

# Modeling and Analysis of Switching Diffusion Systems: Past-Dependent Switching with a Countable State Space\*

Dang Hai Nguyen<sup>†</sup>      George Yin<sup>‡</sup>

June 19, 2017

## Abstract

Motivated by networked systems in random environment and controlled hybrid stochastic dynamic systems, this work focuses on modeling and analysis of a class of switching diffusions consisting of continuous and discrete components. Novel features of the models include the discrete component taking values in a countably infinite set, and the switching depending on the value of the continuous component involving past history. In this work, the existence and uniqueness of solutions of the associated stochastic differential equations are obtained. In addition, Markov and Feller properties of a function-valued stochastic process associated with the hybrid diffusion are also proved. In particular, when the switching rates depend only on the current state, strong Feller properties are obtained. These properties will pave a way for future study of control design and optimization of such dynamic systems.

**Keywords.** Switching diffusion, past-dependent switching, countable state space, existence and uniqueness of solution, Feller property.

**Mathematics Subject Classification.** 93E03, 60J60, 60H10, 92D25.

**Running Title.** Past-Dependent-Switching Diffusion Systems

---

\*This research was supported in part by the National Science Foundation under grant DMS-1207667.

<sup>†</sup>Department of Mathematics, Wayne State University, Detroit, MI 48202, dangnh.maths@gmail.com.

<sup>‡</sup>Department of Mathematics, Wayne State University, Detroit, MI 48202, gyin@math.wayne.edu.

# 1 Introduction

Owing to the demand of modeling, analysis, and computation of complex networked systems, much attention has been devoted to building more realistic dynamic system models. It has been well recognized that in many real-world applications, traditional models using continuous processes represented by solutions to deterministic differential equations and stochastic differential equations alone are often inadequate. Arising from control engineering, queueing networks, manufacturing and production planning, parameter estimation, filtering of dynamic systems, ecological and biological systems, and financial engineering, etc., numerous complex systems contain both continuous dynamics and discrete events. The discrete events in these systems are not normally representable by solutions of the usual differential equations. Because of the demand, switching diffusions (also known as hybrid switching diffusions) have drawn growing and resurgent attention. A switching diffusion is a two-component process  $(X(t), \alpha(t))$  in which the continuous component  $X(t)$  evolves according to the diffusion process whose drift and diffusion coefficients depend on the state of  $\alpha(t)$ , whereas  $\alpha(t)$  takes values in a set consisting of isolated points. Because of their importance, many papers have been devoted to such hybrid dynamic systems; see [10, 20, 26, 27] and the references therein. In their comprehensive treatment of hybrid switching diffusions, Mao and Yuan [14] focused on  $\alpha(t)$  being a continuous-time and homogeneous Markov chain independent of the Brownian motion and the generator of the Markov chain being a constant matrix. Realizing the need, treating the two components jointly, Yin and Zhu [25] extended the study to the Markov process  $(X(t), \alpha(t))$  by allowing the generator  $\alpha(t)$  to depend on the current state  $X(t)$ . Properties of the underlying process including recurrence, positive recurrence, ergodicity, Feller properties, stability, and invariance among others were investigated. Such study provides us with a clear picture of the underlying processes. Nevertheless, in both of the aforementioned books and most related papers to date, the switching process  $\alpha(t)$  is assumed to have a finite state space. One question naturally arises. What happens if the switching process has a countable state space? Much of the argument in [25] relies on the interplay of stochastic processes and the associated systems of partial differential equations. Because the state space of  $\alpha(t)$  was assumed to be a finite set, one can essentially treat a system of partial differential equations with a finite number of equations. When we consider problems involving a countable state space, the number of equations becomes infinite. Much more complex situation is encountered. Different methods have to be developed to treat the systems.

There are plenty of real-world applications involving such switching diffusions. Perhaps, one of the most widely used control models in the literature is the so-called LQG (linear quadratic Gaussian regulator) problem; see [2, pp.165-166] for a traditional model. However, for many new applications in networked systems, it has been found that in addition to the random noise represented by Brownian type of disturbances, there is a source of randomness owing to the presence of random environment that can be modeled by a continuous-time Markov chain. Let  $\alpha(t)$  be a continuous-time Markov chain with state space  $\mathbb{Z}_+$  (the set of positive integers) and generator  $Q$ . Consider the controlled dynamic system

$$\begin{aligned} dX(t) &= [A(\alpha(t))X(t) + B(\alpha(t))u(t)]dt + \sigma(\alpha(t))dW(t), \\ X(s) &= x, \text{ for } s \leq t \leq T, \end{aligned} \tag{1.1}$$

where  $X(t) \in \mathbb{R}^{n_1}$  is the continuous state variable,  $u(t) \in \mathbb{R}^{n_2}$  is the control,  $A(i) \in \mathbb{R}^{n_1 \times n_1}$  and  $B(i) \in \mathbb{R}^{n_1 \times n_2}$  are well defined and have finite values for each  $i \in \mathbb{Z}_+$ . One may wish to find the

optimal control  $u(\cdot)$  so that the expected quadratic cost function

$$J(s, i, x, u(\cdot)) = E \left[ \int_s^T [X^\top(t)M(\alpha(t))X(t) + u^\top(t)N(\alpha(t))u(t)]dt + X^\top(T)DX(T) \right] \quad (1.2)$$

is minimized. The use of  $\alpha(t)$  stems from the formulation of discrete events, and the use of  $\mathbb{Z}_+$  enlarges the applicability of previous consideration of finite state space cases. Switched dynamic systems can also be found in, for example, modeling impatient customers and customer abandonment of Markov-modulated service speeds in the heavy-traffic regime and the many-server systems in the Halfin-Whitt regime and the non-degenerate slowdown regime; see [4]. We also refer the reader to Whitt [24] for further reading on limit results in queueing theory and many references therein. In fact, in most of the queueing models, the discrete set is countable rather than finite.

Two more dynamic systems are to be presented in the next section, in which the main interests are to find long-term behavior and control design in an ecological system and to find optimal strategies under long-run average criteria for a pollution management problem. In order to study the aforesaid problems, we first need to ensure that the systems under consideration have unique solutions and that the solutions possess good properties. Motivated by these examples, we take up the challenge of considering a nonlinear hybrid diffusion  $(X(t), \alpha(t))$  whose discrete component  $\alpha(t)$  has an infinite state space in this paper. Moreover, in lieu of allowing the switching process to depend on the current state  $X(t)$  only, we assume that it is past dependent. That is, we assume that the generator of  $\alpha(t)$  depends on the past history of the continuous process. This paper provides conditions for the existence and uniqueness of the solutions for given initial data, and to demonstrate the Markov-Feller property of a function-valued stochastic process associated with the equation. Our study will build a bridge for future study on related control systems.

The rest of the paper is organized as follows. The formulation of hybrid switching diffusions with past-dependent switching and countably many possible switching locations is given in Section 2. The existence and uniqueness of solutions to the stochastic equations are then proved under suitable conditions in Section 3. Section 4 studies the Markov and Feller properties of a function-valued stochastic process associated with our equation. The proof for the Feller property is rather complex because the state space of  $\alpha(t)$  is infinite, the space of continuous functions is not locally compact, and we do not assume uniform continuity of the switching intensities. In section 5, the strong Feller property of the hybrid diffusion without past-dependent switching is given. Section 6 provides further remarks and points out future research directions. Finally, we provide the proofs of some technical results in an appendix.

## 2 Formulation

Let  $r$  be a fixed positive number. Denote by  $\mathcal{C}([a, b], \mathbb{R}^n)$  the set of  $\mathbb{R}^n$ -valued continuous functions defined on  $[a, b]$ . In what follows, we mainly work with  $\mathcal{C}([-r, 0], \mathbb{R}^n)$ , and simply denote it by  $\mathcal{C} := \mathcal{C}([-r, 0], \mathbb{R}^n)$ . Denote by  $|x|$  the Euclidean norm of  $x \in \mathbb{R}^n$ . For  $\phi \in \mathcal{C}$ , we use the norm  $\|\phi\| = \sup\{|\phi(t)| : t \in [-r, 0]\}$ . For  $y(\cdot) \in \mathcal{C}([-r, \infty), \mathbb{R}^n)$  and  $t \geq 0$ , we denote by  $y_t$  the so-called segment function (or memory segment function)  $y_t(\cdot) := y(t + \cdot) \in \mathcal{C}$ . Let  $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$  be a complete filtered probability space with the filtration  $\{\mathcal{F}_t\}_{t \geq 0}$  satisfying the usual condition, i.e., it is increasing and right continuous while  $\mathcal{F}_0$  contains all  $\mathbb{P}$ -null sets. Let  $W(t)$  be an  $\mathcal{F}_t$ -adapted and  $\mathbb{R}^d$ -valued Brownian motion. Suppose  $b(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{Z}_+ \rightarrow \mathbb{R}^n$  and  $\sigma(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{Z}_+ \rightarrow \mathbb{R}^{n \times d}$ ,

where  $\mathbb{Z}_+ = \mathbb{N} \setminus \{0\} = \{1, 2, \dots\}$ , the set of positive integers. Consider the two-component process  $(X(t), \alpha(t))$ , where  $\alpha(t)$  is a pure jump process taking value in  $\mathbb{Z}_+$ , and  $X(t)$  satisfies

$$dX(t) = b(X(t), \alpha(t))dt + \sigma(X(t), \alpha(t))dW(t). \quad (2.1)$$

We assume that if  $\alpha(t-) := \lim_{s \rightarrow t-} \alpha(s) = i$ , then it can switch to  $j$  at  $t$  with intensity  $q_{ij}(X_t)$  where  $q_{ij}(\cdot) : \mathcal{C} \rightarrow \mathbb{R}$ . When  $q_i(\phi) := \sum_{j=1, j \neq i}^{\infty} q_{ij}(\phi)$  is uniformly bounded in  $(\phi, i) \in \mathcal{C} \times \mathbb{Z}_+$ , and  $q_i(\cdot)$  and  $q_{ij}(\cdot)$  are continuous, one may view the aforementioned assumption as

$$\begin{aligned} \mathbb{P}\{\alpha(t + \Delta) = j | \alpha(t) = i, X_s, \alpha(s), s \leq t\} &= q_{ij}(X_t)\Delta + o(\Delta) \text{ if } i \neq j \text{ and} \\ \mathbb{P}\{\alpha(t + \Delta) = i | \alpha(t) = i, X_s, \alpha(s), s \leq t\} &= 1 - q_i(X_t)\Delta + o(\Delta). \end{aligned} \quad (2.2)$$

However, when  $q_i(\phi)$  and  $q_{ij}(\phi)$  are either discontinuous or unbounded, it does not seem appropriate to use (2.2) to model the switching intensity. To formulate the problem in a general setting without the boundedness and continuity assumptions mentioned above, we construct  $\alpha(t)$  as the solution to a stochastic differential equation with respect to a Poisson random measure. We elaborate on the idea below. Let  $\mathbf{p}(dt, dz)$  be a Poisson random measure with intensity  $dt \times \mathbf{m}(dz)$  and  $\mathbf{m}$  is the Lebesgue measure on  $\mathbb{R}$  such that  $\mathbf{p}(\cdot, \cdot)$  is independent of the Brownian motion  $W(\cdot)$ . Let  $\tilde{\mathbf{p}}$  be the Poisson point process associated with  $\mathbf{p}(\cdot, \cdot)$  (see e.g., [22]). Then  $\tilde{\mathbf{p}}$  can lie in a set  $A$  with intensity  $\mathbf{m}(A)$ , that is, the expected number of Poisson points lying in  $A$  during the period  $dt$  is  $dt \times \mathbf{m}(A)$ . Using this fact, for each  $i \in \mathbb{Z}$ , we can construct disjoint sets  $\{\Delta_{ij}(\phi), j \neq i\}$  such that  $\mathbf{m}(\Delta_{ij}(\phi)) = q_{ij}(\phi)$ . Let  $\tilde{\mathbf{p}}$  govern the switching of  $\alpha(t)$  in the manner that if  $\alpha(t-) = i$  and there is a Poisson point in  $\Delta_{ij}(X_t)$  at time  $t$ , then  $\alpha(t) = j$ . If  $\alpha(t-) = i$  and there is no Poisson point in  $\cup_{j \neq i} \Delta_{ij}(X_t)$  at time  $t$ ,  $\alpha(t)$  remains  $i$ . Using this idea, we formulate the equation for  $\alpha(t)$  as follows. For each function  $\phi : [-r, 0] \rightarrow \mathbb{R}^n$ , and  $i \in \mathbb{Z}_+$ , let  $\Delta_{ij}(\phi), j \neq i$  be the consecutive left-closed and right-open intervals of the real line, each having length  $q_{ij}(\phi)$ . That is,

$$\Delta_{i1}(\phi) = [0, q_{i1}(\phi)), \quad \Delta_{ij}(\phi) = \left[ \sum_{k=1, k \neq i}^{j-1} q_{ik}(\phi), \sum_{k=1, k \neq i}^j q_{ik}(\phi) \right), j > 1, j \neq i.$$

Define  $h : \mathcal{C} \times \mathbb{Z}_+ \times \mathbb{R} \mapsto \mathbb{R}$  by  $h(\phi, i, z) = \sum_{j=1, j \neq i}^{\infty} (j - i) \mathbf{1}_{\{z \in \Delta_{ij}(\phi)\}}$ , where  $\mathbf{1}_{\{z \in \Delta_{ij}(\phi)\}} = 1$  if  $z \in \Delta_{ij}$ , otherwise  $\mathbf{1}_{\{z \in \Delta_{ij}(\phi)\}} = 0$ , is the indicator function. The process  $\alpha(t)$  can be defined as a solution to

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

The pair  $(X(t), \alpha(t))$  is therefore a solution to the system of equations

$$\begin{cases} dX(t) = b(X(t), \alpha(t))dt + \sigma(X(t), \alpha(t))dW(t) \\ d\alpha(t) = \int_{\mathbb{R}} h(X_t, \alpha(t-), z) \mathbf{p}(dt, dz). \end{cases} \quad (2.3)$$

A strong solution to (2.3) on  $[0, T]$  with initial data  $(\phi, i_0)$  being  $\mathcal{C} \times \mathbb{Z}_+$ -valued and  $\mathcal{F}_0$ -measurable random variable, is an  $\mathcal{F}_t$ -adapted process  $(X(t), \alpha(t))$  such that

- $X(t)$  is continuous and  $\alpha(t)$  is cadlag (right continuous with left limits) almost surely (a.s.).
- $X(t) = \phi(t)$  for  $t \in [-r, 0]$  and  $\alpha(0) = i_0$

- $(X(t), \alpha(t))$  satisfies (2.3) for all  $t \in [0, T]$  a.s.

