

Title	Adaptive mesh construction for the numerical solution of stochastic differential equations with Markovian switching
Authors	O'Donovan, Kate
Publication date	2024
Original Citation	O'Donovan, K. 2024. Adaptive mesh construction for the numerical solution of stochastic differential equations with Markovian switching. MRes Thesis, University College Cork.
Type of publication	Masters thesis (Research)
Rights	© 2024, Kate O'Donovan. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-10-07 00:17:14
Item downloaded from	https://hdl.handle.net/10468/17417

Ollscoil na hÉireann, Corcaigh
National University of Ireland, Cork



**Adaptive Mesh Construction for the
Numerical Solution of Stochastic Differential
Equations with Markovian Switching**

A thesis presented by

Kate O'Donovan

for the degree of

Masters by Research

University College Cork

School of Mathematical Sciences

Head of School: Dr. Kevin Hayes

Supervisor: Dr. Cónall Kelly

2024

Contents

1	Introduction	5
2	Mathematical preliminaries	10
2.1	Probability spaces	10
2.2	Random variables	11
2.3	Expectation and conditional expectation	11
2.4	Stochastic processes and filtrations	13
2.5	Standard Brownian motion	14
2.6	Stochastic integrals and Itô processes	15
2.7	Stochastic differential equations	17
2.7.1	Strong solution of Itô type SDEs	17
2.7.2	Existence and uniqueness of solutions of Itô type SDEs	18
2.7.3	Moment properties of the solutions of SDEs	19
3	Numerical Methods for SDEs	21
3.1	The Euler-Maruyama method	21
3.2	The Milstein method	22
3.3	Strong and weak convergence	23
3.4	Strong convergence of the EM method over a random mesh .	24
4	Adaptive numerical solution of SDEs with Markovian switching	32
4.1	SDEs with Markovian switching	33
4.1.1	Existence and uniqueness of solutions of Itô type SDEwMS	34
4.1.2	Characterisation of waiting times	35
4.2	Constructing the numerical method	36

4.2.1	Structure of an adaptive mesh	36
4.2.2	An explicit discretisation scheme for an SDEwMS . .	37
4.2.3	A hybrid adaptive numerical scheme: general form .	38
4.2.4	A hybrid adaptive numerical scheme: particular form	40
4.3	Main theorem on strong convergence	40
5	Numerical Analysis	48
5.1	Implementation of an adaptive numerical method for SDEwMS	48
5.2	Example: a stochastic model of telomere length with Marko- vian switching	49
5.2.1	Numerical Results: Fixed Initial Length L_0	50
5.2.2	Numerical Results: Uniformly Distributed Initial Lengths L_0	50
6	Conclusions	55

DECLARATION

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism and intellectual property.

Abstract

In this dissertation we demonstrate an approach to the numerical solution of nonlinear stochastic differential equations with Markovian switching. Such equations describe the stochastic dynamics of processes where the drift and diffusion coefficients are subject to random state changes according to a Markov chain with finite state space. We propose a variant of the Jump Adapted-Adaptive approach introduced by K, Lord, & Sun (2024) to construct nonuniform meshes for explicit numerical schemes that adjust timesteps locally to rapid changes in the numerical solution and which also incorporate the switching times of an underlying Markov chain as mesh-points. It is shown that a hybrid scheme using such a mesh that combines an efficient explicit method (to be used frequently) and a potentially inefficient backstop method (to be used occasionally) will display strong convergence in mean-square of order δ if both methods satisfy a mean-square consistency condition of the same order in the absence of switching. We demonstrate the construction of an order $\delta = 1$ method of this type and apply it to generate empirical distributions of a nonlinear SDE model of telomere length in DNA replication.

Chapter 1

Introduction

In many branches of mathematics, systems are subject to various types of uncertainty or “noise”. This white noise can be reasonably mathematically interpreted as the derivative of Brownian motion. As we will see, Brownian motion is nowhere differentiable. We will use the stochastic nature of Brownian motion to define the Itô stochastic integral, which will allow us to analyse and solve systems with noise. Such systems are known as stochastic differential equations (SDEs). SDEs assume that the coefficients of the functions in the equation are constants, this is not always the case. It has been shown that in some situations these coefficients can be modelled as random processes, for example the average rate of return and volatility in the Black-Scholes model (See Mao & Yuan [14, Chapter 1]). One type of random fluctuation in coefficients is Markovian switching. In this case the coefficient functions switch from their current state to the next according to the evolution of a Markov chain.

In this dissertation we examine a method of constructing adaptive meshes for the efficient numerical solution of d -dimensional system of stochastic differential equations with Markovian switching (SDEwMS) of the form

$$\begin{aligned} dX(t) &= f(X(t), r(t))dt + g(X(t), r(t))dW(t), \quad 0 \leq t \leq T, \\ X(0) &= x_0 \in \mathbb{R}^d. \end{aligned} \tag{1.1}$$

In (1.1), the drift and diffusion coefficients f and g take the form

$$f : \mathbb{R}^d \times \mathbb{S} \rightarrow \mathbb{R}^d, \quad \text{and} \quad g : \mathbb{R}^d \times \mathbb{S} \rightarrow \mathbb{R}^{d \times m}.$$

The stochastic dynamics of a process governed by (1.1) are influenced by two independent sources of noise. The process r dictates changes occurring at random intervals in the functional form of f and g , and is described by a scalar continuous time Markov chain taking values in the set $\mathbb{S} = \{1, 2, \dots, L\}$. The process W is an m -dimensional Brownian motion and captures the effect of a diffusion process on trajectories. We characterise both mathematically in Chapters 2 and 4.

We are motivated by the following example where a scalar SDE is used to model the shortening over time of telomere length in DNA replication in jackdaws:

$$dL(t) = -(c + aL(t)^2) dt + \sqrt{\frac{1}{3}aL(t)^3} dW(t), \quad (1.2)$$

where W is a scalar standard Brownian motion. As presented in [2], the model arises as the Langevin equation associated with the Fokker-Planck PDE

$$\frac{\partial p}{\partial t} = \frac{\partial}{\partial l} \left\{ \left(c + \frac{1}{2}al^2 \right) p \right\} + \frac{1}{2} \frac{\partial^2}{\partial l^2} \left\{ \left(\frac{1}{3}al^3 \right) p \right\},$$

which governs the probability density function $p(t, l)$ for the telomere length l at time t . By simulating an ensemble of trajectories of (1.2) we can produce a density histogram of lengths $l = L$ at a fixed time $t = T$ which serves to numerically approximate $p(T, L)$. The parameters c and a control the underlying decay rate of the length and the intensity at which random breaks occur in the telomere respectively.

The general form of the SDE (1.1) without Markovian switching, can be written

$$\begin{aligned} dX(t) &= f(X(t))dt + g(X(t))dW(t), \quad 0 \leq t \leq T, \\ X(0) &= x_0 \in \mathbb{R}^d. \end{aligned} \quad (1.3)$$

We can see that the telomere length SDE (1.2) is a special case of (1.3) with $d = m = 1$, $f(x) = -(c + ax^2)$, and $g(x) = (ax^3/3)^{1/2}$.

Suppose we wish to solve (1.3) numerically over an arbitrary mesh $\{t_n\} = \{0 = t_0, t_1, \dots, t_{N-1}, t_N = T\}$. Over any step $[t_n, t_{n+1}]$, solutions of (1.3) satisfy the integral equation

$$X(s) = X(t_n) + \int_{t_n}^s f(X(r))dr + \int_{t_n}^s g(X(r))dW(r), \quad s \in [t_n, t_{n+1}]. \quad (1.4)$$

By fixing the solution at the beginning of the interval $[t_n, t_{n+1}]$, we achieve a continuous form of the explicit Euler-Maruyama approximation

$$\bar{X}(s) = X_n + \int_{t_n}^s f(\bar{X}(t_n)) dr + \int_{t_n}^s g(\bar{X}(t_n)) dW(r), \quad s \in [t_n, t_{n+1}]. \quad (1.5)$$

Setting $X_n := \bar{X}(t_n)$, we can write the scheme in its computational form as

$$X_{n+1} = X_n + h_{n+1}f(X_n) + g(X_n)\Delta W_{n+1}, \quad n = 0, \dots, N-1, \quad (1.6)$$

where $h_{n+1} = t_{n+1} - t_n$ and $\Delta W_{n+1} = W(t_{n+1}) - W(t_n)$. Suppose there exists a constant maximum stepsize $h_{\max} \in (0, \infty)$ such that $h_n \leq h_{\max}$ for all $n = 0, \dots, N$. We say that solutions of (1.5) (or equivalently (1.6)) converge strongly to solutions of (1.3) in L^p with order δ if and only if

$$\max_{t \in [0, T]} (\mathbb{E} [\|X(t) - \bar{X}(t)\|^p])^{1/p} \leq Ch_{\max}^\delta.$$

If $p = 2$ then we say that (1.5) converges strongly in L_2 , in mean-square. It is the case that the Euler-Maruyama method (1.5) converges in mean-square to (1.3) with order $\delta = 1/2$ when f, g are globally Lipschitz and $h_n \equiv h$.

Hutzenthaler et al [5] confirmed that solutions of (1.6) with constant stepsize $h_n \equiv h$ fails to converge strongly in the strong sense to solutions of (1.3) if either f or g grow superlinearly in norm. Indeed this is the case for the telomere model (1.2). To address this issue, many novel explicit methods have been proposed that converge strongly even for highly nonlinear drift and diffusion coefficients; these include tamed, projected, truncated, and adaptive variants of both Euler-Maruyama and Milstein methods (see, for example [1, 3, 6, 13, 16]). In particular we highlight the adaptive timesteping approach that was introduced for the Euler-Maruyama method in [10] and extended to the Milstein method in [11], where it was used to construct a numerical method converging strongly with order $\delta = 1$ and to generate trajectories of the telomere SDE (1.2).

We wish to extend this capability to models where the coefficients f and g undergo Markovian switching. For example in the case of the telomere SDE (1.2) it may be that each of the model parameters in the pair (c, a) undergo transition between two distinct values according to a Markov process. Alternatively one can view Markovian switching in the model parameters as a form of stress testing to investigate the robustness of sample distributions to random changes in these parameters.

We take as our template the approach to mesh construction in [9], where it was shown that an adaptive numerical scheme that satisfies a particular mean-square consistency bound can be extended to ensure strong convergence in mean-square for SDEs with jump perturbations (SJDEs) as long as the jump times are incorporated into the adaptive mesh in a particular way. In such a case the mesh is adapting both to the evolving behaviour of the approximate trajectory and directly to jumps when they occur. We refer to these as Jump Adapted-Adaptive Methods (JA-AMs).

In this dissertation, we show that an analogue of the type of mesh used for JA-AMs also ensures convergence of explicit numerical schemes satisfying a similar mean-square consistency condition to SDEwMS governed by (1.1), as long as we include the times at which the Markov chain r switches between distinct states. We prove strong convergence of such schemes in mean-square, noting that the major challenge of this analysis is dealing with the fact that the number of Markov transitions on any interval is finite but unbounded, and we handle it in the same manner as in the jump case [9]. Finally we demonstrate the practical application of this approach on a variant of SDE model of telomere shortening (1.2) with stochastic dynamics governed by coefficients that undergo transitions at random times.

In Chapter 2 we provide a complete mathematical description of the setting for SDEs without switching, including of W , and provide conditions on f and g for which unique solutions of such equations exist on $[0, T]$ for $T < \infty$. In Chapter 3 we describe the Euler-Maruyama (EM) and Milstein methods for SDEs over a random time mesh. We define strong and weak convergence for numerical methods and present a proof for strong convergence of the EM method. In Chapter 4 we extend the complete mathematical description of the setting for SDEs to include SDEwMS. We describe how an adaptive mesh can be constructed that adapts both to the nonlinear functional response affecting the numerical solutions of (1.1) and which incorporate the switching times of r . We describe the general form of our adaptive scheme and present a version of it based upon the Milstein method. In Section 4.3 we state and prove our main strong convergence theorem, which in particular confirms that the Milstein scheme applied over this class of adaptive meshes will converge strongly in L^2 with order $\delta = 1$ to solutions of (1.1). Finally in Chapter 5 we describe the practical implementation of the scheme and demonstrate its use to generate sample histograms of telomere length in base pairs using a nonlinear SDEwMS model. The python code used for the analysis in Chapter 5 can be found in the Github repository

<https://github.com/kateo25/SDEs-with-Markovian-Switching>

The main result of this dissertation has been accepted for publication in Delay and Stochastic Differential Equations in Modelling in Finance and Life sciences, and Engineering, ICIAM2023-3 and will be published in summer 2025, see prepublication here [12]

Chapter 2

Mathematical preliminaries

In this section we will give a brief overview of some properties and definitions that will be relevant for the analysis which come later in this dissertation. These include probability spaces, stochastic processes and stochastic differential equations.

2.1 Probability spaces

Probability theory deals with mathematical models of trials whose outcomes depend on chance. All possible outcomes are grouped together to form a set Ω , called the sample space, with typical element $\omega \in \Omega$, referred to as an outcome. Not every subset of Ω is observable or interesting. We will only group observable or interesting subsets of Ω , known as events, together as a family called $\mathcal{F}$.

$\mathcal{F}$ has the following properties

- $\emptyset \in \mathcal{F}$, where $\emptyset$ denotes the empty set;
- $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$, where $A^c = \Omega - A$ is the complement of $A \in \Omega$;
- $\{A_i\}_{i \geq 1} \subset \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

A family with these properties is called a σ -algebra. The pair $(\Omega, \mathcal{F})$ is called a measure space.

2.2 Random variables

A real valued function $X : \Omega \rightarrow \mathbb{R}$ is said to be $\mathcal{F}$ -measurable if

$$\{\omega : X(\omega) \leq a\} \in \mathcal{F}, \quad \text{for all } a \in \mathbb{R}.$$

This function X is also known as a random variable.

If C is a family of subsets of Ω , then there exists a smallest σ -algebra $\sigma(C)$ on Ω which contains C . This $\sigma(C)$ is called the σ -algebra generated by C . If $\Omega = \mathbb{R}^d$ and C is the family of all open sets in $\mathbb{R}^d$ the $\mathcal{B}^d = \sigma(C)$ is called the Borel σ -algebra and the elements of $\mathcal{B}^d$ are called a Borel set. If $X, Y : \Omega \rightarrow \mathbb{R}^d$ are two given functions, Y is $\sigma(X)$ -measurable if and only if there exists a Borel measurable function $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ such that $Y = g(X)$.

