (-2\rho_1 x_i^3 + \lambda_i^1 x_i^2 + \lambda_i^2 x_i + \lambda_i^3), \end{aligned}$$

where c_i^k, λ_i^k ($k = 1, 2, 3, i \in E$) are fixed constants, and λ_i^k are independent of a, b , and x . Since there exists a positive integer L^* such that $-2\delta_2 x_i^3 + \lambda_1 x_i^2 + \lambda_2 x_i + \lambda_3 < 0$ when $x_i \geq L^*$, and the set $\{x \in S: x_i \leq L^* \text{ for some } i \in E\}$ is finite, straightforward calculations yield a constant $b_1 > 0$ such that

$$\sum_{y \in S} q(y | x, a, b)w_1(y) \leq b_1 \leq \frac{1}{2}\alpha w_1(x) + b_1 \quad \text{for all } (x, a, b) \in K. \tag{6.4}$$

Conversely, by (2.5) we have

$$q(x) \leq 5N\rho_2 w_1(x),$$

which together with (6.4) gives Assumption 3.1.

By (6.3) and (6.4), Assumption 5.1 is true.

Finally, to verify Assumption 5.2, we let $w_2(x) := \sum_{i=1}^N x_i^4$ for all $x = (x_i, i \in E)$. Noting that $xy \leq x^2 + y^2$, then as in the proof of (6.4) we can verify Assumption 5.2(d), and by (2.5) we obtain Assumption 5.2(c). Thus, Assumption 5.2 follows because Assumption 5.2(a) and Assumption 5.2(b) are part of the model's description.

Remark 6.1. It should be noted that in Examples 6.1 and 6.2 both the reward and the transition rates are unbounded.

Example 6.3. (*A generalized Potlach process with control parameters.*) The generalized Potlach process (see [3], for instance) is a Q-process generated by the infinitesimal operator L defined by (6.5), below. Here we are interested in the game problem.

Take $S := [0, \infty)^d$ for some integer $d \geq 1$. Then the operator L is defined by

$$Lu(x, a, b) := \sum_{i=1}^d \int_0^\infty \left[u \left(x - e_i x_i + y \sum_{j=1}^d p_{ij} x_i e_j \right) - u(x) \right] b e^{-by} dy \quad \text{for all } x \in S, \quad (6.5)$$

where $a := (p_{ij})$ is a Markov transition matrix on $\{1, 2, \dots, d\}$, e_i is the i th unit vector in \mathbb{R}^d , and b is a parameter. When the process is at state $x = (x_1, \dots, x_d) \in S$, the cost incurred at each component x_i is denoted by $q_i > 0$ for all $i = 1, \dots, d$. Here we interpret the parameters a and b as controlled actions by player 1 and player 2, with values in the sets $A(x)$ and $B(x)$, respectively. Suppose that $A(x)$ is a finite set of Markov transition matrices (p_{ij}) and $B(x) := [1, b^*]$ for each $x \in S$, where $b^* > 1$. For each $C \in \mathcal{B}(S)$, $x \in S$, $a = (p_{ij}) \in A(x)$, and $b \in B(x)$, let

$$\tilde{q}(C | x, a, b) := \sum_{i=1}^d \int_0^\infty \mathbf{1}_{C \setminus \{x\}} \left(x - e_i x_i + y \sum_{j=1}^d p_{ij} x_i e_j \right) b e^{-by} dy. \quad (6.6)$$

Then, by (6.5) together with Lemma 7.2(b), below, we see that the transition rate $q(C | x, a, b)$ is given by

$$q(C | x, a, b) = \tilde{q}(C | x, a, b) - \mathbf{1}_C(x) \tilde{q}(S | x, a, b), \quad (6.7)$$

whereas the reward rate is given by

$$r(x, a, b) := \sum_{i=1}^d \sum_{j=1}^d q_i p_{ij} x_j - b(x_1 + \dots + x_d). \quad (6.8)$$

Let $w_1(x) := 1 + x_1 + x_2 + \dots + x_d$ for all $x = (x_1, \dots, x_d) \in S$. Then by (6.7) we have

$$q(x) = d, \quad (6.9)$$

and so (2.2) holds. Moreover, by (6.5)–(6.7) we have

$$\int_S w_1(y) q(dy | x, a, b) = -\frac{b-1}{b} (x_1 + \dots + x_d) \leq \frac{\alpha}{2} w_1(x) + b^*,$$

which together with (6.9), verifies Assumption 3.1 with $c_1 := \alpha/2$ and $b_1 = b^*$.

By (6.8) we have $|r(x, a, b)| \leq (q_1 + \dots + q_d + b^*) w_1(x)$ for all $x \in S$, $a \in A(x)$, and $b \in B(x)$, which together with the finiteness of $A(x)$ and (6.9) verifies Assumption 5.1.

Obviously, Assumption 5.2 follows from the model's description and the last part of Remark 5.1.

