



Risk-sensitive discounted cost criterion for continuous-time Markov decision processes on a general state space

Subrata Golui¹ · Chandan Pal¹

Received: 14 September 2021 / Revised: 31 December 2021 / Accepted: 17 March 2022 /
Published online: 11 April 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

In this paper, we consider risk-sensitive discounted control problem for continuous-time jump Markov processes taking values in general state space. The transition rates of underlying continuous-time jump Markov processes and the cost rates are allowed to be unbounded. Under certain Lyapunov condition, we establish the existence and uniqueness of the solution to the Hamilton–Jacobi–Bellman equation. Also, we prove the existence of optimal risk-sensitive control in the class of Markov control and completely characterized the optimal control.

Keywords Continuous-time Markov decision process · History-dependent control · General state space · Risk-sensitive discounted criterion · HJB equation · Optimal control

1 Introduction

In this paper, we study the risk-sensitive discounted criterion for continuous-time Markov decision processes (CTMDPs) with Borel state space. In the risk-neutral criterion, the controller wants to optimize the expected value of the total payoff. But in the risk-sensitive criterion, controller considers the expected value of the exponential of the total payoff and so, the risk-sensitive criterion gives a better protection from the risk. Therefore, the risk-sensitive or exponential of integral is a very popular cost criterion due to its applications in many areas such as queueing systems and finance, for more details see Bauerle and Rieder (2014) and Whittle (1990) and the references therein. In the literature, risk-sensitive control problems for CTMDPs are an important

✉ Chandan Pal
cpal@iitg.ac.in

Subrata Golui
golui@iitg.ac.in

¹ Department of Mathematics, Indian Institute of Technology Guwahati, Guwahati, Assam, India

class of stochastic optimal control problems and have been widely studied under different sets of conditions. Finite horizon risk-sensitive CTMDPs for countable state space were studied in Ghosh and Saha (2014), Guo et al. (2019) and Wei (2016) and for infinite horizon risk-sensitive CTMDPs we refer to Ghosh and Saha (2014), Guo and Zhang (2020), Kumar and Pal (2013, 2015), Pal and Pradhan (2019) and Zhang (2017). For important contributions to the risk-sensitive control of discrete-time MDP on a general state space, see Masi and Stettner (2000) and Masi and Stettner (2007). Although risk-sensitive control of CTMDPs on a countable state space have been studied extensively, but the corresponding literature in the context of risk-sensitive control of CTMDPs on a general state space is rather limited. Some exceptions are (Guo and Zhang 2019, 2020; Pal and Pradhan 2019).

In the paper (Pal and Pradhan 2019), the authors studied risk-sensitive control of pure jump processes on a general state space. They considered bounded transition and cost rates and all controls are Markovian. In Pal and Pradhan (2019), the authors proved the HJB characterization of the optimal risk-sensitive control. The boundedness assumption on the transition and cost rates plays a key role in the proof of the existence of the optimal risk-sensitive control in Pal and Pradhan (2019). This boundedness requirement, however, imposes some restrictions in applications, for instance in queueing control and population processes, where the transition and reward/cost rates are usually unbounded. Also, there are many real-life situations where the state space may be uncountable, for example, the chemical reaction model, Gaussian model, etc. One can see Guo and Zhang (2019), Piunovskiy and Zhang (2020) and references therein for the real-life examples. In Guo and Zhang (2019), the author considered the finite-horizon risk-sensitive control problem for CTMDPs on a Borel state space with unbounded transition and cost rates and proved the existence of optimal control via the HJB equation.

In this paper, we study a much more general risk-sensitive control problem for CTMDP with general state space. To the best of our knowledge, this is the first work which deals with infinite horizon discounted risk-sensitive control for CTMDPs on a general state space with unbounded cost and transition rates and the controls can be history-dependent. The main objective of this work is to prove the existence of the solution of the HJB equation and the characterization of the optimal risk-sensitive control. In particular (1) We prove that the HJB equations has a unique solution $\varphi_\alpha \in L_V^\infty([0, 1] \times S)$ satisfying the bounds as in Eq. (3.1) below, where $L_V^\infty([0, 1] \times S)$ is described below. (2) We prove that any measurable minimizer of the HJB equation is optimal, and conversely any optimal control in the class of Markov controls is a minimizer of the HJB Eq. (3.1) below. We first consider for bounded transition and cost rates and establish the existence of a solution to the corresponding HJB equation by Banach's fixed point theorem as in Pal and Pradhan (2019). Then we will relax the bounded hypothesis and extend this result to unbounded transition and cost rates. We characterize the value function via the HJB equation. Also, we prove the existence of an optimal control in the class of Markov control and the HJB characterization of the optimal risk-sensitive control and prove its complete characterization. In Corollary 5.1, we prove that if the cost and transition rates are bounded, then an optimal control exists for our model.

The rest of this article is structured as follows. Section 2 deals with the description of the problem, required notations, some Assumptions, and preliminary results. In Sect. 3, we give a continuity-compactness Assumption and prove the stochastic representation of the solution of the HJB Eq. (3.1). In Sect. 4, we truncate our transition and cost rates and prove the existence of the unique solution to the HJB equation. A complete characterization of optimal control is proven in Sect. 5. In Sect. 6, we illustrate our theory and assumptions by an illustrative example.

2 The control problems

The model of CTMDP is a five-tuple which consists of the following elements:

$$\mathbb{M} := \{S, A, (A(x) \subset A, x \in S), c(x, a), q(\cdot|x, a)\},$$

- a Borel space S , called the state space, whose elements are referred to as states of the system and the corresponding Borel σ -algebra is $\mathcal{B}(S)$. (Throughout the whole paper we consider that for any Borel space X , the corresponding Borel σ -algebra is $\mathcal{B}(X)$.)
- A is the action set, which is assumed to be Borel space with the Borel σ -algebra $\mathcal{B}(A)$.
- for each $x \in S$, $A(x) \in \mathcal{B}(A)$ denotes the set of admissible actions for state x . Let $K := \{(x, a)|x \in S, a \in A(x)\}$, which is a Borel subset of $S \times A$.
- the measurable function $c : K \rightarrow \mathbb{R}_+$ denotes the cost rate function. We require the cost function $c(x, a)$ to measure (or evaluate) the utility of taking action a at state x .
- given any $(x, a) \in K$, the transition rate $q(\cdot|x, a)$ is a Borel measurable signed kernel on S given K . That is, $q(\cdot|x, a)$ satisfies countable additivity; $q(D|x, a) \geq 0$ where $(x, a) \in K, x \notin D$ and $D \in \mathcal{B}(S)$. Moreover, we assume that $q(\cdot|x, a)$ satisfies the following conservative and stable conditions: for any $x \in S$,

$$q(S|x, a) \equiv 0 \quad \text{and} \\ q^*(x) := \sup_{a \in A(x)} q_x(a) < \infty,$$

where $q_x(a) := -q(\{x\}|x, a) \geq 0$. We need the transition rates to specify the random dynamic evolution of the system.

Next, we give an informal description of the evolution of the CTMDPs as follows. The controller observes continuously the current state of the system. When the system is in state $x \in S$ at time $t \geq 0$, he/she chooses action $a_t \in A(x)$ according to some control. As a consequence of this, the following happens:

- the controller incurs an immediate cost at rate $c(x, a_t)$; and
- after a random sojourn time (i.e., the holding time at state x), the system jumps to a set $B \in \mathcal{B}(S) (x \notin B)$ of states with the transition probability $\frac{q(B|x, a_t)}{q_x(a_t)}$

determined by the transition rates $q(dy|x, a_t)$. The distribution function of the sojourn time is $(1 - e^{-\int_t^{t+x} q_x(a_s)ds})$. (see Proposition B.8 in [Guo and Hernandez-Lerma (2009), p. 205] for details).

When the state of the system transits to a new state $y \neq x$, the above procedure is repeated. Thus, the controller tries to minimize his/her costs with respect to some performance criterion $\mathcal{J}_\alpha(\cdot, \cdot, \cdot)$, which in our present case is defined by (2.2), below. To formalize what is described above, below we describe the construction of continuous time Markov decision processes (CTMDPs) under possibly history-dependent controls. To construct the underlying CTMDPs (as in Guo and Piunovskiy 2011; Kitaev 1995; Piunovskiy and Zhang 2011, 2020) we introduce some notations: let $S_\Delta := S \cup \{\Delta\}$ (with some ‘‘isolated’’ state $\Delta \notin S$), $\Omega_0 := (S \times (0, \infty))^\infty$, $\Omega_k := (S \times (0, \infty))^k \times S \times (\{\infty\} \times \{\Delta\})^\infty$ for $k \geq 1$ and $\Omega := \cup_{k=0}^\infty \Omega_k$. Let \mathcal{F} be the Borel σ -algebra on Ω . Then we obtain the measurable space (Ω, \mathcal{F}) . For some $k \geq 1$, and sample $\omega := (x_0, \theta_1, x_1, \dots, \theta_k, x_k, \dots) \in \Omega$, define

$$T_0(\omega) := 0, T_k(\omega) := T_{k-1}(\omega) + \theta_k, T_\infty(\omega) := \lim_{k \rightarrow \infty} T_k(\omega).$$

Using $\{T_k\}$, we define the state process $\{\xi_t\}_{t \geq 0}$ as

$$\xi_t(\omega) := \sum_{k \geq 0} I_{\{T_k \leq t < T_{k+1}\}} x_k + I_{\{t \geq T_\infty\}} \Delta, \text{ for } t \geq 0 \text{ (with } T_0 := 0). \tag{2.1}$$

Here, I_E denotes the indicator function of a set E , and we use the convention that $0 + z :=: z$ and $0z :=: 0$ for all $z \in S_\Delta$. Obviously, $\xi_t(\omega)$ is right-continuous on $[0, \infty)$. We denote $\xi_{t-}(\omega) := \liminf_{s \rightarrow t-} \xi_s(\omega)$. From Eq. (2.1), we see that $T_k(\omega)$ ($k \geq 1$) denotes the k -th jump moment of $\{\xi_t, t \geq 0\}$, x_{k-1} is the state of the process on $[T_{k-1}(\omega), T_k(\omega))$, $\theta_k = T_k(\omega) - T_{k-1}(\omega)$ plays the role of sojourn time at state x_{k-1} , and the sample path $\{\xi_t(\omega), t \geq 0\}$ has at most denumerable states x_k ($k = 0, 1, \dots$). The process after T_∞ is regarded to be absorbed in the state Δ . Thus, let $q(\cdot|\Delta, a_\Delta) := 0$, $A_\Delta := A \cup \{a_\Delta\}$, $A(\Delta) := \{a_\Delta\}$, $c(\Delta, a) := 0$ for all $a \in A_\Delta$, where a_Δ is an isolated point.

To precisely define the criterion, we need to introduce the concept of a control as in Guo et al. (2012), Guo and Piunovskiy (2011) and Kitaev and Rykov (1995). Take the right-continuous σ -algebras $\{\mathcal{F}_t\}_{t \geq 0}$ with $\mathcal{F}_t := \sigma(\{T_k \leq s, \xi_{T_k} \in S\} : 0 \leq s \leq t, k \geq 0)$. For all $t \geq 0$, $\mathcal{F}_{s-} := \bigvee_{0 \leq t' < s} \mathcal{F}_{t'}$. Now define a σ -algebra $\mathcal{P} := \sigma(A \times \{0\}, B \times (s, \infty) : A \in \mathcal{F}_0, B \in \mathcal{F}_{s-})$, which denotes the σ -algebra of predictable sets on $\Omega \times [0, \infty)$ related to $\{\mathcal{F}_t\}_{t \geq 0}$. To complete the specification of a stochastic optimal control problem, we need, of course, to introduce an optimality criterion. This requires to define the class of controls as below.

Definition 2.1 A history-dependent policy $\pi := \{\pi_t(\omega)\}_{t \geq 0}$ is a measurable map from $(\Omega \times [0, \infty), \mathcal{P})$ onto $(A_\Delta, \mathcal{B}(A_\Delta))$ satisfying $\pi_t(\omega) \in A(\xi_{t-}(\omega))$ for all $\omega \in \Omega$ and $t \geq 0$. For notational simplicity, we denote a history-dependent control as $\{\pi_t\}_{t \geq 0}$. The set of all history-dependent controls is denoted by Π . A control $\pi \in \Pi$, is called a Markov if $\pi_t(\omega) = \pi_t(\xi_{t-}(w))$ for every $w \in \Omega$ and $t \geq 0$, where $\xi_{t-}(w) := \lim_{s \uparrow t} \xi_s(w)$. We denote by Π^m the family of all Markov controls.

