Numerical Analysis of Stochastic Differential Equations with Applications in Financial Mathematics and Molecular Dynamics

Ioannis S. Stamatiou

Numerical Analysis of Stochastic Differential Equations with Applications in Financial Mathematics and Molecular Dynamics

Ioannis S. Stamatiou

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgments.

Βεβαιώνω ότι εκτός ειδικής αναφοράς σε εργασίες τρίτων, τα περιεχόμενα αυτής της διατριβής είναι πρωτότυπα και δεν έχουν κατατεθεί αλλού όλα ή εν μέρει για οποιοδήποτε άλλο βαθμό ή πιστοποίηση ή σε άλλο πανεπιστήμιο. Αυτή η διατριβή είναι δικό μου πόνημα και δεν περιέχει αποτελέσματα εργασιών σε συνεργασία με τρίτους, εκτός και αν αναφέρεται μέσα στο κείμενο και στις Ευχαριστίες.

Numerical Analysis of Stochastic Differential Equations with Applications in Financial Mathematics and Molecular Dynamics

Ioannis Sarantis Stamatiou

A thesis presented for the degree of
Doctor of Philosophy

Mathematics: Track in Statistics and Actuarial - Financial Mathematics
University of the Aegean, Greece
2nd March, 2016

Supervisor:	Associate Prof. N. Halidias
Panel:	Assistant Prof. S. Hatjispyros
	Assistant Prof. C. Kountzakis
	Assistant Prof. S. Xanthopoulos
	Prof. Dr A. Neuenkirch (Universität Mannheim)
	Prof. Dr. M.A. Peletier (Technische Universiteit Eindhoven)
	Prof. Dr. L. Szpruch (The University of Edinburgh)

.... But there are folks who want to know, and aren't afraid to look and won't turn tail should they find it - and if they never do, they'll have a good time anyway because nothing, neither the terrible truth nor the absence of it, is going to cheat them out of one honest breath of earth's sweet gas.

- Tom Robbins, *Still Life with Woodpecker*,
Bantam Books (1980), somewhere in Section 92]

ABSTRACT

In this thesis we are interested in the numerical solution of stochastic differential equations (SDE) with solutions in a certain domain. Our goal is to construct explicit numerical schemes that preserve that domain, mainly for cases where the coefficients of the SDEs are non-linear.

It is well known that the forward Euler scheme diverges on super-linear problems and the tamed Euler method does not necessarily preserve the structure of the original problem.

We propose a new numerical scheme, using the semi-discrete method, for various classes of stochastic differential equations. For some super-linear problems (like the Heston 3/2-model) as well as sub-linear (like the CEV model), which appear in the field of financial mathematics, we are able to construct a positivity preserving scheme. Moreover, we apply our method to problems arising in the field of molecular dynamics, where our structure preserving scheme is able to approximate effectively some SDEs which appear after a coarse graining procedure.

We also consider the case of Stochastic Delay Differential Equations (SDDEs) with non-negative solutions. Again we aim for explicit numerical schemes that preserve positivity. We expand the semi-discrete method from the Stochastic Ordinary Differential Equations (SODE) setting and apply it to the constant delay case, for which we prove strong convergence (DGBM model).

Numerical experiments support our theoretical results.

Keywords : Semi-Discrete Method, Super-Linear Drift and Diffusion, Hölder Continuous, 3/2-Model, Order of Convergence, Explicit Numerical Scheme, Mean-Reverting CEV Process, Positivity Preserving, Strong Approximation Error, Stochastic Volatility Model, Stochastic Differential Equations, Stochastic Delay Differential Equations, Monte Carlo Simulation, Numerical Methods

AMS subject classification 2010: 65C30, 65C20, 60H10, 60H35, 65J15

JEL classification: C15, C63, G13

Περίληψη

Σε αυτή τη διατριβή αντικείμενο έρευνας είναι η αριθμητική επίλυση στοχαστικών διαφορικών εξισώσεων (ΣΔΕ), οι οποίες έχουν λύση σε ένα συγκεκριμένο χωρίο. Ο στόχος μας ειναι η κατασκευή άμεσων αριθμητικών σχημάτων τα οποία διατηρούν αυτό το χωρίο, κυρίως σε περιπτώσεις όπου οι συντελεστές των ΣΔΕ είναι μη-γραμμικοί.

Είναι γνωστό ότι το *με βήμα προς τα εμπρός* σχήμα Euler αποκλίνει σε υπεργραμμικά προβλήματα και η *ελεγχόμενη* μέθοδος Euler δε διατηρεί απαραίτητα τη δομή του αρχικού προβλήματος.

Προτείνουμε ένα νέο αριθμητικό σχήμα, χρησιμοποιώντας την Ημι-Διακριτή μέθοδο, για διάφορες κλάσεις στοχαστικών διαφορικών εξισώσεων. Για κάποια υπεργραμμικά προβλήματα (όπως το Heston 3/2-μοντέλο) καθώς και για υπογραμμικά (όπως το CEV μοντέλο), τα οποία εμφανίζονται στο πεδίο των χρηματοοικονομικών μαθηματικών, κατασκευάζουμε ένα αριθμητικό σχήμα το οποίο διατηρεί τη θετικότητα. Παραπέρα, εφαρμόζουμε τη μέθοδο μας σε προβλήματα τα οποία εμφανίζονται στο πεδίο των μοριακών δυναμικών, όπου το προτεινόμενο σχήμα το οποίο διατηρεί τη δομή της αρχικής εξίσωσης προσεγγίζει αποτελεσματικά κάποιες ΣΔΕ οι οποίες προκύπτουν έπειτα από μια διαδικασία απλοποίησης (*coarse graining*).

Θεωρούμε επίσης την περίπτωση Στοχαστικών Διαφορικών Εξισώσεων με Υστέρηση με μη-αρνητικές λύσεις. Ξανά στόχος μας είναι άμεσα αριθμητικά σχήματα τα οποία διατηρούν τη θετικότητα. Επεκτείνουμε την Ημι-Διακριτή μέθοδο από το πλαίσιο των Συνήθων ΣΔΕ στην περίπτωση με σταθερή υστέρηση, όπου και αποδεικνύουμε ισχυρή σύγκλιση (μοντέλο DGBM). Αριθμητικά πειράματα υποστηρίζουν τα θεωρητικά μας αποτελέσματα.

Λέξεις Κλειδιά : Ημι-Διακριτή μέθοδος, Υπερ-γραμμική Τάση και Διάχυση, Hölder Συνεχής, 3/2-Μοντέλο, Τάξη Σύγκλισης, Άμεσο Αριθμητικό Σχήμα, Διαδικασία CEV με ιδιότητα Επαναφοράς στο Μέσο, Διατήρηση Θετικότητας, Ισχυρό Σφάλμα Εκτίμησης, Στοχαστικό Μοντέλο Μεταβλητότητας, Στοχαστικές Διαφορικές Εξισώσεις, Στοχαστικές Διαφορικές Εξισώσεις με Υστέρηση, Προσομοίωση Monte Carlo, Αριθμητικές Μέθοδοι

AMS θεματική κατηγοριοποίηση 2010: 65C30, 65C20, 60H10, 60H35, 65J15
JEL θεματική κατηγοριοποίηση: C15, C63, G13

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The work presented in this thesis is motivated by the following question

How can we *efficiently* approximate solutions of non-linear stochastic differential equations?

The way we interpret the word *efficiently* is not necessarily in a computer-time consumption approach, i.e. the use of a numerical scheme that converges fast to the exact solution of our original stochastic differential equation (SDE), but in a qualitative way, in the sense that we aim for a numerical method that preserves some properties of the solution process of the SDE. In particular, we are interested in non-linear SDEs (solutions of linear SDEs or SDEs reducible to linear have an analytical expression, see [KP95, Sec. 4.4]), and in general in SDEs that have no analytical solution, which nevertheless lies in a certain domain. Therefore, our goal is to construct numerical schemes that preserve the original structure of the SDE at hand, i.e. that lie in the same domain. The main models that we treat, arise from the field of financial mathematics and in that setting the goal is to construct positivity preserving numerical schemes (see Chapters 2, 3 and 4). *En plus*, we apply our proposed method to a class of SDEs with solution in the interval $[-1, 1]$ and appear in molecular dynamics (see Chapter 5). But, why the study of SDEs is important?

In the following section, we give a brief discussion about the concept of SDEs and some properties of their solution processes. Notions of probability theory and stochastic processes are given in Appendix A.

1.1 SDEs: Origin & properties of their solution.

The study of SDEs has been extensive in the last 40 years. We present briefly their formulation and existence and uniqueness theorems. More details can be found in ([Fri75], [Mao97], [Øks03]) and references therein.

A first approach to SDEs is as stochastic analogs of ordinary differential equations (ODEs), where we allow some randomness in the coefficient of the ODE. A classic example (see for instance [Mao97, Sec. 2.1]) is the simple population growth model

$$\begin{aligned} \frac{dN}{dt} &= \alpha(t) \cdot N(t) \\ &= (r(t) + \textbf{noise}) \cdot N(t), \end{aligned} \tag{1.1.1}$$

where $N(t)$ is the size of population at time t, $\alpha(t)$ is the relative rate of growth at time t, which we assume to be subject to random environmental effects described by the term 'noise'. Noise can be represented by a suitable process $(\mathcal{W}_t)$ satisfying some properties (independent and stationary increments with zero mean, cf. [Øks03, Ch. 3]). Writing $\mathcal{W}_t \Delta t = \Delta V_t$ implies that the only such process (V_t) with continuous paths is the Wiener process (W_t) (see Def. A.3.8). Therefore (1.1.1) becomes the stochastic differential equation

$$dN(t) = r(t)N(t)dt + \sigma(t)N(t)dW(t), \tag{1.1.2}$$

with $N(0)$ the initial size of the population and σ an appropriate function. Equation (1.1.2) is in differential form. We prefer to use the stochastic integral representation of the solution process

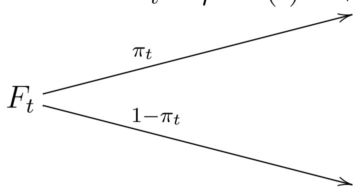
$$N_t = N_0 + \int_0^t r(s)N_s ds + \int_0^t \sigma(s)N_s dW_s, \tag{1.1.3}$$

where $N_0 = N(0)$ and the last integral is a stochastic integral which we interpret in the Itô sense and discuss in Appendix A.4. SDE (1.1.3) is linear, and admits the explicit solution [KP95, Sec. 4.4, p.120]

$$N_t = N_0 \exp\left\{ \int_0^t (r(s) - \frac{1}{2}\sigma^2(s))ds + \int_0^t \sigma(s)dW_s \right\}.$$

Another field of application of SDEs is in control theory, forming the *stochastic control theory*. Suppose that the fortune (F_t) of a person at time t is invested in a portion π_t to a risky investment, say a stock (S_t) and the rest to a safe investment, say a bond (B_t) in the following way

$$dS_t = \mu \cdot S(t)dt + \sigma \cdot S(t)dW(t), \quad \mu > r,\, \sigma \neq 0,$$



$$dB_t = r \cdot B(t)dt, \quad r > 0.$$

If the person has a utility function $U(F_t)$ describing the way that the person is satisfied w.r.t. his fortune at time t, the critical question is about the choice of the optimal portfolio $\pi_t \in [0,1]$ which maximizes the expected utility function of the person at a future time T, i.e. what is $\max_{0\leq \pi_t \leq 1} \mathbb{E}U(F_T)$?

Now, assume that the person above at time $t = 0$ has the choice of buying a unit of the risky asset at terminal time T at a fixed price K. How much should he be willing to pay? The above right (and not obligation) is called European Call Option and the answer to the question was given in [BS73] where the Black-Scholes option price formula was given. Since then, various much more complicated options have been considered in mathematical finance. The theoretical pricing of such options is a non-trivial task in itself. In practice, *strong* approximation schemes are of interest in these situations. Our proposed scheme has that feature, but we postulate the notion of strong approximation in Definition 1.3.10. We just mention that we require the realization to be close and not only the probability distribution as happens with *weak* approximation schemes, see Definition 1.3.11.

1.2 From Itô process to a general type SDE.

We assume the reader is familiar with some probability essentials and some basics of stochastic processes. Nevertheless, we give in Appendix A.2 all the relevant theory. Beginning with a definition of an Itô process and through the Itô formula, we are able to reach to the SDE which is of our main interest.

Let $T > 0$ and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, \mathbb{P})$ be a complete probability space, meaning that the *filtration* $\{\mathcal{F}_t\}_{0\leq t\leq T}$ satisfies the usual conditions, i.e. is right continuous and $\mathcal{F}_0$ includes all $\mathbb{P}$-*null* sets. Let $W_{t,\omega} : [0,T]\times\Omega \to \mathbb{R}^{m\times 1}$

be a m-dimensional *Wiener process* adapted to the filtration $\{\mathcal{F}_t\}_{0\le t\le T}$, i.e. $W_t = (W_t^1, \ldots, W_t^m)^T$ where $(W_t^j), j = 1, \ldots, m$ are independent Brownian motions.

First, we need to define an appropriate space of processes[1].

Definition 1.2.1 *We denote by* $\mathcal{L}^p([0,T];\mathbb{R}^d)$ *the family of all* $\mathbb{R}^d$*-valued* measurable, $\{\mathcal{F}_t\}$-adapted *processes* $\phi = \{\phi(t)\}_{0\le t\le T}$ *such that*

$$\int_0^T |\phi(s)|^p ds < \infty \ \ a.s.$$

□

Now, we define the Itô process [Mao97, Def. 1.6.3].

Definition 1.2.2 *[Itô process] An Itô process is an* $\mathbb{R}^d$*-valued continuous adapted process which satisfies the stochastic integral equation*

$$X_t = X_0 + \int_0^t a(s)ds + \int_0^t b(s)dW_s, \quad t \in [0,T], \tag{1.2.1}$$

where the coefficients $a \in \mathcal{L}^1([0,T];\mathbb{R}^d)$ *and* $b \in \mathcal{L}^2([0,T];\mathbb{R}^{d\times m})$*. The differential form of (1.2.1) is given by*

$$dX_t = a(t)dt + b(t)dW_t, \quad t \in [0,T]. \tag{1.2.2}$$

□

We use Itô stochastic calculus, i.e. we interpret the stochastic integral of (1.2.1) in the Itô sense (see Appendix A.4).

Definition 1.2.3 *[Itô formula] Let* (X_t) *be a d-dimensional Itô process with stochastic differential form (1.2.2). Let* $V \in C^{1,2}([0,T]\times\mathbb{R}^d;\mathbb{R})$*. Then* $V(t,X_t)$ *is again an Itô process which satisfies the following SDE*

$$\begin{aligned} dV(t,X_t) = &\left[\frac{\partial V}{\partial t}(t,X_t) + \frac{\partial V}{\partial X}(t,X_t)a(t) + \frac{1}{2}\text{trace}\left(b^T(t)\frac{\partial^2 V}{\partial X^2}(t,X_t)b(t)\right)\right]dt \\ &+ \frac{\partial V}{\partial X}(t,X_t)b(t)dW_t, \quad t \in [0,T]. \end{aligned}$$

□

[1] We state the following in $\mathbb{R}^d$, even though in the biggest part of this thesis, we treat scalar SDEs ($d = 1$).

We are now ready to state the type of the SDE which is the main subject here; it is the one we numerically approximate in a qualitative sense as will be argued later on.

Consider the following stochastic differential equation (SDE)

$$X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s, \quad t \in [0, T], \tag{1.2.3}$$

where the coefficients $a : [0, T] \times \mathbb{R}^d \to \mathbb{R}^d$ and $b : [0, T] \times \mathbb{R}^d \to \mathbb{R}^{d\times m}$ are measurable functions and $X_0 = (X_0^1, \ldots, X_0^d)^T$ is independent of all $\{W_t\}_{0\leq t\leq T}$.

Definition 1.2.4 *We say that SDE (1.2.3) has a unique strong solution if there is a* predictable *stochastic process* $X : [0, T] \times \Omega \to \mathbb{R}^d$ *such that [Mao97, Def. 2.2.1],*

$$\{a(t, X_t)\} \in \mathcal{L}^1([0, T]; \mathbb{R}^d), \quad \{b(t, X_t)\} \in \mathcal{L}^2([0, T]; \mathbb{R}^{d\times m})$$

and

$$\mathbb{P}\left[X_t = X_0 + \int_0^t a(s, X_s)ds + \int_0^t b(s, X_s)dW_s\right] = 1, \quad \textit{for every } t \in [0, T].$$

□

SDE (1.2.3) has non-autonomous coefficients, i.e. $a(t, x), b(t, x)$ depend explicitly on t. The drift coefficient a is the infinitesimal mean of the process X_t and the diffusion coefficient $\sqrt{bb^T}$ is the infinitesimal standard deviation of the process X_t.

The concept of *strong* solution is that the version of W_t is given in advance and the solution constructed from it is $\mathcal{F}_t$-adapted. If the coefficients a, b are given instead and we search for a pair of processes $((\widetilde{X_t}, \widetilde{W_t}), \mathcal{F}_t)$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ such that (1.2.3) holds then $\widetilde{X_t}$ is a *weak* solution. Heuristically, a strong solution is a functional of the initial condition X_0 and (W_t). A strong solution is also a weak solution, but the converse is not true; a well-known example is Tanaka's equation (cf. [Øks03, Example 5.3.2]) which has the differential form $dX_t = \text{sgn}(X_t)dW_t, X_0 = 0$. We shall focus on SDEs that admit strong solutions, since we are interested in the paths of them and not only their distribution.

1.2.1 Existence & uniqueness of solutions of (1.2.3).

The existence and uniqueness of solution of (1.2.3) is due to a combination of the following conditions:

(a) $||a(t,X_1)-a(t,X_2)||_2^2 \bigvee ||b(t,X_1)-b(t,X_2)||_2^2 \leq C||X_1-X_2||_2^2$, for all $t \in [0,T]$ and $X_1, X_2 \in \mathbb{R}^d$, where $C > 0$, (*Globally Lipschitz*)

(a^*) $||a(t,X_1)-a(t,X_2)||_2^2 \bigvee ||b(t,X_1)-b(t,X_2)||_2^2 \leq C_R||X_1-X_2||_2^2$, for all $t \in [0,T]$ when the d-vectors X_1, X_2 are such that $||X_1||_2 \vee ||X_2||_2 \leq R$ for any $R > 0$ where the quantity C_R depends on R, (*Locally Lipschitz*)

(b) $||a(t,X)||_2^2 \bigvee ||b(t,X)||_2^2 \leq C(1+||X||_2^2)$, for all $(t,X) \in [0,T] \times \mathbb{R}^d$, where the constant $C > 0$, (*Linear growth*)

(b^*) $X^T a(t,X) + \frac{p-1}{2}||b(t,X)||_2^2 \leq C(1+||X||_2^2)$, for all $(t,X) \in [0,T] \times \mathbb{R}^d$ and some $p \geq 2$, where the constant $C > 0$, (*Monotone type condition*)

In particular the pairs $(a)-(b), (a^*)-(b), (a^*)-(b^*)$, imply the existence and uniqueness of the solution [Mao97, Sec. 2.3]. The same conditions hold in ODEs. The next two examples show that Lipschitz continuity or linear growth is essential for the existence and uniqueness of the solution.

Example 1.2.5 *[Explosion in finite time] The SDE $dx_t = x_t^2 dt$, $x_0 = 1$ has solution $x_t = \frac{1}{1-t}$, $t \in [0,1)$.* □

Example 1.2.6 *[Not unique solution] The SDE $dx_t = 3x_t^{2/3} dt$, $x_0 = 0$ has solution $x_t = \mathbb{I}_{(a,\infty)}(t)(t-a)^3$, $t \in [0,\infty)$ for every a.* □

1.2.2 Properties of solutions of (1.2.3).

Condition (b^*) implies in the case $x_0 \in \mathcal{L}^p(\Omega, \mathbb{R}^d), p \geq 2$, the following moment bounds [Mao97, Th. 2.4.1],

$$\mathbb{E}||X_t||_2^p \leq C_{T,p}(1+\mathbb{E}||X_0||_2^p),$$

for all $t \in [0,T]$. Moreover, if the linear growth condition (b) holds we can also obtain uniform bounds for $\mathbb{E}\sup_{0\leq t\leq T}||X_t||_2^p$ for every $p \geq 2$ [Mao97, Th. 2.4.4].

1.3 Motivation.

SDEs of the form (1.2.3) rarely have explicit solutions, thus numerical approximations are necessary for simulations of the paths $X_t(\omega)$, or for approximation of functionals of the form $\mathbb{E}F(X)$, where $F : \mathcal{C}([0,T],\mathbb{R}) \to \mathbb{R}$ can be for example in the area of finance, the discounted payoff of a European type derivative.

Let us recall a definition from [Sch96] concerning the *life time* of numerical solution of SDEs.

Definition 1.3.7 *[Life time of numerical solution] Let $D \subseteq \mathbb{R}^d$ and consider a process (X_t) well-defined* [2] *on the domain $\overline{D}$, with initial condition $X_0 \in \overline{D}$ and such that*

$$\mathbb{P}(\{\omega \in \Omega : X(t,\omega) \notin \overline{D}\}) = 0,$$

for all $t > 0$. A numerical solution $(Y_{t_n})_{n\in\mathbb{N}}$ has a finite life time*, if there exists a stopping time $\tau_n(\omega)$ such that*

$$Y_n := Y_{\tau_n} \notin \overline{D} \quad a.s.$$

Otherwise, we say that it has an eternal life time. □

Equivalently, we say that the numerical integration scheme has an *eternal life time* if

$$\mathbb{P}(Y_{n+1} \in \overline{D} \,\big|\, Y_n \in \overline{D}) = 1. \tag{1.3.1}$$

We discretize $[0,T]$ with steps $\Delta_n := t_{n+1} - t_n$, for $n = 0,\ldots,N-1$, where $0 = t_0 < t_1 < \ldots < t_N = T$. Moreover, let $\Delta W_n := W_{t_{n+1}} - W_{t_n}$ be the increments of the Brownian motion.

The Euler method, applied to the SDE setting, already appeared in the 50's through Maruyama [Mar55] and thereafter has been an extensive study on numerical approximations of solutions of SDEs (we just mention [KN13] for a recent review on numerical methods for SDEs with applications in finance and references therein).

The explicit Euler-Maruyama (EM) scheme for SDE (1.2.3) is given by

$$Y_{n+1}^{EM} = Y_n + a(t_n, Y_n)\Delta_n + b(t_n, Y_n)\Delta W_n,$$

[2] On the complete probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$ the stochastic process $(X_t)_{0\le t\le T}$ takes values in $\overline{D}$, i.e. it is a collection of $\overline{D}$-valued r.vs in Ω.

for $n = 0, \ldots, N-1$, where $Y_0 = X_0$ and $Y_n := Y_{t_n}$. It is clear that the Euler-Maruyama scheme always has a finite life time[3], see e.g. [Kah04, Prop. 4.2]. The next example concerns a well-known process and the lack of the EM method to preserve the domain of its solution.

Example 1.3.8 *[CIR]* The following linear drift model had been initially proposed for the dynamics of the inflation rate by Cox, Ingersoll and Ross [CIR85, (51)] and is thus named CIR. It is used in the field of finance as a description of the stochastic volatility procedure in the Heston model [Hes93], but also belongs to the fundamental family of SDEs that approximate Markov jump processes [EK86]. The CIR model is described by the following SDE,

$$x_t = x_0 + \int_0^t \kappa(\lambda - x_s)ds + \int_0^t \sigma\sqrt{x_s}dW_s, \quad t \in [0, T], \tag{1.3.2}$$

where x_0 is independent of all $\{W_t\}_{t\geq 0}$, $x_0 > 0$ a.s. and the parameters κ, λ, σ are positive. Parameter λ is the level of the interest rate x_t where the drift is zero, meaning that when x_t is below λ the drift is positive, whereas in the other case it is negative. As λ grows, the range of the positive drift becomes wider. Parameter κ defines the slope of the drift. The condition $\kappa > 0$ is necessary for the stationarity of the process x_t. (The stationary distribution of (x_t) is gamma with shape parameter $2\lambda\kappa/\sigma^2$ and scale parameter $\sigma^2/(2\kappa)$. In particular it holds that $\mathbb{E}x_t = x_0 e^{-\kappa t} + \lambda(1 - e^{-\kappa t})$ and $Var x_t = x_0\sigma^2(e^{-\kappa t} - e^{-2\kappa t})/\kappa + \lambda\sigma^2(1 - e^{-\kappa t})^2/(2\kappa)$ c.f. [Shr04, Example 4.4.11]). When κ is negative, the main term of the slope, $-\kappa$, is positive and given the diffusion $\sigma\sqrt{x_t}$, the process x_t blows up. The condition $\sigma^2 < 2\kappa\lambda$ implied by the Feller test [Fel51, Case (ii),p.173] is necessary and sufficient for the process not to reach the boundary zero in finite time. Problem (1.3.2) is meant for non-negative values, since it represents rates or pricing values. Thus 'good' numerical schemes preserve positivity ([AGKR10], [KGR08]). The explicit Euler scheme has not that property, since its increments are conditional Gaussian. For example, the transition probability of the Euler scheme in case of (1.3.2) reads as

$$p(y|x) = \frac{1}{\sqrt{2\pi\sigma^2 x\Delta}} \exp\Big\{ - \frac{(y - (x + \kappa(\lambda - x)\Delta))^2}{2\sigma^2 x\Delta} \Big\}, \qquad y \in \mathbb{R},\ x > 0,$$

[3] We assume that the coefficients $a(t, x)$ and $b(t, x)$ are not simultaneously equal to zero for all (t, x).

thus, even in the first step there is an event of negative values with positive probability. We refer to [KN13], between other papers, that considers Euler type schemes, modifications of them to overcome the above drawback, and the importance of positivity. Thus, for the same problem, the truncated Euler scheme [DD98] has been proposed, as well as a modification of it, [HM05], where in a step the numerical scheme can leave $(0,\infty)$ but is forced to come back in the next steps. For the aforementioned problem there are methods of simulation ([BK06], [MG10]). However, if a full sample path of the SDE has to be simulated or the SDEs under study are a part of a bigger system of SDEs, then numerical schemes are in general more effective. □

The next example is one more non-linear in diffusion model.

Example 1.3.9 *[CEV]* The constant elasticity of the variance model [Cox75] is used for pricing assets and given by the SDE

$$x_t = x_0 + \int_0^t \mu x_s ds + \int_0^t \sigma x_s^\gamma dW_s, \quad t \in [0,T], \tag{1.3.3}$$

where x_0 is independent of $\{W_t\}_{0\leq t\leq T}$, $x_0 > 0$ a.s., $\mu \in \mathbb{R}, \sigma > 0$ and $0 < \gamma \leq 1$. SDE (1.3.3) has a unique strong solution if and only if $\gamma \in [1/2, 1]$ and takes values in $[0,\infty)$. The case $\gamma = 1/2$ corresponds to CIR model (1.3.2), whereas $\gamma = 1$ corresponds to a Brownian motion, i.e. the famous Black-Scholes model [BS73]. □

Therefore, we focus on numerical schemes with *eternal life time*. In [Sch96], where the above issue was originally discussed and further extended to methods of higher order [KS06], the main interest is in the domain $D = \mathbb{R}^+$. We study positivity preserving numerical schemes, but also treat other cases (see Chapter 5).

The second point of interest is in *strong* approximations (mean-square) of (1.2.3), in the case of super- or sub-linear drift and diffusion coefficients. This kind of numerical schemes, whose trajectories (sample paths) are close to those of (1.2.3) have applications in many areas - we discussed some in Section 1.1 but the interested reader is referred for instance to [HJ15, Sec. 4] and references therein - have theoretical interest (they provide fundamental insight for weak-sense schemes) and generally do not involve simulations over long-time periods or of a significant number of trajectories. A criterion of the closeness of the sample paths of (1.2.3) and the approximation process at final time T is the following.

Definition 1.3.10 *[Strong Approximation] A time discrete approximation Y*

with maximum step size Δ converges strongly *to* X *at time* T *if*

$$\lim_{\Delta \to 0} \mathbb{E}|Y_T - X_T|^2 = 0. \tag{1.3.4}$$

□

Sometimes, although not of our interest here, it suffices to have a 'good' approximation of the probability distribution of X_T rather than of its sample paths. This is stated in the following.

Definition 1.3.11 *[Weak Approximation] A time discrete approximation* Y *with maximum step size* Δ converges in the weak sense *to* X *at time* T *w.r.t. a class* $\mathcal{C}$ *of test functions* $\phi : \mathbb{R}^d \to \mathbb{R}$ *if*

$$\lim_{\Delta \to 0} |\mathbb{E}\phi(Y_T) - \mathbb{E}\phi(X_T)|^2 = 0,$$

for all $\phi \in \mathcal{C}$. □

The above approximation is much weaker than the one provided by the strong convergence criterion.

Relation (1.3.4) does not show the rate of convergence.

Definition 1.3.12 *[Order of Strong Convergence] A time discrete approximation* Y *with maximum step size* Δ converges strongly with order γ *to* X *at time* T *if there is a* $C > 0$ *and a* $\Delta^* > 0$ *such that*

$$\sqrt{\mathbb{E}|Y_T - X_T|^2} \leq C \cdot \Delta^\gamma, \tag{1.3.5}$$

for all $\Delta \in (0, \Delta^*)$, *where the constant* C *does not depend on* Δ. □

The order of a numerical scheme is usually less than that of the corresponding deterministic one (when $b = 0$) because of the increments of the Wiener process which are of root mean-square order 1/2, i.e. $\sqrt{\mathbb{E}(\Delta W_n)^2} = \Delta_n^{1/2}$.

Finally, assume the setting (1.2.3) where there is no time dependence in the coefficients a and b, and the resulting SDE is scalar and super-linear, i.e. we consider the following SDE

$$x_t = x_0 + \int_0^t a(x_s)ds + \int_0^t b(x_s)dW_s, \quad t \in [0, T],$$

where $a, b : \mathbb{R} \to \mathbb{R}$ are measurable functions, x_0 is independent of all $\{W_t\}_{0 \leq t \leq T}$ and let the constants $C \geq 1, \beta > \alpha > 1$ be such that

$$(|a(x)| \vee |b(x)|) \geq \frac{|x|^\beta}{C} \text{ and } (|a(x)| \wedge |b(x)|) \leq C|x|^\alpha, \tag{1.3.6}$$

for all $|x| \geq C$. Then the moments of the EM scheme may explode as shown in [HJK11, Th. 1], which we state below.

Theorem 1.3.13 *[Moment Explosion of EM scheme] Assume the setting above. Then, there exists a constant $c \in (1, \infty)$ and a sequence of non-empty events $\Omega_N \in \mathcal{F}, N \in \mathbb{N}$ with $\mathbb{P}(\Omega_N) \geq c^{-N^c}$ and $|Y_N(\omega)| \geq 2^{\alpha^{N-1}}$ for all $\omega \in \Omega_N$ and all $N \in \mathbb{N}$. Furthermore, if $\mathbb{E}|x_T|^p < \infty$ for a $p \in [1, \infty)$, then*

$$\lim_{N\to\infty} \mathbb{E}|x_T - Y_N^{EM}|^p = \infty \text{ and } \lim_{N\to\infty} \mathbb{E}|Y_N^{EM}|^p = \infty.$$

□

In other words, there exists a sequence of events of at least exponentially small probability on which the EM approximations grow at least double-exponentially fast resulting to them being unbounded in the $\mathcal{L}^1$-norm. (The way the EM scheme diverges follows by the inequality $\mathbb{E}|Y_N^{EM}| \geq \mathbb{P}(\Omega_N)|Y_N|$, see proof of [HJK11, Th. 1])

A numerical method that does not explode in the super-linear case is the tamed Euler method, see (2.1.5) suggested in [HJ15, (4)], which is explicit and strongly convergent. Nevertheless, it does not possess an eternal life time.

Therefore, we also aim for a numerical scheme that does not explode. In summary our goal is to construct numerical schemes with eternal life time, that converge strongly to the exact solution and do not explode.

1.4 Content of thesis.

This section explains how the following chapters fit in the picture drawn in the previous sections and especially Section 1.3. It is an outline of the content of this thesis.

Let us rewrite the general SDE (1.2.3) in the one-dimensional case

$$x_t = x_0 + \int_0^t a(s, x_s)ds + \int_0^t b(s, x_s)dW_s, \quad t \in [0, T], \tag{1.4.1}$$

where x_0 is independent of all $\{W_t\}_{0\leq t\leq T}$ and $a, b : [0, T] \times \mathbb{R} \to \mathbb{R}$ are such that (1.4.1) has a unique strong solution. We introduce the auxiliary functions $f(s, r, x, y), g(s, r, x, y) : [0, T]^2 \times \mathbb{R}^2 \to \mathbb{R}$ with $f(s, s, x, x) = a(s, x), g(s, s, x, x) = b(s, x)$, satisfying some local Lipschitz-type conditions, see for instance Assumption 2.2.1. Consider also the equidistant partition

$0 = t_0 < t_1 < ... < t_N = T$ with $\Delta = T/N$. The numerical scheme that we propose, and call semi-discrete (SD), has the following representation in each subinterval $[t_n, t_{n+1}]$, see (2.2.1)

$$y_t = y_{t_n} + \int_{t_n}^{t} f(t_n, s, y_{t_n}, y_s)ds + \int_{t_n}^{t} g(t_n, s, y_{t_n}, y_s)dW_s. \tag{1.4.2}$$

(All the related work concerning the semi-discrete method can be found in [Hal12, Hal14, Hal15d, Hal15c, Hal15b, Hal15a, HS16, HS15, Hal16]). The discretized part of the original SDE is given by the first and third variable of f, g. Note, that by fully discretizing the SDE, i.e. by choosing $f(s, r, x, y) = a(s, x)$ and $g(s, r, x, y) = b(s, x)$, we can reproduce the explicit Euler scheme. Moreover, a main difference of the SD method and all other numerical methods is that in each subinterval we have to solve a new SDE, and not an algebraic equation. The natural question that arises is the following:

How do we choose the auxiliary functions f and g?

The main idea of the SD method is to discretize only partially the original SDE in such a way that the remaining SDE has an explicit solution. Of course in this way, we do not produce a unique numerical scheme. Nevertheless we are able to prove the existence of a numerical scheme that overcomes the problems mentioned before, and make it specific in different cases separately. Thus, our method has the following properties:

($P1$) Converges strongly (in the mean-square sense) to the exact solution;

($P2$) Possesses an eternal life time;

($P3$) Does not explode in some non-linear problems;

($P4$) Is explicit.

In Chapter 2 we apply the SD method to super-linear problems of the form (1.3.6). Assuming moment bounds of the original SDE and the SD approximation we prove the strong convergence of our numerical scheme (1.4.2) to the true solution of (1.4.1). This is stated in Theorem 2.2.2 where we show that

$$\lim_{\Delta \to 0} \mathbb{E} \sup_{0 \le t \le T} |y_t - x_t|^2 = 0. \tag{1.4.3}$$

The kind of discretization that we use in the super-linear examples in Section 2.4 is multiplicative. The new SDE (1.4.2) is linear with exponential solution in each subinterval. The domain of interest is $\mathbb{R}^+$ and in the numerical experiments Section 2.5 we treat the Heston 3/2-model with coefficients of the form $a(s,x) = k_1 x - k_2 x^2$ and $b(s,x) = k_3 x^{3/2}$.

Relation (1.4.3) does not reveal the order of convergence. Nevertheless, since our scheme is first order -we use one stochastic integral in (1.4.2)- we know beforehand that the best we can expect is a numerical scheme with order 1, i.e. of the type (1.3.5) with $\gamma = 1$

$$(\mathbb{E} \sup_{0\leq t\leq T} |y_t - x_t|^2)^{1/2} \leq C \cdot \Delta.$$

We note that even when coefficients a and b are 'good', the same does not hold in general for the auxiliary functions f and g respectively, thus using standard arguments, we cannot estimate the order of convergence.

In Chapter 3 we study sub-linear models with coefficients of the form $a(s,x) = k_1 - k_2 x$ and $b(s,x) = k_3 x^q$ with $1/2 < q < 1$. We know again that $x_t > 0$ a.s. and aim for a scheme that is positivity preserving. Now, we use an additive discretization and are able to show the order of strong convergence which under some assumptions on the coefficients k_i is proved to be $(q - 1/2)/2$, that is (see Theorem 3.2.4)

$$(\mathbb{E} \sup_{0\leq t\leq T} |y_t - x_t|^2)^{1/2} \leq C \cdot \Delta^{(q-1/2)/2}.$$

Now, suppose that the initial condition x_0 in (1.4.1) is replaced by a function $\xi(t)$ with $t \in [-\tau, 0]$, i.e. we have some additional information on previous times, where $\tau > 0$ represents the amount of information available. This is the simplest case of equations called *constant delay differential equations*. We study in Chapter 4 a special model in the above setting called *Delay Geometric Brownian Motion*. This model arises in the area of financial mathematics in evaluating options. We show once more the mean-square convergence of our scheme to the exact solution of the DGBM model, see Theorem 4.2.2.

Finally, Chapter 5 is devoted to a class of SDEs with solutions in a domain other than $\mathbb{R}^+$. In particular the class of SDEs that we study admits solutions that lie in the interval $(-1, 1)$ and the goal is to construct a numerical scheme that preserves that structure satisfying in the same time properties $(P1) - (P4)$. This kind of SDEs appear in the field of molecular dynamics and in particular the so called 3-atom model [LL10, Sec. 4.2].

2. SUPER-LINEAR STOCHASTIC DIFFERENTIAL EQUATIONS

2.1 Introduction.

We [1] assume the setting in Section 1.2, with $d = m = 1$, i.e. let $T > 0$ and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, \mathbb{P})$ be a complete probability space, let $W_{t,\omega} : [0,T]\times\Omega \to \mathbb{R}$

[1] This chapter is based on joint work with Nikolaos Halidias, published in *Comput. Methods Appl. Math. (2015), DOI: 10.1515/cmam-2015-0028* [HS16].

be a one-dimensional Wiener process adapted to the filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ and consider the following stochastic differential equation (SDE),

$$x_t = x_0 + \int_0^t a(s, x_s)ds + \int_0^t b(s, x_s)dW_s, \quad t \in [0, T], \tag{2.1.1}$$

where the coefficients $a, b : [0, T] \times \mathbb{R} \to \mathbb{R}$ are measurable functions such that (2.1.1) has a unique strong solution and x_0 is independent of all $\{W_t\}_{0\leq t\leq T}$. SDE (2.1.1) has non-autonomous coefficients, i.e. $a(t, x), b(t, x)$ depend explicitly on t.

We are interested in numerical approximations of (2.1.1) and in particular in mean-square approximations for all the reasons discussed already in the end of Section 1.2 and in Section 1.3.

We present some models that are super-linear in the drift and diffusion coefficient:

- The 3/2-model [Hes97] or the inverse square root process [AG99], is used for modeling stochastic volatility and reads,

$$x_t = x_0 + \int_0^t (\alpha x_s - \beta x_s^2)ds + \int_0^t \sigma x_s^{3/2} dW_s, \quad t \in [0, T], \tag{2.1.2}$$

  where x_0 is independent of $\{W_t\}_{0\leq t\leq T}$, $x_0 > 0$ a.s. and $\sigma \in \mathbb{R}$. The conditions $\alpha > 0$ and $\beta > 0$ are necessary and sufficient for the stationarity of the process x_t and such that neither zero nor infinity is attainable in finite time [AG99, App. A].

- Super-linear models are models of the form (2.1.1) where one of the coefficients $a(\cdot), b(\cdot)$ is super-linear, i.e. when we have that

$$a(x) \geq \frac{|x|^\beta}{C}, \; b(x) \leq C|x|^\alpha, \text{ for every } \quad |x| \geq C, \tag{2.1.3}$$

  or

$$b(x) \geq \frac{|x|^\beta}{C}, \; a(x) \leq C|x|^\alpha, \text{ for every } \quad |x| \geq C, \tag{2.1.4}$$

  where $\beta > 1, \beta > \alpha \geq 0, C > 0$.

For the Heston 3/2-model there are methods of exact simulation ([BK06], [MG10]). However, if a full sample path of the SDE has to be simulated or the SDEs under study are a part of a bigger system of SDEs, then numerical schemes are in general more effective.

Problem (2.1.2) is meant for non-negative values. Thus for reasons already discussed, see Example 1.3.8 where the inverse process of (2.1.2) is considered, we aim for a positivity preserving scheme, since the explicit EM scheme does not possess that property.

One more drawback, that appears in case of super-linear problems (2.1.3) or (2.1.4), like the special case (2.1.2), is that the moments of the scheme may explode [HJK11, Th. 1]. A method that overcomes this drawback is the tamed Euler method, [HJ15, (4)]. It reads:

$$Y_0^N(\omega) := x_0(\omega)$$

and

(2.1.5)
$$Y_{n+1}^N(\omega) := Y_n^N(\omega) + \frac{T/N \cdot a(Y_n^N(\omega)) + b(Y_n^N(\omega))\Delta W_n(\omega)}{\max\{1, \frac{T}{N} \cdot \left|\frac{T}{N} a(Y_n^N(\omega)) + b(Y_n^N(\omega))\Delta W_n(\omega)\right|\}},$$

for every $n \in \{0, 1, ..., N-1\}, N \in \mathbb{N}$ and all $\omega \in \Omega$ where $\Delta W_n(\omega) := W_{\frac{(n+1)T}{N}}(\omega) - W_{\frac{nT}{N}}(\omega)$. The numerical scheme (2.1.5) is explicit, does not explode and converges strongly to the exact solution x_t of SDE (2.1.1), i.e.,

$$\lim_{N\to\infty} \left(\sup_{0\le t\le T} \mathbb{E}\left|x_t - \overline{Y}_t^N\right|^q \right) = 0, \tag{2.1.6}$$

for some $q > 0$, where $\overline{Y}_t^N := (n + 1 - \frac{tN}{T})Y_n^N + (\frac{tN}{T} - n)Y_{n+1}^N$ are continuous versions of (2.1.5) through linear interpolation. A balanced type scheme is also proposed in [TZ13, (3.1)], which reads

$$Y_{n+1}^N(\omega) := Y_n^N(\omega) + \frac{T/N \cdot a(Y_n^N(\omega)) + b(Y_n^N(\omega))\Delta W_n(\omega)}{1 + \frac{T}{N}\left|a(Y_n^N(\omega))\right| + \left|b(Y_n^N(\omega))\Delta W_n(\omega)\right|}, \tag{2.1.7}$$

where also the mean-square convergence rate is proved to be $1/2$, when the coefficients grow polynomially at infinity and satisfy a one-sided Lipschitz condition [TZ13, Prop. 3.3] in the sense

$$\mathbb{E}\left|x_n - Y_n^N\right|^{2q} \le C(1 + \mathbb{E}|x_0|^{2\gamma q}) \cdot \Delta^q,$$

where $\Delta = T/N, \gamma \geq 1$ and C does not depend on Δ. The stability properties of general tamed Euler schemes of the form

$$Y_{n+1}^N(\omega) := Y_n^N(\omega) + a^{\Delta}(Y_n^N(\omega)) \cdot \Delta + b^{\Delta}(Y_n^N(\omega))\Delta W_n(\omega), \tag{2.1.8}$$

where $a^{\Delta} \to a$ and $b^{\Delta} \to b$ as $\Delta \to 0$, are investigated in [Szp13] and a result of the form (2.1.6) is recovered. Schemes of the form (2.1.8) are also considered in [Sab15] where for the choice

$$a^{\Delta}(t,y) = \frac{a(t,y)}{1+\sqrt{\Delta}|y|^l}, \quad b^{\Delta}(t,y) = \frac{b(t,y)}{1+\sqrt{\Delta}|y|^l}, \tag{2.1.9}$$

where l comes from the polynomial growth of a, a uniform $\mathcal{L}^p$-convergence result is obtained [Sab15, Th. 3],

$$\mathbb{E} \sup_{0 \leq t \leq T} |x_t - Y_n^N(t)|^q \leq C \cdot \Delta^{q/2},$$

where C is independent of Δ and $q < p$. In general, all the above balanced schemes (2.1.5), (2.1.7), (2.1.8) and (2.1.9) that treat no globally Lipschitz coefficients, as well as the ones suggested in [HJ14] are half-order mean-square convergent schemes. Still all of them do not preserve positivity. See also [Zha14] where a first-order mean-square convergent scheme is proposed which reads

$$Y_{n+1}^N(\omega) := Y_n^N(\omega) + \sin(a(Y_n^N(\omega)) \cdot \Delta) + \sin\left(b(Y_n^N(\omega))\Delta W_n(\omega)\right).$$

For the aforementioned reasons there is an interest in the construction of suitable numerical schemes. An attempt to this direction has been made by the first author in [Hal12] and [Hal14] suggesting the semi-discrete method (where, briefly saying, we discretize a part of the SDE). Using this method in [Hal12] the author produced a new numerical scheme (but not unique in this situation) for the first aforementioned problem and proved the strong convergence of the scheme in mean-square sense. Later on, in [Hal14], the author generalized the idea of the semi-discrete method and used this generalization to approximate a class of super-linear problems, suggesting a new numerical scheme that preserves positivity in that case, proving again the strong convergence in the mean-square sense.

A basic feature of the semi-discrete method is that it is explicit, compared to other interesting, but implicit methods ([MS13c],[MS13b]), and converges strongly in the mean-square sense to the exact solution of the original SDE. Moreover, the semi-discrete method preserves positivity [Hal12, Sec. 3] and it does not explode in some super-linear problems [Hal14, Sec. 3].

Here, we generalize further the method to include non-autonomous coefficients, $a(t,x), b(t,x)$ in (2.1.1) and cover cases like that of the Heston 3/2-model. The extension of [Hal14, Th. 1] to time-dependent coefficients is not so difficult, but in order to deal with super-linear diffusion coefficients, like for example of the form $b(t,x) = \beta(t)\cdot x^{3/2}$, we have to use auxiliary functions g that satisfy Assumption 2.2.1 below (cf. [Hal14, Ass. A]).

2.2 The setting and the main result.

Assumption 2.2.1 *Let $f(s,r,x,y), g(s,r,x,y) : [0,T]^2 \times \mathbb{R}^2 \to \mathbb{R}$ be such that*

$$f(s,s,x,x) = a(s,x),\ g(s,s,x,x) = b(s,x),$$

where f, g satisfy the following conditions:

$$\begin{aligned}
|f(s_1,r_1,x_1,y_1) - f(s_2,r_2,x_2,y_2)| &\leq C_R\Big(|s_1-s_2| + |r_1-r_2| \\
&\qquad +|x_1-x_2| + |y_1-y_2|\Big) \\
|g(s_1,r_1,x_1,y_1) - g(s_2,r_2,x_2,y_2)| &\leq C_R\big(|s_1-s_2| + |r_1-r_2| \\
&\qquad +|x_1-x_2| + |y_1-y_2| + \sqrt{|x_1-x_2|}\big),
\end{aligned}$$

for any $R>0$ such that $|x_1| \vee |x_2| \vee |y_1| \vee |y_2| \leq R$, where the constant C_R depends on R. □

Given the equidistant partition $0 = t_0 < t_1 < \ldots < t_N = T$ and $\Delta = T/N$, we propose the semi-discrete numerical scheme

$$y_t = y_{t_n} + \int_{t_n}^{t} f(t_n, s, y_{t_n}, y_s)ds + \int_{t_n}^{t} g(t_n, s, y_{t_n}, y_s)dW_s,\ t \in [t_n, t_{n+1}], \tag{2.2.1}$$

where we assume that for every $n \leq N-1$, (2.2.1) has a unique strong solution and $y_0 = x_0$ a.s. In order to compare with the exact solution x_t, which is a continuous time process, we consider the following interpolation process of the semi-discrete approximation, in a compact form,

$$y_t = y_0 + \int_0^t f(\hat{s}, s, y_{\hat{s}}, y_s)ds + \int_0^t g(\hat{s}, s, y_{\hat{s}}, y_s)dW_s, \tag{2.2.2}$$

where $\hat{s} = t_n$, when $s \in [t_n, t_{n+1})$. The representation (2.2.2) is equivalent to (2.2.1), since for a $t \in [t_n, t_{n+1}]$ we have

$$\begin{aligned}
y_t &= y_0 + \left(\int_0^{t_1} + \int_{t_1}^{t_2} + \ldots + \int_{t_n}^{t}\right) f(\hat{s}, s, y_{\hat{s}}, y_s)ds \\
&\quad + \left(\int_0^{t_1} + \int_0^{t_1} + \ldots + \int_{t_n}^{t}\right) g(\hat{s}, s, y_{\hat{s}}, y_s)dW_s \\
&= \underbrace{y_0 + \int_0^{t_1} f(t_0, s, y_{t_0}, y_s)ds + \int_0^{t_1} g(t_0, s, y_{t_0}, y_s)dW_s}_{y_{t_1}} \\
&\quad + \int_{t_1}^{t_2} f(t_1, s, y_{t_1}, y_s)ds + \int_{t_1}^{t_2} g(t_1, s, y_{t_1}, y_s)dW_s + \ldots \\
&= y_{t_n} + \int_{t_n}^{t} f(t_n, s, y_{t_n}, y_s)ds + \int_{t_n}^{t} g(t_n, s, y_{t_n}, y_s)dW_s.
\end{aligned}$$

The first and third variable in f, g denote the discretized part of the original SDE. We observe from (2.2.2) that in order to solve for y_t, we have to solve an SDE and not an algebraic equation, thus in this context, we cannot reproduce implicit schemes, but we can reproduce the Euler scheme if we choose $f(s, r, x, y) = a(s, x)$ and $g(s, r, x, y) = b(s, x)$.

The numerical scheme (2.2.2) converges to the true solution x_t of SDE (2.1.1) and this is stated in the following, which is our main result.

Theorem 2.2.2 *Suppose Assumption 2.2.1 holds and (2.2.1) has a unique strong solution for every $n \leq N-1$, where $x_0 \in \mathcal{L}^p(\Omega, \mathbb{R}), x_0 > 0$ a.s. Let also*

$$\mathbb{E}(\sup_{0\leq t\leq T} |x_t|^p) \vee \mathbb{E}(\sup_{0\leq t\leq T} |y_t|^p) < A,$$

for some $p > 2$ and $A > 0$. Then the semi-discrete numerical scheme (2.2.2) converges to the true solution of (2.1.1) in the mean-square sense, that is

$$\lim_{\Delta\to 0} \mathbb{E} \sup_{0\leq t\leq T} |y_t - x_t|^2 = 0. \tag{2.2.3}$$

□

Section 2.3 is devoted to the proof of Theorem 2.2.2. Section 2.4 gives applications to super-linear drift and diffusion problems with non-negative solution, one of which includes the Heston 3/2-model. Section 2.5 shows experimentally the order of convergence of the SD method applied to the Heston 3/2-model. The semi-discrete scheme is strongly convergent in the mean-square sense and preserves positivity of the solution.

2.3 Proof of Theorem 2.2.2.

We denote the indicator function of a set A by $\mathbb{I}_A$. The constant C_R may vary from line to line and it may depend apart from R on other quantities, like time T for example, which are all constant, in the sense that we do not let them grow to infinity.

2.3.1 Error Bound for the Explicit Semi-Discrete Scheme.

Lemma 2.3.3 *Let the assumptions of Theorem 2.2.2 hold. Let $R > 0$, and set the stopping time $\theta_R = \inf\{t \in [0,T] : |y_t| > R \text{ or } |y_{\hat{t}}| > R\}$. Then the following estimate holds*

$$\mathbb{E}|y_{s\wedge\theta_R} - y_{\widehat{s\wedge\theta_R}}|^2 \leq C_R\Delta,$$

where C_R does not depend on Δ, implying $\sup_{s\in[t_{n_s},t_{n_s+1}]}\mathbb{E}|y_{s\wedge\theta_R} - y_{\widehat{s\wedge\theta_R}}|^2 = O(\Delta)$ as $\Delta \downarrow 0$. □

Proof of Lemma 2.3.3. Let n_s be an integer such that $s \in [t_{n_s}, t_{n_s+1})$. It holds that

$$\begin{aligned}
|y_{s\wedge\theta_R} - y_{\widehat{s\wedge\theta_R}}|^2 &= \left|\int_{t_{n_s}\wedge\theta_R}^{s\wedge\theta_R} f(\hat{u}, u, y_{\hat{u}}, y_u)du + \int_{t_{n_s}\wedge\theta_R}^{s\wedge\theta_R} g(\hat{u}, u, y_{\hat{u}}, y_u)dW_u\right|^2 \\
&\leq 2\left(\int_{t_{n_s}\wedge\theta_R}^{s\wedge\theta_R} f(\hat{u}, u, y_{\hat{u}}, y_u)du\right)^2 + 2\left(\int_{t_{n_s}\wedge\theta_R}^{s\wedge\theta_R} g(\hat{u}, u, y_{\hat{u}}, y_u)dW_u\right)^2 \\
&\leq 2\Delta\int_{t_{n_s}\wedge\theta_R}^{s\wedge\theta_R} f^2(\hat{u}, u, y_{\hat{u}}, y_u)du + 2\left(\int_{t_{n_s}\wedge\theta_R}^{s\wedge\theta_R} g(\hat{u}, u, y_{\hat{u}}, y_u)dW_u\right)^2 \\
&\leq C_R\Delta^2 + 2\left(\int_{t_{n_s}\wedge\theta_R}^{s\wedge\theta_R} g(\hat{u}, u, y_{\hat{u}}, y_u)dW_u\right)^2,
\end{aligned}$$

where we have used the Cauchy-Schwarz inequality (B.1.1) and Assumption 2.2.1 for the function f. (By the fact that we want the problem (2.1.1) to be well posed and by the conditions on f and g we get that f, g are bounded on bounded intervals.) Taking expectations in the above inequality gives

$$\begin{aligned}\mathbb{E}|y_{s\wedge\theta_R} - y_{\widehat{s\wedge\theta_R}}|^2 &\leq C_R\Delta^2 + 8\mathbb{E}\int_{t_{n_s\wedge\theta_R}}^{t_{n_s+1}\wedge\theta_R} g^2(\hat{u}, u, y_{\hat{u}}, y_u)du \\ &\leq C_R\Delta^2 + C_R\Delta,\end{aligned}$$

where in the first step we have used the BDG inequality (B.3.5) on the diffusion term and in the second step Assumption 2.2.1 for the function g. Thus,

$$\lim_{\Delta\downarrow 0}\frac{\sup_{s\in[t_{n_s},t_{n_s+1}]}\mathbb{E}|y_{s\wedge\theta_R} - y_{\widehat{s\wedge\theta_R}}|^2}{\Delta} \leq C_R,$$

which justifies the $O(\Delta)$ notation, (see for example [Olv97]). □

2.3.2 Convergence of the Semi-Discrete Scheme in $\mathcal{L}^1$.

Proposition 2.3.4 *Let the assumptions of Theorem 2.2.2 hold. Let $R > 0$, and set the stopping time $\theta_R = \inf\{t \in [0,T] : |y_t| > R \text{ or } |x_t| > R\}$. Then we have*

$$\begin{aligned}\sup_{0\leq t\leq T}\mathbb{E}|y_{t\wedge\theta_R} - x_{t\wedge\theta_R}| &\leq \Big[\Big(C_R + \frac{C_R}{me_m}\Big)\sqrt{\Delta} + \Big(\frac{C_R}{me_m} + C_R\Big)\Delta + \frac{C_R}{me_m}\Delta^2 \\ &\quad + \frac{C_R}{m} + e_{m-1}\Big]e^{a_{R,m}T},\end{aligned}$$

for any $m > 1$, where $e_m = e^{-m(m+1)/2}$, $a_{R,m} := C_R + \frac{C_R}{m}$ and C_R does not depend on Δ. It holds that $\lim_{m\uparrow\infty} e_m = 0$. □

Proof of Proposition 2.3.4. Given the non-increasing sequence $\{e_m\}_{m\in\mathbb{N}}$ with $e_m = e^{-m(m+1)/2}$ and $e_0 = 1$, we introduce the following sequence of smooth approximations of $|x|$, (method of Yamada and Watanabe, [YW71])

$$\phi_m(x) = \int_0^{|x|} dy \int_0^y \psi_m(u)du,$$

where the existence of the continuous function $\psi_m(u)$ with $0 \leq \psi_m(u) \leq 2/(mu)$ and support in (e_m, e_{m-1}) is justified by $\int_{e_m}^{e_{m-1}}(du/u) = m/2$. The following relations hold for $\phi_m \in \mathcal{C}^2(\mathbb{R};\mathbb{R})$ with $\phi_m(0)=0$,

$(P1)$ $|x| - e_{m-1} \le \phi_m(x) \le |x|$,

$(P2)$ $|\phi_m'(x)| \le 1$, $x \in \mathbb{R}$,

$(P3)$ $|\phi_m''(x)| \le \frac{2}{m|x|}$, for $e_m < |x| < e_{m-1}$ and $|\phi_m''(x)| = 0$, otherwise.

