

Invariant measures and Euler-Maruyama's approximations of state-dependent regime-switching diffusions*

Jinghai Shao[†]

Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

May 20, 2018

Abstract

Regime-switching processes contain two components: continuous component and discrete component, which can be used to describe a continuous dynamical system in a random environment. Such processes have many different properties than general diffusion processes, and much more difficulties are needed to be overcome due to the intensive interaction between continuous and discrete component. In this work we give conditions for the existence and uniqueness of invariant measures for state-dependent regime-switching diffusion processes. Also, the strong convergence in the L^1 -norm of a numerical approximation is established and its convergence rate is provided. A refined application of Skorokhod's representation of jumping processes plays a substantial role in this work.

AMS subject Classification: 60J60, 65C30, 60H30

Key words: Regime-switching, State-dependent, Euler-Maruyama's approximation, Successful coupling

1 Introduction

The regime-switching diffusion processes have drawn much attention owing to the demand of modeling, analysis and computation of complex dynamical systems. Classical models using deterministic differential equations and stochastic differential equations alone are often inadequate,

*Supported in part by NNSFs of China No. 11771327, 11431014

[†]Email: shaojh@tju.edu.cn

and recently many models having considered the random switching of the environment are extensively proposed and investigated in stochastic control and optimization, queueing networks, filtering of dynamic systems, ecological and biological systems, mathematical finance, risk management etc. See, for instance, [4, 14, 19, 23, 41, 43] and references therein. This kind of process has been studied by Skorokhod [35], where it was called a process with a discrete component to emphasize the difference caused by the application of discrete topology for some component of the investigated process. Precisely, the regime-switching diffusion process (RSDP) concerned in this work has two components $(X(t), \Lambda(t))$. $(X(t))$ is used to describe the continuous dynamical system satisfying the following stochastic differential equation (SDE):

$$dX(t) = b(X(t), \Lambda(t))dt + \sigma(X(t), \Lambda(t))dW(t), \quad t > 0, X_0 = x \in \mathbb{R}^n, \Lambda(0) = i \in \mathcal{S}, \quad (1.1) \quad \boxed{1.1}$$

where $b : \mathbb{R}^n \times \mathcal{S} \rightarrow \mathbb{R}^n$, $\sigma : \mathbb{R}^n \times \mathcal{S} \rightarrow \mathbb{R}^n \otimes \mathbb{R}^n$, and $\mathcal{S} = \{1, 2, \dots, N\}$ with $N < \infty$. $(\Lambda(t))$ is used to describe the switching of regimes or the change of environment in which $(X(t))$ lives. $(\Lambda(t))$ is a jumping process on $\mathcal{S}$ with the transition rate satisfying

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

provided $\Delta \downarrow 0$. When $q_{ij}(x)$ is independent of x for all $i, j \in \mathcal{S}$, $(X(t), \Lambda(t))$ is called a state-independent RSDP or a RSDP with Markovian switching. Otherwise, it is called a state-dependent RSDP.

Although the RSDPs are similar to the well-known diffusion processes with time-dependent coefficients, their properties are quite different from those of the usual diffusion processes. Compared with the diffusion process in a fixed environment, the RSDP owns much more complicated behavior. The random switching of the environment has essential impact on the properties of this system, for example, the properties of recurrence, stability, and tail behavior of the stationary distribution. Pinsky and Scheutzwow in [29] constructed two examples on the half line, which showed that even if the RSDP in every fixed environment is recurrent (or transient), this process itself could be transient (or recurrent respectively) under certain random switching rate of the environment. Similar phenomenon appears in the study of stability of the RSDP, and we refer to the works [2, 3, 15, 20, 30] and references therein for the study of stability of the RSDP. The monographs [24] and [43] provide good summaries of the recent progress in the study of state-independent and state-dependent RSDPs respectively. As shown in [13], [7] for the Ornstein-Uhlenbeck process with Markovian switching, and in [19] for the Cox-Ingeroll-Ross process with Markovian switching, the stationary distributions of the corresponding processes with switching could be heavy-tailed. However, without switching it is well known that the stationary distributions of these processes are all light-tailed. Therefore, the heavy-tailed empirical evidence promotes the application of models with regime-switching.

The recurrent property of RSDP has been extensively investigated; see, for example, [8, 12, 28, 30, 31, 32] for the setting of state-independent switching processes, [8, 12, 31] for the setup of bounded state-dependent switching processes, [22] for the framework of unbounded and state-dependent switching processes. So far, there are several approaches to explore ergodicity for RSDPs; see, for instance, [8, 31] via probabilistic coupling argument, [12, 22] by weak Harris' theorem, [28, 31, 32] based on the theory of M-matrix, Perron-Frobenius theorem and the Fredholm alternative. In particular, to study the ergodicity and stability of RSDP with infinitely countable regimes, we have proposed two methods in [30, 31, 32], i.e. finite partition method based on the M-matrix theory and the principal eigenvalue of bilinear forms method.

Recently, previously introduced RSDPs have been extended in two directions: one is to extend SDEs driven by Brownian motion to those driven by general Lévy processes (e.g. [36, 42, 38]); another is to extend SDEs to functional SDEs (e.g. [23, 34, 6]) or the discrete switching process depending on the past of the continuous process in order to deal with the past dependence of the system in practice (e.g. [26]).

The purpose of current work is two folds:

1. Provide sufficient conditions to ensure the existence and uniqueness of invariant measure for state-dependent RSDPs;
2. Prove the strong convergence of a numerical approximation to the state-dependent RSDP and estimate its convergence rate.

For RSDPs with Markovian switching, these two problems have relatively been well studied. See, for instance, [12, 31, 5] for existence of invariant measures, [40, 25] for numerical approximation of state-independent RSDP under Lipschitz or non-Lipschitz conditions. However, these two problems for the state-dependent RSDPs are not well studied. The difficulty lies in the lack of proper methods to distinguish the role played by the continuous component and the discrete component due to the close interaction between them. At the first stage, by viewing $(X(t), \Lambda(t))$ as a special kind of jump-diffusions, some known results on the Markov processes can be extended to regime-switching diffusion processes. For instance, in [43], some sufficient conditions on the recurrence and stability of state-dependent RSDPs were provided in terms of Foster-Lyapunov conditions. However, it is even harder to find suitable Lyapunov functions for (state-dependent) RSDPs than for diffusion processes without switching, which is essentially due to the complexity of weakly coupled differential systems. See, for instance, [11] and references therein for the connection between regime-switching processes and weakly coupled differential systems. In current work, we propose a new comparison method based on Skorokhod's representation (cf. [35], [43]) to control the evolution of the state-dependent jumping process $(\Lambda(t))$ (see Lemma

2.1 below), then prove the convergence of the distribution of $(X(t), \Lambda(t))$ in the Wasserstein distance by the coupling method which guarantees the existence of a unique invariant measure.

Owing to the complexity of the regime-switching systems, it is usually impossible to give out explicit solutions of such systems. Thus, the numerical approximations are very important alternative of the exact solutions of such systems. However, there was few work besides [39] on the numerical approximation of state-dependent RSDPs. Compared with RSDPs with Markovian switching, the main difficulty for state-dependent RSDPs is caused by the fact that the transition rate matrices of $(\Lambda(t))$ may be different for every step of jump due to its dependence on $(X(t))$. In [39], a new type of numerical approximation $(\tilde{X}^\delta(t), \tilde{\Lambda}^\delta(t))$ was established by constructing a sequence of discrete-time Markov chains, and its weak and strong convergence to the exact solution $(X(t), \Lambda(t))$ was proved. Furthermore, this method was extended to deal with regime-switching jump diffusions in [18]. In [43, Theorem 5.13], the convergence rate of this type of approximation was estimated, i.e. $\mathbb{E} \sup_{0 \leq t \leq T} |X(t) - \tilde{X}^\delta(t)|^2 = O(\delta)$ as $\delta \rightarrow 0$. The following estimate plays the key role in the argument of [43, Theorem 5.13]:

$$\mathbb{E}[\mathbf{1}_{\{\Lambda(s) \neq \Lambda^\delta(s)\}} | \mathcal{F}_{\delta k}] \leq \sum_{i \in \mathcal{S}} \mathbf{1}_{\{\Lambda_k^\delta = i\}} \sum_{j \neq i} [q_{ij}(X(\delta k))(s - \delta k) + o(s - \delta k)]$$

for $s \in [k\delta, (k+1)\delta)$. Note that the right hand side of above estimate depends only on the value of the process $(X(t))$ at time δk . This is obscure since two processes $(X(t))$ and $(\tilde{X}^\delta(t))$ should play a consistent role in this estimate. Here, we use the notation $\Lambda(t)$, $\Lambda^\delta(t)$, $\tilde{X}^\delta(t)$ in lieu of $\alpha(t)$, $\alpha^\varepsilon(t)$, $X^\varepsilon(t)$ in [43]. In current work, based on further investigation on Skorokhod's representation of $(\Lambda(t))$, we can handle the mixture of $(X(t))$ and $(\Lambda(t))$. As an application, we construct a numerical approximation sequence $(X^\delta(t), \Lambda^\delta(t))$ of $(X(t), \Lambda(t))$ by time discretizing, which is a little different to $(\tilde{X}^\delta(t), \tilde{\Lambda}^\delta(t))$ (see Section 3 for details). Then we establish the strong convergence of this numerical approximation and estimate its order of error. However, unlike [39] or [43], we can only deal with regime-switching processes with additive noise. Precisely, we shall show that $\mathbb{E} \sup_{0 \leq t \leq T} |X(t) - X^\delta(t)| \leq C(T)\delta^{1/2}$ in Theorem 3.3 below. The following estimate plays the key role in our method:

$$\int_0^t \mathbb{E}[\mathbf{1}_{\{\Lambda(s) \neq \Lambda^\delta(s)\}}] ds \leq C\delta^{\frac{1}{2}} + C \int_0^t \mathbb{E}|X(s) - X^\delta(s)| ds$$

for some constant $C > 0$.

This paper is organized as follows. In Section 2, we investigate the existence of the invariant measure for state-dependent RSDPs. We apply the coupling method to prove the convergence of the distributions of $(X(t), \Lambda(t))$ in the Wasserstein distance to its unique invariant measure. We construct the coupling by reflection for RSDP. To guarantee this coupling to be successful, we improve the result in [37] by providing weaker and more explicit conditions. Owing to the

state-dependence, the transition rate matrices of the jumping process $(\Lambda(t))$ may be different for every step of jump. The usual technique to handle Markovian switching diffusions, i.e. ensuring first the discrete component meet together, then the continuous component meet together, does not work any more. For the state-dependent case, we have to make two components meet together at the same time, which is the main challenge for state-dependent RSDPs. In order to control the state-dependent jumping process $(\Lambda(t))$, we construct a state-independent Markov chain $(\bar{\Lambda}(t))$ so that almost surely $\Lambda(t) \leq \bar{\Lambda}(t)$ for all $t \geq 0$ and provide explicit condition in terms of $(\bar{\Lambda}(t))$ to control the exponential functional of $(\Lambda(t))$, i.e.

$$\mathbb{E} e^{\int_0^t \lambda_{\Lambda(s)} ds}$$

where $\lambda : \mathcal{S} \rightarrow \mathbb{R}$. The limitation of our construction is that the jumping process for each continuous-state x should be of birth-death form, i.e. $q_{ij}(x) = 0$ for any $i, j \in \mathcal{S}$, $|i - j| \geq 2$, and $x \in \mathbb{R}^n$.

In Section 3, we explore the time discretizing numerical approximation for state-dependent RSDPs. The key point is the estimate given in Lemma 3.2. The strong convergence of such approximation is presented in Theorem 3.3 with the order of error being $1/2$. Note that this order of error consists with the order of error provided by [40] for numerical approximation of Markovian regime-switching diffusion processes.

2 Invariant measures

Consider the state-dependent RSDP $(X(t), \Lambda(t))$ defined by (1.1) and (1.2). The assumptions used in this work on the coefficients and the transition rate matrix are collected as follows.

