CrossMark

Average cost criterion induced by the regular utility function for continuous-time Markov decision processes

Qingda Wei[1] · Xian Chen[2]

© Springer Science+Business Media New York 2017

Abstract In this paper we study the average cost criterion induced by the regular utility function (U-average cost criterion) for continuous-time Markov decision processes. This criterion is a generalization of the risk-sensitive average cost and expected average cost criteria. We first introduce an auxiliary risk-sensitive first passage optimization problem and obtain the properties of the corresponding optimal value function under the slight conditions. Then we show that the pair of the optimal value functions of the risk-sensitive average cost criterion and the risk-sensitive first passage criterion is a solution to the optimality equation of the risk-sensitive average cost criterion allowing the risk-sensitivity parameter to take any nonzero value. Moreover, we have that the optimal value function of the risk-sensitive average cost criterion is continuous with respect to the risk-sensitivity parameter. Finally, we give the connections between the U-average cost criterion and the average cost criteria induced by the identity function and the exponential utility function, and prove the existence of a U-average optimal deterministic stationary policy in the class of all randomized Markov policies.

Keywords Continuous-time Markov decision processes · Regular utility function ·
U-average cost criterion · Optimal policy

1 Introduction

Continuous-time Markov decision processes (CTMDPs) have rich applications in the queueing systems, inventory management, telecommunication, the control of the epidemic, etc;

✉ Xian Chen
chenxian@amss.ac.cn

Qingda Wei
weiqd@hqu.edu.cn

[1] School of Economics and Finance, Huaqiao University, Quanzhou, 362021, People's Republic of China

[2] School of Mathematical Sciences, Xiamen University, Xiamen, 361005, People's Republic of China

🙋 Springer

see, for instance, Puterman (1994), Kitaev and Rykov (1995), Guo and Hernandez-Lerma (2009), Guo et al. (2012) and the references therein. The average cost criterion is a common optimality criterion in the CTMDPs, which includes the expected average cost and the risk-sensitive average cost criteria. For the expected average cost criterion, the decision-maker is supposed to be risk-neutral and there exists a vast amount of literature; see, for instance, Puterman (1994), Kitaev and Rykov (1995), Guo and Hernandez-Lerma (2009), Guo et al. (2012), Wei and Chen (2014) and the extensive references therein. For the risk-sensitive average cost criterion, the exponential utility function is employed to characterize the risk preferences of the decision-maker. More specifically, when the risk-sensitivity parameter of the exponential utility function takes positive (negative) values, the decision-maker is risk-averse (risk-seeking). Although the risk-sensitive average cost criterion for the discrete-time MDPs has been widely studied (see, for instance, Cavazos-Cadena and Fernandez-Gaucherand (2005), Cavazos-Cadena (2010) and Cavazos-Cadena and Hernandez-Hernandez (2011) for the countable state space and Jaśkiewicz (2007) and Di Masi and Stettner (2007) for the uncountable state space), there exists a handful of literature on this criterion for the CTMDPs. Ghosh and Saha (2014) and Wei and Chen (2016) investigate the risk-sensitive average cost criterion with a positive risk-sensitivity parameter for the CTMDPs and obtain the existence of optimal policies via the optimality equation approach. Moreover, to the best of our knowledge, there is no existing literature dealing with the risk-sensitive average cost criterion which allows the risk-sensitivity parameter to take negative values for the CTMDPs.

On the other hand, the risk preferences of the decision-maker may be described neither by the identity function nor by the exponential utility function in the real-world applications. Except the identity function and the exponential utility function, there are other utility functions to describe the risk preferences of the decision-maker, such as the logarithmic utility function, the power utility function, etc. Thus, it is desirable for us to consider the average cost criterion induced by the general utility function. For the discrete-time MDPs, Bäuerle and Rieder (2014) discusses the average cost criterion induced by the power utility function and Cavazos-Cadena and Hernández-Hernández (2016) studies the average cost criterion induced by the regular utility function which is referred to as the U-average cost criterion for simplicity. The U-average cost criterion includes the expected average cost criterion induced by the identity function, the risk-sensitive average cost criterion induced by the exponential utility function and the average cost criterion induced by the logarithmic and power utility functions. For the CTMDPs, as far as we can tell, the discussions on the average cost criterion only focus on the expected average cost and risk-sensitive average cost criteria.

In this paper we study the U-average cost criterion for the CTMDPs. The state space is a finite set and the action space is a Borel space. Since the existence of optimal policies for the U-average cost criterion is closely connected with the risk-sensitive average cost criterion allowing the risk-sensitivity parameter to take any nonzero value, we need to investigate the risk-sensitive average cost criterion at first. Under the optimality conditions of the paper (i.e., the standard continuity-compactness condition and the irreducibility condition), we first show that the simultaneous Doeblin condition holds (see Theorem 3.1). The simultaneous Doeblin condition plays a crucial role in establishing the existence of a solution to the risk-sensitive average cost optimality equation. Then we introduce an auxiliary risk-sensitive first passage optimization problem and obtain the properties of the corresponding optimal value function (see Theorem 3.2). Basing on Theorem 3.2, we have that the pair of the optimal value functions of the risk-sensitive average cost criterion and the risk-sensitive

first passage criterion is a solution to the optimality equation of the risk-sensitive average cost criterion allowing the risk-sensitivity parameter to take any nonzero value (see Theorem 3.3), which generalizes the results in Ghosh and Saha (2014) and Wei and Chen (2016) only allowing the risk-sensitivity parameter to take positive values (see Remarks 3.3 and 3.4). It should be noted that the extension is nontrivial. Moreover, we prove that the optimal value function of the risk-sensitive average cost criterion is continuous with respect to the risk-sensitivity parameter (see Theorem 3.4). Finally, using the results on the expected average cost and risk-sensitive average cost criteria, we show that there exists an optimal deterministic stationary policy in the class of all randomized Markov policies for the U-average cost criterion. Moreover, we have that the optimal value function and the set of all optimal stationary policies for the U-average cost criterion coincide with those for the average cost criterion induced either by the identity function or by the exponential utility function with some risk-sensitivity parameter (see Theorem 4.1).

The rest of this paper is organized as follows. In Section 2, we introduce the control model and the optimality criterion. In Section 3, we present the results on the simultaneous Doeblin condition, the risk-sensitive first passage criterion and the risk-sensitive average cost criterion, whose proofs are given in Sections 5–7. In Section 4, we state and prove the main results on the U-average cost criterion. In Section 8, we conclude with some remarks.

2 Preliminaries

In this section, we introduce the control model and the average cost criterion induced by the regular utility function for the CTMDPs. The control model in this paper is given by

$$\mathcal{M} := \{S, A, (A(i), i \in S), q(j|i, a), c(i, a)\}.$$

- The state space S is a finite set endowed with the discrete topology.
- The action space A is a Borel space with the Borel σ-algebra $\mathcal{B}(A)$.
- The set of all admissible actions in state $i \in S$ denoted by $A(i)$ is a Borel-measurable subset of A. Moreover, we set $K := \{(i, a)|i \in S, a \in A(i)\}$ which stands for the set of all admissible state-action pairs.
- The real-valued measurable transition rate $q(j|i, a)$ satisfies $q(j|i, a) \geq 0$ for all $(i, a) \in K$ and $j \neq i$, and is conservative (i.e., $\sum_{j \in S} q(j|i, a) = 0$ for all $(i, a) \in K$).
- The positive real-valued cost rate function $c(i, a)$ is measurable in $a \in A(i)$ for each $i \in S$.

The evolution of a CTMDP is intuitively described as follows. The state of the dynamical system is observed continuously by a decision-maker. When the state of the system occupies $i \in S$, the decision-maker takes an action a from the set of all admissible actions $A(i)$. As a result of this action, a cost is incurred at the rate $c(i, a)$, and the system stays at state i for a random time following the exponential distribution and then jumps to a new state $j \neq i$ according to some distribution (see Proposition B.8 in Guo and Hernández-Lerna (2009, p.205) for the explicit expressions of the corresponding distributions). When the state of the system transits to the state j, the above procedure is repeated.

Below we formally give a mathematical description.

Let $S_\infty := S \cup \{i_\infty\}$ with an isolated point $i_\infty \notin S$, $\mathbb{R}_+ := (0, +\infty)$, $\Omega^0 := (S \times \mathbb{R}_+)^\infty$, $\Omega := \Omega^0 \cup \{(i_0, \theta_1, i_1, \ldots, \theta_{m-1}, i_{m-1}, \infty, i_\infty, \infty, i_\infty, \ldots)|i_0 \in S, \ i_l \in S, \ \theta_l \in \mathbb{R}_+ \text{ for each } 1 \leq l \leq m - 1, \ m \geq 2\}$, and $\mathcal{F}$ be the Borel σ-algebra of Ω. For each

$\omega = (i_0, \theta_1, i_1, \dots) \in \Omega$, define $X_0(\omega) := i_0$, $T_0(\omega) := 0$, $X_m(\omega) := i_m$, $T_m(\omega) := \theta_1 + \theta_2 + \cdots + \theta_m$ for $m \geq 1$, $T_\infty(\omega) := \lim_{m \to \infty} T_m(\omega)$, and the state process

$$\xi_t(\omega) := \sum_{m \geq 0} I_{\{T_m \leq t < T_{m+1}\}} i_m + I_{\{t \geq T_\infty\}} i_\infty \quad \text{for } t \geq 0,$$

where I_D denotes the indicator function of a set D. The process after T_∞ is regarded to be absorbed in the state i_∞. Hence, we write $q(i_\infty | i_\infty, a_\infty) = 0$, $c(i_\infty, a_\infty) = 0$, $A(i_\infty) := \{a_\infty\}$, $A_\infty := A \cup \{a_\infty\}$, where a_∞ is an isolated point. Let $\mathcal{F}_t := \sigma(\{T_m \leq s, X_m = i\} : i \in S, s \leq t, m \geq 0)$ for $t \geq 0$, $\mathcal{F}_{s-} := \bigvee_{0 \leq t < s} \mathcal{F}_t$, and $\mathcal{P} := \sigma(\{D \times \{0\}, D \in \mathcal{F}_0\} \cup \{D \times (s, \infty), D \in \mathcal{F}_{s-}, s > 0\})$ which denotes the σ-algebra of predictable sets on $\Omega \times [0, \infty)$ related to $\{\mathcal{F}_t\}_{t \geq 0}$.

Before giving the optimality criterion, we need to introduce the following definition of a randomized Markov policy.

Definition 2.1 A $\mathcal{P}$-measurable transition probability $\pi(\cdot | \omega, t)$ on $(A_\infty, \mathcal{B}(A_\infty))$, concentrated on $A(\xi_{t-}(\omega))$ is called a randomized Markov policy if there exists a kernel φ on A_∞ given $S_\infty \times [0, \infty)$ such that $\pi(\cdot | \omega, t) = \varphi(\cdot | \xi_{t-}(\omega), t)$. A policy π is said to be deterministic stationary if there exists a function f on S_∞ satisfying $f(i) \in A(i)$ for all $i \in S_\infty$ and $\pi(\cdot | \omega, t) = \delta_{f(\xi_{t-}(\omega))}(\cdot)$, where $\delta_x(\cdot)$ is the Dirac measure concentrated at the point x.

Let Π and F be the set of all randomized Markov policies and the set of all deterministic stationary policies, respectively.

Given an arbitrary initial state $i \in S$ and any policy $\pi \in \Pi$, employing Theorem 4.27 in Kitaev and Rykov (1995), we obtain the existence of a unique probability measure denoted by P_i^π on $(\Omega, \mathcal{F})$. The notation E_i^π represents the expectation operator with respect to P_i^π.

Let $\mathcal{U}$ be the set of all the real-valued utility functions U on $\mathbb{R}_+$ satisfying the following properties: (i) U has continuous derivatives up to second order; (ii) the first derivative $U'(x)$ is positive for all $x \in \mathbb{R}_+$. For any $U \in \mathcal{U}$, the Arrow-Pratt risk-sensitivity function $\mathcal{A}_U$ is defined by $\mathcal{A}_U(x) := \frac{U''(x)}{U'(x)}$ for all $x \in \mathbb{R}_+$, where U'' denotes the second derivative of U. Below we give the definition of a regular utility function in Cavazos-Cadena and Hernández-Hernández (2016).

Definition 2.2 A utility function $U \in \mathcal{U}$ is said to be regular if $\lambda_U := \lim_{x \to \infty} \mathcal{A}_U(x)$ exists in $\mathbb{R} := (-\infty, \infty)$. The constant λ_U is called the asymptotic risk-sensitivity parameter of the regular utility function U.

Let $\mathcal{U}_r$ be the set of all the regular utility functions in $\mathcal{U}$. For any $U \in \mathcal{U}_r$, $i \in S$ and $\pi \in \Pi$, the average cost criterion induced by the regular utility function U is defined by

$$J_U(i, \pi) := \limsup_{T \to \infty} \frac{1}{T} U^{-1} \left(E_i^\pi \left[U \left(\int_0^T \int_A c(\xi_t, a) \pi(da | \xi_t, t) dt \right) \right] \right), \quad (2.1)$$

where U^{-1} denotes the inverse function of U. In the following, we refer to the average cost criterion defined in Eq. 2.1 as the U-average cost criterion for simplicity.

