The existence of optimal feedback controls for stochastic dynamical systems with regime-switching*

Jinghai Shao†

June 21, 2019

Abstract

In this work we provide explicit conditions on the existence of optimal feedback controls for stochastic processes with regime-switching. We use the compactification method which needs less regularity conditions on the coefficients of the studied stochastic systems. Moreover, the dynamic programming principle is established after showing the continuity of the value function. We have considered the random impact of the environment on the studied stochastic systems by including a regime-switching process on a discrete state space.

AMS subject Classification (2010): 93B52, 60J60, 49K30, 60J27

Keywords: Feedback control, Regime-switching diffusions, Hybrid system, Wasserstein distance

1 Introduction

This work focuses on providing sufficient conditions for the existence of optimal feedback controls for the stochastic control systems with regime-switching. This system contains

*Supported in part by NNSFs of China (No. 11771327, 11431014, 11831014)
†Center for Applied Mathematics, Tianjin University, Tianjin 300072, China. (shaojh@tju.edu.cn.)
two components \((X_t, \Lambda_t)\): the continuous component \((X_t)\) satisfies a stochastic differential equation (SDE) which describes the evolution of the studied dynamical system; the discrete component \((\Lambda_t)\) is a jumping process on a finite state space which describes the random change of the environment in which \((X_t)\) lives. The control policy also owns two terms: one is to control the coefficients of SDEs; another is to control the transition rate matrices of \((\Lambda_t)\). This kind of controls is of great meaning in applications and has not been investigated before. All admissible control policies considered in this paper are in the form of feedback control. We develop the compactification method to provide explicit conditions to guarantee the existence of optimal feedback controls with respect to finite-horizon cost functions. The value function is shown to be continuous and the dynamic programming principle is established.

The existence of optimal feedback controls is a fundamental issue in the study of control theory. This issue is not only theoretical, since it is needed to ensure that the optimization problem is well defined and to allow subsequent analysis of the equations for the value function.

One approach to establish the existence of optimal controls is based on the theory of partial differential equations of dynamic programming; see the early works of Davis \cite{8} and Bismut \cite{4}, Fleming and Rishel \cite{13} or the recent survey Kushner \cite{29} and the references therein. This method has been extensively studied in connection with the theory of Hamilton-Jacobi-Bellman equations, which encounters the restriction of the regularity of corresponding solutions. Another approach is to show directly the compactness of the minimizing sequence of controls. Kushner \cite{27} used the weak convergence of measures to provide a general result on the existence of optimal controls. Also, Haussman and Lepeltier \cite{20}, Haussman and Suo \cite{21, 22} have developed this method to show the existence of optimal controls and even optimal relaxed controls. The advantage of this compactification method is that it requires less regularity of the value function and thus needs only very mild hypothesis on the data. Especially, the works \cite{27, 20, 21, 22} investigated the stochastic open loop problem. Moreover, given the existence of an optimal control, \cite{20} used Krylov’s Markov selection theorem showed the optimal control could be represented as a Markov control. It is a far more trivial task to guarantee the limit of the minimizing sequence being adapted to the stochastic fields generated by the dynamic system which also strongly depends on the limit of the control sequence. In view of this difficulty, the known sufficient conditions on the existence of optimal feedback controls for stochastic control models are mostly provided by the theory of partial differential equations; see, for example, Fleming and Rishel \cite[Chapter VI]{13} and references therein. Besides, Linquist
transformed the feedback control problem into the stochastic open loop problem for a class of linear systems by adding a further restriction on his feedback class. In this work, we shall develop the compactification method to provide sufficient conditions of optimal feedback controls.

The stochastic maximum principle plays a central role in stochastic control theory. It gives necessary conditions for optimal controls. Its first version was established by Kushner [28] where the diffusion coefficients are independent of the controls, and by Peng [34] when the diffusion coefficients depend on the controls. Some advance information about the form of the optimal control is needed to use stochastic maximum principles to find optimal controls in applications. For example, Lü, Wang, and Zhang [31] established the equivalence between the existence of optimal feedback controls for the stochastic linear quadratic control problems and the solvability of the corresponding backward stochastic Riccati equations in some sense.

In this work we are also concerned with the random impact of the environment to the dynamic systems. Variations in the external environment (for example, weather or temperature) can have important effects on the dynamics of the studied systems. For instance, for the ecosystem, certain biological parameters such as the growth rates and the carrying capacities often demonstrate abrupt changes due to the environmental noise. Therefore, it is natural to consider the random changes of the environment in mathematical modeling. Recently, such models are widely applied in stochastic control and optimization, mathematical finance, ecological and biological systems, engineer, etc.; see, for example, [2, 9, 23, 32, 51] amongst others. In view of its wide application, this optimal control problems for regime-switching processes have been studied in the literature; see, for instance, [43, 44, 47, and 50 amongst others. In particular, [43] and [44] investigated the singular control problem for regime-switching processes with Markovian regime-switching. In [43], Song et al. showed that the value function is a viscosity solution of a system of quasi-variational inequalities (QVIs) through proving first the continuity of the value function by exploiting the advantage of a one-dimensional regime-switching diffusion process. For the Markovian regime-switching processes in high dimensional space, by establishing directly a weakly dynamic programming principle instead of proving the continuity of the value function, Song and Zhu in [44] showed directly that the value function is a viscosity solution to a system of QVIs. However, there is no discussion on the optimal control problem for state-dependent regime-switching processes. The interaction between the state process and the switching process makes the optimal control problem more complicated.
The regime-switching diffusion processes \((X_t, \Lambda_t)_{t \geq 0}\) contains two components: the first component \((X_t)_{t \geq 0}\) satisfies the following SDE:

\[
dX_t = b(X_t, \Lambda_t)dt + \sigma(X_t, \Lambda_t)dB_t, \tag{1.1}\]

where \(b : \mathbb{R}^d \times S \to \mathbb{R}^d\), \(\sigma : \mathbb{R}^d \times S \to \mathbb{R}^{d \times d}\), and \((B_t)_{t \geq 0}\) is a standard \(d\)-dimensional Brownian motion; the second component \((\Lambda_t)_{t \geq 0}\) is continuous-time jumping process satisfying

\[
P(\Lambda_{t+\delta} = j | \Lambda_t = i, X_t = x) = \begin{cases} q_{ij}(x)\delta + o(\delta), & \text{if } j \neq i, \\ 1 + q_{ii}(x)\delta + o(\delta), & \text{otherwise}, \end{cases} \tag{1.2}\]

provided \(\delta > 0\). The component \((X_t)_{t \geq 0}\) is used to describe the evolution of a dynamical system, and the component \((\Lambda_t)_{t \geq 0}\) is used to reflect the random switching of the environment where the studied system lives. When the transition rate matrix \((q_{ij}(x))\) depends on \(x\), \((X_t, \Lambda_t)_{t \geq 0}\) is called a state-dependent regime-switching process. When \((q_{ij}(x))\) does not depend on \(x\), then \((\Lambda_t)_{t \geq 0}\) is indeed a continuous-time Markov chain, and is assumed to be independent of the Brownian motion \((B_t)_{t \geq 0}\) as usual. In this case, \((X_t, \Lambda_t)_{t \geq 0}\) is a state-independent regime-switching process, and sometimes called Markovian regime-switching process. Here \(S\) is a denumerable space, and \(\mathcal{B}(S)\) denotes the collection of all measurable sets. When \(S\) is a finite set, various properties of regime-switching processes such as stability, ergodicity, numerical approximation, etc. have been widely studied in the literature; see, e.g. \([32, 33, 51, 39, 49, 51]\) and references therein. When \(S\) is an infinitely countable set, we refer to \([40, 36, 37]\), where two kinds of methods, finite partition method and principle eigenvalue method, were raised to deal with the stability and ergodicity of regime-switching processes.

In this work we use the compactification method to show the existence of the optimal feedback control with respect to a very general finite-horizon cost function. Here our method looks similar to Haussmann and Suo \([21]\), but the technics are quite different. This can be easily seen from the fact that \([21]\) cannot deal with the case that the cost function depends on the terminal value the process, but we can. Moreover, the dynamic programming principle is established based on the continuity of the value function and an application of the selection theorem. To prove the continuity of the value function for the state-dependent regime-switching processes, some elaborate estimates have been established based on Skorokhod’s representation for jumping processes through Poisson random measures. These estimates reflect the essential difference between processes with and without regime-switching.
This paper is organized as follows. In Section 2, we introduce the class of admissible feedback controls and prove the existence of the optimal feedback control by using compactification method. In Section 3, we first prove the continuity of the value function under certain appropriate conditions, then establish the dynamic programming principle.

2 Existence of optimal controls

2.1 Framework and statement of the result

Let $\mathcal{S} = \{1, 2, \ldots, N\}$ with $N < \infty$. $T$ is a positive constant given throughout this work. $U$ is a compact set of, say, $\mathbb{R}^k$ for some $k \in \mathbb{N}$, and $\mathcal{P}(U)$ denotes the collection of all probability measures over $U$. For any two probability measures $\mu$ and $\nu$ in $\mathcal{P}(U)$, their $L^1$-Wasserstein distance is defined as:

$$W_1(\mu, \nu) = \inf_{\Gamma \in \mathcal{C}(\mu, \nu)} \left\{ \int_{U \times U} |x - y| \Gamma(dx, dy) \right\},$$

where $\mathcal{C}(\mu, \nu)$ stands for the set of all couplings of $\mu$ and $\nu$ on $U \times U$.

Let $E$ be a metric space. For $0 \leq a < b \leq T$,

- $\mathcal{C}([a, b]; E)$ is the collection of continuous functions $x : [a, b] \to E$;
- $\mathcal{D}([a, b]; E)$ is the collection of right-continuous functions with left limits $x : [a, b] \to E$.

Denote by $x_{[s,t]}$ the function $x$, in $\mathcal{C}([s, t]; E)$ or $\mathcal{D}([s, t]; E)$ with $s, t \in [0, T]$, and it can be extended to the whole interval $[0, T]$ through the map $\Xi$:

$$ (\Xi x_{[s,t]})_r = \begin{cases} x_s, & \text{if } r \leq s, \\ x_r, & \text{if } s < r < t, \\ x_t, & \text{if } r \geq t. \end{cases} \quad (2.1) $$

Let $\psi : [0, \infty) \to [0, \infty)$ be a measurable function such that

$$ \kappa_1 r \leq \psi(r) \leq \kappa_2 r, \quad \forall r \in [0, T], \quad (2.2) $$

holds for some positive constants $\kappa_1, \kappa_2$. 

A functional $F : [0, T] \times \mathcal{C}([0, T]; \mathbb{R}^d) \times \mathcal{S} \to \mathcal{P}(U)$ is said to be in the class $\Upsilon_\psi$ if for every $t_1, t_2 \in [0, T], x, y \in \mathcal{C}([0, T]; \mathbb{R}^d)$ and $i, j \in \mathcal{S}$, it holds
\begin{equation}
W_1(F(t_1, x, i), F(t_2, y, j)) \leq \psi(|t_1 - t_2| + \|x - y\|_\infty + 1_{i \neq j}),
\end{equation}
where $\|x - y\|_\infty = \sup_{t \in [0, T]} |x_t - y_t|$ is the uniform norm in $\mathcal{C}([0, T]; \mathbb{R}^d)$.

Give a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ endowed with a complete filtration $\{\mathcal{F}_t\}_{t \geq 0}$. Consider the following stochastic dynamical system
\begin{equation}
dX_t = b(X_t, \Lambda_t, \mu_t)dt + \sigma(X_t, \Lambda_t, \mu_t)dB_t,
\end{equation}
where $b : \mathbb{R}^d \times \mathcal{S} \times \mathcal{P}(U) \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \times \mathcal{S} \times \mathcal{P}(U) \to \mathbb{R}^{d \times d}$, and $(B_t)_{t \geq 0}$ is a $d$-dimensional $\mathcal{F}_t$-Brownian motion. Here $(\Lambda_t)_{t \geq 0}$ is a continuous-time jumping process on $\mathcal{S}$ satisfying
\begin{equation}
\mathbb{P}(\Lambda_{t+\delta} = j | \Lambda_t = i, X_t = x, \nu_t = \nu) = \begin{cases}
q_{ij}(x, \nu)\delta + o(\delta), & \text{if } j \neq i, \\
1 + q_{ii}(x, \nu)\delta + o(\delta), & \text{otherwise},
\end{cases}
\end{equation}
provided $\delta > 0$ for every $x \in \mathbb{R}^d$, $\nu \in \mathcal{P}(U)$, $i, j \in \mathcal{S}$. In this controlled system (2.4) and (2.5), we consider two control terms: $\mu_\cdot$ and $\nu_\cdot$, which are both measurable maps from $[0, T]$ to $\mathcal{P}(U)$. The term $\mu_\cdot$ is a kind of classical relaxed control for stochastic dynamical system which has been studied in many works. The term $\nu_\cdot$ is a special control policy for regime-switching processes, which is used to control the transition rate matrices of the jumping process $(\Lambda_t)$. As $(\Lambda_t)$ is a jumping process in the discrete state space, the role played by the control $\nu_t$ is quite different to that played by the term $\mu_t$ in the evolution of the studied dynamic system. This kind of control $\nu_t$ has not been studied in the optimal control problem for regime-switching processes before. In addition, this control is closely related to the control policy used in the study of continuous-time Markov decision process. See [41] for more discussion on their relationship.