For the transition rate matrix $Q(x) := (q_{ij}(x))_{i,j \in \mathcal{S}}$, we shall use the following conditions:

- (Q1) For each $x \in \mathbb{R}^n$, $(q_{ij}(x))$ is conservative and irreducible.
- (Q2) $H := \max_{i \in \mathcal{S}} \sup_{x \in \mathbb{R}^n} q_i(x) < \infty$, where $q_i(x) = \sum_{j \neq i} q_{ij}(x)$ for $i \in \mathcal{S}$, $x \in \mathbb{R}^n$.
- (Q3) There exists a constant c_q so that $|q_{ij}(x) - q_{ij}(y)| \leq c_q |x - y|$, $\forall x, y \in \mathbb{R}^n$, $i, j \in \mathcal{S}$.

Concerning the coefficients of SDE (1.1), we shall use the following conditions:

- (A1) There exist constants $\alpha_i \in \mathbb{R}$, $i \in \mathcal{S}$, such that

$$2\langle x - y, b(x, i) - b(y, i) \rangle + 2\|\sigma(x, i) - \sigma(y, i)\|_{\text{HS}}^2 \leq \alpha_i |x - y|^2, \quad x, y \in \mathbb{R}^n, \quad i \in \mathcal{S}.$$

(A2) There exists a constant C_1 such that

$$|b(x, i)| + \|\sigma(x, i)\|_{\text{HS}} \leq C_1, \quad x \in \mathbb{R}^n, \quad i \in \mathcal{S}.$$

(A3) There exists a constant $C_2 > 0$ such that

$$u^* \sigma(x, i)^* u \geq C_2, \quad \forall u \in \mathbb{R}^n, |u| = 1, \quad x \in \mathbb{R}^n, \quad i \in \mathcal{S}.$$

(A4) There exist some state $i_0 \in \mathcal{S}$, constants $p > 2$, $C_3 > 0$ and $\beta \in \mathbb{R}$ such that

$$\langle x - y, b(x, i_0) - b(y, i_0) \rangle + \|\sigma(x, i_0) - \sigma(y, i_0)\|_{\text{HS}}^2 \leq \beta |x - y|^2 - C_3 |x - y|^p, \quad x, y \in \mathbb{R}^n.$$

The conditions (Q1)-(Q3) and (A1)-(A2) are used to guarantee the existence of unique non-explosive strong solution of (1.1) and (1.2) (cf. for example, [33]). Besides, condition (Q3) also plays important role in the estimation of $\mathbb{P}\left(\int_0^t \mathbf{1}_{\{\Lambda(s) \neq \Lambda'(s)\}} ds\right)$ when studying numerical approximation of state-dependent RSDPs. Condition (A4) is used to guarantee our constructed coupling processes of the state-dependent RSDP to be successful, which improves the result in [37] on successful coupling in two aspects: first, the condition (A4) is more explicit than the condition (T1) in [37], and hence is easier to be verified; second, in this work one only needs that (A4) holds for at least one state of $\mathcal{S}$, however, [37] demands that condition (T1) holds for all states in $\mathcal{S}$.

Next, we introduce Skorokhod's representation of $\Lambda(t)$ in terms of the Poisson random measure as in [35, Chapter II-2.1] or [43]. For each $x \in \mathbb{R}^n$, we construct a family of intervals $\{\Gamma_{ij}(x); i, j \in \mathcal{S}\}$ on the half line in the following manner:

$$\begin{aligned} \Gamma_{12}(x) &= [0, q_{12}(x)) \\ \Gamma_{13}(x) &= [q_{12}(x), q_{12}(x) + q_{13}(x)) \\ &\dots \\ \Gamma_{1N}(x) &= \left[\sum_{j=1}^{N-1} q_{1j}(x), q_1(x) \right) \\ \Gamma_{21}(x) &= [q_1(x), q_1(x) + q_{21}(x)) \\ \Gamma_{23}(x) &= [q_1(x) + q_{21}(x), q_1(x) + q_{21}(x) + q_{23}(x)) \\ &\dots \end{aligned}$$

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 $h : \mathbb{R}^n \times \mathcal{S} \times \mathbb{R} \rightarrow \mathbb{R}$ by

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

Then the process $(\Lambda(t))$ can be expressed by the SDE

$$d\Lambda(t) = \int_{[0,M]} h(X(t), \Lambda(t-), z) N_1(dt, dz), \quad (2.1) \quad \boxed{1.3}$$

where $M = N(N-1)H$, $N_1(dt, dz)$ is a Poisson random measure with intensity $dt \times \mathbf{m}(dz)$, and $\mathbf{m}(dz)$ is the Lebesgue measure on $[0, M]$. Let $p_1(t)$ be the stationary point process associated with the Poisson random measure $N_1(dt, dz)$. Due to the finiteness of $\mathbf{m}(dz)$ on $[0, M]$, there is only finite number of jumps of the process $p_1(t)$ in each finite time interval. Let $0 = \varsigma_0 < \varsigma_1 < \dots < \varsigma_n < \dots$ be the enumeration of all jumps of $p_1(t)$. It holds that $\lim_{n \rightarrow \infty} \varsigma_n = +\infty$ almost surely. Due to (2.1), it follows that, if $\Lambda(0) = i$,

$$\Lambda(\varsigma_1) = i + \sum_{l \in \mathcal{S}} (l - i) \mathbf{1}_{\Gamma_{il}(X(\varsigma_1))}(p_1(\varsigma_1)). \quad (2.2) \quad \boxed{1.4}$$

This yields that $(\Lambda(t))$ has a jump at ς_1 (i.e. $\Lambda(\varsigma_1) \neq \Lambda(\varsigma_1-)$) if $p_1(\varsigma_1)$ belongs to the interval $\Gamma_{il}(X(\varsigma_1))$ for some $l \neq i$. At any other cases, $(\Lambda(t))$ admits no jump at ς_1 . So the set of jumping times of $(\Lambda(t))$ is a subset of $\{\varsigma_1, \varsigma_2, \dots\}$. This fact will be used below without mentioning it again.

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

$$\Omega_1 = \{\omega \mid \omega : [0, \infty) \rightarrow \mathbb{R}^n \text{ is continuous with } \omega(0) = 0\},$$

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 n -dimensional Brownian motion. Let $(\Omega_2, \mathcal{F}_2, \mathbb{P}_2)$ 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 $p_1 : \Omega_2 \rightarrow \Pi_{\mathbb{R}}$ be a Poisson point process with counting measure $N_1(dt, dz)$ on $(0, \infty) \times [0, M]$ defined by

$$N_1((0, t) \times U) = \#\{s \in D_{p_1} \mid s \leq t, p_1(s) \in U\}, \quad t > 0, \quad U \in \mathcal{B}([0, M]), \quad (2.3) \quad \boxed{1.5}$$

and its intensity measure is $dt \times \mathbf{m}(dz)$. Set $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega_1 \times \Omega_2, \mathcal{B}(\Omega_1) \times \mathcal{F}_2, \mathbb{P}_1 \times \mathbb{P}_2)$, 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 $t \mapsto p_1(t, \omega_2)$. Throughout this work, we will work on this probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

To proceed, we provide an estimate on the exponential functional of the state-dependent jumping process by comparing it with a state-independent Markov chain through constructing a coupling process of $(\Lambda(t))$ using its Skorokhod's representation (2.1). This estimate plays an important role in controlling the evolution of this regime-switching system.

t-2.1

Lemma 2.1 (Estimate of exponential functional of $(\Lambda(t))$) Let $\mathcal{S} = \{1, 2, \dots, N\}$ with $2 \leq N < \infty$. Assume $q_{ij}(x) = 0$ for every $i, j \in \mathcal{S}$ with $|i - j| \geq 2$ and every $x \in \mathbb{R}^n$. Assume $(\lambda_i)_{i \in \mathcal{S}}$ is a nondecreasing sequence. Let $\bar{q}_{i,i+1} = \sup_{x \in \mathbb{R}^n} q_{i,i+1}(x)$, $\bar{q}_{i+1,i} = \inf_{x \in \mathbb{R}^n} q_{i+1,i}(x)$, $\bar{q}_i = -\bar{q}_{ii} = \sum_{j \neq i} \bar{q}_{ij}$ for $i \in \mathcal{S}$. Suppose that the matrix $(\bar{q}_{ij})$ is irreducible. Assume

$$\begin{aligned} & \text{for } 1 \leq i \leq N-2, q_{i,i+1}(x) + q_{i+1,i}(x) \text{ is independent of } x, \\ & \bar{q}_{N-1,N} + \bar{q}_{N,N-1} \leq q_{N-1,N}(x) + q_{N,N-1}(x), \quad \forall x \in \mathbb{R}^n. \end{aligned} \quad (2.4) \quad \text{m1}$$

Let

$$\bar{Q}_\lambda = (\bar{q}_{ij}) + \text{diag}(\lambda_1, \dots, \lambda_N),$$

where $\text{diag}(\lambda_1, \dots, \lambda_N)$ denotes the diagonal matrix generated by the vector $(\lambda_1, \dots, \lambda_N)$. Let

$$\bar{\eta} = - \max_{\gamma \in \text{spec } \bar{Q}_\lambda} \text{Re } \gamma.$$

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

$$\mathbb{E} e^{\int_0^t \lambda_{\Lambda(s)} ds} \leq C e^{-\bar{\eta} t} \quad \text{for all } t > 0. \quad (2.5) \quad \text{m2}$$

Proof Corresponding to $(\bar{q}_{ij})$, we define $\bar{\Gamma}_{ij}$ as follows: $\bar{\Gamma}_{12} = [0, \bar{q}_{12})$, $\bar{\Gamma}_{21} = [\bar{q}_{12}, \bar{q}_{12} + \bar{q}_{21})$, $\bar{\Gamma}_{23} = [\bar{q}_{12} + \bar{q}_{21}, \bar{q}_{12} + \bar{q}_{21} + \bar{q}_{23})$,

$$\bar{\Gamma}_{i,i-1} = \left[\sum_{j=1}^{i-1} \bar{q}_j, \sum_{j=1}^{i-1} \bar{q}_j + \bar{q}_{i,i-1} \right), \quad \bar{\Gamma}_{i,i+1} = \left[\sum_{j=1}^{i-1} \bar{q}_j + \bar{q}_{i,i-1}, \sum_{j=1}^i \bar{q}_j \right), \quad i \geq 3.$$

Put

$$\bar{h}(i, z) = \sum_{\ell \in \mathcal{S}} (\ell - i) \mathbf{1}_{\bar{\Gamma}_{i\ell}}(z),$$

and

$$d\bar{\Lambda}(t) = \int_{[0,M]} \bar{h}(\bar{\Lambda}(t-), z) N_1(dt, dz), \quad \bar{\Lambda}(0) = \Lambda(0), \quad (2.6) \quad \text{m3}$$

then $(\bar{\Lambda}(t))$ is a continuous time Markov chain with transition rate matrix $(\bar{q}_{ij})$. Recall that (ς_k) denotes the jumping time of the Poisson point process $(p_1(t))$, so

$$\bar{\Lambda}(\varsigma_{k+1}) = \bar{\Lambda}(\varsigma_k) + \mathbf{1}_{\bar{\Gamma}_{\bar{\Lambda}(\varsigma_k), \bar{\Lambda}(\varsigma_k)+1}}(p_1(\varsigma_{k+1})) - \mathbf{1}_{\bar{\Gamma}_{\bar{\Lambda}(\varsigma_k), \bar{\Lambda}(\varsigma_k)-1}}(p_1(\varsigma_{k+1})). \quad (2.7) \quad \text{m4}$$

Note that (2.4) implies

$$\bar{q}_{i,i+1} + \bar{q}_{i+1,i} = q_{i,i+1}(x) + q_{i+1,i}(x), \quad \forall x \in \mathbb{R}^n, \quad 1 \leq i \leq N-2. \quad (2.8) \quad \text{m5}$$

Indeed, denote by $e_i = q_{i,i+1}(x) + q_{i+1,i}(x)$ for $1 \leq i \leq N-2$. Then, by the definition of $\bar{q}_{ij}$, for any $\varepsilon > 0$, there exists $x_\varepsilon, x'_\varepsilon \in \mathbb{R}^n$ such that

$$\begin{aligned}\bar{q}_{i,i+1} + \bar{q}_{i+1,i} &\leq q_{i,i+1}(x_\varepsilon) + \inf_{x \in \mathbb{R}^n} q_{i+1,i}(x) + \varepsilon \\ &\leq q_{i,i+1}(x_\varepsilon) + q_{i+1,i}(x_\varepsilon) + \varepsilon \\ &= e_i + \varepsilon,\end{aligned}$$

and

$$\begin{aligned}\bar{q}_{i,i+1} + \bar{q}_{i+1,i} &\geq \sup_{x \in \mathbb{R}^n} q_{i,i+1}(x) + q_{i+1,i}(x'_\varepsilon) - \varepsilon \\ &\geq q_{i,i+1}(x'_\varepsilon) + q_{i+1,i}(x'_\varepsilon) - \varepsilon \\ &= e_i - \varepsilon.\end{aligned}$$

Letting $\varepsilon \downarrow 0$, we obtain (2.8).

