



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)**ScienceDirect**

Stochastic Processes and their Applications xx (xxxx) xxx–xxx

**stochastic  
processes  
and their  
applications**
[www.elsevier.com/locate/spa](http://www.elsevier.com/locate/spa)

# Certain properties related to well posedness of switching diffusions

Dang Hai Nguyen<sup>a</sup>, George Yin<sup>a,\*</sup>, Chao Zhu<sup>b</sup>

<sup>a</sup> *Department of Mathematics, Wayne State University, Detroit, MI 48202, United States*

<sup>b</sup> *Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, United States*

Received 10 June 2016; received in revised form 7 January 2017; accepted 6 February 2017

Available online xxxx

## Abstract

This work is devoted to switching diffusions that have two components (a continuous component and a discrete component). Different from the so-called Markovian switching diffusions, in the setup, the discrete component (the switching) depends on the continuous component (the diffusion process). The objective of this paper is to provide a number of properties related to the well posedness. First, the differentiability with respect to initial data of the continuous component is established. Then, further properties including uniform continuity with respect to initial data, and smoothness of certain functionals are obtained. Moreover, Feller property is obtained under only local Lipschitz continuity. Finally, an example of Lotka–Volterra model under regime switching is provided as an illustration.

© 2017 Elsevier B.V. All rights reserved.

MSC: 60J05; 60J60

Keywords: Switching diffusion; Continuous-state dependent switching; Smoothness; Feller property

## 1. Introduction

In the past two decades, a considerable research effort has been devoted to the study of switching diffusion processes that are also called hybrid switching diffusions. Much of the interest stems from pressing need of treating complex systems involving both continuous dynamics representable by using solutions of stochastic differential equations, and discrete

\* Corresponding author.

E-mail addresses: [dangnh.maths@gmail.com](mailto:dangnh.maths@gmail.com) (D.H. Nguyen), [gyin@math.wayne.edu](mailto:gyin@math.wayne.edu) (G. Yin), [zhu@uwm.edu](mailto:zhu@uwm.edu) (C. Zhu).

<http://dx.doi.org/10.1016/j.spa.2017.02.004>

0304-4149/© 2017 Elsevier B.V. All rights reserved.

events that cannot be written as solutions of the usual differential equations. Such hybrid systems are prevalent in a wide range of applications including ecological and biological modeling [21], control systems and filtering [19], economics and finance [9,20], networked systems [8], among others. These switching diffusions can be represented by a two-component process  $(X(t), \alpha(t))$ , where  $X(t)$  is a continuous component (also called “continuous state”) taking values in  $\mathbb{R}$  and  $\alpha(t)$  is a discrete component (also called “discrete state”) taking values in a finite set  $\mathcal{M} = \{1, \dots, m_0\}$ . The interactions of the continuous and discrete components make the models more versatile and suitable for a wide range of applications. On the other hand, these interactions make the analysis of such processes much more difficult. It is interesting to note that such processes are similar to the usual diffusions, but they could behave much differently from the usual diffusion processes qualitatively. For example, suppose that we have two linear diffusions together with a continuous-time Markov chain. The Markov chain serves as a modulating force making the process switch back and forth between these two diffusions. Depending on the switching frequency, even if each of the diffusion is stable, the switched system can be unstable or vice versa; see for example, [15, pp. 229–233]. Likewise, we may have both of the individual diffusions being recurrent, but the switching diffusion is not. In a way, the switching diffusion processes display many peculiar properties.

Because of their importance, hybrid systems in general and switching diffusions in particular have drawn resurgent attentions. For some of the recent progress, we mention the work [2,7,11,13,14,18,22] and references therein. Systematic treatments and comprehensive study of these stochastic processes can be found in [10] and [16]. Both of these references consider systems that have pure jump random switching in addition to the noise processes driven by Brownian motions; the first reference mentioned above concentrates on switching diffusions in which the switching mechanism is given by a continuous-time Markov chain independent of the Brownian motion, whereas the second reference focuses on the switching processes depending on the current state of the diffusions. In what follows, to distinguish these two types of switching diffusions, the first type process is referred to as Markovian switching diffusions (or Markov chain modulated switching diffusions), and the second type process is called continuous-state-dependent switching diffusions.

This paper aims to study continuous and smooth dependence on the initial data of solutions to stochastic differential equations corresponding to continuous-state-dependent switching diffusions. These properties are all related to the well-posedness in certain sense. In the book of Applebaum [1], similar properties for stochastic differential equations, are also referred to as flow properties. These properties vividly highlight the distinctions between Markovian switching diffusions and continuous-state-dependent switching diffusions. For example, it is well-known that a diffusion process is smooth in the  $L^2$  sense with respect to its initial data under suitable conditions; see [6, VII, Section 4]. Such a smoothness property readily carries over to the case of Markovian switching diffusions because the Markov chain is independent of the Brownian motion and the initial state of the continuous component does not influence the dynamics of the Markov chain. In this paper, we show that this phenomenon becomes markedly different in the case of continuous-state-dependent switching diffusions, which is in sharp contrast with that of Markovian switching diffusions and diffusions. In lieu of  $L^2$  convergence, we demonstrate that the smoothness with respect to initial data is in the sense of  $L^p$  for any  $0 < p$  strictly less than  $\lambda$ , with  $\lambda \leq 1$  being the Hölder exponent of the generator for the switching process; see Theorem 3.1 for the precise statement. Moreover, we provide a counterexample to demonstrate that the estimates above is in fact sharp; we cannot expect  $L^1$  convergence, neither can we get  $L^2$  convergence in the current case. Next, we examine uniform estimates of two solutions

if their initial data are close. Furthermore, we establish smoothness of certain functional of switching diffusions. As an application of the uniform estimates, we revisit the issue of Feller property. Although Feller properties for continuous-state-dependent switching diffusions have been established in the literature, see, for example, [13,14] as well as Chapter 2 in [16], this work aims to relax the commonly used global Lipschitz condition and to provide a simple proof.

The rest of the paper is arranged as follows. Section 2 presents the switching diffusion setup. Section 3 focuses on the differentiability of the switching diffusions with respect to the initial data of the continuous state; also presented in this section is a counter-example of the differentiability under  $L^1$  convergence. Section 4 studies further properties, in which Section 4.1 furthers our investigation on uniform estimates on a finite time interval of solutions with different initial data and Section 4.2 investigates smoothness of a functional of the switching diffusions. Using results obtained in Section 3, in Section 5, we first provide an alternative proof of the Feller property under global Lipschitz condition. Then we obtain the Feller property by using only local Lipschitz continuity. Finally, we close the paper with an example in Section 6, a competitive regime-switching Lotka–Volterra model, which illustrates our results.

## 2. Switching diffusions

Let  $(\Omega, \mathcal{F}, \mathcal{F}_t, \mathbb{P})$  be a filtered probability space, where  $\Omega$  is the sample space,  $\mathcal{F}$  is the  $\sigma$ -algebra of subsets,  $\{\mathcal{F}_t\}$  is a filtration (that is,  $\{\mathcal{F}_t\}$  is a family of  $\sigma$ -algebras satisfying  $\mathcal{F}_s \subset \mathcal{F}_t$  for  $s \leq t$ ), and  $\mathbb{P}$  is a probability measure. We assume that  $\mathcal{F}_t$  is complete in that it contains all null sets, and that  $\mathcal{F}_t$  satisfies the usual condition in that  $\mathcal{F}_0$  is complete and  $\{\mathcal{F}_t\}$  is right continuous. Let  $\mathcal{M} = \{1, \dots, m_0\}$  be a finite set, and suppose that  $b(\cdot, \cdot) : \mathbb{R}^r \times \mathcal{M} \rightarrow \mathbb{R}^r$  and  $\sigma(\cdot, \cdot) : \mathbb{R}^r \times \mathcal{M} \rightarrow \mathbb{R}^{r \times d}$ . In this paper, we consider a switching diffusion process, a two-component Markov process  $(X(t), \alpha(t))$  whose generator is given by

$$\begin{aligned} \mathcal{L}f(x, i) &= \nabla f'(x, i)b(x, i) + \text{tr}(\nabla^2 f(x, i)A(x, i)) + Q(x)f(x, \cdot)(i) \\ &= \sum_{k=1}^r b_k(x, i) \frac{\partial f(x, i)}{\partial x_k} + \frac{1}{2} \sum_{k,l=1}^r a_{kl}(x, i) \frac{\partial^2 f(x, i)}{\partial x_k \partial x_l} \\ &\quad + Q(x)f(x, \cdot)(i), \text{ for any } (x, i) \in \mathbb{R}^r \times \mathcal{M}, \end{aligned} \quad (2.1)$$

for a  $C^2$ -function  $f(\cdot, i)$  (whose derivatives with respect to  $x$  up to the second order are continuous) for each  $i \in \mathcal{M}$ , where  $\nabla f(x, i)$  and  $\nabla^2 f(x, i)$  denote the gradient and Hessian of  $f(x, i)$  with respect to  $x$ , respectively, and

$$\begin{aligned} Q(x)f(x, \cdot)(i) &= \sum_{j=1}^{m_0} q_{ij}(x)f(x, j), \text{ and} \\ A(x, i) &= (a_{kl}(x, i)) = \sigma(x, i)\sigma'(x, i) \in \mathbb{R}^{r \times r}. \end{aligned}$$

The dynamics and transition rules of  $(X, \alpha)$  may also be presented as follows. Suppose that  $w(t)$  is an  $\mathbb{R}^d$ -valued Brownian motion,  $\alpha(t)$  is a pure jump process taking value in  $\mathcal{M}$ , and  $X(t)$  satisfies

$$dX(t) = b(X(t), \alpha(t))dt + \sigma(X(t), \alpha(t))dw(t), \quad (2.2)$$

such that the jump intensity of  $\alpha(t)$  depends on the current state of  $X(t)$  in that the generator of  $\alpha(t)$  is given by  $Q(x) = (q_{ij}(x))$  with  $q_{ij}(\cdot) : \mathbb{R}^r \rightarrow \mathbb{R}$  for  $i, j \in \mathcal{M}$ ,  $q_{ij}(x) \geq 0$  for  $i \neq j$ , and  $\sum_j q_{ij}(x) = 0$  for each  $i \in \mathcal{M}$ , satisfying

$$\mathbb{P}\{\alpha(t + \Delta) = j | \alpha(t) = i, X(s), \alpha(s), s \leq t\} = q_{ij}(X(t))\Delta + o(\Delta) \text{ if } i \neq j. \quad (2.3)$$

Alternatively, the evolution of the discrete component  $\alpha(\cdot)$  can be represented by a stochastic integral with respect to a Poisson random measure (see, e.g., [12]). For  $x \in \mathbb{R}^r$  and  $i \in \mathcal{M}$ , let  $\Delta_{ij}(x)$ ,  $j \in \mathcal{M} \setminus \{i\}$  be disjoint intervals of the real line, each having length  $q_{ij}(x)$ . Define a function  $h : \mathbb{R}^r \times \mathcal{M} \times \mathbb{R} \mapsto \mathbb{R}$  by

$$h(x, i, z) = \sum_{j=1}^{m_0} (j - i) I_{\{z \in \Delta_{ij}(x)\}}. \quad (2.4)$$

That is, with the partition  $\{\Delta_{ij}(x) : i, j \in \mathcal{M}\}$  used and for each  $i \in \mathcal{M}$ , if  $z \in \Delta_{ij}(x)$ , then  $h(x, i, z) = j - i$ ; otherwise  $h(x, i, z) = 0$ . Then the dynamics of the discrete component  $\alpha(\cdot)$  can be represented by

$$d\alpha(t) = \int_{\mathbb{R}} h(X(t), \alpha(t-), z) p(dt, dz), \quad (2.5)$$

where  $p(dt, dz)$  is a Poisson random measure with intensity  $dt \times m(dz)$ , and  $m$  is the Lebesgue measure on  $\mathbb{R}$ . The Poisson random measure  $p(\cdot, \cdot)$  is independent of the Brownian motion  $w(\cdot)$ . Thus, the switching process  $(X, \alpha)$  can be presented by the system of stochastic differential equations given by (2.2) and (2.5). Let the initial condition of the switching diffusion be  $(X(0), \alpha(0)) = (x, \alpha) \in \mathbb{R}^r \times \mathcal{M}$ . To ensure the existence and uniqueness of solutions to (2.2) and (2.3), we have the following theorem, which is proved in [16, Theorem 2.7].

**Theorem 2.1.** Assume that  $b(x, i)$  and  $\sigma(x, i)$  are locally Lipschitz in  $x$  for each  $i \in \mathcal{M}$ , and that  $Q(x) = (q_{ij}(x))$  is bounded and continuous. Assume further that there exists a function  $V(\cdot, \cdot) : \mathbb{R}^r \times \mathcal{M} \mapsto \mathbb{R}_+$  that is twice continuously differentiable with respect to  $x \in \mathbb{R}^r$  for each  $i \in \mathcal{M}$  and a constant  $K > 0$  such that

$$\begin{cases} \mathcal{L}V(x, i) \leq KV(x, i) \text{ for all } (x, i) \in \mathbb{R}^r \times \mathcal{M} \\ V_R := \inf\{V(x, i) : |x| \geq R, i \in \mathcal{M}\} \rightarrow \infty \text{ as } R \rightarrow \infty. \end{cases}$$

Then the system given by (2.2) and (2.3) has a unique strong solution for each initial condition.

