Continuous-Time Markov Decision Processes Under the Risk-Sensitive First Passage Discounted Cost Criterion

Qingda Wei[1] · Xian Chen[2]

Abstract

This paper studies the risk-sensitive first passage discounted cost criterion for continuous-time Markov decision processes with the Borel state and action spaces. The cost and transition rates are allowed to be unbounded. We introduce a *new* value iteration to establish the existence of a solution to the risk-sensitive first passage discounted cost optimality equation. Then applying the Feynman–Kac formula, we show that the risk-sensitive first passage discounted cost optimal value function is a unique solution to the risk-sensitive first passage discounted cost optimality equation. Moreover, we derive the existence of a deterministic Markov optimal policy in the class of randomized history-dependent policies. Finally, a cash flow model is given to illustrate the results.

1 Introduction

Markov decision processes have been widely applied to the queueing system, power-managed systems, inventory control, telecommunication, infrastructure surveillance

---

✉ Xian Chen
   chenxian@xmu.edu.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

models, preventive maintenance, and epidemic control, etc. (see [11, 18, 20, 21]). In the standard expected discounted cost criterion, the decision horizon is infinite and the decision-maker is assumed to be risk-neutral (see [11, 14, 18, 20, 21]). From the viewpoint of the risk preference, the decision-maker may be risk-sensitive rather than risk-neutral; see, for instance, [2, 4–10, 15–17, 22, 24–28]. The exponential utility function is a common way of characterizing the risk-sensitivity of the decision-maker in Markov decision processes; see, for instance, [8, 10, 16, 25] for the finite-horizon risk-sensitive cost criterion, [5, 8, 15] for the infinite-horizon risk-sensitive discounted cost criterion and [4–9, 22, 24, 26, 27] for the infinite-horizon risk-sensitive average cost criterion. On the other hand, the decision horizon in [4–10, 15, 16, 22, 24–27] is either finite or infinite. But in the practical applications, the decision-maker may be concerned with the costs before the state of the controlled stochastic system falls into some target set. Thus, the decision horizon is uncertain; see, for instance, [12, 19]. The first passage time to some target set is used to describe the uncertainty of the decision horizon in [12, 19]. More precisely, Guo et al. [12] studied the first passage mean-variance criterion for discounted continuous-time Markov decision processes and [19] investigated the nonzero-sum games under the first passage expected discounted payoff criterion for continuous-time jump processes. Hence, it is meaningful for us to take the risk-sensitivity of the decision-maker and the uncertainty of the decision horizon into consideration. In this paper we intend to investigate continuous-time Markov decision processes under the risk-sensitive first passage discounted cost criterion, which has not been studied yet.

The state and action spaces are Borel spaces, and the cost and transition rates can be possibly unbounded in this paper. In order to ensure the finiteness of the risk-sensitive first passage discounted cost criterion and the existence of optimal policies, we require the drift condition on the transition rate, the growth condition on the cost rate, and the continuity and compactness conditions. First, we derive the bounds on the risk-sensitive first passage discounted cost criterion and the Feynman–Kac formula which is applicable to a class of unbounded functions (see Lemmas 3.1 and 3.3). Then, we construct an approximating sequence of bounded cost rates and introduce a new value iteration [see (4.1)]. Employing the results on the approximating sequence, we obtain the existence of a solution to the risk-sensitive first passage discounted cost optimality equation for the case of the Borel state space and unbounded transition and cost rates. Moreover, using the Feynman–Kac formula, we prove that the risk-sensitive first passage discounted cost optimal value function is a unique solution to the risk-sensitive first passage discounted cost optimality equation. In addition, from the optimality equation we derive the existence of a deterministic Markov optimal policy in the class of randomized history-dependent policies (see Theorem 4.1). Finally, we provide a cash flow model to illustrate the optimality conditions of our results.

Compared with the existing literature on the risk-sensitive discounted continuous-time Markov decision processes, the main contributions of this paper are as follows.

(I) (New method). The existence of a solution to the infinite-horizon risk-sensitive discounted cost optimality equation is obtained in [5] with the Borel state space and bounded cost and transition rates, [8] with the denumerable state space and bounded cost and transition rates, and [15] with the denumerable state space and

unbounded cost and transition rates. The extension of the results in [5, 8, 15] to the case of the Borel state space and unbounded cost and transition rates is nontrivial. We establish the existence of a solution to the optimality equation via introducing a *new* value iteration, which is different from the methods in [5, 8, 15]. More precisely, the existence of a solution to the infinite-horizon risk-sensitive discounted cost optimality equation is obtained in [5, 8, 15] via the fixed point technique. Since the function $\frac{1}{\lambda}$ is singular at the point $\lambda = 0$, the integral in the construction of the fixed point mapping needs to start from some $\varepsilon > 0$ in [5, 8, 15]. Then letting $\varepsilon \to 0$, the existence of a solution to the infinite-horizon risk-sensitive discounted cost optimality equation for the case of bounded cost and transition rates can be derived. The fixed point method requires the boundedness assumption on the cost and transition rates. Thus, the approach in [5, 8] cannot deal with the unbounded cost and transition rates. To investigate the unbounded case, [15] first constructs an approximating sequence of bounded cost and transition rates via the truncation method and then applying the fixed point technique to the approximating sequence. The existence of a solution to the infinite-horizon risk-sensitive discounted cost optimality equation for a denumerable state space and unbounded cost and transition rates is established in [15] by an approximation approach. Because the diagonalization arguments in [15] require the assumption that the state space is a denumerable set, the method in [15] is inapplicable to the case of the Borel state space. To deal with the Borel state space and unbounded cost and transition rates, we introduce a *new* value iteration [see (4.1)]. Our value iteration approach has the following two advantages. On the one hand, we overcome the singular issue from the fixed point method in [5, 8, 15]. The approximation $\varepsilon \to 0$ in [5, 8, 15] is not required for the value iteration method. On the other hand, the value iteration can treat the case of the Borel state space and unbounded transition rates directly without the boundedness assumption as in [5, 8] and the diagonalization arguments and the construction of an approximating sequence of bounded transition rates as in [15].

(II) (Weaker conditions). We obtain the existence result under the conditions weaker than those in [5, 8, 15]. More specifically, we remove the boundedness condition on the cost and transition rates in [5, 8] and the uniform convergence condition of the solution in [8]. We do not require the denumerability assumption on the state space, the uniformly conservative condition on the transition rates, and the additional condition that the constant in the second-order drift condition is less than the discount factor in [15].

The remainder of the paper is organized as follows. In Sect. 2, we introduce the decision model and the risk-sensitive first passage discounted cost criterion. In Sect. 3, we provide the optimality conditions for the existence of optimal policies and prove the Feynman–Kac formula which is applicable to a class of unbounded functions. In Sect. 4, we employ a new value iteration to establish the risk-sensitive first passage discounted cost optimality equation and show the existence of optimal policies. In Sect. 5, we use a cash flow model to illustrate the main results. In Sect. 6, we give the conclusions.

2 The Decision Model

We first give a description of the decision model and then introduce the definition of the policy and the risk-sensitive first passage discounted cost criterion in this section. The decision model contains the following components:

$$\{X, A, (A(x) \subseteq A, x \in X), q(\cdot|x, a), c(x, a)\}.$$

The state space X and action space A are assumed to be Borel spaces endowed with the Borel σ-algebras $\mathcal{B}(X)$ and $\mathcal{B}(A)$, respectively. For each $x \in X$, $A(x) \in \mathcal{B}(A)$ denotes the set of all admissible actions at state x. Let $K := \{(x, a) : x \in X, a \in A(x)\}$ which is supposed to be a Borel measurable subset of $X \times A$. The transition rate q satisfies the following properties: (i) for any $(x, a) \in K$, $q(\cdot|x, a)$ is a signed measure on $\mathcal{B}(X)$ and for any $D \in \mathcal{B}(X)$, $q(D|\cdot)$ is a real-valued measurable function on K; (ii) $0 \le q(D|x, a) < \infty$ for all $(x, a) \in K$ and $x \notin D \in \mathcal{B}(X)$; (iii) $q(X|x, a) = 0$ for all $(x, a) \in K$; and (iv) $q^*(x) := \sup_{a \in A(x)} |q(\{x\}|x, a)| < \infty$ for all $x \in X$. The cost rate c is a nonnegative real-valued measurable function on K.

Next, we describe the construction of the state process. To this end, set $X_\infty := X \cup \{x_\infty\}$ with an isolated point $x_\infty \notin X$, $\mathbb{R}_+ := (0, \infty)$, $\Omega^0 := (X \times \mathbb{R}_+)^\infty$ and $\Omega := \Omega^0 \cup \{(x_0, \theta_1, x_1, \ldots, \theta_n, x_n, \infty, x_\infty, \infty, x_\infty, \ldots)|x_0 \in X, x_l \in X, \theta_l \in \mathbb{R}_+$ for each $1 \le l \le n, n \ge 1\}$. Denote by $\mathcal{F}$ the Borel σ-algebra of Ω. For any $\omega = (x_0, \theta_1, x_1, \ldots) \in \Omega$, define $S_0(\omega) := x_0$, $T_0(\omega) := 0$, $S_n(\omega) := x_n$, $T_n(\omega) := \sum_{i=1}^n \theta_i$ for all $n \ge 1$, $T_\infty(\omega) := \lim_{n \to \infty} T_n(\omega)$ and the state process

$$\xi_t(\omega) := \sum_{n \ge 0} I_{\{T_n \le t < T_{n+1}\}} x_n + I_{\{t \ge T_\infty\}} x_\infty$$

for $t \ge 0$, where I_D stands for the indicator function of a set D. Moreover, let $A(x_\infty) := a_\infty$ with an isolated point $a_\infty \notin A$, $A_\infty := A \cup \{a_\infty\}$, $c(x_\infty, a_\infty) := 0$, $q(x_\infty|x_\infty, a_\infty) := 0$, $\mathcal{F}_t := \sigma(\{T_n \le s, S_n \in D\} : s \le t, D \in \mathcal{B}(X), n \ge 0)$ for all $t \ge 0$, $\mathcal{F}_{s-} := \bigvee_{0 \le t < s} \mathcal{F}_t$, 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 presents the σ-algebra of predictable sets on $\Omega \times [0, \infty)$ with respect to $\{\mathcal{F}_t\}_{t \ge 0}$.

Now we give the definition of a policy.

Definition 2.1 A transition probability $\pi(\cdot|\cdot)$ on A_∞ given $\Omega \times [0, \infty)$ is called a randomized history-dependent policy if for any $(\omega, t) \in \Omega \times [0, \infty)$, $\pi(\cdot|\omega, t)$ is a probability measure supported on $A(\xi_{t-}(\omega))$ and for any $D \in \mathcal{B}(A_\infty)$, $\pi(D|\cdot)$ is $\mathcal{P}$-measurable, where $\xi_{t-} = \lim_{s \uparrow t} \xi_s$. The set of all randomized history-dependent policies is denoted by Π. A policy $\pi \in \Pi$ is said to be deterministic Markov if there exists a measurable mapping $f : [0, \infty) \times X_\infty \to A_\infty$ with $f(t, x) \in A(x)$ for all $(t, x) \in [0, \infty) \times X_\infty$ such that $\pi(\cdot|\omega, t)$ is the Dirac measure at the point $f(t, \xi_{t-}(\omega))$ for all $(\omega, t) \in \Omega \times [0, \infty)$. The set of all deterministic Markov policies is denoted by Π_M.

