May 2021

Regime-Switching Jump Diffusion Processes with Countable Regimes: Feller, Strong Feller, Irreducibility and Exponential Ergodicity

Khwanchai Kunwai

University of Wisconsin-Milwaukee

REGIME-SWITCHING JUMP DIFFUSION PROCESSES WITH COUNTABLE REGIMES: FELLER, STRONG FELLER, IRREDUCIBILITY AND EXPONENTIAL ERGODICITY

by

Khwanchai Kunwai

A Dissertation Submitted in
Partial Fulfillment of the
Requirements of the Degree of

Doctor of Philosophy
in Mathematics

at
The University of Wisconsin-Milwaukee
May 2021

ABSTRACT

REGIME-SWITCHING JUMP DIFFUSION PROCESSES WITH COUNTABLE REGIMES: FELLER, STRONG FELLER, IRREDUCIBILITY AND EXPONENTIAL ERGODICITY

by

Khwanchai Kunwai

The University of Wisconsin-Milwaukee, 2021
Under the Supervision of Professor Chao Zhu

This work is devoted to the study of regime-switching jump diffusion processes in which the switching component has countably infinite regimes. Such processes can be used to model complex hybrid systems in which both structural changes, small fluctuations as well as big spikes coexist and are intertwined. Weak sufficient conditions for Feller and strong Feller properties and irreducibility for such processes are derived; which further lead to Foster-Lyapunov drift conditions for exponential ergodicity. Our results can be applied to stochastic differential equations with non-Lipschitz coefficients. Finally, an application to feedback control problems is presented.

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

First and foremost, I want to thank my adviser, Professor Chao Zhu, for invaluable guidance, advice and encouragement he offered during my study. Without Professor Chao Zhu's support this research would not be possible.

I would like to express my deepest thanks to Professors Richard Stockbridge, Suzanne Boyd, Wei Wei and Jeb Willenbring for serving on my doctoral defense committee and proofreading of my dissertation.

Finally, I would like to thank my parents for their love and support.

Khwanchai Kunwai

Chapter 1

Introduction

1.1 Motivation and Overview

A stochastic model is a tool used to estimate potential outcomes where randomness or uncertainty is presented. Traditionally, stochastic models are constructed based on the continuous dynamics of the outcomes over time. Let's consider the following example of a risky asset. Suppose the stock price S_t satisfies the stochastic differential equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

where μ represents the expected rate of return and σ denotes the market's volatility. Because these parameters are assumed to be constant, this model is good only for a relatively short period of time. Moreover, the movement of an individual stock can be subject to the general trends of the market; bull or bear market. That is; if the the overall market moves up, most stocks go up and if the overall market goes down, most stocks follow. The market trend has a bigger effect on the volatility of the individual stock. To incorporate the broad trend of the stock market we need more realistic and more sophisticated models to capture such complex evolution.

A regime-switching diffusion is a two-component stochastic process $(X(t), \Lambda(t))$ that contains an analog (or continuous state) component $X(t)$ and a switching (or discrete event) component $\Lambda(t)$. Intuitively, the continuous component $X(t)$ represents the state of phenomenon while the discrete component $\Lambda(t)$ refers to the structural changes of the system. In the past decades, this class of processes has received growing attention due to its ability of modeling and analyzing complex systems in which both structural changes and small fluctuations coexist and are intertwined. These processes are widely used in many areas such as economy, financial engineering, risk management, biology, engineering, and etc. For instance, a regime-switching Black-Scholes model is considered in Zhang (2001), in which the continuous component $X(t)$ models the price evolution of a risky asset and the switching component $\Lambda(t)$ delineates the overall economy state. For another example, Ferrari and Rodosthenous (2019) study the optimal control of the debt-to-GDP ratio where $X(t)$ models the level of debt-to-GDP ratio while $\Lambda(t)$ represents the state of macroeconomic conditions. Regime-switching (jump) diffusions are also used in mathematical biology such as the recent paper from Tuong et al. (2019) in which a stochastic SIRS model subject to both white and color noises is analyzed.

Compared to the classical setting when there is no switching component, much care is needed for regime-switching diffusions. This is due to the interactions between the continuous and the switching components. For example, one can combine two stable (or unstable) diffusions to produce an unstable (or stable) regime-switching diffusion; see, for example, Yin and Zhu (2010) and Lawley et al. (2014). In general, a regime-switching diffusion can possess a certain property even though some of its regimes do not. This feature is, of course, not possible in the usual setting. One of the most important problems of great interest is the asymptotic property of such processes. Taking this consideration into account, we will study existence, uniqueness and the rate of convergence of an invariant measure.

Regime-switching diffusion processes with a finite switching state space have been relatively well studied. We refer to Mao and Yuan (2006) and Yin and Zhu (2010) for extensive discussions and applications on this class of processes. To be more precise, Mao and Yuan

(2006) study state-independent regime-switching diffusions; that is, the switching component is a continuous-time Markov chain independent of the continuous component. On the other hand, Yin and Zhu (2010) focus on state-dependent regime-switching diffusions when the switching component is a stochastic process taking values in a finite set and depend on the continuous component. More investigations on this vein can be found in Cloez and Hairer (2015), Nguyen et al. (2017), Shao (2015a), Shao and Xi (2014), Xi (2004, 2009), Zhu and Yin (2009) and references therein. When the switching state space is infinite, this adds more subtlety and difficulty to the analyses as we need to deal with infinite regimes. Moreover, this makes the interactions between the continuous and discrete components much more complicated. Recent developments on the countably infinite case can be found in Nguyen and Yin (2018a,b,c), Shao (2015a,b), Xi and Zhu (2017), Xi et al. (2019).

Regime-switching diffusions become more interesting and challenging when jumps are brought into consideration. This enhances the ability to model more complex and realistic phenomena. We refer to Applebaum (2009) for extensive discussions on jump processes. Continuing on the effort of studying regime-switching diffusions with countable regimes, Xi and Zhu (2017) studied regime-switching diffusions with jumps and derived Feller and strong Feller properties as well as exponential ergodicity of such processes. This paper treats when the coefficients of the associated stochastic differential equations are (locally) Lipschitz. While it is a convenient assumption, it is rather restrictive in many applications. For example, the diffusion coefficients in the Feller branching diffusion and the Cox-Ingersoll-Ross model are only Hölder continuous. For another example, many control and optimization problems often require the handling of systems where the local Lipschitz condition is violated. Recently, Xi et al. (2019) presented non-Lipschitz conditions for existence and uniqueness of nonexplosive strong solution as well as for Feller and strong Feller properties for regime-switching jump diffusion processes. However, with the broad settings in the later paper, the ergodicity has not been investigated yet. So we want to fulfill the gap here as well as improve the results on Feller and strong Feller properties in Xi et al. (2019). Moreover, the irreducibility of theses processes remains in question.

The series of papers Meyn and Tweedie (1992, 1993b,c) provide a powerful criterion for exponential ergodicity of Markov processes. The criterion relies on the existence of a Foster-Lyapunov function and the property that all compact subsets are small in some sense for some skeleton chain. In view of Theorem 3.4 of Meyn and Tweedie (1992), the smallness assumption can be verified by establishing Feller property and φ-irreducibility. Then the existence of a Foster-Lyapunov function becomes the key to establishing exponential ergodicity. Indeed, this type of function plays a vital role in the study of stability and long-term behaviors of stochastic systems; see, for example, Hairer et al. (2011), Khasminskii (2012), Mao and Yuan (2006), Yin and Zhu (2010), to name just a few. However, in practice, it is usually not easy to find such functions. For regime-switching (jump) diffusions with infinite number of switching states, it is even harder to find an appropriate Foster-Lyapunov function due to the interactions between the continuous and the discrete components. In this work we provide weak sufficient conditions for such functions to exist.

In this dissertation, we study state-dependent regime-switching jump diffusion processes with countable regimes and non-Lipschitz coefficients. The contributions of this dissertation can be summarized as follows:

(i) Weak sufficient conditions for Feller and strong Feller properties, (open set) irreducibility, and exponential ergodicity are presented in terms of the coefficients of the associated SDEs and the transition rate matrix of the discrete component Λ.

(ii) Topological and probabilistic concepts of irreducibility are discussed for regime-switching jump diffusions; namely, open set irreducibility and φ-irreducibility.

(iii) Sufficient conditions for the existence of a Foster-Lyapunov function for regime-switching jump diffusions are presented.

(iv) As a result, we can answer three important questions on existence, uniqueness and the rate of convergence to an invariant measure for regime-switching jump diffusions with countable regimes and non-Lipschitz coefficients.

(v) An application to feedback control problems is presented for the demonstration.

This work is organized as follows. In Chapter 2 we review some classical results on stability of Markov processes studied in Meyn and Tweedie (1992, 1993b,c). Then we give the formulation of regime-switching jump diffusion processes. Moreover, we discuss coupling methods for regime-switching jump diffusions. In Chapter 3, Feller and strong Feller properties are investigated by using the coupling methods. Weak sufficient conditions for Feller and strong Feller property are imposed on the coefficients of the associated stochastic differential equations and spelled out in Theorems 3.1.6 and 3.2.8, respectively. In Chapter 4, open set irreducibility is investigated. As an application, we present in Proposition 4.1.12 a set of sufficient conditions under which a unique invariant measure for regime-switching jump diffusions exists. After establishing Feller and strong Feller properties and irreducibility, we derive in Chapter 5 the exponential ergodicity of regime-switching jump diffusions. We construct a Foster-Lyapunov function in Theorem 5.2.4 by incorporating some nice properties of the generator $Q(x)$ of the discrete component Λ. As a result, we can conclude that a regime-switching jump diffusion process can be exponentially ergodic even though some subsystems are not; see Remark 5.2.3. Finally, an application to feedback control problems is present in Chapter 6. For the sake of completeness, some elementary computations are given in Chapter 7.

Chapter 2

Preliminaries and Classical Results

2.1 Mathematical Background

2.1.1 Markov Processes

In this section we briefly review the basic notions and terminology of Markov theory. For the classical discussions we refer to Blumenthal and Getoor (1968), Ethier and Kurtz (1986), Sharpe (1988), and others. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space and let $(E, \mathcal{B}(E))$ be a measurable space where E is a locally compact separable metric space and $\mathcal{B}(E)$ is the Borel σ-algebra of open subsets of E. Let $\Phi := \{\Phi_t : 0 \leq t < \infty\}$ be a continuous time stochastic process defined on $(\Omega, \mathscr{F}, \mathbb{P})$ and take values in the set E. For each $t \geq 0$ we define the sub-σ-algebra $\mathscr{F}_t^{\Phi} = \sigma(\Phi_s : s \leq t)$ of $\mathscr{F}$. Then Φ is called a *Markov process* if

$$\mathbb{P}\{\Phi_{t+s} \in A | \mathscr{F}_t^{\Phi}\} = \mathbb{P}\{\Phi_{t+s} \in A | \Phi_t\} \tag{2.1}$$

for all $s, t \geq 0$ and $A \in \mathcal{B}(E)$. Condition (2.1) is called *Markov property* or *memoryless property*. If $\{\mathscr{F}_t\}_{t \geq 0}$ is a filtration with $\mathscr{F}_t^{\Phi} \subset \mathscr{F}_t$, $t \geq 0$, then Φ is a Markov process with respect to $\{\mathscr{F}_t\}_{t \geq 0}$ if (2.1) holds with $\mathscr{F}_t^{\Phi}$ replaced by $\mathscr{F}_t$. It is clear that if Φ is a Markov process with respect to $\{\mathscr{F}_t\}_{t\geq 0}$, then it is a Markov process. Note that (2.1) implies

$$\mathbb{E}\left[f(\Phi_{t+s}) \in A|\mathscr{F}_t^{\Phi}\right] = \mathbb{E}\left[f(\Phi_{t+s}) \in A|\Phi_t\right]$$

for all $s, t \geq 0$ and bounded and measurable function f on E.

A function $P(t, x, A)$ defined on $[0, \infty) \times E \times \mathcal{B}(E)$ is a *time-homogeneous transition function* if

- $P(t, x, \cdot)$ is a probability measure on $\mathcal{B}(E)$ for all $(t, x) \in [0, \infty) \times E$,
- $P(0, x, \cdot) = \delta_x$ for all $x \in E$ where δ_x is the unit mass at x,
- $P(\cdot, \cdot, A)$ is bounded and measurable function on $[0, \infty) \times E$ for all $A \in \mathcal{B}(E)$,
- for all $s, t \geq 0$, $x \in E$, $A \in \mathcal{B}(E)$

$$P(t+s, x, A) = \int P(s, y, A)P(t, x, dy). \tag{2.2}$$

The relationship (2.2) is called *Chapman-Kolmogorov equation.* A transition function $P(t, x, A)$ is a *transition function for a time homogeneous Markov process* Φ if

$$\mathbb{P}\{\Phi_{t+s} \in A|\mathscr{F}_t^{\Phi}\} = P(s, \Phi_t, A) \tag{2.3}$$

for all $s, t \geq 0$, $A \in \mathcal{B}(E)$, or equivalently, if

$$\mathbb{E}\left[f(\Phi_{t+s}) \in A|\mathscr{F}_t^{\Phi}\right] = \int f(y)P(s, \Phi_t, dy) \tag{2.4}$$

for all $s, t \geq 0$ and bounded and measurable function f on E. The probability measure ν on $\mathcal{B}(E)$ defined by $\nu(A) = \mathbb{P}\{\Phi_0 \in A\}$ is called the *initial distribution* of Φ.

Suppose Φ is a Markov process with respect to a filtration $\{\mathscr{F}_t\}_{t\geq 0}$ such that Φ is $\{\mathscr{F}_t\}$-progressive in the sense that the restriction of Φ to $[0, t] \times \Omega$ is $\mathcal{B}([0, t]) \times \mathscr{F}_t$-measurable for any $t \geq 0$. Then Φ is called *strong Markov process* with respect to $\{\mathscr{F}_t\}_{t\geq 0}$ if, for any $\{\mathscr{F}_t\}$-stopping time τ with $\tau < \infty$ a.s.,

$$\mathbb{P}\{\Phi_{t+\tau} \in A | \mathscr{F}_\tau\} \quad = \quad P(t, \Phi_\tau, A) \tag{2.5}$$

for all $t \geq 0$, $A \in \mathcal{B}(E)$, or equivalently, if

$$\mathbb{E}\left[f(\Phi_{t+\tau}) \in A | \mathscr{F}_\tau\right] \quad = \quad \int f(y) P(t, \Phi_\tau, dy) \tag{2.6}$$

for all $t \geq 0$ and bounded and measurable function f on E. Here we denote $\mathscr{F}_\tau = \{A \in \mathscr{F} : \forall t \geq 0, \{\tau \leq t\} \cap A \in \mathscr{F}_t\}$.

Define the operator P_t which acts on a set of bounded functions f by

$$P_t f(x) := \mathbb{E}\left[f(\Phi_t) | \Phi_0 = x\right] = \int_E f(y) P(t, x, dy).$$

It follows from the Chapman-Kolmogorov property that the family $\{P_t : t \geq 0\}$ is a *semigroup* in the sense that $P_{t+s} = P_t \circ P_s$ for all $s, t \geq 0$. For a σ-finite measure μ on $\mathcal{B}(E)$ we define

$$\mu P_t(A) := \int_E P(t, x, A) \mu(dx), \quad \forall A \in \mathcal{B}(E),\ t \geq 0.$$

Next we define one of the most important tools in Markov theory. The *infinitesimal generator* $\mathscr{A}$ of Φ is defined by

$$\mathscr{A} f(x) := \lim_{t \longrightarrow 0} \frac{P_t f(x) - f(x)}{t} \tag{2.7}$$

provided that the limit exists. This generator measures the expected change of the function f under the dynamic of the Markov process Φ in an infinitesimal time interval.

2.1.2 Feller and Strong Feller Properties

The intuition behind the concepts of Feller and strong Feller Properties is that a slight perturbation of the initial data should result in a small perturbation in the subsequent movement. This is a natural condition in physical or social modeling. In addition, Feller and strong Feller properties are intrinsically related to the existence and uniqueness of invariant measures of the underlying process; see, for example, Meyn and Tweedie (1992, 1993b,c).

We denote by $\mathfrak{B}_b(E)$ the class of bounded and Borel measurable functions on E and by $\mathfrak{C}_b(E)$ the class of bounded and continuous functions on E. Following Dynkin (1965), let us state the following definition.

Definition 2.1.1. *The semigroup $\{P_t : t \geq 0\}$ or the process Φ is said to have Feller property if $\lim_{t \longrightarrow 0} P_t f(x) = f(x)$ for all $f \in \mathfrak{C}_b(E)$ and*

$$P_t : \mathfrak{C}_b(E) \longrightarrow \mathfrak{C}_b(E) \quad \textit{for all} \quad t \geq 0$$

and strong Feller property if

$$P_t : \mathfrak{B}_b(E) \longrightarrow \mathfrak{C}_b(E) \quad \textit{for all} \quad t > 0.$$

2.1.3 Irreducibility

Consider a continuous time Markov process $\Phi := \{\Phi_t : 0 \leq t < \infty\}$ taking values in the state space $(E, \mathcal{B}(E))$ with initial distribution $\nu(A) = \mathbb{P}\{\Phi_0 \in A\}$. The fundamental concept of irreducibility is the idea that all parts of E can be reached by the Markov process. As discussed in Tweedie (1994), connections between topological and probabilistic properties involve irreducibility.

For any measurable set $A \in \mathcal{B}(E)$ we define the *hitting time* and the *occupation time* by

$$\tau_A := \inf\{t \geq 0 : \Phi_t \in A\} \quad \text{and} \quad \eta_A := \int_0^\infty 1_A(\Phi_t) dt,$$

respectively. Topological concepts of irreducibility are defined in terms of the ability of the process to reach open sets, rather than sets of positive measure. To be more precise, Φ is said to be *open set irreducible* if for every point $y \in E$ and any open set A containing y we have

$$\mathbb{P}_x\{\tau_A < \infty\} > 0 \quad \text{for all} \quad x \in E.$$

In this case the point y is called *reachable*. Hence Φ is open set irreducible if every point is reachable.

The process Φ is said to be φ*-irreducible* if there exists a σ-finite measure φ on $\mathcal{B}(E)$ such that for any $x \in E$ we have

$$\varphi(A) > 0 \Longrightarrow \mathbb{E}_x\left[\eta_A\right] > 0.$$

The measure φ is called an *irreducible measure*. This is a concept of probabilistic irreducibility studied in the literature; see, for example, Meyn and Tweedie (1992, 1993b,c), Tweedie (1994).

The connection between these two concepts of irreducibility was studied in Tweedie (1994) for the class of T-process. It was shown in Theorem 3.2 of Tweedie (1994) that any open set irreducible T-process is φ-irreducible. In particular, every open set irreducible Markov chain with strong Feller property is φ-irreducible; see Proposition 6.1.6 of Meyn and Tweedie (2009).

2.1.4 Petite Sets

The concept of petite sets is a generalization of the concept of small sets. It was studied in Meyn and Tweedie (1992) to develop criteria for stability of Markov processes. For a probability distribution a on $\mathbb{R}_+$ we define by

$$K_a(x, A) := \int_{\mathbb{R}_+} a(dt)P(t, x, A)$$

the *transition kernel corresponding to a.* A set $B \in \mathcal{B}(E)$ and a non-trivial sub-probability measure μ on $\mathcal{B}(E)$ are called *petite* if for some probability distribution a we have

$$\mu(\cdot) \leq K_a(x, \cdot) \quad \forall x \in B.$$

In the literature, we sometimes use the term μ_a*-petite* to indicate the sub-probability measure μ and the probability distribution a. This class of petite sets plays a crucial role in the study of stability of Markov processes in both discrete and continuous time cases; see, Meyn and Tweedie (1992, 1993b,c). Furthermore, for discrete time Markov processes with Feller property we have the following result.

Theorem 2.1.2 (Theorem 3.4, Meyn and Tweedie (1992))**.** *Suppose that* $\Phi := \{\Phi_n : n = 0, 1, 2, ...\}$ *is* φ*-irreducible. Then either of the following conditions implies that all compact subsets of* E *are petite:*

(i) Φ *has the Feller property and there exists an open* φ*-positive petite set;*

(ii) Φ *has the Feller property and the support of* φ *has nonempty interior.*

2.1.5 Invariant Measures and Ergodic Theory

To study a Markov process we usually raise questions concerning the existence and uniqueness of an invariant measure. A σ-finite measure π on $\mathcal{B}(E)$ is called *invariant* if

$$\pi(A) = \pi P_t(A) := \int_E P(t, x, A)\pi(dx), \quad \forall A \in \mathcal{B}(E),\ t \geq 0. \tag{2.8}$$

It is called *subinvariant* if we have $\pi \geq \pi P$. It is well known that in finite state space case, that is when E is finite, an invariant measure always exists. However, for infinite state space case, there may be no invariant measures.

Let us recall the concept of Harris recurrence. A Markov process Φ is called *Harris recurrent* if one of the following condition holds:

- there exists a σ-finite measure μ such that

$$\mu(A) > 0 \Longrightarrow \mathbb{P}_x\{\eta_A = \infty\} = 1;$$

- there exists a σ-finite measure φ such that

$$\varphi(A) > 0 \Longrightarrow \mathbb{P}_x\{\tau_A < \infty\} = 1.$$

These conditions are equivalent; see Theorem 1.1 of Meyn and Tweedie (1993a). It is clear from the definition that Harris-recurrent processes are φ-irreducible. It is well known that if Φ is φ-irreducible then a subinvariant measure exists. Moreover, if Φ is Harris recurrent then a unique (up to constant multiples) invariant measure π exists; see, for example, Azéma and Revuz (1991) and Getoor (1979). If the invariant measure π is finite then Φ is called *positive Harris recurrence*. In this case π can be normalized to be a probability measure. We conclude that

$$\text{positive Harris recurrence} \Rightarrow \text{Harris recurrence} \Rightarrow \varphi\text{-irreducibility}.$$

A criterion for positive Harris recurrence was given in Meyn and Tweedie (1993c) in terms of a Foster-Lyapunov drift condition.

Assumption 2.1.3. *There exist constants $c, d > 0$, a function $f \geq 1$, a measurable set*

$C \subset E$, and a function $V \geq 0$ such that

$$\mathscr{A}V(x) \leq -cf(x) + d1_C(x) \tag{2.9}$$

for all $x \in E$.

The function V in (2.9) is called a *Foster-Lyapunov function.* We denote $\pi(f) := \int_E f(x)\pi(dx)$. The following result can be proved by applying the Dynkin's formula to the process $V(\Phi_t)$ above.

Theorem 2.1.4 (Theorem 4.2, Meyn and Tweedie (1993c))**.** *Suppose that Φ is a non-explosive right process. If* (2.9) *holds for a closed petite set $C \subset E$ and V is bounded on C then Φ is positive Harris recurrence and $\pi(f) < \infty$.*

The following result shows that, for Feller processes, the existence of an invariant measure can be obtained without assuming irreducibility.

Theorem 2.1.5 (Theorem 4.5, Meyn and Tweedie (1993c))**.** *Suppose that Φ is a non-explosive right process with Feller property. If* (2.9) *holds for some compact set $C \subset E$, then an invariant probability measure exists, and $\pi(f) \leq d/c$ for any invariant probability π.*

To study long-term behavior of Markov processes, we are interested in the rate of convergence towards invariant measures. To this end, let us give some notations. For any measurable function $f : E \longrightarrow [1, \infty)$ and any signed-measure μ on $\mathcal{B}(E)$, we set

$$||\mu||_f := \sup\{|\mu(g)| : \text{all measurable } g \text{ with } |g| \leq f\}$$

where $\mu(g) := \int_E g(y)\mu(dy)$. We note that the total variation norm $||\mu||_{TV}$ is the special case of $||\mu||_f$ when $f \equiv 1$.

Following the terminology in Meyn and Tweedie (1993c) we say that the process Φ is *ergodic* if an invariant probability measure π exists and

$$\lim_{t\to\infty} ||P(t,x,\cdot) - \pi(\cdot)||_{TV} = 0, \quad \text{for all} \quad x \in E.$$

Moreover, Φ is said to be *f-exponentially ergodic* if there exist a probability measure $\pi(\cdot)$, a constant $\theta \in (0,1)$ and a finite-valued function $\Theta(x)$ such that for all $t \geq 0$ and all $x \in E$ we have

$$||P(t,x,\cdot) - \pi(\cdot)||_f \;\; \leq \;\; \Theta(x)\theta^t.$$

Remark 2.1.6. *In general, the concepts of positive Harris recurrence and ergodicity are not equivalent. If Φ is ergodic then every h-skeleton chain $\{\Phi_{nh} : n = 0, 1, ...\}$ is also ergodic. On the other hand, positive Harris recurrence processes do not have this property. A counter example is a clock process; see Meyn and Tweedie (1993b).*

The following theorem is comparable to Theorem 2.1.4.

Theorem 2.1.7 (Theorem 5.1, Meyn and Tweedie (1993c))**.** *Suppose that Φ is a non-explosive right process and that all compact sets are petite for some h-skeleton chain. If* (2.9) *holds for some compact set C with V bounded on C, then Φ is ergodic.*

To obtain criterion for exponential ergodicity we need a stronger version of condition (2.9). Let us state the following assumption.

Assumption 2.1.8. *There exist constants $c > 0, d < \infty$ and a function $V \geq 0$ such that*

$$\lim_{||x||\to\infty} V(x) = \infty \text{ (norm-like) and } \quad \mathscr{A}V(x) \leq -cV(x) + d \quad \forall x \in E. \tag{2.10}$$

Theorem 2.1.9 (Theorem 6.1, Meyn and Tweedie (1993c))**.** *Suppose that Φ is a right process, and that all compact sets are petite for some h-skeleton chain $\{\Phi_{nh} : n \in \mathbb{Z}_+\}$. If* (2.10) *holds, then Φ is f-exponentially ergodic with $f := 1 + V$.*

We summarize the discussions in section 2.1 in Figure 2.1. For Feller processes, the existence of an appropriate Foster-Lyapunov function becomes the key for deriving the existence and uniqueness as well as the convergence rate of invariant measures. In view of Theorem 2.1.5, the existence of an invariant measure π is guaranteed by condition (2.9). However, the uniqueness and the convergence rate remain questionable. To determine the uniqueness and the convergence rate we need the notion of petite sets. Thanks to Theorem 2.1.2, this can be done if we can verify that the process is φ-irreducibility and if the measure φ has nonempty interior. From Theorem 2.1.7 we obtain the convergence in total variation norm. Furthermore, Theorem 2.1.9 gives the exponential rate of convergence.

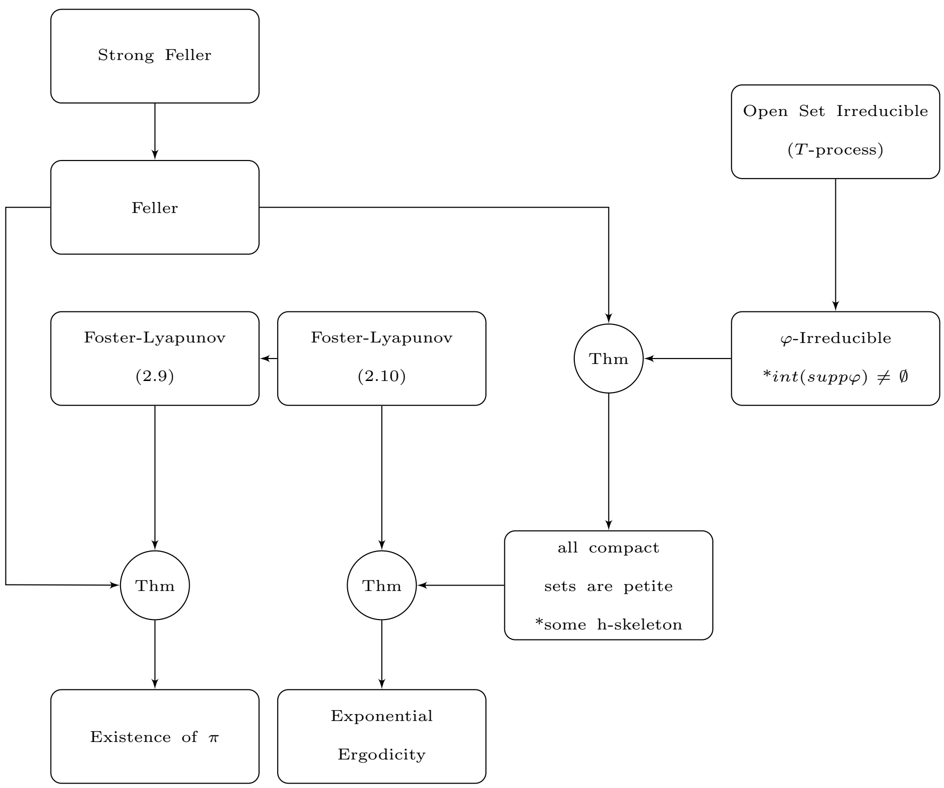

Figure 2.1: Overview of Ergodic Theory

2.2 Regime-Switching Jump Diffusion Processes

2.2.1 Formulation

Let $d \geq 1$ be an integer and $\mathbb{S} = \{1, 2, ...\}$ be the switching state. Throughout this dissertation, we consider an arbitrary filtered measure space $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t\geq 0}, \mathbb{P})$ rich enough to accommodate a standard d-dimensional Brownian motion W. We construct a regime-switching jump diffusion as a continuous-time stochastic process $(X, \Lambda) := \{(X(t), \Lambda(t))\}_{t\geq 0}$ where $(X(t), \Lambda(t)) : \Omega \to \mathbb{R}^d \times \mathbb{S}$ is a random variable for all $t \geq 0$. Let $(U, \mathfrak{U})$ be a measurable space and ν a σ-finite measure on U. Assume that $b : \mathbb{R}^d \times \mathbb{S} \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \times \mathbb{S} \to \mathbb{R}^{d\times d}$ and $c : \mathbb{R}^d \times \mathbb{S} \times U \to \mathbb{R}^d$ are Borel measurable functions. Suppose that (X, Λ) is a right continuous strong Markov process with left-hand limits on $\mathbb{R}^d \times \mathbb{S}$ such that the continuous component X satisfies the following stochastic differential equation (SDE),

$$dX(t) = b(X(t), \Lambda(t))dt + \sigma(X(t), \Lambda(t))dW(t) + \int_U c(X(t^-), \Lambda(t^-), u)\tilde{N}(dt, du), \qquad (2.11)$$

where N is a Poisson random measure on $[0, \infty) \times U$ with intensity $dt\,\nu(du)$ and $\tilde{N}$ is the associated compensated Poisson random measure. As in Nguyen and Yin (2018a,c), Xi and Zhu (2017), and Xi et al. (2019) we suppose that the discrete component Λ is a continuous-time stochastic process taking values in the set $\mathbb{S}$ and generated by the transition rate matrix $Q(x) := (q_{kl}(x))_{k,l\in\mathbb{S}}$. That is Λ satisfies

$$\mathbb{P}\{\Lambda(t+\Delta) = l | \Lambda(t) = k, X(t) = x\} = \begin{cases} q_{kl}(x)\Delta + o(\Delta) & \text{if } k \neq l \\ 1 + q_{kl}(x)\Delta + o(\Delta) & \text{if } k = l, \end{cases} \qquad (2.12)$$

for all $x \in \mathbb{R}^d$. We suppose further that $Q(x)$ is stable and conservative in the sense that for any $x \in \mathbb{R}^d$

$$0 \quad \leq \quad q_{kl}(x) < +\infty \qquad \text{for} \quad k \neq l$$

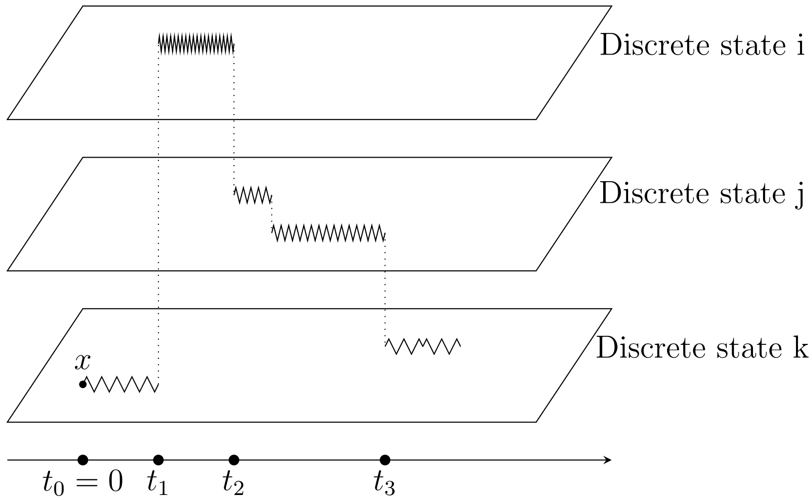

Figure 2.2: Sample Path of Regime-Switching Jump Diffusion

$$\begin{aligned} q_k(x) &:= -q_{kk}(x) < +\infty && \text{(stable)} \\ q_k(x) &= \sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl}(x) && \text{(conservative).} \end{aligned}$$

To obtain the structure of the process Λ, let consider the family of disjoint intervals $\{\Delta_{kl}(x) : k, l \in \mathbb{S}\}$ defined on the positive half of the real line as follows:

$$\begin{aligned} \Delta_{12}(x) &= [0, q_{12}(x)), \\ \Delta_{13}(x) &= [q_{12}(x), q_{12}(x) + q_{13}(x)), \\ &\vdots \\ \Delta_{21}(x) &= [q_1(x), q_1(x) + q_{21}(x)), \\ \Delta_{23}(x) &= [q_1(x) + q_{21}(x), q_1(x) + q_{21}(x) + q_{23}(x)), \\ &\vdots \\ \Delta_{31}(x) &= [q_1(x) + q_2(x), q_1(x) + q_2(x) + q_{31}(x)), \\ &\vdots \end{aligned}$$

where $q_k(x) := \sum_{l \in \mathbb{S} \setminus \{k\}} q_{kl}(x)$ and we set $\Delta_{kl}(x) = \emptyset$ in the case of $q_{kl}(x) = 0$, $k \neq l$. We note that $\{\Delta_{kl}(x) : k, l \in \mathbb{S}\}$ are disjoint intervals and that the length of the interval $\Delta_{kl}(x)$

is equal to $q_{kl}(x)$. Define a function $h : \mathbb{R}^d \times \mathbb{S} \times \mathbb{R}_+ \to \mathbb{R}$ by

$$h(x,k,r) = \sum_{l\in\mathbb{S}} (l-k) 1_{\Delta_{kl}(x)}(r). \tag{2.13}$$

In other words, for each $x \in \mathbb{R}^d$ and $k \in \mathbb{S}$, we set

$$h(x,k,r) = \begin{cases} l-k & \text{if } r \in \Delta_{kl}(x) \\ 0 & \text{otherwise.} \end{cases}$$

As a result, the process $\varLambda$ can be described as a solution to the following stochastic differential equation

$$\varLambda(t) = \varLambda(0) + \int_0^t \int_{\mathbb{R}_+} h(X(s^-), \varLambda(s^-), r) N_1(ds, dr), \tag{2.14}$$

where N_1 is a Poisson random measure on $[0,\infty)\times[0,\infty)$ with characteristic measure $\mathfrak{m}(dz)$, the Lebesgue measure. We refer to Xi et al. (2019) for the existence and uniqueness of non-explosive strong solution to the system of SDEs (2.11) and (2.14).