Denote $\mathcal{E}_t := y_t - x_t$. We have that

$$\mathbb{E}|\mathcal{E}_{t\wedge\theta_R}| \le e_{m-1} + \mathbb{E}\phi_m(\mathcal{E}_{t\wedge\theta_R}). \tag{2.3.1}$$

Applying Itô's formula to the sequence $\{\phi_m\}_{m\in\mathbb{N}}$, we get

$$\begin{aligned}\phi_m(\mathcal{E}_{t\wedge\theta_R}) &= \int_0^{t\wedge\theta_R} \phi_m'(\mathcal{E}_s)(f(\hat{s},s,y_{\hat{s}},y_s) - f(s,s,x_s,x_s))ds + M_t \\ &\quad + \frac{1}{2}\int_0^{t\wedge\theta_R} \phi_m''(\mathcal{E}_s)(g(\hat{s},s,y_{\hat{s}},y_s) - g(s,s,x_s,x_s))^2 ds,\end{aligned}$$

where

$$M_t := \int_0^{t\wedge\theta_R} \phi_m'(\mathcal{E}_u)(g(\hat{u},u,y_{\hat{u}},y_u) - g(u,u,x_u,x_u))dW_u.$$

Assumption 2.2.1 for the functions f, g and the properties of ϕ_m, imply

$$\begin{aligned}&\phi_m(\mathcal{E}_{t\wedge\theta_R}) \le \int_0^{t\wedge\theta_R} C_R\left(|y_{\hat{s}} - x_s| + |\mathcal{E}_s| + |\hat{s} - s|\right)ds + M_t \\ &+\frac{1}{2}\int_0^{t\wedge\theta_R} \frac{2}{m|\mathcal{E}_s|} C_R\left(|y_{\hat{s}} - x_s|^2 + |\mathcal{E}_s|^2 + |y_{\hat{s}} - x_s| + |\hat{s} - s|^2\right)ds \\ \le\ & C_R\int_0^{t\wedge\theta_R} |y_s - y_{\hat{s}}|ds + C_R\int_0^{t\wedge\theta_R}|\mathcal{E}_s|ds + C_R\int_0^{t\wedge\theta_R}|\hat{s} - s|ds + M_t \\ &+\frac{C_R}{m}\int_0^{t\wedge\theta_R} \frac{2|y_s - y_{\hat{s}}|^2 + 3|\mathcal{E}_s|^2 + |y_{\hat{s}} - x_s| + |\hat{s} - s|^2}{|\mathcal{E}_s|}ds \\ \le\ & (C_R + \frac{C_R}{me_m})\int_0^{t\wedge\theta_R}|y_s - y_{\hat{s}}|ds + \frac{C_R}{me_m}\int_0^{t\wedge\theta_R}|y_s - y_{\hat{s}}|^2 ds + M_t \\ &+(C_R + \frac{C_R}{m})\int_0^{t\wedge\theta_R}|\mathcal{E}_s|ds + \frac{C_R}{m} \\ &+\frac{C_R}{me_m}\sum_{k=0}^{[t/\Delta-1]}\int_{t_k}^{t_{k+1}\wedge\theta_R}|t_k - s|^2 ds + C_R\sum_{k=0}^{[t/\Delta-1]}\int_{t_k}^{t_{k+1}\wedge\theta_R}|t_k - s|ds\end{aligned}$$

or

$$\phi_m(\mathcal{E}_{t\wedge\theta_R}) \leq (C_R + \frac{C_R}{me_m}) \int_0^{t\wedge\theta_R} |y_s - y_{\hat{s}}| ds + \frac{C_R}{me_m} \int_0^{t\wedge\theta_R} |y_s - y_{\hat{s}}|^2 ds$$
$$+ (C_R + \frac{C_R}{m}) \int_0^{t\wedge\theta_R} |\mathcal{E}_s| ds + \frac{C_R}{m} + \frac{C_R}{me_m}\Delta^2 + C_R\Delta + M_t.$$

Taking expectations in the above inequality yields

$$\begin{aligned}
\mathbb{E}\phi_m(\mathcal{E}_{t\wedge\theta_R}) &\leq (C_R + \frac{C_R}{me_m}) \int_0^{t\wedge\theta_R} \mathbb{E}|y_s - y_{\hat{s}}| ds + (C_R + \frac{C_R}{m}) \int_0^{t\wedge\theta_R} \mathbb{E}|\mathcal{E}_s| ds \\
&\quad + \frac{C_R}{me_m} \int_0^{t\wedge\theta_R} \mathbb{E}|y_s - y_{\hat{s}}|^2 ds + \frac{C_R}{m} + \frac{C_R}{me_m}\Delta^2 + C_R\Delta + \mathbb{E}M_t \\
&\leq \left(C_R + \frac{C_R}{me_m}\right)\sqrt{\Delta} + \left(\frac{C_R}{me_m} + C_R\right)\Delta + \frac{C_R}{me_m}\Delta^2 + \frac{C_R}{m} \\
&\quad + \left(C_R + \frac{C_R}{m}\right) \int_0^{t\wedge\theta_R} \mathbb{E}|\mathcal{E}_s| ds,
\end{aligned}$$

where we have used Lemma 2.3.3 and the fact that $\mathbb{E}M_t = 0$.(Note that the function $h(u) = \phi'_m(\mathcal{E}_u)(g(\hat{u}, u, y_{\hat{u}}, y_u) - g(u, u, x_u, x_u))$ belongs to the space $\mathcal{M}^2([0, t\wedge\theta_R]; \mathbb{R})$ of real-valued measurable $\mathcal{F}_t$-adapted processes such that $\mathbb{E}\int_0^{t\wedge\theta_R} |h(u)|^2 du < \infty$. Now [Mao97, Th. 1.5.8] implies $\mathbb{E}M_t = 0$.) Thus (2.3.1) becomes

$$\mathbb{E}|\mathcal{E}_{t\wedge\theta_R}| \leq (C_R + \frac{C_R}{me_m})\sqrt{\Delta} + (\frac{C_R}{me_m} + C_R)\Delta + \frac{C_R}{me_m}\Delta^2 + \frac{C_R}{m} + e_{m-1}$$
$$+ \left(C_R + \frac{C_R}{m}\right) \int_0^{t\wedge\theta_R} \mathbb{E}|\mathcal{E}_s| ds$$
$$\leq \left[(C_R + \frac{C_R}{me_m})\sqrt{\Delta} + (\frac{C_R}{me_m} + C_R)\Delta + \frac{C_R}{me_m}\Delta^2 + \frac{C_R}{m} + e_{m-1}\right] e^{a_{R,m}t},$$

where in the last step we have used the Gronwall inequality (B.3.6) and $a_{R,m} = C_R + \frac{C_R}{m}$. Taking the supremum over all $0 \leq t \leq T$ implies the statement of Proposition 2.3.4. □

2.3.3 Convergence of the Semi-Discrete Scheme in $\mathcal{L}^2$.

Set the stopping time $\theta_R = \inf\{t \in [0,T] : |y_t| > R \text{ or } |x_t| > R\}$, for some $R > 0$ big enough. We have that

$$\begin{aligned}
\mathbb{E} \sup_{0\le t\le T} |y_t - x_t|^2 &= \mathbb{E} \sup_{0\le t\le T} |y_t - x_t|^2 \mathbb{I}_{(\theta_R > t)} + \mathbb{E} \sup_{0\le t\le T} |y_t - x_t|^2 \mathbb{I}_{(\theta_R \le t)} \\
&\le \mathbb{E} \sup_{0\le t\le T} |\mathcal{E}_{t\wedge\theta_R}|^2 + \frac{2\delta}{p} \mathbb{E} \sup_{0\le t\le T} |\mathcal{E}_t|^p + \frac{(p-2)}{p\delta^{2/(p-2)}} \mathbb{P}(\theta_R \le T) \\
&\le \mathbb{E} \sup_{0\le t\le T} |\mathcal{E}_{t\wedge\theta_R}|^2 + \frac{2^p\delta}{p} \mathbb{E} \sup_{0\le t\le T} (|y_t|^p + |x_t|^p) + \frac{(p-2)}{p\delta^{2/(p-2)}} \mathbb{P}(\theta_R \le T),
\end{aligned}$$

where in the second step we have applied the Young inequality, see (B.1.2),

$$ab \le \frac{\delta}{r} a^r + \frac{1}{q\delta^{q/r}} b^q,$$

for $a = \sup_{0\le t\le T} |\mathcal{E}_t|^2, b = \mathbb{I}_{(\theta_R \le t)}, r = p/2, q = p/(p-2)$ and $\delta > 0$ and in the third step we have used the elementary inequality $(\sum_{i=1}^n a_i)^p \le n^{p-1} \sum_{i=1}^n a_i^p$, with $n = 2$. In other words,

$$\mathbb{E} \sup_{0\le t\le T} |y_t - x_t|^2 \le \mathbb{E} \sup_{0\le t\le T} |\mathcal{E}_{t\wedge\theta_R}|^2 + \frac{2^{p+1}\delta A}{p} + \frac{(p-2)}{p\delta^{2/(p-2)}} \mathbb{P}(\theta_R \le T), \tag{2.3.2}$$

where A comes from the moment bound assumption. It holds that

$$\begin{aligned}
\mathbb{P}(\theta_R \le T) &\le \mathbb{E}\left(\mathbb{I}_{(\theta_R \le T)} \frac{|y_{\theta_R}|^p}{R^p} \right) + \mathbb{E}\left(\mathbb{I}_{(\theta_R \le T)} \frac{|x_{\theta_R}|^p}{R^p} \right) \\
&\le \frac{1}{R^p} \left(\mathbb{E} \sup_{0\le t\le T} |x_t|^p + \mathbb{E} \sup_{0\le t\le T} |y_t|^p \right) \le \frac{2A}{R^p},
\end{aligned}$$

thus (2.3.2) becomes

$$\mathbb{E} \sup_{0\le t\le T} |y_t - x_t|^2 \le \mathbb{E} \sup_{0\le t\le T} |\mathcal{E}_{t\wedge\theta_R}|^2 + \frac{2^{p+1}\delta A}{p} + \frac{2(p-2)A}{p\delta^{2/(p-2)} R^p}. \tag{2.3.3}$$

We estimate the difference $|\mathcal{E}_{t\wedge\theta_R}|^2$. It holds that

$$
\begin{aligned}
|\mathcal{E}_{t\wedge\theta_R}|^2 = \Big| \int_0^{t\wedge\theta_R} (f(\hat{s}, s, y_{\hat{s}}, y_s) - f(s, s, x_s, x_s))\, ds \\
+ \int_0^{t\wedge\theta_R} (g(\hat{s}, s, y_{\hat{s}}, y_s) - g(s, s, x_s, x_s)) dW_s \Big|^2 \\
\le 2T \int_0^{t\wedge\theta_R} C_R \left(|y_{\hat{s}} - x_s|^2 + |y_s - x_s|^2 + |\hat{s} - s|^2\right) ds + 2|M_t|^2 \\
\le C_R \int_0^{t\wedge\theta_R} |y_s - y_{\hat{s}}|^2 ds + C_R \int_0^{t\wedge\theta_R} |\mathcal{E}_s|^2 ds + C_R \int_0^{t\wedge\theta_R} |\hat{s} - s|^2 ds + 2|M_t|^2,
\end{aligned}
$$

where in the second step we have used the Cauchy-Schwarz inequality and Assumption 2.2.1 for f and

$$
M_t := \int_0^{t\wedge\theta_R} (g(\hat{s}, s, y_{\hat{s}}, y_s) - g(s, s, x_s, x_s))\, dW_s.
$$

Writing once more $\int_0^{t\wedge\theta_R} |\hat{s} - s|^2 = \sum_{k=0}^{[t/\Delta-1]} \int_{t_k}^{t_{k+1}\wedge\theta_R} |t_k - s|^2 ds$, taking the supremum over all $t \in [0, T]$ and then expectations we have

$$
\begin{aligned}
\mathbb{E} \sup_{0\le t\le T} |\mathcal{E}_{t\wedge\theta_R}|^2 &\le C_R \mathbb{E} \left(\int_0^{T\wedge\theta_R} |y_s - y_{\hat{s}}|^2 ds \right) + 2\mathbb{E} \sup_{0\le t\le T} |M_t|^2 \\
&\quad + C_R \int_0^T \mathbb{E} \sup_{0\le l\le s} |\mathcal{E}_{l\wedge\theta_R}|^2 ds + C_R \Delta^2 \\
&\le C_R \int_0^{T\wedge\theta_R} \mathbb{E}|y_s - y_{\hat{s}}|^2 ds + 8\mathbb{E}|M_T|^2 \\
&\quad + C_R \int_0^T \mathbb{E} \sup_{0\le l\le s} |\mathcal{E}_{l\wedge\theta_R}|^2 ds + C_R \Delta^2,
\end{aligned}
\tag{2.3.4}
$$

where in the last step we have used Hölder's inequality (B.1.3) and Doob's martingale inequality with $p = 2$, since M_t is an $\mathbb{R}$-valued martingale that belongs to $\mathcal{L}^2$. It holds that

$$\begin{aligned}
&\mathbb{E}|M_T|^2 := \mathbb{E}\left|\int_0^{T\wedge\theta_R} (g(\hat{s}, s, y_{\hat{s}}, y_s) - g(s, s, x_s, x_s))\, dW_s\right|^2 \\
= \; & \mathbb{E}\left(\int_0^{T\wedge\theta_R} (g(\hat{s}, s, y_{\hat{s}}, y_s) - g(s, s, x_s, x_s))^2\, ds\right) \\
\leq \; & C_R\mathbb{E}\left(\int_0^{T\wedge\theta_R} \left(|y_{\hat{s}} - x_s|^2 + |y_s - x_s|^2 + |y_{\hat{s}} - x_s| + |\hat{s} - s|^2\right) ds\right) \\
\leq \; & C_R\int_0^{T\wedge\theta_R} \mathbb{E}|y_s - y_{\hat{s}}|^2 ds + C_R\int_0^T \mathbb{E}\sup_{0\leq l\leq s}|\mathcal{E}_{l\wedge\theta_R}|^2 ds \\
& + C_R\int_0^{T\wedge\theta_R} \mathbb{E}|y_{\hat{s}} - x_s| ds + C_R\Delta^2,
\end{aligned}$$

where we have used Assumption 2.2.1 for g. Relation (2.3.4) becomes

$$\begin{aligned}
&\mathbb{E}\sup_{0\leq t\leq T}|\mathcal{E}_{t\wedge\theta_R}|^2 \leq C_R\int_0^{T\wedge\theta_R} \mathbb{E}|y_s - y_{\hat{s}}|^2 ds + C_R\int_0^T \mathbb{E}\sup_{0\leq l\leq s}|\mathcal{E}_{l\wedge\theta_R}|^2 ds \\
&+C_R\int_0^{T\wedge\theta_R} (\mathbb{E}|y_s - y_{\hat{s}}| + \mathbb{E}|y_s - x_s|)\, ds + C_R\Delta^2 \\
&\leq C_R\sqrt{\Delta} + C_R\Delta + C_R\Delta^2 + C_R\int_0^T \mathbb{E}\sup_{0\leq l\leq s}|\mathcal{E}_{l\wedge\theta_R}|^2 ds + C_R\int_0^{T\wedge\theta_R} \mathbb{E}|\mathcal{E}_s| ds,
\end{aligned}$$

where we have used Lemma 2.3.3 and Jensen's inequality for the concave function $\phi(x) = \sqrt{x}$. The integrand of the last term is bounded, from Proposition 2.3.4, by

$$K_{R,\Delta,m}(s) := \left[(C_R + \frac{C_R}{me_m})(\sqrt{\Delta} + \Delta) + \frac{C_R}{me_m}\Delta^2 + \frac{C_R}{m} + e_{m-1}\right] e^{a_{R,m}s},$$

where $s \in [0, T\wedge\theta_R]$. Application of the Gronwall inequality implies

$$\mathbb{E}\sup_{0\leq t\leq T}|\mathcal{E}_{t\wedge\theta_R}|^2 \leq \left(C_R\sqrt{\Delta} + C_R\Delta + C_R K_{R,\Delta,m}(T)\right) e^{C_R} \leq C_{R,\Delta,m}.$$

Note that, given $R > 0$, the quantity $C_{R,\Delta,m}$ can be arbitrarily small by choosing big enough m and small enough Δ. Relation (2.3.3) becomes,

$$\begin{aligned}
\mathbb{E}\sup_{0\leq t\leq T}|y_t - x_t|^2 \; &\leq \; C_{R,\Delta,m} + \frac{2^{p+1}\delta A}{p} + \frac{2(p-2)A}{p\delta^{2/(p-2)}R^p} \\
&:= I_1 + I_2 + I_3.
\end{aligned}$$

Given any $\epsilon > 0$, we may first choose δ such that $I_2 < \epsilon/3$, then choose R such that $I_3 < \epsilon/3$, then $m > 1$ and finally Δ such that $I_1 < \epsilon/3$ concluding $\mathbb{E}\sup_{0\le t\le T}|y_t - x_t|^2 < \epsilon$ as required to verify (2.2.3).

2.4 Super-linear examples.

2.4.1 Example I.

We study the numerical approximation of the following SDE:

$$x_t = x_0 + \int_0^t (k_1(s)x_s - k_2(s)x_s^2)ds + \int_0^t k_3(s)x_s^{3/2}\phi(x_s)dW_s,\ t\in[0,T], \tag{2.4.1}$$

where $\phi(\cdot)$ is a locally Lipschitz and bounded function with locally Lipschitz constant C_R^{ϕ}, bounding constant K_ϕ, x_0 is independent of all $\{W_t\}_{0\le t\le T}$, $x_0 \in \mathcal{L}^{4p}(\Omega,\mathbb{R})$ for some $2 < p$ and $x_0 > 0$ a.s., $\mathbb{E}(x_0)^{-2} < A$, $k_1(\cdot), k_2(\cdot), k_3(\cdot)$ are positive and bounded functions with $k_{2,\min} > \frac{7}{2}(K_\phi k_{3,\max})^2$. Model (2.4.1) has super-linear drift and diffusion coefficients.

We propose the following semi-discrete numerical scheme

$$y_t = y_{t_n} + \int_{t_n}^t (k_1(s) - k_2(s)y_{t_n})y_s ds + \int_{t_n}^t k_3(s)\sqrt{y_{t_n}}\phi(y_{t_n})y_s dW_s,\ t\in[t_n,t_{n+1}], \tag{2.4.2}$$

for $n \le T/\Delta$ and $y_0 = x_0$ a.s., or in a more compact form,

$$y_t = y_0 + \int_0^t (k_1(s) - k_2(s)y_{\hat{s}})y_s ds + \int_0^t k_3(s)\sqrt{y_{\hat{s}}}\phi(y_{\hat{s}})y_s dW_s, \tag{2.4.3}$$

where $\hat{s} = t_n$, when $s \in [t_n, t_{n+1})$. The linear SDE (2.4.3) has a solution which, by use of Itô's formula, has the explicit form

$$y_t = x_0 \exp\Big\{\int_0^t (k_1(s) - k_2(s)y_{\hat{s}} - k_3^2(s)\frac{y_{\hat{s}}\phi^2(y_{\hat{s}})}{2})ds + \int_0^t k_3(s)\sqrt{y_{\hat{s}}}\phi(y_{\hat{s}})dW_s\Big\}, \tag{2.4.4}$$

where $y_t = y_t(t_0, x_0)$.

Proposition 2.4.5 *The semi-discrete numerical scheme (2.4.3) converges to the true solution of (2.4.1) in the mean-square sense, that is*

$$\lim_{\Delta\to 0}\mathbb{E}\sup_{0\le t\le T}|y_t - x_t|^2 = 0.$$

□

Proof of Proposition 2.4.5.

In order to prove the proposition, we need to verify the assumptions of Theorem 2.2.2. Let

$$a(s,x) = k_1(s)x - k_2(s)x^2, \qquad f(s,r,x,y) = (k_1(s) - k_2(s)x)y,$$
$$b(s,x) = k_3(s)x^{3/2}\phi(x), \qquad g(s,r,x,y) = k_3(s)\sqrt{x}\phi(x)y.$$

We verify Assumption 2.2.1 for f. Let $R > 0$ such that $|x_1| \vee |x_2| \vee |y_1| \vee |y_2| \vee |s| \vee |r| \leq R$. We have that

$$\begin{aligned}
& |f(s,r,x_1,y_1) - f(s,r,x_2,y_2)| = |(k_1(s)(y_1 - y_2) - k_2(s)(x_1y_1 - x_2y_2)| \\
\leq\ & |k_1(s)||y_1 - y_2| + |k_2(s)|(|x_2||y_1 - y_2| + |y_1||x_1 - x_2|) \\
\leq\ & (|k_{1,\max}| + |k_{2,\max}|R)|y_1 - y_2| + |k_{2,\max}|R|x_1 - x_2| \\
\leq\ & C_R\left(|x_1 - x_2| + |y_1 - y_2|\right),
\end{aligned}$$

thus, Assumption 2.2.1 holds for f with $C_R := |k_{1,\max}| + |k_{2,\max}|R$.

We verify Assumption 2.2.1 for g. Let $R > 0$ such that $|x_1| \vee |x_2| \vee |y_1| \vee |y_2| \vee |s| \vee |r| \leq R$. We have that

$$\begin{aligned}
& |g(s,r,x_1,y_1) - g(s,r,x_2,y_2)| = |k_3(s)\sqrt{x_1}\phi(x_1)y_1 - k_3(s)\sqrt{x_2}\phi(x_2)y_2| \\
\leq\ & |k_3(s)|\Big(\sqrt{x_1}|\phi(x_1)||y_1 - y_2| + |y_2|\Big|\sqrt{x_1}\phi(x_1) - \sqrt{x_1}\phi(x_2) \\
& +\sqrt{x_1}\phi(x_2) - \sqrt{x_2}\phi(x_2)\Big|\Big) \\
\leq\ & |k_{3,\max}|\left(K_\phi\sqrt{R}|y_1 - y_2| + R\sqrt{x_1}|\phi(x_1) - \phi(x_2)| + RK_\phi|\sqrt{x_1} - \sqrt{x_2}|\right) \\
\leq\ & |k_{3,\max}|\left(K_\phi\sqrt{R}|y_1 - y_2| + R^{3/2}C_R^\phi|x_1 - x_2| + RK_\phi\sqrt{|x_1 - x_2|}\right) \\
\leq\ & C_R\left(|x_1 - x_2| + |y_1 - y_2| + \sqrt{|x_1 - x_2|}\right),
\end{aligned}$$

where we have used the fact that the function $\sqrt{x}$ is 1/2-Hölder continuous and $C_R := |k_{3,\max}|\left(C_R^\phi R^{3/2} \vee K_\phi\sqrt{R} \vee K_\phi R\right)$. Thus, Assumption 2.2.1 holds for g. Lemmata 2.4.7 and 2.4.8 complete the proof.

Moment Bound for Original SDE.

Lemma 2.4.6 *[Positivity of (x_t)] In the previous setting it holds that $x_t > 0$ a.s.* □

Proof of Lemma 2.4.6. Set the stopping time $\theta_R = \inf\{t \in [0,T] : x_t^{-1} > R\}$, for some $R > 0$, with the convention that $\inf \emptyset = \infty$. Application of Itô's formula on $(x_{t\wedge\theta_R})^{-2}$ implies,

$$
\begin{aligned}
(x_{t\wedge\theta_R})^{-2} &= (x_0)^{-2} + \int_0^{t\wedge\theta_R} (-2)(x_s)^{-3}(k_1(s)x_s - k_2(s)x_s^2)ds \\
&\quad + \int_0^{t\wedge\theta_R} 3(x_s)^{-4}k_3^2(s)x_s^3\phi^2(x_s)ds + \int_0^{t\wedge\theta_R}(-2)k_3(s)(x_s)^{-3}x_s^{3/2}\phi(x_s)dW_s \\
&\le (x_0)^{-2} + \int_0^{t\wedge\theta_R}(-2k_1(s)x_s^{-2} + 2k_2(s)x_s^{-1} + 3k_3^2(s)K_\phi^2 x_s^{-1})ds \\
&\quad + \int_0^t (-2)k_3(s)x_s^{-3/2}\phi(x_s)\,\mathbb{I}_{(0,t\wedge\theta_R)}(s)dW_s \\
&\le \int_0^{t\wedge\theta_R}[-2k_1(s)x_s^{-2} + (2k_2(s) + 3k_3^2(s)K_\phi^2)\left(x_s^{-1}\mathbb{I}_{(0,1]}(x_s) + x_s^{-1}\mathbb{I}_{(1,\infty]}(x_s)\right)]ds \\
&\quad + (x_0)^{-2} + M_t \\
&\le (2k_{2,\max} + 3k_{3,\max}^2 K_\phi^2)T + \int_0^t (2k_2(s) + 3k_3^2(s)K_\phi^2)x_s^{-2}\,\mathbb{I}_{(0,t\wedge\theta_R)}(s)ds \\
&\quad + (x_0)^{-2} + M_t,
\end{aligned}
$$

where

$$M_t := \int_0^t (-2)k_3(s)x_s^{-3/2}\phi(x_s)\,\mathbb{I}_{(0,t\wedge\theta_R)}(s)dW_s.$$

Taking expectations in the above inequality and using the fact that $\mathbb{E}M_t = 0$, we get that

$$
\begin{aligned}
\mathbb{E}(x_{t\wedge\theta_R})^{-2} &\le \mathbb{E}(x_0)^{-2} + 2k_{2,\max}T + 3k_{3,\max}^2K_\phi^2T + (2k_{2,\max} + 3k_{3,\max}^2K_\phi^2) \\
&\qquad \times \int_0^t \mathbb{E}(x_{s\wedge\theta_R})^{-2}ds \\
&\le \left(\mathbb{E}(x_0)^{-2} + 2k_{2,\max}T + 3k_{3,\max}^2K_\phi^2T\right)e^{(2k_{2,\max}+3k_{3,\max}^2K_\phi^2)T} < C,
\end{aligned}
$$

where we have used Gronwall's inequality with C independent of R. (Note that the function $h(u) = (-2)k_3(u)x_u^{-3/2}\phi(x_u)\,\mathbb{I}_{(0,t\wedge\theta_R)}(u)$ belongs to the space $\mathcal{M}^2([0,t];\mathbb{R})$ thus [Mao97, Th. 1.5.8] implies $\mathbb{E}M_t = 0$.) We have that

$$(x_{t\wedge\theta_R})^{-2} = (x_{\theta_R})^{-2}\mathbb{I}_{(\theta_R\le t)} + (x_t)^{-2}\mathbb{I}_{(t<\theta_R)} = R^2\mathbb{I}_{(\theta_R\le t)} + (x_t)^{-2}\mathbb{I}_{(t<\theta_R)},$$

which implies

$$\mathbb{E}\left(\frac{1}{x_{t\wedge\theta_R}^2}\right) = R^2\mathbb{P}(\theta_R \le t) + \mathbb{E}\left(\frac{1}{x_t^2}\mathbb{I}_{(t<\theta_R)}\right) < C,$$

thus

$$\mathbb{P}(x_t \le 0) = \mathbb{P}\Big(\bigcap_{R=1}^{\infty}\Big\{x_t < \frac{1}{R}\Big\}\Big) = \lim_{R\to\infty}\mathbb{P}\Big(\Big\{x_t < \frac{1}{R}\Big\}\Big) \le \lim_{R\to\infty}\mathbb{P}(\theta_R \le t) = 0.$$

We conclude that $x_t > 0$ a.s. □

Lemma 2.4.7 *In the previous setting it holds that*

$$\mathbb{E}(\sup_{0\le t\le T}(x_t)^p) < A_1,$$

for some $A_1 > 0$ *and any* $2 < p \le k_{2,min}/(K_\phi k_{3,max})^2$. □

Proof of Lemma 2.4.7. In the case all x are outside a finite ball of radius $R > 1$, and $s \in [0, T]$ we have that

$$\begin{aligned} J(s,x) &:= \frac{xa(s,x) + (p-1)b^2(s,x)/2}{1+x^2} \\ &= \frac{x(k_1(s)x - k_2(s)x^2) + (p-1)k_3^2(s)[x^{3/2}\phi(x)]^2/2}{1+x^2} \\ &= \frac{k_1(s)x^2 - k_2(s)x^3 + 0.5(p-1)k_3^2(s)x^3\phi^2(x)}{1+x^2} \\ &\le \frac{k_{1,\max}x^2 + \Big(0.5(p-1)(k_{3,\max}K_\phi)^2 - k_{2,\min}\Big)x^3}{1+x^2} \le k_{1,\max}, \end{aligned}$$

where the last inequality is valid for all p such that $p \le 1+2k_{2,\min}/(K_\phi k_{3,\max})^2$. Thus $J(s,x)$ is bounded for all $(s,x) \in [0,T]\times\mathbb{R}$, since when $|x| \le R$ we have that $J(s,x)$ is finite, say $J(s,x) \le C$. Since C is positive, application of [Mao97, Th. 2.4.1] implies

$$\mathbb{E}(x_t)^p \le 2^{(p-2)/2}(1 + \mathbb{E}(x_0)^p)e^{Cpt},$$

for any $2 < p \le 1+2k_{2,\min}/(K_\phi k_{3,\max})^2$ and all $t \in [0,T]$. Using Itô's formula on $(x_t)^p$, with $p \le k_{2,\min}/(K_\phi k_{3,\max})^2$ (in order to use Doob's martingale inequality later) we have that

$$
\begin{aligned}
(x_t)^p &= (x_0)^p + \int_0^t p(x_s)^{p-1}(k_1(s)x_s - k_2(s)x_s^2)ds \\
&+ \int_0^t \frac{p(p-1)}{2}(x_s)^{p-2}[k_3(s)x_s^{3/2}\phi(x_s)]^2 ds + \int_0^t pk_3(s)(x_s)^{p-1}x_s^{3/2}\phi(x_s)dW_s \\
&\leq (x_0)^p + p\int_0^t \left[k_1(s)(x_s)^p + \left(\frac{p-1}{2}k_{3,\max}^2 K_\phi^2 - k_2\right)(x_s)^{p+1}\right] ds + M_t \\
&\leq (x_0)^p + p\int_0^t k_1(s)(x_s)^p ds + M_t,
\end{aligned}
$$

where $M_t = \int_0^t pk_3(s)\phi(x_s)(x_s)^{p+1/2}dW_s$. Taking the supremum and then expectations in the above inequality we get

$$
\begin{aligned}
\mathbb{E}(\sup_{0\leq t\leq T}(x_t)^p) &\leq \mathbb{E}(x_0)^p + pk_{1,\max}\mathbb{E}\left(\sup_{0\leq t\leq T}\int_0^t (x_s)^p ds\right) + \mathbb{E}\sup_{0\leq t\leq T} M_t \\
&\leq \mathbb{E}(x_0)^p + pk_{1,\max}\int_0^T \mathbb{E}(\sup_{0\leq l\leq s}(x_l)^p)ds + \sqrt{\mathbb{E}\sup_{0\leq t\leq T} M_t^2} \\
&\leq \left(\mathbb{E}(x_0)^p + \sqrt{4\mathbb{E}M_T^2}\right)e^{pk_{1,\max}T} := A_1,
\end{aligned}
$$

where in the last step we have used Doob's martingale inequality to the diffusion term M_t and the Gronwall inequality. (Note that the function $h(u) = pk_3(u)\phi(x_u)(x_u)^{p+1/2}$ belongs to the family $\mathcal{M}^2([0,T];\mathbb{R})$ thus [Mao97, Th. 1.5.8] implies $\mathbb{E}M_t^2 = \mathbb{E}(\int_0^t h(u)dW_u)^2 = \mathbb{E}\int_0^t h^2(u)du$, i.e. $M_t \in \mathcal{L}^2(\Omega;\mathbb{R})$.) □

Moment Bound for Semi-Discrete Approximation.

Lemma 2.4.8 *In the previous setting it holds that*

$$\mathbb{E}(\sup_{0\leq t\leq T}(y_t)^p) < A_2,$$

for some $A_2 > 0$ and for any $2 < p \leq 1/4 + \frac{k_{2,min}}{2(k_{3,max}K_\phi)^2}$. □

Proof of Lemma 2.4.8. Set the stopping time $\theta_R = \inf\{t \in [0,T] : y_t > R\}$, for some $R > 0$, with the convention that $\inf \emptyset = \infty$. Application of Itô's formula on $(y_{t\wedge\theta_R})^q$, with $q = 4p$ implies,

$$\begin{aligned}
&(y_{t\wedge\theta_R})^q = (y_0)^q + \int_0^{t\wedge\theta_R} q(y_s)^{q-1}(k_1(s) - k_2(s)y_{\hat{s}})y_s ds \\
&+ \int_0^{t\wedge\theta_R} \frac{q(q-1)}{2}(y_s)^{q-2}\left[k_3(s)\sqrt{y_{\hat{s}}}\phi(y_{\hat{s}})y_s\right]^2 ds \\
&+ \int_0^{t\wedge\theta_R} qk_3(s)(y_s)^{q-1}\sqrt{y_{\hat{s}}}\phi(y_{\hat{s}})y_s dW_s \\
= \; & (x_0)^q + \int_0^{t\wedge\theta_R} \left(q(k_1(s) - k_2(s)y_{\hat{s}}) + \frac{q(q-1)k_3^2(s)}{2}y_{\hat{s}}\phi^2(y_{\hat{s}})\right)(y_s)^q ds \\
&+ \int_0^{t\wedge\theta_R} qk_3(s)\sqrt{y_{\hat{s}}}\phi(y_{\hat{s}})(y_s)^q dW_s \\
\leq \; & (x_0)^q + q\int_0^t \left[k_1(s) + \left(\frac{q-1}{2}k_{3,\max}^2 K_\phi^2 - k_{2,\min}\right)y_{\hat{s}}\right](y_s)^q \, \mathbb{I}_{(0,t\wedge\theta_R)}(s)ds \\
&+M_t \\
\leq \; & (x_0)^q + q\int_0^t k_1(s)(y_s)^q \, \mathbb{I}_{(0,t\wedge\theta_R)}(s)ds + M_t,
\end{aligned}$$

where the last inequality is valid for $q \leq 1 + 2k_{2,\min}/(k_{3,\max}K_\phi)^2$ and

$$M_t := \int_0^{t\wedge\theta_R} qk_3(s)\sqrt{y_{\hat{s}}}\phi(y_{\hat{s}})(y_s)^q dW_s.$$

Taking expectations and using that $\mathbb{E}M_t = 0$ we get

$$\mathbb{E}(y_{t\wedge\theta_R})^q \leq \mathbb{E}(x_0)^q + qk_{1,\max}\int_0^t \mathbb{E}(y_{s\wedge\theta_R})^q ds.$$

Application of the Gronwall inequality implies

$$\mathbb{E}(y_{t\wedge\theta_R})^q \leq \mathbb{E}(x_0)^q e^{qk_{1,\max}T}.$$

We have that

$$(y_{t\wedge\theta_R})^q = (y_{\theta_R})^q \mathbb{I}_{(\theta_R\leq t)} + (y_t)^q \mathbb{I}_{(t<\theta_R)} = R^q \mathbb{I}_{(\theta_R\leq t)} + (y_t)^q \mathbb{I}_{(t<\theta_R)}.$$

Thus taking expectations in the above inequality and using the estimated upper bound for $\mathbb{E}(y_{t\wedge\theta_R})^q$ we arrive at

$$\mathbb{E}(y_t)^q \mathbb{I}_{(t<\theta_R)} \leq \mathbb{E}(x_0)^q e^{qk_{1,\max}T}$$

and taking limits in both sides as $R \to \infty$ we get that

$$\lim_{R\to\infty} \mathbb{E}(y_t)^q \mathbb{I}_{(t<\theta_R)} \leq \mathbb{E}(x_0)^q e^{qk_{1,\max}T}.$$

Fix t. The sequence $(y_t)^q \mathbb{I}_{(t<\theta_R)}$ is non-decreasing in R since θ_R is increasing in R and $t \wedge \theta_R \to t$ as $R \to \infty$ and $(y_t)^q \mathbb{I}_{(t<\theta_R)} \to (y_t)^q$ as $R \to \infty$, thus the monotone convergence theorem implies

$$\mathbb{E}(y_t)^q \leq \mathbb{E}(x_0)^q e^{qk_{1,\max}T},$$

for any $q \leq 1 + \frac{2k_{2,\min}}{(k_{3,\max}K_\phi)^2}$. Following the same lines as in Lemma 2.4.7, i.e. using again Itô's formula on $(y_t)^p$, taking the supremum and then using Doob's martingale inequality on the diffusion term we obtain the desired result. Note that in this last step we need $2k_{2,\min} > 7(k_{3,\max}K_\phi)^2$. □

Remark 2.4.9

(i) Proposition 2.4.5 implies that our explicit numerical scheme converges in the mean-square sense. Moreover, by (2.4.4) we get that our numerical scheme preserves positivity, which is a desirable modelling property ([AGKR10], [KGR08]). Example (2.4.1) covers the $3/2$*-model (2.1.2), in the case where* $\phi(\cdot), k_1(\cdot), k_2(\cdot), k_3(\cdot)$ *are constant and super-linear problems both in drift and diffusion.*

(ii) Moreover, note that in the analysis that we followed, we did not discretize the coefficients k_i*. In general, by Theorem 2.2.2, we are free to discretize any of the functions* $k_i(\cdot), i = 1, 2, 3,$ *at any degree. Thus, we can fully discretize every* $k_i(\cdot), i = 1, 2, 3,$ *meaning that (2.4.2) will become*

$$y_t = y_{t_n} + \int_{t_n}^t (k_1(t_n) - k_2(t_n)y_{t_n})y_s ds + \int_{t_n}^t k_3(t_n)\sqrt{y_{t_n}}\phi(y_{t_n})y_s dW_s,$$

for $t \in [t_n, t_{n+1}]$*, or semi-discretize every* $k_i(\cdot), i = 1, 2, 3,$

$$y_t = y_{t_n} + \int_{t_n}^t (\hat{k_1}(s, t_n) - \hat{k_2}(s, t_n)y_{t_n})y_s ds + \int_{t_n}^t \hat{k_3}(s, t_n)\sqrt{y_{t_n}}\phi(y_{t_n})y_s dW_s,$$

for $t \in [t_n, t_{n+1}]$*, where* $\hat{k}_i(t,t) = k_i(t), i = 1,2,3$*. The only difference in that situation is that we require* $\hat{k}_i(\cdot,\cdot), i = 1,2,3$ *to be locally Lipschitz in both variables.*

(iii) *One more point of discussion is the dependence on* ω *that we can assume on the coefficients* k_i*. Specifically, we consider the more general SDE*

$$x_t = x_0 + \int_0^t a_\omega(s,x_s)ds + \int_0^t b_\omega(s,x_s)dW_s, \quad t \in [0,T].$$

Then, assuming that it admits a unique strong solution, our method seems to work. In the example discussed here, an extra condition on the coefficients k_i *would be of the form*

$$|k_i(t,\omega)| \le C, t \in [0,T], \omega \in \Omega, i = 1,2,3.$$

(iv) *We illustrate our method in the case* $\phi(x) = \sin(x)$*. Then the diffusion term* $b(x)$ *takes positive and negative values and thus the method presented in [NS14] does not work since it requires* $b(x) > 0$ *in order to use the Lamperti-type transformation; for the same reason the Milstein method [HMS13] fails since [HMS13, Assumption 2.7] is violated. The only method that we know and can be used for this situation is the tamed Euler method ([HJK12], [HJ15]) but the drawback is that it does not preserve positivity.*

Below, we compare our scheme, in the case where $k_1(\cdot), k_2(\cdot), k_3(\cdot)$ *are constant, with the tamed Euler method in [HJ15]. Figure 2.1 shows that for "good" data the two methods are close. Choosing different data, we see that tamed Euler (2.1.5) takes negative values, even in the first step. In particular we see, that by altering the parameters we get the results presented in Table 2.1 and shown in Figure 2.2. Note that if the tamed Euler takes a negative value, it explodes in the next step because of the* $3/2$*-term, while taking the value zero in a step results in zero terms for all the following steps.*

□

2.4.2 Example II.

Consider the following stochastic differential equation:

$$x_t = x_0 + \int_0^t (k_1(s)x_s - k_2(s)x_s^{2r-1})ds + \int_0^t k_3(s)x_s^r dW_s, \quad t \in [0,T], \tag{2.4.5}$$

Set of Parameters $(x_0, k_1, k_2, k_3, \Delta, T)$	Time of first negative step	Value of step
$(1, 1, 1000, 1, 10^{-3}, 1)$	1	−0.18
$(1, 1000, 1, 1, 10^{-3}, 1)$	27	−17.69

Tab. 2.1: Negative values of the tamed Euler scheme (2.1.5) for the Heston 3/2-model.

where x_0 is independent of all $\{W_t\}_{0\leq t\leq T}, x_0 \in \mathcal{L}^p(\Omega,\mathbb{R})$ for some $2 < p \leq \frac{r-1}{4(3-2r)} + \frac{r-1}{2(3-2r)}\frac{k_{2,min}}{(k_{3,max})^2}$ and $x_0 > 0$ a.s., $k_1(\cdot), k_2(\cdot), k_3(\cdot)$ are positive and bounded functions with $2k_{2,\min} > \frac{25-9r}{r-1}k_{3,\max}^2$ and $1 < r < 3/2$.

Lemma 2.4.10 *[Positivity of (x_t)] In the previous setting it holds that $x_t > 0$ a.s.* □

Proof of Lemma 2.4.10. The proof follows the same lines as the proof of Lemma 2.4.6. Nevertheless, we give the details in Appendix C.2. □

The following Lemma shows uniform bounds of p-moments of (x_t).

Lemma 2.4.11 *In the previous setting it holds that*

$$\mathbb{E}(\sup_{0\leq t\leq T}(x_t)^p) < A_1,$$

for some $A_1 > 0$ and any $2 < p \leq \frac{3}{2} - r + \frac{k_{2,min}}{(k_{3,max})^2}$. □

Proof of Lemma 2.4.11. The proof follows the same lines as the proof of Lemma 2.4.7. The details are given in Appendix C.3. □

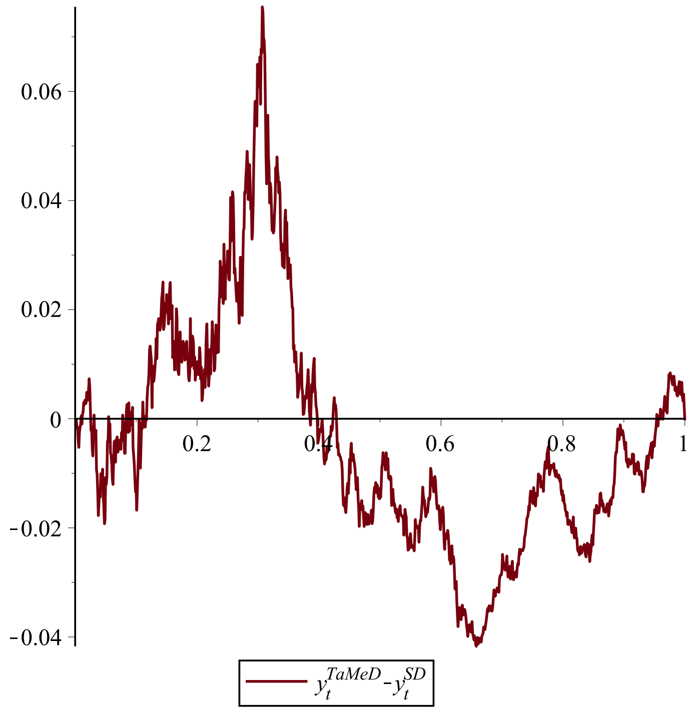
Model (2.4.5) has super-linear drift and diffusion coefficients. We study the numerical approximation of (2.4.5). We propose the following semi-discrete numerical scheme for the transformed process $z_t = (x_t)^{2r-2}$ of (2.4.5),

$$y_t = y_{t_n} + \int_{t_n}^t (K_1(s) - K_2(s)y_{t_n})y_s ds + \int_{t_n}^t K_3(s)\sqrt{y_{t_n}}y_s dW_s,\ t \in [t_n, t_{n+1}],$$

for $n \leq T/\Delta$ and $y_0 = x_0$ a.s., where
(2.4.6)

$$K_1(s) = (2r-2)k_1(s), \quad K_2(s) = (2r-2)k_2(s) - \frac{(2r-2)(2r-3)}{2}k_3^2(s),$$

Fig. 2.1: Difference between the semi-discrete scheme and the tamed Euler scheme (2.1.5) for $x_0 = 1, k_1 = 1, k_2 = 4, k_3 = 1, \Delta = 10^{-3}, T = 1$.



$$K_3(s) = (2r-2)k_3(s), \tag{2.4.7}$$

or in a more compact form,

$$y_t = y_0 + \int_0^t (K_1(s) - K_2(s)y_{\hat{s}})y_s ds + \int_0^t K_3(s)\sqrt{y_{\hat{s}}}y_s dW_s, \tag{2.4.8}$$

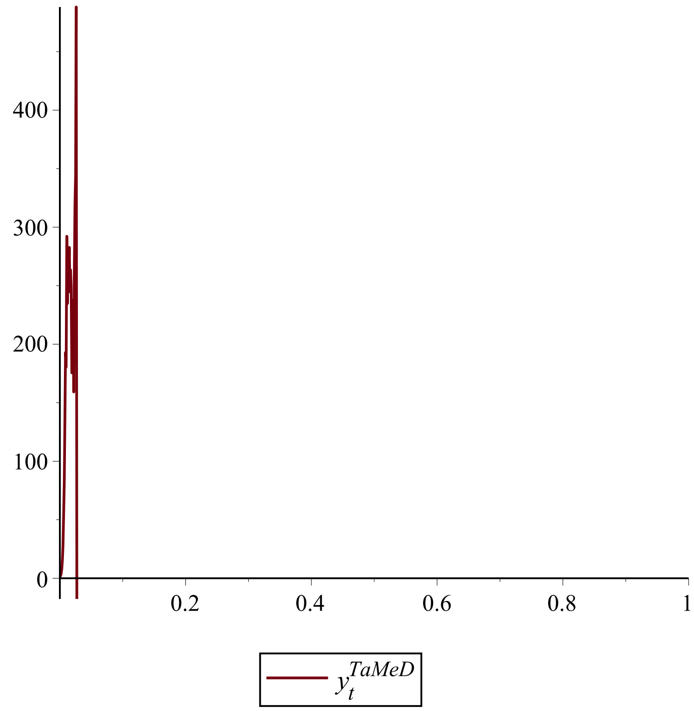
where $\hat{s} = t_n$, when $s \in [t_n, t_{n+1})$. The linear SDE (2.4.8) has a solution which, by use of Itô's formula, has the explicit form

$$y_t = x_0 \exp\Big\{ \int_0^t \Big(K_1(s) - K_2(s)y_{\hat{s}} - K_3^2(s)\frac{y_{\hat{s}}}{2}\Big) ds + \int_0^t K_3(s)\sqrt{y_{\hat{s}}}dW_s \Big\},$$

where $y_t = y_t(t_0, x_0)$.

The transformation of (2.4.5). Application of Itô's formula to the

Fig. 2.2: The tamed Euler method (2.1.5) does not preserve positivity, $x_0 = 1, k_1 = 1000, k_2 = 4, k_3 = 1, \Delta = 10^{-3}, T = 1$.



function $z(t,x) = x^{2r-2}$, implies

$$\begin{aligned} z_t &= z_0 + \int_0^t \big[(2r-2)x_s^{2r-3}(k_1(s)x_s - k_2(s)x_s^{2r-1}) \\ &\quad + \frac{(2r-2)(2r-3)}{2}x_s^{2r-4}k_3^2(s)x_s^{2r}\big]ds + \int_0^t (2r-2)k_3(s)x_s^{2r-3}x_s^r dW_s \\ &= z_0 + \int_0^t \big[k_1(s)(2r-2)x_s^{2r-2} - (2r-2)k_2(s)x_s^{4r-4} \\ &\quad + \frac{(2r-2)(2r-3)}{2}k_3^2(s)x_s^{4r-4}\big]ds + \int_0^t (2r-2)k_3(s)x_s^{3r-3}dW_s \\ &= z_0 + \int_0^t (K_1(s)z_s - K_2(s)z_s^2)ds + \int_0^t K_3(s)z_s^{3/2}dW_s, \end{aligned}$$

where $K_1(\cdot), K_2(\cdot), K_3(\cdot)$ are given by (2.4.6) and (2.4.7).

In order to use Proposition 2.4.5 we have to verify that

$$K_1(s) > 0, \quad K_2(s) > 0, \quad K_3(s) > 0, \quad 2K_{2,\min} > 7K_{3,\max}^2.$$

Since $1 < r < 3/2$ we immediately have $K_1(s) > 0$ and $K_3(s) > 0$. Moreover

$$K_2(s) = (2r-2)k_2(s) - \frac{(2r-2)(2r-3)}{2}k_3^2(s) > \frac{(2r-2)}{2}k_{3,\max}^2(4-2r) > 0,$$

and is easy to see that

$$2K_{2,\min} > 7K_{3,\max}^2.$$

Proposition 2.4.12 *In the previous setting, the following convergence to the true solution of (2.4.5) in the mean-square sense holds,*

$$\lim_{\Delta\to 0} \mathbb{E} \sup_{0\leq t\leq T} |y_t^{\frac{1}{2r-2}} - x_t|^2 = 0. \tag{2.4.9}$$

□

Proof of Proposition 2.4.12.

In order to prove the proposition, we first transform the original SDE (2.4.5) to an SDE of the type (2.4.1), later on verify the assumptions of Example I to use Proposition 2.4.5, and in the end make the necessary arrangements for the approximation of the original SDE.

Convergence Result.