Remark 2.1 Let $Y := \int_0^T \int_A c(\xi_t, a) \pi(da | \xi_t, t) dt$ be the total cost incurred during the finite time interval $[0, T]$. The quantity $U^{-1}(E_i^\pi[U(Y)])$ stands for the certainty equivalent of Y with respect to the utility function U and the decision-maker is indifferent between paying the random cost Y or the certainty equivalent of Y; see the detailed discussions in

Cavazos-Cadena and Fernandez-Gaucherand (2005), Cavazos-Cadena (2010), Cavazos-Cadena and Hernandez-Hernandez (2011), Bauerle and Rieder (2014), and Cavazos-Cadena and Hernandez-Hernandez (2016).

Definition 2.3 A policy $\pi^* \in \Pi$ is said to be U-average optimal if

$$J_U(i, \pi^*) = \inf_{\pi \in \Pi} J_U(i, \pi) =: J_U^*(i)$$

for all $i \in S$. The function J_U^* on S is referred to as the optimal value function of the U-average cost criterion.

Finally, we introduce the expected average cost and risk-sensitive average cost criteria which are the particular cases of the U-average cost criterion and play an important role in proving the existence of U-average optimal policies.

For each real number $\lambda \in \mathbb{R}$, define the real-valued function V_λ on $\mathbb{R}_+$ as follows:

$$V_\lambda(x) = \begin{cases} e^{\lambda x}, & \text{if } \lambda > 0, \\ x, & \text{if } \lambda = 0, \\ -e^{\lambda x}, & \text{if } \lambda < 0, \end{cases} \tag{2.2}$$

for all $x \in \mathbb{R}_+$. It is obvious that V_λ belongs to $\mathcal{U}_r$ for all $\lambda \in \mathbb{R}$. Then for each $\lambda \neq 0$, $i \in S$ and $\pi \in \Pi$, by Eq. 2.1 we have

$$J_{V_\lambda}(i, \pi) = \limsup_{T \to \infty} \frac{1}{\lambda T} \ln E_i^\pi \left[e^{\lambda \int_0^T \int_A c(\xi_t, a)\pi(da|\xi_t, t)dt} \right],$$

which is induced by the exponential utility function and called the risk-sensitive average cost criterion in Ghosh and Saha (2014) and Wei and Chen (2016). For $\lambda = 0$, Eq. 2.1 gives

$$J_{V_0}(i, \pi) = \limsup_{T \to \infty} \frac{1}{T} E_i^\pi \left[\int_0^T \int_A c(\xi_t, a)\pi(da|\xi_t, t)dt \right]$$

for all $i \in S$ and $\pi \in \Pi$, which is induced by the identity function and referred to as the expected average cost criterion; see, for instance, Puterman (1994), Kitaev and Rykov (1995), Guo and Hernandez-Lerma (2009), Guo et al. (2012), and Wei and Chen (2014). Hence, the V_λ-average cost criterion contains the expected average cost and risk-sensitive average cost criteria.

3 The V_λ-average cost criterion

In this section, we aim to give the optimality conditions for the existence of V_λ-average optimal policies and establish the existence of a solution to the optimality equation of the V_λ-average cost criterion. Since the existence of optimal policies and the optimality equation for the V_0-average cost criterion (i.e., the expected average cost criterion) have been well studied (see Puterman (1994), Kitaev and Rykov (1995), Guo and Hernández-Lerma (2009), and Wei and Chen (2014) and the references therein), we mainly focus on the V_λ-average cost criterion for all $\lambda \neq 0$ (i.e., the risk-sensitive average cost criterion) below. To do so, we introduce the following assumption in Wei and Chen (2016), i.e., the usual continuity-compactness condition and the irreducibility condition.

Assumption 3.1 (i) For each $i \in S$, the set $A(i)$ is compact.
(ii) For each $i, j \in S$, the functions $c(i, a)$ and $q(j|i, a)$ are continuous in $a \in A(i)$.

(iii)   For each $f \in F$, the continuous-time Markov chain associated with the transition
         rate $q(\cdot|\cdot, f)$ is irreducible, which means that for any two states $i \neq j$, there exist
         different states $j_1 = i$, $j_2, \ldots, j_m$ such that $q(j_2|j_1, f) \cdots q(j|j_m, f) > 0$, where
         we write $q(j|i, f) := q(j|i, f(i))$.

Remark 3.1 For each $f \in F$, f can be viewed as $\prod_{i \in S} f(i) \in \prod_{i \in S} A(i)$. Thus, by
Assumption 3.1(i) and the Tychonoff theorem, we have that F is compact and metrizable.
    Fix any state $z \in S$ throughout the paper. Define

$$\tau_z := \inf\{t \geq T_1 : \xi_t = z\} \text{ with the convention that } \inf \emptyset := \infty.$$

τ_z is the time of the first entry into the state z after the first transition has occurred. Under
Assumption 3.1, the following result indicates that the simultaneous Doeblin condition (i.e.,
the statement of Theorem 3.1(b)) for the CTMDPs holds.

Theorem 3.1 *Suppose that Assumption 3.1 is satisfied. Then the following assertions hold.*

(a)   *There exists a constant $L \in \mathbb{R}_+$ such that $E_i^f[\tau_z] \leq L$ for all $i \in S$ and $f \in F$.*
(b)   *There exist constants $t_0 \in \mathbb{R}_+$ and $\alpha \in (0, 1)$ such that $P_i^f(\tau_z > t_0) \leq \alpha$ for all $i \in S$
       and $f \in F$.*

Proof See Section 5.   □

Remark 3.2 The assertion of part (a) is equivalent to that of part (b). Indeed, from the proof
of Theorem 3.1, we see that part (a) implies part (b). On the other hand, suppose that part
(b) holds. Then by an induction argument, we have $P_i^f(\tau_z > mt_0) \leq \alpha^m$ for all $i \in S$,
$f \in F$ and $m = 1, 2, \ldots$. Thus, employing the last inequality and Lemma 3.4 in Kallenberg
(2012, p.49), we obtain

$$E_i^f[\tau_z] = \int_0^\infty P_i^f(\tau_z > t)dt = \sum_{m=0}^\infty \int_{mt_0}^{(m+1)t_0} P_i^f(\tau_z > t)dt \leq t_0 \sum_{m=0}^\infty P_i^f(\tau_z > mt_0) \leq \frac{t_0}{1 - \alpha}$$

for all $i \in S$ and $f \in F$. Hence, part (b) implies that part (a) holds with $L = \frac{t_0}{1-\alpha}$.
    To obtain the existence of a solution to the optimality equation of the risk-sensitive aver-
age cost criterion, we need to introduce the following auxiliary risk-sensitive first passage
optimization problem which is of interest on its own.
    For each $i \in S$ and $f \in F$, we set $c(i, f) := c(i, f(i))$. For each $g \in \mathbb{R}$, $\lambda \neq 0$,
$i \in S$ and $f \in F$, the risk-sensitive first passage criterion $h_{g,\lambda}(i, f)$ and the corresponding
optimal value function $h_{g,\lambda}^*(\cdot)$ on S are given by

$$h_{g,\lambda}(i, f) := \frac{1}{\lambda} \ln E_i^f \left[e^{\lambda \int_0^{\tau_z} (c(\xi_t, f) - g)dt} \right] \quad \text{and} \quad h_{g,\lambda}^*(i) := \inf_{f \in F} h_{g,\lambda}(i, f), \qquad (3.1)$$

respectively. Let $G_\lambda := \{g \in \mathbb{R} : h_{g,\lambda}^*(z) \leq 0\}$ for all $\lambda > 0$ and $G_\lambda := \{g \in \mathbb{R} : h_{g,\lambda}^*(z) \geq 0\}$ for all $\lambda < 0$. Moreover, we define

$$g_\lambda^* := \begin{cases} \inf G_\lambda, & \text{if } \lambda > 0, \\ \sup G_\lambda, & \text{if } \lambda < 0. \end{cases} \qquad (3.2)$$

Then we have the following assertions on the risk-sensitive first passage criterion.

Theorem 3.2 *Under Assumption 3.1, the following statements hold for all* $\lambda \neq 0$.

(a) *The set* G_λ *is nonempty.*

(b) *For each* $g \in \mathbb{R}$ *and* $f \in F$, *the function* $h_{g,\lambda}(\cdot, f)$ *on* S *satisfies the following equations*

$$
\begin{cases}
e^{\lambda h_{g,\lambda}(i,f)} = Q(i, f, g, \lambda)\left(q(z|i, f) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda h_{g,\lambda}(j,f)} q(j|i, f)\right) \\
e^{\lambda h_{g,\lambda}(z,f)} = Q(z, f, g, \lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda h_{g,\lambda}(j,f)} q(j|z, f)
\end{cases}
$$

for all $i \in S \setminus \{z\}$, *where we set* $Q(i, f, g, \lambda) := \int_0^\infty e^{\lambda(c(i,f)-g)s+q(i|i,f)s} ds$ *and make a convention that* $0 \cdot \infty := 0$.

(c) *For each* $g \in G_\lambda$, *the function* $h_{g,\lambda}^*$ *on* S *satisfies the following equations*

$$
\begin{cases}
e^{\lambda h_{g,\lambda}^*(i)} = sgn(\lambda) \inf_{a \in A(i)} \left\{ sgn(\lambda) Q(i, a, g, \lambda)\left(q(z|i, a) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|i, a)\right)\right\} \\
e^{\lambda h_{g,\lambda}^*(z)} = sgn(\lambda) \inf_{a \in A(z)} \left\{ sgn(\lambda) Q(z, a, g, \lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|z, a)\right\}
\end{cases}
\tag{3.3}
$$

for all $i \in S \setminus \{z\}$, *where we set* $Q(i, a, g, \lambda) := \int_0^\infty e^{\lambda(c(i,a)-g)s+q(i|i,a)s} ds$ *and* $sgn(\lambda)$ *is the sign function, i.e., if* $\lambda > 0$, $sgn(\lambda) = 1$; *if* $\lambda < 0$, $sgn(\lambda) = -1$. *Moreover, there exists a policy* $f_{g,\lambda}^* \in F$ *with* $f_{g,\lambda}^*(i) \in A(i)$ *attaining the minimum of* Eq. 3.3, *and for any* $f_{g,\lambda}^* \in F$ *with* $f_{g,\lambda}^*(i) \in A(i)$ *attaining the minimum of* Eq. 3.3, *we have* $h_{g,\lambda}(i, f_{g,\lambda}^*) = h_{g,\lambda}^*(i) \in \mathbb{R}$ *and* $Q(i, f_{g,\lambda}^*, g, \lambda) < \infty$ *for all* $i \in S$.

(d) *We have* $g_\lambda^* \in G_\lambda$ *and* $h_{g_\lambda^*,\lambda}^*(z) = 0$.

Proof See Section 6. □

Remark 3.3 The equations in Eq. 3.3 are referred to as the optimality equations of the risk-sensitive first passage criterion. The statements of Theorem 3.2 hold for an arbitrary risk-sensitivity parameter $\lambda \neq 0$ and extend the results in Wei and Chen (2016) for any $\lambda > 0$. Moreover, as can be seen in the proof of Theorem 3.2, the treatment of the case $\lambda < 0$ is more difficult than that of the case $\lambda > 0$. Hence, the extension is nontrivial.

Let $B(S)$ be the set of all real-valued functions on S. Below we state the optimality equation and the existence of optimal policies for the risk-sensitive average cost criterion.

Theorem 3.3 *Suppose that Assumption 3.1 is satisfied. For each* $\lambda \neq 0$, *let* g_λ^* *and* $h_{g_\lambda^*,\lambda}^*$ *be as in* Eqs. 3.1 *and* 3.2. *Then the following assertions hold.*

(a) *The pair* $(g_\lambda^*, h_{g_\lambda^*,\lambda}^*) \in \mathbb{R} \times B(S)$ *satisfies the following optimality equation*

$$
\lambda g_\lambda^* e^{\lambda h_{g_\lambda^*,\lambda}^*(i)} = sgn(\lambda) \inf_{a \in A(i)} \left\{ sgn(\lambda)\left(\lambda c(i, a) e^{\lambda h_{g_\lambda^*,\lambda}^*(i)} + \sum_{j \in S} e^{\lambda h_{g_\lambda^*,\lambda}^*(j)} q(j|i, a)\right)\right\}
\tag{3.4}
$$

for all $i \in S$. *Moreover, there exists* $f_\lambda^* \in F$ *with* $f_\lambda^*(i) \in A(i)$ *attaining the minimum of* Eq. 3.4.

