

# NONZERO-SUM GAMES FOR CONTINUOUS-TIME MARKOV CHAINS WITH UNBOUNDED DISCOUNTED PAYOFFS

XIANPING GUO,\* *CINVESTAV and Zhongshan University*

ONÉSIMO HERNÁNDEZ-LERMA,\*\* *CINVESTAV*

## Abstract

In this paper, we study two-person nonzero-sum games for continuous-time Markov chains with discounted payoff criteria and Borel action spaces. The transition rates are possibly unbounded, and the payoff functions might have neither upper nor lower bounds. We give conditions that ensure the existence of Nash equilibria in stationary strategies. For the zero-sum case, we prove the existence of the value of the game, and also provide a *recursive* way to compute it, or at least to approximate it. Our results are applied to a controlled queueing system. We also show that if the transition rates are *uniformly bounded*, then a continuous-time game is equivalent, in a suitable sense, to a discrete-time Markov game.

*Keywords:* Nonzero-sum game; discounted payoff criterion; Nash equilibrium; controlled Q-process

2000 Mathematics Subject Classification: Primary 91A15; 91A25

## 1. Introduction

Nonzero-sum stochastic dynamic games have been widely studied in the literature. The existing works can be roughly classified into two main groups. The first deals with discrete-time stochastic Markov games (see, for example, [4], [14], [17], [20], [22], and their references); the other with stochastic differential games (e.g. [10], [18], and [24]). In this paper, we study a *third* class of nonzero-sum stochastic games; namely, continuous-time stochastic Markov games in which the state process evolves according to a continuous-time Markov chain, the players can select their actions *continuously* in time, and the transition rates may be *unbounded*. The study of the third class of games is motivated by some applications of stochastic games to queueing systems, telecommunication networks, and other areas; see, for example, [19], [21], [23], [25], and their references.

More precisely, we consider nonzero-sum games for continuous-time Markov chains with a discounted payoff criterion and Borel action spaces, allowing unbounded transition rates and payoff functions that might have neither upper nor lower bounds. We give conditions under which the existence of Nash equilibria in stationary strategies is obtained by using Fan's fixed-point theorem and the optimality equations for the discounted payoff criterion. In the zero-sum case, we furthermore provide a *recursive* way to compute, or at least to approximate, the value

---

Received 5 July 2001; revision received 16 March 2004.

\* Postal address: Department of Mathematics, Zhongshan University, Guangzhou, 510275, PR China.

Email address: mcsghxp@zsu.edu.cn

\*\* Postal address: Department of Mathematics, CINVESTAV IPN, A. Postal 14-740, Mexico DF 07000, Mexico.

Email address: ohermand@math.cinvestav.mx

of the zero-sum game. We illustrate our strategy class with an example in which the players can select their actions continuously in time, and apply our results to a controlled queueing system.

Finally, we also show that if the Markov chain’s transition rates are uniformly bounded, then the continuous-time game is equivalent to a discrete-time Markov game in the sense that both games have the same set of Nash equilibria.

The rest of the paper is organized as follows. In Sections 2 and 3, we introduce the game model and the family of admissible strategies, respectively. The optimality criterion we are concerned with is presented in Section 4. Our main results are stated in Section 5 and illustrated with an example in Section 6. The proofs are all postponed to Section 8, after some technical preliminaries in Section 7. We conclude, in Section 9, with some general remarks.

### 2. The game model

In this section, we introduce the continuous-time two-person nonzero-sum stochastic game model

$$\{S, A, B, K_A, K_B, q, r_1, r_2\}, \tag{2.1}$$

where  $S$  is the state space, a denumerable set, and  $A$  and  $B$  are the *action spaces* for players 1 and 2, respectively. The action spaces are assumed to be Borel spaces. The sets  $K_A \subset S \times A$  and  $K_B \subset S \times B$  are Borel sets that represent the constraint sets for the two players. That is, for each state  $i \in S$ , the  $i$ -section in  $K_A$ , namely

$$A(i) := \{a \in A : (i, a) \in K_A\},$$

represents the set of admissible actions for player 1 in the state  $i$  and, similarly, the  $i$ -section in  $K_B$ ,

$$B(i) := \{b \in B : (i, b) \in K_B\},$$

stands for the family of admissible actions for player 2 in the state  $i$ . Let

$$K := \{(i, a, b) : i \in S, a \in A(i), b \in B(i)\}, \tag{2.2}$$

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

The function  $q$  in (2.1) is the matrix  $[q(j \mid i, a, b)]$  of the game’s transition rates, satisfying  $q(j \mid i, a, b) \geq 0$  for all  $(i, a, b) \in K$  and  $i \neq j$ . This matrix is assumed to be *conservative*, that is

$$\sum_{j \in S} q(j \mid i, a, b) = 0 \quad \text{for each } (i, a, b) \in K, \tag{2.3}$$

and *stable*, that is

$$q(i) := \sup_{a \in A(i), b \in B(i)} q_i(a, b) < \infty \quad \text{for each } i \in S,$$

where  $q_i(a, b) := -q(i \mid i, a, b)$  for all  $a \in A(i)$  and  $b \in B(i)$ . Moreover,  $q(j \mid i, a, b)$  is a measurable function on  $A \times B$  for each fixed  $i, j \in S$ . Finally,  $r_k : K \rightarrow \mathbb{R} := (-\infty, \infty)$  is the payoff function for player  $k = 1, 2$ .

The game is played as follows. The players continuously observe the current state of the game. Whenever the game is in state  $i \in S$  at time  $t \geq 0$ , they independently choose actions  $a_t \in A(i)$  (player 1) and  $b_t \in B(i)$  (player 2) according to some rules. As a consequence of this, the following happens: (i) player  $k$  ( $k = 1, 2$ ) receives an immediate payoff  $r_k(i, a_t, b_t)$ , and (ii) the system moves to a new state  $j \neq i$  with a possibly inhomogeneous transition probability function determined by the transition rates  $q(j \mid i, a_t, b_t)$ . Thus, the goal of the players is to individually maximize, in the sense of Definition 4.1, below, their expected rewards.

### 3. Strategies

We will use the following notation. If  $X$  is a Borel space, we denote by  $\mathcal{B}(X)$  its Borel  $\sigma$ -algebra, and by  $P(X)$  the Borel space of probability measures on  $X$ , endowed with the topology of weak convergence. A *randomized Markov strategy* for player 1 is a family  $\pi^1 = (\pi_t^1, t \geq 0)$  of stochastic kernels such that: (i) for each  $t \geq 0$  and  $i \in S, \pi_t^1(\cdot | i)$  is a probability measure on  $A$  such that  $\pi_t^1(A(i) | i) = 1$ , and (ii) for every  $E \in \mathcal{B}(A)$  and  $i \in S, \pi_t^1(E | i)$  is a Borel-measurable function in  $t \geq 0$ . We denote by  $\Pi_1^M$  the family of all randomized Markov strategies for player 1. Moreover, a strategy  $\pi^1 = (\pi_t^1, t \geq 0) \in \Pi_1^M$  is called *stationary* if, for each  $i \in S$ , there is a probability measure  $\pi^1(\cdot | i) \in P(A(i))$  such that

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

We denote this policy by  $(\pi^1(\cdot | i), i \in S)$ . The set of all stationary policies for player 1 is denoted by  $\Pi_1^S$ . The sets of all randomized Markov strategies  $\Pi_2^M$  and all stationary strategies  $\Pi_2^S$  for player 2 are defined similarly, with  $P(B(i))$  in place of  $P(A(i))$ . Thus, we see that  $\Pi_1^S = \prod_{i \in S} P(A(i))$  and  $\Pi_2^S = \prod_{i \in S} P(B(i))$  are also Borel sets.

For each pair of strategies  $(\pi^1, \pi^2) := ((\pi_t^1, \pi_t^2), t \geq 0) \in \Pi_1^M \times \Pi_2^M$ , the associated transition rates and payoff functions are defined, respectively, as follows: for each  $i, j \in S$  and  $t \geq 0$ ,

$$q(j | i, t, \pi^1, \pi^2) := \int_B \int_A q(j | i, a, b) \pi_t^1(da | i) \pi_t^2(db | i), \tag{3.1}$$

$$r_k(t, i, \pi^1, \pi^2) := \int_B \int_A r_k(i, a, b) \pi_t^1(da | i) \pi_t^2(db | i). \tag{3.2}$$

In particular, when  $\pi^1$  and  $\pi^2$  are both stationary, we write the left-hand sides of (3.1) and (3.2) as  $q(j | i, \pi^1, \pi^2)$  and  $r_k(i, \pi^1, \pi^2)$ , respectively. In addition, the associated  $Q$ -matrix is  $Q(t, \pi^1, \pi^2) := [q(j | i, t, \pi^1, \pi^2)]$ .

