August 2015

A Weak Simpson Method for a Class of Stochastic Differential Equations and Numerical Stability Results

Ram Sharan Adhikari
University of Wisconsin-Milwaukee

A weak Simpson method for a class of stochastic differential equations and numerical stability results

by

Ram Sharan Adhikari

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

Doctor of Philosophy
in
Mathematics

at

The University of Wisconsin–Milwaukee
August 2015

ABSTRACT

A weak Simpson method for a class of stochastic differential equations and numerical stability results

by

Ram Sharan Adhikari

The University of Wisconsin-Milwaukee, 2015
Under the Supervision of Professors Bruce Wade and Chao Zhu

This work proposes a novel weak Simpson method for numerical solution for a class of stochastic differential equations. We show that such a method has weak convergence of order one in general and weak convergence of order three under certain additional assumptions. This work also aims to determine the mean-square stability region of the weak Simpson method for linear stochastic differential equations with multiplicative noises. In this work, a mean-square stability region of the weak Simpson scheme is identified, and stepsizes for the numerical method where errors propagation are under control in well-defined sense are given. The main results are illustrated with numerical examples.

Table of Contents

List of Figures

ACKNOWLEDGMENTS

First and foremost, I would like to express my deepest gratitude and thanks to my advisors, Professors Bruce Wade and Chao Zhu, who guide and encourage me throughout my graduate studies. This dissertation would not be possible without their contributions. I am also very grateful to Professors Richard Stockbridge, Hans Volkmer, and Wei Wei for their help and support.

I would like to thank my parents for their support and encouragement for me to pursue my graduate studies. Last, but not least, I am thankful to my wife, Muna Adhikari Koirala, for supporting me through every stage of this difficult endeavor.

Chapter 1

Introduction

Many real-world dynamics evolve over time in ways that can not be predicted with certainty, for example, the price of an asset in the stock market, the number of claims to an insurance company, the growth of a certain species in the random environment, etc. Stochastic ordinary and partial differential equations naturally arise as realistic yet tractable mathematical models to describe such complicated dynamics subject to various noises. Stochastic differential equations (SDEs) have wide range of applications in many fields such as mathematical biology, population dynamics, protein kinetics and genetics, psychology and neuronal activity, investment finance and option pricing, turbulent diffusion and radio-astronomy, helicopter rotor and satellite orbit stability, seismology and structural mechanics, blood clotting dynamics and cellular energetics.

It is natural that people would like to understand and make inferences about the real-world dynamics by studying the qualitative and quantitative properties of the solutions to the underlying SDEs. Unfortunately, in most of the practical applications, we can not find explicit solutions for the underlying SDEs, as with most ordinary differential equations (ODEs). In such situations, numerical approximation then becomes the one viable approach. For example, it is demonstrated in Bruti-Liberati and Platen (2008) that simulation methods for the approximate solution of stochastic differential equations have become absolutely necessary tools in many areas of application. In fact, numerical solution of SDEs has been an important research area and has drawn continuing attention. We refer to the excellent books Milstein (1995) and Kloeden and Platen (1992) for extensive treatments on numerical solution of SDEs.

We consider the problem of constructing accurate approximations on fixed time intervals to solutions of the following system of SDEs

$$\begin{aligned} X(s) &= x + \int_0^s b(X(r))\,\mathrm{d}r + \sum_{k=1}^{M} \int_0^s \sigma_k(X(r))\nu_k\,\mathrm{d}W_k(r), \quad s \geq 0, \\ X(0) &= x \in \mathbb{R}^d, \end{aligned} \tag{1.1}$$

where $M \in \mathbb{N}$ is a positive integer, $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma_k : \mathbb{R}^d \to \mathbb{R}$, and for $k = 1, \ldots, M$, $\nu_k \in \mathbb{R}^d$, and $W_k(t)$ are independent one dimensional Brownian motions. Here for each k, ν_k represents the direction along which the random noise W_k enters the system (1.1). Suppose the coefficients b and σ_k are measurable and are such that a weak solution to (1.1) exists and is unique in probability law. Typically the coefficients b and σ_k are assumed to satisfy the Lipschitz continuity and the linear growth condition; see, for example, Øksendal (2003) or Karatzas and Shreve (1991).

This work is motivated by Anderson and Mattingly (2011) and improves their weak trapezoidal method. In Anderson and Mattingly (2011), the weak trapezoidal method has weak convergence of order two and seems to require a large number of sample paths (10 million). The key idea behind the weak trapezoidal method in Anderson and Mattingly (2011) is that the solution of (1.1) is equal in distribution to the solution of a differential equation (2.10) driven by a space-time white noise. Then they use the trapezoidal method to approximate the area under the diffusion curve $\sigma_k^2(\cdot)$, which is equivalent in determining the distribution of the diffusion term in (1.1). Motivated by the fact that Simpson's rule is usually an improvement of the trapezoidal rule, we use a Simpson-like rule to approximate the area under the diffusion term in (1.1). In other words, we use a weak Simpson method to approximate the stochastic integral of (1.1) in our algorithm. We show that our method has weak convergence of order one in general and weak convergence of order three under certain additional assumptions (Assumption (A4)). However, we note that our examples in Chapter 3 all demonstrate weak order three convergence, even though they do not necessarily satisfy the Assumption (A4). Moreover, in theses examples, our method requires fewer number (in the order of 50,000) of sample paths compared to Anderson and Mattingly (2011). Unfortunately, at this point, we are not able to prove that the weak Simpson method enjoys weak order three convergence without Assumption (A4).

The algorithm of our method consists of two steps, in the first an explicit Euler-Maruyama type step is used and in the second the resulting fractional point is used in combination with initial point to obtain higher order. We use variable steps to obtain better approximation. The use of different paths for each time step-size make sense in our setting because we are only concerned with the mean of the solution. We can choose any sample $\sqrt{h}N(0,1)$ for the increment $W(t_k) - W(t_{k-1})$. This work develops the method which produce weak approximation rather than strong approximation. Hence, we produce an approximating sample path without giving proper attention to the underlying Wiener process. Our algorithm does not require simulation of the Itô integral.

It is worth noting that there are many other higher order weak Taylor schemes to solve stochastic differential equations, see, for example, Kloeden and Platen (1992). Generally, these higher order methods are much more complicated than ours, and contain a large number of terms, such as all of the multiple Itô integrals of higher multiplicity from Itô-Taylor expansion Kloeden and Platen (1992). For instance, we need to include all of the third order multiple Itô integrals from the Itô-Taylor expansion to construct order three weak Taylor scheme. This makes these methods hard to implement in practice. Compared to those higher order weak Taylor schemes, our algorithm is simple and derivative free, yet still enjoys weak convergence of order three under certain additional assumptions.

Most stochastic differential equations can not be solved explicitly. However, a great deal of useful qualitative information can be obtained about the behavior of their solutions. Asymptotic behavior and the impact of small changes in initial values are of particular interest in applications. We know that if a differential equation is well-posed, then a solution exists and is unique; moreover, the solution is continuous with respect to the initial value in some sense. The concept of stability is an extension of this idea to an infinite time interval (Kloeden and Platen (1992)). In this thesis we shall study mean-square stability and almost sure asymptotic stability of our proposed method in relation to a scalar Itô equation

$$\mathrm{d}X(t) = b(X(t))\,\mathrm{d}t + \sigma(X(t))\,\mathrm{d}W(t), \quad X(0) = x_0 \in \mathbb{R}^d, \tag{1.2}$$

where we assume that (1.2) has a steady solution $X(t) \equiv 0$.

In some sense, the stability of a numerical scheme refers to the conditions under which the impact of an error vanishes asymptotically over time, see, for example, Bruti-Liberati and Platen (2008). Generally, the notion of numerical stability is challenging to quantify. Nevertheless various concepts of numerical stability for different schemes have been extensively studied by many authors. Most of the literature in numerical stability use specially designed test equations, see for instance, Kloeden and Platen (1992). We systemically analyze the stability properties of our scheme for the given family of test equations.

To facilitate the presentation of the thesis, we introduce some notation here. The notation a^T denotes the transpose of a vector or matrix a. For a smooth function $f : \mathbb{R}^d \mapsto \mathbb{R}$, $Df(x) = (\frac{\partial f(x)}{\partial x_1}, \ldots, \frac{\partial f(x)}{\partial x_d})^T$ denotes the gradient of f at x and $D^2 f(x) = (\frac{\partial^2 f(x)}{\partial x_i \partial x_j})$ is the Heissian of f at x. For $\eta, x \in \mathbb{R}^d$, $f'[\eta](x) := \eta \cdot Df(x)$ is the derivative of f in the direction of η evaluated at the point x. And for $\eta, \nu, x \in \mathbb{R}^d$, $f''[\eta, \nu](x) := f'(f'[\eta])[\nu](x)$. In a similar fashion, we define $f'''[\eta, \nu, \xi](x)$ etc. Note that $f''[\eta, \nu](x) = \mathrm{tr}(\eta \nu^T D^2 f(x)) = f''[\nu, \eta](x)$.

A vector $\alpha = (\alpha_1, \ldots, \alpha_d)$ with each component α_i taking values from the set of non-negative integers is called a multi-index. Moreover, we denote $|\alpha| := \alpha_1 + \cdots + \alpha_d$ and $D^\alpha f(x) = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha 1} \ldots \partial x_d^{\alpha d}} f(x)$. We denote the space of continuous and bounded functions whose first through kth order partial derivatives are continuous and bounded by $C^k(\mathbb{R}^d)$, that is,

$$C^k(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \mathbb{R} \text{ s.t. } D^\alpha f \text{ exists, bounded and continuous, }\}.$$

for all multi-index α with $|\alpha| \le k$. In addition, we define the norm of $C^k(\mathbb{R}^d)$ in the following way

$$\|f\|_k = \sup\left\{|D^\alpha f(x)| : x \in \mathbb{R}^d,\ \alpha = (\alpha_1, \ldots, \alpha_d) \text{ is a multi-index with } |\alpha| \le k\right\}. \tag{1.3}$$

Note that when $k = 0$, $C(\mathbb{R}^d) := C^0(\mathbb{R}^d)$ is the family of bounded and continuous functions. Also, $\mathcal{B}(\mathbb{R}^d)$ is the family of real-valued, bounded and Borel measurable functions defined on $\mathbb{R}^d$.

The rest of the thesis is organized as follows. Chapter 2 starts with a discussion on various criteria for convergence. Then we propose our weak Simpson method and describe why and how our method works by considering the simple cases. Next we prove that our method converges with weak order one in general and three under suitable additional conditions. We demonstrate the numerical performance of the weak Simpson method in Chapter 3. In Chapter 4 we first recall the notions of mean-square and asymptotic stability. Then we present a result for linear stability of our scheme for the deterministic case. We also present the conditions for mean-square and asymptotic stability for our scheme as well as the stability regions for our scheme in Chapter 4. Finally we give the proof of Lemmas 2.7 and 2.10 in the Appendix which we need to prove our local approximation theorem.

Chapter 2

The Weak Simpson Method

In this chapter, we discuss various criteria for convergence of numerical schemes in Section 2.1. In addition, the appropriate assumptions as well as a basic moment estimate for the solution to (1.1) are also arranged in Section 2.1. Next we propose the weak Simpson method in Section 2.2. To illustrate the idea of our scheme, we also present the motivation behind our algorithm for a simple case. We also give the theoretical proof for the order of local and global convergence of our proposed method in Section 2.3.

2.1 Preliminaries

The error criteria to be used depend on the type of application. If one is interested in just generating $X(T)$ sufficiently accurately (in the distributional sense), an appropriate error criterion may be

$$\sup_{f\in\mathcal{C}} \left|\mathbb{E}[f(X(T))] - \mathbb{E}[f((Y(N))]\right|,$$

for a suitable class $\mathcal{C}$ of smooth functions, where $Y(.)$ is the simulated path. The accuracy of the sample path approximation can be measured by a criterion such as

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X(t) - Y(t)|\right],$$

assuming that $X(\cdot)$ and $Y(\cdot)$ can be generated on a common probability space, or, for some suitably chosen p, via an L_p error criterion such as

$$\mathbb{E}\Big[\int_0^T |X(t) - Y(t)|^p\Big]dt,$$

$$\mathbb{E}\Big[\sum_{0\le n\le T} |X(n) - Y(n)|^p\Big]$$

in continuous and discrete time respectively (Asmussen and Glynn (2007)). We recall the following definition from Kloeden and Platen (1992).

Definition 2.1. We say that an approximating process Y converges in the strong sense with order $\gamma \in (0, \infty]$ if there exists a finite constant K and a positive constant δ_0 such that

$$\mathbb{E}\big[\,|X(T) - Y(N)|\,\big] \le Kh^{\gamma}, \quad N = \frac{T}{h},$$

for any time discretization stepsize $0 < h < \delta_0$. The strong order of convergence measures the rate at which the "mean of the error" decays as $h \to 0$.

In fact this definition generalizes the standard convergence criterion for ordinary differential equation, reducing to the usual definition when the diffusion coefficient of (1.1) is zero.

Strong convergence allows an accurate approximation to be computed and involves direct simulation of the sample path and demands the approximation be close to that of the Itô process. The order of convergence of strong approximation is sometimes less in the stochastic case than in the corresponding deterministic case, see, for example, Kloeden and Platen (1992). It is also observed in Bruti-Liberati and Platen (2008) that strong explicit methods, particularly, the widely used Euler-Maruyama method, sometimes work unreliably and generate large errors for certain step-sizes.

But if the goal is to have a good approximation of the probability distribution of the solution $X(t)$, individual realizations are not of primary interest. Weak approximations are used in simulating functionals of the form $\mathbb{E}[f(X(T))]$, where $T > 0$ and f is some function. For instance, the arbitrage-free price of a European call option is given by $\mathbb{E}_{\mathbb{Q}}[e^{-r(T-t)}(S(T)-K)^+|\mathcal{F}_t]$, in which $\mathbb{Q}$ is the risk-neutral measure, $r>0$ is the discounting factor, K is the strike price, and $S(T)$ is the price of the underlying asset at time T. For weak approximation, in leu of Definition 2.1 for strong approximation, a less demanding alternative is to measure the rate of decay of the "error of the means". This leads to the concept of weak convergence order. We recall the following definition for weak convergence from Kloeden and Platen (1992).

Definition 2.2. We say that a time discrete approximation Y converges in the weak sense with order $\beta>0$ if for any $f\in C^{2(\beta+1)}(\mathbb{R}^d)$ there exists a finite constant K and a positive constant δ_0 such that

$$|\mathbb{E}[f(X(T))]-\mathbb{E}[f(Y(N))]|\le Kh^{\beta},\quad N=\frac{T}{h} \tag{2.1}$$

for any time discretization with maximum step size $h\in(0,\delta_0)$.

If the stochastic part of the differential equation is zero and the initial value is deterministic, the definition reduces to the usual deterministic convergence criterion for ordinary differential equation and also agrees with the strong convergence criterion.

We state the following standing assumptions throughout the thesis:

(A1) The coefficients of (1.1) satisfy the Lipschitz and linear growth conditions:

$$\begin{aligned}
&|b(x)-b(y)|+\sum_{k=1}^{M}|\sigma_k(x)-\sigma_k(y)|\le\kappa\,|x-y|\,,\\
&|b(x)|+\sum_{k=1}^{M}|\sigma_k(x)|\le\kappa(1+|x|),
\end{aligned} \tag{2.2}$$

for all $k=1,\dots,M$ and $x,y\in\mathbb{R}^d$, where κ is a positive constant.

(A2) For each $k=1,\dots,M$, we have $\inf_{x\in\mathbb{R}^d}\{\sigma_k(x)\}>0$. In addition, there exists a positive constant $\lambda\in(0,1]$ so that for any $x,\xi\in\mathbb{R}^d$ we have

$$\lambda\,|\xi|^2\le\xi^T a(x)\xi\le\lambda^{-1}\,|\xi|^2\,, \tag{2.3}$$

where ξ^T denotes the transpose of ξ and $a(x):=\sum_{k=1}^{M}\sigma_k^2(x)\nu_k\nu_k^T$.

(A3) For all multi-index α with $|\alpha| \le 8$, we have

$$|D^\alpha b(x)| + \sum_{k=1}^{M} |D^\alpha \sigma_k(x)| \le K(1 + |x|^p), \text{ for all } x \in \mathbb{R}^d, \tag{2.4}$$

where K and p are positive numbers.