For any initial state $x \in X$ and an arbitrary policy $\pi \in \Pi$, by Theorem 4.27 in [20] there exists a unique probability measure P_x^π on $(\Omega, \mathcal{F})$. The expectation operator with respect to P_x^π is denoted by E_x^π.

Finally, we introduce the risk-sensitive first passage discounted cost criterion. Fix the target set $B \in \mathcal{B}(X)$ and the discount factor $\alpha > 0$. The first passage time to the target set B is given by

$$\tau_B := \inf\{t \geq 0 : \xi_t \in B\} \text{ with the convention that } \inf \emptyset := \infty.$$

For any risk-sensitivity coefficient $\lambda > 0$, the risk-sensitive first passage discounted cost criterion is defined as follows:

$$J(\lambda, x, \pi) := \frac{1}{\lambda} \ln E_x^\pi \left[e^{\lambda \int_0^{\tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} \right] \tag{2.1}$$

for all $x \in X$ and $\pi \in \Pi$. Taking $B = \emptyset$, we can see that the risk-sensitive first passage discounted cost criterion given by (2.1) reduces to the infinite-horizon risk-sensitive discounted cost criterion in [5, 8, 15]. Let

$$\widehat{J}(\lambda, x, \pi) := E_x^\pi \left[e^{\lambda \int_0^{\tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} \right]$$

for all $x \in X$ and $\pi \in \Pi$. Then we have $J(\lambda, x, \pi) = 0$ and $\widehat{J}(\lambda, x, \pi) = 1$ for all $x \in B$ and $\pi \in \Pi$.

Definition 2.2 A policy $\pi^* \in \Pi$ is said to be optimal if $J(\lambda, x, \pi^*) = \inf_{\pi \in \Pi} J(\lambda, x, \pi) =: J^*(\lambda, x)$ for all $x \in X$. The function J^* is called the risk-sensitive first passage discounted cost optimal value function.

3 Preliminaries

In this section, we give the optimality conditions for the existence of optimal policies. To investigate the possibly unbounded cost and transition rates, we require the following assumption.

Assumption 3.1 There exist a real-valued measurable function $w \geq 1$ on X, constants $\rho > 0, d \geq 0, L > 0, b \geq 0$ and $0 \leq M < \min\left\{\frac{\alpha^2}{\rho\lambda}, \frac{\alpha}{2\lambda}\right\}$ such that

 (i) $\int_X w(y)q(dy|x, a) \leq \rho w(x) + d$ for all $(x, a) \in K$;
 (ii) $q^*(x) \leq Lw(x)$ for all $x \in X$;
 (iii) $c(x, a) \leq M \ln w(x) + b$ for all $(x, a) \in K$.

Remark 3.1 Assumptions 3.1(i) and (ii) are used to ensure the non-explosion of the state process; see, for example, [9–12, 14–16, 19, 25, 27]. Assumption 3.1(iii) is used to guarantee the finiteness of the risk-sensitive first passage discounted cost criterion.

Under Assumption 3.1, we have the following result.

Lemma 3.1 *Under Assumption* 3.1, *the following assertions hold.*

(a) $E_x^\pi[w(\xi_t)] \le e^{\rho t} w(x) + \frac{d}{\rho} e^{\rho t} \le e^{\rho t} \left(1 + \frac{d}{\rho}\right) w(x)$ *for all* $x \in X$ *and* $\pi \in \Pi$.

(b) $\widehat{J}(\lambda, x, \pi) \le R_\lambda w^{\frac{\lambda M}{\alpha}}(x)$ *for all* $x \in X$ *and* $\pi \in \Pi$, *where the constant* $R_\lambda = \frac{\alpha^2 e^{\frac{\lambda b}{\alpha}}}{\alpha^2 - \rho \lambda M} \left(1 + \frac{d}{\rho}\right)^{\frac{\lambda M}{\alpha}}$.

(c) $J(\lambda, x, \pi) \le \frac{1}{\lambda} \ln R_\lambda + \frac{M}{\alpha} \ln w(x)$ *for all* $x \in X$ *and* $\pi \in \Pi$.

Proof (a) The assertion follows from Theorem 3.1 in [14].

(b) Direct computations yield

$$
\begin{aligned}
\widehat{J}(\lambda, x, \pi) &\le E_x^\pi \left[e^{\lambda \int_0^\infty \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} \right] \\
&\le E_x^\pi \left[\int_0^\infty e^{\frac{\lambda}{\alpha} \int_A c(\xi_t, a) \pi(da|\omega, t)} \alpha e^{-\alpha t} dt \right] \\
&\le \alpha e^{\frac{\lambda b}{\alpha}} E_x^\pi \left[\int_0^\infty e^{-\alpha t} w^{\frac{\lambda M}{\alpha}}(\xi_t) dt \right] \\
&\le \alpha e^{\frac{\lambda b}{\alpha}} \int_0^\infty e^{-\alpha t} (E_x^\pi[w(\xi_t)])^{\frac{\lambda M}{\alpha}} dt \\
&\le \alpha e^{\frac{\lambda b}{\alpha}} \left(1 + \frac{d}{\rho}\right)^{\frac{\lambda M}{\alpha}} w^{\frac{\lambda M}{\alpha}}(x) \int_0^\infty e^{-\alpha t} e^{\frac{\rho \lambda M t}{\alpha}} dt \\
&= \frac{\alpha^2 e^{\frac{\lambda b}{\alpha}}}{\alpha^2 - \rho \lambda M} \left(1 + \frac{d}{\rho}\right)^{\frac{\lambda M}{\alpha}} w^{\frac{\lambda M}{\alpha}}(x)
\end{aligned}
$$

for all $x \in X$ and $\pi \in \Pi$, where the second and fourth inequalities are due to the Jensen inequality, the third one follows from Assumption 3.1(iii), and the fifth one is due to part (a).

(c) From part (b), we can directly obtain part (c). □

To obtain the existence of optimal policies, we also need the following assumption.

Assumption 3.2 (i) For each $x \in X$, the set $A(x)$ is compact.

(ii) For each $x \in X$, $c(x, a)$ is lower semi-continuous in $a \in A(x)$ and $\int_X v(y) q(dy|x, a)$ is continuous in $a \in A(x)$ for any bounded measurable function v on X.

(iii) There exist constants $\overline{\rho} > 0$ and $\overline{d} \ge 0$ such that $\int_X w^2(y) q(dy|x, a) \le \overline{\rho} w^2(x) + \overline{d}$ for all $(x, a) \in K$, where the function w on X comes from Assumption 3.1.

Remark 3.2 The continuity and compactness conditions in Assumptions 3.2(i) and (ii) are used to obtain the existence of optimal policies; see, for example, [9–12, 14–16, 19, 25–27]. Assumption 3.2(iii) has been widely used in continuous-time Markov decision processes; see, for example, [9–12, 14–16, 19, 25, 27]. Here, we use it to ensure the integrability in obtaining the Feynman–Kac formula in Lemma 3.3.

Lemma 3.2 *Under Assumptions* 3.2*(i) and (ii), the following statements are true.*

(a) *For any nonnegative real-valued measurable function v on X, $\int_X v(y)q(\mathrm{d}y|x,a)$ is lower semi-continuous in $a \in A(x)$ for all $x \in X$.*

(b) *Let $\{v_n, n \geq 1\}$ be a sequence of nonnegative real-valued measurable functions on X with $\lim_{n\to\infty} v_n = v$. Then, for any $x \in X$ and any sequence $\{a_n, n \geq 1\} \subseteq A(x)$ satisfying $a_n \to a$ as $n \to \infty$, we have*

$$\liminf_{n\to\infty} \int_X v_n(y)q(\mathrm{d}y|x,a_n) \geq \int_X v(y)q(\mathrm{d}y|x,a).$$

Proof (a) Define $v_m(x) := \min\{v(x), m\}$ for all $x \in X$ and $m \geq 1$. Fix any $x \in X$. Let $\{a_n, n \geq 1\} \subseteq A(x)$ be a sequence satisfying $a_n \to a$ as $n \to \infty$. Employing Assumption 3.2(ii), we can get

$$\lim_{n\to\infty} q(\{x\}|x,a_n) = \lim_{n\to\infty} \int_X I_{\{x\}}(y)q(\mathrm{d}y|x,a_n)$$
$$= \int_X I_{\{x\}}(y)q(\mathrm{d}y|x,a) = q(\{x\}|x,a). \qquad (3.1)$$

For each $m \geq 1$, we have

$$\liminf_{n\to\infty} \int_{X\setminus\{x\}} v(y)q(\mathrm{d}y|x,a_n) \geq \liminf_{n\to\infty} \int_{X\setminus\{x\}} v_m(y)q(\mathrm{d}y|x,a_n)$$
$$= \liminf_{n\to\infty} \left[\int_X v_m(y)q(\mathrm{d}y|x,a_n) - v_m(x)q(\{x\}|x,a_n) \right]$$
$$= \int_{X\setminus\{x\}} v_m(y)q(\mathrm{d}y|x,a), \qquad (3.2)$$

where the last equality follows from Assumption 3.2(ii) and (3.1). Thus, letting $m \to \infty$ in (3.2) and using the monotone convergence theorem, we derive $\liminf_{n\to\infty} \int_{X\setminus\{x\}} v(y)q(\mathrm{d}y|x,a_n) \geq \int_{X\setminus\{x\}} v(y)q(\mathrm{d}y|x,a)$. Observe that $\lim_{n\to\infty} v(x)q(\{x\}|x,a_n) = v(x)q(\{x\}|x,a)$. Hence, we can obtain $\liminf_{n\to\infty} \int_X v(y)q(\mathrm{d}y|x,a_n) \geq \int_X v(y)q(\mathrm{d}y|x,a)$. Therefore, $\int_X v(y)q(\mathrm{d}y|x,a)$ is lower semi-continuous in $a \in A(x)$.

(b) For each $m \geq 1$, define $\widetilde{v}_m := \inf_{n\geq m} v_n$. Then, we have

$$\liminf_{n\to\infty} \int_{X\setminus\{x\}} v_n(y)q(\mathrm{d}y|x,a_n) \geq \liminf_{n\to\infty} \int_{X\setminus\{x\}} \widetilde{v}_m(y)q(\mathrm{d}y|x,a_n)$$
$$= \liminf_{n\to\infty} \left[\int_X \widetilde{v}_m(y)q(\mathrm{d}y|x,a_n) - \widetilde{v}_m(x)q(\{x\}|x,a_n) \right] \geq \int_{X\setminus\{x\}} \widetilde{v}_m(y)q(\mathrm{d}y|x,a),$$
$$(3.3)$$

where the last inequality is due to part (a). Letting $m \to \infty$ in (3.3) and using the monotone convergence theorem and (3.1), we can get the statement. $\qquad\square$

Let $\Lambda := \min\left\{\frac{\alpha^2}{\rho M}, \frac{\alpha}{2M}\right\}$ and the function w on X be as in Assumption 3.1. The notation $U_w((0, \Lambda) \times X)$ denotes the set of all real-valued measurable functions v on $(0, \Lambda) \times X$ which satisfy the following properties:

(i) $\sup_{(\lambda,x)\in(0,\Lambda)\times X} \frac{|v(\lambda,x)|}{w(x)} < \infty$;
(ii) for any $x \in X$ and $[\underline{\gamma}, \overline{\gamma}] \subset (0, \Lambda)$, $v(\cdot, x)$ is absolutely continuous on $[\underline{\gamma}, \overline{\gamma}]$ (this implies that the partial derivative of $v(\cdot, x)$ with respect to the variable λ exists almost everywhere (a.e.) $\lambda \in (0, \Lambda)$; see Remark 3.3 for details) and $\lambda \left|\frac{\partial v}{\partial \lambda}(\lambda, x)\right| \leq \overline{R}_v w^2(x)$ a.e. $\lambda \in (0, \Lambda)$ for some positive constant $\overline{R}_v$ independent of λ and x.

