

CONTINUOUS-TIME CONTROLLED MARKOV CHAINS¹

BY XIANPING GUO² AND ONÉSIMO HERNÁNDEZ-LERMA

Zhongshan University and CINVESTAV-IPN

This paper concerns studies on continuous-time controlled Markov chains, that is, continuous-time Markov decision processes with a denumerable state space, with respect to the discounted cost criterion. The cost and transition rates are allowed to be *unbounded* and the action set is a *Borel space*. We first study control problems in the class of deterministic stationary policies and give very weak conditions under which the existence of ε -optimal ($\varepsilon \geq 0$) policies is proved using the construction of a *minimum* Q-process. Then we further consider control problems in the class of randomized Markov policies for (1) *regular* and (2) *nonregular* Q-processes. To study case (1), first we present a new necessary and sufficient condition for a nonhomogeneous Q-process to be regular. This regularity condition, together with the *extended generator* of a nonhomogeneous Markov process, is used to prove the existence of ε -optimal stationary policies. Our results for case (1) are illustrated by a Schlögl model with a controlled diffusion. For case (2), we obtain a similar result using Kolmogorov's forward equation for the *minimum* Q-process and we also present an example in which our assumptions are satisfied, but those used in the previous literature fail to hold.

1. Introduction. In this paper we study continuous-time Markov decision processes (CTMDPs), also known as continuous-time controlled Markov chains, with a discounted cost criterion. A key feature of our control model is that the cost and the transition rates can both be *unbounded*, and that the action (or control) set is a Borel space. Moreover, in contrast to continuous-time *jump* Markov decision processes (see Remark 3.1), which can be reduced to discrete-time problems, our processes can be *continuously* controlled, and usual policies such as *switching controls* are included in our consideration. We first study control problems in the class of deterministic stationary policies and give very weak conditions for the existence of ε -optimal ($\varepsilon \geq 0$) stationary policies. Then we consider control problems in the class of randomized Markov policies for two classes of CTMDPs: regular (or nonexplosive) and nonregular. For each of these classes we give conditions for the existence of (deterministic) stationary policies which are ε -optimal in the set of all randomized Markov policies. We also present

Received August 2001; revised May 2002.

¹Research supported in part by CONACYT Grant 37355-E.

²Also supported in part by Natural Science Foundation of China Grant 19901038 and by the Natural Science Foundation of Guangdong Province, China.

AMS 2000 subject classifications. Primary 93E20; secondary 60J27, 90C40.

Key words and phrases. Nonhomogeneous continuous-time Markov chains, controlled Q-processes, unbounded cost and transition rates, discounted criterion, optimal stationary policies.

a new necessary and sufficient condition for a nonhomogeneous Q-process to be regular.

Continuously controlled Markov processes have been studied by many authors [2, 7, 10–12, 14, 17, 18, 22, 27, 30]; for the *jump* case, see, for instance, [3, 8, 13, 15, 19–21, 23, 25, 26, 28, 29]. However, except for [11, 12, 14], all assume that either the cost or the transition rates are bounded and that the action sets are denumerable, or even finite, as in [7, 18, 22, 27]. Further, in [3, 7, 8, 10–15, 17–23, 25–30] conditions are given for the Q-processes [i.e., a nonhomogeneous transition function associated to given transition rate matrices $Q(t)$, $t \geq 0$] to be regular (i.e., unique and honest), whereas in [2] the treatment is restricted to the class of deterministic stationary policies. On the other hand, the common approach in [7, 10–12, 17, 18, 22, 27] to prove the existence of optimal policies is via Kolmogorov's forward equation, which requires assumptions on the interchange of certain integrals and summations. In this paper we use weaker assumptions to prove the existence of optimal policies using the *extended generator* approach instead of Kolmogorov's forward equation.

As was already mentioned, we first study control problems in the class of deterministic stationary policies and give very weak conditions under which the existence of ε -optimal ($\varepsilon \geq 0$) policies is proved using the construction of a *minimum* Q-process (Theorem 3.2). Then we further consider control problems in the class of randomized Markov policies for two classes of CTMDPs: (1) regular, and (2) nonregular. In case (1), we first give a necessary and sufficient condition for a nonhomogeneous Q-process to be regular (see Proposition 2.1). This condition is different from those in [1, 6, 9–12, 17, 16]. Then, based on this regularity condition and using the *extended generator* of a nonhomogeneous Markov process, we prove the existence of ε -optimal stationary policies (Theorem 3.3). This result, which includes all of the main results in [7, 10–12, 17, 18, 22, 27], is illustrated by considering the Schlögl model [5, 24] with a controlled diffusion (Example 4.1). Next we consider the nonregular case (2). In this case, to prove the existence of ε -optimal stationary policies we use Kolmogorov's forward equation for a nonhomogeneous *minimum* Q-process (Theorems 3.4 and 3.5). These results are illustrated with an example (Example 4.2) in which all of our assumptions are satisfied, but a Q-process is *not* unique, and the control set is a nondenumerable Borel space. Therefore, whereas a Q-process is unique in [3, 7, 8, 10–15, 17–23, 25–30] and the action sets in [17] are finite, the conditions used in the previous literature fail to hold in our example.

The rest of this paper is organized as follows. In Section 2 we introduce the control model and the optimal control problem with which we are concerned, as well as a regularity criterion (Proposition 2.1) for a nonhomogeneous Q-process. Our main results on the existence of ε -optimal ($\varepsilon \geq 0$) stationary policies are all stated in Section 3; see Theorems 3.1–3.5. Whereas these results require lengthy preliminaries, their proofs are postponed to Sections 5–8. In Section 4 we present the two examples briefly mentioned in the previous paragraph.

2. The optimal control problem. The optimal control model we are concerned with is of the form $(S, A, \{A(i), i \in S\}, c, Q)$ the elements of which have the following properties.

PROPERTY M₁. The state space S is a denumerable set.

PROPERTY M₂. The action space A is a Borel space, endowed with the Borel σ -algebra $\mathcal{B}(A)$, and for each state $i \in S$, $A(i) \in \mathcal{B}(A)$ stands for the set of feasible actions in i . Let

$$K := \{(i, a) \mid i \in S, a \in A(i)\}.$$

PROPERTY M₃. The real-valued function $c : K \rightarrow R$ denotes the cost rate. For each $i \in S$, $c(i, a)$ is assumed to be measurable in $a \in A(i)$.

PROPERTY M₄. $Q = [q(j|i, a)]$ is a Q-matrix with $q(j|i, a) \geq 0$ for all $(i, a) \in K$ and $i \neq j$, which is supposed to be *conservative*, that is,

$$\sum_{j \in S} q(j|i, a) = 0 \quad \forall (i, a) \in K,$$

and *stable*, that is,

$$m(i) := \sup_{a \in A(i)} q_i(a) < \infty \quad \forall i \in S,$$

where $q_i(a) := -q(i|i, a) = \sum_{j \neq i} q(j|i, a)$ for all (i, a) in K . In addition, $q(j|i, \cdot)$ is measurable in $a \in A(i)$ for each fixed $i, j \in S$.

To introduce the optimal control problem we are interested in, we first introduce the class of admissible policies.

DEFINITION 2.1. Let F be the family of functions $f : S \rightarrow A$ such that $f(i) \in A(i)$ for all $i \in S$ and let Φ be the family of functions $\pi_t(B|i)$ from $S \times \mathcal{B}(A) \times [0, \infty)$ to $[0, 1]$ such that:

1. for each $i \in S$ and $t \geq 0$, $\pi_t(\cdot|i)$ is a probability measure on $\mathcal{B}(A)$ with $\pi_t(A(i)|i) = 1$;
2. for each $i \in S$ and $B \in \mathcal{B}(A)$, $t \mapsto \pi_t(B|i)$ is a Lebesgue measurable function on $[0, \infty)$.

Then a family $\pi = \{\pi_t, t \geq 0\}$ in Φ is called a *randomized Markov policy*. If in addition there is a function $f \in F$ such that $\pi_t(\cdot|i)$ is concentrated at $f(i)$ for all $i \in S$ and $t \geq 0$, then π is said to be a (deterministic) *stationary policy* or a *switching control* because it prescribes a control only at times when the system changes its state. In the latter case π will be identified with f , and so F will be

regarded as the family of stationary policies. The class of randomized Markov policies is denoted by Π_m .

For each policy $\pi = (\pi_t, t \geq 0) \in \Pi_m$, the associated transition and cost rates are defined, respectively, as

$$(2.1) \quad q_{ij}(t, \pi) := \int_A q(j|i, a) \pi_t(da|i) \quad \text{for } i, j \in S, t \geq 0,$$

$$(2.2) \quad c(t, i, \pi) := \int_A c(i, a) \pi_t(da|i) \quad \text{for } i \in S, t \geq 0.$$

In particular, when $\pi = f \in F$, we write $q_{ij}(t, \pi)$ and $c(t, i, \pi)$ as $q(j|i, f(i))$ and $c(i, f(i))$, respectively.