We use the following inequality implied by the mean value theorem

$$|y_t^{\frac{1}{2r-2}} - x_t| = |y_t^{\frac{1}{2r-2}} - z_t^{\frac{1}{2r-2}}| \leq \frac{1}{2r-2}\left(|y_t|^{\frac{1}{2r-2}-1} + |z_t|^{\frac{1}{2r-2}-1}\right)|z_t - y_t|,$$

to get

$$|y_t^{\frac{1}{2r-2}} - x_t|^2 \leq \frac{2}{(2r-2)^2}\left(|y_t|^{\frac{3-2r}{r-1}} + |z_t|^{\frac{3-2r}{r-1}}\right)|z_t - y_t|^2.$$

Set the stopping time $\theta_R = \inf\{t \in [0,T] : |y_t| > R \text{ or } |x_t| > R\}$, for some $R > 0$ big enough. Taking the supremum and then expectations in the above inequality yields,

$$\mathbb{E}\sup_{0\le t\le T}|y_t^{\frac{1}{2r-2}}-x_t|^2 \le c_r\Big[\mathbb{E}\sup_{0\le t\le T}\left(|y_{t\wedge\theta_R}|^{\frac{3-2r}{r-1}}+|z_{t\wedge\theta_R}|^{\frac{3-2r}{r-1}}\right)|z_{t\wedge\theta_R}-y_{t\wedge\theta_R}|^2$$
$$+\mathbb{E}\sup_{0\le t\le T}\left(|y_t|^{\frac{3-2r}{r-1}}+|z_t|^{\frac{3-2r}{r-1}}\right)|z_t-y_t|^2\mathbb{I}_{(\theta_R\le t)}\Big]$$
$$\le c_{r,R}\mathbb{E}\sup_{0\le t\le T}|z_{t\wedge\theta_R}-y_{t\wedge\theta_R}|^2+c_r\frac{2\delta}{p}\mathbb{E}\sup_{0\le t\le T}\left(|y_t|^{\frac{3-2r}{r-1}}+|z_t|^{\frac{3-2r}{r-1}}\right)^{p/2}|z_t-y_t|^p$$
$$+c_r\frac{(p-2)}{p\delta^{2/(p-2)}}\mathbb{P}(\theta_R\le T),$$

where in the second step we have applied the Young inequality,

$$ab\le\frac{\delta}{w}a^w+\frac{1}{q\delta^{q/w}}b^q,$$

for $a=\sup_{0\le t\le T}\left(|y_t|^{\frac{3-2r}{r-1}}+|z_t|^{\frac{3-2r}{r-1}}\right)|z_t-y_t|^2, b=\mathbb{I}_{(\theta_R\le t)}, w=p/2, q=p/(p-2), \delta>0$, and

$$c_r=\frac{2}{(2r-2)^2},\quad c_{r,R}=2c_rR^{\frac{3-2r}{r-1}}.$$

(For all $t<\theta_R$ it holds that $|x_t|\le R$ or $|z_t|\le R$.) It holds that

$$\begin{aligned}\mathbb{P}(\theta_R\le T) &\le \mathbb{E}\left(\mathbb{I}_{(\theta_R\le T)}\frac{|y_{\theta_R}|^p}{R^p}\right)+\mathbb{E}\left(\mathbb{I}_{(\theta_R\le T)}\frac{|x_{\theta_R}|^p}{R^p}\right)\\ &\le \frac{1}{R^p}\left(\mathbb{E}\sup_{0\le t\le T}|y_t|^p+\mathbb{E}\sup_{0\le t\le T}|x_t|^p\right)\le\frac{2A}{R^p},\end{aligned}$$

where A is the maximum of the bounding moment constants of (y_t) and (x_t). Moreover, we have that,

$$\mathbb{E}\sup_{0\le t\le T}\left(|y_t|^{\frac{3-2r}{r-1}}+|z_t|^{\frac{3-2r}{r-1}}\right)^{\frac{p}{2}}|z_t-y_t|^p\le 2^{\frac{3p}{2}-2}\mathbb{E}\sup_{0\le t\le T}(|y_t|^{\frac{(3-2r)p}{2(r-1)}}+|z_t|^{\frac{(3-2r)p}{2(r-1)}})$$
$$\times(|z_t|^p+|y_t|^p)$$
$$\le 2^{\frac{3p}{2}-2}\mathbb{E}\sup_{0\le t\le T}(|y_t|^{\frac{(3-2r)p}{2(r-1)}}|z_t|^p+|y_t|^{(\frac{3-2r}{2(r-1)}+1)p}+|z_t|^{\frac{(3-2r)p}{2(r-1)}}|y_t|^p+|z_t|^{(\frac{3-2r}{2(r-1)}+1)p})$$
$$\le 2^{\frac{3p}{2}-2}\mathbb{E}\sup_{0\le t\le T}(\frac{|y_t|^{\frac{3-2r}{r-1}p}}{2}+\frac{|z_t|^{2p}}{2}+|y_t|^{\frac{p}{2(r-1)}}+\frac{|z_t|^{\frac{3-2r}{r-1}p}}{2}+\frac{|y_t|^{2p}}{2}+|z_t|^{\frac{p}{2(r-1)}}),$$

where we have used again the Young inequality. When $\frac{5}{4} < r < \frac{3}{2}$ we have that $\frac{3-2r}{r-1} < \frac{1}{2(r-1)} < 2$, thus it suffices to bound the moments of $|z_t|^{2p}$ and $|y_t|^{2p}$. Note that by Lemma 2.4.7 the uniform bound for the moment of $(z_t)^{2p}$ holds when $2 < p \leq \frac{k_{2,min}}{2(k_{3,max})^2}$ and by Lemma 2.4.8 the uniform bound for the moment of $(y_t)^{2p}$ is valid for any $2 < p \leq \frac{1}{8} + \frac{k_{2,min}}{4(k_{3,max})^2}$, thus for $2 < p \leq \frac{k_{2,min}}{2(k_{3,max})^2} \bigwedge \frac{1}{8} + \frac{k_{2,min}}{4(k_{3,max})^2}$ we get that

$$\mathbb{E} \sup_{0 \leq t \leq T} \left(|z_t|^{2p} + |y_t|^{2p} \right) < A,$$

for some $A > 0$. (We also have to ensure that Lemma 2.4.11 holds, thus we have to choose p such that $2 < p \leq \frac{3}{2} - r + \frac{k_{2,min}}{(k_{3,max})^2} \bigwedge \frac{k_{2,min}}{2(k_{3,max})^2} \bigwedge \frac{1}{8} + \frac{k_{2,min}}{4(k_{3,max})^2}$ or equivalently we have to choose p such that $2 < p \leq \frac{1}{8} + \frac{k_{2,min}}{4(k_{3,max})^2}$ whose existence is ensured by the condition $2k_{2,min} \geq 15(k_{3,max})^2$.)

In the case $1 < r < \frac{5}{4}$ it suffices to bound the moments of $|z_t|^{\frac{3-2r}{r-1}p}$ and $|y_t|^{\frac{3-2r}{r-1}p}$. Again by Lemma 2.4.7 the uniform bound for the moment of $|z_t|^{\frac{3-2r}{r-1}p}$ holds when $2 < p \leq \frac{r-1}{3-2r} \frac{k_{2,min}}{(k_{3,max})^2}$ and by Lemma 2.4.8 the uniform bound for the moment of $|y_t|^{\frac{3-2r}{r-1}p}$ is valid for any $2 < p \leq \frac{r-1}{4(3-2r)} + \frac{r-1}{2(3-2r)} \frac{k_{2,min}}{(k_{3,max})^2}$, thus for $2 < p \leq \frac{r-1}{3-2r} \frac{k_{2,min}}{(k_{3,max})^2} \bigwedge \frac{r-1}{4(3-2r)} + \frac{r-1}{2(3-2)r} \frac{k_{2,min}}{(k_{3,max})^2}$ we get that

$$\mathbb{E} \sup_{0 \leq t \leq T} \left(|z_t|^{\frac{3-2r}{r-1}p} + |y_t|^{\frac{3-2r}{r-1}p} \right) < A,$$

for some $A > 0$. (We also have to ensure that Lemma 2.4.11 holds, thus we have to choose p such that $2 < p \leq \frac{3}{2} - r + \frac{k_{2,min}}{(k_{3,max})^2} \bigwedge \frac{r-1}{3-2r} \frac{k_{2,min}}{(k_{3,max})^2} \bigwedge \frac{r-1}{4(3-2r)} + \frac{r-1}{2(3-2r)} \frac{k_{2,min}}{(k_{3,max})^2}$ or equivalently we have to choose p such that $2 < p \leq \frac{r-1}{4(3-2r)} + \frac{r-1}{2(3-2r)} \frac{k_{2,min}}{(k_{3,max})^2}$ whose existence is ensured by the condition $2k_{2,min} \geq \frac{25-9r}{r-1}(k_{3,max})^2$.)Thus, by the two preceding parenthetical remarks and the condition

$$2k_{2,min} \geq \left(\frac{25-9r}{r-1} \bigvee 15 \right) (k_{3,max})^2$$

or equivalently

$$2k_{2,min} \geq \frac{25-9r}{r-1} (k_{3,max})^2$$

we get the bound

$$\mathbb{E} \sup_{0 \leq t \leq T} \left(|y_t|^{\frac{3-2r}{r-1}} + |z_t|^{\frac{3-2r}{r-1}} \right)^{p/2} |z_t - y_t|^p < C(p)A,$$

where $C(p)$ is a constant depending on p. Collecting all the estimates together, we get

$$\begin{aligned}\mathbb{E}\sup_{0\le t\le T}|y_t^{\frac{1}{2r-2}}-x_t|^2 \quad\le\quad & c_{r,R}\mathbb{E}\sup_{0\le t\le T}|z_{t\wedge\theta_R}-y_{t\wedge\theta_R}|^2+c_r\frac{C(p)A}{p}\delta\\ & +c_r\frac{2(p-2)A}{p}\frac{1}{\delta^{2/(p-2)}R^p}:=I_1+I_2+I_3.\end{aligned}$$

Given any $\epsilon>0$, we may first choose δ such that $I_2<\epsilon/3$, then choose R such that $I_3<\epsilon/3$, and finally Δ such that $I_1<\epsilon/3$, which is justified by Proposition 2.4.5 to get that $\mathbb{E}\sup_{0\le t\le T}|y_t^{\frac{1}{2r-2}}-x_t|^2<\epsilon$, as required to verify (2.4.9).

Remark 2.4.13 *Proposition 2.4.12 implies that our explicit numerical scheme converges in the mean-square sense. Moreover, we get that our numerical scheme preserves positivity. SDE (2.4.5) covers super-linear problems both in drift and diffusion.* □

2.4.3 Example III.

Consider the following stochastic differential equation:

$$x_t=x_0+\int_0^t(k_1(s)x_s-k_2(s)x_s^q)ds+\int_0^t k_3(s)x_s^r\phi(x_s)dW_s,\ t\in[0,T], \tag{2.4.10}$$

where $\phi(\cdot)$ is a locally Lipschitz and bounded function with locally Lipschitz constant C_R^ϕ, bounding constant K_ϕ, x_0 is independent of all $\{W_t\}_{0\le t\le T}$, $x_0\in\mathcal{L}^p(\Omega,\mathbb{R})$ for every $2<p,\mathbb{E}|\ln x_0|<\infty$ and $x_0>0$ a.s., $k_1(\cdot),k_2(\cdot),k_3(\cdot)$ are positive and bounded functions and q is odd with $q>2r-1$ where $3/2<r<2$. The above conditions on the parameters imply the uniform bound of $|x_t|^p$ as shown in the following result.

Lemma 2.4.14 *[Moment bound for original SDE] In the previous setting it holds that*

$$\mathbb{E}(\sup_{0\le t\le T}|x_t|^p)<A_1,$$

for some $A_1>0$ and every $p>2$. □

Proof of Lemma 2.4.14. The proof follows the same lines as the proof of Lemma 2.4.6. Nevertheless, we present the proof in Appendix C.5. □

Model (2.4.10) has super-linear drift and diffusion coefficients. We study the numerical approximation of (2.4.10). We propose the following semi-discrete numerical scheme for (2.4.10)

$$y_t = y_{t_n} + \int_{t_n}^{t} (k_1(s) - k_2(s) y_{t_n}^{q-1}) y_s ds + \int_{t_n}^{t} k_3(s) y_{t_n}^{r-1} \phi(y_{t_n}) y_s dW_s, \tag{2.4.11}$$

where $t \in [t_n, t_{n+1}]$, for $n \leq T/\Delta$ and $y_0 = x_0$ a.s.; (2.4.11) in more compact form reads

$$y_t = y_0 + \int_{0}^{t} (k_1(s) - k_2(s) y_{\hat{s}}^{q-1}) y_s ds + \int_{0}^{t} k_3(s) y_{\hat{s}}^{r-1} \phi(y_{\hat{s}}) y_s dW_s, \tag{2.4.12}$$

where $\hat{s} = t_n$, when $s \in [t_n, t_{n+1})$. The linear SDE (2.4.12) has a solution which, by use of Itô's formula, has the explicit form [KP95, Ch. 4.4, (4.10)]

$$y_t = x_0 \exp\{\int_0^t (k_1(s) - k_2(s) y_{\hat{s}}^{q-1} - k_3^2(s) \frac{y_{\hat{s}}^{2r-2} \phi^2(y_{\hat{s}})}{2}) ds + \int_0^t k_3(s) y_{\hat{s}}^{r-1} \phi(y_{\hat{s}}) dW_s\},$$

where $y_t = y_t(t_0, x_0)$.

Proposition 2.4.15 *The following convergence to the true solution of (2.4.10) in the mean-square sense holds,*

$$\lim_{\Delta \to 0} \mathbb{E} \sup_{0 \leq t \leq T} |y_t - x_t|^2 = 0.$$

□

Proof of Proposition 2.4.15.

In order to prove the proposition, we just need to verify the assumptions of Theorem 2.2.2. Let

$$a(s,x) = k_1(s)x - k_2(s)x^q, \qquad f(s,r,x,y) = (k_1(s) - k_2(s)x^{q-1})y$$
$$b(s,x) = k_3(s)x^r \phi(x), \qquad g(s,r,x,y) = k_3(s)x^{r-1}\phi(x)y.$$

We verify Assumption 2.2.1 for f. The conditions on the parameters imply that $q > 2$. Let $R > 0$ such that $|x_1| \vee |x_2| \vee |y_1| \vee |y_2| \vee |s| \vee |r| \leq R$. We have that

$$
\begin{aligned}
&|f(s,r,x_1,y_1)-f(s,r,x_2,y_2)| = |(k_1(s)(y_1-y_2)-k_2(s)(x_1^{q-1}y_1-x_2^{q-1}y_2)| \\
&\le |k_1(s)||y_1-y_2| + |k_{2,\max}|(|x_2|^{q-1}|y_1-y_2| + |y_1||x_1^{q-1}-x_2^{q-1}|) \\
&\le (|k_{1,\max}| + |k_{2,\max}|R^{q-1})|y_1-y_2| + |k_{2,\max}|R|x_1^{q-1}-x_2^{q-1}| \\
&\le (|k_{1,\max}| + |k_{2,\max}|R^{q-1})|y_1-y_2| + 2|k_{2,\max}|(q-1)R^{q-1}|x_1-x_2| \\
&\le C_R\left(|x_1-x_2| + |y_1-y_2|\right),
\end{aligned}
$$

where we have applied the mean value theorem for the function x^{q-1}, thus Assumption 2.2.1 holds for f with $C_R := (|k_{1,\max}|+|k_{2,\max}|R^{q-1}) \vee (2|k_{2,\max}|(q-1)R^{q-1})$.

We verify Assumption 2.2.1 for g. Since $1/2 < r-1 < 1$ we have that $g_1(x) = x^{r-1}$ is locally 1/2-Hölder continuous in x, i.e.

$$
|g_1(x_1) - g_1(x_2)| \le C_R\sqrt{|x_1-x_2|}. \tag{2.4.13}
$$

Let $R>0$ such that $|x_1|\vee|x_2|\vee|y_1|\vee|y_2|\vee|s|\vee|r| \le R$. We have that

$$
\begin{aligned}
&|g(s,r,x_1,y_1)-g(s,r,x_2,y_2)| = \left|k_3(s)x_1^{r-1}\phi(x_1)y_1 - k_3(s)x_2^{r-1}\phi(x_2)y_2\right| \\
&\le |k_{3,\max}|\big(|x_1|^{r-1}|\phi(x_1)||y_1-y_2| + |y_2|\big|x_1^{r-1}\phi(x_1) - x_1^{r-1}\phi(x_2) \\
&\quad + x_1^{r-1}\phi(x_2) - x_2^{r-1}\phi(x_2)\big|\big) \\
&\le |k_{3,\max}|\left(K_\phi R^{r-1}|y_1-y_2| + R|x_1|^{r-1}|\phi(x_1)-\phi(x_2)| + RK_\phi|x_1^{r-1}-x_2^{r-1}|\right) \\
&\le |k_{3,\max}|\left(K_\phi R^{r-1}|y_1-y_2| + R^rC_R^\phi|x_1-x_2| + RK_\phi\sqrt{|x_1-x_2|}\right) \\
&\le C_R\left(|x_1-x_2| + |y_1-y_2| + \sqrt{|x_1-x_2|}\right),
\end{aligned}
$$

where we have used (2.4.13) and $C_R := |k_{3,\max}|\left(C_R^\phi R^r \vee K_\phi R^{r-1} \vee K_\phi R\right)$. Thus, Assumption 2.2.1 holds for g. Lemmata 2.4.14 and 2.4.17 complete the proof.

Lemma 2.4.16 *[Positivity of (x_t)] In the previous setting it holds that $x_t > 0$ a.s.* □

Proof of Lemma 2.4.16. One can use again the arguments in Lemma 2.4.6 applying Itô's formula on $(x_t)^{-2}$. We present an alternative proof in Appendix C.4. □

Lemma 2.4.17 *[Moment bound for Semi-Discrete approximation] In the previous setting it holds that*

$$\mathbb{E}(\sup_{0\leq t\leq T}(y_t)^p) < A_2,$$

for some $A_2 > 0$ and for every $p > 2$. □

Proof of Lemma 2.4.17. The lemma is proved in much the same way as Lemma 2.4.8; see also Appendix C.6. □

2.5 Numerical Experiments.

We study the numerical approximation of the following SDE,

$$x_t = x_0 + \int_0^t (k_1 x_s - k_2 x_s^2)ds + \int_0^t k_3 x_s^{3/2} dW_s, \quad t \in [0, T], \tag{2.5.1}$$

where x_0 is independent of all $\{W_t\}_{0\leq t\leq T}, x_0 \in \mathcal{L}^{4p}(\Omega, \mathbb{R})$ for some $2 < p$ and $x_0 > 0$ a.s., $\mathbb{E}(x_0)^{-2} < A$, k_1, k_2, k_3 are positive constants with $k_2 > \frac{7}{2}(k_3)^2$. Model (2.5.1) has super-linear drift and diffusion coefficients.

In Proposition 2.4.5 we have shown that the semi-discrete numerical scheme[2] (in a more general setting with time-varying coefficients)

$$y_t^{SD} = y_{t_n} + \int_{t_n}^t (k_1 - k_2 y_{t_n}) y_s ds + \int_{t_n}^t k_3 \sqrt{y_{t_n}} y_s dW_s, \quad t \in [t_n, t_{n+1}], \tag{2.5.2}$$

for $n \leq T/\Delta$ and $y_0 = x_0$ a.s., or in a more compact form,

$$y_t^{SD} = y_0 + \int_0^t (k_1 - k_2 y_{\hat{s}}) y_s ds + \int_0^t k_3 \sqrt{y_{\hat{s}}} y_s dW_s, \tag{2.5.3}$$

where $\hat{s} = t_n$, when $s \in [t_n, t_{n+1})$, converges to the true solution of (2.5.1) in the mean-square sense, that is

$$\lim_{\Delta\to 0} \mathbb{E} \sup_{0\leq t\leq T} |y_t^{SD} - x_t|^2 = 0. \tag{2.5.4}$$

Relation (2.5.4) does not show the order of convergence. In the following, we compute experimentally the order of convergence.

[2] The existence and uniqueness of y_t^{SD} is shown in Appendix C.1.

The linear SDE (2.5.3) has a solution which, by use of Itô's formula, has the explicit form

$$y_t^{SD} = x_0 \exp\left\{ \int_0^t \left(k_1 - k_2 y_{\hat{s}} - k_3^2 \frac{y_{\hat{s}}}{2} \right) ds + \int_0^t k_3 \sqrt{y_{\hat{s}}} dW_s \right\}, \tag{2.5.5}$$

where $y_t = y_t(t_0, x_0)$. The semi-discrete scheme preserves positivity, which is a desirable modeling property.

In order to estimate the endpoint error $\epsilon = \mathbb{E}|y_T - x_T|$, where x_T is the exact solution of (2.5.1) and y_T is the semi-discrete approximation (2.5.5), we follow a standard procedure [KPS03, Sec. 3.3]. We compute M batches of L simulation paths. Each batch is estimated by

$$\hat{\epsilon_j} = \frac{1}{L} \sum_{i=1}^{L} |y_T^{i,j} - x_T^{i,j}|$$

and the Monte Carlo estimator of the error

$$\hat{\epsilon} = \frac{1}{M} \sum_{j=1}^{M} \hat{\epsilon}_j = \frac{1}{ML} \sum_{j=1}^{M} \sum_{i=1}^{L} |y_T^{i,j} - x_T^{i,j}|,$$

requires $M \cdot L$ Monte Carlo sample paths. When the batch size averages $L \geq 15$ they can be considered as Gaussian. A $100(1-\alpha)\%$ confidence interval for the error ϵ is determined by endpoints of the form

$$\hat{\epsilon} \pm t_{1-\alpha, M-1} \cdot \sqrt{\frac{1}{M(M-1)} \sum_{j=1}^{M} (\hat{\epsilon}_j - \hat{\epsilon})^2}.$$

We simulate $100 \cdot 100 = 10000$ paths[3]. The choice for $L = 100$ is considered in [KPS03, p.118]. We should not forget to change the student t-test quantile $t_{1-\alpha, M-1}$ when we change the number M of batches or the significance level α. Table 2.2 shows values of t-test quantiles for different values of M and α. In the experiments we consider 98% confidence intervals.

We discretize with a number of steps in power of 2. The iterative SD-procedure reads

$$y_{t_{n+1}}^{SD} = y_{t_n} \exp\left\{ \left(k_1 - k_2 y_{t_n} - \frac{k_3^2 y_{t_n}}{2} \right) \Delta + k_3 \sqrt{y_{t_n}} \Delta W_n \right\},$$

[3] We simulate with 3.06GHz Intel Pentium, 1.49GB of RAM in Matlab $R2014b$ Software. The effort made is just for the purpose of the order of convergence and not for the efficiency of the computer code-time.

M	10	20	30	40	60	100	200
$t_{0.9,M-1}$	1.83	1.73	1.70	1.68	1.67	1.66	1.65
$t_{0.95,M-1}$	2.26	2.09	2.05	2.02	2.00	1.98	1.97
$t_{0.98,M-1}$	2.82	2.54	2.46	2.43	2.39	**2.36**	2.35
$t_{0.99,M-1}$	3.25	2.86	2.76	2.71	2.66	2.63	2.60

Tab. 2.2: t-test quantiles, batches, level of confidence.

for $n = 0, \ldots, N-1$, where $\Delta W_n := W_{t_{n+1}} - W_{t_n}$ are the increments of the Brownian motion.

We want to compare our results with another method that preserves positivity. This is an implicit Milstein scheme proposed in [HMS13, Sec. 2.2], which takes the form

$$y^{HMS}_{t_{n+1}} = \frac{1}{2(k_2 + \frac{3}{4}(k_3)^2)\Delta}\Big(-(1-k_1\Delta)$$
$$+\sqrt{(1-k_1\Delta)^2 + 4(k_2 + \frac{3}{4}(k_3)^2)\Delta(y_{t_n} + k_3 y^{3/2}_{t_n}\Delta W_n + \frac{3}{4}(k_3)^2 y^2_{t_n}(\Delta W_n)^2)}\Big).$$

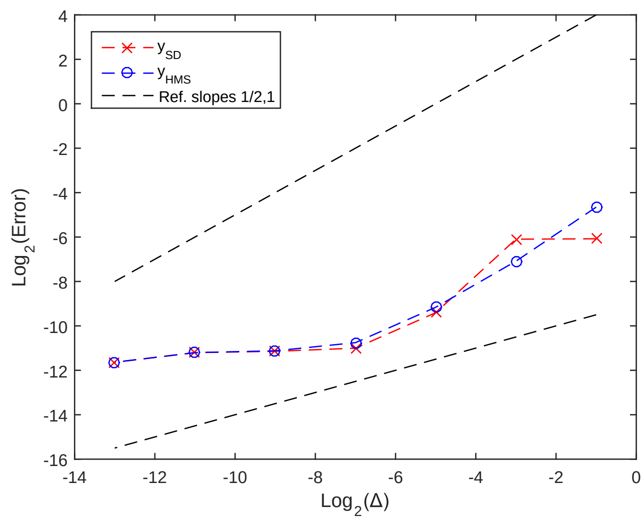
As a reference solution, we take the value of y^{HMS}_T at $\Delta = 2^{-14}$, as in the numerical experiment in [HMS13, Sec. 4.1], and in the second experiment y^{SD}_T at $\Delta = 2^{-14}$, since we have shown by (2.5.4) that it strongly converges to the exact solution. We plot in a $\log_2$-$\log_2$ scale. The results are presented in Tables 2.3 and 2.4 and they are also shown in Figure 2.3 for the first experiment (as the situation is quite similar for the other experiment).

Step Δ	98% SD-Error	98% HMS-Error
2^{-1}	$\mathbf{0.01478805} \pm 1.36 \cdot 10^{-5}$	$0.03968881 \pm 1.38 \cdot 10^{-5}$
2^{-3}	$0.01461442 \pm 1.43 \cdot 10^{-5}$	$\mathbf{0.0073555} \pm 1.43 \cdot 10^{-5}$
2^{-5}	$\mathbf{0.00147411} \pm 1.19 \cdot 10^{-5}$	$0.00174620 \pm 1.18 \cdot 10^{-5}$
2^{-7}	$\mathbf{0.0004872} \pm 7.92 \cdot 10^{-6}$	$0.0005787 \pm 9.96 \cdot 10^{-6}$
2^{-9}	$\mathbf{0.00044181} \pm 7.35 \cdot 10^{-6}$	$0.00044875 \pm 8.11 \cdot 10^{-6}$
2^{-11}	$\mathbf{0.00042386} \pm 7.62 \cdot 10^{-6}$	$0.00042411 \pm 7.63 \cdot 10^{-6}$
2^{-13}	$\mathbf{0.0003137} \pm 5.16 \cdot 10^{-6}$	$0.00031381 \pm 5.15 \cdot 10^{-6}$

Tab. 2.3: Error & step size of SD and HMS approximation of (2.5.1) with HMS exact solution and 32 digits of accuracy.

The following points of discussion are worth mentioning.

Fig. 2.3: SD and HMS method applied to SDE (2.5.1) with parameters $k_1 = 0.1, k_2 = \frac{\lambda}{2}(k_3)^2, k_3 = \sqrt{0.2}, \lambda = 700, x_0 = 1, T = 1$ and 32 digits of accuracy.



- The SD method and the HMS method are very close, with SD performing slightly better, except only for the step size $\Delta = 2^{-3}$. The same situation appears in both cases, i.e. independently of the choice of the reference solution, which is a positive feature of SD.

- A linear regression with the method of least squares fit, in the case one considers only the first four points with steps $\Delta = 2^{-1}, 2^{-3}, 2^{-5}, 2^{-7}$, produced values consistent with the strong order of convergence equal to 1 for both SD and HMS methods, whereas considering all the seven points, values close to $1/2$. Table 2.5 presents the exact values of order of convergence. We see that the order of convergence of SD for problem (2.5.1) is at least $1/2$.

- The confidence intervals are of such an order that indicates that we do not need to increase the number of batches M. All the above calculations are made evaluating with 32 digits.

- For small Δ it may happen that the global error will begin to increase as Δ is further decreased [KPS03, p.97]. This effect is due to the roundoff error which influences the calculated global error. In practice,

Step Δ	98% SD-Error	98% HMS-Error
2^{-1}	$\mathbf{0.01478262} \pm 1.36 \cdot 10^{-5}$	$0.03969424 \pm 1.38 \cdot 10^{-5}$
2^{-3}	$0.01460899 \pm 1.43 \cdot 10^{-5}$	$\mathbf{0.00736093} \pm 1.43 \cdot 10^{-5}$
2^{-5}	$\mathbf{0.0014687} \pm 1.18 \cdot 10^{-5}$	$0.00175161 \pm 1.18 \cdot 10^{-5}$
2^{-7}	$\mathbf{0.00048522} \pm 7.89 \cdot 10^{-6}$	$0.00058162 \pm 9.99 \cdot 10^{-6}$
2^{-9}	$\mathbf{0.00044126} \pm 7.35 \cdot 10^{-6}$	$0.00044942 \pm 8.15 \cdot 10^{-6}$
2^{-11}	$\mathbf{0.00042361} \pm 7.62 \cdot 10^{-6}$	$0.00042413 \pm 7.63 \cdot 10^{-6}$
2^{-13}	$\mathbf{0.00031367} \pm 5.14 \cdot 10^{-6}$	$0.00031384 \pm 5.12 \cdot 10^{-6}$

Tab. 2.4: Error & step size of SD and HMS approximation of (2.5.1) with SD exact solution and 32 digits of accuracy.

No	Order of SD with HMS ref.sol. (with SD ref.sol.)	Order of HMS with HMS ref.sol. (with SD ref.sol.)
4	0.904 (0.905)	1.019 (1.017)
7	0.511 (0.511)	0.556 (0.556)

Tab. 2.5: Order of convergence of SD and HMS approximation of (2.5.1) with HMS (SD) exact solution with 32 digits of accuracy.

that implies the existence of a minimum step size $\Delta_{\min}$, for each initial value problem, below which the accuracy of the approximations through a specific method cannot be improved.

- Convergence of a numerical scheme does not alone guarantee its practical value [KPS03, p.129]. It may be numerical unstable. Moreover, in practice, the computer time consumed to provide a desired level of accuracy, is of great importance. As mentioned in Footnote 3, we do not claim that the SD method performs well in that aspect, because of the exponential calculations involved. However, it seems that it can reach accuracy up to four digits, as fast as the HMS method.

- We would like to see the impact of the parameter λ. The SD method, seems to work, with the theoretical proof shown in Section 2.4.1, when λ is over 7. What happens below that range? The HMS method works for λ over 1/2. Moreover, as noted in Remark 2.4.9(iv), our method can cover more general cases, in contrast to HMS, by introducing the function $\phi(\cdot)$ in the diffusion part, or/and by assuming random coefficients $k_1(\cdot), k_2(\cdot), k_3(\cdot)$.

In the following, we present the situation when we change the parameters of SDE (2.5.1) in such a way that we are closer to the theoretical acceptable range (by lowering λ to 70). The error now is bigger and the rate of convergence drops to a half, for both the SD and HMS method. To be more precise we present the results in Tables 2.6 and 2.7.

Step Δ	98% SD-Error	98% HMS-Error
2^{-1}	$0.10228515 \pm 3.53 \cdot 10^{-4}$	$\mathbf{0.10105429} \pm 3.51 \cdot 10^{-4}$
2^{-3}	$\mathbf{0.01790118} \pm 3.38 \cdot 10^{-4}$	$0.02708432 \pm 3.39 \cdot 10^{-4}$
2^{-5}	$\mathbf{0.01243608} \pm 2.21 \cdot 10^{-4}$	$0.01352563 \pm 2.26 \cdot 10^{-4}$
2^{-7}	$\mathbf{0.01218537} \pm 2.29 \cdot 10^{-4}$	$0.01229934 \pm 2.31 \cdot 10^{-4}$
2^{-9}	$0.0122866 \pm 2.19 \cdot 10^{-4}$	$\mathbf{0.01228005} \pm 2.19 \cdot 10^{-4}$
2^{-11}	$0.01140109 \pm 2.15 \cdot 10^{-4}$	$\mathbf{0.01138454} \pm 2.18 \cdot 10^{-4}$
2^{-13}	$0.00869149 \pm 1.47 \cdot 10^{-4}$	$\mathbf{0.00868346} \pm 1.49 \cdot 10^{-4}$

Tab. 2.6: Error & step size of SD and HMS approximation of (2.5.1) with HMS exact solution and 32 digits of accuracy when $\lambda = 70$.

No	Order of SD	Order of HMS
4	0.487	0.506
7	0.214	0.237

Tab. 2.7: Order of convergence of SD and HMS approximation of (2.5.1) with HMS exact solution and 32 digits of accuracy when $\lambda = 70$.

Finally, we present the case with $\lambda = 7$. The rate of convergence drops dramatically. To be more precise the order of the SD method becomes 0.03 and the order of HMS 0.034.

- Regarding the tamed Euler method, a major drawback is that it does not preserve positivity. The proposed method SD and the implicit Milstein scheme HMS behave in a similar way for small values of Δ and retain this similarity as we lower the parameter λ close to its critical value. Nevertheless, the errors grow and the order of convergence drops as λ changes and the reason for that behavior is the fact that at the critical λ value we have moments explosions of the original process (x_t).

3. SUB-LINEAR STOCHASTIC DIFFERENTIAL EQUATIONS

3.1 Introduction.

Consider[1] the following stochastic models in Itô form

$$
\begin{cases} S_t = S_0 + \int_0^t \mu \cdot S_u du + \int_0^t (V_u)^p \cdot S_u dW_u, & t \in [0,T], \\ V_t = V_0 + \int_0^t (k_1 - k_2 V_s) ds + \int_0^t k_3 (V_s)^q d\widetilde{W}_s & t \in [0,T], \end{cases} \tag{3.1.1}
$$

[1] This chapter is based on joint work with Nikolaos Halidias, published in Journal of Probability and Statistics [HS15].

where S_t represents the underlying financially observable variable, V_t is the instantaneous volatility when $p = 1$ or the instantaneous variance when $p = 1/2$ and the Wiener processes $(W_t), (\widetilde{W}_t)$ have correlation ρ.

We assume that (V_t) is a *mean-reverting* CEV process of the above form, with the coefficients $k_i > 0$ for $i = 1, 2, 3$ and $q > 1/2$, since the process (V_t) has to be non-negative. To be more precise the above restriction on q implies that (V_t) is positive, i.e. 0 is unattainable, as well as non-explosive, i.e. ∞ is unattainable, as can be verified by the Feller's classification of boundaries, see Appendix F.1. (In particular, we get for the dynamics of the mean-reverting CEV process (V_t) of (3.1.1) a boundary behavior which is determined by the scale function (F.1.2) which reads

$$\begin{aligned} s(x) &= \int_c^x \exp\Big\{ -2\int_c^y \frac{k_1 - k_2 z}{(k_3)^2 z^{2q}} dz \Big\} dy \\ &= C\int_c^x \exp\Big\{ -\frac{2k_1}{(k_3)^2(1-2q)} y^{1-2q} + \frac{2k_2}{(k_3)^2(2-2q)} y^{2-2q} \Big\} dy, \end{aligned}$$

for any $x \in I$, where $C > 0$. Let $I = (0, \infty)$ and take $c = 1$. We compute

$$s(0+) = -C\int_0^1 \exp\{ -\frac{2k_1}{(k_3)^2(1-2q)} y^{1-2q} + \frac{2k_2}{(k_3)^2(2-2q)} y^{2-2q} \} dy = -\infty,$$

when $2q > 1$, thus by [KS88, Prop. 5.22c] we have that $\mathbb{P}(\inf_{0\leq t} V_t > 0) = 1$.) The steady-state level of V_t is[2] k_1/k_2 and the rate of mean-reversion is k_2.

System (3.1.1) for $p = q = 1/2$ is the Heston model. When $q = 1$ we get the Brennan-Schwartz model [BS80, Sec. II] that despite its simple form, cannot provide analytical expressions for S_t.

Process (V_t) for $q = 1/2$, is know as the CIR process, see Example 1.3.8, has received a lot of attention and we just mention the latest two contributions to the study of such processes (see [Alf13], [Hal15c] and references therein).

Process (V_t) for $1/2 \leq q \leq 1$ has been also considered for the dynamics of the short-term interest rate [CKLS92, (1)]. The stationary distribution of the process has also been derived in [AP07, Prop 2.2].

We aim for a positivity preserving scheme for the process (V_t). The scheme that we propose, and denote semi-discrete (SD), preserves the analytical property of (V_t) staying positive. The idea of the semi-discrete method is

[2] It holds that $\mathbb{E}V_t \to k_1/k_2$ as $t \to \infty$.

that we discretize a part of the original SDE and then apply Itô's formula (cf. [Hal12] where the method originally appeared and [Hal14], [Hal15d], [HS16]). The explicit Euler scheme fails to preserve positivity, as well as the standard Milstein scheme. We intend to apply the semi-discrete method for the numerical approximation of (V_t) in model (3.1.1) with $1/2 < q < 1$ and compare with other positivity preserving methods such as the balanced implicit method (BIM) (introduced by [MPS98, (3.2)] with the positivity preserving property [KS06, Sec. 5] and its stability properties [AK06]; see also [AK12] for an extended balanced method with better stability behavior) and the balanced Milstein method (BMM) [KS06, Th. 5.9].[3] Finally, we approximate the stochastic volatility model (3.1.1) with $p = 1/2$. In [KJ06] a thorough treatment can be found, where also another stochastic volatility model is suggested.

Section 3.2 provides the setting and the main results, Theorems 3.2.2 and 3.2.4, concerning the $\mathcal{L}^2$-convergence of the proposed semi-discrete method to the true solution of mean-reverting CEV processes of the form of the stochastic volatility in (3.1.1). The rate of mean-square convergence in Theorem 3.2.2 is logarithmic and in Theorem 3.2.4 is polynomial with magnitude $\frac{1}{2}(q-\frac{1}{2})$. The main ingredient of the approach we adopt, inspired by [Hal15c], is a change of the initial Brownian motion (W_t) to another Brownian motion $(\hat{W}_t)$ justified by Lévy's martingale characterization of Brownian motion, see Theorem A.3.9.

Section 3.3 is devoted to the logarithmic rate of convergence of the proposed semi-discrete scheme, while Section 3.4 concerns the proof of the polynomial rate of convergence. In Section 3.5 we briefly discuss the case where we do not alter the initial Brownian motion (W_t). This approach produces reduced convergence rate. Finally, Section 3.6 presents illustrative figures where the behavior of the proposed scheme, regarding the order of convergence, is shown and a comparison with BIM and BMM schemes is given. In Section 3.7 we treat the full model (3.1.1) for a special case. Concluding remarks are in Section 3.8 and in Appendix D we briefly present numerical schemes for the integration of the variance-volatility process (V_t).

[3] We give in Appendix D the form of all the above schemes for the approximation of (V_t).

3.2 The setting and the main results.

We consider the following SDE

$$x_t = x_0 + \int_0^t (k_1 - k_2 x_s)ds + \int_0^t k_3 (x_s)^q dW_s, \quad t \in [0,T], \tag{3.2.1}$$

where k_1, k_2, k_3 are positive and $1/2 < q < 1$. Then, Feller's test implies that there is a unique strong solution such that $x_t > 0$ a.s. when $x_0 > 0$ a.s. Let

$$f_\theta(x,y) = \underbrace{k_1 - k_2(1-\theta)x - \frac{(k_3)^2}{4(1+k_2\theta\Delta)}x^{2q-1} - k_2\theta y}_{f_1(x,y)} + \underbrace{\frac{(k_3)^2}{4(1+k_2\theta\Delta)}x^{2q-1}}_{f_2(x)} \tag{3.2.2}$$

and

$$g(x,y) = k_3 x^{q-\frac{1}{2}}\sqrt{y}, \tag{3.2.3}$$

where $f(x,x) = a(x) = k_1 - k_2 x$ and $g(x,x) = b(x) = k_3 x^q$.

Let the partition $0 = t_0 < t_1 < \ldots < t_N = T$ with $\Delta = T/N$ and consider the following process

$$y_t^{SD}(q) = y_{t_n} + f_1(y_{t_n}, y_t)\cdot\Delta + \int_{t_n}^t f_2(y_{t_n})ds + \int_{t_n}^t \text{sgn}(z_s) g(y_{t_n}, y_s) dW_s,$$

with $y_0 = x_0$ a.s. or more explicitly

$$\begin{aligned} y_t^{SD}(q) &= y_{t_n} + \left(k_1 - k_2(1-\theta)y_{t_n} - \frac{(k_3)^2}{4(1+k_2\theta\Delta)}(y_{t_n})^{2q-1} - k_2\theta y_t\right)\cdot\Delta \\ &\quad + \int_{t_n}^t \frac{(k_3)^2}{4(1+k_2\theta\Delta)}(y_{t_n})^{2q-1}ds + k_3(y_{t_n})^{q-\frac{1}{2}}\int_{t_n}^t \text{sgn}(z_s)\sqrt{y_s}dW_s, \end{aligned} \tag{3.2.4}$$

for $t \in (t_n, t_{n+1}]$, where $\theta \in [0,1]$ represents the level of implicitness and

$$z_t = \sqrt{y_n} + \frac{k_3}{2(1+k_2\theta\Delta)}(y_{t_n})^{q-\frac{1}{2}}(W_t - W_{t_n}), \tag{3.2.5}$$

with

$$y_n := y_{t_n}\left(1 - \frac{k_2\Delta}{1+k_2\theta\Delta}\right) + \frac{k_1\Delta}{1+k_2\theta\Delta} - \frac{(k_3)^2}{4(1+k_2\theta\Delta)^2}(y_{t_n})^{2q-1}\Delta. \tag{3.2.6}$$

Process (3.2.4) is well defined when $y_n \geq 0$ and this is true when

$$\frac{1}{(1+k_2\theta\Delta)}(k_3)^2 \leq 4(k_2 \wedge k_1) \text{ and } \Delta(2-\theta) \leq \frac{1}{k_2}.$$

Furthermore, (3.2.4) has jumps at nodes t_n. Solving for y_t, we end up with the following explicit scheme

$$\begin{aligned} y_t^{SD}(q) &= y_n + \int_{t_n}^t \frac{(k_3)^2}{4(1+k_2\theta\Delta)^2}(y_{t_n})^{2q-1}ds \\ &\quad + \frac{k_3}{1+k_2\theta\Delta}(y_{t_n})^{q-\frac{1}{2}} \int_{t_n}^t \operatorname{sgn}(z_s)\sqrt{y_s}dW_s, \end{aligned} \tag{3.2.7}$$

with solution in each step given by [KP95, (4.39), p.123]

$$y_t^{SD}(q) = (z_t)^2,$$

which has the pleasant feature $y_t^{SD}(q) \geq 0$.

Inspired by [Hal15c] we remove the term $\operatorname{sgn}(z_s)$ from (3.2.4) by considering the process

$$\widetilde{W}_t := \int_0^t \operatorname{sgn}(z_s)dW_s,$$

which is a martingale with quadratic variation $< \widetilde{W}_t, \widetilde{W}_t >= t$ and thus a standard Brownian motion w.r.t. its own filtration, justified by Lévy's theorem. Therefore, the compact form of (3.2.4) becomes

$$\begin{aligned} y_t^{SD} &= x_0 + \int_0^t (k_1 - k_2(1-\theta)y_{\hat{s}} - k_2\theta y_{\widetilde{s}})\, ds \\ &\quad + \int_t^{t_{n+1}} \left(k_1 - k_2(1-\theta)y_{t_n} - \frac{(k_3)^2}{4(1+k_2\theta\Delta)}(y_{t_n})^{2q-1} - k_2\theta y_t\right) ds \\ &\quad + k_3 \int_0^t (y_{\hat{s}})^{q-\frac{1}{2}}\sqrt{y_s}d\widetilde{W}_s, \end{aligned} \tag{3.2.8}$$

for $t \in (t_n, t_{n+1}]$ where $\hat{s} = t_j, s \in (t_j, t_{j+1}]$, $j = 0, \ldots, n$ and

$$\widetilde{s} = \begin{cases} t_{j+1}, & \text{for } s \in [t_j, t_{j+1}], \\ t, & \text{for } s \in [t_n, t] \end{cases} \quad j = 0, \ldots, n-1.$$

Consider also the process

$$\widetilde{x}_t = x_0 + \int_0^t (k_1 - k_2\widetilde{x}_s)ds + \int_0^t k_3(\widetilde{x}_s)^q d\widetilde{W}_s, \quad t \in [0, T]. \tag{3.2.9}$$

The process (x_t) of (3.2.1) and the process $(\widetilde{x}_t)$ of (3.2.9) have the same distribution. We show in the following that $\mathbb{E}\sup_{0\leq t\leq T}|y_t^{SD}(q)-\widetilde{x}_t|^2\to 0$ as $\Delta\downarrow 0$ thus the same holds for the unique solution of (3.2.1), i.e. $\mathbb{E}\sup_{0\leq t\leq T}|y_t^{SD}(q)-x_t|^2\to 0$ as $\Delta\downarrow 0$. To simplify notation we write $\widetilde{W},(\widetilde{x}_t)$ as $W,(x_t)$. We end up with the following explicit scheme

(3.2.10)
$$y_t^{SD}(q)=y_n+\int_{t_n}^{t}\frac{(k_3)^2}{4(1+k_2\theta\Delta)^2}(y_{t_n})^{2q-1}ds+\frac{k_3}{1+k_2\theta\Delta}(y_{t_n})^{q-\frac{1}{2}}\int_{t_n}^{t}\sqrt{y_s}dW_s,$$

where y_n is as in (3.2.6).

Assumption 3.2.1 *Let the parameters k_1,k_2,k_3 be positive such that*

$$\frac{1}{(1+k_2\theta\Delta)}(k_3)^2\leq 4(k_2\wedge k_1)$$

and consider $\Delta>0$ such that $\Delta(2-\theta)<\frac{1}{k_2}$, for $\theta\in[0,1]$. Moreover assume $x_0>0$ a.s. and $\mathbb{E}(x_0)^p<A$ for some $p\geq 4$. □

Theorem 3.2.2 *[Logarithmic rate of convergence] Let Assumption 3.2.1 hold. The semi-discrete scheme (3.2.10) converges to the true solution of (3.2.1) in the mean-square sense with rate given by*

$$\mathbb{E}\sup_{0\leq t\leq T}|y_t^{SD}(q)-x_t|^2\leq\frac{C}{\sqrt{\ln(\Delta)^{-1}}}, \tag{3.2.11}$$

where C is independent of Δ and given by

$$C:=32\sqrt{\frac{6}{\epsilon}}(k_3)^4T^2e^{6T^2(k_2)^2+k_2T},$$

where $0<\epsilon<q-\frac{1}{2}$. □

Assumption 3.2.3 *Let Assumption 3.2.1 hold where now $x_0\in\mathbb{R}$ and $x_0>0$.* □

Theorem 3.2.4 *[Polynomial rate of convergence] Let Assumption 3.2.3 hold. The semi-discrete scheme (3.2.10) converges to the true solution of (3.2.1) in the mean-square sense with rate given by,*

$$\mathbb{E}\sup_{0\leq t\leq T}|y_t^{SD}(q)-x_t|^2\leq C\Delta^{(q-\frac{1}{2})}, \tag{3.2.12}$$

where

$$C := 16(k_3)^2 T\sqrt{A_2(x_0+k_1T)^2}\sqrt{\hat{A}_{4q-2}}\left(2e^{6(k_2)^2T^2}+\frac{C_{HK}}{\epsilon-1}(x_0)^{(1-q)\nu(\lambda)}\right),$$

C_{HK} is a constant described in (3.4.7), λ is an appropriately chosen positive parameter which satisfies (3.4.8) and always exist, $\nu(\lambda) := \frac{\lambda}{2(1-q)^2(k_3)^2} - 1$ and $\epsilon > 1$. □

In the following sections we write for simplicity y_t^{SD} or y_t for $y_t^{SD}(q)$.

3.3 Logarithmic Rate of Convergence.

3.3.1 Moment Bounds.

Lemma 3.3.5 *[Moment bound for SD approximation] It holds that*

$$\mathbb{E}\sup_{0\le t\le T}(y_t)^p \le A_p\mathbb{E}(x_0+k_1T)^p,$$

for any $p > 2$, where $A_p := \exp\left\{\frac{p(p-1)}{2}(k_3)^2\left(\frac{p-1}{2p}+\frac{2^{p-1}}{p}\right)T\right\}$. □

Proof of Lemma 3.3.5. We first observe that (y_t) is bounded in the following way

$$\begin{aligned} 0 \le y_t &\le x_0 + \int_0^t k_1 ds + k_3\int_0^t (y_{\hat{s}})^{q-\frac{1}{2}}\sqrt{y_s}dW_s \\ &\le x_0 + k_1T + k_3\int_0^t (y_{\hat{s}})^{q-\frac{1}{2}}\sqrt{y_s}dW_s := u_t \end{aligned}$$

a.s., where the lower bound comes from the construction of (y_t) and the upper bound follows from a comparison theorem. We will bound (u_t) and therefore (y_t), since $0 \le y_t \le u_t$ a.s. Set the stopping time $\tau_R := \inf\{t \in [0,T] : u_t > R\}$, for $R > 0$ with the convention $\inf\emptyset = \infty$. Application of Itô's formula on $(u_{t\wedge\tau_R})^p$ implies

$$\begin{aligned}
&(u_{t\wedge\tau_R})^p = (x_0+k_1T)^p + \frac{p(p-1)}{2}(k_3)^2\int_0^{t\wedge\tau_R}(u_s)^{p-2}(y_{\hat{s}})^{2q-1}y_s ds \\
&+pk_3\int_0^{t\wedge\tau_R}(u_s)^{p-1}(y_{\hat{s}})^{q-\frac{1}{2}}\sqrt{y_s}dW_s \\
\leq\ & (x_0+k_1T)^p + \frac{p(p-1)}{2}(k_3)^2\int_0^{t\wedge\tau_R}(u_s)^{p-1}(y_{\hat{s}})^{2q-1}ds + M_t \\
\leq\ & (x_0+k_1T)^p + \frac{p(p-1)}{2}(k_3)^2\int_0^{t\wedge\tau_R}(\frac{p-1}{2p}(u_s)^p + \frac{2^{p-1}}{p}(y_{\hat{s}})^{(2q-1)p})ds + M_t \\
\leq\ & (x_0+k_1T)^p + \frac{p(p-1)}{2}(k_3)^2\left(\frac{p-1}{2p}+\frac{2^{p-1}}{p}\right)\int_0^{t\wedge\tau_R}(u_s)^p ds + M_t,
\end{aligned}$$

where in the second step we have used that $0 \leq y_t \leq u_t$, in the third step the inequality

$$x^{p-1}y \leq \epsilon\frac{p-1}{p}x^p + \frac{1}{p\epsilon^{p-1}}y^p,$$

valid for $x \wedge y \geq 0$ and $p > 1$ with $\epsilon = \frac{1}{2}$, in the final step the fact $\frac{1}{2} < q < 1$ and

$$M_t := pk_3\int_0^{t\wedge\tau_R}(u_s)^{p-1}(y_{\hat{s}})^{q-\frac{1}{2}}\sqrt{y_s}dW_s.$$

Taking expectations in the above inequality and using that M_t is a local martingale vanishing at 0, we get

$$\begin{aligned}
\mathbb{E}(u_{t\wedge\tau_R})^p &\leq \mathbb{E}(x_0+k_1T)^p + \frac{p(p-1)}{2}(k_3)^2\left(\frac{p-1}{2p}+\frac{2^{p-1}}{p}\right)\int_0^t \mathbb{E}(u_{s\wedge\tau_R})^p ds \\
&\leq \mathbb{E}(x_0+k_1T)^p \exp\left\{\frac{p(p-1)}{2}(k_3)^2\left(\frac{p-1}{2p}+\frac{2^{p-1}}{p}\right)T\right\} \\
&\leq A_p\mathbb{E}(x_0+k_1T)^p,
\end{aligned}$$

where we have applied the Gronwall inequality (B.3.6).We have that

$$(y_{t\wedge\tau_R})^p = (y_{\tau_R})^p\mathbb{I}_{\{\tau_R\leq t\}} + (y_t)^p\mathbb{I}_{\{t<\tau_R\}} \geq (y_t)^p\mathbb{I}_{\{t<\tau_R\}},$$

thus taking expectations in the above inequality and using the estimated upper bound for $\mathbb{E}(u_{t\wedge\tau_R})^p$ we arrive at

$$\mathbb{E}(y_t)^p\mathbb{I}_{\{t<\tau_R\}} \leq \mathbb{E}(y_{t\wedge\tau_R})^p \leq \mathbb{E}(u_{t\wedge\tau_R})^p \leq A_p\mathbb{E}(x_0+k_1T)^p,$$

and taking the limit as $R \to \infty$, we get

$$\lim_{R\to\infty} \mathbb{E}(y_t)^p \mathbb{I}_{\{t<\tau_R\}} \leq A_p \mathbb{E}(x_0 + k_1 T)^p.$$

Let us fix t. The sequence of stopping times τ_R is increasing in R and $t \wedge \tau_R \to t$ as $R \to \infty$, thus the sequence $(y_t)^p \mathbb{I}_{\{t<\tau_R\}}$ is non-decreasing in R and $(y_t)^p \mathbb{I}_{\{t<\tau_R\}} \to (y_t)^p$ as $R \to \infty$. Application of the monotone convergence theorem, see Theorem B.1.1 implies

$$\mathbb{E}(y_t)^p \leq A_p \mathbb{E}(x_0 + k_1 T)^p,$$

for any $p > 2$. Using again Itô's formula on $(u_t)^p$, taking the supremum and then using Doob's martingale inequality on the diffusion term we bound $\mathbb{E} \sup_{0\leq t\leq T}(u_t)^p$ and thus $\mathbb{E} \sup_{0\leq t\leq T}(y_t)^p$. □

Lemma 3.3.6 *[Error bound for SD scheme] Let n_s be an integer such that $s \in [t_{n_s}, t_{n_s+1}]$. Then*

$$\mathbb{E}|y_s - y_{\hat{s}}|^p \leq \hat{A}_p \Delta^{p/2}, \qquad \mathbb{E}|y_s - y_{\tilde{s}}|^p < \widetilde{A}_p \Delta^{p/2},$$

for any $p > 0$, where the positive quantities $\hat{A}_p, \widetilde{A}_p$ do not depend on Δ. □

Proof of Lemma 3.3.6. First we take a $p \geq 2$. It holds that

$$
\begin{aligned}
|y_s - y_{\hat{s}}|^p &= \Big| \int_{t_{n_s}}^{s} \big(k_1 - k_2(1-\theta)y_{\hat{u}} - k_2\theta y_{\tilde{u}}\big) du + \int_{t_{n_s}}^{t_{n_s+1}} k_2\theta y_{\hat{s}} du \\
&- \int_{s}^{t_{n_s+1}} k_2\theta y_s du + \int_{s}^{t_{n_s}} \left(k_1 - k_2(1-\theta)y_{t_{n_s}} - \frac{(k_3)^2(y_{t_{n_s}})^{2q-1}}{4(1+k_2\theta\Delta)}\right) du \\
&+ k_3 \int_{t_{n_s}}^{s} (y_{\hat{u}})^{q-\frac{1}{2}} \sqrt{y_u} dW_u \Big|^p \\
\leq\ & 5^{p-1}\Big(\Big| \int_{t_{n_s}}^{s} (k_1 - k_2(1-\theta)y_{\hat{u}} - k_2\theta y_{\tilde{u}})\, du \Big|^p + (k_2)^p\theta^p(y_{\hat{s}})^p(t_{n_s+1} - t_{n_s})^p \\
&+ (k_2)^p\theta^p(y_s)^p(t_{n_s+1} - s)^p + \left| \int_{s}^{t_{n_s}} (k_1 - k_2(1-\theta)y_{t_{n_s}} - \frac{(k_3)^2(y_{t_{n_s}})^{2q-1}}{4(1+k_2\theta\Delta)}) du \right|^p \\
&+ (k_3)^p \Big| \int_{t_{n_s}}^{s} (y_{\hat{u}})^{q-\frac{1}{2}} \sqrt{y_u} dW_u \Big|^p \Big) \\
\leq\ & 5^{p-1}\Big(|t_{n_s} - s|^{p-1} \int_{t_{n_s}}^{s} |k_1 - k_2(1-\theta)y_{\hat{u}} - k_2\theta y_{\tilde{u}}|^p\, du \\
&+ (k_2)^p\theta^p\left((y_{\hat{s}})^p + (y_s)^p\right)\Delta^p + \left| k_1 - k_2(1-\theta)y_{t_{n_s}} - \frac{(k_3)^2(y_{t_{n_s}})^{2q-1}}{4(1+k_2\theta\Delta)} \right|^p \Delta^p \\
&+ (k_3)^p \Big| \int_{t_{n_s}}^{s} (y_{\hat{u}})^{q-\frac{1}{2}} \sqrt{y_u} dW_u \Big|^p \Big),
\end{aligned}
$$

where we have used the Cauchy-Schwarz inequality. Taking expectations in the above inequality and using Lemma 3.3.5 and the BDG inequality (B.3.5) on the diffusion term we conclude

$$
\mathbb{E}|y_s - y_{\hat{s}}|^p \leq \hat{A}_p \Delta^{p/2}, \tag{3.3.1}
$$

where the positive quantity $\hat{A}_p$ except on p, depends also on the parameters k_1, k_2, k_3, θ, q, but not on Δ. Now, for $0 < p < 2$ we get

$$
\mathbb{E}|y_s - y_{\hat{s}}|^p \leq \left(\mathbb{E}|y_s - y_{\hat{s}}|^2\right)^{p/2} \leq \hat{A}_p \Delta^{p/2},
$$

where we have used Jensen's inequality for the concave function $\phi(x) = x^{p/2}$. Following the same lines, we can show that

$$
\mathbb{E}|y_s - y_{\tilde{s}}|^p \leq \widetilde{A}_p \Delta^{p/2},
$$

for any $0 < p$, where the positive quantity $\widetilde{A}_p$ except on p, depends also on the parameters k_1, k_2, k_3, θ, q, but not on Δ. $\square$

For the rest of this section we rewrite again the compact form of (3.2.8) in the following way

$$y_t^{SD} = \underbrace{x_0 + \int_0^t f_\theta(y_{\hat{s}}, y_{\tilde{s}})ds + \int_0^t g(y_{\hat{s}}, y_s)dW_s}_{h_t} + \int_t^{t_{n+1}} f_1(y_{t_n}, y_t)ds, \tag{3.3.2}$$

where $f_\theta(\cdot,\cdot)$ is given by (3.2.2) and the auxiliary process (h_t) is close to (y_t) as shown in the next result.