The feedback controls studied in this work are introduced as follows.

**Definition 2.1** For each $(s, x, i) \in [0, T) \times \mathbb{R}^d \times \mathcal{S}$, an admissible control is a term $\alpha = (\mu_t, \nu_t)_{t \in [s, T]}$ satisfying
\begin{enumerate}
\item There exist two functionals $F$ and $G$ in $\Upsilon_\psi$ such that
\[
\mu_t = F(t, \Xi X_{[s, t]}, \Lambda_t), \quad \nu_t = G(t, \Xi X_{[s, t]}, \Lambda_t), \quad t \in (s, T];
\]
\item $(X_t, \Lambda_t)_{t \in [s, T]}$ satisfies (2.4) and (2.5) with $X_s = x$, $\Lambda_s = i$.
\end{enumerate}
Denote by $\Pi_{s,x,i}$ the collection of all admissible controls with initial value $(X_s, \Lambda_s) = (x, i)$ for $(s, x, i) \in [0, T) \times \mathbb{R}^d \times S$. The class $\Pi_{s,x,i}$ contains many interesting controls, especially, it contains the path dependent feedback controls on the component $(X_t)$. The restriction that $F, G \in \Upsilon_\psi$ mainly aims to ensure the existence of strong solution of the controlled system (2.4) and (2.5). This restriction can be weaken to the request of the existence of strong solution. See Subsection 2.2 for an extension of $\Pi_{s,x,i}$ to a more general set of feedback control policies $\tilde{\Pi}_{s,x,i}$. However, we prefer to using the set $\Pi_{s,x,i}$ here because of its clarity of the structure of the control policies.

Haussmann and Suo [21] assumed the existence of martingale solution of the corresponding stochastic dynamical system and proved the existence of optimal control which is not necessary a feedback control policy. In contrast to [21], some explicit conditions on the coefficients of the studied system (2.4) and (2.5) will be presented below to ensure the existence of strong solution of the studied system. By taking advantage of this property, we can show the existence of the optimal feedback controls.

Given two lower semi-continuous functions $f : [0, T] \times \mathbb{R}^d \times S \times \mathcal{P}(U) \times \mathcal{P}(U) \to [0, \infty)$ and $g : \mathbb{R}^d \to [0, \infty)$, the expected cost relative to the control $\alpha$ is defined by

$$J(s, x, i, \alpha) = \mathbb{E} \left[ \int_s^T f(t, X_t, \Lambda_t, \mu_t, \nu_t) dt + g(X_T) \right].$$

(2.6)

The corresponding value function is defined by

$$V(s, x, i) = \inf_{\alpha \in \Pi_{s,x,i}} J(s, x, i, \alpha).$$

(2.7)

An admissible control $\alpha^* \in \Pi_{s,x,i}$ is called optimal, if it holds

$$V(s, x, i) = J(s, x, i, \alpha^*).$$

(2.8)

The hypotheses on the coefficients of $(X_t, \Lambda_t)$ are listed as follows in order to ensure the existence of strong solution $(X_t, \Lambda_t)$ satisfying (2.4) and (2.5).

(H1) There exists a constant $C_1 > 0$ such that

$$|b(x, i, \mu) - b(y, i, \nu)|^2 + \|\sigma(x, i, \mu) - \sigma(y, i, \nu)\|^2 \leq C_1 (|x - y|^2 + W_1(\mu, \nu)^2)$$

for $x, y \in \mathbb{R}^d, i \in S, \mu, \nu \in \mathcal{P}(U)$, where $|x|^2 = \sum_{k=1}^d x_k^2$, $\|\sigma\|^2 = \text{tr}(\sigma\sigma')$, and $\sigma'$ denotes the transpose of the matrix $\sigma$. 
(H2) For every $x \in \mathbb{R}^d$, $\nu \in \mathcal{P}(U)$, $(q_{ij}(x, \nu))$ is conservative, i.e. $q_i(x, \nu) = \sum_{j \neq i} q_{ij}(x, \nu)$ for every $i \in S$. Moreover, $M := \sup_{x \in \mathbb{R}^d, \nu \in \mathcal{P}(U)} \max_{i \in S} q_i(x, \nu) < \infty$.

(H3) There exists a constant $C_2 > 0$ such that for every $i, j \in S$, $x, y \in \mathbb{R}^d$, $\mu, \nu \in \mathcal{P}(U)$,
$$|q_{ij}(x, \mu) - q_{ij}(y, \nu)| \leq C_2(|x - y| + W_1(\mu, \nu)).$$

(H4) $U \subset \mathbb{R}^k$ is compact for some $k \in \mathbb{N}$.

Our first main result of this work is as follows.

**Theorem 2.2** Assume conditions (H1)-(H4) hold. Then for every $(s, x, i) \in [0, T] \times \mathbb{R}^d \times S$, there exists an optimal admissible control $\alpha^* \in \Pi_{s, x, i}$ corresponding to the value function $V(s, x, i)$.

Note that the assumptions (H1)-(H3) ensure that SDEs (2.4) and (2.5) admit a unique strong solution, which is proved in Appendix (Proposition 3.6 below). The Lipschitz conditions can be replaced by some non-Lipschitz conditions to ensure the existence of strong solutions for such kind of system. See, for instance, [38] for the existence of strong solutions of state-dependent regime-switching processes from the viewpoint of SDEs, and [35] for the existence of strong solution of stochastic functional differential equations under non-Lipschitz conditions.

### 2.2 Proof of Theorem 2.2

Before proving Theorem 2.2, we make some necessary preparations. Let $\mathcal{P}(U)$ be endowed with the $L^1$-Wasserstein distance. $\mathcal{C}([0, T]; \mathbb{R}^d)$ is endowed with the uniform topology, and $\mathcal{D}([0, T]; \mathcal{P}(U))$, $\mathcal{D}([0, T]; \mathcal{S})$ are endowed with pseudopath topology which makes $\mathcal{D}([0, T]; \mathcal{P}(U))$ and $\mathcal{D}([0, T]; \mathcal{S})$ to be Polish spaces. Let
$$\mathcal{Y} = \mathcal{C}([0, T]; \mathbb{R}^d) \times \mathcal{D}([0, T]; \mathcal{S}) \times \mathcal{D}([0, T]; \mathcal{P}(U)) \times \mathcal{D}([0, T]; \mathcal{P}(U)),$$
and $\mathcal{Y}$ the Borel $\sigma$-field, $\mathcal{Y}_t$ the $\sigma$-fields up to time $t$. Then, as a product space endowed with the product topology, $\mathcal{Y}$ is a Polish space.

In the argument of Theorem 2.2, we shall consider the tightness of the distributions of admissible controls by transforming them into the canonical space $\mathcal{Y}$ via a measurable map $\Psi$. For an admissible control $\alpha = (\mu, \nu)$ in $\Pi_{s, x, i}$, $\Psi : \Omega \rightarrow \mathcal{Y}$ is defined by
$$\Psi(\omega) = (X_t(\omega), \Lambda_t(\omega), \mu_t(\omega), \nu_t(\omega))_{t \in [0, T]}.$$
Here, $X_r(\omega) := x$, $\Lambda_r(\omega) := i$, $\mu_r(\omega) := \mu_s$, and $\nu_r(\omega) := \nu_s$ for $r \in [0, s]$. Let $R = \mathbb{P} \circ \Psi^{-1}$ be the corresponding probability measure on $\mathcal{Y}$ associated with the control $\alpha = (\mu, \nu)$.

As a preparation, we introduce Skorokhod’s representation of $(\Lambda_t)$ in terms of the Poisson random measure as in [42, Chapter II-2.1] or [51], which plays an important role in this work, especially when we study the continuity of the value function in the next section.

For each $x \in \mathbb{R}^n$ and $\nu \in \mathcal{P}(U)$, we construct a family of intervals $\{\Gamma_{ij}(x, \nu); i, j \in S\}$ on the half line in the following manner:

$\Gamma_{12}(x, \nu) = [0, q_{12}(x, \nu))$
$\Gamma_{13}(x, \nu) = [q_{12}(x, \nu), q_{12}(x, \nu) + q_{13}(x, \nu))$

$\cdots \cdots$
$\Gamma_{1N}(x, \nu) = [\sum_{j=1}^{N-1} q_{1j}(x, \nu), q_{1}(x, \nu))$
$\Gamma_{21}(x, \nu) = [q_{1}(x, \nu), q_{1}(x, \nu) + q_{21}(x, \nu))$
$\Gamma_{23}(x, \nu) = [q_{1}(x, \nu) + q_{21}(x, \nu), q_{1}(x, \nu) + q_{21}(x, \nu) + q_{23}(x, \nu))$

and so on. Therefore, we obtain a sequence of consecutive, left-closed, right-open intervals $\Gamma_{ij}(x, \nu)$, each having length $q_{ij}(x, \nu)$. For convenience of notation, we set $\Gamma_{ii}(x, \nu) = \emptyset$ and $\Gamma_{ij}(x, \nu) = \emptyset$ if $q_{ij}(x, \nu) = 0$. Define a function $\vartheta : \mathbb{R}^n \times S \times \mathcal{P}(U) \times \mathbb{R} \to \mathbb{R}$ by

$\vartheta(x, i, \nu, z) = \sum_{l \in S} (l - i) 1_{\Gamma_{il}(x, \nu)}(z)$.

Then the process $(\Lambda_t)$ can be expressed by the following SDE

$$d\Lambda_t = \int_{[0,H]} \vartheta(X_t, \Lambda_{t}, \nu_{t}, z) N_1(dt, dz), \quad (2.9)$$

where $H = N(N-1)M$, $N_1(dt, dz)$ is a Poisson random measure with intensity $dt \times \mathbf{m}(dz)$, and $\mathbf{m}(dz)$ is the Lebesgue measure on $[0, H]$. Here we also assume that the Poisson random measure $N_1$ and the Brownian motion $(B_t)$ are mutually independent. Let $p_1(t)$ be the stationary point process corresponding to the Poisson random measure $N_1(dt, dz)$. Due to the finiteness of $\mathbf{m}(dz)$ on $[0, H]$, there is only finite number of jumps of the process $p_1(t)$ in each finite time interval. Let $0 = \varsigma_0 < \varsigma_1 < \ldots < \varsigma_n < \ldots$ be the enumeration of all jumps of $p_1(t)$. It holds that $\lim_{n \to \infty} \varsigma_n = +\infty$ almost surely. Due to (2.9), it follows that, if $\Lambda_0 = i$,

$$\Lambda_{\varsigma_1} = i + \sum_{l \in S} (l - i) 1_{\Gamma_{il}(X_{\varsigma_1}, \nu_{\varsigma_1})}(p_1(\varsigma_1)). \quad (2.10)$$
This yields that $\Lambda_t$ has a jump at $s_1$ (i.e., $\Lambda_{s_1} \neq \Lambda_{s_1-}$) if $p_1(s_1)$ belongs to the interval $\Gamma_l(X_{s_1}, \nu_{s_1})$ for some $l \neq i$. At any other cases, $\Lambda_t$ admits no jump at $s_1$. So the set of jumping times of $\Lambda_t$ is a subset of \{s_1, s_2, \ldots\}. This fact will be used below without mentioning it again.

**Proof of Theorem 2.2**

If $V(s, x, i) = \infty$, then according to the definition of $V$, any admissible control $\alpha$ will be optimal. Hence, we only need to consider the case $V(s, x, i) < \infty$. To simplify the notation, we consider only $s = 0$, and more general cases for $s \in (0, T]$ can be proved in the same way with suitable modification. The proof is separated into three steps.

**Step 1.** There exists a sequence of admissible controls $\alpha_n = (\mu^{(n)}_t, \nu^{(n)}_t)$ in $\Pi_{0, x, i}$ such that

$$\lim_{n \to \infty} J(0, x, i, \alpha_n) = V(0, x, i) < \infty. \quad (2.11)$$

Denote by $(X_t^{(n)}, \Lambda_t^{(n)})$ and $(F^{(n)}, G^{(n)})$ respectively the controlled system and functionals in $\Upsilon_\psi$ associated with $\alpha_n$. So,

$$\mu^{(n)}_t = F^{(n)}(t; \Xi X_{[0, t]}, \Lambda_t^{(n)}), \quad \nu^{(n)}_t = G^{(n)}(t, \Xi X_{[0, t]}, \Lambda_t^{(n)}), \quad t \in (0, T]. \quad (2.12)$$

Let $R_n, n \geq 1$, be the joint distribution of $(X_t^{(n)}, \Lambda_t^{(n)}, \mu^{(n)}_t, \nu^{(n)}_t)_{t \in [0, T]}$, which is a sequence of probability measures in the canonical space $\mathcal{Y}$. In this step we aim to prove the tightness of $(R_n)_{n \geq 1}$. Denote respectively by $\mathcal{L}^n_X, \mathcal{L}^n_\Lambda, \mathcal{L}^n_\mu,$ and $\mathcal{L}^n_\nu$ the marginal distribution of $R_n$ for $n \geq 1$.