We will show in the Appendix that the solution  $(X(t), \alpha(t))$  to (2.3), satisfies (2.2) under suitable conditions. Let  $f(\cdot, \cdot) : \mathbb{R}^n \times \mathbb{Z}_+ \mapsto \mathbb{R}$  be twice continuously differentiable in  $x$  and bounded in  $(x, i) \in \mathbb{R}^n \times \mathbb{Z}_+$ . We define the “operator”  $\mathcal{L}f(\cdot, \cdot) : \mathcal{C} \times \mathbb{Z}_+ \mapsto \mathbb{R}$  by

$$\begin{aligned}
\mathcal{L}f(\phi, i) &= \nabla f(\phi(0), i) b(\phi(0), i) + \frac{1}{2} \text{tr} \left( \nabla^2 f(\phi(0), i) A(\phi(0), i) \right) \\
&\quad + \sum_{j=1, j \neq i}^{\infty} q_{ij}(\phi) [f(\phi(0), j) - f(\phi(0), i)] \\
&= \sum_{k=1}^n b_k(\phi(0), i) f_k(\phi(0), i) + \frac{1}{2} \sum_{k, l=1}^n a_{kl}(\phi(0), i) f_{kl}(\phi(0), i) \\
&\quad + \sum_{j=1, j \neq i}^{\infty} q_{ij}(\phi) [f(\phi(0), j) - f(\phi(0), i)],
\end{aligned} \tag{2.4}$$

where  $b(x, i) = (b_1(x, i), \dots, b_n(x, i))^\top$ ,  $\nabla f(x, i) = (f_1(x, i), \dots, f_n(x, i)) \in \mathbb{R}^{1 \times n}$  and  $\nabla^2 f(x, i) = (f_{ij}(x, i))_{n \times n}$  are the gradient and Hessian of  $f(x, i)$  with respect to  $x$ , respectively, with

$$\begin{aligned}
f_k(x, i) &= (\partial / \partial x_k) f(x, i), \quad f_{kl}(x, i) = (\partial^2 / \partial x_k \partial x_l) f(x, i), \quad \text{and} \\
A(x, i) &= (a_{kl}(x, i))_{n \times n} = \sigma(x, i) \sigma^\top(x, i),
\end{aligned}$$

with  $z^\top$  denoting the transpose of  $z$ . Suppose that  $(X(t), \alpha(t))$  satisfies (2.3) and that for any  $T > 0$ ,

$$\sup_{t \in [0, T]} \{q_{\alpha(t)}(X_t)\} < \infty \text{ a.s.} \tag{2.5}$$

Let  $\xi_k = \inf\{t > 0 : q_{\alpha(t)}(X_t) \geq k\}$ ,  $k \in \mathbb{Z}_+$ . By noting that  $h(X_t, z) = 0$  if  $z \notin [0, q_{\alpha(t)}(X_t))$  and that

$$\int_{\mathbb{R}} [f(\phi(0), i + h(\phi, i, z)) - f(\phi(0), i)] \mathbf{m}(dz) = \sum_{j=1, j \neq i}^{\infty} q_{ij}(\phi) [f(\phi(0), j) - f(\phi(0), i)],$$

we have from Itô's formula (see [1, Theorem 4.4.7]) that

$$\begin{aligned}
&f(X(t \wedge \xi_k), \alpha(t \wedge \xi_k)) - f(X(0), \alpha(0)) \\
&= \int_0^{t \wedge \xi_k} \mathcal{L}f(X_s, \alpha(s-)) ds + \int_0^{t \wedge \xi_k} \nabla f(X(s), \alpha(s-)) \sigma(X(s), \alpha(s-)) dW(s) \\
&\quad + \int_0^{t \wedge \xi_k} \int_0^k [f(X(s), \alpha(s-) + h(X_s, \alpha(s-), z)) - f(X(s), \alpha(s-))] \mu(ds, dz) \\
&= \int_0^{t \wedge \xi_k} \mathcal{L}f(X_s, \alpha(s-)) ds + \int_0^{t \wedge \xi_k} \nabla f(X(s), \alpha(s-)) \sigma(X(s), \alpha(s-)) dW(s) \\
&\quad + \int_0^{t \wedge \xi_k} \int_{\mathbb{R}} [f(X(s), \alpha(s-) + h(X_s, \alpha(s-), z)) - f(X(s), \alpha(s-))] \mu(ds, dz),
\end{aligned}$$

where  $\mu(ds, dz)$  is the compensated Poisson random measure given by

$$\mu(ds, dz) = \mathbf{p}(ds, dz) - \mathbf{m}(dz) ds.$$

Under condition (2.5), there exists a random integer  $k_0 = k_0(\omega)$  such that  $t \wedge \xi_k = t$  for any  $k > k_0$ . As a result,

$$f(X(t), \alpha(t)) - f(X(0), \alpha(0)) = \int_0^t \mathcal{L}f(X_s, \alpha(s-))ds + M_1(t) + M_2(t) \text{ a.s.}, \quad (2.6)$$

where  $M_1(\cdot)$  and  $M_2(\cdot)$  are local martingales, defined by

$$\begin{aligned} M_1(t) &= \int_0^t \nabla f(X(s), \alpha(s-))\sigma(X(s), \alpha(s-))dW(s), \\ M_2(t) &= \int_0^t \int_{\mathbb{R}} [f(X(s), \alpha(s-) + h(X_s, \alpha(s-), z)) - f(X(s), \alpha(s-))] \mu(ds, dz). \end{aligned} \quad (2.7)$$

It should be noted that  $\mathcal{L}$  is not the generator of the Markov process  $(X_t, \alpha(t))$ . However this operator is very useful for analyzing the process  $(X(t), \alpha(t))$ . In view of (2.7), if  $\tau_1 \leq \tau_2$  are stopping times that are bounded above by  $T$  a.s., and  $f(\cdot, \cdot)$  and  $\mathcal{L}f(\cdot, \cdot)$  are bounded and (2.5) holds, then

$$\mathbb{E}f(X(\tau_2), \alpha(\tau_2)) = \mathbb{E}f(X(\tau_1), \alpha(\tau_1)) + \mathbb{E} \int_{\tau_1}^{\tau_2} \mathcal{L}f(X_t, \alpha(t-))dt.$$

**Remark 2.1.** If  $\alpha(t)$  depends on the continuous state, but there is no past dependence (that is,  $X_t$  is replaced by  $X(t)$  in (2.2), and  $\phi$  and  $\phi(0)$  are replaced by the current state  $X(t) = x$  in (2.4), respectively), then  $\mathcal{L}$  is indeed the generator of the process  $(X(t), \alpha(t))$ . Even in this case, the current paper settles the matter of the state space of the switching process being countable thus generalizes the study of finite state space cases as considered in [25].

**Example 2.2.** This example stems from applications in ecological systems and biological control. Consider the evolution of two interacting species. One is micro, which is described by a logistic differential equation perturbed by a white noise. The other is macro, we assume that its number of individuals follows a birth-death process. Let  $X(t)$  be the density of the micro species and  $\alpha(t)$  the population of the macro species. The life cycle of a micro species is usually very short, so it is reasonable to assume that the evolution of  $X(t)$  can be described by the following past-independent equation

$$dX(t) = X(t)[a(\alpha(t)) - b(\alpha(t))X(t)]dt + \sigma(\alpha(t))X(t)dW(t), \quad (2.8)$$

where  $a(i), b(i), \sigma(i)$  are positive constants for each  $i \in \mathbb{Z}_+$ .

On the other hand, the reproduction process of  $\alpha(t)$  is assumed to be non-instantaneous. More precisely, suppose the reproduction depends on the period of time from egg formation to hatching, say  $r$ . Then we have

$$d\alpha(t) = \int_{\mathbb{R}} h(X_t, \alpha(t-), z)\mathbf{p}(dt, dz), \quad (2.9)$$

where  $h(\phi, i, z) = \sum_{j=1, j \neq i}^{\infty} (j-i)\mathbf{1}_{\{z \in \Delta_{ij}(\phi)\}}$ ,  $\Delta_{i,i+1}(\phi) = [0, \beta_i(\phi))$ ,  $\Delta_{i,i-1}(\phi) = [0, \delta_i(\phi))$ ,  $\Delta_{i,j}(\phi) = \emptyset$  if  $j \notin \{i-1, i, i+1\}$  or  $i = 0$ . Usually  $\beta_i(\phi), \delta_i(\phi)$  can be given in the integral form  $\beta_i(\phi) = \int_{-r}^0 \tilde{\beta}_i(t)\phi(t)dt$ ,  $\delta_i(\phi) = \int_{-r}^0 \tilde{\delta}_i(t)\phi(t)dt$ , for some appropriate weighting functions  $\tilde{\beta}_i, \tilde{\delta}_i$ . As can be seen from the above, the switching process at  $t$  in fact depends on past history of the state  $X(\cdot)$ . Investigating the interactions between the two species are very important to biological control. A basic biological control problem aims to choose a suitable living organism to control a

particular pest (see e.g., [5, 12]). This chosen organism might be a predator, parasite, or disease, which will attack the harmful insect. To design and evaluate effectiveness of a biological control, some questions should be answered first. For example, under which conditions the species will be permanent forever or they will extinct at some instance? Whether or not there is an invariant measure associated with the system under consideration. Mathematically, these questions are related to the stability and ergodicity of the corresponding stochastic systems, which will be studied in a future paper.

**Example 2.3.** Pollution management is vitally important and has a significant impact on environment. A major issue is concerned with the tradeoff of pollution accumulation and consumption, which affects environmental policy making. Following the seminal paper of Keeler et al. [9], much work has been devoted to the study of optimal control of dynamic economic systems. In [8], Kawaguchi and Morimoto treated a pollution accumulation problem of maximizing the long-run average welfare using a controlled diffusion model. Assume that an economy consumes some good and meanwhile generates pollution. The pollution stock is gradually degraded and its instantaneous growth rate incorporates a random disturbance with mean zero and constant variance. The social welfare is defined by the utility of the consumption net of the disutility of pollution. The problem is to find optimal consumption strategies for the society in the long-run average sense. Departing from their formulation, we consider an extension of their model. Suppose that there is a switching process  $\alpha(t)$  taking values in  $\mathbb{Z}_+$  such that  $\alpha(t)$  represents the level of pollution at time  $t$ . Assume that the stock of pollution at time  $t$  is given by  $X(t)$ , a real-valued process, and there is a positive real-valued function  $\rho(\cdot)$  so that for each  $i \in \mathbb{Z}_+$ , the rate of pollution decay is  $\rho(i)$ . The consumption rate (or flow of pollution) is a control process, which is denoted by  $c(t)$  at time  $t$ ; the social utility function of the consumption  $c$  is denoted by  $U(c)$ , whereas the social disutility of the pollution stock  $x$  is  $D(x)$ . We say that the consumption rate is admissible if it is  $\mathcal{F}_t$ -measurable, where  $\mathcal{F}_t = \{(X(s), \alpha(s)) : s \leq t\}$  such that  $0 \leq c(t) \leq K_0$  for some  $K_0 > 0$ . The ultimate objective is to maximize the long-run average welfare

$$J(c(\cdot)) = \liminf_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \int_0^T [U(c(t)) - D(X(t))] dt, \quad (2.10)$$

subject to

$$dX(t) = [c(t) - \rho(\alpha(t))X(t)]dt + \sigma(X(t), \alpha(t))dW(t). \quad (2.11)$$

Assume that the pollution level  $\alpha(\cdot)$  satisfies the conditions (2.2). First, it is reasonable that the level of pollution can be modeled by a continuous-time process taking values in  $\mathbb{Z}_+$ . Second, to be more realistic, the pollution level depends on the pollution stock  $X(t)$  as well as some past history as given in (2.2). As another generalization of [8], we assume that  $\sigma$  in fact depends on  $(X(t), \alpha(t))$ , and the switching rate depends on some past history of the pollution stock  $X(\cdot)$  as in (2.2), and  $\sigma^2(x, i) > 0$  for each  $i \in \mathbb{Z}_+$ . Treating the optimal pollution management problem, it is natural to consider the replacement of the average in (2.10) by the average with respect to an invariant measure (if it exists) of the controlled systems. To do so, we need to make sure that (2.11) indeed has an invariant measure. Before this matter can be settled, we need to show that the system has a unique solution for each initial data, and the solution possesses certain desired properties such as Markov and Feller properties.

### 3 Existence and Uniqueness of Solutions

We are now in a position to prove the existence and uniqueness in the strong sense of a solution with given initial data under suitable conditions. We give several sets of conditions. The main reason is due to the past dependence and the use of  $\mathbb{Z}_+$ . First in contrast to the case of switching process staying in a finite set, care needs to be exercised regarding uniformity with respect to the switching set. Second, the past dependence requires careful handling of the use of Lipschitz continuity etc. and the uniformity with respect to the element in the corresponding function spaces. Depending on the preference, Assumptions 3.1 allows certain bounds to be dependent of the switching state  $i$ , but uniform in the variable in the function space, whereas Assumption 3.2 requires uniformity in the bounds w.r.t.  $i$ , but requires the past dependent part be localized. Assumptions 3.3 and 3.4 relax the Lipschitz condition to local Lipschitz together with certain growth conditions presented by using bounds with the help of Lyapunov functions.

**Assumption 3.1.** Assume the following conditions hold.

- (i) For each  $i \in \mathbb{Z}_+$ , there is a positive constant  $L_i$  such that

$$|b(x, i) - b(y, i)| + |\sigma(x, i) - \sigma(y, i)| \leq L_i |x - y| \forall x, y \in \mathbb{R}^n.$$

- (ii)  $q_{ij}(\phi)$  is measurable in  $\phi \in \mathcal{C}$  for all  $i$  and  $j \in \mathbb{Z}_+$ . Moreover,

$$M := \sup_{\phi \in \mathcal{C}, i \in \mathbb{Z}_+} \{q_i(\phi)\} < \infty.$$

**Theorem 3.1.** Under Assumption 3.1, for each initial data  $(\xi, i_0)$ , there exists a unique solution  $(X(t), \alpha(t))$  to (2.3).

*Proof.* It is well-known that part (i) of Assumption 3.1 guarantees the existence and uniqueness of strong solutions to the following diffusion

$$dY(t) = b(Y(t), i)dt + \sigma(Y(t), i)dW(t) \quad \text{for each } i \in \mathbb{Z}_+. \quad (3.1)$$

Then, given a stopping time  $\tau$  and an  $\mathcal{F}_\tau$ -measurable  $\mathbb{R}^n$ -valued random variable  $y = y(\tau)$  (depending on  $\tau$ ), there exists a unique strong solution to (3.1) in  $[\tau, \infty)$  satisfying  $Y(\tau) = y(\tau)$  (see [14, Remark 3.10]). We can now construct the solution to (2.3) with initial data  $(\xi, i_0)$  by the interlacing procedure similar to [1, Chapter 5]. Let  $\tilde{Y}^{(0)}(t), t \geq 0$  be the solution with initial data  $\xi(0)$  to

$$d\tilde{Y}^{(0)}(t) = b(\tilde{Y}^{(0)}(t), i_0)dt + \sigma(\tilde{Y}^{(0)}(t), i_0)dW(t).$$

We also set  $\tilde{Y}^{(0)}(t) = \xi(t)$  for  $t \in [-\tau, 0]$ . Let

$$\begin{aligned} \tau_1 &= \inf\{t > 0 : \int_0^t \int_{\mathbb{R}} h(\tilde{Y}_s^{(0)}, i_0, z) \mathbf{p}(ds, dz) \neq 0\} \quad \text{and} \\ i_1 &= i_0 + \int_0^{\tau_1} \int_{\mathbb{R}} h(\tilde{Y}_s^{(0)}, i_0, z) \mathbf{p}(ds, dz), \end{aligned}$$

and  $\tilde{Y}^{(1)}(t), t \geq \tau_1$  be the solution with  $\tilde{Y}_{\tau_1}^{(1)} = \tilde{Y}_{\tau_1}^{(0)}$  to

$$d\tilde{Y}^{(1)}(t) = b(\tilde{Y}^{(1)}(t), i_1)dt + \sigma(\tilde{Y}^{(1)}(t), i_1)dW(t). \quad (3.2)$$



Define

$$\begin{aligned}\tau_2 &= \inf\{t > \tau_1 : \int_{\tau_1}^t \int_{\mathbb{R}} h(\tilde{Y}_s^{(1)}, i_1, z) \mathbf{p}(ds, dz) \neq 0\} \quad \text{and} \\ i_2 &= i_1 + \int_{\tau_1}^{\tau_2} \int_{\mathbb{R}} h(\tilde{Y}_s^{(1)}, i_1, z) \mathbf{p}(ds, dz).\end{aligned}$$

Note that, in the notation above,  $\tilde{Y}_t^{(k)}$  is the function  $s \in [-r, 0] \mapsto \tilde{Y}^{(k)}(t+s)$ . Continuing this procedure, let  $\tau_\infty = \lim_{k \rightarrow \infty} \tau_k$  and set

$$X(t) = \tilde{Y}^{(k)}(t), \quad \alpha(t) = i_k \text{ if } \tau_k \leq t < \tau_{k+1}. \quad (3.3)$$

Clearly,  $X(t)$  satisfies that for every  $t \geq 0$ ,

$$\begin{cases} X(t \wedge \tau_k) = X(0) + \int_0^{t \wedge \tau_k} [b(X(s), \alpha(s)) ds + \sigma(X(s), \alpha(s)) dW(t)] \\ \alpha(t \wedge \tau_k) = i_0 + \int_0^{t \wedge \tau_k} \int_{\mathbb{R}} h(X_s, \alpha(s-), z) \mathbf{p}(ds, dz). \end{cases} \quad (3.4)$$

