2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Hybrid Switching Diffusions: Continuity and Differentiability

G. Yin and C. Zhu

Abstract— **This work focuses on hybrid switching diffusion systems. After presenting the basic models, a difficult and fundamental problem, namely, well posedness is addressed. In particular, continuous and smooth dependence on initial data is treated. The main ideas are presented, whereas the verbatim proofs can be found in [35].**

Key Words. Hybrid switching diffusion, weak continuity, smooth dependence.

1. Introduction

Because of their wide range of applicability, hybrid switching diffusion systems have drawn growing attention especially in the fields of control and optimization recently, Such systems are capable of describing complex systems and their inherent uncertainty and randomness in the environment; the formulation is versatile and provides more opportunity for realistic modeling, but adds substantial difficulties in the analysis. Much of the study comes from applications arising in control engineering, manufacturing systems, estimation, identification, and filtering, two-time-scale systems, and financial engineering; see for example, [11], [14], [20], [21], [22], [26], [28], [30], [37], among others. Random-switching processes are used to model demand rate or machine capacity in production planning, to describe the volatility changes over time to capture discrete shifts such as market trends and interest rates etc. in finance and insurance, and to model time-varying parameter for network problems.

Many real world applications in the new era require complex models, in which the traditional dynamic system setup using continuous dynamics given by differential equations alone is inadequate. For example, one of the early efforts of using such hybrid models for options in a financial market can be traced back to Barone-Adesi and Whaley [1], in which both the return rate and the volatility rate of a stock depend on a continuous-time Markov chain. For instance, in the simplest case, a stock market may be considered to have two "modes" or "regimes," up and down, resulted from the state of the underlying economy, the general mood of investors in the market, and so on. The rationale is that in the different regimes, the volatility and return rates are substantially different. Introducing hybrid models makes it possible to describe stochastic volatility in a relatively simple manner. That is, it is much simpler than the so-called stochastic volatility models, which requires the augmentation by another diffusion process to describe the volatility. Consider another example of a wireless communication network. The performance analysis of an adaptive linear multiuser detector in a cellular direct-sequence code-division multiple-access wireless network with changing user activity due to an admission or access controller at the base station. Under suitable conditions, an associated optimization problem leads to a switching diffusion limit [29]. To summarize, a center piece in the applications mentioned above is a two-component Markov process, a continuous component and a discrete-event component.

Because of the salient features of the coexistence of continuous dynamics and discrete events and their interactions, switching diffusions have drawn resurgent attentions. In addition, there have been increasing demands for modeling large-scale and complex systems, designing optimal controls, and conducting optimization tasks. In the traditional setup, the design of a feedback controller is based on a plant with fixed parameters, which is inadequate when the actual system differs from the assumed nominal model. Much effort has been directed to the design of more "robust" controls in recent years. Studies of hybrid systems with Markov regime switching contribute to this end. Various regime-switching models have been proposed and examined. The so-called jump linear systems, widely used in engineering, have been studied in Mariton [20]; controllability and stabilizability of such systems are treated in Ji and Chizeck [14]. Estimation problems are considered in Sworder and Boyd [26]. Manufacturing and production planning under the framework of hierarchical structure are covered in Sethi and Zhang [21], and Sethi, Zhang, and Zhang [22]. Financial engineering applications and the use of hybrid geometric Brownian motion models can be found in [30], [37], among others. To reduce the complexity, effort has been made to deal with hybrid systems with regime switching by means of time-scale separation in Yin and Zhang [31]; see also [28]. Hybrid systems have received increasing attention in recent years. Other than the switching diffusion systems mentioned above, one may find the formulation in somewhat different setup in Bensoussan and Menaldi [2]. A treatment of stochastic hybrid systems with applications to communication networks is in Hespanha [10]. Recently, stability of hybrid switching diffusion processes have received much attention; see, for example, [15], [17], [19], [39], [40], [42] and the references therein. Numerical methods for switching diffusions and related control problems can be found in [24], [25], [27], [38]. For references on stochastic control and controlled Markov processes, we refer the reader to Fleming and Rishel

Department of Mathematics, Wayne State University, Detroit, MI 48202, gyin@math.wayne.edu. This research was supported in part by the National Science Foundation under DMS-0907753, and in part by the Air Force Office of Scientific Research under FA9550-10-1-0210.

Department of Mathematical Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53201, zhu@uwm.edu. This research was supported in part by the National Science Foundation under DMS-1108782, and in part by a grant from the UWM Research Growth Initiative.

[8], Krylov [16], and Yong and Zhou [36].

Although it is seemingly not much different from a diffusion, even if the switching is a finite-state Markov chain independent of the Brownian motion (subsequently referred to as Markov-modulated switching diffusions), the interactions due to the switching process makes the analysis much more difficult. One of the main features of our recent works [32], [33], [34], [35], [39], [40] is: A large part of it deals with the switching component depending on the continuous component. This is often referred to as state-dependent switching process in what follows. For such processes, properties as recurrence, ergodicity were considered in Zhu and Yin [39], and strong Feller and weak stabilization were treated in [40]. As demonstrated in Zhu, Yin, and Song [42], in the fully degenerate case, the classical Hartman–Grobman theorem has new twist when random switching is considered. In any event, for state-dependent switching diffusions, the analysis is much more difficult than that of the Markovian-switching cases. Although the existence and uniqueness of the solutions of switching diffusions can be obtained [34], the well posedness turns out to much more difficult to deal with. Here, by well posedness we refer to the continuous and smooth dependence of initial data. The main difficulty lies in that since the switching component depends on the continuous component. The sample paths of switching diffusion with different initial data on the continuous component will be infinitely often different. In this paper, we take up this issue and present a positive answer to the proposed question.

