Weak and Strong Approximation of the Log-Heston Model by Euler-Type Methods and Related Topics

Inauguraldissertation
zur Erlangung des akademischen Grades
eines Doktors der Naturwissenschaften
der Universität Mannheim

vorgelegt von

Annalena Mickel
aus Büdingen

Mannheim, 2022

Dekan:	Dr. Bernd Lübcke, Universität Mannheim
Referent:	Professor Dr. Andreas Neuenkirch, Universität Mannheim
Korreferent:	Professor Dr. Andreas Rößler, Universität zu Lübeck
Korreferent:	Dr. habil. Larisa Yaroslavtseva, Universität Graz

Tag der mündlichen Prüfung: 17.03.2023

Abstract

This thesis deals with the weak and strong numerical approximation of so-called *stochastic volatility models*. In particular, the focus is on the log-Heston model and its associated Euler methods, for which there have been only a few convergence results with a polynomial rate in the literature so far. The biggest challenge here is the approximation of the CIR process, which models the stochastic variance and whose diffusion coefficient is not Lipschitz continuous.
We first study the weak order of convergence of two Euler methods that keep the approximation of the CIR process positive. When the Feller index ν of the CIR process is greater than one, weak convergence of order one is obtained as under standard assumptions. For $\nu \leq 1$ we obtain a weak order of convergence of $\nu - \varepsilon$ for $\varepsilon > 0$ arbitrarily small. For the L^1-error for a large class of Euler methods, we can recover the order 1/2 obtained under standard assumptions under the condition $\nu > 1$. Moreover, we prove that this is already the optimal L^1-convergence order for the log-Heston model. Finally, in the last part of this dissertation we deal with the optimal L^2 approximation of more general stochastic volatility models.

Zusammenfassung

Diese Dissertation befasst sich mit der schwachen und starken numerischen Approximation von sogenannten *Stochastischen Volatilitätsmodellen.* Im Fokus stehen hierbei konkret das log-Heston-Modell und die zugehörigen Euler-Verfahren, für die es in der Literatur bisher nur weniger Konvergenzresultate mit polynomieller Rate gab. Die größte Herausforderung stellt hierbei die Approximation des CIR-Prozesses dar, welcher die stochastische Varianz modelliert und dessen Diffusionskoeffizient nicht Lipschitz-stetig ist.
Wir untersuchen zunächst die schwache Konvergenzordnung von zwei Euler-Verfahren, die die Approximation des CIR-Prozesses positiv halten. Wenn der Feller-Index ν des CIR-Prozesses größer als eins ist, so ergibt sich eine schwache Konvergenz der Ordnung eins wie unter Standardannahmen. Für $\nu \leq 1$ erhalten wir eine schwache Ordnung von $\nu - \varepsilon$ für $\varepsilon > 0$ beliebig klein. Für den L^1-Fehler können wir für eine große Klasse von Euler-Verfahren die Ordnung 1/2, die unter Standardannahmen erreicht wird, unter der Bedingung $\nu > 1$ wiederherstellen. Zudem beweisen wir, dass dies bereits die optimale L^1-Konvergenzrate für das log-Heston Modell ist. Im letzten Teil dieser Dissertation beschäftigen wir uns schließlich mit der optimalen L^2-Approximation von allgemeineren Stochastischen Volatilitätsmodellen.

Acknowledgments

This thesis was written during my time as a researcher at the University of Mannheim and as a member of the DFG Research Training Group 1953 "Statistical Modeling of Complex Systems and Processes".

I started my PhD studies during the first Covid-19 "lockdown" and the pandemic had a major impact on the following three years. Under these difficult circumstances I was greatly supported by my supervisor Prof. Dr. Andreas Neuenkirch. I am deeply grateful to him for his excellent guidance, advice and the extremely valuable lessons in academic writing and presenting.

I am thankful to Prof. Dr. Andreas Rößler and Dr. Larisa Yaroslavtseva for being the co-referees of this thesis and the examination committee for kindly agreeing to evaluate my thesis defense.

I am very thankful for my colleagues at the University of Mannheim, especially Alex, André and Thomas, who made my office days extremely enjoyable.

Last but not least I would like to thank my family and my partner for their enormous support and for eliminating my insecurities.

Contents

Chapter 1

Introduction

This thesis deals with the numerical approximation of the Heston model from [36]. Its dynamics are described by the following two stochastic differential equations (SDEs) which typically model an asset price S and its variance V:

$$dS_t = rS_t dt + \sqrt{V_t} S_t \Big(\rho dW_t + \sqrt{1-\rho^2} dB_t\Big),$$
$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t} dW_t.$$

The Heston model is an extension of the famous Black-Scholes model. The latter assumed the variance (or respectively the volatility which is its square root) to be deterministic. In contrast to this, the Heston model falls into the class of *stochastic volatility models* because the variance is modeled here as a stochastic process.
Although a very classical model in mathematical finance, even very simple time-discrete simulation methods for the Heston model are not well understood. The main reason for this is the second SDE. Its diffusion coefficient is not globally Lipschitz continuous and therefore standard textbook results cannot be applied. Our main focus is on the analysis of explicit Euler methods which are arguably the simplest time-discrete simulations schemes for SDEs. Despite this fact, weak and strong convergence results for Euler methods in the context of the Heston model are rare and often do not match observations from numerical experiments. As a consequence, Euler schemes are used in practice without theoretical guarantees. In this thesis, we try to close some of these gaps. Apart from the analysis of explicit Euler methods, we also provide new results for the implicit Milstein method and we present new results concerning the optimal approximation of more general stochastic volatility models.

1.1 Outline

This thesis is structured as follows: In Chapter 2, we introduce the Heston model and its properties. Thereafter, we present some popular simulation algorithms for it in Chapter 3. Chapter 4 lays the groundwork for our analysis of explicit Euler schemes. Here, their properties in the context of the Heston model are analyzed. Chapter 5 establishes the connection between the solution of the log-Heston SDE and its associated partial differential equation (PDE). Our first main result is then carried out in Chapter 6 where we analyze the weak convergence behavior of Euler and Milstein-type discretizations of the log-Heston model. Chapters 7 and 8 examine upper and lower bounds for its L^1-approximation. We support our theoretical results by numerical simulations carried out in Chapter 9. Finally, we turn to the analysis of more general stochastic volatility models in Chapter 10. Here, we deal with the question of their optimal L^2-approximation. We summarize our findings in the last chapter.

1.2 Notation

For a multi-index $l = (l_1, ..., l_d) \in \mathbb{N}^d$, we define $|l| = \sum_{j=1}^{d} l_j$ and for $y \in \mathbb{R}^d$, we define $\partial_y^l = \partial_{y_1}^{l_1} \cdots \partial_{y_d}^{l_d}$. Moreover, we denote by $|y|$ the standard Euclidean norm in $\mathbb{R}^d$. Let $\mathcal{D} \subset \mathbb{R}^d$ be a domain and $q \in \mathbb{N}$. $C^q(\mathcal{D};\mathbb{R})$ is the set of all functions on $\mathcal{D}$ which are q-times continuously differentiable. $C^q_{pol}(\mathcal{D};\mathbb{R})$ is the set of functions $g \in C^q(\mathcal{D};\mathbb{R})$ such that there exist $C, a > 0$ for which

$$|\partial_y^l g(y)| \leq C(1 + |y|^a) \quad y \in \mathcal{D}, |l| \leq q.$$

We set $C^q_{pol,T}(\mathcal{D};\mathbb{R})$ the set of functions $v \in C^{\lfloor q/2 \rfloor, q}_{pol}([0,T) \times \mathcal{D};\mathbb{R})$ such that there exist $C, a > 0$ for which

$$\sup_{t<T} |\partial_t^k \partial_y^l v(t,y)| \leq C(1 + |y|^a) \quad y \in \mathcal{D}, 2k + |l| \leq q.$$

For $\varepsilon \in (0,1)$, we denote by $C^{q+\varepsilon}(\mathcal{D};\mathbb{R})$ the set of all functions from $C^q(\mathcal{D};\mathbb{R})$ in which partial derivatives of order q are Hölder-continuous of order ε, and $C_c^{q+\varepsilon}(\mathcal{D};\mathbb{R})$ is the set of all functions from $C^{q+\varepsilon}(\mathcal{D};\mathbb{R})$ which have compact support.

We use the notation x^+ to denote the positive part of x: $x^+ = \max\{x, 0\}$.

Constants, which depend only on the parameters of the respective SDE such as T, x_0, v_0, κ, θ, σ and ρ in the case of the Heston model, will be denoted in the following by C, regardless of their value. Other dependencies will be denoted by subscripts, i.e. $C_{h,\beta}$ means that this constant depends additionally on the function h and the parameter β. Moreover, the value of all constants can change from line to line.

Throughout almost all of the chapters we require the following well-known Burkholder-Davis-Gundy (BDG) inequalities, see e.g. Theorem 3.28 in Chapter III of [49].

Proposition 1.1. *Let $M = (M_t)_{t\in[0,T]}$ be a continuous martingale and $\alpha > 0$. Then, there exist constants $c_\alpha, C_\alpha > 0$ such that*

$$c_\alpha \, \mathbb{E}\left[\langle M \rangle_t^\alpha\right] \leq \mathbb{E}\left[\sup_{u\in[0,t]} |M_u|^{2\alpha}\right] \leq C_\alpha \, \mathbb{E}\left[\langle M \rangle_t^\alpha\right], \qquad t \in [0,T].$$

Chapter 2

The Heston Model

The Heston model was proposed by Steven L. Heston in 1993 [36]. It is given by the SDEs

$$\begin{aligned} dS_t &= rS_t dt + \sqrt{V_t} S_t \Big(\rho dW_t + \sqrt{1-\rho^2} dB_t \Big), && S_0 = s, \\ dV_t &= \kappa(\theta - V_t)dt + \sigma \sqrt{V_t} dW_t, && V_0 = v, \end{aligned}$$

with $\kappa, \theta, \sigma > 0$, $r \in \mathbb{R}$, $\rho \in [-1, 1]$, $T > 0$ and independent one-dimensional Brownian motions $W = (W_t)_{t\in[0,T]}$, $B = (B_t)_{t\in[0,T]}$ which are defined on a filtered probability space $\left(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, \mathbb{P}\right)$ and the filtration satisfies the usual conditions. Furthermore, the initial values $s, v > 0$ are assumed to be deterministic. Here $(S_t)_{t\in[0,T]}$ models the price of an asset and $(V_t)_{t\in[0,T]}$ its variance, which is given by the so called Cox–Ingersoll–Ross (CIR) process. Usually, the log-Heston model instead of the Heston model is considered in numerical practice. We therefore set $X_t := \log(S_t)$. This yields the SDEs

$$\begin{aligned} dX_t &= \left(r - \frac{1}{2} V_t \right) dt + \sqrt{V_t} \Big(\rho dW_t + \sqrt{1-\rho^2} dB_t \Big), && X_0 = x, \\ dV_t &= \kappa(\theta - V_t)dt + \sigma \sqrt{V_t} dW_t, && V_0 = v, \end{aligned} \tag{2.1}$$

by a simple application of the Itō formula. Note that the square root coefficient is not globally Lipschitz continuous. Thus, the (log-)Heston SDE does not satisfy the standard assumptions for the numerical analysis of SDEs.

The Heston model is a natural extension of the celebrated Black-Scholes model because it considers a stochastic volatility rather than a constant one. As a consequence, the Heston model takes the asymmetry and excess kurtosis of financial asset returns into account which are typically observed in real market data. The analysis of the Heston model is not only of theoretical relevance. With the rise of volatility trading in financial markets, stochastic volatility models are becoming more important.

2.1 The CIR Process

The CIR process is the solution to the following SDE:

$$dV_t = \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t, \qquad V_0 = v > 0. \tag{2.2}$$

It was first used by Cox, Ingersoll and Ross [20] to model short term interest rates. In this thesis, we assume the parameters to be strictly positive. They can be interpreted as follows: θ is the long run mean of the process, κ is its speed of mean reversion and σ is its volatility. Since the coefficients of SDE (2.2) are continuous and of linear growth, a weak solution exists (e.g. Theorem 2.4 in [44]). By the Yamada-Watanabe condition (e.g. Theorem IV 3.2 in [44]), pathwise uniqueness holds. Since the existence of a weak solution and pathwise uniqueness imply the existence of a strong solution (e.g. Chapter IX, Theorem 1.7 in [68]), we know that SDE (2.2) has a unique strong solution. The CIR process has an important relation with the squared Bessel process which is the unique strong solution of

$$dZ_t = \delta dt + 2\sqrt{Z_t}dW_t, \qquad Z_0 = z > 0$$

where we assume $\delta > 0$. From e.g. Proposition 6.3.1.1 in [45], we know that the CIR process can be expressed as a squared Bessel process with $\delta = \frac{4\kappa\theta}{\sigma^2}$ degrees of freedom and the following space-time changes:

$$V_t = e^{-\kappa t} Z_{\frac{\sigma^2}{4k}(e^{\kappa t}-1)}.$$

Groundbreaking work to understand the dynamics of squared Bessel processes was already done in the 1950s when Feller (e.g. in [27]) studied the parabolic PDE

$$u_t = (axu)_{xx} - \left((bx + c)u\right)_x, \qquad 0 < x < \infty. \tag{2.3}$$

Here, a, b, c are constants and Feller only assumed $a > 0$. Equation (2.3) can be seen as a Kolmogorov forward (or Fokker-Planck) equation for an SDE with drift coefficient $bx + c$ and diffusion coefficient $\sqrt{2ax}$. For given initial conditions, Feller showed that the only norm preserving solution of (2.3) (that leads to a transition density of the associated stochastic process) has to have $c \geq 0$ and a flux zero at the origin, i.e.

$$\lim_{x\to 0} -\left((axu(t,x))_x - (bx + c)u(t,x)\right) = 0.$$

This means that a reflecting barrier condition has to be imposed. For $c > a$, this solution vanishes at $x = 0$. Furthermore, Feller derived the Laplace transform of the transition density. For the squared Bessel process we obtain $a = 2$, $b = 0$ and $c = \delta$ and for the CIR process, we have $a = \frac{1}{2}\sigma^2$, $b = -\kappa$ and $c = \kappa\theta$. Because of Feller's work, the ratio $\frac{c}{a} = \frac{2\kappa\theta}{\sigma^2}$ is often called *Feller index* in the context of the CIR process.
To exactly determine the behavior of the squared Bessel process at 0 and ∞, we need to distinguish the cases $0 < \delta < 2$, $\delta = 2$ and $\delta > 2$ and we set

$$\tilde{T} = \inf\{t \geq 0 : Z_t \notin (0, \infty)\}$$

to be the exit time from $(0, \infty)$. For all choices of δ the infinite point ∞ is a natural boundary, it cannot be reached in finite time (see e.g. [50] Chapter 15.6). For $0 < \delta < 2$ we have $\mathbb{P}\left(\tilde{T} < \infty\right) = 1$ and

$$P\left(\lim_{t \uparrow \tilde{T}} Z_t = 0\right) = \mathbb{P}\left(\sup_{0 \leq t < \tilde{T}} Z_t < \infty\right) = 1$$

by applying Theorem 5.29 and Proposition 5.22 from [49]. The point 0 is instantaneously reflecting (see Chapter XI, Proposition (1.5) in [68]). This means that the time spent by Z in the point 0 has Lebesgue measure 0. For $\delta = 2$, we obtain

$$\mathbb{P}\left(\tilde{T} = \infty\right) = \mathbb{P}\left(\sup_{0 \leq t < \infty} Z_t = \infty\right) = \mathbb{P}\left(\inf_{0 \leq t < \infty} Z_t = 0\right) = 1$$

by Proposition 5.22 from [49]. Finally for $\delta > 2$, we have $\mathbb{P}\left(\tilde{T} = \infty\right) = 1$ by Theorem 5.29 from [49] and

$$P\left(\lim_{t \to \infty} Z_t = \infty\right) = \mathbb{P}\left(\sup_{0 \leq t < \infty} Z_t > 0\right) = 1$$

by Proposition 5.22 from [49]. From this, we can deduce the following well-known proposition for the CIR process:

Proposition 2.1. *We denote the Feller index by*

$$\nu := \frac{2\kappa\theta}{\sigma^2}.$$

For $\nu \geq 1$, the solution of the CIR process is strictly positive, i.e.

$$\mathbb{P}\left(V_t \in (0, \infty), \forall t \geq 0\right) = 1.$$

For $0 < \nu < 1$, we have

$$\mathbb{P}\left(V_t \in [0, \infty), \forall t \geq 0\right) = 1,$$

the origin is attainable but instantaneously reflecting.

The Feller index also plays an important role when we look at the moments of the CIR process.

Proposition 2.2. *The CIR process has bounded moments, it holds that*

$$\mathbb{E}\left[\sup_{t \in [0,T]} V_t^p\right] < \infty$$

for all $p \geq 1$ and

$$\sup_{t \in [0,T]} \mathbb{E}\left[V_t^p\right] < \infty$$

for all $p > -\nu$.

Proof. The proof of the second statement can be found in Section 3 of [24] where the results of [42] are used. To show the first assertion, we use Jensen's inequality and the BDG inequality. Let $p \geq 2$, then

$$\begin{aligned}
\mathbb{E}\left[\sup_{t\in[0,T]} V_t^p\right] &= \mathbb{E}\left[\sup_{t\in[0,T]}\left|v + \int_0^t \kappa(\theta - V_s)ds + \sigma\int_0^t \sqrt{V_s}dW_s\right|^p\right] \\
&\leq C_p\left(v^p + \mathbb{E}\left[\left|\sup_{t\in[0,T]}\int_0^t \kappa(\theta - V_s)ds\right|^p\right] + \mathbb{E}\left[\sup_{t\in[0,T]}\left|\sigma\int_0^t \sqrt{V_s}dW_s\right|^p\right]\right) \\
&\leq C_p\left(v^p + T^p + \mathbb{E}\left[\left(\int_0^T V_s ds\right)^p\right] + \mathbb{E}\left[\left(\int_0^T V_s ds\right)^{p/2}\right]\right) \\
&\leq C_p\left(1 + \sup_{s\in[0,T]}\mathbb{E}\left[V_s^p\right] + \sup_{s\in[0,T]}\mathbb{E}\left[V_s^{p/2}\right]\right) \\
&\leq C_p
\end{aligned}$$

by the second statement. The case $p \in [1,2)$ follows by the Lyapunov inequality. □

Furthermore, we have the following L^p-result for the increments of the CIR process.

Lemma 2.3. *For all $p \geq 1$ there exist a constant $C_p > 0$, such that*

$$\sup_{0\leq s<t\leq T}\mathbb{E}\left[\frac{|V_t - V_s|^p}{|t-s|^{p/2}}\right] \leq C_p$$

Proof. First, let $p \geq 2$. Then we have with the BDG and the Hölder inequality

$$\begin{aligned}
\mathbb{E}\left[\frac{|V_t - V_s|^p}{|t-s|^{p/2}}\right] &= \frac{1}{|t-s|^{p/2}}\mathbb{E}\left[\left|\int_s^t \kappa\left(\theta - V_u\right)du + \sigma\int_s^t \sqrt{V_u}dW_u\right|^p\right] \\
&\leq \frac{2^{p-1}}{|t-s|^{p/2}}\left(\mathbb{E}\left[\left|\int_s^t \kappa\left(\theta - V_u\right)du\right|^p\right] + \mathbb{E}\left[\left|\sigma\int_s^t \sqrt{V_u}dW_u\right|^p\right]\right) \\
&\leq \frac{C_p}{|t-s|^{p/2}}\left(|t-s|^p + |t-s|^{p-1}\int_s^t \mathbb{E}\left[V_u^p\right]du + \mathbb{E}\left[\left|\int_s^t V_u du\right|^{\frac{p}{2}}\right]\right) \\
&\leq \frac{C_p}{|t-s|^{p/2}}\left(|t-s|^p + |t-s|^{\frac{p}{2}-1}\int_s^t \mathbb{E}\left[V_u^{\frac{p}{2}}\right]du\right) \\
&\leq \frac{C_p}{|t-s|^{p/2}}\left(|t-s|^p + |t-s|^{p/2}\right) \\
&\leq C_p\left(T^{p/2} + 1\right) \\
&\leq C_p,
\end{aligned}$$

where we used Proposition 2.2. The case $p \in [1,2)$ then follows by the Lyapunov inequality. □

As already mentioned, Feller derived the Laplace transform of the transition density which is the solution of (2.3). By inverting this Laplace transform it is possible to find the conditional distribution of the CIR process (e.g. [6], [20]).

Proposition 2.4. *Let F_{χ^2} be the cumulative distribution function of the non-central chi-squared distribution with non-centrality parameter λ and d degrees of freedom, i.e.*

$$F_{\chi^2}(x; d, \lambda) = \sum_{i=0}^{\infty} e^{-\frac{\lambda}{2}} \frac{\left(\frac{\lambda}{2}\right)^i}{i!} \frac{\int_0^x t^{\frac{d}{2}+i-1} e^{-\frac{t}{2}} dt}{2^{\frac{d}{2}+i} \Gamma\left(i + \frac{d}{2}\right)}$$

where Γ is the gamma function. Let $0 \leq s < t \leq T$. Conditional on V_s, V_t is distributed as $\frac{\sigma^2\left(1-e^{-\kappa(t-s)}\right)}{4\kappa}$ times a non-central chi-squared distributed random variable with $\frac{4\kappa\theta}{\sigma^2}$ degrees of freedom and non-centrality parameter $\frac{4\kappa e^{-\kappa(t-s)}}{\sigma^2(1-e^{-\kappa(t-s)})} V_s$, i.e.

$$P\left(V_t \leq x | V_s\right) = F_{\chi^2}\left(\frac{4\kappa}{\sigma^2 e^{-\kappa(t-s)}} x; \frac{4\kappa\theta}{\sigma^2}, \frac{4\kappa e^{-\kappa(t-s)}}{\sigma^2(1-e^{-\kappa(t-s)})} V_s\right).$$

Since we know the conditional distribution from Proposition 2.4, we can calculate the expectation and variance of V_t given V_s.

Corollary 2.5. *Let $0 \leq s < t \leq T$. Conditional on V_s, the expectation and variance of V_t are*

$$\begin{aligned} \mathbb{E}\left[V_t | V_s\right] &= \theta + (V_s - \theta) e^{-\kappa(t-s)}, \\ Var\left(V_t | V_s\right) &= \frac{V_s \sigma^2 e^{-\kappa(t-s)}}{\kappa}\left(1 - e^{-\kappa(t-s)}\right) + \frac{\theta\sigma^2}{2\kappa}\left(1 - e^{-\kappa(t-s)}\right)^2. \end{aligned}$$

Proof. Let Y be a non-central chi-squared distributed random variable with d degrees of freedom and non-centrality parameter λ. Then,

$$\mathbb{E}\left[Y\right] = d + \lambda, \qquad Var\left(Y\right) = 2(d + 2\lambda).$$

Now, easy calculations give us

$$\begin{aligned} \mathbb{E}\left[V_t | V_s\right] &= \frac{\sigma^2\left(1 - e^{-\kappa(t-s)}\right)}{4\kappa}\left(\frac{4\kappa\theta}{\sigma^2} + \frac{4\kappa e^{-\kappa(t-s)}}{\sigma^2(1-e^{-\kappa(t-s)})} V_s\right) \\ &= \theta\left(1 - e^{-\kappa(t-s)}\right) + V_s e^{-\kappa(t-s)} \end{aligned}$$

and

$$\begin{aligned} Var\left(V_t | V_s\right) &= \frac{\sigma^4\left(1 - e^{-\kappa(t-s)}\right)^2}{8\kappa^2}\left(\frac{4\kappa\theta}{\sigma^2} + \frac{8\kappa e^{-\kappa(t-s)}}{\sigma^2(1-e^{-\kappa(t-s)})} V_s\right) \\ &= \frac{\theta\sigma^2\left(1 - e^{-\kappa(t-s)}\right)^2}{2\kappa} + \frac{V_s\sigma^2 e^{-\kappa(t-s)}}{\kappa}\left(1 - e^{-\kappa(t-s)}\right). \end{aligned}$$

□

2.2 The CEV Process

The constant elasticity of volatility (CEV) process is given by the solution of

$$dV_t = \kappa (\theta - V_t)\, dt + \sigma V_t^{\gamma} dW_t, \qquad V_0 = v_0$$

where $\gamma \in \left[\frac{1}{2}, 1\right)$. It is a generalization of the CIR process and can be used together with the Heston price process for a generalized version of the Heston model. Similar to the CIR process, the CEV process has a unique, strong solution by the Yamada-Watanade condition. For $\gamma \in \left(\frac{1}{2}, 1\right)$ the CEV process has some desirable properties. In contrast to the case $\gamma = \frac{1}{2}$, it is then always strictly positive with no restrictions on the parameters, i.e.

$$\mathbb{P}(V_t > 0, \forall t > 0) = 1$$

for $\nu > 0$ (see e.g. [7]). Furthermore, it has bounded moments for the whole parameter range (see [9]).

Proposition 2.6. *For the CEV process with $\gamma \in \left(\frac{1}{2}, 1\right)$, it holds that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} V_t^p\right] < \infty, \qquad\qquad \sup_{t\in[0,T]} \mathbb{E}\left[V_t^{-p}\right] < \infty$$

for all $p \geq 0$.

By similar calculations as in Lemma 2.3, we get the following result using Proposition 2.6.

Lemma 2.7. *For all $p \geq 1$ there exist a constant $C_p > 0$, such that*

$$\sup_{0\leq s<t\leq T} \mathbb{E}\left[\frac{|V_t - V_s|^p}{|t-s|^{p/2}}\right] \leq C_p.$$

2.3 The Price Process

The price process of the Heston model is given by the solution of the SDE

$$dS_t = rS_t dt + \sqrt{V_t} S_t \left(\rho dW_t + \sqrt{1-\rho^2} dB_t\right)$$

where $(V_t)_{t\in[0,T]}$ is the solution of the CIR process. In the generalized Heston model, the latter would be the CEV process. Here, $r \in \mathbb{R}$ models the risk-free interest rate and $\rho \in [-1, 1]$ the correlation of the price and the variance process. The Brownian motions W and B are independent. The parameter r is often omitted since the transformation $\hat{S}_t = e^{-rt} S_t$ by the Itō formula leads to a Heston model where $r = 0$. From [7], we have the following result:

Proposition 2.8. *The process $(S_t)_{t\in[0,T]}$ can neither reach 0 nor ∞ in finite time. In the case of the CIR process $\left(\gamma = \frac{1}{2}\right)$, $(S_t)_{t\in[0,T]}$ is a martingale. For $\gamma \in \left(\frac{1}{2}, 1\right)$, the price process is a martingale for $\rho \leq 0$ and a strict supermartingale for $\rho > 0$.*

In [7] it is shown that the moments of the price process can become infinite in finite time. Rewriting the results in terms of the correlation parameter ρ leads to the following proposition.

Proposition 2.9. *Define*

$$T^\star(p) := \inf\left\{t \geq 0, \mathbb{E}\left[S_t^p\right] = \infty\right\}$$

for $p \in (1, \infty)$. For $\gamma = \frac{1}{2}$, we have

$$T^\star(p) = \infty \iff \rho \leq -\sqrt{\frac{p-1}{p}} + \frac{\kappa}{2\sigma p}.$$

For $\gamma \in \left(\frac{1}{2}, 1\right)$, it holds that

$$T^\star(p) = \infty \text{ if } \rho < -\sqrt{\frac{p-1}{p}}.$$

For $\rho = 0$, we have $T^\star(p) < \infty$ for all $p > 1$ and for $\rho > 0$, we have $T^\star(p) < \infty$ if $p > \left(1 - \frac{\rho^2}{2}\right)^{-1}$.

Recall that an application of Itō's formula with $X_t = \log(S_t)$ gives

$$dX_t = \left(r - \frac{1}{2}V_t\right)dt + \sqrt{V_t}\left(\rho dW_t + \sqrt{1-\rho^2}dB_t\right). \tag{2.4}$$

The solution of this SDE with $X_0 = x$ is the log-Heston price (or the generalized log-Heston price). Looking at the integral representation of (2.4) and applying the Burkholder-Davis-Gundy inequality and Proposition 2.2 (or Proposition 2.6), we get that the moments of the log-price process are bounded.

Proposition 2.10. *It holds that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t|^p\right] < \infty \quad \forall p \geq 1.$$

Analogously to Lemma 2.3, we can prove the following L^p result for the log-price process.

Lemma 2.11. *For all $p \geq 1$ there exist a constant $C_p > 0$, such that*

$$\sup_{0\leq s<t\leq T} \mathbb{E}\left[\frac{|X_t - X_s|^p}{|t-s|^{p/2}}\right] \leq C_p.$$

Furthermore, (2.4) is a representation of the log-price process which only depends on the volatility $\sqrt{V}$ since S cancels out. These two properties are favorable for the numerical analysis.

Chapter 3

Simulation Algorithms for the log-Heston Model

Consider the SDE in $\mathbb{R}^d$

$$dY_t = b(Y_t)dt + \sigma(Y_t)dW_t, \qquad Y_0 = y \in \mathbb{R}^d \tag{3.1}$$

with drift and diffusion coefficients $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ where $(W_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion. Furthermore, assume that (3.1) has a unique strong solution. The calculation of

$$p := \mathbb{E}\left[\phi(Y_T)\right] \tag{3.2}$$

for functions $\phi : \mathbb{R}^d \to \mathbb{R}$ is of great interest in many applications especially in mathematical finance where (3.2) represents the fair price of an option and ϕ plays the role of a (discounted) payoff function. In general, it is not possible to calculate (3.2) exactly and the value has to be estimated. A standard method is to simulate sample paths of the corresponding SDE and to use a (Multilevel) Monte Carlo estimator.
In the case of the Heston model, we are interested in the approximation of

$$\mathbb{E}\left[g(S_T, V_T)\right]$$

with $g : [0,\infty) \times [0,\infty) \to \mathbb{R}$. Since most of the time we will use the log-Heston price, we replace g by a function $f : \mathbb{R} \times [0,\infty) \to \mathbb{R}$ with $f(x, v) = g(exp(x), v)$. The value of interest is then

$$\mathbb{E}\left[f(X_T, V_T)\right]. \tag{3.3}$$

Note that in many financial applications the value of f only depends on X_T.