Note that for Markovian switching diffusions, the switching process is a homogeneous continuous-time Markov chain with a constant matrix as its generator, and the Markov chain and the Brownian motion are independent. In lieu of such a structure, we are dealing with a much more complex system. Starting from the next section, we focus on deriving certain smoothness and continuity properties of the underlying processes.

### 3. Differentiability with respect to initial data

This section is devoted to differentiability with respect to initial data of the switching processes, or equivalently, the smoothness of the solutions of the switched stochastic differential equations with respect to the initial data of the continuous component. This section is divided into two parts. The first part derives the main result. The second part presents a counter example indicating our result is sharp.

#### 3.1. Differentiability

In [17, Theorem 4.2], we stated that  $X(t)$  is twice continuously differentiable in mean square under certain conditions. There is an error in the proof of [17, Lemma 4.3]. In particular, in the last line of equation (4.13) on page 2421 of the aforementioned paper, a diagonal entry term of

the form  $\tilde{q}_{(j,l)(j,l)}(x, \tilde{x})$  in the generator of the coupling was inadvertently left out in our proof, resulting an error in the proof of the said theorem. We correct this error here and demonstrate that the switching diffusions do possess the smoothness properties. However the results about the differentiability in mean square is in general *not achievable* if the generator  $Q(x) = (q_{ij}(x))$  of  $\alpha(t)$  depends on the current state of  $X(t)$ ; see the simple yet illuminating example in Section 3.2. The differentiability in mean square needs to be replaced by differentiability in  $p$ th moment for an appropriate  $p$ . It turns out that the proof is interesting in its own right. In a way, it really displays certain aspects of the salient features of the continuous state-dependent switching processes.

If we consider a Markovian switching diffusion process, the switching times and hence the switching process can be generated beforehand because  $\alpha(t)$  is independent of the Brownian motion. As can be seen in the proof of this section, the main difficulty we face is that the switching times depend on the continuous component  $X(t)$ . To illustrate, consider a one-dimensional  $X(t)$  as an example. Denote the solution of (2.2) with two different initial data for the continuous component  $(X(0), \alpha(0)) = (x, \alpha)$  and  $(X(0), \alpha(0)) = (\tilde{x}, \alpha) = (x + \Delta, \alpha)$  by  $X^{x,\alpha}(t)$  and  $X^{\tilde{x},\alpha}(t)$ , respectively. The differentiability is concerned with the limit of the difference quotient  $\frac{X^{\tilde{x},\alpha}(t) - X^{x,\alpha}(t)}{\tilde{x} - x}$ . In the Markovian switching diffusions, the difference  $X^{\tilde{x},\alpha}(t) - X^{x,\alpha}(t)$  can be calculated much the same way as in the diffusion case in [6, VIII, Section 4, pp. 403–412]. That is, we can simply subtract one from the other. The  $\alpha(t)$  does not really come into the picture because even  $x \neq \tilde{x}$ , the sample paths of  $\alpha(t)$  are the same. For our continuous-state dependent switching, care must be taken. The analysis is more delicate in places because  $\alpha^{x,\alpha}(t)$  and  $\alpha^{\tilde{x},\alpha}(t)$  can take different values infinitely often. In the analysis to follow, one of the main insights is the use of the first time when the switching processes  $\alpha^{x,\alpha}(t)$  and  $\alpha^{\tilde{x},\alpha}(t)$  are different.

To proceed, we first setup the notation. For a multi-index or a vector  $\beta = (\beta_1, \dots, \beta_r)$  with nonnegative integer entries, put  $|\beta| = \sum_{i=1}^r \beta_i$  and define

$$D_x^\beta = \frac{\partial^{|\beta|}}{\partial x_1^{\beta_1} \dots \partial x_r^{\beta_r}}.$$

Next we define the  $L^p$  differentiability of a smooth random function  $\Phi(x_1, \dots, x_r, t)$ . Its partial derivative in probability with respect to  $x_i$  is defined as a random variable  $\Psi(x_1, \dots, x_r, t)$  such that

$$\begin{aligned} & \frac{1}{h} \left( \Phi(x_1, \dots, x_i + h, \dots, x_r, t) - \Phi(x_1, \dots, x_i, \dots, x_r, t) \right) \\ & \rightarrow \Psi(x_1, \dots, x_r, t) \text{ in probability} \end{aligned}$$

as  $h \rightarrow 0$ . If for some  $p > 0$ ,

$$\begin{aligned} & \mathbb{E} \left| \frac{1}{h} \left( \Phi(x_1, \dots, x_i + h, \dots, x_r, t) - \Phi(x_1, \dots, x_i, \dots, x_r, t) \right) - \Psi(x_1, \dots, x_r, t) \right|^p \\ & \rightarrow 0 \text{ as } h \rightarrow 0 \end{aligned}$$

we say that  $\Phi(x_1, \dots, x_r, t)$  has partial derivative with respect to  $x_i$  in  $L^p$ . We proceed to obtain the smoothness of the switching diffusion with respect to the initial data in the  $L^p$  sense.

**Theorem 3.1.** Assume that  $b(x, i)$  and  $\sigma(x, i)$  are Lipschitz in  $x$  for each  $i \in \mathcal{M}$ . Let  $(X^{x,\alpha}(t), \alpha^{x,\alpha}(t))$  be the solution to the system given by (2.2) and (2.3) with initial data  $(X(0), \alpha(0)) = (x, \alpha)$ . Assume that for each  $i \in \mathcal{M}$ ,  $b(\cdot, i)$  and  $\sigma(\cdot, i)$  have continuous partial

derivatives with respect to the variable  $x$  up to the second order and that

$$|D_x^\beta b(x, i)| + |D_x^\beta \sigma(x, i)| \leq K_0(1 + |x|^{\gamma_0}), \quad (3.1)$$

where  $K_0$  and  $\gamma_0$  are positive constants and  $\beta$  is a multi-index with  $|\beta| \leq 2$ .

- (a) Suppose that  $Q(x) = (q_{ij}(x))$  is bounded and continuous. Then  $X^{x,\alpha}(t)$  is twice continuously differentiable with respect to  $x$  in probability.
- (b) Replace the conditions for  $Q(x)$  in the assumptions of part (a) by  $Q(x)$  is bounded and  $q_{kj}(x)$  is locally Hölder continuous with Hölder exponent  $\lambda$  for some  $\lambda \in (0, 1]$  in that there are  $K_1 > 0$  and  $\gamma_1 > 0$  such that

$$\sum_{k,j \in \mathcal{M}} |q_{kj}(x) - q_{kj}(y)| \leq K_1(1 + |x|^{\gamma_1} + |y|^{\gamma_1})|x - y|^\lambda, \quad \forall x, y \in \mathbb{R}^r. \quad (3.2)$$

Then  $X^{x,\alpha}(t)$  is twice continuously differentiable in  $L^p$  with respect to  $x$  for any  $0 < p < \lambda$ .

**Proof.** For ease of presentation and without loss of generality, we prove the result when  $X(t)$  is one dimensional. Fix  $(x, \alpha) \in \mathbb{R} \times \mathcal{M}$  and  $T > 0$ . Let  $(X(t), \alpha(t))$  be the switching-diffusion process satisfying (2.2) and (2.3) with initial condition  $(x, \alpha)$  respectively. Likewise, let  $(\tilde{X}(t), \tilde{\alpha}(t))$  be the solution process with initial condition  $(\tilde{x}, \alpha)$ , where  $\tilde{x} = x + \Delta$  for  $0 < |\Delta| \ll 1$ . Consider the joint process  $(X(t), \tilde{X}(t), \alpha(t), \tilde{\alpha}(t))$ . By the basic coupling method (see e.g., [3, p. 11]), we can consider them as the solutions to

$$\begin{cases} dX(t) = b(X(t), \alpha(t))dt + \sigma(X(t), \alpha(t))dw(t) \\ d\tilde{X}(t) = b(\tilde{X}(t), \tilde{\alpha}(t))dt + \sigma(\tilde{X}(t), \tilde{\alpha}(t))dw(t) \end{cases} \quad (3.3)$$

with initial condition  $(x, x + \Delta, \alpha, \alpha)$ , where  $(\alpha(t), \tilde{\alpha}(t))$  has the generator  $\tilde{Q}(X(t), \tilde{X}(t))$  which is defined by

$$\begin{aligned} \tilde{Q}(x, \tilde{x})\tilde{f}(k, l) &= \sum_{(j,i) \in \mathcal{M} \times \mathcal{M}} \tilde{q}_{(k,l)(j,i)}(x, \tilde{x}) (\tilde{f}(j, i) - \tilde{f}(k, l)) \\ &= \sum_{j \in \mathcal{M}} [q_{kj}(x) - q_{lj}(\tilde{x})]^+ (\tilde{f}(j, l) - \tilde{f}(k, l)) \\ &\quad + \sum_{j \in \mathcal{M}} [q_{lj}(\tilde{x}) - q_{kj}(x)]^+ (\tilde{f}(k, j) - \tilde{f}(k, l)) \\ &\quad + \sum_{j \in \mathcal{M}} [q_{kj}(x) \wedge q_{lj}(\tilde{x})] (\tilde{f}(j, j) - \tilde{f}(k, l)). \end{aligned} \quad (3.4)$$

Define

$$\tau^\Delta = \inf\{t \geq 0 : \alpha(t) \neq \tilde{\alpha}(t)\}. \quad (3.5)$$

Since  $\alpha(t)$  and  $\tilde{\alpha}(t)$  are the same up to  $\tau^\Delta$ , and  $b(\cdot, i)$  and  $\sigma(\cdot, i)$  are Lipschitz continuous for each  $i \in \mathcal{M}$ , by standard arguments (see e.g. [10, Lemma 3.3]), we obtain that

$$\mathbb{E} \sup_{0 \leq t \leq T \wedge \tau^\Delta} \{|X(t) - \tilde{X}(t)|^2\} \leq K|x - \tilde{x}|^2. \quad (3.6)$$

Recall from [17, Lemma 3.2] that for any  $m > 0$  and  $0 < R < \infty$ , there is a  $C_m(R, T) > 0$  such that for any pair  $(x_0, \alpha_0)$  with  $|x_0| \leq R$  and  $\alpha_0 \in \mathcal{M}$ ,

$$\mathbb{E} \sup_{t \in [0, T]} \{|X^{x_0, \alpha_0}(t)|^m\} \leq C_m(R, T). \quad (3.7)$$

First, we show that

$$\mathbb{P}\{\tau^\Delta \leq T\} \rightarrow 0 \text{ as } \Delta \rightarrow 0.$$

Let  $\tilde{f}(k, l) = \mathbf{1}_{\{k \neq l\}}$ , where  $\mathbf{1}_A$  is the indicator of the set  $A$ . By the definition of the function  $\tilde{f}$ , we have

$$\begin{aligned} \tilde{Q}(x, \tilde{x})\tilde{f}(k, k) &= \sum_{j \in \mathcal{M}, j \neq k} [q_{kj}(x) - q_{kj}(\tilde{x})]^+ + \sum_{j \in \mathcal{M}, j \neq k} [q_{kj}(\tilde{x}) - q_{kj}(x)]^+ \\ &= \sum_{j \in \mathcal{M}, j \neq k} |q_{kj}(x) - q_{kj}(\tilde{x})| =: \rho(x, \tilde{x}, k). \end{aligned} \quad (3.8)$$

Applying the generalized Itô formula to (3.3) and noting that  $\alpha(t) = \tilde{\alpha}(t)$ ,  $t < \tau^\Delta$ , we obtain that

$$\begin{aligned} \mathbb{P}\{\tau^\Delta \leq T\} &= \mathbb{E}\tilde{f}(\alpha(T \wedge \tau^\Delta), \tilde{\alpha}(T \wedge \tau^\Delta)) \\ &= \mathbb{E} \int_0^{T \wedge \tau^\Delta} \tilde{Q}(X(t), \tilde{X}(t))\tilde{f}(\alpha(t), \tilde{\alpha}(t))dt \\ &= \mathbb{E} \int_0^{T \wedge \tau^\Delta} \rho(X(t), \tilde{X}(t), \alpha(t))dt. \end{aligned} \quad (3.9)$$

In view of (3.6),

$$\sup_{t \in [0, T \wedge \tau^\Delta]} |X(t) - \tilde{X}(t)| \rightarrow 0 \text{ in probability as } \Delta \rightarrow 0. \quad (3.10)$$

By the dominated convergence theorem, it follows from (3.10) and boundedness of  $\rho(\cdot, \cdot, \cdot)$  that

$$\lim_{\Delta \rightarrow 0} \mathbb{P}\{\tau^\Delta \leq T\} = \lim_{\Delta \rightarrow 0} \mathbb{E} \int_0^{T \wedge \tau^\Delta} \rho(X(t), \tilde{X}(t), \alpha(t))dt = 0. \quad (3.11)$$