Moreover, by (2.8) and the definition of $\bar{q}_{ij}$, it holds that for every $x \in \mathbb{R}^n$,

$$\begin{aligned}\sum_{j=1}^{i-1} q_j(x) + q_{i,i-1}(x) &= \sum_{j=1}^{i-1} (q_{j,j+1}(x) + q_{j+1,j}(x)) \\ &= \sum_{j=1}^{i-1} (\bar{q}_{j,j+1} + \bar{q}_{j+1,j}) = \sum_{j=1}^{i-1} \bar{q}_j + \bar{q}_{i,i-1}, \quad 2 \leq i \leq N-1,\end{aligned}$$

and

$$\sum_{j=1}^i q_j(x) = \sum_{j=1}^{i-1} (\bar{q}_{j,j+1} + \bar{q}_{j+1,j}) + q_{i,i+1}(x) \leq \sum_{j=1}^{i-1} (\bar{q}_{j,j+1} + \bar{q}_{j+1,j}) + \bar{q}_{i,i+1}, \quad 1 \leq i \leq N-1.$$

Therefore, for every $x \in \mathbb{R}^n$,

$$\Gamma_{i,i+1}(x) \subset \bar{\Gamma}_{i,i+1}, \quad 1 \leq i \leq N-1; \quad \bar{\Gamma}_{i,i-1} \subset \Gamma_{i,i-1}(x), \quad 2 \leq i \leq N.$$

Case 1: $\bar{\Lambda}(\varsigma_k) = \Lambda(\varsigma_k)$. For simplicity of notation, denote $\bar{\Lambda}(\varsigma_k) = \Lambda(\varsigma_k) = i$.

- If $\Lambda(\varsigma_{k+1}) = i+1$, then it must hold $p_1(\varsigma_{k+1}) \in \Gamma_{i,i+1}(X(\varsigma_{k+1}))$, and further $p_1(\varsigma_{k+1}) \in \bar{\Gamma}_{i,i+1}$. Thanks to (2.7), $\bar{\Lambda}(\varsigma_{k+1}) = i+1 = \Lambda(\varsigma_{k+1})$.
- If $\bar{\Lambda}(\varsigma_{k+1}) = i-1$, then $p_1(\varsigma_{k+1}) \in \bar{\Gamma}_{i,i-1}$. As $\bar{\Gamma}_{i,i-1} \subset \Gamma_{i,i-1}(X(\varsigma_{k+1}))$, we have $p_1(\varsigma_{k+1}) \in \Gamma_{i,i-1}(X(\varsigma_{k+1}))$ and $\Lambda(\varsigma_{k+1}) = i-1$. Therefore, $\Lambda(\varsigma_{k+1}) = \bar{\Lambda}(\varsigma_{k+1}) = i-1$.

Consequently, if $\bar{\Lambda}(\varsigma_k) = \Lambda(\varsigma_k)$, we always have $\bar{\Lambda}(\varsigma_{k+1}) \geq \Lambda(\varsigma_{k+1})$.

Case 2: $\bar{\Lambda}(\varsigma_k) > \Lambda(\varsigma_k)$. As the processes $(\Lambda(t))$ and $(\bar{\Lambda}(t))$ can both jump forward or backward at most 1, we only need to consider the situation that $\Lambda(\varsigma_k) = i - 1$ and $\bar{\Lambda}(\varsigma_k) = i$ for some $i \in \mathcal{S}$. For other cases, it obviously holds $\bar{\Lambda}(\varsigma_{k+1}) \geq \Lambda(\varsigma_{k+1})$.

- If $\Lambda(\varsigma_{k+1}) = i + 1$, then $p_1(\varsigma_{k+1}) \in \Gamma_{i,i+1}(X(\varsigma_{k+1}))$, and hence $p_1(\varsigma_{k+1}) \in \bar{\Gamma}_{i,i+1}$. This implies that $\bar{\Lambda}(\varsigma_{k+1}) = \bar{\Lambda}(\varsigma_k) = i + 1$.

Therefore, when $\bar{\Lambda}(\varsigma_k) > \Lambda(\varsigma_k)$, it must hold $\bar{\Lambda}(\varsigma_{k+1}) \geq \Lambda(\varsigma_{k+1})$.

In all, we have shown that

$$\Lambda(t) \leq \bar{\Lambda}(t) \quad a.s.. \quad (2.9) \quad \boxed{\text{com-1}}$$

According to (2.9), and invoking the monotonicity of $(\lambda_i)_{i \in \mathcal{S}}$, it holds $\mathbb{E}e^{\int_0^t \lambda_{\Lambda(s)} ds} \leq \mathbb{E}e^{\int_0^t \lambda_{\bar{\Lambda}(s)} ds}$. Applying [7, Proposition 4.1], there exists a constant $C > 0$ such that

$$\mathbb{E}e^{\int_0^t \lambda_{\Lambda(s)} ds} \leq \mathbb{E}e^{\int_0^t \lambda_{\bar{\Lambda}(s)} ds} \leq Ce^{-\bar{\eta}t}, \quad t > 0.$$

The proof is complete. $\square$

Remark 2.2 If (λ_i) in Lemma 2.1 is a nonincreasing sequence, one can modify the definition of the auxiliary Markov chain to provide an upper control of $\mathbb{E}[\exp(\int_0^t \lambda_{\Lambda(s)} ds)]$. Precisely, define $\tilde{q}_{i,i+1} = \inf_{x \in \mathbb{R}^n} q_{i,i+1}(x)$, $\tilde{q}_{i+1,i} = \sup_{x \in \mathbb{R}^n} q_{i+1,i}(x)$ for $1 \leq i \leq N - 1$. Suppose $(\tilde{q}_{ij})$ is irreducible and

$$\begin{aligned} &\text{for } 1 \leq i \leq N - 2, q_{i,i+1}(x) + q_{i+1,i}(x) \text{ is independent of } x, \\ &\tilde{q}_{N-1,N} + \tilde{q}_{N,N-1} \geq q_{N-1,N}(x) + q_{N,N-1}(x), \quad \forall x \in \mathbb{R}^n. \end{aligned} \quad (2.10) \quad \boxed{\text{low-est}}$$

Then, similar to $(\bar{\Lambda}(t))$, we can define a Markov process $(\tilde{\Lambda}(t))$ associated with $(\tilde{q}_{ij})$ satisfying $\tilde{\Lambda}(0) = \Lambda(0)$ such that

$$\Lambda(t) \geq \tilde{\Lambda}(t), \quad \forall t \geq 0, \quad a.s.. \quad (2.11) \quad \boxed{\text{low}}$$

Applying [7, Proposition 4.1] again, one can derive the control of $\mathbb{E}[\exp(\int_0^t \lambda_{\Lambda(s)} ds)]$ from above as desired. Furthermore, if (λ_i) is not monotone, one can reorder the set $\mathcal{S}$ according to the order of (λ_i) . Then, one needs to check the conditions in Lemma 2.1 on $(q_{ij}(x))$ in this new order so that one can construct the auxiliary Markov chain to control $\mathbb{E}[\exp(\int_0^t \lambda_{\Lambda(s)} ds)]$.

Example 2.1 The following two transition rate matrices satisfy condition (2.4) in Lemma 2.1.

$$1. \quad Q(x) = \begin{pmatrix} -\theta_1 - \sin^2 x & \theta_1 + \sin^2 x \\ \theta_2 + \cos^2 x & -\theta_2 - \cos^2 x \end{pmatrix}, \quad \theta_1, \theta_2 > 1.$$

$$2. \quad Q(x) = \begin{pmatrix} -\theta_1 + \frac{x^2}{1+x^2} & \theta_1 - \frac{x^2}{1+x^2} \\ \theta_2 + \frac{x^2}{1+x^2} & -\theta_2 - \frac{x^2}{1+x^2} \end{pmatrix}, \quad \theta_1, \theta_2 > 1.$$

In this work, the existence and uniqueness of the invariant measure for $(X(t), \Lambda(t))$ is deduced by analyzing the convergence of its distribution in the Wasserstein distance along the line of [12] and [32]. However, the dependence of the transition rate of $(\Lambda(t))$ on the process $(X(t))$ makes it more difficult to ensure the coupling process to be successful. Now we introduce our coupling process for $(X(t), \Lambda(t))$ and prove it will be successful after some necessary preparations.

Let $(X^{x,i}(t), \Lambda^{x,i}(t))$ and $(X^{y,j}(t), \Lambda^{y,j}(t))$ denote the solutions of (1.1) and (1.2) starting from (x, i) and (y, j) respectively. We introduce the coupling by reflection as follows: set

$$a(x, i) = \sigma(x, i)\sigma(x, i)^*, \quad a(x, i, y, j) = \begin{pmatrix} a(x, i) & c(x, i, y, j) \\ c(x, i, y, j)^* & a(y, j) \end{pmatrix} \quad (2.12) \quad \boxed{\text{e-1}}$$

for $x, y \in \mathbb{R}^n$, $i, j \in \mathcal{S}$, where

$$c(x, i, y, j) = \sigma(x, i)(I - 2\bar{u}\bar{u}^*)\sigma(y, j)^*,$$

$\bar{u} = (x - y)/|x - y|$ and I denotes identity matrix. Here A^* denotes the transpose of the matrix A . The coupling by reflection for diffusion processes is due to Lindvall and Rogers [21], and we refer the readers to Chen and Li [10] which showed that such coupling is a successful coupling. The choice of $c(x, i, y, j)$ ensures that $a(x, i, y, j)$ is nonnegative definite. Let

$$H(x, y) = (I - 2\bar{u}\bar{u}^*), \quad \text{and} \quad G(x, i, y, j) = \begin{pmatrix} \sigma(x, i) & 0 \\ \sigma(y, j)H(x, y) & 0 \end{pmatrix}.$$

Then the matrix $G(x, i, y, j)$ satisfies $G(x, i, y, j)G^*(x, i, y, j) = a(x, i, y, j)$. Consider the following SDEs:

$$d \begin{pmatrix} X(t) \\ Y(t) \end{pmatrix} = \begin{pmatrix} b(X(t), \Lambda(t)) \\ b(Y(t), \Lambda'(t)) \end{pmatrix} + G(X(t), \Lambda(t), Y(t), \Lambda'(t))d\tilde{W}(t), \quad (2.13) \quad \boxed{\text{coup-1}}$$

where $(\tilde{W}(t))$ denotes the $2n$ -dimensional Wiener process;

$$\begin{aligned} d\Lambda(t) &= \int_{[0, M]} h(X(t), \Lambda(t-), z)N_1(dt, dz), \\ d\Lambda'(t) &= \int_{[0, M]} h(Y(t), \Lambda'(t-), z)N_2(dt, dz), \end{aligned} \quad (2.14) \quad \boxed{\text{coup-2}}$$

satisfying $(X(0), \Lambda(0)) = (x, i)$ and $(Y(0), \Lambda'(0)) = (y, j)$, where $N_1(dt, dz)$ and $N_2(dt, dz)$ are mutually independent Poisson random measures with intensity measure $dt \times \mathbf{m}(dz)$. The existence of solution to SDEs (2.13) and (2.14) can be established in the same way as (1.1) and (1.2).

