STABILIZATION OF REGIME-SWITCHING PROCESSES BY FEEDBACK CONTROL BASED ON DISCRETE TIME OBSERVATIONS II: STATE-DEPENDENT CASE *

JINGHAI SHAO† AND FUBAO XI‡

Abstract. This work investigates the almost sure stabilization of a class of regime-switching systems based on discrete-time observations of both continuous and discrete components. It develops Shao's work [SIAM J. Control Optim., 55(2017), pp. 724–740] in two aspects: first, to provide sufficient conditions for almost sure stability in lieu of moment stability; second, to investigate a class of state-dependent regime-switching processes instead of state-independent ones. To realize these developments, we establish an estimation of the exponential functional of Markov chains based on the spectral theory of linear operator. Moreover, through constructing order-preserving coupling processes based on Skorokhod's representation of jumping process, we realize the control from up and below of the evolution of state-dependent switching process by state-independent Markov chains.

Key words. Stability, Regime-switching, State-dependent, Feedback control, Discrete-time observations

AMS subject classifications. 60H10, 93D15, 60J10

1. Introduction.

This work is concerned with the stability of the following regime-switching process:

$$\mathrm{d}X(t) = \big[a(X(t),\Lambda(t)) - b(\Lambda(\delta(t)))X(\delta(t))\big]\mathrm{d}t + \sigma(X(t),\Lambda(t))\mathrm{d}W(t), \tag{1.1}$$

where $\delta(t) = [t/\tau]\tau$, $[t/\tau]$ denotes the integer part of the number t/τ, τ is a positive constant, and $(W(t))$ is a d-dimensional Wiener process. Here $(\Lambda(t))$ is a continuous time jumping process on $\mathcal{S} = \{1, 2, \ldots, M\}$, $M < +\infty$, satisfying

$$\mathbb{P}(\Lambda(t+\Delta) = j | \Lambda(t) = i,\, X(t) = x) = \begin{cases} q_{ij}(x)\Delta + o(\Delta), & \text{if } i \neq j, \\ 1 + q_{ii}(x)\Delta + o(\Delta), & \text{if } i = j, \end{cases} \tag{1.2}$$

provided $\Delta \downarrow 0$, where $0 \leq q_{ij}(x) < +\infty$ for all $i, j \in \mathcal{S}$ with $i \neq j$. As usual, we assume that for each $x \in \mathbb{R}^d$, the Q-matrix $Q_x = \big(q_{ij}(x)\big)$ is conservative; namely $q_i(x) := -q_{ii}(x) = \sum_{j\in\mathcal{S}\setminus\{i\}} q_{ij}(x)$ for all $i \in \mathcal{S}$. Equation (1.1) is a type of stochastic functional differential equation. In current work we shall provide sufficient conditions to ensure the almost sure stability of the system (1.1) and (1.2).

Regime-switching processes have drawn much attention due to the demand of modeling, analysis and computation of complex dynamical systems, and have been widely used in mathematical finance, engineering, biology etc (see, e.g. the monographs [16, 32]). Compared with the classical stochastic processes without switching, regime-switching processes

*Supported in part by the National Natural Science Foundation of China (Nos. 11671034, 11771327, 11431014, 11831014)

†Center for Applied Mathematics, Tianjin University, Tianjin 300072, China (shaojh@tju.edu.cn)

‡School of Mathematics, Beijing Institute of Technology, Beijing 100081, China. (xifb@bit.edu.cn)

can reflect the random change of the environment in which the concerning system lives. Then there are many new difficulties and phenomena appeared in the study of regime-switching processes. See, for instance, [3, 12, 16, 24, 31, 32] and references therein on the stability of such system; [4, 7, 18, 24, 19, 20] on the recurrence of such system; [8, 2, 10] on the heavy or light tail behavior of the invariant probability measure of such system. Besides, there are some literature on the regime-switching processes driven by Lévy processes [5, 27, 30, 31]. Recently, there are also some studies on regime-switching stochastic functional differential equations, e.g. [1, 14, 15, 22, 33, 17].

Our motivation to study the equation (1.1) is to stabilize an unstable system (1.1) with $b = 0$ based on discrete time observations of $(X(t))$ and $(\Lambda(t))$. Such stabilization problem for regime-switching processes was first raised by Mao [14] for the sake of saving cost and being more realistic. There, Mao investigated the mean-square stability of the following controlled system:

$$\mathrm{d}X(t) = (a(X(t), \Lambda(t)) - b(X(\delta(t), \Lambda(t))))\mathrm{d}t + \sigma(X(t), \Lambda(t))\mathrm{d}W(t),$$

where $(\Lambda(t))$ is a continuous time Markov chain independent of the Wiener process $(W(t))$. Subsequently, many works were devoted to developing this stabilization problem. See, for example, [15, 33]. Especially, in [33], some sufficient conditions were provided to ensure this system to be almost surely stable. Inspired by these works, [22] investigated the stability of such kind of system not only based on discrete time observations of $(X(t))$ but also according to discrete time observations of $(\Lambda(t))$. This needs to overcome the essential difference between the path property of $(X(t))$ and $(\Lambda(t))$. For the continuous process $(X(t))$, since $X(t-\tau)$ tends to $X(t)$ as $\tau \to 0$, the difference between $X(t)$ and $X(t-\tau)$ can be controlled when τ is sufficiently small. However, for the jumping process $(\Lambda(t))$, if a jump happens at time t, then $\Lambda(t)$ and $\Lambda(t-) := \lim_{s\uparrow t}\Lambda(s)$ is different, and there is no way to control the difference between $\Lambda(t-\tau)$ and $\Lambda(t)$ by letting τ being sufficiently small. Therefore, in [22] Shao takes advantage of the independence of $(W(t))$ and $(\Lambda(t))$ to control the evolution of $(X(t))$ through the long time behavior of $(\Lambda(t))$ and $(\Lambda(n\tau))_{n\geq 0}$. Precisely, it was shown that

$$\mathbb{E}|X(t)|^2 \leq |X(0)|^2\mathbb{E}\Big[\mathrm{e}^{\int_0^t f(\Lambda(r))+K_\tau g(\Lambda(\delta(r)))\mathrm{d}r}\Big], \tag{1.3}$$

where $f, g : \mathcal{S} \to \mathbb{R}$, K_τ is a constant related to τ. As an embedded Markov chain, $(\Lambda(n\tau))_{n\geq 0}$ has the same stationary distribution as that of $(\Lambda(t))$. However, as mentioned in [22, Remark 3.3], the following kind of quantity cannot be handled at that time in order to show the almost sure stability

$$\mathbb{E}\int_0^\infty \mathrm{e}^{\int_0^t g(\Lambda(\delta(s)))\mathrm{d}s}\mathrm{d}t < \infty,$$

which could be dealt with in current work under the help of spectral theory of linear operator (see Lemma 2.3 below).

In this work, we will overcome two difficulties to establish the almost sure stability of $(X(t), \Lambda(t))$ given by (1.1) and (1.2). First, via the spectral theory of linear operators,

we provide estimates from upper and lower the exponential functional of Markov chain $(Y_n)_{n\geq 0}$. Second, through using Skorokhod's representation of jumping processes or constructing order-preserving coupling, we can control the evolution of state-dependent jumping process $(\Lambda(t))$ by some auxiliary state-independent Markov chains. Moreover, to ensure the existence of the system $(X(t), \Lambda(t))$ satisfying (1.1) and (1.2), and the existence of order-preserving couplings, a general result on the existence of regime-switching stochastic functional differential equation is established. In particular, we only assume that $x \mapsto q_{ij}(x)$ is continuous and $q_i(x)$ is of polynomial growth for $i, j \in \mathcal{S}$ (see Theorem 2.1 below for details).

The remainder of this paper is arranged as follows: Section 2 presents some necessary preparation results concerning the existence of solution for state-dependent regime-switching stochastic functional differential equations, estimate of exponential functional of Markov chains, and constructing auxiliary Markov chains to control the evolution of state-dependent jumping process $(\Lambda(t))$. Section 3 studies the almost sure stability for a class of regime-switching systems based on discrete-time observations. Using the technique used in [33], we can prove our main result, Theorem 3.4, of this work. Finally, the proof of the existence and uniqueness of solution for regime-switching stochastic functional differential equations is appended in Appendix A.

2. Preliminary results. Let us begin this section with the existence and uniqueness of above system (1.1) and (1.2) which can be viewed as a regime-switching stochastic functional differential equations (SFDEs). Here we collect the conditions used in this work on the coefficients of (1.1) and transition rate matrix $(q_{ij}(x))$. We present a result on the existence and uniqueness of solution for a more general regime-switching SFDE:

$$\mathrm{d}X(t) = \big[a(X(t), \Lambda(t)) - b(X(\delta(t)), \Lambda(\delta(t)))\big]\mathrm{d}t + \sigma(X(t), \Lambda(t))\mathrm{d}W(t). \tag{2.1}$$

Suppose the coefficients $a(\cdot,\cdot) : \mathbb{R}^d \times \mathcal{S} \to \mathbb{R}^d$, $b(\cdot,\cdot) : \mathbb{R}^d \times \mathcal{S} \to \mathbb{R}^d$ and $\sigma(\cdot,\cdot) : \mathbb{R}^d \times \mathcal{S} \to \mathbb{R}^{d\times d}$ satisfy the following conditions.

(**H1**) There exist nonnegative functions $C(\cdot)$ and $c(\cdot)$ on $\mathcal{S}$ such that

$$c(i)|x|^2 \leq 2\langle a(x,i), x\rangle + \|\sigma(x,i)\|_{\mathrm{HS}}^2 \leq C(i)|x|^2, \quad (x,i) \in \mathbb{R}^d \times \mathcal{S},$$

where $\|\sigma(x,i)\|_{\mathrm{HS}}^2 = \mathrm{trace}(\sigma\sigma^*)(x,i)$, and σ^* denotes the transpose of the matrix σ.

(**H2**) There exists a positive constant $\bar{K}$ such that

$$|a(x,i)-a(y,i)|+|b(x,i)-b(y,i)|+\|\sigma(x,i)-\sigma(y,i)\|_{\mathrm{HS}} \leq \bar{K}|x-y|,\ x,y \in \mathbb{R}^d,\ i \in \mathcal{S}.$$

Moreover, let the Q-matrix $Q_x = \big(q_{ij}(x)\big)$ satisfy the following conditions:

(**Q1**) $x \mapsto q_{ij}(x)$ is continuous for every $i, j \in \mathcal{S}$.

(**Q2**) $H := \sup_{x\in\mathbb{R}^d} \max_{i\in\mathcal{S}} q_i(x) < \infty$.

The condition (**Q2**) is used in the control of the evolution of $(\Lambda(t))$ through Markov chains. If only for the aim of existence and uniqueness of the dynamical system $(X(t), \Lambda(t))$, we can use a weaker condition as follows:

(**Q2'**) $q_i(x) = \sum_{j\in\mathcal{S}\setminus\{i\}} q_{ij}(x) \le K_0(1+|x|^{\kappa_0})$ for every $(x,i)\in\mathbb{R}^d\times\mathcal{S}$, where K_0 and κ_0 are positive constants.

THEOREM 2.1. *Assume conditions (**H1**), (**H2**), (**Q1**) and (**Q2'**) hold. Then there exists a unique nonexplosive solution (X,Λ) to (2.1) and (1.2).*

In order to preserve the flow of presentation, we defer the proof of Theorem 2.1 to Appendix A. We provide a very explicit construction of the solution (X,Λ) to regime-switching SFDE (2.1) and (1.2) and prove the nonexplosiveness of the solution. Compared with the corresponding results in [21, 22] where $x\mapsto q_{ij}(x)$ is assumed to be Lipschitzian, here we only suppose $x\mapsto q_{ij}(x)$ to be continuous. This greatly simplifies the conditions to be verified so that the coupling process constructed below exists. On the other hand, contrary to the usual boundedness assumption (**Q2**) imposed in the previous works such as [22, 28, 32] etc., here the functions $q_{ij}(x)$ in the Q-matrix $Q_x = \big(q_{ij}(x)\big)$ may be unbounded. Hence the construction of the solution in Theorem 2.1 is of interest by itself.

In a similar way, one can establish the existence and uniqueness of solution for another widely studied SFDE with regime-switching. To do so, we introduce some notations. Let $\mathscr{C}$ denote the continuous path space $C([-r,0];\mathbb{R}^d)$ endowed with the uniform topology, i.e., $\|\xi\|_\infty = \sup_{-r\le s\le 0}|\xi(s)|$ for $\xi\in\mathscr{C}$, where $r\ge 0$ is a constant. Consider the following SFDE:

$$\mathrm{d}X(t) = b(X_t,\Lambda(t),\Lambda(t-r))\mathrm{d}t + \sigma(X_t,\Lambda(t),\Lambda(t-r))\mathrm{d}W(t) \tag{2.2}$$

with $X_0=\xi\in\mathscr{C}$, $\Lambda(0)=i\in\mathcal{S}$, and $(\Lambda(t))$ still satisfies (1.2) as above. Here, $b:\mathscr{C}\times\mathcal{S}\times\mathcal{S}\to\mathbb{R}^d$, $\sigma:\mathscr{C}\times\mathcal{S}\times\mathcal{S}\to\mathbb{R}^{d\times d}$, and $X_t\in\mathscr{C}$ is defined by $X_t(\theta)=X(t+\theta)$ for $\theta\in[-r,0]$. Here we regard that $\Lambda(t-r)=i$ for $t-r<0$ when $\Lambda(0)=i$. The following conditions guarantee the existence and uniqueness of the process (X,Λ) satisfying (2.2) and (1.2).

(**A1**) $b(\cdot,i,j)$ and $\sigma(\cdot,i,j)$ are bounded on bounded subset of $\mathscr{C}$ for every $i,j\in\mathcal{S}$. Moreover, there exists a positive constant K_1 such that

$$2\langle b(\xi,i,j)-b(\eta,i,j),\xi(0)-\eta(0)\rangle + \|\sigma(\xi,i,j)-\sigma(\eta,i,j)\|_{\mathrm{HS}}^2 \le K_1\|\xi-\eta\|_\infty^2$$

for all $\xi,\eta\in\mathscr{C}$, $i,j\in\mathcal{S}$, where $\langle\cdot,\cdot\rangle$ denotes the Euclidean inner product in $\mathbb{R}^d$, $\|\sigma\|_{\mathrm{HS}}^2=\mathrm{trace}(\sigma\sigma^*)$ for $\sigma\in\mathbb{R}^{d\times d}$, σ^* denotes the transpose of σ.

(**A2**) There exists a positive constant K_2 such that $2\langle b(\xi,i,j),\xi(0)\rangle + \|\sigma(\xi,i,j)\|_{\mathrm{HS}}^2 \le K_2(1+\|\xi\|_\infty^2)$ for all $\xi\in\mathscr{C}$ and $i,j\in\mathcal{S}$.