A large number of research articles has been published on the efficient simulation of the log-Heston price. The main difficulty of the log-Heston Model is to efficiently simulate the CIR process. In several articles (e.g. [14], [69], [74]) direct simulation via the non-central chi-square distribution is used. Some researchers proposed algorithms to approximate this distribution for a faster simulation (e.g. [6], [72]). Time-discrete

methods such as the Euler scheme are nevertheless popular for the CIR process because of their simplicity and fast computation times. Due to the square root in its diffusion term, a well-defined approximation scheme of this kind must preserve the positivity of the CIR process. Since the standard Euler scheme does not have this property, several Euler-type schemes were proposed that avoid negative values (see [67] for a summary and a numerical comparison). For the log-price process, most of the simulation methods then use a simple Euler or trapezoidal scheme.
In the next three sections, we present a number of simulation schemes for the log-Heston model that were proposed in the scientific literature. Let us remark that this presentation is by no means complete. Rather, it should give an impression of the challenges that arise from simulating the log-Heston price and its variance. At the end of this chapter, we give a brief summary of the standard and the multilevel Monte Carlo estimator for the value (3.2). We also explain why this motivates the weak and strong error estimation in Chapters 6 and 7.

3.1 (Almost) Exact Simulation Methods

Broadia and Kaya [14] were the first ones to develop an exact simulation method for the log-Heston Model. Although their approach is very valuable from a theoretical point of view, it comes with the disadvantage of high computational costs and is therefore considerably slower than other algorithms. First, they simulate the CIR process from the non-central chi-squared distribution. Looking at the SDEs (2.1), one crucial idea of their simulation is then to substitute the integral equation for $(V_t)_{t\in[0,T]}$ into the equation for $(X_t)_{t\in[0,T]}$. For any $s,t\in[0,T]$ with $s<t$, we have

$$\int_s^t V_u dW_u = \frac{1}{\sigma}\left(V_t - V_s - \kappa\theta(t-s) + \kappa\int_s^t V_u du\right).$$

Since the term on the left side also appears in the integral equation of X, we can substitute it as follows:

$$\begin{aligned} X_t = X_s + r(t-s) &+ \frac{\rho}{\sigma}\left(V_t - V_s - \kappa\theta(t-s)\right) \\ &+ \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)\int_s^t V_u du + \sqrt{1-\rho^2}\int_s^t \sqrt{V_u} dB_u. \end{aligned} \tag{3.4}$$

Now, the Brownian motion W disappears from this equation which is very convenient for the simulation and also for the theoretical analysis. Later, we refer to this as the *Broadie-Kaya trick*.

For the next step, Broadie and Kaya derived the Laplace transform of $\int_s^t V_u du$ and calculated the characteristic function and the cumulative distribution function from there. Then, the latter is evaluated by a trapezoidal rule with a finite step size (which leads to discretization and truncation errors). To sample now from the distribution of

$\int_s^t V_u du$, they use the inverse transformation method. This again causes an error since either Newton's method, a bisection search or a similar method has to be applied.

Having generated samples of $\int_s^t V_u du$, the simulation of (3.4) is now straightforward. Since $(V_t)_{t\in[0,T]}$ and $(B_t)_{t\in[0,T]}$ are independent, we can generate a normal random variable with mean 0 and variance $\int_s^t V_u du$ for the last integral from (3.4).

Nevertheless, the algorithm has only a theoretical relevance since the simulation of $\int_s^t V_u du$ is very costly which is due to the evaluation of its characteristic function. In [69], the computational time of the Broadie/Kaya algorithm is reduced by precaching values for the characteristic function. Still, this simulation method is not widely used in practice.

3.2 Semi-Exact Simulation Methods

The class of semi-exact simulation methods for the log-Heston model mostly contains algorithms that simulate the CIR process exactly or approximately from the non-central chi-square distribution and use a simple Euler or trapezoidal discretization for the log-price process.
We denote the discretization grid as

$$0 = t_0 < t_1 < ... < t_N = T$$

and the increment of the Brownian motions as

$$\Delta_k W = W_{t_{k+1}} - W_{t_k}, \quad \Delta_k B = B_{t_{k+1}} - B_{t_k}$$

for $k \in \{0, ..., N-1\}$. Here, $N \in \mathbb{N}$ is the number of time-steps. By using the Broadie-Kaya trick and discretizing (3.4) with the Euler scheme, we get the following iteration for $k \in \{0, ..., N-1\}$:

$$\begin{aligned} x_{k+1} = x_k + r(t_{k+1} - t_k) + \frac{\rho}{\sigma}\left(v_{k+1} - v_k - \kappa\theta(t_{k+1} - t_k)\right) + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) v_k(t_{k+1} - t_k) \\ + \sqrt{1-\rho^2}\sqrt{v_k}\Delta_k B \end{aligned} \tag{3.5}$$

where we set $x_0 = \log(S_0)$. The values v_k are here simulated from the conditional distribution of the CIR process. In [54], we presented and analyzed a so-called *semi-trapezoidal scheme* for the log-price process:

$$\begin{aligned} x_{k+1} &= x_k + r(t_{k+1} - t_k) + \frac{\rho}{\sigma}\left(v_{k+1} - v_k - \kappa\theta(t_{k+1} - t_k)\right) \\ &\quad + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)\frac{1}{2}\left(v_{k+1} + v_k\right)\left(t_{k+1} - t_k\right) + \sqrt{1-\rho^2}\sqrt{v_k}\Delta_k B. \end{aligned} \tag{3.6}$$

Many algorithms (e.g. [6], [72], [74]) use a full trapezoidal discretization of X as follows:

$$\begin{aligned} x_{k+1} = x_k + r(t_{k+1} - t_k) + \frac{\rho}{\sigma}\left(v_{k+1} - v_k - \kappa\theta(t_{k+1} - t_k)\right) \\ + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)\frac{1}{2}\left(v_{k+1} + v_k\right)\left(t_{k+1} - t_k\right) + \sqrt{1-\rho^2}\frac{1}{2}\left(\sqrt{v_{k+1}} + \sqrt{v_k}\right)\Delta_k B. \end{aligned} \tag{3.7}$$

One disadvantage of the exact simulation from the non-central chi-squared distribution is that it heavily depends on the Feller index ν. Recall that ν determines the degrees of freedom of the non-central chi squared distribution (see Proposition 2.4) and as a result, low Feller indices cause long computational times. Therefore, many schemes were proposed that simulate the CIR process by approximating the non-central chi-squared distribution. The QE-scheme from [6] starts with $v_0 = V_0$ and simulates v_{k+1} depending on the value of v_k either as a moment-matched squared Gaussian random variable or as an ordinary chi-squared random variable. The latter is used for low values of v_k. The log-price process is then simulated according to (3.7). The NCI-scheme from [72] simulates from the non-central chi-squared distribution via direct inversion and uses precaching. Again, the full trapezoidal discretization is used for the log-price process.

3.3 Time-Discrete Simulation Methods

Even though a lot of research concerning the development of exact simulation methods has been carried out, simple time-discrete simulation methods for the log-Heston model are a highly relevant topic for researchers since they are not only interesting from a scientific point of view but also very relevant for practical use. Since this chapter is restricted to the presentation of the different schemes, a survey of the respective weak and strong convergence results from the literature will be given in Chapter 6 and Chapter 7.

3.3.1 Explicit Euler schemes

Euler schemes are very popular in practice since they are very easy to implement. The challenge of simulating the log-Heston model with Euler schemes is once again the simulation of the CIR process. A naive Euler scheme would look like this:

$$v_{k+1} = v_k + \kappa\left(\theta - v_k\right)\left(t_{k+1} - t_k\right) + \sigma\sqrt{v_k}\Delta_k W, \quad v_0 = V_0$$

This leads to

$$\mathbb{P}\left(v_{k+1} < 0\right) = \Phi\left(\frac{v_k - \kappa(\theta - v_k)(t_{k+1} - t_k)}{\sigma\sqrt{v_k(t_{k+1} - t_k)}}\right)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. Therefore, the probability of simulating a negative value during the iteration is strictly positive and the scheme is not well defined due to the square root coefficient. To prevent this, the Euler scheme must be "fixed". A summary of the existing Euler schemes for the CIR process and a numerical comparison can be found in [67], where a general framework for Euler schemes for the CIR process is proposed as

$$\begin{aligned} \bar{v}_{k+1} &= f_1(\bar{v}_k) + \kappa\left(\theta - f_2(\bar{v}_k)\right)(t_{k+1} - t_k) + \sigma\sqrt{f_3(\bar{v}_k)}\Delta_k W \\ v_{k+1} &= f_3(\bar{v}_{k+1}) \end{aligned} \tag{3.8}$$

where $\bar{v}_0 = v_0 = V_0$ and suitable functions f_i that are chosen from

$$\begin{aligned} \texttt{id} : \mathbb{R} \to \mathbb{R}, &\quad \texttt{id}(x) = x, \\ \texttt{abs} : \mathbb{R} \to [0, \infty), &\quad \texttt{abs}(x) = x^+, \\ \texttt{sym} : \mathbb{R} \to [0, \infty), &\quad \texttt{sym}(x) = |x|. \end{aligned}$$

Table 3.1 shows all Euler schemes that are presented in [67] in detail.

Scheme	$f_1(x)$	$f_2(x)$	$f_3(x)$
Absorption (AE)	$(x)^+$	$(x)^+$	$(x)^+$
Symmetrization (SE)	$\|x\|$	$\|x\|$	$\|x\|$
Higham and Mao (HM)	x	x	$\|x\|$
Partial Truncation Euler (PTE)	x	x	$(x)^+$
Full Truncation Euler (FTE)	x	$(x)^+$	$(x)^+$

Table 3.1: Euler schemes from [67].

The full truncation Euler was introduced in the same paper. The origin of the Euler with absorption fix is unknown, the symmetrized Euler was analyzed in [10]. The scheme from Higham and Mao was first analyzed in [37] and the partial truncation Euler was first introduced in [23]. The log-price process can then be discretized with the standard Euler scheme:

$$x_{k+1} = x_k + \left(r - \frac{1}{2}v_k\right)(t_{k+1} - t_k) + \sqrt{v_k}\left(\rho\Delta_k W + \sqrt{1-\rho^2}\Delta_k B\right) \tag{3.9}$$

with $x_0 = \log(S_0)$.

3.3.2 Milstein schemes

The CIR process can also be discretized with an implicit Milstein scheme.

$$\begin{aligned} v_{k+1} = v_k &+ \kappa\left(\theta - v_{k+1}\right)(t_{k+1} - t_k) + \sigma\sqrt{v_k}\Delta_k W \\ &+ \frac{\sigma^2}{4}\left((\Delta_k W)^2 - (t_{k+1} - t_k)\right) \end{aligned}$$

with $v_0 = V_0$. This can be rewritten to

$$v_{k+1} = \frac{1}{1+\kappa(t_{k+1}-t_k)} \left(\left(\sqrt{v_k} + \frac{\sigma}{2}\Delta_k W \right)^2 + \left(\kappa\theta - \frac{\sigma^2}{4} \right) (t_{k+1}-t_k) \right) \tag{3.10}$$

where we can immediately see that this scheme is positivity preserving and therefore well-defined for $\nu \geq \frac{1}{2}$. In [5], this scheme was combined with the standard Euler scheme (3.9) for the log-price process. In [48] the authors propose the discretization

$$\begin{aligned} x_{k+1} = x_k + r\,(t_{k+1}-t_k) - \frac{1}{4}\,(v_{k+1}+v_k)\,(t_{k+1}-t_k) + \rho\sqrt{v_k}\Delta_k W \\ + \sqrt{1-\rho^2}\frac{1}{2}\left(\sqrt{v_k}+\sqrt{v_{k+1}}\right)\Delta_k B + \frac{1}{4}\rho\sigma\left((\Delta_k W)^2 - (t_{k+1}-t_k) \right) \end{aligned}$$

for the log-price process. Together, this is called the IJK-IMM scheme. In [32], the following truncated Milstein scheme for the CIR process was analyzed:

$$\begin{aligned} v_{k+1} = \Bigg(\Bigg(\max\left\{ \sqrt{\frac{\sigma^2}{4}(t_{k+1}-t_k)}, \sqrt{\max\left\{ \frac{\sigma^2}{4}(t_{k+1}-t_k), v_k \right\}} + \frac{\sigma}{2}\Delta_k W \right\} \Bigg) \Bigg)^2 \\ + \left(\kappa(\theta - v_k) - \frac{\sigma^2}{4} \right)(t_{k+1}-t_k) \Bigg)^+ . \end{aligned} \tag{3.11}$$

This scheme is well-defined for the whole parameter range.

3.3.3 Drift-implicit Euler schemes

Another way to discretize the CIR process is to look first at its Lamperti transformation. This was first proposed in [1]. Therefore, we consider the process $Z_t = \sqrt{V_t}$. With the Itō formula, we obtain

$$dZ_t = \left(\frac{4\kappa\theta - \sigma^2}{8}\frac{1}{Z_t} - \frac{\kappa}{2}Z_t \right) dt + \frac{\sigma}{2}dW_t, \quad Z_0 = \sqrt{V_0}.$$

The drift-implicit Euler scheme for this process is given by

$$\begin{aligned} z_{k+1} &= z_k + \left(\frac{4\kappa\theta - \sigma^2}{8}\frac{1}{z_{k+1}} - \frac{\kappa}{2}z_{k+1} \right)(t_{k+1}-t_k) + \frac{\sigma}{2}\Delta_k W. \\ v_{k+1} &= z_{k+1}^2 \end{aligned} \tag{3.12}$$

with $z_0 = \sqrt{v_0} = \sqrt{V_0}$. The first line of (3.12) can be rewritten as

$$z_{k+1} = \frac{z_k + \frac{\sigma}{2}\Delta_k W}{2+\kappa\,(t_{k+1}-t_k)} + \sqrt{\frac{\left(z_k + \frac{\sigma}{2}\Delta_k W\right)^2}{(2+\kappa\,(t_{k+1}-t_k))^2} + \frac{\left(\kappa\theta - \frac{\sigma^2}{4}\right)(t_{k+1}-t_k)}{2+\kappa\,(t_{k+1}-t_k)}}. \tag{3.13}$$

Again, this is well-defined and positivity preserving if $\nu \geq \frac{1}{2}$. In [4], the authors propose a method to approximate $p = \mathbb{E}\left[h(S_T)\right]$ especially for discontinuous functions h. They prove that

$$p = \mathbb{E}\left[\frac{H(S_T)}{S_T}\Pi\right]$$

where $H : \mathbb{R}^+ \to \mathbb{R}$ is the antiderivative of h and

$$\Pi = 1 + \frac{1}{T\sqrt{1-\rho^2}}\int_0^T \frac{1}{\sqrt{V_t}}dB_t$$

is a Malliavin weight. They use the drift-implicit Euler for the Lamperti transformation of the CIR process and the standard Euler for the price process and the Malliavin weight.

3.4 Monte Carlo Methods

The standard Monte Carlo algorithm is a straightforward way to approximate the expectation of a random variable which can be simulated exactly or approximately. In the first case, the standard estimator for (3.2) is given by

$$\hat{p}_M = \frac{1}{M}\sum_{i=1}^{M}\phi\left(Y_T^{(i)}\right)$$

where $Y_T^{(i)}$ for $i = 1, ..., M$ are M iid copies of Y_T. We can also define the Monte Carlo estimator if we do not simulate Y_T exactly. Let y_N be an approximation of Y_T which was simulated via some time-discrete scheme with N time steps. Then, the standard Monte Carlo estimator for (3.2) is given by

$$\hat{p}_{N,M} = \frac{1}{M}\sum_{i=1}^{M}\phi\left(y_N^{(i)}\right)$$

where $y_N^{(i)}, i = 1, ..., M$ are iid copies of y_N. The root mean square or L^2-error of the estimator depends on the variance of the estimator and its bias:

$$\begin{aligned} rmsq\left(\hat{p}_{N,M}\right) = ||p - \hat{p}_{N,M}||_{L^2} &= \mathbb{E}[|p - \hat{p}_{N,M}|^2]^{1/2} \\ &= \left(Var(\hat{p}_{N,M}) + |\mathbb{E}[p - \hat{p}_{N,M}]|^2\right)^{1/2} \\ &= \left(\frac{1}{M}Var\left(\phi\left(y_N\right)\right) + |\mathbb{E}[p - \hat{p}_{N,M}]|^2\right)^{1/2} \end{aligned}$$

In the case of an exact simulation, the bias is zero and the error only depends on the first term, i.e.

$$rmsq\left(\hat{p}_M\right) = \frac{1}{\sqrt{M}}\sqrt{Var\left(\phi(Y_T)\right)}.$$

The computational costs of the exact estimator are $O(M)$. If we want to achieve an accuracy of

$$rmsq\left(\hat{p}_M\right) \leq \varepsilon,$$

we have $cost(\hat{p}_M) = O(\varepsilon^{-2})$. This leads to the following error-cost relation:

$$||p - \hat{p}_M||_{L^2} \leq C_{var} \cdot cost(\hat{p}_M)^{-\frac{1}{2}}.$$

assuming that $Var\left(\phi(Y_T)\right)$ is bounded. For non-exact schemes, the weak convergence order α plays an important role for the relation of the root mean square error and the computational costs. The *weak error* is defined by

$$e_{weak}(N) := |\mathbb{E}\left[\phi\left(y_N\right)\right] - \mathbb{E}\left[\phi\left(Y_T\right)\right]|.$$

We say that a scheme has a weak convergence order α if

$$e_{weak}(N) \leq C_\alpha \cdot N^{-\alpha} \tag{3.14}$$

for an $\alpha \in [0, \infty)$ and a constant $C_\alpha > 0$ which does not depend on N. More precisely, the weak convergence order is the largest α for which (3.14) holds. The computational costs of the standard Monte Carlo method for time-discrete schemes are $O\left(N \cdot M\right)$. To illustrate the impact of α in our Monte Carlo simulation let us now assume that we have an equidistant discretization, i.e

$$t_k = \frac{kT}{N},\ k = 0, ..., N$$

and that $Var(\phi(y_N))$ is bounded. Balancing of N and M leads to an optimal choice of $M = \lceil N^{2\alpha} \rceil$ (see e.g. [25]). Again, if we want to achieve an accuracy of ε, i.e.

$$rmsq\left(\hat{p}_{N,M}\right) \leq \varepsilon,$$

the computational costs of the estimator behave in the following way:

$$cost(\hat{p}_{N,M}) = O\left(\varepsilon^{-2-\frac{1}{\alpha}}\right).$$

The relation between L^2-error and computational costs can then be described as

$$||p - \hat{p}_{N,M}||_{L^2} \leq C_{\alpha,var} \cdot cost(\hat{p}_{N,M})^{-\frac{\alpha}{1+2\alpha}}.$$

This emphasizes how important it is to know the weak convergence order for non-exact schemes. Low weak error orders will slow down the convergence speed of the Monte Carlo estimator drastically.

The efficiency of the standard Monte Carlo method can be significantly improved by combining standard Monte Carlo estimators of different step sizes. This idea was first used in the context of parametrical integration problems in [35]. Let $L > 2$ be the number of levels, $0 \leq N_0 < N_1 < ... < N_L$ the number of steps and $M_0, M_1, ..., M_L$ the number of Monte Carlo samples for every level. The Multilevel Monte Carlo (MLMC) estimator $\hat{p}_{ML}$ is defined as

$$\hat{p}_{ML}^{M_0} = \frac{1}{M_0}\sum_{i=1}^{M_0} \phi\left(y_{N_0}^{(i)}\right)$$

$$\hat{p}_{ML}^{M_l} = \frac{1}{M_l}\sum_{i=1}^{M_l} \left(\phi\left(y_{N_l}^{(i)}\right) - \phi\left(y_{N_{l-1}}^{(i)}\right)\right), l = 1, ..., L$$

$$\hat{p}_{ML} = \sum_{l=0}^{L} \hat{p}_{ML}^{M_l}.$$

In [28] this was first used for SDEs. Simulating $\phi\left(y_{N_l}^{(i)}\right)$ and $\phi\left(y_{N_{l-1}}^{(i)}\right)$ from the same Brownian path guarantees a low variance of the estimator. The computational costs of the MLMC method are proportional to the total number of discretization steps:

$$cost(\hat{p}_{ML}) = C\left(M_0 N_0 + \sum_{l=1}^{L} M_l(N_l + N_{l-1})\right).$$

For MLMC, the knowledge of the strong error is crucial. There does not exist a uniform definition of the strong error in the literature. We can analyze the global error which is

$$e_{strong}^{(1a)} := \mathbb{E}\left[\sup_{t\in[0,T]} |Y_t - \hat{y}_t|^p\right]^{\frac{1}{p}}, \quad p \geq 1$$

where $(\hat{y}_t)_{t\in[0,T]}$ is a time-continuous version of the numerical scheme. Sometimes it is easier to study

$$e_{strong}^{(1b)} := \sup_{t\in[0,T]} \mathbb{E}\left[|Y_t - \hat{y}_t|^p\right]^{\frac{1}{p}}, \quad p \geq 1.$$

Note, that it holds that $e_{strong}^{(1b)} \leq e_{strong}^{(1a)}$. Furthermore, we can look at the maximal error in the discretization points, that is

$$e_{strong}^{(2a)} := \mathbb{E}\left[\max_{k\in\{0,...,N\}} |Y_{t_k} - \hat{y}_{t_k}|^p\right]^{\frac{1}{p}}, \quad p \geq 1$$

or alternatively,

$$e_{strong}^{(2b)} := \max_{k\in\{0,...,N\}} \mathbb{E}\left[|Y_{t_k} - \hat{y}_{t_k}|^p\right]^{\frac{1}{p}}, \quad p \geq 1. \tag{3.15}$$

Analogously, we have $e_{strong}^{(2b)} \leq e_{strong}^{(2a)}$. For the MLMC estimator, the knowledge of (3.15) for $p = 2$ is sufficient. Let α be defined as in (3.14) and let the L^2-error at the terminal time T be of order β, i.e.

$$\mathbb{E}\left[|Y_T - y_N|^2\right] \leq C_\beta N^{-2\beta}$$

with $C_\beta > 0$, $\alpha \geq 1/2$ and $\beta \geq 0$. Furthermore, we assume that the function ϕ is globally Lipschitz continuous. Then, the number of levels L and the number of paths N_l for each level can be chosen in such a way that

$$||p - \hat{p}_{ML}||_{L^2} \leq \varepsilon$$

and there exists a constant $C > 0$ such that

$$cost(\hat{p}_{ML}) \leq C \begin{cases} \varepsilon^{-2} & \text{if } \beta > 1/2 \\ (\log(\varepsilon))^2 \varepsilon^{-2} & \text{if } \beta = 1/2 \\ \varepsilon^{-2-\frac{1-2\beta}{\alpha}} & \text{if } \beta < 1/2 \end{cases}$$

(see Theorem 3.1 in [28]). So for $\beta > 1/2$, the Multilevel Monte Carlo recovers the optimal convergence rate of the standard estimator even for a non-exact simulation of the SDE. Consequently, the knowledge of α and β is crucial for the efficient computation of (3.2) and in particular of (3.3) in the case of the log-Heston model.

Chapter 4

Properties of Explicit Euler Schemes for the log-Heston Model

In this chapter, we study the properties of the explicit Euler schemes from Equations (3.8) with f_i given by

$$f_1 = \mathtt{id}, \quad f_2 \in \{\mathtt{id}, \mathtt{abs}, \mathtt{sym}\}, \quad f_3 \in \{\mathtt{abs}, \mathtt{sym}\} \tag{4.1}$$

or

$$f_1 = f_2 = f_3 \in \{\mathtt{abs}, \mathtt{sym}\}. \tag{4.2}$$

These include all schemes from Table 3.1. The first set of conditions modifies the coefficients of the CIR process to deal with negative values which may arise in the computation. For example, $\sqrt{v_k}$ is replaced by $\sqrt{v_k^+}$ or $\sqrt{|v_k|}$. After the approximation $\bar{v}_{k+1}$ has been computed, f_3 is again applied to obtain v_{k+1}, since $\bar{v}_{k+1}$ may be still negative. The second set of conditions is different. Here after each Euler step, `sym` or `abs` is applied to avoid negative values.
In this chapter, we prove important properties of these two cases of explicit Euler schemes that are crucial for our proofs in Chapters 6 and 7. We need the notation $n(t) := \max\{k \in \{0, ..., N\} : t_k \leq t\}$ and $\eta(t) := t_{n(t)}$.

4.1 Euler Schemes - Case I

For the choice (4.1), the time-continuous extensions of $(v_k)_{k\in\{0,...,N\}}$ which are denoted by $\bar{v} = (\bar{v}_t)_{t\in[0,T]}$ and $\hat{v} = (\hat{v}_t)_{t\in[0,T]}$ read as

$$\begin{aligned} \bar{v}_t &= \bar{v}_{\eta(t)} + \int_{\eta(t)}^{t} \kappa(\theta - f_2(\bar{v}_{\eta(s)}))ds + \sigma \int_{\eta(t)}^{t} \sqrt{f_3(\bar{v}_{\eta(s)})}dW_s, \qquad t \in [0, T], \\ \hat{v}_t &= f_3(\bar{v}_t), \end{aligned} \tag{4.3}$$

with $f_2 \in \{\mathtt{id}, \mathtt{abs}, \mathtt{sym}\}$, $f_3 \in \{\mathtt{abs}, \mathtt{sym}\}$ and $\bar{v}_0 = v$. Note that f_2 and f_3 are globally Lipschitz continuous with Lipschitz constant $L = 1$ and satisfy

$$|x - f_i(y)| \le |x - y|, \qquad x \ge 0, y \in \mathbb{R}, \quad i = 2, 3.$$

Moreover note that

$$\sqrt{|f_i(x)|} \le 1 + |x|, \qquad x \in \mathbb{R}, \quad i = 1, 2, 3.$$

The next lemma shows that the moments of $(\bar{v})_{t\in[0,T]}$ are bounded. Furthermore, we have the same smoothness result as for the CIR process in Lemma 2.3.

Lemma 4.1. *Let $p \ge 1$. There exists a constant $C_p > 0$ such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |\bar{v}_t|^p\right] \le C_p.$$

Furthermore, we also have

$$\sup_{0\le s<t\le T} \mathbb{E}\left[\frac{|\bar{v}_t - \bar{v}_s|^p}{|t-s|^{p/2}}\right] < \infty.$$

Proof. For the first term, we prove that

$$\sup_{t\in[0,T]} \mathbb{E}\left[|\bar{v}_t|^p\right] < \infty. \tag{4.4}$$

Let $p \ge 2$ and let τ_n be the stopping time defined by $\tau_n := \inf\{0 < t < T; \bar{v}_t \ge n\}$ with $\inf\{\emptyset\} = 0$. Then, since $f_i(\bar{v}_{\eta(t)}) \le |\bar{v}_{\eta(t)}|$ for $i \in \{2, 3\}$

$$\begin{aligned}
\mathbb{E}\left[|\bar{v}_{t\wedge\tau_n}|^p\right] &\le \mathbb{E}\left[\left|v_0 + \int_0^{t\wedge\tau_n} \kappa(\theta - f_2(\bar{v}_{\eta(s)}))ds + \sigma\int_0^{t\wedge\tau_n} \sqrt{f_3(\bar{v}_{\eta(s)})}dW_s\right|^p\right] \\
&\le C_p\left(v_0^p + \mathbb{E}\left[\left|\int_0^{t\wedge\tau_n} \kappa(\theta - f_2(\bar{v}_{\eta(s)}))ds\right|^p\right]\right. \\
&\qquad \left. + \mathbb{E}\left[\left|\sigma\int_0^{t\wedge\tau_n} \sqrt{f_3(\bar{v}_{\eta(s)})}dW_s\right|^p\right]\right) \\
&\le C_p\left(1 + \mathbb{E}\left[\left|\int_0^{t\wedge\tau_n} \kappa(\theta - f_2(\bar{v}_{\eta(s)}))ds\right|^p\right]\right. \\
&\qquad \left. + \mathbb{E}\left[\left|\sigma^2\int_0^{t\wedge\tau_n} (1 + |\bar{v}_{\eta(s)}|)^2 ds\right|^{p/2}\right]\right) \\
&\le C_p\left(1 + \mathbb{E}\left[\int_0^{t\wedge\tau_n} |\bar{v}_{\eta(s)}|^p ds\right] + \mathbb{E}\left[\left|\int_0^{t\wedge\tau_n} |\bar{v}_{\eta(s)}|^2 ds\right|^{p/2}\right]\right) \\
&\le C_p\left(1 + \mathbb{E}\left[\int_0^{t\wedge\tau_n} |\bar{v}_{\eta(s)}|^p ds\right]\right)
\end{aligned}$$

by applications of the Hölder and the BDG inequality. Therefore, we have shown that

$$\mathbb{E}\left[|\bar{v}_{t\wedge\tau_n}|^p\right] \leq C_p\left(1+\int_0^t \mathbb{E}\left[|\bar{v}_{\eta(s)\wedge\tau_n}|^p\right] ds\right)$$

and we consequently obtain

$$\sup_{s\in[0,t]} \mathbb{E}\left[|\bar{v}_{s\wedge\tau_n}|^p\right] \leq C_p\left(1+\int_0^t \sup_{u\in[0,s]} \mathbb{E}\left[|\bar{v}_{u\wedge\tau_n}|^p\right] ds\right).$$

The Gronwall inequality now yields

$$\sup_{t\in[0,T]} \mathbb{E}\left[|\bar{v}_{t\wedge\tau_n}|^p\right] \leq C_p$$

where C_p does not depend on n. Taking the limit $n\to\infty$, we obtain (4.4). For $p\in[1,2)$ (4.4) follows then by the Lyapunov inequality. Since we have

$$\sup_{t\in[0,T]} |\bar{v}_t|^p \leq C_p\left(1+\int_0^T \left|\bar{v}_{\eta(s)}\right|^p ds + \sup_{t\in[0,T]}\left|\int_0^t \sigma\sqrt{f_3(\bar{v}_{\eta(s)})}dW_s\right|^p\right)$$

the assertion now follows from the properties of f_3, an application of the BDG inequality and Equation (4.4). For the second statement, we begin again with $p\geq 2$. We have

$$\begin{aligned}
\mathbb{E}\left[|\bar{v}_t-\bar{v}_s|^p\right] &\leq \mathbb{E}\left[\left|\int_s^t \kappa\left(\theta - f_2\left(\bar{v}_{\eta(u)}\right)\right)du + \sigma\int_s^t \sqrt{f_3\left(\bar{v}_{\eta(u)}\right)}dW_u\right|^p\right]\\
&\leq C_p\left(\mathbb{E}\left[\left|\int_s^t \kappa\left(\theta - f_2\left(\bar{v}_{\eta(u)}\right)\right)du\right|^p\right] + \mathbb{E}\left[\left|\sigma\int_s^t \sqrt{f_3\left(\bar{v}_{\eta(u)}\right)}dW_u\right|^p\right]\right)\\
&\leq C_p\Bigg(|t-s|^p + |t-s|^{p-1}\int_s^t \mathbb{E}\left[\left|\bar{v}_{\eta(u)}\right|^p\right]du\\
&\qquad + |t-s|^{p/2-1}\int_s^t \mathbb{E}\left[\left|\bar{v}_{\eta(u)}\right|^{p/2}\right]du\Bigg)\\
&\leq C_p\left(|t-s|^p + |t-s|^{p/2}\right)
\end{aligned}$$

by using Hölder's inequality, (4.4) and the properties of f_2 and f_3. It follows that

$$\sup_{0\leq s<t\leq T} \mathbb{E}\left[\frac{|\bar{v}_t-\bar{v}_s|^p}{|t-s|^{p/2}}\right] \leq C_p\left(T^{p/2}+1\right).$$

Again the application of the Lyapunov inequality for $p\in[1,2)$ finishes the proof. □

4.2 Euler Schemes - Case II

For the choice (4.2), we obtain the symmetrized Euler (SE) and the Euler with absorption fix (AE). We can write the time-continuous extension $\hat{v}^{sym} = (\hat{v}_t^{sym})_{t\in[0,T]}$ of the SE as

$$\hat{v}_t^{sym} = \left| \hat{v}_{\eta(t)}^{sym} + \kappa \left(\theta - \hat{v}_{\eta(t)}^{sym} \right) (t - \eta(t)) + \sigma \sqrt{\hat{v}_{\eta(t)}^{sym}} \left(W_t - W_{\eta(t)} \right) \right|$$

and the time-continuous extension $\hat{v}^{abs} = (\hat{v}_t^{abs})_{t\in[0,T]}$ of the AE as

$$\hat{v}_t^{abs} = \left(\hat{v}_{\eta(t)}^{abs} + \kappa \left(\theta - \hat{v}_{\eta(t)}^{abs} \right) (t - \eta(t)) + \sigma \sqrt{\hat{v}_{\eta(t)}^{abs}} \left(W_t - W_{\eta(t)} \right) \right)^+ .$$

Now, let $\star \in \{sym, abs\}$. We define

$$z_t^\star := \hat{v}_{\eta(t)}^\star + \kappa\big(\theta - \hat{v}_{\eta(t)}^\star\big)(t - \eta(t)) + \sigma \sqrt{\hat{v}_{\eta(t)}^\star} \left(W_t - W_{\eta(t)} \right) \tag{4.5}$$

and use the Tanaka-Meyer formula (see e.g. equation 7.9 in Chapter III in [49]) for $\hat{v}_t^{sym} = \left| z_t^{sym} \right|$ and for $\hat{v}_t^{abs} = \left(z_t^{abs} \right)^+$ to obtain

$$\begin{aligned} \hat{v}_t^{sym} = \hat{v}_{\eta(t)}^{sym} &+ \int_{\eta(t)}^t sign\left(z_s^{sym}\right) \kappa \left(\theta - \hat{v}_{\eta(s)}^{sym} \right) ds + \sigma \int_{\eta(t)}^t sign\left(z_s^{sym}\right) \sqrt{\hat{v}_{\eta(s)}^{sym}} dW_s \\ &+ \left(L_t^0(z^{sym}) - L_{\eta(t)}^0(z^{sym}) \right), \qquad t \in [0,T], \end{aligned}$$

and

$$\begin{aligned} \hat{v}_t^{abs} = \ \hat{v}_{\eta(t)}^{abs} &+ \int_{\eta(t)}^t \mathbb{1}_{\{z_s^{abs}>0\}} \kappa \left(\theta - \hat{v}_{\eta(s)}^{abs} \right) ds + \sigma \int_{\eta(t)}^t \mathbb{1}_{\{z_s^{abs}>0\}} \sqrt{\hat{v}_{\eta(s)}^{abs}} dW_s \\ &+ \frac{1}{2} \left(L_t^0(z^{abs}) - L_{\eta(t)}^0(z^{abs}) \right), \qquad t \in [0,T]. \end{aligned}$$

Here $L^0(z^\star) = (L_t^0(z^\star))_{t\in[0,T]}$ is the local time of $z^\star$ in $z = 0$. For almost all $\omega \in \Omega$ the map $[0,T] \ni t \mapsto [L_t^0(z^\star)](\omega) \in \mathbb{R}$ is continuous and non-decreasing with $L_0^0(z) = 0$. See e.g. Theorem 7.1 in chapter III of [49].