Moreover, if (3.2) is satisfied, then by Hölder's inequality, estimates (3.2), and (3.7), we have

$$\begin{aligned} \mathbb{P}\{\tau^\Delta \leq T\} &= \mathbb{E} \int_0^{T \wedge \tau^\Delta} \rho(X(t), \tilde{X}(t), \alpha(t))dt \\ &\leq \mathbb{E} \int_0^{T \wedge \tau^\Delta} K_1(1 + |X(t)|^{\gamma_1} + |\tilde{X}(t)|^{\gamma_1})|X(t) - \tilde{X}(t)|^\lambda dt \\ &\leq K_2 T \mathbb{E} \sup_{t \leq T \wedge \tau^\Delta} (1 + |X(t)|^{\gamma_1} + |\tilde{X}(t)|^{\gamma_1})|X(t) - \tilde{X}(t)|^\lambda \text{ (for some } K_2 > 0) \\ &\leq K_2 T \left( \mathbb{E} \sup_{t \leq T \wedge \tau^\Delta} (1 + |X(t)|^{\gamma_1} + |\tilde{X}(t)|^{\gamma_1})^{2\frac{2-\lambda}{2-\lambda}} \right)^{\frac{2-\lambda}{2}} \left( \mathbb{E} \sup_{t \leq T \wedge \tau^\Delta} |X(t) - \tilde{X}(t)|^2 \right)^{\frac{\lambda}{2}} \\ &\leq \tilde{K}_\lambda |\Delta|^\lambda \text{ (for some } \tilde{K}_\lambda > 0). \end{aligned} \quad (3.12)$$

As in [17], put  $Z^\Delta(t) := \frac{\tilde{X}(t) - X(t)}{\Delta}$  for  $t \geq 0$ . Then we have

$$\begin{aligned} Z^\Delta(t \wedge \tau^\Delta) &= 1 + \frac{1}{\Delta} \int_0^{t \wedge \tau^\Delta} [b(\tilde{X}(s), \alpha(s)) - b(X(s), \alpha(s))]ds \\ &\quad + \frac{1}{\Delta} \int_0^{t \wedge \tau^\Delta} [\sigma(\tilde{X}(s), \alpha(s)) - \sigma(X(s), \alpha(s))]dw(s). \end{aligned}$$



Now, we evaluate the drift:

$$\begin{aligned}
 & \frac{1}{\Delta} \int_0^{t \wedge \tau^\Delta} [b(\tilde{X}(s), \alpha(s)) - b(X(s), \alpha(s))] ds \\
 &= \frac{1}{\Delta} \int_0^t \mathbf{1}_{\{s \leq \tau^\Delta\}} [b(\tilde{X}(s), \alpha(s)) - b(X(s), \alpha(s))] ds \\
 &= \frac{1}{\Delta} \int_0^t \mathbf{1}_{\{s \leq \tau^\Delta\}} \left( \int_0^1 \frac{d}{dv} b(X(s) + v(\tilde{X}(s) - X(s)), \alpha(s)) dv \right) ds \\
 &= \int_0^t Z^\Delta(s) \mathbf{1}_{\{s \leq \tau^\Delta\}} \int_0^1 b_x(X(s) + v(\tilde{X}(s) - X(s)), \alpha(s)) dv ds. \\
 &= \int_0^t Z^\Delta(s \wedge \tau^\Delta) \mathbf{1}_{\{s \leq \tau^\Delta\}} \int_0^1 b_x(X(s) + v(\tilde{X}(s) - X(s)), \alpha(s)) dv ds. \quad (3.13)
 \end{aligned}$$

In view of (3.13) and a similar evaluation for the diffusion part, we have that  $U_\Delta(t) = Z(t \wedge \tau^\Delta)$  satisfies the equation

$$U_\Delta(t) = 1 + \int_0^t A_\Delta(s) U_\Delta(s) ds + \int_0^t B_\Delta(s) U_\Delta(s) dw(s)$$

where

$$\begin{aligned}
 A_\Delta(s) &:= \mathbf{1}_{\{s \leq \tau^\Delta\}} \int_0^1 b_x(X(s) + v(\tilde{X}(s) - X(s)), \alpha(s)) dv, \\
 B_\Delta(s) &:= \mathbf{1}_{\{s \leq \tau^\Delta\}} \int_0^1 \sigma_x(X(s) + v(\tilde{X}(s) - X(s)), \alpha(s)) dv.
 \end{aligned}$$

By (3.10) and (3.11), as  $\Delta \rightarrow 0$ ,

$$A_\Delta(s) \rightarrow b_x(X(s), \alpha(s)) \text{ and } B_\Delta(s) \rightarrow \sigma_x(X(s), \alpha(s)) \text{ in probability.}$$

In view of [5, Theorem 5.2.2],

$$\mathbb{E} \left| \frac{\tilde{X}(T \wedge \tau^\Delta) - X(T \wedge \tau^\Delta)}{\Delta} - \xi(T) \right|^2 = \mathbb{E} |U_\Delta(t) - \xi(T)|^2 \rightarrow 0 \text{ as } \Delta \rightarrow 0 \quad (3.14)$$

where  $\xi(t)$  is the solution to

$$\xi(t) = 1 + \int_0^t b_x(X(s), \alpha(s)) \xi(s) ds + \int_0^t \sigma_x(X(s), \alpha(s)) \xi(s) dw(s). \quad (3.15)$$

Note that

$$\begin{aligned}
 \frac{\tilde{X}(T) - X(T)}{\Delta} &= \frac{\tilde{X}(T \wedge \tau^\Delta) - X(T \wedge \tau^\Delta)}{\Delta} \\
 &\quad + \mathbf{1}_{\{\tau^\Delta \leq T\}} \frac{\tilde{X}(T) - \tilde{X}(T \wedge \tau^\Delta) - X(T) + X(T \wedge \tau^\Delta)}{\Delta}. \quad (3.16)
 \end{aligned}$$

In light of (3.11),

$$\mathbf{1}_{\{\tau^\Delta \leq T\}} \frac{\tilde{X}(T) - \tilde{X}(T \wedge \tau^\Delta) - X(T) + X(T \wedge \tau^\Delta)}{\Delta} \rightarrow 0 \text{ in probability as } \Delta \rightarrow 0. \quad (3.17)$$

It follows from (3.14), (3.16), and (3.17) that

$$\frac{\tilde{X}(T) - X(T)}{\Delta} \rightarrow \xi(T) \text{ in probability as } \Delta \rightarrow 0. \quad (3.18)$$

This proves part (a) of the theorem.



Next we prove part (b) of the theorem. For  $p < \lambda$ , letting  $\theta = \frac{\lambda-p}{2}$ , we have

$$\begin{aligned} & \mathbb{E} \mathbf{1}_{\{\tau^\Delta \leq T\}} |\tilde{X}(T) - \tilde{X}(T \wedge \tau^\Delta) - X(T) + X(T \wedge \tau^\Delta)|^p \\ & \leq 2^p \mathbb{E} \mathbf{1}_{\{\tau^\Delta \leq T\}} \left( \sup_{t \in [0, T]} \{|X(t)| + |\tilde{X}(t)|\} \right)^p \\ & \leq 2^p (\mathbb{E} \mathbf{1}_{\{\tau^\Delta \leq T\}})^{\frac{p+\theta}{\lambda}} \left( \mathbb{E} \left( \sup_{t \in [0, T]} \{|X(t)| + |\tilde{X}(t)|\} \right)^{\frac{p\lambda}{\theta}} \right)^{\frac{\theta}{\lambda}} \\ & \leq \tilde{K}_2 |\Delta|^{p+\theta} \text{ (for some } \tilde{K}_2 \text{ using (3.7) and (3.12)),} \end{aligned} \quad (3.19)$$

which implies that

$$\mathbb{E} \mathbf{1}_{\{\tau^\Delta \leq T\}} \left| \frac{\tilde{X}(T) - \tilde{X}(T \wedge \tau^\Delta) - X(T) + X(T \wedge \tau^\Delta)}{\Delta} \right|^p \rightarrow 0 \text{ as } \Delta \rightarrow 0. \quad (3.20)$$

As a result of (3.14), (3.16) and (3.20), we obtain that

$$\mathbb{E} \left| \frac{\tilde{X}(T) - X(T)}{\Delta} \rightarrow \xi(T) \right|^p \rightarrow 0 \text{ as } \Delta \rightarrow 0. \quad (3.21)$$

Thus,  $X(t)$  is differentiable in  $L^p$  for  $p < \lambda$  if (3.2) is satisfied. Likewise, it can be shown that  $X(t)$  is twice differentiable in probability and in  $L^p$  for  $p < \lambda$  under the condition (3.2).  $\square$

**Remark 3.2.** In the proof, we use the boundedness condition on  $Q(x)$  to derive (3.11) from (3.10) using the dominated convergence theorem. If we assume the polynomial growth  $|q_{ij}(x)| \leq K_0(1 + |x|^{\gamma_0})$ , we will have  $\rho(x, \tilde{x}, k) \leq \tilde{K}_0(1 + |x| + |\tilde{x}|)^{\gamma_0}$  for some  $\tilde{K}_0 > 0$ . Thus, we have

$$\mathbb{E} \left( \int_0^{T \wedge \tau^\Delta} \rho(X(t), \tilde{X}(t), \alpha(t)) dt \right)^2 \leq T^2 \tilde{K}_0^2 \mathbb{E} \sup_{t \in [0, T]} (1 + |X(t)| + |\tilde{X}(t)|)^{2\gamma_0},$$

which results in the uniform integrability of  $\left\{ \int_0^{T \wedge \tau^\Delta} \rho(X(t), \tilde{X}(t), \alpha(t)) dt : \Delta \in (0, 1] \right\}$ . As a result, we derive from (3.10) and the Vitali convergence theorem that (3.11) still holds if we replace the boundedness of  $q_{ij}(x)$  by the condition  $|q_{ij}(x)| \leq K_0(1 + |x|^{\gamma_0})$ . All the remaining arguments in the proof of Theorem 3.1 still carry over.

### 3.2. A counterexample for smoothness under $L^1$ convergence

The smoothness with respect to the initial data was obtained in the last section. The convergence is in the sense of  $L^p$  convergence. One immediate question is: Can we do better? Is it possible to get, for example,  $L^1$  convergence? In general, this question has a negative answer. The intuitive arguments are as follows.  $X(t)$  and  $\tilde{X}(t)$  evolve similarly up to  $\tau^\Delta$ , the moment  $\alpha(t)$  and  $\tilde{\alpha}(t)$  switch apart. Thus, if  $T \leq \tau^\Delta$ , it is easy to estimate  $\frac{X(T) - \tilde{X}(T)}{\Delta}$  using the (local) Lipschitz condition of the coefficients. However, if  $T > \tau^\Delta$ , we cannot expect that  $\frac{X(T) - \tilde{X}(T)}{\Delta}$  is bounded since  $X(t)$  and  $\tilde{X}(t)$  follow different equations after  $\tau^\Delta$ . Using this observation, in the proof of Theorem 3.1, we decompose  $\frac{X(T) - \tilde{X}(T)}{\Delta}$  into two parts in (3.16). The first term on the right-hand side of (3.16) converges to  $\xi(T)$  in  $L^2$  while the second term converges to 0 in probability. Thus, if  $\frac{X(T) - \tilde{X}(T)}{\Delta}$  were to converge in  $L^1$  as  $\Delta \rightarrow 0$ , we would expect that

$$\mathbf{1}_{\{\tau^\Delta \leq T\}} \frac{\tilde{X}(T) - \tilde{X}(T \wedge \tau^\Delta) - X(T) + X(T \wedge \tau^\Delta)}{\Delta} \rightarrow 0 \text{ in } L^1.$$

Since  $X(t)$  and  $\tilde{X}(t)$  evolve completely differently after  $\tau^\Delta$ , we cannot expect that  $X(T) - \tilde{X}(T) \rightarrow 0$  as  $\Delta \rightarrow 0$  in the event  $\{T > \tau^\Delta\}$ . Thus,

$$\frac{\tilde{X}(T) - \tilde{X}(T \wedge \tau^\Delta) - X(T) + X(T \wedge \tau^\Delta)}{\Delta} = O(\Delta^{-1}) \text{ as } \Delta \rightarrow 0 \text{ if } T > \tau^\Delta. \quad (3.22)$$

On the other hand, it follows from (3.8), (3.9), and the Hölder continuity (3.2) that

$$\mathbb{P}\{T \geq \tau^\Delta\} = O(\Delta^\lambda) \text{ as } \Delta \rightarrow 0. \quad (3.23)$$

As a result of (3.22) and (3.23), we can see that  $\mathbf{1}_{\{\tau^\Delta \leq T\}} \frac{\tilde{X}(T) - \tilde{X}(T \wedge \tau^\Delta) - X(T) + X(T \wedge \tau^\Delta)}{\Delta}$  does not in general converge to 0 in  $L^p$  for  $p \geq \lambda$ . Because the Hölder exponent  $\lambda$  cannot exceed 1 except for the case when the  $Q$ -matrix is constant, we cannot, in general, obtain the  $L^1$  convergence.