Let $Q(t, \pi) := [q_{ij}(t, \pi)]$ be a transition rate matrix. Any transition function $P(\pi) := \{\tilde{p}(s, i, t, j, \pi)\}$, possibly substochastic, with transition rates $q_{ij}(t, \pi)$ is called a *Q-process*. To guarantee the existence of such processes we will restrict ourselves to control policies in the class Π defined as

$$\Pi := \{\pi \in \Pi_m : q_{ij}(t, \pi) \text{ is continuous in } t \text{ for each fixed } i, j \in S\}.$$

Observe that Π contains F and, on the other hand, by Property M₄, $Q(t, \pi)$ is conservative and stable. Hence, for each $\pi \in \Pi$, the existence of a Q-process is indeed guaranteed, but, as is well known [1, 6, 9, 16], it is not necessarily unique. We shall denote by $Q(\pi)$ the set of all possible Q-processes associated to $\pi \in \Pi$ and denote by

$$P^{\min}(s, t, \pi) := (p^{\min}(s, i, t, j, \pi), i, j \in S) \quad \text{for } t \geq s \geq 0$$

the *minimum* Q-process in $Q(\pi)$, where $p^{\min}(s, i, t, j, \pi)$ is the transition probability from state i at time s to state j at time t and satisfies that for all $i, j \in S$ and $t \geq s \geq 0$, $p^{\min}(s, i, t, j, \pi) \leq \tilde{p}(s, i, t, j, \pi)$ for any Q-process $\{\tilde{p}(s, i, t, j, \pi)\}$. For future reference it is convenient to recall that the minimum process can be constructed as follows [1, 6, 9, 16]. For $t \geq s \geq 0$, $n \geq 0$ and $u \geq 0$, let

$$(2.3) \quad D(u, \pi) := \text{diag}(q_i(u, \pi), i \in S) \quad \text{with } q_i(u, \pi) := -q_{ii}(u, \pi),$$

$$(2.4) \quad P_0(s, t, \pi) := \Delta(s, t, \pi) = \text{diag}\left(\exp\left(-\int_s^t q_i(u, \pi) du\right), i \in S\right),$$

$$(2.5) \quad P_{n+1}(s, t, \pi) := \int_s^t \Delta(s, u, \pi) (Q(u, \pi) + D(u, \pi)) P_n(u, t, \pi) du.$$

Then

$$(2.6) \quad P^{\min}(s, t, \pi) = \sum_{n=0}^{\infty} P_n(s, t, \pi).$$

In the remainder of the paper, a real-valued function on S is regarded as a column vector, and operations on matrices and vectors are all componentwise. Moreover, I is the identity matrix, and $\mathbf{1}$ and $\mathbf{0}$ are the vectors with all of its components 1 and 0, respectively.

As mentioned in the Introduction, there are many *sufficient* conditions for a Q-process to be *regular*, that is, $P^{\min}(s, t, \pi)\mathbf{1} = \mathbf{1}$ for each $t \geq s \geq 0$. Here we obtain the following general results.

PROPOSITION 2.1. *For any policy $\pi \in \Pi$:*

- (a) $P^{\min}(s, t, \pi) = \int_s^t P^{\min}(s, v, \pi)Q(v, \pi)dv + I \forall t \geq s \geq 0$;
- (b) *the corresponding Q-process is regular if and only if*

$$(2.7) \quad \left[\int_s^t P^{\min}(s, v, \pi)Q(v, \pi)dv \right] \mathbf{1} = \mathbf{0} \quad \forall t \geq s \geq 0;$$

- (c) *if $\int_s^t P^{\min}(s, v, \pi)q(v, \pi)dv < \infty$ for all $t \geq s \geq 0$, then (2.7) holds, where $q(v, \pi)$ is the vector with components $q(v, \pi)(i) := q_i(v, \pi)$ for all $i \in S$.*

PROOF. See Section 8. \square

Note that Proposition 2.1 gives conditions for a nonhomogeneous Q-process to be regular. These conditions are different from those in [1, 6, 9–12, 17, 16].

Now we define the *discounted cost criterion* V_α , where $\alpha > 0$ is a given discount factor.

For each $\pi \in \Pi$, $s \geq 0$, $i \in S$, and each $P(\pi) := \{\tilde{p}(s, i, t, j, \pi)\}$ in $Q(\pi)$, let

$$(2.8) \quad V_\alpha(P(\pi), s, i) := \int_s^\infty e^{-\alpha(t-s)} \sum_{j \in S} \tilde{p}(s, i, t, j, \pi) c(t, j, \pi) dt.$$

Then, for a given family $\bar{\Pi} \subset \Pi$, $i \in S$ and $s \geq 0$, the *minimum value function* is defined as

$$V_{*\bar{\Pi}}(s, i) := \inf_{\substack{P(\pi) \in Q(\pi) \\ \pi \in \bar{\Pi}}} V_\alpha(P(\pi), s, i).$$

[Observe that in this definition it does not suffice to take the infimum over all $\pi \in \bar{\Pi}$ because the discounted costs $V_\alpha(P^1(\pi), \cdot, \cdot)$ and $V_\alpha(P^2(\pi), \cdot, \cdot)$ may differ for two different Q-processes $P^1(\pi)$ and $P^2(\pi)$ associated to the same policy π .]

DEFINITION 2.2. For each $\varepsilon \geq 0$, a policy $\pi^* \in \bar{\Pi}$ is called ε -optimal in $\bar{\Pi}$ if there exists a Q-process $P(\pi^*) \in Q(\pi^*)$ such that $V_\alpha(P(\pi^*), s, i) \leq V_{*\bar{\Pi}}(s, i) + \varepsilon$ for all $i \in S$ and $s \geq 0$. A 0-optimal policy in $\bar{\Pi}$ is said to be optimal in $\bar{\Pi}$.

For each $\pi \in \Pi$, $i \in S$ and $s \geq 0$, let

$$(2.9) \quad V_\alpha(\pi, s, i) := V_\alpha(P^{\min}(\pi), s, i),$$

$$(2.10) \quad V_\alpha^*(s, i) := \inf_{f \in F} V_\alpha(f, s, i),$$

$$(2.11) \quad V_\alpha^*(i) := V_\alpha^*(0, i).$$

For a stationary policy $f \in F$, the associated minimum Q-process is homogeneous [see (2.3)–(2.6)] and so we have $V_\alpha(f, i) := V_\alpha(f, 0, i) = V_\alpha(f, s, i)$ for all $s \geq 0$ and $i \in S$. Therefore, by (2.9)–(2.11),

$$(2.12) \quad V_\alpha^*(0, i) \equiv V_\alpha^*(s, i) \geq V_{*\Pi}(s, i) \quad \forall i \in S \text{ and } s \geq 0.$$

In Theorems 3.3–3.5, we show that in fact the equality holds in (2.12). To do so, we will use several combinations of the following conditions.

ASSUMPTION I. *For each $\pi \in \Pi$, (2.7) holds.*

ASSUMPTION A. *$c(i, a)$ is nonnegative for each $i \in S$ and $a \in A(i)$.*

REMARK 2.1. (a) Obviously, under Assumption A, by (2.9) and the definition of the minimum Q-processes we have $\inf_{\pi \in \Pi} V_\alpha(\pi, s, i) \leq V_\alpha(P(\pi), s, i)$ for all $i \in S$, $s \geq 0$, $P(\pi) \in Q(\pi)$ and $\pi \in \Pi$. Hence, from now on we can restrict ourselves to the class of the minimum Q-processes $\{P^{\min}(s, t, \pi), \pi \in \Pi\}$.

(b) If the Q-process associated to a given policy is regular, then, without loss of generality, in (2.8) we may replace the cost rate c with $c + k$ for any constant k . Therefore, under Assumption I, the condition “ $c \geq 0$ ” in Assumption A can be weakened to “ c is bounded below.”

ASSUMPTION B. *There exists $\hat{f} \in F$ such that, for all $\pi \in \Pi$, $i \in S$ and $t \geq s \geq 0$:*

$$(i) \quad \sum_{j \in S} p^{\min}(s, i, t, j, \pi) V_\alpha(\hat{f}, j) < \infty;$$

$$(ii) \quad \lim_{t \rightarrow \infty} e^{-\alpha t} \sum_{j \in S} p^{\min}(s, i, t, j, \pi) V_\alpha(\hat{f}, j) = 0.$$

REMARK 2.2. (a) Assumption B implies that for all $\pi \in \Pi$, $i \in S$ and $t \geq s \geq 0$,

$$\sum_{j \in S} p^{\min}(s, i, t, j, \pi) V_\alpha^*(j) < \infty$$

and

$$\lim_{t \rightarrow \infty} e^{-\alpha t} \sum_{j \in S} p^{\min}(s, i, t, j, \pi) V_\alpha^*(j) = 0.$$

(b) By Remark 2.1(a) and Definition 2.2, if $V_\alpha(f, i) = \infty$, then such a policy f is *not* optimal. Thus, without loss of generality, from now on we will further assume that $V_\alpha(f, i) < \infty$ for each $f \in F$ and $i \in S$.

(c) It is trivially verified that each of the sets of hypotheses in [7, 10–12, 18, 22, 27] ensures that our Assumptions A, B and I hold. Moreover, we do not require the assumptions on the interchange of integrals and summations used in [7, 10–12, 18, 22, 27, 30]. Finally, observe that Assumption B holds if either c is bounded or the hypotheses in Lemma 4.1 are satisfied.

ASSUMPTION C. For each $i, j \in S$, (i) $A(i)$ is compact and (ii) the functions $c(i, a)$, $q(j|i, a)$ and $\sum_{j \in S} q(j|i, a)V_\alpha(\tilde{f}, j)$ are continuous in $a \in A(i)$ for some $\tilde{f} \in F$.

REMARK 2.3. (a) By Lemmas A.2 and A.3 in [4], Assumption C implies that for each $i \in S$, $\sum_{j \in S} q(j|i, a)V_\alpha^*(j)$ is continuous in $a \in A(i)$.

(b) Assumption C is similar to the compactness–continuity conditions used for discrete-time Markov decision processes; see [3, 15, 17, 23, 25].

3. Existence of optimal policies. In this section we only state our main results; the proofs are presented in later sections (as indicated below). The first of these results gives two equivalent forms, (3.1) and (3.2), of the so-called (α -discounted cost) *optimality equation*.

THEOREM 3.1. *The two optimality equations*

$$(3.1) \quad \alpha u(i) = \inf_{a \in A(i)} \left\{ c(i, a) + \sum_{j \in S} q(j|i, a)u(j) \right\}, \quad i \in S,$$

$$(3.2) \quad u(i) = \inf_{a \in A(i)} \left\{ \frac{c(i, a)}{\alpha + q_i(a)} + \frac{1}{\alpha + q_i(a)} \sum_{j \neq i} q(j|i, a)u(j) \right\}, \quad i \in S,$$

are equivalent, in the sense that if a real-valued function u on S satisfies one of them, then it satisfies the other.

