Large Fluctuations of Stochastic Differential Equations with Regime Switching: Applications to Simulation and Finance

Terry Lynch
B.Sc.

A Dissertation Submitted for the Degree of Doctor of Philosophy
Dublin City University

Supervisor:
Dr. John Appleby
School of Mathematical Sciences
Dublin City University

September 2010

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Doctor of Philosophy in Mathematics is entirely my own work, that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge breach any law of copyright, and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed: ________________________________

ID Number: 52012472

Date: September 2nd 2010

Acknowledgements

I would like to thank my supervisor, Dr. John Appleby, for his endless support and enthusiasm throughout my research. His guidance and insight have been invaluable to me during my time as a postgraduate student and also as an undergraduate.

Thanks is also due to my friends and colleagues (both past and present) in the School of Mathematics for making DCU such a pleasant place to work.

I would also like to thank my parents, Dolores and Patrick, and all of my brothers as well as relatives and friends for all of the support and encouragement they have given me over the years.

This research was supported by the Embark Initiative operated by the Irish Research Council for Science, Engineering and Technology (IRCSET). I thank them for their contribution. I am also indebted to the School of Mathematics in DCU for their financial support during the final year.

Contents

List of Figures

Abstract

This thesis deals with the asymptotic behaviour of various classes of stochastic differential equations (SDEs) and their discretisations. More specifically, it concerns the largest fluctuations of such equations by considering the rate of growth of the almost sure running maxima of the solutions.

The first chapter gives a brief overview of the main ideas and motivations for this thesis. Chapter 2 examines a class of nonlinear finite–dimensional SDEs which have mean–reverting drift terms and bounded noise intensity or, by extension, unbounded noise intensity. Equations subject to Markovian switching are also studied, allowing the drift and diffusion coefficients to switch randomly according to a Markov jump process. The assumptions are motivated by the large fluctuations experienced by financial markets which are subjected to random regime shifts. We determine sharp upper and lower bounds on the rate of growth of the large fluctuations of the process by means of stochastic comparison methods and time change techniques.

Chapter 3 applies similar techniques to a variant of the classical Geometric Brownian Motion (GBM) market model which is subject to random regime shifts. We prove that the model exhibits the same long–run growth properties and deviations from the trend rate of growth as conventional GBM.

The fourth chapter examines the consistency of the asymptotic behaviour of a discretisation of the model detailed in Chapter 3. More specifically, it is shown that the discrete approximation to the stock price grows exponentially and that the large fluctuations from this exponential growth trend are governed by a Law of the Iterated Logarithm.

The results about the asymptotic behaviour of discretised SDEs found in Chapter 4, rely on the use of an exponential martingale inequality (EMI). Chapter 5 considers a discrete version of the EMI driven by independent Gaussian sequences. Some extensions, applications and ramifications of the results are detailed.

The final chapter uses the EMI developed in Chapter 5 to analyse the asymptotic behaviour of discretised SDEs. Two different methods of discretisation are considered: a standard Euler–Maruyama method and an implicit split–step variant of Euler–Maruyama.

Chapter 1

Introduction

This thesis examines the almost sure asymptotic growth rate of the large fluctuations of various classes of stochastic differential equations (SDEs) including equations with Markovian switching and discrete–time approximations of such equations. While Mao and Yuan, [62], have studied the asymptotic behaviour of SDEs with Markovian switching using an exponential martingale and Gronwall inequality approach, this thesis adds to the existing literature by (a) considering a stochastic comparison approach along with a powerful theorem of Motoo, [65], and (b) considering non–linear equations in finite dimensions. Moreover, this thesis examines the large fluctuations of discretised SDEs using the exponential martingale and Gronwall inequality techniques commonly used in continuous–time.

Typically, we characterise the size of these fluctuations by finding upper and lower estimates on the rate of growth of the *running maxima* $t \mapsto \sup_{0\leq s\leq t}|X(s)|$, where $\{X(t)\}_{t\geq 0}$ is the solution of the SDE

$$dX(t) = f(X(t))dt + g(X(t))dB(t), \quad t \geq 0.$$

Here f is known as the *drift* coefficient and g is known as the *diffusion* or *noise* coefficient. Our aim is to find constants C_1 and C_2 and an increasing function $\rho : (0,\infty) \to (0,\infty)$ for which $\rho(t) \to \infty$ as $t \to \infty$ such that

$$0 < C_2 \leq \limsup_{t\to\infty} \frac{\sup_{0\leq s\leq t}|X(s)|}{\rho(t)} \leq C_1, \quad \text{a.s.} \tag{1.0.1}$$

We will refer to such a function ρ as the *essential growth rate* of the largest deviations of the process, with the constants C_1, C_2 being the upper and lower orders of magnitude. Since it can be shown (see, for example, [53]) that

$$\limsup_{t\to\infty} \frac{|X(t)|}{\rho(t)} = \limsup_{t\to\infty} \frac{\sup_{0\leq s\leq t}|X(s)|}{\rho(t)}, \tag{1.0.2}$$

for convenience we will in fact state our results in the manner of the former. In applications, the size of the large fluctuations may represent the largest bubble or crash in a financial market, the largest epidemic in a disease model or a population explosion in an ecological model.

The second chapter considers the size of the large fluctuations of a general class of finite–dimensional SDEs which have stationary solutions. Our focus centres on equations in which the drift term tends to stabilise the solution (we refer to this as *mean–reversion*) and in which the intensity of the stochastic perturbation is bounded (which we refer to as *bounded noise*). These assumptions are suitable for modelling volatilities in a self–regulating economic system which is subjected to persistent stochastic shocks.

We emphasise the importance of the degree of nonlinearity in f in producing the essential growth rate ρ in (1.0.1). To be precise, the largest fluctuations are determined via a scalar function $\Phi(x) := \int_1^x \phi(u)du$, where ϕ determines the degree of nonlinearity and mean–reversion in f.

Our results are then extended to equations which contain Markovian switching features, meaning that the drift and diffusion coefficients can change randomly according to a Markov jump process. In particular, we study an SDE of the form

$$dX(t) = f(X(t), Y(t))\,dt + g(X(t), Y(t))\,dB(t), \quad t \geq 0,$$

where Y is an irreducible Markov chain with finite state space $\mathbb{S}$. The rationale for this in finance is that market sentiment occasionally changes (and often quite rapidly), leading to differing volatility or growth rates. Similarly, observations in financial market econometrics suggest that security prices often move from bearish to bullish (or other) regimes. These regime switches are modelled by the presence of the Markov process Y.

The addition of Markovian switching to the SDE does not play a significant role in determining ρ, the essential rate of growth of the fluctuations of the SDE. It will however have an impact on the constants C_1 and C_2 in (1.0.1), thereby changing the *size* of the largest fluctuations.

Recently, there has been increasing attention devoted to hybrid systems, in which continuous dynamics are intertwined with discrete events. One of the distinct features of such systems is that the underlying dynamics are subject to changes with respect to certain configurations. A convenient way of modelling these dynamics is to use continuous–time

Markov chains to delineate many practical systems where they may experience abrupt changes in their structure and parameters. Such hybrid systems have been considered for the modelling of electric power systems by Willsky and Levy [81] as well as for the control of a solar thermal central receiver by Sworder and Rogers [78]. Athans [12] suggested to use hybrid systems control-related issues in Battle Management Command, Control and Communications (BM/C^3) systems. Sethi and Zhang used Markovian structure to describe hierarchical control of manufacturing systems [74]. Yin and Zhang examined probabilistic structure and developed a two-time-scale approach for control of hybrid dynamic systems [83]. Optimal control of switching diffusions and applications to manufacturing systems were studied in Ghosh, Arapostathis, and Marcus [28] and [29]. In addition, Markovian hybrid systems have also been used in emerging applications in financial engineering [82, 84, 86] and gene regulation [35]. For a detailed treatment of hybrid stochastic differential equations we refer the reader to [62].

After having considered equations with bounded noise, it is a natural question to ask whether or not we can allow the noise to be unbounded while still maintaining similar results. To that end, Chapter 2 also considers equations in which the intensity of the noise term is *unbounded* in the sense that $\lim_{\|x\|\to\infty} \|g(x)\| = +\infty$. We emphasise the importance of the degree of nonlinearity in *both* f and g in producing the essential growth rate ρ in (1.0.1). To be precise, the large fluctuations are determined by the scalar function $\Psi := \int_1^x \phi(u)/\gamma^2(u)du$, where ϕ determines the degree of nonlinearity and mean–reversion in f while γ characterises the degree of nonlinearity in the diffusion g.

Although this research into equations with unbounded noise is substantial, due to the similarities with the equations with bounded noise, we include it only as a subsection and we state without proof some of the main results and methods.

Having considered equations with Markovian switching (which can be used to model rapid financial market changes) in Chapter 2, we then turn our attention to applying these ideas and techniques to a financial market model. This leads us to Chapter 3 where we consider a special class of one–dimensional SDEs which contain Markovian switching and we explore its financial market applications. For this class of SDE, both g and xf are uniformly bounded above and below. We show that the largest deviations of the solution obeys a Law of the Iterated Logarithm, i.e. that the growth function ρ in (1.0.1) takes the form $\sqrt{2t\log\log t}$. Moreover, in the case when the diffusion coefficient depends *only* on the switching parameter, say $g(x,y)=\gamma(y)$, it is shown that

$$\limsup_{t\to\infty}\frac{|X(t)|}{\sqrt{2t\log\log t}}=\sigma_*, \quad \text{a.s.},$$

where $\sigma_*^2=\sum_{j\in\mathbb{S}}\gamma^2(j)\pi_j$ and $\pi=(\pi_j)_{j\in\mathbb{S}}$ is the stationary distribution of the Markov chain Y. These large deviation results are then applied to a security price model, where the security price S obeys

$$dS(t)=\mu S(t)\,dt+S(t)\,dX(t), \quad t\geq 0,$$

where μ is the instantaneous mean rate of growth of the price. This is a variant of the classical Geometric Brownian Motion (GBM) model in which the stock price is the solution of an SDE where the driving Brownian motion is replaced by a semi–martingale which depends on a continuous–time Markov chain. Despite the presence of the Markov process (which introduces *regime shifts*) and an X–dependent drift term (which introduces *market inefficiency*) we can still deduce that the new market model enjoys some of the properties of standard GBM models. In this chapter we also investigate a simple two–state volatility model and show how our results can be implemented in this case.

The introduction of a market model in Chapter 3 raises the question of how this model could be implemented in practice. Chapter 4 facilitates this by considering a discretisation of the model found in Chapter 3. It is shown that one can discretise the model in such a way that the asymptotic behaviour of the discretised model is consistent with that of the continuous–time model of Chapter 3.

Unlike in Chapters 2 and 3, where the proofs rely on stochastic comparison techniques and Motoo's theorem, the proofs for the discrete equations in Chapter 4 use *exponential martingale inequality* (EMI) and *Gronwall inequality* techniques, similar to those used in [54]. We must use these alternative techniques because the proof of Motoo's theorem (a key element of our continuous–time proofs) hinges on an analysis of the excursions of solutions of SDEs which cannot easily be applied in discrete time.

Although there are many discrete versions of the Gronwall inequality, the same is not true of a discrete–time EMI. Nevertheless, a general discrete–time EMI was published by Bercu and Touati, [13], which depends on both the *total* and *predictable* quadratic variations of the martingale (in contrast to the continuous–time EMI which depends only on the predictable quadratic variation). This discrete–time EMI is used to obtain the results in Chapter 4. However, a comparison of the results found in Chapter 3 with their discrete–time analogues in Chapter 4 reveals that the discrete–time results are inferior, due to the use of the general EMI of Bercu and Touati. In Chapter 5 we develop a special class of discrete–time EMI for martingales driven by Gaussian sequences (which naturally arise from an Euler–Maruyama discretisation method). This EMI depends only on the predictable quadratic variation (just as in the continuous–time EMI) and using this EMI instead of the more general EMI of Bercu and Touati yields results which are directly comparable to their continuous–time counterparts.

Having developed a suitable discrete–time EMI, which is very effective in determining the asymptotic behaviour of discretised SDEs, we then return to the asymptotic analysis of discretised SDEs which was started in Chapter 4. In Chapter 6 we consider a different class of SDEs than those considered in Chapter 4, and moreover we consider two different methods of discretisation. While Chapter 4 considers only an Euler–Maruyama discretisation of the SDE, Chapter 6 also considers a *split–step* implicit variant of Euler–Maruyama. On implementing each method, we generally obtain results which are natural discrete analogues of (1.0.1) and are of the form

$$0 < C_2(h) \leq \limsup_{n\to\infty} \frac{|X_h(n)|}{\rho(nh)} \leq C_1(h), \quad \text{a.s.},$$

where h represents the fixed step–size used to produce the discretised process $X_h(n)$. While both discretisation methods obtain similar results, in terms of the asymptotic behaviour of the discretised SDE, they both have benefits and drawbacks which are detailed throughout the chapter.

Mathematical Preliminaries

In this section we define the standard notation used in this thesis as well as useful results used throughout.

1.0.1 Deterministic Preliminaries

Real spaces & vector notation. Let $\mathbb{R}$ denote the set of real numbers and $\mathbb{R}^+$ the set of non-negative real numbers. We denote by $\mathbb{Z}$ the set of all integers, by $\mathbb{N}$ the set of natural numbers (excluding zero) and by $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. For two numbers $x, y \in \mathbb{R}$, $x \vee y$ denotes the maximum of x and y while $x \wedge y$ denotes the minimum of x and y. For any number $x \in \mathbb{R}$, $|x|$ denotes the absolute value of x while $\lfloor x \rfloor$ denotes the integer part of x. Moreover, for any $x \in \mathbb{R}$ we denote $(x)_+ = \max\{a, 0\}$.

Let $\mathbb{R}^d$ denote the set of d–dimensional vectors with entries in $\mathbb{R}$. Vectors $A \in \mathbb{R}^d$ are thought of as column ones. The transpose of a vector $A \in \mathbb{R}^d$ is denoted by A^T and can be thought of as a row vector. Denote by e_i the i^{th} standard basis vector in $\mathbb{R}^d$ with unity in the i^{th} component and zeros elsewhere. Denote by $\langle A, B \rangle$ the standard inner product of vectors $A, B \in \mathbb{R}^d$ and the standard *Euclidean norm,* $\|\cdot\|$, for a vector $A = (a_1, \dots, a_n)^T$ is given by $\|A\|^2 = \sum_{i=1}^n a_i^2$. Moreover we define other norms in $\mathbb{R}^d$ such as the 1–norm, $\|A\|_1 = \sum_{j=1}^d |a_j|$, and the infinity norm (or max norm), $\|A\|_\infty = \max_{1 \le j \le d} |a_j|$. By norm equivalence, there exist numbers $0 < K_1(d) \le K_2(d) < +\infty$ such that

$$K_1(d)\|A\| \le \|A\|_1 \le K_2(d)\|A\|, \qquad A \in \mathbb{R}^d,$$

and the same applies to the infinity norm. We also use the *Cauchy–Schwarz inequality*

$$|\langle A, B \rangle| \le \|A\|\|B\|, \qquad A, B \in \mathbb{R}^d.$$

Matrix notation. Let $\mathbb{R}^{d \times r}$ be the space of $d \times r$ matrices with real entries where I is the *identity matrix* . Let $diag(\alpha_1, \alpha_2, \dots, \alpha_d)$ denote the $d \times d$ matrix with entries

$a_1, a_2, \ldots, a_n$ along the main diagonal and 0 elsewhere. The transpose of a matrix A is denoted by A^T. The *Frobenius norm* of a matrix $A = (a_{ij}) \in \mathbb{R}^{d\times r}$ is denoted $\|A\|_F^2$ and is defined by $\|A\|_F^2 = \sum_{i=1}^{d}\sum_{j=1}^{r} a_{ij}^2$.

Functional notation. We record here some notation for real–valued functions which prove useful throughout the thesis. The deterministic indicator function $1_{\mathbb{N}} : \mathbb{N}_0 \to \{0,1\}$ is defined by

$$1_{\mathbb{N}}(x) = \begin{cases} 1, & \text{if } x \in \mathbb{N}, \\ 0, & \text{if } x = 0. \end{cases}$$

If two functions f, g are asymptotic to each other in the sense that $\lim_{x\to\infty} \frac{f(x)}{g(x)} = 1$, then we use the notation $f \sim g$. We use sgn to denote the signum function, so that $\text{sgn}(x) = 1$ if $x > 0$, $\text{sgn}(x) = -1$ for $x < 0$ and $\text{sgn}(x) = 0$ if $x = 0$. The family of Borel measurable functions $h : [a,b] \to \mathbb{R}^d$ with $\int_a^b |h(x)|^p \, dx < \infty$ are denoted $L^p([a,b];\mathbb{R}^d)$. Finally, $C^1(\mathbb{R})$ is the subspace of $\mathbb{R}$ consisting of continuous functions.

1.0.2 Stochastic Preliminaries

A brief overview of the basic theory concerning stochastic processes is given in this subsection. For a more detailed review see texts such as Mao [54] or Karatzas & Shreve [46].

Probability spaces. We consider the *probability triple* $(\Omega, \mathcal{F}, \mathbb{P})$. Here Ω denotes the sample space where each outcome in Ω is denoted by ω. The family $\mathcal{F}$ is a *σ–algebra* and any set which belongs to $\mathcal{F}$ is said to be *$\mathcal{F}$–measurable* . A *probability measure* $\mathbb{P}$ on the space $(\Omega, \mathcal{F})$ is a function $\mathbb{P} : \mathcal{F} \to [0,1]$. If an event has probability 1 then we say that it is an *almost sure* event and we use the shorthand a.s. A *filtration* $\{\mathcal{F}(t)\}_{t\geq 0}$ is an increasing set of σ–algebras in $\mathcal{F}$. The filtration at time t represents all of the information available up to time t. The filtered probability space is denoted by $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t\geq 0}, \mathbb{P})$.

Standard Brownian Motion. If $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t\geq 0}, \mathbb{P})$ is a filtered probability space then a 1–dimensional standard Brownian motion $\{B(t)\}_{t\geq 0}$ is a process which has the following properties: $B(0) = 0$; the increment $B(t) - B(s)$ is normally distributed with mean 0 and variance $t - s$ where $0 \leq s < t < \infty$; the increment $B(t) - B(s)$ is independent of $\mathcal{F}_s$ where $0 \leq s < t < \infty$.

The Brownian motion has many important properties, such as:

- $\{-B(t)\}$ is a Brownian motion with respect to the same filtration $\{\mathcal{F}(t)\}$,
- $\{B(t)\}$ is a continuous square–integrable martingale with quadratic variation given by $\langle B\rangle(t) = t$ for all $t \geq 0$,
- for almost every $\omega \in \Omega$, the Brownian sample path $t \mapsto B(t, \omega)$ is nowhere differentiable.

Extensions of probability spaces. Let $X = \{X(t), \mathcal{F}(t); 0 \leq t < \infty\}$ be an adapted process on some $(\Omega, \mathcal{F}, \mathbb{P})$. We may need a d–dimensional Brownian motion independent of X, but because $(\Omega, \mathcal{F}, \mathbb{P})$ may not be rich enough to support this Brownian motion, we must extend the probability space to construct this.

Let $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{\mathbb{P}})$ be another probability space, on which we consider a d–dimensional Brownian motion $\hat{B} = \{B(t), \hat{\mathcal{F}}(t); 0 \leq t < \infty\}$, set $\tilde{\Omega} := \Omega \times \hat{\Omega}$ (where $\times$ signifies the Cartesian product), $\tilde{\mathbb{P}} := \mathbb{P} \times \hat{\mathbb{P}}$ (where $\times$ in this case is a product measure), $\tilde{\mathcal{G}} := \mathcal{F} \otimes \hat{\mathcal{F}}$, and define a new filtration by $\tilde{\mathcal{G}}(t) := \mathcal{F}(t) \otimes \hat{\mathcal{F}}(t)$. Here $\mathcal{F} \otimes \mathcal{G}$ defines the product σ–field formed from the σ–fields $\mathcal{F}$ and $\mathcal{G}$, i.e. $\mathcal{F} \otimes \mathcal{G} := \sigma(A \times B; A \in \mathcal{F}, B \in \mathcal{G})$. The new filtration may not satisfy the usual conditions, so we augment it and make it right–continuous by defining $\tilde{\mathcal{F}}(t) := \cap_{s>t}\sigma(\mathcal{G}(s) \cup \mathcal{N})$ where $\mathcal{N}$ is the collection of $\tilde{\mathbb{P}}$–null sets in $\tilde{\mathcal{G}}$. We also complete $\tilde{\mathcal{G}}$ by defining $\tilde{\mathcal{F}} = \sigma(\tilde{\mathcal{G}} \cup \mathcal{N})$. We may extend X and B to $\{\tilde{\mathcal{F}}(t)\}$–adapted processes on

$(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ by defining for $(\omega, \hat{\omega}) \in \tilde{\Omega}$,

$$\tilde{X}\big(t, (\omega, \hat{\omega})\big) = X(t, \omega), \qquad \tilde{B}\big(t, (\omega, \hat{\omega})\big) = B(t, \hat{\omega}).$$

Then $\tilde{B} = \{\tilde{B}(t), \tilde{\mathcal{F}}(t); 0 \le t < \infty\}$ is a d–dimensional Brownian motion independent of $\tilde{X} = \{\tilde{X}(t), \tilde{\mathcal{F}}(t); 0 \le t < \infty\}$.

Stochastic indicator function. If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability triple and A is an event in the σ–algebra $\mathcal{F}$, we denote by $I_A : \Omega \to \{0, 1\}$ the indicator random variable of A, so that

$$I_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A, \\ 0, & \text{if } \omega \notin A. \end{cases}$$

Borel–Cantelli Lemma. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability triple. The first Borel–Cantelli lemma states that if $(A_n : n \ge 1)$ is a sequence of events such that each $A_n \in \mathcal{F}$ and $\sum_{n=1}^{\infty} \mathbb{P}[A_n] < \infty$, then

$$\mathbb{P}[A_n, i.o.] = 0,$$

where $\{A_n, i.o.\}$ is the event that the events A_n are realised infinitely often. The second Borel–Cantelli lemma states that if $(A_n : n \ge 1)$ is a sequence of independent events such that each $A_n \in \mathcal{F}$ and $\sum_{n=1}^{\infty} \mathbb{P}[A_n] = \infty$, then

$$\mathbb{P}[A_n, i.o.] = 1.$$

Sequences of normal random variables Here we state some useful properties of normal random variables.

Let Φ be the distribution of a standard normal (i.e., $\mathcal{N}(0,1)$) random variable N, so that $\Phi(x) := \mathbb{P}[N \le x] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du$, $x \in \mathbb{R}$. Mill's estimate gives us that

$$\frac{1}{\sqrt{2\pi}} \frac{x}{x^2+1} e^{-\frac{x^2}{2}} \le 1 - \Phi(x) \le \frac{1}{\sqrt{2\pi}} \frac{1}{x} e^{-\frac{x^2}{2}}, \quad x > 0. \tag{1.0.3}$$

The following can be obtained from the textbook [75].

Lemma 1.0.1. *If* $Z = \{Z(n) : n \geq 0\}$ *is a sequence of standard normal random variables then*

$$\limsup_{n\to\infty} \frac{|Z(n)|}{\sqrt{2\log n}} \leq 1, \quad a.s.,$$

and if, moreover, the random variables are independent then

$$\limsup_{n\to\infty} \frac{|Z(n)|}{\sqrt{2\log n}} = 1, \quad a.s. \tag{1.0.4}$$

Proof. For every $\varepsilon > 0$, Mill's estimate gives

$$\mathbb{P}[|Z(n)| > \sqrt{2(1+\varepsilon)\log n}] \leq \frac{2}{\sqrt{2\pi}} \frac{1}{\sqrt{2(1+\varepsilon)\log n}} \frac{1}{n^{1+\varepsilon}}.$$

Since the right–hand side is a summable sequence, by the first Borel-Cantelli lemma and by letting $\varepsilon \downarrow 0$ through the rational numbers, we have $\limsup_{n\to\infty} |Z(n)|/\sqrt{2\log n} \leq 1$, a.s. Moreover, if the sequence $Z(n)$ is independent then both sides of Mill's estimate gives

$$\lim_{n\to\infty} \frac{\mathbb{P}[|Z(n)| > \sqrt{2\log n}]}{\frac{1}{\sqrt{\pi}}\frac{1}{\sqrt{\log n}}\frac{1}{n}} = 1.$$

Since the denominator is not a summable sequence, by the second Borel–Cantelli lemma it follows that $\limsup_{n\to\infty} |Z(n)|/\sqrt{2\log n} \geq 1$. Combining both, in the case of independence, we have the equality (1.0.4). □

Law of the Iterated Logarithm (LIL). The following result is one of the most important results on the asymptotic behaviour of standard Brownian motions,

$$\limsup_{t\to\infty} \frac{|B(t)|}{\sqrt{2t\log\log t}} = 1, \quad \text{a.s.}$$

This theorem shows that for any $\varepsilon > 0$ there exists a positive random variable t_ε such that for almost every $\omega \in \Omega$, the Brownian sample path $t \mapsto B(t,\omega)$ is within the interval $\pm(1+\varepsilon)\sqrt{2t\log\log t}$ whenever $t \geq t_\varepsilon(\omega)$.

Markov Chains. Let Y be a continuous–time Markov chain with state space $\mathbb{S}$. We assume that the state space of the Markov chain is finite, say $\mathbb{S} = \{1, 2, \cdots, N\}$. Let the Markov chain have generator $\Gamma = (\gamma_{ij})_{N\times N}$ where

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

and $\Delta > 0$. Here $\gamma_{ij} \geq 0$ is the transition rate from i to j if $i \neq j$ while $\gamma_{ii} = -\sum_{j\neq i} \gamma_{ij}$. It is known (see e.g. [4]) that almost every sample path of $Y(t)$ is a right-continuous step function with a finite number of jumps in any finite subinterval of $[0, \infty)$. As a standing hypothesis we assume in this paper that the Markov chain is *irreducible*. This is equivalent to the condition that for any $i, j \in \mathbb{S}$, one can find finite numbers $i_1, i_2, \cdots, i_k \in \mathbb{S}$ such that $\gamma_{i,i_1}\gamma_{i_1,i_2}\cdots\gamma_{i_k,j} > 0$. Note that Γ always has an eigenvalue 0. The algebraic interpretation of irreducibility is $\text{rank}(\Gamma) = N - 1$. Under this condition, the Markov chain has a unique stationary (probability) distribution $\pi = (\pi_1, \pi_2, \cdots, \pi_N) \in \mathbb{R}^{1\times N}$ which can be determined by solving the following linear equation

$$\pi\Gamma = 0 \quad \text{subject to } \sum_{j=1}^{N} \pi_j = 1 \text{ and } \pi_j > 0 \quad \forall j \in \mathbb{S}. \tag{1.0.5}$$

Moreover, the Markov chain has the very nice ergodic property which states that for any mapping $\phi : \mathbb{S} \to \mathbb{R}$,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t \phi(Y(s))ds = \sum_{j=1}^{N} \phi(j)\pi_j \quad a.s. \tag{1.0.6}$$

In our analysis in this thesis, we will generally have a continuous–time process driven by a Brownian motion and for analytical purposes it is convenient to assume that the Markov process Y is independent of the Brownian motion B. In such a situation, the filtration $\{\mathcal{F}_t\}_{t\geq 0}$ we will work on is the augmentation under $\mathbb{P}$ of the natural filtration generated by the Brownian motion and the Markov chain.

Martingales. The stochastic process $M = \{M(t)\}_{t\geq 0}$ defined on the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}(t)\}_{t\geq 0}, \mathbb{P})$ is said to be a martingale with respect to the filtration $\{\mathcal{F}(t)\}_{t\geq 0}$ if $M(t)$ is $\mathcal{F}(t)$–measurable for all $t \geq 0$, $\mathbb{E}[|M(t)|] < \infty$ for all $t \geq 0$ and

$$\mathbb{E}[M(t)|\mathcal{F}(s)] = M(s), \quad \text{a.s., for all } 0 \leq s \leq t.$$

Doob's continuous martingale representation theorem. Suppose M is a continuous local martingale defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, and the *quadratic variation* $\langle M \rangle$ is an absolutely continuous function of t for $\mathbb{P}$–almost every ω. Then there is an extended space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a one–dimensional Brownian motion $W = \{W(t), \tilde{\mathcal{F}}_t; 0 \leq t < \infty\}$ and an $\tilde{\mathcal{F}}_t$–adapted process X with $\tilde{\mathbb{P}}$–a.s.

$$\int_0^t X^2(s)\, ds < \infty, \quad 0 \leq t < \infty,$$

such that we have the representations $\tilde{\mathbb{P}}$–a.s.

$$M(t) = \int_0^t X(s)\, dW(s), \qquad \langle M \rangle(t) = \int_0^t X^2(s)\, ds, \quad 0 \leq t < \infty.$$

In the proof of the above martingale representation theorem (which can be found in [46], Theorem 3.4.2), the new Brownian motion W is constructed by a continuous local martingale with respect to the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and another Brownian motion, say $\hat{B}$, which is defined on the extended part of $(\Omega, \mathcal{F}, \mathbb{P})$ in $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$. Moreover, $\hat{B}$ is independent of M. Therefore in this thesis, any conclusion made with respect to the extended measure $\tilde{\mathbb{P}}$ about the underlying process with diffusion M defined on $(\Omega, \mathcal{F}, \mathbb{P})$ coincides with that with measure $\mathbb{P}$. Therefore we do not make explicit reference to the probability spaces when stating results.

Stationarity. The following definitions are taken from [31]. The process $U = \{U(t) : t \geq 0\}$, taking values in $\mathbb{R}$, is called *strongly* stationary if the families

$$\{U(t_1), U(t_2), \ldots, U(t_n)\} \quad \text{and} \quad \{U(t_1 + h), U(t_2 + h), \ldots, U(t_n + h)\}$$

have the same joint distribution for all $t_1, t_2, \ldots, t_n$ and $h > 0$. Note that if U is strongly stationary then $U(t)$ has the same distribution for all t.

The process $U = \{U(t) : t \geq 0\}$ is called *weakly* stationary if, for all t_1, t_2 and $h > 0$,

$$\mathbb{E}[U(t_1)] = \mathbb{E}[U(t_2)] \quad \text{and} \quad \mathrm{Cov}[U(t_1), U(t_2)] = \mathrm{Cov}[U(t_1 + h), U(t_2 + h)].$$

Moreover the autocovariance function, $\mathrm{Cov}[U(t), U(t+h)]$, of a weakly stationary process is a function of h only.

1.0.3 Large fluctuations and recurrence of scalar diffusion processes

Here we list some results that are useful in determining the large fluctuations of scalar SDEs using a stochastic comparison approach. Moreover, we can apply these techniques to multi–dimensional equations by first applying a transformation to reduce the equation to a scalar one. Let $\{X(t)\}_{t\geq 0}$ be the scalar solution to the one–dimensional stochastic differential equation

$$dX(t) = b(X(t))\, dt + \sigma(X(t))\, dB(t), \tag{1.0.7}$$

where b is the *drift* coefficient and $\sigma \neq 0$ is the *diffusion* coefficient.

Definition 1.0.1. A *weak solution* in the interval $(0, \infty)$ of equation (1.0.7) is a triple $(X, B), (\Omega, \mathcal{F}, \mathbb{P}), \{\mathcal{F}(t)\}_{t\geq 0}$, with $(\Omega, \mathcal{F}, \mathbb{P})$ and $\{\mathcal{F}(t)\}_{t\geq 0}$ as defined earlier, where:

1. $X = \{X(t), \mathcal{F}(t); 0 \leq t < \infty\}$ is a continuous, adapted $\mathbb{R}^+$-valued process with $X(0) \in (0, \infty)$ and $B = \{B(t), \mathcal{F}(t); 0 \leq t < \infty\}$ is a standard one–dimensional Brownian motion,

2. with $\{l_n\}_{n=1}^{\infty}$ and $\{r_n\}_{n=1}^{\infty}$ strictly monotone sequences satisfying $0 < l_n < r_n < \infty$, $\lim_{n\to\infty} l_n = 0$, $\lim_{n\to\infty} r_n = \infty$ and $S_n := \inf\{t \geq 0 : X(t) \notin (l_n, r_n)\}$ for $n \geq 1$, the following equations hold:

(i) $\mathbb{P}\big[\int_0^{t\wedge S_n}\{|b(X(s))| + \sigma^2(X(s))\}\,ds < \infty\big] = 1; \quad \forall\, 0 \le t < \infty,$

(ii) $\mathbb{P}\big[X(t\wedge S_n) = X(0) + \int_0^t b(X(s))I_{\{s\le S_n\}}\,ds$

$+ \int_0^t \sigma(X(s))I_{\{s\le S_n\}}\,dB(s);\ \forall\, 0\le t<\infty\big] = 1$ valid for every $n \ge 1$.

For more details on the properties of weak solutions see [46].

Scale function and speed measure. Let $I := (l, r)$ with $-\infty \le l < r \le \infty$ and let $b : I \to \mathbb{R}$ and $\sigma : I \to \mathbb{R}$. Moreover, let b and σ satisfy the *non–degeneracy* and *local integrability* conditions:

$$\sigma^2(x) > 0, \quad \forall\, x \in I; \tag{1.0.8}$$

$$\forall\, x \in I, \quad \exists\, \varepsilon > 0 \quad \text{such that} \int_{x-\varepsilon}^{x+\varepsilon} \frac{1+|b(y)|}{\sigma^2(y)}\,dy < \infty. \tag{1.0.9}$$

Under these conditions the scale function and speed measure of X are defined by

$$p(x) = \int_a^x e^{-2\int_a^y \frac{b(u)}{\sigma^2(u)}\,du}\,dy, \qquad a \in \mathbb{R}, \tag{1.0.10}$$

$$m(dx) = \frac{2}{\sigma^2(x)}.\frac{1}{p'(x)}\,dx, \qquad x \in I. \tag{1.0.11}$$

These functions help us to establish the recurrence and stationarity of X in I by Feller's test for explosions (cf. [46]).

Recurrence. A process $\{X(t)\}_{t\ge0}$ is a.s. *recurrent* on, for example $(0,\infty)$, when there exists an a.s. event Ω^* such that for every $t \ge 0$, $\omega \in \Omega^*$ and $x^* \in (0,\infty)$, there exists a $t^*(\omega) > t$ such that $X(t^*(\omega), \omega) = x^*$. Thus, *any* level in $(0,\infty)$ will be attained infinitely many times and there is no "last time" at which a level is attained. The requirements for recurrence are detailed in the following theorem.

Theorem 1.0.1. *Suppose $b, \sigma \in C([0,\infty), \mathbb{R})$, and that X is the weak solution of* (1.0.7) *in* $(0,\infty)$ *with a deterministic initial condition $X(0) \in (0,\infty)$. Assume that σ satisfies*

(1.0.8) *and that both* b *and* σ *satisfy* (1.0.9). *Suppose further that* X *has scale function* p *and speed measure* m *defined by* (1.0.10) *and* (1.0.11) *respectively. Then:*

1. *if* $p(0+) = -\infty$, $p(\infty-) = +\infty$ *and* $m(0,\infty) < +\infty$,

   then X *is recurrent on* $(0,\infty)$.

2. *if* $p(0+) > -\infty$, $p(\infty-) = +\infty$, $m(\{0\}) = 0$ *and* $m[0,\infty) < +\infty$,

   then X *is recurrent on* $[0,\infty)$ *with a reflecting boundary at* 0.

A proof of the recurrence theorem can be found in Chapter 4.12 of [45]. For a more in-depth study of reflecting boundaries we refer the reader to Chapter 7.3 in [70].

Motoo's Theorem This is an important tool for determining the largest deviations for stationary solutions of scalar autonomous stochastic differential equations. We state it here for future use:

Theorem 1.0.2. *Suppose* $b, \sigma \in C([0,\infty), \mathbb{R})$, $\sigma^2(x) > 0$ *for all* $x > 0$ *and that* X *is the weak solution of* (1.0.7) *in* $(0,\infty)$ *with a deterministic initial condition* $X(0) \in (0,\infty)$. *Suppose further that* X *has scale function* p *defined by* (1.0.10).
Then if X *is recurrent we get*

$$\mathbb{P}\left[\limsup_{t\to\infty} \frac{X(t)}{h(t)} \geq 1\right] = 1 \quad or \quad 0,$$

depending on whether

$$\int_c^\infty \frac{1}{p(h(t))}\,dt = \infty \quad or \quad \int_c^\infty \frac{1}{p(h(t))}\,dt < \infty$$

for some $c \in \mathbb{R}$, *where* $h : (0,\infty) \to (0,\infty)$ *is an increasing function with* $h(t) \to \infty$ *as* $t \to \infty$.

A proof of Motoo's theorem can be found in, for example, [45]. In fact, Motoo's theorem can be used to prove the Law of the Iterated Logarith, see [65].

Chapter 2

On the Growth of the Extreme Fluctuations of SDEs with Markovian Switching

2.1 Introduction

In this chapter, we study the almost sure asymptotic growth rate of the *running maxima* $t \mapsto \sup_{0\le s\le t}\|X(s)\|$, where $\{X(t)\}_{t\ge 0}$ is the solution of a finite–dimensional stochastic differential equation (SDE). We study two classes of SDEs: autonomous SDEs and SDEs with Markovian switching.

Since our interest is focussed on unbounded solutions, we consider cases where X obeys

$$\lim_{t\to\infty}\sup_{0\le s\le t}\|X(s)\| = \infty, \quad \text{a.s.}$$

This stipulation covers both recurrent and growing processes, but we make assumptions which ensure that the processes are mean–reverting (in a sense to be later described). In fact, we impose conditions which guarantee that $\liminf_{t\to\infty}\|X(t)\| = 0$, and $\limsup_{t\to\infty}\|X(t)\| = +\infty$, thus ensuring that $\|X\|$ is fluctuating. We characterise the size of these fluctuations by finding upper and lower estimates on the rate of growth of the running maxima. Thus, we find constants C_1 and C_2 and an increasing function $\rho : (0,\infty) \to (0,\infty)$ for which $\rho(t) \to \infty$ as $t \to \infty$ such that

$$0 < C_2 \le \limsup_{t\to\infty}\frac{\|X(t)\|}{\rho(t)} \le C_1, \quad \text{a.s.} \tag{2.1.1}$$

The proofs rely on time change and comparison arguments, constructing upper and lower bounds on $\|X\|$ which are recurrent and stationary processes. The large deviations of these processes are determined by means of a classical theorem of Motoo [65]. In the case when Markovian switching is also present, we ensure that these comparison processes have dynamics which are independent of the switching process.

The first type of equation studied is

$$dX(t) = f(X(t))dt + g(X(t))\,dB(t), \tag{2.1.2}$$

where $f : \mathbb{R}^d \to \mathbb{R}^d$, $g : \mathbb{R}^d \to \mathbb{R}^{d\times r}$ and B is an r–dimensional standard Brownian motion. We also study the stochastic differential equation with Markovian switching

$$dX(t) = f(X(t), Y(t))dt + g(X(t), Y(t))\,dB(t), \tag{2.1.3}$$

where Y is a Markov chain with finite state space $\mathbb{S}$, and $f : \mathbb{R}^d\times\mathbb{S} \to \mathbb{R}^d$, $g : \mathbb{R}^d\times\mathbb{S} \to \mathbb{R}^{d\times r}$ and B is again an r–dimensional Brownian motion.

Our main results in this chapter focus on equations in which the drift term tends to stabilise the solutions (we refer to this phenomenon as *mean–reversion*) and in which the stochastic perturbation has bounded intensity (which we refer to as *bounded* noise). However, our results extend to the case where the stochastic perturbation has unbounded intensity.

These assumptions are suitable for modelling a self–regulating economic system which is subjected to persistent stochastic shocks which (roughly speaking) are stationary processes. By studying finite–dimensional equations, we are able to see how the size of the large fluctuations propagate through the system, and how the interactions between various components of the system influence the dynamics. In fact, we pay particular attention to equations in which the most influential factor driving each component of the process is the degree of mean–reversion of that component on itself. These results therefore find application to models of the spot interest rates of many currency areas which have strong economic (particularly trading) links; the volatilities of many stocks trading on the same exchange, or in the same economic sector; or the prices of a basket of complementary or substitute goods which are subject to stationary shocks on either the supply or demand side. Deterministic nonlinear models of this type in the theory of general equilibrium which exhibit global stability include [63]. Examples of scalar interest rate models can be found in [76] and stochastic volatility models in, for example, [26, 42, 69].

Stochastic differential equations with stationary solutions have found favour in modelling the evolution of the volatility of risky assets. This is in part because they can produce "heavy tails" in the distribution of the returns of risky assets present in real markets, see e.g. [32, 66]. In fact, the rate of decay of the tails in the stationary distribution of the volatility can be related directly to the rate of decay of the tails of the asset returns' distribution. Moreover, it is well–known from the one–dimensional theory of SDEs that there is a direct relationship between the rate of decay of the tails of the distributions of a stationary solution of an autonomous SDE and the rate of growth of the a.s. running maxima of the solution, see for example [15]. Thus, our analysis facilitates in the investigation of heavy tailed returns' distributions in stochastic volatility market models in which many assets are traded. Furthermore, one can still analyse the large fluctuations (and thereby the tails of the distributions) in the case when the market can switch between various regimes, [21].

By keeping the intensity of the stochastic term bounded, we are able to study more directly the impact of different restoring forces of the system towards its equilibrium value. The strength of the restoring force is characterised here by a scalar function $\phi : [0,\infty) \to [0,\infty)$ with $x\phi(x) \to \infty$ as $x \to \infty$, where

$$\limsup_{\|x\|\to\infty} \frac{\langle x, f(x)\rangle}{\|x\|\phi(\|x\|)} \leq -c_2 \in (-\infty, 0). \tag{2.1.4}$$

Therefore, the strength of the mean–reversion is greater the more rapidly that $x \mapsto x\phi(x)$

increases. We ensure that the degree of nonlinearity in f is characterised by ϕ also by means of the assumption

$$\limsup_{\|x\|\to\infty} \frac{|\langle x, f(x)\rangle|}{\|x\|\phi(\|x\|)} \le c_1 \in (0,\infty). \tag{2.1.5}$$

In our main result, we show how the function ρ in (2.1.1) depends directly on ϕ. Therefore, up to the constants C_1 and C_2 in (2.1.1), we are able to characterise the rate of growth of the largest fluctuations of the solutions. Moreover, we can show that these recover the best possible results that are available in the one–dimensional case. As might be expected, the weaker the strength of the mean–reversion, the more slowly that $x \mapsto x\phi(x)$ increases, which leads to a more rapid rate of growth of $\rho(t) \to \infty$ as $t \to \infty$; consequently, as we might expect, weak mean–reversion results in large fluctuations in the solution. The contribution here, of course, is our ability to quantify the relation between the degree of mean–reversion and the size of the fluctuations.

We also study the large fluctuations of the equation (2.1.3) with Markovian switching in this chapter. There has been a lot of work done on the stability and stabilisation of such equations, as seen in [1, 2, 3, 30, 49, 55, 59, 72, 73, 85]. However, to the authors' knowledge less is known about the asymptotic behaviour (and in particular the large deviations) of unstable equations. Despite this, an interesting contribution to the theory of SDEs with Markovian switching in which solutions are not converging to a point equilibrium is given in [51].

In [51], it is shown that highly nonlinear equations suitable for modelling population dynamics exhibit stationary–like behaviour, possessing bounded time average second moments and being stochastically ultimately bounded. Indeed such results should enable upper bounds on the pathwise growth of the running maxima to be established by means of standard Borel–Cantelli and interpolation arguments.

In [60] for example, conditions are given under which an SDE with Markovian switching of the form (2.1.3) admits an asymptotically stationary solution. The analysis in this chapter relates closely to [60] and [51]: we determine the size of large fluctuations but for a more general class of problems.

In this chapter, we emphasise the importance of the degree of nonlinearity in f in producing the essential growth rate ρ in (2.1.1), as the presence of Markovian switching does not seem to play a significant role in determining ρ. However, this does not mean that the switching process does not play a significant role in influencing the size of the largest fluctuation up to a given time. We conjecture that the switching process may have a significant impact on the constants C_1 and C_2 in (2.1.1), thereby changing significantly the size of the largest fluctuations compared to equations which have the same degree of nonlinearity, but are not subject to switching. Some evidence of this conjecture appears in Chapter 3, in which the essential rate of growth of the running maxima of a non–stationary process is governed by the Law of the Iterated Logarithm, but the constants C_1 and C_2 (which are equal) depend on the stationary distribution of the switching process.

In our analysis in this chapter, we focus on equations possessing *stationary* solutions, or which are in some sense close to equations possessing a stationary solution. Some analysis on the large fluctuations of a particular class of scalar SDEs which have dynamics close to a *non–stationary* process is presented in Chapter 3. For the proofs in this chapter we reduce the SDE to a scalar equation by means of time and coordinate changes and use a combination of stochastic comparison techniques and Motoo's theorem (cf. [65]) to determine the asymptotic behaviour. On the other hand, while Chapter 3 deals with a very special class of nonlinear functions f and g, in this chapter we consider much more general equations.

The chapter is organised as follows. Section 2.2 details the method of proof used in this chapter which is an alternative to the deterministic methods used by Mao in [54] for example. We give a brief overview of a useful class of functions in Section 2.3. A synopsis and discussion of the main results for equations without switching is given in Section 2.4 while the extensions to equations with switching and to equations with unbounded noise are given in Sections 2.5 and 2.6 respectively. Proofs can be found from Section 2.7 onwards.

2.2 Stochastic comparison technique

To prove the results in this chapter we use techniques which rely on stochastic comparison principles and Motoo's theorem. The first step of this technique is to reduce the d–dimensional equation (2.1.2) to a scalar equation, using Itô's formula, to which we can apply the stochastic comparison theorem detailed below. The idea is to manufacture a scalar comparison process which has the same diffusion coefficient as the equation we wish to compare it to, while uniform bounds (in the space variable) on the drift coefficient allows us to create an *upper* comparison process or a *lower* comparison process. This then allows us to analyse the asymptotic behaviour of the comparison processes (using Motoo's theorem) rather than analysing the original process. By construction the comparison processes will have recurrent and stationary solutions, a requirement of Motoo's theorem.

The stochastic comparison theorem is stated here and its proof can be found in Section 2.9.

Theorem 2.2.1. *Let B be a one–dimensional $\mathcal{F}(t)$–adapted Brownian motion and suppose that X_1 and X_2 are $\mathcal{F}(t)$–adapted processes restricted to $[0,\infty)$ which obey*

$$X_i(t) = X_i(0) + \int_0^t \beta_i(s)\,ds + \int_0^t \sigma(X_i(s))\,dB(s), \quad t \geq 0, \quad i = 1, 2,$$

where the β_i *are also* $\mathcal{F}(t)$*–adapted. Suppose also that there exists* $b : \mathbb{R}^+ \to \mathbb{R}$ *with*

$$\beta_1(t) \geq b(X_1(t)), \quad b(X_2(t)) \geq \beta_2(t), \quad t \geq 0. \tag{2.2.1}$$

Suppose further that $X_1(0) \geq X_2(0)$*, a.s. and that for every* $n \in \mathbb{N}$ *there exists* $K_n > 0$ *such that*

$$|\sigma(x) - \sigma(y)| \leq K_n\sqrt{|x-y|}, \quad \text{for all } x, y \in [0, n], \tag{2.2.2}$$

$$|b(x) - b(y)| \leq K_n|x - y|, \quad \text{for all } x, y \in [0, n]. \tag{2.2.3}$$

Define $\tau_n^{(1)} = \inf\{t \geq 0 : X_1(t) = n\}$ *and* $\tau_n^{(2)} = \inf\{t \geq 0 : X_2(t) = n\}$ *and assume that either* $\tau_n^{(1)} < +\infty$ *or* $\tau_n^{(2)} < +\infty$ *a.s. Then* $X_1(t) \geq X_2(t)$ *for all* $t \geq 0$ *a.s.*

2.3 Regular Variation

In this chapter, some of our analysis is facilitated by the use of *regularly varying* functions, see [14]. We give some of their properties in this section. In its basic form, regular variation may be viewed as the study of relations such as

$$\lim_{x\to\infty} \frac{f(\lambda x)}{f(x)} = \lambda^\zeta \in (0, \infty) \quad \forall\, \lambda > 0,$$

where f is a positive measurable function and we say that f is regularly varying at infinity with index ζ, i.e. $f \in RV_\infty(\zeta)$. By the representation theorem (Thm 1.3.1 in [14]), if $f \in RV_\infty(\zeta)$ then there exists a measurable function c and a continuous function b such that $f(x) = c(x)\exp\{\int_1^x b(u)/u\,du\}$ for $x \geq 1$, where $c(x) \to c \in (0, \infty)$ and $b(x) \to \zeta$ as $x \to \infty$. Taking logs and using L'Hôpital's rule, we get the following useful result

$$\frac{\log f(x)}{\log x} = \frac{\log c(x)}{\log x} + \frac{\int_1^x \frac{b(u)}{u}\,du}{\log x} \to \zeta \quad \text{as } x \to \infty. \tag{2.3.1}$$

A positive function f defined on some neighbourhood of infinity *varies smoothly* with index $\alpha \in \mathbb{R}$, denoted $f \in SV_\infty(\alpha)$, if $h(x) := \log f(e^x)$ is C^∞, and

$$\lim_{x\to\infty} h'(x) = \alpha, \quad \lim_{x\to\infty} h^{(n)}(x) = 0 \quad n = 2, 3, \ldots$$

From the definition of h, it can easily be shown that $h'(\log(x)) = xf'(x)/f(x)$. Therefore, for a smoothly varying continuous function $f \in SV_\infty(\alpha)$,

$$\lim_{x\to\infty} \frac{xf'(x)}{f(x)} = \alpha \quad \text{and moreover,} \quad SV_\infty(\alpha) \subset RV_\infty(\alpha). \tag{2.3.2}$$

The above limit also allows us to determine whether the function f is increasing or decreasing. More precisely, if $f(x) > 0$ for $x > 0$ and $\alpha > 0$, then we must also have $f'(x) > 0$, and so f is increasing. In fact, by Theorem 1.5.3 in [14], any function f varying regularly with non–zero exponent is asymptotic to a monotone function.

Also of great importance is the fact that

$$f \in RV_\infty(\zeta) \Rightarrow F(x) := \int_1^x f(u)\,du \in RV_\infty(\zeta + 1). \tag{2.3.3}$$

To show this result, note that if it were true we should have $xF'(x)/F(x) = xf(x)/F(x) \to \zeta + 1$ as $x \to \infty$. Applying Karamata's theorem (Thm 1.5.11 in [14]) with $\sigma = 0$ gives us precisely this result, i.e. $xf(x)/\int_1^x f(u)\,du \to \zeta + 1$ as $x \to \infty$, where $f \in RV_\infty(\zeta)$ and is locally bounded on $[1, \infty)$ and $\zeta > -1$.

One theorem which is of particular use is the *smooth variation theorem*, (see Thm 1.8.2, [14] for proof).

Theorem 2.3.1. *If $f \in RV_\infty(\alpha)$, then there exists $f_1, f_2 \in SV_\infty(\alpha)$ with $f_1 \sim f_2$ and $f_1 \le f \le f_2$ on some neighbourhood of infinity. In particular, if $f \in RV_\infty(\alpha)$ there exists $g \in SV_\infty(\alpha)$ with $g \sim f$.*

Also, the following theorem gives a very useful property of the *inverse*, (see Thm 1.8.5, [14] for proof)

Theorem 2.3.2. *If $f \in SV_\infty(\alpha)$ with $\alpha > 0$ then, on some neighbourhood of infinity, f possesses an inverse function $g \in SV_\infty(1/\alpha)$ with $f(g(x)) = g(f(x)) = x$.*

Combining both of these theorems we get the following lemma:

Lemma 2.3.1. *If there exists a continuous and positive function $f \in RV_\infty(\zeta)$ with $\zeta > -1$, then $F(x) := \int_1^x f(u)du$ possesses an inverse function $F^{-1} \in RV_\infty(\frac{1}{\zeta+1})$.*

Proof of Lemma 2.3.1. We have $f \in RV_\infty(\zeta)$ and, by (2.3.3), $F \in RV_\infty(\zeta+1)$. Moreover, by Theorem 2.3.1 there exists $F_1 \in SV_\infty(\zeta+1)$ such that $F(x)/F_1(x) \to 1$ as $x \to \infty$. So $\forall\, \varepsilon \in (0,1)$ there exists $x(\varepsilon) > 0$ such that

$$(1-\varepsilon)F_1(x) < F(x) < (1+\varepsilon)F_1(x), \quad \text{for } x > x(\varepsilon). \tag{2.3.4}$$

Note that F^{-1} exists and is increasing since f is positive and, by Theorem 2.3.2, there exists $F_1^{-1} \in SV_\infty(\frac{1}{\zeta+1})$. Applying F^{-1} to (2.3.4) we have

$$F^{-1}\big((1-\varepsilon)F_1(x)\big) < x < F^{-1}\big((1+\varepsilon)F_1(x)\big), \quad \text{for } x > x(\varepsilon).$$

Taking the left hand side of the inequality, let $y = (1-\varepsilon)F_1(x)$. Then $F^{-1}(y) < F_1^{-1}(\frac{y}{1-\varepsilon})$ for $y > (1-\varepsilon)F_1(x(\varepsilon))$. Similarly, taking the right hand side of the inequality, let $z = (1+\varepsilon)F_1(x)$. Then $F^{-1}(z) > F_1^{-1}(\frac{z}{1+\varepsilon})$ for $z > (1+\varepsilon)F_1(x(\varepsilon))$. Combine both of these by letting $u := \max(y,z) > (1+\varepsilon)F_1(x(\varepsilon))$ and divide across by $F_1^{-1}(u)$ to get

$$\frac{F_1^{-1}(\frac{u}{1+\varepsilon})}{F_1^{-1}(u)} < \frac{F^{-1}(u)}{F_1^{-1}(u)} < \frac{F_1^{-1}(\frac{u}{1-\varepsilon})}{F_1^{-1}(u)}.$$

Since $F_1^{-1}(\lambda u)/F_1^{-1}(u) \to \lambda^{\frac{1}{\zeta+1}}$ as $u \to \infty$ we can let $\varepsilon \to 0$ to get $F^{-1}(u)/F_1^{-1}(u) \to 1$ as $u \to \infty$. Therefore, as $F_1^{-1} \in SV_\infty(\frac{1}{\zeta+1}) \Rightarrow F_1^{-1} \in RV_\infty(\frac{1}{\zeta+1})$, it follows that $F^{-1} \in RV_\infty(\frac{1}{\zeta+1})$ also. □

2.4 Main Results: Equations without Switching

Let $f : \mathbb{R}^d \to \mathbb{R}^d$ and $g : \mathbb{R}^d \to \mathbb{R}^{d\times r}$ be continuous functions obeying local Lipschitz continuity conditions. Let $X(0) = x_0$ and consider the SDE given by

$$dX(t) = f(X(t))\,dt + g(X(t))\,dB(t), \qquad t \geq 0. \tag{2.4.1}$$

We make the standing assumption throughout the chapter that f and g obey this continuity restriction. For economy of exposition these assumptions are not explicitly repeated in the statement of theorems in this chapter. Under these conditions, there exists a unique local solution of (2.4.1).

We write $f_i(x) = \langle f(x), \mathbf{e}_i\rangle$, $i = 1, \dots, d$ and $g_{ij}(x)$ to be the (i, j)–th entry of the $d \times r$ matrix g with real–valued entries. Then the i^{th} component of (2.4.1) is

$$dX_i(t) = f_i(X(t))dt + \sum_{j=1}^{r} g_{ij}(X(t))dB_j(t). \tag{2.4.2}$$

2.4.1 Statement of main results

In what follows, it is convenient to introduce a function ϕ with the following properties:

$$\phi : [0,\infty) \to (0,\infty) \text{ and } x\phi(x) \to \infty \text{ as } x \to \infty, \tag{2.4.3a}$$

$$\phi \text{ is locally Lipschitz continuous on } [0,\infty). \tag{2.4.3b}$$

We often request that ϕ and f possess the following properties also:

$$\text{there exists } c_1 > 0 \text{ such that } \limsup_{\|x\|\to\infty} \frac{|\langle x, f(x)\rangle|}{\|x\|\phi(\|x\|)} \le c_1, \tag{2.4.4}$$

$$\text{there exists } c_2 > 0 \text{ such that } \limsup_{\|x\|\to\infty} \frac{\langle x, f(x)\rangle}{\|x\|\phi(\|x\|)} \le -c_2. \tag{2.4.5}$$

We define the function Φ according to

$$\Phi(x) = \int_1^x \phi(u)\,du, \quad x \ge 1. \tag{2.4.6}$$

Since $x\phi(x) \to \infty$ as $x \to \infty$ it follows that $\Phi(x) \to \infty$ as $x \to \infty$. Therefore, since Φ is increasing, Φ^{-1} exists and $\Phi^{-1}(x) \to \infty$ as $x \to \infty$ also.

We suppose that the noise is *bounded* by imposing the following hypotheses:

$$\text{there exists } K_2 > 0 \text{ such that } \|g(x)\|_F \le K_2, \quad \text{where} \quad \|g(0)\|_F > 0, \tag{2.4.7}$$

$$\text{there exists } K_1 > 0 \text{ such that } \inf_{\|x\|\in\mathbb{R}^d/\{0\}} \frac{\sum_{j=1}^{r}\left(\sum_{i=1}^{d} x_i g_{ij}(x)\right)^2}{\|x\|^2} \ge K_1^2. \tag{2.4.8}$$

Observe that by Cauchy–Schwarz, condition (2.4.8) implies $\|g(x)\|_F^2 \geq K_1^2$.

As mentioned before, under the local Lipschitz continuity conditions on f and g, there exists a unique *local* solution of (2.4.1). However we can now show, using the additional hypotheses above, that in fact there exists a unique *global* solution to (2.4.1).

Note that by (2.4.5) there exists x_1 such that $\langle x, f(x)\rangle < 0$ for all $\|x\| \geq x_1$ and by (2.4.7), $\|g(x)\|_F \leq K_2$ for all $x \in \mathbb{R}^d$. Thus, $\sup_{\|x\|\geq x_1}\{\langle x, f(x)\rangle + \frac{1}{2}\|g(x)\|_F^2\} \leq \frac{1}{2}K_2^2$ and, by continuity, $\sup_{\|x\|\leq x_1}\{\langle x, f(x)\rangle + \frac{1}{2}\|g(x)\|_F^2\} =: C(x_1)$. Combining both,

$$\sup_{x\in\mathbb{R}^d}\left\{\langle x, f(x)\rangle + \frac{1}{2}\|g(x)\|_F^2\right\} \leq \frac{1}{2}K_2^2 + C(x_1) < +\infty.$$

As a result of this global one–sided bound, Theorem 3.6 in [54] states that there exists a unique global solution to equation (2.4.1).

We are now in a position to state our main results. Our first result shows that when the noise is bounded, and f obeys the upper bound (2.4.4), a lower bound on the rate of growth of the running maxima of $\|X\|$ can be obtained.

Theorem 2.4.1. *Suppose there exists a function ϕ satisfying* (2.4.3)*, and that ϕ and f satisfy* (2.4.4)*, and that g obeys* (2.4.7) *and* (2.4.8)*. Then X, the unique adapted continuous solution satisfying* (2.4.1)*, satisfies for any $\varepsilon \in (0,1)$*

$$\limsup_{t\to\infty} \frac{\|X(t)\|}{\Phi^{-1}\left(\frac{K_1^2(1-\varepsilon)}{2c_1}\log t\right)} \geq 1 \quad a.s.\ on\ \Omega_\varepsilon, \tag{2.4.9}$$

where Φ is defined by (2.4.6) *and Ω_ε is an almost sure event.*

The next result shows that when the noise is bounded, and f obeys the mean–reversion property (2.4.5), an upper bound on the rate of growth of the running maxima of $\|X\|$ can be obtained.

Theorem 2.4.2. *Suppose there exists a function ϕ satisfying* (2.4.3)*, and that ϕ and f satisfy* (2.4.5)*, and that g obeys* (2.4.7) *and* (2.4.8)*. Then X, the unique adapted contin-*

uous solution satisfying (2.4.1)*, satisfies for any* $\varepsilon \in (0,1)$

$$\limsup_{t\to\infty} \frac{\|X(t)\|}{\Phi^{-1}\left(\frac{K_2^2(1+\varepsilon)}{2c_2}\log t\right)} \leq 1 \quad a.s. \text{ on } \Omega_\varepsilon, \tag{2.4.10}$$

where Φ *is defined by* (2.4.6) *and* Ω_ε *is an almost sure event.*

Observe that results (2.4.9) and (2.4.10) do not preclude the case where $\|X(t)\|$ is growing (i.e. $\|X(t)\| \to \infty$ as $t \to \infty$) at a rate characterised by $\Phi^{-1}(c\log t)$. However, the next theorem shows that the behaviour of (2.4.9) and (2.4.10) arises from the *fluctuations* of $\|X\|$ rather than the *growth* of $\|X\|$. Indeed, it is Theorem 2.4.3 which allows us to claim that these are results about the growth of large fluctuations.

Theorem 2.4.3. *If X, the unique adapted continuous solution satisfying* (2.4.1)*, satisfies Theorems 2.4.1 and 2.4.2, then* $\|X\|$ *is recurrent on* $(0,\infty)$*. Furthermore, X obeys*

$$\liminf_{t\to\infty}\|X(t)\| = 0, \ a.s. \text{ and } \limsup_{t\to\infty}\|X(t)\| = +\infty, \ a.s.$$

Taking Theorems 2.4.1 and 2.4.2 together, in the special case where ϕ is a regularly varying function, we obtain the following result which characterises the essential almost sure rate of growth of the running maxima of $\|X\|$.

Theorem 2.4.4. *Suppose there exists a function* $\phi \in RV_\infty(\zeta)$ *satisfying* (2.4.3)*, and that* ϕ *and* f *satisfy* (2.4.4) *and* (2.4.5)*, and that* g *obeys* (2.4.7) *and* (2.4.8)*. Then* X*, the unique adapted continuous solution satisfying* (2.4.1)*, satisfies*

$$\left(\frac{K_1^2}{2c_1}\right)^{\frac{1}{\zeta+1}} \leq \limsup_{t\to\infty}\frac{\|X(t)\|}{\Phi^{-1}(\log t)} \leq \left(\frac{K_2^2}{2c_2}\right)^{\frac{1}{\zeta+1}} \quad a.s., \tag{2.4.11}$$

where Φ *is defined by* (2.4.6) *and* $\zeta > -1$*.*

Remark 2.4.1. It is interesting to ask whether the asymptotic estimate in (2.4.11) is sharp. Although this is a difficult question to address in general, we supply now a scalar example which demonstrates that, in some cases at least, the asymptotic estimate (2.4.11) is unimprovable.

Let $c > 0$ and $K > 0$ and consider a simple one-dimensional Ornstein–Uhlenbeck process

$$dX(t) = -cX(t)dt + K\,dB(t), \quad t \geq 0.$$

In the notation of this section, and Theorem 2.4.4 in particular, we have that $d = r = 1$, $f(x) = -cx$, and $g(x) = K$. This implies that $c_1 = c_2 = c$, $K_1 = K_2 = K$, and that $\phi(x) = x$ so $\phi \in \text{RV}_\infty(1)$. This means that $\zeta = 1$. Thus $\Phi(x) = x^2/2$ and $\Phi^{-1}(x) = \sqrt{2x}$. Then applying Theorem 2.4.4 we recover the well–known result

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{2\log t}} = \left(\frac{K^2}{2c}\right)^{\frac{1}{2}}, \quad \text{a.s.}$$

Remark 2.4.2. It is worth mentioning that we prefer hypotheses of the type (2.4.4) and (2.4.5) on f, as opposed to global estimates, because we only require control on the drift for large values of $\|x\|$ in order to obtain asymptotic results. Intuitively, we would not expect the behaviour of the drift for small and moderate values of $\|x\|$ to have an impact on the large deviations, so it is natural not to require hypotheses which explicitly deal with these moderate values of $\|x\|$. As a result of this we can obtain sharper asymptotic estimates, in particular we can obtain better estimates on the constants c_1 and c_2 on the right hand side of (2.4.4) and (2.4.5). The downside is that the proofs become slightly more cumbersome as we have to ensure that the drift is well behaved for small and moderate values of $\|x\|$.

Remark 2.4.3. We remark that hypotheses (2.4.7) and (2.4.8) on g are satisfied for certain equations with additive or bounded noise. For instance, consider the case $g(x) = \Sigma\,\theta(x)$ where $\theta : \mathbb{R}^d \to \mathbb{R}$ is a locally Lipschitz continuous function such that there exists $\theta_1, \theta_2 \in (0,\infty)$ with $\theta_1 \leq |\theta(x)| \leq \theta_2$ for all $x = (x_1, x_2, \ldots, x_d)^T \in \mathbb{R}^d$. Also, Σ is a $d \times r$ matrix ($d \leq r$) such that $\Sigma \neq 0$ and the nullspace of Σ^T, denoted $\text{null}(\Sigma^T)$, contains only the zero vector, where the nullspace is the solution set of $\Sigma^T x = 0$. Under these conditions, (2.4.7) and (2.4.8) hold. Also, if θ is constant then we have additive noise, otherwise we have bounded noise.

To demonstrate that (2.4.7) and (2.4.8) do in fact hold, note that on the one hand $\|g(x)\|_F = |\theta(x)|.\|\Sigma\|_F$ but since $\|\Sigma\|_F$ is constant and $|\theta|$ is bounded above, it follows that $\|g(x)\|_F$ is bounded above as required in (2.4.7). On the other hand,

$$\frac{\sum_{j=1}^{r}\left(\sum_{i=1}^{d} x_i g_{ij}(x)\right)^2}{\|x\|^2} = \theta^2(x)\frac{\sum_{j=1}^{r}(\sum_{i=1}^{d} x_i \Sigma_{ij})^2}{\|x\|^2} = \theta^2(x)\frac{\|\Sigma^T x\|^2}{\|x\|^2}.$$

As mentioned before, $|\theta|$ is bounded below so we just require $\inf_{\|x\|\neq 0} \|\Sigma^T x\|/\|x\| > 0$ in order for (2.4.8) to hold. However, the only way that we would not have a positive lower bound here is if there exists $y \in \mathbb{R}^d/\{0\}$ such that $\Sigma^T y = 0$. In other words, we require $\Sigma^T y \neq 0$ for all $y \in \mathbb{R}^d/\{0\}$. This means that the unique solution of $\Sigma^T y = 0$ must be $y = 0$ and this is equivalent to $\text{null}(\Sigma^T) = \{0\}$.

Note that in the $d = r$ case, where Σ is a square matrix, $\text{null}(\Sigma^T) = \{0\}$ is true if and only if Σ^T is invertible, which is true if and only if Σ is invertible.

If $d < r$ then Σ^T is an $r \times d$ matrix giving rise to the system $\Sigma^T x = b$, for some $b \in \mathbb{R}^d$, which has more equations than unknowns. Let Σ_1 be a $d \times d$ matrix formed by taking any d rows of Σ^T in such a way that Σ_1 is invertible. Then, after row reduction, the first d rows of Σ^T will be the $d \times d$ identity matrix and the remaining $(r - d)$ rows will have all zero entries. Thus, by well–known matrix properties, the system $\Sigma^T x = 0$ has the unique solution $x = 0$, which guarantees $\text{null}(\Sigma^T) = \{0\}$.

If $d > r$, then Σ^T is an $r \times d$ matrix giving rise to the system $\Sigma^T x = b$ with fewer equations than unknowns. Thus, by well–known matrix properties, the system $\Sigma^T x = 0$ has a nontrivial solution: that is, a solution other than the zero vector. Therefore, $\text{null}(\Sigma^T) \neq \{0\}$ and so (2.4.8) does not hold in the case when $d > r$.

Remark 2.4.4. In Theorem 2.4.4 we have proved a result of the form

$$0 < C_1 \leq \limsup_{t\to\infty} \frac{\|X(t)\|}{\rho(t)} \leq C_2 < +\infty, \quad \text{a.s.} \tag{2.4.12}$$

where $\rho(t) \to \infty$ as $t \to \infty$. In an application to a system in economics or population biology, where each *component* of the process represents a quantity of interest, it is reasonable to ask what the size of the *largest component* of the system is, rather than focussing on the Euclidean norm, which may not be as scientifically relevant. Indeed, focussing on the size of the large deviations of the biggest component gives an idea of the most extreme behaviour of the system as a whole, and thereby helps in understanding 'worst case scenarios' for the system. Equation (2.4.12) enables us to prove that the largest component also has an essential growth rate ρ. This is a simple consequence of the fact that the max norm and Euclidean norm in $\mathbb{R}^d$ are equivalent, and related by $\frac{1}{\sqrt{d}}\|x\| \le \max_{1\le i\le d}|x_i| \le \|x\|$. Thus, combining this with (2.4.12), we can get a result of the form

$$0 < \frac{1}{\sqrt{d}}C_1 \le \limsup_{t\to\infty} \frac{\max_{1\le j\le d}|X_j(t)|}{\rho(t)} \le C_2 < +\infty, \quad \text{a.s.}$$

Remark 2.4.5. Returning to (2.4.11), note that $\Phi \in RV_\infty(\zeta+1)$ by (2.3.3) and $\Phi^{-1} \in RV_\infty(\frac{1}{\zeta+1})$ by Lemma 2.3.1. Now, using the fact that $\log\Phi^{-1}(\log t)/\log\log t \to \frac{1}{\zeta+1}$ as $t \to \infty$ by (2.3.1), we take logs in (2.4.11) to get the following *exact* rate of growth for $\zeta > -1$,

$$\limsup_{t\to\infty} \frac{\log\|X(t)\|}{\log\log t} = \frac{1}{\zeta+1}, \quad \text{a.s.}$$

In the case where $\zeta = -1$, although Theorem 2.4.4 does not apply, in many cases we can still get bounds on the asymptotic behaviour by making an appropriate transformation. Consider, for example, $\phi(x) = \log x/x$. Then $\phi \in RV_\infty(-1)$ and satisfies $x\phi(x) \to \infty$. It can easily be shown that $\Phi(x) = \frac{1}{2}(\log x)^2$ and $\Phi^{-1}(x) = e^{\sqrt{2x}}$. Then, following from the results of Theorems 2.4.1 and 2.4.2, we take logs and let $\varepsilon \to 0$ to get

$$\frac{K_1}{\sqrt{c_1}} \le \limsup_{t\to\infty} \frac{\log\|X(t)\|}{\sqrt{\log t}} \le \frac{K_2}{\sqrt{c_2}}, \quad \text{a.s.}$$

2.4.2 Remarks on restrictions on the hypotheses

The results of Theorems 2.4.1, 2.4.2 and in turn Theorem 2.4.4, can be established under the hypotheses (2.4.3) through to (2.4.8). However, it is reasonable to ask whether these hypotheses can be relaxed while still proving a result on large deviations. By considering some examples we demonstrate that, without further analysis, certain hypotheses cannot be easily relaxed while maintaining an asymptotic relation such as (2.4.11). In each of the following examples we assume that one of the key hypotheses is false, and from that one can show that the solution will not obey Theorem 2.4.1.

Take the simple one–dimensional analogue of (2.4.1),

$$dX(t) = f(X(t))dt + g(X(t))dB(t), \tag{2.4.13}$$

where $f : \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$. In Example 2.4.1 below we consider a situation where conditions (2.4.7) and (2.4.8) do not hold, and in Examples 2.4.2 and 2.4.3 we consider a situation where $x\phi(x) \to L < +\infty$. Although we can provide rigorous justifications for the following examples, we choose to omit the details.

Example 2.4.1. Let X be the unique adapted continuous solution satisfying (2.4.13). Let $f(x) = -\phi(x)$ where ϕ satisfies (2.4.3) and let g be a continuous positive function such that $g(x) \to 0$ as $x \to \infty$.

Then g does not satisfy the inequality (2.4.8) and moreover X does not satisfy (2.4.11) for $\phi \in RV_\infty(\zeta),\ \zeta > -1$.

Example 2.4.2. Let X be the unique adapted continuous solution satisfying (2.4.13) and assume that the conditions of Theorem 2.4.1 hold, except that $g(x) = 1$ and $f(x) = -\phi(x)$ where ϕ satisfies

$$x\phi(x) \to L \text{ as } x \to \infty, \quad \text{for } \ \frac{1}{2} < L < +\infty.$$

Then Theorem 2.4.1 *does not* hold, and moreover there exists a sufficiently small $\varepsilon \in (0,1)$ such that

$$\mathbb{P}\Big[\limsup_{t\to\infty} \frac{|X(t)|}{\Phi^{-1}\big(\frac{K_1^2(1-\varepsilon)}{2c_1}\log t\big)} \geq 1\Big] = 0.$$

Example 2.4.3. Let X be the unique adapted continuous solution satisfying (2.4.13) and assume that the conditions of Theorem 2.4.1 hold, except that $g(x) = 1$ and $f(x) = -\phi(x)$ where ϕ satisfies

$$\phi(x) = \frac{Lx}{1+x^2} \quad \text{for} \quad 0 < L < \frac{1}{2}.$$

Then Theorem 2.4.1 *does not* hold, and moreover there exists a sufficiently small $\varepsilon \in (0,1)$ such that

$$\mathbb{P}\Big[\limsup_{t\to\infty} \frac{|X(t)|}{\Phi^{-1}\big(\frac{K_1^2(1-\varepsilon)}{2c_1}\log t\big)} \geq 1\Big] = 0.$$

2.4.3 Asymptotically diagonal systems

We next consider a typical situation in which conditions of the form (2.4.4) and (2.4.5) hold. Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be given by $f(x) = -\varphi(x) + \psi(x)$ for $x \in \mathbb{R}^d$, where $\varphi, \psi : \mathbb{R}^d \to \mathbb{R}^d$. The function φ has the form $\varphi(x_1, x_2, \ldots, x_d) = \sum_{j=1}^d \varphi_j(x_j)\mathbf{e}_j$, where each $\varphi_j : \mathbb{R} \to \mathbb{R}$.

Suppose that $\phi : [0,\infty) \to (0,\infty)$ is such that

$$\phi \in RV_\infty(\zeta) \text{ is locally Lipschitz continuous and } \lim_{x\to\infty} x\phi(x) = \infty. \tag{2.4.14}$$

Moreover, the scalar function ϕ determines the asymptotics of f as follows:

$$\text{for every } j = 1, \ldots, d, \text{ there is } \alpha_j \in (0,\infty) \text{ s.t. } \liminf_{|x|\to\infty} \frac{\text{sgn}(x)\varphi_j(x)}{\phi(|x|)} = \alpha_j; \tag{2.4.15}$$

$$\text{for each } j = 1, \ldots, d \text{ there exists } \beta_j > 0 \text{ s.t. } \limsup_{|x|\to\infty} \frac{\text{sgn}(x)\varphi_j(x)}{\phi(|x|)} = \beta_j; \tag{2.4.16}$$

$$\lim_{\|x\|\to\infty} \frac{|\psi_j(x)|}{\phi(\|x\|)} = 0. \tag{2.4.17}$$

These conditions on ϕ, φ and ψ enable us to verify the conditions on f required to determine good upper and lower estimates on the rate of growth of the almost sure running maxima of the SDE.