Denote by $C_c^2(\mathbb{R}^d)$ the set of twice continuously differentiable functions with compact support defined on $\mathbb{R}^d$. Given a function $f : \mathbb{R}^d \times \mathbb{S} \to \mathbb{R}$ with $f(\cdot, k) \in C_c^2(\mathbb{R}^d)$ for each $k \in \mathbb{S}$. The infinitesimal generator of the regime-switching jump diffusion $(X, \varLambda)$ is given by

$$\mathscr{A} f(x,k) := \mathscr{A}_d f(x,k) + \mathscr{A}_j f(x,k) + Q(x) f(x,k), \tag{2.15}$$

where

$$\mathscr{A}_d f(x,k) := \frac{1}{2}\operatorname{tr}\left(a(x,k)\nabla^2 f(x,k)\right) + \langle b(x,k), \nabla f(x,k)\rangle, \tag{2.16}$$

$$\mathscr{A}_j f(x,k) := \int_U \left(f(x+c(x,k,u),k) - f(x,k) - \langle \nabla f(x,k), c(x,k,u)\rangle\right)\nu(du), \tag{2.17}$$

$$Q(x) f(x,k) := \sum_{l\in\mathbb{S}} q_{kl}(x)\left[f(x,l) - f(x,k)\right]. \tag{2.18}$$

In (2.16)–(2.17) and hereafter, we denote by $\nabla f(x,k) := (\frac{\partial}{\partial x_1} f(x,k), ..., \frac{\partial}{\partial x_d} f(x,k))^T$ the gradient and by $\nabla^2 f(x,k) := [\frac{\partial^2}{\partial x_i \partial x_j} f(x,k)]_{i,j}$ the Hessian matrix of f with respect to x. Moreover, we denote by $\langle\ ,\ \rangle$ the ordinary inner product in $\mathbb{R}^d$. For any square matrix A we denote by $tr(A)$ the trace and by A^T the transpose of A, respectively. Then the Hilbert–Schmidt norm of A is given by $|A| := \sqrt{tr(AA^T)}$.

We finish off this section by presenting some interesting examples of regime-switching diffusion processes and their applications in mathematical finance, economy and public health.

Example 2.2.1. *The regime-switching Black-Scholes model is studied in Zhang (2001). Let $\alpha(\cdot)$ be a Markov chain taking values in a finite set $\{1, 2, ..., N\}$ and generated by the stable and conservative transition rate matrix $Q = (q_{kl})_{1\leq k,l\leq N}$. We regard $\alpha(\cdot)$ as the market-trend indicator process. We denote by $S(t)$ the stock's price at time $t \geq 0$. Suppose that $S(t)$ satisfies*

$$\begin{cases} dS(t) = S(t)[\mu(\alpha(t))dt + \sigma(\alpha(t))dW(t)] \\ S(0) = S_0 \end{cases}$$

where $S_0 > 0$ is the initial price, $\mu(\cdot)$ is the expected return, $\sigma(\cdot)$ is the stock volatility, and W is a one dimensional standard Brownian motion. Here we suppose that α is independent of the Brownian motion W.

Example 2.2.2. *Consider the following debt-to-GDP ratio model studied in Ferrari and Rodosthenous (2019). Suppose W is a one dimensional standard Brownian motion and Λ is a continuous-time Markov chain taking values in a finite set $\{1, 2, ..., N\}$. Suppose further that Λ is independent of W, irreducible and generated by a stable and conservative transition rate matrix $Q := (q_{kl})_{1\leq k,l\leq N}$.*

Assume that the nominal debt D_t grows at time $t \geq 0$ at rate $r + \lambda_{\Lambda_t}$ and satisfies

$$dD_t = (r + \lambda_{\Lambda_t})D_t dt$$

where $r > 0$ is fixed interest rate on government debt and λ_{Λ_t} is the additional interest rate the government has to pay when the macroeconomic conditions are in state $\Lambda_t \in \{1, 2, ..., N\}$. Assume that the country's GDP Ψ_t evolves the stochastic differential equation

$$d\Psi_t = g\Psi_t dt + \sigma\Psi_t dW_t$$

where g is the growth rate of the GDP. Then the Debt-to-GDP ratio $X := D/\Psi$ satisfies the stochastic differential equation

$$dX_t = (r + \lambda_{\Lambda_t} - g + \sigma^2)X_t dt + \sigma X_t dW_t.$$

One can show that the solution to this stochastic differential equation is given by

$$X_t^{x,i} = x \exp^{(r-g+\frac{1}{2}\sigma^2)t+\int_0^t \lambda_{\Lambda_t^i} ds+\sigma W_t}, \quad X_0 = x, \Lambda_0 = i.$$

Example 2.2.3. *The following regime switching SIRS (Susceptible-Infected-Removed-Susceptible) model is studied in Tuong et al. (2019). This is an epidemiological model which classifies individuals into compartments of susceptible, infectious, removed with permanent acquired immunity and susceptible due to the loss of immunity of the removed individuals. Let $\xi = \{\xi_t : t \geq 0\}$ be a right continuous Markov chain taking values in $\{1, 2, ..., m\}$. Consider the model*

$$\begin{cases} dS(t) = [-S(t)I(t)F_1(S(t), I(t), \xi_t) + \mu(\xi_t)(K - S(t)) + \gamma_1(\xi_t)R(t)]dt \\ \qquad\qquad - S(t)I(t)F_2(S(t), I(t), \xi_t)dW(t) \\ dI(t) = [S(t)I(t)F_1(S(t), I(t), \xi_t) - (\mu(\xi_t) + \rho(\xi_t) + \gamma_2(\xi_t))I(t)]dt \\ \qquad\qquad + S(t)I(t)F_2(S(t), I(t), \xi_t)dW(t) \\ dR(t) = [\gamma_2(\xi_t)I(t) - (\mu(\xi_t) + \gamma_1(\xi_t))R(t)]dt \end{cases}$$

where W is a one dimensional standard Brownian motion, F_1, F_2 are positive and locally Lipschitz functions on $[0,\infty)^2 \times \{1, 2..., m\}$ which represent incidence rates. The constant K is a carrying capacity and the parameters $\mu, \rho, \gamma_1, \gamma_2$ are the per capita disease-free death rate, the excess per capita natural death rate of infective class, the per capita loss immunity and return to the susceptible class of infective class, and the per capita recovery rate of the infected individuals, respectively.

2.3 Coupling Methods

2.3.1 Classical Constructions of Couplings

Coupling method is a very powerful tool used to compare two stochastic processes. We will use this method to derive Feller and Strong Feller properties of the regime-switching jump diffusion process (X, Λ). We refer to Lindvall (1992) and Chen and Li (1989) for extensive discussion on coupling method.

Given two processes $\Phi^1 = \{\Phi_t^1 : 0 \leq t < \infty\}$ and $\Phi^2 = \{\Phi_t^2, : 0 \leq t < \infty\}$ taking values in $(E_1, \mathcal{B}_1(E))$ and $(E_2, \mathcal{B}_2(E))$, respectively. Let $\mathbb{P}_1$ and $\mathbb{P}_2$ denote distributions for Φ^1 and Φ^2, respectively. A process $(\Gamma^1, \Gamma^2) := \{(\Gamma_t^1, \Gamma_t^2) : 0 \leq t < \infty\}$ valued in $(E_1 \times E_2, \mathcal{B}_1(E) \otimes \mathcal{B}_2(E))$ with distribution $\tilde{\mathbb{P}}$ is called a *coupling of* Φ^1 *and* Φ^2 if

$$\tilde{\mathbb{P}}\{(\Gamma_t^1, \Gamma_t^2) \in B_1 \times E_2\} = \mathbb{P}_1\{\Phi_t^1 \in B_1\} \quad \text{and} \quad \tilde{\mathbb{P}}\{(\Gamma_t^1, \Gamma_t^2) \in E_1 \times B_2\} = \mathbb{P}_2\{\Phi_t^2 \in B_2\}$$

for all $t \geq 0$ and $B_1 \in \mathcal{B}_1(E), B_2 \in \mathcal{B}_2(E)$. That is, $\tilde{\mathbb{P}}$ has marginals $\mathbb{P}_1$ and $\mathbb{P}_2$. It is worth noting that a coupling is not unique.

Let Φ^1 and Φ^2 be strong Markov processes defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and take values in a state space $(E, \mathcal{B}(E))$. Suppose that (Γ^1, Γ^2) is a coupling process for Φ^1

and Φ^2. A stopping T is called a *coupling time* if

$$\Gamma_t^1 = \Gamma_t^2 \quad \forall t \geq T.$$

If T is a coupling time then we obtain the following *coupling inequality*

$$||\mathbb{P}\{\Phi_t^2 \in \cdot\} - \mathbb{P}\{\Phi_t^1 \in \cdot\}||_{TV} := \sup_{A \in \mathscr{F}} |\mathbb{P}\{\Phi_t^2 \in A\} - \mathbb{P}\{\Phi_t^1 \in A\}| \leq 2\mathbb{P}\{t < T\}. \tag{2.19}$$

Indeed, for any $A \in \mathscr{F}$ and $t \geq 0$ we have

$$\begin{aligned}
& |\mathbb{P}\{\Phi_t^2 \in A\} - \mathbb{P}\{\Phi_t^1 \in A\}| \\
= & \ | \big(\mathbb{P}\{\Phi_t^2 \in A, t < T\} + \mathbb{P}\{\Phi_t^2 \in A, t \geq T\}\big) - \big(\mathbb{P}\{\Phi_t^1 \in A, t < T\} + \mathbb{P}\{\Phi_t^1 \in A, t \geq T\}\big) | \\
= & \ | \big(\mathbb{P}\{\Gamma_t^2 \in A, t < T\} + \mathbb{P}\{\Gamma_t^2 \in A, t \geq T\}\big) - \big(\mathbb{P}\{\Gamma_t^1 \in A, t < T\} + \mathbb{P}\{\Gamma_t^1 \in A, t \geq T\}\big) | \\
= & \ | \big(\mathbb{P}\{\Gamma_t^2 \in A, t < T\} + \mathbb{P}\{\Gamma_t^2 \in A, t \geq T\}\big) - \big(\mathbb{P}\{\Gamma_t^1 \in A, t < T\} + \mathbb{P}\{\Gamma_t^2 \in A, t \geq T\}\big) | \\
= & \ |\mathbb{P}\{\Gamma_t^2 \in A, t < T\} - \mathbb{P}\{\Gamma_t^1 \in A, t < T\}| \\
\leq & \ 2\mathbb{P}\{t < T\}.
\end{aligned}$$

By taking supremum over $A \in \mathscr{F}$ we obtain (2.19).

Consider a diffusion process $X := \{X_t : 0 \leq t < \infty\}$ in $\mathbb{R}^d$ satisfying the SDE

$$dX_t = b(X_t)dt + \sigma(X_t)dW_t$$

where $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d \times d}$ are Borel measurable functions and W is a d-dimensional Brownian motion. Assume further that X is a strong Markov process. For $x, y \in \mathbb{R}^d$ we denote by X^x and X^y the processes X started at x and y, respectively. Then we can define a coupling process for X^x and X^y as follows. Define the *coupling time* T by

$$T := \inf\{t \geq 0 : X_t^x = X_t^y\}$$

and

$$X'_t = \begin{cases} X^x_t & \text{if } t < T \\ X^y_t & \text{if } t \geq T. \end{cases} \quad (2.20)$$

Then (X', X^y) is a coupling process for X^x and X^y. This construction is known as *basic*

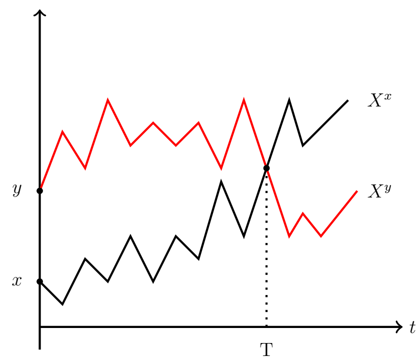

Figure 2.3: Basic Coupling

coupling. Let $a(x) := \sigma(x)\sigma(x)^T$. We denote the generator of X by

$$\mathcal{L}f(x) = \frac{1}{2}\mathrm{tr}\big(a(x)\nabla^2 f(x)\big) + \langle b(x), \nabla f(x)\rangle$$

where $\nabla f(x) := (\frac{\partial}{\partial x_1} f(x), ..., \frac{\partial}{\partial x_d} f(x))^T$ the gradient and by $\nabla^2 f(x) := [\frac{\partial^2}{\partial x_i \partial x_j} f(x)]_{i,j}$ the Hessian matrix of f with respect to x. Then the coefficients of the basic coupling operator are given by

$$a(x,y) = \begin{pmatrix} a(x) & \sigma(x)\sigma(y)^T \\ \sigma(y)\sigma(x)^T & a(y) \end{pmatrix}, \quad b(x,y) = \begin{pmatrix} b(x) \\ b(y) \end{pmatrix}.$$

To be more precise, the generator for the basic coupling is given by

$$\mathcal{L}f(x,y) = \frac{1}{2}\mathrm{tr}\big(a(x,y)\nabla^2 f(x,y)\big) + \langle b(x,y), \nabla f(x,y)\rangle. \quad (2.21)$$

Another useful coupling technique is *coupling by reflection*. We refer to Chen (2004), Chen and Li (1989), Lindvall and Rogers (1986) for the discussions on this method. As in Lindvall and Rogers (1986), given a diffusion process X satisfying the SDE

$$dX_t \;=\; b(X_t)dt + \sigma(X_t)dW_t.$$

The idea is to construct another diffusion process Y with the same generator as X but started at $y \neq x$ and solve the SDE

$$dY_t \;=\; b(Y_t)dt + \sigma(Y_t)dW_t'$$

where $dW_t' = g(x,y)dW_t$ and $g(x,y) = I - 2(x-y)(x-y)^T/|x-y|^2$. The generator for the coupling by reflection is the same as in (2.21) but with the following coefficients

$$a(x,y) = \begin{pmatrix} a(x) & g(x,y) \\ g(x,y)^T & a(y) \end{pmatrix}, \quad b(x,y) = \begin{pmatrix} b(x) \\ b(y) \end{pmatrix}.$$

2.3.2 Coupling Methods for Regime-switching Jump Diffusions

Consider regime-switching diffusion (X, Λ) of the system of SDEs (2.11) and (2.14) and two distinct initial conditions (x,i) and (y,j). We use the constructions discussed in the previous section to construct a coupling process for $(X^{(x,i)}, \Lambda^{(x,i)})$ and $(X^{(y,j)}, \Lambda^{(y,j)})$.

We use basic coupling to derive Feller property. To this end, let us first construct a basic coupling operator $\widetilde{\mathscr{A}}$ for $\mathscr{A}$. For $f(x,i,y,j) \in C_c^2(\mathbb{R}^d \times \mathbb{S} \times \mathbb{R}^d \times \mathbb{S})$, we define

$$\widetilde{\mathscr{A}} f(x,i,y,j) := \big[\widetilde{\Omega}_{\mathrm{d}} + \widetilde{\Omega}_{\mathrm{j}} + \widetilde{\Omega}_{\mathrm{s}}\big] f(x,i,y,j), \tag{2.22}$$

where $\widetilde{\Omega}_{\mathrm{d}}$, $\widetilde{\Omega}_{\mathrm{j}}$, and $\widetilde{\Omega}_{\mathrm{s}}$ are defined as follows. For $x, y \in \mathbb{R}^d$ and $i, j \in \mathbb{S}$, we set $a(x,i) =$

$\sigma(x,i)\sigma(x,i)^T$ and

$$a(x,i,y,j) = \begin{pmatrix} a(x,i) & \sigma(x,i)\sigma(y,j)^T \\ \sigma(y,j)\sigma(x,i)^T & a(y,j) \end{pmatrix}, \quad b(x,i,y,j) = \begin{pmatrix} b(x,i) \\ b(y,j) \end{pmatrix}.$$

Then we define

$$\widetilde{\Omega}_{\mathrm{d}} f(x,i,y,j) := \frac{1}{2}\mathrm{tr}\big(a(x,i,y,j)D^2 f(x,i,y,j)\big) + \langle b(x,i,y,j), Df(x,i,y,j)\rangle, \tag{2.23}$$

$$\begin{aligned} \widetilde{\Omega}_{\mathrm{j}} f(x,i,y,j) := \int_U \big[& f(x + c(x,i,u), i, z + c(y,j,u), j) - f(x,i,y,j) \\ & - \langle D_x f(x,i,y,j), c(x,i,u)\rangle - \langle D_y f(x,i,y,j), c(y,j,u)\rangle\big]\nu(du), \end{aligned} \tag{2.24}$$

where $Df(x,i,y,j) = (D_x f(x,i,y,j), D_y f(x,i,y,j))^T$ is the gradient and $D^2 f(x,i,y,j)$ is the Hessian matrix of f with respect to the variables x and y, and

$$\begin{aligned} \widetilde{\Omega}_{\mathrm{s}} f(x,i,y,j) := & \sum_{l\in\mathbb{S}} [q_{il}(x) - q_{jl}(y)]^+ (f(x,l,y,j) - f(x,i,y,j)) \\ & + \sum_{l\in\mathbb{S}} [q_{jl}(y) - q_{il}(x)]^+ (f(x,i,y,l) - f(x,i,y,j)) \\ & + \sum_{l\in\mathbb{S}} [q_{il}(x) \wedge q_{jl}(y)](f(x,l,y,l) - f(x,i,y,j)). \end{aligned} \tag{2.25}$$

For any function $f : \mathbb{R}^d \times \mathbb{R}^d \mapsto \mathbb{R}$, let $\widetilde{f} : \mathbb{R}^d \times \mathbb{S} \times \mathbb{R}^d \times \mathbb{S} \mapsto \mathbb{R}$ be defined by $\widetilde{f}(x,i,y,j) := f(x,y)$. Now we denote for each $k \in \mathbb{S}$

$$\widetilde{\mathcal{L}}_k f(x,y) = (\widetilde{\Omega}_{\mathrm{d}}^{(k)} + \widetilde{\Omega}_{\mathrm{j}}^{(k)}) f(x,y) := (\widetilde{\Omega}_{\mathrm{d}} + \widetilde{\Omega}_{\mathrm{j}})\widetilde{f}(x,k,y,k), \quad \forall f \in C_c^2(\mathbb{R}^d \times \mathbb{R}^d).$$

To derive the strong Feller property we need another coupling method. Motivated by Priola and Wang (2006) we construct the operator $\widehat{\mathscr{A}}$ as follows. Let $\lambda_R > 0$ be a constant depending on $R > 0$ and σ_{λ_R} be the unique symmetric nonnegative definite matrix-valued function such that $\sigma^2_{\lambda_R}(x,k) = a(x,k) - \lambda_R I$. This will be determined more in Assumption

(3.2.2); in particular, see condition (3.28). To this end, we let

$$\widehat{a}(x,i,y,j) := \begin{pmatrix} a(x,i) & \hat{g}(x,i,y,j) \\ \hat{g}(x,i,y,j)^T & a(y,j) \end{pmatrix} \quad \text{and} \quad b(x,i,y,j) := \begin{pmatrix} b(x,i) \\ b(y,j) \end{pmatrix}$$

where

$$\hat{g}(x,i,y,j) := \lambda_R(I - 2u(x,y)u(x,y)^T) + \sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j)^T,$$

and $u(x,y) := \frac{x-y}{|x-y|}$. Then the coupling operator $\widehat{\mathscr{A}}$ for (2.15) is given by

$$\widehat{\mathscr{A}} f(x,i,y,j) := [\widehat{\Omega}_{\mathrm{d}} + \tilde{\Omega}_{\mathrm{j}} + \tilde{\Omega}_{\mathrm{s}}] f(x,i,y,j), \quad f \in C_c^2(\mathbb{R}^d \times \mathbb{S} \times \mathbb{R}^d \times \mathbb{S}), \tag{2.26}$$

where

$$\widehat{\Omega}_{\mathrm{d}} f(x,i,y,j) = \frac{1}{2}\mathrm{tr}\left(\widehat{a}(x,i,y,j) D^2 f(x,i,y,j)\right) + \langle b(x,i,y,j), Df(x,i,y,j)\rangle, \tag{2.27}$$

and $\tilde{\Omega}_{\mathrm{j}}$ and $\tilde{\Omega}_{\mathrm{s}}$ are defined as in (2.24) and (2.25), respectively. In addition, for each $k \in \mathbb{S}$ and any $F \in C_c^2(\mathbb{R}^d \times \mathbb{R}^d)$, we write $f(x,k,y,k) := F(x,y)$ and denote

$$\widehat{\mathcal{L}}_k F(x,y) = [\widehat{\Omega}_{\mathrm{d}}^{(k)} + \tilde{\Omega}_{\mathrm{j}}^{(k)}] f(x,k,y,k) := \widehat{\mathscr{A}} f(x,k,y,k). \tag{2.28}$$

Remark 2.3.1. *We note that if $\sigma_{\lambda_R} \equiv 0$ then $\widehat{\mathscr{A}}$ is the generator of the coupling by reflection as discussed in the previous section.*

Chapter 3

Feller and Strong Feller Properties

We denote by $\mathfrak{B}_b(\mathbb{R}^d \times \mathbb{S})$ the set of all bounded and Borel measurable functions on $\mathbb{R}^d \times \mathbb{S}$ and by $\mathfrak{C}_b(\mathbb{R}^d \times \mathbb{S})$ the set of all bounded and continuous functions on $\mathbb{R}^d \times \mathbb{S}$. Suppose $(X, \Lambda) := \{(X(t), \Lambda(t) : t \geq 0\}$ is a solution to the system (2.11) and (2.14). For any $t \geq 0, x \in \mathbb{R}^d$ and $k \in \mathbb{S}$, define the operator

$$P_t f(x,k) := \mathbb{E}_{x,k}\left[f(X(t), \Lambda(t))\right] = \mathbb{E}\left[f(X(t), \Lambda(t)) | (X(0), \Lambda(0)) = (x,k)\right] \tag{3.1}$$

for $f \in \mathfrak{B}_b(\mathbb{R}^d \times \mathbb{S})$. Note also that $|P_t f(x,k)| = |\mathbb{E}_{x,k}\left[P_t f(x,k)\right]| \leq \mathbb{E}_{x,k}\left[|P_t f(x,k)|\right] \leq ||f||_\infty$ and hence $||P_t f||_\infty \leq ||f||_\infty$. In other words, P_t is a bounded operator $P_t : \mathfrak{B}_b(\mathbb{R}^d \times \mathbb{S}) \to \mathfrak{B}_b(\mathbb{R}^d \times \mathbb{S})$ and the family $\{P_t\}_{t \geq 0}$ forms a semigroup of bounded linear operators on $\mathfrak{B}_b(\mathbb{R}^d \times \mathbb{S})$. The semigroup $\{P_t\}_{t \geq 0}$ or the corresponding process (X, Λ) is said to have *Feller property* if $P_t : \mathfrak{C}_b(\mathbb{R}^d \times \mathbb{S}) \to \mathfrak{C}_b(\mathbb{R}^d \times \mathbb{S})$ and $\lim_{t \downarrow 0} P_t f(x,k) = f(x,k)$ for all $f \in C_b(\mathbb{R}^d \times \mathbb{S})$ and $(x,k) \in \mathbb{R}^d \times \mathbb{S}$ and *strong Feller property* if $P_t : \mathfrak{B}_b(\mathbb{R}^d \times \mathbb{S}) \to \mathfrak{C}_b(\mathbb{R}^d \times \mathbb{S})$ for all $t > 0$.

In Xi et al. (2019), the authors study multidimensional regime-switching jump diffusion processes with non-Lipschitz coefficients and countably many switching states. The authors imposed Assumptions 2.1 and 2.2 in order to obtain the existence and the uniqueness of a non-explosive strong solution to the system of SDEs (2.11) and (2.14). However, to derive Feller property of such processes, the authors further imposed other assumptions. So, our goal aims to show that the Assumptions 2.1 and 2.2 imposed in Xi et al. (2019) are indeed sufficient to obtain the Feller property of regime-switching jump diffusion processes.

To study Feller and strong Feller properties as well as irreducibility it is sufficient to consider weak solutions of the system (2.11) and (2.14) instead of the strong ones. We make the following standing assumption throughout this work:

Assumption 3.0.2. *For any* $(x,k) \in \mathbb{R}^d \times \mathbb{S}$, *the system of stochastic differential equations* (2.11) *and* (2.14) *has a non-explosive weak solution* $(X^{(x,k)}, \Lambda^{(x,k)})$ *with initial condition* (x,k) *and the solution is unique in the sense of probability law.*

3.1 Feller Property

In this section, we derive Feller property of (X, Λ). To begin let us state the following assumptions.

Assumption 3.1.1. *Assume the following conditions hold.*

(i) If $d = 1$, *then there exist a positive number* δ_0 *and a nondecreasing and concave function* $\rho : [0,\infty) \to [0,\infty)$ *satisfying*

$$\int_{0+} \frac{dr}{\rho(r)} = \infty, \tag{3.2}$$

such that for all $k \in \mathbb{S}, R > 0$ *and* $x, z \in \mathbb{R}$ *with* $|x| \vee |z| \le R$ *and* $|x - z| \le \delta_0$,

$$\operatorname{sgn}(x-z)(b(x,k) - b(z,k)) \le \kappa_R \rho(|x-z|), \tag{3.3}$$

$$|\sigma(x,k) - \sigma(z,k)|^2 + \int_U |c(x,k,u) - c(z,k,u)|^2 \nu(du) \le \kappa_R |x-z|, \tag{3.4}$$

where κ_R *is a positive constant and* $\operatorname{sgn}(a) = 1_{\{a>0\}} - 1_{\{a\le 0\}}$. *In addition, for each* $k \in \mathbb{S}$, *the function* c *satisfies that*

$$\text{the function } x \mapsto x + c(x,k,u) \text{ is nondecreasing for all } u \in U; \tag{3.5}$$

or, there exists some $\beta > 0$ *such that*

$$|x - z + \theta(c(x,k,u) - c(z,k,u))| \geq \beta|x - z|, \forall (x,z,u,\theta) \in \mathbb{R} \times \mathbb{R} \times U \times [0,1]. \quad (3.6)$$

(ii) If $d \geq 2$*, then there exist a positive number* δ_0 *and a nondecreasing and concave function* $\rho : [0,\infty) \to [0,\infty)$ *satisfying*

$$0 < \rho(r) \leq (1+r)^2 \rho(r/(1+r)) \text{ for all } r > 0, \text{ and } \int_{0^+} \frac{dr}{\rho(r)} = \infty, \quad (3.7)$$

so that for all $k \in \mathbb{S}, R > 0$ *and* $x, z \in \mathbb{R}^d$ *with* $|x| \vee |z| \leq R$ *and* $|x - z| \leq \delta_0$*,*

$$\begin{aligned} &2\langle x - z, b(x,k) - b(z,k)\rangle + |\sigma(x,k) - \sigma(z,k)|^2 \\ &\qquad + \int_U |c(x,k,u) - c(z,k,u)|^2 \nu(du) \leq \kappa_R \rho(|x-z|^2), \end{aligned} \quad (3.8)$$

where κ_R *is a positive constant.*

Assumption 3.1.2. *For each* $k \in \mathbb{S}$*, there exists a concave function* $\gamma_k : \mathbb{R}_+ \mapsto \mathbb{R}_+$ *with* $\gamma(0) = 0$ *such that for all* $x, y \in \mathbb{R}^d$ *with* $|x| \vee |y| \leq R$*, there exists a positive constant* κ_R *(which, without loss of generality, can be assumed to be the same positive constant as in* (3.3) *and* (3.4)*) such that*

$$\sum_{l \in \mathbb{S}\setminus\{k\}} |q_{kl}(x) - q_{kl}(y)| \leq \kappa_R \gamma_k(|x - y|). \quad (3.9)$$

Remark 3.1.3. *We note that examples of functions that satisfy condition* (3.2) *or* (3.7) *include* $\rho(r) = r$*,* $\rho(r) = r\log(1/r)$*,* $\rho(r) = r\log(\log(1/r))$*, and* $\rho(r) = r\log(1/r)\log(\log(1/r))$ *for* $r \in (0, \delta_0)$ *with* δ_0 *small enough. When* $\rho(r) = r$ *Assumption 3.1.1 reduces to the usual local Lipschitz condition. For other choices of* $\rho(r)$ *Assumption 3.1.1 allows the coefficients of* (2.11) *to be non-Lipschitz. This enables us to work with regime-switching (jump) diffusions with non-Lipschitz coefficients. This is an important result in Xi et al. (2019) which provides us the opportunity to explore a larger class of regime-switching (jump) diffusion processes.*

Note also that conditions (3.3), (3.4) *and* (3.8) *require the coefficients to be continuous in a small neighborhood of the diagonal line* $x = z$ *in* $\mathbb{R}^d \otimes \mathbb{R}^d$ *with* $|x| \vee |z| \le R$ *for each* $R > 0$.

Remark 3.1.4. *Assumption 3.1.1 is comparable to the corresponding assumption in Xi et al. (2019), except that the non-local term in* (3.4) *and* (3.8) *only requires the regularity of* $\int_U |c(x,k,u) - c(z,k,u)|^2 \nu(du)$. *In Xi et al. (2019), the corresponding term is* $\int_U [|c(x,k,u) - c(z,k,u)|^2 \wedge 4|x-z| \cdot |c(x,k,u) - c(z,k,u)|] \nu(du)$.