For any compact metric space Y , let $P(Y)$ denote the space of probability measures on Y with Prohorov topology. Under Assumption 2.1 below, for any initial state $x \in S$ and any control $\pi \in \Pi$, Theorem 4.27 in Kitaev and Rykov (1995) yields the existence of a unique probability measure denoted by P_x^π on (Ω, \mathcal{F}) . Let E_x^π be the expectation operator with respect to P_x^π . Fix any discounted factor $\alpha > 0$. For any $\pi \in \Pi$ and $x \in S$, the risk-sensitive discounted criterion is defined as

$$\mathcal{J}_\alpha(\theta, x, \pi) := \frac{1}{\theta} \log \left\{ E_x^\pi \left[\exp \left(\theta \int_0^\infty e^{-\alpha t} c(\xi_t, \pi_t) dt \right) \right] \right\}, \tag{2.2}$$

provided that the integral is well defined, where $\{\xi_t\}_{t \geq 0}$ is the Markov process corresponding to $\pi = \{\pi_t\}_{t \geq 0} \in \Pi$ and $\theta \in (0, 1]$ denotes a risk-sensitive parameter and the limiting case of $\theta \rightarrow 0$ is the risk-neutral case. For each $x \in S$, let

$$\mathcal{J}_\alpha^*(\theta, x) = \inf_{\pi \in \Pi} \mathcal{J}_\alpha(\theta, x, \pi).$$

A control $\pi^* \in \Pi$ is said to be optimal if $\mathcal{J}_\alpha(\theta, x, \pi^*) = \mathcal{J}_\alpha^*(\theta, x)$ for all $x \in S$. The objective of this paper is to provide conditions for the existence of optimal control and introduce a HJB characterization of such control.

Since the logarithm is an increasing function, instead of studying $\mathcal{J}_\alpha(\theta, x, \pi)$, we will consider $\tilde{\mathcal{J}}_\alpha(\theta, x, \pi)$ on $[0, 1] \times S \times \Pi$ defined by

$$\tilde{\mathcal{J}}_\alpha(\theta, x, \pi) := E_x^\pi \left[\exp \left(\theta \int_0^\infty e^{-\alpha t} c(\xi_t, \pi_t) dt \right) \right]. \tag{2.3}$$

Obviously, $\tilde{\mathcal{J}}_\alpha(\theta, x, \pi) \geq 1$ for $(\theta, x) \in [0, 1] \times S$ and $\pi \in \Pi$, and we have π^* is optimal if and only if $\inf_{\pi \in \Pi} \tilde{\mathcal{J}}_\alpha(\theta, x, \pi) = \tilde{\mathcal{J}}_\alpha(\theta, x, \pi^*) =: \tilde{\mathcal{J}}_\alpha^*(\theta, x) \forall x \in S$. Since the rates $q(dy|x, a)$ and costs $c(x, a)$ are allowed to be unbounded, we next give conditions for the non-explosion of $\{\xi_t, t \geq 0\}$ and finiteness of $\mathcal{J}_\alpha(\theta, x, \pi)$, which had been widely used in CTMDPs; see, for instance, (Guo and Hernandez-Lerma 2009; Guo et al. 2012; Guo and Liao 2019; Guo and Piunovskiy 2011; Prieto-Rumeau and Hernandez-Lerma 2012) and references therein.

Assumption 2.1 There exists a real-valued Borel measurable function $\mathcal{V} \geq 1$ on S and constants $\rho_0 > 0, M_0 > 0, L_0 \geq 0$ and $0 < \rho_1 < \min\{\alpha, \rho_0^{-1}\alpha^2\}$ such that

- (i) $\int_S \mathcal{V}(y)q(dy|x, a) \leq \rho_0 \mathcal{V}(x) \quad \forall (x, a) \in K;$
- (ii) $\sup_{a \in A(x)} q_x(a) \leq M_0 \mathcal{V}(x) \quad \forall x \in S;$
- (iii) $\sup_{a \in A(x)} c(x, a) \leq \rho_1 \log \mathcal{V}(x) + L_0 \quad \forall x \in S.$

Remark 2.1 (a) Note that, when the transition rates are bounded i.e., $\sup_{x \in S} q^*(x) < \infty$, Assumptions 2.1 (i) and (ii) are satisfied by taking a suitable constant value of $\mathcal{V}(x)$.

(b) Under Assumption 2.1 (iii) the criterion (2.3) is well defined and finite; see Proposition 2.1(c) below.

Proposition 2.1 *Grant Assumption 2.1. Then for any control $\pi \in \Pi$ and $(\theta, x) \in [0, 1] \times S$, the following results are true:*

- (a) $P_x^\pi(T_\infty = \infty) = 1$, $P_x^\pi(\xi_0 = x) = 1$, and $P_x^\pi(\xi_t \in S) = 1$ for all $t \geq 0$;
- (b) $E_x^\pi[\mathcal{V}(\xi_t)] \leq e^{\rho_0 t} \mathcal{V}(x)$ for all $t \geq 0$;
- (c) *We have*

$$\tilde{J}_\alpha(\theta, x, \pi) \leq \frac{\alpha^2}{\alpha^2 - \rho_0 \rho_1 \theta} e^{\theta L_0/\alpha} [\mathcal{V}(x)]^{\frac{\rho_1 \theta}{\alpha}} \leq \frac{\alpha^2}{\alpha^2 - \rho_0 \rho_1} e^{L_0/\alpha} \mathcal{V}(x).$$

Also, we get

$$\mathcal{J}_\alpha^*(\theta, x) \leq \log\left(\frac{\alpha^2}{\alpha^2 - \rho_0 \rho_1 \theta}\right) + \frac{L_0}{\alpha} + \frac{\rho_1}{\alpha} \log \mathcal{V}(x) \quad \forall \theta \in (0, 1], x \in S. \quad (2.4)$$

Proof For parts (a) and (b), see, Guo et al. (2012) and Guo and Piunovskiy (2011, Theorem 3.1).

Proof of part (c): Observe that $d(-e^{-\alpha t})$ is a probability measure on $[0, \infty)$. For any $\pi \in \Pi$ and $(\theta, x) \in [0, 1] \times S$, by (2.3) and Jensen’s inequality we have

$$\begin{aligned} \tilde{J}_\alpha(\theta, x, \pi) &= E_x^\pi \left[\exp\left(\int_0^\infty \frac{\theta}{\alpha} c(\xi_t, \pi_t) d(-e^{-\alpha t})\right) \right] \\ &\leq E_x^\pi \left[\int_0^\infty \exp\left(\frac{\theta}{\alpha} c(\xi_t, \pi_t)\right) d(-e^{-\alpha t}) \right]. \end{aligned}$$

By Assumption 2.1 and part (b) we obtain

$$\begin{aligned} \tilde{J}_\alpha(\theta, x, \pi) &\leq E_x^\pi \left[\int_0^\infty \exp\left(\frac{\theta}{\alpha} (\rho_1 \log \mathcal{V}(\xi_t) + L_0)\right) d(-e^{-\alpha t}) \right] \\ &= e^{\theta L_0/\alpha} \left[\int_0^\infty E_x^\pi \left(\mathcal{V}(\xi_t)^{\frac{\rho_1 \theta}{\alpha}} \right) d(-e^{-\alpha t}) \right] \\ &\leq e^{\theta L_0/\alpha} \left[\int_0^\infty (E_x^\pi [\mathcal{V}(\xi_t)])^{\frac{\rho_1 \theta}{\alpha}} d(-e^{-\alpha t}) \right] \quad (\text{since } \rho_1 \theta < \alpha) \\ &\leq \alpha e^{\theta L_0/\alpha} [\mathcal{V}(x)]^{\frac{\rho_1 \theta}{\alpha}} \left[\int_0^\infty \exp\left(\frac{\rho_0 \rho_1 \theta t}{\alpha} - \alpha t\right) dt \right] \\ &= \frac{\alpha^2}{\alpha^2 - \rho_0 \rho_1 \theta} e^{\theta L_0/\alpha} [\mathcal{V}(x)]^{\frac{\rho_1 \theta}{\alpha}}, \end{aligned}$$

where the last equality holds due to the fact that $\rho_0 \rho_1 \theta < \alpha^2$.

Next observe that $\sup_{\theta \in [0, 1]} \tilde{J}_\alpha^*(\theta, x) \leq \frac{\alpha^2}{\alpha^2 - \rho_0 \rho_1} e^{L_0/\alpha} \mathcal{V}(x)$, and

$$\sup_{\theta \in (0, 1)} \mathcal{J}_\alpha^*(\theta, x) = \sup_{\theta \in (0, 1)} \frac{1}{\theta} \log \tilde{J}_\alpha^*(\theta, x) \leq \sup_{\theta \in (0, 1)} \frac{1}{\theta} \left(\log \frac{\alpha^2}{\alpha^2 - \rho_0 \rho_1 \theta} \right) + \frac{L_0}{\alpha} + \frac{\rho_1}{\alpha} \log \mathcal{V}(x).$$

Also, doing a simple and direct calculation, we achieve (2.4). □

In Ghosh and Saha (2014) and Kumar and Pal (2013), the authors used the Dynkin’s formula within the class of Markov controls by using the Markov property of the state process $\{\xi_t\}_{t \geq 0}$. But this Markov property may fail to hold when we study within the class of history-dependent controls, and consequently, here we can’t directly apply the Dynkin formula. Hence we assume the following condition, so that we can apply the Dynkin’s formula for a large enough class of functions, which had been widely used in CTMDPs; see, for instance, (Guo and Liao 2019; Guo et al. 2019; Guo and Zhang 2019).

Assumption 2.2 The Borel measurable function $\mathcal{V}^2 \geq 1$ on S satisfies the following Lyapunov condition

$$\int_S q(dy|x, a)\mathcal{V}^2(y) \leq \rho_2\mathcal{V}^2(x) + b_0 \quad \forall (x, a) \in K,$$

for some constants $0 < \rho_2 < \alpha$ and $b_0 \geq 0$. Here \mathcal{V} is as in Assumption 2.1.

We now introduce some frequently used notations.

- $C_c^\infty(a, b)$ denotes the set of all infinitely differentiable functions on (a, b) with compact support.
- Let $A_{as}([0, 1] \times S)$ denote the space of all functions which are real-valued and differentiable almost everywhere with respect to the first variable $\theta \in [0, 1]$. Given any real-valued function $W \geq 1$ on S and any Borel set X , a real-valued function φ on $X \times S$ is called W bounded if $\|\varphi\|_W^\infty := \sup_{(\theta, x) \in X \times S} \frac{|\varphi(\theta, x)|}{W(x)} < \infty$.

Denote $B_W(X \times S)$ the Banach space of all W -bounded functions. When $W \equiv 1$, $B_1([0, 1] \times S)$ is the space of all bounded functions on $[0, 1] \times S$. Now define $L_W^\infty([0, 1] \times S) := \{\varphi : [0, 1] \times S \rightarrow \mathbb{R} : \varphi \in B_W([0, 1] \times S) \cap A_{as}([0, 1] \times S)\}$.

3 Stochastic representation of a solution to the HJB equation

In this section, we prove that if the HJB equation for the cost criterion (2.3) has a solution then we will give a stochastic representation of that solution. Using dynamic programming heuristics, the HJB equations for the discounted cost criterion (2.3) is given by

$$\begin{cases} \alpha\theta \frac{\partial \varphi_\alpha}{\partial \theta}(\theta, x) = \inf_{a \in A(x)} \left[\int_S q(dy|x, a)\varphi_\alpha(\theta, y) + \theta c(x, a)\varphi_\alpha(\theta, x) \right], \\ 1 \leq \varphi_\alpha(\theta, x) \leq \frac{\alpha^2}{\alpha^2 - \rho_0\rho_1\theta} e^{\theta L_0/\alpha} (\mathcal{V}(x))^{\frac{\rho_1\theta}{\alpha}} \text{ for } (\theta, x) \in [0, 1] \times S, \end{cases} \tag{3.1}$$

for each $x \in S$ and a.e. $\theta \in [0, 1]$ where the upper bound of $\varphi_\alpha(\theta, x)$ is motivated by Proposition 2.1.

Remark 3.1 To prove the existence of an optimal control for bounded cost and transition rates, in Pal and Pradhan (2019), the authors studied the following HJB equation having a solution $\phi_\alpha(\theta, x)$ on $[0, 1] \times S$ such that

$$\begin{cases} \alpha\theta \frac{\partial \phi_\alpha}{\partial \theta}(\theta, x) = \inf_{a \in A(x)} \left[\int_S q(dy|x, a)\phi_\alpha(\theta, y) + \theta c(x, a)\phi_\alpha(\theta, x) \right], & \text{for } (\theta, x) \in [0, 1] \times S, \\ \lim_{\theta \rightarrow 0} \phi_\alpha(\theta, x) = 1 & \text{uniformly in } x \in S. \end{cases} \tag{3.2}$$

From the arguments for the existence of a unique solution to the Eq. (3.2), it is necessary to have $\phi_\alpha(\theta, x)$ converges to 1 uniformly in x as $\theta \rightarrow 0$. But, it is not true in general when the cost and transition rates are unbounded; for more details see Example 3.2 in Guo and Liao (2019). In this article we replace the uniform convergence condition by the above new one.

To ensure the existence of an optimal control, in addition to Assumptions 2.1 and 2.2, we also need the following continuity and compactness conditions.

Assumption 3.1 The following conditions hold:

- (i) for each $x \in S$, the set $A(x)$ is compact;
- (ii) for any fixed $x \in S$, the function $c(x, a)$ is continuous in $a \in A(x)$;
- (iii) for any given $x \in S$, the function $\int_S \mathcal{V}(y)q(dy|x, a)$ is continuous in $a \in A(x)$, where \mathcal{V} is introduced in Assumption 2.1.