Lemma 2.4.1. *Let* $f = -\varphi + \psi$*, and let* φ *and* ψ *obey* (2.4.15) *and* (2.4.17)*. If* ϕ *obeys* (2.4.14)*, then there exists* $\alpha^* > 0$ *such that*

$$\limsup_{\|x\|\to\infty} \frac{\langle x, f(x)\rangle}{\|x\|\phi(\|x\|)} \leq -\alpha^*. \tag{2.4.18}$$

Lemma 2.4.2. *Let* $f = -\varphi + \psi$*, and let* φ *and* ψ *obey* (2.4.16) *and* (2.4.17)*. If* ϕ *obeys* (2.4.14)*, then there exists* $\beta^* > 0$ *such that*

(i) If $\phi \in RV_\infty(\zeta)$*,* $\zeta > -1$*, then*

$$\limsup_{\|x\|\to\infty} \frac{|\langle x, f(x)\rangle|}{\|x\|\phi(\|x\|)} \leq \beta^*. \tag{2.4.19}$$

(ii) If $\phi \in RV_\infty(-1)$*, and there exists* ϕ_1 *with* $\phi_1(x)/\phi(x) \to 1$ *as* $x \to \infty$ *such that* $x \mapsto x\phi_1(x)$ *is non–decreasing, then* (2.4.19) *holds.*

The conditions (2.4.15) and (2.4.16) ensure that the mean–reverting part of f has strength of mean–reversion $\phi(|x|)$ in each component, while condition (2.4.17) means that the other terms are of a smaller order of magnitude for large $\|x\|$. In some sense, it means that the system is asymptotically diagonal for large $\|x\|$.

The condition (2.4.14) essentially restricts our attention to problems where the strength of mean–reversion $\phi(x)$ is no greater than $|x|^\gamma$ for any $\gamma > -1$. Condition (2.4.14) holds for many ϕ: $\phi_1(x) = (1+x)^\gamma \log^\beta(2+x)$; $\phi_2(x) = (1+x)^\gamma$; $\phi_3(x) = [\log\log(e^2+x)]^\beta$ satisfy (2.4.14) for instance, for any $\beta > 0$, $\gamma > -1$. If $\phi(x) = e^{\gamma|x|}$ for $\gamma > 0$, then (2.4.14) does not hold.

2.5 Extensions to equations with Markovian switching

In this section, we consider the asymptotic behaviour of a finite–dimensional autonomous SDE with Markovian switching. Let Y be a continuous–time Markov chain with state space $\mathbb{S}$, and let B be a standard r–dimensional Brownian motion independent of Y. We assume that the state space of the Markov chain is finite, say $\mathbb{S} = \{1, 2, \cdots, N\}$ and the Markov chain has generator $\Gamma = (\gamma_{ij})_{N\times N}$. As a standing hypothesis we assume in this chapter that the Markov chain is *irreducible*. Under this condition, the Markov chain has a unique stationary (probability) distribution $\pi = (\pi_1, \pi_2, \cdots, \pi_N) \in \mathbb{R}^{1\times N}$ which can be determined by solving the following linear equation

$$\pi\Gamma = 0 \quad \text{subject to} \sum_{j=1}^{N} \pi_j = 1 \text{ and } \pi_j > 0 \quad \forall j \in \mathbb{S}. \tag{2.5.1}$$

Let $f : \mathbb{R}^d \times \mathbb{S} \to \mathbb{R}^d$ and $g : \mathbb{R}^d \times \mathbb{S} \to \mathbb{R}^{d\times r}$ be continuous functions obeying local Lipschitz continuity conditions. Then for all $\|x\| \vee \|u\| \leq n$ and for all $y \in \mathbb{S}$,

$$\|f(x, y) - f(u, y)\| \vee \|g(x, y) - g(u, y)\| \leq K_n\|x - u\|, \tag{2.5.2}$$

for every $n \in \mathbb{N}$. Let $X(0) = x_0$ and consider the SDE with Markovian switching

$$dX(t) = f(X(t), Y(t))dt + g(X(t), Y(t))\, dB(t). \tag{2.5.3}$$

We make the standing assumption that f and g obey this continuity restriction, and that Y is an irreducible continuous–time Markov chain with finite state space $\mathbb{S}$. Under these conditions there exists a unique local solution of (2.5.3).

We write $f_i(x, y) = \langle f(x, y), \mathbf{e}_i\rangle$, $i = 1, \ldots, d$ and $g_{ij}(x, y)$ to be the (i, j)–th entry of the $d \times r$ matrix g with real–valued entries. The i^{th} component of (2.5.3) is

$$dX_i(t) = f_i(X(t), Y(t))dt + \sum_{j=1}^{r} g_{ij}(X(t), Y(t))\, dB_j(t). \tag{2.5.4}$$

Our hypotheses here are direct analogues of the non–switching hypotheses, (2.4.4) through to (2.4.8), with the inclusion of an extra switching parameter y, over which we take the supremum or infimum.

Once again we characterise the nonlinearity of the drift coefficient f via a scalar function ϕ which satisfies (2.4.3) and we suppose that ϕ and f possess the following properties also:

$$\text{there exists } c_1 > 0 \text{ such that } \limsup_{\|x\|\to\infty} \left\{ \sup_{y\in\mathbb{S}} \frac{|\langle x, f(x,y)\rangle|}{\|x\|\phi(\|x\|)} \right\} \le c_1, \tag{2.5.5}$$

$$\text{there exists } c_2 > 0 \text{ such that } \limsup_{\|x\|\to\infty} \left\{ \sup_{y\in\mathbb{S}} \frac{\langle x, f(x,y)\rangle}{\|x\|\phi(\|x\|)} \right\} \le -c_2. \tag{2.5.6}$$

As before we define the function Φ according to (2.4.6).

We suppose that the noise is *bounded* by imposing the following hypotheses:

$$\text{there exists } K_2 > 0 \text{ and } K_0 > 0 \text{ such that } \|g(x,y)\|_F \le K_2 \ \forall\ y \in \mathbb{S} \tag{2.5.7}$$

$$\text{where } K_0 \le \|g(0,y)\|_F \quad \forall\ y \in \mathbb{S},$$

$$\text{there exists } K_1 > 0 \text{ such that } \inf_{\substack{\|x\|\in\mathbb{R}^d/\{0\} \\ y\in\mathbb{S}}} \frac{\sum_{j=1}^r \left(\sum_{i=1}^d x_i g_{ij}(x,y)\right)^2}{\|x\|^2} \ge K_1^2. \tag{2.5.8}$$

Under the local Lipschitz continuity conditions on f and g, there exists a unique local solution of (2.5.3). However we can again show that in fact there exists a unique global solution to (2.5.3). Using (2.5.6), (2.5.7) and the fact that the state space of the Markov chain in finite, it can be shown analogously to the non–switching case that

$$\sup_{y\in\mathbb{S}} \Big\{ \sup_{x\in\mathbb{R}^d} \{\langle x, f(x,y)\rangle + \frac{1}{2}\|g(x,y)\|_F^2\}\Big\} < +\infty.$$

As a result of this global one–sided bound, by Theorem 3.18 in [62] there exists a unique global solution to equation (2.5.3).

We could now state *exact* analogues of Theorems 2.4.1, 2.4.2 and 2.4.3 in the case where the equation contains Markovian switching. However, to avoid repetition we choose not to state them. Nonetheless, in order to give the reader an idea of how such results would be proven we give the statement of the analogy to Theorem 2.4.4 and an extract of its proof.

In the special case where ϕ is a regularly varying function, we obtain the following result which characterises the essential almost sure rate of growth of the running maxima of $\|X\|$ in the case when the process experiences Markovian switching.

Theorem 2.5.1. *Suppose there exists a function* $\phi \in RV_\infty(\zeta)$ *satisfying* (2.4.3), *and that* ϕ *and* f *satisfy* (2.5.5) *and* (2.5.6), *and that* g *obeys* (2.5.7) *and* (2.5.8). *Then* X, *the unique adapted continuous solution satisfying* (2.5.3), *satisfies*

$$\left(\frac{K_1^2}{2c_1}\right)^{\frac{1}{\zeta+1}} \leq \limsup_{t\to\infty} \frac{\|X(t)\|}{\Phi^{-1}(\log t)} \leq \left(\frac{K_2^2}{2c_2}\right)^{\frac{1}{\zeta+1}} \quad a.s., \tag{2.5.9}$$

where Φ *is defined by* (2.4.6) *and* $\zeta > -1$.

2.6 Extensions to equations with unbounded noise

We can extend (2.4.1) and (2.5.3) to the case of *unbounded* noise by also characterising the degree of nonlinearity in g via a scalar function γ which obeys similar properties to those which ϕ obeys in (2.4.3). We would allow g to be unbounded by replacing conditions (2.4.7) and (2.4.8) with:

$$\text{there exists } K_2 > 0 \text{ such that } \limsup_{\|x\|\to\infty} \frac{\|g(x)\|_F}{\gamma(\|x\|)} \leq K_2, \text{ where } \|g(0)\|_F > 0, \tag{2.6.1}$$

$$\text{there exists } K_1 > 0 \text{ such that } \liminf_{\|x\|\to\infty} \frac{\sum_{j=1}^r \left(\sum_{i=1}^d x_i g_{ij}(x)\right)^2}{\|x\|^2\gamma^2(\|x\|)} \geq K_1^2. \tag{2.6.2}$$

Since we are still interested in recurrent processes, we would need the strength of the mean–reversion to be in some sense stronger than the noise intensity. For this purpose we would impose a condition of the form

$$\frac{x\phi(x)}{\gamma^2(x)} \to \infty \quad \text{as } x \to \infty. \tag{2.6.3}$$

In this case, we find that if the function Ψ is defined by

$$\Psi(x) := \int_1^x \frac{\phi(u)}{\gamma^2(u)}\, du, \tag{2.6.4}$$

then, roughly speaking, all of the main results in the bounded case can be generalised to cover the case of unbounded noise by using the auxiliary function Ψ in place of Φ. In particular, in the special case where the ratio ϕ/γ^2 is a regularly varying function, we would get the following analogue of Theorem 2.4.4 which characterises the essential rate of growth of the largest fluctuations of an SDE with unbounded diffusion coefficient.

Theorem 2.6.1. *Suppose there exists functions ϕ and γ obeying* (2.6.3) *and that the ratio $\phi/\gamma^2 \in RV_\infty(\zeta)$. Suppose further that ϕ and f satisfy* (2.4.4) *and* (2.4.5) *and that γ and g satisfy* (2.6.1) *and* (2.6.2)*. Then X, the unique adapted continuous solution satisfying* (2.4.1)*, satisfies*

$$\left(\frac{K_1^2}{2c_1}\right)^{\frac{1}{\zeta+1}} \leq \limsup_{t\to\infty} \frac{\|X(t)\|}{\Psi^{-1}(\log t)} \leq \left(\frac{K_2^2}{2c_2}\right)^{\frac{1}{\zeta+1}} \quad a.s.,$$

where Ψ obeys (2.6.4) *and $\zeta > -1$.*

The proof of this theorem is similar in spirit to the proof of Theorem 2.4.4 and for that reason is not stated.

2.7 Proofs of Results from Section 2.4

Proof of Theorem 2.4.1. Before we begin, note that we often use similar notation from proof to proof for the purpose of clarity and consistency. In some cases, notation actually carries over from one proof to another and this will be specified.

The first step of this proof is to apply a time–change and a transformation to (2.4.1) in order to obtain a 1–dimensional equation with a square root diffusion term. This will allow us to apply the stochastic comparison theorem (Theorem 2.2.1) and will ensure that the diffusion coefficient satisfies (2.2.2). Define

$$G(x) = \begin{cases} \frac{\sqrt{\sum_{j=1}^{r}(\sum_{i=1}^{d} x_i g_{ij}(x))^2}}{\|x\|} & x \neq 0 \\ K_2 \geq c \geq K_1 & x = 0. \end{cases} \tag{2.7.1}$$

Note that by (2.4.7),(2.4.8) and the Cauchy–Schwarz inequality,

$$\begin{aligned} K_1 \leq G(x) &= \frac{\sqrt{\sum_{j=1}^r (\sum_{i=1}^d x_i g_{ij}(x))^2}}{\|x\|} \leq \frac{\sqrt{\sum_{i=1}^d x_i^2 \sum_{j=1}^r \sum_{i=1}^d g_{ij}^2(x)}}{\|x\|} \\ &= \frac{\|x\|.\|g(x)\|_F}{\|x\|} = \|g(x)\|_F \leq K_2 \quad \text{for } x \neq 0. \end{aligned} \tag{2.7.2}$$

Also define θ by

$$\theta(t) = \int_0^t G^2(X(s))\, ds, \quad t \geq 0.$$

Then $\lim_{t\to\infty} \theta(t) = \infty$. Since $t \mapsto \theta(t)$ is increasing, we may define the stopping time τ by $\tau(t) = \inf\{s \geq 0 : \theta(s) > t\}$ so that $\tau(t) = \theta^{-1}(t)$. Define $\tilde{X}(t) = X(\tau(t))$ for $t \geq 0$ and define $\mathcal{G}(t) = \mathcal{F}(\tau(t))$ for all $t \geq 0$ (where $(\mathcal{F}(t))_{t\geq 0}$ is the original filtration). Then $\tilde{X}$ is $\mathcal{G}(t)$–adapted. Furthermore, applying this time change to (2.4.2) we have

$$\tilde{X}_i(t) = X_i(\tau(t)) = X_i(0) + \int_0^{\tau(t)} f_i(X(s))\, ds + M_i(t) \tag{2.7.3}$$

where

$$M_i(t) = \int_0^{\tau(t)} \sum_{j=1}^r g_{ij}(X(s))\, dB_j(s). \tag{2.7.4}$$

Note that $M = (M_1, M_2, \ldots, M_d)^T$ is a d–dimensional $\mathcal{G}(t)$–local martingale.

Now, to deal with the Riemann integral term in (2.7.3), we use Problem 3.4.5 from [46], which states that if N_i is a bounded measurable function and $[a, b] \subset [0, \infty)$ then $\int_a^b N_i(s)\, d\theta(s) = \int_{\theta(a)}^{\theta(b)} N_i(\tau(s))\, ds$. In this case we set

$$N_i(t) = f_i(X(t))/G^2(X(t))$$

and as $d\theta(t) = G^2(X(t))\, dt$, we obtain

$$\begin{aligned} \int_0^{\tau(t)} f_i(X(s))\, ds &= \int_0^{\tau(t)} N_i(s)\, d\theta(s) \\ &= \int_{\theta(0)}^{\theta(\tau(t))} N_i(\tau(s))\, ds = \int_0^t f_i(\tilde{X}(s))/G^2(\tilde{X}(s))\, ds. \end{aligned} \tag{2.7.5}$$

To deal with the martingale term in (2.7.3), we note that the cross variation of M is given by

$$\begin{aligned}\langle M_i, M_m\rangle(t) &= \int_0^{\tau(t)} \sum_{j=1}^{r} g_{ij}(X(s))g_{mj}(X(s))\,ds \\ &= \int_0^{t} \sum_{j=1}^{r} g_{ij}(\tilde{X}(s))g_{mj}(\tilde{X}(s))/G^2(\tilde{X}(s))\,ds,\end{aligned}$$

where we employ the method used to deduce (2.7.5) to obtain the last equality. Thus by Theorem 3.4.2 in [46], there is an extension $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ of $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a d–dimensional Brownian motion $\tilde{B} = \{(\tilde{B}_1(t), \tilde{B}_2(t), \ldots, \tilde{B}_d(t))^T; \tilde{\mathcal{G}}(t); 0 \leq t < +\infty\}$ such that

$$M_i(t) = \int_0^{t} \sum_{j=1}^{r} g_{ij}(\tilde{X}(s))/G(\tilde{X}(s))\,d\tilde{B}_j(s), \quad \widetilde{\mathbb{P}}\text{–a.s.} \tag{2.7.6}$$

The filtration $\tilde{\mathcal{G}}(t)$ in the extended space is such that $\tilde{X}$ is $\tilde{\mathcal{G}}(t)$–adapted.

For reasons of clarity and economy, from this point onward we do not specify the probability measure with respect to which such events are almost sure. Later in the proof we will reverse the time change in order to deal with the original process X. Although the time change is random, the fact that $K_1^2 t \leq \theta(t) \leq K_2^2 t,\ t \geq 0$ ensures that $\tilde{X}(t) = X(\theta^{-1}(t))$ captures the most important aspects of the growth of the running maxima of $\|X(t)\|$. Moreover, almost sure results about the growth rate of the fluctuations of $t \mapsto \|\tilde{X}(t)\|$ still correspond to almost sure results about the growth rate of the fluctuations of $t \mapsto \|X(t)\|$ because $(\tilde{\Omega}, \tilde{\mathcal{F}}(t), \tilde{\mathbb{P}}), (\tilde{\mathcal{G}}(t))_{t\geq 0}$ is an extension of $(\Omega, \mathcal{F}(t), \mathbb{P}), (\mathcal{F}(t))_{t\geq 0}$.

Thus by (2.7.5), (2.7.6) and (2.7.3) we get

$$d\tilde{X}_i(t) = \frac{f_i(\tilde{X}(t))}{G^2(\tilde{X}(t))}\,dt + \frac{1}{G(\tilde{X}(t))}\sum_{j=1}^{r} g_{ij}(\tilde{X}(t))\,d\tilde{B}_j(t).$$

Next, to simplify notation, define $a : \mathbb{R}^d \to \mathbb{R}^+$ by

$$a(x) = \sum_{j=1}^{r}\left(\sum_{i=1}^{d} x_i g_{ij}(x)\right)^2. \tag{2.7.7}$$

By (2.4.8), $a(x) > 0 \ \forall\ x \neq 0$. Define for $j = 1, \dots, r$ the functions $A_j : \mathbb{R}^d \to \mathbb{R}$ by

$$A_j(x) = \frac{1}{\sqrt{a(x)}} \sum_{i=1}^{d} x_i g_{ij}(x), \quad x \neq 0$$

and $A_j(x) = 1/\sqrt{r}$ for $x = 0$. Then

$$\frac{1}{G(x)} \sum_{i=1}^{d} x_i g_{ij}(x) = \|x\| A_j(x), \quad x \in \mathbb{R}^d, \tag{2.7.8}$$

$$\sum_{j=1}^{r} A_j^2(x) = 1, \quad x \in \mathbb{R}^d. \tag{2.7.9}$$

Now applying Itô's rule to $\tilde{Z}(t) := \|\tilde{X}(t)\|^2$ we get

$$d\tilde{Z}(t) = \left[\frac{2\langle \tilde{X}(t), f(\tilde{X}(t))\rangle + \|g(\tilde{X}(t))\|_F^2}{G^2(\tilde{X}(t))}\right] dt + 2\sum_{j=1}^{r} \left(\frac{1}{G(\tilde{X}(t))} \sum_{i=1}^{d} \tilde{X}_i(t) g_{ij}(\tilde{X}(t)) \right) d\tilde{B}_j(t)$$

so by (2.7.8) and $\|\tilde{X}(t)\| = \sqrt{\tilde{Z}(t)}$ we have

$$d\tilde{Z}(t) = \left[\frac{2\langle \tilde{X}(t), f(\tilde{X}(t))\rangle + \|g(\tilde{X}(t))\|_F^2}{G^2(\tilde{X}(t))}\right] dt + 2\sqrt{\tilde{Z}(t)} \sum_{j=1}^{r} A_j(\tilde{X}(t))\, d\tilde{B}_j(t). \tag{2.7.10}$$

Finally define

$$\tilde{W}(t) = \int_0^t \sum_{j=1}^{r} A_j(\tilde{X}(s))\, d\tilde{B}_j(s), \quad t \geq 0.$$

By (2.7.9) and e.g. [46, Theorem 3.3.16], $\tilde{W}$ is a standard 1–dimensional Brownian motion adapted to $\tilde{\mathcal{G}}(t)$ such that

$$d\tilde{Z}(t) = \left[\frac{2\langle \tilde{X}(t), f(\tilde{X}(t))\rangle + \|g(\tilde{X}(t))\|_F^2}{G^2(\tilde{X}(t))}\right] dt + 2\sqrt{\tilde{Z}(t)} d\tilde{W}(t). \tag{2.7.11}$$

The time–change and transformation is now complete. The next step is to derive a lower bound on the drift coefficient of (2.7.11) in order to create a lower comparison process. We can then apply the comparison principle.

For $y \in \mathbb{R}^d$, define the functions $D : \mathbb{R}^d \to \mathbb{R}$ and $\Delta_- : \mathbb{R}^+ \to \mathbb{R}$ by

$$D(y) = \frac{2\langle y, f(y)\rangle + \|g(y)\|_F^2}{G^2(y)} \quad \text{and} \quad \Delta_-(x) = \min_{\|y\|=x} D(y).$$

Then for $y \in \mathbb{R}^d$, $D(y) \geq \min_{\|u\|=\|y\|} D(u) = \Delta_-(\|y\|)$. Thus, a lower bound on Δ_- represents a lower bound on the drift coefficient of (2.7.11). Note that Δ_- is continuous on $(0, \infty)$ and is potentially discontinuous at zero. However, it can be defined at zero. We construct a locally Lipschitz continuous function $\phi_-^{(\varepsilon)} : \mathbb{R}^+ \to \mathbb{R}$ such that

$$\Delta_-(x) + \phi_-^{(\varepsilon)}(x) > 0, \; x \geq 0. \tag{2.7.12}$$

Then for $x \in \mathbb{R}^d$,

$$D(x) + \phi_-^{(\varepsilon)}(\|x\|) \geq \Delta_-(\|x\|) + \phi_-^{(\varepsilon)}(\|x\|) > 0 \tag{2.7.13}$$

and so from (2.7.11) we will have

$$\begin{aligned} d\tilde{Z}(t) &= \Big[-\phi_-^{(\varepsilon)}(\|\tilde{X}(t)\|) + \big\{D(\tilde{X}(t)) + \phi_-^{(\varepsilon)}(\|\tilde{X}(t)\|)\big\}\Big]\,dt + 2\sqrt{\tilde{Z}(t)}\,d\tilde{W}(t) \\ &= \Big[-\phi_-^{(\varepsilon)}\big(\sqrt{\tilde{Z}(t)}\big) + D_{1,\varepsilon}(t)\Big]\,dt + 2\sqrt{\tilde{Z}(t)}\,d\tilde{W}(t) \end{aligned} \tag{2.7.14}$$

where $D_{1,\varepsilon}(t) := D(\tilde{X}(t)) + \phi_-^{(\varepsilon)}(\|\tilde{X}(t)\|)$ is an adapted process such that $D_{1,\varepsilon}(t) > 0 \; \forall \; t > 0$ a.s. by (2.7.13). We construct, as our comparison process,

$$d\underline{Z}_\varepsilon(t) = -\phi_-^{(\varepsilon)}\big(\sqrt{\underline{Z}_\varepsilon(t)}\big)\,dt + 2\sqrt{\underline{Z}_\varepsilon(t)}\,d\tilde{W}(t), \quad t \geq 0 \tag{2.7.15}$$

where $0 \leq \underline{Z}_\varepsilon(0) \leq \|\tilde{X}(0)\|^2 = \tilde{Z}(0)$. We will later show, using stochastic comparison techniques, that $\tilde{Z}(t) \geq \underline{Z}_\varepsilon(t)$ for all $t \geq 0$ almost surely.

Now we return to the construction of the function $\phi_-^{(\varepsilon)}$, the mean–reverting drift coefficient of the lower comparison process (2.7.15). This will effectively act as a lower bound on the drift coefficient of (2.7.11). However, this construction is made more delicate by the fact that our hypothesis (2.4.4) is an asymptotic hypothesis rather than a global one. This means that although we have estimates on f for large values of $\|x\|$, we require extra estimates for small and moderate values of $\|x\|$.

First note that for $\|y\| \neq 0$ we have $\|g(y)\|_F^2/G^2(y) \geq 1$ by (2.7.2). Define the constant $K_3 = \min\{1, \|g(0)\|_F^2/G^2(0)\}$. Then $\|g(y)\|_F^2/G^2(y) \geq K_3$ for all $y \in \mathbb{R}^d$ and moreover K_3

is strictly positive since $G(0) > 0$ by definition and $\|g(0)\|_F > 0$ by (2.4.7).

For an estimate on f observe that since f is continuous, by the Cauchy–Schwarz inequality, $\lim_{\|x\|\to 0} |\langle x, f(x)\rangle| = 0$. Therefore, for every $\varepsilon \in (0, 1 \wedge \frac{1}{4}K_1^2 K_3)$ there exists $0 < X_2(\varepsilon) < 1$ such that $|\langle x, f(x)\rangle| \le \varepsilon$ for all $\|x\| \le X_2(\varepsilon)$. Let $y \in \mathbb{R}^d$ such that $\|y\| \le X_2(\varepsilon)$. Then $2\langle y, f(y)\rangle \ge -2\varepsilon$. Thus, using (2.7.2),

$$D(y) = \frac{2\langle y, f(y)\rangle}{G^2(y)} + \frac{\|g(y)\|_F^2}{G^2(y)} \ge \frac{-2\varepsilon}{K_1^2} + K_3 \ge \frac{1}{2}K_3 =: 2\phi_* > 0.$$

Hence for $x \le X_2(\varepsilon)$,

$$\Delta_-(x) = \min_{\|y\|=x} D(y) \ge \min_{\|y\|=x} 2\phi_* = 2\phi_* > 0.$$

So this gives us an estimate for Δ_- on an interval close to zero. We now look for an estimate on an interval away from zero. From condition (2.4.4) it follows that for every $\varepsilon > 0$ there exists $X_1(\varepsilon) > 1$ such that $|\langle x, f(x)\rangle| \le c_1(1+\varepsilon)\|x\|\phi(\|x\|)$ for $\|x\| > X_1(\varepsilon)$. Therefore,

$$\langle x, f(x)\rangle \ge -c_1(1+\varepsilon)\|x\|\phi(\|x\|) \quad \text{for } \|x\| > X_1(\varepsilon).$$

Let $y \in \mathbb{R}^d$ such that $\|y\| > X_1(\varepsilon)$. Then using (2.7.2),

$$D(y) = \frac{2\langle y, f(y)\rangle}{G^2(y)} + \frac{\|g(y)\|_F^2}{G^2(y)} \ge \frac{-2c_1(1+\varepsilon)}{K_1^2}\|y\|\phi(\|y\|) + 1.$$

Hence for $x > X_1(\varepsilon)$,

$$\Delta_-(x) = \min_{\|y\|=x} D(y) \ge \frac{-2c_1(1+\varepsilon)}{K_1^2} x\phi(x) + 1. \tag{2.7.16}$$

And so we have an estimate for Δ_- on an interval away from zero. We are now in a position to construct the drift function $\phi_-^{(\varepsilon)}$ for the comparison process (2.7.15). However, because $X_2(\varepsilon) < X_1(\varepsilon)$, we must carefully bridge the gap between the estimate close to zero ($x \le X_2(\varepsilon)$) and the estimate away from zero ($x \ge X_1(\varepsilon)$) while ensuring that $\phi_-^{(\varepsilon)}$ is continuous and that it is a uniform bound for $-\Delta_-$.

If there exists $X_3'(\varepsilon) \in (X_2(\varepsilon), X_1(\varepsilon))$ such that $-\Delta_-(X_3'(\varepsilon)) = -\phi_*$ then define $X_3(\varepsilon) = X_3'(\varepsilon)$. Otherwise, define $X_3(\varepsilon) = \frac{1}{2}(X_2(\varepsilon) + X_1(\varepsilon))$. Define

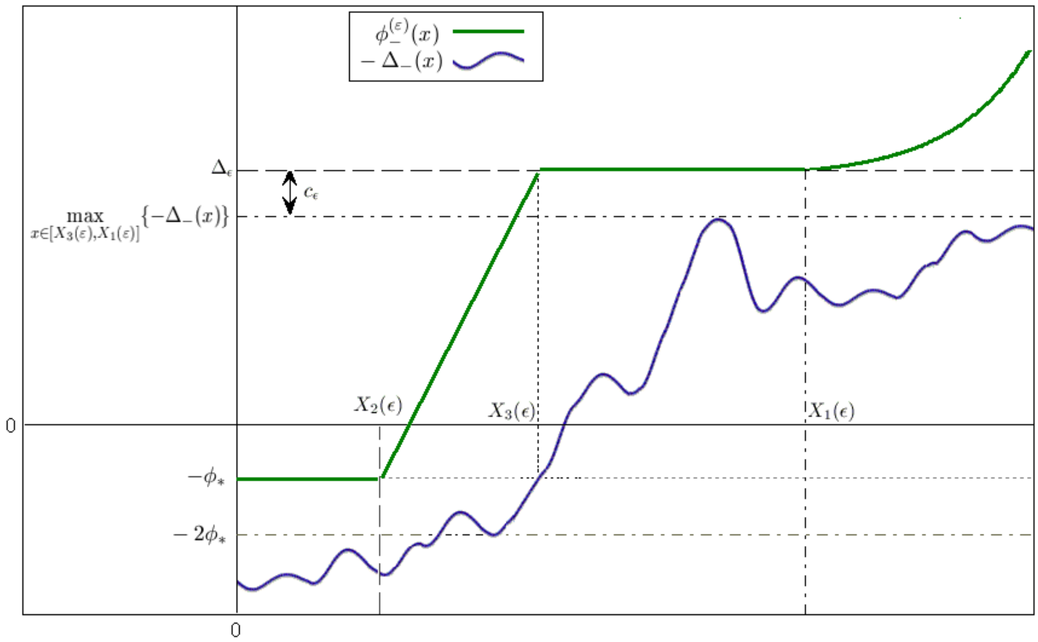
$$\phi_2(\varepsilon) = \frac{2c_1(1+\varepsilon)}{K_1^2} X_1(\varepsilon)\phi(X_1(\varepsilon)) - 1 - \Big[\max_{x\in[X_3,X_1]}\{-\Delta_-(x)\} \vee -\phi_*\Big]$$

and let $\alpha_\varepsilon = |\phi_2(\varepsilon)| + 1$, $c_\varepsilon = \alpha_\varepsilon + \phi_2(\varepsilon)$, and $\Delta_\varepsilon = c_\varepsilon + \left[\max_{x\in[X_3,X_1]}\{-\Delta_-(x)\} \vee -\phi_*\right]$. Note that $\alpha_\varepsilon \geq 1$, $c_\varepsilon \geq 1$ and $\Delta_\varepsilon > 0$. Finally, define

$$\phi_-^{(\varepsilon)}(x) = \begin{cases} -\phi_* & 0 \leq x \leq X_2(\varepsilon) \\ -\phi_* + \frac{\Delta_\varepsilon + \phi_*}{X_3(\varepsilon) - X_2(\varepsilon)}(x - X_2(\varepsilon)) & X_2(\varepsilon) < x \leq X_3(\varepsilon) \\ \Delta_\varepsilon & X_3(\varepsilon) < x \leq X_1(\varepsilon) \\ \alpha_\varepsilon - 1 + \frac{2c_1(1+\varepsilon)}{K_1^2} x\phi(x) & x > X_1(\varepsilon). \end{cases} \tag{2.7.17}$$

A visualisation of this drift function is given in Figure 2.1 below.

Figure 2.1: Bounding drift coefficient



Note that $\phi_-^{(\varepsilon)}$ is locally Lipschitz continuous on $[0,\infty)$ since it is locally Lipschitz continuous on each sub–interval. Now, it remains to check that $\phi_-^{(\varepsilon)}(x) + \Delta_-(x) > 0$ as required by condition (2.7.12). For $x \in [0, X_2(\varepsilon)]$, since $\Delta_-(x) \geq 2\phi_*$,

$$\phi_-^{(\varepsilon)}(x) + \Delta_-(x) = -\phi_* + \Delta_-(x) \geq -\phi_* + 2\phi_* = \phi_* > 0.$$

For $x \in (X_2(\varepsilon), X_3(\varepsilon)]$, since $\phi_-^{(\varepsilon)}$ is increasing on this interval and $\Delta_-(x) \geq \phi_*$,

$$\phi_-^{(\varepsilon)}(x) + \Delta_-(x) > \phi_-^{(\varepsilon)}(X_2(\varepsilon)) + \phi_* = -\phi_* + \phi_* = 0.$$

For $x \in (X_3(\varepsilon), X_1(\varepsilon)]$, since $\Delta_\varepsilon > -\Delta_-(x)$ by construction,

$$\phi_-^{(\varepsilon)}(x) + \Delta_-(x) = \Delta_\varepsilon + \Delta_-(x) > 0.$$

By (2.7.16), on the interval $(X_1(\varepsilon), \infty)$ we have

$$\phi_-^{(\varepsilon)}(x) + \Delta_-(x) = \alpha_\varepsilon - 1 + \frac{2c_1(1+\varepsilon)}{K_1^2} x\phi(x) + \Delta_-(x) \geq \alpha_\varepsilon > 0.$$

Hence, $\phi_-^{(\varepsilon)}(x) + \Delta_-(x) > 0 \ \forall\ x \geq 0$. Therefore, in summary, for every $\varepsilon \in (0, 1 \wedge \frac{1}{4}K_1^2 K_3)$ there exist $0 < X_2(\varepsilon) < X_1(\varepsilon), \alpha_\varepsilon \geq 1, \Delta_\varepsilon > 0$ and $X_3(\varepsilon) \in (X_2(\varepsilon), X_1(\varepsilon))$ such that the function $\phi_-^{(\varepsilon)}$ defined by (2.7.17) is locally Lipschitz continuous on $\mathbb{R}^+$ and obeys $\phi_-^{(\varepsilon)}(x) + \Delta_-(x) > 0$ for $x \geq 0$.

We construct the process $\underline{Z}_\varepsilon$ with $\underline{Z}_\varepsilon(0) \leq \tilde{Z}(0)$ and

$$d\underline{Z}_\varepsilon(t) = \left[-\phi_-^{(\varepsilon)}\left(\sqrt{|\underline{Z}_\varepsilon(t)|}\right)\right] dt + 2\sqrt{|\underline{Z}_\varepsilon(t)|} d\tilde{W}(t). \tag{2.7.18}$$

However, we must first show the non–negativity of this process so that we can drop the absolute values. Let $\tau_0 = \inf\{t \geq 0 : \underline{Z}_\varepsilon = 0\}$. Since $\underline{Z}_\varepsilon(0) \geq 0$, we have $\underline{Z}_\varepsilon(t) \geq 0$ for $t \in [0, \tau_0]$. Define now $\tau^{(1)} = \inf\{t \geq \tau_0 : |\underline{Z}_\varepsilon(t)| = (X_2(\varepsilon))^2\}$ and let Z_0 be defined by

$$Z_0(t \wedge \tau^{(1)}) = \int_{\tau_0}^{t\wedge\tau^{(1)}} 2\sqrt{|Z_0(s)|}\, d\tilde{W}(s), \quad t \geq \tau_0. \tag{2.7.19}$$

Then Z_0 has the unique solution $Z_0(t) = 0 \ \forall\ t \geq \tau_0$ a.s. Notice also that

$$\underline{Z}_\varepsilon(t \wedge \tau^{(1)}) = \underline{Z}_\varepsilon(\tau_0) + \int_{\tau_0}^{t\wedge\tau^{(1)}} -\phi_-^{(\varepsilon)}\left(\sqrt{|\underline{Z}_\varepsilon(s)|}\right) ds + \int_{\tau_0}^{t\wedge\tau^{(1)}} 2\sqrt{|\underline{Z}_\varepsilon(s)|} d\tilde{W}(s).$$

However, by construction, on the interval $\tau_0 \le t \le \tau^{(1)}$ we have $\sqrt{|\underline{Z}_\varepsilon(t)|} \le X_2(\varepsilon)$. Thus, $-\phi_-^{(\varepsilon)}(\sqrt{|\underline{Z}_\varepsilon(t)|}) = \phi_* > 0$ on the interval $\tau_0 \le t \le \tau^{(1)}$ and moreover $\underline{Z}_\varepsilon(\tau_0) = 0$ by the definition of τ_0. Therefore

$$\underline{Z}_\varepsilon(t \wedge \tau^{(1)}) = \int_{\tau_0}^{t\wedge\tau^{(1)}} \phi_*\, ds + \int_{\tau_0}^{t\wedge\tau^{(1)}} 2\sqrt{|\underline{Z}_\varepsilon(s)|}d\tilde{W}(s). \tag{2.7.20}$$

By applying a stochastic comparison principle to (2.7.19) and (2.7.20) we can conclude that $\underline{Z}_\varepsilon(t \wedge \tau^{(1)}) \ge Z_0(t\wedge\tau^{(1)}) = 0$ for $t \ge \tau_0$ a.s. Hence $\underline{Z}_\varepsilon(t) \ge 0$ a.s. for $t \in [\tau_0, \tau^{(1)})$. By iteration, define $\tau_1 = \inf\{t \ge \tau^{(1)} : \underline{Z}_\varepsilon(t) = 0\}$. Then $\underline{Z}_\varepsilon(t) \ge 0$ for $t \in [\tau^{(1)}, \tau_1]$. Define $\tau^{(2)} = \inf\{t \ge \tau_1 : |\underline{Z}_\varepsilon(t)| = (X_2(\varepsilon))^2\}$ and let Z_1 be defined by

$$Z_1(t\wedge\tau^{(2)}) = \int_{\tau_1}^{t\wedge\tau^{(2)}} 2\sqrt{|Z_1(s)|}d\tilde{W}(s), \quad t \ge \tau_1.$$

Then Z_1 has the unique solution $Z_1(t) = 0\ \forall\ t \ge \tau_1$ a.s. Continue as above to prove that $\underline{Z}_\varepsilon(t) \ge 0$ a.s. for $t \in [\tau_1, \tau^{(2)})$. By induction we can show that $\underline{Z}_\varepsilon(t) \ge 0$ for all $t \ge 0$, a.s. This proves the non–negativity of the lower bound and allows us to drop the absolute value signs in (2.7.18), so that $\underline{Z}_\varepsilon$ actually obeys (2.7.15).

We now apply Theorem 2.2.1 to (2.7.14) and (2.7.15). First note that the Brownian motion $\tilde{W}$ is $\tilde{\mathcal{G}}(t)$–adapted. If we set $b(x) = -\phi_-^{(\varepsilon)}(\sqrt{x})$ then condition (2.2.3) is satisfied since $\phi_-^{(\varepsilon)}$ is locally Lipschitz continuous and is constant on $[0, X_2(\varepsilon)]$. Also, set $\beta_1(t)$ equal to the drift coefficient of (2.7.14) and then $\beta_1(t) \ge b(\tilde{Z}(t))$ since $D_{1,\varepsilon}(t) > 0$ by (2.7.13), and moreover, $\beta_1(t)$ is $\tilde{\mathcal{G}}(t)$–adapted. Set $\beta_2(t)$ equal to the drift coefficient of (2.7.15) so that $b(\underline{Z}_\varepsilon(t)) = \beta_2(t)$, and again $\beta_2(t)$ is $\tilde{\mathcal{G}}(t)$–adapted. Next, condition (2.2.2) is trivially satisfied with $\sigma(x) = 2\sqrt{x}$. Finally, since we can (and will) prove independently that $\underline{Z}_\varepsilon$ is recurrent on $[0,\infty)$, it follows that $\tau_n^{(2)} = \inf\{t > 0 : \underline{Z}_\varepsilon(t) = n\} < +\infty$ a.s. Therefore we can apply Theorem 2.2.1 to conclude that for all $t \ge 0$, $\tilde{Z}(t) \ge \underline{Z}_\varepsilon(t)$ a.s.

Now we can approximate a lower bound on the asymptotic behaviour of $\tilde{Z}$ by getting a lower bound on the asymptotic behaviour of $\underline{Z}_\varepsilon$, satisfying

$$d\underline{Z}_\varepsilon(t) = \left[-\phi_-^{(\varepsilon)}\big(\sqrt{\underline{Z}_\varepsilon(t)}\big)\right] dt + 2\sqrt{\underline{Z}_\varepsilon(t)}d\tilde{W}(t).$$

However, in order to apply Theorem 1.0.2 to determine the asymptotic behaviour, we must first show that the conditions of Theorem 1.0.1 are satisfied.

The scale function, defined by (1.0.10), is given by

$$p_{\underline{Z}_\varepsilon}(x) = \int_{X_2^2(\varepsilon)}^x e^{-2\int_{X_2^2(\varepsilon)}^y \frac{-\phi_-^{(\varepsilon)}(\sqrt{z})}{4z}dz} dy \tag{2.7.21}$$

and observe that this can also be written

$$\begin{aligned} p_{\underline{Z}_\varepsilon}(x) &= \int_{X_2^2(\varepsilon)}^{X_1^2(\varepsilon)} e^{\frac{1}{2}\int_{X_2^2(\varepsilon)}^y \frac{\phi_-^{(\varepsilon)}(\sqrt{z})}{z}dz} dy + \int_{X_1^2(\varepsilon)}^x e^{\frac{1}{2}\int_{X_2^2(\varepsilon)}^{X_1^2(\varepsilon)} \frac{\phi_-^{(\varepsilon)}(\sqrt{z})}{z}dz} e^{\frac{1}{2}\int_{X_1^2(\varepsilon)}^y \frac{\phi_-^{(\varepsilon)}(\sqrt{z})}{z}dz} dy \\ &= A_1(\varepsilon) + A_2(\varepsilon)\int_{X_1^2(\varepsilon)}^x e^{\frac{1}{2}\int_{X_1^2(\varepsilon)}^y \frac{\phi_-^{(\varepsilon)}(\sqrt{z})}{z}dz} dy \end{aligned} \tag{2.7.22}$$

where $A_1(\varepsilon)$ and $A_2(\varepsilon)$ are positive bounded measurable functions.

Then on the interval $0 < x \le X_2^2(\varepsilon)$ where $\phi_-^{(\varepsilon)}(\sqrt{x}) = -\phi_* < 0$, (2.7.21) gives

$$p'_{\underline{Z}_\varepsilon}(x) = e^{-2\int_{X_2^2(\varepsilon)}^x \frac{-\phi_-^{(\varepsilon)}(\sqrt{u})}{4u}du} = e^{\frac{1}{2}\int_x^{X_2^2(\varepsilon)} \frac{\phi_*}{u}du} = \Big(\frac{x}{X_2^2(\varepsilon)}\Big)^{\frac{-\phi_*}{2}}. \tag{2.7.23}$$

Let $0 < \delta' < y \le X_2^2(\varepsilon)$. Then the speed measure is given by

$$m(\delta', y) = \int_{\delta'}^y \frac{2}{4z}\Big(\frac{z}{X_2^2(\varepsilon)}\Big)^{\frac{\phi_*}{2}} dz = \frac{1}{\phi_*}(X_2^2(\varepsilon))^{-\frac{\phi_*}{2}}\left[y^{\frac{\phi_*}{2}} - (\delta')^{\frac{\phi_*}{2}}\right].$$

So now, for $y \le X_2^2(\varepsilon)$,

$$m(0,y) = \lim_{\delta'\to 0^+} m(\delta', y) = \frac{1}{\phi_*}(X_2^2(\varepsilon))^{\frac{-\phi_*}{2}} y^{\frac{\phi_*}{2}} < +\infty$$

and $m(\{0\}) = \lim_{y\to 0^+} m(0,y) = 0$. On the interval $(X_2^2(\varepsilon), X_1^2(\varepsilon)]$, the speed measure is finite due to the continuity of $p'_{\underline{Z}_\varepsilon}$. On the interval $x > X_1^2(\varepsilon)$, (2.7.22) gives

$$p'_{\underline{Z}_\varepsilon}(x) = A_2(\varepsilon)e^{\frac{1}{2}\int_{X_1^2(\varepsilon)}^x \phi_-^{(\varepsilon)}(\sqrt{u}).\frac{1}{u}du}$$

which by substitution becomes

$$p'_{\underline{Z}_\varepsilon}(x) = A_2(\varepsilon)e^{\int_{X_1(\varepsilon)}^{\sqrt{x}} \frac{\phi_-^{(\varepsilon)}(v)}{v}\,dv}. \tag{2.7.24}$$

Since $\phi_-^{(\varepsilon)}(x)/x\phi(x) \to 2c_1(1+\varepsilon)/K_1^2$ as $x \to \infty$ and $x\phi(x) \to \infty$ as $x \to \infty$,

$$\log p'_{\underline{Z}_\varepsilon}(x) = \log A_2(\varepsilon) + \int_{X_1(\varepsilon)}^{\sqrt{x}} \frac{\phi_-^{(\varepsilon)}(v)}{v}\,dv = \log A_2(\varepsilon) + \int_{X_1(\varepsilon)}^{\sqrt{x}} \frac{\phi_-^{(\varepsilon)}(v)}{v\phi(v)}.v\phi(v).\frac{1}{v}\,dv$$

and thus $\log p'_{\underline{Z}_\varepsilon}(x)$ tends to infinity as $x \to \infty$. Moreover, by L'Hôpital's rule,

$$\begin{aligned}\lim_{x\to\infty} \frac{\log p'_{\underline{Z}_\varepsilon}(x)}{\log x} &= \lim_{x\to\infty} \frac{\int_{X_1(\varepsilon)}^{\sqrt{x}} \frac{\phi_-^{(\varepsilon)}(v)}{v}\,dv}{\log x} = \lim_{x\to\infty} \frac{\frac{1}{2}x^{-1}\phi_-^{(\varepsilon)}(\sqrt{x})}{x^{-1}} \\ &= \lim_{x\to\infty} \frac{c_1(1+\varepsilon)}{K_1^2}\sqrt{x}\phi(\sqrt{x}) + \frac{\alpha_\varepsilon - 1}{2} = \infty.\end{aligned}$$

Therefore there exists $x_* > X_1^2(\varepsilon)$ such that for $x > x_*$ we have $\log p'_{\underline{Z}_\varepsilon}(x)/\log x > 1$. Then for $x > x_*$ we have $p'_{\underline{Z}_\varepsilon}(x) > x$. Finally, looking at the speed measure we have

$$m(X_1^2(\varepsilon),\infty) = \frac{1}{2}\int_{X_1^2(\varepsilon)}^{\infty} \frac{1}{x}.\frac{1}{p'_{\underline{Z}_\varepsilon}(x)}\,dx < \frac{1}{2}\int_{X_1^2(\varepsilon)}^{x_*} \frac{1}{x}.\frac{1}{p'_{\underline{Z}_\varepsilon}(x)}\,dx + \frac{1}{2}\int_{x_*}^{\infty} \frac{1}{x^2}\,dx < +\infty.$$

Since the speed measure is finite on each interval it follows that $m[0,\infty) < +\infty$. Moreover, as $p'_{\underline{Z}_\varepsilon}(x) > x$ for all $x > x_*$, it follows that $p_{\underline{Z}_\varepsilon}(\infty-) = +\infty$. By (2.7.21), (2.7.23) and the fundamental theorem of calculus, for $x \le X_2^2(\varepsilon)$,

$$p_{\underline{Z}_\varepsilon}(x) = \int_{X_2^2(\varepsilon)}^{x} p'_{\underline{Z}_\varepsilon}(y)\,dy = -\int_{x}^{X_2^2(\varepsilon)} \Big(\frac{y}{X_2^2(\varepsilon)}\Big)^{-\frac{\phi_*}{2}}\,dy = \frac{(X_2^2(\varepsilon))^{\frac{\phi_*}{2}}}{1-\frac{\phi_*}{2}}x^{1-\frac{\phi_*}{2}} - \frac{X_2^2(\varepsilon)}{1-\frac{\phi_*}{2}}$$

and so $p_{\underline{Z}_\varepsilon}(0+) > -\infty$ since $1-\phi_*/2 > 0$. It can also easily be shown that the local integrability and non–degeneracy conditions of Theorem 1.0.1 are satisfied. Therefore, since $m(\{0\}) = 0, p_{\underline{Z}_\varepsilon}(0+) > -\infty$, $m[0,\infty) < +\infty$ and $p_{\underline{Z}_\varepsilon}(\infty-) = +\infty$, the process $\underline{Z}_\varepsilon$ is recurrent on $[0,\infty)$ with a reflecting boundary at zero. Thus, Motoo's theorem can be applied and for that we need to find a function $h_-^{(\varepsilon)}$ such that $\int_T^\infty p_{\underline{Z}_\varepsilon}^{-1}(h_-^{(\varepsilon)}(t))\,dt = \infty$. First we simplify our estimate on the scale function $p_{\underline{Z}_\varepsilon}$. Define $\Phi(x) = \int_1^x \phi(v)\,dv$ and note that Φ is increasing. Then by (2.7.22), for $x > X_1^2(\varepsilon)$,

$$p_{\underline{Z}_\varepsilon}(x) = A_1(\varepsilon) + A_2(\varepsilon)\int_{X_1^2(\varepsilon)}^{x} e^{\int_{X_1(\varepsilon)}^{\sqrt{y}} \frac{\alpha_\varepsilon - 1}{v} + \frac{2c_1(1+\varepsilon)}{K_1^2}\phi(v)\,dv} dy$$
$$= A_1(\varepsilon) + K_4(\varepsilon)\int_{X_1^2(\varepsilon)}^{x} y^{\frac{\alpha_\varepsilon-1}{2}} e^{\frac{2c_1(1+\varepsilon)}{K_1^2}\Phi(\sqrt{y})} dy,$$

where $K_4(\varepsilon) := A_2(\varepsilon)(X_1(\varepsilon))^{1-\alpha_\varepsilon} e^{-\frac{2c_1(1+\varepsilon)}{K_1^2}\Phi(X_1(\varepsilon))} > 0$. Then since Φ is increasing

$$p_{\underline{Z}_\varepsilon}(x) \le A_1(\varepsilon) + K_4(\varepsilon) e^{\frac{2c_1(1+\varepsilon)}{K_1^2}\Phi(\sqrt{x})}\int_{X_1^2(\varepsilon)}^{x} y^{\frac{\alpha_\varepsilon-1}{2}}\,dy$$
$$\le A_1(\varepsilon) + \frac{2K_4(\varepsilon)}{\alpha_\varepsilon+1}x^{\frac{\alpha_\varepsilon+1}{2}} e^{\frac{2c_1(1+\varepsilon)}{K_1^2}\Phi(\sqrt{x})} \le K_5(\varepsilon)x^{\frac{\alpha_\varepsilon+1}{2}} e^{\frac{2c_1(1+\varepsilon)}{K_1^2}\Phi(\sqrt{x})}, \qquad (2.7.25)$$

where $K_5(\varepsilon) := A_1(\varepsilon) + \frac{2K_4(\varepsilon)}{\alpha_\varepsilon+1} > 0$ and we have used the fact that $1 \le x^{\frac{\alpha_\varepsilon+1}{2}}$ and $1 \le \exp[\frac{2c_1(1+\varepsilon)}{K_1^2}\Phi(\sqrt{x})]$ in the last step. Since the exponential term is dominant in (2.7.25), we find it convenient to absorb the $K_5(\varepsilon)x^{\frac{\alpha_\varepsilon+1}{2}}$ term into the exponent for x large enough.

Since $\Phi(\sqrt{x}) \to \infty$ as $x \to \infty$ we have, by L'Hôpital's rule:

$$\lim_{x\to\infty}\frac{\log K_5(\varepsilon) + \frac{\alpha_\varepsilon+1}{2}\log x}{\frac{2c_1}{K_1^2}\Phi(\sqrt{x})} = \lim_{x\to\infty}\frac{\frac{\alpha_\varepsilon+1}{2x}}{\frac{c_1}{K_1^2\sqrt{x}}\phi(\sqrt{x})} = \lim_{x\to\infty}\frac{(\alpha_\varepsilon+1)K_1^2}{2c_1}\frac{1}{\sqrt{x}\phi(\sqrt{x})} = 0.$$

Therefore, for any $\varepsilon > 0$ there exists $X_4(\varepsilon) > 0$ such that for all $x \ge X_4(\varepsilon)$

$$\log K_5(\varepsilon) + \frac{\alpha_\varepsilon+1}{2}\log x \le \frac{2c_1\varepsilon}{K_1^2}\Phi(\sqrt{x}) \quad \text{and thus} \quad K_5(\varepsilon)x^{\frac{\alpha_\varepsilon+1}{2}} \le e^{\frac{2c_1\varepsilon}{K_1^2}\Phi(\sqrt{x})}.$$

Set $X_5(\varepsilon) = \max(X_1^2(\varepsilon), X_4(\varepsilon))$. Then for $x \ge X_5(\varepsilon)$,

$$p_{\underline{Z}_\varepsilon}(x) \le K_5(\varepsilon)x^{\frac{\alpha_\varepsilon+1}{2}} e^{\frac{2c_1(1+\varepsilon)}{K_1^2}\Phi(\sqrt{x})} \le e^{\frac{2c_1(1+2\varepsilon)}{K_1^2}\Phi(\sqrt{x})}. \qquad (2.7.26)$$

Let $T_1(\varepsilon) = e^{\frac{2c_1(1+2\varepsilon)}{K_1^2}\Phi(\sqrt{X_5(\varepsilon)})} > 1$. Then we can define

$$h_-^{(\varepsilon)}(t) = \left[\Phi^{-1}\left(\frac{K_1^2}{2c_1(1+2\varepsilon)}\log t\right)\right]^2, \quad \text{for } t \ge T_1(\varepsilon).$$

Moreover, since Φ^{-1} is increasing

$$\sqrt{h_-^{(\varepsilon)}(t)} = \Phi^{-1}\left(\frac{K_1^2}{2c_1(1+2\varepsilon)}\log t\right) \ge \Phi^{-1}\left(\frac{K_1^2}{2c_1(1+2\varepsilon)}\log T_1(\varepsilon)\right) = \sqrt{X_5(\varepsilon)}$$

and so $h_-^{(\varepsilon)}(t) \geq X_5(\varepsilon)$ and we can make use of (2.7.26). Hence, for $t \geq T_1(\varepsilon)$,

$$p_{\underline{Z}_\varepsilon}(h_-^{(\varepsilon)}(t)) \leq e^{\frac{2c_1(1+2\varepsilon)}{K_1^2}\Phi(\sqrt{h_-^{(\varepsilon)}(t)})} = t.$$

Thus for all $t \geq T_1(\varepsilon)$,

$$\int_{T_1(\varepsilon)}^{\infty} \frac{1}{p_{\underline{Z}_\varepsilon}(h_-^{(\varepsilon)}(t))}\,dt \geq \int_{T_1(\varepsilon)}^{\infty} \frac{1}{t}\,dt = \infty.$$

Therefore, by Theorem 1.0.2, there exists an a.s. event Ω_ε such that

$$\limsup_{t\to\infty} \frac{\underline{Z}_\varepsilon(t)}{h_-^{(\varepsilon)}(t)} \geq 1, \quad \text{a.s. on } \Omega_\varepsilon.$$

So, by the comparison principle

$$\limsup_{t\to\infty} \frac{\tilde{Z}(t)}{h_-^{(\varepsilon)}(t)} \geq \limsup_{t\to\infty} \frac{\underline{Z}_\varepsilon(t)}{h_-^{(\varepsilon)}(t)} \geq 1, \quad \text{a.s. on } \Omega_\varepsilon.$$

By taking square roots we get

$$\limsup_{t\to\infty} \frac{\|X(\tau(t))\|}{\Phi^{-1}\left(\frac{K_1^2}{2c_1(1+2\varepsilon)}\log t\right)} = \limsup_{t\to\infty} \frac{\sqrt{\tilde{Z}(t)}}{\sqrt{h_-^{(\varepsilon)}(t)}} \geq 1, \quad \text{a.s. on } \Omega_\varepsilon.$$

Recalling that $\tau(t) = \theta^{-1}(t)$ and that $\theta(t) \to \infty$ as $t \to \infty$, we let $T = \tau(t)$ to get

$$\limsup_{T\to\infty} \frac{\|X(T)\|}{\Phi^{-1}\left(\frac{K_1^2}{2c_1(1+2\varepsilon)}\log \theta(T)\right)} = \limsup_{t\to\infty} \frac{\|X(\tau(t))\|}{\Phi^{-1}\left(\frac{K_1^2}{2c_1(1+2\varepsilon)}\log t\right)} \geq 1, \text{ a.s. on } \Omega_\varepsilon.$$

Since $\theta(T) = \int_0^T G^2(X(s))\,ds$ and $K_1^2 \leq G^2(x) \leq K_2^2$ we have $\theta(T) \geq K_1^2 T$. Thus, for $T \geq \max(\theta^{-1}(T_1(\varepsilon)), 1/K_1^2)$,

$$\limsup_{T\to\infty} \frac{\|X(T)\|}{\Phi^{-1}\left(\frac{K_1^2}{2c_1(1+2\varepsilon)}\log K_1^2 T\right)} \geq \limsup_{T\to\infty} \frac{\|X(T)\|}{\Phi^{-1}\left(\frac{K_1^2}{2c_1(1+2\varepsilon)}\log \theta(T)\right)} \geq 1, \text{ a.s. on } \Omega_\varepsilon.$$

Note that for every $\varepsilon \in (0,1)$ there exists $T_2(\varepsilon) > 0$ such that

$$(1-\varepsilon)\log T \leq \log K_1^2 T \leq (1+\varepsilon)\log T \quad \text{for } T \geq T_2(\varepsilon).$$

Hence, for $T \geq \max(\theta^{-1}(T_1(\varepsilon)), T_2(\varepsilon), 1/K_1^2)$,

$$\limsup_{T\to\infty} \frac{\|X(T)\|}{\Phi^{-1}\left(\frac{K_1^2(1-\varepsilon)}{2c_1(1+2\varepsilon)}\log T\right)} \geq 1, \quad \text{a.s on } \Omega_\varepsilon.$$

Since the above holds for all ε sufficiently small, it also holds with ε replaced by $\varepsilon/(3-2\varepsilon)$. This proves the result. □

Proof of Theorem 2.4.2. As mentioned in the proof of Theorem 2.4.1, certain notation is common to both proofs while other notation is re–used for clarity and consistency. Assume the latter unless otherwise stated.

Now, following the same method as in the proof of the lower bound, we arrive at

$$d\tilde{Z}(t) = \left[\frac{2\langle \tilde{X}(t), f(\tilde{X}(t))\rangle + \|g(\tilde{X}(t))\|_F^2}{G^2(\tilde{X}(t))}\right] dt + 2\sqrt{\tilde{Z}(t)}d\tilde{W}(t) \tag{2.7.27}$$

where G is defined by (2.7.1) and satisfies (2.7.2). The next step is to derive an upper bound on the drift coefficient of (2.7.27) in order to create an upper comparison process. We can then apply the comparison principle.

For $y \in \mathbb{R}^d$, define the functions $D : \mathbb{R}^d \to \mathbb{R}$ and $\Delta_+ : \mathbb{R}^+ \to \mathbb{R}$ by

$$D(y) = \frac{2\langle y, f(y)\rangle + \|g(y)\|_F^2}{G^2(y)} \quad \text{and} \quad \Delta_+(x) = \max_{\|y\|=x} D(y).$$

Then for $y \in \mathbb{R}^d$, $D(y) \leq \max_{\|u\|=\|y\|} D(u) = \Delta_+(\|y\|)$. Thus, an upper bound on Δ_+ represents an upper bound on the drift coefficient of (2.7.27). Note that Δ_+ is continuous on $(0,\infty)$ and is potentially discontinuous at zero. However, it can be defined at zero. We construct a locally Lipschitz continuous function $\phi_+^{(\varepsilon)} : \mathbb{R}^+ \to \mathbb{R}$ such that

$$\Delta_+(x) + \phi_+^{(\varepsilon)}(x) < 0, \; x \geq 0. \tag{2.7.28}$$

Then for $x \in \mathbb{R}^d$,

$$D(x) + \phi_+^{(\varepsilon)}(\|x\|) \leq \Delta_+(\|x\|) + \phi_+^{(\varepsilon)}(\|x\|) < 0 \tag{2.7.29}$$

and so from (2.7.27) we will have

$$\begin{aligned} d\tilde{Z}(t) &= \left[-\phi_+^{(\varepsilon)}(\|\tilde{X}(t)\|) + \left\{D(\tilde{X}(t)) + \phi_+^{(\varepsilon)}(\|\tilde{X}(t)\|)\right\}\right] dt + 2\sqrt{\tilde{Z}(t)}d\tilde{W}(t) \\ &= \left[-\phi_+^{(\varepsilon)}\left(\sqrt{\tilde{Z}(t)}\right) + D_{2,\varepsilon}(t)\right] dt + 2\sqrt{\tilde{Z}(t)}\, d\tilde{W}(t) \end{aligned} \tag{2.7.30}$$

where $D_{2,\varepsilon}(t) := D(\tilde{X}(t)) + \phi_+^{(\varepsilon)}(\|\tilde{X}(t)\|)$ is an adapted process such that $D_{2,\varepsilon}(t) < 0 \ \forall\ t \geq 0$ a.s. by (2.7.29) We construct, as our comparison process,

$$d\bar{Z}_\varepsilon(t) = -\phi_+^{(\varepsilon)}\big(\sqrt{\bar{Z}_\varepsilon(t)}\big)\,dt + 2\sqrt{\bar{Z}_\varepsilon(t)}\,d\tilde{W}(t),\ t \geq 0 \tag{2.7.31}$$

where $\bar{Z}_\varepsilon(0) \geq \|\tilde{X}(0)\|^2 = \tilde{Z}(0)$. We will later show, using stochastic comparison techniques, that $\bar{Z}_\varepsilon(t) \geq \tilde{Z}(t)$ for all $t \geq 0$ almost surely.

Now we return to the construction of the function $\phi_+^{(\varepsilon)}$, the mean–reverting drift coefficient of the upper comparison process (2.7.31). The construction will be similar to that of Figure 2.1 and again it is made more delicate by the fact that our hypothesis (2.4.5) is an asymptotic hypothesis rather than a global one. Define the constant $K_4 = \max\{\|g(0)\|_F^2/G^2(0), K_2^2/K_1^2\}$. Since $K_2 \geq K_1$ by (2.7.2) it follows that $K_4 \geq 1$ and moreover $\|g(y)\|_F^2/G^2(y) \leq K_4$ for all $y \in \mathbb{R}^d$.

Since f is continuous, by the Cauchy–Schwarz inequality we have $\lim_{\|x\|\to 0} |\langle x, f(x)\rangle| = 0$. Therefore, for every $\varepsilon \in (0, 1 \wedge \frac{1}{2}K_1^2 K_4)$, there exists $0 < X_2(\varepsilon) < 1$ such that $|\langle x, f(x)\rangle| \leq \varepsilon$ for all $\|x\| \leq X_2(\varepsilon)$. Let $y \in \mathbb{R}^d$ such that $\|y\| \leq X_2(\varepsilon)$. Then $2\langle y, f(y)\rangle \leq 2\varepsilon$. Thus, using the above and the fact that $G^2(y) \geq K_1^2$ by (2.4.8), we have

$$D(y) = \frac{2\langle y, f(y)\rangle}{G^2(y)} + \frac{\|g(y)\|_F^2}{G^2(y)} \leq \frac{2\varepsilon}{K_1^2} + K_4 \leq 2K_4 =: \frac{1}{2}\phi_* > 0.$$

Hence, for $x \leq X_2(\varepsilon)$,

$$\Delta_+(x) = \max_{\|y\|=x} D(y) \leq \max_{\|y\|=x} \frac{1}{2}\phi_* = \frac{1}{2}\phi_*.$$

So this gives us an estimate for Δ_+ on an interval close to zero. We now look for an estimate on an interval away from zero. From condition (2.4.5) it follows that for every $\varepsilon > 0$, there exists $X_1(\varepsilon) > 1$ such that $\langle x, f(x)\rangle < -c_2(1-\varepsilon)\|x\|\phi(\|x\|)$ for $\|x\| > X_1(\varepsilon)$. Let $y \in \mathbb{R}^d$ such that $\|y\| > X_1(\varepsilon)$. Then using (2.7.2),

$$D(y) = \frac{2\langle y, f(y)\rangle}{G^2(y)} + \frac{\|g(y)\|_F^2}{G^2(y)} \leq \frac{-2c_2(1-\varepsilon)}{K_2^2}\|y\|\phi(\|y\|) + \frac{K_2^2}{K_1^2}.$$

Hence for $x > X_1(\varepsilon)$,

$$\Delta_+(x) = \max_{\|y\|=x} D(y) \le \frac{-2c_2(1-\varepsilon)}{K_2^2} x\phi(x) + \frac{K_2^2}{K_1^2}. \tag{2.7.32}$$

And so we have an estimate for Δ_+ on an interval away from zero. We are now in a position to construct the drift function $\phi_+^{(\varepsilon)}$ for the comparison process (2.7.31). Once again, because $X_2(\varepsilon) < X_1(\varepsilon)$ we must carefully bridge the gap between the estimate close to zero and the estimate away from zero while ensuring that $\phi_+^{(\varepsilon)}$ is continuous and that it is a uniform bound for $-\Delta_+$.

If there exists $X_3'(\varepsilon) \in (X_2(\varepsilon), X_1(\varepsilon))$ such that $-\Delta_+(X_3'(\varepsilon)) = -\phi_*$ then define $X_3(\varepsilon) = X_3'(\varepsilon)$. Otherwise, define $X_3(\varepsilon) = \frac{1}{2}(X_2(\varepsilon) + X_1(\varepsilon))$. Define

$$\phi_3(\varepsilon) = \frac{2c_2(1-\varepsilon)}{K_2^2} X_1(\varepsilon)\phi(X_1(\varepsilon)) - \frac{K_2^2}{K_1^2} - \left[\min_{x\in[X_3,X_1]}\{-\Delta_+(x)\} \wedge -\phi_*\right]$$

and let $c_\varepsilon = |\phi_3(\varepsilon)| + 1$, $\alpha_\varepsilon = c_\varepsilon + \phi_3(\varepsilon)$, and $\Delta_\varepsilon = -c_\varepsilon + \left[\min_{x\in[X_3,X_1]}\{-\Delta_+(x)\} \wedge -\phi_*\right]$. Note that $c_\varepsilon \ge 1$, $\alpha_\varepsilon \ge 1$ and $\Delta_\varepsilon < -\phi_*$. Finally, define

$$\phi_+^{(\varepsilon)}(x) = \begin{cases} -\phi_* & 0 \le x \le X_2(\varepsilon) \\ -\phi_* + \frac{\Delta_\varepsilon + \phi_*}{X_3(\varepsilon) - X_2(\varepsilon)}(x - X_2(\varepsilon)) & X_2(\varepsilon) < x \le X_3(\varepsilon) \\ \Delta_\varepsilon & X_3(\varepsilon) < x \le X_1(\varepsilon) \\ \frac{2c_2(1-\varepsilon)}{K_2^2} x\phi(x) - \frac{K_2^2}{K_1^2} - \alpha_\varepsilon & x > X_1(\varepsilon). \end{cases} \tag{2.7.33}$$

Note that $\phi_+^{(\varepsilon)}$ is locally Lipschitz continuous on $[0,\infty)$ since it is locally Lipschitz continuous on each subinterval. Now it remains to check that $\phi_+^{(\varepsilon)}(x) + \Delta_+(x) < 0$ as required by condition (2.7.28). For $x \in [0, X_2(\varepsilon)]$, since $\Delta_+(x) \le \phi_*/2$,

$$\phi_+^{(\varepsilon)}(x) + \Delta_+(x) = -\phi_* + \Delta_+(x) \le -\phi_* + \frac{1}{2}\phi_* = -\frac{1}{2}\phi_* < 0.$$

For $x \in (X_2(\varepsilon), X_3(\varepsilon)]$, since $\phi_+^{(\varepsilon)}$ is decreasing on this interval and $\Delta_+(x) \le \phi_*$,

$$\phi_+^{(\varepsilon)}(x) + \Delta_+(x) < \phi_+^{(\varepsilon)}(X_2(\varepsilon)) + \phi_* = -\phi_* + \phi_* = 0.$$

For $x \in (X_3(\varepsilon), X_1(\varepsilon)]$, since $\Delta_\varepsilon < -\Delta_+(x)$ by construction,

$$\phi_+^{(\varepsilon)}(x) + \Delta_+(x) = \Delta_\varepsilon + \Delta_+(x) < 0.$$

By (2.7.32), on the interval $(X_1(\varepsilon), \infty)$ we have

$$\phi_+^{(\varepsilon)}(x) + \Delta_+(x) = \frac{2c_2(1-\varepsilon)}{K_2^2} x\phi(x) - \frac{K_2^2}{K_1^2} - \alpha_\varepsilon + \Delta_+(x) \leq -\alpha_\varepsilon < 0.$$

Hence, $\phi_+^{(\varepsilon)}(x) + \Delta_+(x) < 0$ for all $x \geq 0$. Therefore, in summary, for every $\varepsilon \in (0, 1 \wedge \frac{1}{2}K_1^2 K_4)$ there exist $0 < X_2(\varepsilon) < X_1(\varepsilon)$, $\alpha_\varepsilon \geq 1$, $\Delta_\varepsilon < -\phi_*$ and $X_3(\varepsilon) \in (X_2(\varepsilon), X_1(\varepsilon))$ such that the function $\phi_+^{(\varepsilon)}$ defined by (2.7.33) is locally Lipschitz continuous on $\mathbb{R}^+$ and obeys $\phi_+^{(\varepsilon)}(x) + \Delta_+(x) < 0$ for $x \geq 0$.

We construct the process $\bar{Z}_\varepsilon$ with $\bar{Z}_\varepsilon(0) \geq \tilde{Z}(0) \geq 0$ and

$$d\bar{Z}_\varepsilon(t) = \left[-\phi_+^{(\varepsilon)}\left(\sqrt{|\bar{Z}_\varepsilon(t)|}\right)\right] dt + 2\sqrt{|\bar{Z}_\varepsilon(t)|}d\tilde{W}(t). \tag{2.7.34}$$

By analogy to the proof of the lower bound, we can prove the non–negativity of this process allowing us to drop the absolute value signs in (2.7.34), so that $\bar{Z}_\varepsilon$ actually obeys (2.7.31). The only difference here is that the process $\bar{Z}_\varepsilon$ will never hit zero since, as will soon be shown, it is recurrent on $(0, \infty)$.

We now apply Theorem 2.2.1 to (2.7.30) and (2.7.31). First note that the Brownian motion $\tilde{W}$ is $\tilde{\mathcal{G}}(t)$–adapted. If we set $b(x) = -\phi_+^{(\varepsilon)}(\sqrt{x})$ then condition (2.2.3) is satisfied since $\phi_+^{(\varepsilon)}$ is locally Lipschitz continuous and is constant on $[0, X_2(\varepsilon)]$. Also, set $\beta_1(t)$ equal to the drift coefficient of (2.7.31) and then $\beta_1(t) = b(\bar{Z}_\varepsilon(t))$, and moreover, $\beta_1(t)$ is $\tilde{\mathcal{G}}(t)$–adapted. Set $\beta_2(t)$ equal to the drift coefficient of (2.7.30) and so $b(\tilde{Z}(t)) \geq \beta_2(t)$ since $D_{2,\varepsilon}(t) < 0$ by (2.7.29), and again $\beta_2(t)$ is $\tilde{\mathcal{G}}(t)$–adapted. Next, condition (2.2.2) is trivially satisfied with $\sigma(x) = 2\sqrt{x}$. Finally, since we can (and will) prove independently that $\bar{Z}_\varepsilon$ is recurrent on $(0, \infty)$, it follows that $\tau_n^{(1)} = \inf\{t > 0 : \bar{Z}_\varepsilon(t) = n\} < +\infty$ a.s. Therefore we can apply Theorem 2.2.1 to conclude that for all $t \geq 0$, $\tilde{Z}(t) \leq \bar{Z}_\varepsilon(t)$ a.s.

Now we can approximate an upper bound on the asymptotic behavior of $\tilde{Z}$ by getting an upper bound on the asymptotic behaviour of $\bar{Z}_\varepsilon$, satisfying

$$d\bar{Z}_\varepsilon(t) = \left[-\phi_+^{(\varepsilon)}\big(\sqrt{\bar{Z}_\varepsilon(t)}\big)\right] dt + 2\sqrt{\bar{Z}_\varepsilon(t)}d\tilde{W}(t).$$

However, in order to apply Theorem 1.0.2 to determine the asymptotic behaviour, we must first show that the conditions of Theorem 1.0.1 are satisfied.

The scale function, defined by (1.0.10), is given by

$$p_{\bar{Z}_\varepsilon}(x) = \int_{X_2^2(\varepsilon)}^{x} e^{-2\int_{X_2^2(\varepsilon)}^{y} \frac{-\phi_+^{(\varepsilon)}(\sqrt{z})}{4z} dz} dy \tag{2.7.35}$$

and observe that this can also be written

$$\begin{aligned} p_{\bar{Z}_\varepsilon}(x) =& \int_{X_2^2(\varepsilon)}^{X_1^2(\varepsilon)} e^{\frac{1}{2}\int_{X_2^2(\varepsilon)}^{y} \frac{\phi_+^{(\varepsilon)}(\sqrt{z})}{z} dz} dy + \int_{X_1^2(\varepsilon)}^{x} e^{\frac{1}{2}\int_{X_2^2(\varepsilon)}^{X_1^2(\varepsilon)} \frac{\phi_+^{(\varepsilon)}(\sqrt{z})}{z} dz} e^{\frac{1}{2}\int_{X_1^2(\varepsilon)}^{y} \frac{\phi_+^{(\varepsilon)}(\sqrt{z})}{z} dz} dy \\ &= A_1(\varepsilon) + A_2(\varepsilon) \int_{X_1^2(\varepsilon)}^{x} e^{\frac{1}{2}\int_{X_1^2(\varepsilon)}^{y} \frac{\phi_+^{(\varepsilon)}(\sqrt{z})}{z} dz} dy \end{aligned} \tag{2.7.36}$$

where $A_1(\varepsilon)$ and $A_2(\varepsilon)$ are positive bounded measurable functions.