We first prove that $(\mathcal{L}^n_\Lambda)_{n \geq 1}$ is tight by using Kurtz’s tightness criterion (cf. [10, Theorem 8.6, p.137]). As $\mathcal{S}$ is a finite set, we only need to show there exists a sequence of nonnegative random variable $\gamma_n(\delta)$ such that

$$\mathbb{E}\left[\mathbf{1}_{\Lambda^{(n)}_{t+u} \neq \Lambda^{(n)}_t} | \mathcal{F}_t\right] \leq \mathbb{E}\left[\gamma_n(\delta) | \mathcal{F}_t\right], \quad 0 \leq t \leq T, \quad 0 \leq u \leq \delta, \quad (2.13)$$

and $\lim_{\delta \downarrow 0} \sup_n \mathbb{E}[\gamma_n(\delta)] = 0$. Due to (H2), the boundedness of $(q_{ij}(x, \nu))$ implies

$$\mathbb{P}(\Lambda^{(n)}_t = \Lambda^{(n)}_t, \forall r \in [t, t + u]) \geq \mathbb{E}\left[\exp \left(- \sup_{x \in \mathbb{R}^d, \nu \in \mathcal{P}(U)} \max_{j \in \mathcal{S}} q_{ij}(x, \nu) u\right)\right]$$

$$\geq \exp(-M u).$$

Then, for every $0 \leq u \leq \delta$,

$$\mathbb{E}\left[\mathbf{1}_{\Lambda^{(n)}_{t+u} \neq \Lambda^{(n)}_t} | \mathcal{F}_t\right] \leq 1 - \mathbb{P}(\Lambda^{(n)}_t = \Lambda^{(n)}_t, \forall r \in [t, t + u])$$

$$\leq 1 - e^{-Mu} =: \gamma_n(\delta). \quad (2.14)$$
It is clear that \( \lim_{\delta \downarrow 0} \sup_n \mathbb{E}[\gamma_n(\delta)] = 0 \) and (2.13) is verified. We conclude that \((L_n^\alpha)_{n \geq 1}\) is tight.

For any \(0 \leq t_1 < t_2 \leq T\), by the definition of admissible control, we have
\[
W_1(\mu_{t_1}, \mu_{t_2}) = W_1(F^{(n)}(t_1, \Xi X^{(n)}_{[0,t_1]}, \Lambda^{(n)}_{t_1}), F^{(n)}(t_2, \Xi X^{(n)}_{[t_1,t_2]}, \Lambda^{(n)}_{t_2}))
\leq \psi(|t_2 - t_1| + \sup_{t_1 \leq t \leq t_2} |X^{(n)}_t - X^{(n)}_{t_1}| + \mathbb{1}_{\Lambda^{(n)}_{t_1} \neq \Lambda^{(n)}_{t_2}})
\leq \kappa(2(|t_2 - t_1| + \sup_{t_1 \leq t \leq t_2} |X^{(n)}_t - X^{(n)}_{t_1}| + \mathbb{1}_{\Lambda^{(n)}_{t_1} \neq \Lambda^{(n)}_{t_2}})).
\]

Therefore,
\[
\mathbb{E} W_1(\mu_{t_1}, \mu_{t_2}) \leq \kappa(2(|t_2 - t_1| + \mathbb{E} \sup_{t_1 \leq t \leq t_2} |X^{(n)}_t - X^{(n)}_{t_1}| + \mathbb{P}(\Lambda^{(n)}_{t_1} \neq \Lambda^{(n)}_{t_2})).
\]

(2.15)

Since \(U\) is compact, \((\mathcal{P}(U), W_1)\) is a compact Polish space (cf. e.g. [1]). This implies that the diameter of \(\mathcal{P}(U)\) is finite. Namely, there exists a constant \(K > 0\) such that
\[
\text{Diam}(\mathcal{P}(U)) := \sup_{\mu, \nu \in \mathcal{P}(U)} W_1(\mu, \nu) \leq K.
\]

Hence, the global Lipschitz condition (H1) implies the linear growth condition, i.e. there exists a \(C > 0\) such that \(|b(x, i, \mu)| + \left\|\sigma(x, i, \mu)\right\| \leq C(1 + |x|)\) for every \(i \in S, \mu \in \mathcal{P}(U)\), which leads to
\[
\mathbb{E}\left[\sup_{0 \leq t \leq T} |X^{(n)}_t|^2\right] \leq C(T, x), \quad n \geq 1,
\]
where \(C(T, x)\) is a constant depending on \(T\) and the initial value \(x\) (cf. for example, [33, Lemma 3.1, p.28]). By Burkholder-Davis-Gundy’s inequality, it holds
\[
\mathbb{E}\left[\sup_{t_1 \leq t \leq t_2} |X^{(n)}_t - X^{(n)}_{t_1}|\right]
\leq \mathbb{E}\left[\int_{t_1}^{t_2} |b(X^{(n)}_s, \Lambda^{(n)}_s, \mu^{(n)}_s)| ds \right] + C\mathbb{E}\left[\left(\int_{t_1}^{t_2} \left\|\sigma(X^{(n)}_s, \Lambda^{(n)}_s, \mu^{(n)}_s)\right\|^2 ds\right)^{1/2}\right]
\leq C\mathbb{E}\left[\int_{t_1}^{t_2} (1 + |X^{(n)}_s|^2)^{1/2} ds\right] + C\mathbb{E}\left[\left(\int_{t_1}^{t_2} 1 + |X^{(n)}_s|^2 ds\right)^{1/2}\right]
\leq C \max\{|t_2 - t_1|, \sqrt{|t_2 - t_1|}\},
\]

where \(C\) denotes a positive constant whose value may be different from line to line. By taking expectation in (2.14), we have
\[
\mathbb{P}(\Lambda^{(n)}_{t_1} \neq \Lambda^{(n)}_{t_2}) \leq 1 - e^{-M(t_2 - t_1)}.
\]
Inserting these estimates into (2.15), we obtain that there exists a $C > 0$ such that
\[
\mathbb{E}W_1(\mu_t^{(n)}, \mu_t^{(n)}) \leq C \text{ max}\{\left| t_2 - t_1 \right|, \sqrt{\left| t_2 - t_1 \right|}\}. \tag{2.17}
\]

Applying [3, Theorem 12.3], we obtain that the set $(\mathcal{L}_{\mu}^n)_{n \geq 1}$ is tight. Analogously, we can show that $(\mathcal{L}_{\nu}^n)_{n \geq 1}$ is tight.

By Itô’s formula, for $0 \leq t_1 < t_2 \leq T$,
\[
\mathbb{E}|X_{t_2}^{(n)} - X_{t_1}^{(n)}|^4 \\
\leq 8\mathbb{E}\left| \int_{t_1}^{t_2} b(X_r^{(n)}, \Lambda_r^{(n)}, \mu_r^{(n)})\,dr \right|^4 + 8\mathbb{E}\left| \int_{t_1}^{t_2} \sigma(X_r^{(n)}, \Lambda_r^{(n)}, \mu_r^{(n)})\,dB_r \right|^4 \\
\leq 8(t_2 - t_1)^3\mathbb{E}\left| \int_{t_1}^{t_2} |b(X_r^{(n)}, \Lambda_r^{(n)}, \mu_r^{(n)})|^4\,dr + 288(t_2 - t_1)\mathbb{E}\int_{t_1}^{t_2} |\sigma(X_r^{(n)}, \Lambda_r^{(n)}, \mu_r^{(n)})|^4\,dr \right| \\
\leq C(t_2 - t_1)\int_{t_1}^{t_2} (1 + \mathbb{E}|X_r^{(n)}|^4)\,dr.
\]

Applying condition (H1) again, we have $\int_0^T \mathbb{E}|X_t|^4\,dt \leq C$ for some constant $C$, independent of $n$, (cf. [33, Theorem 3.20]). Furthermore, invoking the fact $X_0 = x$, we conclude that $(\mathcal{L}_{\Lambda}^n)_{n \geq 1}$ is tight by virtue of [3, Theorem 12.3].

**Step 2.** Because all the marginal distributions of $R_n$, $n \geq 1$ are tight, we get $R_n$, $n \geq 1$ is tight as well. Indeed, for any $\varepsilon > 0$, there exist compact subsets $K_1 \subset C([0, T]; \mathbb{R}^n)$, $K_2 \subset \mathcal{D}([0, T]; S)$, and $K_3, K_4 \subset \mathcal{U}$ such that for every $n \geq 1$,
\[
\min\{\mathcal{L}_{B}^n(K_1), \mathcal{L}_{\Lambda}^n(K_2), \mathcal{L}_{\Lambda}^n(K_3), \mathcal{L}_{\mu}^n(K_4)\} \geq 1 - \varepsilon.
\]

This yields that
\[
R_n(K_1 \times K_2 \times K_3 \times K_4) \geq 1 - \mathcal{L}_{\Lambda}^n(K_1^c) - \mathcal{L}_{\Lambda}^n(K_2^c) - \mathcal{L}_{\mu}^n(K_3^c) - \mathcal{L}_{\nu}^n(K_4^c) \geq 1 - 4\varepsilon.
\]

So $(R_n)_{n \geq 1}$ is tight.

As a consequence of the tightness of $(R_n)_{n \geq 1}$, up to extracting a subsequence, we have $R_n$ converges weakly to some probability measure $R_0$ on $\mathcal{Y}$. Since $\mathcal{Y}$ is a Polish space, according to Skorokhod’s representation theorem (cf. [10], Theorem 1.8, p.102), there exists a probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ on which defined a sequence of random variables $Y_n = (X_t^{(n)}, \Lambda_t^{(n)}, \mu_t^{(n)}, \nu_t^{(n)})_{t \in [0, T]} \in \mathcal{Y}$, $n \geq 0$, with the distribution $R_n$, $n \geq 0$, respectively such that
\[
\lim_{n \to \infty} Y_n = Y_0, \quad \mathbb{P}'\text{-a.s..} \tag{2.18}
\]
In this step we want to show that $Y_0 = (X_t^{(0)}, \Lambda_t^{(0)}, \mu_t^{(0)}, \nu_t^{(0)})$ is also associated with an admissible control.

For $0 \leq t_1 < t_2 < \ldots < t_k \leq T$, define the projection map $\pi_{t_1 \ldots t_k} : \mathcal{D}([0, T]; \mathcal{S}) \to \mathcal{S}^k$ by

$$\pi_{t_1 \ldots t_k} (\Lambda) = (\Lambda_{t_1}, \ldots, \Lambda_{t_k}).$$

Let $\mathcal{T}_0$ consist of those $t \in [0, T]$ for which the projection $\pi_t : \mathcal{D}([0, T]; \mathcal{S}) \to \mathcal{S}$ is continuous except at points from a set of $R_0$-measure 0. For $t \in [0, T]$, $t \in \mathcal{T}_0$ if and only if $R_0(J_t) = 0$, where

$$J_t = \{ \Lambda \in \mathcal{D}([0, T]; \mathcal{S}); \Lambda_t \neq \Lambda_{t-} \}.$$ 

Also, $0, T \in \mathcal{T}_0$ by convention. It is known that the complement of $\mathcal{T}_0$ in $[0, T]$ is at most countable (cf. [3, p. 124]). So, for every bounded function $h$ on $\mathcal{S}$,

$$\lim_{n \to \infty} \int_s^t h(\Lambda_r^{(n)}) dr = \int_s^t h(\Lambda_r^{(0)}) dr, \quad 0 \leq s < t \leq T, \ \mathbb{P}$$.a.s.