THEOREM 2.2. *Assume conditions (**A1**), (**A2**), (**Q1**) and (**Q2'**) hold. Then there exists a unique nonexplosive solution (X,Λ) to regime-switching SFDE (2.2) and (1.2).*

Theorem 2.2 can be proved by using the same idea of the argument of Theorem 2.1, and hence the proof will be omitted.

In the remainder of this section, we shall present two kinds of preparation results: in the first place, we establish an estimation of exponential functional of a discrete-time Markov chain by the spectrum analysis method; in the second place, we construct two auxiliary Markov chains to control from upper and lower the evolution of the state-dependent jumping process $(\Lambda(t))$.

First, let us consider a time-homogeneous Markov chain $(Y_n)_{n\geq 0}$ on the state space $\mathcal{S}=\{1,\dots,M\}$ with $1<M<\infty$. Denote

$$P_{ij}=\mathbb{P}(Y_1=j|Y_0=i),\qquad i,j\in\mathcal{S}.$$

Assume the transition matrix $P=(P_{ij})$ is a positive matrix, i.e., $P_{ij}>0$, $\forall\, i,j\in\mathcal{S}$. Let $(\theta(i))_{i\in\mathcal{S}}$ be a series of real numbers. Put

$$\widetilde{P}_{ij}=e^{\theta(i)}P_{ij},\qquad i,j\in\mathcal{S},\qquad \widetilde{P}=(\widetilde{P}_{ij})_{i,j\in\mathcal{S}}.$$

Denote $\mathrm{Spec}(\widetilde{P})$ the spectrum of the linear operator $\widetilde{P}$. Let

$$\lambda_1=\max\{\mathrm{Re}(\lambda);\lambda\in\mathrm{Spec}(\widetilde{P})\},$$

where $\mathrm{Re}(\lambda)$ stands for the real part of the eigenvalue λ.

LEMMA 2.3. *Let $\theta:\mathcal{S}\to\mathbb{R}$. Then there exist two positive constants K_3, K_4 such that*

$$K_3\lambda_1^n\leq\mathbb{E}_\mu\Big[\exp\Big\{\sum_{k=0}^{n-1}\theta(Y_k)\Big\}\Big]\leq K_4\lambda_1^n$$

for every initial probability distribution μ of $(Y_n)_{n\geq 0}$ when n is large enough.

Proof. According to the Perron-Frobenius theorem, due to the positivity of $\widetilde{P}$, which follows directly from the positivity of P, λ_1 is a simple eigenvalue of $\widetilde{P}$, and all the magnitudes of other eigenvalues of $\widetilde{P}$ are strictly smaller than λ_1. Invoking the spectral theory for linear operator in a finite dimensional Banach space (cf. Dunford and Schwartz [9, Chapter VII, Theorem 8]), there exists a family of linear operator $\{E(\lambda);\ \lambda\in\mathrm{Spec}(\widetilde{P})\}$ satisfying $E(\lambda)^2=E(\lambda)$, $E(\lambda)E(\tilde{\lambda})=0$ if $\lambda\neq\tilde{\lambda}$, and $I=\sum_{\lambda\in\mathrm{Spec}(\widetilde{P})}E(\lambda)$ such that

$$\widetilde{P}^n=\lambda_1^nE(\lambda_1)+\sum_{\lambda\in\mathrm{Spec}(\widetilde{P})\setminus\{\lambda_1\}}\sum_{i=0}^{v(\lambda)-1}\frac{(\widetilde{P}-\lambda I)^i}{i!}f^{(i)}(\lambda)E(\lambda),\tag{2.3}$$

where $v(\lambda)$ denotes the index of the eigenvalue λ, which is a constant less than $2M$; the function f is given by $f(x)=x^n$ and $f^{(i)}$ denotes the i-th order derivative of f. We can rewrite the terms in the summation as

$$\begin{aligned}\frac{(\widetilde{P}-\lambda I)^i}{i!}f^{(i)}(\lambda)E(\lambda)&=\frac{n!}{i!(n-i)!}\lambda^{n-i}(\widetilde{P}-\lambda I)^iE(\lambda)\\&=\lambda_1^n\frac{n!}{i!(n-i)!}\Big(\frac{\lambda}{\lambda_1}\Big)^{n-i}\Big(\frac{\widetilde{P}-\lambda I}{\lambda_1}\Big)^iE(\lambda).\end{aligned}$$

Note that for any $\lambda\in\mathrm{Spec}(\widetilde{P})$ with $\lambda\neq\lambda_1$, it holds $|\lambda|<\lambda_1$. Therefore, for any fixed $i<n$,

$$\lim_{n\to\infty}\frac{n!}{i!(n-i)!}\Big(\frac{|\lambda|}{\lambda_1}\Big)^{n-i}=0.$$

Consequently, by (2.3), for any initial probability measure μ of (Y_n) on $\mathcal{S}$,

$$\mu\widetilde{P}^n\mathbf{1}=\lambda_1^n\mu E(\lambda_1)\mathbf{1}+\sum_{\lambda\in\mathrm{Spec}(\widetilde{P})\setminus\{\lambda_1\}}\sum_{i=0}^{v(\lambda)-1}\mu\Big[\frac{(\widetilde{P}-\lambda I)^i}{i!}f^{(i)}(\lambda)E(\lambda)\mathbf{1}\Big].$$

Hence, there exist two positive constants K_3, K_4 independent of the initial distribution μ such that

$$K_3\lambda_1^n \le \mu\widetilde{P}^n\mathbf{1} \le K_4\lambda_1^n, \qquad \text{for } n \text{ large enough.}$$

Invoking the definition of $\widetilde{P}$, this can be written in the expectation form as

$$K_3\lambda_1^n \le \mathbb{E}_\mu\Big[\exp\Big\{\sum_{k=0}^{n-1}\theta(Y_k)\Big\}\Big] \le K_4\lambda_1^n, \qquad \text{for } n \text{ large enough,} \tag{2.4}$$

which is just desired conclusion. □

Next, employing the idea of Shao [23], we go to construct two auxiliary continuous-time Markov chains $(\bar{\Lambda}(t))$ and $(\Lambda^*(t))$ such that $\Lambda^*(t) \le \Lambda(t) \le \bar{\Lambda}(t)$, $t \ge 0$, a.s., under some appropriate conditions. Our stochastic comparisons are based on Skorokhod's representation of $(\Lambda(t))$ in terms of the Poisson random measure by following the line of [26, Chapter II-2.1] or [32]. To focus on the idea, we first consider the special situation that $\mathcal{S}$ consists of only two points, i.e., $\mathcal{S} = \{1, 2\}$. To do so, we further assume the following condition holds:

(**Q3**) For each $x \in \mathbb{R}^d$, the Q-matrix $Q_x = \big(q_{ij}(x)\big)$ is irreducible.

According to the conservativeness of Q_x, one has $q_1(x) := -q_{11}(x) = q_{12}(x)$, $x \in \mathbb{R}^d$. For each $x \in \mathbb{R}^d$, let

$$\Gamma_{12}(x) := [0, q_{12}(x)) \qquad \text{and} \qquad \Gamma_{21}(x) := [q_1(x), q_1(x) + q_{21}(x)).$$

Obviously, the length of $\Gamma_{12}(x)$ and $\Gamma_{21}(x)$ is $q_{12}(x)$ and $q_{21}(x)$, respectively. Define a function $\mathbb{R}^d \times \mathcal{S} \times \mathbb{R} \ni (x, i, u) \mapsto h(x, i, u) \in \mathbb{R}$ by

$$h(x, i, u) = (-1)^{1+i}\Big\{\mathbf{1}_{\{i=1\}}\mathbf{1}_{\Gamma_{ii+1}(x)}(u) + \mathbf{1}_{\{i=2\}}\mathbf{1}_{\Gamma_{ii-1}(x)}(u)\Big\}.$$

Then, as in the proof of Theorem 2.1, $(\Lambda(t))$ solves the following stochastic differential equation (SDE for short)

$$\mathrm{d}\Lambda(t) = \int_{[0,L]} h(X(t), \Lambda(t-), u)N(\mathrm{d}t, \mathrm{d}u), \quad t > 0, \quad \Lambda(0) = i_0 \in \mathcal{S}. \tag{2.5}$$

Herein, $L := 2H$ with H being introduced in condition (**Q2**) and $N(\mathrm{d}t, \mathrm{d}u)$ stands for a Poisson random measure with intensity $\mathrm{d}t \times \mathbf{m}(\mathrm{d}u)$, in which $\mathbf{m}(\mathrm{d}u)$ signifies the Lebesgue measure on $[0, L]$. Let $p(t)$ be the stationary Poisson point process corresponding to the Poisson random measure $N(\mathrm{d}t, \mathrm{d}u)$ so that $N([0, t) \times A) = \sum_{s \le t}\mathbf{1}_A(p(s))$ for $A \in \mathscr{B}(\mathbb{R})$.

Due to the finiteness of $\mathbf{m}(\mathrm{d}u)$ on $[0, L]$, there is only finite number of jumps of the process $(p(t))$ in each finite time interval. Let $0 = \zeta_0 < \zeta_1 < \cdots < \zeta_n < \cdots$ be the enumeration of all jumps of $(p(t))$. It holds that $\lim_{n\to\infty}\zeta_n = +\infty$ a.s. From (2.5), it follows that

$$\Lambda(t) = i_0 + \sum_{s \le t} h(X(s), \Lambda(s-), p(s))\mathbf{1}_{[0,L]}(p(s)), \quad t > 0, \quad i_0 \in \mathcal{S},$$

which implies that $(\Lambda(t))$ may have a jump at only ζ_i (i.e. $\Lambda(\zeta_i) \ne \Lambda(\zeta_i-)$) provided that $p(\zeta_i) \in [0, L]$. So the collection of all jumping times of $(\Lambda(t))$ is a subset of $\{\zeta_1, \zeta_2, \cdots\}$. Subsequently, this basic fact will be used frequently without mentioning it again.

Let

$$\bar{q}_{12} := \sup_{x\in\mathbb{R}^d} q_{12}(x),\ \bar{q}_{21} := \inf_{x\in\mathbb{R}^d} q_{21}(x),\ \bar{q}_1 := -\bar{q}_{11} := \bar{q}_{12},\ \bar{q}_2 := -\bar{q}_{22} := \bar{q}_{21}, \tag{2.6}$$

and

$$q^*_{12} := \inf_{x\in\mathbb{R}^d} q_{12}(x),\ q^*_{21} := \sup_{x\in\mathbb{R}^d} q_{21}(x),\ q^*_1 := -q^*_{11} := q^*_{12},\ q^*_2 := -q^*_{22} := q^*_{21}. \tag{2.7}$$

Let

$$\bar{\Gamma}_{12} := [0, \bar{q}_{12}),\ \bar{\Gamma}_{21} := [\bar{q}_{12},\ \bar{q}_{12}+\bar{q}_{21}),\ \Gamma^*_{12} := [0, q^*_{12}),\ \Gamma^*_{21} := [q^*_{12}, q^*_{12}+q^*_{21}).$$

Using the same Poisson random measure $N(\mathrm{d}t, \mathrm{d}u)$ given in (2.5), we define two auxiliary Markov chains $(\bar{\Lambda}(t))$ and $(\Lambda^*(t))$ by the following SDEs:

$$\mathrm{d}\bar{\Lambda}(t) = \int_{[0,L]} \bar{g}(\bar{\Lambda}(t-), u)N(\mathrm{d}t, \mathrm{d}u), \qquad t > 0, \qquad \bar{\Lambda}(0) = \Lambda(0), \tag{2.8}$$

and

$$\mathrm{d}\Lambda^*(t) = \int_{[0,L]} g^*(\Lambda^*(t-), u)N(\mathrm{d}t, \mathrm{d}u), \qquad t > 0, \qquad \Lambda^*(0) = \Lambda(0), \tag{2.9}$$

where, for $i \in \mathcal{S}$,

$$\bar{g}(i,u) = (-1)^{1+i}\Big\{\mathbf{1}_{\{i=1\}}\mathbf{1}_{\bar{\Gamma}_{ii+1}}(u) + \mathbf{1}_{\{i=2\}}\mathbf{1}_{\bar{\Gamma}_{ii-1}}(u)\Big\}, \quad u \in [0, L],$$

and

$$g^*(i,u) = (-1)^{1+i}\Big\{\mathbf{1}_{\{i=1\}}\mathbf{1}_{\Gamma^*_{ii+1}}(u) + \mathbf{1}_{\{i=2\}}\mathbf{1}_{\Gamma^*_{ii-1}}(u)\Big\}, \quad u \in [0, L].$$

Then, according to Skorokhod's representation, $(\bar{\Lambda}_t)$ and (Λ^*_t) are continuous-time Markov chains on $\mathcal{S} = \{1, 2\}$ generated by the Q-matrices $\bar{Q} = (\bar{q}_{ij})_{1\le,i,j\le 2}$ and $Q^* = (q^*_{ij})_{1\le,i,j\le 2}$, respectively.

LEMMA 2.4. *(i) If* $\bar{q}_{21} > 0$ *and*

$$\bar{q}_{12} + \bar{q}_{21} \le q_{12}(x) + q_{21}(x), \qquad x \in \mathbb{R}^d, \tag{2.10}$$

then $\Lambda(t) \le \bar{\Lambda}(t)$ *for all* $t \ge 0$ *a.s.*

(ii) If $q^*_{12} > 0$ *and*

$$q^*_{12} + q^*_{21} \ge q_{12}(x) + q_{21}(x), \qquad x \in \mathbb{R}^d, \tag{2.11}$$

then $\Lambda^*(t) \le \Lambda(t)$ *for all* $t \ge 0$ *a.s.*

Proof. Here we include a proof for the completeness and the ease of the readers. We shall only prove assertion (i) as assertion (ii) can be proved in a similar way. Since there is no jump during the open interval (ζ_k, ζ_{k+1}), we only need to prove (i) at ζ_k, $k \ge 1$. To this aim, we consider separately three different cases.

Case 1: $\Lambda(\zeta_k)=\bar{\Lambda}(\zeta_k)=1$**,** $k\geq 1$**.** In this case, we deduce from (2.5) and (2.8) that

$$\Lambda(\zeta_{k+1})=1+\mathbf{1}_{\Gamma_{12}(X(\zeta_{k+1}))}(p(\zeta_{k+1})) \quad \text{and} \quad \bar{\Lambda}(\zeta_{k+1})=1+\mathbf{1}_{\bar{\Gamma}_{12}}(p(\zeta_{k+1})).$$

According to the notion of $\bar{q}_{12}$, one clearly has $q_{12}(x)\leq \bar{q}_{12}$, $x\in\mathbb{R}^d$, which implies that $\Gamma_{12}(X(\zeta_{k+1}))\subset\bar{\Gamma}_{12}$, a.s. Whence, $\Lambda(\zeta_{k+1})\leq\bar{\Lambda}(\zeta_{k+1})$, a.s.