Remark 6.2. It should be noted that in Example 6.3 the state space is *not* denumerable, and the reward rates have *neither upper nor lower bounds*; see (6.8). Nevertheless, by the discussion above, in Example 6.3 our Assumptions 3.1, 5.1, and 5.2 are satisfied.

Example 6.4. In Example 6.3, we further suppose that $A(x) = \{a_1, a_2\}$ for each $x \in S$. Then, we define nonstationary Markov strategies $\tilde{\pi}^1 = (\tilde{\pi}_t^1, t \geq 0)$ and $\tilde{\pi}^2 = (\tilde{\pi}_t^2, t \geq 0)$ as

$$\tilde{\pi}_t^1(a \mid x) = \begin{cases} \exp(-\text{Tr}(a_1)(t + 1)w_1(x)) & \text{if } a = a_1, \\ 1 - \exp(-\text{Tr}(a_1)(t + 1)w_1(x)) & \text{if } a = a_2, \end{cases} \tag{6.10}$$

with $\text{Tr}(a_1)$ denoting the *trace* of a (transition) matrix a_1 , and

$$\tilde{\pi}_t^2(C \mid x) = \begin{cases} 1 - \exp(-\mathbf{1}_C(1)(t + 1)w_1(x)) & \text{for } C \in \mathcal{B}(B(x)), \\ \exp(-\mathbf{1}_C(b^*)(t + 1)w_1(x)) & \text{for } C \in \mathcal{B}(B(x)), \end{cases} \tag{6.11}$$

respectively.

Moreover, let $\Pi_1 := \Pi_1^s \cup \{\tilde{\pi}^1\}$, $\Pi_2 := \Pi_2^s \cup \{\tilde{\pi}^2\}$. By (6.6), (6.7), (6.10), (6.11), and (3.1) we see that Π_1 and Π_2 satisfy the requirements in Definition 3.1, and $\Pi_i \supset \Pi_i^s$, with $\Pi_i \neq \Pi_i^s$, for $i = 1, 2$.

7. Proof of Theorems 5.1 and 5.2

In this section we present the proof of Theorems 5.1 and 5.2. To do so, we first introduce some results which are already known but we state them here for completeness and ease of reference.

Lemma 7.1. *Suppose that Assumptions 5.2(a), 5.2(b), and 5.2(c) hold. Then we have the following.*

(a) *For each fixed $x \in S$ and $u \in B_{w_1}(S)$, the functions*

$$r(x, \phi, \psi) \quad \text{and} \quad \int_S u(y) P(dy \mid x, \phi, \psi)$$

(recall (5.2) and (5.4)) are continuous in $(\phi, \psi) \in P(A(x)) \times P(B(x))$.

(b) *If a real-valued function $g(\phi, \psi)$ is continuous in $(\phi, \psi) \in P(A(x)) \times P(B(x))$, then the function $g^*(\phi)$ on $P(A(x))$ defined as*

$$g^*(\phi) := \inf_{\psi \in P(B(x))} g(\phi, \psi)$$

is also continuous in $\phi \in P(A(x))$.

Proof. (a) By Assumptions 5.2(b) and 5.2(c) we see that $r(x, \phi, \psi)$ and $P(C \mid x, \phi, \psi)$, for each $C \in \mathcal{B}(S)$, are continuous. Then, the desired conclusion follows from Lemma 8.3.7 of [14] and Assumption 5.2(c).

(b) By Assumption 5.2(a), $P(B(x))$ is compact. Thus, $g^*(\phi)$ is well defined for each $\phi \in P(A(x))$. Similarly, by Assumption 5.2(a) we also have $P(A(x))$ is compact. Hence, $g(\phi, \psi)$ is uniformly continuous on the compact metric space $P(A(x)) \times P(B(x))$, and so part (b) follows from well-known facts.

Lemma 7.2. *Suppose that Assumptions 3.1, 5.2(b), 5.2(c), and 5.2(d) hold. Then, for each $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$, $x \in S$, and $s \geq 0$, the following statements hold.*

(a) *For each $u \in B_{w_1+w_2}(S)$ and $t \geq v \geq s$,*

(i)

$$\|E_{s,x}^{\pi^1, \pi^2} |u(x(t))|\|_{w_1+w_2} \leq \frac{b_1 + b_2 + |c_1| + c_2}{|c_1| + c_2} \|u\|_{w_1+w_2} e^{(|c_1|+c_2)(t-s)};$$

(ii)

$$\lim_{t \searrow s} \exp\left(\int_s^v q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta\right) E_{v,y}^{\pi^1, \pi^2} u(x(t)) = \lim_{t \searrow s} E_{s,y}^{\pi^1, \pi^2} u(x(t)) = u(y)$$

for all $y \in S$.