Combining this with the almost sure convergence of $(X_t^{(n)})_{t \in [0, T]}$ to $(X_t^{(0)})_{t \in [0, T]}$, and $(\mu_t^{(n)})_{t \in [0, T]}$ to $(\mu_t^{(0)})_{t \in [0, T]}$ in $\mathcal{C}([0, T]; \mathbb{R}^d)$ and $\mathcal{C}([0, T]; \mathcal{P}(U))$ respectively, by passing $n$ to $\infty$ in the equation

$$X_t^{(n)} = x + \int_0^t b(X_r^{(n)}, \Lambda_r^{(n)}, \mu_r^{(n)}) dr + \int_0^t \sigma(X_r^{(n)}, \Lambda_r^{(n)}, \mu_r^{(n)}) dB_r,$$

we obtain that

$$X_t^{(0)} = x + \int_0^t b(X_r^{(0)}, \Lambda_r^{(0)}, \mu_r^{(0)}) dr + \int_0^t \sigma(X_r^{(0)}, \Lambda_r^{(0)}, \mu_r^{(0)}) dB_r.$$

In terms of Skorokhod’s representation (2.9) for jumping process $(\Lambda_t^{(n)})$, we have

$$\Lambda_t^{(n)} = i + \int_0^t \int_{[0, H]} \vartheta(X_r^{(n)}, \Lambda_r^{(n)}, \mu_r^{(n)}, z) N_1 (dr, dz).$$

Since $(x, \nu) \mapsto q_{ij}(x, \nu)$ is continuous for every $i, j \in \mathcal{S}$, one gets $1_{\Gamma_{ij}(y, \nu')}(z)$ tends to $1_{\Gamma_{ij}(x, \nu)}(z)$ as $|y - x| \to 0$ and $W_1(\nu', \nu) \to 0$. Letting $n \to \infty$ in (2.20), we obtain

$$\Lambda_t^{(0)} = i + \int_0^t \int_{[0, H]} \vartheta(X_r^{(0)}, \Lambda_r^{(0)}, \mu_r^{(0)}, z) N_1 (dr, dz).$$

13
By Skorokhod’s representation (2.9), this yields that
\[ P(\Lambda_t^{(0)} = j | \Lambda_t^{(0)} = i, X_t^{(0)} = x, \nu_t^{(0)} = \nu) = \begin{cases} 
q_{ij}(x, \nu)\delta + o(\delta), & \text{if } i \neq j, \\
1 + q_{ii}(x, \nu)\delta + o(\delta), & \text{otherwise,}
\end{cases} \]
provided \( \delta > 0 \). Moreover, there is no \( t_0 \in [0, T] \) such that \( P'(\Lambda_{t_0}^{(0)} \neq \Lambda_{t_0-}^{(0)}) > 0 \), which means that \( T_0 = [0, T] \). Hence, \( \lim_{n \to \infty} \Lambda_{tn}^{(n)} = \Lambda_t^{(0)} \) \( \mathbb{P}' \)-a.s. for every \( t \in [0, T] \).

For each \( n \geq 1 \), \( F^{(n)} \) is a function from \([0, T] \times \mathcal{C}([0, T]; \mathbb{R}^d) \times \mathcal{S}\) to a compact space \((\mathcal{P}(U), W_1)\). By the Ascoli-Arzelà theorem (in a generalized form), up to extracting a subsequence, there exist functionals \( F^{(0)} \) and \( G^{(0)} \) in \( \mathcal{Y}_\psi \) such that \( F^{(n)} \) and \( G^{(n)} \) converge uniformly to \( F^{(0)} \) and \( G^{(0)} \) respectively as \( n \to \infty \). Joining with the fact that \( Y_n \) converges almost surely to \( Y_0 \), we obtain from (2.12) that
\[
\mu_t^{(0)} = F^{(0)}(t, \Xi X_t^{(0)}; \Lambda_t^{(0)}), \quad \nu_t^{(0)} = G^{(0)}(t, \Xi X_t^{(0)}; \Lambda_t^{(0)}), \quad t \in (0, T].
\]
(2.22)

Consequently, by (2.19), (2.21) and (2.22), we conclude that \( \alpha_0 := (\mu_t^{(0)}, \nu_t^{(0)}) \) associated with \( Y_0 = (X_t^{(0)}, \Lambda_t^{(0)}, \mu_t^{(0)}, \nu_t^{(0)}) \) is an admissible control in \( \Pi_{0,x,i} \).

**Step 3.** By (2.11) and the lower semi-continuity of \( f \) and \( g \), we have
\[
V(0, x, i) = \lim_{n \to \infty} J(0, x, i, \alpha_n)
= \lim_{n \to \infty} \mathbb{E}_{\mathbb{P}'} \left[ \int_0^T f(t, X_t^{(n)}; \Lambda_t^{(n)}, \mu_t^{(n)}, \nu_t^{(n)})dt + g(X_T^{(n)}) \right]
\geq \mathbb{E}_{\mathbb{P}'} \left[ \int_0^T f(t, X_t^{(0)}; \Lambda_t^{(0)}, \mu_t^{(0)}, \nu_t^{(0)})dt + g(X_T^{(0)}) \right]
= J(0, x, i, \alpha_0)
\geq V(0, x, i).
\]

Therefore, \( \alpha_0 \) is an optimal admissible control. The proof of this theorem is complete. \( \square \)

Next, we generalize the class of feedback control policies. Let \( \tilde{\psi} : [0, \infty) \to [0, \infty) \) be a measurable function such that
\[
\lim_{r \downarrow 0} \tilde{\psi}(r) = 0.
\]

**Definition 2.3** A feedback control \( \alpha = (\mu_t, \nu_t)_{t \in [s, T]} \) is said to be in the class \( \bar{\Pi}_{s,x,i} \) for \((s, x, i) \in [0, T) \times \mathbb{R}^d \times \mathcal{S}\), if it satisfies
(1) $\mu : [s, T] \to \mathcal{P}(U)$, $\nu : [s, T] \to \mathcal{P}(U)$ are measurable curves such that for every $t_1, t_2 \in [s, T]$,

$$W_1(\mu_{t_1}, \mu_{t_2}) \leq \tilde{\psi}(|t_1 - t_2|), \ W_1(\nu_{t_1}, \nu_{t_2}) \leq \tilde{\psi}(|t_1 - t_2|) \ \text{a.s.}$$

(2) SDEs (2.4) and (2.5) admit a strong solution $(X_t, \Lambda_t)$ with initial value $(X_s, \Lambda_s) = (x, i)$ under the control $(\mu_t, \nu_t)_{t \in [s, T]}$.

(3) $\mu_t$ and $\nu_t$ are adapted to the $\sigma$-fields $\mathcal{F}_t = \sigma\{(X_u, \Lambda_u); s \leq u \leq t\}$. Here and in the sequel the overline in $\sigma\{(X_u, \Lambda_u); s \leq u \leq t\}$ means the completion of the $\sigma\{(X_u, \Lambda_u); s \leq u \leq t\}$.

Note that according to the probability measure theory, the condition that $\mu_t$ is adapted to $\mathcal{F}_t = \sigma\{(X_u, \Lambda_u); s \leq u \leq t\}$ implies that there is a measurable function $F_t$ such that $\mu_t = F_t(X_{[s,t]}, \Lambda_{[s,t]})$ almost surely. Therefore, condition (3) in Definition 2.3 ensures that the control policies $\mu_t$ and $\nu_t$ are a kind of feedback control.

**Theorem 2.4** Assume (H1)-(H4) hold. Then for every $(s, x, i) \in [0, T) \times \mathbb{R}^d \times S$, there exists an optimal feedback control $\alpha^* \in \tilde{\Pi}_{s,x,i}$ such that

$$V(s, x, i) := \inf_{\alpha \in \tilde{\Pi}_{s,x,i}} J(s, x, i, \alpha) = J(s, x, i, \alpha^*),$$

where $J(s, x, i, \alpha)$ is given by (2.6).

**Proof.** We follow the same procedure as Theorem 2.2 to prove this theorem by pointing out the different part.

Consider the nontrivial case $V(0, x, i) < \infty$, and there exists a sequence $\alpha_n = (\mu_{(n)}, \nu^{(n)})$ in $\tilde{\Pi}_{0,x,i}$ such that

$$V(0, x, i) = \lim_{n \to \infty} J(0, x, i, \alpha_n).$$

Denote still by $R_n$ the joint distribution of $(X_{t}^{(n)}, \Lambda_{t}^{(n)}, \mu_{t}^{(n)}, \nu^{(n)})$ in $\mathcal{Y}$ associated with $\alpha_n$. The tightness of $\mathcal{L}_X^n$ and $\mathcal{L}_X^n$ can be proved similar to the argument of Theorem 2.2. Moreover, the tightness of $\mathcal{L}_\mu^n$ and $\mathcal{L}_\nu^n$ follows immediately from the observation that the set

$$\mathcal{U} := \{\mu : [0, T] \to \mathcal{P}(U); W_1(\mu_{t_1}, \mu_{t_2}) \leq \tilde{\psi}(|t_1 - t_2|), \ t_1, t_2 \in [0, T]\}$$
is a compact subset of $C([0, T]; \mathcal{P}(U))$ by virtue of the Ascoli-Arzelà theorem. Consequently, the set of distributions $R_n$, $n \geq 1$, is tight, and up to taking a subsequence, there is a probability measure $R_0$ in $\mathcal{Y}$ such that

$$R_n \text{ weakly converges to } R_0 \text{ as } n \to \infty.$$ 

The enlargement of the set of control policies makes the proof of this part easier, however, the difficulty lies in how to ensure the limit still being a feedback control.

According to Skorokhod's representation theorem, there is a new probability space $(\Omega', \mathcal{F}', \mathbb{P}')$ and a sequence of random variables $Y_n = (X_t^{(n)}, \Lambda_t^{(n)}, \mu_t^{(n)}, \nu_t^{(n)})_{t \in [0, T]}$ in $\mathcal{Y}$ with distribution $R_n$, $n \geq 0$, such that

$$\lim_{n \to \infty} Y_n = Y_0, \quad \mathbb{P}'-\text{a.s..}$$

Analogous to the argument of Theorem 2.2, we can show (2.19) and (2.21) still hold for current $(X_t^{(0)}, \Lambda_t^{(0)}, \mu_t^{(0)}, \nu_t^{(0)})$. The main different part is to show $\mu_t^{(0)}$ and $\nu_t^{(0)}$ are adapted to the $\sigma$-fields generated by $(X_r^{(0)}, \Lambda_r^{(0)})$ up to time $t$.

To this aim, we adopt the notation in the study of backward martingale to define

$$\mathcal{F}^{X_{\Lambda}}_{-n,t} = \sigma\{ (X_r^{(m)}, \Lambda_{r}^{(m)}); m \geq n, r \in [0, t] \}.$$ 

Then

$$\mathcal{F}^{X_{\Lambda}}_{-\infty,t} \supset \mathcal{F}^{X_{\Lambda}}_{-1,t} \supset \cdots \supset \mathcal{F}^{X_{\Lambda}}_{-n,t} \supset \mathcal{F}^{X_{\Lambda}}_{-n-1,t} \supset \cdots.$$ 

Put $\mathcal{F}^{X_{\Lambda}}_{-\infty,t} = \bigcap_{n \geq 1} \mathcal{F}^{X_{\Lambda}}_{-n,t}$. $\mathcal{F}^{X_{\Lambda}}_{-\infty,t}$ is easily checked to be a $\sigma$-field which concerns only the limit behavior of the sequence $(X_r^{(n)}, \Lambda_r^{(n)})_{r \in [0, t]}$ as $n$ tends to $\infty$. Moreover, since $\lim_{n \to \infty} \Lambda_t^{(n)} = \Lambda_t^{(0)}$ and $\lim_{n \to \infty} X_t^{(n)} = X_t^{(0)}$ a.s. for every $t \in [0, T]$, it holds

$$\mathcal{F}^{X_{\Lambda}}_{-\infty} = \sigma\{ (X_r^{(0)}, \Lambda_r^{(0)}); r \in [0, t] \}.$$ 

Define $\mathcal{F}^{\mu}_{-n} = \sigma\{ \mu_t^{(m)}, m \geq n \}$. Due to (3) of Definition 2.3, $\mu_t^{(n)}$ is in $\mathcal{F}^{X_{\Lambda}}_{-n}$ for each $n \geq 1$, and hence $\mathcal{F}^{\mu}_{-n} \subset \mathcal{F}^{X_{\Lambda}}_{-n,t}$. Therefore, it follows from the fact $\lim_{n \to \infty} W_1(\mu_t^{(n)}, \mu_t^{(0)}) = 0$ a.s. that

$$\sigma\{ \mu_t^{(0)} \} \subset \bigcap_{n \geq 1} \mathcal{F}^{\mu}_{-n,t} \subset \bigcap_{n \geq 1} \mathcal{F}^{X_{\Lambda}}_{-n,t} = \mathcal{F}^{X_{\Lambda}}_{-\infty,t} = \sigma\{ (X_r^{(0)}, \Lambda_r^{(0)}); r \in [0, t] \},$$

which means that $\mu_t^{(0)}$ is adapted to $\mathcal{F}_t^{(0)} := \sigma\{ (X_r^{(0)}, \Lambda_r^{(0)}); r \in [0, t] \}$ as desired. Similarly, we can show that $\nu_t^{(0)}$ is also adapted to $\mathcal{F}_t^{(0)}$. 

