CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

| | |
|---|---|
| Title | The role of adaptivity in a numerical method for the Cox-Ingersoll-Ross model |
| Authors | Maulana, Heru |
| Publication date | 2022-07-04 |
| Original Citation | Maulana, H. 2022. The role of adaptivity in a numerical method for the Cox-Ingersoll-Ross model. PhD Thesis, University College Cork. |
| Type of publication | Doctoral thesis |
| Rights | © 2022, Heru Maulana. - https://creativecommons.org/licenses/by-nc-nd/4.0/ |
| Download date | 2026-11-01 05:48:40 |
| Item downloaded from | https://hdl.handle.net/10468/13582 |



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Ollscoil na hÉireann, Corcaigh
**National University of Ireland, Cork**




# The role of adaptivity in a numerical method for the Cox-Ingersoll-Ross model

Thesis presented on behalf of
**Heru Maulana**
**117227380**

for the degree of
**Doctor of Philosophy**

**University College Cork**
**School of Mathematical Sciences**

**Head of School:** Dr. Kevin Hayes
**Supervisor:** Dr. Cónall Kelly

**2022**

# Contents

**Contents** iii

**List of Figures** vii

**List of Tables** ix

**1 Introduction** 1
- 1.1 Discussion of the problem . . . . . 1
- 1.2 Mathematical Preliminaries . . . . . 6

**2 An Adaptive Numerical Method** 9
- 2.1 Framework for a random mesh . . . . . 10
- 2.2 Adaptive timestepping strategy . . . . . 11
- 2.3 The adaptive numerical method with backstop . . . . . 13

**3 Main Results** 17
- 3.1 Strong convergence of the adaptive method . . . . . 17
- 3.2 On the positivity of the adaptive method . . . . . 28
  - 3.2.1 Probability of positivity over a single step . . . . . 28
  - 3.2.2 Probability of positivity over a full trajectory . . . . . 29

**4 Numerical simulation** 37
- 4.1 Strong convergence and efficiency . . . . . 38
- 4.2 Parameter dependence of the strong convergence rate . . . . . 40

**Bibliography** 45

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

*Cónall Kelly, on behalf of Heru Maulana*

# Abstract

We demonstrate the effectiveness of an adaptive explicit Euler method for the approximate solution of the Cox-Ingersoll-Ross model. This relies on a class of path-bounded timestepping strategies which work by reducing the stepsize as solutions approach a neighbourhood of zero. The method is hybrid in the sense that a convergent backstop method is invoked if the timestep becomes too small, or to prevent solutions from overshooting zero and becoming negative.

Under parameter constraints that imply Feller's condition, we prove that such a scheme is strongly convergent, of order at least 1/2. Control of the strong error is important for multi-level Monte Carlo techniques. Under Feller's condition we also prove that the probability of ever needing the backstop method to prevent a negative value can be made arbitrarily small. Numerically, we compare this adaptive method to fixed step implicit and explicit schemes, and a novel semi-implicit adaptive variant.

We observe that the adaptive approach leads to methods that are competitive in a domain that extends beyond Feller's condition, indicating suitability for the modelling of stochastic volatility in Heston-type asset models.

# List of Figures

3.1 Surface plots of probabilities of solutions of (2.2) with (2.3) and timestepping strategy (2.7) becoming negative over a single step, for $h_{\max} \in [0.01, 1]$ and $Y = y \in [h_{\min}, 1.5]$ representing the value of the solution before taking the step. We take $\rho = 2^6$. In **(a)** Assumption 1 holds. In **(b)** Assumption 1 does not hold. . . . . 30

4.1 Convergence and efficiency of methods applied to (1.1) with $\lambda = 0.05$, $\sigma = 0.2$, $Y_0 = 0.02$. $\kappa = 2$ in **(a)** and **(b)**, and $\kappa = 1$ in **(c)** and **(d)**. Reference lines of slope 1/2 and 1 are provided. . . . . 39

4.2 Estimation of strong convergence order for methods applied to (1.1) with $\lambda = 0.05$, $Y_0 = 0.02$, $\kappa = 2$ in **(a)**, and $\kappa = 0.2$ in **(b)**. Here, $a = \sigma^2/(2\kappa\lambda)$. In both cases, Assumption 1 holds to the left of the vertical line at $a = 0.25$, and Feller's condition holds to the left of the vertical line at $a = 1$. . . . . 42

4.3 Percentage of times the backstop was invoked in the production of Figure 4.2 in order to **(a)** avoid a negative value or **(b)** bound the timestep from below by $h_{\min}$. . . . . 43

# List of Tables

1.1 Explicit Euler-Maruyama variants. . . . . 3

3.1 Bounds on $h_{\max}$ ensuring positivity (without the use of the backstop) of trajectories of (2.11) with probability at least $1 - \varepsilon$ where the path-bounded timestepping strategy satisfies Definition 7. Assumption 1 is satisfied ($\kappa\lambda > 2\sigma^2$) for tables in the left column, and violated ($\kappa\lambda < 2\sigma^2$) for tables in the right column. . . . . 35

# Statement by the Supervisor

Heru Maulana was a lecturer in the Mathematics Department of the Universitas Negeri Padang in West Sumatra, Indonesia. In 2017 he was awarded a prestigious scholarship from the Indonesian Government's Endowment Fund for Education (LPDP) to attend UCC to pursue his doctorate in the School of Mathematical Sciences. He registered initially on October 1, 2017. Sadly, Heru fell ill at the end of his third year and passed away in December 2020.

The work contained in this thesis develops an adaptive explicit Euler Maruyama method for numerically solving the Cox-Ingersoll-Ross (CIR) model. It demonstrates theoretical strong convergence on a restricted parameter set and investigates numerically strong convergence across a broader parameter range, as well as the capacity of the method to preserve the domain of the true solution with high probability. Had Heru's study continued into his fourth year, he would have incorporated this method into the implementation of a Monte Carlo approach to the valuation of employee stock options where the underlying asset followed a Heston model with a CIR volatility process.

All material contained in Chapters 1–4 was published online as a peer-reviewed article in the Journal of Computational and Applied Mathematics in February 2022, with the print version to appear in August 2022. For the full reference see [19]. Heru presented this work in a School seminar to an audience of experts in March 2020. He had completed all coursework requirements for the degree. As Heru's supervisor I have edited it into thesis format and submit it to be examined for the posthumous award of the degree of PhD.

*Cónall Kelly, School of Mathematical Sciences, UCC*

# Acknowledgements

We wish to thank Professor Gabriel Lord for his valuable support and input into the project. We are also grateful to Professor Alexandra Rodkina, of the University of the West Indies at Mona, Jamaica, and Ms Fandi Sun, of Heriot-Watt University, Edinburgh, UK, for their helpful suggestions.

Dedicated to the memory of Heru Maulana.

# Chapter 1

# Introduction

## 1.1 Discussion of the problem

The Cox-Ingersoll-Ross (CIR) process, used for example in the pricing of interest rate derivatives and as a model of stochastic volatility, is given by the following Itô-type stochastic differential equation (SDE),

$$dX(t) = \kappa\left(\lambda - X(t)\right)dt + \sigma\sqrt{X(t)}dW(t),\ t \in [0,T];\ X(0) = X_0 > 0, \qquad (1.1)$$

where $W(t)$ is a Wiener Process, $\kappa$, $\lambda$, and $\sigma$ are positive parameters, and $T \in [0,\bar{T}]$ for some fixed $\bar{T} < \infty$. Solutions of (1.1) are almost surely (a.s.) non-negative: in general they can achieve a value of zero but will be reflected back into the positive half of the real line immediately. Moreover, if $2\kappa\lambda \geq \sigma^2$, referred to as the Feller condition, solutions will be a.s. positive. No closed form solution of (1.1) is available, though $X(t)$ has (conditional upon $X(s)$ for $0 \leq s < t$) a non-central chi-square distribution with $\lim_{t\to\infty}\mathbb{E}[X(t)] = \lambda$ and $\lim_{t\to\infty}\mathrm{Var}[X(t)] = \lambda\sigma^2/2\kappa$; see [8].

For Monte Carlo estimation, exact sampling from the known conditional dis-

tribution of $X(t)$ is possible but computationally inefficient and potentially restrictive if the Wiener process of (1.1) is correlated with that of another process: see [1, 6, 13, 24]. Consequently a substantial literature has developed on the efficient numerical approximation of solutions of (1.1); we now highlight the parts which are relevant to our analysis.

An approach that seeks to directly discretise (1.1) using some variant of the explicit Euler-Maruyama method leads to schemes of the form

$$\begin{aligned} \widetilde{V}_{n+1} &= g_0\left(\widetilde{V}_n\right) + \Delta t\kappa\left(\lambda - g_1\left(\widetilde{V}_n\right)\right) + \sigma\sqrt{g_2\left(\widetilde{V}_n\right)}\Delta W_n; \\ V_{n+1} &= g_3\left(\widetilde{V}_{n+1}\right); \quad \widetilde{V}_0 = V_0, \end{aligned} \tag{1.2}$$

for given functions $g_0$, $g_1$,$g_2$ and $g_3$. These functions are selected to ensure that the diffusion coefficient remains real-valued (so that (1.2) is well defined), and to preserve the non-negativity of solutions. This approach seeks to accommodate the non-Lipschitz (square-root) diffusion, which facilitates overshoot when the solutions are close to zero, but it introduces additional bias to the approximation. A survey of choices common in practice may be found in [24], and we present a similar selection in Table 1.1 using the convention $x^+ := \max\{0, x\}$. We highlight in particular the fully truncated method proposed in Lord et al [24]. While it was shown in that article that the method is strongly convergent in $L_1$, the rate of strong convergence has only been recently proved by Cozma & Reisinger [9], who demonstrated a strong order of convergence 1/2 in $L^p$, in the case where $2\kappa\lambda > 3\sigma^2$, for $2 \leq p < 2\kappa\lambda/\sigma^2 - 1$. This method preserves the positivity of the underlying solutions of (1.1), and the authors of [9] state that it is arguably the most widely used in practice.

An alternative approach is to transform (1.1) before discretisation to make

| Method | $g_0(x)$ | $g_1(x)$ | $g_2(x)$ | $g_3(x)$ |
|---|---|---|---|---|
| Explicit Euler | $x$ | $x$ | $x$ | $x$ |
| Partially Truncated [10] | $x$ | $x$ | $x^+$ | $x$ |
| Fully Truncated [24] | $x$ | $x^+$ | $x^+$ | $x^+$ |
| Higham & Mao [14] | $x$ | $x$ | $\lvert x\rvert$ | $x$ |

Table 1.1: Explicit Euler-Maruyama variants.

the diffusion coefficient globally Lipschitz. For example, applying the Lamperti transform $Y = \sqrt{X}$ yields an auxiliary SDE in $Y$ with a state independent and therefore globally Lipschitz diffusion, but a drift coefficient that is unbounded when solutions are in a neighbourhood of zero. This approach is effective: a fully implicit Euler discretisation over a uniform mesh that preserves positivity of solutions was proposed in [1] and shown to have uniformly bounded moments. A continuous time extension interpolating linearly between mesh points was shown to have strong $L_p$ order of convergence $1/2$ (up to a factor of $\sqrt{|\log(h)|}$) in [11] when $2\kappa\lambda > p\sigma^2$, a continuous-time variant based on the same implicit discretisation was shown to have strong $L_p$ convergence of order 1 when $4\kappa\lambda > 3p\sigma^2$ in [3], and in [7] a variant which discretised the transformed SDE for $Y$ with an explicit projection method was shown to give strong $L_2$ convergence of order 1 when $2\kappa\lambda > 5\sigma^2$, of order $1/2$ when $3\sigma^2 < 2\kappa\lambda \leq 5\sigma^2$, and with an order on the interval $(1/6, (2\kappa\lambda - \sigma^2)/(4\kappa\lambda + 2\sigma^2))$ when $2\sigma^2 < 2\kappa\lambda \leq 3\sigma^2$.

It is important to emphasise the distinction between weakly and strongly convergent numerical methods. Weakly convergence methods may be sufficient for the Monte Carlo estimation of some derivatives, and methods which converge weakly with high order for (1.1) are known, see for example [2]. However strongly convergent methods are required in order to take advantage of Multi-Level Monte Carlo variance reduction techniques; see [12].

In this article we show that a strongly convergent numerical scheme can be

constructed by an application of the Lamperti transform to (1.1) followed by an explicit Euler-Maruyama discretisation over a procedurally generated adaptive mesh. The purpose of the adaptivity is to manage the nonlinear drift response of the discrete transformed system (rather than local error control). A framework for this was introduced in [18] for SDEs with one-sided Lipschitz drift and globally Lipschitz diffusion and extended to allow for monotone coefficients and a Milstein-type discretisation in [17, 20] respectively. This framework imposes maximum and minimum timesteps $h_{\max}$ and $h_{\min}$ in a fixed ratio $\rho$ and requires the use of a backstop numerical method in the event that the timestepping strategy attempts to select a stepsize below $h_{\min}$. The introduction of [18] provides a comprehensive review of the adaptive literature for SDEs.

As in [20], we will use here path-bounded strategies, this time designed to increase the density of timesteps when solutions approach zero, and we additionally require the backstop method to retake a step when the adaptive strategy overshoots the singularity at zero in the transformed equation. This latter is carried out without discarding samples from the Brownian path (preserving the trajectory), and without bridging (preserving efficiency).

