A Study of Approximate Descriptions of a Random Evolution

by

Henrike Köpke
B.Sc., Technical University of Munich, 2011

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCES

in the Department of Mathematics and Statistics

A Study of Approximate Descriptions of a Random Evolution

by

Henrike Köpke
B.Sc., Technical University of Munich, 2011

Supervisory Committee

Dr. Reinhard Illner, Co-supervisor
(Department of Mathematics and Statistics)

Dr. Anthony Quas, Co-supervisor
(Department of Mathematics and Statistics)

Supervisory Committee

Dr. Reinhard Illner, Co-supervisor
(Department of Mathematics and Statistics)

Dr. Anthony Quas, Co-supervisor
(Department of Mathematics and Statistics)

ABSTRACT

We consider a dynamical system that undergoes frequent random switches according to Markovian laws between different states and where the associated transition rates change with the position of the system. These systems are called random evolutions; in engineering they are also known as stochastic switching systems. Since these kinds of dynamical systems combine deterministic and stochastic features, they are used for modelling in a variety of fields including biology, economics and communication networks. However, to gather information on future states, it is useful to search for alternative descriptions of this system. In this thesis, we present and study a partial differential equation of Fokker-Planck type and a stochastic differential equation that both serve as approximations of a random evolution. Furthermore, we establish a link between the two differential equations and conclude our analysis on the approximations of the random evolution with a numerical case study.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In a variety of fields such as biology, health, telecommunication, politics and economics, we find processes that feature both deterministic and stochastic components. An example for this can be found in the field of mental health. Individuals who suffer from bipolar disorder experience intense mood swings of high (mania) and low (depression) episodes. Manic and depressive states may last from a few days to a few months and can be intercut by periods of “normal” mood. The transitions between the different mood states occur all of a sudden and present major challenges for the affected person.

In this thesis, we consider a type of a dynamical system that features a system of differential equations, whose field undergoes fast and random switches between different states. In particular, the transition rates associated with the system are dependent on its current position. These systems are called random evolutions and have been intensively studied since 1971 [6]. In engineering, they are also known as stochastic switching systems [2] or stochastic hybrid systems.

An example where a random evolution is used for mathematical modelling is the transitional-transcriptional oscillator (TTO) in cell biology. Here, transcription and translation in the cell are presented by systems of ordinary differential equations for both messenger RNA and protein concentrates, but the onset and the termination of transcription can be best modelled as a stochastic process. A precise analysis of this model is explored in [3].

For the prediction of future states, it is useful to find alternative representations of random evolutions. To compute the expectation, the variance and higher order moments of the process, one would need to calculate the combined probability density of the different states of the system, but it is difficult in general to derive information on this probability density directly from the given representation of the system.

In this thesis we explore a variety of alternative descriptions for a particular type of a random evolution. An exact representation of the investigated system will be provided via a set of Kolmogorov master equations [7]. However, this approach can be rather impractical in an environment where switches between the different regimes occur frequently. A suitable approximation of the set of Kolmogorov master equations can be obtained via a Fokker-Planck equation, which has been initiated in [8].

The random evolution can also be approximated via a stochastic differential equation. Since stochastic differential equations are used for models in a randomly behaving environment, they qualify as a suitable approximation to these special kind of dynamical systems.

1.1 Description of the Switching Model

We consider a two-state continuous-time stochastic process $\{I(t), t \geq 0\}$ that alternates between zero and one. The switching process $\{X(t), t \geq 0\}$, dependent on the stochastic process $\{I(t), t \geq 0\}$ is governed by a system of ordinary differential equations:

$$\dot{X}(t) = \begin{cases} f_0(X(t)), & \text{if } I(t) = 0 \\ f_1(X(t)), & \text{if } I(t) = 1 \end{cases}.$$

We observe that the dynamics between consecutive switches is deterministic, but jumps to the other state occur randomly. The transition probabilities associated with $I(t)$ are defined in the following way:

$$P\left[I(t+h) = 1 | I(t) = 0, X(t) = x\right] = \frac{\lambda_0(x)h}{\epsilon} + o(h)$$
$$P\left[I(t+h) = 0 | I(t) = 1, X(t) = x\right] = \frac{\lambda_1(x)h}{\epsilon} + o(h),$$

where $q_{01}(x)$ indicates the probability of a switch from state zero to state one and $q_{10}(x)$ describes how fast a switch from state one to state zero occurs. Since the processes $X(t)$ and $I(t)$ are mutually dependent, we observe that the pair $(I(t), X(t))_{t\geq 0}$ is a Markov process. The transition rates are scaled by a small value of ϵ to obtain a high rate of switchings between the two states. Furthermore, we assume that the transition rates are bounded by the positive constants C_1 and C_2 such that for $j \in \{0, 1\}$:

$$C_1 \leq \lambda_j(x) \leq C_2.$$

A sample path of the evolution of $I(t)$ and the associated switching process $X(t)$ can be found in Figure 1.1.

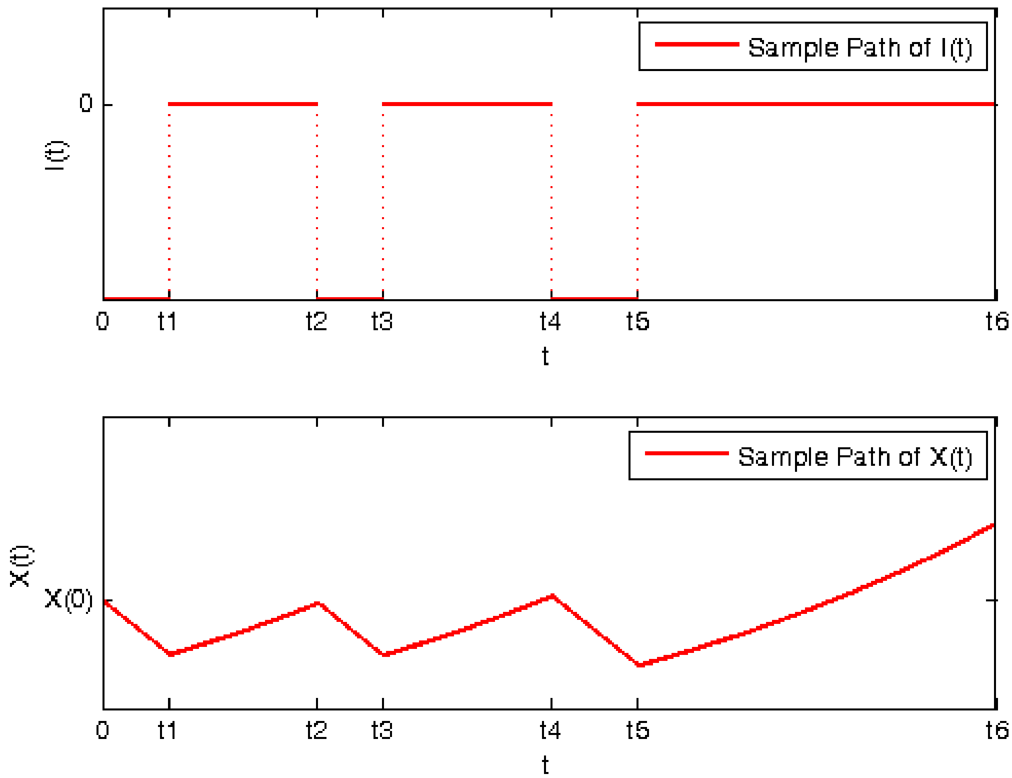

Figure 1.1: Sample path of the evolution of the stochastic process $I(t)$ and its associated switching process $X(t)$

For this particular example we chose:

$$f_0(x) = -1 \text{ and } f_1(x) = \frac{x}{2}.$$

We observe that at the beginning the process $I(t)$ is in state zero and switches to state one at time t_1. At the same time, the switching process evolves with the solution of the ordinary differential equation $f_0(x)$ until t_1. Similarly, between the times t_1 and t_2, the stochastic process $I(t)$ stays in state one, while for the switching process, the right hand side of the differential equation $\dot{x} = f_1(x)$ is solved until time t_2. Given this procedure the switching process continues to evolve up to t_6, where the switches between the two states are governed by $I(t)$. In addition, we define the time it takes for the process $I(t)$ starting in state zero to return to state zero after it had switched to state one as a "cycle".

1.2 Organization of this Thesis

The goal of this thesis is the derivation and testing of approximations of the switching process introduced in Section 1.1 for a small value of ϵ. In Chapter 2, we will present a short introduction to the theory of stochastic differential equations needed for the basic understanding of this work. Before we derive the desired stochastic differential equation, we review in Chapter 3 an alternative description of the switching process via a partial differential equation of Fokker-Planck type. This approach has been studied in [8], where the solution of the partial differential equation presents an approximation to the density of the switching process. In Chapter 4, we focus on the derivation of the desired stochastic differential equation. The solution of the stochastic differential equation serves as an approximation of the distribution of the switching process. The advantage of this approach is that the stochastic differential equation also provides an estimation of the switching differential equation itself. In Chapter 5, we will investigate how the two approximations of the switching process relate to each other. This can be best studied by comparing the two Fokker-Planck equations which result from the two approaches. We will conclude our analysis on the approximations of the switching process in a high rate of switching environment with a numerical case study. Under various parameters we will study the accuracy of the two approximations of the switching process by comparing them to the original switching process.

Chapter 2

Stochastic Differential Equations

Alongside stochastic integrals, stochastic differential equations (SDE) belong to the area of stochastic calculus. The Japanese mathematician Kiyoshi Itô was without a doubt a pioneer in this field. Stochastic differential equations are widely used to model systems with random behaviour. Originally they were developed to describe trajectories of diffusion processes with known drift and diffusion parameters [1]. Nowadays, stochastic differential equations are probably best known for their relevance and application in finance, where they are used to model stock prices and for option pricing models. Moreover, option prices are computed with the help of the Black-Scholes formula, for which Robert C. Merton and Myron S. Scholes were honoured with a Nobel prize in economics in 1997 [10].

2.1 Brownian Motion

Stochastic processes with the probability that given the present state the future is independent of the past are called Markov processes.

Definition 2.1.0.1 (Markov Process)**.** *Suppose that* $\{J(t), t \geq 0\}$ *is a continuous-time stochastic process taking values in the set* $S \subset N$*. The stochastic process* $J(t)$ *is* *said to be a continuous-time Markov chain, if for all* $s, t \geq 0$*:*

$$P\left(J(s+t)=j|J(s)=i, J(u)=j(u), 0 \leq u < s\right) = P\left(J(t+s)=j|J(s)=i\right). \tag{2.1}$$

The property in (2.1) is the Markov property, meaning that future states only depend on the present state [14].

A prominent example of a Markov process is the random walk. The random walk anecdote describes the troubles of a drunken soldier to find his way home at night through the empty town. He either goes one step to the north, the south, the east or the west and since his mind is veiled by alcohol, the soldier forgets after each step that he takes, the direction he came from. Therefore, he is only aware of his present state and decides from there which direction to continue.
The limiting case of a random walk is Brownian motion, where the time steps are of size $\Delta t \to 0$ and the step lengths are of size $\Delta x \to 0$. Brownian motion, named after its discoverer, the English botanist Robert Brown, describes the motion of a small particle in a liquid or a gas [15]. The properties of Brownian motion are summarized below:

Definition 2.1.0.2 (Brownian Motion)**.** *A stochastic process* $\{W(t), t \geq 0\}$ *is called a Brownian motion if it satisfies the following properties:*

1. $W(0) = 0$.
2. $\{W(t), t \geq 0\}$ *has continuous sample paths.*
3. $\{W(t), t \geq 0\}$ *has independent and stationary increments.*
4. $\{W(t), t \geq 0\}$ *is normally distributed with mean* 0 *and variance* $\sigma^2 t$.

We note, that standard Brownian motion is a Brownian motion with variance parameter $\sigma = 1$. In addition, there exists Brownian motion with a drift μ. In this case, the Brownian motion is normally distributed with mean μt and variance $\sigma^2 t$.

Remark 2.1.0.3. *A random variable X is a normal random variable with mean μ and variance σ, $X \sim \Phi(\mu, \sigma^2)$, if it has the probability density function [15]:*

$$f(t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(t-\mu)^2}{2\sigma^2}}$$

Taking a closer look at Definition 2.1.0.2, the third property states that for $0 < t_1 < t_2 < t_3 < ... < t_n$ the respective increments:

$$W_{t_1}, W_{t_2} - W_{t_1}, ..., W_{t_n} - W_{t_{n-1}},$$

are independently distributed random variables.
Furthermore, for $t > s$ the increment $W_t - W_s$ is distributed with:

$$W_t - W_s = W_{t-s} \sim \Phi\left(0, \sigma^2 \left(t - s\right)\right).$$

A graphical example of a one-dimensional Brownian motion is presented in Figure 2.1. The plot presents six sample trajectories of a Brownian motion with no drift, starting at the value $W(0) = 0$. The black line is drawn to represent the distribution of the Brownian motion at time t^*, determined by its transition density $p_{W^{0,0}_{t^*}}(t^*, x)$:

$$p_{W^{0,0}_{t^*}}(t^*, x) = \frac{1}{\sqrt{2\pi t^*}} e^{-\frac{x^2}{2t^*}}.$$

The derivative of Brownian motion dW is described via a white noise process [1]. White noise processes play an important role in the definition of stochastic differential

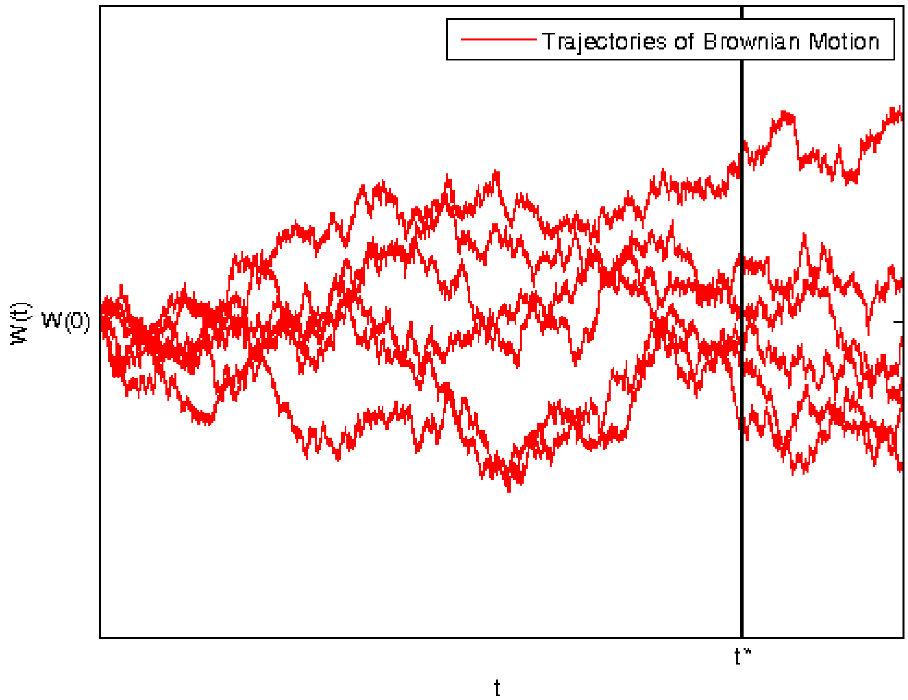

Figure 2.1: Sample trajectories of the one-dimensional Brownian motion

equations.

2.2 Stochastic Differential Equations and Stochastic Integrals

Stochastic differential equations usually consist of a deterministic and a random component. For this let us examine an example of a simplified population growth model [11]. Since the growth of a population over time is subject to random fluctuations, it can be best modelled via a stochastic differential equation.

Let us assume that in the beginning, the size of the population of a certain area is given by $N(0) = N_0$ and at time t the size of a population is $N(t)$. We denote the growth rate of the population with $a(t)$ which is the difference between the birth and death rates.

The evolution of the population over time is given by:

$$\frac{dN}{dt} = a(t)N(t).$$

However, it might be somewhat difficult to determine $a(t)$ since birth and death rates are subject to random fluctuations. Therefore, we represent the growth rate $a(t)$ as follows:

$$a(t) = r(t) + \text{ ``noise'' },$$

where $r(t)$ is a deterministic variable and the noise indicates the random fluctuations that need to be considered in this model. In particular, we observe:

$$\frac{dN}{dt} = (r(t) + \text{ noise }) N(t) = r(t)N(t) + \text{ noise } N(t),$$

which is equivalent to:

$$dN = r(t)N(t)dt + N(t) \underbrace{\text{noise } dt}_{:=dW_t}.$$

The first term in the right hand side of the equation above presents an ordinary differential equation. On the other hand, the term $\frac{dW_t}{dt}$ is represented by a white noise process [1].

The equation above can be written as a stochastic differential equation in the following sense:

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t, \tag{2.2}$$

where $f(t, X_t)$ is the drift parameter and $g(t, X_t)$ is the diffusion parameter of the stochastic differential equation. We can interpret the drift parameter as the slope of the stochastic differential equation, while the diffusion parameter adds the jaggedness to the stochastic differential equation.

The solution of the stochastic differential equation in (2.2) with initial value X_{t_0} is given in the following integral form:

$$X_t = X_{t_0} + \int_{t_0}^{t} f\left(s, X_s\right) ds + \int_{t_0}^{t} g\left(s, X_s\right) dW_s, \tag{2.3}$$

where the integral associated with the diffusion parameter $g(s, X_s)$ is defined by:

$$\int_{t_0}^{t} g(s, X_s) dW_s \equiv \lim_{n\to\infty} \sum_{i=1}^{n} g(t_{i-1}, X_{t_{i-1}}) \left(W_{t_i} - W_{t_{i-1}}\right),$$

with $\max_{i\in\{1,\dots,n\}} \left(t_i - t_{i-1}\right) \to 0$ and $t_0 < t_1 < \dots < t_n = t$. We will not discuss stochastic integrals any further in this thesis; a good introduction to this topic can be found in [1].

2.3 Itô's Lemma and Stochastic Differentiation Rules

Differentiation of stochastic processes follow slightly different differentiation laws than the ones in traditional calculus. For this, let us consider the stochastic differential equation as presented in (2.2) and the stochastic process Y_t, defined by:

$$Y_t = u\left(t, X_t\right).$$

One might assume that the stochastic differential associated with the stochastic process Y_t is computed via:

$$dY_t = u_t\left(t, X_t\right) + u_x\left(t, X_t\right) dX.$$

This, however, is wrong [4]. Therefore, let us take a more careful look at dY_t. Using a formal Taylor expansion on Y_t to compute its stochastic differential, we find:

$$dY = u_t\left(t, X_t\right)dt + u_x\left(t, X_t\right)dX + \frac{1}{2}u_{tt}\left(t, X_t\right)dt^2 + \frac{1}{2}u_{xx}\left(t, X_t\right)\left(dX\right)^2 + \ldots .$$

In addition, we also assume that dt is small enough, so that $dt^{\frac{3}{2}} \approx 0$. With the stochastic differential equation in (2.2), we observe:

$$\begin{aligned} dY =& u_t\left(t, X_t\right)dt + u_x\left(t, X_t\right)\left(f(t, X_t)dt + g(t, X_t)dW_t\right) \\ &+ \frac{1}{2}u_{xx}\left(t, X_t\right)\left(f(t, X_t)dt + g(t, X_t)dW_t\right)^2 + \ldots . \end{aligned}$$

As $dW_t = \left(W_{t+dt} - W_t\right) \sim \Phi\left(0, dt\right)$, we conclude $dW_t \approx \sqrt{dt}$. Therefore, we will replace the term $\left(dW_t\right)^2$ by dt. However, one needs to keep in mind that dW_t is random, but dt is deterministic. By collecting all terms smaller than those of order $O(dt^{\frac{3}{2}})$ we find:

$$\begin{aligned} dY =& u_t\left(t, X_t\right)dt + u_x\left(t, X_t\right)f(t, X_t)dt + u_x\left(t, X_t\right)g(t, X_t)dW_t \\ &+ \frac{1}{2}u_{xx}\left(t, X_t\right)\left(g(t, X_t)\right)^2 dt + O\left(dt^{\frac{3}{2}}\right). \end{aligned}$$

By neglecting all terms of size $O\left(dt^{\frac{3}{2}}\right)$ and higher, we have:

$$dY = \left(u_t\left(t, X_t\right) + u_x\left(t, X_t\right)f(t, X_t) + \frac{1}{2}u_{xx}\left(t, X_t\right)\left(g(t, X_t)\right)^2\right)dt + u_x\left(t, X_t\right)g(t, X_t)dW_t.$$

One may observe from these computations that the additional term

$$\frac{1}{2}u_{xx}\left(t, X_t\right)\left(g(t, X_t)\right)^2 dt$$

arises in the differential dY_t. We thus conclude that the differentials of stochastic processes follow their own differentiation laws. Itô's lemma can be regarded as a chain rule of stochastic processes [1]:

Theorem 2.3.0.4 (Itô's Lemma)**.** *Let $u = u(t,x)$ be a continuous function defined on $[t_0, T] \times R^d$ with the continuous partial derivatives: $\frac{\partial}{\partial t}u(t,x) = u_t$, $\frac{\partial}{\partial x_i}u(t,x) = u_{x_i}$ and $\frac{\partial^2}{\partial x_i \partial x_j}u(t,x) = u_{x_i x_j}$. If the stochastic differential of the stochastic process X_t is defined on $[t_0, T]$ by:*

$$dX_t = f(t, X_t)\, dt + g(t, X_t)\, dW_t,$$

the stochastic process $Y_t = u(t, X_t)$ defined on $[t_0, T]$ with the initial value $Y_{t_0} = u(0, X_{t_0})$ is associated with the stochastic differential:

$$dY_t = \left(u_t(t, X_t) + u_x(t, X_t) f(t, X_t) + \frac{1}{2} u_{xx}(t, X_t) \left(g(t, X_t) \right)^2 \right) dt + u_x(t, X_t) g(t, X_t) dW_t.$$

The formal proof of Itô's lemma is quite technical and is omitted here. The proof can be found in [1] or in any other stochastic calculus book.

Remark 2.3.0.5. *Assuming that dt is small enough, so that $dt^{\frac{3}{2}} \approx 0$, the terms dt^2, dW^2 and $dt \cdot dW$ in Itô's lemma are computed according to the following multiplication table:*

$\bullet$	dt	dW
dt	0	0
dW	0	dt

Table 2.1: Multiplication table for Itô's Lemma

2.4 Existence and Uniqueness of the Solution of a Stochastic Differential Equation

Under certain conditions one can prove that the stochastic differential equation in (2.2) has a unique solution which satisfies the stochastic integral:

$$X_t = X_{t_0} + \int_{t_0}^{t} f\left(s, X_s\right) ds + \int_{t_0}^{t} g\left(s, X_s\right) dW_s.$$

Ordinary differential equations are a special case of stochastic differential equations with $g(t, X_t) \equiv 0$.

According to the Picard-Lindelöf theorem, a unique and global solution of the ordinary differential equation:

$$\dot{x}(t) = f\left(t, x(t)\right),$$

with initial condition $x(t_0) = x_0$ exists in $[t_0, T] \times R$, if the function $f\left(t, x(t)\right)$ is Lipschitz continuous and bounded in $[t_0, T] \times R$. For this, we have a look at two deterministic examples:

- Let us consider the ordinary differential equation with initial value $x(0) = 0$:

$$dx = 3x^{\frac{2}{3}} dt.$$

  The function $f(x) = 3x^{\frac{2}{3}}$ is not Lipschitz continuous for all intervals which include the value $x = 0$. The solutions of this differential equation on $[0, T]$ for any $c > 0$ are given by:

$$x(t) = \begin{cases} 0 & , \text{if } 0 \leq t \leq c \\ (t-c)^3 & , \text{if } c < t \leq T \end{cases}.$$

- Let us consider the ordinary differential equation with initial value $x(0) = 1$:

$$dx = x^2 dt.$$

Since the function $f(x) = x^2$ is Lipschitz continuous in a bounded interval, the unique solution of the differential equation, for $t \in [0, 1)$, is given by:

$$x(t) = \frac{1}{1-t}.$$

However, with $t \to 1$, we observe that $x(t) \to \infty$. In order to obtain a global solution of the differential equation one would need to consider additional restrictions on the growth of $f(x) = x^2$.