(b) *For any* $f_\lambda^* \in F$ *with* $f_\lambda^*(i) \in A(i)$ *attaining the minimum of* Eq. 3.4, *we have*

$$
J_{V_\lambda}^*(i) = J_{V_\lambda}(i, f_\lambda^*) = \lim_{T \to \infty} \frac{1}{\lambda T} \ln E_i^{f_\lambda^*}\left[e^{\lambda \int_0^T c(\xi_t, f_\lambda^*) dt}\right] = g_\lambda^*
$$

for all $i \in S$. *Hence, the policy* f_λ^* *is* V_λ-*average optimal.*

Proof The assertions follow from the above Theorem 3.2, the Feynman-Kac formula and the similar techniques of Theorem 3.2 in Wei and Chen (2016).   □

Remark 3.4 Theorem 3.3 establishes the existence of a solution to the optimality equation and the existence of an optimal stationary policy for the V_λ-average cost criterion with an arbitrary risk-sensitivity parameter $\lambda \neq 0$, which generalizes those in Ghosh and Saha (2014) and Wei and Chen (2016). More precisely, the risk-sensitivity parameter λ is positive in Ghosh and Saha (2014) and Wei and Chen (2016) and satisfies some additional condition that $\lambda \max_{(i,a)\in K} c(i, a) < b$ (for some constant $b > 0$) in Ghosh and Saha (2014). Moreover, to the best of our knowledge, the risk-sensitive average cost criterion with a negative risk-sensitivity parameter has not been studied in the existing literature.

Finally, we give the following statement on the continuity of $J_{V_\lambda}^*(i)$ in $\lambda \in \mathbb{R}$, which plays a crucial role in the study on the existence of U-average optimal policies.

Theorem 3.4 *Suppose that Assumption 3.1 holds. Then for each $i \in S$, $J_{V_\lambda}^*(i)$ is continuous in $\lambda \in \mathbb{R}$.*

Proof See Section 7.   □

4 The existence of U-average optimal policies

In this section, we show the existence of optimal policies for the U-average cost criterion induced by a regular utility function U with the asymptotic risk-sensitivity parameter λ_U.

Below we state the main results on the U-average cost criterion.

Theorem 4.1 *Suppose that Assumption 3.1 is satisfied. Let $g_{\lambda_U}^*$ be as in* Eq. 3.2 *with λ_U in lieu of λ for all $\lambda_U \neq 0$ and $g_0^* := \lim_{\lambda \to 0} g_\lambda^*$. Then the following assertions hold.*

(a) $J_U^*(i) = J_{V_{\lambda_U}}^*(i) = g_{\lambda_U}^*$ *for all $i \in S$.*

(b)   *For each $f \in F$ and $i \in S$, $J_U(i, f) = J_{V_{\lambda_U}}(i, f) = \lim_{T\to\infty} \frac{1}{\lambda_U T} \ln E_i^f \left[e^{\lambda_U \int_0^T c(\xi_t, f)dt} \right]$*

   for all $\lambda_U \neq 0$ and $J_U(i, f) = J_{V_0}(i, f) = \lim_{T\to\infty} \frac{1}{T} E_i^f \left[\int_0^T c(\xi_t, f)dt \right]$ for $\lambda_U = 0$. Moreover, the limits are independent of the state $i \in S$.

(c)   *For any $\lambda_U \neq 0$ and $f_{\lambda_U}^* \in F$ attaining the minimum of* Eq. 3.4 *with λ_U in lieu of λ, we have $J_U^*(i) = J_U(i, f_{\lambda_U}^*) = g_{\lambda_U}^*$ for all $i \in S$. For $\lambda_U = 0$, there exist $(g_0^*, h_0) \in \mathbb{R} \times B(S)$ and a policy $f_0^* \in F$ satisfying*

$$g_0^* = \inf_{a \in A(i)} \left\{ c(i, a) + \sum_{j \in S} h_0(j) q(j|i, a) \right\} \tag{4.1}$$
$$= c(i, f_0^*) + \sum_{j \in S} h_0(j) q(j|i, f_0^*)$$

   for all $i \in S$. Moreover, for any $f_0^* \in F$ attaining the minimum of Eq. 4.1, *we have $J_U^*(i) = J_U^*(i, f_0^*) = g_0^*$ for all $i \in S$. Hence, the policy $f_{\lambda_U}^*$ is U-average optimal.*

Proof (a)   Fix any $i \in S$, $\pi \in \Pi$ and $\eta > 0$. The relation $\lim_{x \to \infty} \mathcal{A}_U(x) = \lambda_U$ implies that there exists a positive constant x_0 satisfying

$$\lambda_U - \eta \leq \mathcal{A}_U(x) \leq \lambda_U + \eta \text{ for all } x > x_0. \tag{4.2}$$

Note that $\min_{(i,a) \in K} c(i, a) > 0$. Thus, there exists a positive constant T^* such that

$$Y := \int_0^T \int_A c(\xi_t, a)\pi(da|\xi_t, t)dt > x_0 \text{ for all } T > T^*.$$

Let $V_{\lambda_U - \eta}$ and $V_{\lambda_U + \eta}$ be the regular utility functions given by Eq. 2.2. Then we have

$$\mathcal{A}_{V_{\lambda_U - \eta}}(x) = \lambda_U - \eta \text{ and } \mathcal{A}_{V_{\lambda_U + \eta}}(x) = \lambda_U + \eta \text{ for all } x \in \mathbb{R}_+. \tag{4.3}$$

Moreover, direct calculations yield that $E_i^\pi[U(Y)]$, $E_i^\pi[V_{\lambda_U - \eta}(Y)]$ and $E_i^\pi[V_{\lambda_U + \eta}(Y)]$ are finite. Thus, employing Theorem 4.1 in Cavazos-Cadena and Hernández-Hernández (2016), Eq. 4.2 and 4.3, we obtain $J_{V_{\lambda_U - \eta}}(i, \pi) \leq J_U(i, \pi) \leq J_{V_{\lambda_U + \eta}}(i, \pi)$, which gives

$$J^*_{V_{\lambda_U - \eta}}(i) \leq J^*_U(i) \leq J^*_{V_{\lambda_U + \eta}}(i). \tag{4.4}$$

Hence, letting $\eta \to 0$ in Eq. 4.4 and using Theorems 3.3(b) and 3.4, we get the desired result.

(b)   Fix any $f \in F$. Let $\mathcal{M}_f$ be the control model in which we take $A(i) = \{f(i)\}$ for all $i \in S$ and the other components are the same as in the model $\mathcal{M}$. Then it is obvious that the model $\mathcal{M}_f$ satisfies Assumption 3.1. Thus, part (b) follows directly from part (a), Lemma 3.1(b) in Guo and Hernández-Lerma (2009) and Theorem 3.3.

(c)   By part (a), Theorem 7.8 in Guo and Hernández-Lerma (2009) and Theorem 3.3 we have $J^*_U(i) = J^*_{V_{\lambda_U}}(i) = J_{V_{\lambda_U}}(i, f^*_{\lambda_U}) = g^*_{\lambda_U}$ for all $i \in S$, which together with part (b) implies the assertion. □

Remark 4.1 (a)   Theorem 4.1 indicates that the optimal value function of the U-average cost criterion induced by a regular utility function U with the asymptotic risk-sensitivity parameter λ_U is a constant and equals the optimal value function of the V_{λ_U}-average cost criterion. Moreover, the set of all U-average optimal stationary policies coincides with the set of all V_{λ_U}-average optimal stationary policies. Hence, we can compute a U-average optimal policy and the optimal value function of the U-average cost criterion via the policy iteration algorithms given in Ghosh and Saha (2014) for the risk-sensitive average cost criterion with the risk-sensitivity parameter $\lambda_U \neq 0$ or in Guo and Hernández-Lerma (2009) for the expected average cost criterion with $\lambda_U = 0$.

(b)   Besides the risk-sensitive average cost criterion induced by the exponential utility function and the expected average cost criterion induced by the identity function, the U-average cost criterion includes other average cost criteria, such as those induced by the logarithmic utility function $W(x) = \ln x$ and the power utility function $U_\beta(x) := x^\beta$ ($\beta > 0$) for all $x \in \mathbb{R}_+$. Obviously, we have that the utility functions W and U_β are regular with the asymptotic risk-sensitivity parameters $\lambda_W = \lambda_{U_\beta} = 0$. Thus, Theorem 4.1 implies $J^*_W(i) = J^*_{U_\beta}(i) = g^*_0$ for all $i \in S$. That is, under Assumption 3.1, the optimal value functions J^*_W and $J^*_{U_\beta}$ are independent of the state variable, and equal the optimal value function of the expected average cost criterion which is risk-neutral.

5 Proof of Theorem 3.1

In this section, we give the proof of Theorem 3.1.

Proof (a)   By Assumption 3.1(iii) and the finiteness of S, for each $f \in F$, the continuous-time Markov chain associated with the transition rate $q(\cdot|\cdot, f)$ has a unique invariant probability measure denoted by μ_f, which satisfies $\sum_{i \in S} q(j|i, f)\mu_f(i) = 0$ for all $j \in S$. Below we show that μ_f is continuous in $f \in F$. In fact, let $\{f_n, n \geq 1\} \subseteq F$ be an arbitrary sequence converging to $f \in F$. Note that $0 \leq \mu_{f_n}(j) \leq 1$ for all $j \in S$ and $n \geq 1$. Fix any $i \in S$ and choose any convergent subsequence $\{\mu_{f_{n_l}}(i), l \geq 1\}$ of $\{\mu_{f_n}(i), n \geq 1\}$. Let $\lim_{l \to \infty} \mu_{f_{n_l}}(i) =: \mu(i)$. Moreover, there exists a subsequence of $\{n_l\}$ (still denoted by $\{n_l\}$) such that $\lim_{l \to \infty} \mu_{f_{n_l}}(j) =: \widetilde{\mu}(j)$ for all $j \in S$ and $\widetilde{\mu}(i) = \mu(i)$. Thus, we have $0 \leq \widetilde{\mu}(j) \leq 1$, $\sum_{j \in S} \widetilde{\mu}(j) = 1$, and

$$\sum_{k \in S} q(j|k, f)\widetilde{\mu}(k) = \lim_{l \to \infty} \sum_{k \in S} q(j|k, f_{n_l})\mu_{f_{n_l}}(k) = 0.$$

Hence, by the uniqueness of the invariant probability measure, we obtain $\widetilde{\mu}(j) = \mu_f(j)$ for all $j \in S$. Therefore, the continuity of μ_f in $f \in F$ follows from the fact that any convergent subsequence $\{\mu_{f_{n_l}}(i), l \geq 1\}$ of $\{\mu_{f_n}(i), n \geq 1\}$ has the same limit $\mu_f(i)$. Set

$$\widetilde{g} := \sup_{f \in F} \frac{E_z^f \left[\int_0^{\tau_z}(1 - I_z(\xi_t))dt \right]}{E_z^f[\tau_z]}.$$

Then direct calculations give

$$\widetilde{g} = \sup_{f \in F} \frac{E_z^f \left[\int_{T_1}^{\tau_z}(1 - I_z(\xi_t))dt \right]}{E_z^f[\tau_z]} = \sup_{f \in F} \frac{E_z^f[\tau_z] - E_z^f[T_1]}{E_z^f[\tau_z]} = \sup_{f \in F} \left(1 - \mu_f(z)\right),$$

where the last equality is due to Proposition 2.1 in Anderson (1991, p.213) and Proposition B.8 in Guo and Hernández-Lerma (2009, p.205). Thus, by the compactness of F and the continuity of $\mu_f(z)$ in $f \in F$, there exists $\widetilde{f} \in F$ such that $\widetilde{g} = 1 - \mu_{\widetilde{f}}(z)$. Moreover, by Assumption 3.1(iii), we have $0 < \mu_{\widetilde{f}}(z) \leq 1$, which implies $0 \leq \widetilde{g} < 1$. Define

$$\widetilde{h}(i) := E_i^{\widetilde{f}} \left[\int_0^{\tau_z}(1 - I_z(\xi_t) - \widetilde{g})dt \right] \quad \text{for all } i \in S.$$

Then we have

$$\widetilde{h}(z) = 0 \text{ and } \widetilde{h}(i) \geq 0 \text{ for all } i \in S \setminus \{z\}. \tag{5.1}$$

Next, we show that $\widetilde{h} \in B(S)$ and

$$\widetilde{g} = \sup_{a \in A(i)} \left\{ 1 - I_z(i) + \sum_{j \in S} \widetilde{h}(j)q(j|i, a) \right\} \quad \text{for all } i \in S. \tag{5.2}$$