Then on the interval $0 < x \le X_2^2(\varepsilon)$ where $\phi_+^{(\varepsilon)}(\sqrt{x}) = -\phi_* < 0$, (2.7.35) gives

$$p'_{\bar{Z}_\varepsilon}(x) = e^{-2\int_{X_2^2(\varepsilon)}^{x} \frac{-\phi_+^{(\varepsilon)}(\sqrt{u})}{4u} du} = e^{\frac{1}{2}\int_{x}^{X_2^2(\varepsilon)} \frac{\phi_*}{u} du} = \left(\frac{x}{X_2^2(\varepsilon)}\right)^{\frac{-\phi_*}{2}}. \tag{2.7.37}$$

Let $0 < \delta' < y \le X_2^2(\varepsilon)$. Then the speed measure is given by

$$m(\delta', y) = \int_{\delta'}^{y} \frac{2}{4z}\left(\frac{z}{X_2^2(\varepsilon)}\right)^{\frac{\phi_*}{2}} dz = \frac{1}{\phi_*}(X_2^2(\varepsilon))^{\frac{-\phi_*}{2}} \left[y^{\frac{\phi_*}{2}} - (\delta')^{\frac{\phi_*}{2}}\right].$$

So now, for $y \le X_2^2(\varepsilon)$,

$$m(0, y) = \lim_{\delta' \to 0^+} m(\delta', y) = \frac{1}{\phi_*}(X_2^2(\varepsilon))^{\frac{-\phi_*}{2}} y^{\frac{\phi_*}{2}} < +\infty$$

On the interval $(X_2^2(\varepsilon), X_1^2(\varepsilon)]$, the speed measure is finite due to the continuity of $p'_{\bar{Z}_\varepsilon}$ and on the interval $x > X_1^2(\varepsilon)$, (2.7.36) gives

$$p'_{\bar{Z}_\varepsilon}(x) = A_2(\varepsilon) e^{\frac{1}{2}\int_{X_1^2(\varepsilon)}^{x} \phi_+^{(\varepsilon)}(\sqrt{u}).\frac{1}{u} du},$$

which by substitution becomes

$$p'_{\bar{Z}_\varepsilon}(x) = A_2(\varepsilon) e^{\int_{X_1(\varepsilon)}^{\sqrt{x}} \frac{\phi_+^{(\varepsilon)}(v)}{v}\, dv}. \tag{2.7.38}$$

Since $\phi_+^{(\varepsilon)}(x)/x\phi(x) \to 2c_2(1-\varepsilon)/K_2^2$ as $x \to \infty$ and $x\phi(x) \to \infty$ as $x \to \infty$,

$$\log p'_{\bar{Z}_\varepsilon}(x) = \log A_2(\varepsilon) + \int_{X_1(\varepsilon)}^{\sqrt{x}} \frac{\phi_+^{(\varepsilon)}(v)}{v}\, dv = \log A_2(\varepsilon) + \int_{X_1(\varepsilon)}^{\sqrt{x}} \frac{\phi_+^{(\varepsilon)}(v)}{v\phi(v)}.v\phi(v).\frac{1}{v}\, dv$$

and thus $\log p'_{\bar{Z}_\varepsilon}(x)$ tends to infinity as $x \to \infty$. Moreover by L'Hôpital's rule,

$$\begin{aligned}\lim_{x\to\infty} \frac{\log p'_{\bar{Z}_\varepsilon}(x)}{\log x} &= \lim_{x\to\infty} \frac{\int_{X_1(\varepsilon)}^{\sqrt{x}} \frac{\phi_+^{(\varepsilon)}(v)}{v}\, dv}{\log x} = \lim_{x\to\infty} \frac{\frac{1}{2}x^{-1}\phi_+^{(\varepsilon)}(\sqrt{x})}{x^{-1}} \\ &= \lim_{x\to\infty} \frac{c_2(1-\varepsilon)}{K_2^2}\sqrt{x}\phi(\sqrt{x}) - \frac{K_2^2}{2K_1^2} - \frac{\alpha_\varepsilon}{2} = \infty.\end{aligned}$$

Therefore there exists $x_* > X_1^2(\varepsilon)$ such that for $x > x_*$ we have $\log p'_{\bar{Z}_\varepsilon}(x)/\log x > 1$. Then for $x > x_*$ we have $p'_{\bar{Z}_\varepsilon}(x) > x$. Finally, looking at the speed measure we have

$$m(X_1^2(\varepsilon), \infty) = \frac{1}{2}\int_{X_1^2(\varepsilon)}^{\infty} \frac{1}{x}.\frac{1}{p'_{\bar{Z}_\varepsilon}(x)}\, dx < \frac{1}{2}\int_{X_1^2(\varepsilon)}^{x_*} \frac{1}{x}.\frac{1}{p'_{\bar{Z}_\varepsilon}(x)}\, dx + \frac{1}{2}\int_{x_*}^{\infty} \frac{1}{x^2}\, dx < +\infty.$$

Since the speed measure is finite on each interval it follows that $m(0,\infty) < +\infty$. Moreover, as $p'_{\bar{Z}_\varepsilon}(x) > x$ for all $x \geq x_*$, it follows that $p_{\bar{Z}_\varepsilon}(\infty-) = +\infty$. By (2.7.35), (2.7.37) and the fundamental theorem of calculus, for $x \leq X_2^2(\varepsilon)$,

$$p_{\bar{Z}_\varepsilon}(x) = \int_{X_2^2(\varepsilon)}^{x} p'_{\bar{Z}_\varepsilon}(y)\, dy = -\int_x^{X_2^2(\varepsilon)} \Big(\frac{y}{X_2^2(\varepsilon)}\Big)^{-\frac{\phi_*}{2}}\, dy = \frac{(X_2^2(\varepsilon))^{\frac{\phi_*}{2}}}{1-\frac{\phi_*}{2}} x^{1-\frac{\phi_*}{2}} - \frac{X_2^2(\varepsilon)}{1-\frac{\phi_*}{2}}$$

and so $p_{\bar{Z}_\varepsilon}(0+) = -\infty$ since $1 - \phi_*/2 < 0$. It can also easily be shown that the local integrability and non–degeneracy conditions of Theorem 1.0.1 are satisfied. Therefore, since $m(0,\infty) < +\infty$, $p_{\bar{Z}_\varepsilon}(0+) = -\infty$ and $p_{\bar{Z}_\varepsilon}(\infty-) = +\infty$, the process $\bar{Z}_\varepsilon$ is recurrent on $(0,\infty)$. Thus Motoo's theorem can be applied and for that we need to find a function $h_+^{(\varepsilon)}$ such that $\int_T^\infty p_{\bar{Z}_\varepsilon}^{-1}(h_+^{(\varepsilon)}(t))\, dt < \infty$. First we simplify our estimate on the scale function $p_{\bar{Z}_\varepsilon}$. Define $\Phi(x) = \int_1^x \phi(v)\, dv$ and note that Φ is increasing. Then by (2.7.36), for $x > X_1^2(\varepsilon)$,

$$\begin{aligned}p_{\bar{Z}_\varepsilon}(x) &= A_1(\varepsilon) + A_2(\varepsilon)\int_{X_1^2(\varepsilon)}^{x} e^{\int_{X_1(\varepsilon)}^{\sqrt{y}} \frac{2c_2(1-\varepsilon)}{K_2^2}\phi(v) - (\frac{K_2^2}{K_1^2}+\alpha_\varepsilon)\frac{1}{v}\, dv}\, dy \\ &\geq K_5(\varepsilon)\int_{X_1^2(\varepsilon)}^{x} y^{-(\frac{K_2^2}{2K_1^2}+\frac{\alpha_\varepsilon}{2})} e^{\frac{2c_2(1-\varepsilon)}{K_2^2}\Phi(\sqrt{y})}\, dy,\end{aligned} \tag{2.7.39}$$

where $K_5(\varepsilon) := A_2(\varepsilon)X_1(\varepsilon)^{(\frac{K_2^2}{K_1^2}+\alpha_\varepsilon)}e^{-\frac{2c_2(1-\varepsilon)}{K_2^2}\Phi(X_1(\varepsilon))} > 0$ and we have used the fact that $A_1(\varepsilon) \geq 0$ in the last step. Since the exponential term is dominant in (2.7.39), we once again find it convenient to absorb the other terms into the exponent for y large enough.

As $\Phi(\sqrt{y}) \to \infty$ as $y \to \infty$ we have, by L'Hopital's rule:

$$\lim_{y\to\infty} \frac{\log\frac{1}{K_5(\varepsilon)} + (\frac{K_2^2}{2K_1^2} + \frac{\alpha_\varepsilon}{2})\log y}{\frac{2c_2}{K_2^2}\Phi(\sqrt{y})} = \lim_{y\to\infty} \frac{\frac{1}{y}(\frac{K_2^2}{2K_1^2} + \frac{\alpha_\varepsilon}{2})}{\frac{c_2}{K_2^2\sqrt{y}}\phi(\sqrt{y})} = \lim_{y\to\infty} \frac{K_2^2(\frac{K_2^2}{K_1^2} + \alpha_\varepsilon)}{2c_2\sqrt{y}\phi(\sqrt{y})} = 0.$$

Therefore, for any $\varepsilon > 0$ there exists $X_4(\varepsilon) > X_1^2(\varepsilon)$ such that for $y \geq X_4(\varepsilon)$

$$\log\frac{1}{K_5(\varepsilon)} + (\frac{K_2^2}{2K_1^2} + \frac{\alpha_\varepsilon}{2})\log y \leq \frac{2c_2\varepsilon}{K_2^2}\Phi(\sqrt{y}) \text{ and thus } K_5(\varepsilon)y^{-(\frac{K_2^2}{2K_1^2}+\frac{\alpha_\varepsilon}{2})} \geq e^{-\frac{2c_2\varepsilon}{K_2^2}\Phi(\sqrt{y})}.$$

Then by (2.7.39) for $x \geq X_4(\varepsilon) + 1 > X_1^2(\varepsilon)$,

$$\begin{aligned} p_{\bar{Z}_\varepsilon}(x) &\geq K_5(\varepsilon)\int_{X_1^2(\varepsilon)}^{X_4(\varepsilon)} y^{-(\frac{K_2^2}{2K_1^2}+\frac{\alpha_\varepsilon}{2})} e^{\frac{2c_2(1-\varepsilon)}{K_2^2}\Phi(\sqrt{y})}\,dy + \int_{X_4(\varepsilon)}^{x} e^{\frac{2c_2(1-2\varepsilon)}{K_2^2}\Phi(\sqrt{y})}\,dy \\ &\geq \int_{x-1}^{x} e^{\frac{2c_2(1-2\varepsilon)}{K_2^2}\Phi(\sqrt{y})}\,dy \geq e^{\frac{2c_2(1-2\varepsilon)}{K_2^2}\Phi(\sqrt{x-1})}. \end{aligned} \tag{2.7.40}$$

Let $T_1(\varepsilon) = e^{\frac{2c_2(1-2\varepsilon)}{K_2^2\sqrt{1+\varepsilon}}\Phi(\sqrt{X_4(\varepsilon)})} > 1$. Then we can define

$$h_+^{(\varepsilon)}(t) = \left[\Phi^{-1}\Big(\frac{K_2^2\sqrt{1+\varepsilon}}{2c_2(1-2\varepsilon)}\log t\Big)\right]^2 + 1, \quad \text{for } t \geq T_1(\varepsilon).$$

Moreover, since Φ^{-1} is increasing

$$\sqrt{h_+^{(\varepsilon)}(t) - 1} = \Phi^{-1}\Big(\frac{K_2^2\sqrt{1+\varepsilon}}{2c_2(1-2\varepsilon)}\log t\Big) \geq \Phi^{-1}\Big(\frac{K_2^2\sqrt{1+\varepsilon}}{2c_2(1-2\varepsilon)}\log T_1(\varepsilon)\Big) = \sqrt{X_4(\varepsilon)}$$

and so $h_+^{(\varepsilon)}(t) \geq X_4(\varepsilon) + 1$ and we can make use of (2.7.40). Hence, for $t \geq T_1(\varepsilon)$,

$$p_{\bar{Z}_\varepsilon}(h_+^{(\varepsilon)}(t)) \geq e^{\frac{2c_2(1-2\varepsilon)}{K_2^2}\Phi(\sqrt{h_+^{(\varepsilon)}(t)-1})} = t^{\sqrt{1+\varepsilon}}.$$

Thus for all $t \geq T_1(\varepsilon)$,

$$\int_{T_1(\varepsilon)}^{\infty} \frac{1}{p_{\bar{Z}_\varepsilon}(h_+^{(\varepsilon)}(t))}\,dt \leq \int_{T_1(\varepsilon)}^{\infty} \frac{1}{t^{\sqrt{1+\varepsilon}}}\,dt < +\infty.$$

Therefore, by Theorem 1.0.2, there exists an a.s. event Ω_ε such that

$$\limsup_{t\to\infty} \frac{\bar{Z}_\varepsilon(t)}{h_+^{(\varepsilon)}(t)} \le 1 \quad \text{a.s. on } \Omega_\varepsilon.$$

So, by the comparison principle

$$\limsup_{t\to\infty} \frac{\tilde{Z}(t)}{h_+^{(\varepsilon)}(t)} \le \limsup_{t\to\infty} \frac{\bar{Z}_\varepsilon(t)}{h_+^{(\varepsilon)}(t)} \le 1 \quad \text{a.s. on } \Omega_\varepsilon.$$

By taking square roots, and noting that $h_+^{(\varepsilon)}(t) - 1$ behaves asymptotically like $h_+^{(\varepsilon)}(t)$,

$$\limsup_{t\to\infty} \frac{\|X(\tau(t))\|}{\Phi^{-1}\left(\frac{K_2^2\sqrt{1+\varepsilon}}{2c_2(1-2\varepsilon)}\log t\right)} = \limsup_{t\to\infty} \frac{\sqrt{\tilde{Z}(t)}}{\sqrt{h_+^{(\varepsilon)}(t)}} \le 1 \quad \text{a.s. on } \Omega_\varepsilon.$$

Recalling that $\tau(t) = \theta^{-1}(t)$ and that $\theta(t) \to \infty$ as $t \to \infty$, we let $T = \tau(t)$ to get

$$\limsup_{T\to\infty} \frac{\|X(T)\|}{\Phi^{-1}\left(\frac{K_2^2\sqrt{1+\varepsilon}}{2c_2(1-2\varepsilon)}\log\theta(T)\right)} = \limsup_{t\to\infty} \frac{\|X(\tau(t))\|}{\Phi^{-1}\left(\frac{K_2^2\sqrt{1+\varepsilon}}{2c_2(1-2\varepsilon)}\log t\right)} \le 1, \text{ a.s. on } \Omega_\varepsilon.$$

Since $\theta(T) = \int_0^T G^2(X(s))\,ds$ and $K_1^2 \le G^2(x) \le K_2^2$ we have $K_1^2 T \le \theta(T) \le K_2^2 T$. Thus, for $T \ge \max(\theta^{-1}(T_1(\varepsilon)), 1/K_1^2)$, since Φ^{-1} is increasing

$$\limsup_{T\to\infty} \frac{\|X(T)\|}{\Phi^{-1}\left(\frac{K_2^2\sqrt{1+\varepsilon}}{2c_2(1-2\varepsilon)}\log K_2^2 T\right)} \le \limsup_{T\to\infty} \frac{\|X(T)\|}{\Phi^{-1}\left(\frac{K_2^2\sqrt{1+\varepsilon}}{2c_2(1-2\varepsilon)}\log\theta(T)\right)} \le 1, \text{ a.s. on } \Omega_\varepsilon.$$

Note that for every $\varepsilon > 0$ there exists $T_2(\varepsilon) > 0$ such that

$$\sqrt{1-\varepsilon}\,\log T \le \log K_2^2 T \le \sqrt{1+\varepsilon}\,\log T \quad \text{for } T \ge T_2(\varepsilon).$$

Hence, for $T \ge \max(\theta^{-1}(T_1(\varepsilon)), T_2(\varepsilon), 1/K_1^2)$,

$$\limsup_{T\to\infty} \frac{\|X(T)\|}{\Phi^{-1}\left(\frac{K_2^2(1+\varepsilon)}{2c_2(1-2\varepsilon)}\log T\right)} \le 1, \quad \text{a.s. on } \Omega_\varepsilon.$$

Since the above holds for all ε sufficiently small, it also holds with ε replaced by $\varepsilon/(3+2\varepsilon)$. This proves the result. □

Proof of Theorem 2.4.3. We prove by contradiction. Let $y > 0$ and recall from Theorems 2.4.1 and 2.4.2 that $\|\tilde{X}(t)\|^2 \ge \underline{Z}_\varepsilon(t)$ for all $t \ge 0$ a.s. and $\|\tilde{X}(t)\|^2 \le \bar{Z}_\varepsilon(t)$ for all

$t \geq 0$ a.s. We aim to prove that there exists a time T_y such that $\|\tilde{X}(T_y)\| = y$ for any y. There are three cases to consider:

Case 1: Let $\|\tilde{X}(0)\|^2 > y^2$ and assume that there does not exists a $t_y > 0$ such that $\|\tilde{X}(t_y)\|^2 = y^2$. Then we must have $\|\tilde{X}(t)\|^2 > y^2$ for all $t \geq 0$. However, it was shown in the proof of Theorem 2.4.2 that $\bar{Z}_\varepsilon$ is recurrent on $(0, \infty)$ and so there exists $\bar{t}_y > 0$ such that $\bar{Z}_\varepsilon(\bar{t}_y) = y^2$. Then we have $y^2 = \bar{Z}_\varepsilon(\bar{t}_y) \geq \|\tilde{X}(\bar{t}_y)\|^2 > y^2$, a contradiction. Hence, there must exist a $t_y < \bar{t}_y$ such that $\|\tilde{X}(t_y)\|^2 = y^2$.

Case 2: Let $\|\tilde{X}(0)\|^2 < y^2$ and assume that there does not exist a $t_y > 0$ such that $\|\tilde{X}(t_y)\|^2 = y^2$. Then we must have $\|\tilde{X}(t_y)\|^2 < y^2$ for all $t \geq 0$. However, it was shown in the proof of Theorem 2.4.1 that $\underline{Z}_\varepsilon$ is recurrent on $[0, \infty)$ and so there exists $\underline{t}_y > 0$ such that $\underline{Z}_\varepsilon(\underline{t}_y) = y^2$. Then we have $y^2 = \underline{Z}_\varepsilon(\underline{t}_y) \leq \|\tilde{X}(\underline{t}_y)\|^2 < y^2$, a contradiction. Hence, there must exist a $t_y < \underline{t}_y$ such that $\|\tilde{X}(t_y)\|^2 = y^2$.

Case 3: Let $\|\tilde{X}(0)\|^2 = y^2$ and assume that there does not exist a $t_y > 0$ such that $\|\tilde{X}(t_y)\|^2 = y^2$. Then we must have either: a) $\|\tilde{X}(t)\|^2 < y^2$ for all $t > 0$, or b) $\|\tilde{X}(t)\|^2 > y^2$ for all $t > 0$.

Proceeding as in case 1 or 2 above, we can show that this is impossible and that there must exist a $t_y > 0$ such that $\|\tilde{X}(t_y)\|^2 = y^2$.

So in each case above we have shown that there exists a $t_y > 0$ such that $\|\tilde{X}(t_y)\|^2 = y^2$. Reversing the time change, we have that for all $y > 0$ there exists $t_y > 0$ such that $\|X(\tau(t_y))\|^2 = y^2$. Now let $T_y = \tau(t_y)$ and take square roots, so that for all $y > 0$ there exists $T_y > 0$ such that $\|X(T_y)\| = y$.

Moreover, since $\underline{Z}_\varepsilon(t) \leq \|\tilde{X}(t)\|^2$ and $\limsup_{t\to\infty} \underline{Z}_\varepsilon(t) = +\infty$ a.s., it follows that $\limsup_{t\to\infty} \|\tilde{X}(t)\|^2 = +\infty$ a.s. and thus $\limsup_{t\to\infty} \|X(t)\| = +\infty$ a.s. Similarly, since $\|\tilde{X}(t)\|^2 \leq \bar{Z}_\varepsilon(t)$ and $\liminf_{t\to\infty} \bar{Z}_\varepsilon(t) = 0$ a.s., it follows that $\liminf_{t\to\infty} \|\tilde{X}(t)\|^2 = 0$ a.s. and thus $\liminf_{t\to\infty} \|X(t)\| = 0$ a.s. This proves the recurrence of $\|X\|$ on $(0, \infty)$. □

Proof of Theorem 2.4.4. From Theorem 2.4.1 we have

$$\limsup_{t\to\infty} \frac{\|X(t)\|}{\Phi^{-1}\left(\frac{K_1^2(1-\varepsilon)}{2c_1}\log t\right)} \geq 1 \quad \text{a.s. on } \Omega_\varepsilon$$

and since $\phi \in RV_\infty(\zeta)$ it follows by (2.3.3) that $\Phi \in RV_\infty(\zeta+1)$ and, by Lemma 2.3.1, $\Phi^{-1} \in RV_\infty(\frac{1}{\zeta+1})$. Thus,

$$\lim_{t\to\infty} \frac{\Phi^{-1}\left(\frac{K_1^2(1-\varepsilon)}{2c_1}\log t\right)}{\Phi^{-1}(\log t)} = \lim_{x\to\infty} \frac{\Phi^{-1}\left(\frac{K_1^2(1-\varepsilon)}{2c_1}x\right)}{\Phi^{-1}(x)} = \left(\frac{K_1^2(1-\varepsilon)}{2c_1}\right)^{\frac{1}{\zeta+1}}.$$

Therefore,

$$\limsup_{t\to\infty} \frac{\|X(t)\|}{\Phi^{-1}(\log t)} \geq \left(\frac{K_1^2(1-\varepsilon)}{2c_1}\right)^{\frac{1}{\zeta+1}} \quad \text{a.s. on } \Omega_\varepsilon.$$

Using the fact that the intersection of almost sure events is itself almost sure, we can let $\varepsilon \to 0$ through the rational numbers to get

$$\limsup_{t\to\infty} \frac{\|X(t)\|}{\Phi^{-1}(\log t)} \geq \left(\frac{K_1^2}{2c_1}\right)^{\frac{1}{\zeta+1}} \quad \text{a.s.} \tag{2.7.41}$$

Similarly from Theorem 2.4.2 we get

$$\limsup_{t\to\infty} \frac{\|X(t)\|}{\Phi^{-1}(\log t)} \leq \left(\frac{K_2^2}{2c_2}\right)^{\frac{1}{\zeta+1}} \quad \text{a.s.} \tag{2.7.42}$$

Finally, combining (2.7.41) and (2.7.42) we get the desired result. □

2.7.1 Proofs of Results from Subsection 2.4.3

Before we prove Lemma 2.4.1 we first state and prove a useful auxiliary result.

Lemma 2.7.1. *Let $\phi \in RV_\infty(\zeta)$. Then there exists $C_+ > 0$ such that*

$$\liminf_{\|x\|\to\infty} \frac{\|x\|_\infty \phi(\|x\|_\infty)}{\|x\|\phi(\|x\|)} = C_+.$$

Proof of Lemma 2.7.1. There are two cases to consider: 1) $\phi \in RV_\infty(\zeta)$, $\zeta \neq 0$ and 2) $\phi \in RV_\infty(0)$.

Case 1: By norm equivalence in $\mathbb{R}^d$, there exists constants c_1 and c_2 such that $c_1\|x\| \leq$

$\|x\|_\infty \le c_2\|x\|$ for $x \in \mathbb{R}^d$. Since $\phi \in RV_\infty(\zeta)$ there exists, by Theorem 2.3.1, $\phi_1 \in SV_\infty(\zeta)$ such that $\phi_1(x)/\phi(x) \to 1$ as $x \to \infty$, $\phi_1 \in C^1(0,\infty)$ and $\lim_{x\to\infty} x\phi_1'(x)/\phi_1(x) = \zeta$, by (2.3.2). Since $\zeta \neq 0$, it follows that there exists x_1 such that for $x > x_1$, either

$$(i)\begin{cases}\phi_1'(x) > 0\\ \phi_1(x) > 0\end{cases} \text{if } \zeta > 0, \quad \text{or} \quad (ii)\begin{cases}\phi_1'(x) < 0\\ \phi_1(x) > 0\end{cases} \text{if } \zeta < 0.$$

We consider each of these cases in turn.

Case 1(i): Let $\|x\| > x_1/c_1$. Then $x_1 < c_1\|x\| \le \|x\|_\infty \le c_2\|x\|$ and so, since ϕ_1 is increasing, $\phi_1(c_1\|x\|) \le \phi_1(\|x\|_\infty) \le \phi_1(c_2\|x\|)$ for $\|x\| \ge x_1/c_1$. Thus,

$$\begin{aligned}\liminf_{\|x\|\to\infty} \frac{\|x\|_\infty \phi(\|x\|_\infty)}{\|x\|\phi(\|x\|)} &= \liminf_{\|x\|\to\infty} \frac{\|x\|_\infty}{\|x\|}\frac{\phi(\|x\|_\infty)}{\phi_1(\|x\|_\infty)}\frac{\phi_1(\|x\|_\infty)}{\phi_1(\|x\|)}\frac{\phi_1(\|x\|)}{\phi(\|x\|)}\\ &\ge c_1 \liminf_{\|x\|\to\infty}\frac{\phi_1(c_1\|x\|)}{\phi_1(\|x\|)} = c_1^{\zeta+1}\end{aligned}$$

since $\phi(x)/\phi_1(x) \to 1$ as $x \to \infty$ and $\phi_1 \in RV_\infty(\zeta)$.

Case 1(ii): Again we have that $x_1 < c_1\|x\| \le \|x\|_\infty \le c_2\|x\|$ for $\|x\| \ge x_1/c_1$. Then, since ϕ_1 is decreasing, $\phi_1(c_1\|x\|) \ge \phi_1(\|x\|_\infty) \ge \phi_1(c_2\|x\|)$ for $\|x\| \ge x_1/c_1$. Thus,

$$\begin{aligned}\liminf_{\|x\|\to\infty} \frac{\|x\|_\infty \phi(\|x\|_\infty)}{\|x\|\phi(\|x\|)} &= \liminf_{\|x\|\to\infty} \frac{\|x\|_\infty}{\|x\|}\frac{\phi(\|x\|_\infty)}{\phi_1(\|x\|_\infty)}\frac{\phi_1(\|x\|_\infty)}{\phi_1(\|x\|)}\frac{\phi_1(\|x\|)}{\phi(\|x\|)}\\ &\ge c_1 \liminf_{\|x\|\to\infty}\frac{\phi_1(c_2\|x\|)}{\phi_1(\|x\|)} = c_1 c_2^{\zeta}\end{aligned}$$

since $\phi(x)/\phi_1(x) \to 1$ as $x \to \infty$ and $\phi_1 \in RV_\infty(\zeta)$.

Case 2: Since $\phi \in RV_\infty(0)$ it follows that $x\phi(x) \in RV_\infty(1)$. Therefore, there exists $\phi_2 \in SV_\infty(1)$ such that $x\phi(x)/\phi_2(x) \to 1$ as $x \to \infty$, $\phi_2 \in C^1(0,\infty)$ and $x\phi_2'(x)/\phi_2(x) \to 1$ as $x \to \infty$. Therefore $\phi_2'(x) > 0$ for all $x > x_1$ since $\phi_2(x) > 0$. Hence, ϕ_2 is increasing on $[x_1,\infty)$. Let $\|x\| > x_1/c_1$. Then $x_1 < c_1\|x\| \le \|x\|_\infty \le c_2\|x\|$ and so, since ϕ_2 is increasing, $\phi_2(c_1\|x\|) \le \phi_2(\|x\|_\infty) \le \phi_2(c_2\|x\|)$ for $\|x\| \ge x_1/c_1$. Thus,

$$\begin{aligned}\liminf_{\|x\|\to\infty} \frac{\|x\|_\infty \phi(\|x\|_\infty)}{\|x\|\phi(\|x\|)} &= \liminf_{\|x\|\to\infty} \frac{\|x\|_\infty \phi(\|x\|_\infty)}{\phi_2(\|x\|_\infty)}\frac{\phi_2(\|x\|_\infty)}{\phi_2(\|x\|)}\frac{\phi_2(\|x\|)}{\|x\|\phi(\|x\|)}\\ &\ge \liminf_{\|x\|\to\infty}\frac{\phi_2(c_1\|x\|)}{\phi_2(\|x\|)} = c_1\end{aligned}$$

since $x\phi(x)/\phi_2(x) \to 1$ as $x \to \infty$ and $\phi_2 \in RV_\infty(1)$.

Therefore, in each case one can find a C_+ such that Lemma 2.7.1 holds. □

Proof of Lemma 2.4.1. Observe first that by the definition of φ and ψ we have that $\langle x, f(x)\rangle = -\sum_{j=1}^d x_j\varphi_j(x_j) + \sum_{j=1}^d x_j\psi_j(x)$. Due to (2.4.15), for every $\varepsilon \in (0,1)$ there exists $x_j^*(\varepsilon) > 0$ such that

$$\frac{\operatorname{sgn}(x)\varphi_j(x)}{\phi(|x|)} \geq \alpha_j(1-\varepsilon), \quad |x| \geq x_j^*(\varepsilon).$$

Since $x = |x|\operatorname{sgn}(x)$, we get

$$x\varphi_j(x) \geq \alpha_j(1-\varepsilon)|x|\phi(|x|) \text{ for all } |x| \geq x_j^*(\varepsilon).$$

Moreover, if we define $D_j(\varepsilon) := \max_{|x|<x_j^*(\varepsilon)} \left|\alpha_j(1-\varepsilon)|x|\phi(|x|) - x\varphi_j(x)\right|$ then

$$x\varphi_j(x) \geq -D_j(\varepsilon) + \alpha_j(1-\varepsilon)|x|\phi(|x|) \quad \forall\, x \in \mathbb{R}. \tag{2.7.43}$$

Next, by virtue of (2.4.17), for every $\varepsilon \in (0,1)$ there is an $x_j(\varepsilon) > 0$ such that for all $\|x\| \geq x_j(\varepsilon)$ we have $|\psi_j(x)| \leq \varepsilon\phi(\|x\|)$. Let $X(\varepsilon) = \max_{j=1,\dots,d} x_j(\varepsilon)$. Then for $\|x\| \geq X(\varepsilon)$ we have

$$|\psi_j(x)| \leq \varepsilon\phi(\|x\|) \text{ for } j = 1,\dots,d. \tag{2.7.44}$$

Define $D(\varepsilon) = \sum_{j=1}^d D_j(\varepsilon)$ and $\alpha' = \min_{j=1,\dots,d}\alpha_j > 0$. Thus by (2.7.43) and (2.7.44) we have, for $\|x\| \geq X(\varepsilon)$,

$$\begin{aligned}
\langle x, f(x)\rangle &= -\sum_{j=1}^d x_j\varphi_j(x_j) + \sum_{j=1}^d x_j\psi_j(x) \\
&\leq \sum_{j=1}^d D_j(\varepsilon) - \sum_{j=1}^d \alpha_j(1-\varepsilon)|x_j|\phi(|x_j|) + \sum_{j=1}^d |x_j||\psi_j(x)| \\
&\leq D(\varepsilon) - \alpha'(1-\varepsilon)\sum_{j=1}^d |x_j|\phi(|x_j|) + \sum_{j=1}^d |x_j|\{\varepsilon\phi(\|x\|)\} \\
&= D(\varepsilon) - \alpha'(1-\varepsilon)\sum_{j=1}^d |x_j|\phi(|x_j|) + \varepsilon\phi(\|x\|)\|x\|_1.
\end{aligned}$$

Note that $\sum_{j=1}^d |x_j|\phi(|x_j|) \geq \|x\|_\infty \phi(\|x\|_\infty)$ for $x = (x_1, x_2, \ldots, x_d)^T$ and, by norm equivalence, $\|x\|_1 \leq c\|x\|$. Therefore, for $\|x\| \geq X(\varepsilon)$,

$$\langle x, f(x)\rangle \leq D(\varepsilon) - \alpha'(1-\varepsilon)\|x\|_\infty \phi(\|x\|_\infty) + c\varepsilon\|x\|\phi(\|x\|).$$

Finally, by Lemma 2.7.1 and the fact that $D(\varepsilon)$ is constant,

$$\begin{aligned}
\limsup_{\|x\|\to\infty} \frac{\langle x, f(x)\rangle}{\|x\|\phi(\|x\|)} &\leq \limsup_{\|x\|\to\infty} \frac{D(\varepsilon)}{\|x\|\phi(\|x\|)} + \limsup_{\|x\|\to\infty} \frac{-\alpha'(1-\varepsilon)\|x\|_\infty\phi(\|x\|_\infty)}{\|x\|\phi(\|x\|)} + c\varepsilon \\
&= -\alpha'(1-\varepsilon) \liminf_{\|x\|\to\infty} \frac{\|x\|_\infty\phi(\|x\|_\infty)}{\|x\|\phi(\|x\|)} + c\varepsilon \\
&\leq -\alpha'(1-\varepsilon)C_+ + c\varepsilon
\end{aligned}$$

Since this estimate holds for every $\varepsilon \in (0,1)$, by letting $\varepsilon \to 0$, we get

$$\limsup_{\|x\|\to\infty} \frac{\langle x, f(x)\rangle}{\|x\|\phi(\|x\|)} \leq -\alpha' C_+ =: -\alpha^*,$$

as required in (2.4.18). □

Proof of Lemma 2.4.2. First of all, if $\zeta > -1$ then it is trivial to find a function ϕ_1 with $\phi_1(x)/\phi(x) \to 1$ as $x \to \infty$ such that $x \mapsto x\phi_1(x)$ is non–decreasing, and this will be shown at the end of the proof. If $\phi \in RV_\infty(-1)$ then this may not always hold but can be verified directly in some cases. We begin by proving that (2.4.19) can be established assuming the existence of a ϕ_1 as specified above.

If $\phi(x)/\phi_1(x) \to 1$ as $x \to \infty$, we have, by (2.4.16),

$$\limsup_{|x|\to\infty} \frac{\text{sgn}(x)\varphi_j(x)}{\phi_1(|x|)} = \beta_j$$

Therefore for every $\varepsilon \in (0,1)$ there exists $x_j^*(\varepsilon) > 0$ such that $x\varphi_j(x) < \beta_j(1+\varepsilon)|x|\phi_1(|x|)$ for all $x > x_j^*(\varepsilon)$. Also there is $x_j^{**}(\varepsilon) > 0$ such that

$$|x|\varphi_j(x) = x\,\text{sgn}(x)\varphi_j(x) > \beta_j(1+\varepsilon)x\phi_1(|x|) = -\beta_j(1+\varepsilon)|x|\phi_1(|x|), \quad x < -x_j^{**}(\varepsilon).$$

Hence, $x < -x_j^{**}(\varepsilon)$ implies $|x\varphi_j(x)| < \beta_j(1+\varepsilon)|x|\phi_1(|x|)$. Let $x_j(\varepsilon) = (x_j^*(\varepsilon) \vee x_j^{**}(\varepsilon))$. Then

$$|x\varphi_j(x)| < \beta_j(1+\varepsilon)|x|\phi_1(|x|), \quad |x| > x_j(\varepsilon).$$

Since φ is continuous, there exist $\tilde{\Phi}_j(\varepsilon)$ such that

$$|x\varphi_j(x)| \leq \tilde{\Phi}_j(\varepsilon), \quad |x| \leq x_j(\varepsilon).$$

Therefore for every $\varepsilon \in (0,1)$ we have

$$|x\varphi_j(x)| < \tilde{\Phi}_j(\varepsilon) + \beta_j(1+\varepsilon)|x|\phi_1(|x|), \quad x \in \mathbb{R}.$$

Moreover because $\phi(x)/\phi_1(x) \to 1$ as $x \to \infty$, by (2.4.17) we have $\|\psi(x)\|_1/\phi_1(\|x\|) \to 0$ as $\|x\| \to \infty$. Therefore, for every $\varepsilon \in (0,1)$, there exists $x(\varepsilon) > 0$ such that $\|\psi(x)\|_1 \leq \varepsilon\phi_1(\|x\|)$ for all $\|x\| > x(\varepsilon)$. Now, as we had in the previous proof,

$$\langle x, f(x)\rangle = -\sum_{j=1}^{d} x_j\varphi_j(x_j) + \sum_{j=1}^{d} x_j\psi_j(x).$$

Therefore for all $x := (x_1, x_2, \ldots, x_d)^T \in \mathbb{R}^d$ such that $\|x\| > x(\varepsilon)$, and $\beta' = \max_{j=1,\ldots,d}\beta_j$, by using the monotonicity of $x \mapsto x\phi_1(x)$ on $\mathbb{R}^+$, we get

$$\begin{aligned}
|\langle x, f(x)\rangle| &\leq \sum_{j=1}^{d} |x_j\varphi_j(x_j)| + \sum_{j=1}^{d} |x_j||\psi_j(x)| \\
&\leq \sum_{j=1}^{d} \tilde{\Phi}_j(\varepsilon) + \sum_{j=1}^{d} \beta_j(1+\varepsilon)|x_j|\phi_1(|x_j|) + \sum_{j=1}^{d} |x_j||\psi_j(x)| \\
&\leq \sum_{j=1}^{d} \tilde{\Phi}_j(\varepsilon) + \beta'(1+\varepsilon)\sum_{j=1}^{d} |x_j|\phi_1(|x_j|) + \|x\|\|\psi(x)\|_1 \\
&\leq \sum_{j=1}^{d} \tilde{\Phi}_j(\varepsilon) + \beta'(1+\varepsilon)\sum_{j=1}^{d} \|x\|\phi_1(\|x\|) + \varepsilon\|x\|\phi_1(\|x\|).
\end{aligned}$$

Therefore, as $x\phi_1(x) \to \infty$ as $x \to \infty$, and $\phi_1(x)/\phi(x) \to 1$ as $x \to \infty$ we have

$$\limsup_{\|x\|\to\infty} \frac{|\langle x, f(x)\rangle|}{\|x\|\phi(\|x\|)} \leq \beta'(1+\varepsilon)d + \varepsilon.$$

Letting $\varepsilon \to 0^+$, we get

$$\limsup_{\|x\|\to\infty} \frac{|\langle x, f(x)\rangle|}{\|x\|\phi(\|x\|)} \le d\beta' =: \beta^*.$$

So we have shown that we get the correct result provided ϕ_1 exists.

Now we return to show the existence of a ϕ_1 with the appropriate properties. As mentioned earlier, to the authors' knowledge, there is no simple and direct way of choosing ϕ_1 if $\zeta = -1$. In the case when $\phi \in \text{RV}_\infty(\zeta)$ for $\zeta > -1$, let $\phi_0(x) := x\phi(x)$. Then $\phi_0 \in \text{RV}_\infty(\zeta+1)$. Since $\zeta + 1 > 0$, there exists a differentiable $\phi_2 \in \text{RV}_\infty(\zeta+1)$ and $x_1 > 0$ such that $\phi_2(x)/\phi_0(x) \to 1$ as $x \to \infty$, and $\phi_2'(x) > 0$, $\phi_2(x) > 1$ for $x > x_1$. Now define ϕ_1 by

$$\phi_1(x) = \begin{cases} \phi_2(x)/x, & x > x_1 \\ \phi_2(x_1)/x_1, & 0 \le x \le x_1 \end{cases}$$

Hence $x\phi_1(x)$ is asymptotic to $\phi_2(x)$ and $x \mapsto x\phi_1(x)$ is non–decreasing. Moreover,

$$\lim_{x\to\infty} \frac{\phi_1(x)}{\phi(x)} = \lim_{x\to\infty} \frac{x\phi_1(x)}{x\phi(x)} = \lim_{x\to\infty} \frac{\phi_2(x)}{\phi_0(x)} = 1.$$

□

2.8 Proofs of Results from Section 2.5

Proof of Theorem 2.5.1. In order to prove Theorem 2.5.1 we would first need to prove analogies to Theorems 2.4.1 and 2.4.2 which apply to equations containing Markovian switching, specifically (2.5.3). Here we give an idea of how one would adapt the proof of Theorem 2.4.1 to incorporate the switching parameter.

Using similar methods to the non–switching case, we arrive at

$$d\tilde{Z}(t) = \left[\frac{2\langle \tilde{X}(t), f(\tilde{X}(t), \tilde{Y}(t))\rangle + \|g(\tilde{X}(t), \tilde{Y}(t))\|_F^2}{G^2(\tilde{X}(t), \tilde{Y}(t))}\right] dt + 2\sqrt{\tilde{Z}(t)}d\tilde{W}(t) \qquad (2.8.1)$$

where

$$G(x,y) = \begin{cases} \frac{\sqrt{\sum_{j=1}^r (\sum_{i=1}^d x_i g_{ij}(x,y))^2}}{\|x\|} & x \neq 0 \\ K_2 \geq c \geq K_1 & x = 0. \end{cases}$$

Let $u \in \mathbb{R}^d$ and $w \in \mathbb{S}$ and define the continuous function $D^* : \mathbb{R}^d \times \mathbb{S} \to \mathbb{R}$ by

$$D^*(u,w) = \frac{2\langle u, f(u,w)\rangle + \|g(u,w)\|_F^2}{G^2(u,w)}.$$

By (2.5.7),(2.5.8) and Cauchy–Schwarz, for all $w \in \mathbb{S}$ and for $u \neq 0$,

$$\begin{aligned} K_1^2 \leq G^2(u,w) &= \frac{\sum_{j=1}^r \left(\sum_{i=1}^d u_i g_{ij}(u,w)\right)^2}{\|u\|^2} \\ &\leq \frac{\sum_{i=1}^d u_i^2 \sum_{j=1}^r \sum_{i=1}^d g_{ij}^2(u,w)}{\|u\|^2} = \|g(u,w)\|_F^2 \leq K_2^2 \end{aligned}$$

and so $\|g(u,w)\|_F^2/G^2(u,w) \geq 1$, for all $u \in \mathbb{R}^d/\{0\}$, $w \in \mathbb{S}$. If $u = 0$, then, by (2.5.7) and the fact that G is defined to be constant at zero,

$$\frac{\|g(0,w)\|_F^2}{G^2(0,w)} = \frac{\|g(0,w)\|_F^2}{c^2} \geq \frac{K_0^2}{c^2}.$$

Define $K_3 = \min\{1, K_0^2/c^2\} > 0$. Then $\|g(u,w)\|_F^2/G^2(u,w) \geq K_3$, $\forall\, u \in \mathbb{R}^d$, $w \in \mathbb{S}$.

By Cauchy–Schwarz, $|\langle u, f(u,w)\rangle| \leq \|u\|.\|f(u,w)\|$. Define $F_j(u) = \|f(u,w_j)\|$. Then by (2.5.2), for any u, v with $\|u\|, \|v\| \leq n$ we have

$$|F_j(u) - F_j(v)| = \big|\|f(u,w_j)\| - \|f(v,w_j)\|\big| \leq \|f(u,w_j) - f(v,w_j)\| \leq K_n\|u - v\|.$$

Thus, the function $F_j : \mathbb{R}^d \to \mathbb{R}$ is locally Lipschitz continuous for each $j = 1, \ldots, N$. Let $F(u) := \max_j \|f(u,w_j)\|$. Then $u \mapsto F(u)$ is continuous since $\mathbb{S}$ is finite. Therefore for every $\varepsilon \in (0, 1 \wedge \frac{1}{4}K_1^2 K_3)$ sufficiently small there is an $X_2(\varepsilon) < 1$ such that $\|u\|F(u) < \varepsilon$ for all $\|u\| < X_2(\varepsilon)$. Then $|\langle u, f(u,w)\rangle| \leq \|u\|.\|f(u,w)\| \leq \|u\|F(u) < \varepsilon$ for all $w \in \mathbb{S}$ and for $\|u\| \leq X_2(\varepsilon)$. Thus, $\forall\ w \in \mathbb{S}$ and for $\|u\| \leq X_2(\varepsilon)$,

$$D^*(u,w) = \frac{2\langle u, f(u,w)\rangle}{G^2(u,w)} + \frac{\|g(u,w)\|_F^2}{G^2(u,w)} \geq \frac{-2\varepsilon}{K_1^2} + K_3 \geq \frac{K_3}{2} =: 2\phi_* > 0.$$

Define the function $D : \mathbb{R}^d \to \mathbb{R}$ by $D(u) = \min_{w\in\mathbb{S}} D^*(u,w)$. Then D is continuous since D^* is continuous. Moreover, $D^*(u,w) \geq D(u)$ for all $w \in \mathbb{S}$. Also, define the function $\Delta_- : \mathbb{R}^+ \to \mathbb{R}$ by $\Delta_-(x) = \min_{\|u\|=x} D(u)$. Then Δ_- is continuous since D is continuous. Moreover, $D(u) \geq \min_{\|j\|=\|u\|} D(j) = \Delta_-(\|u\|)$. Then, for $x \leq X_2(\varepsilon)$,

$$\Delta_-(x) = \min_{\|u\|=x} D(u) \geq \min_{\|u\|=x} 2\phi_* = 2\phi_* > 0.$$

This gives us an estimate for Δ_- on an interval close to zero. We now look for an estimate on an interval away from zero. From condition (2.5.5), it follows that $\forall\ \varepsilon \in (0,1)$ there exists $X_1(\varepsilon) > 1$ such that $\forall\ \|x\| > X_1(\varepsilon)$ and $\forall\ y \in \mathbb{S}$, $|\langle x, f(x,y)\rangle| \leq c_1(1+\varepsilon)\|x\|\phi(\|x\|)$. Since $x \neq 0$ we can use the fact that $\|g(x,y)\|_F^2/G^2(x,y) \geq 1$. Let $u \in \mathbb{R}^d$ such that $\|u\| > X_1(\varepsilon)$ and let $w \in \mathbb{S}$. Then,

$$D^*(u,w) = \frac{2\langle u, f(u,w)\rangle}{G^2(u,w)} + \frac{\|g(u,w)\|_F^2}{G^2(u,w)} \geq \frac{-2c_1(1+\varepsilon)}{K_1^2}\|u\|\phi(\|u\|) + 1.$$

Recall that $D(u) = \min_{w\in\mathbb{S}} D^*(u,w)$. Then for $\|u\| > X_1(\varepsilon)$ and $\forall\ w \in \mathbb{S}$,

$$D^*(u,w) \geq D(u) \geq \frac{-2c_1(1+\varepsilon)}{K_1^2}\|u\|\phi(\|u\|) + 1.$$

Also, recall $\Delta_-(x) = \min_{\|u\|=x} D(u)$. Then for $x > X_1(\varepsilon)$,

$$\Delta_-(x) = \min_{\|u\|=x} D(u) \geq \frac{-2c_1(1+\varepsilon)}{K_1^2} x\phi(x) + 1.$$

And this gives us our estimate for Δ_- on an interval away from zero. Note that the estimates for $\Delta_-(x)$ are the same as in the proof of the non–switching case so we can continue as per that proof to construct the function $\phi_-^{(\varepsilon)}$ such that $\Delta_-(x) + \phi_-^{(\varepsilon)}(x) > 0$ for $x \geq 0$. Then rewrite (2.8.1) as

$$d\tilde{Z}(t) = \left[-\phi_-^{(\varepsilon)}(\|\tilde{X}(t)\|) + D_{1,\varepsilon}(t)\right]dt + 2\sqrt{\tilde{Z}(t)}d\tilde{W}(t)$$

where, by definition,

$$D_{1,\varepsilon}(t) = D^*(\tilde{X}(t), \tilde{Y}(t)) + \phi_-^{(\varepsilon)}(\|\tilde{X}(t)\|) \geq D(\tilde{X}(t)) + \phi_-^{(\varepsilon)}(\|\tilde{X}(t)\|)$$

$$\geq \Delta_-(\|\tilde{X}(t)\|) + \phi_-^{(\varepsilon)}(\|\tilde{X}(t)\|) > 0.$$

Finally, construct the comparison process

$$d\underline{Z}_\varepsilon(t) = -\phi_-^{(\varepsilon)}\left(\sqrt{|\underline{Z}_\varepsilon(t)|}\right)dt + 2\sqrt{|\underline{Z}_\varepsilon(t)|}d\tilde{W}(t), \quad t \geq 0.$$

The proof now follows that of the non–switching case, since the switching dependency has been removed. We would then arrive at

$$\limsup_{t\to\infty} \frac{\|X(t)\|}{\Phi^{-1}\left(\frac{K_1^2(1-\varepsilon)}{2c_1}\log t\right)} \geq 1, \quad \text{a.s. on } \Omega_\varepsilon,$$

where Ω_ε is an almost sure event. Using similar methods to remove the switching dependency from an analogue of the proof of Theorem 2.4.2 we would get

$$\limsup_{t\to\infty} \frac{\|X(t)\|}{\Phi^{-1}\left(\frac{K_2^2(1+\varepsilon)}{2c_2}\log t\right)} \leq 1, \quad \text{a.s. on } \Omega_\varepsilon$$

We could then combine these two results in exactly the same way as was done in the proof of Theorem 2.4.4 to obtain the desired result (2.5.9). □

2.9 Proof of Stochastic Comparison Theorem

Proof of Theorem 2.2.1. Define $\tau_n := \inf\{t \geq 0 : X_1(t) = n \text{ or } X_2(t) = n\}$ and let $\Delta_n(t) = X_2(t\wedge\tau_n) - X_1(t\wedge\tau_n)$ for $t \geq 0$. Thus $\tau_n = \tau_n^{(1)} \wedge \tau_n^{(2)}$ where $\tau_n^{(1)}$ and $\tau_n^{(2)}$ are defined in the statement of Theorem 2.2.1. Also define $\Delta_n^+(t) = (\Delta_n(t))_+$ for all $t \geq 0$. By defining M by

$$M(t) = \int_0^{t\wedge\tau_n} \{\sigma(X_2(s)) - \sigma(X_1(s))\}\, dB(s)$$

we have

$$\Delta_n(t) = X_2(0) - X_1(0) + \int_0^{t\wedge\tau_n} \{\beta_2(s) - \beta_1(s)\}\, ds + M(t)$$

and therefore by the Tanaka–Meyer formula (see [46], Chapter 3, Section 7) we obtain

$$\Delta_n^+(t) = (X_2(0) - X_1(0))_+ + \int_0^{t\wedge\tau_n} 1_{(0,\infty)}(\Delta_n(s))\,\{\beta_2(s) - \beta_1(s)\}\,ds$$
$$+ \int_0^{t\wedge\tau_n} 1_{(0,\infty)}(\Delta_n(s))\,\{\sigma(X_2(s)) - \sigma(X_1(s))\}\,dB(s) + \Lambda_t^0(\Delta_n), \quad (2.9.1)$$

where $\Lambda_t^0(\Delta_n)$ is the local time of Δ_n at zero. Next, following Exercise 3.7.12 in [46], we show that $\Lambda_t^0(\Delta_n) = 0$ for all $t \geq 0$ a.s. Towards this end, we note that $M(t) = M(\tau_n)$ for all $t \geq \tau_n$ and also that $\langle M\rangle(t) = \langle M\rangle(\tau_n)$ for all $t \geq \tau_n$ a.s. Define $\rho_0(x) = x$ for $x \geq 0$. Then for $t \geq \tau_n$ we have

$$\int_0^t \frac{1_{(0,\infty)}(\Delta_n(s))}{\rho_0(\Delta_n(s))}\,d\langle M\rangle(s) = \int_0^{\tau_n} \frac{1_{(0,\infty)}(\Delta_n(s))}{\rho_0(\Delta_n(s))}\,d\langle M\rangle(s),$$

so therefore for any $t \geq 0$ we have

$$\int_0^t \frac{1_{(0,\infty)}(\Delta_n(s))}{\rho_0(\Delta_n(s))}\,d\langle M\rangle(s) \leq \int_0^{\tau_n} \frac{1_{(0,\infty)}(\Delta_n(s))}{\rho_0(\Delta_n(s))}\,d\langle M\rangle(s).$$

For $s \leq \tau_n$ we have that $\langle M\rangle(s) = \int_0^s \{\sigma(X_2(u)) - \sigma(X_1(u))\}^2\,du$, so by (2.2.2) we get

$$\begin{aligned}\int_0^{\tau_n} \frac{1_{(0,\infty)}(\Delta_n(s))}{\rho_0(\Delta_n(s))}\,d\langle M\rangle(s) &= \int_0^{\tau_n} \frac{1_{(0,\infty)}(\Delta_n(s))}{\rho_0(\Delta_n(s))}\,\{\sigma(X_2(s)) - \sigma(X_1(s))\}^2\,ds \\ &\leq \int_0^{\tau_n} \frac{1_{(0,\infty)}(\Delta_n(s))}{\rho_0(\Delta_n(s))} K_n^2 |X_2(s) - X_1(s)|\,ds \\ &= K_n^2 \int_0^{\tau_n} \frac{1_{(0,\infty)}(\Delta_n(s))}{\rho_0(\Delta_n(s))} |\Delta_n(s)|\,ds \leq K_n^2 \tau_n.\end{aligned}$$

By hypothesis either $\tau_n^{(1)} < +\infty$ a.s. or $\tau_n^{(2)} < +\infty$ a.s., so $\tau_n < +\infty$ a.s., and

$$\int_0^t \frac{1_{(0,\infty)}(\Delta_n(s))}{\rho_0(\Delta_n(s))}\,d\langle M\rangle(s) \leq K_n^2 \tau_n < +\infty, \quad \text{a.s.}$$

By the occupation density formula (Theorem 3.7.1 (iii) in [46]) we have

$$\int_0^t \frac{1_{(0,\infty)}(\Delta_n(s))}{\rho_0(\Delta_n(s))} d\langle M\rangle(s) = 2\int_0^\infty \frac{1}{\rho_0(a)} \Lambda_t^a(\Delta_n)\,da. \quad (2.9.2)$$

By the right continuity in a of $\Lambda_t^a(\Delta_n)$, if $\Lambda_t^0(\Delta_n) > 0$, the right–hand side of (2.9.2) is infinite, because $\int_{(0,\epsilon)} 1/\rho_0(a)\,da = \infty$ and $\rho_0(a) > 0$ for $a > 0$. But this introduces a contradiction because the left–hand side of (2.9.2) is finite. Therefore $\Lambda_t^0(\Delta_n) = 0$ for all $t \geq 0$ a.s., and so, returning to (2.9.1), by using the fact that $X_1(0) \geq X_2(0)$ we get

$$\Delta_n^+(t) = \int_0^{t\wedge\tau_n} 1_{(0,\infty)}(\Delta_n(s))\,\{\beta_2(s) - \beta_1(s)\}\,ds + \int_0^{t\wedge\tau_n} 1_{(0,\infty)}(\Delta_n(s))\,\{\sigma(X_2(s)) - \sigma(X_1(s))\}\,dB(s). \quad (2.9.3)$$

Now using (2.2.1), we have $\beta_2(s) - \beta_1(s) \leq b(X_2(s)) - b(X_1(s)) \leq |b(X_2(s)) - b(X_1(s))|$. We consider two cases: if $\Delta_n(s) > 0$ then by (2.2.3) we have, for $s \leq t \wedge \tau_n \leq \tau_n$,

$$\beta_2(s) - \beta_1(s) \leq K_n|X_2(s) - X_1(s)| = K_n|X_2(s\wedge\tau_n) - X_1(s\wedge\tau_n)| = K_n|\Delta_n(s)| = K_n\Delta_n^+(s).$$

Therefore if $\Delta_n(s) > 0$ and $s \leq t \wedge \tau_n$ we have

$$1_{(0,\infty)}(\Delta_n(s))\,\{\beta_2(s) - \beta_1(s)\} = \beta_2(s) - \beta_1(s) \leq K_n\Delta_n^+(s).$$

On the other hand, if $\Delta_n(s) \leq 0$ then for $s \leq t \wedge \tau_n \leq \tau_n$

$$1_{(0,\infty)}(\Delta_n(s))\,\{\beta_2(s) - \beta_1(s)\} = 0 = K_n\Delta_n^+(s).$$

So in both cases we have $1_{(0,\infty)}(\Delta_n(s))\,\{\beta_2(s) - \beta_1(s)\} \leq K_n\Delta_n^+(s)$. Applying this in (2.9.3) we arrive at

$$\begin{aligned}\Delta_n^+(t) &\leq K_n \int_0^{t\wedge\tau_n} \Delta_n^+(s)\,ds + \int_0^{t\wedge\tau_n} 1_{(0,\infty)}(\Delta_n(s))\,\{\sigma(X_2(s)) - \sigma(X_1(s))\}\,dB(s)\\ &\leq K_n \int_0^{t} \Delta_n^+(s)\,ds + \int_0^{t\wedge\tau_n} 1_{(0,\infty)}(\Delta_n(s))\,\{\sigma(X_2(s)) - \sigma(X_1(s))\}\,dB(s).\end{aligned}$$

Applying the optional sampling theorem we arrive at

$$\mathbb{E}[\Delta_n^+(t)] \leq K_n \int_0^t \mathbb{E}[\Delta_n^+(s)]\,ds, \quad t \geq 0.$$

The function $t \mapsto \mathbb{E}[\Delta_n^+(t)]$ is continuous so by Gronwall's inequality, $\mathbb{E}[\Delta_n^+(t)] = 0$ for each $t \geq 0$ and $\Delta_n^+(t) = 0$ for each fixed $t \geq 0$. However, as $t \mapsto \Delta_n^+(t)$ is continuous we have $\Delta_n^+(t) = 0$ for all $t \geq 0$ a.s. This implies that $X_2(t \wedge \tau_n) - X_1(t \wedge \tau_n) = \Delta_n(t) \leq 0$ for all $t \geq 0$ a.s. Hence the event $\Omega_n := \{\omega : X_2(t \wedge \tau_n, \omega) \leq X_1(t \wedge \tau_n, \omega) \text{ for all } t \geq 0\}$ is almost sure. Let $\Omega^* = \cap_{n \in \mathbb{N}} \Omega_n$. Then Ω^* is almost sure and if $\tau_n \to \infty$ as $n \to \infty$ a.s. we have $\Omega^* = \{\omega : X_2(t, \omega) \leq X_1(t, \omega) \text{ for all } t \geq 0\}$, proving the result. To justify that $\lim_{n\to\infty} \tau_n = \infty$ a.s. we simply note that neither X_1 nor X_2 explode in finite time. $\square$

Chapter 3

The Size of the Largest Fluctuations in a Market Model with Markovian Switching

3.1 Introduction

In Chapter 2 we examined the large fluctuations of a general class of *stationary* SDEs; this chapter applies similar techniques (the stochastic comparison principle and Motoo's Theorem) to a class of *non-stationary* SDEs which can be used to build a financial market model. These notions of stationarity and non-stationarity are defined in the preliminaries. The motivation for studying equations with switching in a financial setting comes from observations in financial market econometrics which suggest that security prices often move from bearish to bullish (or other) regimes. These regimes are modelled by the presence of the Markov process Y. One of the seminal contributions on the econometric analysis of financial time series subject to these regime shifts is [33], and a recent monograph covering this topic, amongst others, is [27]. Moreover, examples of stochastic volatility (SV) models with switching can be found in [22], [23] and [79]. Numerical methods for such SV models with Markovian switching are examined in [57]. Interest rate models with switching arise in [76].

In contrast to Chapter 2, this chapter deals with scalar non–autonomous SDEs with Markovian switching of the form

$$dX(t) = f(X(t), Y(t), t)\, dt + g(X(t), Y(t), t)\, dB(t), \tag{3.1.1}$$

where $g(x, y, t)$ and $xf(x, y, t)$ are uniformly bounded above and below in (x, y, t), and Y

is an irreducible continuous–time Markov chain with finite state space $\mathbb{S}$ independent of the Brownian motion B. If the lower bound on $xf(x,y,t)$ is sufficiently large, we show that X obeys upper and lower laws of the iterated logarithm, in the sense that

$$\sqrt{K_2} \le \limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{2t\log\log t}} \le \sqrt{K_1}, \quad \text{a.s.}$$

where $g^2(x,y,t) \in [K_2, K_1]$. In the case when g additionally obeys $g(x,y,t) = \gamma(y)$, it can be shown that

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{2t\log\log t}} = \sigma_*, \quad \text{a.s.} \tag{3.1.2}$$

where $\sigma_*^2 = \sum_{j\in\mathbb{S}} \gamma^2(j)\pi_j$ and $\pi = (\pi_j)_{j\in\mathbb{S}}$ is the stationary distribution of Y. The proofs rely on time change and comparison arguments to construct upper and lower bounds on $|X|$. However, in contrast to the proofs in Chapter 2, the equations must undergo changes of time and scale in order to transform them into stationary processes. The large deviations of the stationary processes are then determined by Motoo's Theorem. These large deviation results are then applied to a security price model, where the security price S obeys

$$dS(t) = \mu S(t)\,dt + S(t)\,dX(t), \quad t \ge 0, \tag{3.1.3}$$

and X obeys (3.1.1). We assume that the movement between regimes is not influenced by the stock price or returns by presuming that Y and the driving Brownian motion B are independent.

The classical Geometric Brownian Motion model of stock evolution assumes that the market is informationally efficient, following forms of the Efficient Market Hypothesis (EMH). A classical statement and discussion about the EMH and its ramifications may be found in e.g., Fama [24] or the volume edited by Cootner [17]. However, in recent times, econometric evidence suggesting that financial markets might be inefficient has accumulated (see e.g., [50]). The equation (3.1.3) models an inefficient market, since the increments of the cumulative returns process $\mu t + X(t)$ are not independent. However, the fact that $xf(x, y, t)$ is uniformly bounded means that the process X does not depart too much (in some sense) from Brownian motion, thereby placing limits on the inefficiency of the market, particularly when the price departs too far from its trend rate of growth. Therefore, the assumption that $xf(x, y, t)$ be bounded can be seen as hypothesising that the market is not "too inefficient". Although the informational inefficiency of a market can be proven directly for certain models (see [7]), the same argument does not hold for the model described by (3.1.3) due to the unobservable switching parameter.

Despite the presence of regime shifts and inefficiency, we can still deduce that the new market model enjoys some of the properties of standard GBM models. Having established the existence of a trend rate of growth in the price, we use results about the solution of (3.1.1) to show that the large deviations of the price from this trend rate of growth obey a law of the iterated logarithm, just as in standard models. Finally, although the returns are non–Gaussian, we can nevertheless show that the running maxima of the returns have the same almost sure rate of growth as those of a stationary Gaussian process.

This chapter also considers the size of the largest fluctuations of the returns process $R_\delta(t) = \log\big(S(t)/S(t-\delta)\big)$, $t \geq \delta$, over $\delta > 0$ time units. It is shown that when the diffusion coefficient is of the form $g(x, y, t) = \gamma(y)$, then under certain conditions on the drift coefficient we have

$$\limsup_{t\to\infty} \frac{|R_\delta(t)|}{\sqrt{2\log t}} = \sigma_H \sqrt{\delta}, \quad \text{a.s.}, \tag{3.1.4}$$

where $\sigma_H = \max_{j\in\mathbb{S}} |\gamma(j)| > 0$.

When compared to (3.1.2) this result (3.1.4) reveals an interesting phenomenon, unique to equations with switching. The magnitiude of the fluctuations in (3.1.2) is determined by an "average" of the different volatility levels that the process can switch between, weighted by the stationary probability distribution of the Markov chain. On the other hand, the magnitude of the fluctuations of the returns over a short time period, (3.1.4), is determined only by the largest value of the diffusion coefficient, namely σ_H. This indicates that the returns fluctuate most during the periods of highest volatility, something that one would intuitively expect. This effect cannot be found in equations without switching, as in the constant diffusion case we would have $\sigma_* = \sigma_H = \sigma$ and there would be no distinction. Both results, (3.1.2) and (3.1.4), highlight the benefits of considering the special case where the diffusion coefficient depends only on the switching parameter, i.e. that $g(x, y, t) = \gamma(y)$. In this case the switching process has a *direct* influence on the magnitude of the large fluctuations, while in Chapter 2 for example it was not clear that the switching would have such an effect on the magnitude of the fluctuations.

The chapter is organised as follows. Useful mathematical preliminaries for this chapter are detailed in Section 3.2 while the main results on iterated logarithm growth rates for the solution of (3.1.1) are given in Section 3.3. In Section 3.4, these results are applied to a stock price model. The proofs of all results are postponed to the final two sections.

3.2 Mathematical Preliminaries

3.2.1 Markov Chains and Jump Processes

Let Y be a continuous–time Markov chain with state space $\mathbb{S}$. To make our theory more understandable, we assume the state space of the Markov chain is finite, say $\mathbb{S} = \{1, 2, \cdots, N\}$. As a standing hypothesis we assume in this chapter that the Markov chain is *irreducible*.

For $n = 0, 1, 2, \ldots$, denote by $M_i(n)$ the length of the n^{th} visit to state i, by $T_i(n)$ the time of the n^{th} return to i and by $L_i(n)$ the length of the n^{th} excursion to i. For a visualisation of these quantities see Figure 3.1. By the strong Markov property (of a continuous jump chain) at the stopping times $T_i(n)$, we find that

$$L_i(1), L_i(2), \dots \text{ are independent and identically distributed with mean } m_i, \tag{3.2.1}$$

and that

$$M_i(1), M_i(2), \dots \text{ are independent and identically distributed with mean } 1/q_i, \tag{3.2.2}$$

where $q_i = \sum_{j \neq i} \gamma_{ij}$ is the off–diagonal row sum of the transition matrix. Moreover, since the M_i's (the holding times, or waiting times) are exponentially and identically distributed in a Markov jump process, we know that

$$\text{the sequence of random variables } M_i(1), M_i(2), \dots \text{ has finite variance } \sigma^2_{M_i}. \tag{3.2.3}$$

An important technical requirement in our analysis is that the second moments of the excursion lengths $L_i(n)$ are finite for all $i \in \mathbb{S}$ and $n \in \mathbb{N}$. Since the process Y is time–homogeneous, the lengths of the excursions $L_i(n)$ are identically distributed for all n. Moreover, since the excursion time $L_i(1)$ is simply the first passage time to state i from state i, it suffices to show that the second moment of the first passage time from i to i is finite for all $i \in \mathbb{S}$. What follows is doubtless well–known to researchers in Markov jump processes and Markov chains, but may be less well–known to those whose backgrounds are in stochastic differential equations, such as the author, and is therefore included for readers with similar backgrounds.

The connection between the distributions of first passage times and waiting (or holding) times for semi–Markov processes was established by Pyke [68]. He established that the first passage distribution function G_{ij} from state i to state j could be written in terms of the waiting time distributions F_{ij} (where i is the state currently occupied, and j the state to be visited next) according to

$$G_{ij}(t) = F_{ij}(t) + \sum_{k=1, k\neq j}^{N} \int_{(0,t]} F_{ik}(t-s)\, dG_{kj}(s), \quad t \geq 0. \tag{3.2.4}$$

See [68, Lemma 3.2] or [43, Lemma 1.1]. The jump process Y is a special type of semi–Markov process with exponentially distributed holding times, so in our case the F_{ij}'s are exponential distributions. Therefore the second moments of the distributions F_{ij} are finite. A consequence of this and the renewal equation (3.2.4) is that the second moments of the G_{ij}'s are finite for all $i, j \in \mathbb{S}$. See e.g., [43, Lemma 2.1], where a formula relating the moments of the F's and G's is deduced. Therefore, by invoking this theory, we have that

$$\text{the sequence of random variables } L_i(1), L_i(2), \ldots \text{ has finite variance } \sigma^2_{L_i}. \tag{3.2.5}$$

3.2.2 Stochastic comparison for equations with non-stationary solutions

The method of proof employed in Chapter 2 relies on creating comparison processes (which have stationary solutions) to which we can apply Motoo's Theorem. This theorem allows us to determine the *exact* asymptotic growth rate of the running maxima of a stationary (or asymptotically stationary) process governed by an autonomous SDE.

However, the use of stochastic comparison principles does not guarantee that the comparison processes will have stationary solutions: in fact in this chapter we deal with equations with non–stationary solutions. Nonetheless it is possible in some cases to apply a change in both time and scale to an equation with non–stationary solutions to transform it into an equation with stationary solutions. The asymptotic behaviour of a process transformed in such a way can then be determined by Motoo's theorem.

Take, for example, a simple 1–dimensional Ornstein–Uhlenbeck process governed by

$$dX(t) = -X(t)\,dt + dB(t), \quad t \geq 0, \quad \text{where } X(0) = 0.$$

It can be shown that this process has stationary solutions and that Motoo's theorem can be applied. On the other hand, we can solve this equation explicitly to get

$$X(t) = e^{-t}\int_0^t e^s\,dB(s) = e^{-t}M(t), \tag{3.2.6}$$

where we define the martingale $M(t) := \int_0^t e^s\,dB(s)$ with quadratic variation given by

$$\langle M\rangle(t) = \int_0^t e^{2s}\,ds = \frac{1}{2}(e^{2t}-1).$$

Then by the martingale time–change theorem (cf., e.g., Theorem 3.4.6 in [46]), we may define a new Brownian motion B^* by $B^*(\langle M\rangle(t)) = M(t)$. Therefore, (3.2.6) becomes

$$X(t) = e^{-t}B^*(\langle M\rangle(t)) = e^{-t}B^*\big(\frac{1}{2}(e^{2t}-1)\big).$$

The significance of this relation is that the Brownian motion itself, $B^*(t)$, is non–stationary but by applying the change of time, $(e^{2t}-1)/2$, and change of scale, e^{-t}, it is transformed into the process $X(t)$ which has stationary solutions.

For some of the proofs in this chapter, we employ a similar change of time and scale before applying Motoo's theorem.

3.3 Statement and Discussion of Main Results

In this section we give sufficient conditions ensuring law of the iterated logarithm–type behaviour for the solution of (3.3.1).

Let $f, g : \mathbb{R}\times\mathbb{S}\times[0,\infty)\to\mathbb{R}$ be continuous functions obeying local Lipschitz continuity and linear growth conditions. Let $X(0) = x_0$ and consider the stochastic differential equation with Markovian switching given by

$$dX(t) = f(X(t), Y(t), t)\,dt + g(X(t), Y(t), t)\,dB(t). \tag{3.3.1}$$

We assume, unless otherwise stated, that there exists $\rho > 0$ such that

$$xf(x, y, t) \le \rho \quad \text{for all } (x, y, t) \in \mathbb{R}\times\mathbb{S}\times[0,\infty), \tag{3.3.2}$$

and that f is globally bounded in the sense that

$$|f(x, y, t)| \le \bar{F} < +\infty, \quad \text{for all } (x, y, t) \in \mathbb{R}\times\mathbb{S}\times[0,\infty). \tag{3.3.3}$$

Under the above conditions, there is a unique continuous and adapted process which satisfies (3.3.1) (see e.g. [62]). We make the standing assumption throughout the chapter that f and g obey these continuity and growth restrictions, and that Y is an irreducible continuous–time Markov chain with finite state space $\mathbb{S}$. For economy of exposition these assumptions are not explicitly repeated in the statement of theorems in this chapter.

The first two theorems deal with upper and lower estimates on the asymptotic growth rate of the running maxima.

Theorem 3.3.1. *Let X be the unique adapted continuous solution satisfying* (3.3.1) *and let f obey* (3.3.2)*. If there exist positive real numbers K_1 and K_2 such that*

$$K_2 \le g^2(x, y, t) \le K_1, \quad \text{for all } (x, y, t) \in \mathbb{R} \times \mathbb{S} \times [0, \infty) \tag{3.3.4}$$

then X satisfies

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{2t \log\log t}} \le \sqrt{K_1}, \quad a.s. \tag{3.3.5}$$

The result and hypotheses of this theorem are similar to those of a theorem in Mao [52], in which no switching process is present. Here in Theorem 3.3.1, a sharper upper bound on the solution is obtained in that if one were to apply Mao's theorem in this case we would get $\sqrt{K_1}\sqrt{e}$ on the right–hand side of (3.3.5). The sharper bound comes at the expense of a two–sided bound on the diffusion coefficient g. The proof in [52] employs martingale and integral inequalities, while Theorem 3.3.1 is proven by means of a comparison result. An advantage of this comparison approach is that a similar argument also yields a lower estimate on the large fluctuations of the solution, which we have been unable to obtain using the methods in [52].

Theorem 3.3.2. *Let X be the unique adapted continuous solution satisfying* (3.3.1)*. If there exist real numbers K_1 and K_2 such that* (3.3.4) *holds, and there is an $L \in \mathbb{R}$ such* *that*

$$\inf_{(x,y,t)\in\mathbb{R}\times\mathbb{S}\times[0,\infty)} \frac{xf(x,y,t)}{g^2(x,y,t)} =: L > -\frac{1}{2}, \tag{3.3.6}$$

then X satisfies

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{2t\log\log t}} \geq \sqrt{K_2}, \quad a.s. \tag{3.3.7}$$

We can combine the arguments used to prove these results to obtain a general result on the exact size of the large fluctuations, under the assumption that the diffusion coefficient depends only on the process Y. The result plays a role later in the chapter when we consider applications of these pathwise large deviation results to finance.

Corollary 3.3.1. *Let X be the unique continuous adapted process satisfying the equation*

$$dX(t) = f(X(t), Y(t), t)\, dt + \gamma(Y(t))\, dB(t), \tag{3.3.8}$$

where $\gamma : \mathbb{S} \to \mathbb{R} \setminus \{0\}$ and $X(0) = x_0$. If there exists a real number $\rho > 0$ such that

$$\sup_{(x,y,t)\in\mathbb{R}\times\mathbb{S}\times[0,\infty)} \frac{xf(x,y,t)}{\gamma^2(y)} \leq \rho \quad \text{and} \quad \inf_{(x,y,t)\in\mathbb{R}\times\mathbb{S}\times[0,\infty)} \frac{xf(x,y,t)}{\gamma^2(y)} > -\frac{1}{2}, \tag{3.3.9}$$

then

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{2t\log\log t}} = \sigma_*, \quad a.s. \tag{3.3.10}$$

where

$$\sigma_*^2 = \sum_{j\in\mathbb{S}} \gamma^2(j)\pi_j, \tag{3.3.11}$$

and π is the stationary probability distribution of Y defined by (1.0.5)*.*

The first condition in (3.3.9) is equivalent to (3.3.2). The second condition is more subtle. Although it is sufficient to establish an iterated logarithm–type result, it is not a necessary condition to do so: Theorem 3.3.3 which follows justifies the second part of this remark. However, examples of equations (3.3.1) exist in which the second condition in (3.3.9) is false, and the solutions do not obey iterated logarithm type growth bounds. We supply such an example now.

Example 3.3.1. Suppose in (3.3.1) that $f(x, y, t) = f(x)$ and that $g(x, y, t) = \sigma \neq 0$, and let f obey $\lim_{x\to\infty} xf(x) = \lim_{x\to-\infty} xf(x) = L < -\sigma^2/2$. Then, provided f is continuous, the first condition in (3.3.9) is true, but $\inf_{x\in\mathbb{R}} xf(x) < -\sigma^2/2$, and so the second condition in (3.3.9) is false. Routine calculations show that the conditions of Motoo's theorem hold. Moreover, by determining the asymptotic behaviour of the scale function, we can use Motoo's theorem to show that

$$\limsup_{t\to\infty} \frac{\log |X(t)|}{\log t} \quad \text{exists a.s.}$$

is deterministic and is strictly less than 1/2. Therefore a solution of (3.3.1) under these conditions cannot obey the law of the iterated logarithm. It can be seen that the second part of condition (3.3.9) is quite a sharp hypothesis, since in the case that $L > -\sigma^2/2$ we can find functions f such that the second part of (3.3.9) holds, and hence the law of the iterated logarithm holds also.

Remark 3.3.1. We observe that (3.3.10) provides an *exact* rate of growth of the running maxima of $|X|$. This is in contrast with the results of Theorems 3.3.1 and 3.3.2, in which only *bounds* on the growth rate are determined. We also notice that the presence of the switching process Y influences the rate of growth, because the value of σ_* in (3.3.11) depends on the stationary distribution of Y. On the other hand, it is not immediately clear from Theorems 3.3.1 and 3.3.2 that the switching process can influence the asymptotic behaviour so directly, because the bounds on the diffusion coefficients K_1 and K_2 are independent of the switching state Y. Finally, not only is the a.s. rate of growth of the running maxima deterministic, but it also can be computed explicitly once the generator of Y and the diffusion coefficient γ are known. The stronger conclusion of Corollary 3.3.1 relies upon the stronger assumption that the diffusion coefficient depends only on the Markov process Y.

In Theorems 3.3.1, 3.3.2 and in Corollary 3.3.1, we assume that f obeys a pointwise bound that depends on x. We can allow f to violate such a bound, provided any "spikes" that may be present in f are sufficiently narrow. This is achieved by the choice of hypothesis (3.3.12) in the statement of Theorem 3.3.3 below.

Theorem 3.3.3. *Let X be the unique continuous adapted process satisfying* (3.3.1)*, with $X(0) = x_0$. If there exist positive real numbers K_1, K_2 such that* (3.3.4) *holds, and there is a locally Lipschitz continuous function $\tilde{f}$ such that*

$$\frac{|f(x,y,t)|}{g^2(x,y,t)} \leq \tilde{f}(x), \quad \tilde{f} \in \mathcal{L}^1(\mathbb{R};\mathbb{R}^+), \tag{3.3.12}$$

then X almost surely obeys

$$\frac{\sqrt{K_2}e^{-2\sup_{x\in\mathbb{R}}\int_0^x(-\tilde{f}(y))dy}}{e^{-2\int_0^\infty(-\tilde{f}(y))dy}} \leq \limsup_{t\to\infty}\frac{X(t)}{\sqrt{2t\log\log t}} \leq \frac{\sqrt{K_1}e^{-2\inf_{x\in\mathbb{R}}\int_0^x\tilde{f}(y)dy}}{e^{-2\int_0^\infty\tilde{f}(y)dy}} \tag{3.3.13a}$$

$$\frac{-\sqrt{K_1}e^{-2\inf_{x\in\mathbb{R}}\int_0^x(-\tilde{f}(y))dy}}{e^{2\int_{-\infty}^0(-\tilde{f}(y))dy}} \leq \liminf_{t\to\infty}\frac{X(t)}{\sqrt{2t\log\log t}} \leq \frac{-\sqrt{K_2}e^{-2\sup_{x\in\mathbb{R}}\int_0^x\tilde{f}(y)dy}}{e^{2\int_{-\infty}^0\tilde{f}(y)dy}}. \tag{3.3.13b}$$

We notice in this result that both positive and negative large fluctuations obey an iterated logarithm growth bound: this contrasts with the results of Theorem 3.3.1, 3.3.2 and Corollary 3.3.1, in which the growth bounds are for the absolute value of the process. While the estimates on the normalising constants $\sqrt{K_1}$ and $\sqrt{K_2}$ in Theorems 3.3.1 and 3.3.2 are sharper than those obtained in Theorem 3.3.3, we are able to dispense with the pointwise bounds required in (3.3.9).