Then $(X(t), \Lambda(t), Y(t), \Lambda'(t))$ is called a coupling by reflection of the processes $(X^{x,i}(t), \Lambda^{x,i}(t))$ and $(X^{y,j}(t), \Lambda^{y,j}(t))$.

t-1.2 **Lemma 2.3** Assume that (Q1)-(Q3) and (A1), (A2) hold. Suppose $(\alpha_i)_{i \in \mathcal{S}}$ is a nondecreasing sequence. $(\bar{q}_{ij})$ is defined as in Lemma 2.1. Assume the conditions on (q_{ij}) and $(\bar{q}_{ij})$ in Lemma 2.1 still hold. Let

$$Q_\alpha = (\bar{q}_{ij}) + \text{diag}(\alpha_1, \dots, \alpha_N), \quad \text{and} \quad \eta_\alpha = - \max_{\zeta \in \text{spec}(Q_\alpha)} \text{Re } \zeta.$$

Assume

$$\eta_\alpha > 0. \quad (2.15) \quad \text{con-1}$$

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

$$\mathbb{E}|X(t) - Y(t)|^2 \leq C|x - y|^2 e^{-\eta_\alpha t}, \quad t > 0. \quad (2.16) \quad 2.1$$

Proof For simplicity of notation, set $Z(t) = X(t) - Y(t)$. According to the construction of $a(x, i, y, j)$, it holds

$$\begin{aligned} & \text{tr}(a(x, i, y, j)) \\ &= \text{tr}(\sigma(x, i)\sigma(x, i)^* + \sigma(y, j)\sigma(y, j)^* - 2\sigma(x, i)\sigma(y, j)^*) + 4 \frac{(x - y)^*}{|x - y|} \sigma(y, j)^* \sigma(x, i) \frac{(x - y)}{|x - y|} \\ &= \|\sigma(x, i) - \sigma(y, j)\|_{\text{HS}}^2 + 4 \frac{(x - y)^*}{|x - y|} \sigma(y, j)^* \sigma(x, i) \frac{(x - y)}{|x - y|}. \end{aligned}$$

By (A1), (A2) and Itô's formula, we obtain, for any $\gamma > 0$,

$$\begin{aligned} d|Z(t)|^2 &= \{2\langle Z(t), b(X(t), \Lambda(t)) - b(Y(t), \Lambda'(t)) \rangle \\ &\quad + \text{tr}(a(X(t), \Lambda(t), Y(t), \Lambda'(t)))\} dt + dM_t \\ &\leq \left\{ \alpha_{\Lambda(t)} |Z(t)|^2 + 2\langle Z(t), b(Y(t), \Lambda(t)) - b(Y(t), \Lambda'(t)) \rangle \right. \\ &\quad \left. + 2\|\sigma(Y(t), \Lambda(t)) - \sigma(Y(t), \Lambda'(t))\|_{\text{HS}}^2 \right. \\ &\quad \left. + 4 \frac{(X(t) - Y(t))^*}{|X(t) - Y(t)|} \sigma(Y(t), \Lambda(t))^* \sigma(X(t), \Lambda(t)) \frac{(X(t) - Y(t))}{|X(t) - Y(t)|} \right\} dt + dM_t \\ &\leq \{(\gamma + \alpha_{\Lambda(t)})|Z(t)|^2 + \frac{4C_1^2}{\gamma} + 12C_1^2\} dt + dM_t, \end{aligned} \quad (2.17) \quad 2.2$$

where (M_t) is a martingale with $M_0 = 0$. By Lemma 2.1, we can define a Markov chain $(\bar{\Lambda}(t))$ associated with the transition rate matrix $(\bar{q}_{ij})$ such that $\alpha_{\Lambda(t)} \leq \alpha_{\bar{\Lambda}(t)}$ for all $t > 0$ almost surely. Hence, for every $\lambda > 0$,

$$\mathbb{E}_{\mathbb{P}_1} \left[e^{-\lambda t} |Z(t)|^2 \right]$$

$$\begin{aligned}
&\leq |x - y|^2 + \int_0^t (4\gamma^{-1} + 12)C_1^2 e^{-\lambda s} ds + \mathbb{E}_{\mathbb{P}_1} \int_0^t (-\lambda + \gamma + \alpha_{\Lambda(s)}) e^{-\lambda s} |Z(s)|^2 ds \\
&\leq |x - y|^2 + \frac{(4\gamma^{-1} + 12)C_1^2}{\lambda} + \int_0^t (-\lambda + \gamma + \alpha_{\bar{\Lambda}(s)}) e^{-\lambda s} \mathbb{E}_{\mathbb{P}_1} |Z(s)|^2 ds.
\end{aligned}$$

By Gronwall's inequality, we get

$$e^{-\lambda t} \mathbb{E}_{\mathbb{P}_1} |Z(t)|^2 \leq \left(|x - y|^2 + \frac{(4\gamma^{-1} + 12)C_1^2}{\lambda} \right) e^{\int_0^t (-\lambda + \gamma + \alpha_{\bar{\Lambda}(s)}) ds}.$$

Then, taking expectation w.r.t. $\mathbb{P}_2$ in both sides of the previous inequality and applying [7, Proposition 4.1], we obtain that there exists a $C > 0$ such that

$$\mathbb{E}|Z(t)|^2 \leq C(|x - y|^2 + \frac{(4\gamma^{-1} + 12)C_1^2}{\lambda}) e^{(\gamma - \eta_\alpha)t}, \quad t > 0. \quad (2.18) \quad \boxed{2.3}$$

By the arbitrariness of γ and λ , letting first $\lambda \rightarrow +\infty$ then $\gamma \downarrow 0$ in (2.18), we obtain that

$$\mathbb{E}|Z(t)|^2 \leq C|x - y|^2 e^{-\eta_\alpha t}, \quad (2.19) \quad \boxed{2.4}$$

and further that

$$\sup_{t>0} \mathbb{E}|Z(t)|^2 < \infty$$

due to the positiveness of η_α . □

t-1.3 **Lemma 2.4** *Under the same assumptions and notation of Lemma 2.3, it holds that*

$$\sup_{t \geq 0} \mathbb{E}|X^{x,i}(t)|^2 \leq C(1 + |x|^2), \quad x \in \mathbb{R}^n, i \in \mathcal{S}, \quad (2.20) \quad \boxed{3.1}$$

where C is a constant.

Proof Note that condition **(A1)** implies that for any $\varepsilon > 0$, there exists a constant $C_\varepsilon > 0$ such that

$$2\langle x, b(x, i) \rangle + \|\sigma(x, i)\|_{\text{HS}}^2 \leq C_\varepsilon + (\varepsilon + \alpha_i)|x|^2, \quad x \in \mathbb{R}^n, i \in \mathcal{S}. \quad (2.21) \quad \boxed{3.1.5}$$

By (2.21) and applying Itô's formula to $X(t) = X^{x,i}(t)$, we get

$$d|X(t)|^2 \leq (C_\varepsilon + (\varepsilon + \alpha_{\Lambda(t)})|X(t)|^2)dt + 2\langle X(t), \sigma(X(t), \Lambda(t))dW(t) \rangle.$$

For every $\lambda > 0$,

$$d[e^{-\lambda t}|X(t)|^2] \leq e^{-\lambda t} \{-\lambda|X(t)|^2 + C_\varepsilon + (\varepsilon + \alpha_{\Lambda(t)})|X(t)|^2\}dt + 2e^{-\lambda t} \langle X(t), \sigma(X(t), \Lambda(t))dW(t) \rangle.$$

Taking expectation in both sides w.r.t. $\mathbb{P}_1$ and noting $\alpha_{\Lambda(t)} \leq \alpha_{\bar{\Lambda}(t)}$ a.s. by Lemma 2.1, we can deduce that

$$\begin{aligned} e^{-\lambda t} \mathbb{E}_{\mathbb{P}_1} |X(t)|^2 &\leq |x|^2 + \frac{C_\varepsilon}{\lambda} + \mathbb{E}_{\mathbb{P}_1} \left\{ \int_0^t (\varepsilon + \alpha_{\Lambda(s)} - \lambda) e^{-\lambda s} |X(s)|^2 ds \right\} \\ &\leq |x|^2 + \frac{C_\varepsilon}{\lambda} + \int_0^t (\varepsilon + \alpha_{\bar{\Lambda}(s)} - \lambda) e^{-\lambda s} \mathbb{E}_{\mathbb{P}_1} |X(s)|^2 ds. \end{aligned} \quad (2.22) \quad \boxed{3.3}$$

Using Gronwall's inequality, this yields

$$e^{-\lambda t} \mathbb{E}_{\mathbb{P}_1} |X(t)|^2 \leq (|x|^2 + \frac{C_\varepsilon}{\lambda}) e^{\int_0^t (\varepsilon + \alpha_{\bar{\Lambda}(s)} - \lambda) ds}.$$

Therefore, by [7, Proposition 4.1], there exists a constant C such that

$$\mathbb{E} |X(t)|^2 \leq (|x|^2 + \frac{C_\varepsilon}{\lambda}) \mathbb{E} e^{\int_0^t \varepsilon + \alpha_{\bar{\Lambda}(s)} ds} \leq C (|x|^2 + \frac{C_\varepsilon}{\lambda}) e^{-(\eta_\alpha - \varepsilon)t}. \quad (2.23) \quad \boxed{3.4}$$

Setting $\varepsilon = \frac{1}{2}\eta_\alpha > 0$, we can deduce from (2.23) that the desired estimate (2.20) holds. $\square$

Next, we go to show that our constructed coupling process is a successful coupling process.

t-1.4 **Lemma 2.5** *Assume the conditions of Lemma 2.3 are fulfilled. In addition, assume **(A3)** holds for $i \in \mathcal{S}$ and **(A4)** holds for $i_0 = 1$. Then the coupling $(X(t), \Lambda(t), Y(t), \Lambda'(t))$ determined by (2.13) and (2.14) is a successful coupling, that is,*

$$\tau_c := \inf\{t > 0; (X(t), \Lambda(t)) = (Y(t), \Lambda'(t))\} < \infty \quad a.s.$$

Proof If $(\Lambda(0), \Lambda'(0)) \neq (1, 1)$, the proof is divided into three steps. Otherwise, we can start directly from the second step below.

Step 1: Let

$$\tau = \inf\{t \geq 0; \Lambda(t) = \Lambda'(t) = 1\}, \quad (2.24) \quad \boxed{4.1}$$

and we shall first show the stopping time τ is almost surely finite. Recall that

$$\bar{q}_{i,i+1} = \sup_{x \in \mathbb{R}^n} q_{i,i+1}(x), \quad \bar{q}_{i+1,i} = \inf_{x \in \mathbb{R}^n} q_{i+1,i}(x), \quad 1 \leq i \leq N-1,$$

and $\bar{h}(i, z) = \sum_{\ell \in \mathcal{S}} (\ell - i) \mathbf{1}_{\bar{\Gamma}_{i\ell}}(z)$, which are defined in Lemma 2.1. Define

$$\begin{aligned} d\Lambda^{(1)}(t) &= \int_{[0,M]} \bar{h}(\Lambda^{(1)}(t-), z) N_1(dt, dz), \quad \Lambda^{(1)}(0) = \Lambda(0), \\ d\Lambda^{(2)}(t) &= \int_{[0,M]} \bar{h}(\Lambda^{(2)}(t-), z) N_2(dt, dz), \quad \Lambda^{(2)}(0) = \Lambda'(0). \end{aligned} \quad (2.25) \quad \boxed{4.2}$$

According to (2.9) in Lemma 2.1, it holds almost surely $\Lambda(t) \leq \Lambda^{(1)}(t)$ and $\Lambda'(t) \leq \Lambda^{(2)}(t)$, $t \geq 0$. The mutual independence of $N_1(dt, dz)$ and $N_2(dt, dz)$ yields that $(\Lambda^{(1)}(t))$ and $(\Lambda^{(2)}(t))$ are also mutually independent. Put

$$\tau' = \inf\{t \geq 0; \Lambda^{(1)}(t) = \Lambda^{(2)}(t) = 1\}.$$

Then it is easy to see that

$$\tau \leq \tau', \quad a.s. \quad (2.26) \quad \boxed{4.3}$$

Moreover, $(\Lambda^{(1)}(t), \Lambda^{(2)}(t))$ is an independent coupling corresponding to the operator $\bar{Q} = (\bar{q}_{ij})$ and itself (cf. for instance, [9]). Due to the irreducibility of $\bar{Q}$ and the finiteness of $\mathcal{S} \times \mathcal{S}$, there exists a positive constant θ such that

$$\mathbb{P}(\tau' \geq t) \leq e^{-\theta t}, \quad t > 0.$$

Invoking (2.26), it holds that

$$\mathbb{P}(\tau \geq t) \leq \mathbb{P}(\tau' \geq t) \leq e^{-\theta t}, \quad t > 0, \quad (2.27) \quad \boxed{4.4}$$

and hence $\mathbb{P}(\tau = \infty) = 0$.