Remark 3.2 Assumptions 3.1 (i)–(iii) are commonly used to find an optimal control for continuous-time MDP, see Guo and Hernandez-Lerma (2009), Guo and Liao (2019), Guo et al. (2019), Guo and Piunovskiy (2011) and Guo and Zhang (2019). Also, note that if Assumption 3.1 (iii) is satisfied, then for any given $x \in S$, the function $\int_S u(y)q(dy|x, a)$ is continuous in $a \in A(x)$ for each function $u \in B_{\mathcal{V}}(S)$.

In the next theorem we show that if the HJB equation has a solution then its stochastic representation is equal to the value function corresponding to the cost criterion (2.3).

Theorem 3.1 Under Assumptions 2.1, 2.2, and 3.1 suppose that the HJB Eq. (3.1) has a solution $\varphi_\alpha \in L^\infty_{\mathcal{V}}([0, 1] \times S)$ satisfying the bounds as in Eq. (3.1). Then, for all $(\theta, x) \in [0, 1] \times S$, we have the probabilistic representation of φ_α as

$$\varphi_\alpha(\theta, x) = \inf_{\pi \in \Pi} E_x^\pi \left[\exp \left(\theta \int_0^\infty e^{-\alpha t} c(\xi_t, \pi_t) dt \right) \right] \tag{3.3}$$

i.e., $\varphi_\alpha(\theta, x) = \tilde{J}_\alpha^*(\theta, x)$ for all $(\theta, x) \in [0, 1] \times S$.

Proof First, we see that

$$\left[\theta c(x, a)\varphi_\alpha(\theta, x) + \int_S q(dy|x, a)\varphi_\alpha(\theta, y) \right]$$

is continuous in $a \in A(x)$ and $A(x)$ is compact. So by measurable selection theorem, [Bertsekas and Shreve (1996), Proposition 7.33], there exists a measurable function $f^* : [0, 1] \times S \rightarrow A$ such that

$$\begin{aligned} & \inf_{a \in A(x)} \left[\theta c(x, a) \varphi_\alpha(\theta, x) + \int_S q(dy|x, a) \varphi_\alpha(\theta, y) \right] \\ &= \left[\theta c(x, f^*(\theta, x)) \varphi_\alpha(\theta, x) + \int_S q(dy|x, f^*(\theta, x)) \varphi_\alpha(\theta, y) \right]. \end{aligned} \tag{3.4}$$

Let

$$\pi^* : S \times \mathbb{R}_+ \rightarrow A$$

be defined by

$$\pi_t^*(x) := f^*(\theta e^{-\alpha t}, x).$$

Now we observe from Eq. (3.1) that for any $x \in S, a \in A(x)$ and a.e. $\theta \in [0, 1]$ that

$$-\alpha \theta \frac{\partial \varphi_\alpha}{\partial \theta}(\theta, x) + \left[\int_S q(dy|x, a) \varphi_\alpha(\theta, y) + \theta c(x, a) \varphi_\alpha(\theta, x) \right] \geq 0. \tag{3.5}$$

For any history-dependent control $\pi \in \Pi$ and $\theta \in [0, 1]$, let $\{\xi_t, t \geq 0\}$ be the corresponding process, and define $\theta(t) := \theta e^{-\alpha t}$. Now for each $\omega \in \Omega$, by Eq. (3.5), we get for a.e. $s \geq 0$,

$$-\alpha \theta(s) \frac{\partial \varphi_\alpha}{\partial \theta}(\theta(s), \xi_s) + \left[\int_S q(dy|\xi_s, \pi_s) \varphi_\alpha(\theta(s), y) + \theta(s) c(\xi_s, \pi_s) \varphi_\alpha(\theta(s), \xi_s) \right] \geq 0. \tag{3.6}$$

Define a function $g : [0, \infty) \times S \times \Omega \rightarrow [0, \infty)$ by

$$g(t, x, \omega) := \exp\left(\int_0^t \theta(s) c(\xi_s, \pi_s) ds\right) \varphi_\alpha(\theta(t), x).$$

In view of Assumptions 2.1 and 2.2, we have

$$\begin{aligned} & E_x^\pi \left[\exp\left(\int_0^t 2e^{-\alpha s} c(\xi_s, \pi_s) ds\right) \right] \\ & \leq E_x^\pi \left[\exp\left(\int_0^\infty \frac{2}{\alpha} c(\xi_s, \pi_s) d(-e^{-\alpha s})\right) \right] \\ & \leq E_x^\pi \left[\int_0^\infty \exp\left(\frac{2}{\alpha} c(\xi_s, \pi_s)\right) d(-e^{-\alpha s}) \right] \\ & \leq E_x^\pi \left[\int_0^\infty \exp\left(\frac{2}{\alpha} (\rho_1 \log \mathcal{V}(\xi_s) + L_0)\right) d(-e^{-\alpha s}) \right] \end{aligned}$$

$$\begin{aligned}
 & \text{(by Assumption 2.1)} \\
 & = e^{2L_0/\alpha} \left[\int_0^\infty E_x^\pi \left(\mathcal{V}(\xi_s)^{\frac{2\rho_1}{\alpha}} \right) d(-e^{-\alpha s}) \right] \\
 & \leq \alpha e^{2L_0/\alpha} \left(\mathcal{V}^2(x) + \frac{b_0}{\rho_2} \right) \left[\int_0^\infty e^{\rho_2 s - \alpha s} ds \right] \\
 & = \frac{\alpha e^{2L_0/\alpha}}{\alpha - \rho_2} \left(\mathcal{V}^2(x) + \frac{b_0}{\rho_2} \right), \tag{3.7}
 \end{aligned}$$

where second inequality is obtained by using Jensen’s inequality.

Hence $E_x^\pi \left[\exp \left(\int_0^t 2e^{-\alpha s} c(\xi_s, \pi_s) ds \right) \right] < \infty$ for all $x \in S$ and $t \in (0, \infty)$. Thus, using the extension of Dynkin formula in Guo et al. (2019), Theorem 3.1 to the function g , we have

$$\begin{aligned}
 & E_x^\pi [g(t, \xi_t, \omega)] - \varphi_\alpha(\theta, x) \\
 & = E_x^\pi \left\{ \int_0^t \exp \left(\int_0^s \theta(v) c(\xi_v, \pi_v) dv \right) \times \left[-\alpha \theta(s) \frac{\partial \varphi_\alpha}{\partial \theta}(\theta(s), \xi_s) \right. \right. \\
 & \quad \left. \left. + \int_S q(dy | \xi_s, \pi_s) \varphi_\alpha(\theta(s), y) + \theta(s) c(\xi_s, \pi_s) \varphi_\alpha(\theta(s), \xi_s) \right] ds \right\}. \tag{3.8}
 \end{aligned}$$

Now from (3.6) and (3.8), we have

$$\varphi_\alpha(\theta, x) \leq E_x^\pi \left[\exp \left(\int_0^t \theta(s) c(\xi_s, \pi_s) ds \right) \varphi_\alpha(\theta(t), \xi_t) \right]. \tag{3.9}$$

Given any $p > 1$, let $q > 1$ such that $\frac{1}{p} + \frac{1}{q} = 1$, by Holder’s inequality we have

$$\begin{aligned}
 \varphi_\alpha(\theta, x) & \leq E_x^\pi \left[\exp \left(\int_0^t \theta(s) c(\xi_s, \pi_s) ds \right) \varphi_\alpha(\theta(t), \xi_t) \right] \\
 & \leq \left\{ E_x^\pi \left[\exp \left(p \int_0^t \theta(s) c(\xi_s, \pi_s) ds \right) \right] \right\}^{1/p} \times \left\{ E_x^\pi [\varphi_\alpha^q(\theta(t), \xi_t)] \right\}^{1/q} \\
 & =: T_1(p, t) \cdot T_2(q, t). \tag{3.10}
 \end{aligned}$$

For $T_2(q, t) := \{E_x^\pi[\varphi_\alpha^q(\theta(t), \xi_t)]\}^{1/q}$, by the upper bound of φ_α in (3.1), we have

$$\varphi_\alpha(\theta(t), \xi_t) = \varphi_\alpha(\theta e^{-\alpha t}, \xi_t) \leq \frac{\alpha^2}{\alpha^2 - \theta e^{-\alpha t} \rho_0 \rho_1} \exp \left(\frac{\theta e^{-\alpha t} L_0}{\alpha} \right) [\mathcal{V}(\xi_t)]^{\frac{\rho_1 \theta e^{-\alpha t}}{\alpha}}.$$

If $t > \alpha^{-1} \log(\theta q \rho_1 / \alpha)$ then $\theta e^{-\alpha t} q \rho_1 / \alpha < 1$. Applying Jensen’s inequality and Proposition 2.1(b), we get

$$\begin{aligned}
 T_2(q, t) &\leq \left\{ E_x^\pi \left[\left(\frac{\alpha^2}{\alpha^2 - \theta e^{-\alpha t} \rho_0 \rho_1} \right)^q \exp\left(\frac{q\theta e^{-\alpha t} L_0}{\alpha} \right) [\mathcal{V}(\xi_t)]^{\frac{q\rho_1\theta e^{-\alpha t}}{\alpha}} \right] \right\}^{1/q} \\
 &= \frac{\alpha^2}{\alpha^2 - \theta e^{-\alpha t} \rho_0 \rho_1} \exp\left(\frac{\theta e^{-\alpha t} L_0}{\alpha} \right) \left[E_x^\pi [\mathcal{V}^{\frac{q\rho_1\theta e^{-\alpha t}}{\alpha}}(\xi_t)] \right]^{\frac{1}{q}} \\
 &\leq \frac{\alpha^2}{\alpha^2 - \theta e^{-\alpha t} \rho_0 \rho_1} \exp\left(\frac{\theta e^{-\alpha t} L_0}{\alpha} \right) [E_x^\pi(\mathcal{V}(\xi_t))]^{\frac{\rho_1\theta e^{-\alpha t}}{\alpha}} \\
 &\leq \frac{\alpha^2}{\alpha^2 - \theta e^{-\alpha t} \rho_0 \rho_1} \exp\left(\frac{\theta e^{-\alpha t}}{\alpha} (L_0 + \rho_0 \rho_1 t) \right) \mathcal{V}^{\frac{\theta e^{-\alpha t} \rho_1}{\alpha}}(x) =: T_3(t).
 \end{aligned}
 \tag{3.11}$$

Next take $t \rightarrow \infty$ and get

$$T_1(p, t) \rightarrow \left\{ E_x^\pi \left[\exp\left(p \int_0^\infty \theta(s)c(\xi_s, \pi_s) ds \right) \right] \right\}^{1/p} \text{ and } T_3(t) \rightarrow 1. \tag{3.12}$$

By (3.10), (3.11) and (3.12) we obtain

$$\varphi_\alpha(\theta, x) \leq \left\{ E_x^\pi \left[\exp\left(p\theta \int_0^\infty e^{-\alpha t} c(\xi_t, \pi_t) dt \right) \right] \right\}^{1/p}.$$

Now, take the limit as $p \downarrow 1$ and get the result

$$\varphi_\alpha(\theta, x) \leq E_x^\pi \left[\exp\left(\theta \int_0^\infty e^{-\alpha t} c(\xi_t, \pi_t) dt \right) \right].$$

Since $\pi \in \Pi$ is an arbitrary control, we have

$$\varphi_\alpha(\theta, x) \leq \inf_{\pi \in \Pi} E_x^\pi \left[\exp\left(\theta \int_0^\infty e^{-\alpha t} c(\xi_t, \pi_t) dt \right) \right]. \tag{3.13}$$

Using (3.1), (3.4) and (3.10), we can show that

$$E_x^{\pi^*} \left[\exp\left(\int_0^t \theta(s)c(\xi_s, \pi_s^*(\xi_{s-})) ds \right) \varphi_\alpha(\theta(t), \xi_t) \right] = \varphi_\alpha(\theta, x). \tag{3.14}$$

Now, using the lower bound of φ_α in (3.1) and Fatou’s lemma, we obtain

$$\begin{aligned}
 &\liminf_{t \rightarrow \infty} E_x^{\pi^*} \left[\exp\left(\int_0^t \theta(s)c(\xi_s, \pi_s^*(\xi_{s-})) ds \right) \varphi_\alpha(\theta(t), \xi_t) \right] \\
 &\geq \liminf_{t \rightarrow \infty} E_x^{\pi^*} \left[\exp\left(\int_0^t \theta(s)c(\xi_s, \pi_s^*(\xi_{s-})) ds \right) \right] \\
 &\geq E_x^{\pi^*} \left[\liminf_{t \rightarrow \infty} \exp\left(\int_0^t \theta(s)c(\xi_s, \pi_s^*(\xi_{s-})) ds \right) \right] \\
 &= \tilde{J}_\alpha(\theta, x, \pi^*).
 \end{aligned}
 \tag{3.15}$$

From (3.14) and (3.15), we have

$$\tilde{J}_\alpha(\theta, x, \pi^*) \leq \varphi_\alpha(\theta, x).$$

Thus

$$\inf_{\pi \in \Pi} \tilde{J}_\alpha(\theta, x, \pi) \leq \tilde{J}_\alpha(\theta, x, \pi^*) \leq \varphi_\alpha(\theta, x). \tag{3.16}$$