3.4 Application to Financial Market Models

In this section, we consider the application of the results from the previous section to a variant of Geometric Brownian Motion (GBM) which includes Markovian switching. In the first subsection, we state and discuss some properties of standard models, and then do likewise for analogous results for the switching model. These results concentrate on the long run growth rate, the size of the largest departures from the trend, and the large fluctuations of the incremental returns. In the second subsection, we specialise our results to a market in which there are only two regimes of "high" and "low" volatility.

3.4.1 Discussion of main results

We begin by reviewing briefly some mathematical and economic properties of GBM. GBM is one of the canonical models used to describe the stochastic evolution of asset prices (see e.g., Karatzas and Shreve [47]), and is behind the classical Black–Scholes–Merton option pricing formula (see e.g., Merton [64]). This work has given rise to a great variety of alternative market models and has led to an explosion in the variety of financial instruments that can be priced; a flavour of this activity can be gleaned from the popular textbook [41].

As is well–known, GBM can be characterised as the unique solution of the linear stochastic differential equation

$$dS^*(t) = \mu S^*(t)\,dt + \sigma S^*(t)\,dB(t), \quad t \geq 0, \tag{3.4.1}$$

where $S^*(0) > 0$. In the context of financial economics, μ is the instantaneous mean rate of growth of the price, and σ its instantaneous volatility. The importance of the GBM model is embodied by the following fact: if security returns are stationary and independent (so that the market is informationally efficient) and the stock price process S^* varies continuously in continuous time, then S^* *must* obey (3.4.1). It is well–known that the logarithm of S^* is a Brownian motion with drift, having mean and variance at time t of $(\mu - \frac{\sigma^2}{2})t$ and $\sigma^2 t$ respectively, and that S^* grows exponentially according to

$$\lim_{t\to\infty} \frac{\log S^*(t)}{t} = \mu - \frac{1}{2}\sigma^2, \quad \text{a.s.} \tag{3.4.2}$$

Furthermore the maximum size of the large deviations from this growth trend obey the law of the iterated logarithm:

$$\limsup_{t\to\infty} \frac{|\log S^*(t) - (\mu - \frac{1}{2}\sigma^2)t|}{\sqrt{2t\log\log t}} = |\sigma|, \quad \text{a.s.} \tag{3.4.3}$$

Before discussing other properties of S^*, we explore the significance and implications of the result (3.4.3) in terms of finance. Since S^* represents the price of a *risky* asset, we cannot expect that S^* grows at *exactly* the rate $\exp[(\mu-\sigma^2/2)t]$ as $t \to \infty$. Indeed, as real stock prices experience departures from such steady growth rates (for example in market crashes or bubbles), it is advantageous for any model of these prices to also have this property and to be able to determine how *large* these bubbles or crashes are likely to be from the perspective of both long–term investment and portfolio management.

This leads us to consider the size of the *largest* fluctuations from the trend rate of growth. We can study these large fluctuations by first removing the exponential trend from the stock price, leaving us with the process $\log S^*(t) - (\mu - \sigma^2/2)t$, which gives the logarithm of the departure from the trend. The largest deviations of this departure obey a law of the iterated logarithm, according to (3.4.3). In terms of the stock price itself, roughly speaking, this means that the stock can be *bigger* than the smooth exponential trend by a factor of $\exp[\sigma\sqrt{2t\log\log t}]$, or can be *smaller* by a factor of $\exp[-\sigma\sqrt{2t\log\log t}]$ as $t \to \infty$, a.s.

Moreover the δ–increments of $\log S^*$ are stationary and Gaussian, with the mean and variance of the increments depending linearly on δ. These δ–increments, defined by $R^*_\delta(t) = \log(S^*(t)/S^*(t-\delta))$, therefore obey

$$\limsup_{t\to\infty} \frac{|R^*_\delta(t)|}{\sqrt{2\log t}} = |\sigma|\sqrt{\delta}, \quad \text{a.s.} \tag{3.4.4}$$

In the following section, we propose a variant of (3.4.1) in which the stock price S is the solution of

$$dS(t) = \mu S(t)\,dt + S(t)\,dX(t) \quad t \geq 0. \tag{3.4.5}$$

Here the driving Brownian motion in (3.4.1) has been replaced by a semi–martingale X which partly depends on a continuous–time Markov chain. The model departs from (3.4.1) in that the returns are no longer stationary nor independent. Note that if the cumulative returns on the security with price $S = \{S(t) : t \geq 0\}$ up to time t are defined by $R(t)$, then

$$R(t) = \log\big(S(t)/S(0)\big), \quad t \geq 0, \tag{3.4.6}$$

and the (log) returns of the security over the time interval $[t-\delta, t]$ are defined by

$$R_\delta(t) = R(t) - R(t-\delta) = \log(S(t)/S(t-\delta)), \quad t \geq \delta. \tag{3.4.7}$$

With these definitions we show that the processes S and R_δ obey analogous properties to (3.4.2), (3.4.3) and (3.4.4). Therefore, the stock price process grows exponentially, experiences large deviations from the trend growth rate of iterated logarithm type, and incremental returns have the same rate of growth as those of stationary Gaussian processes, despite R_δ being non–Gaussian. The above claims are made precise in the following Theorems in this section, whose proofs are supplied in Section 6.

Theorem 3.4.1. *Let Y be a continuous–time Markov process with state space $\mathbb{S}$. Let X be the unique continuous adapted process governed by*

$$dX(t) = f(X(t), Y(t), t)\,dt + \sigma\,dB(t), \quad t \geq 0, \tag{3.4.8}$$

with $X(0) = 0$. Let $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}\setminus\{0\}$, and S be the unique continuous adapted process defined by (3.4.5)*, with $S(0) = s_0 > 0$. Suppose that f obeys* (3.3.3) *and that* (3.3.9) *holds in this special case where $\gamma(y) = \sigma$ for all y. Then:*

(i)

$$\lim_{t\to\infty} \frac{\log S(t)}{t} = \mu - \frac{\sigma^2}{2}, \quad a.s.$$

(ii)

$$\limsup_{t\to\infty} \frac{|\log S(t) - (\mu - \frac{\sigma^2}{2})t|}{\sqrt{2t\log\log t}} = |\sigma|, \quad a.s. \tag{3.4.9}$$

(iii) If R_δ is given by (3.4.7)*, then for each* $0 < \delta < \infty$

$$\limsup_{t\to\infty} \frac{|R_\delta(t)|}{\sqrt{2\log t}} = |\sigma|\sqrt{\delta}, \quad a.s.$$

Despite the presence of the Markov process Y (which introduces regime shifts) and the X–dependent drift term f in (3.4.8) (which introduces inefficiency), we see that S obeys the same asymptotic properties as S^*, namely (3.4.2), (3.4.3) and (3.4.4). These properties of S^* are shared by S because condition (3.3.9) guarantees that f becomes small for large values of X, thereby forcing S and S^* to remain close, in some sense. Indeed, if f is identically zero, we see that S and S^* actually coincide.

On the other hand, the analysis is now more complicated because the increments are (in general) neither independent nor Gaussian, and it is not possible to write down an explicit formula for S in terms of B and Y. This complication is worthwhile, however, because it stems from the addition of inefficiency and regime shifts into the market model.

3.4.2 State–independent diffusion coefficient

We now return to the special case where the diffusion coefficient depends *only* on the switching process Y. Let $f : \mathbb{R} \times \mathbb{S} \times [0,\infty) \to \mathbb{R}$ and $\gamma : \mathbb{S} \to \mathbb{R}$ be continuous functions obeying local Lipschitz continuity and linear growth conditions. Let $X(0) = x_0$ and consider the stochastic differential equation with Markovian switching given by

$$dX(t) = f(X(t), Y(t), t)\, dt + \gamma(Y(t))\, dB(t). \tag{3.4.10}$$

Under the above conditions, there is a unique continuous and adapted process which satisfies (3.4.10) (see e.g. [62])

This is an important special case for two related economic reasons. The first is the principal economic rationale for switching models in finance: namely that market sentiment occasionally changes, leading to differing volatility or growth rates. The incorporation of sentiment in this manner is one of the important motivations behind the discipline of behavioural finance (see e.g., the survey paper [25]). Secondly, it makes the volatility a stochastic process which cannot be explained purely in terms of the current market returns. This places the model within the framework of stochastic volatility (SV) models, particularly as the volatility process is stationary and ergodic. One of the first such SV models was presented in [42], and a recent textbook devoted to stochastic volatility models is [26]. A common feature of SV models is that the volatility is described by the stationary solution of a stochastic differential equation driven by a Brownian motion which is correlated with, but not equal to, the Brownian motion that drives the stock price. In our case, although we only have one Brownian motion, we have two sources of randomness in the security (the other being the switching process). This renders the market incomplete, as there are more sources of randomness than tradable securities. In the model proposed here the volatility is also a stationary stochastic process, but unlike processes in SV models, it can assume only finitely many values, does not change from instant to instant, and is also uncorrelated with the Brownian motion which drives the stock price. However, if employed to price options, the model analysed here should lead to both incomplete markets and the presence of volatility smiles. Volatility smiles have been shown to exist for other stochastic volatility models in which the volatility assumes a finite number of values (see e.g., Renault and Touzi [69]).

The first result shows that when the volatility depends on the switching process alone, there is a well–defined growth rate, and the fluctuations around this growth rate still obey a law of the iterated logarithm.

Theorem 3.4.2. *Let S be the unique continuous adapted process governed by* (3.4.5) *with $S(0) = s_0 > 0$, where X is defined by* (3.4.10)*, with $X(0) = 0$ and $\gamma : \mathbb{S} \to \mathbb{R} \setminus \{0\}$. Suppose that f obeys* (3.3.9)*. Then*

(i)

$$\lim_{t\to\infty} \frac{\log S(t)}{t} = \mu - \frac{\sigma_*^2}{2}, \quad a.s.$$

(ii)

$$\limsup_{t\to\infty} \frac{|\log S(t) - (\mu t - \frac{1}{2}\int_0^t \gamma^2(Y(s))\,ds)|}{\sqrt{2t\log\log t}} = \sigma_*, \quad a.s., \tag{3.4.11}$$

where $\sigma_* > 0$ is defined by (3.3.11)*.*

Before proceeding further, we pause to examine the relevance of (3.4.11) and its connection with (3.4.9). The limit in (3.4.11) gives, at least superficially, a weaker result than the limit in (3.4.9). As explained earlier, (3.4.9) can be interpreted in terms of the size of the fluctuations of the price around its *deterministic* exponential rate of growth $G(t) := \exp[(\mu - \sigma^2/2)t]$. Hence the log trend is $\log G(t) = (\mu - \sigma^2/2)t$, so (3.4.9) can be written

$$\limsup_{t\to\infty} \frac{|\log S(t) - \log G(t)|}{\sqrt{2t\log\log t}} = |\sigma|, \quad \text{a.s.}$$

Similarly, (3.4.11) can be written in this form with σ_* being the limit on the right–hand side and the log trend, $\log G_*(t)$, in this case is *stochastic* and given by

$$\log G_*(t) = \mu t - \frac{1}{2}\int_0^t \gamma^2(Y(s))ds. \tag{3.4.12}$$

The fact that G_* is stochastic does not by itself create a difficulty in (3.4.11) but rather the fact that it depends on the switching process Y which cannot be observed directly from market data. Therefore it is certainly more cumbersome, and perhaps infeasable, to remove this stochastic growth trend as easily as in (3.4.9).

However, we now show that it is possible to recover the full strength of (3.4.9) by introducing a *deterministic* log trend $\log G_1(t) := (\mu - \sigma_*^2/2)t$.

Theorem 3.4.3. *Let S be the unique continuous adapted process governed by* (3.4.5) *with $S(0) = s_0 > 0$, where X satisfies* (3.4.10)*, $\gamma : \mathbb{S} \to \mathbb{R}$ and f obeys* (3.3.9)*. Let Y be a stationary jump process with finite, irreducible state space $\mathbb{S}$. Then, using the ergodic theorem for Markov jump processes,*

$$\sigma_* - \frac{1}{2}\sum_{i\in\mathbb{S}}\gamma^2(i)\beta_i \leq \limsup_{t\to\infty}\frac{|\log S(t) - (\mu - \frac{1}{2}\sigma_*^2)t|}{\sqrt{2t\log\log t}} \leq \sigma_* + \frac{1}{2}\sum_{i\in\mathbb{S}}\gamma^2(i)\beta_i, \quad a.s., \tag{3.4.13}$$

where $\beta_i = (\sigma_{M_i} + \pi_i\sigma_{L_i})/\sqrt{m_i} > 0$ is deterministic and σ_* obeys (3.3.11).

With this result we can interpret a large value of σ_* as giving rise to larger fluctuations from the *deterministic* exponential growth trend $\exp[(\mu - \sigma_*^2/2)t]$.

Theorem 3.4.3 above relies on the rate of convergence of the ergodic theorem for Markov jump processes. We now state this ergodic theorem and its associated rate of convergence in the following result.

Proposition 3.4.1. *Let $\mathbb{S}$ be a finite, irreducible state space, let $\gamma : \mathbb{S} \to \mathbb{R}$ and let Y be a stationary jump process. Then by the ergodic theorem for Markov jump processes*

$$\lim_{t\to\infty}\frac{1}{t}\int_0^t \gamma^2(Y(s))\,ds = \sigma_*^2 = \sum_{j\in\mathbb{S}}\gamma^2(j)\pi_j,$$

where π is the stationary probability distribution of Y defined by (1.0.5)*. Moreover,*

$$\limsup_{t\to\infty}\frac{t}{\sqrt{2t\log\log t}}\left|\frac{1}{t}\int_0^t \gamma^2(Y(s))ds - \sigma_*^2\right| \leq \sum_{i\in\mathbb{S}}\gamma^2(i)\beta_i \quad a.s., \tag{3.4.14}$$

where $\beta_i = (\sigma_{M_i} + \pi_i\sigma_{L_i})/\sqrt{m_i}$ is deterministic.

3.4.3 Large fluctuations of δ–returns

In this subsection, we explore further the case when X is given by (3.4.10), in which the diffusion coefficient depends only on the switching process Y. We associate the state

$H \in \mathbb{S}$ with the largest value of the diffusion coefficient, so that

$$\sigma_H = \max_{j \in \mathbb{S}} |\gamma(j)| > 0. \tag{3.4.15}$$

The state H corresponds to the highest volatility state of the market.

We are interested in the large fluctuations of the returns over $\delta > 0$ time units in continuous time

$$R_\delta(t) := \log(S(t)/S(t-\delta)), \quad t \geq \delta. \tag{3.4.16}$$

The increments R_δ closely relate to the increments of Brownian motion because

$$R_\delta(t) = \log(S(t)/S(t-\delta)) = X(t) - X(t-\delta) - \int_{t-\delta}^{t} \{\mu - \frac{1}{2}\gamma^2(Y(s))\}\, ds, \tag{3.4.17}$$

and due to the fact that the integrand above is bounded, the big fluctuations in the increments R_δ will come from the big fluctuations in the increments of X which in turn are caused by the big fluctuations of Brownian increments. Thus we would expect to see fluctuations in R_δ similar to those of Browian motion, i.e. of the order $\sqrt{2\log t}$. This is confirmed in the following theorem.

Theorem 3.4.4. *Let $\delta > 0$. Let f satisfy* (3.3.9) *and* (3.3.3) *and let Y be an irreducible continuous–time Markov jump process with finite state space $\mathbb{S}$. Let X be the unique adapted continuous solution to* (3.4.10) *and let S satisfy* (3.4.5). *Then R_δ, defined by* (3.4.16), *obeys*

$$\limsup_{t\to\infty} \frac{|R_\delta(t)|}{\sqrt{2\log t}} = \sigma_H\sqrt{\delta}, \quad a.s. \tag{3.4.18}$$

This suggests that the "high" volatility periods are entirely responsible for the largest fluctuations in the absolute δ–returns. This phenomenon cannot be observed from (3.3.10) and (3.4.13) which deal with the cumulative returns, which include accumulated contributions from high, moderate and low volatility periods.

3.4.4 Results for a two–state volatility model

We continue our analysis of the case when X is given by (3.4.10), in which the diffusion coefficient depends only on the switching process Y. In this example, Y is a two-state Markov jump process. To capture this in the notation of the previous subsection we let the state space $\mathbb{S} = \{H, L\}$ so the diffusion coefficient can take the values $\gamma(H) = \sigma_H$ or $\gamma(L) = \sigma_L$. This represents a market model where the volatility can be either "high" or "low", with values $\sigma_H > \sigma_L > 0$ respectively. The generator of Y, denoted Γ, is given by

$$\Gamma = \begin{pmatrix} -\gamma_1 & \gamma_1 \\ \gamma_2 & -\gamma_2 \end{pmatrix}$$

where γ_1 is the rate of transition from the high state to the low state, and γ_2 is the transition rate from the low state to the high state. In a typical situation one would have $\gamma_2 < \gamma_1$ so that the process spends more time in the low volatility state in the long run. We give calculations and interpretations in this case and we note that this can easily be generalised to a finite number of volatility levels. However, econometric evidence indicates that a two–state model is very often sufficient.

Recalling Corollary 3.3.1 we have that

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{2t\log\log t}} = \sigma_* \quad \text{a.s., where } \sigma_*^2 = \sum_{j\in\mathbb{S}} \gamma^2(j)\pi_j \tag{3.4.19}$$

and $\pi = (\pi_H, \pi_L)$ can be found by solving $\pi\Gamma = 0$ (or equivalently $-\pi_H\gamma_1 + \pi_L\gamma_2 = 0$) subject to the constraint $\pi_H + \pi_L = 1$. Solving these equations we arrive at

$$\pi_H = \frac{\gamma_2}{\gamma_1 + \gamma_2}, \quad \pi_L = \frac{\gamma_1}{\gamma_1 + \gamma_2}.$$

Thus, σ_*^2 is now simply the weighted average of the different volatility levels

$$\sigma_*^2 = \sigma_H^2 \frac{\gamma_2}{\gamma_1 + \gamma_2} + \sigma_L^2 \frac{\gamma_1}{\gamma_1 + \gamma_2}.$$

As mentioned earlier, if $\gamma_2 < \gamma_1$ then more weight will be placed on the lower volatility regime as more time will be spent in the low volatility state. This means that σ_* will be small and thus the fluctuations of $|X|$, given by (3.4.19), will be relatively small. On the other hand, if π_H is relatively close to unity then σ_* can be quite large, and thus periods in the high volatility regime can have a big impact on the fluctuations. Moreover, if σ_* is large then the growth rate, given by $\mu - \sigma_*^2/2$, is reduced. These important features are somewhat concealed in the statement of (3.4.19).

3.5 Proofs of Results from Section 3.3

Proof of Theorem 3.3.1. Applying Itô's formula to (3.3.1) we get

$$dX^2(t) = \left[2X(t)f(X(t),Y(t),t) + g^2(X(t),Y(t),t)\right] dt \\ + 2X(t)g(X(t),Y(t),t)\, dB(t), \quad t \geq 0. \quad (3.5.1)$$

Let N be the local martingale defined by $N(t) = \int_0^t 2X(s)g(X(s),Y(s),s)\, dB(s)$. It has quadratic variation given by $\langle N\rangle(t) = \int_0^t 4X^2(s)g^2(X(s),Y(s),s)\, ds$. Then by Doob's martingale representation theorem (see Chapter 1 or Theorem 3.4.2 in [46]), there exists another Brownian motion β in an extended probability space with measure $\widetilde{\mathbb{P}}$ such that

$$N(t) = \int_0^t 2|g(X(s),Y(s),s)|\sqrt{X^2(s)}d\beta(s) \quad \widetilde{\mathbb{P}}\text{-a.s.}$$

Now let $Z(t) = X^2(t)$ and let $\phi(t) = 2X(t)f(X(t),Y(t),t) + g^2(X(t),Y(t),t)$ so that we can write equation (3.5.1) as

$$dZ(t) = \phi(t)\, dt + 2|g(X(t),Y(t),t)|\sqrt{Z(t)}\, d\beta(t). \quad (3.5.2)$$

Let $M(t) = \int_0^t |g(X(s),Y(s),s)|\, d\beta(s)$, so $\langle M\rangle(t) = \int_0^t g^2(X(s),Y(s),s)\, ds$. Then by the martingale time–change theorem (cf. e.g., Theorem 3.4.6 in [46]), we may define a new Brownian motion $\tilde{\beta}$ by $\tilde{\beta}(\langle M\rangle(t)) = M(t)$ and the stopping time τ by $\tau(t) = \inf\{s > 0 : \langle M\rangle(s) > t\}$. Since $g^2(x,y,t) \geq K_2 > 0$ it follows that $\langle M\rangle$ is increasing and $\lim_{t\to\infty}\langle M\rangle(t) = \infty$. Thus $\langle M\rangle^{-1}$ exists and in fact $\tau(t) = \langle M\rangle^{-1}(t)$. Moreover, $M(\tau(t)) = \tilde{\beta}(t)$ and we introduce the processes $\tilde{X}(t) = X(\tau(t))$, $\tilde{Y}(t) = Y(\tau(t))$ and $\tilde{Z}(t) = Z(\tau(t))$. So now, applying this time-change to (3.5.2) we get:

$$\tilde{Z}(t) = Z(\tau(t)) = Z(\tau(0)) + \int_0^{\tau(t)} \phi(s)\,ds + \int_0^{\tau(t)} 2|g(X(s),Y(s),s)|\sqrt{Z(s)}\,d\beta(s). \quad (3.5.3)$$

To deal with the stochastic integral above, we use Proposition 3.4.8 from [46], which states that if $\tilde{\eta}(t) = \eta(\tau(t))$ and η is $\mathcal{F}^{\beta}$–adapted, then $\int_0^{\tau(s)} \eta(u)\,dM(u) = \int_0^s \tilde{\eta}(u)\,d\tilde{\beta}(u)$. In this case, we set $\eta(t) = 2\sqrt{Z(t)}$ and set M equal to the martingale defined above. Therefore

$$\int_0^{\tau(t)} 2\sqrt{Z(s)}\,|g(X(s),Y(s),s)|\,d\beta(s) = \int_0^{\tau(t)} 2\sqrt{Z(s)}\,dM(s) = \int_0^t 2\sqrt{\tilde{Z}(s)}\,d\tilde{\beta}(s).$$

To deal with the Riemann integral term in (3.5.3), we use Problem 3.4.5 from [46], which states that if G is a bounded measurable function, and $[a,b] \subset [0,\infty)$ then we have $\int_a^b G(s)\,d\langle M\rangle(s) = \int_{\langle M\rangle(a)}^{\langle M\rangle(b)} G(\tau(s))\,ds$. In this case, we set

$$G(t) = \phi(t)/g^2(X(t),Y(t),t)$$

and as $d\langle M\rangle(t) = g^2(X(t),Y(t),t)\,dt$, we obtain

$$\begin{aligned}\int_0^{\tau(t)} \phi(s)\,ds &= \int_0^{\tau(t)} G(s)\,d\langle M\rangle(s) = \int_{\langle M\rangle(0)}^{\langle M\rangle(\tau(t))} G(\tau(s))\,ds \\ &= \int_0^t \frac{\tilde{\phi}(s)}{g^2(\tilde{X}(s),\tilde{Y}(s),\tau(s))}\,ds,\end{aligned}$$

where $\tilde{\phi}(t) = \phi(\tau(t))$. So we can now write (3.5.3) as:

$$\tilde{Z}(t) = \tilde{Z}(0) + \int_0^t \frac{\tilde{\phi}(s)}{g^2(\tilde{X}(s),\tilde{Y}(s),\tau(s))}\,ds + \int_0^t 2\sqrt{\tilde{Z}(s)}\,d\tilde{\beta}(s). \quad (3.5.4)$$

Now, using conditions (3.3.2) and (3.3.4), it is easy to see that the drift coefficient of (3.5.4) is bounded above by $(K_2 + 2\rho)/K_2$. Define the process which is uniquely determined by the stochastic differential equation

$$dU(t) = C_u\, dt + 2\sqrt{|U(t)|}\, d\tilde{\beta}(t), \quad \text{a.s.} \tag{3.5.5}$$

with $U(0) \geq \tilde{Z}(0) \geq 0$, where $C_u = (K_2 + 2\rho)/K_2$. We will now show, using a stochastic comparison technique, that for all $t \geq 0$, $\tilde{Z}(t) \leq U(t)$ a.s.

First, we must show that $U(t)$ is positive so that we can drop the absolute value in (3.5.5). We apply the stochastic comparison theorem (Theorem 2.2.1) to (3.5.5) and to the equation $dU_1(t) = 2\sqrt{|U_1(t)|}\, d\tilde{\beta}(t)$ with $U_1(0) = 0$; this shows that $U(t) \geq U_1(t)$ a.s., and since the process U_1 has the unique solution $U_1(t) = 0$, it follows that $U(t) \geq 0$ a.s. Therefore $U(t)$ in fact obeys

$$dU(t) = C_u\, dt + 2\sqrt{U(t)}\, d\tilde{\beta}(t), \quad \text{a.s.} \tag{3.5.6}$$

Finally, we can apply the comparison theorem to (3.5.4) and (3.5.6) to conclude that for all $t \geq 0$, $\tilde{Z}(t) \leq U(t)$ a.s. Now we can approximate an upper bound for $\tilde{Z}$ by getting an upper bound for U. However, before we do that we will apply a time-change and a change of scale to U to get a process with finite speed measure, using the techniques mentioned in Subsection 3.2.2. Consider $V(t) = e^{-t}U(e^t - 1)$. By using the product rule and introducing a new Brownian motion $\bar{\beta}$ on an extended space, we can show that

$$dV(t) = [-V(t) + C_u]\, dt + 2\sqrt{V(t)}\, d\bar{\beta}(t). \tag{3.5.7}$$

A scale function of V is given by $p_V(x) = \mu \int_a^x e^{y/2} y^{-C_u/2}\, dy$ for $a > C_u$ and $\mu := e^{-a/2} a^{C_u/2}$. One can check that V satisfies Theorem 1.0.1. Hence Theorem 1.0.2 can be applied to V. Notice that for all $y \geq a$, $y \mapsto e^{y/2} y^{-C_u/2}$ is increasing. By L'Hôpital's rule

$$\lim_{x\to\infty} \frac{p_V(x)}{\mu e^{x/2} x^{-C_u/2}} = \lim_{x\to\infty} \left(\frac{C_u}{2x} + \frac{1}{2}\right)^{-1} = 2,$$

and so there exists an x_0 such that for $x \geq x_0$ we have $p_V(x) \geq \mu e^{x/2} x^{-C_u/2}$. Let $\beta > 1$

and define $h(t) = 2\beta \log t$. Hence for $t \geq e^{x_0/2\beta}$,

$$\frac{1}{p_V(h(t))} \leq \frac{1}{\mu} e^{-\beta \log t}(2\beta \log t)^{\frac{C_u}{2}}$$

and for any $\varepsilon \in (0, \beta - 1)$ there exists t_ε such that for all $t \geq t_\varepsilon \vee e^{x_0/2\beta} =: t^*_{\varepsilon,\beta}$

$$\frac{\log\left(p_V^{-1}(h(t))\right)}{\log t} \leq -\beta + \varepsilon.$$

Thus, we can apply Motoo's theorem since

$$\int_{t^*_{\varepsilon,\beta}}^{\infty} \frac{1}{p_V(h(s))}\, ds \leq \int_{t^*_{\varepsilon,\beta}}^{\infty} \frac{1}{s^{\beta-\varepsilon}}\, ds < +\infty.$$

Therefore $\limsup_{t\to\infty} V(t)/2\log t \leq \beta$ a.s. and letting $\beta \downarrow 1$ through the rational numbers,

$$\limsup_{t\to\infty} \frac{V(t)}{2\log t} \leq 1, \quad \text{a.s.}$$

Using the fact that $V(t) = e^{-t}U(e^t - 1)$, we find that

$$\limsup_{t\to\infty} \frac{U(t)}{2t\log\log t} \leq 1, \quad \text{a.s.}$$

So

$$\limsup_{t\to\infty} \frac{Z(\tau(t))}{2t\log\log t} = \limsup_{t\to\infty} \frac{\tilde{Z}(t)}{2t\log\log t} \leq 1, \quad \text{a.s.}$$

By definition, $\tau(t) = \langle M\rangle^{-1}(t)$ and $\tau(\cdot)$ is monotone, so it follows that

$$\limsup_{t\to\infty} \frac{Z(t)}{2\langle M\rangle(t)\log\log\langle M\rangle(t)} \leq 1, \quad \text{a.s.} \tag{3.5.8}$$

Since $K_2 t \leq \langle M\rangle(t) \leq K_1 t$, $t \geq 0$, we can show that

$$\lim_{t\to\infty} \frac{\log\log\langle M\rangle(t)}{\log\log t} = 1 \quad \text{and} \quad \frac{t}{\langle M\rangle(t)} \geq \frac{t}{K_1 t} = \frac{1}{K_1}, \text{ a.s. for all } t > 0.$$

Therefore (3.5.8) implies $\limsup_{t\to\infty} Z(t)/(2t\log\log t) \leq K_1$ a.s. By taking square roots on both sides we get the assertion. □

Proof of Theorem 3.3.2. Following the same argument as the previous proof, we arrive at (3.5.4). Therefore

$$d\tilde{Z}(t) = \frac{\phi(\tau(t))}{g^2(\tilde{X}(t), \tilde{Y}(t), \tau(t))}\,dt + 2\sqrt{\tilde{Z}(t)}\,d\tilde{\beta}(t).$$

By (3.3.6), it is easy to see that the drift coefficient of the above equation can be bounded below by $C_l := L + \frac{1}{2}$. Consider the process governed by the following equation

$$dU(t) = C_l\,dt + 2\sqrt{|U(t)|}\,d\tilde{\beta}(t)$$

with $U(0) \leq \tilde{Z}(0)$. Then it can be shown, using the same method as used in the previous proof, that for all $t \geq 0$, $\tilde{Z}(t) \geq U(t) \geq 0$. Applying changes in both time and scale again, let $V(t) = e^{-t}U(e^t - 1)$ to get

$$dV(t) = (-V(t) + C_l)\,dt + 2\sqrt{V(t)}\,d\bar{\beta}(t) \quad t \geq 0.$$

We proceed as before; the process V obeys Theorem 1.0.1, and so we may apply Theorem 1.0.2 to it. Since a scale function of V is given by $p_V(x) = \mu\int_a^x e^{\frac{1}{2}y}y^{-C_l/2}\,dy$ for $a > C_l$ and $\mu = e^{-a/2}a^{C_l/2}$, then by L'Hôpital's Rule $\lim_{x\to\infty} p_V(x)/e^{x/2} = 0$. This implies that there exists $x_* > 0$ such that for all $x > x_*$, $p_V(x) < e^{x/2}$. Hence if we let $h(t) = 2\log t$ and $t_* = e^{x_*/2}$, then for all $t > t_*$ we have $p_V(h(t)) < t$ and thus $\int_{t_*}^\infty 1/p_V(h(s))\,ds > \int_{t_*}^\infty 1/s\,ds = \infty$. Therefore, by Motoo's theorem, $\limsup_{t\to\infty} V(t)/2\log t \geq 1$ a.s. Since $V(t) = e^{-t}U(e^t - 1)$, we get $\limsup_{t\to\infty} U(t)/(2t\log\log t) \geq 1$ a.s. Since $\tilde{Z}(t) \geq U(t)$, we get $\limsup_{t\to\infty} \tilde{Z}(t)/(2t\log\log t) \geq 1$ a.s. Hence, as in the previous proof, we have

$$\limsup_{t\to\infty} \frac{Z(t)}{2\langle M\rangle(t)\log\log\langle M\rangle(t)} \geq 1, \quad \text{a.s.} \tag{3.5.9}$$

Proceeding as in the end of the last proof, we get the desired result (3.3.7). □

Proof of Corollary 3.3.1. By (3.5.8) and (3.5.9), as $Z(t) = X^2(t)$, we have

$$\limsup_{t\to\infty} \frac{X^2(t)}{2\langle M\rangle(t)\log\log\langle M\rangle(t)} = 1, \quad \text{a.s.} \tag{3.5.10}$$

By analogy to the proof of Theorem 3.3.1, we have $\langle M\rangle(t) = \int_0^t \gamma^2(Y(s))\,ds$ and since γ is bounded it follows that

$$\lim_{t\to\infty} \frac{\log\log t}{\log\log \langle M\rangle(t)} = 1.$$

By Proposition 3.4.1 it follows that $\lim_{t\to\infty}\langle M\rangle(t)/t = \sigma_*^2$ a.s., which together with (3.5.10) gives

$$1 = \limsup_{t\to\infty} \frac{X^2(t)}{2t\log\log t}.\frac{t}{\langle M\rangle(t)}.\frac{\log\log t}{\log\log\langle M\rangle t} = \limsup_{t\to\infty} \frac{X^2(t)}{2t\log\log t}.\frac{1}{\sigma_*^2}, \quad \text{a.s.}$$

Taking square roots yields the desired result. □

Proof of Theorem 3.3.3. For all $t \geq 0$

$$X(t) = x_0 + \int_0^t f(X(s), Y(s), s)ds + \int_0^t g(X(s), Y(s), s)\,dB(s).$$

Let $M_1(t) = \int_0^t g(X(s), Y(s), s)\,dB(s)$, so $\langle M_1\rangle(t) = \int_0^t g^2(X(s), Y(s), s)\,ds$. Hence for all $t \geq 0$, $K_2 t \leq \langle M_1\rangle(t) \leq K_1 t$ and $\lim_{t\to\infty}\langle M_1\rangle(t) = \infty$ almost surely. Moreover $\langle M_1\rangle$ is increasing on $(0,\infty)$ and admits an inverse. Again we use the time-change theorem for martingales: for each $0 \leq t < \infty$, define the stopping time $\lambda(t) := \inf\{s > 0 : \langle M_1\rangle(s) > t\}$. Thus $\langle M_1\rangle(\lambda(t)) = t$ and $\lambda(t) = \langle M_1\rangle^{-1}(t)$. A process defined by $W(t) := M(\lambda(t)), \forall\, t \geq 0$ is a standard Brownian motion with respect to the filtration $\mathcal{G}(t) := \mathcal{F}(\lambda(t))$. Therefore, as in the proof of Theorem 3.3.1, we get

$$\begin{aligned}\widetilde{X}(t) := X(\lambda(t)) &= x_0 + \int_0^{\lambda(t)} f(X(s), Y(s), s)ds + \int_0^{\lambda(t)} g(X(s), Y(s), s)dB(s)\\ &= x_0 + \int_0^t \frac{f(\widetilde{X}(s), \widetilde{Y}(s), \lambda(s))}{g^2(\widetilde{X}(s), \widetilde{Y}(s), \lambda(s))}ds + W(t)\end{aligned}$$

where $\widetilde{Y}(t) := Y(\lambda(t))$. Due to (3.3.12), we have

$$\forall\,(x, y, t) \in \mathbb{R}\times\mathbb{S}\times[0,\infty), \quad -\tilde{f}(x) \leq \frac{f(x,y,t)}{g^2(x,y,t)} \leq \tilde{f}(x).$$

Consider two processes Z_1 and Z_2 governed by the following two equations, for $t \geq 0$

$$dZ_1(t) = \tilde{f}(Z_1(t))\,dt + dW(t), \quad dZ_2(t) = -\tilde{f}(Z_2(t))\,dt + dW(t)$$

with $Z_2(0) \le x_0 \le Z_1(0)$. Then again by the stochastic comparison theorem, we can show that for all $t \ge 0$, $Z_2(t) \le \widetilde{X}(t) \le Z_1(t)$ a.s. Consider the scale function of Z_1 defined as the following

$$p_{Z_1}(x) = \int_0^x e^{-2\int_0^y \tilde{f}(z)dz} dy, \quad x \in \mathbb{R}.$$

Then $p_{Z_1} \in C^2(\mathbb{R};\mathbb{R})$ and for all $x \in \mathbb{R}$, we have

$$p'_{Z_1}(x)\tilde{f}(x) + \frac{1}{2}p''_{Z_1}(x) = 0. \tag{3.5.11}$$

This second order differential equation has solution $p'_{Z_1}(x) = \exp[-2\int_0^x \tilde{f}(s)ds]$. Since $\tilde{f} \in \mathcal{L}^1$, there exist real numbers k_1, k_2 such that $\int_0^\infty \tilde{f}(z)dz = k_1$ and $\int_{-\infty}^0 \tilde{f}(z)dz = k_2$, which implies $\lim_{x\to\infty} p'_{Z_1}(x) = e^{-2k_1}$ and $\lim_{x\to-\infty} p'_{Z_1}(x) = e^{2k_2}$. Moreover, by L'Hôpital's Rule,

$$\lim_{x\to\infty} \frac{p_{Z_1}(x)}{x} = e^{-2k_1}, \quad \lim_{x\to-\infty} \frac{p_{Z_1}(x)}{x} = e^{2k_2}. \tag{3.5.12}$$

So $p_{Z_1}(\infty) = \infty$ and $p_{Z_1}(-\infty) = -\infty$ and by Proposition 5.5.22 in [46], the process Z_1 is recurrent in the sense that $\limsup_{t\to\infty} Z_1(t) = \infty$ and $\liminf_{t\to\infty} Z_1(t) = -\infty$ a.s. Let $H(t) = p_{Z_1}(Z_1(t))$. Then by Itô's Rule and (3.5.11)

$$dH(t) = p'_{Z_1}(Z_1(t))dW(t), \quad t \ge 0,$$

with $H(0) = p_{Z_1}(Z_1(0))$. This technique is known as the method of removal of drift and can be found in Chapter 5 of [46]. Now since p_{Z_1} is strictly increasing, the above equation can be written as

$$dH(t) = l(H(t))dW(t), \quad t \ge 0,$$

where $l(x) = p'_{Z_1}(p^{-1}_{Z_1}(x))$, for all $x \in \mathbb{R}$. H is also a recurrent process on $\mathbb{R}$. Moreover, (3.5.12) gives

$$\lim_{t\to\infty} \frac{\sup_{0\le s\le t} H(s)}{\sup_{0\le s\le t} Z_1(s)} = e^{-2k_1} \quad \text{and} \quad \lim_{t\to\infty} \frac{\inf_{0\le s\le t} H(s)}{\inf_{0\le s\le t} Z_1(s)} = e^{2k_2}, \quad \text{a.s.} \tag{3.5.13}$$

For each $t \geq 0$, define the continuous local martingale Q given by

$$Q(t) := \int_0^t l(H(s))\, dW(s),$$

which has quadratic variation $\langle Q\rangle(t) := \int_0^t l^2(H(s))\, ds$. Thus $\langle Q\rangle'(t) > 0$ for $t > 0$ and $\langle Q\rangle$ is an increasing function. Now

$$\inf_{x\in\mathbb{R}} l^2(x) = \inf_{x\in\mathbb{R}} p_Z'(p_Z^{-1}(x))^2 = \inf_{x\in\mathbb{R}} e^{-4\int_0^{p_Z^{-1}(x)} \tilde{f}(z)dz} = e^{-4\sup_{x\in\mathbb{R}}\int_0^x \tilde{f}(z)dz} > 0.$$

Similarly, $\sup_{x\in\mathbb{R}} l^2(x) = e^{-4\inf_{x\in\mathbb{R}}\int_0^x \tilde{f}(z)dz} < \infty$. Let $l_1^2 = \inf_{x\in\mathbb{R}} l^2(x)$ and let $l_2^2 = \sup_{x\in\mathbb{R}} l^2(x)$, so for all $t \geq 0$,

$$l_1^2 t \leq \langle Q\rangle(t) \leq l_2^2 t, \tag{3.5.14}$$

which implies $\lim_{t\to\infty}\langle Q\rangle(t) = \infty$. Now define, for each $0 \leq s < \infty$, the stopping time $\kappa(s) = \inf\{t \geq 0; \langle Q\rangle(t) > s\}$. It is obvious that κ is continuous and tends to infinity almost surely. Furthermore $\langle Q\rangle(\kappa(t)) = t$, and $\kappa^{-1}(t) = \langle Q\rangle(t)$ for $t \geq 0$. Then the time-changed process $\widetilde{W}(t) := Q(\kappa(t))$ is a standard one-dimensional Brownian motion with respect to the filtration $\mathcal{J}(t) := \mathcal{G}(\kappa(t))$. Hence we have

$$\widetilde{H}(t) := H(\kappa(t)) = H(\kappa(0)) + \int_0^{\kappa(t)} l(H(s))dW(s) = \widetilde{H}(0) + \widetilde{W}(t)$$

where $\widetilde{H}$ is $\mathcal{J}(t)$-adapted. So the law of the iterated logarithm holds for $\widetilde{H}$, that is

$$1 = \limsup_{t\to\infty} \frac{H(\kappa(t))}{\sqrt{2t\log\log t}} = \limsup_{t\to\infty} \frac{H(t)}{\sqrt{2\langle Q\rangle(t)\log\log\langle Q\rangle(t)}}, \quad \text{a.s.}$$

Note by (3.5.14) for all $t \geq 0$, that $\log l_1^2 + \log t \leq \log\langle Q\rangle(t) \leq \log l_2^2 + \log t$, so we have

$$\lim_{t\to\infty} \frac{\log\log\langle Q\rangle(t)}{\log\log t} = 1, \quad \text{a.s.}$$

which implies

$$\limsup_{t\to\infty} \frac{H(t)}{\sqrt{2\langle Q\rangle(t)\log\log t}} = 1, \quad \text{a.s.}$$

Similarly, by the law of the iterated logarithm,

$$\liminf_{t\to\infty} \frac{H(t)}{\sqrt{2\langle Q\rangle(t)\log\log t}} = -1, \quad \text{a.s.}$$

Now as $\langle Q\rangle(t) \le l_2^2 t$, we have

$$\limsup_{t\to\infty} \frac{H(t)}{\sqrt{2t\log\log t}} = \limsup_{t\to\infty} \sqrt{\frac{\langle Q\rangle(t)}{t}} \cdot \frac{H(t)}{\sqrt{2\langle Q\rangle(t)\log\log t}} \le l_2, \quad \text{a.s.,}$$

and similarly

$$\liminf_{t\to\infty} \frac{H(t)}{\sqrt{2t\log\log t}} \le -l_1, \quad \text{a.s.}$$

Combining the above results with (3.5.13), we get

$$\limsup_{t\to\infty} \frac{Z_1(t)}{\sqrt{2t\log\log t}} \le e^{2k_1} l_2, \quad \text{and} \quad \liminf_{t\to\infty} \frac{Z_1(t)}{\sqrt{2t\log\log t}} \le -e^{-2k_2} l_1, \quad \text{a.s.}$$

which implies

$$\begin{aligned}\limsup_{t\to\infty} \frac{X(\lambda(t))}{\sqrt{2t\log\log t}} &= \limsup_{t\to\infty} \frac{\widetilde{X}(t)}{\sqrt{2t\log\log t}} \le \limsup_{t\to\infty} \frac{Z_1(t)}{\sqrt{2t\log\log t}} \\ &\le \frac{e^{-2\inf_{x\in\mathbb{R}}\int_0^x \tilde{f}(y)\,dy}}{e^{-2\int_0^\infty \tilde{f}(y)\,dy}}, \quad \text{a.s.,}\end{aligned}$$

and

$$\liminf_{t\to\infty} \frac{X(\lambda(t))}{\sqrt{2t\log\log t}} \le \frac{-e^{-2\sup_{x\in\mathbb{R}}\int_0^x \tilde{f}(y)dy}}{e^{2\int_{-\infty}^0 \tilde{f}(y)dy}}, \quad \text{a.s.}$$

By an analogous argument to that given in the proof of Theorem 3.3.1, we get

$$\begin{aligned}\limsup_{t\to\infty} \frac{X(t)}{\sqrt{2t\log\log t}} &\le \frac{\sqrt{K_1}e^{-2\inf_{x\in\mathbb{R}}\int_0^x \tilde{f}(y)\,dy}}{e^{-2\int_0^\infty \tilde{f}(y)dy}}, \quad \text{a.s.,} \\ \liminf_{t\to\infty} \frac{X(t)}{\sqrt{2t\log\log t}} &\le \frac{-\sqrt{K_2}e^{-2\sup_{x\in\mathbb{R}}\int_0^x \tilde{f}(y)dy}}{e^{2\int_{-\infty}^0 \tilde{f}(y)dy}}, \quad \text{a.s.}\end{aligned}$$

By considering Z_2 in a similar manner, we deduce the lower estimates on lim sup and lim inf of X in (3.3.13). □

3.6 Proofs of Results from Section 3.4

Proof of Theorem 3.4.1. Combining (3.4.8) and (3.4.5), we have

$$dS(t) = [\mu S(t) + f(X(t), Y(t), t)S(t)]\, dt + \sigma S(t)\, dB(t) \quad t \geq 0.$$

Thus $S(t) = s_0 e^{(\mu - \frac{\sigma^2}{2})t + X(t)}$, $t \geq 0$, which implies $\log S(t)/t = \log s_0/t + \mu - \sigma^2/2 + X(t)/t$. Now by Corollary 3.3.1, we have $\lim_{t\to\infty} X(t)/t = 0$, a.s. Therefore by letting $t \to \infty$, the first part of the conclusion is obtained. Since $X(t) = \log S(t) - \log s_0 - (\mu - \frac{\sigma^2}{2})t$, $t \geq 0$, by applying Corollary 3.3.1 in the simple case in which $\gamma(y) = \sigma$ for all $y \in \mathbb{S}$, we get the second part of the conclusion. For the third part, we observe that the assertion is equivalent to

$$\limsup_{t\to\infty} \frac{|X_\delta(t)|}{\sqrt{2\log t}} = |\sigma|\sqrt{\delta}, \quad \text{a.s.}$$

where $X_\delta(t) = \int_{t-\delta}^{t} f(X(s), Y(s), s)\, ds + \sigma(B(t) - B(t-\delta))$. Now since for all $(x, y, t) \in \mathbb{R} \times \mathbb{S} \times \mathbb{R}^+$ we have $-\rho\sigma^2/|x| < f(x, y, t) < \rho\sigma^2/|x|$ by (3.3.9), then for any $y \in \mathbb{S}$ we have $\lim_{|x|\to\infty} f(x, y, t) = 0$. Also, because f is globally bounded by (3.3.3), we have

$$\lim_{t\to\infty} \frac{\int_{t-\delta}^{t} f(X(s), Y(s), s) ds}{\sqrt{2\log t}} = 0, \quad \text{a.s.}$$

Hence it remains to show that

$$\limsup_{t\to\infty} \frac{|B(t) - B(t-\delta)|}{\sqrt{2\log t}} = \sqrt{\delta}, \quad \text{a.s.}$$

Consider $Z_\delta(n) := (B(n\delta) - B((n-1)\delta))/\sqrt{\delta}$, $n \in \mathbb{N}$. Then $\{Z_\delta(n)\}_{n\in\mathbb{N}}$ is a sequence of independent standard normal random variables. Thus, by Lemma 1.0.1

$$\limsup_{n\to\infty} \frac{|Z_\delta(n)|}{\sqrt{2\log n}} = 1, \quad \text{a.s.} \tag{3.6.1}$$

It immediately follows that

$$\limsup_{t\to\infty} \frac{|B(t) - B(t-\delta)|}{\sqrt{2\log t}} \geq \limsup_{n\to\infty} \frac{|B(n\delta) - B((n-1)\delta)|}{\sqrt{2\log n}} = \sqrt{\delta}, \quad \text{a.s.} \tag{3.6.2}$$

For the upper estimate, by the triangle inequality

$$|B(t)-B(t-\delta)| \le |B(t)-B(n^{1-\varepsilon})|+|B(t-\delta)-B(n^{1-\varepsilon}-\delta)|+|B(n^{1-\varepsilon})-B(n^{1-\varepsilon}-\delta)| \tag{3.6.3}$$

where $\varepsilon \in (0,1)$. We now consider the first term on the right-hand side of the above inequality. By properties of Brownian motions,

$$\begin{aligned}\mathbb{P}[\sup_{n^{1-\varepsilon}\le t\le (n+1)^{1-\varepsilon}} |B(t)-B(n^{1-\varepsilon})|>1] &= 2\mathbb{P}[\sup_{0\le t\le (1-\varepsilon)\hat{n}^{-\varepsilon}} B(t)>1] \\ &= 4\mathbb{P}[B((1-\varepsilon)\hat{n}^{-\varepsilon})>1] = 4\Big(1-\Phi(\frac{1}{\sqrt{(1-\varepsilon)\hat{n}^{-\varepsilon}}})\Big),\end{aligned}$$

where $\hat{n}\in[n,n+1]$. Again by Mill's estimate and the Borel-Cantelli lemma, we have

$$\limsup_{n\to\infty} \max_{t\in[n^{1-\varepsilon},(n+1)^{1-\varepsilon}]} |B(t)-B(n^{1-\varepsilon})| \le 1 \quad \text{a.s., and} \tag{3.6.4}$$

$$\limsup_{n\to\infty} \max_{t\in[n^{1-\varepsilon},(n+1)^{1-\varepsilon}]} |B(t-\delta)-B(n^{1-\varepsilon}-\delta)| \le 1, \quad \text{a.s.} \tag{3.6.5}$$

Again it can be shown using Lemma 1.0.1 that

$$\limsup_{n\to\infty} \frac{|B(n^{1-\varepsilon})-B(n^{1-\varepsilon}-\delta)|}{\sqrt{2\log n}} \le \sqrt{\delta}, \quad \text{a.s.} \tag{3.6.6}$$

Therefore, combining the results from (3.6.3) to (3.6.6), for almost all $\omega\in\Omega$, if $n^{1-\varepsilon}\le t\le (n+1)^{1-\varepsilon}$, then for n sufficiently large

$$\begin{aligned}\frac{|B(t)-B(t-\delta)|}{\sqrt{2\log t}} \le & \frac{1}{\sqrt{2(1-\varepsilon)\log n}}\Big[|B(t)-B(n^{1-\varepsilon})| \\ & + |B(t-\delta)-B(n^{1-\varepsilon}-\delta)| + |B(n^{1-\varepsilon})-B(n^{1-\varepsilon}-\delta)|\Big]\end{aligned}$$

which implies $\limsup_{t\to\infty}|B(t)-B(t-\delta)|/(\sqrt{2\log t}) \le \sqrt{\delta}/\sqrt{1-\varepsilon}$ a.s. Finally, letting $\varepsilon\to 0$ through the rational numbers, we obtain

$$\limsup_{t\to\infty} \frac{|B(t)-B(t-\delta)|}{\sqrt{2\log t}} \le \sqrt{\delta}, \quad \text{a.s.} \tag{3.6.7}$$

Combining this with (3.6.2) completes the proof. □

Proof of Theorem 3.4.2. To show the statements in part (i), we observe that

$$\log S(t) = \log S(0) + \mu t - \int_0^t \frac{1}{2}\gamma^2(Y(s))ds + X(t).$$

which implies

$$\frac{\log S(t)}{t} = \frac{\log S(0)}{t} + \mu - \frac{1}{2t}\int_0^t \gamma^2(Y(s))ds + \frac{X(t)}{t}.$$

Now by Corollary 3.3.1, we have $\lim_{t\to\infty} X(t)/t = 0$, a.s. while by the ergodic property of the Markov chain,

$$\lim_{t\to\infty} \frac{1}{2t}\int_0^t \gamma^2(Y(s))ds = \frac{\sigma_*^2}{2} \quad \text{a.s.}$$

Therefore by letting $t \to \infty$, the first assertion in part (i) is obtained. Since

$$\log S(t) - \Big(\mu t - \frac{1}{2}\int_0^t \gamma^2(Y(s))ds\Big) = \log S(0) + X(t),$$

also by Corollary 3.3.1, we get the second assertion in part (ii). □

Proof of Theorem 3.4.3. This proof follows from Proposition 3.4.1 and (3.4.11). We have

$$\begin{aligned}
\sigma_* &= \limsup_{t\to\infty} \frac{|\log S(t) - \mu t + \frac{1}{2}\sigma_*^2 t + \frac{1}{2}\int_0^t \gamma^2(Y(s))ds - \frac{1}{2}\sigma_*^2 t|}{\sqrt{2t\log\log t}} \\
&\le \limsup_{t\to\infty} \frac{|\log S(t) - (\mu - \frac{1}{2}\sigma_*^2)t|}{\sqrt{2t\log\log t}} + \limsup_{t\to\infty} \frac{\frac{1}{2}t|\frac{1}{t}\int_0^t \gamma^2(Y(s))ds - \sigma_*^2|}{\sqrt{2t\log\log t}} \\
&\le \limsup_{t\to\infty} \frac{|\log S(t) - (\mu - \frac{1}{2}\sigma_*^2)t|}{\sqrt{2t\log\log t}} + \frac{1}{2}\sum_{i\in\mathbb{S}} \gamma^2(i)\beta_i.
\end{aligned}$$

Therefore we get one part of the assertion,

$$\limsup_{t\to\infty} \frac{|\log S(t) - (\mu - \frac{1}{2}\sigma_*^2)t|}{\sqrt{2t\log\log t}} \ge \sigma_* - \frac{1}{2}\sum_{i\in\mathbb{S}} \gamma^2(i)\beta_i, \quad \text{a.s.} \tag{3.6.8}$$

Similarly, for the second part of the assertion we have

$$\log S(t) - (\mu - \frac{1}{2}\sigma_*^2)t = \log S(t) - \big(\mu t - \frac{1}{2}\int_0^t \gamma^2(Y(s))ds\big) + \frac{1}{2}\big(\sigma_*^2 t - \int_0^t \gamma^2(Y(s))ds\big).$$

Thus, by (3.4.11) and (3.4.14) we get

$$\limsup_{t\to\infty} \frac{|\log S(t) - (\mu - \frac{1}{2}\sigma_*^2)t|}{\sqrt{2t\log\log t}} \le \sigma_* + \frac{1}{2}\sum_{i\in\mathbb{S}} \gamma^2(i)\beta_i, \quad \text{a.s.} \tag{3.6.9}$$

Combining (3.6.8) and (3.6.9) gives the desired conclusion (3.4.13). □

Before we prove Proposition 3.4.1, we state and prove the following lemma.

Lemma 3.6.1. *If $a(t), V(t)$ and $U(t)$, $t \ge 0$, are three continuous processes such that $V(t) \le a(t) \le U(t)$ where*

$$\limsup_{t\to\infty} |V(t)| \le v \quad \textit{and} \quad \limsup_{t\to\infty} |U(t)| \le u,$$

then

$$\limsup_{t\to\infty} |a(t)| \le \max(v, u).$$

Proof. If we have $V(t) \le a(t) \le U(t)$, then $a^2(t) \le \max(U^2(t), V^2(t))$ and $|a(t)| \le \max(|U(t)|, |V(t)|)$. Thus

$$\begin{aligned}\limsup_{t\to\infty} |a(t)| &\le \limsup_{t\to\infty} \big[\max(|U(t)|, |V(t)|)\big] \\ &\le \max\big(\limsup_{t\to\infty} |U(t)|, \limsup_{t\to\infty} |V(t)|\big) \le \max(v, u),\end{aligned}$$

as required. □

Proof of Proposition 3.4.1. The first part of this proof is modelled on a similar proof in [67]. Suppose that Y is recurrent and fix a state i. Then $\big(Y(t)\big)_{t\ge 0}$ hits i with probability 1 and the long–run proportion of time in i equals the long–run proportion of time in i after first hitting i. In other words, without loss of generality, we start in state i.

Denote by $M_i(n)$ the length of the n^{th} visit to i, by $T_i(n)$ the time of the n^{th} return to i and by $L_i(n)$ the length of the n^{th} excursion to i. Thus for $n = 0, 1, 2, \ldots$, setting

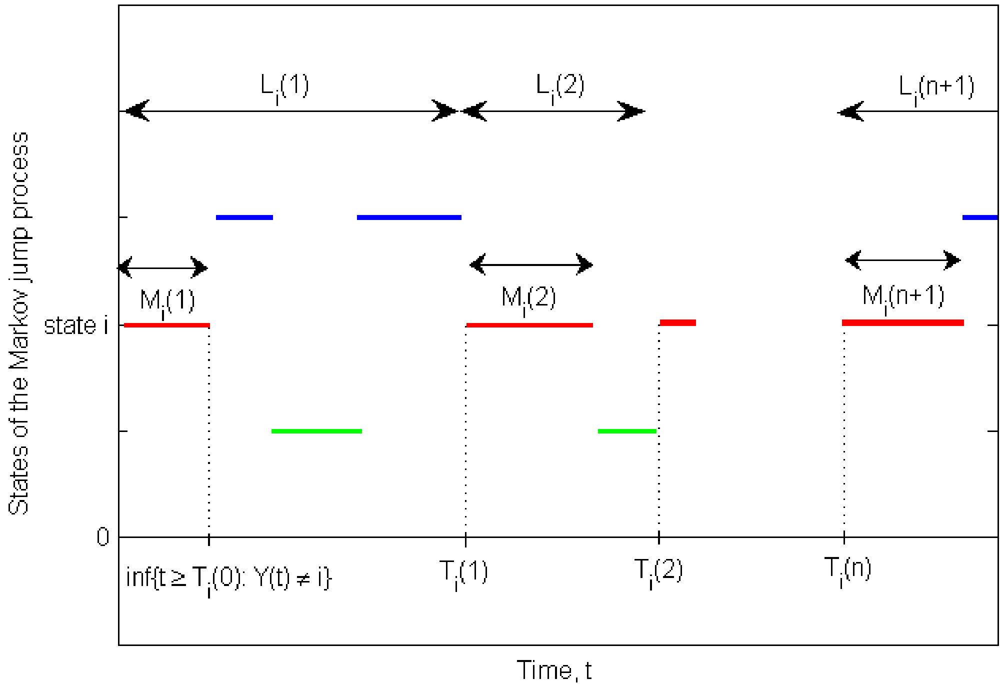
$T_i(0) = 0$, we have

$$M_i(n+1) = \inf\{t > T_i(n) : Y(t) \neq i\} - T_i(n),$$

$$T_i(n+1) = \inf\{t > T_i(n) + M_i(n+1) : Y(t) = i\},$$

$$L_i(n+1) = T_i(n+1) - T_i(n).$$

Figure 3.1: Excursions of a Markov jump process



In summary, the length of the $(n+1)^{\text{th}}$ visit is the difference between the time of the n^{th} return and the time of next exit from i. Then the time of the $(n+1)^{\text{th}}$ return to i is the next time that $Y(t) = i$ which must at least be greater than the time of the previous return and its holding time. Finally, the length of the excursions from i is the distance between two consecutive hitting times of i. Moreover, the time of the n^{th} return to i is equal to the sum of the n excursions to i since $L_i(1) + \cdots + L_i(n) = T_i(n) - T_i(0) = T_i(n)$.

By the strong Markov property (of the jump process) at the stopping times $T_i(n)$ for $n \geq 0$ we find that $L_i(1), L_i(2), \ldots$ are independent and identically distributed with mean m_i, and that $M_i(1), M_i(2), \ldots$ are independent and identically distributed with mean $1/q_i$. Here $q_i = \sum_{j \neq i} \gamma_{ij}$ is the off–diagonal row sum of the transition matrix. Hence, by the strong law of large numbers

$$\frac{L_i(1) + \cdots + L_i(n)}{n} \to m_i \quad \text{and} \quad \frac{M_i(1) + \cdots + M_i(n)}{n} \to \frac{1}{q_i} \quad \text{as } n \to \infty.$$

Therefore, with probability 1,

$$\frac{M_i(1) + \cdots + M_i(n)}{L_i(1) + \cdots + L_i(n)} \to \frac{1}{m_i q_i} \quad \text{as } n \to \infty.$$

Moreover, we note that $T_i(n)/T_i(n+1) \to 1$ a.s. as $n \to \infty$. Now consider $\int_0^t 1_{\{Y(s)=i\}} ds$ for $T_i(n) \leq t < T_i(n+1)$. Since $T_i(n) > \inf\{t > T_i(n-1) : Y(t) \neq i\}$ we have

$$\begin{aligned}\int_0^t 1_{\{Y(s)=i\}} ds &\geq \int_{T_i(0)}^{\inf\{t > T_i(0) : Y(s) \neq i\}} 1 \, ds + \cdots + \int_{T_i(n-1)}^{\inf\{t > T_i(n-1) : Y(s) \neq i\}} 1 \, ds \\ &= M_i(1) + M_i(2) + \cdots + M_i(n).\end{aligned}$$

Similarly we can show that

$$\int_0^t 1_{\{Y(s)=i\}} ds \leq M_i(1) + M_i(2) + \cdots + M_i(n+1).$$

Combining these estimates along with the fact that $1/T_i(n+1) \leq 1/t \leq 1/T_i(n)$, gives

$$\frac{M_i(1) + \cdots + M_i(n)}{T_i(n+1)} \leq \frac{1}{t} \int_0^t 1_{\{Y(s)=i\}} ds \leq \frac{M_i(1) + \cdots + M_i(n+1)}{T_i(n)}. \tag{3.6.10}$$

Now we multiply above and below by $T_i(n)$ on the left–hand side and $T_i(n+1)$ on the right–hand side and use the fact that $T_i(n) = L_i(1) + \cdots + L_i(n)$ to get

$$\begin{aligned}\frac{T_i(n)}{T_i(n+1)} \frac{M_i(1) + \cdots + M_i(n)}{L_i(1) + \cdots + L_i(n)} &\leq \frac{1}{t} \int_0^t 1_{\{Y(s)=i\}} ds \\ &\leq \frac{T_i(n+1)}{T_i(n)} \frac{M_i(1) + \cdots + M_i(n+1)}{L_i(1) + \cdots + L_i(n+1)}\end{aligned}$$

so on letting $t \to \infty$ (and thus $n \to \infty$) we have, with probability 1

$$\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds \to \frac{1}{m_i q_i}. \tag{3.6.11}$$

Since Y is irreducible and the state space $\mathbb{S}$ is finite, we are in the positive recurrent case and we can write

$$\begin{aligned}\Big|\frac{1}{t}\int_0^t \gamma^2(Y(s))ds - \sigma_*^2\Big| &= \Big|\sum_{i\in\mathbb{S}}\frac{1}{t}\gamma^2(i)\int_0^t 1_{\{Y(s)=i\}}ds - \sum_{i\in\mathbb{S}}\pi_i\gamma^2(i)\Big| \\ &= \Big|\sum_{i\in\mathbb{S}}\gamma^2(i)\Big(\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds - \pi_i\Big)\Big| \\ &\le \max_{j\in\mathbb{S}}\gamma^2(j)\sum_{i\in\mathbb{S}}\Big|\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds - \pi_i\Big|\end{aligned} \tag{3.6.12}$$

where $\pi_i = 1/(m_i q_i)$. By (3.6.11), for all $\varepsilon > 0$ there exists $T = T(\omega)$ sufficiently large such that for $t \ge T(\omega)$

$$\sum_{i\in\mathbb{S}}\Big|\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds - \pi_i\Big| \le \varepsilon/\max_{j\in\mathbb{S}}\gamma^2(j)$$

and thus we have, for $t \ge T(\omega)$,

$$\Big|\frac{1}{t}\int_0^t \gamma^2(Y(s))ds - \sigma_*^2\Big| < \varepsilon,$$

which establishes the desired convergence. To prove the second assertion we must determine the rate of this convergence, which is ultimately determined by the rate of convergence of $t^{-1}\int_0^t 1_{\{Y(s)=i\}}ds$ to $1/(m_i q_i)$. For each $t \ge 0$ there exists $n = n(t) \in \mathbb{N}$ such that $T_i(n) \le t < T_i(n+1)$ and $n(t) \to \infty$ as $t \to \infty$ a.s. So by (3.6.10),

$$\begin{aligned}\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds - \frac{1}{m_i q_i} &\le \frac{\sum_{k=1}^{n+1} M_i(k)}{T_i(n)} - \frac{1}{m_i q_i} \\ &= \frac{\sum_{k=1}^{n+1}(M_i(k) - 1/q_i)}{T_i(n)} - \frac{1}{q_i}\Big[\frac{1}{m_i} - \frac{n+1}{T_i(n)}\Big].\end{aligned} \tag{3.6.13}$$

Also, using the fact that $T_i(n) = \sum_{k=1}^{n} L_i(k)$ we have

$$\frac{1}{m_i} - \frac{n+1}{T_i(n)} = \frac{T_i(n) - (n+1)m_i}{m_i T_i(n)} = \frac{\sum_{k=1}^{n}(L_i(k) - m_i)}{m_i T_i(n)} - \frac{1}{T_i(n)}.$$

Combining this with (3.6.13) we see that

$$\frac{t}{\sqrt{2t\log\log t}}\Big(\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds - \frac{1}{m_iq_i}\Big) \le \frac{\sum_{k=1}^{n+1}(M_i(k)-1/q_i)}{\sqrt{2t\log\log t}}\frac{t}{T_i(n)}$$
$$-\frac{1}{q_im_i}\frac{\sum_{k=1}^{n}(L_i(k)-m_i)}{\sqrt{2t\log\log t}}\frac{t}{T_i(n)} + \frac{1}{q_i}\frac{t}{T_i(n)}\frac{1}{\sqrt{2t\log\log t}} =: U_i(t).$$

By (3.2.3) and (3.2.5), the random variables $M_i(k)-1/q_i$ and $L_i(k)-m_i$ are both independent and identically distributed with zero means and finite variances $\sigma^2_{M_i}$ and $\sigma^2_{L_i}$ respectively. Since $n:[0,\infty)\to\mathbb{N}$ is surjective we have by the law of the iterated logarithm,

$$\limsup_{t\to\infty}\frac{|\sum_{k=1}^{n(t)+1}(M_i(k)-1/q_i)|}{\sqrt{2n(t)\log\log n(t)}} \le \sigma_{M_i}, \quad \limsup_{t\to\infty}\frac{|\sum_{k=1}^{n(t)}(L_i(k)-m_i)|}{\sqrt{2n(t)\log\log n(t)}} \le \sigma_{L_i}, \text{ a.s.}$$

Thus, since $n(t)\to\infty$ and $n(t)/t\to 1/m_i$ as $t\to\infty$, these sequences obey

$$\limsup_{t\to\infty}\frac{|\sum_{k=1}^{n(t)+1}(M_i(k)-1/q_i)|}{\sqrt{2t\log\log t}} \le \frac{\sigma_{M_i}}{\sqrt{m_i}}, \quad \limsup_{t\to\infty}\frac{|\sum_{k=1}^{n(t)}(L_i(k)-m_i)|}{\sqrt{2t\log\log t}} \le \frac{\sigma_{L_i}}{\sqrt{m_i}}, \tag{3.6.14}$$

with probability one. So, using $T_i(n)\le t<T_i(n+1)$ and $T_i(n+1)/T_i(n)\to 1$ as $t\to\infty$, we get

$$\limsup_{t\to\infty}|U_i(t)| \le \frac{\sigma_{M_i}}{\sqrt{m_i}} + \frac{1}{q_im_i}\frac{\sigma_{L_i}}{\sqrt{m_i}} = \frac{1}{\sqrt{m_i}}\big(\sigma_{M_i}+\pi_i\sigma_{L_i}\big).$$

Similarly, by (3.6.10) we get the lower bound

$$\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds - \frac{1}{m_iq_i} \ge \frac{\sum_{k=1}^{n}M_i(k)}{T_i(n+1)} - \frac{1}{m_iq_i}$$
$$= \frac{\sum_{k=1}^{n}(M_i(k)-1/q_i)}{T_i(n+1)} - \frac{1}{q_i}\Big[\frac{1}{m_i} - \frac{n}{T_i(n+1)}\Big]. \tag{3.6.15}$$

Also, using the fact that $T_i(n+1)=\sum_{k=1}^{n+1}L_i(k)$ we have

$$\frac{1}{m_i} - \frac{n}{T_i(n+1)} = \frac{T_i(n+1)-nm_i}{m_iT_i(n+1)} = \frac{\sum_{k=1}^{n+1}(L_i(k)-m_i)}{m_iT_i(n+1)} + \frac{1}{T_i(n+1)}$$

Combining this with (3.6.15) we see that

$$\frac{t}{\sqrt{2t\log\log t}}\Big\{\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds - \frac{1}{m_i q_i}\Big\} \geq \frac{\sum_{k=1}^{n}(M_i(k)-1/q_i)}{\sqrt{2t\log\log t}}\frac{t}{T_i(n+1)}$$
$$-\frac{1}{q_i m_i}\frac{\sum_{k=1}^{n+1}(L_i(k)-m_i)}{\sqrt{2t\log\log t}}\frac{t}{T_i(n+1)} - \frac{1}{q_i}\frac{t}{T_i(n+1)}\frac{1}{\sqrt{2t\log\log t}} =: V_i(t).$$

Again the random variables $M_i(k)-1/q_i$ and $L_i(k)-m_i$ are both independent and identically distributed with zero means and finite variances $\sigma^2_{M_i}$ and $\sigma^2_{L_i}$ respectively and they obey the law of the iterated logarithm as per (3.6.14). So, using $T_i(n) \leq t < T_i(n+1)$ and $T_i(n+1)/T_i(n) \to 1$ as $t\to\infty$, we get

$$\limsup_{t\to\infty}|V_i(t)| \leq \frac{\sigma_{M_i}}{\sqrt{m_i}} + \frac{1}{q_i m_i}\frac{\sigma_{L_i}}{\sqrt{m_i}} = \frac{1}{\sqrt{m_i}}\big(\sigma_{M_i}+\pi_i\sigma_{L_i}\big).$$

Thus, by Lemma 3.6.1,

$$\limsup_{t\to\infty}\frac{t}{\sqrt{2t\log\log t}}\Big|\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds - \pi_i\Big| \leq \frac{1}{\sqrt{m_i}}\big(\sigma_{M_i}+\pi_i\sigma_{L_i}\big) =: \beta_i, \quad \text{a.s.} \quad (3.6.16)$$

Returning to (3.6.12) we have

$$\Big|\frac{1}{t}\int_0^t \gamma^2(Y(s))ds - \sigma_*^2\Big| \leq \sum_{i\in\mathbb{S}}\gamma^2(i)\Big|\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds - \pi_i\Big|$$

and hence, since the sum has finitely many terms,

$$\limsup_{t\to\infty}\frac{t}{\sqrt{2t\log\log t}}\Big|\frac{1}{t}\int_0^t \gamma^2(Y(s))ds - \sigma_*^2\Big|$$
$$\leq \sum_{i\in\mathbb{S}}\gamma^2(i)\limsup_{t\to\infty}\frac{t}{\sqrt{2t\log\log t}}\Big|\frac{1}{t}\int_0^t 1_{\{Y(s)=i\}}ds - \pi_i\Big|.$$

Finally, by (3.6.16) we get the desired conclusion. □

Before we prove Theorem 3.4.4 we state a useful lemma.