PROOF. See Section 5. \square

REMARK 3.1. The optimality equation (3.2) for our control model is different from that for jump Markov decision processes (see, e.g., [3, 23]) because of the denominator $\alpha + q_i(a)$.

THEOREM 3.2. (a) *Suppose that $\inf_{a \in A(i)} c(i, a) =: c(i) > 0$ for all $i \in S$. Then:*

- (a₁) V_α^* satisfies the optimality equation (3.1) and
 (a₂) for each $\varepsilon > 0$ there exists a stationary policy which is ε -optimal in F .

(b) If Assumptions A and C hold, then there exists a stationary policy which is optimal in F .

PROOF. See Section 5. \square

REMARK 3.2. (a) Theorem 3.2 gives, in particular, conditions for the existence of stationary policies optimal in $F \subset \Pi$. Examples for which all hypotheses in Theorem 3.2 are satisfied are easily given.

(b) If, in addition, Assumption I holds, then the conditions “ $c(i) > 0$ for all $i \in S$ ” in Theorem 3.2(a) can be weakened to “ $c(i, a)$ is bounded below;” see Remark 2.1(b).

The following result gives conditions for optimality in all of Π and for the equality in (2.12).

THEOREM 3.3. *Suppose that Assumptions I, A and B hold.*

(a) *If $\sum_{j \in S} q(j|i, a)V_\alpha^*(j)$ is bounded in $a \in A(i)$ for each $i \in S$, then:*

- (a₁) $V_\alpha^*(i) = V_\alpha^*(s, i) = V_{*\Pi}(s, i)$ for all $i \in S$ and $s \geq 0$, and
 (a₂) for each $\varepsilon > 0$, there exists a stationary policy which is ε -optimal in Π .

(b) *If in addition Assumption C holds, then so does (a₁) and, moreover, there exists a stationary policy which is optimal in Π .*

PROOF. See Section 6. \square

REMARK 3.3. Theorem 3.3 is similar to the one obtained for discrete-time Markov decision processes (see, e.g., Theorem 7.3.6 in Puterman [23]). However, we need here the additional Assumption B because without it we can only obtain $V_\alpha^* \geq TV_\alpha^*$ [with T as in (5.13)] instead of $V_{*\Pi} \geq TV_{*\Pi}$. The latter inequality is obvious in the discrete-time case.

The proof (in Section 6) of Theorem 3.3 uses the *extended generator* of a nonhomogeneous continuous-time Markov process. However, in the following results we will eliminate Assumption I, so that the Q-processes may be nonregular (i.e., we might have $P^{\min}(s, t, \pi)\mathbf{1} \neq \mathbf{1}$; see, e.g., [1, 6]), and the extended generator approach is no longer applicable. Hence, we now replace Assumption I with the following.

ASSUMPTION I'. *There exists a policy $\hat{g} \in F$ such that for all $i \in S, t > s \geq 0$ and $\pi \in \Pi$,*

$$(3.3) \quad \int_s^t \sum_{k \in S} P^{\min}(s, i, u, k, \pi) \sum_{j \in S} |q_{kj}(u, \pi)| V_\alpha(\hat{g}, j) du < \infty.$$

REMARK 3.4. (a) Each of the sets of hypotheses in [7, 10–12, 17, 18, 22, 27] implies Assumption I'.

(b) Examples are given in [11, 12] for which Assumption I' holds, and in which the transition and cost rates are both *unbounded*.

THEOREM 3.4. *If Assumptions I', A, B and C hold, then:*

- (a) V_α^* satisfies the optimality equation (3.1),
- (b) $V_\alpha^*(i) = V_{*\Pi}(s, i)$ for all $s \geq 0$ and $i \in S$ and
- (c) there exists a stationary policy which is optimal in Π .

PROOF. See Section 7. \square

Without Assumption C, Theorem 3.4 becomes as follows.

THEOREM 3.5. *If Assumptions I', A and B hold, and $c(i) = \inf_{a \in A(i)} c(i, a) > 0$ for all $i \in S$, then:*

- (a) V_α^* satisfies the optimality equation (3.1),
- (b) $V_\alpha^*(i) = V_{*\Pi}(s, i)$ for all $s \geq 0$ and $i \in S$ and
- (c) for each $\varepsilon > 0$, there exists an stationary policy which is ε -optimal in Π .

PROOF. See Section 7. \square

4. Examples. In this section we give two examples to illustrate our results presented in Section 3.

EXAMPLE 4.1 (A controlled Schlögl model). The Schlögl model [5, 24] is a model of a chemical reaction with diffusion in a container, and it is a typical model of nonequilibrium systems. Here we are interested in the following optimal control problem.

The container is supposed to consist of a finite number of small vessels numbered $1, 2, \dots, |E|$. Let $E := \{1, 2, \dots, |E|\}$. The states of the system are vectors $i = (i(u), u \in E)$, where $i(u) \geq 0$ is the number of particles in vessel u . Thus, the state space is $S := Z_+^{|E|}$. In each vessel $u \in E$, there is a reaction described by a birth–death process, whose birth and death rates are given, respectively, by the first two lines in (4.2), where the λ 's are given positive constants, and $e_u \in S$ is the unit vector with components δ_{uk} , the Kronecker delta, for $k = 1, \dots, |E|$. On the other hand, between any two vessels, say u and v , there is a diffusion with rate $p(u, v)$. Here we interpret the transition probability matrix $a := [p(u, v) : u, v \in E]$ as the *control parameter*, which takes value in the *action space* $A(i) \equiv A$ for all $i \in S$, that is, A consists of all the transition matrices $a = [p(u, v) : u, v \in E]$. Moreover, when using the diffusion rate $p(u, v)$, the

decision maker incurs a cost $\hat{c}(u, v) \geq 0$, so that the total cost when the state is $i \in S$ and the control action is $a \in A$ turns out to be

$$c(i, a) := \sum_{u, v \in E} i(u) p(u, v) \hat{c}(u, v).$$

To summarize, let

$$(4.1) \quad b(n, k) := \begin{cases} n! [k!(n-k)!]^{-1}, & \text{if } n \geq k, \\ 0, & \text{otherwise,} \end{cases}$$

be the binomial coefficient. Then our Schlögl model with a controlled diffusion can be expressed as (S, A, c, Q) , with S, A and c as in the previous paragraph, and Q -matrix $[q(j|i, a)]$ defined as follows: for $j \neq i$ and $a = [p(u, v)]$ in A ,

$$(4.2) \quad q(j|i, a) := \begin{cases} \lambda_1 b(i(u), 2) + \lambda_3, & \text{if } j = i + e_u, \\ \lambda_2 b(i(u), 3) + \lambda_4 i(u), & \text{if } j = i - e_u, \\ i(u) p(u, v), & \text{if } j = i - e_u + e_v, \\ 0, & \text{otherwise,} \end{cases}$$

and $q_i(a) := -q(i|i, a) = \sum_{j \neq i} q(j|i, a)$.

To prove the existence of an optimal policy π^* we will use the following result from [12].

LEMMA 4.1 ([12], Lemma 2). *Consider the general control model in Section 2 and suppose that there exist nonnegative functions w_1, w_2, \dots, w_N on S such that for each $i \in S$ and $a \in A(i)$, $W(i) := w_1(i) + \dots + w_N(i) \geq 1$ and, moreover,*

$$(4.3) \quad \sum_{j \in S} q(j|i, a) w_n(j) \leq w_{n+1}(i) \quad \text{for } n = 1, \dots, N-1,$$

$$(4.4) \quad \sum_{j \in S} q(j|i, a) w_N(j) \leq 0.$$

Then, for any $\pi \in \Pi$, $0 \leq s \leq t < \infty$ and $i \in S$,

$$(4.5) \quad P^{\min}(s, t, \pi) W \leq \sum_{k=1}^N \frac{1}{(k-1)!} (t-s)^{(k-1)} W,$$

$$(4.6) \quad \int_s^\infty e^{-\alpha(t-s)} P^{\min}(s, t, \pi) W dt \leq \left(\sum_{k=1}^N \alpha^{-k} \right) W.$$

REMARK 4.1. Lemma 4.1 may be used to verify Assumption B and the finiteness of $V_\alpha(f, \cdot)$ if $|c(i, a)|C \leq W(i)$ for all $i \in S$ and $a \in A(i)$ and some constant C .