Case 2: $\Lambda(\zeta_k)=\bar{\Lambda}(\zeta_k)=2$**,** $k\geq 1$**.** Concerning such case, we also obtain from (2.5) and (2.8) that

$$\Lambda(\zeta_{k+1})=2-\mathbf{1}_{\Gamma_{21}(X(\zeta_{k+1}))}(p(\zeta_{k+1})) \quad \text{and} \quad \bar{\Lambda}(\zeta_{k+1})=2-\mathbf{1}_{\bar{\Gamma}_{21}}(p(\zeta_{k+1})). \tag{2.12}$$

If $p_1(\zeta_{k+1})\notin\bar{\Gamma}_{21}$, then, from (2.12), one has $\bar{\Lambda}(\zeta_{k+1})=2$ so that $\Lambda(\zeta_{k+1})\leq\bar{\Lambda}(\zeta_{k+1})$ due to the fact that $\Lambda(\zeta_{k+1})\leq 2$. Next, we proceed to deal with the case $p(\zeta_{k+1})\in\bar{\Gamma}_{21}$, which of course leads to $\bar{\Lambda}(\zeta_{k+1})=1$ in view of (2.12), and $\bar{q}_{12}\leq p(\zeta_{k+1})<\bar{q}_{12}+\bar{q}_{21}$. Employing the assumption (2.10) and utilizing the fact that $q_{12}(X(\zeta_{k+1}))\leq\bar{q}_{12}$, we arrive at $q_{12}(X(\zeta_{k+1}))\leq p(\zeta_{k+1})<q_{12}(X(\zeta_{k+1}))+q_{21}(X(\zeta_{k+1}))$, namely, $p(\zeta_{k+1})\in\Gamma_{12}(X(\zeta_{k+1}))$. As a consequence, $\Lambda(\zeta_{k+1})=\bar{\Lambda}(\zeta_{k+1})=1$.

Case 3: $\Lambda(\zeta_k)=1$**,** $\bar{\Lambda}(\zeta_k)=2$**,** $k\geq 1$**.** For this setup, it follows from (2.5) and (2.8) that

$$\Lambda(\zeta_{k+1})=1+\mathbf{1}_{\Gamma_{12}(X(\zeta_{k+1}))}(p(\zeta_{k+1})) \quad \text{and} \quad \bar{\Lambda}(\zeta_{k+1})=2-\mathbf{1}_{\bar{\Gamma}_{21}}(p(\zeta_{k+1})). \tag{2.13}$$

From (2.13), it is easy to see that $\Lambda(\zeta_{k+1})\leq\bar{\Lambda}(\zeta_{k+1})$ if $p(\zeta_{k+1})\notin\bar{\Gamma}_{21}$. Now, if $p(\zeta_{k+1})\in\bar{\Gamma}_{21}$, then we infer that $\bar{\Lambda}(\zeta_{k+1})=1$ and that $\bar{q}_{12}\leq p(\zeta_{k+1})<\bar{q}_{12}+\bar{q}_{21}$. Hence, one has $p(\zeta_{k+1})\geq q_{12}(X(\zeta_{k+1}))$; in other words, $p(\zeta_{k+1})\notin\Gamma_{12}(X(\zeta_{k+1}))$. As a result, we obtain from (2.13) that $\Lambda(\zeta_{k+1})=\bar{\Lambda}(\zeta_{k+1})=1$.

The desired result follows immediately by summing up the above three cases. □

We now proceed to the general situation that $\mathcal{S}$ could own more than two states, i.e., $\mathcal{S}=\{1,\ldots,M\}$. To this end, we employ the coupling method and especially construct the so-called order-preserving couplings. To do so, we need some preparation.

Let $\Lambda^{(1)}$ and $\Lambda^{(2)}$ be two continuous-time Markov chains defined by two generators $Q^{(1)}=\big(q_{ij}^{(1)}\big)$ and $Q^{(2)}=\big(q_{ij}^{(2)}\big)$ on the state space $\mathcal{S}$, respectively. Note that $Q^{(1)}$ and $Q^{(2)}$ are called Q-matrices in [6, 34, 35]. A continuous-time Markov chain $(\Lambda^{(1)},\Lambda^{(2)})$ on the product space $\mathcal{S}\times\mathcal{S}$ is called a coupling of $\Lambda^{(1)}$ and $\Lambda^{(2)}$, if the following marginality holds for any $t\geq 0$, $i_1,i_2\in\mathcal{S}$ and $B_1,B_2\subset\mathcal{S}$,

$$\begin{aligned}&\mathbb{P}^{(i_1,i_2)}\big((\Lambda^{(1)}(t),\Lambda^{(2)}(t))\in B_1\times\mathcal{S}\big)=\mathbb{P}^{(i_1)}(\Lambda^{(1)}(t)\in B_1),\\&\mathbb{P}^{(i_1,i_2)}\big((\Lambda^{(1)}(t),\Lambda^{(2)}(t))\in\mathcal{S}\times B_2\big)=\mathbb{P}^{(i_2)}(\Lambda^{(2)}(t)\in B_2),\end{aligned} \tag{2.14}$$

where the superscript in $\mathbb{P}^{(i_1,i_2)}$ and $\mathbb{P}^{(i_1)}$ is used to emphasize the initial value of the corresponding process. From [6] we know that constructing a coupling Markov chain $(\Lambda^{(1)},\Lambda^{(2)})$ is equivalent to constructing a coupling generator $\widetilde{Q}$ on the finite state space $\mathcal{S}\times\mathcal{S}$, and such a generator $\widetilde{Q}$ is called a coupling of $Q^{(1)}$ and $Q^{(2)}$. For given two generators (or two Q-matrices), one can construct many their coupling generators (or coupling Q-matrices); see [6] for more examples and explanation. In what follows we are especially interested in

the order-preserving couplings. On the product space $\mathcal{S}\times\mathcal{S}$, an order-preserving coupling $\widetilde{Q}$ of $Q^{(1)}$ and $Q^{(2)}$ means that the corresponding Markov chain generated by $\widetilde{Q}$ satisfies

$$\mathbb{P}^{(i_1,i_2)}\big(\Lambda^{(1)}(t)\le\Lambda^{(2)}(t),\ \forall\, t\ge 0\big)=1,\quad i_1\le i_2\in\mathcal{S}. \tag{2.15}$$

See [6, Chapter 5] for the details about the coupling Q-matrices and related materials. For more general case, the construction of order-preserving couplings was studied in [34, 35]. In particular, we have the following lemma from the aforementioned three references.

LEMMA 2.5. *If the generators $Q^{(1)}$ and $Q^{(2)}$ on $\mathcal{S}$ satisfy that*

$$\begin{aligned}&\sum_{l\ge m}q^{(1)}_{i_1l}\le\sum_{l\ge m}q^{(2)}_{i_2l}\quad \textit{for all } i_1,i_2,m\in\mathcal{S}\ \textit{with } i_1\le i_2<m\quad \textit{and}\\ &\sum_{l\le m}q^{(1)}_{i_1l}\ge\sum_{l\le m}q^{(2)}_{i_2l}\quad \textit{for all } i_1,i_2,m\in\mathcal{S}\ \textit{with } m<i_1\le i_2,\end{aligned} \tag{2.16}$$

there exists an order-preserving coupling Q-matrix $\widetilde{Q}$ on $\mathcal{S}\times\mathcal{S}$ and hence (2.15) holds.

ASSUMPTION 2.6. Assume that there exists a generator $\bar{Q}=(\bar{q}_{i,j})$ on $\mathcal{S}$ such that the following bounds hold:

$$\begin{aligned}&\sup_{x\in\mathbb{R}^d}\sum_{l\ge m}q_{i_1l}(x)\le\sum_{l\ge m}\bar{q}_{i_2l}\ \text{for all } i_1,i_2,m\in\mathcal{S}\ \text{with}\ \ i_1\le i_2<m\ \ \text{and}\\ &\inf_{x\in\mathbb{R}^d}\sum_{l\le m}q_{i_1l}(x)\ge\sum_{l\le m}\bar{q}_{i_2l}\ \text{for all } i_1,i_2,m\in\mathcal{S}\ \text{with}\ \ m<i_1\le i_2,\end{aligned} \tag{2.17}$$

where the matrix $\big(q_{ij}(x)\big)$ is given in (1.2).

If two generators $Q^{(1)}$ and $Q^{(2)}$ satisfy (2.16), we simply write $Q^{(1)}\preceq Q^{(2)}$. For convenience, with a slight abuse of notation, we denote the matrix $\big(q_{ij}(x)\big)$ by Q_x. So Assumption 2.6 means that for each $x\in\mathbb{R}^d$, $Q_x\preceq\bar{Q}$. By Lemma 2.5, for each $x\in\mathbb{R}^d$, there exists an order-preserving coupling of Q_x and $\bar{Q}$ given in Assumption 2.6. Namely, for each $x\in\mathbb{R}^d$, there exists a Q-matrix on $\mathcal{S}\times\mathcal{S}$ such that this Q-matrix is an order-preserving coupling of Q_x and $\bar{Q}$. In fact, such an order-preserving coupling was constructed explicitly in [34, 35]; see also [6, p. 221]. For definiteness, we choose one such coupling and denote it by $\widetilde{Q}(x)=\big(\widetilde{q}(i,j;m,n)(x)\big)$, which can be expressed explicitly. For the sake of completeness and also certain subsequent application, we sketch the construction of the coupling $\widetilde{Q}(x)$ here though a method which is essentially not new (cf. [34, 35]).

As mentioned in [34], we can define the basic coupling of Q_x and $\bar{Q}$ for the points $(i,j)\in\mathcal{S}\times\mathcal{S}$ with $i>j$ as follows:

$$\begin{cases}\widetilde{q}(i,j;m,n)(x)=\big(q_{ik}(x)-\bar{q}_{jk}\big)^+, & m=k,n=j,k\ne i,\\ \widetilde{q}(i,j;m,n)(x)=\big(\bar{q}_{jk}-q_{ik}(x)\big)^+, & m=i,n=k,k\ne j,\\ \widetilde{q}(i,j;m,n)(x)=q_{ik}(x)\wedge\bar{q}_{jk}, & m=k,n=k,(k,k)\ne(i,j),\\ \widetilde{q}(i,j;m,n)(x)=0, & \text{other}\ \ (m,n)\ne(i,j),\end{cases} \tag{2.18}$$

and $\widetilde{q}(i,j;i,j)(x):=-\sum_{(m,n)\ne(i,j)}\widetilde{q}(i,j;m,n)(x)$.

Next, we construct the order-preserving coupling for the points $(i,j)\in\mathcal{S}\times\mathcal{S}$ with $i\le j$, which is the key point to construct a coupling $(\Lambda,\bar\Lambda)$ so that $\mathbb{P}(\Lambda(t)\le\bar\Lambda(t),\ \forall\, t\ge 0)=1$. For each $n\in\mathcal{S}$, define

$$\begin{aligned} a_{nn}(x)&=\begin{cases} q_{in}(x), & n\neq i,\\ 0, & n=i,\end{cases}\\ b_{nn}(x)&=\begin{cases} \bar q_{jn}, & n\neq j,\\ 0, & n=j.\end{cases}\end{aligned}\tag{2.19}$$

Then, define the sequences $\{a_{mn}(x)\}$, $\{b_{mn}(x)\}$ $(m\le n, n\in\mathcal{S})$: for $m=n-1,\, n-2,\,\cdots,\, 1$ successively as

$$\begin{aligned} a_{mn}(x)&=\big(a_{m,n-1}(x)\big)^+-\big(b_{m,n-1}(x)\big)^+,\\ b_{mn}(x)&=\big(b_{m+1,n}(x)\big)^+-\big(a_{m+1,n}(x)\big)^+.\end{aligned}\tag{2.20}$$

Here and hereafter, $a^+=\max\{a,0\}$. Let us give some explanation on this definition procedure. Clearly, we can define $a_{12}(x)$ and $b_{12}(x)$ with (2.20) by the well-defined $a_{11}(x)$, $b_{11}(x)$, $a_{22}(x)$ and $b_{22}(x)$. Suppose the $\{a_{m'n'}(x)\}$, $\{b_{m'n'}(x)\}$ $(m'\le n')$ have been defined successively for $n'=1,\,2,\,\ldots,\,n-1$. With (2.20) we can further define the case of $n'=n$. Although $b_{nn}(x)$ is independent of x, $b_{mn}(x)$ is x-dependent in general. Finally, with the well-defined sequences $\{a_{mn}(x)\}$, $\{b_{mn}(x)\}$ $(m\le n, n\in\mathcal{S})$, the desired coupling is given by

$$\begin{cases} \widetilde q(i,j;m,n)(x)=\big(a_{mn}(x)\big)^+\wedge\big(b_{mn}(x)\big)^+, & m\le n, m\neq i, n\neq j,\\ \widetilde q(i,j;\ i,\ n)(x)=\bar q_{jn}-\displaystyle\sum_{1\le m\le n,\, m\neq i}\widetilde q(i,j;m,n)(x), & i\le n\neq j,\\ \widetilde q(i,j;m,j)(x)=q_{im}(x)-\displaystyle\sum_{n\ge m,\, n\neq j}\widetilde q(i,j;m,n)(x), & i\neq m\le j,\\ \widetilde q(i,j;m,n)(x)=0, & \text{other}\quad (m,n)\neq(i,j),\end{cases}\tag{2.21}$$

and

$$\widetilde q(i,j;i,j)(x)=-\sum_{(m,n)\neq(i,j),\, m\le n}\widetilde q(i,j;m,n)(x).$$

For every fixed $x\in\mathbb{R}^d$, let $(\Lambda,\bar\Lambda)$ be the continuous-time Markov chain determined by the coupling operator $\widetilde Q$ defined in (2.21) with $\Lambda(0)\le\bar\Lambda(0)$. As shown in [34], the construction of $\widetilde q(i,j;m,n)$ guarantees that the process Λ can never jump to the front of $\bar\Lambda$ a.s., i.e. $\mathbb{P}(\Lambda(t)\le\bar\Lambda(t),\ \forall\, t\ge 0)=1$.

Consequently, the order-preserving coupling $\big(X,\Lambda,\bar\Lambda\big)$ is constructed as follows. Let X satisfy SDE (1.1) and $(\Lambda,\bar\Lambda)$ be a jumping process on $\mathcal{S}\times\mathcal{S}$ with $(\Lambda(0),\bar\Lambda(0))=(i_0,j_0)$ satisfying

$$\begin{aligned}&\mathbb{P}\{(\Lambda(t+\Delta),\bar\Lambda(t+\Delta))=(m,n)|(\Lambda(t),\bar\Lambda(t))=(i,j),X(t)=x\}\\ &\quad=\begin{cases}\widetilde q(i,j;m,n)(x)\Delta+o(\Delta), & \text{if}\,(m,n)\neq(i,j),\\ 1+\widetilde q(i,j;m,n)(x)\Delta+o(\Delta), & \text{if}\,(m,n)=(i,j),\end{cases}\end{aligned}\tag{2.22}$$

provided $\Delta\downarrow 0$. Here $\widetilde q(i,j;m,n)$ is given by (2.18) if $i_0>j_0$ and is given by (2.21) if $i_0\le j_0$.