Step 2: Using the notation introduced in (2.13), let $(X^{(1)}(t), Y^{(1)}(t))$ be the solution of the following SDE:

$$d \begin{pmatrix} X^{(1)}(t) \\ Y^{(1)}(t) \end{pmatrix} = \begin{pmatrix} b(X^{(1)}(t), 1) \\ b(Y^{(1)}(t), 1) \end{pmatrix} dt + G(X^{(1)}(t), 1, Y^{(1)}(t), 1) d\tilde{W}(t), \quad (2.28) \quad \boxed{4.5}$$

satisfying $(X^{(1)}(0), Y^{(1)}(0)) = (x, y)$, which is the process corresponding to $(X(t), Y(t))$ in the fixed environment $(i, j) = (1, 1)$. We shall use the criteria established in [10] to verify the successfulness of this coupling. To estimate the coupling time, as done in [10], we introduce the following notation:

$$\begin{aligned} A(x, y) &= a(x, 1) + a(y, 1) - 2c(x, 1, y, 1), \\ B(x, y) &= \langle x - y, (b(x, 1) - b(y, 1))(x - y) \rangle, \\ \bar{A}(x, y) &= \langle (x - y), A(x, y)(x - y) \rangle / |x - y|^2, \quad x \neq y. \end{aligned}$$

By the condition **(A3)**, it holds

$$\begin{aligned} \inf_{|x-y|=r} \bar{A}(x, y) &= \inf_{|x-y|=r} |(\sigma(x, 1) - \sigma(y, 1))\bar{u}|^2 + 4(\bar{u}^* \sigma^*(x, 1)\bar{u})(\bar{u}^* \sigma(y, 1)^* \bar{u}) \\ &\geq 4C_2^2, \end{aligned}$$

where $\bar{u} = (x - y)/|x - y|$. According to the condition (2),

$$\begin{aligned} & \sup_{|x-y|=r} \frac{\operatorname{tr}(A(x, y)) - \bar{A}(x, y) + 2B(x, y)}{\bar{A}(x, y)} \\ & \leq \sup_{|x-y|=r} \frac{\beta|x-y|^2 - C_3|x-y|^p}{\bar{A}(x, y)} - 1 \leq \frac{\beta r^2 - C_3 r^p}{4C_2^2}. \end{aligned}$$

Set $\alpha(r) = 4C_2^2$, $\gamma(r) = \frac{\beta r^2 - C_3 r^p}{4C_2^2}$, and

$$C(r) = \exp \left[\int_1^r \frac{\gamma(u)}{u} du \right].$$

Analogous to [10, Theorems 4.2 and 5.1], for positive integers ℓ and k , set

$$\begin{aligned} \tau_c^{(1)} &= \inf \{t \geq 0; X^{(1)}(t) = Y^{(1)}(t)\}, \\ S_\ell &= \inf \{t \geq 0; |X^{(1)}(t) - Y^{(1)}(t)| > \ell\}, \\ T_k &= \inf \{t \geq 0; |X^{(1)}(t) - Y^{(1)}(t)| < 1/n\}. \end{aligned}$$

Put $T_{k,\ell} = T_k \wedge S_\ell$, and

$$F_{k,\ell}(r) = - \int_{1/k}^r C(s)^{-1} \left(\int_s^\ell \frac{C(u)}{\alpha(u)} du \right) ds.$$

Then it holds

$$\begin{aligned} -\infty &< F_{k,\ell}(r) \leq 0, \quad F'_{k,\ell}(r) \leq 0, \\ F''_{k,\ell}(r) + \frac{F'_{k,\ell}(r)\gamma(r)}{r} &= \frac{1}{\alpha(r)}. \end{aligned}$$

Applying Dynkin's formula, we get that

$$\begin{aligned} & \mathbb{E}_{x,y} F_{k,\ell}(|X^{(1)}(t \wedge T_{k,\ell}) - Y^{(1)}(t \wedge T_{k,\ell})|) - F_{k,\ell}(|x - y|) \\ &= \frac{1}{2} \mathbb{E}_{x,y} \int_0^{t \wedge T_{k,\ell}} \bar{A}(X^{(1)}(s), Y^{(1)}(s)) F''_{k,\ell}(|Z^{(1)}(s)|) + F'_{k,\ell}(|Z^{(1)}(s)|) [\operatorname{tr} A(X^{(1)}(s), Y^{(1)}(s)) \\ & \quad - \bar{A}(X^{(1)}(s), Y^{(1)}(s)) + 2B(X^{(1)}(s), Y^{(1)}(s))] / |Z^{(1)}(s)| \, ds \\ & \geq \frac{1}{2} \mathbb{E}_{x,y} (t \wedge T_{k,\ell}). \end{aligned}$$

Letting $t \rightarrow \infty$, this yields that

$$\mathbb{E}_{x,y} T_{k,\ell} \leq -2F_{k,\ell}(|x - y|).$$

Set

$$F(r) = \lim_{k \rightarrow \infty} \lim_{\ell \rightarrow \infty} F_{k,\ell} = - \int_0^r C(s)^{-1} \left(\int_s^\infty \frac{C(u)}{\alpha(u)} du \right) ds.$$

Letting $\ell \rightarrow \infty$ and then $k \rightarrow \infty$, we obtain

$$\mathbb{E}_{x,y} \tau_c^{(1)} \leq -2F(|x-y|). \quad (2.29) \quad \boxed{4.6}$$

It is simple to check that

$$C(s)^{-1} \int_s^\infty \frac{C(u)}{\alpha(u)} du \sim s^{1-p}, \quad \text{as } s \rightarrow \infty.$$

As $p > 2$, this yields that

$$\lim_{r \rightarrow \infty} F(r) = - \int_0^\infty C(s)^{-1} \left(\int_s^\infty \frac{C(u)}{\alpha(u)} du \right) ds > -\infty,$$

and further

$$\sup_{x,y \in \mathbb{R}^n} \mathbb{E}_{x,y} \tau_c^{(1)} < \infty. \quad (2.30) \quad \boxed{4.7}$$

Therefore, by Chebyshev's inequality, there exists $t_0 > 0$ such that for any initial point (x, y) ,

$$\mathbb{P}(\tau_c^{(1)} < t_0) \geq \frac{1}{2}. \quad (2.31) \quad \boxed{4.8}$$

Step 3: Define

$$\eta_1 = \inf\{t \geq 0; (\Lambda(t), \Lambda'(t)) \neq (\Lambda(0), \Lambda'(0))\}.$$

By (2.14) and the property of Poisson point process, it is easy to see that $\eta_1 \geq \varsigma_1^{(1)} \wedge \varsigma_1^{(2)}$, where $\varsigma_1^{(1)}$ and $\varsigma_1^{(2)}$ are the first jumping times of the Poisson point processes $(p_1(t))$ and $(p_2(t))$ associated with the Poisson random measures $N_1(dt, dz)$ and $N_2(dt, dz)$ respectively. So

$$\mathbb{P}(\eta_1 \geq t) \geq \mathbb{P}(\varsigma_1^{(1)} \geq t) \mathbb{P}(\varsigma_1^{(2)} \geq t) = e^{-2Mt}, \quad t > 0.$$

Set $\zeta_0 = 0$,

$$\begin{aligned} \zeta_1 &= \inf\{t \geq 0; (\Lambda(t), \Lambda'(t)) \neq (\Lambda(0), \Lambda'(0))\}, \\ \zeta_{2m} &= \inf\{t \geq \zeta_{2m-1}; (\Lambda(t), \Lambda'(t)) = (\Lambda(0), \Lambda'(0))\}, \\ \zeta_{2m+1} &= \inf\{t \geq \zeta_{2m}; (\Lambda(t), \Lambda'(t)) \neq (\Lambda(0), \Lambda'(0))\}, \quad m = 1, 2, \dots \end{aligned}$$

We have the following estimate on the coupling time τ_c :

$$\begin{aligned} \mathbb{P}^{(x,1,y,1)}(\tau_c \in [0, \zeta_1]) &= \mathbb{P}^{(x,1,y,1)}(\tau_c \in [0, \eta_1]) \\ &\geq \mathbb{P}^{(x,1,y,1)}(\eta_1 \geq t_0) \mathbb{P}^{(x,1,y,1)}(\tau_c \in [0, \eta_1] | \eta_1 \geq t_0) \\ &\geq \mathbb{P}^{(x,1,y,1)}(\eta_1 \geq t_0) \mathbb{P}^{(x,y)}(\tau_c^{(1)} < t_0) \\ &\geq e^{-2Mt_0}/2 =: \delta_2 > 0, \end{aligned} \quad (2.32) \quad \boxed{4.9}$$

where t_0 is determined by (2.31) and is independent of the initial point of $(X^{(1)}(t), Y^{(1)}(t))$. Therefore,

$$\begin{aligned}
\mathbb{P}^{(x,i,y,j)}(\tau_c = \infty) &= \mathbb{P}^{(x,i,y,j)}(\mathbf{1}_{\{\tau < \infty\}} \mathbb{P}^{(X(\tau), \Lambda(\tau), Y(\tau), \Lambda'(\tau))}(\tau_c = \infty)) \\
&\leq \mathbb{P}^{(x,i,y,j)}\left(\mathbf{1}_{\{\tau < \infty\}} \mathbb{P}^{(X(\tau), 1, Y(\tau), 1)}\left(\tau_c \notin \bigcup_{m=0}^K [\zeta_{2m}, \zeta_{2m+1})\right)\right) \\
&\leq \mathbb{P}^{(x,i,y,j)}\left(\mathbf{1}_{\{\tau < \infty\}} \mathbb{P}^{(X(\tau), 1, Y(\tau), 1)}\left(\tau_c \notin \bigcup_{m=0}^{K-1} [\zeta_{2m}, \zeta_{2m+1})\right) \right. \\
&\quad \left. \cdot \mathbb{P}^{(X(\zeta_{2K}), 1, Y(\zeta_{2K}), 1)}(\tau_c \notin [0, \zeta_1])\right) \\
&\leq \mathbb{P}^{(x,i,y,j)}\left(\mathbf{1}_{\{\tau < \infty\}} \mathbb{P}^{(X(\tau), 1, Y(\tau), 1)}\left(\tau_c \notin \bigcup_{m=0}^{K-1} [\zeta_{2m}, \zeta_{2m+1})\right)\right)(1 - \delta_2) \\
&\leq (1 - \delta_2)^{K+1},
\end{aligned} \tag{2.33} \tag{4.10}$$

where in the last step we have used the estimate (2.32) recursively. Letting K tend to ∞ , we finally get the desired estimate that $\mathbb{P}^{(x,i,y,j)}(\tau_c = \infty) = 0$, and complete the proof. $\square$

Remark 2.6 In [37], together with F. Xi, we have discussed the question on the existence of successful couplings for state-dependent regime-switching processes. In that work, we imposed a condition (Assumption 2.4 (i) therein) which means in some sense that the corresponding coupling process in every fixed environment is uniformly successful relative to the initial points. This condition is weakened in Lemma 2.5 by supposing only the corresponding coupling process to be uniformly successful in a fixed environment.

Now we introduce the Wasserstein distance used in this work. Define the distance ρ on the space $\mathbb{R}^n \times \mathcal{S}$ by

$$\rho((x, i), (y, j)) = \mathbf{1}_{i \neq j} + |x - y|, \quad x, y \in \mathbb{R}^n, \quad i, j \in \mathcal{S}.$$

The Wasserstein distance between every two probability measures ν_1, ν_2 on $\mathbb{R}^n \times \mathcal{S}$ is defined by

$$W_\rho(\nu_1, \nu_2) = \inf_{\pi \in \mathcal{C}(\nu_1, \nu_2)} \left\{ \int_{(\mathbb{R}^n \times \mathcal{S})^2} \rho((x, i), (y, j)) d\pi((x, i), (y, j)) \right\}, \tag{2.34} \tag{wass}$$

where $\mathcal{C}(\nu_1, \nu_2)$ denotes the set of all couplings of ν_1 and ν_2 on $(\mathbb{R}^n \times \mathcal{S})^2$. This kind of Wasserstein distance has been used in [32] to investigate the recurrent property of regime-switching diffusion process. [12] used further a truncation from above on ρ to define the Wasserstein distance.

t-2.2

Theorem 2.7 *Let $(X(t), \Lambda(t))$ be the solution of (1.1) and (1.2) with $N < \infty$. Under the same conditions of Lemma 2.5, there exists a unique invariant probability measure μ on $\mathbb{R}^n \times \mathcal{S}$ such*

that $\mu P_t = \mu$ for every $t > 0$, and

$$\lim_{t \rightarrow \infty} W_\rho(\delta_{(x,i)} P_t, \mu) = 0 \quad \text{for any } (x, i) \in \mathbb{R}^n \times \mathcal{S}.$$

Proof In order to estimate the Wasserstein distance between $\delta_{(x,i)} P_t$ and $\delta_{(y,j)} P_t$ with $i \neq j$, we use the coupling process determined by (2.13) and (2.14).