Indeed, for each $i \in S \setminus \{z\}$, by the strong Markov property, direct calculations yield

$$
\begin{aligned}
\widetilde{h}(i) &= E_i^{\widetilde{f}}\left[\int_0^{\tau_z}(1 - I_z(\xi_t) - \widetilde{g})dt\, I_{\{\tau_z = T_1\}}\right] + E_i^{\widetilde{f}}\left[\int_0^{\tau_z}(1 - I_z(\xi_t) - \widetilde{g})dt\, I_{\{\tau_z > T_1\}}\right] \\
&= (1 - \widetilde{g})E_i^{\widetilde{f}}[T_1 I_{\{\tau_z = T_1\}}] + (1 - \widetilde{g})E_i^{\widetilde{f}}[T_1 I_{\{\tau_z > T_1\}}] \\
&\quad + E_i^{\widetilde{f}}\left[I_{\{\tau_z > T_1\}}E_i^{\widetilde{f}}\left[\int_{T_1}^{\tau_z}(1 - I_z(\xi_t) - \widetilde{g})dt\,\Big|\mathcal{F}_{T_1}\right]\right] \\
&= (1 - \widetilde{g})E_i^{\widetilde{f}}[T_1] + E_i^{\widetilde{f}}[I_{\{\tau_z > T_1\}}\widetilde{h}(\xi_{T_1})] \\
&= -\frac{1 - \widetilde{g}}{q(i|i, \widetilde{f})} - \sum_{j \in S \setminus \{i, z\}}\frac{\widetilde{h}(j)q(j|i, \widetilde{f})}{q(i|i, \widetilde{f})},
\end{aligned}
\tag{5.3}
$$

where the fourth equality follows from Proposition B.8 in Guo and Hernández-Lerma (2009, p.205). Similarly, we have

$$
\begin{aligned}
\widetilde{h}(z) &= -\widetilde{g}E_z^{\widetilde{f}}[T_1 I_{\{\tau_z > T_1\}}] + E_z^{\widetilde{f}}[I_{\{\tau_z > T_1\}}\widetilde{h}(\xi_{T_1})] \\
&= \frac{\widetilde{g}}{q(z|z, \widetilde{f})} - \sum_{j \in S \setminus \{z\}}\frac{\widetilde{h}(j)q(j|z, \widetilde{f})}{q(z|z, \widetilde{f})}.
\end{aligned}
\tag{5.4}
$$

For any $i \neq z$, Assumption 3.1(iii) implies that there exist different states $j_1 = z$, $j_2, \ldots, j_m = i$ such that $q(j_{n+1}|j_n, \widetilde{f}) > 0$ for all $n = 1, 2, \ldots, m - 1$. Then by Eq. 5.4 we obtain $\widetilde{h}(j_2) < \infty$, which together with Eq. 5.3 and an induction argument gives $\widetilde{h}(i) < \infty$ for all $i \in S$. Hence, we get $\widetilde{h} \in B(S)$. Employing Eqs. 5.1, 5.3 and 5.4, we have

$$
\widetilde{g} = 1 - I_z(i) + \sum_{j \in S}\widetilde{h}(j)q(j|i, \widetilde{f})
\tag{5.5}
$$

$$
\leq \sup_{a \in A(i)}\left\{1 - I_z(i) + \sum_{j \in S}\widetilde{h}(j)q(j|i, a)\right\}
\tag{5.6}
$$

for all $i \in S$. On the other hand, fix any $(k, a) \in K$ and define

$$
\Phi(i) := \begin{cases} \widetilde{g} - 1 + I_z(i) - \sum_{j \in S}\widetilde{h}(j)q(j|i, a), & \text{if } i = k, \\ 0, & \text{otherwise.} \end{cases}
\tag{5.7}
$$

Obviously, we get $\Phi \in B(S)$. Let $\widehat{f} \in F$ be a policy with $\widehat{f}(k) = a$ and $\widehat{f}(i) = \widetilde{f}(i)$ for all $i \in S \setminus \{k\}$. Then by Eqs. 5.5 and 5.7, we obtain

$$
\widetilde{g} = 1 - I_z(i) + \Phi(i) + \sum_{j \in S}\widetilde{h}(j)q(j|i, \widehat{f})
$$

for all $i \in S$. Thus, using the last equality and the Dynkin formula, we have

$$
\widetilde{g}T = E_i^{\widehat{f}}\left[\int_0^T(1 - I_z(\xi_t) + \Phi(\xi_t))dt\right] + E_i^{\widehat{f}}[\widetilde{h}(\xi_T)] - \widetilde{h}(i),
$$

which together with the fact that $\widetilde{h} \in B(S)$ gives

$$
\widetilde{g} = \lim_{T \to \infty}\frac{1}{T}E_i^{\widehat{f}}\left[\int_0^T(1 - I_z(\xi_t) + \Phi(\xi_t))dt\right] = 1 - \mu_{\widehat{f}}(z) + \Phi(k)\mu_{\widehat{f}}(k)
\tag{5.8}
$$

for all $i \in S$. Note that $\mu_{\widehat{f}}(k) > 0$ and $\widetilde{g} \geq 1 - \mu_{\widehat{f}}(z)$. Hence, by Eq. 5.8 we obtain $\Phi(k) \geq 0$. Therefore, we get

$$\widetilde{g} \geq \sup_{a \in A(i)} \left\{ 1 - I_z(i) + \sum_{j \in S} \widetilde{h}(j)q(j|i,a) \right\}$$

for all $i \in S$, which together with Eq. 5.6 implies Eq. 5.2. Fix any $i \in S$ and $f \in F$ below. By Eqs. 5.1 and 5.2 we obtain

$$\widetilde{h}(i) \geq -\frac{1 - I_z(i) - \widetilde{g}}{q(i|i,f)} - \sum_{j \in S \setminus \{i,z\}} \frac{\widetilde{h}(j)q(j|i,f)}{q(i|i,f)}.$$

Then employing the last inequality, we have

$$\widetilde{h}(\xi_{T_m}) \geq E_i^f \left[\int_{T_m}^{T_{m+1}} (1 - I_z(\xi_t) - \widetilde{g})dt \Big| \mathcal{F}_{T_m} \right] + E_i^f \left[\widetilde{h}(\xi_{T_{m+1}}) I_{\{\tau_z > T_{m+1}\}} \Big| \mathcal{F}_{T_m} \right]$$

(5.9)

for all $m = 0, 1, \dots$. Thus, using Eq. 5.9 and an induction argument, we get

$$\widetilde{h}(i) \geq -\frac{1 - I_z(i) - \widetilde{g}}{q(i|i,f)} + (1 - \widetilde{g}) \sum_{l=1}^{n} E_i^f [\theta_{l+1} I_{\{\tau_z > T_l\}}] + E_i^f \left[\widetilde{h}(\xi_{T_{n+1}}) I_{\{\tau_z > T_{n+1}\}} \right]$$

for all $n = 1, 2, \dots$, which together with Eq. 5.1 yields

$$\widetilde{h}(i) + \frac{1 - I_z(i) - \widetilde{g}}{q(i|i,f)} \geq (1 - \widetilde{g}) \sum_{l=1}^{\infty} E_i^f [\theta_{l+1} I_{\{\tau_z > T_l\}}].$$

Hence, by the last inequality we obtain

$$\sum_{l=2}^{\infty} E_i^f [\theta_l I_{\{\tau_z \geq T_l\}}] \leq \frac{1}{1 - \widetilde{g}} \left[\max_{i \in S} \widetilde{h}(i) - \min_{(i,a) \in K} \frac{\widetilde{g}}{q(i|i,a)} \right] := L_1.$$

(5.10)

Observe that

$$E_i^f [\tau_z] = \sum_{l=1}^{\infty} E_i^f [\theta_l I_{\{\tau_z \geq T_l\}}] = -\frac{1}{q(i|i,f)} + \sum_{l=2}^{\infty} E_i^f [\theta_l I_{\{\tau_z \geq T_l\}}],$$

which together with Eq. 5.10 implies $E_i^f [\tau_z] \leq L_1 - \min_{(i,a) \in K} \frac{1}{q(i|i,a)}$. Therefore, the assertion holds with $L := L_1 - \min_{(i,a) \in K} \frac{1}{q(i|i,a)}$.

(b)   By part (a) we have

$$P_i^f (\tau_z > t) \leq \frac{E_i^f [\tau_z]}{t} \leq \frac{L}{t}$$

for all $i \in S$, $f \in F$ and $t > 0$. Moreover, there exists $t_0 \in \mathbb{R}_+$ such that $\frac{L}{t_0} \in (0, 1)$. Hence, part (b) holds with $\alpha := \frac{L}{t_0}$.   □

6 Proof of Theorem 3.2

In this section, we present the proof of Theorem 3.2.

Proof The statements for the case $\lambda > 0$ follow from Theorem 3.1 in Wei and Chen (2016). Below we only need to prove the case $\lambda < 0$.

(a) Set $\widetilde{M} := \min_{(i,a) \in K} c(i,a)$. Thus, we obtain $h_{\widetilde{M}, \lambda}(i, f) \geq 0$ for all $i \in S$ and $f \in F$, which gives $h^*_{\widetilde{M}, \lambda}(z) \geq 0$. Therefore, the set G_λ is nonempty.

(b) From the proof of Theorem 3.1(b) in Wei and Chen (2016), we see that part (b) also holds for the case $\lambda < 0$.

(c) Fix any $g \in G_\lambda$. Set $\overline{c} := \max_{(i,a) \in K} |c(i,a) - g|$. For each $i \in S \setminus \{z\}$, $f \in F$ and $m \geq 1$, direct calculations yield

$$
\begin{aligned}
e^{\lambda h_{g,\lambda}(z,f)} &\geq E_z^f \left[e^{\lambda \int_0^{\tau_z} (c(\xi_t, f) - g) dt} I_{\{\xi_{T_l} \neq i, z, 1 \leq l \leq m-1, \xi_{T_m} = i\}} \right] \\
&= E_z^f \left[I_{\{\xi_{T_l} \neq i, z, 1 \leq l \leq m-1, \xi_{T_m} = i\}} e^{\lambda \int_0^{T_m} (c(\xi_t, f) - g) dt} E_z^f \left[e^{\lambda \int_{T_m}^{\tau_z} (c(\xi_t, f) - g) dt} \Big| \mathcal{F}_{T_m} \right] \right] \\
&= e^{\lambda h_{g,\lambda}(i,f)} E_z^f \left[I_{\{\xi_{T_l} \neq i, z, 1 \leq l \leq m-1, \xi_{T_m} = i\}} e^{\lambda \int_0^{T_m} (c(\xi_t, f) - g) dt} \right] \\
&\geq e^{\lambda h_{g,\lambda}(i,f)} E_z^f \left[I_{\{\xi_{T_l} \neq i, z, 1 \leq l \leq m-1, \xi_{T_m} = i\}} e^{\lambda \overline{c} T_m} \right],
\end{aligned}
$$

which together with Eq. 3.1 and the definition of G_λ gives

$$
e^{\lambda h_{g,\lambda}(i,f)} E_z^f \left[I_{\{\xi_{T_l} \neq i, z, 1 \leq l \leq m-1, \xi_{T_m} = i\}} e^{\lambda \overline{c} T_m} \right] \leq e^{\lambda h^*_{g,\lambda}(z)} \leq 1. \tag{6.1}
$$

Suppose that $\sup_{f \in F} e^{\lambda h_{g,\lambda}(i,f)} = \infty$. Then there exists a sequence $\{f_n, n \geq 1\} \subseteq F$ such that $e^{\lambda h_{g,\lambda}(i,f_n)} \geq n$ for all $n \geq 1$. Thus, by Eq. 6.1 we obtain

$$
\lim_{n \to \infty} E_z^{f_n} \left[I_{\{\xi_{T_l} \neq i, z, 1 \leq l \leq m-1, \xi_{T_m} = i\}} e^{\lambda \overline{c} T_m} \right] = 0 \tag{6.2}
$$

for all $m \geq 1$. Because F is compact, there exists a subsequence of $\{f_n, n \geq 1\}$ (denoted by the same sequence) such that f_n converges to some $\overline{f} \in F$, i.e.,

$$
f_n(j) \to \overline{f}(j) \text{ for all } j \in S \text{ as } n \to \infty. \tag{6.3}
$$

Moreover, for each $m \geq 1$, by Proposition B.8 in Guo and Hernández-Lerma (2009, p.205) we have

$$
\begin{aligned}
&E_z^{f_n} \left[I_{\{\xi_{T_l} \neq i, z, 1 \leq l \leq m-1, \xi_{T_m} = i\}} e^{\lambda \overline{c} T_m} \right] \\
&= \sum_{j_1 \in S \setminus \{i\}, j_{l+1} \in S \setminus \{j_l, i, z\}, l = 1, 2, \ldots, m-2} \left(-\frac{q(j_1 | z, f_n)}{q(z | z, f_n) + \lambda \overline{c}} \right) \prod_{l=1}^{m-2} \left(-\frac{q(j_{l+1} | j_l, f_n)}{q(j_l | j_l, f_n) + \lambda \overline{c}} \right) \\
&\quad \times \left(-\frac{q(i | j_{m-1}, f_n)}{q(j_{m-1} | j_{m-1}, f_n) + \lambda \overline{c}} \right)
\end{aligned}
$$

for all $n \geq 1$, which together with Assumption 3.1(ii), Eqs. 6.2 and 6.3 implies

$$
E_z^{\overline{f}} \left[I_{\{\xi_{T_l} \neq i, z, 1 \leq l \leq m-1, \xi_{T_m} = i\}} e^{\lambda \overline{c} T_m} \right] = 0. \tag{6.4}
$$

On the other hand, Assumption 3.1(iii) gives that there exist different states $k_0 = z$, $k_1, \ldots, k_{\widetilde{m}} = i$ such that $q(k_{n+1} | k_n, \overline{f}) > 0$ for all $n = 0, 2, \ldots, \widetilde{m} - 1$. Thus, we get

$$
E_z^{\overline{f}} \left[I_{\{\xi_{T_l} \neq i, z, 1 \leq l \leq \widetilde{m}-1, \xi_{T_{\widetilde{m}}} = i\}} e^{\lambda \overline{c} T_{\widetilde{m}}} \right] \geq \prod_{l=0}^{\widetilde{m}-1} \left(-\frac{q(k_{l+1} | k_l, \overline{f})}{q(k_l | k_l, \overline{f}) + \lambda \overline{c}} \right) > 0,
$$

which contradicts (6.4). Hence, we obtain

$$e^{\lambda h_{g,\lambda}^*(i)} = \sup_{f \in F} e^{\lambda h_{g,\lambda}(i,f)} < \infty \quad \text{for all } i \in S. \tag{6.5}$$