It is well-known that under Assumption (A1), the stochastic differential equation (1.1) has a unique strong solution; see, for example, Karatzas and Shreve (1991), Øksendal (2003) or Yin and Zhu (2010). Moreover, we have the following moment estimate:

Lemma 2.3 (Yin and Zhu (2010))**.** *Assume* (A1)*. Let $T > 0$ be fixed. Then for any positive constant p, we have*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X^x(t)|^p\right] \le C < \infty, \quad x \in \mathbb{R}^d \times \mathcal{M}, \tag{2.5}$$

where the constant C satisfies $C = C(x, T, p) > 0$ and X^x denotes the solution to (1.1) *with initial condition $x \in \mathbb{R}^d$.*

Remark 2.4. We note that Assumptions (A1)–(A3) are slightly weaker then those in Anderson and Mattingly (2011), where it is assumed that $b, \sigma_k, k = 1, \dots, M$ are bounded with bounded and continuous partial derivatives up to the sixth order and that $\inf_{x\in\mathbb{R}^d}\{\sigma_k(x)\} > 0$ for each k. Also, (2.3) plays an important role in a certain Gaussian tail estimate in the proof of Lemma 2.10.

2.2 The Algorithm

The weak Simpson method can be summarized as follows. Let $T > 0$ and $\Pi := \{0 = t_0 < t_1 < ... < t_N = T\}$ be a subdivision of $[0, T]$. Let $\{\eta_{1k}^{(i)}, \eta_{2k}^{(i)} : i \in \mathbb{N}, k \in \{1, 2, \dots, M\}\}$ be a collection of mutually independent normal random variables with mean zero and variance 1. Fix $\theta \in (0, 1)$ and define

$$\alpha_1 = \frac{5}{12\theta(1-\theta)} \text{ and } \alpha_2 = \alpha_1 - 1 = \frac{5 - 12\theta + 12\theta^2}{12\theta(1-\theta)}. \tag{2.6}$$

In this work, we take constant discretization stepsize $h = T/N$ and so $t_i = ih$ for $i = 0, 1, \ldots, N$. Let $Y_0 = X(0) = x_0$ and, for $i = 1, 2, \ldots, N$, we repeat the following steps:

Step 1.

$$Y_i^* = Y_{i-1} + b(Y_{i-1})\theta h + \sum_{k=1}^{M} \sigma_k(Y_{i-1})\nu_k \eta_{1k}^{(i)} \sqrt{\theta h}. \tag{2.7}$$

Step 2.

$$\begin{aligned} Y_i =& Y_i^* + (\alpha_1 b(Y_i^*) - \alpha_2 b(Y_{i-1}))(1-\theta)h \\ &+ \sum_{k=1}^{M} \sqrt{[\alpha_1 \sigma_k^2(Y_i^*) - \alpha_2 \sigma_k^2(Y_{i-1})]^+}\ \nu_k\, \eta_{2k}^{(i)} \sqrt{(1-\theta)h}. \end{aligned} \tag{2.8}$$

We call such an algorithm the *weak Simpson method*. To motivate such a name, let us temporarily ignore the diffusion terms in (1.1). In addition, if we take $\theta = \frac{1}{2}$, then $\alpha_1 = \frac{5}{3}$, $\alpha_2 = \frac{2}{3}$, and (2.7) reduces to $Y_i^* = Y_{i-1} + \frac{h}{2} b(Y_{i-1})$. Next we insert it into (2.8) to obtain

$$Y_i = Y_{i-1} + \frac{h}{6}[b(Y_{i-1}) + 4b(Y_i^*) + b(Y_i^*)]. \tag{2.9}$$

On the other hand, the Simpson rule approximates the deterministic integral $\int_{t_{i-1}}^{t_i} b(Y(s))\,\mathrm{d}s$ by

$$\begin{aligned} \int_{t_{i-1}}^{t_i} b(Y(s))\,\mathrm{d}s &\approx \frac{t_i - t_{i-1}}{6}\left[b(Y(t_{i-1})) + 4b\left(Y\left(\frac{t_{i-1}+t_i}{2}\right)\right) + b(Y(t_i))\right] \\ &= \frac{h}{6}\left[b(Y(t_{i-1})) + 4b\left(Y\left(t_{i-1} + \frac{h}{2}\right)\right) + b(Y(t_i))\right]. \end{aligned}$$

Compare this with the second term of the right-hand side of (2.9), and notice that

$$Y_i^* = Y_{i-1} + \frac{h}{2} b(Y(t_{i-1})) \approx Y\left(t_{i-1} + \frac{h}{2}\right).$$

Therefore our algorithm (2.7)–(2.8) is similar to the deterministic Simpson rule, though we use the θ-midpoint value $b(Y_i^*)$ instead of the terminal value $b(Y(t_i))$ in (2.9). As a result, we have an explicit scheme.

To further illustrate the idea behind the algorithm (2.7)–(2.8), we note from Anderson and Mattingly (2011) that the solution of (1.1) is equivalent in distribution to

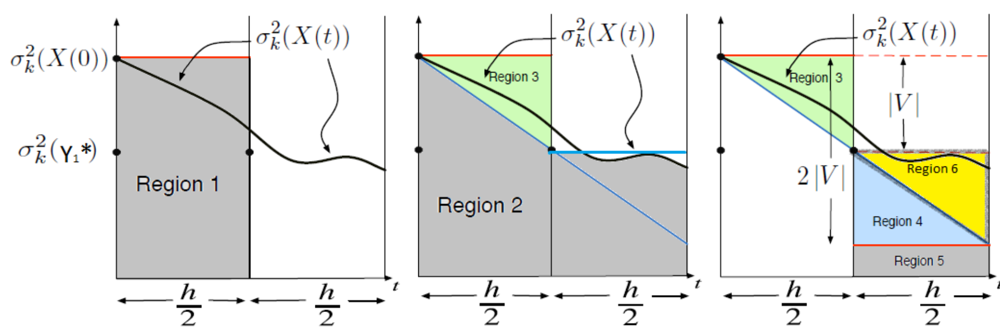
$$X(t) = x + \int_0^t b(X(s))\,\mathrm{d}s + \sum_{k=1}^{M} \nu_k \int_0^\infty \int_0^t I_{[0,\sigma_k^2(X(s))]}(u)\xi_k(\mathrm{d}u \times \,\mathrm{d}s), \qquad (2.10)$$

where ξ_k is a time-space white noise for each $k = 1, \dots, M$. Recall that an $\mathbb{F}$-adapted centered Gaussian random field $\{\xi_k(t,x)\}_{t\geq 0, x\in\mathbb{R}^+}$ is a time-space white noise if

$$\mathbb{E}[\xi(s,y)\xi(t,x)] = \delta(t-s)\delta(x-y), \qquad \text{for all } t,s \geq 0 \text{ and } x,y \in \mathbb{R}^+,$$

where $\delta(\cdot)$ is the delta function. In particular, it follows that if A, B are disjoint subsets of $[0,\infty)^2$, then $\xi_k(A)$ and $\xi_k(B)$ are independent normal random variables with means 0 and variances $|A|$ and $|B|$, respectively, where $|\cdot|$ denotes the Lebesgue measure on $[0,\infty)^2$. We refer to Walsh (1986) for introduction to space-time white noise and stochastic partial differential equations.

To illustrate the idea, let us fix $\theta = \frac{1}{2}$ and modify a figure from Anderson and Mattingly (2011) as in Figure 2.1.



(a) First step (b) Desired second step (c) Used second step

Figure 2.1: A graphical illustration of weak Simpson scheme for $\theta = \frac{1}{2}$, where $V = \sigma_k^2(X(0)) - \sigma_k^2(Y_1^*)$

For simplicity, we take $i = 1$ in (2.7) and (2.8). In order to approximate the diffusion term in (2.10) over the interval $[0,h]$, we must approximate $\xi_k(A_{[0,h]}(\sigma_k^2))$, where $A_{[0,h]}(\sigma_k^2)$ is the region under the curve $\sigma_k^2(X(t))$ for $0 \leq t \leq h$. Since ξ_k

is a space-time white noise, $\xi_k(A_{[0,h]}(\sigma_k^2))$ is normally distributed with mean zero and variance equals the area of the region $A_{[0,h]}(\sigma_k^2)$. Thus it is enough to find an accurate approximation to the area of the region $A_{[0,h]}(\sigma_k^2)$.

We would like to approximate the area under the curve $\sigma_k^2(X(t))$ using the Simpson rule. Recall that for a positive function $f(x)$, the Simpson rule to approximate $\int_a^b f(x)\,\mathrm{d}x$ which gives the area of the region under the curve $y = f(x)$ between a and b is given by

$$\begin{aligned} I(f) &= \frac{h}{6}\left[f(a) + 4f\Big(\frac{a+b}{2}\Big) + f(b)\right] \\ &= \frac{1}{3}\cdot\frac{h}{2}(f(a)+f(b)) + \frac{2}{3}\cdot hf\Big(\frac{a+b}{2}\Big) \\ &= \frac{1}{3}(\text{Area of } BCDF) + \frac{2}{3}(\text{Area of } ACDE), \end{aligned} \tag{2.11}$$

where we put $h = b - a$, $BCDF$ is the trapezoid with base $[a, b]$ and heights $f(a)$ and $f(b)$, and $ACDE$ is the rectangle with base $[a, b]$ and height $f(\frac{a+b}{2})$; see the illustration in Figure 2.2.

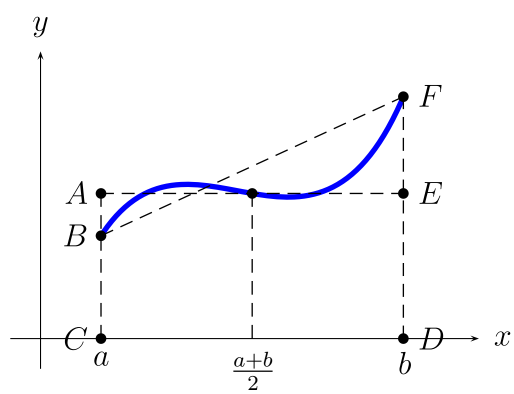


Figure 2.2: The Simpson Rule

To approximate the area of the region $A_{[0,h]}(\sigma_k^2)$, we first note that

$$\xi_k(\text{Region 1}) \overset{d}{=} N\Big(0, \sigma_k^2(X(0))\frac{h}{2}\Big) \overset{d}{=} \sigma_k(X(0))\sqrt{\frac{h}{2}}N(0,1),$$

which is equivalent in distribution to the summand of the right-hand side of (2.7) in Step 1 of our algorithm. Here and throughout the thesis, $N(\mu, \sigma^2)$ denotes a normal distribution with mean μ and variance σ^2.

Now using the estimated θ-midpoint Y_1^* from Step 1, we approximate the area under the curve $\sigma_k^2(X(t))$. Since we used the area of the green shaded region (Region 3 in Figure 2.1 (b)) in (2.7) of Step 1, we need to discard this area in Step 2. Observe that Region 3 and the blue shaded region (Region 4 in Figure 2.1 (c)) have equal areas. Thus in Step 2, we do not add the area of Region 4. This is equivalent to discarding the area of Region 3.

Since Region 5 in Figure 2.1 (c) is the part of both the rectangle and the trapezoid, we take the whole region under consideration.

$$\begin{aligned}\xi_k(\text{Region 5}) &\stackrel{d}{=} N\Big(0, \Big[\sigma_k^2(X(0)) - 2(\sigma_k^2(X(0)) - \sigma_k^2(Y_1^*))\Big]\frac{h}{2}\Big)\\ &\stackrel{d}{=} N\Big(0, \Big[2\sigma_k^2(Y_1^*) - \sigma_k^2(X(0))\Big]\frac{h}{2}\Big).\end{aligned}$$

On the other hand, Region 6 in Figure 2.1 (c) is part of the rectangle only, we take $\frac{2}{3}$ of the area of Region 6 only:

$$\begin{aligned}\xi_k\Big(\frac{2}{3}\text{Region 6}\Big) &\stackrel{d}{=} N\Big(0, \frac{2}{3}\cdot\frac{1}{2}\cdot(\sigma_k^2(X(0)) - \sigma_k^2(Y_1^*))\cdot\frac{h}{2}\Big)\\ &\stackrel{d}{=} N\Big(0, \frac{1}{3}(\sigma_k^2(X(0)) - \sigma_k^2(Y_1^*))\frac{h}{2}\Big).\end{aligned}$$

Note that Region 5 and Region 6 are disjoint. Thus we have

$$\begin{aligned}&\xi_k(\text{Region 5}) + \xi_k\Big(\frac{2}{3}\text{ Region 6}\Big)\\ &\quad\stackrel{d}{=} N\Big(0, \Big[2\sigma_k^2(Y_1^*) - \sigma_k^2(X(0))\Big]\frac{h}{2}\Big) + N\Big(0, \frac{1}{3}(\sigma_k^2(X(0)) - \sigma_k^2(Y_1^*))\frac{h}{2}\Big)\\ &\quad\stackrel{d}{=} N\Big(0, \Big(\frac{5}{3}\sigma_k^2(Y_1^*) - \frac{2}{3}\sigma_k^2(X(0))\Big)\frac{h}{2}\Big)\\ &\quad\stackrel{d}{=} \sqrt{\frac{5}{3}\sigma_k^2(Y_1^*) - \frac{2}{3}\sigma_k^2(X(0))}\sqrt{\frac{h}{2}}N(0,1).\end{aligned}$$

Note that for $\theta = \frac{1}{2}$ we have $\alpha_1 = \frac{5}{3}$ and $\alpha_2 = \frac{2}{3}$. Again, we find that $\xi_k(\text{Region 5}) + \xi_k(\frac{2}{3}\text{ Region 6})$ is equivalent in distribution to the summand of the right-hand side of (2.8) of Step 2. Therefore it is reasonable to anticipate that the weak Simpson algorithm shall work well. We note that the weak trapezoidal method in Anderson and Mattingly (2011) does not consider Region 6.

2.3 The Weak Convergence Rate

Define the Markov semigroup $\mathcal{P}_t : \mathcal{B}(\mathbb{R}^d) \to \mathcal{B}(\mathbb{R}^d)$ related to (1.1) by

$$(\mathcal{P}_t f)(x) :\overset{\text{def}}{=} \mathbb{E}_x[f(X(t))], t \geq 0 \tag{2.12}$$

where $X(0) = x$ and Markov semigroup $P_h : \mathcal{B}(\mathbb{R}^d) \to \mathcal{B}(\mathbb{R}^d)$ associated with single full step size h of the weak Simpson method (2.7)–(2.8) by

$$(P_h f)(y) :\overset{\text{def}}{=} \mathbb{E}_y[f(Y_1)], \tag{2.13}$$

where $Y_0 = y$. Since $\|f\|_0 = \sup\{|f(x)| : x \in \mathbb{R}^d\}$, it follows that $\|\mathcal{P}_t f\|_0 \leq \|f\|_0$ and similarly $\|P_h f\|_0 \leq \|f\|_0$.

The following Proposition is from Anderson and Mattingly (2011).

Proposition 2.5. *If $b, \sigma_1, \ldots, \sigma_M \in C^k$, then for any $0 < t \leq T$ and $k \in \mathbb{N}$, there exists a constant $C = C(T, k, b, \sigma) > 0$ such that $\|\mathcal{P}_t f\|_k \leq C\|f\|_k$.*

As in Anderson and Mattingly (2011), we define the induced operator norm for any linear operator $L : C^k \to C^l$ by

$$\|L\|_{k\to l} = \sup_{f\in C^k, f\neq 0} \frac{\|Lf\|_l}{\|f\|_k}.$$

Then it follows from Proposition 2.5 that $\|\mathcal{P}_t\|_{k\to k} \leq C$. In particular, we have

$$\|\mathcal{P}_t\|_{0\to 0} = \sup_{f\in C^0, f\neq 0} \frac{\|\mathcal{P}_t f\|_0}{\|f\|_0} \leq \sup_{f\in C^0, f\neq 0} \frac{\|f\|_0}{\|f\|_0} = 1.$$

Similarly, $\|P_h\|_{0\to 0} \leq 1$.

In terms of the induced operator norm, (2.1) can be rewritten equivalently as

$$\left\|\mathcal{P}_T - P_h^N\right\|_{k\to 0} \leq Ch^\beta. \tag{2.14}$$

The following theorems give respectively the weak local and global convergence rate of the weak Simpson method (2.7)–(2.8) for general case.

Theorem 2.6 (Local Approximation General Case)**.** *Assume* (A1)–(A3)*. Then there exist a constant κ so that*

$$\|\mathcal{P}_h - P_h\|_{4\to 0} \le \kappa h^2, \quad \textit{for all } h > 0 \textit{ sufficiently small.}$$

The proof of Theorem 2.6 depends on the following lemma, whose proof is given in Appendix A.