PROPOSITION 4.1. *For the Schlögl model with a controlled diffusion, Assumptions I, A, B and C hold. Therefore (by Theorem 3.3), there exists a stationary policy optimal in Π .*

PROOF. To verify Assumption I we will use Lemma 4.1 with $N = 2$, $w_1(i) := \sum_{u \in E} i(u)^3$, and $w_2(\cdot) \equiv$ a constant K specified below. Thus, note that for all $i \in S$ and $a \in A$,

$$\begin{aligned} & \sum_{j \in S} q(j|i, a)w_1(j) \\ & \leq \sum_{u \in E} [\lambda_1 b(i(u), 2) + \lambda_4][3i(u)^2 + 3i(u) + 1] \\ & \quad + \sum_{u \in E} [\lambda_2 b(i(u), 3) + \lambda_3][-3i(u)^2 + 3i(u) - 1] + 4|E| \sum_{u \in E} (i(u)^4 + i(u)^2) \\ & =: \sum_{u \in E} [-3\lambda_2 i(u)^5 + \beta_1(u)i(u)^4 + \dots + \beta_4(u)i(u) + \beta_5(u)], \end{aligned}$$

where $i(u) \geq 3$ for each $u \in E$ and the $\beta_k(u)$ are fixed constants independent of a and i . Since the set $\{i \in S : i(u) \leq 2 \text{ for some } u \in E\}$ is finite because of the finiteness of the set E and $\sum_{u \in E} [-3\lambda_2 i(u)^5 + \beta_1(u)i(u)^4 + \dots + \beta_4(u)i(u) + \beta_5(u)] < 0$ when $i(u)$ is sufficiently large, straightforward calculations yield a constant $K > 0$ such that

$$\sum_{j \in S} q(j|i, a)w_1(j) \leq K \quad \forall i \in S \text{ and } a \in A.$$

Let $w_2 \equiv K$. As the Q-matrix is conservative, $\sum_{j \in S} q(j|i, a)w_2(j) = 0$ for all $i \in S$ and $a \in A$. Hence, by Lemma 4.1, for any $\pi \in \Pi$, $0 \leq s \leq t < \infty$ and $i \in S$ we have

$$P^{\min}(s, t, \pi)(w_1 + w_2) \leq (1 + (t - s))(w_1 + w_2),$$

which together with (4.2) and (4.1) gives

$$\int_s^t P^{\min}(s, v, \pi)m \, dv \leq 5(\lambda_1 + \dots + \lambda_4) \int_s^t P^{\min}(s, v, \pi)(w_1 + w_2) \, dv < \infty,$$

where $m(i)$ is the function in Property M₄ (see Section 2). Thus, Assumption I follows from Proposition 2.1(c).

On the other hand, since E is finite, $|c| := \sum_{u, v \in E} \hat{c}(u, v)$ is finite and so is $c(i, a) \leq |c|w_1(i)$. This inequality, together with Lemma 4.1, yields Assumptions A and B. Finally, using again that E is finite, the action space A is compact. Thus, Assumption C is obviously valid. \square

EXAMPLE 4.2. Consider a controlled pure-birth process with state space $S = \{0, 1, \dots\}$, and action sets $A(i) \equiv A$, a nonempty compact metric space. The cost rate is given by $c(i, a) := (1/2^{i+1})r(i, a)$, where $r(i, a)$ is a nonnegative, continuous and bounded function. For any $i \in S$ and $a \in A(i)$, let

$$q(i+1|i, a) := -q(i|i, a) := 2^i h(a) \quad \text{and} \quad q(j|i, a) := 0 \quad \text{if } j \neq i \text{ or } i+1,$$

where h is a continuous function on A and $h(a) > 0$ for all $a \in A$.

For Example 4.2 we can derive the following conclusions.

PROPOSITION 4.2. *For any $\pi \in \Pi$ and $0 \leq s \leq t < \infty$, if $i > j$, then $P^{\min}(s, i, t, j, \pi) = 0$.*

PROOF. For given $\pi \in \Pi$, $0 \leq s \leq t < \infty$ and $n \geq 0$, let $P_n(s, t, \pi) := \{p_{i,j}^n(s, t, \pi) : i, j \in S\}$ be as in (2.4) and (2.5). By the construction of $P^{\min}(s, i, t, j, \pi)$ [see (2.4)–(2.6)], it suffices to prove that

$$(4.7) \quad p_{i,j}^n(s, t, \pi) = 0 \quad \forall n \geq 0, \text{ when } i > j.$$

We will prove (4.7) by induction. From (2.4)–(2.6), it is obvious that (4.7) is valid for $n = 0$. Suppose now that (4.7) holds for some $n \geq 0$. For $i > j$, from (2.5) and the induction hypothesis we get

$$\begin{aligned} p_{i,j}^{n+1}(s, t, \pi) &= \int_s^t \sum_{k \neq i} \exp\left(-\int_s^u q_i(v, \pi) dv\right) q_{ik}(u) p_{k,j}^n(u, t, \pi) du \\ &= \int_s^t \exp\left(-\int_s^u q_i(v, \pi) dv\right) q_{i,i+1}(u, \pi) p_{i+1,j}^n(u, t, \pi) du = 0, \end{aligned}$$

which means that (4.7) holds for $n + 1$, and so is for all $n \geq 0$. \square

PROPOSITION 4.3. *Assumptions Y, A, B and C hold and, therefore, Theorem 3.4 is applicable to Example 4.2.*

PROOF. Obviously, Assumptions A, B and C hold. To verify (3.3) first note that taking $\hat{g} \in F$ and $i \in S$, by (2.8), (2.9) and Proposition 4.2 we have

$$\begin{aligned} (4.8) \quad V_\alpha(\hat{g}, i) &= V_\alpha(\hat{g}, 0, i) = \int_0^\infty \sum_{j \in S} e^{-\alpha t} p^{\min}(0, i, t, j, \hat{g}) c(j, \hat{g}) dt \\ &\leq \int_0^\infty \sum_{j \geq i} e^{-\alpha t} p^{\min}(0, i, t, j, \hat{g}) \frac{1}{2^{j+1}} \|r\| dt \\ &\leq \int_0^\infty \sum_{j \geq i} e^{-\alpha t} \frac{1}{2^{j+1}} \|r\| dt = \frac{\|r\|}{\alpha 2^i} < \infty. \end{aligned}$$

Now let $\|h\| := \max\{h(a) : a \in A\}$. Then, for any $\pi \in \Pi$, $0 \leq s < t$ and $i \in S$, from (4.8) we have that

$$\begin{aligned}
 & \int_s^t \sum_{k \in S} p^{\min}(s, i, u, k, \pi) \sum_{j \in S} |q_{kj}(u, \pi)| V_\alpha(\hat{g}, j) du \\
 & \leq \int_s^t \sum_{k \in S} p^{\min}(s, i, u, k, \pi) 2^{k+1} \frac{2\|h\|\|r\|}{\alpha 2^k} du \\
 (4.9) \quad & \leq \int_s^t \sum_{k \in S} p^{\min}(s, i, u, k, \pi) \frac{4\|h\|\|r\|}{\alpha} du \\
 & \leq \frac{4\|h\|\|r\|}{\alpha} (t - s) < \infty.
 \end{aligned}$$

Hence Assumption I' holds. \square

PROPOSITION 4.4. *For any $f \in F$, a Q-process associated to $Q(f)$ is not unique.*

PROOF. Note that $Q(t, f) \equiv Q(f)$ is a pure-birth Q-matrix and that

$$(4.10) \quad \sum_{i=1}^{\infty} \frac{1}{q_i(f)} \leq \sum_{i=1}^{\infty} \frac{1}{2^i \min\{h(a) : a \in A\}} < \infty.$$

By (4.10) and Lemma 2.1.1 in [2] (or Theorem 12.8.1 in [16]) we obtain that a Q-process corresponding to $Q(f)$ is not unique and so it is not regular. \square

REMARK 4.2. Propositions 4.3 and 4.4 show that for Example 4.2 our assumptions are satisfied, whereas the conditions in [3, 7, 8, 10–15, 17–23, 25–30] fail to hold because the hypotheses in these references except [17] guarantee the regularity of a Q-process, whereas in [17] the feasible action sets are finite.

5. Proof of Theorems 3.1 and 3.2. The proof of Theorem 3.2 is based on the following lemma.

LEMMA 5.1. *If Assumption A holds, then for each $f \in F$:*

(a) $V_\alpha(f, i)$ is the minimal nonnegative solution of the equation

$$(5.1) \quad \alpha u(i) = c(i, f(i)) + \sum_{j \in S} q(j|i, f(i))u(j), \quad i \in S;$$

(b) if a nonnegative real-valued function u on S satisfies

$$\alpha u(i) \geq c(i, f(i)) + \sum_{j \in S} q(j|i, f(i))u(j) \quad \forall i \in S,$$

then $u \geq V_\alpha(f)$.

PROOF. Choose an arbitrary $f \in F$. By Property M_4 , $q_i(f) < \infty$. Thus we see that (5.1) and

$$(5.2) \quad u(i) = \frac{c(i, f(i))}{q_i(f) + \alpha} + \frac{1}{q_i(f) + \alpha} \sum_{j \neq i} q(j|i, f(i))u(j), \quad i \in S,$$

where $q_i(f) := -q(i|i, f(i))$, are equivalent in the sense that they have the same real-valued solutions $u(\cdot)$.

Therefore, we may consider (5.2) instead of (5.1). Let δ_{ij} be the Kronecker delta and let

$$(5.3) \quad \phi_{ij}^{(n+1)}(f) := \begin{cases} \frac{\delta_{ij}}{\alpha + q_i(f)}, & \text{for } n = 0, \\ \frac{1}{\alpha + q_i(f)} \left[\delta_{ij} + \sum_{k \neq i} q(k|i, f(i))\phi_{kj}^{(n)}(f) \right], & \text{for } n \geq 1. \end{cases}$$

Note that $\phi_{ij}^{(n)}(f)$ is nondecreasing in n . As $c(i, a)$ is nonnegative, by Assumption A, from the theory of continuous-time Markov processes (e.g., see pages 121–122 in [1]) and the Fubini theorem we have

$$(5.4) \quad \begin{aligned} V_\alpha(f, i) &= \sum_{j \in S} \int_0^\infty e^{-\alpha t} p^{\min}(0, i, t, j, f) c(j, f(j)) dt \\ &= \sum_{j \in S} \left[\lim_{n \rightarrow \infty} \phi_{ij}^{(n)}(f) \right] c(j, f(j)) \geq 0. \end{aligned}$$

By (5.4) and the monotone convergence theorem, we have

$$(5.5) \quad V_\alpha(f, i) = \lim_{n \rightarrow \infty} \sum_{j \in S} \phi_{ij}^{(n)}(f) c(j, f(j)).$$

For any $n \geq 1$, from (5.3) we can derive that

$$(5.6) \quad \begin{aligned} &\sum_{j \in S} \phi_{ij}^{(n+1)}(f) c(j, f(j)) \\ &= \sum_{j \in S} \frac{1}{\alpha + q_i(f)} \left[\delta_{ij} + \sum_{k \neq i} q(k|i, f(i))\phi_{kj}^{(n)}(f) \right] c(j, f(j)) \\ &= \frac{c(i, f(i))}{\alpha + q_i(f)} + \frac{1}{\alpha + q_i(f)} \sum_{k \neq i} q(k|i, f(i)) \left[\sum_{j \in S} \phi_{kj}^{(n)}(f) c(j, f(j)) \right], \end{aligned}$$

and letting $n \rightarrow \infty$ in (5.6), from the monotone convergence theorem we obtain

$$V_\alpha(f, i) = \frac{c(i, f(i))}{\alpha + q_i(f)} + \frac{1}{\alpha + q_i(f)} \sum_{j \neq i} q(j|i, f(i)) V_\alpha(f, j).$$

Thus $\{V_\alpha(f, i), i \in S\}$ is a nonnegative solution to (5.2).

