1-1-2014

On Switching Diffusions: The Feynman-Kac Formula And Near-Optimal Controls

Nicholas Baran
Wayne State University,

ON SWITCHING DIFFUSIONS: THE FEYNMAN-KAC FORMULA AND NEAR-OPTIMAL CONTROLS

by

NICHOLAS BARAN

DISSERTATION

Submitted to the Graduate School

of Wayne State University,

Detroit, Michigan

in partial fulfillment of the requirements

for the degree of

DOCTOR OF PHILOSOPHY

2014

MAJOR: MATHEMATICS

Approved by:

Advisor Date

OVERVIEW

This dissertation concerns switching diffusions or hybrid-switching diffusions in two different contexts. Accordingly, it is divided into two main chapters. The setup and formulation is slightly different in each context so we choose to present the necessary background for each context in their respective chapter as opposed to here in the introduction. In the first chapter we consider the so-called Feynman-Kac formula(s). These formulas provide stochastic representations for solutions to partial differential equations and are now standard in virtually any introductory text to stochastic differential equations; see for example [13] or [7]. Here we verify these formulas in the context of switching diffusions for boundary value problems, initial boundary value problems, and for the initial value problem. This work can also be found in [3]. In the second chapter we switch gears a bit and consider a problem in stochastic optimal control. Namely we show, under fairly broad conditions, the existence of a near-optimal control for a dynamical system driven by wideband noise in the presence of regime switching. The use of a wideband noise is actually motivated by modeling and applications as true white noise is often not encountered in reality. Furthermore, such systems can be accurately approximated using approximations to white noise and wideband noise turns out to be a suitable candidate for this purpose. To carry out our program we actually recast the setting using a relaxed control formulation, an idea that was introduced in [16]; the stochastic version was introduced in [5]. Recast in this setting, the computational difficulty is greatly reduced. Furthermore, using weak convergence methods we obtain controls that are nearly optimal for the original system. The outline and motivation for this program was inspired by the work in [10] and many of the ideas and related results presented here can also be found in [19].

TABLE OF CONTENTS

Chapter 1

The Feynman Kac Formula

1.1 Introduction

Let $(\Omega, \mathcal{F}, P)$ be a probability space, and $\{\mathcal{F}_t\}$ be a filtration on this space satisfying the usual condition (i.e., $\mathcal{F}_0$ contains all the null sets and the filtration $\{\mathcal{F}_t\}$ is right continuous). The probability space $(\Omega, \mathcal{F}, P)$ together with the filtration $\{\mathcal{F}_t\}$ is denoted by $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$. Suppose that $\alpha(\cdot)$ is a stochastic process with right-continuous sample paths (or a pure jump process), finite-state space $\mathcal{M} = \{1, \ldots, m_0\}$, and x-dependent generator $Q(x)$ so that for a suitable function $f(\cdot, \cdot)$,

$$Q(x)f(x,\cdot)(i) = \sum_{j\in\mathcal{M},\ j\neq i} q_{ij}(x)(f(x,j) - f(x,i)), \qquad \text{for each } i \in \mathcal{M}. \tag{1.1}$$

Assume throughout the paper that $Q(x)$ satisfies the q-property [22]. That is, $Q(x) = (q_{ij}(x))$ satisfies:

(i) $q_{ij}(x)$ is Borel measurable and uniformly bounded for all i, $j \in \mathcal{M}$ and $x \in \mathbb{R}^n$;

(ii) $q_{ij}(x) \geq 0$ for all $x \in \mathbb{R}^n$ and $j \neq i$; and

(iii) $q_{ii}(x) = -\sum_{j\neq i} q_{ij}(x)$ for all $x \in \mathbb{R}^n$ and $i \in \mathcal{M}$.

Let $w(\cdot)$ be an $\mathbb{R}^n$-valued standard Brownian motion defined on $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}, P)$, $b(\cdot,\cdot) : \mathbb{R}^n \times \mathcal{M} \to \mathbb{R}^n$, and $\sigma(\cdot,\cdot) : \mathbb{R}^n \times \mathcal{M} \to \mathbb{R}^n \times \mathbb{R}^n$ such that the two-component process $(X(\cdot), \alpha(\cdot))$, satisfies

$$\begin{aligned}&dX(t) = b(X(t), \alpha(t))dt + \sigma(X(t), \alpha(t))dw(t),\\&(X(0), \alpha(0)) = (x, i)\end{aligned} \tag{1.2}$$

and

$$P\{\alpha(t+\delta) = j|\alpha(t) = i,\ X(s),\ \alpha(s),\ s \le t\} = q_{ij}(X(t))\delta + o(\delta),\ i \ne j. \tag{1.3}$$

The process given by (2.2) and (1.3) is called a switching diffusion or a regime-switching diffusion.

Now, before carrying out our analysis, we state a theorem regarding existence and uniqueness of the solution of the aforementioned stochastic differential equation, which will be important in what follows.

Theorem 1.1.1. (Yin and Zhu [22]) *Let $x \in \mathbb{R}^n$, $\mathcal{M} = \{1, \dots, m_0\}$, and $Q(x) = (q_{ij}(x))$ be an $m_0 \times m_0$ matrix satisfying the q-property. Consider the two component process $Y(t) = (X(t), \alpha(t))$ given by* (2.2) *with initial data (x, i). Suppose that $Q(\cdot) : \mathbb{R}^n \to \mathbb{R}^{m_0 \times m_0}$ is bounded and continuous, and that the functions $b(\cdot, \cdot)$ and $\sigma(\cdot, \cdot)$ satisfy*

$$|b(x, i)| + |\sigma(x, i)| \le K(1 + |x|), \qquad i \in \mathcal{M}, \tag{1.4}$$

for some constant $K > 0$, and for each $N > 1$, there exists a positive constant M_N such that for all $i \in \mathcal{M}$ and all $x, y \in \mathbb{R}^n$ with $|x| \vee |y| \le M_N$,

$$|b(x, i) - b(y, i)| \vee |\sigma(x, i) - \sigma(y, i)| \le M_N|x - y|, \tag{1.5}$$

where $a \vee b = \max(a, b)$ for $a, b \in \mathbb{R}$. Then there exists a unique solution to (2.2) *in which the evolution of the discrete component is given by* (1.3).

1.2 Itô's Formula

Consider $(X(t), \alpha(t))$ given in (2.2) and let $a(x,i) = \sigma(x,i)\sigma'(x,i)$, where $\sigma'(x,i)$ denotes the transpose of $\sigma(x,i)$. Given any function $g(\cdot,i) \in C^2(\mathbb{R}^n)$ with $i \in \mathcal{M}$, define $\mathcal{L}$ by

$$\mathcal{L}g(x,i) := \frac{1}{2}\mathrm{tr}(a(x,i)D^2g(x,i)) + b'(x,i)Dg(x,i) + Q(x)g(x,\cdot)(i), \tag{1.6}$$

where $Dg(\cdot,i) = (\frac{\partial g}{\partial x_1}, \ldots, \frac{\partial g}{\partial x_n})$, $D^2g(\cdot,i)$ denotes the Hessian of $g(\cdot,i)$, and $Q(x)g(x,\cdot)(i)$ is given by (1). The choice for $\mathcal{L}$ will become clear momentarily.

It turns out that the evolution of the discrete component can be represented as a stochastic integral with respect to a Poisson random measure $\mathfrak{p}(dt,dz)$, whose intensity is $dt \times m(dz)$, where $m(\cdot)$ is the Lebesgue measure on $\mathbb{R}$. We have

$$d\alpha(t) = \int_{\mathbb{R}} h(X(t), \alpha(t-), z)\mathfrak{p}(dt,dz), \tag{1.7}$$

where h is an integer valued function; furthermore, this representation is equivalent to (3). For details, we refer the reader to [15] and [22].

We now state (generalized) Itô's formula. For each $i \in \mathcal{M}$ and $g(\cdot,i) \in C^2(\mathbb{R}^n)$, we have

$$g(X(t),\alpha(t)) - g(X(0),\alpha(0)) = \int_0^t \mathcal{L}g(X(s),\alpha(s))ds + M_1(t) + M_2(t) \tag{1.8}$$

where

$$M_1(t) = \int_0^t \langle Dg(X(s),\alpha(s)), \sigma(X(s),\alpha(s)\rangle dw(s),$$

$$M_2(t) = \int_0^t \int_{\mathbb{R}} [g(X(s), \alpha(0) + h(X(s),\alpha(s),z)) - g(X(s),\alpha(s))]\mu(ds,dz).$$

The compensated or centered Poisson measure $\mu(ds, dz) = \mathfrak{p}(ds, dz) - ds \times m(dz)$ is a martingale measure. For $t \geq 0$, and $g(\cdot, i) \in C_0^2$ (the collection of C^2 functions with compact support) for each $i \in \mathcal{M}$,

$$E^{x,i} g(X(t), \alpha(t)) - g(x, i) = E^{x,i} \int_0^t \mathcal{L} g(X(s), \alpha(s)) ds, \tag{1.9}$$

where $E^{x,i}$ denotes the expectation with initial data $(X(0), \alpha(0)) = (x, i)$. The above equation is known as Dynkin's formula. The condition $g \in C_0^2$ ensures that

$$g(X(t), \alpha(t)) - g(x, i) - \int_0^t \mathcal{L} g(X(s), \alpha(s)) ds \ \text{ is a martingale.}$$

Furthermore, one can show that $\mathcal{L}$ agrees with its classical interpretation, as the (infinitesimal) generator of the process $(X(t), \alpha(t))$ given by

$$\mathcal{L} g(x, i) = \lim_{t \downarrow 0} \frac{E^{x,i}[g(X(t), \alpha(t))] - g(x, i)}{t}. \tag{1.10}$$

To see this, pick t sufficiently small so that $\alpha(t)$ agrees with the initial data. Then it follows that

$$\begin{aligned} &\frac{1}{t} \int_0^t \mathcal{L} g(X(s), \alpha(s)) ds \\ &\quad = \frac{1}{t} \int_0^t \mathcal{L} g(X(s), i) ds \to \mathcal{L} g(x, i), \qquad t \to 0 \end{aligned}$$

by continuity. Hence by multiplying by t^{-1} then letting t tend to zero, one gets

$$\left| \frac{1}{t} E \int_0^t \mathcal{L} g(X(s), \alpha(s)) ds - \mathcal{L} g(x, i) \right| \to 0, \ \text{as } t \to 0,$$

and consequently (1.10).

Noting (1.9), when the deterministic time t is replaced by a stopping time τ satisfying $\tau < \infty$ w.p.1 (recalling that $g(\cdot, i) \in C_0^2$), then

$$E^{x,i} g(X(\tau), \alpha(\tau)) - g(x, i) = E^{x,i} \int_0^\tau \mathcal{L} g(X(s), \alpha(s)) ds. \tag{1.11}$$

Note that if τ is the first exit time of the process from a bounded domain satisfying $\tau < \infty$ w.p.1, then Dynkin's formula holds for any $g(\cdot, i) \in C^2$ and each $i \in \mathcal{M}$ without the compact support assumption. To proceed, we obtain the following system of Kolmogorov backward equations for switching diffusions; see also [21].

Theorem 1.2.1. (Kolmogorov Backward Equation) *Suppose* $g(\cdot, i) \in C_0^2(\mathbb{R}^n)$, *for* $i \in \mathcal{M}$, *and define*

$$u(x, t, i) = E^{x,i}[g(X(t), \alpha(t))]. \tag{1.12}$$

Then u *satisfies*

$$\begin{cases} \dfrac{\partial u}{\partial t} = \mathcal{L} u & \text{for } \ t > 0, \ x \in \mathbb{R}^n, \ i \in \mathcal{M} \\ u(x, 0, i) = g(x, i) & \text{for } x \in \mathbb{R}^n, \ i \in \mathcal{M} \end{cases} \tag{1.13}$$

A proof of the theorem can be found in [21, Theorem 5.2]; see also Theorem 5.1 in the aforementioned reference.