LEMMA 2.7. *Suppose that (**H1**), (**H2**), (**Q1**), (**Q2'**) and Assumption 2.6 hold. Then the coupling process* $\big(X,\Lambda,\bar{\Lambda}\big)$ *satisfying* (1.1) *and* (2.22) *exists, and further*

$$\mathbb{P}^{(x,i,j)}\big(\Lambda(t)\le\bar{\Lambda}(t),\ \forall\, t\ge 0\big)=1,\quad x\in\mathbb{R}^d,\quad i\le j\in\mathcal{S}. \tag{2.23}$$

In addition, suppose that $\bar{Q}$ *is irreducible, then the invariant measure* $\bar{\mu}=(\bar{\mu}_1,\dots,\bar{\mu}_M)$ *of* $\bar{\Lambda}$ *exists, and for each increasing function* h *on* $\mathcal{S}$ *and each* $(x,i)\in\mathbb{R}^d\times\mathcal{S}$,

$$\mathbb{P}^{(x,i)}\bigg(\limsup_{t\to\infty}\frac{1}{t}\int_0^t h(\Lambda(s))\mathrm{d}s\le\sum_{m\in\mathcal{S}}h(m)\bar{\mu}_m\bigg)=1, \tag{2.24}$$

for every initial value $\big(X(0),\Lambda(0)\big)=(x,i)$.

Proof. By the definitions of $a_{nn}(x)$, $a_{mn}(x)$ and $\widetilde{q}(i,j;m,n)$, it is easy to check the validation of conditions (**Q1**) and (**Q2'**) for $\widetilde{Q}(x)$. Therefore, Theorem 2.1 ensures that the system $(X,\Lambda,\bar{\Lambda})$ satisfying (1.1) and (2.22) exists. Although the transition rate matrix of $(\Lambda,\bar{\Lambda})$ now depends on X which is time varying, the construction of $\widetilde{Q}_x$ still can ensure that $\Lambda(t)$ cannot jump to the front of $\bar{\Lambda}(t)$ a.s. if $\Lambda(0)\le\bar{\Lambda}(0)$. Hence, (2.23) holds.

Using the right continuity of $\big(X,\Lambda,\bar{\Lambda}\big)$, from (2.23) we obtain that for each given increasing function h on $\mathcal{S}$, $x\in\mathbb{R}^d$ and $i\le j\in\mathcal{S}$,

$$\mathbb{P}^{(x,i,j)}\bigg(\frac{1}{t}\int_0^t h(\Lambda(s))\mathrm{d}s\le\frac{1}{t}\int_0^t h(\bar{\Lambda}(s))\mathrm{d}s,\quad\forall\, t\ge 0\bigg)=1. \tag{2.25}$$

Since $\bar{Q}$ is irreducible and $\mathcal{S}$ is a finite state space, the associated Markov chain $\bar{\Lambda}$ is ergodic with the invariant probability measure given by $\bar{\mu}=(\bar{\mu}_1,\dots,\bar{\mu}_M)$. By the ergodic property of the Markov chain, we have

$$\mathbb{P}\bigg(\lim_{t\to\infty}\frac{1}{t}\int_0^t h(\bar{\Lambda}(s))\mathrm{d}s=\sum_{m\in\mathcal{S}}h(m)\bar{\mu}_m\bigg)=1. \tag{2.26}$$

For any arbitrarily given $\delta>0$, by Egorov's theorem, we then get

$$\mathbb{P}\bigg(\frac{1}{t}\int_0^t h(\bar{\Lambda}(s))\mathrm{d}s\to\sum_{m\in\mathcal{S}}h(m)\bar{\mu}_m\ \text{ uniformly as }\ t\to\infty\bigg)\ge 1-\delta. \tag{2.27}$$

Thus, for any given $\varepsilon>0$, there exists a $T(\varepsilon)>0$ such that

$$1-\delta\le\mathbb{P}\bigg(\frac{1}{t}\int_0^t h(\bar{\Lambda}(s))\mathrm{d}s\le\sum_{m\in\mathcal{S}}h(m)\bar{\mu}_m+\varepsilon\ \text{ for all }\ t\ge T(\varepsilon)\bigg). \tag{2.28}$$

Therefore, combining (2.25) and (2.28), we derive that for every increasing function h on $\mathcal{S}$, $x\in\mathbb{R}^d$ and $i\le j\in\mathcal{S}$,

$$\mathbb{P}^{(x,i,j)}\bigg(\frac{1}{t}\int_0^t h(\Lambda(s))\mathrm{d}s\le\sum_{m\in\mathcal{S}}h(m)\bar{\mu}_m+\varepsilon\ \text{ for all }\ t\ge T(\varepsilon)\bigg)\ge 1-\delta, \tag{2.29}$$

which yields that

$$\mathbb{P}^{(x,i)}\bigg(\frac{1}{t}\int_0^t h(\Lambda(s))\mathrm{d}s\le\sum_{m\in\mathcal{S}}h(m)\bar{\mu}_m+\varepsilon\ \text{ for all }\ t\ge T(\varepsilon)\bigg)\ge 1-\delta \tag{2.30}$$

since the left hand side of (2.29) does not depend on j. So,

$$\mathbb{P}^{(x,i)}\left(\limsup_{t\to\infty}\frac{1}{t}\int_0^t h(\Lambda(s))\mathrm{d}s \le \sum_{m\in\mathcal{S}} h(m)\bar{\mu}_m + \varepsilon\right) \ge 1-\delta. \tag{2.31}$$

Finally, letting ε and δ tend to 0, we arrive at (2.24). ∎

Such a Markov chain $\bar{\Lambda}$ having the properties stated in Lemma 2.7 will be called a upper control Markov chain. It will serve as an *upper envelop* as mentioned above. On the other hand, a lower control Markov chain Λ^* can be constructed under proper conditions.

Assumption 2.8. Assume that there exists a generator $Q^* = \big(q^*_{i,j}\big)$ on $\mathcal{S}$ such that the following bounds hold:

$$\begin{aligned}
&\sum_{j\ge m} q^*_{i_1,j} \le \inf_{x\in\mathbb{R}^d}\sum_{j\ge m} q_{i_2 j}(x) \quad \text{for all} \quad i_1 \le i_2 < m \quad \text{and}\\
&\sum_{j\le m} q^*_{i_1,j} \ge \sup_{x\in\mathbb{R}^d}\sum_{j\le m} q_{i_2 j}(x) \quad \text{for all} \quad m < i_1 \le i_2.
\end{aligned} \tag{2.32}$$

For each $x\in\mathbb{R}^d$, an order-preserving coupling of Q^* and $Q(x)$ can be constructed explicitly (see [6, p. 221]). Actually, replace the $(q_{ij}(x))$ and $(\bar{q}_{i,j})$ in (2.19)–(2.21) by $(q^*_{i,j})$ and $(q_{ij}(x))$ respectively, we can construct a coupling operator $\widetilde{Q}(x) = (\widetilde{q}(i,j;m,n)(x))$ for each $x\in\mathbb{R}^d$ and $i\ge j$. When $i<j$, we still use the basic coupling given in (2.18) by replacing $\bar{q}_{j,k}$ with $q^*_{j,k}$. We now construct an order-preserving coupling process $\big(X,\Lambda^*,\Lambda\big)$ as follows. Let X satisfy SDE (1.1) and (Λ^*,Λ) be a jumping process on $\mathcal{S}\times\mathcal{S}$ with $(\Lambda^*(0),\Lambda(0)) = (i_0,j_0)$ satisfying

$$\begin{aligned}
&\mathbb{P}\{(\Lambda^*(t+\Delta),\Lambda(t+\Delta)) = (m,n) | (\Lambda^*(t),\Lambda(t)) = (i,j), X(t)=x\}\\
&= \begin{cases} \widehat{q}(i,j;m,n)(x)\Delta + o(\Delta), & \text{if } (m,n)\ne(i,j),\\ 1+\widehat{q}(i,j;m,n)(x)\Delta + o(\Delta), & \text{if } (m,n)=(i,j),\end{cases}
\end{aligned} \tag{2.33}$$

provided $\Delta\downarrow 0$. Similar to Lemma 2.7, we can prove the following lemma.

Lemma 2.9. *Suppose that (**H1**), (**H2**), (**Q1**), (**Q2'**) and Assumption 2.8 hold. Then the coupling process $\big(X,\Lambda^*,\Lambda\big)$ satisfying (1.1) and (2.22) exists, and further*

$$\mathbb{P}^{(x,j,i)}\big(\Lambda(t)\ge\Lambda^*(t),\ \forall\, t\ge 0\big) = 1 \text{ with } (X(0),\Lambda^*(0),\Lambda(0)) = (x,j,i),\ i\ge j. \tag{2.34}$$

In addition, suppose that Q^* is irreducible, then for each increasing function h on $\mathcal{S}$ and each $(x,i)\in\mathbb{R}^d\times\mathcal{S}$,

$$\mathbb{P}^{(x,i)}\left(\liminf_{t\to\infty}\frac{1}{t}\int_0^t h(\Lambda(s))\mathrm{d}s \ge \sum_{m\in\mathcal{S}} h(m)\mu^*_m\right) = 1, \tag{2.35}$$

where $\mu^* = (\mu^*_1,\dots,\mu^*_M)$ is the invariant probability measure associated with Λ^*.

In what follows, we provide two concrete examples to illustrate the application of order-preserving couplings to construct upper control and lower control Markov chains for the jump component of state-dependent regime-switching processes.

EXAMPLE 2.10. Consider the case $d = 1$, $\mathcal{S} = \{1, 2\}$. Let Q_x in (1.2) be given by

$$Q_x = \big(q_{ij}(x)\big) = \begin{pmatrix} \sin^2 x - 2 & 2 - \sin^2 x \\ 1 + |\cos x| & -1 - |\cos x| \end{pmatrix}.$$

Meanwhile, we choose

$$\bar{Q} = (\bar{q}_{i,j}) = \begin{pmatrix} -2 & 2 \\ 1 & -1 \end{pmatrix} \qquad \text{and} \qquad Q^* = (q^*_{i,j}) = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}.$$

It can be verified that $Q_x \preceq \bar{Q}$ and $Q^* \preceq Q(x)$ for all $x \in \mathbb{R}$ and that both $\bar{Q}$ and Q^* are irreducible and their associated invariant probability measures are given by $\bar{\mu} = (\bar{\mu}_1, \bar{\mu}_2) = (1/3, 2/3)$ and $\mu^* = (\mu^*_1, \mu^*_2) = (2/3, 1/3)$ respectively. Thus, for the system (X, Λ) satisfying (1.1) and (1.2) and any given increasing function h on $\mathcal{S} = \{1, 2\}$, by virtue of Lemma 2.7, we have

$$\mathbb{P}^{(x,i)}\left(\limsup_{t\to\infty} \frac{1}{t}\int_0^t h(\Lambda(s))\mathrm{d}s \le \frac{h(1) + 2h(2)}{3}\right) = 1; \tag{2.36}$$

and by virtue of Lemma 2.9, we have

$$\mathbb{P}^{(x,i)}\left(\liminf_{t\to\infty} \frac{1}{t}\int_0^t h(\Lambda(s))\mathrm{d}s \ge \frac{2h(1) + h(2)}{3}\right) = 1. \tag{2.37}$$

EXAMPLE 2.11. Consider the case $d = 1$, $\mathcal{S} = \{1, 2, 3\}$. Let $Q(x)$ in (1.2) be defined by

$$Q(x) = \big(q_{ij}(x)\big) = \begin{pmatrix} -3 - |\cos x| + \sin^2 x & 1 + |\cos x| & 2 - \sin^2 x \\ 1 + \frac{x^2}{1+x^2} & -2 - \frac{x^2}{1+x^2} & 1 \\ 2 + |\sin x| & 1 + \frac{|x|}{1+|x|} & -3 - |\sin x| - \frac{|x|}{1+|x|} \end{pmatrix}.$$

Meanwhile, we choose

$$\bar{Q} = (\bar{q}_{i,j}) = \begin{pmatrix} -4 & 2 & 2 \\ 1 & -3 & 2 \\ 2 & 1 & -3 \end{pmatrix} \quad \text{and} \quad Q^* = (q^*_{i,j}) = \begin{pmatrix} -2 & 1 & 1 \\ 3 & -3 & 0 \\ 3 & 2 & -5 \end{pmatrix}.$$

For any $x \in \mathbb{R}$, it is easy to see that

$$\begin{aligned} &q_{12}(x) + q_{13}(x) \le \bar{q}_{1,2} + \bar{q}_{1,3}, \quad q_{13}(x) \le \bar{q}_{1,3}, \\ &q_{13}(x) \le \bar{q}_{2,3}, \quad q_{23}(x) \le \bar{q}_{2,3}, \\ &q_{21}(x) \ge \bar{q}_{2,1}, \quad q_{21}(x) \ge \bar{q}_{3,1}, \\ &q_{31}(x) + q_{32}(x) \ge \bar{q}_{3,1} + \bar{q}_{3,2}, \quad q_{31}(x) \ge \bar{q}_{3,1} \end{aligned}$$

and that

$$\begin{aligned} &q^*_{1,2} + q^*_{1,3} \le q_{12}(x) + q_{13}(x), \quad q^*_{1,3} \le q_{13}(x), \\ &q^*_{1,3} \le q_{23}(x), \quad q^*_{2,3} \le q_{23}(x), \\ &q^*_{2,1} \ge q_{21}(x), \quad q^*_{2,1} \ge q_{31}(x), \\ &q^*_{3,1} + q^*_{3,2} \ge q_{31}(x) + q_{32}(x), \quad q^*_{3,1} \ge q_{31}(x). \end{aligned}$$

Hence, we get that $Q(x) \preceq \bar{Q}$ and $Q^* \preceq Q(x)$ for all $x \in \mathbb{R}$, it is easy to see that

$$\begin{aligned}
&q_{12}(x) + q_{13}(x) \le q^*_{1,2} + q^*_{1,3}, \quad q_{13}(x) \le q^*_{1,3},\\
&q_{13}(x) \le q^*_{2,3}, \quad q_{23}(x) \le q^*_{2,3},\\
&q_{21}(x) \ge q^*_{2,1}, \quad q_{21}(x) \ge q^*_{3,1},\\
&q_{31}(x) + q_{32}(x) \ge q^*_{3,1} + q^*_{3,2}, \quad q_{31}(x) \ge q^*_{3,1}
\end{aligned}$$

and that

$$\begin{aligned}
&q_*(1,2) + q_*(1,3) \le q_{12}(x) + q_{13}(x), \quad q_*(1,3) \le q_{13}(x),\\
&q_*(1,3) \le q_{23}(x), \quad q_*(2,3) \le q_{23}(x),\\
&q_*(2,1) \ge q_{21}(x), \quad q_*(2,1) \ge q_{31}(x),\\
&q_*(3,1) + q_*(3,2) \ge q_{31}(x) + q_{32}(x), \quad q_*(3,1) \ge q_{31}(x).
\end{aligned}$$