The rest of the paper is arranged as follows. The mathematical formulation of the switching diffusions is given next. Then Section 3 is devoted to two main issues, namely weak continuity and smooth dependence on the initial data (to be more precisely on the continuous component).

2. Switching Diffusions

We begin this section with some intuitive descriptions and examples. Then we present the mathematics models.

What is a Switching Diffusion? Roughly speaking, switching diffusions are systems involve both diffusion and discrete jumps. In our setup, the discrete events are modeled by a finite-state process, whereas the continuous events are diffusion like processes. Consider a switching diffusion that consists of three diffusions sitting on three parallel plans. The discrete event is a three-state jump process. We denote the pair of processes by (continuous process, discrete event) = $(X(t), \gamma(t))$. Suppose that initially, the process is at $(X(0), \gamma(0)) = (x, 1)$. The discrete event process sojourns in discrete state 1 for a random duration; during this period, the continuous component evolves according to the diffusion process specified by the drift and diffusion coefficients associated with discrete state 1 until a jump takes place for the discrete component. At random moment τ_1, a jump to discrete state 3 occurs. Then the continuous component evolves according to the diffusion process whose drift and diffusion coefficients are determined by discrete event 3. The process wanders around in the third plan until another random jump time τ_2. At τ_2, the system switches to the second parallel plan and follows another diffusion with different drift and diffusion coefficients and so on.

Examples of Switching Diffusion Systems. To demonstrate the utility of the switching diffusion models, this section provides a number of examples.

Example 2.1:. The well-known Lotka-Volterra models concern ecological population modeling. The models have been extensively studied in the literature. When two or more species live in proximity and share the same basic requirements, they usually compete for resources, food, habitat, or territory. Both deterministic and stochastic models of the Lotka-Volterra systems have been studied extensively.

It has been noted that the growth rates and the carrying capacities are often subject to environmental noise. Moreover, the qualitative changes of the growth rates and the carrying capacities form an essential aspect of the dynamics of the ecosystem. These changes usually cannot be described by the traditional (deterministic or stochastic) Lotka-Volterra models. For instance, the growth rates of some species in the rainy season will be much different from those in the dry season. Moreover, the carrying capacities often vary according to the changes in nutrition and food resources. Similarly, the interspecific or intraspecific interactions differ in different environment.

Therefore, it is natural to consider a stochastic Lotka-Volterra ecosystem in random environment that can be formulated by use of an additional factor process. For $i = 1, \ldots, r$, let $x_i(t)$ denote the population size of the ith species in the ecosystem at time t. Consider the following stochastic differential equation with regime switching

$$\begin{aligned} dx_i(t) &= x_i(t)\Big\{ \Big[b_i(\gamma(t)) - \sum_{j=1}^{r} a_{ij}(\gamma(t))x_j(t)\Big]dt \\ &\qquad + \sigma_i(\gamma(t)) \circ dw_i(t)\Big\},\ i = 1, \ldots, r, \end{aligned} \tag{2.1}$$

where $w(t) = (w_1(t), \ldots, w_r(t))'$ is an r-dimensional standard Brownian motion, $\gamma(t) \in \mathcal{S} = \{1, \ldots, m_0\}$ is a continuous time Markov chain describing the random environments, and for $\gamma \in \mathcal{S}$, $b(\gamma) = (b_1(\gamma), \ldots, b_r(\gamma))'$, $A(\gamma) = (a_{ij}(\gamma))$, $\Sigma(\gamma) = \mathrm{diag}(\sigma_1(\gamma), \ldots, \sigma_r(\gamma))$ represent different growth rates, community matrices, and noise intensities in different external environments, respectively, and z' denotes the transpose of z. The above formulation is seen to be in the form of Stratonovich. This form is often considered to be more suitable for environmental modeling.

Regime-switching stochastic Lotka-Volterra models have drawn much attention lately. For instance, the study of trajectory behavior of Lotka-Volterra competition bistable systems and systems with telegraph noises, stochastic population dynamics under regime switching, the dynamics of a population in a Markovian environment, the evolution of a system composed of two predator-prey deterministic systems described by Lotka-Volterra equations in random environment were investigated by a host of researchers; see [41] for an extensive list of references. In the absence of regime switching, the system is completely modeled by stochastic time evolution in a fixed environment. The results in a fixed environment correspond to ours in the case when the Markov chain has only one state or the Markov chain always stays in the fixed state (environment). When random environments are considered, the system's qualitative behavior can be drastically different; see [41].