Remark 1. We illustrate the proof of the theorem using the idea as in [13, p. 140].

Fix $t > 0$. Then, using (1.10) and the Markov property, we have

$$\begin{aligned}
&\frac{E^{x,i}[u(X(r),t,\alpha(r))] - u(x,t,i)}{r} \\
&= \frac{E^{x,i}[E^{X(r),\alpha(r)}[g(X(t),\alpha(t))]] - E^{x,i}[g(X(t),\alpha(t))]}{r} \\
&= \frac{E^{x,i}[E^{x,i}[g(X(t+r),\alpha(t+r))|\mathcal{F}_r] - E^{x,i}[g(X(t),\alpha(t))]}{r} \\
&= \frac{E^{x,i}[g(X(t+r),\alpha(t+r))] - E^{x,i}[g(X(t),\alpha(t))]}{r} \\
&= \frac{u(x,t+r,i) - u(x,t,i)}{r} \to \frac{\partial u}{\partial t}(x,t,i) \quad \text{ as } \ r \downarrow 0.
\end{aligned}$$

Thus, by the definition of $\mathcal{L}$, (1.13) is satisfied.

We now state the Feynman-Kac Formula, which is a generalization of the Kolmogorov Backward equation.

Theorem 1.2.2. (The Feynman-Kac Formula) *Suppose $g(\cdot,i) \in C_0^2(\mathbb{R}^n)$, and let $c(\cdot,i) \in C(\mathbb{R}^n)$ be bounded; $i \in \mathcal{M}$. Define*

$$v(x,t,i) = E^{x,i}\left[\exp\left(-\int_0^t c(X(s),\alpha(s))ds\right)g(X(t),\alpha(t))\right]. \tag{1.14}$$

Then v satisfies

$$\begin{cases} \dfrac{\partial v}{\partial t} = \mathcal{L}v - cv & \text{for } \ t > 0, \ x \in \mathbb{R}^n, \ i \in \mathcal{M} \\ v(x,0,i) = g(x,i) & \text{for } x \in \mathbb{R}^n, \ i \in \mathcal{M} \end{cases} \tag{1.15}$$

Proof. To simplify the notation, let

$$Y(t) = g(X(t),\alpha(t)), \qquad Z(t) = \exp\left(-\int_0^t c(X(s),\alpha(s))ds\right).$$

Now, following the argument in Remark 1, we fix $t > 0$. We have

$$
\begin{aligned}
&\frac{E^{x,i}[v(X(r),t,\alpha(r))] - v(x,t,i)}{r} \\
&= \frac{E^{x,i}[E^{X(r),\alpha(r)}[Z(t)Y(t)]] - E^{x,i}[Z(t)Y(t)]}{r} \\
&= \frac{E^{x,i}[E^{x,i}[\exp\left(-\int_0^t c(X(s+r),\alpha(s+r))ds\right)Y(t+r)|\mathcal{F}_r]] - E^{x,i}[Z(t)Y(t)]}{r} \\
&= \frac{E^{x,i}[E^{x,i}[\exp\left(-\int_r^{t+r} c(X(s),\alpha(s))ds\right)Y(t+r)|\mathcal{F}_r]] - E^{x,i}[Z(t)Y(t)]}{r} \\
&= \frac{E^{x,i}[Z(t+r)\exp\left(\int_0^r c(X(s),\alpha(s))ds\right)Y(t+r)] - E^{x,i}[Z(t)Y(t)]}{r} \\
&= \frac{E^{x,i}[Z(t+r)Y(t+r)] - E^{x,i}[Z(t)Y(t)]}{r} \\
&\quad + \frac{E^{x,i}[Z(t+r)Y(t+r)\left\{\exp\left(\int_0^r c(X(s),\alpha(s))ds\right) - 1\right\}]}{r} \\
&= \frac{v(x,t+r,i) - v(x,t,i)}{r} \\
&\quad + \frac{E^{x,i}[Z(t+r)Y(t+r)\left\{\exp\left(\int_0^r c(X(s),\alpha(s))ds\right) - 1\right\}]}{r}.
\end{aligned}
$$

First, clearly,

$$\frac{v(x,t+r,i) - v(x,t,i)}{r} \to \frac{\partial v}{\partial t}(x,t,i), \qquad r \downarrow 0.$$

Furthermore, we claim

$$\frac{E^{x,i}[Z(t+r)Y(t+r)\left\{\exp\left(\int_0^r c(X(s),\alpha(s))ds\right) - 1\right\}]}{r} \to c(x,i)v(x,t,i).$$

To verify this claim, first note that

$$Z(t+r)Y(t+r) \to Z(t)Y(t), \qquad r \downarrow 0,$$

by continuity. Now, if we let

$$f(r) = \exp\left(\int_0^r c(X(s),\alpha(s))ds\right),$$

for r sufficiently small. Denote the first jump time of $\alpha(\cdot)$ by τ_1. With $\alpha(0) = i$, for any $t \in [0, \tau_1)$, $\alpha(t) = i$. It follows that

$$f(r) = \exp\left(\int_0^r c(X(s),i)ds\right),\ r \in [0,\tau_1).$$

Hence f is differentiable at the origin and

$$\frac{d}{dt}f(0) = f(0)c(X(0),i) = c(x,i).$$

This in turn yields that

$$\begin{aligned}
&Z(t+r)Y(t+r)\cdot\frac{1}{r}\left(\exp\left(\int_0^r c(X(s),\alpha(s))ds\right)-1\right)\\
&\quad = Z(t+r)Y(t+r)\left(\frac{f(r)-f(0)}{r}\right) \to Z(t)Y(t)c(x,i), \qquad r\downarrow 0.
\end{aligned}$$

Furthermore, the assumptions on the functions $c(\cdot,i)$ and $g(\cdot,i)$ ensure that this forms a bounded sequence, so we may apply the bounded convergence theorem to yield

$$\begin{aligned}
&\lim_{r\downarrow 0} E^{x,i}\left[Z(t+r)Y(t+r)\frac{1}{r}\left(\exp\left(\int_0^r c(X(s),\alpha(s))ds\right)-1\right)\right]\\
&\quad = E^{x,i}\left[\lim_{r\downarrow 0} Z(t+r)Y(t+r)\frac{1}{r}\left(\exp\left(\int_0^r c(X(s),\alpha(s))ds\right)-1\right)\right]\\
&\quad = E^{x,i}[Z(t)Y(t)c(x,i)] = c(x,i)E^{x,i}[Z(t)Y(t)] = c(x,i)v(x,t,i)
\end{aligned}$$

as claimed. This completes the proof. $\square$

So we have seen that the functions given by (1.12) and (1.14) necessarily satisfy certain initial value problems. The remainder of this section will be dedicated to giving stochastic representations for solutions to certain partial differential equations (PDEs) related to the operator $\mathcal{L}$.

1.3 Dirichlet Problem

Let $O \subset \mathbb{R}^n$, be a bounded open set and consider the following Dirichlet problem:

$$\begin{cases} \mathcal{L}u(x,i) + c(x,i)u(x,i) = \psi(x,i) & \text{in } O \times \mathcal{M} \\ u(x,i) = \varphi(x,i) & \text{on } \partial O \times \mathcal{M}, \end{cases} \tag{1.16}$$

where ∂O denotes the boundary of O. To proceed, we impose the following conditions:

(A1) Assume the following conditions hold

1. $\partial O \in C^2$
2. for some $1 \le j \le r$, and all $i \in \mathcal{M}$, $\min_{x \in \bar{O}} a_{jj}(x,i) > 0$
3. $a(\cdot,i)$ and $b(\cdot,i)$ are uniformly Lipschitz continuous in $\bar{O}$ for each $i \in \mathcal{M}$
4. $c(x,i) \le 0$ and $c(\cdot,i)$ is uniformly Hölder continuous in $\bar{O}$ for each $i \in \mathcal{M}$
5. $\psi(\cdot,i)$ is uniformly continuous in $\bar{O}$ and $\varphi(\cdot,i)$ is continuous on ∂O, both for each $i \in \mathcal{M}$

It follows that under (A1), the system of boundary value problems has a unique solution; see [2] or [12]. Our goal is to derive a stochastic representation for this problem, similar to the Feynman-Kac formula. In order to achieve this, we need the following lemma.

Lemma 1.3.1. *Suppose that* $\tau = \inf\{t \geq 0 : X^x(t) \notin O\}$. *That is,* τ *is the first exit time from the open set* O *of the switching diffusion given in* (2.2) *and* (1.3). *Then* $\tau < \infty$ *w.p.*1.

Proof. We use the idea as in [2]. Consider a function $V : \mathbb{R}^n \times \mathcal{M} \to \mathbb{R}$ defined by

$$V(x,i) = -A\exp(\lambda x_1), \qquad A, \lambda > 0, \ i \in \mathcal{M}.$$

Clearly $V(\cdot, i) \in C^\infty(O)$ and since V is independent of $i \in \mathcal{M}$,

$$Q(x)V(x,\cdot)(i) = \sum_{i \neq j} q_{ij}(x)(V(x,j) - V(x,i)) = 0,$$

and thus

$$\mathcal{L}V(x,i) = -A\exp(\lambda x_1)\left[\frac{1}{2}a_{11}\lambda^2 + b_1\lambda\right].$$

Note that as long as $\lambda > \dfrac{-b_1}{2a_{11}}$, it follows that $\mathcal{L}V(x,i) < 0$. Hence, by choosing λ and $A = A(\lambda)$ sufficiently large, we can make $\mathcal{L}V(x,i) \leq -1$ for each $i \in \mathcal{M}$. As the function $V(\cdot, i)$ and its derivatives w.r.t. x are bounded on $\bar{O}$, we may apply Dynkin's formula to yield:

$$\begin{aligned} E^{x,i}V(X(t\wedge\tau), \alpha(t\wedge\tau)) - V(x,i) &= E^{x,i}\int_0^{t\wedge\tau} \mathcal{L}V(X(s),\alpha(s))ds \\ &\leq -E^{x,i}(t\wedge\tau), \end{aligned}$$

where $E^{x,i}$ denotes the expectation taken with $(X(0),\alpha(0)) = (x,i)$. This yields that

$$E^{x,i}(t\wedge\tau) \leq V(x,i) - E^{x,i}V(X(t\wedge\tau),\alpha(t\wedge\tau)) \leq 2\max_{x\in\bar{O}, i\in\mathcal{M}} |V(x,i)| < \infty.$$

Taking the limit as $t \to \infty$, and using the monotone convergence theorem yields $E^{x,i}\tau < \infty$, which in turn leads to $\tau < \infty$ w.p.1. $\square$

Similar to the Feynman-Kac formula, the following result is true.