Lemma 3.6.2. *Suppose Y is an irreducible Markov jump process with generator Γ and finite state space $\mathbb{S}$, and moreover that Y is stationary. Let $H\in\mathbb{S}$ and $\delta>0$. Define for $n\geq 0$*

$$A_n = \{Y(s) = H \textit{ for all } s\in[n\delta,(n+1)\delta]\}.$$

Then

$$\lim_{m\to\infty} Cov(I_{A_n}, I_{A_{n+m}}) = 0. \tag{3.6.17}$$

Proof of Lemma 3.6.2. Since Y is stationary, it has a stationary distribution π such that $\pi\Gamma = 0$ and $\sum_{j\in\mathbb{S}} \pi_j = 1$ with each $\pi_j > 0$. We note that $\mathbb{P}[A_n] = \mathbb{E}[I_{A_n}]$ and that $\mathbb{P}[A_n \cap A_{n+m}] = \mathbb{E}[I_{A_n} I_{A_{n+m}}]$. Therefore

$$\text{Cov}(I_{A_n}, I_{A_{n+m}}) = \mathbb{P}[A_n \cap A_{n+m}] - \mathbb{P}[A_n]\mathbb{P}[A_{n+m}]. \tag{3.6.18}$$

We compute $\mathbb{P}[A_n]$ and $\mathbb{P}[A_n \cap A_{n+m}]$. First, as Y is stationary and has exponential holding times, we have

$$\begin{aligned}
\mathbb{P}[A_n] &= \mathbb{P}\big[\{Y(s) = H \text{ for all } s \in (n\delta, (n+1)\delta]\} \cap \{Y(n\delta) = H\}\big] \\
&= \mathbb{P}\big[Y(s) = H \text{ for all } s \in (n\delta, (n+1)\delta] | Y(n\delta) = H\big]\mathbb{P}[Y(n\delta) = H] \\
&= \mathbb{P}\big[Y(s) = H \text{ for all } s \in (n\delta, (n+1)\delta] | Y(n\delta) = H\big]\pi_H = e^{\gamma_{HH}\delta}\pi_H =: \pi(\delta).
\end{aligned}$$

Similarly, due to the stationarity of Y, we find $\mathbb{P}[A_{n+m}] = \pi(\delta)$ also. Thus by the Markov property we have for $m \geq 1$,

$$\begin{aligned}
&\mathbb{P}[A_n \cap A_{n+m}] = \mathbb{P}[A_{n+m}|A_n]\mathbb{P}[A_n] \\
&\quad = \mathbb{P}\Big[Y(s) = H \text{ for all } s \in [(n+m)\delta, (n+m+1)\delta]\Big| \\
&\qquad\qquad Y(s) = H \text{ for all } s \in [n\delta, (n+1)\delta]\Big]\pi(\delta) \\
&\quad = \mathbb{P}\Big[Y(s) = H \text{ for all } s \in [(n+m)\delta, (n+m+1)\delta]\Big|Y((n+1)\delta) = H\Big]\pi(\delta) \\
&\quad = \mathbb{P}\Big[\{Y(s) = H \text{ for all } s \in ((n+m)\delta, (n+m+1)\delta]\} \\
&\qquad\qquad \cap \{Y((n+m)\delta) = H\}\Big|Y((n+1)\delta) = H\Big]\pi(\delta).
\end{aligned}$$

Finally, using properties of conditional probability and the Markov property we get

$$\begin{aligned}\mathbb{P}[A_n \cap A_{n+m}] &= \mathbb{P}\Big[Y(s) = H \text{ for all } s \in ((n+m)\delta, (n+m+1)\delta]\Big| \\ &\qquad \{Y((n+m)\delta) = H, Y((n+1)\delta) = H\}\Big] \\ &\qquad\qquad \times \mathbb{P}\Big[Y((n+m)\delta) = H \Big| Y((n+1)\delta) = H\Big]\pi(\delta) \\ &= \mathbb{P}\Big[Y(s) = H \text{ for all } s \in ((n+m)\delta, (n+m+1)\delta]\Big| \\ &\qquad Y((n+m)\delta) = H\Big] P_{HH}((m-1)\delta)\pi(\delta) = e^{\gamma_{HH}\delta} P_{HH}((m-1)\delta)\pi(\delta),\end{aligned}$$

where $P_{ij}(t) = \mathbb{P}[Y(t+s) = j | Y(s) = i]$. Since Y is irreducible and finite, it follows from Theorem 3.6.2 in [67] that

$$\lim_{m\to\infty} P_{HH}((m-1)\delta) = \pi_H. \tag{3.6.19}$$

This implies

$$\lim_{m\to\infty} \mathbb{P}[A_n \cap A_{n+m}] = e^{\gamma_{HH}\delta}\pi_H \pi(\delta).$$

Therefore, returning to (3.6.18) we have

$$\lim_{m\to\infty} \mathrm{Cov}(I_{A_n}, I_{A_{n+m}}) = e^{\gamma_{HH}\delta}\pi_H\pi(\delta) - \pi(\delta)^2 = 0,$$

where we used the definition of $\pi(\delta)$ at the last step. This completes the proof. □

Proof of Theorem 3.4.4. Applying Theorem 4.3 in [9] in the simple case where the diffusion coefficient is $t-$ and $X-$independent, gives the upper bound

$$\limsup_{t\to\infty} \frac{|R_\delta(t)|}{\sqrt{2\log t}} \leq \sigma_H\sqrt{\delta}, \quad \text{a.s.} \tag{3.6.20}$$

We are lead to prove (3.4.18) by the following argument. First $R_\delta(t)$ is given by (3.4.17), so because the limits are finite we have

$$\limsup_{t\to\infty} \frac{|R_\delta(t)|}{\sqrt{2\log t}} = \limsup_{t\to\infty} \frac{|X(t-\delta) - X(t)|}{\sqrt{2\log t}}, \quad \text{a.s.}$$

and since

$$X(t) - X(t-\delta) = \int_{t-\delta}^{t} f(X(s), Y(s))\, ds + \int_{t-\delta}^{t} \gamma(Y(s))\, dB(s), \quad t \geq \delta$$

we have, using the fact that f is globally bounded by (3.3.3),

$$\limsup_{t\to\infty} \frac{|R_\delta(t)|}{\sqrt{2\log t}} = \limsup_{t\to\infty} \frac{|\int_{t-\delta}^{t} \gamma(Y(s))\, dB(s)|}{\sqrt{2\log t}}, \quad \text{a.s.}$$

In particular, with $U_n = \int_{n\delta}^{(n+1)\delta} \gamma(Y(s))\, dB(s)$ we have

$$\limsup_{t\to\infty} \frac{|R_\delta(t)|}{\sqrt{2\log t}} \geq \limsup_{n\to\infty} \frac{|R_\delta((n+1)\delta)|}{\sqrt{2\log((n+1)\delta)}} = \limsup_{n\to\infty} \frac{|U_n|}{\sqrt{2\log n}}. \tag{3.6.21}$$

Since Y is stationary, the probability that $Y(n\delta) = H$ is π_H. Define the event $A_n := \{Y(s) = H, \text{ for all } s \in [n\delta, (n+1)\delta]\}$. Then

$$\mathbb{P}[A_n] = \mathbb{P}[Y(n\delta) = H]\mathbb{P}[\text{no jump from state } H \text{ for at least } \delta \text{ time units}]$$
$$= \pi_H e^{\gamma_{HH}\delta} =: \pi(\delta).$$

Note also that the process $\{I_{A_n} : n \geq 1\}$ is stationary and that by Lemma 3.6.2 we have $\text{Cov}(I_{A_n}, I_{A_{n+m}}) \to 0$ as $m \to \infty$. Define $T_n = \sum_{j=1}^{n} I_{A_j}$. By Theorem 9.5.2 in [31] there exists a random variable W such that $\lim_{n\to\infty} T_n/n = W$ where $\mathbb{E}[W] = \pi(\delta)$ and by Theorem 9.5.3 in [31], $\mathbb{E}[(T_n/n - W)^2] \to 0$ as $n \to \infty$. Therefore, using Problem 9.7.9 in [31] along with the fact that $\text{Cov}(I_{A_n}, I_{A_{n+m}}) \to 0$ as $m \to \infty$ it can be shown that $T_n/n \to \pi(\delta)$ as $n \to \infty$ a.s. Let $L_n = \min\{l \geq n : \sum_{j=1}^{l} I_{A_j} = n\}$. By definition $I_{A_{L_n}} = 1$. Then if we consider the collection of $\{U_j : j = 1, \ldots, n\}$ for which $I_{A_j} = 1$ we have

$$\max_{1\leq j\leq n} |U_j| \geq \max_{1\leq k\leq T_n} |U_{L_k}|.$$

Next, if $I_{A_n} = 1$ then $Y(s) = H$ for all $s \in [n\delta, (n+1)\delta]$ and thus we have $U_n = \int_{n\delta}^{(n+1)\delta} \gamma(H)\, dB(s) = \gamma(H)(B((n+1)\delta) - B(n\delta))$. Without loss of generality we consider the case when $\gamma(H) > 0$. If $\gamma(H) < 0$ then we can redefine the Brownian motion as

$B_- = -B$ and proceed as in the case where $\gamma(H) > 0$. Hence, since $\gamma(H) = \sigma_H$, we get

$$\max_{1\le j\le n} |U_j| \ge \max_{1\le k\le T_n} |U_{L_k}| = \max_{1\le k\le T_n} |\sigma_H\{B((L_k+1)\delta) - B(L_k\delta)\}|.$$

Therefore with $\xi(k) := B((L_k+1)\delta) - B(L_k\delta)$ we have

$$\limsup_{n\to\infty} \frac{\max_{1\le j\le n} |U_j|}{\sqrt{2\log n}} \ge \limsup_{n\to\infty} \frac{\max_{1\le k\le T_n} |\sigma_H \xi(k)|}{\sqrt{2\log T_n}} \cdot \sqrt{\frac{\log T_n}{\log n}}$$
$$= \limsup_{n\to\infty} \frac{\max_{1\le k\le T_n} |\sigma_H \xi(k)|}{\sqrt{2\log T_n}} = \sigma_H \limsup_{n\to\infty} \frac{\max_{1\le k\le n} |\xi(k)|}{\sqrt{2\log n}},$$

where we used the fact that $T_n \to \infty$ as $n \to \infty$ a.s. at the last step. Since B and Y are independent, it follows that $L = \{L_n : n \ge 1\}$ and B are independent. Let $m \in \mathbb{N}$ and $k_1 < k_2 < \dots k_m$. Then, because $L_{k+1} - L_k \ge 1$, we have

$$\mathbb{P}[\xi(k_1) \le x_1, \xi(k_2) \le x_2, \dots, \xi(k_{m-1}) \le x_{m-1}, \xi(k_m) \le x_m]$$
$$= \sum_{n_1<n_2\dots<n_m} \mathbb{P}[\xi(k_1) \le x_1, \xi(k_2) \le x_2, \dots, \xi(k_{m-1}) \le x_{m-1}, \xi(k_m) \le x_m|$$
$$L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_{m-1}} = n_{m-1}, L_{k_m} = n_m]$$
$$\times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_{m-1}} = n_{m-1}, L_{k_m} = n_m]$$
$$= \sum_{n_1<n_2\dots<n_m} \mathbb{P}\big[B((L_{k_1}+1)\delta) - B(L_{k_1}\delta) \le x_1, \dots, B((L_{k_m}+1)\delta) - B(L_{k_m}\delta) \le x_m|$$
$$L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_{m-1}} = n_{m-1}, L_{k_m} = n_m\big]$$
$$\times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_{m-1}} = n_{m-1}, L_{k_m} = n_m]$$
$$= \sum_{n_1<n_2\dots<n_m} \mathbb{P}\big[B(n_1+1)\delta) - B(n_1\delta) \le x_1, \dots, B((n_m+1)\delta) - B(n_m\delta) \le x_m|$$
$$L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_{m-1}} = n_{m-1}, L_{k_m} = n_m\big]$$
$$\times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_{m-1}} = n_{m-1}, L_{k_m} = n_m]$$
$$= \sum_{n_1<n_2\dots<n_m} \mathbb{P}\big[B(n_1+1)\delta) - B(n_1\delta) \le x_1, \dots, B((n_m+1)\delta) - B(n_m\delta) \le x_m\big]$$
$$\times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_{m-1}} = n_{m-1}, L_{k_m} = n_m],$$

where we have used the independence of the B's and L's at the last step. Since $\delta > 0$ and we have $1 + n_i \leq n_{i+1}$, it follows that $(n_i + 1)\delta \leq n_{i+1}\delta$. Therefore, because there is no overlap of the Brownian increments, each of the random variables $B((n_i+1)\delta) - B(n_i\delta)$ for $i = 1, \ldots, m$ are independently and identically normally distributed with zero mean and variance δ. Therefore if Φ_δ is the distribution function of a standardised normal random variable, we have

$$\begin{aligned}
&\mathbb{P}[\xi(k_1) \leq x_1, \xi(k_2) \leq x_2, \ldots, \xi(k_{m-1}) \leq x_{m-1}, \xi(k_m) \leq x_m] \\
&= \sum_{n_1 < n_2 \cdots < n_m} \mathbb{P}\big[B(n_1+1)\delta) - B(n_1\delta) \leq x_1\big] \ldots \mathbb{P}\big[B((n_m+1)\delta) - B(n_m\delta) \leq x_m\big] \\
&\qquad\qquad \times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \ldots, L_{k_{m-1}} = n_{m-1}, L_{k_m} = n_m] \\
&= \sum_{n_1 < n_2 \cdots < n_m} \Phi_\delta(x_1) \cdots \Phi_\delta(x_m) \times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \ldots, L_{k_m} = n_m] \\
&= \prod_{i=1}^m \Phi_\delta(x_i) \sum_{n_1 < n_2 \cdots < n_m} \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \ldots, L_{k_m} = n_m] = \prod_{i=1}^m \Phi_\delta(x_i).
\end{aligned}$$

Thus $\{\xi(k) : k \geq 1\}$ is a sequence of independent and identically distributed normal random variables with mean zero and variance δ. Therefore, by Lemma 1.0.1 and Lemma 3.1 in [6],

$$\limsup_{n\to\infty} \frac{\max_{1\leq k\leq n} |\xi(k)|}{\sqrt{2\log n}} = \limsup_{n\to\infty} \frac{|\xi(n)|}{\sqrt{2\log n}} = \sqrt{\delta}, \quad \text{a.s.}$$

Hence

$$\Lambda := \limsup_{n\to\infty} \frac{\max_{1\leq j\leq n} |U_j|}{\sqrt{2\log n}} \geq \sigma_H \sqrt{\delta}, \quad \text{a.s.}$$

This implies, again by Lemma 3.1 in [6], that

$$\limsup_{n\to\infty} \frac{|U_n|}{\sqrt{2\log n}} = \Lambda \geq \sigma_H \sqrt{\delta}, \quad \text{a.s.} \tag{3.6.22}$$

Combining (3.6.20), (3.6.21) and (3.6.22) gives (3.4.18). □

Chapter 4

Asymptotic Consistency in the Large Fluctuations of Discretised Market Models with Markovian Switching

4.1 Introduction

Chapter 3 examined the application of large deviation results to a variant of the Geometric Brownian Motion market model containing Markovian switching. This naturally leads to the question of whether or not these results can be recovered in a discrete–time setting or whether such results can be reliably captured by simulation. The first paper to appear in the literature in relation to the numerical simulation of SDEs with Markovian switching is [61]. The emphasis of this paper lies in error analysis, and in it they show the strong convergence of the numerical method (in this case Euler–Maruyama) to the exact solution. In recent years there has also been attention devoted to the question of whether or not properties of the solution of an SDE are preserved under a discretisation, for example in [39] and [58]. While the emphasis in these papers is on preserving mean–square stability and preserving stationarity, we devote our attention to preserving asymptotic behaviour.

In this chapter we study the discretisations of the type of SDEs with Markovian switching found in Chapter 3, although for simplicity we restrict our attention to autonomous equations. Moreover, we concentrate on the special case where the diffusion coefficient depends only on the switching parameter. More specifically, we study the discretisation of an SDE of the form

$$dX(t) = f(X(t), Y(t))\, dt + \gamma(Y(t))\, dB(t), \quad t \geq 0, \tag{4.1.1}$$

where $\gamma(y)$ and $xf(x, y)$ are uniformly bounded above and below, and Y is an irreducible continuous–time Markov chain with finite state space $\mathbb{S}$ independent of the Brownian motion B. We then examine the influence that X has on the discretisation of the process S which is governed by

$$dS(t) = \mu S(t)\, dt + S(t)\, dX(t), \quad t \geq 0, \tag{4.1.2}$$

and X obeys (4.1.1). Again, S may be thought of as a security price. (4.1.1) and (4.1.2) are motivated by observations from financial market econometrics that security prices often move from bearish to bullish (or other) regimes. These regimes are modelled by the presence of the Markov process Y. Asymptotic properties of the continuous–time model described by (4.1.1) and (4.1.2) are examined in Chapter 3.

This chapter shows that it is possible to discretise (4.1.1) and (4.1.2), by explicit Euler–Maruyama methods, in such a way that the almost sure asymptotic behaviour of the discretisation mimics that of the continuous–time equation, at least for all sufficiently small uniform step sizes h. To make our discussion more precise, recall that

$$\sigma_*^2 = \sum_{j \in \mathbb{S}} \gamma^2(j) \pi_j, \tag{4.1.3}$$

where $\pi = (\pi_j)_{j \in \mathbb{S}}$ is the stationary distribution of Y. We know from Chapter 3 that the continuous–time stock price obeys

$$\lim_{t \to \infty} \frac{1}{t} \log S(t) = \mu - \frac{1}{2}\sigma_*^2, \quad \text{a.s.}, \tag{4.1.4}$$

and that there exist constants $C_1, C_2 > 0$ such that

$$C_2 \leq \limsup_{t \to \infty} \frac{|\log S(t) - (\mu - \frac{1}{2}\sigma_*^2)t|}{\sqrt{2t \log \log t}} \leq C_1, \quad \text{a.s.} \tag{4.1.5}$$

We have also shown that if R_δ, the returns process, is defined for $\delta > 0$ by

$$R_\delta(t) := \log(S(t)/S(t-\delta)), \quad t \geq \delta, \tag{4.1.6}$$

then

$$\limsup_{t\to\infty} \frac{|R_\delta(t)|}{\sqrt{2\log t}} = \sigma_H \sqrt{\delta}, \quad \text{a.s.}, \tag{4.1.7}$$

where $\sigma_H = \max_{j\in\mathbb{S}} |\gamma(j)|$. It is shown in this chapter that each of these results is recovered appropriately for the discretisations of (4.1.1) and (4.1.2). More precisely we prove that if we take a h–uniform time discretisation, then the discretised stock price S_h obeys

$$\lim_{n\to\infty} \frac{1}{nh} \log S_h(n) = \mu - \frac{1}{2}\sigma_*^2, \quad \text{a.s.} \tag{4.1.8}$$

Moreover, we show that there exists a constant $C'(h) > 0$ such that

$$\limsup_{n\to\infty} \frac{|\log S_h(n) - (\mu - \frac{1}{2}\sigma_*^2)nh|}{\sqrt{2nh\log\log nh}} \leq C'(h), \quad \text{a.s.}, \tag{4.1.9}$$

and that the discrete returns $R_{\delta,h}$ defined by

$$R_{\delta,h}(n) = \log(S_h(n)/S_h(n - \Delta(h,\delta)))$$

where $\Delta(h,\delta)$ is the smallest integer greater than or equal to δ/h, obey

$$\limsup_{n\to\infty} \frac{|R_{\delta,h}(n)|}{\sqrt{2\log nh}} = \sigma_H \sqrt{h\Delta(h,\delta)}, \quad \text{a.s.}, \tag{4.1.10}$$

where σ_H is as defined earlier. Moreover the constant on the right–hand side of (4.1.10) converges to the constant on the right–hand side of (4.1.7) as $h \to 0$. The asymptotic results (4.1.8)–(4.1.10) are clearly discrete analogues of (4.1.4)–(4.1.7).

In this chapter, the methods of discretisation and discussion of the main results are given in Section 4.2, with proofs being postponed to Section 4.3.

4.2 Discrete–Time Processes

4.2.1 Discretisation of the continuous–time Markov chain

We seek to approximate $Y(t)$, as defined in Chapter 1, at the set of uniformly spaced time points $t = nh$ for $n \geq 0$ and some fixed $h > 0$. We suppose that

$$0 < h < h_1 := \min_{i \in \mathbb{S}} \frac{1}{\sum_{j \neq i} \gamma_{ij}}. \tag{4.2.1}$$

A consequence of irreducibility is that the denominator in each fraction is positive, so h_1 is finite. Thus $\gamma_{ii} < 0$. Now, as $h < h_1$, for each $i \in \mathbb{S}$ we have

$$h < \min_{i \in \mathbb{S}} \frac{1}{\sum_{j \neq i} \gamma_{ij}} \leq \frac{1}{\sum_{j \neq i} \gamma_{ij}} = \frac{1}{-\gamma_{ii}}.$$

Hence $1 + \gamma_{ii} h > 0$, and clearly we also have $1 + \gamma_{ii} h < 1$. For $i \neq j$ we have $\gamma_{ij} h \geq 0$ and as $\gamma_{ij} \leq \sum_{k \neq i} \gamma_{ik}$ we have $1/\gamma_{ij} \geq 1/\sum_{k \neq i} \gamma_{ik} = 1/-\gamma_{ii} > h$, so $h\gamma_{ij} < 1$. Therefore the $N \times N$ matrix $P(h)$ defined by

$$P(h) = I_N + h\Gamma \tag{4.2.2}$$

has $P_{ij}(h) = h\gamma_{ij} \in (0,1)$ for $i \neq j$ and $P_{ii}(h) = 1 + h\gamma_{ii} \in (0,1)$. Moreover for each $i \in \mathbb{S}$ we have

$$\sum_{j=1}^{N} P_{ij}(h) = \sum_{j \neq i} h\gamma_{ij} + 1 + h\gamma_{ii} = 1 + h\Big(\sum_{j \neq i} \gamma_{ij} + \gamma_{ii}\Big) = 1.$$

Therefore for $h \in (0, h_1)$ we have that $P(h)$ is an $N \times N$ stochastic matrix. We now define the discrete–time and time–homogeneous Markov chain $Y_h = \{Y_h(n) : n \geq 0\}$ so that $Y_h(0) = Y(0)$, where $P(h)$ is the one–step transition matrix of Y_h, namely

$$P_{ij}(h) = \mathbb{P}[Y_h(n+1) = j | Y_h(n) = i].$$

Note also that if π is the vector representing the stationary distribution of Y then $\pi(h) = \pi$ obeys

$$\pi(h) - \pi(h)P(h) = \pi - \pi(I_N + h\Gamma) = -h\pi\Gamma = 0, \tag{4.2.3}$$

by (1.0.5). Since $\sum_{j=1}^{N}\pi_j = 1$, it follows that π is also a stationary distribution of Y_h.

We now show that irreducibility of Y implies that of Y_h. The irreducibility of Y implies that for every $i, j \in \mathbb{S}$, one can find finite numbers $i_1, i_2, \cdots, i_k \in \mathbb{S}$ such that $\gamma_{i,i_1}\gamma_{i_1,i_2}\cdots\gamma_{i_k,j} > 0$. In this product, we can choose without loss of generality to have $i \neq i_1$, $i_l \neq i_{l+1}$ for $l = 1, \ldots, k-1$, $i_k \neq j$. Therefore we have $\gamma_{i,i_1} > 0$, $\gamma_{i_1,i_2} > 0, \ldots, \gamma_{i_k,j} > 0$ with $i \neq i_1$, $i_l \neq i_{l+1}$ for $l = 1, \ldots, k-1$, and $i_k \neq j$. Thus we have $P_{i,i_1}(h) = \gamma_{i,i_1}h > 0$, $P_{i_1,i_2}(h) = \gamma_{i_1,i_2}h > 0, \ldots, P_{i_k,j}(h) = \gamma_{i_k,j}h > 0$ with $i \neq i_1$, $i_l \neq i_{l+1}$ for $l = 1, \ldots, k-1$, and $i_k \neq j$ and therefore, by Theorem 1.2.1 in [67], Y_h is irreducible. Since the state space is finite, it follows that the chain is non–null persistent (cf. e.g., [31, Lemma 6.3.5]) and since it is irreducible, by e.g., [31, Theorem 6.4.3] there is a unique stationary distribution of Y_h. By e.g., [16, Theorem 5.7], the finiteness of the state space $\mathbb{S}$ and irreducibility of Y_h we have that

$$\lim_{n\to\infty}[P(h)^n]_{ij} = \pi_j^*(h),$$

where the limit is independent of i, $\pi_j^*(h) > 0$ for each $j \in \mathbb{S}$ and $\sum_{j\in\mathbb{S}}\pi_j^*(h) = 1$. Moreover by e.g., [16, Proposition 5.8] we have that $\pi^*(h) = \pi^*(h)P(h)$. Therefore $\pi^*(h)$ is a stationary distribution of Y_h. However, as the stationary distribution must be unique, and we already know from (4.2.3) that π is a stationary distribution, we have that $\pi^*(h) = \pi$, and so

$$\lim_{n\to\infty}[P(h)^n]_{ij} = \pi_j, \text{ with the limit being independent of } i.$$

We summarise the above discussion by stating a Theorem.

Theorem 4.2.1. *Let Y be a continuous time Markov chain with finite state space $\mathbb{S} = \{1, \ldots, N\}$. Suppose that Y has generator Γ and is irreducible. Let $h < h_1$ where $h_1 > 0$ is defined by* (4.2.1). *Then*

(i) Y has a unique stationary distribution $\pi \in \mathbb{R}^{1\times N}$ given by (1.0.5).

(ii) *The* $N \times N$ *matrix* $P(h)$ *defined by* (4.2.2) *is stochastic.*

(iii) *If* $Y_h = \{Y_h(n) : n \geq 0\}$ *is a discrete–time and time homogeneous Markov chain with state space* $\mathbb{S}$ *and with one–step transition matrix* $P(h)$*, then* Y_h *is irreducible and has unique stationary distribution* π*.*

(iv) *If* π *is given by* (1.0.5)*, and* $P(h)$ *by* (4.2.2) *then for every* $i, j \in \mathbb{S}$ *we have*

$$\lim_{n\to\infty} [P(h)^n]_{ij} = \pi_j, \text{ with the limit being independent of } i. \tag{4.2.4}$$

(v) *If* $\mathbb{P}[Y(0) = i] = \pi_i$ *for all* $i \in \mathbb{S}$*, and we define* $Y_h(0) = Y(0)$*, then both* Y *and* Y_h *are strictly stationary.*

A proof of this theorem is omitted as the details were presented above.

Remark 4.2.1. It is worth noting that this method of discretising the Markov chain, namely (4.2.2), is different to that used by Higham, Mao and Yuan in [39] for example, where the discretisation takes the form

$$P_Y(h) = e^{h\Gamma}. \tag{4.2.5}$$

Although this discretisation method (4.2.5) does not require an initial step–size restriction on h, it does however require more computational effort compared to the method described in (4.2.2). In fact, (4.2.2) represents the first two terms in the Taylor expansion of (4.2.5), while treating the remaining terms as negligible for small enough step–size. Nonetheless, the actual simulation of the discrete Markov chain follows the same procedure as outlined in [39]. Throughout the chapter we use (4.2.2) as an alternative to (4.2.5), although we are free to use (4.2.5) if we wish to remove the initial step–size restriction.

Remark 4.2.2. It is clear then that the discrete chain Y_h with transition matrix (4.2.2) represents an *approximation* to the original chain Y, it is not an exact discrete representation. However, as shown above, $P(h)$ defined in (4.2.2) is a stochastic matrix and it preserves the correct stationary distribution of the original chain. Moreover, the approximation is good for sufficiently small $h > 0$ since

$$\|P_Y(h) - P(h)\| = \|e^{h\Gamma} - (I_N + h\Gamma)\| \leq \sum_{n=2}^{\infty} \frac{h^n \|\Gamma\|^n}{n!} = O(h^2). \tag{4.2.6}$$

In fact, we can approximate (4.2.5) by taking the first $m+1$ terms in the Taylor expansion and such a matrix will still preserve the stationary distribution of the original chain. So, provided such a matrix is a stochastic matrix (which may be arranged for sufficiently small h) this will approximate the original chain where the error is of order $O(h^{m+1})$.

In a similar fashion to the continuous–time case in Chapter 3, we denote by $T_i(r)$ the time of the r^{th} return to state i, by $S_i(r)$ the length of the r^{th} excursion to i and by $V_i(r)$ the number of visits to i before time r. One can visualise these quantities in a similar way to Figure 3.1. By Lemma 1.5.1 in [67],

$$\text{the non-negative random variables } S_i(1), S_i(2), \dots \text{ are i.i.d with mean } m_{S_i}. \tag{4.2.7}$$

Since the chain is time–homogenous, for each $i \in \mathbb{S}$ the lengths of the excursions $S_i(r)$ are identically distributed for all r. The finiteness of the first and second moments of the length of the excursions is a consequence of analysis of e.g., Hunter [44] and Kemeny and Snell [48]. Alongside the fact that the second moments are finite, *formulae* for these finite moments of passage times between any two states are deduced in [44, Theorem 7.3.10] and in [48, Theorem 4.5.1], under the assumptions that the Markov chain is irreducible and has a finite state space. Both of these stipulations are satisfied by our discretised chain. Since the excursion time S_i is simply the passage time to state i from state i, we can therefore assume that

$$\text{the sequence of random variables } S_i(1), S_i(2), \dots \text{ has finite variance } \sigma^2_{S_i}. \tag{4.2.8}$$

4.2.2 Main Results

We are now in a position to state our main results. We consider the typical Euler–Maruyama discretisation of the SDE (4.1.1), which takes the form

$$X_h(n+1) = X_h(n) + hf(X_h(n), Y_h(n)) + \sqrt{h}\gamma(Y_h(n))\xi(n+1), \ n \geq 0, \tag{4.2.9}$$

where h is the step size and ξ is a sequence of independent standard normal random variables. We assume that there exists $\rho > 0$ such that

$$xf(x,y) \leq \rho \quad \text{for all } (x,y) \in \mathbb{R} \times \mathbb{S}, \tag{4.2.10}$$

and that f is globally bounded in the sense that

$$|f(x,y)| \leq \bar{f} < +\infty, \quad \text{for all } (x,y) \in \mathbb{R} \times \mathbb{S}. \tag{4.2.11}$$

To motivate the discretisation of the stock price (4.1.2), note that Itô's rule gives

$$d\log S(t) = \big[\mu + f(X(t), Y(t)) - \frac{1}{2}\gamma^2(Y(t))\big]dt + \gamma(Y(t))dB(t).$$

We define the gains process G by

$$G(t) := \log\frac{S(t)}{S(0)} = \int_0^t \mu + f(X(s), Y(s)) - \frac{1}{2}\gamma^2(Y(s))\,ds + \int_0^t \gamma(Y(s))\,dB(s).$$

Then, a discretisation of this gains process is given by

$$\begin{aligned} G_h(n+1) = G_h(n) + h\big[\mu + f(X_h(n), Y_h(n)) - \frac{1}{2}\gamma^2(Y_h(n))\big] \\ + \sqrt{h}\gamma(Y_h(n))\xi(n+1), \quad n \geq 0 \end{aligned} \tag{4.2.12}$$

where the discretised stock price process obeys

$$S_h(n) = S_h(0)\exp[G_h(n)], \quad n \geq 0; \quad S_h(0) > 0. \tag{4.2.13}$$

We note that one nice by–product of this discretisation is that the discretised stock prices are automatically positive, almost surely.

Theorem 4.2.2. *Let $h < h_1$ where h_1 is defined by* (4.2.1). *Let f satisfy* (4.2.10) *and* (4.2.11) *and let $\gamma : \mathbb{S} \to \mathbb{R}$. Then X_h, the unique adapted solution satisfying* (4.2.9), *satisfies*

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{2nh\log\log nh}} \leq \sigma_*\sqrt{2e}, \quad a.s.$$

where σ_* is defined by (4.1.3).

Notice that this result gives the same rate of growth but not the same constant as the equivalent continuous–time result (3.3.10) in Chapter 3, as this result contains an extra $\sqrt{2e}$ term. However, this difference between discrete–time and continuous–time results is not a consequence of the discretisation process. Rather it is a consequence of the fact that we use a different method of proof based on the exponential martingale inequality and Gronwall's Lemma (as opposed to the stochastic comparison methods used in Chapter 3). A continuous–time version of Theorem 4.2.2 was first established by an exponential martingale and Gronwall lemma proof in [52]: the proof of Theorem 4.2.2 is modelled on the argument in that work.

Moreover, the reason we have the factor of two (in contrast to the $\sqrt{e}$ term that one would normally expect when using this method of proof) is that the discrete–time analogue of the exponential martingale inequality (see Lemma 4.3.1) contains two quadratic variation terms instead of the usual one.

In light of Theorem 4.2.2 we get the following result for the trend rate of growth of the stock price process (4.2.13).

Theorem 4.2.3. *Let $h < h_1$ where h_1 is defined by* (4.2.1). *Let S_h be the discrete–time stock price process given by* (4.2.13) *where X_h is given by* (4.2.9). *Let f obey* (4.2.10) *and* (4.2.11) *and let $\gamma : \mathbb{S} \to \mathbb{R}$. Then S_h obeys*

$$\lim_{n\to\infty} \frac{1}{nh} \log S_h(n) = \mu - \frac{1}{2}\sigma_*^2,$$

where σ_* *is defined by* (4.1.3).

We also get a result concerning the deviations around this trend rate of growth.

Theorem 4.2.4. *Let* $h < h_1$ *where* h_1 *is defined by* (4.2.1). *Let* f *obey* (4.2.10) *and* (4.2.11) *and let* S_h *be the security price model given by* (4.2.13) *where* X_h *satisfies* (4.2.9). *Then, using the ergodic theorem for Markov chains,*

$$\limsup_{n\to\infty} \frac{|\log S_h(n) - (\mu - \frac{1}{2}\sigma_*^2)nh|}{\sqrt{2nh\log\log nh}} \le \sigma_*\sqrt{2e} + \frac{1}{2}\sum_{i\in\mathbb{S}}\gamma^2(i)\sqrt{h}\alpha_i, \quad a.s.,$$

where σ_* *is defined by* (4.1.3) *and* $\alpha_i = \sigma_{S_i}/(m_{S_i}\sqrt{m_{S_i}})$ *is deterministic.*

Remark 4.2.3. On first impressions it appears as though the $\sqrt{h}$ term on the right–hand side above will cause the summation to go to zero, which would be inconsistent with the continuous–time counterpart (3.4.13). However, a simulation of the problem (see Appendix A) reveals evidence which suggests that in fact $\sqrt{h}\alpha_i \to \alpha_i^*$ as $h \to 0$ where α_i^* is finite.

Also, we are unable to obtain a lower bound on the fluctuations (in contrast to the continuous–time equivalent (3.4.13) in Chapter 3) because we do not have an exact fluctuations result corresponding to Corollary 3.3.1 in the continuous case.

We now state the ergodic theorem and its associated rate of convergence, as used in the previous theorem, in the following discrete–time analogue of Proposition 3.4.1.

Proposition 4.2.1. *Let* $h < h_1$ *where* h_1 *is defined by* (4.2.1). *Let* $\mathbb{S}$ *be a finite, irreducible state space, let* $\gamma : \mathbb{S} \to \mathbb{R}$ *and let* Y_h *be a stationary discrete–time Markov chain. Then by the ergodic theorem*

$$\lim_{n\to\infty} \frac{1}{nh}\sum_{j=1}^{n}\gamma^2(Y_h(j))h = \sigma_*^2 = \sum_{i\in\mathbb{S}}\gamma^2(i)\pi_i.$$

Moreover, the rate of this convergence is given by

$$\limsup_{n\to\infty} \frac{nh}{\sqrt{2nh\log\log nh}}\left|\frac{1}{nh}\sum_{j=1}^{n}\gamma^2(Y_h(j))h - \sigma_*^2\right| \le \sum_{i\in\mathbb{S}}\gamma^2(i)\sqrt{h}\alpha_i \quad a.s.,$$

where $\alpha_i = \sigma_{S_i}/(m_{S_i}\sqrt{m_{S_i}})$ *is deterministic.*

4.2.3 Large fluctuations of the discretised δ–returns

In this subsection, we examine the large fluctuations of the δ–returns associated with the discretised stock price process. We let X be given by (4.1.1) and X_h be given by (4.2.9). Again the diffusion coefficient in (4.1.1) depends only on the N–state Markov jump process Y. As before we let Y_h be the discrete–time Markov chain which is a discretisation of Y.

We have shown in Chapter 3 that when Y is a stationary irreducible Markov jump process independent of B, and X is the unique adapted continuous solution to (4.1.1) and S satisfies (4.1.2), then R_δ, defined by (4.1.6), satisfies

$$\limsup_{t\to\infty} \frac{|R_\delta(t)|}{\sqrt{2\log t}} = \sigma_H\sqrt{\delta}, \quad \text{a.s.,} \tag{4.2.14}$$

where $\sigma_H = \max_{j\in\mathbb{S}} |\gamma(j)| > 0$. We now wish to show that when the δ–returns are appropriately discretised, the asymptotic behaviour of R_δ captured by (4.2.14) is recovered in discrete–time. Let $\Delta = \Delta(h,\delta) \in \mathbb{N}_0$ be such that $\Delta(h,\delta) = \lceil \frac{\delta}{h} \rceil$ so

$$\delta/h \le \Delta(h,\delta) < \delta/h + 1. \tag{4.2.15}$$

Now we define the discrete–time approximation to the δ return by

$$R_{\delta,h}(n) := \log(S_h(n)/S_h(n-\Delta(h,\delta))), \quad n \ge \Delta(h,\delta) \tag{4.2.16}$$

so that $R_{\delta,h}(n)$ approximates $R_\delta(nh)$. We see from (4.2.12) and (4.2.13) that for $n \ge \Delta$

$$\begin{aligned} R_{\delta,h}(n) = G_h(n) - G_h(n-\Delta) &= \sum_{j=1}^{\Delta} \big[G_h(n+1-j) - G_h(n-j)\big] \\ &= \sum_{j=1}^{\Delta} h\big\{\mu + f(X_h(n-j), Y_h(n-j)) - \frac{1}{2}\gamma^2(Y_h(n-j))\big\} \\ &\quad + \sum_{j=1}^{\Delta} \sqrt{h}\gamma(Y_h(n-j))\xi(n+1-j). \end{aligned}$$

It proves convenient to introduce the process $V_\Delta(n)$ by

$$V_\Delta(n) = \sum_{l=n-\Delta}^{n-1} \gamma(Y_h(l))\xi(l+1), \quad n \ge \Delta, \tag{4.2.17}$$

so that $R_{\delta,h}$ is given by

$$R_{\delta,h}(n) = \sum_{l=n-\Delta}^{n-1} h\big[\mu + f(X_h(l), Y_h(l)) - \frac{1}{2}\gamma^2(Y_h(l))\big] + \sqrt{h}V_\Delta(n), \quad n \geq \Delta. \tag{4.2.18}$$

We have the following result concerning the large fluctuations of $R_{\delta,h}$ which corresponds to (4.2.14) in the continuous–time case.

Theorem 4.2.5. *Let f obey* (4.2.10) *and* (4.2.11) *and let Y be an irreducible N–state Markov jump process. Let $h < h_1$ and let Y_h be the discrete Markov chain defined in Theorem 4.2.1. Suppose that S_h is given by* (4.2.13)*. Let $\delta > 0$, and suppose that $\Delta(h, \delta) \in \mathbb{N}$ is defined by* (4.2.15)*. Then $R_{\delta,h}$ defined by* (4.2.16) *obeys*

$$\limsup_{n\to\infty} \frac{|R_{\delta,h}(n)|}{\sqrt{2\log nh}} = \sigma_H \sqrt{h\Delta(h,\delta)}, \quad a.s., \tag{4.2.19}$$

where $\sigma_H = \max_{j\in\mathbb{S}} |\gamma(j)|$.

Note that (4.2.19) shows that the discretised returns $R_{\delta,h}$ defined by (4.2.16) have the same exact a.s. power logarithmic growth in time of its large fluctuations as described by (4.2.14) which is experienced by the continuous–time process R_δ defined by (4.1.6). Moreover, as the time step h tends to zero, the growth rates of the discrete process converges to that of the continuous process since the coefficient $c(h) := \sigma_H \sqrt{h\Delta(h,\delta)}$ on the righthand side of (4.2.19) converges to the coefficient $c = \sigma_H\sqrt{\delta}$ on the righthand side of (4.2.14), because by (4.2.15) we have $\delta \leq h\Delta(h,\delta) < \delta + h$.

4.3 Proofs of Results from Section 4.2

4.3.1 Preliminaries

We first state discrete–time analogues of the well–known Exponential Martingale Inequality and Gronwall Inequality. These will be useful in proving the subsequent results.

Lemma 4.3.1. *Let M be a locally square–integrable martingale with predictable quadratic variation $\langle M\rangle_n = \sum_{k=1}^n \mathbb{E}[(\Delta M_k)^2|\mathcal{F}_{k-1}]$ and total quadratic variation given by $[M]_n = \sum_{k=1}^n(\Delta M_k)^2$. Then, for any $\alpha,\beta>0$ and $N\in\mathbb{N}$ we have*

$$\mathbb{P}\Big[\max_{1\le n\le N}\{M_n-\frac{\alpha}{2}([M]_n+\langle M\rangle_n)\}\ge\beta\Big]\le e^{-\alpha\beta}.$$

Proof. By Lemma B.1 in [13] it follows that for all $\alpha>0$ and $n\ge 0$,

$$V_\alpha(n):=\exp\big[\alpha M_n-\frac{\alpha^2}{2}([M]_n+\langle M\rangle_n)\big]$$

is a positive supermartingale with $\mathbb{E}[V_\alpha(n)]\le 1$. Then by the supermartingale inequality (see for example [62]) we have, for any $c>0$,

$$c\,\mathbb{P}\Big[\max_{1\le n\le N}V_\alpha(n)\ge c\Big]\le\mathbb{E}[V_0]=1$$

and hence

$$\mathbb{P}\Big[\max_{1\le n\le N}\exp[\alpha M_n-\frac{\alpha^2}{2}([M]_n+\langle M\rangle_n)]\ge c\Big]\le\frac{1}{c}.$$

Now, taking logs and dividing by α we obtain

$$\mathbb{P}\Big[\max_{1\le n\le N}\big\{M_n-\frac{\alpha}{2}([M]_n+\langle M\rangle_n)\big\}\ge\frac{1}{\alpha}\log c\Big]\le\frac{1}{c}.$$

Finally, letting $c>0$ be such that $(\log c)/\alpha=\beta$ we obtain the desired result. □

A more detailed discussion of this discrete–time EMI is postponed to the next chapter.

Lemma 4.3.2. *Let $a>0$ and $c>0$. Let $y(\cdot)$ and $b(\cdot)$ be nonnegative sequences. If*

$$y(n)\le a+c\sum_{j=0}^{n-1}b(j)y(j),\quad n\ge 1,$$

then

$$y(n)\le a\prod_{j=0}^{n-1}(1+c\,b(j)),\quad n\ge 1.$$

Proof. Let $Y(n) := a + c\sum_{j=0}^{n-1} b(j)y(j)$ for $n \geq 1$ where, by the summation convention, $Y(0) = a$ and $Y(n) \geq a$. Then $y(n) \leq Y(n)$ for $n \geq 1$. Now since $Y(n+1) - Y(n) = c\,b(n)y(n) \leq c\,b(n)Y(n)$ we have $Y(n+1) \leq (1 + c\,b(n))Y(n)$ and by iteration $Y(n+1) \leq Y(0)\prod_{j=0}^{n}(1 + c\,b(j))$, from which the result follows. □

4.3.2 Ergodic theorem for a product of white noise and a Markov chain

In the proof of Theorem 4.2.2 we encounter a term which involves the product of a white noise term and a Markov chain. Here we introduce some auxiliary results which will help us to deal with such terms as they arise. We assume that

$$\zeta = \{\zeta(n) : n \in \mathbb{N}\} \text{ and } \eta = \{\eta(n) : n \in \mathbb{N}\} \text{ are independent processes}$$
$$\text{i.e., } \zeta(n) \text{ and } \eta(m) \text{ are independent for each } n, m \in \mathbb{N} \quad (4.3.1)$$

as well as

$$\zeta \text{ is a sequence of i.i.d. non–negative random variables with finite mean } \mu_\zeta \quad (4.3.2)$$

and

$$\eta \text{ is an irreducible, stationary Markov chain with transition probability}$$
$$\text{matrix } P, \text{ on a finite state space } \mathbb{S} \subset (0, \infty). \quad (4.3.3)$$

If we denote the stationary distribution of η by π, then

$$\mu_\eta := \mathbb{E}[\eta(1)] = \sum_{j \in \mathbb{S}} j\pi_j. \quad (4.3.4)$$

Lemma 4.3.3. *Suppose that the processes ζ and η obey* (4.3.1), (4.3.2) *and* (4.3.3). *Then the process $U := \zeta\eta$ is both strictly and weakly stationary.*

Lemma 4.3.4. *If η is a sequence which obeys* (4.3.3) *and $\lim_{n\to\infty}\left[P^n(h)\right]_{ij} = \pi_j$, then $Cov(\eta(0), \eta(n)) \to 0$ as $n \to \infty$.*

Proposition 4.3.1. *Suppose that the processes* ζ *and* η *obey conditions* (4.3.1), (4.3.2) *and* (4.3.3) *and that* η *satisfies Lemma 4.3.4. Then the process* $U = \zeta\eta$ *obeys*

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} U(j) = \mu_\zeta \sum_{j\in\mathbb{S}} j\pi_j = \mu_\zeta\mu_\eta \quad a.s. \tag{4.3.5}$$

Proof of Lemma 4.3.3. Suppose momentarily that U is strictly stationary. Since ζ and η are independent and non–negative, we have

$$\mathbb{E}[U(n)] = \mathbb{E}[\eta(n)\zeta(n)] = \mathbb{E}[\eta(n)]\mathbb{E}[\zeta(n)] = \mu_\eta\mu_\zeta =: \mu_U.$$

Next, we consider $\text{Cov}(U(n), U(n+k))$ for $k \geq 0$ and $n \in \mathbb{N}$. By definition

$$\text{Cov}(U(n), U(n+k)) = \mathbb{E}[(U(n) - \mu_U)(U(n+k) - \mu_U)],$$

if it exists. We will show that $\mathbb{E}[U(n)U(n+k)]$ exists and then use this to compute $\text{Cov}(U(n), U(n+k)) = \mathbb{E}[U(n)U(n+k)] - \mu_U^2$. Since $U(n)$ is non–negative for each n, $\mathbb{E}[U(n)U(n+k)]$ exists, though may possibly be infinite. Now by the independence of η and ζ and the non–negativity of η and ζ we have

$$\mathbb{E}[U(n)U(n+k)] = \mathbb{E}[\eta(n)\zeta(n)\eta(n+k)\zeta(n+k)] = \mathbb{E}[\eta(n)\eta(n+k)]\mathbb{E}[\zeta(n)\zeta(n+k)].$$

Finally, using the fact that ζ is a sequence of independent random variables with finite mean, we get $\mathbb{E}[U(n)U(n+k)] = \mathbb{E}[\eta(n)\eta(n+k)]\mu_\zeta^2$. This quantity is finite because η assumes only a finite number of values. Therefore

$$\text{Cov}(U(n), U(n+k)) = \mathbb{E}[\eta(n)\eta(n+k)]\mu_\zeta^2 - \mu_U^2 = \mathbb{E}[\eta(n)\eta(n+k)]\mu_\zeta^2 - \mu_\zeta^2\mu_\eta^2.$$

Hence we have $\text{Cov}(U(n), U(n+k)) = \mu_\zeta^2\text{Cov}(\eta(n), \eta(n+k))$. Since η is strictly stationary and assumes only finitely many values, it has finite variance and is therefore weakly stationary. Therefore $\text{Cov}(\eta(n), \eta(n+k)) = \text{Cov}(\eta(0), \eta(k))$, and so $\text{Cov}(U(n), U(n+k)) = \mu_\zeta^2\text{Cov}(\eta(0), \eta(k)) =: \rho_U(k)$. Therefore U is weakly stationary since the covariance depends only on k.

We now turn to the proof that U is strictly stationary. We wish to prove for any $n \in \mathbb{N}$, any collection of non–negative integers $j_1, \ldots, j_n$ and $j \geq 0$ and any $x_1, \ldots, x_n \in \mathbb{R}$ that

$$\mathbb{P}[U(j_1) \leq x_1, U(j_2) \leq x_2, U(j_3) \leq x_3, \ldots, U(j_n) \leq x_n]$$
$$= \mathbb{P}[U(j_1 + j) \leq x_1, U(j_2 + j) \leq x_2, U(j_3 + j) \leq x_3, \ldots, U(j_n + j) \leq x_n]. \quad (4.3.6)$$

Let each ζ have distribution function F. We evaluate the lefthand side of (4.3.6) and deduce by analogy a formula for the righthand side; it will transpire that these formulae will be equal by virtue of the stationarity of η. By definition

$$\mathbb{P}[U(j_1) \leq x_1, U(j_2) \leq x_2, U(j_3) \leq x_3, \ldots, U(j_n) \leq x_n]$$
$$= \mathbb{P}[\eta(j_1)\zeta(j_1) \leq x_1, \eta(j_2)\zeta(j_2) \leq x_2, \eta(j_3)\zeta(j_3) \leq x_3, \ldots, \eta(j_n)\zeta(j_n) \leq x_n].$$

The Law of Total probability gives

$$\mathbb{P}[U(j_1) \leq x_1, U(j_2) \leq x_2, U(j_3) \leq x_3, \ldots, U(j_n) \leq x_n]$$
$$= \sum_{l_1 \in \mathbb{S}} \sum_{l_2 \in \mathbb{S}} \cdots \sum_{l_n \in \mathbb{S}} \mathbb{P}[\eta(j_1)\zeta(j_1) \leq x_1, \eta(j_2)\zeta(j_2) \leq x_2, \ldots, \eta(j_n)\zeta(j_n) \leq x_n |$$
$$\eta(j_1) = l_1, \eta(j_2) = l_2, \ldots, \eta(j_n) = l_n] \times \mathbb{P}[\eta(j_1) = l_1, \eta(j_2) = l_2, \ldots, \eta(j_n) = l_n].$$

Therefore as each $l_j \in \mathbb{S}$ is positive, we have

$$\mathbb{P}[U(j_1) \leq x_1, U(j_2) \leq x_2, U(j_3) \leq x_3, \ldots, U(j_n) \leq x_n]$$
$$= \sum_{l_1 \in \mathbb{S}} \sum_{l_2 \in \mathbb{S}} \cdots \sum_{l_n \in \mathbb{S}} \mathbb{P}[\zeta(j_1) \leq x_1/l_1, \zeta(j_2) \leq x_2/l_2, \ldots, \zeta(j_n) \leq x_n/l_n |$$
$$\eta(j_1) = l_1, \eta(j_2) = l_2, \ldots, \eta(j_n) = l_n] \times \mathbb{P}[\eta(j_1) = l_1, \eta(j_2) = l_2, \ldots, \eta(j_n) = l_n].$$

By (4.3.1) we get

$$\mathbb{P}[U(j_1) \leq x_1, U(j_2) \leq x_2, U(j_3) \leq x_3, \dots, U(j_n) \leq x_n]$$
$$= \sum_{l_1 \in \mathbb{S}} \sum_{l_2 \in \mathbb{S}} \cdots \sum_{l_n \in \mathbb{S}} \mathbb{P}[\zeta(j_1) \leq x_1/l_1, \zeta(j_2) \leq x_2/l_2, \dots, \zeta(j_n) \leq x_n/l_n]$$
$$\times \mathbb{P}[\eta(j_1) = l_1, \eta(j_2) = l_2, \dots, \eta(j_n) = l_n],$$

and the fact that ζ is a sequence of i.i.d. random variables with distribution function F implies

$$\mathbb{P}[U(j_1) \leq x_1, U(j_2) \leq x_2, U(j_3) \leq x_3, \dots, U(j_n) \leq x_n]$$
$$= \sum_{l_1 \in \mathbb{S}} \cdots \sum_{l_n \in \mathbb{S}} F(x_1/l_1)F(x_2/l_2)\dots F(x_n/l_n)\mathbb{P}[\eta(j_1) = l_1, \dots, \eta(j_n) = l_n]. \quad (4.3.7)$$

In the same manner we have

$$\mathbb{P}[U(j_1 + j) \leq x_1, U(j_2 + j) \leq x_2, U(j_3 + j) \leq x_3, \dots, U(j_n + j) \leq x_n]$$
$$= \sum_{l_1 \in \mathbb{S}} \cdots \sum_{l_n \in \mathbb{S}} F(x_1/l_1)F(x_2/l_2)\dots F(x_n/l_n)\mathbb{P}[\eta(j_1 + j) = l_1, \dots, \eta(j_n + j) = l_n],$$

which by the stationarity of η gives

$$\mathbb{P}[U(j_1 + j) \leq x_1, U(j_2 + j) \leq x_2, U(j_3 + j) \leq x_3, \dots, U(j_n + j) \leq x_n]$$
$$= \sum_{l_1 \in \mathbb{S}} \cdots \sum_{l_n \in \mathbb{S}} F(x_1/l_1)F(x_2/l_2)\dots F(x_n/l_n)\mathbb{P}[\eta(j_1) = l_1, \dots, \eta(j_n) = l_n].$$

Comparing this with (4.3.7) gives (4.3.6). □

Proof of Lemma 4.3.4. We note that as η is stationary, we have $\mathbb{P}[\eta(0) = i] = \pi_i$ and $\mathbb{P}[\eta(n) = i] = \pi_i$. Therefore

$$\mathbb{E}[\eta(n)] = \mathbb{E}[\eta(0)] = \sum_{i \in \mathbb{S}} i\pi_i. \quad (4.3.8)$$

Also we have

$$\mathbb{E}[\eta(0)\eta(n)] = \sum_{i\in\mathbb{S}}\sum_{j\in\mathbb{S}} ij\mathbb{P}[\eta(0)=i,\eta(n)=j] = \sum_{i\in\mathbb{S}}\sum_{j\in\mathbb{S}} ij\mathbb{P}[\eta(n)=j|\eta(0)=i]\mathbb{P}[\eta(0)=i]$$
$$= \sum_{i\in\mathbb{S}} i\pi_i \sum_{j\in\mathbb{S}} j\mathbb{P}[\eta(n)=j|\eta(0)=i].$$

Let P be the transition matrix associated with the Markov chain η. Then $\mathbb{P}[\eta(n) = j|\eta(0) = i] = [P^n]_{ij}$. Now, by assumption we have $\lim_{n\to\infty}[P^n]_{ij} = \pi_j$. Therefore $\mathbb{P}[\eta(n) = j|\eta(0) = i] \to \pi_j$ as $n \to \infty$. Since the state space $\mathbb{S}$ is finite, we have

$$\lim_{n\to\infty} \mathbb{E}[\eta(0)\eta(n)] = \lim_{n\to\infty} \sum_{i\in\mathbb{S}} i\pi_i \sum_{j\in\mathbb{S}} j\mathbb{P}[\eta(n)=j|\eta(0)=i]$$
$$= \sum_{i\in\mathbb{S}} i\pi_i \sum_{j\in\mathbb{S}} j \lim_{n\to\infty} \mathbb{P}[\eta(n)=j|\eta(0)=i] = \sum_{i\in\mathbb{S}} i\pi_i \sum_{j\in\mathbb{S}} j\pi_j = \left(\sum_{i\in\mathbb{S}} i\pi_i\right)^2.$$

Combining this with (4.3.8) we have

$$\lim_{n\to\infty} \text{Cov}(\eta(0),\eta(n)) = \lim_{n\to\infty} \mathbb{E}[\eta(0)\eta(n)] - \mathbb{E}[\eta(0)]^2 = 0,$$

whence the result. □

Proof of Proposition 4.3.1. By Lemma 4.3.3 we have that U is strictly stationary. Thus it follows from the ergodic theorem for strictly stationary sequences (see e.g., Theorem 9.5.2 in [31]) that there exists a random variable U^* with $\mathbb{E}[U^*] = \mathbb{E}[U]$ and

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} U(j) = U^*, \quad \text{a.s.},$$

and in mean. Since $\mathbb{E}[U(j)] = \mu_\zeta\mu_\eta = \mu_U$ we have

$$\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \{U(j) - \mu_\zeta\mu_\eta\} = U^* - \mu_\zeta\mu_\eta =: V_*, \quad \text{a.s.},$$

and the random variables $V(j) := U(j) - \mu_U$ in the summand have zero mean. Next notice that V is also weakly stationary with autocovariance function ρ_V, where $\rho_V(k) = \rho_U(k) = \mu_\zeta^2\text{Cov}(\eta(0),\eta(k))$ and so $\rho_V(k) \to 0$ as $k \to \infty$. Therefore

$$\lambda := \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \rho_V(j) = 0.$$

Thus by Lemma 9.5.12 in [31] we have that $\mathbb{E}[V_*] = 0$ and $\mathbb{E}[V_*^2] = \lambda = 0$, and so $V_* = 0$ a.s., from which we can deduce (4.3.5). □

4.3.3 Proofs of main results

Proof of Theorem 4.2.2. Fix $h \in (0, h_1)$, where h_1 is defined by (4.2.1). By squaring (4.2.9) and by adding and subtracting $h\gamma^2(Y_h(n))$ we have, for $n \geq 0$,

$$X_h^2(n+1) - X_h^2(n) = \lambda_h(X_h(n), Y_h(n)) + 2h\sqrt{h} f(X_h(n), Y_h(n))\gamma(Y_h(n))\xi(n+1) \\ + 2\sqrt{h} X_h(n)\gamma(Y_h(n))\xi(n+1) + h\gamma^2(Y_h(n))[\xi^2(n+1) - 1], \quad (4.3.9)$$

where $\lambda_h(x, y) := 2hxf(x, y) + h\gamma^2(y) + h^2 f^2(x, y)$. Then define

$$\Delta M_h^{(1)}(n+1) = 2h\sqrt{h} f(X_h(n), Y_h(n))\gamma(Y_h(n))\xi(n+1), \quad n \geq 0,$$

$$\Delta M_h^{(2)}(n+1) = h\gamma^2(Y_h(n))[\xi^2(n+1) - 1], \quad n \geq 0,$$

$$\Delta\theta_h(n+1) = 2\sqrt{h} X_h(n)\gamma(Y_h(n))\xi(n+1), \quad n \geq 0.$$

Thus $M_h^{(1)}(n+1) := \sum_{j=0}^n \Delta M_h^{(1)}(j+1)$, $M_h^{(2)}(n+1) := \sum_{j=0}^n \Delta M_h^{(2)}(j+1)$ and $\theta_h(n+1) := \sum_{j=0}^n \Delta\theta_h(j+1)$ are martingales with respect to the natural filtration generated by the ξ's. Returning to (4.3.9) and summing on both sides then gives, for $n \geq 0$,

$$X_h^2(n+1) - X_h^2(0) = \sum_{j=0}^n \lambda_h(X_h(n), Y_h(n)) + M_h^{(1)}(n+1) + M_h^{(2)}(n+1) + \theta_h(n+1). \quad (4.3.10)$$

Note that the martingales $M_h^{(1)}$ and $M_h^{(2)}$ have predictable quadratic variation

$$\langle M_h^{(1)}\rangle(n+1) = \sum_{j=0}^n 4h^3 f^2(X_h(j), Y_h(j))\gamma^2(Y_h(j)) \leq (n+1)4h^3 \bar{f}^2 \bar{\gamma}^2,$$

$$\langle M_h^{(2)}\rangle(n+1) = \sum_{j=0}^n h^2\gamma^4(Y_h(j))c_* \leq (n+1)h^2\bar{\gamma}^4 c_*,$$

where $\bar{f}$ is defined by (4.2.11), $\bar{\gamma} := \max_{j\in\mathbb{S}} \gamma(Y_h(j))$ and $c_* := \text{Var}[\xi^2(j+1) - 1] < +\infty$. By Section 12.14 in [80] it follows that

$$\lim_{n\to\infty} \frac{M_h^{(k)}(n+1)}{\langle M_h^{(k)}\rangle(n+1)} = 0, \quad \text{a.s. for } k = 1, 2.$$

Therefore we have

$$\lim_{n\to\infty}\frac{M_h^{(k)}(n+1)}{n+1} = \lim_{n\to\infty}\frac{M_h^{(k)}(n+1)}{\langle M_h^{(k)}\rangle(n+1)}.\frac{\langle M_h^{(k)}\rangle(n+1)}{n+1} = 0, \quad \text{a.s. for } k=1,2. \quad (4.3.11)$$

We now apply the exponential martingale inequality to the martingale $\theta_h(n+1)$ which has predictable quadratic variation given by

$$\begin{aligned}\langle\theta_h\rangle(n+1) &= \sum_{j=0}^{n}\mathbb{E}[(\Delta\theta_h(j+1))^2|\mathcal{F}_j] = \sum_{j=0}^{n}4h\gamma^2(Y_h(j))X_h^2(j)\mathbb{E}[\xi^2(j+1)|\mathcal{F}_j]\\ &= \sum_{j=0}^{n}4h\gamma^2(Y_h(j))X_h^2(j)\mathbb{E}[\xi^2(j+1)] = \sum_{j=0}^{n}4h\gamma^2(Y_h(j))X_h^2(j),\end{aligned}$$

and has total quadratic variation given by

$$[\theta_h](n+1) = \sum_{j=0}^{n}(\Delta\theta_h(j+1))^2 = \sum_{j=0}^{n}4h\gamma^2(Y_h(j))X_h^2(j)\xi^2(j+1), \quad n\ge 0.$$

Hence the sum of the quadratic variations is given by

$$\langle\theta_h\rangle(n+1)+[\theta_h](n+1) = \sum_{j=0}^{n}4h\gamma^2(Y_h(j))X_h^2(j)\big[1+\xi^2(j+1)\big], \quad \text{for } n\ge 0. \quad (4.3.12)$$

Applying Lemma 4.3.1 where $\beta>0$ and $\tau>1$ are arbitrary constants we have that for all $n\in\mathbb{N}$,

$$\mathbb{P}\Big[\max_{1\le m\le\lfloor\tau^n\rfloor}\Big\{\theta_h(m)-\frac{\beta}{2\tau^n}\sum_{j=0}^{m-1}4h\gamma^2(Y_h(j))X_h^2(j)\big[1+\xi^2(j+1)\big]\Big\}\ge\frac{\tau^{n+1}}{\beta}\log n\Big]\le\frac{1}{n^\tau},$$

where $\lfloor\cdot\rfloor$ signifies the integer part. The Borel–Cantelli lemma then yields that for almost all $\omega\in\Omega$, where $\mathbb{P}[\Omega]=1$, there is a random integer $n_0=n_0(\omega,h)$ sufficiently large such that $\lfloor\tau^{n_0-1}\rfloor>e^1$ and for $n\ge n_0$ we have

$$\theta_h(m)\le\frac{\tau^{n+1}}{\beta}\log n+\frac{2\beta}{\tau^n}\sum_{j=0}^{m-1}h\gamma^2(Y_h(j))X_h^2(j)[1+\xi^2(j+1)], \quad 1\le m\le\lfloor\tau^n\rfloor. \quad (4.3.13)$$

Recall from (4.3.10) that

$$X_h^2(m) = X_h^2(0)+\sum_{j=0}^{m-1}\lambda_h(X_h(j),Y_h(j))+M_h^{(1)}(m)+M_h^{(2)}(m)+\theta_h(m), \ m\ge 1. \quad (4.3.14)$$

Then define

$$T_h(m) := X_h^2(0) + \sum_{j=0}^{m-1} \lambda_h(X_h(j), Y_h(j)) + M_h^{(1)}(m) + M_h^{(2)}(m),$$

and note that by (4.2.10) and (4.2.11) we get

$$\lambda_h(x,y) = 2hxf(x,y) + h^2 f^2(x,y) + h\gamma^2(y) \le 2h\rho + h^2\bar{f}^2 + h\bar{\gamma}^2 := \bar{\lambda}_h.$$

Therefore, as a result of (4.3.11) it follows that

$$\limsup_{m\to\infty} \frac{T_h(m)}{m} \le \bar{\lambda}_h, \quad \text{a.s.}$$

Thus, for each fixed $h \in (0, h_1)$ there is an $m_1(h,\omega) \in \mathbb{N}$ such that $T_h(m) \le 2\bar{\lambda}_h m$ for $m \ge m_1(h,\omega)$. Moreover, on the finite set $m \in \{1, \dots, m_1(h,\omega)-1\}$ there exists a constant $T^*(h,\omega) < +\infty$ such that $T_h(m) \le T^*(h,\omega)$. Combining both of these estimates we have $T_h(m) \le T^*(h,\omega) + 2\bar{\lambda}_h m$ for $m \ge 1$. Using this bound, along with (4.3.13), (4.3.14) and the definition of T_h we have, for $n \ge n_0$ and $1 \le m \le \lfloor \tau^n \rfloor \le \tau^n$,

$$X_h^2(m) \le T^*(h,\omega) + 2\bar{\lambda}_h\tau^n + \frac{\tau^{n+1}}{\beta}\log n + \frac{2\beta}{\tau^n}\sum_{j=0}^{m-1} hX_h^2(j)\gamma^2(Y_h(j))[1+\xi^2(j+1)].$$

Following the notation of Lemma 4.3.2, set $y(m) := X_h^2(m)$, $a_h(n) := T^*(h,\omega) + 2\bar{\lambda}_h\tau^n + \beta^{-1}\tau^{n+1}\log n$, $c(n) := 2\beta\tau^{-n}$ and $b_h(j) := h\gamma^2(Y_h(j))[1+\xi^2(j+1)]$. Therefore, for $n \ge n_0$,

$$y(m) \le a_h(n) + c(n)\sum_{j=0}^{m-1} b_h(j)y(j), \quad 1 \le m \le \lfloor \tau^n \rfloor,$$

and so we can then apply Lemma 4.3.2 to conclude that for $n \ge n_0$,

$$X_h^2(m) \le a_h(n)\prod_{j=0}^{m-1}\big(1 + c(n)b_h(j)\big), \quad 1 \le m \le \lfloor \tau^n \rfloor.$$

Then using the fact that $1 + x \le e^x$ for any $x \ge 0$ we get, for $n \ge n_0$ and $1 \le m \le \lfloor \tau^n \rfloor$,

$$\begin{aligned} X_h^2(m) &\le a_h(n)\exp\left[2h\beta\tau^{-n}\sum_{j=0}^{m-1}\gamma^2(Y_h(j))\big(1+\xi^2(j+1)\big)\right] \\ &\le a_h(n)\exp\left[2h\beta\frac{1}{\lfloor \tau^n \rfloor}\sum_{j=0}^{\lfloor \tau^n \rfloor - 1}\gamma^2(Y_h(j))\big(1+\xi^2(j+1)\big)\right]. \end{aligned}$$

Now let $1 \leq \lfloor \tau^{n-1} \rfloor \leq m \leq \lfloor \tau^n \rfloor$. Then for $n \geq n_0$,

$$\frac{X_h^2(m)}{2mh\log\log m} \leq \frac{a_h(n)}{\beta^{-1}\tau^{n+1}\log n}\frac{1}{2\beta h}\frac{\tau^{n+1}}{\lfloor \tau^{n-1} \rfloor}\frac{\log n}{\log\log \lfloor \tau^{n-1} \rfloor} \times \exp\Big[2h\beta\frac{1}{\lfloor \tau^n \rfloor}\sum_{j=0}^{\lfloor \tau^n \rfloor -1}\gamma^2(Y_h(j))\big(1+\xi^2(j+1)\big)\Big]. \quad (4.3.15)$$

Notice that $\lim_{n\to\infty} a_h(n)/(\beta^{-1}\tau^{n+1}\log n) = 1$ and since $\tau^{n-1}-1 \leq \lfloor \tau^{n-1} \rfloor \leq \tau^{n-1}$ we have $\lim_{n\to\infty}\tau^{n+1}/\lfloor \tau^{n-1} \rfloor = \tau^2$. Moreover,

$$\log\left[(n-1)\log\tau + \log\left(1-1/\tau^{n-1}\right)\right] \leq \log\log\lfloor \tau^{n-1} \rfloor \leq \log(n-1)+\log\log\tau$$

and thus dividing by $\log n$ and taking the limit as $n \to \infty$ we find that

$$\lim_{n\to\infty}\frac{\log\log\lfloor \tau^{n-1} \rfloor}{\log n} = 1.$$

Moreover, by Propositions 4.2.1 and 4.3.1 we have

$$\lim_{n\to\infty}\frac{1}{\lfloor \tau^n \rfloor}\sum_{j=0}^{\lfloor \tau^n \rfloor -1}\gamma^2(Y_h(j))\big[1+\xi^2(j+1)\big] = \lim_{n\to\infty}\frac{1}{\lfloor \tau^n \rfloor}\sum_{j=0}^{\lfloor \tau^n \rfloor -1}\gamma^2(Y_h(j)) + \lim_{n\to\infty}\frac{1}{\lfloor \tau^n \rfloor}\sum_{j=0}^{\lfloor \tau^n \rfloor -1}\gamma^2(Y_h(j))\xi^2(j+1) = \sigma_*^2 + \sigma_*^2.1 = 2\sigma_*^2. \quad (4.3.16)$$

Finally, returning to (4.3.15) and using the fact that $n \to \infty$ as $m \to \infty$ we obtain

$$\limsup_{m\to\infty}\frac{X_h^2(m)}{2mh\log\log m} \leq \frac{\tau^2}{2\beta h}e^{2\beta h(2\sigma_*^2)}, \quad \text{a.s.}$$

Letting $\tau \to 1$ and choosing $\beta = (4h\sigma_*^2)^{-1}$ we get

$$\limsup_{m\to\infty}\frac{X_h^2(m)}{2mh\log\log mh} = \limsup_{m\to\infty}\frac{X_h^2(m)}{2mh\log\log m} \leq 2\sigma_*^2 e, \quad \text{a.s.}$$

Taking square roots on both sides gives the desired result. □

Proof of Theorem 4.2.3. By (4.2.13) we have

$$\log S_h(n) - \big(\mu - \frac{1}{2}\sigma_*^2\big)nh = \log S_h(0) + G_h(n) - \big(\mu - \frac{1}{2}\sigma_*^2\big)nh \quad (4.3.17)$$

and using (4.2.9) and (4.2.12) we have

$$G_h(n+1) - X_h(n+1) = G_h(n) - X_h(n) + h[\mu - \frac{1}{2}\gamma^2(Y_h(n))].$$

Define the process $H_h(n) := G_h(n) - X_h(n)$, then we have $G_h(n) = H_h(n) + X_h(n)$ where $H_h(n+1) = H_h(n) + h[\mu - \frac{1}{2}\gamma^2(Y_h(n))]$. Summing on both sides we find that

$$H_h(N) - H_h(0) = Nh\mu - \frac{h}{2}\sum_{n=0}^{N-1}\gamma^2(Y_h(n))$$

where $H_h(0) = G_h(0) - X_h(0) = \log(S_h(0)/S_h(0)) - X_h(0) = -X_h(0)$. Returning to (4.3.17),

$$\begin{aligned}\log S_h(n) - \big(\mu - \frac{1}{2}\sigma_*^2\big)nh &= \log S_h(0) + H_h(0) + nh\mu - \frac{h}{2}\sum_{j=0}^{n-1}\gamma^2(Y_h(j)) \\ &\quad + X_h(n) - \big(\mu - \frac{1}{2}\sigma_*^2\big)nh \\ &= \log S_h(0) - X_h(0) - \frac{nh}{2}\Big[\frac{1}{nh}\sum_{j=0}^{n-1}\gamma^2(Y_h(j))h - \sigma_*^2\Big] + X_h(n). \qquad (4.3.18)\end{aligned}$$

Finally, dividing by nh and using Proposition 4.2.1 and Theorem 4.2.2, we have

$$\lim_{n\to\infty}\frac{\log S_h(n)}{nh} - \big(\mu - \frac{1}{2}\sigma_*^2\big) = 0,$$

which gives the desired result. □

Proof of Theorem 4.2.4. Recalling (4.3.18) we have

$$\log S_h(n) - \big(\mu - \frac{\sigma_*^2}{2}\big)nh = \log S_h(0) - X_h(0) - \frac{nh}{2}\Big[\frac{1}{nh}\sum_{j=0}^{n-1}\gamma^2(Y_h(j))h - \sigma_*^2\Big] + X_h(n).$$

Thus, by Theorem 4.2.2 and Proposition 4.2.1 we get

$$\limsup_{n\to\infty}\frac{|\log S_h(n) - \big(\mu - \frac{1}{2}\sigma_*^2\big)nh|}{\sqrt{2nh\log\log nh}} \le \frac{1}{2}\sum_{i\in\mathbb{S}}\gamma^2(i)\sqrt{h}\alpha_i + \sigma_*\sqrt{2e}, \quad \text{a.s.},$$

which gives the desired result. □

Before we prove Proposition 4.2.1, we state without proof the following discrete–time analogue of Lemma 3.6.1 of Chapter 3.

Lemma 4.3.5. *If* $a(n), R(n)$ *and* $L(n)$, $n \geq 0$, *are three sequences such that* $R(n) \leq a(n) \leq L(n)$ *where*

$$\limsup_{n\to\infty} |R(n)| \leq r \quad \text{and} \quad \limsup_{n\to\infty} |L(n)| \leq l,$$

then

$$\limsup_{n\to\infty} |a(n)| \leq \max(r, l).$$

Proof of Proposition 4.2.1. The first part of this proof is modelled on a similar proof in [67]. Define by $T_i(r)$ the time of the r^{th} return to state i, by $S_i(r)$ the length of the r^{th} excursion to i and by $V_i(r)$ the number of visits to i before time r. Thus for $r = 0, 1, 2, \ldots$, setting $T_i(0) = 0$ we have

$$T_i(r+1) = \inf\{n \geq T_i(r) + 1 : Y_h(n) = i\},$$

$$S_i(r) = T_i(r) - T_i(r-1),$$

$$V_i(r) = \sum_{k=0}^{r-1} I_{\{Y_h(k)=i\}}.$$

In other words, the time of the $(r+1)^{\text{th}}$ return to i is the next time that $Y_h(n) = i$ which must be at least one time step after the previous return to i. Then the length of each excursion is the distance between consecutive return times. Note that $V_i(n)/n$ gives the proportion of time before n spent in state i.

Without loss of generality, suppose that $Y_h(n)$ is recurrent and fix a state i. For $T = T_i$ we have $\mathbb{P}[T < \infty] = 1$ since the process must return to i in finite time. By the strong Markov property $Y_h(T+n)$, $n \geq 0$, is independent of $Y_h(0), Y_h(1), \ldots, Y_h(T)$ and the long–run proportion of time spent in i is the same for $Y_h(T+n)$ and for $Y_h(n)$. In other words, we can assume that we start in state i. By Lemma 1.5.1 in [67], the non–negative random variables $S_i(1), S_i(2), \ldots,$ are independent and identically distributed with $\mathbb{E}[S_i(r)] = m_{S_i}$. Now

$$S_i(1) + S_i(2) + \cdots + S_i(V_i(n) - 1) \leq n - 1,$$

the left–hand side being the time of the last visit to i before n. Also

$$S_i(1) + S_i(2) + \cdots + S_i(V_i(n)) \geq n,$$

the left–hand side being the time of the first visit to i after $n-1$. Hence

$$\frac{S_i(1) + \cdots + S_i(V_i(n) - 1)}{V_i(n)} < \frac{n}{V_i(n)} \leq \frac{S_i(1) + \cdots + S_i(V_i(n))}{V_i(n)}. \tag{4.3.19}$$

By the strong law of large numbers

$$\mathbb{P}\Big[\frac{S_i(1) + \cdots + S_i(n)}{n} \to m_{S_i} \quad \text{as } n \to \infty\Big] = 1$$

and since the Markov process is recurrent, $\mathbb{P}[V_i(n) \to \infty \quad \text{as } n \to \infty] = 1$. So letting $n \to \infty$ (and thus $V_i(n) \to \infty$) we get, for each fixed state i,

$$\mathbb{P}\Big[\frac{V_i(n)}{n} \to \frac{1}{m_{S_i}} \quad \text{as } n \to \infty\Big] = 1. \tag{4.3.20}$$

Assume now that $Y_h(n)$ has stationary probability distribution $\pi_i = 1/m_{S_i}$. Then

$$\begin{aligned} \Big|\frac{1}{nh}\sum_{k=0}^{n-1}\gamma^2(Y_h(k))h - \sigma_*^2\Big| &= \Big|\frac{1}{n}\sum_{i\in\mathbb{S}}\gamma^2(i)\sum_{k=0}^{n-1} I_{\{Y_h(k)=i\}} - \sum_{i\in\mathbb{S}}\gamma^2(i)\pi_i\Big| \qquad (4.3.21) \\ &= \Big|\sum_{i\in\mathbb{S}}\gamma^2(i)\Big(\frac{V_i(n)}{n} - \pi_i\Big)\Big| \leq \max_{j\in\mathbb{S}}\gamma^2(j)\sum_{i\in\mathbb{S}}\Big|\frac{V_i(n)}{n} - \pi_i\Big| \end{aligned}$$

By (4.3.20) there exists $N = N(\omega)$ sufficiently large such that for $n \geq N(\omega)$

$$\sum_{i\in\mathbb{S}}\Big|\frac{V_i(n)}{n} - \pi_i\Big| < \varepsilon / \max_{j\in\mathbb{S}}\gamma^2(j)$$

and thus we have, for $n \geq N(\omega)$,

$$\Big|\frac{1}{nh}\sum_{k=0}^{n-1}\gamma^2(Y_h(k))h - \sigma_*^2\Big| < \varepsilon,$$

which establishes the desired convergence. To prove the second assertion we must determine the rate of this convergence, which is ultimately determined by the rate of convergence of $n/V_i(n)$ to m_{S_i}. By (4.3.19) we have

$$\frac{n}{V_i(n)} - m_{S_i} \leq \frac{\sum_{j=1}^{V_i(n)}(S_i(j) - m_{S_i})}{V_i(n)}$$

and so we get the upper bound

$$\frac{n}{\sqrt{2n\log\log n}}\Big\{\frac{n}{V_i(n)} - m_{S_i}\Big\} \le \frac{\sum_{j=1}^{V_i(n)}(S_i(j)-m_{S_i})}{\sqrt{2n\log\log n}}.\frac{n}{V_i(n)} =: L_i(n).$$

Similarly, by (4.3.19)

$$\frac{n}{V_i(n)} - m_{S_i} \ge \frac{\sum_{j=1}^{V_i(n)-1}(S_i(j)-m_{S_i})}{V_i(n)} - \frac{m_{S_i}}{V_i(n)}$$

and so we get the lower bound

$$\begin{aligned}\frac{n}{\sqrt{2n\log\log n}}\Big\{\frac{n}{V_i(n)} - m_{S_i}\Big\} \ge& \frac{\sum_{j=1}^{V_i(n)-1}(S_i(j)-m_{S_i})}{\sqrt{2n\log\log n}}.\frac{n}{V_i(n)}\\ &- \frac{m_{S_i}}{V_i(n)/n}.\frac{1}{\sqrt{2n\log\log n}} =: R_i(n).\end{aligned}$$

By (4.2.7) and (4.2.8), the random variables $S_i(j) - m_{S_i}$ are independent and identically distributed with zero mean and finite variance $\sigma^2_{S_i}$, so they obey the law of the iterated logarithm as follows

$$\limsup_{n\to\infty}\frac{|\sum_{j=1}^{n}(S_i(j)-m_{S_i})|}{\sqrt{2n\log\log n}} = \sigma_{S_i}, \quad \text{a.s.} \tag{4.3.22}$$

Since $V_i(n)\to\infty$ as $n\to\infty$ we have

$$\limsup_{n\to\infty}\frac{|\sum_{j=1}^{V_i(n)}(S_i(j)-m_{S_i})|}{\sqrt{2V_i(n)\log\log V_i(n)}} \le \sigma_{S_i}, \quad \text{a.s.,} \tag{4.3.23}$$

and since $V_i(n)/n \to 1/m_{S_i}$ as $n\to\infty$ we get

$$\limsup_{n\to\infty}\frac{|\sum_{j=1}^{V_i(n)}(S_i(j)-m_{S_i})|}{\sqrt{2n\log\log n}} \le \frac{\sigma_{S_i}}{\sqrt{m_{S_i}}}, \quad \text{a.s.}$$

The same result also holds with $V_i(n)$ replaced with $V_i(n)-1$. Therefore we have

$$\limsup_{n\to\infty}|L_i(n)| \le \frac{\sigma_{S_i}}{\sqrt{m_{S_i}}}m_{S_i}, \quad \text{a.s.,}$$

and since m_{S_i} is finite,

$$\limsup_{n\to\infty}|R_i(n)| \le \frac{\sigma_{S_i}}{\sqrt{m_{S_i}}}m_{S_i}, \quad \text{a.s.}$$

Thus, by Lemma 4.3.5

$$\limsup_{n\to\infty} \frac{n}{\sqrt{2n\log\log n}}\Big|\frac{n}{V_i(n)} - m_{S_i}\Big| \le \frac{\sigma_{S_i}}{\sqrt{m_{S_i}}} m_{S_i}, \quad \text{a.s.} \tag{4.3.24}$$

However, from this we must now determine the rate of convergence of $V_i(n)/n$ to $1/m_{S_i}$.