The theory of ordinary differential equations can be extended to stochastic differential equations.

Theorem 2.4.0.6 (Existence-and-Uniqueness Theorem of a Stochastic Differential Equation)**.** *Consider the stochastic differential as defined in (2.2) with initial value* $X(t_0) = X_{t_0}$ *and* $t_0 \leq t \leq T < \infty$*, where* W_t *is a Brownian motion process and* X_{t_0} *is a random variable independent of* $(W_t - W_{t_0})$*. We also assume that the drift function* $f(t, x)$ *and the diffusion function* $g(t, x)$ *are defined and measurable on* $[t_0, T]$ *and satisfy the following properties:*

1. ***Lipschitz condition:*** *There exists a positive constant* $C > 0$ *so that for all* $x, y \in R$ *and* $t \in [t_0, T]$*, the drift* $f(t, x)$ *and diffusion* $g(t, x)$ *satisfy:*

$$|f(t,x) - f(t,y)| + |g(t,x) - g(t,y)| \leq C\,|x-y|\,. \tag{2.4}$$

2. ***Restriction on growth:*** *There exists a positive constant* $C > 0$*, such that for*

all $x \in R$ and $t \in [t_0, T]$:

$$|f(t,x)|^2 + |g(t,x)|^2 \leq C^2 \left(1 + |x|^2\right). \tag{2.5}$$

Then the stochastic differential equation has a unique solution X_t on $[t_0, T]$ that is almost surely continuous and which satisfies the initial condition $X(t_0) = X_{t_0}$.

The complete proof of Theorem 2.4.0.6 can be found in [1]. Note that these existence and uniqueness conditions are sufficient, but not necessary.

2.5 The Solution of a Stochastic Differential Equation as a Markov and as a Diffusion Process

The solution of a stochastic differential equation is a stochastic process on the interval $[t_0, T]$ that can be regarded as a set of distributions [1]:

$$P\left[X_{t_1} \in B_1, ..., X_{t_n} \in B_n\right] = P_{t_1,...,t_n}(B_1, ..., B_n).$$

The advantage of a Markov process is that we can obtain $P\left[X_{t_1} \in B_1, ..., X_{t_n} \in B_n\right]$ from its initial distribution:

$$P\left[X_{t_0} \in B\right] = P_{t_0}(B)$$

and its transition probability for $t_0 \leq s \leq t \leq T$:

$$P\left(X_t \in B | X_s = x_s\right) = P(s, x, t, B).$$

Therefore it might be useful to analyze under which conditions a stochastic differential equation is a Markov process.
For this we consider the stochastic differential as defined in (2.2). In addition, we consider the same differential equation on the interval $[s, T]$ with initial value $X_s = x_s$. Hence its equivalent integral form for $t_0 \leq s \leq t \leq T$ is given by:

$$X_t = x_s + \int_s^t f(u, X_u)\, du + \int_s^t g(u, X_u)\, dW_u. \tag{2.6}$$

Theorem 2.5.0.7 (The Solution of a Stochastic Differential Equation as a Markov Process)**.** *We assume that the stochastic differential equation as defined in (2.2) satisfies the existence and uniqueness conditions of Theorem 2.4.0.6. Then the solution of the stochastic differential equation X_t for arbitrary initial values is a Markov process on $[t_0, T]$ with the initial distribution:*

$$P\left[X_{t_0} \in B\right] = P_{t_0}(B)$$

and whose transition probabilities are given by:

$$P(s, x, t, B) = P\left(X_t \in B | X_s = x_s\right) = P\left[X_t(s, x) \in B\right],$$

where $X_t(s, x)$ is the unique solution of (2.6).

Therefore we observe that if a stochastic differential equation has a unique solution X_t, the stochastic process X_t is necessarily a Markov process.
Diffusion processes are a special class of Markov processes with continuous sample paths that describe the random displacement of a particle in a fluid or a gas. The most prominent example of a diffusion process is Brownian motion. A diffusion process is defined as follows [9]:

Definition 2.5.0.8 (Diffusion Process)**.** *A Markov process with transition probabilities $P(s,x;t,B)$ is called a diffusion process, if the following three limits exists for all $\epsilon > 0$, $s \geq 0$ and $x \in R$:*

$$\lim_{t\to s}\frac{1}{t-s}\int_{|y-x|>\epsilon} p(s,x;t,y)dy = 0 \tag{2.7}$$

$$\lim_{t\to s}\frac{1}{t-s}\int_{|y-x|<\epsilon} (y-x)p(s,x;t,y)dy = f(s,x) \tag{2.8}$$

$$\lim_{t\to s}\frac{1}{t-s}\int_{|y-x|<\epsilon} (y-x)^2 p(s,x;t,y)dy = (g(s,x))^2. \tag{2.9}$$

In this case, $f(s,x)$ and $g(s,x)$ are called the drift and the diffusion parameter.

In particular, Condition (2.7) prevents the process from having sudden large jumps. With the definition of the expected value, Condition (2.8) implies:

$$f(s,x) = \lim_{t\to s}\frac{1}{t-s}E\left(X_t - X_s | X_s = x_s\right)$$

and similarly, Condition (2.9) implies:

$$(g(s,x))^2 = \lim_{t\to s}\frac{1}{t-s}E\left((X_t - X_s)^2 | X_s = x_s\right).$$

The drift parameter presents the instantaneous rate of change in the mean of X_t, given that $X_s = x_s$, whereas the squared diffusion measures the instantaneous rate of change of the squared fluctuations of X_t, given that $X_s = x_s$.

Additionally, the transition density $p(s,x;t,y)$ of a diffusion process satisfies the two partial differential equations:

$$\partial_t p(s,x;t,y) + \partial_y\left\{f(t,y)\cdot p(s,x;t,y)\right\} - \frac{1}{2}\partial_{yy}\left\{(g(t,y))^2\cdot p(s,x;t,y)\right\} = 0 \tag{2.10}$$

$$\partial_s p(s,x;t,y) + f(s,x)\cdot\partial_x p(s,x;t,y) + \frac{1}{2}(g(s,x))^2\cdot\partial_{xx} p(s,x;t,y) = 0. \tag{2.11}$$

In (2.10) we assume that the pair of variables (s, x) is fixed. This partial differential equation is called Kolmogorov forward equation or Fokker-Planck equation. The partial differential equation in (2.11) is called Kolmogorov backward equation, where we assume that the pair of the variables (t, y) is fixed.
Since diffusion processes are Markov processes with continuous sample paths, one might suppose that under additional conditions on the drift and the diffusion parameter, the solution of a stochastic differential equation is a diffusion process.

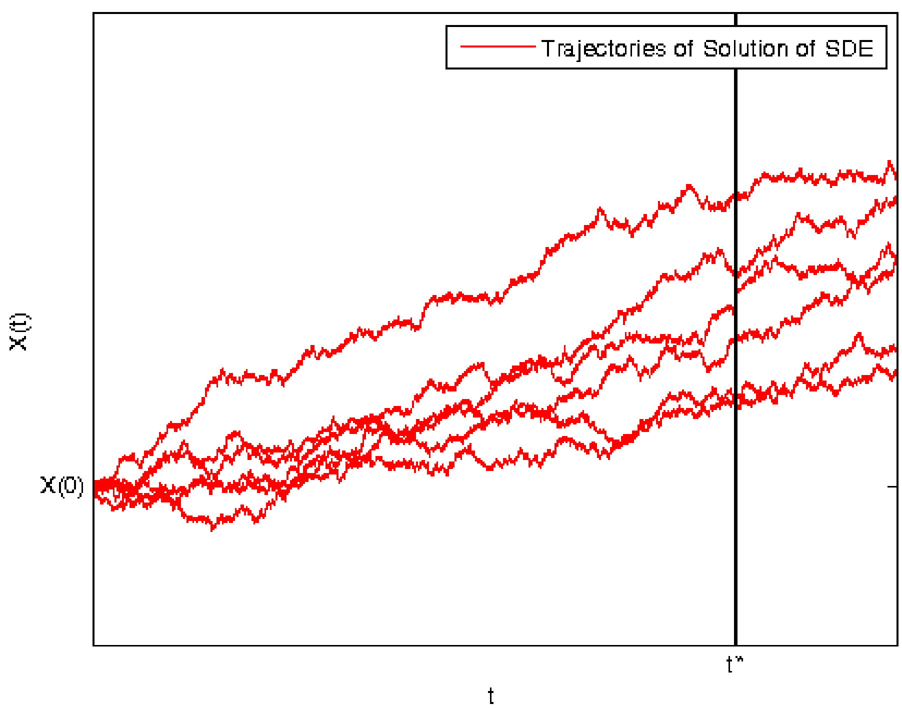

Figure 2.2: Sample trajectory of the solution of the stochastic differential equation

Theorem 2.5.0.9 (The Solution of a Stochastic Differential Equation as a Diffusion Process)**.** *We assume that the conditions of the existence and uniqueness theorem (Theorem 2.4.0.6) hold for the stochastic differential equation:*

$$dX_t = f(t, X_t)dt + g(t, X_t)dW_t,$$

with the initial condition $X_{t_0} = x_{t_0}$ for $t_0 \leq t \leq T$. If the drift parameter $f(t, X_t)$ and the diffusion parameter $g(t, X_t)$ of the stochastic differential equation are continuous, the solution of the stochastic differential equation for any fixed initial value x_{t_0} is a diffusion process on $[t_0, T]$ with drift $f(t, x)$ and diffusion $g(t, x)$.

A proof of this theorem can be found in [9]. Moreover, if the solution of a stochastic differential equation is a diffusion process, its transition density $p(s, x; t, y)$ satisfies the Fokker-Planck equation as defined in (2.10) and the Kolmogorov backward equation given in (2.11).

The sample paths of the trajectories of the solution of a stochastic differential equaltion can be found in Figure 2.2. The plot shows six sample trajectories of the stochastic differential equation starting at $X(0) = 1$ with drift $f = 5x$ and constant diffusion parameter $g = \frac{1}{2}$ up to the time $T = 0.01$. Since the solution of the stochastic differential is a diffusion process, we obtain its transition density $p(s, x; t^*, y^*)$ by solving the Kolmogorov backward equation or the Fokker-Planck equation up to time t^*.

The solution of a stochastic differential equation is governed by its drift and diffusion parameters. As mentioned in Section 2.2, we can interpret the drift parameter f as the “slope” of the stochastic differential equation. The diffusion term g adds the random fluctuations to the behaviour of the stochastic differential equation.

2.6 The Relation between Stochastic Differential Equations and the Fokker-Planck Equation

The Fokker-Planck equation was first discovered by the Dutch physician Adriaan D. Fokker in 1914 and later (in 1917) by the German physicist Max Planck to describe the motion of a particle in a fluid or in a gas [13]. From the previous section, we observe that the transition density p of the solution of a stochastic differential equation with continuous drift and diffusion parameters satisfies the Fokker-Planck equation. The equivalence of a stochastic differential equation and the Fokker-Planck equation can be shown with the help of Itô's lemma.

For this, recall that $p = p(s, x_s; t, x)$ is the transition density of the solution of the stochastic differential equation given $X(s) = x_s$, as introduced in Section 2.5. Furthermore, the associated transition distribution can be obtained from:

$$P\left(X_t \in \Omega | X_s = x_s\right) = \int_\Omega p(s, x_s; t, x)dx.$$

Let us assume that u is an arbitrary twice differentiable function, so that the application of Itô's lemma to the stochastic process $u(X_t)$ provides:

$$\begin{aligned} u(X_t) =& u(X_0) + \int_0^t \left(u_x(X_u)f(X_u) + \frac{1}{2}u_{xx}(X_u)(g(X_u))^2 \right) du \\ &+ \int_0^t u_x(X_u)g(X_u)dW_u. \end{aligned}$$

By taking the expected value of $u(X_t)$ and differentiating it with respect to t, we find:

$$\begin{aligned} \frac{d}{dt}\left\{\mathrm{E}\left[u(X_t)\right]\right\} =& \frac{d}{dt}\left\{\mathrm{E}\left[\int_0^t \left(u_x(X_u)f(X_u) + \frac{1}{2}u_{xx}(X_u)(g(X_u))^2 \right) du\right]\right\} \\ &+ \frac{d}{dt}\left\{\mathrm{E}\left[\int_0^t u_x(X_u)g(X_u)dW_u\right]\right\}. \end{aligned}$$

Moreover, by taking the derivative inside of the expectation, we compute for:

$$\begin{aligned} &\frac{d}{dt}\left\{\mathrm{E}\left[\int_0^t \left(u_x(X_u)f(X_u) + \frac{1}{2}u_{xx}(X_u)(g(X_u))^2 \right) du\right]\right\} \\ &= \mathrm{E}\left[\frac{d}{dt}\left\{\int_0^t \left(u_x(X_u)f(X_u) + \frac{1}{2}u_{xx}(X_u)(g(X_u))^2 \right) du\right\}\right] \\ &= \mathrm{E}\left[u_x(X_t)f(X_t) + \frac{1}{2}u_{xx}(X_t)(g(X_t))^2\right]. \end{aligned}$$

Similarly, by taking the derivative inside the expectation in $\frac{d}{dt}\left\{\mathrm{E}\left[\int_0^t u_x(X_u)g(X_u)dW_u\right]\right\}$, we find:

$$\begin{aligned}\frac{d}{dt}\left\{\mathrm{E}\left[\int_0^t u_x(X_u)g(X_u)dW_u\right]\right\} &= \mathrm{E}\left[\frac{d}{dt}\left\{\int_0^t u_x(X_u)g(X_u)dW_u\right\}\right] \\ &= \mathrm{E}\left[\lim_{h\to 0}\int_t^{t+h}\frac{u_x(X_u)g(X_u)}{h}dW_u\right].\end{aligned}$$

Using the definition of the stochastic integral as presented in Section 2.2, we obtain:

$$\mathrm{E}\left[\lim_{h\to 0}\int_t^{t+h}\frac{u_x(X_u)g(X_u)}{h}dW_u\right] = \mathrm{E}\left[u_x(X_t)g(X_t)\left(\lim_{h\to 0}\frac{W_{t+h}-W_t}{h}\right)\right].$$

By conditioning on $X_t = y$, we further observe:

$$\begin{aligned}\mathrm{E}\left[u_x(X_t)g(X_t)\frac{W_{t+h}-W_t}{h}\right] =&\mathrm{E}\left[\mathrm{E}\left[u_x(X_t)g(X_t)\left(\lim_{h\to 0}\frac{W_{t+h}-W_t}{h}\right)\Big|X_t = y\right]\right] \\ =&\mathrm{E}\left[u_x(y)g(y)\underbrace{\mathrm{E}\left[\left(\lim_{h\to 0}\frac{W_{t+h}-W_t}{h}\right)\Big|X_t = y\right]}_{:=0}\right] = 0.\end{aligned}$$

Combining our previous steps, we conclude:

$$\frac{d}{dt}\left\{\mathrm{E}\left[u(X_t)\right]\right\} = E\left[u_x(X_t)f(X_t) + \frac{1}{2}u_{xx}(X_t)(g(X_t))^2\right].$$

In addition, we assume that the transition density $p(s, x_s; t, x)$ of the solution of the stochastic differential equation X_t satisfies:

$$\lim_{x\to\pm\infty} p(s, x_s; t, x) = 0 \text{ and } \lim_{x\to\pm\infty} \partial_x\left\{p(s, x_s; t, x)\right\} = 0.$$

With

$$\partial_t \left\{ \mathrm{E}\left[u(X_t) \right] \right\} = \partial_t \left\{ \int_{-\infty}^{\infty} u(x) p(s, x_s; t, x) dx \right\} = \int_{-\infty}^{\infty} u(x) p_t(s, x_s, ; t, x) dx$$

and

$$\mathrm{E}\left[u_x(X_t) f(X_t) + \frac{1}{2} u_{xx}(X_t) (g(X_t))^2 \right] = \int_{-\infty}^{\infty} \left(u'(x) f(x) + \frac{1}{2} u''(x) (g(x))^2 \right) p(s, x_s; t, x) dx,$$

we observe:

$$\int_{-\infty}^{\infty} u(x) \partial_t p(s, x_s; t, x) dx = \int_{-\infty}^{\infty} \left(u'(x) f(x) + \frac{1}{2} u''(x) (g(x))^2 \right) p(s, x_s; t, x) dx.$$

Integration by parts of the expression above gives:

$$\begin{aligned} \int_{-\infty}^{\infty} u(x) \partial_t p(s, x_s; t, x) dx = & \underbrace{u(x) f(x) p(s, x_s; t, x) \Big|_{-\infty}^{\infty}}_{\to 0} - \int_{-\infty}^{\infty} u(x) \partial_x \left\{ f(x) p(s, x_s; t, x) \right\} dx \\ & + \underbrace{\frac{1}{2} u'(x) (g(x))^2 p(s, x_s; t, x) \Big|_{-\infty}^{\infty}}_{\to 0} \\ & - \frac{1}{2} \int_{-\infty}^{\infty} u'(x) \partial_x \left\{ (g(x))^2 p(s, x_s; t, x) \right\} dx. \end{aligned}$$

Further integration of the integral $\frac{1}{2} \int_{-\infty}^{\infty} u'(x) \partial_x \left\{ (g(x))^2 p(s, x_s; t, x) \right\} dx$ yields:

$$\begin{aligned} \int_{-\infty}^{\infty} u(x) \partial_t p(s, x_s; t, x) dx = & - \int_{-\infty}^{\infty} u(x) \partial_x \left\{ f(x) p(s, x_s; t, x) \right\} dx \\ & - \underbrace{\frac{1}{2} u(x) \partial_x \left\{ (g(x))^2 p(s, x_s; t, x) \right\} \Big|_{-\infty}^{\infty}}_{\to 0} \\ & + \frac{1}{2} \int_{-\infty}^{\infty} u(x) \partial_{xx} \left\{ (g(x))^2 p(s, x_s; t, x) \right\} dx. \end{aligned}$$

Since u is an arbitrary twice differentiable function, we can therefore deduce that p satisfies the Fokker-Planck equation [5]:

$$\partial_t p + \partial_x \{f(x)p\} - \frac{1}{2}\partial_{xx} \left\{(g(x))^2 p\right\} = 0.$$

The connection between the stochastic differential equations and the Fokker-Planck equation will become more important in Chapter 5, as we analyze how the two derived approximations of the stochastic switching process $X(t)$ relate to each other.

Chapter 3

Approach via Partial Differential Equations

3.1 The Setting

We return to the mathematical model of the system of switched ordinary differential equations as introduced in Section 1.1. If $\{I(t), t \geq 0\}$ is a stochastic process that alternates between the states 0 and 1, the evolution of the switching process $\{X(t), t \geq 0\}$ dependent on $I(t)$, is described by:

$$\dot{X}(t) = \begin{cases} f_0(X(t)) & \text{if } I(t) = 0 \\ f_1(X(t)) & \text{if } I(t) = 1, \end{cases}$$

where we assume that f_0 and f_1 are smooth functions. In particular, if $I(t) = 0$ we say the system is in state zero and if $I(t) = 1$ the system is said to be in state one. We also assume that the transition rates associated with the stochastic process $I(t)$

are given by:

$$q_{01}(\xi) := \lim_{h \to 0} \left(\frac{P(I(t+h) = 1 | I(t) = 0, X(t) \in \Omega)}{h} \right) = \frac{\lambda_0(\xi) + O(\Delta\xi)}{\epsilon}$$
$$q_{10}(\xi) := \lim_{h \to 0} \left(\frac{P(I(t+h) = 0 | I(t) = 1, X(t) \in \Omega)}{h} \right) = \frac{\lambda_1(\xi) + O(\Delta\xi)}{\epsilon},$$

with $\Omega = [\xi, \xi + \Delta\xi)$. Thus, we observe that $q_{01}(\xi)$ and $q_{10}(\xi)$ are dependent on the position of $X(t)$ and are scaled by a sufficiently small parameter $\epsilon > 0$ to ensure a high volume of switches between the two respective states. From these transition rates we deduce the respective transition probabilities:

$$P(I(t+h) = 1 | I(t) = 0, X(t) \in \Omega) = \frac{1}{\epsilon}(\lambda_0(\xi) + O(\Delta\xi))h + o(h) \tag{3.1}$$
$$P(I(t+h) = 0 | I(t) = 1, X(t) \in \Omega) = \frac{1}{\epsilon}(\lambda_1(\xi) + O(\Delta\xi))h + o(h), \tag{3.2}$$

where (3.1) indicates the probability that a switch from state zero to state one occurs and (3.2) gives the probability that the system switches from state one to state zero within $[t, t+h]$.

In addition, we note that the pair of stochastic processes $(I(t), X(t), t \geq 0)_{t \geq 0}$ is a Markov process. Thus $I(t)$ and $X(t)$ are mutually dependent, but $(I(t), X(t), t \geq 0)_{t \geq 0}$ is a Markov process.

Remark 3.1.0.10. *The conditional probabilities for $\Omega = [\xi, \xi + \Delta\xi)$, are defined with $j \in \{0, 1\}$ by:*

$$P(I(t+h) = j | I(t) = 1-j, X(t) \in \Omega) = \frac{P(I(t+h) = j, I(t) = 1-j, X(t) \in \Omega)}{P(I(t) = 1-j, X(t) \in \Omega)}.$$

The aim of this chapter is to derive a partial differential equation whose solution gives an accurate approximation of the switching process $X(t)$ for small values of the scaling parameter ϵ. For this, we will first derive a set of Kolmogorov master equations that presents an exact description of $X(t)$. The master equations, which are commonly used in physics and other related fields, determine the time evolution of an individual state of a system of differential equations where switches between the respective states occur randomly. The solutions of the set of Kolmogorov master equations will give the respective probability densities that the switching process $X(t)$ is in state zero and in state one at the fixed time T.

However, in the end we would like to observe the effects of the switches on $X(t)$, rather than determining in which state the system is at the time T. In addition, we will also observe that the Kolmogorov master equations can be quite difficult to work with for small values of ϵ. Therefore, we will combine the two Kolmogorov master equations and apply an asymptotic expansion (Section 3.2) to derive a suitable approximation of the switching process $X(t)$.

For the analytic derivations throughout this chapter, we assume:

- There exist sufficiently smooth probability densities $p_j(x,t)$ for $j \in \{0,1\}$

$$P(I(t) = j, X(t) \in \Omega) = \int_\Omega p_j(x,t)dx,$$

  with $\Omega = [\xi, \xi + \Delta\xi)$.

- The smooth transition probabilities are defined in (3.1) and (3.2). We assume that the functions $\lambda_0(x)$ and $\lambda_1(x)$ that are associated with the transition rates are bounded by the positive constants C_1 and C_2, such that for $j \in \{0,1\}$:

$$C_1 \leq \lambda_j(\xi) \leq C_2.$$

- The solution operators of the system of differential equations:

$$\dot{X}(t) = f_j(X(t)), \tag{3.3}$$

for $j \in \{0, 1\}$ with the respective initial conditions $T_j(0) = I$ are given by:

$$\frac{d}{dt}T_j(t)(x) = (f_j)(T_j(t)(x)).$$

Furthermore, the unique solutions of the differential equation in (3.3) for $j \in \{0, 1\}$ are given by:

$$X(t) = T_j(t)x_0,$$

where $X(0) = x_0$ presents the initial value of the switching system.

3.2 The Kolmogorov Master Equation

The Kolmogorov master equations study the evolution of the probability densities associated with the solutions of the switching system. In particular, we study $p_0(x, t)$ and $p_1(x, t)$ which describe the respective densities that the switching system is at a certain time t in the state zero or the state one.

Theorem 3.2.0.11. *Consider the system of switching differential equations:*

$$\dot{X}(t) = \begin{cases} f_0(X(t)) & \text{if } I(t) = 0 \\ f_1(X(t)) & \text{if } I(t) = 1 \end{cases},$$

which evolves according to the smooth functions f_0 and f_1, dependent on the state of the Markov chain $\{I(t), t \geq 0\}$. The transition probabilities associated with the switching process $X(t)$ are presented in (3.1) and (3.2). Then, the system of switching *differential equations is described via the set of Kolmogorov master equations:*

$$\partial_t p_0 + \partial_x \left\{ f_0 p_0 \right\} = \frac{1}{\epsilon} \left(\lambda_1 p_1 - \lambda_0 p_0 \right)$$
$$\partial_t p_1 + \partial_x \left\{ f_1 p_1 \right\} = \frac{1}{\epsilon} \left(\lambda_0 p_0 - \lambda_1 p_1 \right).$$

Proof. To prove this assertion we will focus on the calculation of the Kolmogorov master equation associated with p_1, since the derivation of the Kolmogorov master equation associated with p_0 follows the same principle. For this, we first compute:

$$P\left(X(t+h) \in \Omega, I(t+h) = 1 \right),$$

which indicates the probability distribution that the system is in state one at time $t+h$. From this probability distribution we will deduce the probability density $p_1(\xi, t+h)$, whose derivative gives the desired Kolmogorov master equation associated with p_1.