Any transition function  $p(s, i, t, j, \pi^1, \pi^2)$  (with  $i, j \in S$  and  $t \geq s \geq 0$ ) for which  $\{Q(t, \pi^1, \pi^2), t \geq 0\}$  are the transition rate matrices is called a  $Q$ -process. To guarantee the existence of such processes, we will restrict ourselves to control strategies in the classes  $\Pi_1$  and  $\Pi_2$  defined as follows.

**Definition 3.1.** For  $k = 1, 2, \Pi_k$  is any subset of randomized Markov strategies for player  $k$  such that  $\Pi_k$  contains  $\Pi_k^S$  and satisfies either

$$\begin{aligned} \Pi_1 &:= \{\pi^1 \in \Pi_1^M : \\ &\quad q(j | i, t, \pi^1, \pi^2) \text{ is continuous in } t \text{ for each fixed } i, j \in S \text{ and } \pi^2 \in \Pi_2^M\} \end{aligned}$$

or

$$\begin{aligned} \Pi_2 &:= \{\pi^2 \in \Pi_2^M : \\ &\quad q(j | i, t, \pi^1, \pi^2) \text{ is continuous in } t \text{ for each fixed } i, j \in S \text{ and } \pi^1 \in \Pi_1^M\}, \end{aligned}$$

as appropriate.

Observe that  $\Pi_1 \times \Pi_2$  contains  $\Pi_1^S \times \Pi_2^S$  and, as the matrix  $[q(j | i, a, b)]$  is conservative and stable, that  $Q(t, \pi^1, \pi^2)$  is also conservative and stable. Hence, for each  $\pi^1 \in \Pi_1$  and  $\pi^2 \in \Pi_2$ , the existence of a  $Q$ -process is indeed guaranteed but, as is well known [1], [2], [16], it is not necessarily unique. Thus, to guarantee the uniqueness of a  $Q$ -process, throughout this paper we make the following assumption.

**Assumption 3.1.** *There exist a sequence  $\{S_m, m \geq 1\}$  of subsets of  $S$ , a nonnegative function  $R \geq 1$  on  $S$ , and positive constants  $c$  and  $d$  such that*

1.  $S_m \uparrow S$  and  $\sup_{i \in S_m} q(i) < \infty$  for each  $m \geq 1$ ;
2.  $\lim_{m \rightarrow \infty} [\inf_{j \notin S_m} R(j)] = \infty$ ; and
3.  $\sum_{j \in S} q(j \mid i, a, b)R(j) \leq cR(i) + d$  for all  $(i, a, b) \in K$  (the set defined in (2.2)).

**Remarks 3.1.** (a) For the case of uniformly bounded transition rates (i.e.  $\sup_{i \in S} q(i) < \infty$  as in, for example, [19], [21], and [23]), Assumptions 3.1.1 and 3.1.2 are not required because they are only used to guarantee the uniqueness of a Q-process. For the case of *unbounded* transition rates, Assumption 3.1 is a variant of hypotheses used in [5], [6], [7], [8], [9], and [15] for continuous-time Markov decision (or control) processes. Thus, sufficient conditions and examples satisfying Assumption 3.1 can be given as in the latter references, for instance.

(b) Under Assumption 3.1, it follows from Lemma 7.1, below, that a Q-process with transition rate matrices  $Q(t, \pi^1, \pi^2)$  is unique and honest (i.e. regular). Thus, we denote by  $\{x(t, \pi^1, \pi^2)\}$  the associated, and possibly inhomogeneous, Markov process with the given transition rates  $Q(t, \pi^1, \pi^2)$ , and write the unique and honest transition function as  $p(s, i, t, j, \pi^1, \pi^2)$ . Furthermore, for each initial state  $i \in S$  at time  $s$ , we use  $E_{s,i}^{\pi^1, \pi^2}$  to denote the probability expectation operator determined by  $p(s, i, t, j, \pi^1, \pi^2)$ .

#### 4. Discounted payoff criteria

For each  $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$ ,  $(s, i) \in \bar{S} := [0, \infty) \times S$ , and a fixed discount factor  $\alpha > 0$ , the discounted payoff criterion  $V_\alpha^k(s, i, \pi^1, \pi^2)$ ,  $k = 1, 2$ , is defined as

$$V_\alpha^k(s, i, \pi^1, \pi^2) := E_{s,i}^{\pi^1, \pi^2} \left[ \int_s^\infty e^{-\alpha(t-s)} r_k(t, x(t, \pi^1, \pi^2), \pi^1, \pi^2) dt \right]. \tag{4.1}$$

For the criterion (4.1) to be well defined and finite, we make the following assumption.

**Assumption 4.1.** *With  $c$  and  $R$  as in Assumption 3.1,*

1. *either  $c \leq 0$ , or  $\alpha - c > 0$  when  $c > 0$ ; and*
2. *there exist nonnegative constants  $M_1$  and  $M_2$  such that*

$$|r_k(i, a, b)| \leq M_1 + M_2R(i) \quad \text{for all } (i, a, b) \in K \text{ and } k = 1, 2.$$

**Remark 4.1.** Assumption 4.1.2 holds trivially if the payoff functions  $r_k$  are uniformly bounded (i.e.  $\sup_{(i,a,b) \in K, k=1,2} |r^k(i, a, b)| < \infty$ ). In Section 6, we give an example for which both Assumptions 3.1 and 4.1 hold and the payoff functions and transition rates are all unbounded.

We now introduce our optimality criterion.

**Definition 4.1.** A pair of strategies  $(\pi^{*1}, \pi^{*2}) \in \Pi_1 \times \Pi_2$  is called a *Nash* (or *noncooperative*) *equilibrium* for the discounted criterion (4.1) if

$$V_\alpha^1(s, i, \pi^{*1}, \pi^{*2}) \geq V_\alpha^1(s, i, \pi^1, \pi^{*2}) \quad \text{for all } (s, i) \in \bar{S} \text{ and } \pi^1 \in \Pi_1 \tag{4.2}$$

and

$$V_\alpha^2(s, i, \pi^{*1}, \pi^{*2}) \geq V_\alpha^2(s, i, \pi^{*1}, \pi^2) \quad \text{for all } (s, i) \in \bar{S} \text{ and } \pi^2 \in \Pi_2. \tag{4.3}$$

In the *zero-sum* case, i.e. when  $r_1 = -r_2$ , a Nash equilibrium is also known as a *saddle point*.

Our aim is to establish, under certain assumptions, the existence of a Nash equilibrium  $(\pi^{*1}, \pi^{*2})$  in the set  $\Pi_1^s \times \Pi_2^s$  of pairs of stationary strategies.

### 5. Existence of Nash equilibria

To ensure the existence of a Nash equilibrium, in addition to Assumptions 3.1 and 4.1 we also need the following conditions.

**Assumption 5.1.** 1. For each  $i \in S$ ,  $A(i)$  and  $B(i)$  are compact.

2. The payoff function  $r_k(i, a, b)$  is continuous in  $(a, b) \in A(i) \times B(i)$  for each  $i \in S$  and  $k = 1, 2$ .

3. For each  $i \in S$ , the function  $\sum_{j \in S} q(j | i, a, b)u(j)$  is continuous in  $(a, b) \in A(i) \times B(i)$  for each bounded function  $u$  on  $S$  and also for  $u = R$ , as in Assumption 3.1.

4. There exist a nonnegative function  $R'$  on  $S$  and constants  $c' > 0, d' \geq 0$ , and  $M' > 0$  such that

$$q(i)R(i) \leq M'R'(i)$$

$$\sum_{j \in S} q(j | i, a, b)R'(j) \leq c'R'(i) + d'$$

for all  $(i, a, b) \in K$ , with  $R$  as in Assumption 3.1.

**Remark 5.1.** Assumptions 5.1.1–5.1.3 consist of standard continuity–compactness hypotheses for continuous-time Markov control processes; see, e.g. [5], [6], [7], [8], [9], [11], [12], [15], [23], and [26]. Assumption 5.1.4 is necessary for the application of Dynkin’s formula, and it can be dropped when  $q(i)$  is bounded in  $i \in S$ . On the other hand, if the  $r_k(i, a, b), k = 1, 2$ , are uniformly bounded in  $K$ , then the continuity condition for  $u = R$  in Assumption 5.1.3 is not required.