Lemma 2.7. *Assume* (A1)–(A3)*. Then for all $h > 0$ sufficiently small and $f \in C^4$, we have*

$$\mathbb{E}\left[f(Y_1^*) + (Bf)(Y_1^*)(1-\theta)h\right] = f(x_0) + (Af)(x_0)h + O(h^2), \tag{2.15}$$

where

$$(Af)(x) = f'[b(x)](x) + \frac{1}{2}\sum_{k=1}^{M} \sigma_k^2(x) f''[\nu_k, \nu_k](x), \tag{2.16}$$

$$(Bf)(x) = f'[\alpha_1 b(Y_1^*) - \alpha_2 b(x_0)](x) + \frac{1}{2}\sum_{k=1}^{M}[\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)]^+ f''[\nu_k, \nu_k](x). \tag{2.17}$$

Proof of Theorem 2.6. We need to show that for any $f \in C^4$ and $h > 0$ sufficiently small, there exists a constant $\kappa > 0$ so that

$$|\mathbb{E}_{x_0}[f(Y_1)] - \mathbb{E}_{x_0}[f(X(h))]| \le \kappa \|f\|_4 h^2, \quad \text{for all } x_0 \in \mathbb{R}^d. \tag{2.18}$$

To this end, we consider the stochastic differential equation

$$\mathrm{d}y(t) = b(x_0)\,\mathrm{d}t + \sum_{k=1}^{M} \sigma_k(x_0)\nu_k\,\mathrm{d}W_k(t), \quad t \ge 0, \quad y(0) = x_0. \tag{2.19}$$

Then we have

$$y(\theta h) = x_0 + b(x_0)\theta h + \sum_{k=1}^{M} \sigma_k(x_0)\nu_k(W_k(\theta h) - W_k(0)). \tag{2.20}$$

Since $W_k(\theta h) - W_k(0) \stackrel{d}{=} N(0, \theta h) \stackrel{d}{=} N(0,1)\sqrt{\theta h}$, we see that (2.7) in Step 1 of the weak Simpson algorithm produces a value Y_1^* which is equal to $y(\theta h)$ in distribution.

Let B_1 is the infinitesimal generator of (2.19). That is, for all f sufficiently smooth, we define

$$(B_1 f)(x) = f'[b(x_0)](x) + \frac{1}{2}\sum_{k=1}^{M} \sigma_k(x_0)^2 f''[\nu_k, \nu_k](x). \tag{2.21}$$

Similarly, (2.8) in Step 2 produces a value Y_1 that is equivalent to $z(h)$ in distribution, where $z(t)$ solves the stochastic differential equation

$$\begin{cases} \mathrm{d}z(t) = (\alpha_1 b(Y_1^*) - \alpha_2 b(x_0))\,\mathrm{d}t + \displaystyle\sum_{k=1}^{M} \sqrt{[\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)]^+}\nu_k\,\mathrm{d}W_k(t), t \geq \theta h, \\ z(\theta h) = Y_1^*. \end{cases} \tag{2.22}$$

Recall the definitions of the operators A and B in (2.16) and (2.17), respectively. Note that A is the infinitesimal generator of (1.1), and B the infinitesimal generator for the process (2.22).

Let $\mathcal{F}_t$ denote the filtration generated by the Brownian motion processes $W_k(t)$ in (2.22). Then

$$\mathbb{E}[f(z(h))] = \mathbb{E}[\mathbb{E}[f(z(h))|\mathcal{F}_{\theta h}]] \stackrel{\text{def}}{=} \mathbb{E}[\mathbb{E}_{\theta h}[f(z(h))]], \tag{2.23}$$

where we defined $\mathbb{E}_{\theta h}[\cdot] \stackrel{\text{def}}{=} \mathbb{E}[\cdot|\mathcal{F}_{\theta h}]$. Let $z(h)$ be the solution to (2.22). Since $f \in C^4$, we can use Dynkin's formula repeatedly to obtain

$$\begin{aligned} &\mathbb{E}_{\theta h}[f(z(h))] \\ &\quad = f(Y_1^*) + \int_{\theta h}^{h} \mathbb{E}_{\theta h}[(Bf)(z(s))]\,\mathrm{d}s \\ &\quad = f(Y_1^*) + (Bf)(Y_1^*)(1-\theta)h + \int_{\theta h}^{h}\int_{\theta h}^{s} \mathbb{E}_{\theta h}[(B^2 f)(z(r))]\,\mathrm{d}r\,\mathrm{d}s. \end{aligned} \tag{2.24}$$

The term $(B^2 f)(z(r))$ in the last integral above depends only on the first four derivatives of f. Therefore, using the fact that $f \in C^4$ and Lemma 2.3, we obtain

$$\left| \int_{\theta h}^{h}\int_{\theta h}^{s} \mathbb{E}_{\theta h}[(B^2 f)(z(r))]\,\mathrm{d}r\,\mathrm{d}s \right| \leq C_1 \|f\|_4 h^2 \tag{2.25}$$

for some constant C_1 independent of h.

Recall that Y_1 of (2.8) and $z(h)$ of (2.22) have the same distribution and, in particular, we have $\mathbb{E}_{x_0}[f(Y_1)] = \mathbb{E}_{x_0}[f(z(h))]$. Then it follows from (2.23), (2.24), and (2.25) that

$$\begin{aligned}\mathbb{E}_{x_0}[f(Y_1)] &= \mathbb{E}_{x_0}[\mathbb{E}_{\theta h}[f(z(h))]] \\ &= \mathbb{E}_{x_0}\left[f(Y_1^*) + (Bf)(Y_1^*)(1-\theta)h + O(h^2)\right].\end{aligned}$$

On the other hand, proceeding as above and applying Dynkin's formula to (1.1) repeatedly gives

$$\mathbb{E}_{x_0}[f(X(h))] = f(x_0) + (Af)(x_0)h + O(h^2).$$

Then (2.18) follows from Lemma 2.7 and the above two displayed equations. This completes the proof of the theorem. □

Theorem 2.8 (Global Approximation General Case)**.** *Assume* (A1)–(A3)*. Then for any* $T > 0$ *there exists a constant* $C(T) > 0$ *such that*

$$\sup_{0\le nh\le T} ||\mathcal{P}_{nh} - P_h^n||_{4\to 0} \le C(T)h. \tag{2.26}$$

Proof. Let us first observe the following nested sum:

$$\begin{aligned}P_h^n - \mathcal{P}_{nh} &= P_h^1\mathcal{P}_{h(n-1)} - P_h^0\mathcal{P}_{nh} \\ &\quad + P_h^2\mathcal{P}_{h(n-2)} - P_h^1\mathcal{P}_{h(n-1)} + P_h^3\mathcal{P}_{h(n-3)} - P_h^2\mathcal{P}_{h(n-2)} \\ &\quad + \cdots + P_h^{n-1}\mathcal{P}_{h(n-(n-1))} - P_h^{n-2}\mathcal{P}_{h(n-(n-2))} \\ &\quad + P_h^n\mathcal{P}_{h(n-n)} - P_h^{n-1}\mathcal{P}_{h(n-(n-1))} \\ &= \sum_{k=1}^{n} P_h^{k-1}(P_h - \mathcal{P}_h)\mathcal{P}_{h(n-k)}.\end{aligned}$$

Thus it follows that

$$\begin{aligned}||\mathcal{P}_{nh} - P_h^n||_{4\to 0} &= ||\sum_{k=1}^{n} P_h^{k-1}(P_h - \mathcal{P}_h)\mathcal{P}_{h(n-k)}||_{4\to 0} \\ &\le \sum_{k=1}^{n} ||P_h^{k-1}(P_h - \mathcal{P}_h)\mathcal{P}_{h(n-k)}||_{4\to 0} \\ &= \sum_{k=1}^{n} ||P_h^{k-1}||_{0\to 0}||(P_h - \mathcal{P}_h)||_{4\to 0}||\mathcal{P}_{h(n-k)}||_{4\to 4}.\end{aligned}$$

Using the fact that $\sup_{0\le s\le T}\|\mathcal{P}_s\|_{4\to 4}\le C'(T)$ and $\|P_h^k\|_{0\to 0}\le 1$ and using theorem 2.6 for any n with $0\le nh\le T$,

$$\|\mathcal{P}_{nh}-P_h^n\|_{4\to 0}\le \sum_{k=1}^{n} C'(T)Kh^2 = nC'(T)Kh^2\le KTC'(T)h = C(T)h,$$

where C' and C are positive constants independent of h. This completes the proof of the theorem. □

Next we state the following assumption, which is necessary to show that the weak Simpson method (2.7)–(2.8) has weak convergence order three.

(A4) For any sufficiently smooth functions $f:\mathbb{R}^d\mapsto\mathbb{R}$, we have

$$\begin{aligned}&(A^2f)(x_0)=(B_1^2f)(x_0),\ \text{ and}\\ &(A^3f)(x_0)=(B_1^2(Af))(x_0)=(B_1g)(x_0)=(B_1(A(B_1f)))(x_0)=(B_1^3f)(x_0),\end{aligned} \tag{2.27}$$

for all $x_0\in\mathbb{R}$, where

$$\begin{aligned} g(x) &= f''[b(x),b(x)](x)+\sum_{k=1}^{M}\sigma_k^2(x)f'''[\nu_k,\nu_k,b(x)](x)\\ &\quad+\frac{1}{4}\sum_{k,j=1}^{M}\sigma_j^2(x)\sigma_k^2(x)f^{(4)}[\nu_k,\nu_k,\nu_j,\nu_j](x),\end{aligned}$$

and we define $(A^nf)(x)=(A(A^{n-1}f))(x)$ for any integer $n\ge 2$, and similarly for B_1 and B.

The following theorems give respectively the weak local and global convergence rate of the weak Simpson method (2.7)–(2.8) under Assumptions (A1)–(A4).

Theorem 2.9 (Local Approximation)**.** *Assume* (A1)–(A4)*. Then there exists a constant κ_1 so that*

$$\|\mathcal{P}_h-P_h\|_{8\to 0}\le\kappa_1h^4,\qquad \textit{for all } h>0 \textit{ sufficiently small.}$$

The proof of Theorem 2.9 depends on the following lemma, whose proof is relegated to Appendix A.

Lemma 2.10. *Assume* (A1)–(A4). *Then for all $h > 0$ sufficiently small and $f \in C^8$, we have*

$$\begin{aligned}
&\mathbb{E}\left[f(Y_1^*) + (Bf)(Y_1^*)(1-\theta)h + (B^2 f)(Y_1^*)\frac{(1-\theta)^2 h^2}{2} + (B^3 f)(Y_1^*)\frac{(1-\theta)^3 h^3}{6}\right] \\
&\quad = f(x_0) + (Af)(x_0)h + (A^2 f)(x_0)\frac{h^2}{2} + (A^3 f)(x_0)\frac{h^3}{6} + O(h^4),
\end{aligned} \tag{2.28}$$

Proof of Theorem 2.9. We need to show that for any $f \in C^8$ and $h > 0$ sufficiently small, there exists a constant $\kappa_1 > 0$ so that

$$|\mathbb{E}_{x_0}[f(Y_1)] - \mathbb{E}_{x_0}[f(X(h))]| \le \kappa_1 \|f\|_8 h^4, \qquad \text{for all } x_0 \in \mathbb{R}^d. \tag{2.29}$$

Since $f \in C^8$, we can use Dynkin's formula repeatedly to obtain

$$\begin{aligned}
&\mathbb{E}_{\theta h}[f(z(h))] \\
&\quad = f(Y_1^*) + \int_{\theta h}^{h} \mathbb{E}_{\theta h}[(Bf)(z(s))]\,\mathrm{d}s \\
&\quad = f(Y_1^*) + (Bf)(Y_1^*)(1-\theta)h + \int_{\theta h}^{h}\int_{\theta h}^{s} \mathbb{E}_{\theta h}[(B^2 f)(z(r))]\,\mathrm{d}r\,\mathrm{d}s \\
&\quad = f(Y_1^*) + (Bf)(Y_1^*)(1-\theta)h + (B^2 f)(Y_1^*)\frac{(1-\theta)^2 h^2}{2} \\
&\qquad + \int_{\theta h}^{h}\int_{\theta h}^{s}\int_{\theta h}^{r} \mathbb{E}_{\theta h}[(B^3 f)(z(u))]\,\mathrm{d}u\,\mathrm{d}r\,\mathrm{d}s \\
&\quad = f(Y_1^*) + (Bf)(Y_1^*)(1-\theta)h + (B^2 f)(Y_1^*)\frac{(1-\theta)^2 h^2}{2} \\
&\qquad + (B^3 f)(Y_1^*)\frac{(1-\theta)^3 h^3}{6} + \int_{\theta h}^{h}\int_{\theta h}^{s}\int_{\theta h}^{r}\int_{\theta h}^{v} \mathbb{E}_{\theta h}[(B^4 f)(z(w))]\,\mathrm{d}w\,\mathrm{d}u\,\mathrm{d}r\,\mathrm{d}s.
\end{aligned} \tag{2.30}$$

The term $(B^4 f)(z(w))$ in the last integral above depends only on the first eight derivatives of f. Therefore, using the fact that $f \in C^8$ and Lemma 2.3, we obtain

$$\left|\int_{\theta h}^{h}\int_{\theta h}^{s}\int_{\theta h}^{r}\int_{\theta h}^{v} \mathbb{E}_{\theta h}[(B^4 f)(z(w))]\,\mathrm{d}w\,\mathrm{d}u\,\mathrm{d}r\,\mathrm{d}s\right| \le C_1' \|f\|_8 h^4 \tag{2.31}$$

for some constant C_1' independent of h.

Recall that Y_1 of (2.8) and $z(h)$ of (2.22) have the same distribution and, in particular, we have $\mathbb{E}_{x_0}[f(Y_1)] = \mathbb{E}_{x_0}[f(z(h))]$. Then it follows from (2.23), (2.30), and (2.31) that

$$\begin{aligned}\mathbb{E}_{x_0}[f(Y_1)] &= \mathbb{E}_{x_0}[\mathbb{E}_{\theta h}[f(z(h))]] \\ &= \mathbb{E}_{x_0}\Bigg[f(Y_1^*) + (Bf)(Y_1^*)(1-\theta)h + (B^2 f)(Y_1^*)\frac{(1-\theta)^2h^2}{2} \\ &\qquad\qquad + (B^3 f)(Y_1^*)\frac{(1-\theta)^3h^3}{6} + O(h^4)\Bigg].\end{aligned}$$

On the other hand, proceeding as above and applying Dynkin's formula to (1.1) repeatedly gives

$$\mathbb{E}_{x_0}[f(X(h))] = f(x_0) + (Af)(x_0)h + (A^2f)(x_0)\frac{h^2}{2} + (A^3f)(x_0)\frac{h^3}{6} + O(h^4).$$

Then (2.29) follows from Lemma 2.10 and the above two displayed equations. This completes the proof of the theorem. □

With Theorem 2.9 in our hands, we can proceed to derive the global weak convergence rate for the weak Simpson method (2.7)–(2.8).

Theorem 2.11. *Assume* (A1)–(A4). *Then for any* $T > 0$ *there exists a constant* $C''(T) > 0$ *such that*

$$\sup_{0\le nh\le T} ||\mathcal{P}_{nh} - P_h^n||_{8\to 0} \le C''(T)h^3. \tag{2.32}$$

The proof of Theorem 2.11 is similar to that of Theorem 2.8. We shall omit the details here.

Chapter 3

Examples

To illustrate the main results in Section 2.3, we present three examples in this section. We start with the one-dimensional geometric Brownian motion (3.1) in Example 3.1, so that we can compare the numerical computations using the weak Simpson method (2.7)–(2.8) with the theoretical values. Next we consider two nonlinear two-dimensional SDEs which are investigated in Anderson and Mattingly (2011). In particular, we want to compare the performance of the weak Simpson method with that of the weak trapezoidal method proposed in Anderson and Mattingly (2011).