Step 1:

For the calculation of $P\left(X(t+h) \in \Omega, I(t+h) = 1 \right)$, let us assume that $\Omega = [\xi, \xi + \Delta\xi)$. We define the probabilities P_{01} and P_{11} by:

$$P_{01} := P(X(t+h) \in \Omega, I(t+h) = 1, I(t) = 0)$$
$$P_{11} := P(X(t+h) \in \Omega, I(t+h) = 1, I(t) = 1).$$

We further obtain by adding the probabilities P_{01} and P_{10}:

$$P(X(t+h) \in \Omega, I(t+h) = 1) = P_{01} + P_{11}. \tag{3.4}$$

This equality holds, since the switching system can only be either in state zero or in state one at time t. The probability P_{11} indicates that either no switch to state zero or an even number of switches occur in the time interval $[t, t+h]$. The latter case is dismissed, since the probability of this event is of size $O(h^2)$. Moreover, we can show that for $j \in \{0, 1\}$:

$$P(X(t+h) \in \Omega, I(t) = j) = P(X(t+h) \in \Omega, I(t+h) = j) + O(h). \tag{3.5}$$

For this, we express (3.4) with the associated conditional probabilities as follows:

$$\begin{aligned}
&P(X(t+h) \in \Omega, I(t+h) = j) \\
&= P(I(t+h) = j | I(t) = j, X(t+h) \in \Omega) \cdot P(X(t+h) \in \Omega, I(t) = j) \\
&+ P(I(t+h) = j | I(t) = 1-j, X(t+h) \in \Omega) \cdot P(X(t+h) \in \Omega, I(t) = 1-j).
\end{aligned}$$

We also observe that the transition probability as defined in (3.1) is of size $O\left(h\right)$. We conclude then:

$$\begin{aligned}
&P(X(t+h) \in \Omega, I(t+h) = j) \\
&= P(I(t+h) = j | X(t+h) \in \Omega, I(t) = j) \cdot P(X(t+h) \in \Omega, I(t) = j) + O(h).
\end{aligned}$$

On the other hand from the transition probabilities of $I(t)$, we find:

$$P(I(t+h) = j | X(t+h) \in \Omega, I(t) = j) = 1 - O(h).$$

Hence, by combining the results from above and rearranging of terms we obtain (3.5). We will use this result to further calculate the probabilities P_{01} and P_{10}.

By expressing P_{01} with conditional probabilities we find:

$$P_{01} = P(X(t+h) \in \Omega, I(t) = 0)P(I(t+h) = 1|X(t+h) \in \Omega, I(t) = 0).$$

In addition with the result of (3.5), we have:

$$P_{01} = (P(X(t+h) \in \Omega, I(t+h) = 0) + O(h)) \cdot P(I(t+h) = 1|X(t+h) \in \Omega, I(t) = 0).$$

Therefore, with the assumptions made on the probability density and with the transition probability as defined in (3.2), we conclude:

$$P_{01} = \left(\int_\Omega p_0(x, t+h)dx + O(h) \right) \cdot \left(\frac{1}{\epsilon}(\lambda_0(\xi) + O(\Delta\xi))h + o(h) \right).$$

Using conditional probabilities for the representation of P_{11}, gives:

$$P_{11} = P(X(t+h) \in \Omega, I(t) = 1)P(I(t+h) = 1|X(t+h) \in \Omega, I(t) = 1).$$

Recall that P_{11} indicates that no switches occur between the times t and $t+h$ (since we ignore the case that more than one switch occurs in $[t, t+h]$). The dynamics of the time interval $[t, t+h]$ are therefore given by T_1 and from Section 3.1 we recall that the unique solution of $X(t+h)$ is: $X(t+h) = T_1(h)X(t)$. With the properties of T_1, we have that $X(t+h) \in \Omega$ is equivalent to $X(t) \in T_1(-h)\Omega$. Consequently:

$$P_{11} = P(X(t) \in T_1(-h)\Omega, I(t) = 1)P(I(t+h) = 1|X(t) \in T_1(-h)\Omega, I(t) = 1).$$

With $P(I(t+h)=1|X(t)\in T_1(-h)\Omega, I(t)=1) = 1-P(I(t+h)=0|X(t)\in T_1(-h)\Omega, I(t)=1)$ and the assumption on the transition probability in (3.2), we rewrite the expression above as:

$$P_{11} = \left(\int_{T_1(-h)\Omega} p_1(x,t)dx\right)\cdot\left(1-\frac{1}{\epsilon}\lambda_1\left(T_1(-h)\xi+O(\Delta\xi)\right)h+o(h)\right).$$

We now aim to find a suitable substitution of the integral:

$$\int_{T_1(-h)\Omega} p_1(x,t)dx.$$

By setting $x = T_1(-h)z$, we observe for the differential equation in (3.3): $\dot{T}_1(-h)z = -f_1(T_1(-h)z)$. Including a Taylor expansion on $f_1(T_1(-s)z)$, we further obtain:

$$\begin{aligned} T_1(-h)z &= z-\int_0^h f_1(T_1(-s)z)ds \approx z-\int_0^h \left(f_1(z)+Df_1(z)(-s)\right)ds \\ &= z-hf_1(z)+o(h). \end{aligned}$$

Since $dx = (1-h\partial_z f_1(z))dz$, our chosen substitution yields:

$$\int_{T_1(-h)\Omega} p_1(x,t)dx = \int_\Omega p_1(z-hf_1(z),t)(1-h\partial_z f_1(z))dz + o(h).$$

On the other hand, since $T_1(-h)\xi = \xi + O\left(h\right)$, we find:

$$P_{11} = \left(\int_{T_1(-h)\Omega} p_1(x,t)dx\right)\cdot\left(1-\frac{1}{\epsilon}\left(\lambda_1\left(\xi\right)+O(\Delta\xi)\right)h+o(h)\right).$$

Combining our previous steps and the results for P_{01}, we conclude:

$$
\begin{aligned}
&P(X(t+h)\in\Omega, I(t+h)=1)\\
&=\left(\int_\Omega p_0(x,t+h)dx+O(h)\right)\cdot\left(\frac{1}{\epsilon}(\lambda_0(\xi)+O(\Delta\xi))h+o(h)\right)\\
&+\left(\int_\Omega p_1(z-hf_1(z),t)(1-h\partial_z f_1(z))dz\right)\cdot\left(1-\frac{1}{\epsilon}\left(\lambda_1\left(\xi\right)+O(\Delta\xi)\right)h+o(h)\right).
\end{aligned}
$$

Step 2:

We now focus on the computation of the density $p_1(x,t+h)$ via differentiation of $P(X(t+h)\in\Omega, I(t+h)=1)$. Further along and given the result of $p_1(x,t+h)$, we will derive the associated Kolmogorov master equation associated with state one. From the assumptions on the probability densities, we observe:

$$
P(X(t+h)\in\Omega, I(t+h)=1)=\int_\Omega p_1(x,t+h)dx.
$$

By combining the results from above, we obtain:

$$
\begin{aligned}
&\int_\Omega p_1(x,t+h)dx\\
&=\left(\int_\Omega p_0(x,t+h)dx+O(h)\right)\cdot\left(\frac{1}{\epsilon}(\lambda_0(\xi)+O(\Delta\xi))h+o(h)\right)\\
&+\left(\int_\Omega p_1(z-hf_1(z),t)(1-h\partial_z f_1(z))dz\right)\cdot\left(1-\frac{1}{\epsilon}\left(\lambda_1\left(\xi\right)+O(\Delta\xi)\right)h+o(h)\right).
\end{aligned}
$$

Differentiation of this expression and letting: $\Delta\xi\to 0$, gives:

$$
\begin{aligned}
p_1(\xi,t+h)=&\left(p_0(\xi,t+h)+O(h)\right)\left(\frac{1}{\epsilon}\lambda_0(\xi)h+o(h)\right)\\
&+p_1(\xi-hf_1(\xi),t)\left(1-h\partial_\xi f_1(\xi)\right)\cdot\left(1-\frac{1}{\epsilon}\lambda_1\left(\xi\right)h+o(h)\right).
\end{aligned}
$$

We determine the probability density $p_1(\xi, t+h)$ by rearranging terms:

$$p_1(\xi, t+h) = \frac{1}{\epsilon}\lambda_0(\xi)p_0(\xi, t+h)h + p_1(\xi - hf_1(\xi), t) \cdot \left(1 - \frac{1}{\epsilon}\lambda_1(\xi)h - h\partial_\xi f_1(\xi)\right) + o(h).$$

Recall that $p_1(\xi, t+h)$ represents the density that the switching system is in state one at time $t+h$. The respective Kolmogorov master equation will be obtained from $\partial_t p_1(\xi, t)$. We compute $\partial_t p_1(\xi, t)$ as follows:

$$\begin{aligned}
\partial_t p_1(\xi, t) &= \lim_{h\to 0} \left\{ \frac{(p_1(\xi, t+h) - p_1(\xi, t))}{h} \right\} \\
&= \lim_{h\to 0} \left\{ \frac{p_1(\xi - hf_1(\xi), t) - p_1(\xi, t)}{h} \right\} - p_1(\xi, t)(\partial_\xi f_1(\xi)) \\
&\quad + \frac{1}{\epsilon}\left(\lambda_0(\xi)p_0(\xi, t) - \lambda_1(\xi)p_1(\xi, t)\right) \\
&= -\partial_\xi \left\{ f_1(\xi)p_1(\xi, t) \right\} + \frac{1}{\epsilon}\left(\lambda_0(\xi)p_0(\xi, t) - \lambda_1(\xi)p_1(\xi, t)\right).
\end{aligned}$$

Consequently, the Kolmogorov master equation associated with p_1 is:

$$\partial_t p_1 + \partial_x \left\{ f_1 p_1 \right\} = \frac{1}{\epsilon}\left(\lambda_0 p_0 - \lambda_1 p_1\right). \tag{3.6}$$

Similarly, we obtain for the Kolmogorov master equation associated with p_0:

$$\partial_t p_0 + \partial_x \left\{ f_0 p_0 \right\} = \frac{1}{\epsilon}\left(\lambda_1 p_1 - \lambda_0 p_0\right). \tag{3.7}$$

□

The terms on the right hand side of (3.6) and (3.7) can present problems to work with numerically. By choosing the scaling parameter ϵ sufficiently small, we would obtain a large number of switchings between the two respective states. To avoid this problem, we aim to find a suitable approximation of the system of Kolmogorov master equations. This approximation can be derived via a Chapman-Enskog expansion, which we will be studying in the following section.

3.3 Chapman-Enskog Equations

Theorem 3.3.0.12. *Consider the Kolmogorov master equations as presented in (3.6) and (3.7) and define the combined density p by:*

$$p := p_0 + p_1.$$

Then the partial differential equation:

$$\partial_t p + \partial_x \left\{ \frac{\lambda_0 f_1 + \lambda_1 f_0}{\lambda_0 + \lambda_1} p - \epsilon \frac{\lambda_0 (f_0 - f_1)}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_1}{\lambda_0 + \lambda_1} f_0 p \right)_x + \epsilon \frac{\lambda_1 (f_0 - f_1)}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_0}{\lambda_0 + \lambda_1} f_1 p \right)_x \right\} = 0.$$

provides an approximation of the set of Kolmogorov master equations correct to order $O(\epsilon^2)$.

Proof. Our objective is to find an approximation of the Kolmogorov master equations, where we choose the scaling parameter ϵ to be sufficiently small, so that a large number of switchings between the states zero and one occur. We would also like to obtain the combined state density of the switching process $X(t)$, rather than the respective distributions indicating that the switching system is in state one or state zero at a fixed time T. Therefore, we introduce the new variable p, which is the sum of the two densities p_0 and p_1. In addition, we also introduce an exchange term d to eliminate the terms of size $O(\epsilon)$. We then rewrite the Kolmogorov master equations in terms of the variables p and d. We are then going to apply a Chapman-Enskog expansion on the exchange term d which will yield a partial differential equation that is only dependent on the unknown variable p.

Step 1:

We define the new density variable p as:

$$p = p_0 + p_1. \tag{3.8}$$

In addition, we introduce the exchange term d as follows:

$$d = \frac{1}{\lambda_0 + \lambda_1} \left(\lambda_0 p_0 - \lambda_1 p_1 \right). \tag{3.9}$$

By setting $s_0 := \frac{\lambda_0}{\lambda_0 + \lambda_1}$ and $s_1 := \frac{\lambda_1}{\lambda_0 + \lambda_1}$ we abbreviate:

$$d = s_0 p_0 - s_1 p_1.$$

Adding the Kolmogorov master equations in (3.6) and (3.7), we obtain:

$$\partial_t p_0 + \partial_t p_1 + \partial_x \{f_0 p_0\} + \partial_x \{f_1 p_1\} = \frac{1}{\epsilon} \left(\lambda_1 p_1 - \lambda_0 p_0 \right) + \frac{1}{\epsilon} \left(\lambda_0 p_0 - \lambda_1 p_1 \right),$$

which yields:

$$\partial_t \underbrace{\{p_0 + p_1\}}_{=p} + \partial_x \{f_0 p_0 + f_1 p_1\} = 0.$$

Multiplying (3.6) and (3.7) with the respective terms s_0 and s_1, we observe for the system of Kolmogorov master equations:

$$s_0 \partial_t p_0 + s_0 \partial_x \{f_0 p_0\} - s_1 \partial_t p_1 - s_1 \partial_x \{f_1 p_1\} = \frac{s_0}{\epsilon} \left(\lambda_1 p_1 - \lambda_0 p_0 \right) - \frac{s_1}{\epsilon} \left(\lambda_0 p_0 - \lambda_1 p_1 \right).$$

This expression simplifies to:

$$\partial_t \underbrace{\{s_0 p_0 - s_1 p_1\}}_{=d} + s_0 \partial_x \{f_0 p_0\} - s_1 \partial_x \{f_1 p_1\} = -\frac{\lambda_0 + \lambda_1}{\epsilon} \underbrace{(s_0 p_0 - s_1 p_1)}_{=d}.$$

By setting $m := p_0 f_0 + p_1 f_1$, the system of Kolmogorov master Equations in terms of p and d is equivalent to:

$$\partial_t p + \partial_x m = 0 \tag{3.10}$$

and

$$\partial_t d + s_0 \partial_x \{p_0 f_0\} - s_1 \partial_x \{p_1 f_1\} = -\frac{(\lambda_0 + \lambda_1)}{\epsilon} d. \tag{3.11}$$

However (3.10) and (3.11) are still dependent on the probability densities p_0 and p_1. Thus, we will need to eliminate these variables and replace them with the introduced p and d. For this, we assume that there exists $\alpha \in R$ and $\beta \in R$, such that:

$$m = \alpha p + \beta d.$$

With (3.8) and (3.9) we obtain that the transition densities p_0 and p_1 satisfy:

$$p_0 = d + s_1 p \quad \text{and} \quad p_1 = -d + s_0 p.$$

Since $m = p_0 f_0 + p_1 f_1$, we compute α and β as follows:

$$\begin{aligned} p_0 f_0 + p_1 f_1 &= (d + s_1 p) f_0 + (-d + s_0 p) f_1 \\ &= \underbrace{(f_0 s_1 + s_0 f_1)}_{:=\alpha} p + \underbrace{(f_0 - f_1)}_{:=\beta} d \end{aligned}$$

Therefore, we observe that with $\alpha = s_0 f_1 + s_1 f_0$ and $\beta = f_0 - f_1$ the partial differential equation in (3.10) transforms into:

$$\partial_t p + \partial_x \left\{ (s_1 f_0 + s_0 f_1) p + (f_0 - f_1) d \right\} = 0. \tag{3.12}$$

Similarly, we obtain for the partial differential equation in (3.11) with $p_0 = d + s_1 p$ and $p_1 = -d + s_0 p$:

$$\partial_t d + s_0 \partial_x \left\{ (d + s_1 p) f_0 \right\} - s_1 \partial_x \left\{ (-d + s_0 p) f_1 \right\} = -\frac{\lambda_0 + \lambda_1}{\epsilon} d. \tag{3.13}$$

Step 2:

With the system of partial differential equations established in (3.12) and (3.13), the set of Kolmogorov master equations is dependent on the variables ρ and d. We can now eliminate the exchange term d via a Chapman-Enskog expansion. For this, we assume that d can be represented as an analytic function, such that for $\epsilon > 0$:

$$d = \sum_{n=0}^{\infty} \epsilon^n d_n.$$

With $d \approx d_0 + \epsilon d_1$, we will obtain an approximation of the set of Kolmogorov master equations that is correct to order $O\left(\epsilon^2\right)$. From (3.13) and with $d = \sum_{n=0}^{\infty} \epsilon^n d_n$, we find:

$$\left(\sum_{k=0}^{\infty} \epsilon^k d_k \right)_t + s_0 \left(\left(\sum_{k=0}^{\infty} \epsilon^k d_k + s_1 p \right) f_0 \right)_x - s_1 \left(\left(- \sum_{k=0}^{\infty} \epsilon^k d_k + s_0 p \right) f_1 \right)_x = -\frac{\lambda_0 + \lambda_1}{\epsilon} \left(\sum_{k=0}^{\infty} \epsilon^k d_k \right).$$

Matching the coefficients of order $O(\frac{1}{\epsilon})$ yields $d_0 = 0$, since:

$$0 = -\frac{\lambda_0 + \lambda_1}{\epsilon} d_0.$$

Furthermore, by matching the coefficients of order $O(\epsilon^0)$, we observe:

$$s_0(f_0 s_1 p)_x - s_1(f_1 s_0 p)_x = -(\lambda_0 + \lambda_1) d_1.$$

With the resubstitution of the forms $s_0 = \frac{\lambda_0}{\lambda_0+\lambda_1}$ and $s_1 = \frac{\lambda_1}{\lambda_0+\lambda_1}$, we determine d_1 as follows:

$$d_1 = \frac{\lambda_1}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_0}{\lambda_0 + \lambda_1} f_1 p \right)_x - \frac{\lambda_0}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_1}{\lambda_0 + \lambda_1} f_0 p \right)_x .$$

Inserting the approximation $d \approx d_0 + \epsilon d_1$ for (3.12) with d_0 and d_1 obtained from above, we finally arrive at:

$$\partial_t p + \partial_x \left\{ \frac{\lambda_0 f_1 + \lambda_1 f_0}{\lambda_0 + \lambda_1} p - \epsilon \frac{\lambda_0 (f_0 - f_1)}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_1}{\lambda_0 + \lambda_1} f_0 p \right)_x + \epsilon \frac{\lambda_1 (f_0 - f_1)}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_0}{\lambda_0 + \lambda_1} f_1 p \right)_x \right\} = 0. \tag{3.14}$$

□

We have derived a partial differential equation that gives a suitable approximation of the set of Kolmogorov master equations for sufficiently small values of ϵ. We can also represent (3.14) as a Fokker-Planck equation, which will be done in Chapter 5. The solution of this Fokker-Planck equation will then give the combined state density of the switching system $X(t)$ for a fixed time $T > 0$.

Chapter 4

Approach via a Stochastic Differential Equation

4.1 The Setting

In the last chapter, we have obtained a suitable approximation of the switching process $X(t)$ via a partial differential equation in a high switching environment. We are now going to approach the problem from another angle. Our objective in this chapter is to derive a stochastic differential equation whose solution gives an approximation of the switching process $X(t)$ within a small time interval, such that:

$$0 \ll \epsilon \ll dt \ll 1.$$

The fixed time interval length under which the switching process $X(t)$ evolves is denoted by dt. In particular we choose the time interval length small enough so that $X(t)$ does not move too far, but large enough such that many switchings will occur. Recall from Section 1.1 that a cycle is the time it takes for the process $I(t)$ starting in state zero, to return to this state after it had switched to state one.

We define dT_i for $i \in \mathbb{N}$ as the period of time the system stays in a particular state until it switches to the other state. In particular, the length of the time period before the first switching occurs is denoted by dT_1. Similarly, the length of the time period between the occurrence of the second switching and after the first switching is given by dT_2. Furthermore, the sum $dT_1 + dT_2$ represents the length of the first cycle and the sum $dt_{i-1} + dt_i$ would represent the length of the $\frac{i}{2}^{th}$ cycle.
To approximate the time interval length dt, we derive an estimate for the number of cycles before dt. For this, we choose an even integer m such that $\mathrm{E}[dT] \approx dt$ with:

$$dT = dT_1 + dT_2 + ... + dT_m.$$

The length of the time between two consecutive switchings dT_i is approximately distributed according to an exponential and independent random variable, for odd i, with:

$$dT_i \sim \mathrm{Exp}(q_{01}(X_{i-1})). \tag{4.1}$$

Similarly, the approximate distribution of the length of the time intervals dT_i, for even i, is given by:

$$dT_i \sim \mathrm{Exp}(q_{10}(X_{i-1})). \tag{4.2}$$

Remark 4.1.0.13. *A random variable is an exponential random variable with the constant rate parameter* $\lambda > 0$, $X \sim \mathrm{Exp}(\lambda)$, *if it has the probability density function:*

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \textit{if } f(t) \geq 0 \\ 0 & \textit{if } f(t) < 0 \end{cases}.$$

Furthermore, with $X \sim \text{Exp}(\lambda)$*, we have:*

$$\lambda X \sim \text{Exp}(1).$$

Note, that $q_{01}(X_{i-1})$ indicates the transition rate for a switch from state zero to state one and $q_{10}(X_{i-1})$ describes how fast a switch from state one to state zero occurs at the position X_{i-1}. The approximation of the distribution of the sub-time interval lengths dT_i for $i \in \{1, .., m\}$ is only valid for sufficiently small values of ϵ, as we do not take into account that the value of $X(t)$ changes between consecutive switches. Since we can calculate the expected time per cycle, we will first derive an approximation to the number of cycles through which the system passes by dt. This number will be given by $\frac{m}{2}$. Having derived a quantity for $\frac{m}{2}$, we can then apply the Central Limit Theorem to the approximate switching differential equation:

$$dX_{uncor} = f_0(X_0)dT_1 + f_1(X_1)dT_2 + ... + f_1(X_{m-1})dT_m,$$

which is associated with the switching process. From this limit, we obtain a stochastic differential equation whose solution provides a first estimation of the distribution of the switching process. Note that dX_{uncor} presents only an approximation of the actual switching differential equation, as it assumes that the value of $X(t)$ stays constant between two consecutive switches.

This procedure, however, introduces a systematic error that needs to be corrected. By summing $\frac{m}{2}$ cycles, we end up at the point (dT, dX_{uncor}). On the other hand, we want to estimate the distribution of the x-coordinate dX at the exact time dt. We use the formula $dX = dX_{uncor} - f(X)\,(dT - dt)$ as an estimation of dX. This method is presented in more detail in Figure 4.1.