To state our results, we need the following notation. For the function  $R \geq 1$  in Assumption 3.1, we define the weighted supremum norm  $\|\cdot\|_R$  for real-valued functions  $u$  on  $S$  by

$$\|u\|_R := \sup_{i \in S} \{R(i)^{-1}|u(i)|\};$$

we define the Banach space

$$B(S) := \{u : \|u\|_R < \infty\};$$

and we define the operator  $T$  on  $B(S)$  as

$$Tu(i) := \sup_{\phi \in P(A(i))} \inf_{\psi \in P(B(i))} \left\{ \frac{r_1(i, \phi, \psi)}{\alpha + m(i)} + \frac{m(i)}{\alpha + m(i)} \sum_{j \in S} \left[ \frac{q(j | i, \phi, \psi)}{m(i)} + \delta_{ij} \right] u(j) \right\} \quad (5.1)$$

for all  $i \in S$  and  $u \in B(S)$ , where  $m(i)$  is such that  $m(i) > q(i)$  for all  $i \in S$ .

**Theorem 5.1.** *Suppose that Assumptions 3.1, 4.1, and 5.1 hold. Then the following statements also hold.*

(a) *There exists a pair  $(\pi^{*1}, \pi^{*2}) \in \Pi_1^s \times \Pi_2^s$  such that*

$$\begin{aligned} & \alpha V_\alpha^1(i, \pi^{*1}, \pi^{*2}) \\ &= \sup_{\pi^1 \in \Pi_1^s} \left\{ r_1(i, \pi^1, \pi^{*2}) + \sum_{j \in S} q(j \mid i, \pi^1, \pi^{*2}) V_\alpha^1(j, \pi^{*1}, \pi^{*2}) \right\} \text{ for all } i \in S, \end{aligned} \tag{5.2}$$

$$\begin{aligned} & \alpha V_\alpha^2(i, \pi^{*1}, \pi^{*2}) \\ &= \sup_{\pi^2 \in \Pi_2^s} \left\{ r_2(i, \pi^{*1}, \pi^2) + \sum_{j \in S} q(j \mid i, \pi^{*1}, \pi^2) V_\alpha^2(j, \pi^{*1}, \pi^{*2}) \right\} \text{ for all } i \in S. \end{aligned} \tag{5.3}$$

(b) *The pair  $(\pi^{*1}, \pi^{*2})$  obtained in part (a) is a Nash equilibrium. Moreover, a pair of strategies  $(\pi^{*1}, \pi^{*2}) \in \Pi_1^s \times \Pi_2^s$  is a Nash equilibrium if and only if (5.2) and (5.3) are satisfied.*

(c) *If, in addition,  $r_1 = -r_2$  (i.e. the zero-sum case), then the function*

$$V(i) := \sup_{\pi^1 \in \Pi_1} \inf_{\pi^2 \in \Pi_2} V_\alpha^1(s, i, \pi^1, \pi^2) = \inf_{\pi^2 \in \Pi_2} \sup_{\pi^1 \in \Pi_1} V_\alpha^1(s, i, \pi^1, \pi^2) \text{ for all } i \in S, \tag{5.4}$$

*which is called the value of the zero-sum game, is independent of  $s \geq 0$ . Furthermore,  $V(\cdot)$  can be computed to be*

$$V(i) = \lim_{n \rightarrow \infty} u_n(i) \text{ for all } i \in S, \tag{5.5}$$

*where  $u_n(i) := Tu_{n-1}(i)$  for all  $n \geq 1$ , with  $u_0(i) := -[M_1/\alpha + dM_2/(\alpha(\alpha - c)) + M_2R(i)/(\alpha - c)]$ . (Here,  $R(i)$ ,  $c$ ,  $d$ ,  $M_1$ , and  $M_2$  are as in Assumptions 3.1 and 4.1.)*

*Proof.* See Section 8.

We now show the relationship between discrete- and continuous-time stochastic Markov games. To this end, we introduce the following notation. Suppose that the transition rates  $q(j \mid i, a, b)$  are uniformly bounded, i.e.  $\|q\| := \sup_{i \in S} q(i) < \infty$ . Then, we can define a stochastic kernel  $P(j \mid i, a, b)$  on  $S$ , given  $K$  and new payoff functions  $r_k^d(i, a, b)$ , as follows: for each  $(i, a, b) \in K$ ,  $j \in S$ , and  $k = 1, 2$ , let

$$P(j \mid i, a, b) := \frac{q(j \mid i, a, b)}{1 + \|q\|} + \delta_{ij}, \quad r_k^d(i, a, b) := \frac{r_k(i, a, b)}{\alpha + 1 + \|q\|}, \quad \beta = \frac{1 + \|q\|}{\alpha + 1 + \|q\|}. \tag{5.6}$$

From (2.3) and (5.6) we see that  $P(j \mid i, a, b)$  is indeed a stochastic kernel on  $S$ , given  $K$ . We now define a *discrete-time* (two-person nonzero-sum) Markov game model with transition probabilities  $P(j \mid i, a, b)$ , payoffs  $r_k^d(i, a, b)$ , and discount factor  $\beta$ :

$$\{S, A, B, K_A, K_B, P(j \mid i, a, b), r_1^d(i, a, b), r_2^d(i, a, b), \beta\} \tag{5.7}$$

with  $S, A, B, K_A$ , and  $K_B$  as in (2.1).

For the discrete-time model (5.7) with discount factor  $\beta$  (cf. [20] and [22]), we can also define the discounted payoff criterion and then a Nash equilibrium, as in Section 4.

**Theorem 5.2.** *Suppose that Assumptions 3.1, 4.1, and 5.1 hold and, in addition, that  $\|q\| < \infty$ . Then, the continuous-time model (2.1) and the discrete-time model (5.7) have the same Nash equilibria. In other words, if a pair of stationary strategies is a Nash equilibrium for one of the models, then it is so for the other.*

*Proof.* See Section 8.

**Remark 5.2.** Theorem 5.2 shows the relationship between the two classes of Markov games in discrete and continuous time, under the condition  $\|q\| < \infty$ . This condition is very important because it is used to define the discount factor  $\beta$  for the discrete-time model (5.7). However, such a condition may of course not hold in general; see Remark 6.1(b), below. Therefore, in our continuous-time model (2.1) with *unbounded* transition rates, the existence of a Nash equilibrium cannot be obtained by introducing a discrete-time model such as (5.7).

### 6. Examples

In this section, we apply our results to a controlled queueing system in Example 6.1, and then illustrate our classes of strategies with Example 6.2.

#### 6.1. Example 6.1

Consider a single-server queueing system in which the state variable denotes the total number of jobs (in service and waiting in the queue) at each time  $t \geq 0$ . There are ‘natural’ arrival and service rates, say  $\lambda$  and  $\mu$ , respectively, in addition to service parameters  $u(a)$  controlled by player 1, and arrival parameters  $v(b)$  controlled by player 2. When the state of the system is  $i \in S := \{0, 1, \dots\}$ , player 1 takes an action  $a$  from a given set  $A(i) \subset A$ , which may increase ( $u(a) \geq 0$ ) or decrease ( $u(a) \leq 0$ ) the service rate. These actions produce a reward or cost denoted by  $c_1(a) \geq 0$  or  $c_1(a) \leq 0$ , respectively, per unit time. Similarly, when the state is  $i \in S$ , player 2 takes an action  $b$ , from a set  $B(i) \subset B$ , that rejects ( $v(b) \leq 0$ ) or admits ( $v(b) \geq 0$ ) customers, and these actions result in a reward or cost rate  $c_2(b) \geq 0$  or  $c_2(b) \leq 0$ , respectively. In addition, we assume that player  $k$  ( $k = 1, 2$ ) gets a reward  $p_k i$  or incurs a cost  $p_k i$  for each unit of time during which the system remains in state  $i$ . Here,  $p_k > 0$  is a fixed reward fee, and  $p_k < 0$  a fixed cost fee, per customer, for player  $k$ .

We now formulate this model as a continuous-time Markov game. The corresponding transition rates  $q(j | i, a, b)$  and payoff functions  $r_k(i, a, b)$  for player  $k$  are as follows. For  $i = 0, a \in A(0)$ , and  $b \in B(0)$ ,

$$q(1 | 0, a, b) = -q(0 | 0, a, b) := u(a) + v(b)$$

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

$$q(j | i, a, b) = \begin{cases} \mu i + u(a) & \text{if } j = i - 1, \\ -(\mu + \lambda)i - u(a) - v(b) & \text{if } j = i, \\ \lambda i + v(b) & \text{if } j = i + 1, \\ 0 & \text{otherwise,} \end{cases} \tag{6.1}$$

$$r_1(i, a, b) = p_1 i - c_1(a), \quad r_2(i, a, b) = p_2 i - c_2(b) \quad \text{for } (i, a, b) \in K, \tag{6.2}$$

with  $K$  as in (2.2).