A probability measure P on a measurable space $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \rightarrow [0, 1]$ such that

- $P(\Omega) = 1$;
- for any disjoint sequence $(A_i)_{i \geq 1} \subseteq \mathcal{F}$,

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

The triple $(\Omega, \mathcal{F}, P)$ is called a probability space. We set

$$\bar{\mathcal{F}} = \{A \subset \Omega : \text{there exists } B, C \in \mathcal{F} \text{ such that } B \subset A \subset C \text{ and } P(B) = P(C)\}.$$

Then $\bar{\mathcal{F}}$ is a σ -algebra and is called the completion of $\mathcal{F}$. If $\mathcal{F} = \bar{\mathcal{F}}$, the probability space $(\Omega, \mathcal{F}, P)$ is said to be complete.

An indicator function $\mathcal{I}_A$ of the set $A \subset \Omega$ is defined by

$$\mathcal{I}_A(w) = \begin{cases} 1, & \omega \in A; \\ 0, & \omega \notin A. \end{cases}$$

2.3 Expectation and conditional expectation

Next we will look at conditional expectation of random variables, this is crucial for later proofs relating to the main results of this dissertation. First let's define the expectation of a random variable.

Let $(\Omega, \mathcal{F}, P)$ be a probability space, then if X is a real-valued random variable and is integrable with respect to the probability measure P , then the number

$$\mathbb{E}[X] = \int_{\Omega} X(\omega) dP(\omega).$$

is called the expectation of X with respect to P . This integral is known as a Lebesgue integral.

For $p \in (0, \infty)$ let $L^p = L^p(\Omega, \mathbb{R}^d)$ be a family of $\mathbb{R}^d$ -valued random variables X with $\mathbb{E}[\|X\|^p] < \infty$. In L^1 , we have $\|\mathbb{E}[X]\| \leq \mathbb{E}[\|X\|]$, and a very useful inequality known as Hölder's inequality follows.

Lemma 2.3.1. (Hölder's inequality)

Let $X \in L^p$ and $Y \in L^q$, suppose $p > 1$ and $\frac{1}{p} + \frac{1}{q} = 1$. Then

$$\|\mathbb{E}[X^T Y]\| \leq (\mathbb{E}[\|X\|^p])^{\frac{1}{p}} (\mathbb{E}[\|Y\|^q])^{\frac{1}{q}}. \quad (2.1)$$

Hölder's inequality (2.1) implies that if $X \in L^p$ then for $0 < r < p < \infty$

$$(\mathbb{E}[\|X\|^r])^{\frac{1}{r}} \leq (\mathbb{E}[\|X\|^p])^{\frac{1}{p}}.$$

Let $X \in L^1(\Omega; \mathbb{R})$ and let $\mathcal{G} \subset \mathcal{F}$ be a sub- σ -algebra of $\mathcal{F}$ so $(\Omega, \mathcal{G})$ is a measure space. In general this X is not $\mathcal{G}$ -measurable, so we seek an integrable $\mathcal{G}$ -measurable random variable Y such that

$$\mathbb{E}[\mathcal{I}_G Y] = \mathbb{E}[\mathcal{I}_G X],$$

and

$$\int_G Y(\omega) dP(\omega) = \int_G X(\omega) dP(\omega), \quad \text{for all } G \in \mathcal{G}.$$

By the Radon-Nikodym theorem, there exists only one such Y , almost surely unique. This is called the conditional expectation of X under the condition $\mathcal{G}$

$$Y = \mathbb{E}[X|\mathcal{G}].$$

If $\mathcal{G}$ is the σ -algebra generated by Y then

$$\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X|Y],$$

and it follows that

$$\mathbb{E}[\mathbb{E}[X|\mathcal{G}]] = \mathbb{E}[X].$$

Some other useful properties from conditional expectations include:

- $\|\mathbb{E}[X|\mathcal{G}]\| \leq \mathbb{E}[\|X\||\mathcal{G}]$.
- $\mathcal{G} = \{\emptyset, \Omega\}$, $\mathbb{E}[X|\mathcal{G}] = \mathbb{E}[X]$.
- When X is $\mathcal{G}$ -measurable, we can take out what is known, $\mathbb{E}[X|\mathcal{G}] = X$.
- Linearity of conditional expectation implies $\mathbb{E}[aX+bY|\mathcal{G}] = a\mathbb{E}[X|\mathcal{G}] + b\mathbb{E}[Y|\mathcal{G}]$, for $a, b \in \mathbb{R}$.
- The tower property states that if $\mathcal{G}_1 \subset \mathcal{G}_2 \subset \mathcal{F}$ then $\mathbb{E}[\mathbb{E}[X|\mathcal{G}_2]|\mathcal{G}_1] = \mathbb{E}[X|\mathcal{G}_1]$.

2.4 Stochastic processes and filtrations

Stochastic processes are families of random variables, as defined below, used to mathematically model systems that have a random element. They form an integral part of the stochastic differential equations (SDEs) we will be studying in section 2.7.

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A family $(X(t))_{t \in [0, T]}$ of $\mathbb{R}^d$ -valued random variables is called a stochastic process on $[0, T]$ with state space $\mathbb{R}^d$. For each fixed $t \in [0, T]$ we have a random variable

$$\Omega \ni \omega \rightarrow X_t(\omega) \in \mathbb{R}^d,$$

and for each fixed $\omega \in \Omega$ we have a function

$$[0, T] \ni t \rightarrow X_t(\omega) \in \mathbb{R}^d.$$

which is called a sample path of the process.

A filtration can be thought of as a mathematical representation of the flow of time as it appears to an observer of a stochastic process. A precise mathematical definition follows. A filtration is a family $(\mathcal{F})_{t \geq 0}$ of increasing sub- σ -algebras of $\mathcal{F}$ i.e. $\mathcal{F}_t \subset \mathcal{F}_s \subset \mathcal{F}$ for all $0 \leq t < s \leq \infty$. The filtration is said to be right-continuous if $\mathcal{F}_t = \bigcap_{s > t} \mathcal{F}_s$ for all $t \geq 0$. When the probability space is complete the filtration is said to satisfy the *usual conditions* if it is right continuous and $\mathcal{F}_0$ contains all P-null sets. A probability space may be equipped with a filtration denoted as $(\Omega, \mathcal{F}, (\mathcal{F})_{t \geq 0}, P)$, we will refer to this as a filtered probability space, or simply a quadruple.

An $\mathbb{R}^d$ -valued stochastic process $(X(t))_{t \geq 0}$ is said to be $\mathcal{F}_t$ -adapted if for every t , X_t is $\mathcal{F}_t$ -measurable.

2.5 Standard Brownian motion

Brownian motion is one of the specific stochastic process we will use throughout this dissertation for the construction of SDEs. Brownian motion is named after Robert Brown, a scottish botanist who first described the process after observing the movement of pollen grains suspended in water. It is often used in models for financial and biological processes.

Let $(\Omega, \mathcal{F}, (\mathcal{F})_{t \in [T, 0]}, P)$ be a filtered probability space. A one-dimensional standard Brownian motion on this space is a real-valued continuous $\mathcal{F}_t$ -adapted process $(W(t))_{t \in [0, T]}$ with the following properties

1. $W(0) = 0$ a.s.
2. For $0 \leq s < t < \infty$, the increment $W(t) - W(s)$ is normally distributed, with respect to the measure P , with mean zero and variance $t - s$.
3. For $0 \leq s < t < \infty$, the increment $W(t) - W(s)$ is independent of $\mathcal{F}_s$.

If W is a Brownian motion and $0 \leq t_0 < t_1 < \dots < t_k < \infty$, then the increments $W(t_i) - W(t_{i-1})$, where $1 \leq i \leq k$, are mutually pairwise independent. The distribution of $W(t) - W(s)$ depends only on the difference $t - s$, and we say that the Brownian motion has stationery increments.

Notice the above definition is associated with a specific filtration $\mathcal{F}_t$, it is possible to define W with an associated filtration in which case property three above is replaced by the fact that W has independent increments. In this case let $\mathcal{F}_t^W$ be the σ -algebra generated by $\{W(s) : 0 \leq s < t\}$, this induces a natural filtration $(\mathcal{F}_t^W)_{t \in [0, T]}$. If it's then the case that $\mathcal{F}_t^W \subset \mathcal{F}$ for $t > 0$, and $W(t) - W(s)$ is independent of $\mathcal{F}_s$ when $0 \leq s < t < \infty$, then W is a Brownian motion with respect to the filtration $(\mathcal{F}_t)_{t \in [0, T]}$.

Next we look at m -dimensional Brownian motion

Definition 2.5.1. Let $\{W(t) = W^1(t), \dots, W^m(t)\}_{t \in [0, T]}$ be an m -dimensional process. This process is called an m -dimensional Brownian motion if every $\{W^i(t)\}_{t \in [0, T]}$ is a one-dimensional Brownian motion and $\{W^1(t)\}_{t \in [0, T]}, \dots, \{W^m(t)\}_{t \in [0, T]}$ are mutually pairwise independent.

2.6 Stochastic integrals and Itô processes

Before we begin to look at stochastic differential equations (SDEs), let's first look at stochastic integrals and Itô process. As Brownian motion is nowhere differentiable we cannot define a stochastic integral in the ordinary way. For this we will need Itô calculus, namely the Itô stochastic integral and Itô's formula, which were first defined by K. Itô in 1949. In this section we will follow the treatment in Mao's book [14].

Let $(\Omega, \mathcal{F}, (\mathcal{F})_{t \in [0, T]}, P)$ be a complete, filtered probability space satisfying the usual conditions.

The following definition will be useful for later definitions of Itô integrals

Definition 2.6.1. Let $0 \leq a < b < \infty$. The set of all real-valued measurable $\mathcal{F}_t$ -adapted processes $f = (f(t))_{a \leq t \leq b}$ such that

$$\|f\|_{a,b}^2 = \mathbb{E} \left[\int_a^b |f(t)|^2 dt \right] < \infty,$$

will be denoted by $\mathcal{M}^2([a, b]; \mathbb{R})$.

Here we say $\mathcal{L}^p(\mathbb{R}_+, \mathbb{R})$ is the family of $\mathbb{R}^d$ -valued, $\mathcal{F}_t$ -adapted processes $(f(t))_{t \in [0, T]}$ such that $\int_a^b |f(t)|^p dt < \infty$.

Next we must define a simple process, which we will use to construct Itô's integral

Definition 2.6.2. A real valued stochastic process $g = (g(t))_{a \leq t \leq b}$ is called a simple process if there exists a partition $a = t_0 < t_1 < \dots < t_k = b$ of $[a, b]$, and bounded random variables ξ_i , $0 \leq i \leq k-1$ such that ξ_i is $\mathcal{F}_{t_i}$ -measurable and

$$g(t) = \xi_0 I_{[t_0, t_1]}(t) + \sum_{i=1}^{k-1} \xi_i I_{(t_i, t_{i+1}]}(t), \quad (2.2)$$

the family of all processes of this type, will be denoted as $\mathcal{M}_0([a, b]; \mathbb{R})$.

Here we define the first part of Itô's integral

Definition 2.6.3. For a simple process g with the form of (2.2) in $\mathcal{M}_0([a, b]; \mathbb{R})$ we call

$$\int_a^b g(t) dW_t = \sum_{i=0}^{k-1} \xi_i (W(t_{i+1}) - W(t_i)),$$

the stochastic integral of g with respect to the Brownian motion W or the Itô integral.

Next we will need this lemma to complete our definition of the Itô integral

Lemma 2.6.1. For any $f \in \mathcal{M}^2([a, b]; \mathbb{R})$ there exists a sequence $\{g_n\}$ of simple processes such that

$$\lim_{n \rightarrow \infty} E \int_a^b |f(t) - g_n(t)|^2 dt = 0.$$

The definition of the Itô integral is then as follows

Definition 2.6.4. Let $f \in \mathcal{M}^2([a, b]; \mathbb{R})$. The Itô integral with respect to W is defined by

$$\int_a^b f(t) dW(t) = \lim_{n \rightarrow \infty} \int_a^b g_n(t) dW(t) \text{ in } L^2(\Omega, \mathbb{R}),$$

where $\{g_n\}$ is a simple process and lemma 2.6.1 holds and $L^2(\Omega, \mathbb{R})$ is the family of $\mathbb{R}^d$ -valued random variables ξ with $\mathbb{E} \|\xi\| < \infty$ and we say that $x(t)$ has the stochastic differential $dX(t)$ on $t \geq 0$ given by

$$dX(t) = f(t)dt + g(t)dW(t).$$

An Itô process is defined as

Definition 2.6.5. A one-dimensional Itô process is a continuous adapted process $X(t)$ on $t \geq 0$ of the form

$$X(t) = X(0) + \int_0^t f(s)ds + \int_0^t g(s)dW(s),$$

where $f \in \mathcal{L}^1(\mathbb{R}_+, \mathbb{R})$ and $g \in \mathcal{L}^2(\mathbb{R}_+, \mathbb{R})$.

And finally we now introduce the one dimensional Itô formula. The Itô formula is the stochastic version of the chain rule for Itô integrals, it is very useful in stochastic analysis and for evaluating Itô integrals.

Definition 2.6.6. Let $X(t)$ be an Itô process as defined in Definition 2.6.5 and let $V \in C^{2,1}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$. Then $V(X(t), t)$ is an Itô process with the stochastic differential given by

$$dV(X(t), t) = [V_t(X(t), t) + V_x(X(t), t)f(t) + \frac{1}{2}V_{xx}(X(t), t)g^2(t)]dt + V_x(X(t), t)g(t)dW(t) \quad \text{a.s.} \quad (2.3)$$

Itô's isometry relates the second moment of an Itô integral to a Riemann integral and will be used when developing the properties of the SDE trajectories between random timesteps. Here we state the one-dimensional version as this is what we will use in our proof.

Lemma 2.6.2. (Itô's isometry)

Let $X \in L^2$ and W be a scalar standard Brownian motion

$$\mathbb{E} \left[\left| \int_s^t (X(u))dW(u) \right|^2 \right] = \int_s^t \mathbb{E} [|X(u)|^2] du. \quad (2.4)$$

2.7 Stochastic differential equations

Let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ be a complete, filtered probability space satisfying the usual conditions.