Hence, we get that $Q(x) \preceq Q^*$ and $Q_* \preceq Q(x)$ for all $x \in \mathbb{R}^1$. Clearly, both $\bar{Q}$ and Q^* are irreducible and their associated invariant probability measures are given by $\bar{\mu} = (\bar{\mu}_1, \bar{\mu}_2, \bar{\mu}_3) = (7/25, 8/25, 2/5)$ and $\mu^* = (\mu^*_1, \mu^*_2, \mu^*_3) = (3/5, 7/25, 3/25)$ respectively. Thus, for the system (X, Λ) satisfying (1.1) and (1.2) and any given increasing function h on $\mathcal{S} = \{1, 2, 3\}$, by virtue of Lemma 2.7, we have

$$\mathbb{P}^{(x,i)}\left(\limsup_{t\to\infty} \frac{1}{t}\int_0^t h(\Lambda(s))\mathrm{d}s \le \frac{7h(1) + 8h(2) + 10h(3)}{25}\right) = 1; \tag{2.38}$$

and by virtue of Lemma 2.9, we have

$$\mathbb{P}^{(x,i)}\left(\liminf_{t\to\infty} \frac{1}{t}\int_0^t h(\Lambda(s))\mathrm{d}s \ge \frac{15h(1) + 7h(2) + 3h(3)}{25}\right) = 1. \tag{2.39}$$

3. Almost sure stability of regime-switching SFDE. To make the idea clear, in this work we shall study the stability of a regime-switching system with linear feedback control as we did in [22]. Recall the equation satisfied by $(X(t))$, i.e.

$$\mathrm{d}X(t) = \big[a(X(t), \Lambda(t)) - b(\Lambda(\delta(t)))X(\delta(t))\big]\mathrm{d}t + \sigma(X(t), \Lambda(t))\mathrm{d}W(t), \;\; X(0) = x \in \mathbb{R}^d. \tag{3.1}$$

We would like to point out that these constants $b(i)$, $i \in \mathcal{S}$ need not to be all positive. In order to show the almost sure stability of the system (3.1) and (1.2), we shall apply the method in [33], and the key point is to prove

$$\int_0^\infty \mathbb{E}|X(t)|^2\mathrm{d}t < \infty.$$

As mentioned in [22, Remark 3.3], we cannot show the finiteness of the following quantity at that time:

$$\mathbb{E}\int_0^\infty \mathrm{e}^{\int_0^t g(\Lambda(\delta(s)))\mathrm{d}s}\mathrm{d}t < \infty$$

for a general function g. However, with the help of Lemma 2.3, we can handle this quantity now. From now on, we suppose the coefficient $a(\cdot,\cdot) : \mathbb{R}^d \times \mathcal{S} \to \mathbb{R}^d$ satisfies the following condition.

(**H3**) There exists a positive constant M_a such that $|a(x, i)| \le M_a|x|$ for all $(x, i) \in \mathbb{R}^d \times \mathcal{S}$.

To make our computation below more precisely, we give out a more explicit construction of the probability space used in the sequel. Let

$$\Omega_1 = \{\omega|\ \omega : [0,\infty) \to \mathbb{R}^d \text{ is continuous with } \omega(0) = 0\},$$

which is endowed with the locally uniform convergence topology and the Wiener measure $\mathbb{P}_1$ so that the coordinate process $W(t,\omega) := \omega(t)$, $t \geq 0$, is a standard d-dimensional Brownian motion. Let $(\Omega_2, \mathscr{F}_2, \mathbb{P}_2)$ be a probability space and $\Pi_{\mathbb{R}}$ be the totality of point functions on $\mathbb{R}$. For a point function $(p(t))$, D_p denotes its domain, which is a countable subset of $[0,\infty)$. Let $p : \Omega_2 \to \Pi_{\mathbb{R}}$ be a Poisson point process with counting measure $N(\mathrm{d}t, \mathrm{d}z)$ on $(0,\infty) \times \mathbb{R}_+$ defined by

$$N((0,t) \times U) = \#\{s \in D_p|\ s \leq t,\ p(s) \in U\}, \quad t > 0, \quad U \in \mathscr{B}(\mathbb{R}_+), \tag{3.2}$$

and its intensity measure is $\mathrm{d}t \times \mathfrak{m}(\mathrm{d}z)$. Set $(\Omega, \mathscr{F}, \mathbb{P}) = (\Omega_1 \times \Omega_2, \mathscr{B}(\Omega_1) \times \mathscr{F}_2, \mathbb{P}_1 \times \mathbb{P}_2)$, then under $\mathbb{P} = \mathbb{P}_1 \times \mathbb{P}_2$, for $\omega = (\omega_1, \omega_2)$, $t \mapsto \omega_1(t)$ is a Wiener process, which is independent of the Poisson point process $t \mapsto p(t, \omega_2)$. Throughout this work, we will work on this probability space $(\Omega, \mathscr{F}, \mathbb{P})$. Define

$$\mathbb{E}^N[\,\cdot\,](\omega_2) = \mathbb{E}[\,\cdot\,|\mathscr{F}_2](\omega_2)$$

to be the conditional expectation with respect to the σ-algebra $\mathscr{F}_2$.

LEMMA 3.1. *Let $(X(t), \Lambda(t))$ be the solution of (3.1) and (1.2). Suppose conditions **(H1)**-**(H3)**, **(Q1)**, **(Q2)** and **(Q3)** hold. Set*

$$K(\tau) = 2\tau(2\bar{C} + M_a + \bar{b})\mathrm{e}^{(2\bar{C}+3M_a+\bar{b})\tau}, \tag{3.3}$$

where $\bar{C} = \max_{i\in\mathcal{S}} C(i)$ and $\bar{b} = \max_{i\in\mathcal{S}} b(i)$. Moreover, assume τ is sufficiently small so that $K(\tau) < 1$. Then it holds

$$\mathbb{E}^N|X(t) - X(\delta(t))|^2(\omega_2) \leq \frac{K(\tau)}{1 - K(\tau)}\mathbb{E}^N|X(t)|^2(\omega_2). \tag{3.4}$$

Proof. For any $t \geq 0$, there is a unique integer $\upsilon \geq 0$ for $t \in [\upsilon\tau, (\upsilon+1)\tau)$. Moreover, $\delta(s) = \upsilon\tau$ for $s \in [\upsilon\tau, t]$. From (3.1) we have that

$$\begin{aligned} X(t) - X(\delta(t)) &= X(t) - X(\upsilon\tau) \\ &= \int_{\upsilon\tau}^t \big[a(X(s), \Lambda(s)) - b(\Lambda(\delta(s)))X(\delta(s))\big]\mathrm{d}s \\ &\quad + \int_{\upsilon\tau}^t \sigma(X(s), \Lambda(s))\mathrm{d}w_1(s). \end{aligned}$$

According to the condition (**H1**), it follows from this and Itô's formula that

$$\begin{aligned} &|X(t) - X(\delta(t))|^2 \\ &\quad = \int_{\delta(t)}^t 2\langle X(s) - X(\delta(s)), a(X(s), \Lambda(s)) - b(\Lambda(\delta(s)))X(\delta(s))\rangle \mathrm{d}s \end{aligned}$$

$$+\int_{\delta(t)}^{t}\|\sigma(X(s),\Lambda(s))\|_{\mathrm{HS}}^{2}\mathrm{d}s+\int_{\delta(t)}^{t}2\langle X(s)-X(\delta(s)),\sigma(X(s),\Lambda(s))\mathrm{d}w_1(s)\rangle.$$

Taking the conditional expectation w.r.t. $\mathscr{F}_2$ on both sides of previous equality and using the independence of $(w_1(t))$ and $(w_2(t))$, we get

$$\begin{aligned}
&\mathbb{E}^N[|X(t)-X(\delta(t))|^2](\omega_2)\\
&\le \mathbb{E}^N\Big[\int_{\delta(t)}^{t} C(\Lambda(s))|X(s)|^2+2M_a|X(s)||X(s)-X(\delta(s))|\mathrm{d}s\Big](\omega_2)\\
&\quad+\mathbb{E}^N\Big[\int_{\delta(t)}^{t} 2b(\Lambda(\delta(s)))|X(\delta(s))||X(s)-X(\delta(s))|\mathrm{d}s\Big](\omega_2)\\
&\le \mathbb{E}^N\Big[\int_{\delta(t)}^{t} \big\{2\bar{C}(|X(s)-X(\delta(s))|^2+|X(\delta(s))|^2)\\
&\quad+M_a(3|X(s)-X(\delta(s))|^2+|X(\delta(s))|^2)\\
&\quad+\bar{b}|X(s)-X(\delta(s))|^2+\bar{b}|X(\delta(s))|^2\big\}\mathrm{d}s\Big](\omega_2)\\
&\le (2\bar{C}+M_a+\bar{b})\tau\mathbb{E}^N[|X(\delta(t))|^2](\omega_2)\\
&\quad+\mathbb{E}^N\Big[\int_{\delta(t)}^{t}(2\bar{C}+3M_a+\bar{b})|X(s)-X(\delta(s))|^2\mathrm{d}s\Big](\omega_2).
\end{aligned}$$

By virtue of Gronwall's inequality,

$$\begin{aligned}
&\mathbb{E}^N|X(t)-X(\delta(t))|^2(\omega_2)\\
&\le (2\bar{C}+M_a+\bar{b})\tau\mathbb{E}^N[|X(\delta(t))|^2](\omega_2)\mathrm{e}^{\int_{\delta(t)}^{t}\left(2\bar{C}+3M_a+\bar{b}\right)\mathrm{d}s}\\
&\le 2\tau(2\bar{C}+M_a+\bar{b})\mathbb{E}^N[|X(t)-X(\delta(t))|^2+|X(t)|^2](\omega_2)\mathrm{e}^{\int_{\delta(t)}^{t}\left(2\bar{C}+3M_a+\bar{b}\right)\mathrm{d}s},
\end{aligned}$$

which yields immediately the desired conclusion. □

In the following, we consider only the case $\mathcal{S}=\{1,2\}$ and use Lemma 2.4 to construct the control Markov chains from upper and below. For the case $\mathcal{S}$ containing more than two states, we can use Lemma 2.7 and Lemma 2.9 to construct the control Markov chains, then follow the same line as the case $\mathcal{S}=\{1,2\}$ to derive the corresponding results.

Recall that $(\bar{q}_{ij})$ and (q^*_{ij}) are defined by (2.6) and (2.7). Suppose that (2.10), (2.11) and $\bar{q}_{21}$, $q^*_{12}>0$ hold. Then, according to Lemma 2.4, it holds

$$\Lambda^*(t)\le\Lambda(t)\le\bar{\Lambda}(t),\quad t\ge 0,\quad \text{a.s.} \tag{3.5}$$

Define

$$\begin{aligned}
\bar{P}_{ij}&=\mathbb{P}(\bar{\Lambda}(\tau)=j|\bar{\Lambda}(0)=i)\\
P^*_{ij}&=\mathbb{P}(\Lambda^*(\tau)=j|\Lambda^*(0)=i),\qquad i,\,j\in\mathcal{S}.
\end{aligned}$$

For a function $b(\cdot):\mathcal{S}\to\mathbb{R}$, define

$$\bar{P}(b)=\big(\mathrm{e}^{b(i)}\bar{P}_{ij}\big),\quad P^*(b)=\big(\mathrm{e}^{b(i)}P^*_{ij}\big).$$

Then the corresponding first eigenvalues of the linear operators $\bar{P}_b$, P_b^* are denoted by

$$\begin{aligned}\bar{\lambda}_1(b) &= \max\{\mathrm{Re}(\lambda); \lambda \in \mathrm{Spec}(\bar{P}_b)\},\\ \lambda_1^*(b) &= \max\{\mathrm{Re}(\lambda); \lambda \in \mathrm{Spec}(P_b^*)\}.\end{aligned} \tag{3.6}$$

The Markov chain $(\bar{\Lambda}(n\tau))_{n\geq 0}$ is a skeleton Markov chain of $(\bar{\Lambda}(t))$. If we denote $\bar{f}_{ij}(t) = \mathbb{P}(\bar{\Lambda}(t) = j|\bar{\Lambda}(0) = i)$, $i, j \in \mathcal{S}$, $t \geq 0$, then

$$\bar{P}_{ij} = \mathbb{P}(\bar{\Lambda}(\tau) = j|\bar{\Lambda}(0) = i) = \bar{f}_{ij}(\tau).$$

It is known (cf. e.g. [6, Chapter 4]) that $\bar{f}_{ij}(t)$ satisfies the following equation

$$\bar{f}_{ij}(t) = \mathrm{e}^{\bar{q}_i t}\delta_{ij} + \sum_{k\neq j}\int_0^t \bar{f}_{ik}(t-s)\bar{q}_{kj}\mathrm{e}^{-\bar{q}_j s}\mathrm{d}s, \tag{3.7}$$

where $\delta_{ij} = 1$ if $i = j$; otherwise, $\delta_{ij} = 0$. Since $(\bar{\Lambda}(t))$ is assumed to be irreducible, this yields that $\bar{P}_{ij} > 0$ for all $i, j \in \mathcal{S}$, which means that the transition matrix $\bar{P}$ of $(\bar{\Lambda}(n\tau))_{n\geq 0}$ is positive. Therefore, Lemma 2.3 can be applied to $(\bar{\Lambda}(n\tau))_{n\geq 0}$. Similar deduction holds for $(\Lambda^*(n\tau))_{n\geq 0}$.

LEMMA 3.2. *There exist two constants $\widetilde{K}_1$, $\widetilde{K}_2 > 0$ such that for any initial point $i_0 \in \mathcal{S}$ of $\bar{\Lambda}$ and Λ^*, it holds that*

$$\widetilde{K}_1(\bar{\lambda}_1(b))^t \leq \mathbb{E}_{i_0}\Big[\exp\Big\{\int_0^t b(\bar{\Lambda}(\delta(s)))\mathrm{d}s\Big\}\Big] \leq \widetilde{K}_2(\bar{\lambda}_1(b))^t, \tag{3.8}$$

$$\widetilde{K}_1(\lambda_1^*(b))^t \leq \mathbb{E}_{i_0}\Big[\exp\Big\{\int_0^t b(\Lambda^*(\delta(s)))\mathrm{d}s\Big\}\Big] \leq \widetilde{K}_2(\lambda_1^*(b))^t, \tag{3.9}$$

for t large enough.