Lemma 3.3.7 *[Moment bounds involving the auxiliary process] For any $s \in [0,T]$ it holds that*

$$\mathbb{E}|h_s - y_s|^p \leq C_p\Delta^p, \quad \mathbb{E}|h_s|^p \leq C_h, \tag{3.3.3}$$

and for $s \in [t_n, t_{n+1}]$ we have that

$$\mathbb{E}|h_s - y_{\hat{s}}|^p \leq \hat{C}_p\Delta^{p/2}, \qquad \mathbb{E}|h_s - y_{\tilde{s}}|^p \leq \widetilde{C}_p\Delta^{p/2},$$

for any $p > 0$, where the positive quantities $C_p, \hat{C}_p, \widetilde{C}_p, C_h$ do not depend on Δ. □

Proof of Lemma 3.3.7. We have that

$$|h_s - y_s|^p = \left|\int_s^{t_{n+1}} f_1(y_{t_n}, y_s)du\right|^p \leq |t_{n+1} - s|^p |f_1(y_{t_n}, y_s)|^p,$$

for any $p > 0$, where we have used (3.3.2). Using Lemma 3.3.5 we get the left part of (3.3.3). Now for $p \geq 2$ and noting that

$$\begin{aligned} \mathbb{E}|h_s|^p &\leq 2^{p-1}\mathbb{E}|h_s - y_s|^p + 2^{p-1}\mathbb{E}|y_s|^p \\ &\leq 2^{p-1}C_p\Delta^p + 2^{p-1}A_p\mathbb{E}(x_0 + k_1T)^p \leq C_h, \end{aligned}$$

we get the right part of (3.3.3), where we have used Lemma 3.3.5. The case $0 < p < 2$ follows by Jensen's inequality as in Lemma 3.3.6.

Furthermore, for $s \in [t_n, t_{n+1}]$ and $p > 2$ we derive that

$$\begin{aligned} \mathbb{E}|h_s - y_{\hat{s}}|^p &\leq 2^{p-1}\mathbb{E}|h_s - y_s|^p + 2^{p-1}\mathbb{E}|y_s - y_{\hat{s}}|^p \\ &\leq 2^{p-1}C_p\Delta^p + 2^{p-1}\hat{A}_p\Delta^{p/2} \leq \hat{C}_p\Delta^{p/2} \end{aligned}$$

where we have used (3.3.1) and in the same manner

$$\mathbb{E}|h_s - y_{\tilde{s}}|^p \leq 2^{p-1}C_p\Delta^p + 2^{p-1}\widetilde{A}_p\Delta^{p/2} \leq \widetilde{C}_p\Delta^{p/2}.$$

The case $0 < p < 2$ follows by Jensen's inequality. □

3.3.2 Convergence of the auxiliary process (h_t) to (x_t) in $\mathcal{L}^1$.

We will use the representation (3.3.2) and write

$$h_t - x_t = \int_0^t \left(f_\theta(y_{\hat{s}}, y_{\tilde{s}}) - f_\theta(x_s, x_s)\right) ds + \int_0^t \left(g(y_{\hat{s}}, y_s) - g(x_s, x_s)\right) dW_s.$$

Proposition 3.3.8 *Let Assumption 3.2.1 hold. Then we have*

$$\sup_{0\le t\le T} \mathbb{E}|h_t - x_t| \le \left(J_3 \frac{\Delta^{q-\frac{1}{2}}}{me_m} + 2(k_3)^2 T \frac{1}{m} \right) e^{k_2 T}, \tag{3.3.4}$$

for any $m > 1$, where $e_m = e^{-m(m+1)/2}$ and

$$J_3 := 2(k_3)^2 T \sqrt{A_2 \mathbb{E}(x_0 + k_1 T)^2} \sqrt{\hat{A}_{4q-2}}.$$

□

Proof of Proposition 3.3.8. We use the method of Yamada and Watanabe [YW71] as in the proof of Proposition 2.3.4. We have that

$$\mathbb{E}|h_t - x_t| \le e_{m-1} + \mathbb{E}\phi_m(h_t - x_t), \tag{3.3.5}$$

where ϕ_m is the sequence of approximations of $|x|$. Moreover we find that

$$\begin{aligned} f_\theta(y_{\hat{s}}, y_{\tilde{s}}) - f_\theta(x_s, x_s) &= (k_1 - k_2(1-\theta)y_{\hat{s}} - k_2\theta y_{\tilde{s}}) - (k_1 - k_2 x_s) \\ &= -k_2(1-\theta)(y_{\hat{s}} - x_s) - k_2\theta(y_{\tilde{s}} - x_s) \\ &= k_2(1-\theta)(h_s - y_{\hat{s}}) + k_2\theta(h_s - y_{\tilde{s}}) - k_2(h_s - x_s) \end{aligned} \tag{3.3.6}$$

and

$$\begin{aligned} |g(y_{\hat{s}}, y_s) - g(x_s, x_s)|^2 &= |k_3(y_{\hat{s}})^{q-\frac{1}{2}}\sqrt{y_s} - k_3(x_s)^q|^2 \\ &\le (k_3)^2 \left(\sqrt{y_s}\left((y_{\hat{s}})^{q-\frac{1}{2}} - (y_s)^{q-\frac{1}{2}}\right) + ((y_s)^q - (x_s)^q) \right)^2 \\ &\le 2(k_3)^2 \left(y_s \left((y_{\hat{s}})^{q-\frac{1}{2}} - (y_s)^{q-\frac{1}{2}}\right)^2 + ((y_s)^q - (x_s)^q)^2 \right) \\ &\le 2(k_3)^2 \left(y_s|y_{\hat{s}} - y_s|^{2q-1} + (\sqrt{|y_s - x_s|})^2 \right) \\ &\le 2(k_3)^2 \left(y_s|y_{\hat{s}} - y_s|^{2q-1} + |h_s - y_s| + |h_s - x_s| \right), \end{aligned} \tag{3.3.7}$$

where we have used properties of Hölder continuous functions and namely the fact that x^q is q-Hölder continuous for $q \leq 1$, i.e. $|x^q - y^q| \leq |x-y|^q$, and that x^q is 1/2-Hölder continuous since $q > 1/2$. Application of Itô's formula to the sequence $\{\phi_m\}_{m\in\mathbb{N}}$, implies

$$\begin{aligned}
\phi_m(h_t - x_t) &= \int_0^t \phi'_m(h_s - x_s)(f_\theta(y_{\hat{s}}, y_{\tilde{s}}) - f_\theta(x_s, x_s))ds + M_t \\
&+\frac{1}{2}\int_0^t \phi''_m(h_s - x_s)(g(y_{\hat{s}}, y_{\tilde{s}}) - g(x_s, x_s))^2 ds \\
\leq & \int_0^t \left(k_2(1-\theta)|h_s - y_{\hat{s}}| + k_2\theta|h_s - y_{\tilde{s}}| + k_2|h_s - x_s|\right) ds + M_t \\
&+\int_0^t \frac{2(k_3)^2}{m|h_s - x_s|}\left(y_s|y_{\hat{s}} - y_s|^{2q-1} + |h_s - y_s| + |h_s - x_s|\right) ds \\
\leq & \; k_2(1-\theta)\int_0^t |h_s - y_{\hat{s}}|ds + k_2\theta\int_0^t |h_s - y_{\tilde{s}}|ds + \frac{2(k_3)^2}{me_m}\int_0^t |h_s - y_s|ds \\
&+k_2\int_0^t |h_s - x_s|ds + M_t + \frac{2(k_3)^2}{me_m}\int_0^t y_s|y_{\hat{s}} - y_s|^{2q-1}ds + \frac{2(k_3)^2 T}{m},
\end{aligned}$$

where in the second step we have used (3.3.6) and (3.3.7) and the properties of ϕ_m and

$$M_t := \int_0^t \phi'_m(h_u - x_u)(g(y_{\hat{u}}, y_{\tilde{u}}) - g(x_u, x_u))dW_u.$$

Taking expectations in the above inequality yields

$$
\begin{aligned}
& \mathbb{E}\phi_m(h_t - x_t) \leq k_2(1-\theta)\int_0^t \mathbb{E}|h_s - y_{\hat{s}}|ds + k_2\theta\int_0^t \mathbb{E}|h_s - y_{\widetilde{s}}|ds \\
& +\frac{2(k_3)^2}{me_m}\int_0^t \mathbb{E}|h_s - y_s|ds + \frac{2(k_3)^2}{me_m}\int_0^t \mathbb{E}y_s|y_{\hat{s}} - y_s|^{2q-1}ds + \frac{2(k_3)^2T}{m} \\
& +k_2\int_0^t \mathbb{E}|h_s - x_s|ds \\
\leq\ & k_2(1-\theta)T\hat{C}_1\sqrt{\Delta} + k_2\theta T\widetilde{C}_1\sqrt{\Delta} + \frac{2(k_3)^2TC_1}{me_m}\Delta + k_2\int_0^t \mathbb{E}|h_s - x_s|ds \\
& +\frac{2(k_3)^2}{me_m}\int_0^t \sqrt{\mathbb{E}(y_s)^2}\sqrt{\mathbb{E}|y_{\hat{s}} - y_s|^{4q-2}}ds + \frac{2(k_3)^2T}{m} \\
\leq\ & k_2T((1-\theta)\hat{C}_1 + \theta\widetilde{C}_1)\sqrt{\Delta} + \frac{2(k_3)^2TC_1}{me_m}\Delta + k_2\int_0^t \mathbb{E}|h_s - x_s|ds \\
& +\frac{2(k_3)^2T}{me_m}\sqrt{A_2\mathbb{E}(x_0 + k_1T)^2}\sqrt{\hat{A}_{4q-2}}\Delta^{q-\frac{1}{2}} + \frac{2(k_3)^2T}{m},
\end{aligned}
$$

where we have used Lemma 3.3.7 in the second step and Hölder's inequality, Lemmata 3.3.5 and 3.3.6 in the third step and the fact that $\mathbb{E}M_t = 0$ (The function $d(u) = \phi'_m(h_u - x_u)(g(y_{\hat{u}}, y_{\widetilde{u}}) - g(x_u, x_u))$ belongs to the space $\mathcal{M}^2([0,t];\mathbb{R})$ of real-valued measurable $\mathcal{F}_t$-adapted processes such that $\mathbb{E}\int_0^t |d(u)|^2du < \infty$; thus [Mao97, Th. 1.5.8] implies $\mathbb{E}M_t = 0$.) Thus (3.3.5) becomes

$$
\begin{aligned}
\mathbb{E}|h_t - x_t| &\leq e_{m-1} + J_1\sqrt{\Delta} + 2(k_3)^2TC_1\frac{\Delta}{me_m} + J_3\frac{\Delta^{q-\frac{1}{2}}}{me_m} + 2(k_3)^2T\frac{1}{m} \\
&\quad +k_2\int_0^t \mathbb{E}|h_s - x_s|ds \\
&\leq J_3\frac{\Delta^{q-\frac{1}{2}}}{me_m} + 2(k_3)^2T\frac{1}{m} + k_2\int_0^t \mathbb{E}|h_s - x_s|ds \\
&\leq \left(J_3\frac{\Delta^{q-\frac{1}{2}}}{me_m} + 2(k_3)^2T\frac{1}{m}\right)e^{k_2t},
\end{aligned}
$$

where in the second step we have used the asymptotic relations, $\Delta^\kappa = o(\Delta^{q-\frac{1}{2}})$ as $\Delta \downarrow 0$ for any $\kappa \geq 1/2$, $e_{m-1} = o(\frac{1}{m})$ as $m \to \infty$, $\sqrt{\Delta} = o(\frac{\Delta^\kappa}{me_m})$ for any $\kappa \leq 1$ as $m \to \infty$, in the last step we have used the Gronwall inequality and J_3 is as defined in Proposition 3.3.8 while

$$J_1 := k_2T((1-\theta)\hat{C}_1 + \theta\widetilde{C}_1).$$

Taking the supremum over all $0 \le t \le T$ gives (3.3.4). $\square$

3.3.3 Convergence of the auxiliary process (h_t) to (x_t) in $\mathcal{L}^2$.

Proposition 3.3.9 *Let Assumption 3.2.1 hold. Then we have*

$$\mathbb{E}\sup_{0\le t\le T}|h_t - x_t|^2 \le \frac{C_\epsilon}{\sqrt{\ln(\Delta)^{-1}}}, \tag{3.3.8}$$

where C_ϵ is independent of Δ and given by $C_\epsilon := 32\sqrt{\frac{3}{2\epsilon}}(k_3)^4T^2e^{6T^2(k_2)^2+k_2T}$, where $0<\epsilon<q-\frac{1}{2}$. $\square$

Proof of Proposition 3.3.9. We estimate the difference $|\mathcal{E}_t|^2 := |h_t - x_t|^2$. It holds that

$$\begin{aligned}
|\mathcal{E}_t|^2 &= \Big|\int_0^t (f_\theta(y_{\hat{s}}, y_{\widetilde{s}}) - f_\theta(x_s, x_s))\,ds + \int_0^t (g(y_{\hat{s}}, y_s) - g(x_s, x_s))\,dW_s\Big|^2 \\
&\le 2T\int_0^t (k_2(1-\theta)|h_s - y_{\hat{s}}| + k_2\theta|h_s - y_{\widetilde{s}}| + k_2|\mathcal{E}_s|)^2\,ds + 2|M_t|^2 \\
&\le 6T(k_2)^2(1-\theta)^2\int_0^t |h_s - y_{\hat{s}}|^2 ds + 6T(k_2)^2\theta^2\int_0^t |h_s - y_{\widetilde{s}}|^2 ds \\
&\quad +6T(k_2)^2\int_0^t |\mathcal{E}_s|^2 ds + 2|M_t|^2,
\end{aligned}$$

where in the second step we have used the Cauchy-Schwarz inequality and (3.3.6) and

$$M_t := \int_0^t (g(y_{\hat{u}}, y_{\widetilde{u}}) - g(x_u, x_u))dW_u.$$

Taking the supremum over all $t \in [0,T]$ and then expectations we have

$$
\begin{aligned}
\mathbb{E}\sup_{0\le t\le T}|\mathcal{E}_t|^2 &\le 6T(k_2)^2(1-\theta)^2\int_0^T \mathbb{E}|h_s-y_{\hat{s}}|^2ds + 6T(k_2)^2\theta^2\int_0^T \mathbb{E}|h_s-y_{\tilde{s}}|^2ds \\
&\quad +6T(k_2)^2\int_0^T \mathbb{E}\sup_{0\le l\le s}|\mathcal{E}_l|^2ds + 2\mathbb{E}\sup_{0\le t\le T}|M_t|^2 \\
&\le 6T^2(k_2)^2((1-\theta)^2\hat{A}_2+\theta^2\widetilde{A}_2)\Delta + 6T(k_2)^2\int_0^T \mathbb{E}\sup_{0\le l\le s}|\mathcal{E}_l|^2ds \\
&\quad +8\mathbb{E}|M_T|^2, \qquad (3.3.9)
\end{aligned}
$$

where in the second step we have used Lemma 3.3.6 and Doob's martingale inequality with $p=2$, since M_t is an $\mathbb{R}$-valued martingale that belongs to $\mathcal{L}^2$. We find that

$$
\begin{aligned}
\mathbb{E}|M_T|^2 &= \mathbb{E}\left|\int_0^T |g(y_{\hat{s}},y_s)-g(x_s,x_s)|dW_s\right|^2 = \mathbb{E}\int_0^T |g(y_{\hat{s}},y_s)-g(x_s,x_s)|^2ds \\
&\le 2(k_3)^2\mathbb{E}\left(\int_0^T \left(y_s|y_{\hat{s}}-y_s|^{2q-1}+|h_s-y_s|+|h_s-x_s|\right)ds\right) \\
&\le 2(k_3)^2\int_0^T \mathbb{E}\left(y_s|y_{\hat{s}}-y_s|^{2q-1}\right)ds + 2(k_3)^2\int_0^T \mathbb{E}|h_s-y_s|ds \\
&\quad +2(k_3)^2\int_0^T \mathbb{E}|\mathcal{E}_s|ds,
\end{aligned}
$$

where we have used (3.3.7). Now, Lemmata 3.3.5, 3.3.6 and 3.3.7 imply

$$
\begin{aligned}
\mathbb{E}|M_T|^2 &\le J_6\sqrt{\Delta^{2q-1}} + 2(k_3)^2TC_1\Delta + 2(k_3)^2\int_0^T \mathbb{E}|\mathcal{E}_s|ds \\
&\le J_6\Delta^{q-\frac{1}{2}} + 2(k_3)^2\int_0^T \mathbb{E}|\mathcal{E}_s|ds,
\end{aligned}
$$

where we have used the asymptotic relations, $\Delta^l = o(\Delta^{q-\frac{1}{2}})$ for all $l\ge\frac{1}{2}$ as $\Delta\downarrow 0$ and the quantity J_6 is given by $J_6 := 2(k_3)^2T\sqrt{A_2\mathbb{E}(x_0+k_1T)^2}\sqrt{\hat{A}_{4q-2}}$.

Relation (3.3.9) becomes

$$\begin{aligned} & \mathbb{E}\sup_{0\le t\le T}|\mathcal{E}_t|^2 \le 8J_6\Delta^{q-\frac{1}{2}} + J_5\Delta + 6T(k_2)^2\int_0^T \mathbb{E}\sup_{0\le l\le s}|\mathcal{E}_l|^2 ds \\ & +16(k_3)^2\int_0^T \mathbb{E}\sup_{0\le l\le s}|\mathcal{E}_l| ds \\ \le\; & 8J_6\Delta^{q-\frac{1}{2}} + 16(k_3)^2 T\left(J_3\frac{\Delta^{q-\frac{1}{2}}}{me_m} + 3(k_3)^2 T\frac{1}{m}\right)e^{k_2 T} \\ & +6T(k_2)^2\int_0^T \mathbb{E}\sup_{0\le l\le s}|\mathcal{E}_l|^2 ds \\ \le\; & 16(k_3)^2 T J_3 e^{k_2T+6T^2(k_2)^2}\frac{\Delta^{q-\frac{1}{2}}}{me_m} + 32(k_3)^4T^2e^{k_2T+6T^2(k_2)^2}\frac{1}{m}, \end{aligned}$$

where we have used Proposition 3.3.8 in the second step with the sequence e_m as defined there, the Gronwall inequality in the last step and the asymptotic relation $\Delta^\kappa = o(\frac{\Delta^\kappa}{me_m})$ as $m \to \infty$, for any $\kappa > 0$ and J_5 is independent of Δ and given by $J_5 := 6T^2(k_2)^2((1-\theta)^2\hat{A}_2 + \theta^2\widetilde{A}_2)$.

We take $m = \sqrt{\ln\Delta^{-\lambda}}$, with $\lambda > 0$ to be specified soon and note that $e^{\sqrt{\ln\Delta^{-\lambda}}} = o(\Delta^{-\lambda})$ as $\Delta \downarrow 0$, since $e^{\sqrt{\ln n}} = o(n)$ as $n \to \infty$. Moreover we have that

$$\frac{\Delta^{q-\frac{1}{2}}}{e_m} = \frac{\Delta^{q-\frac{1}{2}}}{e^{-\frac{m^2}{2}}}e^{\frac{m}{2}} = \frac{\Delta^{q-\frac{1}{2}}}{e^{-\frac{\ln\Delta^{-\lambda}}{2}}}e^{\frac{1}{2}\sqrt{\ln\Delta^{-\lambda}}} = \Delta^{q-\frac{1}{2}-\frac{3\lambda}{2}}\frac{e^{\frac{1}{2}\sqrt{\ln\Delta^{-\lambda}}}}{\Delta^{-\lambda}}.$$

Now, since $q > \frac{1}{2}$ there is an $\epsilon > 0$ small enough such that $q - \frac{1}{2} - \epsilon > 0$. We take $\lambda = \frac{2\epsilon}{3}$ and conclude that

$$\frac{\Delta^{q-\frac{1}{2}}}{e_m} = \Delta^{q-\frac{1}{2}-\epsilon}\frac{e^{\frac{1}{2}\sqrt{\ln\Delta^{-\frac{2\epsilon}{3}}}}}{\Delta^{-\frac{2\epsilon}{3}}} \to 0,$$

as $\Delta \to 0$ which in turn implies the asymptotic relation $\frac{\Delta^{q-\frac{1}{2}}}{me_m} = o(\frac{1}{m})$ as $m \to \infty$, with the logarithmic rate stated before. We finally arrive at

$$\mathbb{E}\sup_{0\le t\le T}|\mathcal{E}_t|^2 \le 32(k_3)^4T^2e^{k_2T+6T^2(k_2)^2}\frac{1}{\sqrt{\ln\Delta^{-\frac{2\epsilon}{3}}}},$$

by taking $0 < \epsilon < q - \frac{1}{2}$, which implies (3.3.8). □

3.3.4 Proof of Theorem 3.2.2.

In order to finish the proof of Theorem 3.2.2 we just use the triangle inequality, Lemma 3.3.7 and Proposition 3.3.9 to get

$$\begin{aligned}\mathbb{E}\sup_{0\le t\le T}|y_t-x_t|^2 &\le 2\mathbb{E}\sup_{0\le t\le T}|h_t-y_t|^2+2\mathbb{E}\sup_{0\le t\le T}|\mathcal{E}_t|^2\\ &\le 2C_2\Delta^2+2\frac{C_\epsilon}{\sqrt{\ln\Delta^{-1}}}\le\frac{C}{\sqrt{\ln\Delta^{-1}}},\end{aligned}$$

where $C=C(k_2,k_3,\epsilon,T)$, is given in the statement of Theorem 3.2.2.

3.4 Polynomial Rate of Convergence.

We work with the stochastic time change inspired by [Ber04]. We define the process

$$\gamma(t):=\int_0^t\frac{128(k_3)^2q^2}{\left[(y_s)^{1-q}+(x_s)^{1-q}\right]^2}ds$$

and the stopping time

$$\tau_l:=\inf\{s\in[0,T]:6T(k_2)^2s+\gamma(s)\ge l\}.$$

The process $\gamma(t)$ is well defined since $x_t>0$ a.s. and $y_t\ge 0$ (see Section 3.2).

The difference $|\mathcal{E}_t|^2:=|h_t-x_t|^2$ is estimated as in Section 3.3 and we get, as in (3.3.9), that

$$\mathbb{E}\sup_{0\le t\le\tau}|\mathcal{E}_t|^2\le J_5\Delta+6T(k_2)^2\int_0^\tau\mathbb{E}\sup_{0\le l\le s}|\mathcal{E}_l|^2ds+8\mathbb{E}|M_\tau|^2, \tag{3.4.1}$$

where τ a stopping time and J_5 independent of Δ is as in proof of Proposition 3.3.9. The main difference here will be the estimation of the last term in (3.4.1). The approach in Section 3.3 resulted in the $\mathcal{L}^1$ estimation $\mathbb{E}|\mathcal{E}_t|$ where we used the Yamada-Watanabe approach. Now, we use the Berkaoui approach. Relation (3.3.7) becomes

$$\begin{aligned}|g(y_{\hat{s}},y_s)-g(x_s,x_s)|^2&\le 2(k_3)^2\left(y_s|y_{\hat{s}}-y_s|^{2q-1}+|(y_s)^q-(x_s)^q|^2\right)\\ &\le 2(k_3)^2\left(y_s|y_{\hat{s}}-y_s|^{2q-1}\right)+|(y_s)^q-(x_s)^q|^2\left((y_s)^{1-q}+(x_s)^{1-q}\right)^2\frac{(\gamma_s)'}{64q^2}\\ &\le 2(k_3)^2\left(y_s|y_{\hat{s}}-y_s|^{2q-1}\right)+\frac{1}{8}\left(|h_s-y_s|^2+|\mathcal{E}_s|^2\right)(\gamma_s)',\end{aligned}$$

where we have used the inequality

$$|a^q - b^q|(a^{1-q} + b^{1-q}) \leq 2q|a-b|,$$

valid for all $a \geq 0, b \geq 0$ and $\frac{1}{2} \leq q \leq 1$. Consequently, we get the upper bound

$$\begin{aligned}
\mathbb{E}|M_\tau|^2 &:= \mathbb{E}\left|\int_0^\tau |g(y_{\hat{s}}, y_s) - g(x_s, x_s)| dW_s\right|^2 \\
&\leq J_6 \Delta^{q-\frac{1}{2}} + \frac{1}{8}\int_0^\tau \mathbb{E}|h_s - y_s|^2 (\gamma_s)' ds + \frac{1}{8}\int_0^\tau \mathbb{E}|\mathcal{E}_s|^2 (\gamma_s)' ds \\
&\leq J_6 \Delta^{q-\frac{1}{2}} + \frac{1}{8}\int_0^\tau \sqrt{\mathbb{E}|h_s - y_s|^4}\sqrt{\mathbb{E}((\gamma_s)')^2} ds + \frac{1}{8}\int_0^\tau \mathbb{E}|\mathcal{E}_s|^2 (\gamma_s)' ds,
\end{aligned}$$

where we used Hölder's inequality; J_6 independent of Δ is as in the proof of Proposition 3.3.9. Relation (3.4.1) becomes

$$\begin{aligned}
& \mathbb{E}\sup_{0\leq t\leq\tau} |\mathcal{E}_t|^2 \leq 8J_6\Delta^{q-\frac{1}{2}} + 6T(k_2)^2 \int_0^\tau \mathbb{E}\sup_{0\leq l\leq s}|\mathcal{E}_l|^2 ds \\
& + \int_0^\tau \sqrt{\mathbb{E}|h_s - y_s|^4}\sqrt{\mathbb{E}((\gamma_s)')^2} ds + \int_0^\tau \mathbb{E}|\mathcal{E}_s|^2 (\gamma_s)' ds \\
\leq\ & 8J_6\Delta^{q-\frac{1}{2}} + \sqrt{C_4}\Delta^2 \int_0^\tau \sqrt{\mathbb{E}\left(\frac{128(k_3)^2 q^2}{[(y_s)^{1-q} + (x_s)^{1-q}]^2}\right)^2} ds \\
& + \int_0^\tau \mathbb{E}\sup_{0\leq l\leq s}|\mathcal{E}_l|^2 (6T(k_2)^2 s + \gamma_s)' ds \\
\leq\ & 8J_6\Delta^{q-\frac{1}{2}} + \sqrt{C_4}128(k_3)^2 q^2 \Delta^2 \int_0^\tau \sqrt{\mathbb{E}\left(\frac{1}{(x_s)^{2-2q}}\right)} ds \\
& + \int_0^\tau \mathbb{E}\sup_{0\leq l\leq s}|\mathcal{E}_l|^2 (6T(k_2)^2 s + \gamma_s)' ds,
\end{aligned}$$

where we have used Lemma 3.3.7 in the second step. At this point we want to estimate the inverse moments of (x_t) and to do so we consider the transformation $v = x^{2-2q}$ and apply Itô's formula to get

$$\begin{aligned}
v_t &= v_0 + \int_0^t (\underbrace{(1-2q)(1-q)(k_3)^2}_{K_0} + \underbrace{2(1-q)k_1}_{K_1}(v_s)^{\frac{1-2q}{2-2q}} - \underbrace{2(1-q)k_2}_{K_2} v_s) ds \\
&\quad + \int_0^t \underbrace{2k_3(1-q)}_{K_3}\sqrt{v_s} dW_s,
\end{aligned}$$

for $t \in [0,T]$, where $v_0 = (x_0)^{2-2q} > 0$. Denote the drift coefficient of the process (v_t) by $a(v_t)$ and consider the function

$$\alpha(v) := a(v) - \lambda + K_2 v + \underbrace{\frac{(2q-1)(\lambda+K_0)^{\frac{1}{2q-1}}}{(k_1)^{\frac{2-2q}{2q-1}}}}_{\eta(\lambda)} v,$$

where $\lambda \geq 0$. Some elementary calculations show that this function attains its minimum at $v^* := \left(\frac{k_1(2q-1)}{\eta(\lambda)}\right)^{2-2q}$ and $\alpha(v^*) = 0$, thus

$$a(v) \geq \lambda - (K_2 + \eta(\lambda))\, v.$$

Consider the process $(\zeta_t(\lambda))$ defined through

$$\zeta_t(\lambda) = \zeta_0 + \int_0^t (\lambda - (K_2 + \eta(\lambda))\, \zeta_s) ds + \int_0^t K_3 \sqrt{\zeta_s} dW_s, \tag{3.4.2}$$

for $t \in [0,T]$ with $\zeta_0(\lambda) = v_0$. Process (3.4.2) is a square root diffusion process and when $\frac{2\lambda}{(K_3)^2} - 1 \geq 0$ or

$$\lambda \geq 2(1-q)^2 (k_3)^2, \tag{3.4.3}$$

the process is a CIR process which remains positive if $\zeta_0(\lambda) > 0$. By a comparison theorem [KS88, Prop. 5.2.18] we obtain that $v_t \geq \zeta_t(\lambda) > 0$ a.s. or $(v_t)^{-1} \leq (\zeta_t(\lambda))^{-1}$ a.s. or equivalently $(x_t)^{2q-2} \leq (\zeta_t(\lambda))^{-1}$ a.s. The inverse moment bounds of $(\zeta_t(\lambda))$ follow by [DNS11, (3.1)]

$$\sup_{t\in[0,T]} \mathbb{E}(\zeta_t(\lambda))^p < \infty, \qquad \text{for } p > -2\frac{\lambda}{K_3^2}$$

by choosing big enough λ and particularly such that (3.4.3) holds strictly. Therefore,

$$\mathbb{E} \sup_{0\leq t\leq \tau} |\mathcal{E}_t|^2 \leq 8J_6 \Delta^{q-\frac{1}{2}} + \int_0^\tau \mathbb{E} \sup_{0\leq l\leq s} |\mathcal{E}_l|^2 (6T(k_2)^2 s + \gamma_s)' ds. \tag{3.4.4}$$

Relation (3.4.4) for $\tau = \tau_l$ implies

$$\begin{aligned}
\mathbb{E} \sup_{0\leq t\leq \tau_l} (\mathcal{E}_t)^2 &\leq 8J_6 \Delta^{q-\frac{1}{2}} + \int_0^{\tau_l} \mathbb{E} \sup_{0\leq l\leq s} (\mathcal{E}_l)^2 (6T(k_2)^2 s + \gamma_s)' ds \\
&\leq 8J_6 \Delta^{q-\frac{1}{2}} + \int_0^{l} \mathbb{E} \sup_{0\leq j\leq u} (\mathcal{E}_{\tau_j})^2 du \\
&\leq 8J_6 e^l \Delta^{q-\frac{1}{2}}, \qquad (3.4.5)
\end{aligned}$$

where in the last step we have used Gronwall's inequality. Using again relation (3.4.4) for $\tau = T$ and under the change of variables $u = 6T(k_2)^2 s + \gamma_s$ we get

$$
\begin{aligned}
&\mathbb{E} \sup_{0\le t\le T} (\mathcal{E}_t)^2 \le 8J_6 \Delta^{q-\frac{1}{2}} + \int_0^{6(k_2)^2T^2+\gamma_T} \mathbb{E} \sup_{0\le j\le u} (\mathcal{E}_{\tau_j})^2 du \\
\le\ & 8J_6\Delta^{q-\frac{1}{2}} + \int_0^{\infty} \mathbb{E}\left(\sup_{0\le j\le u} (\mathbb{I}_{\{6(k_2)^2T^2+\gamma_T\ge u\}}\mathcal{E}_{\tau_j})^2\right) du \\
\le\ & 8J_6\Delta^{q-\frac{1}{2}} + \int_0^{6(k_2)^2T^2} \mathbb{E} \sup_{0\le j\le u} (\mathcal{E}_{\tau_j})^2 du \\
& + \int_{6(k_2)^2T^2}^{\infty} \mathbb{P}(6(k_2)^2T^2+\gamma_T \ge u)\mathbb{E}\left(\sup_{0\le j\le u} (\mathcal{E}_{\tau_j})^2 \big| \{6(k_2)^2T^2+\gamma_T\ge u\}\right) du \\
\le\ & 8J_6\Delta^{q-\frac{1}{2}} + 8J_6 e^{6(k_2)^2T^2}\Delta^{q-\frac{1}{2}} + \int_0^{\infty} \mathbb{P}(\gamma_T\ge u)\mathbb{E} \sup_{0\le j\le u} (\mathcal{E}_{\tau_j})^2 du \\
\le\ & 16J_6 e^{6(k_2)^2T^2}\Delta^{q-\frac{1}{2}} + 8J_6\Delta^{q-\frac{1}{2}} \int_0^{\infty} \mathbb{P}(\gamma_T\ge u)e^u du,
\end{aligned}
$$

where in the last steps we have used (3.4.5). We proceed by showing that $u \to \mathbb{P}(\gamma_T \ge u)e^u \in \mathcal{L}^1(\mathbb{R}_+)$. Markov's inequality implies

$$\mathbb{P}(\gamma_T \ge u) \le e^{-\epsilon u}\mathbb{E}(e^{\epsilon\gamma_T}),$$

for any $\epsilon > 0$. The following bound holds

$$\gamma_T = \int_0^T \frac{128(k_3)^2q^2}{\left[(y_s)^{1-q} + (x_s)^{1-q}\right]^2} ds \le 128(k_3)^2q^2 \int_0^T (x_s)^{2q-2} ds,$$

thus

$$\mathbb{E}(e^{\epsilon\gamma_T}) \le \mathbb{E}\left(e^{\epsilon 128(k_3)^2q^2\int_0^T (x_s)^{2q-2}ds}\right), \tag{3.4.6}$$

where $-1 < 2q-2 < 0$. It remains to bound the exponential inverse moments of (x_t) defined through the stochastic integral equation (3.2.1). Exponential inverse moments for the CIR process are known [HK08, Th. 3.1] and are given by

$$\mathbb{E}e^{\delta\int_0^t(\zeta_s(\lambda))^{-1}ds} \le C_{HK}(\zeta_0)^{-\frac{1}{2}(\nu(\lambda) - \sqrt{\nu(\lambda)^2 + 8\frac{\delta}{(K_3)^2}})}, \tag{3.4.7}$$

for $0 \le \delta \le \left(\frac{2\lambda}{K_3^2} - 1\right)^2 \frac{K_3^2}{8} =: \nu(\lambda)^2 \frac{K_3^2}{8}$, where the positive constant C_{HK} is explicitly given in [HK08, (10)] depends on the parameters k_2, k_3, T, q, but is independent of ζ_0. Thus the other condition that we require for parameter λ is

$$\lambda \ge 2(1-q)\sqrt{2\delta}(k_3) + 2(1-q)^2(k_3)^2. \tag{3.4.8}$$

When (3.4.8) is satisfied then (3.4.3) is satisfied too; thus there is actually no restriction on the coefficient δ in (3.4.7) since we can always choose appropriately a λ such that (3.4.8) holds. Relation (3.4.6) becomes

$$\mathbb{E}(e^{\epsilon\gamma_T}) \le \mathbb{E}\left(e^{\epsilon 128(k_3)^2 q^2 \int_0^T (v_s)^{-1} ds}\right) \le \mathbb{E}\left(e^{\epsilon 128(k_3)^2 q^2 \int_0^T (\zeta_s(\lambda))^{-1} ds}\right). \tag{3.4.9}$$

We therefore require that

$$128(k_3)^2 q^2 \epsilon \le (\nu(\lambda))^2 \frac{K_3^2}{8} \tag{3.4.10}$$

and can always find a $\epsilon > 1$, such the above relation holds by choosing appropriately λ as discussed before. Relation (3.4.9) becomes

$$\mathbb{E}(e^{\epsilon\gamma_T}) \le C_{HK}(\zeta_0)^{-\frac{\nu(\lambda)}{2}},$$

and therefore

$$\mathbb{P}(\gamma_T \ge u) \le C_{HK}(x_0)^{(1-q)\nu(\lambda)} e^{-\epsilon u},$$

where λ is chosen such that (3.4.10) holds with $\epsilon > 1$. We conclude

$$\begin{aligned}\mathbb{E} \sup_{0\le t\le T} (\mathcal{E}_t)^2 &\le 16 J_6 e^{6(k_2)^2 T^2} \Delta^{q-\frac{1}{2}} + 8 J_6 C_{HK}(x_0)^{(1-q)\nu(\lambda)} \Delta^{q-\frac{1}{2}} \int_0^\infty e^{(1-\epsilon)u} du \\ &\le C \cdot \Delta^{q-\frac{1}{2}},\end{aligned}$$

by choosing $\epsilon > 1$, where $C = C(k_1, k_2, k_3, T, q, \epsilon) := 8J_6(2e^{6(k_2)^2T^2} + \frac{C_{HK}}{\epsilon - 1}(x_0)^{(1-q)\nu(\lambda)})$, is as given in statement of Theorem 3.2.4.

3.5 Alternative approach with reduced rate of convergence.

In this section we briefly discuss the case where instead of (3.2.8) we use directly (3.2.4). Then, Lemmata 3.3.5, 3.3.6 and 3.3.7 still hold, i.e. the moment bounds and error bounds of (y_t^{SD}), as well as the moment bounds involving the auxiliary process (h_t) are true. The proof of the convergence results follow the same lines as in Sections 3.3 and Section 3.4. The error $\mathcal{E}_t := h_t - x_t$ now reads

$$h_t - x_t = \int_0^t (f_\theta(y_{\hat{s}}, y_{\tilde{s}}) - f_\theta(x_s, x_s))\, ds + \int_0^t \mathrm{sgn}(z_s)\, (g(y_{\hat{s}}, y_s) - g(x_s, x_s))\, dW_s.$$

The main difference is in the estimation (3.3.7) that now becomes

$$\begin{aligned}|\mathrm{sgn}(z_s) g(y_{\hat{s}}, y_s) - g(x_s, x_s)|^2 \quad \leq \quad & 3(k_3)^2 \Big((y_{\hat{s}})^{2q-1} y_s (\mathrm{sgn}(z_s) - 1)^2 \\ & + y_s |y_{\hat{s}} - y_s|^{2q-1} + |h_s - y_s| + |h_s - x_s| \Big).\end{aligned}$$

The first term on the right-hand side of the above inequality containing the $\mathrm{sgn}(z_s)$ will contribute in a negative way to the rate of convergence. We do not give all the details, but just mention that in order to bound the expectation of that term, which can be done in the following way,

$$\begin{aligned} & \mathbb{E}(y_{\hat{s}})^{2q-1} y_s |\mathrm{sgn}(z_s) - 1|^2 = \mathbb{E}\left(4 (y_{t_n})^{2q-1} y_s \mathbb{I}_{\{z_s \leq 0\}}\right) \\ \leq \; & 4\mathbb{E}\left|(y_{t_n})^{2q-1} y_s - (y_{t_n})^{2q}\right| \\ & + 4\mathbb{E}\left((y_{t_n})^{2q} \mathbb{I}_{\{z_s \leq 0\}} \mathbb{I}_{\{y_{t_n} \leq \Delta^{1-2\xi}\}}\right) + 4\mathbb{E}\left((y_{t_n})^{2q} \mathbb{I}_{\{z_s \leq 0\}} \mathbb{I}_{\{y_{t_n} > \Delta^{1-2\xi}\}}\right) \\ \leq \; & 4\mathbb{E}\left|(y_{t_n})^{2q-1}(y_s - y_{t_n})\right| + 4\Delta^{2q-4q\xi} \\ & + 4\sqrt{\mathbb{E}(y_{t_n})^{4q}} \sqrt{\mathbb{P}\left(\{z_s \leq 0\} \cap \{y_{t_n} > \Delta^{1-2\xi}\}\right)}, \end{aligned}$$

we need to estimate the probability of z_t being negative when at the same time $y_{t_n} > \Delta^{1-2\xi}$, for $0 < \xi < \frac{1}{2}$.

Lemma 3.5.10 *For every $t \in [t_n, t_{n+1}]$ it holds*

$$\mathbb{P}\left(\{z_t \leq 0\} \cap \{y_{t_n} > \Delta^{1-2\xi}\}\right) \leq C_{k_2,k_3,\theta,\Delta} \sqrt{\Delta}, \tag{3.5.1}$$

where $C_{k_2,k_3,\theta,\Delta} := \frac{k_3}{\sqrt{1 - k_2(2-\theta)\Delta}}$ and $\Delta(2-\theta) < \frac{1}{k_2}$ and $\frac{(k_3)^2}{(1+k_2\theta\Delta)} \leq 4k_2$. Relation (3.5.1) implies that $\mathbb{P}\left(\{z_t \leq 0\} \cap \{y_{t_n} > \Delta^{1-2\xi}\}\right) = O(\sqrt{\Delta})$, as $\Delta \downarrow 0$. □

Proof of Lemma 3.5.10. By the definition (3.2.5) of (z_t) for $t \in [t_n, t_{n+1}]$ and for $0 < \xi < \frac{1}{2}$, we have for the following event $A := \{z_t \leq 0\} \cap \{y_{t_n} > \Delta^{1-2\xi}\}$ that

$$\begin{aligned} A &= \left\{ (y_{t_n})^{q-\frac{1}{2}}(W_t - W_{t_n}) \leq -\frac{2(1+k_2\theta\Delta)}{k_3}\sqrt{y_n} \right\} \cap \{y_{t_n} > \Delta^{1-2\xi}\} \\ &\subseteq A_1 \cup A_2, \end{aligned} \tag{3.5.2}$$

where

$$A_1 := \left\{ W_t - W_{t_n} \leq -\frac{2(1+k_2\theta\Delta)}{k_3}\sqrt{y_n}(y_{t_n})^{-q+\frac{1}{2}} \right\} \cap \{y_{t_n} \geq 1\},$$

and

$$A_2 := \left\{ W_t - W_{t_n} \leq -\frac{2(1+k_2\theta\Delta)}{k_3}\sqrt{y_n}(y_{t_n})^{-q+\frac{1}{2}} \right\} \cap \{1 > y_{t_n} > \Delta^{1-2\xi}\}.$$

The following inclusion relations hold for the event A_1,

$$\begin{aligned} A_1 &\subseteq \{\Delta W_n \leq -\frac{2(1+k_2\theta\Delta)}{k_3(y_{t_n})^{q-\frac{1}{2}}}\sqrt{y_{t_n}(1 - \frac{k_2\Delta}{1+k_2\theta\Delta}) - \frac{(k_3)^2\Delta(y_{t_n})^{2q-1}}{4(1+k_2\theta\Delta)^2}}\} \\ &\quad \cap\{y_{t_n} \geq 1\} \\ &\subseteq \left\{ \Delta W_n \leq -\frac{2(1+k_2\theta\Delta)}{k_3}\sqrt{\frac{1-k_2(2-\theta)\Delta}{1+k_2\theta\Delta} - \frac{(k_3)^2\Delta}{4(1+k_2\theta\Delta)^2}} \right\} \\ &\subseteq \left\{ \frac{\Delta W_n}{\sqrt{t-t_n}} \leq -\frac{2}{k_3}\frac{\sqrt{(1-k_2(2-\theta)\Delta)(1+k_2\theta\Delta)}}{\sqrt{t-t_n}} \right\} \end{aligned}$$

when $\Delta(2-\theta) < \frac{1}{k_2}$ and $\frac{(k_3)^2}{(1+k_2\theta\Delta)} \leq 4k_2$, where $\Delta W_n := W_t - W_{t_n}$. We obtain

$$\mathbb{P}(G \leq -\beta) = \int_{-\infty}^{-\beta} \frac{1}{\sqrt{2\pi}} e^{-u^2/2} du \leq \int_{-\infty}^{-\beta} e^{-u^2/2} du = \int_{\beta}^{\infty} e^{-u^2/2} du \leq \frac{e^{-(\beta)^2/2}}{\beta}, \tag{3.5.3}$$

for every standard normal random variable G, where in the last step we have used [KS88, (9.20), p.112] valid for $\beta > 0$. Using the fact that $\frac{\Delta W_n}{\sqrt{t-t_n}}$ is a standard normal r.v. and ignoring the exponential term in (3.5.3), since its exponent is negative, we get that

$$\mathbb{P}(A_1) \leq \frac{k_3}{2\sqrt{(1-k_2(2-\theta)\Delta)}}\sqrt{t-t_n} \leq C_{k_2,k_3,\theta,\Delta}\sqrt{\Delta}. \tag{3.5.4}$$

The following inclusion relations hold for the event A_2,

$$\begin{aligned}
&A_2\\
&\subseteq \left\{\Delta W_n \le -\frac{2(1+k_2\theta\Delta)}{k_3(y_{t_n})^{q-\frac{1}{2}}}\sqrt{y_{t_n}\frac{1-k_2(1-\theta)\Delta}{1+k_2\theta\Delta}+\frac{k_1\Delta}{1+k_2\theta\Delta}-\frac{(k_3)^2\Delta(y_{t_n})^{2q-1}}{4(1+k_2\theta\Delta)^2}}\right\}\\
&\cap\{1>y_{t_n}>\Delta^{1-2\xi}\}\\
&\subseteq \left\{\Delta W_n \le -\frac{2(1+k_2\theta\Delta)}{k_3}\sqrt{\Delta^{1-2\xi}\frac{1-k_2(1-\theta)\Delta}{1+k_2\theta\Delta}+(k_1-\frac{(k_3)^2}{4(1+k_2\theta\Delta)})\frac{\Delta}{1+k_2\theta\Delta}}\right\}\\
&\subseteq \left\{\frac{\Delta W_n}{\sqrt{t-t_n}} \le -\frac{2}{k_3}\frac{\sqrt{(1-k_2(1-\theta)\Delta)(1+k_2\theta\Delta)}}{\sqrt{t-t_n}}\Delta^{\frac{1}{2}-\xi}\right\}
\end{aligned}$$

when $\Delta(1-\theta)<\frac{1}{k_2}$ and $\frac{(k_3)^2}{(1+k_2\theta\Delta)}\le 4k_1$. Using again (3.5.3) we have that

$$\begin{aligned}
\mathbb{P}(A_2) &\le \frac{k_3}{2\sqrt{(1-k_2(1-\theta)\Delta)}}\sqrt{t-t_n}\Delta^{\xi-\frac{1}{2}}e^{-\frac{2}{(k_3)^2}\frac{(1-k_2(1-\theta)\Delta)(1+k_2\theta\Delta)}{\sqrt{t-t_n}}\Delta^{1-2\xi}}\\
&\le \frac{k_3}{2\sqrt{(1-k_2(1-\theta)\Delta)}}\Delta^{\xi}e^{-\frac{2}{(k_3)^2}(1-k_2(1-\theta)\Delta)(1+k_2\theta\Delta)\Delta^{-2\xi}}. \qquad (3.5.5)
\end{aligned}$$

Taking probabilities in the inclusion relation (3.5.2) and using (3.5.4) and (3.5.5) we get

$$\begin{aligned}
\mathbb{P}(A) &\le \mathbb{P}(A_1)+\mathbb{P}(A_2)\\
&\le C_{k_2,k_3,\theta,\Delta}\sqrt{\Delta}+\frac{k_3}{2\sqrt{(1-k_2(1-\theta)\Delta)}}\Delta^{\xi}e^{-\frac{2}{(k_3)^2}(1-k_2(1-\theta)\Delta)(1+k_2\theta\Delta)\Delta^{-2\xi}}\\
&\le C_{k_2,k_3,\theta,\Delta}\sqrt{\Delta},
\end{aligned}$$

since $\Delta^{\xi}e^{-\Delta^{-2\xi}}=o(\sqrt{\Delta})$ as $\Delta\downarrow 0$. Finally, note that $C_{k_2,k_3,\theta,\Delta}=\frac{k_3}{\sqrt{1-k_2(2-\theta)\Delta}}\to k_3$ as $\Delta\downarrow 0$ which justifies the $O(\sqrt{\Delta})$ notation. □

Applying Lemma 3.5.10 we obtain for $s\in[t_n,t_{n+1}]$ that

$$\begin{aligned}
&\mathbb{E}(y_{\hat{s}})^{2q-1}y_s|\mathrm{sgn}(z_s)-1|^2 \le 4\sqrt{\mathbb{E}(y_{t_n})^{4q-2}}\sqrt{\mathbb{E}|y_s-y_{t_n}|^2}+4\Delta^{2q-4q\xi}\\
&+4\sqrt{\mathbb{E}(y_{t_n})^{4q}}\sqrt{C_{k_2,k_3,\theta,\Delta}\sqrt{\Delta}}\\
\le\ & 4\sqrt{A_{4q-2}\mathbb{E}(x_0+k_1T)^{4q-2}}\sqrt{\hat{A}_2}\sqrt{\Delta}+4\Delta^{2q-4q\xi}\\
&+4\sqrt{A_{4q}\mathbb{E}(x_0+k_1T)^{4q}}\sqrt{C_{k_2,k_3,\theta,\Delta}}\Delta^{\frac{1}{4}},
\end{aligned}$$

where we have used Lemmata 3.3.6 and 3.3.5 in the final step. For $\xi = \frac{1}{2} - \frac{1}{16q}$ we get the estimate
(3.5.6)

$$\mathbb{E}(y_{\hat{s}})^{2q-1} y_s |\mathrm{sgn}(z_s) - 1|^2 \leq 4 \left(\sqrt{A_{4q}\mathbb{E}(x_0 + k_1 T)^{4q} C_{k_2,k_3,\theta,\Delta}} \vee 1 \right) \Delta^{\frac{1}{4}},$$

which in turn implies

$$\sup_{0\leq t\leq T} \mathbb{E}|\mathcal{E}_t| \leq \left(J_2 \frac{\Delta^{\frac{1}{4}}}{m e_m} + J_3^* \frac{\Delta^{q-\frac{1}{2}}}{m e_m} + 3(k_3)^2 T \frac{1}{m} \right) e^{k_2 T}, \tag{3.5.7}$$

$$J_2 := 12(k_3)^2 T (\sqrt{A_{4q}\mathbb{E}(x_0 + k_1 T)^{4q} C_{k_2,k_3,\theta,\Delta}} \vee 1)$$

and $J_3^* := (3/2)J_3$. We use the process $\overline{\gamma}(t) := (192/128)\gamma(t)$ and following the same lines as in Section 3.4 we conclude

$$\mathbb{E} \sup_{0\leq t\leq T} |y_t^{SD}(q) - x_t|^2 \leq C \Delta^{(q-\frac{1}{2})\wedge\frac{1}{4}},$$

where

$$C := 8(J_2 \vee J_3^*) \left(2e^{6(k_2)^2 T^2} + \frac{C_{HK}}{\epsilon - 1} (x_0)^{(1-q)\nu(\lambda)} \right).$$

3.6 Numerical Experiments.

We discretize the interval $[0, T]$ with a number of steps in power of 2. The semi-discrete (SD) scheme is given by

$$y_{t_{n+1}}^{SD} = \left(\sqrt{ y_{t_n} \left(1 - \frac{k_2 \Delta}{1 + k_2\theta\Delta} \right) + \frac{k_1 \Delta}{1 + k_2\theta\Delta} - \frac{(k_3)^2 \Delta}{4(1 + k_2\theta\Delta)^2} (y_{t_n})^{2q-1} } + \frac{k_3}{2(1 + k_2\theta\Delta)} (y_{t_n})^{q-\frac{1}{2}} \Delta W_n \right)^2, \tag{3.6.1}$$

for $n = 0, \ldots, N-1$, where $\Delta W_n := W_{t_{n+1}} - W_{t_n}$ are the increments of the Brownian motion which are Gaussian random variables with $\Delta W_n \sim \mathcal{N}(0, \Delta)$.

The ALF (Alfonsi) scheme [Alf13, Sec. 3] is an implicit scheme which requires solving the non-linear equation
(3.6.2)

$$Y_{n+1} = y_{t_n} + (1-q)(k_1 (Y_{n+1})^{\frac{-q}{1-q}} - k_2 Y_{n+1} - \frac{q(k_3)^2}{2} (Y_{n+1})^{-1})\Delta + k_3(1-q)\Delta W_n,$$

and then computing $y_{t_{n+1}}^{ALF} = (Y_{n+1})^{\frac{1}{1-q}}$. The estimation of Y_{n+1} in (3.6.2) can be done for example with Newton's method, but requires a small enough Δ.[4] We also consider a scheme recently proposed in [Hal15c] using again the SD method, but in a different way,

(3.6.3)
$$y_{t_{n+1}}^{HAL}(q) = \left| (y_{t_n}(1-k_2\Delta) + k_1\Delta - \frac{q(k_3)^2\Delta}{2}(y_{t_n})^{2q-1})^{1-q} + k_3(1-q)\Delta W_n \right|^{\frac{1}{1-q}},$$

for $n = 0, \ldots, N-1$. Note the similarity in the expressions of (3.6.3) and the SD scheme (3.6.1) proposed here. This is not strange, because they both rely in the same way of splitting the drift coefficient. In particular, in the explicit HAL scheme, the following process is considered

$$y_t^{HAL}(q) = y_{t_n} + \widetilde{f}_1(y_{t_n}) \cdot \Delta + \int_{t_n}^{t} \widetilde{f}_2(y_s)ds + \int_{t_n}^{t} \mathrm{sgn}(z_s)\widetilde{g}(y_s)dW_s, \tag{3.6.4}$$

for $t \in (t_n, t_{n+1}]$ with $y_0 = x_0$ a.s. where now

$$f(x) = \underbrace{k_1 - k_2x - \frac{q(k_3)^2}{2}x^{2q-1}}_{\widetilde{f}_1(x)} + \underbrace{\frac{q(k_3)^2}{2}x^{2q-1}}_{\widetilde{f}_2(x)}. \quad \widetilde{g}(x) = k_3x^q$$

and

$$z_t = \left(y_{t_n}(1-k_2\Delta) + k_1\Delta - \frac{q(k_3)^2\Delta}{2}(y_{t_n})^{2q-1} \right)^{1-q} + k_3(1-q)(W_t - W_{t_n}).$$

A comparison with (3.2.2) and (3.2.3) shows that $\widetilde{f}_2(x) = 2qf_2(x)$ and $\widetilde{g}(x) = g(x,x)$, for $\theta = 0$. We write (3.6.4) again as

$$\begin{aligned} y_t^{HAL}(q) &= y_{t_n} + \left(k_1 - k_2y_{t_n} - \frac{q(k_3)^2}{2}(y_{t_n})^{2q-1} \right)\Delta + \int_{t_n}^{t} \frac{q(k_3)^2}{2}(y_s)^{2q-1}ds \\ &\quad + k_3\int_{t_n}^{t} \mathrm{sgn}(z_s)(y_s)^q dW_s \end{aligned} \tag{3.6.5}$$

and the process (3.6.5) is well defined when

$$(k_3)^2 \leq \frac{2}{q}k_1 \text{ and } \Delta \leq \frac{2}{2k_2 + q(k_3)^2}. \tag{3.6.6}$$

[4] In the CIR case, i.e. when $q = 1/2$ (3.6.2) simplifies to a solution of a quadratic equation.