From (3.13) and (3.16), we have (3.3). □

4 The existence of solution to the HJB equation

In this Section, we prove that the Eq. (3.1) is the HJB equation for the α discounted cost (2.3) and the Eq. (3.1) has a solution in $L^\infty_{\mathcal{V}}([0, 1] \times S)$. We now proceed to make a rigorous analysis of the above. First, we prove a lemma about the existence of a solution for the HJB equation for bounded transition and cost rates; see Lemma 4.1 below. Then in Theorem 4.1, we relax these boundedness condition and prove the existence of a solution to the HJB Eq. (3.1). For that, we first truncate our transition and cost rates which plays a crucial role to derive the HJB equations and find the solution. Fix any $n \geq 1, 0 < \delta < 1$. For each $n \geq 1, x \in S, a \in A(x)$, let $A_n(x) := A(x), S_n := \{x \in S | \mathcal{V}(x) \leq n\}$, and $K_n := \{(x, a) | x \in S_n, a \in A_n(x)\}$. Moreover for each $x \in S, a \in A_n(x)$ define

$$q^{(n)}(dy|x, a) := \begin{cases} q(dy|x, a) & \text{if } x \in S_n, \\ 0 & \text{if } x \notin S_n \end{cases} \tag{4.1}$$

and

$$c_n(x, a) := \begin{cases} c(x, a) \wedge \min\{n, \rho_1 \ln \mathcal{V}(x) + L_0\} & \text{if } x \in S_n, \\ 0 & \text{if } x \notin S_n. \end{cases} \tag{4.2}$$

Lemma 4.1 *Grant Assumptions 2.1, 2.2 and 3.1. Then, there exists a unique function $\varphi_\alpha^{(n,\delta)}$ (depending on n, δ) in $L^\infty_{\mathcal{V}}([0, 1] \times S)$ for which the followings are true :*

- (1) $\varphi_\alpha^{(n,\delta)} \in B_1([0, 1] \times S)$ is a bounded solution to the following differential equations (DEs) for all $x \in S$ and a.e. $\theta \in (\delta, 1]$:

$$\begin{cases} \alpha \theta \frac{\partial \varphi_\alpha^{(n,\delta)}}{\partial \theta}(\theta, x) = \inf_{a \in A(x)} \left[\theta c_n(x, a) \varphi_\alpha^{(n,\delta)}(\theta, x) + \int_S q^{(n)}(dy|x, a) \varphi_\alpha^{(n,\delta)}(\theta, y) \right] \\ \varphi_\alpha^{(n,\delta)}(\delta, x) = e^{n\delta/\alpha}. \end{cases} \tag{4.3}$$

- (2) $\varphi_\alpha^{(n,\delta)}(\theta, x)$ has a stochastic representation as follows: for each $x \in S$ and a.e. $\theta \in (\delta, 1]$,

$$\varphi_\alpha^{(n,\delta)}(\theta, x) = \inf_{\pi \in \Pi} E_x^\pi \left[e^{n\delta/\alpha} \exp \left(\theta \int_0^{T_\delta(\theta)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \right], \tag{4.4}$$

where $T_\delta(\theta) := \alpha^{-1} \log(\theta/\delta)$ and $\{\xi_t^{(n)}\}_{t \geq 0}$ is the process corresponding to the $q^{(n)}(\cdot|x, a)$.

Proof (1) Since $S_n := \{x \in S | \mathcal{V}(x) \leq n\}$, by Assumption 2.1(ii), we see that $q_x^{(n)}(a) := \int_{S/\{x\}} q^{(n)}(dy|x, a)$ is bounded. So we can use the Lyapunov function $V \equiv 1$ such that $\int_S q^{(n)}(dy|x, a)V(y) \leq \rho_0 V(x)$, and $\bar{q}^{(n)} := \sup_{(x,a) \in K} q_x^{(n)}(a) < \infty$. Now let us define a nonlinear operator T on $B_1([0, 1] \times S)$ as follows:

$$Tu(\theta, x) = e^{\delta n/\alpha} + \frac{1}{\alpha} \int_\delta^\theta \inf_{a \in A(x)} \left[\frac{1}{s} \int_S q^{(n)}(dy|x, a)u(s, y) + c_n(x, a)u(s, x) \right] ds,$$

where $u \in B_1([0, 1] \times S)$ and $(\theta, x) \in [\delta, 1] \times S$. By using the Assumption 2.1 and the fact that c_n is bounded, we obtain

$$\begin{aligned} & \sup_{\theta \in [\delta, 1]} \sup_{x \in S} |Tu(\theta, x)| \\ & \leq e^{\delta n/\alpha} + \frac{1}{\alpha} \int_\delta^1 \sup_{a \in A(x)} \left\{ \frac{1}{s} \sup_{x \in S} \left[\int_S |q^{(n)}(dy|x, a)| |u(s, y)| \right] + n \sup_{x \in S} |u(s, x)| \right\} ds \\ & \leq e^{\delta n/\alpha} + \frac{\|u\|_1^\infty}{\alpha} \left\{ \int_\delta^1 \sup_{a \in A(x)} \frac{1}{s} \sup_{x \in S} \left(2q_x^{(n)}(a) \right) ds + n(1 - \delta) \right\} \\ & \leq e^{\delta n/\alpha} + \frac{1}{\alpha} \left[(-2)\bar{q}^{(n)} \log \delta + n(1 - \delta) \right] \|u\|_1^\infty. \end{aligned}$$

Therefore, T is a nonlinear operator from $B_1([0, 1] \times S)$ to $B_1([0, 1] \times S)$. For any $g_1, g_2 \in B_1([0, 1] \times S)$ and $\theta \in [\delta, 1]$, we have

$$\begin{aligned} \sup_{x \in S} |Tg_1(t, x) - Tg_2(t, x)| & \leq \frac{1}{\alpha} \int_\delta^t \left(2\bar{q}^{(n)}/s + n \right) \sup_{x \in S} |g_1(s, x) - g_2(s, x)| ds \\ & \leq \frac{1}{\alpha} \left[2\bar{q}^{(n)}(\log t - \log \delta) + n(t - \delta) \right] \|g_1 - g_2\|_1^\infty. \end{aligned} \tag{4.5}$$

Now, we prove the following:

$$\begin{aligned} & \sup_{x \in S} |T^l g_1(t, x) - T^l g_2(t, x)| \\ & \leq \frac{\|g_1 - g_2\|_1^\infty}{\alpha^l \cdot l!} \left[2\bar{q}^{(n)}(\log t - \log \delta) + n(t - \delta) \right]^l \quad \forall l \geq 1. \end{aligned} \tag{4.6}$$

By (4.5) and (4.6) we have

$$\begin{aligned} & \sup_{x \in S} |T^{l+1} g_1(t, x) - T^{l+1} g_2(t, x)| \\ & \leq \frac{1}{\alpha} \int_{\delta}^t \left(2\bar{q}^{(n)}/s + n \right) \sup_{x \in S} |T^l g_1(s, x) - T^l g_2(s, x)| ds \\ & \leq \frac{\|g_1 - g_2\|_1^\infty}{\alpha^{l+1} \cdot l!} \int_{\delta}^t \left(2\bar{q}^{(n)}/s + n \right) \left[2\bar{q}^{(n)}(\log s - \log \delta) + n(s - \delta) \right]^l ds \\ & = \frac{\|g_1 - g_2\|_1^\infty}{\alpha^{l+1} \cdot (l + 1)!} \left[2\bar{q}^{(n)}(\log t - \log \delta) + n(t - \delta) \right]^{l+1}. \end{aligned}$$

Since $\sum_{k \geq 1} \frac{1}{\alpha^k \cdot k!} \left[-2\bar{q}^{(n)} \log \delta + n(1 - \delta) \right]^k < \infty$, there exists some m such that $\beta := \frac{1}{\alpha^m \cdot m!} \left[-2\bar{q}^{(n)} \log \delta + n(1 - \delta) \right]^m < 1$, which implies that $\|T^m g_1 - T^m g_2\|_1^\infty \leq \beta \|g_1 - g_2\|_1^\infty$. Therefore, T is a m -step contraction operator on $B_1([0, 1] \times S)$. So, by Banach fixed point theorem, there exists a unique bounded function $\varphi_\alpha^{(n, \delta)} \in B_1([0, 1] \times S)$ (depending on (n, δ)) such that $T\varphi_\alpha^{(n, \delta)} = \varphi_\alpha^{(n, \delta)}$; that is,

$$\begin{aligned} \varphi_\alpha^{(n, \delta)}(\theta, x) &= e^{\delta n/\alpha} + \frac{1}{\alpha} \int_{\delta}^{\theta} \inf_{a \in A(x)} \left[\frac{1}{s} \int_S q^{(n)}(dy|x, a) \varphi_\alpha^{(n, \delta)}(s, y) \right. \\ & \quad \left. + c_n(x, a) \varphi_\alpha^{(n, \delta)}(s, x) \right] ds. \end{aligned}$$

Also note that $\varphi_\alpha^{(n, \delta)}(\delta, x) = e^{\delta n/\alpha}$. Hence by using (4.1), (4.2) and the above equation, we have $\varphi_\alpha^{(n, \delta)} \in L^\infty_{\mathcal{V}}([0, 1] \times S)$ and it satisfies equation (4.3).

(2) First we see that

$$\left[\theta c_n(x, a) \varphi_\alpha^{(n, \delta)}(\theta, x) + \int_S q^{(n)}(dy|x, a) \varphi_\alpha^{(n, \delta)}(\theta, y) \right]$$

is continuous in $a \in A(x)$ and $A(x)$ is compact. So by measurable selection theorem (Bertsekas and Shreve 1996), Proposition 7.33, there exists a measurable function $f^{*\delta} : [0, 1] \times S \rightarrow A$ such that

$$\begin{aligned} & \inf_{a \in A(x)} \left[\theta c_n(x, a) \varphi_\alpha^{(n, \delta)}(\theta, x) + \int_S q^{(n)}(dy|x, a) \varphi_\alpha^{(n, \delta)}(\theta, y) \right] \\ & = \left[\theta c_n(x, f^{*\delta}(\theta, x)) \varphi_\alpha^{(n, \delta)}(\theta, x) + \int_S q^{(n)}(dy|x, f^{*\delta}(\theta, x)) \varphi_\alpha^{(n, \delta)}(\theta, y) \right]. \end{aligned} \tag{4.7}$$

Let

$$\pi^{*\delta} : S \times \mathbb{R}_+ \rightarrow A$$

be defined by

$$\pi_t^{*\delta}(x) := f^{*\delta}(\theta e^{-\alpha t}, x).$$

Let $\theta(t) := \theta e^{-\alpha t}$ for $t \in [0, \infty)$. Since c_n and $\varphi_\alpha^{(n,\delta,k)}$ are bounded, by Dynkin’s formula we get

$$\begin{aligned} & E_x^\pi \left[\exp \left(\int_0^{T_\delta(\theta)} \theta(s) c_n(\xi_s^{(n)}, \pi_s) ds \right) \varphi_\alpha^{(n,\delta)} \left(\theta(T_\delta), \xi_{T_\delta}^{(n)} \right) \right] - \varphi_\alpha^{(n,\delta)}(\theta, x) \\ &= E_x^\pi \left\{ \int_0^{T_\delta(\theta)} \left[-\alpha \theta(s) \frac{\partial \varphi_\alpha^{(n,\delta)}}{\partial \theta}(\theta(s), \xi_s^{(n)}) + \int_S q^{(n)}(dy | \xi_s^{(n)}, \pi_s) \varphi_\alpha^{(n,\delta)}(\theta(s), y) \right. \right. \\ &\quad \left. \left. + \theta(s) c_n(\xi_s^{(n)}, \pi_s) \varphi_\alpha^{(n,\delta)}(\theta(s), \xi_s^{(n)}) \right] \times \exp \left(\int_0^s \theta(v) c_n(\xi_v^{(n)}, \pi_v) dv \right) ds \right\}. \end{aligned} \tag{4.8}$$

By using (4.3) and (4.8), we obtain

$$E_x^\pi \left[\exp \left(\int_0^{T_\delta(\theta)} \theta(s) c_n(\xi_s^{(n)}, \pi_s) ds \right) \varphi_\alpha^{(n,\delta)} \left(\theta(T_\delta), \xi_{T_\delta}^{(n)} \right) \right] \geq \varphi_\alpha^{(n,\delta)}(\theta, x).$$

Since $\pi \in \Pi$ is an arbitrary control and $\varphi_\alpha^{(n,\delta)}(\theta(T_\delta(\theta)), \xi_{T_\delta}^{(n)}) = e^{n\delta/\alpha}$, we have

$$\varphi_\alpha^{(n,\delta)}(\theta, x) \leq \inf_{\pi \in \Pi} E_x^\pi \left[e^{n\delta/\alpha} \exp \left(\int_0^{T_\delta(\theta)} \theta(s) c_n(\xi_s^{(n)}, \pi_s) ds \right) \right]. \tag{4.9}$$

Using Eqs. (4.3), (4.7) and (4.8), we can show that

$$\varphi_\alpha^{(n,\delta)}(\theta, x) = E_x^{\pi^{*\delta}} \left[e^{n\delta/\alpha} \exp \left(\int_0^{T_\delta(\theta)} \theta(s) c_n(\xi_s^{(n)}, \pi_s^{*\delta}(\xi_s^{(n)})) ds \right) \right].$$

Therefore

$$\varphi_\alpha^{(n,\delta)}(\theta, x) \geq \inf_{\pi \in \Pi} E_x^\pi \left[e^{n\delta/\alpha} \exp \left(\int_0^{T_\delta(\theta)} \theta(s) c_n(\xi_s^{(n)}, \pi_s) ds \right) \right]. \tag{4.10}$$

Therefore, from (4.9) and (4.10), we obtain (4.4). This completes the proof. □

Theorem 4.1 *Grant Assumptions 2.1, 2.2 and 3.1. Then the HJB Eq. (3.1) has a unique solution $\varphi_\alpha \in L^\infty([0, 1] \times S)$ satisfying $1 \leq \varphi_\alpha(\theta, x) \leq \frac{\alpha^2 e^{\theta L_0/\alpha}}{\alpha^2 - \rho_0 \rho_1 \theta} (\mathcal{V}(x)) \frac{\rho_1 \theta}{\alpha}$ for all $(\theta, x) \in [0, 1] \times S$.*

Proof First note that $\varphi_\alpha^{(n,\delta)}$ is the solution to the Eq.(4.3), which depends on two parameters n, δ . We prove this theorem in two steps.