Proof. We only prove (3.8), and the rest assertion can be proved in the same way. Applying Lemma 2.3, we obtain that for t large enough,

$$\begin{aligned}\mathbb{E}_{i_0}\Big[\exp\Big\{\int_0^t b(\bar{\Lambda}(\delta(s)))\mathrm{d}s\Big\}\Big] &= \mathbb{E}_{i_0}\Big[\exp\Big\{\sum_{k=0}^{[t/\tau]-1}\big(b(\bar{\Lambda}(k\tau))\tau + b(\bar{\Lambda}([t/\tau]))(t-[t/\tau]\tau)\big)\Big\}\Big]\\ &\leq \mathrm{e}^{\tau\max_i b(i)}\mathbb{E}_{i_0}\Big[\exp\Big\{\sum_{k=0}^{[t/\tau]-1} b(\bar{\Lambda}(k\tau))\tau\Big\}\Big]\\ &\leq \mathrm{e}^{\tau\max_i b(i)}K_2\big(\min\{\bar{\Lambda}_{1,b},1\}\big)^{-\tau}(\bar{\lambda}_1(b))^t.\end{aligned}$$

Analogously,

$$\mathbb{E}_{i_0}\Big[\exp\Big\{\int_0^t b(\bar{\Lambda}(\delta(s)))\mathrm{d}s\Big\}\Big] \geq \mathrm{e}^{\tau(\min_i b(i)\wedge 0)}K_1\big(\max\{\bar{\lambda}_{1,b},1\}\big)^{-\tau}(\bar{\lambda}_1(b))^t.$$

Inequality (3.8) follows from the finiteness of the cardinality of $\mathcal{S}$. ∎

LEMMA 3.3. *Under the conditions **(H1)**-**(H3)**, **(Q1)**, **(Q2)** and **(Q3)**, suppose further that (2.10), (2.11) and $\bar{q}_{21}$, $q_{12}^* > 0$ hold. Assume the functions $b(\cdot)$, $C(\cdot)$, $c(\cdot)$ on $\mathcal{S}$ are all non-decreasing. Then for $\varepsilon \in (0,1)$,*

$$\mathbb{E}[|X(t)|^2] \leq |x_0|^2\mathbb{E}\Big[\mathrm{e}^{\int_0^t (C(\bar{\Lambda}(r))-2b(\Lambda^*(\delta(r)))+2\sqrt{\frac{K(\tau)}{1-K(\tau)}}b(\bar{\Lambda}(\delta(r))))\mathrm{d}r}\Big] \tag{3.10}$$

and

$$\mathbb{E}[|X(t)|^2] \geq |x_0|^2 \mathbb{E}\Big[\mathrm{e}^{\int_0^t c(\Lambda^*(r))-2b(\bar{\Lambda}(\delta(r)))-2\sqrt{\frac{K(\tau)}{1-K(\tau)}}b(\bar{\Lambda}(\delta(r)))\mathrm{d}r}\Big]. \tag{3.11}$$

Proof. It follows from Itô's formula and condition (**H1**) that

$$\begin{aligned}
\mathrm{d}|X(t)|^2 &= \big(2\langle X(t), a(X(t),\Lambda(t))\rangle + \|\sigma(X(t),\Lambda(t))\|_{\mathrm{HS}}^2\big)\mathrm{d}t \\
&\quad - 2b(\Lambda(\delta(t)))\langle X(t), X(\delta(t))\rangle \mathrm{d}t + 2\langle X(t), \sigma(X(t),\Lambda(t))\mathrm{d}w_1(t)\rangle \\
&\leq \big(C(\Lambda(t))|X(t)|^2 - 2b(\Lambda(\delta(t)))\langle X(t), X(\delta(t))\rangle\big)\mathrm{d}t \\
&\quad + 2\langle X(t), \sigma(X(t),\Lambda(t))\mathrm{d}w_1(t)\rangle \\
&\leq \big(C(\Lambda(t)) - (2-\varepsilon)b(\Lambda(\delta(t)))\big)|X(t)|^2\mathrm{d}t \\
&\quad + \frac{1}{\varepsilon}b(\Lambda(\delta(t)))|X(t)-X(\delta(t))|^2\mathrm{d}t \\
&\quad + 2\langle X(t), \sigma(X(t),\Lambda(t))\mathrm{d}w_1(t)\rangle, \text{ for every } \varepsilon\in(0,1).
\end{aligned}$$

Taking conditional expectation $\mathbb{E}^N[\,\cdot\,]$ on both sides, then applying (3.5) and the non-decreasing property of functions $b(\cdot)$ and $C(\cdot)$, we obtain that for $0 \leq s < t$,

$$\begin{aligned}
&\mathbb{E}^N[|X(t)|^2](\omega_2) - \mathbb{E}^N[|X(s)|^2](\omega_2) \\
&\leq \mathbb{E}^N\Big[\int_s^t \Big\{\big(C(\Lambda(r)) - (2-\varepsilon)b(\Lambda(\delta(r)))\big)|X(r)|^2 \\
&\qquad + \frac{1}{\varepsilon}b(\Lambda(\delta(r)))|X(r)-X(\delta(r))|^2\Big\}\mathrm{d}r\Big](\omega_2) \\
&\leq \mathbb{E}^N\Big[\int_s^t \Big\{\big(C(\bar{\Lambda}(r)) - (2-\varepsilon)b(\Lambda^*(\delta(r)))\big)|X(r)|^2 \\
&\qquad + \frac{1}{\varepsilon}b(\bar{\Lambda}(\delta(r)))|X(r)-X(\delta(r))|^2\Big\}\mathrm{d}r\Big](\omega_2) \\
&= \int_s^t \Big\{\big(C(\bar{\Lambda}(r)) - (2-\varepsilon)b(\Lambda^*(\delta(r)))\big)\mathbb{E}^N\big[|X(r)|^2\big](\omega_2) \\
&\qquad + \frac{1}{\varepsilon}b(\bar{\Lambda}(\delta(r)))\mathbb{E}^N\big[|X(r)-X(\delta(r))|^2\big](\omega_2)\Big\}\mathrm{d}r,
\end{aligned}$$

where in the last equality we used the fact that the processes $(\bar{\Lambda}(t))$ and $(\Lambda^*(t))$ are fixed once ω_2 is given. Invoking the estimate in Lemma 3.1, this yields

$$\begin{aligned}
&\mathbb{E}^N[|X(t)|^2](\omega_2) - \mathbb{E}^N[|X(s)|^2](\omega_2) \\
&\leq \int_s^t \Big\{\Big(C(\bar{\Lambda}(r)) - (2-\varepsilon)b(\Lambda^*(\delta(r))) + \frac{K(\tau)}{\varepsilon(1-K(\tau))}b(\bar{\Lambda}(\delta(r)))\Big)\mathbb{E}^N[|X(r)|^2](\omega_2)\Big\}\mathrm{d}r.
\end{aligned} \tag{3.12}$$

Note that as a function of ε,

$$C(\bar{\Lambda}(r)) - (2-\varepsilon)b(\Lambda^*(\delta(r))) + \frac{K(\tau)}{\varepsilon(1-K(\tau))}b(\bar{\Lambda}(\delta(r)))$$

takes its maximal value at

$$\varepsilon = \sqrt{\frac{K(\tau)}{1-K(\tau)}}\cdot\sqrt{\frac{b(\bar{\Lambda}(\delta(r)))}{b(\Lambda^*(\delta(r)))}}.$$

Hence, it follows from (3.12) that

$$\begin{aligned}&\mathbb{E}^N[|X(t)|^2](\omega_2)-\mathbb{E}^N[|X(s)|^2](\omega_2)\\&\leq\int_s^t\Big\{\Big(C(\bar\Lambda(r))-2b(\Lambda^*(\delta(r)))+2\sqrt{\frac{K(\tau)}{1-K(\tau)}}\sqrt{b(\bar\Lambda(\delta(r)))b(\Lambda^*(\delta(r)))}\Big)\mathbb{E}^N[|X(r)|^2](\omega_2)\Big\}\mathrm{d}r.\end{aligned}$$

Since $b(\Lambda^*(\delta(r)))\leq b(\bar\Lambda(\delta(r)))$ for all $r\geq 0$, we then have that

$$\begin{aligned}&\mathbb{E}^N[|X(t)|^2](\omega_2)-\mathbb{E}^N[|X(s)|^2](\omega_2)\\&\leq\int_s^t\Big\{\Big(C(\bar\Lambda(r))-2b(\Lambda^*(\delta(r)))+2\sqrt{\frac{K(\tau)}{1-K(\tau)}}b(\bar\Lambda(\delta(r)))\Big)\mathbb{E}^N[|X(r)|^2](\omega_2)\Big\}\mathrm{d}r.\end{aligned}\tag{3.13}$$

Set $u(t)(\omega_2)=\mathbb{E}^N[|X(t)|^2](\omega_2)$, and

$$g(r)=C(\bar\Lambda(r))-2b(\Lambda^*(\delta(r)))+2\sqrt{\frac{K(\tau)}{1-K(\tau)}}b(\bar\Lambda(\delta(r))).$$

(3.13) can be rewritten in the form

$$u(t)(\omega_2)-u(s)(\omega_2)\leq\int_s^t g(r)u(r)(\omega_2)\mathrm{d}r.\tag{3.14}$$

As the function $r\to g(r)$ needs not to be differentiable, we can use the trick as in [23, Lemma 3.2] to derive from Gronwall's inequality that

$$u(t)(\omega_2)\leq u(0)(\omega_2)\mathrm{e}^{\int_0^t g(r)\mathrm{d}r}.\tag{3.15}$$

Then we get the desired upper estimate (3.10) after taking expectation w.r.t. $\mathbb{E}[\cdot]$ on both sides of (3.15).

The lower estimate (3.11) can be deduced by the same method. Actually, it follows from Itô's formula and condition (**H1**) that

$$\begin{aligned}\mathrm{d}|X(t)|^2&=\big(2\langle X(t),a(X(t),\Lambda(t))\rangle+\|\sigma(X(t),\Lambda(t))\|_{\mathrm{HS}}^2\big)\mathrm{d}t\\&\quad-2b(\Lambda(\delta(t)))\langle X(t),X(\delta(t))\rangle\mathrm{d}t+2\langle X(t),\sigma(X(t),\Lambda(t))\mathrm{d}w_1(t)\rangle\\&\geq\big(c(\Lambda(t))-(2+\varepsilon)b(\Lambda(\delta(t)))\big)|X(t)|^2\mathrm{d}t\\&\quad-\frac{1}{\varepsilon}b(\Lambda(\delta(t)))|X(t)-X(\delta(t))|^2\mathrm{d}t\\&\quad+2\langle X(t),\sigma(X(t),\Lambda(t))\mathrm{d}w_1(t)\rangle,\ \text{for every } \varepsilon\in(0,1).\end{aligned}$$

In what follows, the difference is that we shall use the transform $c(\Lambda(r))\geq c(\Lambda^*(r))$ instead of $C(\Lambda(r))\leq C(\bar\Lambda(r))$ and corresponding transform for $b(\Lambda(r))$ in this case. However, these details are omitted. ∎

In order to estimate the long time behavior of the quantity

$$\mathbb{E}\mathrm{e}^{p\int_0^t C(\bar\Lambda(r))\mathrm{d}r},\qquad p>0,$$

we present the estimate established in [2] after introducing some necessary notation. Let

$$\bar{Q}_p = \bar{Q} + p\mathrm{diag}(C(1), \dots, C(M)),$$

where $\mathrm{diag}(C(1), \dots, C(M))$ denotes the diagonal matrix generated by $(C(1), \dots, C(M))$. Put

$$\eta_{p,C} = -\max\{\mathrm{Re}(\gamma);\ \gamma \in \mathrm{Spec}(\bar{Q}_p)\}. \tag{3.16}$$

According to [2, Proposition 4.1], for any $p > 0$, there exist two positive constants $\kappa_1(p)$, $\kappa_2(p)$ such that for any initial point $i_0 \in \mathcal{S}$,

$$\kappa_1(p)\mathrm{e}^{-\eta_{p,C}t} \le \mathbb{E}_{i_0}\Big[\mathrm{e}^{p\int_0^t C(\bar{\Lambda}(r))\mathrm{d}r}\Big] \le \kappa_2(p)\mathrm{e}^{-\eta_{p,C}t}, \quad t > 0. \tag{3.17}$$

Now we formulate our main result.

THEOREM 3.4. *Suppose the conditions in Lemma 3.3 hold. In addition, assume*

$$\eta_{3,C} > 0, \quad \lambda_1^*(-6b) < 1, \quad \textit{and} \quad \bar{\lambda}_1\Big(6\sqrt{\frac{K(\tau)}{1-K(\tau)}}b\Big) < 1. \tag{3.18}$$

Then for every initial value $(X(0), \Lambda(0)) = (x_0, i_0) \in \mathbb{R}^d \times \mathcal{S}$ *of* (1.1) *and* (1.2)*, it holds*

$$\lim_{t\to\infty} X(t) = 0, \quad a.s. \tag{3.19}$$

Proof. Applying Lemma 3.2, (3.17) and Hölder's inequality to the estimate (3.10), we get

(3.20)

$$\begin{aligned}\mathbb{E}[|X(t)|^2] &\le |x_0|^2\Big(\mathbb{E}\mathrm{e}^{3\int_0^t C(\bar{\Lambda}(r))\mathrm{d}r}\Big)^{\frac{1}{3}}\Big(\mathbb{E}\mathrm{e}^{-6\int_0^t b(\Lambda^*(\delta(r)))\mathrm{d}r}\Big)^{\frac{1}{3}}\Big(\mathbb{E}\mathrm{e}^{6\sqrt{\frac{K(\tau)}{1-K(\tau)}}\int_0^t b(\bar{\Lambda}(\delta(r)))\mathrm{d}r}\Big)^{\frac{1}{3}}\\ &\le |x_0|^2\big(\kappa_2(3)\widetilde{K}_2^2\big)^{\frac{1}{3}}\mathrm{e}^{-\frac{\eta_{3,C}t}{3}}\big(\lambda_1^*(-6b)\big)^{\frac{t}{3}}\Big(\bar{\lambda}_1\Big(6\sqrt{\frac{K(\tau)}{1-K(\tau)}}b\Big)\Big)^{\frac{t}{3}}.\end{aligned}$$

Then, it is easy to check that under the condition (3.18),

$$\int_0^\infty \mathbb{E}\big[|X(t)|^2\big]\mathrm{d}t < \infty, \tag{3.21}$$

and there exists a constant $C > 0$ such that

$$\mathbb{E}[|X(t)|^2] \le C \quad \text{for all} \quad t \ge 0. \tag{3.22}$$

By Itô's formula, we obtain that

$$\begin{aligned}&\mathbb{E}[|X(t_2)|^2] - \mathbb{E}[|X(t_1)|^2]\\ &\quad= \mathbb{E}\int_{t_1}^{t_2}\big(2\langle X(s), a(X(s),\Lambda(s)) - b(\Lambda(\delta(s)))X(\delta(s))\rangle\\ &\qquad\qquad + \|\sigma(X(s),\Lambda(s))\|_{\mathrm{HS}}^2\big)\mathrm{d}s\end{aligned}$$

for any $0 \le t_1 < t_2 < \infty$. Thus, by virtue of the condition (**H1**) and (3.22), there exists a generic constant $C > 0$ such that

$$\big|\mathbb{E}\big[|X(t_2)|^2\big] - \mathbb{E}[|X(t_1)|^2]\big| \le C(t_2 - t_1).$$