In this section, we provide an example showing that the process  $X^{x,\alpha}(t)$  is not differentiable in  $L^2$  or in  $L^1$  although the coefficients  $b(\cdot, \cdot)$ ,  $\sigma(\cdot, \cdot)$  and  $Q(x)$  are smooth and bounded. In order to simplify the calculations, we consider an example in which the  $Q$ -matrix is reducible and there is no diffusion part. It is possible to construct a counter-example with irreducible  $Q$ -matrix and nondegenerate diffusion. However, it will involve more cumbersome and tedious calculations. It appears to be more instructive to construct an example with structure as simple as possible to highlight the distinctions of  $x$ -dependent switching and Markovian switching. Thus, we omit such an example here.

Let  $\mathcal{M} = \{1, 2\}$  and consider the equation

$$dX(t) = b(\alpha(t))dt \quad (3.24)$$

where  $b(1) = 0$  and  $b(2) = 1$ . Suppose that the switching process  $\alpha(t)$  has generator  $Q = \begin{pmatrix} -f(x) & f(x) \\ 0 & 0 \end{pmatrix}$ , where  $f(x)$  is a smooth positive function with compact support and  $f(x) = x$  for  $x \in [1, 2]$ . Let  $\Delta_{12}(x) = [0, f(x))$ ,  $\Delta_{21}(x) = \emptyset$ . The process  $\alpha(t)$  can be defined as the solution to

$$d\alpha(t) = \int_{\mathbb{R}} h(X(t), \alpha(t-), z) p(dt, dz)$$

where  $h(x, 1, z) = \mathbf{1}_{\{z \in \Delta_{12}(x)\}}$  and  $h(x, 2, z) = 0$ ,  $p(dt, dz)$  is a Poisson random measure with intensity  $dt \times m(dz)$  and  $m$  is the Lebesgue measure on  $\mathbb{R}$ .

Let  $y > 0$  and  $(X^{y,1}(t), \alpha^{y,1}(t))$  be the solution with initial data  $(y, 1)$ . Let  $\tau^y = \inf\{t \geq 0 : \alpha^{y,1}(t) = 2\}$ . For  $x \in [0, 1]$  we have  $X^{1+x,1}(t) = 1 + x$ ,  $\forall t \in [0, \tau^{1+x}]$ , and  $\alpha^{1+x}(t)$  stays in state 2 once it jumps into 2 since the jump intensity  $q_{21}(x) = 0$  for any  $x \in \mathbb{R}$ . Thus,

$$X^{1+x,1}(T) = 1 + x + T - T \wedge \tau^{x+1} \text{ and} \\ X^{1,1}(T) = 1 + T - T \wedge \tau^1.$$

As a result,

$$Z^{1+x} := \frac{X^{1+x,1}(T) - X^{1,1}(T)}{x} = 1 + \frac{[T \wedge \tau^1 - T \wedge \tau^{x+1}]}{x}. \quad (3.25)$$

If  $Z^{1+x}$  converges in  $L^1$  to a variable  $Z_0$ , then the sequence  $Z^{1+\frac{1}{n}}$  must be a Cauchy sequence in  $L^1$ . Then it follows that,

$$\mathbb{E}|Z^{1+\frac{2}{n}} - Z^{1+\frac{1}{n}}| \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (3.26)$$

Now, we show that (3.26) cannot be satisfied. Let

$$\begin{aligned} Y_0 &= \inf \left\{ t \geq 0 : \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \in [0,1)\}} \mathbf{p}(ds, dz) \neq 0 \right\}, \\ Y_1 &= \inf \left\{ t \geq 0 : \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \in [1, 1+\frac{1}{n})\}} \mathbf{p}(ds, dz) \neq 0 \right\}, \\ Y_2 &= \inf \left\{ t \geq 0 : \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \in [1+\frac{1}{n}, 1+\frac{2}{n})\}} \mathbf{p}(ds, dz) \neq 0 \right\}. \end{aligned}$$

Since  $[0, 1)$ ,  $[1, 1 + \frac{1}{n})$ ,  $[1 + \frac{1}{n}, 1 + \frac{2}{n})$  are disjoint sets, we have that  $Y_0, Y_1, Y_2$  are three independent exponential random variables with parameter  $1, \frac{1}{n}, \frac{1}{n}$ , respectively. Note that

$$\tau^{1+x} = \inf \left\{ t \geq 0 : \int_0^t \int_{\mathbb{R}} \mathbf{1}_{\{z \in [0, 1+x)\}} \mathbf{p}(ds, dz) \neq 0 \right\}. \quad (3.27)$$

Thus,  $\tau^1 = Y_0$ ,  $\tau^{1+\frac{1}{n}} = Y_0 \wedge Y_1$ , and  $\tau^{1+\frac{2}{n}} = Y_0 \wedge Y_1 \wedge Y_2$ . From (3.25), we have

$$\left| Z^{1+\frac{2}{n}} - Z^{1+\frac{1}{n}} \right| = \left| n[T \wedge \tau^1 - T \wedge \tau^{1+\frac{1}{n}}] - \frac{n}{2}[T \wedge \tau^1 - T \wedge \tau^{1+\frac{2}{n}}] \right|. \quad (3.28)$$

Let  $A = \left\{ Y_1 \leq Y_2, Y_1 < \frac{T}{3}, Y_0 \in \left[ \frac{2T}{3}, T \right] \right\}$ . When  $\omega \in A$ ,  $T \wedge \tau^1 = \tau^1 = Y_0$ ,  $T \wedge \tau^{1+\frac{1}{n}} = T \wedge \tau^{1+\frac{2}{n}} = Y_1$ . As a result, when  $\omega \in A$ ,

$$\left| Z^{1+\frac{2}{n}} - Z^{1+\frac{1}{n}} \right| = \left| n(Y_0 - Y_1) - \frac{n}{2}(Y_0 - Y_1) \right| = \frac{n}{2}(Y_0 - Y_1) \geq \frac{Tn}{6}. \quad (3.29)$$

By direct calculation,

$$\begin{aligned} \mathbb{P}(A) &= \mathbb{P} \left\{ Y_1 \leq Y_2, Y_1 < \frac{T}{3} \right\} \times \mathbb{P} \left\{ Y_0 \in \left[ \frac{2T}{3}, T \right] \right\} \\ &= \frac{1}{n} \left( 1 - \exp \left( - \left( 1 + \frac{2}{n} \right) \frac{T}{3} \right) \right) \times \left( \exp \left( - \frac{2T}{3} \right) - \exp(-T) \right). \end{aligned} \quad (3.30)$$

In view of (3.29) and (3.30),

$$\begin{aligned} \mathbb{E} \left| Z^{1+\frac{2}{n}} - Z^{1+\frac{1}{n}} \right| &\geq \mathbb{E} \mathbf{1}_A \left| Z^{1+\frac{2}{n}} - Z^{1+\frac{1}{n}} \right| \geq \frac{Tn}{6} \mathbb{P}(A) \\ &= \frac{T}{6} \left( 1 - \exp \left( - \left( 1 + \frac{2}{n} \right) \frac{T}{3} \right) \right) \times \left( \exp \left( - \frac{2T}{3} \right) - \exp(-T) \right) \\ &\rightarrow \frac{T}{6} \left( 1 - \exp \left( - \frac{T}{3} \right) \right) \times \left( \exp \left( - \frac{2T}{3} \right) - \exp(-T) \right) \neq 0 \text{ as } n \rightarrow \infty. \end{aligned}$$

As a result,  $\{Z^{1+\frac{1}{n}} : n \in \mathbb{Z}_+\}$  is not a Cauchy sequence in  $L^1$ . Thus neither is it Cauchy in  $L^p$ ,  $p > 1$ . In the example, the continuous-state dependence makes the switching diffusions markedly different from that of the Markov modulated switching diffusions.

#### 4. Further properties

Continuing on our investigation, we derive further properties in this section. It is arranged in two subsections.

#### 4.1. Uniform continuity with respect to initial data

This section aims to obtain uniform estimates on a finite interval for the difference of two solutions  $X^{\tilde{x},\alpha}(t) - X^{x,\alpha}(t)$  with distinct initial data on the continuous component. Again, such a property is distinctly different from that of the Markov modulated switching diffusions; see [Remark 4.2](#) for details.

**Theorem 4.1.** Assume that  $b(x, i)$  and  $\sigma(x, i)$  are globally Lipschitz in  $x$  (with Lipschitz constant  $\kappa$ ) for each  $i \in \mathcal{M}$ , and that  $Q(x) = (q_{ij}(x))$  is bounded and for some  $\gamma_1 > 0$  we have

$$\sum_{k,j \in \mathcal{M}} |q_{kj}(x) - q_{kj}(y)| \leq K_1(1 + |x|^{\gamma_1} + |y|^{\gamma_1})|x - y|, \quad \forall x, y \in \mathbb{R}^r. \quad (4.1)$$

Then, there exists a constant  $C_T$  depending only on  $T$ ,  $K_1$ , and  $\kappa$  such that for any  $x, \tilde{x} \in \mathbb{R}^r$ , and  $\alpha \in \mathcal{M}$ , we have

$$\mathbb{E} \sup_{t \in [0, T]} |X^{\tilde{x},\alpha}(t) - X^{x,\alpha}(t)| \leq C_T(1 + |x|^{\gamma_1+2} + |\tilde{x}|^{\gamma_1+2})|\tilde{x} - x|. \quad (4.2)$$

**Proof.** By [\[22, Proposition 3.5\]](#), for any  $m \in \mathbb{Z}_+$ , there is a constant  $C_m$  depending only on the Lipschitz constant  $\kappa$  such that

$$\mathbb{E} \left( \sup_{t \in [0, T]} |X^{x_0, \alpha_0}(t)|^m \right) \leq C_m(1 + |x_0|^m) \exp(C_m T) \quad \forall (x_0, \alpha_0) \in \mathbb{R}^r \times \mathcal{M}, T > 0. \quad (4.3)$$

Let  $(X(t), \tilde{X}(t), \alpha(t), \tilde{\alpha}(t))$  be as in the proof of [Theorem 3.1](#). Again, denote  $\Delta = \tilde{x} - x$  and recall the definition of  $\tau^\Delta$  in [\(3.5\)](#). Then

$$\sup_{t \in [0, T]} |X(t) - \tilde{X}(t)| \leq \sup_{t \in [0, T \wedge \tau^\Delta]} |X(t) - \tilde{X}(t)| + \sup_{t \in (T \wedge \tau^\Delta, T]} |X(t) - \tilde{X}(t)|.$$

Hence

$$\begin{aligned} & \mathbb{E} \left( \sup_{t \in [0, T]} |X(t) - \tilde{X}(t)| \right) \\ & \leq \mathbb{E} \left( \sup_{t \in [0, T \wedge \tau^\Delta]} |X(t) - \tilde{X}(t)| \right) + \mathbb{E} \left( \mathbf{1}_{\{\tau^\Delta \leq T\}} \sup_{t \in (T \wedge \tau^\Delta, T]} |X(t) - \tilde{X}(t)| \right). \end{aligned} \quad (4.4)$$

Let  $\mathcal{F}_{T \wedge \tau^\Delta}$  be the  $\sigma$ -algebra generated by the processes  $(X(t), \tilde{X}(t), \alpha(t), \tilde{\alpha}(t))$  up to the time  $T \wedge \tau^\Delta$ . We have that

$$\begin{aligned} & \mathbb{E} \left[ \mathbf{1}_{\{\tau^\Delta \leq T\}} \sup_{t \in (T \wedge \tau^\Delta, T]} |X(t) - \tilde{X}(t)| \right] \\ & = \mathbb{E} \left[ \mathbb{E} \left( \mathbf{1}_{\{\tau^\Delta \leq T\}} \sup_{t \in (T \wedge \tau^\Delta, T]} |X(t) - \tilde{X}(t)| \middle| \mathcal{F}_{T \wedge \tau^\Delta} \right) \right] \\ & = \mathbb{E} \left[ \mathbf{1}_{\{\tau^\Delta \leq T\}} \mathbb{E} \left( \sup_{t \in (T \wedge \tau^\Delta, T]} |X(t) - \tilde{X}(t)| \middle| \mathcal{F}_{T \wedge \tau^\Delta} \right) \right] \\ & \leq \mathbb{E} \left[ \mathbf{1}_{\{\tau^\Delta \leq T\}} \mathbb{E} \left( \sup_{t \in (T \wedge \tau^\Delta, T]} \{|X(t)| + |\tilde{X}(t)|\} \middle| \mathcal{F}_{T \wedge \tau^\Delta} \right) \right] \end{aligned}$$

$$\begin{aligned} &\leq \mathbb{E} \left[ \mathbf{1}_{\{\tau^\Delta \leq T\}} \mathbb{E} \left( \sup_{t \in (T \wedge \tau^\Delta, T \wedge \tau^\Delta + T]} \{|X(t)| + \tilde{X}(t)|\} \middle| \mathcal{F}_{T \wedge \tau^\Delta} \right) \right] \\ &\leq 2C_1 e^{C_1 T} \mathbb{E} [\mathbf{1}_{\{\tau^\Delta \leq T\}} (1 + |X(T \wedge \tau^\Delta)| + |\tilde{X}(T \wedge \tau^\Delta)|)], \end{aligned} \quad (4.5)$$

where the last inequality follows from the strong Markov property of  $(X(t), \tilde{X}(t), \alpha(t), \tilde{\alpha}(t))$  and (4.3). Note that the coupled process  $(X(t), \tilde{X}(t), \alpha(t), \tilde{\alpha}(t))$  is the solution to (3.3). For  $k, l \in \mathcal{M}$ , let  $\tilde{A}_{kl}$  be the generator of the diffusion

$$\begin{cases} dY(t) = b(Y(t), k)dt + \sigma(Y(t), k)dw(t) \\ d\tilde{Y}(t) = b(\tilde{Y}(t), l)dt + \sigma(\tilde{Y}(t), l)d\tilde{w}(t). \end{cases} \quad (4.6)$$