Theorem 1.3.1. *Suppose that* (A1) *holds. Then with* τ *as in the previous Lemma, the solution of the system of boundary value problems* (1.16) *is given by*

$$\begin{aligned} u(x,i) = E^{x,i}&\left[\varphi(X(\tau),\alpha(\tau))\exp\left(\int_0^\tau c(X(s),\alpha(s))ds\right)\right] \\ &- E^{x,i}\left[\int_0^\tau \psi(X(t),\alpha(t))\exp\left(\int_0^t c(X(s),\alpha(s))ds\right)dt\right]. \end{aligned} \tag{1.17}$$

Proof. We apply Itô's formula to the switching process

$$\tilde{u}(X(t),t,\alpha(t)) := u(X(t),\alpha(t))\exp\left(\int_0^t c(X(s),\alpha(s))ds\right).$$

To simplify notation we let

$$Z(t) = \exp\left(\int_0^t c(X(s),\alpha(s))ds\right).$$

We have

$$\begin{aligned} &E^{x,i}u(X(t\wedge\tau),\alpha(t\wedge\tau))Z(t\wedge\tau) - u(x,i) \\ &\quad = E^{x,i}\int_0^{t\wedge\tau}\left(\frac{\partial}{\partial s}+\mathcal{L}\right)\big\{u(X(s),\alpha(s))Z(s)\big\}ds \\ &\quad = E^{x,i}\int_0^{t\wedge\tau} Z(s)\big\{u(X(s),\alpha(s))c(X(s),\alpha(s)) + \mathcal{L}u(X(s),\alpha(s))\big\}ds \\ &\quad = E^{x,i}\int_0^{t\wedge\tau} Z(s)\psi(X(s),\alpha(s))ds. \end{aligned}$$

Taking the limit as $t \to \infty$ and noting the boundary conditions, (1.17) follows. $\square$

1.4 Initial Boundary Value Problem

Consider next the initial boundary value problem given by

$$\begin{cases} [\mathcal{L} + \frac{\partial}{\partial t}]u(x,t,i) + c(x,t,i)u(x,t,i) = \psi(x,t,i) & \text{in } O \times [0,T) \times \mathcal{M} \\ u(x,T,i) = \varphi(x,i) & \text{in } O \times \mathcal{M} \\ u(x,t,i) = \phi(x,t,i) & \text{on } \partial O \times [0,T] \times \mathcal{M} \end{cases} \tag{1.18}$$

where O is the same as before and

$$\mathcal{L}f(x,t,i) = \frac{1}{2}\mathrm{tr}(a(x,t,i)D^2 f(x,t,i)) + b'(x,t,i)Df(x,t,i) + Q(x)f(x,t,\cdot)(i) \tag{1.19}$$

We assume the following conditions:

(A2) Assume the following conditions hold.

1. $\langle a(x,t,i)y, y\rangle \geq \kappa |y|^2$, for each $i \in \mathcal{M}$ and for $y \in \mathbb{R}^n$, $(\kappa > 0)$,
2. $a_{lk}(\cdot,\cdot,i)$, $b_l(\cdot,\cdot,i)$ are uniformly Lipschitz continuous in $\bar{O} \times [0,T]$, for each $i \in \mathcal{M}$,
3. $c(\cdot,\cdot,i)$ and $\psi(\cdot,\cdot,i)$ are uniformly Hölder continuous in $\bar{O} \times [0,T]$, for each $i \in \mathcal{M}$,
4. $\varphi(\cdot,i)$ is continuous on $\bar{O}$, $\phi(\cdot,\cdot,i)$ is continuous on $\partial O \times [0,T]$, for each $i \in \mathcal{M}$, where ∂O denotes the boundary of O,
5. $\varphi(x,i) = \phi(x,T,i)$, for $x \in \partial O$.

Under (A2) it follows that the system of initial-boundary value problems has a unique solution; see [2] or [12]. In order to get a stochastic representation for the solution, we also require the drift and diffusion coefficients of u to be Lipschitz continuous in the time variable; namely we require

$$|b(x,t,i)-b(x,s,i)| \vee |\sigma(x,t,i)-\sigma(x,s,i)| \le K(|t-s|), \quad i \in \mathcal{M},$$

in addition to (2.4) and (1.5).

Now for $(x,t,i) \in O \times [0,T) \times \mathcal{M}$, consider the switching SDE given by

$$dX(s) = b(X(s),s,\alpha(s))ds + \sigma(X(s),s,\alpha(s))dw(s), \qquad s \in [t,T], \tag{1.20}$$

with initial data $(X(t),\alpha(t)) = (x,i)$. If we let $\sigma(x,t,i)$ be the square root of $a(x,t,i)$, then the following is true.

Theorem 1.4.1. *Suppose that* (A2) *holds. Then the solution of the system of initial value problems in* (1.18) *is given by*

$$\begin{aligned} u(x,t,i) &= E^{x,i}\left[I_{\{\tau<T\}}\phi(X(\tau),\tau,\alpha(\tau))\exp\left(\int_t^\tau c(X(r),r,\alpha(r))dr\right)\right] \\ &\quad + E^{x,i}\left[I_{\{\tau=T\}}\varphi(X(T),\alpha(T))\exp\left(\int_t^T c(X(r),r,\alpha(r))dr\right)\right] \\ &\quad - E^{x,i}\left[\int_t^{\tau\wedge T}\psi(X(s),s,\alpha(s))\exp\left(\int_t^s c(X(r),r,\alpha(r))dr\right)ds\right]. \end{aligned} \tag{1.21}$$

Proof. Proceeding similarly to the previous theorem, we apply Itô's formula to the process

$$u(X(s),s,\alpha(s))\exp\left(\int_t^s c(X(r),r,\alpha(r))dr\right), \qquad s \in [t,T].$$

To simplify notation we let

$$Z_t(s) = \exp\left(\int_t^s c(X(r),r,\alpha(r))dr\right).$$

We have

$$
\begin{aligned}
&E^{x,i}u(X(\tau\wedge T),\tau\wedge T,\alpha(\tau\wedge T))Z_t(\tau\wedge T)-u(x,t,i)\\
&\quad= E^{x,i}\int_t^{\tau\wedge T}\left(\frac{\partial}{\partial s}+\mathcal{L}\right)\big\{u(X(s),s,\alpha(s))Z_t(s)\big\}ds\\
&\quad= E^{x,i}\int_t^{\tau\wedge T} Z_t(s)\big\{u(X(s),s,\alpha(s))c(X(s),s,\alpha(s))+\mathcal{L}u(X(s),s,\alpha(s))\big\}ds\\
&\quad= E^{x,i}\int_t^{\tau\wedge T} Z_t(s)\psi(X(s),s,\alpha(s))ds.
\end{aligned}
$$

If we note that

$$
\begin{aligned}
u(X(\tau\wedge T),\tau\wedge T,\alpha(\tau\wedge T))Z_t(\tau\wedge T) &= \begin{cases} u(X(\tau),\tau,\alpha(\tau))Z_t(\tau), & \tau<T\\[1ex] u(X(T),T,\alpha(T))Z_t(T), & \tau=T\end{cases}\\
&= \begin{cases} \phi(X(\tau),\tau,\alpha(\tau))Z_t(\tau), & \tau<T\\[1ex] \varphi(X(T),\alpha(T))Z_t(T), & \tau=T,\end{cases}
\end{aligned}
$$

then by replacing the correct value for

$$
u(X(\tau\wedge T),\tau\wedge T,\alpha(\tau\wedge T))Z_t(\tau\wedge T)
$$

in the above derivation, one gets (1.21). $\square$

1.4.1 Cauchy Problem

If we let $O = \mathbb{R}^n$ in the initial value problem (17) of the previous section, we get the Cauchy problem:

$$\begin{cases} [\mathcal{L} + \frac{\partial}{\partial t}]u(x,t,i) + c(x,t,i)u(x,t,i) = \psi(x,t,i) & \text{in } \mathbb{R}^n \times [0,T) \times \mathcal{M} \\ u(x,T,i) = \varphi(x,i) & \text{in } \mathbb{R}^n \times \mathcal{M} \end{cases} \tag{1.22}$$

To proceed we impose the following condition:

(A3) Assume the following conditions hold.

1. The functions $a_{lk}(\cdot,\cdot,i)$, $b_l(\cdot,\cdot,i)$ are bounded in $\mathbb{R}^n \times [0,T]$ and uniformly Lipschitz continuous in (x,t,i) in compact subsets of $\mathbb{R}^n \times [0,T] \times \mathcal{M}$, for each $i \in \mathcal{M}$.

2. The functions $a_{lk}(\cdot,\cdot,i)$ are Hölder continuous in x, uniformly with respect to (x,t,i) in $\mathbb{R}^n \times [0,T] \times \mathcal{M}$, for each $i \in \mathcal{M}$.

3. The function $c(\cdot,\cdot,i)$ is bounded in $\mathbb{R}^n \times [0,T]$ and uniformly Hölder continuous in (x,t,i) in compact subsets of $\mathbb{R}^n \times [0,T] \times \mathcal{M}$, for each $i \in \mathcal{M}$.

4. The function $\psi(\cdot,\cdot,i)$ is continuous in $\mathbb{R}^n \times [0,T]$, for each $i \in \mathcal{M}$, Hölder continuous in x with respect to $(x,t,i) \in \mathbb{R}^n \times [0,T] \times \mathcal{M}$, and

$$|\psi(x,t,i)| \leq K(1 + |x|^p), \qquad \text{in } \mathbb{R}^n \times [0,T] \times \mathcal{M}.$$

5. The function $\varphi(\cdot,i)$ is continuous in $\mathbb{R}^n$, for each $i \in \mathcal{M}$, and $|\varphi(x,i)| \leq K(1 + |x|^p)$, where K and p are positive constants.

Under (A3) it follows that the Cauchy problem has a unique solution; see [2] or [12]. Moreover, the following is true.

Theorem 1.4.2. *Suppose that* (A3) *holds. Then the solution of the Cauchy problem in* (1.22) *is given by*

$$\begin{aligned} u(x,t,i) = E^{x,i}&\left[\varphi(X(T),\alpha(T))\exp\left(\int_t^T c(X(s),s,\alpha(s))ds\right)\right] \\ &- E^{x,i}\left[\int_t^T \psi(X(s),s,\alpha(s))\exp\left(\int_t^s c(X(r),r,\alpha(r))dr\right)ds\right]. \end{aligned} \tag{1.23}$$

Proof. As before, by Itô's formula, one has

$$\begin{aligned} &E^{x,i}u(X(T),T,\alpha(T))Z_t(T) - u(x,t,i) \\ &\quad = E^{x,i}\int_t^T \left(\frac{\partial}{\partial s} + \mathcal{L}\right)\left\{u(X(s),s,\alpha(s))Z_t(s)\right\}ds. \end{aligned}$$

Now, proceeding as in the proof of the initial-boundary value problem, we get (1.23).

□

Remark 2. Note by taking $c = \psi = 0$, we see that the Kolmogorov Backward Equation is a special case of the Cauchy problem by replacing u by:

$$\tilde{u}(x,t,i) := u(x,T-t,i).$$

So we have shown that even in the presence of regime switching, the generalizations and extensions of the Feynman-Kac formula remain valid. We close this chapter with a few examples.

1.4.2 Examples

Example 1. Let $O \subset \mathbb{R}^n$ be an open set and consider the following weakly coupled system:

$$\begin{cases} \triangle u(x,1) + q_{11}(x)u(x,1) + q_{12}(x)u(x,2) = \psi(x,1) & \text{in } O \\ \triangle u(x,2) + q_{21}(x)u(x,1) + q_{22}(x)u(x,2) = \psi(x,2) & \text{in } O \\ u(x,1) = u(x,2) = 0 & \text{on } \partial O. \end{cases} \tag{1.24}$$

Where $Q(x) = \begin{pmatrix} q_{11}(x) & q_{12}(x) \\ q_{21}(x) & q_{22}(x) \end{pmatrix}$ satisfies the q-property. Such systems are studied in [17]. It follows that this Dirichlet problem has the unique solution

$$u(x,i) = -E^{x,i}\left[\int_0^\tau \psi(x + B(t), \alpha(t))dt\right],$$

where $B(t)$ is a standard, n-dimensional Brownian motion and $\alpha(t)$ is a two-state, discrete process with generator $Q(x)$.