For $\kappa \in (0, 1)$, it holds that

$$\begin{aligned} W_\rho(\delta_{(x,i)} P_t, \delta_{(y,j)} P_t) &\leq \mathbb{E}[|X(t) - Y(t)| + \mathbf{1}_{\{\Lambda(t) \neq \Lambda'(t)\}}] \\ &= \mathbb{E}[(|X(t) - Y(t)| + \mathbf{1}_{\{\Lambda(t) \neq \Lambda'(t)\}}) \mathbf{1}_{\{\tau < \kappa t\}}] \\ &\quad + \mathbb{E}[(|X(t) - Y(t)| + \mathbf{1}_{\{\Lambda(t) \neq \Lambda'(t)\}}) \mathbf{1}_{\{\tau \geq \kappa t\}}] \\ &\leq \mathbb{E}[\mathbf{1}_{\{\tau < \kappa t\}} \mathbb{E}[(|X(t) - Y(t)| + \mathbf{1}_{\{\Lambda(t) \neq \Lambda'(t)\}}) | \mathcal{F}_\tau]] \\ &\quad + \mathbb{E}[(1 + |X(t)| + |Y(t)|) \mathbf{1}_{\{\tau \geq \kappa t\}}] \\ &\leq \mathbb{E}[\mathbf{1}_{\{\tau \leq \kappa t\}} \mathbb{E}[|X(t) - Y(t)| | \mathcal{F}_\tau]] + \mathbb{E}[\mathbf{1}_{\{\tau \leq \kappa t\}} \mathbf{1}_{\{\Lambda(t) \neq \Lambda'(t)\}}] \\ &\quad + \sqrt{\mathbb{E}[(1 + |X(t)| + |Y(t)|)^2]} \sqrt{\mathbb{P}(\tau \geq \kappa t)} \\ &\leq c(1 + |x| + |y|)(e^{-\frac{1}{2}\theta\kappa t} + e^{-\frac{\eta\alpha}{2}(1-\kappa)t}) + \mathbb{E}[\mathbf{1}_{\{\tau \leq \kappa t\}} \mathbf{1}_{\{\Lambda(t) \neq \Lambda'(t)\}}], \end{aligned}$$

where in the last step we have used the estimates (2.27), Lemma 2.3, and Lemma 2.4.

Note that after the stopping τ , the processes $(\Lambda(t))$ and $(\Lambda'(t))$ do not necessarily move together due to the dependence of $(q_{ij}(x))$ on the component x . But, after the coupling time τ_c defined in Lemma 2.5, the processes $(X(t), \Lambda(t))$ and $(Y(t), \Lambda'(t))$ must move together. Hence, by Lemma 2.5, we have

$$\begin{aligned} \mathbb{E}[\mathbf{1}_{\{\tau \leq \kappa t\}} \mathbf{1}_{\{\Lambda(t) \neq \Lambda'(t)\}}] &\leq \mathbb{E}[\mathbf{1}_{\{\tau_c \geq t\}} \mathbf{1}_{\{\Lambda(t) \neq \Lambda'(t)\}}] + \mathbb{E}[\mathbf{1}_{\{\tau_c < t\}} \mathbf{1}_{\{\Lambda(t) \neq \Lambda'(t)\}}] \\ &\leq \mathbb{E}[\mathbf{1}_{\{\tau_c \geq t\}}] \longrightarrow 0, \quad \text{as } t \rightarrow \infty. \end{aligned} \tag{2.35} \quad \boxed{5.1}$$

Consequently, we have

$$\lim_{t \rightarrow \infty} W_\rho(\delta_{(x,i)} P_t, \delta_{(y,j)} P_t) = 0. \tag{2.36} \quad \boxed{5.2}$$

According to Lemma 2.4, $\mathbb{E}[|X_t|^2]$ is bounded for all $t > 0$, which yields that the family of probability measures $(\delta_{(x,i)} P_t)_{t>0}$ is tight. Moreover, this yields that $(\delta_{(x,i)} P_t)_{t>0}$ is uniformly integrable w.r.t. the Euclidean metric $|\cdot|$ in $\mathbb{R}^n$. Hence, $(\delta_{(x,i)} P_t)_{t>0}$ is compact in $\mathcal{P}(\mathbb{R}^n)$ w.r.t. the Wasserstein distance W_ρ (cf. for instance, [1, Proposition 7.1.5]). There exists a subsequence $(\delta_{(x,i)} P_{t_k})_{k \geq 1}$ with $t_k \rightarrow \infty$ as $k \rightarrow \infty$ converging to some probability measure μ on $\mathbb{R}^n$. Moreover, (2.36) implies that for all $(y, j) \in \mathbb{R}^n \times \mathcal{S}$, $\delta_{(y,j)} P_t$ converges in W_ρ -metric to μ as $k \rightarrow \infty$, and further that $\nu_0 P_{1,t_k} := \sum_{j \in \mathcal{S}} \int_{\mathbb{R}^n} P_{1,t}^{y,j} \nu_0(dy)$ converges in W_ρ -metric to μ for every probability measure ν_0 on $\mathbb{R}^n$ satisfying $\int_{\mathbb{R}^n} |y| \nu_0(dy) < \infty$. Invoking Lemma 2.3, we get that for every $s > 0$, $\delta_{(x,i)} P_s P_{t_k}$ converges in W_ρ -metric to μ . Since $\delta_{(x,i)} P_s P_{t_k} = \delta_{(x,i)} P_{t_k} P_s$ and the latter term converges weakly to μP_s , this yields that $\mu P_s = \mu$. Hence, μ is the unique invariant measure of the process $(X(t), \Lambda(t))$. $\square$

3 Euler-Maruyama's approximation

Due to the complexity of the regime-switching systems, numerical approximation is frequently an important alternative of closed-form solutions of such systems. Being extremely important, numerical methods have drawn much attention. Starting from the work [40], numerical approximation of state-independent regime-switching processes has been studied. See also [25]. Besides, the approximation of the invariant measures was investigated in [5]. Unlike the state-independent regime-switching diffusions, less result is known for the state-dependent case since the transition rate matrix of the switching process may be different at every jumping step due to its dependence on the continuous-state process. To overcome the complex caused by the mixture of $(\Lambda(t))$ and $(X(t))$, [39] used the local analysis to construct a sequence of discrete-time process to approximate in weak and strong topology to the state-dependent regime-switching diffusions. Later, in [43, Theorem 5.13], the convergence rate of this approximation in strong topology was given. In this work, we investigate a numerical approximation $(X^\delta(t), \Lambda^\delta(t))$, defined by (3.1) and (3.2) below, of the state-dependent RSDP $(X(t), \Lambda(t))$, and show its strong convergence in L^1 -norm. The order of error is also estimated which is consistent with the order obtained in [40] for the state-independent RSDP in suitable sense. Our approximation sequence $(X^\delta(t))$ is just a time discretizing of $(X(t))$, which is a little different to that defined in [39]. Precisely, in lieu of discretizing a Brownian motion, [39] generated a sequence of independent and identically distributed normal random variables to approximate Brownian motion. Meanwhile, [39] used $I + \delta Q(x)$ to approximate $\exp(\delta Q(x))$ to facilitate the computation. However, we use $\exp(\delta Q(x))$ directly in the computation. Similar to the previous section, our approach here also strongly depends on the refined estimate of the switching process based on its Skorokhod's representation.

Consider the following numerical approximate solution to equations (1.1) and (1.2): for $\delta \in (0, 1)$,

$$dY(t) = b(Y(t_\delta), \Lambda'(t_\delta))dt + \sigma(Y(t_\delta), \Lambda'(t_\delta))dW(t), \quad (3.1) \quad \boxed{\text{s-3.1}}$$

$$\Lambda'(t) = i + \int_0^t \int_{[0, M]} h(Y(s_\delta), \Lambda'(s-), z) N_1(ds, dz), \quad (3.2) \quad \boxed{\text{s-3.2}}$$

where $t_\delta = [t/\delta]\delta$, $[t/\delta]$ denotes the integer part of t/δ , $N_1(dt, dz)$ is a Poisson random measure used in (2.1) to determine the process $(\Lambda(t))$ with $\Lambda(0) = i$. Here and in the sequel, for the ease of notation, we use $(Y(t), \Lambda'(t))$ instead of $(X^\delta(t), \Lambda^\delta(t))$ to denote the numerical approximation of $(X(t), \Lambda(t))$ for some given $\delta \in (0, 1)$. Then, by Skorokhod's representation, it holds

$$\mathbb{P}(\Lambda'(t + \Delta) = k | \Lambda'(t) = j, Y(t_\delta) = y) = \begin{cases} q_{jk}(y)\Delta + o(\Delta), & k \neq j, \\ 1 + q_{jj}(y)\Delta + o(\Delta), & k = j, \end{cases} \quad (3.3) \quad \boxed{\text{s-3.2b}}$$

provided $\Delta \downarrow 0$. Set $(Y(0), \Lambda'(0)) = (X(0), \Lambda(0)) = (x, i)$. Note that $(\Lambda'(t))$ is a continuous time jumping process whose transition rate depends on the process $(Y(t))$. In (3.1), the evolution of

$Y(t)$ depends only on the embedded chain $(\Lambda'(k\delta))_{k \geq 1}$ of the process $(\Lambda'(t))$, which coincides with the numerical approximation to state-independent regime-switching process studied in [24, Chapter 4].

In this section, we will use the following conditions on σ and b :

(H1) There exists a constant $C_4 > 0$ such that

$$|b(x, i) - b(y, i)| \leq C_4 |x - y|, \quad x, y \in \mathbb{R}^n, \quad i \in \mathcal{S}.$$

(H2) $\sigma(x, i) \equiv \sigma$ is a constant matrix, and $|b(x, i)| \leq C_5$.

Moreover, it is easy to see that under the conditions (Q1)-(Q3), **(H1)** and **(H2)** the existence of the solution of (3.1) and (3.2) is easily established by considering recursively these equations for $t \in [k\delta, (k+1)\delta)$, $k \geq 0$.

The main difficult and different part to study the time discretizing numerical approximation of state-dependent regime-switching diffusions against the state-independent ones is the requirement of the estimation of the term

$$\int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s)) ds. \quad (3.4) \quad \boxed{\text{o-1}}$$

We shall use Skorokhod's representation to provide a suitable estimate of (3.4). Let us recall Skorokhod's representations of $(\Lambda(t))$ and $(\Lambda'(t))$ defined by (2.1) and (3.2):

$$\begin{aligned} \Lambda(t) &= i + \int_0^t \int_{[0, M]} h(X(s), \Lambda(s-), z) N_1(ds, dz), \\ \Lambda'(t) &= i + \int_0^t \int_{[0, M]} h(Y(s_\delta), \Lambda'(s-), z) N_1(ds, dz). \end{aligned}$$

To make our calculation clear, we present a more concrete construction of the Poisson point process $(p_1(t))$ associated with the Poisson random measure $N_1(dt, dz)$ introduced in Section 1 (cf. for example [27, Chapter 1]).

Let ξ_i , $i = 1, 2, \dots$, be random variables supported on $[0, M]$ satisfying $\mathbb{P}_2(\xi_i \in dx) = \mathbf{m}(dx)/M$. Let τ_i , $i = 1, 2, \dots$, be nonnegative random variables such that $\mathbb{P}_2(\tau_i > t) = \exp(-tM)$, $t \geq 0$. Suppose that (ξ_i) , (τ_i) are mutually independent. Set

$$\begin{aligned} \varsigma_1 &= \tau_1, \quad \varsigma_2 = \tau_1 + \tau_2, \dots, \varsigma_k = \tau_1 + \dots + \tau_k, \dots, \\ D_{p_1} &= \{\varsigma_1, \varsigma_2, \dots, \varsigma_k, \dots\}, \end{aligned}$$

and

$$p_1(\varsigma_k) = \xi_k, \quad k = 1, 2, \dots$$

Then $(p_1(t))$ is a Poisson point process as desired. Let $N(t) = \#\{k; \varsigma_k \leq t\}$ standing for the number of jumps of the process $(p_1(t))$ before time t , which is a Poisson process with rate M .