We prove, when $\kappa\lambda > 2\sigma^2$, that the order of strong convergence in $L_2$ is at least $1/2$. This parameter constraint implies the Feller condition and is technical, ensuring the finiteness of sufficiently many conditional inverse moments of solutions of (1.1) (as described by [5]). We separately prove that, under exactly the Feller condition, the probability of invoking the backstop method to avoid negative values can be made arbitrarily small by choosing $h_{\max}$ sufficiently small, and provide a practical method for doing so given a user defined tolerance level. The proof relies upon a finite partitioning of the sample space of trajectories induced by $h_{\max}$ and $h_{\min}$, which allows us to handle the randomness of the number of

timesteps via the Law of Total Probability.

Numerically we compare the convergence and efficiency of our hybrid adaptive method with a semi-implicit adaptive variant, the fixed-step explicit method due to [14], and the transformed implicit fixed-step method proposed and analysed in [1, 11, 3], examining the parameter dependence of the numerical order of convergence in each case. The numerical convergence rates of adaptive methods are seen to outperform those of fixed-step methods over the entire domain where Feller's condition holds. Indeed we observe polynomial orders of convergence beyond this domain, indicating that these methods are also applicable to modelling stochastic volatility processes, for example in a Heston model.

Our results extend naturally to variants of (1.1) with time dependent parameters (see Glasserman [13]) subject to the existence of inverse moments in that setting, and to shifted CIR models such as those found in [27] which allow for negative interest rates.

The structure of the thesis is as follows. In Chapter 1.2 we give the form of the SDE governing the Lamperti transform of (1.1), specify the constraints placed upon the parameters for the main strong convergence theorem, and examine the availability of conditional moment and inverse moment bounds under these constraints. In Chapter 2 we set up the framework for our random mesh, characterise the class of path-bounded timestepping strategies and define our adaptive numerical method. In Chapter 3 we present the two main theorems on strong convergence and positivity, providing illustrative examples in the latter case. Finally in Chapter 4 we numerically compare convergence and efficiency of several commonly used methods.

## 1.2 Mathematical Preliminaries

Throughout the thesis we let $(\mathcal{F}_t)_{t\geq 0}$ be the natural filtration of $W$. By using Itô's formula and applying the transformation $Y = \sqrt{X}$ we get,

$$dY(t) = \left(\frac{4\kappa\left(\lambda - X_t\right) - \sigma^2}{8\sqrt{X_t}}\right) dt + \frac{\sigma}{2} dW_t,\ t \in [0,T]; \quad Y(0) = \sqrt{X_0} \in \mathbb{R}^+.$$

By then setting,

$$\alpha = \frac{4\kappa\lambda - \sigma^2}{8},\ \beta = \frac{-\kappa}{2},\ \gamma = \frac{\sigma}{2},$$

we can write,

$$dY(t) = \left(\frac{\alpha}{Y(t)} + \beta Y(t)\right) dt + \gamma dW_t,\ t \in [0,T]; \quad Y(0) = \sqrt{X_0} \in \mathbb{R}^+, \tag{1.3}$$

where $f(y) = \alpha/y + \beta y$ is not globally Lipschitz continuous, but when $\alpha > 0$ it satisfies a one-sided Lipschitz condition with constant $\beta < 0$:

$$[f(x) - f(y)](x - y) \leq \beta(x - y)^2, \text{ for all } x, y \in \mathbb{R}^+,$$

which can be seen by noting that

$$f(x) - f(y) = (x - y)\left[\beta - \frac{\alpha}{xy}\right].$$

Meanwhile the diffusion coefficient $g(y) = \gamma$ is constant and therefore globally Lipschitz continuous. The SDE (1.3) has integral form

$$Y(t) = Y(0) + \int_0^t \left(\frac{\alpha}{Y(s)} + \beta Y(s)\right) ds + \int_0^t \gamma dW_t, \quad t \geq 0. \tag{1.4}$$

In order to ensure the a.s. positivity of solutions of (1.3) and the boundedness

of certain inverse moments of solutions of (1.3), we will also need the following assumption:

**Assumption 1.** *Suppose that*

$$\kappa\lambda > 2\sigma^2. \tag{1.5}$$

Eq. (1.5) implies the Feller Condition ($2\kappa\lambda \geq \sigma^2$; see, for example [25, Chapter 9.9.2, p. 308]), which ensures that solutions of (1.1), and therefore (1.3), remain positive with probability one:

$$\mathbb{P}[Y(t) > 0,\ t \geq 0] = 1.$$

Assumption 1 provides inverse moment bounds as follows:

**Lemma 2.** *Let $(Y(t))_{t\in[0,T]}$ be a solution of (1.3), where Assumption 1 holds, and let $0 \leq t < s \leq T$. For any $Y(0) > 0$, and for $1 \leq p \leq 6$, there exists $C(p,T) > 0$ such that*

$$\mathbb{E}\left[\frac{1}{Y(s)^p}\middle|\mathcal{F}_t\right] \leq \frac{C(p,T)}{Y(t)^p}, \quad a.s. \tag{1.6}$$

*Proof.* Let $(X(t))_{t\in[0,T]}$ be a solution of (1.1) where Assumption 1 holds. By Lemma A.1 in Bossy & Diop [5],

$$\mathbb{E}\left[\frac{1}{X(t)}\right] \leq \frac{e^{\kappa t}}{X_0} \quad \text{and} \quad \mathbb{E}\left[\frac{1}{X(t)^p}\right] \leq \frac{C(p,T)}{X_0^p}, \tag{1.7}$$

for some $C(p,T)$ and any $p$ such that $1 < p < \frac{2\kappa\lambda}{\sigma^2} - 1$. Assumption 1 ensures that $\frac{2\kappa\lambda}{\sigma^2} - 1 > 3$, and since $Y(t) = \sqrt{X(t)}$, (1.6) follows by Lemma A.1 in [5] as it applies to conditional expectations, the former requiring an additional application of Jensen's inequality to the first inequality in (1.7). □

We also need the following bounds on positive moments of solutions of (1.3),

which apply under Feller's condition and in particular under Assumption 1.

**Lemma 3.** *Let $(Y(t))_{t\in[0,T]}$ be a solution of (1.3), where Assumption 1 holds, and let $0 \leq t \leq T$. For any $Y(0) > 0$ and any $p > 0$, there exist constants $M_{1,p}, M_{2,p} < \infty$, such that*

$$\mathbb{E}\left[\sup_{u\in[0,T]} Y(u)^p \middle| \mathcal{F}_t\right] \leq M_{1,p}(1 + Y(t)^p), \quad a.s., \tag{1.8}$$

*and*

$$\mathbb{E}\left[\sup_{u\in[0,T]} Y(u)^p\right] \leq M_{2,p}. \tag{1.9}$$

*Proof.* The proof of (1.8) is an application of [5, Lemma 2.1] to conditional expectations requiring an invocation of Jensen's inequality when $0 < p < 4$. Eq. (1.9) is provided by [11, Lemma 3.2]. □

Finally, we will make frequent use of the following elementary inequalities: for $n \in \mathbb{N}$ and $a_1, \ldots, a_n \in \mathbb{R}$ and $p \geq 0$,

$$\sqrt{|a_1 + a_2|} \leq \sqrt{|a_1|} + \sqrt{|a_2|}; \tag{1.10}$$

$$|a_1 a_2| \leq \frac{1}{2}(a_1^2 + a_2^2); \tag{1.11}$$

$$(a_1 + \ldots + a_n)^p \leq n^p(|a_1|^p + \cdots + |a_n|^p). \tag{1.12}$$

# Chapter 2

# An Adaptive Numerical Method

The article [18] provides a framework within which to construct timestepping strategies for an adaptive explicit Euler-Maruyama numerical scheme applied to nonlinear Itô-type SDEs of the form

$$dX(t) = f(X(t))dt + g(X(t))dB(t), \quad t \in [0, T], \tag{2.1}$$

over a random mesh $\{t_n\}_{n\in\mathbb{N}}$ on the interval $[0, T]$ given by,

$$Y_{n+1} = Y_n + h_{n+1}f(Y_n) + g(Y_n)(W(t_{n+1}) - W(t_n)), \tag{2.2}$$

where $\{h_n\}_{n\in\mathbb{N}}$ is a sequence of random timesteps and $\{t_n = \sum_{i=1}^{n} h_i\}_{n=1}^{N}$ with $t_0 = 0$, so that $t_{n+1} > t_n$ for each $n$. The choice of indexing ensures consistency of notation between $t_n$ and $h_n$ and the random time step $h_{n+1}$ is determined by $Y_n$.

Our proposed timestepping strategy will reflect dynamical considerations specific to the transformed CIR model (1.3) corresponding to

$$f(y) = \frac{\alpha}{y} + \beta y, \qquad g(y) = \gamma. \tag{2.3}$$

## 2.1 Framework for a random mesh

**Definition 4** (See [23]). *Suppose that each member of the sequence $\{t_n\}_{n\in\mathbb{N}}$ is an $\mathcal{F}_t$-stopping time: i.e. $\{t_n \le t\} \in \mathcal{F}_t$, for all $t \ge 0$, where $(\mathcal{F}_t)_{t\ge 0}$ is the natural filtration of $W$. We may then define a discrete time filtration $\{\mathcal{F}_{t_n}\}_{n\in\mathbb{N}}$ by*

$$\mathcal{F}_{t_n} = \{A \in \mathcal{F} : A \cap \{t_n \le t\} \in \mathcal{F}_t\}, n \in \mathbb{N}.$$

**Assumption 5.** *Each $h_n$ is $\mathcal{F}_{t_{n-1}}$-measurable, and $N$ is a random integer such that,*

$$N = \max\{n \in \mathbb{N} : t_{n-1} < T\} \text{ and } t_N = T,$$

*and the length of maximum and minimum stepsizes satisfy $h_{\max} = \rho h_{\min}$, for some $1 < \rho < \infty$, and*

$$h_{\min} \le h_n \le h_{\max} \le 1. \tag{2.4}$$

**Remark 6.** *The lower bound $h_{\min}$ ensures that a simulation over the interval $[0, T]$ can be completed in a finite number of timesteps, and the upper bound $h_{\max}$ prevents stepsizes from becoming too large. The latter is used as a convergence parameter in our examination of the strong convergence of the adaptive method. The random variable $N$ cannot take values outside the finite set $\{N_{\min}, \dots, N_{\max}\}$, where $N_{\min} := \lfloor T/h_{\max} \rfloor$ and $N_{\max} := \lceil T/h_{\min} \rceil$.*

$\triangle W_{n+1} := W(t_{n+1} - W(t_n)$ is a Wiener increment over a random interval the length of which depends on $Y_n$, through which it depends on $\{W(s),\ s \in [0, t_n]\}$. Therefore $\triangle W_{n+1}$ is not independent of $\mathcal{F}_{t_n}$; indeed it is not necessarily normally distributed. Since $h_{n+1}$ is a bounded $\mathcal{F}_{t_n}$-stopping time and $\mathcal{F}_t$-measurable, then $W(t_{n+1}) - W(t_n)$ is $\mathcal{F}_{t_n}$-conditionally normally distributed, by Doob's optional sampling theorem (see for example [28]), and for all $p \ge 2$ there exists $\upsilon_p < \infty$

such that

$$\begin{aligned}
\mathbb{E}[W(t_{n+1}) - W(t_n)|\mathcal{F}_{t_n}] &= 0, \quad a.s.; \\
\mathbb{E}[|W(t_{n+1}) - W(t_n)|^2|\mathcal{F}_{t_n}] &= h_{n+1}, \quad a.s.; \\
\mathbb{E}\left[\left|\int_{t_n}^{s} dW(r)\right|^p \middle| \mathcal{F}_{t_n}\right] &= \upsilon_p |s - t_n|^{\frac{p}{2}}, \quad a.s.
\end{aligned} \tag{2.5}$$

## 2.2 Adaptive timestepping strategy

To ensure strong convergence, our strategy is to reduce the size of each timestep if discretised solutions attempt to enter a neighbourhood of zero. If we wish to control the likelihood of invoking the backstop to avoid negative values, we will also reduce the timestep when solutions grow large.

**Definition 7** (A path-bounded time-stepping strategy)**.** *Let $\{Y_n\}_{n\in\mathbb{N}}$ be a solution of* (2.2)*. We say that $\{h_n\}_{n\in\mathbb{N}}$ is a path-bounded time-stepping strategy for* (2.2) *if the conditions of Assumption 5 are satisfied and there exist real non-negative constants $0 \leq Q < R$ (where $R$ may be infinite if $Q \neq 0$) such that whenever $h_{\min} \leq h_n \leq h_{\max}$,*

$$Q \leq |Y_n| < R, \quad n = 0, \ldots, N-1. \tag{2.6}$$

We now give two examples of path-bounded strategies that are valid for (2.2), the first with $R$ infinite (which, in conjunction with a suitable backstop method, is sufficient to ensure strong convergence), and the second with $R$ finite (which is useful if we also wish to minimise the use of the backstop to ensure positivity).