Moreover, Assumption 3.1.2 is weaker than that in Xi et al. (2019). Indeed, that paper assumes that $Q(x) = (q_{kl}(x))$ *satisfies*

$$\sum_{l \in \mathbb{S} \setminus \{k\}} |q_{kl}(x) - q_{kl}(y)| \le \kappa_R \, \rho\left(\frac{|x-y|}{1+|x-y|} \right), \text{ for each } k \in \mathbb{S},$$

for all $x, y \in \mathbb{R}^d$ *with* $|x| \vee |y| \le R$, *in which* $\kappa_R > 0$ *and* ρ *is an increasing and concave function satisfying* (3.7). *In contrast, the function* γ_k *in Assumption 3.1.2 may depend on* k, *and is only required to be concave with* $\gamma_k(0) = 0$. *In particular, the non-integrability condition* $\int_{0+} \frac{dr}{\rho(r)} = \infty$ *is dropped. This relaxation is significant and renders that the analyses in Xi et al. (2019) are not applicable.*

We introduce the following notations. Let $(X(\cdot), \Lambda(\cdot), \tilde{X}(\cdot), \tilde{\Lambda}(\cdot))$ denote the coupling process corresponding to the operator $\widetilde{\mathscr{A}}$ with initial condition (x,k,z,k), in which $\delta_0 > |x-z| > 0$, and δ_0 is the positive constant in Assumption 3.1.1. For any $R > 0$, let

$$\tau_R := \inf\{t \ge 0 : |\tilde{X}(t)| \vee |X(t)| \vee |\tilde{\Lambda}(t)| \vee |\Lambda(t)| > R\}. \tag{3.10}$$

In view of Assumption 3.0.2, $\lim_{R\to\infty} \tau_R = \infty$ a.s. Also denote $\Delta_t = \tilde{X}(t) - X(t)$ and

$$S_{\delta_0} := \inf\{t \ge 0 : |\Delta_t| > \delta_0\} = \inf\{t \ge 0 : |\tilde{X}(t) - X(t)| > \delta_0\}. \tag{3.11}$$

In addition, let

$$\zeta := \inf\{t \ge 0 : \Lambda(t) \ne \tilde{\Lambda}(t)\} \tag{3.12}$$

denote the first time when the switching components Λ and $\tilde{\Lambda}$ differ.

Lemma 3.1.5. *Under Assumption 3.1.1, the following assertion holds:*

$$\lim_{|\tilde{x}-x|\to 0} \mathbb{E}[|\Delta_{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta}|] = 0, \quad \forall t \geq 0. \tag{3.13}$$

Proof. We will prove the lemma separately for the cases $d = 1$ and $d \geq 2$.

Case (i)*: $d = 1$.* Let $\{a_n\}$ be a strictly decreasing sequence of real numbers satisfying $a_0 = 1$, $\lim_{n\to\infty} a_n = 0$, and $\int_{a_n}^{a_{n-1}} \frac{\mathrm{d}r}{r} = n$ for each $n \geq 1$. For each $n \geq 1$, let ρ_n be a nonnegative continuous function with support on (a_n, a_{n-1}) so that

$$\int_{a_n}^{a_{n-1}} \rho_n(r)\mathrm{d}r = 1 \text{ and } \rho_n(r) \leq 2(nr)^{-1} \text{ for all } r > 0.$$

For $x \in \mathbb{R}$, define

$$\psi_n(x) = \int_0^{|x|} \int_0^y \rho_n(z)\mathrm{d}z\mathrm{d}y. \tag{3.14}$$

We can immediately verify that ψ_n is even and twice continuously differentiable, with

$$\psi_n'(r) = \operatorname{sgn}(r) \int_0^{|r|} \rho_n(z)\mathrm{d}z = \operatorname{sgn}(r)|\psi_n'(r)|, \tag{3.15}$$

and

$$|\psi_n'(r)| \leq 1, \quad 0 \leq |r|\psi_n''(r) = |r|\rho_n(|r|) \leq \frac{2}{n}, \quad \text{and} \quad \lim_{n\to\infty} \psi_n(r) = |r| \tag{3.16}$$

for $r \in \mathbb{R}$. Furthermore, for each $r > 0$, the sequence $\{\psi_n(r)\}_{n\geq 1}$ is nondecreasing. Note also that for each $n \in \mathbb{N}$, ψ_n, ψ_n', and ψ_n'' all vanish on the interval $(-a_n, a_n)$. Moreover the classical arguments using Assumption 3.1.1 (i), (3.15) and (3.16) reveal that

$$\begin{aligned}\widetilde{\mathcal{L}}_k\psi_n(x-z) = &\frac{1}{2}\psi_n''(x-z)|\sigma(x,k)-\sigma(z,k)|^2 + \psi_n'(x-z)(b(x,k)-b(z,k)) \\ &+ \int_U [\psi_n(x-z+c(x,k,u)-c(z,k,u)) \\ &\qquad - \psi_n(x-z) - \psi_n'(x-z)(c(x,k,u)-c(z,k,u))]\nu(\mathrm{d}u)\end{aligned}$$

$$\leq K\frac{\kappa_R}{n} + \kappa_R\varrho(|x-z|), \tag{3.17}$$

for all x, z with $|x|\vee|z| \leq R$ and $0 < |x-z| \leq \delta_0$, where K is a positive constant independent of R and n. Then it follows that

$$\begin{aligned}
\mathbb{E}[\psi_n(\Delta_{t\wedge S_{\delta_0}\wedge\tau_R\wedge\zeta})] &= \mathbb{E}[\psi_n(\tilde{X}(t\wedge S_{\delta_0}\wedge\tau_R\wedge\zeta) - X(t\wedge S_{\delta_0}\wedge\tau_R\wedge\zeta))] \\
&= \psi_n(\tilde{x}-x) + \mathbb{E}\bigg[\int_0^{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta} \widetilde{\mathcal{L}}_k\psi_n(\tilde{X}(s)-X(s))\mathrm{d}s\bigg] \\
&\leq \psi_n(|\Delta_0|) + \mathbb{E}\bigg[\int_0^{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta}\Big(\kappa_R\varrho(|\Delta_s|) + K\frac{\kappa_R}{n}\Big)\mathrm{d}s\bigg] \\
&\leq \psi_n(|\Delta_0|) + K\frac{\kappa_R}{n}t + \kappa_R\int_0^t \rho\big(\mathbb{E}[|\Delta_{s\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta}|]\big)\mathrm{d}s,
\end{aligned}$$

where the first inequality follows from (3.17) and the second inequality follows from the concavity of ρ and Jensen's inequality. Then we use the monotone convergence theorem and (3.16) to derive

$$\mathbb{E}[|\Delta_{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta}|] \leq |\Delta_0| + \kappa_R\int_0^t \rho(\mathbb{E}[|\Delta_{s\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta}|])ds.$$

Let $u(t) := \mathbb{E}[|\Delta_{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta}|]$. Then u satisfies

$$0 \leq u(t) \leq v(t) := |\Delta_0| + \kappa_R\int_0^t \rho(u(s))ds.$$

Define the function $\Gamma(r) := \int_1^r \frac{ds}{\rho(s)}$ for $r > 0$. Thanks to (3.2), we can verify that Γ is nondecreasing and satisfies $\Gamma(r) > -\infty$ for all $r > 0$ and $\lim_{r\to 0}\Gamma(r) = -\infty$. Then we have

$$\begin{aligned}
\Gamma(u(t)) \leq \Gamma(v(t)) &= \Gamma(|\Delta_0|) + \int_0^t \Gamma'(v(s))v'(s)ds = \Gamma(|\Delta_0|) + \kappa_R\int_0^t \frac{\rho(u(s))}{\rho(v(s))}ds \\
&\leq \Gamma(|\Delta_0|) + \kappa_R\int_0^t 1ds = \Gamma(|\Delta_0|) + \kappa_R t,
\end{aligned}$$

where we use the assumption that ρ is nondecreasing to obtain the last inequality. Taking the limit $|\Delta_0| = |\tilde{x} - x| \to 0$ we have $\Gamma(u(t)) \to -\infty$ since $\lim_{r\to 0}\Gamma(r) = -\infty$. Moreover, since $\Gamma(r) > -\infty$ for all $r > 0$ we must have $\lim_{|\tilde{x}-x|\to 0} u(t) = 0$. This gives (3.13) as desired.

Case (ii) $d \geq 2$. Consider the function $f(x,z) := |x-z|^2$. Then Assumption 3.1.1 (ii) implies that

$$\begin{aligned}\widetilde{\mathcal{L}}_k f(x,z) &= 2\langle x-z, b(x,k)-b(z,k)\rangle + |\sigma(x,k)-\sigma(z,k)|^2 + \int_U |c(x,k,u)-c(z,k,u)|^2\nu(du)\\ &\leq \kappa_R \rho(|x-z|^2),\end{aligned}$$

for all $x, z \in \mathbb{R}^d$ with $|x| \vee |z| \leq R$ and $|x-z| \leq \delta_0$. Consequently

$$\begin{aligned}\mathbb{E}\big[|\Delta_{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta}|^2\big] &= \mathbb{E}[f(\tilde{X}(t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta), X(t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta))]\\ &= f(\tilde{x},x) + \mathbb{E}\bigg[\int_0^{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta} \widetilde{\mathcal{L}}_k f(\tilde{X}(s), X(s))ds\bigg]\\ &\leq |\Delta_0| + \mathbb{E}\bigg[\int_0^{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta} \kappa_R\rho(|\tilde{X}(s)-X(s)|^2)ds\bigg]\\ &\leq |\Delta_0| + \kappa_R\int_0^t \rho(\mathbb{E}[|\Delta_{s\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta}|^2])ds,\end{aligned}$$

where the last inequality follows from the concavity of ρ and Jensen's inequality. Using the same argument as that in Case (i), we can show that $\lim_{|\tilde{x}-x|\to 0}\mathbb{E}[|\Delta_{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta}|^2] = 0$; which, together with Hölder's inequality, leads to (3.13).

Combining the two cases completes the proof. □

Now we are ready to show that the process (X, Λ) has Feller property.

Theorem 3.1.6. *Under Assumptions 3.1.1 and 3.1.2, the process* (X,Λ) *has Feller property.*

Proof. We need to show that for each $(x,k) \in \mathbb{R}^d \times \mathbb{S}$ and each $f \in C_b(\mathbb{R}^d \times \mathbb{S})$, the limit $(P_t f)(\tilde{x},\tilde{k}) \to (P_t f)(x,k)$ as $(\tilde{x},\tilde{k}) \to (x,k)$ holds for all $t \geq 0$. Since $\mathbb{S} = \{1,2,...\}$ has a discrete topology, it is enough to consider only $(\tilde{x},k) \to (x,k)$. First, observe that

$$|(P_t f)(\tilde{x},\tilde{k}) - (P_t f)(x,k)| = |\mathbb{E}[f(\tilde{X}(t),\tilde{\Lambda}(t))] - \mathbb{E}[f(X(t),\Lambda(t))]|$$

$$
\begin{aligned}
&\leq |\mathbb{E}[f(\tilde{X}(t),\tilde{\Lambda}(t))] - \mathbb{E}[f(\tilde{X}(t),\Lambda(t))]| \\
&\quad + |\mathbb{E}[f(\tilde{X}(t),\Lambda(t))] - \mathbb{E}[f(X(t),\Lambda(t))]| \\
&= |\mathbb{E}[(f(\tilde{X}(t),\tilde{\Lambda}(t)) - f(\tilde{X}(t),\Lambda(t)))1_{\{\zeta\leq t\}}]| \\
&\quad + |\mathbb{E}[(f(\tilde{X}(t),\tilde{\Lambda}(t)) - f(\tilde{X}(t),\Lambda(t)))1_{\{\zeta> t\}}]| \\
&\quad + |\mathbb{E}[f(\tilde{X}(t),\Lambda(t))] - \mathbb{E}[f(X(t),\Lambda(t))]| \\
&\leq 2||f||_\infty \mathbb{P}\{\zeta\leq t\} + |\mathbb{E}[f(\tilde{X}(t),\Lambda(t))] - \mathbb{E}[f(X(t),\Lambda(t))]|.
\end{aligned} \tag{3.18}
$$

We will show that both terms on the right-hand side of (3.18) converge to 0 as $\tilde{x}\to x$.

Consider the function $\Xi(x,k,z,l) := 1_{\{k\neq l\}}$. It follows directly from the definition that

$$
\widetilde{\mathscr{A}}\Xi(x,k,z,l) = \widetilde{\Omega}_{\rm s}\Xi(x,k,z,l) \leq 0, \text{ if } k\neq l.
$$

When $k=l$, we have from (3.9) that

$$
\begin{aligned}
\widetilde{\mathscr{A}}\Xi(x,k,z,l) &= \widetilde{\Omega}_{\rm s}\Xi(x,k,z,k) \\
&= \sum_{i\in\mathbb{S}}[q_{ki}(x)-q_{ki}(z)]^+(1_{\{i\neq k\}} - 1_{\{k\neq k\}}) + \sum_{i\in\mathbb{S}}[q_{ki}(z)-q_{ki}(x)]^+(1_{\{i\neq k\}} - 1_{\{k\neq k\}}) \\
&\leq \sum_{i\in\mathbb{S}, i\neq k} |q_{ki}(x)-q_{ki}(z)| \leq \kappa_R\gamma_k(|x-y|).
\end{aligned}
$$

Hence

$$
\widetilde{\mathscr{A}}\Xi(x,k,z,l) \leq \kappa_R\gamma_k(|x-y|) \tag{3.19}
$$

for all $k,l\in\mathbb{S}$ and $x,z\in\mathbb{R}^d$ with $|x|\vee|z|\leq R$. Note that $\zeta\leq t\wedge\tau_R\wedge S_{\delta_0}$ if and only if $\tilde{\Lambda}(t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta)\neq\Lambda(t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta)$. Thus we can use (3.19) to compute

$$
\begin{aligned}
&\mathbb{P}\{\zeta\leq t\wedge\tau_R\wedge S_{\delta_0}\} \\
&\quad = \mathbb{E}[\Xi(\tilde{X}(t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta), \tilde{\Lambda}(t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta), X(t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta), \Lambda(t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta))]
\end{aligned}
$$

$$
\begin{aligned}
&= \Xi(\tilde{x}, k, x, k) + \mathbb{E}\bigg[\int_0^{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta} \widetilde{\mathscr{A}}\Xi(\tilde{X}(s), \tilde{\Lambda}(s), X(s), \Lambda(s))ds\bigg] \\
&\le \kappa_R \mathbb{E}\left[\int_0^{t\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta} \gamma_k(|\tilde{X}(s) - X(s)|)ds\right] \\
&\le \kappa_R \int_0^t \mathbb{E}[\gamma_k(|\tilde{X}(s\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta) - X(s\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta)|)]ds \\
&\le \kappa_R \int_0^t \gamma_k\big(\mathbb{E}[|\Delta_{s\wedge\tau_R\wedge S_{\delta_0}\wedge\zeta}|]\big)ds,
\end{aligned}
$$

where the last inequality follows from the assumption that γ_k is concave. Then it follows from (3.13), the assumption that $\gamma_k(0) = 0$, and the bounded convergence theorem that

$$
\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{\zeta \le t\wedge\tau_R\wedge S_{\delta_0}\} = 0. \tag{3.20}
$$

Note also that on the set $\{S_{\delta_0} \le t\wedge\zeta\wedge\tau_R\}$ we have $\delta_0 \le |\Delta_{S_{\delta_0}\wedge t\wedge\zeta\wedge\tau_R}|$. This implies

$$
\delta_0 \mathbb{P}\{S_{\delta_0} \le t\wedge\zeta\wedge\tau_R\} \le \mathbb{E}[|\Delta_{t\wedge S_{\delta_0}\wedge\zeta\wedge\tau_R}| 1_{\{S_{\delta_0}\le t\wedge\zeta\wedge\tau_R\}}] \le \mathbb{E}[|\Delta_{t\wedge S_{\delta_0}\wedge\zeta\wedge\tau_R}|].
$$

Therefore, it follows from (3.13) that

$$
\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{S_{\delta_0} \le t\wedge\zeta\wedge\tau_R\} = 0. \tag{3.21}
$$

Fix an arbitrary positive number ϵ. From (3.13) we have

$$
\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{|\Delta_{t\wedge S_{\delta_0}\wedge\tau_R\wedge\zeta}| > \epsilon\} \le \lim_{|\tilde{x}-x|\to 0} \frac{\mathbb{E}\big[|\Delta_{t\wedge S_{\delta_0}\wedge\tau_R\wedge\zeta}|\big]}{\epsilon} = 0 \tag{3.22}
$$

Since $\lim_{R\to\infty} \tau_R = \infty$ a.s., we can choose R sufficiently large so that

$$
\mathbb{P}\{\tau_R < t\} < \epsilon. \tag{3.23}
$$

Then

$$
\begin{aligned}
\mathbb{P}\{|\Delta_t| > \varepsilon\} &= \mathbb{P}\{|\Delta_t| > \varepsilon, \tau_R < t\} + \mathbb{P}\{|\Delta_t| > \varepsilon, \tau_R \geq t, \zeta \leq t \wedge S_{\delta_0} \wedge \tau_R\} \\
&\quad + \mathbb{P}\{|\Delta_t| > \varepsilon, \tau_R \geq t, \zeta > t \wedge S_{\delta_0} \wedge \tau_R, S_{\delta_0} \leq t \wedge \tau_R \wedge \zeta\} \\
&\quad + \mathbb{P}\{|\Delta_t| > \varepsilon, \tau_R \geq t, \zeta > t \wedge S_{\delta_0} \wedge \tau_R, S_{\delta_0} > t \wedge \tau_R \wedge \zeta\} \\
&\leq \epsilon + \mathbb{P}\{\zeta \leq t \wedge S_{\delta_0} \wedge \tau_R\} + \mathbb{P}\{S_{\delta_0} \leq t \wedge \tau_R \wedge \zeta\} + \mathbb{P}\{|\Delta_t| > \varepsilon, t \leq S_{\delta_0} \wedge \tau_R \wedge \zeta\} \\
&\leq \epsilon + \mathbb{P}\{\zeta \leq t \wedge S_{\delta_0} \wedge \tau_R\} + \mathbb{P}\{S_{\delta_0} \leq t \wedge \tau_R \wedge \zeta\} + \mathbb{P}\{|\Delta_{t \wedge S_{\delta_0} \wedge \tau_R \wedge \zeta}| > \varepsilon\}.
\end{aligned}
$$

From (3.20)–(3.22) we have

$$
\lim_{|\tilde{x}-x| \to 0} \mathbb{P}\{|\Delta_t| > \varepsilon\} \leq \epsilon.
$$

Since ϵ is arbitrary, we conclude that $\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{|\Delta_t| > \varepsilon\} = 0$. In other words, $\tilde{X}(t) \to X(t)$ in probability as $\tilde{x} \to x$. With the metric d on $\mathbb{R}^d \times \mathbb{S}$ defined by $d((x,i),(y,j)) := |x-y| + 1_{\{i \neq j\}}$, we see immediately that $(\tilde{X}(t), \Lambda(t)) \to (X(t), \Lambda(t))$ in probability as $\tilde{x} \to x$. Because the function f is continuous, we also have $f(\tilde{X}(t), \Lambda(t)) \to f(X(t), \Lambda(t))$ in probability as $\tilde{x} \to x$. Then the bounded convergence theorem implies

$$
|\mathbb{E}[f(\tilde{X}(t), \Lambda(t))] - \mathbb{E}[f(X(t), \Lambda(t))]| \to 0 \text{ as } \tilde{x} \to x. \tag{3.24}
$$

Next, we show that $\lim_{\tilde{x}\to x} \mathbb{P}\{\zeta \leq t\} = 0$ holds. Recall that R is chosen so that (3.23) holds. Then we compute

$$
\begin{aligned}
\mathbb{P}\{\zeta \leq t\} &= \mathbb{P}\{\zeta \leq t, \tau_R < t\} + \mathbb{P}\{\zeta \leq t, \tau_R \geq t\} \\
&\leq \mathbb{P}\{\tau_R < t\} + \mathbb{P}\{\zeta \leq t, \tau_R \geq t, S_{\delta_0} \leq t \wedge \zeta\} + \mathbb{P}\{\zeta \leq t, \tau_R \geq t, S_{\delta_0} > t \wedge \zeta\} \\
&\leq \epsilon + \mathbb{P}\{S_{\delta_0} \leq t \wedge \zeta \wedge \tau_R\} + \mathbb{P}\{\zeta \leq t \wedge \tau_R \wedge S_{\delta_0}\}.
\end{aligned}
$$

It then follows from (3.20) and (3.21) that $\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{\zeta \leq t\} \leq \epsilon$. Again since ϵ is arbitrary, we conclude that

$$\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{\zeta \le t\} = 0. \tag{3.25}$$

Finally we plug (3.24) and (3.25) into (3.18) to complete the proof. □

3.2 Strong Feller Property

To facilitate future presentations, we introduce the following notations. For any $x, y \in \mathbb{R}^d$ and $i, j \in \mathbb{S}$, put

$$\begin{aligned}
A(x,i,y,j) &:= a(x,i) + a(y,j) - 2\hat{g}(x,i,y,j),\\
\bar{A}(x,i,y,j) &:= \frac{1}{|x-y|^2}\langle x-y, A(x,i,y,j)(x-y)\rangle,\\
B(x,i,y,j) &:= \langle x-y, b(x,i) - b(y,j)\rangle.
\end{aligned}$$

We first obtain following lemma whose proof involves elementary and straightforward computations; see section 7.1. Similar computations can be found in Chen and Li (1989) and Priola and Wang (2006).

Lemma 3.2.1. *For each $x, y \in \mathbb{R}^d$ and $i, j \in \mathbb{S}$, we have*

(i) $\widehat{a}(x,i,y,j)$ *is symmetric and uniformly positive definite,*

(ii) $\mathrm{tr}A(x,i,y,j) = |\sigma_{\lambda_R}(x,i) - \sigma_{\lambda_R}(y,j)|^2 + 4\lambda_R$, *and*

(iii) $\bar{A}(x,i,y,j) \ge 4\lambda_R$.

To derive the strong Feller property we need stronger conditions than those in Assumption 3.1.1.

Assumption 3.2.2. *For every $k \in \mathbb{S}$ the following assertions hold:*

(i) For every $R > 0$ there exits a constant $\lambda_R > 0$ such that

$$\langle \xi, a(x,k)\xi\rangle \geq \lambda_R|\xi|^2, \qquad \xi \in \mathbb{R}^d, \tag{3.26}$$

for all $x \in \mathbb{R}^d$ with $|x| \leq R$, where $a(x,k) := \sigma(x,k)\sigma(x,k)^T$.

(ii) There exits a nonnegative function $g \in C(0,\infty)$ satisfying

$$\int_0^1 g(r)dr < \infty, \tag{3.27}$$

and

$$2\langle x-z, b(x,k)-b(z,k)\rangle + |\sigma_{\lambda_R}(x,k) - \sigma_{\lambda_R}(z,k)|^2 \\ + \int_U |c(x,k,u) - c(z,k,u)|^2\nu(du) \leq 2\kappa_R|x-z|g(|x-z|), \tag{3.28}$$

for all $x, z \in \mathbb{R}^d$ with $|x| \vee |z| \leq R$ and $|x-z| \leq \delta_0$, where δ_0 is a positive constant and σ_{λ_R} the unique symmetric nonnegative definite matrix-valued function such that $\sigma^2_{\lambda_R}(x,k) = a(x,k) - \lambda_R I$.

Remark 3.2.3. *Note that $\langle \xi, a(x,k)\xi\rangle \geq 0$ for all $\xi \in \mathbb{R}^d$. However, to obtain the strong Feller property we need this to be strictly positive. So we require the elliptic condition (3.26). If we set $\lambda_R = 0$ and $g(r) = r$ then condition (3.28) reduces to condition (3.8) with $\rho(r) = r$. In this case, the coefficients satisfy the usual local Lipschitz condition.*

Assumption 3.2.2 improves significantly over those in the literature such as Shao (2015b), Xi and Zhu (2017), which require Lipschitz condition for the coefficients of the associated stochastic differential equations. By contrast, (3.28) places very mild conditions on the coefficients; it allows to treat for example the case of Hölder continuous coefficients by taking $g(r) = r^{-p}$ for $0 \leq p < 1$; see Example 3.3.2.

Moreover, the condition (3.28) significantly weakens the requirement in Xi et al. (2019)

as we do not require g to satisfy $\lim_{r\to 0} g(r) = 0$; see Lemma 4.4 of Xi et al. (2019). Therefore, Theorem 3.2.8 improves Theorem 4.8 of Xi et al. (2019) on strong Feller property.

As in Section 3.1, we will use the coupling method to prove Theorem 3.2.8. Now, let $\phi \in C^2([0,\infty))$. As in Chen and Li (1989), for each $k \in \mathbb{S}$ and all $x, z \in \mathbb{R}^d$ with $x \neq z$, we can verify that

$$\begin{aligned}\widehat{\Omega}_{\mathrm{d}}^{(k)}\phi(|x-z|) = &\frac{\phi''(|x-z|)}{2}\bar{A}(x,k,z,k)\\ &+\frac{\phi'(|x-z|)}{2|x-z|}[\mathrm{tr}A(x,k,z,k)-\bar{A}(x,k,z,k)+2B(x,k,z,k)].\end{aligned} \tag{3.29}$$

Moreover, we have

$$\begin{aligned}\tilde{\Omega}_{\mathrm{j}}^{(k)}\phi(|x-z|) = \int_U &(\phi(|x+c(x,k,u)-z-c(z,k,u)|)-\phi(|x-z|)\\ &-\frac{\phi'(|x-z|)}{|x-z|}\langle x-z, c(x,k,u)-c(z,k,u)\rangle)\nu(du).\end{aligned} \tag{3.30}$$

Motivated by Priola and Wang (2006), we consider the function G given by

$$G(r) := \int_0^r \exp\left\{-\int_0^s \frac{\kappa_R}{2\lambda_R}g(w)dw\right\}\int_s^1 \exp\left\{\int_0^v \frac{\kappa_R}{2\lambda_R}g(u)du\right\}dvds, \qquad r\in[0,1]$$

where g is the function given in Assumption 3.2.2 (ii). Since $g \geq 0$, we see that

$$G'(r) = e^{-\int_0^r \frac{\kappa_R}{2\lambda_R}g(w)dw}\int_r^1 e^{\int_0^v \frac{\kappa_R}{2\lambda_R}g(u)du}dv \geq 0, \quad \text{and} \quad G''(r) = -1-\frac{\kappa_R}{2\lambda_R}g(r)G'(r) \leq 0. \tag{3.31}$$

Note also that G is concave and $\lim_{r\to 0} G(r) = 0$. Since $G'(0) \geq 1$ and $G(0) = 0$, there exists a constant $\alpha \in (0,1)$ so that

$$r \ \leq \ G(r) \qquad \text{for all} \quad r \in [0,\alpha]. \tag{3.32}$$

Lemma 3.2.4. *Suppose Assumptions 3.2.2 holds. Then for any $R > 0$ and $k \in \mathbb{S}$ there exits* *a positive constant* $\beta_R > 0$ *such that*

$$\widehat{\mathcal{L}}_k G(|x-z|) \le -\beta_R \tag{3.33}$$

for all $x, z \in \mathbb{R}^d$ *with* $|z| \vee |x| \le R$ *and* $0 < |x-z| \le \alpha \wedge \delta_0$*, where* $\alpha > 0$ *is given in* (3.32).

Proof. In view of (3.29), it follows from (3.31) that

$$\begin{aligned}
&\widehat{\Omega}_{\mathrm{d}}^{(k)} G(|x-z|) \\
&= \frac{G''(|x-z|)}{2}\bar{A}(x,k,z,k) + \frac{G'(|x-z|)}{2|x-z|}[\mathrm{tr}A(x,k,z,k) - \bar{A}(x,k,z,k) + 2B(x,k,z,k)] \\
&\le \frac{G''(|x-z|)}{2}4\lambda_R + \frac{G'(|x-z|)}{2|x-z|}[|\sigma_{\lambda_R}(x,k) - \sigma_{\lambda_R}(z,k)|^2 + 2B(x,k,z,k)] \\
&= 2\lambda_R\left(-1 - \frac{\kappa_R}{2\lambda_R}g(|x-z|)F'(|x-z|)\right) \\
&\qquad + \frac{G'(|x-z|)}{2|x-z|}[|\sigma_{\lambda_R}(x,k) - \sigma_{\lambda_R}(z,k)|^2 + 2B(x,k,z,k)] \\
&= -2\lambda_R + \left(-\kappa_R g(|x-z|) + \frac{|\sigma_{\lambda_R}(x,k) - \sigma_{\lambda_R}(z,k)|^2 + 2B(x,k,z,k)}{2|x-z|}\right)G'(|x-z|).
\end{aligned} \tag{3.34}$$

Since the function G is concave, we have $G(r_1) - G(r_0) \le G'(r_0)(r_1 - r_0)$ for all $r_0, r_1 \ge 0$. Take $r_0 = |x-z|$ and $r_1 = |x + c(x,k,u) - z - c(z,k,u)|$ to obtain

$$\begin{aligned}
&G(|x + c(x,k,u) - z - c(z,k,u)|) - G(|x-z|) - \frac{G'(|x-z|)}{|x-z|}\langle x-z, c(x,k,u) - c(z,k,u)\rangle \\
&\le G'(|x-z|)\left(|x + c(x,k,u) - z - c(z,k,u)| - |x-z| - \frac{\langle x-z, c(x,k,u) - c(z,k,u)\rangle}{|x-z|}\right).
\end{aligned}$$

Furthermore, with $a := x - z$ and $b := c(x,k,u) - c(z,k,u)$, we can verify directly that

$$|a+b| - |a| - \frac{\langle a,b\rangle}{|a|} = \frac{-(|a+b| - |a|)^2 + |b|^2}{2|a|} \le \frac{|b|^2}{2|a|}.$$

Hence it follows that

$$
\begin{aligned}
&G(|x+c(x,k,u)-z-c(z,k,u)|)-G(|x-z|)-\frac{G'(|x-z|)}{|x-z|}\langle x-z, c(x,k,u)-c(z,k,u)\rangle\\
&\quad\le\left(\frac{|c(x,k,u)-c(z,k,u)|^2}{2|x-z|}\right)G'(|x-z|).
\end{aligned}
$$

Then we have

$$
\tilde{\Omega}_{\mathrm{j}}^{(k)}G(|x-z|)\le G'(|x-z|)\int_U\frac{|c(x,k,u)-c(z,k,u)|^2}{2|x-z|}\nu(du). \tag{3.35}
$$

From (3.34) and (3.35), we see that

$$
\begin{aligned}
\widehat{\mathcal{L}}_kG(|x-z|)&=[\widehat{\Omega}_{\mathrm{d}}^{(k)}+\tilde{\Omega}_{\mathrm{j}}^{(k)}]G(|x-z|)\\
&\le -2\lambda_R+G'(|x-z|)\bigg(-\kappa_Rg(|x-z|)+\frac{|\sigma_{\lambda_R}(x,i)-\sigma_{\lambda_R}(z,j)|^2+2B(x,k,z,k)}{2|x-z|}\\
&\qquad\qquad+\int_U\frac{|c(x,k,u)-c(z,k,u)|^2}{2|x-z|}\nu(du)\bigg)\\
&\le -2\lambda_R.
\end{aligned}
$$

The proof is complete. □

Throughout the rest of the section, we use the following notations. For any $x,\tilde{x}\in\mathbb{R}^d$ and $k\in\mathbb{S}$, denote by $(X(\cdot),\Lambda(\cdot),\tilde{X}(\cdot),\tilde{\Lambda}(\cdot))$ the process corresponding to the coupling operator $\widehat{\mathscr{A}}$ with initial condition $(x,k,\tilde{x},k)$. As in Section 3.1, $\Delta_t:=\tilde{X}(t)-X(t), t\ge 0$. Let τ_R, S_{δ_0}, and ζ be defined as in (3.10), (3.11), and (3.12), respectively. In addition, for each $n\in\mathbb{N}$, we define

$$
T_n:=\inf\left\{t\ge 0:|X(t)-\tilde{X}(t)|<\frac{1}{n}\right\}. \tag{3.36}
$$

Then $\lim_{n\to\infty}T_n=T$, where T is the coupling time given by

$$
T:=\inf\{t\ge 0: X(t)=\tilde{X}(t)\}. \tag{3.37}
$$

Lemma 3.2.5. *Suppose Assumption 3.2.2 holds. Then the following assertions hold for every* $t \geq 0$*:*

$$\lim_{|\tilde{x}-x|\to 0} \mathbb{E}[G(|\Delta_{t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta}|)] = 0. \tag{3.38}$$

$$\lim_{|\tilde{x}-x|\to 0} \mathbb{E}[G(|\Delta_{t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta-}|)] = 0. \tag{3.39}$$