Example 3.1. We consider the one-dimensional geometric Brownian motion in the following form

$$\begin{aligned} &\mathrm{d}X(t) = \lambda X(t)\,\mathrm{d}t + \mu X(t)\,\mathrm{d}W(t), \\ &X(0) = X(0) \in \mathbb{R}, \end{aligned} \tag{3.1}$$

where λ and μ are real constants. The solution to (3.1) is $X(t) = X(0)e^{(\lambda-\frac{1}{2}\mu^2)t+\mu W(t)}$ and we have

$$\mathbb{E}[X(t)] = \mathbb{E}[X(0)]e^{\lambda t}, \quad \mathbb{E}[(X(t)^2] = \mathbb{E}[X(0)^2]e^{(2\lambda+\mu^2)t}.$$

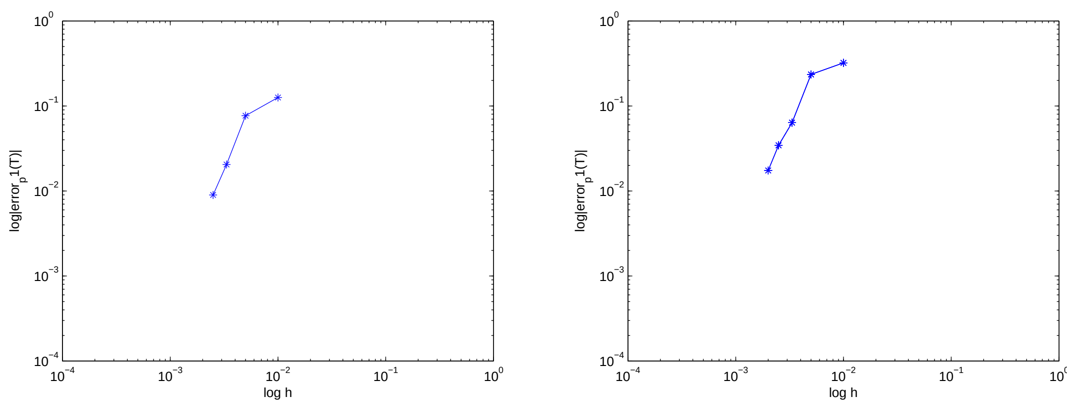
We test the performance of the weak Simpson method (2.7)–(2.8) for (3.1). We take $T = 1$, $\lambda = 3$, $\mu = 0.3$, and $X(0) = 1$ in (3.1). Also, we use stepsizes $h_p = \frac{1}{100p}$, for $p = 1, \ldots, 4$ to generate $N = 48,200$ sample paths of (3.1) using the the weak Simpson algorithm (2.7)–(2.8). We then compute

$$\text{Error}_p^1(1) = \mathbb{E}[X(1)] - \frac{1}{N}\sum_{k=1}^{N} Y_{N_p}^{k,h_p}. \tag{3.2}$$

Again, we test the performance of the weak Simpson method (2.7)–(2.8) for (3.1). We take $T = 1$, $\lambda = 2$, $\mu = 0.1$, and $X(0) = 1$ in (3.1). Also, we use step sizes $h_p = \frac{1}{100p}$, for $p = 1, \dots, 5$ to generate $N = 45000$ sample paths of (3.1) using the the weak Simpson algorithm (2.7)–(2.8). We then compute

$$\mathrm{Error}_p^2(1) = \mathbb{E}[X(1)^2] - \frac{1}{N}\sum_{i=1}^{N}\left(Y_{N_p}^{k,h_p}\right)^2. \tag{3.3}$$

where $N_p = \frac{1}{h_p}$, and $Y_{N_p}^{k,h_p}$ is the value obtained using the weak Simpson method (2.7)–(2.8) in the kth simulation with step size h_p. The results are plotted in Figure 3.1 in the log-log scale. Note that in both plots of Figure 3.1, the best fit lines have slope 3 as the step size h tends to zero, as expected.



(a) $\mathrm{Error}_p^1(1)$ of (3.2) (b) $\mathrm{Error}_p^2(1)$ of (3.3)

Figure 3.1: Log-Log plots of error versus step-size for Example 3.1

The stepsizes and corresponding errors for (3.2) are shown in TABLE 3.1. The slopes of the line joining the points $(-2.30, -1.115)$ and $(-2.48, -1.69)$, $(-2.48, -1.69)$ and $(-2.60, -2.05)$ are 3.19 and 3 respectively; see the fifth column of Table 3.1. Hence we observe that the weak order of convergence of our algorithm is empirically three.

Again, the stepsizes and corresponding errors for (3.3) are shown in TABLE 3.2. The slopes of the line joining the points $(-2.3010, -.6293)$ and $(-2.4771, -1.1972)$,

Table 3.1: Log-log table for stepsizes and errors.

step-size	error	log(step-size)	log(error)	slope
0.0100	0.1256	-2.00	-0.901	
				0.713
0.0050	0.0767	-2.30	-1.115	
				3.194
0.0033	0.0205	-2.48	-1.69	
				3.000
0.0025	0.0090	-2.60	-2.05	

Table 3.2: Log-log table for stepsizes and errors.

step-size	error	log(stepsize)	log(error)	slope
0.0100	0.3212	-2.0000	-0.4932	
				0.452
0.0500	0.2348	-2.3010	-0.6293	
				3.22
0.0033	0.0635	-2.4771	-1.1972	
				2.12
0.0025	0.0345	-2.6021	-1.4622	
				3.06
0.0020	0.0174	-2.6990	-1.7595	

$(-2.4771, -1.1972)$ and $(-2.6021, -1.4622)$, $(-2.6021, -1.4622)$ and $(-2.6990, -1.7595)$ are $3.22, 2.12$ and 3.06 respectively. Hence we observe that the weak order of convergence of our algorithm is empirically three.

Example 3.2. Here we investigate the same example considered in Anderson and Mattingly (2011).

$$\begin{bmatrix} \mathrm{d}X_1(t) \\ \mathrm{d}X_2(t) \end{bmatrix} = \begin{bmatrix} X_1(t) \\ 0 \end{bmatrix} \mathrm{d}t + X_1(t) \begin{bmatrix} 0 \\ 1 \end{bmatrix} \mathrm{d}W_1(t) + \frac{1}{10} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \mathrm{d}W_2(t), \tag{3.4}$$

where $W_1(t)$ and $W_2(t)$ are independent standard one-dimensional Brownian motion processes. In our system of SDE $b_1(x) = x_1, b_2(x) = 0, \sigma_1(x) = x_1, \sigma_2(x) = \frac{1}{10}$, $\nu_1 = [0, 1]^T$ and $\nu_2 = [1, 1]^T$. The system (3.4) can be rewritten in component wise as

$$\mathrm{d}X_1(t) = X_1(t)\,\mathrm{d}t + \frac{1}{10}\,\mathrm{d}W_2(t), \tag{3.5}$$

$$\mathrm{d}X_2(t) = X_1(t)\,\mathrm{d}W_1(t) + \frac{1}{10}\,\mathrm{d}W_2(t). \tag{3.6}$$

To solve (3.5), let $f(t, x) = xe^{-t}$. Using Itô's formula,

$$\mathrm{d}\big(X_1(t)e^{-t}\big) = -X_1(t)e^{-t}\,\mathrm{d}t + e^{-t}\,\mathrm{d}X_1(t) = \frac{1}{10}e^{-t}\,\mathrm{d}W_2(t). \tag{3.7}$$

Thus it follows that

$$X_1(t) = X_1(0)e^t + \frac{1}{10}\int_0^t e^{(t-s)}\,\mathrm{d}W_2(s).$$

Then we have

$$X_1(t)^2 = X_1(0)^2 e^{2t} + \frac{1}{100}\left(\int_0^t e^{(t-s)}\,\mathrm{d}W_2(s)\right)^2 + \frac{1}{5}X_1(0)\int_0^t e^{(2t-s)}\,\mathrm{d}W_2(s),$$

from which we deduce

$$\begin{aligned}\mathbb{E}[X_1(t)^2] &= \mathbb{E}[X_1(0)^2]e^{2t} + \frac{1}{100}\mathbb{E}\int_0^t e^{2(t-s)}\,\mathrm{d}s \\ &= \mathbb{E}[X_1(0)^2]e^{2t} + \frac{1}{200}e^{2t} - \frac{1}{200}. \end{aligned} \tag{3.8}$$

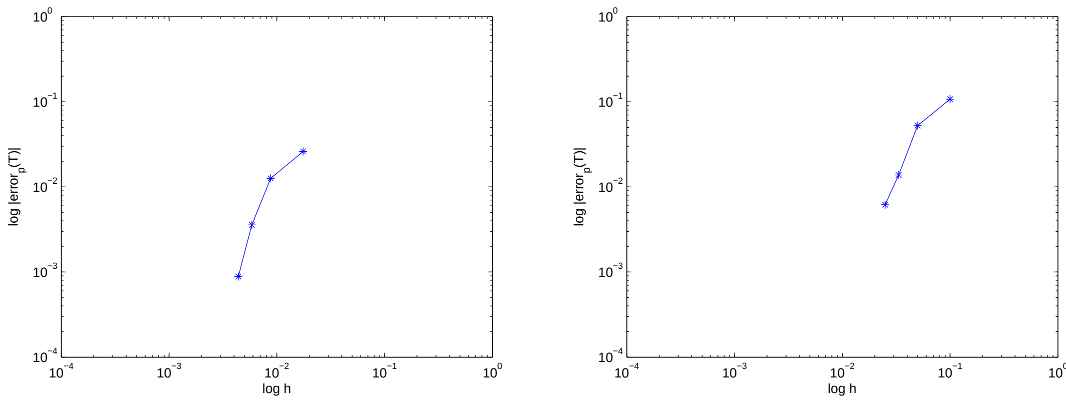
We compute 50,000 different discretized Brownian paths with step size $h_p = \frac{1}{55p}$ for $1 \le p \le 4$ and initial condition $X_1(0) = X_2(0) = 1$. We computed

$$\mathrm{Error}_p(1) = \mathbb{E}[X_1(1)^2] - \frac{1}{5\times 10^4}\sum_{k=1}^{5\times 10^4}\left(Y_{1,N_p}^{k,h_p}\right)^2, \tag{3.9}$$

where Y_{1,N_p}^{k,h_p} is the numerically simulated sample path and $\mathbb{E}[X_1(t)^2]$ is from (3.8). The resulting error is displayed in Figure 3.2(a) where we have plotted the weak error against h on log-log scale. The stepsizes and corresponding errors for (3.9) are shown in TABLE 3.3. The slopes of the line joining the points $(-2.0414, -1.8996)$ and $(-2.2175, -2.4437), (-2.2175, -2.4437)$ and $(-2.3424, -3.0458)$ are 3.08 and 4.82 respectively. Hence it is observed that the slope of the best fit line is empirically three.

Table 3.3: Log-log table for step-sizes and errors.

step-size	error	log(step-size)	log(error)	slope
0.0182	0.0261	-1.7404	-1.5834	
				1.05
0.0091	0.0126	-2.0414	-1.8996	
				3.08
0.0061	0.0036	-2.2175	-2.4437	
				4.82
0.0045	0.0009	-2.3424	-3.0458	



(a) $\text{Error}_p(1)$ of (3.9) in Example 3.2 (b) $\text{Error}_p(2)$ of (3.11) in Example 3.2

Figure 3.2: Log-Log plots of error versus step-size for Examples 3.2

Next we solve (3.6). Consider the function $f(t,x) = x^2$. Using Itô's formula, we have,

$$\begin{aligned} \mathrm{d}X_2(t)^2 &= 2X_2(t)\,\mathrm{d}X_2(t) + \mathrm{d}X_2(t)\cdot\mathrm{d}X_2(t) \\ &= 2X_1(t)X_2(t)\,\mathrm{d}W_1(t) + \frac{1}{5}X_2(t)\,\mathrm{d}W_2(t) + X_1(t)^2\,\mathrm{d}t + \frac{1}{100}\,\mathrm{d}t \end{aligned}$$

Thus

$$\begin{aligned} X_2(t)^2 = X_2(0)^2 &+ \int_0^t 2X_1(s)X_2(s)\,\mathrm{d}W_1(s) \\ &+ \frac{1}{5}\int_0^t X_2(s)\,\mathrm{d}W_2(s) + \int_0^t \big(X_1(s)^2 + \frac{1}{100}\big)\,\mathrm{d}s. \end{aligned}$$

Note that

$$\begin{aligned} \mathbb{E}\Big[\Big(\int_0^t X_1(s)X_2(s)\,\mathrm{d}W_1(s)\Big)^2\Big] &= \mathbb{E}\Big[\int_0^t X_1(s)^2X_2(s)^2\,\mathrm{d}s\Big] \\ &\le \mathbb{E}\Big[\int_0^t \frac{1}{2}\left(X_1(s)^4 + X_2(s)^4\right)\,\mathrm{d}s\Big] < \infty, \end{aligned}$$

and

$$\mathbb{E}\Big[\Big(\int_0^t X_2(s)\,\mathrm{d}W_2(s)\Big)^2\Big] = \mathbb{E}\Big[\int_0^t X_2(s)^2\,\mathrm{d}s\Big] < \infty.$$

Therefore, the expected value of the stochastic integrals $\int_0^t X_1(s)X_2(s)\,\mathrm{d}W_1(s)$ and $\int_0^t X_2(s)\,\mathrm{d}W_2(s)$ are zero. Hence

$$\mathbb{E}[X_2(t)^2] = \mathbb{E}[X_2(0)^2] + \int_0^t \mathbb{E}[X_1(s)^2]\,\mathrm{d}s + \int_0^t \left(\frac{1}{100}\right)\mathrm{d}s.$$

Furthermore, using (3.8), we arrive that

$$\mathbb{E}[X_2(t)^2] = \mathbb{E}[X_2(0)^2] - \frac{1}{2}\mathbb{E}[X_1(0)^2] + \frac{1}{400}e^{2t}(200\mathbb{E}[X_1(0)^2]+1) + \frac{t}{200} - \frac{1}{400}. \tag{3.10}$$

Next we use (3.10) to compute the error

$$\mathrm{Error}_p(2) = \mathbb{E}[X_2(1)^2] - \frac{1}{N}\sum_{k=1}^{N}\left(Y_{2,N_p}^{k,h_p}\right)^2, \tag{3.11}$$

by generating $N = 46,300$ sample paths of (3.4), where $h_p = \frac{1}{10p}$ for $p = 1,\ldots,4$, $N_p = \frac{1}{h_p}$, and $(Y_{1,N_p}^{k,h_p}, Y_{2,N_p}^{k,h_p})' \in \mathbb{R}^2$ is the approximated value of (3.4) obtained using the weak Simpson method (2.7)–(2.8) in the kth simulation, for $1 \le k \le N$. The resulting error is displayed in Figure 3.2 (b), where we have plotted the weak error against h_p on log-log scale. We observe that the slope of the best fit line in Figure 3.2 (b) is three as the the step size h tends to zero. The stepsizes and corresponding errors for (3.11) are shown in TABLE 3.4. The slopes of the line joining the points $(-1.301, -1.279)$ and $(-1.477, -1.86)$, $(-1.477, -1.86)$ and $(-1.602, -2.208)$ are 3.30 and 2.784 respectively. Hence it is observed that the slope of the best fit line is empirically three.

Table 3.4: Log-log table for step-sizes and errors.

step-size	error	log(step-size)	log(error)	slope
0.0100	0.1075	-1.000	-0.969	1.03
0.0050	0.0526	-1.301	-1.279	3.30
0.0033	0.0138	-1.477	-1.860	2.78
0.0025	0.0062	-1.602	-2.208	

Using (3.8) and (3.10), we have

$$\mathbb{E}[X(t)^2] = \frac{1}{2}\mathbb{E}[X_1(0)^2](3e^{2t}-1) + \mathbb{E}[X_2(0)^2] + \frac{1}{400}(3e^{2t}+2t-3). \tag{3.12}$$

We use (3.12) to compute the error

$$\text{Error}_p(1) = \mathbb{E}[X(1)^2] - \frac{1}{N}\sum_{k=1}^{N}\left(Y_{N_p}^{k,h_p}\right)^2, \tag{3.13}$$

by generating $N = 46,000$ sample paths of (3.4), where $h_p = \frac{1}{25p}$ for $p = 1, \ldots, 4$, $N_p = \frac{1}{h_p}$, and $(Y_{N_p}^{k,h_p})' \in \mathbb{R}^2$ is the approximated value of (3.4) obtained using the weak Simpson method (2.7)–(2.8) in the kth simulation, for $1 \leq k \leq N$. The resulting error is displayed in Figure 3.3(a), where we have plotted the weak error against h_p on log-log scale. We observe that the slope of the best fit line in Figure 3.3(a) is empirically three as the the step size h tends to zero. The stepsizes and corresponding errors for (3.13) are shown in TABLE 3.5 . The slopes of the

Table 3.5: Log-log table for step-sizes and errors.

step-size	error	log(step-size)	log(error)	slope
0.040	0.1307	-1.398	-0.884	
0.020	0.0405	-1.699	-1.393	1.69
0.013	0.0084	-1.875	-2.076	3.88
0.010	0.0011	-2.000	-2.959	7.06

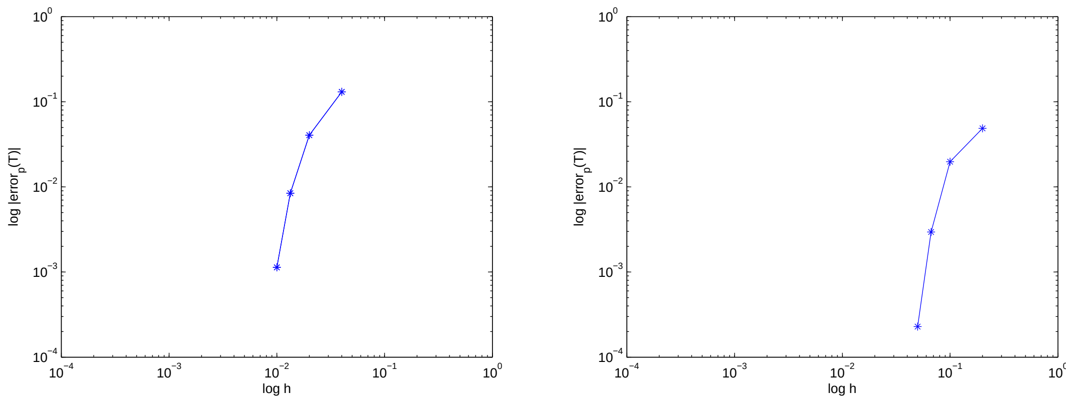
line joining the points $(-1.398, -0.884)$ and $(-1.699, -1.393), (-1.699, -1.393)$ and $(-1.875, -2.076), (-1.875, -2.076)$ and $(-2.000, -2.959)$ are $1.69, 3.88$ and 7.06 respectively. Hence we observe that the weak order of convergence of our algorithm is empirically three.