**Lemma 8.** *Let $\{Y_n\}_{n\in\mathbb{N}}$ be a solution of* (2.2)*, and let $\{h_n\}_{n\in\mathbb{N}}$ be a time-stepping*

*strategy that satisfies Assumption 5. If* $\{h_n\}_{n\in\mathbb{N}}$ *satisfies, for some* $r \geq 1$,

$$h_{n+1} := \max\left(h_{\min}, h_{\max} \cdot \min\{1, |Y_n|^r\}\right), \quad n \in \mathbb{N}, \tag{2.7}$$

*or*

$$h_{n+1} = \max\left(h_{\min}, h_{\max} \cdot \min(|Y_n|^r, |Y_n|^{-r})\right), \quad n \in \mathbb{N}, \tag{2.8}$$

*then it is path-bounding for* (2.2) *in the sense of Definition 7.*

*Proof.* Suppose that (2.7) holds and so $h_{n+1} \geq h_{\min}$. When $|Y_n| < 1$,

$$h_{n+1} \leq h_{\max}|Y_n|^r \Leftrightarrow \frac{1}{|Y_n|^r} \leq \frac{h_{\max}}{h_{n+1}} \leq \frac{h_{\max}}{h_{\min}} = \rho, \quad n \in \mathbb{N},$$

and when $|Y_n| \geq 1$ it is obvious that $\frac{1}{|Y_n|} \leq 1$, so we also have $\frac{1}{|Y_n|^r} \leq \rho < \infty$. Hence, when using the strategy defined by (2.7),

$$|Y_n| \geq \frac{1}{\rho^{1/r}}, \quad \frac{1}{|Y_n|} \leq \rho^{1/r}, \quad n \in \mathbb{N},$$

so that (2.6) holds with $Q = 1/\rho^{1/r}$ and $R = \infty$.

We can similarly show that (2.8) is path-bounding for (2.2), with $Q = 1/\rho^{1/r}$ and $R = \rho^{1/r}$. □

Note that for the strategies defined by (2.7) and (2.8), solutions of (2.2) cannot enter the neigbourhood $(\pm 1/\rho^{1/r})$, and therefore terms of the sequence $(1/|Y_n|)_{n\in\mathbb{N}}$ are uniformly bounded from above. This has the effect of controlling inverse moments of the solutions of (2.2). Moreover for (2.7), when (2.3) holds,

$$f(Y_n^2) = \frac{\alpha}{|Y_n|^2} + \beta|Y_n|^2 \leq \alpha\rho^{1/r} + \beta|Y_n|^2,$$

and therefore that strategy is admissible in the sense of [18, Definition 2.2] with $R_1 = \alpha\rho^{1/r}$ and $R_2 = \beta$. Similarly, (2.8) is admissible with $R_1 = \alpha\rho^{1/r} + \beta\rho^{-1/r}$ and $R_2 = 0$.

## 2.3 The adaptive numerical method with backstop

We consider an adaptive scheme based upon the following explicit Euler-Maruyama discretisation of (1.3) over a random mesh given by,

$$Y_{n+1} = Y_n + h_{n+1}\left(\frac{\alpha}{Y_n} + \beta Y_n\right) + \gamma \Delta W_{n+1}. \tag{2.9}$$

where the timestep sequence is constructed according to (2.7). For $s \in [t_n, t_{n+1})$, the continuous version is given by

$$\widetilde{Y}(s) = Y_n + \int_{t_n}^{s}\left(\frac{\alpha}{Y_n} + \beta Y_n\right) dr + \gamma \int_{t_n}^{s} dW(r), \tag{2.10}$$

so that $\widetilde{Y}(t_n) = Y_n$ for each $n \in \mathbb{N}$.

We combine this scheme with a positivity-preserving backstop scheme that is to be applied if the timestepping strategy attempts to select a timestep below $h_{\min}$ (in which case we choose $h_{n+1} = h_{\min}$) or if the current selected timestep and subsequently observed Brownian increment $\triangle W_{n+1}$ would result in the approximation becoming negative. First, we define a map representing the explicit Euler scheme:

**Definition 9.** *Define the map* $\theta : \mathbb{R}^3 \to \mathbb{R}$ *such that*

$$\theta(y, z, h) := y + h\left(\frac{\alpha}{y} + \beta y\right) + \gamma z,$$

*so that, if* $\{Y_n\}_{n\in\mathbb{N}}$ *is defined by* (2.9)*, then*

$$Y_{n+1} = \theta(Y_n, \Delta W_{n+1}, h_{n+1}),\ n \in \mathbb{N}.$$

Next we characterise the map associated with the backstop method:

**Definition 10.** *Define the backstop map* $\varphi : \mathbb{R}^3 \to \mathbb{R}$ *so that, for* $s \in [t_n + h_{\min}, t_n + h_{\max}]$*, it satisfies*

$$\mathbb{E}\left[\left|\varphi\left(\bar{Y}_n, \int_{t_n}^{s} dW(r), \int_{t_n}^{s} dr\right) - Y(s)\right|^2 \middle| \mathcal{F}_{t_n}\right] - \left|\bar{Y}_n - Y(t_n)\right|^2$$
$$\leq C_1 \int_{t_n}^{s} \mathbb{E}\left[|\bar{Y}(r) - Y(r)|^2 | \mathcal{F}_{t_n}\right] dr + C_2 |s - t_n|^{3/2},\ n \in \mathbb{N},\ a.s., \quad (2.11)$$

*for some non-negative constants* $C_1$ *and* $C_2$*, independent of* $N$*, and*

$$\varphi\left(\bar{Y}_n, \int_{t_n}^{s} dW(r), \int_{t_n}^{s} dr\right) > 0 \quad a.s., \quad \bar{Y}_n > 0. \quad (2.12)$$

*where where* $\bar{Y}_n := \bar{Y}(t_n)$*, and* $\bar{Y}$ *is the continuous form of our hybrid scheme constructed in next, in Definition 11.*

**Definition 11.** *Define the sequence of functions* $\{(\bar{Y}(s))_{s\in[t_n,t_{n+1})}\}_{n\in\mathbb{N}}$ *obeying*

$$\begin{aligned}\bar{Y}(s) = \theta\left(\bar{Y}_n, \int_{t_n}^{s} dW(r), \int_{t_n}^{s} dr\right)\mathcal{I}_{\{h_{\min}<h_{n+1}\leq h_{\max}\}\cap\{Y_{n+1}>0\}} \\ + \varphi\left(\bar{Y}_n, \int_{t_n}^{s} dW(r), \int_{t_n}^{s} dr\right)\mathcal{I}_{\{h_{n+1}=h_{\min}\}\cap\{Y_{n+1}>0\}} \\ + \varphi\left(\bar{Y}_n, \int_{t_n}^{s} dW(r), \int_{t_n}^{s} dr\right)\mathcal{I}_{\{h_{\min}<h_{n+1}\leq h_{\max}\}\cap\{Y_{n+1}<0\}}, \quad (2.13)\end{aligned}$$

*for* $s \in [t_n, t_{n+1})$, $n \in \mathbb{N}$, *where* $\{h_n\}_{n\in\mathbb{N}}$ *satisfies the conditions of Assumption 5.*

In practice, rather than checking (2.11) directly, we use as our backstop a method that is known to be positivity preserving and strongly convergent of order at least 1/2. In Chapter 4 we use the transformed fully implicit method proposed by [1]; one could also choose the fully truncated method [24].

**Remark 12.** *Since the events* $\{Y_{n+1} < 0\}$ *and* $\{Y_{n+1} > 0\}$ *are* $\mathcal{F}_{t_{n+1}}$*-measurable but not* $\mathcal{F}_{t_n}$*-measurable, if a negative value of* $Y_{n+1}$ *is observed following a step of length* $h_{n+1}$ *we must retake the step using the backstop method, which will ensure positivity over that step by* (2.12)*. This introduces an element of backtracking into the algorithm, but as long as the originally computed stepsize* $h_{n+1}$ *and Brownian increment are retained we can stay on the same trajectory while avoiding the use of a Brownian bridge.*

*Theorem 18 in Chapter 3.2, illustrated by Example 19, demonstrates that it is always possible to choose* $h_{\max}$ *to ensure that this particular use of the backstop can be avoided with probability* $1 - \varepsilon$*, for arbitrarily small* $\varepsilon \in (0, 1)$*, on each trajectory.*

# Chapter 3

# Main Results

In this chapter, we first demonstrate strong convergence of solutions of (2.13) to those of (1.3) under Assumption 1 and a path-bounded timestepping strategy. Second, we investigate the likelihood that the adaptive part of the method generates a negative value (triggering the use of the backstop to ensure positivity) and show how $h_{\max}$ may be chosen to control the probability of this occurring.

## 3.1 Strong convergence of the adaptive method

**Lemma 13.** *Let $(Y(t))_{t\in[0,T]}$ be a solution of (1.3) and let $\{t_n\}_{n\in\mathbb{N}}$ be a random mesh such that each $t_n$ is an $\mathcal{F}_t$-stopping time. Fix $n \in \mathbb{N}$ and suppose that $t_n \leq s \leq T$, where $T \in [0, \bar{T}]$. Then, for any $1 \leq p \leq 6$, we have*

$$\mathbb{E}\left[|Y(s) - Y(t_n)|^p \middle| \mathcal{F}_{t_n}\right] \leq 2^p \gamma^p \upsilon_p |s - t_n|^{p/2} + \bar{L}_{n,p}|s - t_n|^p, \ a.s.,$$

*where*

$$\bar{L}_{n,p} := 2^{2p}\left(\alpha^p \frac{C(p,T)}{Y(t_n)^p} + |\beta|^p M_{1,p}(1 + Y(t_n)^p)\right)$$

*is an $\mathcal{F}_{t_n}$-measurable random variable with finite expectation, and $C(T)$, $M_{1,p}$ are the constants defined by* (1.6) *and* (1.8) *in the statements of Lemmas 2 and 3 respectively .*

*Proof.* Solutions of (1.3) satisfy the integral equation

$$Y(s) = Y(t_n) + \int_{t_n}^{s} \left( \frac{\alpha}{Y(u)} + \beta Y(u) \right) du + \int_{t_n}^{s} \gamma dW(u), \quad t_n \leq s \leq T,$$

and therefore

$$Y(s) - Y(t_n) = \int_{t_n}^{s} \left( \frac{\alpha}{Y(u)} + \beta Y(u) \right) du + \gamma \left( W(s) - W(t_n) \right), \quad t_n \leq s \leq T.$$

Using the triangle and Cauchy-Schwarz inequalities, and the elementary inequality (1.12) with $n = 2$,

$$\begin{aligned}
&|Y(s) - Y(t_n)|^p \\
&\leq 2^p \left| \int_{t_n}^{s} \left( \frac{\alpha}{Y(u)} + \beta Y(u) \right) du \right|^p + 2^p \gamma^p |W(s) - W(t_n)|^p \\
&\leq 2^p |s - t_n|^{p-1} \int_{t_n}^{s} \left| \frac{\alpha}{Y(u)} + \beta Y(u) \right|^p du + 2^p \gamma^p |W(s) - W(t_n)|^p \\
&\leq 2^{2p} |s - t_n|^{p-1} \left( \int_{t_n}^{s} \frac{\alpha^p}{Y(u)^p} du + \int_{t_n}^{s} |\beta|^p Y(u)^p du \right) \\
&\qquad\qquad + 2^p \gamma^p |W(s) - W(t_n)|^p,
\end{aligned}$$

for $s \in [t_n, T]$. Now apply conditional expectations on both sides with respect to

$\mathcal{F}_{t_n}$ and (2.5) to get,

$$\begin{aligned}\mathbb{E}\left[|Y(s)-Y(t_n)|^p\middle|\mathcal{F}_{t_n}\right] &\leq 2^p\gamma^p\mathbb{E}\left[|W(s)-W(t_n)|^p\middle|\mathcal{F}_{t_n}\right]\\ &+2^{2p}|s-t_n|^{p-1}\left(\mathbb{E}\left[\int_{t_n}^s\frac{\alpha^p}{Y(u)^p}du\middle|\mathcal{F}_{t_n}\right]+\mathbb{E}\left[\int_{t_n}^s|\beta|^pY(u)^pdu\middle|\mathcal{F}_{t_n}\right]\right)\\ &\leq 2^p\gamma^p\upsilon_p|s-t_n|^{p/2}+2^{2p}|s-t_n|^{p-1}\left(\alpha^p\int_{t_n}^s\frac{C(p,T)}{Y(t_n)^p}du\right.\\ &\left.+|\beta|^p\int_{t_n}^sM_{1,p}(1+Y(t_n)^p)du\right),\quad a.s,\end{aligned}$$

where we have used (1.6) and (1.8) from the statement of Lemma 2 at the last step. Therefore

$$\begin{aligned}\mathbb{E}\left[|Y(s)-Y(t_n)|^p\middle|\mathcal{F}_{t_n}\right] &\leq 2^p\gamma^p\upsilon_p|s-t_n|^{p/2}\\ &+2^{2p}\left(\alpha^p\frac{C(p,T)}{Y(t_n)^p}+\beta M_{1,p}(1+Y(t_n)^p)\right)|s-t_n|^p,\ a.s,\end{aligned}$$

as required.