The aim now is to find conditions under which there exists a Nash equilibrium in the set of stationary strategies. To do so, we consider the following conditions.

**Condition 6.1.** (a) For each  $i \in S$ ,  $\mu i + u(a) \geq 0$  for all  $a \in A(i)$  and  $\lambda i + v(b) \geq 0$  for all  $b \in B(i)$ .

(b) Either  $0 \leq \lambda \leq \mu$ , or  $\alpha > \lambda - \mu$  when  $\lambda > \mu$ .

**Condition 6.2.** The action sets  $A$  and  $B$  are metric spaces, and  $A(i)$  and  $B(i)$  are compact for each  $i \in S$ .

**Condition 6.3.** The functions  $c_1(a)$ ,  $c_2(b)$ ,  $u(a)$ , and  $v(b)$  are bounded in the supremum norm and continuous on their respective domains.

Under these conditions, we obtain the following result.

**Proposition 6.1.** Under Conditions 6.1, 6.2, and 6.3, the above queueing system satisfies Assumptions 3.1, 4.1, and 5.1. Therefore (by Theorem 5.1), there exists a Nash equilibrium in the set of stationary strategies.

*Proof.* We shall first verify that Assumption 3.1 is satisfied. Let  $S_m := \{0, 1, \dots, m\}$  for each  $m \geq 1$ ,  $R(i) := i + 1$  for all  $i \in S$ ,  $\|u\| := \sup_{a \in A} |u(a)|$ ,  $\|v\| := \sup_{b \in B} |v(b)|$ ,  $\|c_1\| := \sup_{a \in A} |c_1(a)|$ , and  $\|c_2\| := \sup_{b \in B} |c_2(b)|$ . Then Assumptions 3.1.1 and 3.1.2 are obviously true and, for each  $(i, a, b) \in K$ , from (6.1) we have

$$\begin{aligned} \sum_{j \in S} q(j \mid i, a, b)R(j) &= (\lambda - \mu)i - u(a) + v(b) \\ &\leq (\lambda - \mu)R(i) + (\lambda + \mu + \|u\| + \|v\|), \quad i \geq 1, \\ \sum_{j \in S} q(j \mid i, a, b)R(j) &= u(a) + v(b) \\ &\leq R(0) + (\lambda + \mu + \|u\| + \|v\|), \quad i = 0. \end{aligned}$$

These inequalities yield Assumption 3.1.3 with  $c = \lambda - \mu$ ,  $d = \lambda + \mu + \|u\| + \|v\|$ , and  $R(i) = i + 1$ . To verify Assumption 4.1, let  $p^* := \max\{|p_1|, |p_2|\}$ . Then  $|r_k(i, a, b)| \leq p^*R(i) + \|c_1\| + \|c_2\|$  for all  $i \in S$ , which, together with Condition 6.1(b), gives Assumption 4.1. Finally, from Conditions 6.1–6.3 together with (6.1) and (6.2), we see that Assumptions 5.1.1–5.1.3 hold. Moreover, as in the verification of Assumption 3.1 (replacing  $R(i)$  with  $R'(i) := (i + 1)^2$ ), we also see that Assumption 5.1.4 is satisfied and, so, Assumption 5.1 follows in its entirety.

**Remark 6.1.** (a) This example improves upon one given in [6] for zero-sum games, in which (i) the payoff function was assumed to be bounded below and (ii) the condition  $\lambda \leq \mu$  was necessary.

(b) It should be noted that, in Example 6.1, both the reward and transition rates are unbounded when  $\lambda$  (or  $\mu$ ) is positive, and that the transition rates are uniformly bounded when  $\lambda = \mu = 0$ .

In the following example, we will further show that the players can select their actions *continuously* in time.

**6.2. Example 6.2**

In Example 6.1, take arbitrary actions  $a_k^i$ ,  $k = 1, 2$ , from  $A(i)$  and  $b_k^i$ ,  $k = 1, 2$ , from  $B(i)$ , for each  $i \in S$ , and then define the *inhomogeneous* (randomized) Markov strategies

$\tilde{\pi}^1 = (\tilde{\pi}_t^1, t \geq 0)$  and  $\tilde{\pi}^2 = (\tilde{\pi}_t^2, t \geq 0)$  as, respectively,

$$\tilde{\pi}_t^1(E \mid i) = \begin{cases} \frac{1}{2}e^{-\rho it} & \text{if } E = \{a_1^i\}, \\ 1 - \frac{1}{2}e^{-\rho it} & \text{if } E = \{a_2^i\}, \\ 0 & \text{otherwise,} \end{cases} \tag{6.3}$$

with a fixed constant  $\rho > 0$ , and

$$\tilde{\pi}_t^2(E \mid i) = \begin{cases} \sin^2(it) & \text{if } E = \{b_1^i\}, \\ \cos^2(it) & \text{if } E = \{b_2^i\}, \\ 0 & \text{otherwise.} \end{cases} \tag{6.4}$$

Moreover, let  $\Pi_1 := \Pi_1^s \cup \{\tilde{\pi}^1\}$  and  $\Pi_2 := \Pi_2^s \cup \{\tilde{\pi}^2\}$ . By (6.1), (6.3), (6.4), and (3.1), we see that  $\Pi_1$  and  $\Pi_2$  satisfy the requirements of Definition 3.1, and that  $\Pi_1 \supset \Pi_1^s$ ,  $\Pi_1 \neq \Pi_1^s$ ,  $\Pi_2 \supset \Pi_2^s$ , and  $\Pi_2 \neq \Pi_2^s$ .

**Remark 6.2.** Obviously, when using the inhomogeneous strategies, the players select their actions *continuously* in time  $t \geq 0$ .

### 7. Technical preliminaries

In this section, we present some results needed to prove Theorems 5.1 and 5.2. Some of these results are already known, but we state them here for completeness and ease of reference.

**Lemma 7.1.** *Suppose that Assumption 3.1 holds. Then the following statements also hold.*

- (a) *For each  $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$ , the associated  $Q$ -process is regular.*
- (b) *If, in addition, Assumption 4.1 holds, then there exist nonnegative constants  $M_3$  and  $M_4$  (which only depend on the constants  $c, d$ , and  $\alpha$ ) such that, for each  $(\pi^1, \pi^2) \in \Pi_1 \times \Pi_2$  and  $(s, i) \in \tilde{S}$ ,*

$$|V_\alpha^k(s, i, \pi^1, \pi^2)| \leq M_3 + M_4 R(i) \quad \text{for } k = 1, 2. \tag{7.1}$$

*Proof.* The lemma follows from [5, Theorems 3.1(a) and 3.2(b)].

In the proof of Theorem 5.1, we will use some results on continuous-time Markov control processes. To state these results we introduce the following notation. For each  $i, j \in S$ ,  $\psi \in P(B(i))$ , and stationary strategies  $\pi^k = (\pi^k(\cdot \mid i), i \in S) \in \Pi_k^s, k = 1, 2$ , let

$$\begin{aligned} V_*^1(i, \pi^2) &:= \sup_{\pi^1 \in \Pi_1} V_\alpha^1(0, i, \pi^1, \pi^2), \\ V_*^2(i, \pi^1) &:= \sup_{\pi^2 \in \Pi_2} V_\alpha^2(0, i, \pi^1, \pi^2), \\ q(j \mid i, \pi^1, \psi) &:= \int_{B(i)} \int_{A(i)} q(j \mid i, a, b) \pi^1(da \mid i) \psi(db), \end{aligned} \tag{7.2}$$

$$r_1(i, \pi^1, \psi) := \int_{B(i)} \int_{A(i)} r_1(i, a, b) \pi^1(da \mid i) \psi(db). \tag{7.3}$$

Taking  $\phi \in P(A(i))$  rather than  $\psi \in P(B(i))$ , we define  $q(j \mid i, \phi, \pi^2)$  and  $r_2(i, \phi, \pi^2)$  similarly.