Example 3.3. Here we consider the system

$$\begin{aligned}\begin{bmatrix}\mathrm{d}X_1(t)\\ \mathrm{d}X_2(t)\end{bmatrix} = \begin{bmatrix}-X_2(t)\\ X_1(t)\end{bmatrix}\mathrm{d}t + \sqrt{\frac{\sin^2(X_1(t)+X_2(t))}{t+1}}\begin{bmatrix}1\\0\end{bmatrix}\mathrm{d}W_1(t)\\ + \sqrt{\frac{\cos^2(X_1(t)+X_2(t))}{t+1}}\begin{bmatrix}0\\1\end{bmatrix}\mathrm{d}W_2(t).\end{aligned} \tag{3.14}$$

where $W_1(t)$ and $W_2(t)$ are independent standard one-dimensional Brownian motions. The system (3.14) is similar to the one considered in Anderson and Mattingly (2011). Then $\mathbb{E}[|X(t)|^2]$ can be calculated as

$$\mathbb{E}[|X(t)|^2] = \mathbb{E}[X(0)^2] + \log(1+t). \tag{3.15}$$



(a) $\mathrm{Error}_p(1)$ of (3.13) in Example 3.2 (b) $\mathrm{Error}_p(1)$ of (3.16) in Example 3.3

Figure 3.3: Log-Log plots of error versus step-size for Examples 3.2 and 3.3

We generate $N = 51,000$ different sample paths of (3.14) using the weak Simpson algorithm (2.7)–(2.8) with step sizes $h = \frac{1}{5p}$ for $1 \leq p \leq 4$ and initial condition $X(0) = (1,1)'$. We then compute

$$\mathrm{Error}_p(1) = \mathbb{E}[|X(1)|^2] - \frac{1}{N}\sum_{k=1}^{N}\left|Y_{N_p}^{k,h_p}\right|^2, \tag{3.16}$$

where $N_p = 1/h_p$, for $1 \leq k \leq 51,000$, $Y_{N_p}^{k,h_p} \in \mathbb{R}^2$ is the kth numerical value of (3.14) obtained from the weak Simpson method and $\mathbb{E}[|X(1)|^2]$ is from (3.15). The result of numerical experiment is shown in Figure 3.3 (b), where we have plotted the error against h on log-log scale. It is observed that the slope of the best fit line is empirically three as the the step size h tends to zero.

Remark 3.4. We note that the weak trapezoidal method in Anderson and Mattingly (2011) gives a weak convergence order two; and seems to require 10 million sample paths to obtain such a convergence order for Examples 3.2 and 3.3. In contrast, the weak Simpson method (2.7)–(2.8) gives a weak convergence order three and requires only approximately 5×10^4 sample paths. Therefore it is a substantial improvement over the method in Anderson and Mattingly (2011).

Remark 3.5. The computation of the error may be sometimes influenced by different kind of errors like sampling error, random number bias and rounding error.

In our algorithm we generate more than one sample of random numbers. So, for a large number of sample paths there is a greater chance of dependency in the samples that might degrade the order of convergence of our algorithm.

Chapter 4

Numerical Mean-square And Asymptotic Stability Analysis For The Weak Simpson Method

The concept of weak convergence given in Definition 2.2 concerns the accuracy of a numerical method over a finite interval $[0, T]$ for small step sizes h. However, in many applications the long-term behavior of an SDE is of interest. The concept of numerical stability means whether a numerical solution can keep a similar asymptotic property as $n \to \infty$, when it is applied to the stable SDEs. The stability of various stochastic processes has been extensively studied by many authors; see for instance Khasminskii (2012), Kushner (1967), Mao and Yuan (2006), Meyn and Tweedie (2009), Yin and Zhu (2010) and references therein. Assume that a unique solution $X(t) = X(t; x_0)$ for (1.2) exists for all initial condition $x_0 \in \mathbb{R}^d$ and $t \geq 0$. Furthermore, we assume that $b(0) = \sigma(0) = 0$ and hence $X(t) \equiv 0$ is a steady solution to (1.2). Often, 0 is called an equilibrium point of (1.2). We recall the definition for p^{th}-mean stability and asymptotically stable in pth-mean from Kloeden and Platen (1992).

Definition 4.1. The steady solution $X_t \equiv 0$ is called *stable in pth mean* if for every $\epsilon > 0$, there exists a $\delta = \delta(\epsilon) > 0$ such that $\mathbb{E}[|X(t, x_0)|^p] < \epsilon$ for all $t \geq 0$ and $|x_0| < \delta$. The steady solution $X_t \equiv 0$ is *asymptotically stable in pth mean* if it is stable in pth mean and in addition, there exists a $\delta_0 > 0$ such that

$$\lim_{t\to\infty} \mathbb{E}[|X(t,x_0)|^p] = 0 \quad \text{for all} \quad |x_0| < \delta_0.$$

We call stability in mean and mean-square stability for $p = 1$ and $p = 2$ respectively.

Let us also recall the following definition of asymptotic stability from Bruti-Liberati and Platen (2008).

Definition 4.2. The steady solution $X(t) \equiv 0$ is called *almost surely asymptotically stable* if

$$\mathbb{P}\left(\lim_{t\to\infty} |X(t;x_0)| = 0\right) = 1, \text{ for all } x_0 \in \mathbb{R}.$$

In simulations and numerical approximations, roundoff and truncation errors, sampling error, random number bias, etc. are common. The utility of a numerical method depends upon its ability to control the propagation of such errors in extended time horizon. Concerning the long-time behavior or stability analysis of numerical schemes, the following two questions are fundamental:

(i) Do the numerical solutions of SDEs preserve stability properties of the original SDEs? And if the answer is yes,

(ii) For what range of step sizes h so that the numerical solutions are stable in appropriate senses?

These question have received a lot of attention; some recent developments in this line of research can be found in Higham (2000a,b, 2001), Saito and Mitsui (1996, 2002) and references therein.

In this work, we are concerned with mean-square and almost surely asymptotic stability analysis for the weak Simpson method (2.7)–(2.8). As in the aforementioned references on numerical stability, we will focus on the linear test equation

$$X(t) = X(0) + \int_0^t \lambda X(s)\,\mathrm{d}s + \int_0^t \mu X(s)\,\mathrm{d}W(s), \quad t \geq 0, \tag{4.1}$$

for real or complex constants λ and μ. Here the underlying idea is one that has proved valuable throughout many areas of numerical analysis- study a numerical method on a test problem which is simple enough to allow analysis to be performed, but which retains features present in more general problems of interest. In this thesis we work on the linear, scalar, and autonomous test problem, and the property under consideration is the stability of the trivial or steady solution.

4.1 Linear Stability Analysis For Deterministic Case

We start with the deterministic case when $\mu = 0$ and hence (4.1) reduces to

$$\begin{aligned} \frac{\mathrm{d}X(t)}{\mathrm{d}t} &= \lambda X(t), \quad t > 0 \\ X(0) &= x \neq 0. \end{aligned} \tag{4.2}$$

Here $\lambda \in \mathbb{C}$ is a complex constant. The solution to (4.2) is $X(t) = xe^{\lambda t}$ and hence $\lim_{t\to+\infty} X(t) = 0$ if and only if $\lambda \in \mathbb{C}^-$, where $\mathbb{C}^-$ denotes the left-half complex plane. This is the stability region for (4.2).

The weak Simpson method (2.7)–(2.8) applied to (4.2) produces the recurrence

$$Y_n = \left(1 + \lambda h + \frac{5}{12}\lambda^2 h^2\right) Y_{n-1}. \tag{4.3}$$

Then it follows from (4.3) that

$$\lim_{n\to+\infty} Y_n = 0 \text{ if and only if } \left|1 + \lambda h + \frac{5}{12}\lambda^2 h^2\right| < 1 \tag{4.4}$$

Therefore for a given step size $h > 0$, the stability region for the weak Simpson method is

$$S_w = \left\{\lambda \in \mathbb{C} : \left|1 + \lambda h + \frac{5}{12}\lambda^2 h^2\right| < 1\right\}. \tag{4.5}$$

It is more common to speak of the region of absolute stability as a region in the complex λh-plane. Setting $z = \lambda h = x + iy$ in (4.5) and detailed calculation gives

$$\begin{aligned} S_w = \big\{(x, y) \in \mathbb{R}^2 : \; & 25x^4 + 25y^4 + 50x^2y^2 + 120x^3 \\ & + 120xy^2 + 264x^2 + 24y^2 + 288x < 0\big\}. \end{aligned} \tag{4.6}$$

The stability domain (4.6) of weak Simpson method for the deterministic linear test equation is shown in Figure 4.1.

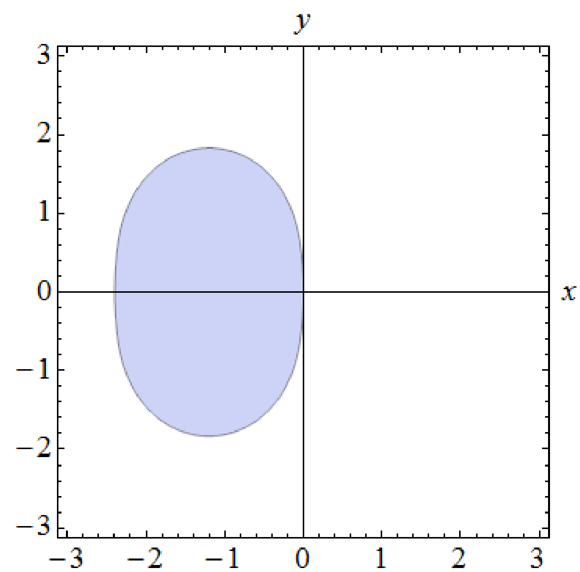


Figure 4.1: Stability domain for weak-Simpson method (blue shaded region)

4.2 Mean-square Stability Analysis

Returning to the SDE (4.1), where we assume that λ and μ are real constants and that $X(0) \neq 0$, since the solution is $X(t) = X(0)e^{(\lambda - \frac{1}{2}\mu^2)t + \mu W(t)}$, we have

$$\lim_{t \to +\infty} \mathbb{E}[|X(t)|^2] = 0 \text{ if and only if } 2\lambda + \mu^2 < 0, \tag{4.7}$$

and

$$\lim_{t \to +\infty} |X(t)| = 0 \text{ with probability 1 if and only if } \lambda - \frac{1}{2}\mu^2 < 0. \tag{4.8}$$

It is clear from (4.7) and (4.8) that mean-square stability implies asymptotic stability but not vice versa. Note that if $W = \{W(t) : 0 \leq t < \infty\}$ is a Brownian motion, so is the process $-W = \{-W(t) : 0 \leq t < \infty\}$. Thus we can and will assume without loss of generality in the rest of the subsection that $\mu > 0$. For ease of later presentation, we denote by

$$S_P = \{(\lambda, \mu) \in \mathbb{R} : 2\lambda + \mu^2 < 0\},$$

the set of ordered pairs of real parameters (λ, μ) so that the trivial solution of (4.1) is mean-square stable.

Applying the weak Simpson method to (4.1) produces the following iterative sequence:

$$Y_n = \left[A + B\eta_1^{(n)}\right] Y_{n-1} + \sqrt{(1-\theta)h\left[C + D\eta_1^{(n)} + E\left(\eta_1^{(n)}\right)^2\right]^+}\,\eta_2^{(n)}\,|Y_{n-1}|\,, \tag{4.9}$$

for $n = 1, 2, \ldots$, where $\{\eta_1^{(n)}, \eta_2^{(n)}, n = 1, 2, \ldots\}$ are mutually independent Gaussian random variables with mean zero and variance one, and

$$\begin{aligned} A &:= 1 + \lambda h + \frac{5}{12}\lambda^2 h^2, \\ B &:= \mu\sqrt{\theta h}\left(1 + \frac{5}{12\theta}\lambda h\right), \\ C &:= \mu^2\left(1 + \alpha_1\lambda^2\theta^2h^2 + 2\alpha_1\lambda\theta h\right) \\ D &:= 2\alpha_1\mu^3(\lambda\theta h + 1)\sqrt{\theta h}, \\ E &:= \alpha_1\mu^4\theta h > 0. \end{aligned} \tag{4.10}$$

Notice C is positive when $\lambda \geq 0$. When $\lambda < 0$, it is easy to see that C is positive when $0 < h < \frac{1}{\lambda\theta}\left(-1 + \sqrt{\alpha_2/\alpha_1}\right)$.

The sequence Y_n of (4.9) is mean-square stable if $\lim_{n\to\infty} \mathbb{E}[|Y_n|^2] = 0$ (Higham (2000b)). By the construction, $Y_{n-1}, \eta_1^{(n)}$ and $\eta_2^{(n)}$ are mutually independent. Thus it follows from (4.9) that

$$\mathbb{E}[|Y_n|^2] = \mathbb{E}[|Y_{n-1}|^2]\left(A^2 + B^2 + (1-\theta)h\mathbb{E}\left[\left[C + D\eta_1^{(n)} + E\left(\eta_1^{(n)}\right)^2\right]^+\right]\right). \tag{4.11}$$

Lemma 4.3. *Assume either one of the following is true:*

$$\lambda \geq 0, \text{ and } h > 0; \tag{4.12}$$

$$\lambda < 0, \text{ and } 0 < h < \frac{1}{\lambda\theta}\left(-1 + \sqrt{\alpha_2/\alpha_1}\right). \tag{4.13}$$

Then we have

$$\mathbb{E}\Big[[C + D\eta_1^{(n)} + E(\eta_1^{(n)})^2]^+\Big] \le C + E + o(h^2). \tag{4.14}$$

Proof. Consider the function $g(x) := C + Dx + Ex^2$, $x \in \mathbb{R}$. As we observed before, under condition (4.12) or (4.13), both C and E are positive. By straightforward computations, $g(x) > 0$ for $x \in (-\infty, x_1) \cup (x_2, \infty)$, where

$$x_1 = -\frac{1 + \sqrt{\alpha_2/\alpha_1} + \lambda\theta h}{\mu\sqrt{\theta h}} < x_2 = \frac{-1 + \sqrt{\alpha_2/\alpha_1} - \lambda\theta h}{\mu\sqrt{\theta h}}.$$

Note from (2.6) that $0 < \alpha_2/\alpha_1 < 1$. Thus $-1+\sqrt{\alpha_2/\alpha_1} < 0$ and $x_2 < 0$. Moreover, $x_2 \to -\infty$ as $h \downarrow 0$. Let $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, x \in \mathbb{R}$ denote the probability density function of a standard normal random variable. Since $\eta_1^{(n)}$ is normally distributed with mean 0 and variance 1, we compute

$$\begin{aligned}
&\mathbb{E}\Big[[C + D\eta_1^{(n)} + E(\eta_1^{(n)})^2]^+\Big] \\
&= \int_{-\infty}^{x_1} (C + Dx + Ex^2)\varphi(x)\,\mathrm{d}x + \int_{x_2}^{\infty} (C + Dx + Ex^2)\varphi(x)\,\mathrm{d}x \\
&\le C + E\int_{-\infty}^{\infty} x^2\varphi(x)\,\mathrm{d}x + D\int_{-\infty}^{x_1} x\varphi(x)\,\mathrm{d}x + D\int_{x_2}^{\infty} x\varphi(x)\,\mathrm{d}x \\
&= C + E - D\int_{x_1}^{x_2} x\varphi(x)\,\mathrm{d}x \\
&= C + E + \frac{D}{\sqrt{2\pi}}\left[\exp\left\{-\frac{x_2^2}{2}\right\} - \exp\left\{-\frac{x_1^2}{2}\right\}\right].
\end{aligned}$$