□

**Lemma 14.** *Let $(Y(t))_{t\in[0,T]}$ be a solution of* (1.3) *and let Assumption 1 hold. Let $\{t_n\}_{n\in\mathbb{N}}$ arise from the adaptive timestepping strategy satisfying* (2.6) *in Definition 7 for some $0<Q<R$ , and formulate the Taylor expansion of $f(Y(s))$ around $Y(t_n)$, where $f$ is as given in* (2.3)*, as*

$$f(Y(s))=f(Y(t_n))+R_f(s,t_n,Y(t_n)),\quad s\in[t_n,t_{n+1}],\tag{3.1}$$

*where*

$$R_f(s,t_n,Y(t_n))\ =\ \int_0^1Df\left(Y(t_n)+\tau(Y(s)-Y(t_n))\right)(Y(s)\ -\ Y(t_n))d\tau.\tag{3.2}$$

*For any $1 \leq p \leq 3$, the $p^{th}$ conditional moment of $R_f(s, t_n, Y(t_n))$ satisfies*

$$\mathbb{E}\left[|R_f|^p \middle| \mathcal{F}_{t_n}\right] \leq K_{n,p} h_{n+1}^{p/2}, \quad a.s., \tag{3.3}$$

*where $K_{n,p}$ is an a.s. finite and $\mathcal{F}_{t_n}$-measurable random variable given by*

$$K_{n,p} = 2^p |\beta|^p \left(2^p \gamma^p \upsilon_p + \bar{L}_{n,p} h_{n+1}^{p/2}\right) + 2^p \frac{\alpha^p \sqrt{C(p,T)}}{Y(t_n)^{2p}} \left(2^p \gamma^p \upsilon_{2p}^{1/2} + \bar{L}_{n,2p}^{1/2} h_{n+1}^{p/2}\right).$$

*Moreover, there exists $K_p$ independent of $n$ such that*

$$K_p := \mathbb{E}[K_{n,p}] < \infty. \tag{3.4}$$

*Proof.* By direct substitution of $f(y)$ from (2.3) into (3.2), evaluating the integral in $\tau$, and taking the $p^{th}$-moment conditional upon $\mathcal{F}_{t_n}$, we get

$$\mathbb{E}\left[|R_f|^p \middle| \mathcal{F}_{t_n}\right] = \mathbb{E}\left[\left|(Y(s) - Y(t_n))\left(\beta - \frac{\alpha}{Y(s)Y(t_n)}\right)\right|^p \middle| \mathcal{F}_{t_n}\right].$$

Using the triangle inequality and (1.12) we get

$$\mathbb{E}\left[|R_f|^p \middle| \mathcal{F}_{t_n}\right] \leq 2^p |\beta|^p \mathbb{E}\left[|Y(s) - Y(t_n)|^p \middle| \mathcal{F}_{t_n}\right] + \frac{2^p \alpha^p}{Y(t_n)^p} \mathbb{E}\left[\left|\left((Y(s) - Y(t_n)) \cdot \frac{1}{Y(s)}\right)\right|^p \middle| \mathcal{F}_{t_n}\right]. \tag{3.5}$$

Next apply Lemma 13 followed by the Cauchy-Schwarz inequality to get

$$\begin{aligned}\mathbb{E}[|R_f|^p\big|\mathcal{F}_{t_n}] &\leq 2^p|\beta|^p h_{n+1}^{p/2}\left(2^p\gamma^p\upsilon_p+\bar{L}_{n,p}h_{n+1}^{p/2}\right)\\ &\quad+2^p\frac{\alpha^p}{Y(t_n)^p}\sqrt{\mathbb{E}[|Y(s)-Y(t_n)|^{2p}|\mathcal{F}_{t_n}]}\sqrt{\mathbb{E}\left[\frac{1}{|Y(s)|^{2p}}\middle|\mathcal{F}_{t_n}\right]}\\ &\leq 2^p|\beta|^p h_{n+1}^{p/2}\left(2^p\gamma^p\upsilon_p+\bar{L}_{n,p}h_{n+1}^{p/2}\right)\\ &\quad+2^p\frac{\alpha^p\sqrt{C(p,T)}}{Y(t_n)^{2p}}\sqrt{\mathbb{E}[|Y(s)-Y(t_n)|^{2p}|\mathcal{F}_{t_n}]}.\end{aligned}$$

Again applying Lemma 13 and the elementary inequality (1.10) this becomes

$$\begin{aligned}\mathbb{E}[|R_f|^p\big|\mathcal{F}_{t_n}] &\leq 2^p|\beta|^p h_{n+1}^{p/2}\left(2^p\gamma^p\upsilon_p+\bar{L}_{n,p}h_{n+1}^{p/2}\right)\\ &\quad+2^p h_{n+1}^{p/2}\frac{\alpha^p\sqrt{C(p,T)}}{Y(t_n)^{2p}}\left(2^p\gamma^p\upsilon_{2p}^{1/2}+\bar{L}_{n,2p}^{1/2}h_{n+1}^{p/2}\right),\end{aligned}$$

from which the statement of the Lemma follows when we observe that the a.s. finiteness of $K_{n,p}$ is ensured by Assumption 1, and (3.4) is ensured by Lemmas 2 & 3. □

**Remark 15.** *It is also possible to estimate the second expectation in* (3.5) *by an application of Itô's formula, rather than the Cauchy-Schwarz inequality. However, this increases the maximum number of finite inverse moments of* $Y$ *required from* $2p$ *to* $3p$ *and does not improve the order of the bound* (3.3).

**Lemma 16.** *Let* $(Y(t_n))_{t_n\in[0,T]}$ *be the solution of* (1.3) *with initial value* $Y(0)=Y_0=\sqrt{X_0}$. *Let* $\left(\widetilde{Y}(s)\right)_{s\in[t_n,t_{n+1}]}$ *be a solution of* (2.10) *over the interval* $[t_n,t_{n+1}]$ *and* $\{h_n\}_{n\in\mathbb{N}}$ *be a sequence of random timesteps defined by* (1.3) *and* $\{t_n=\sum_{i=1}^n h_i\}_{n=1}^N$ *with* $t_0=0$. *Then for* $n,p\in\mathbb{N}$, *there exists an* $\mathcal{F}_{t_n}$*-measurable*

*random variable* $\bar{K}_n$ *with finite expectation* $K_n := \mathbb{E}[\bar{K}_n] < \infty$ *such that*

$$\mathbb{E}\left[E(t_{n+1})^2 \middle| \mathcal{F}_{t_n}\right] - E(t_n)^2 \leq \int_{t_n}^{t_{n+1}} \mathbb{E}\left[E(r)^2 | \mathcal{F}_{t_n}\right] + \bar{K}_n h_{n+1}^2, \quad a.s., \tag{3.6}$$

*where the error* $E(s) := Y(s) - \widetilde{Y}(s)$, $s \in [t_n, t_{n+1}]$.

*Proof.* For $s \geq t_n$, we subtract (2.10) from (1.4) to get

$$\begin{aligned} E(s) &= Y(s) - \widetilde{Y}(s) \\ &= \left[Y(t_n) + \int_{t_n}^{s} f(Y(r))dr + \gamma \int_{t_n}^{s} dW(r)\right] \\ &\qquad - \left[Y_n + \int_{t_n}^{s} f(Y_n)dr + \gamma \int_{t_n}^{s} dW(r)\right] \\ &= E(t_n) + \int_{t_n}^{s} \tilde{f}(Y(r), Y_n)dr, \end{aligned} \tag{3.7}$$

where $f$ is defined as in (2.3) and $\tilde{f}(Y(r), Y_n) = f(Y(r)) - f(Y_n)$. Applying the Itô formula and setting $s = t_{n+1}$, we can write,

$$E(t_{n+1})^2 = E(t_n)^2 + 2\int_{t_n}^{t_{n+1}} E(r)\tilde{f}(Y(r), Y_n)dr.$$

By (3.1) in the statement of Lemma 14,

$$\tilde{f}(Y(r), Y_n) = \tilde{f}(Y(t_n), Y_n) + R_f(r, t_n, Y(t_n),$$

where $R_f$ is defined in (3.2). This, and an application of (1.11) gives

$$\begin{aligned}
&E(t_{n+1})^2 - E(t_n)^2 \\
&= 2\int_{t_n}^{t_{n+1}} E(r)R_f(r,t_n,Y(t_n))dr + 2\int_{t_n}^{t_{n+1}} E(r)\tilde{f}(Y(t_n),Y_n)dr \\
&\leq \int_{t_n}^{t_{n+1}} E(r)^2 dr + \int_{t_n}^{t_{n+1}} R_f(r,t_n,Y(t_n))^2 dr \\
&\qquad\qquad +2\int_{t_n}^{t_{n+1}} E(r)\tilde{f}(Y(t_n),Y_n)dr. \qquad (3.8)
\end{aligned}$$

Consider the third term on the RHS of (3.8), and substitute (3.7) into the integrand:

$$\begin{aligned}
&\int_{t_n}^{t_{n+1}} E(r)\tilde{f}(Y(t_n),Y_n)dr \\
&= E(t_n)\tilde{f}(Y(t_n),Y_n)\int_{t_n}^{t_{n+1}} dr + \tilde{f}(Y(t_n),Y_n)^2 \int_{t_n}^{t_{n+1}}\int_{t_n}^{r} du\,dr \\
&\qquad +\tilde{f}(Y(t_n),Y_n)\int_{t_n}^{t_{n+1}}\int_{t_n}^{r} R_f(u,t_n,Y(t_n))du\,dr \\
&\leq \beta\int_{t_n}^{t_{n+1}} E(t_n)^2 dr + \tilde{f}(Y(t_n),Y_n)^2 h_{n+1}^2 \\
&\qquad +\tilde{f}(Y(t_n),Y_n)\int_{t_n}^{t_{n+1}}\int_{t_n}^{r} R_f(u,t_n,Y(t_n))du\,dr \\
&\leq \tilde{f}(Y(t_n),Y_n)^2 h_{n+1}^2 \\
&\qquad +\tilde{f}(Y(t_n),Y_n)\int_{t_n}^{t_{n+1}}\int_{t_n}^{r} R_f(u,t_n,Y(t_n))du\,dr. \qquad (3.9)
\end{aligned}$$

Now substitute (3.9) into (3.8), to get

$$\begin{aligned}
E(t_{n+1})^2 - E(t_n)^2 &\leq \int_{t_n}^{t_{n+1}} E(r)^2 dr + \int_{t_n}^{t_{n+1}} R_f(r,t_n,Y(t_n))^2 dr \\
&+ 2\tilde{f}(Y(t_n),Y_n)^2 h_{n+1}^2 + 2\tilde{f}(Y(t_n),Y_n)\int_{t_n}^{t_{n+1}}\int_{t_n}^{r} R_f(u,t_n,Y(t_n))du\,dr. \qquad (3.10)
\end{aligned}$$

Apply expectations to both sides of (3.10), conditional upon $\mathcal{F}_{t_n}$, to get

$$\begin{aligned}\mathbb{E}\left[E(t_{n+1})^2|\mathcal{F}_{t_n}\right] - E(t_n)^2 \leq \int_{t_n}^{t_{n+1}} \mathbb{E}[E(r)^2|\mathcal{F}_{t_n}]dr \\ + \int_{t_n}^{t_{n+1}} \mathbb{E}[R_f(r,t_n,Y(t_n))^2|\mathcal{F}_{t_n}]dr + 2\tilde{f}(Y(t_n),Y_n)^2 h_{n+1}^2 \\ + 2\tilde{f}(Y(t_n),Y_n)\int_{t_n}^{t_{n+1}}\int_{t_n}^{r} \mathbb{E}[R_f(u,t_n,Y(t_n))|\mathcal{F}_{t_n}]du\,dr, \quad a.s.\end{aligned}$$

Apply the bound (3.3) in the statement of Lemma 14 with $p = 1, 2$ to get

$$\begin{aligned}\mathbb{E}\left[E(t_{n+1})^2|\mathcal{F}_{t_n}\right] - E(t_n)^2 \leq \int_{t_n}^{t_{n+1}} \mathbb{E}[E(r)^2|\mathcal{F}_{t_n}]dr \\ + (2\tilde{f}(Y(t_n),Y_n) + K_{n,2})h_{n+1}^2 + 2\tilde{f}(Y(t_n),Y_n)K_{n,1}h_{n+1}^{5/2}, \quad a.s.\end{aligned}$$

Since, by (2.4) in the statement of Assumption 5, $h_{n+1} \leq h_{\max} \leq 1$, the statement of the Lemma now follows, with $\bar{K}_n := 2\tilde{f}(Y(t_n),Y_n) + K_{n,2} + 2\tilde{f}(Y(t_n),Y_n)K_{n,1}$.

To see that $\mathbb{E}[\bar{K}_n] < \infty$, note that by (2.6) in Definition 7, and since $\beta < 0$, $f(Y_n) \leq \alpha/Q$. The finiteness of $\mathbb{E}[K_{n,2}]$ and $\mathbb{E}[K_{n,1}]$ is given by (1.6) in the statement of Lemma 2 with $p = 6, 3$ respectively, along with (1.8) in the statement of Lemma 3.