Example 2. Let

$$L_i = \frac{1}{2}\text{tr}(a(x,i)D^2 g(x,i)) + b'(x,i)Dg(x,i); \quad i = 1, 2,$$

and consider the following stationary system; found in [6].

$$\begin{cases} L_1 u(x,1) + q_{11}(x)u(x,1) + q_{12}(x)u(x,2) = 0 & \text{in } O \\ L_2 u(x,2) + q_{21}(x)u(x,1) + q_{22}(x)u(x,2) = 0 & \text{in } O \\ u(x,i) = \varphi(x,i) & \text{on } \partial O. \end{cases}$$

It follows that the solution of the above problem has the form:

$$u(x,i) = E^{x,i}\varphi(X(\tau),\alpha(\tau))\exp\left\{\int_0^\tau \tilde{q}(X(s),\alpha(s))ds\right\},$$

where $\tilde{q}(x,i) = q_{ii}(x) + q_{ij}(x)$ and $\alpha(t)$ is a two-state process satisfying:

$$P\{\alpha(t+\delta) = j|\alpha(t) = i,\ X(s),\ \alpha(s),\ s \le t\} = q_{ij}(X(t))\delta + o(\delta).$$

Hence if the generator $Q(x) = \begin{pmatrix} q_{11}(x) & q_{12}(x) \\ q_{21}(x) & q_{22}(x) \end{pmatrix}$ satisfies the q-property, then it follows that $\tilde{q}(x,i) = 0$ for all x, so the solution reduces to the form:

$$u(x,i) = E^{x,i}\varphi(X(\tau),\alpha(\tau)),$$

which agrees with the solution to the Dirichlet problem given by:

$$\begin{cases} \mathcal{L}u(x,i) = 0 & \text{in } O \times \{1,2\} \\ u(x,i) = \varphi(x,i) & \text{on } \partial O \times \{1,2\}. \end{cases}$$

Chapter 2

Near-Optimal Controls of Systems with Regime Switching

2.1 Introduction

Let $\mathcal{M} = \{1, \ldots, m_0\}$ be the state space of a discrete event process $\alpha^\epsilon(\cdot)$, that is a continuous-time Markov chain with $\epsilon > 0$. Let $z(\cdot)$ be an $\mathbb{R}^n$-valued stationary process that is independent of $\alpha^\epsilon(\cdot)$. Let $\mathcal{U}$ be our control space and suppose that $\mathcal{U}$ is a compact subset of $\mathbb{R}^{n_0}$. Let $b(\cdot,\cdot,\cdot) : \mathbb{R}^n \times \mathcal{M} \times \mathcal{U} \to \mathbb{R}^n$ and let $\sigma(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$. We consider the following controlled diffusion process with switching

$$
\begin{aligned}
&X^\epsilon(t) = x + \int_0^t b(X^\epsilon(s), \alpha^\epsilon(s), u^\epsilon(s))ds + \frac{1}{\sqrt{\epsilon}} \int_0^t \sigma(X^\epsilon(s), z^\epsilon(s))ds, \\
&\alpha^\epsilon(0) = i,
\end{aligned}
\tag{2.1}
$$

where $u^\epsilon(\cdot)$ takes values in $\mathcal{U}$. Our goal is to find the optimal control, $u^\epsilon(\cdot)$, so that the cost function (for some $\tilde{T} > 0$),

$$
J^\epsilon(u^\epsilon) = J^\epsilon(x, i, u^\epsilon) = E^{x,i} \int_0^{\tilde{T}} C(X^\epsilon(s), \alpha^\epsilon(s), u^\epsilon(s))ds \tag{2.2}
$$

is minimized. Here $E^{x,i}$ denotes the expectation with respect to the probability law of $(X^\epsilon(t), \alpha^\epsilon(t))$ with initial data $(X^\epsilon(0), \alpha^\epsilon(0)) = (x, i)$.

2.2 Relaxed Controls

We now introduce the relaxed control formulation. Denote the σ-algebra of Borel subsets of $\mathcal{U}$ by $\mathcal{B}(\mathcal{U})$. Let $M(\infty)$ denote the set of measures, $m(\cdot)$, defined on the Borel sets of $\mathcal{U} \times [0, \infty)$ that satisfy $m(\mathcal{U} \times [0, t]) = t$, for all $t \geq 0$. We say that a random $M(\infty)$-valued measure $m(\cdot)$ is an admissible relaxed control if for each $B \in \mathcal{B}(\mathcal{U})$, the function defined by $m(B, t) := m(B \times [0, t])$ is $\mathcal{F}_t$ adapted. Equivalently, one could say that $m(\cdot)$ is a relaxed control if $\int_0^t f(s, c)m(ds \times dc)$ is progressively measurable with respect to $\{\mathcal{F}_t\}$ for every bounded and continuous function $f(\cdot, \cdot)$.

It can be shown that if $m(\cdot)$ is an admissible relaxed control, then there is a measure-valued function $m_t(\cdot)$ so that $m_t(c)dt = m(dt \times dc)$ and, for a smooth function $f(\cdot)$, we have

$$\int f(s, c)m(ds \times dc) = \int ds \int_{\mathcal{U}} f(s, c)m_s(dc)$$

To proceed, we topologize $M(\infty)$ as follows. Let $\{f_{k_i}(\cdot) : i < \infty\}$ be a countable dense set of continuous functions on $\mathcal{U} \times [0, k]$, for each k. Let

$$\langle m, f \rangle = \int f(s, c)m(ds \times dc),$$

$$d(m_1, m_2) = \sum_{k=1}^{\infty} \frac{1}{2^k} d_k(m_1, m_2),$$

where

$$d_k(m_1, m_2) = \sum_{i=1}^{\infty} \frac{1}{2^i} \frac{|(m_1 - m_2, f_{k_i})|}{1 + |(m_1 - m_2, f_{k_i})|}$$

If a sequence of measures $\{m_k(\cdot)\}$ in $M(\infty)$ converges weakly to a measure $m(\cdot)$, we will denote this by $m_k(\cdot) \Rightarrow m(\cdot)$. This setup is fairly standard and can be found in several other sources, for instance [20].

We say that an ordinary admissible control $u(\cdot)$ is a feedback control if there is a $\mathcal{U}$-valued Borel measurable function $u_0(\cdot)$ so that $u(t) = u_0(x(t))$ for almost all ω and t. For each x, let $m_f(x, \cdot)$ be a probability measure on $(\mathcal{U}, \mathcal{B}(\mathcal{U}))$ and suppose that for each $B \in \mathcal{B}(\mathcal{U})$, $m_f(\cdot, B)$ is Borel measurable as a function of x. If for almost all ω and t we have that $m_t(\cdot) = m_f(x(t), \cdot)$, then $m(\cdot)$ is said to be a relaxed feedback control.

We now rewrite the system using the relaxed control formulation. We have

$$\begin{aligned} X^\epsilon(t) &= x_0 + \int_0^t \int_{\mathcal{U}} b(X^\epsilon(s), \alpha^\epsilon(s), c) m_s^\epsilon(dc) ds + \frac{1}{\sqrt{\epsilon}} \int_0^t \sigma(X^\epsilon(s), z^\epsilon(s)) ds \\ \alpha^\epsilon(0) &= i_0 \in \mathcal{M}, \end{aligned} \tag{2.3}$$

where $m^\epsilon(\cdot)$ is the relaxed control. Our goal is to choose the optimal control $m^\epsilon(\cdot)$ so that the cost function

$$J^\epsilon(m^\epsilon(\cdot)) = J^\epsilon(x_0, i_0, u^\epsilon(\cdot)) = E^{x_0, i_0} \int_0^{\tilde{T}} \int_{\mathcal{U}} C(X^\epsilon(s), \alpha^\epsilon(s), c) m_s^\epsilon(dc) ds \tag{2.4}$$

is minimized.

Assumptions

To proceed we make the following assumptions.

(A1) $\alpha^\epsilon(\cdot)$ is a continuous-time Markov chain with state space $\mathcal{M}$ and generator $Q^\epsilon(t) = (q_{ij}^\epsilon(t))$ given by

$$Q^\epsilon(t) = \frac{\tilde{Q}(t)}{\epsilon} + \hat{Q}(t),$$

where both $\tilde{Q}(t)$ and $\hat{Q}(t)$ are bounded and Borel measurable generators of continuous Markov chains such that $\tilde{Q}(t) = \text{diag}(\tilde{Q}^1(t), \ldots, \tilde{Q}^k(t))$ where each $\tilde{Q}^j(t)$ is weakly irreducible with quasi-stationary distribution
$v^j(t) = (v_1^j(t), \ldots, v_{m_j}^j(t)) \in \mathbb{R}^{1\times m_j}$.

(A2) The functions $b(\cdot)$ and $\sigma(\cdot)$ satisfy: For each $\alpha \in \mathcal{M}$, $b(\cdot, \alpha, c)$ and $\sigma(\cdot, z)$ are defined and Borel measurable on $[0, T] \times \mathbb{R}$ such that, for each x and $y \in \mathbb{R}^n$, the following Lipschitz condition holds.

$$|b(y, \alpha, c) - b(x, \alpha, c)| \leq L|x - y|,$$

$$|\sigma(y, z) - \sigma(x, z)| \leq L|x - y|,$$

for each α, c, and z.

(A3) The process $z^\epsilon(t) = z(t/\epsilon)$, where $z(\cdot)$ is a stationary process, independent of $\alpha(\cdot)$, with mean $Ez(s) = 0$. It is a strong mixing process with mixing measure $\phi(\cdot)$ so that the process is bounded, right continuous, and satisfies $\int_0^\infty \phi^{1/2}(s)ds < \infty$.

(A4) There is a positive integer p_0 so that for each c and $\alpha \in \mathcal{M}$ we have

$$|C(x, \alpha, c)| \leq \kappa(1 + |x|^{p_0}),$$

where κ is an arbitrary positive constant.

We also assume that $\mathcal{M}$ is nearly completely decomposable, an idea introduced in [1]. This means that our state space $\mathcal{M}$ can be decomposed into subspaces

$$\mathcal{M} = \mathcal{M}_1 \cup \mathcal{M}_2 \cup \cdots \cup \mathcal{M}_l$$

with the elements of $\mathcal{M}_i$ labeled as

$$\mathcal{M}_i = \{s_{i1}, \dots, s_{im_i}), \quad i = 1, \dots, l,$$

where the probability of leaving a particular subspace, $\mathcal{M}_j$, is small once it has been entered. However, as the name suggests, the subspaces are not completely isolated but communicate via the slow part of the generator $\hat{Q}(t)$. In addition, we may reduce much of the computational complexity by treating all of the states in each subspace as a single state. Precisely, we define a new process given by

$$\bar{\alpha}^\epsilon(t) = j \;\text{ if }\; \alpha^\epsilon(t) \in \mathcal{M}_j. \tag{2.5}$$

Note that $\bar{\alpha}^\epsilon(\cdot)$ takes values in $\bar{\mathcal{M}} = \{1, \dots, l\}$.

Remark. Let us make the following observations.

- Note that $\bar{\alpha}^\epsilon(\cdot)$ is not a Markov process. However, as it was proved in Chapter 5 of [20], $\bar{\alpha}^\epsilon(\cdot)$ converges weakly to a Markov chain $\bar{\alpha}(\cdot)$ such that the generator of $\bar{\alpha}(\cdot)$ is given by

$$\bar{Q} = \text{diag}(\nu^1(t), \dots, \nu^l(t))\hat{Q}(t)\text{diag}(\mathbb{1}_{m_1}, \dots, \mathbb{1}_{m_l}). \tag{2.6}$$

- One of the difficulties that we are facing is that system (2.1) in general is non-Markovian. This is because that $z^\epsilon(t)$ is generally non-Markovian. Nevertheless, $z^\epsilon(\cdot)$ is a so-called wideband noise process that approximates the white noise. Recall that a wideband noise process is one whose band width goes to ∞ and hence an approximation of the white noise. We will use the methods of averaging to obtain a limit the system. In fact, we examine the pair of processes

$(X^\epsilon(\cdot), \bar{\alpha}^\epsilon(\cdot))$ and show the limit is a Markovian. This limit system will lead to a controlled martingale problem.

- Using the near-optimal control of the limit system, we construct controls for the original systems. Then we show that the control so constructed is nearly optimal.