2.7.1 Strong solution of Itô type SDEs

Let $W(t) = (W_1(t), \dots, W_m(t))^T$, $t \in [0, T]$ be an m -dimensional Brownian motion defined on the space. Let $0 \leq t_0 < T < \infty$. Let x_0 be an $\mathcal{F}_{t_0}$ -measurable $\mathbb{R}^d$ -valued random variable such that $\mathbb{E}|x_0|^2 < \infty$. Let $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $g : \mathbb{R}^d \rightarrow \mathbb{R}^{d \times m}$ be both Borel measurable. Consider an SDE of Itô type

$$dX(t) = f(X(t))dt + g(X(t))dB(t), \quad 0 \leq t \leq T, \quad (2.5)$$

with initial value $X(0) = x_0$

Definition 2.7.1. An $\mathbb{R}^d$ -valued stochastic process $(X(t))_{t \in [0, T]}$ is called a solution of equation (2.5) if it has the following properties:

1. $(X(t))_{t \in [0, T]}$ is continuous and $\mathcal{F}_t$ -adapted.
2. $(f(X(t)))_{t \in [0, T]} \in \mathcal{L}^1([0, T]; \mathbb{R}^d)$ and $(g(X(t)))_{t \in [0, T]} \in \mathcal{L}^2([0, T]; \mathbb{R}^{d \times m})$.
3. for any $t \in [0, T]$, equation

$$X(t) = X(0) + \int_0^t f(X(s))ds + \int_0^t g(X(s))dB(s),$$

holds with probability 1.

2.7.2 Existence and uniqueness of solutions of Itô type SDEs

We identify the minimal conditions on the drift, f , and diffusion, g , coefficients that will ensure the existence of unique solutions of (2.5) over each interval $[0, T]$. We require that both satisfy a local Lipschitz condition:

Assumption 2.7.1. For every integer $k \geq 1$, there exists a positive constant h_k such that, for all $x, y \in \mathbb{R}^d$ with $|x| \vee |y| \leq k$

$$\|f(x) - f(y)\|^2 \vee \|g(x) - g(y)\|^2 \leq \bar{h}_k \|x - y\|^2,$$

Consider also the following linear growth condition:

Assumption 2.7.2. f and g satisfy a linear growth bound if for all $x \in \mathbb{R}^d$ there exists a constant $0 < K_3 < \infty$ such that

$$\|f(x)\|^2 \vee \|g(x)\|^2 \leq K_3(1 + \|x\|^2).$$

Under Assumptions 2.7.1 and 2.7.2 there exists a unique solution to (2.5) on each $[0, T]$: See Mao [14, Theorem 3.16].

It is possible to replace the linear growth condition with a weaker constraint. If Assumption 2.7.1 holds, so that f and g satisfy a local Lipschitz condition, and they also satisfy the following monotonicity condition

Assumption 2.7.3. There exists a positive constant K such that for all $(x) \in \mathbb{R}^d$

$$x^T f(x) + \frac{1}{2} \|g(x)\|^2 \leq K(1 + \|x\|^2).$$

Then there exists a unique solution X to Eq. (2.5) on each $[0, T]$.

2.7.3 Moment properties of the solutions of SDEs

Under the same conditions (Assumption 2.7.1 and either Assumption 2.7.2 or 2.7.3) solutions of (2.5) also satisfy the following uniform moment bounds and Hölder regularity bound.

Assumption 2.7.4. Let X be the solution of (2.5) with $d = m = 1$. Suppose f and g satisfy assumption 2.7.2. Then for any $0 < p < \infty$ there exists a constant $0 < k_3 < \infty$ such that

$$\mathbb{E} \left[\sup_{r \in [0, T]} |X(r)|^p \right] \leq k_3. \quad (2.6)$$

The Hölder regularity bound describes the variation in solutions of SDEs as measured in L^2

Lemma 2.7.1. (Hölder regularity bound)

Let X be the solution to (2.5) and suppose f and g are globally Lipschitz continuous. Then there exists a constant $0 < k_4 < \infty$ such that

$$\mathbb{E}[|X(t) - X(s)|^2] \leq k_4[t - s]$$

where $0 \leq s < t \leq T$

Proof. Using Lemma 2.6.2 and Assumption 2.7.2 we will prove a special case

where $f(x) = 0$

$$\begin{aligned}\mathbb{E} [|X(t) - X(s)|^2] &= \mathbb{E} \left[\left| \int_s^t G(X(u)) dW(u) \right|^2 \right] \\ &\leq \int_s^t \mathbb{E} [G^2(X(u))] du \\ &\leq \int_s^t \mathbb{E} [k_1(1 + X^2(u))] du\end{aligned}$$

Now applying the moment bound (2.6) in Assumption 2.7.4

$$\begin{aligned}\mathbb{E} [|X(t) - X(s)|^2] &\leq k_1 \int_s^t \mathbb{E} \left[1 + \sup_{r \in (s,t)} |X(r)|^2 \right] du \\ &\leq k_4 \int_s^t du = k_4 |t - s|\end{aligned}$$

□

Chapter 3

Numerical Methods for SDEs

In this chapter we will define the Euler-Maruyama (EM) and Milstein methods for the numerical approximation of the solution to SDEs. As it is not always possible to compute the exact solution of an SDE, numerical methods have been developed to overcome this. Here we will define the EM and Milstein methods over a random time mesh, meaning the time steps between each application of the methods is a random variable. We then prove the order-1/2 strong convergence of the EM method to the true solution of an SDE over this random mesh. This shows that the expectation of the error between the exact solution and the approximated solution is bounded by a polynomial function of the maximum stepsize of order 1/2.

3.1 The Euler-Maruyama method

Let's consider an Itô process $(x(t))_{t \in [0, T]}$ which satisfies the scalar SDE

$$X(t) = x_0 + \int_0^t f(X(s))ds + \int_0^t g(X(s))dW(s), \quad 0 \leq t \leq T. \quad (3.1)$$

with the initial value $X(0) = x_0$, where f and g satisfy Assumptions 2.7.1 and 2.7.2, and $(f(X(t)))_{t \in [0, T]} \in \mathcal{L}^1([0, T]; \mathbb{R}^d)$ and $(g(X(t)))_{t \in [0, T]} \in \mathcal{L}^2([0, T]; \mathbb{R}^{d \times m})$ and let $(W(t))_{t \in [0, T]}$ be a standard Brownian motion.

To apply a numerical method like the EM method, we need to start by constructing a discrete time approximation of the process. First we create a discretisation of time $\{0 = t_0 < \dots < t_n < \dots < t_N = T\}$. Let $\{h_{n+1}\}_{n \in \mathbb{N}}$ be

a sequence of strictly positive random time steps, where $h_{n+1} = t_{n+1} - t_n$, and have corresponding random times $\{t_n = \sum_{i=0}^n h_i\}_{n \in \mathbb{N}}$. We set $t_0 = 0$ and $t_N = T$, where N is a.s. finite. We also assume each h_{n+1} is $\mathcal{F}_{t_n}$ -measurable, and there exists a maximum timestep $h_{\max}$. This is what we will refer to as a random mesh.

The random increments of the Brownian motion are defined as follows

$$\Delta W_{n+1} = W(t_{n+1}) - W(t_n),$$

for $n = 0, 1, \dots, N$. As these increments are independent random variables with conditional mean $\mathbb{E}[\Delta W_{n+1} | \mathcal{F}_{t_n}] = 0$ and conditional variance $\mathbb{E}[(\Delta W_{n+1})^2 | \mathcal{F}_{t_n}] = h_n$. We can generate the Brownian path sequentially using independent Gaussian pseudo-random numbers for the increments in our method.

Then the Euler-Mayumama method for the SDE (3.1) is defined as

$$X_{n+1} = X_n + f(X_n)h_{n+1} + g(X_n)\Delta W_{n+1}, \quad (3.2)$$

for $n = 0, 1, \dots, N - 1$. The continuous version of this approximation for all $s \in [t_n, t_{n+1}]$ is

$$\bar{X}(s) = X_n + \int_{t_n}^s f(\bar{X}(t_n))dr + \int_{t_n}^s g(\bar{X}(t_n))dW(r). \quad (3.3)$$

3.2 The Milstein method

As with the EM method we can construct a discrete time version of the Milstein method, including the independent increments for the Brownian motion. In the case where $d = m = 1$, the discrete form of the Milstein method for the SDE (3.1) is as follows

$$\begin{aligned} X_{n+1} = X_n + f(X_n)h_{n+1} + g(X_n)\Delta W_{n+1} \\ + \frac{1}{2}g'(X_n)g(X_n)(\Delta W_{n+1}^2 - h_{n+1}), \end{aligned} \quad (3.4)$$

with corresponding continuous version

$$\begin{aligned} \bar{X}(s) = X_n + \int_{t_n}^s f(\bar{X}(t_n))dr \\ + \int_{t_n}^s g(\bar{X}(t_n))dW(r) + \int_{t_n}^s \int_{t_n}^r g'(\bar{X}(t_n))g(\bar{X}(t_n))dW(p)dW(r). \end{aligned} \quad (3.5)$$

3.3 Strong and weak convergence

As the EM and Milstein methods are approximations, there will be differences between the approximate and true solutions. If the error created by these methods can be made arbitrarily small when we reduce the maximum stepsize $h_{\max}$ and based on the rate at which the error decreases, we have methods of quantifying these errors, namely strong and weak convergence. Broadly speaking strong convergence is of the average error on each individual trajectory and weak convergence is the error of the estimated expected value, see Kelly [7, Section 8.3].

Definition 3.3.1. Let $X(t)$ be the true solution of the SDE (3.1) and suppose $(X_n)_{n=0}^N$ is a numerical approximation over a random mesh $\{0 = t_0 < \dots < t_n < \dots < t_N = T\}$, where $h_{n+1} = t_{n+1} - t_n$ and $h_n \leq h_{\max}$ for all $n \in 0, 1 \dots N$.

1. $\{X_n\}_{n=0}^N$ converges weakly with order δ to the true solution of the SDE (3.1) if there exists constants $\bar{h}, C > 0$ such that

$$|E[P(X_N)] - E[P(X(T))]| \leq Ch_{\max}^\delta \quad \text{for all } h_{\max} \leq \bar{h}$$

where C may depend on T but not on $h_{\max}$, P is some real-valued test function satisfying appropriate conditions.

2. Fix $p \geq 1$. $\{X_n\}^N$ converges strongly in p^{th} moment with order δ to the true solution of the SDE (3.1) if there exists constants $h_{\max}, C > 0$ such that

$$\left(E[|X_N - X(T)|^p]^{\frac{1}{p}}\right) \leq Ch_{\max}^\delta \quad \text{for all } h_{\max} \leq \bar{h}$$

where C may depend on T but not on h .

3.4 Strong convergence of the EM method over a random mesh

Before we begin our proof of strong convergence, we first need to introduce Gronwall's inequality. Gronwall-type inequalities are widely used in the analysis of SDEs to prove results on existence and uniqueness of solutions, boundedness, stability etc.

Lemma 3.4.1. (Gronwall's Inequality)

Let $T > 0$ and $c \geq 0$. Let u be a Borel measurable bounded nonnegative function on $[0, T]$, and let v be a nonnegative integrable function on $0, T$. If

$$u(t) \leq c + \int_0^t v(s)u(s)ds \quad \text{for all } 0 \leq t \leq T,$$

then

$$u(t) \leq c \exp \left(\int_0^t v(s)ds \right) \quad \text{for all } 0 \leq t \leq T. \quad (3.6)$$

We will make frequent use of the following elementary inequalities;

$$ab \leq \frac{1}{2}a^2 + \frac{1}{2}b^2 \quad (3.7)$$

$$(a + b)^2 \leq 2a^2 + 2b^2 \quad (3.8)$$

Theorem 3.4.1. *Let $(X(t))_{t \in [0, T]}$ be a solution of (3.1) with initial value $X(0) = x_0 \in \mathbb{R}^d$ and suppose that Assumptions 2.7.1, 2.7.3 and 2.7.4 hold. Let $\bar{X}$ be an numerical scheme as characterised in Section 3.1 with $\bar{X}_0 = x_0$ with $\delta = \frac{1}{2}$. Then there exists a constant $C(T) > 0$ such that*

$$\max_{t \in [0, T]} \left(\mathbb{E} \left[\|X(t) - \bar{X}(t)\|^2 \right] \right)^{\frac{1}{2}} \leq C(T) h_{\max}^{\delta}, \quad (3.9)$$

where C does not depend on $h_{\max}$.

Proof. For all $s \in [t_n, t_{n+1}]$, the solution of the SDE (3.1) is

$$X(s) = X(t_n) + \int_{t_n}^s f(X(r))dr + \int_{t_n}^s g(X(r))dW(r). \quad (3.10)$$

The continuous approximation of the SDE (3.1) by the EM method (3.3) is

$$\bar{X}(s) = X(t_n) + \int_{t_n}^s f(\bar{X}(t_n))dr + \int_{t_n}^s g(\bar{X}(t_n))dW(r). \quad (3.11)$$

Then subtracting (3.11) from (3.10), where $E(s)$ is the error at time s , we have

$$\begin{aligned} E(s) &:= X(s) - \bar{X}(s) \\ &= X(t_n) + \int_{t_n}^s f(X(r))dr + \int_{t_n}^s g(X(r))dW(r) \\ &\quad - \left(\bar{X}(t_n) + \int_{t_n}^s f(\bar{X}(t_n))dr + \int_{t_n}^s g(\bar{X}(t_n))dW(r) \right) \\ &= X(t_n) - \bar{X}(t_n) + \int_{t_n}^s f(X(r))dr - \int_{t_n}^s f(\bar{X}(t_n))dr \\ &\quad + \int_{t_n}^s g(X(r))dW(r) - \int_{t_n}^s g(\bar{X}(t_n))dW(r) \\ &= E(t_n) + \int_{t_n}^s (f(X(r)) - f(\bar{X}(t_n)))dr \\ &\quad + \int_{t_n}^s (g(X(r)) - g(\bar{X}(t_n)))dW(r). \end{aligned}$$

Simplifying gives us

$$\begin{aligned} E(s) &= E(t_n) + \int_{t_n}^s (f(X(r)) - f(\bar{X}(t_n)))dr \\ &\quad + \int_{t_n}^s (g(X(r)) - g(\bar{X}(t_n)))dW(r), \quad (3.12) \end{aligned}$$

for $s \in [t_n, t_{n+1}]$.