□

**Theorem 17.** *Let $(Y(t))_{t\in[0,T]}$ be the solution of (1.3) with initial value $Y(0) = Y_0 = \sqrt{X_0}$, and suppose that Assumption 1 holds. Let $(\bar{Y}(t))_{t\in[0,T]}$ be a solution of (2.13) with initial value $\bar{Y}(0) = Y(0)$ and path-bounded timestepping strategy $\{h_n\}_{n\in\mathbb{N}}$ satisfying the conditions of Definition 7 for some $0 < Q < R$, with $R$*

*possibly infinite. There exists $C > 0$, independent of $h_{\max}$, such that*

$$\mathbb{E}[|Y(T) - \bar{Y}(T)|^2] \leq C h_{\max}.$$

*Proof.* From (3.6) in the statement of Lemma 16, we have that when $h_{n+1} \geq h_{\min}$,

$$\mathbb{E}\left[E(t_{n+1})^2 | \mathcal{F}_{t_n}\right] - E(t_n)^2 \leq \int_{t_n}^{t_{n+1}} \mathbb{E}\left[E(r)^2 | \mathcal{F}_{t_n}\right] dr + \bar{K}_n h_{n+1}^2, \quad a.s. \tag{3.11}$$

Suppose that $h_{n+1} < h_{\min}$ and $Y_{n+1}$ is generated from $Y_n$ via an application of the backstop method over a single step of length $h_{\min}$. This corresponds to single application of the map $\varphi$ in Definition 11 and therefore the relation (2.11) holds.

We now combine (2.11) and (3.11) to generate a single one-step error estimate for the hybrid method given by (2.13). Define the positive constant $\Gamma_1 = C_1 \vee 1$ and $\mathcal{F}_{t_n}$-measurable random variable $\bar{\Gamma}_{n,2} = C_2 \vee \bar{K}_n$, for $n \in \mathbb{N}$. Noting again that, by (2.4) in the statement of Assumption 5, $h_{n+1} \leq h_{\max} \leq 1$, we see that (2.13) satisfies, on almost all trajectories,

$$\mathbb{E}\left[E(t_{n+1})^2 | \mathcal{F}_{t_n}\right] - E(t_n)^2 \leq \Gamma_1 \int_{t_n}^{t_{n+1}} \mathbb{E}\left[E(r)^2 | \mathcal{F}_{t_n}\right] dr + \bar{\Gamma}_{n,2} h_{n+1}^2. \tag{3.12}$$

Sum both sides of (3.12) over $n = 0, \ldots, N-1$ and take expectations:

$$\begin{aligned} &\mathbb{E}\left[\sum_{n=0}^{N-1} \left(\mathbb{E}\left[E(t_{n+1})^2 | \mathcal{F}_{t_n}\right] - E(t_n)^2\right)\right] \\ &\qquad \leq \Gamma_1 \mathbb{E}\left[\sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \mathbb{E}\left[E(r)^2 | \mathcal{F}_{t_n}\right] dr\right] + \mathbb{E}\left[\sum_{n=0}^{N-1} \bar{\Gamma}_{n,2} h_{n+1}^2\right]. \end{aligned} \tag{3.13}$$

Consider first the LHS of (3.13). Since $N$ is a $\mathcal{F}_{t_n}$-stopping time, the event

$\{N \leq n\} \in \mathcal{F}_{t_n}$. Moreover $N \leq N_{\max}$. So we can write

$$\begin{aligned}
&\mathbb{E}\left[\sum_{n=0}^{N-1}\left(\mathbb{E}\left[E(t_{n+1})^2|\mathcal{F}_{t_n}\right] - E(t_n)^2\right)\right] \\
&= \sum_{n=0}^{N_{\max}-1} \mathbb{E}\left[\left(\mathbb{E}\left[E(t_{n+1})^2|\mathcal{F}_{t_n}\right] - E(t_n)^2\right)\mathcal{I}_{\{N\geq n+1\}}\right] \\
&= \sum_{n=0}^{N_{\max}-1} \left(\mathbb{E}\left[E(t_{n+1})^2\mathcal{I}_{\{N\geq n+1\}}|\mathcal{F}_{t_n}\right] - E(t_n)^2\mathcal{I}_{\{N\geq n+1\}}\right) \\
&= \sum_{n=0}^{N_{\max}-1} \left(\mathbb{E}\left[E(t_{n+1})^2\mathcal{I}_{\{N\geq n+1\}}\right] - \mathbb{E}\left[E(t_n)^2\mathcal{I}_{\{N\geq n+1\}}\right]\right) \\
&= \mathbb{E}\left[E(t_N)^2\right] = \mathbb{E}\left[E(T)^2\right]. \qquad (3.14)
\end{aligned}$$

Similarly, the RHS of (3.13) can be written

$$\begin{aligned}
&\Gamma_1\mathbb{E}\left[\sum_{n=0}^{N-1}\int_{t_n}^{t_{n+1}} \mathbb{E}\left[E(r)^2|\mathcal{F}_{t_n}\right]dr + \sum_{n=0}^{N-1}\bar{\Gamma}_{n,2}h_{n+1}^2\right] \\
&\qquad = \Gamma_1\underbrace{\mathbb{E}\left[\sum_{n=0}^{N_{\max}-1}\int_{t_n}^{t_{n+1}} \mathbb{E}\left[E(r)^2|\mathcal{F}_{t_n}\right]\mathcal{I}_{\{N\geq n+1\}}dr\right]}_{(I)} \\
&\qquad\qquad + \underbrace{\mathbb{E}\left[\sum_{n=0}^{N_{\max}-1}\bar{\Gamma}_{n,2}h_{n+1}^2\mathcal{I}_{\{N\geq n+1\}}\right]}_{(II)}.
\end{aligned}$$

Consider first (I). Since $t_n$, $t_{n+1}$, and $\mathcal{I}_{\{N\geq n+1\}}$ are $\mathcal{F}_{t_n}$-measurable we can bring $\mathcal{I}_{\{N\geq n+1\}}$ inside the conditional expectation, and exchange the order of integration

and conditional expectation as follows

$$\begin{aligned}
&\mathbb{E}\left[\sum_{n=0}^{N_{\max}-1}\int_{t_n}^{t_{n+1}}\mathbb{E}\left[E(r)^2|\mathcal{F}_{t_n}\right]\mathcal{I}_{\{N\geq n+1\}}dr\right]\\
&= \sum_{n=0}^{N_{\max}-1}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\mathbb{E}\left[E(r)^2|\mathcal{F}_{t_n}\right]\mathcal{I}_{\{N\geq n+1\}}dr\right]\\
&= \sum_{n=0}^{N_{\max}-1}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}\mathbb{E}\left[E(r)^2\mathcal{I}_{\{N\geq n+1\}}|\mathcal{F}_{t_n}\right]dr\right]\\
&= \sum_{n=0}^{N_{\max}-1}\mathbb{E}\left[\mathbb{E}\left[\int_{t_n}^{t_{n+1}}E(r)^2\mathcal{I}_{\{N\geq n+1\}}dr\middle|\mathcal{F}_{t_n}\right]\right]\\
&= \sum_{n=0}^{N_{\max}-1}\mathbb{E}\left[\int_{t_n}^{t_{n+1}}E(r)^2\mathcal{I}_{\{N\geq n+1\}}dr\right]\\
&= \mathbb{E}\left[\sum_{n=0}^{N_{\max}-1}\int_{t_n}^{t_{n+1}}E(r)^2\mathcal{I}_{\{N\geq n+1\}}dr\right]\\
&= \mathbb{E}\left[\sum_{n=0}^{N_{\max}-1}\int_{t_n}^{t_{n+1}}E(r)^2\mathcal{I}_{\{r\leq T\}}dr\right]\\
&= \mathbb{E}\left[\int_0^T E(r)^2dr\right]=\int_0^T\mathbb{E}\left[E(r)^2\right]dr. \qquad (3.15)
\end{aligned}$$

Finally consider (II). We have, since $\mathbb{E}\left[\bar{\Gamma}_{n,2}\mathcal{I}_{\{N\geq n+1\}}\right]\leq\mathbb{E}\left[\bar{\Gamma}_{n,2}\right]\leq C_2\vee K_2=:\Gamma_2$ for all $n=0,\ldots,N_{\max}-1$, $N_{\max}=\lceil T/h_{\min}\rceil$, and by Assumption 2.4, $\rho h_{\min}=h_{\max}\leq 1$,

$$\begin{aligned}
\mathbb{E}\left[\sum_{n=0}^{N_{\max}-1}\bar{\Gamma}_{n,2}h_{n+1}^2\mathcal{I}_{\{N\geq n+1\}}\right] &\leq \mathbb{E}\left[h_{\max}^2\sum_{n=0}^{N_{\max}-1}\bar{\Gamma}_{n,2}\mathcal{I}_{\{N\geq n+1\}}\right]\\
&= h_{\max}^2\sum_{n=0}^{N_{\max}-1}\mathbb{E}\left[\bar{\Gamma}_{n,2}\mathcal{I}_{\{N\geq n+1\}}\right]\\
&\leq h_{\max}^2\Gamma_2N_{\max}\leq h_{\max}^2\Gamma_2\left(\frac{T}{h_{\min}}+1\right)\\
&\leq (\rho T+1)\Gamma_2h_{\max}. \qquad (3.16)
\end{aligned}$$

Substituting (3.14), (3.15), and (3.16) back into (3.13) we get

$$\mathbb{E}[E(T)^2] \leq \Gamma_1 \int_0^T \mathbb{E}[E(r)^2]dr + (\rho T + 1)\Gamma_2 h_{\max}.$$

Since this inequality holds if $T$ is varied continuously over $[0, \bar{T}]$, for any $\bar{T} < \infty$ (see [17] for a demonstration) an application of Gronwall's inequality gives the result. □

## 3.2 On the positivity of the adaptive method

Now, we assume Feller's condition ($2\kappa\lambda \geq \sigma^2$) to ensure that solutions of (1.1) remain a.s. positive, but we do not require that Assumption 1 holds.

### 3.2.1 Probability of positivity over a single step

Consider the timestepping strategy defined by (2.7) with $r = 1$, satisfying Definition 7 with $R = \infty$. The probability of solutions of (2.9) becoming negative after a single step with this strategy, and hence triggering a use of the backstop method, is given by

$$\mathbb{P}\left[Y_{k+1} < 0 | Y_k = y > 0\right] = \Phi\left(a(y)\right), \quad y > h_{\min},$$

where $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-s^2/2} ds$, and

$$a(y) = \frac{-y - \left(\frac{\alpha}{y} + \beta y\right) h_{\max}(1 \wedge y^r)}{\gamma \sqrt{h_{\max}(1 \wedge y^r)}}.$$

Figure 3.1 presents two surface plots of these one-step probabilities against $y$ and $h_{\max}$. Feller's condition is satisfied in both cases. In Figure 3.1 (a), when

Assumption 1 holds, we see that the probability of invoking the backstop to avoid a negative value is highest when $h_{\max}$ is large and $Y_n$ is close to or above 1, in which case the timestepping strategy will tend to select $h_{n+1}$ to be close to $h_{\max}$. This probability drops off rapidly as $h_{\max}$ reduces. In Figure 3.1 (b), when Assumption 1 does not hold, the highest probabilities of invoking the backstop for preserving positivity when $Y_n$ is close to $h_{\min}$. However the maximum probability is significantly lower than any seen in Figure 3.1 (a).

### 3.2.2 Probability of positivity over a full trajectory

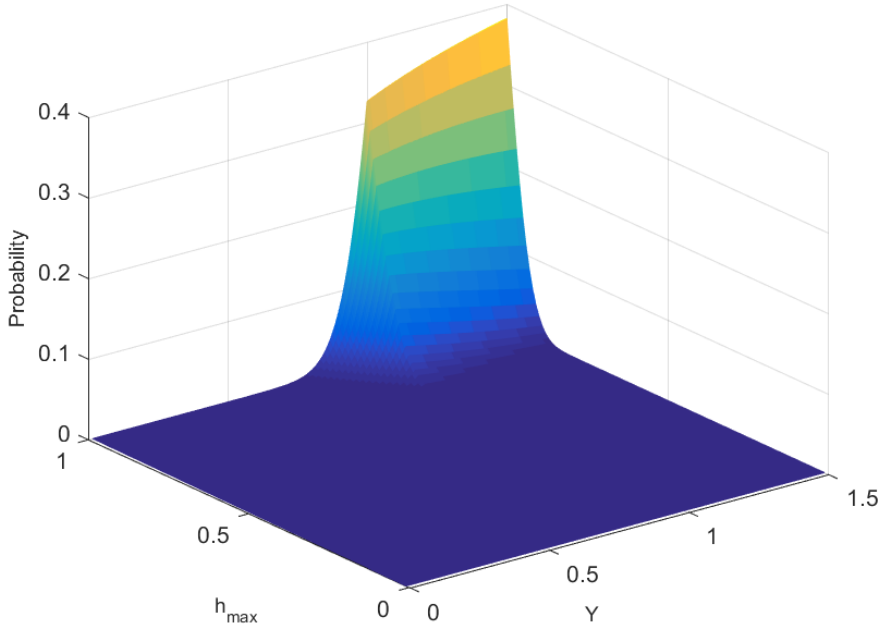
If we require path-bounded strategies where $R < \infty$, it is possible to derive an upper limit on $h_{\max}$ that is sufficient, over the entire trajectory, to keep the probability of needing the backstop scheme to prevent a negative value below some arbitrarily small tolerance. Our analysis reworks and extends the approach taken in the proof of [21, Theorem 4.3], using adaptive timestepping to handle unboundedness in the drift term. By allowing the use of the backstop to ensure a minimum timestep, along with an application of the Law of Total Probability, we can avoid fixing the random number of steps $N$.