The proof will be complete if we can show that

$$\frac{D}{\sqrt{2\pi}}\left[\exp\left\{-\frac{x_2^2}{2}\right\} - \exp\left\{-\frac{x_1^2}{2}\right\}\right] = o(h^2) \text{ as } h \downarrow 0.$$

To this end, we note from the expression for D in (4.10) and the fact that $x_1 < x_2 < 0$ that

$$\left|\frac{D}{\sqrt{2\pi}h^2}\left[\exp\left\{-\frac{x_2^2}{2}\right\} - \exp\left\{-\frac{x_1^2}{2}\right\}\right]\right| \le K_1 \frac{e^{-\frac{x_2^2}{2}}}{h^{\frac{3}{2}}} \le K_2 \frac{h^{-\frac{3}{2}}}{e^{\frac{K_3}{h}}} \to 0,$$

as $h \downarrow 0$, where K_1, K_2, and K_3 are positive constants independent of h. This completes the proof of the lemma. $\square$

Putting (4.14) into (4.11) and using the expressions for A, B, C, D, E in (4.10), detailed computations reveal that

$$\begin{aligned}\mathbb{E}[|Y_n|^2] &< \mathbb{E}[|Y_{n-1}|^2]\left(A^2+B^2+(1-\theta)h(C+E+o(h^2))\right)\\ &= \mathbb{E}[|Y_{n-1}|^2]\left[1+(2\lambda+\mu^2)h+\frac{1}{12}[22\lambda^2+20\lambda\mu^2+5\mu^4]h^2+o(h^3)\right].\end{aligned} \tag{4.15}$$

Furthermore, for the expression inside the brackets of the right-hand side of (4.15), we notice that

$$22\lambda^2+20\lambda\mu^2+5\mu^4 = 22\Big(\lambda+\frac{5}{11}\mu^2\Big)^2+\frac{5}{11}\mu^4>0. \tag{4.16}$$

Next we compute the discriminant

$$\Delta = \left(2\lambda+\mu^2\right)^2-4\left(\frac{1}{12}\right)\left(22\lambda^2+20\lambda\mu^2+5\mu^4\right) = -\frac{10}{3}\left[\Big(\lambda+\frac{2}{5}\mu^2\Big)^2+\frac{1}{25}\mu^4\right]<0.$$

Thus it follows that for any $h>0$,

$$1+(2\lambda+\mu^2)h+\frac{1}{12}[22\lambda^2+20\lambda\mu^2+5\mu^4]h^2>0.$$

Putting this observation into (4.15), we obtain a condition for mean-square stability of the weak Simpson method (2.7)–(2.8) for (4.1)

$$1+(2\lambda+\mu^2)h+\frac{1}{12}[22\lambda^2+20\lambda\mu^2+5\mu^4]h^2<1. \tag{4.17}$$

Since $h>0$, and noting (4.16), we can rewrite equation (4.17) as

$$0<h<\frac{-12(2\lambda+\mu^2)}{22\lambda^2+20\lambda\mu^2+5\mu^4}. \tag{4.18}$$

Note that when $\lambda \geq 0$, the set of h that satisfies (4.18) is an empty set. On the other hand, when $\lambda<0$, $2\lambda+\mu^2<0$, and $h>0$ satisfies (4.13) and (4.18), then the weak Simpson method (2.7)–(2.8) is mean-square stable for (4.1). In other words, given $(\lambda,\mu)\in S_P$, the combination of (4.13) and (4.18):

$$0<h<\min\left\{\frac{-12(2\lambda+\mu^2)}{22\lambda^2+20\lambda\mu^2+5\mu^4}, \frac{1}{\lambda\theta}\left(-1+\sqrt{\alpha_2/\alpha_1}\right)\right\}. \tag{4.19}$$

gives a sufficient condition for mean-square stability of the weak Simpson method (2.7)–(2.8) for (4.1).

Conversely, suppose the weak Simpson method with discretization stepsize $h > 0$ is mean-square stable for (4.1). Note from (4.11) that

$$\mathbb{E}[|Y_n|^2] \geq \mathbb{E}[|Y_{n-1}|^2](A^2 + B^2) = \mathbb{E}[|Y_{n-1}|^2](1 + (2\lambda + \mu^2\theta)h + O(h^2)).$$

Thus for the weak Simpson method to be mean-square stable, λ, μ, and θ necessarily satisfy $2\lambda + \mu^2\theta \leq 0$.

We summarize the above discussion into the following theorem:

Theorem 4.4. *The following assertions are true:*

(a) Given $(\lambda, \mu) \in S_P$, the weak Simpson method is mean-square stable if the discretization stepsize h satisfies (4.19). Therefore the mean square stability of the process (4.1) implies the mean square stability of the weak Simpson method if the discretization stepsize h satisfies (4.19).

(b) Conversely, if the weak Simpson method with discretization stepsize $h > 0$ is mean-square stable for (4.1), then the parameters λ and μ of (4.1) satisfies $2\lambda + \mu^2\theta \leq 0$.

We can visualize the stability region when

$$\frac{-12(2\lambda + \mu^2)}{22\lambda^2 + 20\lambda\mu^2 + 5\mu^4} < \frac{1}{\lambda\theta}\left(-1 + \sqrt{\alpha_2/\alpha_1}\right). \tag{4.20}$$

It is common in the literature to visualize the region of stability in the xy-plane, in which $x = \lambda h$ and $y = \mu^2 h > 0$. Therefore using (4.17), we have

$$S_M := \left\{(x, y) \in \mathbb{R}^2 : 22x^2 + 20xy + 5y^2 + 24x + 12y < 0\right\}. \tag{4.21}$$

Note that since $22x^2 + 20xy + 5y^2 = 22(x + \frac{5}{11}y)^2 + \frac{5}{11}y^2 > 0$, for any $(x, y) \in S_M$, we necessarily have $24x + 12y < 0$ or $y < -2x$. See Figure 4.2 for the plot of the mean-square stability domain for the weak Simpson method S_M.

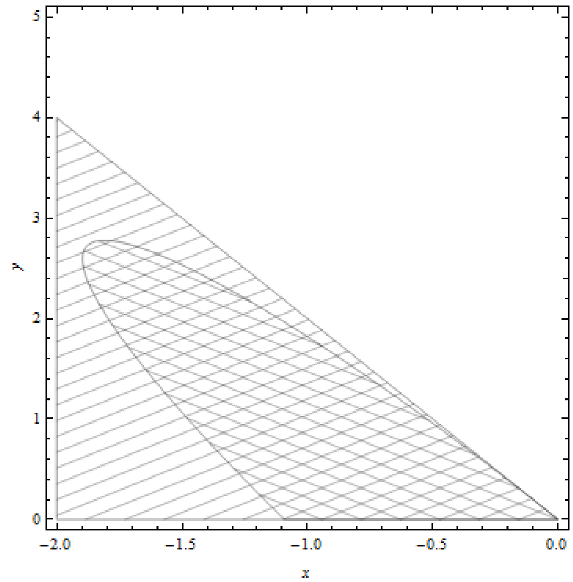


Figure 4.2: Real mean-square stability domain for weak Simpson method (crossed hashing)

Example 4.5. Again, we test the mean-square stability over $[0, 30]$ with non-random initial value $X_0 = 1$. We take $\lambda = -2$ and $\mu = \sqrt{2}$ in (4.1). We have $2\lambda + \mu^2 < 0$ and hence thanks to (4.7), the trivial solution of (4.1) is mean-square stable. The left-hand side of (4.20) is equal to $\frac{6}{7}$ and the right-hand side of (4.20) is equal to 0.37. We apply weak Simpson method to simulate 45000 discrete sample paths of (4.1) for stepsizes $h = 1, \frac{1}{2}, \frac{1}{4}$. The stepsize $h = \frac{1}{4}$ satisfy (4.20) but not the stepsizes $h = 1$ and $h = \frac{1}{2}$. Therefore the weak Simpson method is mean-square stable for (4.1) when $h = \frac{1}{4}$. We plot the sample average of Y_n^2 against $t_n := nh$ with logarithmically scaled y-axis in Figure 4.3. The numerical experiments indicates that weak Simpson method is mean-square stable for $h = \frac{1}{2}$ or $\frac{1}{4}$, and unstable for $h = 1$. These observations are consistent with Theorem 4.4.

4.3 Almost Sure Asymptotic Stability Analysis

The sequence (4.9) is almost surely asymptotically stable if $\lim_{n\to\infty} |Y_n| = 0$ with probability one. The weak-Simpson method is almost surely asymptotically stable

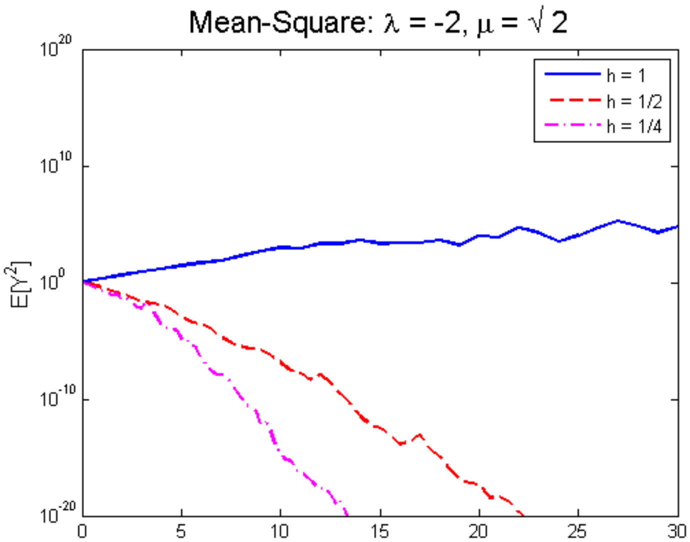


Figure 4.3: Mean-square stability test

if it produces an almost surely asymptotically stable sequence.

Let us first quote the following lemma from Higham (2000b).

Lemma 4.6. *Given a sequence of real-valued, nonnegative, independent and identically distributed random variables Z_n, consider the sequence of random variables $\{Y_n\}_{n\geq 1}$ defined by*

$$Y_n = Y_0 \prod_{i=0}^{n-1} Z_i, \tag{4.22}$$

where $Y_0 \geq 0$ and $Y_0 \neq 0$ with probability one. Suppose that the random variables $\log(Z_i)$ are square-integrable. Then

$$\lim_{n\to\infty} Y_n = 0 \text{ with probability one } \iff \mathbb{E}[\log(Z_i)] < 0.$$

In order to apply lemma 4.6, as in Higham (2000b) we take

$$Z_i = \left|\left[A + B\eta_1^{(i)}\right]\right| + \left|\sqrt{(1-\theta)h\left[C + D\eta_1^{(i)} + E\left(\eta_1^{(i)}\right)^2\right]^+}\eta_2^{(i)}\right|$$

where A, B, C, D and E are constants defined in (4.10) and $\eta_1^{(i)}$ and $\eta_2^{(i)}$ are mutually independent normal random variables with mean zero and variance one. Furthermore assume that $\log(Z_i)$ are square integrable. We see that a sufficient condition for almost surely asymptotic stability is

$$\mathbb{E}\left[\log\left(\left|\left[A+B\eta_1^{(i)}\right]\right|+\left|\sqrt{(1-\theta)h\left[C+D\eta_1^{(i)}+E\left(\eta_1^{(i)}\right)^2\right]^+}\eta_2^{(i)}\right|\right)\right]<0. \quad (4.23)$$

We use the parameters $\lambda = \frac{1}{3}$ and $\mu = \sqrt{3}$ to test the asymptotic stability of our scheme. The SDE is asymptotically stable but not mean-square stable for $\lambda = \frac{1}{3}$ and $\mu = \sqrt{3}$. We integrate over $[0, 600]$ using weak Simpson method for step sizes $h = 1, \frac{1}{2}, \frac{1}{4}$. The plot of $|Y_n|$ versus t_j is shown in Figure 4.4. We see that the solutions for $h = \frac{1}{2}, \frac{1}{4}$ decay to zero as t increases, in a manner consistent with the theory.

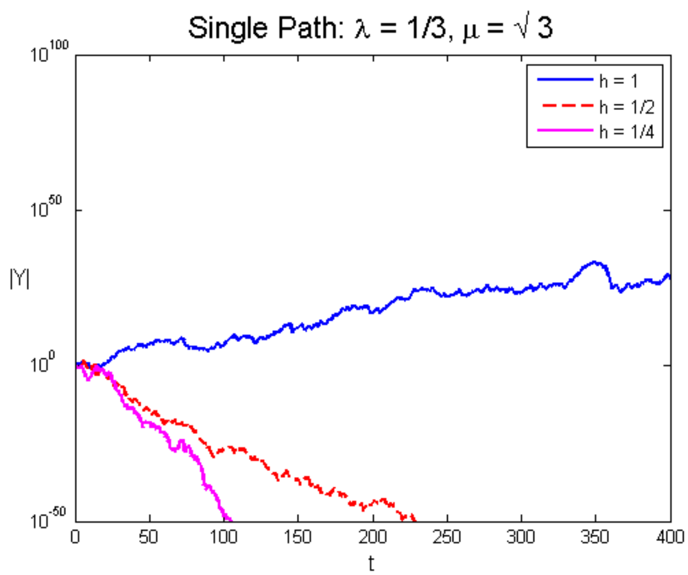


Figure 4.4: Asymptotic stability test

Appendix A

The Proofs of Lemmas 2.7 and 2.10

The proofs of Lemmas 2.7 and 2.10 depend on the following Lemmas.

Lemma A.1. *Let X, Z, W, and Y be real valued random variables on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Let $p, q > 1$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then the following assertions are true:*

(i) if $\|XY\|_{L^p(\Omega)} < \infty$, then

$$\mathbb{E}\big[\big|YX^+ - YX\big|\big] \leq \|XY\|_{L^p(\Omega)}(\mathbb{P}\{X < 0\})^{\frac{1}{q}},$$

(ii) if $\|ZXY\|_{L^p(\Omega)} < \infty$, then

$$\mathbb{E}\left[\big|ZY^+X^+ - ZYX\big|\right] \leq \|XYZ\|_{L^p(\Omega)}(\mathbb{P}\{X < 0 \text{ or } Z < 0\})^{\frac{1}{q}},$$

(iii) if $\|WXYZ\|_{L^p(\Omega)} < \infty$, then

$$\mathbb{E}\big[\big|ZW^+X^+Y^+ - ZWXY\big|\big] \leq \|WYXZ\|_{L^p(\Omega)}(\mathbb{P}\{W < 0 \text{ or } X < 0 \text{ or } Y < 0\})^{\frac{1}{q}}.$$

Proof. Let $A = \{X < 0\}$. Observe that $YX^+ - YX = 0$ on A^c and $YX^+ - YX = -YX$ on A. Thus it follows from the Hölder inequality that

$$\mathbb{E}\left[|YX^+ - YX|\right] = \mathbb{E}[|-YXI_A|] \leq \|XY\|_{L^p(\Omega)}(\mathbb{P}\{X < 0\})^{\frac{1}{q}},$$

proving the first assertion.