In particular,

$$\lim_{|\tilde{x}-x|\to 0} \mathbb{E}[|\Delta_{t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta-}|] = 0, \tag{3.40}$$

where $\bar{\delta} := \delta_0 \wedge \alpha$*,* δ_0 *is the constant given in Assumption 3.2.2 (ii), and* $\alpha \in (0,1)$ *is the constant given in* (3.32)*.*

Proof. Assume without loss of generality that $\bar{\delta} \geq |x - \tilde{x}| > 0$. We apply Itô's formula to the process $G(|\tilde{X}(\cdot) - X(\cdot)|) = G(|\Delta_\cdot|)$:

$$\begin{aligned}\mathbb{E}[G(|\Delta_{t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta}|)] &= G(|\Delta_0|) + \mathbb{E}\left[\int_0^{t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta} \widehat{\mathcal{L}}G(|\Delta_s|)ds\right] \\ &\leq G(|\Delta_0|) - \beta_R\mathbb{E}[t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta],\end{aligned}$$

where the last inequality follows from (3.33). Hence

$$\mathbb{E}[G(|\Delta_{t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta}|)] + \beta_R\mathbb{E}[t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta] \leq G(|\Delta_0|) = G(|x-\tilde{x}|).$$

Since $\lim_{r\to 0} G(r) = 0$, then we obtain (3.38). The same argument implies (3.39). Since $|\Delta_{t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta-}| \leq \bar{\delta} \leq \alpha$, it follows from (3.32) that

$$|\Delta_{t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta-}| \leq G(|\Delta_{t\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta-}|)$$

and therefore (3.40) follows. □

Lemma 3.2.6. *Suppose Assumptions 3.1.2 and 3.2.2 hold. Then for any $t \geq 0$ we have*

$$\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{\zeta \leq t\} = 0. \tag{3.41}$$

Proof. Given $\epsilon > 0$. Choose R sufficiently large so that $\mathbb{P}\{\tau_R \leq t\} < \epsilon$. Observe that

$$\begin{aligned}
\mathbb{P}\{\zeta \leq t\} &= \mathbb{P}\{\zeta \leq t, \tau_R < t\} + \mathbb{P}\{\zeta \leq t, \tau_R \geq t\} \\
&\leq \mathbb{P}\{\tau_R < t\} + \mathbb{P}\{\zeta \leq t, \tau_R \geq t, S_{\bar{\delta}} \leq t \wedge \zeta\} + \mathbb{P}\{\zeta \leq t, \tau_R \geq t, S_{\bar{\delta}} > t \wedge \zeta\} \\
&\leq \epsilon + \mathbb{P}\{\zeta \leq t, \tau_R \geq t, S_{\bar{\delta}} \leq t \wedge \zeta\} + \mathbb{P}\{\zeta \leq t, \tau_R \geq t, S_{\bar{\delta}} > t \wedge \zeta\} \\
&\leq \epsilon + \mathbb{P}\{S_{\bar{\delta}} \leq t \wedge \zeta \wedge \tau_R\} + \mathbb{P}\{\zeta \leq t \wedge \tau_R \wedge S_{\bar{\delta}}\}. \qquad (3.42)
\end{aligned}$$

As in the proof of Theorem 3.1.6, condition (3.9) enables us to derive

$$\mathbb{P}\{\zeta \leq t \wedge \tau_R \wedge S_{\bar{\delta}}\} \leq \kappa_R \int_0^t \gamma_k(\mathbb{E}[|\Delta_{s\wedge\tau_R\wedge S_{\bar{\delta}}\wedge\zeta^-}|])ds.$$

Furthermore, the limit in (3.40) implies

$$\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{\zeta \leq t \wedge \tau_R \wedge S_{\bar{\delta}}\} = 0. \tag{3.43}$$

Note that on the set $\{S_{\bar{\delta}} \leq t \wedge \zeta \wedge \tau_R\}$ we have $\bar{\delta} \leq |\Delta_{S_{\bar{\delta}}\wedge t\wedge\zeta\wedge\tau_R}|$. Since G is increasing, we obtain

$$0 < G(\bar{\delta}) \leq G(|\Delta_{t\wedge S_{\bar{\delta}}\wedge\zeta\wedge\tau_R}|).$$

Thus

$$G(\bar{\delta})\mathbb{P}\{S_{\bar{\delta}} \leq t \wedge \zeta \wedge \tau_R\} \leq \mathbb{E}[G(|\Delta_{t\wedge S_{\bar{\delta}}\wedge\zeta\wedge\tau_R}|)1_{\{S_{\bar{\delta}}\leq t\wedge\zeta\wedge\tau_R\}}] \leq \mathbb{E}[G(|\Delta_{t\wedge S_{\bar{\delta}}\wedge\zeta\wedge\tau_R}|)].$$

This, together with (3.38), implies that

$$\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{S_{\bar{\delta}} \le t \wedge \zeta \wedge \tau_R\} = 0. \tag{3.44}$$

In view of (3.42), it follows from (3.43) and (3.44) that $\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{\zeta \le t\} \le \epsilon$. Since ϵ is arbitrary we obtain (3.41). □

Lemma 3.2.7. *Suppose Assumptions 3.1.2 and 3.2.2 hold. Then for every $t \ge 0$ we have*

$$\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{t < T\} = 0. \tag{3.45}$$

Proof. We may assume without loss of generality that $\bar{\delta} \ge |x - \tilde{x}| > \frac{1}{n_0} > 0$ for some $n_0 \in \mathbb{N}$. Let $\epsilon > 0$ and choose a sufficiently large R so that $\mathbb{P}\{\tau_R \le t\} < \epsilon$. For each $n \ge n_0$, we define T_n and T as in (3.36) and (3.37), respectively. We first observe that

$$\begin{aligned}
\mathbb{P}\{t < T\} &= \mathbb{P}\{t < T, \tau_R < t\} + \mathbb{P}\{t < T, \tau_R \ge t\} \\
&\le \mathbb{P}\{\tau_R < t\} + \mathbb{P}\{t < T, \tau_R \ge t, S_{\bar{\delta}} < t\} + \mathbb{P}\{t < T, \tau_R \ge t, S_{\bar{\delta}} \ge t\} \\
&\le \epsilon + \mathbb{P}\{S_{\bar{\delta}} \le t \wedge T \wedge \tau_R\} + \mathbb{P}\{t \le T \wedge \tau_R \wedge S_{\bar{\delta}}\} \\
&= \epsilon + \mathbb{P}\{S_{\bar{\delta}} \le t \wedge T \wedge \tau_R, S_{\bar{\delta}} \le \zeta\} + \mathbb{P}\{S_{\bar{\delta}} \le t \wedge T \wedge \tau_R, S_{\bar{\delta}} > \zeta\} \\
&\quad + \mathbb{P}\{t \le T \wedge \tau_R \wedge S_{\bar{\delta}}, t < \zeta\} + \mathbb{P}\{t \le T \wedge \tau_R \wedge S_{\bar{\delta}}, t \ge \zeta\} \\
&\le \epsilon + \mathbb{P}\{S_{\bar{\delta}} \le t \wedge T \wedge \tau_R \wedge \zeta\} + \mathbb{P}\{\zeta < S_{\bar{\delta}} \wedge t \wedge T \wedge \tau_R\} \\
&\quad + \mathbb{P}\{t \le T \wedge \tau_R \wedge S_{\bar{\delta}} \wedge \zeta\} + \mathbb{P}\{\zeta \le t\} \\
&\le \epsilon + \mathbb{P}\{S_{\bar{\delta}} \le t \wedge T \wedge \tau_R \wedge \zeta\} + \mathbb{P}\{\zeta \le t\} + \mathbb{P}\{t \le T \wedge \tau_R \wedge S_{\bar{\delta}} \wedge \zeta\} + \mathbb{P}\{\zeta \le t\} \\
&= \epsilon + \mathbb{P}\{S_{\bar{\delta}} \le t \wedge T \wedge \tau_R \wedge \zeta\} + \mathbb{P}\{t \le T \wedge \tau_R \wedge S_{\bar{\delta}} \wedge \zeta\} + 2\mathbb{P}\{\zeta \le t\} \\
&\le \epsilon + \mathbb{P}\{S_{\bar{\delta}} \le T \wedge \tau_R \wedge \zeta\} + \frac{\mathbb{E}[T \wedge \tau_R \wedge S_{\bar{\delta}} \wedge \zeta]}{t} + 2\mathbb{P}\{\zeta \le t\}.
\end{aligned} \tag{3.46}$$

Note that on the set $\{S_{\bar{\delta}} \le T_n \wedge \tau_R \wedge \zeta\}$ we have $\bar{\delta} \le |\Delta_{S_{\bar{\delta}} \wedge T_n \wedge \tau_R \wedge \zeta}|$. Since G is increasing, we have $0 < G(\bar{\delta}) \le G(|\Delta_{S_{\bar{\delta}} \wedge T_n \wedge \tau_R \wedge \zeta}|)$. Thus

$$\begin{aligned} G(\bar{\delta})\mathbb{P}\{S_{\bar{\delta}} \le T_n \wedge \tau_R \wedge \zeta\} &\le \mathbb{E}[G(|\Delta_{S_{\bar{\delta}} \wedge T_n \wedge \tau_R \wedge \zeta}|)1_{\{S_{\bar{\delta}} \le T_n \wedge \tau_R \wedge \zeta\}}] \le \mathbb{E}[G(|\Delta_{S_{\bar{\delta}} \wedge T_n \wedge \tau_R \wedge \zeta}|)] \\ &= G(|x - \tilde{x}|) + \mathbb{E}\bigg[\int_0^{S_{\bar{\delta}} \wedge T_n \wedge \tau_R \wedge \zeta} \widehat{\mathcal{L}}_k G(|\Delta_s|)ds\bigg] \\ &\le G(|x - \tilde{x}|) - \beta_R \mathbb{E}[T_n \wedge \tau_R \wedge S_{\bar{\delta}} \wedge \zeta], \end{aligned}$$

where that last inequality follows from (3.33). Hence

$$G(\bar{\delta})\mathbb{P}\{S_{\bar{\delta}} \le T_n \wedge \tau_R \wedge \zeta\} + \beta_R \mathbb{E}[T_n \wedge \tau_R \wedge S_{\bar{\delta}} \wedge \zeta] \le G(|x - \tilde{x}|).$$

Passing to the limit as $n \to \infty$, we obtain

$$G(\bar{\delta})\mathbb{P}\{S_{\bar{\delta}} \le T \wedge \tau_R \wedge \zeta\} + \beta_R \mathbb{E}[T \wedge \tau_R \wedge S_{\bar{\delta}} \wedge \zeta] \le G(|x - \tilde{x}|).$$

In view of (3.46), we see that

$$\begin{aligned} \mathbb{P}\{t < T\} &\le \epsilon + \mathbb{P}\{S_{\bar{\delta}} \le T \wedge \tau_R \wedge \zeta\} + \frac{\mathbb{E}[T \wedge \tau_R \wedge S_{\bar{\delta}} \wedge \zeta]}{t} + 2\mathbb{P}\{\zeta \le t\} \\ &\le \epsilon + \frac{G(|x - \tilde{x}|)}{G(\bar{\delta})} + \frac{G(|x - \tilde{x}|)}{t\beta} + 2\mathbb{P}\{\zeta \le t\}. \end{aligned}$$

From (3.41) and the fact that $\lim_{|\tilde{x} - x| \to 0} G(|x - \tilde{x}|) = 0$, we have $\lim_{|\tilde{x} - x| \to 0} \mathbb{P}\{t < T\} \le \epsilon$. Since ϵ was arbitrary we obtain (3.45). □

Now we are ready to present the proof of the main theorem of this section.

Theorem 3.2.8. *Suppose Assumptions 3.1.2 and 3.2.2 hold. Then the process (X, Λ) has strong Feller property.*

Proof. Given $x \in \mathbb{R}^d$ and $k \in \mathbb{S}$. We want to show that for every bounded Borel measurable function f on $\mathbb{R}^d$ the limit $(P_t f)(\tilde{x}, \tilde{k}) \to (P_t f)(x, k)$ as $(\tilde{x}, \tilde{k}) \to (x, k)$ holds for all $t > 0$. Since $\mathbb{S} = \{1, 2, ...\}$ has a discrete topology, we may consider only when $(\tilde{x}, k) \to (x, k)$; that is, when $\tilde{k} = k$.

For any given $\epsilon > 0$ we can choose a sufficiently large R so that $\mathbb{P}\{\tau_R \leq t\} < \epsilon$. Let $\tilde{x} \in \mathbb{R}^d$ be such that $\bar{\delta} \geq |x - \tilde{x}| > 0$, where $\bar{\delta} := \delta_0 \wedge \alpha$. Denote the coupling process corresponding to the coupling operator $\tilde{\mathcal{L}}$ with initial condition $(x, k, \tilde{x}, k)$ by $(X(t), \Lambda(t), \tilde{X}(t), \tilde{\Lambda}(t))$. Let

$$\tilde{T} := \inf\{t \geq 0 : (X(t), \Lambda(t)) = (\tilde{X}(t), \tilde{\Lambda}(t))\} \tag{3.47}$$

be the coupling time of $(X(t), \Lambda(t))$ and $(\tilde{X}(t), \tilde{\Lambda}(t))$. Recall the stopping time T defined in (3.37). We make the following observations:

(i) $T \leq \tilde{T}$, and

(ii) $T < \zeta$ implies $T = \tilde{T}$.

Then we have

$$\begin{aligned}
1_{\{t<\tilde{T}\}} &= 1_{\{t<T\}} + 1_{\{T\leq t<\tilde{T}\}} \\
&= 1_{\{t<T\}} + 1_{\{T\leq t<\tilde{T},\zeta\leq t\}} + 1_{\{T\leq t<\tilde{T},\zeta>t\}} \\
&\leq 1_{\{t<T\}} + 1_{\{\zeta\leq t\}} + 1_{\{T\leq t,t<\zeta,t<\tilde{T}\}} \\
&\leq 1_{\{t<T\}} + 1_{\{\zeta\leq t\}} + 1_{\{T<\zeta,t<\tilde{T}\}} \\
&\leq 1_{\{t<T\}} + 1_{\{\zeta\leq t\}} + 1_{\{T=\tilde{T},t<\tilde{T}\}} \\
&\leq 1_{\{t<T\}} + 1_{\{\zeta\leq t\}} + 1_{\{t<T\}} \\
&= 2 \cdot 1_{\{t<T\}} + 1_{\{\zeta\leq t\}}.
\end{aligned}$$

It follows that

$$\begin{aligned}
|(P_t f)(\tilde{x}, \tilde{k}) - (P_t f)(x, k)| &= |\mathbb{E}[f(\tilde{X}(t), \tilde{\Lambda}(t))] - \mathbb{E}[f(X(t), \Lambda(t))]| \\
&\leq \mathbb{E}[|f(\tilde{X}(t), \tilde{\Lambda}(t)) - f(X(t), \Lambda(t))| 1_{\{t<\tilde{T}\}}] \\
&\quad + \mathbb{E}[|f(\tilde{X}(t), \tilde{\Lambda}(t)) - f(X(t), \Lambda(t))| 1_{\{t\geq\tilde{T}\}}]
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E}[|f(\tilde{X}(t), \tilde{\Lambda}(t)) - f(X(t), \Lambda(t))|1_{\{t<\tilde{T}\}}] \\
&\le 2||f||_\infty \mathbb{E}[1_{\{t<\tilde{T}\}}] \\
&\le 2||f||_\infty \mathbb{E}[2 \cdot 1_{\{t<T\}} + 1_{\{\zeta \le t\}}] \\
&= 4||f||_\infty \mathbb{P}\{t < T\} + 2||f||_\infty \mathbb{P}\{\zeta \le t\}.
\end{aligned}
$$

A combination of (3.41) and (3.45) then gives

$$
\lim_{|\tilde{x}-x|\to 0} |(P_t f)(\tilde{x}, \tilde{k}) - (P_t f)(x, k)| = 0.
$$

This establishes the strong Feller property and completes the proof. □

3.3 Examples

Example 3.3.1. *Consider the following SDE*

$$
\begin{aligned}
&dX(t) = b(X(t), \Lambda(t))dt + \sigma(X(t), \Lambda(t))dW(t) + \int_U c(X(t^-), \Lambda(t^-), u)\tilde{N}(dt, du), \\
&X(0) = x \in \mathbb{R}^2,
\end{aligned}
\tag{3.48}
$$

where W is a standard 2-dimensional Brownian motion, $\tilde{N}$ is the associated compensated Poisson random measure on $[0, \infty) \times U$ with intensity $dt\nu(du)$ in which $U = \{u \in \mathbb{R}^2 : 0 < |u| < 1\}$ and $\nu(du) := \frac{du}{|u|^{2+\delta}}$ for some $\delta \in (0, 2)$. The coefficients of (3.48) are given by

$$
\sigma(x, k) = (|x| + 1)I, \quad b(x, k) = -kx, \quad c(x, k, u) = \gamma\sqrt{k}|u|x
$$

where $\gamma > 0$ is a constant so that $\gamma^2 \int_U |u|^2 \nu(du) = 2$. The component Λ is the continuous-time stochastic process taking values in $\mathbb{S} = \{1, 2, \dots\}$ generated by $Q(x) = (q_{kl}(x))$ where

$$q_{kl}(x) = \begin{cases} \frac{k}{2^{l+k}} \frac{1}{(1+l|x|^2)} & \text{if } k \neq l \\ -\sum_{l\neq k} q_{kl}(x) & \text{otherwise.} \end{cases}$$

We make the following observations.

(i) Assumption 3.0.2 is satisfied. In deed, the coefficients satisfy all assumptions in Theorem 2.5 of Xi et al. (2019) with $\zeta(r) = 1$ and $\rho(r) = r$. So a unique non-explosive strong solution exists.

(ii) To verify Assumption 3.1.2 we compute

$$\begin{aligned} \sum_{l\in\mathbb{S}\setminus\{k\}} |q_{kl}(x) - q_{kl}(y)| &= \sum_{l\in\mathbb{S}\setminus\{k\}} \left| \frac{k}{2^{l+k}} \frac{1}{(1+l|x|^2)} - \frac{k}{2^{l+k}} \frac{1}{(1+l|y|^2)} \right| \\ &\leq \sum_{l\in\mathbb{S}} \frac{l}{2^l} \frac{||y|^2 - |x|^2|}{(1+l|x|^2)(1+l|y|^2)} \\ &= \sum_{l\in\mathbb{S}} \frac{l}{2^l} \frac{(|y|+|x|)||y|-|x||}{(1+l|x|^2)(1+l|y|^2)} \\ &\leq \sum_{l\in\mathbb{S}} \frac{l}{2^l} |y-x| \\ &= |x-y|, \end{aligned}$$

where the last inequality follows from the triangle inequality $||y| - |x|| \leq |x-y|$ and the observation that

$$\frac{|y|+|x|}{(1+l|x|^2)(1+l|y|^2)} \leq \frac{|y|}{1+l|y|^2} + \frac{|x|}{1+l|x|^2} \leq \frac{|y|}{1+|y|^2} + \frac{|x|}{1+|x|^2} \leq \frac{1}{2} + \frac{1}{2} = 1.$$

(iii) Assumption 3.2.2 also holds. Indeed, since $a(x,k) = \sigma(x,k)\sigma^T(x,k) = (|x|+1)^2 I$, for each $R > 0$, we can take $\lambda_R = 1$ and $\sigma_{\lambda_R}(x,k) = ((|x|+1)^2 - 1)^{\frac{1}{2}}$ for all $(x,k) \in \mathbb{R} \times \mathbb{S}$.

Then

$$\langle \xi, a(x,k)\xi\rangle = \langle \xi, (|x|+1)^2 I\xi\rangle = (|x|+1)^2 I|\xi|^2 \geq \lambda_R|\xi|^2, \qquad \xi \in \mathbb{R}.$$

This implies condition (3.26)*. Moreover, for all* $x, z \in \mathbb{R}$ *with* $|x| \vee |z| \leq R$ *and* $k \in \mathbb{S}$ *we have*

$$\begin{aligned}
&|\sigma_{\lambda_R}(x,k) - \sigma_{\lambda_R}(z,k)|^2 + 2\langle x-z, b(x,k)-b(z,k)\rangle + \int_U |c(x,k,u)-c(z,k,u)|^2\nu(du)\\
&= 2(\sqrt{(|x|+1)^2-1} - \sqrt{(|z|+1)^2-1})^2 - 2k|x-z|^2 + k|x-z|^2\gamma^2\int_U |u|^2\nu(du)\\
&= 2(\sqrt{(|x|+1)^2-1} - \sqrt{(|z|+1)^2-1})^2\\
&\leq 4|(|x|+1)^2 - (|z|+1)^2|\\
&= 4||x|^2 + 2|x| - |z|^2 - 2|z||\\
&\leq 4||x|^2 - |z|^2| + 8||x|-|z||\\
&= 4(|x|+|z|)||x|-|z|| + 8||x|-|z||\\
&\leq 8(R+1)|x-z|\\
&= 2\kappa_R|x-z|g(|x-z|)
\end{aligned}$$

where the first inequality follows from the inequality $|\sqrt{a}-\sqrt{b}|^2 \leq 2|a-b|$ *and we take* $\kappa_R = 4(R+1)$ *and* $g(r) = 1$*.*

Therefore, the process (X,Λ) *given by* (3.48) *is strong Feller continuous by Theorem 3.2.8.*

Example 3.3.2. *Consider the following SDE*

$$dX(t) = b(X(t),\Lambda(t))dt + \sigma(X(t),\mathit{\Lambda}(t))dW(t) + \int_U c(X(t^-),\mathit{\Lambda}(t^-),u)\tilde{N}(dt,du), \qquad (3.49)$$

$X(0) = x \in \mathbb{R}$ *where* W *is a standard* 1*-dimensional Brownian motion,* $\tilde{N}$ *is the associated compensated Poisson random measure on* $[0,\infty)\times U$ *with intensity* $dt\nu(du)$ *in which* $U =$

$\{u \in \mathbb{R} : 0 < |u| < 1\}$ *and* $\nu(du) := \frac{du}{|u|^2}$. *Note that* ν *is a* σ*-finite measure on* U *with* $\nu(U) = \infty$. *Suppose that the coefficients of* (3.49) *are given by*

$$\sigma(x,k) = x^{\frac{2}{3}} + 1, \quad b(x,k) = -\frac{x}{2k^2}, \quad c(x,k,u) = \frac{\gamma}{k}ux, \qquad (x,k) \in \mathbb{R} \times \mathbb{S},$$

where $\gamma = \frac{1}{\sqrt{2}}$. *The component* Λ *is the continuous-time stochastic process taking values in* $\mathbb{S} = \{1, 2, \dots\}$ *generated by* $Q(x) = (q_{kl}(x))$ *where*

$$q_{kl}(x) = \begin{cases} \frac{k}{3^{l+k}} \frac{1}{(1+l|x|^2)} & \text{if } k \neq l \\ -\sum_{l \neq k} q_{kl}(x) & \text{otherwise.} \end{cases}$$

We make the following observations.

(i) As in Example 3.3.1 we can show that the coefficients satisfy all conditions of Theorem 2.5 of Xi et al. (2019) and hence Assumption 3.0.2 is verified.

(ii) Similarly, we verify Assumption 3.1.2 by computing

$$\begin{aligned}
\sum_{l \in \mathbb{S} \setminus \{k\}} |q_{kl}(x) - q_{kl}(y)| &= \sum_{l \in \mathbb{S} \setminus \{k\}} \left| \frac{k}{3^{l+k}} \frac{1}{(1+l|x|^2)} - \frac{k}{3^{l+k}} \frac{1}{(1+l|y|^2)} \right| \\
&= \frac{k}{3^k} \sum_{l \in \mathbb{S} \setminus \{k\}} \frac{1}{3^l} \left| \frac{1}{1+l|x|^2} - \frac{1}{1+l|y|^2} \right| \\
&\leq \sum_{l \in \mathbb{S}} \frac{l}{3^l} \frac{||y|^2 - |x|^2|}{(1+l|x|^2)(1+l|y|^2)} \\
&= \sum_{l \in \mathbb{S}} \frac{l}{3^l} \frac{(|y|+|x|)||y|-|x||}{(1+l|x|^2)(1+l|y|^2)} \\
&\leq \sum_{l \in \mathbb{S}} \frac{l}{3^l} |y - x| \\
&= \frac{3}{4}|x - y|.
\end{aligned}$$

(iii) Assumption 3.2.2 also holds. Indeed, since $a(x,k) = \sigma^2(x,k) = x^{\frac{4}{3}} + 2x^{\frac{2}{3}} + 1$, *for each*

$R > 0$, we can take $\lambda_R = 1$ and $\sigma_{\lambda_R}(x,k) = (x^{\frac{4}{3}} + 2x^{\frac{2}{3}})^{\frac{1}{2}}$ for all $(x,k) \in \mathbb{R} \times \mathbb{S}$. Hence

$$\langle \xi, a(x,k)\xi \rangle = \xi(x^{\frac{4}{3}} + 2x^{\frac{2}{3}} + 1)\xi \geq \lambda_R |\xi|^2, \qquad \xi \in \mathbb{R}^d.$$

This verifies (3.26). Moreover, for all $x, z \in \mathbb{R}$ with $|x| \vee |z| \leq R$ and $k \in \mathbb{S}$ we have

$$\begin{aligned}
&|\sigma_{\lambda_R}(x,k) - \sigma_{\lambda_R}(z,k)|^2 + 2\langle x - z, b(x,k) - b(z,k)\rangle + \int_U |c(x,k,u) - c(z,k,u)|^2 \nu(du) \\
&\leq 2(z^{\frac{2}{3}} + x^{\frac{2}{3}} + 2)(z^{\frac{2}{3}} - x^{\frac{2}{3}}) - \frac{1}{k^2}|x - z|^2 + \frac{1}{2k^2}|x - z|^2 \\
&\leq 4(R^{\frac{2}{3}} + 1)|x - z|^{\frac{2}{3}} \\
&= 4(R^{\frac{2}{3}} + 1)|x - z| g(|x - z|),
\end{aligned}$$

where $g(r) = r^{-\frac{1}{3}}$. Note that the function g satisfies (3.27).

In view of Theorem 3.2.8, the process (X, Λ) has strong Feller property. Moreover, this example shows that our results can be applied to stochastic differential equations with non-Lipschitz coefficients. As discussed in Remark 3.1.4, we only requires the regularity of $\int_U |c(x,k,u) - c(z,k,u)|^2 \nu(du)$ not $\int_U [|c(x,k,u) - c(z,k,u)|^2 \wedge |x - z| \cdot |c(x,k,u) - c(z,k,u)|] \nu(du)$. Note also that the function g does not satisfy $\lim_{r \to 0} g(r) = 0$. So the result in Xi et al. (2019) can not be applied to this particular example. This shows that Theorem 3.2.8 makes a tremendous improvement over many existing results in the literature.

Chapter 4

Irreducibility

In this chapter we aim to answer the questions on existence and uniqueness of an invariant measure. Irreducibility plays an important role in establishing the uniqueness of an invariant measure; see, for example, Cerrai (2001) and Hairer (2016). Unfortunately, irreducibility of regime-switching jump diffusions has not been systematically investigated in the literature yet. In this work, we derive the irreducibility for regime-switching jump diffusions (Theorem 4.1.10) by using an important identity concerning the transition probability of such processes. An intermediate step, which is interesting in its own right, is to show that the sub-systems consisting of jump diffusions are irreducible under weaker conditions than those in the recent papers such as Qiao (2014) and Xi and Zhu (2019). We present in Proposition 4.1.12 a set of sufficient conditions under which a unique invariant measure for regime-switching jump diffusions exists. Finally, we show that the process (X, Λ) is φ-irreducible and proceed further to obtain the property that all compact subsets of $\mathbb{R}^d \times \mathbb{S}$ are petite for every h-skeleton chain $\{(X(nh), \Lambda(nh)) : n = 0, 1, \dots\}$ of (X, Λ).

4.1 Open Set Irreducibility

4.1.1 Open Set Irreducibility

Denote the transition probability of the process $(X, \mathit{\Lambda})$ by

$$P(t,(x,k),B\times\{l\}) := P_t 1_{B\times\{l\}}(x,k) = \mathbb{P}\{(X(t),\mathit{\Lambda}(t))\in B\times\{l\}|(X(0),\mathit{\Lambda}(0))=(x,k)\},$$

for $B\in\mathfrak{B}(\mathbb{R}^d)$ and $l\in\mathbb{S}$. The process (X,Λ) or the semigroup $\{P_t\}_{t\geq 0}$ of (3.1) is said to be *open set irreducible* or *irreducible* if for any $t>0$ and $(x,k)\in\mathbb{R}^d\times\mathbb{S}$

$$P(t,(x,k),B\times\{l\})>0$$

for all $l\in\mathbb{S}$ and all nonempty open sets $B\in\mathfrak{B}(\mathbb{R}^d)$.

We make the following assumptions:

Assumption 4.1.1. *For each $k\in\mathbb{S}$ and $x\in\mathbb{R}^d$, the stochastic differential equation*

$$\begin{aligned} X^{(k)}(t) = x &+ \int_0^t b(X^{(k)}(s),k)ds + \int_0^t \sigma(X^{(k)}(s),k)dW(s) \\ &+ \int_0^t\int_U c(X^{(k)}(s-),k,u)\widetilde{N}(ds,du) \end{aligned} \tag{4.1}$$

has a non-explosive weak solution $X^{(k)}$ with initial condition x and the solution is unique in the sense of probability law.

Assumption 4.1.2. *For any $x\in\mathbb{R}^d$ and $k\in\mathbb{S}$, we have*

$$2\langle x, b(x,k)\rangle \leq \kappa\left(|x|^2+1\right), \quad |\sigma(x,k)|^2 + \int_U |c(x,k,u)|^2\nu(du) \leq \kappa\left(|x|^2+1\right). \tag{4.2}$$

Assumption 4.1.3. *(i) There exists positive constant λ such that for each $x\in\mathbb{R}^d$ and*

$k \in \mathbb{S}$, we have

$$\langle \xi, a(x,k)\xi \rangle \quad \geq \quad \lambda |\xi|^2, \qquad \xi \in \mathbb{R}^d. \tag{4.3}$$

(ii) There exits a nonnegative function $g \in C(0,\infty)$ satisfying

$$\int_0^1 g(r)dr < \infty, \tag{4.4}$$

and

$$\begin{aligned} 2\langle x - z, b(x,k) - b(z,k)\rangle + |\sigma_\lambda(x,k) - \sigma_\lambda(z,k)|^2 \\ + \int_U |c(x,k,u) - c(z,k,u)|^2 v(du) \leq 2\kappa |x - z| g(|x - z|), \end{aligned} \tag{4.5}$$

for all $x, z \in \mathbb{R}^d$ with $|x - z| \leq \delta_0$, where δ_0 is a positive constant and σ_λ the unique symmetric nonnegative definite matrix-valued function such that $\sigma_\lambda^2(x,k) = a(x,k) - \lambda I$.

Assumption 4.1.4. *(i) There exists a positive constant $\kappa_0 > 0$ such that*

$$0 \leq q_{kl}(x) \leq \kappa_0 l 3^{-l} \tag{4.6}$$

for all $x \in \mathbb{R}^d$ and $k \neq l \in \mathbb{S}$.

(ii) For any $k, l \in \mathbb{S}$, there exist $k_0, k_1, ..., k_n \in \mathbb{S}$ with $k_i \neq k_{i+1}, k_0 = k$, and $k_n = l$ such that the set $\{x \in \mathbb{R}^d : q_{k_i k_{i+1}}(x) > 0\}$ has positive Lebesgue measure for all $i = 0, 1, \ldots, n-1$.