The reader can compare again with (3.2.4) for $\theta = 0$. Solving for y_t, we end up with $y_t^{HAL}(q) = |z_t|^{\frac{1}{1-q}}$. The main result in [Hal15c] is

$$\mathbb{E}|y_t^{HAL} - x_t|^2 \leq C \cdot \Delta^{2q(q-\frac{1}{2})},$$

when (3.6.6) holds, implying a rate of convergence at least $q(q - \frac{1}{2})$ which is bigger than the rate of convergence of the SD scheme proposed here which is at least $\frac{1}{2}(q - \frac{1}{2})$ (see Th. 3.2.4).

We also consider two more linear-implicit schemes that were stated in the introduction and discussed in Appendix D. Namely, we compare with the balanced implicit method (BIM) with appropriate weight functions to guarantee positivity ([KS06, Th. 5.9]), which reads

$$y_{t_{n+1}}^{BIM}(q) = \frac{y_{t_n} + k_1\Delta + k_3(y_{t_n})^q(\Delta W_n + |\Delta W_n|)}{1 + k_2\Delta + k_3(y_{t_n})^{q-1}|\Delta W_n|},$$

and the balanced Milstein method (BMM) with the suggested weight functions [KS06, Th. 5.9] that is given by

(3.6.7)
$$y_{t_{n+1}}^{BMM}(q) = \frac{y_{t_n} + (k_1 + (\Theta - 1)k_2 y_{t_n})\Delta + k_3(y_{t_n})^q\Delta W_n + \frac{q(k_3)^2}{2}(y_{t_n})^{2q-1}(\Delta W_n)^2}{1 + \Theta k_2\Delta + \frac{q(k_3)^2}{2}|y_{t_n}|^{2q-2}\Delta}.$$

We take the relaxation parameter Θ to be $1/2$ as recommended in [KS06, (5.10)].

We aim to show experimentally the order of convergence for the above positivity preserving methods for the estimation of the true solution of the CEV model (3.2.1), i.e. the semi-discrete methods SD (3.6.1) and the HAL scheme (3.6.3), as well as the implicit ALF scheme (3.6.2) and the linear-implicit schemes BIM and BMM. The choice of the parameters is the same as in [KJ06, Fig. 6] with $k_3 = 0.4$. In particular $(x_0, k_1, k_2, k_3, q, T) = (\frac{1}{16}, \frac{1}{16}, 1, 0.4, \frac{3}{4}, 1)$.

Furthermore, we would also like to reveal the dependence of the order of the semi-discrete methods on q, i.e. we want to verify our theoretical results and in particular the order shown in Theorem 3.2.4. We take the level of implicitness of SD method (3.6.1) to be $\theta = 1$, i.e. we consider the fully implicit scheme. We also discuss about the fully explicit scheme, that is, when $\theta = 0$, but also an intermediate scheme $\theta = 1/2$, in Section 3.7.

We want to estimate the endpoint $\mathcal{L}^2$-norm $\epsilon = \sqrt{\mathbb{E}|y^{(\Delta)}(T) - x_T|^2}$, of the difference between the numerical scheme evaluated at step size Δ and the exact solution of (3.2.1). For that purpose, we compute M batches of L simulation paths, where each batch is estimated by $\hat{\epsilon}_j = \frac{1}{L}\sum_{i=1}^{L} |y_{i,j}^{(\Delta)}(T) - y_{i,j}^{(ref)}(T)|^2$ and the Monte Carlo estimator of the error is

$$\hat{\epsilon} = \sqrt{\frac{1}{ML}\sum_{j=1}^{M}\sum_{i=1}^{L} |y_{i,j}^{(\Delta)}(T) - y_{i,j}^{(ref)}(T)|^2},$$

and requires $M \cdot L$ Monte Carlo sample paths. The reference solution is evaluated at step size 2^{-14} of the numerical scheme. For the SD case, we have shown in Theorems 3.2.2 and 3.2.4 that it strongly converges to the exact solution. We simulate $100 \cdot 100 = 10000$ paths, where the choice for $L = 100$ is as in [KPS03, p.118]. The choice of the number of trajectories $M \cdot L = 10^4$ is also considered in [TZ13, Sec. 5] where a fundamental mean-square theorem is proved for SDEs with super-linear growing coefficients satisfying a one-side Lipschitz condition, but unfortunately it is not positivity preserving. Of course, the number of Monte Carlo paths has to be sufficiently large, so as not to significantly hinder the mean-square errors.

We plot in a $\log_2$-$\log_2$ scale and error bars represent 98%-confidence intervals. The results are shown in Table 3.1 and Figure 3.1. Table 3.1 does not present the computed Monte Carlo errors with 98%-confidence, since they were at least 9 times smaller that the mean-square errors.

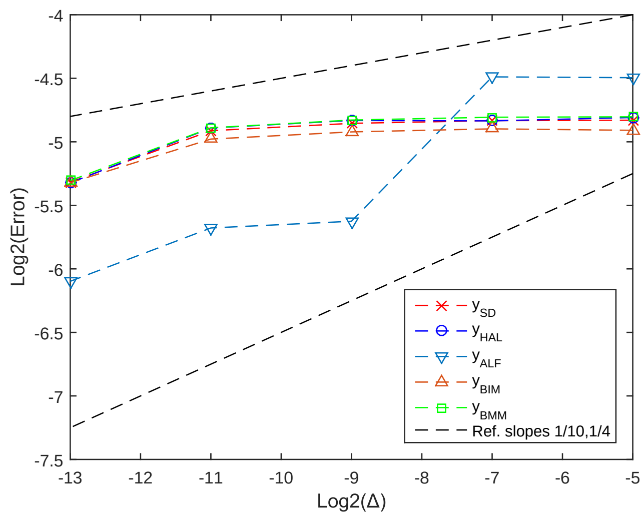
Step Δ	SD($\theta = 1$)	HAL	ALF	BIM	BMM
2^{-5}	0.0352	0.0357	0.0443	**0.0333**	0.0358
2^{-7}	0.0351	0.035	0.0445	**0.0335**	0.0357
2^{-9}	0.0346	0.0351	**0.0203**	0.033	0.0352
2^{-11}	0.0332	0.0337	**0.0195**	0.0317	0.0337
2^{-13}	0.0251	0.025	**0.0146**	0.025	0.0253

Tab. 3.1: Error and step size of fully implicit SD, HAL,ALF, BIM and BMM scheme for (3.2.1) with $(x_0, k_1, k_2, k_3, q, T) = (\frac{1}{16}, \frac{1}{16}, 1, 0.4, \frac{3}{4}, 1)$.

In Table 3.3 we present the computational times,[5] of fully implicit SD, HAL, ALF, BIM and BMM, for the same problem. Figure 3.2 shows the

[5] We simulate with 3.06GHz Intel Pentium, 1.49GB of RAM in Matlab $R2014b$ Software. The random number generator is Mersenne Twister. The evaluated times do not include the random number generation time, since all the methods we compare, involve the same amount of random numbers.

Fig. 3.1: Convergence of fully implicit SD, HAL, ALF, BIM and BMM schemes applied to SDE (3.2.1) with parameters $x_0 = k_1 = \frac{1}{16}, k_2 = 1, k_3 = 0.4, q = 3/4$ and $T = 1$.



Step Δ	SD($\theta = 1$)-Rate	HAL-Rate	ALF-Rate	BIM-Rate	BMM-Rate
2^{-7}	0.002	0.013	-0.003	-0.006	0.002
2^{-9}	0.107	-0.002	0.568	0.012	0.011
2^{-11}	0.029	0.029	0.026	0.028	0.031
2^{-13}	0.203	0.216	0.21	0.172	0.208

Tab. 3.2: Experimental rates of fully implicit SD, HAL,ALF, BIM and BMM scheme for (3.2.1) with $(x_0, k_1, k_2, k_3, q, T) = (\frac{1}{16}, \frac{1}{16}, 1, 0.4, \frac{3}{4}, 1)$.

relation between the error and computer time consumption. As one can see from Table 3.3 the CPU times for ALF are at least 1000 times bigger than the other schemes, thus we choose in Figure 3.2 to restrict our attention to the rest of the methods.

We show, in Table 3.4, the $\mathcal{L}^2$-distance between our proposed method and the other methods for the numerical approximation of (3.2.1). We work as before and estimate the distance

$$d(G, H) = \sqrt{\frac{1}{ML} \sum_{j=1}^{M} \sum_{i=1}^{L} |y_{i,j}^{(\Delta,G)}(T) - y_{i,j}^{(\Delta,H)}(T)|^2}, \tag{3.6.8}$$

Step Δ	Implicit SD	HAL	ALF	BIM	BMM
2^{-5}	**0.0000130**	0.0000164	0.0221883	0.0000174	0.0000196
2^{-7}	**0.0000422**	0.0000558	0.0841705	0.0000584	0.0000657
2^{-9}	**0.0001586**	0.0002137	0.2453943	0.0002207	0.0002482
2^{-11}	**0.0006243**	0.0008437	0.9768619	0.0008703	0.0009795
2^{-13}	**0.0024975**	0.0033977	3.9096332	0.0034785	0.0039143

Tab. 3.3: Average computational time (in seconds) for a path, for different discretizations, for all considered positivity preserving methods for the mean-reverting CEV process (3.2.1) with $q = 3/4$.

between method G and H, by considering sufficient small Δ, and in particular for $\Delta = 10^{-2}, 10^{-3}, 10^{-4}$.

Step Δ	d(SD,HAL)	d(SD,ALF)	d(SD,BIM)	d(SD,BMM)
10^{-2}	0.0005727	0.0716140	0.0038373	0.0005312
10^{-3}	0.0001577	0.0286630	0.0013460	0.0001564
10^{-4}	0.0000498	0.0283117	0.0004448	0.0000498

Tab. 3.4: The $\mathcal{L}^2$-distance between all the considered numerical schemes applied to SDE (3.2.1) with parameter set $x_0 = k_1 = \frac{1}{16}, k_2 = 1, k_3 = 0.4, q = 3/4$ and $T = 1$.

Finally, we examine the behavior of all the methods for a value of the parameter q close to $1/2$. The results are shown in Table 3.5.

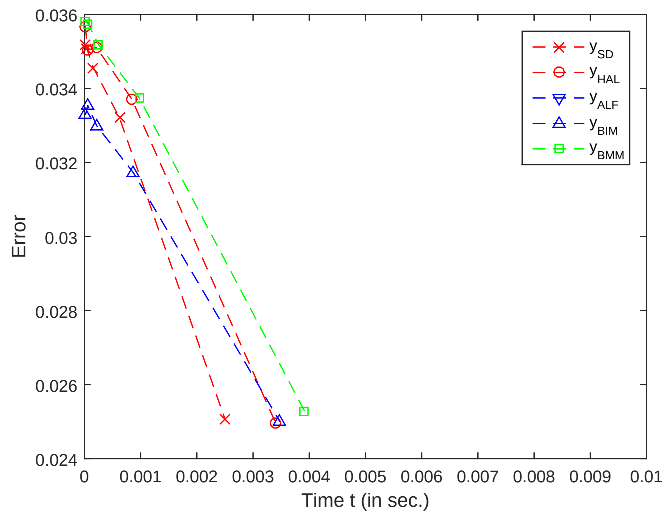
Step Δ	SD($\theta = 1$)	Rate	HAL	Rate	BIM	Rate	BMM	Rate
2^{-10}	0.05913	–	0.05967	–	**0.05582**	–	0.05966	–
2^{-11}	0.05818	0.023	0.05867	0.024	**0.05503**	0.021	0.05867	0.024
2^{-12}	0.05458	0.092	0.05499	0.094	**0.0518**	0.088	0.05499	0.096
2^{-13}	0.04407	0.309	0.04429	0.312	**0.04244**	0.287	0.04429	0.312

Tab. 3.5: Error and step size of fully implicit SD, HAL, BIM and BMM scheme for (3.2.1) with $(x_0, k_1, k_2, k_3, q, T) = (\frac{1}{16}, \frac{1}{16}, 1, 0.4, 0.55, 1)$.

The following points of discussion are worth mentioning.

- The performance of all methods, as shown in Table 3.1 and Figure 3.1, implies, in terms of error estimates, that the implicit ALF scheme performs better, for values of discretization steps $\Delta \leq 2^{-9}$. All the other methods, i.e. the semi-discrete SD and HAL, the BIM and BMM

Fig. 3.2: Strong convergence error of the mean-reverting CEV process (3.2.1) as a function of CPU time (in sec) using positivity preserving schemes SD,HAL, ALF, BIM and BMM with $x_0 = k_1 = \frac{1}{16}, k_2 = 1, k_3 = 0.4, q = 3/4, T = 1$ and 32 digits of accuracy.



have a similar behavior for all values of Δ w.r.t. error estimation as Figure 3.1 shows. The similarity of SD, HAL, BIM and BMM is also indicated in Table 3.4, where we see how close they are w.r.t. the $\mathcal{L}^2$-norm. Nevertheless, Table 3.4 also shows that in order to get an accu racy to at least two decimal digits, which in practice may be adequate concerning that we want for example to evaluate an option and thus our results are in euros, we are free to use any of the above available methods. We may then choose the fastest one, as will be discussed later on.

- The experimental strong order of convergence of implicit SD for problem (3.2.1) is 1/5 (at least $1/2(q - 1/2) = 1/8$ as shown theoretically and presented in Table 3.1). We also see that all methods converge with similar orders and the theoretically rate 1 of the ALF method [Alf13] does not hold for these values of Δ. Thus, again we see that the rate in practical situations does not necessarily matter, if one has to consider very small values of Δ to achieve it. Moreover, we present in Table 3.6 the performance of the explicit SD method and see that it is very close to the implicit, which is of course natural to happen.

Step Δ	98%-SD-Error($\theta = 0$)	Rate
2^{-7}	0.0344244	–
2^{-9}	0.0342415	0.0038
2^{-11}	0.0331273	0.0239
2^{-13}	0.0250195	0.2025

Tab. 3.6: The performance of fully explicit SD scheme (3.6.1) applied to SDE (3.2.1) with parameter set $(x_0, k_1, k_2, k_3, q, T) = (\frac{1}{16}, \frac{1}{16}, 1, 0.4, \frac{3}{4}, 1)$

- Table 3.5 concerns the case where the parameter q is 0.55. We do not present the ALF method since it required smaller values of Δ. All the methods again behave quite the same, with the BIM performing better w.r.t. error estimation.

- In practice, the computer time consumed to provide a desired level of accuracy, is of great importance. Especially, in financial applications, a scheme is considered better when except of its accuracy, it is implemented faster. As mentioned before, the SD method as well as the HAL method performs well in that aspect, compared to the implicit ALF method, which requires the estimation of a root of a non-linear equation in each step and is therefore time consuming. This is presented in Table 3.3 and Figure 3.2 which illustrates the advantage of the semi-discrete method SD, performing slightly better than HAL and BMM, better than BIM, and of course a lot better compared with ALF (over 1000 times quicker to achieve an accuracy of almost two decimal digits.) Moreover, the explicit SD, performs slightly better in that aspect, as shown in Table 3.7.

- A negative step of a numerical method appears when the computer-generated random variable exceeds a certain threshold, which tends to increase as the step size Δ decreases. Thus, the undesirable effect of negative values that are produced by some numerical schemes (such as the explicit Euler (EM) and standard Milsten (M)), tends to disappear, since after a certain small step size, the threshold exceeds the maximum standard normal random number attainable by the computer system.

Step Δ	Time/Path(in sec): Fully Explicit SD	(Implicit)
2^{-5}	0.000013	(0.000013)
2^{-7}	0.0000411	(0.0000422)
2^{-9}	0.0001545	(0.0001586)
2^{-11}	0.0006048	(0.0006243)
2^{-13}	0.0024319	(0.0024975)

Tab. 3.7: Average computational time for a path (in seconds) for fully explicit SD method for $q = 3/4$.

3.7 Approximation of Stochastic Model (3.1.1).

So far we have focused on the process (V_t), which is one part of the two-dimensional system (3.1.1). Nevertheless, it can be treated independently, since the only way that it interacts with the process (S_t) is through the correlation ρ of the Wiener processes. First we apply Itô's formula on $\ln(S_t)$ to get,

$$(3.7.1)\quad \ln S_t = \ln S_0 + \int_0^t \mu du - \frac{1}{2}\int_0^t (V_u)^{2p} du + \int_0^t (V_u)^p dW_u, \quad t \in [0, T].$$

Then, we consider two different schemes for the integration of (3.7.1).[6] The first is the EM scheme which reads

$$(3.7.2)\qquad \ln S^{EM}_{t_{n+1}} = \ln S_{t_n} + \mu\Delta - \frac{1}{2}(V_{t_n})^{2p}\Delta + (V_{t_n})^p \Delta W_n,$$

has strong convergence order 1/2 and is easy to implement. The second scheme, which is based on an interpolation of the drift term and an interpolation of the diffusion term, considering decorrelation of the diffusion term, including a higher order Milstein term [KJ06, Sec. 4.2], is denoted IJK and is given by [KJ06, (137)]

[6] The reason for not considering other schemes such as the two-dimensional Milstein is that they generally are time consuming, since they involve additional random number generation for the approximation of double Wiener integrals.

$$\ln S^{IJK}_{t_{n+1}} = \ln S_{t_n} + \mu\Delta - \frac{1}{4}\left((V_{t_n})^{2p} + (V_{t_{n+1}})^{2p}\right)\Delta + \rho(V_{t_n})^p\Delta\widetilde{W}_n$$
$$+\frac{1}{2}((V_{t_n})^p + (V_{t_{n+1}})^p)(\Delta W_n - \rho\Delta\widetilde{W}_n) \quad (3.7.3)$$
$$+\frac{1}{2}\rho p k_3 (V_{t_n})^{q+p-1}((\Delta\widetilde{W}_n)^2 - \Delta).$$

We therefore consider the EM scheme (3.7.2) combined with SD (3.6.1), the IJK scheme (3.7.3) combined with SD (3.6.1) and compare with the case where the stochastic variance ($p = \frac{1}{2}$) is integrated with BMM scheme (3.6.7), for three different correlation parameters, $\rho = 0, \rho = -0.4$ and $\rho = -0.8$ with $S_0 = 100, \mu = 0.05$, as in [KJ06, Sec. 5]. We present in Tables 3.8, 3.9 and 3.10 and Figures 3.3, 3.4 and 3.5, the errors, in the sense of distance (3.6.8), for all the above considered ways of numerical integration of process (S_t), for different step sizes, as well as the average computational time (in seconds) consumed for each discretization.

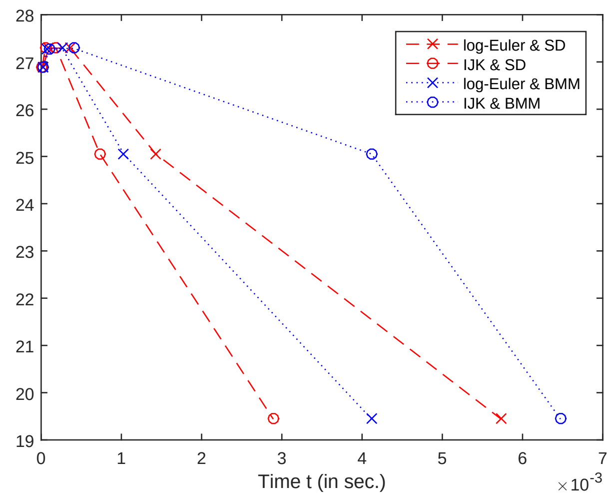
Step Δ	EM&SD($\theta = 0.5$)	IJK&SD($\theta = 0.5$)	EM&BMM($\Theta = 0.5$)	IJK& BMM($\Theta = 0.5$)
2^{-5}	26.901 (0.0000261)	26.901 (**0.0000159**)	26.891 (0.00002)	26.890 (0.0000294)
2^{-7}	27.288 (0.0000919)	27.288 (**0.0000492**)	27.277 (0.0000676)	27.277 (0.0001043)
2^{-9}	27.298 (0.0003595)	27.297 (**0.0001843**)	27.289 (0.0002610)	27.288 (0.0004081)
2^{-11}	25.057 (0.0014255)	25.058 (**0.0007309**)	25.051 (0.0010309)	25.051 (0.0016191)
2^{-13}	19.441 (0.0057322)	19.441 (**0.0028928**)	19.442 (0.0041177)	19.442 (0.0064721)

Tab. 3.8: 98%-Error, step size and average computational time of numerical integration of process (S_t) using log-Euler or IJK method with SD or BMM scheme for (3.1.1) with $x_0 = k_1 = \frac{1}{16}, k_2 = 1, k_3 = 0.4, S_0 = 100, \mu = 0.05, q = 3/4, T = 1$ and correlation $\rho = 0$.

Step Δ	EM&SD($\theta = 0.5$)	IJK&SD($\theta = 0.5$)	EM&BMM($\Theta = 0.5$)	IJK& BMM($\Theta = 0.5$)
2^{-5}	26.382 (0.0000266)	26.331 (**0.0000161**)	26.372 (0.0000202)	26.324 (0.00003)
2^{-7}	26.448 (0.0000951)	26.396 (**0.000005**)	26.439 (0.0000691)	26.389 (0.0001081)
2^{-9}	25.951 (0.0003631)	25.909 (**0.000184**)	25.944 (0.0002606)	25.904 (0.0004131)
2^{-11}	24.540 (0.0014506)	24.494 (**0.0007355**)	24.531 (0.0010378)	24.486 (0.0016495)
2^{-13}	18.738 (0.0060748)	18.749 (**0.0030185**)	18.735 (0.0042868)	18.747 (0.0068395)

Tab. 3.9: 98%-Error, step size and average computational time of numerical integration of process (S_t) using log-Euler or IJK method with SD or BMM scheme for (3.1.1) with $x_0 = k_1 = \frac{1}{16}, k_2 = 1, k_3 = 0.4, S_0 = 100, \mu = 0.05, q = 3/4, T = 1$ and correlation $\rho = -0.4$

Fig. 3.3: Strong convergence error of the financial underlying process (S_t), as a function of CPU time (in sec) using log-Euler or IJK method with SD or BMM scheme for (3.1.1) with $x_0 = k_1 = \frac{1}{16}, k_2 = 1, k_3 = 0.4, S_0 = 100, \mu = 0.05, q = 3/4, T = 1$ and correlation $\rho = 0$.

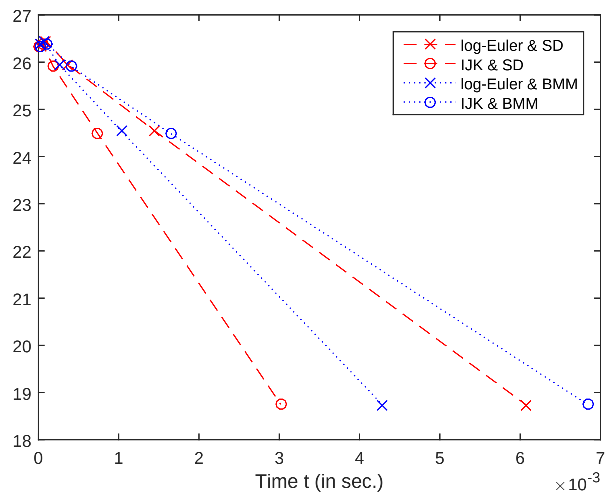


Figures 3.3, 3.4 and 3.5 indicate that in all cases the favorable choice is to integrate (S_t) using IJK method combined with the SD scheme for (V_t) in model (3.1.1). The IJK-SD approximation of system (3.1.1) seems to be the better one, w.r.t. CPU time, for every correlation coefficient considered.

3.8 Conclusion.

In this chapter, we exploit further the semi-discrete method (SD), originally appeared in [Hal12], to numerically approximate stochastic processes that appear in financial mathematics and are meant to be non-negative. In [HS16] we examined the Heston 3/2-model, that is a mean-reverting process with super-linear diffusion, described by a SDE of the form (3.2.1) with $q = 3/2$. Now, we deal with SDEs with sub-linear diffusion coefficients of the type $(x_t)^q$ with $1/2 < q < 1$. These kinds of SDEs, called mean-reverting CEV processes, appear in stochastic models, where they represent the instantaneous volatility-variance of an underlying financially observable variable. We prove theoretically the strong convergence of our proposed SD scheme,

Fig. 3.4: Strong convergence error of the financial underlying process (S_t), as a function of CPU time (in sec) using log-Euler or IJK method with SD or BMM scheme for (3.1.1) with $x_0 = k_1 = \frac{1}{16}, k_2 = 1, k_3 = 0.4, S_0 = 100, \mu = 0.05, q = 3/4, T = 1$ and correlation $\rho = -0.4$.



revealing the order of convergence. The resulting polynomial rate is shown in Theorem 3.2.2.We present a comparative study between various positivity preserving schemes and the SD method seems to be the best w.r.t. CPU time consumption. The advantage of the SD method here is that although implicit, has an explicit formula and thus requires fewer arithmetic operations and consequently less computational time. Moreover, our method can cover cases where (3.2.1) has time varying coefficients, i.e. $k_1(t), k_2(t), k_3(t)$.

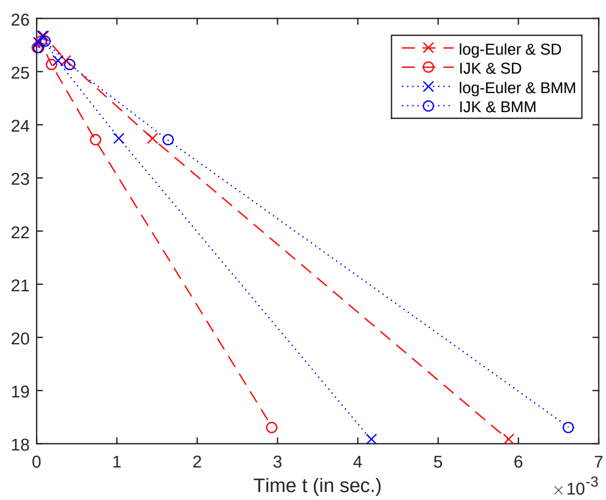
We also treat the two-dimensional stochastic volatility model (3.1.1). In order to do that, we actually integrate the process $\ln(S_t)$ which satisfies a SDE of the form (3.7.1) and in the end transform back for (S_t). We only consider two different schemes for the integration of $\ln(S_t)$, namely the Euler-Maruyama (EM) scheme, which is easy to implement and the IJK scheme [KJ06, (137)] which is shown to be the most efficient method, robust and simple as EM [KJ06]. We do not apply other two-dimensional schemes, such as for example the Milstein scheme, since they are in general time consuming, as they involve approximations of double Wiener integrals which require additional random number generation. We therefore combine the EM scheme with SD ((3.7.2) & (3.6.1)), the IJK scheme with SD ((3.7.3) & (3.6.1))

Step Δ	EM&SD($\theta = 0.5$)	IJK&SD($\theta = 0.5$)	EM&BMM($\Theta = 0.5$)	IJK& BMM($\Theta = 0.5$)
2^{-5}	25.552 (0.0000263)	25.455 (0.0000159)	25.541 (0.0000199)	25.449 (0.0000296)
2^{-7}	25.670 (0.0000932)	25.569 (0.0000494)	25.659 (0.0000678)	25.564 (0.0001059)
2^{-9}	25.217 (0.0003622)	25.137 (0.0001835)	25.208 (0.0002595)	25.132 (0.0004111)
2^{-11}	23.743 (0.0014407)	23.711 (0.0007306)	23.734 (0.0010307)	23.707 (0.0016376)
2^{-13}	18.082 (0.005871)	18.316 (0.0029312)	18.078 (0.0041637)	18.312 (0.0066239)

Tab. 3.10: 98%-Error, step size and average computational time of numerical integration of process (S_t) using log-Euler or IJK method with SD or BMM scheme for (3.1.1) with $x_0 = k_1 = \frac{1}{16}, k_2 = 1, k_3 = 0.4, S_0 = 100, \mu = 0.05, q = 3/4, T = 1$ and correlation $\rho = -0.8$.

and compare with the case where the stochastic variance $(p = \frac{1}{2})$ is integrated with BMM scheme (3.6.7), for three different correlation parameters, $\rho = 0, \rho = -0.4$ and $\rho = -0.8$ with $S_0 = 100, \mu = 0.05$, as in [KJ06, Sec. 5]. The combination IJK with SD seems to be the most favorable w.r.t. CPU time, for all the cases.

Fig. 3.5: Strong convergence error of the financial underlying process (S_t), as a function of CPU time (in sec) using log-Euler or IJK method with SD or BMM scheme for (3.1.1) with $x_0 = k_1 = \frac{1}{16}, k_2 = 1, k_3 = 0.4, S_0 = 100, \mu = 0.05, q = 3/4, T = 1$ and correlation $\rho = -0.8$.



4. CONSTANT DELAY DIFFERENTIAL EQUATIONS

4.1 Introduction.

We [1] assume the setting in Section 1.2, with $d = m = 1$, i.e. let $T > 0$ and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, \mathbb{P})$ be a complete probability space and let $W_{t,\omega} : [0,T] \times \Omega \to \mathbb{R}$ be a one-dimensional Wiener process adapted to the filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$. Consider the following stochastic delay differential equation (SDDE),

$$x_t = \begin{cases} \xi_0 + \int_0^t a(x_{s-\tau})x_s ds + \int_0^t b(x_{s-\tau})x_s dW_s, & t \in [0,T], \\ \xi(t), & t \in [-\tau, 0], \end{cases} \tag{4.1.1}$$

where the coefficients $a, b \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$,[2] the function $\xi \in \mathcal{C}([-\tau, 0], (0,\infty))$ and $\tau > 0$ is a positive constant which represents the delay.

[1] This chapter is based on unpublished work.

[2] $\mathcal{C}(A, B)$ the space of continuous functions $\phi : A \to B$ with norm $\|\phi\| = \sup_{u\in A} \phi(u)$.

SDDE (4.1.1) is called Delay Geometric Brownian Motion (DGBM), as in [MS13a], and is used in financial modeling. In that setting $b(\cdot)$ is called the volatility function and $\xi(\cdot)$ represents the initial data.

Consider the transformation $z(x,t) = \ln(e^{-rt}x) = -rt + \ln x$, $t \geq 0$, where $r > 0$.[3] Then Itô's formula implies

$$\begin{aligned} z_t &= z_0 + \int_0^t \left[-r + \frac{1}{x_s} a(x_{s-\tau}) x_s + \frac{1}{2}(-\frac{1}{x_s^2}) b^2(x_{s-\tau}) x_s^2 \right] ds + \int_0^t \frac{1}{x_s} b(x_{s-\tau}) x_s dW_s \\ &= \ln \xi_0 + \int_0^t \left(-r + a(x_{s-\tau}) - \frac{1}{2} b^2(x_{s-\tau}) \right) ds + \int_0^t b(x_{s-\tau}) dW_s. \end{aligned}$$

Therefore we introduce the function

$$\theta(x) = \frac{1}{2} b(x) + \frac{-r + a(x) - \frac{1}{2} b^2(x)}{b(x)} = \frac{a(x) - r}{b(x)},$$

which satisfies the Novikov[4] condition and the equivalent martingale measure $\widetilde{\mathbb{P}}$[5] with Radon-Nikodym derivative w.r.t. $\mathbb{P}$, restricted to the maturity-time σ-algebra, given by

$$\frac{d\widetilde{\mathbb{P}}}{d\mathbb{P}}\Big|_{\mathcal{F}_T} = \exp\left\{ -\frac{1}{2} \int_0^T \theta^2(x_s) ds - \int_0^T \theta(x_s) dW_s \right\}.$$

Under the above Girsanov transformation of measure [KS88, Sec. 3.5] we have

$$x_t = \begin{cases} \xi_0 + \int_0^t r x_s ds + \int_0^t b(x_{s-\tau}) x_s d\widetilde{W}_s, & t \in [0,T], \\ \xi(t), & t \in [-\tau, 0], \end{cases} \tag{4.1.2}$$

where $\widetilde{W}_t := W_t + \int_0^t \theta(x_s) ds$ lives in the space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \widetilde{\mathbb{P}})$.

SDDE (4.1.2) which describes the DGBM (4.1.1) has a unique global positive solution [MS13a, Th. 2.1] which can be computed, conditionally

[3] $z_t = \ln(e^{-rt} x_t)$ represents the log-price of the discounted asset and r the risk-free interest rate.

[4] An Itô integrable function $\theta(t,\omega) : [0,\infty) \times \Omega \to \mathbb{R}$, i.e. an $\mathcal{F}_t$-adapted measurable function with $\mathbb{E}\left(\int_0^T \theta^2(s,\omega) ds\right) < \infty$, satisfies the Novikov condition if $\mathbb{E}\left(\exp\{\int_0^T \theta^2(s,\omega) ds\}\right) < \infty$.

[5] $\widetilde{\mathbb{P}}$ is the risk-neutral measure.

on time-lagged information, but its unconditional distribution is not known when $t > \tau$. Thus, numerical approximations of (4.1.2) are necessary for simulations of the paths $x_t(\omega)$, as well as for approximation of functionals of the form $\mathbb{E}F(x)$, where $F : \mathcal{C}([0,T],\mathbb{R}) \to \mathbb{R}$ can be the expected payoff of an option. To simplify notation, in the following, we write (W_t) for $(\widetilde{W_t})$.

4.2 The setting and the main result.

Assumption 4.2.1 *Let the following conditions hold:*

(i) $b(\cdot)$ *is bounded, i.e. there is a* $C_b > 0$ *such that* $b(x) \leq C_b$ *for every* $x \geq 0$*;*

(ii) $b(\cdot)$ *is locally Lipschitz, that is*

$$|b(x) - b(y)| \leq C_R^b |x-y|,$$

for any $R > 0$ *such that* $|x| \vee |y| \leq R$*, where the constant* C_R^b *depends on* R*;*

(iii) $\xi(\cdot)$ *is Hölder continuous with order* γ*, where* $0 < \gamma \leq 1/2$*, that is*

$$\sup_{-\tau<u<v\leq 0} \frac{|\xi(v)-\xi(u)|}{(v-u)^\gamma} := C_\xi < \infty.$$

□

Let the observation time T be a multiple of τ, i.e. $T = N_0\tau$, where $N_0 \in \mathbb{N}$. We discretize the interval $[-\tau, T]$ with equidistant steps of size $\Delta = \tau/l$ for $l \in \{2, 3, \ldots\}$ and $t_n = n\Delta = n\tau/l$, where $n = -l, -l+1, \ldots, N$, i.e.

$$t_{-l} = -\tau < t_{-l+1} < \ldots < t_0 = 0 < \ldots < t_N = T,$$

with $N = l \cdot N_0$. Thus, $0 < \Delta < 1$.

The interpolation process of the Euler-Maruyama (EM) approximation proposed in [MS13a, Th. 6.2] reads

$$y_t^{EM} = \begin{cases} y_{t_n} + \int_{t_n}^t r y_{t_n} ds + \int_{t_n}^t b(y_{t_n-\tau}) y_{t_n} dW_s, & t \in [t_n, t_{n+1}], \\ \xi(t), & t \in [-\tau, 0], \end{cases}$$

Under Assumption 4.2.1 it holds [MS13a, Th. 6.2]

$$\lim_{\Delta\to 0}\mathbb{E}\sup_{0\le t\le T}|y_t^{EM}-x_t|^2=0.$$

A drawback of the above scheme is that there is a positive probability of taking negative values.

We propose the following semi-discrete numerical scheme

$$y_t^{SD}=\begin{cases} y_{t_n}+\int_{t_n}^t ry_s ds+\int_{t_n}^t b(y_{t_n-\tau})y_s dW_s, & t\in[t_n,t_{n+1}],\\ \xi(t), & t\in[-\tau,0],\end{cases} \tag{4.2.1}$$

or in compact form

$$y_t^{SD}=\begin{cases} \xi_0+\int_0^t ry_s ds+\int_0^t b(y_{\hat{s}-\tau})y_s dW_s, & t\in[0,T],\\ \xi(t), & t\in[-\tau,0],\end{cases} \tag{4.2.2}$$

where $\hat{s}=t_n$, when $s\in[t_n,t_{n+1})$. Thus, we discretize only the diffusion term of the SDDE (4.1.2), and to be more precise the volatility function $b(\cdot)$. We observe from (4.2.1) that in order to solve for y_t, we have to solve at each step an SDDE and not an algebraic equation. Note that we can reproduce the EM scheme if we fully discretize the drift and diffusion term of (4.1.2).

The linear SDDE (4.2.2) has an explicit solution

$$y_t^{SD}=\begin{cases} \xi_0\exp\left\{\int_0^t\left(r-b^2(y_{\hat{s}-\tau})/2\right)ds+\int_0^t b(y_{\hat{s}-\tau})dW_s\right\}, & t\in[0,T],\\ \xi(t), & t\in[-\tau,0].\end{cases}$$

The numerical scheme (4.2.2) converges to the true solution x_t of SDDE (4.1.2) and this is stated in the following, which is our main result.

Theorem 4.2.2 *Suppose Assumption 4.2.1 holds. Then the semi-discrete numerical scheme (4.2.2) converges to the true solution of (4.1.2) in the mean-square sense, that is*

$$\lim_{\Delta\to 0}\mathbb{E}\sup_{0\le t\le T}|y_t^{SD}-x_t|^2=0. \tag{4.2.3}$$

□

Section 4.3 concerns the proof of Theorem 4.2.2, which shows the strong convergence of the SD method.

Remark 4.2.3 *In Assumption 4.2.1(iii), we have assumed that* $0<\gamma\le 1/2$. *This is in accordance with the setting of [MS13a, Th. 6.2].* □

4.3 Strong Convergence of the SD scheme.

We denote the indicator function of a set A by $\mathbb{I}_A$. First we state moment bounds of the solution of the original process (x_t) of (4.1.2) and then prove moment bounds of the semi-discrete approximation (y_t^{SD}) given by (4.2.2). Later on, we provide an auxiliary result, that holds between the step process $(y_{\hat{t}}^{SD})$ and its continuous time approximation (y_t^{SD}) until the time of explosion θ_R. The process (y_t^{SD}) may also appear as (y_t).

4.3.1 Moment Bounds for original process and Semi-Discrete approximation.

Lemma 4.3.4 *[Moment bound of (x_t)] Let Assumption 4.2.1(i). The following bound is true*

$$\mathbb{E}(\sup_{0\leq t\leq T}(x_t)^p) < C_x(p),$$

for some $C_x(p) > 0$ and any $p \geq 1$ where

$$C_x(p) = 3\cdot 2^{p/2-1}(1+\|\xi\|^p)\exp\{p(2r\vee C_b + (33p-1)r^2\vee C_b^2)T\},$$

or

$$C_x(p) = \xi^p(0)\left(2+\frac{9p^2C_b^2}{p(r+0.5(p-1)C_b^2)}\right)\exp\{p(r+0.5(p-1)C_b^2)T\}.$$

□

Proof of Lemma 4.3.4. The first bound is given in [Mao97, Th. 5.4.1] and the sharper is as stated in [MS13a, Th. 2.4]. □

Lemma 4.3.5 *Let Assumptions 4.2.1(i) and 4.2.1(iii) hold. The following bound is true*

$$\mathbb{E}(\sup_{-\tau\leq t\leq T}(y_t^{SD})^p) < C_y(p),$$

for some $C_y(p) > 0$ and any $p > 2$. □

Proof of Lemma 4.3.5. Set the stopping time $\theta_R = \inf\{t \in [0,T] : y_t > R\}$, for some $R > \|\xi\| > 0$, with the convention that $\inf \emptyset = \infty$. Application of Itô's formula on $(y_{t\wedge\theta_R})^p$, implies,

$$\begin{aligned}
&(y_{t\wedge\theta_R})^p = (\xi_0)^p + \int_0^{t\wedge\theta_R} p(y_s)^{p-1} r y_s ds \\
&+ \int_0^{t\wedge\theta_R} \frac{p(p-1)}{2}(y_s)^{p-2}\left[b(y_{\hat{s}-\tau})y_s\right]^2 ds + \int_0^{t\wedge\theta_R} p(y_s)^{p-1} b(y_{\hat{s}-\tau}) y_s dW_s \\
= \; &(\xi_0)^p + \int_0^{t\wedge\theta_R} (pr + \frac{p(p-1)}{2} b^2(y_{\hat{s}-\tau}))(y_s)^p ds + \int_0^{t\wedge\theta_R} pb(y_{\hat{s}-\tau})(y_s)^p dW_s \\
\leq \; &(\xi_0)^p + \left(pr + \frac{p(p-1)}{2} C_b^2\right) \int_0^t (y_s)^p \mathbb{I}_{(0,t\wedge\theta_R)}(s) ds + M_t,
\end{aligned}$$

where the last inequality is valid for any $p > 2$ and

$$M_t := \int_0^{t\wedge\theta_R} pb(y_{\hat{s}-\tau})(y_s)^p dW_s.$$

Taking expectations and using that $\mathbb{E}M_t = 0$ we get

$$\begin{aligned}
\mathbb{E}(y_{t\wedge\theta_R})^p &\leq \mathbb{E}(\xi_0)^p + \left(pr + \frac{p(p-1)}{2} C_b^2\right) \int_0^t \mathbb{E}(y_{s\wedge\theta_R})^p ds \\
&\leq \mathbb{E}(\xi_0)^p e^{\left(pr + \frac{p(p-1)}{2} C_b^2\right)t},
\end{aligned}$$

where we have applied the Gronwall inequality. We have that

$$(y_{t\wedge\theta_R})^p = (y_{\theta_R})^p \mathbb{I}_{(\theta_R \leq t)} + (y_t)^p \mathbb{I}_{(t<\theta_R)} = R^p \mathbb{I}_{(\theta_R \leq t)} + (y_t)^p \mathbb{I}_{(t<\theta_R)},$$

thus taking expectations in the above inequality and using the estimated upper bound for $\mathbb{E}(y_{t\wedge\theta_R})^q$ we arrive at

$$\mathbb{E}(y_t)^p \mathbb{I}_{(t<\theta_R)} \leq \mathbb{E}(\xi_0)^p e^{\left(pr + \frac{p(p-1)}{2} C_b^2\right)t},$$

and taking limits in both sides as $R \to \infty$ we get that

$$\lim_{R\to\infty} \mathbb{E}(y_t)^p \mathbb{I}_{(t<\theta_R)} \leq \mathbb{E}(\xi_0)^p e^{\left(pr + \frac{p(p-1)}{2} C_b^2\right)t}.$$

Fix t. The sequence $(y_t)^p\mathbb{I}_{(t<\theta_R)}$ is non-decreasing in R since θ_R is increasing in R and $t\wedge\theta_R\to t$ as $R\to\infty$ and $(y_t)^p\mathbb{I}_{(t<\theta_R)}\to(y_t)^p$ as $R\to\infty$, thus the monotone convergence theorem, Theorem B.1.1, implies

$$\mathbb{E}(y_t)^p \leq \mathbb{E}(\xi_0)^p e^{\left(pr+\frac{p(p-1)}{2}C_b^2\right)t},$$

for any $p>2$. Thus for any $t_1\in[0,T]$, we have

$$\sup_{0\leq t\leq t_1}\mathbb{E}(y_t)^p \leq \mathbb{E}(\xi_0)^p e^{\left(pr+\frac{p(p-1)}{2}C_b^2\right)t_1},$$

which implies

$$\begin{aligned}\sup_{-\tau\leq t\leq t_1}\mathbb{E}(y_t)^p &= \sup_{-\tau\leq t\leq 0}\mathbb{E}|\xi_t|^p\bigvee\sup_{0\leq t\leq t_1}\mathbb{E}(y_t)^p\\ &\leq \|\xi\|^p\mathbb{E}(\xi_0)^p e^{\left(pr+\frac{p(p-1)}{2}C_b^2\right)t_1}.\end{aligned}$$

Using again Itô's formula on $(y_t)^p$, taking the supremum and then expectations we have that

$$\begin{aligned}&\mathbb{E}(\sup_{0\leq t\leq T}(y_t)^p) \leq \mathbb{E}(\xi_0)^p + \left(pr+\frac{p(p-1)}{2}C_b^2\right)\mathbb{E}\left(\sup_{0\leq t\leq T}\int_0^t(y_s)^p ds\right)\\ &+\mathbb{E}\sup_{0\leq t\leq T}M_t\\ \leq\ &\mathbb{E}(\xi_0)^p + \left(pr+\frac{p(p-1)}{2}C_b^2\right)\int_0^T\mathbb{E}(\sup_{0\leq l\leq s}(y_l)^p)ds + \sqrt{\mathbb{E}\sup_{0\leq t\leq T}M_t^2}\\ \leq\ &\left(\mathbb{E}(\xi_0)^p+\sqrt{4\mathbb{E}M_T^2}\right)e^{\left(pr+\frac{p(p-1)}{2}C_b^2\right)T},\end{aligned}$$

where in the last step we have used Doob's martingale inequality to the diffusion term M_t[6] and Gronwall's inequality. Thus

$$\mathbb{E}(\sup_{-\tau\leq t\leq T}(y_t^{SD})^p) \leq \left(\mathbb{E}(\xi_0)^p+\sqrt{4\mathbb{E}M_T^2}\right)e^{\left(pr+\frac{p(p-1)}{2}C_b^2\right)T}\bigvee\|\xi\|^p := C_y(p).$$

□

[6] The function $h(u)=pb(y_{\hat{u}-\tau})(y_u)^p$ belongs to the family $\mathcal{M}^2([0,T];\mathbb{R})$ thus [Mao97, Th. 1.5.8] implies $\mathbb{E}M_t^2=\mathbb{E}(\int_0^t h(u)dW_u)^2=\mathbb{E}\int_0^t h^2(u)du$, i.e. $M_t\in\mathcal{L}^2(\Omega;\mathbb{R})$.