Remark 3.3 Let $v \in U_w((0, \Lambda) \times X)$ and fix any $x \in X$. Note that $(0, \Lambda) = \bigcup_{n=1}^{\infty} \left[\frac{1}{n}, \Lambda - \frac{1}{n}\right]$. For each $n \geq 1$, since $v(\cdot, x)$ is absolutely continuous on $\left[\frac{1}{n}, \Lambda - \frac{1}{n}\right]$, there exists $O_{x,n} \subset (0, \Lambda)$ with Lebesgue measure zero such that the partial derivative of $v(\cdot, x)$ with respect to the variable λ exists for all $\lambda \in \left[\frac{1}{n}, \Lambda - \frac{1}{n}\right] \setminus O_{x,n}$. Set $O_x := \bigcup_{n=1}^{\infty} O_{x,n}$. Then, the Lebesgue measure of O_x is zero. Moreover, the partial derivative of $v(\cdot, x)$ with respect to the variable λ exists for all $\lambda \in (0, \Lambda) \setminus O_x$. Hence, the partial derivative of $v(\cdot, x)$ with respect to the variable λ exists a.e. $\lambda \in (0, \Lambda)$.

Inspired by Theorem 3.1 in [10] and Lemma 3.2 in [16], we have the following Feynman–Kac formula which plays a key role in proving the existence of optimal policies.

Lemma 3.3 *Suppose that Assumptions 3.1(i), (ii) and 3.2(iii) hold. Then, for any bounded measurable function r on K, $v \in U_w((0, \Lambda) \times X)$, $T \geq 0$, $\lambda \in (0, \Lambda)$, $\pi \in \Pi$ and stopping time η, we have*

$$E_x^{\pi}\left[e^{\lambda \int_0^{T\wedge\eta} \int_A e^{-\alpha t} r(\xi_t, a)\pi(da|\omega,t)dt} v(\lambda e^{-\alpha(T\wedge\eta)}, \xi_{T\wedge\eta})\right] - v(\lambda, x)$$

$$= E_x^{\pi}\left[\int_0^{T\wedge\eta} e^{\lambda \int_0^s \int_A e^{-\alpha u} r(\xi_u, a)\pi(da|\omega,u)du}\right.$$

$$\left(\lambda e^{-\alpha s}\int_A r(\xi_s, a)\pi(da|\omega, s)v(\lambda e^{-\alpha s}, \xi_s) - \alpha\lambda e^{-\alpha s}\frac{\partial v}{\partial \lambda}(\lambda e^{-\alpha s}, \xi_s)\right.$$

$$\left.+ \int_X \int_A v(\lambda e^{-\alpha s}, y)q(dy|\xi_s, a)\pi(da|\omega, s)\right)ds\bigg]$$

for all $x \in X$, where $T \wedge \eta := \min\{T, \eta\}$.

Proof Fix any $x \in X$, $\pi \in \Pi$, $T \geq 0$, $\lambda \in (0, \Lambda)$ and $v \in U_w((0, \Lambda) \times X)$. Let $\|r\| := \sup_{(x,a)\in K} |r(x, a)|$, $\|v\|_w := \sup_{(\lambda,x)\in(0,\Lambda)\times X} \frac{|v(\lambda,x)|}{w(x)}$ and $G(\omega, t, x) := e^{\lambda \int_0^t \int_A e^{-\alpha s} r(\xi_s, a)\pi(da|\omega,s)ds} v(\lambda e^{-\alpha t}, x)$ for all $t \in [0, T]$. By the Dynkin formula, we have

$$E_x^{\pi}[w(\xi_{T\wedge\eta})] = w(x) + E_x^{\pi}\left[\int_0^{T\wedge\eta} \int_X \int_A w(y)q(dy|\xi_t, a)\pi(da|\omega, t)dt\right]. \quad (3.4)$$

Employing Assumptions 3.1(i), (ii), 3.2(iii), (3.4) and Lemma 3.1(a), we can derive

$$E_x^\pi \left[|G(\omega, T \wedge \eta, \xi_{T \wedge \eta})| \right] \le \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} E_x^\pi[w(\xi_{T \wedge \eta})]$$

$$\le \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} w(x) + \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} E_x^\pi \left[\int_0^T \int_X \int_A w(y)|q(dy|\xi_t, a)|\pi(da|\omega, t)dt \right]$$

$$\le \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} w(x) + \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} \int_0^T E_x^\pi \left[\int_X \int_A w(y)q(dy|\xi_t, a)\pi(da|\omega, t) + 2q^*(\xi_t)w(\xi_t) \right] dt$$

$$\le \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} w(x) + \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} \int_0^T E_x^\pi \left[\rho w(\xi_t) + d + 2Lw^2(\xi_t) \right] dt$$

$$\le \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} w(x) + \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} T \left[\rho e^{\rho T} \left(1 + \frac{d}{\rho} \right) w(x) + d + 2L \left(e^{\overline{\rho}T} w^2(x) + \frac{\overline{d}}{\rho} e^{\overline{\rho}T} \right) \right],$$

$$E_x^\pi \left[\int_0^{T \wedge \eta} e^{\lambda \int_0^s \int_A e^{-\alpha u} r(\xi_u, a)\pi(da|\omega, u)du} e^{-\alpha s} \int_A |r(\xi_s, a)|\pi(da|\omega, s)|v(\lambda e^{-\alpha s}, \xi_s)|ds \right]$$

$$\le \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} \|r\| \int_0^T E_x^\pi[w(\xi_s)]ds \le \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} \|r\| T e^{\rho T} \left(1 + \frac{d}{\rho} \right) w(x),$$

$$E_x^\pi \left[\int_0^{T \wedge \eta} e^{\lambda \int_0^s \int_A e^{-\alpha u} r(\xi_u, a)\pi(da|\omega, u)du} e^{-\alpha s} \lambda \left| \frac{\partial v}{\partial \lambda}(\lambda e^{-\alpha s}, \xi_s) \right| ds \right]$$

$$\le \overline{R}_v e^{\frac{\lambda\|r\|}{\alpha}} \int_0^T E_x^\pi[w^2(\xi_s)]ds \le \overline{R}_v e^{\frac{\lambda\|r\|}{\alpha}} T \left[e^{\overline{\rho}T} w^2(x) + \frac{\overline{d}}{\rho} e^{\overline{\rho}T} \right],$$

and

$$E_x^\pi \left[\int_0^{T \wedge \eta} e^{\lambda \int_0^s \int_A e^{-\alpha u} r(\xi_u, a)\pi(da|\omega, u)du} \int_X \int_A |v(\lambda e^{-\alpha s}, y)||q(dy|\xi_s, a)|\pi(da|\omega, s)ds \right]$$

$$\le \|v\|_w e^{\frac{\lambda\|r\|}{\alpha}} \int_0^T E_x^\pi \left[\int_X \int_A w(y)q(dy|\xi_s, a)\pi(da|\omega, s) + 2q^*(\xi_s)w(\xi_s) \right] ds$$

$$\le |v\|_w e^{\frac{\lambda\|r\|}{\alpha}} T \left[\rho e^{\rho T} \left(1 + \frac{d}{\rho} \right) w(x) + d + 2L \left(e^{\overline{\rho}T} w^2(x) + \frac{\overline{d}}{\rho} e^{\overline{\rho}T} \right) \right].$$

Because $v(\cdot, x)$ is absolutely continuous on $[\lambda e^{-\alpha T}, \lambda]$, we can obtain

$$G(\omega, T \wedge \eta, \xi_{T \wedge \eta}) = v(\lambda, x) + \int_0^{T \wedge \eta} G'(\omega, t, \xi_t)dt + \sum_{k \ge 1} \int_{(0, T \wedge \eta]} \Delta G(\omega, t, \xi_t)\delta_{T_k}(dt),$$

(3.5)

where $\Delta G(\omega, t, \xi_t) := G(\omega, t, \xi_t) - G(\omega, t-, \xi_{t-})$, G' denotes the derivative of G with respect to the variable t and $\delta_s(\cdot)$ represents the Dirac measure concentrated at s. Hence, using (3.5) and following the similar arguments of Theorem 3.1 in [10] or Lemma 3.2 in [16], we can get the assertion.   □

4 The Main Results

In this section, we show the existence of a unique solution to the risk-sensitive first passage discounted cost optimality equation and the existence of optimal policies. To this end, we introduce the following *new* value iteration.

Let $m(x) := Lw(x)$ for all $x \in X$, where the constant L and the function w on X are as in Assumption 3.1. Set

$$c_n(x, a) := \min\{c(x, a), n\} \text{ and } Q(dy|x, a) := \frac{q(dy|x, a)}{m(x)} + \delta_x(dy)$$

for all $(x, a) \in K$ and $n \geq 1$. For each $n \geq 1$, define $v_n^{(0)}(\lambda, x) := 1$ for all $(\lambda, x) \in (0, \Lambda) \times X$ and

$$
\begin{cases}
v_n^{(k+1)}(\lambda, x) := (\alpha\lambda)^{-\frac{m(x)}{\alpha}} \int_0^\lambda (\alpha t)^{\frac{m(x)}{\alpha}} \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^{(k)}(t, x) \right. \\
\qquad\qquad \left. + \frac{m(x)}{\alpha t} \int_X v_n^{(k)}(t, y) Q(dy|x, a) \right\} dt \text{ for all } (\lambda, x) \in (0, \Lambda) \times B^c \\
v_n^{(k+1)}(\lambda, x) := 1 \text{ for all } (\lambda, x) \in (0, \Lambda) \times B.
\end{cases}
$$

$$(4.1)$$

for all $k \geq 0$, where B^c stands for the complement of B with respect to X.

We have the following lemma which plays a key role in obtaining the existence of a unique solution to the risk-sensitive first passage discounted cost optimality equation.

Lemma 4.1 *Suppose that Assumptions 3.1 and 3.2 are satisfied. Then, the following statements hold for all $n \geq 1$.*

(a) *There exists a bounded measurable function v_n^* on $(0, \Lambda) \times X$ satisfying the following equation:*

$$
\begin{cases}
v_n^*(\lambda, x) = (\alpha\lambda)^{-\frac{m(x)}{\alpha}} \int_0^\lambda (\alpha t)^{\frac{m(x)}{\alpha}} \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^*(t, x) \right. \\
\qquad\qquad \left. + \frac{m(x)}{\alpha t} \int_X v_n^*(t, y) Q(dy|x, a) \right\} dt \; for \; all \; (\lambda, x) \in (0, \Lambda) \times B^c \\
v_n^*(\lambda, x) = 1 \; for \; all \; (\lambda, x) \in (0, \Lambda) \times B.
\end{cases}
$$

Moreover, we have $1 \leq v_n^*(\lambda, x) \leq e^{\frac{\lambda \|c_n\|}{\alpha}} \leq e^{\frac{\Lambda \|c_n\|}{\alpha}}$ for all $(\lambda, x) \in (0, \Lambda) \times X$, where $\|c_n\| := \max_{(x,a) \in K} c_n(x, a)$.