Example 2.2:. Consider a regime-switching linear system

$$\dot{x}(t) = A(\gamma(t))x(t) + B(\gamma(t))u(t),$$

where $\gamma(t)$ is a continuous-time Markov chain taking values in a finite set $\mathcal{S} = \{1, \dots, m_0\}$, $A(i)$ and $B(i)$ for $i \in \mathcal{S}$ are matrices with compatible dimensions, and $u(\cdot)$ is the control. In lieu of one linear system, we have a number of systems interconnected through the Markov chain. Such systems have enjoyed numerous applications in emerging application areas as financial engineering, wireless communications, as well as in existing applications. A class of important problems concerns the asymptotic behavior of such systems when they are in operations for a long time. Very often, in many engineering problems, one is more interested in the system stability. Much interest lies in finding admissible controls so that the resulting system will be stabilized. In [20], stabilization for robust controls of jump linear quadratic (LQ) control problems was treated. In [14], both controllability and stabilizability of jump linear LQ systems were considered. In [4], adaptive LQG problems with finite-state process parameters were treated. Additional difficulties come when the switching process $\gamma(t)$ cannot be observed, but we can only observe it with a noise added. That is, we can observe

$$dX(t) = [A(\gamma(t))X(t) + B(\gamma(t))u(t)]dt + dw(t).$$

For such partially observed systems, it is natural to use nonlinear filtering techniques. The associated filter is the well-known Wonham filter, which is one of a handful of finite dimensional filters in existence. Stabilization of linear systems with hidden Markov chains were considered in [5], [7]. In both of these references, averaging criteria were used for the purpose of stabilization;

$$\limsup_{t\to\infty} E[|X(t)|^2 + |u(t)|^2] < \infty,$$

in [7], whereas stabilization under

$$\limsup_{t\to\infty} \frac{1}{t} E \int_0^t [|X(s)|^2 + |u(s)|^2] ds < \infty$$

was considered in [5]. A question of considerable practical interest is: Can we design controls so that the resulting system will be stable in the almost sure sense. Using Wonham filters and converting the partial observed system to an equivalent fully observed system, almost sure stabilizing controls were found under the criterion

$$\limsup_{t\to\infty} \frac{1}{t} \log |X(t)| \le 0 \quad \text{almost surely} \tag{2.2}$$

in [3]. The main idea lies in analyzing the sample path properties using a suitable Liapunov function. We refer the reader to the reference given above for further details.

Formulation. Let $(\Omega, \mathcal{G}, \{\mathcal{G}_t\}_{t\ge 0}, P)$ be a complete probability space with a filtration $\{\mathcal{G}_t\}_{t\ge 0}$ satisfying the usual condition (i.e., it is right continuous with $\mathcal{G}_0$ containing all P-null sets). Let $x \in \mathbb{R}^r$, $\mathcal{S} = \{1, \dots, m_0\}$, and $\Psi(x) = (\psi_{ij}(x))$ an $m_0 \times m_0$ matrix depending on x satisfying the q-property, which means that for any $x \in \mathbb{R}^r$, $\psi_{ij}(x) \ge 0$ for $i \ne j$ and $\sum_{j=1}^{m_0} \psi_{ij}(x) = 0$. For any twice continuously differentiable function $h(\cdot, i)$, $i \in \mathcal{S}$, define $\mathcal{L}$ by

$$\begin{aligned}\mathcal{L}h(x,i) \ &= \frac{1}{2}\mathrm{tr}(a(x,i)\nabla^2 h(x,i)) \\ &\quad + b'(x,i)\nabla h(x,i) + \Psi(x)h(x,\cdot)(i),\end{aligned} \tag{2.3}$$

where $\nabla h(\cdot, i)$ and $\nabla^2 h(\cdot, i)$ denote the gradient and Hessian of $h(\cdot, i)$, respectively, $b'(x,i)\nabla h(x,i)$ denotes the usual inner product on $\mathbb{R}^r$, and for each $i \in \mathcal{S}$

$$\Psi(x)h(x,\cdot)(i) \ = \sum_{j\in\mathcal{S}} \psi_{ij}(x)(h(x,j) - h(x,i)). \tag{2.4}$$

Consider a Markov process $Y(t) = (X(t), \gamma(t))$, whose associated operator is given by $\mathcal{L}$. Note that $Y(t)$ has two components, an r-dimensional continuous component $X(t)$ and a discrete component $\gamma(t)$ taking values in $\mathcal{S} = \{1, \dots, m_0\}$.