Remark 4.1.5. *We note that Assumption 4.1.3 is comparable to Assumption 3.2.2 in which we require the constant λ to be uniform for all $x, z \in \mathbb{R}^d$.*

Condition (4.6) is imposed to facilitate the technical analyses as we need to express the *transition probability $P(t,(x,k),\cdot)$ in a suitable form; see the proof of Theorem 4.8 of Xi et al. (2019).*

In order to obtain the irreducibility of the process (X,Λ) we first show that, for any given $k\in\mathbb{S}$, the process $X^{(k)}$ of (4.1) is strong Feller and irreducible. Then we use a result in Xi et al. (2019) to write $P(t,(x,k),B\times\{l\})$ as a convergent series in terms of sub-transition probabilities of the killed processes $\tilde{X}^{(j)}, j\in\mathbb{S}$ and the transition rates $q_{jl}(x)$. Denote the transition probability of the process $X^{(k)}$ by

$$P^{(k)}(t,x,B):=\mathbb{P}\{X^{(k)}(t)\in B|X^{(k)}(0)=x\},\quad B\in\mathfrak{B}(\mathbb{R}^d).$$

The corresponding semigroup $\{P_t^{(k)}\}_{t\geq 0}$ is said to be irreducible if $P^{(k)}(t,x,B)>0$ for all nonempty open set $B\subset\mathbb{R}^d$. We next kill the process $X^{(k)}$ with killing rate $q_k(\cdot)$ and denote the killed process by $\tilde{X}^{(k)}$, that is, we define

$$\tilde{X}^{(k)}(t)=\begin{cases}X^{(k)}(t) & \text{if}\quad t<\tau,\\ \partial & \text{if}\quad t\geq\tau,\end{cases}$$

where $\tau:=\inf\{t\geq 0:\Lambda(t)\neq\Lambda(0)\}$ and ∂ is a cemetery point added to $\mathbb{R}^d$. Then the semigroup of the killed process $\tilde{X}^{(k)}$ is given by

$$\tilde{P}_t^{(k)}f(x):=\mathbb{E}_x[f(\tilde{X}^{(k)}(t))]=\mathbb{E}\bigg[f(X^{(k)}(t))\exp\bigg\{\int_0^t q_{kk}(X^{(k)}(s))ds\bigg\}|X^{(k)}(0)=x\bigg],$$

where $f\in\mathfrak{B}_b(\mathbb{R}^d)$. We also denote its sub-transition probability by

$$\tilde{P}^{(k)}(t,x,B):=\mathbb{E}_x[1_B(\tilde{X}^{(k)}(t))]=\mathbb{P}\{\tilde{X}^{(k)}(t)\in B|\tilde{X}^{(k)}(0)=x\},\quad B\in\mathfrak{B}(\mathbb{R}^d).$$

We first obtain the following lemma. It is worth mentioning that this lemma still holds when we replace Assumption 4.1.3 by Assumption 3.2.2.

Lemma 4.1.6. *Under Assumptions 4.1.1 and 4.1.3, the semigroup $\{P_t^{(k)}\}_{t\geq 0}$ is strong Feller.*

Proof. Let $(\widetilde{X}^{(k)}, X^{(k)})$ be the coupling process corresponding to $\widehat{\mathcal{L}}_k$ of (2.28) with initial condition $(\tilde{x}, x)$. Suppose without loss of generality that $0 < |\tilde{x} - x| < \delta_0$, where δ_0 is the positive constant in Assumption 3.2.2. Define $T := \inf\{t \geq 0 : \widetilde{X}(t) = X(t)\}$. Using very similar calculations as those in the proof of Lemma 3.2.7, we can show that $\lim_{|\tilde{x}-x|\to 0} \mathbb{P}\{t < T\} = 0$. Then it follows that for any $f \in \mathscr{B}_b(\mathbb{R}^d)$ and $t > 0$, we have

$$|(P_t^{(k)} f)(\tilde{x}) - (P_t^{(k)} f)(x)| = |\mathbb{E}[f(\tilde{X}^{(k)}(t))] - \mathbb{E}[f(X^{(k)}(t))]| \leq 2||f||_\infty \mathbb{P}\{t < T\} \to 0,$$

as $\tilde{x} - x \to 0$. This implies that $P_t^{(k)} f$ is a continuous function and hence completes the proof. □

Lemma 4.1.7. *Suppose that Assumptions 4.1.2 and 4.1.3 (i) hold. Then for every $T > 0$ there exists a constant $K := K(T, X(0)) > 0$ so that*

$$\mathbb{E}[|X(t)|^2] \leq K \tag{4.7}$$

for all $t \in [0, T]$.

Proof. For any $R > 0$, we set $\tau_R := \inf\{t \geq 0 : |X(t)| > R\}$. Also note that $\lim_{R\to\infty} \tau_R = \infty$ a.s. By Ito's formula and mean theorem we see that

$$\begin{aligned}
&\mathbb{E}[|X(t \wedge \tau_R)|^2] \\
= \; & |X(0)|^2 + \mathbb{E}\left[\int_0^{t\wedge\tau_R} 2\langle X_s, b(X(s), k)\rangle ds\right] + \mathbb{E}\left[\int_0^{t\wedge\tau_R} ||\sigma(X(s), k)||^2 ds\right] \\
&+\mathbb{E}\left[\int_0^{t\wedge\tau_R} \int_U \left(|X(s) + c(X(s), k, u)|^2 - |X_s|^2 - 2\langle X(s), c(X(s), k, u)\rangle\right) \nu(du) ds\right] \\
= \; & |X(0)|^2 + \mathbb{E}\left[\int_0^{t\wedge\tau_R} 2\langle X_s, b(X(s), k)\rangle ds\right] + \mathbb{E}\left[\int_0^{t\wedge\tau_R} ||\sigma(X(s), k)||^2 ds\right] \\
&+\mathbb{E}\left[\int_0^{t\wedge\tau_R} \int_U |c(X(s), k, u)|^2 \nu(du) ds\right] \\
\leq \; & |X(0)|^2 + \mathbb{E}\left[\int_0^{t\wedge\tau_R} \kappa\left(|X(s)|^2 + 1\right) ds\right].
\end{aligned}$$

where we use (4.2) to obtain the inequality. Then

$$\mathbb{E}[1+|X(t\wedge\tau_R)|^2] \le 1+|X(0)|^2+\kappa\mathbb{E}\bigg[\int_0^{t\wedge\tau_R}(1+|X(s)|^2)ds\bigg].$$

Now letting $R\to\infty$ and using Fatou's lemma and the monotone convergence theorem to obtain

$$\mathbb{E}[1+|X(t)|^2] \le 1+|X(0)|^2+\kappa\int_0^t \mathbb{E}\Big[(1+|X(s)|^2)\Big]ds.$$

So we can use Gronwall's inequality to obtain

$$\begin{aligned}\mathbb{E}[1+|X(t)|^2] &\le [1+|X(0)|^2]e^{\kappa t}\\ &\le [1+|X(0)|^2]e^{\kappa T}.\end{aligned}$$

Therefore,

$$\mathbb{E}[|X(t)|^2] \le [1+|X(0)|^2]e^{\kappa T}-1.$$

□

To derive irreducibility for the semigroup $\{P_t^{(k)}\}_{t\ge0}$, we consider the function F given by

$$F(r) := \int_0^{\frac{r}{1+r}} e^{-\int_0^s g(w)dw}ds, \qquad r\in[0,\infty) \tag{4.8}$$

where g is the function given in Assumption 3.2.2(ii). Since $g\ge 0$, we see that

$$0\le F(r)\le \frac{r}{1+r}\le 1 \tag{4.9}$$

$$0\le F'(r) = \frac{1}{(1+r)^2}e^{-\int_0^{\frac{r}{1+r}}g(w)dw} \le \frac{1}{(1+r)^2}\le 1 \tag{4.10}$$

$$0\ge F''(r) = -\frac{2}{(1+r)^3}e^{-\int_0^{\frac{r}{1+r}}g(w)dw} - \frac{g(\frac{r}{1+r})}{(1+r)^4}e^{-\int_0^{\frac{r}{1+r}}g(w)dw} = -\left[\frac{2}{1+r}+\frac{g(\frac{r}{1+r})}{(1+r)^2}\right]F'(r). \tag{4.11}$$

In addition, for any $x \in \mathbb{R}^d$, we have

$$\nabla F(|x|^2) = 2F'(|x|^2)x, \qquad \nabla^2 F(|x|^2) = 4F''(|x|^2)xx^T + 2F'(|x|^2)I.$$

Lemma 4.1.8. *Under Assumptions 4.1.2 and 4.1.3, the semigroup $\{P_t^{(k)}\}_{t\geq 0}$ is irreducible.*

Proof. Let $T > 0, r > 0$ and $x, a \in \mathbb{R}^d$ be arbitrary but fixed. We will show that

$$P^{(k)}(T, x, B(a; r)) = \mathbb{P}\{|X^{(k)}(T) - a| < r | X^{(k)}(0) = x\} > 0,$$

or equivalently, $\mathbb{P}\{|X^{(k)}(T) - a| \geq r | X^{(k)}(0) = x\} < 1$. Let us choose some $t_0 \in (0, T)$. For any $n \in \mathbb{N}$, we set $X_n^{(k)}(t_0) := X^{(k)}(t_0) 1_{\{|X^{(k)}(t_0)| \leq n\}}$. Since $\lim_{r\to 0} F(r) = 0$ and $0 \leq F \leq 1$, the bounded convergence implies that

$$\lim_{n\to\infty} \mathbb{E}[F(|X_n^{(k)}(t_0) - X^{(k)}(t_0)|^2)] = 0. \tag{4.12}$$

For $t \in [t_0, T]$, define

$$J^n(t) := \frac{T-t}{T-t_0} X_n^{(k)}(t_0) + \frac{t-t_0}{T-t_0} a, \quad \text{and} \quad h^n(t) := \frac{a - X_n^{(k)}(t_0)}{T-t_0} - b(J^n(t), k).$$

We see that $J^n(t_0) = X_n^{(k)}(t_0)$ and $J^n(T) = a$. In addition, J^n satisfies the following stochastic differential equation

$$J^n(t) = X_n^{(k)}(t_0) + \int_{t_0}^t b(J^n(s), k) ds + \int_{t_0}^t h^n(s) ds, \qquad t \in [t_0, T].$$

Consider the stochastic differential equation

$$\begin{aligned} Y(t) = X^{(k)}(t_0) &+ \int_{t_0}^t [b(Y(s), k) + h^n(s)] ds + \int_{t_0}^t \sigma(Y(s), k) dW(s) \\ &+ \int_{t_0}^t \int_U c(Y(s), k, u) \tilde{N}(ds, du), \qquad t \in [t_0, T]. \end{aligned} \tag{4.13}$$

Also denote $\Delta_t := Y(t) - J^n(t)$ for $t \in [t_0, T]$. Note that $\Delta_{t_0} = X^{(k)}(t_0) - X_n^{(k)}(t_0)$ and $\Delta_T = Y(T) - a$. In addition, Δ_t satisfies the stochastic differential equation

$$\begin{aligned}\Delta_t = \Delta_{t_0} &+ \int_{t_0}^t [b(Y(s),k) - b(J^n(s),k)]ds + \int_{t_0}^t \sigma(Y(s),k)dW(s) \\ &+ \int_{t_0}^t \int_U c(Y(s),k,u)\tilde{N}(ds,du).\end{aligned}$$

Consequently the generator of the process Δ_t is given by

$$\begin{aligned}\mathcal{L}f(x) &= \mathcal{L}_{\mathrm{d}}f(x) + \mathcal{L}_{\mathrm{j}}f(x) \\ &:= \frac{1}{2}\mathrm{tr}\left(\sigma(Y(s),k)\sigma(Y(s),k)^T\nabla^2 f(x)\right) + \langle b(Y(s),k) - b(J^n(s),k), \nabla f(x)\rangle \\ &\quad + \int_U \left(f(x + c(Y(s),k,u)) - f(x) - \langle \nabla f(x), c(Y(s),k,u)\rangle\right)\nu(du), \quad f \in C_c^2(\mathbb{R}^d).\end{aligned}$$

We compute

$$\begin{aligned}\mathcal{L}_{\mathrm{d}}F(|\Delta_s|^2) &= \frac{1}{2}\mathrm{tr}\left(\sigma(Y(s),k)\sigma(Y(s),k)^T\nabla^2 F(|\Delta_s|^2)\right) + \langle b(Y(s),k) - b(J^n(s),k), \nabla F(|\Delta_s|^2)\rangle \\ &= \frac{1}{2}\mathrm{tr}\left(\sigma(Y(s),k)\sigma(Y(s),k)^T\left[4F''(|\Delta_s|^2)\Delta_s\Delta_s^T + 2F'(|\Delta_s|^2)I\right]\right) \\ &\quad + \langle b(Y(s),k) - b(J^n(s),k), 2F'(|\Delta_s|)\Delta_s\rangle \\ &= 2F''(|\Delta_s|^2)|\Delta_s^T\sigma(Y(s),k)|^2 + F'(|\Delta_s|)|\sigma(Y(s),k)|^2 \\ &\quad + 2F'(|\Delta_s|^2)\langle b(Y(s),k) - b(J^n(s),k), \Delta_s\rangle \\ &\le F'(|\Delta_s|)\left[|\sigma(Y(s),k)|^2 + 2\langle b(Y(s),k) - b(J^n(s),k), \Delta_s\rangle\right] \\ &\le |\sigma(Y(s),k)|^2 + 2\langle b(Y(s),k) - b(J^n(s),k), \Delta_s\rangle,\end{aligned}$$

where the inequalities follow from (4.10) and (4.11). Likewise, the concavity of F leads to

$$\begin{aligned}\mathcal{L}_{\mathrm{j}}F(|\Delta_s|^2) &= \int_U \left(F(|\Delta_s + c(Y(s),k,u)|^2) - F(|\Delta_s|^2) - \langle \nabla F(|\Delta_s|^2), c(Y(s),k,u)\rangle\right)\nu(du) \\ &\le \int_U \left[F'(|\Delta_s|^2)[|\Delta_s + c(Y(s),k,u)|^2 - |\Delta_s|^2] - 2F'(|\Delta_s|^2)\langle \Delta_s, c(Y(s),k,u)\rangle\right]\nu(du)\end{aligned}$$

$$= \int_U F'(|\Delta_s|^2)|c(Y(s),k,u)|^2\nu(du)$$
$$\leq \int_U |c(Y(s),k,u)|^2\nu(du).$$

Therefore by adding the above two inequalities, we have

$$\mathcal{L}F(|\Delta_s|^2) \leq |\sigma(Y(s),k)|^2 + 2\langle b(Y(s),k) - b(J^n(s),k), \Delta_s\rangle + \int_U |c(Y(s),k,u)|^2\nu(du).$$

Furthermore, when $|Y(s)| \leq R$, $|J^n(s)| \leq R$ and $|\Delta_s| \leq \delta_0$, we can use (4.2) and (3.28) to obtain

$$\mathcal{L}F(|\Delta_s|^2) \leq \kappa(|Y(s)|^2+1) + 2\kappa|\Delta_s|g(|\Delta_s|) \leq K_0 + \kappa|Y(s)|^2,$$

where $K_0 = \kappa + 2\kappa \max_{r\in[0,\delta_0]}\{rg(r)\} < \infty$. In view of (4.2) and (4.7), we can use the standard arguments to show that $\mathbb{E}[\sup_{t_0\leq s\leq T}|Y(s)|^2] \leq K$, where K is a positive constant independent of t_0. For any $R>0$, we define $\tau_R := \inf\{t \geq t_0 : |Y(t)| \vee |J^n(t)| > R\} \wedge T$ and $S_{\delta_0} := \inf\{t \geq t_0 : |Y(t) - J^n(t)| \geq \delta_0\} \wedge T$. Then we can compute

$$\begin{aligned}
\mathbb{E}[F(|\Delta_{T\wedge\tau_R\wedge S_{\delta_0}}|^2)] &= \mathbb{E}[F(|\Delta_{t_0}|^2)] + \mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R\wedge S_{\delta_0}} \mathcal{L}F(|\Delta_{s^-}|^2)ds\right] \\
&\leq \mathbb{E}[F(|\Delta_{t_0}|^2)] + \mathbb{E}\left[\int_{t_0}^{T\wedge\tau_R\wedge S_{\delta_0}} (K_R + \kappa|Y(s^-)|^2)ds\right] \\
&\leq \mathbb{E}[F(|\Delta_{t_0}|^2)] + K_0(T-t_0) + \mathbb{E}\left[\int_{t_0}^{T} \kappa|Y(s)|^2 ds\right] \\
&\leq \mathbb{E}[F(|\Delta_{t_0}|^2)] + (K_0+K)(T-t_0).
\end{aligned} \tag{4.14}$$

Next we show that

$$\mathbb{E}[F(|\Delta_T|^2)] \leq \frac{1}{F(\delta_0^2)}\mathbb{E}[F(|\Delta_{T\wedge S_{\delta_0}}|^2)]. \tag{4.15}$$

Indeed, we observe that $|\Delta_{T\wedge S_{\delta_0}\wedge\tau_R}| \geq \delta_0$ on the set $\{S_{\delta_0} < T\wedge\tau_R\}$. Since F is increasing, we have $F(\delta_0^2) \le F(|\Delta_{T\wedge S_{\delta_0}}|^2)$. This together with the fact that $0 \le F \le 1$ give the following

$$\begin{aligned}
&\frac{\mathbb{E}[F(|\Delta_{T\wedge\tau_R\wedge S_{\delta_0}}|^2)]}{F(\delta_0^2)} - \mathbb{E}[F(|\Delta_{T\wedge\tau_R}|^2)] \\
&= \frac{\mathbb{E}[F(|\Delta_{T\wedge\tau_R\wedge S_{\delta_0}}|^2)1_{\{T\wedge\tau_R\le S_{\delta_0}\}}] + \mathbb{E}[F(|\Delta_{T\wedge\tau_R\wedge S_{\delta_0}}|^2)1_{\{T\wedge\tau_R> S_{\delta_0}\}}]}{F(\delta_0^2)} - \mathbb{E}[F(|\Delta_{T\wedge\tau_R}|^2)] \\
&\ge \frac{\mathbb{E}[F(|\Delta_{T\wedge\tau_R}|^2)1_{\{T\wedge\tau_R\le S_{\delta_0}\}}] + F(\delta_0^2)\mathbb{P}\{T\wedge\tau_R > S_{\delta_0}\}}{F(\delta_0^2)} - \mathbb{E}[F(|\Delta_{T\wedge\tau_R}|^2)] \\
&\ge \mathbb{P}\{T\wedge\tau_R > S_{\delta_0}\} + \mathbb{E}[F(|\Delta_{T\wedge\tau_R}|^2)1_{\{T\wedge\tau_R\le S_{\delta_0}\}}] - \mathbb{E}[F(|\Delta_{T\wedge\tau_R}|^2)] \\
&= \mathbb{P}\{T\wedge\tau_R > S_{\delta_0}\} - \mathbb{E}[F(|\Delta_{T\wedge\tau_R}|^2)1_{\{T\wedge\tau_R> S_{\delta_0}\}}] \\
&\ge \mathbb{P}\{T\wedge\tau_R > S_{\delta_0}\} - \mathbb{E}[1\cdot 1_{\{T\wedge\tau_R> S_{\delta_0}\}}] = 0.
\end{aligned}$$

Consequently we have $\mathbb{E}[F(|\Delta_{T\wedge\tau_R}|^2)] \le \frac{\mathbb{E}[F(|\Delta_{T\wedge\tau_R\wedge S_{\delta_0}}|^2)]}{F(\delta_0^2)}$. Since $\lim_{R\to\infty}\tau_R = \infty$ a.s. and $0 \le F \le 1$, the bounded convergence theorem gives (4.15).

Recall that Y satisfies the stochastic differential equation (4.13) for $t \in [t_0, T]$. For $t \in [0, t_0]$, we define $Y(t) := X^{(k)}(t)$ and $X^{(k)}(t)$ is the weak solution to (4.1) with initial condition x. Then the process Y satisfies the following stochastic differential equation:

$$Y(t) = x + \int_0^t [b(Y(s), k) + h^n(s)1_{\{s>t_0\}}]ds + \int_0^t \sigma(Y(s), k)dW(s) + \int_0^t \int_U c(Y(s), k, u)\tilde{N}(ds, du)$$

for $t \in [0, T]$. Next we set

$$\begin{aligned}
H(t) &:= 1_{\{t>t_0\}}\sigma^{-1}(Y(t), k)h^n(t), \\
M(t) &:= \exp\left\{ \int_0^t \langle H(s), dW(s)\rangle - \frac{1}{2}\int_0^t |H(s)|^2 ds \right\}.
\end{aligned}$$

As argued in Qiao (2014), it follows from (4.3) that $|H(t)|^2$ is bounded and hence M is a martingale under $\mathbb{P}$ by Novikov's criteria. Moreover, $\mathbb{E}[M(T)] = 1$. Define

$$\mathbb{Q}(B) := \mathbb{E}[M(T)1_{\{B\}}], \quad B \in \mathcal{F}_T$$

$$\tilde{W}(t) := W(t) + \int_0^t H(s)ds.$$

It follows from Theorem 132 of Situ (2005) that $\mathbb{Q}$ is a probability measure, $\tilde{W}$ is a $\mathbb{Q}$-Brownian motion and $\tilde{N}(dt,du)$ is a $\mathbb{Q}$-compensated Poisson random measure with compensator $dt\nu(du)$. Furthermore, under the measure $\mathbb{Q}$, Y solves the following stochastic differential equation

$$Y(t) = x + \int_0^t b(Y(s),k)ds + \int_0^t \sigma(Y(s),k)d\tilde{W}(s) + \int_0^t \int_U c(Y(s),k,u)\tilde{N}(ds,du)$$

for $t \in [0,T]$. By the uniqueness in law of the solution to the SDE, we have that the law of $\{X^{(k)}(t) : t \in [0,T]\}$ under $\mathbb{P}$ is the same as the law of $\{Y(t) : t \in [0,T]\}$ under $\mathbb{Q}$. In particular, we have $\mathbb{P}\{|X^{(k)}(T) - a| \geq r | X^{(k)}(0) = x\} = \mathbb{Q}\{|Y(T) - a| \geq r | Y(0) = x\}$. Since $\mathbb{P}$ and $\mathbb{Q}$ are equivalent, the desired assertion $\mathbb{P}\{|X^{(k)}(T) - a| \geq r | X^{(k)}(0) = x\} = \mathbb{Q}\{|Y(T) - a| \geq r | Y(0) = x\} < 1$ will follow if we can show that $\mathbb{P}\{|Y(T) - a| \geq r | Y(0) = x\} < 1$. To this end, for any $\varepsilon > 0$, we first choose an $R > 0$ sufficiently large so that $\mathbb{P}\{\tau_R < T\} < \varepsilon$. Next we use the facts that the function F is bounded and increasing, (4.15) and (4.14) to compute

$$\begin{aligned}
\mathbb{P}\{|Y(T) - a| \geq r | Y(0) = x\} &= \mathbb{P}\{|Y(T) - a|^2 \geq r^2 | Y(0) = x\} \\
&= \mathbb{P}\{F(|Y(T) - a|^2) \geq F(r^2) | Y(0) = x\} \\
&\leq \frac{\mathbb{E}[F(|Y(T) - a|^2)]}{F(r^2)} = \frac{\mathbb{E}[F(|\Delta_T|^2)]}{F(r^2)} \\
&\leq \frac{\mathbb{E}[F(|\Delta_{T \wedge S_{\delta_0}}|^2)]}{F(r^2)F(\delta_0^2)} \\
&= \frac{\mathbb{E}[F(|\Delta_{T \wedge S_{\delta_0} \wedge \tau_R}|^2) 1_{\{\tau_R \geq T \wedge S_{\delta_0}\}}] + \mathbb{E}[F(|\Delta_{T \wedge S_{\delta_0}}|^2) 1_{\{\tau_R < T \wedge S_{\delta_0}\}}]}{F(r^2)F(\delta_0^2)} \\
&\leq \frac{\mathbb{E}[F(|\Delta_{t_0}|^2)] + (K_0 + K)(T - t_0) + \mathbb{P}\{\tau_R < T\}}{F(r^2)F(\delta_0^2)} \\
&\leq \frac{\mathbb{E}[F(|\Delta_{t_0}|^2)] + (K_0 + K)(T - t_0) + \varepsilon}{F(r^2)F(\delta_0^2)}.
\end{aligned}$$

Note that by virtue of (4.12), $\mathbb{E}[F(|\Delta_{t_0}|^2)] \to 0$ as $n \to \infty$. Therefore we can choose n sufficiently large and t_0 close enough to T to make the last term less than 1 as desired. □

Remark 4.1.9. *While irreducibility for jump diffusions has been considered in the literature such as Qiao (2014), Xi and Zhu (2019), it is worth pointing out that Assumption 3.2.2(ii) is much weaker than Assumptions (H_1') and (H_f') of Qiao (2014) and Assumption 2.5 of Xi and Zhu (2019). In particular, as we mentioned in Remark 3.2.3, Assumption 4.1.3(ii) allows to treat SDEs with merely Hölder continuous coefficients. The relaxations make the analyses more involved and subtle than those in the literature.*

Theorem 4.1.10. *Suppose that Assumptions 4.1.1, 4.1.2, 4.1.3, and 4.1.4 hold. Then the semigroup $\{P_t\}_{t\geq 0}$ of* (3.1) *is irreducible.*

Proof. Given $t > 0$ and $(x, k) \in \mathbb{R}^d \times \mathbb{S}$. We want to show that $P(t, (x, k), B \times \{l\}) > 0$ for all $l \in \mathbb{S}$ and all open sets $B \in \mathfrak{B}(\mathbb{R}^d)$ with positive Lebesgue measure. Under Assumption 4.1.4 and from Lemma 4.1.6, as in the proof of Theorem 4.8 of Xi et al. (2019), we can write

$$\begin{aligned} &P(t, (x, k), B \times \{l\}) \\ &= \delta_{kl}\tilde{P}^{(k)}(t, x, B) + \sum_{m=1}^{\infty} \underset{0<t_1<\cdots<t_m<t}{\int \cdots \int} \sum_{\substack{l_0, l_1, l_2, \ldots, l_m \in \mathbb{S} \\ l_i \neq l_{i+1}, l_0 = k, l_m = l}} \underset{\mathbb{R}^d}{\int} \cdots \underset{\mathbb{R}^d}{\int} \tilde{P}^{(l_0)}(t_1, x, dy_1) q_{l_0 l_1}(y_1) \\ &\quad \times \tilde{P}^{(l_1)}(t_2 - t_1, y_1, dy_2) \cdots q_{l_{m-1} l_m}(y_m) \tilde{P}^{(l_m)}(t - t_m, y_m, B) dt_1 dt_2 \cdots dt_m, \end{aligned} \tag{4.16}$$

where δ_{kl} is the Kronecker symbol. From Assumption 4.1.4 (ii), we know that the set $\{y \in \mathbb{R}^d : q_{l_i l_{i+1}}(y) > 0\}$ has positive Lebesgue measure. Then it suffices to show that $\tilde{P}^{(k)}(s, y, B) > 0$ for all $k \in \mathbb{S}$, $s > 0$ and all open sets $B \in \mathfrak{B}(\mathbb{R}^d)$ with positive Lebesgue measure. We calculate

$$\begin{aligned} \tilde{P}^{(k)}(s, y, B) &= \mathbb{P}\{\tilde{X}_y^{(k)}(s) \in B\} \\ &= \mathbb{E}_k \left[1_B(X_y^{(k)}(s)) \exp\left(-\int_0^s q_k(X_y^{(k)}(r)) dr \right) \right] \\ &\geq \mathbb{E}_k \left[1_B(X_y^{(k)}(s)) e^{-M} \right] \end{aligned}$$

$$\geq e^{-M}\mathbb{P}\{X_y^{(k)}(s) \in B\}$$
$$= e^{-M}P^{(k)}(s, y, B).$$

From Lemma 4.1.8, the semigroup associated with the process $X^{(k)}$ is irreducible and therefore $P^{(k)}(s, y, B) > 0$. This completes the proof. □

4.1.2 Existence and Uniqueness of Invariant Measures

In this section we study existence and uniqueness of an invariant measure of the semigroup $\{P_t\}_{t\geq 0}$. We first obtain the following result which can be proved in the same manner as in Proposition 6.1 of Xi and Zhu (2017). Similar result for the finite regimes case can also be found in Theorem 3.3 of Xi (2004).

Proposition 4.1.11. *Suppose Assumptions 3.1.1 and 3.1.2 hold. In addition, assume there exist constants $\alpha, \beta > 0$, a compact subset $C \subset \mathbb{R}^d$, a compact subset $N \subset \mathbb{S}$, a measurable function $f : \mathbb{R}^d \times \mathbb{S} \to [1, \infty)$, and a twice continuously differentiable function $V : \mathbb{R}^d \times \mathbb{S} \to [0, \infty)$ such that*

$$\mathscr{A}V(x, k) \leq -\alpha f(x, k) + \beta 1_{C\times N}(x, k), \quad \forall (x, k) \in \mathbb{R}^d \times \mathbb{S}. \tag{4.17}$$

Then the semigroup $\{P_t\}_{t\geq 0}$ of (3.1) *has an invariant probability measure π.*

Proof. Thanks to Theorem 3.1.6, under Assumptions 3.1.1 and 3.1.2, the semigroup P_t possesses Feller property. From (4.17) we observe that

$$\begin{aligned}
0 &\leq \mathbb{E}_{(x,k)}\left[V(X(t\wedge\tau_R), \Lambda(t\wedge\tau_R))\right] \\
&= V(x, k) + \mathbb{E}_{(x,k)}\left[\int_0^{t\wedge\tau_R} \mathscr{A}V(X(s), \Lambda(s))ds\right] \\
&\leq V(x, k) + \mathbb{E}_{(x,k)}\left[\int_0^{t\wedge\tau_R} \left(-\alpha f(X(s), \Lambda(s)) + \beta I_{C\times N}(X(s), \Lambda(s))\right) ds\right] \\
&= V(x, k) - \alpha\mathbb{E}_{(x,k)}\left[\int_0^{t\wedge\tau_R} f(X(s), \Lambda(s))ds\right] + \beta\mathbb{E}_{(x,k)}\left[\int_0^{t\wedge\tau_R} I_{C\times N}(X(s), \Lambda(s))ds\right].
\end{aligned}$$

Since $f \geq 1$ we have

$$\begin{aligned}
\alpha \mathbb{E}_{(x,k)}\left[t \wedge \tau_R\right] &\leq \alpha \mathbb{E}_{(x,k)}\left[\int_0^{t \wedge \tau_R} f(X(s), \Lambda(s)) ds\right] \\
&\leq V(x,k) + \beta \int_0^{t \wedge \tau_R} \mathbb{E}_{(x,k)}\left[I_{C \times N}(X(s), \Lambda(s))\right] ds \\
&\leq V(x,k) + \beta \int_0^t P(s,(x,k), C \times N) ds.
\end{aligned}$$

Letting $R \longrightarrow \infty$ we obtain

$$\alpha t \leq V(x,k) + \beta \int_0^t P(s,(x,k), C \times N) ds$$

and hence

$$\frac{\alpha}{\beta} \leq \frac{V(x,k)}{\beta t} + \frac{1}{t} \int_0^t P(s,(x,k), C \times N) ds.$$

This implies that

$$\frac{\alpha}{\beta} \leq \liminf_{t \longrightarrow \infty} \frac{1}{t} \int_0^t P(s,(x,k), C \times N) ds. \tag{4.18}$$

As in the proof of Theorem 4.5 of Meyn and Tweedie (1993c), the author states the following result from Foguel (1969) and Stettner (1986). For any Feller process, there are two mutually exclusive possibilities: either an invariant probability measure exists, or

$$\lim_{t \longrightarrow \infty} \sup_{\mu} \frac{1}{t} \int_0^t \int P(s,(x,k), C \times N) \mu(dx, dk) ds = 0 \tag{4.19}$$

for any compact set $C \times N \subset \mathbb{R}^d \times \mathbb{S}$, where the supremum is taken over all initial distributions μ on the state space $\mathbb{R}^d \times \mathbb{S}$. From (4.18), we know that (4.19) is impossible, then an invariant probability measure π exists. $\square$

Proposition 4.1.12. *Suppose Assumptions 3.1.2, 4.1.1, 4.1.2, 4.1.3, and 4.1.4 hold. If* *there exists a twice continuously differentiable function* $V : \mathbb{R}^d \times \mathbb{S} \to [0, \infty)$ *such that* (4.17) *holds, then the the semigroup* $\{P_t\}_{t\geq 0}$ *of* (3.1) *has a unique invariant measure.*

Proof. The existence follows from Proposition 4.1.11. The semigroup $\{P_t\}_{t\geq 0}$ is strong Feller and irreducible by Theorems 3.2.8 and 4.1.10, respectively. From the classical result; see, for example, Cerrai (2001) and Hairer (2016), if a semigroup $\{P_t\}_{t\geq 0}$ is irreducible with strong Feller property then it can admit at most one invariant measure. This completes the proof. □

4.2 φ-Irreducibility and Petite Sets

Let $h > 0$ be a constant and consider the h-skeleton chain $\{(X(nh), \Lambda(nh)) : n = 0, 1, \dots\}$. It is worth notice that the transition kernel of the embedded chain is given by $P(h, (x, k), A)$. As in Meyn and Tweedie (1992, 1993b) we say that the h-skeleton chain $\{(X(nh), \Lambda(nh)) : n = 0, 1, \dots\}$ is *φ-irreducible* if φ is a σ-finite measure on $\mathfrak{B}(\mathbb{R}^d \times \mathbb{S})$ and

$$\varphi(A) > 0 \Rightarrow \sum_{n=1}^{\infty} P(nh, (x, k), A) > 0 \quad \text{for all} \quad (x, k) \in \mathbb{R}^d \times \mathbb{S}.$$

We obtain the following result as a direct consequence of Propositions 6.1.5 and 6.1.6 of Meyn and Tweedie (2009). For the sake of completeness we give the proof here.