(b) $v_n^*(\lambda, x) = \inf_{\pi \in \Pi} E_x^\pi \left[e^{\lambda \int_0^{\tau_B} \int_A e^{-\alpha t} c_n(\xi_t, a) \pi(da|\omega, t) dt} \right]$ *for all $(\lambda, x) \in (0, \Lambda) \times X$.*

Proof (a) Fix any $n \geq 1$. By (4.1), Assumption 3.2, Lemma 3.2, Lemma 8.3.8 in [18] and an induction argument, we see that $v_n^{(k)}$ is measurable on $(0, \Lambda) \times X$ for all $k \geq 0$.

Below, we prove the following fact that

$$v_n^{(k)} \le v_n^{(k+1)} \text{ for all } k \ge 0. \tag{4.2}$$

Indeed, by (4.1) direct computations imply

$$v_n^{(1)}(\lambda, x) \ge (\alpha\lambda)^{-\frac{m(x)}{\alpha}} \int_0^\lambda (\alpha t)^{\frac{m(x)}{\alpha}} \frac{m(x)}{\alpha t} dt = v_n^{(0)}(\lambda, x)$$

for all $(\lambda, x) \in (0, \Lambda) \times B^c$. Thus, (4.2) is true for $k = 0$. Suppose that (4.2) holds for some $k_0 \ge 0$. Then, it follows from (4.1) and the induction hypothesis that

$$
\begin{aligned}
v_n^{(k_0+2)}(\lambda, x) &\ge (\alpha\lambda)^{-\frac{m(x)}{\alpha}} \int_0^\lambda (\alpha t)^{\frac{m(x)}{\alpha}} \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^{(k_0)}(t, x) \right. \\
&\quad \left. + \frac{m(x)}{\alpha t} \int_X v_n^{(k_0)}(t, y) Q(dy|x, a) \right\} dt \\
&= v_n^{(k_0+1)}(\lambda, x)
\end{aligned}
$$

for all $(\lambda, x) \in (0, \Lambda) \times B^c$. Hence, (4.2) holds for $k = k_0 + 1$. Therefore, we can derive (4.2) from the induction argument. Let $v_n^*(\lambda, x) := \lim_{k \to \infty} v_n^{(k)}(\lambda, x)$ for all $(\lambda, x) \in (0, \Lambda) \times X$. Then, we have $v_n^*(\lambda, x) = 1$ for all $(\lambda, x) \in (0, \Lambda) \times B$. For each $k \ge 0$ and $(t, x) \in (0, \Lambda) \times B^c$, by Assumption 3.2, Lemma 3.2 and Theorem 2.43 in [1], there exists $a_{t,x}^{(k)} \in A(x)$ such that

$$
\begin{aligned}
&\inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^{(k)}(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^{(k)}(t, y) Q(dy|x, a) \right\} \\
&= \frac{1}{\alpha} c_n(x, a_{t,x}^{(k)}) v_n^{(k)}(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^{(k)}(t, y) Q(dy|x, a_{t,x}^{(k)}). \tag{4.3}
\end{aligned}
$$

Because $A(x)$ is compact, there exists a subsequence of $\{a_{t,x}^{(k)}, k \ge 0\}$ (denoted by the same sequence) such that $a_{t,x}^{(k)}$ converges to some $\widetilde{a}_{t,x} \in A(x)$. Employing (4.1) and an induction argument, we see that

$$1 \le v_n^{(k)}(t, x) \le e^{\frac{\lambda \|c_n\|}{\alpha}} \tag{4.4}$$

for all $(t, x) \in (0, \Lambda) \times B^c$ and $k \ge 0$. Then, it follows from (4.3), (4.4), Assumption 3.2 and Lemma 3.2 that

$$
\begin{aligned}
&\lim_{k \to \infty} \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^{(k)}(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^{(k)}(t, y) Q(dy|x, a) \right\} \\
&\ge \frac{1}{\alpha} c_n(x, \widetilde{a}_{t,x}) v_n^*(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^*(t, y) Q(dy|x, \widetilde{a}_{t,x}) \\
&\ge \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^*(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^*(t, y) Q(dy|x, a) \right\} \tag{4.5}
\end{aligned}
$$

for all $(t, x) \in (0, \Lambda) \times B^c$. Moreover, we have

$$\inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^{(k)}(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^{(k)}(t, y) Q(\mathrm{d}y | x, a) \right\}$$
$$\leq \frac{1}{\alpha} c_n(x, a) v_n^{(k)}(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^{(k)}(t, y) Q(\mathrm{d}y | x, a)$$

for all $k \geq 0$, which yields

$$\lim_{k \to \infty} \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^{(k)}(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^{(k)}(t, y) Q(\mathrm{d}y | x, a) \right\}$$
$$\leq \frac{1}{\alpha} c_n(x, a) v_n^*(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^*(t, y) Q(\mathrm{d}y | x, a)$$

for all $(t, x) \in (0, \Lambda) \times B^c$ and $a \in A(x)$. Thus, from the last inequality and (4.5), we get

$$\lim_{k \to \infty} \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^{(k)}(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^{(k)}(t, y) Q(\mathrm{d}y | x, a) \right\}$$
$$= \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^*(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^*(t, y) Q(\mathrm{d}y | x, a) \right\} \qquad (4.6)$$

for all $(t, x) \in (0, \Lambda) \times B^c$. Hence, using (4.1), (4.4), (4.6) and the monotone convergence theorem, we derive

$$v_n^*(\lambda, x) = (\alpha \lambda)^{-\frac{m(x)}{\alpha}} \int_0^\lambda (\alpha t)^{\frac{m(x)}{\alpha}} \inf_{a \in A(x)}$$
$$\left\{ \frac{1}{\alpha} c_n(x, a) v_n^*(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^*(t, y) Q(\mathrm{d}y | x, a) \right\} \mathrm{d}t$$

for all $(\lambda, x) \in (0, \Lambda) \times B^c$. Furthermore, by (4.4) we can obtain $1 \leq v_n^*(\lambda, x) \leq e^{\frac{\lambda \|c_n\|}{\alpha}} \leq e^{\frac{\Lambda \|c_n\|}{\alpha}}$ for all $(\lambda, x) \in (0, \Lambda) \times B^c$.

(b) The assertion is obviously true for all $(\lambda, x) \in (0, \Lambda) \times B$. Below, we show that the assertion is true for all $(\lambda, x) \in (0, \Lambda) \times B^c$. Fix any $n \geq 1$. For each $x \in B^c$, by part (a) we have

$$(\alpha \lambda)^{\frac{m(x)}{\alpha}} v_n^*(\lambda, x) = \int_0^\lambda (\alpha t)^{\frac{m(x)}{\alpha}} \inf_{a \in A(x)}$$
$$\left\{ \frac{1}{\alpha} c_n(x, a) v_n^*(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^*(t, y) Q(\mathrm{d}y | x, a) \right\} \mathrm{d}t \quad (4.7)$$

for all $\lambda \in (0, \Lambda)$. For any $x \in B^c$ and $[\underline{\gamma}, \overline{\gamma}] \subset (0, \Lambda)$, employing Theorem 3.11, Exercise 22 in [23, p. 130, 149] and (4.7), we see that $v_n^*(\cdot, x)$ is absolutely continuous on $[\underline{\gamma}, \overline{\gamma}]$. Thus, for each $x \in B^c$, the partial derivative of $v_n^*(\cdot, x)$ with respect to the

variable λ exists a.e. $\lambda \in (0, \Lambda)$. Calculating the derivative with respect to the variable λ in (4.7), for each $x \in B^c$, we derive

$$
\begin{aligned}
&(\alpha\lambda)^{\frac{m(x)}{\alpha}} \frac{m(x)}{\alpha\lambda} v_n^*(\lambda, x) + (\alpha\lambda)^{\frac{m(x)}{\alpha}} \frac{\partial v_n^*}{\partial\lambda}(\lambda, x) \\
&= (\alpha\lambda)^{\frac{m(x)}{\alpha}} \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^*(\lambda, x) + \frac{m(x)}{\alpha\lambda} \int_X v_n^*(\lambda, y) Q(dy|x, a) \right\} \\
&= (\alpha\lambda)^{\frac{m(x)}{\alpha}} \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^*(\lambda, x) + \frac{m(x)}{\alpha\lambda} \left(\frac{1}{m(x)} \int_X v_n^*(\lambda, y) q(dy|x, a) + v_n^*(\lambda, x) \right) \right\}
\end{aligned}
\tag{4.8}
$$

a.e. $\lambda \in (0, \Lambda)$. Then, using (4.8), for each $x \in B^c$, there exists $O_x \subset (0, \Lambda)$ with Lebesgue measure zero such that

$$
\frac{\partial v_n^*}{\partial\lambda}(\lambda, x) = \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^*(\lambda, x) + \frac{1}{\alpha\lambda} \int_X v_n^*(\lambda, y) q(dy|x, a) \right\}
\tag{4.9}
$$

for all $\lambda \in O_x^c$, where O_x^c denotes the complement of O_x with respect to $(0, \Lambda)$. Thus, it follows from (4.9), Assumption 3.1 and part (a) that

$$
\lambda \left| \frac{\partial v_n^*}{\partial\lambda}(\lambda, x) \right| \leq \frac{\Lambda}{\alpha} \|c_n\| e^{\frac{\Lambda\|c_n\|}{\alpha}} + \frac{2}{\alpha} e^{\frac{\Lambda\|c_n\|}{\alpha}} q^*(x) \leq \left(\frac{\Lambda}{\alpha} \|c_n\| e^{\frac{\Lambda\|c_n\|}{\alpha}} + \frac{2}{\alpha} e^{\frac{\Lambda\|c_n\|}{\alpha}} L \right) w(x)
$$

for all $x \in X$ and $\lambda \in O_x^c$. Hence, employing Lemma 3.3 we can get

$$
\begin{aligned}
&E_x^\pi \left[e^{\lambda \int_0^{T \wedge \tau_B} \int_A e^{-\alpha t} c_n(\xi_t, a) \pi(da|\omega, t) dt} v_n^*(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right] - v_n^*(\lambda, x) \\
&= E_x^\pi \left[\int_0^{T \wedge \tau_B} e^{\lambda \int_0^s \int_A e^{-\alpha u} c_n(\xi_u, a) \pi(da|\omega, u) du} \left(\lambda e^{-\alpha s} \int_A c_n(\xi_s, a) \pi(da|\omega, s) v_n^*(\lambda e^{-\alpha s}, \xi_s) \right. \right. \\
&\qquad \left. \left. - \alpha\lambda e^{-\alpha s} \frac{\partial v_n^*}{\partial\lambda}(\lambda e^{-\alpha s}, \xi_s) + \int_X \int_A v_n^*(\lambda e^{-\alpha s}, y) q(dy|\xi_s, a) \pi(da|\omega, s) \right) ds \right]
\end{aligned}
\tag{4.10}
$$

for all $x \in X$, $\lambda \in (0, \Lambda)$, $T \geq 0$ and $\pi \in \Pi$. By Assumption 3.2, Lemma 3.2 and Lemma 8.3.8 in [18], there exists a measurable mapping $f_n^* : (0, \Lambda) \times X \to A$ with $f_n^*(\lambda, x) \in A(x)$ for all $(\lambda, x) \in (0, \Lambda) \times X$ such that

$$
\begin{aligned}
&\inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^*(\lambda, x) + \frac{1}{\alpha\lambda} \int_X v_n^*(\lambda, y) q(dy|x, a) \right\} \\
&= \frac{1}{\alpha} c_n(x, f_n^*(\lambda, x)) v_n^*(\lambda, x) + \frac{1}{\alpha\lambda} \int_X v_n^*(\lambda, y) q(dy|x, f_n^*(\lambda, x))
\end{aligned}
\tag{4.11}
$$