Recall that the process $Y(t) = (X(t), \gamma(t))$ may be described by the following pair of equations:

$$\begin{aligned}&dX(t) = b(X(t), \gamma(t))dt + \sigma(X(t), \gamma(t))dw(t), \\ &X(0) = x, \ \gamma(0) = \gamma,\end{aligned} \tag{2.5}$$

and

$$\begin{aligned}&P\{\gamma(t+\delta) = j | \gamma(t) = i, X(s), \gamma(s), s \le t\} \\ &\quad = \psi_{ij}(X(t))\delta + o(\delta), \ i \ne j,\end{aligned} \tag{2.6}$$

where $w(t)$ is a d-dimensional standard Brownian motion, $b(\cdot,\cdot) : \mathbb{R}^r \times \mathcal{S} \mapsto \mathbb{R}^r$, and $\sigma(\cdot,\cdot) : \mathbb{R}^r \times \mathcal{S} \mapsto \mathbb{R}^{r\times d}$ satisfying $\sigma(x,i)\sigma'(x,i) = a(x,i)$. Note that (2.5) depicts the system dynamics and (2.6) delineates the probability structure of the jump process. Note that if $\gamma(\cdot)$ is a continuous-time Markov chain independent of the Brownian motion $w(\cdot)$ and $\Psi(x) = \Psi$ or $\Psi(x) = \Psi(t)$ (independent of x), then equation (2.5) together with the generator Ψ or $\Psi(t)$ is sufficient to characterize the underlying process. As long as there is an x-dependence, equation (2.6) is needed in delineating the dynamics of the switching diffusion. By considering x-dependent generator $\Psi(x)$, our model provides more realistic formulation allowing the switching component depending on the continuous states. This, in turn, allows the coupling and correlation between $X(t)$ and $\gamma(t)$.

The evolution of the discrete component or the switching process $\gamma(\cdot)$ can be represented by a stochastic integral with respect to a Poisson random measure $\mathfrak{p}(dt, dz)$, whose intensity is $dt \times m(dz)$, where $m(\cdot)$ is the Lebesgue measure on $\mathbb{R}$. The compensated or centered Poisson measure $\mu(ds, dz) = \mathfrak{p}(ds, dz) - ds \times m(dz)$ is a martingale measure. We refer to [18], [23], [34] for details.

Similar to diffusions, for each $\imath \in \mathcal{S}$ and each $f(\cdot, \imath) \in C^2$, a result known as generalized Itô's lemma (see [18], [23]) reads

$$f(X(t),\gamma(t))-f(X(0),\gamma(0)) \\ =\int_0^t \mathcal{L}f(X(s),\gamma(s))ds+M_1(t)+M_2(t), \tag{2.7}$$

where $\mathcal{L}$ is the operator defined in (2.3), and

$$M_1(t)=\int_0^t \big\langle \nabla f(X(s),\gamma(s)),\sigma(X(s),\gamma(s))dw(s)\big\rangle,$$
$$M_2(t)=\int_0^t\int_{\mathbb{R}}\big[f(X(s),\gamma(0)+h(X(s),\gamma(s),z)) \\ -f(X(s),\gamma(s))\big]\mu(ds,dz),$$

where $\langle z,y\rangle$ denotes the usual inner product on $\mathbb{R}^r$ and h is an integer-valued function (see [34] for more details).

In view of the generalized Itô formula,

$$M_f(t)=f(X(t),\gamma(t))-f(X(0),\gamma(0)) \\ -\int_0^t \mathcal{L}f(X(s),\gamma(s))ds \tag{2.8}$$

is a (local) martingale. Similar to the case of diffusion processes, we can define the corresponding notion of solution of martingale problem accordingly.

Theorem 2.3: Let $x\in\mathbb{R}^r$, $\mathcal{S}=\{1,\dots,m_0\}$, and $\Psi(x)=(\psi_{ij}(x))$ an $m_0\times m_0$ matrix depending on x satisfying the q-property. Consider the two-component process $Y(t)=(X(t),\gamma(t))$ given by (2.5)–(2.6) with initial data (x,γ). Suppose that $\Psi(\cdot):\mathbb{R}^r\mapsto\mathbb{R}^{m_0\times m_0}$ is a bounded and continuous function, that the functions $b(\cdot,\cdot)$ and $\sigma(\cdot,\cdot)$ satisfy

$$|b(x,\gamma)|+|\sigma(x,\gamma)|\le K_0(1+|x|),\quad \gamma\in\mathcal{S}, \tag{2.9}$$

for some $K_0>0$ and that for every integer $N\ge 1$, there exists a positive constant M_N such that for all $i\in\mathcal{S}$ and all $x,y\in\mathbb{R}^r$ with $|x|\vee|y|\le M_N$,

$$|b(x,i)-b(y,i)|\vee|\sigma(x,i)-\sigma(y,i)|\le M_N|x-y|. \tag{2.10}$$

Then there exists a unique solution $(X(t),\gamma(t))$ to equation (2.5) with given initial data in which the evolution of the jump process is specified by (2.6).

The proof of the theorem is deemed to be well known. For brevity, the detailed proof is omitted. Instead, we make the following remarks. There are a number of possible proofs. For example, the existence can be obtained as in [23, pp. 103–104]. Viewing the switching diffusion as a special case of a jump-diffusion process prove the existence and uniqueness using [13, Section III.2]. Another possibility is to use a martingale problem formulation together with utilization of truncations and stopping times as in [12, Chapter IV]. In [27], we proposed and analyzed a couple of numerical approximation algorithms for approximating solutions of switching diffusions. We showed that the interpolations of the iterates converge weakly to the switching diffusion by a martingale problem formulation. Then using Lipschitz continuity and the weak convergence, we further obtain the strong convergence of the approximations. As a byproduct, we also obtained the existence and uniqueness of the solution.