4.3.2 Error Bound for the explicit Semi-Discrete scheme.

Lemma 4.3.6 *Let Assumption 4.2.1(i) hold and* $0 < |\xi\| < \infty$. *The following estimate holds*

$$\mathbb{E}|y_s^{SD} - y_{\widehat{s}}^{SD}|^p \leq C_p \Delta^{p/2},$$

where C_p *does not depend on* Δ, *implying* $\sup_{s\in[t_{n_s},t_{n_s+1}]} \mathbb{E}|y_s^{SD} - y_{\widehat{s}}^{SD}|^p = O(\Delta^{p/2})$ *as* $\Delta \downarrow 0$. □

Proof of Lemma 4.3.6. First, we fix a $p \geq 2$. Let the integer n_s be such that $s \in [t_{n_s}, t_{n_s+1})$. It holds that

$$\begin{aligned}
|y_s - y_{\widehat{s}}|^p &= \left| \int_{t_{n_s}}^{s} r y_u du + \int_{t_{n_s}}^{s} b(y_{\hat{u}-\tau}) y_u dW_u \right|^p \\
&\leq 2^{p-1} \left| \int_{t_{n_s}}^{s} r y_u du \right|^p + 2^{p-1} \left| \int_{t_{n_s}}^{s} b(y_{\hat{u}-\tau}) y_u dW_u \right|^p \\
&\leq 2^{p-1} (s - t_{n_s})^{p-1} r^p \int_{t_{n_s}}^{s} (y_u)^p du + 2^{p-1} \left| \int_{t_{n_s}}^{s} b(y_{\hat{u}-\tau}) y_u dW_u \right|^p,
\end{aligned}$$

where we have used the Cauchy-Schwarz inequality. Taking expectations in the above inequality gives

$$\begin{aligned}
\mathbb{E}|y_s - y_{\widehat{s}}|^p &\leq 2^{p-1} \Delta^{p-1} r^p \int_{t_{n_s}}^{s} \mathbb{E}(y_u)^p du \\
&\quad + 2^{p-1} \underbrace{\left(\frac{p^{p+1}}{2(p-1)^{p-1}} \right)^{p/2}}_{C_p^*} \mathbb{E} \left| \int_{t_{n_s}}^{t_{n_s+1}} b^2(y_{\hat{u}-\tau}) (y_u)^2 du \right|^{p/2} \\
&\leq 2^{p-1} r^p C_y(p) \Delta^p + 2^{p-1} C_p^* \Delta^{\frac{p-2}{2}} (C_b)^p \mathbb{E} \int_{t_{n_s}}^{t_{n_s+1}} |y_u|^p du \\
&\leq 2^{p-1} r^p C_y(p) \Delta^p + 2^{p-1} C_p^* C_y(p) \Delta^{p/2} \\
&\leq C\Delta^p + C\Delta^{p/2},
\end{aligned}$$

where we have used twice Hölder's inequality and the BDG inequality (B.3.5) on the diffusion term, Assumption 4.2.1(i) and the moment bounds of $(y_t)^p$ which are valid for any $2 \leq p$ by Lemma 4.3.5. Thus,

$$\lim_{\Delta \downarrow 0} \frac{\sup_{s\in[t_{n_s},t_{n_s+1}]} \mathbb{E}|y_s - y_{\widehat{s}}|^p}{\Delta^{p/2}} \leq C_p,$$

which justifies the $O(\Delta^{p/2})$ notation. Now for $0<p<2$ we have that

$$\mathbb{E}|y_s - y_{\hat{s}}|^p \leq \left(\mathbb{E}|y_s - y_{\hat{s}}|^2\right)^{p/2} \leq C_p\Delta^{p/2},$$

where we have used Jensen's inequality for the concave function $\phi(x) = x^{p/2}$. □

4.3.3 Convergence of the Semi-Discrete scheme in $\mathcal{L}^2$.

Set the stopping time $\theta_R = \inf\{t \in [0,T] : |y_t| > R \,\text{or}\, |x_t| > R\}$, for some $R>0$ big enough. We have that

$$\begin{aligned}
\mathbb{E}\sup_{0\leq t\leq T}|y_t - x_t|^2 &= \mathbb{E}\sup_{0\leq t\leq T}|y_t - x_t|^2\mathbb{I}_{(\theta_R>t)} + \mathbb{E}\sup_{0\leq t\leq T}|y_t - x_t|^2\mathbb{I}_{(\theta_R\leq t)}\\
&\leq \mathbb{E}\sup_{0\leq t\leq T}|y_{t\wedge\theta_R} - x_{t\wedge\theta_R}|^2 + \frac{2\delta}{p}\mathbb{E}\sup_{0\leq t\leq T}|y_t - x_t|^p\\
&\quad + \frac{(p-2)}{p\delta^{2/(p-2)}}\mathbb{P}(\theta_R\leq T),
\end{aligned}$$

where we have applied the Young inequality,

$$ab \leq \frac{\delta}{r}a^r + \frac{1}{q\delta^{q/r}}b^q,$$

for $a = \sup_{0\leq t\leq T}|y_t - x_t|^2, b = \mathbb{I}_{(\theta_R\leq t)}, r = p/2, q = p/(p-2)$ and $\delta>0$, or

$$\begin{aligned}
\mathbb{E}\sup_{0\leq t\leq T}|y_t - x_t|^2 &\leq \mathbb{E}\sup_{0\leq t\leq T}|y_{t\wedge\theta_R} - x_{t\wedge\theta_R}|^2 + \frac{2^p\delta}{p}\mathbb{E}\sup_{0\leq t\leq T}(|y_t|^p + |x_t|^p)\\
&\quad + \frac{(p-2)}{p\delta^{2/(p-2)}}\mathbb{P}(\theta_R\leq T)\\
&\leq \mathbb{E}\sup_{0\leq t\leq T}|y_{t\wedge\theta_R} - x_{t\wedge\theta_R}|^2\\
&\quad + \frac{2^p\delta(C_y(p)+C_x(p))}{p} + \frac{(p-2)}{p\delta^{2/(p-2)}}\mathbb{P}(\theta_R\leq T), \qquad (4.3.1)
\end{aligned}$$

where we have used the elementary inequality $(\sum_{i=1}^n a_i)^p \leq n^{p-1}\sum_{i=1}^n a_i^p$, with $n=2$, and $C_x(p), C_y(p)$ stand for the moment bounds of $(x_t), (y_t^{SD})$ given in Lemmata 4.3.4 and 4.3.5. It holds that

$$\begin{aligned}
\mathbb{P}(\theta_R\leq T) &\leq \mathbb{E}\left(\mathbb{I}_{(\theta_R\leq T)}\frac{|y_{\theta_R}|^p}{R^p}\right) + \mathbb{E}\left(\mathbb{I}_{(\theta_R\leq T)}\frac{|x_{\theta_R}|^p}{R^p}\right)\\
&\leq \frac{1}{R^p}\left(\mathbb{E}\sup_{0\leq t\leq T}|x_t|^p + \mathbb{E}\sup_{0\leq t\leq T}|y_t|^p\right) \leq \frac{C_x(p)\vee C_y(p)}{R^p},
\end{aligned}$$

thus (4.3.1) becomes

$$\mathbb{E}\sup_{0\le t\le T}|y_t - x_t|^2 \le \mathbb{E}\sup_{0\le t\le T}|y_{t\wedge\theta_R} - x_{t\wedge\theta_R}|^2 + \frac{2^p\delta C_x(p)\vee C_y(p)}{p} + \frac{2(p-2)C_x(p)\vee C_y(p)}{p\delta^{2/(p-2)}R^p}. \tag{4.3.2}$$

We estimate the difference $(e_{t\wedge\theta_R})^2 := |y_{t\wedge\theta_R} - x_{t\wedge\theta_R}|^2$. It holds that

$$\begin{aligned}(e_{t\wedge\theta_R})^2 &= \left|\int_0^{t\wedge\theta_R}(ry_s - rx_s)ds + \int_0^{t\wedge\theta_R}(b(y_{\hat{s}-\tau})y_s - b(x_{s-\tau})x_s)\,dW_s\right|^2\\ &\le 2T\int_0^{t\wedge\theta_R} r^2|y_s - x_s|^2 ds + 2\left|\int_0^{t\wedge\theta_R}(b(y_{\hat{s}-\tau})y_s - b(x_{s-\tau})x_s)\,dW_s\right|^2\\ &\le 2r^2T\int_0^{t\wedge\theta_R}(e_s)^2 ds + 2|M_t|^2,\end{aligned}$$

where in the second step we have used the Cauchy-Schwarz inequality and

$$M_t := \int_0^{t\wedge\theta_R}(b(y_{\hat{s}-\tau})y_s - b(x_{s-\tau})x_s)\,dW_s.$$

Taking the supremum over all $t \in [0,T]$ and then expectations we have

$$\begin{aligned}\mathbb{E}\sup_{0\le t\le T}(e_{t\wedge\theta_R})^2 &\le 2r^2T\int_0^T \mathbb{E}\sup_{0\le l\le s}(e_{l\wedge\theta_R})^2 ds + 2\mathbb{E}\sup_{0\le t\le T}|M_t|^2\\ &\le 2r^2T\int_0^T \mathbb{E}\sup_{0\le l\le s}(e_{l\wedge\theta_R})^2 ds + 8\mathbb{E}|M_T|^2,\end{aligned} \tag{4.3.3}$$

where in the last step we have used Hölder's inequality and Doob's martingale inequality with $p = 2$, since M_t is an $\mathbb{R}$-valued martingale that belongs to $\mathcal{L}^2$. It holds that

$$\begin{aligned}\mathbb{E}|M_T|^2 &:= \mathbb{E}\left|\int_0^{T\wedge\theta_R}(b(y_{\hat{s}-\tau})y_s - b(x_{s-\tau})x_s)\,dW_s\right|^2\\ &= \mathbb{E}\left(\int_0^{T\wedge\theta_R}(b(y_{\hat{s}-\tau})y_s - b(x_{s-\tau})x_s)^2\,ds\right)\\ &\le \mathbb{E}\left(\int_0^{T\wedge\theta_R}\left(2|b(y_{\hat{s}-\tau})|^2|y_s - x_s|^2 + 2|x_s|^2|b(y_{\hat{s}-\tau}) - b(x_{s-\tau})|^2\right)ds\right)\\ &\le 2C_b^2\int_0^{T\wedge\theta_R}\mathbb{E}\sup_{0\le l\le s}|e_{l\wedge\theta_R}|^2 ds + 2(C_R^b)^2\mathbb{E}\left(\int_0^{T\wedge\theta_R}(x_s)^2|y_{\hat{s}-\tau} - x_{s-\tau}|^2 ds\right),\end{aligned}$$

where we have used Assumption 4.2.1(ii) and C_R^b is the local Lipschitz constant of b or

$$\begin{aligned}\mathbb{E}|M_T|^2 \;\leq\; & 2C_b^2\int_0^{T\wedge\theta_R}\mathbb{E}\sup_{0\leq l\leq s}(e_{l\wedge\theta_R})^2 ds\\ & +4(C_R^b)^2\int_0^{T\wedge\theta_R}\sqrt{\mathbb{E}(x_s)^4}\sqrt{\mathbb{E}|y_{\hat{s}-\tau}-y_{s-\tau}|^4}ds\\ & +4(C_R^b)^2\mathbb{E}\left(\int_0^{T\wedge\theta_R}(x_s)^2|y_{s-\tau}-x_{s-\tau}|^2 ds\right), \qquad (4.3.4)\end{aligned}$$

The second integral of (4.3.4) is estimated by the following

$$\begin{aligned}&\int_0^{T\wedge\theta_R}\sqrt{\mathbb{E}(x_s)^4}\sqrt{\mathbb{E}|y_{\hat{s}-\tau}-y_{s-\tau}|^4}ds \leq \sqrt{C_x(4)}\int_0^{T}\sqrt{\mathbb{E}|y_{\hat{s}\wedge\theta_R-\tau}-y_{s\wedge\theta_R-\tau}|^4}ds\\ \leq&\sqrt{C_x(4)}\int_0^{T}\sqrt{\mathbb{E}|y_{(\hat{s}-\tau)\wedge\theta_R}-y_{(s-\tau)\wedge\theta_R}|^4}ds\\ \leq&\sqrt{C_x(4)}(\int_0^{\tau}\sqrt{|\xi(\hat{s}-\tau)-\xi(s-\tau)|^4}ds+\int_\tau^{T}\sqrt{\mathbb{E}|y_{(\hat{s}-\tau)\wedge\theta_R}-y_{(s-\tau)\wedge\theta_R}|^4}ds)\\ \leq&\sqrt{C_x(4)}\tau\sup_{-\tau<\hat{s}<s\leq 0}|\xi(\hat{s})-\xi(s)|^2+\sqrt{C_x(4)}\int_0^{T-\tau}\sqrt{\mathbb{E}|y_{\hat{s}\wedge\theta_R}-y_{s\wedge\theta_R}|^4}ds\\ \leq&\tau\sqrt{C_x(4)}(C_\xi)^{2\gamma}\Delta^{2\gamma}+\sqrt{C_x(4)}\sqrt{C_4}T\Delta,\end{aligned}$$

where in the last step we have used Assumption 4.2.1(iii) and Lemma 4.3.6, $C_x(4)$ as in Lemma 4.3.4, C_ξ is the Hölder constant of ξ and C_4 is as in Lemma 4.3.6. The last integral of (4.3.4) is estimated by the following

$$\begin{aligned}&\mathbb{E}\int_0^{T\wedge\theta_R}(x_s)^2|y_{s-\tau}-x_{s-\tau}|^2 ds\leq\mathbb{E}\int_{-\tau}^{T\wedge\theta_R-\tau}(x_s)^2|y_s-x_s|^2 ds\\ \leq\;&\mathbb{E}\int_{-\tau}^{0}(x_s)^2(e_s)^2ds+\mathbb{E}\int_0^{T\wedge\theta_R-\tau}(x_s)^2(e_s)^2 ds\\ \leq\;&\sup_{-\tau<\hat{s}<s\leq 0}|\xi(\hat{s})-\xi(s)|^2\mathbb{E}\int_{-\tau}^{0}(x_s)^2ds+R^2\int_0^T\mathbb{E}\sup_{0\leq l\leq s}(e_{l\wedge\theta_R})^2ds\\ \leq\;&\tau(C_\xi)^{2\gamma}C_x(2)\Delta^{2\gamma}+R^2\int_0^T\mathbb{E}\sup_{0\leq l\leq s}(e_{l\wedge\theta_R})^2ds,\end{aligned}$$

where in the last step we have used Assumption 4.2.1(iii), $C_x(2)$ is as in Lemma 4.3.4 and C_ξ is the Hölder constant of ξ. Relation (4.3.3) becomes

$$\begin{aligned}
&\mathbb{E}\sup_{0\le t\le T}(e_{t\wedge\theta_R})^2 \le 32\sqrt{C_x(4)}\sqrt{C_4}TC_b^2(C_R^b)^2\Delta \\
&+64\tau C_b^2(C_R^b)^2(C_\xi)^{2\gamma}(\sqrt{C_x(4)}\vee C_x(2))\Delta^{2\gamma} \\
&+(2r^2T+16C_b^2+32R^2(C_R^b)^2)\int_0^T \mathbb{E}\sup_{0\le l\le s}|y_{l\wedge\theta_R}-x_{l\wedge\theta_R}|^2 ds \\
\le\ & \widetilde{C}_3(C_R^b)^2\Delta^{2\gamma}+(\widetilde{C}_4+32R^2(C_R^b)^2)\int_0^T \mathbb{E}\sup_{0\le l\le s}(e_{l\wedge\theta_R})^2 ds,
\end{aligned}$$

where the constants $\widetilde{C}_3$ and $\widetilde{C}_4$ are given by

$$\widetilde{C}_3 := 64\tau C_b^2(C_\xi)^{2\gamma}(\sqrt{C_x(4)}\vee C_x(2)), \quad \widetilde{C}_4 := 2r^2T+16C_b^2.$$

Application of the Gronwall inequality implies

$$\mathbb{E}\sup_{0\le t\le T}(e_{t\wedge\theta_R})^2 \le \widetilde{C}_3(C_R^b)^2\Delta^{2\gamma}e^{(\widetilde{C}_4+32R^2(C_R^b)^2)T}.$$

Relation (4.3.2) becomes,

$$\begin{aligned}
\mathbb{E}\sup_{0\le t\le T}(e_t)^2 \ \le\ & \widetilde{C}_3(C_R^b)^2\Delta^{2\gamma}e^{(\widetilde{C}_4+32R^2(C_R^b)^2)T}+\frac{2^p\delta C_x(p)\vee C_y(p)}{p} \\
&+\frac{2(p-2)C_x(p)\vee C_y(p)}{p\delta^{2/(p-2)}R^p} := I_1+I_2+I_3.
\end{aligned}$$

Given any $\epsilon>0$, we may first choose δ such that $I_2<\epsilon/3$, then choose R such that $I_3<\epsilon/3$, and finally Δ such that $I_1<\epsilon/3$ concluding $\mathbb{E}\sup_{0\le t\le T}(e_t)^2<\epsilon$ as required to verify (4.2.3).

4.4 Comments and future work.

- Assumption 4.2.1(i), for the boundedness of the volatility function $b(\cdot)$ does not seem unrealistic as shown in [BBF02, (7)], [CLV99] where the 'local' volatility function is approximated with splines and [DFW98, (6)-(9)], where the Deterministic Volatility Function (DVF) hypotheses is made and the following four models are compared

$$\begin{array}{llcl}
\text{(Model 0)} & b(t,x) & = & 0.01 \bigvee a_0 \\
\text{(Model 1)} & b(t,x) & = & 0.01 \bigvee (a_0 + a_1 x + a_2 x^2) \\
\text{(Model 2)} & b(t,x) & = & 0.01 \bigvee (a_0 + a_1 x + a_2 x^2 + a_3 t + a_5 xt) \\
\text{(Model 3)} & b(t,x) & = & 0.01 \bigvee (a_0 + a_1 x + a_2 x^2 + a_3 t + a_4 t^2 + a_5 xt)
\end{array}$$

Model 0 is the Black-Scholes case. The rest of the models are in quadratic form and vary only with the index level (Model 1), or depend also on time. Time variation seems important and specifically the cross-product term xt provides better explanatory power [DFW98, p. 2072].

The above-mentioned (DVF) approach is a simple way to explain the volatility 'smile' preserving the arbitrage argument. Between other attempts in that direction, are the stochastic volatility models of Hull-White [HW87], where the price of a call option on a security (with price S_t) is derived

$$\begin{cases} S_t = S_0 + \int_0^t a(s, S_u, b_u) S_u du + \int_0^t b_u S_u dW_u, & t \in [0,T], \\ \\ V_t = V_0 + \int_0^t \mu(u, V_u) V_u du + \int_0^t \xi(u, V_u) V_u d\widetilde{W}_u & t \in [0,T], \end{cases}$$

where $V_t = (b_t)^2$ is the instantaneous variance and the Wiener processes $W_t, \widetilde{W_t}$ have correlation ρ, and Heston model [Hes93, (1)-(2)]

$$\begin{cases} S_t = S_0 + \int_0^t a \cdot S_u du + \int_0^t b_u S_u dW_u, & t \in [0,T], \\ \\ b_t = b_0 - \int_0^t \mu b_u du + \int_0^t \xi d\widetilde{W}_u & t \in [0,T], \end{cases}$$

where the Wiener processes $W_t, \widetilde{W_t}$ have correlation ρ. Thus, the volatility process b_t follows an Ornstein-Uhlenbeck process and Itô's lemma shows that the instantaneous variance $V_t = (b_t)^2$ satisfies the following square-root process

$$V_t = V_0 + \int_0^t (\xi^2 - 2\mu V_s) ds + \int_0^t 2\xi \sqrt{V_s} d\widetilde{W}_s, \quad t \in [0,T],$$

which has been studied lately in [Hal15d].

- In [MS13a], where the DGBM model was proposed, the delay effect on various options is studied, as in European options and barrier options. In particularly under Assumption 4.2.1(ii), it is shown that it is robust, in the sense that if we consider another time lag $\widetilde{\tau}$ and the corresponding SDDE

$$\widetilde{x}_t = \begin{cases} \xi_0 + \int_0^t r\widetilde{x}_s ds + \int_0^t b(\widetilde{x}_{s-\widetilde{\tau}})\widetilde{x}_s dW_s, & t \in [0,T], \\ \xi(t), & t \in [-\widetilde{\tau}, 0], \end{cases}$$

then for the price of a European call option $C_\tau := e^{-rT}\mathbb{E}(x_T - K)^+$, with exercise price K and expiry date T, or for an up-and-out call option $B_\tau := e^{-rT}\mathbb{E}(x_T - K)^+\mathbb{I}_{\{0 \le x_t \le B,\, 0 \le t \le T\}}$ with given barrier B, the following continuity property holds

$$\lim_{\tau - \widetilde{\tau} \to 0} |C_\tau - C_{\widetilde{\tau}}| = 0, \qquad \lim_{\tau - \widetilde{\tau} \to 0} |B_\tau - B_{\widetilde{\tau}}| = 0.$$

Assuming thus that the evolution of the asset price is described by the DGBM model, which has the above robust property and having established the strong convergence of SD method in Theorem 4.2.2 , we can now use our positivity preserving numerical scheme to evaluate the expected payoff of the above-mentioned options. In that way the European call option and the barrier option can be approximated by

$$\widetilde{C}_\tau := e^{-rT}\mathbb{E}(y_T - K)^+ \tag{4.4.1}$$

and

$$\widetilde{B}_\tau := e^{-rT}\mathbb{E}(y_T - K)^+\mathbb{I}_{\{0 \le y_t \le B,\, 0 \le t \le T\}} \tag{4.4.2}$$

accordingly, where y_t is given by (4.2.2). In particular, the following result holds.

Proposition 4.4.7 *In the framework of the DGBM model, and under Assumption 4.2.1, the following approximations are true*

$$\lim_{\Delta \to 0} |C_\tau - \widetilde{C}_\tau| = 0, \qquad \lim_{\Delta \to 0} |B_\tau - \widetilde{B}_\tau| = 0, \tag{4.4.3}$$

where y_t is the SD method proposed in (4.2.2) and C_τ, B_τ are the expected payoffs of the European call option and up-and-out call option *given as the proposed approximations (4.4.1) and (4.4.2) but with the process* (x_t) *in place of* (y_t). □

For sake of completeness we highlight the proof in Appendix E, where we follow the ideas of [HM05].

- We may consider a variable interest rate, i.e. the transformation

$$z(x,t) = \ln(e^{-\int_0^t r(s)ds}x) = -\int_0^t r(s)ds + \ln x,\ t \geq 0.$$

and/or a variable delay setting, i.e. the SDDE

$$x_t = \begin{cases} \xi_0 + \int_0^t a(x_{\delta_1(s)})x_s ds + \int_0^t b(x_{\delta_2(s)})x_s dW_s, & t \in [0,T], \\ \xi(t), & t \in [-L,0], \end{cases}$$

where $\delta_1(\cdot), \delta_2(\cdot)$ are $\mathcal{F}_0$-measurable functions with $\delta_i(t) \leq n\tau$, when $t \in [n\tau \wedge T, (n+1)\tau \wedge T)$, for $n \in \mathbb{N} \cup \{0\}$, and $\tau \in (0,T], i = 1,2$ and $L > 0$ represents the level of past data available on ξ. The case $\delta_i(t) = t - \tau$, describes constant delay models whereas the case $\delta_i(t) = \left\lfloor \frac{t}{\tau} \right\rfloor \tau$, variable step-function delay models [AHMP07, Model (22)].

- We may consider the Kuchler-Platen setting, i.e. the SDDE

$$x_t = \begin{cases} \xi_0 + \int_0^t a(s, x_s, x_{s-\tau})ds + \int_0^t b(s, x_s, x_{s-\tau})dW_s, & t \in [0,T], \\ \xi(t), & t \in [-\tau,0], \end{cases}$$

as was introduced in [KP00].

- We can also work on other SDDEs with non-negative solution as (4.4.4)

$$x_t = \begin{cases} \xi_0 + \int_0^t \left(a(x_s)^\beta + b(x_{s-\tau})^\beta\right) ds + \int_0^t \sigma(s)dW_s, & t \in [0,T], \\ \xi(t), & t \in [-\tau,0], \end{cases}$$

where $\xi \in (\mathcal{C}[-\tau,0],\mathbb{R})$. Problem (4.4.4) has solution with $x_t \geq 0$ a.s when $a > |b| \geq 0, \sigma \in (\mathcal{C}[0,\infty),\mathbb{R})$ and $\beta > 1$ is a quotient of odd integers.

5. STRUCTURE PRESERVING NUMERICAL SCHEME IN MOLECULAR DYNAMICS

5.1 Introduction.

We [1] assume the setting in Section 1.2, with $d = m = 1$, i.e. let $T > 0$ and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\leq t\leq T}, \mathbb{P})$ be a complete probability space and let $W_{t,\omega} : [0,T]\times\Omega \to \mathbb{R}$ be a one-dimensional Wiener process adapted to the filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$. We are interested in the numerical approximation of the following scalar stochastic differential equation (SDE),

$$x_t = x_0 + \int_0^t \left(\phi(x_s)\sqrt{1-x_s^2} - \frac{c^2}{2}x_s\right) ds + c\int_0^t \sqrt{1-x_s^2}dW_s, \tag{5.1.1}$$

where $c > 0$ is a positive constant and $\phi(\cdot)$ is a bounded and Lipschitz continuous function with bounding constant K_ϕ and Lipschitz constant L_ϕ. A boundary classification result, see Appendix F.1, implies that $-1 < x_t < 1$

[1] This chapter is based on joint (unpublished) work with Mark Peletier and Upanshu Sharma during my Erasmus visit at the Department of Mathematics and Computer Science, Technische Universiteit Eindhoven. I would like to thank Erasmus for the financial support.

a.s. when $x_0 \in (-1,1)$. We therefore aim for a numerical scheme which apart from strongly converging to the true solution of (5.1.1), produces values in the same domain, i.e. $(-1,1)$. In other words, we are interested in numerical schemes that have an *eternal life time*, see Definition 1.3.7 and the equivalent statement (1.3.1), which we repeat here:

Definition 5.1.1 *[Eternal Life time of numerical solution] Let $D \subseteq \mathbb{R}^d$ and consider a process (X_t) well defined on the domain $\overline{D}$, with initial condition $X_0 \in \overline{D}$ and such that*

$$\mathbb{P}(\{\omega \in \Omega : X(t,\omega) \in \overline{D}\}) = 1,$$

for all $t > 0$. A numerical solution $(Y_{t_n})_{n\in\mathbb{N}}$ has an eternal life time *if*

$$\mathbb{P}(Y_{n+1} \in \overline{D} \,\big|\, Y_n \in \overline{D}) = 1.$$

□

In [Sch96] the main interest is in the domain $D = \mathbb{R}^+$. Moreover, it is clear that the Euler-Maruyama scheme has always a finite life time, see e.g. [Kah04, Prop. 4.2].

The proposed SD iterative scheme for the numerical approximation of (5.1.1) reads

$$y^{SD}_{t_{n+1}} = \cos\left(-c\Delta W_n + \arccos(y_{t_n} + \phi(y_{t_n})\sqrt{1 - y^2_{t_n}} \cdot \Delta)\right), \tag{5.1.2}$$

where $\Delta W_n := W_{t_{n+1}} - W_{t_n}$, are the increments of the Wiener process and the discretization step Δ is such that (5.1.2) is well-defined. By construction, the SD scheme (5.1.2) possesses an eternal life time.

An attempt in that direction, i.e. in constructing explicit numerical schemes with an eternal life time, has been made in [Hal14]. The domain of the original SDE is $\mathbb{R}^+$ and for a class of super-linear problems, a positivity preserving scheme is suggested that is strongly convergent in the mean-square sense. In Chapter 2 another class of one-dimensional SDEs with non-negative solutions is treated, which covers cases like that of the Heston 3/2-model, a popular model in the field of financial mathematics which is also super-linear. The case of sub-linearities covered in Chapter 3 is also treated in [Hal15d] and [HS15] where the domain is still $\mathbb{R}^+$.

The purpose of this chapter is to generalize further the method to preserve the structure of the original SDE. In the previous chapters, the suggested schemes preserve positivity; all the quantities appearing belong to the field of finance and are meant to be non-negative. The application that motivated us now, comes from physics, and it is a model about the evolution of the trigonometric function *cosine* of an angle between 3 atoms of the form (5.1.1). We are able in that case to preserve the domain of the original process, by applying appropriately the SD method: we use an additive semi-discretization of the drift coefficient. The SD method is problem dependent, so there is not a unique way to treat all models.

The semi-discrete method was originally proposed in [Hal12]. Briefly saying, a part of the SDE is discretized in a certain way such that the resulting SDE to be solved has an analytical solution (see details in Section 5.2). This is also a special feature of the method, since in the derivation of it, instead of an algebraic equation a new SDE has to be solved. The SD method can also reproduce the Euler scheme.

In Section 5.2 we provide the setting and the main goal which concerns the mean-square convergence of the proposed structure-preserving SD scheme (5.1.2) for the approximation of a modification of (5.1.1) with dynamics described by $\widehat{W}$ (see (5.2.5)). We chose the dynamics (5.1.1), because that naturally arose from our application, which was the main motivation for us. Therefore, Section 5.3 is devoted to the description of the particular form of SDE that we study, and which was implied by a specific model, from the field of molecular dynamics, called a 3-atom model [LL10, Sec. 4.2].

In Section 5.4 some results are stated and proved concerning the proposed method. Finally, in Section 5.5 we make some numerical experiments.

5.2 The setting and the main goal.

Consider the partition $0 = t_0 < t_1 < \ldots < t_N = T$ with uniform discretization step $\Delta = T/N$ and the following process

(5.2.1)
$$y_t^{SD} = y_{t_n} + \int_{t_n}^{t_{n+1}} \phi(y_{t_n})\sqrt{1-y_{t_n}^2}ds + \int_{t_n}^{t} \frac{-c^2}{2} y_s ds + c\int_{t_n}^{t} \sqrt{1-y_s^2}\,\mathrm{sgn}(z_s)dW_s,$$

for $t \in (t_n, t_{n+1}]$, with $y_0 = x_0$ a.s. and

$$z_t = \sin\left(-c\Delta W_n^t + \arccos\left(y_{t_n} + \phi(y_{t_n})\sqrt{1-y_{t_n}^2}\cdot\Delta\right)\right), \tag{5.2.2}$$

where $\Delta W_n^t := W_t - W_{t_n}$. Process (5.2.1) has jumps at nodes t_n and the solution in each step is given by, see Appendix F.2,

(5.2.3)
$$y_t^{SD} = \cos\left(-c\Delta W_n^t + \arccos(y_{t_n} + \phi(y_{t_n})\sqrt{1-y_{t_n}^2}\cdot\Delta)\right),\ t\in(t_n, t_{n+1}],$$

which has the pleasant feature that $y_t^{SD} \in [-1,1]$. Process (5.2.3) is well defined when

$$\left|y_{t_n} + \phi(y_{t_n})\sqrt{1-y_{t_n}^2}\cdot\Delta\right| \leq 1. \tag{5.2.4}$$

Therefore, we assume the following condition for the well-posedness of the SD scheme (5.2.3).

Assumption 5.2.2 *Let the discretization step Δ be such that (5.2.4) holds.* □

Remark 5.2.3 *Note that in general the discretization step Δ satisfying (5.2.4) is a r.v. depending on ω. The ω-dependence is inherited through the increments $\Delta W_n(\omega)$ which in turn affect the sequence $(y_{t_n})_{n\in\mathbb{N}}$. Nevertheless, in the application considered in Section 5.5, the step Δ is not a r.v. but a fixed sufficiently small number.* □

Now, we work as in Section 3.2, i.e. we consider the process

$$\widehat{W}_t := \int_0^t \operatorname{sgn}(z_s)dW_s, \tag{5.2.5}$$

which is a martingale with quadratic variation $<\widehat{W}_t, \widehat{W}_t>= t$ and thus a standard Brownian motion w.r.t. its own filtration, justified by Lévy's theorem (see Theorem A.3.9) and consequently (5.2.1) becomes

$$y_t^{SD} = y_{t_n} + \phi(y_{t_n})\sqrt{1-y_{t_n}^2}\cdot\Delta + \int_{t_n}^t(-\frac{c^2}{2})y_s ds + c\int_{t_n}^t \sqrt{1-y_s^2}d\widehat{W}_s, \tag{5.2.6}$$

Moreover, consider

$$\widehat{x}_t = x_0 + \int_0^t \left(\phi(\widehat{x}_s)\sqrt{1-\widehat{x}_s^2} - \frac{c^2}{2}\widehat{x}_s\right)ds + c\int_0^t \sqrt{1-\widehat{x}_s^2}d\widehat{W}_s. \tag{5.2.7}$$

The process (x_t) of (5.1.1) and the process $(\widehat{x}_t)$ of (5.2.7) have the same distribution. Our main goal is to deduce an estimate of the form

$$\lim_{\Delta\to 0}\mathbb{E}\sup_{0\le t\le T}|y_t^{SD}-\widehat{x}_t|^2=0.$$

We are not able to achieve that result and restrict ourselves to a numerical application of the method. To simplify notation we write $\widehat{W}, (\widehat{x}_t)$ as $W, (x_t)$.

Theorem 5.2.4 *[Polynomial rate of convergence] Let Assumption 5.2.2 hold. Then, the semi-discrete scheme (5.2.3) converges in the mean-square sense to the true solution of (5.2.7), that is*

$$\lim_{\Delta\to 0}\mathbb{E}\sup_{0\le t\le T}|y_t^{SD}-x_t|^2=0$$

□

Remark 5.2.5 *The main convergence result given in Theorem 5.2.4 is still true in the case where $\phi(\cdot)$ is γ-Hölder continuous function with Hölder exponent $1/2\le\gamma\le 1$ and Hölder constant L_ϕ. (The case $\gamma=1$ corresponds to the Lipschitz continuous case treated here.) The main difference now is that we first need an $\mathcal{L}^1$-estimate. This can be done following the Yamada-Watanabe approach as in Section 2.3.2 or Section 3.3.2.* □

5.3 Transformation of the 3-atom model.

In this section we give the details of the 3-atom model from [LL10, Sec. 4.2]. Consider 3 atoms in $\mathbb{R}^2$, whose positions are given by $A=(a_1,a_2)$, $B=(b_1,b_2)$, $C=(c_1,c_2)$. We define $X^0=(a_2-b_2)$, $X^1=(a_1-b_1)$, $X^2=(c_1-b_1)$, $X^3=(c_2-b_2)$. The state space of the dynamics is $\mathbb{R}^4\ni X^T=(X^0,X^1,X^2,X^3)$. The full dynamics is described by the following SDE,

$$dX_t=-\nabla V^\varepsilon(X_t)dt+\sqrt{\frac{2}{\beta}}dW_t,$$

in differential form,[2] where the potential $V^\varepsilon:\mathbb{R}^4\to\mathbb{R}$ depends on $\varepsilon>0$, the initial condition $X_{t=0}=X_0$ is independent of ε and W_t denotes a 4-d

[2] These dynamics are called overdamped Langevin.

Brownian motion. The potential V^ε is explicitly given by

$$\begin{aligned} V^\varepsilon(X) &= \frac{1}{2\varepsilon}(|\underbrace{\sqrt{(X^0)^2+(X^1)^2}-l_{eq}}_{q_{AB}}|^2+|\underbrace{\sqrt{(X^2)^2+(X^3)^2}-l_{eq}}_{q_{BC}}|^2) \\ &+\mathcal{W}(\theta_{ABC}(X)), \end{aligned}$$

where l_{eq} is a given equilibrium distance and θ_{ABC} is the angle given by the relation

$$\theta_{ABC}(X) = \arccos\left(\frac{X^1X^2+X^0X^3}{\sqrt{(X^0)^2+(X^1)^2}\sqrt{(X^2)^2+(X^3)^2}}\right).$$

Further the function $\mathcal{W}:\mathbb{R}\to\mathbb{R}$, which is in this case a three-body potential, is a double-well potential given by

$$\mathcal{W}\circ\theta = \frac{k_\theta}{2}\left((\theta-\theta_m)^2-\delta\theta^2\right)^2,$$

where the wells of $\mathcal{W}$ are located at $\theta=\theta_m\pm\delta\theta$. Note that $\nabla V^\varepsilon:\mathbb{R}^4\to\mathbb{R}^4$ is explicitly given by $\frac{1}{\varepsilon}q_{AB}\nabla q_{AB}+\frac{1}{\varepsilon}q_{BC}\nabla q_{BC}+\nabla\mathcal{W}\circ\theta=$

$$\nabla V^\varepsilon = \begin{pmatrix} \frac{1}{\varepsilon}\left(\sqrt{(X^0)^2+(X^1)^2}-l_{eq}\right)\frac{X^0}{\sqrt{(X^0)^2+(X^1)^2}}+\frac{\partial\mathcal{W}}{\partial X^0} \\ \frac{1}{\varepsilon}\left(\sqrt{(X^0)^2+(X^1)^2}-l_{eq}\right)\frac{X^1}{\sqrt{(X^0)^2+(X^1)^2}}+\frac{\partial\mathcal{W}}{\partial X^1} \\ \frac{1}{\varepsilon}\left(\sqrt{(X^2)^2+(X^3)^2}-l_{eq}\right)\frac{X^2}{\sqrt{(X^2)^2+(X^3)^2}}+\frac{\partial\mathcal{W}}{\partial X^2} \\ \frac{1}{\varepsilon}\left(\sqrt{(X^2)^2+(X^3)^2}-l_{eq}\right)\frac{X^3}{\sqrt{(X^2)^2+(X^3)^2}}+\frac{\partial\mathcal{W}}{\partial X^3} \end{pmatrix}.$$

As suggested in [LL10, Sec. 4.2], we simplify the computations by making the system invariant under rigid body motion, which can be accomplished by setting $B=(b_1,b_2)$ to $(0,0)$ and $A\cdot e_y=a_2$ to 0 which implies $X^0=0$.

Thus, simplifying according to the suggestion above, we consider $\mathbb{R}^3\ni$

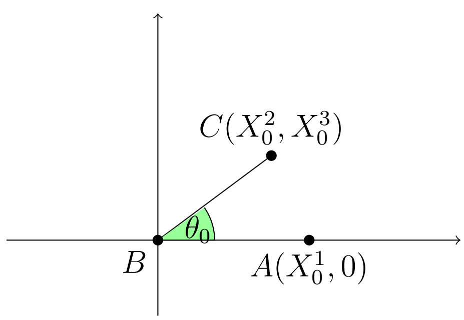
$X^T = (X^1, X^2, X^3)$ where now $\nabla V^\varepsilon : \mathbb{R}^3 \to \mathbb{R}^3$ is given by

$$\nabla V^\varepsilon = \begin{pmatrix} \frac{1}{\varepsilon}\left(\frac{|X^1| - l_{eq}}{|X^1|}\right) X^1 + \frac{\partial \mathcal{W}}{\partial X^1} \\ \frac{1}{\varepsilon}\left(\frac{\sqrt{(X^2)^2 + (X^3)^2} - l_{eq}}{\sqrt{(X^2)^2 + (X^3)^2}}\right) X^2 + \frac{\partial \mathcal{W}}{\partial X^2} \\ \frac{1}{\varepsilon}\left(\frac{\sqrt{(X^2)^2 + (X^3)^2} - l_{eq}}{\sqrt{(X^2)^2 + (X^3)^2}}\right) X^3 + \frac{\partial \mathcal{W}}{\partial X^3} \end{pmatrix}.$$

and

$$\theta_{ABC} = \arccos\left(\frac{X^1 X^2}{|X^1|\sqrt{(X^2)^2 + (X^3)^2}}\right).$$

Below is a plot of the original configuration of the 3-atom model.



Moreover $\frac{\partial \theta}{\partial X^1} = 0$, thus the corresponding SDE for X^1 can be solved independently (numerically or analytically if we know that $X^1 > 0$ a.s. or $X^1 < 0$ a.s.) and has the stochastic integral representation

$$X_t^1 = X_0^1 + \int_0^t \frac{1}{\varepsilon}\left(\frac{l_{eq} - |X_s^1|}{|X_s^1|}\right) X_s^1 ds + \sqrt{\frac{2}{\beta}} W_t^1, \tag{5.3.1}$$

where $X^1_0 = a_1$. For the other two components we can write

$$X^2_t = X^2_0 + \int_0^t \frac{1}{\varepsilon}\frac{l_{eq} - \sqrt{(X^2_s)^2 + (X^3_s)^2}}{\sqrt{(X^2_s)^2 + (X^3_s)^2}} X^2_s ds + \int_0^t f(\theta)\frac{X^1_s}{|X^1_s|}\frac{X^3_s}{|X^3_s|}\frac{X^3_s}{(X^2_s)^2 + (X^3_s)^2} ds + \sqrt{\frac{2}{\beta}} W^2_t \tag{5.3.2}$$

$$X^3_t = X^3_0 + \int_0^t \frac{1}{\varepsilon}\frac{l_{eq} - \sqrt{(X^2_s)^2 + (X^3_s)^2}}{\sqrt{(X^2_s)^2 + (X^3_s)^2}} X^3_s ds - \int_0^t f(\theta)\frac{X^1_s}{|X^1_s|}\frac{X^3_s}{|X^3_s|}\frac{X^2_s}{(X^2_s)^2 + (X^3_s)^2} ds + \sqrt{\frac{2}{\beta}} W^3_t, \tag{5.3.3}$$

where the third-order polynomial $f = \mathcal{W}'$ is given by

$$f(\theta) = 2k_\theta \cdot \theta^3 - 6k_\theta\theta_m \cdot \theta^2 + k_\theta(4\theta_m^2 + 2\widetilde{\theta}) \cdot \theta - 2k_\theta\theta_m\widetilde{\theta}, \quad \widetilde{\theta} = 2\theta_m\theta_0 - \theta_0^2.$$

5.3.1 *Scalar-valued coarse-graining map.*

We consider the transformation $\xi(X_t) = \cos\theta_t$, that is $\xi(X) = \mathrm{sgn}(X^1)X^2[(X^2)^2 + (X^3)^2]^{-1/2}$. Itô's formula implies

$$\xi(X_t) = \xi(X_0) + \int_0^t \left(-(\nabla\xi)(\nabla V^\epsilon)(X_s) + \frac{1}{\beta}\Delta\xi(X_s)\right) ds + \sqrt{\frac{2}{\beta}}\int_0^t |\nabla\xi(X_s)| d\widetilde{W}_s, \tag{5.3.4}$$

where

$$\widetilde{W}_t = \int_0^t \frac{\nabla\xi}{|\nabla\xi|}(X_s) dW_s, \tag{5.3.5}$$

is a scalar Brownian motion. We have that

$$\begin{aligned} |\nabla\xi(X)|^2 &= \left(\frac{\partial\xi}{\partial X^1}\right)^2 + \left(\frac{\partial\xi}{\partial X^2}\right)^2 + \left(\frac{\partial\xi}{\partial X^3}\right)^2 \\ &= 0 + \left(\frac{X^1}{|X^1|}\left(\frac{1}{\sqrt{(X^2)^2 + (X^3)^2}} - \frac{(X^2)^2}{((X^2)^2 + (X^3)^2)^{3/2}}\right)\right)^2 \\ &\quad + \left(-\frac{X^1}{|X^1|}\frac{X^2X^3}{((X^2)^2 + (X^3)^2)^{3/2}}\right)^2 \\ &= \frac{(X^3)^2}{((X^2)^2 + (X^3)^2)^2}, \end{aligned}$$

Moreover, $\frac{\partial^2 \cos\theta}{\partial (X^1)^2} = 0$,

$$\begin{aligned}\frac{\partial^2 \cos\theta}{\partial (X^2)^2} &= \frac{\partial}{\partial X^2}\left(\frac{X^1}{|X^1|}\frac{1}{\sqrt{(X^2)^2+(X^3)^2}}\frac{(X^3)^2}{(X^2)^2+(X^3)^2}\right)\\ &= -3\mathrm{sgn}(X^1)X^2(X^3)^2\left((X^2)^2+(X^3)^2\right)^{-5/2}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial^2 \cos\theta}{\partial (X^3)^2} &= \frac{\partial}{\partial X^3}\left(-\frac{X^1}{|X^1|}\frac{1}{\sqrt{(X^2)^2+(X^3)^2}}\frac{X^2X^3}{(X^2)^2+(X^3)^2}\right)\\ &= -\mathrm{sgn}(X^1)X^2[\left((X^2)^2+(X^3)^2\right)^{-3/2} - 3(X^3)^2\left((X^2)^2+(X^3)^2\right)^{-5/2}]\\ &= -\mathrm{sgn}(X^1)X^2\left((X^2)^2+(X^3)^2\right)^{-5/2}\left((X^2)^2-2(X^3)^2\right),\end{aligned}$$

so that

$$\Delta\xi(X) = -\frac{\mathrm{sgn}(X^1)X^2}{((X^2)^2+(X^3)^2)^{3/2}} = -\frac{\xi(X)}{(X^2)^2+(X^3)^2} = -\frac{\xi(X)|\nabla\xi(X)|}{|X^3|}.$$

Furthermore,

$$\begin{aligned}(\nabla\xi)\nabla(\mathcal{W}\circ\theta) &= f(\theta)\left(\frac{\partial\cos\theta}{\partial X^2}\cdot\frac{\partial\theta}{\partial X^2}+\frac{\partial\cos\theta}{\partial X^3}\cdot\frac{\partial\theta}{\partial X^3}\right)\\ &= f(\theta)\Bigg(-\frac{\mathrm{sgn}(X^1)(X^3)^2}{((X^2)^2+(X^3)^2)^{3/2}}\cdot\frac{\mathrm{sgn}(X^1)\mathrm{sgn}(X^3)X^3}{(X^2)^2+(X^3)^2}\\ &\qquad -\frac{\mathrm{sgn}(X^1)X^2X^3}{((X^2)^2+(X^3)^2)^{3/2}}\cdot\frac{\mathrm{sgn}(X^1)\mathrm{sgn}(X^3)X^2}{(X^2)^2+(X^3)^2}\Bigg)\\ &= f(\theta)\frac{\mathrm{sgn}(X^3)}{((X^2)^2+(X^3)^2)^{5/2}}\left(-(X^3)^3-(X^2)^2X^3\right)\\ &= -f(\theta)\frac{|X^3|}{((X^2)^2+(X^3)^2)^{3/2}} = -f(\arccos\xi(X))\frac{\xi(X)|\nabla\xi(X)|}{X^2}.\end{aligned}$$

Finally,

$$
\begin{aligned}
&-(\nabla\xi)(\nabla V^{\varepsilon})(X) = \frac{1}{\varepsilon}\frac{\partial \cos\theta}{\partial X^2}\cdot\left(\frac{l_{eq}-\sqrt{(X^2)^2+(X^3)^2}}{\sqrt{(X^2)^2+(X^3)^2}}\right)X^2 \\
&+\frac{1}{\varepsilon}\frac{\partial \cos\theta}{\partial X^3}\cdot\left(\frac{l_{eq}-\sqrt{(X^2)^2+(X^3)^2}}{\sqrt{(X^2)^2+(X^3)^2}}\right)X^3-(\nabla\xi)\nabla\mathcal{W}(\theta) \\
= \quad &\frac{1}{\varepsilon}\frac{l_{eq}-\sqrt{(X^2)^2+(X^3)^2}}{\sqrt{(X^2)^2+(X^3)^2}}\left(\frac{\operatorname{sgn}(X^1)(X^3)^2}{((X^2)^2+(X^3)^2)^{3/2}}X^2-\frac{\operatorname{sgn}(X^1)X^2X^3}{((X^2)^2+(X^3)^2)^{3/2}}X^3\right) \\
&+f(\arccos\xi(X))\frac{\xi(X)|\nabla\xi(X)|}{X^2} = f(\arccos\xi(X))\frac{\xi(X)|\nabla\xi(X)|}{X^2}.
\end{aligned}
$$

The choice for the reaction coordinate ξ is such that it is orthogonal to the stiff terms of potential energy q_{AB} and q_{BC} since $(\nabla\xi)(\nabla q_{AB}) = 0 = (\nabla\xi)(\nabla q_{BC})$.

The corresponding SDE for $\xi_t = \cos\theta_t$ becomes

(5.3.6)

$$
\xi_t = \xi_0 + \int_0^t \underbrace{f(\arccos\xi_s)\frac{\xi_s|\nabla\xi(X_s)|}{X_s^2} - \frac{1}{\beta}\frac{\xi_s|\nabla\xi(X_s)|}{|X_s^3|}}_{a(\xi_s,X_s)} ds + \sqrt{\frac{2}{\beta}}\int_0^t \underbrace{|\nabla\xi(X_s)|}_{b(\xi_s,X_s)} d\widetilde{W}_s.
$$

SDE (5.3.6) is not closed. Gyöngy [Gyö86, (1.3)] suggested the following closing procedure

$$
\widetilde{\xi}_t = \xi_0 + \int_0^t \underbrace{\mathbb{E}\left(a(\xi_s,X_s)|\xi_s=\widetilde{\xi}_s\right)}_{\widetilde{a}(s,\widetilde{\xi}_s)} ds + \sqrt{\frac{2}{\beta}}\int_0^t \underbrace{\sqrt{\mathbb{E}\left(b^2(\xi_s,X_s)|\xi_s=\widetilde{\xi}_s\right)}}_{\widetilde{b}(s,\widetilde{\xi}_s)} d\widetilde{W}_s,
$$

where now the drift and diffusion coefficients are non-autonomous, i.e. they depend explicitly on time. A way to get an autonomous equation again is proposed in [LL10, (23)] where the authors consider

(5.3.7)

$$
\overline{\xi}_t = \xi_0 + \int_0^t \underbrace{\mathbb{E}_\mu\left(a(\xi,X)|\xi=\overline{\xi}_s\right)}_{\overline{a}(\overline{\xi}_s)} ds + \sqrt{\frac{2}{\beta}}\int_0^t \underbrace{\sqrt{\mathbb{E}_\mu\left(b^2(\xi,X)|\xi=\overline{\xi}_s\right)}}_{\overline{b}(\overline{\xi}_s)} d\widetilde{W}_s,
$$

where now the expectation is w.r.t. the measure μ_z given by
(5.3.8)
$$d\mu_z = \underbrace{\left(\int_{\{X\in\mathbb{R}^3:\xi(X)=z\}} e^{-\beta V^{\varepsilon}(X)}|\nabla\xi(X)|^{-1}d\sigma_z\right)^{-1}}_{C_z} e^{-\beta V^{\varepsilon}(X)}|\nabla\xi(X)|^{-1}d\sigma_z$$

and σ_z is the Hausdorff measure on the submanifold $\Sigma_z = \{X\in\mathbb{R}^3 : \xi(X) = z\}$ of $\mathbb{R}^3$, induced by the Hausdorff measure in the ambient Euclidean space $\mathbb{R}^3 \supset \Sigma_z$. The first factor in (5.3.8) is a normalizing constant, depending on z.

The expressions for the expectations of the above coefficients w.r.t. the measure μ_z read

$$\begin{aligned}
\mathbb{E}_\mu\left(a(\xi,X)|\xi=z\right) &= \int_{\Sigma_z}|\nabla\xi(X)|\left(f(\arccos z)\frac{z}{X^2}-\frac{1}{\beta}\frac{z}{|X^3|}\right)d\mu_z\\
&= C_z\left(f(\arccos z)z\int_{\Sigma_z}\frac{1}{X^2}e^{-\beta V^{\varepsilon}(X)}d\sigma_z-\frac{1}{\beta}z\int_{\Sigma_z}\frac{1}{|X^3|}e^{-\beta V^{\varepsilon}(X)}d\sigma_z\right)\\
&= C_z f(\arccos z)z\int_{\Sigma_z}\frac{1}{X^2}e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]}d\sigma_z\\
&\quad -\frac{C_z}{\beta}z\int_{\Sigma_z}\frac{1}{|X^3|}e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]}d\sigma_z,
\end{aligned}$$

and for the diffusion coefficient it holds that

$$\begin{aligned}
\mathbb{E}_\mu\left(b^2(\xi,X)|\xi=z\right) &= \int_{\Sigma_z}|\nabla\xi(X)|^2d\mu_z\\
&= C_z\int_{\Sigma_z}\frac{|X^3|}{(X^2)^2+(X^3)^2}e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]}d\sigma_z,
\end{aligned}$$

where $\Sigma_z = \{X\in\mathbb{R}^3 : \xi(X)=z\} = \{X\in\mathbb{R}^3 : \frac{\operatorname{sgn}(X^1)X^2}{\sqrt{(X^2)^2+(X^3)^2}} = z\}$ and

$$C_z = \int_{\Sigma_z}\frac{(X^2)^2+(X^3)^2}{|X^3|}e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]}d\sigma_z.$$

We decompose $\Sigma_z = \Sigma_z^+\cup\Sigma_z^-$ where $\Sigma_z^\pm = \{(j,k)\in\mathbb{R}^2 : j = \pm z\sqrt{j^2+k^2}\}$ and $(X^1,X^2,X^3) = (x,j,k)$. Therefore we have that,

$$\int_{\Sigma_z}\frac{1}{X^2}e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]}d\sigma_z = 2\int_{x>0}e^{-\frac{\beta}{2\varepsilon}(x-l_{eq})^2}dx\int_{\Sigma_z^+}\frac{1}{j}e^{-\frac{\beta}{2\varepsilon}(\sqrt{j^2+k^2}-l_{eq})^2}d\sigma_z^+,$$

and analogously get that

$$\int_{\Sigma_z} \frac{(X^2)^2+(X^3)^2}{|X^3|} e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]} d\sigma_z = 2\int_{x>0} e^{-\frac{\beta}{2\varepsilon}(x-l_{eq})^2}dx \times \int_{\Sigma_z^+} \frac{j^2+k^2}{|k|} e^{-\frac{\beta}{2\varepsilon}(\sqrt{j^2+k^2}-l_{eq})^2} d\sigma_z^+.$$

It holds that

$$\begin{aligned}
\frac{\int_{\Sigma_z} \frac{1}{X^2} e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]} d\sigma_z}{\int_{\Sigma_z} \frac{(X^2)^2+(X^3)^2}{|X^3|} e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]} d\sigma_z} &= \frac{\int_{\Sigma_z^+} \frac{1}{j} e^{-\frac{\beta}{2\varepsilon}(\sqrt{j^2+k^2}-l_{eq})^2} d\sigma_z^+}{\int_{\Sigma_z^+} \frac{j^2+k^2}{|k|} e^{-\frac{\beta}{2\varepsilon}(\sqrt{j^2+k^2}-l_{eq})^2} d\sigma_z^+} \\
&= \frac{\int_0^\infty \frac{1}{rz} e^{-\frac{\beta}{2\varepsilon}(r-l_{eq})^2} r dr}{\int_0^\infty \frac{r^2}{r\sqrt{1-z^2}} e^{-\frac{\beta}{2\varepsilon}(r-l_{eq})^2} r dr} \\
&= \frac{\sqrt{1-z^2}}{z}\frac{J_0}{J_2},
\end{aligned}$$

where we made the transformation $(j,k) = (r\cos\phi, r\sin\phi)$ and

$$J_n := \int_0^\infty r^n e^{-\frac{\beta}{2\varepsilon}(r-l_{eq})^2} dr.$$

Moreover,

$$\begin{aligned}
\frac{\int_{\Sigma_z} \frac{1}{|X^3|} e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]} d\sigma_z}{\int_{\Sigma_z} \frac{(X^2)^2+(X^3)^2}{|X^3|} e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]} d\sigma_z} &= \frac{\int_0^\infty \frac{1}{r\sqrt{1-z^2}} e^{-\frac{\beta}{2\varepsilon}(r-l_{eq})^2} r dr}{\int_0^\infty \frac{r^2}{r\sqrt{1-z^2}} e^{-\frac{\beta}{2\varepsilon}(r-l_{eq})^2} r dr} \\
&= \frac{J_0}{J_2},
\end{aligned}$$

which implies

$$\overline{a}(z) = \frac{J_0}{J_2} f(\arccos z) z \frac{\sqrt{1-z^2}}{z} - \frac{J_0}{J_2}\frac{1}{\beta} z = \overline{C} f(\arccos z)\sqrt{1-z^2} - \frac{\overline{C}}{\beta} z,$$

with $\overline{C} = J_0/J_2$. Furthermore,

$$\begin{aligned}
\int_{\Sigma_z} |\nabla\xi(X)|^2 d\mu_z &= \frac{\int_{\Sigma_z} \frac{|X^3|}{(X^2)^2+(X^3)^2} e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]} d\sigma_z}{\int_{\Sigma_z} \frac{(X^2)^2+(X^3)^2}{|X^3|} e^{-\beta[\frac{q_{AB}^2(X)}{2\varepsilon}+\frac{q_{BC}^2(X)}{2\varepsilon}]} d\sigma_z} \\
&= \frac{\int_0^\infty \frac{r\sqrt{1-z^2}}{r^2} e^{-\frac{\beta}{2\varepsilon}(r-l_{eq})^2} r dr}{\int_0^\infty \frac{r^2}{r\sqrt{1-z^2}} e^{-\frac{\beta}{2\varepsilon}(r-l_{eq})^2} r dr} = (1-z^2)\frac{J_0}{J_2},
\end{aligned}$$

which implies $\overline{b}(z) = \sqrt{\overline{C}(1-z^2)}$. The effective dynamics (5.3.7) now read
(5.3.9)

$$\overline{\xi}_t = \xi_0 + \overline{C}\int_0^t \left(f(\arccos \overline{\xi}_s)\sqrt{1-\overline{\xi}_s^2} - \frac{1}{\beta}\overline{\xi}_s \right) ds + \sqrt{\frac{2\overline{C}}{\beta}} \int_0^t \sqrt{1-\overline{\xi}_s^2} d\widetilde{W}_s.$$

SDE (5.3.9) is of the form of (5.1.1) with $c = \sqrt{2\overline{C}\beta^{-1}}$ and $\phi(x) = f(\arccos x)$.

5.4 Local error of the SD method.

In this Section we provide uniform moment bounds for the original SDE and the SD scheme as well as the local error of the proposed scheme. We remind here that for notational reasons the processes (W_t, x_t) stand for $(\widehat{W}_t, \widehat{x}_t)$.

Lemma 5.4.6 *[Moment bounds for original problem and SD approximation] Let Assumption 5.2.2 hold. Then*

$$\mathbb{E} \sup_{0\leq t\leq T} |x_t|^p \bigvee \mathbb{E} \sup_{0\leq t\leq T} |y_t|^p \leq 1,$$

for any $p > 0$. □

Proof of Lemma 5.4.6. The result is trivial since we already know that (x_t) satisfying (5.2.7) has the property $x_t \in D$ when $x_0 \in D, D = (-1,1)$, by Appendix F.1 and regarding the bounds for the SD approximation (5.2.1), it is clear, by its form (5.2.3), that they are valid. □

For the rest of this section we write (5.2.1) in a compact form, introducing an auxiliary process (h_t) as
(5.4.1)

$$y_t^{SD} = \underbrace{x_0 + \int_0^t \Phi(y_{\hat{s}}, y_s)ds + \int_0^t g(y_s)dW_s}_{h_t} + \int_t^{t_{n+1}} \phi(y_{t_n})\sqrt{1-y_{t_n}^2}ds$$

where $\hat{s} = t_n$ when $s \in [t_n, t_{n+1})$ and $\Phi(a,b) = \phi(a)\sqrt{1-a^2} - c^2 b/2$ and $g(a) = c\sqrt{1-a^2}$.

By the above representation, the form of the discretization becomes apparent. We only discretized the drift coefficient of (5.2.7) in an additive way, so that the remaining part $-(c^2/2)y_t dt$ combined with the diffusion part

$g(y_t)dW_t$ produces the analytic solution (5.2.3). The next result concerns the local error of the proposed scheme.

Lemma 5.4.7 *[Error bound for SD scheme] Let Assumption 5.2.2 hold and n_s be an integer such that $s \in [t_{n_s}, t_{n_s+1}]$. There is a $K_p > 0$, which does not depend on Δ, such that*

$$\mathbb{E}|y_s - y_{\hat{s}}|^p \leq K_p \Delta^{p/2},$$

for any $p > 0$. and for any $s \geq 0$. □

Proof of Lemma 5.4.7. First we take a $p \geq 2$. Representation (5.2.1) yields

$$\begin{aligned}
|y_s - y_{\hat{s}}|^p &= \Big| \int_{t_{n_s}}^{t_{n_s+1}} \phi(y_{\hat{u}}) \sqrt{1 - y_{\hat{u}}^2} du + \int_{t_{n_s}}^{s} \frac{-c^2}{2} y_u du + c \int_{t_{n_s}}^{s} \sqrt{1 - y_u^2} dW_u \Big|^p \\
&\leq 3^{p-1} \left(\left| \int_{t_{n_s}}^{t_{n_s+1}} \phi(y_{t_{n_s}}) \sqrt{1 - y_{t_{n_s}}^2} du \right|^p + \left| \int_{t_{n_s}}^{s} \frac{c^2}{2} y_u du \right|^p + \left| \int_{t_{n_s}}^{s} c \sqrt{1 - y_u^2} dW_u \right|^p \right) \\
&\leq 3^{p-1} |K_\phi|^p | \sqrt{1 - y_{t_{n_s}}^2} |^p \Delta^p + 3^{p-1} \frac{c^{2p}}{2^p} \int_{t_{n_s}}^{s} |y_u|^p du + 3^{p-1} |c|^p \left| \int_{t_{n_s}}^{s} \sqrt{1 - y_u^2} dW_u \right|^p
\end{aligned}$$

where we have used the Cauchy-Schwarz inequality and $K_\phi = \sup |\phi|$. Taking expectations in the above inequality and using Lemma 5.4.6 and the BDG inequality (B.3.5) on the diffusion term we conclude, as in the proof of Lemma 4.3.6,

$$\mathbb{E}|y_s - y_{\hat{s}}|^p \leq K_p \Delta^{p/2},$$

where the positive quantity K_p depends on p and on the parameters c, K_ϕ but not on Δ. The case $0 < p < 2$ follows by Jensen's inequality since

$$\mathbb{E}|y_s - y_{\hat{s}}|^p \leq \left(\mathbb{E}|y_s - y_{\hat{s}}|^2 \right)^{p/2} \leq K_p \Delta^{p/2}.$$

□

Remark 5.4.8 *Note that Lemmata 5.4.6 and 5.4.7 are valid when $\phi(\cdot)$ is only bounded; we do not need to assume anything more for $\phi(\cdot)$, i.e. Lipschitz continuity.* □

5.5 Numerical Experiments.

Let a_0, a_1, a_2 be such that $f(z) = z^3 + a_2 z^2 + a_1 z + a_0$ where $f(\cdot)$ is the derivative of the potential $\mathcal{W}$ in Section 5.3. The continuous form of the SD scheme (5.2.3) now reads

$$y_t^{SD} = \cos\left(-c\Delta\widetilde{W}_n^t + \arccos(y_{t_n} + \overline{C}f(\arccos y_{t_n})\sqrt{1-y_{t_n}^2}\cdot\Delta)\right),$$

where $\Delta\widetilde{W}_n^t = \widetilde{W}_t - \widetilde{W}_{t_n}$ and is well defined when

$$|y_{t_n} + \overline{C}f(\arccos y_{t_n})\sqrt{1-y_{t_n}^2}\cdot\Delta| \leq 1.$$

We first prove a lemma which concerns the well-posedness of our proposed scheme in that case, i.e. we want to examine when Assumption 5.2.2 holds.