(b) *For each $u \in B_{w_1}(S)$, let $L^{\pi^1, \pi^2} u(s, x) := \lim_{t \downarrow 0} t^{-1} [E_{s,x}^{\pi^1, \pi^2} u(x(s+t)) - u(x)]$. Then*

(i) $L^{\pi^1, \pi^2} u(s, x) = \int_S u(y)q(dy \mid x, \pi_s^1, \pi_s^2),$

(ii)

$$\begin{aligned} & \|E_{s,x}^{\pi^1, \pi^2} |L^{\pi^1, \pi^2} u(t, x(t))|\|_{w_1+w_2} \\ & \leq \frac{\|u\|_{w_1} (|c_1| + c_2 + b_1 + b_2 + 2M_2)^2}{|c_1| + c_2} e^{(|c_1|+c_2)(t-s)} \end{aligned}$$

for all $t \geq s$.

Proof. (a) Let $w := w_1 + w_2$, $c := |c_1| + c_2$, and $b := b_1 + b_2$. For each $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$, $y \in S$ and $t \geq v \geq s \geq 0$, by Assumptions 3.1 and 5.2(d) as well as Lemma 3.2 we have

$$\begin{aligned} \int_S w(z)p(v, y, t, dz; \pi^1, \pi^2) & \leq e^{c(t-s)}w(y) + \frac{b}{c}(e^{c(t-s)} - 1) \\ & \leq \frac{c + b}{c}e^{c(t-s)}w(y) \quad \text{for all } v \in [s, t], \end{aligned} \tag{7.1}$$

which together with $v := s$ and noting that $|u(x)| \leq w(x)\|u\|_w$ gives (a)(i).

Noting that $q(\{x\} \mid x, \pi_t^1, \pi_t^2) \leq 0$, by (3.3) we have

$$\begin{aligned} w(y) & \geq p(v, y, t, y; \pi^1, \pi^2)w(y) \\ & \geq \exp\left(\int_v^t q(\{y\} \mid y, \pi_\delta^1, \pi_\delta^2) d\delta\right)w(y) \\ & \geq \exp\left(\int_s^t q(\{y\} \mid y, \pi_\delta^1, \pi_\delta^2) d\delta\right)w(y) \quad \text{for all } v \in [s, t]. \end{aligned} \tag{7.2}$$

This inequality and (7.1) give

$$\begin{aligned} 0 & \leq \exp\left(\int_s^v q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta\right) \int_{S \setminus \{y\}} w(z)p(v, y, t, dz; \pi^1, \pi^2) \\ & \leq \exp\left(\int_s^v q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta\right) \\ & \quad \times \left(e^{c(t-s)}w(y) + \frac{b}{c}(e^{c(t-s)} - 1) - w(y)p(v, y, t, y; \pi^1, \pi^2)\right). \end{aligned} \tag{7.3}$$

Letting $t \searrow s$, from (7.2) and (7.3) we have

$$\begin{aligned} & \lim_{t \searrow s} \left(\exp \left(\int_s^v q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta \right) p(v, y, t, y; \pi^1, \pi^2) - 1 \right) w(y) \\ &= \lim_{t \searrow s} \exp \left(- \int_s^v q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta \right) \int_{S \setminus \{y\}} w(z) p(v, y, t, dz; \pi^1, \pi^2) \\ &= 0. \end{aligned}$$

Thus, for each $u \in B_w(S)$,

$$\begin{aligned} 0 &\leq \limsup_{t \searrow s} \left| \exp \left(\int_s^v q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta \right) \int_S u(z) p(v, y, t, dz; \pi^1, \pi^2) - u(y) \right| \\ &\leq \|u\|_w \limsup_{t \searrow s} \left[1 - \exp \left(\int_s^v q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta \right) p(v, y, t, y; \pi^1, \pi^2) \right] w(y) \\ &\quad + \|u\|_w \limsup_{t \searrow s} \exp \left(\int_s^v q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta \right) \int_{S \setminus \{y\}} w(z) p(v, y, t, dz; \pi^1, \pi^2) \\ &= 0. \end{aligned}$$

This implies (a)(ii).

(b) Let γ be as in the proof of Lemma 3.2(b). Under Assumption 3.1, by Fubini’s theorem, (3.2), and (3.3) we have

$$\begin{aligned} & E_{s,x}^{\pi^1, \pi^2} u(x(t)) - u(x) \\ &= \int_s^t \int_S \exp \left(\int_s^v q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta \right) E_{v,y}^{\pi^1, \pi^2} u(x(t)) \gamma(dy \mid x, \pi_v^1, \pi_v^2) dv \\ &\quad + \exp \left(\int_s^t q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta \right) u(x) - u(x). \end{aligned} \tag{7.4}$$

From part (a) and Lemma 7.1, above, together with [14, Lemma 8.3.7], we have

$$\begin{aligned} & \lim_{t \downarrow s} \int_S \exp \left(\int_s^v q(\{x\} \mid x, \pi_\delta^1, \pi_\delta^2) d\delta \right) E_{v,y}^{\pi^1, \pi^2} u(x(t)) \gamma(dy \mid x, \pi_v^1, \pi_v^2) \\ &= \int_S u(y) \gamma(dy \mid x, \pi_s^1, \pi_s^2). \end{aligned} \tag{7.5}$$

Finally, divide both sides of (7.4) by $t - s$, and then use (7.5) to obtain (b)(i).