16
At last, similar to Step 3 in the argument of Theorem 2.2, it follows from the lower semi-continuity of $f$ and $g$ that

$$V(0, x, i) = \lim_{n \to \infty} J(0, x, i, \alpha_n) = J(0, x, i, \alpha_0).$$

Hence, $\alpha_0 = (\mu^{(0)}, \nu^{(0)})$ is an optimal feedback control. The proof is complete. \hfill \Box

3 Dynamic programming principle

In this section we first study the continuity of the value function $V(s, x, i)$, then establish the dynamic programming principle for $V(s, x, i)$. In this part, compared with the optimal control problem for stochastic processes without regime-switching, the regime-switching process $(\Lambda_t)$ plays an important role in studying the continuity of the value function. The application of Skorokhod’s representation for the jumping process $(\Lambda_t)$ helps us to separate the intensive interaction between $(X_t)$ and $(\Lambda_t)$ in some sense.

3.1 Continuity of value functions for processes with regime-switching

**Theorem 3.1** Assume (H1)-(H4) hold and $\sigma(x, i, \mu) = \sigma(x)$ for all $(x, i, \mu) \in \mathbb{R}^d \times S \times \mathcal{P}(U)$. Suppose that there exist constants $C_3 > 0$, $p \geq 1$ and $C_4 > 0$ such that

$$|f(t, x, i, \mu, \nu) - f(t, y, j, \mu', \nu')| \leq C_3(|x - y| + 1_{i \neq j} + W_1(\mu, \mu') + W_1(\nu, \nu'))$$

and

$$|g(x) - g(y)| \leq C_3|x - y|, \quad \forall t \in [0, T], x, y \in \mathbb{R}^d, i, j \in S, \mu, \mu', \nu, \nu' \in \mathcal{P}(U),$$

and

$$|f(t, x, i, \mu, \nu)| \leq C_4(1 + |x|^p), \quad \forall t \in [0, T], x \in \mathbb{R}^d, i \in S, \mu, \nu \in \mathcal{P}(U). \quad (3.1)$$

Then the value function $(s, x, i) \mapsto V(s, x, i)$ is continuous in $[0, T] \times \mathbb{R}^d \times S$.

**Proof.** Since the topology of $S$ is the discrete topology, it is necessary to consider the continuity of $(s, x) \mapsto V(s, x, i)$, which will be proved by studying first the continuity of $s \mapsto V(s, x, i)$ and then the continuity of $x \mapsto V(s, x, i)$.
Given \((x, i) \in \mathbb{R}^d \times S\), for any two \(s, s' \in [0, T]\) with \(s \neq s'\), we go to estimate \(|V(s, x, i) - V(s', x, i)|\). According to Theorem 2.2, there exists an optimal admissible control \(\alpha^* = (\mu, \nu)\) in \(\Pi_{s, x, i}\) with \(\mu_t = F(t,\Xi X_{[s,t]}, \Lambda_t), \nu_t = G(t,\Xi X_{[s,t]}, \Lambda_t)\) for some \(F, G \in \Upsilon_\psi\) such that

\[
V(s, x, i) = \mathbb{E}\left[\int_s^T f(r, X_r, \Lambda_r, \mu_r, \nu_r) dr + g(X_T)\right]. \tag{3.2}
\]

Let us consider a transformation of \((X_t, \Lambda_t, \mu_t)\) under a time-shift. Let \(\Delta s = s' - s\). Define

\[
\tilde{X}_t = X_{t-\Delta s}, \ \tilde{\Lambda}_t = \Lambda_{t-\Delta s}, \ \tilde{\mu}_t = \mu_{t-\Delta s}, \ \tilde{\nu}_t = \nu_{t-\Delta s}, \ \tilde{B}_t = B_{t-\Delta s}, \ t \geq 0 \vee \Delta s. \tag{3.3}
\]

It is easy to check that \((\tilde{X}_t, \tilde{\Lambda}_t)\) satisfies the following equations

\[
d\tilde{X}_t = b(\tilde{X}_t, \tilde{\Lambda}_t, \tilde{\mu}_t)dt + \sigma(\tilde{X}_t)d\tilde{B}_t,
\]

and

\[
\mathbb{P}(\tilde{\Lambda}_{t+\delta} = j| \tilde{\Lambda}_t = i', \ \tilde{X}_t = \tilde{x}, \tilde{\nu}_t = \tilde{\nu}) = \begin{cases} q_{ij}(\tilde{x}, \tilde{\nu})\delta + o(\delta), & \text{if } j \neq i', \\ 1 + q_{ij}(\tilde{x}, \tilde{\nu})\delta + o(\delta), & \text{otherwise}, \end{cases}
\]

with \(\tilde{X}_{s'} = X_s = x, \ \tilde{\Lambda}_{s'} = \Lambda_s = i\). In order to guarantee that \((\tilde{\mu}, \tilde{\nu}, s', x, i)\) is an admissible control, we need to find two functionals \(\tilde{F}, \tilde{G} \in \Upsilon_\psi\) such that

\[
\tilde{\mu}_t = \tilde{F}(t, \Xi \tilde{X}_{[s', t]}, \tilde{\Lambda}_t), \ \tilde{\nu}_t = \tilde{G}(t, \Xi \tilde{X}_{[s', t]}, \tilde{\Lambda}_t), \ t \geq s'. \tag{3.4}
\]

Indeed, we define a shift operator \(\theta_{\Delta s}\) on the path space \(C([0, T]; \mathbb{R}^d)\) or \(D([0, T]; S)\) by

\[(\theta_{\Delta s}x)_r = x_{r+\Delta s}, \text{ if } r + \Delta s \geq 0; \ (\theta_{\Delta s}x)_r = x_0, \text{ otherwise.}\]

Then we can choose

\[
\tilde{F}(t, x, j) := F(t - \Delta s, (\theta_{\Delta s}x)_r, j), \ \tilde{G}(t, x, j) := G(t - \Delta s, (\theta_{\Delta s}x)_r, j)
\]

for \(t \in [0, T], x, j \in C([0, T]; \mathbb{R}^d)\) and \(j \in S\). It is easy to check that \(\tilde{F}, \tilde{G}\) are in \(\Upsilon_\psi\). In addition, by noting

\[
\Xi \tilde{X}_{[s', t]}(r) = \begin{cases} X_s \quad r < s', \\ X_{r-\Delta s} \quad r \in [s', t], \\ X_{t-\Delta s} \quad r > t, \end{cases}
\]

18
we can check directly that (3.4) holds.

By the definition of the value function and (3.2), it holds
\[ V(s',x,i) - V(s,x,i) \]
\[ \leq \mathbb{E}\left[ \int_s^{s'} f(r,\bar{X}_r,\bar{\Lambda}_r,\bar{\mu}_r,\bar{\nu}_r)dr + g(\bar{X}_t) - \int_s^T f(r,X_r,\Lambda_r,\mu_r,\nu_r)dr - g(X_T) \right] \]
\[ \leq \mathbb{E}\left[ \int_s^{s'} f(r,X_r,\Lambda_r,\mu_r,\nu_r)dr \right] + \mathbb{E}\left[ \int_{T-\Delta s}^T f(r,X_r,\Lambda_r,\mu_r,\nu_r)dr \right] \]
\[ + \mathbb{E}[|g(X_{T-\Delta s}) - g(X_T)|] \]
\[ \leq 2C_4|\Delta s| \sup_{r\in[0,T+1]} \mathbb{E}|X_r|^p + C_3 \mathbb{E}[|X_T - X_{T-\Delta s}|]. \tag{3.5} \]

Condition (H1) and the compactness of the space \((\mathcal{P}(U), W_1)\) yield that there exists a constant \(C > 0\) such that
\[ |b(x',i',\mu)| + \|\sigma(x',i',\mu)\| \leq C(1 + |x'|), \quad \forall x' \in \mathbb{R}^d, i' \in S, \mu \in \mathcal{P}(U). \]

Then it follows that
\[ \mathbb{E}\left[ \sup_{0\leq t\leq T+1} |X_t|^p \right] \leq C(T, x, p), \tag{3.6} \]
for some constant \(C(T, x, p) > 0\). See, for instance, [33, Theorem 3.24] for such kind of estimate. Moreover,
\[ \mathbb{E}|X_{T-\Delta s} - X_T| \leq \mathbb{E}\left[ \int_{T-\Delta s}^T b(X_r,\Lambda_r,\mu_r)dr + \int_{T-\Delta s}^T \sigma(X_r)dB_r \right] \]
\[ \leq C \max\{|\Delta s|, \sqrt{|\Delta s|}\}, \tag{3.7} \]
where \(C > 0\) is a generic constant. Inserting the estimates (3.6) and (3.7) into (3.5), we get
\[ V(s',x,i) - V(s,x,i) \leq C \max\{|\Delta s|, \sqrt{|\Delta s|}\}. \tag{3.8} \]
By the symmetric position of \(s\) and \(s'\), (3.8) further leads to
\[ |V(s',x,i) - V(s,x,i)| \leq C \max\{|\Delta s|, \sqrt{|\Delta s|}\}. \tag{3.9} \]

Next, we go to estimate \(V(s,x',i) - V(s,x,i)\) for \(x, x' \in \mathbb{R}^d\). Consider the following SDEs
\[ dX'_t = b(X'_t,\Lambda'_t,\mu'_t)dt + \sigma(X'_t)dB_t, \quad X'_s = x', \tag{3.10} \]
and
\[ d\Lambda_t' = \int_{[0,T]} \vartheta(X_t', \Lambda_{t-}', \nu_{t-}', z) \mathcal{N}_1 (dt, dz), \quad \Lambda_0' = i, \quad (3.11) \]

where
\[ \mu' = F(t, \Xi X_{[s,t]}, \Lambda_t'), \quad \nu' = G(t, \Xi X_{[s,t]}, \Lambda_t'), \quad t \in [s, T]. \]

Due to (H1)-(H3), SDEs (3.10) and (3.11) admit a unique strong solution \((X_t', \Lambda_t')\) with initial value \((x', i)\). Then, \(\alpha' = (\mu', \nu')\) is an admissible control in \(\Pi_{s, x', i}\). Thus,
\[ V(s, x', i) \leq \mathbb{E} \left[ \int_s^T f(r, X_r', \Lambda_r', \mu_r', \nu_r') dr + g(X_T') \right]. \quad (3.12) \]

Moreover, by (H1) and \(F, G \in \mathcal{Y}_\psi\), we get
\[ |X_t' - X_t| \]
\[ \leq |x' - x| + \int_s^t |b(X_r', \Lambda_r', \mu_r') - b(X_r, \Lambda_r, \mu_r)| dr + \left| \int_s^t (\sigma(X_r') - \sigma(X_r)) dB_r \right| \]
\[ \leq |x' - x| + \int_s^t C_1 (|X_r' - X_r| + 1_{\Lambda_r' \neq \Lambda_r} + W_1(\mu_r', \mu_r)) dr + \left| \int_s^t (\sigma(X_r') - \sigma(X_r)) dB_r \right| \quad (3.13) \]
\[ \leq |x' - x| + \int_s^t C_1 (|X_r' - X_r| + (1 + \kappa_2) 1_{\Lambda_r' \neq \Lambda_r} + \kappa_2 \sup_{s \leq u \leq t} |X_u' - X_u|) dr \]
\[ + \left| \int_s^t (\sigma(X_r') - \sigma(X_r)) dB_r \right|. \]

Applying Burkholder-Davis-Gundy’s inequality,
\[ \mathbb{E} \sup_{s \leq u \leq t} \left| \int_s^u (\sigma(X_r') - \sigma(X_r)) dB_r \right| \leq \mathbb{E} \left[ \left( \int_s^t |\sigma(X_r') - \sigma(X_r)|^2 dr \right)^{\frac{1}{2}} \right] \]
\[ \leq C_1 \mathbb{E} \left[ \left( \int_s^t |X_r' - X_r|^2 dr \right)^{\frac{1}{2}} \right] \]
\[ \leq \frac{1}{2} \mathbb{E} \left[ \sup_{s \leq u \leq t} |X_u' - X_u| + \frac{C_2^2}{2} \mathbb{E} \int_s^t |X_r' - X_r| dr \right] \]

Combining this with (3.13) and Lemma 3.2 below, we obtain that
\[ \mathbb{E} \sup_{s \leq u \leq t} |X_u' - X_u| \leq 2|x' - x| + C \int_s^t \mathbb{E} \sup_{s \leq u \leq r} |X_u' - X_u| dr, \]
where \(C = 2C_1(1 + \kappa_2) + 2C_1C_2N^2T(1 + \kappa_2) + C_1^2\). Hence, Gronwall’s inequality yields that
\[ \mathbb{E} \sup_{s \leq u \leq t} |X_u' - X_u| \leq 2|x' - x|e^{C(t-s)}. \quad (3.14) \]
Invoking (3.2) and (3.12),
\[
V(s, x', i) - V(s, x, i) \\
\leq \mathbb{E} \left[ \int_s^T (f(r, X'_r, \Lambda'_r, \mu'_r, \nu'_r) - f(r, X_r, \Lambda_r, \mu_r, \nu_r))dr + g(X'_T) - g(X_T) \right] \\
\leq C_3 \left( \int_s^T \mathbb{E}(|X'_r - X_r| + 1_{\Lambda'_r \neq \Lambda_r} + W_1(\mu'_r, \mu_r) + W_1(\nu'_r, \nu_r))dr \right) + C_3 \mathbb{E}|X'_T - X_T| \\
\leq C_3 \left( \int_s^T (1 + 2\kappa_2)\mathbb{E}\left( \sup_{s \leq u \leq r} |X'_u - X_u| + 1_{\Lambda'_r \neq \Lambda_r} \right)dr \right) + C_3 \mathbb{E}|X'_T - X_T|.
\]

By using Lemma 3.2, we get
\[
V(s, x', i) - V(s, x, i) \leq C \int_s^t \mathbb{E} \sup_{s \leq u \leq r} |X'_u - X_u|dr + C_3 \mathbb{E}|X'_T - X_T| \quad (3.15)
\]
where \( C = C_2C_3N^2T(1 + \kappa_2)(1 + 2\kappa_2) + C_3(1 + 2\kappa_2) \). Thus, by (3.14), we finally obtain that
\[
V(s, x', i) - V(s, x, i) \leq C(T, N, C_2, C_3, \kappa_2)|x' - x|. \quad (3.16)
\]
Since the position of \( x \) and \( x' \) in (3.16) is symmetric, we have
\[
|V(s, x', i) - V(s, x, i)| \leq C(T, N, C_2, C_3, \kappa_2)|x' - x|, \quad (3.17)
\]
where \( C(T, N, C_2, C_3, \kappa_2) \) is a positive constant independent of \( s \). Hence, \( x \mapsto V(s, x, i) \) is uniformly continuous relative to \( s \).