To show that  $X(t)$  is a global solution, we need only prove that  $\tau_\infty = \infty$  a.s. For any  $T > 0$ ,

$$\begin{aligned}\mathbb{P}\{\tau_k \leq T\} &= \mathbb{P}\left\{ \int_0^{T \wedge \tau_k} \int_{\mathbb{R}} \mathbf{1}_{\{z \in [0, q_{\alpha(s-)}(X_s)]\}} \mathbf{p}(ds, dz) = k \right\} \\ &\leq \mathbb{P}\left\{ \int_0^{T \wedge \tau_k} \int_{\mathbb{R}} \mathbf{1}_{\{z \in [0, M]\}} \mathbf{p}(ds, dz) \geq k \right\} \\ &\leq \mathbb{P}\left\{ \int_0^T \int_{\mathbb{R}} \mathbf{1}_{\{z \in [0, M]\}} \mathbf{p}(ds, dz) \geq k \right\} \\ &= \sum_{l=k}^{\infty} e^{-MT} \frac{(MT)^l}{l!}.\end{aligned} \quad (3.5)$$

It follows that  $\mathbb{P}\{\tau_k \leq T\} \rightarrow 0$  as  $k \rightarrow \infty$ . As a result  $\tau_\infty = \infty$  a.s. By this construction, it can be seen that  $X(t)$  is continuous and  $\alpha(t)$  is cadlag almost surely. The uniqueness of  $(X(t), \alpha(t))$  follows from the uniqueness of  $\tilde{Y}^{(k)}(t)$  on  $[\tau_k, \tau_{k+1}]$  and the uniqueness of  $i_k$  defined by

$$i_k = i_{k-1} + \int_{\tau_{k-1}}^{\tau_k} \int_{\mathbb{R}} h(\tilde{Y}_t^{(k-1)}, i_{k-1}, z) \mathbf{p}(dt, dz).$$

This concludes the proof. □

**Assumption 3.2.** Assume the following conditions hold.

(i) There is a positive constant  $L$  such that

$$|b(x, i) - b(y, i)| + |\sigma(x, i) - \sigma(y, i)| \leq L|x - y|, \quad \forall x, y \in \mathbb{R}^n, i \in \mathbb{Z}_+.$$

(ii)  $q_{ij}(\phi)$  is measurable in  $\phi \in \mathcal{C}$  for each  $(i, j) \in \mathbb{Z}_+^2$ . Moreover, for any  $H > 0$ ,

$$M_H := \sup_{\phi \in \mathcal{C}, \|\phi\| \leq H, i \in \mathbb{Z}_+} \{q_i(\phi)\} < \infty.$$

**Remark 3.2.** We can use either Assumption 3.1 or Assumption 3.2 to obtain the existence and uniqueness of solutions to (2.3). Recall that now  $\mathbb{Z}_+$  is a countable set, so care must be taken to distinct it with a finite state case. In Assumption 3.1, the Lipschitz constants of  $b(\cdot, i), \sigma(\cdot, i)$  depend on  $i$ , and  $q_i(\phi)$  is assumed to be bounded uniformly in  $(\phi, i) \in \mathcal{C} \times \mathbb{Z}_+$ . In contrast, the uniform boundedness of  $q_i(\phi)$  is relaxed, but the Lipschitz constant of  $b(\cdot, i), \sigma(\cdot, i)$  is assumed to be in  $i \in \mathbb{Z}_+$ .

**Theorem 3.3.** *Under Assumption 3.2, for each initial data  $(\xi, i_0)$ , there exists a unique solution  $(X(t), \alpha(t))$  to (2.3).*

*Proof.* Without loss of generality, we may assume that  $(\xi, i_0)$  is bounded, since we can use the truncation method in [3, Theorem 3 in §6] to obtain the result for general  $(\xi, i_0)$  once we have proved for the case  $(\xi, i_0)$  being bounded. Construct the process  $(X(t), \alpha(t))$  as in the proof of Theorem 3.1. We need to show that  $\tau_\infty = \infty$  a.s. Following the proof of [13, Lemma 3.2, p. 51], there is a  $K = K(T)$  such that

$$\mathbb{E} \left( \sup_{0 \leq t \leq T \wedge \tau_k} |X(t)|^2 \right) \leq K \forall k \in \mathbb{Z}_+.$$

As a result, for any  $\varepsilon > 0$ , there is an  $H_\varepsilon$  such that

$$\mathbb{P}\{\|X_t\| \leq H_\varepsilon \forall t \in [0, T \wedge \tau_k]\} > 1 - \frac{\varepsilon}{2}. \quad (3.6)$$

Let  $\eta_{H_\varepsilon} = \inf\{t \geq 0 : \|X_t\| \geq H_\varepsilon\}$  and  $M_{H_\varepsilon} = \sup_{\phi \in \mathcal{C}, \|\phi\| \leq H_\varepsilon, i \in \mathbb{Z}_+} \{q_i(\phi)\} < \infty$ . Then

$$\begin{aligned} \mathbb{P}\{\tau_k \leq T \wedge \eta_{H_\varepsilon}\} &= \mathbb{P}\left\{ \int_0^{T \wedge \tau_k \wedge \eta_{H_\varepsilon}} \int_{\mathbb{R}} \mathbf{1}_{\{z \in [0, q_{\alpha(s)}(X_s)]\}} \mathbf{p}(ds, dz) = k \right\} \\ &\leq \mathbb{P}\left\{ \int_0^{T \wedge \tau_k \wedge \eta_{H_\varepsilon}} \int_{\mathbb{R}} \mathbf{1}_{\{z \in [0, M_{H_\varepsilon}]\}} \mathbf{p}(ds, dz) \geq k \right\} \\ &\leq \mathbb{P}\left\{ \int_0^T \int_{\mathbb{R}} \mathbf{1}_{\{z \in [0, M_{H_\varepsilon}]\}} \mathbf{p}(ds, dz) \geq k \right\} \\ &= e^{-M_{H_\varepsilon} T} \sum_{l=k}^{\infty} \frac{(M_{H_\varepsilon} T)^l}{l!}. \end{aligned} \quad (3.7)$$

For sufficiently large  $k$ , we have

$$\mathbb{P}\{\tau_k \leq T \wedge \eta_{H_\varepsilon}\} \leq e^{-M_{H_\varepsilon} T} \sum_{l=k}^{\infty} \frac{(M_{H_\varepsilon} T)^l}{l!} \leq \frac{\varepsilon}{2}. \quad (3.8)$$

From (3.6) and (3.8),  $\mathbb{P}\{\tau_k \geq T\} \geq \mathbb{P}(\{\tau_k \wedge T < \eta_{H_\varepsilon}\} \cap \{\tau_k > T \wedge \eta_{H_\varepsilon}\}) \geq 1 - \varepsilon$  for sufficiently large  $k$ . Thus, we obtain that  $\mathbb{P}\{\tau_\infty \geq T\} \geq 1 - \varepsilon$ . It holds for every  $T > 0$  and  $\varepsilon > 0$ , so we obtain the desired result.  $\square$

**Remark 3.4.** To obtain the existence and uniqueness of solutions, Assumptions 3.1 and 3.2 can be relaxed by replacing the global Lipschitz conditions with local Lipschitz conditions together with Lyapunov-type functions. To be specific, let  $V(\cdot) : \mathbb{R}^n \mapsto \mathbb{R}$  be twice continuously differentiable in  $x$ . For each  $i \in \mathbb{Z}_+$ , let  $\mathcal{L}_i V(x) = \nabla V(x) b(x, i) + \frac{1}{2} \text{tr} \left( \nabla^2 V(x) A(x, i) \right)$ . For instance (1) of Assumption 3.1 and (1) of Assumption 3.2 can be replaced by the following Assumptions 3.3 and 3.4, respectively.

**Assumption 3.3.** Assume the following conditions hold.

(i) For each  $H > 0$ ,  $i \in \mathbb{Z}_+$ , there is a positive constant  $L_{H,i}$  such that

$$|b(x, i) - b(y, i)| + |\sigma(x, i) - \sigma(y, i)| \leq L_{H,i}|x - y|, \quad \forall |x|, |y| \leq H, i \in \mathbb{Z}_+.$$

(ii) For each  $i \in \mathbb{Z}_+$ , there exist a twice continuously differentiable function  $V_i(x)$  and a constant  $C_i > 0$  such that

$$\lim_{R \rightarrow \infty} \left( \inf\{V_i(x) : |x| \geq R\} \right) = \infty \quad \text{and} \quad \mathcal{L}_i V_i(x) \leq C_i(1 + V_i(x)) \quad \forall x \in \mathbb{R}^n.$$

**Assumption 3.4.** Assume the following conditions hold.

(i) For each  $H > 0$ ,  $i \in \mathbb{Z}_+$ , there is a positive constant  $L_{H,i}$  such that

$$|b(x, i) - b(y, i)| + |\sigma(x, i) - \sigma(y, i)| \leq L_{H,i}|x - y| \quad \forall |x|, |y| \leq H, i \in \mathbb{Z}_+.$$

(ii) There exist a twice continuously differentiable function  $V(x)$  and a constant  $C > 0$  independent of  $i \in \mathbb{Z}_+$  such that

$$\lim_{R \rightarrow \infty} \left( \inf\{V(x) : |x| \geq R\} \right) = \infty \quad \text{and} \quad \mathcal{L}_i V(x) \leq C(1 + V(x)) \quad \forall x \in \mathbb{R}^n, i \in \mathbb{Z}_+.$$

**Theorem 3.5.** For given initial data  $(\xi, i_0)$ , there exists a unique solution  $(X(t), \alpha(t))$  to (2.3) if either of the following conditions is satisfied

- Assumption 3.3 and (ii) of Assumption 3.1,
- Assumption 3.4 and (ii) of Assumption 3.2.

*Proof.* It is well known that Assumption 3.3 guarantees the existence and uniqueness of solutions to (3.1). Hence, if (ii) in Assumption 3.1 is satisfied, we can prove the desired result by using the proof of Theorem 3.1. Now, suppose Assumption 3.4 and (ii) of Assumption 3.2 hold. Similar to the proof of Theorem 3.3, we can assume that  $(\xi, i_0)$  is bounded. Consider  $X(t)$  and define  $\tau_k$  as in the proof of Theorem 3.1. Then  $X(t)$  is the solution with initial data  $(\xi, i_0)$  to (2.3) on  $[0, T \wedge \tau_k]$  for any  $T > 0, k \in \mathbb{Z}_+$ . We have from the generalized Itô formula that

$$\begin{aligned} \mathbb{E}V(X(T \wedge \tau_k \wedge \eta_H)) &= \mathbb{E}V(\xi(0), i_0) + \mathbb{E} \int_0^{T \wedge \tau_k \wedge \eta_H} \mathcal{L}_i V(X(t), \alpha(t-)) dt \\ &\leq \mathbb{E}V(\xi(0), i_0) + C \mathbb{E} \int_0^{T \wedge \tau_k \wedge \eta_H} (1 + V(X(t))) dt, \end{aligned}$$

where  $\eta_H = \inf\{t \geq 0 : |X(t)| > H\}$ . Using the estimate above and the argument in [14, Theorem 3.19], we can show that

$$\mathbb{E}V(T \wedge \tau_k \wedge \eta_H) \leq K = K(\xi, T) \quad \forall H > 0, k \in \mathbb{Z}_+.$$

In view of the property  $\lim_{R \rightarrow \infty} \left( \inf\{V(x) : |x| \geq R\} \right) = \infty$ , for any  $\varepsilon > 0$ , there is  $H_\varepsilon > 0$  such that

$$\mathbb{P}\{\eta_{H_\varepsilon} > T \wedge \tau_k\} > 1 - \frac{\varepsilon}{2} \quad \forall k \in \mathbb{Z}_+.$$

Then, proceeding similarly as in the proof of Theorem 3.3 yields the existence and uniqueness of solutions with initial data  $(\xi, i_0)$  to (2.3).  $\square$

**Example 3.6.** (cont. of Example 2.2) We come back to Example 2.2 in Section 1. We want to show that  $X(t) > 0$  for all  $t \geq 0$  under certain conditions. To proceed, we can set  $Y(t) = \ln X(t)$  to obtain

$$dY(t) = [a(\alpha(t)) - \frac{\sigma^2(\alpha(t))}{2} - b(\alpha(t)) \exp(Y(t))]dt + \sigma(\alpha(t))dW(t). \quad (3.9)$$

To demonstrate (2.8) and (2.9) has a unique solution with  $X(t) > 0$  for all  $t \geq 0$ , it is equivalent to show that (3.9) and (2.9) has a strong solution on  $[0, \infty)$ . Let  $V(y) = e^y + e^{-y}$ . By direct calculation,

$$\begin{aligned} \mathcal{L}_i V(y) &= b(i) + (\sigma^2(i) - a(i))e^{-y} + a(i)e^y - b(i)e^{2y} \\ &\leq c(i) + (\sigma^2(i) - a(i))V(y), \end{aligned}$$

where  $c(i) = \max_{y \in \mathbb{R}} \{b(i) + (2a(i) - \sigma^2(i))e^y - b(i)e^{2y}\}$ . Applying Theorem 3.3, we can see that the equation has a unique solution if one of the following is satisfied

- $\beta_i(\phi) + \delta_i(\phi)$  is bounded uniformly in  $\phi \in \mathcal{C}_+ := \{\psi \in \mathcal{C} : \psi(t) > 0 \forall t \in [-r, 0]\}$  and  $i \in \mathbb{Z}_+$ .
- $c(i)$  and  $\sigma^2(i) - a(i)$  are bounded above uniformly and for each  $i \in \mathbb{Z}_+$ ,  $\beta_i(\phi) + \delta_i(\phi)$  is bounded in each compact subset of  $\phi \in \mathcal{C}_+$ .

It can be shown by applying the result of the next section that the process  $(Y_t, \alpha(t))$  has the Markov-Feller property if  $\beta_i(\cdot)$  and  $\delta_i(\cdot)$  are continuous in addition to one of the above conditions.

**Example 3.7.** (cont. of Example 2.3) To study the long-run average control problem in Example 2.3, it is important to make sure that the system under consideration processes ergodicity. Before the ergodicity can be verified, we need (2.11) has a unique solution for each initial condition. Denote the control set by  $\tilde{K}$  and assume it is a compact and convex set. Using a relaxed control representation  $m_t(\cdot)$  (see [11]) to represent the consumption rate  $c(\cdot)$ , we can rewrite (2.11) as

$$dX(t) = \left[ \int_{\tilde{K}} c(u)m_t(du) - \rho(\alpha(t))X(t) \right]dt + \sigma(X(t), \alpha(t))dW(t). \quad (3.10)$$

Assume that for each  $i \in \mathbb{Z}_+$ ,  $\sigma(x, i)$  satisfies the conditions in Assumption 3.1 (i), and  $Q(\phi)$  satisfies Assumption 3.1 (ii). Then the conditions of Theorem 3.1 are all verified. As a result, (3.10) has a unique solution for each initial condition.

## 4 Markov and Feller Properties

This section establishes the Markov and Feller properties of the process  $(X_t, \alpha(t))$ . While the Markov property can be derived by the well-known arguments, it requires much more efforts to obtain the Feller property. As already seen in the previous section, the past dependence and the use of  $\mathbb{Z}_+$  make the analysis more complex than that of the switching diffusions with diffusion-dependent switching living in a finite set. To overcome the difficulties, in this section, we carry out the analysis by introducing some auxiliary or intermediate processes. First, it would be better if we could untangle the past dependence of the switching process and the infinity of the cardinality of its state space. For this purpose, we introduce a continuous-time Markov chain independent of the past and continuous state; we call this process  $\gamma(t)$ . Then naturally, associated with  $\gamma(t)$ , we examine a pair of process  $(Z(t), \gamma(t))$ . Even after this introduction, in the analysis, we still need to look into the details of the switching process  $\alpha(t)$  such as when it jumps and the post jump

location etc. To do so, we introduce another auxiliary process  $Y(t)$ , which is a “fixed”- $i$  process. We then have another pair of processes  $(Y(t), \beta(t))$  to deal with. These auxiliary processes help us to establish the desired results. Their connections and interactions will be further specified in what follows.