We also need a preliminary result, which was first shown in [31].

t-3.1 **Lemma 3.1** *Assume the conditions (Q2), (Q3) hold. Denote $A\Delta B = (A \setminus B) \cup (B \setminus A)$ for Borel measurable sets A and B , and $|A\Delta B|$ its Lebesgue measure. Then*

$$|\Gamma_{ij}(x)\Delta\Gamma_{ij}(y)| \leq \tilde{K}|x - y|, \quad \text{for every } i, j \in \mathcal{S}, \quad (3.5) \quad \text{o-2}$$

where $\tilde{K} = 2(N - 1)Nc_q + 1$.

Since the proof of this lemma is similar to that of [31, Lemma 2.1], we omit it here.

t-3.2 **Lemma 3.2** *Assume (Q1)-(Q3) and (H1) hold. Assume that there exists some constant $\|\sigma\|_{\text{Lip}}$ such that*

$$\|\sigma(x, i) - \sigma(y, i)\|_{\text{HS}} \leq \|\sigma\|_{\text{Lip}}|x - y|, \quad x, y \in \mathbb{R}^n, \quad i \in \mathcal{S}. \quad (3.6) \quad \text{sig-lip}$$

Let $(X(t), \Lambda(t))$ and $(Y(t), \Lambda'(t))$ be determined by (1.1), (1.2) and (3.1), (3.2) respectively. Then, for any $t > 0$, there exists a positive constant $C = C(t, \tilde{K}, C_4, \|\sigma\|_{\text{Lip}})$ independent of δ such that

$$\int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s))ds \leq C\delta^{\frac{1}{2}} + C \int_0^t \mathbb{E}|X(s) - Y(s)|ds. \quad (3.7) \quad \text{s-3.9}$$

Proof We divide this proof into three steps, and in the proof we shall use C to denote a generic constant whose value may be different from line to line.

Step 1: For $t \in (0, \delta]$, noting that $\Lambda(0) = \Lambda'(0) = i$, we have

$$\begin{aligned} \mathbb{P}(\Lambda(t) \neq \Lambda'(t)) &= \mathbb{P}(\Lambda(t) \neq \Lambda'(t), N(t) \geq 1) \\ &= \mathbb{P}(\Lambda(t) \neq \Lambda'(t), N(t) = 1) + \mathbb{P}(\Lambda(t) \neq \Lambda'(t), N(t) \geq 2). \end{aligned}$$

For the first term, it is easy to check that there is some $\tilde{C} > 0$ so that

$$\begin{aligned} \mathbb{P}(\Lambda(t) \neq \Lambda'(t), N(t) \geq 2) &\leq \mathbb{P}(N(\delta) \geq 2) = \sum_{k=2}^{\infty} \frac{(M\delta)^k}{k!} e^{-M\delta} \\ &= 1 - e^{-M\delta} - M\delta e^{-M\delta} \leq \tilde{C}\delta^2. \end{aligned} \quad (3.8) \quad \text{s-3.9.5}$$

To deal with the second term, since $\Lambda(0) = \Lambda'(0)$, we get

$$\{\omega; \Lambda(t) \neq \Lambda'(t), N(t) = 1\}$$

$$= \{\omega; \tau_1 \leq t, \tau_1 + \tau_2 > t, p_1(\tau_1) \notin \cup_{j \in \mathcal{S}} (\Gamma_{ij}(X(\tau_1)) \cap \Gamma_{ij}(Y(\tau_1\delta)))\}.$$

Hence,

$$\begin{aligned} \mathbb{P}(\Lambda(t) \neq \Lambda'(t), N(t) = 1) &= \int_0^t \mathbb{P}(\Lambda(t) \neq \Lambda'(t), \tau_1 \in ds, \tau_2 > t) \\ &= \int_0^t \mathbb{P}\left(\xi_1 \notin \bigcup_{j \in \mathcal{S}} (\Gamma_{ij}(X(s)) \cap \Gamma_{ij}(Y(s\delta))), \tau_1 \in ds\right) e^{-M(t-s)}. \end{aligned}$$

By virtue of Lemma 3.1, the Lebesgue measure of $\Gamma_{ij}(x)\Delta\Gamma_{ij}(y)$ can be controlled by $|x - y|$. Hence,

$$\mathbb{P}(\xi_1 \in \Gamma_{ij}(X(s))\Delta\Gamma_{ij}(Y(s\delta)) | \tau_1 \in ds) \leq \frac{\tilde{K}}{M} \mathbb{E}|X(s) - Y(s\delta)|, \quad (3.9) \quad \boxed{\text{s-3.10}}$$

where we have used the fact that both $X(s)$ and $Y(s)$ are independent of ξ_1 under the condition $\tau_1 = s$. Indeed, as $\tau_1 = s \in (0, \delta)$, we have

$$\begin{aligned} X(s) &= x + \int_0^s b(X(r), i)dr + \int_0^s \sigma dW(r), \\ Y(s) &= x + \int_0^s b(x, i)dr + \int_0^s \sigma dW(r). \end{aligned}$$

Above equations show that $X(s)$ and $Y(s)$ are completely determined by $(W(r), 0 \leq r \leq s)$. Then the independence between $(W(t))$ and ξ_1 yields that both $X(s)$ and $Y(s)$ are independent of ξ_1 . Consequently, for $t \in (0, \delta]$,

$$\mathbb{P}(\Lambda(t) \neq \Lambda'(t)) \leq \tilde{C}\delta^2 + \tilde{K} \int_0^\delta \mathbb{E}|X(s) - Y(s\delta)|ds. \quad (3.10) \quad \boxed{\text{s-3.11}}$$

Step 2: We proceed to estimating $\mathbb{P}(\Lambda(k\delta) \neq \Lambda'(k\delta))$ for $k \geq 2$ recursively. Denote by $N([s, t])$ the number of jumps of $(p_1(t))$ during the period of $[s, t]$. Note that $(p_1(t))$ is a stationary point process. Let τ_1^δ be the first jumping time of $(p_1(t))$ after time δ , then τ_1^δ has the same law as τ_1 , i.e. $\mathbb{P}_2(\tau_1^\delta > s) = \exp(-Ms)$ for $s \geq 0$. We have

$$\begin{aligned} &\mathbb{P}(\Lambda(2\delta) \neq \Lambda'(2\delta) | \Lambda(\delta) = \Lambda'(\delta)) \\ &= \mathbb{P}(\Lambda(2\delta) \neq \Lambda'(2\delta), N([\delta, 2\delta]) \geq 2 | \Lambda(\delta) = \Lambda'(\delta)) \\ &\quad + \mathbb{P}(\Lambda(2\delta) \neq \Lambda'(2\delta), N([\delta, 2\delta]) = 1 | \Lambda(\delta) = \Lambda'(\delta)) \\ &\leq \mathbb{P}(N([\delta, 2\delta]) \geq 2) + \mathbb{P}(\Lambda(2\delta) \neq \Lambda'(2\delta), N([\delta, 2\delta]) = 1 | \Lambda(\delta) = \Lambda'(\delta)) \\ &\leq \tilde{C}\delta^2 + \int_\delta^{2\delta} \mathbb{P}(\xi_1 \notin \cup_{j \in \mathcal{S}} (\Gamma_{\Lambda(\delta)j}(X(s)) \cap \Gamma_{\Lambda'(\delta)j}(Y(s\delta))), \tau_1^\delta \in ds) e^{-M(2\delta-s)} \end{aligned}$$

$$\leq \tilde{C}\delta^2 + \tilde{K} \int_{\delta}^{2\delta} \mathbb{E}|X(x) - Y(s_{\delta})|ds$$

Combining with the estimation in Step 1, we obtain that

$$\begin{aligned} & \mathbb{P}(\Lambda(2\delta) \neq \Lambda'(2\delta)) \\ & \leq \mathbb{P}(\Lambda(2\delta) \neq \Lambda'(2\delta) | \Lambda(\delta) = \Lambda'(\delta)) + \mathbb{P}(\Lambda(\delta) \neq \Lambda'(\delta)) \\ & \leq \tilde{K} \int_{\delta}^{2\delta} \mathbb{E}|X(s) - Y(s_{\delta})|ds + \tilde{C}\delta^2 + \mathbb{P}(\Lambda(\delta) \neq \Lambda'(\delta)) \\ & \leq \tilde{K} \int_0^{2\delta} \mathbb{E}|X(s) - Y(s_{\delta})|ds + 2\tilde{C}\delta^2. \end{aligned}$$

Deducing recursively, we have

$$\mathbb{P}(\Lambda(k\delta) \neq \Lambda'(k\delta)) \leq \tilde{K} \int_0^{k\delta} \mathbb{E}|X(s) - Y(s_{\delta})|ds + k\tilde{C}\delta^2, \quad k \geq 1. \quad (3.11) \quad \boxed{\text{s-3.12}}$$

Step 3: For $t > 0$, for the convenience of notation we denote by $t_k = k\delta$ for $k \leq N(t)$ and $t_{N(t)+1} = t$. Then,

$$\begin{aligned} & \int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s) | \Lambda(s_{\delta}) = \Lambda'(s_{\delta}))ds \\ & \leq \int_0^t \left(\tilde{K} \int_0^{\delta} \mathbb{E}|X(s_{\delta} + r) - Y(s_{\delta})|dr + \tilde{C}\delta^2 \right) ds \\ & = \tilde{K} \sum_{k=0}^{N(t)} \int_{t_k}^{t_{k+1}} \int_0^{\delta} \mathbb{E}|X(s_{\delta} + r) - Y(s_{\delta})|dr ds + \tilde{C}\delta^2 t \\ & = \tilde{K} \sum_{k=0}^{N(t)} \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} \mathbb{E}|X(r) - Y(r_{\delta})|dr ds + \tilde{C}\delta^2 t \\ & = \tilde{K}\delta \int_0^t \mathbb{E}|X(s) - Y(s_{\delta})|ds + \tilde{C}\delta^2 t. \end{aligned} \quad (3.12) \quad \boxed{\text{s-3.12.5}}$$

Therefore, by (3.11) and (3.12),

$$\begin{aligned} & \int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s))ds \\ & = \int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s), \Lambda(s_{\delta}) = \Lambda'(s_{\delta}))ds + \int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s), \Lambda(s_{\delta}) \neq \Lambda'(s_{\delta}))ds \\ & \leq \int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s) | \Lambda(s_{\delta}) = \Lambda'(s_{\delta}))ds + \int_0^t \mathbb{P}(\Lambda(s_{\delta}) \neq \Lambda'(s_{\delta}))ds \end{aligned}$$

$$\begin{aligned}
&\leq \int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s) | \Lambda(s_\delta) = \Lambda'(s_\delta)) ds + \sum_{k=0}^K \delta \mathbb{P}(\Lambda(k\delta) \neq \Lambda'(k\delta)) \\
&\leq \int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s) | \Lambda(s_\delta) = \Lambda'(s_\delta)) ds + \sum_{k=0}^K \delta \left[\tilde{K} \int_0^{k\delta} \mathbb{E}|X(s) - Y(s_\delta)| ds + k\tilde{C}\delta^2 \right] \\
&\leq \int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s) | \Lambda(s_\delta) = \Lambda'(s_\delta)) ds + \frac{\tilde{C}\delta(t+\delta)t}{2} + (t+\delta)\tilde{K} \int_0^t \mathbb{E}|X(s) - Y(s_\delta)| ds \\
&\leq \tilde{K}(t+2\delta) \int_0^t \mathbb{E}|X(s) - Y(s_\delta)| ds + \tilde{C}\delta^2 t + \frac{\tilde{C}\delta(t+\delta)t}{2}.
\end{aligned}$$

Furthermore, by the conditions **(H1)** and (3.6), it is standard to show (cf. e.g. [24, Lemma 4.1])