for all $(\lambda, x) \in (0, \Lambda) \times X$. Let $\pi_n^*(\cdot|\omega, t) := \delta_{f_n^*(\lambda e^{-\alpha t}, \xi_{t-})}(\cdot)$ for all $t \geq 0$. Using (4.9)–(4.11) we can obtain

$$
v_n^*(\lambda, x) = E_x^{\pi_n^*} \left[e^{\lambda \int_0^{T \wedge \tau_B} e^{-\alpha t} c_n(\xi_t, f_n^*(\lambda e^{-\alpha t}, \xi_t)) dt} v_n^*(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right]
\tag{4.12}
$$

for all $x \in B^c$, $\lambda \in (0, \Lambda)$ and $T \geq 0$. Note that part (a) implies $v_n^*(0, x) = 1$ for all $x \in X$. Employing (4.12) we have

$$
\begin{aligned}
v_n^*(\lambda, x) = E_x^{\pi_n^*} & \left[e^{\lambda \int_0^{T \wedge \tau_B} e^{-\alpha t} c_n(\xi_t, f_n^*(\lambda e^{-\alpha t}, \xi_t)) \mathrm{d}t} v_n^*(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) I_{\{\tau_B < \infty\}} \right] \\
& + E_x^{\pi_n^*} \left[e^{\lambda \int_0^{T} e^{-\alpha t} c_n(\xi_t, f_n^*(\lambda e^{-\alpha t}, \xi_t)) \mathrm{d}t} v_n^*(\lambda e^{-\alpha T}, \xi_T) I_{\{\tau_B = \infty\}} \right] \quad (4.13)
\end{aligned}
$$

for all $x \in B^c$, $\lambda \in (0, \Lambda)$ and $T \geq 0$. Therefore, letting $T \to \infty$ in (4.13), by part (a) and the dominated convergence theorem, we can derive

$$
\begin{aligned}
v_n^*(\lambda, x) & = E_x^{\pi_n^*} \left[e^{\lambda \int_0^{\tau_B} e^{-\alpha t} c_n(\xi_t, f_n^*(\lambda e^{-\alpha t}, \xi_t)) \mathrm{d}t} I_{\{\tau_B < \infty\}} \right] \\
& \quad + E_x^{\pi_n^*} \left[e^{\lambda \int_0^{\infty} e^{-\alpha t} c_n(\xi_t, f_n^*(\lambda e^{-\alpha t}, \xi_t)) \mathrm{d}t} I_{\{\tau_B = \infty\}} \right] \\
& = E_x^{\pi_n^*} \left[e^{\lambda \int_0^{\tau_B} e^{-\alpha t} c_n(\xi_t, f_n^*(\lambda e^{-\alpha t}, \xi_t)) \mathrm{d}t} \right] \\
& \geq \inf_{\pi \in \Pi} E_x^{\pi} \left[e^{\lambda \int_0^{\tau_B} \int_A e^{-\alpha t} c_n(\xi_t, a) \pi(\mathrm{d}a|\omega, t) \mathrm{d}t} \right] \quad (4.14)
\end{aligned}
$$

for all $x \in B^c$ and $\lambda \in (0, \Lambda)$. On the other hand, it follows from (4.9) and (4.11) that

$$
v_n^*(\lambda, x) \leq E_x^{\pi} \left[e^{\lambda \int_0^{T \wedge \tau_B} \int_A e^{-\alpha t} c_n(\xi_t, a) \pi(\mathrm{d}a|\omega, t) \mathrm{d}t} v_n^*(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right]
$$

for all $x \in B^c$, $\lambda \in (0, \Lambda)$, $T \geq 0$ and $\pi \in \Pi$. Thus, using the last inequality and the similar arguments as (4.14), we can get

$$
v_n^*(\lambda, x) \leq E_x^{\pi} \left[e^{\lambda \int_0^{\tau_B} \int_A e^{-\alpha t} c_n(\xi_t, a) \pi(\mathrm{d}a|\omega, t) \mathrm{d}t} \right] \quad (4.15)
$$

for all $x \in B^c$, $\lambda \in (0, \Lambda)$ and $\pi \in \Pi$. Hence, the assertion follows from (4.14) and (4.15).   □

Denote by $V_w((0, \Lambda) \times X)$ the set of all nonnegative real-valued measurable functions $v \in U_w((0, \Lambda) \times X)$ and $v(\lambda, x) \leq R_\lambda w^{\frac{\lambda M}{\alpha}}(x)$ for all $(\lambda, x) \in (0, \Lambda) \times X$, where the constant R_λ is as in Lemma 3.1. Employing Lemma 4.1, we can derive the existence of a unique solution in $V_w((0, \Lambda) \times X)$ to the risk-sensitive first passage discounted cost optimality equation and the existence of optimal policies in the following theorem.

Theorem 4.1 *Under Assumptions 3.1 and 3.2, the following assertions are true.*

(a) *There exists a measurable function $v^* \in V_w((0, \Lambda) \times X)$ satisfying $v^*(\lambda, x) = 1$ for all $(\lambda, x) \in (0, \Lambda) \times B$ and*

$$
\frac{\partial v^*}{\partial \lambda}(\lambda, x) = \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c(x, a) v^*(\lambda, x) + \frac{1}{\alpha \lambda} \int_X v^*(\lambda, y) q(\mathrm{d}y|x, a) \right\} \quad (4.16)
$$

for all $x \in B^c$ and a.e. $\lambda \in (0, \Lambda)$.

(b) *There exists a measurable mapping $f^* : (0, \Lambda) \times X \to A$ with $f^*(\lambda, x) \in A(x)$ for all $(\lambda, x) \in (0, \Lambda) \times X$ such that*

$$\inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c(x, a) v^*(\lambda, x) + \frac{1}{\alpha \lambda} \int_X v^*(\lambda, y) q(dy|x, a) \right\}$$
$$= \frac{1}{\alpha} c(x, f^*(\lambda, x)) v^*(\lambda, x) + \frac{1}{\alpha \lambda} \int_X v^*(\lambda, y) q(dy|x, f^*(\lambda, x))$$

for all $(\lambda, x) \in (0, \Lambda) \times X$. Let $\pi^*(\cdot|\omega, t) := \delta_{f^*(\lambda e^{-\alpha t}, \xi_{t-})}(\cdot)$ for all $t \geq 0$. Then, we have $v^*(\lambda, x) = \widehat{J}(\lambda, x, \pi^*) = \inf_{\pi \in \Pi} \widehat{J}(\lambda, x, \pi)$ for all $(\lambda, x) \in (0, \Lambda) \times X$. Hence, there exists a deterministic Markov optimal policy under the risk-sensitive first passage discounted cost criterion.

(c) *If there exists a measurable function $v \in V_w((0, \Lambda) \times X)$ satisfying $v(\lambda, x) = 1$ for all $(\lambda, x) \in (0, \Lambda) \times B$ and (4.16), then we have $v(\lambda, x) = \inf_{\pi \in \Pi} \widehat{J}(\lambda, x, \pi)$ for all $(\lambda, x) \in (0, \Lambda) \times X$.*

Proof (a) By Lemmas 3.1 and 4.1, we can obtain

$$1 \leq v_n^*(\lambda, x) \leq \inf_{\pi \in \Pi} E_x^\pi \left[e^{\lambda \int_0^{\tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} \right] \leq R_\lambda w^{\frac{\lambda M}{\alpha}}(x) \qquad (4.17)$$

for all $(\lambda, x) \in (0, \Lambda) \times X$ and $n \geq 1$. From Lemma 4.1(b), we see that v_n^* is nondecreasing in n. Let $v^*(\lambda, x) := \lim_{n \to \infty} v_n^*(\lambda, x)$ for all $(\lambda, x) \in (0, \Lambda) \times X$. Lemma 4.1(a) gives $v^*(\lambda, x) = 1$ for all $(\lambda, x) \in (0, \Lambda) \times B$. Employing the similar arguments as (4.6), we can get

$$\lim_{n \to \infty} \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c_n(x, a) v_n^*(t, x) + \frac{m(x)}{\alpha t} \int_X v_n^*(t, y) Q(dy|x, a) \right\}$$
$$= \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c(x, a) v^*(t, x) + \frac{m(x)}{\alpha t} \int_X v^*(t, y) Q(dy|x, a) \right\} \qquad (4.18)$$

for all $(t, x) \in (0, \Lambda) \times B^c$. Moreover, using Lemma 4.1(a), (4.18) and the monotone convergence theorem, we derive

$$v^*(\lambda, x) = (\alpha \lambda)^{-\frac{m(x)}{\alpha}} \int_0^\lambda (\alpha t)^{\frac{m(x)}{\alpha}} \inf_{a \in A(x)}$$
$$\left\{ \frac{1}{\alpha} c(x, a) v^*(t, x) + \frac{m(x)}{\alpha t} \int_X v^*(t, y) Q(dy|x, a) \right\} dt \qquad (4.19)$$

for all $(t, x) \in (0, \Lambda) \times B^c$. Below, we show that $v^* \in V_w((0, \Lambda) \times X)$. In fact, (4.17) implies

$$1 \leq v^*(\lambda, x) \leq R_\lambda w^{\frac{\lambda M}{\alpha}}(x) \qquad (4.20)$$

for all $(\lambda, x) \in (0, \Lambda) \times X$. For each $x \in B^c$, by (4.19) we get

$$(\alpha\lambda)^{\frac{m(x)}{\alpha}} v^*(\lambda, x) = \int_0^\lambda (\alpha t)^{\frac{m(x)}{\alpha}} \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c(x, a) v^*(t, x) + \frac{m(x)}{\alpha t} \int_X v^*(t, y) Q(dy|x, a) \right\} dt \quad (4.21)$$

for all $\lambda \in (0, \Lambda)$. For any $x \in B^c$ and $[\underline{\gamma}, \overline{\gamma}] \subset (0, \Lambda)$, using Theorem 3.11, Exercise 22 in [23, p.130, 149] and (4.21), we have that $v^*(\cdot, x)$ is absolutely continuous on $[\underline{\gamma}, \overline{\gamma}]$. Similar to the calculations as (4.9), employing (4.21) we can obtain

$$\frac{\partial v^*}{\partial \lambda}(\lambda, x) = \inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c(x, a) v^*(\lambda, x) + \frac{1}{\alpha\lambda} \int_X v^*(\lambda, y) q(dy|x, a) \right\} \quad (4.22)$$

a.e. $\lambda \in (0, \Lambda)$. Thus, by (4.22) and Assumption 3.1, for any $x \in B^c$, we have

$$\lambda \left| \frac{\partial v^*}{\partial \lambda}(\lambda, x) \right| \leq \frac{\Lambda}{\alpha}(M \ln w(x) + b) R_\lambda w(x) + \frac{R_\lambda}{\alpha} \int_X w(y)|q(dy|x, a)|$$

$$\leq \frac{\Lambda}{\alpha}(M + b) R_\lambda w^2(x) + \frac{R_\lambda}{\alpha} \left[\int_X w(y) q(dy|x, a) + 2q^*(x) w(x) \right]$$

$$\leq \left[\frac{\Lambda}{\alpha}(M + b) R_\Lambda + \frac{R_\Lambda}{\alpha}(\rho + d + 2L) \right] w^2(x)$$

a.e. $\lambda \in (0, \Lambda)$. Note that $v^*(\lambda, x) = 1$ and $\frac{\partial v^*}{\partial \lambda}(\lambda, x) = 0$ for all $(\lambda, x) \in (0, \Lambda) \times B$. Hence, we get $v^* \in V_w((0, \Lambda) \times X)$.