First, note that the Brownian motion and the Poisson point process associated to  $\mathbf{p}(dt, dz)$  possess stationary strong Markov property, that is, for any finite stopping time  $\eta$ ,  $\{W^*(t)\}_{t \geq 0} = \{W(t+\eta) - W(\eta)\}_{t \geq 0}$  is an  $\mathcal{F}_t^*$ -Brownian motion and  $\mathbf{p}^*([t, t+s] \times U) = \mathbf{p}([t+\eta, t+s+\eta] \times U)$  is a Poisson random measure with density  $dt \times \mathbf{m}(dz)$  (see [22, Theorem 101]). Hence, by standard arguments, we can obtain the following theorem whose proof is omitted. In fact, the theorem can be proved essentially by imitating the proof in [17, Chap. 5], [1, Chap. 6], or [15, Chap. 7].

**Theorem 4.1.** *Assume that the hypotheses of Theorem 3.1, or Theorem 3.3, or Theorem 3.5 are satisfied. Let  $(X(t), \alpha(t))$  be a solution to (2.3). Then  $(X_t, \alpha(t))$  is a homogeneous strong Markov process taking value in  $\mathcal{C} \times \mathbb{Z}_+$  with transition probabilities*

$$P(\phi, i, t, A \times \{j\}) = \mathbb{P}\{X_t^{\phi, i} \in A, \alpha(t) = j\},$$

where  $X^{\phi, i}(t)$  is the solution to (2.3) with initial data  $(\phi, i) \in \mathcal{C} \times \mathbb{Z}_+$ .

We proceed with obtaining the Feller property of  $(X_t, \alpha(t))$ . Assuming that the hypotheses of Theorem 3.1, or Theorem 3.3, or Theorem 3.5 are satisfied leads to the existence and uniqueness of strong solutions. Next, we introduce an auxiliary hybrid diffusion with Markov switching. Let  $\gamma^i(t)$  be a Markov chain starting at  $i$  with generator  $\tilde{Q} = (\rho_{ij})$  for  $(i, j) \in \mathbb{Z}_+ \times \mathbb{Z}_+$ , where  $\rho_{ii} = -1$  and  $\rho_{ij} = 2^{-j}$  if  $j < i$  and  $\rho_{ij} = 2^{-j+1}$  if  $j > i$ , that is,

$$\tilde{Q} = \begin{pmatrix} -1 & 1/2 & 1/4 & \cdots \\ 1/2 & -1 & 1/4 & \cdots \\ 1/2 & 1/4 & -1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

We recursively define a sequence of stopping times  $\{\theta_k^i\}$  with  $\theta_k^i$  being the first jump time of  $\gamma^i(t)$  after  $\theta_{k-1}^i$  as follows

$$\theta_0^i = 0, \theta_k^i = \inf\{t > \theta_{k-1}^i : \gamma^i(t) \neq \gamma^i(\theta_{k-1}^i)\}, k \in \mathbb{Z}_+.$$

For  $(\phi, i) \in \mathcal{C} \times \mathbb{Z}_+$ , let  $Z^{\phi, i}(t)$  be the solution to

$$dZ(t) = b(Z(t), \gamma(t))dt + \sigma(Z(t), \gamma(t))dW(t), t \geq 0$$

satisfying  $Z^{\phi, i}(t) = \phi(t)$  in  $[-r, 0]$  and  $\gamma(0) = i$ . Similar to Girsanov’s theorem, which tells us how to convert an Itô process to a Brownian motion under a change of measure, we aim to establish a change of measure allowing us to “convert” a hybrid diffusion with past-dependent switching to a hybrid diffusion with Markov switching. To establish such a change of measure, we need to find the distribution of jump times of  $\alpha(t)$ . Because of the interactions between  $\alpha(t)$  and  $X(t)$ , we need to introduce another auxiliary (or intermediate) process, which helps to examine the distribution of the jump times of  $\alpha(t)$ . Let  $(Y^{\phi, i}(t), \beta^{\phi, i}(t))$  be the solution to

$$\begin{cases} dY(t) = b(Y(t), i)dt + \sigma(Y(t), i)dW(t), t \geq 0 \\ d\beta(t) = \int_{\mathbb{R}} h(Y_t, \beta(t-), z)\mathbf{p}(dt, dz), t \geq 0 \end{cases} \quad (4.1)$$

satisfying  $Y^{\phi,i}(t) = \phi(t)$  in  $[-r, 0]$  and  $\beta^{\phi,i}(0) = i$ . By the definition,  $\alpha^{\phi,i}(t) = \beta^{\phi,i}(t)$ ,  $X^{\phi,i}(t) = Y^{\phi,i}(t)$  up to the first jump time of the two process  $\alpha(t)$  and  $\beta(t)$ . There is an advantage working with  $(Y^{\phi,i}(t), \beta^{\phi,i}(t))$ . Unlike the pair  $(X(t), \alpha(t))$  in which  $\alpha(t)$  depends on the continuous state, the process  $Y^{\phi,i}(t)$  evolving for a fixed discrete state  $i$  that does not depend on  $\beta^{\phi,i}(t)$ . Thus, it is easier to examine, for example, the first jump time of  $\beta^{\phi,i}(t)$  (or  $\alpha^{\phi,i}(t)$ ).

Next we recursively define sequences of stopping times associated with  $\beta(t)$  and  $\alpha(t)$  so that  $\lambda_k^{\phi,i}$  and  $\tau_k^{\phi,i}$  are the first jump times of the processes  $\beta^{\phi,i}(t)$  and  $\alpha^{\phi,i}(t)$  after  $\lambda_{k-1}^{\phi,i}$  and  $\tau_{k-1}^{\phi,i}$ , respectively. More specifically, for  $k \in \mathbb{Z}_+$ , let

$$\lambda_0^{\phi,i} = 0, \lambda_k^{\phi,i} = \inf\{t > \lambda_{k-1}^{\phi,i} : \beta^{\phi,i}(t) \neq \beta^{\phi,i}(\lambda_{k-1}^{\phi,i})\}, i \in \mathbb{Z}_+.$$

and

$$\tau_0^{\phi,i} = 0, \tau_k^{\phi,i} = \inf\{t > \tau_{k-1}^{\phi,i} : \alpha^{\phi,i}(t) \neq \alpha^{\phi,i}(\tau_{k-1}^{\phi,i})\}, i \in \mathbb{Z}_+.$$

To simplify the notation, we put

$$\alpha_k^{\phi,i} := \alpha^{\phi,i}(\tau_k^{\phi,i}), \beta_k^{\phi,i} := \beta^{\phi,i}(\lambda_k^{\phi,i}), \gamma_k^i := \gamma^i(\theta_k^i),$$

and

$$X_{(k)}^{\phi,i} := X_{\tau_k^{\phi,i}}^{\phi,i}, Y_{(k)}^{\phi,i} := Y_{\lambda_k^{\phi,i}}^{\phi,i}, Z_{(k)}^{\phi,i} := Z_{\theta_k^i}^{\phi,i},$$

where we use the subscript  $k$  with parentheses to avoid confusion with the function-valued processes  $X_t^{\phi,i}, Y_t^{\phi,i}, Z_t^{\phi,i}$  at  $t = k$ .

**Lemma 4.2.** *Let  $g : \mathcal{C} \times \mathbb{R}_+ \times \mathbb{Z}_+ \mapsto \mathbb{R}$  be a bounded and measurable function, and  $\mathcal{F}_T^W$  be the  $\sigma$ -algebra generated by  $\{W(t), t \in [0, T]\}$ . The following assertions hold:*

- (i)  $\mathbb{P}(\{\lambda_1^{\phi,i} > t\} | \mathcal{F}_T^W) = \mathbb{E}[\mathbf{1}_{\{\lambda_1^{\phi,i} > t\}} | \mathcal{F}_T^W] = \exp(-\int_0^t q_i(Y_s^{\phi,i}) ds) \quad \forall t \in [0, T].$
- (ii)  $\mathbb{E}\left[g(Y_{(1)}^{\phi,i}, \lambda_1^{\phi,i}, \beta_1^{\phi,i}) \mathbf{1}_{\{\lambda_1^{\phi,i} \leq T\}} \middle| \mathcal{F}_T^W\right] = \sum_{j=1, j \neq i}^{\infty} \int_0^T g(Y_t, t, j) q_{ij}(Y_t) \exp(-\int_0^t q_i(Y_s) ds) dt.$

As indicated previously, it is difficult to estimate the difference of  $X_t^{\phi_1,i}$  and  $X_t^{\phi_2,i}$  because the states of  $\alpha^{\phi_1,i}(t)$  and  $\alpha^{\phi_2,i}(t)$  may differ significantly due to the continuous state dependence. In contrast, it is considerably easier to compare  $Z_t^{\phi_1,i}$  and  $Z_t^{\phi_2,i}$  because of the continuous-state-dependent switching is replaced by the continuous-state-independent Markov chain. With help of the intermediate process  $(Y(t), \beta(t))$  and Lemma 4.2, we obtain the following change of measure formula, which is a bridge to connect the continuous-state-dependent and continuous-state-independent processes.

**Proposition 4.3.** *For any  $T > 0$ , let  $f(\cdot, \cdot) : \mathcal{C} \times \mathbb{Z}_+ \mapsto \mathbb{R}$  be a bounded continuous function. For any  $l = 0, 1, \dots$ , any  $i_k \in \mathbb{Z}_+$  with  $i_k \neq i_{k+1}$  and  $k = 1, \dots, l+1$ , and any  $(\phi, i) \in \mathcal{C} \times \mathbb{Z}_+$ ,*

$$\begin{aligned} & \mathbb{E}\left[f(X_T^{\phi,i}, \alpha^{\phi,i}(T)) \mathbf{1}_{\{\tau_l^{\phi,i} \leq T < \tau_{l+1}^{\phi,i}\}} \prod_{k=1}^l \mathbf{1}_{\{\alpha_k^{\phi,i} = i_k\}}\right] \\ &= \exp(T) \mathbb{E}\left[f(Z_T^{\phi,i}, i_l) \mathbf{1}_{\{\theta_l^i \leq T < \theta_{l+1}^i\}} \prod_{k=1}^l \left(\mathbf{1}_{\{\gamma_k^i = i_k\}} \frac{q_{i_k i_{k+1}}(Z_{(k)}^{\phi,i})}{\rho_{i_k i_{k+1}}}\right) \exp\left\{-\int_0^T q_{\gamma^i(s)}(Z_s^{\phi,i}) ds\right\}\right]. \end{aligned} \tag{4.2}$$

**Remark 4.4.** The proofs of Lemma 4.2 and Proposition 4.3 will be given in the Appendix. We are now in a position to prove the Feller property for the solution to (2.3). In addition to the sufficient conditions for the existence and uniqueness of solution, we prove the Feller property of the solution only with an additional condition that  $q_{ij}(\phi)$  is continuous in  $\phi$  for any  $i, j \in \mathbb{Z}_+$ . There are some difficulties because the process  $\{X_t\}$  takes value in an infinite dimensional Banach space and the switching  $\{\alpha(t)\}$  has an infinite state space. Moreover, although we suppose that  $q_{ij}(\phi)$  is continuous, neither the uniform continuity in  $\phi \in \mathcal{C}$  nor equi-continuity in  $i, j \in \mathbb{Z}_+$  is assumed. Because of these difficulties, we divide the proof into several steps. First, we make the following assumptions, which will be relaxed later.

**Assumption 4.1.** Assume the following conditions hold.

- (i) For each  $i \in \mathbb{Z}_+$ ,  $b(x, i)$  and  $\sigma(x, i)$  are Lipschitz continuous functions that are vanishing outside  $\{x : |x| \leq R\}$  for some  $R > 0$ .
- (ii)  $M := \sup\{q_i(\phi) : i \in \mathbb{Z}_+, \phi \in \mathcal{C}\} < \infty$ .
- (iii) For each  $i, j \in \mathbb{Z}_+, j \neq i$ ,  $q_i(\cdot)$  and  $q_{ij}(\cdot)$  are continuous on  $\mathcal{C}$ .

Before applying (4.2) to prove the continuous dependence of  $u_f(\phi, i) = \mathbb{E}_{\phi, i} f(X_T, \alpha(T))$  on  $(\phi, i)$ , we first need the following lemma.

**Lemma 4.5.** *Assume that Assumption 4.1 is satisfied. Let  $(\phi_0, i_0) \in \mathcal{C} \times \mathbb{Z}_+$  with  $\|\phi_0\| \leq R$  and  $T > 0$ . For each  $\Delta > 0$ , there exist  $m = m(\Delta) \in \mathbb{Z}_+$ ,  $n_m = n_m(\Delta) \in \mathbb{Z}_+$ , and  $d_m = d_m(\Delta) > 0$  such that*

$$\mathbb{P}\left(\{\tau_{m+1}^{\phi, i_0} > T\} \cap \{\alpha^{\phi, i_0}(t) \in N_{n_m}, \forall t \in [0, T]\}\right) \geq 1 - \Delta, \quad \forall \|\phi - \phi_0\| < d_m,$$

where  $N_k = \{1, \dots, k\}$ .

This lemma allows us to confine our attention to a finite subset of  $\mathbb{Z}_+$  (the state space of  $\alpha^{\phi, i_0}(\cdot)$ ) and a finite number of jumps when  $\phi$  is close to  $\phi_0$ . It is a crucial step in providing some uniform estimates because we do not assume the equi-continuity of  $q_{ij}(\cdot)$  in either  $i$  or  $j$ . Since the switching intensity of  $\alpha^{\phi, i_0}(t)$  depends on  $X_t^{\phi, i_0}$ , in order to obtain Lemma 4.5, we need to show that with an arbitrarily large probability,  $X_t^{\phi, i_0}, t \in [0, T]$  belongs to a compact set in  $\mathcal{C}$  for any  $\phi$  sufficiently close to  $\phi_0$ . Note that under some suitable conditions, sample paths of a diffusion process in a finite interval  $[0, T]$  are Hölder continuous. Thus, it is easy to find a compact set in which sample paths of a diffusion process lie with a large probability. Our arguments rely on this fact. However, the initial data  $\phi$  of our process  $X(t)$  does not always satisfy the Hölder continuity. Moreover,  $X(t)$  depends on the state of  $\alpha(t)$ . We therefore need to introduce the following operator, which is motivated by merging trajectories of  $X(t)$  at jump times. For  $\mathcal{A}, \mathcal{B} \subset \mathcal{C}$ , we define the set of continuous functions that are formed by merging functions in  $\mathcal{A}$  and  $\mathcal{B}$  as follows.

$$\begin{aligned} \mathcal{A} \uplus \mathcal{B} &:= \mathcal{A} \cup \mathcal{B} \cup \{\psi \in \mathcal{C} : \exists \psi_1 \in \mathcal{A}, \psi_2 \in \mathcal{B}, s \in [0, r] \text{ such that} \\ &\quad \psi(t) = \psi_1(s + t) \forall t \in [-r, -s], \psi(t) = \psi_2(t + s - r) \forall t \in [-s, 0]\}. \end{aligned}$$

By virtue of the Arzelá-Ascoli theorem, if  $\mathcal{A}$  and  $\mathcal{B}$  are compact, so is  $\mathcal{A} \uplus \mathcal{B}$ . Using this fact and the Hölder continuity of sample paths of a diffusion process, we can find a suitable compact set to which  $X_t^{\phi, i_0}, t \in [0, T]$ , belongs with a large probability for any  $\phi$  which is sufficiently close to  $\phi_0$ .

Then, Lemma 4.5 can be proved. The details of the proof are postponed to the Appendix. Now, we point out some nice properties of the diffusion process with Markov switching  $(Z(t), \gamma(t))$ , which are useful to compare the sample paths of  $Z(t)$  with different initial values.