Moreover, for each $u \in B_{w_1}(S)$ and $t \geq 0$, by Assumption 5.2(d) we have

$$\begin{aligned} \left| \int_S u(y) q(dy \mid x, \pi_t^1, \pi_t^2) \right| &\leq \|u\|_{w_1} \left(\int_S w_1(y) \gamma(dy \mid x, \pi_t^1, \pi_t^2) + q(x) w_1(x) \right) \\ &\leq \|u\|_{w_1} (c_1 w_1(x) + b_1 + 2M_2 w_2(x)) \\ &\leq \|u\|_{w_1} (|c_1| + b_1 + 2M_2) (w_1(x) + w_2(x)), \end{aligned}$$

which together with part (a)(i) gives (b)(ii).

Remark 7.1. For each fixed $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$, Lemma 7.2(b) shows that L^{π^1, π^2} is the extended infinitesimal operator corresponding to $p(s, x, t, C; \pi^1, \pi^2)$, and also that the domain of L^{π^1, π^2} contains $B_{w_1}(S)$.

Lemma 7.3. Suppose that Assumptions 3.1, 5.1, 5.2(c), and 5.2(d) hold. Choose arbitrary $u \in B_{w_1}(S)$ and $(\pi^1, \pi^2) = (\pi_t^1, \pi_t^2) \in \Pi_1 \times \Pi_2$.

(a) If, for each $x \in S$ and $t \geq 0$,

$$\alpha u(x) \geq r(x, \pi_t^1, \pi_t^2) + \int_S u(y)q(dy \mid x, \pi_t^1, \pi_t^2),$$

then $u(x) \geq J(x, \pi^1, \pi^2)$ for all $x \in S$.

(b) If, for each $x \in S$ and $t \geq 0$,

$$\alpha u(x) \leq r(x, \pi_t^1, \pi_t^2) + \int_S u(y)q(dy \mid x, \pi_t^1, \pi_t^2),$$

then $u(x) \leq J(x, \pi^1, \pi^2)$ for all $x \in S$.

(c) If $(\pi^1, \pi^2) \in \Pi_1^s \times \Pi_2^s$, then $J(x, \pi^1, \pi^2)$ is the unique solution to the equation

$$\alpha u(x) = r(x, \pi^1, \pi^2) + \int_S u(y)q(dy \mid x, \pi^1, \pi^2) \quad \text{for all } x \in S.$$

Proof. (a) By Lemma 7.2 and Dynkin’s formula we have

$$\begin{aligned} e^{-\alpha t} \mathbb{E}_x^{\pi^1, \pi^2} u(x(t)) - u(x) &= \mathbb{E}_x^{\pi^1, \pi^2} \int_0^t e^{-\alpha v} (L^{\pi^1, \pi^2} u(v, x(v)) - \alpha u(x(v))) \, dv \\ &\leq - \int_0^t \mathbb{E}_x^{\pi^1, \pi^2} e^{-\alpha v} r(x(v), \pi_v^1, \pi_v^2) \, dv. \end{aligned} \tag{7.6}$$

Conversely, by Lemma 3.2 and Assumption 5.1 we have

$$\begin{aligned} 0 &\leq \lim_{t \rightarrow \infty} e^{-\alpha t} \mathbb{E}_x^{\pi^1, \pi^2} |u(x(t))| \\ &\leq \|u\|_{w_1} \lim_{t \rightarrow \infty} [\mathbb{E}_x^{\pi^1, \pi^2} e^{-\alpha t} w_1(x(t))] \\ &\leq \|u\|_{w_1} \lim_{t \rightarrow \infty} [e^{-(\alpha - c_1)t} w_1(x) + \frac{b_1}{|c_1|} e^{-(\alpha - c_1)t} + \frac{b_1}{|c_1|} e^{-\alpha t}] \\ &= 0 \quad \text{for all } x \in S, \end{aligned}$$

which together with (7.6) gives (a). The proof of (b) is similar.

To prove (c), we define an operator T_{π^1, π^2} (similar to T in (5.3)) on $B_{w_1}(S)$ as follows: for $u \in B_{w_1}(S)$ and $x \in S$,

$$T_{\pi^1, \pi^2} u(x) := \frac{r(x, \pi^1, \pi^2)}{\alpha + m(x)} + \frac{m(x)}{\alpha + m(x)} \int_S u(y) P(dy \mid x, \pi^1, \pi^2),$$

with $P(\cdot \mid x, \phi, \psi)$ as defined in (5.4). Obviously, the operator T_{π^1, π^2} is monotone. Let $v_0 := u_0$ be as in (5.5) and let

$$v_n := T_{\pi^1, \pi^2}^n v_0 \quad \text{for } n = 1, 2, \dots$$

By Assumptions 5.1, 5.2(c), and 5.2(d), the functions v_0 and v_n are well defined for each $n \geq 1$. Moreover, under Assumptions 3.1(a) and 5.1(a) straightforward calculations give