Step 1: In the first step, we construct a solution $\varphi_\alpha^{(n)}(\cdot, x)$ from $\varphi_\alpha^{(n,\delta)}(\cdot, x)$ by passing the limit as $\delta \rightarrow 0$, such that $\varphi_\alpha^{(n)}(\cdot, x)$ is an absolutely continuous function and satisfies the following DEs:

$$\left\{ \begin{aligned} \alpha\theta \frac{\partial \varphi_\alpha^{(n)}}{\partial \theta}(\theta, x) &= \inf_{a \in A(x)} \left[\int_S q^{(n)}(dy|x, a) \varphi_\alpha^{(n)}(\theta, y) + \theta c_n(x, a) \varphi_\alpha^{(n)}(\theta, x) \right], \\ &x \in S, \text{ a.e., } \theta \in [0, 1], \\ 1 \leq \varphi_\alpha^{(n)}(\theta, x) &\leq \frac{\alpha^2 e^{\theta L_0/\alpha}}{\alpha^2 - \rho_0 \rho_1 \theta} (\mathcal{V}(x)) \frac{\rho_1 \theta}{\alpha} \quad \forall (\theta, x) \in [0, 1] \times S. \end{aligned} \right. \tag{4.11}$$

Given $0 < \delta < 1$ and $1 \leq n < \infty$ by (4.4) and $\sup_{(x,a) \in K} c_n(x, a) \leq n$, we have

$$\varphi_\alpha^{(n,\delta)}(\theta, x) \leq e^{2n/\alpha}, \quad x \in S, \theta \in [\delta, 1].$$

Next, we extend the domain of $\varphi_\alpha^{(n,\delta)}$ to $[0, 1] \times S$ by

$$\bar{\varphi}_\alpha^{(n,\delta)}(\theta, x) = \begin{cases} \varphi_\alpha^{(n,\delta)}(\theta, x), & \delta \leq \theta \leq 1 \quad \forall x \in S \\ e^{n\delta/\alpha}, & 0 \leq \theta < \delta \quad \forall x \in S. \end{cases}$$

We consider the following expression, for any given $\pi \in \Pi, x \in S, \theta, \theta_0 \in [\delta, 1]$:

$$\begin{aligned} &\left| E_x^\pi \left[e^{n\delta/\alpha} \exp\left(\theta \int_0^{T_\delta(\theta)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt\right) \right] \right. \\ &\quad \left. - E_x^\pi \left[e^{n\delta/\alpha} \exp\left(\theta_0 \int_0^{T_\delta(\theta_0)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt\right) \right] \right| \\ &\leq P_1 + P_2, \end{aligned}$$

where

$$\begin{aligned} P_1 &:= \left| E_x^\pi \left[e^{n\delta/\alpha} \exp\left(\theta \int_0^{T_\delta(\theta)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt\right) \right] \right. \\ &\quad \left. - E_x^\pi \left[e^{n\delta/\alpha} \exp\left(\theta_0 \int_0^{T_\delta(\theta)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt\right) \right] \right|, \end{aligned}$$

and

$$\begin{aligned} P_2 &:= \left| E_x^\pi \left[e^{n\delta/\alpha} \exp\left(\theta_0 \int_0^{T_\delta(\theta)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt\right) \right] \right. \\ &\quad \left. - E_x^\pi \left[e^{n\delta/\alpha} \exp\left(\theta_0 \int_0^{T_\delta(\theta_0)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt\right) \right] \right|. \end{aligned}$$

Consider $c \wedge d := \min\{c, d\}$ and $c \vee d := \max\{c, d\}$. Then for fix $n \geq 1$; we have

$$\int_0^{T_\delta(\theta)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \leq n \int_0^{T_\delta(\theta)} e^{-\alpha t} dt \leq \frac{n}{\alpha}$$

and

$$\begin{aligned} \int_{T_\delta(\theta \wedge \theta_0)}^{T_\delta(\theta \vee \theta_0)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt &\leq n \int_{T_\delta(\theta \wedge \theta_0)}^{T_\delta(\theta \vee \theta_0)} e^{-\alpha t} dt \\ &= \frac{n}{\alpha} [\exp(-\alpha T_\delta(\theta \wedge \theta_0)) - \exp(-\alpha T_\delta(\theta \vee \theta_0))] \leq \frac{\delta n |\theta_0 - \theta|}{\alpha \theta \theta_0}. \end{aligned}$$

Using the above results and knowing the fact that $e^{bz} - 1 \leq (e^b - 1)z$ for all $z \in [0, 1]$ and $b > 0$, we obtain

$$\begin{aligned} P_1 &= e^{n\delta/\alpha} E_x^\pi \left[\exp\left((\theta \wedge \theta_0) \int_0^{T_\delta(\theta)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \right. \\ &\quad \times \left. \left(\exp\left(|\theta_0 - \theta| \int_0^{T_\delta(\theta)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) - 1 \right) \right] \\ &\leq e^{2n/\alpha} E_x^\pi \left[\exp\left(|\theta_0 - \theta| \int_0^{T_\delta(\theta)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) - 1 \right] \\ &\leq e^{2n/\alpha} \left(\exp\left(\frac{n}{\alpha} |\theta_0 - \theta| \right) - 1 \right) \\ &\leq e^{2n/\alpha} (e^{n/\alpha} - 1) |\theta_0 - \theta|. \end{aligned}$$

Similarly for P_2 we have

$$\begin{aligned} P_2 &= e^{n\delta/\alpha} E_x^\pi \left[\exp\left(\theta_0 \int_0^{T_\delta(\theta \wedge \theta_0)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \right. \\ &\quad \times \left. \left(\exp\left(\theta_0 \int_{T_\delta(\theta \wedge \theta_0)}^{T_\delta(\theta \vee \theta_0)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) - 1 \right) \right] \\ &\leq e^{2n/\alpha} E_x^\pi \left[\exp\left(\theta_0 \int_{T_\delta(\theta \wedge \theta_0)}^{T_\delta(\theta \vee \theta_0)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) - 1 \right] \\ &\leq e^{2n/\alpha} \left(\exp\left(\frac{n\delta |\theta - \theta_0|}{\alpha \theta} \right) - 1 \right) \\ &\leq e^{2n/\alpha} (e^{n/\alpha} - 1) |\theta_0 - \theta|. \end{aligned}$$

Hence for all $(\theta, x) \in [0, 1] \times S$, we have

$$|\bar{\varphi}_\alpha^{(n,\delta)}(\theta_0, x) - \bar{\varphi}_\alpha^{(n,\delta)}(\theta, x)| \leq 2e^{2n/\alpha} (e^{n/\alpha} - 1) |\theta - \theta_0|. \tag{4.12}$$

Now we want to show that $\bar{\varphi}_\alpha^{(n,\delta)}$ is decreasing as $\delta \rightarrow 0$ for any (θ, x) . For a fixed $\alpha > 0$ and $\varepsilon > 0$ small enough, consider $\bar{\varphi}_\alpha^{(n,\delta+\varepsilon)}(\theta, x) - \bar{\varphi}_\alpha^{(n,\delta)}(\theta, x)$ and assume that $h_\delta := e^{\frac{n\delta}{\alpha}}$. By measurable selection theorem we get the minimizer $\pi^{*(\delta+\varepsilon)}$ like in Eq. (3.4), corresponding to $\bar{\varphi}_\alpha^{(n,\delta+\varepsilon)}$ such that the followings cases hold.

Case 1. If $\delta + \varepsilon < \theta$ then

$$\begin{aligned} & \bar{\varphi}_\alpha^{(n,\delta+\varepsilon)}(\theta, x) - \bar{\varphi}_\alpha^{(n,\delta)}(\theta, x) \\ &= E_x^{\pi^{*(\delta+\varepsilon)}} \left[h_{\delta+\varepsilon} \exp \left(\theta \int_0^{T_{\delta+\varepsilon}} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t^{*(\delta+\varepsilon)}(\xi_{t-}^{(n)})) dt \right) \right] \\ & \quad - \inf_{\pi \in \Pi} E_x^\pi \left[h_\delta \exp \left(\theta \int_0^{T_\delta} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \right] \\ &\geq h_\delta E_x^{\pi^{*(\delta+\varepsilon)}} \left[\exp \left(\theta \int_0^{T_{\delta+\varepsilon}} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t^{*(\delta+\varepsilon)}(\xi_{t-}^{(n)})) dt \right) \right. \\ & \quad \left. \times \left\{ h_\varepsilon - \exp \left(\theta \int_{T_{\delta+\varepsilon}}^{T_\delta} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t^{*(\delta+\varepsilon)}(\xi_{t-}^{(n)})) dt \right) \right\} \right] \\ &\geq h_\delta E_x^{\pi^{*(\delta+\varepsilon)}} \left[\exp \left(\theta \int_0^{T_{\delta+\varepsilon}} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t^{*(\delta+\varepsilon)}(\xi_{t-}^{(n)})) dt \right) \left\{ h_\varepsilon - \exp \left(\theta \int_{T_{\delta+\varepsilon}}^{T_\delta} e^{-\alpha t} ndt \right) \right\} \right] \\ &= h_\delta E_x^{\pi^{*(\delta+\varepsilon)}} \left[\exp \left(\theta \int_0^{T_{\delta+\varepsilon}} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t^{*(\delta+\varepsilon)}(\xi_{t-}^{(n)})) dt \right) \left\{ h_\varepsilon - \exp \left(\frac{n\theta(e^{-\alpha T_{\delta+\varepsilon}} - e^{-\alpha T_\delta})}{\alpha} \right) \right\} \right] \\ &= 0. \end{aligned}$$

Case 2. $\delta < \theta \leq \delta + \varepsilon$

$$\begin{aligned} & \bar{\varphi}_\alpha^{(n,\delta+\varepsilon)}(\theta, x) - \bar{\varphi}_\alpha^{(n,\delta)}(\theta, x) \\ &= h_{\delta+\varepsilon} - E_x^{\pi^{*\delta}} \left[h_\delta \exp \left(\theta \int_0^{T_\delta} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t^{*\delta}(\xi_{t-}^{(n)})) dt \right) \right] \\ &= h_\delta \left[h_\varepsilon - E_x^{\pi^{*\delta}} \left[\exp \left(\theta \int_0^{T_\delta} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t^{*\delta}(\xi_{t-}^{(n)})) dt \right) \right] \right] \\ &\geq h_\delta \left[h_\varepsilon - \exp \left(\theta \int_0^{T_\delta} e^{-\alpha t} ndt \right) \right] \\ &= h_\delta \left[h_\varepsilon - e^{n\theta \frac{(1-e^{-\alpha T_\delta})}{\alpha}} \right] \geq 0. \end{aligned}$$

Case 3. $\theta \leq \delta$

$$\bar{\varphi}_\alpha^{(n,\delta+\varepsilon)}(\theta, x) - \bar{\varphi}_\alpha^{(n,\delta)}(\theta, x) = h_{\delta+\varepsilon} - h_\delta = h_\delta(h_\varepsilon - 1) = h_\delta(e^{\frac{n\varepsilon}{\alpha}} - 1) \geq 0.$$

Hence $\bar{\varphi}_\alpha^{(n,\delta)}(\theta, x)$ is increasing in δ for any $(\theta, x) \in [0, 1] \times S$. Now from (4.12), we know that for each $x \in S$, $\bar{\varphi}_\alpha^{(n,\delta)}(\cdot, x)$ is Lipschitz continous in $\theta \in [0, 1]$. Also, $\bar{\varphi}_\alpha^{(n,\delta)}(\theta, x)$ is increasing in δ for any $(\theta, x) \in [0, 1] \times S$ and bounded above (since $\bar{\varphi}_\alpha^{(n,\delta)}(\theta, x) \leq e^{2n/\alpha}$, $x \in S, \theta \in [\delta, 1]$), therefore there exists a function $\varphi_\alpha^{(n)}$ on $[0, 1] \times S$ that is continuous with respect to $\theta \in [0, 1]$, such that along a subsequence

$\delta_m \rightarrow 0$, we have $\lim_{m \rightarrow \infty} \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) = \varphi_\alpha^{(n)}(\theta, x)$ and for any fixed $x \in S$ this convergence is uniform in $\theta \in [0, 1]$.