**Lemma 4.6.** *Fix  $i_0 \in \mathbb{Z}_+$ . For each  $k \in \mathbb{Z}_+$  and  $\varepsilon > 0$ , there is an  $\hbar_k = \hbar_k(\varepsilon) > 0$  such that*

$$\mathbb{P}\left\{\sup_{t,s \in [0, T \wedge \iota_k], 0 < t-s < \hbar_k} \frac{|Z^{\phi, i_0}(t) - Z^{\phi, i_0}(s)|}{(s-t)^{0.25}} \leq 4\right\} > 1 - \varepsilon \forall |\phi(0)| \leq R,$$

and

$$\mathbb{E}\left[\sup_{t \in [0, T \wedge \iota_k]} |Z^{\phi, i_0} - Z^{\psi, i_0}|^2\right] \leq \bar{C}|\phi - \psi|^2,$$

where  $\iota_k = \inf\{t > 0 : \gamma^{i_0}(t) > k\}$  and  $\bar{C}$  is some positive constant.

*Proof.* Since  $b(x, i)$  and  $\sigma(x, i)$  are Lipschitzian in  $x$  uniformly in  $N_k$ , by standard arguments (e.g., [14, Theorem 3.23]), we can show that

$$\mathbb{E}|x(t \wedge \iota_k) - x(s \wedge \iota_k)|^6 < \tilde{C}_k(t-s)^3, \forall 0 \leq s \leq t \leq T.$$

Using the Kolmogorov-Centsov theorem, we obtain the first inequality. The details are similar to the proof of Lemma 4.5 in the Appendix. The second claim is proved in the same manner as that of [25, Lemma 2.14].  $\square$

Having Lemmas 4.5 and 4.6, we are ready to use the change of measure (4.2) to prove the Feller property of  $(X_t, \alpha(t))$  under Assumption 4.1.

**Proposition 4.7.** *Suppose that Assumption 4.1 is satisfied. Let  $f(\cdot, \cdot) : \mathcal{C} \times \mathbb{Z}_+ \mapsto \mathbb{R}$  be continuous and bounded. Then for any  $T > 0$ ,  $u_f(\phi, i) = \mathbb{E}f(X_T^{\phi, i}, \alpha^{\phi, i}(T))$  is a continuous function in  $\phi \in \mathcal{C}$ .*

*Proof.* We suppose without loss of generality that  $|f(\phi, i)| \leq 1 \forall (\phi, i) \in \mathcal{C} \times \mathbb{Z}_+$ . Fix  $(\phi_0, i_0) \in \mathcal{C} \times \mathbb{Z}_+$ . We show that for any  $\Delta > 0$ , there exists  $d^* = d^*(\Delta, \phi_0, i_0) > 0$  such that

$$|\mathbb{E}f(X_T^{\phi, i_0}, \alpha^{\phi, i_0}(T)) - \mathbb{E}f(X_T^{\phi_0, i_0}, \alpha^{\phi_0, i_0}(T))| \leq 3\Delta \forall \|\phi - \phi_0\| < d^*. \quad (4.3)$$

In view of Lemma 4.5, there are  $m, n_m \in \mathbb{Z}_+$ , and  $d_m > 0$  such that

$$\mathbb{P}\left(\{\tau_{m+1}^{\phi, i_0} > T\} \cap \{\alpha^{\phi, i_0}(t) \in N_{n_m}, \forall t \in [0, T]\}\right) \geq 1 - \Delta \forall \|\phi - \phi_0\| < \frac{d_m}{2}. \quad (4.4)$$

Let  $\varepsilon = \varepsilon(\Delta) > 0$  (to be specified later). Let  $\hbar_k$  be as in Lemma 4.6. Denote

$$\tilde{\mathcal{H}} = \left\{ \psi(\cdot) \in \mathcal{C} : \|\psi\| \leq R + 1 \text{ and } \sup_{t,s \in [-r, 0], 0 < t-s < \hbar_{n_m}} \frac{|\psi(s) - \psi(t)|}{(s-t)^{0.25}} \leq 4 \right\}$$

and  $\tilde{\mathcal{K}} = \{\phi_0\} \uplus \tilde{\mathcal{H}}$ . By the compactness of  $\tilde{\mathcal{K}}$ , there is a  $\tilde{d}_m > 0$  such that

$$\|q_{ij}(\psi) - q_{ij}(\phi)\| < \varepsilon, \quad |f(\psi, i) - f(\phi, i)| < \varepsilon \quad (4.5)$$

if  $\phi \in \tilde{\mathcal{K}}, i, j \in N_{n_m}$  and  $\|\psi - \phi\| < \tilde{d}_m$ . In view of Lemma 4.6, we can choose  $\hat{d}_m > 0$  such that

$$\mathbb{P}\left\{\sup_{t \in [0, T \wedge \iota_k]} \|Z_t^{\phi, i_0} - Z_t^{\phi_0, i_0}\| \leq \tilde{d}_m\right\} < \varepsilon \text{ if } \|\phi - \phi_0\| \leq \hat{d}_m. \quad (4.6)$$



Let  $A^\phi$  be the event  $\{\tau_m^{\phi,i_0} \leq T < \tau_{m+1}^{\phi,i_0}, \iota_{n_m} > T\}$  and  $l(T)$  be the number of jumps up to time  $T$  of  $\gamma^{i_0}(t)$ . It follows from Proposition 4.3 that

$$\begin{aligned} & \mathbb{E}[f(X_T^{\phi,i_0}, \alpha^{\phi,i_0}(T))\mathbf{1}_{A^\phi}] - \mathbb{E}[f(X_T^{\phi_0,i_0}, \alpha^{\phi_0,i_0}(T))\mathbf{1}_{A^{\phi_0}}] \\ &= \exp(T)\mathbb{E}\left[\mathbf{1}_{\{l(T) < m+1, \iota_{n_m} > T\}}\left[g(Z^{\phi,i_0}(\cdot), \gamma^{i_0}(\cdot)) - g(Z^{\phi_0,i_0}(\cdot), \gamma^{i_0}(\cdot))\right]\right], \end{aligned} \quad (4.7)$$

where

$$g(Z^{\phi,i_0}(\cdot), \gamma^{i_0}(\cdot)) = f(Z_T^{\phi,i_0}, \gamma^{i_0}(T)) \prod_{k=1}^{l(T)} \frac{q_{\gamma_k^{i_0} \gamma_{k+1}^{i_0}}(Z_{(k+1)}^{\phi,i_0})}{\rho_{\gamma_k^{i_0} \gamma_{k+1}^{i_0}}} \exp\left\{-\int_0^T q_{\gamma^{i_0}(s)}(Z_s^{\phi,i_0}) ds\right\}.$$

Let  $D_m^\phi$  be the event

$$D_m^\phi := \left\{ \sup_{t \in [0, T \wedge \iota_k]} \|Z_t^{\phi,i_0} - Z_t^{\phi_0,i_0}\| \leq \tilde{d}_m \right\} \cap \left\{ \sup_{t,s \in [0, T \wedge \iota_k], 0 < t-s < h_{n_m}} \frac{|Z^{\phi_0,i_0}(s) - Z^{\phi_0,i_0}(t)|}{(s-t)^{0.25}} \leq 4 \right\}.$$

Using (4.7) and the estimates in [25, Lemma 2.17], we obtain for  $l \geq 1$ ,

$$\begin{aligned} & \left| \mathbb{E}[f(X_T^{\phi,i_0}, \alpha^{\phi,i_0}(T))\mathbf{1}_{A^\phi}] - \mathbb{E}[f(X_T^{\phi_0,i_0}, \alpha^{\phi_0,i_0}(T))\mathbf{1}_{A^{\phi_0}}] \right| \\ & \leq K\mathbb{E}\left[\mathbf{1}_{\{\theta_l^{i_0} \leq T < \theta_{l+1}^{i_0}, \iota_{n_m} > T\}} \times \sup_{i \in N_{n_m}} |f(Z_T^{\phi,i_0}, i) - f(Z_T^{\phi_0,i_0}, i)|\right] \\ & \quad + K\mathbb{E}\left[\mathbf{1}_{\{\theta_l^{i_0} \leq T < \theta_{l+1}^{i_0}, \iota_{n_m} > T\}} \times \sup_{t \in [0, T], i, j \in N_{n_m}} |q_{ij}(Z_t^{\phi,i_0}) - q_{ij}(Z_t^{\phi_0,i_0})|\right] \\ & \leq K\mathbb{E}\left[\mathbf{1}_{\{\theta_l^{i_0} \leq T < \theta_{l+1}^{i_0}, \iota_{n_m} > T\}} \mathbf{1}_{D_m^\phi} \times \sup_{i \in N_{n_m}} |f(Z_T^{\phi,i_0}, i) - f(Z_T^{\phi_0,i_0}, i)|\right] \\ & \quad + K\mathbb{E}\left[\mathbf{1}_{\{\theta_l^{i_0} \leq T < \theta_{l+1}^{i_0}, \iota_{n_m} > T\}} \mathbf{1}_{D_m^\phi} \times \sup_{t \in [0, T], i, j \in N_{n_m}} |q_{ij}(Z_t^{\phi,i_0}) - q_{ij}(Z_t^{\phi_0,i_0})|\right] \\ & \quad + 2K(M+1)\mathbb{P}(\Omega \setminus D_m^\phi), \end{aligned}$$

where  $K$  is a constant depending only on  $T, m, n_m$ .

Note that if  $\omega \in \{\theta_l^{i_0} \leq T < \theta_{l+1}^{i_0}, \iota_{n_m} > T\} \cap D_m^\phi$ , then  $Z_t^{\phi_0,i_0} \in \tilde{\mathcal{K}}$  and  $\|Z_t^{\phi,i_0} - Z_t^{\phi_0,i_0}\| \leq \tilde{d}_m \forall t \in [0, T]$  which implies in view of (4.5) that

$$\sup_{i \in N_{n_m}} |f(Z_T^{\phi,i_0}, i) - f(Z_T^{\phi_0,i_0}, i)| + \sup_{t \in [0, T], i, j \in N_{n_m}} |q_{ij}(Z_t^{\phi,i_0}) - q_{ij}(Z_t^{\phi_0,i_0})| < 2\varepsilon.$$

On the other hand, Lemma 4.6 and (4.6) imply that

$$\mathbb{P}(\Omega \setminus D_m^\phi) \leq 3\varepsilon \text{ if } \|\phi - \phi_0\| \leq \hat{d}_m.$$

Hence for  $\|\phi - \phi_0\| \leq \hat{d}_m$ , we have that

$$\left| \mathbb{E}[f(X_T^{\phi,i_0}, \alpha^{\phi,i_0}(T))\mathbf{1}_{A^\phi}] - \mathbb{E}[f(X_T^{\phi_0,i_0}, \alpha^{\phi_0,i_0}(T))\mathbf{1}_{A^{\phi_0}}] \right| \leq 2K(4+3M)\varepsilon, \quad (4.8)$$

Note that

$$\mathbb{P}(\Omega \setminus A^\phi) = \mathbb{P}(\{\tau_{m+1}^{\phi,i_0} < T\} \cup \{\alpha^{\phi,i_0}(t) \notin N_{n_m} \text{ for some } t \in [0, T]\}) < \Delta,$$

which implies

$$\left| \mathbb{E}[f(X_T^{\phi, i_0}, \alpha^{\phi, i_0}(T)) \mathbf{1}_{\Omega \setminus A^\phi}] - \mathbb{E}[f(X_T^{\phi_0, i_0}, \alpha^{\phi_0, i_0}(T)) \mathbf{1}_{\Omega \setminus A^\phi}] \right| \leq 2\Delta. \quad (4.9)$$

Choosing  $\varepsilon = \frac{\Delta}{2K(4+3M)}$ , we have from (4.8) and (4.9) that

$$\left| \mathbb{E}[f(X_T^{\phi, i_0}, \alpha^{\phi, i_0}(T))] - \mathbb{E}[f(X_T^{\phi_0, i_0}, \alpha^{\phi_0, i_0}(T))] \right| \leq 3\Delta$$

if  $\|\phi - \phi_0\| < d^* := \frac{d_m}{2} \wedge \widehat{d}_m$ .  $\square$

With the above technical preparations, we are now in a position to prove the main theorem of this section. By using truncation arguments, we can obtain the Feller property of  $(X_t, \alpha(t))$  even if  $b(\cdot, i), \sigma(\cdot, i)$  do not vanish outside a bounded region and  $q_i(\phi)$  is not bounded. The precise condition is given below.

**Theorem 4.8.** *Let either Assumption 3.1 or Assumption 3.2 be satisfied. Assume further that  $q_{ij}(\cdot)$  is a continuous function for any  $i, j \in \mathbb{Z}_+$ . Then the solution to (2.3) has the Feller property.*

*Proof.* Let  $f(\cdot, \cdot) : \mathcal{C} \times \mathbb{Z}_+ \mapsto \mathbb{R}$  be a continuous function with  $|f(\phi, i)| \leq 1 \forall (\phi, i) \in \mathcal{C} \times \mathbb{Z}_+$ . Fix  $R > 0, T > 0$ . Suppose that  $\|\phi_0\| < R$ . Under the hypotheses of Theorem 3.1, or Theorem 3.3, or Theorem 3.5, it is shown in the proofs of those theorems that for any  $\varepsilon > 0$ , there exists an  $\tilde{R} > 0$  such that

$$\mathbb{P}\{\|X_t^{\phi, i_0}\| \leq \tilde{R}\} > 1 - \frac{\varepsilon}{8} \forall \|\phi\| \leq R + 1. \quad (4.10)$$

Let  $\Phi(x) : \mathbb{R}^n \mapsto \mathbb{R}$  be a twice continuously differentiable satisfying  $\Phi(x) = 1$  if  $|x| \leq \tilde{R}$  and  $\Phi(x) = 0$  if  $|x| \geq \tilde{R} + 1$ . Let  $(\tilde{X}_t^{\phi, i_0}, \tilde{\alpha}^{\phi, i_0}(t))$  be the solution with initial data  $(\phi, i_0)$  to

$$\begin{cases} d\tilde{X}(t) = \Phi(\tilde{X}(t))b(\tilde{X}(t), \tilde{\alpha}(t))dt + \Phi(\tilde{X}(t))\sigma(\tilde{X}(t), \tilde{\alpha}(t))dW(t) \\ d\tilde{\alpha}(t) = \int_{\mathbb{R}} h(\tilde{X}_t, \tilde{\alpha}(t-), z)\mathbf{p}(dt, dz). \end{cases} \quad (4.11)$$

Then  $(\tilde{X}^{\phi, i_0}(t), \tilde{\alpha}^{\phi, i_0}(t)) = (X^{\phi, i_0}(t), \alpha^{\phi, i_0}(t))$  up to the time that  $\|X_t^{\phi, i_0}\| > \tilde{R}$ , which combined with (4.10) implies

$$\mathbb{P}\{\tilde{\Omega}_{\phi, i_0}\} > 1 - \frac{\varepsilon}{8}, \quad \forall \|\phi\| < R$$

where  $\tilde{\Omega}_{\phi, i_0} := \{\tilde{X}_T^{\phi, i_0} = X_T^{\phi, i_0}, \tilde{\alpha}^{\phi, i_0}(T) = \alpha^{\phi, i_0}(T)\}$ . As a result, if  $\|\phi\| < R$ , we have

$$\begin{aligned} & \left| \mathbb{E}f(X_T^{\phi, i_0}, \alpha^{\phi, i_0}(T)) - \mathbb{E}f(\tilde{X}_T^{\phi, i_0}, \tilde{\alpha}^{\phi, i_0}(T)) \right| \\ & \leq \mathbb{E} \left[ \mathbf{1}_{\tilde{\Omega}_{\phi, i_0}^c} \left| f(X_T^{\phi, i_0}, \alpha^{\phi, i_0}(T)) - f(\tilde{X}_T^{\phi, i_0}, \tilde{\alpha}^{\phi, i_0}(T)) \right| \right] \quad (\text{with } \tilde{\Omega}_{\phi, i_0}^c = \Omega \setminus \tilde{\Omega}_{\phi, i_0}) \\ & \leq 2\mathbb{P} \left( \mathbf{1}_{\tilde{\Omega}_{\phi, i_0}^c} \right) \leq 2\frac{\varepsilon}{8} = \frac{\varepsilon}{4}. \end{aligned} \quad (4.12)$$

It follows from Proposition 4.7 that there exists a  $\delta \in (0, 1)$  such that if  $\|\phi - \phi_0\| < \delta$ , we have

$$\left| \mathbb{E}f(\tilde{X}_T^{\phi, i_0}, \tilde{\alpha}^{\phi, i_0}(T)) - \mathbb{E}f(\tilde{X}_T^{\phi_0, i_0}, \tilde{\alpha}^{\phi_0, i_0}(T)) \right| < \frac{\varepsilon}{2}. \quad (4.13)$$

Since  $\|f\| \leq 1$ , we can easily obtain from (4.12) and (4.13) that

$$\begin{aligned}
& \left| \mathbb{E}f(X_T^{\phi, i_0}, \alpha^{\phi, i_0}(T)) - \mathbb{E}f(X_T^{\phi_0, i_0}, \alpha^{\phi_0, i_0}(T)) \right| \\
& \leq \left| \mathbb{E}f(\tilde{X}_T^{\phi, i_0}, \tilde{\alpha}^{\phi, i_0}(T)) - \mathbb{E}f(\tilde{X}_T^{\phi_0, i_0}, \tilde{\alpha}^{\phi_0, i_0}(T)) \right| \\
& \quad + \left| \mathbb{E}f(X_T^{\phi, i_0}, \alpha^{\phi, i_0}(T)) - \mathbb{E}f(\tilde{X}_T^{\phi, i_0}, \tilde{\alpha}^{\phi, i_0}(T)) \right| \\
& \quad + \left| \mathbb{E}f(X_T^{\phi_0, i_0}, \alpha^{\phi_0, i_0}(T)) - \mathbb{E}f(\tilde{X}_T^{\phi_0, i_0}, \tilde{\alpha}^{\phi_0, i_0}(T)) \right| \\
& < \frac{\varepsilon}{2} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon, \quad \text{if } \|\phi - \phi_0\| < \delta.
\end{aligned}$$

The proof of the theorem is complete.  $\square$

## 5 Feller Property of Hybrid Diffusion without Past Dependence

Now, we suppose that the  $q_{ij}, i, j \in \mathbb{Z}_+$  associated with  $\alpha(t)$  depend only the current state of  $X(t)$ . To be more precise  $q_{ij}(\cdot)$  is a function from  $\mathbb{R}^n$  to  $\mathbb{R}$  for each  $(i, j) \in \mathbb{Z}_+ \times \mathbb{Z}_+$ . As a special case of the hybrid diffusion with past-dependent switching, we obtain the following theorem.