Consequently, it follows from (3.9) and (3.17) that
\[
|V(s, x, i) - V(s', x, i)| \leq |V(s, x, i) - V(s', x, i)| + |V(s', x, i) - V(s', x', i)| \\
\leq C \left( \max\{|s - s'|, \sqrt{|s - s'|}\} + |x - x'| \right). \quad (3.18)
\]
We conclude that \( (s, x, i) \mapsto V(s, x, i) \) is continuous. \( \square \)

**Lemma 3.2** Under the same assumptions and notation as Theorem 3.1, it holds
\[
\int_0^t \mathbb{P}(\Lambda'_s \neq \Lambda_s)ds \leq N(N - 1)C_2(1 + \kappa_2)t \int_0^t \mathbb{E}|X_s - X'_s|ds, \quad t \in [0, T]. \quad (3.19)
\]

**Proof.** For the clarity of the calculation, we present a explicit construction of the Poisson random measure \( N_1(dt, dz) \) used in Skorokhod's representation for jumping processes.

21
Let $\xi_k, k = 1, 2, \ldots$ be random variables supported on $[0, H]$ with
$$\mathbb{P}(\xi_k \in dx) = m(dx)/H,$$
where $m(dx)$ stands for the Lebesgue measure on $[0, H]$ and $H = N(N - 1)M$. Let $\tau_k, k = 1, 2, \ldots$ be nonnegative variables such that $\mathbb{P}(\tau_k > t) = \exp(-Ht), t \geq 0$. Set
$$\zeta_1 = \tau_1, \zeta_2 = \tau_1 + \tau_2, \ldots, \zeta_k = \tau_1 + \tau_2 + \ldots + \tau_k, \quad k > 2,$$
and
$$D_{pt} = \{\zeta_1, \zeta_2, \ldots, \zeta_k, \ldots\}.$$
Define $p_1(\zeta_k) = \xi_k, k \geq 1,$ and
$$N_1((0, t] \times A) = \#\{s \in D_{pt}, s \leq t, p_1(s) \in A\}, \quad t > 0, A \in \mathcal{B}(\mathbb{R}).$$
Then $p_1(\cdot)$ is a Poisson process and $N_1(dt, dz)$ is a Poisson random measure with intensity $dtdm(dz)$.

For simplicity of notation, put $q_{ij}(t) = q_{ij}(X_t, \nu_t)$, $q'_{ij}(t) = q_{ij}(X'_t, \nu'_t), i, j \in S$, and $Q_t = (q_{ij}(t)), Q'_t = (q_{ij}(t))$. Denote $\Gamma_{ij}(t) = \Gamma_{ij}(X_t, \nu_t)$ and $\Gamma'_{ij}(t) = \Gamma_{ij}(X'_t, \Lambda'_t)$ where $\Gamma_{ij}(x, \nu)$ is defined in the beginning of Subsection 2.2. According to Skorokhod’s representation theorem, the process $(\Lambda'_t)$ given by (3.11) satisfies
$$\mathbb{P}(\Lambda'_{t+\delta} = j|\Lambda'_t = \iota', X'_t = x', \nu'_t = \nu') = \begin{cases} q_{\iota'j}(x', \nu')\delta + o(\delta), & \text{if } j \neq \iota', \\ 1 + q_{\iota'\iota}(x', \nu')\delta + o(\delta), & \text{otherwise}, \end{cases}$$

Let $\Gamma_{ij}(t)\Delta\Gamma'_{ij}(t) = (\Gamma_{ij}(t)\setminus\Gamma'_{ij}(t)) \cup (\Gamma'_{ij}(t)\setminus\Gamma_{ij}(t))$. By virtue of the construction of $\Gamma_{ij}(t)$ and $\Gamma'_{ij}(t)$,
$$m(\Gamma_{ij}(t)\Delta\Gamma'_{ij}(t)) \leq \left| \sum_{k=1}^{i-1} q_k(t) + \sum_{k=1, k \neq i}^{j-1} q_{ik}(t) - \sum_{k=1}^{i-1} q'_k(t) - \sum_{k=1, k \neq i}^{j-1} q'_{ik}(t) \right|$$
$$+ \left| \sum_{k=1}^{i-1} q_k(t) + \sum_{k=1, k \neq i}^{j} q_{ik}(t) - \sum_{k=1}^{i-1} q'_k(t) - \sum_{k=1, k \neq i}^{j} q'_{ik}(t) \right|$$
$$\leq 2N \max_{k \in S} \left\{ \sum_{j \neq k} \left| q_{kj}(t) - q'_{jk}(t) \right| \right\}$$
$$\leq 2C_2N(N - 1)((1 + \kappa_2)|X_t - X'_t| + \kappa_21_{\Lambda_t \neq \Lambda'_t}),$$

22
where in the last step we have used (H3) and $G \in \Upsilon_\psi$.

For $\delta \in (0, 1)$ and $s > 0$, denote by $s_\delta = \lfloor s/\delta \rfloor$ the integer part of $s/\delta$. Let $N(t) = N_1((0, t] \times S)$. For every $t \in (0, \delta]$, since $\Lambda'_0 = \Lambda_0 = i$,

$$
P(\Lambda_t \neq \Lambda'_t) = P(\Lambda'_t \neq \Lambda_t, N(t) \geq 1)
= P(\Lambda'_t \neq \Lambda_t, N(t) = 1) + P(\Lambda'_t \neq \Lambda_t, N(t) = 2).
$$

So, there is a constant $C > 0$ such that

$$
P(N(t) \geq 2) \leq P(N(\delta) \geq 2) = 1 - e^{-H\delta} - H\delta e^{-H\delta} \leq C\delta^2. \quad (3.20)
$$

On the other hand, by the mutually independence of the Brownian motion $(B_t)$ and the Poisson process $(\mu_t(t))$,

$$
P(\Lambda'_t \neq \Lambda_t, N(t) = 1) = \int_0^t P(\Lambda'_t \neq \Lambda_t, \tau_1 \in ds, \tau_2 > t-s)
= \int_0^t \mathbb{E}\left[\mathbb{E}\left[1_{\left\{\xi_j \notin J \cap \{t\} \cap \{t+1\}\left(1 + \kappa_2\right)\int_0^t \mathbb{E}\left[\left|X_s - X'_s\right|\right]e^{-Hs}e^{-H(t-s)}ds
\leq 2N(N-1)C(1 + \kappa_2)\int_0^t \mathbb{E}\left[\left|X_s - X'_s\right|\right]ds,
\right)\right]^{\mathbb{F}_s}\right]
\leq 2N(N-1)C(1 + \kappa_2)\int_0^t \mathbb{E}\left[\left|X_s - X'_s\right|\right]ds,
\]$$

where $\mathbb{F}_B^s = \sigma(B_r; 0 \leq r \leq s)$, the $\sigma$-algebra generated by B.M. $(B_r)$ up to time $s$. Hence,

$$
P(\Lambda'_t \neq \Lambda_t) \leq C\delta^2 + 2N(N-1)C(1 + \kappa_2)\int_0^t \mathbb{E}\left[\left|X_s - X'_s\right|\right]ds, \quad 0 < t \leq \delta. \quad (3.21)
$$

Note that the estimate is independent of the common initial value of $(\Lambda'_t)$ and $(\Lambda_t)$. Hence, by the same method, we can get that

$$
P(\Lambda'_t \neq \Lambda'_t | \Lambda'_t = \Lambda_t) \leq C\delta^2 + 2N(N-1)C(1 + \kappa_2)\int_0^\delta \mathbb{E}\left[\left|X_s - X'_s\right|\right]ds.
$$

Thus,

$$
P(\Lambda'_t \neq \Lambda'_t) = P(\Lambda'_t \neq \Lambda'_t, \Lambda'_t = \Lambda_t) + P(\Lambda'_t \neq \Lambda'_t, \Lambda'_t \neq \Lambda_t)
\leq P(\Lambda'_t \neq \Lambda'_t | \Lambda'_t = \Lambda_t) + P(\Lambda'_t \neq \Lambda_t)$$

23
\[
\leq 2C\delta^2 + 2N(N-1)C_2(1+\kappa_2) \int_0^{2\delta} \mathbb{E}[|X_s - X'_s|]ds.
\]

Deducing inductively, we obtain that
\[
\mathbb{P}(\Lambda'_{k\delta} \neq \Lambda_{k\delta}) \leq kC\delta^2 + 2N(N-1)C_2(1+\kappa_2) \int_0^{k\delta} \mathbb{E}[|X_s - X'_s|]ds, \quad k \geq 2. \tag{3.22}
\]

By virtue of (3.21) and (3.22), we have that for \(t > 0\),
\[
\int_0^t \mathbb{P}(\Lambda'_s \neq \Lambda_s)ds \\
= \int_0^t \mathbb{P}(\Lambda'_s \neq \Lambda_s, \Lambda'_{s\delta} = \Lambda_{s\delta})ds + \int_0^t \mathbb{P}(\Lambda'_s \neq \Lambda_s, \Lambda'_{s\delta} \neq \Lambda_{s\delta})ds \\
\leq \int_0^t \mathbb{P}(\Lambda'_s \neq \Lambda_s)\mathbb{P}(\Lambda'_{s\delta} = \Lambda_{s\delta})ds + \int_0^t \mathbb{P}(\Lambda'_{s\delta} \neq \Lambda_{s\delta})ds \\
\leq \int_0^t (C\delta^2 + 2N(N-1)C_2(1+\kappa_2) \int_{s\delta}^{s} \mathbb{E}[|X_r - X'_r|]dr)ds + \sum_{k=1}^{K} \mathbb{P}(\Lambda'_{k\delta} \neq \Lambda_{k\delta})\delta \\
\leq C\delta^2 t + 2N(N-1)C_2(1+\kappa_2) \int_0^t \int_{s\delta}^{s\delta} \mathbb{E}[|X_r - X'_r|]drds + \frac{C\delta^2}{2}K(K+1) \\
+ 2N(N-1)C_2(1+\kappa_2)\delta \sum_{k=1}^{K} \int_0^{k\delta} \mathbb{E}[|X_s - X'_s|]ds,
\]

where \(K = \left[ \frac{t}{\delta} \right] + 1\). Letting \(\delta \downarrow 0\), we obtain that
\[
\int_0^t \mathbb{P}(\Lambda'_s \neq \Lambda_s)ds \leq N(N-1)C_2(1+\kappa_2)t \int_0^t \mathbb{E}[|X_s - X'_s|]ds,
\]

which is the desired estimate (3.19). \(\square\)

### 3.2 Dynamic programming principle

In this subsection, we go to establish the dynamic programming principle for the optimal control problem associated with the value function \(V(s, x, i)\). To this end, the key point is to establish a measurable selection of the optimal feedback control relative to the initial values. We adopt the method of Stroock and Varadhan [45] in the study of measurable choices on separable metric space.
For \((s, x, i) \in [0, T] \times \mathbb{R}^d \times \mathcal{S}\), define
\[
\Pi_{s,x,i}^0 = \{ \alpha \in \Pi_{s,x,i}; J(s, x, i, \alpha) = V(s, x, i) \},
\] (3.23)
and \(\mathcal{R}_{s,x,i}\) = \{the distribution of \((X, \Lambda, \mu, \nu)\) in \(\mathcal{Y}\); \((X, \Lambda, \mu, \nu)\) is associated with \(\alpha \in \Pi_{s,x,i}\}\}. Similar to \(\mathcal{R}_{s,x,i}\), define \(\mathcal{R}_{s,x,i}^0\) by replace \(\alpha \in \Pi_{s,x,i}\) there with \(\alpha \in \Pi_{s,x,i}^0\).