2.3 Limit Systems

Let $D([0,\infty) : S)$ denote the set of cadlag functions from $[0,\infty)$ to S. A cadlag function is one that is right-continuous and has left hand limits. A detailed presentation of their weak convergence properties can be found in [4]. We now present the main result of this section.

Theorem 2.3.1. *Suppose that assumptions (A1)-(A3) hold. Then* $\{X^\epsilon(\hat{m}^\epsilon(\cdot),\cdot), \bar{\alpha}^\epsilon(\cdot), \hat{m}(\cdot)\}$ *is tight in* $D([0,\infty) : \mathbb{R}^n \times \mathcal{M}) \times M(\infty)$ *where* $D([0,\infty) : \mathbb{R}^n \times \mathcal{M})$ *denotes the set of cad-lag functions from* $[0,\infty)$ *to* $\mathbb{R}^n \times \mathcal{M}$. *Suppose that* $(X^\epsilon(\hat{m}^\epsilon(\cdot),\cdot), \bar{\alpha}^\epsilon(\cdot), \hat{m}^\epsilon(\cdot))$ *converges weakly to* $(X(\hat{m}(\cdot),\cdot), \bar{\alpha}(\cdot), \hat{m}(\cdot))$. *Then there exists an* $\mathbb{R}^n$*-valued standard Brownian motion* $w(\cdot)$ *such that* $w(\cdot)$ *and* $\bar{\alpha}(\cdot)$ *are mutually independent,* $\hat{m}(\cdot)$ *is admissible with respect to* $(w(\cdot), \bar{\alpha}(\cdot))$, *and*

$$X(t) = x_0 + \int_0^t \int_{\mathcal{U}} \bar{b}(X(s), \bar{\alpha}(s), c) m_s(dc) + \int_0^t \bar{\sigma}(X(s)) dw, \tag{2.7}$$

where $\bar{\sigma}(x)$ *is the square root of* $R(x)$ *with* $R(x)$ *defined by*

$$\begin{aligned} R(x) = E\sigma(x, z(0))\sigma'(x, z(0)) + \int_0^\infty E\sigma(x, z(t))\sigma'(x, z(0)) dt \\ + \int_0^\infty E\sigma(x, z(0))\sigma'(x, z(t)) dt, \end{aligned} \tag{2.8}$$

and

$$\bar{b}(x,i,u) = \sum_{j=1}^{m_i} v_j^i(t) b(x, s_{ij}, u) + \int_0^\infty E[D\sigma(x, z(s))\sigma(x, z(0))]ds. \qquad (2.9)$$

Remark. Note that corresponding to (2.7) or corresponding to the limit switching diffusion, there is a generator given by

$$\mathcal{G}f(x,i) = \mathcal{L}f(x,i) + \bar{Q}(t)f(x,\cdot)(i), \quad \text{for} \quad i \in \bar{\mathcal{M}},$$

where $\mathcal{L}$ and Q are given by

$$\mathcal{L}f(x,i) = \frac{1}{2}\mathrm{tr}[\sigma(x)\sigma'(x)D^2 f(x,i)] + \int_{\mathcal{U}} \bar{b}'(x,i,c)Df(x,i)m_t(dc), , i \in \bar{\mathcal{M}},$$

$$\bar{Q}(t)f(x,\cdot)(i) = \sum_{j \in \bar{\mathcal{M}}} q_{ij}(t)f(x,j) = \sum_{j \neq i, j \in \bar{\mathcal{M}}} q_{ij}(t)[f(x,j) - f(x,i)], \ i \in \bar{\mathcal{M}},$$

and where $\bar{Q}(t) = (\bar{q}_{ij}(t))$ given by (2.6).

To prove Theorem 2.3.1, we carry out a series of tasks. To begin, define

$$\Phi_f(t) = f(X(t), \bar{\alpha}(t)) - f(X(0)), \bar{\alpha}(0)) - \int_0^t \mathcal{G}f(X(s), \bar{\alpha}(s))ds, \qquad (2.10)$$

for and $f(\cdot, i)$ that is twice continuously differentiable with compact support for each $i \in \bar{\mathcal{M}}$. Following the classical approach, we hope to show that $\Phi_f(\cdot)$ is a martingale for each $i \in \bar{\mathcal{M}}$, and for each $f(\cdot, i) \in C_0^2$, where C_0^2 denotes the set of C^2 functions with compact support. This will in turn show that $(X(\cdot), \bar{\alpha}(\cdot))$ is a solution to the martingale problem with operator $\mathcal{G}$. Since we do not assume the process is bounded, we first approach the problem from the angle of truncation. Then, once we have verified the claims for the truncated process, we show that these claims also remain valid when the truncated process is replaced with the untruncated process. To proceed, let

$K > 0$ be an arbitrary constant and define a truncation function

$$h_K(x) = \begin{cases} 1, & \text{if } |x| \le K, \\ 0, & \text{if } |x| \ge K+1, \end{cases}$$

such that $h_K(x)$ is sufficiently smooth for $K \le |x| \le K+1$. We define the truncated version of $X^\epsilon(\cdot)$ by

$$\begin{aligned} X_K^\epsilon(t) = x_0 &+ \int_0^t \int_{\mathcal{U}} b_K(X_K^\epsilon(s), \alpha^\epsilon(s), c) \hat{m}_s^\epsilon(dc) ds \\ &+ \frac{1}{\sqrt{\epsilon}} \int_0^t \sigma_K(X_K^\epsilon(s), z^\epsilon(s)) ds, \end{aligned} \tag{2.11}$$

where $b_K = bh_K$ and $\sigma_K = \sigma h_K$ are the truncated versions of b and σ. Now, instead of the original martingale problem, we aim to show that

$$\Phi_{f,K}(t) = f(X(t), \bar{\alpha}(t)) - f(X(0), \bar{\alpha}(0)) - \int_0^t \mathcal{G}_K f(X(s), \bar{\alpha}(s)) ds, \tag{2.12}$$

is a martingale, where $\mathcal{G}_K$ is essentially $\mathcal{G}$ but b and σ have been replaced with their truncated counterparts. To proceed, we need the following lemmas.

Lemma 2.3.1. *Under the conditions of Theorem 2.3.1, $\{\hat{m}^\epsilon(\cdot)\}$ is tight in $M(\infty)$ and $\{\bar{\alpha}^\epsilon(\cdot)\}$ is tight in $D([0,\infty) : \mathcal{M})$.*

Proof. First note that since $\mathcal{U} \times [0, \infty)$ is a Polish space (separable, complete, and metrizable), we have that $\{\hat{m}^\epsilon(\cdot)\}$ is tight in $M(\infty)$. Furthermore, from our assumptions, it follows that $\{\bar{\alpha}^\epsilon(\cdot)\}$ is tight in $D([0,\infty) : \mathcal{M})$ from [20]. □

Now, to proceed, we need a different notion of convergence, that is weaker than strong convergence; see [8]. We say that f is the p-limit of $\{f_k\}$ and write

$$p - \lim_{k\to\infty} f_k = f,$$

if and only if

$$\sup_k |f_k| < \infty$$

and

$$\lim_{k\to\infty} E(|f_k(t) - f(t)|) = 0 \qquad \text{for all } t$$

Now, let $\mathcal{F}_t^\epsilon$ be the minimal sigma algebra so that $\{\alpha^\epsilon(s), m^\epsilon(\cdot), z^\epsilon(s) : s \le t\}$ is $\mathcal{F}_t^\epsilon$-measurable and let E_t^ϵ denote the conditional expectation with respect to $\mathcal{F}_t^\epsilon$. Given an operator $\mathcal{A}^\epsilon$, we say that a function $f(\cdot)$ is in the domain of $\mathcal{A}^\epsilon$, and $\mathcal{A}^\epsilon f = g$, if

$$p - \lim_{\Delta\to 0} \left| \frac{E_t^\epsilon f(t+\Delta) - f(t)}{\Delta} - g(t) \right| = 0.$$

Two consequences of this definition are that given $f(\cdot)$ is in the domain of $\mathcal{A}^\epsilon$, we have that

$$f(t) - \int_0^t \mathcal{A}^\epsilon f(s) ds$$

is a martingale, and

$$E_t^\epsilon f(t+s) - f(t) = \int_t^{t+s} E_t^\epsilon \mathcal{A}^\epsilon f(r) dr \qquad \text{w.p.1},$$

see [20] or [9] for more details. We now state a theorem, from [11], which will be used soon to verify the tightness of certain sequences.

Theorem 2.3.2. *(Kushner [11]) Suppose that $Y^\epsilon(\cdot)$ has sample paths in $D([0,\infty)$:*

$\mathbb{R}^n$) *and suppose that*

$$\lim_{N\to\infty}\limsup_{\epsilon\to 0} P(\sup_{t\in[0,T]}|Y^\epsilon(t)|\geq N)=0, \quad \textit{for each} \quad T<\infty.$$

Given $f^\epsilon(\cdot)$ *in the domain of* $\mathcal{A}^\epsilon$*, if for each* $T<\infty$*,* $\{\mathcal{A}^\epsilon f^\epsilon(t) : \epsilon>0, t\leq T\}$ *is uniformly integrable and for each* $\lambda>0$*,*

$$\lim_{\epsilon\to 0} P(\sup_{t\leq T}|f^\epsilon(t)-f(Y^\epsilon(t))|\geq\lambda)=0,$$

then $\{Y^\epsilon(\cdot)\}$ *is tight in* $D([0,\infty):\mathbb{R}^n)$*.*

We now return to our problem.

Lemma 2.3.2. *Under the conditions of Theorem 2.3.1,* $\{(X_K^\epsilon(\cdot),\bar{\alpha}^\epsilon(\cdot))\}$ *is tight in* $D([0,\infty):\mathbb{R}^n\times\bar{\mathcal{M}})$*.*

Proof. We verify this result using an Aldous-like tightness criterion. For $i\in\bar{\mathcal{M}}$ and $\eta(\cdot,i)\in C_0^2$ define

$$\hat{\eta}(x,\alpha)=\sum_{i\in\bar{\mathcal{M}}}\eta(x,i)I_{\{\alpha\in\mathcal{M}_i\}}, \quad \text{for } \alpha\in\mathcal{M}; \tag{2.13}$$

note that $\hat{\eta}(x,\alpha)$ is identically zero if α does not belong to the class $\mathcal{M}_i$.