$$\begin{aligned}
 v_1(x) &\geq -\frac{|r(x, \pi^1, \pi^2)|}{\alpha + m(x)} \\
 &\quad - \frac{m(x)}{\alpha + m(x)} \left(\frac{b_1 M_1}{\alpha(\alpha - c_1)} + \frac{M_1 w_1(x)}{\alpha - c_1} + \frac{M_1 \int_S w_1(y) q(dy \mid x, \pi^1, \pi^2)}{(\alpha - c_1)m(x)} \right) \\
 &\geq -\frac{M_1 w_1(x)}{\alpha + m(x)} - \frac{m(x)}{\alpha + m(x)} \left(\frac{b_1 M_1}{\alpha(\alpha - c_1)} + \frac{M_1 w_1(x)}{\alpha - c_1} + \frac{M_1(c_1 w_1(x) + b_1)}{(\alpha - c_1)m(x)} \right) \\
 &= -\frac{b_1 M_1}{\alpha(\alpha - c_1)} - \frac{M_1 w_1(x)}{\alpha - c_1} \\
 &= v_0(x) \quad (\text{by (5.5)}). \tag{7.7}
 \end{aligned}$$

Therefore,

$$v_0 \leq v_1 \leq \dots \leq v_n \dots,$$

and so $v_n \uparrow v$ for some function $v \geq v_n$ for all $n \geq 0$. Conversely, by induction and a straightforward calculation as in the proof of (7.7), we have

$$|v_n(x)| \leq \frac{b_1 M_1}{\alpha(\alpha - c_1)} + \frac{M_1 w_1(x)}{\alpha - c_1} \leq \frac{(\alpha + b_1) M_1}{\alpha(\alpha - c_1)} w_1(x) =: Y(x) \quad \text{for all } n \geq 0 \text{ and } x \in S,$$

and

$$\begin{aligned}
 \left| \int_S Y(y) P(dy \mid x, \pi^1, \pi^2) \right| &\leq \frac{(\alpha + b_1) M_1}{\alpha(\alpha - c_1)} \int_S w_1(y) P(dy \mid x, \pi^1, \pi^2) \\
 &\leq \frac{(\alpha + b_1) M_1}{\alpha(\alpha - c_1)} (w_1(x) + c_1 w_1(x) + b_1).
 \end{aligned}$$

Thus, v is in the space $B_{w_1}(S)$, and it follows, from the Dominated Convergence Theorem, that

$$v(x) = \frac{r(x, \pi^1, \pi^2)}{\alpha + m(x)} + \frac{m(x)}{\alpha + m(x)} \int_S v(y) P(dy \mid x, \pi^1, \pi^2) \quad \text{for all } x \in S,$$

or, equivalently (by (5.4)),

$$\alpha v(x) = r(x, \pi^1, \pi^2) + \int_S v(y) q(dy \mid x, \pi^1, \pi^2) \quad \text{for all } x \in S.$$

This equation together with (a) and (b) gives $v(x) = J(x, \pi^1, \pi^2)$. Thus, $J(x, \pi^1, \pi^2)$ is a solution to the equation in part (c). The uniqueness follows from (a) and (b) also.

Proof of Theorem 5.1. (a) By Assumption 5.1, (4.4), and Lemma 3.2, a straightforward calculation yields part (a).

(b) Obviously, the operator T in (5.3) is monotone. By Assumptions 5.1 and 5.2, u_0 and $u_n = T^n u_0$ are also well defined for each $n \geq 1$. Thus, under Assumptions 3.1(a) and 5.1(a),

as in the proof of (7.7), a straightforward calculation yields

$$\begin{aligned} u_1(x) &\geq -\frac{M_1 w_1(x)}{\alpha + m(x)} - \frac{m(x)}{\alpha + m(x)} \left(\frac{b_1 M_1}{\alpha(\alpha - c_1)} + \frac{M_1 w_1(x)}{\alpha - c_1} + \frac{M_1(c_1 w_1(x) + b_1)}{(\alpha - c_1)m(x)} \right) \\ &= -\frac{b_1 M_1}{\alpha(\alpha - c_1)} - \frac{M_1 w_1(x)}{\alpha - c_1} \\ &= u_0(x). \end{aligned}$$

Therefore,

$$u_0 \leq u_1 \leq \dots \leq u_n \dots,$$

and so $u_n \uparrow u^*$ for some function $u^* \geq u_n$ for all $n \geq 0$. Hence, for a moment assuming that u^* is in $B_{w_1}(S)$, we have $Tu^* \geq Tu_n = u_{n+1}$ for all $n \geq 1$, which gives

$$Tu^* \geq u^*. \tag{7.8}$$

Conversely, to prove that u^* is indeed in $B_{w_1}(S)$ it suffices to show that

$$|u_n(x)| \leq \frac{b_1 M_1}{\alpha(\alpha - c_1)} + \frac{M_1 w_1(x)}{\alpha - c_1} \leq \frac{(\alpha + b_1)M_1}{\alpha(\alpha - c_1)} w_1(x) \quad \text{for all } n \geq 0 \text{ and } x \in S.$$

This, however, follows by induction and a straightforward calculation as in the proof of (7.7).