It is known that \(\mathcal{P}(\mathcal{Y})\) is a Polish space endowed with \(L^1\)-Wasserstein distance \(W_1\), which is defined as follows: for any \(\tilde{R}_1\) and \(\tilde{R}_2\) in \(\mathcal{P}(\mathcal{Y})\), define
\[
W_{1,Y}(\tilde{R}_1, \tilde{R}_2) = \inf_{\Gamma \in \mathcal{E}(\tilde{R}_1, \tilde{R}_2)} \left\{ \int_{\mathcal{Y} \times \mathcal{Y}} \rho((x, i, \mu, \nu), (x', i', \mu', \nu')) d\Gamma \right\},
\]
where
\[
\rho((x, i, \mu, \nu), (x', i', \mu', \nu')) = |x - x'| + 1_{i \neq i'} + W_1(\mu, \mu') + W_1(\nu, \nu').
\]
As a subset of \(\mathcal{P}(\mathcal{Y})\), \(\mathcal{R}_{s,x,i}^0\) is closed under the metric \(W_{1,Y}\). Analogous to the argument of Theorem 2.2, we can show that \(\mathcal{R}_{s,x,i}^0\) is tight. By Prohorov’s theorem, \(\mathcal{R}_{s,x,i}^0\) is a compact set in \(\mathcal{P}(\mathcal{Y})\).

Denote by \(\text{Comp}(\mathcal{P}(\mathcal{Y}))\) the space of all compact subsets of \(\mathcal{P}(\mathcal{Y})\), and define a metric \(\text{dist}(K_1, K_2)\) between two points \(K_1, K_2\) of \(\text{Comp}(\mathcal{P}(\mathcal{Y}))\) by
\[
\text{dist}(K_1, K_2) = \inf \{ \varepsilon > 0; K_1 \subset K^\varepsilon, K_2 \subset K_1^\varepsilon \}.
\]
Here, for all set \(A \in \text{Comp}(\mathcal{P}(\mathcal{Y}))\),
\[
A^\varepsilon := \{ \tilde{\nu} \in \mathcal{P}(\mathcal{Y}); W_{1,Y}(\tilde{\nu}, \tilde{\mu}) < \varepsilon \text{ for some } \tilde{\mu} \in A \}.
\]

**Proposition 3.3** Assume all the assumptions of Theorem 3.1 hold. Then \(\mathcal{R}^0 : [0, T] \times \mathbb{R}^d \times \mathcal{S} \to \text{Comp}(\mathcal{P}(\mathcal{Y}))\) is Borel measurable. Moreover, there exists a measurable selector \(H\) of \(\mathcal{R}^0\), i.e. \(H(s, x, i) \in \mathcal{R}^0_{s,x,i}\), and \(H : [0, T] \times \mathbb{R}^d \times \mathcal{S} \to \mathcal{P}(\mathcal{Y})\) is Borel measurable.

**Proof.** According to [45, Lemma 12.1.8], it is sufficient to show that for \((s_n, x_n, i) \to (s, x, i)\) as \(n \to \infty\), there exists a subsequence \(R_{n_k} \in \mathcal{R}^0_{s_{n_k}, x_{n_k}, i}\) and \(R_0 \in \mathcal{R}^0_{s,x,i}\) such that \(R_{n_k}\) converges weakly to \(R_0\) as \(k \to \infty\).

Since \((s_n, x_n, i)\) converges to \((s, x, i)\), we can prove the tightness of \((R_n)_{n \geq 1}\) similar to the argument of Theorem 2.2. Then there exists a subsequence \((R_{n_k})_{k \geq 1}\) and \(R_0 \in \mathcal{R}^0_{s,x,i}\) such that \(R_{n_k}\) converges weakly to \(R_0\) as \(k \to \infty\), and \(R_0\) is the joint distribution of
\((X^{(0)}, \Lambda^{(0)}, \mu^{(0)}, \nu^{(0)})\). By the continuity of the cost functions \(f\) and \(g\), using Theorem 3.1, we have

\[
V(s, x, i) = \lim_{k \to \infty} V(s_{n_k}, x_{n_k}, i)
\]

\[
= \lim_{k \to \infty} \mathbb{E} \left[ \int_{s_{n_k}}^{T} f(t, X_t^{(n_k)}, \Lambda_t^{(n_k)}, \mu_t^{(n_k)}, \nu_t^{(n_k)}) \, dt + g(X_T^{(n_k)}) \right]
\]

\[
\geq \mathbb{E} \left[ \int_{s}^{T} f(t, X_t^{(0)}, \Lambda_t^{(0)}, \mu_t^{(0)}, \nu_t^{(0)}) \, dt + g(X_T^{(0)}) \right]
\]

\[
\geq V(s, x, i).
\]

Therefore, \(R_0\) belongs to \(\mathcal{R}^0_{s,x,i}\) and this yields immediately \((s, x, i) \mapsto \mathcal{R}^0_{s,x,i}\) is measurable. Moreover, according to [45, Theorem 12.1.10], there exists a measurable selector \(H : [0, T] \times \mathbb{R}^d \times \mathcal{S} \to \mathcal{P}(\mathcal{Y})\) such that \(H(s, x, i) \in \mathcal{R}^0_{s,x,i}\). \(\square\)

To proceed, we adopt the method and notation of [22] to establish the dynamic programming principle. According to [22, Lemma 3.3, Corollary 3.9], under the help of the selection theorem established in Proposition 3.3, the following result holds.

**Lemma 3.4** For every \(R \in \mathcal{R}_{s,x,i}\), \(s < t \leq T\), there exists a unique probability measure on \(\mathcal{Y}\), denoted by \(R \otimes_t Q\), such that

1. \(R \otimes_t Q(A) = R(A), \forall A \in \tilde{\mathcal{Y}}_t\).

2. The regular conditional probability distribution of \(R \otimes_t Q\) with respect to \(\tilde{\mathcal{Y}}_t\) is \(Q_t\), where

\[
Q_t = H(t, X_t, \Lambda_t),
\]

and \(H\) is given by Proposition 3.3.

3. \(R \otimes_t Q\) is the distribution of the process \((X, \Lambda, \mu, \nu)\) associated with some \(\alpha = (\mu, \nu) \in \Pi_{s,x,i}\).

**Theorem 3.5** Assume all the conditions of Theorem 3.1 are still valid. Then for \(s < t \leq T\),

\[
V(s, x, i) = \inf \left\{ \mathbb{E} \left[ \int_{s}^{t} f(r, X_r, \Lambda_r, \mu_r) \, dr + V(t, X_t, \Lambda_t) \right] ; \alpha \in \Pi_{s,x,i} \right\},
\]

(3.24)
Proof. Let \( \alpha \in \Pi_{s,x,i} \) and denote by \( R \) the distribution of \((X, \Lambda, \mu, \nu)\) in \( \mathcal{V} \) associated with \( \alpha \). By Lemma 3.4, there exists an \( \tilde{\alpha} = (\tilde{\mu}, \tilde{\nu}, s, x, i) \in \Pi_{s,x,i} \) associated with \( R \otimes_t Q \). Then,

\[
V(s, x, i) \\
\leq \mathbb{E} \left[ \int_s^T f(r, \tilde{X}_r, \tilde{\Lambda}_r, \tilde{\mu}_r, \tilde{\nu}_r)dr + g(\tilde{X}_T) \right] \\
= \mathbb{E} \left[ \int_s^t f(r, \tilde{X}_r, \tilde{\Lambda}_r, \tilde{\mu}_r, \tilde{\nu}_r)dr + \int_t^T f(r, \tilde{X}_r, \tilde{\Lambda}_r, \tilde{\mu}_r, \tilde{\nu}_r)dr + g(\tilde{X}_T) \right] \\
= \mathbb{E} \left[ \int_s^t f(r, X_r, \Lambda_r, \mu_r, \nu_r)dr + \mathbb{E} \left[ \int_t^T f(r, \tilde{X}_r, \tilde{\Lambda}_r, \tilde{\mu}_r, \tilde{\nu}_r)dr + g(\tilde{X}_T) \bigg| \mathcal{F}_t \right] \right] \\
= \mathbb{E} \left[ \int_s^t f(r, X_r, \Lambda_r, \mu_r, \nu_r)dr + V(\tau, X_\tau, \Lambda_\tau) \right].
\]

In the second equality of the previous equation, we have used that before \( t \), \( \tilde{\alpha} \) coincides with \( \alpha \), and after \( t \), coincides with the measurable selector \( H(t, X_t, \Lambda_t) \). The arbitrariness of \( \alpha \in \Pi_{s,x,i} \) yields that

\[
V(s, x, i) \leq \inf \left\{ \mathbb{E} \left[ \int_s^t f(r, X_r, \Lambda_r, \mu_r, \nu_r)dr + V(t, X_t, \Lambda_t) \right] ; \ \alpha \in \Pi_{s,x,i} \right\}. \tag{3.25}
\]

On the other hand, by Theorem 2.2, there exists an optimal admissible control \( \alpha^* = (\mu^*, \nu^*, s, x, i) \in \Pi_{s,x,i} \). Denote by \((X^*, \Lambda^*)\) the processes associated with \( \alpha^* \). Then,

\[
V(s, x, i) = \mathbb{E} \left[ \int_s^T f(t, X^*_t, \Lambda^*_t, \mu^*_t, \nu^*_t)dt + g(X^*_T) \right] \\
= \mathbb{E} \left[ \int_s^t f(r, X^*_r, \Lambda^*_r, \mu^*_r, \nu^*_r)dr + \int_t^T f(r, X^*_r, \Lambda^*_r, \mu^*_r, \nu^*_r)dr + g(X^*_T) \right] \\
\geq \mathbb{E} \left[ \int_s^t f(r, X^*_r, \Lambda^*_r, \mu^*_r, \nu^*_r)dr + V(t, X^*_t, \Lambda^*_t) \right] \tag{3.26}\\n\geq \inf \left\{ \mathbb{E} \left[ \int_s^t f(r, X_r, \Lambda_r, \mu_r)dr + V(t, X_t, \Lambda_t) \right] ; \ \alpha \in \Pi_{s,x,i} \right\}.
\]

Consequently, the dynamic programming principle (3.24) has been established following from (3.25) and (3.26). \( \square \)
Appendix

We shall provide an argument on the existence and uniqueness of strong solutions for the controlled system \((2.4)\) and \((2.5)\) under the admissible control \(\alpha \in \Pi_{s,x,i}\).

**Proposition 3.6** Assume (H1)-(H4) hold, then for each \(\alpha \in \Pi_{s,x,i}\) the controlled system \((2.4)\) and \((2.5)\) admits a unique non-explosive strong solution.

**Proof.** For \(\alpha \in \Pi_{s,x,i}\), there exist \(F\) and \(G\) in \(\Upsilon_\psi\) such that
\[
\mu_t = F(t, \Xi_{[s,t]}, \Lambda_t), \quad \nu_t = G(t, \Xi_{[s,t]}, \Lambda_t), \quad t \in (s,T].
\]
Rewrite
\[
\tilde{b}(t, X_{[s,t]}, \Lambda_t) = b(X_t, \Lambda_t, F(t, \Xi_{[s,t]}, \Lambda_t)), \quad \tilde{\sigma}(t, X_{[s,t]}, \Lambda_t) = \sigma(X_t, \Lambda_t, G(t, \Xi_{[s,t]}, \Lambda_t)).
\]
Then, \((X_t, \Lambda_t)\) satisfies a stochastic functional differential equation (SFDE) with continuous coefficients under the conditions (H1)-(H3) and \(F, G \in \Upsilon_\psi\). Namely,
\[
\begin{align*}
\text{d}X_t &= \tilde{b}(t, X_{[s,t]}, \Lambda_t)\text{d}t + \tilde{\sigma}(t, X_{[s,t]}, \Lambda_t)\text{d}B_t, \\
\text{d}\Lambda_t &= \int_{[0,H]} \vartheta(t, \Lambda_{t-}, G(t, \Xi_{[s,t]}, \Lambda_{t-}), z)N_1(\text{d}t, \text{d}z),
\end{align*}
\]
where \(\vartheta(x, i, \nu, z) = \sum_{\ell \in S} (\ell - i)1_{\Gamma_{i\nu}(x,\nu)}(z)\). The proof of the existence of a weak solution of \((3.27)\) is standard (see, e.g. [35, Theorem 4.2] for SFDEs without switching, and see [38, Theorem 2.3] for the technique to deal with the switching). According to the Yamada-Watanabe principle, we only need to verify the pathwise uniqueness to show the existence of the unique strong solution.