Using Definition 2.6.6 and choosing $V(X) = x^2$ this gives us $V_x(X) = 2x$

and $V_{xx}(X) = 2$. Using this with the error calculation (3.12) we have

$$\begin{aligned} dE(t)^2 &= [2E(r) (f(X(r)) - f(\bar{X}(t_n))) + (g(X(r)) - g(\bar{X}(t_n)))^2] dt \\ &\quad + 2E(t) (g(X(t)) - g(\bar{X}(t_n))) dW(t), \end{aligned}$$

which when written in integral form becomes

$$\begin{aligned} E(s)^2 &= E(t_n)^2 + \int_{t_n}^s [2E(r) (f(X(r)) - f(\bar{X}(t_n))) + (g(X(r)) - g(\bar{X}(t_n)))^2] dr \\ &\quad + \int_{t_n}^s 2E(r) (g(X(r)) - g(\bar{X}(t_n))) dW(r) \\ &= E(t_n)^2 + 2 \int_{t_n}^s E(r) (f(X(r)) - f(\bar{X}(t_n))) dr \\ &\quad + \int_{t_n}^s (g(X(r)) - g(\bar{X}(t_n)))^2 dr \\ &\quad + 2 \int_{t_n}^s E(r) (g(X(r)) - g(\bar{X}(t_n))) dW(r). \quad (3.13) \end{aligned}$$

Now separating the RHS of (3.13) and looking at the second term we have

$$2 \int_{t_n}^s E(r) (f(X(r)) - f(\bar{X}(t_n))) dr, \quad (3.14)$$

then using (3.7) we can rewrite (3.14) as

$$\begin{aligned} &2 \int_{t_n}^s E(r) (f(X(r)) - f(\bar{X}(t_n))) dr \\ &\leq 2 \int_{t_n}^s \frac{1}{2} [E(r)]^2 + \frac{1}{2} [f(X(r)) - f(\bar{X}(t_n))]^2 dr \\ &\leq \int_{t_n}^s E(r)^2 dr + \int_{t_n}^s [f(X(r)) - f(\bar{X}(t_n))]^2 dr. \quad (3.15) \end{aligned}$$

Focusing on the second term on the RHS of (3.15)

$$\int_{t_n}^s [f(X(r)) - f(\bar{X}(t_n))]^2 dr,$$

this can be rewritten as

$$\int_{t_n}^s [f(X(t_n)) - f(\bar{X}(t_n)) + f(X(r)) - f(X(t_n))]^2 dr, \quad (3.16)$$

using (3.8) on (3.16) gives us

$$\begin{aligned} & \int_{t_n}^s [f(X(t_n)) - f(\bar{X}(t_n)) + f(X(r)) - f(X(t_n))]^2 dr \\ & \leq 2 \int_{t_n}^s [f(X(t_n)) - f(\bar{X}(t_n))]^2 dr + 2 \int_{t_n}^s [f(X(r)) - f(X(t_n))]^2 dr. \end{aligned}$$

Next looking at the third term on the RHS of (3.13) we have

$$\int_{t_n}^s (g(X(r)) - g(\bar{X}(t_n)))^2 dr,$$

similarly to above this can be rewritten as

$$\int_{t_n}^s [g(X(t_n)) - g(\bar{X}(t_n)) + g(X(r)) - g(X(t_n))]^2 dr. \quad (3.17)$$

Then applying (3.8) to (3.17) gives us

$$\begin{aligned} & \int_{t_n}^s [g(X(t_n)) - g(\bar{X}(t_n)) + g(X(r)) - g(X(t_n))]^2 dr \\ & \leq 2 \int_{t_n}^s [g(X(t_n)) - g(\bar{X}(t_n))]^2 dr + 2 \int_{t_n}^s [g(X(r)) - g(X(t_n))]^2 dr. \end{aligned}$$

Now going back to (3.13), and rewriting to include these workings, it becomes

$$\begin{aligned} E(s)^2 & \leq E(t_n)^2 + \int_{t_n}^s E(r)^2 dr \\ & + 2 \int_{t_n}^s [f(X(t_n)) - f(\bar{X}(t_n))]^2 dr + 2 \int_{t_n}^s [f(X(r)) - f(X(t_n))]^2 dr \\ & + 2 \int_{t_n}^s [g(X(t_n)) - g(\bar{X}(t_n))]^2 dr + 2 \int_{t_n}^s [g(X(r)) - g(X(t_n))]^2 dr \\ & + 2 \int_{t_n}^s E(r) (g(X(r)) - g(\bar{X}(t_n))) dW(r). \end{aligned}$$

Taking the conditional expectation, with respect to $\mathcal{F}_n$, of both sides

$$\begin{aligned}
\mathbb{E} [E(s)^2 | \mathcal{F}_n] &\leq \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] + \mathbb{E} \left[\int_{t_n}^s E(r)^2 dr \middle| \mathcal{F}_n \right] \\
&+ \mathbb{E} \left[2 \int_{t_n}^s [f(X(t_n)) - f(\bar{X}(t_n))]^2 dr \middle| \mathcal{F}_n \right] + \mathbb{E} \left[2 \int_{t_n}^s [f(X(r)) - f(X(t_n))]^2 dr \middle| \mathcal{F}_n \right] \\
&+ \mathbb{E} \left[2 \int_{t_n}^s [g(X(t_n)) - g(\bar{X}(t_n))]^2 dr \middle| \mathcal{F}_n \right] + \mathbb{E} \left[2 \int_{t_n}^s [g(X(r)) - g(X(t_n))]^2 dr \middle| \mathcal{F}_n \right] \\
&+ \mathbb{E} \left[2 \int_{t_n}^s E(r) (g(X(r)) - g(\bar{X}(t_n))) dW(r) \middle| \mathcal{F}_n \right]. \quad (3.18)
\end{aligned}$$

Looking at each term on the RHS of (3.18) separately, we can begin to bring the expectation inside the integral

$$\begin{aligned}
&2 \int_{t_n}^s \mathbb{E} [(f(X(t_n)) - f(\bar{X}(t_n)))^2 | \mathcal{F}_n] dr \\
&+ 2 \int_{t_n}^s \mathbb{E} [(f(X(r)) - f(X(t_n)))^2 | \mathcal{F}_n] dr. \quad (3.19)
\end{aligned}$$

Now changing s to t_{n+1} and applying the global Lipschitz bound to (3.19) we have

$$\begin{aligned}
&2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(f(X(t_n)) - f(\bar{X}(t_n)))^2 | \mathcal{F}_n] dr \\
&+ 2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(f(X(r)) - f(X(t_n)))^2 | \mathcal{F}_n] dr \leq 2k \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(t_n) - \bar{X}(t_n))^2 | \mathcal{F}_n] dr \\
&+ 2k \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(r) - X(t_n))^2 | \mathcal{F}_n] dr. \quad (3.20)
\end{aligned}$$

As $E(t_n) = X(t_n) - \bar{X}(t_n)$ we can rewrite (3.20) and applying Lemma 2.7.1 which gives us

$$\begin{aligned}
&2k \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(t_n) - \bar{X}(t_n))^2 | \mathcal{F}_n] dr + 2k \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(r) - X(t_n))^2 | \mathcal{F}_n] dr \\
&\leq 2k \int_{t_n}^{t_{n+1}} \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] dr + 2c \int_{t_n}^{t_{n+1}} |r - t_n| dr.
\end{aligned}$$

Finally resolving the integrals we see that

$$\begin{aligned} 2k \int_{t_n}^{t_{n+1}} \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] dr + 2c \int_{t_n}^{t_{n+1}} |r - t_n| dr \\ \leq 2k \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] h_{n+1} + 2k_1 h_{n+1}^2. \end{aligned}$$

Likewise using the same process for the next term on the RHS of (3.18) we will find

$$\begin{aligned} 2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(g(X(t_n)) - g(\bar{X}(t_n)))^2 | \mathcal{F}_n] dr \\ + 2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(g(X(r)) - g(X(t_n)))^2 | \mathcal{F}_n] dr \\ \leq 2k \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] h_{n+1} + 2k_1 h_{n+1}^2. \end{aligned}$$

Rewriting (3.18) with these workings, and as the expectation of the Itô integrals is zero, we now have

$$\begin{aligned} \mathbb{E} [E(t_{n+1})^2 | \mathcal{F}_{n+1}] \leq \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] + \\ \int_{t_n}^{t_{n+1}} \mathbb{E} [E(r)^2 | \mathcal{F}_n] dr + 4k \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] h_{n+1} + 4k_1 h_{n+1}^2. \end{aligned} \quad (3.21)$$

The last term of (3.21) to be resolved is

$$\int_{t_n}^{t_{n+1}} \mathbb{E} [E(r)^2 | \mathcal{F}_n] dr, \quad (3.22)$$

remembering that $E(r) = X(r) - \bar{X}(r)$ we can do a similar rearranging as before such that

$$\begin{aligned} \int_{t_n}^{t_{n+1}} \mathbb{E} [E(r)^2 | \mathcal{F}_n] dr &= \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(r) - \bar{X}(r))^2 | \mathcal{F}_n] dr \\ &= \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(r) - \bar{X}(t_n) + \bar{X}(t_n) - \bar{X}(r))^2 | \mathcal{F}_n] dr \\ &= 2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(r) - \bar{X}(t_n))^2 | \mathcal{F}_n] dr + 2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(\bar{X}(t_n) - \bar{X}(r))^2 | \mathcal{F}_n] dr. \end{aligned} \quad (3.23)$$

Now looking at the the second term on the RHS of (3.23) and applying Lemma 2.7.1 we have

$$\begin{aligned}
2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(\bar{X}(t_n) - \bar{X}(r))^2 | \mathcal{F}_n] dr &= 2 \int_{t_n}^{t_{n+1}} \mathbb{E} [|\bar{X}(t_n) - \bar{X}(r)|^2 | \mathcal{F}_n] dr \\
&\leq 2c_1 \int_{t_n}^{t_{n+1}} |r - t_n| dr \\
&\leq 2c_1 h_{n+1}^2.
\end{aligned}$$

Going back to the first term on the RHS of (3.23) we can rewrite as before, and apply Lemma 2.7.1 again, which gives us

$$\begin{aligned}
2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(r) - \bar{X}(t_n))^2 | \mathcal{F}_n] dr &= \\
&2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(t_n) - \bar{X}(t_n) + X(r) - X(t_n))^2 | \mathcal{F}_n] dr \\
&\leq 2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(t_n) - \bar{X}(t_n))^2 | \mathcal{F}_n] dr + 2 \int_{t_n}^{t_{n+1}} \mathbb{E} [(X(r) - X(t_n))^2 | \mathcal{F}_n] dr \\
&\leq 4c_2 \int_{t_n}^{t_{n+1}} \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] dr + 4c_2 \int_{t_n}^{t_{n+1}} |r - t_n| dr \\
&\leq 4c_2 \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] h_{n+1} + 4c_2 h_{n+1}^2.
\end{aligned}$$

Now putting these workings together, (3.22) can be written as

$$\int_{t_n}^{t_{n+1}} \mathbb{E} [E(r)^2 | \mathcal{F}_n] dr \leq 4c_2 \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] h_{n+1} + (4c_2 + 2c_1) h_{n+1}^2.$$

Next we can substitute our workings into (3.21), we can now compute the following bound on the local error of the scheme in mean square

$$\mathbb{E} [E(t_{n+1})^2 | \mathcal{F}_{n+1}] \leq \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] + k_2 \mathbb{E} [E(t_n)^2 | \mathcal{F}_n] h_{n+1} + k_3 h_{n+1}^2. \quad (3.24)$$

Using the tower property of conditional expectations from Section 2.3, we condition the equation with respect to $\sigma(\{\emptyset, \Omega\})$ so we have

$$\begin{aligned}
\mathbb{E} \left[\mathbb{E} [E(t_{n+1})^2 | \mathcal{F}_{n+1}] \middle| \sigma(\{\emptyset, \Omega\}) \right] &\leq \mathbb{E} \left[\mathbb{E} [E(t_n)^2 | \mathcal{F}_n] \middle| \sigma(\{\emptyset, \Omega\}) \right] \\
&\quad + k_2 \mathbb{E} \left[\mathbb{E} [E(t_n)^2 | \mathcal{F}_n] \middle| \sigma(\{\emptyset, \Omega\}) \right] h_{n+1} \\
&\quad + k_3 h_{n+1}^2,
\end{aligned}$$

rearranging we see that

$$\mathbb{E}[E(t_{n+1})^2] - \mathbb{E}[E(t_n)^2] \leq k_2 \mathbb{E}[E(t_n)^2] h_{n+1} + k_3 h_{n+1}^2.$$

Next we can compute the cumulative effect of the local errors over the interval $[0, T]$, summing both sides over $n = 0, \dots, N-1$, and remembering $\sum_{i=0}^n h_i = T$, gives us

$$\sum_{n=0}^{N-1} (\mathbb{E}[E(t_{n+1})^2] - \mathbb{E}[E(t_n)^2]) \leq \sum_{n=0}^{N-1} k_2 \mathbb{E}[E(t_n)^2] h_{\max} + k_2 T h_{\max}. \quad (3.25)$$

Notice that the LHS of (3.25) is a telescoping sum, so when we write it out in full we see that all of the terms, bar the first and the last terms, will cancel

$$\sum_{n=0}^{N-1} (\mathbb{E}[E(t_{n+1})^2] - \mathbb{E}[E(t_n)^2]) = \mathbb{E}[E(t_N)^2] - \mathbb{E}[E(t_0)^2].$$

As $E(t_0)^2 = 0$ we have

$$\mathbb{E}[E(t_N)^2] \leq \sum_{n=0}^{N-1} \mathbb{E}[E(t_n)^2] k_4 h_{\max} + k_3 h_{\max},$$

where $k_3 = k_2 T$ and $k_4 = k_3 T$.

Now applying the Gronwall inequality (Lemma 3.4.1) with $u_n = \mathbb{E}[E(t_n)^2]$, $v_n = k_4 h_{\max}$, and $\alpha = k_3 h_{\max}$ we have

$$\mathbb{E}[E(t_N)^2] \leq k_3 h_{\max} e^{\sum_{n=0}^{N-1} h k_4} = k_3 h_{\max} e^{T k_4} = k_5 h_{\max},$$

where $k_5 = k_3 e^{T k_4}$.

Taking the square root of both sides and recalling that $t_n = T$ gives the result. $\square$

Chapter 4

Adaptive numerical solution of SDEs with Markovian switching

Now that we have established the fundamentals of strongly convergent methods for Itô-type SDEs, we will extend to the case of SDEs with Markovian switching (SDEwMS). In this chapter we begin by defining an SDEwMS and the general form of a solution to such an equation. We then expand on the definitions and conditions outlined in Chapter 2 to include the Markov chain.

An SDEwMS is an SDE where the coefficient functions are subject to random switches according to the evolution of a Markov chain. All states are predefined in a state space with each state corresponding to a switch of the coefficient functions.

We will show that a numerical method satisfying a one step bound of the form (3.24) for an SDE without switching will, when applied to an SDEwMS, demonstrate strong convergence.