3. Weak Continuity and Smooth Dependence on the Initial Data

A. Weak Continuity

Definition 3.1: Recall that a stochastic process $Y(t)$ with right continuous sample paths is said to be *weakly continuous* or *continuous in probability* at t if for any $\varepsilon>0$,

$$\lim_{\delta\to 0}P(|Y(t+\delta)-Y(t)|\ge\varepsilon)=0. \tag{3.1}$$

It is *mean square continuous* at t if

$$\lim_{\delta\to 0}E|Y(t+\delta)-Y(t)|^2=0. \tag{3.2}$$

The process $Y(t)$ is said to be continuous in probability in the interval $[0,T]$ (or in short continuous in probability if the interval $[0,T]$ is clearly understood), if it is continuous in probability at every $t\in[0,T]$. Likewise it is continuous in mean square if it is continuous in mean square at every $t\in[0,T]$.

Theorem 3.2: Suppose that the conditions of Theorem 2.3 are satisfied. Then the process $Y(t)=(X(t),\gamma(t))$ is continuous in probability and continuous in mean square.

Idea of Proof. We first consider the case when $\Psi(x)=\Psi$ and $\gamma(\cdot)$ is a continuous time Markov chain independent of the Brownian motion $w(\cdot)$. Note that for any $t\ge 0$, $\gamma(t)=\sum_{i=1}^{m_0} iI_{\{\gamma(t)=i\}}=\chi(t)(1,\dots,m_0)'$, where

$$\chi(t)=(I_{\{\gamma(t)=1\}},\dots,I_{\{\gamma(t)=m_0\}})\in\mathbb{R}^{1\times m_0}. \tag{3.3}$$

Using the fact that $\chi(t+\delta)-\chi(t)-\int_t^{t+\delta}\chi(s)\Psi ds$ is a martingale, we can show that $|\int_t^{t+\delta}\chi(s)\Psi ds|=O(\delta)$ and hence $E\,|\gamma(t+\delta)-\gamma(t)|^2\le\delta$. Then it follows that

$$E|(X(t+\delta),\gamma(t+\delta))-(X(t),\gamma(t))|^2\to 0 \text{ as } \delta\to 0.$$

Next, for the general case where $\gamma(\cdot)$ is generated by $\Psi(x)$, we can show that $|\int_t^{t+\delta}\chi(s)\Psi(X(s))ds|=O(\delta)$ a.s. Consequently, as in the special case above, we again have $E\,|\gamma(t+\delta)-\gamma(t)|^2\le\delta$ and hence $\lim_{\delta\to 0}E|(X(t+\delta),\gamma(t+\delta))-(X(t),\gamma(t))|^2=0$.

B. Smooth Dependence on the Initial Data x

Dealing with a continuous-time dynamic systems modeled by differential equations together with appropriate initial data, the well-posedness is crucial. As time-honored phenomena, the well-posedness appears in ordinary differential equations, partial differential equations together with initial and/or boundary data, stochastic differential equations, and stochastic differential equations with Markovian switching. A problem for the switching diffusion is well posed if there is a unique solution for the initial value problem and the solution continuously depends on the initial data.

Recall that a vector $\beta=(\beta_1,\dots,\beta_r)$ with nonnegative integer components is referred to as a multi-index. Put $|\beta|=\beta_1+\dots+\beta_r$, and define D_x^β as

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

Definition 3.3: Suppose that $\Phi(x_1,\dots,x_r)$ is a random function. Its partial derivative in mean square with respect to x_i for some $1\le i\le r$ is defined as the random variable $\widetilde{\Phi}(x_1,\dots,x_r)$ such that

$$\lim_{\Delta x_i\to 0} E\left|\frac{1}{\Delta x_i}\Big[\Phi(x_1,\dots,x_i+\Delta x_i,\dots,x_r) -\Phi(x_1,\dots,x_i,\dots,x_r)\Big]-\widetilde{\Phi}(x_1,\dots,x_r)\right|^2=0.$$

When the mean square partial derivative exists, we normally write it as

$$\widetilde{\Phi}(x_1,\dots,x_r)=\frac{\partial\Phi(x_1,\dots,x_r)}{\partial x_i}=\Phi_{x_i}(x_1,\dots,x_r).$$

Theorem 3.4: Assume the conditions of Theorem 2.3 with the modification of the local Lipschitz condition replaced by a global Lipschitz condition. Let $(X^{x,\gamma}(t),\gamma^{x,\gamma}(t))$ be the solution to the system given by (2.5) and (2.6). Assume that for each $i\in\mathcal{S}$, $b(\cdot,i)$ and $\sigma(\cdot,i)$ have continuous partial derivatives with respect to the variable x up to the second order and that

$$\left|D_x^\beta b(x,i)\right|+\left|D_x^\beta\sigma(x,i)\right|\le K_0(1+|x|^p), \tag{3.4}$$

where K_0 and p are positive constants and β is a multi-index with $|\beta|\le 2$. Then $X^{x,\gamma}(t)$ is twice continuously differentiable in mean square with respect to x.