Then the generator  $\tilde{\mathcal{L}}$  of  $(X(t), \tilde{X}(t), \alpha(t), \tilde{\alpha}(t))$  is given by

$$\tilde{\mathcal{L}}f(x, \tilde{x}, k, l) = \tilde{A}_{kl}f(x, \tilde{x}, k, l) + \tilde{Q}(x, \tilde{x})f(x, \tilde{x}, k, l). \quad (4.7)$$

Define  $U(x, \tilde{x}, k, l) = V(k, l)(1 + |x|^2 + |\tilde{x}|^2) = \mathbf{1}_{\{k \neq l\}}(1 + |x|^2 + |\tilde{x}|^2)$ . Clearly, when  $k = l$ ,  $U(x, \tilde{x}, k, k) = 0$ , for all  $x$  and  $\tilde{x}$ . Thus,  $\tilde{A}_{kk}U(x, \tilde{x}, k, k) = 0$ , which combined with (3.8) implies that

$$\begin{aligned} \tilde{\mathcal{L}}U(x, \tilde{x}, k, k) &= \tilde{Q}(x, \tilde{x})U(x, \tilde{x}, k, k) \\ &= (1 + |x|^2 + |\tilde{x}|^2)\tilde{Q}(x, \tilde{x})V(k, k) = (1 + |x|^2 + |\tilde{x}|^2)\rho(x, \tilde{x}, k), \end{aligned} \quad (4.8)$$

where  $\rho(x, \tilde{x}, k)$  is defined in (3.8). Since all moments of  $X(t)$  and  $\tilde{X}(t)$  are bounded, applying Dynkin's formula, we have

$$\begin{aligned} &\mathbb{E}U(X(T \wedge \tau^\Delta), \tilde{X}(T \wedge \tau^\Delta), \alpha(T \wedge \tau^\Delta), \tilde{\alpha}(T \wedge \tau^\Delta)) \\ &= U(x, \tilde{x}, \alpha, \tilde{\alpha}) + \mathbb{E} \int_0^{T \wedge \tau^\Delta} \tilde{\mathcal{L}}U(X(s), \tilde{X}(s), \alpha(s), \tilde{\alpha}(s))ds \\ &= \mathbb{E} \int_0^{T \wedge \tau^\Delta} (1 + |X(s)|^2 + |\tilde{X}(s)|^2)\rho(X(s), \tilde{X}(s), \alpha(s))ds, \end{aligned} \quad (4.9)$$

where the last equality above follows from (4.8) and the fact  $\alpha(t) = \tilde{\alpha}(t)$  if  $t < \tau^\Delta$ . Then by Hölder's inequality, estimates (4.1), and (4.3), we have

$$\begin{aligned} &\mathbb{E} \int_0^{T \wedge \tau^\Delta} (1 + |X(t)|^2 + |\tilde{X}(t)|^2)\rho(X(t), \tilde{X}(t), \alpha(t))dt \\ &\leq \mathbb{E} \int_0^{T \wedge \tau^\Delta} \tilde{K}(1 + |X(t)|^{\gamma_1+2} + |\tilde{X}(t)|^{\gamma_1+2})|X(t) - \tilde{X}(t)|dt \quad (\text{for some } \tilde{K} > 0) \\ &\leq \tilde{K}T \mathbb{E} \sup_{t \leq T \wedge \tau^\Delta} \{(1 + |X(t)|^{\gamma_1+2} + |\tilde{X}(t)|^{\gamma_1+2})|X(t) - \tilde{X}(t)|\} \\ &\leq \tilde{K}T \left( \mathbb{E} \sup_{t \leq T \wedge \tau^\Delta} \{(1 + |X(t)|^{\gamma_1+2} + |\tilde{X}(t)|^{\gamma_1+2})^2\} \right)^{\frac{1}{2}} \left( \mathbb{E} \sup_{t \leq T \wedge \tau^\Delta} \{|X(t) - \tilde{X}(t)|^2\} \right)^{\frac{1}{2}} \\ &\leq \tilde{K}_{1,T}(1 + |x|^{\gamma_1+2} + |\tilde{x}|^{\gamma_1+2})|\Delta| \quad (\text{for some } \tilde{K}_{1,T} > 0), \end{aligned} \quad (4.10)$$

where the last inequality follows from (3.6) and (4.3).

In view of (4.5), (4.9), and (4.10),

$$\begin{aligned}
 & \mathbb{E} \left( \mathbf{1}_{\{\tau^\Delta \leq T\}} \sup_{t \in [T \wedge \tau^\Delta, T]} |X(t) - \tilde{X}(t)| \right) \\
 & \leq 2C_1 e^{C_1 T} \mathbb{E} \left( \mathbf{1}_{\{\tau^\Delta \leq T\}} (1 + |X(T \wedge \tau^\Delta)| + |\tilde{X}(T \wedge \tau^\Delta)|) \right) \\
 & \leq 6C_1 e^{C_1 T} \mathbb{E} \left( \mathbf{1}_{\{\tau^\Delta \leq T\}} (1 + |X(T \wedge \tau^\Delta)|^2 + |\tilde{X}(T \wedge \tau^\Delta)|^2) \right) \\
 & \leq 6C_1 e^{C_1 T} \mathbb{E} U(X(T \wedge \tau^\Delta), \tilde{X}(T \wedge \tau^\Delta), \alpha(T \wedge \tau^\Delta), \tilde{\alpha}(T \wedge \tau^\Delta)) \\
 & \leq 6C_1 e^{C_1 T} \tilde{K}_{1,T} (1 + |x|^{\gamma_1+2} + |\tilde{x}|^{\gamma_1+2}) |\Delta|. \tag{4.11}
 \end{aligned}$$

The proof is complete by applying (4.11) and (3.6) to (4.3).  $\square$

**Remark 4.2.** In contrast to switching diffusions with Markovian switching, in which the switching is independent of the Brownian motion, in our case, we have to estimate  $X(t) - \tilde{X}(t)$  after  $\tau^\Delta$  if  $\tau^\Delta \leq T$ . Note that the difference  $|X(t) - \tilde{X}(t)|$  after  $\tau^\Delta$  cannot be estimated by something related to  $|x - \tilde{x}|$  since after  $\tau^\Delta$ , the evolutions of  $X(t)$  and  $\tilde{X}(t)$  are quite different. We can only estimate  $|X(t) - \tilde{X}(t)|$  by  $|X(t)| + |\tilde{X}(t)|$  for  $t > \tau^\Delta$ . For this reason, in general, it does not seem that  $C_T(1 + |x|^{\gamma_1+2} + |\tilde{x}|^{\gamma_1+2})$  in (4.2) can be replaced with a constant  $K$  independent of  $x$  and  $\tilde{x}$ , even under the condition that  $Q(x)$  is globally Lipschitz. However, if we assume  $b$  and  $\sigma$  are bounded,  $C_T(1 + |x|^{\gamma_1+2} + |\tilde{x}|^{\gamma_1+2})$  in (4.2) can be replaced by a constant  $K$ . Finally, we remark that if the condition for  $Q(x)$  is reduced to

$$\begin{aligned}
 & \sum_{k,j \in \mathcal{M}} |q_{kj}(x) - q_{kj}(y)| \leq K_1(1 + |x|^{\gamma_1} + |y|^{\gamma_1})|x - y|^\lambda, \\
 & \text{for all } x, y \in \mathbb{R}^r \text{ and some } 0 < \lambda \leq 1
 \end{aligned}$$

then we can obtain

$$\mathbb{E} \sup_{t \in [0, T]} |X^{\tilde{x}, \alpha}(t) - X^{x, \alpha}(t)| \leq C_T(1 + |x|^{\gamma_1+2} + |\tilde{x}|^{\gamma_1+2})|\tilde{x} - x|^\lambda.$$

#### 4.2. Smoothness of a functional of the switching diffusion

Continuing on our investigations, this section obtains smoothness of nonlinear functional of the switching diffusions. Once again, the continuous-state dependent switching presents much difficulty. Our task is to untangle the process  $\alpha^{x, \alpha}(t)$  and  $\alpha^{\tilde{x}, \alpha}(t)$  as in the previous sections. As alluded to in Remark 4.2, the difficulty due to the continuous state dependent switching is particularly pronounced.

**Theorem 4.3.** For each  $i \in \mathcal{M}$ , assume that  $b(x, i)$  and  $\sigma(x, i)$  are globally Lipschitz in  $x$  (with Lipschitz constant  $\kappa$ ), that  $b(\cdot, i)$ ,  $\sigma(\cdot, i)$ ,  $q_{ij}(\cdot) \in C^2$ , that there is a  $\phi(\cdot, i) \in C^2$ , and that

$$|D_x^\beta b(x, i)| + |D_x^\beta \sigma(x, i)| + |D_x^\beta \phi(x, i)| \leq K(1 + |x|^\gamma), i \in \mathcal{M}, |\beta| \leq 2. \tag{4.12}$$

for some positive constant  $\gamma$ . Assume further that  $|D^\beta q_{ij}(\cdot)|$  are Lipschitz and bounded uniformly by some constant  $M$  for  $|\beta| \leq 2$ . Then,  $u(t, x, i) = \mathbb{E}[\phi(X^{x, i}(t), \alpha^{x, i}(t))]$  is twice continuously differentiable with respect to the variable  $x$ .

**Proof.** Again, for simplicity, we work out the details for the one-dimensional case. Denote by  $(\partial/\partial x)\phi(\cdot, i)$  for each  $i \in \mathcal{M}$ , and  $(d/dx)q_{ij}(\cdot)$  the derivatives of  $\phi(\cdot, i)$  and  $q_{ij}(\cdot)$  with respect

to  $x$ , respectively. Let  $\chi(t)$  be the Markov chain in  $\mathcal{M}$  with generators  $\hat{q}_{ij} = 1$  if  $i \neq j$ ,  $\hat{q}_{ii} = -m_0 + 1$ . Let  $\theta_n^{i_0}$  be the  $n$ th jump time of  $\chi(t)$  given that  $\chi(0) = i_0$ . Let  $Z^{x_0, i_0}(t)$  be the solution with initial value  $Z^{x_0, i_0}(0) = x_0$  to

$$dZ(t) = b(Z(t), \chi(t))dt + \sigma(Z(t), \chi(t))dw(t). \quad (4.13)$$

Let  $\tau_n^{x_0, i_0}$  be the  $n$ th jump time of  $\alpha^{x_0, i_0}(t)$ . By a change of measure (see [4]), we have

$$\begin{aligned} & \mathbb{E} \left[ \phi(X^{x_0, i_0}(T), \alpha^{x_0, i_0}(T)) \mathbf{1}_{\{\tau_n^{x_0, i_0} \leq T < \tau_{n+1}^{x_0, i_0}\}} \prod_{k=1}^n \mathbf{1}_{\{\alpha^{x_0, i_0}(\tau_k^{x_0, i_0}) = i_k\}} \right] \\ &= \exp((m_0 - 1)T) \mathbb{E} \left[ \phi(Z^{x_0, i_0}(T), i_n) \mathbf{1}_{\{\theta_n^{i_0} \leq T < \theta_{n+1}^{i_0}\}} \left( \prod_{k=1}^n \mathbf{1}_{\{\chi(\theta_k^{i_0}) = i_k\}} \right) \right. \\ & \quad \times \exp \left\{ - \int_0^T q_{\chi(s)}(Z^{x_0, i_0}(s)) ds \right\} \prod_{k=0}^{n-1} \left( q_{i_k i_{k+1}}(Z^{x_0, i_0}(\theta_{k+1}^{i_0})) \right) \Big], \end{aligned} \quad (4.14)$$

where  $q_i(x) = \sum_{j \neq i} q_{ij}(x)$ . Now, fix  $x$  and  $\alpha = i_0$  and let  $(X(t), \alpha(t))$ ,  $(\tilde{X}(t), \tilde{\alpha}(t))$  be the switching-diffusion processes satisfying (2.2) and (2.3) with initial conditions  $(x, \alpha)$  and  $(\tilde{x}, \tilde{\alpha})$ , respectively, where  $\tilde{x} = x + \Delta$ ,  $\tilde{\alpha} = i_0$ . Likewise, let  $Z(t)$ ,  $\tilde{Z}(t)$  be the solutions to (4.13) with initial conditions  $x, \tilde{x}$  respectively. By standard arguments (e.g. [10, Lemma 3.3] and [17, Lemma 3.2]), we have

$$\begin{cases} \mathbb{E} \sup_{t \in [0, T]} |\tilde{Z}(t) - Z(t)|^4 \leq K \Delta^4 \text{ for some } K > 0, \\ \mathbb{E} \sup_{t \in [0, T]} |Z(t)|^m \leq K_m (1 + |x|^m) \text{ for any } m > 0. \end{cases} \quad (4.15)$$