**Lemma 7.2.** *Suppose that Assumptions 3.1, 4.1, and 5.1 hold. Then the following statements also hold.*

(a) *For each pair  $(\pi^1, \pi^2) \in \Pi_1^s \times \Pi_2^s$  and  $k = 1, 2$ ,*

$$V_\alpha^k(s, i, \pi^1, \pi^2) = V_\alpha^k(0, i, \pi^1, \pi^2) =: V_\alpha^k(i, \pi^1, \pi^2) \text{ for all } i \in S \text{ and } s \geq 0.$$

*Moreover,  $V_\alpha^k(i, \pi^1, \pi^2)$  is the unique solution in  $B(S)$  of the equation*

$$\alpha u(i) = r_k(i, \pi^1, \pi^2) + \sum_{j \in S} q(j | i, \pi^1, \pi^2)u(j) \text{ for all } i \in S.$$

(b) *For each strategy  $\pi^2 \in \Pi_2^s$ , there exists a strategy  $\pi^{*1} \in \Pi_1^s$  such that*

$$\begin{aligned} V_\alpha^1(i, \pi^{*1}, \pi^2) &= r_1(i, \pi^{*1}, \pi^2) + \sum_{j \in S} q(j | i, \pi^{*1}, \pi^2)V_\alpha^1(j, \pi^{*1}, \pi^2) \\ &= \sup_{\pi^1 \in \Pi_1^s} \left\{ r_1(i, \pi^1, \pi^2) + \sum_{j \in S} q(j | i, \pi^1, \pi^2)V_\alpha^1(j, \pi^1, \pi^2) \right\} \text{ for all } i \in S, \end{aligned}$$

*and  $V_*^1(\cdot, \pi^2)$  is the unique solution in  $B(S)$  of the optimality equation*

$$\alpha u(i) = \sup_{\pi^1 \in \Pi_1^s} \left\{ r_1(i, \pi^1, \pi^2) + \sum_{j \in S} q(j | i, \pi^1, \pi^2)u(j) \right\} \text{ for all } i \in S.$$

(c) *Similarly, for each strategy  $\pi^1 \in \Pi_1^s$ , there exists a strategy  $\pi^{*2} \in \Pi_2^s$  such that*

$$\begin{aligned} \alpha V_\alpha^2(i, \pi^1, \pi^{*2}) &= r_2(i, \pi^1, \pi^{*2}) + \sum_{j \in S} q(j | i, \pi^1, \pi^{*2})V_\alpha^2(j, \pi^1, \pi^{*2}) \\ &= \sup_{\pi^2 \in \Pi_2^s} \left\{ r_2(i, \pi^1, \pi^2) + \sum_{j \in S} q(j | i, \pi^1, \pi^2)V_\alpha^2(j, \pi^1, \pi^2) \right\} \text{ for all } i \in S, \end{aligned}$$

*and  $V_*^2(\cdot, \pi^1)$  is the unique solution in  $B(S)$  of the optimality equation*

$$\alpha u(i) = \sup_{\pi^2 \in \Pi_2^s} \left\{ r_2(i, \pi^1, \pi^2) + \sum_{j \in S} q(j | i, \pi^1, \pi^2)u(j) \right\} \text{ for all } i \in S.$$

(d) *For each  $\pi^1 = (\pi_r^1, t \geq 0) \in \Pi_1$  and  $\pi^2 = (\pi_r^2, t \geq 0) \in \Pi_2$ , if there exists a function  $u \in B(S)$  such that*

$$\alpha u(i) \geq r_k(t, i, \pi^1, \pi^2) + \sum_{j \in S} q(j | i, t, \pi^1, \pi^2)u(j) \text{ for all } i \in S \text{ and } t \geq 0,$$

*then  $u(i) \geq V_\alpha^k(s, i, \pi^1, \pi^2)$  for all  $(s, i) \in \bar{S}$ , with  $k = 1, 2$ .*

(e)  *$V_\alpha^k(i, \pi^1, \pi^2)$  is continuous in  $(\pi^1, \pi^2) \in \Pi_1^s \times \Pi_2^s$  for each fixed  $i \in S$  and  $k = 1, 2$ .*

*Proof.* Part (a) can be proved similarly to Lemma 6.2(c) of [5], whereas part (d) follows from [5, Lemma 6.2(a)] and parts (b) and (c) follow from [5, Theorem 3.2]. To prove part (e), it suffices to consider the case  $k = 1$  because the case  $k = 2$  is, of course, similar. Let  $(\pi_n^1, \pi_n^2) \rightarrow (\pi^1, \pi^2)$  and  $\limsup_{n \rightarrow \infty} V_\alpha^1(i, \pi_n^1, \pi_n^2) =: v(i)$  for all  $i \in S$ . Then, there exists a subsequence  $\{(\pi_m^1, \pi_m^2)\}$  of  $\{(\pi_n^1, \pi_n^2)\}$  such that

$$\lim_{m \rightarrow \infty} V_\alpha^1(i, \pi_m^1, \pi_m^2) = v(i) \quad \text{and} \quad \lim_{m \rightarrow \infty} r_1(i, \pi_m^1, \pi_m^2) = r_1(i, \pi^1, \pi^2) \quad \text{for all } i \in S, \tag{7.4}$$

where the latter equality follows from Assumption 5.1.2 and the definition of weak convergence of probability measures. From (7.4) and (7.1), we find that

$$|v(i)| \leq M_3 + M_4 R(i) \quad \text{for all } i \in S. \tag{7.5}$$

Moreover, using part (a), for each  $i \in S$  and  $m \geq 1$  we have

$$\alpha V_\alpha^1(i, \pi_m^1, \pi_m^2) = r_1(i, \pi_m^1, \pi_m^2) + \sum_{j \in S} q(j | i, \pi_m^1, \pi_m^2) V_\alpha^1(j, \pi_m^1, \pi_m^2),$$

which can be expressed as

$$\begin{aligned} V_\alpha^1(i, \pi_m^1, \pi_m^2) &= \frac{r_1(i, \pi_m^1, \pi_m^2)}{\alpha - q(i | i, \pi_m^1, \pi_m^2)} \\ &+ \frac{1}{\alpha - q(i | i, \pi_m^1, \pi_m^2)} \sum_{j \neq i} q(j | i, \pi_m^1, \pi_m^2) V_\alpha^1(j, \pi_m^1, \pi_m^2). \end{aligned}$$

Hence, by (7.4), (7.5), Assumption 5.1, and Lemma 8.3.7 of [13], we obtain

$$v(i) = \frac{r_1(i, \pi^1, \pi^2)}{\alpha - q(i | i, \pi^1, \pi^2)} + \frac{1}{\alpha - q(i | i, \pi^1, \pi^2)} \sum_{j \neq i} q(j | i, \pi^1, \pi^2) v(j),$$

that is,

$$\alpha v(i) = r_1(i, \pi^1, \pi^2) + \sum_{j \in S} q(j | i, \pi^1, \pi^2) v(j),$$

which, together with the ‘uniqueness of solutions’ in part (a), implies that

$$\limsup_{n \rightarrow \infty} V_\alpha^1(i, \pi_n^1, \pi_n^2) = v(i) = V_\alpha^1(i, \pi^1, \pi^2) \quad \text{for all } i \in S. \tag{7.6}$$

Similarly, we obtain

$$\liminf_{n \rightarrow \infty} V_\alpha^1(i, \pi_n^1, \pi_n^2) = V_\alpha^1(i, \pi^1, \pi^2) \quad \text{for all } i \in S. \tag{7.7}$$

Thus, from (7.6) and (7.7),

$$\lim_{n \rightarrow \infty} V_\alpha^1(i, \pi_n^1, \pi_n^2) = V_\alpha^1(i, \pi^1, \pi^2) \quad \text{for all } i \in S,$$

and part (e) follows for  $k = 1$ .

**Lemma 7.3.** *Suppose that Assumptions 3.1, 4.1, and 5.1 hold. Then the following statements also hold.*

- (a) *The functions  $r_k(i, \pi^1, \pi^2)$  and  $\sum_{j \in S} q(j \mid i, \pi^1, \pi^2) V_\alpha^k(j, \pi^1, \pi^2)$  are continuous in  $(\pi^1, \pi^2) \in \Pi_1^S \times \Pi_2^S$  for each fixed  $i \in S$  and  $k = 1, 2$ .*
- (b) *If a real-valued function  $f(\pi^1, \pi^2)$  is continuous in  $(\pi^1, \pi^2) \in \Pi_1^S \times \Pi_2^S$ , then the function  $f^*(\pi^1)$  on  $\Pi_1^S$ , defined as*

$$f^*(\pi^1) := \sup_{\pi^2 \in \Pi_2^S} f(\pi^1, \pi^2),$$

*is continuous in  $\pi^1 \in \Pi_1^S$ .*

*Proof.* (a) By Assumptions 5.1.2 and 5.1.3, we see that  $r_k(i, \pi^1, \pi^2)$  and  $q(j \mid i, \pi^1, \pi^2)$  are continuous for each  $i, j \in S$ . Thus, by Assumption 5.1.1, the desired conclusion follows from our Lemma 7.2(e) and Lemma 8.3.7 of [13].