We can rewrite both schemes as

$$\begin{aligned} \hat{v}_t^\star = \hat{v}_{\eta(t)}^\star &+ \int_{\eta(t)}^t \kappa \left(\theta - \hat{v}_{\eta(s)}^\star \right) ds + \sigma \int_{\eta(t)}^t \sqrt{\hat{v}_{\eta(s)}^\star} dW_s \\ &- 2c^\star \sigma \int_{\eta(t)}^t \mathbb{1}_{\{z_s^\star \le 0\}} \sqrt{\hat{v}_{\eta(s)}^\star} dW_s - 2c^\star \int_{\eta(t)}^t \mathbb{1}_{\{z_s^\star \le 0\}} \kappa \left(\theta - \hat{v}_{\eta(s)}^\star \right) ds \\ &+ c^\star \left(L_t^0(z^\star) - L_{\eta(t)}^0(z^\star) \right), \qquad t \in [0,T], \end{aligned} \tag{4.6}$$

with $c^{sym} = 1$ and $c^{abs} = \frac{1}{2}$.

The Euler schemes in this section also have bounded moments and increments.

Lemma 4.2. *Let $\star \in \{sym, abs\}$ and $p \geq 1$. Then, there exists a $C_p > 0$ such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |\hat{v}_t^\star|^p\right] \leq C_p. \tag{4.7}$$

Furthermore, we have

$$\sup_{t\in[0,T]} \mathbb{E}\left[\frac{|\hat{v}_t^\star - \hat{v}_{\eta(t)}^\star|^p}{|t-\eta(t)|^{p/2}}\right] < \infty.$$

Proof. The proof of (4.7) can be found in Lemma 2.1 in [10] for the symmetrized Euler scheme and can be obtained analogously for the absorbed Euler scheme. For the second statement, we give a proof for the absorbed Euler. The proof for the symmetrized Euler can be done analogously. We drop the *abs*-label to simplify the notation. Since

$$\left|(v+z)^+ - v\right| \leq |z|$$

for $v > 0$ and $z \in \mathbb{R}$, we have that

$$\begin{aligned}
\left|\hat{v}_t - \hat{v}_{\eta(t)}\right|^p &= \left|\left(\hat{v}_{\eta(t)} + \kappa\left(\theta - \hat{v}_{\eta(t)}\right)(t-\eta(t)) + \sigma\sqrt{\hat{v}_{\eta(t)}}\left(W_t - W_{\eta(t)}\right)\right)^+ - \hat{v}_{\eta(t)}\right|^p \\
&\leq \left|\kappa\left(\theta - \hat{v}_{\eta(t)}\right)(t-\eta(t)) + \sigma\sqrt{\hat{v}_{\eta(t)}}\left(W_t - W_{\eta(t)}\right)\right|^p \\
&\leq 2^{p-1}(t-\eta(t))^p\left|\kappa\left(\theta - \hat{v}_{\eta(t)}\right)\right|^p + 2^{p-1}\sigma^p\hat{v}_{\eta(t)}^{p/2}\left|W_t - W_{\eta(t)}\right|^p.
\end{aligned}$$

Using (4.7), we obtain

$$\mathbb{E}\left[\left|\kappa\left(\theta - \hat{v}_{\eta(t)}\right)\right|^p\right] \leq C_p$$

and

$$\mathbb{E}\left[\hat{v}_{\eta(t)}^{p/2}\left|W_t - W_{\eta(t)}\right|^p\right] \leq \left(\mathbb{E}\left[\hat{v}_{\eta(t)}^p\right]\right)^{\frac{1}{2}}\left(\mathbb{E}\left[\left|W_t - W_{\eta(t)}\right|^{2p}\right]\right)^{\frac{1}{2}} \leq C_p\,(t-\eta(t))^{\frac{p}{2}}.$$

Therefore, we have

$$\mathbb{E}\left[\left|\hat{v}_t - \hat{v}_{\eta(t)}\right|^p\right] \leq C_p\left(|t-\eta(t)|^p + |t-\eta(t)|^{p/2}\right)$$

and the statement follows. □

For the remainder of this section, we will assume that the discretization is equidistant. So, our discretization grid is defined as

$$t_k = k\Delta t, \quad k = 0, \ldots, N$$

with $\Delta t := T/N$. We are now interested in the probability of $z_t^\star$ becoming less or equal to 0. The next lemma is similar to Lemma 3.7 in [10].

Lemma 4.3. *Let $\Delta t < \frac{1}{\kappa}$. We have*

$$P(z_t^\star \le 0) \le \mathbb{E}\left[\exp\left(-\frac{\hat{v}^\star_{\eta(t)}(1-\kappa\Delta t)^2}{2\sigma^2\Delta t}\right)\right], \qquad t \in [0,T],$$

for $\star \in \{sym, abs\}$.

Proof. First, note that

$$P(z_t^{sym} = 0 | \hat{v}^{sym}_{\eta(t)} = y) = P(z_t^{abs} = 0 | \hat{v}^{abs}_{\eta(t)} = y) = 0, \qquad y \ge 0,\, t \in (0,T],$$

and so $P(z_t^\star = 0) = 0$ for all $t \in [0,T]$, and $\star \in \{sym, abs\}$. Therefore, we only need to consider $P(z_t^\star < 0)$. By the definition of z in (4.5), we have

$$P(z_t < 0 | \hat{v}^\star_{\eta(t)} = y) \le P\left(W_t - W_{\eta(t)} < \frac{-y(1-\kappa(t-\eta(t))) - \kappa\theta(t-\eta(t))}{\sigma\sqrt{y}}\right)$$

for $y > 0$ and

$$P(z_t < 0 | \hat{v}_{\eta(t)} = 0) = 0.$$

For a centered Gaussian random variable G with variance $\zeta^2 > 0$, it holds that

$$P(G < \beta) \le \exp\left(-\frac{\beta^2}{2\zeta^2}\right)$$

for $\beta < 0$. Therefore, we have

$$\begin{aligned}
P(z_t \le 0) &\le \mathbb{E}\left[\exp\left(-\frac{\left(\hat{v}^\star_{\eta(t)}(1-\kappa(t-\eta(t))) + \kappa\theta(t-\eta(t))\right)^2}{2\sigma^2\hat{v}^\star_{\eta(t)}(t-\eta(t))}\right)\mathbb{1}_{\{\hat{v}^\star_{\eta(t)}>0\}}\right] \\
&\le \mathbb{E}\left[\exp\left(-\frac{\left(\hat{v}^\star_{\eta(t)}(1-\kappa(t-\eta(t)))\right)^2}{2\sigma^2\hat{v}^\star_{\eta(t)}(t-\eta(t))}\right)\right] \\
&= \mathbb{E}\left[\exp\left(-\frac{\hat{v}^\star_{\eta(t)}(1-\kappa(t-\eta(t)))^2}{2\sigma^2(t-\eta(t))}\right)\right].
\end{aligned}$$

Since

$$\frac{(1-\kappa(t-\eta(t)))^2}{t-\eta(t)} \ge \frac{(1-\kappa\Delta t)^2}{\Delta t}, \qquad t \in [0,T] \setminus \{t_0, t_1, \ldots, t_N\}$$

the assertion follows.

□

For the further control of $P(z_t^\star \le 0)$ we will need the following technical result on a sequence that was analyzed by Cozma and Reisinger in [21]. We are now giving a different and simplified bound which is crucial for our error analysis.

Lemma 4.4. *Suppose that $\Delta t < \frac{1}{\kappa}$ and set*

$$\alpha_N = \frac{1-\kappa\Delta t}{2}.$$

(i) Consider the sequence $(c_j)_{0\leq j\leq N}$ with

$$c_0 = \alpha_N, \qquad c_1 = \alpha_N - \alpha_N^2, \qquad c_{j+1} = c_j^2 + \alpha_N - \alpha_N^2, \quad j = 1,\dots,N-1.$$

Then, we have

$$c_j \leq 1 - \alpha_N - \frac{\varepsilon(1-\varepsilon)}{1+\varepsilon(j-1)}, \qquad j = 1,\dots,N,$$

for all $\varepsilon \in (0, 1/2]$.
(ii) Define the sequence $(a_j)_{0\leq j\leq N}$ by

$$a_j = \frac{2(\alpha_N - c_j)}{\sigma^2\Delta t}, \qquad j = 0,\dots,N.$$

Then, we have $a_j \geq 0$ for $j = 0,\dots,N$. Moreover, let $\varepsilon \in (0, 1/2]$ and

$$c = \exp\left(\kappa\left(\nu T + \frac{2v_0}{\sigma^2}\right)\right)\left(\max\left\{1, \frac{\sigma^2\nu}{v_0 e}\right\}\right)^{\nu}. \tag{4.8}$$

Then, we have

$$\exp\left(-\kappa\theta\sum_{j=0}^{k-1} a_{j+1}\Delta t\right)\exp\left(-v_0 a_{k+1}\right) \leq c\left(\frac{\Delta t}{\varepsilon}\right)^{\nu(1-\varepsilon)}$$

for all $k = 1,\dots,N$.

Proof. (i) Since $\Delta t \in (0, \frac{1}{\kappa})$, we know that

$$\alpha_N = \frac{1-\kappa\Delta t}{2} < \frac{1}{2}, \qquad \alpha_N > \frac{1-\kappa\frac{1}{\kappa}}{2} = 0$$

and therefore $\alpha_N \in (0, 1/2)$. Now let $\varepsilon \in (0, 1/2]$. We show that

$$c_j \leq 1 - \alpha_N - \frac{1-\varepsilon}{j-1+\varepsilon^{-1}}, \qquad j = 1,\dots,N,$$

by induction. For $j = 1$, we have

$$c_1 = \alpha_N - \alpha_N^2 = 1 - \alpha_N - (1-\alpha_N)^2 \leq 1 - \alpha_N - \frac{1}{4} \leq 1 - \alpha_N - \frac{1-\varepsilon}{\varepsilon^{-1}},$$

since $1/4 \geq (1-\varepsilon)\varepsilon$.

Suppose that the statement holds for a fixed $j \in \{1, ..., N\}$. Then, we have

$$
\begin{aligned}
c_{j+1} = \alpha_N - \alpha_N^2 + c_j^2 \leq & \alpha_N - \alpha_N^2 + \left(1 - \alpha_N - \frac{1-\varepsilon}{j-1+\varepsilon^{-1}}\right)^2 \\
= & \alpha_N - \alpha_N^2 + 1 - 2\alpha_N - 2\frac{1-\varepsilon}{j-1+\varepsilon^{-1}} + \alpha_N^2 + 2\alpha_N \frac{1-\varepsilon}{j-1+\varepsilon^{-1}} \\
& + \frac{(1-\varepsilon)^2}{(j-1+\varepsilon^{-1})^2} \\
= & 1 - \alpha_N - 2(1-\alpha_N)\frac{1-\varepsilon}{j-1+\varepsilon^{-1}} + \frac{(1-\varepsilon)^2}{(j-1+\varepsilon^{-1})^2}.
\end{aligned}
$$

For the statement to be true, it must hold that

$$
\frac{1-\varepsilon}{j-1+\varepsilon^{-1}} - \frac{(1-\varepsilon)^2}{(j-1+\varepsilon^{-1})^2} \geq \frac{1-\varepsilon}{j+\varepsilon^{-1}}
$$

since $2(1-\alpha_N) \in (1, 2)$. This can be verified by a simple computation.

(ii) Since $c_{j+1} = c_j^2 + \alpha_N - \alpha_N^2$ and $c_0 = \alpha_N$, $c_1 = \alpha_N - \alpha_N^2 \leq \alpha_N$, we can establish by induction that $c_j \leq \alpha_N$. Since

$$
a_j = \frac{2(\alpha_N - c_j)}{\sigma^2 \Delta t}
$$

we therefore have $a_j \geq 0$ for $j = 0, \ldots, N$. It follows that

$$
\begin{aligned}
-\kappa\theta \sum_{j=0}^{k-1} a_{j+1}\Delta t = \frac{2\kappa\theta}{\sigma^2}\sum_{j=0}^{k-1}(c_{j+1} - \alpha_N) \leq & \frac{2\kappa\theta}{\sigma^2}\sum_{j=0}^{k-1}\left(1 - 2\alpha_N - \frac{1-\varepsilon}{j+\varepsilon^{-1}}\right) \\
\leq & \frac{2\kappa\theta}{\sigma^2}\int_0^k \left(1 - 2\alpha_N - \frac{1-\varepsilon}{j+\varepsilon^{-1}}\right) dj \\
= & \frac{2\kappa\theta(1-\varepsilon)}{\sigma^2}\left(\ln(\varepsilon^{-1}) - \ln(k+\varepsilon^{-1})\right) \\
& + \frac{2\kappa\theta}{\sigma^2}\kappa\Delta t k \\
\leq & \nu(1-\varepsilon)\ln\left(\frac{1}{1+\varepsilon k}\right) + \kappa\nu T.
\end{aligned}
$$

Using the definition of a_{k+1} and α_N, as well as the estimate for c_{k+1} from (i) we obtain

$$
\begin{aligned}
\exp\left(-v_0 a_{k+1}\right) = & \exp\left(\frac{2v_0}{\sigma^2\Delta t}\left(c_{k+1} - \alpha_N\right)\right) \\
\leq & \exp\left(\frac{2v_0}{\sigma^2\Delta t}\left(1 - 2\alpha_N - \frac{\varepsilon(1-\varepsilon)}{1+\varepsilon k}\right)\right) \\
\leq & \exp\left(\frac{2v_0\kappa}{\sigma^2}\right)\exp\left(-\frac{2v_0\varepsilon(1-\varepsilon)}{\sigma^2}\frac{1}{\Delta t}\frac{1}{1+\varepsilon k}\right) \\
\leq & \exp\left(\frac{2v_0\kappa}{\sigma^2}\right)\exp\left(-\frac{v_0\varepsilon}{\sigma^2}\frac{1}{\Delta t}\frac{1}{1+\varepsilon k}\right),
\end{aligned}
$$

since we have $\varepsilon \in (0, 1/2]$. Thus, we obtain

$$\begin{aligned}
&\exp\left(-\kappa\theta\sum_{j=0}^{k-1} a_{j+1}\Delta t\right)\exp\left(-v_0 a_{k+1}\right)\\
&\quad\leq \exp\left(\kappa\left(\nu T + \frac{2v_0}{\sigma^2}\right)\right)\exp\left(-\frac{v_0\varepsilon}{\sigma^2}\frac{1}{\Delta t}\frac{1}{1+\varepsilon k}\right)\left(\frac{1}{1+\varepsilon k}\right)^{\nu(1-\varepsilon)}\\
&\quad= \exp\left(\kappa\left(\nu T + \frac{2v_0}{\sigma^2}\right)\right)\exp\left(-\frac{v_0\varepsilon}{\sigma^2}\frac{1}{\Delta t}\frac{1}{1+\varepsilon k}\right)\left(\frac{v_0\varepsilon}{\sigma^2}\frac{1}{\Delta t}\frac{1}{1+\varepsilon k}\frac{\sigma^2\Delta t}{v_0\varepsilon}\right)^{\nu(1-\varepsilon)}.
\end{aligned}$$

The inequality

$$x^\alpha \exp(-x) \leq \alpha^\alpha \exp(-\alpha), \qquad \alpha > 0,\, x > 0,$$

and using again that $\varepsilon \in (0, 1/2]$ now yield

$$\begin{aligned}
&\exp\left(-\kappa\theta\sum_{j=0}^{k-1} a_{j+1}\Delta t\right)\exp\left(-v_0 a_{k+1}\right)\\
&\quad\leq \exp\left(\kappa\left(\nu T + \frac{2v_0}{\sigma^2}\right)\right)\left(\frac{\nu(1-\varepsilon)}{e}\right)^{\nu(1-\varepsilon)}\left(\frac{\sigma^2\Delta t}{v_0\varepsilon}\right)^{\nu(1-\varepsilon)}\\
&\quad\leq \exp\left(\kappa\left(\nu T + \frac{2v_0}{\sigma^2}\right)\right)\left(\frac{\sigma^2\nu}{v_0 e}\right)^{\nu(1-\varepsilon)}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu(1-\varepsilon)}\\
&\quad\leq \exp\left(\kappa\left(\nu T + \frac{2v_0}{\sigma^2}\right)\right)\left(\max\left\{1, \frac{\sigma^2\nu}{v_0 e}\right\}\right)^{\nu}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu(1-\varepsilon)},
\end{aligned}$$

which finishes the proof. □

The next proposition gives an upper bound for the expression from Lemma 4.3. It plays the same role as Lemma 3.6 in [10] and in comparison to this lemma it removes the restriction on ν and also obtains a better estimate in terms of ν for $P(z_t^\star \leq 0)$.

Proposition 4.5. *For $\Delta t < \frac{1}{\kappa}$ and $\varepsilon \in (0, 1/2]$ we have that*

$$\mathbb{E}\left[\exp\left(-\frac{\hat{v}_{t_k}^\star(1-\kappa\Delta t)^2}{2\sigma^2\Delta t}\right)\right] \leq c\left(\frac{\Delta t}{\varepsilon}\right)^{\nu(1-\varepsilon)}, \qquad k = 0, \ldots, N, \tag{4.9}$$

and

$$P(z_t^\star \leq 0) \leq c\left(\frac{\Delta t}{\varepsilon}\right)^{\nu(1-\varepsilon)}, \qquad t \in [0, T] \setminus \{t_0, t_1, \ldots, t_N\}, \tag{4.10}$$

for $\star \in \{sym, abs\}$, where c is given by (4.8).

Proof. Lemma 4.3 and (4.9) directly give (4.10). So it remains to show (4.9).
The first step of this proof is to describe a sequence $(a_j)_{0\le j\le N}$ whose first element is equal to $\frac{(1-\kappa\Delta t)^2}{2\sigma^2\Delta t}$ and which has some suitable properties to bound the term on the left side of (4.9). Suppose that $\Delta t < \frac{1}{\kappa}$. Define the sequence $(a_j)_{0\le j\le N}$ as in the previous Lemma, i.e.

$$a_j = \frac{2(\alpha_N - c_j)}{\sigma^2\Delta t}$$

with

$$c_0 = \alpha_N, \qquad c_1 = \alpha_N - \alpha_N^2, \qquad c_{j+1} = c_j^2 + \alpha_N - \alpha_N^2, \quad j = 1, \dots, N-1,$$

and $\alpha_N = \frac{1-\kappa\Delta t}{2}$. In particular, we have $a_0 = 0$,

$$a_1 = \frac{2\alpha_N^2}{\sigma^2\Delta t} = \frac{(1-\kappa\Delta t)^2}{2\sigma^2\Delta t}$$

and

$$\begin{aligned} a_{j+1} = \frac{2(\alpha_N - c_{j+1})}{\sigma^2\Delta t} = \frac{2(\alpha_N^2 - c_j^2)}{\sigma^2\Delta t} &= \frac{4\alpha_N(\alpha_N - c_j) - 2(\alpha_N - c_j)^2}{\sigma^2\Delta t} \\ &= 2\alpha_N a_j - \frac{1}{2}a_j^2\sigma^2\Delta t. \end{aligned}$$

Next, we take a look at

$$\mathbb{E}\left[\exp\left(-\hat{v}^\star_{t_k} a_i\right)\right] = \mathbb{E}\left[\mathbb{E}\left[\exp\left(-\hat{v}^\star_{t_k} a_i\right)\big|\mathcal{F}_{t_{k-1}}\right]\right]$$

and bound the inner expectation, using that $|v| \ge v$ and $v^+ \ge v$, respectively. We have

$$\begin{aligned} &\mathbb{E}\left[\exp\left(-\hat{v}^\star_{t_k} a_i\right)\big|\mathcal{F}_{t_{k-1}}\right] \\ &\quad\le \mathbb{E}\left[\exp\left(-a_i\left(\kappa\theta\Delta t + \hat{v}^\star_{t_{k-1}}(1-\kappa\Delta t) + \sigma\sqrt{\hat{v}^\star_{t_{k-1}}}\left(W_{t_k} - W_{t_{k-1}}\right)\right)\right)\big|\mathcal{F}_{t_{k-1}}\right] \\ &\quad= \exp\left(-a_i\left(\kappa\theta\Delta t + \hat{v}^\star_{t_{k-1}}(1-\kappa\Delta t)\right)\right)\mathbb{E}\left[\exp\left(-a_i\sigma\sqrt{\hat{v}^\star_{t_{k-1}}}\left(W_{t_k} - W_{t_{k-1}}\right)\right)\big|\mathcal{F}_{t_{k-1}}\right] \\ &\quad= \exp\left(-a_i\left(\kappa\theta\Delta t + \hat{v}^\star_{t_{k-1}}(1-\kappa\Delta t)\right)\right)\exp\left(\frac{1}{2}a_i^2\sigma^2\hat{v}^\star_{t_{k-1}}\Delta t\right). \end{aligned}$$

Since

$$a_{i+1} = a_i(1-\kappa\Delta t) - \frac{1}{2}a_i^2\sigma^2\Delta t,$$

it follows

$$\mathbb{E}\left[\exp\left(-\hat{v}^\star_{t_k} a_i\right)\right] \le \exp\left(-a_i\kappa\theta\Delta t\right)\mathbb{E}\left[\exp\left(-\hat{v}^\star_{t_{k-1}} a_{i+1}\right)\right].$$

Plugging in a_1 and applying this upper bound k times, we arrive at

$$\begin{aligned}\mathbb{E}\left[\exp\left(-\frac{\hat{v}^\star_{t_k}(1-\kappa\Delta t)^2}{2\sigma^2\Delta t}\right)\right] &= \mathbb{E}\left[\exp\left(-\hat{v}^\star_{t_k} a_1\right)\right] \\ &\leq \exp\left(-\kappa\theta\sum_{j=0}^{k-1} a_{j+1}\Delta t\right)\exp\left(-v_0 a_{k+1}\right).\end{aligned}$$

The assertion now follows from the second part of Lemma 4.4. □

We now need an upper bound for the expected local time that $z^\star$ spends in 0. Our proof follows similar ideas as the proof of Proposition 3.5 in [10] but adds the results from Proposition 4.5 to obtain a better convergence estimate.