Next, we will show that (7.8) holds with equality. Indeed, for each fixed $n \geq 1$, $x \in S$, and $\phi \in P(A(x))$, we have already proved that u_n is in $B_{w_1}(S)$, whereas, by Assumption 5.2(a), $P(B(x))$ is compact. Thus, by Lemma 7.1, there exists a $\psi_n^* \in P(B(x))$, which may also depend on x and ϕ , such that

$$\begin{aligned} u_{n+1}(x) &= \max_{\phi \in P(A(x))} \min_{\psi \in P(B(x))} \left\{ \frac{r(x, \phi, \psi)}{\alpha + m(x)} + \frac{m(x)}{\alpha + m(x)} \int_S u_n(y) P(dy | x, \phi, \psi) \right\} \\ &\geq \min_{\psi \in P(B(x))} \left\{ \frac{r(x, \phi, \psi)}{\alpha + m(x)} + \frac{m(x)}{\alpha + m(x)} \int_S u_n(y) P(dy | x, \phi, \psi) \right\} \\ &= \frac{r(x, \phi, \psi_n^*)}{\alpha + m(x)} + \frac{m(x)}{\alpha + m(x)} \int_S u_n(y) P(dy | x, \phi, \psi_n^*). \end{aligned} \tag{7.9}$$

Since $P(B(x))$ is compact, without loss of generality we may suppose that $\psi_n^* \rightarrow \psi^* \in P(B(x))$. Therefore, as $u_0 \leq u_n \uparrow u^* \in B_{w_1}(S)$, by the ‘extended Fatou lemma’, Lemma 8.3.7(b) of [14], and our Lemma 7.1, letting $n \rightarrow \infty$ in (7.9) we obtain

$$\begin{aligned} u^*(x) &\geq \frac{r(x, \phi, \psi^*)}{\alpha + m(x)} + \frac{m(x)}{\alpha + m(x)} \int_S u^*(y) P(dy | x, \phi, \psi^*) \\ &\geq \min_{\psi \in P(B(x))} \left\{ \frac{r(x, \phi, \psi)}{\alpha + m(x)} + \frac{m(x)}{\alpha + m(x)} \int_S u^*(y) P(dy | x, \phi, \psi) \right\}. \end{aligned} \tag{7.10}$$

As (7.10) holds for all $\phi \in P(A(x))$ and $x \in S$, we conclude that

$$u^* \geq Tu^*,$$

which together with (7.8) gives $u^* = Tu^*$. This completes the proof of part (b), because using (5.4) it is easily seen that the equation $u^* = Tu^*$ is equivalent to (5.6).

(c) Since $r(x, \phi, \psi)$ and $q(dy | x, \phi, \psi)$ are bilinear in ϕ and ψ , by (5.6), Fan’s minimax theorem in [4], and the minimax measurable selection theorem in [21] and [26], there exists a pair of strategies $(\pi_*^1, \pi_*^2) \in \Pi_1^s \times \Pi_2^s$ that satisfies (5.7)–(5.9), i.e.

$$\alpha u^*(x) = r(x, \pi_*^1, \pi_*^2) + \int_S u^*(y)q(dy | x, \pi_*^1, \pi_*^2) \tag{7.11}$$

$$= \max_{\phi \in P(A(x))} \left\{ r(x, \phi, \pi_*^2) + \int_S u^*(y)q(dy | x, \phi, \pi_*^2) \right\} \tag{7.12}$$

$$= \min_{\psi \in P(B(x))} \left\{ r(x, \pi_*^1, \psi) + \int_S u^*(y)q(dy | x, \pi_*^1, \psi) \right\}.$$

(d) Since $u^* \in B_{w_1}(S)$, by (7.11) and Lemma 7.3 we have

$$J(x, \pi_*^1, \pi_*^2) = u^*(x) \quad \text{for all } x \in S. \tag{7.13}$$

Now fix π_*^2 . Then, for any strategy $\pi^1 = (\pi_t^1) \in \Pi_1$, we have $\pi_t^1(\cdot | x) \in P(A(x))$ for all $t \geq 0$ and $x \in S$, and so, from (5.1), (5.2), and (7.12) we have

$$\alpha u^*(x) \geq r(x, \pi_t^1, \pi_*^2) + \int_S u^*(y)q(dy | x, \pi_t^1, \pi_*^2) \quad \text{for all } x \in S.$$

The latter inequality and Lemma 7.3(a) yield

$$u^*(x) \geq J(x, \pi^1, \pi_*^2) \quad \text{for all } \pi^1 \in \Pi_1 \text{ and } x \in S,$$

which together with (7.13) implies that

$$u^*(x) = \sup_{\pi^1 \in \Pi_1} J(x, \pi^1, \pi_*^2) \geq \inf_{\pi^2 \in \Pi_2} \sup_{\pi^1 \in \Pi_1} J(x, \pi^1, \pi^2) = U(x) \quad \text{for all } x \in S. \tag{7.14}$$

A similar argument, replacing π_*^2 with π_*^1 gives

$$u^*(x) = \inf_{\pi^2 \in \Pi_2} J(x, \pi_*^1, \pi^2) \leq \sup_{\pi^1 \in \Pi_1} \inf_{\pi^2 \in \Pi_2} J(x, \pi^1, \pi^2) = L(x) \quad \text{for all } x \in S. \tag{7.15}$$

By (7.15), (7.14), and (4.2) we obtain $J(x, \pi_*^1, \pi_*^2) = U(x) = L(x) = u^*(x) =: V(x)$ for all $x \in S$, which together with (7.13) gives part (d).