Namely, $\mathbb{E}[|X(t)|^2]$ is uniformly continuous with respect to t over $\mathbb{R}_+$. Hence, it follows from (3.21) that

$$\lim_{t\to\infty} \mathbb{E}\big[|X(t)|^2\big] = 0. \tag{3.23}$$

Now we can completely follow the proof line of [33, Theorem 3.4] to show that

$$\lim_{t\to\infty} X(t) = 0, \quad \text{a.s.}$$

The details are omitted. □

REMARK 3.5. *Note that in Lemma 3.2 and Theorem 3.4, we have assumed the non-decreasing property of the functions $b(\cdot)$, $C(\cdot)$, $c(\cdot)$ on $\mathcal{S}$. This is a technical assumption to simplify our presentation. Without this monotone condition, after doing some necessary rearrangement of the states of $\mathcal{S}$, our results remain valid. Precisely, for instance, in order to control $\exp\big(\int_0^t C(\Lambda(s))\mathrm{d}s\big)$, one can first reorder the set $\mathcal{S}$ so that $C(\cdot)$ is non-decreasing. Of course, under this new order, the original Q matrix becomes a new form $\tilde{Q} = (\tilde{q}_{ij})$, while $\tilde{Q}$ remains to be conservative and irreducible. Hence, we can define the corresponding Markov chains $(\bar{\Lambda}(t))$ and $(\Lambda^*(t))$ associated with $\tilde{Q}$. Then, applying Lemmas 2.4, 2.7, 2.9, one can control $\exp\big(\int_0^t C(\Lambda(s))\mathrm{d}s\big)$ from upper and below.*

Appendix A: Proof of Theorem 2.1. According to [26] and [32], $(\Lambda(t))$ can be represented in terms of Poisson random measure. This representation will play an important role in this work. For the sake of clarity in the presentation and calculation, we introduce the following construction of the probability space which will be used throughout this work. Let

$$\Omega^{(1)} = \{\omega |\, \omega : [0,\infty) \to \mathbb{R}^d \text{ is continuous with } \omega_0 = 0\},$$

which is endowed with the locally uniform convergence topology and the Wiener measure $\mathbb{P}^{(1)}$ so that the coordinate process $W(t,\omega) := \omega(t)$, $t \ge 0$ is a standard d-dimensional Wiener process on $(\Omega^{(1)}, \mathcal{F}^{(1)}, \{\mathcal{F}_t^{(1)}\}_{t\ge0}, \mathbb{P}^{(1)})$. Let $(\Omega^{(2)}, \mathcal{F}^{(2)}, \{\mathcal{F}_t^{(2)}\}_{t\ge0}, \mathbb{P}^{(2)})$ be a complete probability space with a filtration $\{\mathcal{F}_t^{(2)}\}_{t\ge0}$ satisfying the usual conditions, and let $\{\xi_n\}$ be a sequence of independent exponentially distributed random variables with mean 1 on $(\Omega^{(2)}, \mathcal{F}^{(2)}, \{\mathcal{F}_t^{(2)}\}_{t\ge0}, \mathbb{P}^{(2)})$. Define

$$\Omega = \Omega^{(1)} \times \Omega^{(2)}, \quad \mathcal{F} = \mathcal{F}^{(1)} \times \mathcal{F}^{(2)}, \quad \mathcal{F}_t = \mathcal{F}_t^{(1)} \times \mathcal{F}_t^{(2)}, \quad \mathbb{P} = \mathbb{P}^{(1)} \times \mathbb{P}^{(2)},$$

and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\ge0}, \mathbb{P})$ is just the probability space used throughout this appendix. The proof of Theorem 2.1 is a little long, so we separate it into two steps.

Step 1: Construction of solution. Fix some $(x,i) \in \mathbb{R}^d \times \mathcal{S}$ and consider the following SFDE

$$\mathrm{d}X^{(i)}(t) = \big[a(X^{(i)}(t),i) - b(X^{(i)}(\delta(t)),i)\big]\mathrm{d}t + \sigma(X^{(i)}(t),i)\mathrm{d}W(t) \tag{3.24}$$

with $X^{(i)}(0)=x$. By virtue of conditions (**H1**) and (**H2**), we can prove that equation (3.24) admits a unique nonexplosive solution $X^{(i)}(t)$ by using the Picard iterations following the line of [13, Chapter 5, Theorem 2.2]. Then, we have

$$\mathbb{P}\big(\lim_{m\to\infty}\tau_{O_m}=\infty\big)=1,$$

where $\tau_{O_m}:=\inf\{t\geq 0:|X^{(i)}(t)|\geq m\}$ for $m\geq 1$. Recall that $\{\xi_n\}$ is a sequence of independent mean 1 exponentially distributed random variables introduced above. Let

$$\tau_1=\theta_1:=\inf\Big\{t\geq 0:\int_0^t q_i(X^{(i)}(s))\mathrm{d}s>\xi_1\Big\}, \tag{3.25}$$

so we have

$$\mathbb{P}\big(\tau_1>t|\mathcal{F}_t^{(1)}\big)=\mathbb{P}\Big(\xi_1\geq\int_0^t q_i(X^{(i)}(s))\mathrm{d}s\Big|\mathcal{F}_t^{(1)}\Big)=\exp\Big\{-\int_0^t q_i(X^{(i)}(s))\mathrm{d}s\Big\}. \tag{3.26}$$

Then, it follows from condition (**Q2'**) that for some $m\geq|x|+1$,

$$\begin{aligned}\mathbb{P}\big(\tau_1>t\big)&=\mathbb{E}\exp\Big\{-\int_0^t q_i(X^{(i)}(s))\mathrm{d}s\Big\}\\&\geq\mathbb{E}\Big(\mathbf{1}_{\{\tau_{O_m}\geq t\}}\exp\Big\{-\int_0^t q_i(X^{(i)}(s))\mathrm{d}s\Big\}\Big)\\&=\mathbb{E}\Big(\mathbf{1}_{\{\tau_{O_m}\geq t\}}\exp\Big\{-\int_0^{t\wedge\tau_{O_m}} q_i(X^{(i)}(s))\mathrm{d}s\Big\}\Big)\\&\geq\mathbb{P}\big(\{\tau_{O_m}\geq t\}\big)\exp\Big\{-K_0\big(1+m^{\kappa_0}\big)t\Big\}.\end{aligned} \tag{3.27}$$

Since both terms of the product on the last line tend to 1 as $t\downarrow 0$, one gets $\mathbb{P}\big(\tau_1>0\big)=1$. We define a process $(X,\Lambda)\in\mathbb{R}^d\times\mathcal{S}$ on $[0,\tau_1]$ as follows:

$$X(t)=X^{(i)}(t)\text{ for all }t\in[0,\tau_1],\text{ and }\Lambda(t)=i\text{ for all }t\in[0,\tau_1). \tag{3.28}$$

Moreover, we define $\Lambda(\tau_1)\in\mathcal{S}$ according to the probability distribution:

$$\mathbb{P}\big(\Lambda(\tau_1)=j|\mathcal{F}_{\tau_1-}\big)=\frac{q_{ij}(X(\tau_1))}{q_i(X(\tau_1))}(1-\delta_{ij})\mathbf{1}_{\{q_i(X(\tau_1))>0\}}+\delta_{ij}\mathbf{1}_{\{q_i(X(\tau_1))=0\}}. \tag{3.29}$$

Consequently, the unique solution (X,Λ) of (2.1) and (1.2) has been constructed on the temporal interval $[0,\tau_1]$.

Next, we construct the solution after τ_1. To this end, we introduce

$$\eta_1:=\inf\{t\geq\tau_1;\delta(t)\geq\tau_1\},$$

which satisfies $\tau_1\leq\eta_1<+\infty$. We divide the interval $[\tau_1,+\infty)$ into two parts $[\tau_1,\eta_1)$ and $[\eta_1,+\infty)$ and consider the construction of solution separately in each subinterval. When $\eta_1=\tau_1$, we only need to consider the construction of solution on $[\eta_1,+\infty)$.

Let $\widehat{X}$ satisfy

$$\mathrm{d}\widehat{X}(t)=\big[a(\widehat{X}(t),\Lambda(\tau_1))-b(X^{(i)}(\delta(t+\tau_1)),i)\big]\mathrm{d}t+\sigma(\widehat{X}(t),\Lambda(\tau_1))\mathrm{d}\widetilde{W}(t),\ t\geq 0, \tag{3.30}$$

with $\widehat{X}(0) = X(\tau_1)$; and let $\widehat{X}'$ satisfy

$$\mathrm{d}\widehat{X}'(t) = \big[a(\widehat{X}'(t), \Lambda(\tau_1)) - b(\widehat{X}'(\delta(t)), \Lambda(\tau_1))\big]\mathrm{d}t + \sigma(\widehat{X}'(t), \Lambda(\tau_1))\mathrm{d}\widetilde{W}(t),\ t \geq 0, \tag{3.31}$$

with $\widehat{X}'(0) = X(\tau_1)$, where $\widetilde{W}(t) = W(t+\tau_1) - W(\tau_1)$. Conditions (**H1**) and (**H2**) ensure that SFDE (3.30) admits a unique nonexplosive solution, and so does the SFDE (3.31). Let

$$\widetilde{X}(t) = \widehat{X}(t) \quad \text{for } t \in [0, \eta_1 - \tau_1); \quad \widetilde{X}(t) = \widehat{X}'(t) \quad \text{for } t \in [\eta_1 - \tau_1, +\infty). \tag{3.32}$$

and

$$\theta_2 := \inf\Big\{t \geq 0 : \int_0^t q_{\Lambda(\tau_1)}(\widetilde{X}(s))\mathrm{d}s > \xi_2\Big\}. \tag{3.33}$$

As argued in (3.26), we have

$$\begin{aligned}\mathbb{P}\big(\theta_2 > t|\mathcal{F}_{\tau_1+t}\big) &= \mathbb{P}\Big(\xi_2 \geq \int_0^t q_{\Lambda(\tau_1)}(\widetilde{X}(s))\mathrm{d}s\Big|\mathscr{F}_{\tau_1+t}\Big)\\ &= \exp\Big\{-\int_0^t q_{\Lambda(\tau_1)}(\widetilde{X}(s))\mathrm{d}s\Big\}.\end{aligned}$$

Once again, we can derive from condition (**Q2'**) that $\mathbb{P}\big(\theta_2 > 0\big) = 1$. Then we let

$$\tau_2 := \tau_1 + \theta_2 = \theta_1 + \theta_2 \tag{3.34}$$

and define (X, Λ) on $[\tau_1, \tau_2]$ by

$$X(t) = \widetilde{X}(t - \tau_1) \text{ for } t \in [\tau_1, \tau_2], \quad \Lambda(t) = \Lambda(\tau_1) \text{ for } t \in [\tau_1, \tau_2), \tag{3.35}$$

and

$$\begin{aligned}&\mathbb{P}\big(\Lambda(\tau_2) = l|\mathcal{F}_{\tau_2-}\big)\\ &= \frac{q_{\Lambda(\tau_1),l}(X(\tau_2))}{q_{\Lambda(\tau_1)}(X(\tau_2))}(1 - \delta_{\Lambda(\tau_1),l})\mathbf{1}_{\{q_{\Lambda(\tau_1)}(X(\tau_2))>0\}} + \delta_{\Lambda(\tau_1),l}\mathbf{1}_{\{q_{\Lambda(\tau_1)}(X(\tau_2))=0\}}.\end{aligned} \tag{3.36}$$

Following this procedure, we can further define (X, Λ) on the interval $[\tau_n, \tau_{n+1})$ inductively for $n \geq 3$, where τ_n is defined similarly as (3.25) and (3.34). This "interlacing procedure" uniquely determines a process $(X, \Lambda) \in \mathbb{R}^d \times \mathcal{S}$ for all $t \in [0, \tau_\infty)$, where

$$\tau_\infty = \lim_{n\to\infty} \tau_n. \tag{3.37}$$

Since the sequence τ_n is strictly increasing, the limit $\tau_\infty \leq \infty$ exists. Hence, the process (X, Λ) constructed above can be regarded as the unique solution to SFDE (2.1) and (1.2) on $[0, \tau_\infty)$.

Step 2: Nonexplosion of solution. What is left to complete the proof of Theorem 2.1 is to show $\mathbb{P}(\tau_\infty = \infty) = 1$, which is also the most delicate and difficult part of the argument.

First, we show that the evolution of the discrete component Λ can be represented as a stochastic integral with respect to a Poisson random measure, which is sometimes called

Skorokhod's representation of Λ. In view of [26, Section II-2.1] or [32, Section 2.2], for each $x \in \mathbb{R}^d$, construct a family of intervals $\{\Gamma_{ij}(x);\ i,j \in \mathcal{S}\}$ on the positive half line in the following manner:

$$\begin{aligned}
\Gamma_{12}(x) &= [0, q_{12}(x)) \\
\Gamma_{13}(x) &= [q_{12}(x), q_{12}(x) + q_{13}(x)) \\
&\cdots \\
\Gamma_{1M}(x) &= \Big[\sum_{j=2}^{M-1} q_{1j}(x), q_1(x) \Big) \\
\Gamma_{21}(x) &= [q_1(x), q_1(x) + q_{21}(x)) \\
\Gamma_{23}(x) &= [q_1(x) + q_{21}(x), q_1(x) + q_{21}(x) + q_{23}(x)) \\
&\cdots
\end{aligned}$$

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

$$h(x, i, z) = \sum_{j \in \mathcal{S}} (j - i) \mathbf{1}_{\Gamma_{ij}(x)}(z).$$

Namely, for each $x \in \mathbb{R}^d$ and $i \in \mathcal{S}$, we set $h(x, i, z) = j - i$ if $z \in \Gamma_{ij}(z)$ for some $j \neq i$; otherwise $h(x, i, z) = 0$.

Put $\lambda(t) := \int_0^t q_{\Lambda(s)}(X(s))\mathrm{d}s$ and $n(t) := \max\{n \in \mathbb{N} : \xi_1 + \cdots + \xi_n \leq \lambda(t)\}$ for all $t \in [0, \tau_\infty)$, where $\{\xi_n, n = 1, 2, \dots\}$ is the sequence of independent exponential random variables with mean 1 introduced above. Then in view of (3.25), (3.26), (3.33), and (3.34), the process $\{n(t \wedge \tau_\infty), t \geq 0\}$ is a counting process that counts the number of switches for the component Λ. We can regard $n(\cdot)$ as a nonhomogeneous Poisson process with random intensity function $q_{\Lambda(t)}(X(t))$, $t \in [0, \tau_\infty)$.