(b) By Assumption 3.2(ii) and Lemma 3.2, we see that for each $(\lambda, x) \in (0, \Lambda) \times X$, $\frac{1}{\alpha} c(x, a) v^*(\lambda, x) + \frac{1}{\alpha\lambda} \int_X v^*(\lambda, y) q(dy|x, a)$ is lower semi-continuous in $a \in A(x)$. Then, using Assumption 3.2(i) and Lemma 8.3.8 in [18], we can derive the existence of a measurable mapping $f^* : (0, \Lambda) \times X \to A$ with $f^*(\lambda, x) \in A(x)$ for all $(\lambda, x) \in (0, \Lambda) \times X$ which satisfies

$$\inf_{a \in A(x)} \left\{ \frac{1}{\alpha} c(x, a) v^*(\lambda, x) + \frac{1}{\alpha\lambda} \int_X v^*(\lambda, y) q(dy|x, a) \right\}$$

$$= \frac{1}{\alpha} c(x, f^*(\lambda, x)) v^*(\lambda, x) + \frac{1}{\alpha\lambda} \int_X v^*(\lambda, y) q(dy|x, f^*(\lambda, x)) \quad (4.23)$$

for all $(\lambda, x) \in (0, \Lambda) \times X$. Let $\pi^*(\cdot|\omega, t) := \delta_{f^*(\lambda e^{-\alpha t}, \xi_{t-})}(\cdot)$ for all $t \geq 0$. Part (a) and Lemma 3.3 give

$$E_x^{\pi^*} \left[e^{\lambda \int_0^{T \wedge \tau_B} e^{-\alpha t} c_n(\xi_t, f^*(\lambda e^{-\alpha t}, \xi_t)) dt} v^*(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right] - v^*(\lambda, x)$$

$$= E_x^{\pi^*} \left[\int_0^{T \wedge \tau_B} e^{\lambda \int_0^s e^{-\alpha u} c_n(\xi_u, f^*(\lambda e^{-\alpha u}, \xi_u)) du} \left(\lambda e^{-\alpha s} c_n(\xi_s, f^*(\lambda e^{-\alpha s}, \xi_s)) v^*(\lambda e^{-\alpha s}, \xi_s) \right. \right.$$

$$\left. \left. - \alpha\lambda e^{-\alpha s} \frac{\partial v^*}{\partial \lambda}(\lambda e^{-\alpha s}, \xi_s) + \int_X v^*(\lambda e^{-\alpha s}, y) q(dy|\xi_s, f^*(\lambda e^{-\alpha s}, \xi_s)) \right) ds \right] \quad (4.24)$$

for all $x \in X$, $\lambda \in (0, \Lambda)$, $T \geq 0$ and $n \geq 1$. From part (a) and (4.23), for each $x \in B^c$, we have

$$\frac{\partial v^*}{\partial \lambda}(\lambda, x) \geq \frac{1}{\alpha} c_n(x, f^*(\lambda, x)) v^*(\lambda, x) + \frac{1}{\alpha \lambda} \int_X v^*(\lambda, y) q(\mathrm{d}y | x, f^*(\lambda, x))$$

a.e. $\lambda \in (0, \Lambda)$. Thus, it follows from (4.20), (4.24) and the last inequality that

$$
\begin{aligned}
v^*(\lambda, x) &\geq E_x^{\pi^*} \left[e^{\lambda \int_0^{T \wedge \tau_B} e^{-\alpha t} c_n(\xi_t, f^*(\lambda e^{-\alpha t}, \xi_t)) \mathrm{d}t} v^*(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right] \\
&\geq E_x^{\pi^*} \left[e^{\lambda \int_0^{T \wedge \tau_B} e^{-\alpha t} c_n(\xi_t, f^*(\lambda e^{-\alpha t}, \xi_t)) \mathrm{d}t} \right]
\end{aligned}
\tag{4.25}
$$

for all $x \in X$, $\lambda \in (0, \Lambda)$, $T \geq 0$ and $n \geq 1$. Letting $n \to \infty$ in (4.25), we get

$$v^*(\lambda, x) \geq E_x^{\pi^*} \left[e^{\lambda \int_0^{T \wedge \tau_B} e^{-\alpha t} c(\xi_t, f^*(\lambda e^{-\alpha t}, \xi_t)) \mathrm{d}t} \right]$$

for all $x \in X$, $\lambda \in (0, \Lambda)$ and $T \geq 0$. Moreover, letting $T \to \infty$ in the last inequality, the monotone convergence theorem implies

$$v^*(\lambda, x) \geq E_x^{\pi^*} \left[e^{\lambda \int_0^{\tau_B} e^{-\alpha t} c(\xi_t, f^*(\lambda e^{-\alpha t}, \xi_t)) \mathrm{d}t} \right] \geq \inf_{\pi \in \Pi} \widehat{J}(\lambda, x, \pi) \tag{4.26}$$

for all $x \in X$ and $\lambda \in (0, \Lambda)$.

Below, we show that

$$v^*(\lambda, x) \leq \inf_{\pi \in \Pi} \widehat{J}(\lambda, x, \pi) \text{ for all } x \in X \text{ and } \lambda \in (0, \Lambda). \tag{4.27}$$

Let $\widehat{c}(x) := \max_{a \in A(x)} c(x, a)$ for all $x \in X$, $Y_k := \{x : \widehat{c}(x) > k\}$ and $\eta_{Y_k} := \inf\{t \geq 0 : \xi_t \in Y_k\}$ for all $k \geq 1$. Then, for any $n > k$, we can obtain

$$
\begin{aligned}
&E_x^{\pi} \left[e^{\lambda \int_0^{T \wedge \tau_B \wedge \eta_{Y_k}} \int_A e^{-\alpha t} c_n(\xi_t, a) \pi(\mathrm{d}a | \omega, t) \mathrm{d}t} v^*(\lambda e^{-\alpha(T \wedge \tau_B \wedge \eta_{Y_k})}, \xi_{T \wedge \tau_B \wedge \eta_{Y_k}}) \right] - v^*(\lambda, x) \\
&= E_x^{\pi} \left[\int_0^{T \wedge \tau_B \wedge \eta_{Y_k}} e^{\lambda \int_0^s \int_A e^{-\alpha u} c_n(\xi_u, a) \pi(\mathrm{d}a | \omega, u) \mathrm{d}u} \left(\lambda e^{-\alpha s} \int_A c_n(\xi_s, a) \pi(\mathrm{d}a | \omega, s) v^*(\lambda e^{-\alpha s}, \xi_s) \right. \right. \\
&\quad \left. \left. - \alpha \lambda e^{-\alpha s} \frac{\partial v^*}{\partial \lambda}(\lambda e^{-\alpha s}, \xi_s) + \int_X \int_A v^*(\lambda e^{-\alpha s}, y) q(\mathrm{d}y | \xi_s, a) \pi(\mathrm{d}a | \omega, s) \right) \mathrm{d}s \right] \\
&= E_x^{\pi} \left[\int_0^{T \wedge \tau_B \wedge \eta_{Y_k}} e^{\lambda \int_0^s \int_A e^{-\alpha u} c(\xi_u, a) \pi(\mathrm{d}a | \omega, u) \mathrm{d}u} \left(\lambda e^{-\alpha s} \int_A c(\xi_s, a) \pi(\mathrm{d}a | \omega, s) v^*(\lambda e^{-\alpha s}, \xi_s) \right. \right. \\
&\quad \left. \left. - \alpha \lambda e^{-\alpha s} \frac{\partial v^*}{\partial \lambda}(\lambda e^{-\alpha s}, \xi_s) + \int_X \int_A v^*(\lambda e^{-\alpha s}, y) q(\mathrm{d}y | \xi_s, a) \pi(\mathrm{d}a | \omega, s) \right) \mathrm{d}s \right] \geq 0 \quad (4.28)
\end{aligned}
$$

for all $x \in X$, $\lambda \in (0, \Lambda)$, $T \geq 0$, $k \geq 1$ and $\pi \in \Pi$, where the first equality follows from part (a) and Lemma 3.3, and the inequality is due to part (a). Employing (4.28)

we derive

$$
v^*(\lambda, x) \leq E_x^\pi \left[e^{\lambda \int_0^{T \wedge \tau_B \wedge \eta_{Y_k}} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} v^*(\lambda e^{-\alpha(T \wedge \tau_B \wedge \eta_{Y_k})}, \xi_{T \wedge \tau_B \wedge \eta_{Y_k}}) \right]
$$
(4.29)

for all $x \in X$, $\lambda \in (0, \Lambda)$, $T \geq 0$, $k \geq 1$ and $\pi \in \Pi$. By Lemma 4.1(b), we see that $v^*(\lambda, x)$ is nondecreasing in λ for all $x \in X$. Moreover, we have

$$
e^{\lambda \int_0^{T \wedge \tau_B \wedge \eta_{Y_k}} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} v^*(\lambda e^{-\alpha(T \wedge \tau_B \wedge \eta_{Y_k})}, \xi_{T \wedge \tau_B \wedge \eta_{Y_k}})
$$
$$
\leq e^{\lambda \int_0^T \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} v^*(\lambda, \xi_{T \wedge \tau_B \wedge \eta_{Y_k}})
$$
$$
\leq R_\lambda e^{\lambda \int_0^T \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} w^{\frac{\lambda M}{\alpha}}(\xi_{T \wedge \tau_B \wedge \eta_{Y_k}})
$$
$$
\leq R_\lambda^2 e^{2\lambda \int_0^T \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} + w^{\frac{2\lambda M}{\alpha}}(\xi_{T \wedge \tau_B \wedge \eta_{Y_k}})
$$
$$
\leq R_\lambda^2 e^{2\lambda \int_0^T \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} + w(\xi_{T \wedge \tau_B \wedge \eta_{Y_k}})
$$
(4.30)

for all $\lambda \in (0, \Lambda)$, $T \geq 0$ and $k \geq 1$, where the second inequality follows from (4.20). From the Dynkin formula, we can get

$$
E_x^\pi[w(\xi_{T \wedge \tau_B \wedge \eta_{Y_k}})] = w(x) + E_x^\pi \left[\int_0^{T \wedge \tau_B \wedge \eta_{Y_k}} \int_X \int_A w(y) q(dy|\xi_t, a) \pi(da|\omega, t) dt \right]
$$
(4.31)

for all $x \in X$, $\pi \in \Pi$, $T \geq 0$ and $k \geq 1$. Using Assumptions 3.1, 3.2 and Lemma 3.1(a), we obtain

$$
E_x^\pi \left[\int_0^T \int_X \int_A w(y) |q(dy|\xi_t, a)| \pi(da|\omega, t) dt \right]
$$
$$
\leq \int_0^T E_x^\pi[\rho w(\xi_t) + d + 2q^*(\xi_t) w(\xi_t)] dt
$$
$$
\leq (\rho + d + 2L) \int_0^T E_x^\pi[w^2(\xi_t)] dt \leq (\rho + d + 2L) T \left[e^{\overline{\rho} T} w^2(x) + \frac{\overline{d}}{\overline{\rho}} e^{\overline{\rho} T} \right]
$$
(4.32)