Lemma 5.5.9 *The SD scheme (5.2.3) is well defined for all Δ such that*

$$\overline{C}\Delta < \frac{2\sqrt{2}+\pi-2}{\sqrt{2}\pi(\pi^2+2a_1)} \bigwedge \frac{4+\pi^2}{-(4a_2+a_0)\pi^2-4a_0} \bigwedge \frac{1}{\pi(\pi^2+a_1)} \bigwedge \frac{1}{-a_2\pi^2-a_0}.$$

□

Proof of Lemma 5.5.9. We have to show that (5.2.4) holds for appropriate $\Delta > 0$ and all $y \in [-1,1]$. If $y=-1$ or $y=1$ then (5.2.4) holds trivially for all $\Delta > 0$. We therefore examine the cases $-1 < y \leq 0$ and $0 \leq y < 1$. Also, note that $a_2 < 0$ when $\theta_m > 0$, that $a_1 > 0, a_0 < 0$ when $\theta_0(2-\frac{\theta_0}{\theta_m}) > 0$ and $\overline{C} > 0$. We will use the following inequality [Zhu09, Th. 5],

$$\frac{3(2\sqrt{1-y})^p}{(2\sqrt{2})^p + (\sqrt{1+y})^p} < (\arccos y)^p < \frac{(2\pi\sqrt{1-y})^p}{(2\sqrt{2})^p + (\pi^p-2^p)(\sqrt{1+y})^p},$$

valid for any $0 < y < 1$ and $p \geq 1$. The above relation and the property $\arccos(-y) = \pi - \arccos(y)$ imply

$$\left[\pi - \frac{2\pi\sqrt{1+y}}{((2\sqrt{2})^p + (\pi^p-2^p)(\sqrt{1-y})^p)^{1/p}}\right]^p < (\arccos y)^p$$

and

$$(\arccos y)^p < \left[\pi - \frac{3^{1/p}(2\sqrt{1+y})}{((2\sqrt{2})^p + (\sqrt{1-y})^p)^{1/p}}\right]^p,$$

for any $-1 < y < 0$ and $p \geq 1$.
Case $0 \leq y < 1$: Then

$$\begin{aligned}
& y + \overline{C} f(\arccos y)\sqrt{1-y^2}\Delta \leq y + \overline{C}((\arccos y)^3 + a_1 \arccos y)\sqrt{1-y^2}\Delta \\
\leq\ & y + \overline{C}\Big(\frac{8\pi^3(\sqrt{1-y})^3}{16\sqrt{2} + (\pi^3-8)(\sqrt{1+y})^3} + a_1 \frac{2\pi\sqrt{1-y}}{2\sqrt{2} + (\pi-2)\sqrt{1+y}}\Big)\sqrt{1-y^2}\Delta \\
\leq\ & y + \overline{C}\Big(\frac{8\pi^3(1-y)^2}{16\sqrt{2} + (\pi^3-8)(\sqrt{1+y})^3} + a_1 \frac{2\pi(1-y)}{2\sqrt{2} + (\pi-2)\sqrt{1+y}}\Big)\sqrt{1+y}\Delta \\
\leq\ & y + (1-y)\overline{C}\frac{\sqrt{2}\pi(\pi^2+2a_1)}{2\sqrt{2}+\pi-2}\Delta \leq 1,
\end{aligned}$$

when $\overline{C}\Delta \leq \frac{2\sqrt{2}+\pi-2}{\sqrt{2}\pi(\pi^2+2a_1)}$. Moreover,

$$\begin{aligned}
& y + \overline{C} f(\arccos y)\sqrt{1-y^2}\Delta \geq y + \overline{C}(a_2(\arccos y)^2 + a_0)\sqrt{1-y^2}\Delta \\
\geq\ & y + \overline{C}\left(a_2 \frac{4\pi^2(1-y)}{8 + (\pi^2-4)(1+y)} + a_0\right)\Delta \\
\geq\ & y\left(1 - 2\overline{C}a_2\Delta\right) + \overline{C}\left(a_2\frac{4\pi^2}{4+\pi^2} + a_0\right)\Delta \\
\geq\ & \overline{C}\frac{(4a_2+a_0)\pi^2 + 4a_0}{4+\pi^2}\Delta \geq -1,
\end{aligned}$$

when $\overline{C}\Delta \leq \frac{4+\pi^2}{-(4a_2+a_0)\pi^2 - 4a_0}$.
Case $-1 < y < 0$: It holds

$$\begin{aligned}
& y + \overline{C} f(\arccos y)\sqrt{1-y^2}\Delta \leq y + \overline{C}((\arccos y)^3 + a_1 \arccos y)\sqrt{1-y^2}\Delta \\
\leq\ & y + \overline{C}\left((\pi - \frac{3^{1/3}2\sqrt{1+y}}{\left(16\sqrt{2} + (\sqrt{1-y})^3\right)^{1/3}})^3 + a_1(\pi - \frac{6\sqrt{1+y}}{2\sqrt{2}+\sqrt{1-y}})\right)\Delta \\
\leq\ & \overline{C}\left(\pi^3 + a_1\pi\right)\Delta \leq 1,
\end{aligned}$$

when $\overline{C}\Delta \leq \frac{1}{\pi^3 + a_1\pi}$. Finally

$$\begin{aligned}
& y + \overline{C} f(\arccos y)\sqrt{1-y^2}\Delta \geq y + \overline{C}(a_2(\arccos y)^2 + a_0)\sqrt{1-y^2}\Delta \\
\geq\ & y + \overline{C}\left(a_2\left(\pi - \frac{2\sqrt{3}\sqrt{1+y}}{\sqrt{9-y}}\right)^2 + a_0\right)(1+y)\Delta \\
\geq\ & y + \overline{C}\left(a_2\pi^2 + a_0\right)(1+y)\Delta \geq -1,
\end{aligned}$$

when $\overline{C}\Delta \le \frac{-1}{a_2\pi^2+a_0}$. □

Lemma 5.5.9 suggests the following assumption for the SD scheme to be well-posed.

Assumption 5.5.10 *Let the discretization step* Δ *be such that*

$$\overline{C}\Delta < \frac{2\sqrt{2}+\pi-2}{\sqrt{2}\pi(\pi^2+2a_1)} \bigwedge \frac{4+\pi^2}{-(4a_2+a_0)\pi^2-4a_0} \bigwedge \frac{1}{\pi(\pi^2+a_1)} \bigwedge \frac{1}{-a_2\pi^2-a_0}.$$

□

The SD iterative scheme for the numerical approximation of (5.3.9) reads

$$y^{SD}_{t_{n+1}} = \cos\left(-c\Delta\widetilde{W}_n + \arccos(y_{t_n} + \overline{C}f(\arccos y_{t_n})\sqrt{1-y^2_{t_n}}\Delta)\right),$$

where $\Delta\widetilde{W}_n := \widetilde{W}_{t_{n+1}} - \widetilde{W}_{t_n}$, are the increments of the Wiener process.

We consider the configuration as in [LL10, Sec. 5.6], i.e.

$$\mathcal{W}(\theta) = \frac{k_\theta}{2}(\theta-\theta_0)^2,\ k_\theta = 208,\ l_{eq} = 1,\ \beta = 1, \epsilon = 10^{-3},$$

where we now take the initial angle $\theta_0 = 3$. Then $f(\theta) = \mathcal{W}'(\theta) = k_\theta\theta - k_\theta\theta_0$. The effective dynamics are

(5.5.1)
$$\overline{\xi}_t = \xi_0 + \overline{C}\int_0^t \left(k_\theta(\arccos(\overline{\xi}_s)-\theta_0)\sqrt{1-\overline{\xi}^2_s} - \overline{\xi}_s\right) ds + \sqrt{2\overline{C}}\int_0^t \sqrt{1-\overline{\xi}^2_s}\, d\widetilde{W}_s,$$

where $\overline{C} = 0.999$. We want to compare our proposed SD scheme

(5.5.2)
$$y^{SD}_{t_{n+1}} = \cos\left(-\sqrt{2\overline{C}}\Delta\widetilde{W}_n + \arccos(y_n + \overline{C}k_\theta(\arccos(y_n)-\theta_0)\sqrt{1-y^2_n}\Delta)\right),$$

with the EM scheme which reads

(5.5.3)
$$y^{EM}_{n+1} = y_n + \overline{C}\left(k_\theta(\arccos(y_n)-\theta_0)\sqrt{1-y^2_n} - y_n\right)\cdot\Delta + \sqrt{2\overline{C}}\sqrt{1-y^2_n}\Delta\widetilde{W}_n,$$

where $\Delta\widetilde{W}_n = \frac{\mathrm{sgn}(X^1)}{\sqrt{(X^2)^2+(X^3)^2}}X^3(\mathrm{sgn}(X^3)\Delta W^2_n - X^2)\Delta W^3_n$ and $y_n = y_{t_n}$.

According to the notation introduced in the beginning of the section, we have that $a_2 = 0, a_1 = k_\theta$ and $a_0 = -k_\theta \theta_0$. Note, that according to Assumption 5.5.10, the SD scheme (5.5.2) is well-posed for all Δ such that

$$\Delta < \frac{1}{\overline{C}} \left[\frac{2\sqrt{2} + \pi - 2}{\sqrt{2}\pi(\pi^2 + 2k_\theta)} \bigwedge \frac{1}{k_\theta \theta_0} \bigwedge \frac{1}{\pi(\pi^2 + k_\theta)} \right].$$

Since the EM scheme has a finite life time, in order to be well posed, we examine the following modification of (5.5.3)

$$\begin{aligned} y^{+EM}_{t_{n+1}} &= y_{t_n} + \overline{C} \left(k_\theta (\arccos(y_{t_n}) - \theta_0) \sqrt{(1 - y^2_{t_n})^+} - y_{t_n} \right) \cdot \Delta \\ &\quad + \sqrt{2\overline{C}} \sqrt{(1 - y^2_{t_n})^+} \Delta \widetilde{W}_n. \end{aligned} \tag{5.5.4}$$

Below, we make a simple numerical experiment to compare the EM scheme (5.5.4) with the proposed SD (5.5.2). For the implementation of the SD method, we have to assume that $\Delta < 0.0021 \bigwedge 0.0016 \bigwedge 0.00146$, thus the step $\Delta = 10^{-3}$ is sufficient. Figure 5.1 shows that EM produces values outside the interval $[-1, 1]$, even when the time $T = 1$, where SD by its construction does not exhibit that behavior.

Therefore, in order to make a comparative result of SD scheme with EM, we have to consider the following modification of the EM scheme, a stopped EM scheme (sEM), which is structure preserving,

$$y^{sEM}_{t_{n+1}} = \begin{cases} -1, & y^{+EM}_{t_{n+1}} < -1, \\ 1, & y^{+EM}_{t_{n+1}} > 1, \\ y^{+EM}_{t_{n+1}}, & \text{otherwise.} \end{cases} \tag{5.5.5}$$

If at some time t_k, the EM scheme drops below 1, that is $y^{+EM}_{t_k} < -1$, then the stopped EM scheme $y^{sEM}_{t_k} = -1$ and in the next step we have

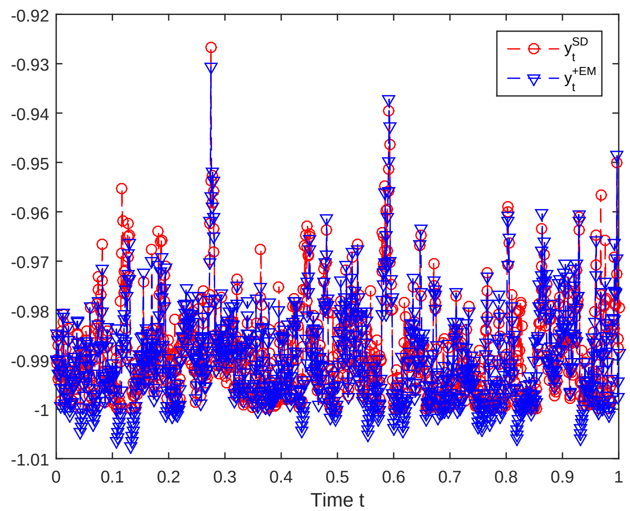
$$y^{sEM}_{t_{k+1}} = y^{sEM}_{t_k}(1 - \overline{C}\Delta) = -1 + \overline{C}\Delta \leq 1,$$

for $\Delta \leq 2/\overline{C}$. Moreover, for the case where EM exceeds the upper boundary 1 at time t_k,

$$y^{sEM}_{t_{k+1}} = y^{sEM}_{t_k}(1 - \overline{C}\Delta) = 1 - \overline{C}\Delta \geq -1,$$

for $\Delta \leq 2/\overline{C}$. Thus, the stopped EM scheme (5.5.5) is well defined for $\Delta \leq 2/\overline{C}$.

Fig. 5.1: Comparison of a path of (y_t^{+EM}) and (y_t^{SD}) at step $\Delta = 10^{-3}$, with $\theta_0 = 3$. Euler method produces values outside the interval $[-1, 1]$.



We aim to show experimentally the order of convergence of structure preserving methods for the estimation of the true solution of (5.5.1). Therefore, we consider the semi-discrete method (5.5.2). We want to verify our theoretical results and in particular the order shown in Theorem 5.2.4. Moreover, we would like to compare SD with the EM modification (5.5.4), even though it is not structure preserving,and with the stopped EM scheme (5.5.5) in terms of error estimation and computer time consumption.

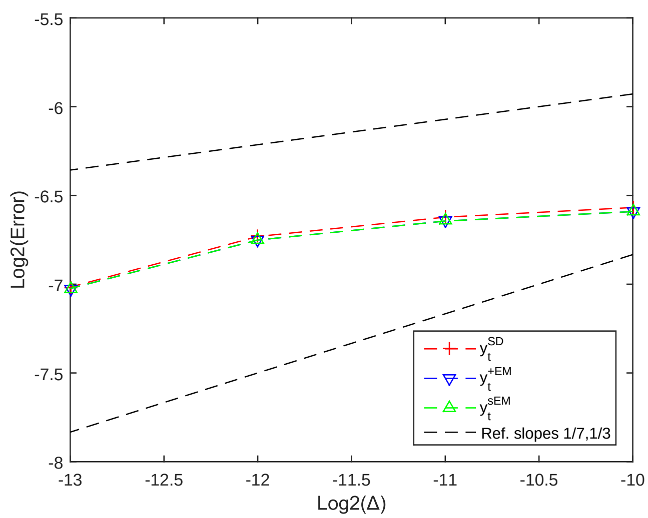
We estimate the endpoint $\mathcal{L}^2$-norm $\epsilon = \sqrt{\mathbb{E}|y^{(\Delta)}(T) - \overline{\xi}_T|^2}$, of the difference between the numerical scheme evaluated at step size Δ and the exact solution of (5.5.1). For that purpose, we compute M batches of L simulation paths, where each batch is estimated by $\hat{\epsilon_j} = \frac{1}{L}\sum_{i=1}^{L} |y_{i,j}^{(\Delta)}(T) - y_{i,j}^{(ref)}(T)|^2$ and the Monte Carlo estimator of the error is

$$\hat{\epsilon} = \sqrt{\frac{1}{ML}\sum_{j=1}^{M}\sum_{i=1}^{L} |y_{i,j}^{(\Delta)}(T) - y_{i,j}^{(ref)}(T)|^2}, \tag{5.5.6}$$

and requires $M \cdot L$ Monte Carlo sample paths. The reference solution is evaluated at step size 2^{-14} of the numerical scheme. For the proposed SD scheme, we have shown in Theorems 5.2.4 that it strongly converges to the exact solution, so we take that as a reference solution. We simulate $100 \cdot 100 = 10^4$ paths, where the choice for $L = 100$ is as in [KPS03, p.118]. Of course, the number of Monte Carlo paths has to be sufficiently large, so as not to significantly hinder the mean-square errors.

We plot in a $\log_2$-$\log_2$ scale and error bars represent 98%-confidence intervals. The results are shown in Table 5.1 and Figure 5.2. Table 5.1 does not present the computed Monte Carlo errors with 98% confidence, since they were at least 10 times smaller that the mean-square errors.

Fig. 5.2: Convergence of SD, +EM and sEM applied to SDE (5.5.1) with parameters $(\theta_0, k_\theta, \overline{C}, T) = (3, 208, 0.999, 1)$ with 32 digits of accuracy.



Step Δ	SD-Error	Rate	+EM-Error	Rate	sEM-Error	Rate
2^{-10}	0.010459	–	**0.010333**	–	0.010654	–
2^{-11}	0.010263	0.0273	0.010042	0.0412	**0.009945**	**0.0994**
2^{-12}	0.009470	**0.116**	0.009333	0.1056	**0.009269**	0.1016
2^{-13}	0.007674	**0.3034**	0.007645	0.2878	**0.007637**	0.2794

Tab. 5.1: Error and step size of SD, +EM and sEM scheme for (5.5.1) with $(\theta_0, k_\theta, \overline{C}, T) = (3, 208, 0.999, 1)$ and 32 digits of accuracy.

In Table 5.2 we present the computational times,[3] of the explicit numerical

[3] We simulate with 3.06GHz Intel Pentium, 1.49GB of RAM in Matlab $R2014b$ Software.

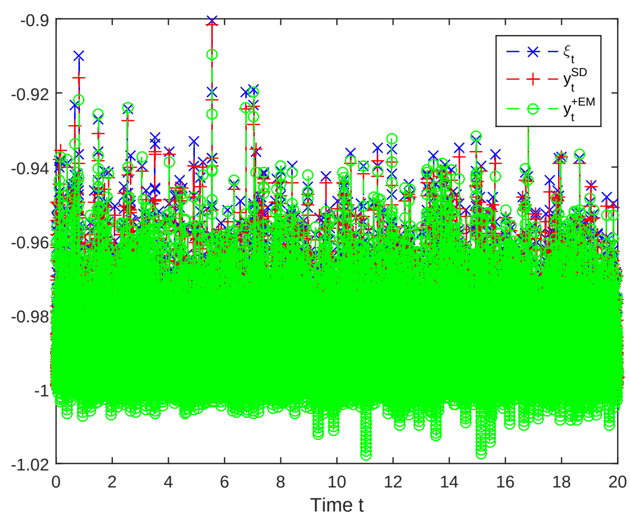
schemes SD, the modification of EM and the stopped sEM scheme for the same problem.

Step Δ	SD	+EM	sEM
2^{-10}	0.000715	0.001917	0.000565
2^{-11}	0.001398	0.003602	0.001088
2^{-12}	0.002768	0.007231	0.002191
2^{-13}	0.005385	0.014094	0.004183

Tab. 5.2: Average computational time for a path (in seconds) for the selected schemes.

In Figure 5.3 we illustrate a path of the solution $\cos(\theta_{ABC}(X_t))$ where X_t solves the 3-dimensional system approximated by Euler-Maruyama(EM) method with discretization step $\Delta = 10^{-3}$ and the effective dynamics (5.5.1) computed again with SD scheme (5.5.2) and EM scheme (5.5.3), taking into account the path of W_t considered for the solution process X_t.

Fig. 5.3: A sample path of the transformation $\xi_t = \cos(\theta_{ABC}(X_t))$ of the solution process and the effective dynamics y_t^{SD}, y_t^{+EM}, at step $\Delta = 10^{-3}$.



As we saw before in Figure 5.1, EM produced negative values, even for time horizon $T = 1$, so by increasing the integration time T, to 20 in this

The random number generator is Mersenne Twister.

case, results in an increase of the probability of such events. In particular, Figure 5.3 shows this oscillatory behavior.

We prefer not to compare with implicit methods, if any, which are structure preserving, since in principle they require more computational time. Nevertheless, we present one such method. The balanced Milstein method (BMM), proposed in [KS06], is given by the following linear implicit relation

$$\begin{aligned} y_{t_{n+1}}^{BMM} &= y_{t_n} + \overline{C}\left(k_\theta(\arccos(y_{t_n}) - \theta_0)\sqrt{1-y_{t_n}^2} - y_{t_n}\right)\Delta \\ &\quad + \sqrt{2\overline{C}}\sqrt{1-y_{t_n}^2}\Delta\widetilde{W}_n - \overline{C}y_{t_n}\left((\Delta\widetilde{W}_n)^2 - \Delta\right) \\ &\quad + \left(d^0(y_{t_n})\Delta_n + d^1(y_{t_n})((\Delta\widetilde{W}_n)^2 - \Delta)\right)(y_{t_n} - y_{t_n+1}), \end{aligned}$$

for $n = 0, \dots, N-1$ where d^0 and d^1 are appropriate weight functions. Rearranging leads to

$$\begin{aligned} &(1 + d^0(y_x)\Delta_n + d^1(y_x)((\Delta\widetilde{W}_n)^2 - \Delta))y_{t_{n+1}}^{BMM} = x \\ &+(\overline{C}k_\theta(\arccos(x) - \theta_0)\sqrt{1-x^2} + (d^0(x) - d^1(x))x)\Delta + g(\Delta\widetilde{W}_n), \end{aligned}$$

where $x = y_{t_n}$ and $g(z) = \sqrt{2\overline{C}}\sqrt{1-x^2}\cdot z + (d^1(x) - \overline{C})x \cdot z^2$. The BMM scheme is able to preserve positivity for suitable d^0 and d^1 but it is not clear if there are functions d^0 and d^1 such that starting with an $x \in (-1, 1)$ we have that $y_{t_{n+1}}^{BMM} \in [-1, 1]$ too.

Moreover, there exist other interesting balanced schemes (tamed schemes), of the form [Sab15]

$$y_{t_{n+1}}^{TAMED} = x + \frac{\left(\overline{C}k_\theta(\arccos(x) - \theta_0)\sqrt{1-x^2} - x\right)\Delta + \sqrt{2\overline{C}}\sqrt{1-x^2}\Delta\widetilde{W}_n}{1 + \left|\overline{C}k_\theta(\arccos(x) - \theta_0)\sqrt{1-x^2} - x\right|\Delta^\beta + 2\overline{C}(1-x^2)\Delta^\beta}$$

or [Zha14]

$$y_{t_{n+1}}^{SIN} = x + \sin[(\overline{C}k_\theta(\arccos(x) - \theta_0)\sqrt{1-x^2} - x)\Delta] + \sin[\sqrt{2\overline{C}}\sqrt{1-x^2}\Delta\widetilde{W}_n],$$

but still not clear, whether they possess eternal lifetime or not.

Therefore, we choose the best candidate of the numerical methods presented above and estimate the error produced by the coarse graining procedure. In order to approximate the Coarse Graining Error (CGE), we use again the same Monte Carlo procedure, where now (5.5.6) reads

$$\widehat{CGE} = \sqrt{\frac{1}{ML}\sum_{j=1}^{M}\sum_{i=1}^{L}|y_{i,j}^{(\Delta)}(T) - \widehat{\cos(\theta(T))}|^2},$$

and consider again $ML = 10^4$ Monte Carlo sample paths. Here $\widehat{\cos(\theta_t)}$ is an approximation of the cosine of the angle

$$\cos\theta_{ABC}(t) = \frac{X_t^1 X_t^2}{|X_t^1|\sqrt{(X_t^2)^2 + (X_t^3)^2}}.$$

We use the EM scheme for the approximation of the 3-dimensional $X^T = (X^1, X^2, X^3)$. In particular, we can implement, as discussed earlier in Section 5.3, any numerical scheme for the scalar SDE (5.3.1), which refers to the evolution of the first coordinate X^1, and then approximate independently the system (X^2, X^3). We choose the EM scheme, since it is easy to implement. It reads

$$_{em}Y_{t_{n+1}}^1 = Y_{t_n}^1\left(1 + \frac{1}{\varepsilon}\left(\frac{l_{eq} - |Y_{t_n}^1|}{|Y_{t_n}^1|}\right)Y_{t_n}^1\Delta\right) + \sqrt{\frac{2}{\beta}}\Delta W_n^1,$$

$$\begin{aligned} _{em}Y_{t_{n+1}}^2 &= Y_{t_n}^2\Bigg[1 + \Bigg(\frac{1}{\varepsilon}\frac{l_{eq} - \sqrt{(Y_{t_n}^2)^2 + (Y_{t_n}^3)^2}}{\sqrt{(Y_{t_n}^2)^2 + (Y_{t_n}^3)^2}}Y_{t_n}^2 \\ &\quad + f(\theta)\frac{Y_{t_n}^1}{|Y_{t_n}^1|}\frac{Y_{t_n}^3}{|Y_{t_n}^3|}\frac{Y_{t_n}^3}{(Y_{t_n}^2)^2 + (Y_{t_n}^3)^2}\Bigg)\Delta\Bigg] + \sqrt{\frac{2}{\beta}}\Delta W_n^2, \end{aligned}$$

$$\begin{aligned} _{em}Y_{t_{n+1}}^3 &= Y_{t_n}^3\Bigg[1 + \Bigg(\frac{1}{\varepsilon}\frac{l_{eq} - \sqrt{(Y_{t_n}^2)^2 + (Y_{t_n}^3)^2}}{\sqrt{(Y_{t_n}^2)^2 + (Y_{t_n}^3)^2}}Y_{t_n}^3 \\ &\quad + f(\theta)\frac{Y_{t_n}^1}{|Y_{t_n}^1|}\frac{Y_{t_n}^3}{|Y_{t_n}^3|}\frac{Y_{t_n}^2}{(Y_{t_n}^2)^2 + (Y_{t_n}^3)^2}\Bigg)\Delta\Bigg] + \sqrt{\frac{2}{\beta}}\Delta W_n^3, \end{aligned}$$

where $\Delta W_n^i := W_{t_{n+1}}^i - W_{t_n}^i$ are the increments of the Brownian motions (W_t^i), $i = 1..3$.

Finally, to get an impression of the difference in computer time consumption, between numerically solving the original system (5.3.1) – (5.3.3) and the effective dynamics (5.5.1) using SD (or sEM) we present Table 5.4. We also include the times for the new Wiener process (5.3.5).

Step Δ	CGE Error using SD	Rate
2^{-10}	0.010611	–
2^{-11}	0.010268	0.0474
2^{-12}	0.009499	0.1123
2^{-13}	0.007671	0.3083

Tab. 5.3: Coarse-graining error estimates using EM for the original system (5.3.1)-(5.3.3) of X and the SD scheme (5.5.2) for the evolution of the effective dynamics (5.5.1) with 32 digits of accuracy.

Step Δ	$\widehat{\cos(\theta(T))}$	SD + (New Wiener Process)
2^{-10}	0.041374	$0.000715 + 0.000298$
2^{-11}	0.076626	$0.001398 + 0.000583$
2^{-12}	0.147262	$0.002768 + 0.001141$
2^{-13}	0.284851	$0.005385 + 0.002200$

Tab. 5.4: Average computational time for a path (in seconds) for the original system (5.3.1)-(5.3.3) of X and the SD scheme (5.5.2) for the evolution of the effective dynamics (5.5.1) with 32 digits of accuracy.

5.6 Conclusion.

In this note, we propose a new explicit numerical scheme for a class of scalar SDEs that appear in the field of molecular dynamics, after a coarse-graining procedure. The qualitative feature of the scheme is its ability to preserve the domain of the original scalar SDE, which in the specific case studied here, is $D = [-1, 1]$. In other words, our scheme possesses an eternal life time. Unfortunately, we are not able to prove strong convergence of the proposed scheme and we restrict ourselves with an application in the numerical experiment Section. Our first goal is to prove a convergence result of the SD scheme.

In previous works concerning the SD method, we have mainly focused on SDEs with non-negative solutions which appear in the field of financial mathematics. We want to exploit further the main idea of the SD method, to be able first to retain some features, as the structure preserving property, but in the same time approximate efficiently the SDE at hand and within reasonable time limits.

The semi-discrete method is problem dependent and at least to us, there is no unified way of applying it. Therefore, we treat each problem separately. The following figure is representative of the general situation we want to handle.

$$\underbrace{X_t}_{Full\, dynamics\,(5.3.1)-(5.3.3)} \longrightarrow \underbrace{\xi(X_t)}_{Push-forward\, dynamics\,(5.3.4)} \longrightarrow \underbrace{\bar{\xi}_t}_{Effective\, dynamics\,(5.3.7)}$$

Here, we mainly dealt with the efficient numerical approximation of the effective dynamics $(\bar{\xi}_t)$ and then estimated the coarse-graining error (CGE) between $\xi(X_t)$ and $(\bar{\xi}_t)$, by computing the error $\mathbb{E}\sup_{\{0\leq t\leq T\}}|\xi(\hat{X}_t)-\bar{\xi}_t|^2$, where for the approximation of the 3-dimensional $(\hat{X}_t)$ we used the EM scheme and for the approximation of the effective dynamics the SD scheme. Moreover, we did that for a particular choice of $(\xi) = \cos(\theta_{ABC})$. We would like to answer the following questions:

Question 1 : Can we generalize these ideas (estimates) for a general scalar transformation $\xi(X)$ for the same problem?

Question 2 : Can we generalize these ideas (estimates) for a vector-valued coarse-graining map $\xi(X)$, by considering for example a 4-atom model?

Furthermore, we used the EM scheme for the numerical approximation of (X_t) in order to get an estimate for $\xi(X)$. Thus,

Question 3 : Can we also improve the estimation of the full dynamics (X_t) using another scheme, which may preserve qualitative features, of the problem, when the EM fails to do so?

APPENDIX

A. STOCHASTIC CALCULUS: BASICS

We briefly review the basic notations of probability theory and stochastic processes in order to define the Brownian motion and henceforth the stochastic integral. For more details we refer to [Mao97, Ch. 1], [Øks03, Ch. 2,3,4], [Fri75, Ch. 1,3,4], [KS88], [JP03] and references therein.

A.1 From a measurable space to a complete probability space.

Let $\Omega \neq \emptyset$ be the set of all possible outcomes (events) of trials of our mathematical model. We are interested in a group of such events, i.e. a family of subsets of Ω, which we denote by $\mathcal{F}$ and call a σ-algebra.

Definition A.1.1 *[σ-algebra] A family $\mathcal{F}$ of subsets of Ω is called a σ-field or a σ-algebra if*

- $\varnothing \in \mathcal{F}$, *where* $\varnothing$ *is the empty set;*
- $A \in \mathcal{F}$ *implies* $A^C \in \mathcal{F}$, *where* $A^C = \Omega \setminus A$ *is the complement of* A*;*
- $A_1, A_2, \ldots \in \mathcal{F}$, *implies* $\cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

□

We call the pair $(\Omega, \mathcal{F})$ a *measurable space* and all elements of $\mathcal{F}$ are called $\mathcal{F}$-measurable sets. For a family $\mathcal{C}$ of subsets of Ω there exists the smallest σ-algebra containing $\mathcal{C}$, which we denote by $\sigma(\mathcal{C})$ and call the *σ-algebra generated by $\mathcal{C}$*. In the case $\Omega = \mathbb{R}^d$ and $\mathcal{C}$ is the family of all open sets in $\mathbb{R}^d$, the generated σ-algebra by $\mathcal{C}$, denoted by $\mathcal{B}^d = \sigma(\mathcal{C})$, is called *Borel σ-algebra* and its elements *Borel sets*.

A function $X : \Omega \to \mathbb{R}^d$ is said to be *$\mathcal{F}$-measurable* if all X_i, $i = 1, \dots, d$ are $\mathcal{F}$-measurable (*random variables*), i.e. if for every i it holds:

$$\{\omega : X_i(\omega) \leq a\} \in \mathcal{F} \text{ for every } a \in \mathbb{R}.$$

Now we are ready to equip our measurable space with a *probability measure.*

Definition A.1.2 *[Probability measure] A function $\mathbb{P} : \mathcal{F} \to [0, 1]$ is called a* probability measure *on the measurable space $(\Omega, \mathcal{F})$ if*

- $\mathbb{P}(\Omega) = 1$*;*
- *It holds $\mathbb{P}(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$ for any sequence $(A_i)_{i \geq 1} \subset \mathcal{F}$ of disjoint sets, i.e. such that $A_i \cap A_j = \emptyset, i \neq j$. (Countable Additivity)*

□

The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is called a *probability space.* Let

$$\overline{\mathcal{F}} := \{A \subset \Omega : \exists L, U \in \mathcal{F} \text{ such that } L \subset A \subset U \text{ with } \mathbb{P}(L) = \mathbb{P}(U)\}.$$

Then $\overline{\mathcal{F}}$ is a σ-algebra called the *completion* of $\mathcal{F}$. The probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called *complete* if $\mathcal{F} = \overline{\mathcal{F}}$.

Finally we introduce the notion of *filtration.*

Definition A.1.3 *[Filtration] A collection $\{\mathcal{F}_t\}_{t \geq 0}$ of increasing σ-algebras of $\mathcal{F}$, i.e. such that $\mathcal{F}_s \subset \mathcal{F}_t$ for all $0 \leq s < t < \infty$ is called a filtration. It is* right continuous *if $\mathcal{F}_t = \mathcal{F}_{t^+} := \cap_{s > t} \mathcal{F}_s$ for all $t \geq 0$.* □

Heuristically, the filtration tells us about future time information, that is when we are at time t we know for every set in $\mathcal{F}_t$ whether ω belongs to that set.

When the probability $(\Omega, \mathcal{F}, \mathbb{P})$ is complete, we say that the filtration $\{\mathcal{F}_t\}_{t \geq 0}$ satisfies the *usual conditions* if it is right continuous and includes all $\mathbb{P}$-null sets.

A.2 About Stochastic Processes.

We always work with a complete probability space equipped with a filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions, i.e. with the quadraple $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ which is also called a *stochastic basis*. Actually, in the main corpse of this thesis we consider the filtration $\{\mathcal{F}_t\}_{0\leq t\leq T}$ for $T > 0$.

Definition A.2.4 *[Stochastic Process] A collection of $\mathbb{R}^d$-valued random variables $\{X_t\}_{t\in I}$ is called a* stochastic process *with index set I and state space $\mathbb{R}^d$.* □

The index set we consider here is $I = [0, T]$ for $T > 0$ or in general $I = \mathbb{R}^+ = [0, \infty)$. For each fixed $t \in I$ we have that $X_t(\omega) : \Omega \to \mathbb{R}^d$ is a random variable whereas for each fixed $\omega \in \Omega$ the function $X_t(\omega) : I \to \mathbb{R}^d$ is called a *sample path* of the process.

In the following, we introduce various stochastic processes.

The $\mathbb{R}^d$-valued process $\{X_t\}_{t\geq 0}$ is called:

- *Continuous* (resp. *right continuous*, *left continuous*) if for almost all $\omega \in \Omega$ the function $X_t(\omega)$ is continuous (resp. right continuous, left continuous) on $t \geq 0$;
- *Càdlàg* (continue à droite limite à gauche) if it right continuous and for almost all $\omega \in \Omega$ the left limit $\lim_{s\uparrow t} X_s(\omega)$ exists and is finite for all $t > 0$;
- *Integrable* if X_t is an integrable r.v. for every $t \geq 0$;
- $\mathcal{F}_t$*-adapted* if X_t is $\mathcal{F}_t$-measurable for every $t \geq 0$;
- *Measurable* if regarded as a function of two variables, that is $X(t, \omega) : \mathbb{R}^+ \times \Omega \to \mathbb{R}^d$ is $\mathcal{B}(\mathbb{R}^+) \times \mathcal{F}$-measurable;
- *Progressive* if for every $T \geq 0$ it holds that $X(t, \omega) : [0, T] \times \Omega \to \mathbb{R}^d$ is $\mathcal{B}([0, T]) \times \mathcal{F}_T$-measurable;
- *Optional* (resp. *Predictable*) if it is $\mathcal{O}$-(resp. $\mathcal{P}$-)measurable, where $\mathcal{O}$(resp. $\mathcal{P}$) denotes the smallest σ-algebra on $\mathbb{R}^+ \times \Omega$ w.r.t. which every càdlàg adapted process (resp. left-continuous process) is a measurable function of (t, w).

A random variable $\tau : \Omega \to [0, \infty]$ is called a *stopping time* if $\{\omega : \tau(\omega) \leq t\} \in \mathcal{F}_t$ for every $t \geq 0$. We present a result concerning stopping times (see e.g. [Mao97, Th. 1.3.2]).

Theorem A.2.5 *[First Exit Time] Let $D \subset \mathbb{R}^d$ be an open set and define*

$$\tau := \inf\{t \geq 0 : X_t \notin D\},$$

with the convention $\inf \emptyset = \infty$. *Then τ is a stopping time called the* first exit time *from D.* □

Two important classes of adapted integrable stochastic processes are *martingales* and *Markov processes*. The process $\{M_t\}_{t\geq 0}$ is called a martingale (w.r.t. $\mathcal{F}_t$) if

$$\mathbb{E}(M_t|\mathcal{F}_s) = M_s \text{ a.s. for all } 0 \leq s < t < \infty \tag{A.2.1}$$

and $\{M_t\}_{t\geq 0}$ is called a Markov process whenever for given Borel measurable function $f(\cdot)$ it holds

$$\mathbb{E}(f(M_t)|\mathcal{F}_s) = \mathbb{E}(f(M_t)|M_s) \text{ a.s. for all } 0 \leq s < t < \infty.$$

If we replace the equality sign in (A.2.1) with $\leq$ we have a *supermartingale*, whereas the sign $\geq$ corresponds to a *submartingale*. A process that is either supermartingale or submartingale is called *semimartingale*.

Relation (A.2.1) suggests that by considering s as the current time, then the expected value of the process in a future time t conditional to the current information, is equal to the current value. This is a picture of a *fair game* where we can not lose or win in average. This property is used in the modeling of no-arbitrage in financial mathematics. We refer to [Man09] for a note about the origin of the word martingale.

Finally we introduce the notion of *quadratic variation* and *quadratic covariation*.

Definition A.2.6 *[Quadratic Variation] The quadratic variation of a stochastic process X_t with continuous sample paths $t \to X_t(\omega)$ is defined as the limit*

$$\langle X \rangle_T = \lim_{n\to\infty} \sum_{i=1}^{\infty} \left(X_{T_i^n}(\omega) - X_{T_{i-1}^n}(\omega) \right)^2.$$

□

Intuitively, the above limit can be written as a 'second order' integral

$$\langle X \rangle_T = \text{``}\int_0^T (dX_s(\omega))^2\text{''},$$

or in a differential form

$$d\langle X \rangle_t = \text{``}dX_t(\omega)dX_t(\omega)\text{''}.$$

A stochastic process whose trajectories $t \to X_t(\omega)$ are differentiable for almost all ω, satisfies $\langle X \rangle_t = 0$. In the case X is a deterministic process $t \to t$, such that $dX_t = dt$, we get the classic differential calculus result

$$dtdt = 0.$$

Definition A.2.7 *[Quadratic Covariation] The quadratic covariation of the stochastic processes X_t and Y_t with continuous sample paths is defined as the limit*

$$\langle X, Z \rangle_T = \lim_{n\to\infty} \sum_{i=1}^{\infty} \left(X_{T_i^n}(\omega) - X_{T_{i-1}^n}(\omega) \right) \left(Z_{T_i^n}(\omega) - Z_{T_{i-1}^n}(\omega) \right).$$

□

As before, we can intuitively write it as a 'second order' integral or in differential form

$$\langle X, Z \rangle_T = \text{``}\int_0^T dX_s(\omega)dZ_s(\omega)\text{''}, \text{ or } d\langle X, Z \rangle_t = \text{``}dX_t(\omega)dZ_t(\omega)\text{''}.$$

A.3 The Wiener Process.

The general type SDE (1.2.3) which we rewrite in differential form, highlighting the dependence on ω,

(A.3.2) $$dX_t(\omega) = a(t, X_t(\omega))dt + b(t, X_t(\omega))dW_t(\omega),$$

is defined through the increments $dW_t(\omega)$ of a process with continuous trajectories. This is the *Wiener process*[1] one of the two most important stochastic processes in the field of probability.[2]

[1] Also known as *Brownian Motion*. We shall adopt the name Wiener.
[2] The other stochastic process is the *Poisson process*, see e.g. [Bil86, Sec. 23].

The Wiener process has independent and Gaussian increments. This is rigorously stated in the following.

Definition A.3.8 *[Wiener Process] Let* $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{P})$ *be a stochastic basis and* $\{W_t\}_{t\geq 0}$ *be a real-valued continuous* $\mathcal{F}_t$*-adapted process. We call* $\{W_t\}_{t\geq 0}$ *a one-dimensional Wiener process if it has the following properties*

- $W_0 = 0$ *a.s.;*
- $W_t(\omega) - W_s(\omega)$ *is independent of* $\mathcal{F}_s$;
- $W_t(\omega) - W_s(\omega) \sim N(0, t-s)$, *(Gaussian increments),*

for all $0 \leq s < t < \infty$. *A* m*-dimensional process* $\{W_t\}_{t\geq 0} = (W_t^1, \dots, W_t^m)_{t\geq 0}$ *is called a* m*-dimensional Wiener process if each* $\{W_t^i\}$ *is a one-dimensional Wiener process and* $\{W_t^1\}, \dots, \{W_t^m\}$ *are independent.* □

As a consequence the Wiener process $\{W_t\}_{t\geq 0}$ satisfies the following:

- $W_u(\omega) - W_t(\omega)$ is independent of $W_t(\omega) - W_s(\omega)$, (independent increments),
- $W_{t+h}(\omega) - W_{s+h}(\omega) \sim W_t(\omega) - W_s(\omega)$, (stationary increments),

for all $0 \leq s < t < u < \infty$ and $h > 0$.

It can be shown that the Wiener process is a continuous square-integrable martingale with quadratic variation $\langle W \rangle_t = t$, for every $t \geq 0$ which can also be written, in a more 'relaxed' way as

$$dW_t(\omega)dW_t(\omega) = dt.$$

The above result is due to the fact that the Wiener process moves that fast, so that the second order terms can not be regarded as negligible. On the contrary, a process with differentiable trajectories can not move that fast, and therefore second order terms do not contribute.

Moreover, we have that $\langle W, t \rangle_t = 0$, for every $t \geq 0$ which can be rewritten 'informally' as

$$dW_t(\omega)dt = 0.$$

The trajectories of the Wiener process are almost nowhere differentiable. In particular, they are of unbounded variation and consequently the derivative $\dot{W}_t(\omega) = dW_t(\omega)/dt$ does not exist (see [Bre68].)

The m-dimensional Wiener process $\{W_t\}_{t\geq 0} = (W_t^1, \dots, W_t^m)_{t\geq 0}$ is a m-dimensional continuous martingale with joint quadratic variation given by $\langle W^i, W^j \rangle_t = \delta_{ij} t$ for $1 \leq i, j \leq m$, where δ_{ij} is the Dirac delta function, i.e.

$$\delta_{ij} = \begin{cases} 1, & \text{when } i = j, \\ 0, & \text{otherwise.} \end{cases}$$

The above property characterizes the Wiener process among continuous martingales as shown by Lévy's characterization theorem of Brownian motion, [KS88, Th. 3.16, p.157].

Theorem A.3.9 *[Lévy] Let M_t be a m-dimensional process, which is a martingale w.r.t. the filtration $\{\mathcal{F}_t\}_{t\geq 0}$, with continuous paths, $M_0 = 0$ a.s. and $\langle M^i, M^j \rangle_t = \delta_{ij} t$ for $1 \leq i, j \leq m$. Then M_t is a m-dimensional Wiener process w.r.t. $\{\mathcal{F}_t\}$.* □

A.4 Itô Integral.

The integral representation of (A.3.2) is

$$X_t(\omega) = X_0(\omega) + \int_0^t a(s, X_s(\omega))ds + \int_0^t b(s, X_s(\omega))dW_s(\omega). \tag{A.4.3}$$

Now we have to define the stochastic integral $\int_0^t b(s, X_s(\omega))dW_s(\omega)$. This integral can not be defined in the ordinary way as a Stieltjes integral for every path, since the variation of the paths is unbounded. However, under 'reasonable' assumptions ([Øks03, Ch. 3]), we can define the integral for a large class of stochastic processes, in a Stieltjes way, where now the integral depends on the intermediate points of the partitions used in the corresponding limit. In particular, we consider the interval $[0, T]$, and the following partition which depends on an integer n,

$$T_i^n = \min_{i=0,1,\dots,\infty} \left\{ T, \frac{i}{2^n} \right\}.$$

Note that $T_i^n = T$ for every $i > 2^n T$. The bigger the n the better discrete approximation of the continuous interval $[0, T]$.

Definition A.4.10 *We denote by $\mathcal{M}^2([0, T]; \mathbb{R})$ the family of all real-valued measurable, $\{\mathcal{F}_t\}$-adapted processes $\phi = \{\phi(t)\}_{0\leq t\leq T}$ such that*

$$||\phi||_{0,T}^2 = \mathbb{E} \int_0^T |\phi_s(\omega)|^2 ds < \infty.$$

We say that ϕ and $\hat{\phi}$ are equivalent if $||\phi - \hat{\phi}||^2_{0,T} = 0$. □

For $\phi \in \mathcal{M}^2([0,T];\mathbb{R})$ we define the integral as,

$$I(T) := \int_0^T \phi_s(\omega) dW_s(\omega) = \lim_{n\to\infty} \sum_{i=0}^{\infty} \phi_{t_i^n}(\omega) \left[W^n_{T_{i+1}}(\omega) - W^n_{T_i}(\omega) \right],$$

where t_i^n is an intermediate point of $[T_i^n, T_{i+1}^n)$. The choice $t_i^n := T_i^n$, i.e. the left endpoint of the above interval, corresponds to the *Itô integral*, whereas considering the midpoint $t_i^n := (T_i^n + T_{i+1}^n)/2$, defines the *Stratonovich integral*.

The Itô integral has interesting probabilistic properties - it is a martingale - but in the calculus it defines the classic chain rule is not valid. On the other hand, the Stratonovich integral, even though with less probabilistic interest, retains the chain rule and is preferred as regards the properties of the paths.

We present a classic example of a stochastic integral calculated in the Itô and Stratonovich sense

$$\int_0^t W_s(\omega) dW_s(\omega) = \begin{cases} \frac{W_t^2(\omega)}{2} - \frac{t}{2} & \text{Itô,} \\ \frac{W_t^2(\omega)}{2} & \text{Stratonovich.} \end{cases}$$

Usually the sign "$\circ$" is used in the Stratonovich case to distinguish between the two integrals $\int_0^t W_s(\omega) \circ dW_s(\omega)$. It is also possible to transform a SDE written in Itô form to one in Stratonovich form and conversely [Øks03, Ch. 3]

$$\begin{aligned} dX_t(\omega) &= a(t, X_t(\omega))dt + b(t, X_t(\omega)) \circ dW_t(\omega) \\ &= \hat{a}(t, X_t(\omega))dt + b(t, X_t(\omega))dW_t(\omega), \end{aligned}$$

where

$$\hat{a}(t,x) = a(t,x) + \frac{1}{2} b(t,x) \frac{\partial b}{\partial x}(t,x).$$

In the case where $b(t,x) = b(t)$ the drift coefficients of the corresponding Itô and Stratonovich SDEs are the same.

In general, stochastic integrals are defined in a Lebesgue way instead of a Riemann-Stieltjes way. We define the stochastic integral first for simple processes and then take the 'limit'. Finally we can extend the definition to the multi-dimensional case. For a detailed study of stochastic integration and its connection with SDEs we refer to [Øks03] and [RW87].

We collect some properties of the Itô integral $I(t) := \int_0^t \phi_s(\omega) dW_s(\omega)$:

- It is continuous as a function of t;
- It is $\mathcal{F}_t$-adapted;
- It is linear and additive;
- It is martingale with $\mathbb{E}(I(t)|\mathcal{F}_0) = 0$;
- $\mathbb{E}(I^2(t)|\mathcal{F}_0) = \int_0^t \mathbb{E}(\phi_s^2(\omega)|\mathcal{F}_0)ds$ (Itô isometry)

B. SOME USEFUL INEQUALITIES-RESULTS

In this section we collect some inequalities frequently used throughout this thesis as well as some theorems mainly from the field of probability and stochastic processes.

B.1 Probability Related Results.

First, we present the *Cauchy-Schwarz inequality* (see [HLP52, Th. 181])

$$\left(\int XY ds\right)^2 \leq \left(\int X^2 ds\right)\left(\int Y^2 ds\right), \tag{B.1.1}$$

where we have suppressed the limits of integration and the *Young inequality* (see [HLP52, Th. 61])

$$ab \leq \frac{a^r}{r} + \frac{b^q}{q}, \tag{B.1.2}$$

for non-negative a, b and conjugate exponents r, q.

In the following, we present *Hölder's inequality* (see [HLP52, Th. 189]) and a simple application of it:

For $X \in \mathcal{L}^p, Y \in \mathcal{L}^q$ and r, p, q such that $1/p + 1/q = 1/r$ with $0 < q < \infty, 0 < r \leq p$ it holds that

$$\|X \cdot Y\|^p_{\mathcal{L}^r(\Omega;\mathbb{R})} \leq \|X\|^p_{\mathcal{L}^p(\Omega;\mathbb{R})} \|Y\|^p_{\mathcal{L}^q(\Omega;\mathbb{R})}. \tag{B.1.3}$$

For $r = 1$ and $p > 1$ inequality (B.1.3) becomes

$$|\mathbb{E}(X^T \cdot Y)| \leq (\mathbb{E}|X|^p)^{1/p}(\mathbb{E}|Y|^q)^{1/q}$$

and for $Y \equiv 1$ (B.1.3) implies

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

Now, we state two integration convergence theorems, see e.g. [Mao97, Th. 1.2.2 & 1.2.3].

Theorem B.1.1 *[Monotone Convergence] For an increasing sequence $\{X_n\}$ of non-negative random variables it holds*

$$\lim_{n\to\infty} \mathbb{E}X_n = \mathbb{E}(\lim_{n\to\infty} X_n).$$

□

Theorem B.1.2 *[Dominated Convergence] Let $\{X_n\} \in \mathcal{L}^p(\Omega, \mathbb{R}^d)$ and $Y \in \mathcal{L}^p(\Omega, \mathbb{R})$ for some $p \geq 1$ such that $|X_n| \leq Y$ a.s. and let $\{X_n\}$ converge* in probability *to X, i.e. for every $\epsilon > 0, \mathbb{P}(\omega : |X_n(\omega) - X(\omega)| > \epsilon) \to 0$ as $n \to \infty$. Then*

- $X \in \mathcal{L}^p(\Omega; \mathbb{R}^d)$;
- $\{X_n\}$ *converges to X in $\mathcal{L}^p$, i.e. $\mathbb{E}|X_n - X|^p \to 0$ as $n \to \infty$;*
- $\lim_{n\to\infty} \mathbb{E}X_n = \mathbb{E}X$.

□

B.2 Stochastic Processes Related Results.

The following inequality is an application of the well-known *Doob inequality*, see e.g. [KS88, Th. 1.3.8].

Theorem B.2.3 *[Doob's Martingale Inequality] Let $\{M_t\}_{t\geq 0}$ be a martingale such that $M_t \in \mathcal{L}^p(\Omega, \mathbb{R}^d)$ for some $p > 1$. Then*

$$\mathbb{E} \sup_{0\leq t\leq T} |M_t|^p \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}|M_T|^p, \qquad \text{(B.2.4)}$$

for $T > 0$. □

B.3 Moment-Integral Inequalities.

The next result is known as Burkholder-Davis-Gundy inequality, see e.g. [Mao97, Th. 1.7.3], [KS88, Th. 3.28].

Theorem B.3.4 *[BDG Inequality] Let $\phi \in \mathcal{L}^2(\mathbb{R}^+; \mathbb{R}^{d\times m})$. Then for every $p > 0$ there exist positive constants c_p, C_p such that*

$$c_p \mathbb{E} \left| \int_0^T \phi_s^2 ds \right|^{p/2} \leq \mathbb{E} \sup_{0\leq t\leq T} \left| \int_0^t \phi_s dW_s \right|^p \leq C_p \mathbb{E} \left| \int_0^T \phi_s^2 ds \right|^{p/2}, \tag{B.3.5}$$

for $T \geq 0$. □

The universal constants c_p, C_p may be chosen in the way shown in Table B.1.

Values of p	c_p	C_p
$(0,2)$	$(p/2)^p$	$(32/p)^{p/2}$
2	1	4
$(2,\infty)$	$(2p)^{-p/2}$	$[p^{p+1}/2(p-1)^{p-1}]^{p/2}$

Tab. B.1: Universal constants in the BDG inequality.

The following integral inequality has been used in theory of ODEs and SDEs for the proof of existence, uniqueness, boundness and comparison results between other applications. It is also important in this thesis. It goes back to 1919, see [Gro19, (7)].

Theorem B.3.5 *[Gronwall's Inequality] Let $u(\cdot)$ be a Borel measurable bounded non-negative function on $[0,T]$, where $T > 0$ and let $v(\cdot)$ be a nonnegative integrable function on $[0,T]$. If*

$$u(t) \leq c + \int_0^t v(s)u(s)ds,$$

for every $0 \leq t \leq T$ where $c \geq 0$, then

$$u(t) \leq c \exp\left\{ \int_0^t v(s)ds \right\}. \tag{B.3.6}$$

□

C. SUPPLEMENTARY FOR CHAPTER 2.

C.1 Existence and Uniqueness of y_t^{SD} for the Heston 3/2-model.

C.1.1 Uniqueness of solution of y_t^{SD}.

Let $y_t, \hat{y}_t$ be two solutions of SDE (2.5.3) with same initial condition, i.e. with $y_0 = \hat{y}_0$. By Lemma 2.4.8 they both belong to the space $\mathcal{M}^2([0,T];\mathbb{R})$ of measurable $\mathcal{F}_t$-adapted processes z such that

$$\mathbb{E}\int_0^T |z_s|^2 ds < \infty.$$

Set the stopping times $\theta_R^i = \inf\{t \in [t_{i-1}, t_i] : |y_t| > R\}$ and $\hat{\theta}_R^i = \inf\{t \in [t_{i-1}, t_i] : |\hat{y}_t| > R\}$ for some $R > 0$ big enough and consider the stopping times $\tau_R^i = \theta_R^i \wedge \hat{\theta}_R^i$, for $i = 1, ..., N$. Take $t \in [0, t_1]$ and $e_{t\wedge\tau_R^1} := y_{t\wedge\tau_R^1} - \hat{y}_{t\wedge\tau_R^1}$.