Adapting the proofs of [5, Theorems 5.5.2, 5.5.3] with a slight modification in replacing fixed times by stopping times, we can show that

$$\mu_\Delta := \sup_{\{\vartheta \in T\}} \frac{\mathbb{E} |\tilde{Z}(\vartheta) - Z(\vartheta) - \eta(\vartheta)\Delta|^2}{\Delta^2} \rightarrow 0 \text{ as } \Delta \rightarrow 0, \quad (4.16)$$

where  $\eta(t)$  is the derivative of  $Z(t)$  with respect to the initial condition  $x$  which exists in  $L^2$ -sense and  $T$  is the set of all stopping times that are bounded above by  $T$  almost surely. Note that  $\eta(t)$  satisfies

$$\eta(t) = 1 + \int_0^t \eta(s) b_x(Z(s), \chi(s)) ds + \int_0^t \eta(s) \sigma_x(Z(s), \chi(s)) dw(s).$$

Using the above equation, (4.15), the Burkholder–Davis–Gundy inequality, Gronwall's inequality and standard arguments in [17, Lemma 3.2] or [5, Theorem 5.2.2], we can show that

$$\mathbb{E} \sup_{t \in [0, T]} |\eta(t)|^2 \leq C(T, x), \quad (4.17)$$

where  $C(T, x)$  is a constant depending on  $T$  and  $x$ . By Taylor's expansion, it can be easily shown that

$$\begin{aligned} & \left| \exp \left\{ - \int_0^T q_{\chi(s)}(\tilde{Z}_s) ds \right\} - \exp \left\{ - \int_0^T q_{\chi(s)}(Z_s) ds \right\} \right. \\ & \quad \left. + \left( \int_0^T (\tilde{Z}(s) - Z(s)) \frac{d}{dz} q_{\chi(s)}(Z(s)) ds \right) \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \right| \end{aligned}$$



$$\leq M \int_0^T |\tilde{Z}(s) - Z(s)|^2 ds. \quad (4.18)$$

In view of (4.18)

$$\begin{aligned} & \left| \frac{1}{\Delta} \left( \exp \left\{ - \int_0^T q_{\chi(s)}(\tilde{Z}_s) ds \right\} - \exp \left\{ - \int_0^T q_{\chi(s)}(Z_s) ds \right\} \right) \right. \\ & \quad \left. + \left( \int_0^T \eta(s) \frac{d}{dz} q_{\chi(s)}(Z(s)) ds \right) \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \right| \\ & \leq \frac{1}{\Delta} \left| \exp \left\{ - \int_0^T q_{\chi(s)}(\tilde{Z}_s) ds \right\} - \exp \left\{ - \int_0^T q_{\chi(s)}(Z_s) ds \right\} \right| \\ & \quad + \left| \left( \int_0^T (\tilde{Z}(s) - Z(s)) \frac{d}{dz} q_{\chi(s)}(Z(s)) ds \right) \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \right| \\ & \quad + \left| \left( \int_0^T \left[ \eta(s) - \frac{\tilde{Z}(s) - Z(s)}{\Delta} \right] \frac{d}{dz} q_{\chi(s)}(Z(s)) ds \right) \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \right| \\ & \leq \frac{M}{\Delta} \int_0^T (|\tilde{Z}(s) - Z(s)|^2 + |\tilde{Z}(s) - Z(s) - \Delta \eta(s)|) ds. \end{aligned} \quad (4.19)$$

We have from Taylor's expansion

$$\begin{aligned} & \left| \frac{q_{i_k i_{k+1}}(\tilde{Z}(\theta_{k+1})) - q_{i_k i_{k+1}}(Z(\theta_{k+1}))}{\Delta} - \eta(\theta_{k+1}) \frac{d}{dz} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \right| \\ & \leq \frac{1}{\Delta} \left| q_{i_k i_{k+1}}(\tilde{Z}(\theta_{k+1})) - q_{i_k i_{k+1}}(Z(\theta_{k+1})) - (\tilde{Z}(\theta_{k+1}) - Z(\theta_{k+1})) \frac{d}{dz} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \right| \\ & \quad + \frac{1}{\Delta} |\tilde{Z}(\theta_{k+1}) - Z(\theta_{k+1}) - \Delta \eta(\theta_{k+1})| \left| \frac{d}{dz} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \right| \\ & \leq \frac{M}{\Delta} (|\tilde{Z}(\theta_{k+1}) - Z(\theta_{k+1})|^2 + |\tilde{Z}(\theta_{k+1}) - Z(\theta_{k+1}) - \Delta \eta(\theta_{k+1})|). \end{aligned} \quad (4.20)$$

Similar to [5, (5.22) on p. 123], we have

$$\begin{aligned} \psi_\Delta &:= \left| \frac{\phi(\tilde{Z}(T), \chi(T)) - \phi(Z(T), \chi(T))}{\Delta} - \eta(T) \frac{\partial}{\partial z} \phi(Z(T), \chi(T)) \right| \\ &\rightarrow 0 \text{ in } L^2(\Omega) \text{ as } \Delta \rightarrow 0. \end{aligned} \quad (4.21)$$

To simplify notation, we drop the superscript of  $\theta_n^{i_0}$  and for each  $n \in \mathbb{Z}_+$ , let

$$A_n := \{(i_1, \dots, i_{n+1}) : i_k \in \mathcal{M}, i_k \neq i_{k-1}, k = 1, \dots, n+1\}.$$

For each  $\mathbf{I} = (i_1, \dots, i_{n+1}) \in A_n$ , let

$$f_{\mathbf{I}} := \mathbf{1}_{\{\theta_n \leq T < \theta_{n+1}\}} \left( \prod_{k=1}^n \mathbf{1}_{\{\chi(\theta_k) = i_k\}} \right).$$

We compute

$$\begin{aligned} & \left| \frac{1}{\Delta} \left[ \phi(\tilde{Z}(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(\tilde{Z}(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(\tilde{Z}(\theta_{k+1})) \right. \right. \\ & \quad \left. \left. - \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \right] \right| \end{aligned}$$

$$\begin{aligned}
& -\eta(T) \frac{\partial}{\partial z} \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \\
& - \sum_{j=0}^{n-1} \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \eta(\theta_{j+1}) \frac{d}{dz} q_{i_j i_{j+1}}(Z(\theta_{j+1})) \\
& \times \prod_{k=0, k \neq j}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \\
& + \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \left( \int_0^T \eta(s) \frac{d}{dz} q_{\chi(s)}(Z(s)) ds \right) \\
& \times \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \Big| \\
& \leq C f_{\mathbf{I}} M^n \left[ \psi_{\Delta} + (1 + |\tilde{Z}(T)|^{\gamma} + |Z(T)|^{\gamma}) \right. \\
& \times \int_0^T \frac{|\tilde{Z}(s) - Z(s)|^2 + |\tilde{Z}(s) - Z(s) - \Delta \eta(s)|}{\Delta} ds \\
& + (1 + |\tilde{Z}(T)|^{\gamma} + |Z(T)|^{\gamma}) \\
& \times \sum_{j=0}^n \frac{|\tilde{Z}(\theta_{j+1}) - Z(\theta_{j+1})|^2 + |\tilde{Z}(\theta_{j+1}) - Z(\theta_{j+1}) - \Delta \eta(\theta_{j+1})|}{\Delta} \Big]. \tag{4.22}
\end{aligned}$$

Let  $\hat{\zeta}_n(T)$  be defined by

$$\begin{aligned}
\hat{\zeta}_n(T) &= \sum_{\mathbf{I} \in A_n} \left[ \eta(T) \frac{\partial}{\partial z} \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \right. \\
& + \sum_{j=0}^{n-1} \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \eta(\theta_{j+1}) \frac{d}{dz} q_{i_j i_{j+1}}(Z(\theta_{j+1})) \\
& \times \prod_{k=0, k \neq j}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \\
& - \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \left( \int_0^T \eta(s) \frac{d}{dz} q_{\chi(s)}(Z(s)) ds \right) \\
& \times \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \Big].
\end{aligned}$$

Note that  $\sum_{\mathbf{I} \in A_n} f_{\mathbf{I}} = 1$ . It then follows from (4.22) that for each  $n \in \mathbb{Z}_+$ ,  $\hat{\zeta}_n(t)$  such that

$$\begin{aligned}
& \mathbb{E} \sum_{\mathbf{I} \in A_n} \left| \frac{1}{\Delta} \left[ \phi(\tilde{Z}(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(\tilde{Z}(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(\tilde{Z}(\theta_{k+1})) \right. \right. \\
& \quad \left. \left. - \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \right] - \hat{\zeta}_n(T) \right| \\
& \leq M^n \mathbb{E} \mathbf{1}_{\{\theta_n \leq T < \theta_{n+1}\}} \psi_{\Delta}
\end{aligned}$$

$$\begin{aligned}
& + \frac{C}{\Delta} M^n \mathbb{E} \mathbf{1}_{\{\theta_n \leq T < \theta_{n+1}\}} (1 + |\tilde{Z}(T)|^\gamma + |Z(T)|^\gamma) \\
& \times \int_0^T \left( |\tilde{Z}(s) - Z(s)|^2 + |\tilde{Z}(s) - Z(s) - \Delta\eta(s)| \right) ds \\
& + \frac{C}{\Delta} M^n \mathbb{E} \mathbf{1}_{\{\theta_n \leq T < \theta_{n+1}\}} (1 + |\tilde{Z}(T)|^\gamma + |Z(T)|^\gamma) \\
& \times \sum_{j=0}^n \left( |\tilde{Z}(\theta_{j+1}) - Z(\theta_{j+1})|^2 + |\tilde{Z}(\theta_{j+1}) - Z(\theta_{j+1}) - \Delta\eta(\theta_{j+1})| \right). \tag{4.23}
\end{aligned}$$

By Hölder's inequality and the fact that

$$\mathbb{P}\{\theta_n \leq T < \theta_{n+1}\} = \frac{\exp(-m_0 T)(m_0 T)^n}{n!}, \tag{4.24}$$

we can obtain the following estimate:

$$\begin{aligned}
\mathbb{E} \mathbf{1}_{\{\theta_n \leq T < \theta_{n+1}\}} \psi_\Delta & \leq \left( \mathbb{P}\{\theta_n \leq T < \theta_{n+1}\} \mathbb{E} \psi_\Delta^2 \right)^{\frac{1}{2}} \\
& \leq \left( \frac{\exp(-m_0 T)(m_0 T)^n}{n!} \right)^{\frac{1}{2}} (\mathbb{E} \psi_\Delta^2)^{\frac{1}{2}}. \tag{4.25}
\end{aligned}$$

Similarly, with the aid of Hölder's and Minkowski's inequalities, we have

$$\begin{aligned}
& \frac{1}{\Delta} \mathbb{E} \mathbf{1}_{\{\theta_n \leq T < \theta_{n+1}\}} (1 + |\tilde{Z}(T)|^\gamma + |Z(T)|^\gamma) \\
& \times \int_0^T \left( |\tilde{Z}(s) - Z(s)|^2 + |\tilde{Z}(s) - Z(s) - \Delta\eta(s)| \right) ds \\
& \leq \left( \mathbb{P}\{\theta_n \leq T < \theta_{n+1}\} \right)^{\frac{1}{4}} \left( 1 + (\mathbb{E} |\tilde{Z}(T)|^{4\gamma})^{\frac{1}{4}} + (\mathbb{E} |Z(T)|^{4\gamma})^{\frac{1}{4}} \right) \\
& \times \frac{1}{\Delta} \left[ \int_0^T \left( \mathbb{E} |\tilde{Z}(s) - Z(s)|^4 \right)^{\frac{1}{2}} ds + \int_0^T \left( \mathbb{E} |\tilde{Z}(s) - Z(s) - \Delta\eta(s)|^2 \right)^{\frac{1}{2}} ds \right] \\
& \leq K \left( \frac{\exp(-m_0 T)(m_0 T)^n}{n!} \right)^{\frac{1}{4}} (\Delta + \mu^{\frac{1}{2}}_\Delta) \text{ (by (4.15) and (4.16))}. \tag{4.26}
\end{aligned}$$

Likewise,

$$\begin{aligned}
& \frac{1}{\Delta} \mathbb{E} \left[ \mathbf{1}_{\{\theta_n \leq T < \theta_{n+1}\}} (1 + |\tilde{Z}(T)|^\gamma + |Z(T)|^\gamma) \right. \\
& \times \left. \sum_{j=0}^n \left( |\tilde{Z}(\theta_{j+1}) - Z(\theta_{j+1})|^2 + |\tilde{Z}(\theta_{j+1}) - Z(\theta_{j+1}) - \Delta\eta(\theta_{j+1})| \right) \right] \\
& \leq n K \left( \frac{\exp(-m_0 T)(m_0 T)^n}{n!} \right)^{\frac{1}{4}} (\Delta + \mu^{\frac{1}{2}}_\Delta). \tag{4.27}
\end{aligned}$$

Applying (4.25), (4.26) and (4.27) to (4.23) to obtain

$$\begin{aligned}
& \mathbb{E} \sum_{\mathbf{I} \in A_n} \left| \frac{1}{\Delta} \left[ \phi(\tilde{Z}(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(\tilde{Z}(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(\tilde{Z}(\theta_{k+1})) \right. \right. \\
& \left. \left. - \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \right] - \hat{\zeta}_n(T) \right|
\end{aligned}$$

$$\leq (n+1)\tilde{K}M^n \left( \frac{\exp(-m_0T)(m_0T)^n}{n!} \right)^{\frac{1}{4}} (\mu_{\Delta}^{\frac{1}{2}} + (\mathbb{E}\psi_{\Delta}^2)^{\frac{1}{2}} + \Delta) \quad (4.28)$$

for some  $\tilde{K}$  independent of  $n$ . It is not difficult to show that

$$\sum_{n=0}^{\infty} (n+1)\tilde{K}M^n \left( \frac{\exp(-m_0T)(m_0T)^n}{n!} \right)^{\frac{1}{4}} < \infty. \quad (4.29)$$

Moreover, by virtue of (4.16) and (4.21), we have

$$\lim_{\Delta \rightarrow 0} (\mu_{\Delta}^{\frac{1}{2}} + (\mathbb{E}\psi_{\Delta}^2)^{\frac{1}{2}} + \Delta) = 0.$$

Thus,

$$\begin{aligned} \mathbb{E} \sum_{n=0}^{\infty} \left| \sum_{\mathbf{I} \in A_n} \frac{1}{\Delta} \left[ \phi(\tilde{Z}(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(\tilde{Z}(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(\tilde{Z}(\theta_{k+1})) \right. \right. \\ \left. \left. - \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \right] - \hat{\zeta}_n(T) \right| \rightarrow 0 \end{aligned} \quad (4.30)$$

as  $\Delta \rightarrow 0$ .