Consider the operator $\mathcal{G}^\epsilon$ given by:

$$\mathcal{G}^\epsilon\hat{\eta}(x,\alpha)=L^\epsilon\hat{\eta}(x,\alpha)+\sum_{\beta\in\mathcal{M}}q_{\alpha\beta}^\epsilon(t)\hat{\eta}(x,\beta), \quad \text{for } \alpha\in\mathcal{M}, \tag{2.14}$$

where

$$L^\epsilon\hat{\eta}(x,\alpha)=D\hat{\eta}(x,\alpha)'\left[\int_{\mathcal{U}}b(x,\alpha,c)m_t^\epsilon(dc)+\frac{1}{\sqrt{\epsilon}}\sigma(x,z)\right].$$

Furthermore, let L_K^ϵ denote the truncated version of L^ϵ where b and σ are replaced with b_K and σ_K respectively. Note that, from our definition of $\hat{\eta}$, it follows that

$$\begin{aligned}\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t)) &:= \sum_{i\in\bar{\mathcal{M}}} \eta(X_K^\epsilon(t), i) I_{\{\bar{\alpha}^\epsilon(t)=i\}} \\ &= \eta(X_K^\epsilon(t), \bar{\alpha}^\epsilon(t)).\end{aligned}$$

We have

$$\begin{aligned}L_K^\epsilon \hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t)) = D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))'[\int_{\mathcal{U}} b_K(X_K^\epsilon(t), \alpha^\epsilon(t), c) m_t^\epsilon(dc) \\ + \frac{1}{\sqrt{\epsilon}} \sigma_K(X_K^\epsilon(t), z^\epsilon(t))].\end{aligned} \tag{2.15}$$

Now, for $T > 0$ with $t \leq T$, define the first perturbation of $\hat{\eta}$ by

$$\hat{\eta}_1^\epsilon(t) = \hat{\eta}_1^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) = \frac{1}{\sqrt{\epsilon}} \int_t^T E_t^\epsilon[D\hat{\eta}(X_K^\epsilon(s), \alpha^\epsilon(s))' \sigma_K(X_K^\epsilon(s), z^\epsilon(s))] ds$$

Using a change of variables we obtain

$$\hat{\eta}_1^\epsilon(t) = \sqrt{\epsilon} \int_{t/\epsilon}^{T/\epsilon} E_t^\epsilon\big[D\hat{\eta}(X_K^\epsilon(s), \alpha^\epsilon(s))' \sigma_K(X_K^\epsilon(s), z(s))\big] ds.$$

Furthermore, using the mixing condition on $z(\cdot)$, the independence of $\alpha^\epsilon(\cdot)$ and $z(\cdot)$, and the fact that $X_K^\epsilon(\cdot)$ is bounded, it follows that

$$\hat{\eta}_1^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) = O(\sqrt{\epsilon}). \tag{2.16}$$

Therefore,

$$\sup_{t\leq T} |\hat{\eta}_1^\epsilon(t)| \xrightarrow{p} 0, \quad \text{as } \epsilon \to 0.$$

Now, using some algebraic manipulation, it follows that

$$\begin{aligned} L_K^\epsilon \hat{\eta}_1^\epsilon(t) &= -\frac{1}{\sqrt{\epsilon}} D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \sigma_K(X_K^\epsilon(t), z^\epsilon(t)) \\ &\quad + \frac{1}{\sqrt{\epsilon}} \int_t^T \left[D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' E_t^\epsilon \sigma_K(X_K^\epsilon(t), z^\epsilon(s))\right]_x' \dot{X}_K^\epsilon(t) ds \end{aligned} \tag{2.17}$$

So, if we define $\hat{\eta}^\epsilon(t) = \hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t)) + \hat{\eta}_1^\epsilon(t)$, we get

$$\begin{aligned} L_K^\epsilon \hat{\eta}^\epsilon(t) &= L_K^\epsilon \hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t)) + L_K^\epsilon \hat{\eta}_1^\epsilon(t) \\ &= D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \int_{\mathcal{U}} b_K(X_K^\epsilon(t), \alpha^\epsilon(t), c) m_t^\epsilon(dc) \\ &\quad + \int_{t/\epsilon}^{T/\epsilon} E_t^\epsilon \big[D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \sigma_K(X_K^\epsilon(t), z(s))\big]_x' \sigma_K(X_K^\epsilon(t), z^\epsilon(t)) ds \\ &\quad + \sqrt{\epsilon} \int_{t/\epsilon}^{T/\epsilon} E_t^\epsilon [D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \sigma_K(X_K^\epsilon(t), z(s))]_x' \\ &\qquad \times \int_{\mathcal{U}} b_K(X_K^\epsilon(t), \alpha^\epsilon(t), c) m_t^\epsilon(dc). \end{aligned} \tag{2.18}$$

We now aim to analyze and bound each of the terms in the previous equation. First, as $\hat{\eta}$ is smooth with compact support and b_K is bounded, it follows that

$$|D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))| \cdot \left| \int_{\mathcal{U}} b_K(X_K^\epsilon(t), \alpha^\epsilon(t), c) m_t^\epsilon(dc) \right|$$

is uniformly bounded. By the mixing condition on $z(\cdot)$, it follows that

$$\left| \int_{t/\epsilon}^{T/\epsilon} E_t^\epsilon \left[D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t)) \sigma_K(X_K^\epsilon(t), z(s))\right]_x' \sigma_K(X_K^\epsilon(t), z^\epsilon(t)) ds \right|$$

is also uniformly bounded. Furthermore, in a similar fashion, we have

$$\sqrt{\epsilon}\left|\int_{t/\epsilon}^{T/\epsilon} E_t^\epsilon\left[D\hat{\eta}(X_K^\epsilon(t),\alpha^\epsilon(t))\sigma_K(X_K^\epsilon(t),z(s))\right]_x' \int_{\mathcal{U}} b_K(X_K^\epsilon(t),\alpha^\epsilon(t),c)m_t^\epsilon(dc)\right|$$

$$= O(\sqrt{\epsilon})$$

Due to (A2) and the definition of $\hat{\eta}(\cdot)$, we get that

$$\begin{aligned} &Q^\epsilon(t)[\hat{\eta}(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t)) + \hat{\eta}_1^\epsilon(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))] \\ &= \hat{Q}(t)[\hat{\eta}(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t)) + \hat{\eta}_1^\epsilon(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))] \end{aligned} \tag{2.19}$$

as a result of the fact that

$$\tilde{Q}(t)\mathbf{1} = 0,$$

where $\mathbf{1} = \mathrm{diag}(\mathbb{1}_{m_1},\ldots,\mathbb{1}_{m_l})$, and $\mathbb{1}_\ell$ is an ℓ-dimensional vector with all entries being 1. Thus,

$$p - \lim \hat{Q}(t)\hat{\eta}_1^\epsilon(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t)) \to 0,$$

as $\epsilon \to 0$. Combining all our estimates for (2.18), we see that $\{\mathcal{G}_K^\epsilon \hat{\eta}^\epsilon(t)\}$ is a uniformly integrable family. In addition, the process $\{X_K^\epsilon(\cdot)\}$ clearly satisfies the first condition of Theorem 2.3.3 due to their truncations. Finally, using (2.16), we see that all the conditions of Theorem 2.3.2 are satisfied, and, therefore, it follows that $\{(X_K^\epsilon(\cdot),\alpha^\epsilon(t))\}$ is tight. □

We have now shown that $\{X_K^\epsilon(\cdot)\}$, $\{\bar{\alpha}^\epsilon(\cdot)\}$, and $\{\hat{m}^\epsilon(\cdot)\}$ are tight in $D([0,\infty):\mathbb{R}^n)$, $D([0,\infty):\bar{\mathcal{M}})$, and $M(\infty)$, respectively. Our next order of business is to analyze the limit we observe under the weak convergence. Citing a result found in [8], if we can show that $\hat{\eta}(\cdot,\cdot) \in D(\mathcal{G}_K^\epsilon)$, it will follow that

$$\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t)) - \int_0^t \mathcal{G}_K^\epsilon \hat{\eta}(X_K^\epsilon(s), \alpha^\epsilon(s))ds$$

is a martingale. This will also imply that

$$E_t^\epsilon \hat{\eta}(X_K^\epsilon(t+s), \alpha^\epsilon(t+s)) - \hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t)) = \int_t^{t+s} E_t^\epsilon \mathcal{G}_K^\epsilon \hat{\eta}(X_K^\epsilon(r), \alpha^\epsilon(r))dr \quad \text{w.p.1}$$

Let

$$\theta^\epsilon(x, \alpha, z^\epsilon(t)) := \int_{t/\epsilon}^{T/\epsilon} [E_t^\epsilon \hat{\eta}^\epsilon(x,\alpha)\sigma_K(x, z(s))]_x' \, \sigma_K(x, z^\epsilon(t))ds,$$

and with $\hat{\eta}_1^\epsilon(\cdot)$ as before, we define the second perturbation

$$\hat{\eta}_2^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) := \int_t^T \left[E_t^\epsilon \theta^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t), z^\epsilon(s)) - E\theta^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t), z^\epsilon(s))\right] ds. \tag{2.20}$$

By substituting the definition of θ^ϵ into the second permutation, we see that

$$\begin{aligned}\hat{\eta}_2^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) = \epsilon \int_{t/\epsilon}^{T/\epsilon} \int_r^{T/\epsilon} &[E_t^\epsilon [D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))'\sigma_K(X_K^\epsilon(t), z(s))]_x' \sigma_K(X_K^\epsilon(t), z(r)) \\ &- E[E_t^\epsilon D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))'\sigma_K(X_K^\epsilon(t), z(s))]_x' \sigma_K(X_K^\epsilon(t), z(r))]dsdr.\end{aligned}$$

Furthermore, we have

$$\begin{aligned} p - \lim \hat{\eta}_1^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) = 0, \\ p - \lim \hat{\eta}_2^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) = 0, \end{aligned} \tag{2.21}$$

by letting $\epsilon \to 0$.

Now we define

$$\hat{\eta}^\epsilon(x,\alpha) = \hat{\eta}(x,\alpha) + \hat{\eta}_1^\epsilon(x,\alpha) + \hat{\eta}_2^\epsilon(x,\alpha). \tag{2.22}$$

Applying L_K^ϵ to $\hat{\eta}_2^\epsilon(\cdot,\cdot)$ yields

$$
\begin{aligned}
& L_K^\epsilon \hat{\eta}_2^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) \\
& = -\theta^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t), z^\epsilon(t)) \\
& \quad + \int_{t/\epsilon}^{T/\epsilon} E[D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \sigma_K(X_K^\epsilon(t), z(s))]_x' \sigma_K(X_K^\epsilon(t), z(r)) ds dr \\
& \quad + \epsilon \int_{t/\epsilon}^{T/\epsilon} \int_r^{T/\epsilon} \Big[E_t^\epsilon [D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \sigma_K(X_K^\epsilon(t), z(s))]_x' \sigma_K(X_K^\epsilon(t), z(r)) \\
& \qquad - E[D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \sigma_K(X_K^\epsilon(t), z(s))]_x \sigma_K(X_K^\epsilon(t), z(r)) \Big]_x' \dot{X}_K^\epsilon(t) ds dr.
\end{aligned}
\tag{2.23}
$$

Combining everything we have used so far, we see that

$$
\begin{aligned}
& \mathcal{G}_K^\epsilon \hat{\eta}^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) \\
& \quad = L_K^\epsilon \left[\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t)) + \hat{\eta}_1^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) + \hat{\eta}_2^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) \right] \\
& \qquad + \sum_{j \in \mathcal{M}} q_{ij}^\epsilon(t) \left[\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t)) + \hat{\eta}_1^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) + \hat{\eta}_2^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) \right] \\
& \quad = D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \int_{\mathcal{U}} b_K(X_K^\epsilon(t), \alpha^\epsilon(t), c) m_t^\epsilon(dc) \\
& \qquad + \int_{t/\epsilon}^{T/\epsilon} E[D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \sigma_K(X_K^\epsilon(t), z(s))]_x' \sigma_K(X_K^\epsilon(t), z^\epsilon(t)) ds \\
& \qquad + \sum_{j \in \mathcal{M}} q_{ij}^\epsilon(t) \left[\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t)) + \hat{\eta}_1^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) + \hat{\eta}_2^\epsilon(X_K^\epsilon(t), \alpha^\epsilon(t)) \right] \\
& \qquad + e^\epsilon(t),
\end{aligned}
\tag{2.24}
$$

where $e^\epsilon(\cdot)$ is an error term satisfying

$$
p - \lim e^\epsilon(t) \to 0, \quad \text{as} \quad \epsilon \to 0.
$$

As before in (2.19), we see that

$$\begin{aligned}&Q^\epsilon(t)\left[\hat\eta(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))+\hat\eta_1^\epsilon(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))+\hat\eta_2^\epsilon(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))\right]\\&=\hat Q(t)\left[\hat\eta(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))+\hat\eta_1^\epsilon(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))+\hat\eta_2^\epsilon(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))\right].\end{aligned}$$

In addition, equations (2.21) yields

$$p-\lim\hat Q(t)\left[\hat\eta_1^\epsilon(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))+\hat\eta_2^\epsilon(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))\right]=0.$$

Thus, the only term that requires our attention is $\hat Q(t)\hat\eta(X_K^\epsilon(t),\cdot)(\alpha^\epsilon(t))$.