Let $\psi \in C_c^\infty(0, 1)$, then we have

$$\begin{aligned}
 & - \int_0^1 \alpha \frac{d(\theta\psi)}{d\theta}(\theta) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) d\theta = \int_0^1 \alpha \theta \frac{\partial \bar{\varphi}_\alpha^{(n, \delta_m)}}{\partial \theta}(\theta, x) \psi(\theta) d\theta \\
 & = \int_0^1 \inf_{a \in A(x)} \left[\theta c_n(x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) + \int_S q^{(n)}(dy|x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, y) \right] \psi(\theta) d\theta \\
 & - \int_0^{\delta_m} \inf_{a \in A(x)} \left[\theta c_n(x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) + \int_S q^{(n)}(dy|x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, y) \right] \psi(\theta) d\theta \\
 & = \int_0^1 \inf_{a \in A(x)} \left[\theta c_n(x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) + \int_S q^{(n)}(dy|x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, y) \right] \psi(\theta) d\theta \\
 & - \int_0^{\delta_m} \inf_{a \in A(x)} \left[\theta c_n(x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) \right] \psi(\theta) d\theta. \tag{4.13}
 \end{aligned}$$

Now take $\tau(x) := M_0 \mathcal{V}(x)$ and define

$$Q^{(n)}(dy|x, a) := \delta_x(dy) + \frac{q^{(n)}(dy|x, a)}{\tau(x)}$$

for all $(x, a) \in K$ where $\delta_x(\cdot)$ is the Dirac measure concentrated at x . We see that under Assumption 2.1, $Q^{(n)}$ is a stochastic kernel on S given K . Then (4.13) can be written as

$$\begin{aligned}
 & - \int_0^1 \left\{ \frac{\alpha}{\tau(x)} \frac{d(\theta\psi)}{d\theta} \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) - \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) \psi(\theta) \right\} d\theta \\
 & = \int_0^1 \inf_{a \in A(x)} \left[\frac{\theta}{\tau(x)} c_n(x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) + \int_S Q^{(n)}(dy|x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, y) \right] \psi(\theta) d\theta \\
 & - \frac{1}{\tau(x)} \int_0^{\delta_m} \inf_{a \in A(x)} \left[\theta c_n(x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) \right] \psi(\theta) d\theta. \tag{4.14}
 \end{aligned}$$

Now

$$\begin{aligned}
 & \left| \inf_{a \in A(x)} \left[\frac{\theta}{\tau(x)} c_n(x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, x) + \int_S Q^{(n)}(dy|x, a) \bar{\varphi}_\alpha^{(n, \delta_m)}(\theta, y) \right] \psi(\theta) \right| \\
 & \leq |\psi(\theta)| \sup_{a \in A(x)} \left[\frac{\theta}{\tau(x)} |c_n(x, a)| |\bar{\varphi}_\alpha^{(n, \delta)}(\theta, x)| + \int_S Q^{(n)}(dy|x, a) |\bar{\varphi}_\alpha^{(n, \delta)}(\theta, y)| \right] \\
 & \leq \frac{\alpha^2}{\alpha^2 - \rho_0 \rho_1 \theta} e^{\theta L_0 / \alpha} \sup_{a \in A(x)} \left[\frac{\theta}{\tau(x)} n \mathcal{V}^{\frac{\rho_1 \theta}{\alpha}}(x) + \int_S Q^{(n)}(dy|x, a) \mathcal{V}^{\frac{\rho_1 \theta}{\alpha}}(y) \right] |\psi(\theta)| \\
 & \leq \frac{\alpha^2}{\alpha^2 - \rho_0 \rho_1 \theta} e^{\theta L_0 / \alpha} \sup_{a \in A(x)} \left[\frac{\theta}{\tau(x)} n \mathcal{V}(x) + \int_S Q^{(n)}(dy|x, a) \mathcal{V}(y) \right] |\psi(\theta)|
 \end{aligned}$$

$$\begin{aligned} &\leq \frac{\alpha^2}{\alpha^2 - \rho_0\rho_1\theta} e^{\theta L_0/\alpha} \left[\frac{\theta}{\tau(x)} n\mathcal{V}(x) + \mathcal{V}(x) + \rho_0 \frac{\mathcal{V}(x)}{\tau(x)} \right] |\psi(\theta)| \\ &= \frac{\alpha^2}{\alpha^2 - \rho_0\rho_1\theta} e^{\theta L_0/\alpha} \left[\frac{\theta}{\tau(x)} n\mathcal{V}(x) + \mathcal{V}(x) + \frac{\rho_0}{M_0} \right] |\psi(\theta)|. \end{aligned} \tag{4.15}$$

Since for each fixed $x \in S$, $A(x)$ is compact, there exist a subsequence of $\{m\}$, by abuse of notation, we denote the same sequence and $a^* \in A(x)$ such that $\lim_{m \rightarrow \infty} a_m^* = a^*$. Now, from (4.13), for any $a \in A(x)$, we have

$$\begin{aligned} &- \int_0^1 \left\{ \frac{\alpha}{\tau(x)} \frac{d(\theta\psi)}{d\theta} \bar{\varphi}_\alpha^{(n,\delta_m)}(\theta, x) - \bar{\varphi}_\alpha^{(n,\delta_m)}(\theta, x) \psi(\theta) \right\} d\theta \\ &= \int_0^1 \left[\frac{\theta}{\tau(x)} c_n(x, a_m^*) \bar{\varphi}_\alpha^{(n,\delta_m)}(\theta, x) + \int_S \mathcal{Q}^{(n)}(dy|x, a_m^*) \bar{\varphi}_\alpha^{(n,\delta_m)}(\theta, y) \right] \psi(\theta) d\theta \\ &\quad - \frac{1}{\tau(x)} \int_0^{\delta_m} \inf_{a \in A(x)} \left[\theta c_n(x, a) \bar{\varphi}_\alpha^{(n,\delta_m)}(\theta, x) \right] \psi(\theta) d\theta. \end{aligned} \tag{4.16}$$

So, by Lemma 8.3.7 in Hernandez-Lerma and Lasserre (1999) (Hernandez-Lerma and Lasserre 1999) taking limit as $m \rightarrow \infty$ in (4.16), we get

$$\begin{aligned} &- \int_0^1 \left\{ \frac{\alpha}{\tau(x)} \frac{d(\theta\psi)}{d\theta}(\theta) \varphi_\alpha^{(n)}(\theta, x) - \varphi_\alpha^{(n)}(\theta, x) \psi(\theta) \right\} d\theta \\ &\geq \int_0^1 \left[\frac{\theta}{\tau(x)} c_n(x, a^*) \varphi_\alpha^{(n)}(\theta, x) + \int_S \mathcal{Q}^{(n)}(dy|x, a^*) \varphi_\alpha^{(n)}(\theta, y) \right] \psi(\theta) d\theta. \end{aligned}$$

Hence

$$\begin{aligned} &- \int_0^1 \left\{ \frac{\alpha}{\tau(x)} \frac{d(\theta\psi)}{d\theta}(\theta) \varphi_\alpha^{(n)}(\theta, x) - \varphi_\alpha^{(n)}(\theta, x) \psi(\theta) \right\} d\theta \\ &\geq \inf_{a \in A(x)} \int_0^1 \left[\frac{\theta}{\tau(x)} c_n(x, a) \varphi_\alpha^{(n)}(\theta, x) + \int_S \mathcal{Q}^{(n)}(dy|x, a) \varphi_\alpha^{(n)}(\theta, y) \right] \psi(\theta) d\theta. \end{aligned} \tag{4.17}$$

But

$$\begin{aligned} &- \int_0^1 \left\{ \frac{\alpha}{\tau(x)} \frac{d(\theta\psi)}{d\theta} \bar{\varphi}_\alpha^{(n,\delta_m)}(\theta, x) - \bar{\varphi}_\alpha^{(n,\delta_m)}(\theta, x) \psi(\theta) \right\} d\theta \\ &\leq \int_0^1 \left[\frac{\theta}{\tau(x)} c_n(x, a) \bar{\varphi}_\alpha^{(n,\delta_m)}(\theta, x) + \int_S \mathcal{Q}^{(n)}(dy|x, a) \bar{\varphi}_\alpha^{(n,\delta_m)}(\theta, y) \right] \psi(\theta) d\theta \\ &\quad - \frac{1}{\tau(x)} \int_0^{\delta_m} \inf_{a \in A(x)} \left[\theta c_n(x, a) \bar{\varphi}_\alpha^{(n,\delta_m)}(\theta, x) \right] \psi(\theta) d\theta. \end{aligned}$$

By analogous arguments, we get

$$\begin{aligned}
 & - \int_0^1 \left\{ \frac{\alpha}{\tau(x)} \frac{d(\theta\psi)}{d\theta}(\theta) \varphi_\alpha^{(n)}(\theta, x) - \varphi_\alpha^{(n)}(\theta, x) \psi(\theta) \right\} d\theta \\
 & \leq \inf_{a \in A(x)} \int_0^1 \left[\frac{\theta}{\tau(x)} c_n(x, a) \varphi_\alpha^{(n)}(\theta, x) + \int_S Q^{(n)}(dy|x, a) \varphi_\alpha^{(n)}(\theta, y) \right] \psi(\theta) d\theta.
 \end{aligned}
 \tag{4.18}$$

From (4.17) and (4.18), we get

$$\begin{aligned}
 & - \int_0^1 \left\{ \frac{\alpha}{\tau(x)} \frac{d(\theta\psi)}{d\theta}(\theta) \varphi_\alpha^{(n)}(\theta, x) - \varphi_\alpha^{(n)}(\theta, x) \psi(\theta) \right\} d\theta \\
 & = \inf_{a \in A(x)} \int_0^1 \left[\frac{\theta}{\tau(x)} c_n(x, a) \varphi_\alpha^{(n)}(\theta, x) + \int_S Q^{(n)}(dy|x, a) \varphi_\alpha^{(n)}(\theta, y) \right] \psi(\theta) d\theta.
 \end{aligned}
 \tag{4.19}$$

Thus we obtain

$$\begin{aligned}
 & - \int_0^1 \alpha \frac{d(\theta\psi)}{d\theta}(\theta) \varphi_\alpha^{(n)}(\theta, x) d\theta \\
 & = \inf_{a \in A(x)} \int_0^1 \left[\theta c_n(x, a) \varphi_\alpha^{(n)}(\theta, x) + \int_S q^{(n)}(dy|x, a) \varphi_\alpha^{(n)}(\theta, y) \right] \psi(\theta) d\theta.
 \end{aligned}$$

Hence

$$\alpha \theta \frac{\partial \varphi_\alpha^{(n)}}{\partial \theta}(\theta, x) = \inf_{a \in A(x)} \left[\theta c_n(x, a) \varphi_\alpha^{(n)}(\theta, x) + \int_S q^{(n)}(dy|x, a) \varphi_\alpha^{(n)}(\theta, y) \right] \text{ a.e. } \theta \in [0, 1]$$

in the sense of distribution. Now for $\theta \in [\delta_m, 1]$, by using (4.4) and Proposition 2.1, we have

$$\begin{aligned}
 \varphi_\alpha^{(n, \delta_m)}(\theta, x) &= \inf_{\pi \in \Pi} E_x^\pi \left[e^{n\delta_m/\alpha} \exp\left(\theta \int_0^{T_{\delta_m}(\theta)} e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \right] \\
 &\leq e^{n\delta_m/\alpha} \inf_{\pi \in \Pi} E_x^\pi \left[\exp\left(\theta \int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \right] \\
 &\leq e^{n\delta_m/\alpha} \inf_{\pi \in \Pi} E_x^\pi \left[\exp\left(\theta \int_0^\infty e^{-\alpha t} c(\xi_t^{(n)}, \pi_t) dt \right) \right] \\
 &\leq e^{n\delta_m/\alpha} \frac{\alpha^2 e^{\theta L_0/\alpha}}{\alpha^2 - \rho_0 \rho_1 \theta} (\mathcal{V}(x))^{\frac{\rho_1 \theta}{\alpha}}
 \end{aligned}$$

Note that $\varphi_\alpha^{(n, \delta_m)} \rightarrow \varphi_\alpha^{(n)}$ as $m \rightarrow \infty$. Thus, letting $m \rightarrow \infty$ in the above equation, we obtain

$$1 \leq \varphi_\alpha^{(n)}(\theta, x) \leq \frac{\alpha^2 e^{\theta L_0/\alpha}}{\alpha^2 - \rho_0 \rho_1 \theta} (\mathcal{V}(x))^{\frac{\rho_1 \theta}{\alpha}}. \tag{4.20}$$