It holds that

$$\begin{aligned}
|e_{t\wedge\tau_R^1}|^2 &= \Big|\int_0^{t\wedge\tau_R^1} (f(\hat{s},s,y_{\hat{s}},y_s)-f(\hat{s},s,\hat{y}_{\hat{s}},\hat{y}_s))\,ds \\
&\quad + \int_0^{t\wedge\tau_R^1} (g(\hat{s},s,y_{\hat{s}},y_s)-g(\hat{s},s,\hat{y}_{\hat{s}},\hat{y}_s))\,dW_s\Big|^2 \\
&\leq 2t_1\int_0^{t\wedge\tau_R^1} \Big|f(\hat{s},s,y_{\hat{s}},y_s)-f(\hat{s},s,\hat{y}_{\hat{s}},\hat{y}_s)\Big|^2 ds + 2|M_t|^2 \\
&\leq 6t_1C_R^2\int_0^{t\wedge\tau_R^1} \big(|y_{\hat{s}}-\hat{y}_{\hat{s}}|^2+|y_s-\hat{y}_s|^2+|y_{\hat{s}}-\hat{y}_{\hat{s}}|\big)\,ds + 2|M_t|^2 \\
&\leq 6t_1C_R^2\int_0^{t} |e_{s\wedge\tau_R^1}|^2 ds + 2|M_t|^2,
\end{aligned}$$

where in the second step we have used the Cauchy-Schwarz inequality, in the third step the elementary inequality $(\sum_{i=1}^3 a_i)^2 \leq 3\sum_{i=1}^3 a_i^2$, for appropriate a_i and Assumption 2.2.1 for f, in the last step the fact that $\hat{s}=0$, when $s\in[0,t_1]$, and the equality in the initial conditions $y_0=\hat{y}_0$. Furthermore,

$$M_t := \int_0^{t\wedge\tau_R^1} (g(\hat{s},s,y_{\hat{s}},y_s)-g(\hat{s},s,\hat{y}_{\hat{s}},\hat{y}_s))\,dW_s.$$

Taking the supremum over all $t\in[0,t_1]$ and then expectations we have

$$\begin{aligned}
\mathbb{E}\sup_{0\leq t\leq t_1}|e_{t\wedge\tau_R^1}|^2 &\leq 6t_1C_R^2\mathbb{E}\sup_{0\leq t\leq t_1}\left(\int_0^{t\wedge\tau_R^1}|y_s-\hat{y}_s|^2ds\right) + 2\mathbb{E}\sup_{0\leq t\leq t_1}|M_t|^2 \\
&\leq 6t_1C_R^2\int_0^{t_1}\mathbb{E}\sup_{0\leq l\leq s}|e_{l\wedge\tau_R^1}|^2ds + 2\mathbb{E}|M_{t_1}|^2, \qquad \text{(C.1.1)}
\end{aligned}$$

where we have used Doob's maximal inequality with $p = 2$, since M_t is an $\mathbb{R}$-valued martingale that belongs to $\mathcal{L}^2$. Moreover, we have that

$$\begin{aligned}
\mathbb{E}|M_{t_1}|^2 &:= \mathbb{E}\left|\int_0^{t_1\wedge\tau_R^1} \left(g(\hat{s}, s, y_{\hat{s}}, y_s) - g(\hat{s}, s, \hat{y}_{\hat{s}}, \hat{y}_s)\right) dW_s\right|^2 \\
&= \mathbb{E}\left(\int_0^{t_1\wedge\tau_R^1} \left(g(\hat{s}, s, y_{\hat{s}}, y_s) - g(\hat{s}, s, \hat{y}_{\hat{s}}, \hat{y}_s)\right)^2 ds\right) \\
&\leq 3C_R^2\mathbb{E}\left(\int_0^{t_1\wedge\tau_R^1} \left(|y_0 - \hat{y}_0|^2 + |y_s - \hat{y}_s|^2 + |y_0 - \hat{y}_0|\right) ds\right) \\
&\leq 3C_R^2\int_0^{t_1\wedge\tau_R^1} \mathbb{E}|y_s - \hat{y}_s|^2 ds \leq 3C_R^2\int_0^{t_1} \mathbb{E}\sup_{0\leq l\leq s}|e_{l\wedge\tau_R^1}|^2 ds,
\end{aligned}$$

where we have used Assumption 2.2.1 for g. Thus relation (C.1.1) becomes

$$\mathbb{E}\sup_{0\leq t\leq t_1}|e_{t\wedge\tau_R^1}|^2 \leq (6t_1C_R^2 + 3C_R^2)\int_0^{t_1}\mathbb{E}\sup_{0\leq l\leq s}|e_{l\wedge\tau_R^1}|^2 ds,$$

which by use of Gronwall's inequality gives

$$\mathbb{E}\sup_{0\leq t\leq t_1}|e_{t\wedge\tau_R^1}|^2 = 0.$$

Following the same arguments we can show that

$$\mathbb{E}\sup_{0\leq t\leq t_1}|e_{t\wedge\tau_R^i}|^2 = 0,$$

for every integer $1 \leq i \leq N$.[1] Thus, if we drop the index i from the stopping times with the meaning that $\theta_R = \inf\{t \in [0, T] : |y_t| > R\}$ and $\hat{\theta}_R = \inf\{t \in [0, T] : |\hat{y}_t| > R\}$ for some $R > 0$ big enough and consider the stopping time $\tau_R = \theta_R \wedge \hat{\theta}_R$, we have that

$$\mathbb{E}\sup_{0\leq t\leq T}|e_{t\wedge\tau_R}|^2 \leq \sum_{i=1}^{N}\mathbb{E}\sup_{t_{i-1}\leq t\leq t_i}|e_{t\wedge\tau_R^i}|^2 = 0.$$

Hence, $y_t = \hat{y}_t$ for all $0 \leq t \leq T$ a.s. which proves that the solution of SDE (2.5.3), and in general of SDE (2.2.1) when it exists, is unique.

[1] For $i = 2$ just use the same ideas as for $i = 1$ and the other cases follow exactly the same way using in every step the result of the previous step.

C.1.2 Existence of solution of y_t^{SD}.

We will show the existence of the solution of SDE (2.5.2) for $n = 0$ and the same procedure can be followed to show the existence of the solution of SDE (2.5.2) for every integer $n = 1, .., N-1$, i.e. the existence of the solution of SDE (2.5.3). Application of Itô's formula to $\ln y_t$, for $0 \le t \le t_1$ implies

$$\begin{aligned} \ln y_t = \ln y_0 + \int_0^t \frac{1}{y_s}(k_1(s) - k_2(s)y_0)y_s ds + \frac{1}{2}\int_0^t \left(-\frac{1}{y_s^2}\right) k_3^2(s)y_0 y_s^2 ds \\ + \int_0^t \frac{1}{y_s} k_3(s) y_0 y_s dW_s \\ = \ln y_0 + \int_0^t \left(k_1(s) - k_2(s)y_0 - \frac{k_3^2(s)}{2}\sqrt{y_0}\right) ds + \int_0^t k_3(s)\sqrt{y_0} dW_s. \end{aligned}$$

Now take the exponential of both sides of (2.4.4) with $\hat{s} = 0$ in the case $0 \le t \le t_1$ to verify that (2.5.5) is indeed a solution of SDE (2.5.2) for $n = 0$.

C.2 Proof of Lemma 2.4.10.

Set the stopping time $\theta_R = \inf\{t \in [0,T] : x_t^{-1} > R\}$, for some $R > 0$, with the convention that $\inf \emptyset = \infty$. Application of Itô's formula on $(x_{t\wedge\theta_R})^{-2}$ implies,

$$\begin{aligned} (x_{t\wedge\theta_R})^{-2} = (x_0)^{-2} + \int_0^{t\wedge\theta_R} (-2)x_s^{-3}(k_1(s)x_s - k_2(s)x_s^{2r-1})ds \\ + \int_0^{t\wedge\theta_R} \frac{(-2)(-3)}{2}(x_s)^{-4}k_3^2(s)x_s^{2r} ds + \int_0^{t\wedge\theta_R} (-2)k_3(s)(x_s)^{-3}x_s^r dW_s \\ = (x_0)^{-2} + \int_0^{t\wedge\theta_R} (-2)k_1(s)x_s^{-2} + 2k_2(s)x_s^{2r-4} + 3k_3^2(s)x_s^{2r-4})ds \\ + \int_0^t (-2)k_3(s)x_s^{r-3}\,\mathbb{I}_{(0,t\wedge\theta_R)}(s)dW_s \\ = \int_0^{t\wedge\theta_R} \left(-2k_1(s)x_s^{-2} + (2k_2(s) + 3k_3^2(s))\left(x_s^{2r-4}\mathbb{I}_{(0,1]}(x_s) + x_s^{2r-4}\mathbb{I}_{(1,\infty]}(x_s)\right)\right) ds \\ +(x_0)^{-2} + M_t \\ \le (x_0)^{-2} + (2k_{2,\max} + 3k_{3,\max}^2)T + \int_0^t (2k_2(s) + 3k_3^2(s))x_s^{-2}\,\mathbb{I}_{(0,t\wedge\theta_R)}(s)ds + M_t, \end{aligned}$$

where

$$M_t := \int_0^t (-2)k_3(s)x_s^{r-3}\,\mathbb{I}_{(0,t\wedge\theta_R)}(s)dW_s.$$

Taking expectations in the above inequality and using the fact that $\mathbb{E}M_t = 0$,[2] we get that

$$\begin{aligned}\mathbb{E}(x_{t\wedge\theta_R})^{-2} &\leq \mathbb{E}(x_0)^{-2} + (2k_{2,\max} + 3k_{3,\max}^2)T + (2k_{2,\max} + 3k_{3,\max}^2)\int_0^t \mathbb{E}(x_{s\wedge\theta_R})^{-2}ds\\ &\leq \left(\mathbb{E}(x_0)^{-2} + 2k_{2,\max}T + 3k_{3,\max}^2T\right)e^{(2k_2+3k_3^2)T} < C,\end{aligned}$$

where we have used Gronwall's inequality with C independent of R. We have that

$$(x_{t\wedge\theta_R})^{-2} = (x_{\theta_R})^{-2}\mathbb{I}_{(\theta_R\leq t)} + (x_t)^{-2}\mathbb{I}_{(t<\theta_R)} = R^2\mathbb{I}_{(\theta_R\leq t)} + (x_t)^{-2}\mathbb{I}_{(t<\theta_R)}. \tag{C.2.2}$$

By relation (C.2.2) we have that,

$$\mathbb{E}\left(\frac{1}{x_{t\wedge\theta_R}^2}\right) = R^2\mathbb{P}(\theta_R\leq t) + \mathbb{E}\left(\frac{1}{x_t^2}\mathbb{I}_{(t<\theta_R)}\right) < C,$$

thus

$$\mathbb{P}(x_t\leq 0) = \mathbb{P}\left(\bigcap_{R=1}^{\infty}\left\{x_t < \frac{1}{R}\right\}\right) = \lim_{R\to\infty}\mathbb{P}\left(\left\{x_t < \frac{1}{R}\right\}\right) \leq \lim_{R\to\infty}\mathbb{P}(\theta_R\leq t) = 0.$$

We conclude that $x_t > 0$ a.s.

C.3 *Proof of Lemma 2.4.11.*

We work as in the proof of Lemma 2.4.7. In particular, we first get the bound

$$J(s,x) \leq \frac{k_{1,\max}x^2 + \left(0.5(p-1)(k_{3,\max})^2 - k_{2,\min}\right)x^{2r}}{1+x^2} \leq k_{1,\max},$$

valid for all p such that $p \leq 1 + 2k_{2,\min}/(k_{3,\max})^2$, where $J(s,x)$ is as in the proof of Lemma 2.4.7, which in turn implies,

$$\mathbb{E}(x_t)^p \leq 2^{(p-2)/2}(1+\mathbb{E}(x_0)^p)e^{Cpt},$$

[2] The function $h(u) = (-2)k_3(u)x_u^{r-3}\,\mathbb{I}_{(0,t\wedge\theta_R)}(u)$ belongs to the space $\mathcal{M}^2([0,t];\mathbb{R})$ thus [Mao97, Th. 1.5.8] implies $\mathbb{E}M_t = 0$.

for any $2 < p \leq 1+2k_{2,\min}/(k_{3,\max})^2$ and all $t \in [0,T]$. Using Itô's formula on $(x_t)^p$, with $p \leq \frac{3}{2} - r + \frac{k_{2,\min}}{(k_{3,\max})^2}$ (in order to use Doob's martingale inequality later) we have that

$$\begin{aligned}(x_t)^p &\leq (x_0)^p + p\int_0^t \left[k_1(s)(x_s)^p + \left(\frac{p-1}{2}k_{3,\max}^2 K_\phi^2 - k_2\right)(x_s)^{p+2r-2}\right] ds + M_t \\ &\leq (x_0)^p + p\int_0^t k_1(s)(x_s)^p ds + M_t,\end{aligned}$$

where $M_t = \int_0^t pk_3(s)(x_s)^{p+2r-1}dW_s$. Taking the supremum and then expectations in the above inequality we get

$$\mathbb{E}(\sup_{0\leq t\leq T}(x_t)^p) \leq \left(\mathbb{E}(x_0)^p + \sqrt{4\mathbb{E}M_T^2}\right)e^{pk_{1,\max}T} := A_1,$$

where in the last step we have used Doob's martingale inequality to the diffusion term M_t[3] and Gronwall's inequality.

C.4 Proof of Lemma 2.4.16.

Set the stopping time $\theta_R = \inf\{t \in [0,T] : x_t^{-1} > R\}$, for some $R > 0$, with the convention that $\inf\emptyset = \infty$. Application of Itô's formula on $\ln x_{t\wedge\theta_R}$ implies,

$$\begin{aligned}\ln x_{t\wedge\theta_R} &= \ln x_0 + \int_0^{t\wedge\theta_R}\frac{1}{x_s}(k_1(s)x_s - k_2(s)x_s^q)ds \\ &\quad + \int_0^{t\wedge\theta_R}\left(-\frac{1}{x_s^2}\right)k_3^2(s)x_s^{2r}\phi^2(x_s)ds + \int_0^{t\wedge\theta_R}\frac{1}{x_s}k_3(s)x_s^r\phi(x_s)dW_s \\ &= \ln x_0 + \int_0^{t\wedge\theta_R}\left(k_1(s) - k_2(s)x_s^{q-1} - k_3^2(s)x_s^{2r-2}\phi^2(x_s)\right)ds \\ &\quad + \int_0^{t\wedge\theta_R}k_3(s)x_s^{r-1}\phi(x_s)dW_s.\end{aligned}$$

[3] The function $h(u) = pk_3(u)\phi(x_u)(x_u)^{p+2r-1}$ belongs to the family $\mathcal{M}^2([0,T];\mathbb{R})$ thus [Mao97, Th. 1.5.8] implies $\mathbb{E}M_t^2 = \mathbb{E}(\int_0^t h(u)dW_u)^2 = \mathbb{E}\int_0^t h^2(u)du$, i.e. $M_t \in \mathcal{L}^2(\Omega;\mathbb{R})$.

Taking absolute values in the above equality, then expectations, and using Jensen's inequality and then Itô's isometry on the diffusion term M_t, we get

$$\begin{aligned}
\mathbb{E}|\ln x_{t\wedge\theta_R}| &\leq \mathbb{E}|\ln x_0| + T(|k_{1,\max}| + |k_{2,\max}|\mathbb{E}\sup_{0\leq t\leq T}|x_t|^{q-1} \\
&\quad + |k_{3,\max}|^2 K_\phi^2 \mathbb{E}\sup_{0\leq t\leq T}|x_t|^{2r-2}) + \mathbb{E}|M_t| \\
&\leq \mathbb{E}|\ln x_0| + (|k_{1,\max}| + (|k_{2,\max}| + |k_{3,\max}|^2)A_1 + |k_{3,\max}|^2 K_\phi^2)T \\
&\quad + \sqrt{4\mathbb{E}M_T^2} < C,
\end{aligned}$$

where A_1 is as in Lemma 2.4.14 and $M_t := \int_0^t k_3(s)x_s^{r-1}\phi(x_s)\,\mathbb{I}_{(0,t\wedge\theta_R)}(s)dW_s$. Now, we proceed as in Lemmata 2.4.6 and 2.4.10, to get $\mathbb{P}(\theta_R \leq t) \downarrow 0$ as $R \to \infty$ and then conclude that $\mathbb{P}(x_t \leq 0) \leq 0$, i.e. $x_t > 0$ a.s.

C.5 Proof of Lemma 2.4.14.

In the case all x are outside a finite ball of radius $R > 1$, and $s \in [0, T]$ we have that

$$\begin{aligned}
J(s,x) &= \frac{x(k_1(s)x - k_2(s)x^q) + (p-1)k_3^2(s)[x^r\phi(x)]^2/2}{1+x^2} \\
&= \frac{k_1(s)x^2 - k_2(s)x^{q+1} + 0.5(p-1)k_3^2(s)x^{2r}\phi^2(x)}{1+x^2} \leq k_{1,\max},
\end{aligned}$$

where the the last inequality is valid for all $p > 2$ and we have used $q+1 > 2r$ and that q is odd. Thus $J(s,x)$ is bounded for all $(s,x) \in [0,T]\times\mathbb{R}$, since when $|x| \leq R$ we have that $J(s,x)$ is finite, say $J(s,x) \leq C$. Application of [Mao97, Th. 2.4.1] implies

$$\mathbb{E}|x_t|^p \leq 2^{(p-2)/2}(1 + \mathbb{E}|x_0|^p)e^{Cpt},$$

for any $2 < p$ and all $t \in [0,T]$. Using Itô's formula on $|x_t|^p$, we have that

$$\begin{aligned} |x_t|^p &= |x_0|^p + \int_0^t \frac{p}{2}\left(|x_s|^{p-2} + (p-2)|x_s|^{p-4}x_s^2\right)[k_3(s)x_s^r\phi(x_s)]^2 ds \\ &\quad + \int_0^t p|x_s|^{p-2}x_s(k_1(s)x_s - k_2(s)x_s^q)ds + \int_0^t pk_3(s)|x_s|^{p-2}x_s x_s^r\phi(x_s)dW_s \\ &\leq |x_0|^p + p\int_0^t \left[k_1(s) - k_2(s)(x_s)^{q-1} + \frac{p-1}{2}k_3^2(s)K_\phi^2(x_s)^{2r-2}\right]|x_s|^p ds \\ &\quad + \underbrace{\int_0^t pk_3(s)\phi(x_s)|x_s|^p(x_s)^{r-1}dW_s}_{M_t} \\ &\leq |x_0|^p + C\int_0^t |x_s|^p ds + M_t, \end{aligned}$$

where we have used that $0 < 2r-2 < q-1$ and q is odd. Taking the supremum and then expectations in the above inequality we get

$$\begin{aligned} \mathbb{E}(\sup_{0\leq t\leq T}|x_t|^p) &\leq \mathbb{E}|x_0|^p + C\mathbb{E}\left(\sup_{0\leq t\leq T}\int_0^t |x_s|^p ds\right) + \mathbb{E}\sup_{0\leq t\leq T} M_t \\ &\leq \mathbb{E}|x_0|^p + C\int_0^t \mathbb{E}(\sup_{0\leq l\leq s}|x_l|^p)ds + \sqrt{\mathbb{E}\sup_{0\leq t\leq T} M_t^2} \\ &\leq \left(\mathbb{E}|x_0|^p + \sqrt{4\mathbb{E}M_T^2}\right)e^{CT} := A_1, \end{aligned}$$

where in the last step we have used Doob's martingale inequality to the diffusion term M_t[4] and the Gronwall inequality.

C.6 Proof of Lemma 2.4.17.

Set the stopping time $\theta_R = \inf\{t \in [0,T] : y_t > R\}$, for some $R > 0$, with the convention that $\inf\emptyset = \infty$. Application of Itô's formula on $(y_{t\wedge\theta_R})^p$, implies,

[4] The function $h(u) = pk_3(u)\phi(x_u)|x_u|^p x_s^{r-1}$ belongs to the family $\mathcal{M}^2([0,T];\mathbb{R})$ thus [Mao97, Th. 1.5.8] implies $\mathbb{E}M_t^2 = \mathbb{E}(\int_0^t h(u)dW_u)^2 = \mathbb{E}\int_0^t h^2(u)du$, i.e. $M_t \in \mathcal{L}^2(\Omega;\mathbb{R})$.

$$
\begin{aligned}
&(y_{t\wedge\theta_R})^p = (y_0)^p + \int_0^{t\wedge\theta_R} pk_3(s)(y_s)^{p-1}y_{\hat{s}}^{r-1}\phi(y_{\hat{s}})y_s dW_s \\
&+ \int_0^{t\wedge\theta_R} p(y_s)^{p-1}(k_1(s) - k_2(s)y_{\hat{s}}^{q-1})y_s + \frac{p(p-1)}{2}(y_s)^{p-2}\left[k_3(s)y_{\hat{s}}^{r-1}\phi(y_{\hat{s}})y_s\right]^2 ds \\
= \; & (x_0)^p + \int_0^{t\wedge\theta_R}\left(p(k_1(s) - k_2(s)y_{\hat{s}}^{q-1}) + \frac{p(p-1)k_3^2(s)}{2}y_{\hat{s}}^{2r-2}\phi^2(y_{\hat{s}})\right)(y_s)^p ds \\
&+ \int_0^{t\wedge\theta_R} pk_3(s)y_{\hat{s}}^{r-1}\phi(y_{\hat{s}})(y_s)^p dW_s \\
\leq \; & \int_0^t \left[-k_2(s)(y_{\hat{s}})^{q-1} + \frac{p-1}{2}k_{3,\max}^2 K_\phi^2 y_{\hat{s}}^{2r-2} + k_{1,\max}\right](y_s)^p\,\mathbb{I}_{(0,t\wedge\theta_R)}(s)ds \\
&+(x_0)^p + M_t \\
\leq \; & (x_0)^p + C\int_0^t (y_s)^p\,\mathbb{I}_{(0,t\wedge\theta_R)}(s)ds + M_t,
\end{aligned}
$$

where we have used that $q-1 > 2r-2 > 1$, the last inequality is valid for $p > 2$, the constant C is independent of R and $M_t := \int_0^{t\wedge\theta_R} pk_3(s)y_{\hat{s}}^{r-1}\phi(y_{\hat{s}})(y_s)^p dW_s$. Taking expectations and using that $\mathbb{E}M_t = 0$ we get

$$
\begin{aligned}
\mathbb{E}(y_{t\wedge\theta_R})^p &\leq \mathbb{E}(x_0)^p + C\int_0^t \mathbb{E}(y_{s\wedge\theta_R})^p ds \\
&\leq \mathbb{E}(x_0)^p e^{CT},
\end{aligned}
$$

where in the second step we have applied Gronwall's inequality. We have that

$$
(y_{t\wedge\theta_R})^p = (y_{\theta_R})^p\mathbb{I}_{(\theta_R\leq t)} + (y_t)^p\mathbb{I}_{(t<\theta_R)} = R^p\mathbb{I}_{(\theta_R\leq t)} + (y_t)^p\mathbb{I}_{(t<\theta_R)},
$$

thus taking expectations in the above inequality and using the estimated upper bound for $\mathbb{E}(y_{t\wedge\theta_R})^p$ we arrive at

$$
\mathbb{E}(y_t)^p\mathbb{I}_{(t<\theta_R)} \leq \mathbb{E}(x_0)^p e^{CT}
$$

and taking limits in both sides as $R\to\infty$ we get that

$$
\lim_{R\to\infty}\mathbb{E}(y_t)^p\mathbb{I}_{(t<\theta_R)} \leq \mathbb{E}(x_0)^p e^{CT}.
$$

Fix t. The sequence $(y_t)^p\mathbb{I}_{(t<\theta_R)}$ is non-decreasing in R since θ_R is increasing in R and $t\wedge\theta_R \to t$ as $R\to\infty$ and $(y_t)^p\mathbb{I}_{(t<\theta_R)} \to (y_t)^p$ as $R\to\infty$, thus the monotone convergence theorem [Mao97, Th. 1.2.2] implies

$$
\mathbb{E}(y_t)^p \leq \mathbb{E}(x_0)^p e^{CT},
$$

for any $2 < p$. Following the same lines as in Lemma 2.4.14, i.e. using again Itô's formula on $(y_t)^p$, taking the supremum and then using Doob's martingale inequality on the diffusion term we obtain the desired result.

D. NUMERICAL SCHEMES FOR THE INTEGRATION OF THE VARIANCE-VOLATILITY PROCESS (V_T).

We consider a partition of the time interval $[0, T]$ with $0 = t_0 < t_1 < ... < t_N = T$ and discretization steps $\Delta_n := t_{n+1} - t_n$ for $n = 0, \dots, N-1$. Moreover, we denote by $\Delta W_n := W_{t_{n+1}} - W_{t_n}$ the increments of the Brownian motion. We show in the following subsections some numerical schemes for the approximation of

$$V_t = V_0 + \int_0^t (k_1 - k_2 V_s) ds + \int_0^t k_3 (V_s)^q dW_s, \quad t \in [0, T]$$

and make some brief comments on them. We also denote $V_n := V_{t_n}$.

D.1 Standard Euler-Maruyama scheme.

The explicit Euler-Maruyama (EM) scheme for the process (V_t) is given by

$$V_{n+1}^{EM} = V_n + (k_1 - k_2 V_n)\Delta_n + k_3 (V_n)^q \Delta W_n, \tag{D.1.1}$$

for $n = 0, \dots, N-1$. Clearly $\mathbb{P}(V_{n+1} < 0 | V_n > 0) > 0$, thus the EM scheme can produce negative values with positive probability, or in the notion of [Sch96] we say that (D.1.1) has a finite life time.

Standard Milstein scheme.

The standard one dimensional Milstein (M) scheme contains some extra terms derived by Itô-Taylor expansion [KP95, Sec. 5], and applied to (V_t)

reads

$$V_{n+1}^{M} = V_n + (k_1 - k_2 V_n)\Delta_n + k_3 (V_n)^q \Delta W_n + \frac{1}{2}(k_3)^2 q (V_n)^{2q-1}\left((\Delta W_n)^2 - \Delta_n\right),$$

for $n = 0, \ldots, N-1$ where we have retained terms of order (Δ_n). Again (M) scheme has a finite life time.

D.2 Balanced Implicit Method.

The balanced implicit method (BIM) [MPS98, (3.2)] was the first attempt to treat the problem of invariance-preserving of specific domains of the underlying process and reads

$$\begin{aligned} V_{n+1}^{BIM} &= V_n + (k_1 - k_2 V_n)\Delta_n + k_3 (V_n)^q \Delta W_n \\ &\quad + \left(c^0(V_n)\Delta_n + c^1(V_n)|\Delta W_n|\right)(V_n - V_{n+1}), \end{aligned}$$

for $n = 0, \ldots, N-1$, where c^0 and c^1 are appropriate weight functions. The choice $c^0(x) = k_2$ and $c^1(x) = k_3 x^{q-1}$ preserves positivity [KS06, Sec. 5]. Rearranging the above equation, we get the expression

$$V_{n+1}^{BIM} = \frac{V_n + k_1\Delta_n + k_3(V_n)^q(\Delta W_n + |\Delta W_n|)}{1 + k_2\Delta_n + k_3(V_n)^{q-1}|\Delta W_n|}.$$

D.3 Balanced Milstein Method.

The balanced Milstein method (BMM), was proposed in [KS06], for an improvement of the BIM in the stability behavior as well as in the rate of convergence. It is given by the following linear implicit relation

$$\begin{aligned} V_{n+1}^{BMM} &= V_n + (k_1 - k_2 V_n)\Delta_n + k_3 (V_n)^q \Delta W_n \\ &\quad + \frac{1}{2}(k_3)^2 q (V_n)^{2q-1}\left((\Delta W_n)^2 - \Delta_n\right) \\ &\quad + \left(d^0(V_n)\Delta_n + d^1(V_n)((\Delta W_n)^2 - \Delta_n)\right)(V_n - V_{n+1}), \end{aligned}$$

for $n = 0, \ldots, N-1$, where d^0 and d^1 are appropriate weight functions. The choice $d^0(x) = \Theta k_2 + \frac{1}{2}(k_3)^2 q |x|^{2q-2}$, where $\Theta \in [0,1]$ and $d^1(x) = 0$ implies an eternal life time for the scheme [KS06, Th. 5.9], in the sense that $\mathbb{P}(V_{n+1} > 0 | V_n > 0) = 1$. The step sizes Δ_n have to be such that $\Delta_n < \frac{2q-1}{2qk_2(1-\Theta)}$. The relaxation parameter resembles to the implicitness parameter

(θ in our notation). For $\Theta = 1$ there is no restriction in the step size, but it is recommended when possible [KS06, Rem. 5.10] to take $\Theta = 1/2$. Rearranging with the above specifications leads to

$$V_{n+1}^{BMM} = \frac{V_n + (k_1 - (1-\Theta)k_2 V_n)\Delta_n + k_3 (V_n)^q \Delta W_n + \frac{1}{2}(k_3)^2 q (V_n)^{2q-1}(\Delta W_n)^2}{1 + \Theta k_2 \Delta_n + \frac{1}{2}(k_3)^2 q |V_n|^{2q-2}\Delta_n}.$$

Finally, the proposed semi-discrete (SD) scheme reads

$$\begin{aligned} V_{n+1}^{SD} &= \Bigg(\sqrt{V_n\left(1 - \frac{k_2\Delta}{1+k_2\theta\Delta}\right) + \frac{k_1\Delta}{1+k_2\theta\Delta} - \frac{(k_3)^2\Delta}{4(1+k_2\theta\Delta)^2}(V_n)^{2q-1}} \\ &\quad + \frac{k_3}{2(1+k_2\theta\Delta)}(V_n)^{q-\frac{1}{2}}\Delta W_n\Bigg)^2. \end{aligned}$$

Increasing the time horizon T results in an increase of the percentage of negative paths of EM and M. On the other hand BIM, BMM and of course SD are not affected by that, since they preserve their positivity on any interval $[0, T]$.

E. SUPPLEMENTARY FOR CHAPTER 4.

E.1 Proof of Proposition 4.4.7.

For the first relation of (4.4.3) we have that

$$\begin{aligned} |C_\tau - \widetilde{C}_\tau| &= e^{-rT}|\mathbb{E}(x_T - K)^+ - \mathbb{E}(y_T - K)^+| \\ &\leq \mathbb{E}\left|(x_T - K)^+ - (y_T - K)^+\right| \\ &\leq \mathbb{E}|x_T - y_T| \\ &\leq \sqrt{\mathbb{E}|x_T - y_T|^2} \to 0 \text{ as } \Delta \downarrow 0, \end{aligned} \tag{E.1.1}$$

by Theorem 4.2.2.

The other relations require a little more care. We will only sketch the proof as one can follow [HM05, Th. 5.1], where the result is in the setting of SODEs, nevertheless the main idea works also here. Therefore, setting $A := \{0 \leq x_t \leq B,\, 0 \leq t \leq T\}$ and $\widetilde{A} := \{0 \leq y_t \leq B,\, 0 \leq t \leq T\}$, we have that,

$$\begin{aligned} |B_\tau - \widetilde{B}_\tau| &\leq e^{-rT}\mathbb{E}\left|(x_T - K)^+\mathbb{I}_A - (y_T - K)^+\mathbb{I}_{\widetilde{A}}\right| \\ &\leq \mathbb{E}\left(\left|(x_T - K)^+ - (y_T - K)^+\right|\mathbb{I}_{A\cap\widetilde{A}}\right) + \mathbb{E}\left((x_T - K)^+\mathbb{I}_{A\cap\widetilde{A}^c}\right) \\ &\quad + \mathbb{E}\left((y_T - K)^+\mathbb{I}_{A^c\cap\widetilde{A}}\right) \\ &\leq \mathbb{E}\left(|x_T - y_T|\mathbb{I}_{A\cap\widetilde{A}}\right) + (B - K)\left(\mathbb{P}(A\cap\widetilde{A}^c) + \mathbb{P}(A^c\cap\widetilde{A})\right) \\ &\leq \sqrt{\mathbb{E}|x_T - y_T|^2} + (B - K)\left(\mathbb{P}(A\cap\widetilde{A}^c) + \mathbb{P}(A^c\cap\widetilde{A})\right), \end{aligned} \tag{E.1.2}$$

where M^c denotes the complement of a set M. The estimation of the above probabilities boils down to the estimation[1] of $\mathbb{E}\sup_{0\leq t\leq T}|y_t - x_t|^2$. Thus,

[1] Details can be found in the proof of [HM05, Th. 5.1].

using Theorem 4.2.2 one can show the second relation of (4.4.3).

F. SUPPLEMENTARY FOR CHAPTER 5

F.1 Boundary classification of one-dimensional time-homogeneous SDEs.

Let us now recall some results [KS88, Sec. 5.5] concerning the boundary behavior of SDEs of the form,

$$dX_t = a(X_t)dt + b(X_t)dW_t. \tag{F.1.1}$$

Let $I = (l, r)$ be an interval with $-\infty \leq l < r \leq \infty$ and define the exit time from I to be

$$S := \inf\{t \geq 0 : X_t \notin (l, r)\}.$$

Let also the coefficients of (F.1.1) satisfy the following conditions

$$b^2(x) > 0, \quad \forall x \in I, \text{ (Non Degeneracy), (ND),}$$

$$\forall x \in I, \ \exists \epsilon > 0 : \int_{x-\epsilon}^{x+\epsilon} \frac{1 + |a(y)|}{b^2(y)} dy < \infty, \text{ (Local Integrability), (LI).}$$

Then for $c \in I$, we can define the scale function

$$s(x) := \int_c^x e^{-2\int_c^y \frac{a(z)}{b^2(z)}dz} dy, \tag{F.1.2}$$

whose behavior at the endpoints of I determines the boundary behavior of (X_t) [KS88, Prop. 5.22]. In particular, we get for the transformation

$\xi(X) = \cos\theta$ in Section 5.3.1, that the effective dynamics (5.3.9) have a boundary behavior which is determined by the scale function

$$\begin{aligned} s(x) &= \int_c^x \exp\Big\{2\overline{C}\int_c^y -\frac{f(\arccos z)\sqrt{1-z^2}+\frac{1}{\beta}z}{2(1-z^2)\overline{C}\beta^{-1}}dz\Big\}dy \\ &= \int_c^x \exp\Big\{\int_c^y -\frac{\beta f(\arccos z)}{\sqrt{1-z^2}}+\frac{z}{1-z^2}dz\Big\}dy \\ &= \int_c^x \exp\Big\{\int_c^y -\frac{\beta f(\arccos z)}{\sqrt{(1-z)(1+z)}}+\frac{1}{2}\frac{1}{1-z}-\frac{1}{2}\frac{1}{1+z}dz\Big\}dy, \end{aligned}$$

for any $x \in I$. Let $I = (-1,1)$ and take $c = 0$. We compute

$$s(1-) = \int_0^1 \frac{1}{\sqrt{(1-y)(1+y)}}\exp\Big\{\int_0^y -\frac{\beta f(\arccos z)}{\sqrt{(1-z)(1+z)}}dz\Big\}dy = \infty,$$

and

$$\begin{aligned} s((-1)+) &= -\int_{-1}^0 \frac{1}{\sqrt{(1-y)(1+y)}}\exp\Big\{\int_0^y -\frac{\beta f(\arccos z)}{\sqrt{(1-z)(1+z)}}dz\Big\}dy \\ &= -\infty, \end{aligned}$$

thus by [KS88, Prop. 5.22a] we have that $\mathbb{P}(S = \infty) = 1$ that is $\mathbb{P}(-1 < \overline{\xi}_t < 1) = 1$.

F.2 Solution process of stochastic integral equation (5.2.1).

We will show that the process (5.2.3) for $n = 0$, is the solution of the stochastic integral equation (5.2.1) for $n = 0$, that is

$$y_t^{SD} = \cos\left(-cW_t + \arccos(Y_0)\right), \tag{F.2.1}$$

satisfies

$$y_t^{SD} = Y_0 + \int_0^t (-\frac{c^2}{2})y_s ds + c\int_0^t \sqrt{1-y_s^2}\,d\widehat{W}_s,$$

for $t \in (0, t_1]$, with

$$Y_0 := |x_0 + \phi(x_0)\sqrt{1-x_0^2}\cdot\Delta| \le 1.$$

Relations (5.2.5) and (5.2.2) yield

$$d\widehat{W}_t := \operatorname{sgn}(z_t)dW_t,$$

where

$$z_t = \sin\left(-c\Delta W + \arccos(Y_0)\right).$$

The cases for $n = 1, \ldots, N-1$ follow with the appropriate modifications.

We can write the increment of the Wiener process as

$$dW_t = 0 \cdot dt + 1 \cdot dW_t,$$

and view (F.2.1) as a function of W_t, i.e. $y = V(W)$ with

$$\begin{aligned} \frac{dy}{dW} &= -\sin\left(-cW + \arccos(Y_0)\right) \cdot (-c) \\ &= c\sqrt{1-y^2}\operatorname{sgn}\Big(\sin\left(-cW + \arccos(Y_0)\right)\Big), \end{aligned}$$

and

$$\begin{aligned} \frac{d^2y}{dW^2} &= -c^2\cos\left(-cW + \arccos(Y_0)\right) \\ &= -c^2 y. \end{aligned}$$

Application of Itô's formula implies

$$\begin{aligned} dy_t &= \frac{1}{2}V''(W_t)dt + V'(W_t)dW_t \\ &= -\frac{c^2}{2}y_t dt + c\sqrt{1-y_t^2}\,d\widehat{W}_t. \end{aligned}$$

SUMMARY OF NOTATION & ABBREVIATION

Ω	State Space
$\mathbb{R}^+$	$[0, \infty)$
$\mathbb{I}_A$	Indicator of set A
$\{\mathcal{F}_t\}_{t\geq 0}$	Filtration
$\mathbb{P}$	Probability Measure on the measurable space $(\Omega, \mathcal{F})$
$W(t,\omega)$ or $W_t(\omega)$ or W_t	Wiener Process
$\mathcal{L}^p([0,T];\mathbb{R}^d)$	Family of all $\mathbb{R}^d$-valued *measurable*, $\{\mathcal{F}_t\}$-*adapted* processes $\phi = \{\phi(t)\}_{0\leq t\leq T}$ such that $\int_0^T \lvert\phi(s)\rvert^p ds < \infty$ a.s.
$\mathcal{M}^p([0,T];\mathbb{R}^d)$	Family of all processes $\phi \in \mathcal{L}^p([0,T];\mathbb{R}^d)$ such that $\mathbb{E}\int_0^T \lvert\phi(s)\rvert^p ds < \infty$
$\mathcal{C}^l(A;B)$	Family of all continuous functions from A to B with continuous derivatives up to order l
$\mathcal{G}_{a,b}$	Generator corresponding to a SDE with drift coeffiicient a and diffusion coefficient b.
G_b	Noise operator corresponding to a SDE with diffusion coefficient b.
a^T, A^T	Transpose of vector a and matrix A
$x \vee y$	The maximum of x, y

a.s.	almost surely
cf.	*conferre*≡compare
e.g.	*exempli gratia*≡for example
i.e.	*id est*≡that is
r.v.	random variable
w.r.t.	with respect to
Ch.	Chapter
Def.	Definition
Fig.	Figure
Prop.	Proposition
Rem.	Remark
Sec.	Section
Th.	Theorem

INDEX

CURRICULUM VITAE

List of Publications

N. Halidias and I.S. Stamatiou. (2016). On the Numerical Solution of Some Non-Linear Stochastic Differential Equations using the Semi-Discrete Method. *Computational Methods in Applied Mathematics, 16(1), 105-132*

N. Halidias and I.S. Stamatiou (2015). Approximating Explicitly the Mean-Reverting CEV Process. 2015 *Journal of Probability and Statistics, 20 pages, DOI: 10.1155/2015/513137*

Education

B.C., Financial & Management Engineering, Oct. 2013
University of Aegean, School of Business

M.Sc., Statistics & Operational Research, Nov. 2006
National and Kapodistrian University of Athens, School of Sciences, Mathematics Department

B.C., Mathematics, Oct. 2004
National and Kapodistrian University of Athens, School of Sciences, Mathematics Department

Experience

Teaching Assistant from 2011 in the following subjects: *Actuarial Mathematics, Insurance Mathematics, Modeling Extreme Events, Calculus, Applied Linear Algebra, Financial Mathematics, Mortality Analysis*

Member of local organizing committee of 6th and 7th Conference in Actuarial Sciences & Finance on Samos.

ACKNOWLEDGMENTS

I would like to express the deepest gratitude and immeasurable appreciation to my advisor, Associate Prof. Nikolaos Halidias, for his guidance, patience and continuous support of my Ph.D study.

My sincere thanks also goes to the rest of the (Greek) committee members: Assistant Prof. Stylianos Xanthopoulos and Assistant Prof. Christos Kountzakis for our collaboration (and not only) with the teaching procedure in the Mathematics department in Samos, as well as to Assistant Prof. Spyros Hatjispyros for his support.

In addition, I would like to thank the (external) committee members: Prof. Dr. Andreas Neuenkirch, Prof. Dr. Lukasz Szpruch and Prof. Dr. Mark Peletier. It is an honor to have you all in my panel. I had the chance to meet and collaborate with Prof. Dr. Mark Peletier (I also thank the Erasmus+ Studies grant for making this possible) during my visit to the Technische Universiteit Eindhoven. Thank you Mark mostly for your time, for making helpful comments on an earlier draft of the thesis, for your advice and interest.

I would certain be remiss not to mention and sincerely thank my family (*George, Eleni, Sarantis & the rest*), Provatari Ground-Flour Team (*Dimitrakopoulou, Kaloudis, Merkatas, Sagianou, Tachtsis*), Cosmology Team (*Kolionis, Trachilis*) and my beloved friends (*Akis,Cool,Jakob,Mavros, Miltron, Pletsis, Sotiris, Theodoras, Vaggelino,*..., 16*g & D.D.*).

It was a unique experience, a journey full of uncertainty most of the times, but with moments of infinite joy, here in Samos, the Greek island of the Aegean sea which gave birth to Pythagoras, Epicurus and Aristarchus.

Ioannis S. Stamatiou, February 2016.

BIBLIOGRAPHY

[AG99] D.-H. Ahn and B. Gao. A parametric nonlinear model of term structure dynamics. *The Review of Financial Studies*, 12(4):721–762, 1999.

[AK06] J. Alcock and B. K. A note on the Balanced Method. *BIT*, 46(4):689–710, 2006.

[AK12] J. Alcock and B. K. Stable strong order 1.0 schemes for solving stochastic ordinary differential equations. *BIT Numer. Math.*, 52(3):539–557, 2012.

[Alf13] A. Alfonsi. Strong order one convergence of a drift implicit Euler scheme: Application to the CIR process. *Statistics & Probability Letters*, 83(2):602–607, 2013.

[AP07] L. Andersen and V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11(1):29–50, 2007.

[AGKR10] J. Appleby, M. Guzowska, C. Kelly, and A. Rodkina. Preserving positivity in solutions of discretised stochastic differential equations. *Applied Mathematics and Computation*, 217(2):763–774, 2010.

[AHMP07] M. Arriojas, Y. Hu, S.-E. Mohammed, and G. Pap. A delayed Black and Scholes formula. *Stochastic Analysis and Applications*, 25(2):471–492, 2007.

[BBF02] H. Berestycki, J. Busca, and I. Florent. Asymptotics and calibration of local volatility models. *Quantitative finance*, 2(1):61–69, 2002.

[Ber04] A. Berkaoui. Euler scheme for solutions of stochastic differential equations with non-Lipschitz coefficients. *Portugaliae Mathematica*, 61(4):461–478, 2004.

[Bil86] P. Billingsley. *Probability and measure*. John Willey, 2nd edition, 1986.

[BS73] F. Black and M. Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[Bre68] L. Breiman. *Probability*. Addison-Wesley, 1968.

[BS80] M. Brennan and E. Schwartz. Analyzing convertible bonds. *Journal of Financial and Quantitative analysis*, 15(4):907–929, 1980.

[BK06] M. Broadie and Ö.. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research*, 54(2):217–231, 2006.

[CKLS92] K. Chan, G. A. Karolyi, F. Longstaff, and A. Sanders. An empirical comparison of alternative models of the short-term interest rate. *The Journal of Finance*, 47(3):1209–1227, 1992.

[CLV99] T. Coleman, Y. Li, and A. Verma. Reconstructing the unknown local volatility function. *Journal of Computational Finance*, 2(3):77–102, 1999.

[Cox75] J. Cox. Notes on option pricing I: Constant elasticity of variance diffusions. *Unpublished note, Stanford University, Graduate School of Business*, 1975.

[CHJ13] S. Cox, M. Hutzenthaler, and A. Jentzen. Local Lipschitz continuity in the initial value and strong completeness for nonlinear stochastic differential equations. *arXiv preprint arXiv:1309.5595*, 2013.

[CIR85] J. Cox, J. Ingersoll, and S. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.

[DD98] G. Deelstra and F. Delbaen. Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Applied Stochastic Models and Data Analysis*, 14(1):77–84, 1998.

[DNS11] S. Dereich, A. Neuenkirch, and L. Szpruch. An Euler-type method for the strong approximation of the cox–ingersoll–ross process. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, page rspa20110505. The Royal Society, 2011.

[DFW98] B. Dumas, J. Fleming, and R. Whaley. Implied volatility functions: Empirical tests. *The Journal of Finance*, 53(6):2059–2106, 1998.

[EK86] S. Ethier and T. Kurtz. Markov Processes: Characterization and Convergence, 1986.

[Fel51] W. Feller. Two singular diffusion problems. *Annals of Mathematics*, 54:173–182, 1951.

[Fri75] A. Friedman. *Stochastic Differential Equations and Applications*, volume 1. Academic Press, 1975.

[Gro19] T. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Annals of Mathematics*, 20:292–296, 1919.

[Gyö86] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an Itô differential. *Probability theory and related fields*, 71(4):501–516, 1986.

[Hal12] N. Halidias. Semi-discrete approximations for stochastic differential equations and applications. *International Journal of Computer Mathematics*, 89(6):780–794, 2012.

[Hal14] N. Halidias. A novel approach to construct numerical methods for stochastic differential equations. *Numerical Algorithms*, 66(1):79–87, 2014.

[Hal15a] N. Halidias. Constructing positivity preserving numerical schemes for the two-factor CIR model. *Monte Carlo Methods and Applications*, 21(4):313–323, 2015.

[Hal15b] N. Halidias. Construction of positivity preserving numerical schemes for some multidimensional stochastic differential equations. *Discrete and Continuous Dynamical System Series B*, 20(1):153–160, 2015.

[Hal15c] N. Halidias. An explicit and positivity preserving numerical scheme for the mean reverting CEV model. *Japan Journal of Industrial and Applied Mathematics*, 32(2):545–552, 2015.

[Hal15d] N. Halidias. A new numerical scheme for the CIR process. *Monte Carlo Methods and Applications*, 2015.

[Hal16] N. Halidias. On construction of boundary preserving numerical schemes. *arXiv preprint arXiv:1601.07864*, 2016.

[HS15] N. Halidias and I. Stamatiou. Approximating Explicitly the Mean-Reverting CEV Process. *Journal of Probability and Statistics*, Article ID 513137, 20 pages, 2015.

[HS16] N. Halidias and I. Stamatiou. On the Numerical Solution of Some Non-Linear Stochastic Differential Equations Using the Semi-Discrete Method. *Computational Methods in Applied Mathematics*, 16(1):105–132, 2016.

[HLP52] G. Hardy, J. Littlewood, and G. Pólya. *Inequalities*. Cambridge university press, 1952.

[Hes93] S. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies*, 6(2):327–343, 1993.

[Hes97] S. Heston. A simple new formula for options with stochastic volatility. *preprint, http://ssrn.com/abstract=86074*, 1997.

[HM05] D. Higham and X. Mao. Convergence of Monte Carlo simulations involving the mean-reverting square root process. *Journal of Computational Finance*, 8(3):35–61, 2005.

[HMS13] D. Higham, X. Mao, and L. Szpruch. Convergence, nonnegativity and stability of a new Milstein scheme with applications to finance. *Discrete & Continuous Dynamical Systems-Series B.*, 18(8):1–18, 2013.

[HW87] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *The Journal of Finance*, 42(2):281–300, 1987.

[HK08] T. Hurd and A. Kuznetsov. Explicit formulas for Laplace transforms of stochastic integrals. *Markov Processes and Related Fields*, 14(2):277–290, 2008.

[HJ14] M. Hutzenthaler and A. Jentzen. On a perturbation theory and on strong convergence rates for stochastic ordinary and partial differential equations with non-globally monotone coefficients. *arXiv preprint arXiv:1401.0295*, 2014.

[HJ15] M. Hutzenthaler and A. Jentzen. Numerical approximations of stochastic differential equations with non-globally Lipschitz continuous coefficients. *to appear in Memoirs of the American Mathematical Society*, 236(1112), 2015.

[HJK11] M. Hutzenthaler, A. Jentzen, and P. Kloeden. Strong and weak divergence in finite time of Euler's method for stochastic differential equations with non-globally Lipschitz continuous coefficients. In *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, volume 467, pages 1563–1576. The Royal Society, 2011.

[HJK12] M. Hutzenthaler, A. Jentzen, and P. Kloeden. Strong convergence of an explicit numerical method for SDEs with nonglobally Lipschitz continuous coefficients. *The Annals of Applied Probability*, 22(4):1611–1641, 2012.

[JP03] J. Jacod and P. E. Protter. *Probability essentials.* Springer Science & Business Media, 2003.

[Kah04] C. Kahl. Positive numerical integration of stochastic differential equations. *University of Wuppertal, Research Group Numerical Analysis*, 2004.

[KGR08] C. Kahl, M. Günther, and T. Rossberg. Structure preserving stochastic integration schemes in interest rate derivative modeling. *Applied Numerical Mathematics*, 58(3):284–295, 2008.

[KJ06] C. Kahl and P. Jäckel. Fast strong approximation Monte Carlo schemes for stochastic volatility models. *Quantitative Finance*, 6(6):513–536, 2006.

[KS06] C. Kahl and H. Schurz. Balanced Milstein methods for ordinary SDEs. *Monte Carlo Methods and Applications*, 12(2):143–170, 2006.

[KS88] I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus.* Springer-Verlag, New York, 1988.

[KN13] P. Kloeden and A. Neuenkirch. Convergence of numerical methods for stochastic differential equations in mathematical finance. *Recent Developments in Computational Finance: Foundations, Algorithms and Applications*, pages 49–80, 2013.

[KP95] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*, volume 23. Springer-Verlag, Berlin, corrected 2nd printing, 1995.

[KPS03] P. Kloeden, E. Platen, and H. Schurz. Numerical solution of SDE through computer experiments, 2003.

[KP00] U. Küchler and E. Platen. Strong discrete time approximation of stochastic differential equations with time delay. *Mathematics and Computers in Simulation*, 54(1):189–205, 2000.

[LL10] F. Legoll and T. Lelièvre. Effective dynamics using conditional expectations. *Nonlinearity*, 23(9):2131–2163, September 2010.

[Man09] R. Mansuy. The Origins of the Word "Martingale". *Electronic Journal for History of Probability and Statistics*, 5(1):1–10, 2009.

[Mao97] X. Mao. *Stochastic differential equations and applications.* Horwood Publishing, Chichester, 1997.

[MS13a] X. Mao and S. Sabanis. Delay geometric Brownian motion in financial option valuation. *Stochastics An International Journal of Probability and Stochastic Processes*, 85(2):295–320, 2013.

[MS13b] X. Mao and L. Szpruch. Strong convergence and stability of implicit numerical methods for stochastic differential equations with non-globally Lipschitz continuous coefficients. *Journal of Computational and Applied Mathematics*, 238:14–28, 2013.

[MS13c] X. Mao and L. Szpruch. Strong convergence rates for backward Euler–Maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients. *Stochastics: An International Journal of Probability and Stochastic Processes*, 85(1):144–171, 2013.

[MG10] R. Makarov and D. Glew. Exact simulation of Bessel diffusions. *Monte Carlo Methods and Applications*, 16(3-4):283–306, 2010.

[Mar55] G. Maruyama. Continuous Markov processes and stochastic equations. *Rendiconti del Circolo Matematico di Palermo*, 4(1):48–90, 1955.

[MPS98] G. Milstein, E. Platen, and H. Schurz. Balanced implicit methods for stiff stochastic systems. *SIAM Journal on Numerical Analysis*, 35(3):1010–1019, 1998.

[NS14] A. Neuenkirch and L. Szpruch. First order strong approximations of scalar SDEs defined in a domain. *Numerische Mathematik*, 128(1):103–136, 2014.

[Øks03] B. Øksendal. *Stochastic differential equations.* Springer, 2003.

[Olv97] F. Olver. *Asymptotics and Special Functions.* AKP classics, A K Peters, Wellesley., 1997.

[RW87] L. C. G. Rogers and D. Williams. *Diffusions, Markov Processes and Martingales: Itô calculus*, volume 2. Wiley & Sons, 1987.

[Sab15] S. Sabanis. Euler approximations with varying coefficients: the case of superlinearly growing diffusion coefficients. *arXiv preprint arXiv:1308.1796v3*, 2015.

[Sch96] H. Schurz. Numerical regularization for SDEs: Construction of nonnegative solutions. *Dynamic Systems and Applications*, 5(3):323–351, 1996.

[Shr04] S. Shreve. *Stochastic Calculus for Finance II Continuous-Time Models.* Springer, 2004.

[Szp13] L. Szpruch. V-stable tamed Euler schemes. *arXiv preprint arXiv:1310.0785v1*, 2013.

[TZ13] M. Tretyakov and Z. Zhang. A fundamental mean-square convergence theorem for SDEs with locally Lipschitz coefficients and its applications. *SIAM Journal on Numerical Analysis*, 51(6):3135–3162, 2013.

[YW71] T. Yamada and S. Watanabe. On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University*, 11(1):155–167, 1971.

[Zha14] Z. Zhang. New explicit balanced schemes for SDEs with locally Lipschitz coefficients. *arXiv preprint arXiv:1402.3708*, 2014.

[Zhu09] L. Zhu. A source of inequalities for circular functions. *Computers & Mathematics with Applications*, 58(10):1998–2004, 2009.