**Theorem 5.1.** *Assume that  $q_{ij}(\cdot)$  is a continuous function for any  $i, j \in \mathbb{Z}_+$ . Assume further that one of the following conditions is satisfied:*

- (A) *Assumption 3.3 and  $q_i(y) = \sum_{j \neq i} q_{ij}(y)$  is bounded uniformly in  $(i, y) \in \mathbb{Z}_+ \times \mathbb{R}^n$ .*
- (B) *Assumption 3.4 and  $q_i(y)$  is bounded uniformly in  $(i, y) \in \mathbb{Z}_+ \times K$  for each compact subset  $K$  of  $\mathbb{R}^n$ .*

*Then the unique solution to (2.3) is a Markov process having the Feller property.*

**Remark 5.2.** If for each discrete state  $i \in \mathbb{Z}_+$ , the diffusion  $Y^{(i)}(t)$ , which is the solution process to

$$dY^{(i)}(t) = b(Y^{(i)}(t), i)dt + \sigma(Y^{(i)}(t), i)dW(t) \quad (5.1)$$

has the strong Feller property, we do not need the continuity of  $q_{ij}(\cdot)$  to get the Feller property of  $(X(t), \alpha(t))$ . In fact, we will obtain a stronger result, namely, the strong Feller property. The condition for the strong Feller property of  $Y^{(i)}(t)$  is essentially the ellipticity of  $A(x, i)$  or the Hörmander condition for hyperellipticity (see e.g., [16, 23]).

**Theorem 5.3.** *Assume that  $q_{ij}(\cdot)$  is measurable for any  $i, j \in \mathbb{Z}_+$  and either (A) or (B) in Theorem 5.1 holds. If for each  $i \in \mathbb{Z}_+$ , the solution process  $Y^{(i)}(t)$  to (5.1) has the strong Feller property, then the unique solution to (2.3) has the strong Feller property, that is, for each bounded measurable function  $g(y, i) : \mathbb{R}^n \times \mathbb{Z}_+ \rightarrow \mathbb{R}$ , the function  $(x, i) \rightarrow \mathbb{E}g(X^{x, i}(T), \alpha^{x, i}(T))$  is continuous for each  $T > 0$ .*

*Proof.* We assume without loss of generality that  $|g(z, i)| \leq 1 \forall z \in \mathbb{R}^n, i \in \mathbb{Z}_+$ . Let  $Y^{y, i}(t)$  be the solution with initial data  $y$  to (5.1). Fix  $(x, i) \in \mathbb{R}^n \times \mathbb{Z}_+$  and  $\varepsilon > 0$ . Under the assumption (A) or (B), we can show that for each  $x \in \mathbb{R}^n$ , there is a  $K > 0$  satisfying

$$\mathbb{P}((\Omega_\varepsilon^{y, i})^c) < \frac{\varepsilon}{8} \forall y \in B(x, 1) := \{z : |x - z| < 1\}, \quad (5.2)$$

where  $\Omega_\varepsilon^{y,i} = \{|Y^{y,i}(t)| \leq K \forall t \in [0, 1]\}$ ,  $(\Omega_\varepsilon^{y,i})^c = \Omega \setminus \Omega_\varepsilon^{y,i}$ . Let  $M = \sup_{i \in \mathbb{Z}_+, |z| \leq K} q_i(z)$  and  $t_0 \in (0, 1)$  satisfying  $1 - \exp\{-Mt_0\} < \frac{\varepsilon}{16}$ . It follows from (5.2) and (i) of Lemma 4.2 that

$$\mathbb{P}\{\tau^{y,i} > t_0\} > 1 - \frac{3\varepsilon}{16} \forall y \in B(x, 1), \quad (5.3)$$

where

$$\tau^{y,i} := \inf\{t > 0 : \alpha^{y,i}(t) \neq i\} = \inf\left\{t > 0 : \int_0^t \int_{\mathbb{R}} h(Y^{y,i}(s), i, u) \mathbf{p}(ds, du) \neq 0\right\}.$$

Denote  $\Phi(y, i) := \mathbb{E}g(X^{y,i}(T - t_0), \alpha^{y,i}(T - t_0))$ . The condition  $|g(y, i)| \leq 1$  implies  $|\Phi(y, i)| \leq 1$  for all  $y \in \mathbb{R}^n, i \in \mathbb{Z}_+$ . By the strong Feller property of  $Y^{(i)}(t)$ , there is a  $\delta > 0$  such that

$$|\mathbb{E}\Phi(Y^{y,i}(t_0), i) - \mathbb{E}\Phi(Y^{x,i}(t_0), i)| \leq \frac{\varepsilon}{4} \forall y \in B(x, \delta). \quad (5.4)$$

By the strong Markov property of  $X(t)$ , we have

$$\begin{aligned} \mathbb{E}g(X^{y,i}(T), \alpha^{y,i}(T)) &= \mathbb{E}\Phi(X^{y,i}(t_0), \alpha^{y,i}(t_0)) \\ &= \mathbb{E}[\mathbf{1}_{\{\tau^{y,i} > t_0\}} \Phi(X^{y,i}(t_0), \alpha^{y,i}(t_0))] + \mathbb{E}[\mathbf{1}_{\{\tau^{y,i} \leq t_0\}} \Phi(X^{y,i}(t_0), \alpha^{y,i}(t_0))]. \end{aligned} \quad (5.5)$$

Applying (i) of Lemma 4.2, we obtain

$$\begin{aligned} &\mathbb{E}[\mathbf{1}_{\{\tau^{y,i} > t_0\}} \Phi(X^{y,i}(t_0), \alpha^{y,i}(t_0))] \\ &= \mathbb{E}[\Phi(Y^{y,i}(t_0), \alpha^{y,i}(t_0)) \exp\{-\int_0^{t_0} q_i(Y^{y,i}(s)) ds\}] \\ &= \mathbb{E}[\Phi(Y^{y,i}(t_0), \alpha^{y,i}(t_0))] + \mathbb{E}\left[\mathbf{1}_{(\Omega_\varepsilon^{y,i})^c} \Phi(Y^{y,i}(t_0), \alpha^{y,i}(t_0)) \left(\exp\{-\int_0^{t_0} q_i(Y^{y,i}(s)) ds\} - 1\right)\right] \\ &\quad + \mathbb{E}\left[\mathbf{1}_{\Omega_\varepsilon^{y,i}} \Phi(Y^{y,i}(t_0), \alpha^{y,i}(t_0)) \left(\exp\{-\int_0^{t_0} q_i(Y^{y,i}(s)) ds\} - 1\right)\right]. \end{aligned} \quad (5.6)$$

Note that if  $|g(z, i)| \leq 1 \forall z \in \mathbb{R}^n, i \in \mathbb{Z}_+$  then  $|\Phi(z, i)| \leq 1 \forall z \in \mathbb{R}^n, i \in \mathbb{Z}_+$ . We have the following estimates for  $y \in B(x, \delta)$  using (5.2), (5.3), (5.4), and the fact that  $|g(z, i)|, |\Phi(z, i)| \leq 1 \forall z \in \mathbb{R}^n, i \in \mathbb{Z}_+$ .

$$\left| \mathbb{E}\left[\mathbf{1}_{(\Omega_\varepsilon^{y,i})^c} \Phi(Y^{y,i}(t_0), \alpha^{y,i}(t_0)) \left(\exp\{-\int_0^{t_0} q_i(Y^{y,i}(s)) ds\} - 1\right)\right] \right| \leq \mathbb{P}((\Omega_\varepsilon^{y,i})^c) \leq \frac{\varepsilon}{8}, \quad (5.7)$$

$$\left| \mathbb{E}\left[\mathbf{1}_{\Omega_\varepsilon^{y,i}} \Phi(Y^{y,i}(t_0), \alpha^{y,i}(t_0)) \left(\exp\{-\int_0^{t_0} q_i(Y^{y,i}(s)) ds\} - 1\right)\right] \right| \leq 1 - \exp(-Mt_0) \leq \frac{\varepsilon}{16}, \quad (5.8)$$

$$\left| \mathbb{E}\left[\mathbf{1}_{\{\tau^{y,i} \leq t_0\}} \Phi(X^{y,i}(t_0), \alpha^{y,i}(t_0))\right] \right| \leq \mathbb{P}\{\tau^{y,i} \leq t_0\} \leq \frac{3\varepsilon}{16}. \quad (5.9)$$

Applying estimates (5.4), (5.7), (5.8), and (5.9) to (5.5) and (5.6), we have

$$\left| \mathbb{E}\Phi(X^{y,i}(T), \alpha^{y,i}(T)) - \mathbb{E}\Phi(X^{x,i}(T), \alpha^{x,i}(T)) \right| \leq \varepsilon, \quad \forall y \in B(x, \delta).$$

The proof is complete.  $\square$

**Remark 5.4.** Sufficient conditions for the strong Feller property of  $(X(t), \alpha(t))$ , in which the rates of switching  $q_{ij}$  for  $i, j \in \mathbb{Z}_+$  depend only on the current continuous state  $X(t)$ , was obtained in Shao [19]. However, the conditions there are restrictive. To obtain the strong Feller property, it is assumed in [19] that  $q_{ij}(x), b(x, i)$  and  $\sigma(x, i)$  are Lipschitz in  $x$  uniformly in  $i \in \mathbb{Z}_+$ . The ellipticity of  $A(x, i)$  is also assumed to be uniform in  $(x, i) \in \mathbb{R}^n \times \mathbb{Z}_+$ . Moreover, it is assumed that  $q_{ij}(x) = 0$  if  $|i - j| < k$  for some constant  $k$ . It can be seen that our conditions in this paper are much more relaxed compared with the aforementioned reference.

## 6 Further Remarks

This paper has been devoted to modeling and analysis of switching diffusions in which past-dependent switching processes and countable switching sets are considered. Many problems remain open. Based on our results, one may consider such properties as recurrence, ergodicity, and stability. Future work may also be directed to the study of switching diffusions in which the drift and diffusions are also past dependent. It is important to work out all the details of the control problems presented in the previous sections, which may open up a new avenue for investigation of a wide range of control and optimization problems involving switching diffusions that are treated in this paper.

## A Appendix

This section is devoted to the proofs of some technical results. To simplify the notation, we denote by  $\mathbb{P}_{\phi, i}$  the probability measure conditional on the initial value  $(\phi, i)$ , that is, for any  $t > 0$ ,

$$\begin{aligned}\mathbb{P}_{\phi, i}\{(X_t, \alpha(t)) \in \cdot\} &= \mathbb{P}\{(X_t^{\phi, i}, \alpha^{\phi, i}(t)) \in \cdot\}, \\ \mathbb{P}_{\phi, i}\{(Y_t, \beta(t)) \in \cdot\} &= \mathbb{P}\{(Y_t^{\phi, i}, \beta^{\phi, i}(t)) \in \cdot\},\end{aligned}$$

and

$$\mathbb{P}_{\phi, i}\{(Z_t, \gamma(t)) \in \cdot\} = \mathbb{P}\{(Z_t^{\phi, i}, \gamma^{\phi, i}(t)) \in \cdot\}.$$

Let  $\mathbb{E}_{\phi, i}$  be the expectation associated with  $\mathbb{P}_{\phi, i}$ . First, we prove the following result.

**Lemma A.1.** *Let either Assumption 3.1 or Assumption 3.3 combined with (ii) of Assumption 3.1 be satisfied. Assume further that  $q_i(\cdot), q_{ij}(\cdot)$  are continuous functions in  $\mathcal{C}$  for each  $i, j \in \mathbb{Z}_+$ . Then the solution  $(X_t, \alpha(t))$  to (2.3) satisfies (2.1) and (2.2).*

*Proof.* It is clear that the solution  $(X_t, \alpha(t))$  to (2.3) satisfies (2.1). Fix  $\phi \in \mathcal{C}$ ,  $i, j \in \mathbb{Z}_+, i \neq j$ . Applying the generalized Itô formula to the function  $V(\psi, k) = 0$  if  $k \neq j$  and  $V(\psi, j) = 1$  we have

$$\mathbb{P}_{\phi, i}\{\alpha(\Delta) = j\} = \mathbb{E}_{\phi, i} V(X_\Delta, \alpha(\Delta)) = \mathbb{E}_{\phi, i} \int_0^\Delta q_{\alpha(t), j}(X_t) dt, \quad \text{for } \Delta > 0,$$

where  $q_{ii}(\phi) := -q_i(\phi) = -\sum_{j \neq i} q_{ij}(\phi)$ .<sup>1</sup> Since  $\alpha(t)$  is cadlag and  $X(t)$  is continuous,  $\lim_{t \rightarrow 0^+} \alpha(t) = i$  and  $\lim_{t \rightarrow 0^+} X_t = \phi$   $\mathbb{P}_{\phi, i}$ -a.s. In light of the continuity of  $q_{ij}(\cdot)$  we obtain  $\lim_{t \rightarrow 0^+} q_{\alpha(t), j}(X_t) = q_{ij}(\phi)$   $\mathbb{P}_{\phi, i}$  - a.s. which implies that

$$\lim_{\Delta \rightarrow 0^+} \frac{1}{\Delta} \int_0^\Delta q_{\alpha(t), j}(X_t) dt = q_{ij}(\phi) \mathbb{P}_{\phi, i} - \text{a.s.}$$

---

<sup>1</sup>  $q_{ii}(\phi)$  is not defined in the journal version (SICON) of this paper

Since  $q_{ij}(\cdot)$  is uniformly bounded, so is  $\frac{1}{\Delta} \int_0^\Delta q_{\alpha(t),j}(X_t)dt$ . By virtue of the Lebesgue dominated convergence theorem, we have

$$\lim_{\Delta \rightarrow 0^+} \frac{\mathbb{P}_{\phi,i}\{\alpha(\Delta) = j\}}{\delta} = \lim_{\Delta \rightarrow 0^+} \mathbb{E}_{\phi,i} \left( \frac{1}{\Delta} \int_0^\Delta q_{\alpha(t),j}(X_t)dt \right) = q_{ij}(\phi). \quad (\text{A.1})$$

In the same manner, applying the generalized Itô formula to the function  $V(\psi, k) = 1$  if  $k \neq i$  and  $V(\psi, i) = 0$ , we obtain that

$$\lim_{\Delta \rightarrow 0^+} \frac{1 - \mathbb{P}_{\phi,i}\{\alpha(\Delta) = i\}}{\delta} = q_i(\phi). \quad (\text{A.2})$$

The proof is complete by noting that (2.2) follows from (A.1) and (A.2) and the Markov property of  $(X(t), \alpha(t))$ .  $\square$

Next, we provide the proofs of some results in Section 4.