Notice that

$$\frac{n}{\sqrt{2n\log\log n}}\Big|\frac{V_i(n)}{n} - \frac{1}{m_{S_i}}\Big| = \frac{n}{\sqrt{2n\log\log n}} \frac{|m_{S_i} - n/V_i(n)|}{m_{S_i} n/V_i(n)}$$

and so

$$\limsup_{n\to\infty} \frac{n}{\sqrt{2n\log\log n}}\Big|\frac{V_i(n)}{n} - \frac{1}{m_{S_i}}\Big| = \limsup_{n\to\infty} \frac{n}{\sqrt{2n\log\log n}}\Big|m_{S_i} - \frac{n}{V_i(n)}\Big|\frac{1}{m_{S_i}}\frac{V_i(n)}{n}.$$

Therefore, by (4.3.24) and (4.3.20)

$$\limsup_{n\to\infty} \frac{n}{\sqrt{2n\log\log n}}\Big|\frac{V_i(n)}{n} - \frac{1}{m_{S_i}}\Big| \le \frac{\sigma_{S_i}}{m_{S_i}\sqrt{m_{S_i}}} =: \alpha_i, \quad \text{a.s.} \tag{4.3.25}$$

Returning to (4.3.21) we have

$$\Big|\frac{1}{nh}\sum_{k=0}^{n-1}\gamma^2(Y_h(k))h - \sigma_*^2\Big| \le \sum_{i\in\mathbb{S}}\gamma^2(i)\Big|\frac{V_i(n)}{n} - \pi_i\Big|$$

and hence, since the sum has finitely many terms,

$$\limsup_{n\to\infty} \frac{nh}{\sqrt{2nh\log\log n}}\Big|\frac{1}{nh}\sum_{k=0}^{n-1}\gamma^2(Y_h(k))h - \sigma_*^2\Big|$$
$$\le \sum_{i\in\mathbb{S}}\gamma^2(i)\sqrt{h}\limsup_{n\to\infty}\frac{n}{\sqrt{2n\log\log n}}\Big|\frac{V_i(n)}{n} - \pi_i\Big|. \tag{4.3.26}$$

Finally, using (4.3.25) we get the desired result. □

Before we can prove Theorem 4.2.5 we need to state and prove two useful results in the form of Lemma 4.3.6 and Lemma 4.3.7.

Lemma 4.3.6. *Suppose W is an irreducible Markov chain with finite state space $\mathbb{S}$ and matrix of one–step transition probabilities P, and moreover that W is stationary. Let $H \in \mathbb{S}$ and $\Delta \in \mathbb{N}$. Define for $n \ge 1$*

$$A_n = \{W(j) = H \text{ for all } j \in \{(n-1)\Delta, \dots, n\Delta - 1\}\}.$$

Then

$$\lim_{m\to\infty} Cov(I_{A_n}, I_{A_{n+m}}) = 0. \tag{4.3.27}$$

The proof is almost identical to the analogous result, Lemma 3.6.2, in the continuous–time case and hence is omitted.

Lemma 4.3.7. *Suppose that* $\Delta \geq 1$ *and that*

$$\xi = \big(\xi(l)\big)_{l\geq 0} \text{ is a sequence of independent standard normal random variables} \tag{4.3.28}$$

and that Y_h *is a discrete Markov chain with state space* $\mathbb{S} = \{1, \dots, N\}$ *such that* $H \in \mathbb{S}$ *obeys* $\sigma_H = \max_{j\in\mathbb{S}} |\gamma(j)|$. *Let* $V_\Delta(n)$ *be defined by* (4.2.17). *Then*

$$\limsup_{n\to\infty} \frac{|V_\Delta(n)|}{\sqrt{2\log n}} \leq \sigma_H\sqrt{\Delta}, \quad a.s. \tag{4.3.29}$$

Proof. By (4.2.17) and the fact that $\sigma_H = \max_{j\in\mathbb{S}} |\gamma(j)| > 0$ we have, for $n \geq \Delta$

$$|V_\Delta(n)| \leq \sum_{l=n-\Delta}^{n-1} \sigma_H |\xi(l+1)|.$$

Let $\lambda > 0$. Then as $\xi(l)$ are independent and identically distributed we have

$$\mathbb{E}[e^{\lambda|V_\Delta(n)|}] \leq \mathbb{E}\left[e^{\lambda\sigma_H \sum_{l=n-\Delta}^{n-1} |\xi(l+1)|}\right] = \prod_{l=n-\Delta}^{n-1} \mathbb{E}\left[e^{\lambda\sigma_H|\xi(l+1)|}\right] = \left(\mathbb{E}\left[e^{\lambda\sigma_H|\xi|}\right]\right)^{\Delta},$$

where ξ is normally distributed with zero mean and unit variance. Note that since $x \mapsto e^{\alpha|x|}e^{-x^2/2}$ is even for any $\alpha > 0$, we have

$$\begin{aligned}\mathbb{E}[e^{\alpha|\xi|}] &= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{\alpha|x|}e^{-x^2/2}\,dx = 2\frac{1}{\sqrt{2\pi}}\int_0^{\infty} e^{\alpha x - x^2/2}\,dx \\ &= 2e^{\alpha^2/2}\frac{1}{\sqrt{2\pi}}\int_{-\alpha}^{\infty} e^{-u^2/2}\,du \leq 2e^{\alpha^2/2}.\end{aligned}$$

Thus $\mathbb{E}[e^{\lambda|V_\Delta(n)|}] \leq 2^\Delta e^{\Delta\lambda^2\sigma_H^2/2}$ for $n \geq \Delta$. Let $x_n > 0$; then by Markov's inequality we have

$$\mathbb{P}[|V_\Delta(n)| \geq x_n] \leq e^{-\lambda x_n}\mathbb{E}[e^{\lambda|V_\Delta(n)|}] \leq e^{-\lambda x_n} 2^\Delta e^{\Delta\lambda^2\sigma_H^2/2}.$$

Set $\lambda = x_n/(\Delta\sigma_H^2)$, so that

$$\mathbb{P}[|V_\Delta(n)| \geq x_n] \leq 2^\Delta e^{-\lambda x_n} e^{\Delta\lambda^2\sigma_H^2/2} = 2^\Delta e^{-x_n^2/(2\Delta\sigma_H^2)}.$$

Let $\varepsilon > 0$, and set $x_n = \sqrt{2(1+\varepsilon)}\sqrt{\Delta}\sigma_H\sqrt{\log(n+1)}$ for $n \geq \Delta$. Then

$$\mathbb{P}\Big[|V_\Delta(n)| \geq \sqrt{2(1+\varepsilon)}\sqrt{\Delta}\sigma_H\sqrt{\log(n+1)}\Big] \leq 2^\Delta(n+1)^{-(1+\varepsilon)}, \quad n \geq \Delta.$$

Therefore by the Borel–Cantelli lemma we have for every $\varepsilon > 0$ that there is an almost sure event Ω_ε such that

$$\limsup_{n\to\infty} \frac{|V_\Delta(n)|}{\sigma_H\sqrt{2\Delta\log(n+1)}} \leq \sqrt{1+\varepsilon}, \quad \text{on } \Omega_\varepsilon.$$

Let $\Omega^* = \cap_{\varepsilon\in(0,1)\cap\mathbb{Q}}\Omega_\varepsilon$; then Ω^* is an almost sure event and we have

$$\limsup_{n\to\infty} \frac{|V_\Delta(n)|}{\sqrt{2\log n}} \leq \sigma_H\sqrt{\Delta}, \quad \text{a.s. on } \Omega^*,$$

as required in (4.3.29). □

We are now in a position to prove Theorem 4.2.5.

Proof of Theorem 4.2.5. Note that the upper bound obtained in Lemma 4.3.7, together with (4.2.18) and the fact that f and γ are globally bounded, gives the inequality

$$\limsup_{n\to\infty} \frac{|R_{\delta,h}(n)|}{\sqrt{2\log n}} \leq \sigma_H\sqrt{h}\sqrt{\Delta}, \quad \text{a.s.} \tag{4.3.30}$$

It remains to prove that

$$\limsup_{n\to\infty} \frac{|R_{\delta,h}(n)|}{\sqrt{2\log n}} \geq \sigma_H\sqrt{h}\sqrt{\Delta}, \quad \text{a.s.} \tag{4.3.31}$$

By (4.2.18) we have

$$\limsup_{n\to\infty} \frac{|R_{\delta,h}(n)|}{\sqrt{2\log n}} = \sqrt{h}\limsup_{n\to\infty} \frac{|V_\Delta(n)|}{\sqrt{2\log n}}, \quad \text{a.s.}$$

In particular, with $U_n := V_\Delta(n\Delta) = \sum_{l=(n-1)\Delta}^{n\Delta-1} \gamma(Y_h(l))\xi(l+1)$, we have

$$\limsup_{n\to\infty} \frac{|R_{\delta,h}(n)|}{\sqrt{2\log n}} \geq \sqrt{h}\limsup_{n\to\infty} \frac{|V_\Delta(n\Delta)|}{\sqrt{2\log n}} = \sqrt{h}\limsup_{n\to\infty} \frac{|U_n|}{\sqrt{2\log n}}. \tag{4.3.32}$$

Since Y_h is stationary, the probability that $Y_h((n-1)\Delta) = H$ is π_H. Define the event $A_n := \{Y_h(j) = H, \text{ for all } j \in \{(n-1)\Delta, \dots, n\Delta - 1\}\}$. The process $\{I_{A_n} : n \geq 1\}$ is stationary so it can be shown in a similar manner to the proof of Lemma 3.6.2 that $\mathbb{P}[A_n] = \mathbb{E}[I_{A_n}] =: \pi(\Delta)$ and $\text{Cov}(I_{A_n}, I_{A_{n+m}}) \to 0$ as $m \to \infty$. Define $T_n = \sum_{j=1}^{n} I_{A_j}$. Then using the same argument as used in the proof of Theorem 3.4.4 we have $T_n/n \to \pi(\Delta)$ as $n \to \infty$ a.s. Let $L_n = \min\{l \geq n : \sum_{j=1}^{l} I_{A_j} = n\}$. By definition $I_{A_{L_n}} = 1$. Then if we consider the collection of $\{U_j : j = 1, \dots, n\}$ for which $I_{A_j} = 1$ we have $\max_{1\leq j\leq n} |U_j| \geq \max_{1\leq k\leq T_n} |U_{L_k}|$.

Next, if $I_{A_n} = 1$ then $Y_h(j) = H$ for all $j \in \{(n-1)\Delta, \dots, n\Delta - 1\}$ and we have $U_n = \sum_{l=(n-1)\Delta}^{n\Delta-1} \gamma(H)\xi(l+1) = \gamma(H)\sum_{l=(n-1)\Delta}^{n\Delta-1} \xi(l+1)$. Without loss of generality we consider the case where $\gamma(H) > 0$. If $\gamma(H) < 0$ then we can redefine the standard normal random variables as $\xi_- = -\xi$ and proceed as in the case $\gamma(H) > 0$. Hence we get

$$\begin{aligned}\max_{1\leq j\leq n} |U_j| \geq \max_{1\leq k\leq T_n} |U_{L_k}| &= \max_{1\leq k\leq T_n} \left| \sigma_H \sum_{l=(L_k-1)\Delta}^{L_k\Delta-1} \xi(l+1) \right| \\ &= \sigma_H \max_{1\leq k\leq T_n} \left| \sum_{l=(L_k-1)\Delta}^{L_k\Delta-1} \xi(l+1) \right|.\end{aligned}$$

Define $\zeta(k) := \sum_{l=(L_k-1)\Delta}^{L_k\Delta-1} \xi(l+1)$. Then we have

$$\begin{aligned}\limsup_{n\to\infty} \frac{\max_{1\leq j\leq n} |U_j|}{\sqrt{2\log n}} &\geq \sigma_H \limsup_{n\to\infty} \frac{\max_{1\leq k\leq T_n} |\zeta(k)|}{\sqrt{2\log T_n}} \cdot \sqrt{\frac{\log T_n}{\log n}} \\ &= \sigma_H \limsup_{n\to\infty} \frac{\max_{1\leq k\leq T_n} |\zeta(k)|}{\sqrt{2\log T_n}} = \sigma_H \limsup_{n\to\infty} \frac{\max_{1\leq k\leq n} |\zeta(k)|}{\sqrt{2\log n}}, \quad (4.3.33)\end{aligned}$$

where we used the fact that $T_n \to \infty$ as $n \to \infty$ a.s. Since ξ and Y_h are independent, it follows that $L = \{L_n : n \geq 1\}$ and ξ are independent. Let $m \in \mathbb{N}$ and $k_1 < k_2 < \dots k_m$.

Then, because $L_{k+1} - L_k \geq 1$, we have

$$\begin{aligned}
&\mathbb{P}[\zeta(k_1) \leq x_1, \zeta(k_2) \leq x_2, \dots, \zeta(k_m) \leq x_m] \\
&= \sum_{n_1 < n_2 \cdots < n_m} \mathbb{P}\big[\zeta(k_1) \leq x_1, \zeta(k_2) \leq x_2, \dots, \zeta(k_m) \leq x_m \big| \\
&\qquad L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_m} = n_m\big] \times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_m} = n_m] \\
&= \sum_{n_1 < n_2 \cdots < n_m} \mathbb{P}\Bigg[\sum_{l=(L_{k_1}-1)\Delta}^{L_{k_1}\Delta - 1} \xi(l+1) \leq x_1, \dots, \sum_{l=(L_{k_m}-1)\Delta}^{L_{k_m}\Delta-1} \xi(l+1) \leq x_m \Bigg| \\
&\qquad L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_m} = n_m\Bigg] \times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_m} = n_m] \\
&= \sum_{n_1 < n_2 \cdots n_{m-1} < n_m} \mathbb{P}\Bigg[\sum_{l=(n_1-1)\Delta}^{n_1\Delta - 1} \xi(l+1) \leq x_1, \dots, \sum_{l=(n_m-1)\Delta}^{n_m\Delta-1} \xi(l+1) \leq x_m \Bigg| \\
&\qquad L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_m} = n_m\Bigg] \times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_m} = n_m] \\
&= \sum_{n_1 < n_2 \cdots < n_m} \mathbb{P}\Bigg[\sum_{l=(n_1-1)\Delta}^{n_1\Delta - 1} \xi(l+1) \leq x_1, \dots, \sum_{l=(n_m-1)\Delta}^{n_m\Delta-1} \xi(l+1) \leq x_m \Bigg] \\
&\qquad \times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_m} = n_m].
\end{aligned}$$

where we have used the independence of the ξ's and L's at the last step. Since $\Delta \geq 1$ and in the sum we have $1+n_i \leq n_{i+1}$, it follows that $(n_{i+1}-1)\Delta > n_i\Delta - 1$. Therefore, because there is no overlap from one sum to the next, each of the random variables $\sum_{l=(n_i-1)\Delta}^{n_i\Delta-1} \xi(l+1)$ for $i = 1, \dots, m$ are independent and identically normally distributed with zero mean and variance Δ. Therefore if Φ_Δ is the distribution function of a standardised normal random variable, we have

$$\begin{aligned}
&\mathbb{P}[\zeta(k_1) \leq x_1, \zeta(k_2) \leq x_2, \dots, \zeta(k_m) \leq x_m] \\
&= \sum_{n_1 < n_2 \cdots < n_m} \mathbb{P}\left[\sum_{l=(n_1-1)\Delta}^{n_1\Delta-1} \xi(l+1) \leq x_1\right] \cdots \mathbb{P}\left[\sum_{l=(n_m-1)\Delta}^{n_m\Delta-1} \xi(l+1) \leq x_m\right] \\
&\qquad \times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_m} = n_m] \\
&= \sum_{n_1 < n_2 \cdots < n_m} \Phi_\Delta(x_1) \cdots \Phi_\Delta(x_m) \times \mathbb{P}[L_{k_1} = n_1, L_{k_2} = n_2, \dots, L_{k_m} = n_m] = \prod_{i=1}^{m} \Phi_\Delta(x_i).
\end{aligned}$$

Thus $\{\zeta(k) : k \geq 1\}$ is a sequence of independent and identically distributed normal random variables with mean zero and variance Δ. Therefore, by Lemma 1.0.1 and Lemma 3.1 in [6],

$$\limsup_{n\to\infty} \frac{\max_{1\leq k\leq n} |\zeta(k)|}{\sqrt{2\log n}} = \limsup_{n\to\infty} \frac{|\zeta(n)|}{\sqrt{2\log n}} = \sqrt{\Delta}, \quad \text{a.s.}$$

Hence, combining this with (4.3.33) gives

$$\Lambda := \limsup_{n\to\infty} \frac{\max_{1\leq j\leq n} |U_j|}{\sqrt{2\log n}} \geq \sigma_H \sqrt{\Delta}, \quad \text{a.s.}$$

This implies that

$$\limsup_{n\to\infty} \frac{|U_n|}{\sqrt{2\log n}} = \Lambda \geq \sigma_H \sqrt{\Delta}, \quad \text{a.s.} \tag{4.3.34}$$

Combining (4.3.30), (4.3.32) and (4.3.34) gives (4.3.31) and hence (4.2.19). □

Chapter 5

A Discrete Exponential Martingale Inequality for Martingales driven by Gaussian Sequences

5.1 Introduction

The aim of this chapter is to develop a discrete version of the exponential martingale inequality (EMI) which can be applied specifically to martingales driven by Gaussian sequences. A comparison of Theorem 4.2.2 in Chapter 4 and Corollary 3.3.1 in Chapter 3 reveals that the discrete version of the result is inferior due to an extra factor of two (the extra $\sqrt{e}$ naturally arises from the alternative method of proof). Upon analysing the proof of Theorem 4.2.2, it becomes clear that the extra factor of 2 arises from the duplication of the σ_*^2 term in (4.3.16). This duplication is directly linked to the sum of the two quadratic variations in the discrete EMI that was used in Chapter 4. It seems reasonable then to believe that if the discrete EMI depended *only* on the predictable quadratic variation then the duplication in (4.3.16) would not appear and we would not get the extra factor of 2 in the final estimate.

This motivates the need for a discrete EMI, containing only one quadratic variation term, which can be used to estimate more accurately the size of the large fluctuations of the solutions of discretisations of stochastic differential equations (SDEs). The EMI, together with a Gronwall inequality argument, was first successfully applied to estimate these large fluctuations in Mao [52]. More recently, extensions of Mao's result to a wider class of SDEs appear in [5], and to stochastic delay differential equations in [56]. Mao's results are collected in [54, Chapter 2, Section 5].

This EMI–Gronwall technique can be readily applied (even to non–autonomous SDEs) and gives excellent upper bounds on the size of fluctuations of SDEs, as evidenced by results obtained in [5] by both comparison principle and EMI arguments. The comparison results rely on the powerful theorem of Motoo [65] which we used in Chapters 2 and 3 to give very precise upper and lower bounds on the size of the pathwise large fluctuations of scalar autonomous SDEs which possess recurrent or asymptotically stationary solutions. However, the proof of Motoo's theorem hinges on an analysis of the excursions of solutions of SDEs which cannot easily be applied in discrete time, and therefore to discretisations of the SDE.

However, based on the evidence of Chapter 4 and by scrutinising the proofs in [54, Chapter 2, Section 5], it is apparent that asymptotic estimates for the large fluctuations of an Euler–Maruyama scheme would yield results consistent with those obtained in the continuous–time case, provided an appropriate version of the EMI is employed. In this case, we would deem an EMI to be appropriate if the estimate on the martingale depended solely on its predictable quadratic variation, and allowed us to recover an estimate consistent with that obtained in continuous time. We are however unaware of a result in the literature that fulfills these two requirements. A significant literature on exponential inequalities already exists, and we refer the reader to work of De La Peña and co-authors [18, 19, 20].

In our main result (Theorem 5.2.4) we establish such an appropriate EMI for discrete–time martingales which are driven by Gaussian sequences. We specialise to this class of martingales because they can be used to approximate Itô integrals in stochastic Euler methods. We propose, in Chapter 6, to apply Theorem 5.2.4 systematically to study the large fluctuations of stochastic Euler schemes. Results which apply *existing* EMIs to stochastic Euler schemes, and which highlight current limitations, were presented in Chapter 4.

As we believe that Theorem 5.2.4 may be of independent interest and utility in stochastic numerical analysis in particular, we state and prove it in a more general context here. In particular, it might be of interest to extend our result to martingales driven by e.g., heavy tailed sequences of independent and identically distributed random variables with a known moment generating function. In this way, one might develop a useful tool to estimate the large fluctuations of discretisations of stochastic differential equations driven by Lévy processes, by once again imitating the Gronwall–EMI programme outlined in the works cited above.

The chapter is organised as follows. A synopsis and discussion of existing EMIs as well as a special case of the EMI for martingales driven by Gaussian sequences is given in Section 5.2. The proof of the main theorem is given in Section 5.3 while an alternative proof is given in Section 5.4.

5.2 Statement and discussion of main results

5.2.1 Existing Exponential Martingale Inequalities

In this section we state some existing and well–known exponential martingale inequalities which we will compare and contrast with our main result.

In the first instance, we consider a result for continuous–time. Let $d \in \mathbb{N}$. Denote the complete filtered probability space by $(\tilde{\Omega}, \tilde{\mathcal{F}}, (\tilde{\mathcal{F}}(t))_{t\geq 0}, \tilde{\mathbb{P}})$. Here the filtration $\tilde{\mathcal{F}}(t)$ is such that $B = (B_1, \ldots, B_d)$ is a d–dimensional Brownian motion adapted to $(\tilde{\mathcal{F}}(t))_{t\geq 0}$. We denote by $\mathcal{L}^2(\mathbb{R}^+; \mathbb{R}^d)$ the space of $\mathbb{R}^d$–valued measurable and $(\tilde{\mathcal{F}}(t))_{t\geq 0}$–adapted processes V obeying $\int_0^T |V(s)|^2\, ds < +\infty$ $\tilde{\mathbb{P}}$–a.s. for every $T > 0$. If M is a local martingale in

$\mathcal{L}^2(\mathbb{R}^+;\mathbb{R})$ given by

$$M(t) = \sum_{i=1}^{d} \int_0^t U_i(s)\, dB_i(s), \quad t \geq 0, \tag{5.2.1}$$

where each U_i is in $\mathcal{L}^2(\mathbb{R}^+;\mathbb{R})$, then the quadratic variation of M is the process denoted by $\langle M \rangle$ where

$$\langle M \rangle(t) = \sum_{i=1}^{d} \int_0^t U_i(s)^2\, ds, \quad t \geq 0.$$

We are now in a position to state the well–known continuous–time exponential martingale inequality, found for example in [54].

Theorem 5.2.1. *Let $U = (U_1, \ldots, U_d) \in \mathcal{L}^2(\mathbb{R}^+;\mathbb{R}^d)$ and let $B = (B_1, \ldots, B_d)$ be a d–dimensional $\tilde{\mathcal{F}}(t)$–Brownian motion. Let M be the local martingale in $\mathcal{L}^2(\mathbb{R}^+;\mathbb{R})$ given by* (5.2.1) *with quadratic variation $\langle M \rangle$. Then for any $T, \alpha, \beta > 0$ we have*

$$\tilde{\mathbb{P}}\Big[\sup_{0 \leq t \leq T} \big\{ M(t) - \frac{\alpha}{2} \langle M \rangle(t) \big\} \geq \beta \Big] \leq e^{-\alpha\beta}.$$

On the other hand, exponential martingale inequalities also exist for discrete–time martingales. The following example can be developed from work of Bercu and Touati [13]. In this case we work on the complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}(n))_{n\in\mathbb{N}_0}, \mathbb{P})$. We denote by $\ell^2(\mathbb{N}_0;\mathbb{R}^d)$ the space of $\mathbb{R}^d$–valued measurable $(\mathcal{F}(n))_{n\in\mathbb{N}_0}$–adapted processes $V = \{V(n) : n \in \mathbb{N}_0\}$ obeying $\sum_{n=0}^N |V(n)|^2 < +\infty$ $\mathbb{P}$–a.s. for every $N \in \mathbb{N}$. If $M \in \ell^2(\mathbb{N}_0;\mathbb{R})$ is an $\mathcal{F}(n)$–martingale, which is null at zero, we define its *predictable quadratic variation* $\langle M \rangle = \{\langle M \rangle(n) : n \in \mathbb{N}\}$ by

$$\langle M \rangle(n) = \sum_{k=1}^{n} \mathbb{E}\big[(M(k) - M(k-1))^2 \big| \mathcal{F}(k-1)\big],$$

and its *total quadratic variation* $[M] = \{[M](n) : n \in \mathbb{N}\}$ by

$$[M](n) = \sum_{k=1}^{n} (M(k) - M(k-1))^2.$$

Theorem 5.2.2. *Let $M \in \ell^2(\mathbb{N}_0;\mathbb{R})$ be an $\mathcal{F}(n)$–martingale, null at zero, which has predictable quadratic variation $\langle M\rangle$ and total quadratic variation $[M]$. Then, for any $\alpha,\beta>0$ and $N\in\mathbb{N}$ we have*

$$\mathbb{P}\Big[\max_{1\le n\le N}\big\{M(n)-\frac{\alpha}{2}\big([M](n)+\langle M\rangle(n)\big)\big\}\ge\beta\Big]\le e^{-\alpha\beta}.$$

An outline proof of this theorem was given in Chapter 4. We note that in Theorem 5.2.2 the estimate on M depends on both the predictable and the total quadratic variations; by contrast, in Theorem 5.2.1 the estimate on M depends only on one quadratic variation.

5.2.2 Statement of Main Result

We now develop a discrete exponential martingale inequality which depends only on the predictable quadratic variation. Let $(\Omega,\mathcal{F},(\mathcal{F}(n))_{n\ge0},\mathbb{P})$ be a complete filtered probability space and suppose that $U=\{U(n):n\ge0\}$ is an $\mathbb{R}^d$–valued and $\mathcal{F}(n)$–adapted process given by $U(n)=\sum_{i=1}^d U_i(n)\underline{e}_i$, where $\underline{e}_i$ is the unit vector with i^{th} entry one and zeros elsewhere.

Assumption 5.2.3. *We suppose that $\xi=\{\xi(n):n\ge1\}$ is an $\mathbb{R}^d$–valued and $\mathcal{F}(n)$–adapted process denoted by $\xi(n)=\sum_{i=1}^d\xi_i(n)\underline{e}_i$ where the vectors $\big(\xi(n)\big)_{n\ge1}$ are independent and $\big(\xi_i(n)\big)_{i=1,\dots,d}$ are independent and identically distributed standard normal random variables for each fixed $n\ge1$.*

This assumption on the independent Gaussian sequences mimics the presence of a d–dimensional Brownian motion in Theorem 5.2.1. In this sense, if $h>0$, $\sqrt{h}\xi_i(n+1)$ can be seen as the increment of Brownian motion over the period $[nh,(n+1)h]$, i.e., $\xi_i(n+1)=\big(B_i((n+1)h)-B_i(nh)\big)/\sqrt{h}$.

In many situations, it is convenient to let the filtration $(\mathcal{F}(n))_{n\ge0}$ be that which is naturally generated by ξ i.e., $\mathcal{F}(n)=\sigma\big(\{\xi(j):1\le j\le n\}\big)$ so that $(\mathcal{F}(n))_{n\ge0}$ is the natural filtration of ξ. Certainly, if we choose the initial value $U(0)$ to be deterministic, then we will use such a filtration $(\mathcal{F}(n))_{n\geq 0}$. However, to maintain generality we may allow the initial value $U(0)$ to be random and independent of $(\mathcal{F}(n))_{n\geq 0}$ in which case we define the filtration to be the combination of the natural filtration of ξ along with the σ–algebra generated by the initial value $U(0)$.

Theorem 5.2.4. *Suppose ξ satisfies Assumption 5.2.3 and let $U \in \ell^2(\mathbb{N}_0;\mathbb{R}^d)$. Define the local martingale $M \in \ell^2(\mathbb{N}_0;\mathbb{R})$, which is null at zero, by*

$$M(n) = \sum_{j=0}^{n-1}\sum_{i=1}^{d} U_i(j)\xi_i(j+1), \quad n \geq 1.$$

Then for any $\alpha, \beta > 0$ and $N \in \mathbb{N}$, we have

$$\mathbb{P}\Big[\max_{1\leq n\leq N}\{M(n) - \frac{\alpha}{2}\langle M\rangle(n)\} \geq \beta\Big] \leq e^{-\alpha\beta}.$$

Remark 5.2.1. We prove Theorem 5.2.4 by showing directly that the discrete–time martingale M has all of the necessary properties in order for the above conclusion to hold. Alternatively, Theorem 5.2.4 can be proven using a shorter, less direct approach which involves sampling a *continuous–time* martingale driven by a Brownian motion to which the continuous–time EMI can be applied. For completeness we give *both* proofs.

Remark 5.2.2. By mimicing the direct proof of Theorem 5.2.4 it is possible to formulate EMIs when the white noise sequence ξ is a sequence of independent and identically distributed zero mean random variables, each of which has a common moment generating function φ defined on an interval $I \subseteq \mathbb{R}$. We state a representative result here in the scalar case. Let $(U(j))_{j\geq 0}$ be adapted to the filtration generated by ξ and define the local martingale $M \in \ell^2(\mathbb{N}_0;\mathbb{R})$, which is null at zero, by

$$M(n) = \sum_{j=0}^{n-1} U(j)\xi(j+1), \quad n \geq 1.$$

Let $\alpha > 0$, $N \in \mathbb{N}$ and suppose that $\alpha|U(n)| \in I$ for $n = 0, \dots, N$ a.s. Then one could prove in a similar manner to the proof of Theorem 5.2.4 that

$$\mathbb{P}\Big[\max_{1\le n\le N}\{M(n) - \frac{1}{\alpha}\sum_{j=0}^{n-1}\log\varphi(\alpha U(j))\} \ge \beta\Big] \le e^{-\alpha\beta}.$$

In fact, we recover Theorem 5.2.4 in the scalar case by noticing that $\varphi : \mathbb{R} \to \mathbb{R}$ given by $\varphi(\alpha) = e^{\alpha^2/2}$ is the moment generating function of a standard normal random variable.

5.3 Proof of Theorem 5.2.4

For $\theta \in \mathbb{N}$, define the stopping time

$$\tau_\theta = \inf\big\{n \ge 0 : A(n) > \theta \quad \text{or} \quad \sum_{i=1}^{d}|U_i(n)| > \theta\big\} \tag{5.3.1}$$

where $A(n) = \sum_{j=0}^{n-1}\big|\sum_{i=1}^{d}U_i(j)\xi_i(j+1)\big|$ for $n \ge 1$ and $A(0) = 0$. Note that since $A(n)$ is $\mathcal{F}(n)$–measurable and $\sum_{i=1}^{d}|U_i(n)|$ is $\mathcal{F}(n)$–measurable, it follows that $\{\tau_\theta \le n\} \in \mathcal{F}(n)$, so τ_θ is indeed a stopping time. Let $S_1(j) := \sum_{i=1}^{d}U_i(j)\xi_i(j+1)$ and $S_2(j) := \sum_{i=1}^{d}U_i^2(j)$ and define for $n \ge 0$,

$$X_\theta(n) = \exp\Big[I_{\{\tau_\theta>0\}}1_{\mathbb{N}}(n)\Big(\alpha\sum_{j=0}^{n\wedge\tau_\theta-1}S_1(j) - \frac{1}{2}\alpha^2\sum_{j=0}^{n\wedge\tau_\theta-1}S_2(j)\Big)\Big], \tag{5.3.2}$$

where α is a positive constant. To show that X_θ is a martingale we show that $\mathbb{E}[X_\theta(n)] < +\infty$ and that $\mathbb{E}[X_\theta(n+1)|\mathcal{F}(n)] = X_\theta(n)$. In the trivial cases where either $\tau_\theta = 0$ or $n = 0$ we have that $X_\theta(n) = 1$ and so clearly $\mathbb{E}[X_\theta(n)] < +\infty$. Otherwise we have $\tau_\theta > 0,\ n \ge 1$ and

$$X_\theta(n) = \exp\Big[\alpha\sum_{j=0}^{n\wedge\tau_\theta-1}S_1(j) - \frac{1}{2}\alpha^2\sum_{j=0}^{n\wedge\tau_\theta-1}S_2(j)\Big] \le \exp\big[\alpha\sum_{j=0}^{n\wedge\tau_\theta-1}|S_1(j)|\big].$$

Here there are two cases to consider: $\tau_\theta > n$ or $\tau_\theta \le n$. If $\tau_\theta > n$ then

$$\sum_{j=0}^{n\wedge\tau_\theta-1}|S_1(j)| = \sum_{j=0}^{n-1}|S_1(j)| = \sum_{j=0}^{n-1}\big|\sum_{i=1}^{d}U_i(j)\xi_i(j+1)\big| = A(n) \le \theta,$$

by the definition of the stopping time τ_θ. Thus $X_\theta(n) \le e^{\alpha\theta}$ when $\tau_\theta > n$. If $\tau_\theta \le n$ then

$$\sum_{j=0}^{n\wedge\tau_\theta-1} |S_1(j)| = \sum_{j=0}^{\tau_\theta-1} \Big|\sum_{i=1}^{d} U_i(j)\xi_i(j+1)\Big| = A(\tau_\theta-1) + \Big|\sum_{i=1}^{d} U_i(\tau_\theta-1)\xi_i(\tau_\theta)\Big|$$
$$\le \theta + \sum_{i=1}^{d} |U_i(\tau_\theta-1)||\xi_i(\tau_\theta)| \le \theta + \theta\sum_{i=1}^{d}|\xi_i(\tau_\theta)|,$$

where in the last step we have used the fact that $|U_i(\tau_\theta-1)| \le \theta$, since the definition of τ_θ gives $\sum_{i=1}^{d}|U_i(\tau_\theta-1)| \le \theta$. Therefore, for $\tau_\theta \le n$,

$$X_\theta(n) \le \exp\Big[\alpha\theta + \alpha\theta\sum_{i=1}^{d}|\xi_i(\tau_\theta)|\Big] \le \exp\Big[\alpha\theta + \alpha\theta\max_{1\le j\le n}\sum_{i=1}^{d}|\xi_i(j)|\Big]. \tag{5.3.3}$$

Since $X_\theta(n) \le e^{\alpha\theta}$ when $\tau_\theta > n$ and $\alpha\theta\max_{1\le j\le n}\sum_{i=1}^{d}|\xi_i(j)|$ is positive, it follows that the estimate in (5.3.3) holds for all values of n. Thus it remains to show that $\mathbb{E}[\exp\{\alpha\theta\max_{1\le j\le n}\sum_{i=1}^{d}|\xi_i(j)|\}]$ is finite. Define $\zeta(j) := \sum_{i=1}^{d}|\xi_i(j)|$. Then $\zeta(j)_{j\ge 0}$ are independent random variables since $\xi_i(n), \xi_{i'}(m)$ are independent for $m \ne n$ and for all i, i'. Note that since each $\xi_i(j)$ is independent and normally distributed we have

$$\mathbb{E}\big[e^{\alpha\theta|\xi_i(j)|}\big] = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{\alpha\theta|x|}e^{-x^2/2}\,dx = \frac{2}{\sqrt{2\pi}} e^{\alpha^2\theta^2/2}\int_{-\alpha\theta}^{+\infty} e^{-u^2/2}\,du$$
$$\le \frac{2}{\sqrt{2\pi}} e^{\alpha^2\theta^2/2}\int_{-\infty}^{+\infty} e^{-u^2/2}\,du = 2e^{\alpha^2\theta^2/2} < +\infty,$$

and as a result,

$$\mathbb{E}\big[\exp\{\alpha\theta\zeta(j)\}\big] = \prod_{i=1}^{d}\mathbb{E}\big[\exp\{\alpha\theta|\xi_i(j)|\}\big] \le \prod_{i=1}^{d} 2e^{(\alpha\theta)^2/2} < +\infty.$$

Therefore,

$$\mathbb{E}\big[\exp\{\alpha\theta\max_{1\le j\le n}\zeta(j)\}\big] \le \mathbb{E}\big[\exp\{\alpha\theta\sum_{j=1}^{n}\zeta(j)\}\big] = \prod_{j=1}^{n}\mathbb{E}\big[\exp\{\alpha\theta\zeta(j)\}\big]$$
$$\le \prod_{j=1}^{n}\prod_{i=1}^{d} 2e^{(\alpha\theta)^2/2} = \big(2^d e^{d(\alpha\theta)^2/2}\big)^n < +\infty.$$

Thus, we have shown that $\mathbb{E}[X_\theta(n)] < +\infty$. Now we aim to show that $\mathbb{E}[X_\theta(n+1)|\mathcal{F}(n)] = X_\theta(n)$. From (5.3.2) we have

$$\begin{aligned}
X_\theta(n+1) &= \exp\Big[I_{\{\tau_\theta>0\}}1_{\mathbb{N}}(n+1)\Big(\alpha\sum_{j=0}^{(n+1)\wedge\tau_\theta-1} S_1(j) - \frac{1}{2}\alpha^2\sum_{j=0}^{(n+1)\wedge\tau_\theta-1} S_2(j)\Big)\Big] \\
&= X_\theta(n)\Big\{\exp\Big[I_{\{\tau_\theta>0\}}\alpha\big(1_{\mathbb{N}}(n+1)\sum_{j=0}^{(n+1)\wedge\tau_\theta-1} S_1(j) - 1_{\mathbb{N}}(n)\sum_{j=0}^{n\wedge\tau_\theta-1} S_1(j)\big) \\
&\quad - \frac{\alpha^2}{2}I_{\{\tau_\theta>0\}}\big(1_{\mathbb{N}}(n+1)\sum_{j=0}^{(n+1)\wedge\tau_\theta-1} S_2(j) - 1_{\mathbb{N}}(n)\sum_{j=0}^{n\wedge\tau_\theta-1} S_2(j)\big)\Big]\Big\}. \quad (5.3.4)
\end{aligned}$$

For $n > 0$ there are three cases to consider. If $1 \le \tau_\theta \le n$ it follows that $X_\theta(n+1) = X_\theta(n)$. Moreover, if $\tau_\theta = 0$ then $X_\theta(n+1) = 1 = X_\theta(n)$ also. If $\tau_\theta \ge n+1$ then $X_\theta(n+1) = X_\theta(n)\exp[\alpha S_1(n) - \frac{1}{2}\alpha^2 S_2(n)]$. Therefore, for $n > 0$ we can write (5.3.4) more concisely as

$$X_\theta(n+1) = X_\theta(n)\exp\big[I_{\{\tau_\theta>n\}}(\alpha S_1(n) - \frac{1}{2}\alpha^2 S_2(n))\big].$$

Moreover, if we define $V_i(n) := U_i(n)I_{\{\tau_\theta>n\}}$ for $n \ge 0$ and observe that $I^2_{\{\tau_\theta>n\}} = I_{\{\tau_\theta>n\}}$ we have

$$X_\theta(n+1) = X_\theta(n)\exp\big[\alpha\sum_{i=1}^d V_i(n)\xi_i(n+1) - \frac{1}{2}\alpha^2\sum_{i=1}^d V_i^2(n)\big], \quad n \ge 1.$$

Note that $V_i(n)$ is $\mathcal{F}(n)$–measurable because $U_i(n)$ is $\mathcal{F}(n)$–measurable and $I_{\{\tau_\theta>n\}}$ is $\mathcal{F}(n)$–measurable due to the fact that τ_θ is a stopping time. Moreover, if $\tau_\theta > n$ then by the definition of τ_θ we have $\sum_{i=1}^d |U_i(n)| \le \theta$ and thus $|V_i(n)| = |U_i(n)| \le \theta$ for each $i = 1, \dots, d$. If $\tau_\theta \le n$ then $V_i(n) = 0 \le \theta$ for each $i = 1, \dots, d$. Hence, $|V_i(n)| \le \theta$ for all $n \ge 0$ and for each $i = 1, \dots, d$. Now, since $V_i(n)$ is bounded and $\mathcal{F}(n)$–measurable and $\mathbb{E}[X_\theta(n)] < +\infty$, we have for $n \ge 1$,

$$\mathbb{E}[X_\theta(n+1)|\mathcal{F}(n)] = X_\theta(n)\exp\big(-\frac{\alpha^2}{2}\sum_{i=1}^d V_i^2(n)\big)\mathbb{E}\big[\exp\big(\alpha\sum_{i=1}^d V_i(n)\xi_i(n+1)\big)\big|\mathcal{F}(n)\big].$$

By Assumption 5.2.3, $\big(\xi_i(n+1)\big)_{i=1}^d$ are mutually independent $N(0,1)$ random variables, so each $\xi_i(n+1)$ is independent of $\mathcal{F}(n)$. If we define the moment generating function

$\varphi_i(\lambda) := \mathbb{E}[\exp(\lambda\xi_i(n+1))] = \exp(\frac{1}{2}\lambda^2)$, then since $V_i(n)$ is bounded we have

$$\mathbb{E}[\exp\big(\alpha\sum_{i=1}^{d} V_i(n)\xi_i(n+1)\big)|\mathcal{F}(n)] = \prod_{i=1}^{d}\varphi_i(\alpha V_i(n)) = \exp\big(\frac{1}{2}\alpha^2\sum_{i=1}^{d} V_i^2(n)\big).$$

Thus, $\mathbb{E}[X_\theta(n+1)|\mathcal{F}(n)] = X_\theta(n)$ for $n \geq 1$. To deal with the case when $n = 0$, we note that

$$X_\theta(1) = \exp\Big[I_{\{\tau_\theta>0\}}\big(\alpha\sum_{j=0}^{1\wedge\tau_\theta-1} S_1(j) - \frac{1}{2}\alpha^2\sum_{j=0}^{1\wedge\tau_\theta-1} S_2(j)\big)\Big].$$

Here, there are two cases to consider. If $\tau_\theta = 0$ then $X_\theta(1) = 1$. If $\tau_\theta \geq 1$ then $X_\theta(1) = \exp[\alpha S_1(0) - \frac{1}{2}\alpha^2 S_2(0)]$. This can be written more concisely, for any τ_θ, as

$$\begin{aligned} X_\theta(1) &= \exp\Big[I_{\{\tau_\theta>0\}}\big(\alpha\sum_{i=1}^{d} U_i(0)\xi_i(1) - \frac{1}{2}\alpha^2\sum_{i=1}^{d} U_i^2(0)\big)\Big] \\ &= \exp\big[\alpha\sum_{i=1}^{d} V_i(0)\xi_i(1) - \frac{1}{2}\alpha^2\sum_{i=1}^{d} V_i^2(0)\big], \end{aligned}$$

where $V_i(0)$ is as defined earlier. So again, since $V_i(0)$ is bounded,

$$\begin{aligned} \mathbb{E}[X_\theta(1)|\mathcal{F}(0)] &= \mathbb{E}\big[\exp\big(\alpha\sum_{i=1}^{d} V_i(0)\xi_i(1) - \frac{1}{2}\alpha^2\sum_{i=1}^{d} V_i^2(0)\big)|\mathcal{F}(0)\big] \\ &= \exp\big(-\frac{\alpha^2}{2}\sum_{i=1}^{d} V_i^2(0)\big)\prod_{i=1}^{d}\varphi_i(\alpha V_i(0)) = 1 = X_\theta(0). \end{aligned}$$

Therefore, $\big(X_\theta(n)\big)_{n\geq 0}$ is a positive $\mathcal{F}(n)$–martingale. Thus it follows by Theorem 1.3.8 in [54] that

$$\mathbb{P}[\max_{0\leq n\leq N} X_\theta(n) \geq e^{\alpha\beta}] \leq e^{-\alpha\beta}\mathbb{E}[X_\theta(N)] = e^{-\alpha\beta},$$

where $\mathbb{E}[X_\theta(N)] = \mathbb{E}[X_\theta(0)] = 1$. Hence, taking logs and dividing by α we have

$$\mathbb{P}\big[\max_{0\leq n\leq N} I_{\{\tau_\theta>0\}}1_{\mathbb{N}}(n)\big(\sum_{j=0}^{n\wedge\tau_\theta-1} S_1(j) - \frac{\alpha}{2}\sum_{j=0}^{n\wedge\tau_\theta-1} S_2(j)\big) \geq \beta\big] \leq e^{-\alpha\beta}. \tag{5.3.5}$$

Define for $n \geq 0$,

$$\tilde{X}_\theta(n) = I_{\{\tau_\theta>0\}}1_{\mathbb{N}}(n)\big(\sum_{j=0}^{n\wedge\tau_\theta-1} S_1(j) - \frac{\alpha}{2}\sum_{j=0}^{n\wedge\tau_\theta-1} S_2(j)\big),$$

and define its limit

$$\lim_{\theta\to\infty} \tilde{X}_\theta(n) = 1_{\mathbb{N}}(n)\big(\sum_{j=0}^{n-1} S_1(j) - \frac{\alpha}{2}\sum_{j=0}^{n-1} S_2(j)\big) =: \tilde{X}(n).$$

Define $\bar{X}_\theta(N) = \max_{0\le n\le N} \tilde{X}_\theta(n)$ so that (5.3.5) gives $\mathbb{P}[\bar{X}_\theta(N) \ge \beta] \le e^{-\alpha\beta}$ and

$$\lim_{\theta\to\infty} \bar{X}_\theta(N) = \lim_{\theta\to\infty} \max_{0\le n\le N} \tilde{X}_\theta(n) = \max_{0\le n\le N} \tilde{X}(n) =: \bar{X}(N).$$

Therefore, for any $\varepsilon > 0$,

$$\mathbb{P}[\bar{X}(N) \ge \beta] \le \mathbb{P}[\bar{X}(N) - \bar{X}_\theta(N) \ge \varepsilon] + \mathbb{P}[\bar{X}_\theta(N) \ge \beta - \varepsilon]$$
$$\le \mathbb{P}[|\bar{X}(N) - \bar{X}_\theta(N)| > \varepsilon/2] + e^{-\alpha(\beta-\varepsilon)}.$$

However, since $\lim_{\theta\to\infty} \bar{X}_\theta(N) = \bar{X}(N)$, we have that for all $\varepsilon > 0$

$$\lim_{\theta\to\infty} \mathbb{P}[|\bar{X}(N) - \bar{X}_\theta(N)| > \varepsilon/2] = 0,$$

and hence $\mathbb{P}[\bar{X}(N) \ge \beta] = \lim_{\theta\to\infty} \mathbb{P}[\bar{X}(N) \ge \beta] \le e^{-\alpha(\beta-\varepsilon)}$. Letting $\varepsilon \to 0$ yields

$$\mathbb{P}\big[\max_{0\le n\le N} \tilde{X}(n) \ge \beta\big] \le e^{-\alpha\beta}.$$

Finally, since $\tilde{X}(0) = 0$ and $\beta > 0$ we have

$$e^{-\alpha\beta} \ge \mathbb{P}\big[\max_{0\le n\le N} \tilde{X}(n) \ge \beta\big] = \mathbb{P}[\max_{1\le n\le N} \tilde{X}(n) \ge \beta]$$
$$= \mathbb{P}\big[\max_{1\le n\le N} \big(\sum_{j=0}^{n-1} S_1(j) - \frac{\alpha}{2}\sum_{j=0}^{n-1} S_2(j)\big) \ge \beta\big],$$

which completes the proof.

5.4 Alternative proof of Theorem 5.2.4

For simplicity we consider the scalar case. Let $\xi = \{\xi(n) : n \ge 1\}$ be a sequence of independent and identically distributed standard normal random variables and let $U = \{U(n) : n \ge 0\}$ be a stochastic process which is adapted to the natural filtration of the Gaussian sequence on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{P}, (\mathcal{F}(n))_{n\geq 0})$. Here the filtration is $\mathcal{F}(n) := \sigma\{\xi(j) : 1 \leq j \leq n\}$ and we make the simplifying assumption that $U(0)$ is deterministic. Since $U(n)$ is $\mathcal{F}(n)$–measurable for each $n \geq 1$, by the classical Doob-Dynkin lemma (see [54] page 4), for each $n \geq 1$ there exists a deterministic and measurable function $f_n : \mathbb{R}^n \to \mathbb{R}$ such that $U(n) = f_n(\xi(1), \xi(2), \ldots, \xi(n))$.

Let $B(t)$, $t \geq 0$, be a standard Brownian motion (with $B(0) = 0$) on a probability space $(\tilde{\Omega}, \tilde{\mathcal{F}}, \tilde{\mathbb{P}})$ with the natural filtration $(\tilde{\mathcal{F}}(t))_{t\geq 0}$. Hence $\{B(n) - B(n-1) : n \geq 1\}$ and $\{\xi(n) : n \geq 1\}$ have the same probability distributions (although they are defined on different probability spaces). That is, for any $N \geq 1$ and any real numbers $c_1, \cdots, c_N$,

$$\tilde{\mathbb{P}}\big[B(n) - B(n-1) \leq c_n,\ 1 \leq n \leq N\big] = \mathbb{P}\big[\xi(n) \leq c_n,\ 1 \leq n \leq N\big].$$

Define $\tilde{U}(n) = f_n(B(1) - B(0), B(2) - B(1), \ldots, B(n) - B(n-1))$ for $n \geq 1$ and $\tilde{U}(0) = U(0)$. Then $\{\xi(n), U(n) : n \geq 1\}$ and $\{B(n) - B(n-1), \tilde{U}(n) : n \geq 1\}$ have the same probability distributions. Now extend $\tilde{U}$ to continuous–time according to

$$\tilde{U}(t) = \tilde{U}(\lfloor t \rfloor), \quad t \geq 0$$

where $\lfloor t \rfloor \in \mathbb{N}_0$ denotes the integer part (or floor) of $t \geq 0$. This process is $\tilde{\mathcal{F}}(t)$–adapted. Define the stopping time $\tau(\theta)$ by

$$\tau(\theta) = \inf\left\{ t \geq 0 : \left| \int_0^t \tilde{U}(s)\, dB(s) \right| + \int_0^t \tilde{U}^2(s)\, ds \geq \theta \right\}$$

Moreover, let $\alpha > 0$, $N \in \mathbb{N}$ and consider the $\tilde{\mathcal{F}}(t)$–adapted process $X_\theta = \{X_\theta(t) : 0 \leq t \leq N\}$ given by

$$X_\theta(t) = \exp\left(\alpha \int_0^{t\wedge\tau(\theta)} \tilde{U}(s)\, dB(s) - \frac{1}{2}\alpha^2 \int_0^{t\wedge\tau(\theta)} \tilde{U}^2(s)\, ds \right), \quad 0 \leq t \leq N.$$

By Theorem 1.7.4 in [54], we find that $X_\theta(t)$ is a $\tilde{\mathcal{F}}(t)$ martingale. Therefore Doob's martingale inequality applies and we have

$$\tilde{\mathbb{P}}[\max_{0\leq t\leq N} X_\theta(n) \geq e^{\alpha\beta}] \leq e^{-\alpha\beta}.$$

Proceeding as in the proof of Theorem 1.7.4 in [54], by taking $\theta \to \infty$ and using the fact that $\tau(\theta) \uparrow \infty$ as $\theta \to \infty$ we obtain

$$\tilde{\mathbb{P}}\left[\max_{0\le t\le N}\left\{\int_0^t \tilde{U}(s)\,dB(s) - \frac{1}{2}\alpha\int_0^t \tilde{U}^2(s)\,ds\right\} \ge \beta\right] \le e^{-\alpha\beta}. \tag{5.4.1}$$

Define

$$F(t) := \int_0^t \tilde{U}(s)\,dB(s) - \frac{1}{2}\alpha\int_0^t \tilde{U}^2(s)\,ds, \quad t \in [0,N].$$

Since $\tilde{U}(t) = \tilde{U}(\lfloor t\rfloor)$ we have for all $t \ge 0$ and for each $n \in \{0,\dots,N\}$,

$$\begin{aligned} F(n) &= \int_0^n \tilde{U}(s)\,dB(s) - \frac{1}{2}\alpha\int_0^n \tilde{U}^2(s)\,ds \\ &= \sum_{j=0}^{n-1}\int_j^{j+1} \tilde{U}(s)\,dB(s) - \frac{1}{2}\alpha\sum_{j=0}^{n-1}\int_j^{j+1} \tilde{U}^2(s)\,ds \\ &= \sum_{j=0}^{n-1}\tilde{U}(j)\big(B(j+1)-B(j)\big) - \frac{1}{2}\alpha\sum_{j=0}^{n-1}\tilde{U}^2(j). \end{aligned}$$

Now, because $\max_{t\in[0,N]} F(t) \ge \max_{n\in\{1,\dots,N\}} F(n)$ it follows for any $\beta > 0$ that

$$\left\{\max_{n\in\{1,\dots,N\}} F(n) \ge \beta\right\} \subseteq \left\{\max_{t\in[0,N]} F(t) \ge \beta\right\}$$

and so, using (5.4.1) we get

$$\tilde{\mathbb{P}}\left[\max_{n\in\{1,\dots,N\}}\left\{\sum_{j=0}^{n-1}\tilde{U}(j)(B(j+1)-B(j)) - \frac{1}{2}\alpha\sum_{j=0}^{n-1}\tilde{U}^2(j)\right\} \ge \beta\right] \le e^{-\alpha\beta}.$$

But, recalling that $\{\xi(n), U(n) : n \ge 1\}$ and $\{B(n)-B(n-1), \tilde{U}(n) : n \ge 1\}$ have the same probability distributions, we have

$$\begin{aligned} &\mathbb{P}\left[\max_{n\in\{1,\dots,N\}}\left\{\sum_{j=0}^{n-1}U(j)\xi(j+1) - \frac{1}{2}\alpha\sum_{j=0}^{n-1}U^2(j)\right\} \ge \beta\right] \\ &\qquad = \tilde{\mathbb{P}}\left[\max_{n\in\{1,\dots,N\}}\left\{\sum_{j=0}^{n-1}\tilde{U}(j)(B(j+1)-B(j)) - \frac{1}{2}\alpha\sum_{j=0}^{n-1}\tilde{U}^2(j)\right\} \ge \beta\right]. \end{aligned}$$

Hence

$$\mathbb{P}\left[\max_{n\in\{1,\dots,N\}}\left\{\sum_{j=0}^{n-1}U(j)\xi(j+1) - \frac{1}{2}\alpha\sum_{j=0}^{n-1}U^2(j)\right\} \ge \beta\right] \le e^{-\alpha\beta}$$

as required.

Chapter 6

On the Pathwise Large Fluctuations of Discretised SDEs

6.1 Introduction

Use of more general exponential martingale inequalities (EMIs) tends, as shown in Chapter 4, to make it more difficult to obtain asymptotic estimates for the discretisations of SDEs which correspond to those of the underlying continuous–time equation. Moreover, these estimates can be inferior to their continuous counterparts. For these reasons, in Chapter 5 we developed a discrete EMI for martingales driven by Gaussian sequences as this is the type of martingale which occurs as a result of a typical Euler–Maruyama discretisation method applied to an SDE driven by standard Brownian motion. This EMI is more suitable for estimating the fluctuations of martingales which may be viewed as discretisations of Itô integrals.

In this chapter we once again consider the asymptotic behaviour of discretisations of SDEs but, in contrast to Chapter 4, we now utilise the discrete EMI of Chapter 5. More specifically, we study the asymptotic behaviour of the numerical solution of the non–autonomous SDE given by

$$dX(t) = f(X(t),t)\,dt + g(X(t),t)\,dB(t), \tag{6.1.1}$$

with drift coefficient $f : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ and diffusion coefficient $g : \mathbb{R} \times [0,\infty) \to \mathbb{R}$. Note that one could also consider the discretisation of an SDE with Markovian switching but since our attention in this chapter is on numerical analysis, we consider only non–switching

SDEs. In particular we attempt to find deterministic upper and lower estimates on the rate of growth of the *running maxima* $t \mapsto \sup_{0\le s\le t}|X(s)|$ by finding constants C_u and C_l and a function $\phi : (0,\infty) \to (0,\infty)$ such that

$$0 < C_l \le \limsup_{t\to\infty} \frac{|X(t)|}{\phi(t)} \le C_u, \quad \text{a.s.} \tag{6.1.2}$$

As before, we refer to such a function ϕ as the *essential rate of growth* of the largest deviations.

Asymptotic properties of the continuous–time model described by (6.1.1) are examined in [54] while some new continuous–time results are developed in Section 6.2 of this chapter. The results in [54] are achieved mainly through the combination of the *exponential martingale inequality* and *Gronwall's inequality.* The main aim of this chapter is to extend these methods to discrete–time and to obtain results which are consistent with the continuous–time counterparts.

Before stating comparable results for discrete problems, we give a rough indication of the main technical challenges that our proofs entail. We establish counterparts to the continuous–time results by mimicking each of the techniques used in their proofs.

Firstly, we obviate any difficulties relating to discrete Itô formulae by confining attention to equations where, in continuous time, the square of the process is considered. Secondly, good discrete analogues of Gronwall's inequality already exist enabling us, as in continuous time, to deal easily with the final estimation of the almost sure growth bound. The other key element of the proofs is the use of an exponential martingale inequality. In the discrete proofs in this chapter, we now use the more refined version of the discrete EMI developed in Chapter 5. Consideration of the inequality shows that it retains all of the features of the continuous EMI, and in particular, depends only on the predictable quadratic variation.

On implementing this programme, we generally obtain results which are natural discrete analogues of (6.1.2) and are of the form

$$0 < C_l(h) \leq \limsup_{n\to\infty} \frac{|X_h(n)|}{\phi(nh)} \leq C_u(h), \quad \text{a.s.}, \tag{6.1.3}$$

where h represents the fixed step–size used to produce the discretised process $X_h(n)$. Here, the limiting constants C_u and C_l may be h–dependent but will be arbitrarily close to their continuous–time counterparts provided h is small enough.

Over the last ten years a literature on the dynamic asymptotic consistency of SDEs has developed, among others [8, 10, 11, 34, 37, 38, 71]. The results (6.1.2) and (6.1.3) can be thought of in that context, as they show that the asymptotic behaviour of the SDE and its discretisation are consistent with each other (particularly as the step size $h \to 0$). However, most of this literature deals with SDEs and their discretisations which have asymptotically stable solutions (in either a p–th mean or a pathwise sense), as in [40] and [58]. Moreover, in [40] it is shown that certain properties can be lost under an Euler–Maruyama discretisation. To the best of the author's knowledge, less is known about the dynamic asymptotic consistency of discretisations of SDEs which have fluctuating solutions. This chapter attempts, at least in part, to fill this gap.

We consider two alternative methods of discretising (6.1.1). The first is the standard Euler–Maruyama method, while the second is an implicit variant of Euler–Maruyama introduced in [36] as the split–step backward Euler method. For the standard explicit Euler-Maruyama method, we obtain asymptotic behaviour of the form (6.1.3). However in general this holds provided that the step size $h > 0$ is chosen sufficiently small and that some assumptions are imposed on the drift coefficient f which are not required in order to prove (6.1.2). On the other hand the method is easy to implement and the discretisation has, in common with the underlying SDE, a unique solution.

In contrast, the implicit split–step variant of the Euler–Maruyama method does not necessarily require a restriction on the step size h, nor does it require such additional assumptions on the drift. However, the method is more difficult to implement than the explicit scheme and without an extra assumption on f (such as a one–sided Lipschitz condition), a unique solution of the split–step scheme cannot be guaranteed. Nevertheless we are able to show, as in deterministic work by Stuart and Humphries [77], that all solutions have the appropriate long–run behaviour.

Throughout the chapter our results are divided into two categories. In one case we impose conditions on f so that the drift is mean–reverting and close to being linear. This makes our solutions asymptotically similar to an Ornstein–Uhlenbeck process. Here our results, which we call *O–U type results*, are consistent with the asymptotic behaviour of the well-known Ornstein–Uhlenbeck process in the sense that the essential rate of growth is of the order $\phi(t) = \sqrt{\log t}$. Moreover we are able to obtain lower bounds on the large deviations of the running maxima which were not found in [54].

Our other class of results we call *Iterated Logarithm type results* since our solutions obey the Law of the Iterated Logarithm. That is, the essential growth rate is of the order $\phi(t) = \sqrt{2t \log\log t}$. Here we impose conditions on f which ensure that (6.1.1) is close to Brownian motion in the sense that the drift coefficient is small, especially for large values of X. It is interesting to remark that discrete results in this case can be proven without any restrictions on h.

To begin with we state results for the simplified constant diffusion case, i.e., where $g(x,t) = \sigma$ for all $(x,t) \in \mathbb{R} \times [0,\infty)$. We then extend the results to the more general case (6.1.1) where the diffusion coefficient is globally bounded, and has identifiable asymptotic behaviour as $|x| \to \infty$. A new feature of the proofs in continuous time, as well as the corresponding discrete results, is that the asymptotic estimates on the large fluctuations depend only on the behaviour of f and g as $|x| \to \infty$. This leads to sharper results than those obtained in Mao [52], in which related *global bounds* on the coefficients lead to asymptotic estimates which depend on these global bounds.

The chapter is organised as follows. In Section 6.2 we establish results for continuous–time equations obeying either Ornstein–Uhlenbeck type growth bounds, or iterated logarithm type growth bounds. In Section 6.3 we consider an Euler–Maruyama discretisation of such equations while in Section 6.4 we consider a split–step discretisation method. Proofs are postponed to the final three sections: proofs of the continuous–time results can be found in Section 6.5, proofs of exponential martingale estimates are in Section 6.6 and proofs in the discrete cases can be found in Section 6.7.

6.2 Continuous–Time Processes

Let $X(0) = x_0$ and consider the stochastic differential equation given by

$$dX(t) = f(X(t), t)\,dt + \sigma\,dB(t), \quad t \geq 0, \tag{6.2.1}$$

where $\sigma \in \mathbb{R}/\{0\}$ and $f : \mathbb{R} \times [0, \infty) \to \mathbb{R}$. We assume throughout the chapter, without further repetition, that f is locally Lipschitz continuous in the sense that there exists a constant $M_n > 0$ such that

$$|f(x,t) - f(y,t)| \leq M_n|x - y|, \quad \text{for all } |x|, |y| \leq n \text{ and } t \geq 0. \tag{6.2.2}$$

For economy of exposition this assumption is not explicitly repeated in the statement of theorems in this chapter. We sometimes ask that f obeys the following linear growth condition.

Assumption 6.2.1. *There exists a positive constant* Γ *such that*

$$\limsup_{|x|\to\infty} \frac{|f(x,t)|}{|x|} \leq \Gamma, \quad \textit{uniformly for all } t \geq 0. \tag{6.2.3}$$

Moreover for every $A > 0$, there exists $c_1(A) < \infty$ such that

$$\sup_{|x|\le A} |f(x,t)| \le c_1(A), \quad \textit{uniformly for all } t \ge 0.$$

Remark 6.2.1. Note that (6.2.3) implies that for all $\upsilon \in (0,1)$ there exists $X(\upsilon) > 0$, independent of t, such that $|f(x,t)| \le \Gamma(1+\upsilon)|x|$ for $|x| \ge X(\upsilon)$. Moreover, for $|x| < X(\upsilon)$ we have $|f(x,t)| \le c_1(X(\upsilon))$. Define $\Gamma_0 := c_1(X(\upsilon))$ and $\Gamma_1 := \Gamma(1+\upsilon)$. Then combining both estimates, for all $x \in \mathbb{R}$, we have

$$|f(x,t)| \le \Gamma_0 + \Gamma_1|x| \quad \text{uniformly for all } t \ge 0. \tag{6.2.4}$$

By Assumption 6.2.1 and the fact that f is locally Lipschitz continuous, there is a unique continuous and adapted process which satisfies (6.2.1) (see e.g. [54]).

6.2.1 O–U type results

In order to introduce the main ideas in this section we first set out the conditions on the drift which produce Ornstein–Uhlenbeck type behaviour in the solution.

Assumption 6.2.2. *There exists a positive constant γ such that*

$$\limsup_{|x|\to\infty} \frac{xf(x,t)}{x^2} \le -\gamma, \quad \textit{uniformly for all } t \ge 0. \tag{6.2.5}$$

Moreover for every $A > 0$, there exists $c_2(A) < \infty$ such that

$$\sup_{|x|\le A} xf(x,t) \le c_2(A), \quad \textit{uniformly for all } t \ge 0.$$

Remark 6.2.2. Note that (6.2.5) implies that for all $\eta \in (0,1)$ there exists $X(\eta) > 0$, independent of t, such that $xf(x,t) \le -\gamma(1-\eta)x^2$ for $|x| \ge X(\eta)$. Moreover, for $|x| < X(\eta)$,

$$\begin{aligned} xf(x,t) &\le c_2(X(\eta)) = c_2(X(\eta)) + \gamma(1-\eta)x^2 - \gamma(1-\eta)x^2 \\ &\le c_2(X(\eta)) + \gamma(1-\eta)X^2(\eta) - \gamma(1-\eta)x^2. \end{aligned}$$

Define $\rho_1 := c_2(X(\eta))+\gamma(1-\eta)X^2(\eta)$ and $\gamma_1 := \gamma(1-\eta)$. Then combining both estimates, for all $x \in \mathbb{R}$, we have

$$xf(x,t) \leq \rho_1 - \gamma_1 x^2 \quad \text{uniformly for all } t \geq 0. \tag{6.2.6}$$

In our proofs throughout this chapter we will be using (6.2.4) and (6.2.6) for notational convenience and for the fact that assumptions of that form are consistent with those used by Mao in [54]. However, at the final stage of our proofs we will re–introduce the fact that $\gamma_1 = \gamma(1-\eta)$ and $\Gamma_1 = \Gamma(1+\nu)$ for η, ν arbitrarily small and we can then allow $\eta, \nu \to 0$ to obtain results which depend on the asymptotic estimates (6.2.3) and (6.2.5). The reason for this is subtle yet worthwhile. Although we could simply make global assumptions on f of the form (6.2.4) and (6.2.6), these might not accurately reflect the asymptotic behaviour of f, since the conditions must hold for *all* values of x. By instead using asymptotic conditions of the form (6.2.3) and (6.2.5), we are isolating the asymptotic behaviour of f as $|x| \to \infty$ since it should only be these values which contribute to the asymptotic behaviour of the process.

To that end, we now demonstrate that an asymptotic condition of the form (6.2.5) can give a sharper result than if one were to use a global condition of the form (6.2.6). This result mirrors a similar theorem in Mao, [54].

Theorem 6.2.3. *Let X be the unique continuous adapted solution to* (6.2.1) *and let f obey Assumption 6.2.2. Then,*

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{\log t}} \leq |\sigma|\sqrt{\frac{e}{\gamma}}, \quad a.s. \tag{6.2.7}$$

It is important to note that if one were to use the global condition (6.2.6) then we would get the same result, but with γ_1 instead of γ. However, since $\gamma \geq \gamma_1$ it follows that (6.2.7) is slightly sharper than that which is obtained under (6.2.6).

Under Assumption 6.2.1 we can also get a complementary lower bound on the large fluctuations of $|X|$, a result which has not appeared in the literature to date.

Theorem 6.2.4. *Let X be the unique continuous adapted solution to* (6.2.1) *and let f obey Assumption 6.2.1. Then,*

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{\log t}} \geq \frac{|\sigma|}{2\sqrt{2\Gamma}}, \quad a.s. \tag{6.2.8}$$

Theorems 6.2.3 and 6.2.4 are reasonably sharp in the sense that in the simple linear case where $f(x,t) = -\gamma x$, it is well–known that $\limsup_{t\to\infty} |X(t)|/\sqrt{\log t} = |\sigma|/\sqrt{\gamma}$ a.s., (see [54] for example) and the hypotheses (6.2.4) and (6.2.6) effectively represent a perturbation from the linear case. In fact, the essential growth rate is of the right order ($\phi(t) = \sqrt{\log t}$) and the constants on the right–hand sides of (6.2.7) and (6.2.8) are of the right "dimension" in that they are determined by the diffusion coefficient divided by the square root of the linearity coefficient, roughly speaking.

6.2.2 Iterated Logarithm type results

One can also derive iterated logarithm type growth bounds on (6.2.1) under the following condition.

Assumption 6.2.5. *There exists a positive constant ρ such that for all $x \in \mathbb{R}$,*

$$xf(x,t) \leq \rho \quad \textit{uniformly for all } t \geq 0. \tag{6.2.9}$$

Under Assumption 6.2.5, (6.2.1) has a unique continuous adapted solution X. Again, in [54] it was shown that under condition (6.2.9) on f we have

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{2t\log\log t}} \leq |\sigma|\sqrt{e}, \quad \text{a.s.} \tag{6.2.10}$$

A discrete analogue of this result was first attempted in Chapter 4. However, the use of a general discrete–time exponential martingale inequality gave us a result which was less sharp than (6.2.10). In the next section we use the specialised discrete exponential martingale inequality from Chapter 5 to obtain a sharper discrete analogue of (6.2.10).

6.2.3 General Diffusion Coefficient

We can also generalise the noise term to study an equation of the form

$$dX(t) = f(X(t),t)\,dt + g(X(t),t)\,dB(t), \quad t \geq 0, \tag{6.2.11}$$

where $g : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ obeys the following hypothesis.

Assumption 6.2.6. *There exists a positive constant K such that*

$$\limsup_{|x|\to\infty} |g(x,t)| \leq K, \quad \textit{uniformly for all } t \geq 0. \tag{6.2.12}$$

Moreover, we assume that there exists $\kappa > 0$ such that for all $x, y \in \mathbb{R}$

$$\big(g(x,t) - g(y,t)\big)^2 \leq \kappa(x-y)^2 \quad \textit{uniformly for all } t \geq 0. \tag{6.2.13}$$

Under Assumption 6.2.6 on g and Assumption 6.2.5 on f, there is a unique continuous and adapted process which satisfies (6.2.11).

Remark 6.2.3. Note that (6.2.12) implies that for all $\varepsilon \in (0,1)$ there exists $X(\varepsilon) > 0$, independent of t, such that $|g(x,t)| \leq K(1+\varepsilon)$ for $|x| \geq X(\varepsilon)$. Moreover, by the continuity of g, for $|x| < X(\varepsilon)$ there exists c_3 such that $|g(x,t)| \leq c_3(X(\varepsilon)) < \infty$. Define $K_1 := c_3(X(\varepsilon)) + K(1+\varepsilon)$. Then combining both estimates, for all $x \in \mathbb{R}$, we have

$$|g(x,t)| \leq K_1 \quad \text{uniformly for all } t \geq 0. \tag{6.2.14}$$

In some instances in our proofs it may be more convenient to use the global bound (6.2.14) when dealing with terms which are not asymptotically dominant. It turns out that such terms do not affect the final asymptotic result. However, when dealing with terms which will contribute to the final asymptotic result we will instead use the more accurate asymptotic bound on g given by (6.2.12).

In [54], Mao makes the point that under condition (6.2.14) it is clear that the upper bound results (6.2.7) and (6.2.10) still hold for the solution of equation (6.2.11) with the corresponding $|\sigma|$ replaced by K_1. In fact, it could easily be shown using (6.2.12) that the upper bound results (6.2.7) and (6.2.10) still hold for the solution of equation (6.2.11) with the corresponding $|\sigma|$ replaced by K, though the details are omitted. We now show that the lower bound result (6.2.8) also holds for the solution of (6.2.11) provided g is uniformly bounded *below.*

Theorem 6.2.7. *Let X be the unique continuous adapted solution to* (6.2.11)*. Let f obey Assumption 6.2.1 and let g be Lipschitz continuous and obey*

$$|g(x,t)| \geq K_2 \quad \textit{for all } (x,t) \in \mathbb{R} \times [0,\infty).$$

Then,

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{\log t}} \geq \frac{K_2}{2\sqrt{2\Gamma}}, \quad a.s. \tag{6.2.15}$$

6.3 Euler–Maruyama Discretisation Scheme

Having stated results for continuous–time processes in the previous section, we now consider discrete–time analogues of such results. This is achieved through discretisation of the process $X(t)$. We first consider an Euler–Maruyama discretisation of the SDE (6.2.1), which takes the form

$$X_h(n+1) = X_h(n) + hf(X_h(n), nh) + \sqrt{h}\sigma\xi(n+1), \; n \geq 0, \tag{6.3.1}$$

where $h > 0$ is the fixed step size and ξ is a sequence of independent standard normal random variables. To avoid repetition, we assume throughout the rest of the chapter that

$\big(\xi(n)\big)_{n\geq 1}$ is a sequence of i.i.d. $\mathcal{N}(0,1)$ random variables. This is a commonly used explicit discretisation method and under our assumptions on f there exists a unique solution to (6.3.1).

6.3.1 O–U type results

We are now in a position to state our main results in discrete–time. The following is an analogue of Theorem 6.2.3.

Theorem 6.3.1. *Let f obey Assumptions 6.2.1 and 6.2.2 and let $h < \min(\gamma/\Gamma^2, 1/(2\gamma))$. If X_h is the unique adapted solution to* (6.3.1), *then*

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{\log nh}} \leq |\sigma|\sqrt{\frac{e}{\gamma}\,C(h)}, \quad a.s., \tag{6.3.2}$$

where $C(h) > 1$ and $C(h) \to 1$ as $h \to 0$. Moreover, we can write $C(h)$ explicitly as

$$C(h) := \Big[\frac{2}{1 - 2h\gamma + 2h^2\Gamma^2} - 1\Big]\Big(\frac{\gamma}{\gamma - h\Gamma^2}\Big).$$

Remark 6.3.1. Notice that in contrast to Theorem 6.2.3, we must assume *both* Assumptions 6.2.1 and 6.2.2 in order to prove Theorem 6.3.1. Later in the chapter we will show that a different discretisation method allows us to drop Assumption 6.2.1.