Similar to (4.28) and (4.29), by using Holder's inequality, (4.17), and (4.24), we can obtain that

$$\mathbb{E} \left| \sum_{n=0}^{\infty} \hat{\zeta}_n(T) \right| \leq \hat{C}(T, x) < \infty. \quad (4.31)$$

As a result of (4.30) and (4.31), we deduce

$$\begin{aligned} & \frac{1}{\Delta} [u(T, x + \Delta, i) - u(T, x, i)] \\ &= \frac{1}{\Delta} \mathbb{E} (\phi(\tilde{X}(T), \tilde{\alpha}(T)) - \phi(X(T), \alpha(T))) \\ &= \frac{1}{\Delta} \mathbb{E} \sum_{n=0}^{\infty} \sum_{\mathbf{I} \in A_n} \left[ \phi(\tilde{Z}(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(\tilde{Z}(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(\tilde{Z}(\theta_{k+1})) \right. \\ & \quad \left. - \phi(Z(T), i_n) f_{\mathbf{I}} \exp \left\{ - \int_0^T q_{\chi(s)}(Z(s)) ds \right\} \prod_{k=0}^{n-1} q_{i_k i_{k+1}}(Z(\theta_{k+1})) \right] \\ & \rightarrow \mathbb{E} \sum_{n=0}^{\infty} \hat{\zeta}_n(T) \text{ as } \Delta \rightarrow 0. \end{aligned} \quad (4.32)$$

We have therefore proved that  $u(t, x, i)$  is differentiable with respect to  $x$  with

$$\frac{\partial}{\partial x} u(t, x, i) = \mathbb{E}_{x,i} \sum_{n=0}^{\infty} \hat{\zeta}_n(T).$$

With Taylor's expansion up to the second order and using the same method, we can show that  $u(t, x, i)$  is twice differentiable with respect to  $x$ .  $\square$

**Remark 4.4.** If (4.12) holds for any  $\beta \leq n$ , then we can use a change of measure argument, Taylor's expansion, and arguments in the proof of Theorem 4.3 to obtain differentiability up to order  $n$  of  $u(t, x, i)$ . However, the estimates would be more complicated.

## 5. Feller property under non-global Lipschitz condition

This section is devoted to obtaining Feller properties of switching diffusions under non-Lipschitz condition. There has been much work on Feller properties of switching diffusions in the literature. However, to the best of our knowledge, all of the work up to date has been concentrated on the case under a global Lipschitz condition; see for example, [13,14,16]. When the global Lipschitz condition is violated, will the processes still possess Feller property? We address this issue in what follows.

**Theorem 5.1.** Assume that the hypothesis of Theorem 2.1 is satisfied. Then the solution process  $(X(t), \alpha(t))$  for the system given by (2.2) and (2.3) is a Markov–Feller process.

**Proof.** The Markov property follows from standard arguments. In what follows, we focus on the proof of the Feller property. First, we suppose that  $b(x, i)$  and  $\sigma(x, i)$  are Lipschitz in  $x$  for each  $i \in \mathcal{M}$ . In fact, the Feller property was obtained in [16, Section 2.5] under global Lipschitz condition. The proof was rather long. Here, using our results of the current paper, we provide an alternative proof. Let  $f(\cdot, \cdot) : \mathbb{R}^r \times \mathcal{M} \mapsto \mathbb{R}$  be a bounded and continuous function. Fix  $(x, \alpha) \in \mathbb{R}^r \times \mathcal{M}$  and  $t > 0$ . Let  $\{x_n\}$  be any sequence converging to  $x$  as  $n \rightarrow \infty$ . By the definition of the limit superior of a sequence, since  $f(\cdot, \cdot)$  is bounded, we can always extract a subsequence  $\{\mathbb{E}f(X^{x_{n_k}, \alpha}(t), \alpha^{x_{n_k}, \alpha}(t))\}$  from  $\{\mathbb{E}f(X^{x_n, \alpha}(t), \alpha^{x_n, \alpha}(t))\}$  such that

$$\lim_{k \rightarrow \infty} \mathbb{E}f(X^{x_{n_k}, \alpha}(t), \alpha^{x_{n_k}, \alpha}(t)) = \limsup_{n \rightarrow \infty} \mathbb{E}f(X^{x_n, \alpha}(t), \alpha^{x_n, \alpha}(t)). \quad (5.1)$$

In view of (3.10) and (3.11),  $(X^{x_{n_k}, \alpha}(t), \alpha^{x_{n_k}, \alpha}(t))$  converges to  $(X^{x, \alpha}(t), \alpha^{x, \alpha}(t))$  in probability as  $k \rightarrow \infty$ . Then it has a subsequence converging almost surely to  $(X^{x, \alpha}(t), \alpha^{x, \alpha}(t))$ . Without loss of generality, we may assume that  $(X^{x_{n_k}, \alpha}(t), \alpha^{x_{n_k}, \alpha}(t))$  converges almost surely to  $(X^{x, \alpha}(t), \alpha^{x, \alpha}(t))$  as  $k \rightarrow \infty$ . Then the dominated convergence theorem and (5.1) imply that

$$\limsup_{n \rightarrow \infty} \mathbb{E}f(X^{x_n, \alpha}(t), \alpha^{x_n, \alpha}(t)) = \lim_{k \rightarrow \infty} \mathbb{E}f(X^{x_{n_k}, \alpha}(t), \alpha^{x_{n_k}, \alpha}(t)) = \mathbb{E}f(X^{x, \alpha}(t), \alpha^{x, \alpha}(t)).$$

Likewise,

$$\liminf_{n \rightarrow \infty} \mathbb{E}f(X^{x_n, \alpha}(t), \alpha^{x_n, \alpha}(t)) = \mathbb{E}f(X^{x, \alpha}(t), \alpha^{x, \alpha}(t)).$$

Thus, we obtain the Feller property under the condition that  $b(x, i)$  and  $\sigma(x, i)$  are Lipschitz in  $x$  for each  $i \in \mathcal{M}$ .

Next we relax the condition and assume only the local Lipschitz continuity as in the statement of the theorem. It is proved in [16, Theorem 2.7] that for any  $R > 0$ ,  $\varepsilon > 0$ , and  $t > 0$ , there is an  $H_R > 0$  such that

$$\mathbb{P}\{|X^{x, \alpha}(s)| < H_R \text{ for all } s \in [0, t]\} > 1 - \varepsilon \text{ if } |x| \leq R. \quad (5.2)$$

Now fix  $(x, \alpha) \in \mathbb{R}^r \times \mathcal{M}$  and  $R > |x| + 1$ . Let  $f(\cdot, \cdot) : \mathbb{R}^r \times \mathcal{M} \mapsto \mathbb{R}$  be a continuous function satisfying  $|f(x, \alpha)| \leq 1$  for all  $(x, \alpha) \in \mathbb{R}^r \times \mathcal{M}$ . Let  $\psi(\cdot) : \mathbb{R}^r \mapsto [0, 1]$  be a smooth function

with compact support satisfying  $\psi(x) = 1$  if  $|x| \leq H_R$ . By the first part of this proof, the process  $(\hat{X}(t), \hat{\alpha}(t))$  satisfying

$$\begin{cases} d\hat{X}(t) = \psi(\hat{X}(t))b(\hat{X}(t), \hat{\alpha}(t))dt + \psi(\hat{X}(t))\sigma(\hat{X}(t), \hat{\alpha}(t))dw(t), \\ \mathbb{P}\{\hat{\alpha}(t + \Delta) = j | \hat{\alpha}(t) = i, \hat{X}(s), \hat{\alpha}(s), s \leq t\} = q_{ij}(\hat{X}(t))\Delta + o(\Delta) \text{ if } i \neq j \end{cases} \quad (5.3)$$

has the Feller property. Thus, there exists some  $\delta \in (0, 1)$  such that

$$\left| \mathbb{E}f(\hat{X}^{x+h, \alpha}(t), \hat{\alpha}^{x+h, \alpha}(t)) - \mathbb{E}f(\hat{X}^{x, \alpha}(t), \hat{\alpha}^{x, \alpha}(t)) \right| < \varepsilon \text{ for any } h \in \mathbb{R}^r, |h| \leq \delta. \quad (5.4)$$

where  $(\hat{X}^{x, \alpha}(t), \hat{\alpha}^{x, \alpha}(t))$  denotes the solution of (5.3) with initial value  $(x, \alpha)$ . By the definition of  $\psi(x)$  and (5.2), we have that

$$\mathbb{P}\{X^{x+h, \alpha}(t) = \hat{X}^{x+h, \alpha}(t), \alpha^{x+h, \alpha}(t) = \hat{\alpha}^{x+h, \alpha}(t)\} > 1 - \varepsilon \text{ if } |h| \leq \delta. \quad (5.5)$$

In view of (5.4) and (5.5) and the assumption that  $|f(x, \alpha)| \leq 1$  for all  $(x, \alpha) \in \mathbb{R}^r \times \mathcal{M}$ , we obtain

$$\left| \mathbb{E}f(X^{x+h, \alpha}(t), \alpha^{x+h, \alpha}(t)) - \mathbb{E}f(X^{x, \alpha}(t), \alpha^{x, \alpha}(t)) \right| < 7\varepsilon \text{ for any } h \in \mathbb{R}^r, |h| \leq \delta. \quad (5.6)$$

The Feller property is therefore proved.  $\square$

## 6. An example

As an application of the well-posedness properties studied in the previous sections, this section deals with a competitive Lotka–Volterra system with regime switching. Such a model and many of its variants were extensively investigated in the literature; we refer the reader to [21] and many references therein for the recent developments.

**Example 6.1.** Consider a stochastic competitive Lotka–Volterra model with regime switching

$$\begin{aligned} dX_i(t) = & X_i(t) \left[ b_i(\alpha(t)) - \sum_{j=1}^r a_{ij}(\alpha(t))X_j(t) \right] dt \\ & + X_i(t)\sigma_i(\alpha(t))dW_i(t), \quad i = 1, \dots, r, \end{aligned} \quad (6.1)$$

where  $b_i(\cdot), \sigma_i(\cdot), i \in \{1, \dots, r\}, a_{ij}(\cdot), i, j \in \{1, \dots, r\}$  are functions from  $\mathcal{M}$  to  $\mathbb{R}$  and  $a_{ii}(k) > 0, a_{ij}(k) \geq 0, i, j \in \{1, \dots, r\}, k \in \mathcal{M}, W_i(t), i \in \{1, \dots, r\}$  are Brownian motions,  $\alpha(t)$  is the switching process taking value in  $\mathcal{M} = \{1, \dots, m_0\}$  with generators  $Q(x) = (q_{ij}(x))_{m_0 \times m_0}$ . Assume that  $q_{ij}(\cdot), i, j \in \mathcal{M}$  are bounded and continuous. (6.1) can be written in the matrix form

$$dX(t) = \text{diag}(X(t)) [b(\alpha(t)) - A(\alpha(t))X(t)] dt + \text{diag}(X(t))\text{diag}(\sigma(\alpha(t)))dW(t), \quad (6.2)$$

where  $b(k) = (b_1(k), \dots, b_r(k)), A(k) = (a_{ij}(k))_{r \times r}, \sigma(k) = (\sigma_1(k), \dots, \sigma_r(k))$  and  $W(t) = (W_1(t), \dots, W_r(t))$ . The model (6.1) with Markovian switching, that is, when  $Q(x)$  is a constant matrix, was studied in [21]. Although we are considering a more complex model with state-dependent switching, following the proofs of [21, Theorems 2.1, 3.1], we can still obtain that

- For any  $x \in \mathbb{R}_+^{r, \circ} := \{(x_1, \dots, x_r) : x_i > 0, i = 1, \dots, r\}$ , there exists a unique global solution  $(X^{x, \alpha}(t), \alpha^{x, \alpha}(t))$  with  $X^{x, \alpha}(t) = (X_1^{x, \alpha}(t), \dots, X_r^{x, \alpha}(t))$  to (6.1) with initial value  $x$ . Moreover,

$$\mathbb{P}\{X^x(t) \in \mathbb{R}_+^{r, \circ} \forall t \geq 0\} = 1.$$

- For any  $m > 0$ , there exists a constant  $K_m > 0$  such that

$$\mathbb{E}|X^x(t)|^m \leq K_m(1 + |x|^m) \text{ for all } x \in \mathbb{R}_+^{n,o}, t \geq 0, \quad (6.3)$$

where we use the norm  $|x| = \sum_{i=1}^r |x_i|$  for  $x = (x_1, \dots, x_r) \in \mathbb{R}^r$ .