At time dt the switching differential equation arrives at dX. Likewise, at time dT the

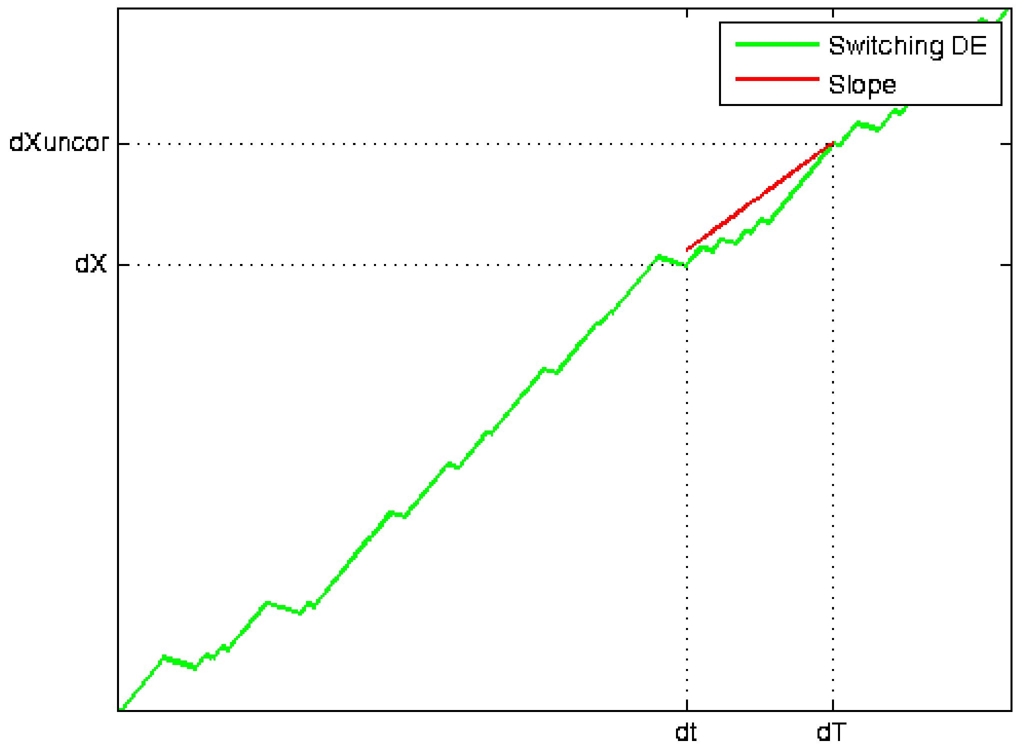

Figure 4.1: Slope argument used to find a correction of stochastic differential equation

switching differential equation is at dX_{uncor}. In Figure 4.1, we observe an example of overestimating the fixed dt with dT. The green curve represents the sample path of the switching differential equation. The red curve represents the "slope" of the switching differential equation between the time intervals dt and dT. We define this slope $f(X)$ as follows:

$$f(X) = \frac{\mathrm{E}[dX_{uncor}]}{\mathrm{E}[dT]}.$$

In order to find a correction for the stochastic differential equation derived with the Central Limit Theorem, we will subtract the overestimated time interval according to the following slope argument:

$$f(X) \approx \frac{dX_{uncor} - dX}{dT - dt}.$$

Then the corrected stochastic differential equation dX will be obtained via the formula:

$$dX \approx dX_{uncor} - f(X)\left(dT - dt\right),$$

whose solution gives an approximation of the switching process, correct to order $O\left(\epsilon\right)$. At the end of Section 4.3 we will find an expression for dX of the form:

$$dX \approx f(X)dt + g(X)dW_t.$$

4.2 Derivation of the First Version of the Stochastic Differential Equation

For our computations throughout this section, we review an important limit theorem in probability theory: the Central Limit Theorem [15]:

Theorem 4.2.0.14 (Central Limit Theorem)**.** *Let* $\{Y_i, i \geq 1\}$ *be a sequence of identically and independently distributed random variables with mean* $\mathrm{E}\left[Y_i\right] = \mu$ *and variance* $\mathrm{Var}\left[Y_i\right] = \sigma^2$*, for all* $i \geq 1$*. Then:*

$$\lim_{n \to \infty} P\left\{\frac{Y_1 + \ldots + Y_n - n\mu}{\sigma\sqrt{n}} \leq y\right\} = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{y} e^{-\frac{s^2}{2}} ds = \Phi\left(y\right).$$

We first aim to derive an estimate for the number of cycles $\frac{m}{2}$ we need to add to approximate the fixed time length dt. In particular, we choose $\frac{m}{2}$ to ensure that $E[dT] \approx dt$. Using the result for $\frac{m}{2}$, we apply the Central Limit Theorem to the switching differential equation to obtain the "uncorrected" stochastic differential equation.

Step 1:

We assume that we can estimate the fixed time interval length dt via the sum of m exponentially and independently distributed sub-time intervals length dT_i for $i \in \{1, ..., m\}$ with:

$$dT = dT_1 + dT_2 + ... + dT_m,$$

such that $\mathrm{E}[dT] \approx dt$. According to (4.1) and (4.2), we observe that dT_i for $i \in \{1, ..., m\}$ and $j \in \{0, 1\}$ is independently distributed and generated by:

$$dT_i \sim \mathrm{Exp}\left(\frac{\lambda_j(X_{i-1})}{\epsilon}\right).$$

Furthermore, by setting:

$$Y_i := \frac{\lambda_j(X_{i-1})}{\epsilon} dT_i,$$

for $i \in \{1, ..., m\}$, we conclude with Remark 4.1.0.13:

$$Y_i \sim \mathrm{Exp}(1).$$

Hence, we can rewrite dT as:

$$dT = \frac{\epsilon}{\lambda_0(X_0)} Y_1 + \frac{\epsilon}{\lambda_1(X_1)} Y_2 + ... + \frac{\epsilon}{\lambda_0(X_{m-2})} Y_{m-1} + \frac{\epsilon}{\lambda_1(X_{m-1})} Y_m.$$

We derive an estimate for $\frac{m}{2}$ by computing the expected value of dT. For this it is rather beneficial that the respective random generated sub-time intervals lengths dT_i for $i \in \{1, ..., m\}$ are identically and independently distributed. We can solve this problem by including a Taylor expansion that evolves around the points $\lambda_j(X_{i-1})$ for $j \in \{0, 1\}$ and $i \in \{1, ..., m\}$. (For further details, see Appendix A.)

Letting $X_0 \to X$, this results in the following approximation:

$$dT \approx \frac{\epsilon}{\lambda_0(X)} Y_1 + \frac{\epsilon}{\lambda_1(X)} Y_2 + ... + \frac{\epsilon}{\lambda_1(X)} Y_m.$$

We define the random generated cycles for $k \in \left\{1, ..., \frac{m}{2}\right\}$ by:

$$d\tilde{T}_k := \frac{\epsilon}{\lambda_0(X)} Y_{2k-1} + \frac{\epsilon}{\lambda_1(X)} Y_{2k}.$$

Therefore, we can rewrite dT as:

$$dT \approx d\tilde{T}_1 + d\tilde{T}_2 + ... + d\tilde{T}_{\frac{m}{2}}.$$

Since the cycles $d\tilde{T}_k$ for $k \in \left\{1, ..., \frac{m}{2}\right\}$ are identically and independently distributed with mean:

$$\mathrm{E}[d\tilde{T}_k] = \left(\frac{\epsilon}{\lambda_0(X)} + \frac{\epsilon}{\lambda_1(X)}\right)$$

we observe:

$$dt \approx \mathrm{E}[dT] \approx \frac{m}{2} \left(\frac{\epsilon}{\lambda_0(X)} + \frac{\epsilon}{\lambda_1(X)}\right),$$

from which we conclude the estimated number of cycles to add to approximate dt:

$$\frac{m}{2} \approx \frac{\lambda_0(X)\lambda_1(X)}{\epsilon\left(\lambda_0(X) + \lambda_1(X)\right)} dt. \tag{4.3}$$

Step 2:

Let us assume that the switching process $X(t)$ starts in state zero, where it stays for dT_1 until it switches to state one. There it stays for an additional time period dT_2, until it switches back to state zero. Given this procedure, the approximate switching differential equation evolves according to:

$$dX_{uncor} = f_0(X_0)dT_1 + f_1(X_1)dT_2 + ... + f_1(X_{m-1})dT_m.$$

Additionally, from (4.1) and (4.2) we observe:

$$dX_{uncor} = \epsilon \frac{f_0(X_0)}{\lambda_0(X_0)} Y_1 + \epsilon \frac{f_1(X_1)}{\lambda_1(X_1)} Y_2 + ... + \epsilon \frac{f_1(X_{m-1})}{\lambda_1(X_{m-1})} Y_m,$$

Recall that dX_{uncor} only reflects the approximate behaviour of the switching differential equation, since we assume here that the value of $X(t)$ stays constant between consecutive switches.

In order to apply the Central Limit Theorem to dX_{uncor}, the respective random variables $\frac{f_j(X_{i-1})}{\lambda_j(X_{i-1})} Y_i$, for $i \in \{1, ..., m\}$ and $j \in \{0, 1\}$, need to be identically and independently distributed. By including a Taylor expansion on $\frac{f_j(X_{i-1})}{\lambda_j(X_{i-1})}$, for $i \in \{1, ..., m\}$ and $j \in \{0, 1\}$, as presented in Appendix B and letting $X_0 \to X$, we have:

$$dX_{uncor} \approx \epsilon \frac{f_0(X)}{\lambda_0(X)} Y_1 + \epsilon \frac{f_1(X)}{\lambda_1(X)} Y_2 + ... + \epsilon \frac{f_1(X)}{\lambda_1(X)} Y_m.$$

To compute the limit of dX_{uncor} we introduce the new random variable:

$$d\tilde{X}_k := \epsilon \frac{f_0(X)}{\lambda_0(X)} Y_{2k-1} + \epsilon \frac{f_1(X)}{\lambda_1(X)} Y_{2k},$$

for $k \in \left\{1, ..., \frac{m}{2}\right\}$, whose expectation and variance are given by:

$$\mathrm{E}[d\tilde{X}_k] = \epsilon \left(\frac{f_0(X)}{\lambda_0(X)} + \frac{f_1(X)}{\lambda_1(X)} \right) \text{ and } \mathrm{Var}[d\tilde{X}_k] = \epsilon^2 \left(\frac{(f_0(X))^2}{(\lambda_0(X))^2} + \frac{(f_1(X))^2}{(\lambda_1(X))^2} \right).$$

With the introduced random variables $d\tilde{X}_k$, for $k \in \left\{1, ..., \frac{m}{2}\right\}$, we can rewrite dX_{uncor}

as follows:

$$dX_{uncor} \approx d\tilde{X}_1 + d\tilde{X}_2 + ... + d\tilde{X}_{\frac{m}{2}}.$$

Then, the application of the Central Limit Theorem for large $\frac{m}{2}$ yields:

$$dX_{uncor} \approx \Phi\left(\frac{m}{2}\epsilon\left(\frac{f_0(X)}{\lambda_0(X)} + \frac{f_1(X)}{\lambda_1(X)}\right), \frac{m}{2}\epsilon^2\left(\frac{(f_0(X))^2}{(\lambda_0(X))^2} + \frac{(f_1(X))^2}{(\lambda_1(X))^2}\right)\right).$$

Substituting $\frac{m}{2}$ with the result in (4.3) we obtain a stochastic differential equation of the following form:

$$\begin{aligned} dX_{uncor} \approx & \frac{\lambda_0(X)f_1(X) + \lambda_1(X)f_0(X)}{\lambda_0(X) + \lambda_1(X)}dt \\ & + \sqrt{\frac{\epsilon}{\lambda_0(X) + \lambda_1(X)}\left(\frac{\lambda_1(X)}{\lambda_0(X)}(f_0(X))^2 + \frac{\lambda_0(X)}{\lambda_1(X)}(f_1(X))^2\right)}dW_t. \end{aligned}$$

Consequently, from the application of the Central Limit Theorem, we derived a stochastic differential equation whose solution gives a first approximation of the solution of the switching differential equation. However, this method is prone to errors. By adding $\frac{m}{2}$ random generated cycles to approximate the fixed time interval length dt, we generally either overestimate or underestimate dt, which also shows in the accuracy of the approximation of the switching differential equation. Therefore, we are going to correct the stochastic differential equation in the following section, so that its solution gives an improved estimation of the distribution of the solution of the switching differential equation.

4.3 Correction to the Diffusion Parameter

Theorem 4.3.0.15. *Consider the switching process $\{X(t), t \geq 0\}$ as introduced in Section 1.1. Then the solution of the stochastic differential equation:*

$$dX = \frac{\lambda_0(X)f_1(X) + \lambda_1(X)f_0(X)}{\lambda_0(X) + \lambda_1(X)}dt + \sqrt{\frac{2\epsilon\lambda_0(X)\lambda_1(X)}{(\lambda_0(X) + \lambda_1(X))^3}}\,(f_0(X) - f_1(X))\,dW_t$$

gives an approximation of the distribution of the switching process correct to order $O(\epsilon)$.

Proof. In order to prove this assertion, we first calculate the difference between the random generated time interval length dT and the fixed time interval length dt. We are then going to use the slope argument as introduced in Section 4.1 to derive a formula that presents a correction for the stochastic differential equation from Section 4.2. Using the result of this procedure, we calculate the "corrected" stochastic differential equation.

Step 2:

In the previous section we estimated the time interval length dt by adding $\frac{m}{2}$ cycles as follows:

$$dT = d\tilde{T}_1 + d\tilde{T}_2 + ... + d\tilde{T}_{\frac{m}{2}}.$$

Recall, that the random variables $d\tilde{T}_k$ for $k \in \left\{1, ..., \frac{m}{2}\right\}$ are distributed according to:

$$d\tilde{T}_k \sim \left(\frac{\epsilon}{\lambda_0(X)}\text{Exp}(1) + \frac{\epsilon}{\lambda_1(X)}\text{Exp}(1)\right).$$

with mean and variance given by:

$$\text{E}[d\tilde{T}_k] = \epsilon\left(\frac{1}{\lambda_0(X)} + \frac{1}{\lambda_1(X)}\right) \text{ and } \text{Var}[d\tilde{T}_k] = \epsilon^2\left(\frac{1}{(\lambda_0(X))^2} + \frac{1}{(\lambda_1(X))^2}\right).$$

Application of the Central Limit Theorem to dT for large $\frac{m}{2}$ yields:

$$dT \approx \Phi\left(\frac{m}{2}\left(\frac{\epsilon}{\lambda_0(X)} + \frac{\epsilon}{\lambda_1(X)}\right), \frac{m}{2}\left(\frac{\epsilon^2}{(\lambda_0(X))^2} + \frac{\epsilon^2}{(\lambda_1(X))^2}\right)\right).$$

With $\frac{m}{2}$ as given in (4.3), we further reduce dT to:

$$dT \approx \Phi\left(dt, \frac{\epsilon dt}{\lambda_0(X) + \lambda_1(X)}\left(\frac{\lambda_1(X)}{\lambda_0(X)} + \frac{\lambda_0(X)}{\lambda_1(X)}\right)\right).$$

Moreover, we observe that the difference between the random generated time length dT and the fixed time length dt is given by:

$$dT - dt \doteq \sqrt{\frac{\epsilon dt}{\lambda_0(X) + \lambda_1(X)}\left(\frac{\lambda_1(X)}{\lambda_0(X)} + \frac{\lambda_0(X)}{\lambda_1(X)}\right)}\Phi(0,1). \tag{4.4}$$

Step 2:

Our aim is now to derive a formula via a slope argument which will improve the accuracy of the approximation of the switching differential equation with the stochastic differential equation. We define the slope of the switching process by:

$$f(X) = \frac{\mathrm{E}[dX_{uncor}]}{\mathrm{E}[dT]},$$

which coincides with the drift parameter of the stochastic differential equation dX_{uncor}, since $\mathrm{E}[dT] \approx dt$. Therefore, the slope of the switching process is:

$$f(X) = \frac{\lambda_1(X) f_0(X) + \lambda_0(X) f_1(X)}{\lambda_0(X) + \lambda_1(X)}. \tag{4.5}$$

Note that here we are not taking the fluctuations into account that arise from the variance of the switching process. However, since the variance of the switching process is of size $O\left(\sqrt{\epsilon}\right)$, this approximation is still valid for a sufficiently small choice of ϵ.

Since we calculate the slope of the switching process with:

$$f(X) \approx \frac{dX_{uncor} - dX}{dT - dt},$$

as introduced in Section 4.1, we observe for dX:

$$dX \approx dX_{uncor} - f(X)\left(dT - dt\right), \tag{4.6}$$

which provides us with the formula for the correction of the stochastic differential equation.

Step 3:

We now compute the corrected stochastic differential equation using the results obtained from the last two steps. Recall that the uncorrected stochastic differential equation is given in the following form:

$$dX_{uncor} \approx f(X)dt + \sqrt{\text{Var}[dX_{uncor}]}\Phi(0,1),$$

where the variance parameter of dX_{uncor} is defined by:

$$\text{Var}[dX_{uncor}] = \frac{\epsilon}{\lambda_0(X) + \lambda_1(X)}\left(\frac{\lambda_1(X)}{\lambda_0(X)}(f_0(X))^2 + \frac{\lambda_0(X)}{\lambda_1(X)}(f_1(X))^2\right). \tag{4.7}$$

The difference between the random generated time interval length dT and the fixed time interval length dt is:

$$dT - dt = \sqrt{\text{Var}[dT]}\Phi(0,1),$$

with:

$$\text{Var}[dT] = \frac{\epsilon dt}{\lambda_0(X) + \lambda_1(X)} \left(\frac{\lambda_1(X)}{\lambda_0(X)} + \frac{\lambda_0(X)}{\lambda_1(X)} \right). \tag{4.8}$$

Therefore, by combining the results from above, we rewrite (4.6) as follows:

$$dX \approx f(X)dt + \sqrt{\text{Var}[dX_{uncor}]}\Phi(0,1) - f(X)\sqrt{\text{Var}[dT]}\Phi(0,1).$$

Since dX and dX_{uncor} have the same drift term, we compute the variance of the corrected stochastic differential equation dX with:

$$\text{Var}[dX] = \text{Var}[dX_{uncor}] + (f(X))^2\text{Var}[dT] - 2f(X)\text{Cov}[dX_{uncor}, dT].$$

At the end, we aim to find a diffusion parameter for dX of the following form:

$$g(X)dW_t = \sqrt{\text{Var}[dX]}\Phi(0,1). \tag{4.9}$$

As the values of the variances of dX_{uncor} and dT are presented in (4.7) and (4.8) respectively, it is necessary to calculate the covariance of dX_{uncor} and dT.

Remark 4.3.0.16. *The covariance of two random variables X and Y, denoted by* $\text{Cov}(X,Y)$, *is defined by [15]:*

$$\text{Cov}(X,Y) = \text{E}[(X - \text{E}[X])(Y - \text{E}[Y])] = \text{E}[XY] - \text{E}[X]\text{E}[Y].$$

Furthermore, for any random variable X, Y, Z and constant c, the following properties are satisfied:

1. $\text{Cov}(X,X) = \text{Var}(X)$,

2. $\text{Cov}(X,Y) = \text{Cov}(Y,X)$,

3. $\text{Cov}(cX,Y) = c\text{Cov}(X,Y)$

4. $\text{Cov}(X,Y+Z) = \text{Cov}(X,Y) + \text{Cov}(X,Z)$.

Since the random variable dX_{uncor} and dT consist of the sum of $\frac{m}{2}$ respective independently distributed random variables $d\tilde{X}_k$ and $d\tilde{T}_k$ for $k \in \left\{1, ..., \frac{m}{2}\right\}$, we have:

$$\begin{aligned}\text{Cov}[dX_{uncor}, dT] =& \frac{m}{2}\text{Cov}[d\tilde{X}_1, d\tilde{T}_1] \\ =& \epsilon^2 \frac{m}{2}\text{Cov}\left[\frac{f_0(X)}{\lambda_0(X)}Y_1 + \frac{f_1(X)}{\lambda_1(X)}Y_2, \frac{1}{\lambda_0(X)}Y_1 + \frac{1}{\lambda_1(X)}Y_2\right],\end{aligned}$$

where $Y_1 \sim \text{Exp}(1)$ and $Y_2 \sim \text{Exp}(1)$. We further deduce:

$$\begin{aligned}\text{Cov}[dX_{uncor}, dT] =& \epsilon^2 \frac{m}{2}\text{Cov}\left[\frac{f_0(X)}{\lambda_0(X)}Y_1, \frac{1}{\lambda_0(X)}Y_1\right] + \epsilon^2 \frac{m}{2}\text{Cov}\left[\frac{f_0(X)}{\lambda_0(X)}Y_1, \frac{1}{\lambda_1(X)}Y_2\right] \\ &+ \epsilon^2 \frac{m}{2}\text{Cov}\left[\frac{f_1(X)}{\lambda_1(X)}Y_2, \frac{1}{\lambda_0(X)}Y_1\right] + \epsilon^2 \frac{m}{2}\text{Cov}\left[\frac{f_1(X)}{\lambda_1(X)}Y_2, \frac{1}{\lambda_1(X)}Y_2\right].\end{aligned}$$

Since $Y_1 \sim \text{Exp}(1)$ and $Y_2 \sim \text{Exp}(1)$ are independent random variables, their joint covariances $\text{Cov}\left[Y_1, Y_2\right]$ and $\text{Cov}\left[Y_2, Y_1\right]$ disappear. However, since $\text{Cov}\left[Y_1, Y_1\right] = \text{Var}[Y_1]$ and $\text{Cov}\left[Y_2, Y_2\right] = \text{Var}[Y_2]$, we can reduce the expression above to:

$$\text{Cov}[dX_{uncor}, dT] = \epsilon^2 \frac{m}{2}\left(\frac{f_0(X)}{(\lambda_0(X))^2} + \frac{f_1(X)}{(\lambda_1(X))^2}\right). \tag{4.10}$$

Combining the results of (4.3), (4.7), (4.8) and (4.10), we calculate the variance of

dX as follows:

$$\begin{aligned}\text{Var}[dX] =&\frac{\epsilon dt}{\lambda_0(X)+\lambda_1(X)}\left(\frac{\lambda_1(X)}{\lambda_0(X)}(f_0(X))^2+\frac{\lambda_0(X)}{\lambda_1(X)}(f_1(X))^2\right)\\ &+\frac{\epsilon(f(X))^2 dt}{\lambda_0(X)+\lambda_1(X)}\left(\frac{\lambda_1(X)}{\lambda_0(X)}+\frac{\lambda_0(X)}{\lambda_1(X)}\right)\\ &-\left(\frac{2\epsilon f(X)dt}{\lambda_0(X)+\lambda_1(X)}\right)\left(\frac{\lambda_1(X)}{\lambda_0(X)}f_0(X)+\frac{\lambda_0(X)}{\lambda_1(X)}f_1(X)\right).\end{aligned}$$

This further reduces itself to:

$$\text{Var}[dX] =\frac{\epsilon dt}{\lambda_0(X)+\lambda_1(X)}\left(\frac{\lambda_1(X)}{\lambda_0(X)}\left(f_0(X)-f(X)\right)^2+\frac{\lambda_0(X)}{\lambda_1(X)}\left(f_1(X)-f(X)\right)^2\right).$$

With $f(X)$ as presented in (4.5), we observe:

$$\text{Var}[dX]=\frac{2\epsilon\lambda_0(X)\lambda_1(X)}{(\lambda_0(X)+\lambda_1(X))^3}\left(f_0(X)-f_1(X)\right)^2 dt.$$

Since the diffusion parameter of dX is given in the form of (4.9), we obtain a corrected stochastic differential equation of the following form:

$$dX\approx\left(\frac{\lambda_1(X)f_0(X)+\lambda_0(X)f_1(X)}{\lambda_0(X)+\lambda_1(X)}\right)dt+\sqrt{\frac{2\epsilon\lambda_0(X)\lambda_1(X)}{\left(\lambda_0(X)+\lambda_1(X)\right)^3}}\left(f_0(X)-f_1(X)\right)dW.$$

□

4.4 Comments on the Correction of the Stochastic Differential Equation

In the previous two sections we have derived two stochastic differential equations with different diffusion parameters that serve as an approximation of the switching differential equation. Using the Central Limit Theorem with the estimate $\frac{m}{2}$, we obtained an approximation of the switching differential equation that is prone to errors, since it is either overestimating or underestimating the fixed time length dt. To improve the accuracy of this approximation, we used a slope argument to find a correction of the stochastic differential equation.

We will now consider a specific numerical example to shortly discuss the accuracy of the approximations with the two stochastic differential equations. For this let us take a look at the plot in Figure 4.2. In this plot we obtain the distribution of the solution of the switching differential equation, as well as the densities of the solutions of the uncorrected and corrected stochastic differential equation at $T = 0.001$.

We observe major differences in the diffusion parameter of the red and the green density curve. While the red curve presents a suitable approximation of the solution of the switching differential equation, the green curve seems to only coincide with the drift term of the blue histogram. This is not surprising as the two stochastic differential equations have the same drift term, but different diffusion parameters. The distribution of the green curve is much more expanded than the distribution of the red curve, which directly results from the overestimation or underestimation of the fixed time interval length dt.