Proposition 4.6. *Let $\beta > 0$, $\delta > 0$, $\varepsilon \in (0, 1/2]$, $\Delta t \leq \frac{1}{2\kappa}$ and $\star \in \{sym, abs\}$. Then, there exist constants $C_\delta > 0$ and $C_{\beta,\delta} > 0$ such that*

$$\mathbb{E}\left[L^0_t(z^\star) - L^0_{\eta(t)}(z^\star)\right] \leq C_\delta \Delta t \left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{1+\delta}}, \qquad t \in [0, T],$$

and

$$\mathbb{E}\left[\left|L^0_t(z^\star) - L^0_{\eta(t)}(z^\star)\right|^{1+\beta}\right]^{\frac{1}{1+\beta}} \leq C_{\beta,\delta}\left(\Delta t\right)^{\frac{1}{(1+\beta)^2}}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{(1+\delta)(1+\beta)^2}}.$$

Proof. (i) To simplify the notation, we drop the $\star$-label. By the occupation time formula, see e.g. Theorem 7.1 in chapter III of [49], we have for any $t \in [t_k, t_{k+1}]$ and for any non-negative Borel-measurable function $\phi : \mathbb{R} \to \mathbb{R}$ that P-a.s

$$\int_{\mathbb{R}} \phi(x)\left(L^x_t(z) - L^x_{t_k}(z)\right)dx = \int_{t_k}^t \phi(z_s)d\langle z\rangle_s = \sigma^2\int_{t_k}^t \phi(z_s)\hat{v}_{t_k}ds.$$

Here $L^x(z)$ is the local time of z in $x \in \mathbb{R}$. Since

$$\mathbb{P}^{z_s|\hat{v}_{\eta(s)}=y} = \mathcal{N}\left(y + \kappa(\theta - y)(s - \eta(s)), \sigma^2 y(s - \eta(s))\right)$$

we have for any $y > 0$ that

$$\begin{aligned}&\int_{\mathbb{R}} \phi(x)\mathbb{E}\left[L^x_t(z) - L^x_{t_k}(z)\,|\hat{v}_{t_k} = y\right]dx = \sigma^2\int_{t_k}^t y\mathbb{E}\left[\phi(z_s)|\hat{v}_{t_k} = y\right]ds \\ &= \sigma\int_{\mathbb{R}}\phi(x)\int_{t_k}^t \frac{\sqrt{y}}{\sqrt{2\pi(s-t_k)}}\exp\left(-\frac{(x - y - \kappa(\theta-y)(s-t_k))^2}{2\sigma^2 y(s-t_k)}\right)dsdx.\end{aligned}$$

Since the above equation holds for any non-negative Borel-measurable function ϕ, we must have that

$$\begin{aligned}
&\mathbb{E}\left[L_t^x(z) - L_{t_k}^x(z)\,|\hat{v}_{t_k} = y\right] \\
&\quad = \sigma \int_{t_k}^t \frac{\sqrt{y}}{\sqrt{2\pi(s-t_k)}} \exp\left(-\frac{(x-y-\kappa(\theta-y)(s-t_k))^2}{2\sigma^2 y(s-t_k)}\right) ds
\end{aligned}$$

for any $x \in \mathbb{R}$. Setting $x = 0$ yields

$$\begin{aligned}
\mathbb{E}\left[L_t^0(z) - L_{t_k}^0(z)\,|\hat{v}_{t_k} = y\right] &= \sigma \int_{t_k}^t \frac{\sqrt{y}}{\sqrt{2\pi(s-t_k)}} \exp\left(-\frac{(y+\kappa(\theta-y)(s-t_k))^2}{2\sigma^2 y(s-t_k)}\right) ds \\
&\leq \sigma \int_{t_k}^t \frac{\sqrt{y}}{\sqrt{2\pi(s-t_k)}} \exp\left(-\frac{y(1-\kappa(s-t_k))^2}{2\sigma^2(s-t_k)}\right) ds.
\end{aligned}$$

Since for $\delta > 0$ there exist a $c_\delta > 0$ such that $\sqrt{b}\exp(-b) \leq c_\delta \exp\left(-\frac{b}{1+\delta}\right)$ for all $b \geq 0$, we have

$$\frac{\sqrt{y}(1-\kappa(s-t_k))}{\sqrt{2\sigma^2(s-t_k)}} \exp\left(-\frac{y(1-\kappa(s-t_k))^2}{2\sigma^2(s-t_k)}\right) \leq c_\delta \exp\left(-\frac{y(1-\kappa(s-t_k))^2}{2\sigma^2(s-t_k)(1+\delta)}\right).$$

Moreover, since $1-\kappa(s-t_k) \in [1/2, 1]$ we obtain

$$\frac{\sqrt{y}}{\sqrt{s-t_k}} \exp\left(-\frac{y(1-\kappa(s-t_k))^2}{2\sigma^2(s-t_k)}\right) \leq \sqrt{8}\sigma c_\delta \exp\left(-\frac{y(1-\kappa(s-t_k))^2}{2\sigma^2(s-t_k)(1+\delta)}\right).$$

It follows

$$\begin{aligned}
\mathbb{E}\left[L_t^0(z) - L_{t_k}^0(z)\,|\mathcal{F}_{t_k}\right] &\leq \sigma \int_{t_k}^t \frac{\sqrt{\hat{v}_{t_k}}}{\sqrt{2\pi(s-t_k)}} \exp\left(-\frac{\hat{v}_{t_k}(1-\kappa(s-t_k))^2}{2\sigma^2(s-t_k)}\right) ds \\
&\leq c_\delta \frac{2\sigma}{\sqrt{\pi}} \int_{t_k}^t \exp\left(-\frac{\hat{v}_{t_k}(1-\kappa(s-t_k))^2}{2\sigma^2(s-t_k)(1+\delta)}\right) ds \\
&\leq c_\delta \frac{2\sigma}{\sqrt{\pi}} \int_{t_k}^t \exp\left(-\frac{\hat{v}_{t_k}(1-\kappa\Delta t)^2}{2\sigma^2\Delta t(1+\delta)}\right) ds.
\end{aligned}$$

Now, the Lyapunov inequality and Proposition 4.5 yield

$$\begin{aligned}
\mathbb{E}\left[L_t^0(z) - L_{t_k}^0(z)\right] &\leq C_\delta \int_{t_k}^t \mathbb{E}\left[\exp\left(-\frac{\hat{v}_{t_k}(1-\kappa\Delta t)^2}{2\sigma^2\Delta t(1+\delta)}\right)\right] ds \\
&= C_\delta \int_{t_k}^t \mathbb{E}\left[\exp\left(-\frac{\hat{v}_{t_k}(1-\kappa\Delta t)^2}{2\sigma^2\Delta t}\right)^{\frac{1}{1+\delta}}\right] ds \\
&\leq C_\delta \int_{t_k}^t \left(\mathbb{E}\left[\exp\left(-\frac{\hat{v}_{t_k}(1-\kappa\Delta t)^2}{2\sigma^2\Delta t}\right)\right]\right)^{\frac{1}{1+\delta}} ds \\
&\leq C_\delta \Delta t \left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{1+\delta}}.
\end{aligned}$$

(ii) For the second statement note first that

$$\begin{aligned}
&\mathbb{E}\left[\left|L_t^0(z) - L_{\eta(t)}^0(z)\right|^{1+\beta}\right]^{\frac{1}{1+\beta}} \\
&\quad = \mathbb{E}\left[\left(L_t^0(z) - L_{\eta(t)}^0(z)\right)^{\frac{1}{1+\beta}}\left(L_t^0(z) - L_{\eta(t)}^0(z)\right)^{\beta+1-\frac{1}{1+\beta}}\right]^{\frac{1}{1+\beta}} \\
&\quad \leq \left(\mathbb{E}\left[L_t^0(z) - L_{\eta(t)}^0(z)\right]\right)^{\frac{1}{(1+\beta)^2}}\left(\mathbb{E}\left[\left(L_t^0(z) - L_{\eta(t)}^0(z)\right)^{\frac{(\beta+1)^2-1}{\beta}}\right]\right)^{\frac{\beta}{(1+\beta)^2}}
\end{aligned}$$

by Hölder's inequality. Now, consider first $z = z^{sym}$ and note that

$$\begin{aligned}
&\mathbb{E}\left[\left|L_t^0(z) - L_{\eta(t)}^0(z)\right|^p\right] \\
&\quad = \mathbb{E}\left[\left|\hat{v}_t - \hat{v}_{\eta(t)} - \int_{\eta(t)}^t sign(z_s)\,\kappa\left(\theta - \hat{v}_{\eta(s)}\right) ds - \sigma\int_{\eta(t)}^t sign(z_s)\sqrt{\hat{v}_{\eta(s)}}dW_s\right|^p\right] \\
&\quad \leq 3^{p-1}\Bigg(\mathbb{E}\left|\hat{v}_t - \hat{v}_{\eta(t)}\right|^p + \mathbb{E}\left[\left|\int_{\eta(t)}^t sign(z_s)\,\kappa\left(\theta - \hat{v}_{\eta(s)}\right) ds\right|^p\right] \\
&\qquad\qquad + \mathbb{E}\left[\left|\sigma\int_{\eta(t)}^t sign(z_s)\sqrt{\hat{v}_{\eta(s)}}dW_s\right|^p\right]\Bigg)
\end{aligned}$$

for since $|x+y+z|^p \leq 3^{p-1}(|x|^p+|y|^p+|z|^p)$ for $x, y, z \in \mathbb{R}$, $p \geq 1$. We can conclude from Lemma 4.2, the Hölder inequality and the Burkholder-Davis-Gundy inequality that

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left|L_t^0(z) - L_{\eta(t)}^0(z)\right|^p\right] < \infty.$$

The case $z = z^{abs}$ can be done analogously. Applying the estimate from the first part, we obtain

$$\mathbb{E}\left[\left|L_t^0(z)-L_{\eta(t)}^0(z)\right|^{1+\beta}\right]^{\frac{1}{1+\beta}} \leq C_\beta\left(\mathbb{E}\left[L_t^0(z)-L_{\eta(t)}^0(z)\right]\right)^{\frac{1}{(1+\beta)^2}}$$
$$\leq C_{\beta,\delta}\,(\Delta t)^{\frac{1}{(1+\beta)^2}}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{(1+\delta)(1+\beta)^2}}.$$

□

The following lemma gives a control of the non-martingale terms, which arise additionally in the expansion of SE and AE.

Lemma 4.7. *Let $\Delta t \leq \frac{1}{2\kappa}$, $\varepsilon \in (0,1/2]$, $\beta > 0$ and $\star \in \{sym, abs\}$. Moreover, let $g:\mathbb{R}^2 \to \mathbb{R}$ be bounded and $h:\mathbb{R}\to\mathbb{R}$ be of linear growth. Then we have*

$$\sup_{t\in[0,T]}\mathbb{E}\left[\left|\int_0^t g(V_u,\hat{v}_u^\star)h(\hat{v}_{\eta(u)}^\star)\mathbb{1}_{\{z_u^\star\leq 0\}}du\right|\right] \leq C_{g,h,\beta}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{1+\beta}}$$

and

$$\sup_{t\in[0,T]}\mathbb{E}\left[\left|\int_0^t g(V_u,\hat{v}_u^\star)dL_u^0(z^\star)\right|\right] \leq C_{g,\beta}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{1+\beta}}.$$

Proof. For the first assertion note that

$$\sup_{t\in[0,T]}\mathbb{E}\left[\left|\int_0^t g(V_u,\hat{v}_u^\star)h(\hat{v}_{\eta(u)}^\star)\mathbb{1}_{\{z_u^\star\leq 0\}}du\right|\right]$$
$$\leq C_h\|g\|_\infty\int_0^T\mathbb{E}\left[\left(1+\sup_{t\in[0,T]}|\hat{v}_t^\star|\right)\mathbb{1}_{\{z_u^\star\leq 0\}}\right]du$$

with $\|g\|_\infty = \sup_{x,y\in\mathbb{R}}|g(x,y)|$. An application of Hölder's inequality together with Lemma 4.2 yields

$$\sup_{t\in[0,T]}\mathbb{E}\left[\left|\int_0^t g(V_u,\hat{v}_u^\star)h(\hat{v}_{\eta(u)}^\star)\mathbb{1}_{\{z_u^\star\leq 0\}}du\right|\right] \leq C_{h,g,\beta}\int_0^T(P(z_u^\star\leq 0))^{\frac{1}{1+\beta}}du$$

for all $\beta > 0$. Proposition 4.5 implies now that

$$\sup_{t\in[0,T]}\mathbb{E}\left[\left|\int_0^t g(V_u,\hat{v}_u^\star)h(\hat{v}_{\eta(u)}^\star)\mathbb{1}_{\{z_u^\star\leq 0\}}du\right|\right] \leq C_{h,g,\beta}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{1+\beta}}.$$

For the second assertion, we note that the integral under consideration is a pathwise Riemann-Stieltjes integral, since $L^0(z^\star)$ is positive and non-decreasing with $L_0^0(z^\star)=0$. We then have

$$-\|g\|_\infty L_T^0(z^\star) \leq \int_0^t g(V_u,\hat{v}_u^\star)dL_u^0(z^\star) \leq \|g\|_\infty L_T^0(z^\star), \qquad t\in[0,T].$$

It follows

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left|\int_0^t g(V_u,\hat{v}_u^\star)dL_u^0(z^\star)\right|\right] \leq \|g\|_\infty \sum_{k=0}^{N-1} \mathbb{E}\left[L^0_{t_{k+1}}(z^\star) - L^0_{t_k}(z^\star)\right]$$

and Proposition 4.6 gives

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left|\int_0^t g(V_u,\hat{v}_u^\star)dL_u^0(z^\star)\right|\right] \leq C_{g,\beta}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{1+\beta}}$$

which finishes the proof. □

4.3 The Euler Scheme for the Log-Price Process

The time-continuous extension $\hat{x} = (\hat{x}_t)_{t\in[0,T]}$ of the Euler scheme for the log-price process in the Heston model is given by

$$\begin{aligned}\hat{x}_t = \hat{x}_{\eta(t)} + \left(r - \frac{1}{2}\hat{v}_{\eta(t)}\right)(t-\eta(t)) + \rho\sqrt{\hat{v}_{\eta(t)}}\left(W_t - W_{\eta(t)}\right) \\ + \sqrt{1-\rho^2}\sqrt{\hat{v}_{\eta(t)}}\left(B_t - B_{\eta(t)}\right).\end{aligned} \tag{4.11}$$

For $(\hat{v}_t)_{t\in[0,T]}$, we can choose one of the previously introduced schemes for the CIR process. We have the same results concerning the moment stability and the local smoothness as before.

Lemma 4.8. *Let $p \geq 1$. For the Euler scheme* (4.11) *together with the scheme* (4.3) *or* (4.6)*, there exists $C_p > 0$ such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |\hat{x}_t|^p\right] \leq C_p.$$

and

$$\sup_{0\leq s<t\leq T} \mathbb{E}\left[\frac{|\hat{x}_t - \hat{x}_s|^p}{|t-s|^{p/2}}\right] < \infty.$$

Proof. Using the bounded moment results from Lemma 4.1 and Lemma 4.2, both statements follow again by standard computations. For the first term, we use the Hölder and the BDG inequality. Let $p \geq 2$. Again, the case $p \in [1,2)$ for both terms follows by the

Lyapunov inequality. Then,

$$
\begin{aligned}
&\mathbb{E}\left[\sup_{t\in[0,T]}|\hat{x}_t|^p\right]\\
&\quad= \mathbb{E}\left[\sup_{t\in[0,T]}\left|x_0 + rt - \frac{1}{2}\int_0^t \hat{v}_{\eta(s)}ds + \rho\int_0^t \sqrt{\hat{v}_{\eta(s)}}dW_s + \sqrt{1-\rho^2}\int_0^t \sqrt{\hat{v}_{\eta(s)}}dB_s\right|^p\right]\\
&\quad\le C_p\left(1 + x_0^p + \mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \hat{v}_{\eta(s)}ds\right|^p\right]\right.\\
&\qquad\qquad \left.+\mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \sqrt{\hat{v}_{\eta(s)}}dW_s\right|^p\right] + \mathbb{E}\left[\sup_{t\in[0,T]}\left|\int_0^t \sqrt{\hat{v}_{\eta(s)}}dB_s\right|^p\right]\right)\\
&\quad\le C_p\left(1 + x_0^p + T^{p-1}\int_0^T \mathbb{E}\left[\left|\hat{v}_{\eta(s)}\right|^p\right]ds + T^{p/2-1}\int_0^T \mathbb{E}\left[\left|\hat{v}_{\eta(s)}\right|^{p/2}\right]ds\right)\\
&\quad\le C_p.
\end{aligned}
$$

Furthermore, we have that

$$
\begin{aligned}
&\mathbb{E}\left[|\hat{x}_t - \hat{x}_s|^p\right]\\
&\le \mathbb{E}\left[\left|r(t-s) - \frac{1}{2}\int_s^t \hat{v}_{\eta(u)}du + \rho\int_s^t \sqrt{\hat{v}_{\eta(u)}}dW_u + \sqrt{1-\rho^2}\int_s^t \sqrt{\hat{v}_{\eta(u)}}dB_u\right|^p\right]\\
&\le C_p\left(|t-s|^p + \mathbb{E}\left[\left|\int_s^t \hat{v}_{\eta(u)}du\right|^p\right] + \mathbb{E}\left[\left|\int_s^t \sqrt{\hat{v}_{\eta(u)}}dW_u\right|^p\right] + \mathbb{E}\left[\left|\int_s^t \sqrt{\hat{v}_{\eta(u)}}dB_u\right|^p\right]\right)\\
&\le C_p\left(|t-s|^p + |t-s|^{p-1}\int_s^t \mathbb{E}\left[\left|\hat{v}_{\eta(u)}\right|^p\right]du + |t-s|^{p/2-1}\int_s^t \mathbb{E}\left[\left|\hat{v}_{\eta(u)}\right|^{p/2}\right]du\right)\\
&\le C_p\left(|t-s|^p + |t-s|^{p/2}\right),
\end{aligned}
$$

from which the second statement follows. □

Chapter 5

Regularity Results for the Kolmogorov backward PDE

There is a rich connection between partial differential equations (PDEs) and SDEs which was, amongst others, studied by Kolmogorov and Feller. It is a now classical technique to use this connection to study the weak error of numerical approximations which was introduced in Section 3.4. Solutions of elliptic and parabolic PDEs can be represented as expectations of stochastic functionals. One of the most famous results for this connection is the Feynman-Kac theorem (see e.g. Theorem 5.7.6 in [49]). In the case of the Heston model, classical results do not apply. Therefore, we will present a result by Briani et al. [13] which links the solution of the log-Heston SDE with the solution of a degenerate parabolic PDE.

First, we present a theorem from [66] which establishes the connection between PDEs and SDEs under standard textbook assumptions. We assume that we have a stochastic process $Y = (Y_t)_{t\geq 0}$ with state space $[0,T] \times \mathbb{R}^d$ and Lipschitz continuous drift and diffusion coefficients $b : \mathbb{R}^d \to \mathbb{R}^d$, $\sigma : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ which is the unique strong solution of

$$dY_t = b(Y_t)dt + \sigma(Y_t)dW_t, \quad Y_0 = y \in \mathbb{R}^d.$$

Again, $(W_t)_{t\in[0,T]}$ is an m-dimensional Brownian motion in this scenario. Moreover, we denote by $Y_t^{s,x}$ the solution at time $t > 0$ which starts in x at time $s \leq t$. The infinitesimal generator $\mathcal{L}$ of Y is defined by

$$(\mathcal{L}f)(y) = \lim_{t\downarrow 0} \frac{\mathbb{E}\left[f\left(Y_t^{0,y}\right)\right] - f(y)}{t}.$$

Here, we denote by $\mathcal{D}_{\mathcal{L}}$ the set of functions $f : \mathbb{R}^d \to \mathbb{R}$ for which the above limit exists for all $y \in \mathbb{R}^d$. If we have $f \in C_c^2\left(\mathbb{R}^d\right)$ then $f \in \mathcal{D}_{\mathcal{L}}$ and the generator has the form

$$(\mathcal{L}f)(y) = \sum_{i=1}^d b_i(y)\frac{\partial f}{\partial y_i} + \frac{1}{2}\sum_{i,j=1}^d \left(\sigma\sigma^T\right)_{i,j}(y)\frac{\partial^2 f}{\partial y_i \partial y_j}, \tag{5.1}$$

see e.g. Theorem 7.3.3 in [66]. The right side of (5.1) is called the *second order differential operator associated with the drift vector b and the diffusion matrix σ.* The next result is Theorem 8.1.1 from [66] which establishes the connection between SDE and PDE solutions.

Theorem 5.1. *Let Y be as defined above with infinitesimal generator $\mathcal{L}$ and let $f \in C_c^2(\mathbb{R}^d)$.*

(i) Define

$$u(t,y) = \mathbb{E}\left[f\left(Y_t^{0,y}\right)\right]. \tag{5.2}$$

Then, $u(t,\cdot) \in \mathcal{D}_{\mathcal{L}}$ for each $t \in [0,T]$ and

$$\begin{aligned} u_t - \mathcal{L}u &= 0, \qquad t \in (0,T], y \in \mathbb{R}^d \\ u(0,y) &= f(y), \qquad y \in \mathbb{R}^d. \end{aligned} \tag{5.3}$$

where $\mathcal{L}$ is applied to the function $y \to u(t,y)$. Equation (5.3) is called Kolmogorov backward equation.

(ii) Moreover, if $w(t,x) \in C^{1,2}([0,T] \times \mathbb{R}^d)$ is a bounded function satisfying (5.3) then $w(t,y) = u(t,y)$ given in (5.2).

Remark 5.2. *In financial applications f is often considered as a (discounted) payoff function which is applied at the final time point T. It is therefore useful to perform a time-shift from t to $T-t$. Because of the Markov property of our solution Y, Equation (5.2) then changes to*

$$v(t,y) := u(T-t,y) = \mathbb{E}\left[f\left(Y_T^{t,y}\right)\right]$$

and v satisfies

$$\begin{aligned} v_t + \mathcal{L}v &= 0, \qquad t > 0, y \in \mathbb{R}^d \\ v(T,y) &= f(y), \qquad y \in \mathbb{R}^d. \end{aligned}$$

For our main proof in Chapter 6 we need a similar result as Theorem 5.1 (i) for the Heston model. This was given by Briani et al. in [13].

Proposition 5.3 (Briani, Caramellino, Terenzi (2021))**.** *Let $q \in \mathbb{N}$ and suppose that $\partial_x^{2j} f \in C_{pol}^{q-j}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ for every $j = 0, 1, ..., q$. Set*

$$u(t,x,v) = \mathbb{E}\left[f\left(X_T^{t,x,v}, V_T^{t,v}\right)\right].$$

Then, $u \in C^q_{pol,T}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$*. Moreover, the following stochastic representation holds: For* $m + 2n \leq 2q$

$$\begin{aligned}
&\partial_x^m \partial_v^n u(t,x,v) \\
&\quad = \mathbb{E}\left[e^{-n\kappa(T-t)} \partial_x^m \partial_v^n f\left(X_T^{n,t,x,v}, V_T^{n,t,v} \right) \right] \\
&\quad\quad + n\mathbb{E}\left[\int_t^T e^{-n\kappa(T-s)} \left[\frac{1}{2}\partial_x^{m+2}\partial_v^{n-1} u + \frac{1}{2}\partial_x^{m+1}\partial_v^{n-1} u \right] \left(s, X_s^{n,t,x,v}, V_s^{n,t,v} \right) ds \right]
\end{aligned}$$

where $\partial_x^m \partial_v^{n-1} u = 0$ *when* $n = 0$ *and* $\left(X^{n,t,x,v}, V^{n,t,v} \right)$*,* $n \geq 0$*, denotes the solution to the log-Heston SDE starting in* (x, v) *at time* t *with parameters:*

$$\rho_n = \rho \qquad r_n = r + n\rho\sigma \qquad \kappa_n = \kappa \qquad \theta_n = \theta + \frac{n\sigma^2}{2\kappa} \qquad \sigma_n = \sigma.$$

In particular, if $q \geq 2$ *then* $u \in C^{1,2}_{pol}([0,T] \times \mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ *solves the PDE*

$$\begin{cases} \partial_t u(t,x,v) + (\mathcal{A}u)(t,x,v) = 0 & (t,x,v) \in [0,T) \times \mathbb{R} \times \mathbb{R}_+ \\ u(T,x,v) = f(x,v) & (x,v) \in \mathbb{R} \times \mathbb{R}_+ \end{cases} \tag{5.4}$$

where $\mathcal{A}$ *is the second order differential operator associated with the log-Heston SDE, i.e.*

$$\begin{aligned}
(\mathcal{A}u)(t,x,v) = \; & -\frac{v}{2} u_x(t,x,v) + \kappa(\theta - v) u_v(t,x,v) \\
& + \frac{v}{2}\left(u_{xx}(t,x,v) + 2\rho\sigma u_{xv}(t,x,v) + \sigma^2 u_{vv}(t,x,v) \right).
\end{aligned}$$

Remark 5.4. *Briani et al. prove this proposition for functions* f *that fulfill* $\partial_x^{2j} f \in C^{p,q-j}_{pol}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ *which means that they are additionally in* L^p*. They show that then* $u \in C^{p,q}_{pol,T}(\mathbb{R} \times \mathbb{R}_+, \mathbb{R})$ *holds. As they stated in Remark 5.4 of [13], the Proposition also holds if one drops the* L^p*-property.*

Remark 5.5. *Proposition 5.3 tells us that for the weak error analysis, we need test functions (e.g. payoff functions)* $f \in C^{2q}_{pol}$ *with* $q \geq 2$ *to get a solution* u *of the Kolmogorov backward equation that is q-times differentiable and polynomially bounded.*

For our weak error analysis in Chapter 6, we need an additional result for

$$u^\gamma(t,x,v) := u(t,x,v+\gamma)$$

where $\gamma \in [0,1]$. A direct application of Proposition 5.3 gives:

Lemma 5.6. *Let* $f \in C^6_{pol}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$*. Then, there exist* $C_f > 0$ *and* $a > 0$ *such that*

$$\sup_{\gamma \in [0,1]} \sup_{t \in [0,T)} |\partial_x^l \partial_v^m u^\gamma(t,x,v)| \leq C_f(1 + |x|^a + |1+v|^a), \quad x \in \mathbb{R}, v \geq 0,$$

if $l + 2m \leq 6$*.*

Moreover, note that the function u^γ satisfies by construction the PDE

$$\begin{aligned} u_t^\gamma(t,x,v) + (\mathcal{A}u^\gamma)(t,x,v) &= \frac{\gamma}{2}(\mathcal{R}u^\gamma)(t,x,v), \\ u^\gamma(T,x,v) &= f(x,v+\gamma), \end{aligned}$$

where

$$\begin{aligned} (\mathcal{R}u^\gamma)(t,x,v) = u_x^\gamma(t,x,v) &+ 2\kappa u_v^\gamma(t,x,v) - u_{xx}^\gamma(t,x,v) \\ &- 2\rho\sigma u_{xv}^\gamma(t,x,v) - \sigma^2 u_{vv}^\gamma(t,x,v). \end{aligned}$$

Lemma 5.6 then yields

$$\sup_{\gamma\in[0,1]}\sup_{t\in[0,T)} |(\mathcal{R}u^\gamma)(t,x,v)| \le C_f(1+|x|^a+|1+v|^a), \quad x\in\mathbb{R}, v\ge 0, \tag{5.5}$$

$$\sup_{\gamma\in[0,1]}\sup_{t\in[0,T)} \left|\frac{\partial}{\partial x}(\mathcal{R}u^\gamma)(t,x,v)\right| \le C_f(1+|x|^a+|1+v|^a), \quad x\in\mathbb{R}, v\ge 0, \tag{5.6}$$

$$\sup_{\gamma\in[0,1]}\sup_{t\in[0,T)} \left|\frac{\partial}{\partial v}(\mathcal{R}u^\gamma)(t,x,v)\right| \le C_f(1+|x|^a+|1+v|^a), \quad x\in\mathbb{R}, v\ge 0. \tag{5.7}$$

for $f \in C^6_{pol}(\mathbb{R}\times\mathbb{R}_+;\mathbb{R})$.

Chapter 6

Weak Convergence

As we have seen before, the knowledge of the weak error

$$e_{weak}(N) = |\mathbb{E}\left[f(x_N, v_N) - f(X_T, V_T)\right]|$$

plays an important role for Monte Carlo simulations. Despite the apparent simplicity of the Heston model, only a few results concerning the weak convergence order of its discretizations have been proven. Since the square root function which appears in both of the SDEs is non-Lipschitz, standard results cannot be applied. Additionally, all of the proposed time-discrete schemes for the CIR process require "fixes" to preserve the positivity of the scheme either for the full parameter regime (e.g. explicit Euler schemes) or when the Feller index is low (e.g. drift-implicit Milstein and drift-implicit Euler). This makes their analysis even more challenging.

The first proof for a full discretization of the Heston model can be found in [5] where the implicit Milstein scheme (3.10) for the CIR process and the Euler scheme (3.9) for the log-price process were analyzed. Under the assumption that $\nu > 2$, Altmayer and Neuenkirch prove weak convergence order 1 for functions f which are twice continuously differentiable with compact support and which have a Hölder-continuous second derivative of order $\varepsilon > 0$. The assumptions on the function f arise from using the results from [26] where the regularity of the solution of certain degenerate parabolic PDEs was studied. The article [74] analyzes a semi-exact scheme where the CIR process is simulated exactly from the non-central chi-squared distribution and the log price process is discretized with the trapezoidal scheme from Equation (3.7). Here, a weak convergence order of 2 is proven for polynomials for the whole parameter range, i.e. $\nu > 0$.
The (positivity preserving) weak approximation of the CIR process has been studied by Alfonsi in [1, 2]. In particular, weak first and second order schemes have been derived in these references. The article [10] studies the weak error of the symmetrized Euler scheme for the CIR process.

First, we present some of our own results concerning the weak convergence of semi-exact discretization schemes. In the main part of this chapter, we analyze the weak

convergence rate of two Euler type discretization schemes, the symmetrized Euler (SE) and the Euler with absorption fix (AE) which were presented in Chapter 3 and analyzed in Chapter 4. For these two schemes, we prove a weak convergence order of 1 for $\nu > 1$ and a weak convergence of order $\nu - \varepsilon$ for arbitrarily small $\varepsilon > 0$ for the case $\nu \leq 1$. These results have been published in [56]. Then, we extend the findings from [5] for the implicit Milstein scheme using the results from Chapter 5. Finally, we give an overview of all schemes and their convergence rates that can be proven with the presented techniques and results from this chapter. In our analysis, we observe the usual trade-off between the smoothness assumption on f and the restrictions on the Feller index ν.
Recall for the following results that our discretization grid is

$$0 = t_0 < t_1 < \ldots < t_N = T$$

and that we denote $n(t) := \max\{k \in \{0, ..., N\} : t_k \leq t\}$ and $\eta(t) := t_{n(t)}$.

6.1 Semi-Exact Discretization Schemes

Inspired by the two results which we presented in the introduction of this chapter, we analyzed a semi-exact discretization scheme in [54]. We assumed that the CIR process can be simulated exactly and studied the Euler and semi-trapezoidal discretization for the log-price process from Equations (3.5) and (3.6). Using the results from [26], we could prove a weak convergence order of 1 for both schemes.

Theorem 6.1. *Let* $\varepsilon > 0$ *and*

$$\Delta t = \max_{k \in \{0,...,N-1\}} |t_{k+1} - t_k| .$$

Let the variance process be simulated exactly, i.e. $v_k = V_{t_k}$ *for* $k \in \{0, \ldots, N\}$*, and the log-price process be discretized as in* (3.5) *or* (3.6)*.*
(i) If $f \in C_c^{2+\varepsilon}(\mathbb{R} \times [0, \infty); \mathbb{R})$ *and* $\nu > \frac{3}{2}$*, then both schemes satisfy*

$$\limsup_{N \to \infty} N \left| \mathbb{E}\left[f(x_N, v_N)\right] - \mathbb{E}\left[f(X_T, V_T)\right] \right| < \infty.$$

(ii) If $f \in C_c^{4+\varepsilon}(\mathbb{R} \times [0, \infty); \mathbb{R})$*, then both schemes satisfy*

$$\limsup_{N \to \infty} N \left| \mathbb{E}\left[f(x_N, v_N)\right] - \mathbb{E}\left[f(X_T, V_T)\right] \right| < \infty.$$

Note that it was possible to drop the restrictions on the Feller index in the second case. Using the results from [13] which we stated in Proposition 5.3, we could give an error expansion for both schemes.