4.1 SDEs with Markovian switching

First, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, P)$ be a complete filtered probability space satisfying the usual conditions. We are now back in $\mathbb{R}^d$ and the process $W = (W_1(t), \dots, W_m(t))^T$ is an m -dimensional standard Brownian motion defined on this probability space. Let $r(t)$, $t \geq 0$ be a right-continuous Markov chain on the same probability space, taking values in $\mathbb{S}$ with generator $\Gamma = (\gamma_{ij})_{L \times L}$ given by

$$\mathbb{P}[r(t + \Delta) = j | r(t) = i] = \begin{cases} \gamma_{ij}\Delta + o(\Delta), & \text{if } i \neq j; \\ 1 + \gamma_{ii}\Delta + o(\Delta), & \text{if } i = j; \end{cases} \quad (4.1)$$

where $\Delta > 0$ and $o(\Delta)$ are higher order terms of Δ . Here $\gamma_{ij} \geq 0$ is the transition rate from i to j if $i \neq j$ while

$$\gamma_{ii} = - \sum_{j \neq i} \gamma_{ij}. \quad (4.2)$$

We assume that the standard Brownian Motion W and the Markov chain r are both $\mathcal{F}_t$ -adapted, and for the latter the initial value $r(0) = r_0$ is an $\mathcal{F}_0$ -measurable random variable taking values in $\mathbb{S}$.

The generator matrix Γ for a state space $\mathbb{S} = \{1, \dots, L\}$ is

$$\Gamma = \begin{bmatrix} \gamma_{11} & \gamma_{12} & \dots & \gamma_{1L} \\ \gamma_{21} & \gamma_{22} & \dots & \gamma_{2L} \\ & & \ddots & \\ \gamma_{L1} & \gamma_{L2} & \dots & \gamma_{LL} \end{bmatrix} \quad (4.3)$$

Consider an SDE with Markovian switching (SDEwMS) of the form

$$dX(t) = f(X(t), r(t))dt + g(X(t), r(t))dW(t), \quad 0 \leq t \leq T \quad (4.4)$$

with initial data $X(0) = x_0 \in L^2_{\mathcal{F}_{t_0}}(\Omega; \mathbb{R}^d)$ and $r(0) = r_0$, where r_0 is an $\mathbb{S}$ -valued $\mathcal{F}_{t_0}$ -measurable random variable and

$$f : \mathbb{R}^d \times \mathbb{S} \rightarrow \mathbb{R}^d \quad \text{and} \quad g : \mathbb{R}^d \times \mathbb{S} \rightarrow \mathbb{R}^{d \times m}$$

Definition 4.1.1. An $\mathbb{R}^d$ -valued stochastic process $(X(t))_{t \in [0, T]}$ is called a solution of equation (4.4) if it has the following properties:

1. $(X(t))_{t \in [0, T]}$ is continuous and $\mathcal{F}_t$ -adapted
2. $(f(X(t), t, r(t)))_{t \in [0, T]} \in \mathcal{L}^1([0, T]; \mathbb{R}^d)$ while
 $(g(x(t), t, r(t)))_{t \in [0, T]} \in \mathcal{L}^2([0, T]; \mathbb{R}^{d \times m})$
3. for any $t \in [0, T]$, equation

$$X(t) = X(0) + \int_0^t f(X(s), s, r(s)) ds + \int_0^t g(X(s), s, r(s)) dW(s) \quad (4.5)$$

holds with probability 1.

4.1.1 Existence and uniqueness of solutions of Itô type SDEwMS

Here we will extend the minimal conditions defined for f and g in Chapter 2, that ensure the existence of unique solutions over each interval $[0, T]$, to include Markovian switching. We require that both satisfy a local Lipschitz condition:

Assumption 4.1.1. For every integer $k \geq 1$, there exists a positive constant h_k such that, for all $i \in \mathbb{S}$ and those $x, y \in \mathbb{R}^d$ with $|x| \vee |y| \leq k$

$$\|f(x, i) - f(y, i)\|^2 \vee \|g(x, i) - g(y, i)\|^2 \leq \bar{h}_k \|x - y\|^2, \quad (4.6)$$

Consider also the following linear growth condition:

Assumption 4.1.2. f and g satisfy a linear growth bound if for all $(x, i) \in \mathbb{R}^d \times \mathbb{S}$ there exists a constant $0 < K_3 < \infty$ such that

$$\|f(x, i)\|^2 \vee \|g(x, i)\|^2 \leq K_3(1 + \|x\|^2). \quad (4.7)$$

Under Assumptions 4.1.1 and 4.1.2 there exists a unique solution to (4.4) on each $[0, T]$: See Mao & Yuan [14, Theorem 3.16].

It is possible to replace the linear growth condition with a weaker constraint. If Assumption 4.1.1 holds, so that f and g satisfy a local Lipschitz condition, and they also satisfy the following monotonicity condition

Assumption 4.1.3. There exists a positive constant K such that for all $(x, i) \in \mathbb{R}^d \times [0, T] \times \mathbb{S}$,

$$x^T f(x, i) + \frac{1}{2} \|g(x, i)\|^2 \leq K(1 + \|x\|^2). \quad (4.8)$$

Then there exists a unique solution X to Eq. (4.4) on each $[0, T]$. Under the same conditions (Assumption 4.1.1 and either Assumption 4.1.2 or 4.1.3) solutions of (4.4) also satisfy uniform moment bounds and Hölder regularity bounds as outlined in Chapter 2.

4.1.2 Characterisation of waiting times

From (4.2) we see that $\sum_{j=i}^L \gamma_{ij} = 1$ for all $i = 1, \dots, L$. Denote the k^{th} time at which the Markov chain switches between distinct states as τ_k . Then $(\tau_i)_{i \geq 0}$ is an increasing sequence of non-negative random variables which can be written

$$(\tau_i)_{i \geq 0} = \sum_{k=0}^i \pi_k, \quad (4.9)$$

where π_k denotes the waiting time between the k^{th} and $(k+1)^{th}$ such transitions. In practice, as long as we keep track of the current state $i \in \mathbb{S}$ of the Markov chain, these waiting times may be computed as the holding times associated with each i , and these are modelled as exponentially distributed random variables with state-dependent rates λ_i : see Section 5.1.

Associated with the Markov chain r is a process $(\bar{N}_t)_{t \geq 0}$ that counts the number of switches that take place over the interval $[0, T]$ for all $t \geq 0$, written

$$\bar{N}_t := \#\{i \geq 1, \tau_i \in [0, T]\}. \quad (4.10)$$

Since the state space $\mathbb{S}$ is a finite set, $\bar{N}_t$ is almost surely (a.s) finite on any interval $[0, T]$: see for example [15, Theorem 2.7.1]. However $\bar{N}_t(\omega)$ is not uniformly bounded from above across the sample space Ω for any $t \in [0, T]$, and this presents a significant challenge in our main convergence result. A key to overcoming this challenge is the following.

Assumption 4.1.4. We suppose that $(\mathcal{F}_t)_{t \geq 0}$ can be decomposed such that $\mathcal{F}_t = \sigma(\mathcal{G}_t \cup \mathcal{H}_T)$ for all $t \geq 0$, where W is adapted to the filtration $(\mathcal{G}_t)_{t \geq 0}$, r is adapted to the filtration $(\mathcal{H}_t)_{t \geq 0}$, and moreover the σ -algebra $\mathcal{G}_t$ is independent of $\mathcal{H}_T$ for all $t \in [0, T]$.

Assumption 4.1.4 holds if r is independent of W , and in practice it allows us to simulate trajectories of (4.4) by precomputing trajectories of r and incorporating switching times into an adaptive mesh.

4.2 Constructing the numerical method

4.2.1 Structure of an adaptive mesh

Let $\{h_{n+1}\}_{n \in \mathbb{N}}$ be a sequence of strictly positive random timesteps with corresponding random times $\{t_n := \sum_{i=1}^n h_i\}_{n \in \mathbb{N} \setminus \{0\}}$, where $t_0 = 0$. We set up a framework by which nodes in this mesh can be generated procedurally as a function of the solution at the start of each step, while also ensuring that the timepoints at which the Markov chain switches between distinct states are included.

Assumption 4.2.1. For each term of the sequence of random timesteps $\{h_{n+1}\}_{n \in \mathbb{N}}$ at least one of the following three possibilities holds:

1. there are constant values $h_{\max} > h_{\min} > 0$, $\rho > 1$ such that $h_{\max} = \rho h_{\min}$, and

$$0 < h_{\min} \leq h_{n+1} \leq h_{\max} \leq 1;$$

2. there exists a switching time τ_k , as defined in Section 2, such that $t_{n+1} = \tau_k$, so that

$$0 < h_{n+1} = \tau_k - t_n \leq h_{\max} \leq 1;$$

3. $t_{n+1} = T$, so that

$$0 < h_{n+1} = T - t_n \leq h_{\max} \leq 1.$$

In addition we assume each h_{n+1} is $\mathcal{F}_{t_n}$ -measurable.

It was shown in [9] that as long as h_1 can be computed from the initial data, and successive adaptive timesteps h_{k+1} are computed on each trajectory as a deterministic function of the numerical approximation at time t_k ,

then each t_n on a mesh satisfying the conditions of Assumption 4.2.1 is an $\mathcal{F}_t$ -stopping time.

Definition 4.2.1. Let $N^{(t)}$ be a random integer such that

$$N^{(t)} := \max\{n \in \mathbb{N} \setminus \{0\} : t_{n-1} < t\}, \quad (4.11)$$

and let $N = N^{(t)}$ and $t_N = T$, so that T is always the last point on the mesh

Let Definition 4.2.1 be satisfied with $N_{\min}^{(t)} := \lfloor t/h_{\max} \rfloor$ and $N_{\max}^{(t)} := \lceil t/h_{\min} + \bar{N}_t \rceil$. We ensure that we reach the final time by taking $h_N = T - t_{N-1}$ as our final step, and use the backstop method if $h_N < h_{\min}$. Recall that $\bar{N}_t < \infty$ a.s for all $t \in [0, T]$, where $\mathbb{S}$ is a finite set, therefore $N_{\max}^{(t)}$ is an a.s finite random variable.

4.2.2 An explicit discretisation scheme for an SDEwMS

While the main result in this dissertation applies for any strongly convergent numerical method, in this section we give the example of the explicit Euler-Maruyama and Milstein methods, and show how may be adapted to numerically solve an SDEwMS.

The explicit Euler-Maruyama scheme

We may extend (3.2), presented in Chapter 3, to solve an SDEwMS over an adaptive mesh $\{t_n\}$ as

$$X_{n+1} = X_n + h_{n+1}f(X_n, r(t_n)) + g(X_n, r(t_n))\Delta W_{n+1}. \quad (4.12)$$

Then the continuous approximation for all $s \in [t_n, t_{n+1}]$ is

$$\bar{X}(s) = X_n + \int_{t_n}^s f(\bar{X}(t_n), r(t_n))dr + \int_{t_n}^s g(\bar{X}(t_n), r(t_n))dW(r). \quad (4.13)$$

Notice that since switching times are included in the adaptive mesh as constructed in Section 4.2.1 we can simply fix the value of the Markov chain r to the value observed at the beginning of each step.

The Milstein scheme

Again, we extend (3.4), in the case where $d = m = 1$, to solve an SDEwMS over an adaptive mesh $\{t_n\}$ as

$$\begin{aligned} X_{n+1} = X_n &+ h_{n+1}f(X_n, r(t_n)) + g(X_n, r(t_n))\Delta W_{n+1} \\ &+ \frac{1}{2}g'(X_n, r(t_n))g(X_n, r(t_n))(\Delta W_{n+1}^2 - h_{n+1}), \end{aligned} \quad (4.14)$$

with corresponding continuous version

$$\begin{aligned} \bar{X}(s) = X_n &+ \int_{t_n}^s f(\bar{X}(t_n), r(t_n))dr + \int_{t_n}^s g(\bar{X}(t_n), r(t_n))dW(r) \\ &+ \int_{t_n}^s \int_{t_n}^r g'(\bar{X}(t_n), r(t_n))g(\bar{X}(t_n), r(t_n))dW(p)dW(r). \end{aligned} \quad (4.15)$$

A full description of the Milstein scheme applied to (4.4) in the more general d -dimensional case with m independent Brownian noise terms requires the simulation of Lévy areas. Details are in [8]. When $m = d = 1$, the Milstein method (3.5) may be expressed as a map $\psi : \mathbb{R} \times \mathbb{S} \times [0, T] \times [0, h_{\max}] \rightarrow \mathbb{R}$ over each step $[t_n, t_{n+1}]$:

$$\begin{aligned} \psi(x, i, t_n, s - t_n) &:= x + (s - t_n)f(x, i) \\ &+ \int_{t_n}^s g(x, i)dW(v) + \int_{t_n}^s \int_{t_n}^v g'(x, i)g(x, i)dW(p)dW(v). \end{aligned} \quad (4.16)$$

Finally, note that if we modify the second term on the RHS of (4.14) to be $h_{n+1}f(X_{n+1})$ we arrive at the implicit Milstein scheme, which will be used later in Chapter 5. A strong convergence analysis of this scheme for SDEs with non-globally Lipschitz continuous coefficients is contained in [4].

4.2.3 A hybrid adaptive numerical scheme: general form

Constructing our scheme we will use two maps $\mathcal{M} : \mathbb{R}^d \times \mathbb{S} \times [0, T] \times [0, h_{\max}] \rightarrow \mathbb{R}^d$ and $\mathcal{B} : \mathbb{R}^d \times \mathbb{S} \times [0, T] \times [0, h_{\max}] \rightarrow \mathbb{R}^d$ where $\mathcal{M}$ is the main or default map and $\mathcal{B}$ the backstop map. $\mathcal{M}$ corresponds to an efficient scheme which is convergent as an adaptive method and $\mathcal{B}$ corresponds

to a method which is convergent as a fixed step method but which may be inefficient or may induce distortions in solution dynamics, with the aim that $\mathcal{B}$ will be rarely used.

We characterise $\mathcal{M}$ and $\mathcal{B}$ through their action over a single step, requiring them to satisfy a certain mean-square consistency bound.

Definition 4.2.2. Let $\{h_{n+1}\}_{n \in \mathbb{N}}$ satisfy Assumption 4.2.1. We define the continuous form of hybrid adaptive numerical scheme associated with the timestepping strategy $\{h_{n+1}\}_{n \in \mathbb{N}}$ to be

$$\begin{aligned} Y(s) = & \mathcal{M}(Y(t_n), r(t_n), t_n, s - t_n) \cdot \mathcal{I}_{\{h_{\min} < h_{n+1} \leq h_{\max}\}} \\ & + \mathcal{B}(Y(t_n), r(t_n), t_n, s - t_n) \cdot \mathcal{I}_{\{h_{n+1} \leq h_{\min}\}}. \end{aligned} \quad (4.17)$$

Now define the global error of the scheme

Definition 4.2.3. For any $s \in [0, T]$, we define the global error acquired by the scheme (4.17) to be

$$E(s) := X(s) - Y(s) \quad (4.18)$$

where $X(s)$ satisfies (4.4) and $Y(s)$ satisfies (4.17).