From this example, we conclude that the corrected stochastic differential equation provides a better estimation of the solution of the switching differential equation. A numerical case study on the accuracy of the approximation of the solution of the switching differential equation with the corrected stochastic differential equation will be further discussed Chapter 6.

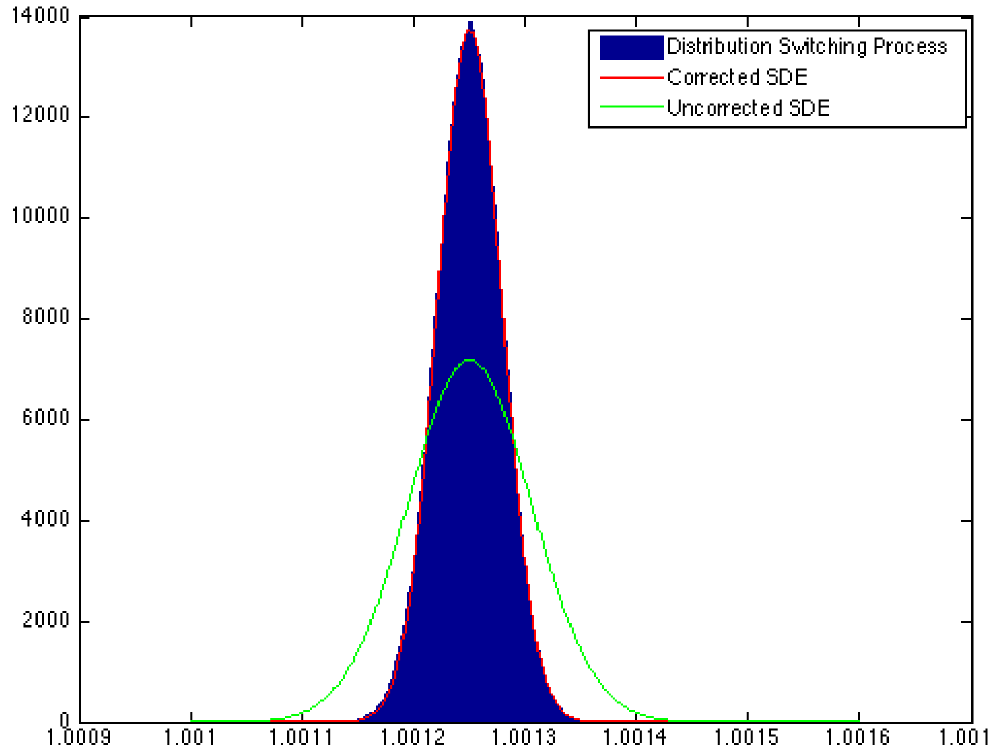

Figure 4.2: Histogram of the solution of the switching differential equation with the densities of the solution of the corrected (red curve) and the density of the solution of the uncorrected stochastic differential equation (green curve)

Chapter 5

Connection between the two Approaches

As a result of the analysis in the previous two chapters, we obtain two differential equations that present an approximate description of the switching process, if switches between the states zero and one occur in a frequent manner. The solution of the partial differential equation derived in Chapter 3, replicates the distribution of the switching process for a fixed time $T > 0$. In addition, from the stochastic differential equation derived in Chapter 4, we receive information on the behaviour of the sample paths of the switching differential equation itself.

We will now investigate how the two approaches of Chapter 3 and 4 relate to each other. This can best be analyzed by comparing the two Fokker-Planck equations resulting from the two approaches. Therefore, we will first rewrite the partial differential equation from Chapter 3 in Fokker-Planck form. For the stochastic differential equation, we recall that if its unique solution has continuous drift and diffusion parameters, it is a diffusion process. Moreover, the transition density of the solution of the stochastic differential equation with continuous drift and diffusion satisfies a

Fokker-Planck equation.

5.1 The Fokker-Planck Equation of the Partial Differential Equation Approach

Recall that in Chapter 3 we obtained the following partial differential equation:

$$\partial_t p + \partial_x \left\{ \frac{\lambda_0 f_1 + \lambda_1 f_0}{\lambda_0 + \lambda_1} p - \epsilon \frac{\lambda_0 (f_0 - f_1)}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_1 f_0 p}{\lambda_0 + \lambda_1} \right)_x + \epsilon \frac{\lambda_1 (f_0 - f_1)}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_0 f_1 p}{\lambda_0 + \lambda_1} \right)_x \right\} = 0, \tag{5.1}$$

whose solution provides an approximation of the switching process $X(t)$. In particular, this partial differential equation is a diffusion equation with a drift term. Ideally, we would like to modify (5.1), so that we obtain a Fokker-Planck equation with drift f_{pde} and diffusion g_{pde}:

$$\partial_t p + \partial_x \left\{ f_{\text{pde}} p \right\} - \frac{1}{2} \partial_{xx} \left\{ g_{\text{pde}}^2 p \right\} = 0.$$

For this, let us rewrite (5.1) as follows:

$$\begin{aligned} \partial_t p &+ \partial_x \left\{ \frac{\lambda_1 f_0 + \lambda_0 f_1}{\lambda_0 + \lambda_1} p \right\} \\ &- \epsilon \partial_x \left\{ \frac{\lambda_0 (f_0 - f_1)}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_1 f_0 p}{\lambda_0 + \lambda_1} \right)_x - \frac{\lambda_1 (f_0 - f_1)}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_0 f_1 p}{\lambda_0 + \lambda_1} \right)_x \right\} = 0. \end{aligned} \tag{5.2}$$

To represent the expression above in Fokker-Planck form, we consider the derivative:

$$\partial_{xx} \left\{ \frac{\lambda_0 \lambda_1}{(\lambda_0 + \lambda_1)^3} (f_0 - f_1)^2 p \right\} = \partial_{xx} \left\{ \frac{\lambda_0 (f_0 - f_1)}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_1 f_0 p}{\lambda_0 + \lambda_1} \right) - \frac{\lambda_1 (f_0 - f_1)}{(\lambda_0 + \lambda_1)^2} \left(\frac{\lambda_0 f_1 p}{\lambda_0 + \lambda_1} \right) \right\}.$$

Furthermore, we observe via differentiation with respect to x:

$$\partial_{xx}\left\{\frac{\lambda_0\lambda_1}{(\lambda_0+\lambda_1)^3}(f_0-f_1)^2p\right\} = \partial_x\left\{\left(\frac{\lambda_0(f_0-f_1)}{(\lambda_0+\lambda_1)^2}\right)_x\left(\frac{\lambda_1 f_0 p}{\lambda_0+\lambda_1}\right)+\frac{\lambda_0(f_0-f_1)}{(\lambda_0+\lambda_1)^2}\left(\frac{\lambda_1 f_0 p}{\lambda_0+\lambda_1}\right)_x\right\}$$
$$-\partial_x\left\{\left(\frac{\lambda_1(f_0-f_1)}{(\lambda_0+\lambda_1)^2}\right)_x\left(\frac{\lambda_0 f_1 p}{\lambda_0+\lambda_1}\right)+\frac{\lambda_1(f_0-f_1)}{(\lambda_0+\lambda_1)^2}\left(\frac{\lambda_0 f_1 p}{\lambda_0+\lambda_1}\right)_x\right\}.$$

Conveniently, we notice that the second and the forth term of the right hand side in the equation above also occur in (5.2). Therefore, by adding and subtracting the first and the third term of the expression above to (5.2), we incorporate $\partial_{xx}\left\{\frac{\lambda_0\lambda_1}{(\lambda_0+\lambda_1)^3}(f_0-f_1)^2p\right\}$ into the partial differential equation in (5.1).

Therefore, we observe for (5.2):

$$\partial_t p+\partial_x\left\{\left(\frac{\lambda_1 f_0+\lambda_0 f_1}{\lambda_0+\lambda_1}+\left(\frac{\lambda_0(f_0-f_1)}{(\lambda_0+\lambda_1)^2}\right)_x\frac{\epsilon\lambda_1 f_0}{\lambda_0+\lambda_1}-\left(\frac{\lambda_1(f_0-f_1)}{(\lambda_0+\lambda_1)^2}\right)_x\frac{\epsilon\lambda_0 f_1}{\lambda_0+\lambda_1}\right)p\right\}$$
$$-\frac{1}{2}\partial_{xx}\left\{\frac{2\epsilon\lambda_0\lambda_1}{(\lambda_0+\lambda_1)^3}(f_0-f_1)^2p\right\}=0.$$

We conclude that with a particular choice of adding and subtracting terms to our original partial differential equation, we can represent (5.1) in a Fokker-Planck form with the drift parameter $f_{\text{pde}} := f_{\text{pde}}(x)$:

$$f_{\text{pde}} := \frac{\lambda_1 f_0}{\lambda_0+\lambda_1}\left(1+\epsilon\left(\frac{\lambda_0(f_0-f_1)}{(\lambda_0+\lambda_1)^2}\right)_x\right)+\frac{\lambda_0 f_1}{\lambda_0+\lambda_1}\left(1-\epsilon\left(\frac{\lambda_1(f_0-f_1)}{(\lambda_0+\lambda_1)^2}\right)_x\right) \quad (5.3)$$

and the diffusion parameter $g_{\text{pde}} := g_{\text{pde}}(x)$:

$$g_{\text{pde}} := \sqrt{\frac{2\epsilon\lambda_0\lambda_1}{(\lambda_0+\lambda_1)^3}}(f_0-f_1). \quad (5.4)$$

We also notice that g_{pde} coincides with the diffusion term of our derived stochastic differential equation.

5.2 The Fokker-Planck Equation that Results from our Derived Stochastic Differential Equation

Recall that the transition densities of a diffusion process satisfy the Kolmogorov backward equation and the Fokker-Planck equation. For this matter, we will now analyze under which conditions the solution of our derived stochastic differential equation is a diffusion process.

From the existence and uniqueness theorem of the solution of a stochastic differential equations in Section 2.4, we obtain that a stochastic differential equation with drift parameter f and diffusion parameter g has a unique solution, if there exists a positive constant C, so that the Lipschitz condition:

$$|f(x) - f(y)| + |g(x) - g(y)| \leq C\,|x - y|\,,$$

and the restriction on growth condition:

$$|f(x)|^2 + |g(x)|^2 \leq C^2\left(1 + |x|^2\right)$$

is satisfies for all $x, y \in R$. Furthermore, as discussed in Section 2.5, the solution of a stochastic differential equation is a diffusion process if f and g are continuous. Then the transition densities p of the solution of the stochastic differential equation satisfy the Fokker-Planck equation:

$$\partial_t p + \partial_x\left\{fp\right\} - \frac{1}{2}\partial_{xx}\left\{g^2 p\right\} = 0.$$

Recall from Chapter 4, that the stochastic differential equation has the drift parameter $f := f(x)$:

$$f := \frac{\lambda_0 f_1 + \lambda_1 f_0}{\lambda_0 + \lambda_1} \tag{5.5}$$

and the diffusion parameter $g := g(x)$:

$$g := \sqrt{\frac{2\epsilon\lambda_0\lambda_1}{(\lambda_0 + \lambda_1)^3}} \left(f_0 - f_1\right). \tag{5.6}$$

We can verify the Lipschitz condition and the restriction on growth condition for the stochastic differential equation with drift (5.3) and diffusion (5.4), if we assume:

- The functions $f_0(x)$ and $f_1(x)$ are bounded and differentiable with bounded derivatives $f_0'(x)$ and $f_1'(x)$.

- The functions $\lambda_0(x)$ and $\lambda_1(x)$ associated with the rates of change are bounded by positive constants C_1 and C_2 for $j \in \{0, 1\}$ with:

$$C_1 \leq \lambda_j(x) \leq C_2.$$

  In addition, their respective derivatives $\lambda_0'(x)$ and $\lambda_1'(x)$ exist and are bounded.

(The calculations are presented in detail in Appendix C.)
Moreover, from the assumptions made on $f_j(x)$ and $\lambda_j(x)$, for $j \in \{0, 1\}$, one can obtain the continuity of the drift f and the diffusion g. We can conclude, therefore, that the solution of the stochastic differential equation is a diffusion process and its transition density p satisfies the Fokker-Planck equation:

$$\partial_t p + \partial_x \left\{ \frac{\lambda_0 f_1 + \lambda_1 f_0}{\lambda_0 + \lambda_1} p \right\} - \frac{1}{2}\partial_{xx} \left\{ \frac{2\epsilon\lambda_0\lambda_1}{(\lambda_0 + \lambda_1)^3} \left(f_0 - f_1\right)^2 p \right\} = 0.$$

5.3 Comparing the two Fokker-Planck Equations

Combining the results of the two previous sections, we are now able to analyze how the two obtained Fokker-Planck equations relate to each other. As already observed earlier, the diffusion parameters g and g_{pde} of the two respective Fokker-Planck equations are identical. However, we notice that there is a slight difference between the two drift parameters f_{pde} and f, which is given by:

$$f_{\text{pde}} - f = \frac{\epsilon\lambda_1 f_0}{\lambda_0+\lambda_1}\left(\frac{\lambda_0\left(f_0-f_1\right)}{\left(\lambda_0+\lambda_1\right)^2}\right)_x - \frac{\epsilon\lambda_0 f_1}{\lambda_0+\lambda_1}\left(\frac{\lambda_1\left(f_0-f_1\right)}{\left(\lambda_0+\lambda_1\right)^2}\right)_x .$$

The two Fokker-Planck equations would be identical, if we would choose constant sample functions for $\lambda_j(x)$ and $f_j(x)$, for $j \in \{0,1\}$. On the other hand, with a sufficiently small choice of the scaling parameter ϵ, we might not recognize a visible gap between the two drift parameters f_{pde} and f. Nevertheless, it might be still of advantage to analyze the cause of the difference in the two drift terms.

For this, let us have a another glance at the derivations of the partial differential equation in Chapter 3 and the stochastic differential equation in Chapter 4. In the approach of Chapter 3 we first compute the respective evolutions of the state densities of switching system in the time interval $[t, t+h]$. Calculating the derivative of these densities with $h \to 0$, we obtained the respective Kolmogorov master equations which are dependent on the choice of ϵ. Since for our purposes we typically aim to choose a small value of ϵ, we applied a Chapman-Enskog expansion to derive a partial differential equation whose solution gives an approximation of the density of the switching process up to order $O\left(\epsilon^2\right)$.

In contrast for the approach of Chapter 4 we studied the evolution of the sample paths of the switching differential equation for a fixed time interval length dt, such that:

$$0 \ll \epsilon \ll dt \ll 1.$$

From the switching differential equation we derived a first version of a stochastic differential equation via the application of the Central Limit Theorem. The accuary of this stochastic differential equation was improved via a slope argument, which yields a correct estimation of the switching process to order $O(\epsilon)$. However, in order to use the Central Limit Theorem, we need to ensure that the random generated time intervals dT_i, for $i \in \{1, ..., m\}$, are identically and independently distributed. In the computations of Appendix B, we omit the expression:

$$\begin{aligned}
\mathrm{E}\left[dX_b\right] = &\left(\frac{f_0(X)}{\lambda_0(X)}\right)_x \frac{\lambda_0(X)(\lambda_1(X))^2 f_0(X) + (\lambda_0(X))^2 \lambda_1(X) f_1(X)}{2\left(\lambda_0(X)+\lambda_1(X)\right)^2} dt^2 \\
&+ \left(\frac{f_1(X)}{\lambda_1(X)}\right)_x \frac{\lambda_0(X)(\lambda_1(X))^2 f_0(X) + (\lambda_0(X))^2 \lambda_1(X) f_1(X)}{2\left(\lambda_0(X)+\lambda_1(X)\right)^2} dt^2 \\
&+ \left(\frac{f_1(X)}{\lambda_1(X)}\right)_x \frac{\lambda_1(X) f_0(X) - \lambda_0(X) f_1(X)}{2(\lambda_0(X)+\lambda_1(X))} \epsilon dt \\
&- \left(\frac{f_0(X)}{\lambda_0(X)}\right)_x \frac{\lambda_1(X) f_0(X) + \lambda_0(X) f_1(X)}{2(\lambda_0(X)+\lambda_1(X))} \epsilon dt
\end{aligned}$$

due to the specific constraints on the scaling parameter ϵ and the time interval length dt. If we would want to obtain an additional drift term of size $O(\epsilon)$ from the stochastic differential equation, we would have to take the terms of size $O(dt^2)$ into account. In contrast, in the derivation of the Kolmogorov master equations in Chapter 3 we sent h (in Chapter 4 we denote the time interval length by dt) to zero but kept $\epsilon > 0$ fixed. Only after the Kolmogorov master equations were established did we consider the asymptotics as $\epsilon \to 0$. In conclusion, the two approaches from Chapter 3 and Chapter 4 correspond to different orders of the limit $h \to 0$ ($dt \to 0$) and $\epsilon \to 0$. Thus, it is not surprising that we arrive at slightly different asymptotic approximations.

Chapter 6

Numerical Simulations

We conclude the analysis of this thesis with a numerical case study on the two derived approximate descriptions of the switching process. For this, we will first compare the distribution of the solution of the switching differential equation with the solution of the stochastic differential equation and then with the solution of the partial differential equation. In addition, we will conduct later on a long time analysis on the behaviour of the solution of the switching differential equation along with its two approximate descriptions under various choices of the scaling parameter ϵ.

6.1 Testing the Diffusion Parameter of the Stochastic Differential Equation

In this section, we compare the distribution of the solution of the switching differential equation with the density of the solution of the stochastic differential equation. In particular, we investigate whether the diffusion parameter of the stochastic differential equation gives a good estimation of the diffusion term of the switching differential equation.

Recall that the approximate switching differential equation is given in the following form:

$$dX = f_0(X_0)dT_1 + f_1(X_1)dT_2 + ... + f_1(X_{m-1})dT_m$$

with $dT_i \sim \text{Exp}\left(\frac{\lambda_j(X_{i-1})}{\epsilon}\right)$ for $i \in \{1, ..., m\}$ and $j \in \{0, 1\}$. To obtain a proper representation of the distribution of the solution of the switching differential equation, we will compute a histogram consisting of $1,000,000$ samples of the random variable $X(T)$, where T is fixed during the experiment. In particular, for each individual point we solve the switching differential equation starting at $t = 0$ up until T. Since the time interval lengths dT_i for $i \in \{1, ..., m\}$ are exponentially and independently distributed random variables, we will generally overestimate T when solving the switching differential equation. Therefore, we terminate the numerical computation at a T^*, such that $T^* \geq T$, and subtract the overestimated time $T^* - T$ by a similar method as the one introduced in Section 4.3.

From Chapter 4 we also recall that the stochastic differential equation providing an approximation of the switching differential equation is given in the following form:

$$dX = f(X)dt + g(X)dW_t,$$

with the drift parameter:

$$f(X) = \frac{\lambda_0(X)f_1(X) + \lambda_1(X)f_0(X)}{\lambda_0(X) + \lambda_1(X)}$$

and the diffusion parameter:

$$g(X) = \sqrt{\frac{2\epsilon\lambda_0(X)\lambda_1(X)}{(\lambda_0(X) + \lambda_1(X))^3}}\left(f_0(X) - f_1(X)\right).$$

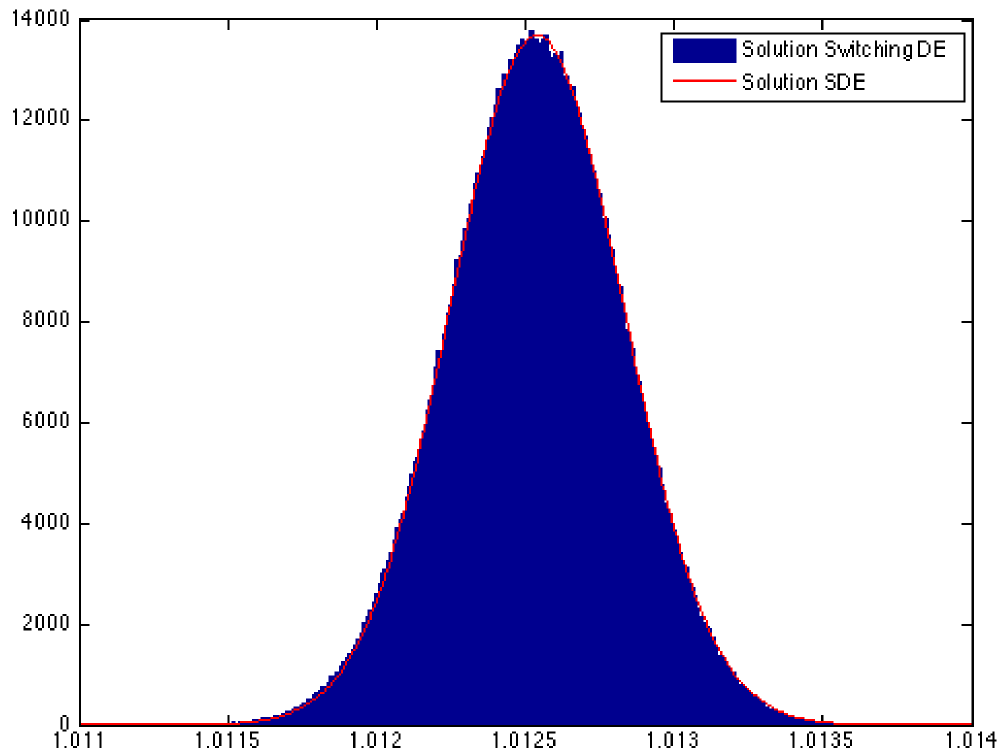

Figure 6.1: Histogram of the solution of switching differential equation with $1,000,000$ simulation points and the density of the solution of the stochastic differential equation with drift μ_X and diffusion σ_X at $T_1 = 0.01$.

Furthermore, the distribution of the stochastic differential equation follows a normal curve with drift parameter $f(X)$ and diffusion parameter $g(X)$. Thus, we calculate the solution of the stochastic differential equation by computing the drift parameter μ_X and the diffusion parameter σ_X (where $\dot{\mu}_X = f(X)$ and $\dot{\sigma}_X = g(X)$) according to the explicit Euler method with the initial value $X(0) = 1$ until the fixed time T.
To measure the accuracy of the approximation of the switching process, we overlay the density curve of the solution of the stochastic differential equation with mean μ_X and variance σ_X on top of the histogram of the solution of the switching differential equation.

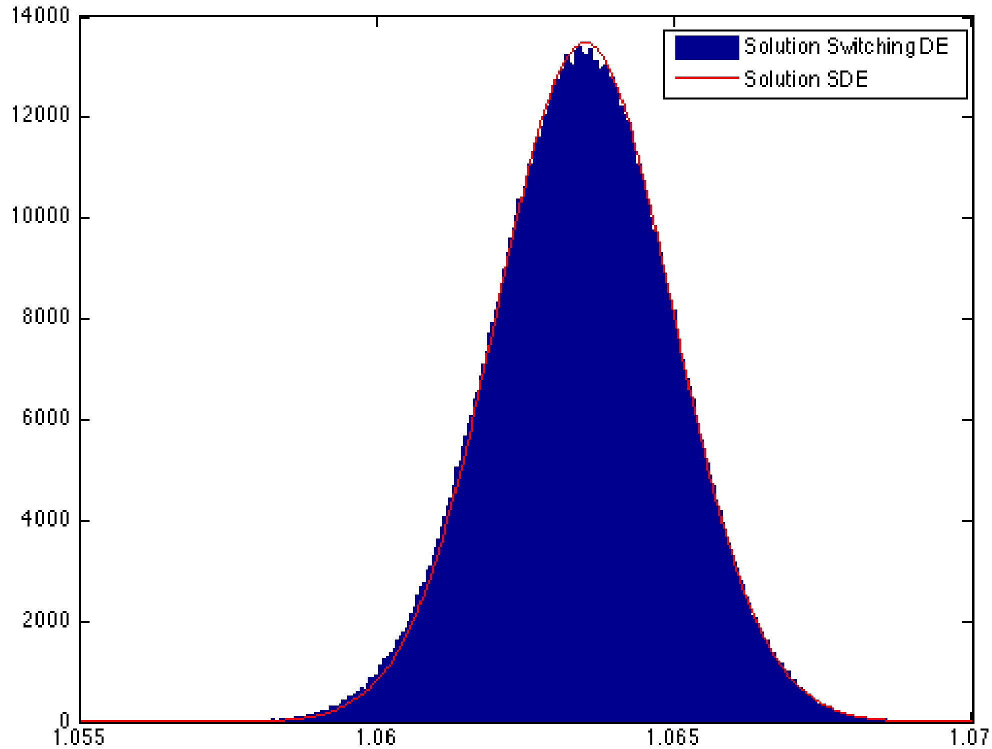

Figure 6.2: Histogram of the solution of switching differential equation with $1,000,000$ simulation points and the density of the solution of the stochastic differential equation with drift μ_X and diffusion σ_X at $T_2 = 0.05$.