As $\{X_K^\epsilon(\cdot),\bar\alpha^\epsilon(\cdot),\hat m^\epsilon(\cdot)\}$ is tight, we may select a weakly convergent subsequence that we still denote by $\{X_K^\epsilon(\cdot),\bar\alpha^\epsilon(\cdot),\hat m^\epsilon(\cdot)\}$, for notational simplicity. Using the Skorohod representation we may assume, with a slight abuse of notation, that $(X_K^\epsilon(\cdot),\bar\alpha^\epsilon(\cdot),\hat m^\epsilon(\cdot))$ converges to $(X_K(\cdot),\bar\alpha(\cdot),\hat m(\cdot))$ with probability 1. Finally, we use an idea from [11] to complete the proof. For arbitrary $t>0$, $s>0$, bounded and continuous $f(\cdot)$, and positive integer N_0, we show that

$$\begin{aligned}Ef(X_K(t),\bar\alpha(t):t_l<t,l\le N_0)\Big[&\eta(X_K(t+s),\bar\alpha(t+s))-\eta(X_K(t),\bar\alpha(t))\\&-\int_t^{t+s}\mathcal G_K\eta(X_K(r),\bar\alpha(r))dr\Big]=0.\end{aligned}$$

This implies that

$$\eta(X_K(t),\bar\alpha(t))-\int_t^{t+s}\mathcal G_K\eta(X_K(r),\bar\alpha(r))dr$$

is a martingale, completing the proof.

First, we have

$$\begin{aligned}
&\int_t^{t+s} \hat{Q}(r)\hat{\eta}(X_K^\epsilon(r), \alpha^\epsilon(r))dr \\
&\quad = \int_t^{t+s} \sum_{i\in\bar{\mathcal{M}}} \sum_{j\in\mathcal{M}_i} \hat{Q}(r)\hat{\eta}(X_K^\epsilon(t), s_{ij}) I_{\{\alpha^\epsilon(r)=s_{ij}\}} dr \\
&\quad = \int_t^{t+s} \sum_{i\in\bar{\mathcal{M}}} \sum_{j\in\mathcal{M}_i} \hat{Q}(r)\hat{\eta}(X_K^\epsilon(t), s_{ij}) v_j^i(r) I_{\{\bar{\alpha}^\epsilon(r)=i\}} dr \\
&\qquad + \int_t^{t+s} \sum_{i\in\bar{\mathcal{M}}} \sum_{j\in\mathcal{M}_i} \hat{Q}(r)\hat{\eta}(X_K^\epsilon(t), s_{ij}) [I_{\{\alpha^\epsilon(r)=s_{ij}\}} - v_j^i(r) I_{\{\bar{\alpha}^\epsilon(r)=i\}}] dr.
\end{aligned} \tag{2.25}$$

Now, by Theorem 5.52 in [22], for each $i \in \bar{\mathcal{M}}$ and $j = 1, 2, \ldots, m_i$,

$$E\left(\int_t^{t+s} [I_{\{\alpha^\epsilon(r)=s_{ij}\}} - v_j^i(r) I_{\{\bar{\alpha}^\epsilon(r)=i\}}] dr\right)^2 \to 0 \quad \text{as} \quad \epsilon \to 0.$$

This clearly implies that we may omit the last line of (2.25) in our analysis. We have

$$\begin{aligned}
&\lim_{\epsilon\to 0} Ef(X_K^\epsilon(t), \bar{\alpha}^\epsilon(t) : t_l < t, l \le N_0) \left[\int_t^{t+s} \hat{Q}(r)\hat{\eta}(X_K^\epsilon(r), \alpha^\epsilon(r))dr\right] \\
&= \lim_{\epsilon\to 0} f(X_K^\epsilon(t), \bar{\alpha}^\epsilon(t) : t_l < t, l \le N_0) \left[\int_t^{t+s} [v(t)\hat{Q}(r)\mathbf{1}]\eta(X_K^\epsilon(r), \bar{\alpha}(r))dr\right] \\
&= Ef(X_K(t), \bar{\alpha}(t) : t_l < t, l \le N_0) \left[\int_t^{t+s} \bar{Q}(r)\eta(X_K(r), \bar{\alpha}(r))dr\right].
\end{aligned} \tag{2.26}$$

Furthermore, from the weak convergence and the Skorohod representation, we have

$$
\begin{aligned}
&Ef(X_K^\epsilon(t), \bar{\alpha}^\epsilon(t) : t_l < t, l \le N_0) \\
&\qquad \times \left[\int_t^{t+s} D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \int_{\mathcal{U}} b_K(X_K^\epsilon(t), \alpha^\epsilon(t), c) m_r^\epsilon(dc) dr \right] \\
&\quad = Ef(f(X_K^\epsilon(t), \bar{\alpha}^\epsilon(t) : t_l < t, l \le N_0) \\
&\qquad \times \Big[\sum_{i \in \bar{\mathcal{M}}} \sum_{j \in \mathcal{M}_i} \int_t^{t+s} D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \int_{\mathcal{U}} b_K(X_K^\epsilon(t), s_{ij}, c) m_r^\epsilon(dc) v_j^i(r) I_{\{\bar{\alpha}^\epsilon(r) = i\}}] dr \\
&\qquad + \sum_{i \in \bar{\mathcal{M}}} \sum_{j \in \mathcal{M}_i} \int_t^{t+s} D\hat{\eta}(X_K^\epsilon(t), \alpha^\epsilon(t))' \int_{\mathcal{U}} b_K(X_K^\epsilon(t), s_{ij}, c) m_r^\epsilon(dc) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \times [I_{\{\bar{\alpha}^\epsilon(r) = i\}} - v_J^i(t) I_{\{\bar{\alpha}^\epsilon(r) = i\}}] dr \Big] \\
&\quad \to Ef(X_K(t), \bar{\alpha}(t), t_l < t, l \le N_0) \\
&\qquad\qquad\qquad \times \left[\int_t^{t+s} D\eta(X_K(t), \bar{\alpha}(t))' \int_{\mathcal{U}} b_K(X_K(t), \bar{\alpha}(t), c) m_r^\epsilon(dc) dr \right].
\end{aligned}
\tag{2.27}
$$

We have now shown that the following lemma is true.

Lemma 2.3.3. *Under the conditions of Theorem 2.3.1, the pair $(X_K^\epsilon(\cdot), \bar{\alpha}^\epsilon(\cdot))$ converges weakly to $(X_K(\cdot), \bar{\alpha}(\cdot))$, and this process is the solution to the martingale problem with operator $\mathcal{G}_K$.*

Finally, we must show that our original (untruncated) process $(X(\cdot), \bar{\alpha}(\cdot))$ is the solution of the martingale problem with operator $\mathcal{G}$.

Lemma 2.3.4. *Under the conditions of Theorem 2.3.1, the pair $(X^\epsilon(\cdot), \bar{\alpha}^\epsilon(\cdot))$ converges weakly to $(X(\cdot), \bar{\alpha}(\cdot))$, and this process is the solution to the martingale problem with operator $\mathcal{G}$.*

Proof. First, we have seen that $\Phi_{f,K}(\cdot)$ are martingales with respect to the sigma-algebra $\mathcal{F}_t = \sigma\{X_K(s), \bar{\alpha}(s), \hat{m}(B \times [0, s]) : \text{for Borel sets } B \text{ with } s \le t\}$. Therefore, there exists a standard Brownian motion $w_K(\cdot)$ that is adapted to $\mathcal{F}_t$ such that

$$dX_K(t) = \int_{\mathcal{U}} b_K(X_K(t), \bar{\alpha}(t), c)\hat{m}_t(dc) + \bar{\sigma}(X_K(t))dw_K(t) \tag{2.28}$$

If we let $K \to \infty$, we see that $(X(\cdot), \bar{\alpha}(\cdot))$ satisfies equation (2.22) but with untruncated functions. Furthermore, because $(X^\epsilon(\cdot), \bar{\alpha}^\epsilon(\cdot), \hat{m}^\epsilon(\cdot))$ converges weakly to $(X(\cdot), \bar{\alpha}(\cdot), \hat{m}(\cdot))$, the result contained in Theorem 2.3.1 now follows. □

Proposition 2.3.1. *Assume that $(X^\epsilon(\cdot), \bar{\alpha}^\epsilon(\cdot), m^\epsilon(\cdot))$ converges weakly to $(X(\cdot), \bar{\alpha}(\cdot), m(\cdot))$ where $m(\cdot)$ is admissible with respect to $w(\cdot)$. Then, as $\epsilon \to 0$,*

$$\begin{aligned} J^\epsilon(x, i, m^\epsilon(\cdot)) &\to J(x, i, m(\cdot)) \\ &= E^{x,i} \int_0^{\tilde{T}} \int_{\mathcal{U}} \bar{C}(X(t), \bar{\alpha}(t), c) m_t(dc)dt, \end{aligned} \tag{2.29}$$

where

$$\bar{C}(x, i, c) = \sum_{j=1}^{m_i} v_j^i(t) C(x, s_{ij}, c).$$

Proof. First, by applying assumption (A4) and the Cauchy-Schwarz inequality we obtain

$$\begin{aligned} &E\left|\int_0^{\tilde{T}} \int_{\mathcal{U}} C(X^\epsilon(t), s_{ij}, c) \left[I_{\{\alpha^\epsilon(t)=s_{ij}\}} - v_j^i(t) I_{\{\bar{\alpha}^\epsilon(t)=i\}}\right] m_t^\epsilon(dc)dt\right|^2. \\ &\le \left[\int_0^T \kappa^2(1 + E|X^\epsilon(t)|^{2p_0})dt\right] E\left[\int_0^{\tilde{T}} \left[I_{\{\alpha^\epsilon(t)=s_{ij}\}} - v_j^i(t) I_{\{\bar{\alpha}^\epsilon(t)=i\}}\right]^2 dt\right] \\ &\le \left[\int_0^T \kappa^2(1 + E|X^\epsilon(t)|^{2p_0})dt\right] E\left[\int_0^{\tilde{T}} \left[I_{\{\alpha^\epsilon(t)=s_{ij}\}} - v_j^i(t) I_{\{\bar{\alpha}^\epsilon(t)=i\}}\right] dt\right]^2 \\ &= O(\epsilon) \to 0 \text{ as } \epsilon \to 0, \end{aligned} \tag{2.30}$$

by Theorem 5.25 in [20]. Because $(X^\epsilon(\cdot), \bar{\alpha}^\epsilon(\cdot))$ converges weakly to $(X(\cdot), \bar{\alpha}(\cdot))$, using the Skorohod representation we may assume, abusing notation slightly, that $(X^\epsilon(\cdot), \bar{\alpha}^\epsilon(\cdot))$ converges to $(X(\cdot), \bar{\alpha}(\cdot))$ w.p.1. Furthermore, we have

$$
\begin{aligned}
&\lim_{\epsilon\to 0} J^\epsilon(x,\iota,m^\epsilon)\\
&\quad= \lim_{\epsilon\to 0} E^{x,\iota}\int_0^{\tilde{T}}\int_{\mathcal{U}} C(X^\epsilon(t),\alpha^\epsilon(t),c)m_t^\epsilon(dc)dt\\
&\quad= \lim_{\epsilon\to 0}\sum_{i=1}^{l}\sum_{j=1}^{m_i} E^{x,\iota}\int_0^{\tilde{T}}\int_{\mathcal{U}} C(X^\epsilon(t),s_{ij},c)I_{\{\alpha^\epsilon(t)=s_{ij}\}}m_t^\epsilon(dc)dt\\
&\quad= \lim_{\epsilon\to 0}\sum_{i=1}^{l}\sum_{j=1}^{m_i} E^{x,\iota}\int_0^{\tilde{T}}\int_{\mathcal{U}} C(X^\epsilon(t),s_{ij},c)\left[I_{\{\alpha^\epsilon(t)=s_{ij}\}} - v_j^i(t)I_{\{\bar{\alpha}^\epsilon(t)=i\}}\right]m_t^\epsilon(dc)dt\\
&\qquad+ \lim_{\epsilon\to 0}\sum_{i=1}^{l}\sum_{j=1}^{m_i} E^{x,\iota}\int_0^{\tilde{T}}\int_{\mathcal{U}} C(X^\epsilon(t),s_{ij},c)v_j^i(t)I_{\{\bar{\alpha}^\epsilon(t)=i\}}m_t^\epsilon(dc)dt\\
&\quad= \sum_{i=1}^{l}\sum_{j=1}^{m_i} E^{x,\iota}\int_0^{\tilde{T}}\int_{\mathcal{U}} C(X(t),s_{ij},c)v_j^i(t)I_{\{\bar{\alpha}(t)=i\}}m_t(dc)dt\\
&\quad= \sum_{i=1}^{l}\sum_{j=1}^{m_i} E^{x,\iota}\int_0^{\tilde{T}}\int_{\mathcal{U}} \bar{C}(X(t),i,c)I_{\{\bar{\alpha}(t)=i\}}m_t(dc)dt\\
&\quad= E^{x,\iota}\int_0^{\tilde{T}}\int_{\mathcal{U}} \bar{C}(X(t),\bar{\alpha}(t),c)m_t(dc)dt.
\end{aligned}
\tag{2.31}
$$