Then we make the following assumption

Assumption 4.2.2. For $s \in [t_n, t_{n+1}]$, $n \in \mathbb{N}$ and any $x, y \in \mathbb{R}^d$, let

$$\begin{aligned} E_{\mathcal{M}}^{x,y,i}(s) &:= x - \mathcal{M}(y, i, t_n, s - t_n); \\ E_{\mathcal{B}}^{x,y,i}(s) &:= x - \mathcal{B}(y, i, t_n, s - t_n). \end{aligned} \quad (4.19)$$

Assume that $E_{\mathcal{M}}^{x,y,i}$ and $E_{\mathcal{B}}^{x,y,i}$ satisfy, for some $\delta > 0$, both $\psi = \mathcal{M}$ or $\psi = \mathcal{B}$, any $i \in \mathbb{S}$, and any a.s. finite $\mathcal{F}_{t_n}$ -measurable $\mathbb{R}^d$ -valued random variables A_n, B_n

$$\begin{aligned} \mathbb{E}[|E_{\psi}^{A_n, B_n, i}(t_{n+1})|^2 | \mathcal{F}_{t_n}] &\leq \|A_n - B_n\|^2 \\ &+ \Gamma_1 \int_{t_n}^{t_{n+1}} \mathbb{E}[|E_{\psi}^{A_n, B_n, i}(r)|^2 | \mathcal{F}_{t_n}] dr + \Gamma_{2,n} h_{n+1}^{2\delta+1}. \end{aligned} \quad (4.20)$$

In (4.20), for any $h_{n+1} \leq h_{\min}$ in the case where $\psi = \mathcal{B}$ and for any $h_{\min} < h_{n+1} \leq h_{\max}$ in the case where $\psi = \mathcal{M}$, $\Gamma_1 < \infty$ is a constant, $\Gamma_{2,n}$ is a scalar $\mathcal{F}_{t_n}$ -measurable random variable with finite expectation denoted Γ_2 , and both Γ_1 and $\Gamma_{2,n}$ are independent of $h_{\max}$.

4.2.4 A hybrid adaptive numerical scheme: particular form

To ensure the appropriate bound to satisfy Assumption 4.2.2 we use a particular kind of adaptive time-stepping strategy such that, whenever $h_{\min} < h_{n+1} \leq h_{\max}$,

$$\|Y(t_n)\| < R. \quad (4.21)$$

Lemma 4.2.1. Let $Y(s)$ be the hybrid scheme (4.17). Fix $n = 0, \dots, N-1$ and $k > 0$. Then, in the event that $h_{\min} < h_{n+1} \leq h_{\max}$, h_{n+1} satisfies (4.21)

$$h_{n+1} = \left(h_{\min} \vee \left(\frac{h_{\max}}{\|Y(t_n)\|^{1/k}} \wedge h_{\max} \right) \right) \wedge (\tau_{\bar{N}_{t_n}+1} - t_n) \quad (4.22)$$

For a proof, see Lemma 4.2 in [9]. Since the mesh $\{t_n\}$ has been constructed so that the value of the Markov chain r cannot change between meshpoints, conditions under which the Milstein map (4.16) satisfies (4.20) over a single step with $\delta = 1$ are independent of i and the same as those provided in [9, Section 4].

4.3 Main theorem on strong convergence

For any adaptive numerical scheme composed of main and backstop maps that satisfy the bound (4.20) in Assumption 4.2.2, our main result demonstrates strong L^2 convergence when extended to an SDEwMS via a switching time adapted mesh. The proof is a modification of that of [9], though the change of setting from SJDEs to SDEwMS allows for the analysis to be simplified. We also point to the Milstein method 4.15 as an example of a method leading to a specific scheme of order $\delta = 1$.

Theorem 4.3.1. *Let $(X(t))_{t \in [0, T]}$ be a solution of (4.4) with initial value $X(0) = x_0 \in \mathbb{R}^d$ and suppose that Assumptions 4.1.1, 4.1.2 and 4.1.3 hold. Let Y be an adaptive numerical scheme as characterised in Definition 4.2.2 with $Y_0 = x_0$ such that Assumption 4.2.2 holds with some $\delta > 0$. Then there exists a constant $C(T) > 0$ such that*

$$\max_{t \in [0, T]} \left(\mathbb{E} \left[\|X(t) - Y(t)\|^2 \right] \right)^{\frac{1}{2}} \leq C(T) h_{\max}^{\delta}, \quad (4.23)$$

where C does not depend on $h_{\max}$.

Proof. The proof naturally divides into four steps which successively show how conditional mean-square errors accumulate, first over the steps between successive switching times, then from the initial time to the final switching time before a fixed time t , then from the last switching time to t , and finally over the full interval $[0, t]$. We may then apply the Gronwall inequality (see Lemma 3.4.1) to complete the proof.

Step 1: cumulative error bound over the interval between successive switches. Over a single step, we obtain from map error in Assumption the following error bound regardless of whether the scheme uses the main map or the backstop map over a single step of length $0 < h_{n+1} \leq h_{\max}$, and regardless of whether or not a change in the state of the Markov chain r occurs at either endpoint of the timestep: a.s,

$$\mathbb{E}[\|E(t_{n+1})\|^2 | \mathcal{F}_{t_n}] \leq \|E(t_n)\|^2 + \Gamma_1 \int_{t_n}^{t_{n+1}} \mathbb{E}[\|E(u)\|^2 | \mathcal{F}_{t_n}] du + \Gamma_{2,n} h_{n+1}^{2\delta+1}. \quad (4.24)$$

Let τ_k and τ_{k+1} be successive switching times. Subtract the first term on the RHS of ref one step error from both sides of the inequality and sum

both sides over all the steps from τ_k to τ_{k+1} . Then we get

$$\begin{aligned}
\mathcal{K}_k &:= \sum_{n=N^{(\tau_k)} }^{N^{(\tau_{k+1})}-1} \left(\mathbb{E}[\|E(t_{n+1})\|^2 | \mathcal{F}_{t_n}] - \|E(t_n)\|^2 \right) \mathcal{I}_{\{N^{(\tau_{k+1})} > n\}} \\
&\leq \sum_{n=N^{(\tau_k)} }^{N^{(\tau_{k+1})}-1} \left(\Gamma_1 \int_{t_n}^{t_{n+1}} \mathbb{E}[\|E(u)\|^2 | \mathcal{F}_{t_n}] du + \Gamma_{2,n} h_{n+1}^{2\delta+1} \right) \mathcal{I}_{\{N^{(\tau_{k+1})} > n\}} \\
&= \Gamma_1 \sum_{n=N^{(\tau_k)} }^{N^{(\tau_{k+1})}-1} \int_{t_n}^{t_{n+1}} \mathbb{E}[\|E(u)\|^2 | \mathcal{F}_{t_n}] \mathcal{I}_{\{N^{(\tau_{k+1})} > n\}} du \\
&\quad + \sum_{n=N^{(\tau_k)} }^{N^{(\tau_{k+1})}-1} \Gamma_{2,n} h_{n+1}^{2\delta+1} \mathcal{I}_{\{N^{(\tau_{k+1})} > n\}}, \quad a.s.
\end{aligned}$$

By Definition, the filtration $\mathcal{F}_{t_n}$ may be written as $\mathcal{F}_{t_{N^{(u)}-1}}$, for all $u \in [t_n, t_{n+1})$ and $n = 0, \dots, N^{(u)} - 1$, so we can rewrite the inequality as

$$\mathcal{K}_k \leq \Gamma_1 \int_{\tau_k}^{\tau_{k+1}} \mathbb{E}[\|E(u)\|^2 | \mathcal{F}_{t_{N^{(u)}-1}}] du + \sum_{n=N^{(\tau_k)} }^{N^{(\tau_{k+1})}-1} \Gamma_{2,n} h_{n+1}^{2\delta+1}, \quad a.s. \quad (4.25)$$

where all indicator variables on the RHS have been bounded above by 1.

Step 2: cumulative error from time $t = 0$ to the time of the last switch before time t . Fix any $t \in [0, T]$. Summing over k for all switching times on the interval $(0, t]$ and denoting $\mathcal{Q}(t) := \sum_{k=0}^{\bar{N}_t-1} \mathcal{K}_k$, we have a.s.,

$$\mathcal{Q}(t) \leq \Gamma_1 \sum_{k=0}^{\bar{N}_t-1} \int_{\tau_k}^{\tau_{k+1}} \mathbb{E}[\|E(u)\|^2 | \mathcal{F}_{t_{N^{(u)}-1}}] du + \sum_{k=0}^{\bar{N}_t-1} \sum_{n=N^{(\tau_k)} }^{N^{(\tau_{k+1})}-1} \Gamma_{2,n} h_{n+1}^{2\delta+1}. \quad (4.26)$$

Set $\tau_0 = 0$ to include time from 0 to the first switch τ_1 and take the expectation with respect to $\mathcal{H}_T$

$$\begin{aligned}
\mathbb{E}[\mathcal{Q}(t) | \mathcal{H}_T] &\leq \Gamma_1 \mathbb{E} \left[\sum_{k=0}^{\bar{N}_t-1} \int_{\tau_k}^{\tau_{k+1}} \mathbb{E}[\|E(u)\|^2 | \mathcal{F}_{t_{N^{(u)}-1}}] du \middle| \mathcal{H}_T \right] \\
&\quad + \mathbb{E} \left[\sum_{k=0}^{\bar{N}_t-1} \sum_{n=N^{(\tau_k)} }^{N^{(\tau_{k+1})}-1} \Gamma_{2,n} h_{n+1} \middle| \mathcal{H}_T \right] h_{\max}^{2\delta}, \quad a.s.
\end{aligned}$$

Since $\bar{N}_t$ is $\mathcal{H}_T$ measurable we can take all of the summations on the RHS out of the conditional expectations.

$$\begin{aligned} \mathbb{E}[\mathcal{Q}(t)|\mathcal{H}_T] &\leq \Gamma_1 \sum_{k=0}^{\bar{N}_t-1} \mathbb{E} \left[\int_{\tau_k}^{\tau_{k+1}} \mathbb{E} [\|E(u)\|^2 | \mathcal{F}_{t_{N(u)-1}}] du | \mathcal{H}_T \right] \\ &\quad + \sum_{k=0}^{\bar{N}_t-1} \mathbb{E} \left[\sum_{n=N(\tau_k)}^{N(\tau_{k+1})-1} \Gamma_{2,n} h_{n+1} \middle| \mathcal{H}_T \right] h_{\max}^{2\delta}, \quad a.s. \end{aligned}$$

Since $\mathcal{H}_T \subseteq \mathcal{F}_{t_{N(u)-1}}$ for any $u \in [0, T]$, an application of the tower property of conditional expectations yields a.s,

$$\begin{aligned} &\mathbb{E}[\mathcal{Q}(t)|\mathcal{H}_T] \\ &\leq \Gamma_1 \sum_{k=0}^{\bar{N}_t-1} \int_{\tau_k}^{\tau_{k+1}} \mathbb{E} [\|E(u)\|^2 | \mathcal{H}_T] du + \sum_{k=0}^{\bar{N}_t-1} \mathbb{E} \left[\sum_{n=N(\tau_k)}^{N(\tau_{k+1})-1} \Gamma_{2,n} h_{n+1} \middle| \mathcal{H}_T \right] h_{\max}^{2\delta} \\ &\leq \Gamma_1 \int_0^{\tau_{\bar{N}_t}} \mathbb{E} [\|E(u)\|^2 | \mathcal{H}_T] du + \sum_{k=0}^{\bar{N}_t-1} \mathbb{E} \left[\sum_{n=N(\tau_k)}^{N(\tau_{k+1})-1} \Gamma_{2,n} h_{n+1} \middle| \mathcal{H}_T \right] h_{\max}^{2\delta}, \quad a.s. \end{aligned} \tag{4.27}$$

Step 3: error bound from the time of the last switch to time t .

Fix $t \in [0, T]$. If $t_{N(t)} = t$ this step can be omitted. Therefore suppose that $t_{N(t)} < t$. Replacing t_n with $t_{N(t)-1}$ and t_{n+1} with t in ref one step error, we have a bound on the error over the last step to time t as

$$\begin{aligned} \mathbb{E} [\|E(t)\|^2 | \mathcal{F}_{t_{N(t)-1}}] &\leq \|E(t_{N(t)-1})\|^2 \\ &\quad + \Gamma_1 \int_{t_{N(t)-1}}^t \mathbb{E} [\|E(u)\|^2 | \mathcal{F}_{t_{N(t)-1}}] du + \Gamma_{2,n} |t - t_{N(t)-1}|^{2\delta+1}. \end{aligned} \tag{4.28}$$

Now summing up to t with the error bound over the last step satisfying ref and taking the expectation conditioned on $\mathcal{H}_T$ we have, for all $t \in [0, T]$ and

a.s,

$$\begin{aligned}
& \mathcal{R}(t) \\
&:= \mathbb{E} \left[\sum_{n=N^{(\tau_{\bar{N}_t})}}^{N^{(t)}-2} (\mathbb{E} [\|E(t_{n+1})\|^2 | \mathcal{F}_{t_n}] - \|E(t_n)\|^2) \mathcal{I}_{\{N^{(t)} > n+1\}} \right. \\
&\quad \left. + \mathbb{E} [\|E(t)\|^2 | \mathcal{F}_{t_{N^{(t)}-1}}] - \|E(t_{N^{(t)}-1})\|^2 \right] \Big| \mathcal{H}_T \\
&\leq \Gamma_1 \mathbb{E} \left[\sum_{n=N^{(\tau_{\bar{N}_t})}}^{N^{(t)}-2} \int_{t_n}^{t_{n+1}} \mathbb{E} [\|E(u)\|^2 | \mathcal{F}_{t_n}] \mathcal{I}_{\{N^{(t)} > n+1\}} du \right. \\
&\quad \left. + \int_{t_{N^{(t)}-1}}^t \mathbb{E} [\|E(u)\|^2 | \mathcal{F}_{t_{N^{(t)}-1}}] du \right] \Big| \mathcal{H}_T \\
&\quad + h_{\max}^{2\delta} \mathbb{E} \left[\sum_{n=N^{(\tau_{\bar{N}_t})}}^{N^{(t)}-2} \Gamma_{2,n} h_{n+1} + \Gamma_{2,N^{(t)}-1} |t - t_{N^{(t)}-1}| \right] \Big| \mathcal{H}_T. \quad (4.29)
\end{aligned}$$