Let \((X'_t, \Lambda'_t)\) be another solution of \((3.27)\) with initial value \((X'_s, \Lambda'_s) = (x, i) = (X_s, \Lambda_s)\). Then
\[
\begin{align*}
\text{d}(X_t - X'_t) &= \left(\tilde{b}(t, X_{[s,t]}, \Lambda_t) - \tilde{b}(t, X'_{[s,t]}, \Lambda'_t)\right)\text{d}t \\
&\quad + \left(\tilde{\sigma}(t, X_{[s,t]}, \Lambda_t) - \tilde{\sigma}(t, X'_{[s,t]}, \Lambda'_t)\right)\text{d}B(t), \\
\text{d}(\Lambda_t - \Lambda'_t) &= \int_{[0,H]} \left(\vartheta(X_t, \Lambda_{t-}, G(t, \Xi_{[s,t]}, \Lambda_{t-}), z) \\
&\quad - \vartheta(X'_t, \Lambda'_{t-}, G(t, \Xi_{[s,t]}, \Lambda'_{t-}), z)\right)N_1(\text{d}t, \text{d}z).
\end{align*}
\]
Recall the construction of the Poisson random measure $N_1(dt, dz)$ in the argument of Lemma 3.2, and similarly we can define the jumping time $\zeta_1, \zeta_2, \ldots$ after the Poisson process after time $s$. For $s < t < \zeta_1$, it holds $\Lambda_t = \Lambda_t' = i$, and
\[
\begin{align*}
\begin{aligned}
d(X_t - X_t') &= \left(\tilde{b}(t, X_{[s,t]}, i) - \tilde{b}(t, X_{[s,t]}, i')\right)dt + \left(\tilde{\sigma}(t, X_{[s,t]}, i) - \tilde{\sigma}(t, X_{[s,t]}, i')\right)dB_t.
\end{aligned}
\end{align*}
\]
By (H1)-(H3), the fact $F, G \in \Psi$, and Burkholder-Davis-Gundy’s inequality, we obtain
\[
\begin{align*}
\mathbb{E} \sup_{s \leq r < t \wedge \zeta_1} |X_r - X_r'|^2 &\leq C \int_s^t \mathbb{E}\left[ \sup_{s \leq u < r \wedge \zeta_1} |X_u - X_u'|^2 \right]dr \\
&+ C\mathbb{E}\left[ \int_s^t 2\langle X_u - X_u', (\tilde{\sigma}(u, X_{[s,u]}, i) - \tilde{\sigma}(u, X_{[s,u]}, i'))dB_u \rangle \right] \\
&\leq C \int_s^t \mathbb{E}\left[ \sup_{s \leq u < r \wedge \zeta_1} |X_u - X_u'|^2 \right]dr \\
&+ C\mathbb{E}\left[ \int_s^t 1 \left( \int_s^t \sup_{s \leq u < r \wedge \zeta_1} |X_u - X_u'|^2 dr \right) \right]^{\frac{1}{2}} \\
&\leq C \int_s^t \mathbb{E}\left[ \sup_{s \leq u < r \wedge \zeta_1} |X_u - X_u'|^2 \right]dr \\
&+ \frac{1}{2} \mathbb{E}\sup_{s \leq r < t \wedge \zeta_1} |X_r - X_r'|^2 + C\mathbb{E}\left[ \int_s^t \sup_{s \leq u < r \wedge \zeta_1} |X_u - X_u'|^2 dr \right],
\end{align*}
\]
where $t \wedge \zeta_1 = \min\{t, \zeta_1\}$, and $C$ is a positive constant whose value may be different from line to line. By Gronwall’s inequality, this leads to
\[
\mathbb{E} \sup_{s \leq r < t \wedge \zeta_1} |X_r - X_r'|^2 \leq 0,
\]
which means that $X_t \equiv X_t'$ for all $t \in [s, \zeta_1]$ almost surely. By virtue of the continuity of the paths of $t \mapsto X_t$ and $t \mapsto X_t'$, we further get
\[
X_t \equiv X_t', \quad \forall t\in[s, \zeta_1] \text{ a.s.}
\]
Invoking equation (3.27),
\[
\Lambda_{\zeta_1} = i + \int_s^{\zeta_1} \int_{[0,H]} \psi(X_r, i, G(r, X_{[s,r]}t), z) N_1(dr, dz)
\]
\[i + \int_{s}^{\zeta_1} \int_{[0,H]} \vartheta(X'_t, i, G(r, \Xi_{(s,r)}), z) N_1(dr, dz) \]
\[= \Lambda'_{\zeta_1}.\]

Therefore, we have proved that \((X_t, \Lambda_t) \equiv (X'_t, \Lambda'_t)\) for \(t \in [s, \zeta_1]\). By repeating this procedure, we can show \((X_t, \Lambda_t) \equiv (X'_t, \Lambda'_t)\) for \(t \in [\zeta_{k-1}, \zeta_k], k \geq 2\), a.s.. Since condition (H2) ensures that there is only finite number of jumps for the Poisson random measure \(N_1(dt, dz)\) during \([0, T]\), we further obtain that \((X_t, \Lambda_t) \equiv (X'_t, \Lambda'_t)\) for \(t \in [s, T]\) a.s.. □

References

[1] L. Ambrosio, N. Gigli, G. Savaré, Gradient flows in metric spaces and in the space of probability measures, Lectures in Mathematics ETH Zürich. Birkhäuser Verlag, Basel, 2005.

[2] J. Bao, J. Shao, Permanence and extinction of regime-switching predator-prey models, SIAM J. Math. Anal. 48 (2016), 725-739.

[3] P. Billingsley, convergence of probability measures, John Wiley, New York, 1968.

[4] J.M. Bismut, Théorie probabiliste du contrôle des diffusions, Mem. Amer. Math. Soc. 4, 1976.

[5] V. Bogachev, Measure theory, Vol. II, Springer-Verlag Berlin Heidelberg, 2007.

[6] M.-F. Chen, From Markov chains to non-equilibrium particle systems, 2nd ed. Singapore: World Scientific, 2004.

[7] J. Daleckij, M. Krein, Stability of solutions of differential equations in Banach space, Amer. Math. Soc. Transl. 43, 1974.

[8] M.H.A. Davis, On the existence of optimal policies in stochastic control, SIAM J. Control Optim. 11 (1973), 587-594.

[9] N.H. Du, H.N. Dang, G. Yin, Conditions for permanence and ergodicity of certain stochastic predator-prey models, J. Appl. Probab. 53 (2015), 187-202.

[10] S.N. Ethier, T.G. Kurtz, Markov processes characterization and convergence, Wiley, New York, 1986.
[11] W. Feller, On the integro-differential equations of purely discontinuous Markoff processes. Trans. Amer. Math. Soc. 48 (1940), 488-515.

[12] W.H. Fleming, M. Nisio, On the existence of optimal stochastic controls, J. Math. and Mechanics, 15 (1966), 777-794.

[13] W.H. Fleming, R.W. Rishel, Deterministic and stochastic optimal control, Springer-Verlag Berlin Heidelberg New York, 1975.

[14] X.P. Guo, Continuous-time Markov decision processes with discounted rewards: the case of Polish spaces, Math. Oper. Res. 32 (2007), 73-87.

[15] X.P. Guo, O. Hernández-Lerma, Continuous-time controlled Markov chains, Ann. Appl. Prob. 13 (2003), 363-388.

[16] X.P. Guo, O. Hernández-Lerma, Continuous-time Markov decision processes. Theory and applications, Springer-Verlag, Berlin, 2009.

[17] X.P. Guo, M. Vykertas, Y. Zhang, Absorbing continuous-time Markov decision processes with total cost criteria, Adv. Appl. Prob. 45 (2003), 490-519.

[18] U.G. Haussmann, General necessary conditions for optimal control of stochastic systems, Math. Programming Study, 6 (1976), 30-48.

[19] U.G. Haussmann, On the stochastic maximum principle, SIAM J. Control Optim. 16 (1978), 236-251.

[20] U.G. Haussmann, J.P. Lepeltier, On the existence of optimal control, SIAM J. Control Optim. 28 (1990), 851-902.

[21] U.G. Haussmann, W. Suo, Singular optimal stochastic controls I: existence, SIAM J. Control Optim. 33 (1995), 916-936.

[22] U.G. Haussmann, W. Suo, Singular optimal stochastic controls II: Dynamic programming, SIAM J. Control Optim. 33 (1995), 937-959.

[23] T. Hou, J. Shao, Heavy tail and light tail of Cox-Ingersoll-Ross processes with regime-switching, to appear in Sci. China Math. 2019 or arXiv:1709.01691

[24] N. Ikeda, S. Watanabe, Stochastic differential equations and diffusion processes, 2nd ed. North-Holland Mathematical Library. 24, 1989.
[25] M. Kitaev, V.V. Rykov, Controlled queueing systems, CRC Press, Boca Raton, FL. 1995.

[26] H.J. Kushner, On the stochastic maximum principle: fixed time of control, J. Math. Anal. Appl. 11 (1965), 78-92.

[27] H.J. Kushner, Existence results for optimal stochastic controls, J. Optim. Theory Appl., 15 (1975), 347-359.

[28] H.J. Kushner, Necessary conditions for continuous parameter stochastic optimization problems, SIAM J. Control Optim. 10 (1972), 550-565.

[29] H.J. Kushner, A partial history of the early development of continuous-time nonlinear stochastic systems theory, Automatica J. IFAC 50 (2014), no. 2, 303-334.

[30] A. Linquist, On feedback control of linear stochastic systems, SIAM J. Control 11 (1973), 323-343.

[31] Q. Lü, T. Wang, X. Zhang, Characterization of optimal feedback for stochastic linear quadratic control problems, Probability, Uncertainty and Quantitative Risk, (2017), 2:11.

[32] X. Mao, Stabilization of continuous-time hybrid stochastic differential equations by discrete time feedback control, Automatica J. IFAC, 49 (2013), 3677-3681.

[33] X. Mao, C. Yuan, Stochastic Differential Equations with Markovian Switching, Imperial College Press, London, 2006.

[34] S. Peng, A general stochastic maximum principle for optimal control problems, SIAM J Control Optim. 28 (1990), 966-979.

[35] J. Shao, F. Wang, C. Yuan, Harnack inequalities for stochastic (functional) differential equations with non-Lipschitzian coefficients, Electron. J. Probab. 17 (2012), no. 100, 1-18.

[36] J. Shao, Criteria for transience and recurrence of regime-switching diffusion processes, Electron. J. Probab. 20 (2015), no. 63, 1-15.

[37] J. Shao, Ergodicity of regime-switching diffusions in Wasserstein distances, Stoch. Proc. Appl. 125 (2015), 739-758.
[38] J. Shao, Strong solutions and strong Feller properties for regime-switching diffusion processes in an infinite state space, SIAM J. Control Optim. 53 (2015), 2462-2479.

[39] J. Shao, Invariant measures and Euler-Maruyama’s approximations of state-dependent regime-switching diffusions. SIAM J. Control Optim. 56 (2018), no. 5, 3215-3238.

[40] J. Shao, F. Xi, Stability and recurrence of regime-switching diffusion processes, SIAM J. Control Optim. 52 (2014), 3496-3516.

[41] J. Shao, K. Zhao, The existence of optimal control for continuous-time Markov decision processes in random environments, submitted 2019.

[42] A. Skorokhod, Asymptotic Methods in the Theory of Stochastic Differential Equations, American Mathematical Society, Providence, RI. 1989.

[43] Q. Song, R. Stockbridge, C. Zhu, On optimal harvesting problems in random environments, SIAM J. Control Optim. 49 (2011), 859-889.

[44] Q. Song, C. Zhu, On singular control problems with state constraints and regime-switching: a viscosity solution approach, Automatica J. IFAC 70 (2016), 66-73.

[45] D.W. Stroock, S.R.S. Varadhan, Multidimensional diffusion processes, Springer-Verlag, New York, 1979.

[46] C. Villani, Optimal transport, old and new, Grundlehren der mathematischen Wissenschaften, vol. 338, Springer Berlin Heidelberg, 2009.

[47] X. Zhang, T. Siu, Q. Meng, Portfolio selection in the enlarged Markovian regime-switching market. SIAM J. Control Optim. 48 (2009/10), 3368-3388.

[48] H. Zhang, X. Zhang, Second-order necessary conditions for stochastic optimal control problems, SIAM Review, 60 (2018), 139-178.

[49] F. Xi, C. Zhu, On Feller and strong Feller properties and exponential ergodicity of regime-switching jump diffusion processes with countable regimes. SIAM J. Control Optim. 55 (2017), 1789-1818.

[50] G. Yin, X.Y. Zhou, Markowitz’s mean-variance portfolio selection with regime-switching: A continuous-time model, SIAM J. Control Optim., 42 (2003), 1466-1482.
[51] G. Yin, C. Zhu, Hybrid switching diffusions: properties and applications, Vol. 63, Stochastic Modeling and Applied Probability, Springer, New York. 2010.