*Proof of Lemma 4.2.* To prove claim (i), we apply the generalized Itô formula to  $V(j) = 1$  if  $j = i$ , and  $V(j) = 0$  if  $j \neq i$ . We have

$$V(\beta(\lambda_1 \wedge t)) = - \int_0^{\lambda_1 \wedge t} q_i(Y_s)ds + \int_0^{\lambda_1 \wedge t} \int_{\mathbb{R}} \left( V(i + h(Y_t, i, z)) - 1 \right) \mu(ds, dz).$$

Since  $W(\cdot)$  is independent of the Poisson random measure, taking the conditional expectation with respect to  $\mathcal{F}_T^W$  yields

$$\begin{aligned} \mathbb{E}_{\phi,i} [\mathbf{1}_{\{\lambda_1 > t\}} | \mathcal{F}_T^W] &= \mathbb{E}_{\phi,i} [V(\lambda_1 \wedge t) | \mathcal{F}_T^W] = - \mathbb{E}_{\phi,i} \left[ \int_0^{\lambda_1 \wedge t} q_i(Y_s)ds | \mathcal{F}_T^W \right] + 1 \\ &= - \mathbb{E}_{\phi,i} \left[ \int_0^t q_i(Y_s)ds \mathbf{1}_{\{\lambda_1 > s\}} | \mathcal{F}_T^W \right] + 1 \\ &= - \int_0^t q_i(Y_s) \mathbb{E}_{\phi,i} [\mathbf{1}_{\{\lambda_1 > s\}} | \mathcal{F}_T^W] ds + 1. \end{aligned}$$

Hence,

$$\frac{d}{dt} \mathbb{E}_{\phi,i} [\mathbf{1}_{\{\lambda_1 > t\}} | \mathcal{F}_T^W] = -q_i(Y_t) \mathbb{E}_{\phi,i} [\mathbf{1}_{\{\lambda_1 > t\}} | \mathcal{F}_T^W].$$

Since  $\mathbb{E}_{\phi,i} [\mathbf{1}_{\{\lambda_1 > 0\}} | \mathcal{F}_T^W] = 1$ , we obtain

$$\mathbb{P}_{\phi,i} (\{\lambda_1 > t\} | \mathcal{F}_T^W) = \mathbb{E}_{\phi,i} [\mathbf{1}_{\{\lambda_1 > t\}} | \mathcal{F}_T^W] = \exp \left( - \int_0^t q_i(Y_s)ds \right). \quad (\text{A.3})$$

Now we prove claim (ii). First, we try to find the distribution of  $(\lambda_1, \beta_1)$  conditioned on  $\mathcal{F}_T^W$  when  $\lambda_1 \in [0, T]$ . Fix  $j \neq i$  and let  $f(t, k) : [0, T] \times \mathbb{Z}_+ \rightarrow \mathbb{Z}_+$  be any bounded measurable function satisfying  $f(t, k) = 0$  if  $k \neq j$ . Applying the generalized Itô formula, we obtain

$$\begin{aligned} f(\lambda_1 \wedge T, \beta(\lambda_1 \wedge T)) &= \int_0^{\lambda_1 \wedge T} q_{ij}(Y_t) f(t, j) dt \\ &\quad + \int_0^{\lambda_1 \wedge T} \int_{\mathbb{R}} \left( f(s, i + h(Y_t, i, z)) - f(Y_t, i) \right) \mu(ds, dz). \end{aligned}$$

Since  $W(\cdot)$  is independent of the Poisson random measure, taking the conditional expectation with respect to  $\mathcal{F}_T^W$ , we have

$$\begin{aligned}
\mathbb{E}_{\phi,i}[f(\lambda_1 \wedge T, \beta(\lambda_1 \wedge T)) | \mathcal{F}_T^W] &= \mathbb{E}_{\phi,i} \left[ \int_0^{\lambda_1 \wedge T} q_{ij}(Y_t) f(t, j) dt \middle| \mathcal{F}_T^W \right] \\
&= \mathbb{E}_{\phi,i} \left[ \int_0^T q_{ij}(Y_t) f(t, j) dt \mathbf{1}_{\{\lambda_1 > t\}} dt \middle| \mathcal{F}_T^W \right] \\
&= \int_0^T q_{ij}(Y_t) f(t, j) \mathbb{E}[\mathbf{1}_{\{\lambda_1 > t\}} | \mathcal{F}_T^W] dt \\
&= \int_0^T q_{ij}(Y_t) f(t, j) \exp\left(-\int_0^t q_i(Y_s) ds\right) dt.
\end{aligned}$$

As a result, for  $t \in [0, T]$ ,

$$\mathbb{P}_{\phi,i}\{\lambda_1 \in dt, \beta(\lambda_1) = j | \mathcal{F}_T^W\} = q_{ij}(Y_t) \exp\left(-\int_0^t q_i(Y_s) ds\right) dt.$$

Thus,

$$\begin{aligned}
&\mathbb{E}_{\phi,i} \left[ g(Y_{(1)}, \lambda_1, \beta_1) \mathbf{1}_{\{\lambda_1 \leq T\}} \middle| \mathcal{F}_T^W \right] \\
&= \sum_{j=1, j \neq i}^{\infty} \int_0^T g(Y_t, t, j) \mathbb{P}_{\phi,i}\{\lambda_1 \in dt, \beta(\lambda_1) = j | \mathcal{F}_T^W\} \\
&= \sum_{j=1, j \neq i}^{\infty} \int_0^T g(Y_t, t, j) q_{ij}(Y_t) \exp\left(-\int_0^t q_i(Y_s) ds\right) dt
\end{aligned}$$

as desired.  $\square$

*Proof of Proposition 4.3.* First, we prove (4.2) for the case  $l = 0$ . Since  $(X_t, \alpha(t)) = (Y_t, \lambda(t))$  up to the moment  $\alpha_1 = \lambda_1$ , we have

$$\begin{aligned}
\mathbb{E}_{\phi,i} \left[ f(X_T, \alpha(T)) \mathbf{1}_{\{\tau_1 > T\}} \right] &= \mathbb{E}_{\phi,i} \left[ f(Y_T, i) \mathbf{1}_{\{\lambda_1 > T\}} \right] \\
&= \mathbb{E}_{\phi,i} \left[ \mathbb{E}_{\phi,i} \left( f(Y_T, i) \mathbf{1}_{\{\lambda_1 > T\}} \middle| \mathcal{F}_T^W \right) \right] = \mathbb{E}_{\phi,i} \left[ f(Y_T, i) \mathbb{E}_{\phi,i}(\mathbf{1}_{\{\lambda_1 > T\}} | \mathcal{F}_T^W) \right] \quad (\text{A.4}) \\
&= \mathbb{E}_{\phi,i} \left[ f(Y_T, i) \exp\left(-\int_0^T q_i(Y_s) ds\right) \right],
\end{aligned}$$

where the last equality is consequence of (i) of Lemma 4.2. Since  $\gamma(\cdot)$  and  $Y(\cdot)$  are independent,  $Z_t = Y_t$  up to the moment  $\theta_1$  and  $\mathbb{P}_{\phi,i}\{\theta_1 > T\} = \exp(-T)$ , we obtain

$$\begin{aligned}
&\mathbb{E}_{\phi,i} \left( f(Z_T, \gamma(T)) \mathbf{1}_{\{\theta_1 > T\}} \exp\left\{-\int_0^T q_i(Z_s) ds\right\} \right) \\
&= \mathbb{E}_{\phi,i} \left( f(Y_T, i) \mathbf{1}_{\{\theta_1 > T\}} \exp\left\{-\int_0^T q_i(Y_s) ds\right\} \right) \\
&= \mathbb{P}_{\phi,i}\{\theta_1 > T\} \mathbb{E}_{\phi,i} \left( f(Y_T, i) \exp\left\{-\int_0^T q_i(Y_s) ds\right\} \right) \\
&= \exp(-T) \mathbb{E}_{\phi,i} \left( f(Y_T, i) \exp\left\{-\int_0^T q_i(Y_s) ds\right\} \right).
\end{aligned} \quad (\text{A.5})$$

From (A.4) and (A.5), we have for  $t \in [0, T]$  that

$$\mathbb{E}f(X_T, \alpha(T))\mathbf{1}_{\{\tau_1 > T\}} = \exp(T)\mathbb{E}\left(f(Z_T, \gamma(T))\mathbf{1}_{\{\theta_1 > T\}} \exp\left\{-\int_0^T q_i(Z_s)ds\right\}\right). \quad (\text{A.6})$$

We now prove (4.2) for  $l = 1$ . Let  $g(\phi, t, i), \tilde{g}(\phi, t, i) : \mathcal{C} \times [0, \infty) \times \mathbb{Z}_+ \rightarrow \mathbb{R}$  be bounded measurable functions and  $g(\phi, t, i) = \tilde{g}(\phi, t, i) = 0$  if  $t > T$ . It follows from (ii) of Lemma 4.2 that

$$\begin{aligned} \mathbb{E}_{\phi, i}g(X_{(1)}, \tau_1, \alpha_1) &= \mathbb{E}_{\phi, i}g(Y_{(1)}, \lambda_1, \beta_1) \\ &= \sum_{i_1 \neq i} \int_0^T \mathbb{E}_{\phi, i}\left(g(Y_t, t, i_1)q_{ii_1}(Y_t) \exp\left(-\int_0^t q_i(Y_s)ds\right)\right)dt. \end{aligned} \quad (\text{A.7})$$

On the other hand,

$$\begin{aligned} \mathbb{E}_{\phi, i}\left[\tilde{g}(Z_{(1)}, \theta_1, \gamma_1) \exp\left(-\int_0^{\theta_1} q_i(Z_s)ds\right)\right] \\ &= \mathbb{E}_{\phi, i}\left[\tilde{g}(Y_{(1)}, \theta_1, \gamma_t) \exp\left(-\int_0^{\theta_1} q_i(Y_s)ds\right)\right] \\ &= \sum_{i_1 \neq i} \int_0^T \mathbb{E}_{\phi, i}\left(\tilde{g}(Y_t, t, i_1) \exp\left(-\int_0^t q_i(Y_s)ds\right)\right)\mathbb{P}_{\phi, i}\{\theta_1 \in dt, \gamma_1 = i_1\} \\ &= \sum_{i_1 \neq i} \int_0^T \mathbb{E}_{\phi, i}\left(\tilde{g}(Y_t, t, i_1) \exp\left(-\int_0^t q_i(Y_s)ds\right)\right)\rho_{ii_1} \exp(-t)dt. \end{aligned} \quad (\text{A.8})$$

Substituting  $\tilde{g}(\phi, t, i) = g(\phi, t, i) \exp(t) \times \frac{q_{ii}(\phi)}{\rho_{ii}}$  into (A.8), we have

$$\begin{aligned} \mathbb{E}_{\phi, i}\left[g(Z_{(1)}, \theta_1, \gamma_1) \exp(\theta_1) \times \frac{q_{i\gamma_1}(Z_s)}{\rho_{i\gamma_1}} \exp\left(-\int_0^{\theta_1} q_i(Z_s)ds\right)\right] \\ &= \sum_{i_1 \neq i} \int_0^T \mathbb{E}_{\phi, i}\left[g(Y_t, t, i_1) \exp(t) \frac{q_{ii_1}(Y_t)}{\rho_{ii_1}} \exp\left(-\int_0^t q_i(Y_s)ds\right)\right]\rho_{ii_1} \exp(-t)dt \\ &= \sum_{i_1 \neq i} \int_0^T \mathbb{E}_{\phi, i}\left[g(Y_t, t, i_1)q_{ii_1}(Y_t) \exp\left(-\int_0^t q_i(Y_s)ds\right)\right]dt. \end{aligned} \quad (\text{A.9})$$

It follows from (A.7) and (A.9) that

$$\begin{aligned} \mathbb{P}_{\phi, i}\{\tau_1 \in dt, \alpha_1 = i_1, X_{(1)} \in d\phi_1\} \\ &= \mathbb{E}_{\phi, i}\left[\mathbf{1}_{\{\theta_1 \in dt, \gamma_1 = i_1, Z_{(1)} \in d\phi_1\}} \exp(t) \times \frac{q_{ii}(Z_t)}{\rho_{ii_1}} \exp\left(-\int_0^t q_i(Z_s)ds\right)\right]. \end{aligned} \quad (\text{A.10})$$

We now use the strong Markov property of  $(X_t, \alpha(t))$  and  $(Z_t, \gamma(t))$ , (A.10) as well as (A.6) with



$\phi, i, T$  replaced by  $\phi_1, i_1, T - t$ , respectively;

$$\begin{aligned}
& \mathbb{E}_{\phi, i} f(X_T, \alpha(T)) \mathbf{1}_{\{\tau_1 \leq T < \tau_2\}} \mathbf{1}_{\{\alpha_1 = i_1\}} \\
&= \int_0^T \int_{\mathcal{C}} \left[ \mathbb{P}\{\tau_1 \in dt, \alpha_1 = i_1, X_{(1)} \in d\phi_1\} \times \mathbb{E}_{\phi_1, i_1} f(X_{T-t}, \alpha(T-t)) \mathbf{1}_{\{\tau_1 > T-t\}} \right] \\
&= \int_0^T \int_{\mathcal{C}} \left[ \mathbb{E}_{\phi, i} \left( \mathbf{1}_{\{\theta_1 \in dt, \gamma_1 = i_1, Z_{(1)} \in d\phi_1\}} \exp(t) \times \frac{q_{ii_1}(Z_s)}{\rho_{ii_1}} \exp\left(-\int_0^t q_i(Z_s) ds\right) \right) \right. \\
&\quad \left. \times \exp(T-t) \mathbb{E}_{\phi_1, i_1} f(Z_{T-t}, \gamma(T-t)) \mathbf{1}_{\{\theta_1 > T-t\}} \exp\left\{-\int_0^{T-t} q_{i_1}(Z_s) ds\right\} \right] \quad (\text{A.11}) \\
&= \exp(T) \mathbb{E}_{\phi, i} \left[ f(Z_T, \gamma(T)) \mathbf{1}_{\{\theta_1 \leq T < \theta_2\}} \mathbf{1}_{\{\gamma_1 = i_1\}} \exp\left(-\int_{\theta_1}^T q_{i_1}(Z_s) ds\right) \right. \\
&\quad \left. \frac{q_{ii_1}(Z_{(1)})}{\rho_{ii_1}} \exp\left(-\int_0^{\theta_1} q_i(Z_s) ds\right) \right].
\end{aligned}$$

We have already proved (4.2) for  $l = 0, 1$ . Using the same argument, the induction, and the strong Markov property of  $(X_t, \alpha(t))$  and  $(Z_t, \gamma(t))$ , we can obtain (4.2) for any  $l \in \mathbb{Z}_+$ .  $\square$

*The proof of Lemma 4.5.* By (3.5), we can find  $m \in \mathbb{Z}_+$  such that

$$\mathbb{P}_{\phi, i_0} \{\tau_{m+1} < T\} < \frac{\Delta}{2}, \quad \forall (\phi, i) \in \mathcal{C} \times \mathbb{Z}_+. \quad (\text{A.12})$$

Now, let  $\varepsilon = \varepsilon(\Delta) > 0$  (to be specified later). In view of [13, Theorem 4.3, p. 61], for each  $i \in \mathbb{Z}_+$ , there is a constant  $C_i$  such that

$$\mathbb{E}_{\phi, i_0} |Y(t) - Y(s)|^6 \leq C_i |t - s|^3 \forall t, s \in [0, T], \forall \|\phi\| \leq R + 1. \quad (\text{A.13})$$