Summing the integrals on the RHS of ref, bounding the indicators on the RHS above by 1, and again applying the tower property of conditional expectations we get

$$\begin{aligned}
\mathcal{R}(t) &\leq \Gamma_1 \int_{\tau_{\bar{N}_t}}^t \mathbb{E} [\|E(u)\|^2 | \mathcal{H}_T] du \\
&\quad + h_{\max}^{2\delta} \mathbb{E} \left[\sum_{n=N^{(\tau_{\bar{N}_t})}}^{N^{(t)}-2} \Gamma_{2,n} h_{n+1} + \Gamma_{2,N^{(t)}-1} |t - t_{N^{(t)}-1}| \right] \Big| \mathcal{H}_T, \quad t \in [0, T], \text{ a.s.} \\
&\hspace{15em} (4.30)
\end{aligned}$$

Step 4: cumulative error bound from time $t = 0$ to time t . Adding the error bound associated with the steps that occur after the last switch (4.30) to the error bound at the last switch (4.27) we have a.s,

$$\begin{aligned} \mathbb{E} [\mathcal{Q}(t) | \mathcal{H}_T] + \mathcal{R}(t) &\leq \Gamma_1 \int_0^t \mathbb{E} [\|E(u)\|^2 | \mathcal{H}_T] du \\ &+ h_{\max}^{2\delta} \mathbb{E} \left[\sum_{n=0}^{N^{(t)}-2} \Gamma_{2,n} h_{n+1} + \Gamma_{2,N^{(t)}-1} |t - t_{N^{(t)}-1}| \middle| \mathcal{H}_T \right], \quad t \in [0, T]. \quad (4.31) \end{aligned}$$

Next lets simplify the LHS of (4.31):

$$\begin{aligned} &\mathbb{E} \left[\mathcal{Q}(t) \middle| \mathcal{H}_T \right] + \mathcal{R}(t) \\ &= \mathbb{E} \left[\sum_{k=0}^{\bar{N}_t-1} \left(\sum_{n=N^{(\tau_k)}-1}^{N^{(\tau_{k+1})}-1} \left(\mathbb{E} [\|E(t_{n+1})\|^2 | \mathcal{F}_{t_n}] - \|E(t_n)\|^2 \right) \mathcal{I}_{\{N^{(\tau_{k+1})} > n\}} \right. \right. \\ &\quad \left. \left. + \sum_{n=N^{(\tau_{\bar{N}_t)}}-1}^{N^{(t)}-2} \left(\mathbb{E} [\|E(t_{n+1})\|^2 | \mathcal{F}_{t_n}] - \|E(t_n)\|^2 \right) \mathcal{I}_{\{N^{(t)} > n+1\}} \right) \right. \\ &\quad \left. + \mathbb{E} [\|E(t)\|^2 | \mathcal{F}_{t_{N^{(t)}-1}}] - \|E(t_{N^{(t)}-1})\|^2 \middle| \mathcal{H}_T \right] \\ &= \mathbb{E} \left[\sum_{n=0}^{N^{(t)}-2} \left(\mathbb{E} [\|E(t_{n+1})\|^2 | \mathcal{F}_{t_n}] - \|E(t_n)\|^2 \right) \mathcal{I}_{\{N^{(t)} > n+1\}} \right. \\ &\quad \left. + \mathbb{E} [\|E(t)\|^2 | \mathcal{F}_{t_{N^{(t)}-1}}] - \|E(t_{N^{(t)}-1})\|^2 \middle| \mathcal{H}_T \right], \quad t \in [0, T], \text{ a.s.} \end{aligned}$$

As $N_{\max}^{(t)} = \lceil t/h_{\min} + \bar{N}_t \rceil$ is $\mathcal{H}_T$ measurable, we can bound $N^{(t)}$ by $N_{\max}^{(t)}$ and move the sum out of the conditional expectation. Again applying the

tower property and now the telescoping sum with $E(0) = 0$ we see that

$$\begin{aligned}
& \mathbb{E} [\mathcal{Q}(t) | \mathcal{H}_T] + \mathcal{R}(t) \\
&= \sum_{n=0}^{N_{\max}^{(t)}-2} \mathbb{E} \left[\mathbb{E} [\|E(t_{n+1})\|^2 \mathcal{I}_{\{N^{(t)} > n+1\}} | \mathcal{F}_{t_n}] - \|E(t_n)\|^2 \mathcal{I}_{\{N^{(t)} > n+1\}} \right. \\
&\quad \left. + \mathbb{E} [\|E(t)\|^2 | \mathcal{F}_{t_{N^{(t)}-1}}] - \|E(t_{N^{(t)}-1})\|^2 | \mathcal{H}_t] \right] \\
&= \mathbb{E} [\|E(t_{N^{(t)}-1})\|^2 | \mathcal{H}_T] - \mathbb{E} [\|E(0)\|^2 | \mathcal{H}_T] \\
&\quad + \mathbb{E} [\|E(t)\|^2 | \mathcal{H}_T] - \mathbb{E} [\|E(t_{N^{(t)}-1})\|^2 | \mathcal{H}_T] \\
&= \mathbb{E} [\|E(t)\|^2 | \mathcal{H}_T], \quad t \in [0, T], \text{ a.s.}
\end{aligned}$$

Substituting (4.32) into (4.31) we have

$$\begin{aligned}
\mathbb{E} [\|E(t)\|^2 | \mathcal{H}_T] &\leq \Gamma_1 \int_0^t \mathbb{E} [\|E(u)\|^2 | \mathcal{H}_T] du \\
&\quad + h_{\max}^{2\delta} \mathbb{E} \left[\sum_{n=0}^{N^{(t)}-2} \Gamma_{2,n} h_{n+1} + \Gamma_{2,N^{(t)}-1} |t - t_{N^{(t)}-1}| \middle| \mathcal{H}_T \right], \quad t \in [0, T], \text{ a.s.}
\end{aligned} \tag{4.32}$$

Now let's consider the second term on the RHS of (4.32). On each trajectory $\omega \in \Omega$ we can define a step function

$$\bar{\Gamma}_2(s)(\omega), \quad \text{where } s \in [t_n(\omega), t_{n+1}(\omega)); \quad n = 0, \dots, N^{(t)}(\omega).$$

Then write the summation as an integral and take the expectation to get

$$\begin{aligned}
& \mathbb{E} \left[h_{\max}^{2\delta} \mathbb{E} \left[\sum_{n=0}^{N^{(t)}-2} \Gamma_{2,n} h_{n+1} + \Gamma_{2,N^{(t)}-1} |t - t_{N^{(t)}-1}| \middle| \mathcal{H}_T \right] \right] \\
&= h_{\max}^{2\delta} \mathbb{E} \left[\mathbb{E} \left[\int_0^t \bar{\Gamma}_2(s) ds \middle| \mathcal{H}_T \right] \right].
\end{aligned}$$

As $\bar{\Gamma}_2(s)(\omega) = \Gamma_{2,n}(\omega)$, $s \in [t_n(\omega), t_{n+1}(\omega))$ and $\mathbb{E}[\Gamma_{2,n}] < \Gamma_2 < \infty$ we see that

$$h_{\max}^{2\delta} \mathbb{E} \left[\mathbb{E} \left[\int_0^t \bar{\Gamma}_2(s) ds \middle| \mathcal{H}_T \right] \right] = h_{\max}^{2\delta} \left[\int_0^t \mathbb{E} [\bar{\Gamma}_2(s)] ds \right] \leq h_{\max}^{2\delta} t \Gamma_2. \tag{4.33}$$

Substitute (4.33) back into (4.32) and take expectations on both sides to get

$$\mathbb{E} [\mathbb{E} [\|E(t)\|^2 | \mathcal{H}_T]] \leq \mathbb{E} \left[\Gamma_1 \int_0^t \mathbb{E} [\|E(u)\|^2 | \mathcal{H}_T] du \right] + h_{\max}^{2\delta} t \Gamma_2, \quad (4.34)$$

and again it follows from the tower property that

$$\mathbb{E} [\|E(t)\|^2] \leq \Gamma_1 \int_0^t \mathbb{E} [\|E(u)\|^2] du + h_{\max}^{2\delta} t \Gamma_2, \quad t \in [0, T]. \quad (4.35)$$

Now we can apply the Gronwall inequality (see Lemma 3.4.1) to get

$$\begin{aligned} \left(\mathbb{E} [\|E(t)\|^2] \right)^{\frac{1}{2}} &\leq \left(h_{\max}^{2\delta} t \Gamma_2 \exp \left\{ \Gamma_1 \int_0^t dr \right\} \right)^{\frac{1}{2}} \\ &\leq \sqrt{h_{\max}^{2\delta} t \Gamma_2 \exp \{ \Gamma_1 t \}} \\ &\leq \sqrt{\Gamma_2 t \exp \{ \Gamma_1 t \}} h_{\max}^{\delta} \\ &\leq C(t) h_{\max}^{\delta}, \quad t \in [0, T]. \end{aligned} \quad (4.36)$$

Taking the maximum over $t \in [0, T]$ on both sides, the bound (4.23) and therefore the statement of the theorem follows. $\square$

Chapter 5

Numerical Analysis

In this chapter we will apply the convergent adaptive method outlined in Chapter 4 to a theoretical biological model as described in [2]. The chosen SDE models the shortening of the telomere length found in DNA replication, of a nest of 75 jackdaws over a time period of thirty days. We introduce an element of Markovian switching to the model by allowing the model to switch between the maximum and minimum observed values based on a Markov chain with a specified generator matrix.

5.1 Implementation of an adaptive numerical method for SDEwMS

First, we simulate the trajectory of a given Markov chain r with state space $\mathbb{S}$ and generator matrix Γ (see Eq. (4.3)). For each sampled $\omega \in \Omega$, we wish to compute the value of $r(\tau_i, \omega)$, along with the respective switching times τ_i , for $i = 0, 1, \dots, \bar{N}_T$. For a trajectory that has just transitioned to state S_i , the holding time before transitioning to a different state is an exponentially distributed random variable with parameter λ_i , where $\lambda_i = \gamma_{i1} + \gamma_{i2} + \dots + \gamma_{iL} - \gamma_{ii}$. The conditional transition probabilities are computed using the entries of Γ as

$$p_{ij} = \frac{\gamma_{ij}}{\gamma_{i1} + \gamma_{i2} + \dots + \gamma_{iL} - \gamma_{ii}}.$$

Then for each state i we can create a probability mass function with these transition probabilities and sample from it to simulate which state S_j is chosen at each switching time. We store the switching times $(\tau_i)_{i \geq 0}$ and the states selected $(S_i)_{i \geq 0}$ as two separate vectors for each trajectory ω to use in the main algorithm.

Then we can begin to construct the mesh, with specific timestepping strategy as described in (4.22). The step size, h_n , taken for each step dictates whether we implement the main map $\mathcal{M}$ or the backstop map $\mathcal{B}$ as per (4.17). In this section we choose as our main map the Milstein map (4.16) and as our backstop map the map associated with the implicit Milstein method. Therefore our adaptive method will be strongly convergent in mean-square with order $\delta = 1$.

Running this method for a specified number of trajectories ω with initial starting point L_0 , we plot the final values on each trajectory $L_T(\omega)$ as a density histogram.

5.2 Example: a stochastic model of telomere length with Markovian switching

Consider the following variant of the SDE (1.2)

$$dL(t) = - (C(t) + A(t)L(t)^2) dt + \sqrt{\frac{1}{3}A(t)L(t)^3} dW(t), \quad (5.1)$$

where the random pair $(C(t), A(t))_{t \geq 0}$ evolves according to a Markov chain that is independent of W and with state space $\mathbb{S} = \{1, 2, 3, 4\}$. Here, each $i \in \mathbb{S}$ corresponds to an element of the set $\mathbb{S}' = \{(c_1, a_1), (c_1, a_2), (c_2, a_1), (c_2, a_2)\}$, where a_i and c_j , $i, j = 1, 2$ are constants calculated from data in [2]. There, different estimation methods provided different values for these constants and so we can consider the use of Markovian switching in this context to be a form of robustness testing. The estimates we will use for our numerical analysis are $(c_1, a_1) = (4.5, 0.22 \times 10^{-6})$, and $(c_2, a_2) = (7.5, 0.41 \times 10^{-6})$, and we will switch between the corresponding states in $\mathbb{S}'$ according to r generated with the algorithm described in Section 5.1. For the purposes of

our numerical demonstration the generator is chosen to be

$$\Gamma = \begin{bmatrix} -0.3 & 0.1 & 0.1 & 0.1 \\ 0.1 & -0.3 & 0.1 & 0.1 \\ 0.1 & 0.1 & -0.3 & 0.1 \\ 0.1 & 0.1 & 0.1 & -0.3 \end{bmatrix}$$

For the adaptive mesh, we set $h_{max} = 3 \times 10^{-2}$, $k = 10$ and $\rho = 15$ throughout.

5.2.1 Numerical Results: Fixed Initial Length L_0

For our fixed initial length we have chosen $L_0 = 1000$ and $T = 30$. In Figure 5.1 we display a histogram of 1000 values, with $\mathbb{E}[L_{30}] \approx 814.33$ base pairs (bp).

In Figure 5.2 we display density histograms produced when the parameters are fixed at $(c, a) = (c_1, a_1)$ (top) and $(c, a) = (c_2, a_2)$ (bottom). The means are computed to be $\mathbb{E}[L_{30}] \approx 862.69$ bp and $\mathbb{E}[L_{30}] \approx 770.75$ bp respectively.

5.2.2 Numerical Results: Uniformly Distributed Initial Lengths L_0

Next we randomly generated 1000 initial lengths from a uniform distribution with range $[4000, 8000]$ and ran the algorithm 100 times for each starting length. In Figure 5.3 we display a line graph comparing initial lengths, and for each initial length the final value of a single trajectory as well as the mean value of each end length.