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |Y(s)|^p \right] \leq C, \quad p \geq 1,$$

where $C = C(t, p, C_4, \|\sigma\|_{\text{Lip}})$. Thus, for $s \in (0, t]$,

$$\begin{aligned}
\mathbb{E}|Y(s) - Y(s_\delta)| &\leq \mathbb{E} \int_{s_\delta}^s |b(Y(r_\delta), \Lambda'(r_\delta))| dr + \left(\mathbb{E} \int_{s_\delta}^s \|\sigma(Y(r_\delta), \Lambda'(r_\delta))\|_{\text{HS}}^2 dr \right)^{\frac{1}{2}} \\
&\leq C \int_{s_\delta}^s (1 + |Y(r_\delta)|) dr + \left(\mathbb{E} \int_{s_\delta}^s C(1 + |Y(r_\delta)|^2) dr \right)^{\frac{1}{2}} \\
&\leq C\delta^{\frac{1}{2}}
\end{aligned} \tag{3.13} \quad \boxed{\text{s-3.6}}$$

Invoking the triangle inequality,

$$\int_0^t \mathbb{E}|X(s) - Y(s_\delta)| ds \leq \int_0^t \mathbb{E}|X(s) - Y(s)| ds + C\delta^{\frac{1}{2}}.$$

Consequently,

$$\begin{aligned}
&\int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s)) ds \\
&\leq \tilde{K}(t+2\delta) \int_0^t \mathbb{E}|X(s) - Y(s)| ds + C\tilde{K}(t+2\delta)\delta^{\frac{1}{2}} + \tilde{C}\delta^2 t + \frac{\tilde{C}(t+\delta)t}{2}\delta \\
&\leq C \int_0^t \mathbb{E}|X(s) - Y(s)| ds + C\delta^{\frac{1}{2}},
\end{aligned} \tag{3.14} \quad \boxed{\text{s-3.13}}$$

where $C = C(t, \tilde{K}, C_4, \|\sigma\|_{\text{Lip}})$ independent of δ . The proof is complete. $\square$

t-3.3 **Theorem 3.3** Assume (Q1)-(Q3), **(H1)** and **(H2)** hold. Let $(X(t), \Lambda(t))$ and $(Y(t), \Lambda'(t))$ be determined by (1.1), (1.2) and (3.1), (3.2) respectively. Then for every $T > 0$,

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |X(t) - Y(t)| \right] \leq C\delta^{\frac{1}{2}}, \tag{3.15} \quad \boxed{\text{s-3.3}}$$

for some constant C depending on T and independent of δ . In particular, this implies that

$$\lim_{\delta \rightarrow 0} \mathbb{E} \left[\sup_{0 \leq t \leq T} |X(t) - Y(t)| \right] = 0. \quad (3.16) \quad \boxed{\text{s-3.4}}$$

Proof Let $Z(t) = X(t) - Y(t)$ for $t \geq 0$, then $Z(0) = X(0) - Y(0) = 0$ and

$$Z(t) = \int_0^t b(X(s), \Lambda(s)) - b(Y(s_\delta), \Lambda'(s_\delta)) ds, \quad t > 0.$$

By virtue of **(H1)** and **(H2)**, it holds

$$\begin{aligned} & \mathbb{E} \sup_{0 \leq s \leq t} |Z(s)| \\ & \leq \mathbb{E} \int_0^t |b(X(s), \Lambda(s)) - b(Y(s_\delta), \Lambda'(s_\delta))| ds \\ & \leq \mathbb{E} \int_0^t \left\{ |b(X(s), \Lambda(s)) - b(Y(s), \Lambda(s))| + |b(Y(s), \Lambda(s)) - b(Y(s_\delta), \Lambda(s))| \right. \\ & \quad \left. + |b(Y(s_\delta), \Lambda(s)) - b(Y(s_\delta), \Lambda'(s))| + |b(Y(s_\delta), \Lambda'(s)) - b(Y(s_\delta), \Lambda'(s_\delta))| \right\} ds \\ & \leq \mathbb{E} \int_0^t \left\{ C_4(|Z(s)| + |Y(s) - Y(s_\delta)|) + 2C_5(\mathbf{1}_{\{\Lambda(s) \neq \Lambda'(s)\}} + \mathbf{1}_{\{\Lambda'(s) \neq \Lambda'(s_\delta)\}}) \right\} ds. \end{aligned} \quad (3.17) \quad \boxed{\text{s-3.5}}$$

For $t > 0$, set $K = [t/\delta]$, $t_k = k\delta$ for $k \leq K$ and $t_{K+1} = t$. Then, due to (3.3) and (Q2),

$$\int_0^t \mathbb{E} \mathbf{1}_{\{\Lambda'(s) \neq \Lambda'(s_\delta)\}} ds = \sum_{k=0}^K \int_{t_k}^{t_{k+1}} \mathbb{P}(\Lambda'(s) \neq \Lambda'(t_k)) ds \leq H\delta t + o(\delta). \quad (3.18) \quad \boxed{\text{s-3.7}}$$

According to the Lemma 3.2, there exists a positive constant C depending on t such that

$$\int_0^t \mathbb{P}(\Lambda(s) \neq \Lambda'(s)) ds \leq C\delta^{\frac{1}{2}} + C \int_0^t \mathbb{E}|Z(s)| ds. \quad (3.19) \quad \boxed{\text{s-3.8}}$$

Inserting (3.13), (3.18), (3.19) into (3.17), we obtain

$$\mathbb{E} \left[\sup_{0 \leq s \leq t} |Z(t)| \right] \leq C\delta^{\frac{1}{2}} + C \int_0^t \mathbb{E} \left[\sup_{0 \leq r \leq s} |Z(r)| \right] ds.$$

Then Gronwall's inequality yields that

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} |Z(t)| \right] \leq C(T)\delta^{\frac{1}{2}},$$

which is the desired conclusion. $\square$

References

- [AGS] [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.
- [BBG1996] [2] G.K. Basak, A. Bisi, M.K. Ghosh, Stability of a random diffusion with linear drift, J. Math. Anal. Appl. 202 (1996), 604-622.
- [BBG1999] [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.
- [BS16] [4] J. Bao, J. Shao, Permanence and extinction of regime-switching predator-prey models, SIAM J. Math. Anal. 48 (2016), 725-739.
- [BSY] [5] J. Bao, J. Shao, C. Yuan, Approximation of invariant measures for regime-switching diffusions, Potential Anal. 44 (2016), no. 4, 707-727.
- [BSY17] [6] J. Bao, J. Shao, C. Yuan, Invariant measures for path-dependent random diffusions, arXiv:1706.05638
- [Bar] [7] 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.
- [BL] [8] M. Benaïm, S. Le Borgne, F. Malrieu, P.-A. Zitt, Quantitative ergodicity for some switched dynamical systems, Electron. Commun. Probab. 17 (2012), 1-14.
- [Chen] [9] M.-F. Chen, From Markov chains to non-equilibrium particle systems, 2nd ed. Singapore: World Scientific, 2004.
- [CL] [10] M.-F. Chen, S. Li, Coupling methods for multidimensional diffusion processes, Ann. Probab. 17 (1989), 151-177.
- [CZ] [11] Z.Q. Chen, Z. Zhao, Potential theory for elliptic systems, Ann. Probab. 24 (1996), 293-319.
- [CH] [12] B. Cloez, M. Hairer, Exponential ergodicity for Markov processes with random switching, Bernoulli, 21 (2015), no. 1, 505-536.
- [SY] [13] B. de Saporta, J.-F. Yao, Tail of linear diffusion with Markov switching, Ann. Appl. Probab. 15 (2005), 992-1018.
- [DDY] [14] 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.

- [FC] [15] M. Fragoso, O. Costa, A unified approach for stochastic and mean square stability of continuous-time linear systems with Markovian jumping parameters and additive disturbances, *SIAM J. Control Optim.* 44 (2005), 1165-1191.
- [Gh] [16] M. Ghosh, A. Arapostathis, S. Marcus, Optimal control of switching diffusions with application to flexible manufacturing systems, *SIAM J. Control Optim.* 30 (1992), 1-23.
- [GS] [17] I. Gikhman, A. Skorokhod, The theory of stochastic processes I, II, III, Springer-Verlag, 2004.
- [HYX] [18] T. Hoang, G. Yin, F. Xi, Numerical solutions of regime-switching jump diffusions, *Applied Mathematics and Computation*, 244 (2014), No. 2, 822-835.
- [HS17] [19] T. Hou, J. Shao, Heavy tail and light tail of Cox-Ingersoll-Ross processes with regime-switching, *arXiv:1709.01691*
- [KZY2007] [20] R.Z. Khasminskii, C. Zhu, G. Yin, Stability of regime-switching diffusions, *Stoch. Process. Appl.* 117 (2007), 1037-1051.
- [LR86] [21] T. Lindvall, L.C.G. Rogers, Coupling of multidimensional diffusion processes, *Ann. Probab.* 14 (1986), 860-872.
- [MT] [22] A.J. Majda, X. Tong, Geometric ergodicity for piecewise contracting processes with applications for tropical stochastic lattice models, *Comm. Pure Appl. Math.* 69 (2016), 1110-1153.
- [M13] [23] X. Mao, Stabilization of continuous-time hybrid stochastic differential equations by discrete time feedback control, *Automatica J. IFAC*, 49 (2013), 3677-3681.
- [MY] [24] X. Mao, C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, 2006.
- [MY] [25] X. Mao, C. Yuan, G. Yin, Approximations of Euler-Maruyama type for stochastic differential equations with Markovian switching, under non-Lipschitz conditions. *Journal of Computational and Applied Mathematics*, 205 (2007), 936-948.
- [NY] [26] 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.
- [IW] [27] N. Ikeda, S. Watanabe, *Stochastic differential equations and diffusion processes*, North-Holland Mathematical Library 24. Amsterdam-New York, North-Holland Publ. Co. 1981.
- [PP] [28] M. Pinsky, R. Pinsky, Transience recurrence and central limit theorem behavior for diffusions in random temporal environments, *Ann. Probab.* 21 (1993), 433-452.

- [PS] [29] R. Pinsky, M. Scheutzow, Some remarks and examples concerning the transience and recurrence of random diffusions, *Ann. Inst. Henri. Poincaré*, 28 (1992), 519-536.
- [SX] [30] J. Shao, F. Xi, Stability and recurrence of regime-switching diffusion processes, *SIAM J. Control Optim.* 52 (2014), 3496-3516.
- [Sh15a] [31] J. Shao, Criteria for transience and recurrence of regime-switching diffusion processes, *Electron. J. Probab.* 20 (2015), no. 63, 1-15.
- [Sh15b] [32] J. Shao, Ergodicity of regime-switching diffusions in Wasserstein distances, *Stoch. Proc. Appl.* 125 (2015), 739-758.
- [Sh15c] [33] J. Shao, Strong solutions and strong Feller properties for regime-switching diffusion processes in an infinite state space, *SIAM J. Control Optim.* 53 (2015), no. 4, 2462-2479.
- [Sh17a] [34] J. Shao, Stabilization of regime-switching process by feedback control based on discrete time observations, *SIAM J. Control Optim.* 55 (2017), no. 2, 724-740.
- [Sk] [35] A. Skorokhod, *Asymptotic methods in the theory of stochastic differential equations*, American Mathematical Society, Providence, RI. 1989.
- [Xi] [36] F. Xi, Asymptotic properties of jump-diffusion processes with state-dependent switching, *Stochastic Process. Appl.* 119 (2009), 2198-2221.
- [XS] [37] F. Xi, J. Shao, Successful couplings for diffusion processes with state-dependent switching, *Sci. China Math.* 56 (2013), no. 10, 2135-2144.
- [XZ17] [38] 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), no. 3, 1789-1818.
- [YMYC] [39] G. Yin, X. Mao, C. Yuan, D. Cao, Approximation Methods for Hybrid Diffusion Systems with State-Dependent Switching Processes: Numerical Algorithms and Existence and Uniqueness of Solutions, *SIAM J. Math. Anal.* 41 (2010), 2335-2352.
- [YM04] [40] C. Yuan, X. Mao, Convergence of the Euler-Maruyama method for stochastic differential equations with Markovian switching, *Math. Comput. Simulation*, 64 (2004), 223-235.
- [YML] [41] C. Yuan, X. Mao, J. Lygeros, Stochastic hybrid delay population dynamics: Well-posed models and extinction, *J. Biol. Dyn.* 3 (2009), 1-21.
- [YX2010] [42] G. Yin, F. Xi, Stability of regime-switching jump diffusions, *SIAM J. Contr. Optim.* 48 (2010), 4525-4549.

- [YZ] [43] G. Yin, C. Zhu, Hybrid switching diffusions: properties and applications, Vol. 63, Stochastic Modeling and Applied Probability, Springer, New York. 2010.