Idea of Proof. a.) Without loss of generality, we prove Theorem 3.4 when $X(t)$ is 1-dimensional. Let $\delta\ne 0$ be small and denote $\widetilde{x}=x+\delta$. Let $(X(t),\gamma(t))$ be the switching diffusion process satisfying (2.5) and (2.6) with initial condition (x,γ) and $(\widetilde{X}(t),\widetilde{\gamma}(t))$ be the process starting from $(\widetilde{x},\gamma)$ (i.e., $(X(0),\gamma(0))=(x,\gamma)$) and $(\widetilde{X}(0),\widetilde{\gamma}(0))=(\widetilde{x},\gamma)$ respectively).

Fix any $T>0$ and let $0<t<T$. Put $\Xi^\delta(t)=\Xi^{x,\delta,\gamma}(t):=\frac{\widetilde{X}(t)-X(t)}{\delta}$. Then we have

$$\begin{aligned}\Xi^\delta(t)\ &=1+\eta^\delta(t)\\ &\quad+\frac{1}{\delta}\int_0^t[b(\widetilde{X}(s),\gamma(s))-b(X(s),\gamma(s))]ds\\ &\quad+\frac{1}{\delta}\int_0^t[\sigma(\widetilde{X}(s),\gamma(s))-\sigma(X(s),\gamma(s))]dw(s),\end{aligned} \tag{3.5}$$

where

$$\begin{aligned}\eta^\delta(t)\ &=\frac{1}{\delta}\int_0^t[b(\widetilde{X}(s),\widetilde{\gamma}(s))-b(\widetilde{X}(s),\gamma(s))]ds\\ &\quad+\frac{1}{\delta}\int_0^t[\sigma(\widetilde{X}(s),\widetilde{\gamma}(s))-\sigma(\widetilde{X}(s),\gamma(s))]dw(s).\end{aligned}$$

b.) Show that

$$\lim_{\delta\to 0}E\left[\sup_{0\le t\le T}|\eta^\delta(t)|^2\right]=0. \tag{3.6}$$

A crucial point is to use the technique of basic coupling of Markov processes (see for example, the book of Chen [6, p. 11]). Let $(\gamma(t),\widetilde{\gamma}(t))$ be a discrete random process with a finite state space $\mathcal{S}\times\mathcal{S}$ such that

$$\begin{aligned}&P\Big[(\gamma(t+\delta),\widetilde{\gamma}(t+\delta))=(j,i)\Big|\\ &\qquad\qquad(\gamma(t),\widetilde{\gamma}(t))=(k,l),(X(t),\widetilde{X}(t))=(x,\widetilde{x})\Big]\\ &\quad=\begin{cases}\widetilde{\psi}_{(k,l)(j,i)}(x,\widetilde{x})\delta+o(\delta), & \text{if } (k,l)\ne(j,i),\\ 1+\widetilde{\psi}_{(k,l)(k,l)}(x,\widetilde{x})\delta+o(\delta), & \text{if } (k,l)=(j,i),\end{cases}\end{aligned}$$

where $\delta\to 0$, and the matrix $(\widetilde{\psi}_{(k,l)(j,i)}(x,\widetilde{x}))$ is the basic coupling of matrices $\Psi(x)=(\psi_{kl}(x))$ and $\Psi(\widetilde{x})=(\psi_{kl}(\widetilde{x}))$ satisfying

$$\begin{aligned}&\widetilde{\Psi}(x,\widetilde{x})\widetilde{f}(k,l)\\ &\quad=\sum_{(j,i)\in\mathcal{S}\times\mathcal{S}}\psi_{(k,l)(j,i)}(x,\widetilde{x})(\widetilde{f}(j,i)-\widetilde{f}(k,l))\\ &\quad=\sum_j(\psi_{kj}(x)-\psi_{lj}(\widetilde{x}))^+(\widetilde{f}(j,l)-\widetilde{f}(k,l))\\ &\qquad+\sum_j(\psi_{lj}(\widetilde{x})-\psi_{kj}(x))^+(\widetilde{f}(k,j)-\widetilde{f}(k,l))\\ &\qquad+\sum_j(\psi_{kj}(x)\wedge\psi_{lj}(\widetilde{x}))(\widetilde{f}(j,j)-\widetilde{f}(k,l)),\end{aligned}$$

for any function $\widetilde{f}(\cdot,\cdot)$ defined on $\mathcal{S}\times\mathcal{S}$.

c.) For any fixed $T>0$, we have

$$E\left[\sup_{0\le t\le T}\left|X^{\widetilde{x},\gamma}(t)-X^{x,\gamma}(t)\right|^2\right]\le K\left|\widetilde{x}-x\right|^2, \tag{3.7}$$

where K is a constant depending only on T and the global Lipschitz and the linear growth constant K_0.

d.) Let $\xi(t):=\xi^{x,\gamma}(t)$ be the solution of

$$\begin{aligned}\xi(t)=1\ &+\int_0^t b_x(X(s),\gamma(s))\xi(s)ds\\ &+\int_0^t\sigma_x(X(s),\gamma(s))\xi(s)dw(s),\end{aligned} \tag{3.8}$$

where b_x and σ_x denote the partial derivatives of b and σ with respect to x, respectively. Then (3.5)–(3.7) and [9, Theorem 5.5.2] imply that

$$E\left|\Xi^\delta(t)-\xi(t)\right|^2\to 0\ \text{ as }\ \delta\to 0 \tag{3.9}$$

and $\xi(t)=\xi^{x,\gamma}(t)$ is mean square continuous with respect to x. Therefore, $(\partial/\partial x)X(t)$ exists in the mean square sense and $(\partial/\partial x)X(t)=\xi(t)$. Likewise, we can show that $(\partial^2/\partial x^2)X^{x,\gamma}(t)$ exists in the mean square sense and is mean square continuous with respect to x. The proof of the theorem is thus concluded.