Theorem 6.2. *Suppose that $f \in C^8_{pol}(\mathbb{R} \times [0,\infty); \mathbb{R})$. (i) Then, the Euler scheme* (3.5) *satisfies*

$$\mathbb{E}\left[f(x_N, v_N)\right] - \mathbb{E}\left[f(X_T, V_T)\right] = \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} \mathbb{E}\left[\mathcal{H}(s, t, \hat{x}_s, \hat{x}_t, V_s, V_t)\right] ds dt + O((\Delta t)^2),$$

where

$$\begin{aligned}\mathcal{H}(s, t, \hat{x}_s, \hat{x}_t, V_s, V_t) = &\left(\frac{1}{2} - \frac{\rho\kappa}{\sigma}\right)\big(\kappa(\theta - V_s)u_x(t, \hat{x}_t, V_t) + \sigma^2 V_s u_{xv}(s, \hat{x}_s, V_s)\big) \\ &- \frac{(1-\rho^2)}{2}\big(\kappa(\theta - V_s)u_{xx}(t, \hat{x}_t, V_t) + \sigma^2 V_s u_{xxv}(s, \hat{x}_s, V_s)\big)\end{aligned}$$

and

$$\hat{x}_t = \hat{x}_{\eta(t)} + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) V_{\eta(t)}(t - \eta(t)) + \sqrt{1-\rho^2}\sqrt{V_{\eta(t)}}(B_t - B_{\eta(t)}).$$

In particular, for an equidistant discretization with $t_k = kT/N$, $k = 0, \ldots, N$, we have

$$\lim_{N\to\infty} N\left(\mathbb{E}\left[f(x_N, v_N)\right] - \mathbb{E}\left[f(X_T, V_T)\right]\right) = \frac{T}{2}\int_0^T \mathbb{E}\left[\mathcal{H}(t, t, X_t, X_t, V_t, V_t)\right] dt.$$

Here, u denotes the solution of the associated Kolmogorov PDE, see Equation (5.4).
(ii) For the semi-trapezoidal scheme (3.6)*, we have*

$$\mathbb{E}\left[f(x_N, v_N)\right] - \mathbb{E}\left[f(X_T, V_T)\right] = \sum_{n=0}^{N-1} \int_{t_n}^{t_{n+1}} \int_{t_n}^{t} \mathbb{E}\left[\mathcal{H}(s, t, \hat{x}_s, \hat{x}_t, V_s, V_t)\right] ds dt + O((\Delta t)^2),$$

where

$$\mathcal{H}(s, t, \hat{x}_s, \hat{x}_t, V_s, V_t) = -\frac{(1-\rho^2)}{2}\big(\kappa(\theta - V_s)u_{xx}(t, \hat{x}_t, V_t) + \sigma^2 V_s u_{xxv}(s, \hat{x}_s, V_s)\big)$$

and

$$\hat{x}_t = \hat{x}_{\eta(t)} + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)\frac{V_t + V_{\eta(t)}}{2}(t - \eta(t)) + \sqrt{1-\rho^2}\sqrt{V_{\eta(t)}}\left(B_t - B_{\eta(t)}\right).$$

In particular, for an equidistant discretization $t_k = kT/N$, $k = 0, \ldots, N$, it holds

$$\lim_{N\to\infty} N\left(\mathbb{E}\left[f(x_N, v_N)\right] - \mathbb{E}\left[f(X_T, V_T)\right]\right) = \frac{T}{2}\int_0^T \mathbb{E}\left[\mathcal{H}(t, t, X_t, X_t, V_t, V_t)\right] dt.$$

Here, u denotes again the solution of the associated Kolmogorov PDE as in Equation (5.4).

6.2 Weak Convergence Order of two Euler-Type Discretization Schemes

Now, we turn to a full Euler discretization of the log-Heston model. In particular, we will analyze the scheme (3.8)

$$\begin{aligned} \bar{v}_{k+1} &= f_1(\bar{v}_k) + \kappa\left(\theta - f_2(\bar{v}_k)\right)(t_{k+1} - t_k) + \sigma\sqrt{f_3(\bar{v}_k)}\left(W_{t_{k+1}} - W_{t_k}\right) \\ v_{k+1} &= f_3(\bar{v}_{k+1}) \end{aligned} \tag{6.1}$$

for the choice

$$f_1 = f_2 = f_3 \in \{\mathtt{abs}, \mathtt{sym}\}. \tag{6.2}$$

which are the symmetrized Euler (SE) and the Euler with absorption fix (AE) (see Table 3.1). The price process is discretized as in (3.9), i.e. the standard Euler scheme

$$x_{k+1} = x_k - \frac{1}{2}v_k(t_{k+1} - t_k) + \sqrt{v_k}\left(\rho\left(W_{t_{k+1}} - W_{t_k}\right) + \sqrt{1-\rho^2}\left(B_{t_{k+1}} - B_{t_k}\right)\right). \tag{6.3}$$

Our analysis leads us to the following main result of this chapter:

Theorem 6.3. *Let $f \in C^6_{pol}(\mathbb{R} \times [0,\infty);\mathbb{R})$ and $(v_k, x_k)_{k\in\{0\dots N\}}$ be given by* (6.1)*,* (6.2) *and* (6.3)*. Furthermore, let the discretization grid be $t_k = k\Delta t, k \in \{0,...,N\}$ where $\Delta t = \frac{T}{N}$. Then, we have*

$$\limsup_{N\to\infty} N\, \left|\mathbb{E}\left[f(x_N, v_N)\right] - \mathbb{E}\left[f(X_T, V_T)\right]\right| < \infty$$

if $\nu > 1$ and

$$\limsup_{N\to\infty} N^{\alpha}\, \left|\mathbb{E}\left[f(x_N, v_N)\right] - \mathbb{E}\left[f(X_T, V_T)\right]\right| = 0$$

for all $\alpha \in (0,\nu)$ if $\nu \leq 1$.

Thus, for $\nu > 1$ we have weak convergence order one and for $\nu \leq 1$ we have weak convergence order $\nu - \varepsilon$ for arbitrarily small $\varepsilon > 0$.

Remark 6.4. *The decay in the weak convergence rate for $\nu \leq 1$ is due to the application of Propositions 4.5 and 4.6. However, this decay is also observed in numerical simulations of the respective Euler schemes (see Chapter 9). Interestingly, the convergence order ν also appears for the CIR process in a different context, namely for the L^1-approximation at the terminal time point (see Chapter 8).*

Remark 6.5. *Our analysis unfortunately does not carry over to Euler schemes for the choice*

$$f_1 = \mathtt{id}, \quad f_2 \in \{\mathtt{id}, \mathtt{abs}, \mathtt{sym}\}, \quad f_3 \in \{\mathtt{abs}, \mathtt{sym}\},$$

i.e. schemes that take negative values. As a consequence, the approximation of the CIR component is not bounded from below which prohibits our application of the Kolmogorov PDE and Itō's lemma.

Remark 6.6. *Bally and Talay analyze in [8] the weak error of the Euler scheme for SDEs with C_b^∞-coefficients, i.e. coefficients which are infinitely differentiable and whose derivatives of any order are bounded, that satisfy an additional non-degeneracy condition of Hörmander type (UH). They establish weak order one for the Euler scheme for test functions f that are only measurable and bounded. However, the log-Heston model does not satisfy the above assumptions and an adaptation of the approach of [8] to the log-Heston model leads to the restrictive assumption $\nu > \frac{9}{2}$ in [3].*

6.3 Proof of Theorem 6.3

All preliminary results for the proof were presented in Section 4.2 and Chapter 5. We recall the time-continuous extensions of the SE and AE, i.e.

$$\begin{aligned}\hat{v}_t^\star =& \hat{v}_{\eta(t)}^\star + \int_{\eta(t)}^t \kappa\left(\theta - \hat{v}_{\eta(s)}^\star\right) ds + \sigma \int_{\eta(t)}^t \sqrt{\hat{v}_{\eta(s)}^\star} dW_s \\ &- 2c^\star\sigma \int_{\eta(t)}^t \mathbb{1}_{\{z_s^\star \leq 0\}} \sqrt{\hat{v}_{\eta(s)}^\star} dW_s - 2c^\star \int_{\eta(t)}^t \mathbb{1}_{\{z_s^\star \leq 0\}} \kappa\left(\theta - \hat{v}_{\eta(s)}^\star\right) ds \\ &+ c^\star\left(L_t^0(z^\star) - L_{\eta(t)}^0(z^\star)\right), \qquad t \in [0,T],\end{aligned}$$

with $c^{sym} = 1$ and $c^{abs} = \frac{1}{2}$. We start with the now classical approach of Talay and Tubaro [70]: Since $\mathbb{E}\left[u(T, x_N, v_N)\right] = \mathbb{E}\left[f(x_N, v_N)\right]$ and $u(0, x_0, v_0) = u(0, x, v) = \mathbb{E}\left[f(X_T, V_T)\right]$ the weak error is a telescoping sum of local errors:

$$\left|\mathbb{E}\left[f(x_N, v_N)\right] - \mathbb{E}\left[f(X_T, V_T)\right]\right| = \left|\sum_{n=1}^N \mathbb{E}\left[u(t_n, x_n, v_n) - u(t_{n-1}, x_{n-1}, v_{n-1})\right]\right|.$$

Since $\hat{v}_t^{abs}$ can be zero with positive probability, technical difficulties with the Itō-formula for u at $v = 0$, i.e. at the boundary of the state space, arise. Therefore, we will analyze first

$$\left|\sum_{n=1}^N \mathbb{E}\left[u(t_n, x_n, v_n + \gamma) - u(t_{n-1}, x_{n-1}, v_{n-1} + \gamma)\right]\right|$$

with $\gamma > 0$ and in a second step exploit that

$$\begin{aligned}&\left|\mathbb{E}\left[f(x_N, v_N) - f(X_T, V_T)\right]\right| \\ &\quad = \limsup_{\gamma \searrow 0} \left|\sum_{n=1}^N \mathbb{E}\left[u(t_n, x_n, v_n + \gamma) - u(t_{n-1}, x_{n-1}, v_{n-1} + \gamma)\right]\right|.\end{aligned}$$

This regularization is not required for the symmetrized Euler scheme, but to present both proofs in a concise way, we use it for both schemes.

After the previous preparations, we now apply the Itō formula with $\gamma \in (0,1]$ to the summands of the telescoping sum. Using (4.6) and (4.11) we have

$$
\begin{aligned}
e_n^\gamma &:= \mathbb{E}\left[u^\gamma(t_{n+1},x_{n+1},v_{n+1}) - u^\gamma(t_n,x_n,v_n)\right] \\
&= \int_{t_n}^{t_{n+1}} \mathbb{E}\Bigg[u_t^\gamma(t,\hat{x}_t^\star,\hat{v}_t^\star) - \frac{1}{2}\hat{v}_{\eta(t)}^\star u_x^\gamma(t,\hat{x}_t^\star,\hat{v}_t^\star) + \kappa(\theta - \hat{v}_{\eta(t)}^\star)u_v^\gamma(t,\hat{x}_t^\star,\hat{v}_t^\star) \\
&\qquad + \frac{1}{2}\hat{v}_{\eta(t)}^\star u_{xx}^\gamma(t,\hat{x}_t^\star,\hat{v}_t^\star) + \rho\sigma\hat{v}_{\eta(t)}^\star u_{xv}^\gamma(t,\hat{x}_t^\star,\hat{v}_t^\star) + \frac{1}{2}\sigma^2\hat{v}_{\eta(t)}^\star u_{vv}^\gamma(t,\hat{x}_t^\star,\hat{v}_t^\star)\Bigg]\,dt \\
&\quad - 2c^\star \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\mathbb{1}_{\{z_t\le 0\}}\left(\kappa(\theta - \hat{v}_{\eta(t)}^\star)u_v^\gamma(t,\hat{x}_t^\star,\hat{v}_t^\star) + \rho\sigma\hat{v}_{\eta(t)}^\star u_{xv}^\gamma(t,\hat{x}_t^\star,\hat{v}_t^\star)\right)\right]dt \\
&\quad + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} u_v^\gamma(t,\hat{x}_t^\star,\hat{v}_t^\star)\,dL_t^0(z)\right].
\end{aligned}
$$

Note $t \mapsto L_t(z)$ is pathwise increasing and that $\int_{t_n}^{t_{n+1}} u_v^\gamma(t,\hat{x}_t^\star,\hat{v}_t^\star)dL_t^0(z)$ is a pathwise Riemann-Stieltjes integral. We again drop now the $\star$-label to simplify the notation. Since

$$u_t^\gamma(t,x,v) + (\mathcal{A}u^\gamma)(t,x,v) = \frac{\gamma}{2}(\mathcal{R}u^\gamma)(t,x,v)$$

with

$$
\begin{aligned}
(\mathcal{A}u^\gamma)(t,x,v) = &-\frac{v}{2}u_x^\gamma(t,x,v) + \kappa(\theta - v)u_v^\gamma(t,x,v) \\
&+ \frac{v}{2}\left(u_{xx}^\gamma(t,x,v) + 2\rho\sigma u_{xv}^\gamma(t,x,v) + \sigma^2 u_{vv}^\gamma(t,x,v)\right)
\end{aligned}
$$

and

$$
\begin{aligned}
(\mathcal{R}u^\gamma)(t,x,v) = u_x^\gamma(t,x,v) &+ 2\kappa u_v^\gamma(t,x,v) - u_{xx}^\gamma(t,x,v) \\
&- 2\rho\sigma u_{xv}^\gamma(t,x,v) - \sigma^2 u_{vv}^\gamma(t,x,v)
\end{aligned}
$$

we can write

$$
\begin{aligned}
e_n^\gamma &= \int_{t_n}^{t_{n+1}} \mathbb{E}\Bigg[\frac{\hat{v}_t - \hat{v}_{\eta(t)}}{2}\big(u_x^\gamma(t,\hat{x}_t,\hat{v}_t) + 2\kappa u_v^\gamma(t,\hat{x}_t,\hat{v}_t) - u_{xx}^\gamma(t,\hat{x}_t,\hat{v}_t) \\
&\qquad -2\rho\sigma u_{xv}^\gamma(t,\hat{x}_t,\hat{v}_t) - \sigma^2 u_{vv}^\gamma(t,\hat{x}_t,\hat{v}_t)\big)\Bigg]\,dt \\
&\quad + \frac{\gamma}{2}\int_{t_n}^{t_{n+1}} \mathbb{E}\left[(\mathcal{R}u^\gamma)(t,\hat{x}_t,\hat{v}_t)\right]dt \\
&\quad - 2c^\star \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\mathbb{1}_{\{z_t\le 0\}}\left(\kappa(\theta - \hat{v}_{\eta(t)})u_v^\gamma(t,\hat{x}_t,\hat{v}_t) + \rho\sigma\hat{v}_{\eta(t)}u_{xv}^\gamma(t,\hat{x}_t,\hat{v}_t)\right)\right]dt \\
&\quad + \mathbb{E}\left[\int_{t_n}^{t_{n+1}} u_v^\gamma(t,\hat{x}_t,\hat{v}_t)\,dL_t^0(z)\right] \\
&= e_n^{(1,\gamma)} + e_n^{(2,\gamma)} + e_n^{(3,\gamma)} + e_n^{(4,\gamma)}
\end{aligned}
$$

with

$$
\begin{aligned}
e_n^{(1,\gamma)} &:= \mathbb{E}\left[\int_{t_n}^{t_{n+1}} u_v^\gamma(t,\hat{x}_t,\hat{v}_t)dL_t^0(z)\right],\\
e_n^{(2,\gamma)} &:= -2c^\star \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\mathbb{1}_{\{z_t\le 0\}}\left(\kappa(\theta-\hat{v}_{\eta(t)})u_v^\gamma(t,\hat{x}_t,\hat{v}_t)+\rho\sigma\hat{v}_{\eta(t)}u_{xv}^\gamma(t,\hat{x}_t,\hat{v}_t)\right)\right]dt,\\
e_n^{(3,\gamma)} &:= \frac{1}{2}\int_{t_n}^{t_{n+1}} \mathbb{E}\left[(\hat{v}_t-\hat{v}_{\eta(t)})(\mathcal{R}u^\gamma)(t,\hat{x}_t,\hat{v}_t)\right]dt,\\
e_n^{(4,\gamma)} &:= \frac{\gamma}{2}\int_{t_n}^{t_{n+1}} \mathbb{E}\left[(\mathcal{R}u^\gamma)(t,\hat{x}_t,\hat{v}_t)\right]dt.
\end{aligned}
$$

6.3.1 The first term

Recall that $\int_{t_n}^{t_{n+1}} u_v^\gamma(t,\hat{x}_t,\hat{v}_t)dL_t^0(z)$ is a pathwise Riemann-Stieltjes integral and $L(z)$ is pathwise increasing. Therefore we have

$$
\left|\int_{t_n}^{t_{n+1}} u_v^\gamma(t,\hat{x}_t,\hat{v}_t)dL_t^0(z)\right| \le \sup_{t\in[t_n,t_{n+1}]} |u_v^\gamma(t,\hat{x}_t,\hat{v}_t)| \left(L_{t_{n+1}}(z)-L_{t_n}(z)\right).
$$

With Lemma 5.6 it follows

$$
\left|\int_{t_n}^{t_{n+1}} u_v^\gamma(t,\hat{x}_t,\hat{v}_t)dL_t^0(z)\right| \le C_f \sup_{t\in[0,T]} (1+|\hat{x}_t|^a+|1+\hat{v}_t|^a)\left(L_{t_{n+1}}(z)-L_{t_n}(z)\right).
$$

The Lemmas 4.2 and 4.8 yield the existence of a constant $C_p>0$ such that

$$
\left(\mathbb{E}\left[\left|\sup_{t\in[0,T]} (1+|\hat{x}_t|^a+|1+\hat{v}_t|^a)\right|^p\right]\right)^{1/p} \le C_p,
$$

and Hölder's inequality then gives

$$
e_n^{(1,\gamma)} \le C_{f,\beta}\left(\mathbb{E}\left[\left|L_{t_{n+1}}(z)-L_{t_n}(z)\right|^{1+\beta}\right]\right)^{\frac{1}{1+\beta}}
$$

for $\beta>0$. With Proposition 4.6, we can therefore conclude that

$$
|e_n^{(1,\gamma)}| \le C_{f,\beta,\delta}\,(\Delta t)^{\frac{1}{(1+\beta)^2}}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{(1+\delta)(1+\beta)^2}}, \tag{6.4}
$$

uniformly in $\gamma\in(0,1]$.

6.3.2 The second term

Recall that

$$
e_n^{(2,\gamma)} = -2c^\star \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\mathbb{1}_{\{z_t\le 0\}}\left(\kappa(\theta-\hat{v}_{\eta(t)})u_v^\gamma(t,\hat{x}_t,\hat{v}_t)+\rho\sigma\hat{v}_{\eta(t)}u_{xv}^\gamma(t,\hat{x}_t,\hat{v}_t)\right)\right]dt.
$$

An application of Hölder's inequality yields

$$e_n^{(2,\gamma)} \leq 2c^\star \int_{t_n}^{t_{n+1}} \left(P(z_t \leq 0)\right)^{\frac{1}{1+\beta}} \cdot \left(\mathbb{E}\left[\left(\kappa(\theta - \hat{v}_{\eta(t)}) u_v^\gamma(t, \hat{x}_t, \hat{v}_t) + \rho\sigma\hat{v}_{\eta(t)} u_{xv}^\gamma(t, \hat{x}_t, \hat{v}_t)\right)^{\frac{1+\beta}{\beta}} \right] \right)^{\frac{\beta}{1+\beta}} dt.$$

Lemma 5.6 and the Lemmas 4.2 and 4.8 give that

$$2 \left(\mathbb{E}\left[\left(\kappa(\theta - \hat{v}_{\eta(t)}) u_v(t, \hat{x}_t, \hat{v}_t) + \rho\sigma\hat{v}_{\eta(t)} u_{xv}(t, \hat{x}_t, \hat{v}_t)\right)^{\frac{1+\beta}{\beta}} \right] \right)^{\frac{\beta}{1+\beta}} \leq C_{f,\beta}$$

for $\beta > 0$. Since

$$P(z_t \leq 0) \leq c \left(\frac{\Delta t}{\varepsilon}\right)^{\nu(1-\varepsilon)}, \qquad t \in [0, T],$$

by Proposition 4.5, we end up with

$$|e_n^{(2,\gamma)}| \leq C_{f,\beta} \Delta t \left(\frac{\Delta t}{\varepsilon}\right)^{\nu \frac{1-\varepsilon}{1+\beta}}, \tag{6.5}$$

uniformly in $\gamma \in (0, 1]$.

6.3.3 The third term

Now, we consider

$$e_n^{(3,\gamma)} = \frac{1}{2} \int_{t_n}^{t_{n+1}} \mathbb{E}\left[(\hat{v}_t - \hat{v}_{\eta(t)}) (\mathcal{R}u^\gamma)(t, \hat{x}_t, \hat{v}_t) \right] dt.$$

Due to our assumptions the function $k^\gamma := \mathcal{R}u^\gamma$ belongs to $C^1_{pol,T}$. Using the expression for $\hat{v}_t$ from Equation (4.6) we have

$$\begin{aligned}
&\int_{t_n}^{t_{n+1}} \mathbb{E}\left[(\hat{v}_t - \hat{v}_{\eta(t)}) k^\gamma(t, \hat{x}_t, \hat{v}_t) \right] dt \\
&\quad = \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\left(\int_{\eta(t)}^t \kappa(\theta - \hat{v}_{\eta(s)}) ds + \sigma \int_{\eta(t)}^t \sqrt{\hat{v}_{\eta(s)}} dW_s \right) k^\gamma(t, \hat{x}_t, \hat{v}_t) \right] dt \\
&\quad + \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\left(-2c^\star \sigma \int_{\eta(t)}^t \mathbb{1}_{\{z_s \leq 0\}} \sqrt{\hat{v}_{\eta(s)}} dW_s \right) k^\gamma(t, \hat{x}_t, \hat{v}_t) \right] dt \\
&\quad + \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\left(-2c^\star \int_{\eta(t)}^t \mathbb{1}_{\{z_s \leq 0\}} \kappa\left(\theta - \hat{v}_{\eta(s)}\right) ds \right) k^\gamma(t, \hat{x}_t, \hat{v}_t) \right] dt \\
&\quad + \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\left(L_t^0(z) - L_{\eta(t)}^0(z) \right) k^\gamma(t, \hat{x}_t, \hat{v}_t) \right] dt.
\end{aligned}$$

Looking at the first term, we have using Hölder's inequality, Equation (5.5) and the Lemmas 4.2, 4.8 that

$$\left|\mathbb{E}\left[\left(\int_{\eta(t)}^{t} \kappa(\theta - \hat{v}_{\eta(s)})ds\right) k^{\gamma}(t, \hat{x}_t, \hat{v}_t)\right]\right| \leq C_f \Delta t. \tag{6.6}$$

By an application of the law of total expectation, the Hölder and the Minkowski inequalities we have

$$\begin{aligned}
&\left|\mathbb{E}\left[\int_{\eta(t)}^{t} \sqrt{\hat{v}_{\eta(s)}} dW_s k^{\gamma}(t, \hat{x}_t, \hat{v}_t)\right]\right| \\
&= \left|\mathbb{E}\left[\int_{\eta(t)}^{t} \sqrt{\hat{v}_{\eta(s)}} dW_s \left(k^{\gamma}(t, \hat{x}_t, \hat{v}_t) - k^{\gamma}(t, \hat{x}_{\eta(t)}, \hat{v}_{\eta(t)})\right)\right]\right| \\
&\leq \mathbb{E}\left[\left|\int_{\eta(t)}^{t} \sqrt{\hat{v}_{\eta(s)}} dW_s\right|^2\right]^{\frac{1}{2}} \cdot \left(\mathbb{E}\left[\left|k^{\gamma}(t, \hat{x}_t, \hat{v}_t) - k^{\gamma}(t, \hat{x}_{\eta(t)}, \hat{v}_t)\right|^2\right]^{1/2}\right. \\
&\qquad\qquad \left. + \mathbb{E}\left[\left|k^{\gamma}(t, \hat{x}_{\eta(t)}, \hat{v}_t) - k^{\gamma}(t, \hat{x}_{\eta(t)}, \hat{v}_{\eta(t)})\right|^2\right]^{1/2}\right).
\end{aligned}$$

The mean value theorem now gives

$$k^{\gamma}(t, \hat{x}_t, \hat{v}_t) - k(t, \hat{x}_{\eta(t)}, \hat{v}_t) = \int_0^1 k_x^{\gamma}(t, \lambda \hat{x}_t + (1-\lambda)\hat{x}_{\eta(t)}, \hat{v}_t) d\lambda \, (\hat{x}_t - \hat{x}_{\eta(t)})$$

and so

$$\begin{aligned}
\mathbb{E}\left[\left|k^{\gamma}(t, \hat{x}_t, \hat{v}_t) - k(t, \hat{x}_{\eta(t)}, \hat{v}_t)\right|^2\right]^{1/2} &\leq \int_0^1 \mathbb{E}\left[\left|k_x^{\gamma}(t, \lambda \hat{x}_t + (1-\lambda)\hat{x}_{\eta(t)}, \hat{v}_t)\right|^4\right]^{1/4} d\lambda \\
&\quad \cdot \mathbb{E}\left[\left|\hat{x}_t - \hat{x}_{\eta(t)}\right|^4\right]^{1/4}.
\end{aligned}$$

Equation (5.6) and the Lemmas 4.2, 4.8 imply that

$$\sup_{t \in [0,T]} \int_0^1 \mathbb{E}\left[\left|k_x^{\gamma}(t, \lambda \hat{x}_t + (1-\lambda)\hat{x}_{\eta(t)}, \hat{v}_t)\right|^4\right]^{1/4} d\lambda \leq C_f.$$

Thus, we have again by Lemma 4.8 that

$$\mathbb{E}\left[\left|k^{\gamma}(t, \hat{x}_t, \hat{v}_t) - k^{\gamma}(t, \hat{x}_{\eta(t)}, \hat{v}_t)\right|^2\right]^{1/2} \leq C_f \Delta t^{1/2}.$$

Similarly, we obtain

$$\mathbb{E}\left[\left|k^{\gamma}(t, \hat{x}_{\eta(t)}, \hat{v}_t) - k^{\gamma}(t, \hat{x}_{\eta(t)}, \hat{v}_{\eta(t)})\right|^2\right]^{1/2} \leq C_f \Delta t^{1/2}$$

by Equation (5.7) and the Lemmas 4.2, 4.8. Since

$$\left[\mathbb{E}\left|\int_{\eta(t)}^{t}\sqrt{\hat{v}_{\eta(s)}}dW_s\right|^2\right]^{\frac{1}{2}} \leq C(\Delta t)^{1/2}$$

by Lemma 4.2 and the Itō-isometry, we end up with

$$\left|\mathbb{E}\left[\int_{\eta(t)}^{t}\sqrt{\hat{v}_{\eta(s)}}dW_s k^{\gamma}(t,\hat{x}_t,\hat{v}_t)\right]\right| \leq C_f\Delta t. \tag{6.7}$$

Similarly, we obtain

$$\left|\mathbb{E}\left[\left(-2c^{\star}\sigma\int_{\eta(t)}^{t}\mathbb{1}_{\{z_s\leq 0\}}\sqrt{\hat{v}_{\eta(s)}}dW_s - 2c^{\star}\int_{\eta(t)}^{t}\mathbb{1}_{\{z_s\leq 0\}}\kappa\left(\theta-\hat{v}_{\eta(s)}\right)ds\right)k^{\gamma}(t,\hat{x}_t,\hat{v}_t)\right]\right| \\ \leq C_f\Delta t \tag{6.8}$$

With the Hölder inequality for some $\beta > 0$, we have

$$\left|\int_{t_n}^{t_{n+1}}\mathbb{E}\left[\left(L_t^0(z)-L_{\eta(t)}^0(z)\right)k^{\gamma}(t,\hat{x}_t,\hat{v}_t)\right]dt\right| \\ \leq \int_{t_n}^{t_{n+1}}\mathbb{E}\left[\left|L_t^0(z)-L_{\eta(t)}^0(z)\right|^{1+\beta}\right]^{\frac{1}{1+\beta}}\mathbb{E}\left[|k^{\gamma}(t,\hat{x}_t,\hat{v}_t)|^{\frac{1+\beta}{\beta}}\right]^{\frac{\beta}{1+\beta}}dt.$$

As before, we can show that there exists a constant $C_{f,\beta} > 0$, such that

$$\sup_{t\in[0,T]}\mathbb{E}\left[|k^{\gamma}(t,\hat{x}_t,\hat{v}_t)|^{\frac{1+\beta}{\beta}}\right]^{\frac{\beta}{1+\beta}} \leq C_{f,\beta}.$$

Since

$$\mathbb{E}\left[\left|L_t^0(Z)-L_{\eta(t)}^0(Z)\right|^{1+\beta}\right]^{\frac{1}{1+\beta}} \leq C_{\beta,\delta}\,(\Delta t)^{\frac{1}{(1+\beta)^2}}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{(1+\delta)(1+\beta)^2}},$$

again by Proposition 4.6, we obtain that

$$\left|\int_{t_n}^{t_{n+1}}\mathbb{E}\left[\left(L_t^0(Z)-L_{\eta(t)}^0(Z)\right)k^{\gamma}(t,\hat{x}_t,\hat{v}_t)\right]dt\right| \\ \leq C_{f,\beta,\delta}\Delta t\,(\Delta t)^{\frac{1}{(1+\beta)^2}}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{(1+\delta)(1+\beta)^2}}. \tag{6.9}$$

Summarizing (6.6), (6.7), (6.8) and (6.9) we have shown that

$$|e_n^{(3,\gamma)}| \leq C_{f,\beta,\delta}\Delta t\left(\Delta t + (\Delta t)^{\frac{1}{(1+\beta)^2}}\left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{(1+\delta)(1+\beta)^2}}\right), \tag{6.10}$$

uniformly in $\gamma \in (0,1]$.

6.3.4 The fourth term

Finally, consider

$$e_n^{(4,\gamma)} = \frac{\gamma}{2}\int_{t_n}^{t_{n+1}} \mathbb{E}\left[(\mathcal{R}u^\gamma)(t,\hat{x}_t,\hat{v}_t)\right] dt.$$

Since

$$\sup_{t\in[0,T]} \mathbb{E}\left[(\mathcal{R}u^\gamma)(t,\hat{x}_t,\hat{v}_t)\right] \le C_f$$

due to Equation (5.5) and the Lemmas 4.2, 4.8, we have that

$$\frac{1}{\gamma} e_n^{(4,\gamma)} \le C_f \Delta t, \tag{6.11}$$

uniformly in $\gamma \in (0,1]$.

6.3.5 The conclusion

Recall that $\Delta t = T/N$. Adding the Estimates (6.4), (6.5),(6.10) and (6.11), we have derived that

$$\begin{aligned}|e_n^\gamma| \le{}& C_{f,\beta,\delta}\,(\Delta t)^{\frac{1}{(1+\beta)^2}} \left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{(1+\delta)(1+\beta)^2}} \\ &+ C_{f,\beta}\Delta t \left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{1+\beta}} \\ &+ C_{f,\beta,\delta}\Delta t \left(\Delta t + (\Delta t)^{\frac{1}{(1+\beta)^2}} \left(\frac{\Delta t}{\varepsilon}\right)^{\nu\frac{1-\varepsilon}{(1+\delta)(1+\beta)^2}}\right) \\ &+ C_f\gamma\Delta t.\end{aligned}$$

For any given $\epsilon \in (0,1/2)$ we now can find $\varepsilon \in (0,1/2]$, $\beta > 0$ and $\delta > 0$ such that

$$\frac{1}{(1+\beta)^2} + \nu\frac{1-\varepsilon}{(1+\delta)(1+\beta)^2} \ge 1 + \nu(1-\epsilon)$$

and

$$1 + \nu\frac{1-\varepsilon}{1+\beta} \ge 1 + \nu(1-\epsilon).$$

Consequently, we obtain

$$|e_n^\gamma| \le C_{f,\epsilon}\Delta t\left(\Delta t + (\Delta t)^{\nu(1-\epsilon)}\right) + C_f\gamma\Delta t.$$

and

$$\sum_{n=0}^{N-1} |e_n^\gamma| \le C_f\gamma + C_{f,\epsilon}\left(\Delta t + (\Delta t)^{\nu(1-\epsilon)}\right).$$

Since

$$|\mathbb{E}\left[f(x_N, v_N) - f(X_T, V_T)\right]| \leq \limsup_{\gamma \searrow 0} \sum_{n=0}^{N-1} |e_n^\gamma|$$

we have that

$$|\mathbb{E}\left[f(x_N, v_N) - f(X_T, V_T)\right]| \leq C_{f,\epsilon}\left(\Delta t + (\Delta t)^{\nu(1-\epsilon)}\right),$$

which concludes the proof.