We aim to show that the model (6.1) satisfies the conclusions of the theorems in previous sections whose proofs rely on estimates (3.6) and (3.7). Since the coefficients of (6.1) is not Lipschitz, to obtain the desired results we need to use (6.3) and the following lemma.

**Lemma 6.2.** Let  $R > 0$ . For any  $x, y \in \mathbb{R}_+^{r,o}$  and  $|x| \vee |y| \leq R$ .

$$\mathbb{E} \sup_{t \in [0, T]} \{|X^{x, \alpha}(t \wedge \tau) - X^{y, \alpha}(t \wedge \tau)|^2\} \leq K_{R, T} |x - y|^2. \quad (6.4)$$

where  $K_{R, T}$  depends only on  $R$  and  $T$  and

$$\tau = \inf\{t \geq 0 : \alpha^{x, \alpha}(t) \neq \alpha^{y, \alpha}(t)\}.$$

**Proof.** Using the elementary estimate  $|\text{diag}(x)A(k)x - \text{diag}(y)A(k)y| \leq C_A(|x| + |y|)|x - y|$  for some  $C_A > 0$ , we obtain

$$\begin{aligned} & |X^{x, \alpha}(t \wedge \tau) - X^{y, \alpha}(t \wedge \tau)| \\ & \leq |x - y| + \sum_{i=1}^r \int_0^{t \wedge \tau} |b_i(X_i^{x, \alpha}(s) - X_i^{y, \alpha}(s))| ds \\ & \quad + C_A \int_0^{t \wedge \tau} (|X^{x, \alpha}(s)| + |X^{y, \alpha}(s)|) |X_i^{x, \alpha}(s) - X_i^{y, \alpha}(s)| ds \\ & \quad + \sum_{i=1}^r \left| \int_0^{t \wedge \tau} \sigma_i(\alpha(s))(X_i^{x, \alpha}(s) - X_i^{y, \alpha}(s)) dW_i(s) \right|. \end{aligned}$$

It follows from the Cauchy–Schwarz inequality that

$$\begin{aligned} & |X^{x, \alpha}(t \wedge \tau) - X^{y, \alpha}(t \wedge \tau)|^2 \\ & \leq C|x - y|^2 + C \left( \int_0^{t \wedge \tau} (1 + |X^{x, \alpha}(s)| + |X^{y, \alpha}(s)|) |X^{x, \alpha}(s) - X^{y, \alpha}(s)| ds \right)^2 \\ & \quad + C \sum_{i=1}^r \left| \int_0^{t \wedge \tau} \sigma_i(\alpha(s))(X_i^{x, \alpha}(s) - X_i^{y, \alpha}(s)) dW_i(s) \right|^2 \end{aligned} \quad (6.5)$$

for some constant  $C > 0$ . By the Burkholder–Davis–Gundy Inequality,

$$\begin{aligned} & C \mathbb{E} \sup_{t \in [0, T]} \left\{ \sum_{i=1}^r \left| \int_0^{t \wedge \tau} \sigma_i(\alpha(s))(X_i^{x, \alpha}(s) - X_i^{y, \alpha}(s)) dW_i(s) \right|^2 \right\} \\ & \leq \tilde{C} \mathbb{E} \int_0^{T \wedge \tau} |X^{x, \alpha}(s) - X^{y, \alpha}(s)|^2 ds \end{aligned} \quad (6.6)$$

for some constant  $\tilde{C}$ . In view of Hölder's inequality,

$$\mathbb{E} \sup_{t \in [0, T]} \left\{ C \left( \int_0^{t \wedge \tau} (1 + |X^{x, \alpha}(s)| + |X^{y, \alpha}(s)|) |X^{x, \alpha}(s) - X^{y, \alpha}(s)| ds \right)^2 \right\}$$



$$\begin{aligned}
&= C \mathbb{E} \left( \int_0^{T \wedge \tau} (1 + |X^{x,\alpha}(s)| + |X^{y,\alpha}(s)|) |X^{x,\alpha}(s) - X^{y,\alpha}(s)| ds \right)^2 \\
&\leq C \left[ \mathbb{E} \int_0^{T \wedge \tau} (1 + |X^{x,\alpha}(t)| + |X^{y,\alpha}(t)|)^2 dt \right] \mathbb{E} \int_0^{T \wedge \tau} |X^{x,\alpha}(t) - X^{y,\alpha}(t)|^2 dt \\
&\leq K(1+T)(1+|x|^2+|y|^2) \mathbb{E} \int_0^{T \wedge \tau} |X^{x,\alpha}(t) - X^{y,\alpha}(t)|^2 dt \quad (\text{due to (6.3)}) \\
&\leq K(1+T)(1+|x|^2+|y|^2) \mathbb{E} \int_0^T \sup_{s \in [0,t]} \{ |X^{x,\alpha}(s \wedge \tau) - X^{y,\alpha}(s \wedge \tau)|^2 \} dt \quad (6.7)
\end{aligned}$$

for some  $K > 0$ . Taking the supreme over  $[0, T]$ , followed by taking the expectation on both sides of (6.5), and using (6.6) and (6.7), we have

$$\begin{aligned}
&\mathbb{E} \sup_{t \in [0,T]} \{ |X^{x,\alpha}(t \wedge \tau) - X^{y,\alpha}(t \wedge \tau)|^2 \} \\
&\leq C|x - y|^2 + K(1+T)(1+|x|^2+|y|^2) \mathbb{E} \\
&\quad \times \int_0^T \sup_{s \in [0,t]} \{ |X^{x,\alpha}(s \wedge \tau) - X^{y,\alpha}(s \wedge \tau)|^2 \} dt \quad (6.8)
\end{aligned}$$

for some constant  $K > 0$ . Then (6.4) follows from the Gronwall inequality.  $\square$

Although the coefficients of (6.1) are not globally Lipschitz, the estimates (6.3) and (6.4) are sufficient for us to derive the following results.

**Theorem 6.3.** Assume that  $q_{ij}(\cdot)$ ,  $i, j \in \mathcal{M}$  are bounded and continuous. Let  $(X^{x,\alpha}(t), \alpha^{x,\alpha}(t))$  be the solution to (6.1) and (2.3) with initial value  $(x, \alpha) \in \mathbb{R}_+^{r,\circ} \times \mathcal{M}$ . The following assertions hold:

1.  $X^{x,\alpha}(t)$  is twice continuously differentiable with respect to  $x$  in probability. If in addition,  $q_{kj}(x)$  satisfies (3.2) then  $X^{x,\alpha}(t)$  is twice continuously differentiable in  $L^p$  with respect to  $x$  for any  $0 < p < \lambda$ , where  $\lambda$  is the Hölder exponent in (3.2).
2. If  $q_{kj}(x)$  satisfies (4.1) then for any  $R$  and  $T > 0$ , there is a  $C_{R,T} > 0$  such that for any  $x, \tilde{x} \in \mathbb{R}_+^{r,\circ}$ ,  $|x| \vee |\tilde{x}| \leq R$ , and  $\alpha \in \mathcal{M}$ , we have

$$\mathbb{E} \sup_{t \in [0,T]} |X^{\tilde{x},\alpha}(t) - X^{x,\alpha}(t)| \leq C_{R,T} |\tilde{x} - x|.$$

3. Assume that for each  $i, j \in \mathcal{M}$ ,  $q_{ij}(\cdot) \in C^2$  and  $|D^\beta q_{ij}(\cdot)|$  are Lipschitz and bounded uniformly by some constant  $M$  for  $|\beta| \leq 2$ . Let  $\phi(\cdot, i) \in C^2$  satisfy

$$|D_x^\beta \phi(x, i)| \leq K(1+|x|^\gamma), i \in \mathcal{M}, |\beta| \leq 2.$$

Then,  $u(t, x, i) = \mathbb{E}[\phi(X^{x,i}(t), \alpha^{x,i}(t))]$  is twice continuously differentiable with respect to the variable  $x \in \mathbb{R}_+^{r,\circ}$ .

4. The solution process  $(X(t), \alpha(t))$  for the system given by (6.1) and (2.3) is a Markov–Feller process.

**Proof.** In the proof of Theorem 3.1, we use the global Lipschitz to obtain (3.6). In this example, the constant  $K$ , depending only on  $T$ , in (3.6) is replaced by  $K_{R,T}$ , which depends on both  $R$  and  $T$  (see (6.4)). Although (6.4) is slightly weaker than (3.6), it is still sufficient to follow the proofs of Theorems 3.1 and 4.3 to obtain the first and third claims of Theorem 6.3, respectively. The second claim is derived from Theorem 4.1 with the minor modification that the constant  $C_T$

in (4.2) is replaced by  $C_{R,T}$  because the constant  $K_{R,T}$  in (6.4) depends on  $R$ . The fourth claim follows directly from Theorem 5.1.  $\square$

### Acknowledgment

The first and second authors research was supported in part by the National Science Foundation under grant DMS-1207667.

### References

- [1] D. Applebaum, Lévy Processes and Stochastic Calculus, second ed., Cambridge University Press, Cambridge, 2009.
- [2] J. Bao, J. Shao, C. Yuan, Approximation of invariant measures for regime-switching diffusions, Potential Anal. 44 (2016) 707–727.
- [3] M.F. Chen, From Markov Chains to Non-Equilibrium Particle Systems, World Scientific, Singapore, 2004.
- [4] A. Eizenberg, M. Freidlin, On the Dirichlet problem for a class of second order PDE systems with small parameter, Stoch. Stoch. Rep. 33 (1990) 111–148.
- [5] A. Friedman, Stochastic Differential Equations and Applications, Vols. I and II, Academic Press, New York, 1975.
- [6] I.I. Gihman, A.V. Skorohod, Introduction to the Theory of Random Processes, W.B. Saunders, Philadelphia, PA, 1969.
- [7] G. Huang, M. Mandjes, P. Spreij, Large deviations for Markov-modulated diffusion processes with rapid switching, Stochastic Process. Appl. 126 (6) (2016) 1785–1818.
- [8] V. Krishnamurthy, O.N. Gharehshiran, M. Hamdi, Interactive sensing and decision making in social networks, Found. Trends Signal Process. 7 (1–2) (2015) 1–196.
- [9] R.H. Liu, Optimal investment and consumption with proportional transaction costs in regime-switching model, J. Optim. Theory Appl. 163 (2014) 614–641.
- [10] X. Mao, C. Yuan, Stochastic Differential Equations with Markovian Switching, Imperial College Press, London, UK, 2006.
- [11] J. Shao, Ergodicity of regime-switching diffusions in Wasserstein distances, Stochastic Process. Appl. 125 (2015) 739–758.
- [12] A.V. Skorokhod, Asymptotic Methods in the Theory of Stochastic Differential Equations, in: Translations of Mathematical Monographs, vol. 78, American Mathematical Society, Providence, RI, 1989. Translated from the Russian by H. H. McFaden.
- [13] F.B. Xi, Feller property and exponential ergodicity of diffusion processes with state-dependent switching, Sci. China A 51 (2008) 329–342.
- [14] F. Xi, Asymptotic properties of jump-diffusion processes with state-dependent switching, Stochastic Process Appl. 119 (2009) 2198–2221.
- [15] G. Yin, Q. Zhang, Continuous-Time Markov Chains and Applications: A Two-Time-Scale Approach, second ed., Springer, New York, 2013.
- [16] G. Yin, C. Zhu, Hybrid Switching Diffusions: Properties and Applications, Springer, New York, 2010.
- [17] G. Yin, C. Zhu, Properties of solutions of stochastic differential equations with continuous-state-dependent switching, J. Differential Equations 249 (2010) 2409–2439.
- [18] C. Yuan, X. Mao, Asymptotic stability in distribution of stochastic differential equations with Markovian switching, Stochastic Process. Appl. 103 (2003) 277–291.
- [19] Q. Zhang, Nonlinear filtering and control of a switching diffusion with small observation noise, SIAM J. Control Optim. 36 (1998) 1738–1768.
- [20] Q. Zhang, Stock trading: An optimal selling rule, SIAM J. Control Optim. 40 (2001) 64–87.
- [21] C. Zhu, G. Yin, On hybrid competitive Lotka–Volterra ecosystems, Nonlinear Anal. 71 (12) (2009) 1370–1379.
- [22] C. Zhu, G. Yin, On strong Feller, recurrence, and weak stabilization of regime-switching diffusions, SIAM J. Control Optim. 48 (2009) 2003–2031.