Proposition 4.2.1. *Suppose that Assumptions 3.1.2, 4.1.1, 4.1.2, 4.1.3, and 4.1.4 hold. Then the h-skeleton chain* $\{(X(nh), \Lambda(nh)) : n = 0, 1, \dots\}$ *is φ-irreducible where* $\varphi = P(h, (x, k), \cdot)$.

Proof. Thanks to Theorem 3.2.8 the process (X, Λ) has strong Feller property and so does the chain $\{(X(nh), \Lambda(nh)) : n = 0, 1, \dots\}$. Then $P(h, \cdot, A)$ is lower semicontinuous for every $A \in \mathcal{B}(\mathbb{R}^d \times \mathbb{S})$; see, for example, Proposition 6.1.1 of Meyn and Tweedie (2009). Given a measurable set $A \in \mathcal{B}(\mathbb{R}^d \times \mathbb{S})$ with $P(h, (x, k), A) > 0$. Since $P(h, \cdot, A)$ is lower semicontinuous then there exists a neighborhood U of (x, k) such that $P(h, (z, j), A) > 0$

for all $(z, j) \in U$. From Theorem 4.1.10, the semigroup (P_t) is open set irreducible and hence every point in $\mathbb{R}^d \times \mathbb{S}$ is reachable. In particular, the (x, k) is reachable. For any $(y, i) \in \mathbb{R}^d \times \mathbb{S}$ there exists $n \geq 1$ such that $P(nh, (y, i), U) > 0$. Then we have

$$P((n+1)h, (y, i), A) \geq \int_U P(nh, (y, i), dz \times dj) P(h, (z, j), A) > 0.$$

Summing this up gives $\sum_{n=1}^{\infty} P(nh, (y, i), A) > 0$. This completes the proof. □

As in Meyn and Tweedie (1992, 1993b), a set $B \in \mathfrak{B}(\mathbb{R}^d \times \mathbb{S})$ and a sub-probability measure φ on $\mathfrak{B}(\mathbb{R}^d \times \mathbb{S})$ are called *petite* for the h-skeleton chain $\{(X(nh), \Lambda(nh)) : n = 0, 1, \dots\}$ if for some probability a on $\mathbb{Z}_+$, we have

$$K_a((x, k), \cdot) := \sum_{n=1}^{\infty} a(n) P(nh, (x, k), \cdot) \geq \varphi(\cdot) \quad \text{for all } (x, k) \in B.$$

According to Theorem 2.1.2 (ii), if the semigroup (P_t) is φ-irreducible with Feller property and if the support $\operatorname{supp} \varphi$ has non-empty interior, then all compact subsets of $\mathbb{R}^d \times \mathbb{S}$ are petite. Then we obtain the following result.

Proposition 4.2.2. *Suppose that Assumptions 3.1.2, 4.1.1, 4.1.2, 4.1.3, and 4.1.4 hold. Then all compact sets of $\mathbb{R}^d \times \mathbb{S}$ are petite for any h-skeleton chain of (X, Λ).*

Proof. The h-skeleton chain $\{(X(nh), \Lambda(nh)) : n = 0, 1, \dots\}$ has strong Feller property by Theorem 3.2.8. Moreover, it is φ-irreducible where $\varphi = P(h, (x, k), \cdot)$ by Proposition 4.2.1. Theorem 4.1.10 ensures that this chain is open set irreducible and hence every point in $\mathbb{R}^d \times \mathbb{S}$ is reachable. Then we have $supp(\varphi) = \mathbb{R}^d \times \mathbb{S}$; see Lemma 6.1.4 of Meyn and Tweedie (2009). Therefore, every compact subset of $\mathbb{R}^d \times \mathbb{S}$ is petite by Theorem 2.1.2 (ii). □

4.3 Examples

Example 4.3.1. *Consider the following SDE*

$$
\begin{aligned}
dX(t) &= b(X(t),\Lambda(t))dt + \sigma(X(t),\Lambda(t))dW(t) + \int_U c(X(t^-),\Lambda(t^-),u)\tilde{N}(dt,du),\\
X(0) &= x \in \mathbb{R}^2,
\end{aligned}
\tag{4.20}
$$

where W is a standard 2-dimensional Brownian motion, $\tilde{N}$ is the associated compensated Poisson random measure on $[0,\infty)\times U$ with intensity $dt\nu(du)$ in which $U = \{u \in \mathbb{R}^2 : 0 < |u| < 1\}$ and $\nu(du) := \frac{du}{|u|^{2+\delta}}$ for some $\delta \in (0,2)$. The coefficients of (4.20) *are given by*

$$
\sigma(x,k) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad b(x,k) = -2kx, \qquad c(x,k,u) = \gamma\sqrt{k}|u|x
$$

where γ is a positive constant so that $\gamma^2 \int_U |u|^2\nu(du) = 2$. The component Λ is the continuous-time stochastic process taking values in $\mathbb{S} = \{1,2,\dots\}$ generated by $Q(x) = (q_{kl}(x))$

$$
q_{kl}(x) = \begin{cases} \frac{k}{3^{l+k}}\frac{1}{(1+l|x|^2)} & \text{if } k \neq l \\ -\sum_{l\neq k} q_{kl}(x) & \text{otherwise.} \end{cases}
$$

Detailed calculations reveal that (4.20) *has a unique non-explosive weak solution. Moreover, all assumption in Proposition 4.1.12 are satisfied; that is the solution is strong Feller continuous and irreducible. Next we verify that $V(x,k) := 1 + k|x|^2$ satisfies (4.17) and hence a unique invariant measure exists.*

Observe that $\nabla V(x,k) = 2kx$ and $\nabla^2 V(x,k) = 2kI$. Then we compute

$$
\begin{aligned}
\mathscr{A}V(x,k) := \frac{1}{2}\mathrm{tr}\left(a(x,k)\nabla^2 V(x,k)\right) + \langle b(x,k), \nabla V(x,k)\rangle + \sum_{l\in\mathbb{S}} q_{kl}(x)\left[V(x,l) - V(x,k)\right] \\
+ \int_U \left(V(x + c(x,k,u),k) - V(x,k) - \langle \nabla V(x,k), c(x,k,u)\rangle\right)\nu(du)
\end{aligned}
$$

$$
\begin{aligned}
&\le \frac{1}{2}\mathrm{tr}\left((|x|+1)^2 I\right) + \langle -2kx, 2kx\rangle + \sum_{l\in\mathbb{S}} q_{kl}(x)V(x,l) \\
&\quad + \int_U \left([1+k|x+\gamma\sqrt{k}|u|x|^2] - [1+k|x|^2] - \langle 2kx, \gamma\sqrt{k}|u|x\rangle\right)\nu(du) \\
&= (|x|^2+2|x|+1) - 4k^2|x|^2 + \sum_{l\in\mathbb{S}} \frac{k}{3^{k+l}}\frac{1}{1+l|x|^2}[1+l|x|^2] \\
&\quad + \int_U \left(k(1+\gamma\sqrt{k}|u|)^2|x|^2 - k|x|^2 - 2\gamma k\sqrt{k}|u||x|^2\right)\nu(du) \\
&\le (|x|^2+2|x|+1) - 4k^2|x|^2 + 1 \\
&\quad + \int_U \left(k(1+2\gamma\sqrt{k}|u| + \gamma^2 k|u|^2)|x|^2 - k|x|^2 - 2\gamma k\sqrt{k}|u||x|^2\right)\nu(du) \\
&\le k(|x|^2+2|x|+1) - 4k|x|^2 + 1 + k|x|^2\gamma^2\int_U |u|^2\nu(du) \\
&= k[1+2|x|-2|x|^2] + 1 \\
&= \frac{k[1+2|x|-2|x|^2]+1}{1+k|x|^2}V(x,k).
\end{aligned}
$$

Note that there exists some positive real number r_0 such that for all $|x| > r_0$, we have $1+2|x| \le |x|^2$. Thus it follows that

$$
\mathscr{A}V(x,k) \le \left(-\frac{k|x|^2}{1+k|x|^2} + \frac{1}{1+k|x|^2}\right)V(x,k), \quad \forall (x,k)\in\{x\in\mathbb{R}^d : |x|\ge r_0\}\times\mathbb{S}.
$$

Moreover, for all $k\in\mathbb{S}$ and $|x|\ge r_0$, we have $\frac{k|x|^2}{1+k|x|^2} \ge \frac{|x|^2}{1+|x|^2} \ge \frac{r_0^2}{1+r_0^2} =: 2\alpha > 0$. Also notice that there exists some $r_1 > 0$ such that for all $|x|\ge r_1$ and $k\in\mathbb{S}$, we have $\frac{1}{1+k|x|^2} \le \frac{1}{1+|x|^2} \le \alpha$. Consequently it follows that for some sufficiently large $\beta > 1$, we have

$$
\mathscr{A}V(x,k) \le -\alpha V(x,k) + \beta 1_{C\times N}(x,k), \quad \forall (x,k)\in\mathbb{R}^d\times\mathbb{S},
$$

where $C := \{x\in\mathbb{R}^d : |x| \le r_0 \vee r_1\}$ and $N := \{1\}$. This implies condition (4.17). Thanks to Proposition 4.1.12 there exists a unique invariant probability measure π.

Chapter 5

Exponential Ergodicity

The existence and uniqueness of an invariant measure π for regime-switching diffusion were established in Propositions 4.1.11 and 4.1.12, respectively. In this chapter we study the convergence rate of the transition probability $P(t,(x,k),\cdot)$ to $\pi(\cdot)$. Recall that a σ-finite measure $\pi(\cdot)$ on the Borel σ-algebra $\mathfrak{B}(\mathbb{R}^d \times \mathbb{S})$ is called *invariant* for the semigroup (P_t) if

$$\pi(A) = \pi P_t(A) := \int_{\mathbb{R}^d \times \mathbb{S}} P(t,(x,k),A)\pi(dx,dk) \qquad \forall A \in \mathfrak{B}(\mathbb{R}^d \times \mathbb{S}) \text{ and } t \geq 0.$$

For any function $f : \mathbb{R}^d \times \mathbb{S} \longrightarrow [1,\infty)$ and any signed measure μ on $\mathfrak{B}(\mathbb{R}^d \times \mathbb{S})$, we set

$$\|\mu\|_f := \sup\{|\mu(g)| : \text{ all measurable } g(x,k) \text{ with } |g| \leq f\},$$

where $\mu(g) := \int_{\mathbb{R}^d \times \mathbb{S}} g(x,k)\mu(dx,dk)$. Using the terminology in Meyn and Tweedie (1993c), we say that the process (X,Λ) is *f-exponentially ergodic* if there exist an invariant measure $\pi(\cdot)$, a constant θ in $(0,1)$, and a finite-valued function $\Theta(x,k)$ such that

$$\|P(t,(x,k),\cdot) - \pi(\cdot)\|_f \leq \Theta(x,k)\theta^t, \tag{5.1}$$

for all $t \geq 0$ and all $(x,k) \in \mathbb{R}^d \times \mathbb{S}$.

5.1 Exponential Ergodicity of Regime-Switching Jump Diffusions

We obtain the following result as a direct consequence of Theorem 2.1.9. Similar results can be found in Theorem 6.3 of Xi (2009) and Theorem 6.3 of Xi and Zhu (2017) when the switching state $\mathbb{S}$ is finite and infinite, respectively.

Theorem 5.1.1. *Suppose that all compact subsets of* $\mathbb{R}^d \times \mathbb{S}$ *are petite for some skeleton chain of* $(X(t), \Lambda(t))$. *If there exists a Foster-Lyapunov function* $U : \mathbb{R}^d \times \mathbb{S} \to [0, \infty)$*; that is,* U *satisfies*

(i) $U(x,k) \to \infty \quad \text{as} \quad |x| \vee k \to \infty,$

(ii) $\mathscr{A}U(x,k) \le -\alpha U(x,k) + \beta \quad \text{for all} \quad x \in \mathbb{R}^d, k \in \mathbb{S},$

where $\alpha, \beta > 0$ *are constants, then the process* (X, Λ) *is* f*-exponentially ergodic with* $f(x,k) := U(x,k) + 1$ *and* $\Theta(x,k) = B(U(x,k)+1)$ *where* B *is a finite constant.*

In order to obtain the exponential ergodicity, we still need to determine the existence of a Foster-Lyapunov function. And this will be investigated in the next section under Assumptions 5.2.1 and 5.2.2. However, to keep the flow of the paper let us now state and prove our main result in this section as follows.

Theorem 5.1.2. *Suppose that Assumptions 3.1.2, 4.1.1, 4.1.2, 4.1.3, 4.1.4, 5.2.1, and 5.2.2 hold, then the process* (X, Λ) *is* f*-exponentially ergodic.*

Proof. Thanks to Proposition 4.2.2, under Assumptions 3.1.2, 4.1.1, 4.1.2, 4.1.3, and 4.1.4, all compact sets of $\mathbb{R}^d \times \mathbb{S}$ are petite for some h-skeleton chain. Under Assumptions 5.2.1 and 5.2.2 a Foster-Lyapunov function U exists by Theorem 5.2.4. Then the desired f-exponential ergodicity follows from Theorem 5.1.1 where $f(x,k) := U(x,k) + 1$. □

5.2 Existence of Foster-Lyapunov Functions

Having established sufficient conditions for petite compact subsets of $\mathbb{R}^d \times \mathbb{S}$, it remains to find an appropriate Foster-Lyapunov function. In practice, it is not easy to find the right Foster-Lyapunov function for an underlying regime-switching jump diffusion especially when dealing with countably many regimes. Motivated by the recent paper Nguyen and Yin (2018c) in which the stability of regime-switching diffusion was investigated, we develop a novel approach to construct a Foster-Lyapunov function for regime-switching jump diffusions. Let us briefly sketch the idea here. Suppose that there exists a common "nice" function $V : \mathbb{R}^d \mapsto \mathbb{R}_+$ so that

$$\mathcal{L}_k V(x) \le \alpha_k V(x) + \beta_k, \text{ for all } (x,k) \in \mathbb{R}^d \times \mathbb{S},$$

where α_k and β_k are real numbers. Suppose also that the generator $Q(x)$ of the discrete component is "close" to a *strongly exponentially ergodic* (see Definition 7.2.3) constant q-matrix in the neighborhood of ∞. Then, under some additional assumptions, we construct a Foster-Lyapunov function for the process (X, Λ).

To proceed, we make the following assumptions.

Assumption 5.2.1. *(a) There exists an increasing function $\phi : \mathbb{S} \to [0, \infty)$ such that $\lim_{k\to\infty} \phi(k) = \infty$ and*

$$\sum_{j\in\mathbb{S}} q_{kj}(x)\,[\phi(j) - \phi(k)] \le C_1 - C_2\phi(k) \quad \text{for all } \; k \in \mathbb{S}, x \in \mathbb{R}^d, \tag{5.2}$$

where $C_1 \ge 0$ and $C_2 > 0$ are constants.

(b) There exists a bounded and x-independent q-matrix $\hat{Q} = (\hat{q}_{ij})_{i,j\in\mathbb{S}}$ which is strongly

exponentially ergodic with invariant measure $\nu = (\nu_1, \nu_2, ...)$ *such that*

$$\sup_{k\in\mathbb{S}} \sum_{j\in\mathbb{S}} |q_{kj}(x) - \hat{q}_{kj}| \to 0 \quad as \quad x \to \infty. \tag{5.3}$$

Assumption 5.2.2. *There exists a twice continuously differentiable and norm-like function* $V : \mathbb{R}^d \to [1, \infty)$ *such that for each* $k \in \mathbb{S}$

$$\mathcal{L}_k V(x) \le \alpha_k V(x) + \beta_k \qquad \text{for all } x \in \mathbb{R}^d, \tag{5.4}$$

where $\{\alpha_k\}_{k\in\mathbb{S}}$ *and* $\{\beta_k\}_{k\in\mathbb{S}}$ *are bounded sequences of real numbers such that* $\beta_k \ge 0$ *and*

$$\sum_{k\in\mathbb{S}} \alpha_k \nu_k < 0. \tag{5.5}$$

Remark 5.2.3. *It is worth noticing that if* $\alpha_k < 0$ *in* (5.4) *then* $V(x)$ *is a Foster-Lyapunov function for the corresponding subsystem* $X^{(k)}$ *defined in* (4.1). *However, to obtain the existence of a Foster-Lyapunov function for the process* (X, Λ) *we only require* (5.5) *to be satisfied. In other words, we can still obtain exponential ergodicity of the process* (X, Λ) *as long as "most"of the subsystems* $X^{(k)}$ *are nice in some sense; for example, in this case, the "average"in* (5.5) *is satisfied.*

Theorem 5.2.4. *Suppose that Assumptions 5.2.1 and 5.2.2 hold. Then there exists a Foster-Lyapunov function* $U : \mathbb{R}^d \times \mathbb{S} \mapsto \mathbb{R}_+$ *satisfying the following properties*

(i) $U(x, k) \to \infty$ *as* $|x| \vee k \to \infty$,

(ii) $\mathscr{A} U(x, k) \le -\alpha U(x, k) + \beta$ *for all* $x \in \mathbb{R}^d, k \in \mathbb{S}$,

where $\alpha, \beta > 0$ *are constants.*

Proof. Let $\gamma := -\sum_{k\in\mathbb{S}} \alpha_k \nu_k > 0$. Since $\{\alpha_k\}_{k\in\mathbb{S}}$ is bounded and $\sum_{k\in\mathbb{S}} \nu_k = 1$ then the series

$\sum_{k=1}^{\infty}(\alpha_k + \gamma)\nu_k$ is absolutely convergent and hence

$$\sum_{k=1}^{\infty}(\alpha_k + \gamma)\nu_k = \sum_{k=1}^{\infty}\alpha_k\nu_k + \sum_{k=1}^{\infty}\gamma\nu_k = 0.$$

Since $\hat{Q} = (\hat{q}_{ij})_{i,j\in\mathbb{S}}$ is strongly exponentially ergodic, it follows from Lemma 7.2.6 that there exits a bounded sequence of real numbers $\{\gamma_k : k \in \mathbb{S}\}$ such that

$$\sum_{j\in\mathbb{S}} \hat{q}_{kj}\gamma_j = \alpha_k + \gamma \quad \text{ for all } \ k \in \mathbb{S}. \tag{5.6}$$

Next we choose $p \in (0, 1)$ so that

$$p|\gamma_k| \leq 0.5 \tag{5.7}$$

and

$$p|\gamma_k\alpha_k| \leq 0.5\gamma. \tag{5.8}$$

Define a function $U : \mathbb{R}^d \times \mathbb{S} \to [0, \infty)$ by

$$U(x, k) \;\; := \;\; (1 - p\gamma_k)V^p(x) + \phi(k). \tag{5.9}$$

From (5.7), we see that $U(x, k)$ is nonnegative and satisfies $\lim_{|x|\vee k\to\infty} U(x, k) = \infty$.

The rest of the proof is to verify that condition (ii) holds. To proceed, we compute and estimate each term of the generator

$$\mathscr{A}U(x, k) = \mathscr{A}_d U(x, k) + \mathscr{A}_j U(x, k) + Q(x)U(x, k).$$

First, observe that

$$\nabla U(x,k) = p(1-p\gamma_k)V^{p-1}(x)\nabla V(x)$$

and

$$\nabla^2 U(x,k) = p(1-p\gamma_k)V^{p-1}(x)\nabla^2 V(x) - p(1-p)(1-p\gamma_k)V^{p-2}(x)\nabla V(x)\nabla V(x)^T.$$

Then

$$\begin{aligned}
\mathscr{A}_d U(x,k) &= p(1-p\gamma_k)V^{p-1}(x)\frac{1}{2}\mathrm{tr}\left(a(x,k)\nabla^2 V(x)\right) \\
&\quad -p(1-p)(1-p\gamma_k)V^{p-2}(x)\frac{1}{2}\mathrm{tr}\left(a(x,k)\nabla V(x)\nabla V(x)^T\right) \\
&\quad +p(1-p\gamma_k)V^{p-1}(x)\langle b(x,k),\nabla V(x)\rangle \\
&= p(1-p\gamma_k)V^{p-1}(x)\left[\frac{1}{2}\mathrm{tr}\left(a(x,k)\nabla^2 V(x)\right) + \langle b(x,k),\nabla V(x)\rangle\right] \\
&\quad -p(1-p)(1-p\gamma_k)V^{p-2}(x)\frac{1}{2}|\nabla V(x)^T\sigma(x,k)|^2 \\
&\le p(1-p\gamma_k)V^{p-1}(x)\left[\frac{1}{2}\mathrm{tr}\left(a(x,k)\nabla^2 V(x)\right) + \langle b(x,k),\nabla V(x)\rangle\right]. \quad (5.10)
\end{aligned}$$

To estimate the second term we note that the function $f(r) = r^p$ for $r > 0$ is concave since $0 < p < 1$. Hence $b^p - a^p \le pa^{p-1}[b-a]$ for all $a, b > 0$. By taking $b = V(x + c(x,k,u))$ and $a = V(x)$, we have

$$\begin{aligned}
&U(x + c(x,k,u),k)) - U(x,k) - \langle \nabla U(x,k), c(x,k,u)\rangle \\
&\quad = (1-p\gamma_k)V^p(x + c(x,k,u)) - (1-p\gamma_k)V^p(x) - p(1-p\gamma_k)V^{p-1}(x)\langle \nabla V(x), c(x,k,u)\rangle \\
&\quad \le p(1-p\gamma_k)V^{p-1}(x)\left[V(x + c(x,k,u)) - V(x) - \langle \nabla V(x), c(x,k,u)\rangle\right].
\end{aligned}$$

Hence

$$\mathscr{A}_j U(x,k) \le p(1-p\gamma_k)V^{p-1}(x)\int_U \left[V(x+c(x,k,u)) - V(x) - \langle \nabla V(x), c(x,k,u)\rangle\right]\nu(du). \tag{5.11}$$

Finally, we estimate the last term $Q(x)U(x,k)$. Note that $q_{kj}(x) \ge 0$ for all $k \ne j$. Since ϕ is increasing and satisfies $\phi(k) \to \infty$ as $k \to \infty$, then (5.2) asserts that

$$\begin{aligned}
\sum_{j\in\mathbb{S}} q_{kj}(x)|\phi(j)-\phi(k)| &= \sum_{j<k} q_{kj}(x)|\phi(j)-\phi(k)| + \sum_{j>k} q_{kj}(x)\left[\phi(j)-\phi(k)\right] \\
&\le -\sum_{j<k} q_{kj}(x)\left[\phi(j)-\phi(k)\right] + C_1 - C_2\phi(k) - \sum_{j<k} q_{kj}(x)\left[\phi(j)-\phi(k)\right] \\
&= C_1 - C_2\phi(k) - 2\sum_{j<k} q_{kj}(x)\left[\phi(j)-\phi(k)\right] \\
&< \infty.
\end{aligned}$$

Then $\sum_{j\in\mathbb{S}} q_{kj}(x)\left[\phi(j)-\phi(k)\right]$ is absolutely convergent for each $k \in \mathbb{S}$. Since $\{\gamma_k\}_{k\in\mathbb{S}}$ is a bounded sequence, we have

$$\sum_{j\in\mathbb{S}} |q_{kj}(x)[pV^p(x)(\gamma_k-\gamma_j)]| = pV^p(x)\sum_{j\ne k} q_{kj}(x)|\gamma_k-\gamma_j| \le 2pV^p(x)\sup_{j\in\mathbb{S}}\{|\gamma_j|\}q_k(x) < \infty.$$

Hence $\sum_{j\in\mathbb{S}} q_{kj}(x)\left[pV^p(x)(\gamma_k-\gamma_j)\right]$ is also absolutely convergent. Then we can compute

$$\begin{aligned}
\sum_{j\in\mathbb{S}} q_{kj}(x)\left[pV^p(x)(\gamma_k-\gamma_j)\right] &= pV^p(x)\gamma_k\sum_{j\in\mathbb{S}} q_{kj}(x) - pV^p(x)\sum_{j\in\mathbb{S}} q_{kj}(x)\gamma_j \\
&= 0 - pV^p(x)\sum_{j\in\mathbb{S}} q_{kj}(x)\gamma_j \\
&= -pV^p(x)\sum_{j\in\mathbb{S}} q_{kj}(x)\gamma_j.
\end{aligned}$$

The absolute convergences allow us to rearrange the summands of $Q(x)U(x,k)$ as follows

$$\begin{aligned}
Q(x)U(x,k) &= \sum_{j\in\mathbb{S}} q_{kj}(x)\left[U(x,j)-U(x,k)\right] \\
&= \sum_{j\in\mathbb{S}} q_{kj}(x)\left[(1-p\gamma_j)V^p(x)+\phi(j)-(1-p\gamma_k)V^p(x)-\phi(k)\right] \\
&= \sum_{j\in\mathbb{S}} q_{kj}(x)\left[pV^p(x)(\gamma_k-\gamma_j)+\phi(j)-\phi(k)\right] \\
&= \sum_{j\in\mathbb{S}} q_{kj}(x)\left[pV^p(x)(\gamma_k-\gamma_j)\right]+\sum_{j\in\mathbb{S}} q_{kj}(x)\left[\phi(j)-\phi(k)\right] \\
&= -pV^p(x)\sum_{j\in\mathbb{S}} q_{kj}(x)\gamma_j+\sum_{j\in\mathbb{S}} q_{kj}(x)\left[\phi(j)-\phi(k)\right] \\
&\le -pV^p(x)\sum_{j\in\mathbb{S}} q_{kj}(x)\gamma_j+C_1-C_2\phi(k),
\end{aligned}$$

where we use (5.2) to obtain the inequality. Furthermore, since $\sum_{j\in\mathbb{S}} q_{kj}(x)\gamma_j$ and $\sum_{j\in\mathbb{S}} \hat{q}_{kj}\gamma_j$ are also absolutely convergent, we have

$$\begin{aligned}
Q(x)U(x,k) &\le -pV^p(x)\sum_{j\in\mathbb{S}} q_{kj}(x)\gamma_j+C_1-C_2\phi(k) \\
&= -pV^p(x)\sum_{j\in\mathbb{S}}\left(q_{kj}(x)-\hat{q}_{kj}\right)\gamma_j - pV^p(x)\sum_{j\in\mathbb{S}}\hat{q}_{kj}\gamma_j+C_1-C_2\phi(k) \\
&\le pV^p(x)\sup_{j\in\mathbb{S}}|\gamma_j|\sum_{j\in\mathbb{S}}|q_{kj}(x)-\hat{q}_{kj}| - pV^p(x)\sum_{j\in\mathbb{S}}\hat{q}_{kj}\gamma_j+C_1-C_2\phi(k) \\
&= pV^p(x)\sup_{j\in\mathbb{S}}|\gamma_j|\sum_{j\in\mathbb{S}}|q_{kj}(x)-\hat{q}_{kj}| - pV^p(x)(\alpha_k+\gamma)+C_1-C_2\phi(k), \quad (5.12)
\end{aligned}$$

where we use (5.6) to obtain the last equality. It follows from (5.10)–(5.12) that

$$\begin{aligned}
&\mathscr{A}U(x,k) \\
&\le p(1-p\gamma_k)V^{p-1}(x)\left[\frac{1}{2}\mathrm{tr}\left(a(x,k)\nabla^2V(x)\right)+\langle b(x,k),\nabla V(x)\rangle\right] \\
&\quad + p(1-p\gamma_k)V^{p-1}(x)\int_U\left[V(x+c(x,k,u))-V(x)-\langle\nabla V(x),c(x,k,u)\rangle\right]\nu(du)
\end{aligned}$$

$$
\begin{aligned}
&+ pV^p(x) \sup_{j\in\mathbb{S}} |\gamma_j| \sum_{j\in\mathbb{S}} |q_{kj}(x) - \hat{q}_{kj}| - pV^p(x)(\alpha_k + \gamma) + C_1 - C_2\phi(k) \\
&= p(1 - p\gamma_k)V^{p-1}(x)\,[\mathcal{L}_k V(x)] + pV^p(x) \sup_{j\in\mathbb{S}} |\gamma_j| \sum_{j\in\mathbb{S}} |q_{kj}(x) - \hat{q}_{kj}| \\
&- pV^p(x)(\alpha_k + \gamma) + C_1 - C_2\phi(k) \\
&\le p(1 - p\gamma_k)V^{p-1}(x)\,[\alpha_k V(x) + \beta_k] + pV^p(x) \sup_{j\in\mathbb{S}} |\gamma_j| \sum_{j\in\mathbb{S}} |q_{kj}(x) - \hat{q}_{kj}| \\
&- pV^p(x)(\alpha_k + \gamma) + C_1 - C_2\phi(k) \\
&= p(1 - p\gamma_k)V^p(x)\left[\frac{\beta_k}{V(x)} - \frac{p\alpha_k\gamma_k + \gamma}{1 - p\gamma_k} + \frac{\sup_{j\in\mathbb{S}} |\gamma_j| \sum_{j\in\mathbb{S}} |q_{kj}(x) - \hat{q}_{kj}|}{1 - p\gamma_k}\right] + C_1 - C_2\phi(k),
\end{aligned}
\tag{5.13}
$$

where the last inequality follows from (5.4). Thanks to (5.7), we have $1 - p\gamma_k > 0$ and it is bounded. Let $\delta := 1 \wedge \frac{0.5\gamma}{1-p\gamma_k} \wedge \frac{2C_2}{p}$. Note that $0 < \delta \le \frac{0.5\gamma}{1-p\gamma_k}$. From (5.8), we see that $-(p\gamma_k\alpha_k + \gamma) \le -0.5\gamma$. Hence

$$
-\frac{p\gamma_k\alpha_k + \gamma}{1 - p\gamma_k} \le \frac{-0.5\gamma}{1 - p\gamma_k} \le \frac{-\delta(1 - p\gamma_k)}{1 - p\gamma_k} = -\delta. \tag{5.14}
$$

On the other hand, since V is norm-like and $\{\beta_k\}_{k\in\mathbb{S}}$ is bounded, there exists an $M_1 > 0$ such that

$$
\frac{\beta_k}{V(x)} \le 0.25\delta \quad \text{ for all } |x| \ge M_1 \text{ and } k \in \mathbb{S}. \tag{5.15}
$$

Similarly, we can use (5.3) and (5.7) to find an $M_2 > 0$ such that

$$
\frac{\sup_{j\in\mathbb{S}} |\gamma_j| \sum_{j\in\mathbb{S}} |q_{kj}(x) - \hat{q}_{kj}|}{1 - p\gamma_k} \le 2 \sup_{j\in\mathbb{S}} |\gamma_j| \sum_{j\in\mathbb{S}} |q_{kj}(x) - \hat{q}_{kj}| \le 0.25\delta \tag{5.16}
$$

for all $|x| \ge M_2$ and $k \in \mathbb{S}$. Now plugging (5.14)–(5.16) into (5.13) yields

$$
\mathscr{A}U(x,k) \le p(1 - p\gamma_k)V^p(x)[0.25\delta - \delta + 0.25\delta] + C_1 - C_2\phi(k)
$$

$$
\begin{aligned}
&= -0.5\delta p(1 - p\gamma_k)V^p(x) + C_1 - C_2\phi(k) \\
&= -0.5\delta p\left[(1 - p\gamma_k)V^p(x) + \phi(k)\right] + 0.5\delta p\phi(k) + C_1 - C_2\phi(k) \\
&\leq -0.5\delta p U(x, k) + C_1,
\end{aligned}
$$

for all $|x| \geq M_1 \vee M_2$ and $k \in \mathbb{S}$. Note that we used the fact that $0.5\delta p \leq C_2$ to derive the last inequality. To complete the proof, we choose $\alpha, \beta > 0$ so that

$$
\mathscr{A} U(x, k) \leq -\alpha U(x, k) + \beta
$$

holds for all $x \in \mathbb{R}^d$ and $k \in \mathbb{S}$. This completes the proof. □

Corollary 5.2.4.1. *Suppose that the q-matrix $Q(x) = (q_{kl})$ is constant, irreducible, and strongly exponentially ergodic with invariant measure $\pi = (\pi_1, \pi_2, \dots)$. Then under Assumptions 5.2.1 (a) and 5.2.2, a Foster-Lyapunov function exists.*

5.3 Examples

The strongly exponentially ergodic q-matrix $\hat{Q}$ in Assumption 5.2.1 (b) plays a very crucial role in the proof of Theorem 5.2.4 as it allows us, by Lemma 7.2.6, to find a bounded sequence $\{\gamma_k : k \in \mathbb{S}\}$ satisfying equality (5.6). To demonstrate our results, let us first give some examples of such matrices.