By Eq. 3.1 and part (b), we have

$$\begin{cases} e^{\lambda h_{g,\lambda}^*(i)} \geq Q(i,f,g,\lambda)\Big(q(z|i,f) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda h_{g,\lambda}(j,f)} q(j|i,f)\Big) \\ e^{\lambda h_{g,\lambda}^*(z)} \geq Q(z,f,g,\lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda h_{g,\lambda}(j,f)} q(j|z,f) \end{cases} \tag{6.6}$$

for all $i \in S \setminus \{z\}$ and $f \in F$. Note that Theorem 3.1 implies

$$e^{\lambda h_{g,\lambda}(i,f)} \geq E_i^f\left[e^{\lambda \max_{(i,a) \in K} |c(i,a)-g|\tau_z}\right] > 0 \tag{6.7}$$

for all $i \in S$ and $f \in F$. Thus, employing Eqs. 6.5–6.7 and Assumption 3.1(iii), we get

$$e^{\lambda h_{g,\lambda}(i,f)} < \infty \quad \text{and} \quad Q(i,f,g,\lambda) < \infty \quad \text{for all } i \in S \text{ and } f \in F. \tag{6.8}$$

Moreover, using Eq. 3.1 and part (b) again, we obtain

$$\begin{cases} e^{\lambda h_{g,\lambda}^*(i)} \leq \sup_{a \in A(i)} \Big\{ Q(i,a,g,\lambda)\Big(q(z|i,a) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|i,a)\Big)\Big\} \\ e^{\lambda h_{g,\lambda}^*(z)} \leq \sup_{a \in A(z)} \Big\{ Q(z,a,g,\lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|z,a)\Big\} \end{cases} \tag{6.9}$$

for all $i \in S \setminus \{z\}$. By Theorem 3.1(c) in Wei and Chen (2016), Assumption 3.1(ii) and Eq. 6.5, we see that $Q(i,a,g,\lambda)(q(z|i,a) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|i,a))$ and $Q(z,a,g,\lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|z,a)$ in Eq. 6.9 are continuous in $a \in A(i)$ and $a \in A(z)$, respectively. Thus, the Weierstrass theorem in Aliprantis and Border (2007, p.40) and Assumption 3.1 imply that there exists $f_{g,\lambda}^* \in F$ with $f_{g,\lambda}^*(i) \in A(i)$ attaining the maximum of Eq. 6.9, i.e.,

$$\begin{cases} e^{\lambda h_{g,\lambda}^*(i)} \leq Q(i,f_{g,\lambda}^*,g,\lambda)\Big(q(z|i,f_{g,\lambda}^*) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|i,f_{g,\lambda}^*)\Big) \\ e^{\lambda h_{g,\lambda}^*(z)} \leq Q(z,f_{g,\lambda}^*,g,\lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|z,f_{g,\lambda}^*) \end{cases} \tag{6.10}$$

for all $i \in S \setminus \{z\}$. Hence, by Eq. 6.10, Proposition B.8 in Guo and Hernández-Lerma (2009, p.205) and an induction argument, we have

$$e^{\lambda h_{g,\lambda}^*(i)} \leq \sum_{m=1}^n E_i^{f_{g,\lambda}^*}\left[e^{\lambda \int_0^{T_m}\left(c(\xi_t,f_{g,\lambda}^*)-g\right)dt} I_{\{\tau_z=T_m\}}\right]$$
$$+ E_i^{f_{g,\lambda}^*}\left[e^{\lambda \int_0^{T_n}\left(c(\xi_t,f_{g,\lambda}^*)-g\right)dt} e^{\lambda h_{g,\lambda}^*(\xi_{T_n})} I_{\{\tau_z>T_n\}}\right] \tag{6.11}$$

for all $i \in S \setminus \{z\}$ and $n = 1, 2, \ldots$. Furthermore, it follows from part (b) and the similar arguments of Eq. 6.11 that

$$e^{\lambda h_{g,\lambda}(i,f_{g,\lambda}^*)} = \sum_{m=1}^n E_i^{f_{g,\lambda}^*}\left[e^{\lambda \int_0^{T_m}\left(c(\xi_t,f_{g,\lambda}^*)-g\right)dt} I_{\{\tau_z=T_m\}}\right]$$
$$+ E_i^{f_{g,\lambda}^*}\left[e^{\lambda \int_0^{T_n}\left(c(\xi_t,f_{g,\lambda}^*)-g\right)dt} e^{\lambda h_{g,\lambda}(\xi_{T_n},f_{g,\lambda}^*)} I_{\{\tau_z>T_n\}}\right]$$

for all $i \in S \setminus \{z\}$ and $n = 1, 2, \ldots$. Thus, employing the last equality and the fact that $\min_{i \in S} e^{\lambda h_{g,\lambda}(i, f_{g,\lambda}^*)} > 0$, we obtain

$$\liminf_{n \to \infty} E_i^{f_{g,\lambda}^*} \left[e^{\lambda \int_0^{T_n} \left(c(\xi_t, f_{g,\lambda}^*) - g \right) dt} I_{\{\tau_z > T_n\}} \right] = 0 \tag{6.12}$$

for all $i \in S \setminus \{z\}$. Hence, using Eqs. 6.11 and 6.12, we get

$$e^{\lambda h_{g,\lambda}^*(i)} \le e^{\lambda h_{g,\lambda}(i, f_{g,\lambda}^*)} + \max_{j \in S} e^{\lambda h_{g,\lambda}^*(j)} \liminf_{n \to \infty} E_i^{f_{g,\lambda}^*} \left[e^{\lambda \int_0^{T_n} \left(c(\xi_t, f_{g,\lambda}^*) - g \right) dt} I_{\{\tau_z > T_n\}} \right] = e^{\lambda h_{g,\lambda}(i, f_{g,\lambda}^*)},$$

which together with Eq. 3.1 implies

$$h_{g,\lambda}^*(i) = h_{g,\lambda}(i, f_{g,\lambda}^*) \text{ for all } i \in S \setminus \{z\}. \tag{6.13}$$

Thus, using Eqs. 6.6 and 6.13, we get

$$e^{\lambda h_{g,\lambda}^*(i)} \ge Q(i, f_{g,\lambda}^*, g, \lambda) \left(q(z|i, f_{g,\lambda}^*) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|i, f_{g,\lambda}^*) \right)$$

$$= \sup_{a \in A(i)} \left\{ Q(i, a, g, \lambda) \left(q(z|i, a) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|i, a) \right) \right\},$$

which together with Eq. 6.9 yields

$$e^{\lambda h_{g,\lambda}^*(i)} = \sup_{a \in A(i)} \left\{ Q(i, a, g, \lambda) \left(q(z|i, a) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|i, a) \right) \right\} \tag{6.14}$$

for all $i \in S \setminus \{z\}$. Following the similar arguments of Eqs. 6.13 and 6.14, we have

$$h_{g,\lambda}^*(z) = h_{g,\lambda}(z, f_{g,\lambda}^*) \text{ and } e^{\lambda h_{g,\lambda}^*(z)} = \sup_{a \in A(z)} \left\{ Q(z, a, g, \lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda h_{g,\lambda}^*(j)} q(j|z, a) \right\}. \tag{6.15}$$

Therefore, the function $h_{g,\lambda}^*$ on S satisfies (3.3). Moreover, Eq. 6.7 gives

$$e^{\lambda h_{g,\lambda}^*(i)} = e^{\lambda h_{g,\lambda}(i, f_{g,\lambda}^*)} > 0 \text{ for all } i \in S. \tag{6.16}$$

Hence, by Eqs. 6.5, 6.8, 6.13, 6.15 and 6.16, we see that for any $f_{g,\lambda}^* \in F$ with $f_{g,\lambda}^*(i) \in A(i)$ attaining the minimum of Eq. 3.3, $h_{g,\lambda}(i, f_{g,\lambda}^*) = h_{g,\lambda}^*(i) \in \mathbb{R}$ and $Q(i, f_{g,\lambda}^*, g, \lambda) < \infty$ for all $i \in S$.

(d)   Choose a sequence $\{\overline{g}_n^\lambda, n \ge 1\} \subseteq G_\lambda$ satisfying

$$\overline{g}_n^\lambda \le \overline{g}_{n+1}^\lambda \text{ for all } n \ge 1 \text{ and } \lim_{n \to \infty} \overline{g}_n^\lambda = g_\lambda^*. \tag{6.17}$$

By Eqs. 3.1 and 6.17, we obtain

$$h_{\overline{g}_n^\lambda, \lambda}^*(i) \ge h_{\overline{g}_{n+1}^\lambda, \lambda}^*(i) \ge h_{g_\lambda^*, \lambda}^*(i) \text{ for all } i \in S \text{ and } n \ge 1. \tag{6.18}$$

Set $H_\lambda(i) := \lim_{n \to \infty} h_{\overline{g}_n^\lambda, \lambda}^*(i)$ for all $i \in S$. Then by the definition of G_λ and Eq. 6.18, we have

$$H_\lambda(z) \ge 0 \text{ and } H_\lambda(i) \ge h_{g_\lambda^*, \lambda}^*(i) \text{ for all } i \in S. \tag{6.19}$$

Moreover, employing part (c), we get

$$
\begin{cases}
e^{\lambda h^*_{\overline{g}^\lambda_n,\lambda}(i)} \geq Q(i, f, \overline{g}^\lambda_n, \lambda) \left(q(z|i, f) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda h^*_{\overline{g}^\lambda_n,\lambda}(j)} q(j|i, f) \right) \\
e^{\lambda h^*_{\overline{g}^\lambda_n,\lambda}(z)} \geq Q(z, f, \overline{g}^\lambda_n, \lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda h^*_{\overline{g}^\lambda_n,\lambda}(j)} q(j|z, f)
\end{cases}
$$

for all $n \geq 1$, which together with the Fatou lemma gives

$$
\begin{cases}
e^{\lambda H_\lambda(i)} \geq Q(i, f, g^*_\lambda, \lambda) \left(q(z|i, f) + \sum_{j \in S \setminus \{i,z\}} e^{\lambda H_\lambda(j)} q(j|i, f) \right) \\
e^{\lambda H_\lambda(z)} \geq Q(z, f, g^*_\lambda, \lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda H_\lambda(j)} q(j|z, f)
\end{cases}
\tag{6.20}
$$

for all $i \in S \setminus \{z\}$ and $f \in F$. Thus, using Eq. 6.20 and the similar arguments of Eq. 6.11, we obtain $e^{\lambda H_\lambda(i)} \geq e^{\lambda h^*_{g^*_\lambda,\lambda}(i,f)}$ for all $f \in F$, which together with Eqs. 3.1 and 6.19 yields $h^*_{g^*_\lambda,\lambda}(z) = H_\lambda(z) \geq 0$. Hence, we obtain $g^*_\lambda \in G_\lambda$.