By using (4.1), (4.2), (4.20), and the DE satisfied by $\varphi_\alpha^{(n)}$ (that is just proven), we see that $\varphi_\alpha^{(n)} \in L^\infty_{\mathcal{V}}([0, 1] \times S)$ and it is a solution of (4.11). Thus by closely mimicking the arguments as in Theorem 3.1, one can easily get the stochastic representation of the solution $\varphi_\alpha^{(n)}$, that is

$$\varphi_\alpha^{(n)}(\theta, x) = \inf_{\pi \in \Pi} E_x^\pi \left[\exp \left(\theta \int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \right]. \tag{4.21}$$

Step 2: In this step we prove Theorem 4.1, by passing to the limit as $n \rightarrow \infty$. Now we will prove that for each $x \in S$, $\{\varphi_\alpha^{(n)}\}_{n \geq 1}$ is equicontinuous on $[0, 1]$. We consider the following expression, for any given $\pi \in \Pi, x \in S, \theta, \theta_0 \in [0, 1]$:

$$\left| E_x^\pi \left[\exp \left(\theta \int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \right] - E_x^\pi \left[\exp \left(\theta_0 \int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \right] \right| \leq K_1,$$

where

$$\begin{aligned} K_1 &= E_x^\pi \left[\exp \left((\theta \wedge \theta_0) \int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \times \right. \\ &\quad \left. \left(\exp \left(|\theta_0 - \theta| \int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) - 1 \right) \right] \\ &\leq E_x^\pi \left[\exp \left((\theta \wedge \theta_0) \int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \times \right. \\ &\quad \left. \left(\exp \left(\int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) - 1 \right) |\theta_0 - \theta| \right] \\ &\leq E_x^\pi \left[\exp \left(\int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \times \right. \\ &\quad \left. \left(\exp \left(\int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) |\theta_0 - \theta| \right) \right] \\ &= |\theta_0 - \theta| \times E_x^\pi \left[\exp \left(2 \int_0^\infty e^{-\alpha t} c_n(\xi_t^{(n)}, \pi_t) dt \right) \right] \\ &\leq |\theta_0 - \theta| \times \frac{\alpha e^{2L_0/\alpha}}{\alpha - \rho_2} \left(\mathcal{V}^2(x) + \frac{b_0}{\rho_2} \right). \end{aligned}$$

Here, the first inequality is according to $e^{bz} - 1 \leq (e^b - 1)z$ for all $z \in [0, 1]$ and $b > 0$ and the last inequality follows from (3.7). Therefore, we have

$$\begin{aligned}
 |\varphi_\alpha^{(n)}(\theta_0, x) - \varphi_\alpha^{(n)}(\theta, x)| &\leq \sup_{\pi \in \Pi} |\theta_0 - \theta| \times \frac{\alpha e^{2L_0/\alpha}}{\alpha - \rho_2} \left(\mathcal{V}^2(x) + \frac{b_0}{\rho_2} \right) \\
 &= |\theta_0 - \theta| \times \frac{\alpha e^{2L_0/\alpha}}{\alpha - \rho_2} \left(\mathcal{V}^2(x) + \frac{b_0}{\rho_2} \right). \tag{4.22}
 \end{aligned}$$

By measurable selection theorem, [Bertsekas and Shreve (1996), Proposition 7.33], there exists a measurable function $f^{*n} : [0, 1] \times S \rightarrow A$ such that

$$\begin{aligned}
 \inf_{a \in A(x)} &\left[\theta c_n(x, a) \varphi_\alpha(\theta, x) + \int_S q^{(n)}(dy|x, a) \varphi_\alpha(\theta, y) \right] \\
 &= \left[\theta c_n(x, f^{*n}(\theta, x)) \varphi_\alpha(\theta, x) + \int_S q^{(n)}(dy|x, f^{*n}(\theta, x)) \varphi_\alpha(\theta, y) \right]. \tag{4.23}
 \end{aligned}$$

Let

$$\pi^{*n} : S \times \mathbb{R}_+ \rightarrow A$$

be defined by

$$\pi_t^{*n}(x) := f^{*n}(\theta e^{-\alpha t}, x).$$

Hence by Eq. (4.11), we have a.e. $\theta \in [0, 1]$ and $\forall x \in S$, we have

$$\begin{cases} \alpha \theta \frac{\partial \varphi_\alpha^{(n)}}{\partial \theta}(\theta, x) = \left[\int_S q^{(n)}(dy|x, f^{*n}(\theta, x)) \varphi_\alpha^{(n)}(\theta, y) + \theta c_n(x, f^{*n}(\theta, x)) \varphi_\alpha^{(n)}(\theta, x) \right] \\ 1 \leq \varphi_\alpha^{(n)}(\theta, x) \leq \frac{\alpha^2 e^{\theta L_0/\alpha}}{\alpha^2 - \rho_0 \rho_1 \theta} (\mathcal{V}(x))^{\frac{\rho_1 \theta}{\alpha}} \quad \forall (\theta, x) \in [0, 1] \times S. \end{cases} \tag{4.24}$$

Since $c_n \geq 0$, by (4.21), we see $\varphi_\alpha^{(n)}(\theta, x)$ is increasing in θ . Also we know that $\varphi_\alpha^{(n)}(\theta, x)$ is differentiable a.e. with respect to $\theta \in [0, 1]$. So

$$\frac{\partial \varphi_\alpha^{(n)}}{\partial \theta}(\theta, x) \geq 0 \text{ for a.e. } \theta. \tag{4.25}$$

So, by (4.1), (4.2) and (4.24), for all $x \in S$ and for a.e. θ , we have

$$\begin{cases} -\alpha \theta \frac{\partial \varphi_\alpha^{(n)}}{\partial \theta}(\theta, x) + \left[\int_S q^{(n-1)}(dy|x, f^{*n}(\theta, x)) \varphi_\alpha^{(n)}(\theta, y) + \theta c_{n-1}(x, f^{*n}(\theta, x)) \varphi_\alpha^{(n)}(\theta, x) \right] \\ \leq 0 \quad \text{if } x \in S_{n-1} \end{cases} \tag{4.26}$$

and

$$\begin{cases} -\alpha\theta \frac{\partial \varphi_\alpha^{(n)}}{\partial \theta}(\theta, x) + \left[\int_S q^{(n-1)}(dy|x, f^{*n}(\theta, x)) \varphi_\alpha^{(n)}(\theta, y) + \theta c_{n-1}(x, f^{*n}(\theta, x)) \varphi_\alpha^{(n)}(\theta, x) \right] \\ = -\alpha\theta \frac{\partial \varphi_\alpha^{(n)}}{\partial \theta}(\theta, x) \leq 0 \\ \text{if } x \notin S_{n-1} \text{ (by (4.25)).} \end{cases} \tag{4.27}$$

So, by Dynkin formula, we get

$$E_x^{\pi^{*n}} \left[\exp \left(\theta \int_0^\infty e^{-\alpha t} c_{n-1}(\xi_t^{(n-1)}, \pi_t^{*n}(\xi_t^{(n-1)})) dt \right) \right] \leq \varphi_\alpha^{(n)}(\theta, x) \text{ for all } (\theta, x) \in [0, 1] \times S. \tag{4.28}$$

Also using (4.11) and Dynkin formula (see (3.7) and (3.13)), we have

$$\varphi_\alpha^{(n-1)}(\theta, x) \leq E_x^{\pi^{*n}} \left[\exp \left(\theta \int_0^\infty e^{-\alpha t} c_{n-1}(\xi_t^{(n-1)}, \pi_t^{*n}(\xi_t^{(n-1)})) dt \right) \right]. \tag{4.29}$$

By (4.28) and (4.29), we have $\varphi_\alpha^{(n-1)}(\theta, x) \leq \varphi_\alpha^{(n)}(\theta, x)$.

Hence $\varphi_\alpha^{(n)}(\theta, x)$ is increasing in n for any $(\theta, x) \in [0, 1] \times S$. Now from (4.22), we know that for each $x \in S$, $\varphi^{(n)}(\cdot, x)$ is Lipschitz continuous in $\theta \in [0, 1]$. Also, $\varphi_\alpha^{(n)}(\theta, x)$ is increasing as $n \rightarrow \infty$ for any $(\theta, x) \in [0, 1] \times S$ and bounded above (by (4.20)), therefore there exists a function φ_α on $[0, 1] \times S$ that is continuous with respect to $\theta \in [0, 1]$, such that along a subsequence $n_k \rightarrow \infty$, we have $\lim_{n_k \rightarrow \infty} \varphi_\alpha^{(n_k)}(\theta, x) = \varphi_\alpha(\theta, x)$ and this convergence is uniform in $\theta \in [0, 1]$ for each fixed $x \in S$. Moreover, by (4.20), we have

$$1 \leq \varphi_\alpha(\theta, x) \leq \frac{\alpha^2 e^{\theta L_0/\alpha}}{\alpha^2 - \rho_0 \rho_1 \theta} (\mathcal{V}(x))^{\frac{\rho_1 \theta}{\alpha}}. \tag{4.30}$$

As the proof of equation (4.11) in step 1 (starting from the first equality of (4.13)), we see that φ_α is a solution to the HJB Eq. (3.1). Also by (4.30), we can conclude that $\varphi_\alpha \in L^\infty_{\mathcal{Y}}([0, 1] \times S)$. Finally, the uniqueness of $\varphi_\alpha(\theta, x)$ follows from the stochastic representation in Theorem 3.1. \square

5 The existence of optimal control

In this section, we present the main result of this article. Here we show the existence of an optimal control.

Theorem 5.1 *Suppose that Assumptions 2.1, 2.2 and 3.1 are satisfied. Then, the following assertions hold.*

(1) The HJB Eq. (3.1) has a unique solution $\varphi_\alpha \in L^\infty_{\mathcal{V}}([0, 1] \times S)$ and the solution admits the following representation

$$1 \leq \varphi_\alpha(\theta, x) = \inf_{\pi \in \Pi} E_x^\pi \left[\exp \left(\theta \int_0^\infty e^{-\alpha t} c(\xi_t, \pi_t) dt \right) \right] \leq \frac{\alpha^2 e^{\theta L_0/\alpha}}{\alpha^2 - \rho_0 \rho_1 \theta} (\mathcal{V}(x))^{\frac{\rho_1 \theta}{\alpha}}.$$

(2) There exists a measurable function $f^* : [0, 1] \times S \rightarrow A$ such that

$$\alpha \theta \frac{\partial \varphi_\alpha}{\partial \theta}(\theta, x) = \left[\int_S q(dy|x, f^*(\theta, x)) \varphi_\alpha(\theta, y) + \theta c(x, f^*(\theta, x)) \varphi_\alpha(\theta, x) \right] \text{ a.e. } \theta \in [0, 1]. \tag{5.1}$$

(3) Furthermore an optimal Markov control for the cost criterion (2.2) exists and is given by

$$\tilde{\pi}_t^*(x) := f^*(\theta e^{-\alpha t}, x),$$

where f^* satisfies (5.1).

Proof Part (1) follows from Theorems 3.1 and 4.1.

To prove (2), by Hernandez-Lerma and Lasserre (1999), we first observe that the function

$$\int_S q(dy|x, a) \varphi_\alpha(\theta, y) + \theta c(x, a) \varphi_\alpha(\theta, x)$$

is continuous in $a \in A(x)$ for each given $(\theta, x) \in [0, 1] \times S$. Thus, by the measurable selection theorem (Bertsekas and Shreve 1996), Proposition 7.33 there exists a measurable function f^* satisfying (5.1), and so (2) follows. For part (3), take any f^* that satisfies (5.1). Then by Theorem 3.1, we have $\inf_{\pi \in \Pi} \tilde{J}_\alpha(\theta, x, \pi) = \tilde{J}_\alpha(\theta, x, \tilde{\pi}^*) = \varphi_\alpha(\theta, x)$, which together with (2.2), (2.3) and part (1), we have $\inf_{\pi \in \Pi} \mathcal{J}_\alpha(\theta, x, \pi) = \mathcal{J}_\alpha(\theta, x, \tilde{\pi}^*) = \frac{1}{\theta} \ln \tilde{J}_\alpha(\theta, x, \tilde{\pi}^*) = \frac{1}{\theta} \ln \varphi_\alpha(\theta, x)$. Hence $\tilde{\pi}^*$ is an optimal Markov control. \square

Now we prove the converse of the Theorem 5.1.

Theorem 5.2 Grant Assumptions 2.1, 2.2 and 3.1. Suppose there exists an optimal Markov control for the cost criterion (2.2) and is given by

$$\hat{\pi}_t^*(x) := \tilde{f}^*(\theta e^{-\alpha t}, x),$$

for some measurable function \tilde{f}^* . Then we prove that \tilde{f}^* is a minimizing selector of (3.1).