Given a positive constant $\theta > 0$. Consider the q-matrix $\hat{Q} = (\hat{q}_{ij})$ given by

$$
\hat{q}_{ij} := \begin{cases} -\frac{1}{2}\theta & \text{if} \quad j = 1, i = 1 \\ \frac{1}{2}\theta & \text{if} \quad j = 1, i \neq j \\ \frac{1}{3^{j-1}}\theta & \text{if} \quad j > 1, i \neq j \\ -\frac{3^{j-1}-1}{3^{j-1}}\theta & \text{if} \quad j > 1, i = j, \end{cases}
$$

that is

$$\hat{Q} = \theta \begin{bmatrix} -\frac{1}{2} & \frac{1}{3} & \frac{1}{3^2} & \cdots \\ \frac{1}{2} & -\frac{2}{3} & \frac{1}{3^2} & \cdots \\ \frac{1}{2} & \frac{1}{3} & -\frac{8}{3^2} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}.$$

It is clear that $\nu = (\frac{1}{2}, \frac{1}{3}, \frac{1}{3^2}, ...)$ solves the equation $\nu\hat{Q} = 0$ and $\nu 1 = 1$. Then ν is an invariant probability measure. Solving the Kolmogorov backward equation $\hat{P}'(t) = \hat{Q}\hat{P}(t)$ gives

$$\hat{P}_{ij}(t) = \begin{cases} \frac{1}{2} + \frac{1}{2}e^{-\theta t} & \text{if} \quad j = 1, i = 1 \\ \frac{1}{2} - \frac{1}{2}e^{-\theta t} & \text{if} \quad j = 1, i \neq j \\ \frac{1}{3^{j-1}} - \frac{1}{3^{j-1}}e^{-\theta t} & \text{if} \quad j > 1, i \neq j \\ \frac{1}{3^{j-1}} + \frac{3^{j-1}-1}{3^{j-1}}e^{-\theta t} & \text{if} \quad j > 1, i = j, \end{cases}$$

that is

$$\hat{P}(t) = \begin{bmatrix} \frac{1}{2} + \frac{1}{2}e^{-\theta t} & \frac{1}{3} - \frac{1}{3}e^{-\theta t} & \frac{1}{3^2} - \frac{1}{3^2}e^{-\theta t} & \cdots \\ \frac{1}{2} - \frac{1}{2}e^{-\theta t} & \frac{1}{3} + \frac{2}{3}e^{-\theta t} & \frac{1}{3^2} - \frac{1}{3^2}e^{-\theta t} & \cdots \\ \frac{1}{2} - \frac{1}{2}e^{-\theta t} & \frac{1}{3} - \frac{1}{3}e^{-\theta t} & \frac{1}{3^2} + \frac{8}{3^2}e^{-\theta t} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix}.$$

For each $i \in \mathbb{S}$ and $t \geq 0$, we see that

$$\sum_{j=1}^{\infty} |\hat{P}_{ij}(t) - \nu_j| = \frac{1}{2}e^{-\theta t} + \sum_{j>1, j\neq i} \frac{1}{3^{j-1}}e^{-\theta t} + |\frac{3^{i-1}-1}{3^{i-1}}e^{-\theta t}| \leq \frac{1}{2}e^{-\theta t} + \sum_{j=1}^{\infty} \frac{1}{3^j}e^{-\theta t} + e^{-\theta t} = 2e^{-\theta t}.$$

Then, for arbitrary but fixed $\theta > 0$, any Markov chain generated by $\hat{Q}$ is strongly exponentially ergodic.

Example 5.3.1. *Consider the following SDE*

$$dX(t) = b(X(t), \varLambda(t))dt + \sigma(X(t), \varLambda(t))dW(t) + \int_U c(X(t^-), \varLambda(t^-), u)\tilde{N}(dt, du), \quad (5.17)$$

where W is a standard 2-dimensional Brownian motion, $\tilde{N}$ is the compensated Poisson random measure on $[0, \infty) \times U$ with intensity $dt\nu(du)$ in which $U = \{u \in \mathbb{R}^2 : 0 < |u| < 1\}$ and $\nu(du) := \frac{du}{|u|^{2+\delta}}$ for some $\delta \in (0, 2)$. The coefficients of (5.17) are given by

$$\sigma(x, k) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad b(x, k) = \begin{cases} -x & \text{if } k = 1 \\ \frac{1}{4k}x & \text{if } k \geq 2 \end{cases}, \quad \text{and} \quad c(x, k, u) = \gamma \frac{1}{\sqrt{2k}}|u|x,$$

where γ is a positive constant so that $\gamma^2 \int_U |u|^2 \nu(du) = 1$. The $\varLambda$ component takes value in $\mathbb{S} := \{1, 2, ...\}$ and is generated by $Q(x) = (q_{kj}(x))$ given by

$$q_{kj}(x) := \begin{cases} \frac{1}{2}\frac{k}{k+e^{-|x|^2}} & \text{if } \ j = 1, k \neq j, \\ \frac{1}{3^{j-1}}\frac{k}{k+e^{-|x|^2}} & \text{if } \ j > 1, k \neq j \\ -\sum_{j \neq k} q_{kj}(x) & \text{if } \ k = j. \end{cases}$$

Apparently, (5.17) possesses a unique strong solution $(X, \varLambda) = \{(X(t), \varLambda(t)), 0 \leq t < \infty\}$ (see Theorem 2.5 of Xi et al. (2019)). Assumptions 4.1.1, 4.1.2, and 4.1.4 are trivially satisfied. Next we verify Assumptions 5.2.1 and 5.2.2. To this end, we show that the function $\phi(k) = k$ satisfies (5.2). Indeed, we have

$$\begin{aligned} \sum_{j=1}^{\infty} q_{kj}(x)[\phi(j) - \phi(k)] &= \frac{1}{2}\frac{k}{k + e^{-|x|^2}}[1 - k] + \sum_{j>1} \frac{1}{3^{j-1}}\frac{k}{k + e^{-|x|^2}}[j - k] \\ &\leq \frac{1}{2} - \frac{1}{2} \cdot \frac{1}{2}k + \frac{1}{3} \cdot \frac{k}{k + e^{-|x|^2}} \left[\sum_{j>1} \frac{j}{3^j} - k \sum_{j>1} \frac{1}{3^j}\right] \\ &= \frac{1}{2} - \frac{1}{4}k + \frac{1}{3} \cdot \frac{k}{k + e^{-|x|^2}} \left[(\frac{3}{4} - \frac{1}{3}) - k(\frac{1}{2} - \frac{1}{3})\right] \end{aligned}$$

$$
\begin{aligned}
&= \frac{1}{2} - \frac{1}{4}k + \frac{1}{3} \cdot \frac{k}{k + e^{-|x|^2}} \left[\frac{5}{12} - \frac{1}{6}k\right] \\
&\leq \frac{1}{2} - \frac{1}{4}k + \frac{1}{3} \cdot \frac{5}{12} - \frac{1}{3} \cdot \frac{1}{2} \cdot \frac{1}{6}k \\
&\leq 2 - \frac{5}{18}\phi(k).
\end{aligned}
$$

To show (5.3) we consider the following. Let $\hat{\Lambda}$ be a continuous-time Markov chain with state space $\mathbb{S}$ and generated by $\hat{Q} = \{\hat{q}_{kj}\}$, where

$$
\hat{q}_{ij} := \begin{cases} -\frac{1}{2} & \textit{if} \quad j = 1, i = 1 \\ \frac{1}{2} & \textit{if} \quad j = 1, i \neq j \\ \frac{1}{3^{j-1}} & \textit{if} \quad j > 1, i \neq j \\ -\frac{3^{j-1}-1}{3^{j-1}} & \textit{if} \quad j > 1, i = j. \end{cases}
$$

As shown above, with $\theta = 1$, $\hat{\Lambda}$ is strongly exponentially ergodic with invariant measure $\nu = (\frac{1}{2}, \frac{1}{3}, \frac{1}{3^2}, \dots)$. We see that

$$
\begin{aligned}
\sum_{j\in\mathbb{S}} \left|q_{kj}(x) - \hat{q}_{kj}\right| &= \sum_{j\neq k} \left|q_{kj}(x) - \hat{q}_{kj}\right| + \left|q_{kk}(x) - \hat{q}_{kk}\right| \\
&= \sum_{j\neq k} \left|q_{kj}(x) - \hat{q}_{kj}\right| + \left| - \sum_{j\neq k} q_{kj}(x) + \sum_{j\neq k} \hat{q}_{kj}(x)\right| \\
&\leq \sum_{j\neq k} \left|q_{kj}(x) - \hat{q}_{kj}\right| + \sum_{j\neq k} \left|q_{kj}(x) - \hat{q}_{kj}\right| \\
&= 2\sum_{j\neq k} \left|q_{kj}(x) - \hat{q}_{kj}\right| \\
&= 2\left|\frac{1}{2}\frac{k}{k + e^{-|x|^2}} - \frac{1}{2}\right| + 2\sum_{j>1, j\neq k} \left|\frac{1}{3^{j-1}}\frac{k}{k + e^{-|x|^2}} - \frac{1}{3^{j-1}}\right| \\
&\leq \left|\frac{k}{k + e^{-|x|^2}} - 1\right| + 2\sum_{j\geq 1} \frac{1}{3^j}\left|\frac{k}{k + e^{-|x|^2}} - 1\right| \\
&= 2\left[1 - \frac{k}{k + e^{-|x|^2}}\right] \\
&\leq 2e^{-|x|^2}.
\end{aligned}
$$

This implies that

$$\sup_{k\in\mathbb{S}}\sum_{j\in\mathbb{S}}|q_{kj}(x)-\hat{q}_{kj}|\to 0 \quad as \quad x\to\infty$$

and thus establishing (5.3). *As a result, Assumption 5.2.1 is verified.*

To verify Assumption 5.2.2 we consider function $V(x) := |x|^2$ *and observe that* $\nabla V(x) = 2x$ *and* $\nabla^2 V(x) = 2I$. *We compute*

$$\begin{aligned}
\mathcal{L}_k V(x) &= \frac{1}{2}\operatorname{tr}\big(a(x,k)\nabla^2 V(x)\big) + \langle b(x,k), \nabla V(x,k)\rangle \\
&\quad + \int_U \left(V(x+c(x,k,u)) - V(x) - \langle \nabla V(x), c(x,k,u)\rangle\right)\nu(du) \\
&= 2 + 2\langle b(x,k), x\rangle + \int_U \big(|x+c(x,k,u)|^2 - |x|^2 - 2\langle x, c(x,k,u)\rangle\big)\,\nu(du) \\
&= 2 + 2\langle b(x,k), x\rangle + \frac{|x|^2}{2k} \\
&= \begin{cases} 2 - \frac{3}{2}|x|^2 & if \quad k = 1, \\ 2 + \frac{|x|^2}{k} & if \quad k \geq 2. \end{cases}
\end{aligned}$$

Then $\beta_k = 2$ *for all* $k \in \mathbb{S}$ *and*

$$\alpha_k = \begin{cases} -\frac{3}{2} & if \quad k = 1 \\ \frac{1}{k} & if \quad k \geq 2. \end{cases}$$

We also see that

$$\sum_{k=1}^{\infty}\alpha_k\nu_k = -\frac{3}{2}\cdot\frac{1}{2} + \sum_{k=2}^{\infty}\frac{1}{k}\frac{1}{3^{k-1}} = -\frac{3}{4} + 3[\log(\frac{3}{2}) - \frac{1}{3}] < 0.$$

Then Theorem 5.2.4 ensures the existence of a Foster-Lyapunov function $U(x,k)$. *Moreover, the process* (X,Λ) *is exponential ergodic by Theorem 5.1.2.*

Chapter 6

Application to Feedback Controls

6.1 Problem Formulation

In this section we illustrate an application of Theorem 5.1.2. To proceed, we start with the following system of SDEs

$$\begin{cases} dX(t) = b(X(t), \Lambda(t))dt + \sigma(X(t), \Lambda(t))dW(t) + \int_U c(X(t^-), \Lambda(t^-), u)\tilde{N}(dt, du) \\ \Lambda(t) = \Lambda(0) + \int_0^t \int_{\mathbb{R}_+} h(X(s^-), \Lambda(s^-), r)N_1(ds, dr) \end{cases} \tag{6.1}$$

where b, σ and c are appropriate measurable functions.

Recall that feedback control is any control that depends on the current state of the underlying process. Motivated by the study of feedback controls for weak stabilization studied in Zhu and Yin (2009), we raise and try to answer the following question: If a regime-switching jump diffusion is not exponentially ergodic or even not ergodic, can we find a suitable control so that the controlled regime-switching jump diffusion becomes exponentially ergodic? To this end, we consider the following SDE

$$\begin{aligned} dX(t) = b(X(t), \Lambda(t))dt + \xi(X(t), \Lambda(t))dt + \sigma(X(t), \Lambda(t))dW(t) \\ + \int_U c(X(t^-), \Lambda(t^-), u)\tilde{N}(dt, du), \end{aligned} \tag{6.2}$$

where $\xi:\mathbb{R}^d\times\mathbb{S}\to\mathbb{R}^d$ denotes the feedback control which will be determined later on. We denote by $(\tilde X,\tilde\Lambda)$ the solution to the following system of SDEs

$$\begin{cases} d\tilde X(t)=b(\tilde X(t),\tilde\Lambda(t))dt+\xi(\tilde X(t),\tilde\Lambda(t))dt+\sigma(\tilde X(t),\tilde\Lambda(t))dW(t)\\ +\int_U c(\tilde X(t^-),\tilde\Lambda(t^-),u)\tilde N(dt,du),\\ \tilde\Lambda(t)=\tilde\Lambda(0)+\int_0^t\int_{\mathbb{R}_+}\tilde h(\tilde X(s^-),\tilde\Lambda(s^-),r)N_1(ds,dr)\end{cases} \tag{6.3}$$

where $\tilde h$ can be defined in a similar way to (2.13). In other words, if Λ is determined by the probability rate matrix $Q(x)$ then $\tilde\Lambda$ will be determined by the matrix $Q(\tilde x)$.

In practice, we usually decompose the switching state space $\mathbb{S}$ into the union of two disjoint subsets, namely $\mathbb{S}=\mathbb{S}_{int}\cup\mathbb{S}_{ab}$. To be more precise, $\mathbb{S}_{ab}$ consists of states with an absence of intervention while $\mathbb{S}_{int}$ consists of those states when any intervention can take place. It is reasonable and easy to consider the feedback controls of the form

$$\xi(\tilde X(t),\tilde\Lambda(t))=-L(\tilde\Lambda(t))\tilde X(t), \tag{6.4}$$

where $L(k)\in\mathbb{R}^{d\times d}$ is a constant matrix for $k\in\mathbb{S}_{int}$. Of course, we take $L(k)=0$ for each $k\in\mathbb{S}_{ab}$. For simplicity, we set $\tilde b(y,k):=b(y,k)-L(k)y$. Then (6.3) becomes

$$\begin{cases} d\tilde X(t)=\tilde b(\tilde X(t),\tilde\Lambda(t))dt+\sigma(\tilde X(t),\tilde\Lambda(t))dW(t)+\int_U c(\tilde X(t^-),\tilde\Lambda(t^-),u)\tilde N(dt,du),\\ \tilde\Lambda(t)=\tilde\Lambda(0)+\int_0^t\int_{\mathbb{R}_+}\tilde h(\tilde X(s^-),\tilde\Lambda(s^-),r)N_1(ds,dr).\end{cases} \tag{6.5}$$

We denote by $\tilde X^{(k)}$ a non-explosive solution of the subsystem

$$\begin{aligned}\tilde X^{(k)}(t)=x+\int_0^t\tilde b(\tilde X^{(k)}(s),k)ds+\int_0^t\sigma(\tilde X^{(k)}(s),k)dW(s)\\ +\int_0^t\int_U c(\tilde X^{(k)}(s-),k,u)\widetilde N(ds,du).\end{aligned} \tag{6.6}$$

6.2 Exponential Ergodicity of the Controlled Process

In this section we show how to apply Theorem 5.1.2 to the feedback control problems. Before we proceed one may ask whether the control process ξ in (6.4) is admissible or not. That is, whether the system (6.3) or equivalently (6.5) has a unique non-explosive strong solution. To tackle this issue we need Assumption 3.1.1.

Theorem 6.2.1. *Suppose that Assumptions 3.1.1, 3.1.2, 4.1.2, 4.1.3, 4.1.4, and 5.2.1 hold. Given a constant matrix $L(k)$ for each $k \in \mathbb{S}_{int}$. If there exists a bounded sequence of real numbers $\{\alpha_k\}$ such that*

$$\sum_{k \in \mathbb{S}} \alpha_k \nu_k < 0 \tag{6.7}$$

and

$$(\kappa - \alpha_k)|x|^2 \leq \langle x, L(k)x \rangle \quad \text{for all} \quad x \in \mathbb{R}^d, k \in \mathbb{S}_{int} \cup \mathbb{S}_{ab}, \tag{6.8}$$

where κ is the positive constant given in (4.2), then the controlled process $(\tilde{X}, \tilde{\Lambda})$ is exponentially ergodic.

Proof. In view of Theorem 5.1.2, we proceed as follows. Since the control process ξ defined in (6.4) is linear in the x variable, it is clear that if (6.1) satisfies Assumptions 3.1.1–5.2.1 then so does (6.5). Thanks to Lemma 2.4, Xi et al. (2019), under Assumption 3.1.1 and (4.2), there exists a unique non-explosive strong solution $\tilde{X}^{(k)}$ to the SDE (6.6).

We observe that (4.2), (3.9), (4.6), and (5.2) constitute Assumption 2.1 of Xi et al. (2019) with relaxed condition (3.9). Moreover, Theorem 2.5 of Xi et al. (2019) is still valid under this relaxation. This ensures the existence and uniqueness of a non-explosive strong solution $(\tilde{X}, \tilde{\Lambda})$ to (6.5).

Next, we only need to verify (5.4) for $\tilde{X}$. Consider function $V(x) := |x|^2$. Thanks to

(4.2) and (6.8), we verify that

$$\begin{aligned}
\mathcal{L}_k V(x) &= \frac{1}{2}\mathrm{tr}\left(a(x,k)\nabla^2 V(x)\right) + \langle \tilde{b}(x,k), \nabla V(x)\rangle \\
&\quad + \int_U \left(V(x+c(x,k,u)) - V(x) - \langle \nabla V(x), c(x,k,u)\rangle\right)\nu(du) \\
&= |\sigma(x,k)|^2 + 2\langle x, b(x,k)\rangle + \int_U |c(x,k,u)|^2\nu(du) - 2\langle x, L(k)x\rangle \\
&\le 2\kappa(|x|^2+1) - 2\langle x, L(k)x\rangle \\
&\le 2\alpha_k|x|^2 + 2\kappa \\
&= 2\alpha_k V(x) + 2\kappa.
\end{aligned}$$

This, together with (6.7), implies that $(\tilde{X}, \tilde{\Lambda})$ satisfies Assumption 5.2.2. It follows directly from Theorem 5.1.2 that the controlled process $(\tilde{X}, \tilde{\Lambda})$ is exponentially ergodic. □

As in Zhu and Yin (2009), one of the simplest example of feedback controls is of the following form

$$\xi(x,k) = -\theta(k)Ix,$$

where $\theta(k)$ is a non-negative constant and I is the $d \times d$ identity matrix. That is $L(k)$ takes the form

$$L(k) = \theta(k)I.$$

Since κ is a fixed constant and $\{\alpha_k\}$ is bounded, then (6.8) is immediate if we choose $\theta(k)$ large enough; that is, $\theta(k) \ge \kappa - \alpha_k$. To summarize this discussion, we state the following corollary.

Corollary 6.2.1.1. *Suppose that Assumptions 3.1.1, 3.1.2, 4.1.2, 4.1.3, 4.1.4, and 5.2.1 hold. Assume further that $\sum_{k\in\mathbb{S}_{int}} \nu_k > 0$. Then there exists a feedback control ξ so that the controlled process $(\tilde{X}, \tilde{\Lambda})$ is exponentially ergodic.*

Proof. We take $\alpha_k = \kappa$ for all $k \in \mathbb{S}_{ab}$ and $\alpha_k = -2\kappa / \sum_{k\in\mathbb{S}_{int}} \nu_k$ for all $k \in \mathbb{S}_{int}$. Then $\sum_{k\in\mathbb{S}} \alpha_k \nu_k < 0$. Choose $\theta(k)$ big enough so that $\theta(k) \geq \kappa - \alpha_k$ for $k \in \mathbb{S}_{int}$ and $\theta(k) = 0$ for $k \in \mathbb{S}_{ab}$. So $\xi(x,k) := -\theta(k)Ix$ is the desired feedback control. □

6.3 Examples

Example 6.3.1. *Consider the following SDE*

$$dX(t) = \sigma(X(t), \Lambda(t))dW(t), \tag{6.9}$$

where W is the standard 1-dimensional Brownian motion and $\sigma(x,k) = 1$ for all $(x,k) \in \mathbb{R}\times\mathbb{S}$. Suppose that Λ takes value in $\mathbb{S} = \{1, 2, ...\}$ and is generated by $Q(x) = (q_{kj}(x))$ given by

$$q_{kj}(x) := \begin{cases} \frac{1}{2}\frac{k}{k+e^{-|x|^2}} & \text{if} \quad j = 1, k \neq j, \\ \frac{1}{3^{j-1}}\frac{k}{k+e^{-|x|^2}} & \text{if} \quad j > 1, k \neq j \\ -\sum_{j\neq k} q_{kj}(x) & \text{if} \quad k = j. \end{cases}$$

The solution to (6.9) *is trivially given by $X(t) = W(t)$. It is well known that the Brownian motion is not ergodic. So, the process (X, Λ) is not exponential ergodic. However, all of the Assumptions 3.1.1, 3.1.2, 4.1.2, 4.1.3, 4.1.4, and 5.2.1 are trivially satisfied. By Corollary 6.2.1.1, there exists a feedback control ξ so that the controlled process $(\tilde{X}, \tilde{\Lambda})$ is exponentially ergodic. Indeed, the feed back control ξ is given by $\xi(x,k) = -\theta(k)x$ where $\theta(k)$ is a constant large enough. Then the controlled process is given by*

$$dX(t) = -\theta(\Lambda(t))X(t)dt + dW(t).$$

This is an Ornstein–Uhlenbeck process which is Gaussian. Moreover, one can show that this process is exponentially ergodic.

Example 6.3.2. *Consider the following SDE*

$$dX(t) = b(X(t), \Lambda(t))dt + \sigma(X(t), \Lambda(t))dW(t), \tag{6.10}$$

where W is the standard 2-dimensional Brownian motion. Suppose that Λ takes values in the set $\mathbb{S} := \{1, 2\}$ and is generated by the constant rate matrix

$$Q(x) = \begin{pmatrix} -\frac{3}{2} & \frac{3}{2} \\ \frac{3}{2} & -\frac{3}{2} \end{pmatrix}.$$

Define the coefficients in (6.10) as follows:

$$\sigma(x,k) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad b(x,1) = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} x, \quad b(x,2) = \begin{pmatrix} 2 & 2 \\ -2 & 3 \end{pmatrix} x.$$

For any matrix A, we denote by $\lambda_{min}(A)$ the minimal of the eigenvalues of A. We verify that

$$\lambda_{min}\left(\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} + \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}^T\right) = 2 > 0$$

and

$$\lambda_{min}\left(\begin{pmatrix} 2 & 2 \\ -2 & 3 \end{pmatrix} + \begin{pmatrix} 2 & 2 \\ -2 & 3 \end{pmatrix}^T\right) = 4 > 0.$$

In view of Theorem 4.13 of Zhu and Yin (2009), we conclude that the process (X, Λ) is transient and hence it is not ergodic. One can show that Assumptions 3.1.1, 3.1.2, 4.1.2, 4.1.3, 4.1.4, and 5.2.1 are satisfied. In particular, the matrix Q itself is strongly exponentially *ergodic with invariant measure* $\nu = (1/2, 1/2)$ *and transition matrix*

$$P(t) = \begin{pmatrix} \frac{1}{2} + e^{-3t} & \frac{1}{2} - e^{-3t} \\ \frac{1}{2} - e^{-3t} & \frac{1}{2} + e^{-3t} \end{pmatrix}.$$

It follows from Corollary 6.2.1.1 that there exists a feedback control ξ *for which the controlled process* $(\tilde{X}, \tilde{\Lambda})$ *is exponentially ergodic.*

Chapter 7

Appendix

7.1 Elementary Properties of Coupling Operators

In this section we give the detailed proof of Lemma 3.2.1.

Proof of Lemma 3.2.1. i) It is clear that $a(x,i) = a(x,i)^T$ and $a(y,j) = a(y,j)^T$. Then we have

$$\begin{aligned} a(x,i,y,j)^T &= \begin{pmatrix} a(x,i) & \hat{g}(x,i,y,j) \\ \hat{g}(x,i,y,j)^T & a(y,j) \end{pmatrix}^T \\ &= \begin{pmatrix} a(x,i)^T & (\hat{g}(x,i,y,j)^T)^T \\ \hat{g}(x,i,y,j)^T & a(y,j)^T \end{pmatrix} \\ &= \begin{pmatrix} a(x,i) & \hat{g}(x,i,y,j) \\ \hat{g}(x,i,y,j)^T & a(y,j) \end{pmatrix} \\ &= a(x,i,y,j) \end{aligned}$$

So $a(x,i,y,j)$ is symmetric. To show $a(x,i,y,j)$ is uniformly positive definite, we observe that $(I-2u(x,y)u(x,y)^T)^2 = I$. Let $\begin{pmatrix}\xi\\ \eta\end{pmatrix} = (\xi,\eta)^T = (\xi_1,...,\xi_d,\eta_1,...,\eta_d)^T \in \mathbb{R}^{2d}$. Then

$$\begin{aligned}
&\begin{pmatrix}\xi\\ \eta\end{pmatrix}^T a(x,i,y,j)\begin{pmatrix}\xi\\ \eta\end{pmatrix}\\
= &\begin{pmatrix}\xi\\ \eta\end{pmatrix}^T \begin{pmatrix} a(x,i) & \hat{g}(x,i,y,j)\\ \hat{g}(x,i,y,j)^T & a(y,j)\end{pmatrix}\begin{pmatrix}\xi\\ \eta\end{pmatrix}\\
= &\ \xi^T a(x,i)\xi + \xi^T\hat{g}(x,i,y,j)\eta + \eta^T a(y,j)\eta + \eta^T\hat{g}(x,i,y,j)^T\xi\\
= &\ \xi^T a(x,i)\xi + \xi^T(\lambda_R(I-2u(x,y)u(x,y)^T) + \sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j)^T)\eta\\
&+\eta^T a(y,j)\eta + \eta^T(\lambda_R(I-2u(x,y)u(x,y)^T) + \sigma_{\lambda_R}(y,j)\sigma_{\lambda_R}(x,i)^T)\xi\\
= &\ \xi^T\sigma_{\lambda_R}(x,i)^2\xi + \lambda_R\xi^T\xi + \lambda_R\xi^T(I-2u(x,y)u(x,y)^T)\eta + \xi^T\sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j)^T\eta\\
&+\eta^T\sigma_{\lambda_R}(y,j)^2\eta + \lambda_R\eta^T\eta + \lambda_R\eta^T(I-2u(x,y)u(x,y)^T)\xi + \eta^T\sigma_{\lambda_R}(y,j)\sigma_{\lambda_R}(x,i)^T\xi\\
= &\ \xi^T\sigma_{\lambda_R}(x,i)^2\xi + \eta^T\sigma_{\lambda_R}(y,j)^2\eta + \xi^T\sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j)^T\eta + \eta^T\sigma_{\lambda_R}(y,j)\sigma_{\lambda_R}(x,i)^T\xi\\
&+\lambda_R\left(\xi^T\xi + \eta^T\eta + \xi^T(I-2u(x,y)u(x,y)^T)\eta + \eta^T(I-2u(x,y)u(x,y)^T)\xi\right)\\
= &\ |\sigma_{\lambda_R}(x,i)^2\xi + \sigma_{\lambda_R}(y,j)^2\eta|^2 + \lambda_R[\xi^T(I-2u(x,y)u(x,y)^T)(I-2u(x,y)u(x,y)^T)\xi\\
&+\xi^T(I-2u(x,y)u(x,y)^T)\eta + \eta^T(I-2u(x,y)u(x,y)^T)\xi + \eta^T\eta]\\
= &\ |\sigma_{\lambda_R}(x,i)^2\xi + \sigma_{\lambda_R}(y,j)^2\eta|^2 +\\
&\lambda_R\langle(I-2u(x,y)u(x,y)^T)\xi+\eta, (I-2u(x,y)u(x,y)^T)\xi+\eta\rangle\\
= &\ |\sigma_{\lambda_R}(x,i)^2\xi + \sigma_{\lambda_R}(y,j)^2\eta|^2 + \lambda_R|(I-2u(x,y)u(x,y)^T)\xi+\eta|^2\\
\geq &\ 0.
\end{aligned}$$

ii) Note that for any matrices A and B we have $tr(AB) = tr(BA)$. Then $tr(u(x,y)u(x,y)^T) = tr(u(x,y)^T u(x,y)) = 1$. We obtain

$$trA(x,i,y,j) = tr(a(x,i)) + tr(a(y,j)) - 2tr(\hat{g}(x,i,y,j)$$

$$
\begin{aligned}
&= tr(\sigma_{\lambda_R}(x,i)^2 + \lambda_R I) + tr(\sigma_{\lambda_R}(y,j)^2 + \lambda_R I) \\
&\quad -2tr(\lambda_R I - 2\lambda_R u(x,y)u(x,y)^T + \sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j)^T) \\
&= 2\lambda_R d + tr(\sigma_{\lambda_R}(x,i)^2) + tr(\sigma_{\lambda_R}(y,j)^2) \\
&\quad -2\lambda_R d + 4\lambda_R tr(u(x,y)u(x,y)^T) + 2tr(\sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j)^T) \\
&= tr(\sigma_{\lambda_R}(x,i)^2 + \sigma_{\lambda_R}(y,j)^2 - 2\sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j)^T) + 4\lambda_R \\
&= tr((\sigma_{\lambda_R}(x,i) - \sigma_{\lambda_R}(y,j))^2) + 4\lambda_R \\
&= |\sigma_{\lambda_R}(x,i) - \sigma_{\lambda_R}(y,j)|^2 + 4\lambda_R
\end{aligned}
$$

iii) Note that a $d \times d$-matrix M is symmetric if and only if $\langle Mx, y\rangle = \langle x, My\rangle$ for all $x, y \in \mathbb{R}^d$. Since σ_{λ_R} is symmetric, we observe that

$$
\begin{aligned}
\langle x-y, \sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j)(x-y)\rangle &= \langle \sigma_{\lambda_R}(x,i)(x-y), \sigma_{\lambda_R}(y,j)(x-y)\rangle \\
&= \langle \sigma_{\lambda_R}(y,j)(x-y), \sigma_{\lambda_R}(x,i)(x-y)\rangle \\
&= \langle x-y, \sigma_{\lambda_R}(y,j)\sigma_{\lambda_R}(x,i)(x-y)\rangle.
\end{aligned}
$$

Note that $\sigma_{\lambda_R}(x,i) - \sigma_{\lambda_R}(y,j)$ is also symmetric. This implies

$$
\begin{aligned}
&|(\sigma_{\lambda_R}(x,i) - \sigma_{\lambda_R}(y,j))(x-y)|^2 \\
&= \langle (\sigma_{\lambda_R}(x,i) - \sigma_{\lambda_R}(y,j))(x-y), (\sigma_{\lambda_R}(x,i) - \sigma_{\lambda_R}(y,j))(x-y)\rangle \\
&= \langle x-y, (\sigma_{\lambda_R}(x,i) - \sigma_{\lambda_R}(y,j))(\sigma_{\lambda_R}(x,i) - \sigma_{\lambda_R}(y,j))(x-y)\rangle \\
&= \langle x-y, (\sigma_{\lambda_R}(x,i)^2 + \sigma_{\lambda_R}(y,j)^2 - \sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j) - \sigma_{\lambda_R}(y,j)\sigma_{\lambda_R}(x,i))(x-y)\rangle \\
&= \langle x-y, (\sigma_{\lambda_R}(x,i)^2 + \sigma_{\lambda_R}(y,j)^2 - 2\sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j))(x-y)\rangle.
\end{aligned}
$$

We then obtain

$$
\begin{aligned}
&\bar{A}(x,i,y,j)|x-y|^2 \\
&= \langle x-y, A(x,i,y,j)(x-y)\rangle
\end{aligned}
$$

$$
\begin{aligned}
&= \langle (x-y, (a(x,i)+a(y,j)-2g(x,i,y,j))(x-y)\rangle \\
&= \langle (x-y, (a(x,i)+a(y,j)-2(\lambda_R(I-2u(x,y)u(x,y)^T)+\sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j)^T))(x-y)\rangle \\
&= \langle x-y, (\sigma_{\lambda_R}(x,i)^2+\sigma_{\lambda_R}(y,j)^2-2\sigma_{\lambda_R}(x,i)\sigma_{\lambda_R}(y,j)))(x-y))\rangle + \\
&\quad 4\lambda_R\langle x-y, u(x,y)u(x,y)^T(x-y)\rangle \\
&= |(\sigma_{\lambda_R}(x,i)-\sigma_{\lambda_R}(y,j))(x-y)|^2 + 4\lambda_R\langle x-y, \frac{(x-y)}{|x-y|}\frac{(x-y)^T}{|x-y|}(x-y)\rangle \\
&= |(\sigma_{\lambda_R}(x,i)-\sigma_{\lambda_R}(y,j))(x-y)|^2 + \frac{4\lambda_R}{|x-y|^2}\langle x-y, (x-y)(x-y)^T(x-y)\rangle \\
&= |(\sigma_{\lambda_R}(x,i)-\sigma_{\lambda_R}(y,j))(x-y)|^2 + \frac{4\lambda_R}{|x-y|^2}|x-y|^2\langle x-y, x-y\rangle \\
&= |(\sigma_{\lambda_R}(x,i)-\sigma_{\lambda_R}(y,j))(x-y)|^2 + 4\lambda_R|x-y|^2 \\
&\geq 4\lambda_R|x-y|^2.
\end{aligned}
$$

Hence

$$\bar{A}(x,i,y,j) \geq 4\lambda_R.$$

□

7.2 Strong Exponential Ergodicity of $\hat{\Lambda}$

Definition 7.2.1. *Suppose that a Markov process $\hat{\Lambda}$ is generated by a bounded generator $Q(0)=\{q_{ij}\}_{i,j\in\mathbb{S}}$ ($0 \leq q_{ij} < \infty,\ i \neq j$) which is totally stable and conservative, that is,*

$$q_i := -q_{ii} < \infty \quad \textit{and} \quad \sum_{j\neq i} q_{ij} = q_i \tag{7.1}$$

If the matrix $Q(0)$ is totally stable and conservative, it is sometimes called regular as in Chen (2004).