**Theorem 18.** *Let $\{Y_n\}_{n=0}^N$ be a solution of (2.13), with initial value $Y_0 > 0$, evaluated on a random mesh $\{h_n\}_{n=1}^N$ satisfying the conditions of Definition 7 with $R < \infty$. Suppose also that $Y_0 \in (0, R)$. Then, for each $\varepsilon \in (0, 1)$ there exists $\bar{h}_{\max}(\varepsilon) > 0$ such that, for all $h_{\max} \in (0, \bar{h}_{\max}(\varepsilon))$*

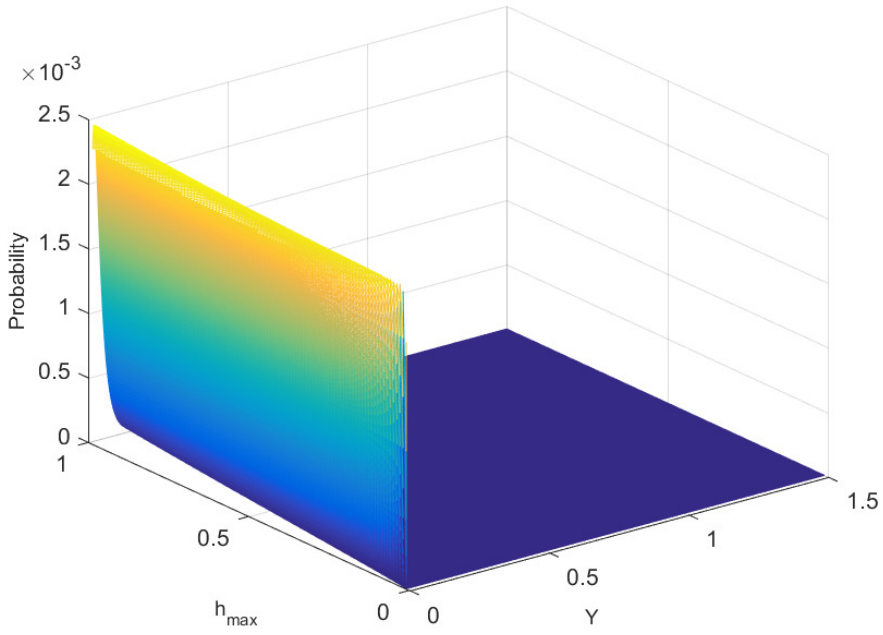
$$\mathbb{P}[\mathcal{R}_N] > 1 - \varepsilon,$$

*where $\mathcal{R}_N := \bigcap_{j=0}^N \{Y_j > 0\}$.*

*Proof.* Since, by (2.12), the backstop method will ensure positivity over a single




**(a)**: $\kappa\lambda > 2\sigma^2$, where $\sigma = 0.2$, $\lambda = 0.05$, $\kappa = 2$




**(b)**: $\kappa\lambda < 2\sigma^2$, where $\sigma = 0.2$, $\lambda = 0.05$, $\kappa = 1$

Figure 3.1: Surface plots of probabilities of solutions of (2.2) with (2.3) and timestepping strategy (2.7) becoming negative over a single step, for $h_{\max} \in [0.01, 1]$ and $Y = y \in [h_{\min}, 1.5]$ representing the value of the solution before taking the step. We take $\rho = 2^6$. In **(a)** Assumption 1 holds. In **(b)** Assumption 1 does not hold.

step if $h_{n+1} = h_{\min}$, the event $\{Y_{n+1} > 0\}$ is equivalent to the following union:

$$\left\{ \left\{ \frac{\triangle W_{n+1}}{\sqrt{h_{n+1}}} > -\frac{1}{\gamma} \left( \frac{Y_n}{h_{n+1}} + \frac{\alpha\sqrt{h_{n+1}}}{Y_n} + \beta\sqrt{h_{n+1}} Y_n \right) \right\} \cap \{h_{n+1} > h_{\min}\} \right\} \cup \{h_{n+1} = h_{\min}\}.$$

Moreover, when $Y_n > 0$ satisfies (2.6) in Definition 7,

$$\left\{ u \in -\frac{1}{\gamma} \left( \frac{Q}{h_{\max}} + \frac{\alpha\sqrt{h_{\min}}}{R} + \beta\sqrt{h_{\max}} R \right) \right\} \subseteq \left\{ u \in -\frac{1}{\gamma} \left( \frac{Y_n}{h_{n+1}} + \frac{\alpha\sqrt{h_{n+1}}}{Y_n} + \beta\sqrt{h_{n+1}} Y_n \right) \right\}. \quad (3.17)$$

For each $i = N_{\min}, \ldots N_{\max}$, define $\Omega_i := \{\omega \in \Omega : N(\omega) = i\}$, so that $\{\Omega\}_{i=N_{\min}}^{N_{\max}}$ is a finite partition of the sample space $\Omega$. On each $\Omega_i$ define the sequence of sub-events

$$\mathcal{R}_n(i) = \{Y_n > 0, Y_{n-1} > 0, \ldots, Y_1 > 0, Y_0 > 0\} \cap \Omega_i, \quad n = 0, 1, \ldots i.$$

Recall that the random variable $\triangle W_{n+1}/\sqrt{h_{n+1}}$ is distributed conditionally upon $\mathcal{F}_{t_n}$ like a standard Normal random variable. Moreover, $\mathcal{R}_n(i) \in \mathcal{F}_{t_n}$ for $n = 0, \ldots, i$ and $i = N_{\min}, \ldots, N_{\max}$. Let $\Phi$ denote the distribution function of a standard Normal random variable, and suppose $\{\xi_n\}_{n\in\mathbb{N}}$ is a sequence of mutually

independent standard normal random variables. Then

$$\begin{aligned}
&\mathbb{P}[Y_{n+1} > 0 | \mathcal{R}_n(i)] \\
&= \frac{\mathbb{P}[\{Y_{n+1} > 0\} \cap \{\mathcal{R}_n(i)\}]}{\mathbb{P}[\mathcal{R}_n(i)]} = \frac{\mathbb{E}\left[\mathbb{E}\left[I_{\{Y_{n+1}>0\}\cap\{\mathcal{R}_n(i)\}} | \mathcal{F}_{t_n}\right]\right]}{\mathbb{P}[\mathcal{R}_n(i)]} \\
&\geq \mathbb{E}\left[\mathbb{E}\left[I_{\{Y_{n+1}>0\}\cap\{\mathcal{R}_n(i)\}} | \mathcal{F}_{t_n}\right]\right] = \mathbb{E}\left[\mathbb{P}[\{Y_{n+1} > 0\} \cap \{\mathcal{R}_n(i)\} | \mathcal{F}_{t_n}]\right] \\
&\geq \mathbb{E}\left[\mathbb{P}[h_{n+1} = h_{\min} | \mathcal{F}_{t_n}]\right] \\
&\qquad + \mathbb{E}\left[\mathbb{P}\left[\frac{\triangle W_{n+1}}{\sqrt{h_{n+1}}} > -\frac{1}{\gamma}\left(\frac{Q}{h_{\max}} + \frac{\alpha\sqrt{h_{\min}}}{R} + \beta\sqrt{h_{\max}}R\right) \middle| \mathcal{F}_{t_n}\right]\right] \\
&\geq \mathbb{E}\left[\mathbb{P}\left[\frac{\triangle W_{n+1}}{\sqrt{h_{n+1}}} > -\frac{1}{\gamma}\left(\frac{Q}{h_{\max}} + \frac{\alpha\sqrt{h_{\min}}}{R} + \beta\sqrt{h_{\max}}R\right) \middle| \mathcal{F}_{t_n}\right]\right] \\
&= \mathbb{P}\left[\xi_{n+1} > -\frac{1}{\gamma}\left(\frac{Q}{h_{\max}} + \frac{\alpha\sqrt{h_{\max}}}{R\sqrt{\rho}} + \beta\sqrt{h_{\max}}R\right)\right] \\
&= 1 - \Phi\left(-\frac{1}{\gamma}\left(\frac{Q}{h_{\max}} + \sqrt{h_{\max}}\left(\frac{\alpha}{R\sqrt{\rho}} + \beta R\right)\right)\right) \\
&= \Phi\left(\frac{1}{\gamma}\left(\frac{Q}{h_{\max}} + \sqrt{h_{\max}}\left(\frac{\alpha}{R\sqrt{\rho}} + \beta R\right)\right)\right), \quad n = 0, \ldots, i-1.
\end{aligned}$$

Since $Y_0 > 0$ we have $\mathbb{P}[\mathcal{R}_0(i)|\Omega_i] = 1$ and therefore, since $\mathcal{R}_n(i) \subseteq \Omega_i$, $\Phi$ takes values on $[0, 1]$, and $i \leq N_{\max}$,

$$\begin{aligned}
\mathbb{P}[\mathcal{R}_i(i)|\Omega_i] \geq \mathbb{P}[\mathcal{R}_i(i)] &= \mathbb{P}\left[\bigcap_{n=0}^{i} \mathcal{R}_n(i)\right] \\
&= \prod_{n=1}^{i} \mathbb{P}[\mathcal{R}_n(i) | \mathcal{R}_{n-1}(i), \ldots, \mathcal{R}_0(i)] \\
&= \prod_{n=0}^{i-1} \mathbb{P}[Y_{n+1} > 0 | \mathcal{R}_n(i)] \\
&\geq \Phi\left(\frac{1}{\gamma}\left(\frac{Q}{h_{\max}} + \sqrt{h_{\max}}\left(\frac{\alpha}{R\sqrt{\rho}} + \beta R\right)\right)\right)^{i} \\
&\geq \Phi\left(\frac{1}{\gamma}\left(\frac{Q}{h_{\max}} + \sqrt{h_{\max}}\left(\frac{\alpha}{R\sqrt{\rho}} + \beta R\right)\right)\right)^{N_{\max}},
\end{aligned}$$

for $i = N_{\min}, \ldots, N_{\max}$. Multiplying through by $\mathbb{P}[\Omega_i]$ and applying the Law of

Total Probability by summing both sides over $i = N_{\min}, \ldots, N_{\max}$ gives

$$\begin{aligned}
\mathbb{P}[\mathcal{R}_N] &= \sum_{i=N_{\min}}^{N_{\max}} \mathbb{P}[\mathcal{R}_i(i)|\Omega_i]\mathbb{P}[\Omega_i] \\
&\geq \sum_{i=N_{\min}}^{N_{\max}} \Phi\left(\frac{1}{\gamma}\left(\frac{Q}{h_{\max}} + \sqrt{h_{\max}}\left(\frac{\alpha}{R\sqrt{\rho}} + \beta R\right)\right)\right)^{N_{\max}} \mathbb{P}[\Omega_i] \\
&= \Phi\left(\frac{1}{\gamma}\left(\frac{Q}{h_{\max}} + \sqrt{h_{\max}}\left(\frac{\alpha}{R\sqrt{\rho}} + \beta R\right)\right)\right)^{N_{\max}}.
\end{aligned}$$

Fix $\varepsilon \in (0,1)$, then for all $h_{\max} \in (0, \bar{h}_{\max}(\varepsilon))$, we have

$$\Phi\left(\frac{1}{\gamma}\left(\frac{Q}{h_{\max}} + \sqrt{h_{\max}}\left(\frac{\alpha}{R\sqrt{\rho}} + \beta R\right)\right)\right)^{N_{\max}} \geq 1 - \varepsilon. \tag{3.18}$$

To (3.18), apply the following inequality due to [26]

$$\frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-s^2/2} ds > \sqrt{1 - e^{-x^2/2}}, \quad x \in \mathbb{R}^+,$$

along with the fact that $N_{\max} = \rho T / h_{\max}$, leading us to seek $h_{\max}$ so that

$$\left(\frac{1}{2} + \frac{1}{2}\sqrt{1 - \exp\left(-\frac{\left(\frac{Q}{h_{\max}} + \sqrt{h_{\max}}\left(\frac{\alpha}{R\sqrt{\rho}} + \beta R\right)\right)^2}{2\gamma^2}\right)}\right)^{\frac{\rho T}{h_{\max}}} \geq 1 - \varepsilon.$$

Thus we derive the bound

$$\begin{aligned}
\tilde{h}_{\max}(\varepsilon) := \sup \Bigg\{ & \bar{h} \in (0,1) : \\
& \frac{Q}{h} + \sqrt{h}\left(\frac{\alpha}{R\sqrt{\rho}} + \beta R\right) \geq \sqrt{\ln\left(1 - (2(1-\varepsilon)^{\frac{h}{\rho T}} - 1)^2\right)^{-2\gamma^2}}, \, h \in (0, \bar{h}) \Bigg\}.
\end{aligned}$$

$\bar{h}_{\max}(\varepsilon)$ is uniquely defined for each $\varepsilon \in (0,1)$ because

$$g(h) := \frac{Q}{h} + \sqrt{h}\left(\frac{\alpha}{R\sqrt{\rho}} + \beta R\right) - \sqrt{-2\gamma^2 \ln\left(1 - (2(1-\varepsilon)^{h/\rho T} - 1)^2\right)} \quad (3.19)$$

is continuous on $\mathbb{R}^+$ with $\lim_{h\to 0^+} g(h) = \infty$, and therefore there is a neighbourhood of zero corresponding to $(0, \bar{h}_{\max}(\varepsilon))$ within which $g$ is positive. □

Note that if we extend the interval of simulation $[0,T]$ and keep $\varepsilon \in (0,1)$ fixed, there will be a corresponding increase in $N_{\max}$ in (3.18). This will lead to a reduction in the bound $\bar{h}_{\max}(\varepsilon)$, in a way that is characterised by (3.19). More generally $g(h)$, as defined by (3.19) in the proof of Theorem 18, provides a practical guide for choosing $h_{\max}$ in order to control the probability of invoking the backstop to avoid negative values.