Proof Since $\hat{\pi}^*$ is optimal for the cost criterion (2.2), therefore we have

$$\inf_{\pi \in \Pi} \tilde{J}_\alpha(\theta, x, \pi) = \tilde{J}_\alpha(\theta, x, \hat{\pi}^*) = \tilde{J}_\alpha^*(\theta, x). \tag{5.2}$$

Now for \tilde{f}^* by Theorem 4.1, there exists a unique solution $\psi_\alpha \in L^\infty([0, 1] \times S)$ for the equation

$$\alpha\theta \frac{\partial \psi_\alpha}{\partial \theta}(\theta, x) = \left[\int_S q(dy|x, \tilde{f}^*(\theta, x))\psi_\alpha(\theta, y) + \theta c(x, \tilde{f}^*(\theta, x))\psi_\alpha(\theta, x) \right], \tag{5.3}$$

for each $x \in S$ and a.e. $\theta \in [0, 1]$, satisfying $1 \leq \psi_\alpha(\theta, x) \leq \frac{\alpha^2 e^{\theta L_0/\alpha}}{\alpha^2 - \rho_0 \rho_1 \theta} (\mathcal{V}(x))^{\frac{\rho_1 \theta}{\alpha}}$ for all $(\theta, x) \in [0, 1] \times S$.

Now by Theorem 3.1, we know that

$$\begin{aligned} 1 \leq \psi_\alpha(\theta, x) &= E_x^{\hat{\pi}^*} \left[\exp \left(\theta \int_0^\infty e^{-\alpha t} c(\xi_t, \hat{\pi}_t^*(\xi_{t-})) dt \right) \right] \\ &\leq \frac{\alpha^2 e^{\theta L_0/\alpha}}{\alpha^2 - \rho_0 \rho_1 \theta} (\mathcal{V}(x))^{\frac{\rho_1 \theta}{\alpha}}. \end{aligned} \tag{5.4}$$

From (5.2) and (5.4), we get

$$\psi_\alpha(\theta, x) = \inf_{\pi \in \Pi} \tilde{J}_\alpha(\theta, x, \pi) = \tilde{J}_\alpha(\theta, x, \hat{\pi}^*) = \tilde{J}_\alpha^*(\theta, x) \text{ for } (\theta, x) \in [0, 1] \times S. \tag{5.5}$$

So, in view of Theorem 3.1, by Eqs. (3.1), (5.3), and (5.5), we conclude that \tilde{f}^* is a minimizing selector of (3.1). \square

When the transition and cost rates are bounded, the existence of an optimal control is ensured by Theorem 5.1.

Corollary 5.1 *Grant Assumption 3.1 ((i)–(ii)). Also, assume that the transition and cost rates are bounded. Then, there exist a unique solution φ_α and an optimal control for the HJB Eq. (3.1).*

Proof Suppose there exist constants L_1 and b_1 , such that $\sup_{(x,a) \in K} q_x(a) \leq L_1$ and

$\sup_{(x,a) \in K} c(x, a) \leq b_1$. First we take the Lyapunov function $\mathcal{V}(x) \equiv P$, for all $x \in$

$S, P \geq 1$, a constant. Now $\int_S \mathcal{V}(y)q(dy|x, a) = \int_S \mathcal{V}^2(y)q(dy|x, a) = 0$, for all $(x, a) \in K$. Now, take $\rho_0 = \alpha, M_0 = L_1$, any real number, $\rho_1 \in (0, \alpha)$, and $L_0 = b_1$. Then Assumption 2.1 is verified. Now for all $x \in S$, take any constants $\rho_2 \in (0, \alpha)$ and $b_0 \in (0, \infty)$. Then Assumption 2.2 holds. Also $\int_S \mathcal{V}(y)q(dy|x, a)$ is continuous in $a \in A(x)$. So, Assumption 3.1 is also true. Then, by Theorem 5.1, we have a unique solution φ_α and an optimal control for the HJB Eq. (3.1). \square

6 Application and example

In this section, we verify the above assumptions with one example, where the transition and cost rates are unbounded.

Example 6.1 The Gaussian Model: Suppose a hunter is hunting outside his house for his manager. Suppose the house is at state 0. A positive state represents the distance from the house to the right, and a negative state represents the distance from the house to the left. Let $S = \mathbb{R}$. If the current position is $x \in S$, the hunter takes an action $a \in A(x)$, then after an exponentially distributed travel time with rate $\lambda(x, a) > 0$, the hunter reaches the new position, and the travel distance follows the normal distribution with mean x and variance σ . (Or we can interpret $\lambda(x, a)$ as the total jump intensity that is an arbitrary measurable positive-valued function on $S \times A$, and the distribution of the state after a jump from $x \in S$ is normal with the variance σ and expectation x .) Also assume that the hunter receives a payoff $c(x, a)$ from his manager for each unit of time he spends there. Let us consider the model as $A_2 := \{S, (A, A(x), x \in S), c(x, a), q(dy|x, a)\}$, where $S = (-\infty, \infty)$. For each $D \in \mathcal{B}(S)$, the transition rate is

$$q(D|x, a) = \lambda(x, a) \left[\int_{y \in D} \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(y-x)^2}{2\sigma^2}} dy - \delta_x(D) \right], \quad x \in S, a \in A(x), \sigma > 0. \tag{6.1}$$

To ensure the existence of an optimal Markov control for the model, we consider the following hypotheses.

- (I) For each fixed $x \in S$, $\lambda(x, a)$ is continuous in $a \in A(x)$ and there exists a positive constant M_1 such that $0 < \sup_{a \in A(x)} \lambda(x, a) \leq M_1(x^2 + 1)$ and $M_1 < \frac{\alpha}{6\sigma^2(\sigma^2 + 1)}$.
- (II) For each $x \in S$, the cost rate $c(x, a)$ is nonnegative and continuous in $a \in A(x)$ and there exists a constant $0 < \rho_1 < \min\{\alpha, \frac{\alpha^2}{M_1\sigma^2}\}$ such that

$$\sup_{a \in A(x)} c(x, a) \leq \rho_1 \log(1 + x^2).$$

- (III) For each fixed $x \in S$, $A(x)$ is a compact subset of the Borel spaces A .

Proposition 6.1 *Under conditions (I)–(III), the above controlled system satisfies the Assumptions 2.1, 2.2, and 3.1. Hence by Theorem 5.1, there exists an optimal Markov control for this model.*

Proof We know $\frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} (y - x)^{2k+1} e^{-\frac{(y-x)^2}{2\sigma^2}} dy = 0$ and $\frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} (y - x)^{2k} e^{-\frac{(y-x)^2}{2\sigma^2}} dy = 1 \cdot 3 \cdot \dots \cdot (2k - 1)\sigma^{2k}$ for all $k = 0, 1, \dots$.

We first verify Assumption 2.1. Let $\mathcal{V}(x) = x^2 + 1$.

$$\int_S \mathcal{V}(y)q(dy|x, a) = \lambda(x, a) \left[\frac{1}{\sqrt{2\pi\sigma}} \int_{-\infty}^{\infty} (y^2 + 1)e^{-\frac{(y-x)^2}{2\sigma^2}} dy - (x^2 + 1) \right]$$

$$\begin{aligned}
 &= \lambda(x, a)\sigma^2 \\
 &\leq M_1\sigma^2\mathcal{V}(x).
 \end{aligned}
 \tag{6.2}$$

Let $\rho_0 = M_1\sigma^2$. Then $\int_S \mathcal{V}(y)q(dy|x, a) \leq \rho_0\mathcal{V}(x)$. Now

$$q^*(x) = \sup_{a \in A(x)} q_x(a) = \sup_{a \in A(x)} \lambda(x, a) \leq M_1(x^2 + 1) = M_1\mathcal{V}(x) \quad \forall x \in S.$$

Now by condition (II), we can write

$$\sup_{a \in A(x)} c(x, a) \leq \rho_1 \log(1 + x^2) + M_1.$$

Observe that by condition (II), $0 < \rho_1 < \min\{\alpha, \rho_0^{-1}\alpha^2\}$. Hence Assumption 2.1 is verified with $M_0 = L_0 = M_1$.

Next we verify Assumption 2.2. For any $x \in S, a \in A(x)$,

$$\begin{aligned}
 \int_S q(dy|x, a)\mathcal{V}^2(y) &= \int_S q(dy|x, a)(1 + y^2)^2 \\
 &= \lambda(x, a) \left[\frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{\infty} (y^2 + 1)^2 e^{-\frac{(y-x)^2}{2\sigma^2}} dy - (x^2 + 1)^2 \right] \\
 &= \lambda(x, a)[1 \cdot 3\sigma^4 + \sigma^2(2 + 6x^2)] \\
 &\leq M_1(x^2 + 1)[3\sigma^4 + \sigma^2(2 + 6x^2)] \\
 &= M_1\sigma^2(x^2 + 1)(3\sigma^2 + 2 + 6x^2) \\
 &\leq 6M_1\sigma^2(x^2 + 1)(x^2 + 1)(\sigma^2 + 1) \\
 &= 6M_1\mathcal{V}^2(x)\sigma^2(\sigma^2 + 1) \\
 &\leq \rho_2\mathcal{V}^2(x) + 1
 \end{aligned}$$

where $\rho_2 = 6M_1\sigma^2(\sigma^2 + 1)$, and $b_0 = 1$. Then by condition (I), we have $0 < \rho_2 < \alpha$. Hence, Assumption 2.2 is verified. Now by conditions (I) and (II) $c(x, a)$ is continuous in $a \in A(x)$. Observe that by condition (I) and (6.2), $\int_S \mathcal{V}(y)q(dy|x, a)$ is continuous in $a \in A(x)$. Hence Assumption 3.1 is also verified. So, by Theorem 5.1, we see that there exists an optimal Markov control for this model. □

Remark 6.1 As we mention in the introduction, there are many real-life applications, where the underlying system dynamic is modeled as a CTMDP, with a Borel state and action spaces as well as cost and transition rates are unbounded, see such a cash-flow problem in Guo and Zhang (2019), [p. 112, Piunovskiy and Zhang (2020)]. Also, there are lots of real-life examples like infrastructure surveillance models [p. 115–116, Piunovskiy and Zhang (2020)], queueing model [p. 192, Piunovskiy and Zhang (2020)], where we see that the state space is uncountable, can be formulated in our set-up.

Acknowledgements We thank the anonymous referees for their valuable comments and helpful suggestions that have improved the presentation of this paper.

References

- Bauerle N, Rieder U (2014) More risk-sensitive Markov decision processes. *Math Oper Res* 39:105–120
- Bertsekas D, Shreve S (1996) *Stochastic optimal control: the discrete-time case*. Academic Press Inc, New York
- Ghosh MK, Saha S (2014) Risk-sensitive control of continuous-time Markov chains. *Stoch* 86:655–675
- Guo XP, Hernandez-Lerma O (2009) *Continuous-time Markov decision processes: theory and applications*. Stochastic modelling and applied probability. Springer, Berlin
- Guo X, Liao ZW (2019) Risk-sensitive discounted continuous-time Markov decision processes with unbounded rates. *SIAM J Control Optim* 57:3857–3883
- Guo X, Piunovskiy A (2011) Discounted continuous-time Markov decision processes with constraints: unbounded transition and loss rates. *Math Oper Res* 36:105–132
- Guo X, Zhang J (2019) Risk-sensitive continuous-time Markov decision processes with unbounded rates and Borel spaces. *Discrete Event Dyn Syst* 29:445–471
- Guo X, Zhang Y (2020) On risk-sensitive piecewise deterministic Markov decision processes. *Appl Math Optim* 81:685–710
- Guo X, Huang Y, Song X (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
- Guo X, Liu Q, Zhang Y (2019) Finite horizon risk-sensitive continuous-time Markov decision processes with unbounded transition and cost rates. *4OR* 17:427–442
- Hernandez-Lerma O, Lasserre J (1999) *Further topics on discrete-time Markov control processes*. Springer, New York
- Kitaev MY (1995) Semi-Markov and jump Markov controlled models: average cost criterion. *SIAM Theory Probab Appl* 30:272–288
- Kitaev MY, Rykov VV (1995) *Controlled queueing systems*. CRC Press, Boca Raton
- Kumar KS, Pal C (2013) Risk-sensitive control of jump process on denumerable state space with near monotone cost. *Appl Math Optim* 68:311–331
- Kumar KS, Pal C (2015) Risk-sensitive control of continuous-time Markov processes with denumerable state space. *Stoch Anal Appl* 33:863–881
- Masi GB, Stettner L (2000) Infinite horizon risk-sensitive control of discrete time Markov processes with small risk. *Syst Control Lett* 40:15–20
- 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
- Pal C, Pradhan S (2019) Risk-sensitive control of pure jump processes on a general state space. *Int J Probab Stoch Process* 91(2):155–174
- Piunovskiy A, Zhang Y (2020) *Continuous-time Markov decision processes*. Springer, Berlin
- Piunovskiy A, Zhang Y (2011) Discounted continuous-time Markov decision processes with unbounded rates: the convex analytic approach. *SIAM J Control Optim* 49:2032–2061
- Prieto-Rumeau T, Hernandez-Lerma O (2012) *Selected topics in continuous-time controlled Markov chains and Markov games*. Imperial College Press, London
- Wei Q (2016) Continuous-time Markov decision processes with risk-sensitive finite-horizon cost criterion. *Math Methods Oper Res* 84:461–487
- Whittle P (1990) *Risk-sensitive optimal control*, Wiley-Inter Science series in systems and optimization. Wiley, Chichester
- Zhang Y (2017) Continuous-time Markov decision processes with exponential utility. *SIAM J Control Optim* 55:2636–2660

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.