Remark 3.5: If $\Psi(x)=\Psi$ and $\gamma(\cdot)$ is a continuous time Markov chain independent of the Brownian motion $w(\cdot)$, then the assertions of Theorem 3.4 can be established using essentially the same arguments as those for the counterparts of diffusion processes (see, for example, [9]). In our setup, however, the state-dependent-switching or the x-dependent $\Psi(x)$ creates much difficulty. The arguments for diffusion processes do not work here because $\widetilde{\gamma}(t)$ may not equal to $\gamma(t)$ for infinitely many $t>0$ even though $\widetilde{\gamma}(0)=\gamma(0)$. Nevertheless, with the coupling method, we are able to show (3.6), which is a key step in the proof of Theorem 3.4.

As a by-product of Theorem 3.4, we can show that the possibly degenerate initial value problem (3.10)–(**??**) has a *classical* solution.

Theorem 3.6: Assume the conditions of Theorem 3.4. In addition, suppose that $|\Psi(x)| \leq K$ for all $x \in \mathbb{R}^r$ and some $K > 0$, that $\eta(\cdot, i) \in C^2$ with $D_x^\theta \eta(\cdot, i)$ being Lipschitz continuous for each $i \in \mathcal{S}$ and multi-index θ with $|\theta| = 2$, and that $\left|D_x^\beta \eta(x, i)\right| \leq K(1 + |x|^p)$, where K and p are positive constants and β is a multi-index with $|\beta| \leq 2$. Then the function v defined by $v(t, x, i) := E_{x,i}[\eta(X(t), \gamma(t))]$ is continuously differentiable w.r.t. the variable t and twice continuously differential w.r.t. the variable x. Moveover, v satisfies the system of Kolmogorov backward equations

$$\begin{aligned} &\frac{\partial v(t, x, i)}{\partial t} = \mathcal{L}v(t, x, i), \quad (t, x, i) \in (0, T] \times \mathbb{R}^r \times \mathcal{S}, \\ &\lim_{t \downarrow 0} v(t, x, i) = \eta(x, i), \quad (x, i) \in \mathbb{R}^r \times \mathcal{S}, \end{aligned} \tag{3.10}$$

where $\mathcal{L}v(t, x, i)$ in (3.10) is to be interpreted as $\mathcal{L}$ applied to the function $(x, i) \mapsto v(t, x, i)$.

4. Further Remarks

After providing an overview of certain features and examples of switching diffusion processes, we examined the smooth dependence on the initial data of the switching diffusions, which is part of the fundamental issue of well posedness. These results are important for control systems. They can be used in control design, optimality seeking, wireless communication networks, stabilization and destabilization of controlled systems, financial engineering applications.

References

[1] G. Barone-Adesi and R. Whaley, Efficient analytic approximation of American option values, *J. Finance*, **42** (1987), 301–320.

[2] A. Bensoussan and J.L. Menaldi, Hybrid control and dynamic programming, *Dynamics Continuous Disc. Impulsive Sys.*, **3** (1997), 395–442.

[3] B. Bercu, F. Dufour, and G. Yin, Almost sure stabilization for feedback controls of regime-switching linear systems with a hidden Markov chain, *IEEE Trans. Automatic Control*, **54** (2009), 2114–2125.

[4] P.E. Caines and H.-F. Chen, Optimal adaptive LQG control for systems with finite state process parameters, *IEEE Trans. Automat. Control*, **30** (1985), 185–189.

[5] P.E. Caines and J.-F. Zhang, On the adaptive control of jump parameter systems via nonlinear filtering, *SIAM J. Control Optim.*, **33** (1995), 1758–1777.

[6] M.-F. Chen, *From Markov Chains to Non-equilibrium Particle Systems*, World Sci., Singapore, 2004.

[7] F. Dufour and P. Bertrand, Stabilizing control law for hybrid modes, *IEEE Trans. Automatic. Control*, **39** (1994), 2354–2357.

[8] W.H. Fleming and R.W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, NY, 1975.

[9] A. Friedman, *Stochastic Differential Equations and Applications*, Vols. I and II, Academic Press, New York, 1975.

[10] J.P. Hespanha, *Stochastic Hybrid Systems: Application to Communication Networks*, Springer, Berlin, 2004.

[11] A.M. Il'in, R.Z. Khasminskii, and G. Yin, Asymptotic expansions of solutions of integro-differential equations for transition densities of singularly perturbed switching diffusions, *J. Math. Anal. Appl.*, **238** (1999), 516–539.