for all $x \in X$, $\pi \in \Pi$ and $T \geq 0$. Thus, we can derive

$$
\lim_{k \to \infty} E_x^\pi[w(\xi_{T \wedge \tau_B \wedge \eta_{Y_k}})] = w(x) + E_x^\pi \left[\int_0^{T \wedge \tau_B} \int_X \int_A w(y) q(dy|\xi_t, a) \pi(da|\omega, t) dt \right]
$$
$$
= E_x^\pi[w(\xi_{T \wedge \tau_B})]
$$
(4.33)

for all $x \in X$, $\pi \in \Pi$ and $T \geq 0$, where the first equality follows from (4.31) and the dominated convergence theorem, and the second one is due to the Dynkin formula. Direct calculations give

$$
\begin{aligned}
&E_x^\pi \left[R_\lambda^2 e^{2\lambda \int_0^T \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} + w(\xi_{T \wedge \tau_B}) \right] \\
&\leq R_\lambda^2 E_x^\pi \left[\int_0^T e^{\frac{2\lambda}{\alpha}(1 - e^{-\alpha T}) \int_A c(\xi_t, a) \pi(da|\omega, t)} \frac{\alpha e^{-\alpha t}}{1 - e^{-\alpha T}} dt \right] + w(x) \\
&\quad + E_x^\pi \left[\int_0^T \int_X \int_A w(y) |q(dy|\xi_t, a)| \pi(da|\omega, t) dt \right] \\
&\leq \frac{R_\lambda^2 \alpha e^{\frac{2\lambda b}{\alpha}}}{1 - e^{-\alpha T}} \int_0^T e^{-\alpha t} E_x^\pi[w(\xi_t)] dt + w(x) + (\rho + d + 2L)T \left[e^{\overline{\rho}T} w^2(x) + \frac{\overline{d}}{\rho} e^{\overline{\rho}T} \right] \\
&\leq \frac{R_\lambda^2 \alpha e^{\frac{2\lambda b}{\alpha}} (e^{(\rho - \alpha)T} - 1)}{(1 - e^{-\alpha T})(\rho - \alpha)} \left(1 + \frac{d}{\rho}\right) w(x) + w(x) + (\rho + d + 2L)T \left[e^{\overline{\rho}T} w^2(x) + \frac{\overline{d}}{\rho} e^{\overline{\rho}T} \right]
\end{aligned}
$$

for all $x \in X$, $\lambda \in (0, \Lambda)$, $\pi \in \Pi$ and $T \geq 0$, where the first inequality follows from the Jensen inequality and (4.33), the second one is due to Assumption 3.1 and (4.32), and the last one follows from Lemma 3.1(a). Hence, by (4.30)–(4.33), the last inequality and the generalized dominated convergence theorem (see Theorem 2.88 in [3]), we have

$$
\begin{aligned}
&\lim_{k \to \infty} E_x^\pi \left[e^{\lambda \int_0^{T \wedge \tau_B \wedge \eta_{Y_k}} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} v^*(\lambda e^{-\alpha(T \wedge \tau_B \wedge \eta_{Y_k})}, \xi_{T \wedge \tau_B \wedge \eta_{Y_k}}) \right] \\
&= E_x^\pi \left[e^{\lambda \int_0^{T \wedge \tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} v^*(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right],
\end{aligned}
$$

which together with (4.29) implies

$$
v^*(\lambda, x) \leq E_x^\pi \left[e^{\lambda \int_0^{T \wedge \tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} v^*(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right] \quad (4.34)
$$

for all $x \in X$, $\lambda \in (0, \Lambda)$, $\pi \in \Pi$ and $T \geq 0$. Let p be an arbitrary constant satisfying $p > 1$, $\frac{\lambda M}{\alpha} < 1$ and $\frac{\lambda M \rho p}{\alpha^2} < 1$. Employing the Hölder inequality, we can derive

$$
\begin{aligned}
&E_x^\pi \left[e^{\lambda \int_0^{T \wedge \tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} v^*(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right] \\
&\leq \left(E_x^\pi \left[e^{\lambda p \int_0^{T \wedge \tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} \right] \right)^{\frac{1}{p}} \left(E_x^\pi \left[v^{* \frac{p}{p-1}}(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right] \right)^{\frac{p-1}{p}}
\end{aligned}
$$

$$(4.35)$$

for all $x \in X$, $\lambda \in (0, \Lambda)$, $\pi \in \Pi$ and $T \geq 0$. Moreover, by part (a) we obtain

$$
E_x^\pi \left[v^{* \frac{p}{p-1}}(\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right]
$$

$$
= E_x^\pi \left[v^{* \frac{p}{p-1}} (\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \left(I_{\{\tau_B \geq T\}} + I_{\{\tau_B < T\}} I_{\{\tau_B < \infty\}} \right) \right]
$$

$$
\leq E_x^\pi \left[v^{* \frac{p}{p-1}} (\lambda e^{-\alpha T}, \xi_T) \right] + E_x^\pi \left[v^{* \frac{p}{p-1}} (\lambda e^{-\alpha \tau_B}, \xi_{\tau_B}) I_{\{\tau_B < \infty\}} \right]
$$

$$
\leq \left(\frac{\alpha^2 e^{\frac{\lambda e^{-\alpha T}}{\alpha}}}{\alpha^2 - \rho \lambda e^{-\alpha T} M} \right)^{\frac{p}{p-1}} \left(1 + \frac{d}{\rho} \right)^{\frac{\lambda e^{-\alpha T} Mp}{\alpha(p-1)}} E_x^\pi \left[w^{\frac{\lambda e^{-\alpha T} Mp}{\alpha(p-1)}} (\xi_T) \right] + 1 \quad (4.36)
$$

for all $x \in X$, $\lambda \in (0, \Lambda)$, $\pi \in \Pi$ and $T \geq 0$. Observe that there exists a constant $\widehat{T} > 0$ satisfying $\frac{\lambda e^{-\alpha T} Mp}{\alpha(p-1)} < 1$ for all $T \geq \widehat{T}$. Thus, using Lemma 3.1(a), (4.36) and the Jensen inequality, we get

$$
E_x^\pi \left[v^{* \frac{p}{p-1}} (\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right]
$$

$$
\leq \left(\frac{\alpha^2 e^{\frac{\lambda e^{-\alpha T}}{\alpha}}}{\alpha^2 - \rho \lambda e^{-\alpha T} M} \right)^{\frac{p}{p-1}} \left(1 + \frac{d}{\rho} \right)^{\frac{\lambda e^{-\alpha T} Mp}{\alpha(p-1)}} (E_x^\pi [w(\xi_T)])^{\frac{\lambda e^{-\alpha T} Mp}{\alpha(p-1)}} + 1
$$

$$
\leq \left(\frac{\alpha^2 e^{\frac{\lambda e^{-\alpha T}}{\alpha}}}{\alpha^2 - \rho \lambda e^{-\alpha T} M} \right)^{\frac{p}{p-1}} \left(1 + \frac{d}{\rho} \right)^{\frac{2\lambda e^{-\alpha T} Mp}{\alpha(p-1)}} e^{\frac{\lambda e^{-\alpha T} M\rho Tp}{\alpha(p-1)}} w^{\frac{\lambda e^{-\alpha T} Mp}{\alpha(p-1)}} (x) + 1,
$$

which together with (4.35) yields

$$
E_x^\pi \left[e^{\lambda \int_0^{T \wedge \tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} v^* (\lambda e^{-\alpha(T \wedge \tau_B)}, \xi_{T \wedge \tau_B}) \right]
$$

$$
\leq \left(E_x^\pi \left[e^{\lambda p \int_0^{T \wedge \tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} \right] \right)^{\frac{1}{p}}
$$

$$
\times \left[\left(\frac{\alpha^2 e^{\frac{\lambda e^{-\alpha T}}{\alpha}}}{\alpha^2 - \rho \lambda e^{-\alpha T} M} \right)^{\frac{p}{p-1}} \left(1 + \frac{d}{\rho} \right)^{\frac{2\lambda e^{-\alpha T} Mp}{\alpha(p-1)}} e^{\frac{\lambda e^{-\alpha T} M\rho Tp}{\alpha(p-1)}} w^{\frac{\lambda e^{-\alpha T} Mp}{\alpha(p-1)}} (x) + 1 \right]^{\frac{p-1}{p}}
$$

$$
(4.37)
$$

for all $x \in X$, $\lambda \in (0, \Lambda)$, $\pi \in \Pi$ and $T \geq \widehat{T}$. Direct calculations give

$$
e^{\lambda p \int_0^{T \wedge \tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} \leq e^{\lambda p \int_0^\infty \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt}
$$

$$
\leq \int_0^\infty e^{\frac{\lambda p}{\alpha} \int_A c(\xi_t, a) \pi(da|\omega, t)} \alpha e^{-\alpha t} dt
$$

$$
\leq \alpha e^{\frac{\lambda bp}{\alpha}} \int_0^\infty e^{-\alpha t} w^{\frac{\lambda Mp}{\alpha}} (\xi_t) dt
$$

for all $T \geq 0$, where the second inequality is due to the Jensen inequality and the third one follows from Assumption 3.1. Moreover, by Lemma 3.1(a) and the Jensen

inequality, we have

$$
E_x^\pi \left[\int_0^\infty e^{-\alpha t} w^{\frac{\lambda M p}{\alpha}}(\xi_t) dt \right]
$$

$$
\leq \int_0^\infty e^{-\alpha t} (E_x^\pi [w(\xi_t)])^{\frac{\lambda M p}{\alpha}} dt \leq \left(1 + \frac{d}{\rho} \right)^{\frac{\lambda M p}{\alpha}} \int_0^\infty e^{-\alpha t} e^{\frac{\rho \lambda M p t}{\alpha}} w^{\frac{\lambda M p}{\alpha}}(x) dt
$$

$$
= \left(1 + \frac{d}{\rho} \right)^{\frac{\lambda M p}{\alpha}} \frac{\alpha}{\alpha^2 - \rho \lambda M p} w^{\frac{\lambda M p}{\alpha}}(x)
$$

for all $x \in X$ and $\pi \in \Pi$. Hence, letting $T \to \infty$ in (4.37) and using (4.34) and the dominated convergence theorem, we can derive

$$
v^*(\lambda, x) \leq 2^{\frac{p-1}{p}} \left(E_x^\pi \left[e^{\lambda p \int_0^{\tau_B} \int_A e^{-\alpha t} c(\xi_t, a) \pi(da|\omega, t) dt} \right] \right)^{\frac{1}{p}}
$$

for all $x \in X$, $\lambda \in (0, \Lambda)$ and $\pi \in \Pi$. Furthermore, letting $p \downarrow 1$ in the last inequality, by the dominated convergence theorem we obtain $v^*(\lambda, x) \leq \widehat{J}(\lambda, x, \pi)$ for all $x \in X$, $\lambda \in (0, \Lambda)$ and $\pi \in \Pi$. Therefore, we can get (4.27). Moreover, from (4.26) and (4.27), we have $v^*(\lambda, x) = \widehat{J}(\lambda, x, \pi^*) = \inf_{\pi \in \Pi} \widehat{J}(\lambda, x, \pi)$ for all $(\lambda, x) \in (0, \Lambda) \times X$. Hence, we see that $\pi^* \in \Pi_M$ is an optimal policy.

(c) The assertion follows from the same arguments as (4.26) and (4.27). $\qquad \square$

Remark 4.1 (a) Theorem 4.1 establishes the existence of a unique solution in $V_w((0, \Lambda) \times X)$ to the risk-sensitive first passage discounted cost optimality equation given by (4.16) and the existence of a deterministic Markov optimal policy for the Borel state space and unbounded cost and transition rates, which extends the existence results of the infinite-horizon risk-sensitive discounted cost criterion in [5] with the Borel state space and bounded cost and transition rates, [8] with the denumerable state space and bounded cost and transition rates, and [15] with the denumerable state space and unbounded cost and transition rates.