**Example 19.** *Consider two adaptive timestepping strategies based on* (2.8) *with* $r = 1$ *and* $\rho = 2^6, 2^8$*, each used to simulate a single trajectory of* (1.3) *over the interval* $[0,1]$ *using the adaptive method* (2.13)*. In each case, we wish to choose* $h_{\max}$ *so that the probability of requiring the backstop in order to avoid negative values on that trajectory is less than* $\varepsilon$*, and this will hold for any* $Y_0 \in (h_{\min}, R)$*.*

*Table 3.1 shows the value of* $\bar{h}(\varepsilon)$ *for a range of tolerances* $\varepsilon$ *for parameter sets where Assumption 1 is satisfied (*$\kappa\lambda > 2\sigma^2$*), and where it is not (*$\kappa\lambda < 2\sigma^2$*). The resulting bounds on* $h_{\max}$ *are determined by substituting all parameters into* (3.19) *and solving* $g(h) = 0$ *for* $h$ *using the* `fsolve` *command in Maple with 20 digits of precision. We report the first 4 significant digits in each case, which is sufficient to illustrate the sensitivity of these bounds to the choice of* $\rho$ *and* $\varepsilon$*.*

$\rho = 2^6$, $Q = 0.015625$, $R = 64$

| | $\varepsilon$ | $\bar{h}_{\max}(\varepsilon)$ |
|---|---|---|
| $\sigma = 0.2$ | $10^{-2}$ | $3.594 \times 10^{-3}$ |
| $\lambda = 0.05$ | $10^{-4}$ | $3.547 \times 10^{-3}$ |
| $\kappa = 2$ | $10^{-6}$ | $3.506 \times 10^{-3}$ |

| | $\varepsilon$ | $\bar{h}_{\max}(\varepsilon)$ |
|---|---|---|
| $\sigma = 0.2$ | $10^{-2}$ | $5.454 \times 10^{-3}$ |
| $\lambda = 0.05$ | $10^{-4}$ | $5.341 \times 10^{-3}$ |
| $\kappa = 1$ | $10^{-6}$ | $5.246 \times 10^{-3}$ |

$\rho = 2^8$, $Q = 0.00390625$, $R = 256$

| | $\varepsilon$ | $\bar{h}_{\max}(\varepsilon)$ |
|---|---|---|
| $\sigma = 0.2$ | $10^{-2}$ | $5.800 \times 10^{-4}$ |
| $\lambda = 0.05$ | $10^{-4}$ | $5.755 \times 10^{-4}$ |
| $\kappa = 2$ | $10^{-6}$ | $5.716 \times 10^{-4}$ |

| | $\varepsilon$ | $\bar{h}_{\max}(\varepsilon)$ |
|---|---|---|
| $\sigma = 0.2$ | $10^{-2}$ | $8.912 \times 10^{-4}$ |
| $\lambda = 0.05$ | $10^{-4}$ | $8.804 \times 10^{-4}$ |
| $\kappa = 1$ | $10^{-6}$ | $8.710 \times 10^{-4}$ |

Table 3.1: Bounds on $h_{\max}$ ensuring positivity (without the use of the backstop) of trajectories of (2.11) with probability at least $1 - \varepsilon$ where the path-bounded timestepping strategy satisfies Definition 7. Assumption 1 is satisfied ($\kappa\lambda > 2\sigma^2$) for tables in the left column, and violated ($\kappa\lambda < 2\sigma^2$) for tables in the right column.

# Chapter 4

# Numerical simulation

Given (1.1) and its associated transformation (1.3), we compare our hybrid adaptive method (2.13), referred to in this chapter as `Explicit Adaptive (EA)`, to a natural semi-implicit variant constructed by replacing the update equation (2.9) with

$$Y_{n+1} = (1 - \beta h_{n+1})^{-1} \left[ Y_n + h_{n+1} \frac{\alpha}{Y_n} + \gamma \triangle W_{n+1} \right], \tag{4.1}$$

referred to in this chapter as `Semi-Implicit Adaptive (SIA)`. In both cases we will use the adaptive timestepping strategy given by (2.7) with $r = 1$ (note that we see similar results when $r = 2$). We also compare to three fixed step methods: the explicit discretisation of (1.1) analysed by [14] given by

$$X_{n+1} = X_n + h\kappa(\lambda - X_n) + \sigma\sqrt{|X_n|}\triangle W_{n+1},$$

referred to in this chapter as `Explicit Fixed (EF)`, the fully truncated method proposed by [24] given by

$$\widetilde{X}_{n+1} = \widetilde{X}_n + h\kappa(\lambda - \widetilde{X}_n^+) + \sigma\sqrt{\widetilde{X}_n^+}\triangle W_{n+1}; \quad X_{n+1} = \widetilde{X}_{n+1}^+; \quad \widetilde{X}_0 = X_0,$$

referred to in this chapter as `Fully Truncated (FT)`, and the drift implicit square root discretisation of (1.3) proposed and analysed in [1, 11, 3], given by

$$Y_{n+1} = \frac{Y_n + \gamma \triangle W_{n+1}}{2(1-\beta h)} + \sqrt{\frac{(Y_n + \gamma \triangle W_{n+1})^2}{4(1-\beta h)^2} + \frac{\alpha h}{1-\beta h}},$$

and referred to in this chapter as `Implicit Fixed (IF)`.

In the first part, we will compare the strong convergence of these methods in the mean stepsize, and the corresponding numerical efficiency. In the second part we explore the dependence on model parameters.

## 4.1 Strong convergence and efficiency

Throughout the section, we take $\rho = 2^6$. We solve using `EA` and `SIA` with values of $h_{\max} = 2^{-i}$, $i = 4, \ldots, 9$ and $M$ sample trajectories to estimate $\sqrt{\mathbb{E}\left[|X(T) - X_N|^2\right]}$, the root mean square error (RMSE), at a final time $T = 1$. To compute error estimates we first generate a reference solution using `IF` over a mesh with stepsize $h = 2^{-25}$, using a Brownian bridge to ensure values for the adaptive approximations are on the reference trajectory. To ensure that we are comparing adaptive and fixed step schemes of similar average cost, when solving using `IF`, `EF`, and `FT` we take as the fixed step $h_{\text{mean}}$ the average of all timesteps $h_n^{(m)}$ taken by `EA` over each path and each realisation $\omega_m$, $m = 1, \ldots, M$ so that

$$h_{\text{mean}} = \frac{1}{M} \sum_{m=1}^{M} \frac{1}{N^{(\omega_m)}} \sum_{n=1}^{N^{(\omega_m)}} h_n^{(\omega_m)}.$$

In Figure 4.1 we examine strong convergence for these methods by plotting RMSE against $h_{\text{mean}}$ with $M = 1000$ on a log-log scale, and efficiency by plotting RMSE against average compute time (cputime) again with $M = 1000$.

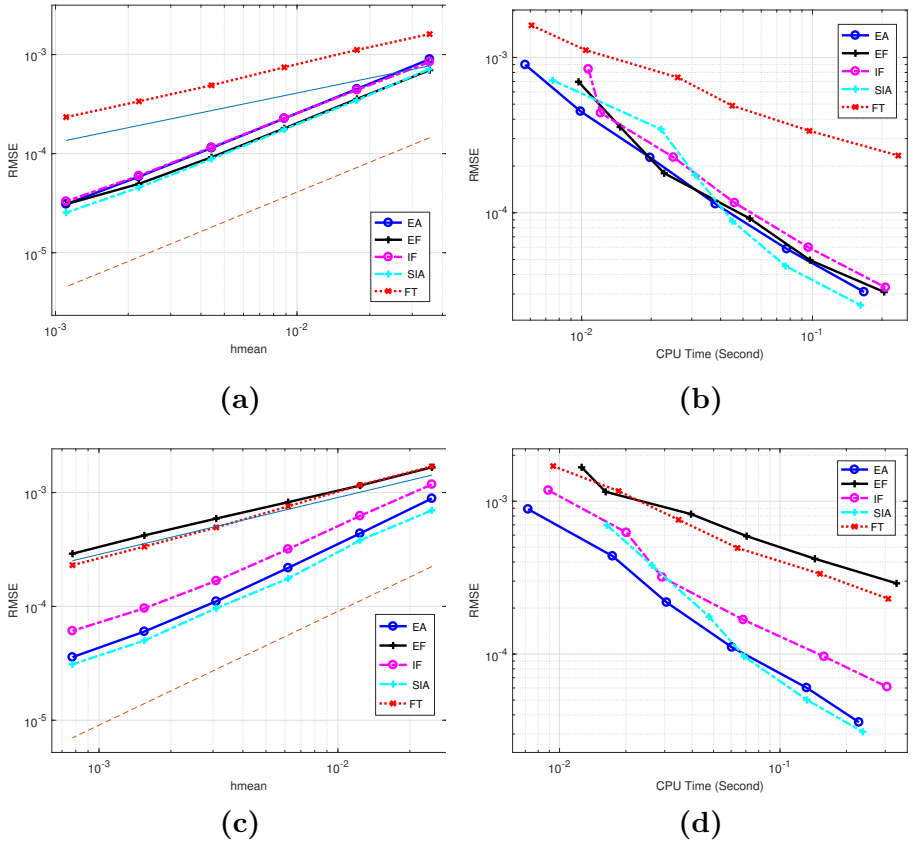



Figure 4.1: Convergence and efficiency of methods applied to (1.1) with $\lambda = 0.05$, $\sigma = 0.2$, $Y_0 = 0.02$. $\kappa = 2$ in **(a)** and **(b)**, and $\kappa = 1$ in **(c)** and **(d)**. Reference lines of slope 1/2 and 1 are provided.

In Figure 4.1 (a), Assumption 1 holds. The estimated error at each value of $h_{\text{mean}}$ is comparable for all methods except `FT`, and the numerical order appears to be close to one. For `FT`, the estimated error at each value of $h_{\text{mean}}$ is higher, and the numerical order appears closer to 1/2. In Figure 4.1 (b) we also see comparable efficiencies as measured by CPU time for this example, again with the exception of `FT`. In Figure 4.1 (c), Assumption 1 does not hold, and we see first that the numerical order of `EF` has reduced, and the estimated error at each value of $h_{\text{mean}}$ is lowest for `EA` and `SIA`, which also demonstrate the fastest CPU times in Figure 4.1 (d) for lower RMSE values.

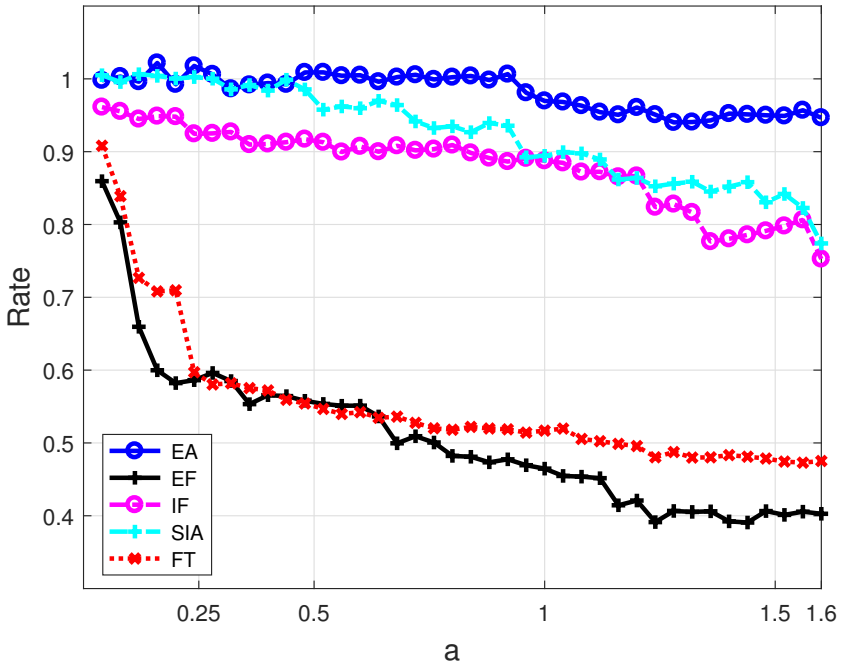
## 4.2 Parameter dependence of the strong convergence rate

Finally, we investigate numerically the dependence of the rate of strong convergence on the value of the parameter $a := \sigma^2/(2\kappa\lambda)$. Note that Assumption 1 corresponds to $a < 0.25$, and Feller's condition corresponds to $a \leq 1$. For 40 uniformly spaced values of $a$ in the interval $[0.04, 1.6]$, we numerically estimate the order of strong convergence in $L_2$ as the slope of the corresponding error over a range of values of $h_{\text{mean}}$ computed by generated strong convergence plots as in Figure 4.1 and using the `polyfit` command in MATLAB to estimate the order of strong convergence for each method. We make the following caveat: for $a \in (1, 1.6]$ all numerical schemes presented here are well-defined stochastic processes, even though the SDE (1.3) is not well defined in that parameter regime. We present the numerically estimated error in terms of $X$ outside the Feller regime using the reference solution generated by `IF`, though it is not known if `IF` converges in that regime with nonzero rate.