Now, let u be a nonnegative solution to (5.2). To prove that $V_\alpha(f, i) \leq u(i)$ for all $i \in S$, by (5.5) it is sufficient to show that

$$(5.7) \quad \sum_{j \in S} \phi_{ij}^{(n)}(f) c(j, f(j)) \leq u(i) \quad \forall i, j \in S \text{ and } n \geq 1.$$

This is true for $n = 1$ because $q(j|i, f(i)) \geq 0$ for $i \neq j$ and $u(j) \geq 0$, and so

$$\begin{aligned} u(i) &= \frac{c(i, f(i))}{\alpha + q_i(f)} + \frac{1}{\alpha + q_i(f)} \sum_{j \neq i} q(j|i, f(i)) u(j) \\ &\geq \frac{c(i, f(i))}{\alpha + q_i(f)} = \sum_{j \in S} \phi_{ij}^{(1)}(f) c(j, f(j)). \end{aligned}$$

To prove (5.7) by induction, suppose that it holds for some $n \geq 0$. Then, by (5.6) and the induction hypothesis, we have

$$\begin{aligned} &\sum_{j \in S} \phi_{ij}^{(m+1)}(f) c(j, f(j)) \\ &= \frac{c(i, f(i))}{\alpha + q_i(f)} + \frac{1}{\alpha + q_i(f)} \sum_{k \neq i} q(k|i, f(i)) \sum_{j \in S} \phi_{kj}^{(m)}(f) c(j, f(j)) \\ &\leq \frac{c(i, f(i))}{\alpha + q_i(f)} + \frac{1}{\alpha + q_i(f)} \sum_{k \neq i} q(k|i, f(i)) u(k) = u(i). \end{aligned}$$

Hence (5.7) is valid for $n + 1$ and (5.7) follows. This completes the proof of (a).

(b) Under the condition in (b), there exists a nonnegative real-valued function g on S such that

$$u(i) = \frac{c(i, f(i)) + g(i)}{\alpha + q_i(f)} + \frac{1}{\alpha + q_i(f)} \sum_{j \neq i} q(j|i, f(i)) u(j), \quad i \in S.$$

From part (a) we have that

$$\begin{aligned} u(i) &\geq \sum_{j \in S} \lim_{n \rightarrow \infty} \phi_{ij}^{(n)}(f) [c(j, f(j)) + g(j)] \\ &\geq \sum_{j \in S} \lim_{n \rightarrow \infty} \phi_{ij}^{(n)}(f) c(j, f(j)) = V_\alpha(f, i) \end{aligned}$$

for all $i \in S$, which yields (b). \square

We are now ready for the proofs of Theorems 3.1 and 3.2.

PROOF OF THEOREM 3.1. Let $u : S \rightarrow R$ be a solution of (3.2) and choose an arbitrary $i \in S$. Then, for any $a \in A(i)$ we have that

$$(5.8) \quad u(i) \leq \frac{c(i, a)}{\alpha + q_i(a)} + \frac{1}{\alpha + q_i(a)} \sum_{j \neq i} q(j|i, a) u(j).$$

Recall from Property M_4 that $m(i) := \sup_{a \in A(i)} q_i(a) < \infty$. Then, since

$$(5.9) \quad |u(i)q_i(a)| \leq |u(i)|m(i) < \infty,$$

(5.8) can be rewritten as

$$\alpha u(i) \leq c(i, a) + \sum_{j \in S} q(j|i, a)u(j) \quad \forall a \in A(i),$$

which yields

$$(5.10) \quad \alpha u(i) \leq \inf_{a \in A(i)} \left\{ c(i, a) + \sum_{j \in S} q(j|i, a)u(j) \right\}.$$

Now choose an arbitrary $\varepsilon > 0$. Then, by (3.2), there exists $a_i \in A(i)$ such that

$$u(i) \geq \frac{c(i, a_i)}{\alpha + q_i(a_i)} + \frac{1}{\alpha + q_i(a_i)} \sum_{j \neq i} q(j|i, a_i)u(j) - \varepsilon.$$

Therefore, by (5.9),

$$\begin{aligned} \alpha u(i) &\geq c(i, a_i) + \sum_{j \in S} q(j|i, a_i)u(j) - \varepsilon(\alpha + q_i(a_i)) \\ &\geq \inf_{a \in A(i)} \left\{ c(i, a) + \sum_{j \in S} q(j|i, a)u(j) \right\} - \varepsilon(\alpha + m(i)), \end{aligned}$$

and letting $\varepsilon \rightarrow 0$, we obtain

$$(5.11) \quad \alpha u(i) \geq \inf_{a \in A(i)} \left\{ c(i, a) + \sum_{j \in S} q(j|i, a)u(j) \right\}.$$

By (5.10) and (5.11) we obtain that u satisfies (3.1). The converse, that (3.1) implies (3.2), is proved similarly. \square

PROOF OF THEOREM 3.2. We will prove (a) and (b) together. By the condition $c(i) > 0$ for all $i \in S$, we have $c(i, f(i)) \geq 0$ for all $i \in S$ and $f \in F$. Then Lemma 5.1(a) yields

$$\begin{aligned} V_\alpha(f, i) &= \frac{c(i, f(i))}{\alpha + q_i(f)} + \frac{1}{\alpha + q_i(f)} \sum_{j \neq i} q(j|i, f(i))V_\alpha(f, j) \\ &\geq \frac{c(i, f(i))}{\alpha + q_i(f)} + \frac{1}{\alpha + q_i(f)} \sum_{j \neq i} q(j|i, f(i))V_\alpha^*(j) \\ &\geq \inf_{a \in A(i)} \left\{ \frac{c(i, a)}{\alpha + q_i(a)} + \frac{1}{\alpha + q_i(a)} \sum_{j \neq i} q(j|i, a)V_\alpha^*(j) \right\}. \end{aligned}$$

Hence

$$(5.12) \quad V_\alpha^*(i) \geq \inf_{a \in A(i)} \left\{ \frac{c(i, a)}{\alpha + q_i(a)} + \frac{1}{\alpha + q_i(a)} \sum_{j \neq i} q(j|i, a) V_\alpha^*(j) \right\}.$$

To prove that in fact (5.12) holds with equality, let $M(S)$ be the family of nonnegative functions v on S , and let $T : M(S) \rightarrow M(S)$ be the operator defined by

$$(5.13) \quad T v(i) = \inf_{a \in A(i)} \left\{ \frac{c(i, a)}{\alpha + q_i(a)} + \frac{1}{\alpha + q_i(a)} \sum_{j \neq i} q(j|i, a) v(j) \right\} \quad \forall i \in S, v \in B(S).$$

Obviously, T is monotone, and by (5.12) we have $V_\alpha^* \geq T V_\alpha^*$. Let $u_0 := V_\alpha^*$ and $u_n = T^n u_0$ for $n \geq 1$. Then $V_\alpha^* \geq u_1 \geq \dots \geq u_n \geq \dots \geq 0$ and so $u_n \downarrow u$ for some function with $V_\alpha^* \geq u \geq 0$, which implies that

$$(5.14) \quad V_\alpha^*(i) \geq u(i) \geq T u(i) \quad \forall i \in S.$$

On the other hand, for any $a \in A(i), i \in S$ and $n \geq 1$ we have

$$(5.15) \quad \begin{aligned} u(i) &\leq u_{n+1}(i) = T u_n(i) \\ &= \inf_{a \in A(i)} \left\{ \frac{c(i, a)}{\alpha + q_i(a)} + \frac{1}{\alpha + q_i(a)} \sum_{j \neq i} q(j|i, a) u_n(j) \right\} \\ &\leq \frac{c(i, a)}{\alpha + q_i(a)} + \frac{1}{\alpha + q_i(a)} \sum_{j \neq i} q(j|i, a) u_n(j). \end{aligned}$$

Now take $f \in F$ such that $f(i) = a$. Since $u_n(i) \leq V_\alpha^*(i) \leq V_\alpha(f, i)$, it follows from Assumption A, Lemmas 5.1(a) and (5.2) that

$$\sum_{j \neq i} q(j|i, a) u_n(j) \leq \sum_{j \neq i} q(j|i, f(i)) V_\alpha(f, j) \leq (\alpha + q_i(f)) V_\alpha(f, i) < \infty$$

for all $n \geq 1$ and $i \in S$. Thus, letting $n \rightarrow \infty$ in (5.15), the dominated convergence theorem gives

$$u(i) \leq \frac{c(i, a)}{\alpha + q_i(a)} + \frac{1}{\alpha + q_i(a)} \sum_{j \neq i} q(j|i, a) u(j).$$

As this holds for all $a \in A(i)$ and $i \in S$, we have $u \leq T u$. The latter inequality and (5.14) imply that $u = T u$. Summarizing, u satisfies (3.2) and $V_\alpha^* \geq u$. Therefore, to prove $V_\alpha^* = T V_\alpha^*$, we only need to prove that $u \geq V_\alpha^*$.