Thus giving the desired result. This completes the proof. □

2.4 Nearly Optimal Control

Starting with the limit system, we can construct optimal of near-optimal controls. Then we use such controls in the original problem. We aim to show that such constructed controls are nearly optimal. The main tool for this purpose is the so-called chattering lemma, which we state now; a proof can be found in [11].

Lemma 2.4.1. *Assume that (A2) holds and that (2.3) has a weak solution for every* *initial condition for the admissible pair* $(m(\cdot), w(\cdot))$*. Given* $T > 0$ *and* $\Delta > 0$*, there exists a finite set* $\mathcal{U}^\Delta = \{a_1^\Delta, a_2^\Delta, \dots, a_{k_\Delta}^\Delta\} \subset \mathcal{U}$*, a* $\delta > 0$*, and a* $\mathcal{U}^\Delta$*-valued ordinary admissible stochastic control* $u^\Delta(\cdot)$ *such that* u^Δ *is constant on each interval* $[i\delta, i\delta+\delta)$ *and for all* m *we have*

$$\begin{aligned} &P_{x,i}^m(\sup_{t\le T}|x(t,u^\Delta) - x(t,m)| > \Delta) \le \Delta, \\ &|J(x,i,m) - J(x,i,u^\Delta)| \le \Delta. \end{aligned} \tag{2.32}$$

If the solution to (2.3) is unique for each admissible control $m(\cdot)$, then the above equation holds for all $m(\cdot)$ simultaneously.

Lemma 2.4.2. *Suppose that* $m(\cdot)$ *is an admissible relaxed control with respect to the Brownian motion* $w(\cdot)$*.*

- *Then there is a nonanticipative solution to (2.3) with* $X(0) = x_0$ *and*

$$E\sup_{t\le T}|X(t)|^2 \le K_0(1+|x_0|^2),$$

  where K_0 *depends only on* T *and the Lipschitz constant of* $b(\cdot)$ *and* $\sigma(\cdot)$*.*

- *Let* $m^n(\cdot)$ *converge weakly to* $m(\cdot)$*, where each* $m^n(\cdot)$ *is an admissible control with respect to the Brownian motion* $w(\cdot)$*. Suppose* $x(m^n(\cdot),\cdot)$ *satisfies (2.3). Then* $(X^n(m^n(\cdot),\cdot), m^n(\cdot))$ *converges weakly to* $(X(m(\cdot),\cdot), m(\cdot))$ *where* $(X(m(\cdot),\cdot), m(\cdot))$ *satisfies (2.3). Furthermore,* $m(\cdot)$ *is admissible with respect to the Brownian motion* $w(\cdot)$*.*

Proof. Proofs of both items appear in several sources; see, for example, [11] and [5]. □

Lemma 2.4.3. *In the class of admissible controls for (2.3), there is an optimal control.*

Proof. Let $m^\delta(\cdot)$ be a sequence of relaxed controls that converge weakly to some $m(\cdot)$ and satisfy $J(m^\delta) \to \inf_{\mathcal{R}} J(m)$, as $\delta \to 0$. Then, by Lemma 2.3.6, if $x(m^\delta, \cdot)$ is a trajectory satisfying (2.7), it follows that $(X^\delta(m^\delta(\cdot), \cdot), m^\delta(\cdot))$ converges weakly to $(X(m(\cdot), \cdot), m(\cdot))$, where $(X(m(\cdot), \cdot), m(\cdot))$ satisfies (2.7) and such that $m(\cdot)$ is admissible with respect to $w(\cdot)$. Hence, $m(\cdot)$ is the desired admissible control. $\square$

Finally, we show that those controls obtained in the limit are near-optimal.

Theorem 2.4.1. *Assume the conditions of Theorem 2.3.1 are satisfied. Suppose that $u^\delta(\cdot)$ is a δ-optimal control of (2.3). Then this control is nearly optimal satisfying*

$$\lim_{\epsilon \to 0} \sup [J^\epsilon(u^\delta) - \inf_{\mathcal{R}^\epsilon} J^\epsilon(m)] \le \delta. \tag{2.33}$$

where $\mathcal{R}^\epsilon$ denotes the class of admissible controls corresponding to (2.1) and (2.2).

Proof. First, by Proposition 2.3.1, we have that

$$J^\epsilon(u^\delta) \to J(u^\delta) \text{ as } \epsilon \to 0.$$

Furthermore, as u^δ is a δ-optimal control, we have

$$J(u^\delta) \ge \inf_{\mathcal{R}} J(m),$$

where $\mathcal{R}$ denotes the class of admissible controls for (2.2). Combining this with the fact that $\hat{m}^\epsilon(\cdot)$ is a Δ_ϵ-optimal control, we obtain

$$\begin{aligned}
\inf_{m\in\mathcal{R}^\epsilon} J^\epsilon(m) + \Delta_\epsilon &\geq J^\epsilon(\hat{m}^\epsilon) \\
&\to J(\hat{m}) \\
&\geq \inf_{m\in\mathcal{R}} J(m) \\
&\geq J(u^\delta) - \delta \\
&= J^\epsilon(u^\delta) + \varrho_\epsilon + \delta,
\end{aligned} \tag{2.34}$$

where ϱ_ϵ is an error term such that $\varrho_\epsilon \to 0$ as $\epsilon \to 0$. Therefore,

$$J^\epsilon(u^\delta) - \inf_{m\in\mathcal{R}^\epsilon} J^\epsilon(m) \leq \Delta_\epsilon - \varrho_\epsilon + \delta, \tag{2.35}$$

from which (2.19) follows after taking $\limsup$ as $\epsilon \to 0$ of both sides. This completes the proof. $\square$

BIBLIOGRAPHY

[1] Ando, Albert, and Franklin M. Fisher. "Near-decomposability, partition and aggregation, and the relevance of stability discussions." International Economic Review 4.1 (1963): 53-67.

[2] A. Friedman, Stochastic Differential Equations and Applications. Dover Publications, 2006.

[3] Baran, Nicholas A., George Yin, and Chao Zhu. "Feynman-Kac formula for switching diffusions: connections of systems of partial differential equations and stochastic differential equations." Advances in Difference Equations 2013.1 (2013): 315.

[4] Billingsley, Patrick. Convergence of probability measures. Vol. 493. John Wiley & Sons, 2009.

[5] Fleming, W. H., and M. Nisio. "On stochastic relaxed control for partially observed diffusions." Nagoya Mathematical Journal 93 (1984): 71-108.

[6] Freidlin, M: Functional Integration and Partial Differential Equations, Princeton University Press, Princeton (1985)

[7] Friedman, Avner. Stochastic differential equations and applications. Courier Dover Publications, 2012.

[8] Kurtz, Thomas G. Semigroups of Conditioned Shifts and Approximation of Markov Processes. The Annals of Probability 3 (1975), no. 4, 618–642. doi:10.1214/aop/1176996305. http://projecteuclid.org/euclid.aop/1176996305.

[9] Kushner, Harold Joseph. Approximation and weak convergence methods for random processes, with applications to stochastic systems theory. Vol. 6. MIT press, 1984.

[10] Kushner, Harold J., and W. Runggaldier. "Nearly optimal state feedback controls for stochastic systems with wideband noise disturbances." SIAM Journal on Control and Optimization 25.2 (1987): 298-315.

[11] Kushner, Harold J. Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems, Birkhäuser, Boston, 1990.

[12] X. Mao, Stochastic Differential Equations and Their Applications. Horwood publishing, 1997.

[13] Øksendal, Bernt. Stochastic differential equations. Springer Berlin Heidelberg, 2003.

[14] M. Rosenblatt. A central limit theorem and a strong mixing condition. Proc. Natl. Acad. Sci. USA 42 (1956) 43-47. MR74711

[15] Skorokhod, A. Asymptotic Methods in the Theory of Stochastic Differential Equations, 2008. [?? Amer. Math. Soc., Providence, RI, 1989.]

[16] Warga, Jack. "Relaxed variational problems." Journal of Mathematical Analysis and Applications 4.1 (1962): 111-128.

[17] Weinberger, HF: Some remarks on invariant sets for systems. Maximum Principles and Eigenvalue Problems in Partial Differential Equations. 189207 (1988) [??reference incomplete]

[18] Yao, David D., Hanqin Zhang, and Xun Yu Zhou, eds. Stochastic Modeling and Optimization: With Applications in Queues, Finance, and Supply Chains. Springer, 2003.

[19] Yin, G., Xianping Guo, Yousef Talafha, Nicholas A. Baran. "Near-Optimal controls of Differential Systems with Switching and Random Jumps Subject to Fast Switching and Wideband Noise Perturbation." Acta Mathematicae Applicatae Sinica, English Series: e13164.

[20] Yin, G. George, and Qing Zhang. Continuous-time Markov chains and applications: a two-time-scale approach. 2nd Edition, Springer, New York, 2013.

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

[22] Yin, G. George, and Chao Zhu. Hybrid switching diffusions: properties and applications. Springer, New York, 2010.

ABSTRACT

ON SWITCHING DIFFUSIONS: THE FEYNMAN-KAC FORMULA AND NEAR-OPTIMAL CONTROLS

by

NICHOLAS BARAN

August 2014

Advisor: Dr. George Yin

Major: Mathematics

Degree: Doctor of Philosophy

We consider diffusions in two different contexts. First, we consider the so-called Feynman-Kac formula(s) for switching diffusions. These formulas provide stochastic representations for solutions of certain weakly coupled elliptical systems of partial differential equations. The formulas are for the boundary value problem, the initial value problem, and the initial boundary value problem. Second, we show the existence of near-optimal controls for a system driven by wideband noise in the presence of regime-switching. Using a relaxed control formulation, together with weak convergence methods, we show that given a stochastic optimal control problem, one may find a control that is near-optimal. The use of wideband noise is inspired by applications.

AUTOBIOGRAPHICAL STATEMENT

Nicholas Baran was born in south Florida and moved to the Detroit area in the year 2001. He graduated from Rochester High School, in Rochester Michigan, and began study at Wayne State University in 2005 in the hopes of studying Pharmacy. After realizing that he was only enjoying his math classes, he quickly switched his major to mathematics. He graduated and from Wayne State University with a B.S. in Mathematics, magna cum laude, in 2009 and began his graduate study shortly thereafter. When not studying or teaching math, he can be found listening, playing, or thinking about music.