(e) This follows from (7.13), (7.14), and (7.15).

Proof of Theorem 5.2. From the proof of Theorem 5.1 and the definition of Δ , (5.10), we see that (b) implies (a). Conversely, if (a) holds, then from Lemma 7.3(c) and Theorem 5.1 we obtain (b). Therefore, (a) and (b) are equivalent.

We next prove that (b) implies (c). Using Lemma 7.2 and Dynkin’s formula, from (5.10) and (5.11) a standard argument gives, for each $x \in S$, $(\pi^1, \pi^2) \in \Pi_1^s \times \Pi_2^s$, and $t' \geq t \geq 0$,

$$E_x^{\pi^1, \pi^2} [M(t', \pi^1, \pi^2) | \mathcal{F}_t] = M(t, \pi^1, \pi^2) + E_x^{\pi^1, \pi^2} \left[\int_t^{t'} e^{-\alpha \tau} \Delta(x(\tau), \pi^1, \pi^2) d\tau \mid \mathcal{F}_t \right]. \tag{7.16}$$

By (7.16) we immediately see that (b) yields (c).

Finally, we shall show that (c) implies (b). First note that taking expectations in both sides of (7.16) gives

$$E_x^{\pi^1, \pi^2} M(t', \pi^1, \pi^2) = E_x^{\pi^1, \pi^2} M(t, \pi^1, \pi^2) + E_x^{\pi^1, \pi^2} \left[\int_t^{t'} \Delta(x(\tau), \pi^1, \pi^2) d\tau \right] \tag{7.17}$$

for all $t' > t \geq 0$ and $(\pi^1, \pi^2) \in \Pi_1^s \times \Pi_2^s$. Therefore, if (c)(i) holds, from (7.17) and the Fubini theorem we obtain

$$\begin{aligned} \int_t^{t'} E_x^{\pi_*^1, \pi_*^2} [\Delta(x(\tau), \pi_*^1, \pi_*^2)] d\tau &= E_x^{\pi_*^1, \pi_*^2} \left[\int_t^{t'} \Delta(x(\tau), \pi_*^1, \pi_*^2) d\tau \right] \\ &= 0 \quad \text{for all } t' > t \geq 0 \text{ and } x \in S. \end{aligned}$$

Hence,

$$E_x^{\pi_*^1, \pi_*^2} \Delta(x(t), \pi_*^1, \pi_*^2) = 0 \quad \text{for all } x \in S \text{ and almost all } t \geq 0.$$

Therefore, there exists a sequence $t_n \downarrow 0$ such that

$$E_x^{\pi_*^1, \pi_*^2} \Delta(x(t_n), \pi_*^1, \pi_*^2) = 0 \quad \text{for all } n \geq 0 \text{ and } x \in S. \tag{7.18}$$

Since $\Delta(\cdot, \pi_*^1, \pi_*^2)$ is in $B_{w_1+w_2}(S)$, by Lemma 7.2(a) we have

$$\lim_{t \rightarrow 0^+} E_x^{\pi_*^1, \pi_*^2} \Delta(x(t), \pi_*^1, \pi_*^2) = \Delta(x, \pi_*^1, \pi_*^2) \quad \text{for all } x \in S. \tag{7.19}$$

Letting $n \rightarrow \infty$ in (7.18), from (7.19) we obtain

$$\Delta(x, \pi_*^1, \pi_*^2) = 0 \quad \text{for all } x \in S. \tag{7.20}$$

Conversely, under (c)(ii), (7.17) gives

$$E_x^{\pi_*^1, \pi_*^2} \left[\int_s^t \Delta(x(\tau), \pi_*^1, \pi_*^2) d\tau \right] \geq 0 \quad \text{for all } t \geq s \geq 0, \ x \in S, \text{ and } \pi^2 \in \Pi_2^s.$$

Hence, as in the proof of (7.20), we can derive that

$$\Delta(x, \pi_*^1, \pi^2) \geq 0 \quad \text{for all } x \in S \text{ and } \pi \in \Pi_2^s. \tag{7.21}$$

Similarly, from (7.17) and (c)(iii) we can obtain

$$\Delta(x, \pi^1, \pi_*^2) \leq 0 \quad \text{for all } x \in S \text{ and } \pi^1 \in \Pi_1^s,$$

which together with (7.20) and (7.21), gives part (b).