Choose $\varepsilon > 0$, and for each $i \in S$, let ε_i be such that $0 < \alpha\varepsilon_i \leq \min\{\alpha\varepsilon, c(i)\}$. Then, as $u = Tu$, from (3.2) we obtain that for each $i \in S$, there exists $f_\varepsilon(i) \in A(i)$ such that

$$\begin{aligned} u(i) &\geq \frac{c(i, f_\varepsilon(i))}{\alpha + q_i(f_\varepsilon)} + \frac{1}{\alpha + q_i(f_\varepsilon)} \sum_{j \neq i} q(j|i, f_\varepsilon(i))u(j) - \frac{\alpha\varepsilon_i}{\alpha + q_i(f_\varepsilon)} \\ &= \frac{c(i, f_\varepsilon(i)) - \alpha\varepsilon_i}{\alpha + q_i(f_\varepsilon)} + \frac{1}{\alpha + q_i(f_\varepsilon)} \sum_{j \neq i} q(j|i, f_\varepsilon(i))u(j). \end{aligned}$$

Noting that $c(i, f_\varepsilon(i)) - \alpha\varepsilon_i \geq 0$ and $\varepsilon_i \leq \varepsilon$ for all $i \in S$, by Lemma 5.1(b) we have

$$u(i) \geq V_\alpha(f_\varepsilon, i) - \varepsilon \geq V_\alpha^*(i) - \varepsilon.$$

Therefore, f_ε is ε -optimal in F and also, letting $\varepsilon \rightarrow 0$, we obtain that $u(i) \geq V_\alpha^*(i)$ for all $i \in S$; that is, $u \geq V_\alpha^*$. This completes the proof of parts (a₁) and (a₂).

By the proof of part (a), there exists a solution u of (3.2) that satisfies $u \leq V_\alpha^*$. Then, under Assumption C, there exists $f^* \in F$ such that

$$u(i) = \frac{c(i, f^*(i))}{\alpha + q_i(f^*)} + \frac{1}{\alpha + q_i(f^*)} \sum_{j \neq i} q(j|i, f^*(i))u(j) \quad \forall i \in S.$$

By Lemma 5.1(b), we have $u \geq V_\alpha(f^*) \geq V_\alpha^*$. Hence $u = V_\alpha(f^*) = V_\alpha^*$, which implies that f^* is optimal in F . \square

6. Proof of Theorem 3.3. We first introduce some general terminology and results on a *noncontrolled, nonhomogeneous continuous-time* Markov process that are needed to prove Theorem 3.3.

Let $p(s, i, t, j)$, defined for $t \geq s \geq 0$ and $i, j \in S$, be the nonhomogeneous transition probability function of a Markov process. We denote by M the linear space of real-valued functions v on $\bar{S} := [0, \infty) \times S$ such that

$$(6.1) \quad \sum_{j \in S} p(s, i, t, j)|v(t, j)| < \infty$$

for each $s \leq t$ and $i \in S$. For each $t \geq 0$ and $v \in M$, let $T_t v$ be the function on \bar{S} defined by

$$(6.2) \quad T_t v(s, i) := \sum_{j \in S} p(s, i, s+t, j)v(s+t, j).$$

By the Chapman–Kolmogorov equation, $\{T_t\}$ is a semigroup of operators on M . Let M_0 be the subset of M consisting of those functions $v \in M$ for which the following hold:

- (a) $\lim_{t \downarrow 0} T_t v(s, i) = v(s, i)$ for every $(s, i) \in \bar{S}$;
 (b) there exists $t_0 > 0$ and $u \in M$ such that

$$T_t |v|(s, i) \leq u(s, i) \quad \forall (s, i) \in \bar{S} \quad \text{and} \quad 0 \leq t \leq t_0.$$

Finally, let $D(L)$ be the set of functions v in M_0 for which the following conditions hold:

- (a) The limit

$$(6.3) \quad Lv(s, i) := \lim_{t \downarrow 0} t^{-1} [T_t v(s, i) - v(s, i)]$$

exists for all $(s, i) \in \bar{S}$;

- (b) $Lv \in M_0$;
 (c) there exist $t_0 > 0$ and $u \in M$ such that

$$t^{-1} |T_t v(s, i) - v(s, i)| \leq u(s, i) \quad \forall (s, i) \in \bar{S}, \quad 0 \leq t \leq t_0.$$

The operator L in (6.3) will be referred to as the *extended generator* of the nonhomogeneous continuous-time Markov process or the semigroup $\{T_t\}$. The set $D(L)$ is called the *domain* of L .

The extended generator L is in fact an extension of the well-known weak infinitesimal operator of $\{T_t\}$ and it has essentially the same properties. For instance:

LEMMA 6.1. (a) If $v \in D(L)$, then $T_t v(s, i) - v(s, i) = \int_0^t T_r (Lv)(s, i) dr$.

(b) For each $v \in D(L)$ and $\alpha > 0$, let $v_\alpha(s, i) := e^{-\alpha s} v(s, i)$. Then v_α is in $D(L)$ and

$$Lv_\alpha(s, i) = e^{-\alpha s} [Lv(s, i) - \alpha v(s, i)] \quad \forall (s, i) \in \bar{S}.$$

(c) Let $q_{ij}(s) := \lim_{h \downarrow 0} [p(s, i, s+h, j) - \delta_{ij}]/h$. If $v(s, i)$ has a partial derivative $v_s(s, i)$ with respect to s and if there exists $u \in M$ such that $|v(s, i)| + |v_s(s, i)| \leq u(s, i)$, then

$$Lv(s, i) = v_s(s, i) + \sum_{j \in S} q_{ij}(s) v(s, j) \quad \forall (s, i) \in \bar{S}.$$

(d) If $v(s, i) \equiv v(i)$ is independent of s , and $\sum_{j \in S} q_{ij}(s) v(j)$ converges and is finite for each $i \in S$ and $s \geq 0$, then

$$Lv(s, i) = \sum_{j \in S} q_{ij}(s) v(j) \quad \forall (s, i) \in \bar{S}.$$

PROOF. Parts (a) and (b) come from Lemma 2.1 in [14], for instance, and parts (c) and (d) follow from Proposition 14.4 in [14]. \square

LEMMA 6.2. Suppose that $\int_0^\infty e^{-\alpha r} T_r c(s, i) dr$ exists for every $\alpha > 0$ and $(s, i) \in \bar{S}$, and let u be a function in $D(L)$.

(a) Suppose that $\lim_{t \rightarrow \infty} e^{-\alpha t} T_t u(s, i) = 0$ for every $(s, i) \in \bar{S}$.

(a₁) If $\alpha u(s, i) = c(s, i) + Lu(s, i)$, then

$$u(s, i) = \int_0^\infty e^{-\alpha r} T_r c(s, i) dr \quad \forall (s, i) \in \bar{S}.$$

(a₂) Fix $\varepsilon \geq 0$. If $\alpha u(s, i) \leq c(s, i) + Lu(s, i) + \varepsilon$ for every $(s, i) \in \bar{S}$, then

$$u(s, i) \leq \int_0^\infty e^{-\alpha r} T_r c(s, i) dr + \alpha^{-1} \varepsilon \quad \forall (s, i) \in \bar{S}.$$

(b) Similarly, for a given $\varepsilon \geq 0$, if u is nonnegative and $\alpha u(s, i) \geq c(s, i) + Lu(s, i) - \varepsilon$ for every $(s, i) \in \bar{S}$, then

$$u(s, i) \geq \int_0^\infty e^{-\alpha r} T_r c(s, i) dr - \alpha^{-1} \varepsilon \quad \forall (s, i) \in \bar{S}.$$

PROOF. (a₁) Suppose that $u \in D(L)$ and $\alpha > 0$. Then, by Lemma 6.1(b), the function $u_\alpha(s, i) := e^{-\alpha s} u(s, i)$ belongs to $D(L)$ and satisfies

$$Lu_\alpha(s, i) = e^{-\alpha s} [Lu(s, i) - \alpha u(s, i)].$$

Thus, Lemma 6.1(a) applied to u_α yields

$$T_t u_\alpha(s, i) - u_\alpha(s, i) = \int_0^t e^{-\alpha(s+r)} T_r [Lu(s, i) - \alpha u(s, i)] dr,$$

which we may rewrite as

$$(6.4) \quad e^{-\alpha t} T_t u(s, i) - u(s, i) = \int_0^t e^{-\alpha r} T_r [Lu(s, i) - \alpha u(s, i)] dr.$$

Hence by the hypotheses on u and (6.4) we get (a₁).

(a₂) From (6.4) we have

$$(6.5) \quad \begin{aligned} e^{-\alpha t} T_t u(s, i) - u(s, i) &\geq - \int_0^t e^{-\alpha r} T_r [c(s, i) + \varepsilon] dr \\ &\geq - \int_0^t e^{-\alpha r} T_r c(s, i) dr - \int_0^\infty e^{-\alpha r} \varepsilon dr, \end{aligned}$$

that is,

$$(6.6) \quad e^{-\alpha t} T_t u(s, i) - u(s, i) \geq - \int_0^t e^{-\alpha r} T_r c(s, i) dr - \alpha^{-1} \varepsilon.$$

Thus, (a₂) follows from (6.6), letting $t \rightarrow \infty$.

(b) As in (6.5), we can obtain

$$(6.7) \quad e^{-\alpha t} T_t u(s, i) - u(s, i) \leq - \int_0^t e^{-\alpha r} T_r [c(s, i) - \varepsilon] dr.$$

Moreover, since $e^{-\alpha t} T_t u(s, i) \geq 0$, from (6.7) we have

$$\begin{aligned} -u(s, i) &\leq -\int_0^t e^{-\alpha r} T_r c(s, i) dr + \int_0^t e^{-\alpha r} \varepsilon dr \\ &\leq -\int_0^t e^{-\alpha r} T_r c(s, i) dr + \alpha^{-1} \varepsilon, \end{aligned}$$

which yields part (b). \square

Now we give our proof of Theorem 3.3, where for any given $\pi \in \Pi$, we denote the above $p(s, i, t, j)$, M , T_t , $D(L)$, L by $p^{\min}(s, i, t, j, \pi)$, M^π , T_t^π , $D(L^\pi)$, L^π , respectively.