(b) Since  $S$  is denumerable, by Assumption 5.1.1 we know that  $\Pi_2^S = \prod_{i \in S} P(B(i))$  is compact. Thus,  $f^*(\pi^1)$  is well defined for each  $\pi^1 \in \Pi_1^S$ . Similarly, by Assumption 5.1.1 we see that  $\Pi_1^S \times \Pi_2^S$  is also compact. Hence,  $f$  is uniformly continuous on the compact metric space  $\Pi_1^S \times \Pi_2^S$  and, so, part (b) follows from well-known facts.

Finally, we recall Fan’s fixed-point theorem.

**Lemma 7.4.** (Fan’s fixed-point theorem.) *Let  $X$  be a locally convex topological linear space and  $Y$  a compact convex set in  $X$ . Let  $F(Y)$  be the family of all closed convex (nonempty) subsets of  $Y$ . Then, for any upper-semicontinuous point-to-set transformation  $f$  from  $Y$  to  $F(Y)$ , there exists a point  $y_0 \in Y$  such that  $y_0 \in f(y_0)$ .*

*Proof.* See [3].

**Remark 7.1.** It is clear that the compact sets  $\Pi_1^S$  and  $\Pi_2^S$  are convex. Moreover, by (3.1) and (3.2), the function

$$\pi^1 \mapsto r_k(i, \pi^1, \pi^2) + \sum_{j \in S} q(j \mid i, \pi^1, \pi^2) u(j) \tag{7.8}$$

is convex in  $\pi^1 \in \Pi_1^S$  for any fixed  $\pi^2 \in \Pi_2^S$ ,  $u \in B(S)$ ,  $i \in S$ , and  $k = 1, 2$ . (This convexity is also true, of course, if we consider (7.8) as a function of  $\pi^2 \in \Pi_2^S$  for fixed  $\pi^1 \in \Pi_1^S$ .)

### 8. Proof of Theorems 5.1 and 5.2

#### 8.1. Proof of Theorem 5.1

8.1.1. *Proof of part (a).* For each fixed  $\pi^2 \in \Pi_2^S$ , Lemma 7.3 gives the existence of a strategy  $\pi^{*1} \in \Pi_1^S$  such that

$$\begin{aligned} & r_1(i, \pi^{*1}, \pi^2) + \sum_{j \in S} q(j \mid i, \pi^{*1}, \pi^2) V_\alpha^1(j, \pi^{*1}, \pi^2) \\ &= \sup_{\pi^1 \in \Pi_1^S} \left\{ r_1(i, \pi^1, \pi^2) + \sum_{j \in S} q(j \mid i, \pi^1, \pi^2) V_\alpha^1(j, \pi^1, \pi^2) \right\} \quad \text{for all } i \in S. \end{aligned}$$

Thus, the set

$$\begin{aligned}
 f(\pi^2) &:= \left\{ \pi^{*1} \in \Pi_1^s : r_1(i, \pi^{*1}, \pi^2) + \sum_{j \in S} q(j \mid i, \pi^{*1}, \pi^2) V_\alpha^1(j, \pi^{*1}, \pi^2) \right. \\
 &= \left. \sup_{\pi^1 \in \Pi_1^s} \left\{ r_1(i, \pi^1, \pi^2) + \sum_{j \in S} q(j \mid i, \pi^1, \pi^2) V_\alpha^1(j, \pi^1, \pi^2) \right\} \text{ for all } i \in S \right\}
 \end{aligned}
 \tag{8.1}$$

is nonempty. Similarly, given any  $\pi^1 \in \Pi_1^s$ , the set

$$\begin{aligned}
 f(\pi^1) &:= \left\{ \pi^{*2} \in \Pi_2^s : r_2(i, \pi^1, \pi^{*2}) + \sum_{j \in S} q(j \mid i, \pi^1, \pi^{*2}) V_\alpha^2(j, \pi^1, \pi^{*2}) \right. \\
 &= \left. \sup_{\pi^2 \in \Pi_2^s} \left\{ r_2(i, \pi^1, \pi^2) + \sum_{j \in S} q(j \mid i, \pi^1, \pi^2) V_\alpha^2(j, \pi^1, \pi^2) \right\} \text{ for all } i \in S \right\}
 \end{aligned}
 \tag{8.2}$$

is nonempty. Moreover, by Lemma 7.3(b),  $f(\pi^1)$  and  $f(\pi^2)$  are compact sets while, by Lemma 7.2(a)–(c) and Remark 7.1, they are convex. We now define the multifunction

$$\Phi : \Pi_1^s \times \Pi_2^s \rightarrow 2^{\Pi_1^s \times \Pi_2^s} \quad \text{by} \quad (\pi^1, \pi^2) \mapsto f(\pi^2) \times f(\pi^1),$$

and show that it is upper semicontinuous. Let  $\{\pi_n^1, \pi_n^2\}$  and  $\{\pi_n^{*1}, \pi_n^{*2}\}$  be such that, for all  $n \geq 1$ ,

- (i)  $\{\pi_n^1, \pi_n^2\}, \{\pi_n^{*1}, \pi_n^{*2}\} \in \Pi_1^s \times \Pi_2^s$ ,
- (ii)  $\lim_{n \rightarrow \infty} (\pi_n^1, \pi_n^2) = (\pi^1, \pi^2)$  and  $\lim_{n \rightarrow \infty} (\pi_n^{*1}, \pi_n^{*2}) = (\pi^{*1}, \pi^{*2})$ , and
- (iii)  $(\pi_n^{*1}, \pi_n^{*2}) \in \Phi(\pi_n^1, \pi_n^2)$ .

We wish to show that  $(\pi^{*1}, \pi^{*2})$  is in  $\Phi(\pi^1, \pi^2)$ . To this end, note that, by (iii),

$$\begin{aligned}
 &r_1(i, \pi_n^{*1}, \pi_n^2) + \sum_{j \in S} q(j \mid i, \pi_n^{*1}, \pi_n^2) V_\alpha^1(j, \pi_n^{*1}, \pi_n^2) \\
 &= \sup_{\pi^1 \in \Pi_1^s} \left\{ r_1(i, \pi^1, \pi_n^2) + \sum_{j \in S} q(j \mid i, \pi^1, \pi_n^2) V_\alpha^1(j, \pi^1, \pi_n^2) \right\}
 \end{aligned}
 \tag{8.3}$$

for all  $n \geq 1$ . Furthermore, by (ii) and Lemma 7.3, letting  $n \rightarrow \infty$  in (8.3) gives

$$\begin{aligned}
 &r_1(i, \pi^{*1}, \pi^2) + \sum_{j \in S} q(j \mid i, \pi^{*1}, \pi^2) V_\alpha^1(j, \pi^{*1}, \pi^2) \\
 &= \sup_{\pi^1 \in \Pi_1^s} \left\{ r_1(i, \pi^1, \pi^2) + \sum_{j \in S} q(j \mid i, \pi^1, \pi^2) V_\alpha^1(j, \pi^1, \pi^2) \right\},
 \end{aligned}$$

which means that  $\pi^{*1}$  is in  $f(\pi^2)$ . Similarly, we find that  $\pi^{*2} \in f(\pi^1)$  and, so,  $(\pi^{*1}, \pi^{*2}) \in \Phi(\pi^1, \pi^2)$ . Thus, the mapping  $\Phi$  is indeed upper semicontinuous. It follows, from Lemma 7.4,

that there exists a pair  $(\pi^{*1}, \pi^{*2}) \in \Pi_1^s \times \Pi_2^s$  such that  $(\pi^{*1}, \pi^{*2}) \in f(\pi^{*2}) \times f(\pi^{*1})$ , which, together with (8.1), (8.2), and Lemma 7.2(a), implies that

$$\alpha V_\alpha^1(i, \pi^{*1}, \pi^{*2}) = \sup_{\pi^1 \in \Pi_1^s} \left\{ r_1(i, \pi^1, \pi^{*2}) + \sum_{j \in S} q(j \mid i, \pi^1, \pi^{*2}) V_\alpha^1(j, \pi^{*1}, \pi^{*2}) \right\} \text{ for all } i \in S \quad (8.4)$$

and

$$\alpha V_\alpha^2(i, \pi^{*1}, \pi^{*2}) = \sup_{\pi^2 \in \Pi_2^s} \left\{ r_2(i, \pi^{*1}, \pi^2) + \sum_{j \in S} q(j \mid i, \pi^{*1}, \pi^2) V_\alpha^2(j, \pi^{*1}, \pi^{*2}) \right\} \text{ for all } i \in S. \quad (8.5)$$

Hence, part (a) follows.