The proof of the second assertion follows from similar arguments. Observe that $ZY^+X^+ - YXZ = -YXZ$ on the set $\{X < 0 \text{ or } Y < 0\}$ and $ZY^+X^+ - YXZ = 0$ on the set $\{X < 0 \text{ or } Y < 0\}^c = \{X \geq 0 \text{ and } Y \geq 0\}$. Thus

$$\begin{aligned}\mathbb{E}[|ZY^+X^+ - ZYX|] &= \mathbb{E}[|-ZYX| I_{\{X<0 \text{ or } Y<0\}}] \\ &\leq \|XYZ\|_{L^p(\Omega)} (\mathbb{P}\{X < 0 \text{ or } Y < 0\})^{\frac{1}{q}}.\end{aligned}$$

For the proof of the third assertion, let $A := \{W < 0 \text{ or } X < 0 \text{ or } Y < 0\}$. Then we have $A^c = \{W \geq 0 \text{ and } X \geq 0 \text{ and } Y \geq 0\}$. Moreover, on the set A^c, $ZW^+X^+Y^+ = ZWXY$. Thus it follows from the Hölder inequality that

$$\mathbb{E}\left[\left|ZW^+X^+Y^+ - ZWXY\right|\right] = \mathbb{E}[|-ZWXYI_A|] \leq \|ZWXY\|_{L^p(\Omega)} \mathbb{P}(A)^{1/q}.$$

$\square$

Lemma A.2. *Suppose Assumptions (A1) and (A2). Then for each* $k = 1, \ldots, M$ *and any* $p > 0$*, there exists an* $h_0 > 0$ *so that*

$$\mathbb{P}\left\{\alpha_1 \sigma_k^2(Y_1^*) - \alpha_2 \sigma_k^2(x_0) < 0\right\} = O(h^p) \quad \textit{for all } 0 < h < h_0, \tag{A.1}$$

where Y_1^* *is given in* (2.7)*.*

Proof. Denote $E_k := \{\alpha_1 \sigma_k^2(Y_1^*) - \alpha_2 \sigma_k^2(x_0) < 0\}$. As we noted before, Y_1^* of (2.7) and $y(\theta h)$ of (2.20) have the same distribution. Thus

$$\begin{aligned}\mathbb{P}(E_k) &= \mathbb{P}\left\{\alpha_1 \sigma_k^2(y(\theta h)) - \alpha_2 \sigma_k^2(x_0) < 0\right\} \\ &= \mathbb{P}\left\{|\sigma_k(y(\theta h))| < \sqrt{\alpha_2/\alpha_1}\, |\sigma_k(x_0)|\right\} \\ &= \mathbb{P}\left\{|\sigma_k(y(\theta h))| - |\sigma_k(x_0)| < \left(\sqrt{\alpha_2/\alpha_1} - 1\right) |\sigma_k(x_0)|\right\}.\end{aligned} \tag{A.2}$$

On the other hand, using the triangle inequality and (2.2), we have

$$|\sigma_k(y(\theta h))| - |\sigma_k(x_0)| \geq -|\sigma_k(y(\theta h)) - \sigma_k(x_0)| \geq -L\,|y(\theta h) - x_0|\,.$$

Putting this into (A.2), we have

$$\begin{aligned}
\mathbb{P}(E_k) &\leq \mathbb{P}\left\{-L\,|y(\theta h)-x_0| \leq \left(\sqrt{\alpha_2/\alpha_1}-1\right)|\sigma_k(x_0)|\right\} = \mathbb{P}\left\{|y(\theta h)-x_0| \geq C\right\} \\
&= \mathbb{P}\left\{\left|b(x_0)\theta h + \sum_{k=1}^{M}\sigma_k(x_0)\nu_k(W_k(\theta h)-W_k(0))\right| \geq C\right\} \\
&\leq \mathbb{P}\left\{\left|\sum_{k=1}^{M}\sigma_k(x_0)\nu_k(W_k(\theta h)-W_k(0))\right| \geq C - |b(x_0)|\,\theta h\right\} \\
&= \mathbb{P}\left\{\left|\sum_{k=1}^{M}\sigma_k(x_0)\nu_k\frac{W_k(\theta h)-W_k(0)}{\sqrt{\theta h}}\right| \geq \frac{C-|b(x_0)|\,\theta h}{\sqrt{\theta h}}\right\},
\end{aligned}$$

where $C := \frac{1}{L}(1-\sqrt{\alpha_2/\alpha_1})\,|\sigma_k(x_0)|$. Note that $C > 0$ since by (2.6), $0 < \alpha_2 < \alpha_1$. For each $k = 1, \ldots, M$ and $h > 0$, $Z_k := \frac{W_k(\theta h)-W_k(0)}{\sqrt{\theta h}}$ has standard normal distribution. This, together with the assumption that $W_1, \ldots, W_M$ are independent Brownian motions, implies that $\sum_{k=1}^{M}\sigma_k(x_0)\nu_k\frac{W_k(\theta h)-W_k(0)}{\sqrt{\theta h}}$ has multivariate normal distribution with mean zero and covariance matrix $\sum_{k=1}^{M}\sigma_k^2(x_0)\nu_k\nu_k^T = a(x_0)$. Now (A.1) follows from (2.3) and the usual Gaussian tail estimation (see, for instance, Theorem 1 of Hüsler et al. (2002)). $\square$

Corollary A.3. *Assume the conditions of Lemma A.2. Suppose $f \in C^8(\mathbb{R}^d)$ and that for all multi-index α with $|\alpha| \leq 8$, we have*

$$|D^\alpha f(x)| \leq K(1+|x|^q) \tag{A.3}$$

for some positive constants K and $q \geq 1$. Then for any $k, j, l = 1, \ldots, M$ and $p \geq 1$, there exists an $h_0 > 0$ so that for all $h \in (0, h_0]$, we have

$$\begin{aligned}
&\mathbb{E}\left[(\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0))^+ f''[\nu_k,\nu_k](Y_1^*)\right] \\
&\quad = \mathbb{E}\left[(\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)) f''[\nu_k,\nu_k](Y_1^*)\right] + O(h^p),
\end{aligned} \tag{A.4}$$

$$\begin{aligned}
&\mathbb{E}\left[[\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)]^+[\alpha_1\sigma_j^2(Y_1^*) - \alpha_2\sigma_j^2(x_0)]^+ f^{(4)}[\nu_k,\nu_k,\nu_j,\nu_j](Y_1^*)\right] \\
&\quad = \mathbb{E}\left[[\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)][\alpha_1\sigma_j^2(Y_1^*) - \alpha_2\sigma_j^2(x_0)] f^{(4)}[\nu_k,\nu_k,\nu_j,\nu_j](Y_1^*)\right] + O(h^p),
\end{aligned} \tag{A.5}$$

and

$$\begin{aligned}
&\mathbb{E}\big[[\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)]^+[\alpha_1\sigma_j^2(Y_1^*) - \alpha_2\sigma_j^2(x_0)]^+[\alpha_1\sigma_l^2(Y_1^*) - \alpha_2\sigma_l^2(x_0)]^+ \\
&\qquad\qquad \times f^{(6)}[\nu_k,\nu_k,\nu_j,\nu_j,\nu_l,\nu_l](Y_1^*)\big] \\
&\quad = \mathbb{E}\big[[\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)][\alpha_1\sigma_j^2(Y_1^*) - \alpha_2\sigma_j^2(x_0)][\alpha_1\sigma_l^2(Y_1^*) - \alpha_2\sigma_l^2(x_0)] \\
&\qquad\qquad \times f^{(6)}[\nu_k,\nu_k,\nu_j,\nu_j,\nu_l,\nu_l](Y_1^*)\big] + O(h^p),
\end{aligned} \tag{A.6}$$

Proof. As observed in the proof of Lemma A.2, Y_1^* is equal to $y(\theta h)$ in distribution, where $y(\theta h)$ is given by (2.20). Therefore, in view of (A.3), the standard arguments as those in Øksendal (2003) or Yin and Zhu (2010) yield

$$\begin{aligned}
&\|(\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0))f''[\nu_k,\nu_k](Y_1^*)\|_{L^p} \\
&= \|(\alpha_1\sigma_k^2(y(\theta h)) - \alpha_2\sigma_k^2(x_0))f''[\nu_k,\nu_k](y(\theta h))\|_{L^p} \\
&\le K < \infty.
\end{aligned}$$

Then (A.4) follows from Lemmas A.1 and A.2.

Observe that

$$\begin{aligned}
&\mathbb{P}\left\{[\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0) < 0] \text{ or } [\alpha_1\sigma_j^2(Y_1^*) - \alpha_2\sigma_j^2(x_0) < 0]\right\} \\
&\quad \le \mathbb{P}\left\{\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0) < 0\right\} + \mathbb{P}\left\{\alpha_1\sigma_j^2(Y_1^*) - \alpha_2\sigma_j^2(x_0) < 0\right\}.
\end{aligned}$$

Then (A.5) follows from a similar argument as above using Lemmas A.1 and A.2.

In a similar fashion, we can establish (A.6). □

Lemma A.4. *For any sufficiently smooth function $f : \mathbb{R}^d \mapsto \mathbb{R}$, we have*

$$\begin{aligned}
B_1^2 f(x) &= f''[b(x_0), b(x_0)](x) + \sum_{k=1}^{M} \sigma_k^2(x_0) f'''[\nu_k,\nu_k,b(x_0)](x) \\
&\quad + \frac{1}{4}\sum_{k,j=1}^{M} \sigma_k^2(x_0)\sigma_j^2(x_0) f^{(4)}[\nu_k,\nu_k,\nu_j,\nu_j](x), \\
A(B_1 f)(x) &= f''[b(x), b(x_0)](x) + \frac{1}{2}\sum_{k=1}^{M} \sigma_k^2(x_0) f'''[\nu_k,\nu_k,b(x)](x) \\
&\quad + \frac{1}{2}\sum_{k=1}^{M} \sigma_k^2(x) f'''[\nu_k,\nu_k,b(x_0)](x)
\end{aligned}$$

$$
\begin{aligned}
&+\frac{1}{4}\sum_{k,j=1}^{M}\sigma_k^2(x)\sigma_j^2(x_0)f^{(4)}[\nu_k,\nu_k,\nu_j,\nu_j](x),\\
B^2f(x) &= f''[\alpha_1 b(Y_1^*)-\alpha_2 b(x_0),\alpha_1 b(Y_1^*)-\alpha_2 b(x_0)](x)\\
&+\sum_{k=1}^{M}(\alpha_1\sigma_k^2(Y_1^*)-\alpha_2\sigma_k^2(x_0))^+f'''[\nu_k,\nu_k,\alpha_1 b(Y_1^*)-\alpha_2 b(x_0)](x)\\
&+\frac{1}{4}\sum_{k,j=1}^{M}(\alpha_1\sigma_k^2(Y_1^*)-\alpha_2\sigma_k^2(x_0))^+(\alpha_1\sigma_j^2(Y_1^*)-\alpha_2\sigma_j^2(x_0))^+\\
&\qquad\times f^{(4)}[\nu_k,\nu_k,\nu_j,\nu_j](x),\\
B_1^3f(x) &= f'''[b(x_0),b(x_0),b(x_0)](x)+\frac{3}{2}\sum_{k=1}^{M}\sigma_k^2(x_0)f^{(4)}[\nu_k,\nu_k,b(x_0),b(x_0)](x)\\
&+\frac{3}{4}\sum_{k,j=1}^{M}\sigma_k^2(x_0)\sigma_j^2(x_0)f^{(5)}[\nu_k,\nu_k,\nu_j,\nu_j,b(x_0)](x)\\
&+\frac{1}{8}\sum_{k,j,l=1}^{M}\sigma_k^2(x_0)\sigma_j^2(x_0)\sigma_l^2(x_0)f^{(6)}[\nu_k,\nu_k,\nu_j,\nu_j,\nu_l,\nu_l](x)
\end{aligned}
$$

and

$$
\begin{aligned}
B^3f(x) &= f'''[\alpha_1 b(Y_1^*)-\alpha_2 b(x_0),\alpha_1 b(Y_1^*)-\alpha_2 b(x_0),\alpha_1 b(Y_1^*)-\alpha_2 b(x_0)](x)\\
&+\frac{3}{2}\sum_{k=1}^{M}(\alpha_1\sigma_k^2(Y_1^*)-\alpha_2\sigma_k^2(x_0))^+\\
&\qquad\times f^{(4)}[\nu_k,\nu_k,\alpha_1 b(Y_1^*)-\alpha_2 b(x_0),\alpha_1 b(Y_1^*)-\alpha_2 b(x_0)](x)\\
&+\frac{3}{4}\sum_{k,j=1}^{M}(\alpha_1\sigma_k^2(Y_1^*)-\alpha_2\sigma_k^2(x_0))^+(\alpha_1\sigma_j^2(Y_1^*)-\alpha_2\sigma_j^2(x_0))^+\\
&\qquad\times f^{(5)}[\nu_k,\nu_k,\nu_j,\nu_j,\alpha_1 b(Y_1^*)-\alpha_2 b(x_0)](x)\\
&+\frac{1}{8}\sum_{k,j,l=1}^{M}(\alpha_1\sigma_k^2(Y_1^*)-\alpha_2\sigma_k^2(x_0))^+(\alpha_1\sigma_j^2(Y_1^*)-\alpha_2\sigma_j^2(x_0))^+\\
&\qquad\times(\alpha_1\sigma_l^2(Y_1^*)-\alpha_2\sigma_l^2(x_0))^+f^{(6)}[\nu_k,\nu_k,\nu_j,\nu_j,\nu_l,\nu_l](x).
\end{aligned}
$$

Proof. This lemma follows from straightforward and tedious calculations. We shall omit the details here. $\square$

Now we prove Lemma 2.7 and then we prove Lemma 2.10.

Proof of Lemma 2.7. We analyze every term on the left hand side of (2.15).

Step 1. Since Y_1^* is equal to $y(\theta h)$ of (2.20) in distribution, and noting that the infinitesimal generator of (2.19) is given by B_1 in (2.21), we can apply the Dynkin formula repeatedly to obtain

$$\begin{aligned}\mathbb{E}[f(Y_1^*)] &= f(x_0) + \int_0^{\theta h} \mathbb{E}[(B_1 f)(Y_1^*)]\,\mathrm{d}s \\ &= f(x_0) + (B_1 f)(x_0)\theta h + \int_0^{\theta h}\int_0^s \mathbb{E}[(B_1^2 f)(Y_1^*(r))]\,dr\,\mathrm{d}s.\end{aligned}$$

The term $(B_1^2 f)(Y_1^*(r))$ in the last integral above depends on the first four derivatives of f. Since $f \in C^4$,

$$\left|\int_0^{\theta h}\int_0^s \mathbb{E}(B_1^2 f)(Y_1^*(r))dr\,ds\right| \le C_2||f||_4 h^2$$

for some constant C_2 independent of h. Thus we have

$$\begin{aligned}\mathbb{E}[f(Y_1^*)] &= f(x_0) + (B_1 f)(x_0)\theta h + O(h^2) \\ &= f(x_0) + (Af)(x_0)\theta h + O(h^2),\end{aligned} \tag{A.7}$$

where the second equality follows from the observation that $(B_1 f)(x_0) = (Af)(x_0)$.

Step 2. Next we deal with the term $\mathbb{E}[(Bf)(Y_1^*)]$. It follows from the definition of B and (A.4) that

$$\begin{aligned}&\mathbb{E}[(Bf)(Y_1^*)] \\ &= \mathbb{E}\Big[\Big(f'[\alpha_1 b(Y_1^*) - \alpha_2 b(x_0)] + \frac{1}{2}\sum_{k=1}^M [\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)]^+ f''[\nu_k,\nu_k]\Big)(Y_1^*)\Big] \\ &= \mathbb{E}\big[f'[\alpha_1 b(Y_1^*) - \alpha_2 b(x_0)](Y_1^*)\big] \\ &\quad + \frac{1}{2}\sum_{k=1}^M \mathbb{E}\Big[[\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)]f''[\nu_k,\nu_k])(Y_1^*)\Big] + O(h) \\ &= (Af)(x_0) + O(h).\end{aligned} \tag{A.8}$$

Step 3. Combining (A.7) and (A.8) gives the desired result. □

Proof of Lemma 2.10. We analyze every term on the left hand side of (2.28).