PROOF OF THEOREM 3.3. (a₁) Under Assumption A, by Theorem 3.2 we have

$$(6.8) \quad \alpha V_\alpha^*(i) \leq c(i, a) + \sum_{j \in S} q(j|i, a) V_\alpha^*(j) \quad \forall a \in A(i) \text{ and } i \in S.$$

Using the condition in (a), for any $\pi \in \Pi$, from (6.8), (2.1) and (2.2) we obtain

$$(6.9) \quad \alpha V_\alpha^*(i) \leq c(t, i, \pi) + \sum_{j \in S} q_{ij}(t, \pi) V_\alpha^*(j) \quad \forall i \in S \text{ and } t \geq 0.$$

By Assumption B(i) and Remark 2.2(a), $V_\alpha^* \in M^\pi$, whereas by the conditions in (a) and Property M₄ as well as (2.1), $\sum_{j \in S} q_{ij}(t, \pi) V_\alpha^*(j) < \infty$. Then by Lemma 6.1(d), $V_\alpha^* \in D(L^\pi)$. Furthermore, by Lemma 6.2(a) and Lemma 6.1(d), from (6.9) we have $V_\alpha^*(i) \leq \int_0^\infty e^{-\alpha r} T_r^\pi c(s, i) dr = V_\alpha(\pi, s, i)$ for all $i \in S$ and $s \geq 0$. Therefore,

$$(6.10) \quad V_\alpha^*(s, i) = V_\alpha^*(i) = V_{*\Pi}(s, i) \quad \forall i \in S \text{ and } s \geq 0,$$

which yields (a₁).

(a₂) As in the proof of Theorem 3.2(b), there exists $f_\varepsilon \in F$ such that

$$(6.11) \quad V_\alpha^*(i) \geq c(i, f_\varepsilon(i)) + \sum_{j \in S} q(j|i, f_\varepsilon(i)) V_\alpha^*(j) - \alpha \varepsilon \quad \forall i \in S.$$

Hence, by Lemma 6.2(b), (6.10) and (6.11), under Assumption B we have $V_\alpha(f_\varepsilon, s, i) \leq V_\alpha^*(s, i) + \varepsilon = V_{*\Pi}(s, i) + \varepsilon$ for all $i \in S$ and $s \geq 0$. Therefore, part (a₂) follows.

(b) Since, under Assumption C the condition in part (a) is satisfied, (6.10) holds. Moreover, Theorem 3.2(b) gives the existence of $f_* \in F$ such that

$$(6.12) \quad V_\alpha^*(i) = c(i, f_*(i)) + \sum_{j \in S} q(j|i, f_*(i)) V_\alpha^*(j) \quad \forall i \in S.$$

By Lemma 6.2(a), (6.10) and (6.12), under Assumption B we have $V_\alpha(f_*, s, i) = V_\alpha^*(i) = V_{*\Pi}(s, i)$ for all $i \in S$ and $s \geq 0$, and (b) follows. \square

7. Proofs of Theorems 3.4 and 3.5. Let $B(S)$ be the family of functions $u : S \rightarrow R$ such that $|u(i)| \leq c(u)V_\alpha^*(i)$ for all $i \in S$ and some constant $c(u)$ that may depend on u . Obviously, $V_\alpha^* \in B(S)$. Furthermore, by Assumption B(ii), (2.10) and (2.11), if u is in $B(S)$, then

$$(7.1) \quad \lim_{T \rightarrow \infty} e^{-\alpha T} P^{\min}(s, T, \pi)|u| = \mathbf{0} \quad \forall \pi \in \Pi \text{ and } s \geq 0.$$

LEMMA 7.1. For any $\pi \in \Pi$, $t \geq s \geq 0$ and $u \in B(S)$, under Assumption I' we have:

- (a) $Q(t, \pi)\mathbf{1} = \mathbf{0}$;
- (b) $P^{\min}(s, t, \pi)[Q(t, \pi)u] = [P^{\min}(s, t, \pi)Q(t, \pi)]u$ for a.e. $t \geq 0$;
- (c) $\frac{\partial}{\partial t} P^{\min}(s, t, \pi) = P^{\min}(s, t, \pi)Q(t, \pi)$ for a.e. $t \geq s$.

PROOF. By (2.1) and Property M₄ in Section 2, we have

$$\sum_{j \in S} \int_A |q(j|i, a)|\pi_t(da|i) = \int_A |2q(i|i, a)|\pi_t(da|i) \leq 2m(i) < \infty.$$

Hence, by the Fubini theorem and (2.1), for any $i \in S$ we have

$$(7.2) \quad \sum_{j \in S} q_{ij}(t, \pi) = \sum_{j \in S} \int_A q(j|i, a)\pi_t(da|i) = \int_A \left(\sum_{j \in S} q(j|i, a) \right) \pi_t(da|i) = 0,$$

which yields (a). Part (b) follows from Assumption I' and the Fubini theorem. Finally, part (c) follows from Proposition 2.1(a). \square

LEMMA 7.2. If Assumptions I' and B hold, then for any $u \in B(S)$, $\pi \in \Pi$, $s \geq 0$ and $\varepsilon \geq 0$:

- (a) if $\alpha u \leq c(t, \pi) + Q(t, \pi)u$ for a.e. $t \geq 0$, then $u \leq V_\alpha(\pi, s)$ and
- (b) if $\alpha u \geq c(t, \pi) + Q(t, \pi)u - \varepsilon\mathbf{1}$ for a.e. $t \geq 0$, then $u \geq V_\alpha(\pi, s) - \alpha^{-1}\varepsilon\mathbf{1}$.

PROOF. (a) By Lemma 7.1, for $T > 0$ we have

$$(7.3) \quad \begin{aligned} & \int_s^T e^{-\alpha(t-s)} P^{\min}(s, t, \pi)c(t, \pi) dt \\ & \geq \int_s^T e^{-\alpha(t-s)} P^{\min}(s, t, \pi)[\alpha u - Q(t, \pi)u] dt \\ & = \left[\int_0^{T-s} P^{\min}(s, s+t, \pi) d(-e^{-\alpha t}) - \int_0^{T-s} e^{-\alpha t} dP^{\min}(s, s+t, \pi) \right] u \\ & = u - e^{-\alpha(T-s)} P^{\min}(s, T, \pi)u. \end{aligned}$$

Letting $T \rightarrow \infty$ in (7.3), by (7.1) we get part (a). The proof of part (b) is similar. \square

PROOF OF THEOREM 3.4. (a) Let T be the operator in (5.13). As in the proof of Theorem 3.2(a), there exists a real-valued function u such that $u = Tu$ and $V_\alpha^* \geq u$. To prove that $u \geq V_\alpha^*$, we see that, under Assumption C, there exists $f^* \in F$ such that

$$u(i) = \frac{c(i, f^*(i))}{\alpha + q_i(f^*)} + \frac{1}{\alpha + q_i(f^*)} \sum_{j \neq i} q(j|i, f^*)u(j), \quad i \in S.$$

Thus, by Lemma 5.1, we have $u \geq V_\alpha(f^*) \geq V_\alpha^*$. Hence, $V_\alpha(f^*) = V_\alpha^* = u$ and V_α^* satisfies (3.1).

(b) By part (a), we have

$$\alpha V_\alpha^*(i) \leq c(i, a) + \sum_{j \in S} q(j|i, a) V_\alpha^*(j) \quad \forall a \in A(i), i \in S,$$

which gives that, for any $\pi \in \Pi$,

$$\alpha V_\alpha^*(i) \leq c(t, i, \pi) + \sum_{j \in S} q_{ij}(t, \pi) V_\alpha^*(j) \quad \forall i \in S \text{ and a.e. } t \geq 0.$$

This implies, by Assumption I' and Lemma 7.2(a), that $V_\alpha^*(i) \leq V_\alpha(\pi, i)$ for all $\pi \in \Pi$. Hence, $V_\alpha^* = V_{*\Pi}$, which yields (b).

(c) In the proof of (a), we have shown that there exists $f^* \in F$ such that $V_\alpha(f^*) = V_\alpha^*$. By part (b), $V_\alpha(f^*) = V_{*\Pi}$, which means that f^* is optimal in Π . \square

PROOF OF THEOREM 3.5. (a) As in the proof of Theorem 3.2(a), there exists a real-valued function u such that $u = Tu$ and $V_\alpha^* \geq u$. To prove that $u \geq V_\alpha^*$, by $c(i) > 0$ for all $i \in S$, as in the proof of Theorem 3.2, we can derive that $u \geq V_\alpha^*$ and so $V_\alpha^* = TV_\alpha^*$, which yields (a), and also that there exists $f^* \in F$ such that

$$(7.4) \quad V_\alpha(f^*, i) \leq V_\alpha^*(i) + \varepsilon \quad \forall i \in S.$$

(b) This can be obtained as in the proof of Theorem 3.4(b).