6.4 Weak Convergence Order of a Milstein-Type Discretization

We can use the techniques of the proof of Theorem 6.3 and the results of Briani et al. [13] to give new results for the scheme from [5] which was mentioned in the introduction of this chapter. We assume that the CIR process is discretized by the implicit Milstein scheme (3.10), i.e.

$$\begin{aligned} v_{k+1} &= v_k + \kappa\left(\theta - v_{k+1}\right)(t_{k+1} - t_k) + \sigma\sqrt{v_k}\Delta_k W + \frac{\sigma^2}{4}\left((\Delta_k W)^2 - (t_{k+1} - t_k)\right) \\ &= \frac{1}{1 + \kappa(t_{k+1} - t_k)}\left(\left(\sqrt{v_k} + \frac{\sigma}{2}\Delta_k W\right)^2 + \left(\kappa\theta - \frac{\sigma^2}{4}\right)(t_{k+1} - t_k)\right) \end{aligned} \tag{6.12}$$

and the price process is discretized by the standard Euler scheme (6.3). As in [5], we define the time-continuous extension of the implicit Milstein scheme as

$$\begin{aligned} \bar{v}_t &= \hat{v}_{\eta(t)} + \int_{\eta(t)}^t \kappa\theta ds + \int_{\eta(t)}^t \left(\sigma\sqrt{\hat{v}_{\eta(t)}} + \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right)\right) dW_s \\ \hat{v}_t &= \frac{1}{1 + \kappa(t - \eta(t))}\bar{v}_t. \end{aligned} \tag{6.13}$$

Before we start, we need some preliminary results for the implicit Milstein scheme. From [5], we have the following Lemma:

Lemma 6.7. *Let $\nu > \frac{1}{2}$. The implicit Milstein scheme has bounded moments, it holds that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} \hat{v}_t^p\right] \leq C$$

for $p \geq 1$ and

$$\sup_{t\in[0,T]} \mathbb{E}\left[\hat{v}_t^{-p}\right] \leq C$$

for $0 \leq p \leq \nu - 1$.

We also need the following results.

Lemma 6.8. *Let $p \geq 1$. We have*

$$\sup_{s,t\in[0,T]} \mathbb{E}\left[\frac{|\hat{v}_t - \hat{v}_s|^p}{|t-s|^{p/2}}\right] < \infty$$

and

$$\sup_{s,t\in[0,T]} \mathbb{E}\left[\frac{|\hat{x}_t - \hat{x}_s|^p}{|t-s|^{p/2}}\right] < \infty.$$

Proof. The proof of the second statement is analogous to the proof of Lemma 4.8 except that we now use the results bounded moment result from Lemma 6.7. For the proof of the first assertion, we rewrite (6.13). For this, we denote additionally $n^+(t) := \min\{k \in \{0, ..., N\} : t_k \geq t\}$ and $\eta^+(t) := t_{n^+(t)}$. First, we have

$$\hat{v}_t = \hat{v}_{\eta(t)} + \int_{\eta(t)}^t \kappa\left(\theta - \hat{v}_t\right) ds + \int_{\eta(t)}^t \left(\sigma\sqrt{\hat{v}_{\eta(t)}} + \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right)\right) dW_s$$

and

$$\hat{v}_{t_k} = \hat{v}_0 + \int_0^{t_k} \kappa\left(\theta - \hat{v}_{\eta^+(s)}\right) ds + \int_0^{t_k} \left(\sigma\sqrt{\hat{v}_{\eta(s)}} + \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right)\right) dW_s$$

for all $k \in \{0, ..., N\}$. Combining both terms, we obtain

$$\begin{aligned}
\hat{v}_t &= \hat{v}_0 + \int_0^{\eta(t)} \kappa\left(\theta - \hat{v}_{\eta^+(s)}\right) ds + \int_{\eta(t)}^t \kappa\left(\theta - \hat{v}_t\right) ds \\
&\quad + \int_0^t \left(\sigma\sqrt{\hat{v}_{\eta(s)}} + \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right)\right) dW_s \\
&= \hat{v}_0 + \int_0^t \kappa\theta ds - \kappa \int_0^t \left(\hat{v}_{\eta^+(s)} \mathbb{1}_{\{s \leq \eta(t)\}} + \hat{v}_t \mathbb{1}_{\{s > \eta(t)\}}\right) ds \\
&\quad + \int_0^t \left(\sigma\sqrt{\hat{v}_{\eta(s)}} + \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right)\right) dW_s
\end{aligned}$$

Now, let $p \geq 2$. We have

$$\begin{aligned}
\mathbb{E}\left[|\hat{v}_t - \hat{v}_s|^p\right] &= \mathbb{E}\left[\left|\kappa\theta(t-s) - \kappa\int_s^t \left(\hat{v}_{\eta^+(u)} \mathbb{1}_{\{u \leq \eta(t)\}} + \hat{v}_t \mathbb{1}_{\{u > \eta(t)\}}\right) du \right.\right. \\
&\quad \left.\left. + \ \sigma \int_s^t \sqrt{\hat{v}_{\eta(u)}} dW_u + \int_s^t \frac{\sigma^2}{2}\left(W_u - W_{\eta(u)}\right) dW_u\right|^p\right] \\
&\leq C_p \left(|t-s|^p + |t-s|^{p-1} \int_s^t \mathbb{E}\left[\left|\hat{v}_{\eta^+(u)} \mathbb{1}_{\{u \leq \eta(t)\}} + \hat{v}_t \mathbb{1}_{\{u > \eta(t)\}}\right|^p\right] du \right. \\
&\quad \left. + \mathbb{E}\left[\left|\int_s^t \hat{v}_{\eta(u)} du\right|^{p/2}\right] + \mathbb{E}\left[\left|\int_s^t \left(W_u - W_{\eta(u)}\right)^2 du\right|^{p/2}\right]\right) \\
&\leq C_p \left(|t-s|^p + |t-s|^{p/2}\right)
\end{aligned}$$

where we used Lemma 6.7 and Hölder's inequality again. The case $p \in [1, 2)$ can be treated by using the Lyapunov inequality. □

By assuming the same regularity for f as in Theorem 6.3, we can prove a weak convergence order of one for the whole parameter range where the Milstein scheme is well-defined.

Proposition 6.9. *Let $f \in C^6_{pol}(\mathbb{R} \times [0,\infty); \mathbb{R})$, $\nu > \frac{1}{2}$ and let $(v_k, x_k)_{k \in \{0,\dots,N\}}$ be given by Equations* (6.12) *and* (6.3)*. Furthermore, we set $\Delta t = \frac{T}{N}$. Then, we have*

$$\limsup_{N\to\infty} N \, |\mathbb{E}[f(x_N, v_N)] - \mathbb{E}[f(X_T, V_T)]| < \infty.$$

Proof. We only need to make some slight changes to the proof from [5] but we present them for completeness. As before, the weak error is a telescoping sum of local errors:

$$|\mathbb{E}[f(x_N, v_N)] - \mathbb{E}[f(X_T, V_T)]| = \left| \sum_{n=0}^{N-1} \mathbb{E}[u(t_{n+1}, x_{n+1}, v_{n+1}) - u(t_n, x_n, v_n)] \right|.$$

Note that v_k for all $k \in \{0, ..., N\}$, is strictly positive since $\nu > \frac{1}{2}$. From [5], we have

$$\begin{aligned} e_n &:= \mathbb{E}[u(t_{n+1}, x_{n+1}, v_{n+1}) - u(t, x_n, v_n)] \\ &= e_n^{(1)} + e_n^{(2)} + e_n^{(3)} \end{aligned}$$

with

$$\begin{aligned} e_n^{(1)} &:= \int_{t_n}^{t_{n+1}} (t - \eta(t)) \mathbb{E}\Bigg[\frac{\kappa^2}{1 + \kappa(t - \eta(t))} (\hat{v}_t - \theta) u_v(t, \hat{x}_t, \hat{v}_t) \\ &\quad + \frac{\kappa}{2} (\theta - \hat{v}_t) (u_x(t, \hat{x}_t, \hat{v}_t) - u_{xx}(t, \hat{x}_t, \hat{v}_t)) - \frac{\rho\sigma\kappa\theta}{1 + \kappa(t - \eta(t))} u_{xv}(t, \hat{x}_t, \hat{v}_t) \\ &\quad - \frac{\sigma^2}{2(1 + \kappa(t - \eta(t)))} \left(\kappa \hat{v}_t + \frac{4\kappa\theta - \sigma^2}{4(1 + \kappa(t - \eta(t)))} \right) u_{vv}(t, \hat{x}_t, \hat{v}_t) \Bigg] dt \\ e_n^{(2)} &:= \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\sqrt{\hat{v}_{\eta(t)}} \left(W_t - W_{\eta(t)} \right) \left(\frac{\sigma}{2} u_x(t, \hat{x}_t, \hat{v}_t) - \frac{\sigma}{2} u_{xx}(t, \hat{x}_t, \hat{v}_t) \right) \right] \\ &\quad + \mathbb{E}\left[\sqrt{\hat{v}_{\eta(t)}} \left(W_t - W_{\eta(t)} \right) \left(- \frac{\rho\sigma^2}{2(1 + \kappa(t - \eta(t)))} u_{xv}(t, \hat{x}_t, \hat{v}_t) \right) \right] dt \\ e_n^{(3)} &:= \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\left(\left(W_t - W_{\eta(t)} \right)^2 - (t - \eta(t)) \right) \left(\frac{\sigma^2}{8} (u_x(t, \hat{x}_t, \hat{v}_t) - u_{xx}(t, \hat{x}_t, \hat{v}_t)) \right) \right] \\ &\quad + \mathbb{E}\Bigg[\left(\left(W_t - W_{\eta(t)} \right)^2 - (t - \eta(t)) \right) \\ &\qquad \cdot \left(- \frac{\sigma^3\rho}{4(1 + \kappa(t - \eta(t)))} u_{xv}(t, \hat{x}_t, \hat{v}_t) \right) \Bigg] dt. \end{aligned}$$

From Proposition 5.3, we know that for $f \in C^6_{pol}$ there exist $C_f > 0$ and $a > 0$ such that

$$\sup_{t\in[0,T)} |\partial_x^l \partial_v^m u(t,x,v)| \le C_f(1+|x|^a+|v|^a), \quad x \in \mathbb{R}, v>0,$$

if $l+2m \le 6$. Now with Lemma 6.7, Lemma 6.8 and the Hölder inequality, we obtain

$$\begin{aligned}\left|e_n^{(1)}\right| &\le C_f \int_{t_n}^{t_{n+1}} (t-\eta(t))\mathbb{E}\left[(1+\hat{v}_t)\left(1+|\hat{x}_t|^a+|\hat{v}_t|^a\right)\right]dt \\ &\le C_f\,(\Delta t)^2 .\end{aligned}$$

By the same arguments, we have that

$$\left|e_n^{(3)}\right| \le C_f\,(\Delta t)^2 .$$

For the second term, we first set

$$k(t,\hat{x}_t,\hat{v}_t) := \frac{\sigma}{2}u_x(t,\hat{x}_t,\hat{v}_t) - \frac{\sigma}{2}u_{xx}(t,\hat{x}_t,\hat{v}_t) - \frac{\rho\sigma^2}{2(1+\kappa(t-\eta(t)))}u_{xv}(t,\hat{x}_t,\hat{v}_t).$$

By an application of the law of total expectation, we get

$$\begin{aligned}e_n^{(2)} &= \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\sqrt{\hat{v}_{\eta(t)}}\left(W_t - W_{\eta(t)}\right)k(t,\hat{x}_t,\hat{v}_t)\right]dt \\ &= \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\sqrt{\hat{v}_{\eta(t)}}\left(W_t - W_{\eta(t)}\right)\left(k(t,\hat{x}_t,\hat{v}_t) - k(t,\hat{x}_{\eta(t)},\hat{v}_{\eta(t)})\right)\right]dt.\end{aligned}$$

And again, by Hölder's and Minkowski's inequality

$$\begin{aligned}&\mathbb{E}\left[\sqrt{\hat{v}_{\eta(t)}}\left(W_t - W_{\eta(t)}\right)\left(k(t,\hat{x}_t,\hat{v}_t) - k(t,\hat{x}_{\eta(t)},\hat{v}_{\eta(t)})\right)\right] \\ &\quad\le \mathbb{E}\left[\left|\sqrt{\hat{v}_{\eta(t)}}\left(W_t - W_{\eta(t)}\right)\right|^2\right]^{1/2} \\ &\quad\cdot\left(\mathbb{E}\left[\left|k(t,\hat{x}_t,\hat{v}_t) - k(t,\hat{x}_{\eta(t)},\hat{v}_t)\right|^2\right]^{1/2} + \mathbb{E}\left[\left|k(t,\hat{x}_{\eta(t)},\hat{v}_t) - k(t,\hat{x}_{\eta(t)},\hat{v}_{\eta(t)})\right|^2\right]^{1/2}\right).\end{aligned}$$

As in the proof of Theorem 6.3 the mean value theorem together with Lemma 6.7 and Lemma 6.8 gives

$$\begin{aligned}&\mathbb{E}\left[\left|k(t,\hat{x}_t,\hat{v}_t) - k(t,\hat{x}_{\eta(t)},\hat{v}_t)\right|^2\right]^{1/2} + \mathbb{E}\left[\left|k(t,\hat{x}_{\eta(t)},\hat{v}_t) - k(t,\hat{x}_{\eta(t)},\hat{v}_{\eta(t)})\right|^2\right]^{1/2} \\ &\quad\le C_f\,(\Delta t)^{1/2}\end{aligned}$$

and therefore,

$$\mathbb{E}\left[\sqrt{\hat{v}_{\eta(t)}}\left(W_t - W_{\eta(t)}\right)\left(k(t,\hat{x}_t,\hat{v}_t) - k(t,\hat{x}_{\eta(t)},\hat{v}_{\eta(t)})\right)\right] \le C_f \Delta t$$

by Lemma 6.7 as well as

$$\left| e_n^{(2)} \right| \leq C_f \left(\Delta t\right)^2 .$$

Summarizing, we have

$$|e_n| \leq C_f \left(\Delta t\right)^2$$

and

$$|\mathbb{E}\left[f(x_N, v_N)\right] - \mathbb{E}\left[f(X_T, V_T)\right]| = \left| \sum_{n=0}^{N-1} e_n \right| \leq C_f \Delta t$$

and the proof is done. □

6.5 An Overview of Weak Convergence Results

Since we presented many different time-discrete schemes for the log-Heston model, Table 6.1 gives an overview of weak convergence rates that were proven so far. We do not claim this table to be complete. However, we can observe a well-known characteristic of weak convergence proofs that involve the CIR process: The trade-off between the regularity of the function f and the restrictions that we impose on the Feller index. Parameter restrictions for the CIR process are usually necessary due to the need of finite negative moments (recall Proposition 2.2). In Chapter 9, we will perform numerical simulations with the schemes that were analyzed in this thesis.

Scheme	Regularity	Parameter range	Order	From
Exact + Euler	$f \in C_c^{2+\varepsilon}$	$\nu > \frac{3}{2}$	1	[54]
Exact + Euler	$f \in C_c^{4+\varepsilon}$	$\nu > 0$	1	[54]
Exact + Semi-Trap.	$f \in C_c^{2+\varepsilon}$	$\nu > \frac{3}{2}$	1	[54]
Exact + Semi-Trap.	$f \in C_c^{4+\varepsilon}$	$\nu > 0$	1	[54]
Exact + Trap.	f polynomial	$\nu > 0$	2	[74]
SE + Euler	$f \in C_{pol}^6$	$\nu > 0$	$\min\{1, \nu - \varepsilon\}$	This thesis
AE + Euler	$f \in C_{pol}^6$	$\nu > 0$	$\min\{1, \nu - \varepsilon\}$	This thesis
Impl. Milst. + Euler	f measurable, bounded	$\nu > \frac{9}{2}$	1	[3]
Impl. Milst. + Euler	$f \in C_c^{2+\varepsilon}$	$\nu > 2$	1	[5]
Impl. Milst. + Euler	$f \in C_{pol}^6$	$\nu > \frac{1}{2}$	1	This thesis

Table 6.1: Overview of weak convergence rates

Chapter 7

L^1-Approximation of the Log-Heston SDE: Upper Bounds

We are now turning to the analysis of the strong convergence of numerical schemes for the Heston model. In particular, we study the L^1-convergence of the explicit Euler schemes from Table 3.1. Explicit Euler schemes for the CIR process and for the full Heston model are popular among practitioners because they are easy to implement and computationally cheap. However, results involving a (polynomial) strong convergence rate for these Euler schemes are rare and usually come along with a strong restriction on the Feller index.
Recall the general framework for explicit Euler schemes for the CIR process

$$\begin{aligned} \bar{v}_{k+1} &= f_1(\bar{v}_k) + \kappa\left(\theta - f_2(\bar{v}_k)\right)(t_{k+1} - t_k) + \sigma\sqrt{f_3(\bar{v}_k)}\left(W_{t_{k+1}} - W_{t_k}\right) \\ v_{k+1} &= f_3(\bar{v}_{k+1}) \end{aligned} \tag{7.1}$$

where we can choose the f_i as

$$f_1 = \texttt{id}, \quad f_2 \in \{\texttt{id}, \texttt{abs}, \texttt{sym}\}, \quad f_3 \in \{\texttt{abs}, \texttt{sym}\} \tag{7.2}$$

or

$$f_1 = f_2 = f_3 \in \{\texttt{abs}, \texttt{sym}\}. \tag{7.3}$$

For the first case, we have the following time-continuous extension from Section 4.1

$$\begin{aligned} \bar{v}_t &= \bar{v}_{\eta(t)} + \int_{\eta(t)}^t \kappa(\theta - f_2(\bar{v}_{\eta(s)}))ds + \sigma\int_{\eta(t)}^t \sqrt{f_3(\bar{v}_{\eta(s)})}dW_s, \\ \hat{v}_t &= f_3\left(\bar{v}_t\right), \end{aligned} \qquad t \in [0, T]. \tag{7.4}$$

For the second case, the time continuous extensions for the SE and the AE are

$$\begin{aligned}\hat{v}_t^\star = \hat{v}_{\eta(t)}^\star &+ \int_{\eta(t)}^t \kappa\left(\theta - \hat{v}_{\eta(s)}^\star\right) ds + \sigma \int_{\eta(t)}^t \sqrt{\hat{v}_{\eta(s)}^\star} dW_s \\ &- 2c^\star \sigma \int_{\eta(t)}^t \mathbb{1}_{\{z_s^\star \leq 0\}} \sqrt{\hat{v}_{\eta(s)}^\star} dW_s - 2c^\star \int_{\eta(t)}^t \mathbb{1}_{\{z_s^\star \leq 0\}} \kappa \left(\theta - \hat{v}_{\eta(s)}^\star\right) ds \\ &+ c^\star \left(L_t^0(z^\star) - L_{\eta(t)}^0(z^\star)\right), \qquad t \in [0,T],\end{aligned} \tag{7.5}$$

with $c^{sym} = 1$ and $c^{abs} = \frac{1}{2}$. This was shown in Section 4.2. The time-continuous extension of the Euler scheme (3.9) of the log-price process is given by

$$\begin{aligned}\hat{x}_t = \hat{x}_{\eta(t)} &+ \left(r - \frac{1}{2}\hat{v}_{\eta(t)}\right)(t - \eta(t)) + \rho\sqrt{\hat{v}_{\eta(t)}}\left(W_t - W_{\eta(t)}\right) \\ &+ \sqrt{1-\rho^2}\sqrt{\hat{v}_{\eta(t)}}\left(B_t - B_{\eta(t)}\right).\end{aligned} \tag{7.6}$$

Furthermore, we also analyze the strong convergence of the implicit Milstein scheme for the CIR process. We recall the time-continuous extension of the implicit Milstein scheme which is

$$\begin{aligned}\hat{v}_t &= \hat{v}_{\eta(t)} + \int_{\eta(t)}^t \kappa(\theta - \hat{v}_t)\, ds + \int_{\eta(t)}^t \left(\sigma\sqrt{\hat{v}_{\eta(t)}} + \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right)\right) dW_s \\ &= \frac{1}{1+\kappa(t-\eta(t))}\left(\hat{v}_{\eta(t)} + \int_{\eta(t)}^t \kappa\theta ds + \int_{\eta(t)}^t \left(\sigma\sqrt{\hat{v}_{\eta(t)}} + \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right)\right) dW_s\right).\end{aligned} \tag{7.7}$$

This chapter is now organized as follows: First, we summarize existing strong approximation results for the CIR process and the log-Heston model. Then, we present some preliminary results which are needed for our proofs. In the third part, we then prove upper bounds for the L^1-approximation of the CIR process by explicit Euler methods and by the implicit Milstein method. Afterwards, we combine these schemes with the explicit Euler scheme for the price process and prove upper bounds for the L^1-error of the full Heston model. Finally, we summarize our results.
In our proofs, we assume an equidistant discretization grid with $\Delta t = \frac{T}{N}$.

7.1 Previous Results

The strong approximation of the CIR process has been intensively studied in the last years. The first works on this topic are [1, 23, 37], which prove strong convergence (without a polynomial rate) of various explicit and implicit schemes using the Yamada-Watanabe approach.

7.1.1 Drift-implicit Euler

One of the schemes of [1] is the drift-implicit Euler scheme which we presented in Section 3.3.3 in Equation (3.13). It is positivity preserving for $\nu \geq \frac{1}{2}$. This scheme turned out to be accessible to a more detailed error analysis, see [2, 24, 43, 62]. In [2], the following time-continuous extension of (3.13) for $t \in [t_k, t_{k+1}]$ was analyzed:

$$\begin{aligned}\hat{z}_t &= \frac{z_k + \frac{\sigma}{2}\left(W_t - W_{t_k}\right)}{2 + \kappa\left(t - t_k\right)} + \sqrt{\frac{\left(z_k + \frac{\sigma}{2}\left(W_t - W_{t_k}\right)\right)^2}{\left(2 + \kappa\left(t - t_k\right)\right)^2} + \frac{\left(\kappa\theta - \frac{\sigma^2}{4}\right)\left(t - t_k\right)}{2 + \kappa\left(t - t_k\right)}} \\ \hat{v}_t &= \hat{z}_t^2.\end{aligned} \tag{7.8}$$

For $\nu > 2$ and $1 \leq p < \frac{2}{3}\nu$, it was then proven that

$$\mathbb{E}\left[\sup_{t \in [0,T]} \left|V_t - \hat{v}_t\right|^p\right]^{1/p} \leq C\Delta t.$$

Another possibility is to look at the linear interpolation between y_k and y_{k+1} which is

$$\begin{aligned}\hat{z}_t &= \frac{t_{k+1} - t}{\Delta t} z_k + \frac{t - t_k}{\Delta t} z_{k+1} \\ \hat{v}_t &= \hat{z}_t^2.\end{aligned} \tag{7.9}$$

In [24], the authors show

$$\mathbb{E}\left[\sup_{t \in [0,T]} \left|V_t - \hat{v}_t\right|^p\right]^{1/p} \leq C_p \sqrt{\left|\log\left(\Delta t\right)\right|}\sqrt{\Delta t}$$

for $\nu > 1$ and $1 \leq p < \nu$. For the same scheme, we also have

$$\mathbb{E}\left[\sup_{t \in [0,T]} \left|V_t - \hat{v}_t\right|^p\right]^{1/p} \leq C_p \left(\Delta t\right)^{\frac{\min\{\nu, 1\} - \frac{1}{2}}{p} - \epsilon}$$

from [43] for $\nu > \frac{1}{2}$ and $p \geq 1$.

7.1.2 (Truncated) Milstein

A breakthrough for the (very challenging) case $\nu \leq 1$ was provided by [32] and [31]. For the truncated Milstein scheme (3.11) from Section 3.3.2 we have

$$\sup_{t \in [0,T]} \mathbb{E}\left[\left|V_t - \hat{v}_t\right|^p\right]^{1/p} \leq C_p \left(\Delta t\right)^{\frac{\min\{\frac{1}{2}, \nu\}}{p} - \epsilon}$$

for $p \geq 1$ from [32]. Here, the continuous-time extension $\hat{v}$ is a constant interpolation, i.e.

$$\hat{v}_t = \hat{v}_{t_k} \qquad t \in [t_k, t_{k+1}).$$

In particular, the truncated Milstein scheme attains L^1-convergence order $\min\{\frac{1}{2}, \nu\} - \epsilon$ for the whole parameter range. For the implicit Milstein scheme (3.10) from Section 3.3.2, we can find strong convergence results in [62]. For the linear interpolated scheme similar as in (7.9), we have

$$\mathbb{E}\left[\sup_{t\in[0,T]} |V_t - \hat{v}_t|^2\right]^{1/2} \leq C\sqrt{|\log(\Delta t)|}\sqrt{\Delta t}$$

for $\nu > 3$. Furthermore,

$$\sup_{k\in\{0,\dots,N\}} \mathbb{E}\left[|V_{t_k} - \hat{v}_{t_k}|\right] \leq C\Delta t,$$

again for $\nu > 3$.

7.1.3 Explicit Euler schemes

In contrast to this, convergence rate results for explicit Euler schemes have been rare. In [9], the authors proved

$$\mathbb{E}\left[\sup_{t\in[0,T]} |V_t - \hat{v}_t|^{2p}\right]^{\frac{1}{2p}} \leq C_p\sqrt{\Delta t}$$

for the Symmetrized Euler (SE) (7.5) under the (strong) restriction

$$\frac{\sigma^2}{8}(\nu - 1)^2 > \kappa(4p-1) \vee (2\sigma(2p-1))^2.$$

For FTE, which is (7.4) with $f_2 = f_3 = \texttt{abs}$, the L^p-convergence order $\frac{1}{2}$ for $2 \leq p < \nu - 1$ and $\nu > 3$ is shown in [21], i.e.

$$\sup_{t\in[0,T]} \mathbb{E}\left[|V_t - \hat{v}_t|^p\right]^{1/p} \leq C_p\sqrt{\Delta t}.$$

Further contributions on the strong approximation of the CIR process can be found in [11, 17, 30].

7.1.4 Full Heston model

We are not aware of any results concerning the strong approximation of the log-Heston model except [4, 51]. In [4] the drift-implicit Euler (3.13) for the CIR process is combined with the Euler discretization (3.9) of the log-Heston process and it is proven that

$$\mathbb{E}\left[|X_T - x_N|^p\right]^{1/p} \leq C_p\sqrt{\Delta t}$$

for $p < \frac{4}{3}\nu$ and $\nu > 2$. The article [51] uses a drift implicit Milstein discretization of the CIR process instead and obtains L^2-convergence for $\nu > 1$ without a rate.

7.2 Preliminaries

In this section, we present some preliminary results that are needed for our main theorems in this chapter. The following lemma gives us a bound for the expected local time in zero of a semimartingale. This is Lemma 5.1 from [22].

Lemma 7.1. *For any $\delta \in (0,1)$ and any real-valued, continuous semimartingale $Y = (Y_t)_{t\in[0,T]}$, we have*

$$\mathbb{E}\left[L_t^0(Y)\right] \leq 4\delta - 2\mathbb{E}\left[\int_0^t \left(\mathbb{1}_{\{Y_s\in(0,\delta)\}} + \mathbb{1}_{\{Y_s>\delta\}} e^{1-\frac{Y_s}{\delta}}\right) dY_s\right] + \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{Y_s>\delta\}} e^{1-\frac{Y_s}{\delta}} d\langle Y\rangle_s\right], \qquad t\in[0,T].$$

The following inequality will be helpful for all proofs in this chapter.

Lemma 7.2. *For $\lambda \in [0,1]$ and $x, y \geq 0$, we have*

$$\left|\sqrt{x} - \sqrt{y}\right| \leq x^{-\frac{1}{2}(1-\lambda)} \left|x-y\right|^{1-\frac{\lambda}{2}}.$$

Proof. For the case $x = 0$ and/or $y = 0$, the inequality holds trivially. By using the binomial expansion, the assertion follows from standard calculations.

$$\begin{aligned}
\left|\sqrt{x} - \sqrt{y}\right| = \left|\sqrt{x} - \sqrt{y}\right|^{\lambda} \left|\sqrt{x} - \sqrt{y}\right|^{1-\lambda} &= \left|\sqrt{x} - \sqrt{y}\right|^{\lambda} \frac{\left|x-y\right|^{1-\lambda}}{(\sqrt{x}+\sqrt{y})^{1-\lambda}} \\
&\leq \left(\sqrt{|x-y|}\right)^{\lambda} \frac{\left|x-y\right|^{1-\lambda}}{(\sqrt{x}+\sqrt{y})^{1-\lambda}} \\
&= \frac{\left|x-y\right|^{1-\frac{\lambda}{2}}}{(\sqrt{x}+\sqrt{y})^{1-\lambda}} \\
&\leq \frac{\left|x-y\right|^{1-\lambda}}{x^{\frac{1}{2}(1-\lambda)}}.
\end{aligned}$$

□

We also need Doob's maximal inequality, see e.g. Theorem 3.8 in Chapter I of [49].

Proposition 7.3. *Let $M = (M_t)_{t\in[0,T]}$ be a continuous martingale and $p > 1$. Then, it holds that*

$$\mathbb{E}\left[\sup_{u\in[0,t]} |M_u|^p\right] \leq \left(\frac{p}{p-1}\right)^p \mathbb{E}\left[|M_t|^p\right], \qquad t\in[0,T].$$

7.3 L^1-approximation of the CIR process

7.3.1 Euler schemes - Case I

We first look at the discretization from (7.4) under the condition $\nu > 1$.