8.1.2. *Proof of part (b).* We first fix  $\pi^{*2} \in \Pi_2^s$ . By (3.1), (3.2), (7.2), (7.3), and (8.4), we have

$$\begin{aligned} \alpha V_\alpha^1(i, \pi^{*1}, \pi^{*2}) &\geq r_1(i, \phi, \pi^{*2}) + \sum_{j \in S} q(j \mid i, \phi, \pi^{*2}) V_\alpha^1(j, \pi^{*1}, \pi^{*2}) \text{ for all } \phi \in P(A(i)), i \in S. \end{aligned} \quad (8.6)$$

Thus, for each  $\pi^1 = (\pi_t^1, t \geq 0) \in \Pi_1$ , since  $\pi_t^1(\cdot \mid i) \in P(A(i))$  for all  $i \in S$  and  $t \geq 0$ , by (3.1), (3.2), and (8.6) we have

$$\begin{aligned} \alpha V_\alpha^1(i, \pi^{*1}, \pi^{*2}) &\geq r_1(t, i, \pi^1, \pi^{*2}) + \sum_{j \in S} q(j \mid i, t, \pi^1, \pi^{*2}) V_\alpha^1(j, \pi^{*1}, \pi^{*2}) \text{ for all } i \in S, \end{aligned}$$

which, together with parts (a) and (d) of Lemma 7.2, gives

$$V_\alpha^1(s, i, \pi^{*1}, \pi^{*2}) = V_\alpha^1(i, \pi^{*1}, \pi^{*2}) \geq V_\alpha^1(s, i, \pi^1, \pi^{*2}) \text{ for all } (s, i) \in \bar{S} \text{ and } \pi^1 \in \Pi_1.$$

From this we recover (4.2). Similarly, by (3.1), (3.2), (7.2), (7.3), and (8.5), we see that (4.3) is also satisfied. Thus,  $(\pi^{*1}, \pi^{*2})$  is a Nash equilibrium. On the other hand, if  $(\pi^{*1}, \pi^{*2})$  is a Nash equilibrium, then by parts (b) and (c) of Lemma 7.2 we see that (5.2) and (5.3) are satisfied. Thus, we have completed the proof of part (b).

8.1.3. *Proof of part (c).* Suppose that  $r_1 = -r_2$ . By parts (a) and (b) of Lemma 7.2, (5.2), and (5.3) we know that (5.4) is true. Thus, to complete the proof of part (c) it only remains to verify (5.5). The operator  $T$  in (5.1) is obviously monotone. Furthermore, by Assumptions 5.1 and 4.1.2,  $u_0$  and  $u_n = T^n u_0$  are well defined for each  $n \geq 1$ . Thus, under Assumptions 3.1.3 and 5.1.2 straightforward calculations give, for each  $i \in S$ ,

$$\begin{aligned} u_1(i) &\geq -\frac{M_1 + M_2 R(i)}{\alpha + m(i)} - \frac{m(i)}{\alpha + m(i)} \left[ \frac{M_1}{\alpha} + \frac{dM_2}{\alpha(\alpha - c)} + \frac{M_2 R(i)}{\alpha - c} + \frac{M_2(cR(i) + d)}{(\alpha - c)m(i)} \right] \\ &= \frac{M_1}{\alpha} + \frac{dM_2}{\alpha(\alpha - c)} + \frac{M_2 R(i)}{\alpha - c} \\ &= u_0(i). \end{aligned} \quad (8.7)$$

Therefore,

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

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

$$Tu^* \geq u^*. \tag{8.8}$$

On the other hand, to prove that  $u^* \in B(S)$ , it suffices to show that

$$|u_n(i)| \leq \frac{M_1}{\alpha} + \frac{dM_2}{\alpha(\alpha - c)} + \frac{M_2R(i)}{\alpha - c} \quad \text{for all } n \geq 0 \text{ and } i \in S.$$

This, however, follows by induction, by an argument similar to that resulting in (8.7).

Next, we will show that (8.8) holds with equality. Indeed, for each fixed  $n \geq 1$ ,  $i \in S$ , and  $\phi \in P(A(i))$ , we have already proved that  $u_n$  is in  $B(S)$  while, by Assumption 5.1.1,  $P(B(i))$  is compact. Thus, by Lemma 7.3, there exists  $\psi_n^* \in P(B(i))$ , which may depend on  $i$  and  $\phi$ , such that

$$\begin{aligned} u_{n+1}(i) &\geq \inf_{\psi \in P(B(i))} \left\{ \frac{r_1(i, \phi, \psi)}{\alpha + m(i)} + \frac{m(i)}{\alpha + m(i)} \sum_{j \in S} \left[ \frac{q(j | i, \phi, \psi)}{m(i)} + \delta_{ij} \right] u_n(j) \right\} \\ &= \frac{r_1(i, \phi, \psi_n^*)}{\alpha + m(i)} + \frac{m(i)}{\alpha + m(i)} \sum_{j \in S} \left[ \frac{q(j | i, \phi, \psi_n^*)}{m(i)} + \delta_{ij} \right] u_n(j). \end{aligned} \tag{8.9}$$

Since  $P(B(i))$  is compact, without loss of generality we may suppose that  $\psi_n^* \rightarrow \psi^*$  in  $P(B(i))$ . Therefore, as  $u_0 \leq u_n \uparrow u^* \in B(S)$ , by the ‘extended Fatou lemma’, i.e. Lemma 8.3.7(b) of [13], and our Lemma 7.1(b), by letting  $n \rightarrow \infty$  in (8.9) we obtain

$$\begin{aligned} u^*(i) &\geq \frac{r_1(i, \phi, \psi^*)}{\alpha + m(i)} + \frac{m(i)}{\alpha + m(i)} \sum_{j \in S} \left[ \frac{q(j | i, \phi, \psi^*)}{m(i)} + \delta_{ij} \right] u^*(j) \\ &\geq \inf_{\psi \in P(B(i))} \left\{ \frac{r_1(i, \phi, \psi)}{\alpha + m(i)} + \frac{m(i)}{\alpha + m(i)} \sum_{j \in S} \left[ \frac{q(j | i, \phi, \psi)}{m(i)} + \delta_{ij} \right] u^*(j) \right\}. \end{aligned} \tag{8.10}$$

As (8.10) holds for all  $\phi \in P(A(i))$  and  $i \in S$ , we conclude that

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

which, together with (8.8), gives  $u^* = Tu^*$ , i.e.

$$\alpha u^*(i) := \sup_{\phi \in P(A(i))} \inf_{\psi \in P(B(i))} \left\{ r_1(i, \phi, \psi) + \sum_{j \in S} q(j | i, \phi, \psi) u^*(j) \right\} \quad \text{for all } i \in S.$$

This equality and Lemma 7.4 imply the existence of  $(\bar{\pi}^1, \bar{\pi}^2) \in \Pi_1^s \times \Pi_2^s$  such that

$$\begin{aligned} \alpha u^*(i) &= \sup_{\phi \in P(A(i))} \left\{ r_1(i, \phi, \bar{\pi}^2) + \sum_{j \in S} q(j | i, \phi, \bar{\pi}^2) u^*(j) \right\} \\ &= \inf_{\psi \in P(B(i))} \left\{ r_1(i, \bar{\pi}^1, \psi) + \sum_{j \in S} q(j | i, \bar{\pi}^1, \psi) u^*(j) \right\}. \end{aligned} \tag{8.11}$$

If we now we fix  $\bar{\pi}^2$  then, for any strategy  $\pi^1 = (\pi_t^1) \in \Pi_1$ , we have  $\pi_t^1(\cdot | i) \in P(A(i))$  for all  $t \geq 0$  and  $i \in S$ . Hence, from (3.1), (3.2), and (8.11), we find that

$$u^*(i) \geq r_1(t, i, \pi^1, \bar{\pi}^2) + \sum_{j \in S} q(j | t, i, \pi^1, \bar{\pi}^2) u^*(j) \quad \text{for all } (t, i) \in \bar{S}.$$

This inequality and Lemma 6.2(a) of [5] yield

$$u^*(i) \geq V_\alpha^1(s, i, \pi^1, \bar{\pi}^2) \quad \text{for all } \pi^1 \in \Pi_1 \text{ and } (s, i) \in \bar{S},$$

which in turn implies that

$$u^*(i) \geq \sup_{\pi^1 \in \Pi_1} \inf_{\pi^2 \in \Pi_2} V_\alpha^1(s, i, \pi^1, \pi^2) \quad \text{for all } (s, i) \in \bar{S}. \tag{8.12}$$

A similar argument, replacing  $\bar{\pi}^2$  with  $\bar{\pi}^1$ , gives

$$u^*(i) \leq \inf_{\pi^2 \in \Pi_2} \sup_{\pi^1 \in \Pi_1} V_\alpha^1(s, i, \pi^1, \pi^2) \quad \text{for all } (s, i) \in \bar{S}. \tag{8.13}$$

From (8.13), (8.12), and (5.4), we obtain the desired result (5.5) with  $V(\cdot) = u^*(\cdot)$ .