And finally in Figure 5.4 we display a histogram which shows the mean change in length of the telomere. We calculated the mean change for all 1000 initial lengths, over 100 simulations. The mean change observed over all initial lengths and across all simulated trajectories was found to be $\mathbb{E}[L_5 - L_{30}] \approx -350.74$ bp. The distribution of the simulated data presented here may be compared to the data presented in [2], and we note a qualitative similarity in the mean change histograms. The difference between our estimated mean change and that given in [2] ($\mathbb{E}[L_5 - L_{30}] \approx -223.68$ bp) may be the result of using a model with Markovian parameters (including

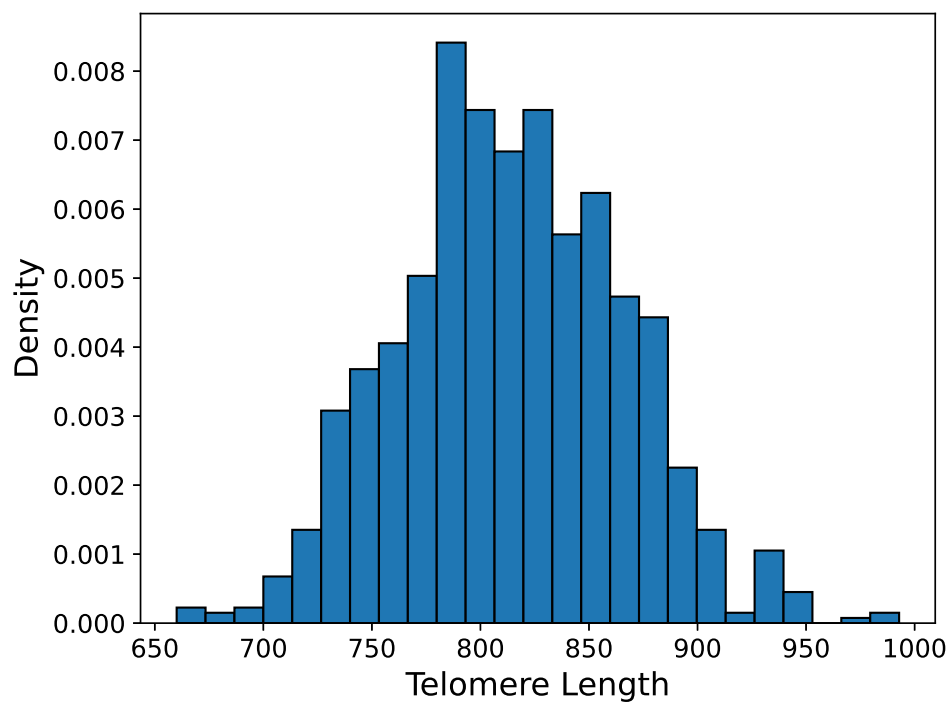


Figure 5.1: Histogram of $L_{30}(\omega)$ for 1000 samples.

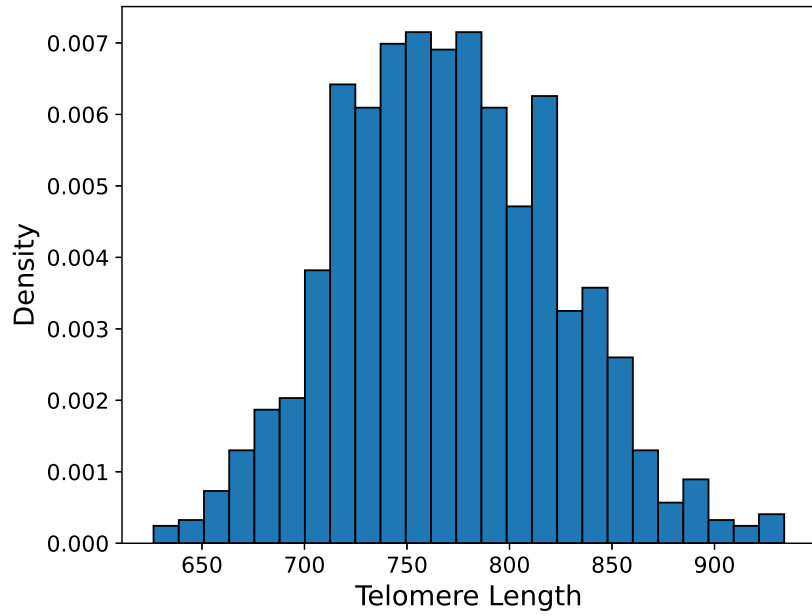
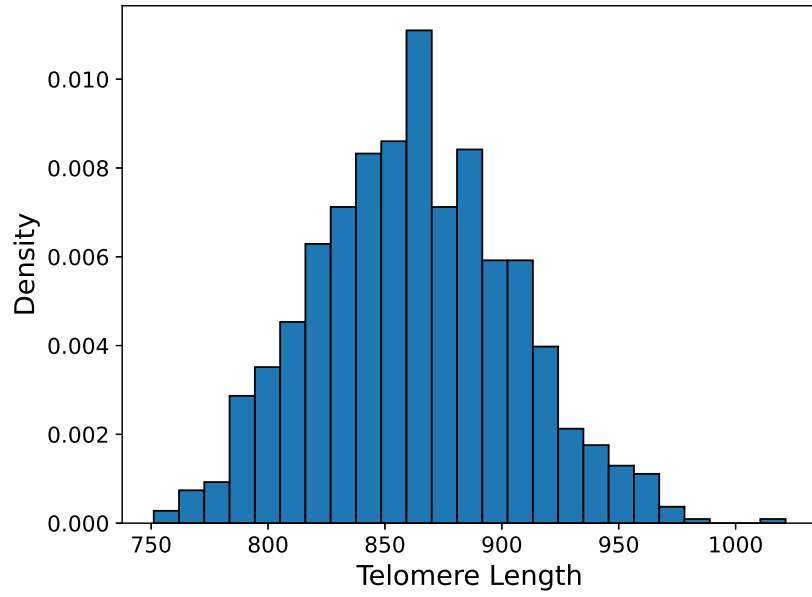


Figure 5.2: Histograms of $L_{30}(\omega)$ for 1000 samples for the SDE without switching and choosing $(c, a) = (c_1, a_1)$ (top) and $(c, a) = (c_2, a_2)$ (bottom).

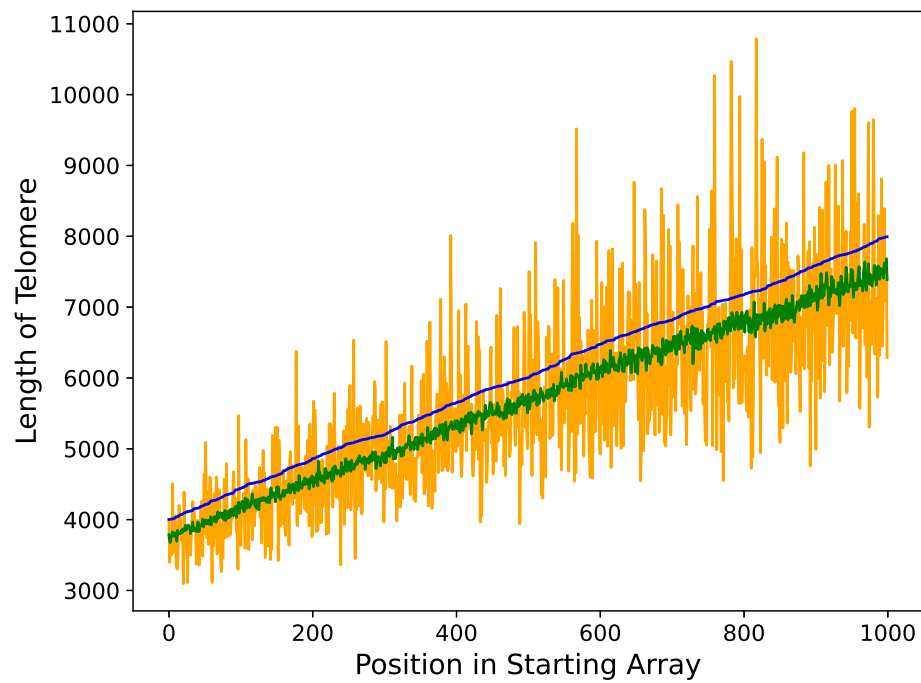


Figure 5.3: Initial telomere lengths ordered (blue), actual final length for a single trajectory (orange), and the mean change in length over all trajectories for a given initial value (green).

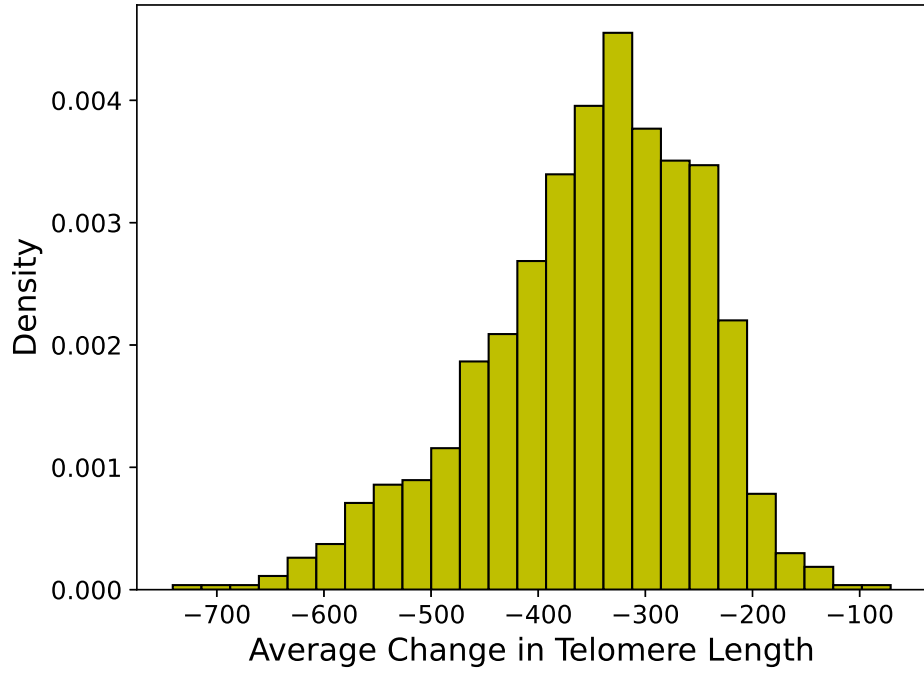


Figure 5.4: Density histogram of the mean change in telomere lengths between day 5 and day 30.

potentially the choice of Γ), as well as possibly reflecting sampling variation and deviations from a uniform distribution of initial values in the real-world data presented in [2]. A more detailed investigation of this is suggested as a future line of research.

Chapter 6

Conclusions

In conclusion having worked up from the construction of a filtered probability space, to the definition and solution of an SDE with Markovian switching, to then proving the strong convergence of a broad class of methods satisfying a one-step error bound, in the non-switching case, over an adaptive mesh, I believe this was a successful and enriching project.

In Chapter 5 many functions were developed to analyse SDEwMS, including a state switching function, a switching-time generator and an SDE solver for SDEs over an adaptive mesh, in the coding language python. The generator matrix (5.2) was chosen as it gave the best number of switching times to conduct a meaningful test of the method within the given time interval. The reasoning behind the choice of $h_{\max}$, k and p were similar, the idea was to minimise the use of the backstop map, ensure a range of values similar to those seen in [2] and to be mindful of computational time. The simulations for 1000 samples with and without switching produced density histograms very similar to those in [2], but it was noted that the mean change was larger than what was seen in the original data. This could be explained by a difference in real world and simulated data or it could be due to the fact that we were switching between the maximum and minimum values for a and c . Further investigation, especially with future models, could focus on the choice of generator matrix and the effect of this on the outputs. It was noted, when running the model with simulated data on single trajectories, that the telemore lengths sometimes finished with a larger value than that of their initial value. Concerned this may be an error in the code, investigation led to Figure 5.3 which shows the average final length across the full 100

simulations ordered and projected onto a graph with the corresponding ordered initial lengths. This showed a noticeable decrease in value, meaning the code was behaving as expected. A possible area to explore is for which initial values, or in which time intervals, does the model stop working. Overall this analysis showed that this new convergent method is accurate and useful.

The highlight of the project was the acceptance of the main result for publication in Delay and Stochastic Differential Equations in Modelling in Finance ICIAM2023-3. The next steps for this project is to attempt to implement this approach in the context of multilevel Monte Carlo methods.

Bibliography

- [1] W.-J. BEYN, E. ISAAK, AND R. KRUSE, *Stochastic C-stability and B-consistency of explicit and implicit Milstein-type schemes*, Journal of Scientific Computing, 70 (2017), pp. 1042–1077.
- [2] J. GRASMAN, H. SALOMONS, AND S. VERHULST, *Stochastic modeling of length-dependent telomere shortening in Corvus monedula*, Journal of Theoretical Biology, 282 (2011), pp. 1–6.
- [3] Q. GUO, W. LIU, X. MAO, AND R. YUE, *The truncated Milstein method for stochastic differential equations with commutative noise*, Journal of Computational and Applied Mathematics, 338 (2018), pp. 298 – 310.
- [4] D. J. HIGHAM, X. MAO, AND L. SZPRUCH, *Convergence, non-negativity and stability of a new Milstein scheme with applications to finance*, Discrete and Continuous Dynamical Systems - B, 18 (2013), pp. 2083–2100.
- [5] M. HUTZENTHALER, A. JENTZEN, AND P. E. KLOEDEN, *Strong and weak divergence in finite time of Euler’s method for stochastic differential equations with non-globally Lipschitz continuous coefficients*, Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences, 467 (2011), pp. 1563–1576.
- [6] M. HUTZENTHALER, A. JENTZEN, AND P. E. KLOEDEN, *Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients*, Annals of Applied Probability, 22 (2012), pp. 1611–1641.
- [7] C. KELLY, *Computation and Simulation for Finance*, Springer, University College Cork, 1 ed., 2024.

- [8] C. KELLY, G. LORD, AND F. SUN, *Strong convergence of an adaptive time-stepping Milstein method for SDEs with monotone coefficients*, BIT Numerical Mathematics, 63 (2023).
- [9] C. KELLY, G. LORD, AND F. SUN, *Strong convergence of a class of adaptive numerical methods for SDEs with jumps*, Accepted in Mathematics and Computer in Simulation, (2024).
- [10] C. KELLY AND G. J. LORD, *Adaptive time-stepping strategies for nonlinear stochastic systems*, IMA Journal of Numerical Analysis, 38 (2018), pp. 1523–1549.
- [11] C. KELLY, G. J. LORD, AND F. SUN, *Strong convergence of an adaptive time-stepping Milstein method for SDEs with monotone coefficients*, BIT Numerical Mathematics, 63 (2023).
- [12] C. KELLY AND K. O'DONOVAN, *Adaptive mesh construction for the numerical solution of stochastic differential equations with markovian switching*, 2024.
- [13] X. MAO, *The truncated Euler–Maruyama method for stochastic differential equations*, Journal of Computational and Applied Mathematics, 290 (2015), pp. 370–384.
- [14] X. MAO AND C. YUAN, *Stochastic Differential Equations with Markovian Switching*, G - Reference, Information and Interdisciplinary Subjects Series, Imperial College Press, 2006.
- [15] J. R. NORRIS, *Markov Chains*, Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 1997.
- [16] X. WANG AND S. GAN, *The tamed Milstein method for commutative stochastic differential equations with non-globally Lipschitz continuous coefficients*, Journal of Difference Equations and Applications, 19 (2013), pp. 466–490.