Suppose that $h^*_{g^*_\lambda,\lambda}(z) > 0$. For each $n \geq 1$, let $\gamma_n := e^{n\lambda h^*_{g^*_\lambda,\lambda}(z)}$. Then we have $\gamma_n \in (0, 1)$ for all $n \geq 1$ and $\lim_{n \to \infty} \gamma_n = 0$. Employing Eq. 6.8 we get $\lambda c(i, f) - \lambda g^*_\lambda + q(i|i, f) < 0$ for all $i \in S$ and $f \in F$. For each $n \geq 1$ and $f \in F$, define the transition rate as follows:

$$
\overline{p}_n(z|z, f) := -\gamma_n e^{\lambda h^*_{g^*_\lambda,\lambda}(z,f)}, \quad \overline{p}_n(j|z, f) := -\frac{\gamma_n e^{\lambda h^*_{g^*_\lambda,\lambda}(j,f)} q(j|z, f)}{\lambda c(z, f) - \lambda g^*_\lambda + q(z|z, f)} \quad \text{for } j \in S \setminus \{z\},
\tag{6.21}
$$

and for any $i \in S \setminus \{z\}$,

$$
\overline{p}_n(i|i, f) := -\gamma_n e^{\lambda h^*_{g^*_\lambda,\lambda}(i,f)}, \quad \overline{p}_n(z|i, f) := -\frac{\gamma_n q(z|i, f)}{\lambda c(i, f) - \lambda g^*_\lambda + q(i|i, f)}, \tag{6.22}
$$

$$
\overline{p}_n(j|i, f) := -\frac{\gamma_n e^{\lambda h^*_{g^*_\lambda,\lambda}(j,f)} q(j|i, f)}{\lambda c(i, f) - \lambda g^*_\lambda + q(i|i, f)} \quad \text{for } j \in S \setminus \{i, z\}. \tag{6.23}
$$

For any initial state $i \in S$ and any policy $f \in F$, the probability measure and expectation operator corresponding to the transition rate $\overline{p}_n(\cdot|\cdot, f)$ defined in Eqs. 6.21–6.23 are denoted by $\overline{P}^f_{i,n}$ and $\overline{E}^f_{i,n}$. For any $\kappa > 0$ and $n \geq 1$, define

$$
\overline{H}_{\kappa,n}(i, f) := \frac{1}{\lambda} \ln \overline{E}^f_{i,n} \left[e^{-\lambda \kappa \tau_z} \right] \quad \text{for all } i \in S \text{ and } f \in F.
$$

Note that Eq. 6.7 gives $e^{\lambda h^*_{g^*_\lambda,\lambda}(i,f)} > 0$ for all $i \in S$ and $f \in F$. Employing Eq. 3.1 and part (c), we have $\min_{i \in S} \inf_{f \in F} e^{-\lambda h^*_{g^*_\lambda,\lambda}(i,f)} = \min_{i \in S} e^{-\lambda h^*_{g^*_\lambda,\lambda}(i)} > 0$. Then by the fact that $\lim_{n \to \infty} \gamma_n = 0$, there exists a positive integer n_1 such that

$$
\gamma_{n_1} < \min_{(i,a) \in K} \left(\lambda g^*_\lambda - \lambda c(i, a) - q(i|i, a) \right) \quad \text{and} \tag{6.24}
$$

$$
\gamma_{n_1} \leq \min_{(i,a) \in K} \left(\lambda g^*_\lambda - \lambda c(i, a) - q(i|i, a) \right) \times \min_{i \in S} \inf_{f \in F} e^{-\lambda h^*_{g^*_\lambda,\lambda}(i,f)}
$$

$$
\leq \left(\lambda g^*_\lambda - \lambda c(i, f) - q(i|i, f) \right) e^{-\lambda h^*_{g^*_\lambda,\lambda}(i,f)} \tag{6.25}
$$

for all $i \in S$ and $f \in F$. Moreover, we have that there exists $f^1 \in F$ satisfying $h_{g_\lambda^*, \lambda}(i, f^1) = \sup_{f \in F} h_{g_\lambda^*, \lambda}(i, f) =: \overline{h}_{g_\lambda^*, \lambda}(i)$ for all $i \in S$. In fact, by part (b) we get

$$
\begin{cases}
e^{\lambda \overline{h}_{g_\lambda^*, \lambda}(i)} \geq \inf_{a \in A(i)} \left\{ Q(i, a, g, \lambda) \left(q(z|i, a) + \sum_{j \in S \setminus \{i, z\}} e^{\lambda \overline{h}_{g_\lambda^*, \lambda}(j)} q(j|i, a) \right) \right\} \\
e^{\lambda \overline{h}_{g_\lambda^*, \lambda}(z)} \geq \inf_{a \in A(z)} \left\{ Q(z, a, g, \lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda \overline{h}_{g_\lambda^*, \lambda}(j)} q(j|z, a) \right\}
\end{cases} \tag{6.26}
$$

for all $i \in S \setminus \{z\}$. Then employing Assumption 3.1, Theorem 3.1(c) in Wei and Chen (2016) and the Weierstrass theorem in Aliprantis and Border (2007, p.40), we obtain the existence of $f^1 \in F$ with $f^1(i) \in A(i)$ attaining the minimum of Eq. 6.26, i.e,

$$
\begin{cases}
e^{\lambda \overline{h}_{g_\lambda^*, \lambda}(i)} \geq Q(i, f^1, g, \lambda) \left(q(z|i, f^1) + \sum_{j \in S \setminus \{i, z\}} e^{\lambda \overline{h}_{g_\lambda^*, \lambda}(j)} q(j|i, f^1) \right) \\
e^{\lambda \overline{h}_{g_\lambda^*, \lambda}(z)} \geq Q(z, f^1, g, \lambda) \sum_{j \in S \setminus \{z\}} e^{\lambda \overline{h}_{g_\lambda^*, \lambda}(j)} q(j|z, f^1)
\end{cases} \tag{6.27}
$$

for all $i \in S \setminus \{z\}$. Thus, using Eq. 6.27 and following the similar arguments of Eq. 6.11, we have

$$
e^{\lambda \overline{h}_{g_\lambda^*, \lambda}(i)} \geq \sum_{m=1}^{n} E_i^{f^1} \left[e^{\lambda \int_0^{T_m} (c(\xi_t, f^1) - g_\lambda^*) dt} I_{\{\tau_z = T_m\}} \right] + E_i^{f^1} \left[e^{\lambda \int_0^{T_n} (c(\xi_t, f^1) - g_\lambda^*) dt} e^{\lambda \overline{h}_{g_\lambda^*, \lambda}(\xi_{T_n})} I_{\{\tau_z > T_n\}} \right]
$$

for all $n = 1, 2, \ldots$, which gives $\overline{h}_{g_\lambda^*, \lambda}(i) \leq h_{g_\lambda^*, \lambda}(i, f^1)$ for all $i \in S \setminus \{z\}$. Hence, we get $\overline{h}_{g_\lambda^*, \lambda}(i) = h_{g_\lambda^*, \lambda}(i, f^1)$ for all $i \in S \setminus \{z\}$. Similarly, by Eq. 6.27 we can obtain $\overline{h}_{g_\lambda^*, \lambda}(z) = h_{g_\lambda^*, \lambda}(z, f^1)$. Furthermore, by Eq. 6.7 we have

$$
\inf_{f \in F} e^{\lambda h_{g_\lambda^*, \lambda}(i, f)} = e^{\lambda \sup_{f \in F} h_{g_\lambda^*, \lambda}(i, f)} = e^{\lambda h_{g_\lambda^*, \lambda}(i, f^1)} > 0, \tag{6.28}
$$

which implies $\zeta(i) := \inf_{f \in F} \frac{1}{\lambda} \overline{p}_{n_1}(i|i, f) > 0$ for all $i \in S$. Hence, for any $\kappa \in \left(0, \min_{i \in S} \zeta(i) \right) =: \overline{O}_{n_1}$, by Eqs. 6.21–6.23 and the similar arguments of Theorem 3.1(b) in Wei and Chen (2016), we get

$$
\begin{cases}
e^{\lambda \overline{H}_{\kappa, n_1}(i, f)} = -\dfrac{1}{\gamma_{n_1} e^{\lambda h_{g_\lambda^*, \lambda}(i, f)} + \lambda \kappa} \left(\dfrac{\gamma_{n_1} q(z|i, f)}{\lambda c(i, f) - \lambda g_\lambda^* + q(i|i, f)} + \sum_{j \in S \setminus \{i, z\}} \dfrac{\gamma_{n_1} e^{\lambda \overline{H}_{\kappa, n_1}(j, f) + \lambda h_{g_\lambda^*, \lambda}(j, f)}}{\lambda c(i, f) - \lambda g_\lambda^* + q(i|i, f)} q(j|i, f) \right) \\
e^{\lambda \overline{H}_{\kappa, n_1}(z, f)} = -\dfrac{1}{\gamma_{n_1} e^{\lambda h_{g_\lambda^*, \lambda}(z, f)} + \lambda \kappa} \sum_{j \in S \setminus \{z\}} \dfrac{\gamma_{n_1} e^{\lambda \overline{H}_{\kappa, n_1}(j, f) + \lambda h_{g_\lambda^*, \lambda}(j, f)}}{\lambda c(z, f) - \lambda g_\lambda^* + q(z|z, f)} q(j|z, f)
\end{cases} \tag{6.29}
$$

for all $i \in S \setminus \{z\}$ and $f \in F$. Below we show that there exist $\widetilde{t_0} \in \mathbb{R}_+$ and $\widetilde{\alpha} \in (0, 1)$ such that

$$
\overline{P}_{i, n_1}^f (\tau_z > \widetilde{t_0}) \leq \widetilde{\alpha} \quad \text{for all } i \in S \text{ and } f \in F. \tag{6.30}
$$

By Theorem 3.1 and an induction argument, we have $P_i^f(\tau_z > nt_0) \le \alpha^n$ for all $n = 1, 2, \ldots$, which implies that there exists a positive integer n_2 such that

$$0 < \alpha^{n_2} \le \frac{1}{3} \text{ and } P_i^f(\tau_z > n_2 t_0) \le \alpha^{n_2} \tag{6.31}$$

for all $i \in S$ and $f \in F$. Let $\varpi_1 := \max_{(i,a) \in K}(-q(i|i,a))$. Then direct calculations yield

$$
\begin{aligned}
P_i^f(T_m < n_2 t_0) &= P_i^f(\theta_1 + \cdots + \theta_m < n_2 t_0) \\
&= \int_{s_1 + \cdots + s_m < n_2 t_0} \sum_{j_1 \in S \setminus \{i\}, j_{l+1} \in S \setminus \{j_l\}, l = 1, \ldots, m-1} e^{q(i|i,f)s_1} q(j_1|i,f) \\
&\quad \times \prod_{l=1}^{m-1} \left(e^{q(j_l|j_l,f)s_{l+1}} q(j_{l+1}|j_l, f) \right) ds_m \cdots ds_1 \\
&\le \varpi_1^m \int_{s_1 + \cdots + s_m < n_2 t_0} ds_m \cdots ds_1 = \frac{(\varpi_1 n_2 t_0)^m}{m!} \tag{6.32}
\end{aligned}
$$

for all $i \in S$ and $f \in F$, where the second equality follows from Proposition B.8 in Guo and Hernández-Lerma (2009, p.205). Thus, using Eq. 6.32 we obtain $\lim_{m \to \infty} \sup_{i \in S, f \in F} P_i^f(T_m < n_2 t_0) = 0$, which gives that there exists a positive integer m^* satisfying

$$P_i^f(T_{m^*} < n_2 t_0) \le \frac{\alpha^{n_2}}{2} \text{ for all } i \in S \text{ and } f \in F. \tag{6.33}$$

For each $f \in F$, let $\{Y_n, n = 0, 1, \ldots\}$ be the embedding Markov chain of the continuous-time Markov chain associated with the transition rate $q(\cdot|\cdot, f)$ and define $\tau_1 := \inf\{n \ge 1 : Y_n = z\}$. Employing Eqs. 6.31 and 6.33 we get

$$
\begin{aligned}
P_i^f(\tau_1 \le m^*) = P_i^f(\tau_z \le T_{m^*}) &\ge P_i^f(\tau_z \le n_2 t_0, T_{m^*} \ge n_2 t_0) \\
&\ge P_i^f(\tau_z \le n_2 t_0) + P_i^f(T_{m^*} \ge n_2 t_0) - 1 \ge \frac{1}{2} \tag{6.34}
\end{aligned}
$$

for all $i \in S$ and $f \in F$. Let $\varpi_2 := \min_{(i,a) \in K}\left(-\frac{\gamma_{n_1}}{\lambda c(i,a) - \lambda g_\lambda^* + q(i|i,a)}\right)$, $\varpi_3 := \min_{i \in S} \inf_{f \in F} e^{\lambda h_{g_\lambda^*, \lambda}(i,f)}$, and $\varpi_4 := \min_{i \in S} \inf_{f \in F} \overline{p}_{n_1}(i|i, f)$. Then Eq. 6.24 and Theorem 3.2(c) imply $\varpi_2 \in (0, 1)$ and $\varpi_4 \in (-\infty, 0)$. By Eqs. 3.1, 6.28 and the fact that $g_\lambda^* \in G_\lambda$, we obtain

$$\varpi_3 = \min_{i \in S} e^{\lambda h_{g_\lambda^*, \lambda}(i, f^1)} > 0 \text{ and } \varpi_3 \le e^{\lambda h_{g_\lambda^*, \lambda}^*(z)} \le 1.$$