As a case study for our numerical computations throughout this section, we choose the following non-constant functions $f_0(x)$ and $f_1(x)$:

$$f_0(x) = -x^2 \text{ and } f_1(x) = 1 + x.$$

For the functions $\lambda_0(x)$ and $\lambda_1(x)$ associated with the positive rates of change $q_{01}(x)$ and $q_{10}(x)$, we select:

$$\lambda_0(x) = x^2 + 2 \text{ and } \lambda_1(x) = 1.$$

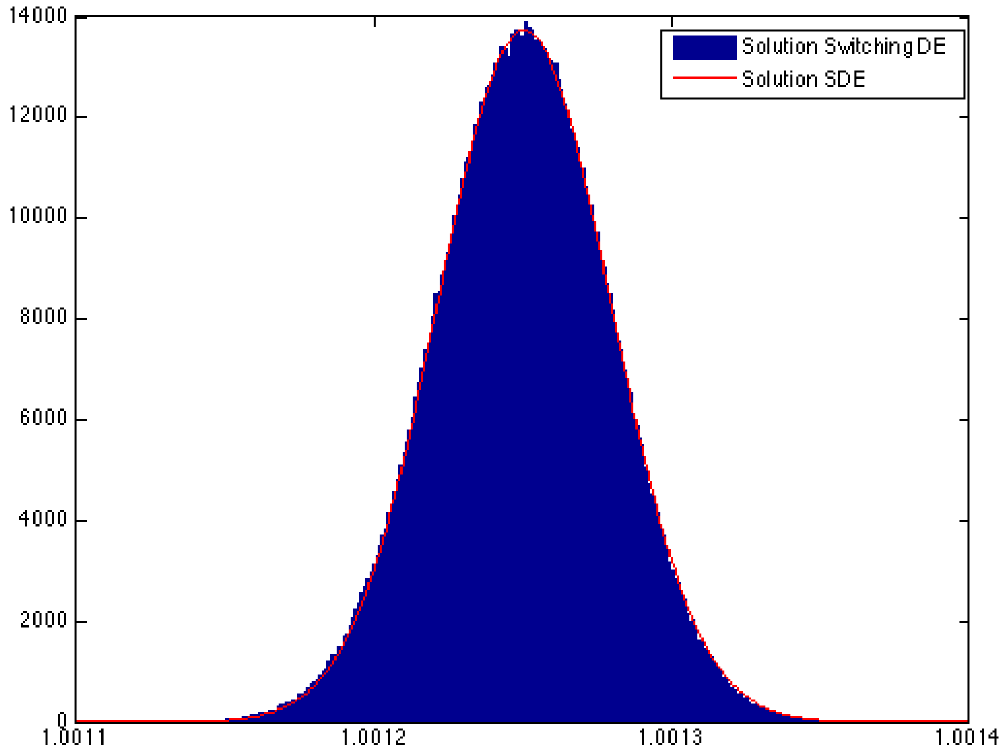

Figure 6.3: Histogram of the solution of switching differential equation with $1,000,000$ simulation points and the density of the solution of the stochastic differential equation with drift μ_X and diffusion σ_X at $T_3 = 0.001$.

We choose the initial value $X(0) = 1$ and solve the two differential equations for the three different times $T_1 = 0.01$, $T_2 = 0.05$ and $T_3 = 0.001$. In addition, we set the scaling parameter ϵ for $l \in \{1, 2, 3\}$ to be:

$$\epsilon = \frac{T_l}{1000}.$$

In the Figures 6.1, 6.2 and 6.3 we find the plots of the numerical results of the solutions of the switching and the stochastic differential equations at the times T_1, T_2 and T_3. We observe that in all of the three plots the histogram of the solution of the switching differential equation is well approximated by the density curve of the solution of the stochastic differential equation.

In Figure 6.1, the drift parameters of the blue histogram and the red normal curve seem to position themselves around the value $x = 1.0125$. We also notice that most of the points of the blue histogram are located within the interval $[1.0115, 1.0135]$. In Figure 6.2, we observe that the drift parameter of the blue histogram and the red curve coincide around $x = 1.0635$ and most of their values are concentrated within the interval $[1.058, 1.069]$. Further on in Figure 6.3, it appears that the mean parameters of the blue histogram and the red density curve seem to center themselves around $x = 1.00125$ and most of their values are located within the interval $[1.00115, 1.00135]$. With the particular choice of the functions $f_j(x)$ and $\lambda_j(x)$, for $j \in \{0, 1\}$, we observe that the drift function:

$$f(x) = \frac{x^3 + 2x + 2}{x^2 + 3}$$

is monotonically increasing. In case of a monotonically decreasing $f(x)$, we would observe that with a longer choice of T, the value of the drift parameter of the solutions of both, the switching and the stochastic differential equations, would get smaller.

In addition, a longer choice of T presents a more extended distribution of the solutions of the switching and the stochastic differential equations. According to our particular choice of ϵ, the scaling parameter is dependent on the selection of the value of T. Hence with a longer value of T more time passes between two consecutive switchings, so that the distribution of the solution of the switching differential equation becomes further spread out. Since the diffusion parameter $g(X)$ of the stochastic differential equation is dependent on ϵ, we therefore observe that a smaller choice of ϵ yields a narrower solution density curve.

To summarize the results of this section, we may conclude that the solution of the switching differential equation is well approximated with the solution of the stochastic differential equation.

On the other hand, the method used to calculate the solution of the switching differential equation might be prone to small numerical errors which we can observe when values of the blue histogram exceed the red normal curve. The error decreases with a smaller value of ϵ. However, for a smaller ϵ, the numerical computations will obviously take longer. In the proceeding two sections we will focus more on the fluctuations between the switching process and its respective approximate descriptions.

6.2 Tests on the Distributions of the Solutions of the Stochastic and the Switching Differential Equations

Having compared the density of the solution of the stochastic differential equation with the solution of the switching differential equation, we are now interested in the actual behaviour of the distribution of the solution of the stochastic differential equation.

For this, we will compute a second histogram for the solution of the stochastic differential equation in a similar manner as performed for the histogram in Section 6.1. In this case, the individual points of the histogram will be calculated according to the Euler-Maruyama method [9].

Definition 6.2.0.17 (Euler-Maruyama Approximation)**.** *We consider a stochastic process* $\{X_t, t_0 \leq t \leq T\}$ *with the initial value* $X_{t(0)} = X_0$*, that satisfies for* $t_0 \leq t \leq T$

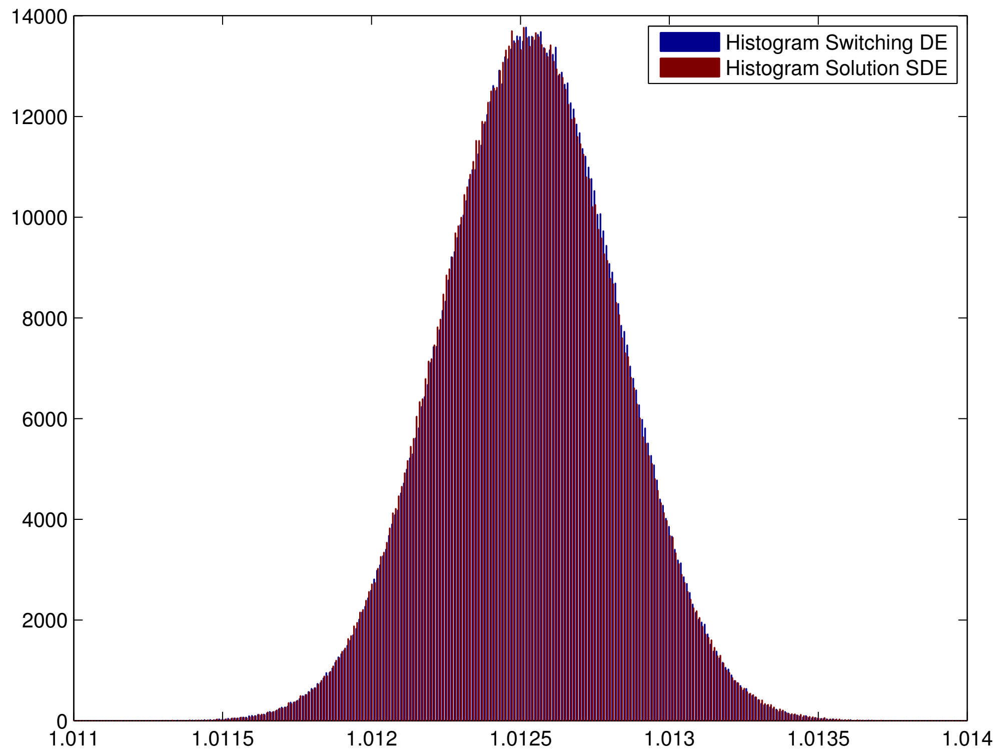

Figure 6.4: Histograms of the solutions of the switching differential equation (blue) and the stochastic differential equation (red) for 1,000,000 simulation points at time $T_1 = 0.01$.

the stochastic differential equation with drift $f(X_t)$ and diffusion $g(X_t)$:

$$dX = f(X_t)dt + g(X_t)dW_t.$$

In addition, we consider the discretization of the time interval $[t_0, T]$ with the equidistant time increments $\Delta t_n = t_{n+1} - t_n = \frac{T}{N}$:

$$t_0 = t_0 < t_1 < ... < t_n < ... < t_N = T.$$

The Euler-Maruyama approximation of the solution of the stochastic differential equa-

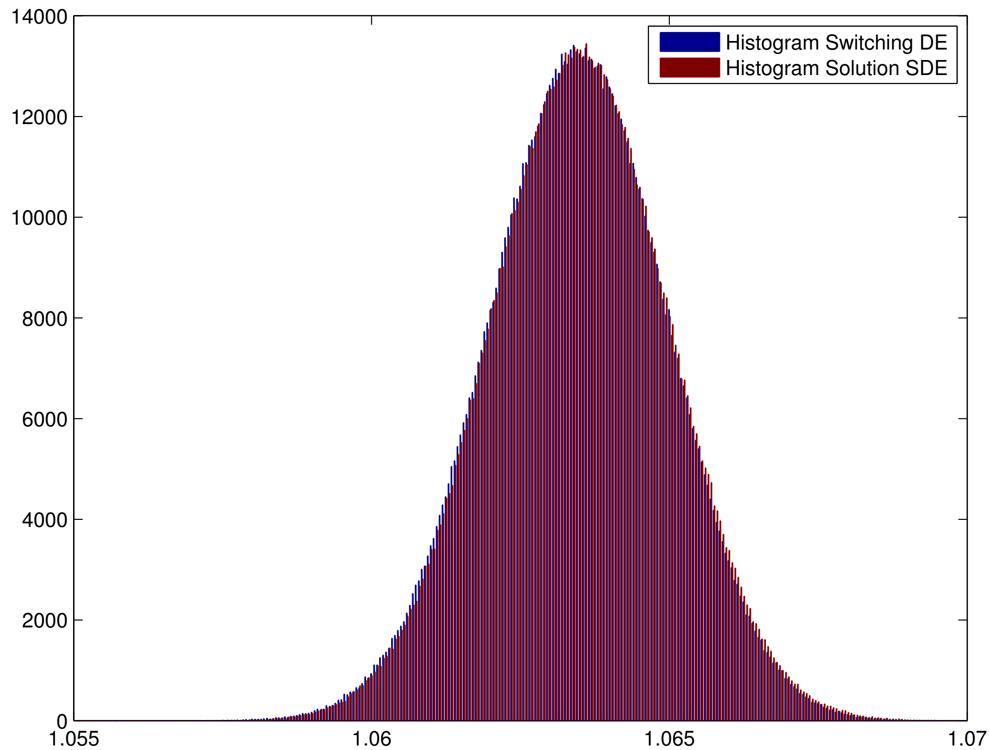

Figure 6.5: Histograms of the solutions of the switching differential equation (blue) and the stochastic differential equation (red) for 1,000,000 simulation points at time $T_2 = 0.05$.

tion is a continuous-time stochastic process

$$Y = \{Y_t, t_0 \leq t \leq T\},$$

that satisfies the following iterative scheme for $n = 0, 1, 2, ..., N-1$ *and* $Y_0 = X_0$*:*

$$Y_{t_{n+1}} = Y_{t_n} + f(Y_{t_n})\Delta t_n + g(Y_{t_n})\Delta W_{t_n},$$

where $\Delta W_{t_n} = W_{t_{n+1}} - W_{t_n}$ *is the random increment. Moreover, the random increment is generated by a Brownian Motion process with* $W_{t_{n+1}} - W_{t_n} \sim \Phi(0, \Delta t_n)$*.*

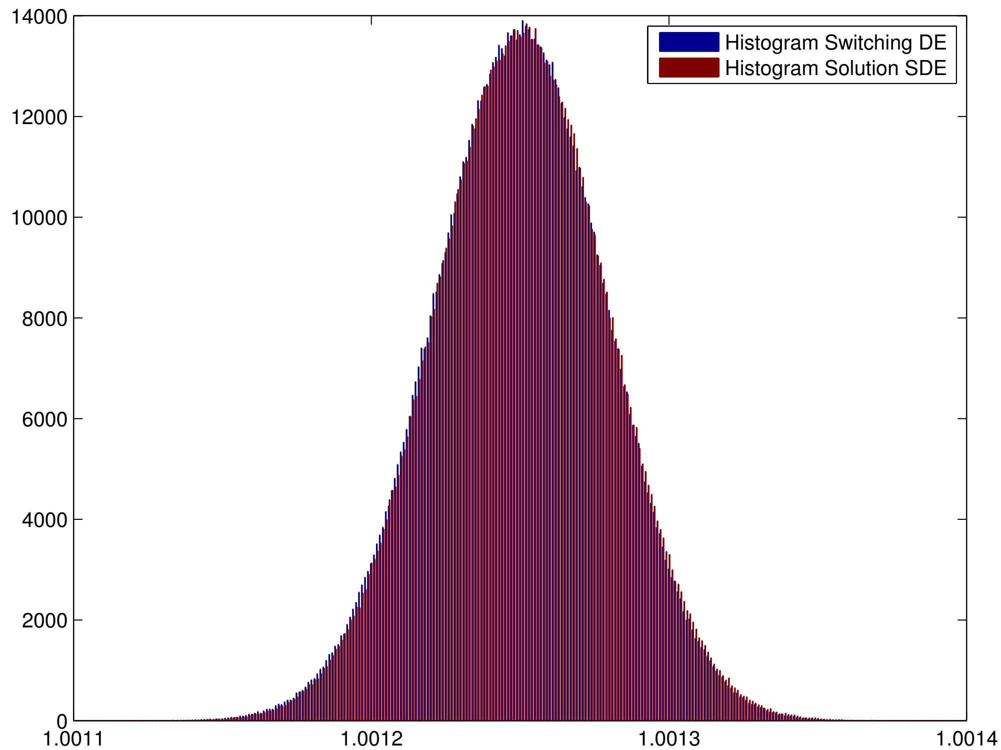

Figure 6.6: Histograms of the solutions of the switching differential equation (blue) and the stochastic differential equation (red) for 1,000,000 simulation points at time $T_3 = 0.001$.

For the numerical computations of this section, we choose the same sample function $f_j(x)$ and $\lambda_j(x)$ for $j \in \{0, 1\}$ as the ones introduced in Section 6.1. We select the respective times T_l, for $l \in \{1, 2, 3\}$, as presented in Section 6.1 to solve the stochastic differential equation, where each of the time intervals $[0, T_l]$ is divided into $1,000$ equidistant sub-time intervals.

The plots in the Figures 6.4, 6.5 and 6.6 present the respective distributions of the solutions of the stochastic differential equation and the switching differential equation for the times $T_1 = 0.01$, $T_2 = 0.05$ and $T_3 = 0.001$ with the scaling parameters $\epsilon = \frac{T_l}{1000}$, for $l \in \{1, 2, 3\}$.

Although it appears that in all of the three plots the distribution of the stochastic differential equation matches the distribution of the solution of the switching differential equation quite well, we observe slight differences between the two distributions. In particular, in Figure 6.4, we notice that in the left upper part of the two distributions the values of the red histogram exceed the values of the blue histogram. Conversely, we find that there are more values of the blue histogram concentrated in the right upper part of the two distributions.

However, in Figure 6.5, it appears that the values of the blue histogram slightly exceed the values of the red histogram on the left hand side of the two distributions. We observe that this trend reverses on the right hand side of the two distributions.

Further on, in Figure 6.6, we observe a similar trend as the one in Figure 6.5. While there tend to be slightly more values of the red histogram on the left hand side of the two distributions located, we observe the opposite trend on the right hand side of the two distributions.

Judging from the three graphs, we may conclude that the differences between the two histograms are small enough to not be really concerned with. In addition, it appears that there is no real connection between the fluctuations of the two distributions for the different plots. Therefore, we may attribute the fluctuations between the two distributions as a result of the random nature of the respective solutions of the switching differential equation and the stochastic differential equation.

From the results of this section and the ones in Section 6.1, we observe that the solution of the stochastic differential equation presents a suitable approximation of the solution of the switching differential equation for our particular choice of T and ϵ. We also remark that the approximation of the switching process via the stochastic differential equation provides information on how the sample paths of the actual switching differential equation evolve.

6.3 Numerical Tests for the Solutions of the Fokker-Planck Equation and the Switching Differential Equation

We now focus our numerical analysis on the accuracy of the estimation of the switching differential equation with the partial differential equation derived in Chapter 3. Recall from Section 5.1 that we can represent this partial differential equation in the following Fokker-Planck form:

$$\partial_t p + \partial_x \left\{ f_{\text{pde}} p \right\} - \frac{1}{2} \partial_{xx} \left\{ \left(g_{\text{pde}} \right)^2 p \right\} = 0. \tag{6.1}$$

with drift parameter f_{pde} and diffusion parameter g_{pde} as defined below:

$$f_{\text{pde}} = \frac{\lambda_1 f_0 + \lambda_0 f_1}{\lambda_0 + \lambda_1} + \frac{\epsilon \lambda_1 f_0}{\lambda_0 + \lambda_1} \left(\frac{\lambda_0 \left(f_0 - f_1 \right)}{\left(\lambda_0 + \lambda_1 \right)^2} \right)_x - \frac{\epsilon \lambda_0 f_1}{\lambda_0 + \lambda_1} \left(\frac{\lambda_1 \left(f_0 - f_1 \right)}{\left(\lambda_0 + \lambda_1 \right)^2} \right)_x$$

and

$$g_{\text{pde}} = \sqrt{\frac{2 \epsilon \lambda_0 \lambda_1}{\left(\lambda_0 + \lambda_1 \right)^3}} \left(f_0 - f_1 \right).$$

In this section, we will solve the Fokker-Planck equation with the finite differences method, which is commonly used to compute the solution of a partial differential equations [12].

For this, we represent the density $p(t, x)$ with a discretization of its values x and t:

$$t_k = t_0 + k \Delta t, \qquad \text{for } k = 0, ..., K$$

$$x_n = t_0 + n \Delta x, \qquad \text{for } n = 0, ..., N,$$

where Δt is the grid spacing for the time axis and Δx is the grid spacing for the space axis. We associate the discretization of $p(t_k, x_n)$ with p_k^n.

We represent the derivative of $\partial_t p(t,x)$ with backward time difference at t_{k+1}:

$$\left.\frac{\partial p}{\partial t}\right|_{n,k} \approx \frac{p_n^{k+1} - p_n^k}{\Delta t}.$$

Since this representation of $\partial_t p(t,x)$ is only correct to the first order, we express the derivatives $\partial_x \left\{f_{\mathrm{pde}}(x)p(t,x)\right\}$ and $\partial_{xx}\left\{g_{\mathrm{pde}}^2(x)p(t,x)\right\}$ with so called central differences:

$$\left.\frac{\partial \left\{f_{\mathrm{pde}} p\right\}}{\partial x}\right|_{n,k} \approx \frac{f_{\mathrm{pde}}(x_{n+1})p_{n+1}^{k+1} - f_{\mathrm{pde}}(x_{n-1})p_{n-1}^{k+1}}{2\Delta x}$$

and

$$\left.\frac{\partial^2 \left\{g_{\mathrm{pde}}^2 p\right\}}{\partial x^2}\right|_{n,k} \approx \frac{g_{\mathrm{pde}}^2(x_{n+1})p_{n+1}^{k+1} - 2g_{\mathrm{pde}}^2(x_n)p_n^{k+1} + g_{\mathrm{pde}}^2(x_{n-1})p_{n-1}^{k+1}}{\left(\Delta x\right)^2}.$$

The advantage of central differences is that they are correct to order 2 and therefore improve the accuracy of the numerical integration scheme. In total, combining our results from above we find for the Fokker-Planck equation in (6.1) using finite differences the following integration scheme:

$$\begin{aligned}\frac{p_n^{k+1} - p_n^k}{\Delta t} + &\frac{f_{\mathrm{pde}}(x_{n+1})p_{n+1}^{k+1} - f_{\mathrm{pde}}(x_{n-1})p_{n-1}^{k+1}}{2\Delta x} \\ &- \frac{1}{2}\frac{g_{\mathrm{pde}}^2(x_{n+1})p_{n+1}^{k+1} - 2g_{\mathrm{pde}}^2(x_n)p_n^{k+1} + g_{\mathrm{pde}}^2(x_{n-1})p_{n-1}^{k+1}}{\left(\Delta x\right)^2} = 0.\end{aligned}$$

Furthermore, solving for p_n^k, we obtain the Backward Time Centred Space representation (BTCS) with $r := \frac{\Delta t}{2\Delta x}$ and $s := \frac{\Delta t}{2(\Delta x)^2}$:

$$\begin{aligned} p_n^k =& p_{n-1}^{k+1} \underbrace{\left(-r f_{\text{pde}}(x_{n-1}) - s g_{\text{pde}}^2(x_{n-1})\right)}_{:=b(x_{n-1})} + p_n^{k+1} \underbrace{\left(1 + 2s g_{\text{pde}}^2(x_n)\right)}_{:=a(x_n)} \\ &+ p_{n+1}^{k+1} \underbrace{\left(r f_{\text{pde}}(x_{n+1}) - s g_{\text{pde}}^2(x_{n+1})\right)}_{:=c(x_{n+1})}. \end{aligned}$$

Since this equation holds for all $n \in \{0, ..., N\}$, the matrix representation of the BTCS integration scheme is given by:

$$\begin{pmatrix} p_0^k \\ p_1^k \\ \vdots \\ p_{N-1}^k \\ p_N^k \end{pmatrix} = \begin{pmatrix} a(x_0) & c(x_1) & & & 0 \\ b(x_0) & a(x_1) & c(x_2) & & \\ & \ddots & \ddots & \ddots & \\ & & b(x_{N-2}) & a(x_{N-1}) & c(x_N) \\ 0 & & & b(x_{N-1}) & a(x_N) \end{pmatrix} \begin{pmatrix} p_0^{k+1} \\ p_1^{k+1} \\ \vdots \\ p_{N-1}^{k+1} \\ p_N^{k+1} \end{pmatrix}.$$

One might wonder why we chose an implicit over an explicit integration scheme in order to solve the Fokker-Planck equation. With an explicit method, we need to enforce an additional constraint on the grid spacings of the time and space axes, known as the Courant-Friedrichs-Lewy condition. For a standard diffusion equation ($f_{\text{pde}} \equiv 0$ and $g_{\text{pde}} \equiv 1$), the Courant-Friedrichs-Lewy condition is:

$$\frac{\Delta t}{(\Delta x)^2} \leq 1.$$

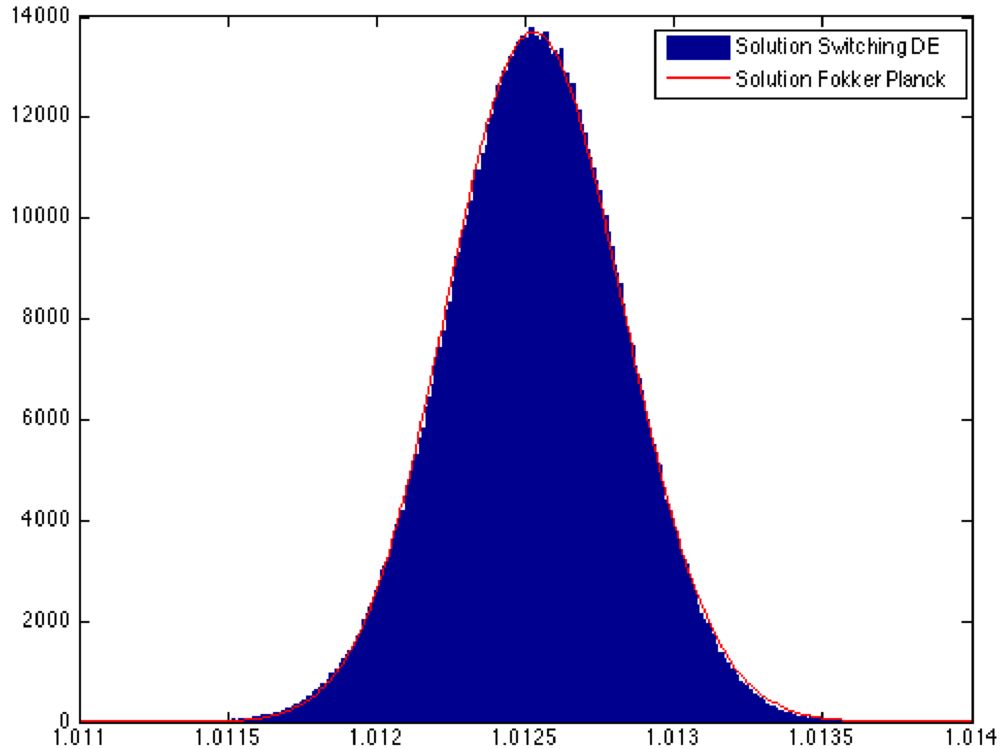

Figure 6.7: Plot of the histogram of the solution of the switching differential equation with $1,000,000$ simulation points and the solution of the Fokker-Planck equation with drift f_{pde} and diffusion g_{pde} (red curve) at time $T_1 = 0.01$.