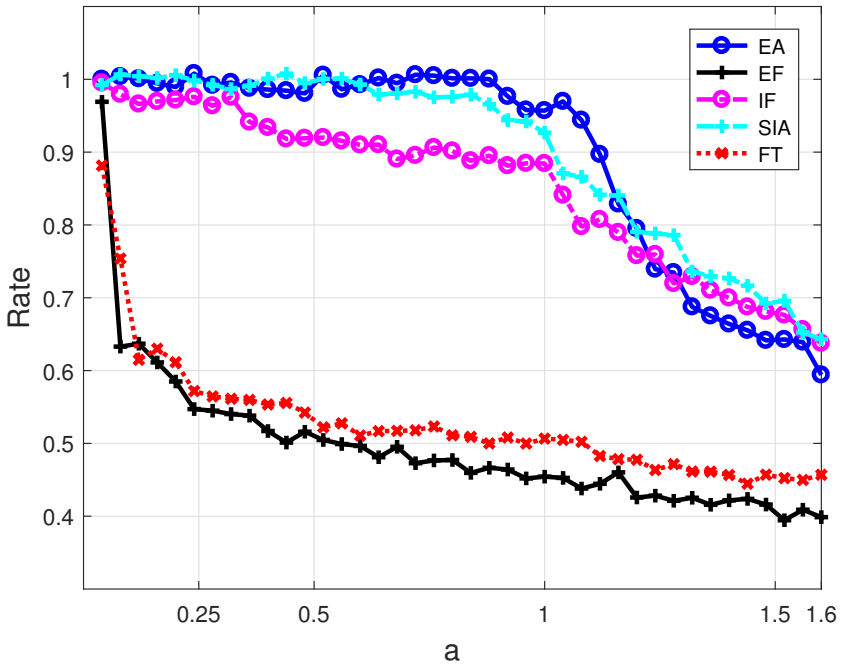
In Figure 4.2 we observe that `EA` maintains the highest numerical order of convergence: at or close to one while Feller's condition holds. The reduction in order outside of this region, which is visible for all methods, occurs more sharply in Figure 4.2 (b) when $\kappa$ is small. This is followed by `SIA`, which maintains a numerical order of convergence close to `EA` when Assumption 1 holds, but reduces more quickly outside this region. The difference is more pronounced in Figure 4.2 (a), and this may be because updates using the `SIA` method are subject to a damping factor $(1 - \beta h_{n+1})^{-1}$, as can be seen in (4.1), which has greater effect for larger values of $\kappa$. Finally, we observe for `EF` and `FT` an uptick in convergence rate as $a$ approaches zero, and this is consistent with the notion that it should display

order one convergence in the absence of noise.

Figure 4.3 demonstrates how frequently the backstop was invoked in the production of Figure 4.2, where we separately track usage to ensure positivity and usage to bound below the stepsize at $h_{\min}$. We see that when Assumption 1 holds, we do not require the backstop to avoid negative values, though as we move to the boundary of that region we do start to use it to bound the stepsize at $h_{\min}$ for a small proportion of steps. Note also that usage to avoid negative values increases with $a$ when $\kappa = 2$ only, whereas usage to bound the timestep increases with $a$ for both $\kappa = 2, 0.2$, more rapidly in the latter case.

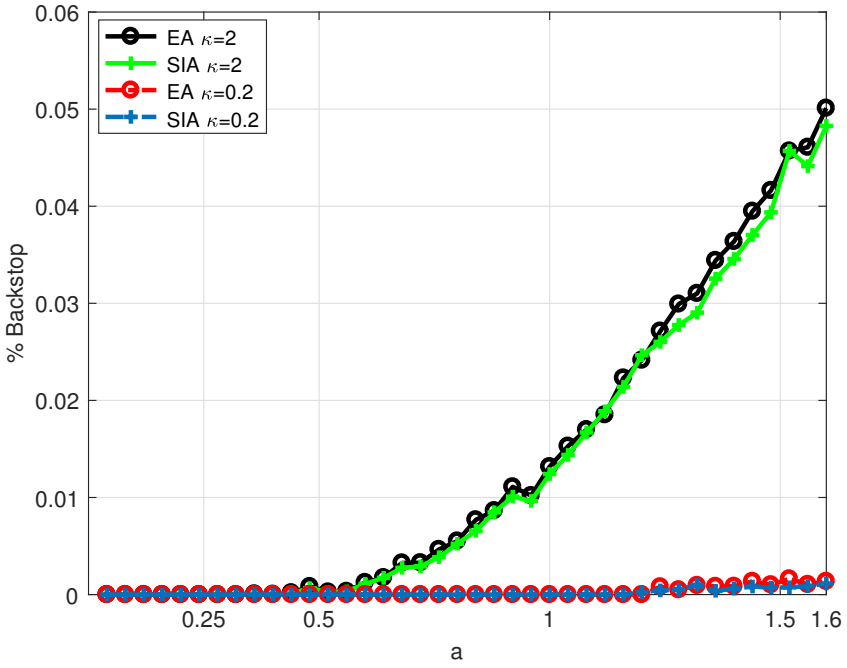



**(a)**

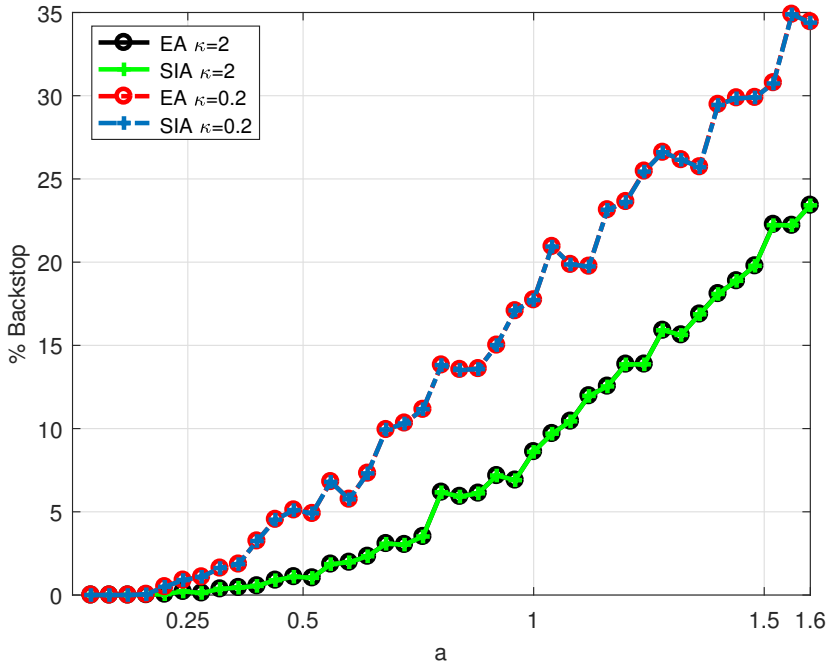



**(b)**

Figure 4.2: Estimation of strong convergence order for methods applied to (1.1) with $\lambda = 0.05$, $Y_0 = 0.02$, $\kappa = 2$ in **(a)**, and $\kappa = 0.2$ in **(b)**. Here, $a = \sigma^2/(2\kappa\lambda)$. In both cases, Assumption 1 holds to the left of the vertical line at $a = 0.25$, and Feller's condition holds to the left of the vertical line at $a = 1$.




**(a)**: Avoid $Y_n < 0$




**(b)**: Ensure $h_n = h_{\min}$

Figure 4.3: Percentage of times the backstop was invoked in the production of Figure 4.2 in order to **(a)** avoid a negative value or **(b)** bound the timestep from below by $h_{\min}$.

# Bibliography


[1] Alfonsi, A. (2005). On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods and Applications*, **11**(4), pp. 355–384.

[2] Alfonsi, A. (2010). High order discretization schemes for the CIR process: application to Affine Term Structure and Heston models. *Mathematics of Computation, American Mathematical Society*, **79**(269), pp. 209–237.

[3] Alfonsi, A. (2013). Strong order one convergence of a drift implicit Euler scheme: Application to the CIR process. *Statistics and Probability Letters*, **83**(2), pp. 602–607.

[4] Berkaoui, A., Bossy, M., and Diop, A. (2008). Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM Probability and Statistics*, **12**(1), pp. 1–11.

[5] Bossy, M., and Diop, A. (2004). An efficient discretisation scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^\alpha$, $\alpha \in [1/2, 1]$. *INRIA working paper*, no. 5396.

[6] Broadie, M., and Kaya, O. (2006). Exact Simulation of Stochastic Volatility and Other Affine Jump Diffusion Processes. *Operations Research*, **54**(2), pp. 217–231.

[7] Chassagneux, J.-F., Jacquier, A., and Mihaylov, I. (2016), An explicit Euler scheme with strong rate of convergence for financial SDEs with non-Lipschitz coefficients. SIAM Journal on Financial Mathematics, **7**(1), pp. 993–1021.

[8] Cox, J. C., Ingersoll, J. E., and Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, **53**(2), pp. 385–407.

[9] Cozma, A., and Reisinger, C. (2020). Strong order 1/2 convergence of full truncation Euler approximations to the Cox-Ingersoll-Ross process. *IMA Journal of Numerical Analysis*, **40**(1), pp. 358–376.

[10] Deelstra, G., and Delbaen, F. (1998). Convergence of Discretized Stochastic (Interest Rate) Process with Stochastic Drift Term. *Applied Stochastic Models and Data Analysis*, **14**(1), pp. 77–84.

[11] Dereich, S., Neuenkirch, A., and Szpruch, L. (2012). An Euler-Type method for the strong approximation of the Cox-Ingersoll-Ross process. *Proceedings of The Royal Society A: Mathematical Physical and Engineering Sciences*, **468**, pp. 1105–1115.

[12] Giles, M. B. (2015). Multilevel Monte Carlo methods. *Acta Numerica*, **24**, pp. 259–328.

[13] Glassermann, P. (2008). *Monte Carlo Methods in Financial Engineering*. New York: Springer.

[14] Higham, D. J., and Mao, X. (2005). Convergence of Monte Carlo simulations involving the mean reverting square root process. *Journal of Computational Finance*, **8**(3), pp. 35–61.

[15] Hutzenthaler, M., Jentzen, A., and Kloeden, P. E. (2012). Strong Convergence of An Explicit Numerical Method for SDEs with Non-Globally Lipschitz Continuous Coefficients. *Annals of Applied Probability*, **22**(4), pp. 1611–1641.

[16] Jafari, M. A., and Abbasian, S. (2017). The Moments for Solution of the Cox-Ingersoll-Ross Interest Rate Model. *Journal of Finance and Economics*, **5**(1), pp. 34–37.

[17] Kelly, C., and Lord, G. J. (2022). Adaptive Euler methods for stochastic systems with non-globally Lipschitz coefficients. *Numerical Algorithms*, **89**, pp. 721–747.

[18] Kelly, C., and Lord, G. J. (2018). Adaptive timestepping strategies for nonlinear stochastic systems. *IMA Journal of Numerical Analysis*, **38**(3), pp. 1523–1549.

[19] Kelly, C., Lord, G. J., and Maulana, H. (2022). The role of adaptivity in a numerical method for the Cox-Ingersoll-Ross model. *Journal of Computational and Applied Mathematics*, **410**(15), 114208.

[20] Kelly, C., Lord, G. J., and Sun, F. (2019). Strong convergence of an adaptive time-stepping Milstein method for SDEs with one-sided Lipschitz drift. arXiv:1909.00099, 20 pages.

[21] Kelly, C., Rodkina, A., and Rapoo, E. M. (2018). Adaptive timestepping for pathwise stability and positivity of strongly discretised nonlinear stochastic differential equations. *Journal of Computational and Applied Mathematics*, **334**, pp. 39–57.

[22] Kloeden, P. E., and Platen, E. (2011). *Numerical Solution of Stochastic Differential Equations*, Stochastic Modeling and Applied Probability. Berlin Heidelberg: Springer.

[23] Liu, W., and Mao, X. (2017). Almost sure stability of the Euler-Maruyama method with random variable stepsize for stochastic differential equations. *Numerical Algorithms*, **74**(2), pp. 573–592.

[24] Lord, R., Koekkoek R., and Van Dijk, D. (2010). A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, **10**(2), pp. 177–194.

[25] Mao, X. (2007). *Stochastic Differential Equations and Applications (2nd Edition)*. Horwood Publishing, Chichester, UK.

[26] Sasvari, Z., and Chen, H. (1999). Tight Bounds for the Normal Distribution: 10611. *The American Mathematical Monthly*, **106**(1), p. 76–76.

[27] Orlando, G., Mininni, R., and Bufalo, M. (2019). Interest rates calibration with a CIR model. *Journal of Risk Finance*, **20**(4), pp. 370–387.

[28] Shiryaev, A. N. (1996). *Probability (2nd edition)*. Berlin: Springer.

[29] Stuart, A. M., and Humphries, A. R. (1996). *Dynamical Systems and Numerical Analysis*. Cambridge University Press.

[30] Zeytun, S., and Gupta, A. (2007). *A Comparative Study of the Vasicek and the CIR Model of the Short Rate*. Berichte des Fraunhofer ITWM. Kaiserslautern, Germany.