Now for any $s < t \in [0, \tau_\infty)$ and $A \in \mathcal{B}(\mathcal{S})$, let

$$p((s,t] \times A) = \sum_{u \in (s,t]} \mathbf{1}_{\{\Lambda(u) \neq \Lambda(u-), \Lambda(u) \in A\}} \text{ and } p(t, A) = p((0,t] \times A).$$

Then we have $p(t \wedge \tau_\infty, \mathcal{S}) = n(t \wedge \tau_\infty)$ and

$$\begin{aligned}
\Lambda(t \wedge \tau_\infty) &= \Lambda(0) + \sum_{k=1}^{\infty} [\Lambda(\tau_k) - \Lambda(\tau_k -)] \mathbf{1}_{\{\tau_k \leq t \wedge \tau_\infty\}} \\
&= \Lambda(0) + \int_0^{t \wedge \tau_\infty} \int_{\mathcal{S}} [j - \Lambda(s-)]\, p(\mathrm{d}s, \mathrm{d}j).
\end{aligned} \tag{3.38}$$

We can also define a Poisson random measure $N(\cdot, \cdot)$ on $[0, \infty) \times \mathbb{R}_+$ by

$$N(t \wedge \tau_\infty, B) := \sum_{j \in \mathcal{S} \cap B} p(t \wedge \tau_\infty, j), \quad \text{for all } t \geq 0 \text{ and } B \in \mathcal{B}(\mathbb{R}_+).$$

Observe that for any $(x,i)\in\mathbb{R}^d\times\mathcal{S}$ and $j\in\mathcal{S}\setminus\{i\}$, we have

$$\mathfrak{m}\{z\in[0,\infty): h(x,i,z)\neq 0\}=q_i(x) \text{ and } \mathfrak{m}\{z\in[0,\infty): h(x,i,z)=j-i\}=q_{ij}(x),$$

where $\mathfrak{m}$ is the Lebesgue measure on $\mathbb{R}_+$. Therefore, we can rewrite (3.29), (3.36) and (3.38) into the following form

$$\Lambda(t\wedge\tau_\infty)=\Lambda(0)+\int_0^{t\wedge\tau_\infty}\int_{\mathbb{R}_+} h(X(s-),\Lambda(s-),z)N(\mathrm{d}s,\mathrm{d}z). \tag{3.39}$$

Second, we shall prove that the process X is nonexplosive. Without loss of generality, we fix the initial value $(X(0),\Lambda(0))=(x,i)\in\mathbb{R}^d\times\mathcal{S}$ and for any integer $m\geq[|x|]+1$, denote by $\widetilde{\tau}_m:=\inf\{t\geq 0:|X(t)|\geq m\}$ the first exit time for the X component from the open ball $O_m:=\{x\in\mathbb{R}^d:|x|<m\}$, and let $\widetilde{\tau}_\infty:=\lim_{m\to\infty}\widetilde{\tau}_m$. We shall prove that

$$\mathbb{P}\big(\widetilde{\tau}_\infty=\infty\big)=1. \tag{3.40}$$

To this end, we first consider the X component process on the temporal interval $[0,\tau)$, where $\tau>0$ is the length of discrete time observation period. Actually, by Itô's formula, we have

$$\begin{aligned}\mathrm{d}|X(t)|^2&=2\langle X(t),a(X(t),\Lambda(t))-b(X(\delta(t)),\Lambda(\delta(t)))\rangle\mathrm{d}t\\&\quad+\|\sigma(X(t),\Lambda(t))\|_{\mathrm{HS}}^2\mathrm{d}t+2\langle X(t),\sigma(X(t),\Lambda(t))\mathrm{d}W(t)\rangle\\&=\big(2\langle X(t),a(X(t),\Lambda(t))\rangle+\|\sigma(X(t),\Lambda(t))\|_{\mathrm{HS}}^2\big)\mathrm{d}t\\&\quad-2\langle X(t),b(X(\delta(t)),\Lambda(\delta(t)))\rangle\mathrm{d}t+2\langle X(t),\sigma(X(t),\Lambda(t))\mathrm{d}W(t)\rangle.\end{aligned} \tag{3.41}$$

Then, by condition (**H2**), there exists a constant $\hat{b}>0$ such that $|b(x,i)|\leq\hat{b}(1+|x|)$ for all $(x,i)\in\mathbb{R}^d\times\mathcal{S}$, which implies further that for $t\in[0,\tau)$,

$$\begin{aligned}|X(t)|^2&=|X(0)|^2+\int_0^t\big(2\langle X(s),a(X(s),\Lambda(s))\rangle+\|\sigma(X(s),\Lambda(s))\|_{\mathrm{HS}}^2\big)\mathrm{d}s\\&\quad-2\int_0^t\langle X(s),b(X(0),\Lambda(0))\rangle\mathrm{d}s+M(t)\\&\leq\big(2\hat{b}^2t+1\big)\big(|X(0)|^2+1\big)+\big(\bar{C}+1\big)\int_0^t|X(s)|^2\mathrm{d}s+M(t),\end{aligned} \tag{3.42}$$

where $M(t)$ is a continuous local martingale and $\bar{C}=\max_{i\in\mathcal{S}}C(i)$. Taking expectations on both sides yields that

$$\mathbb{E}|X(t)|^2\leq\big(2\hat{b}^2t+1\big)\big(\mathbb{E}|X(0)|^2+1\big)+\big(\bar{C}+1\big)\int_0^t\mathbb{E}|X(s)|^2\mathrm{d}s,$$

and then by Gronwall's inequality, it follows that

$$\mathbb{E}|X(t)|^2\leq\big(2\hat{b}^2t+1\big)\big(\mathbb{E}|X(0)|^2+1\big)\mathrm{e}^{(\bar{C}+1)t},\quad\text{for every } t\in[0,\tau). \tag{3.43}$$

Using Fatou's lemma, from (3.43) we also have

$$\mathbb{E}|X(\tau)|^2\leq\big(2\hat{b}^2\tau+1\big)\big(\mathbb{E}|X(0)|^2+1\big)\mathrm{e}^{(\bar{C}+1)\tau}. \tag{3.44}$$

Deducing inductively, we can obtain that for any integer $m \geq 1$,

$$\mathbb{E}|X(m\tau+t)|^2 \leq \big(2\hat{b}^2 t+1\big)\big(\mathbb{E}|X(m\tau)|^2+1\big)\mathrm{e}^{(\bar{C}+1)t} \quad \text{for every } t\in[0,\tau), \tag{3.45}$$

which further implies that $\mathbb{E}|X(t)|^2<\infty$ for any $t\geq 0$. Therefore, the explosion time $\widetilde{\tau}_\infty$ of the component X must be infinity almost surely, and so (3.40) holds.

Third, we go to prove that for any given $m \geq [|x|]+1$,

$$\mathbb{P}\big(\tau_\infty < \widetilde{\tau}_m\big)=0, \quad \text{or equivalently} \quad \mathbb{P}\big(\tau_\infty \geq \widetilde{\tau}_m\big)=1. \tag{3.46}$$

Indeed, for any arbitrarily fixed $m_0 \geq [|x|]+1$, let

$$\widehat{H} := K_0\big(1+m_0^{\kappa_0}\big).$$

Since $|X(s)|\leq m_0$ for all $s\leq \widetilde{\tau}_{m_0}$, so by condition (**Q2'**) we have $q_{ij}(X(s))\leq \widehat{H}$ for all $i,j\in\mathcal{S}$ and all $s\leq\widetilde{\tau}_{m_0}$. Then, it follows from (3.39) that for any $t\leq\tau_\infty$,

$$\begin{aligned}\Lambda(t\wedge\widetilde{\tau}_{m_0}) &= \Lambda(0)+\int_0^{t\wedge\widetilde{\tau}_{m_0}}\int_{\mathbb{R}_+} h(X(s),\Lambda(s-),z)N(\mathrm{d}s,\mathrm{d}z)\\ &= \Lambda(0)+\int_0^{t\wedge\widetilde{\tau}_{m_0}}\int_{[0,M(M-1)\widehat{H}]} h(X(s-),\Lambda(s-),z)N(\mathrm{d}s,\mathrm{d}z),\end{aligned} \tag{3.47}$$

since the integrand $h(X(s),\Lambda(s-),z)$ equals 0 when $s\leq\widetilde{\tau}_{m_0}$ and $z\notin[0,M(M-1)\widehat{H}]$, where constant M is the number of elements in $\mathcal{S}$. For the characteristic measure (i.e., the intensity measure) $\mathfrak{m}(\cdot)$ of the Poisson random measure $N(\cdot,\cdot)$, since $\mathfrak{m}\big([0,M(M-1)\widehat{H}]\big)<\infty$, so the stationary point process corresponding to the above Poisson random measure $N(\cdot,\cdot)$ has only finite *occurrence times* on the temporal interval $[0,t\wedge\widetilde{\tau}_{m_0})$ almost surely. Hence, it follows from (3.47) that the component Λ has only finite jumps (i.e., switches) on the temporal interval $[0,t\wedge\widetilde{\tau}_{m_0})$ almost surely; refer to [28, Proposition 2.1 and Corollary 2.2] for the details. This implies that $\tau_\infty\geq t\wedge\widetilde{\tau}_{m_0}$ almost surely, and then that $\tau_\infty\geq\widetilde{\tau}_{m_0}$ almost surely due to that t is arbitrary. Now we actually have proven (3.46) since the above $m_0\geq[|x|]+1$ is also arbitrary.

Let $A_m:=\big\{\omega\in\Omega:\tau_\infty<\widetilde{\tau}_m\big\}$ for $m\geq[|x|]+1$, and let

$$A=\bigcup_{m=[|x|]+1}^{\infty} A_m, \quad A^c=\Omega\backslash A.$$

It follows from (3.46) that

$$\mathbb{P}(A)=0, \quad \text{and then} \quad \mathbb{P}(A^c)=1. \tag{3.48}$$

Note that $A^c=\{\tau_\infty\geq\widetilde{\tau}_\infty\}$. Thus, by (3.40) we have

$$\mathbb{P}(\tau_\infty=\infty)\geq\mathbb{P}(\widetilde{\tau}_\infty=\infty)=1.$$

Consequently, the solution $(X(t),\Lambda(t))$ is nonexplosive, and so the above interlacing procedure actually determines a process $(X(t),\Lambda(t))$ for all $t\in[0,\infty)$. The proof is thus completed.

Acknowledgement. We are grateful to the editor and referees for their valuable comments on this paper.

REFERENCES

[1] J.H. Bao, J.H. Shao, C.G. Yuan, Approximation of invariant measures for regime-switching diffusions, Potential Anal. 44 (2016), no. 4, 707-727.

[2] J. Bardet, H. Guerin, F. Malrieu, Long time behavior of diffusions with Markov switching, ALEA Lat. Am. J. Probab. Math. Stat. 7 (2010), 151-170.

[3] G.K. Basak, A. Bisi, M.K. Ghosh, Stability of a degenerate diffusions with state-dependent switching, J. Math. Anal. Appl. 240 (1999), 219-248.

[4] M. Benaim, S. Le Borgne, F. Malrieu, P.-A. Zitt, Quantitative ergodicity for some switched dynamical systems, Electron. Commun. Probab. 17 (2012), 1-14.

[5] Z. Chao, K. Wang, C. Zhu, Y.L. Zhu, Almost sure and moment exponential stability of regime-switching jump diffusions, SIAM J. Control Optim. 55 (2017), 3458-3488.

[6] M.F. Chen, From Markov Chains to Non-Equilibrium Particle Systems, Second Edition, World Scientific, Singapore, 2004.

[7] B. Cloez, M. Hairer, Exponential ergodicity for Markov processes with random switching, Bernoulli, 21 (2015), no. 1, 505-536.

[8] B. de Saporta, J.-F. Yao, Tail of linear diffusion with Markov switching, Ann. Appl. Probab. 15 (2005), 992-1018.

[9] N. Dunford, J.T. Schwartz, Linear operators, Part I, Wiley Classics Library. John Wiley & Sons Inc. New York (1988). Reprint of the 1958 original, A Wiley-Interscience Publication.

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

[11] N. Ikeda, S. Watanabe, Stochastic Differential Equations and Diffusion Processes, North-Holland, Amsterdam, 1981.

[12] R.Z. Khasminskii, C. Zhu, G. Yin, Stability of regime-switching diffusions, Stoch. Process. Appl. 117 (2007), 1037-1051.

[13] X.R. Mao, Stochastic Differential Equations and Applications, Horwood Publishing, Chichester, England, 1997.

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

[15] X.R. Mao, W. Liu, L. Hu, Q. Luo, J. Lu, Stabilization of hybrid stochastic differential equations by feedback control based on discrete-time state observations, Systems Control Lett. 73 (2014), 88-95.

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

[17] D. Nguyen, G. Yin, Modeling and analysis of switching diffusion systems: past-dependent switching with a countable state space, SIAM J. Control Optim. 54 (2016), no. 5, 2450-2477.

[18] M. Pinsky, R. Pinsky, Transience recurrence and central limit theorem behavior for diffusions in random temporal environments, Ann. Probab. 21 (1993), 433-452.

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

[20] J.H. Shao, Ergodicity of regime-switching diffusions in Wasserstein distances, Stoch. Proc. Appl. 125 (2015), 739-758.

[21] J.H. 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.

[22] J.H. Shao, Stabilization of regime-switching processes by feedback control based on discrete time observations, SIAM J. Control Optim. 55 (2017), 724-740.

[23] J.H. Shao, Invariant measures and Euler-Maruyama's approximations of state-dependent regime-switching diffusions, SIAM J. Control Optim., 56 (2018), 3215-3238.

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

[25] M.K. von Renesse, M. Scheutzow, Existence and unqueness of solutions of stochastic functional differential equations, Random Oper. Stoch. Equ. 18 (2010), 267-284.

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

[27] F.B. Xi, Asymptotic properties of jump-diffusion processes with state-dependent switching, Stochastic Process. Appl. 119 (2009), 2198-2221.

[28] F.B. Xi, G. Yin, Jump-diffusions with state-dependent switching: existence and uniqueness, Feller property, linearization, and uniform ergodicity, Sci. China Math. 54 (2011), 2651–2667.

[29] F.B. Xi, G. Yin, Almost sure stability and instability for switching-jump-diffusion systems with state-dependent switching, J. Math. Anal. Appl. 400 (2013), 460–474.

[30] F.B. 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), no. 3, 1789-1818.

[31] G. Yin, F.B. Xi, Stability of regime-switching jump diffusions, SIAM J. Control Optim. 48 (2010), 4525-4549.

[32] G. Yin, C. Zhu, Hybrid Switching Diffusions: Properties and Applications, Vol. 63, Stochastic Modeling and Applied Probability, Springer, New York. 2010.

[33] S. You, W. Liu, J. Lu, X.R. Mao, Q. Qiu, Stabilization of hybrid systems by feedback control based on discrete-time state observations, SIAM J. Control Optim. 53 (2015), 905-925.

[34] Y.H. Zhang, Construction of order-preserving coupling for one-dimensional Markov chains (in Chinese), Chin. J. Appl. Prob. Stat. 12 (1996), 376–382.

[35] Y.H. Zhang, Construction of order-preserving coupling for one-dimensional Markov chains (continued, in Chinese), J. Beijing Normal Univ. 34 (1998), 292–296.