**8.2. Proof of Theorem 5.2**

First, let  $R^d(i) := R(i) + d/c$  for all  $i \in S$ . Since the model (2.1) is conservative (see (2.3)), by Assumption 3.1 and (5.6) we have

$$\begin{aligned} \sum_{j \in S} P(j | i, a, b) R^d(j) &= \frac{1}{1 + \|q\|} \sum_{j \in S} q(j | i, a, b) \left( R(j) + \frac{d}{c} \right) + R^d(i) \\ &\leq \frac{1}{1 + \|q\|} (cR(i) + d) + R^d(i) \\ &= \kappa R^d(i) \quad \text{for all } (i, a, b) \in K, \end{aligned}$$

where  $\kappa := (c + 1 + \|q\|)/(1 + \|q\|)$ . Since  $\alpha > c$ , we have

$$\kappa\beta = \frac{c + 1 + \|q\|}{1 + \|q\|} \frac{1 + \|q\|}{\alpha + 1 + \|q\|} = \frac{c + 1 + \|q\|}{\alpha + 1 + \|q\|} < 1.$$

Then, as in the proof of Theorem 5.1(b), but now using Theorem 8.3.6 of [13], there exists a Nash equilibrium  $(\pi^{*1}, \pi^{*2})$  for the model (5.7). On the other hand, since (5.2) and (5.3) are equivalent to

$$\begin{aligned} &V_\alpha^1(i, \pi^{*1}, \pi^{*2}) \\ &= \sup_{\pi^1 \in \Pi_1^s} \left\{ r_1^d(i, \pi^1, \pi^{*2}) + \beta \sum_{j \in S} P(j | i, \pi^1, \pi^{*2}) V_\alpha^1(j, \pi^{*1}, \pi^{*2}) \right\} \quad \text{for all } i \in S \end{aligned}$$

and

$$\begin{aligned} &V_\alpha^2(i, \pi^{*1}, \pi^{*2}) \\ &= \sup_{\pi^2 \in \Pi_2^s} \left\{ r_2^d(i, \pi^{*1}, \pi^2) + \beta \sum_{j \in S} P(j | i, \pi^{*1}, \pi^2) V_\alpha^2(j, \pi^{*1}, \pi^{*2}) \right\} \quad \text{for all } i \in S, \end{aligned}$$

respectively, by Theorem 5.1(b) and Theorem 8.3.6 of [13] we see that Theorem 5.2 holds.

## 9. Concluding remarks

In the previous sections, we have studied two-person nonzero-sum games for continuous-time Markov chains with discounted payoff criteria. Under reasonably mild assumptions, we have shown that the game has a stationary Nash equilibrium. In particular, the transition and reward rates may be unbounded provided that they satisfy the ‘growth’ conditions in Assumptions 3.1.3 and 4.1.2. For the zero-sum case, we proved the existence of the value of the game, and also provided a *recursive* way to compute it, or at least to approximate it. Despite our assumptions, we believe that our formulation and approach are general enough to provide a way to analyse other important problems, such as games with an average payoff criterion, which is an important optimization criterion in many applications (e.g. telecommunication systems). To the best of the authors’ knowledge, such games have not been previously studied for continuous-time Markov chains. Research on this topic is in progress.

To conclude, we should mention that similar results for  $N$ -person nonzero-sum games can be obtained using our proof techniques, with obvious (notational) changes.

## Acknowledgements

This research was partially supported by CONACYT grant 37355-E. The research of the first author was supported by the Natural Science Foundations of China and Guangdong Province, by EYTP, and by NCET.

## References

- [1] ANDERSON, W. J. (1991). *Continuous-Time Markov Chains*. Springer, New York.
- [2] CHUNG, K. L. (1960). *Markov Chains with Stationary Transition Probabilities*. Springer, Berlin.
- [3] FAN, K. (1953). Minimax theorems. *Proc. Nat. Acad. Sci. USA* **39**, 42–47.
- [4] GHOSH, M. K. AND BAGCHI, A. (1998). Stochastic games with average payoff criterion. *Appl. Math. Optimization* **38**, 238–301.
- [5] GUO, X. P. AND HERNÁNDEZ-LERMA, O. (2003). Continuous-time controlled Markov chains with discounted rewards. *Acta Appl. Math.* **79**, 195–216.
- [6] 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.
- [7] GUO, X. AND LIU, K. (2001). A note on optimality conditions for continuous-time Markov decision processes with average cost criterion. *IEEE Trans. Automatic Control* **46**, 1984–1988.
- [8] GUO, X. P. AND ZHU, W. (2002). Denumerable state continuous time Markov decision processes with unbounded cost and transition rates under average criterion. *ANZIAM J.* **43**, 541–557.
- [9] GUO, X. P. AND ZHU, W. (2002). Optimality conditions for CTMDP with average cost criterion. In *Markov Processes and Controlled Markov Chains*, eds Z. T. Hou, J. A. Filar and A. Y. Chen, Kluwer, Dordrecht, pp. 167–188.
- [10] HAMADÉNE, S. (1999). Nonzero sum linear–quadratic stochastic differential games and backward–forward equations. *Stoch. Anal. Appl.* **17**, 117–130.
- [11] HERNÁNDEZ-LERMA, O. (1994). *Lectures on Continuous-Time Markov Control Processes* (Aportaciones Matemáticas **3**). Sociedad Matematica Mexicana, México.
- [12] HERNÁNDEZ-LERMA, O. AND GOVINDAN, T. E. (2001). Nonstationary continuous-time Markov control processes with discounted costs on infinite horizon. *Acta Appl. Math.* **67**, 277–293.
- [13] HERNÁNDEZ-LERMA, O. AND LASSERRE, J. B. (1999). *Further Topics on Discrete-Time Markov Control Processes*. Springer, New York.
- [14] HIMMELBERG, C. J., PARTHASARATHY, T., RAGHAVAN, T. E. S. AND VAN VLECK, F. S. (1976). Existence of  $p$ -equilibrium and optimal stationary strategies in stochastic games. *Proc. Amer. Math. Soc.* **60**, 245–251.
- [15] HOU, Z. T. AND GUO, X. (1998). *Markov Decision Processes*. Science and Technology Press of Hunan, Changsha (in Chinese).
- [16] HOU, Z. T. *et al.* (1994). *The Q-Matrix Problems on Markov Processes*. Science and Technology Press of Hunan, Changsha (in Chinese).
- [17] KARATZAS, I., SHUBIK, M. AND SUDDERTH, W. D. (1994). Construction of stationary Markov equilibria in a strategic market game. *Math. Operat. Res.* **19**, 975–1006.

- [18] KUSHNER, H. J. (2002). Numerical approximations for stochastic differential games. *SIAM J. Control Optimization* **41**, 457–486.
- [19] LEFÈVRE, C. (1981). Optimal control of a birth and death epidemic process. *Operat. Res.* **29**, 971–982.
- [20] 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.
- [21] PUTERMAN, M. L. (1994). *Markov Decision Processes*. John Wiley, New York.
- [22] SENNOTT, L. I. (1994). Nonzero-sum stochastic games with unbounded cost: discounted and average cost cases. *Math. Meth. Operat. Res.* **40**, 145–162.
- [23] SENNOTT, L. I. (1999). *Stochastic Dynamic Programming and the Control of Queueing Systems*. John Wiley, New York.
- [24] USHIDA, K. (1978). On the existence of Nash equilibrium points in  $N$ -person nonzero sum stochastic differential games. *SIAM J. Control Optimization* **16**, 142–149.
- [25] WALRAND, J. (1988). *An Introduction to Queueing Networks*. Prentice-Hall, Englewood Cliffs, New Jersey.
- [26] YUSHKEVICH, A. A. AND FEINBERG, E. A. (1979). On homogeneous Markov model with continuous time and finite or countable state space. *Theory Prob. Appl.* **24**, 156–161.