(b) In [5, 8, 15] the existence of a solution to the infinite-horizon risk-sensitive discounted cost optimality equation is derived by the fixed point technique. The fixed point approach needs the boundedness assumption on the cost and transition rates. Thus, the method in [5, 8] cannot be applied to the case of unbounded cost and transition rates. To deal with the unbounded case, [15] constructs an approximating sequence of bounded cost and transition rates and applies the fixed point technique to the approximating sequence. The diagonalization arguments in [15] require the assumption that the state space is a denumerable set. Thus, the method in [15] is inapplicable to the case of the Borel state space. We introduce a new value iteration given by (4.1) to deal with the Borel state space and unbounded transition rates directly without the boundedness assumption as in [5, 8] and the diagonalization arguments and the construction of an approximating sequence of bounded transition rates as in [15].

(c) We obtain the existence result in Theorem 4.1 under the conditions weaker than those in [5, 8, 15]. More precisely, we remove the boundedness condition on

the cost and transition rates in [5, 8] and the uniform convergence condition of the solution (i.e., $\lim_{\lambda \to 0} v^*(\lambda, x) = 1$ uniformly in $x \in X$) in [8]. We do not require the denumerability assumption on the state space, the uniformly conservative condition on the transition rates (i.e., for any $i \in X$, $\sum_{j \in X} q(j|i, a) = 0$ uniformly in $a \in A(i)$) and the additional condition that the constant $\overline{\rho} < \alpha$ in [15].

5 An Example

In this section, a cash flow model is given to illustrate the main results.

Example 5.1 (A cash flow model in [13]) The amount of the cash in the cash flow model is regarded as the state variable and all the possible states are given by $X = (-\infty, \infty)$. The action a denotes the withdrawal rate of money in cash (if $a < 0$) or the supply rate (if $a > 0$). When the amount of the cash is $x \in X$, the decision-maker takes an action from a given set $A(x) = [\zeta_1(x), \zeta_2(x)]$, where ζ_1 and ζ_2 are measurable functions on X and satisfy $\zeta_1(x) < 0$ and $\zeta_2(x) > 0$ for all $x \in X$. The amount of the cash x and the action a chosen by the decision-maker incur a nonnegative cost $c(x, a)$. Moreover, when the amount of the cash equals x and the action $a \in A(x)$ is taken by the decision-maker, after an exponentially distributed random time with the rate $\kappa(x, a) > 0$, the amount of the cash is changed to a new state following the normal distribution with the mean x and the variance β^2. So the transition rate can be given by

$$q(D|x, a) = \kappa(x, a) \left[\int_{D \setminus \{x\}} \frac{1}{\sqrt{2\pi} \beta} e^{-\frac{(y-x)^2}{2\beta^2}} \, dy - \delta_x(D) \right] \tag{5.1}$$

for all $(x, a) \in K$ and $D \in \mathcal{B}(X)$. Assume that the risk-sensitivity coefficient of the decision-maker is given by $\lambda > 0$. The decision-maker wishes to minimize the risk-sensitive discounted cost before the amount of the cash falls into the target set $B = (-\infty, 0)$.

We consider the following conditions to guarantee the existence of optimal policies for the cash flow model.

(E1) There exists a constant $\widehat{L} > 0$ such that $\kappa(x, a) \leq \widehat{L}(x^2 + 1)$ for all $(x, a) \in K$.
(E2) There exist constants $0 \leq \widehat{M} < \min\left\{\frac{\alpha^2}{\widehat{L}\beta^2\lambda}, \frac{\alpha}{2\lambda}\right\}$ and $\widehat{b} \geq 0$ such that $c(x, a) \leq \widehat{M} \ln(x^2 + 1) + \widehat{b}$ for all $(x, a) \in K$.
(E3) The function $\kappa(x, a)$ is measurable on K and continuous in $a \in A(x)$, and the function $c(x, a)$ is measurable on K and lower semi-continuous in $a \in A(x)$ for all $x \in X$.

Proposition 5.1 *Under conditions (E1)–(E3), Example 5.1 satisfies Assumptions 3.1 and 3.2. Therefore, by Theorem 4.1 there exists a deterministic Markov optimal policy for the cash flow model under the risk-sensitive first passage discounted cost criterion.*

Proof Take $w(x) = x^2 + 1$ for all $x \in X$. By (5.1) and condition (E1), we obtain

$$\int_X w(y)q(dy|x, a) = \beta^2 \kappa(x, a) \le \widehat{L}\beta^2 w(x) \text{ and } q^*(x) = \sup_{a \in A(x)} \kappa(x, a) \le \widehat{L}w(x) \tag{5.2}$$

for all $x \in X$ and $a \in A(x)$. Thus, it follows from (5.2) and condition (E2) that Assumption 3.1 is satisfied with $\rho = \widehat{L}\beta^2$, $d = 0$, $L = \widehat{L}$, $M = \widehat{M}$ and $b = \widehat{b}$. Moreover, from the description of the model and condition (E3), we see that Assumptions 3.2(i) and (ii) hold. Finally, we verify Assumption 3.2(iii). Employing (5.1) and condition (E1), we can derive

$$\int_X w^2(y)q(dy|x, a) = \kappa(x, a)(3\beta^4 + 6\beta^2 x^2 + 2\beta^2) \le \widehat{L}(3\beta^4 + 6\beta^2)w^2(x)$$

for all $(x, a) \in K$. Hence, Assumption 3.2(iii) holds with $\overline{\rho} = \widehat{L}(3\beta^4 + 6\beta^2)$ and $\overline{d} = 0$. □

Remark 5.1 The state space is a Borel space and the transition and cost rates are allowed to be unbounded in Example 5.1. Hence, the conditions in [5, 8, 15] fail to hold because the cost and transition rates are bounded in [5, 8] and the state space is a denumerable set in [15].

6 Conclusions

In this paper, we have investigated continuous-time Markov decision processes under the risk-sensitive first passage discounted cost criterion. The state and action spaces are Borel spaces, and the cost and transition rates can be unbounded. We have introduced a new value iteration to derive the existence of a solution to the risk-sensitive first passage discounted cost optimality equation under the suitable conditions. Moreover, employing the Feynman–Kac formula, we have proved that the risk-sensitive first passage discounted cost optimal value function is a unique solution to the risk-sensitive first passage discounted cost optimality equation. In addition, we have obtained the existence of a deterministic Markov optimal policy in the class of randomized history-dependent policies.

Declaration

Conflict of interest We declare that no conflict of interest exists in this paper.

References

1. Aliprantis, C., Border, K.: Infinite Dimensional Analysis. Springer, New York (2006)
2. Bäuerle, N., Rieder, U.: More risk-sensitive Markov decision processes. Math. Oper. Res. **39**, 105–120 (2014)
3. Bogachev, V.I.: Measure Theory, vol. I. Springer, Berlin (2007)
4. Cavazos-Cadena, R.: Characterization of the optimal risk-sensitive average cost in denumerable Markov decision chains. Math. Oper. Res. **43**, 1025–1050 (2018)
5. Chandan, P., Somnath, P.: Risk sensitive control of pure jump processes on a general state space. Stochastics **91**, 155–174 (2019)
6. Di Masi, G.B., Stettner, Ł: Risk-sensitive control of discrete-time Markov processes with infinite horizon. SIAM J. Control Optim. **38**, 61–78 (1999)
7. Di Masi, G.B., Stettner, Ł: Infinite horizon risk sensitive control of discrete time Markov processes under minorization property. SIAM J. Control Optim. **46**, 231–252 (2007)
8. Ghosh, M., Saha, S.: Risk-sensitive control of continuous time Markov chains. Stochastics **86**, 655–675 (2014)
9. Guo, X., Huang, Y.H.: Risk-sensitive average continuous-time Markov decision processes with unbounded transition and cost rates. J. Appl. Probab. **58**, 523–550 (2021)
10. Guo, X., Liu, Q.L., Zhang, Y.: Finite horizon risk-sensitive continuous-time Markov decision processes with unbounded transition and cost rates. 4OR **17**, 427–442 (2019)
11. Guo, X.P., Hernández-Lerma, O.: Continuous-Time Markov Decision Processes: Theory and Applications. Springer, Berlin (2009)
12. Guo, X.P., Huang, X.X., Zhang, Y.: On the first passage g-mean-variance optimality for discounted continuous-time Markov decision processes. SIAM J. Control Optim. **53**, 1406–1424 (2015)
13. Guo, X.P., Huang, Y.H., Song, X.Y.: Linear programming and constrained average optimality for general continuous-time Markov decision processes in history-dependent policies. SIAM J. Control Optim. **50**, 23–47 (2012)
14. Guo, X.P., Song, X.Y.: Discounted continuous-time constrained Markov decision processes in Polish spaces. Ann. Appl. Probab. **21**, 2016–2049 (2011)
15. Guo, X.P., Liao, Z.W.: Risk-sensitive discounted continuous-time Markov decision processes with unbounded rates. SIAM J. Control Optim. **57**, 3857–3883 (2019)
16. Guo, X.P., Zhang, J.Y.: Risk-sensitive continuous-time Markov decision processes with unbounded rates and Borel spaces. Discrete Event Dyn. Syst. **29**, 445–471 (2019)
17. Hernández-Hernández, D., Marcus, S.I., Fard, P.J.: Analysis of a risk-sensitive control problem for hidden Markov chains. IEEE Trans. Autom. Control **44**, 1093–1100 (1999)
18. Hernández-Lerma, O., Lasserre, J.B.: Further Topics on Discrete-time Markov Control Processes. Springer, New York (1999)
19. Huang, X.X., Liu, Q.L., Guo, X.P.: N-person nonzero-sum games for continuous-time jump processes with varying discount factors. IEEE Trans. Autom. Control **64**, 2037–2044 (2019)
20. Kitaev, M.Y., Rykov, V.V.: Controlled Queueing Systems. CRC Press, Boca Ration (1995)
21. Piunovskiy, A., Zhang, Y.: Continuous-Time Markov Decision Processes: Borel Space Models and General Control Strategies. Springer, Cham (2020)
22. Shen, Y., Stannat, W., Obermayer, K.: Risk-sensitive Markov control processes. SIAM J. Control Optim. **51**, 3652–3672 (2013)
23. Stein, E.M., Shakarchi, R.: Real Analysis: Measure Theory, Integration, and Hilbert Spaces. Princeton University Press, Princeton (2005)
24. Suresh Kumar, K., Pal, C.: Risk-sensitive ergodic control of continuous time Markov processes with denumerable state space. Stoch. Anal. Appl. **33**, 863–881 (2015)
25. Wei, Q.D.: Continuous-time Markov decision processes with risk-sensitive finite-horizon cost criterion. Math. Methods Oper. Res. **84**, 461–487 (2016)
26. Wei, Q.D., Chen, X.: Continuous-time Markov decision processes under the risk-sensitive average cost criterion. Oper. Res. Lett. **44**, 457–462 (2016)
27. Wei, Q.D., Chen, X.: Risk-sensitive average continuous-time Markov decision processes with unbounded rates. Optimization **68**, 773–800 (2019)
28. Zhang, Y.: Continuous-time Markov decision processes with exponential utility. SIAM J. Control Optim. **55**, 2636–2660 (2017)