Let us state the following classical result for continuous-time Markov processes with countable state spaces.

Lemma 7.2.2 (Lemma 4.36, Chen (2004))**.** *Let $Q(0) = \{q_{ij}\}_{i,j\in\mathbb{S}}$ be a regular irreducible Q-matrix. Then the limit*

$$\lim_{t\longrightarrow\infty} P_{ij}(t) =: \pi_j \tag{7.2}$$

exists for all $i, j \in \mathbb{S}$ and the limit is independent of i. Moreover, we have either $\sum_{j\in\mathbb{S}} \pi_j = 1$ or $\sum_{j\in\mathbb{S}} \pi_j = 0$.

In Nguyen and Yin (2018c), the stability of regime-switching diffusion processes $(X(t), \alpha(t))$ was investigated where the second component $\alpha(t)$ takes values in a countable state space. The argument used to obtain the result is based on the existence of a continuous-time Markov chain $\hat{\alpha}$ which posses a certain property called strongly exponential ergodicity.

Definition 7.2.3 (Definition 2.5, Nguyen and Yin (2018c))**.** *Let $\hat{\alpha}(t)$ be a Markov chain generated by a bounded generator $Q(0) = \{\hat{q}_{ij}\}_{i,j\in\mathbb{S}}$ and transition function $\hat{p}_{ij}(t)$. Then $\hat{\alpha}(t)$ is said to be strongly exponentially ergodic if it has an invariant probability measure $\pi = (\pi_1, \pi_2, ...)$ such that for some constants $C, \lambda > 0$, we have*

$$\sum_{j\in\mathbb{S}} |\hat{p}_{ij}(t) - \pi_j| \le Ce^{-\lambda t} \quad \text{for all } i \in \mathbb{S} \ \text{ and all } \ t \ge 0. \tag{7.3}$$

Then we obtain the following lemma as a direct consequence of Theorem 6.1 of Meyn and Tweedie (1993b). Also, a similar result was given in Theorem 7.1 of Meyn and Tweedie (1993b).

Lemma 7.2.4. *Suppose that the Markov process Λ is irreducible. Assume there exists a function $\phi : \mathbb{S} \longrightarrow [0, \infty)$ with $\phi(k) \longrightarrow \infty$ as $k \longrightarrow \infty$ such that*

$$Q\phi(k) := \sum_{j\in\mathbb{S}} q_{kj} [\phi(j) - \phi(k)] \le C - \phi(k) \quad \text{for all } \ k \in \mathbb{S},$$

where C is a constant and Q is the infinitesimal generator of Λ. Then there exist $\beta < 1$ and

$B < \infty$ such that

$$||P^t(k, \cdot) - \pi||_f \;\; \leq \;\; Bf(k)\beta^t, \qquad t \geq 0, k \in \mathbb{S}$$

where $f = \phi + 1$.

Remark 7.2.5. *We note that ϕ is a Foster-Lyapunov function for the discrete component Λ. Also, this lemma says that Λ is f-exponential ergodic where $f = \phi + 1$. For measurable functions $f, g \geq 1$, if $f \geq g$ then it is clear from the definition that $|| \cdot ||_f \geq || \cdot ||_g$. In particular, this lemma implies $||P^t(k, \cdot) - \pi||_1 \leq Bf(k)\beta^t$. However, this result may not give the strong exponential ergodicity of Λ.*

The following lemma plays an important role in the proof of Theorem 5.2.4. It was introduced in Nguyen and Yin (2018c) and was the main tool used to obtain the stability result in Theorem 3.1 of Nguyen and Yin (2018c). For the sake of completeness, let us state and give the detailed proof of this lemma.

Lemma 7.2.6 (Lemma A.1, Nguyen and Yin (2018c))**.** *Suppose that the Markov chain $\hat{\alpha}(t)$ is strongly exponentially ergodic which is generated by a bounded and regular generator $Q(0) = \{\hat{q}_{ij}\}_{i,j \in \mathbb{S}}$ and invariant probability measure $\pi = (\pi_1, \pi_2, ...)$. If $\mathbf{b} = (b_1, b_2, ...)^T$ is bounded satisfying $\pi \mathbf{b} = \sum_{j \in \mathbb{S}} \pi_j b_j = 0$, then there exists a bounded vector $\mathbf{c} = (c_1, c_2, ...)^T$ such that $b_i = \sum_{j \in \mathbb{S}} \hat{q}_{ij} c_j$ for all $i \in \mathbb{S}$.*

Proof. Denote $\hat{P}(t) := [\hat{p}_{ij}(t)]_{i,j}$ the transition matrix. First, let us state the following properties of $\hat{P}(t)$. We know that

$$0 \leq \hat{p}_{ij}(t) \leq 1 \quad \text{for all} \;\; t \geq 0 \;\; \text{and all } i, j \in \mathbb{S} \tag{7.4}$$

$$\sum_{j=1}^{\infty} \hat{p}_{ij}(t) = 1 \quad \text{for all} \;\; t \geq 0 \;\; \text{and all} \;\; i \in \mathbb{S}. \tag{7.5}$$

From Kolmogorov backward equations, we have $Q\hat{P}(t) = \hat{P}'(t)$, that is,

$$\hat{p}'_{ik} = \sum_{j=0}^{\infty} \hat{q}_{ij}\hat{p}_{jk}(t) \quad \text{for all} \quad i, k \in \mathbb{S}. \tag{7.6}$$

To prove this lemma, we will need to apply the Bounded Convergence Theorem. So, for any $0 < T < \infty$, we consider the finite measure space $([0, T], \mathcal{B}([0, T]), d)$ where d denotes the Lebesgue measure. Define $\mathbf{c}(T) := -\int_0^T \hat{P}(t)\mathbf{b}dt$, that is,

$$c_i(T) := -\int_0^T \sum_{k\in\mathbb{S}} \hat{p}_{ik}(t)b_k dt, \qquad i \in \mathbb{S}.$$

First, we show that $\mathbf{c}(T)$ is bounded for all $0 < T < \infty$. We observe

$$\sum_{j\in\mathbb{S}} |\pi_j b_j| \leq \sup_{j\in\mathbb{S}} |b_j| \sum_{j\in\mathbb{S}} |\pi_j| \leq \sup_{j\in\mathbb{S}} |b_j| < \infty$$

and from (7.3) we have

$$\sum_{j\in\mathbb{S}} |\left(\hat{p}_{ij}(t) - \pi_j\right) b_j| \leq \sup_{j\in\mathbb{S}} |b_j| \sum_{j\in\mathbb{S}} |\hat{p}_{ij}(t) - \pi_j| \leq \sup_{j\in\mathbb{S}} |b_j| C e^{-\lambda t} < \infty.$$

These imply that

$$\begin{aligned}
\sum_{j\in\mathbb{S}} [\hat{p}_{ij}(t)b_j] &= \sum_{j\in\mathbb{S}} [\left(\hat{p}_{ij}(t) - \pi_j\right) b_j + \pi_j b_j] \\
&= \sum_{j\in\mathbb{S}} \left(\hat{p}_{ij}(t) - \pi_j\right) b_j + \sum_{j\in\mathbb{S}} \pi_j b_j \\
&= \sum_{j\in\mathbb{S}} \left(\hat{p}_{ij}(t) - \pi_j\right) b_j.
\end{aligned}$$

Then

$$|c_i(T)| = |-\int_0^T \sum_{j\in\mathbb{S}} \hat{p}_{ij}(t) b_j dt|$$

$$
\begin{aligned}
&= |\int_0^T \sum_{j\in\mathbb{S}} (\hat{p}_{ij}(t) - \pi_j)\, b_j dt| \\
&\leq \int_0^T \sum_{j\in\mathbb{S}} |\hat{p}_{ij}(t) - \pi_j||b_j| dt \\
&\leq \sup_{j\in\mathbb{S}} |b_j| \int_0^T \sum_{j\in\mathbb{S}} |\hat{p}_{ij}(t) - \pi_j| dt \\
&\leq \sup_{j\in\mathbb{S}} |b_j| \int_0^T Ce^{-\lambda t} dt \\
&= \sup_{j\in\mathbb{S}} |b_j| \frac{C}{\lambda}[1 - e^{-\lambda T}] \\
&\leq \sup_{j\in\mathbb{S}} |b_j| \frac{C}{\lambda}. \qquad (7.7)
\end{aligned}
$$

Next, we determine the matrix $Q\mathbf{c}(T)$ for each $0 < T < \infty$. We will do this componentwise. Let $i \in \mathbb{S}$ be arbitrary but fixed. Define $f_n(t) := \sum_{j=1}^{n} \hat{q}_{ij} \left(\sum_{k=1}^{\infty} \hat{p}_{jk}(t) b_k \right)$. From (7.4) and (7.5), we see that

$$
\begin{aligned}
|f_n(t)| &= |\sum_{j=1}^{n} \hat{q}_{ij} \left(\sum_{k=1}^{\infty} \hat{p}_{jk}(t) b_k \right)| \\
&= \sum_{j=1}^{n} |\hat{q}_{ij}| \left(\sum_{k=1}^{\infty} |\hat{p}_{jk}(t)||b_k| \right) \\
&\leq \sup_{k\in\mathbb{S}} |b_k| \sum_{j=1}^{n} |\hat{q}_{ij}| \left(\sum_{k=1}^{\infty} |\hat{p}_{jk}(t)| \right) \\
&\leq \sup_{k\in\mathbb{S}} |b_k| \sum_{j=1}^{n} |\hat{q}_{ij}| \\
&\leq \sup_{k\in\mathbb{S}} |b_k| \sum_{j=1}^{\infty} |\hat{q}_{ij}| \\
&\leq \sup_{k\in\mathbb{S}} |b_k| (2\hat{q}_i).
\end{aligned}
$$

The Bounded Convergence Theorem implies that

$$
\sum_{j=1}^{\infty} \hat{q}_{ij} c_j(T) = \sum_{j=1}^{\infty} \hat{q}_{ij} \left[-\int_0^T \sum_{k=1}^{\infty} \hat{p}_{jk}(t) b_k dt \right]
$$

$$
\begin{aligned}
&= -\sum_{j=1}^{\infty}\left[\int_0^T \hat{q}_{ij}\sum_{k=1}^{\infty}\hat{p}_{jk}(t)b_k dt\right] \\
&= -\lim_{n\longrightarrow\infty}\sum_{j=1}^{n}\left[\int_0^T \hat{q}_{ij}\sum_{k=1}^{\infty}\hat{p}_{jk}(t)b_k dt\right] \\
&= -\lim_{n\longrightarrow\infty}\int_0^T\left[\sum_{j=1}^{n}\hat{q}_{ij}\sum_{k=1}^{\infty}\hat{p}_{jk}(t)b_k\right]dt \\
&= -\lim_{n\longrightarrow\infty}\int_0^T\left[f_n(t)\right]dt \\
&= -\int_0^T\left[\lim_{n\longrightarrow\infty}f_n(t)\right]dt \\
&= -\int_0^T\left[\sum_{j=1}^{\infty}\hat{q}_{ij}\left(\sum_{k=1}^{\infty}\hat{p}_{jk}(t)b_k\right)\right]dt. \qquad (7.8)
\end{aligned}
$$

We claim that we can interchange the summation $\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\hat{q}_{ij}\hat{p}_{jk}(t)b_k = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\hat{q}_{ij}\hat{p}_{jk}(t)b_k$. From (7.4) and (7.5), we observe that

$$
\begin{aligned}
\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}|\hat{q}_{ij}\hat{p}_{jk}(t)b_k| &\leq \sup_{k\in\mathbb{S}}|b_k|\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}|\hat{q}_{ij}||\hat{p}_{jk}(t)| \\
&\leq \sup_{k\in\mathbb{S}}|b_k|\sum_{j=1}^{\infty}|\hat{q}_{ij}|\sum_{k=1}^{\infty}\hat{p}_{jk}(t) \\
&\leq \sup_{k\in\mathbb{S}}|b_k|\sum_{j=1}^{\infty}|\hat{q}_{ij}| \\
&\leq \sup_{k\in\mathbb{S}}|b_k|(2\hat{q}_i) \\
&< \infty.
\end{aligned}
$$

Therefore, $\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\hat{q}_{ij}\hat{p}_{jk}(t)b_k$ is absolutely convergence and then

$$
\sum_{j=1}^{\infty}\sum_{k=1}^{\infty}\hat{q}_{ij}\hat{p}_{jk}(t)b_k = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty}\hat{q}_{ij}\hat{p}_{jk}(t)b_k.
$$

In view of (7.8), we have from (7.6) that

$$\begin{aligned}
\sum_{j=1}^{\infty} \hat{q}_{ij} c_j(T) &= -\int_0^T \left[\sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \hat{q}_{ij} \hat{p}_{jk}(t) b_k \right] dt \\
&= -\int_0^T \left[\sum_{k=1}^{\infty} \sum_{j=1}^{\infty} \hat{q}_{ij} \hat{p}_{jk}(t) b_k \right] dt \\
&= -\int_0^T \left[\sum_{k=1}^{\infty} b_k \left(\sum_{j=1}^{\infty} \hat{q}_{ij} \hat{p}_{jk}(t) \right) \right] dt \\
&= -\int_0^T \left[\sum_{k=1}^{\infty} b_k \hat{p}'_{ik}(t) \right] dt. \qquad (7.9)
\end{aligned}$$

Now, let $g_n(t) := \sum_{k=1}^{n} b_k \hat{p}'_{ik}(t)$. Observe that

$$\begin{aligned}
|g_n(t)| &= |\sum_{k=1}^{n} b_k \hat{p}'_{ik}(t)| \\
&\leq \sum_{k=1}^{n} |b_k \hat{p}'_{ik}(t)| \\
&\leq \sup_{k \in \mathbb{S}} |b_k| \sum_{k=1}^{n} |\hat{p}'_{ik}(t)| \\
&= \sup_{k \in \mathbb{S}} |b_k| \sum_{k=1}^{n} | \left(\sum_{j=1}^{\infty} \hat{q}_{ij} \hat{p}_{jk}(t) \right) | \\
&\leq \sup_{k \in \mathbb{S}} |b_k| \sum_{k=1}^{n} \sum_{j=1}^{\infty} |\hat{q}_{ij} \hat{p}_{jk}(t)| \\
&= \sup_{k \in \mathbb{S}} |b_k| \sum_{j=1}^{\infty} \sum_{k=1}^{n} |\hat{q}_{ij} \hat{p}_{jk}(t)| \\
&= \sup_{k \in \mathbb{S}} |b_k| \sum_{j=1}^{\infty} \left(|\hat{q}_{ij}| \sum_{k=1}^{n} |\hat{p}_{jk}(t)| \right) \\
&\leq \sup_{k \in \mathbb{S}} |b_k| \sum_{j=1}^{\infty} |\hat{q}_{ij}| \\
&\leq \sup_{k \in \mathbb{S}} |b_k| (2\hat{q}_i).
\end{aligned}$$

Then, in (7.9), the Bounded Convergence Theorem implies that

$$\begin{aligned}
\sum_{j=1}^{\infty} \hat{q}_{ij} c_j(T) &= -\int_0^T \left[\sum_{k=1}^{\infty} b_k \hat{p}'_{ik}(t)\right] dt \\
&= -\int_0^T \left[\lim_{n \longrightarrow \infty} g_n(t)\right] dt \\
&= -\lim_{n \longrightarrow \infty} \int_0^T \left[\sum_{k=1}^{n} b_k \hat{p}'_{ik}(t)\right] dt \\
&= -\lim_{n \longrightarrow \infty} \sum_{k=1}^{n} \int_0^T \left[b_k \hat{p}'_{ik}(t)\right] dt \\
&= -\sum_{k=1}^{\infty} b_k \int_0^T \left[\hat{p}'_{ik}(t)\right] dt \\
&= \sum_{k=1}^{\infty} b_k \left[\hat{p}_{ik}(0) - \hat{p}_{ik}(T)\right] \\
&= \sum_{k=1}^{\infty} b_k \left[\delta_{ik} - \hat{p}_{ik}(T)\right]. \qquad (7.10)
\end{aligned}$$

Note that

$$\sum_{k=1}^{\infty} |b_k \delta_{ik}| = |b_i| < \infty$$

and

$$\sum_{k=1}^{\infty} |b_k \hat{p}_{ik}(T)| \quad \leq \quad \sup_{k \in \mathbb{S}} |b_k| \sum_{k=1}^{\infty} |\hat{p}_{ik}(T)| \quad \leq \quad \sup_{k \in \mathbb{S}} |b_k| \, [1] \quad < \quad \infty.$$

Hence $\sum_{k=1}^{\infty} b_k \left[\delta_{ik} - \hat{p}_{ik}(T)\right] = \sum_{k=1}^{\infty} b_k \delta_{ik} - \sum_{k=1}^{\infty} b_k \hat{p}_{ik}(T)$. Then, in (7.10), we have

$$\begin{aligned}
\sum_{j=1}^{\infty} \hat{q}_{ij} c_j(T) &= \sum_{k=1}^{\infty} b_k \left[\delta_{ik} - \hat{p}_{ik}(T)\right] \\
&= \sum_{k=1}^{\infty} b_k \left[\delta_{ik}\right] - \sum_{k=1}^{\infty} b_k \left[\hat{p}_{ik}(T)\right]
\end{aligned}$$

$$= \quad b_i - \sum_{k=1}^{\infty} b_k \left[\hat{p}_{ik}(T)\right]. \tag{7.11}$$

We claim that the functions $\sum_{j=1}^{\infty} \hat{q}_{ij} c_j(t)$ and $\sum_{k=1}^{\infty} b_k \left[\hat{p}_{ik}(t)\right]$ are continuous on $[0, T]$. It is done if one of them is continuous. Since $\hat{p}_{ik}(t)$ is continuous, then so is the function

$$h_n(t) := \sum_{k=1}^{n} b_k \hat{p}_{ik}(t), \quad t \in [0, T].$$

We want so show that $h_n(t) \longrightarrow \sum_{k=1}^{\infty} b_k \left[\hat{p}_{ik}(t)\right]$ as $n \longrightarrow \infty$ uniformly on $[0, T]$. Note that

$$\sum_{k=1}^{\infty} |b_k \hat{p}_{ik}(t)| \leq \sup_{k \in \mathbb{S}} |b_k| \sum_{k=1}^{\infty} |\hat{p}_{ik}(t)| \leq \sup_{k \in \mathbb{S}} |b_k| < \infty.$$

This implies that $h_n(t)$ converges to $\sum_{k=1}^{\infty} b_k \left[\hat{p}_{ik}(t)\right]$ on $[0, T]$. Since $[0, T]$ is compact, then the convergence is uniform. Therefor, $\sum_{k=1}^{\infty} b_k \left[\hat{p}_{ik}(t)\right]$ is continuous on $[0, T]$ for all $T > 0$.

In view of (7.7), the limit $\lim_{T \longrightarrow \infty} c_j(T)$ exists (see, for example, Theorem 10.33 Apostol (1974)). Let $c_j := \lim_{T \longrightarrow \infty} c_j(T)$. It follows from (7.7) that c_j is bounded.

In view of (7.11), we have from (7.2) that

$$\begin{aligned}
\sum_{j=1}^{\infty} \hat{q}_{ij} c_j &= \lim_{T \longrightarrow \infty} \left[\sum_{j=1}^{\infty} \hat{q}_{ij} c_j(T)\right] \\
&= \lim_{T \longrightarrow \infty} \left[b_i - \sum_{k=1}^{\infty} b_k \left[\hat{p}_{ik}(T)\right]\right] \\
&= b_i - \lim_{T \longrightarrow \infty} \left[\sum_{k=1}^{\infty} b_k \left[\hat{p}_{ik}(T)\right]\right] \\
&= b_i - \sum_{k=1}^{\infty} b_k \pi_k \\
&= b_i.
\end{aligned}$$

This completes the proof. □

Bibliography

Applebaum, D. (2009). *Lévy processes and stochastic calculus,* volume 116 of Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge, second edition.

Anderson, W.J. (1991). *Continuous-time Markov chains: An applications-oriented approach.* Springer, New York.

Apostol, T (1974). *Mathematical Analysis.*

Azéma, J. Kaplan-Duflo, M. and Revuz, D. (1967). *Measure invariant sur les classes récurrentes des processus de Markov.* Z. Wahrscheinlichkeitsth. 8 :157-181.

Cerrai, S. (2001). *Second order PDE's in finite and infinite dimension: A probabilistic approach, Lectures notes in mathematics.* Springer, Berlin, 1762.

Blumenthal, R. M. and Getoor, R. K. (1968). *Markov Processes and Potential Theory.* Academic Press, New York.

Chen, M.F. (2004). *From Markov chains to non-equilibrium particle systems(second edition).*

Chen, M. F. and Li, S. F. (1989). Coupling methods for multidimensional diffusion processes. *Ann. Probab.*, 17(1):151–177.

Cloez, B. and Hairer, M. (2015). Exponential ergodicity for Markov processes with random switching. *Bernoulli*, 21(1):505–536.

Dynkin, E. (1965). *Markov Processes*, volume 121 of *Die Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, Berlin.

Ethier, S. N. and Kurtz, T. G. (1986). *Markov Processes: Characterization and Convergence.* J. Wiley, New York.

Ferrari, G and Rodosthenous, N. (2019). Optimal control of debt-to-GDP ratio in an N-state regime switching economy. *SIAM J. Control Optim.*, 58(2):755–786.

Foguel, S.R.(1969). Positive operators on $C(X)$. *Proc. Amer. Math. Soc.*, 22:295-297.

Getoor, R. K.(1979). Transience and recurrence of Markov processes. *In Séminaire de Probabilités XIV*, ed. J. Azéma and M. Yor, Springer-Verlag, Berlin, 22:295-297.

Hairer, M. (2016). Convergence of markov processes. Lecture notes, Univ. Warwick. Available in http://www.hairer.org/notes/Convergence.pdf.

Hairer, M., Mattingly, J. C., and Scheutzow, M. (2011). Asymptotic coupling and a general form of Harris's theorem with applications to stochastic delay equations. *Probab. Theory Related Fields*, 149(1-2):233–259.

Karlin, S. and Taylor, H. M. (1975). *A First Course in Stochastic Processes.* Academic Press, second edition.

Khasminskii, R. (2012). *Stochastic stability of differential equations*, volume 66 of *Stochastic Modeling and Applied Probability.* Springer, New York, 2nd edition.

Lawley, S. D., Mattingly, J. C. and Reed, M. C. (2014). Sensitivity to switching rates in stochastically switched ODEs. *Commun. Math. Sci.*, 12(7): 1343–1352.

Lindvall, T. (1992). *Lectures on the Coupling Method.* John Wiley & Sons, New York, 1992.

Lindvall, T. and Rogers, L. C. G. (1986). *Coupling of multidimensional diffusions by reflection.* Ann. Probab. 14:860-872.

Mao, X. and Yuan, C. (2006). *Stochastic differential equations with Markovian switching.* Imperial College Press, London.

Meyn, S. P. and Tweedie, R. L. (1992). Stability of Markovian processes. I. Criteria for discrete-time chains. *Adv. in Appl. Probab.*, 24(3):542–574.

Meyn, S. P. and Tweedie, R. L. (1993a). Generalized Resolvents and Harris Recurrence of Markovian processes. *Contemp. Math.*, 149:227–250.

Meyn, S. P. and Tweedie, R. L. (1993b). Stability of Markovian processes. II. Continuous-time processes and sampled chains. *Adv. in Appl. Probab.*, 25(3):487–517.

Meyn, S. P. and Tweedie, R. L. (1993c). Stability of Markovian processes. III. Foster-Lyapunov criteria for continuous-time processes. *Adv. in Appl. Probab.*, 25(3):518–548.

Meyn, S. P. and Tweedie, R. L. (2009). *Markov chains and Stochastic Stability, second edition.* Cambridge: Cambridge University Press.

Nguyen, D. H. and Yin, G.(2018a). Recurrence and ergodicity of switching diffusions with past-dependent switching having a countable state space. *Potential Anal.*, 48(4):405–435.

Nguyen, D. H. and Yin, G.(2018b). Recurrence for switching diffusion with past dependent switching and countable state space. *Math. Control Relat. Fields*, 8(3-4):879–897.

Nguyen, D. H. and Yin, G.(2018c). Stability of regime-switching diffusion systems with discrete states belonging to a countable set. *SIAM J. Control Optim.*, 56(5):3893–3917.

Nguyen, D. H., Yin, G., and Zhu, C. (2017). Certain properties related to well posedness of switching diffusions. *Stochastic Process. Appl.*, 127:3135–3158.

Norris, J. R. (1997). Markov Chains. *Cambridge University Press.*

Priola, E. and Wang, F.-Y. (2006). Gradient estimates for diffusion semigroups with singular coefficients. *J. Funct. Anal.*, 236(1):244–264.

Qiao, H. (2014). Exponential ergodicity for SDEs with jumps and non-Lipschitz coefficients. *J. Theoret. Probab.*, 27(1):137–152.

Shao, J. (2015a). Ergodicity of regime-switching diffusions in Wasserstein distances. *Stochastic Process. Appl.*, 125(2):739–758.

Shao, J. (2015b). Strong solutions and strong Feller properties for regime-switching diffusion processes in an infinite state space. *SIAM J. Control Optim.*, 53(4):2462–2479.

Shao, J. and Xi, F. (2014). Stability and recurrence of regime-switching diffusion processes. *SIAM J. Control Optim.*, 52(6):3496–3516.

Sharpe, M. (1988). *General Theory of Markov Processes.* Academic Press, New York.

Situ, R. (2005). *Theory of stochastic differential equations with jumps and applications.* Mathematical and Analytical Techniques with Applications to Engineering. Springer, New York.

Stettner, L. (1986). On the existence and uniqueness of invariant measure for continuous-time Markov processes. Technical Report LCDS 86-18, Brown University, Providence, RI, April.

Tuong, T. D., Nguyen, D. H., Dieu, N. T., and Tran, K. (2019). Extinction and permanence in a stochastic SIRS model in regime-switching with general incidence rate. *Nonlinear Anal. Hybrid Syst.*, 34:121–130.

Tweedie, R. L. (1994). Topological conditions enabling use of Harris methods in discrete and continuous time. *Acta. Appl. Math.*, 34(1-2):175–188.

Xi, F. (2004). Stability of a random diffusion with nonlinear drift. *Statist. Probab. Lett.*, 68(3):273–286.

Xi, F. (2009). Asymptotic properties of jump-diffusion processes with state-dependent switching. *Stochastic Process. Appl.*, 119(7):2198–2221.

Xi, F. and Zhu, C. (2017). On Feller and strong Feller properties and exponential ergodicity of regime-switching jump diffusion processes with countable regimes. *SIAM J. Control Optim.*, 55(3):1789–1818.

Xi, F. and Zhu, C. (2019). Jump type stochastic differential equations with non-Lipschitz coefficients: non-confluence, Feller and strong Feller properties, and exponential ergodicity. *J. Differential Equations*, 266(8):4668–4711.

Xi, F., Yin, G., and Zhu, C. (2019). Regime-switching jump diffusions with non-Lipschitz coefficients and countably many switching states: existence and uniqueness, Feller, and strong Feller properties. In *Modeling, stochastic control, optimization, and applications*, volume 164 of *IMA Vol. Math. Appl.*, pages 571–599. Springer, Cham.

Yin, G. and Zhu, C. (2010). *Hybrid Switching Diffusions: Properties and Applications*, volume 63 of *Stochastic Modelling and Applied Probability.* Springer, New York.

Zhang, Q. (2001). Stock trading: an optimal selling rule. *SIAM J. Control Optim.*, 40(1):64–87.

Zhu, C. and Yin, G. (2009). On strong Feller, recurrence, and weak stabilization of regime-switching diffusions. *SIAM J. Control Optim.*, 48(3):2003–2031.

Curriculum Vitae

Khwanchai Kunwai

born May 11th, 1989, Lampang, Thailand

EDUCATION

2016 - 2021
PhD in Mathematics
University of Wisconsin-Milwaukee, USA
Dissertation title: Regime-Switching Jump Diffusion Processes with Countable Regimes: Feller, Strong Feller, Irreducibility and Exponential Ergodicity

2012 - 2015
Master of Science in Mathematics(Honors)
Chiang Mai University, Thailand
Thesis title: Properties of Attractive Points in $CAT(0)$ Spaces

2008 - 2012
Bachelor of Science in Mathematics(Honors)
Chiang Mai University, Thailand
Thesis title: On the Complete Rotations of Cayley Graphs Defined on Left Groups and Right Groups

PROFESSIONAL EXPERIENCE

2020-2021
Teaching Assistance, Department of Mathematical Sciences
University of Wisconsin-Milwaukee, USA

SEMINARS

Regime-Switching Jump Diffusion Processes with Countable Regimes: Feller, Strong Feller, Irreducibility and Exponential Ergodicity, AMS Spring Central Virtual Sectional Meeting, April 2021

Foster-Lyapunov type criterion for non-explosive Markov processes, Graduate Student Colloquium, Department of Mathematical Sciences, UWM, November 2019

PUBLICATIONS

Kunwai, K. (2021). On Foster-Lyapunov criteria for exponential ergodicity of regime-switching jump diffusion processes with countable regimes. *Adv. in Appl. Probab*, in press.

Kunwai, K. and Zhu, C. (2020). On Feller and Strong Feller properties and irreducibility of regime-switching jump diffusion processes with countable regimes. *Nonlinear Anal. Hybrid Syst.*, 38: 100946.

HONORS

2021
Chancellor's Graduate Student Award, Department of Mathematical Sciences
University of Wisconsin-Milwaukee

2020
Chancellor's Graduate Student Award, Department of Mathematical Sciences
University of Wisconsin-Milwaukee

Morris and Miriam Marden Award in Mathematics, Department of Mathematical Sciences
University of Wisconsin-Milwaukee

2019
Research Excellence Award, Department of Mathematical Sciences
University of Wisconsin-Milwaukee

2018
Research Excellence Award, Department of Mathematical Sciences
University of Wisconsin-Milwaukee

2017
Research Excellence Award, Department of Mathematical Sciences
University of Wisconsin-Milwaukee