Thus, for each $i \in S$, $f \in F$ and $m \geq 1$, direct calculations yield

$$
\begin{aligned}
&\overline{P}^f_{i,n_1}(\tau_z \leq n_2 t_0) \\
&\geq \overline{P}^f_{i,n_1}(\tau_z \leq n_2 t_0, \tau_z \leq T_m) \\
&= \sum_{n=1}^m \overline{P}^f_{i,n_1}(\tau_z \leq n_2 t_0, \tau_z = T_n) \\
&= \sum_{n=1}^m \int_{s_1+\cdots+s_n \leq n_2 t_0} \sum_{j_1 \in S\backslash\{i,z\}, j_{l+1} \in S\backslash\{j_l,z\}, l=1,2,\ldots,n-2} e^{\overline{p}_{n_1}(i|i,f)s_1} \overline{p}_{n_1}(j_1|i,f) \\
&\qquad \times \prod_{l=1}^{n-2} e^{\overline{p}_{n_1}(j_l|j_l,f)s_{l+1}} \overline{p}_{n_1}(j_{l+1}|j_l,f) e^{\overline{p}_{n_1}(j_{n-1}|j_{n-1},f)s_n} \overline{p}_{n_1}(z|j_{n-1},f) ds_n \cdots ds_1 \\
&\geq \sum_{n=1}^m \int_{s_1+\cdots+s_n \leq n_2 t_0} \sum_{j_1 \in S\backslash\{i,z\}, j_{l+1} \in S\backslash\{j_l,z\}, l=1,2,\ldots,n-2} \varpi_2 \varpi_3 e^{\varpi_4 s_1} e^{q(i|i,f)s_1} q(j_1|i,f) \\
&\qquad \times \prod_{l=1}^{n-2} \varpi_2 \varpi_3 e^{\varpi_4 s_{l+1}} e^{q(j_l|j_l,f)s_{l+1}} q(j_{l+1}|j_l,f) \varpi_2 e^{\varpi_4 s_n} e^{q(j_{n-1}|j_{n-1},f)s_n} q(z|j_{n-1},f) ds_n \cdots ds_1 \\
&= \sum_{n=1}^m (\varpi_2 \varpi_3)^{n-1} \varpi_2 \int_{s_1+\cdots+s_n \leq n_2 t_0} \sum_{j_1 \in S\backslash\{i,z\}, j_{l+1} \in S\backslash\{j_l,z\}, l=1,2,\ldots,n-2} e^{\varpi_4(s_1+\cdots+s_n)} \\
&\qquad \times e^{q(i|i,f)s_1} q(j_1|i,f) \prod_{l=1}^{n-2} e^{q(j_l|j_l,f)s_{l+1}} q(j_{l+1}|j_l,f) e^{q(j_{n-1}|j_{n-1},f)s_n} q(z|j_{n-1},f) ds_n \cdots ds_1 \\
&\geq (\varpi_2 \varpi_3)^{m-1} \varpi_2 e^{\varpi_4 n_2 t_0} \sum_{n=1}^m \int_{s_1+\cdots+s_n \leq n_2 t_0} \sum_{j_1 \in S\backslash\{i,z\}, j_{l+1} \in S\backslash\{j_l,z\}, l=1,2,\ldots,n-2} e^{q(i|i,f)s_1} q(j_1|i,f) \\
&\qquad \times \prod_{l=1}^{n-2} e^{q(j_l|j_l,f)s_{l+1}} q(j_{l+1}|j_l,f) e^{q(j_{n-1}|j_{n-1},f)s_n} q(z|j_{n-1},f) ds_n \cdots ds_1 \\
&= (\varpi_2 \varpi_3)^{m-1} \varpi_2 e^{\varpi_4 n_2 t_0} P^f_i(\tau_z \leq n_2 t_0, \tau_z \leq T_m), \tag{6.35}
\end{aligned}
$$

where the second equality follows from Proposition B.8 in Guo and Hernández-Lerma (2009, p.205). Note that $\overline{b} := (\varpi_2 \varpi_3)^{m^*-1} \varpi_2 e^{\varpi_4 n_2 t_0} \in (0,1)$. Hence, using Eqs. 6.31, 6.34 and 6.35, we have

$$
\begin{aligned}
\overline{P}^f_{i,n_1}(\tau_z \leq n_2 t_0) &\geq \overline{b} P^f_i(\tau_z \leq n_2 t_0, \tau_z \leq T_{m^*}) \\
&= \overline{b} P^f_i(\tau_z \leq n_2 t_0, \tau_1 \leq m^*) \\
&\geq \overline{b}[P^f_i(\tau_z \leq n_2 t_0) + P^f_i(\tau_1 \leq m^*) - 1] \geq \frac{1}{6}\overline{b}
\end{aligned}
$$

for all $i \in S$ and $f \in F$. Therefore, Eq. 6.30 holds with $\widetilde{t}_0 := n_2 t_0$ and $\widetilde{\alpha} := 1 - \frac{1}{6}\overline{b}$. Furthermore, employing Eq. 6.30 and an induction argument, we obtain

$$
\overline{P}^f_{i,n_1}(\tau_z > n\widetilde{t}_0) \leq \widetilde{\alpha}^n \quad \text{for all } i \in S, \ f \in F \text{ and } n = 1, 2, \ldots. \tag{6.36}
$$

Then for any $\kappa_0 \in \overline{O}_{n_1}$ satisfying $\kappa_0 < \frac{\ln \widetilde{\alpha}}{\lambda \widetilde{t}_0}$, $i \in S$ and $f \in F$, by (6.36) we get

$$
\begin{aligned}
e^{\lambda \overline{H}_{\kappa_0,n_1}(i,f)} &= \sum_{m=0}^{\infty} \overline{E}_{i,n_1}^f \left[e^{-\lambda \kappa_0 \tau_z} I_{\{\tau_z \in (m\widetilde{t}_0, (m+1)\widetilde{t}_0]\}} \right] \\
&\leq \sum_{m=0}^{\infty} e^{-\lambda \kappa_0 (m+1)\widetilde{t}_0} \overline{P}_{i,n_1}^f (\tau_z > m\widetilde{t}_0) \\
&\leq \frac{e^{-\lambda \kappa_0 \widetilde{t}_0}}{1 - \widetilde{\alpha} e^{-\lambda \kappa_0 \widetilde{t}_0}}.
\end{aligned}
$$

Take any $\kappa_1 \in (0, \kappa_0)$ satisfying $\kappa_1 < \min_{(i,a) \in K} \left\{ \frac{1}{\lambda} [\lambda c(i,a) - \lambda g_\lambda^* + q(i|i,a)] \right\}$ and define $\overline{H}_{\kappa_1,n_1}^*(i,f) := \gamma_{n_1} e^{\lambda \overline{H}_{\kappa_1,n_1}(i,f) + \lambda h_{g_\lambda^*,\lambda}^*(i,f)}$ for all $i \in S$ and $f \in F$. Thus, using Eqs. 6.25 and 6.29, we have

$$
\begin{cases}
\overline{H}_{\kappa_1,n_1}^*(i,f) \geq -\frac{1}{\lambda c(i,f) - \lambda g_\lambda^* - \lambda \kappa_1 + q(i|i,f)} \left(\gamma_{n_1} q(z|i,f) + \sum_{j \in S \setminus \{i,z\}} \overline{H}_{\kappa_1,n_1}^*(j,f) q(j|i,f) \right) \\
\overline{H}_{\kappa_1,n_1}^*(z,f) \geq -\frac{1}{\lambda c(z,f) - \lambda g_\lambda^* - \lambda \kappa_1 + q(z|z,f)} \sum_{j \in S \setminus \{z\}} \overline{H}_{\kappa_1,n_1}^*(j,f) q(j|z,f)
\end{cases}
$$

for all $i \in S \setminus \{z\}$ and $f \in F$. Hence, by the last inequalities and the similar arguments of Eq. 6.11, we obtain $\overline{H}_{\kappa_1,n_1}^*(i,f) \geq \gamma_{n_1} e^{\lambda h_{g_\lambda^* + \kappa_1,\lambda}^*(i,f)}$ for all $i \in S$ and $f \in F$, which implies

$$
e^{\lambda h_{g_\lambda^*,\lambda}^*(z)} \sup_{f \in F} e^{\lambda \overline{H}_{\kappa_1,n_1}(z,f)} \geq \sup_{f \in F} e^{\lambda \overline{H}_{\kappa_1,n_1}(z,f) + \lambda h_{g_\lambda^*,\lambda}^*(z,f)} \geq e^{\lambda h_{g_\lambda^* + \kappa_1,\lambda}^*(z)}. \quad (6.37)
$$

Let $\{\rho_m, m \geq 1\} \subseteq (0, \kappa_1)$ be a sequence satisfying $\lim_{m \to \infty} \rho_m = 0$. Note that $0 \leq e^{-\lambda \rho_m \widetilde{t}_0} \widetilde{\alpha} \leq e^{-\lambda \kappa_0 \widetilde{t}_0} \widetilde{\alpha} < 1$ for all $m \geq 1$. Then direct calculations give

$$
\begin{aligned}
0 \leq \sup_{f \in F} e^{\lambda \overline{H}_{\rho_m,n_1}(z,f)} - 1 &= \sup_{f \in F} \sum_{n=1}^{\infty} \frac{(-\lambda \rho_m)^n}{n!} \overline{E}_{z,n_1}^f [\tau_z^n] \\
&= \sup_{f \in F} \sum_{n=1}^{\infty} \frac{(-\lambda \rho_m)^n}{(n-1)!} \int_0^{\infty} t^{n-1} \overline{P}_{z,n_1}^f (\tau_z > t) dt \\
&= \sup_{f \in F} \sum_{n=1}^{\infty} \frac{(-\lambda \rho_m)^n}{(n-1)!} \sum_{l=0}^{\infty} \int_{l\widetilde{t}_0}^{(l+1)\widetilde{t}_0} t^{n-1} \overline{P}_{z,n_1}^f (\tau_z > t) dt \\
&\leq \sup_{f \in F} \sum_{n=1}^{\infty} \frac{(-\lambda \rho_m \widetilde{t}_0)^n}{(n-1)!} \sum_{l=0}^{\infty} (l+1)^{n-1} \overline{P}_{z,n_1}^f (\tau_z > l\widetilde{t}_0) \\
&\leq \sum_{n=1}^{\infty} \frac{(-\lambda \rho_m \widetilde{t}_0)^n}{(n-1)!} \sum_{l=0}^{\infty} (l+1)^{n-1} \widetilde{\alpha}^l \\
&= -\frac{\lambda \rho_m \widetilde{t}_0 e^{-\lambda \rho_m \widetilde{t}_0}}{1 - \widetilde{\alpha} e^{-\lambda \rho_m \widetilde{t}_0}} \quad (6.38)
\end{aligned}
$$

for all $m \geq 1$, where the second equality is due to Lemma 3.4 in Kallenberg (2012, p.49) and the last inequality follows from Eq. 6.36. Thus, employing Eq. 6.38 we obtain $\lim_{m \to \infty} \sup_{f \in F} e^{\lambda \overline{H}_{\rho_m,n_1}(z,f)} = 1$. Hence, for any $\varepsilon \in (0, e^{-\lambda h_{g_\lambda^*,\lambda}^*(z)} - 1]$, there exists a positive integer m_0 such that $\sup_{f \in F} e^{\lambda \overline{H}_{\rho_{m_0},n_1}(z,f)} - 1 \leq \varepsilon$, which together with Eq. 6.37 implies

$$
e^{\lambda h_{g_\lambda^* + \rho_{m_0},\lambda}^*(z)} \leq e^{\lambda h_{g_\lambda^*,\lambda}^*(z)} \sup_{f \in F} e^{\lambda \overline{H}_{\rho_{m_0},n_1}(z,f)} \leq 1. \quad (6.39)
$$

Moreover, by Eq. 6.39 we have $g_\lambda^* + \rho_{m_0} \in G_\lambda$, which leads to a contradiction that $g_\lambda^* + \rho_{m_0} \leq g_\lambda^*$. Therefore, we get $h_{g_\lambda^*, \lambda}^*(z) = 0$. $\square$

7 Proof of Theorem 3.4

In this section, we prove Theorem 3.4.

Proof Fix any $i \in S$ and $\pi \in \Pi$. For any $\lambda_1, \lambda_2 > 0$, we have

$$E_i^\pi \left[e^{\lambda_1 \int_0^T \int_A c(\xi_t, a) \pi(da|\xi_t, t) dt} \right] = E_i^\pi \left[e^{\lambda_2 \int_0^T \int_A c(\xi_t, a) \pi(da|\xi_t, t) dt} e^{(\lambda_1 - \lambda_2) \int_0^T \int_A c(\xi_t, a) \pi(da|\xi_t, t) dt} \right]$$

$$\leq E_i^\pi \left[e^{\lambda_2 \int_0^T \int_A c(\xi_t, a) \pi(da|\xi_t, t) dt} \right] e^{|\lambda_1 - \lambda_2| T \max_{(i,a) \in K} c(i,a)}$$

and

$$E_i^\pi \left[e^{\lambda_1 \int_0^T \int_A c(\xi_t, a) \pi(da|\xi_t, t) dt} \right] \geq E_i^\pi \left[e^{\lambda_2 \int_0^T \int_A c(\xi_t, a) \pi(da|\xi_t, t) dt} \right] e^{-|\lambda_1 - \lambda_2| T \max_{(i,a) \in K} c(i,a)},$$

which give

$$|\lambda_1 J_{V_{\lambda_1}}^*(i) - \lambda_2 J_{V_{\lambda_2}}^*(i)| \leq |\lambda_1 - \lambda_2| \max_{(i,a) \in K} c(i, a).$$