Since one typically chooses Δx small, the numerical computations will consequently become quite time consuming. In contrast, implicit methods are stable even for relatively large time steps.

For the numerical calculations in this section, we select the same sample functions $f_j(x)$ and $\lambda_j(x)$ for $j \in \{0, 1\}$ as introduced in Section 6.1. We once again select the times: $T_1 = 0.01$, $T_2 = 0.05$ and $T_3 = 0.001$ with the respective scaling parameter:

$$\epsilon = \frac{T_l}{1000},$$

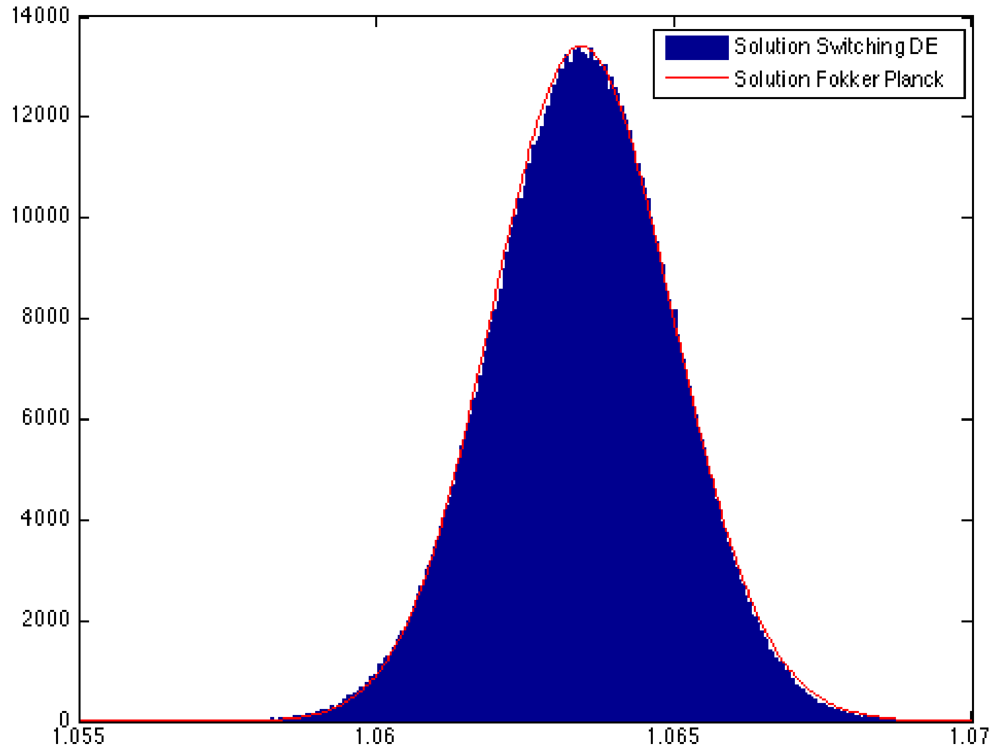

Figure 6.8: Plot of the histogram of the solution of the switching differential equation with $1,000,000$ simulation points and the solution of the Fokker-Planck equation with drift f_{pde} and diffusion g_{pde} (red curve) at time $T_2 = 0.05$.

for $l \in \{1, 2, 3\}$. Since the switching process starts in $X(0) = 1$, the initial distribution of the Fokker-Planck Equation is defined by:

$$p\,(0, x) = \begin{cases} 0 & \text{if } x \neq 0 \\ 1 & \text{if } x = 1 \end{cases}.$$

The plots in the Figures 6.7, 6.8 and 6.9 present the distribution of the solution of the switching differential equation and the density of the solution of the Fokker-Planck equation at the respective times T_1, T_2 and T_3. From the numerical results in all three Figures we conclude that the solution of the switching differential equation seems to

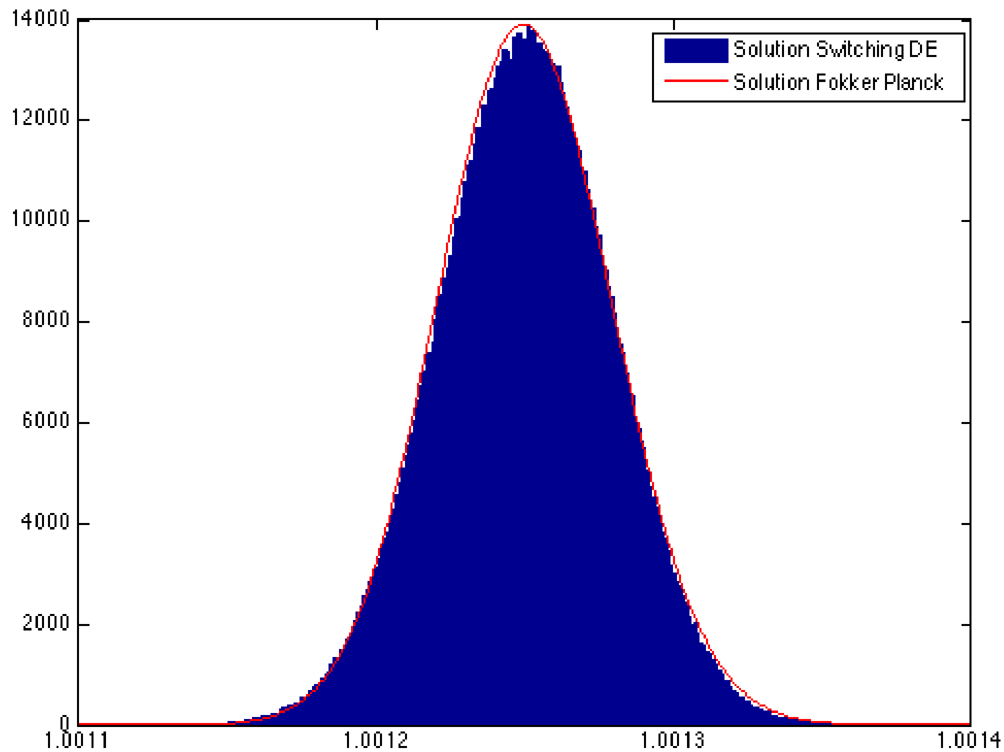

Figure 6.9: Plot of the histogram of the solution of the switching differential equation with $1,000,000$ simulation points and the solution of the Fokker-Planck equation with drift f_{pde} and diffusion g_{pde} (red curve) at time $T_3 = 0.001$.

be well approximated by the density of the solution of the Fokker-Planck equation. However, we notice that there are minor fluctuations in all of the three plots. In particular, we find that there is a gap between the red density curve and the blue histogram on the upper left hand side and the lower right hand side of the two distribution curves in all of the three examples. Conversely, we observe that the values of the blue histogram exceed the red density curve on the lower left hand side and the upper right hand side of the two distributions.

Nevertheless, these numerical errors are small enough so that we can conclude that the solution of the Fokker-Planck equation gives a good approximation of the distribution of the switching process $X(t)$ for our particular choice of ϵ and T. However, it might be of interest to analyze what is the cause of these gaps, since there appears to be a systematic error that is connected to all of the three plots. One possibility might be that the error is linked to the representation of the switching differential equation. Recall that we determined the switching differential equation by:

$$dX = \epsilon \frac{f_0(X_0)}{\lambda_0(X_0)} \mathrm{Exp}(1) + \epsilon \frac{f_1(X_1)}{\lambda_1(X_1)} \mathrm{Exp}(1) + ... + \epsilon \frac{f_1(X_{m-1})}{\lambda_1(X_{m-1})} \mathrm{Exp}(1),$$

whose solution is approximated by the solution of the Fokker-Planck equation at a fixed T. It is worth noting here that this is only an approximate description of the switching differential equation, since we assume that the value of the switching system stays constant between switches. In contrast, in the derivation of the partial differential equation in Chapter 3, we have taking into account that the switching process $X(t)$ is generally not constant between switches.

6.4 A "Long-Time" Comparison of all Three Descriptions

In the previous sections, we have been analyzing the behaviour of our approximations in a relatively short time frame. In addition, it might be of interest to investigate the accuracy of the two approximate descriptions of the switching process with a "large" choice of T and a value of ϵ that is relatively close to T.

For this, we compute the transition density $p(t, x)$ of the solution of the stochastic differential equation by solving the Fokker-Planck equation:

$$\partial_t p + \partial_x \left\{ \frac{\lambda_1 f_0 + \lambda_0 f_1}{\lambda_0 + \lambda_1} p \right\} - \frac{1}{2} \partial_{xx} \left\{ \frac{2\epsilon \lambda_0 \lambda_1}{(\lambda_0 + \lambda_1)^3} (f_0 - f_1)^2 p \right\} = 0.$$

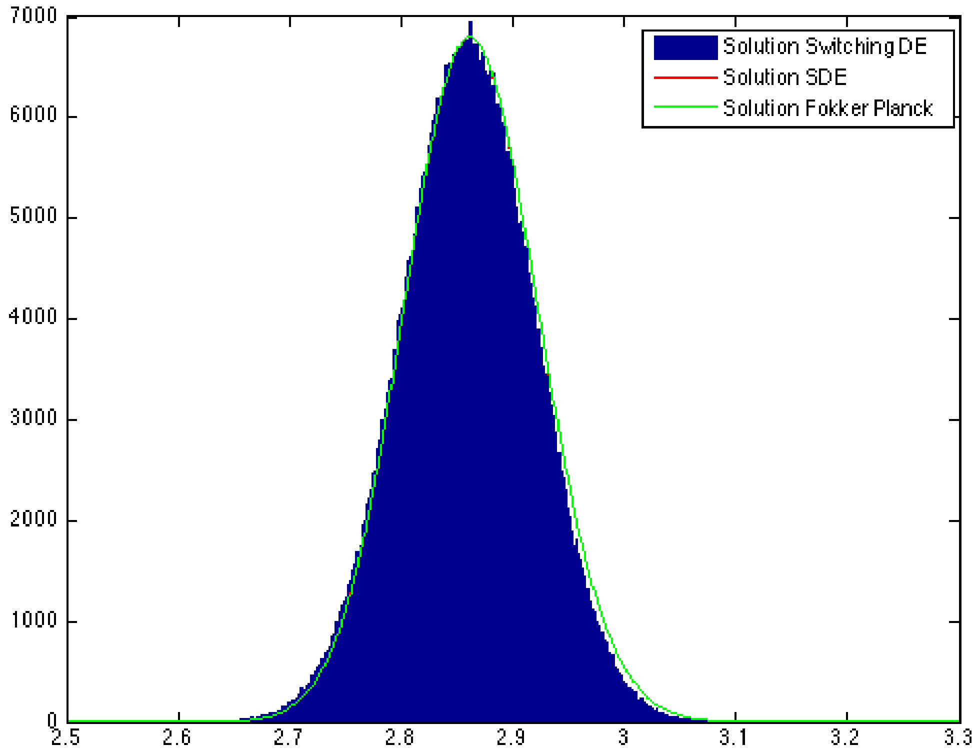

Figure 6.10: Plot of the histogram of the solution of the switching differential equation with $1,000,000$ simulation points and the solutions of the stochastic differential equation (red curve) and the Fokker-Planck equation (green curve) at time $T = 1$ with scaling parameter $\epsilon = 0.001$.

For the partial differential equation of Fokker-Planck type derived in Chapter 3, we compute its solution in the same manner as presented in Section 6.3. To distinguish between the two solved Fokker-Planck equations, we denote the Fokker-Planck equation associated with the stochastic differential equation as "Solution SDE" and the one associated with the partial differential equation, as "Solution Fokker Planck". As a case study, we choose the same sample function $f_j(x)$ and $\lambda_j(x)$ for $j \in \{0, 1\}$ as introduced in Section 6.1. For the time we select: $T = 1$ with the scaling parameters: $\epsilon_1 = 0.001$, $\epsilon_2 = 0.05$ and $\epsilon_3 = 0.02$. The results of the numerical simulations for the respective solutions of the switching differential equation, the stochastic differential

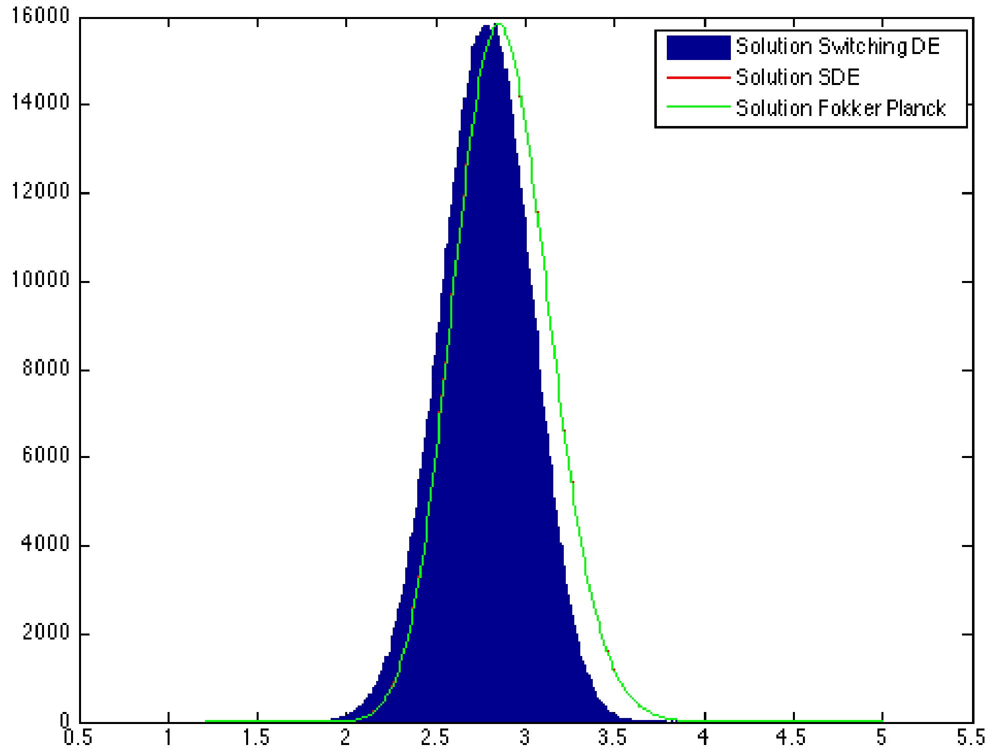

Figure 6.11: Plot of the histogram of the solution of the switching differential equation with $1,000,000$ simulation points and the solutions of the stochastic differential equation (red curve) and the Fokker-Planck equation (green curve) at time $T = 1$ with scaling parameter $\epsilon = 0.02$.

and the Fokker-Planck equation under the given parameters can be obtained in the plots of the Figures 6.10, 6.11 and 6.12.

While the density curves of the solutions of the Fokker-Planck equation and the stochastic differential equation appear to be identical in the plot of Figure 6.10, there is a noticeable gap between the drift terms of the two density curves and the blue histogram. This phenomenon becomes even more obvious, with a larger choice of ϵ. The plot in Figure 6.11 presents the solution of the switching process and the respective solutions of the Fokker-Planck equation and the stochastic differential equation, where 50 switches occur within the time interval $[0, 1]$. The drift term of the solu-

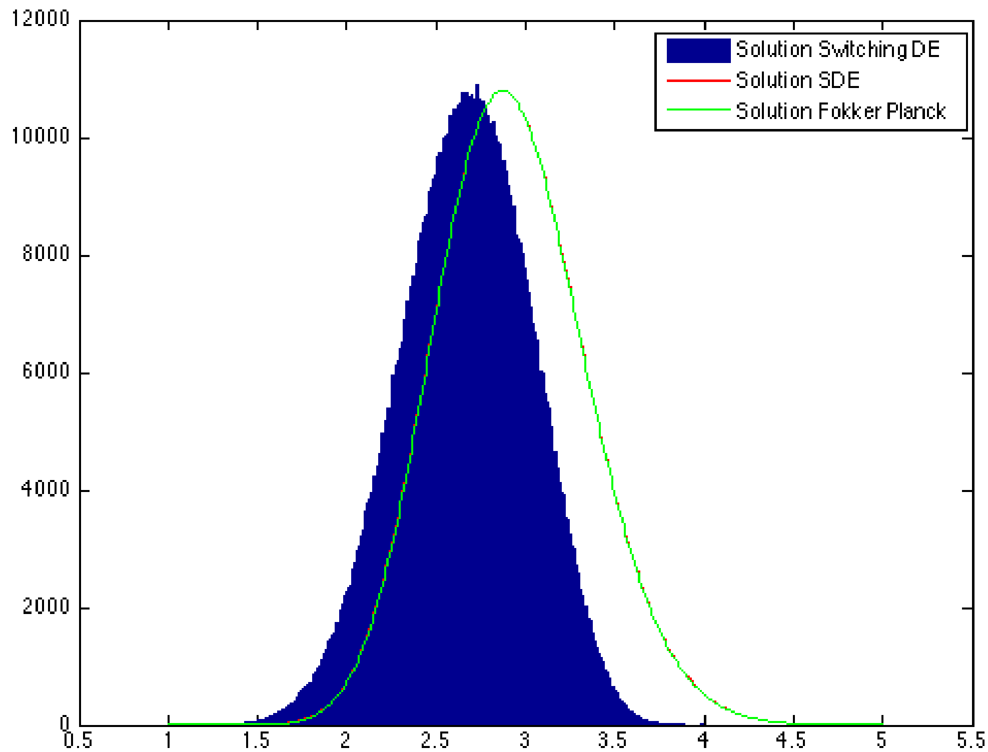

Figure 6.12: Plot of the histogram of the solution of the switching differential equation with $1,000,000$ simulation points and the solutions of the stochastic differential equation (red curve) and the Fokker-Planck equation (green curve) at time $T = 1$ with scaling parameter $\epsilon = 0.05$.

tion of the switching differential equation appears to be of a smaller value than the drift term that results from the two associated approximations. We also notice that the approximation of the diffusion term of the solution of the switching differential equation becomes less accurate with a bigger choice of ϵ, which can be obtained in the plot of Figure 6.12

In this particular plot, we observe that only 20 switches occur within the time interval $[0, 1]$ and both, the drift parameter and the diffusion parameter of the solution of the switching differential equation, are not properly estimated by the respective solutions of the stochastic differential equation and the Fokker-Planck equation. Although the value of ϵ increases with a lower rate of switchings, we do not notice any significant differences in the density curves of the solution of the stochastic differential equation and the solution of the Fokker-Planck equation, which results from our choice of $f_j(x)$ and $\lambda_j(x)$ for $j \in \{0, 1\}$. In particular, we observe for the difference of the two drift parameters $f(x)$ and $f_{\text{pde}}(x)$:

$$f_{\text{pde}}(x) - f(x) = \frac{\epsilon}{(x^2+3)^4}\left(-3x^6 + x^5 - 2x^4 + 5x^3 + 7x^2 + 10x + 6\right).$$

We conclude that with a larger choice of ϵ, the approximations of the distribution of the switching process become less accurate. However, by choosing $\epsilon = \frac{1}{1000}$ as in the case of Figure 6.10, one would expect a better estimation of the distribution of the switching process. Consequently, we also conclude that the accuracy of the approximations of the switching process declines with a larger choice of T.

Going back to our derivations in Chapter 3, one might link the inconsistency between the respective solutions of the Fokker-Planck equation and the switching differential equation to the asymptotic expansion of the exchange term d. Choosing a relatively large value of ϵ, the properties associated with the representation of d will get lost.

In the case of the approach in Chapter 4, we have again a closer look at the additional drift term of the stochastic differential equation that arises in the calculations in Appendix B. Since the additional drift term of size $\mathrm{E}[dX_b] = O\left(\epsilon dt + dt^2\right)$ can not be omitted when choosing $dt = 1$ and large values ϵ, the gap between the drift parameters of the solution of the stochastic differential equation and the solution of the switching process may be attributed to this cause.

Chapter 7

Conclusion

Stochastic switching systems which allow a large number of switches within a small time period have many applications for modelling purposes in a variety of fields such as economics, communication networks and cell biology. However, for future predictions on the evolution of these particular type of dynamical systems or to compute their moments, it is useful to find alternative levels of descriptions that provide information on the combined state densities of the switching system.

The model analyzed throughout this thesis is a two-state stochastic switching system in a high rate switching environment. This system has been studied in various literatures, including [8], where an approximate description of the switching system via a partial differential equation was derived. In particular, one may observe that an exact representation of the switching system is presented via a set of Kolmogorov master equations. However, the solutions of the respective Kolmogorov master equations only provide the density of the individual states of the switching process. In fact it might be also rather difficult to find a numerical solution of the Kolmogorov master equations for a small value of ϵ. Nevertheless this problem can be avoided with a Chapman-Enskog expansion on the set of Kolmogorov master equations. This resulted in a partial differential equation whose solution provides an approximation of the density of the switching process correct to order $O\left(\epsilon^2\right)$.

In this thesis we derived yet another approximation of the switching differential equation via a stochastic differential equation. We initiated our derivations through the use of the Central Limit Theorem to obtain a stochastic differential equation that serves as an approximation of the switching differential equation. However, this method is prone to errors due to the estimation of the fixed time length dt and therefore, we added a correction via a slope argument. This method significantly improved the accuracy of the approximation of the switching differential equation to order $O(\epsilon)$. The advantage of the approximation with the stochastic differential equation is that it not only allows us to receive information on the distribution of the solution of the switching differential equation, but we also get an idea of how the sample paths of the actual switching differential equation evolve.

Although the two approximate descriptions of the switching process are of a very distinctive nature, we observe that they relate to each other. For this, we computed the two Fokker-Planck equations that are associated with the partial and stochastic differential equation. We obtained that their respective diffusion parameters are identical, but the drift parameters of the two Fokker-Planck equations differ in a term of size $O(\epsilon)$. Since we typically choose ϵ to be small, the difference between the two Fokker-Planck equations was barely visible in numerical tests.