It is important to note also that the constant on the right–hand side of (6.3.2) is arbitrarily close to the analogous constant in the continuous-time result (6.2.7), for h small enough. The continuous time version of Theorem 6.3.1 was first established by an exponential martingale and Gronwall inequality proof in [52]: the proof of Theorem 6.3.1 is modelled on the argument in that work except that in the discrete case an extra martingale term arises which must be estimated. This extra ingredient is a result of the inability to accurately replicate Itô's rule in discrete time. Therefore, to prove Theorem 6.3.1 we require new technical results which hinge on a further specialised version of the EMI. These new technical results can be found in Section 6.6.

We now state a discrete–time analogue of Theorem 6.2.4.

Theorem 6.3.2. *Let f satisfy Assumption 6.2.1 and let $h < 2/\Gamma$. If X_h is the unique adapted solution to* (6.3.1)*, then*

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{\log nh}} \geq \frac{|\sigma|}{2\sqrt{2\Gamma}} C(h), \quad a.s., \tag{6.3.3}$$

where $C(h) := \sqrt{2}/\sqrt{2 - h\Gamma/2} > 1$ and $C(h) \to 1$ as $h \to 0$.

Again, the constant on the right–hand side of (6.3.3) is arbitrarily close to the analogous constant in Theorem 6.2.4, for h small enough.

6.3.2 Iterated Logarithm type results

We now look at a discrete–time analogue of (6.2.10). Once again we use Assumption 6.2.5 on f. We do not require that f be linearly bounded as in Assumption 6.2.1, instead we request that f be bounded in the sense that there exists a constant $\bar{f} > 0$ such that for all $x \in \mathbb{R}$

$$|f(x,t)| \leq \bar{f} \quad \text{uniformly for all } t \geq 0. \tag{6.3.4}$$

We then get the following result.

Theorem 6.3.3. *Let f obey* (6.3.4) *and Assumption 6.2.5 and let X_h be the unique adapted solution to* (6.3.1)*. Then*

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{2nh \log\log nh}} \leq |\sigma|\sqrt{e}, \quad a.s. \tag{6.3.5}$$

In this case we are able to recover exactly Mao's result (6.2.10) in discrete–time, contingent on the additional assumption (6.3.4). In particular, we have eliminated the extra factor of 2 which was first established in similar results in Chapter 4. Later in this chapter, we show that an alternative discretisation enables us to prove (6.3.5) without assuming (6.3.4).

6.3.3 General Diffusion Coefficient

We can again generalise the noise term in (6.3.1) to study an equation of the form

$$X_h(n+1) = X_h(n) + hf(X_h(n), nh) + \sqrt{h}g(X_h(n), nh)\xi(n+1), \ n \geq 0, \tag{6.3.6}$$

where $g : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ obeys Assumption 6.2.6. Since f is continuous there is a unique continuous and adapted process which satisfies (6.3.6).

Our next results (of O–U type and Iterated Logarithm type respectively) show that Theorems 6.3.1 and 6.3.3 still hold for the solution of (6.3.6) under Assumption 6.2.6 on g. To date, we have not been able to recover an analogue of the lower bound estimate from Theorem 6.3.2 in the case when we have a non–constant diffusion coefficient.

Theorem 6.3.4. *Let f satisfy Assumptions 6.2.1 and 6.2.2 and let g satisfy Assumption 6.2.6. Let $h < \min(\gamma/\Gamma^2, 1/(2\gamma))$ and let X_h be the unique adapted solution to* (6.3.6)*. Then*

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{\log nh}} \leq K\sqrt{\frac{e}{\gamma}\, C(h)}, \quad a.s.$$

where $C(h) > 1$ and $C(h) \to 1$ as $h \to 0$. Moreover, we can write $C(h)$ explicitly as

$$C(h) := \Big[\frac{2}{1 - 2h\gamma + 2h^2\Gamma^2} - 1\Big]\Big(\frac{\gamma}{\gamma - h\Gamma^2}\Big).$$

Theorem 6.3.5. *Let f obey* (6.3.4) *and Assumption 6.2.5 and let g satisfy Assumption 6.2.6. If X_h is the unique adapted solution to* (6.3.6)*, then*

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{2nh \log\log nh}} \leq K\sqrt{e}, \quad a.s.$$

6.4 Split–Step discretisation scheme

In our efforts to replicate the effect of Itô's rule in discrete–time, we very often encounter additional terms which are estimated with the help of conditions such as (6.2.4) in the

O–U case and (6.3.4) in the Iterated Logarithm case. In particular, when using an Euler–Maruyama scheme such as (6.3.1), our proofs seem to require these estimates on f. We now show that the use of a split–step implicit variant of Euler–Maruyama negates the need for conditions (6.2.4) and (6.3.4). The discretisation is as follows. Set $X_h(0) = X(0)$ and define

$$X_h^*(n) = X_h(n) + hf(X_h^*(n), nh),$$
$$X_h(n+1) = X_h^*(n) + \sigma\sqrt{h}\xi(n+1), \tag{6.4.1}$$

where, as before, $h > 0$ is the fixed step size and ξ is a sequence of independent standard normal random variables. The first step of the method is an implicit equation that must be solved in order to obtain the intermediate approximation $X_h^*(n)$. Having obtained $X_h^*(n)$, adding the appropriate stochastic increment $\sigma\sqrt{h}\xi(n+1)$ produces the next approximation $X_h(n+1)$. We say that (6.4.1) has a solution if there is a pair of processes (X_h, X_h^*) which obey (6.4.1). Such a solution will automatically be global, (i.e., defined for all $n \geq 0$) and will be adapted to the natural filtration generated by the ξ's.

We now consider the existence of solutions to the discretisation method given by (6.4.1). In [36] it is shown that under a global one–sided Lipschitz condition on the drift function f there exists a *unique* solution to (6.4.1) provided the step size h is chosen to be sufficiently small.

Assumption 6.4.1. *Assume that* $f \in C^1(\mathbb{R})$ *and that for all* $x, y \in \mathbb{R}$ *there exists a constant* $c \in \mathbb{R}$ *such that*

$$(x-y)(f(x,t) - f(y,t)) \leq c(x-y)^2, \quad \textit{uniformly for all } t \geq 0. \tag{6.4.2}$$

We refer to this condition as a global one–sided Lipschitz condition because the constant c is independent of x and y in (6.4.2). Although this is weaker than requesting that f

satisfy a global Lipschitz condition, it places a restriction on f on all $\mathbb{R}$, and still excludes some functions f which grow faster than polynomially as $|x| \to \infty$.

Moreover, we do not necessarily require this global one–sided Lipschitz condition if we are willing to sacrifice uniqueness in the solution of (6.4.1). This is in the spirit of generalised dynamical systems considered by Stuart and Humphries [77]. In Section 6.8, we will show that either condition (6.2.6) or condition (6.2.9) is enough to guarantee that (6.4.1) has a (not necessarily unique) solution. However, if uniqueness of the solution of (6.4.1) is required then we are still free to impose condition (6.4.2) on f, at the expense of an a–priori size restriction on h.

We are now in a position to state an analogue of Theorem 6.3.1 under the split–step discretisation scheme. Here, we no longer require Assumption 6.2.1 on f.

Theorem 6.4.2. *Let f satisfy Assumption 6.2.2 and let X_h be any adapted solution to (6.4.1). Then*

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{\log nh}} \le |\sigma|\sqrt{\frac{e}{\gamma}\, C(h)}, \quad a.s., \tag{6.4.3}$$

where $C(h) := 1 + 6h\gamma + 8h^2\gamma^2 > 1$ and $C(h) \to 1$ as $h \to 0$.

Similarly we have an analogue of Theorem 6.3.3 under the split–step discretisation scheme and here we no longer require condition (6.3.4) on f.

Theorem 6.4.3. *Let f satisfy Assumption 6.2.5 and let X_h be any adapted solution to (6.4.1). Then*

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{2nh\log\log nh}} \le |\sigma|\sqrt{e}, \quad a.s. \tag{6.4.4}$$

As mentioned earlier, we may also restate Theorems 6.4.2 and 6.4.3 where X_h is the *unique* adapted solution to (6.4.1) if we also stipulate that f obeys Assumption 6.4.1 and that $h \le (2c+1)^{-1}$ where c arises from (6.4.2).

6.4.1 General Diffusion Coefficient

We can again generalise the noise term in (6.4.1) to study a system of the form

$$X_h^*(n) = X_h(n) + hf(X_h^*(n), nh),$$
$$X_h(n+1) = X_h^*(n) + g(X_h^*(n), nh)\sqrt{h}\xi(n+1), \tag{6.4.5}$$

where $g : \mathbb{R} \times [0, \infty) \to \mathbb{R}$ obeys Assumption 6.2.6. The addition of the non-constant diffusion coefficient will not affect the existence of a solution of (6.4.5) since g is bounded.

We now show that Theorems 6.4.2 and 6.4.3 still hold for the solution of (6.4.5) under Assumption 6.2.6 on g.

Theorem 6.4.4. *Let f satisfy Assumption 6.2.2 and let g satisfy Assumption 6.2.6. If X_h is any adapted solution to* (6.4.5)*, then*

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{\log nh}} \leq K\sqrt{\frac{e}{\gamma}\, C(h)}, \quad a.s.,$$

where $C(h) := 1 + 6h\gamma + 8h^2\gamma^2 > 1$ and $C(h) \to 1$ as $h \to 0$.

Theorem 6.4.5. *Let f satisfy Assumption 6.2.5 and let g satisfy Assumption 6.2.6. If X_h is any adapted solution to* (6.4.5)*, then*

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{2nh\log\log nh}} \leq K\sqrt{e}, \quad a.s.$$

Again, we may restate Theorems 6.4.4 and 6.4.5 where X_h is the *unique* adapted solution to (6.4.5) if we also stipulate that f obeys Assumption 6.4.1 and that $h \leq ((2c+1)\vee 4\kappa)^{-1}$ where c and κ arise from (6.4.2) and (6.2.13) respectively.

6.5 Proofs of Results from Section 6.2

Proof of Theorem 6.2.3. In [54] it was shown that if X is the unique continuous adapted solution to (6.2.1) then under condition (6.2.6) on f we have

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{\log t}} \le |\sigma| \sqrt{\frac{e}{\gamma_1}}, \quad \text{a.s.}$$

Now recalling that $\gamma_1 = \gamma(1-\eta)$ where $\eta \in (0,1)$ we get

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{\log t}} \le |\sigma| \sqrt{\frac{e}{\gamma(1-\eta)}}, \quad \text{a.s.}$$

This estimate holds for all outcomes in an event of probability one, say Ω_η. By considering $\Omega^* := \cap_{p\in\mathbb{N}} \Omega_{\frac{1}{p}}$, i.e. by letting $\eta \to 0$, we get the desired result. □

Proof of Theorem 6.2.4. Let $\mu > 0$. From (6.2.1) we have

$$dX(t) = \big[-\mu X(t) + F(X(t)) \big]\, dt + \sigma\, dB(t),$$

where $F(x) := \mu x + f(x,t)$. We can then show that

$$e^{\mu t} X(t) = X(0) + \int_0^t e^{\mu s} F(X(s))\, ds + M(t), \tag{6.5.1}$$

where $M(t) := \int_0^t \sigma e^{\mu s}\, dB(s)$ is a martingale with quadratic variation given by $\langle M\rangle(t) = \int_0^t \sigma^2 e^{2\mu s}\, ds = \sigma^2(e^{2\mu t}-1)/2\mu$. Moreover, notice that

$$\lim_{t\to\infty} \frac{\langle M\rangle(t)}{e^{2\mu t}} = \lim_{t\to\infty} \frac{\sigma^2}{2\mu}(1-e^{-2\mu t}) = \frac{\sigma^2}{2\mu} \quad \text{and} \quad \lim_{t\to\infty} \frac{\log\log \langle M\rangle(t)}{\log t} = 1.$$

Therefore, using the law of the iterated logarithm for martingales (Exercise 5.1.15, [70])

$$\begin{aligned}\limsup_{t\to\infty} \frac{|M(t)|}{e^{\mu t}\sqrt{\log t}} &= \limsup_{t\to\infty} \frac{\sqrt{2}|M(t)|}{\sqrt{2\langle M\rangle(t) \log\log \langle M\rangle(t)}} \cdot \sqrt{\frac{\langle M\rangle(t)}{e^{2\mu t}} \frac{\log\log \langle M\rangle(t)}{\log t}} \\ &= \sqrt{\frac{\sigma^2}{\mu}}, \quad \text{a.s.}\end{aligned} \tag{6.5.2}$$

Let $c > 0$ and define the event

$$A_c := \left\{ \omega : \limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{\log t}} < c \right\} \tag{6.5.3}$$

and assume that $\mathbb{P}[A_c] > 0$. We will demonstrate that this is impossible for the appropriate choice of c. By (6.5.1) and the triangle inequality it follows that

$$\limsup_{t\to\infty}\frac{|M(t)|}{e^{\mu t}\sqrt{\log t}} \leq \limsup_{t\to\infty}\frac{|X(t)|}{\sqrt{\log t}} + \limsup_{t\to\infty}\frac{|X(0)|}{e^{\mu t}\sqrt{\log t}} + \limsup_{t\to\infty}\frac{\int_0^t e^{\mu s}|F(X(s))|\,ds}{e^{\mu t}\sqrt{\log t}}$$
$$< c + \limsup_{t\to\infty}\frac{\int_0^t e^{\mu s}|F(X(s))|\,ds}{e^{\mu t}\sqrt{\log t}}, \quad \text{a.s. on } A_c. \tag{6.5.4}$$

Using (6.2.4) we have $|F(x)| \leq \mu|x| + |f(x,t)| \leq (\mu+\Gamma_1)|x| + \Gamma_0$ for all x, and so

$$\int_0^t e^{\mu s}|F(X(s))|\,ds \leq (\mu+\Gamma_1)\int_0^t e^{\mu s}|X(s)|\,ds + \frac{\Gamma_0}{\mu}(e^{\mu t}-1). \tag{6.5.5}$$

By (6.5.3) it follows that for every $\varepsilon > 0$ there exists $T_1(\varepsilon,\omega) > 0$ such that for $t \geq T_1(\varepsilon,\omega)$ we have $|X(t,\omega)| \leq c(1+\varepsilon)\sqrt{\log t}$, for all $\omega \in A_c$. Therefore,

$$\int_0^t e^{\mu s}|X(s,\omega)|\,ds \leq \int_0^{T_1(\varepsilon,\omega)} e^{\mu s}|X(s,\omega)|\,ds + c(1+\varepsilon)\int_{T_1(\varepsilon,\omega)}^t e^{\mu s}\sqrt{\log s}\,ds$$
$$\leq X^*(\varepsilon,\omega)\int_0^{T_1(\varepsilon,\omega)} e^{\mu s}\,ds + c(1+\varepsilon)\int_{T_1(\varepsilon,\omega)}^t e^{\mu s}\sqrt{\log s}\,ds, \tag{6.5.6}$$

where $X^*(\varepsilon,\omega) := \max_{0\leq s\leq T_1(\varepsilon,\omega)}|X(s,\omega)| < +\infty$. By L'Hôpital's rule

$$\lim_{t\to\infty}\frac{\int_{T_1(\varepsilon,\omega)}^t e^{\mu s}\sqrt{\log s}\,ds}{e^{\mu t}\sqrt{\log t}} = \frac{1}{\mu}\lim_{t\to\infty}\frac{2t\log t}{\mu^{-1}+2t\log t} = \frac{1}{\mu},$$

and so using this along with equations (6.5.5) and (6.5.6) we get

$$\limsup_{t\to\infty}\frac{\int_0^t e^{\mu s}|F(X(s,\omega))|\,ds}{e^{\mu t}\sqrt{\log t}} \leq \frac{(\mu+\Gamma_1)c(1+\varepsilon)}{\mu}, \quad \text{for all } \omega \in A_c.$$

Moreover, since $\varepsilon > 0$ can be chosen arbitrarily we can allow $\varepsilon \to 0$ through the rationals to get

$$\limsup_{t\to\infty}\frac{\int_0^t e^{\mu s}|F(X(s,\omega))|\,ds}{e^{\mu t}\sqrt{\log t}} \leq \frac{(\mu+\Gamma_1)c}{\mu}, \quad \text{for all } \omega \in A_c.$$

Therefore by (6.5.2) and (6.5.4),

$$\sqrt{\frac{\sigma^2}{\mu}} = \limsup_{t\to\infty}\frac{|M(t)|}{e^{\mu t}\sqrt{\log t}} < c + \frac{(\mu+\Gamma_1)c}{\mu}, \quad \text{a.s. on } A_c.$$

However, we get a contradiction above if we make the choice $c = \sqrt{\sigma^2\mu}/(2\mu+\Gamma_1)$ which means that our assumption that $\mathbb{P}[A_c] > 0$ is incorrect. Thus, $\mathbb{P}[\bar{A}_c] = 1$ and

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{\log t}} \geq c = \frac{|\sigma|\sqrt{\mu}}{2\mu+\Gamma_1}, \quad \text{a.s.}$$

Then choosing $\mu = \Gamma_1/2$ we get

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{\log t}} \geq \frac{|\sigma|}{2\sqrt{2\Gamma_1}}, \quad \text{a.s.}$$

Now recalling that $\Gamma_1 = \Gamma(1+\upsilon)$ where $\upsilon \in (0,1)$ we get

$$\limsup_{t\to\infty} \frac{|X(t)|}{\sqrt{\log t}} \geq \frac{|\sigma|}{2\sqrt{2\Gamma(1+\upsilon)}}, \quad \text{a.s.}$$

This estimate holds for all outcomes in an event of probability one, say Ω_υ. By considering $\Omega^* := \cap_{p\in\mathbb{N}}\Omega_{\frac{1}{p}}$, i.e. by letting $\upsilon \to 0$, we get the desired result. □

Proof of Theorem 6.2.7. Let $\mu > 0$. From (6.2.11) we have

$$dX(t) = \big[-\mu X(t) + F(X(t))\big]\,dt + g(X(t),t)\,dB(t),$$

where $F(x) := \mu x + f(x,t)$. Then by Itô's rule we can show that

$$e^{\mu t}X(t) = X(0) + \int_0^t e^{\mu s}F(X(s))\,ds + M(t),$$

where $M(t) := \int_0^t g(X(s),s)e^{\mu s}\,dB(s)$ is a martingale with quadratic variation given by $\langle M\rangle(t) = \int_0^t g^2(X(s),s)e^{2\mu s}\,ds \geq K_2^2(e^{2\mu t}-1)/2\mu$. Moreover, notice that

$$\lim_{t\to\infty} \frac{\langle M\rangle(t)}{e^{2\mu t}} \geq \lim_{t\to\infty} \frac{K_2^2}{2\mu}(1-e^{-2\mu t}) = \frac{K_2^2}{2\mu} \quad \text{and} \quad \lim_{t\to\infty} \frac{\log\log\langle M\rangle(t)}{\log t} \geq 1.$$

Therefore, using the law of the iterated logarithm for martingales we get

$$\begin{aligned}\limsup_{t\to\infty} \frac{|M(t)|}{e^{\mu t}\sqrt{\log t}} &= \limsup_{t\to\infty} \frac{\sqrt{2}|M(t)|}{\sqrt{2\langle M\rangle(t)\log\log\langle M\rangle(t)}}\cdot\sqrt{\frac{\langle M\rangle(t)}{e^{2\mu t}}\frac{\log\log\langle M\rangle(t)}{\log t}} \\ &\geq \sqrt{\frac{K_2^2}{\mu}}, \quad \text{a.s.}\end{aligned}$$

The proof now follows exactly the same steps as in the proof of Theorem 6.2.4, where σ is replaced by K_2. □

6.6 Proofs of Exponential Martingale Estimates

The following is the (easily derived) moment generating function of a form of Chi–Squared distribution.

Lemma 6.6.1. *Suppose that ξ is a standard normal random variable and that $\zeta = \xi^2 - 1$. Then for $\lambda \in (0, 1/2)$*

$$\varphi(\lambda) := \mathbb{E}[e^{\lambda\zeta}] = e^{-\lambda}\left(\frac{1}{1-2\lambda}\right)^{1/2}. \tag{6.6.1}$$

We will need the following specific form of the exponential martingale inequality.

Lemma 6.6.2. *Let $N \in \mathbb{N}$. Let $(\xi(j))_{j\geq 1}$ be a sequence of independent standard normal random variables and define $\zeta(j) = \xi^2(j) - 1$ for $j \geq 1$. Let $(c_j)_{0\leq j\leq N-1}$ be a deterministic positive sequence. Define $M(0) = 0$ and*

$$M(n) = \sum_{j=0}^{n-1} c_j \zeta(j+1), \quad 1 \leq n \leq N.$$

Suppose that $\lambda_N > 0$ is such that $c_j\lambda_N < 1/2$ for all $0 \leq j \leq N-1$. Then for every $\beta_N > 0$ we have

$$\mathbb{P}\Big[\max_{1\leq n\leq N}\Big\{M(n) + \sum_{j=0}^{n-1} c_j - \frac{1}{2\lambda_N}\sum_{j=0}^{n-1}\log\Big(\frac{1}{1-2\lambda_N c_j}\Big)\Big\} \geq \beta_N\Big] \leq e^{-\lambda_N\beta_N}. \tag{6.6.2}$$

Careful perusal of the argument of Lemma 6.6.2 reveals that much of the argument holds for white noise sequences with general moment generating functions, but to simplify the proof (particularly regarding the domain of definition of the moment generating function in (6.6.1)) we specialise our calculations during the proof to cover merely the moment generating function given in (6.6.1). In the forthcoming Lemma 6.6.3, we apply Lemma 6.6.2 to a particular martingale by choosing carefully the sequences λ_N and β_N in the statement of Lemma 6.6.2.

Proof of Lemma 6.6.2. Let φ be given by (6.6.1). Since $c_j\lambda_N < 1/2$ for $j = 1,\dots,N-1$, it follows that $\varphi(\lambda_N c_j) > 0$ is well–defined and finite for each $j = 1,\dots,N-1$. Define $X(0) = 1$ and

$$X(n) = e^{\lambda_N M(n)} \prod_{j=0}^{n-1} \frac{1}{\varphi(\lambda_N c_j)}, \quad 1 \le n \le N.$$

We prove that X is a positive martingale (with respect to the appropriate filtration). Clearly $X(n)$ is positive a.s. for each $n \in \{1,\dots,N\}$ because $M(n)$ is almost surely finite for each $n \in \{1,\dots,N\}$. Now for $n \in \{1,\dots,N\}$ by the independence of $\zeta(j+1)$ for $j = 0,\dots,N-1$ and the finiteness of $\varphi(\lambda_N c_j)$ we have

$$\mathbb{E}\left[e^{\lambda_N M(n)}\right] = \mathbb{E}\left[\exp\left(\sum_{j=0}^{n-1} \lambda_N c_j \zeta(j+1)\right)\right] = \prod_{j=0}^{n-1} \mathbb{E}\left[\exp\left(\lambda_N c_j \zeta(j+1)\right)\right] = \prod_{j=0}^{n-1} \varphi(\lambda_N c_j),$$

which means that $\mathbb{E}[|X(n)|] = \mathbb{E}[X(n)]$ is finite for $n \in \{1,\dots,N\}$ and moreover $\mathbb{E}[X(n)] = 1$ for all $n \in \{0,\dots,N\}$. We now show that $\mathbb{E}[X(n+1)|\mathcal{F}(n)] = X(n)$ where we define $\mathcal{F}(n) := \sigma\{\zeta(j) : 1 \le j \le n\}$ for $n \in \{1,\dots,N\}$, so that each $\mathcal{F}(n)$ is a (naturally generated) sigma–algebra, and that $(\mathcal{F}(n))_{1\le n\le N}$ is the natural filtration generated by the sequence $(\zeta(n))_{1\le n\le N}$. We may denote by $\mathcal{F}(0)$ the trivial sigma–algebra. With this definition, $\zeta(1)$ is clearly independent of $\mathcal{F}(0)$ and we therefore get

$$\begin{aligned}\mathbb{E}[X(1)|\mathcal{F}(0)] &= \frac{1}{\varphi(\lambda_N c_0)}\mathbb{E}[e^{\lambda_N c_0 \zeta(1)}|\mathcal{F}(0)] = \frac{1}{\varphi(\lambda_N c_0)}\mathbb{E}[e^{\lambda_N c_0 \zeta(1)}] \\ &= \frac{1}{\varphi(\lambda_N c_0)}\varphi(\lambda_N c_0) = 1 = X(0).\end{aligned}$$

For $1 \le n \le N-1$ we have

$$X(n+1) = e^{\lambda_N M(n+1)} \prod_{j=0}^{n} \frac{1}{\varphi(\lambda_N c_j)} = e^{\lambda_N (M(n+1)-M(n))} \cdot \frac{1}{\varphi(\lambda_N c_n)} X(n),$$

so as c_n is deterministic, $\mathbb{E}[X(n)]$ and $\mathbb{E}[X(n)e^{\lambda_N c_n \zeta(n+1)}]$ are the finite expectations of non–negative random variables, $\zeta(n+1)$ is independent of $\mathcal{F}(n)$, and $\varphi(\lambda_N c_n)$ is finite, we get

$$\begin{aligned}\mathbb{E}[X(n+1)|\mathcal{F}(n)] &= \mathbb{E}[e^{\lambda_N(M(n+1)-M(n))}\cdot\frac{1}{\varphi(\lambda_N c_n)}X(n)|\mathcal{F}(n)]\\ &= \frac{1}{\varphi(\lambda_N c_n)}\mathbb{E}[e^{\lambda_N c_n\zeta(n+1)}X(n)|\mathcal{F}(n)] = \frac{1}{\varphi(\lambda_N c_n)}X(n)\mathbb{E}[e^{\lambda_N c_n\zeta(n+1)}|\mathcal{F}(n)]\\ &= \frac{1}{\varphi(\lambda_N c_n)}X(n)\mathbb{E}[e^{\lambda_N c_n\zeta(n+1)}] = X(n).\end{aligned}$$

Therefore we have that $X = \{X(n) : 0 \le n \le N\}$ is a positive martingale relative to $(\mathcal{F}(n))_{0\le n\le N}$ with $\mathbb{E}[X(N)] = 1$. Therefore, for any $\beta_N > 0$, by Doob's martingale inequality we have

$$\mathbb{P}\left[\max_{1\le n\le N} X(n) \ge e^{\lambda_N\beta_N}\right] \le \frac{\mathbb{E}[X(N)]}{e^{\lambda_N\beta_N}} = e^{-\lambda_N\beta_N}.$$

By the definition and positivity of X, and by using the fact that $\lambda_N > 0$ we get

$$\mathbb{P}\left[\max_{1\le n\le N}\left\{M(n) + \frac{1}{\lambda_N}\sum_{j=0}^{n-1}\log\left(\frac{1}{\varphi(\lambda_N c_j)}\right)\right\} \ge \beta_N\right] \le e^{-\lambda_N\beta_N}.$$

Since φ is given by (6.6.1), we have that

$$\frac{1}{\lambda_N}\log\left(\frac{1}{\varphi(\lambda_N c_j)}\right) = c_j + \frac{1}{\lambda_N}\log\left(1-2\lambda_N c_j\right)^{1/2},$$

and inserting this into the last estimate gives (6.6.2), as required. □

Lemma 6.6.2 is now used to establish the following estimate. It concerns a martingale which appears in the proof of Theorem 6.3.1.

Lemma 6.6.3. *Suppose $\sigma^2 > 0$, $h > 0$ and $\alpha_h > 1$. Let $(\xi(j))_{j\ge1}$ be a sequence of independent standard normal random variables. Define $M_h(0) = 0$ and*

$$M_h(n) = \sum_{j=0}^{n-1} h\sigma^2\alpha_h^{j+1}\left(\xi^2(j+1)-1\right), \quad n \ge 1. \tag{6.6.3}$$

Then

$$\limsup_{n\to\infty}\frac{M_h(n)}{\alpha_h^n\log n} \le 2h\sigma^2, \quad a.s. \tag{6.6.4}$$

Note that the normalising constant on the right hand side of (6.6.4) can be made as small as desired if $h > 0$ is chosen sufficiently small. This is precisely the kind of estimate we require to make the analysis of Mao's continuous time result sharp in the context of Theorem 6.3.1. Lemma 6.6.3 yields a sharper (and crucially, h–dependent) estimate than, for example, a simple term–by–term majorisation of the martingale M_h defined by (6.6.3). Such a majorisation leads to the normalising constant tending to a non–trivial limit as $h \to 0$.

Proof of Lemma 6.6.3. Let $N \geq 2$. Define $\zeta(j) = \xi^2(j) - 1$ for $j \geq 1$ and notice that M_h defined by (6.6.3) is a martingale relative to the natural filtration generated by $(\zeta(j))_{j\geq 1}$. Set $c_j = h\sigma^2\alpha_h^{j+1}$ for $0 \leq j \leq N-1$. Then $(c_j)_{0\leq j\leq N-1}$ is a positive sequence and M_h can be written as

$$M_h(n) = \sum_{j=0}^{n-1} c_j \zeta(j+1), \quad n \geq 1.$$

Define, for any fixed $\epsilon > 0$,

$$\lambda_N = \frac{1+\epsilon}{(2+4\epsilon)h\sigma^2}\alpha_h^{-N}, \quad \beta_N = (2+4\epsilon)h\sigma^2\alpha_h^N \log N.$$

Then $\lambda_N > 0$ and $\beta_N > 0$. Also for $j = 0, \ldots, N-1$ as $\alpha_h > 1$ we have

$$c_j\lambda_N = \frac{1+\epsilon}{2+4\epsilon}\alpha_h^{j+1-N} \leq \frac{1+\epsilon}{2+4\epsilon} < \frac{1}{2},$$

and $\lambda_N\beta_N = (1+\epsilon)\log N$. Hence we may apply Lemma 6.6.2 (specifically (6.6.2)) to M_h to obtain

$$\mathbb{P}\left[\max_{1\leq n\leq N}\left\{M_h(n) + \sum_{j=0}^{n-1} c_j - \frac{1}{2\lambda_N}\sum_{j=0}^{n-1}\log\left(\frac{1}{1-2\lambda_N c_j}\right)\right\} \geq \beta_N\right] \leq \frac{1}{N^{1+\epsilon}}.$$

By the Borel–Cantelli lemma, there exists Ω^* with $\mathbb{P}[\Omega^*] = 1$ such that for each $\omega \in \Omega^*$ there is an $N_0 = N_0(\omega) \in \mathbb{N}$ such that

$$\max_{1\leq n\leq N}\left\{M_h(n) + \sum_{j=0}^{n-1} c_j - \frac{1}{2\lambda_N}\sum_{j=0}^{n-1}\log\left(\frac{1}{1-2\lambda_N c_j}\right)\right\} < \beta_N, \quad N \geq N_0.$$

Therefore for $N \geq N_0$ we have

$$M_h(N) + \sum_{j=0}^{N-1} c_j - \frac{1}{2\lambda_N}\sum_{j=0}^{N-1}\log\left(\frac{1}{1-2\lambda_N c_j}\right)$$
$$\leq \max_{1\leq n\leq N}\left\{M_h(n) + \sum_{j=0}^{n-1} c_j - \frac{1}{2\lambda_N}\sum_{j=0}^{n-1}\log\left(\frac{1}{1-2\lambda_N c_j}\right)\right\} \leq \beta_N.$$

Since $c_j > 0$ for $1 \leq j \leq N-1$, and by using the definitions of λ_N and β_N, for $N \geq N_0$ we get the estimate

$$\begin{aligned}
M_h(N) &\leq \beta_N + \frac{1}{2\lambda_N}\sum_{j=0}^{N-1}\log\left(\frac{1}{1-2\lambda_N c_j}\right) - \sum_{j=0}^{N-1} c_j \\
&\leq (2+4\epsilon)h\sigma^2\alpha_h^N \log N + \frac{1+2\epsilon}{1+\epsilon}h\sigma^2\alpha_h^N \sum_{j=0}^{N-1}\log\left(1-\frac{1+\epsilon}{1+2\epsilon}\alpha_h^{-N+j+1}\right)^{-1} \\
&= (2+4\epsilon)h\sigma^2\alpha_h^N \log N + \frac{1+2\epsilon}{1+\epsilon}h\sigma^2\alpha_h^N \sum_{l=0}^{N-1} -\log\left(1-\frac{1+\epsilon}{1+2\epsilon}\alpha_h^{-l}\right).
\end{aligned}$$

Let $x = (1+\epsilon)\alpha_h^{-l}/(1+2\epsilon)$ and note that $x \leq (1+\epsilon)/(1+2\epsilon)$ since $\alpha_h > 1$. Now by Taylor's theorem, for every $x \in (0, (1+\epsilon)/(1+2\epsilon)]$ there is $c_x \in (0, x]$ such that

$$-\log(1-x) = x + \frac{1}{2(1-c_x)^2}x^2 \leq x + \frac{(1+2\epsilon)^2}{2\epsilon^2}x^2,$$

where we have used the fact that $c_x \in (0, (1+\epsilon)/(1+2\epsilon)]$ at the last step. Therefore for $N \geq N_0$, we have

$$\begin{aligned}
M_h(N) &\leq (2+4\epsilon)h\sigma^2\alpha_h^N \log N + \frac{1+2\epsilon}{1+\epsilon}h\sigma^2\alpha_h^N \sum_{l=0}^{N-1}\left\{\frac{1+\epsilon}{1+2\epsilon}\alpha_h^{-l} + \frac{(1+\epsilon)^2}{2\epsilon^2}\alpha_h^{-2l}\right\} \\
&\leq (2+4\epsilon)h\sigma^2\alpha_h^N \log N + h\sigma^2\alpha_h^N \sum_{l=0}^{\infty}\alpha_h^{-l} + \frac{(1+2\epsilon)(1+\epsilon)}{2\epsilon^2}h\sigma^2\alpha_h^N \sum_{l=0}^{\infty}\alpha_h^{-2l} \\
&= (2+4\epsilon)h\sigma^2\alpha_h^N \log N + \frac{h\sigma^2\alpha_h^N}{1-\alpha_h^{-1}} + \frac{(1+2\epsilon)(1+\epsilon)}{2\epsilon^2(1-\alpha_h^{-2})}h\sigma^2\alpha_h^N.
\end{aligned}$$

Recalling that this estimate holds for all N sufficiently large on every sample path of an a.s. event, we see that

$$\limsup_{N\to\infty}\frac{M_h(N)}{\alpha_h^N \log N} \leq (2+4\epsilon)h\sigma^2, \quad \text{a.s.},$$

and finally upon letting $\epsilon \to 0$ the result (6.6.4) follows. This completes the proof. □

6.7 Proofs of Discrete Results from Section 6.3

In order to make our exposition self–contained, we first state the version of discrete Gronwall inequality that we use in this chapter. Since many forms of the Gronwall inequality can be formulated, we omit its proof.

Lemma 6.7.1. *Let $a > 0$ and $c > 0$. Let $y(\cdot)$ and $b(\cdot)$ be nonnegative sequences. If*

$$y(n) \leq a + c \sum_{j=0}^{n-1} b(j) y(j), \quad n \geq 1,$$

then

$$y(n) \leq a \prod_{j=0}^{n-1} (1 + c\, b(j)), \quad n \geq 1.$$

We are now in a position to prove Theorem 6.3.1.

Proof of Theorem 6.3.1. First fix the step size $h < \min(\gamma/\Gamma^2, (2\gamma)^{-1})$ where the constants Γ and γ are defined in Assumption 6.2.1 and Assumption 6.2.2 respectively. Recall that $\Gamma_1 = \Gamma(1 + \upsilon)$ and $\gamma_1 = \gamma(1 - \eta)$ where $\upsilon, \eta \in (0, 1)$ are arbitrary. For simplicity, choose $\upsilon = \eta$ so that $\Gamma_1 = \Gamma(1 + \eta)$. Note that if $h < \gamma/\Gamma^2$ then $h\Gamma^2/\gamma < 1$. Since this inequality is strict it means that for all $\eta \in (0, 1)$ sufficiently small we can ensure that $h\Gamma^2/\gamma < (1 - \eta)/(1 + \eta)^2 < 1$ which implies that $h\Gamma_1^2/\gamma_1 < 1$. Moreover, if $h < 1/(2\gamma)$ then it is immediately true that $h < 1/(2\gamma_1)$ for any value of η. Therefore the step size restriction $h < \min(\gamma/\Gamma^2, (2\gamma)^{-1})$ implies also that $h < \min(\gamma_1/\Gamma_1^2, (2\gamma_1)^{-1})$. This allows us to define $\alpha_h := \left(1 - 2h\gamma_1 + 2h^2\Gamma_1^2\right)^{-1}$ where $\alpha_h > 1$. Squaring (6.3.1) and multiplying across by α_h^{n+1} gives

$$\begin{aligned}\alpha_h^{n+1} X_h^2(n+1) &= \alpha_h^{n+1} X_h^2(n) + 2h\alpha_h^{n+1} X_h(n) f(X_h(n), nh) \\ &\quad + h^2 \alpha_h^{n+1} f^2(X_h(n), nh) + 2\sigma\sqrt{h}\alpha_h^{n+1} X_h(n)\xi(n+1) \\ &\quad + 2\sigma h\sqrt{h}\alpha_h^{n+1} f(X_h(n), nh)\xi(n+1) + \sigma^2 h \alpha_h^{n+1} \xi^2(n+1).\end{aligned}$$

Note that by (6.2.4) and the fact that $2xy \le x^2 + y^2$ for any $x, y \in \mathbb{R}$ we have

$$f^2(x, nh) \le (\Gamma_0 + \Gamma_1|x|)^2 = \Gamma_0^2 + 2\Gamma_0\Gamma_1|x| + \Gamma_1^2 x^2 \le 2\Gamma_0^2 + 2\Gamma_1^2 x^2.$$

Using this and (6.2.6) we get

$$\begin{aligned}\alpha_h^{n+1} X_h^2(n+1) \le\, & \alpha_h^{n+1} X_h^2(n) + \alpha_h^n X_h^2(n) - \alpha_h^n X_h^2(n) + 2h\alpha_h^{n+1}\left[\rho_1 - \gamma_1 X_h^2(n)\right] \\ & + 2h^2\alpha_h^{n+1}\Gamma_0^2 + 2h^2\alpha_h^{n+1}\Gamma_1^2 X_h^2(n) + 2\sigma\sqrt{h}\alpha_h^{n+1} X_h(n)\xi(n+1) \\ & + 2\sigma h\sqrt{h}\alpha_h^{n+1} f(X_h(n), nh)\xi(n+1) + \sigma^2 h\alpha_h^{n+1}\xi^2(n+1).\end{aligned}$$

By the definition of α_h we have

$$\alpha_h^{n+1} X_h^2(n) - \alpha_h^n X_h^2(n) - 2h\alpha_h^{n+1}\gamma_1 X_h^2(n) + 2h^2\alpha_h^{n+1}\Gamma_1^2 X_h^2(n) = 0,$$

so that our equation above becomes

$$\begin{aligned}\alpha_h^{n+1} X_h^2(n+1) \le\, & \alpha_h^n X_h^2(n) + 2h\alpha_h^{n+1}\rho_1 + 2h^2\alpha_h^{n+1}\Gamma_0^2 + \alpha_h^{n+1}\sigma^2 h \\ & + 2\sigma\sqrt{h}\alpha_h^{n+1} F_h(X_h(n))\xi(n+1) + \sigma^2 h\alpha_h^{n+1}[\xi^2(n+1) - 1] \qquad (6.7.1)\end{aligned}$$

where $F_h(x) := x + hf(x, nh)$. Define the martingale differences

$$\Delta M_h(n+1) = \sigma^2 h\alpha_h^{n+1}\left[\xi^2(n+1) - 1\right], \quad n \ge 0,$$

$$\Delta\theta_h(n+1) = 2\sigma\sqrt{h}\alpha_h^{n+1} F_h(X_h(n))\xi(n+1), \quad n \ge 0,$$

so that $M_h(n+1) := \sum_{j=0}^n \Delta M_h(j+1)$ and $\theta_h(n+1) := \sum_{j=0}^n \Delta\theta_h(j+1)$ are martingales. We also define $\lambda_h := 2\rho_1 + \sigma^2 + 2\Gamma_0^2 h$ so that (6.7.1) reduces to

$$\alpha_h^{n+1} X_h^2(n+1) - \alpha_h^n X_h^2(n) \le h\lambda_h\alpha_h^{n+1} + \Delta M_h(n+1) + \Delta\theta_h(n+1).$$

Summing on both sides yields for $n \ge 0$,

$$\alpha_h^{n+1} X_h^2(n+1) - X_h^2(0) \le h\lambda_h \sum_{j=0}^n \alpha_h^{j+1} + M_h(n+1) + \theta_h(n+1). \qquad (6.7.2)$$

We now apply the exponential martingale inequality to the martingale $\theta_h(n+1)$ which has quadratic variation given by

$$\langle\theta_h\rangle(n) = \sum_{j=0}^{n-1} 4\sigma^2 h\alpha_h^{2j+2}F_h^2(X_h(j)), \quad n \geq 1.$$

Applying Theorem 5.2.4, where $\beta > 0$, $\delta > 0$ and $\tau > 1$ are arbitrary constants, we have for all $n \in \mathbb{N}$,

$$\mathbb{P}\Big[\max_{1\leq m\leq \lfloor n\delta\rfloor}\Big\{\theta_h(m) - 2\beta\alpha_h^{-n\delta}\sum_{j=0}^{m-1}\sigma^2 h\alpha_h^{2j+2}F_h^2(X_h(j))\Big\} \geq \beta^{-1}\alpha_h^{n\delta}\tau\log n\Big] \leq \frac{1}{n^\tau}.$$

The Borel–Cantelli lemma then yields that for all $\omega \in \Omega_0$, where $\mathbb{P}[\Omega_0] = 1$, there is a random integer $n_0 = n_0(\omega, h)$ sufficiently large such that for $n \geq n_1 := n_0(\omega, h) \vee \lceil 1/\delta\rceil \vee 2$,

$$\theta_h(m) \leq \beta^{-1}\alpha_h^{n\delta}\tau\log n + 2\beta\alpha_h^{-n\delta}\sum_{j=0}^{m-1}\sigma^2 h\alpha_h^{2j+2}F_h^2(X_h(j)), \quad 1 \leq m \leq \lfloor n\delta\rfloor.$$

For the next segment of the proof, we view the path ω as fixed and suppress it in our notation. The following argument is, for the most part, deterministic in spirit. Using (6.2.4) and (6.2.6) it can be shown that

$$F_h^2(x) \leq 2h\rho_1 + 2h^2\Gamma_0^2 + x^2(1 - 2h\gamma_1 + 2h^2\Gamma_1^2) = (2h\rho_1 + 2h^2\Gamma_0^2) + x^2\alpha_h^{-1}$$

so that for $n \geq n_1$,

$$\theta_h(m) \leq \beta^{-1}\alpha_h^{n\delta}\tau\log n + 2\beta\alpha_h^{-n\delta}\sigma^2 h(2h\rho_1 + 2h^2\Gamma_0^2)\sum_{j=0}^{m-1}\alpha_h^{2j+2} \tag{6.7.3}$$

$$+ 2\beta\alpha_h^{-n\delta}\sigma^2 h\alpha_h^{-1}\sum_{j=0}^{m-1}\alpha_h^{2j+2}X_h^2(j), \qquad 1 \leq m \leq \lfloor n\delta\rfloor.$$

Define

$$T_h(m) := X_h^2(0) + h\lambda_h\sum_{j=0}^{m-1}\alpha_h^{j+1} + 2\beta\alpha_h^{-n\delta}\sigma^2 h(2h\rho_1 + 2h^2\Gamma_0^2)\sum_{j=0}^{m-1}\alpha_h^{2j+2} + M_h(m)$$

so that (6.7.2) along with (6.7.3) gives, for $1 \leq m \leq \lfloor n\delta\rfloor$ and $n \geq n_1$,

$$\alpha_h^m X_h^2(m) \leq T_h(m) + \beta^{-1}\alpha_h^{n\delta}\tau\log n + 2\beta\alpha_h^{-n\delta}\sum_{j=0}^{m-1}\sigma^2 h\alpha_h^{2j+1}X_h^2(j). \tag{6.7.4}$$

Notice that since $m \le \lfloor n\delta \rfloor \le n\delta$ and $2h\alpha_h/(\alpha_h - 1) = 1/(\gamma_1 - h\Gamma_1^2)$ we have

$$T_h(m) \le X_h^2(0) + \frac{\lambda_h}{2(\gamma_1 - h\Gamma_1^2)}\,\alpha_h^m + \frac{\beta\sigma^2(2h\rho_1 + 2h^2\Gamma_0^2)\alpha_h}{(\alpha_h+1)(\gamma_1 - h\Gamma_1^2)}\,\alpha_h^m + M_h(m)$$

and so by Lemma 6.6.3,

$$\limsup_{m\to\infty} \frac{T_h(m)}{\alpha_h^m \log m} \le \limsup_{m\to\infty} \frac{M_h(m)}{\alpha_h^m \log m} \le 2h\sigma^2, \quad \text{a.s. on } \Omega_1.$$

Now consider outcomes in $\Omega_2 := \Omega_0 \cap \Omega_1$ where $\mathbb{P}[\Omega_2] = 1$. For every $\varepsilon \in (0,1)$ there exists a number $m_0(\varepsilon) \in \mathbb{N}$ such that for $m_0(\varepsilon) \le m \le \lfloor n\delta \rfloor$ we have $T_h(m) \le 2h\sigma^2(1 + \varepsilon)\alpha_h^m \log m \le 2h\sigma^2(1+\varepsilon)\alpha_h^{n\delta}\log n\delta$. Moreover, on the finite set $m \in \{1, \dots, m_0(\varepsilon)\}$ there exists a number $T_1(\varepsilon, h) > 0$ such that $T_h(m) \le T_1(\varepsilon, h)$. Altogether, for $1 \le m \le \lfloor n\delta \rfloor$, we have

$$T_h(m) \le T_1(\varepsilon, h) + 2h\sigma^2(1+\varepsilon)\alpha_h^{n\delta}\log n\delta =: T^*_{h,\varepsilon}(n).$$

So now, returning to (6.7.4), for $n \ge n_1$ and $1 \le m \le \lfloor n\delta \rfloor$ we have

$$\alpha_h^m X_h^2(m) \le T^*_{h,\varepsilon}(n) + \beta^{-1}\alpha_h^{n\delta}\tau\log n + 2\beta\alpha_h^{-n\delta}\sum_{j=0}^{m-1}\sigma^2 h\alpha_h^{2j+1}X_h^2(j).$$

Following the notation of Lemma 6.7.1 we set $y(m) := \alpha_h^m X_h^2(m)$, $a_h(n) := T^*_{h,\varepsilon}(n) + \beta^{-1}\alpha_h^{n\delta}\tau\log n$, $c(n) := 2\beta\alpha_h^{-n\delta}$ and $b_h(j) := \sigma^2 h\alpha_h^{j+1}$. Therefore, for $n \ge n_1$,

$$y(m) \le a_h(n) + c(n)\sum_{j=0}^{m-1} b_h(j)y(j), \quad 1 \le m \le \lfloor n\delta \rfloor,$$

and so we can apply Lemma 6.7.1 to conclude that for $n \ge n_1$

$$y(m) \le a_h(n)\prod_{j=0}^{m-1}\big(1 + c(n)b_h(j)\big), \qquad 1 \le m \le \lfloor n\delta \rfloor.$$

Then using the fact that $1 + x \le e^x$ for $x \ge 0$ we get

$$\alpha^m X_h^2(m) \le a_h(n)\exp\Big[2\beta\alpha_h^{-n\delta}\sigma^2 h\sum_{j=0}^{m-1}\alpha_h^{j+1}\Big], \quad 1 \le m \le \lfloor n\delta \rfloor, \quad n \ge n_1.$$

Notice that

$$2h\alpha_h^{-n\delta}\sum_{j=0}^{m-1}\alpha_h^{j+1} \le 2h\alpha_h^{-n\delta}\sum_{j=0}^{\lfloor n\delta\rfloor - 1}\alpha_h^{j+1} \le \frac{2h\alpha_h}{\alpha_h - 1} = \frac{1}{\gamma_1 - h\Gamma_1^2}.$$

Now let n_2 be such that $\lfloor (n_2-1)\delta \rfloor > 1$ and let $n_3 := n_1 \vee n_2$. Then for $1 \le \lfloor (n-1)\delta \rfloor \le m \le \lfloor n\delta \rfloor$ and $n \ge n_3$,

$$\alpha_h^{\lfloor (n-1)\delta \rfloor} X_h^2(m) \le \alpha_h^m X_h^2(m) \le a_h(n) \exp\Big[\frac{\beta\sigma^2}{\gamma_1 - h\Gamma_1^2}\Big],$$

and moreover, since $\lfloor (n-1)\delta \rfloor \ge (n-1)\delta - 1$,

$$\begin{aligned}\frac{X_h^2(m)}{\log m} &\le \frac{a_h(n)}{\alpha_h^{n\delta}\log n\delta}.\frac{\log n\delta}{\log m}\alpha_h^{n\delta - \lfloor (n-1)\delta \rfloor} \exp\Big[\frac{\beta\sigma^2}{\gamma_1 - h\Gamma_1^2}\Big] \\ &\le \frac{a_h(n)}{\alpha_h^{n\delta}\log n\delta}.\frac{\log n\delta}{\log \lfloor (n-1)\delta \rfloor}\alpha_h^{\delta+1} \exp\Big[\frac{\beta\sigma^2}{\gamma_1 - h\Gamma_1^2}\Big]\end{aligned}$$

Now notice that

$$\limsup_{n\to\infty} \frac{a_h(n)}{\alpha_h^{n\delta}\log n\delta} = 2h\sigma^2(1+\varepsilon) + \beta^{-1}\tau \quad \text{and} \quad \lim_{n\to\infty} \frac{\log n\delta}{\log \lfloor (n-1)\delta \rfloor} = 1.$$

Therefore, since $n \to \infty$ as $m \to \infty$ we have

$$\limsup_{m\to\infty} \frac{X_h^2(m)}{\log m} \le (2h\sigma^2(1+\varepsilon) + \beta^{-1}\tau)\alpha_h^{\delta+1} \exp\Big[\frac{\beta\sigma^2}{\gamma_1 - h\Gamma_1^2}\Big].$$

Since this analysis holds on the same event for arbitrary $\varepsilon > 0$, we have

$$\limsup_{m\to\infty} \frac{X_h^2(m)}{\log m} \le (2h\sigma^2 + \beta^{-1}\tau)\alpha_h^{\delta+1} \exp\Big[\frac{\beta\sigma^2}{\gamma_1 - h\Gamma_1^2}\Big].$$

This estimate holds for all outcomes in an event of probability one, say $\Omega_{\tau,\delta}$. By considering $\Omega^* := \cap_{p\in\mathbb{N}}\Omega_{1+\frac{1}{p},\frac{1}{p}}$, i.e. by letting $\tau \to 1$ and $\delta \to 0$, we get

$$\limsup_{m\to\infty} \frac{X_h^2(m)}{\log m} \le (2h\sigma^2 + \beta^{-1})\alpha_h \exp\Big[\frac{\beta\sigma^2}{\gamma_1 - h\Gamma_1^2}\Big], \quad \text{a.s. on } \Omega^*,$$

where Ω^* is an almost sure event. Choose $\beta = (\gamma_1 - h\Gamma_1^2)/\sigma^2$ to get

$$\begin{aligned}\limsup_{m\to\infty} \frac{X_h^2(m)}{\log mh} = \limsup_{m\to\infty} \frac{X_h^2(m)}{\log m} &\le \alpha_h\Big(2h\sigma^2 + \frac{\sigma^2}{\gamma_1 - h\Gamma_1^2}\Big)\, e^1 \\ &= \sigma^2\Big(\frac{1+2h\gamma_1 - 2h^2\Gamma_1^2}{1-2h\gamma_1 + 2h^2\Gamma_1^2}\Big)\Big(\frac{\gamma_1}{\gamma_1 - h\Gamma_1^2}\Big)\frac{e}{\gamma_1}, \quad \text{a.s.}\end{aligned}$$

Finally, recall that $\gamma_1 = \gamma(1-\eta)$ and $\Gamma_1 = \Gamma(1+\eta)$ where $\eta \in (0,1)$ is sufficiently small. Letting $\eta \to 0$ and taking square roots on both sides gives the desired result. □

Proof of Theorem 6.3.2. First fix the step size $h < 2/\Gamma$ where Γ arises from Assumption 6.2.1. Recall that $\Gamma_1 = \Gamma(1+\upsilon)$ where $\upsilon \in (0,1)$ is arbitrary and note that if $h < 2/\Gamma$ then $h\Gamma/2 < 1$. Since this inequality is strict it means that for all $\upsilon \in (0,1)$ sufficiently small we can ensure that

$$\frac{h\Gamma}{2} < \frac{1}{(1+\upsilon)} < 1 \quad \text{which implies that} \quad \frac{h\Gamma_1}{2} < 1.$$

Therefore the step size restriction $h < 2/\Gamma$ implies also that $h < 2/\Gamma_1$. This allows us to define $\alpha_h := 1/(1-h\mu)$ where $\mu := \Gamma_1/2$ and $\alpha_h > 1$. Multiplying (6.3.1) by α_h^{n+1} gives

$$\alpha_h^{n+1} X_h(n+1) = \alpha_h^{n+1} X_h(n) + \alpha_h^{n+1} h f(X_h(n), nh) + \alpha_h^{n+1}\sigma\sqrt{h}\xi(n+1), \quad n \geq 0.$$

Therefore by adding and subtracting terms we can write this as

$$\alpha_h^{n+1} X_h(n+1) - \alpha_h^{n} X_h(n) = h\alpha_h^{n+1} F_h(X_h(n)) + \alpha_h^{n+1}\sigma\sqrt{h}\xi(n+1), \quad n \geq 0,$$

where $F_h(x) := \mu x + f(x,nh)$ and $\alpha_h^n X_h(n)[\alpha_h - 1] - h\mu\alpha_h^{n+1} X_h(n) = 0$ by the construction of α_h. Summing on both sides then yields

$$\alpha_h^{n+1} X_h(n+1) - X_h(0) = h\sum_{j=0}^{n} \alpha_h^{j+1} F_h(X_h(j)) + \sum_{j=0}^{n} \alpha_h^{j+1}\sigma\sqrt{h}\xi(j+1). \tag{6.7.5}$$

Define $Y_h(n+1) := \alpha_h^{-(n+1)} \sum_{j=0}^{n} \alpha_h^{j+1}\sigma\sqrt{h}\xi(j+1)$ for $n \geq 0$, and notice that $Y_h(n+1) \sim \mathcal{N}(0, V(n+1))$ where

$$V(n) := \text{Var}[Y_h(n)] = \frac{\sigma^2 h}{\alpha_h^{2n}} \sum_{j=0}^{n-1} \alpha_h^{2j+2} = \frac{\sigma^2 h}{\alpha_h^{2n}} \left[\frac{\alpha_h^2(\alpha_h^{2n}-1)}{\alpha_h^2 - 1}\right] = \frac{\sigma^2 h \alpha_h^2(1-\alpha_h^{-2n})}{\alpha_h^2 - 1}.$$

Since $V(n)$ is increasing it follows that for any $k \geq 1$ and $m \geq 0$ we have $V(k) \leq V(k+m)$ which in turn means that $V(k) \leq \sqrt{V(k)V(k+m)}$. Now consider $Z_h(k) := Y_h(k)/\sqrt{V(k)}$ for $k \geq 1$. Then for every k, $Z_h(k)$ is a standard normal random variable. Moreover, for $k \geq 1$ and $m \geq 0$,

$$0 \leq \text{Cov}(Z_h(k), Z_h(k+m)) = \frac{\text{Cov}\big(Y_h(k), Y_h(k+m)\big)}{\sqrt{V(k)V(k+m)}} \leq \frac{\text{Cov}(Y_h(k), Y_h(k+m))}{V(k)},$$

where

$$\begin{aligned}\mathrm{Cov}(Y_h(k), Y_h(k+m)) &= \frac{\sigma^2 h}{\alpha_h^k \alpha_h^{k+m}} \mathrm{Cov}\Big(\sum_{j=0}^{k-1} \alpha_h^{j+1}\xi(j+1), \sum_{l=0}^{k+m-1} \alpha_h^{l+1}\xi(l+1)\Big) \\ &= \frac{\sigma^2 h}{\alpha_h^k \alpha_h^{k+m}} \sum_{j=0}^{k-1} \alpha_h^{j+1}\alpha_h^{j+1}.1 = \frac{\sigma^2 h}{\alpha_h^{2k}\alpha_h^m} \frac{\alpha_h^2(\alpha_h^{2k}-1)}{\alpha_h^2-1} = \frac{V(k)}{\alpha_h^m}.\end{aligned}$$

Thus for $k \geq 1$ and $m \geq 0$,

$$\mathrm{Cov}(Z_h(k), Z_h(k+m)) \leq \frac{\mathrm{Cov}(Y_h(k), Y_h(k+m))}{V(k)} = \alpha_h^{-m}.$$

Therefore by Lemma 2 and Lemma 3 in [6], it follows that

$$\limsup_{n\to\infty} \frac{|Z_h(n)|}{\sqrt{2\log n}} = 1, \quad \text{a.s.},$$

and using the fact that $h\alpha_h^2/(\alpha_h^2-1) = 1/\mu(2-h\mu)$ we have

$$\limsup_{n\to\infty} \frac{|Y_h(n)|}{\sqrt{\log nh}} = \limsup_{n\to\infty} \frac{|Z_h(n)|}{\sqrt{2\log n}} \sqrt{2V(n)} = \sqrt{\frac{2\sigma^2}{\mu(2-h\mu)}}, \quad \text{a.s.} \tag{6.7.6}$$

Now let $c > 0$ and define the event

$$A_c := \Big\{\omega : \limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{\log nh}} < c\Big\} \tag{6.7.7}$$

and assume that $\mathbb{P}[A_c] > 0$. We will demonstrate that this is impossible for the appropriate choice of c. By (6.7.5) and the triangle inequality it follows that

$$\begin{aligned}\limsup_{n\to\infty} \frac{|Y_h(n+1)|}{\sqrt{\log(n+1)h}} &\leq \limsup_{n\to\infty} \frac{|X_h(n+1)|}{\sqrt{\log(n+1)h}} + \limsup_{n\to\infty} \frac{|X_h(0)|}{\alpha_h^{n+1}\sqrt{\log(n+1)h}} \\ &\quad + \limsup_{n\to\infty} \frac{h\sum_{j=0}^n \alpha_h^{j+1}|F_h(X_h(j))|}{\alpha_h^{n+1}\sqrt{\log(n+1)h}} \\ &< c + \limsup_{n\to\infty} \frac{h\sum_{j=0}^n \alpha_h^{j+1}|F_h(X_h(j))|}{\alpha_h^{n+1}\sqrt{\log(n+1)h}}, \quad \text{a.s. on } A_c.\end{aligned} \tag{6.7.8}$$

Using (6.2.4) we have $|F_h(x)| \leq \mu|x| + |f(x, nh)| \leq (\mu+\Gamma_1)|x| + \Gamma_0$ for all x, and so

$$h\sum_{j=0}^n \alpha_h^{j+1}|F_h(X_h(j))| \leq h(\mu+\Gamma_1)\sum_{j=0}^n \alpha_h^{j+1}|X_h(j)| + h\Gamma_0\sum_{j=0}^n \alpha_h^{j+1}. \tag{6.7.9}$$

By (6.7.7) it follows that for all $\varepsilon > 0$ there exists $N_1(\varepsilon,\omega) > 0$ such that for $n \geq N_1(\varepsilon,\omega)$ we have $|X_h(n,\omega)| \leq c(1+\varepsilon)\sqrt{\log nh}$, for all $\omega \in A_c$. Therefore,

$$\begin{aligned} h\sum_{j=0}^{n} \alpha_h^{j+1}|X_h(j,\omega)| &\leq h \sum_{j=0}^{N_1(\varepsilon,\omega)-1} \alpha_h^{j+1}|X_h(j,\omega)| + c(1+\varepsilon)h \sum_{N_1(\varepsilon,\omega)}^{n} \alpha_h^{j+1}\sqrt{\log jh} \\ &\leq X_h^*(\varepsilon,\omega)h \sum_{j=0}^{N_1(\varepsilon,\omega)-1} \alpha_h^{j+1} + c(1+\varepsilon)h \sum_{N_1(\varepsilon,\omega)}^{n} \alpha_h^{j+1}\sqrt{\log jh}, \quad (6.7.10) \end{aligned}$$

where $X_h^*(\varepsilon,\omega) := \max_{0\leq j\leq N_1(\varepsilon,\omega)} |X_h(j,\omega)| < +\infty$. Notice that

$$\begin{aligned} h \sum_{j=N_1(\varepsilon,\omega)}^{n} \alpha_h^{j+1}\sqrt{\log jh} &= h\{\alpha_h^{N_1(\varepsilon,\omega)+1}\sqrt{\log N_1(\varepsilon,\omega)h} + \cdots + \alpha_h^{n+1}\sqrt{\log nh}\} \\ &\leq h\sqrt{\log nh}\{\alpha_h^{N_1(\varepsilon,\omega)+1} + \cdots + \alpha_h^{n+1}\} \\ &= \frac{\sqrt{\log nh}}{\mu}(\alpha_h^{n+1} - \alpha_h^{N_1(\varepsilon,\omega)}) \leq \frac{\sqrt{\log nh}}{\mu}\alpha_h^{n+1} \end{aligned}$$

where we have used the fact that $h\alpha_h/(\alpha_h - 1) = 1/\mu$. Therefore,

$$\limsup_{n\to\infty} \frac{h\sum_{j=N_1(\varepsilon,\omega)}^{n} \alpha_h^{j+1}\sqrt{\log jh}}{\alpha_h^{n+1}\sqrt{\log{(n+1)h}}} \leq \frac{1}{\mu},$$

and so using this along with equations (6.7.9) and (6.7.10) we get

$$\limsup_{n\to\infty} \frac{h\sum_{j=0}^{n} \alpha_h^{j+1}|F_h(X_h(j,\omega))|}{\alpha_h^{n+1}\sqrt{\log{(n+1)h}}} \leq \frac{(\mu+\Gamma_1)c(1+\varepsilon)}{\mu}, \quad \text{for all } \omega \in A_c.$$

Moreover, since $\varepsilon > 0$ can be chosen arbitrarily we can allow $\varepsilon \to 0$ through the rationals to get

$$\limsup_{n\to\infty} \frac{h\sum_{j=0}^{n} \alpha_h^{j+1}|F_h(X_h(j,\omega))|}{\alpha_h^{n+1}\sqrt{\log{(n+1)h}}} \leq \frac{(\mu+\Gamma_1)c}{\mu}, \quad \text{for all } \omega \in A_c.$$

Therefore by (6.7.6) and (6.7.8),

$$\sqrt{\frac{2\sigma^2}{\mu(2-h\mu)}} = \limsup_{n\to\infty}\frac{|Y_h(n+1)|}{\sqrt{\log{(n+1)h}}} < c + \frac{(\mu+\Gamma_1)c}{\mu}, \quad \text{a.s. on } A_c.$$

However, we get a contradiction above if we make the choice

$$c = \frac{\sqrt{2\sigma^2\mu/(2-h\mu)}}{2\mu+\Gamma_1},$$

which means that our assumption that $\mathbb{P}[A_c] > 0$ is incorrect. Thus, $\mathbb{P}[\bar{A}_c] = 1$ and

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{\log nh}} \geq c, \quad \text{a.s.}$$

Then recalling that $\mu = \Gamma_1/2$ we get

$$\limsup_{n\to\infty} \frac{|X_h(n)|}{\sqrt{\log nh}} \geq \frac{|\sigma|}{2\sqrt{2 - h\Gamma_1/2}} . \frac{1}{\sqrt{\Gamma_1}}, \quad \text{a.s.}$$

Finally, we reintroduce the fact that $\Gamma_1 = \Gamma(1 + \upsilon)$ where $\upsilon \in (0, 1)$ is sufficiently small. Letting $\upsilon \to 0$ proves the desired result. □

Proof of Theorem 6.3.3. By squaring (6.3.1) we can show that for $n \geq 0$,

$$X_h^2(n+1) - X_h^2(n) = \lambda_h(X_h(n)) + 2\sigma\sqrt{h}F_h(X_h(n))\xi(n+1) + \sigma^2 h[\xi^2(n+1) - 1], \tag{6.7.11}$$

where $\lambda_h(x) := 2hxf(x, nh) + h^2 f^2(x, nh) + \sigma^2 h$ and $F_h(x) := hf(x, nh) + x$. Define the martingale differences

$$\Delta M_h(n+1) := \sigma^2 h[\xi^2(n+1) - 1], \quad n \geq 0,$$

$$\Delta\theta_h(n+1) := 2\sigma\sqrt{h}F_h(X_h(n))\xi(n+1), \quad n \geq 0,$$

so that $M_h(n+1) := \sum_{j=0}^{n} \Delta M_h(j+1)$ and $\theta_h(n+1) := \sum_{j=0}^{n} \Delta\theta_h(j+1)$ are martingales. Thus, returning to (6.7.11) and summing on both sides yields

$$X_h^2(n+1) - X_h^2(0) = \sum_{j=0}^{n} \lambda_h(X_h(j)) + M_h(n+1) + \theta_h(n+1), \quad n \geq 0. \tag{6.7.12}$$

It is clear, by the strong law of large numbers, that there is an almost sure event Ω_1 such that on Ω_1,

$$\lim_{n\to\infty} \frac{M_h(n)}{n} = \sigma^2 h \lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} [\xi^2(j+1) - 1] = \sigma^2 h \mathbb{E}[\xi^2(j+1) - 1] = 0, \tag{6.7.13}$$

where we are using the fact that ξ has a standard normal distribution. We now apply the exponential martingale inequality to the martingale $\theta_h(n+1)$ which has quadratic variation given by

$$\langle \theta_h \rangle(n+1) = \sum_{j=0}^{n} 4\sigma^2 h F_h^2(X_h(j)), \quad n \geq 0.$$

Applying Theorem 5.2.4, where $\beta > 0$ and $\tau > 1$ are arbitrary constants, we have for all $n \in \mathbb{N}$,

$$\mathbb{P}\Big[\max_{1 \leq m \leq \lfloor \tau^n \rfloor} \Big\{\theta_h(m) - \frac{\beta\tau^{-n}}{2} \sum_{j=0}^{m-1} 4\sigma^2 h F_h^2(X_h(j))\Big\} \geq \beta^{-1}\tau^{n+1} \log n\Big] \leq \frac{1}{n^\tau}.$$

The Borel–Cantelli lemma then yields that for all $\omega \in \Omega_0$, where $\mathbb{P}[\Omega_0] = 1$, there is a random integer $n_0 = n_0(\omega, h)$ sufficiently large such that $\lfloor \tau^{n_0 - 1} \rfloor > e^1$ and for $n \geq n_0$ we have

$$\theta_h(m) \leq \beta^{-1}\tau^{n+1} \log n + 2\beta\tau^{-n} \sum_{j=0}^{m-1} \sigma^2 h F_h^2(X_h(j)), \quad 1 \leq m \leq \lfloor \tau^n \rfloor.$$

For the next part of the proof we view the path ω as fixed and suppress it in our notation. The following argument is, for the most part, deterministic in spirit. Using (6.2.9) and (6.3.4) it is clear that

$$F_h^2(x) = h^2 f^2(x, nh) + 2hxf(x, nh) + x^2 \leq h^2 \bar{f}^2 + 2h\rho + x^2,$$

so that for $n \geq n_0$ and $1 \leq m \leq \lfloor \tau^n \rfloor$ we have

$$\theta_h(m) \leq \beta^{-1}\tau^{n+1} \log n + \frac{2\beta}{\tau^n} \sum_{j=0}^{m-1} \sigma^2 h\big(h^2 \bar{f}^2 + 2h\rho\big) + \frac{2\beta}{\tau^n} \sum_{j=0}^{m-1} \sigma^2 h X_h^2(j). \tag{6.7.14}$$

Define

$$T_h(m) := X_h^2(0) + \sum_{j=0}^{m-1} \lambda_h(X_h(j)) + 2\beta\tau^{-n} \sum_{j=0}^{m-1} \sigma^2 h\big(h^2 \bar{f}^2 + 2h\rho\big) + M_h(m),$$

so that (6.7.12) gives, for $1 \leq m \leq \lfloor \tau^n \rfloor$ and $n \geq n_0$,

$$X_h^2(m) \leq T_h(m) + \beta^{-1}\tau^{n+1} \log n + 2\beta\tau^{-n} \sum_{j=0}^{m-1} \sigma^2 h X_h^2(j). \tag{6.7.15}$$

Note that by (6.2.9) and (6.3.4) and the fact that $m \leq \lfloor \tau^n \rfloor \leq \tau^n$ we have

$$T_h(m) \leq X_h^2(0) + (2h\rho + h^2\bar{f}^2 + \sigma^2 h)m + 2\beta\sigma^2 h\big(h^2\bar{f}^2 + 2h\rho\big) + M_h(m),$$

and so by (6.7.13),

$$\limsup_{m\to\infty} \frac{T_h(m)}{m} \leq 2h\rho + h^2\bar{f}^2 + \sigma^2 h := \bar{\lambda}_h, \quad \text{a.s. on } \Omega_0 \cap \Omega_1,$$

where $\mathbb{P}[\Omega_0 \cap \Omega_1] = 1$. Thus, for every $\varepsilon \in (0,1)$ there exists a number $m_0(\varepsilon) \in \mathbb{N}$ such that for $m_0(\varepsilon) \leq m \leq \lfloor \tau^n \rfloor$ we have $T_h(m) \leq \bar{\lambda}_h(1+\varepsilon)m \leq \bar{\lambda}_h(1+\varepsilon)\tau^n$. Moreover, on the finite set $m \in \{1, \ldots, m_0(\varepsilon) - 1\}$ there exists a number $T_1(\varepsilon, h) > 0$ such that $T_h(m) \leq T_1(\varepsilon, h)$. Combining both of these estimates gives,

$$T_h(m) \leq T_1(\varepsilon, h) + \bar{\lambda}_h(1+\varepsilon)\tau^n := T^*_{h,\varepsilon}(n), \quad 1 \leq m \leq \lfloor \tau^n \rfloor.$$

So now, returning to (6.7.15), for $n \geq n_0$ and $1 \leq m \leq \lfloor \tau^n \rfloor$ we have

$$X_h^2(m) \leq T^*_{h,\varepsilon}(n) + \beta^{-1}\tau^{n+1}\log n + 2\beta\tau^{-n}\sum_{j=0}^{m-1}\sigma^2 h X_h^2(j).$$

Following the notation of Lemma 6.7.1 we set $y(m) := X_h^2(m)$, $a_h(n) := T^*_{h,\varepsilon}(n) + \beta^{-1}\tau^{n+1}\log n$, $c(n) := 2\beta\tau^{-n}$ and $b_h := \sigma^2 h$. Therefore, for $n \geq n_0$,

$$y(m) \leq a_h(n) + c(n)\sum_{j=0}^{m-1} b_h y(j), \quad 1 \leq m \leq \lfloor \tau^n \rfloor,$$

and so we can apply Lemma 6.7.1 to conclude that for $n \geq n_0$,

$$y(m) \leq a_h(n)\prod_{j=0}^{m-1}\big(1 + c(n)b_h\big), \quad 1 \leq m \leq \lfloor \tau^n \rfloor.$$

Then using the fact that $1 + x \leq e^x$ for $x \geq 0$ we have for $1 \leq m \leq \lfloor \tau^n \rfloor \leq \tau^n$,

$$X_h^2(m) \leq a_h(n)\exp\Big[2\beta\tau^{-n}\sum_{j=0}^{m-1}\sigma^2 h\Big] \leq a_h(n)\exp\big[2\beta\sigma^2 h\big], \quad n \geq n_0.$$

Now let $1 \leq \lfloor \tau^{n-1} \rfloor \leq m \leq \lfloor \tau^n \rfloor$ where $n \geq n_0$. Then

$$\begin{aligned}\frac{X_h^2(m)}{2mh\log\log m} &\leq \frac{a_h(n)}{\beta^{-1}\tau^{n+1}\log n}\frac{1}{2\beta h}\frac{\tau^{n+1}}{m}\frac{\log n}{\log\log m}\exp\big[2\beta\sigma^2 h\big] \\ &\leq \frac{a_h(n)}{\beta^{-1}\tau^{n+1}\log n}\frac{1}{2\beta h}\frac{\tau^{n+1}}{\lfloor \tau^{n-1} \rfloor}\frac{\log n}{\log\log\lfloor \tau^{n-1} \rfloor}\exp\big[2\beta\sigma^2 h\big].\end{aligned}$$

Notice that

$$\lim_{n\to\infty}\frac{a_h(n)}{\beta^{-1}\tau^{n+1}\log n}=1,\quad \lim_{n\to\infty}\frac{\tau^{n+1}}{\lfloor\tau^{n-1}\rfloor}=\tau^2,\quad \text{and} \lim_{n\to\infty}\frac{\log\log\lfloor\tau^{n-1}\rfloor}{\log n}=1.$$

Therefore, since $n\to\infty$ as $m\to\infty$ we have

$$\limsup_{m\to\infty}\frac{X_h^2(m)}{2mh\log\log m}\le\frac{\tau^2}{2\beta h}e^{2\beta\sigma^2 h},$$

where this estimate holds for all outcomes in an event of probability one, say Ω_τ. By considering $\Omega^* := \cap_{p\in\mathbb{N}}\Omega_{1+\frac{1}{p}}$, i.e. by letting $\tau\to 1$, we get

$$\limsup_{m\to\infty}\frac{X_h^2(m)}{2mh\log\log m}\le\frac{1}{2\beta h}e^{2\beta\sigma^2 h},\quad \text{a.s. on } \Omega^*,$$

where Ω^* is an almost sure event. Finally, choose $\beta = 1/(2h\sigma^2)$ to get

$$\limsup_{m\to\infty}\frac{X_h^2(m)}{2mh\log\log mh}=\limsup_{m\to\infty}\frac{X_h^2(m)}{2mh\log\log m}\le\sigma^2 e^1,\quad \text{a.s.}$$

Taking square roots on both sides gives the desired result. □

Proof of Theorem 6.3.4. First fix the step size $h<\min(\gamma/\Gamma^2,(2\gamma)^{-1})$ where the constants Γ and γ are defined in Assumption 6.2.1 and Assumption 6.2.2 respectively. By the same argument as that used in the proof of Theorem 6.3.1, this step size restriction also implies that $h<\min(\gamma_1/\Gamma_1^2,(2\gamma_1)^{-1})$ which allows us to define $\alpha_h := \left(1-2h\gamma_1+2h^2\Gamma_1^2\right)^{-1}$ where $\alpha_h>1$. Squaring (6.3.6) and multiplying across by α_h^{n+1} gives

$$\begin{aligned}\alpha_h^{n+1}X_h^2(n+1) &= \alpha_h^{n+1}X_h^2(n)+2h\alpha_h^{n+1}X_h(n)f(X_h(n),nh)\\ &\quad + h^2\alpha_h^{n+1}f^2(X_h(n),nh)+2\sqrt{h}\alpha_h^{n+1}X_h(n)g(X_h(n),nh)\xi(n+1)\\ &\quad + 2h\sqrt{h}\alpha_h^{n+1}f(X_h(n),nh)g(X_h(n),nh)\xi(n+1).\\ &\quad + g^2(X_h(n),nh)h\alpha_h^{n+1}\xi^2(n+1)\end{aligned}$$

Note that by (6.2.4) and the fact that $2xy\le x^2+y^2$ for any $x,y\in\mathbb{R}$ we have

$$f^2(x,nh)\le(\Gamma_0+\Gamma_1|x|)^2=\Gamma_0^2+2\Gamma_0\Gamma_1|x|+\Gamma_1^2x^2\le 2\Gamma_0^2+2\Gamma_1^2x^2.$$

Using this and conditions (6.2.6) and (6.2.14) we get

$$\alpha_h^{n+1}X_h^2(n+1) \le \alpha_h^{n+1}X_h^2(n) - \alpha_h^n X_h^2(n) + \alpha_h^n X_h^2(n) + 2h\alpha_h^{n+1}\big[\rho_1 - \gamma_1 X_h^2(n)\big]$$
$$+ 2h^2\alpha_h^{n+1}\Gamma_0^2 + 2h^2\alpha_h^{n+1}\Gamma_1^2 X_h^2(n) + 2\sqrt{h}\alpha_h^{n+1}X_h(n)g(X_h(n),nh)\xi(n+1)$$
$$+ 2h\sqrt{h}\alpha_h^{n+1}f(X_h(n),nh)g(X_h(n),nh)\xi(n+1) + K_1^2 h\alpha_h^{n+1}\xi^2(n+1).$$

By the definition of α_h we have

$$\alpha_h^{n+1}X_h^2(n) - \alpha_h^n X_h^2(n) - 2h\alpha_h^{n+1}\gamma_1 X_h^2(n) + 2h^2\alpha_h^{n+1}\Gamma_1^2 X_h^2(n) = 0,$$

so that our equation above becomes

$$\alpha_h^{n+1}X_h^2(n+1) \le \alpha_h^n X_h^2(n) + 2h\alpha_h^{n+1}\rho_1 + 2h^2\alpha_h^{n+1}\Gamma_0^2 + \alpha_h^{n+1}K_1^2 h$$
$$+ 2\sqrt{h}\alpha_h^{n+1}F_h(X_h(n))\xi(n+1) + K_1^2 h\alpha_h^{n+1}[\xi^2(n+1) - 1] \qquad (6.7.16)$$

where $F_h(x) := xg(x,nh) + hf(x,nh)g(x,nh)$. Define the martingale differences

$$\Delta M_h(n+1) = K_1^2 h\alpha_h^{n+1}\big[\xi^2(n+1) - 1\big], \quad n \ge 0,$$
$$\Delta\theta_h(n+1) = 2\sqrt{h}\alpha_h^{n+1}F_h(X_h(n))\xi(n+1), \quad n \ge 0,$$

so that $M_h(n+1) := \sum_{j=0}^n \Delta M_h(j+1)$ and $\theta_h(n+1) := \sum_{j=0}^n \Delta\theta_h(j+1)$ are martingales. We also define $\lambda_h := 2\rho_1 + K_1^2 + 2\Gamma_0^2 h$ so that (6.7.16) reduces to

$$\alpha_h^{n+1}X_h^2(n+1) - \alpha_h^n X_h^2(n) \le h\lambda_h\alpha_h^{n+1} + \Delta M_h(n+1) + \Delta\theta_h(n+1).$$

Summing on both sides yields for $n \ge 0$,

$$\alpha_h^{n+1}X_h^2(n+1) - X_h^2(0) \le h\lambda_h \sum_{j=0}^n \alpha_h^{j+1} + M_h(n+1) + \theta_h(n+1).$$

We now apply the exponential martingale inequality to the martingale $\theta_h(n+1)$ which has quadratic variation given by

$$\langle\theta_h\rangle(n) = \sum_{j=0}^{n-1} 4h\alpha_h^{2j+2}F_h^2(X_h(j)), \quad n \ge 1.$$

Applying Theorem 5.2.4, where $\beta > 0$, $\delta > 0$ and $\tau > 1$ are arbitrary constants, we have for all $n \in \mathbb{N}$,

$$\mathbb{P}\Big[\max_{1\le m\le \lfloor n\delta\rfloor}\Big\{\theta_h(m) - 2\beta\alpha_h^{-n\delta}\sum_{j=0}^{m-1} h\alpha_h^{2j+2}F_h^2(X_h(j))\Big\} \ge \beta^{-1}\alpha_h^{n\delta}\tau\log n\Big] \le \frac{1}{n^\tau}.$$

The Borel–Cantelli lemma then yields that for all $\omega \in \Omega_0$, where $\mathbb{P}[\Omega_0] = 1$, there is a random integer $n_0 = n_0(\omega, h)$ sufficiently large such that for $n \ge n_1 := n_0(\omega, h) \vee \lceil 1/\delta\rceil \vee 2$,

$$\theta_h(m) \le \beta^{-1}\alpha_h^{n\delta}\tau\log n + 2\beta\alpha_h^{-n\delta}\sum_{j=0}^{m-1} h\alpha_h^{2j+2}F_h^2(X_h(j)), \quad 1 \le m \le \lfloor n\delta\rfloor.$$

Using (6.2.4) and (6.2.6) it can be shown that

$$F_h^2(x) = g^2(x, nh)\big(x + hf(x, nh)\big)^2 \le g^2(x, nh)\big[(2h\rho_1 + 2h^2\Gamma_0^2) + x^2\alpha_h^{-1}\big]$$

Recall from Remark 6.2.3 that for all $\varepsilon \in (0,1)$ there exists $X(\varepsilon) > 0$, independent of t, such that $|g(x,t)| \le K(1+\varepsilon)$ for $|x| \ge X(\varepsilon)$ uniformly in t. Moreover, by the continuity of g, for $|x| < X(\varepsilon)$ we have $|g(x,t)| \le c_3(X(\varepsilon)) < \infty$. Thus,

$$\begin{aligned}\sum_{j=0}^{m-1} h\alpha_h^{2j+2}F_h^2(X_h(j)) &= \sum_{j=0}^{m-1} h\alpha_h^{2j+2}F_h^2(X_h(j))1_{\{|X_h(j)|\ge X(\varepsilon)\}} \\ &\quad + \sum_{j=0}^{m-1} h\alpha_h^{2j+2}F_h^2(X_h(j))1_{\{|X_h(j)|< X(\varepsilon)\}} \\ &\le \sum_{j=0}^{m-1} h\alpha_h^{2j+2}K^2(1+\varepsilon)^2\big[(2h\rho_1 + 2h^2\Gamma_0^2) + X_h^2(j)\alpha_h^{-1}\big] \\ &\quad + \sum_{j=0}^{m-1} h\alpha_h^{2j+2}c_3^2(X(\varepsilon))\big[(2h\rho_1 + 2h^2\Gamma_0^2) + X^2(\varepsilon)\alpha_h^{-1}\big].\end{aligned}$$

By splitting the sum in this way, we can isolate the terms where we can bound g using the asymptotic bound given in Assumption 6.2.6, which only comes into effect for large enough values of $X_h(j)$. This results in an addition term which we can define as $M_h^*(m) := \sum_{j=0}^{m-1} h\alpha_h^{2j+2}c_3^2(X(\varepsilon))\big[(2h\rho_1 + 2h^2\Gamma_0^2) + X^2(\varepsilon)\alpha_h^{-1}\big]$. For $n \ge n_1$ and $1 \le m \le \lfloor n\delta\rfloor$ this gives

$$\theta_h(m) \le \beta^{-1}\alpha_h^{n\delta}\tau\log n + 2\beta\alpha_h^{-n\delta}K^2(1+\varepsilon)^2h(2h\rho_1 + 2h^2\Gamma_0^2)\sum_{j=0}^{m-1}\alpha_h^{2j+2}$$
$$+ 2\beta\alpha_h^{-n\delta}M_h^*(m) + 2\beta\alpha_h^{-n\delta}K^2(1+\varepsilon)^2h\alpha_h^{-1}\sum_{j=0}^{m-1}\alpha_h^{2j+2}X_h^2(j).$$

Note that the only major difference between this estimate and the equivalent estimate (6.7.3) in the proof of Theorem 6.3.1 is the extra term $2\beta\alpha_h^{-n\delta}M_h^*(m)$. However, it can be shown that this term is of order $\alpha_h^{n\delta}$ and so it will be donimated by the term of order $\alpha_h^{n\delta}\log n$ as $n\to\infty$ and it will not contribute to the final asymptotic estimate.