Thus, Theorem 3.3(b) and the last inequality yield that λg_λ^* is continuous in $\lambda \in (0, \infty)$. Hence, by Theorem 3.3(b) again, we obtain that $J_{V_\lambda}^*(i)$ is continuous in $\lambda \in (0, \infty)$. Moreover, employing the same technique above, we can obtain the continuity of $J_{V_\lambda}^*(i)$ in $\lambda \in (-\infty, 0)$. Below we show that $J_{V_\lambda}^*(i)$ is continuous in $\lambda = 0$. Take an arbitrary sequence $\{\lambda_n, n \geq 1\} \subseteq (0, \infty)$ with $\lim_{n \to \infty} \lambda_n = 0$. For each $n \geq 1$, let $g_{\lambda_n}^*$ and $h_{g_{\lambda_n}^*, \lambda_n}^*$ be as in Theorem 3.3 with λ_n in lieu of λ. Set $\widetilde{h}_{g_{\lambda_n}^*, \lambda_n}^*(i) := h_{g_{\lambda_n}^*, \lambda_n}^*(i) - \min_{j \in S} h_{g_{\lambda_n}^*, \lambda_n}^*(j)$ for all $n \geq 1$. Then for each $n \geq 1$, we have $\widetilde{h}_{g_{\lambda_n}^*, \lambda_n}^*(i) \geq 0$ and there exists $i^{\lambda_n} \in S$ satisfying $\widetilde{h}_{g_{\lambda_n}^*, \lambda_n}^*(i^{\lambda_n}) = 0$. Moreover, using Theorem 3.3(a), for each $n \geq 1$, there exists $f_{\lambda_n}^* \in F$ such that

$$\lambda_n g_{\lambda_n}^* e^{\lambda_n \widetilde{h}_{g_{\lambda_n}^*, \lambda_n}^*(i)} = \inf_{a \in A(i)} \left\{ \lambda_n c(i, a) e^{\lambda_n \widetilde{h}_{g_{\lambda_n}^*, \lambda_n}^*(i)} + \sum_{j \in S} e^{\lambda_n \widetilde{h}_{g_{\lambda_n}^*, \lambda_n}^*(j)} q(j|i, a) \right\} \quad (7.1)$$

$$= \lambda_n c(i, f_{\lambda_n}^*) e^{\lambda_n \widetilde{h}_{g_{\lambda_n}^*, \lambda_n}^*(i)} + \sum_{j \in S} e^{\lambda_n \widetilde{h}_{g_{\lambda_n}^*, \lambda_n}^*(j)} q(j|i, f_{\lambda_n}^*). \quad (7.2)$$

Note that Theorem 3.3(b) gives $0 \leq g_{\lambda_n}^* \leq \max_{(i,a) \in K} c(i, a)$ for all $n \geq 1$. Thus, choose any convergent subsequence $\{g_{\lambda_{n_l}}^*, l \geq 1\}$ of $\{g_{\lambda_n}^*, n \geq 1\}$ and denote the corresponding limit by

$$\overline{g} := \lim_{l \to \infty} g_{\lambda_{n_l}}^* \in \left[0, \max_{(i,a) \in K} c(i, a) \right]. \quad (7.3)$$

Furthermore, by the finiteness of S and the compactness of F and $[0, \infty]$, there exists a subsequence of $\{n_l\}$ (still denoted by $\{n_l\}$) such that $i^{\lambda_{n_l}} = i^* \in S$ for all $l \geq 1$ and the limits of the sequences $\{\widetilde{h}_{g_{\lambda_{n_l}}^*, \lambda_{n_l}}^*, l \geq 1\}$ and $\{f_{\lambda_{n_l}}^*, l \geq 1\}$ exist. Set

$$\overline{h}(j) := \lim_{l \to \infty} \widetilde{h}_{g_{\lambda_{n_l}}^*, \lambda_{n_l}}^*(j) \text{ and } f^*(j) := \lim_{l \to \infty} f_{\lambda_{n_l}}^*(j) \text{ for all } j \in S. \quad (7.4)$$

Then we have $\overline{h}(j) \geq 0$ for all $j \in S$ and $\overline{h}(i^*) = 0$. Below we show $\overline{h}(j) < \infty$ for all $j \in S$ by an induction argument. For any $j \neq i^*$, Assumption 3.1(iii) yields that there

exist different states $k_1 = i^*, k_2, \ldots, k_{m'} = j$ such that $q(k_{n+1}|k_n, f^*) > 0$ for all $n = 1, 2, \ldots, m' - 1$. Observe that $\bar{h}(k_1) = \bar{h}(i^*) = 0$. Thus, $\bar{h}(k_n) < \infty$ holds for $n = 1$. Suppose that $\bar{h}(k_{n^*}) < \infty$ for some $n^* \in \{1, 2, \ldots, m' - 1\}$. Employing Eq. 7.2 we obtain

$$g^*_{\lambda_{n_l}} e^{\lambda_{n_l} \widetilde{h}^*_{g^*_{\lambda_{n_l}}, \lambda_{n_l}}(k_{n^*})} = c(k_{n^*}, f^*_{\lambda_{n_l}}) e^{\lambda_{n_l} \widetilde{h}^*_{g^*_{\lambda_{n_l}}, \lambda_{n_l}}(k_{n^*})} + \sum_{j \in S} \frac{e^{\lambda_{n_l} \widetilde{h}^*_{g^*_{\lambda_{n_l}}, \lambda_{n_l}}(j)} - 1}{\lambda_{n_l}} q(j|k_{n^*}, f^*_{\lambda_{n_l}}),$$

which together with the inequality $e^x - 1 \geq x \ (x \geq 0)$ gives

$$g^*_{\lambda_{n_l}} e^{\lambda_{n_l} \widetilde{h}^*_{g^*_{\lambda_{n_l}}, \lambda_{n_l}}(k_{n^*})} \geq c(k_{n^*}, f^*_{\lambda_{n_l}}) e^{\lambda_{n_l} \widetilde{h}^*_{g^*_{\lambda_{n_l}}, \lambda_{n_l}}(k_{n^*})} + \frac{e^{\lambda_{n_l} \widetilde{h}^*_{g^*_{\lambda_{n_l}}, \lambda_{n_l}}(k_{n^*})} - 1}{\lambda_{n_l}} q(k_{n^*}|k_{n^*}, f^*_{\lambda_{n_l}})$$
$$+ \widetilde{h}^*_{g^*_{\lambda_{n_l}}, \lambda_{n_l}}(k_{n^*+1}) q(k_{n^*+1}|k_{n^*}, f^*_{\lambda_{n_l}})$$

for all $l \geq 1$. Then letting $l \to \infty$ in the both sides of the last inequality and using the induction hypothesis, Assumption 3.1(ii), Eqs. 7.3 and 7.4, we get

$$\bar{g} \geq c(k_{n^*}, f^*) + \bar{h}(k_{n^*}) q(k_{n^*}|k_{n^*}, f^*) + \bar{h}(k_{n^*+1}) q(k_{n^*+1}|k_{n^*}, f^*),$$

which implies $\bar{h}(k_{n^*+1}) < \infty$. Hence, by induction we have $\bar{h}(k_n) < \infty$ for all $n = 1, 2, \ldots, m'$. Therefore, we obtain $\bar{h}(j) \in [0, \infty)$ for all $j \in S$. By Eq. 7.1 we get

$$g^*_{\lambda_{n_l}} e^{\lambda_{n_l} \widetilde{h}^*_{g^*_{\lambda_{n_l}}, \lambda_{n_l}}(i)} \leq c(i, a) e^{\lambda_{n_l} \widetilde{h}^*_{g^*_{\lambda_{n_l}}, \lambda_{n_l}}(i)} + \sum_{j \in S} \frac{e^{\lambda_{n_l} \widetilde{h}^*_{g^*_{\lambda_{n_l}}, \lambda_{n_l}}(j)} - 1}{\lambda_{n_l}} q(j|i, a)$$

for all $l \geq 1$ and $a \in A(i)$. Then the last inequality, Eqs. 7.3 and 7.4 yield

$$\bar{g} \leq c(i, a) + \sum_{j \in S} \bar{h}(j) q(j|i, a) \ \text{ for all } a \in A(i). \tag{7.5}$$

Thus, employing Eq. 7.5 and the Dynkin formula, we obtain

$$\bar{g} T \leq E^\pi_i \left[\int_0^T \int_A c(\xi_t, a) \pi(da|\xi_t, t) dt \right] + E^\pi_i [\bar{h}(\xi_T)] - \bar{h}(i)$$

for all $T > 0$, which gives

$$\bar{g} \leq J^*_{V_0}(i). \tag{7.6}$$

On the other hand, using Eq. 7.2 and the similar arguments of Eq. 7.6, we have

$$\bar{g} = J_{V_0}(i, f^*) \geq J^*_{V_0}(i). \tag{7.7}$$

Then combining Eqs. 7.6 and 7.7, we get $\bar{g} = J^*_{V_0}(i)$. Since all the convergent subsequences of $\{g^*_{\lambda_n}, n \geq 1\}$ have the same limit $J^*_{V_0}(i)$, we obtain $\lim_{n \to \infty} g^*_{\lambda_n} = J^*_{V_0}(i)$, which together with Theorem 3.3(b) implies $\lim_{n \to \infty} J^*_{V_{\lambda_n}}(i) = J^*_{V_0}(i)$. Therefore, $J^*_{V_\lambda}(i)$ is right-continuous in $\lambda = 0$. Moreover, following the same technique above, we have that $J^*_{V_\lambda}(i)$ is left-continuous in $\lambda = 0$. Hence, we complete the proof of the theorem. $\square$

8 Concluding remarks

In this paper we have studied the U-average cost criterion for the CTMDPs with a finite state space. Under the continuity-compactness condition and the irreducibility condition, we have shown that the simultaneous Doeblin condition for the CTMDPs holds. Moreover, we

have obtained the optimality equation of the auxiliary risk-sensitive first passage optimization problem and the properties of the corresponding optimal value function for any nonzero risk-sensitivity parameter. Then employing the obtained results on the risk-sensitive first passage criterion, we have established the existence of a solution to the optimality equation of the risk-sensitive average cost criterion allowing the risk-sensitivity parameter to take any nonzero value. Furthermore, we have proven that the optimal value function of the risk-sensitive average cost criterion is continuous with respect to the risk-sensitivity parameter. Finally, we have given the connections between the U-average cost criterion and the average cost criteria induced by the identity function and the exponential utility function, from which the existence of a U-average optimal deterministic stationary policy has been shown. It should be mentioned that the optimality equation of the risk-sensitive average cost criterion allowing the risk-sensitivity parameter to take any nonzero value plays a crucial role in the study of the U-average cost criterion. Hence, when dealing with the U-average cost criterion with a countable state space, the difficulty lies in finding the conditions under which the optimality equation of the risk-sensitive average cost criterion holds for any nonzero risk-sensitivity parameter. In addition, the CTMDPs with the bounded transition rates under the expected discounted cost and expected average cost criteria can be transformed to the equivalent discrete-time MDPs by the uniformization technique. Whether the uniformization technique is applicable to the CTMDPs under the risk-sensitive average cost criterion is a very interesting problem.

References

Aliprantis C, Border K (2007) Infinite dimensional analysis. Springer, New York

Anderson WJ (1991) Continuous-time Markov chains: an applications-oriented approach. Springer, New York

Bäuerle N, Rieder U (2014) More risk-sensitive Markov decision processes. Math Oper Res 39:105–120

Cavazos-Cadena R, Fernández-Gaucherand E (2005) Risk-sensitive optimal control in communicating average Markov decision chains. In: Dror M, L'Ecuyer P, Szidarovsky F (eds) In: Modelling uncertainty: an examination of stochastic theory, methods, and applications. Springer, New York, pp 515–553

Cavazos-Cadena R (2010) Optimality equations and inequalities in a class of risk-sensitive average cost Markov decision chains. Math Meth Oper Res 71:47–84

Cavazos-Cadena R, Hernández-Hernández D (2011) Discounted approximations for risk-sensitive average criteria in Markov decision chains with finite state space. Math Oper Res 36:133–146

Cavazos-Cadena R, Hernández-Hernández D (2016) A characterization of the optimal certainty equivalent of the average cost via the Arrow-Pratt sensitivity function. Math Oper Res 41:224–235

Di Masi GB, Stettner L (2007) Infinite horizon risk sensitive control of discrete time Markov processes under minorization property. SIAM J Control Optim 46:231–252

Ghosh MK, Saha S (2014) Risk-sensitive control of continuous time Markov chains. Stochastics 86:655–675

Guo XP, Hernández-Lerma O (2009) Continuous-time Markov decision processes: theory and applications. Springer, Berlin

Guo XP, Huang YH, Song XY (2012) Linear programming and constrained average optimality for general continuous-time Markov decision processes in history-dependent policies. SIAM J Control Optim 50:23–47

Jaśkiewicz A (2007) Average optimality for risk-sensitive control with general state space. Ann Appl Probab 17:654–675

Kallenberg O (2012) Foundations of modern probability. Springer, New York

Kitaev MY, Rykov VV (1995) Controlled queueing systems. CRC Press, Boca Raton

Puterman ML (1994) Markov decision processes: discrete stochastic dynamic programming. Wiley, New York

Wei QD, Chen X (2014) Strong average optimality criterion for continuous-time Markov decision processes. Kybernetika 50:950–977

Wei QD, Chen X (2016) Continuous-time Markov decision processes under the risk-sensitive average cost criterion. Oper Res Lett 44:457–462

Qingda Wei received the Ph.D. degree in probability and statistics from Sun Yat-Sen University in 2013. He is an associate professor in Huaqiao University. His research interests include Markov decision processes, stochastic games and stochastic optimization.

Xian Chen received the Ph.D. degree in probability and statistics from Academy of Mathematics and Systems Science, Chinese Academy of Sciences in 2014. She is an assistant professor in Xiamen University. Her research interests include Markov decision processes, mathematical biology, Dirichlet forms and stochastic partial differential equations.