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

[13] J. Jacod and A.N. Shiryayev, *Limit Theorems for Stochastic Processes*, Springer-Verlag, New York, 1980.

[14] Y. Ji and H.J. Chizeck, Controllability, stabilizability, and continuous-time Markovian jump linear quadratic control, *IEEE Trans. Automatic Control*, **35** (1990), 777–788.

[15] R.Z. Khasminskii, C. Zhu, and G. Yin, Stability of regime-switching diffusions, *Stochastic Process. Appl.*, **117** (2007), 1037–1051.

[16] N.V. Krylov, *Controlled Diffusion Processes*, Springer-Verlag, New York, 1980.

[17] X. Mao, Stability of stochastic differential equations with Markovian switching, *Stochastic Process. Appl.*, **79** (1999), 45–67.

[18] X. Mao and C. Yuan, *Stochastic Differential Equations with Markovian Switching*, Impreial College Press, London, 2006.

[19] X. Mao, G. Yin, and C. Yuan, Stabilization and destabilization of hybrid systems of stochastic differential equations, *Automatica*, **43** (2007), 264–273.

[20] M. Mariton, *Jump Linear Systems in Automatic Control*, Marcel Dekker, Inc., New York, 1990.

[21] S.P. Sethi and Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhäuser, Boston, 1994.

[22] S.P. Sethi, H. Zhang, and Q. Zhang, *Average-cost Control of Stochastic Manufacturing Systems*, Springer, New York, 2005.

[23] A.V. Skorohod, *Asymptotic Methods in the Theory of Stochastic Differential Equations*, Amer. Math. Soc., Providence, RI, 1989.

[24] Q.S. Song, G. Yin, and Z. Zhang, Numerical method for controlled regime-switching diffusions and regime-switching jump diffusions, *Automatica*, **42** (2006), 1147–1157.

[25] Q.S. Song and G. Yin, Rates of convergence of numerical methods for controlled regime-switching diffusions with stopping times in the costs, *SIAM J. Control Optim.*, **48** (2009), 1831–1857.

[26] D.D. Sworder and J.E. Boyd, *Estimation Problems in Hybrid Systems*, Cambridge University Press, 1999.

[27] G. Yin, X. Mao, C. Yuan, and D. Cao, Approximatio methods for hybrid diffusion systems with state-dependent switching diffusion processes: Numerical alogorithms and existence and uniqueness of solutions, *SIAM J. Math. Anal.*, **41** (2010), 2335–2352.

[28] G. Yin and S. Dey, Weak convergence of hybrid filtering problems involving nearly completely decomposable hidden Markov chains, *SIAM J. Control Optim.*, **41** (2003), 1820–1842.

[29] G. Yin, V. Krishnamurthy, and C. Ion, Regime switching stochastic approximation algorithms with application to adaptive discrete stochastic optimization, *SIAM J. Optim.*, **14** (2004), 1187–1215.

[30] G. Yin, R.H. Liu, and Q. Zhang, Recursive algorithms for stock Liquidation: A stochastic optimization approach, *SIAM J. Optim.*, **13** (2002), 240–263.

[31] G. Yin and Q. Zhang, *Continuous-time Markov Chains and Applications: A Singular Perturbations Approach*, Springer-Verlag, New York, NY, 1998.

[32] G. Yin and C. Zhu, On the notion of weak stability and related issues of hybrid diffusion systems, *Nonlinear Anal.: Hybrid System*, **1** (2007), 173–187.

[33] G. Yin and C. Zhu, Regularity and recurrence of switching diffusions, *J. Syst. Sci. Complexity*, **20** (2007), 273–283.

[34] G. Yin and C. Zhu, *Switching Diffusions: Asymptotic Properties and Applications*, Springer, New York, 2010.

[35] G. Yin and C. Zhu, Properties of solutions of stochastic differential equations with continuous-state-dependent switching, *J. Differential Eqs.*, **249** (2010), 2409–2439.

[36] J. Yong and X.Y. Zhou, *Stochastic Controls: Hamiltonian Systems and HJB Equations*, Springer-Verlag, New York, 1999.

[37] Q. Zhang, Stock trading: An optimal selling rule, *SIAM J. Control Optim.*, **40** (2001), 64–87.

[38] X.Y. Zhou and G. Yin, Markowitz mean-variance portfolio selection with regime switching: A continuous-time model, *SIAM J. Control Optim.*, **42** (2003), 1466–1482.

[39] C. Zhu and G. Yin, Asymptotic properties of hybrid diffusion systems, *SIAM J. Control Optim.*, **46** (2007), 1155–1179.

[40] C. Zhu and G. Yin, On strong Feller, recurrence, and weak stabilization of regime-switching diffusions, *SIAM J. Control Optim.*, **48** (2009), 2003–2031.

[41] C. Zhu and G. Yin, On competitive Lotka-Volterra model in random environments, *J. Math. Anal. Appl.*, **357** (2009), 154–170.

[42] C. Zhu, G. Yin, and Q.S. Song, Stability of random-switching systems of differential equations, *Quarterly Appl. Math.*, **67** (2009), 201–220.