The proof now follows similar steps to the proof of Theorem 6.3.1 where σ is replaced by $K(1+\varepsilon)$ and we will ultimately let $\varepsilon\to 0$ to get the desired result. □

Proof of Theorem 6.3.5. By squaring (6.3.6) we can show that for $n\ge 0$,

$$X_h^2(n+1) \le X_h^2(n) + h^2f^2(X_h(n),nh) + 2hf(X_h(n),nh)g(X_h(n),nh)\sqrt{h}\xi(n+1)$$
$$+ 2hX_h(n)f(X_h(n),nh) + 2X_h(n)g(X_h(n),nh)\sqrt{h}\xi(n+1) + K_1^2h\xi^2(n+1),$$

where we have used (6.2.14). To simplify the notation, define $\lambda_h(x) := 2hxf(x,nh) + h^2f^2(x,nh) + K_1^2h$ and $F_h(x) := hf(x,nh)g(x,nh) + xg(x,nh)$ so that we can write our equation as

$$X_h^2(n+1) - X_h^2(n) \le \lambda_h(X_h(n)) + \Delta M_h(n+1) + \Delta\theta_h(n+1), \quad n\ge 0, \tag{6.7.17}$$

where

$$\Delta M_h(n+1) := K_1^2h[\xi^2(n+1)-1], \quad n\ge 0,$$
$$\Delta\theta_h(n+1) := 2\sqrt{h}F_h(X_h(n))\xi(n+1), \quad n\ge 0.$$

Therefore $M_h(n+1) := \sum_{j=0}^n \Delta M_h(j+1)$ and $\theta_h(n+1) := \sum_{j=0}^n \Delta\theta_h(j+1)$ are martingales. Thus, returning to (6.7.17) and summing on both sides yields

$$X_h^2(n+1) - X_h^2(0) \le \sum_{j=0}^n \lambda_h(X_h(j)) + M_h(n+1) + \theta_h(n+1), \quad n\ge 0.$$

It is clear, by the strong law of large numbers, that

$$\lim_{n\to\infty}\frac{M_h(n)}{n} = K_1^2 h \lim_{n\to\infty}\frac{1}{n}\sum_{j=0}^{n-1}[\xi^2(j+1)-1] = K_1^2 h\mathbb{E}[\xi^2(j+1)-1] = 0,$$

where we are using the fact that ξ has a standard normal distribution. We now apply the exponential martingale inequality to the martingale $\theta_h(n+1)$ which has quadratic variation given by

$$\langle\theta_h\rangle(n+1) = \sum_{j=0}^{n} 4hF_h^2(X_h(j)), \quad n\ge 0.$$

Applying Theorem 5.2.4, where $\beta>0$ and $\tau>1$ are arbitrary constants, we have for all $n\in\mathbb{N}$,

$$\mathbb{P}\Big[\max_{1\le m\le\lfloor\tau^n\rfloor}\Big\{\theta_h(m)-\frac{\beta\tau^{-n}}{2}\sum_{j=0}^{m-1}4hF_h^2(X_h(j))\Big\}\ge \beta^{-1}\tau^{n+1}\log n\Big]\le\frac{1}{n^\tau}.$$

The Borel–Cantelli lemma then yields that for all $\omega\in\Omega_0$, where $\mathbb{P}[\Omega_0]=1$, there is a random integer $n_0=n_0(\omega,h)$ sufficiently large such that $\lfloor\tau^{n_0-1}\rfloor>e^1$ and for $n\ge n_0$ we have

$$\theta_h(m)\le\beta^{-1}\tau^{n+1}\log n+2\beta\tau^{-n}\sum_{j=0}^{m-1}hF_h^2(X_h(j)),\quad 1\le m\le\lfloor\tau^n\rfloor.$$

Using (6.2.9) and (6.3.4) it is clear that

$$F_h^2(x)=g^2(x,nh)\big(hf(x,nh)+x\big)^2\le g^2(x,nh)\big[h^2\bar{f}^2+2h\rho+x^2\big].$$

Recall from Remark 6.2.3 that for all $\varepsilon\in(0,1)$ there exists $X(\varepsilon)>0$, independent of t, such that $|g(x,t)|\le K(1+\varepsilon)$ for $|x|\ge X(\varepsilon)$ uniformly in t. Moreover, by the continuity of g, for $|x|<X(\varepsilon)$ we have $|g(x,t)|\le c_3(X(\varepsilon))<\infty$. Thus,

$$\begin{aligned}
&2\beta\tau^{-n}\sum_{j=0}^{m-1}hF_h^2(X_h(j))\\
&=\frac{2\beta}{\tau^n}\sum_{j=0}^{m-1}hF_h^2(X_h(j))1_{\{|X_h(j)|\ge X(\varepsilon)\}}+\frac{2\beta}{\tau^n}\sum_{j=0}^{m-1}hF_h^2(X_h(j))1_{\{|X_h(j)|<X(\varepsilon)\}}\\
&\le\frac{2\beta}{\tau^n}\sum_{j=0}^{m-1}hK^2(1+\varepsilon)^2\big[h^2\bar{f}^2+2h\rho+X_h^2(j)\big]+\frac{2\beta}{\tau^n}\sum_{j=0}^{m-1}hc_3^2(X(\varepsilon))\big[h^2\bar{f}^2+2h\rho+X^2(\varepsilon)\big].
\end{aligned}$$

Define $M_h^*(m) := 2\beta\tau^{-n}\sum_{j=0}^{m-1} hc_3^2(X(\varepsilon))\big[h^2\bar{f}^2+2h\rho+X^2(\varepsilon)\big]$. Therefore, for $n \geq n_0$ and $1 \leq m \leq \lfloor \tau^n \rfloor$ we have

$$\begin{aligned}\theta_h(m) \leq \beta^{-1}\tau^{n+1}\log n + 2\beta\tau^{-n}\sum_{j=0}^{m-1} hK^2(1+\varepsilon)^2\big(h^2\bar{f}^2+2h\rho\big)\\ + M_h^*(m) + 2\beta\tau^{-n}\sum_{j=0}^{m-1} hK^2(1+\varepsilon)^2X_h^2(j).\end{aligned}$$

Note that the only major difference between this estimate and the equivalent estimate (6.7.14) in the proof of Theorem 6.3.3 is the extra term $M_h^*(m)$. However, it can be shown that this term will be donimated by the term of order $\tau^{n+1}\log n$ as $n \to \infty$ and it will not contribute to the final asymptotic estimate.

The proof now follows similar steps to the proof of Theorem 6.3.3 where σ is replaced by $K(1+\varepsilon)$ and we will ultimately let $\varepsilon \to 0$ to get the desired result. □

6.8 Proofs of Results from Section 6.4

Before proving our main results we first show that even without the one-sided Lipschitz condition (6.4.2) on f, we can prove that there exists a (not necessarily unique) solution to (6.4.1) in both the O–U case and the Iterated Logarithm case. Notice that there will be a solution of (6.4.1) provided that, for every $x \in \mathbb{R}$ and $t \geq 0$, there is a solution $y \in \mathbb{R}$ of

$$y = x + hf(y,t). \tag{6.8.1}$$

First we consider the O–U case where our condition on f is given by (6.2.6).

Lemma 6.8.1. *Let $h > 0$ and suppose that f obeys* (6.2.6). *Then for every $x \in \mathbb{R}$ and $t \geq 0$ there is at least one solution y of* (6.8.1).

Proof. Fix $t \geq 0$ and define for each $x \in \mathbb{R}$ the function $G_x : \mathbb{R} \times [0,\infty) \to \mathbb{R}$

$$G_x(y,t) = y - x - hf(y,t), \quad y \in \mathbb{R},\ t \geq 0.$$

Note that the continuity of f ensures that G_x is continuous for each fixed t. Therefore, if we can find values y_+, y_- such that $G_x(y_+) > 0$ and $G_x(y_-) < 0$ then there must exist a solution $y \in (y_-, y_+)$. Define $y_* := (x^2 + 2h\rho_1)/(1 + 2h\gamma_1)$. Then for $|y| \geq \sqrt{y_*}$

$$yG_x(y,t) = y^2 - xy - hyf(y,t) \geq y^2(\frac{1}{2} + h\gamma_1) - \frac{1}{2}x^2 - h\rho_1 \geq 0$$

where we have used (6.2.6) and the fact that $xy \leq \frac{1}{2}x^2 + \frac{1}{2}y^2$. This means that y and $G_x(y,t)$ have the same sign when $y \geq \sqrt{y_*}$ and $y \leq -\sqrt{y_*}$. Thus, setting $y_+ = \sqrt{y_*} > 0$ and $y_- = -\sqrt{y_*} < 0$ means that $G_x(y_+, t) > 0$ and $G_x(y_-, t) < 0$ for every fixed t and so there exists a solution $y \in (y_-, y_+)$. □

We now consider the Iterated Logarithm case where our condition on f is (6.2.9).

Lemma 6.8.2. *Let $h > 0$ and suppose that f obeys* (6.2.9). *Then for every $x \in \mathbb{R}$ and $t \geq 0$ there is at least one solution y of* (6.8.1).

Proof. Following on from the proof of Lemma 6.8.1, define $y_* := x^2 + 2h\rho$. Then for $|y| \geq \sqrt{y_*}$

$$yG_x(y,t) = y^2 - xy - hyf(y,t) \geq \frac{1}{2}y^2 - \frac{1}{2}x^2 - h\rho \geq 0$$

where we have used (6.2.9) and the fact that $xy \leq \frac{1}{2}x^2 + \frac{1}{2}y^2$. This means that y and $G_x(y,t)$ have the same sign when $y \geq \sqrt{y_*}$ and $y \leq -\sqrt{y_*}$. Thus, setting $y_+ = \sqrt{y_*} > 0$ and $y_- = -\sqrt{y_*} < 0$ means that $G_x(y_+, t) > 0$ and $G_x(y_-, t) < 0$ for every fixed t and so there exists a solution $y \in (y_-, y_+)$. □

Proof of Theorem 6.4.2. Consider the first step of the discretisation given by (6.4.1). Multiplying across by $X_h^*(n)$ on both sides gives

$$\left(X_h^*(n)\right)^2 = X_h(n)X_h^*(n) + hX_h^*(n)f(X_h^*(n), nh), \quad n \geq 0.$$

Then using (6.2.6) and the fact that $xy \leq \frac{1}{2}x^2 + \frac{1}{2}y^2$ for $x, y \in \mathbb{R}$, we can show that

$$\left(X_h^*(n)\right)^2 \leq \frac{X_h^2(n)}{1 + 2h\gamma_1} + \frac{2h\rho_1}{1 + 2h\gamma_1}, \quad n \geq 0. \tag{6.8.2}$$

Now consider the second step of the discretisation given by (6.4.1). Squaring it gives

$$\begin{aligned} X_h^2(n+1) &= \big(X_h^*(n)\big)^2 + 2\sigma\sqrt{h}X_h^*(n)\xi(n+1) + \sigma^2 h\xi^2(n+1) \\ &\le \frac{X_h^2(n)+2h\rho_1}{1+2h\gamma_1} + 2\sigma\sqrt{h}X_h^*(n)\xi(n+1) + \sigma^2 h[\xi^2(n+1)-1] + \sigma^2 h \end{aligned}$$

Define $\alpha_h := 1+2h\gamma_1$, where γ_1 arises from condition (6.2.6). Note that $\alpha_h > 1$ for any $h > 0$ since $\gamma_1 > 0$. Multiplying both sides of the above equation by α_h^{n+1} gives

$$\begin{aligned} \alpha_h^{n+1}X_h^2(n+1) \le \frac{\alpha_h^{n+1}X_h^2(n)}{1+2h\gamma_1} + \alpha_h^n X_h^2(n) - \alpha_h^n X_h^2(n) + \frac{\alpha_h^{n+1}2h\rho_1}{1+2h\gamma_1} + \alpha_h^{n+1}\sigma^2 h \\ + \alpha_h^{n+1}2\sigma\sqrt{h}X_h^*(n)\xi(n+1) + \alpha_h^{n+1}\sigma^2 h[\xi^2(n+1)-1]. \qquad (6.8.3) \end{aligned}$$

However, by the construction of α_h we have $\alpha_h^{n+1}X_h^2(n)/(1+2h\gamma_1) - \alpha_h^n X_h^2(n) = 0$. Defining $\lambda_h := \sigma^2 + 2\rho_1/(1+2h\gamma_1)$ and summing on both sides of (6.8.3) then yields

$$\alpha_h^{n+1}X_h^2(n+1) - X_h^2(0) \le h\lambda_h \sum_{j=0}^{n} \alpha_h^{j+1} + \theta_h(n+1) + M_h(n+1), \quad n \ge 0,$$

where we define the martingales

$$\begin{aligned} \theta_h(n+1) &:= \sum_{j=0}^{n} 2\sigma\sqrt{h}\alpha_h^{j+1}X_h^*(j)\,\xi(j+1), \quad n \ge 0, \\ M_h(n+1) &:= \sum_{j=0}^{n} \sigma^2 h\alpha_h^{j+1}[\xi^2(j+1)-1], \quad n \ge 0. \end{aligned}$$

We now apply the exponential martingale inequality to the martingale $\theta_h(n+1)$ which has quadratic variation given by

$$\langle\theta_h\rangle(n+1) = \sum_{j=0}^{n} 4\sigma^2 h\alpha_h^{2j+2}\big(X_h^*(j)\big)^2.$$

Applying Theorem 5.2.4, where $\beta > 0, \delta > 0$ and $\tau > 1$ are arbitrary constants, we have for all $n \in \mathbb{N}$,

$$\mathbb{P}\Big[\max_{1\le m\le\lfloor n\delta\rfloor}\Big\{\theta_h(m) - \frac{\beta}{2\alpha_h^{n\delta}}\sum_{j=0}^{m-1} 4\sigma^2 h\alpha_h^{2j+2}\big(X_h^*(j)\big)^2\Big\} \ge \frac{\alpha_h^{n\delta}\tau\log n}{\beta}\Big] \le \frac{1}{n^\tau}.$$

The Borel–Cantelli lemma then yields that for all $\omega \in \Omega_0$, where $\mathbb{P}[\Omega_0] = 1$, there is a random integer $n_0 = n_0(\omega, h)$ sufficiently large such that for $n \geq n_1 := n_0 \vee \lceil 1/\delta \rceil \vee 2$, and for $1 \leq m \leq \lfloor n\delta \rfloor$,

$$\begin{aligned}\theta_h(m) &\leq \frac{\alpha_h^{n\delta}\tau \log n}{\beta} + \frac{\beta}{2\alpha_h^{n\delta}} \sum_{j=0}^{m-1} 4\sigma^2 h \alpha_h^{2j+2} \big(X_h^*(j)\big)^2 \\ &\leq \frac{\alpha_h^{n\delta}\tau \log n}{\beta} + \frac{4\beta\sigma^2 h^2 \rho_1}{\alpha_h^{n\delta}} \sum_{j=0}^{m-1} \alpha_h^{2j+1} + \frac{2\beta\sigma^2 h}{\alpha_h^{n\delta}} \sum_{j=0}^{m-1} \alpha_h^{2j+1} X_h^2(j), \end{aligned} \tag{6.8.4}$$

where we have used (6.8.2) in the last step. One can now immediately see the similarites between this estimate and the analogous estimate (6.7.3) arising in the Euler–Maruyama case. Here, our estimate does not depend on the constants Γ_0, Γ_1 of the linear growth bound (6.2.4). Our proof now follows exactly the same steps as in the proof of Theorem 6.3.1 up until

$$\limsup_{m\to\infty} \frac{X_h^2(m)}{\log m} \leq (2h\sigma^2 + \beta^{-1}\tau)\alpha_h^{\delta+1} \exp\big[2\beta\alpha_h^{-n\delta}\sigma^2 h \sum_{j=0}^{m-1} \alpha_h^{j+1}\big], \tag{6.8.5}$$

where, by the definition of α_h and the fact that $m \leq \lfloor n\delta \rfloor \leq n\delta$,

$$2\beta\alpha_h^{-n\delta}\sigma^2 h \sum_{j=0}^{m-1} \alpha_h^{j+1} \leq \frac{2\beta\sigma^2 h \alpha_h^{-m}\alpha_h \alpha_h^m}{\alpha_h - 1} = \frac{\beta\sigma^2\alpha_h}{\gamma_1}.$$

The estimate (6.8.5) holds for all outcomes in an event of probability one, say $\Omega_{\tau,\delta}$. By considering $\Omega^* := \cap_{p\in\mathbb{N}} \Omega_{1+\frac{1}{p},\frac{1}{p}}$, i.e. by letting $\tau \to 1$ and $\delta \to 0$, we get

$$\limsup_{m\to\infty} \frac{X_h^2(m)}{\log m} \leq (2h\sigma^2 + \beta^{-1})\alpha_h \exp\Big[\frac{\beta\sigma^2\alpha_h}{\gamma_1}\Big], \quad \text{a.s. on } \Omega^*$$

where Ω^* is an almost sure event. Finally, choose $\beta = \gamma_1/\sigma^2\alpha_h$ to get

$$\limsup_{m\to\infty} \frac{X_h^2(m)}{\log mh} = \limsup_{m\to\infty} \frac{X_h^2(m)}{\log m} \leq \frac{\sigma^2}{\gamma_1}[1 + 6h\gamma_1 + 8h^2\gamma_1^2]e, \quad \text{a.s.}$$

Once again, we recall from Remark 6.2.2 that $\gamma_1 := \gamma(1-\eta)$ where $\eta \in (0,1)$ is arbitrary. Letting $\eta \to 0$ and taking square roots on both sides gives the desired result. □

Proof of Theorem 6.4.3. Consider the first step of the discretisation given by (6.4.1). Multiplying across by $X_h^*(n)$ on both sides gives

$$\big(X_h^*(n)\big)^2 = X_h(n)X_h^*(n) + hX_h^*(n)f(X_h^*(n), nh), \quad n \geq 0.$$

Then using (6.2.9) and the fact that $xy \leq \frac{1}{2}x^2 + \frac{1}{2}y^2$ for $x, y \in \mathbb{R}$, we can show that

$$\big(X_h^*(n)\big)^2 \leq X_h^2(n) + 2h\rho, \quad n \geq 0. \tag{6.8.6}$$

Now consider the second step of the discretisation given by (6.4.1). Squaring it gives

$$\begin{aligned} X_h^2(n+1) &= \big(X_h^*(n)\big)^2 + 2\sigma\sqrt{h}X_h^*(n)\xi(n+1) + \sigma^2 h\xi^2(n+1) \\ &\leq X_h^2(n) + 2h\rho + \sigma^2 h + 2\sigma\sqrt{h}X_h^*(n)\,\xi(n+1) + \sigma^2 h[\xi^2(n+1) - 1] \end{aligned}$$

Defining $\lambda_h := \sigma^2 h + 2h\rho$ and summing on both sides then yields

$$X_h^2(n+1) - X_h^2(0) \leq \sum_{j=0}^{n} \lambda_h + \theta_h(n+1) + M_h(n+1), \quad n \geq 0,$$

where we define the martingales

$$\begin{aligned} \theta_h(n+1) &:= \sum_{j=0}^{n} 2\sigma\sqrt{h}X_h^*(j)\,\xi(j+1), \quad n \geq 0, \\ M_h(n+1) &:= \sum_{j=0}^{n} \sigma^2 h[\xi^2(j+1) - 1], \quad n \geq 0. \end{aligned}$$

We now apply the exponential martingale inequality to the martingale $\theta_h(n+1)$ which has quadratic variation given by

$$\langle\theta_h\rangle(n+1) = \sum_{j=0}^{n} 4\sigma^2 h\big(X_h^*(j)\big)^2.$$

Applying Theorem 5.2.4, where $\beta > 0$ and $\tau > 1$ are arbitrary constants, we have for all $n \in \mathbb{N}$,

$$\mathbb{P}\Big[\max_{1\leq m\leq\lfloor\tau^n\rfloor}\Big\{\theta_h(m) - \frac{\beta\tau^{-n}}{2}\sum_{j=0}^{m-1} 4\sigma^2 h\big(X_h^*(j)\big)^2\Big\} \geq \beta^{-1}\tau^{n+1}\log n\Big] \leq \frac{1}{n^\tau}.$$

The Borel–Cantelli lemma then yields that for all $\omega \in \Omega_0$, where $\mathbb{P}[\Omega_0] = 1$, there is a random integer $n_0 = n_0(\omega, h)$ sufficiently large such that for $n \geq n_0$, and for $1 \leq m \leq \lfloor \tau^n \rfloor$,

$$\begin{aligned}\theta_h(m) &\leq \beta^{-1}\tau^{n+1}\log n + \frac{\beta\tau^{-n}}{2}\sum_{j=0}^{m-1} 4\sigma^2 h\big(X_h^*(j)\big)^2 \\ &\leq \beta^{-1}\tau^{n+1}\log n + 2\beta\tau^{-n}\sum_{j=0}^{m-1} 2\sigma^2 h^2\rho + 2\beta\tau^{-n}\sum_{j=0}^{m-1}\sigma^2 h X_h^2(j), \end{aligned} \tag{6.8.7}$$

where we have used (6.8.6) in the last step. One can now immediately see the similarites between this estimate and the analogous estimate (6.7.14) arising in the Euler–Maruyama case. Here, our estimate does not depend on the constant $\bar{f}$ of condition (6.3.4). Our proof now follows exactly the same steps as in the proof of Theorem 6.3.3 to get

$$\limsup_{m\to\infty} \frac{X_h^2(m)}{2mh\log\log mh} \leq \sigma^2 e, \quad \text{a.s.}$$

Taking square roots on both sides gives the desired result. □

Before we begin the proofs of Theorems 6.4.4 and 6.4.5 we must first prove auxiliary results.

Lemma 6.8.3. *If f is locally Lipschitz continuous as per* (6.2.2) *and obeys either Assumption 6.2.2 or 6.2.5, then there exists $c_5 > 0$ such that $|f(0,t)| < c_5$ for all $t \geq 0$.*

Proof. Assumption 6.2.2 implies that $f(1,t) \leq c_2(1)$ and $-f(-1,t) \leq c_2(1)$ for all $t \geq 0$. On the other hand Assumption 6.2.5 implies that $f(1,t) \leq \rho$ and that $-f(-1,t) \leq \rho$ for all $t \geq 0$. Therefore, in either case, there exists $c_4 > 0$ such that $f(1,t) \leq c_4$ and $f(-1,t) \geq -c_4$ for all $t \geq 0$. By (6.2.2) there exists $M_1 > 0$ such that

$$|f(1,t) - f(0,t)| \leq M_1 \quad \text{and} \quad |f(-1,t) - f(0,t)| \leq M_1 \quad \text{for all } t \geq 0.$$

Therefore $f(0,t) \leq f(1,t) + M_1 \leq M_1 + c_4$ for all $t \geq 0$. Similarly, we have $f(0,t) \geq f(-1,t) - M_1 \geq -c_4 - M_1$ for $t \geq 0$. The result follows with $c_5 := M_1 + c_4 + 1$. □

In Lemmas 6.8.1 and 6.8.2 it was shown that if f obeys either Assumption 6.2.2 or 6.2.5, then for any $x \in \mathbb{R}$ there exists $x^* = x^*(x,t) \in \mathbb{R}$ such that

$$x^* = x + hf(x^*, t). \tag{6.8.8}$$

In the proof of Theorem 6.4.4 we need to know that in an equation of the form (6.8.8), if $|x|$ is of a certain size then we can also bound $|x^*|$. This connection between x and x^* is detailed in the following lemma.

Lemma 6.8.4. *Let f be locally Lipschitz continuous as per* (6.2.2) *and let f obey either Assumption 6.2.2 or 6.2.5. Suppose that c_5 is defined by Lemma 6.8.3 and that $c \geq c_5$. If $|x| > c$ then for any x^* which obeys* (6.8.8) *we have $|x^*| > H_t(c)$ where H_t is an increasing function which obeys $H_t(c) \to \infty$ as $c \to \infty$.*

Proof. First define the function

$$F_{c,t}(y) := y + h \max_{|u| \leq y} |f(u,t)| - c, \quad \text{for every } t \geq 0,$$

and notice that $y \mapsto F_{c,t}(y)$ is continuous and increasing for every $t \geq 0$. Then $F_{c,t}(c) = h \max_{|u|\leq c} |f(u,t)| \geq 0$ for every $t \geq 0$. Moreover, by Lemma 6.8.3 we have $|f(0,t)| < c_5$ for all $t \geq 0$ and so $F_{c,t}(0) = |f(0,t)| - c < c_5 - c \leq 0$ for every $t \geq 0$. Since the function changes sign it follows that for every $c \geq c_5$ there must exist a $H_t(c) \in (0, c]$ such that $F_{c,t}(H_t(c)) = 0$. Now let $|x| > c$ and suppose that $|x^*| \leq H_t(c)$. Then using the fact that $F_{c,t}(H_t(c)) = 0$ we have

$$\begin{aligned} |x| &= |x^* - hf(x^*,t)| \leq |x^*| + h|f(x^*,t)| \\ &\leq H_t(c) + h \max_{|u| \leq H_t(c)} |f(u,t)| - c + c = c, \end{aligned}$$

which forces a contradiction. Therefore, the assumption that $|x^*| \leq H_t(c)$ is incorrect. Thus, $|x| > c$ implies that $|x^*| > H_t(c)$.

We now show that $c \mapsto H_t(c)$ is increasing and that $H_t(c) \to \infty$ as $c \to \infty$. Let $a_2 > a_1 > c_5$. Then for every $t \geq 0$ we have

$$F_{a_2,t}(y) = y + h \max_{|u|\leq y} |f(u,t)| - a_2 < y + h \max_{|u|\leq y} |f(u,t)| - a_1 = F_{a_1,t}(y).$$

Using this fact we have

$$F_{a_1,t}(H_t(a_1)) = 0 = F_{a_2,t}(H_t(a_2)) < F_{a_1,t}(H_t(a_2)),$$

which means that

$$F_{a_1,t}(H_t(a_1)) < F_{a_1,t}(H_t(a_2)).$$

However, since $y \mapsto F_{c,t}(y)$ is increasing we must have $H_t(a_1) < H_t(a_2)$, which in turn means that H_t is increasing. Now suppose that $\lim_{c\to\infty} H_t(c) = L_t < +\infty$. Since $F_{c,t}(H_t(c)) = 0$ we have

$$H_t(c) + h \max_{|u|\leq H_t(c)} |f(u,t)| = c, \quad \text{for every } t \geq 0.$$

However, taking the limit as $c \to \infty$ on both sides yields a contradiction since the limit on the left–hand side is $L_t + h \max_{|u|\leq L_t} |f(u,t)|$ which is finite for every $t \geq 0$ since f is continuous. Thus, our assumption that $\lim_{c\to\infty} H_t(c) = L_t < +\infty$ was incorrect and so we must have $\lim_{c\to\infty} H_t(c) = \infty$. □

We need one final estimate on H_t^{-1}, to hold independently of t.

Lemma 6.8.5. *Suppose that f is locally Lipschitz continuous as per* (6.2.2) *and obeys either Assumption 6.2.2 or 6.2.5, and let c_5 be the number defined in Lemma 6.8.3. Let $t \geq 0$. If H_t is the invertible function defined in Lemma 6.8.4, then for all $d > c_5$, there exists a finite $F_h(d) > 0$ such that*

$$H_t^{-1}(d) \leq F_h(d), \quad t \geq 0.$$

Proof. For any $t \geq 0$ and $c \geq c_5$ we have by definition that $H_t(c)$ is the unique solution of

$$c = H_t(c) + h \max_{|u| \leq H_t(c)} |f(u,t)|, \quad \text{where } H_t(c) \in (0, c].$$

Let $d > c_5$. Then $d > c_5 \geq H_t(c_5)$. Thus $H_t^{-1}(d) > c_5$. Therefore, for each $d > c_5$, there exists a $c > c_5$ such that $c = H_t^{-1}(d)$. Therefore for each $d > c_5$ we have that

$$H_t^{-1}(d) = d + h \max_{|u| \leq d} |f(u,t)|.$$

By Assumption 6.2.2 or 6.2.5 it follows, in the same way as was shown in the proof of Lemma 6.8.3, that there exists a finite $c_6 = c_6(d) > 0$ such that $df(d,t) \leq c_6(d)$ and $-df(-d,t) \leq c_6(d)$. Therefore

$$f(d,t) \leq c_6(d)/d, \quad f(-d,t) \geq -c_6(d)/d.$$

By condition (6.2.2), for all $|u| \leq d$ there exists $M_d > 0$ such that

$$|f(u,t) - f(d,t)| \leq M_d |u - d| \leq 2dM_d, \quad |f(u,t) - f(-d,t)| \leq M_d |u + d| \leq 2dM_d.$$

Therefore for all $|u| \leq d$ we have

$$f(u,t) \leq f(d,t) + 2dM_d \leq c_6(d)/d + 2dM_d,$$

and

$$f(u,t) \geq f(-d,t) - 2dM_d \geq -c_6(d)/d - 2dM_d.$$

Hence $|f(u,t)| \leq c_6(d)/d + 2dM_d$ for all $t \geq 0$ and $|u| \leq d$. Therefore we have that

$$H_t^{-1}(d) = d + h \max_{|u| \leq d} |f(u,t)| \leq d + h[c_6(d)/d + 2dM_d].$$

Defining $F_h(d) = d + h[c_6(d)/d + 2dM_d]$ yields the desired result. □

Proof of Theorem 6.4.4. Consider the first step of the discretisation given by (6.4.5). Multiplying across by $X_h^*(n)$ on both sides gives

$$\left(X_h^*(n)\right)^2 = X_h(n) X_h^*(n) + h X_h^*(n) f(X_h^*(n), nh), \quad n \geq 0.$$

Then using (6.2.6) and the fact that $xy \le \frac{1}{2}x^2 + \frac{1}{2}y^2$ for $x, y \in \mathbb{R}$, we can show that

$$\big(X_h^*(n)\big)^2 \le \frac{X_h^2(n)}{1+2h\gamma_1} + \frac{2h\rho_1}{1+2h\gamma_1}, \quad n \ge 0. \tag{6.8.9}$$

Now consider the second step of the discretisation given by (6.4.5). Squaring it gives

$$\begin{aligned} X_h^2(n+1) &= \big(X_h^*(n)\big)^2 + 2\sqrt{h}X_h^*(n)g(X_h^*(n),nh)\xi(n+1) + g^2(X_h^*(n),nh)h\xi^2(n+1) \\ &\le \frac{X_h^2(n)+2h\rho_1}{1+2h\gamma_1} + 2\sqrt{h}X_h^*(n)g(X_h^*(n),nh)\xi(n+1) + K_1^2h\xi^2(n+1), \end{aligned}$$

where we have used (6.2.14). Define $\alpha_h := 1+2h\gamma_1$, where γ_1 arises from condition (6.2.6). Note that $\alpha_h > 1$ for any $h > 0$ since $\gamma_1 > 0$. Multiplying both sides of the above equation by α_h^{n+1} gives

$$\begin{aligned} \alpha_h^{n+1}X_h^2(n+1) \le \frac{\alpha_h^{n+1}X_h^2(n)}{1+2h\gamma_1} + \alpha_h^nX_h^2(n) - \alpha_h^nX_h^2(n) + \frac{\alpha_h^{n+1}2h\rho_1}{1+2h\gamma_1} + \alpha_h^{n+1}K_1^2h \\ + \alpha_h^{n+1}2\sqrt{h}X_h^*(n)g(X_h^*(n),nh)\xi(n+1) + \alpha_h^{n+1}K_1^2h[\xi^2(n+1)-1]. \end{aligned}$$

However, by the construction of α_h we have $\alpha_h^{n+1}X_h^2(n)/(1+2h\gamma_1) - \alpha_h^nX_h^2(n) = 0$. Defining $\lambda_h := K_1^2 + 2\rho_1/(1+2h\gamma_1)$ and summing on both sides then yields

$$\alpha_h^{n+1}X_h^2(n+1) - X_h^2(0) \le h\lambda_h\sum_{j=0}^{n}\alpha_h^{j+1} + \theta_h(n+1) + M_h(n+1), \quad n \ge 0,$$

where we define the martingales

$$\theta_h(n+1) := \sum_{j=0}^{n} 2\sqrt{h}\alpha_h^{j+1}X_h^*(j)g(X_h^*(j),jh)\,\xi(j+1), \quad n \ge 0,$$

$$M_h(n+1) := \sum_{j=0}^{n} K_1^2h\alpha_h^{j+1}[\xi^2(j+1)-1], \quad n \ge 0.$$

We now apply the exponential martingale inequality to the martingale $\theta_h(n+1)$ which has quadratic variation given by

$$\langle\theta_h\rangle(n+1) = \sum_{j=0}^{n} 4h\alpha_h^{2j+2}\big(X_h^*(j)\big)^2g^2(X_h^*(j),jh).$$

Applying Theorem 5.2.4, where $\beta > 0, \delta > 0$ and $\tau > 1$ are arbitrary constants, we have for all $n \in \mathbb{N}$,

$$\mathbb{P}\Big[\max_{1\le m\le \lfloor n\delta\rfloor}\Big\{\theta_h(m) - \frac{2h\beta}{\alpha_h^{n\delta}}\sum_{j=0}^{m-1}\alpha_h^{2j+2}\big(X_h^*(j)\big)^2 g^2(X_h^*(j),jh)\Big\} \ge \frac{\tau\log n}{\beta\alpha_h^{-n\delta}}\Big] \le \frac{1}{n^\tau}.$$

The Borel–Cantelli lemma then yields that for all $\omega \in \Omega_0$, where $\mathbb{P}[\Omega_0] = 1$, there is a random integer $n_0 = n_0(\omega, h)$ sufficiently large such that for $n \ge n_1 := n_0 \vee \lceil 1/\delta\rceil \vee 2$, and for $1 \le m \le \lfloor n\delta\rfloor$,

$$\begin{aligned}\theta_h(m) &\le \frac{\tau\log n}{\beta\alpha_h^{-n\delta}} + \frac{2h\beta}{\alpha_h^{n\delta}}\sum_{j=0}^{m-1}\alpha_h^{2j+2}\big(X_h^*(j)\big)^2 g^2(X_h^*(j),jh)\\ &\le \frac{\tau\log n}{\beta\alpha_h^{-n\delta}} + \frac{4\beta h^2K_1^2\rho_1}{\alpha_h^{n\delta}}\sum_{j=0}^{m-1}\alpha_h^{2j+1} + \frac{2\beta h}{\alpha_h^{n\delta}}\sum_{j=0}^{m-1}\alpha_h^{2j+1}g^2(X_h^*(j),jh)X_h^2(j),\end{aligned}$$

where we have used (6.8.9) and the fact that $g^2(x^*, jh) \le K_1^2$.

Recall from Assumption 6.2.6 that for every $\varepsilon \in (0,1)$ there is an $X(\varepsilon) > 0$ such that $g^2(x^*,t) \le K^2(1+\varepsilon)^2$ for all $|x^*| \ge X(\varepsilon)$ and for all $t \ge 0$. Now let $X_1(\varepsilon) := \max(X(\varepsilon), c_5)+\varepsilon$ for every $t \ge 0$, where c_5 is derived in Lemma 6.8.3. Then $X_1(\varepsilon) > X(\varepsilon)$ and $X_1(\varepsilon) > c_5$ for every $t \ge 0$. Now consider

$$\begin{aligned}\sum_{j=0}^{m-1}\alpha_h^{2j+1}g^2(X_h^*(j),jh)X_h^2(j) &= \sum_{j=0}^{m-1}\alpha_h^{2j+1}g^2(X_h^*(j),jh)X_h^2(j)1_{\{|X_h(j)|>H_{jh}^{-1}(X_1(\varepsilon))\}}\\ &\quad+\sum_{j=0}^{m-1}\alpha_h^{2j+1}g^2(X_h^*(j),jh)X_h^2(j)1_{\{|X_h(j)|\le H_{jh}^{-1}(X_1(\varepsilon))\}}\end{aligned}$$

where H_{jh} is defined in Lemma 6.8.4 and is shown to be increasing, resulting in H_{jh}^{-1} being well–defined. By splitting the sum in this way, we aim to isolate the terms where we can bound g using the asymptotic bound given in Assumption 6.2.6, which only comes into effect for large enough values of $X_h^*(j)$. Doing so is more complicated in the split–step case because the terms in the summations above involve a combination of $X_h^*(j)$ and $X_h(j)$ terms. Lemmas 6.8.3, 6.8.4 and 6.8.5 effectively allow us to estimate both simultaneously.

Recall that $H_{jh}(X_1(\varepsilon)) \le X_1(\varepsilon)$ since $X_1(\varepsilon) > c_5$. Therefore by Lemma 6.8.4, in the case when $|X_h(j)| > H_{jh}^{-1}(X_1(\varepsilon)) \ge X_1(\varepsilon) > c_5$, we must have $|X_h^*(j)| > H_{jh}(H_{jh}^{-1}(X_1(\varepsilon))) =$

$X_1(\varepsilon) > X(\varepsilon)$ and so we can bound g using Assumption 6.2.6. In the case when $|X_h(j)| \le H_{jh}^{-1}(X_1(\varepsilon))$, we can simply bound g using the global bound (6.2.14) and we can bound $X_h^2(j)$ using the fact that $H_{jh}^{-1}(X_1(\varepsilon)) \le F_h(X_1(\varepsilon)) < +\infty$ by Lemma 6.8.5. Therefore we have that

$$\sum_{j=0}^{m-1} \alpha_h^{2j+1} g^2(X_h^*(j), jh) X_h^2(j) \le K^2(1+\varepsilon)^2 \sum_{j=0}^{m-1} \alpha_h^{2j+1} X_h^2(j) + M_h^*(m),$$

where we define $M_h^*(m) := K_1^2 F_h^2(X_1(\varepsilon)) \sum_{j=0}^{m-1} \alpha_h^{2j+1}$. Thus, for $n \ge n_1$ and for $1 \le m \le \lfloor n\delta \rfloor$,

$$\theta_h(m) \le \beta^{-1}\tau\alpha_h^{n\delta}\log n + 4\beta h^2 K_1^2 \rho_1 \alpha_h^{-n\delta} \sum_{j=0}^{m-1} \alpha_h^{2j+1} + 2\beta h \alpha_h^{-n\delta} M_h^*(m)$$
$$+ 2\beta h \alpha_h^{-n\delta} K^2 (1+\varepsilon)^2 \sum_{j=0}^{m-1} \alpha_h^{2j+1} X_h^2(j).$$

Note that the only major difference between this estimate and the equivalent estimate (6.8.4) in the proof of Theorem 6.4.2 is the extra term $2\beta h\alpha_h^{-n\delta} M_h^*(m)$. However, it can be shown that this term is of order $\alpha_h^{n\delta}$ and so it will be dominated by the term of order $\alpha_h^{n\delta}\log n$ as $n \to \infty$ and it will not contribute to the final asymptotic estimate. The proof now follows similar steps to the proof of Theorem 6.4.2 to obtain the desired result. □

Proof of Theorem 6.4.5. Consider the first step of the discretisation given by (6.4.5). Multiplying across by $X_h^*(n)$ on both sides gives

$$\left(X_h^*(n)\right)^2 = X_h(n) X_h^*(n) + h X_h^*(n) f(X_h^*(n), nh), \quad n \ge 0.$$

Then using (6.2.9) and the fact that $xy \le \frac{1}{2}x^2 + \frac{1}{2}y^2$ for $x, y \in \mathbb{R}$, we can show that

$$\left(X_h^*(n)\right)^2 \le X_h^2(n) + 2h\rho, \quad n \ge 0. \tag{6.8.10}$$

Now consider the second step of the discretisation given by (6.4.5). Squaring it gives

$$X_h^2(n+1) = \left(X_h^*(n)\right)^2 + 2\sqrt{h} X_h^*(n) g(X_h^*(n), nh)\xi(n+1) + g^2(X_h^*(n), nh) h\xi^2(n+1)$$
$$\le X_h^2(n) + 2h\rho + 2\sqrt{h} X_h^*(n) g(X_h^*(n), nh)\, \xi(n+1) + K_1^2 h \xi^2(n+1),$$

where we have used the fact that $g^2(x,nh) \leq K_1^2$. Defining $\lambda_h := K_1^2 h + 2h\rho$ and summing on both sides then yields

$$X_h^2(n+1) - X_h^2(0) \leq \sum_{j=0}^{n} \lambda_h + \theta_h(n+1) + M_h(n+1), \quad n \geq 0,$$

where we define the martingales

$$\theta_h(n+1) := \sum_{j=0}^{n} 2\sqrt{h} X_h^*(j) g(X_h^*(j), jh)\, \xi(j+1), \quad n \geq 0,$$

$$M_h(n+1) := \sum_{j=0}^{n} K_1^2 h[\xi^2(j+1) - 1], \quad n \geq 0.$$

We now apply the exponential martingale inequality to the martingale $\theta_h(n+1)$ which has quadratic variation given by

$$\langle \theta_h \rangle(n+1) = \sum_{j=0}^{n} 4h\big(X_h^*(j)\big)^2 g^2(X_h^*(j), jh).$$

Applying Theorem 5.2.4, where $\beta > 0$ and $\tau > 1$ are arbitrary constants, we have for all $n \in \mathbb{N}$,

$$\mathbb{P}\Big[\max_{1 \leq m \leq \lfloor \tau^n \rfloor} \Big\{\theta_h(m) - \frac{2h\beta}{\tau^n} \sum_{j=0}^{m-1} \big(X_h^*(j)\big)^2 g^2(X_h^*(j), jh)\Big\} \geq \frac{\tau^{n+1}\log n}{\beta}\Big] \leq \frac{1}{n^\tau}.$$

The Borel–Cantelli lemma then yields that for all $\omega \in \Omega_0$, where $\mathbb{P}[\Omega_0] = 1$, there is a random integer $n_0 = n_0(\omega, h)$ sufficiently large such that for $n \geq n_0$, and for $1 \leq m \leq \lfloor \tau^n \rfloor$,

$$\begin{aligned}\theta_h(m) &\leq \beta^{-1}\tau^{n+1}\log n + \frac{2h\beta}{\tau^n} \sum_{j=0}^{m-1} \big(X_h^*(j)\big)^2 g^2(X_h^*(j), jh) \\ &\leq \beta^{-1}\tau^{n+1}\log n + \frac{2\beta}{\tau^n} \sum_{j=0}^{m-1} 2K_1^2 h^2 \rho + \frac{2\beta h}{\tau^n} \sum_{j=0}^{m-1} g^2(X_h^*(j), jh) X_h^2(j),\end{aligned}$$

where we have used (6.8.10) and the fact that $g^2(x, jh) \leq K_1^2$. Using the same method as used in the proof of Theorem 6.4.4 we can show that

$$\sum_{j=0}^{m-1} g^2(X_h^*(j), jh) X_h^2(j) \leq K^2(1+\varepsilon)^2 \sum_{j=0}^{m-1} X_h^2(j) + M_h^*(m),$$

where we define $M_h^*(m) := K_1^2 F_h^2(X_1(\varepsilon))m$. Thus for $n \geq n_0$,

$$\theta_h(m) \leq \beta^{-1}\tau^{n+1}\log n + \frac{2\beta}{\tau^n}\sum_{j=0}^{m-1} 2K_1^2h^2\rho + 2\beta h\tau^{-n}M_h^*(m) \\ + 2\beta h\tau^{-n}K^2(1+\varepsilon)^2\sum_{j=0}^{m-1} X_h^2(j), \qquad 1 \leq m \leq \lfloor \tau^n \rfloor .$$

Again, the only major difference between this estimate and the equivalent estimate (6.8.7) in the proof of Theorem 6.4.3 is the extra term $2\beta h\tau^{-n}M_h^*(m)$. However, it can be shown that this term will be dominated by the term of order $\tau^{n+1}\log n$ as $n \to \infty$ and it will not contribute to the final asymptotic estimate. The proof now follows similar steps to the proof of Theorem 6.4.3 to obtain the desired result. □

Simulation of a Simple Two–State Markov Jump Process

The need for the simulation of a Markov jump process became evident in light of the analysis in Chapter 4 and in particular, the comments made in Remark 4.2.3. The issue is that the rate of convergence of the ergodic theorem for Markov chains (Proposition 4.2.1) appears to depend upon the step size h of the discretisation process. As a result, the log of the discretised stock price S_h (minus its trend) obeys an iterated logarithm growth bound of the form

$$\limsup_{n\to\infty} \frac{|\log S_h(n) - (\mu - \frac{1}{2}\sigma_*^2)nh|}{\sqrt{2nh\log\log nh}} \le \sigma_*\sqrt{2e} + \frac{1}{2}\sum_{i\in\mathbb{S}} \gamma^2(i)\sqrt{h}\alpha_i, \quad \text{a.s.},$$

as seen in Theorem 4.2.4. On the other hand, the underlying continuous–time stock price S (minus its trend) obeys an iterated logarithm growth bound of the form

$$\limsup_{t\to\infty} \frac{|\log S(t) - (\mu - \frac{1}{2}\sigma_*^2)t|}{\sqrt{2t\log\log t}} \le \sigma_* + \frac{1}{2}\sum_{i\in\mathbb{S}} \gamma^2(i)\beta_i, \quad \text{a.s.},$$

as seen in Theorem 3.4.3. Therefore, if it was true that $\sqrt{h}\alpha_i \to 0$ as $h \to 0$, then the result for the discretised stock price would not be consistent with the result for the continuous–time stock price. This prompted further analysis and led us to attempt to simulate the problem numerically and to give evidence which *suggests* that in fact $\sqrt{h}\alpha_i \to \alpha_i^*$ as $h \to 0$ for some finite value α_i^*. The Matlab code used to produce these simulations was adapted from code originally developed by Craig L. Zirbel and can be found here:

`http://www-math.bgsu.edu/z/ap/`

Due to computational and time constraints, we simplify the problem to study a two–state Markov jump process with generator matrix

$$\Gamma = \begin{pmatrix} -\gamma_1 & \gamma_1 \\ \gamma_2 & -\gamma_2 \end{pmatrix}.$$

Then our discretised Markov chain has a matrix of transition probabilities of the form

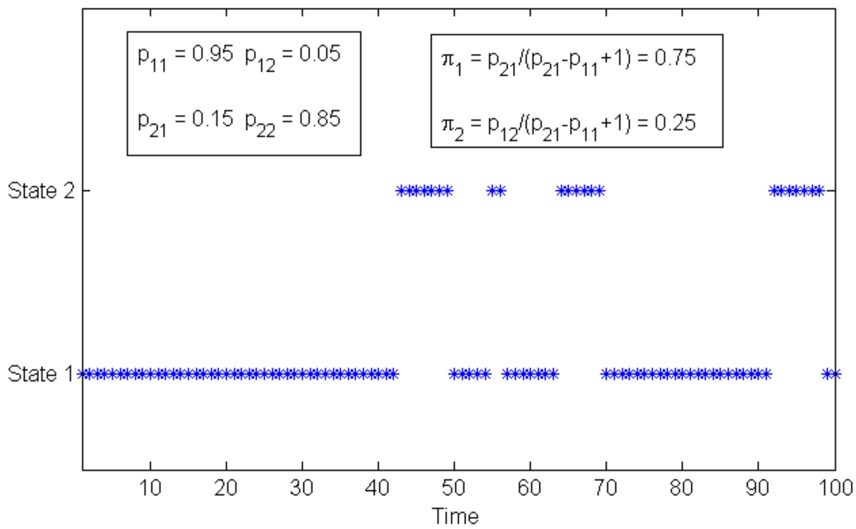
$$P(h) = I + h\Gamma = \begin{pmatrix} 1 - h\gamma_1 & h\gamma_1 \\ h\gamma_2 & 1 - h\gamma_2 \end{pmatrix} \qquad \text{where } h < \max\left(\frac{1}{\gamma_1}, \frac{1}{\gamma_2}\right).$$

As mentioned in Remark 4.2.2, this is an *approximation* of the chain. To begin with, we choose $\gamma_1 = 1, \gamma_2 = 3$ and $h = .05$ so that $1 - h\gamma_1 = 0.95$ meaning that the probability of staying in state 1 is 0.95 and similarly, $1 - h\gamma_2 = 0.85$ meaning that the probability of staying in state 2 is 0.85. Moreover, the stationary distribution $\pi = (\pi_1, \pi_2)$ can be found by solving $\pi\Gamma = 0$ subject to $\pi_1 + \pi_2 = 1$. This gives

$$\pi_1 = \frac{\gamma_2}{\gamma_1 + \gamma_2} = \frac{3}{4} \quad \text{and} \quad \pi_2 = \frac{\gamma_1}{\gamma_1 + \gamma_2} = \frac{1}{4},$$

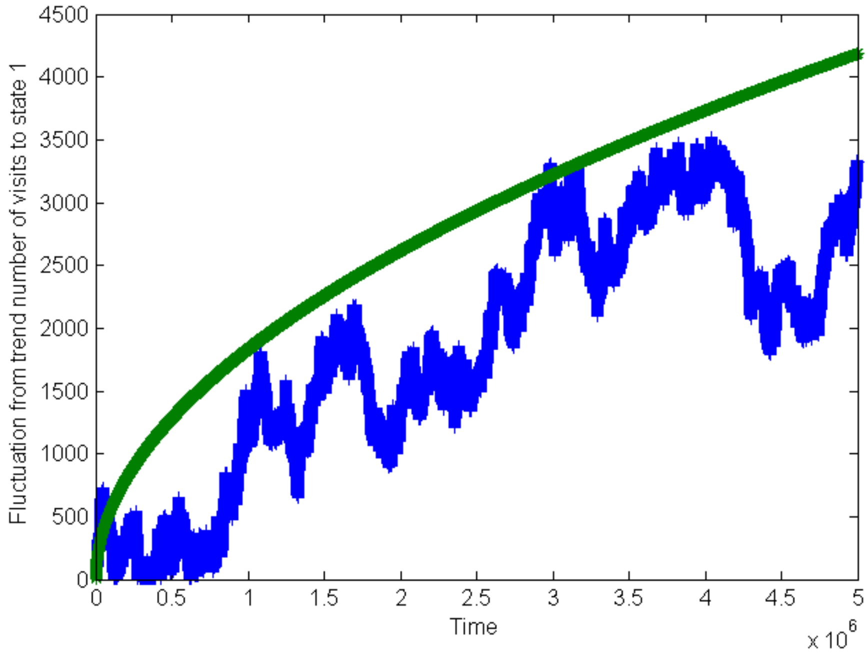
which means that in the long run the process will spend 75% of the time in state 1 and 25% of the time in state 2. The simulation of such a chain over 100 time steps is given below using the notation p_{ij} as the $i - j^{th}$ entry of the transiton matrix $P(h)$.

Figure A.1: Two–state Markov chain



Using this code we can also count the number of times that the chain has been in a particular state up to a given number of time steps and in this two–state example we can restrict our attention to state 1 only. The number of times that the chain has been in a state is crucial to the estimation of the rate of convergence of the ergodic theorem as it corresponds to the quantity $V_i(n)$ in the proof of Proposition 4.2.1. Indeed we see from (4.3.26) that the rate of convergence we require is determined by the rate of convergence of $V_1(n)$ to $n\pi_1$. With this information we can plot $|V_1(n) - n\pi_1|$ alongside $\sqrt{2n \log \log n}$ to see that it does indeed obey an iterated logarithm growth bound as seen in the figure below.

Figure A.2: Fluctuations from trend of $V_1(n)$ vs. $\sqrt{2n \log \log n}$



This simulation is very much consistent with (4.3.25), which states that

$$\limsup_{n\to\infty} \frac{|V_1(n) - \pi_1 n|}{\sqrt{2n \log \log n}} \le \alpha_1, \quad \text{a.s.} \tag{A.0.1}$$

We then want to approximate the magnitude, α_1, of the iterated logarithm growth rate. That is, by considering the running maximum of $V_1(n) - n\pi_1$ divided by $\sqrt{2n \log \log n}$ at the final time step we get an estimate on a value A_1 where

$$\limsup_{n\to\infty} \frac{|V_1(n) - \pi_1 n|}{\sqrt{2n \log \log n}} = A_1.$$

Averaging a number of such calculations gives a value of approximately $A_1 \approx 0.96$. Therefore, in the case when $h = 0.05$ we have $\sqrt{h} A_1(h) \approx 0.215$.

We then repeat such calculations where the step size is smaller by a factor of 100, i.e. $h = 0.0005$. Averaging a number of these calculations gives a value of approximately $A_1(h) \approx 10.5$ which (significantly) is bigger than the previous $A_1(h)$ by a factor of $\sqrt{100}$. Therefore, in this case when $h = 0.0005$ we have $\sqrt{h}A_1(h) \approx 0.235$.

One can continue like this to obtain approximations of similar orders of magnitude. We conclude from this that there exists a finite number A_1^* such that $\sqrt{h}A_1 \to A_1^*$ as $h \to 0$.

Unfortunately, due to the fact that we have to sum along the sequence $V_i(n)$ in (4.3.23) instead of summing along the integers as in (4.3.22), we lose equality in (4.3.23) and consequently in (A.0.1) and thus we cannot be certain that A_1 is actually equal to α_1. Although we are unable to prove it, we conjecture that the sequence $V_i(n)$ is such that equality is in fact preserved in (4.3.23) which would result in equality in (A.0.1). This would mean that the estimate A_1 of the numerical simulation coincides with the bound α_1. Clearly it is far from a solid argument, but we believe that due to the Markov property and the resulting loss of memory, the sequence $V_i(n)$ has sufficient independence and near–linear growth that the equality in (4.3.23) is preserved and that there is a correspondence between the A_1 and α_1.

Naturally, if this is true it would also extend to a Markov process with finite state space $\mathbb{S}$ and to any $i \in \mathbb{S}$. This leads us to believe that for each $i \in \mathbb{S}$ there exists α_i^* such that $\sqrt{h}\alpha_i \to \alpha_i^*$ as $h \to 0$.

Bibliography

[1] M. J. Anabtawi and G. S. Ladde. Convergence and stability analysis of systems of partial differential equations under Markovian structural perturbations: I, II. *Stoch. Anal. Appl.*, 18(4,5):493–524, 671–696, 2000.

[2] M. J. Anabtawi and S. Sathananthan. Stochastic hybrid parabolic systems under jump Markovian perturbations: (I) convergence and stability via Lyapunov functionals. *Dynam. Systems Appl.*, 15(2):269–286, 2006.

[3] M. J. Anabtawi, S. Sathananthan, and G.S. Ladde. Convergence and stability analysis of large-scale parabolic systems under Markovian structural perturbations: I, II. *Int. J. Appl. Math.*, 2(1):57–85, 87–111, 2000.

[4] W. J. Anderson. Continuous-time Markov chains. *Springer, New York*, 1991.

[5] J. A. D. Appleby, X. Mao, and H. Wu. On the size of the largest pathwise deviations of stochastic functional differential equations. *to appear, Nonlinear Studies.*

[6] J. A. D. Appleby, X. Mao, and H. Wu. On the almost sure running maxima of solutions of affine stochastic functional differential equations. *SIAM J. Math. Anal.*, 42(2):646–678, 2010.

[7] J. A. D. Appleby and H. Wu. Solutions of stochastic differential equations obeying the law of the iterated logarithm with applications to financial markets. *Electron. J. Probab*, 14:912–959, 2009.

[8] J.A.D Appleby, J. Cheng, and A. Rodkina. The split-step Euler–Maruyama method preserves asymptotic stability for simulated annealing problems. *Proceedings of Neural, Parallel, and Scientific Computations 4*, page 6pp, 2010.

[9] J.A.D Appleby, T. Lynch, X. Mao, and H. Wu. The size of the largest fluctuations in a market model with Markovian switching. *Communications in Applied Analysis*, 13(2):135–166, 2009.

[10] J.A.D Appleby, A. Rodkina, and D. Mackey. Almost sure polynomial asymptotic stability of stochastic difference equations. *Journal of Mathematical Sciences*, 149(6):1629–1647, 2008.

[11] J.A.D Appleby, A. Rodkina, and M. Riedle. On asymptotic stability of linear stochastic Volterra difference equations with respect to a fading perturbation. *Adv. Stud. Pure Math.*, 53:271–282, 2009.

[12] M. Athans. Command and control (C2) theory: A challenge to control science. *IEEE Trans. Automat. Contr.*, 32:286–293, 1987.

[13] B. Bercu and A. Touati. Exponential inequalities for self-normalized martingales with applications. *Annals of Applied Probability*, 18(5), 2008.

[14] N. H. Bingham, C. M. Goldie, and J. L. Teugels. Regular variation. *Cambridge University Press, Cambridge*, 1987.

[15] M. Borkovec and C. Klüppelberg. Extremal behavior of diffusion models in finance. *Extremes 1*, (1,3):47–80, 379–380, 1998,1999.

[16] Z. Brezeźniak and T. Zastawniak. Basic stochastic processes. *Springer Verlag, London*, 1999.

[17] P. H. Cootner. The random character of stock market prices. *MIT Press, Cambridge*, 1964.

[18] V.H. de la Peña. A general class of exponential inequalities for martingales and ratios. *Ann. Probab.*, 27:537–564, 1999.

[19] V.H. de la Peña, M.J. Klass, and T.L. Lai. Self-normalized processes: exponential inequalities, moment bounds and iterated logarithm laws. *Ann. Probab.*, 32 (3A):1902–1933, 2004.

[20] V.H. de la Peña and G. Pang. Exponential inequalities for self-normalized processes with applications. *Electron. Commun. Probab.*, 14:372–381, 2009.

[21] B. de Saporta and J. F. Yao. Tail of a linear diffusion with Markov switching. *Ann. Appl. Probab.*, 15(1B):992–1018, 2005.

[22] R. J. Elliott, T. K. Siu, and L. Chan. Option pricing for GARCH models with Markov switching. *Int. J. Theor. Appl. Finance*, 9(6):825–841, 2006.

[23] R. J. Elliott, T. K. Siu, and L. Chan. Pricing volatility swaps under Heston's stochastic volatility model with regime switching. *Appl. Math. Finance*, 14(1):41–62, 2007.

[24] E. F. Fama. The behavior of stock-market prices. *J. Business*, 38(1):34–105, 1965.

[25] E. F. Fama. Market efficiency, long-term returns, and behavioral finance. *J. Financial Economics*, 49(3):283–306, 1998.

[26] J. P. Fouque, G. Papanicolaou, and K. R. Sircar. Derivatives in financial markets with stochastic volatility. *Cambridge University Press, Cambridge*, 2000.

[27] P. H. Franses and D. van Dijk. Nonlinear time series models in empirical finance. *Cambridge University Press, Cambridge*, 2000.

[28] M.K. Ghosh, A. Arapostathis, and S.I. Marcus. Optimal control of switching diffusions with application to flexible manufacturing systems. *SIAM J. Control Optim.*, 35:1183–1204, 1993.

[29] M.K. Ghosh, A. Arapostathis, and S.I. Marcus. Ergodic control of switching diffusions. *SIAM J. Control Optim.*, 35:1952–1988, 1997.

[30] J. Golec and S. Sathananthan. Stability analysis of a stochastic logistic model. *Math. Comput. Modelling*, 38(5-6):585–593, 2003.

[31] G. R. Grimmett and D. R. Stirzaker. Probability and random processes. *Oxford University Press, third edition*, 2001.

[32] D. Guillaume, M. Dacorogna, R. Davé, U. Müller, R. Olsen, and O. Pictet. From the bird's eye to the microscope: a survey of new stylized facts of the intra-daily foreign exchange markets. *Finance and Stochastics*, 1:95–129, 1997.

[33] J. D. Hamilton. A new approach to the economic analysis of nonstationary time series subject to changes in regime. *Econometrica*, 57:357–384, 1989.

[34] D.J. Higham. Mean-square and asymptotic stability of numerical methods for stochastic ordinary differential equations. *SIAM J. Numer. Anal.*, 38:753–769, 2000.

[35] D.J. Higham, S. Intep, and X. Mao. Switching and diffusion models for gene regulation networks. *Multiscale Modeling and Simulation*, 8:30–45, 2009.

[36] D.J. Higham, X. Mao, and A.M. Stuart. Strong convergence of Euler-like methods for nonlinear stochastic differential equations. *SIAM J. Numer. Anal.*, 40:1041–1063, 2002.

[37] D.J. Higham, X. Mao, and A.M. Stuart. Exponential mean square stability of numerical solutions to stochastic differential equations. *LMS J. Comput. Math.*, 6:297–313, 2003.

[38] D.J. Higham, X. Mao, and C. Yuan. Almost sure and moment exponential stability in the numerical simulation of stochastic differential equations. *SIAM J. Numer. Anal.*, 45:592–609, 2007.

[39] D.J. Higham, X. Mao, and C. Yuan. Preserving exponential mean-square stability in the simulation of hybrid stochastic differential equations. *Numer. Math.*, 108(2):295–325, 2007.

[40] D.J. Higham, J. Mattingly, and A.M. Stuart. Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stoch. Proc. Appl.*, 101:185–232, 2002.

[41] J. C. Hull. Options, futures and other derivatives. *Prentice Hall, New York, Sixth Edition*, 2005.

[42] J. C. Hull and A. White. The pricing of options on assets with stochastic volatilities. *J. Financial and Quantitative Analysis*, 23:237–251, 1987.

[43] J.J. Hunter. On the moments of Markov renewal processes. *Advances in Applied Probability*, 1(2):188–210, 1969.

[44] J.J. Hunter. Mathematical techniques of applied probability volume 2, discrete time models: Techniques and applications. *Academic Press, New York*, 1983.

[45] K. Itô and H.P. McKean. Diffusion processes and their sample paths. *Springer, Berlin, 2nd Edition*, 1974.

[46] I. Karatzas and S. Shreve. Brownian motion and stochastic calculus. *Springer*, 1991.

[47] I. Karatzas and S. Shreve. Methods of mathematical finance. *Springer-Verlag, New York*, 1998.

[48] J.G. Kemeny and J.L. Snell. Finite Markov chains. *Van Nostrand, New York*, 1960.

[49] G. S. Ladde. Large-scale stochastic hereditary systems under Markovian structural perturbations. I, II, III. *Appl. Math. Stoch. Anal.*, (Art. ID 19871, 67268, 24643.), 2006.

[50] A. Lo and A. C. MacKinlay. A non-random walk down Wall Street. *Princeton University Press*, 1999.

[51] Q. Luo and X. Mao. Stochastic population dynamics under regime switching. *J. Math. Anal. Appl.*, 334(1):69–84, 2007.

[52] X. Mao. Almost sure asymptotic bounds for a class of stochastic differential equations. *Stochastics Stochastics Rep.*, 41(1–2):57–69, 1992.

[53] X. Mao. Exponential stability of stochastic differential equations. *Marcel Dekker*, 1994.

[54] X. Mao. Stochastic differential equations and their applications. *Chichester: Horwood Pub.*, 1997.

[55] X. Mao. Stability of stochastic differential equations with Markovian switching. *Stochastic Process. Appl.*, 79:45–67, 1999.

[56] X. Mao and M.J. Rassias. Almost sure asymptotic estimations for solutions of stochastic differential delay equations. *Int. J. Appl. Math. Stat.*, J07(9):95–109, 2007.

[57] X. Mao, A. Truman, and C. Yuan. Euler-Maruyama approximations in mean-reverting stochastic volatility models under regime-switching. *J. Appl. Math. Stoch. Anal.*, (Art. ID 80967), 2006.

[58] X. Mao, G. Yin, and C. Yuan. Numerical method for stationary distribution of stochastic differential equations with Markovian switching. *J. Comput. Appl. Math.*, 174(1):1–27, 2005.

[59] X. Mao, G. Yin, and C. Yuan. Stabilization and destabilization of hybrid systems of stochastic differential equations. *Automatica J. IFAC*, 43(2):264–273, 2007.

[60] X. Mao and C. Yuan. Asymptotic stability in distribution of stochastic differential equations with Markovian switching. *Stochastic Process. Appl.*, 103(2):277–291, 2003.

[61] X. Mao and C. Yuan. Convergence of the Euler-Maruyama method for stochastic differential equations with Markovian switching. *Math. Comput. Simulation*, 64(2):223–235, 2004.

[62] X. Mao and C. Yuan. Stochastic differential equations with Markovian switching. *Imperial College Press, London*, 2006.

[63] L. McKenzie. Stability of equilibrium and the value of positive excess demand. *Econometrica*, 28(3):606–617, 1960.

[64] R. C. Merton. Theory of rational option pricing. *Bell J. Econom. and Management Sci.*, 4:141–183, 1973.

[65] M. Motoo. Proof of the law of the iterated logarithm through diffusion equation. *Ann. Inst. Statist. Math.*, 10:21–28, 1958.

[66] U. Müller, M. Dacorogna, and O. Pictet. Heavy tails in high-frequency financial data: A practical guide to heavy tails. *Birkhäuser Publishers*, pages 55–79, 1998.

[67] J.R. Norris. Markov chains. *Cambridge University Press*, 1997.

[68] R. Pyke. Markov renewal processes with finitely many states. *Annals of Mathematical Statistics*, 32(4):1243–1259, 1961.

[69] E. Renault and N. Touzi. Option hedging and implied volatilities in a stochastic volatility model. *J. Math. Finance*, 6(3):279–302, 1996.

[70] D. Revuz and M. Yor. Continuous martingales and Brownian motion. *Springer-Verlag, Berlin*, 1999.

[71] A. Rodkina and H. Schurz. Almost sure asymptotic stability of drift-implicit theta-methods for bilinear ordinary stochastic differential equations in $\mathbf{R}^1$. *J. Comput. Appl. Math.*, 180(2):13–31, 2005.

[72] S. Sathananthan and L. H. Keel. Optimal practical stabilization and controllability of systems with Markovian jumps. *Nonlinear Anal.*, 54(6):1011–1027, 2003.

[73] S. Sathananthan and L. H. Keel. Convergence and stability of stochastic hereditary iterative processes. *Dyn. Contin. Discrete Impuls. Syst. Ser. A Math. Anal.*, 11(1):81–104, 2004.

[74] S.P. Sethi and Q. Zhang. Hierarchical decision making in stochastic manufacturing systems. *Birkhäuser, Boston*, 1994.

[75] A.N. Shiryaev. Probability, 2nd. ed. *Springer–Verlag, New York*, 1989.

[76] D. Smith. Markov-switching and stochastic volatility diffusion models of short-term interest rates. *J. Bus. Econom. Statist.*, 20(2):183–197, 2002.

[77] A.M. Stuart and A.R. Humphries. Dynamical systems and numerical analysis. *Cambridge University Press, Cambridge*, 1996.

[78] D.D. Sworder and R.O. Rogers. An LQ-solution to a control problem associated with solar thermal central receiver. *IEEE Trans. Automat. Contr.*, 28:971–978, 1983.

[79] J. M. Vanden. Portfolio insurance and volatility regime switching. *Math. Finance*, 16(2):387–417, 2006.

[80] D. Williams. Probability with martingales. *Cambridge University Press*, 1991.

[81] A.S. Willsky and B.C. Levy. Stochastic stability research for complex power systems. *DOE Contract, LIDS, MIT, Rep. ET-76-C-01-2295*, 1979.

[82] G. Yin, R.H. Liu, and Q. Zhang. Recursive algorithms for stock liquidation: A stochastic optimization approach. *SIAM J. Optim.*, 13:240–263, 2002.

[83] G. Yin and Q. Zhang. Continuous-time Markov chains and applications: A singular perturbation approach. *Springer-Verlag, New York*, 1998.

[84] G. Yin and X.Y. Zhou. Markowitz's mean-variance portfolio selection with regime switching: From discrete-time models to their continuous-time limits. *IEEE Transactions on Automatic Control*, 49:349–360, 2004.

[85] C. Yuan and J. Lygeros. Asymptotic stability and boundedness of delay switching diffusions. *IEEE Trans. Automat. Control*, 51(1):171–175, 2006.

[86] Q. Zhang. Stock trading: An optimal selling rule. *SIAM J. Control. Optim.*, 40:64–87, 2001.