Theorem 7.4. *Let $(\hat{v}_t)_{t\in[0,T]}$ be given by (7.4) and $\nu > 1$. Then, for all $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ such that*

$$\sup_{t\in[0,T]} \mathbb{E}\left[|V_t - \hat{v}_t|\right] \le C_\varepsilon \left(\Delta t\right)^{\frac{1}{2}-\varepsilon}.$$

Proof. Define $e = (e_t)_{t\in[0,T]}$ by $e_t = V_t - \bar{v}_t$.
(i) The Tanaka-Meyer formula, see e.g. equation 7.9 in Chapter III in [49], yields

$$\begin{aligned}
\mathbb{E}\left[|e_t|\right] &= \mathbb{E}\left[\int_0^t sign(e_u) de_u\right] + \mathbb{E}\left[L_t^0(e)\right] \\
&= \mathbb{E}\left[\int_0^t sign(e_u)\left(-\kappa\left(V_u - f_2(\bar{v}_{\eta(u)})\right)\right) du\right] \\
&\quad + \mathbb{E}\left[\int_0^t sign(e_u)\sigma\left(\sqrt{V_u} - \sqrt{f_3(\bar{v}_{\eta(u)})}\right) dW_u\right] + \mathbb{E}\left[L_t^0(e)\right].
\end{aligned}$$

We have

$$\mathbb{E}\left[\int_0^t sign(e_u)\sigma\left(\sqrt{V_u} - \sqrt{f_3(\bar{v}_{\eta(u)})}\right) dW_u\right] = 0$$

due to Proposition 2.2, Lemma 4.1 and the martingale property of the Itō integral. Looking at the first term, we have

$$\begin{aligned}
&\mathbb{E}\left[\int_0^t sign(e_u)\left(-\kappa\left(V_u - f_2(\bar{v}_{\eta(u)})\right)\right) du\right] \\
&= -\kappa\,\mathbb{E}\left[\int_0^t sign(e_u)\left(V_u - f_2(\bar{v}_u)\right) du\right] - \kappa\,\mathbb{E}\left[\int_0^t sign(e_u)\left(f_2(\bar{v}_u) - f_2(\bar{v}_{\eta(u)})\right) du\right] \\
&\le \kappa \int_0^t \mathbb{E}\left[|e_u|\right] du + \kappa \int_0^t \mathbb{E}\left[|\bar{v}_u - \bar{v}_{\eta(u)}|\right] du \\
&\le C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \mathbb{E}\left[|e_u|\right] du
\end{aligned}$$

due to Lemma 4.1 and $|x - f_2(y)| \le |x - y|$ for $x \ge 0, y \in \mathbb{R}$ as well as $|f_2(x) - f_2(y)| \le |x - y|$ for $x, y \in \mathbb{R}$. Therefore, we obtain

$$\sup_{u\in[0,t]} \mathbb{E}\left[|e_u|\right] \le C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du + \mathbb{E}\left[L_t^0(e)\right]. \tag{7.10}$$

(ii) With Lemma 7.1 we can derive a bound for the expected local time in 0 of e. Let $\delta \in (0,1)$, then

$$\begin{aligned}\mathbb{E}\left[L_t^0(e)\right] \leq 4\delta - 2\mathbb{E}\left[\int_0^t \left(\mathbb{1}_{\{e_s\in(0,\delta)\}} + \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\right) de_s\right] \\ + \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\, d\langle e\rangle_s\right].\end{aligned} \tag{7.11}$$

We define $Y_s := \mathbb{1}_{\{e_s\in(0,\delta)\}} + \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}$ and look at the second term of (7.11), i.e. at

$$\mathbb{E}\left[\int_0^t Y_s de_s\right] = -\kappa\,\mathbb{E}\left[\int_0^t Y_s \left(V_s - f_2(\bar{v}_{\eta(s)})\right) ds\right]$$

where we already used the martingale property of the Itō integral. Since $0 \leq Y_s \leq 1$, we obtain proceeding as above that

$$\begin{aligned}\left|\mathbb{E}\left[\int_0^t Y_s de_s\right]\right| &\leq \kappa \int_0^t \mathbb{E}\left[|V_s - \bar{v}_{\eta(s)}|\right] ds \\ &\leq C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \mathbb{E}\left[|e_u|\right] du.\end{aligned} \tag{7.12}$$

The third term of (7.11) can be bounded as follows with Lemma 7.2 (using $\lambda = 0$) and the properties of f_3:

$$\begin{aligned}\frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\, d\langle e\rangle_s\right] &= \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} \sigma^2 \left(\sqrt{V_s} - \sqrt{f_3(\bar{v}_{\eta(s)})}\right)^2 ds\right] \\ &\leq \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} \sigma^2 \frac{\left|V_s - f_3(\bar{v}_{\eta(s)})\right|^2}{V_s} ds\right] \\ &\leq \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} \sigma^2 \frac{\left|V_s - \bar{v}_{\eta(s)}\right|^2}{V_s} ds\right].\end{aligned}$$

With Lemma 4.1, Proposition 2.2, the Minkowski inequality and Hölder's inequality it follows

$$\begin{aligned}
\frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\, d\langle e\rangle_s\right] \leq & \frac{C}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\ \frac{|V_s-\bar{v}_s|^2}{V_s} ds\right] \\
& + \frac{C}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} \frac{\left|\bar{v}_s-\bar{v}_{\eta(s)}\right|^2}{V_s} ds\right] \\
\leq & \frac{C}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\ \frac{|e_s|^2}{V_s} ds\right] \\
& + C\frac{\Delta t}{\delta}\int_0^t \left(\mathbb{E}\left[\frac{1}{V_s^{(1+\nu)/2}}\right]\right)^{\frac{2}{1+\nu}} ds \\
\leq & \frac{C}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} \frac{|e_s|^2}{V_s} ds\right] \\
& + C\left(\frac{\Delta t}{\delta}\right).
\end{aligned} \tag{7.13}$$

Now let $\alpha \in (0,1)$. Since

$$\sup_{s\in[0,T]} \mathbb{E}\left[|e_s|^p\right] \leq C_p$$

for all $p \geq 1$ due to Proposition 2.2 and Lemma 4.1, we have that

$$\begin{aligned}
& \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} \frac{|e_s|^2}{V_s} ds\right] \\
& = \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s\in(\delta,\delta^\alpha)\}} e^{1-\frac{e_s}{\delta}} \frac{|e_s|^2}{V_s} ds\right] + \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s\geq\delta^\alpha\}} e^{1-\frac{e_s}{\delta}} \frac{|e_s|^2}{V_s} ds\right] \\
& \leq \delta^{2\alpha-1}\int_0^t \mathbb{E}\left[\frac{1}{V_s}\right] ds + C\frac{e^{-\delta^{\alpha-1}}}{\delta}\int_0^t \left(\mathbb{E}\left[\frac{1}{V_s^{(1+\nu)/2}}\right]\right)^{\frac{2}{1+\nu}} ds \\
& \leq C_\alpha \delta^{2\alpha-1}
\end{aligned} \tag{7.14}$$

by another application of Hölder's inequality and $\limsup_{\delta\to 0} \frac{e^{-\delta^{\alpha-1}}}{\delta^{2\alpha}} = 0$. Summarizing (7.11) – (7.14) we have shown that

$$\mathbb{E}\left[L_t^0(e)\right] \leq 4\delta + C(\Delta t)^{\frac{1}{2}} + \kappa\int_0^t \mathbb{E}\left[|e_u|\right] du + C_\alpha\delta^{2\alpha-1} + C\frac{\Delta t}{\delta}.$$

(iii) Setting $\delta = (\Delta t)^{1/2}$ gives

$$\begin{aligned}
\mathbb{E}\left[L_t^0(e)\right] & \leq \kappa\int_0^t \mathbb{E}\left[|e_u|\right] du + C_\alpha(\Delta t)^{\alpha-1/2} \\
& \leq \kappa\int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du + C_\alpha(\Delta t)^{\alpha-1/2}.
\end{aligned}$$

Combining this with (7.10) yields

$$\sup_{u\in[0,t]} \mathbb{E}\left[|e_u|\right] \leq C_\alpha (\Delta t)^{\alpha-1/2} + 2\kappa \int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du$$

and choosing $\alpha = 1-\varepsilon$ and an application of Gronwall's lemma gives

$$\sup_{t\in[0,T]} \mathbb{E}\left[|V_t - \bar{v}_t|\right] \leq C_\varepsilon \, (\Delta t)^{\frac{1}{2}-\varepsilon}. \tag{7.15}$$

The assertion now follows since $|x - f_3(y)| \leq |x-y|$ for $x \geq 0$, $y \in \mathbb{R}$. □

Now we study again the discretization from (7.4) but under the condition $\nu \leq 1$.

Proposition 7.5. *Let $(\hat{v}_t)_{t\in[0,T]}$ be given by (7.4) and $\nu \leq 1$. Then, for all $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ such that*

$$\sup_{t\in[0,T]} \mathbb{E}\left[|V_t - \hat{v}_t|\right] \leq C_\varepsilon \, (\Delta t)^{\frac{\nu}{2}-\varepsilon}.$$

Proof. Define again $e = (e_t)_{t\in[0,T]}$ by $e_t = V_t - \bar{v}_t$.
(i) Proceeding as in the proof of Theorem 7.4 we obtain

$$\sup_{u\in[0,t]} \mathbb{E}\left[|e_u|\right] \leq C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du + \mathbb{E}\left[L_t^0(e)\right] \tag{7.16}$$

and

$$\mathbb{E}\left[L_t^0(e)\right] \leq 4\delta + C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \mathbb{E}\left[|e_u|\right] du + \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} d\langle e\rangle_s\right] \tag{7.17}$$

with

$$\langle e\rangle_t = \sigma^2 \int_0^t \left(\sqrt{V_s} - \sqrt{f_3(\bar{v}_{\eta(s)})}\right)^2 ds.$$

(ii) For the remaining term in (7.17) we apply Lemma 7.2 with $\lambda = 1-\nu(1-\zeta)$ for $\zeta \in (0,1)$ and Proposition 2.2, Lemma 4.1, Hölder's and Minkowski's inequality to obtain

$$
\begin{aligned}
&\frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\, d\langle e\rangle_s\right] \\
&\leq \frac{C_\zeta}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\ \frac{|V_s-\bar{v}_s|^{1+\nu(1-\zeta)}}{{V_s}^{\nu(1-\zeta)}} ds\right] \\
&\quad + \frac{C_\zeta}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\ \frac{\left|\bar{v}_s-\bar{v}_{\eta(s)}\right|^{1+\nu(1-\zeta)}}{V_s^{\nu(1-\zeta)}} ds\right] \\
&\leq \frac{C_\zeta}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\ \frac{|e_s|^{1+\nu(1-\zeta)}}{V_s^{\nu(1-\zeta)}} ds\right] \\
&\quad + C_\zeta \frac{(\Delta t)^{(1+\nu(1-\zeta))/2}}{\delta}\int_0^t \left(\mathbb{E}\left[\frac{1}{V_s^{\nu(1-\zeta^2)}}\right]\right)^{\frac{1}{1+\zeta}} ds \\
&\leq \frac{C_\zeta}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\ \frac{|e_s|^{1+\nu(1-\zeta)}}{V_s^{\nu(1-\zeta)}} ds\right] \\
&\quad + C_\zeta \left(\frac{(\Delta t)^{(1+\nu(1-\zeta))/2}}{\delta}\right).
\end{aligned}
$$

Now let again $\alpha \in (0,1)$. Since

$$
\sup_{s\in[0,T]} \mathbb{E}\left[|e_s|^p\right] \leq C_p
$$

for all $p \geq 1$ due to Proposition 2.2 and Lemma 4.1, we have that

$$
\begin{aligned}
&\frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} \frac{|e_s|^{1+\nu(1-\zeta)}}{V_s^{\nu(1-\zeta)}} ds\right] \\
&= \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s\in(\delta,\delta^\alpha)\}} e^{1-\frac{e_s}{\delta}} \frac{|e_s|^{1+\nu(1-\zeta)}}{V_s^{\nu(1-\zeta)}} ds\right] \\
&\quad + \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s\geq\delta^\alpha\}} e^{1-\frac{e_s}{\delta}} \frac{|e_s|^{1+\nu(1-\zeta)}}{V_s^{\nu(1-\zeta)}} ds\right] \\
&\leq \delta^{(1+\nu(1-\zeta))\alpha-1}\int_0^t \mathbb{E}\left[\frac{1}{V_s^{\nu(1-\zeta)}}\right] ds \\
&\quad + C\frac{e^{-\delta^{\alpha-1}}}{\delta}\int_0^t \left(\mathbb{E}\left[\frac{1}{V_s^{\nu(1-\zeta^2)}}\right]\right)^{\frac{1}{1+\zeta}} ds \\
&\leq C_{\zeta,\alpha}\delta^{(1+\nu(1-\zeta))\alpha-1}
\end{aligned}
$$

by another application of Hölder's inequality and Proposition 2.2 as well as $\limsup_{\delta\to 0} \frac{e^{-\delta^{\alpha-1}}}{\delta^{2\alpha}} = 0$.

Summarizing the previous steps we have shown that

$$\begin{aligned}\mathbb{E}\left[L_t^0(e)\right] \leq 4\delta + C_\zeta \frac{(\Delta t)^{(1+\nu(1-\zeta))/2}}{\delta} + C_{\zeta,\alpha}\delta^{(1+\nu(1-\zeta))\alpha-1} \\ + \kappa \int_0^t \mathbb{E}\left[|e_u|\right] du + C(\Delta t)^{1/2}.\end{aligned}$$

Setting $\delta = (\Delta t)^{1/2}$ and $\alpha = 1-\zeta$ gives

$$\mathbb{E}\left[L_t^0(e)\right] \leq \kappa \int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du + C_\zeta\, (\Delta t)^{\nu(1-\zeta)/2} + C_\zeta (\Delta t)^{(\nu(1-\zeta)^2-\zeta)/2}.$$

Combining this with (7.16) yields

$$\sup_{u\in[0,t]} \mathbb{E}\left[|e_u|\right] \leq C_\zeta (\Delta t)^{(\nu(1-\zeta)^2-\zeta)/2} + 2\kappa \int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du.$$

Now, choosing ζ sufficiently small and an application of Gronwall's lemma gives

$$\sup_{t\in[0,T]} \mathbb{E}\left[|V_t - \bar{v}_t|\right] \leq C_\varepsilon\, (\Delta t)^{\frac{\nu}{2}-\varepsilon} \tag{7.18}$$

and the assertion follows since $|x - f_3(y)| \leq |x-y|$ for $x \geq 0$, $y \in \mathbb{R}$. □

7.3.2 Euler schemes - Case II

In this section, we analyze both schemes from (7.5) under the condition $\nu > 1$.

Theorem 7.6. *Let $(\hat{v}_t)_{t\in[0,T]}$ be given by (7.5) and $\nu > 1$. Then, for all $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ such that*

$$\sup_{t\in[0,T]} \mathbb{E}\left[|V_t - \hat{v}_t|\right] \leq C_\varepsilon\, (\Delta t)^{\frac{1}{2}-\varepsilon}.$$

Proof. The proof is very similar to the proof of Theorem 7.4. Differences are only due to the additional terms in the expansion of the schemes and we will give the required additional steps in the following. For $e_t = V_t - \hat{v}_t$ we have

$$\begin{aligned}e_t = &-\int_0^t \kappa\left(V_s - \hat{v}^\star_{\eta(s)}\right) ds + \sigma \int_0^t \left(\sqrt{V_s} - \sqrt{\hat{v}^\star_{\eta(s)}}\right) dW_s \\ &+ 2c^\star\sigma \int_0^t \mathbb{1}_{\{z_s^\star \leq 0\}} \sqrt{\hat{v}^\star_{\eta(s)}} dW_s + 2c^\star \int_0^t \mathbb{1}_{\{z_s^\star\leq 0\}} \kappa\left(\theta - \hat{v}^\star_{\eta(s)}\right) ds \\ &- c^\star L_t^0(z^\star).\end{aligned}$$

(i) The Tanaka-Meyer formula yields

$$
\begin{aligned}
\mathbb{E}\left[|e_t|\right] &= \mathbb{E}\left[\int_0^t sign(e_u) de_u\right] + \mathbb{E}\left[L_t^0(e)\right] \\
&= -\kappa\mathbb{E}\left[\int_0^t sign(e_u)(V_u - \hat{v}^\star_{\eta(u)})du\right] \\
&\quad + \mathbb{E}\left[\sigma \int_0^t sign(e_u)\left(\sqrt{V_u} - \sqrt{\hat{v}^\star_{\eta(u))}}\right) dW_u\right] + \mathbb{E}\left[L_t^0(e)\right] \\
&\quad + \mathbb{E}\left[2c^\star\sigma \int_0^t sign(e_u)\mathbb{1}_{\{z_u^\star \leq 0\}}\sqrt{\hat{v}^\star_{\eta(u)}} dW_u\right] \\
&\quad + \mathbb{E}\left[2c^\star \int_0^t sign(e_u)\mathbb{1}_{\{z_u^\star \leq 0\}}\kappa\left(\theta - \hat{v}^\star_{\eta(u)}\right) du - c^\star \int_0^t sign(e_u) dL_u^0(z^\star)\right].
\end{aligned}
$$

However, Lemma 4.2 and the martingale property of stochastic integrals imply

$$
\mathbb{E}\left[2c^\star\sigma \int_0^t sign(e_u)\mathbb{1}_{\{z_u^\star \leq 0\}}\sqrt{\hat{v}^\star_{\eta(u)}} dW_u\right] = 0
$$

and Lemma 4.7 gives

$$
\left|\mathbb{E}\left[\int_0^t sign(e_u)\left(\mathbb{1}_{\{z_u^\star \leq 0\}} 2\kappa(\theta - \hat{v}^\star_{\eta(u)})du - dL_u^0(z^\star)\right)\right]\right| \leq C_\varepsilon \left(\Delta t\right)^{\nu\frac{1-\varepsilon}{1+\varepsilon}}
$$

by choosing $\beta = \varepsilon$. Thus, we have

$$
\sup_{u\in[0,t]} \mathbb{E}\left[|e_u|\right] \leq C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du + \mathbb{E}\left[L_t^0(e)\right] \tag{7.19}
$$

as in the first step of the previous proof by choosing ε appropriately (and since $\nu > 1$).
(ii) In the same way Lemma 4.7 and Lemma 4.2 also yield

$$
\left|\mathbb{E}\left[\int_0^t Y_s de_s\right]\right| \leq C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \mathbb{E}\left[|e_u|\right] du.
$$

Finally, we have

$$
\langle e\rangle_t = \sigma^2 \int_0^t \left(\sqrt{V_s} - \sqrt{\hat{v}^\star_{\eta(s)}} + 2c^\star\mathbb{1}_{\{z_s^\star \leq 0\}}\sqrt{\hat{v}^\star_{\eta(s)}}\right)^2 ds.
$$

But again Lemma 4.7 with $\beta = \varepsilon$ gives that

$$
\begin{aligned}
\frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s > \delta\}} e^{1-\frac{e_s}{\delta}} d\langle e\rangle_s\right] \leq& \frac{2}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s > \delta\}} e^{1-\frac{e_s}{\delta}}\left(\sqrt{V}_s - \sqrt{\hat{v}_{\eta(s)}}\right)^2\right] ds \\
&+ C_\varepsilon \frac{(\Delta t)^{\nu\frac{1-\varepsilon}{1+\varepsilon}}}{\delta}
\end{aligned}
$$

and proceeding as in the previous proof we obtain that

$$\mathbb{E}\left[L_t^0(e)\right] \leq 4\delta + C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \mathbb{E}\left[|e_u|\right] du + C_\alpha \delta^{2\alpha-1} + C_\varepsilon \frac{(\Delta t)^{\nu\frac{1-\varepsilon}{1+\varepsilon}}}{\delta}.$$

(iii) Setting $\delta = (\Delta t)^{1/2}$ and using $\nu > 1$ now yields

$$\mathbb{E}\left[L_t^0(e)\right] \leq \kappa \int_0^t \mathbb{E}\left[|e_u|\right] du + C_\alpha (\Delta t)^{\alpha-1/2} + C_\varepsilon (\Delta t)^{\frac{1-\varepsilon}{1+\varepsilon}-1/2}.$$

Combining this with (7.19) yields

$$\sup_{u\in[0,t]} \mathbb{E}\left[|e_u|\right] \leq C_\alpha (\Delta t)^{\alpha-1/2} + C_\varepsilon (\Delta t)^{\frac{1-\varepsilon}{1+\varepsilon}-1/2} + 2\kappa \int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du.$$

Choosing $\alpha = 1-\varepsilon$ and observing that $\frac{1}{2} - \varepsilon \geq \frac{1-\varepsilon}{1+\varepsilon} - \frac{1}{2}$ for all $\varepsilon \in (0,1)$ an application of Gronwall's lemma give then

$$\sup_{u\in[0,T]} \mathbb{E}\left[|e_u|\right] \leq C_\varepsilon (\Delta t)^{\frac{1-\varepsilon}{1+\varepsilon}-1/2}$$

and the assertion follows by choosing ε appropriately.

□

Remark 7.7. *We are not able to establish the analogous result to Proposition 7.5 for Case II of the Euler schemes* (7.5)*, since we have in that case*

$$\frac{1}{\delta}\langle e\rangle_t = \frac{\sigma^2}{\delta} \int_0^t \left(\sqrt{V_s} - \sqrt{\hat{v}^\star_{\eta(s)}} + 2c^\star \mathbb{1}_{\{z_s^\star \leq 0\}} \sqrt{\hat{v}^\star_{\eta(s)}}\right)^2 ds$$

instead of

$$\frac{1}{\delta}\langle e\rangle_t = \frac{\sigma^2}{\delta} \int_0^t \left(\sqrt{V_s} - \sqrt{f_3(\overline{v}_{\eta(s)})}\right)^2 ds.$$

The additional term gives a contribution of order $\frac{1}{\delta}\Delta t^{\nu\frac{1-\varepsilon}{1+\varepsilon}}$*, which will lead to a worse error bound than the one given in Proposition 7.5.*

7.3.3 The Implicit Milstein scheme

Proposition 7.8. *Let* $(\hat{v}_t)_{t\in[0,T]}$ *be given by* (7.7) *and* $\nu > 1$*. Then, for all* $\varepsilon > 0$ *there exists a constant* $C_\varepsilon > 0$ *such that*

$$\sup_{t\in[0,T]} \mathbb{E}\left[|V_t - \hat{v}_t|\right] \leq C_\varepsilon \left(\Delta t\right)^{\frac{1}{2}-\varepsilon}.$$

Proof. (i) Again, we denote $n^+(t) := \min\{k \in \{0,...,N\} : t_k > t\}$ and $\eta^+(t) := t_{n^+(t)}$. We define $(e_t)_{t\in[0,T]}$ by $e_t = V_t - \hat{v}_t$. Then,

$$\begin{aligned}
e_{t_{k+1}} &= e_{t_k} + \int_{t_k}^{t_{k+1}} -\kappa\left(V_s - \hat{v}_{t_{k+1}}\right) ds + \int_{t_k}^{t_{k+1}} \sigma\left(\sqrt{V_s} - \sqrt{\hat{v}_{t_k}} - \frac{\sigma^2}{2}(W_s - W_{t_k})\right) dW_s \\
&= e_{t_k} + \int_{t_k}^{t_{k+1}} -\kappa\left(V_s - \hat{v}_s + \hat{v}_s - \hat{v}_{t_{k+1}}\right) ds + \int_{t_k}^{t_{k+1}} \sigma\left(\sqrt{V_s} - \sqrt{\hat{v}_{t_k}}\right) dW_s \\
&\quad - \int_{t_k}^{t_{k+1}} \frac{\sigma^2}{2}(W_s - W_{t_k}) dW_s \\
&= e_{t_k} - \int_{t_k}^{t_{k+1}} \kappa e_s ds + \int_{t_k}^{t_{k+1}} \kappa\left(\hat{v}_{t_{k+1}} - \hat{v}_s\right) ds + \int_{t_k}^{t_{k+1}} \sigma\left(\sqrt{V_s} - \sqrt{\hat{v}_{t_k}}\right) dW_s \\
&\quad - \int_{t_k}^{t_{k+1}} \frac{\sigma^2}{2}(W_s - W_{t_k}) dW_s
\end{aligned}$$

for every $k \in \{0, ..., N-1\}$. Summing over k, we get

$$\begin{aligned}
e_{t_k} = &- \int_0^{t_k} \kappa e_s ds + \int_0^{t_k} \kappa\left(\hat{v}_{\eta^+(s)} - \hat{v}_s\right) ds + \int_0^{t_k} \sigma\left(\sqrt{V_s} - \sqrt{\hat{v}_{\eta(s)}}\right) dW_s \\
&- \int_0^{t_k} \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right) dW_s.
\end{aligned}$$

Analogously, we have

$$\begin{aligned}
e_t = e_{t_k} &- \int_{t_k}^{t} \kappa e_s ds + \int_{t_k}^{t} \kappa\left(\hat{v}_t - \hat{v}_s\right) ds + \int_{t_k}^{t} \sigma\left(\sqrt{V_s} - \sqrt{\hat{v}_{t_k}}\right) dW_s \\
&- \int_{t_k}^{t} \frac{\sigma^2}{2}(W_s - W_{t_k}) dW_s
\end{aligned}$$

for all $t \in [t_k, t_{k+1}]$. Combining the two terms yields

$$\begin{aligned}
e_t = &- \int_0^t \kappa e_s ds + \int_0^{\eta(t)} \kappa\left(\hat{v}_{\eta^+(s)} - \hat{v}_s\right) ds + \int_{\eta(t)}^t \kappa\left(\hat{v}_t - \hat{v}_s\right) ds \\
&+ \int_0^t \sigma\left(\sqrt{V_s} - \sqrt{\hat{v}_{\eta(s)}}\right) dW_s - \int_0^t \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right) dW_s.
\end{aligned} \tag{7.20}$$

Again, the Tanaka-Mayer formula, the Martingale property, Proposition 2.2 and Lemma

6.7 yield

$$\begin{aligned}
\mathbb{E}\left[|e_t|\right] &= \mathbb{E}\left[\int_0^t sign(e_u)de_u\right] + \mathbb{E}\left[L_t^0(e)\right] \\
&= \mathbb{E}\left[-\int_0^t sign(e_s)\kappa e_s ds + \int_0^{\eta(t)} sign(e_s)\kappa\left(\hat{v}_{\eta^+(s)} - \hat{v}_s\right)ds\right] \\
&\quad + \mathbb{E}\left[\int_{\eta(t)}^t sign(e_s)\kappa\left(\hat{v}_t - \hat{v}_s\right)ds\right] - \mathbb{E}\left[\int_0^t \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right)dW_s\right] \\
&\quad + \mathbb{E}\left[L_t^0(e)\right] \\
&\leq \kappa\int_0^t \mathbb{E}\left[|e_s|\right]ds + \kappa\int_0^{\eta(t)} \mathbb{E}\left[\left|\hat{v}_{\eta^+(s)} - \hat{v}_s\right|\right]ds + \kappa\int_{\eta(t)}^t \mathbb{E}\left[|\hat{v}_t - \hat{v}_s|\right]ds \\
&\quad + \mathbb{E}\left[L_t^0(e)\right].
\end{aligned}$$

By Lemma 6.8, we then have

$$\sup_{u\in[0,t]} \mathbb{E}\left[|e_u|\right] \leq C\left(\Delta t\right)^{1/2} + \kappa\int_0^t \sup_{v\in[0,s]} \mathbb{E}\left[|e_v|\right]ds + \mathbb{E}\left[L_t^0(e)\right]. \tag{7.21}$$

(ii) Proceeding as in the proof of Theorem 7.4 we obtain

$$\mathbb{E}\left[L_t^0(e)\right] \leq 4\delta + C(\Delta t)^{\frac{1}{2}} + \kappa\int_0^t \mathbb{E}\left[|e_u|\right]du + \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}}e^{1-\frac{e_s}{\delta}}d\langle e\rangle_s\right] \tag{7.22}$$

with

$$\langle e\rangle_t = \int_0^t \left(\sigma\left(\sqrt{V_s} - \sqrt{\hat{v}_{\eta(s)}}\right) - \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right)\right)^2 ds.$$

For the remaining term in (7.22), we obtain

$$\begin{aligned}
\frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}}e^{1-\frac{e_s}{\delta}}d\langle e\rangle_s\right] &\leq \frac{C}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}}e^{1-\frac{e_s}{\delta}}\sigma^2\left(\sqrt{V_s} - \sqrt{\hat{v}_{\eta(s)}}\right)^2 ds\right] \\
&\quad + \frac{C}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}}e^{1-\frac{e_s}{\delta}}\frac{\sigma^4}{4}\left(W_s - W_{\eta(s)}\right)^2 ds\right].
\end{aligned} \tag{7.23}$$

For the first term in Equation (7.23), we can proceed as before and get with Lemma 6.8 and Proposition 2.2

$$\frac{C}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}}e^{1-\frac{e_s}{\delta}}\sigma^2\left(\sqrt{V_s} - \sqrt{\hat{v}_{\eta(s)}}\right)^2 ds\right] \leq C_\alpha\delta^{2\alpha-1} + C\left(\frac{\Delta t}{\delta}\right)$$

for an $\alpha \in (0,1)$ and for $\nu > 1$. For the second term, we obtain

$$\frac{C}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}}e^{1-\frac{e_s}{\delta}}\frac{\sigma^4}{4}\left(W_s - W_{\eta(s)}\right)^2 ds\right] \leq \frac{C}{\delta}\mathbb{E}\left[\int_0^t \left(W_s - W_{\eta(s)}\right)^2 ds\right] \leq C\left(\frac{\Delta t}{\delta}\right).$$

Therefore, combining all results as before, we have

$$\mathbb{E}\left[L_t^0(e)\right] \leq 4\delta + C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \mathbb{E}\left[|e_u|\right] du + C_\alpha \delta^{2\alpha-1} + C\left(\frac{\Delta t}{\delta}\right).$$

(iii) Setting $\delta = (\Delta t)^{1/2}$ and combining with (7.21), we finally get

$$\sup_{u\in[0,t]} \mathbb{E}\left[|e_u|\right] \leq C_\alpha (\Delta t)^{\alpha-1/2} + 2\kappa \int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du$$

and the assertion follows by choosing $\alpha = 1-\varepsilon$ and an application of Gronwall's lemma.
□

Now we study again the discretization from (7.7) but under the condition $\frac{1}{2} < \nu \leq 1$ where still no truncation is needed.