References

- [1] ANDERSON, W. J. (1991). *Continuous-Time Markov Chains*. Springer, New York.
- [2] BORKAR, V. S. AND GHOSH, M. K. (1992). Stochastic differential games: an occupation measure based approach. *J. Optimization Theory Appl.* **73**, 359–385. (Correction: **88** (1996), 251–252.)
- [3] CHEN, M. F. (2004). *From Markov Chains to Non-Equilibrium Particle Systems*, 2nd edn. World Scientific, Singapore.
- [4] FAN, K. (1953). Minimax theorems. *Proc. Nat. Acad. Sci. USA* **39**, 42–47.

- [5] FELLER, W. (1940). On the integro-differential equations of purely discontinuous Markoff processes. *Trans. Amer. Math. Soc.* **48**, 488–515.
- [6] FILAR, J. A. AND VRIEZE, K. (1997). *Competitive Markov Decision Processes*. Springer, New York.
- [7] GHOSH, M. K. AND KUMAR, K. S. (2002). A stochastic differential game in the orthant. *J. Math. Appl. Appl.* **265**, 12–37.
- [8] GIHMAN, I. I. AND SKOROHOD, A. V. (1979). *Controlled Stochastic Processes*. Springer, New York.
- [9] GUO, X. P. (2007). Continuous-time Markov decision processes with discounted rewards: the case of Polish spaces. *Math. Operat. Res.* **32**, 73–87.
- [10] GUO, X. P. AND HERNÁNDEZ-LERMA, O. (2003). Continuous-time controlled Markov chains. *Ann. Appl. Prob.* **13**, 363–388.
- [11] GUO, X. P. AND HERNÁNDEZ-LERMA, O. (2003). Drift and monotonicity conditions for continuous-time controlled Markov chains an average criterion. *IEEE Trans. Automatic Control* **48**, 236–245.
- [12] GUO, X. P. AND HERNÁNDEZ-LERMA, O. (2003). Zero-sum games for continuous-time Markov chains with unbounded transition and average payoff rates. *J. Appl. Prob.* **40**, 327–345.
- [13] GUO, X. P. AND HERNÁNDEZ-LERMA, O. (2005). Zero-sum continuous-time Markov games with unbounded transition and discounted payoff rates. *Bernoulli* **11**, 1009–1029.
- [14] HERNÁNDEZ-LERMA, O. AND LASSERRE, J. B. (1999). *Further Topics on Discrete-Time Markov Control Processes*. Springer, New York.
- [15] HERNÁNDEZ-LERMA, O. AND LASSERRE, J. B. (2001). Zero-sum stochastic games in Borel spaces: average payoff criterion. *SIAM J. Control Optimization* **39**, 1520–1539.
- [16] JAŚKIEWICZ, A. (2002). Zero-sum semi-Markov games. *SIAM J. Control Optimization* **41**, 723–739.
- [17] KUSHNER, H. J. (2003). Numerical approximations for stochastic differential games: the ergodic case. *SIAM J. Control Optimization* **42**, 1911–1933.
- [18] LAI, H.-C. AND TANAKA, K. (1984). On an N -person noncooperative Markov game with a metric state space. *J. Math. Anal. Appl.* **101**, 78–96.
- [19] LAL, A. K. AND SINHA, S. (1992). Zero-sum two-person semi-Markov games. *J. Appl. Prob.* **29**, 56–72.
- [20] LUND, R. B., MEYN, S. P. AND TWEEDIE, R. L. (1996). Computable exponential convergence rates for stochastically ordered Markov processes. *Ann. Appl. Prob.* **6**, 218–237.
- [21] NOWAK, A. S. (1985). Measurable selection theorems for minimax stochastic optimization problems. *SIAM J. Control Optimization* **23**, 466–476.
- [22] NOWAK, A. S. (2000). Some remarks on equilibria in semi-Markov games. *Appl. Math. (Warsaw)* **27**, 385–394.
- [23] PARTHASARATHY, T. AND SINHA, S. (1989). Existence of equilibrium stationary strategies in nonzero-sum discounted stochastic games with uncountable state space, and state independent transitions. *Internat. J. Game Theory* **18**, 189–194.
- [24] PUTERMAN, M. L. (1994). *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley, New York.
- [25] RAO, M. M. (1995). *Stochastic Processes: General Theory*. Kluwer, Dordrecht.
- [26] RIEDER, U. (1978). On semi-continuous dynamic games. Preprint, Abt. für Mathematik VII, Universität Ulm.
- [27] SCHLÖGL, F. (1972). Chemical reaction models for phase transition. *Z. Phys.* **253**, 147–161.
- [28] SENNOTT, L. I. (1994). Zero-sum stochastic games with unbounded cost: discounted and average cost cases. *Z. Operat. Res.* **39**, 209–225.
- [29] VEGA-AMAYA, O. (2003). Zero-sum average semi-Markov games: fixed-point solutions of the Shapley equation. *SIAM J. Control Optimization* **42**, 1876–1984.