To conclude our analysis throughout this thesis, we implemented a numerical case study on the two approximations of the switching process, where we considered the cases of a high switching volume within a short time period and of a lower switching volume within a longer time frame. For a sufficiently small value of T and ϵ, we observed that the distribution of the solution of the switching differential equation was well approximated by the respective solutions of the partial and stochastic differential equations. However, this did not seem to be the case in the longer time frame, where we studied the behaviour of the approximations of the switching process under various parameters of ϵ. In particular, with a larger value of ϵ, we observed that the estimations of the switching process became less accurate.

We conclude that the derived stochastic differential equation, as well as the partial differential equation, provide very accurate approximations of the distribution of the analyzed switching process $X(t)$ in a short time environment, where switches between the two states occur frequently. From the solution of the partial differential equation we can gather informations on the combined state densities of the switching process. An alternative approximation of the solution of the switching differential equation is given via the solution of the stochastic differential equation, from which we also obtain an approximate description of how the sample paths of the actual switching differential equation evolve.

7.1 Further Direction

The approach of the stochastic differential equation opens many doors for future research. For modelling purposes, a first idea is to extend the switching model to n different states, so that with each switch the system has the choice to jump to $n-1$ other possible states. One could also derive a stochastic differential equation that serves as an approximation of a multi-dimensional two-state switching system.

Finally, it should be possible to find an approximation of a multi-dimensional n-state switching system in a high rate of switching environment through a stochastic differential equation.

Appendix A

Appendix: Taylor Expansion on dT

We study the limit of the random generated time interval length dT:

$$dT = \frac{\epsilon}{\lambda_0(X_0)} Y_1 + \frac{\epsilon}{\lambda_1(X_1)} Y_2 + ... + \frac{\epsilon}{\lambda_1(X_{m-1})} Y_m.$$

To simplify the computation the expectated value of dT, we aim that the respective random variables $\frac{\epsilon}{\lambda_j(X_{i-1})} Y_i$, for all $i \in \{1, ..., m\}$ and $j \in \{0, 1\}$, are identically and independently distributed.

For $\lambda_j(X_{i-1})$ we find the following representation:

$$\lambda_j(X_{i-1}) = \begin{cases} \lambda_0(X_0 + f_0(X_0)dT_1 + f_1(X_1)dT_2 + ... + f_1(X_{i-2})dT_{i-1}), & \text{if } j = 0 \text{ and } i \text{ odd} \\ \lambda_1(X_0 + f_0(X_0)dT_1 + f_1(X_1)dT_2 + ... + f_0(X_{i-2})dT_{i-1}), & \text{if } j = 1 \text{ and } i \text{ even} \end{cases}.$$

Furthermore, by including Taylor expansions that evolve around the points $\lambda_j(X_{i-1})$ for $j \in \{0, 1\}$ and $i \in \{1, ..., m\}$ up to order $O(\epsilon^2)$, we observe:

$$\lambda_j(X_{i-1}) \approx \begin{cases} \lambda_0(X_0) + \lambda_0'(X_0)\left(f_0(X_0)dT_1 + ... + f_1(X_0)dT_{i-1}\right), & \text{if } j = 0 \text{ and } i \text{ odd} \\ \lambda_1(X_0) + \lambda_1'(X_0)\left(f_0(X_0)dT_1 + ... + f_0(X_0)dT_{i-1}\right), & \text{if } j = 1 \text{ and } i \text{ even} \end{cases}.$$

Inserting the presented Taylor expansions into dT, we deduce:

$$
\begin{aligned}
dT =& \frac{\epsilon}{\lambda_0(X_0)} Y_1 + \frac{\epsilon}{\lambda_1(X_0) + \lambda_1'(X_0) f_0(X_0) dT_1} Y_2 + ... \\
& ... + \frac{\epsilon}{\lambda_1(X_0) + \lambda_1'(X_0) \left(f_0(X_0) dT_1 + ... + f_0(X_0) dT_{m-1} \right)} Y_m.
\end{aligned}
$$

Further expansions of dT yields:

$$
\begin{aligned}
dT =& \frac{\epsilon}{\lambda_0(X_0)} Y_1 + \frac{\epsilon}{\lambda_1(X_0)} \left(\frac{1}{1 + \frac{\lambda_1'(X_0)}{\lambda_1(X_0)} f_0(X_0) dT_1} \right) Y_2 \\
& + \frac{\epsilon}{\lambda_0(X_0)} \left(\frac{1}{1 + \frac{\lambda_0'(X_0)}{\lambda_0(X_0)} \left(f_0(X_0) dT_1 + f_1(X_0) dT_2 \right)} \right) Y_3 + ... \\
& ... + \frac{\epsilon}{\lambda_1(X_0)} \left(\frac{1}{1 + \frac{\lambda_1'(X_0)}{\lambda_1(X_0)} \left(f_0(X_0) dT_1 + ... + f_0(X_0) dT_{m-1} \right)} \right) Y_m.
\end{aligned}
$$

By assuming that the random variables:

$$
r_j(X_{i-1}) := \begin{cases} -\frac{\lambda_0'(X_0)}{\lambda_0(X_0)} \left(f_0(X_0) dT_1 + ... + f_1(X_0) dT_{i-1} \right) & \text{if } j = 0 \text{ and } i \text{ odd} \\ -\frac{\lambda_1'(X_0)}{\lambda_1(X_0)} \left(f_0(X_0) dT_1 + ... + f_0(X_0) dT_{i-1} \right) & \text{if } j = 1 \text{ and } i \text{ even} \end{cases},
$$

satisfy $|r_j(X_{i-1})| < 1$ for all $i \in \{1, ..., m\}$ and $j \in \{0, 1\}$, we can use geometric series to rewrite dT. In particular, we observe:

$$
\begin{aligned}
dT =& \frac{\epsilon}{\lambda_0(X_0)} Y_1 + \frac{\epsilon}{\lambda_1(X_0)} \left(1 - \frac{\lambda_1'(X_0)}{\lambda_1(X_0)} f_0(X_0) dT_1 \right) Y_2 \\
& + \frac{\epsilon}{\lambda_0(X_0)} \left(1 - \frac{\lambda_0'(X_0)}{\lambda_0(X_0)} \left(f_0(X_0) dT_1 + f_1(X_0) dT_2 \right) \right) Y_3 + ... \\
& ... + \frac{\epsilon}{\lambda_1(X_0)} \left(1 - \frac{\lambda_1'(X_0)}{\lambda_1(X_0)} \left(f_0(X_0) dT_1 + ... + f_0(X_0) dT_{m-1} \right) \right) Y_m,
\end{aligned}
$$

where we have collected all terms up to order $O(\epsilon^2)$. In addition, we split the sum dT into two disjoint parts with $dT = dT_a - dT_b$, where:

$$dT_a := \frac{\epsilon}{\lambda_0(X_0)}Y_1 + \frac{\epsilon}{\lambda_1(X_0)}Y_2 + ... + \frac{\epsilon}{\lambda_1(X_0)}Y_m$$

and

$$\begin{aligned} dT_b := & \frac{\epsilon\lambda_1'(X_0)}{(\lambda_1(X_0))^2}\left(f_0(X_0)dT_1Y_2 + ... + \left(f_0(X_0)dT_1 + ... + f_1(X_0)dT_{m-1}\right)Y_m\right) \\ & + \frac{\epsilon\lambda_0'(X_0)}{(\lambda_0(X_0))^2}\left(\left(f_0(X_0)dT_1 + f_1(X_0)dT_2\right)Y_3 + ... + \left(f_0(X_0)dT_1 + ... + f_0(X_0)dT_{m-2}\right)Y_{m-1}\right). \end{aligned}$$

Ideally, we would like to see that the sum dT_b disappears, so that $dT \approx dT_a$. For this, we approximate the length of the time intervals dT_i by:

$$dT_i \approx \frac{\epsilon}{\lambda_j(X_0)}Y_i,$$

where all $Y_i \sim \text{Exp}(1)$ for $i \in \{1, ..., m\}$ and $j \in \{0, 1\}$ are independently distributed. Inserting these approximations into dT_b, we deduce:

$$\begin{aligned} dT_b \approx & \frac{\epsilon\lambda_1'(X_0)}{(\lambda_1(X_0))^2}\left(\frac{\epsilon f_0(X_0)}{\lambda_0(X_0)}Y_1Y_2 + ... + \left(\frac{\epsilon f_0(X_0)}{\lambda_0(X_0)}Y_1 + ... + \frac{\epsilon f_1(X_0)}{\lambda_1(X_0)}Y_{m-1}\right)Y_m\right) \\ & + \frac{\epsilon\lambda_0'(X_0)}{(\lambda_0(X_0))^2}\left(\left(\frac{\epsilon f_0(X_0)}{\lambda_0(X_0)}Y_1 + \frac{\epsilon f_1(X_0)}{\lambda_1(X_0)}Y_2\right)Y_3 + ... + \left(\frac{\epsilon f_0(X_0)}{\lambda_0(X_0)}Y_1 + ... + \frac{\epsilon f_1(X_0)}{\lambda_1(X_0)}Y_{m-2}\right)Y_{m-1}\right). \end{aligned}$$

Since $\mathrm{E}[Y_i Y_l] = 1$ for all $i \in \{1, ..., m-1\}$, $l \in \{2, ..., m\}$ with $i < l$, computation of the expectation of dT_b yields:

$$\begin{aligned}\mathrm{E}[dT_b] \approx & \frac{\epsilon^2 \lambda_1'(X_0)}{(\lambda_1(X_0))^2} \left(\frac{f_0(X_0)}{\lambda_0(X_0)} \left(\sum_{k=1}^{\frac{m}{2}} k \right) + \frac{f_1(X_0)}{\lambda_1(X_0)} \left(\sum_{k=1}^{\frac{m-2}{2}} k \right) \right) \\ & + \frac{\epsilon^2 \lambda_0'(X_0)}{(\lambda_0(X_0))^2} \left(\frac{f_0(X_0)}{\lambda_0(X_0)} + \frac{f_1(X_0)}{\lambda_1(X_0)} \right) \left(\sum_{k=1}^{\frac{m-2}{2}} k \right).\end{aligned}$$

Note that the values of the summations in $\mathrm{E}[dT_b]$ are given by:

$$\sum_{k=1}^{\frac{m-2}{2}} k = \frac{1}{2} \left(\frac{m^2}{4} - \frac{m}{2} \right) \text{ and } \sum_{k=1}^{\frac{m}{2}} k = \frac{1}{2} \left(\frac{m^2}{4} + \frac{m}{2} \right).$$

Since the quantity $\frac{m}{2}$ is of size $O\left(\frac{dt}{\epsilon}\right)$, we observe through further computations:

$$\mathrm{E}[dT_b] = O\left(dt^2 + \epsilon \cdot dt\right).$$

From the assumptions made on ϵ and dt in the beginning of Section 4.1, we observe $\mathrm{E}[dT_b] \approx 0$. In the same manner, one can show that $\mathrm{Var}[dT_b] \approx 0$. Therefore, we conclude: $dT \approx dT_a$.

Appendix B

Appendix: Taylor Expansion on dX

We consider the switching differential equation:

$$dX = \epsilon g_0(X_0)Y_1 + \epsilon g_1(X_1)Y_2 + ... + \epsilon g_0(X_{m-2})Y_{m-1} + \epsilon g_1(X_{m-1})Y_m, \qquad \text{(B.1)}$$

where $g_j(X_i) = \frac{f_j(X_i)}{\lambda_j(X_i)}$ for $j \in \{0, 1\}$ and $i \in \{1, ..., m\}$. In the same manner as in Appendix A, we introduce the following Taylor expansions evolving around the points $g_0(X_{i-1})$ and $g_1(X_{i-1})$ up to order $O(\epsilon^2)$:

$$g_j(X_{i-1}) \approx \begin{cases} g_0(X_0) + g_0'(X_0)\left(f_0(X_0)dT_1 + ... + f_1(X_0)dT_{i-1}\right) & \text{if } j = 0 \text{ and } i \text{ odd} \\ g_1(X_0) + g_1'(X_0)\left(f_0(X_0)dT_1 + ... + f_0(X_0)dT_{i-1}\right) & \text{if } j = 1 \text{ and } i \text{ even} \end{cases}.$$

Inserting the Taylor expansions into the expression in $(B.1)$, we find:

$$\begin{aligned} dX \approx & \epsilon \left(g_0(X_0)Y_1 + \left(g_1(X_0) + g_1'(X_0)f_0(X_0)dT_1\right)Y_2\right) + ... \\ & ... + \epsilon \left(\left(g_1(X_0) + g_1'(X_0)\left(f_0(X_0)dT_1 + ... + f_0(X_0)dT_{m-1}\right)\right)Y_m\right). \end{aligned}$$

From the approximation of the distribution of the time lengths dT_i, for $i \in \{1, ..., m\}$ and $j \in \{0, 1\}$ in Appendix A, we observe:

$$dT_i \approx \frac{\epsilon}{\lambda_j(X_0)} Y_i,$$

where all $Y_i \sim \text{Exp}\,(1)$ are independently distributed. Thus, with the approximation of the time interval lengths dT_i we obtain:

$$dX \approx \epsilon \left(g_0(X_0)Y_1 + ... + g_0(X_0)Y_m\right) + \epsilon^2 g_1'(X_0)g_0(X_0)Y_1Y_2 + ... + \epsilon^2 g_1'(X_0)g_0(X_0)Y_{m-1}Y_m. \tag{B.2}$$

Furthermore, we split (B.2) into two disjoint parts dX_a and dX_b with

$$dX_a := \epsilon \left(g_0(X_0)Y_1 + g_1(X_0)Y_2 + ... + g_1(X_0)Y_m\right)$$

and

$$dX_b := \epsilon^2 \left(g_1'(X_0)g_0(X_0)Y_1Y_2 + ... + g_1'(X_0)g_0(X_0)Y_{m-1}Y_m\right),$$

so that $dX = dX_a + dX_b$. Should the term dX_b disappear, we will conclude: $dX \approx dX_a$. Since all $Y_i \sim \text{Exp}(1)$ are independently distributed, we compute: $\text{E}[Y_iY_l] = 1$ for all $i \in \{1, ..., m-1\}$, $l \in \{2, ..., m\}$ with $i < l$. Calculation of the expectation of dX_b, results in:

$$\begin{aligned}\text{E}[dX_b] =& \epsilon^2 g_1'(X_0)g_0(X_0)\left(\sum_{k=1}^{\frac{m}{2}} k\right) \\ &+ \epsilon^2 \left(g_0'(X_0)g_1(X_0) + g_0'(X_0)g_0(X_0) + g_1'(X_0)g_1(X_0)\right)\left(\sum_{k=1}^{\frac{m-2}{2}} k\right).\end{aligned}$$

By a similar argument as the one in Appendix A, we find:

$$\mathrm{E}[dX_b] = O\left(dt^2 + \epsilon \cdot dt\right) \approx 0.$$

In the same manner, one can show that $\mathrm{Var}[dX_b] \approx 0$. Therefore, we conclude: $dX_b \approx 0$, which yields:

$$dX \approx dX_a.$$

Appendix C

Appendix: Existence of a Unique Solution of the SDE

According to Theorem 2.4.0.6, a stochastic differential equation has a unique solution, if the following properties are satisfied:

- **Lipschitz Condition:** There exists a constant $C > 0$, such that for all $x, y \in R$:

$$|f(x) - f(y)| + |g(x) - g(y)| \leq C |x - y| .$$

- **Restriction on Growth:** There exists a constant $C > 0$, such that for all $x \in R$:

$$|f(x)|^2 + |g(x)|^2 \leq C^2 \left(1 + |x|^2\right) .$$

To verify these properties for the stochastic differential equation derived in Chapter 4, we assume:

- The functions $f_0(x)$ and $f_1(x)$ are bounded and differentiable. Their respective derivatives are also bounded.

- From Section 1.1, we obtain that the functions $\lambda_j(x)$ for $j \in \{0,1\}$ associated with the rates of change are bounded by the positive constants C_1 and C_2:

$$C_1 \leq \lambda_j(x) \leq C_2.$$

- Furthermore, the functions $\lambda_j(x)$ for $j \in \{0,1\}$ are differentiable and their respective derivatives are bounded.

Lipschitz Condition:

For the subsequent calculations let us define the functions $s_j(x)$ for $j \in \{0,1\}$ and $\gamma(x)$ as follows:

$$s_j(x) := \frac{\lambda_j(x)}{\lambda_0(x) + \lambda_1(x)}$$

and

$$\gamma(x) := \sqrt{\frac{2\epsilon\lambda_0(x)\lambda_1(x)}{(\lambda_0 + \lambda_1(x))^3}}.$$

Recall that the functions $s_0(x)f_1(x)$ and $s_1(x)f_0(x)$ are Lipschitz continuous, if they are differentiable and their respective derivates are bounded. For the derivative of $s_0(x)f_1(x)$, we compute:

$$s_0'(x)f_1(x) + s_0(x)f_1'(x) = \frac{(\lambda_0(x) + \lambda_1(x))\,\lambda_0(x)f_1'(x) + (\lambda_1(x)\lambda_0'(x) - \lambda_0(x)\lambda_1'(x))\,f_1(x)}{(\lambda_0(x) + \lambda_1(x))^2}$$

Since we assumed that the functions $f_j(x)$ and $\lambda_j(x)$ for $j \in \{0,1\}$ and their respective derivatives are bounded, we conclude that $s_0(x)f_1(x)$ is Lipschitz continuous with the Lipschitz constant:

$$L_0 := \max_{\xi \in R} \left| s_0'(\xi)f_1(\xi) + s_0(\xi)f_1'(\xi) \right|.$$

Analogously, one can show that $s_1(x)f_0(x)$ is Lipschitz continuous with the Lipschitz constant:

$$L_1 := \max_{\xi \in R} \left| s_1'(\xi) f_0(\xi) + s_1(\xi) f_0'(\xi) \right|.$$

To verify the Lipschitz continuity of $\gamma(x)\left(f_0(x) - f_1(x)\right)$, we compute its derivative by:

$$\gamma'(x)\left(f_0(x) - f_1(x)\right) + \gamma(x)\left(f_0'(x) - f_1'(x)\right).$$

Since

$$\gamma'(x) = \sqrt{\frac{\epsilon\left(\lambda_0(x) + \lambda_1(x)\right)^3}{2\lambda_0(x)\lambda_1(x)}} \left(\frac{(\lambda_0(x))^2 \lambda_1'(x) - 2\lambda_0(x)\lambda_1(x)\left(\lambda_1'(x) - \lambda_0'(x)\right) + \lambda_1(x))^2 \lambda_0'(x)}{\left(\lambda_0(x) + \lambda_1(x)\right)^4} \right),$$

one observes that $\gamma'(x)$ is bounded. Therefore, we find that $\gamma(x)\left(f_0(x) - f_1(x)\right)$ is Lipschitz continuous with the Lipschitz constant:

$$\Gamma := \max_{\xi \in R} \left| \gamma'(\xi)\left(f_0(\xi) - f_1(\xi)\right) + \gamma(\xi)\left(f_0'(\xi) - f_1'(\xi)\right) \right|.$$

Combining our previous steps, we conclude that for all $x, y \in R$:

$$\begin{aligned} |f(x) - f(y)| + |g(x) - g(y)| \leq & |s_0(x)f_1(x) + s_1(x)f_0(x) - s_0(y)f_1(y) - s_1(y)f_0(y)| \\ & + |\gamma(x)(f_0(x) - f_1(x)) - \gamma(y)(f_0(y) - f_1(y))| \\ \leq & L_0 |x - y| + L_1 |x - y| + \Gamma |x - y| \\ = & \underbrace{(L_0 + L_1 + \Gamma)}_{:=\tilde{C}_1} |x - y| \end{aligned}$$

Restriction on Growth:

To verify the restriction on growth condition, we compute:

$$
\begin{aligned}
&|s_1(x)f_0(x) + s_0(x)f_1(x)|^2 \\
&= |s_1(x)f_0(x) + s_1(0)f_0(0) - s_1(0)f_0(0) + s_0(x)f_1(x) + s_0(0)f_1(0) - s_0(0)f_1(0)|^2 \\
&\leq (|s_1(x)f_0(x) - s_1(0)f_0(0)| + |s_1(0)f_0(0)| + |s_0(x)f_1(x) - s_0(0)f_1(0)| + |s_0(0)f_1(0)|)^2 .
\end{aligned}
$$

As a result of Lipschitz continuity of $s_1(x)f_0(x)$ and $s_0(x)f_1(x)$, we find:

$$
\begin{aligned}
|s_1(x)f_0(x) + s_0(x)f_1(x)|^2 &\leq (L_0 |x| + |s_1(0)f_0(0)| + L_1 |x| + |s_0(0)f_1(0)|)^2 \\
&\leq \max \left\{ (L_0 + L_1)^2 , (|s_1(0)f_0(0)| + |s_0(0)f_1(0)|)^2 \right\} \cdot (|x| + 1)^2 .
\end{aligned}
$$

With the estimation $2|x| \leq |x|^2 + 1$, the expression above simplifies to:

$$
|s_1(x)f_0(x) + s_0(x)f_1(x)|^2 \leq 2 \max \left\{ (L_0 + L_1)^2 , (|s_1(0)f_0(0)| + |s_0(0)f_1(0)|)^2 \right\} \cdot \left(|x|^2 + 1 \right) .
$$

Analogously, as a result of the Lipschitz continuity of $\gamma(x)\left(f_0(x) - f_1(x)\right)$, we obtain:

$$
|\gamma(x)(f_0(x) - f_1(x)|^2 \leq 2 \max \left\{ \Gamma^2, (|\gamma(0)f_0(0)| + |\gamma(0)f_1(0)|)^2 \right\} \left(1 + |x|^2 \right) .
$$

Therefore, by combining the previous steps we conclude:

$$
|s_1(x)f_0(x) + s_0(x)f_1(x)|^2 + |\gamma(x)(f_0(x) - f_1(x)|^2 \leq \tilde{C_2}^2 \left(1 + |x|^2 \right) ,
$$

where

$$\tilde{C_2}^2 := 2\max\left\{(L_0+L_1)^2, (|s_1(0)f_0(0)|+|s_0(0)f_1(0)|)^2\right\}$$
$$+ 2\max\left\{\Gamma^2, (|\gamma(0)f_0(0)|+|\gamma(0)f_1(0)|)^2\right\}.$$

Setting

$$C := \max\left\{\tilde{C}_1, \tilde{C}_2\right\},$$

the Lipschitz condition and the restriction on growth condition are satisfied for the constant $C > 0$ with the assumptions made on the functions $f_j(x)$ and $\lambda_j(x)$ for $j \in \{0, 1\}$ and their respective derivatives.

Bibliography

[1] L. Arnold. *Stochastic Differential Equations*. Wiley, New York, 1973.

[2] E.-K. Boukas. *Stochastic Switching Systems: Analysis and Design*. Birkhäuser, Boston, 2006.

[3] M. De La Chevrotiere. On a jump markovian model for a gene regulatory network. Master's thesis, University of Victoria, Canada, 2008.

[4] L. C. Evans. An introduction to stochastic differential equations.

[5] C. W. Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Springer-Verlag, New York, 3rd edition, 2004.

[6] R. Griego and R. Hersh. Theory of random evolutions with applications to partial differential equations. *Transactions of the American Mathematical Society*, 156:405–418, May 1971.

[7] R. Illner. Modeling with dynamical systems and kinetic equations. *Rivista di Mathematica della Universita di Parma*, 6(7):207–226, 2007.

[8] R. Illner and M. De La Chevrotiere. On random dynamical systems and levels of their description, 2012.

[9] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23. Springer, Berlin; New York, 3rd edition, 1999.

[10] Nobelprize.org. The prize in economics 1997 - press release, 14 October 1997.

[11] B. Oeksendal. *Stochastic Differential Equations: An Introduction with Applications*. Springer, New York, 6th edition, 2003.

[12] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes: The Art of Scientific Computing (FORTRAN Version)*. Cambridge Press, Cambridge, 1st edition, 1989.

[13] H. Risken. *The Fokker-Planck Equation: Methods of Solution and Applications*, volume 18. Springer-Verlag, Berlin; New York, 2nd edition, 1989.

[14] S. M. Ross. *A First Course in Probability*. Pearson Prentice Hall, 8th edition, 2010.

[15] S. M. Ross. *Introduction to Probability Models*. Elsevier Science, 10th edition, 2010.