(c) Part (c) follows from (b) and (7.4). \square

8. Proof of Proposition 2.1. Multiplying both sides of the equality in part (a) by the vector $\mathbf{1}$, we see that part (b) follows from (a). Moreover, by (7.2), $Q(t, \pi)\mathbf{1} = \mathbf{0}$, which together with the Fubini theorem yields part (c). Hence, we only need to prove part (a), that is,

$$(8.1) \quad P^{\min}(s, t, \pi) = \int_s^t P^{\min}(s, v, \pi)Q(v, \pi)dv + I.$$

To get (8.1) we will first use induction and (2.3)–(2.5) to prove that for $k \geq 0$ and $t > s \geq 0$,

$$(8.2) \quad \begin{aligned} & \int_s^t P_{k+1}(s, v, \pi)D(v, \pi)dv \\ &= \int_s^t P_k(s, v, \pi)(D(v, \pi) + Q(v, \pi))dv - P_{k+1}(s, t, \pi). \end{aligned}$$

Note that by (2.3) and (2.4),

$$(8.3) \quad \int_s^t \Delta(s, u, \pi) D(u, \pi) du = I - \Delta(s, t, \pi).$$

Hence, taking $n = 0$ in (2.5) we get

$$\begin{aligned} & \int_s^t P_1(s, v, \pi) D(v, \pi) dv \\ &= \int_s^t \left[\int_s^v (\Delta(s, u, \pi)(Q(u, \pi) + D(u, \pi))\Delta(u, v, \pi) du \right] D(v, \pi) dv \\ &= \int_s^t \Delta(s, u, \pi)(Q(u, \pi) + D(u, \pi)) \int_u^t \Delta(u, v, \pi) D(v, \pi) dv du \\ &= \int_s^t \Delta(s, u, \pi)(Q(u, \pi) + D(u, \pi))(I - \Delta(u, t, \pi)) du \\ &= \int_s^t \Delta(s, u, \pi)(Q(u, \pi) + D(u, \pi)) du - P_1(s, t, \pi). \end{aligned}$$

Thus, (8.2) holds for $k = 0$. Suppose that (8.2) holds for $k = n$. Then, by (2.5) and the induction hypothesis, we obtain

$$\begin{aligned} & \int_s^t P_{n+2}(s, v, \pi) D(v, \pi) dv \\ &= \int_s^t \left[\int_s^v \Delta(s, u, \pi)(Q(u, \pi) + D(u, \pi)) P_{n+1}(u, v, \pi) du \right] D(v, \pi) dv \\ &= \int_s^t \Delta(s, u, \pi)(Q(u, \pi) + D(u, \pi)) \int_u^t P_{n+1}(u, v, \pi) D(v, \pi) dv du \\ &= \int_s^t \Delta(s, u, \pi)(Q(u, \pi) + D(u, \pi)) \\ & \quad \times \left[\int_u^t P_n(u, v, \pi)(Q(v, \pi) + D(v, \pi)) dv - P_{n+1}(u, t, \pi) \right] du \\ &= \int_s^t \int_u^t \Delta(s, u, \pi)(Q(u, \pi) + D(u, \pi)) P_n(u, v, \pi) \\ & \quad \times (Q(v, \pi) + D(v, \pi)) dv du \\ & \quad - \int_s^t \Delta(s, u, \pi)(Q(u, \pi) + D(u, \pi)) P_{n+1}(u, t, \pi) du \\ &= \int_s^t \left(\int_s^v \Delta(s, u, \pi)(Q(u, \pi) + D(u, \pi)) P_n(u, v, \pi) du \right) \\ & \quad \times (Q(v, \pi) + D(v, \pi)) dv - P_{n+2}(s, t, \pi) \quad [\text{by (2.5) again}] \\ &= \int_s^t P_{n+1}(s, v, \pi)(Q(v, \pi) + D(v, \pi)) dv - P_{n+2}(s, t, \pi), \end{aligned}$$

which means that (8.2) holds for $k = n + 1$. Hence, (8.2) is valid for all $k \geq 0$.

Finally, from (8.2) and (2.6), we can get (8.1) as

$$\begin{aligned} \sum_{n=0}^{\infty} P_{n+1}(s, t, \pi) &= \int_s^t \sum_{n=0}^{\infty} P_n(s, v, \pi) (Q(v, \pi) + D(v, \pi)) dv \\ &\quad - \int_s^t \sum_{n=0}^{\infty} P_{n+1}(s, v, \pi) D(v, \pi) dv. \end{aligned}$$

Hence

$$\begin{aligned} P^{\min}(s, t, \pi) &= \int_s^t P^{\min}(s, v, \pi) (Q(v, \pi) + D(v, \pi)) dv - \int_s^t P^{\min}(s, v, \pi) D(v, \pi) dv \\ &\quad + \int_s^t \Delta(s, v, \pi) D(v, \pi) dv + \Delta(s, t, \pi) \\ &= \int_s^t P^{\min}(s, v, \pi) Q(v, \pi) dv + I \quad [\text{by (8.3)}], \end{aligned}$$

which yields (8.1). \square

REFERENCES

- [1] ANDERSON, W. J. (1991). *Continuous Time Markov Chains*. Springer, New York.
- [2] BATHER, J. (1976). Optimal stationary policies for denumerable Markov chains in continuous-time. *Adv. in Appl. Probab.* **8** 114–158.
- [3] BERTSEKAS, D. P. (1987). *Dynamic Programming: Deterministic and Stochastic Models*. Prentice–Hall, Englewood Cliffs, NJ.
- [4] CAVAZOS-CADENA, R. and GAUCHERAND, E. (1996). Value iteration in a class of average controlled Markov chains with unbounded costs: Necessary and sufficient conditions for pointwise convergence. *J. Appl. Probab.* **33** 986–1002.
- [5] CHEN, M. F. (1990). On three classical problems for Markov chains with continuous time parameters. *J. Appl. Probab.* **28** 305–320.
- [6] CHUNG, K. L. (1960). *Markov Chains with Stationary Transition Probabilities*. Springer, Berlin.
- [7] DONG, Z. Q. (1979). Continuous time Markov decision programming with average reward criterion—countable state and action space. *Sci. Sinica* **SP(II)** 131–148.
- [8] FEINBERG, E. A. (1998). Continuous time discounted jump Markov decision processes: A discrete-event approach. Preprint.
- [9] FELLER, W. (1940). On the integro-differential equations of purely discontinuous Markoff processes. *Trans. Amer. Math. Soc.* **48** 488–515.
- [10] GUO, X. P. and LIU, K. (2001). A note on optimality conditions for continuous-time Markov decision processes with average cost criterion. *IEEE Trans. Automat. Control* **46** 1984–1989.
- [11] GUO, X. P. and ZHU, W. P. (2002). Optimality conditions for CTMDP with average cost criterion. In *Markov Processes and Controlled Markov Chains* (Z. T. Hou, J. A. Filar and A. Y. Chen, eds.) Chap. 10. Kluwer, Dordrecht.

- [12] GUO, X. P. and ZHU, W. P. (2002). Denumerable-state continuous-time Markov decision processes with unbounded transition and reward rates under the discounted criterion. *J. Appl. Probab.* **39** 233–250.
- [13] HAVIV, M. and PUTERMAN, M. L. (1998). Bias optimality in controlled queuing systems. *J. Appl. Probab.* **35** 16–150.
- [14] HERNÁNDEZ-LERMA, O. (1994). *Lectures on Continuous-time Markov Control Processes*. Sociedad Matemática Mexicana, México City.
- [15] HEYMAN, D. P. and SOBEL, M. J. (1984). *Stochastic Models in Operations Research 2*. McGraw–Hill, New York.
- [16] HOU, Z. T. (1994). *The Q-matrix Problems on Markov Processes*. Science and Technology Press of Hunan, Changsha, China (in Chinese).
- [17] HOU, Z. T. and GUO, X. P. (1998). *Markov Decision Processes*. Science and Technology Press of Hunan, Changsha, China (in Chinese).
- [18] KAKUMANU, P. (1971). Continuously discounted Markov decision model with countable state and action spaces. *Ann. Math. Statist.* **42** 919–926.
- [19] LEFÈVRE, C. (1981). Optimal control of a birth and death epidemic process. *Oper. Res.* **29** 971–982.
- [20] LEWIS, M. E. and PUTERMAN, M. (2001). A probabilistic analysis of bias optimality in unichain Markov decision processes. *IEEE Trans. Automat. Control* **46** 96–100.
- [21] LEWIS, M. E. and PUTERMAN, M. (2000). A note on bias optimality in controlled queueing systems. *J. Appl. Probab.* **37** 300–305.
- [22] MILLER, R. L. (1968). Finite state continuous time Markov decision processes with an infinite planning horizon. *J. Math. Anal. Appl.* **22** 552–569.
- [23] PUTERMAN, M. L. (1994). *Markov Decision Processes*. Wiley, New York.
- [24] SCHLÖGL, F. (1972). Chemical reaction models for phase transition. *Z. Phys.* **253** 147–161.
- [25] SENNOTT, L. I. (1999). *Stochastic Dynamic Programming and the Control of Queueing Systems*. Wiley, New York.
- [26] SERFOZO, R. (1981). Optimal control of random walks, birth and death processes, and queues. *Adv. in Appl. Probab.* **13** 61–83.
- [27] SONG, J. S. (1987). Continuous time Markov decision programming with nonuniformly bounded transition rates. *Sci. Sinica* **12** 1258–1267 (in Chinese).
- [28] TIJMS, H. C. (1994). *Stochastic Models: An Algorithmic Approach*. Wiley, Chichester.
- [29] WALRAND, J. (1988). *An Introduction to Queueing Networks*. Prentice–Hall, Englewood Cliffs, NJ.
- [30] YUSHKEVICH, A. A. and FEINBERG, E. A. (1979). On homogeneous Markov model with continuous time and finite or countable state space. *Theory Probab. Appl.* **24** 156–161.

ZHONGSHAN UNIVERSITY
 DEPARTMENT OF STATISTICAL SCIENCE
 GUANGZHOU 510275
 PEOPLES REPUBLIC OF CHINA
 E-MAIL: mcsgxp@zsu.edu.cn

DEPARTAMENTO DE MATEMÁTICAS
 CINVESTAV-IPN
 A. POSTAL 14-740
 MÉXICO D.F. 0700
 MÉXICO
 E-MAIL: ohernand@math.cinvestav.mx