By the Kolmogorov-Centsov theorem (see [7, Theorem 2.8]), there is a positive random variable  $h_i^\phi(\omega)$  such that

$$\mathbb{P}_{\phi, i_0} \left\{ \sup_{t, s \in [0, T], 0 < t - s < h_i^\phi(\omega)} \frac{|Y^{\phi, i}(t) - Y(s)|}{(t - s)^{0.25}} \leq 4 \right\} = 1.$$

Since  $C_i$  in (A.13) does not depend on  $\phi \in \{\psi : \|\psi\| \leq R + 1\}$ , it can be seen from the proof of the Kolmogorov-Centsov theorem that for any  $\varepsilon > 0$ , there is a constant  $h_i > 0$  satisfying

$$\mathbb{P}_{\phi, i_0} \left\{ \sup_{t, s \in [0, T], 0 < t - s < h_i} \frac{|Y(t) - Y(s)|}{(s - t)^{0.25}} \leq 4 \right\} > 1 - \varepsilon, \quad \forall \|\phi\| \leq R + 1. \quad (\text{A.14})$$

Without loss of generality, we can choose  $h_{i+1} < h_i, \forall i \in \mathbb{Z}_+$ . Let

$$\mathcal{H}_{i, T} = \left\{ \psi(\cdot) \in \mathcal{C}([0, T], \mathbb{R}) : \|\psi\| \leq R + 1 \text{ and } \sup_{t, s \in [0, T], 0 < t - s < h_i} \frac{|\psi(s) - \psi(t)|}{(s - t)^{0.25}} \leq 4 \right\},$$

and

$$\mathcal{H}_i = \left\{ \psi(\cdot) \in \mathcal{C} : \|\psi\| \leq R + 1 \text{ and } \sup_{t, s \in [-r, 0], 0 < t - s < h_i} \frac{|\psi(s) - \psi(t)|}{(s - t)^{0.25}} \leq 4 \right\}.$$

Hence  $\mathcal{H}_{i+1, T} \supset \mathcal{H}_{i, T}$  and  $\mathcal{H}_{i+1} \supset \mathcal{H}_i$ . For  $d > 0$  and a compact set  $\mathcal{K} \subset \mathcal{C}$ , we define

$$\mathcal{K}_d := \{\psi \in \mathcal{C} : \exists \phi \in \mathcal{K} \text{ such that } \|\psi - \phi\| < d\}.$$

Define  $\mathcal{K}^0 = \{\psi(\cdot) = \phi_0(\cdot) + c : c \in \mathbb{R}^n, |c| \leq 1\}$ , which is compact, and  $\mathcal{K}^1 = \mathcal{K}^0 \uplus \mathcal{H}_{i_0}$ . For each  $\phi \in \mathcal{K}^1$ , there is  $n_{\phi, i_0} > i_0$  such that

$$\sum_{k=n_{\phi, i_0}+1}^{\infty} q_{i_0, k}(\phi) = q_{i_0}(\phi) - \sum_{k=1, k \neq i_0}^{n_{\phi, i_0}} q_{i_0 k}(\phi) < \frac{\varepsilon}{2}.$$

By the continuous of  $q_{i_0}$  and  $q_{i_0 k}(\phi)$ , there is a  $d_{\phi, i_0} > 0$  such that

$$\sum_{k=n(\phi)+1}^{\infty} q_{i_0, k}(\phi') = q_{i_0}(\phi') - \sum_{k=1, k \neq i_0}^{n_{\phi, i_0}} q_{i_0 k}(\phi') < \varepsilon \forall \|\phi' - \phi\| < d_{\phi, i_0}.$$

Since  $\mathcal{K}^1$  is compact, there exist  $n_1 > 0$  and  $d_1 > 0$  such that

$$\sum_{k=n_1+1}^{\infty} q_{i_0, k}(\phi) < \varepsilon \forall \phi \in \mathcal{K}_{d_1}^1.$$

Define  $\mathcal{K}^2 = \mathcal{K}^1 \uplus \mathcal{H}_{n_1}$ . Using the compactness of  $\mathcal{K}^2$ , there exist  $n_2 > n_1$  and  $d_2 \in (0, d_1]$  such that

$$\sum_{k=n_2+1}^{\infty} q_{i, k}(\phi) < \varepsilon \forall i \in N_{n_1}, \phi \in \mathcal{K}_{d_2}^2.$$

Continuing this way, for  $\mathcal{K}^m = \mathcal{K}^{m-1} \uplus \mathcal{H}_{n_{m-1}}$ , there exists  $n_m > n_{m-1}$  and  $d_m \in (0, d_{m-1}]$  such that

$$\sum_{k=n_m+1}^{\infty} q_{i, k}(\phi) < \varepsilon \forall i \in N_{n_{m-1}}, \phi \in \mathcal{K}_{d_m}^m.$$

Set  $\mathcal{K}^{\phi, 1} = \{\phi\} \uplus \mathcal{H}_{i_0}$  and  $\mathcal{K}^{\phi, k} = \mathcal{K}^{\phi, k-1} \uplus \mathcal{H}_{n_{k-1}}$  for  $\phi \in \mathcal{C}$  and  $k = 2, \dots, m$ . It is not difficult to verify that

$$\mathcal{K}^{\phi, k} \subset \mathcal{K}_{d_k}^k \quad \forall k = 1, \dots, m, \text{ for } \|\phi - \phi_0\| < \frac{d_m}{2}. \quad (\text{A.15})$$

Denote by  $\{Y(\cdot) \in \mathcal{H}_{n_0, T}\}$  the event  $\{t \in [0, T] \mapsto Y(t) \text{ is a function belonging to } \mathcal{H}_{n_0, T}\}$ . Clearly, if  $Y(\cdot) \in \mathcal{H}_{i_0, T}$ , then  $Y_t \in \mathcal{K}^{\phi, 1} \forall t \in [0, T]$ . Thus, we can proceed as follows:

$$\begin{aligned} & \mathbb{P}_{\phi, i_0} \{ \tau_1 \leq T, (X_{(1)}, \alpha_1) \notin \mathcal{K}^{\phi, 1} \times N_{n_1} \} \\ &= \mathbb{P}_{\phi, i_0} \{ \{ \tau_1 \leq T, \alpha_1 > n_1 \} \cup \{ \tau_1 \leq T, X(\tau_1) \notin \mathcal{K}^{\phi, 1} \} \} \\ &= \mathbb{P}_{\phi, i_0} \{ \{ \lambda_1 \leq T, \beta_1 > n_1 \} \cup \{ \lambda_1 \leq T, Y(\lambda_1) \notin \mathcal{K}^{\phi, 1} \} \} \\ &\leq \mathbb{P}_{\phi, i_0} \{ \lambda_1 \leq T, Y(\cdot) \in \mathcal{H}_{n_0, T}, \beta_1 > n_1 \} + \mathbb{P} \{ Y(\cdot) \notin \mathcal{H}_{n_0}^T \} \\ &\leq \mathbb{E}_{\phi, i_0} \left[ \mathbb{E} \left( \mathbf{1}_{\{Y(\cdot) \in \mathcal{H}_{n_0, T}\}} \mathbf{1}_{\{\lambda_1 \leq T, \beta_1 > n_1\}} \mid \mathcal{F}_T^W \right) \right] + \varepsilon \\ &= \mathbb{E}_{\phi, i_0} \left[ \mathbf{1}_{\{Y(\cdot) \in \mathcal{H}_{n_0, T}\}} \mathbb{E} \left( \mathbf{1}_{\{\lambda_1 \leq T, \beta_1 > n_1\}} \mid \mathcal{F}_T^W \right) \right] + \varepsilon \\ &= \mathbb{E}_{\phi, i_0} \left[ \mathbf{1}_{\{Y(\cdot) \in \mathcal{H}_{n_0, T}\}} \int_0^T \sum_{i > n_1} q_{i_0, i}(Y_t) \exp \left( - \int_0^t q_{i_0}(Y_s) ds \right) dt \right] + \varepsilon \\ &\leq \mathbb{E}_{\phi, i_0} \left[ \mathbf{1}_{\{Y(\cdot) \in \mathcal{H}_{n_0, T}\}} \int_0^T \varepsilon dt \right] + \varepsilon \leq (T + 1)\varepsilon. \end{aligned} \quad (\text{A.16})$$

Similarly, if  $(\phi_1, i_1) \in \mathcal{K}^{\phi,1} \times N_1$ , then  $\mathbb{P}_{\phi_1, i_1} \{ \tau_1 \leq T, (X_{(1)}, \alpha_1) \notin \mathcal{K}^{\phi,2} \times N_{n_2} \} \leq (T+1)\varepsilon$ . Using the strong Markov property of  $(X_t, \alpha(t))$ , we obtain

$$\begin{aligned} & \mathbb{P}_{\phi, i_0} \left\{ \tau_1 < T, (X_{(1)}, \alpha_1) \in \mathcal{K}^{\phi,1} \times N_{n_1}, \tau_2 \leq T, (X_{(2)}, \alpha_2) \notin \mathcal{K}^{\phi,2} \times N_{n_2} \right\} \\ & \leq \mathbb{P}_{\phi, i_0} \left\{ \tau_1 < T, (X_{(1)}, \alpha_1) \in \mathcal{K}^{\phi,1} \times N_{n_1} \right\} \\ & \quad \times \mathbb{P}_{\phi, i_0} \left[ \left\{ \tau_2 \leq T + \tau_1, (X_{(2)}, \alpha_2) \notin \mathcal{K}^{\phi,2} \times N_{n_2} \right\} \middle| \tau_1 < T, (X_{(1)}, \alpha_1) \in \mathcal{K}^{\phi,1} \times N_{n_1} \right] \\ & \leq \sup_{(\phi_1, i_1) \in \mathcal{K}_1^{\phi} \times N_{n_1}} \mathbb{P}_{\phi_1, i_1} \left\{ \tau_1 \leq T, (X_{(1)}, \alpha_1) \notin \mathcal{K}^{\phi,2} \times N_{n_2} \right\} \leq (T+1)\varepsilon. \end{aligned}$$

Continuing this way, we can show for any  $k = 1, \dots, m$  that

$$\mathbb{P}_{\phi, i_0} \left\{ \tau_k \leq T, (X_{\tau_k}, \alpha_k) \notin \mathcal{K}_k^{\phi} \times N_{n_k}, (X_{(j)}, \alpha_j) \in \mathcal{K}_j^{\phi} \times N_{n_j}, j = 1, \dots, k-1 \right\} \leq (T+1)\varepsilon. \quad (\text{A.17})$$

Consequently,

$$\mathbb{P}_{\phi, i_0} \left\{ \exists k = 1, \dots, m : \tau_k \leq T \text{ and } (X_{(k)}, \alpha_k) \notin \mathcal{K}_k^{\phi} \times N_{n_k} \right\} \leq (T+1)m\varepsilon.$$

Hence, if we choose  $\varepsilon = \frac{1}{2m(T+1)}\Delta$ ,

$$\begin{aligned} & \mathbb{P}_{\phi, i_0} \left\{ \forall k = 1, \dots, m : \tau_k > T \text{ or } \alpha_k \in N_{n_k} \right\} \\ & \geq \mathbb{P} \left\{ \forall k = 1, \dots, m : \tau_k > T \text{ or } (X_{(k)}, \alpha_k) \in \mathcal{K}_k^{\phi} \times N_{n_k} \right\} \geq 1 - \frac{\Delta}{2}. \end{aligned} \quad (\text{A.18})$$

It follows from (A.12) and (A.18) that

$$\mathbb{P}_{\phi, i_0} \left( \{ \tau_{m+1} > T \} \cap \left\{ \forall k = 1, \dots, m : \tau_k > T \text{ or } \alpha_k \in N_{n_k} \right\} \right) \geq 1 - \Delta.$$

It is easily verified that if  $\omega \in \{ \tau_{m+1} > T \} \cap \{ \forall k = 1, \dots, m : \tau_k > T \text{ or } \alpha_k \in N_{n_k} \}$ , then  $\alpha(t) \in N_{n_m}, \forall t \in [0, T]$ . The assertion of the lemma is proved.  $\square$

## References

- [1] D. Applebaum, *Lévy Processes and Stochastic Calculus*, Cambridge University Press, 2009.
- [2] W.H. Fleming and R.W. Rishel, *Deterministic and Stochastic Optimal Control*, Springer-Verlag, New York, NY, 1975.
- [3] I.I. Gihman, A.V. Skorohod, *Stochastic Differential Equations*, Springer, New York, 1979.
- [4] J. Huang, H. Zhang, and J. Zhang, A unified approach to diffusion analysis of queues with general patience-time distributions, to appear in *Math. Oper. Res.*
- [5] S.-B. Hsu, T.-W. Hwang and Y. Kuang, A ratio-dependent food chain model and its applications to biological control, *Math. Biosc.*, **181**(2003), 55-83
- [6] N. Ikeda, S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, second edition, North-Holland Publishing Co., Amsterdam, 1989.
- [7] I. Karatzas, S.E. Shreve, *Brownian Motion and Stochastic Calculus*, Springer.
- [8] K. Kawaguchi and H. Morimoto, Long-run average welfare in a pollution accumulation model, *J. Econom. Dyn. Control*, **31** (2007), 703-720.

- [9] E. Keeler, M. Spence, and R. Zeckhauser, The optimal control of pollution, *J. Economic Theory* **4** (1972), 19-34.
- [10] R. Z. Khasminskii, G. Yin, C. Zhu, Stability of regime-switching diffusions, *Stochastic Process. Appl.* **117** (2007), no. 8, 1037-1051.
- [11] H.J. Kushner, *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Birkhäuser, Boston, 1990.
- [12] R.F. Luck, Evaluation of natural enemies for biological control: a behavior approach, *Trends Ecol. Evol.* **5** (1990) 196-199
- [13] X. Mao, *Stochastic Differential Equations and Applications*, 2nd Ed., Horwood, Chichester, 2008.
- [14] X. Mao, C. Yuan. *Stochastic Differential Equations with Markovian Switching*, Imperial College Press, London, 2006.
- [15] B. Øksendal, *Stochastic Differential Equations*, Springer, Berlin, 2003.
- [16] D. Nualart, *The Malliavin calculus and related topics* **1995** (2006), Springer, Berlin.
- [17] P. Protter, *Stochastic Integration and Differential Equations*, Springer, New York, 1990.
- [18] A.V. Skorokhod, *Asymptotic Methods in the Theory of Stochastic Differential Equations*, Vol. 78. American Mathematical Soc., 1989.
- [19] J. Shao, Strong solutions and strong Feller properties for regime-switching diffusion processes in an infinite state space *SIAM J. Control Optim.* **53** (2015), no. 4, 2462-2479.
- [20] J. Shao, F. Xi, Strong ergodicity of the regime-switching diffusion processes. *Stochastic Process. Appl.* **123** (2013), no. 11, 3903-3918.
- [21] J. Shao, F. Xi, Stability and recurrence of regime-switching diffusion processes, *SIAM J. Control Optim.* **52** (2014), no. 6, 3496-3516.
- [22] R. Situ, *Theory of Stochastic Differential Equations with Jumps and Applications: Mathematical and Analytical Techniques with Applications to Engineering*, Springer, New York, 2005.
- [23] D. W. Stroock, and S. S. Varadhan. *Multidimensional diffusion processes*, Springer, 2007.
- [24] W. Whitt, *Stochastic-Process Limits*, Springer, New York, 2002.
- [25] G. Yin, C. Zhu, *Hybrid Switching Diffusions: Properties and Applications*, Springer, 2010.
- [26] G. Yin, C. Zhu, Properties of solutions of stochastic differential equations with continuous-state-dependent switching, *J. Differential Equations* **249** (2010), no. 10, 2409-2439.
- [27] C. Zhu, G. Yin, Asymptotic properties of hybrid diffusion systems, *SIAM J. Control Optim.* **46** (2007), no. 4, 1155-1179.
- [28] C. Zhu, G. Yin, On strong Feller, recurrence, and weak stabilization of regime-switching diffusions, *SIAM J. Control Optim.* **48** (2009), no. 3, 2003-2031.