Step 1. Since Y_1^* is equal to $y(\theta h)$ of (2.20) in distribution, and noting that the infinitesimal generator of (2.19) is given by B_1 in (2.21), we can apply the Dynkin formula repeatedly to obtain

$$\begin{aligned}
\mathbb{E}[f(Y_1^*)] &= f(x_0) + \int_0^{\theta h} \mathbb{E}[(B_1 f)(Y_1^*)]\,\mathrm{d}s \\
&= f(x_0) + (B_1 f)(x_0)\theta h + \int_0^{\theta h}\int_0^s \mathbb{E}[(B_1^2 f)(Y_1^*(r))]\; dr\,\mathrm{d}s \\
&= f(x_0) + (B_1 f)(x_0)\theta h + (B_1^2 f)(x_0)\frac{\theta^2 h^2}{2} \\
&\quad + \int_0^{\theta h}\int_0^s\int_0^r \mathbb{E}[(B_1^3 f)(Y_1^*(u))]\,\mathrm{d}u\,\mathrm{d}r\,\mathrm{d}s \\
&= f(x_0) + (B_1 f)(x_0)\theta h + (B_1^2 f)(x_0)\frac{\theta^2 h^2}{2} + (B_1^3 f)(x_0)\frac{\theta^3 h^3}{6} \\
&\quad + \int_0^{\theta h}\int_0^s\int_0^r\int_0^v \mathbb{E}[(B_1^4 f)(Y_1^*(w))]\,\mathrm{d}w\,\mathrm{d}u\,\mathrm{d}r\,\mathrm{d}s.
\end{aligned}$$

The term $(B_1^4 f)(Y_1^*(w))$ in the last integral above depends on the first eight derivatives of f. Since $f \in C^8$,

$$\left|\int_0^{\theta h}\int_0^s\int_0^r\int_0^v \mathbb{E}(B_1^4 f)(Y_1^*(w))dw\,du\,dr\,ds\right| \le C_2'||f||_8 h^4$$

for some constant C_2' independent of h.

$$\begin{aligned}
\mathbb{E}[f(Y_1^*)] &= f(x_0) + (B_1 f)(x_0)\theta h + (B_1^2 f)(x_0)\frac{\theta^2 h^2}{2} + (B_1^3 f)(x_0)\frac{\theta^3 h^3}{6} + O(h^4) \\
&= f(x_0) + (Af)(x_0)\theta h + (B_1^2 f)(x_0)\frac{\theta^2 h^2}{2} + (B_1^3 f)(x_0)\frac{\theta^3 h^3}{6} + O(h^4),
\end{aligned} \tag{A.9}$$

where the second equality follows from the observation that $(B_1 f)(x_0) = (Af)(x_0)$.

Step 2. Next we deal with the term $\mathbb{E}[(Bf)(Y_1^*)]$. It follows from the definition of B and (A.4) that

$$
\begin{aligned}
&\mathbb{E}[(Bf)(Y_1^*)] \\
&\quad = \mathbb{E}\Big[\Big(f'[\alpha_1 b(Y_1^*) - \alpha_2 b(x_0)] + \frac{1}{2}\sum_{k=1}^{M}[\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)]^+ f''[\nu_k,\nu_k]\Big)(Y_1^*)\Big] \\
&\quad = \mathbb{E}\big[f'[\alpha_1 b(Y_1^*) - \alpha_2 b(x_0)](Y_1^*)\big] \\
&\qquad + \frac{1}{2}\sum_{k=1}^{M}\mathbb{E}\Big[[\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0)]f''[\nu_k,\nu_k])(Y_1^*)\Big] + O(h^3) \\
&\quad = \mathbb{E}\big[f'[\alpha_1 b(y(\theta h)) - \alpha_2 b(x_0)](y(\theta h))\big] \\
&\qquad + \frac{1}{2}\sum_{k=1}^{M}\mathbb{E}\Big[[\alpha_1\sigma_k^2(y(\theta h)) - \alpha_2\sigma_k^2(x_0)]f''[\nu_k,\nu_k])(y(\theta h))\Big] + O(h^3).
\end{aligned}
$$

In the above, we again used the fact that Y_1^* and $y(\theta h)$ of (2.20) have the same distribution to obtain the third equality. Moreover, since

$$
\begin{aligned}
f'[\alpha_1 b(y(\theta h)) - \alpha_2 b(x_0)](y(\theta h)) &= \alpha_1 b(y(\theta h))\cdot Df(y(\theta h)) - \alpha_2 b(x_0)\cdot Df(y(\theta h)) \\
&= \alpha_1 f'[b(y(\theta h))](y(\theta h)) - \alpha_2 f'[b(x_0)](y(\theta h)),
\end{aligned}
$$

and for each $k = 1, \ldots, M$,

$$
\begin{aligned}
&[\alpha_1\sigma_k^2(y(\theta h)) - \alpha_2\sigma_k^2(x_0)]f''[\nu_k,\nu_k](y(\theta h)) \\
&\quad = \alpha_1\sigma_k^2(y(\theta h))f''[\nu_k,\nu_k](y(\theta h)) - \alpha_2\sigma_k^2(x_0)f''[\nu_k,\nu_k](y(\theta h)),
\end{aligned}
$$

we have

$$
\begin{aligned}
\mathbb{E}[(Bf)(Y_1^*)] &= \alpha_1\mathbb{E}\Big[f'[b(y(\theta h))](y(\theta h)) + \frac{1}{2}\sum_{k=1}^{M}\sigma_k(y(\theta h))^2 f''[\nu_k,\nu_k])(y(\theta h))\Big] \\
&\qquad - \alpha_2\mathbb{E}\Big[f'[b(x_0)](y(\theta h)) + \frac{1}{2}\sum_{k=1}^{M}\sigma_k(x_0)^2 f''[\nu_k,\nu_k])(y(\theta h))\Big] \\
&= \alpha_1\mathbb{E}[Af(y(\theta h))] - \alpha_2\mathbb{E}[B_1 f(y(\theta h))],
\end{aligned} \tag{A.10}
$$

where A and B_1 are defined in (2.16) and (2.21); they are the infinitesimal generators for the stochastic differential equations (1.1) and (2.19), respectively. Now we apply

Dynkin's formula repeatedly to obtain

$$
\begin{aligned}
\mathbb{E}[Af(y(\theta h))] &= Af(x_0) + \int_0^{\theta h} \mathbb{E}[B_1(Af)(y(s))]\,\mathrm{d}s \\
&= Af(x_0) + B_1(Af)(x_0)\theta h + \int_0^{\theta h}\int_0^{s} \mathbb{E}[B_1^2(Af)(y(r))]\,\mathrm{d}r\,\mathrm{d}s \\
&= Af(x_0) + B_1(Af)(x_0)\theta h + B_1^2(Af)(x_0)\frac{\theta^2 h^2}{2} \\
&\qquad + \int_0^{\theta h}\int_0^{s}\int_0^{r} \mathbb{E}[B_1^3(Af)(y(u))]\,\mathrm{d}u\,\mathrm{d}r\,\mathrm{d}s \\
&= Af(x_0) + B_1(Af)(x_0)\theta h + B_1^2(Af)(x_0)\frac{\theta^2 h^2}{2} + O(h^3).
\end{aligned} \tag{A.11}
$$

Similarly, we have

$$
\mathbb{E}[B_1 f(y(\theta h))] = B_1 f(x_0) + B_1^2 f(x_0)\theta h + B_1^3 f(x_0)\frac{\theta^2 h^2}{2} + O(h^3). \tag{A.12}
$$

Notice that $B_1 f(x_0) = Af(x_0)$ and hence $B_1(Af)(x_0) = A(Af)(x_0) = A^2 f(x_0)$. Using these observations in (A.11) and (A.12) and plugging them into (A.10), and noting $\alpha_1 - \alpha_2 = 1$, we have

$$
\begin{aligned}
\mathbb{E}[(Bf)(Y_1^*)] &= (Af)(x_0) + \alpha_1(A^2 f)(x_0)\theta h - \alpha_2(B_1^2 f)(x_0)\theta h \\
&\qquad + \alpha_1 B_1^2(Af)(x_0)\frac{\theta^2 h^2}{2} - \alpha_2(B_1^3 f)(x_0)\frac{\theta^2 h^2}{2} + O(h^3).
\end{aligned} \tag{A.13}
$$

Step 3. Next we evaluate $\mathbb{E}(B^2 f)(Y_1^*)$. Thanks to Lemma A.4 and Corollary A.3, we have

$$
\begin{aligned}
\mathbb{E}[B^2 f(Y_1^*)] &= \mathbb{E}[f''[\alpha_1 b(Y_1^*) - \alpha_2 b(x_0), \alpha_1 b(Y_1^*) - \alpha_2 b(x_0)](Y_1^*)] \\
&\quad + \sum_{k=1}^{M} \mathbb{E}[(\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0))^+ \\
&\qquad\qquad \times f'''[\nu_k, \nu_k, \alpha_1 b(Y_1^*) - \alpha_2 b(x_0)](Y_1^*)] \\
&\quad + \frac{1}{4}\sum_{k,j=1}^{M} \mathbb{E}[(\alpha_1\sigma_k^2(Y_1^*) - \alpha_2\sigma_k^2(x_0))^+(\alpha_1\sigma_j^2(Y_1^*) - \alpha_2\sigma_j^2(x_0))^+ \\
&\qquad\qquad \times f^{(4)}[\nu_k, \nu_k, \nu_j, \nu_j](Y_1^*)] \\
&= \mathbb{E}[f''[\alpha_1 b(Y_1^*) - \alpha_2 b(x_0), \alpha_1 b(Y_1^*) - \alpha_2 b(x_0)](Y_1^*)]
\end{aligned}
$$

$$+ \sum_{k=1}^{M} \mathbb{E}[(\alpha_1 \sigma_k^2(Y_1^*) - \alpha_2 \sigma_k^2(x_0)) f'''[\nu_k, \nu_k, \alpha_1 b(Y_1^*) - \alpha_2 b(x_0)](Y_1^*)]$$
$$+ \frac{1}{4} \sum_{k,j=1}^{M} \mathbb{E}[(\alpha_1 \sigma_k^2(Y_1^*) - \alpha \sigma_k^2(x_0))(\alpha_1 \sigma_j^2(Y_1^*) - \alpha_2 \sigma_j^2(x_0))$$
$$\times f^{(4)}[\nu_k, \nu_k, \nu_j, \nu_j](Y_1^*)] + O(h^2).$$

Moreover, detailed calculations using Lemma A.4 reveal that

$$\begin{aligned}
&\mathbb{E}[B^2 f(Y_1^*)] \\
&\quad = \mathbb{E}\Bigg[\alpha_2^2 B_1^2 f(Y_1^*) - 2\alpha_1\alpha_2 A(B_1 f)(Y_1^*) + \alpha_1^2 \Bigg[\sum_{k=1}^{M} \sigma_k^2(Y_1^*) f'''[\nu_k, \nu_k, b(Y_1^*)](Y_1^*) \\
&\qquad + f''[b(Y_1^*), b(Y_1^*)](Y_1^*) + \frac{1}{4} \sum_{k,j=1}^{M} \sigma_k^2(Y_1^*) \sigma_j^2(Y_1^*) f^{(4)}[\nu_k, \nu_k, \nu_j, \nu_j](Y_1^*)\Bigg]\Bigg] + O(h^2).
\end{aligned}$$

Next we apply Dynkin's formula repeatedly to obtain

$$\begin{aligned}
&\mathbb{E}[B^2 f(Y_1^*)] \\
&\quad = \mathbb{E}[B^2 f(y(\theta h))] \\
&\quad = \alpha_2^2 \Bigg[B_1^2 f(x_0) + B_1^3 f(x_0)\theta h + \int_0^{\theta h} \int_0^s \mathbb{E}[B_1^4 f(y(r))]\, \mathrm{d}r\, \mathrm{d}s\Bigg] \\
&\qquad - 2\alpha_1\alpha_2 \Bigg[A(B_1 f)(x_0) + B_1(A(B_1 f))(x_0)\theta h + \int_0^{\theta h}\int_0^s \mathbb{E}[B_1^2(A(B_1 f))(y(r))\mathrm{d}r\mathrm{d}s]\Bigg] \\
&\qquad + \alpha_1^2 \Bigg[(B_1^2 f)(x_0) + (B_1^3 f)(x_0)\theta h + \int_0^{\theta h} \int_0^s \mathbb{E}[B_1^4 f(y(r))\Bigg]\, \mathrm{d}r\, \mathrm{d}s] + O(h^2).
\end{aligned}$$

The detailed calculation shows that,

$$\mathbb{E}[B^2 f(Y_1^*)] = (B_1^2 f)(x_0) + (B_1^3 f)(x_0)\theta h + O(h^2). \tag{A.14}$$

Step 4. Proceeding in the same way as above, we have

$$\mathbb{E}(B^3 f)(Y_1^*) = (B_1^3 f)(x_0) + O(h). \tag{A.15}$$

Step 5. Combining (A.9), (A.13), (A.14), (A.15) and observing

$$\alpha_1 \theta(1-\theta)h^2 - \alpha_2 \theta(1-\theta)h^2 + \frac{(1-\theta)^2 h^2}{2} + \frac{\theta^2 h^2}{2} = \frac{h^2}{2}$$

and

$$\frac{\theta^3 h^3}{6} + \alpha_1 \frac{\theta^2(1-\theta)}{2} h^3 - \alpha_2 \frac{\theta^2(1-\theta)}{2} h^3 + \frac{\theta(1-\theta)^2}{2} h^3 + \frac{(1-\theta)^3}{6} h^3 = \frac{h^3}{6}$$

give the desired result. $\square$

Bibliography

Anderson, D. F. and Mattingly, J. C. (2011). A weak trapezoidal method for a class of stochastic differential equations. *Commun. Math. Sci.*, 9(1):301–318.

Asmussen, S. and Glynn, P. W. (2007). *Stochastic simulation: algorithms and analysis*, volume 57 of *Stochastic Modelling and Applied Probability*. Springer, New York.

Atkinson, K. E. (1989). *An introduction to numerical analysis*. John Wiley & Sons, Inc., New York, second edition.

Bruti-Liberati, N. and Platen, E. (2008). Strong predictor-corrector Euler methods for stochastic differential equations. *Stoch. Dyn.*, 8(3):561–581.

Ethier, S. N. and Kurtz, T. G. (1986). *Markov processes: Characterization and Convergence*. Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. John Wiley & Sons Inc., New York.

Gaines, J. G. and Lyons, T. J. (1997). Variable step size control in the numerical solution of stochastic differential equations. *SIAM J. Appl. Math.*, 57(5):1455–1484.

Higham, D. J. (2000a). A-stability and stochastic mean-square stability. *BIT*, 40(2):404–409.

Higham, D. J. (2000b). Mean-square and asymptotic stability of the stochastic theta method. *SIAM J. Numer. Anal.*, 38(3):753–769 (electronic).

Higham, D. J. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Rev.*, 43(3):525–546 (electronic).

Hüsler, J., Liu, R. Y., and Singh, K. (2002). A formula for the tail probability of a multivariate normal distribution and its applications. *Journal of multivariate analysis*, 82(2):422–430.

Karatzas, I. and Shreve, S. E. (1991). *Brownian motion and stochastic calculus*, volume 113 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition.

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

Kloeden, P. E. and Platen, E. (1992). *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin.

Kushner, H. J. (1967). *Stochastic stability and control*, volume 33 of *Mathematics in Science and Engineering*. Academic Press, New York.

Mao, X. and Yuan, C. (2006). *Stochastic differential equations with Markovian switching*. Imperial College Press, London.

Meyn, S. P. and Tweedie, R. L. (2009). *Markov chains and stochastic stability*. Cambridge University Press, Cambridge.

Milstein, G. N. (1995). *Numerical integration of stochastic differential equations*, volume 313 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht.

Øksendal, B. (2003). *Stochastic differential equations, An introduction with applications*. Universitext. Springer-Verlag, Berlin, sixth edition.

Saito, Y. and Mitsui, T. (1996). Stability analysis of numerical schemes for stochastic differential equations. *SIAM J. Numer. Anal.*, 33(6):2254–2267.

Saito, Y. and Mitsui, T. (2002). Mean-square stability of numerical schemes for stochastic differential systems. Vietnam Journal of Mathematics 30(SI):551-560.

Sauer, T. (2012). Numerical solution of stochastic differential equations in finance. In *Handbook of computational finance*, Springer Handb. Comput. Stat., pages 529–550. Springer, Heidelberg.

Talay, D. and Tubaro, L. (1990). Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991).

Walsh, J. B. (1986). *An introduction to stochastic partial differential equations.* Springer, Berlin.

Yin, G. G. and Zhu, C. (2010). *Hybrid Switching Diffusions: Properties and Applications*, volume 63 of *Stochastic Modelling and Applied Probability.* Springer, New York.

Curriculum Vitae

Education

Ph.D. Mathematics, University of Wisconsin-Milwaukee, WI, 2015.

M.S., Mathematics, Western Illinois University, Macomb, Illinois, 2011.

M.Ed. (First Division), Mathematics Education, Tribhuvan University, Nepal, 2004.

B.Ed. (First Division), Mathematics Education, Tribhuvan University, Nepal, 2002.

Research Interests

- Applied Probability, Statistics, Numerical Analysis, Design and evaluation of computational methods, Stochastic Computation, Mathematical Finance, Mathematics Education.

University Experience in Research

- Research Assistant, Department of Mathematics, University of Wisconsin-Milwaukee (2014-2015).

Employment

- Graduate Teaching Assistant, Department of Mathematics, University of Wisconsin Milwaukee (2011-2014).

- Graduate Assistant, Department of Mathematics, Western Illinois University (2009-2011)

- Mathematics Teacher, Fluorescent Higher Secondary School, Baniyatar, Kathmandu, Nepal (2004-2009).