Proposition 7.9. *Let $(\hat{v}_t)_{t\in[0,T]}$ be given by (7.7) and $\frac{1}{2} < \nu \leq 1$. Then, for all $\varepsilon > 0$ there exists a constant $C_\varepsilon > 0$ such that*

$$\sup_{t\in[0,T]} \mathbb{E}\left[|V_t - \hat{v}_t|\right] \leq C_\varepsilon (\Delta t)^{\frac{\nu}{2}-\varepsilon}.$$

(i) Proceeding as in the proof of Theorem 7.8 we obtain

$$\sup_{u\in[0,t]} \mathbb{E}\left[|e_u|\right] \leq C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du + \mathbb{E}\left[L_t^0(e)\right] \tag{7.24}$$

and

$$\mathbb{E}\left[L_t^0(e)\right] \leq 4\delta + C(\Delta t)^{\frac{1}{2}} + \kappa \int_0^t \mathbb{E}\left[|e_u|\right] du + \frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} d\langle e\rangle_s\right] \tag{7.25}$$

with

$$\langle e\rangle_t = \int_0^t \left(\sigma\left(\sqrt{V_s} - \sqrt{\hat{v}_{\eta(s)}}\right) + \frac{\sigma^2}{2}\left(W_s - W_{\eta(s)}\right)\right)^2 ds.$$

(ii) Then, as in the proof of Proposition 7.5 we apply Lemma 7.2 with $\lambda = 1-\nu(1-\zeta)$ for $\zeta \in (0,1)$ in the remaining term of (7.25). With Proposition 2.2, Lemma 6.7, Lemma

6.8, Hölder's and Minkowski's inequality we obtain

$$\begin{aligned}
&\frac{1}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\, d\langle e\rangle_s\right] \\
&\leq \frac{C_\zeta}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\, \frac{|V_s-\hat{v}_s|^{1+\nu(1-\zeta)}}{{V_s}^{\nu(1-\zeta)}} ds\right] \\
&\quad + \frac{C_\zeta}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} \frac{\left|\hat{v}_s-\hat{v}_{\eta(s)}\right|^{1+\nu(1-\zeta)}}{V_s^{\nu(1-\zeta)}} ds\right] + C\left(\frac{\Delta t}{\delta}\right) \\
&\leq \frac{C_\zeta}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}}\, \frac{|e_s|^{1+\nu(1-\zeta)}}{V_s^{\nu(1-\zeta)}} ds\right] \\
&\quad + C_\zeta \frac{(\Delta t)^{(1+\nu(1-\zeta))/2}}{\delta}\int_0^t \left(\mathbb{E}\left[\frac{1}{V_s^{\nu(1-\zeta^2)}}\right]\right)^{\frac{1}{1+\zeta}} ds + C\left(\frac{\Delta t}{\delta}\right) \\
&\leq \frac{C_\zeta}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} \frac{|e_s|^{1+\nu(1-\zeta)}}{V_s^{\nu(1-\zeta)}} ds\right] \\
&\quad + C_\zeta\left(\frac{(\Delta t)^{(1+\nu(1-\zeta))/2}}{\delta}\right) + C\left(\frac{\Delta t}{\delta}\right).
\end{aligned}$$

Now let $\alpha \in (0,1)$. For the first term, we again have

$$\frac{C_\zeta}{\delta}\mathbb{E}\left[\int_0^t \mathbb{1}_{\{e_s>\delta\}} e^{1-\frac{e_s}{\delta}} \frac{|e_s|^{1+\nu(1-\zeta)}}{V_s^{\nu(1-\zeta)}} ds\right] \leq C_{\zeta,\alpha}\delta^{(1+\nu(1-\zeta))\alpha-1}$$

due to Proposition 2.2, Lemma 6.7 and Lemma 6.8 as in the proof of Proposition 7.5. Summarizing the previous steps we have shown that

$$\begin{aligned}
\mathbb{E}\left[L_t^0(e)\right] &\leq 4\delta + C_\zeta \frac{(\Delta t)^{(1+\nu(1-\zeta))/2}}{\delta} + C_{\zeta,\alpha}\delta^{(1+\nu(1-\zeta))\alpha-1} \\
&\quad + \kappa\int_0^t \mathbb{E}\left[|e_u|\right] du + C(\Delta t)^{1/2} + C\left(\frac{\Delta t}{\delta}\right).
\end{aligned}$$

Setting $\delta = (\Delta t)^{1/2}$ and $\alpha = 1-\zeta$ gives

$$\mathbb{E}\left[L_t^0(e)\right] \leq \kappa\int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du + C_\zeta\,(\Delta t)^{\nu(1-\zeta)/2} + C_\zeta(\Delta t)^{(\nu(1-\zeta)^2-\zeta)/2}.$$

Combining this with (7.24) yields

$$\sup_{u\in[0,t]} \mathbb{E}\left[|e_u|\right] \leq C_\zeta(\Delta t)^{(\nu(1-\zeta)^2-\zeta)/2} + 2\kappa\int_0^t \sup_{v\in[0,u]} \mathbb{E}\left[|e_v|\right] du$$

and the assertion follows by choosing ζ sufficiently small and an application of Gronwall's lemma.

7.4 L^1-Approximation of the Heston Model

In this section, we show in particular that the results from Theorem 7.4, Theorem 7.6, Theorem 7.8, Proposition 7.5 and Proposition 7.9 carry over to a discretization of the log-Heston model where the log-price process is additionally discretized with the Euler scheme.

7.4.1 Euler schemes - Case I

The key ingredient here and also for the second case is the observation that two continuous martingales $M = (M_t)_{t\in[0,T]}$ and $\tilde{M} = (\tilde{M}_t)_{t\in[0,T]}$, whose quadratic variation coincides, have equivalent moments. This directly follows from Proposition 1.1.

Theorem 7.10. *Let $(\hat{x}, \hat{v})$ be given by (7.6) and (7.4). Then, for all $\epsilon > 0$ there exists a constant $C_\epsilon > 0$ such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t - \hat{x}_t|\right] \leq C_\epsilon\,(\Delta t)^{\frac{\min\{1,\nu\}}{2}-\epsilon}.$$

Proof. (i) Without loss of generality, we can assume $r = 0$. We have to analyze

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t - \hat{x}_t|\right] = \mathbb{E}\left[\sup_{t\in[0,T]} \left|\frac{1}{2}\int_0^t \left(V_s - \hat{v}_{\eta(s)}\right) ds + \int_0^t \left(\sqrt{V_s} - \sqrt{\hat{v}_{\eta(s)}}\right) dU_s\right|\right]$$

with

$$U_t = \rho W_t + \sqrt{1-\rho^2} B_t.$$

Using Estimate (7.15) from Theorem 7.4, Estimate (7.18) from Proposition 7.5 and Lemma 4.1, we obtain

$$\begin{aligned}\mathbb{E}\left[\sup_{t\in[0,T]} |X_t - \hat{x}_t|\right] &\leq \frac{1}{2}\int_0^T \mathbb{E}\left[|V_s - \hat{v}_s|\right] ds + \frac{1}{2}\int_0^T \mathbb{E}\left[\left|\hat{v}_s - \hat{v}_{\eta(s)}\right|\right] ds + \mathbb{E}\left[\sup_{t\in[0,T]} |M_t|\right] \\ &\leq C_\varepsilon\,(\Delta t)^{\frac{\min\{1,\nu\}}{2}-\varepsilon} + \mathbb{E}\left[\sup_{t\in[0,T]} |M_t|\right]\end{aligned} \tag{7.26}$$

where

$$M_t = \int_0^t \left(\sqrt{V_s} - \sqrt{\hat{v}_{\eta(s)}}\right) dU_s.$$

(ii) Let

$$\tilde{M}_t = \int_0^t \left(\sqrt{V_s} - \sqrt{\hat{v}_{\eta(s)}}\right) dW_s, \qquad t \in [0,T].$$

Clearly, we have

$$\langle M\rangle_t = \langle \tilde{M}\rangle_t, \qquad t \in [0,T],$$

and so Proposition 1.1 yields

$$\mathbb{E}\left[\sup_{t\in[0,T]}|M_t|\right] \leq C_{1/2}\,\mathbb{E}\left[\langle M\rangle_T^{\frac{1}{2}}\right] = C_{1/2}\,\mathbb{E}\left[\langle \tilde{M}\rangle_T^{\frac{1}{2}}\right] \leq \frac{C_{1/2}}{c_{1/2}}\,\mathbb{E}\left[\sup_{t\in[0,T]}|\tilde{M}_t|\right].$$

Now, the Lyapunov inequality and an application of Doob's maximal inequality, i.e. Proposition 7.3, give

$$\mathbb{E}\left[\sup_{t\in[0,T]}|M_t|\right] \leq C_{1/2,\beta}\left(\mathbb{E}\left[|\tilde{M}_T|^{1+\beta}\right]\right)^{1/(1+\beta)}. \tag{7.27}$$

for $\beta > 0$. Using (7.4) and the SDE for the CIR process, we have

$$\tilde{M}_T = \frac{1}{\sigma}\left(V_T - \overline{v}_T + \kappa\int_0^T \left(V_s - f_2(\overline{v}_{\eta(s)})\right)ds\right)$$

and we obtain

$$\begin{aligned}\mathbb{E}\left[|\tilde{M}_T|\right] &\leq \frac{1}{\sigma}\left(\mathbb{E}\left[|V_T - \overline{v}_T|\right] + \kappa\int_0^T \mathbb{E}\left[|V_s - \overline{v}_s|\right]ds + \kappa\int_0^T \mathbb{E}\left[\left|\overline{v}_s - \overline{v}_{\eta(s)}\right|\right]ds\right)\\ &\leq C_\varepsilon\,(\Delta t)^{\frac{\min\{1,\nu\}}{2}-\varepsilon},\end{aligned} \tag{7.28}$$

where we used Theorem 7.4, Proposition 7.5, Lemma 4.1 and the properties of f_2. Moreover, for all $p \geq 1$ there exists a constant $C_p > 0$ such that

$$\mathbb{E}\left[|\tilde{M}_T|^p\right] \leq C_p$$

due to Lemma 4.1 and Proposition 2.2. Thus, a standard application of Hölder's inequality as in the proof of Proposition 4.6, part (ii), yields

$$\mathbb{E}\left[|\tilde{M}_T|^{1+\beta}\right] \leq C_\beta\left(\mathbb{E}\left[|\tilde{M}_T|\right]\right)^{\frac{1}{1+\beta}},$$

which in turn together with (7.27) and (7.28) gives

$$\mathbb{E}\left[\sup_{t\in[0,T]}|M_t|\right] \leq C_{\beta,\varepsilon}\,(\Delta t)^{\left(\frac{\min\{1,\nu\}}{2}-\varepsilon\right)\frac{1}{(1+\beta)^2}}. \tag{7.29}$$

(iii) The assertion follows now from (7.26) and (7.29) by choosing ε and β sufficiently small. □

7.4.2 Euler schemes - Case II

The second case can be treated analogously for $\nu > 1$, except at one point. Here the martingale $\tilde{M}$ is given by

$$\begin{aligned}\tilde{M}_t &:= \int_0^t \left(\sqrt{V_s} - \sqrt{\hat{v}^\star_{\eta(s)}}\right) dW_s \\ &= \frac{1}{\sigma}\left(V_t - \hat{v}^\star_t + \kappa \int_0^t \left(V_s - \hat{v}^\star_{\eta(s)}\right) ds\right) + 2c^\star \int_0^t \mathbb{1}_{\{z^\star_s \le 0\}} \sqrt{\hat{v}^\star_{\eta(s)}} dW_s \\ &\quad + \frac{2c^\star}{\sigma} \int_0^t \mathbb{1}_{\{z^\star_s \le 0\}} \kappa \left(\theta - \hat{v}^\star_{\eta(s)}\right) ds - \frac{c^\star}{\sigma} L^0_t(z^\star), \qquad t \in [0,T].\end{aligned}$$

However, the additional terms can be treated with the Lyapunov inequality and Lemma 4.7 and are (at least) of order $(\Delta t)^{\frac{1}{2}-\varepsilon}$. Using Theorem 7.6 and Lemma 4.2 instead of Theorem 7.4 and Lemma 4.1 and proceeding as in Case I we obtain

$$\mathbb{E}\left[|\tilde{M}_T|\right] \le C_\varepsilon (\Delta)^{\frac{1}{2}-\varepsilon}.$$

Therefore, we also have the following result:

Theorem 7.11. *Let $\nu > 1$ and $(\hat{x}, \hat{v})$ be given by* (7.6) *and* (7.5)*. Then, for all $\epsilon > 0$ there exists a constant $C_\epsilon > 0$ such that*

$$\mathbb{E}\left[\sup_{t \in [0,T]} |X_t - \hat{x}_t|\right] \le C_\epsilon \left(\Delta t\right)^{\frac{1}{2}-\epsilon}.$$

7.4.3 Implicit Milstein and Euler

The Milstein case can be again treated analogously for $\nu > \frac{1}{2}$, except at one point. By rearranging Equation (7.20), the martingale $\tilde{M}$ is given by

$$\begin{aligned}\tilde{M}_t &:= \int_0^t \left(\sqrt{V_s} - \sqrt{\hat{v}_{\eta(s)}}\right) dW_s \\ &= \frac{1}{\sigma}\left(V_t - \hat{v}_t + \kappa \int_0^t (V_s - \hat{v}_s) ds - \kappa \int_0^{\eta(t)} \left(\hat{v}_{\eta^+(s)} - \hat{v}_s\right) ds - \kappa \int_{\eta(t)}^t \left(\hat{v}_t - \hat{v}_s\right) ds\right) \\ &\quad - \int_0^t \frac{\sigma}{2} \left(W_s - W_{\eta(s)}\right) dW_s.\end{aligned}$$

Therefore, $\tilde{M}_T$ is given by

$$\begin{aligned}\tilde{M}_T = \frac{1}{\sigma}&\left(V_T - \hat{v}_T + \kappa \int_0^T (V_s - \hat{v}_s) ds - \kappa \int_0^T \left(\hat{v}_{\eta^+(s)} - \hat{v}_s\right) ds\right) \\ &- \int_0^T \frac{\sigma}{2} \left(W_s - W_{\eta(s)}\right) dW_s.\end{aligned}$$

These terms can be treated with Proposition 7.8, Proposition 7.9, Lemma 6.8 and standard estimations. We can proceed as in Case I and obtain

$$\mathbb{E}\left[|\tilde{M}_T|\right] \leq C_\varepsilon(\Delta)^{\frac{\min\{1,\nu\}}{2}-\varepsilon}.$$

Therefore, we also have the following result:

Proposition 7.12. *Let $(\hat{x}, \hat{v})$ be given by (7.6) and (7.7). Then, for all $\epsilon > 0$ there exists a constant $C_\epsilon > 0$ such that*

$$\mathbb{E}\left[\sup_{t\in[0,T]} |X_t - \hat{x}_t|\right] \leq C_\epsilon \left(\Delta t\right)^{\frac{\min\{1,\nu\}}{2}-\epsilon}.$$

7.5 Summary

Let us briefly summarize our results. We proved an upper bound of the L^1-convergence order of a large class of explicit Euler schemes for the CIR process and for the log-Heston model under the assumption that the Feller index is larger than 1.

Theorem 7.13. *Let $\nu > 1$, $\epsilon > 0$ and $(\hat{v}_{t_k}, \hat{x}_{t_k})_{k\in\{0,\dots,N\}}$ given by Equations (7.1), (7.2), (3.9) or by Equations (7.1), (7.3), (3.9). Then we have*

$$\lim_{N\to\infty} N^{1/2-\epsilon}\left(\max_{k\in\{0,\dots,N\}} \mathbb{E}\left[|X_{t_k} - \hat{x}_{t_k}|\right] + \max_{k\in\{0,\dots,N\}} \mathbb{E}\left[|V_{t_k} - \hat{v}_{t_k}|\right]\right) = 0.$$

This follows directly from Theorem 7.4, Theorem 7.6, Theorem 7.10 and Theorem 7.11. Thus, we recover (up to an arbitrarily small $\epsilon > 0$) the standard convergence order of the Euler scheme for SDEs with globally Lipschitz continuous coefficients. For the Case $\nu \leq 1$, we proved an L^1-convergence order of $\frac{\nu}{2}$ for the first case of Euler methods.

Proposition 7.14. *Let $\nu \leq 1$, $\epsilon > 0$ and $(\hat{v}_{t_k}, \hat{x}_{t_k})_{k\in\{0,\dots,N\}}$ given by Equations (7.1), (7.2), (3.9). Then we have*

$$\lim_{N\to\infty} N^{\nu/2-\epsilon}\left(\max_{k\in\{0,\dots,N\}} \mathbb{E}\left[|X_{t_k} - \hat{x}_{t_k}|\right] + \max_{k\in\{0,\dots,N\}} \mathbb{E}\left[|V_{t_k} - \hat{v}_{t_k}|\right]\right) = 0.$$

This is due to Proposition 7.5 and Theorem 7.10. However, our numerical simulations in Chapter 9 indicate that the sharp rate should be $\min\{\frac{1}{2}, \nu\}$.
For the implicit Milstein scheme in combination with the standard Euler approximation of the log-price process we could prove an L^1-convergence order of $\frac{\min\{\nu,1\}}{2} - \epsilon$ which holds for the whole parameter range where this scheme is positivity preserving.

Proposition 7.15. *Let $\nu > \frac{1}{2}$, $\epsilon > 0$ and $(\hat{v}_{t_k}, \hat{x}_{t_k})_{k\in\{0,\dots,N\}}$ given by Equations (3.10) and (3.9). Then we have*

$$\lim_{N\to\infty} N^{\frac{\min\{\nu,1\}}{2}-\epsilon}\left(\max_{k\in\{0,\dots,N\}} \mathbb{E}\left[|X_{t_k} - \hat{x}_{t_k}|\right] + \max_{k\in\{0,\dots,N\}} \mathbb{E}\left[|V_{t_k} - \hat{v}_{t_k}|\right]\right) = 0.$$

This follows from Proposition 7.8, Proposition 7.9 and Proposition 7.12.

Remark 7.16. *By a standard application of Hölder's inequality, we could deduce L^p-convergence orders for $p > 1$ for all presented schemes. These would be $\frac{1}{2p} - \epsilon$ for the setting of Theorem 7.13, $\frac{\nu}{2p} - \epsilon$ for Proposition 7.14 and $\frac{\min\{\nu,1\}}{2p} - \epsilon$ for Proposition 7.15. However, these bounds are unlikely to be sharp, see e.g. [9], [21], so we do not spell out these results in detail.*

Remark 7.17. *The results of Theorem 7.13 and Proposition 7.14 appear in [55]. This manuscript has been accepted for publication in the Journal of Computational Finance.*

Chapter 8

L^1-Approximation of the Log-Heston SDE: Lower Bounds

In the last chapter, we derived upper bounds for the L^1-approximation of the log-Heston SDE by Euler-type and Implicit Milstein methods. Now we would like to know: Which order is the best possible when we use an equidistant discretization? This question has been answered for the CIR process by the works [33] and [34], which yield

$$\liminf_{N\to\infty} N^{\min\{\nu,1\}} \inf_{u\in\mathcal{U}(N)} \mathbb{E}\left[\left|u(W_{t_1}, W_{t_2}, \dots, W_{t_N}) - V_T\right|\right] > 0,$$

where $\mathcal{U}(N)$ is the set of measurable functions $u : \mathbb{R}^N \to \mathbb{R}$ and we have $t_k = k\frac{T}{N}$ for $k \in \{0, ..., N\}$. Recalling the presented results from Section 7.1, the convergence order of the truncated Milstein scheme for $\nu \leq \frac{1}{2}$ and the order of the drift-implicit Euler for $\nu > 2$ are optimal. However, the optimal approximation of the log-Heston SDE has not been studied yet up to the best of our knowledge.
In this chapter, we show that for $\nu > 1$ and $|\rho| \neq 1$ the convergence orders of the studied schemes from Chapter 7 are optimal, since arbitrary methods that use an equidistant discretization of the driving two-dimensional Brownian motion (W, B) can achieve at most order $\frac{1}{2}$ for the L^1-approximation at the final time point.

Theorem 8.1. *Let $\nu > 1$, $|\rho| \neq 1$, $t_k = k\Delta t$ for $k \in \{0, ..., N\}$ with $\Delta t = \frac{T}{N}$, let $\mathcal{U}(N)$ be the set of measurable functions $u : \mathbb{R}^{2N} \to \mathbb{R}$ and*

$$e(N) = \inf_{u\in\mathcal{U}(N)} \mathbb{E}\left[\left|u(W_{t_1}, W_{t_2}, \dots, W_{t_N}, B_{t_1}, B_{t_2}, \dots, B_{t_N}) - X_T\right|\right].$$

Then we have that

$$\liminf_{N\to\infty} \sqrt{N}\, e(N) \geq \frac{\sigma T}{8}\sqrt{1-\rho^2}.$$

Remark 8.2. *A modified version of this result has been accepted for publication in the proceedings of the 15th International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing, a standard outlet for complexity results. A preprint can be found in [53].*

The pioneering work on optimal approximation of stochastic differential equations is [18]. Clark and Cameron studied in particular the optimal L^2-approximation of

$$\begin{aligned} dX_t &= V_t dB_t, \\ dV_t &= dW_t, \end{aligned} \qquad t \in [0,1],$$

at the final time point by an equidistant discretization of the driving Brownian motion. Here, the optimal method is given by

$$\mathbb{E}\big[X_1 \big| W_{\frac{1}{N}}, ..., W_1, B_{\frac{1}{N}}, ..., B_1\big]$$

and one has

$$\left(\mathbb{E}\left[\Big|X_1 - \mathbb{E}\big[X_1 \big| W_{\frac{1}{N}}, ..., W_1, B_{\frac{1}{N}}, ..., B_1\big]\Big|^2\right]\right)^{1/2} = \frac{1}{2} N^{-1/2}.$$

Since then, a detailed and exhaustive study for the optimal approximation of general SDEs under standard assumptions has been carried out for various error criteria. See e.g. [15, 16, 38, 39, 40, 41, 57, 58, 59, 63, 64, 65] and [60] for a survey.
Recently, the analysis of the optimal approximation of SDEs has been extended to the case of non-standard coefficients. We already mentioned the works [31, 33, 34] which analyze the optimal approximation of the squared Bessel process respectively of the CIR process. In [46, 61, 73] SDEs with arbitrary slow best possible convergence rates are constructed.

8.1 Proof of Theorem 8.1

We will simplify the analysis of

$$e(N) = \inf_{u \in \mathcal{U}} \mathbb{E}\left[\left|u(W_{t_1}, W_{t_2}, \ldots, W_{t_N}, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - X_T\right|\right]$$

in several steps until we end up with the optimal L^1-approximation of $\int_0^T B_t dW_t$ by arbitrary methods, which use an equidistant discretization of B and have complete information of W, i.e. with the analysis of the quantity

$$\inf_{v \in \mathcal{V}} \mathbb{E}\left[\left|v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - \int_0^T B_t dW_t\right|\right],$$

where $\mathcal{V}$ is the set of measurable functions $v : C([0,T];\mathbb{R}) \times \mathbb{R}^N \to \mathbb{R}$. This quantity can be then analyzed in a final step by a symmetrization argument. The latter is a simplified version of Lemma 1 in [46] and is a particular case of the *radius of information* concept in information based complexity, see [71].

8.1.1 Allowing complete information on W

Let

$$\mathcal{G}_N = \sigma\left(W_{t_1}, W_{t_2}, \ldots, W_{t_N}, B_{t_1}, B_{t_2}, \ldots, B_{t_N}\right), \quad \mathcal{H}_N = \sigma\left(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}\right)$$

and

$$\mathcal{Z}_{\mathcal{U}} = \{Z : \Omega \to \mathbb{R} : \ Z \text{ is } \mathcal{G}_N \text{ measurable}\},\ \mathcal{Z}_{\mathcal{V}} = \{Z : \Omega \to \mathbb{R} : \ Z \text{ is } \mathcal{H}_N \text{ measurable}\}.$$

Since

$$\mathcal{Z}_{\mathcal{U}} \subset \mathcal{Z}_{\mathcal{V}}$$

it follows that

$$\begin{aligned} \inf_{Z \in \mathcal{Z}_{\mathcal{U}}} \mathbb{E}\left[|Z - X_T|\right] &= \inf_{u \in \mathcal{U}} \mathbb{E}\left[|u(W_{t_1}, W_{t_2}, \ldots, W_{t_N}, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - X_T|\right] \\ &\geq \inf_{Z \in \mathcal{Z}_{\mathcal{V}}} \mathbb{E}\left[|Z - X_T|\right] = \inf_{v \in \mathcal{V}} \mathbb{E}\left[|v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - X_T|\right], \end{aligned}$$

where $\mathcal{V}$ is as above. Thus, it is sufficient to analyze the quantity

$$\inf_{Z \in \mathcal{Z}_{\mathcal{V}}} \mathbb{E}\left[|Z - X_T|\right] = \inf_{v \in \mathcal{V}} \mathbb{E}\left[|v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - X_T|\right] \tag{8.1}$$

to obtain a lower bound for $e(N)$.

8.1.2 Rewriting X_T and removing the measurable part

Now, we rewrite X_T. Note that the CIR process $V = (V_t)_{t\in[0,T]}$ is $\sigma(W)$-measurable and therefore $\mathcal{H}_N$-measurable as the unique strong solution of SDE (2.2).

Lemma 8.3. *For $\nu > 1$ we have that*

$$X_T = Y_T + \sqrt{1-\rho^2} \int_0^T A_t dB_t - \frac{\sigma}{2}\sqrt{1-\rho^2} \int_0^T B_t dW_t$$

where

$$\begin{aligned} Y_T = x &+ \frac{\rho}{\sigma}\left(V_T - v - \kappa\theta T\right) + \mu T + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) \int_0^T V_u du \\ &+ \sqrt{1-\rho^2}(\sqrt{V_T} B_T - A_T B_T) \end{aligned}$$

and

$$A_t = \frac{4\kappa\theta - \sigma^2}{8} \int_0^t \frac{1}{\sqrt{V_u}}\, du - \frac{\kappa}{2} \int_0^t \sqrt{V_u}\, du, \qquad t \in [0, T].$$

In particular, $A = (A_t)_{t\in[0,T]}$ and Y_T are $\mathcal{H}_N$-measurable.

Proof. Since

$$V_T = v + \int_0^T \kappa(\theta - V_u)du + \sigma \int_0^T \sqrt{V_u} dW_u$$

we have that

$$X_T = Y_T^{(1)} + \sqrt{1-\rho^2} \int_0^T \sqrt{V_u} dB_u \tag{8.2}$$

with

$$Y_T^{(1)} = x + \frac{\rho}{\sigma}(V_T - v - \kappa\theta T) + \mu T + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)\int_0^T V_u du.$$

Since almost all sample paths of V are strictly positive due to $\nu > 1$ we can use Itō's lemma to write

$$\sqrt{V_t} = \sqrt{v} + A_t + \frac{\sigma}{2} W_t, \qquad t \in [0, T], \tag{8.3}$$

where

$$A_t = \frac{4\kappa\theta - \sigma^2}{8} \int_0^t \frac{1}{\sqrt{V_u}} du - \frac{\kappa}{2} \int_0^t \sqrt{V_u} du, \qquad t \in [0, T].$$

So $\sqrt{V}$ is a continuous semi-martingale with representation (8.3). Integration by parts now gives

$$\int_0^T \sqrt{V_u} dB_u = \sqrt{V_T} B_T - \int_0^T B_t dA_t - \frac{\sigma}{2} \int_0^T B_t dW_t$$

and

$$\int_0^T B_t dA_t = A_T B_T - \int_0^T A_t dB_t,$$

respectively. This gives

$$\int_0^T \sqrt{V_u} dB_u = \sqrt{V_T} B_T - B_T A_T + \int_0^T A_t dB_t - \frac{\sigma}{2} \int_0^T B_t dW_t$$

and (8.2) yields

$$X_T = Y_T^{(1)} + Y_T^{(2)} + \sqrt{1-\rho^2}\left(\int_0^T A_t dB_t - \frac{\sigma}{2}\int_0^T B_t dW_t\right)$$

with

$$Y_T^{(2)} = \sqrt{1-\rho^2}(\sqrt{V_T} B_T - B_T A_T),$$

which finishes the proof. $\square$

As a consequence, we have

$$\begin{aligned}\inf_{Z \in \mathcal{Z}_\mathcal{V}} \mathbb{E}\left[|Z - X_T|\right] &= \inf_{Z \in \mathcal{Z}_\mathcal{V}} \mathbb{E}\left[\left|Z - Y_T - \sqrt{1-\rho^2}\left(\int_0^T A_t dB_t - \frac{\sigma}{2}\int_0^T B_t dW_t\right)\right|\right] \\ &= \inf_{\tilde{Z} \in \mathcal{Z}_\mathcal{V}} \mathbb{E}\left[\left|\tilde{Z} - \sqrt{1-\rho^2}\left(\int_0^T A_t dB_t - \frac{\sigma}{2}\int_0^T B_t dW_t\right)\right|\right]\end{aligned}$$

and it remains to analyze

$$\inf_{v\in\mathcal{V}} \mathbb{E}\left[\left|v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - \int_0^T A_t dB_t + \frac{\sigma}{2}\int_0^T B_t dW_t\right|\right]. \tag{8.4}$$

8.1.3 Removing the smooth part

Since $A = (A_t)_{t\in[0,T]}$ is smooth enough, $\int_0^T A_t dB_t$ does not matter asymptotically for our approximation problem.

Lemma 8.4. *Let $\nu > 1$. Then, there exists a constant $C > 0$ such that*

$$\mathbb{E}\left[\left|\int_0^T A_t dB_t - \sum_{i=0}^{N-1} A_{t_i}(B_{t_{i+1}} - B_{t_i})\right|\right] \le C \cdot N^{-1}.$$

Proof. We have

$$A_t = \int_0^t a_u du$$

with

$$a_u = \frac{4\kappa\theta - \sigma^2}{8}\frac{1}{\sqrt{V_u}} - \frac{\kappa}{2}\sqrt{V_u}, \qquad u \in [0,T].$$

Since $\nu > 1$ we have by Lemma 2.2 and Jensen's inequality that

$$\sup_{t\in[0,T]} \mathbb{E}\left[|a_t|^2\right] < \infty, \qquad \sup_{t\in[0,T]} \mathbb{E}\left[|A_t|^2\right] < \infty.$$

The Itō isometry now gives

$$\begin{aligned}\mathbb{E}\left[\left|\int_0^T A_t dB_t - \sum_{i=0}^{N-1} A_{t_i}(B_{t_{i+1}} - B_{t_i})\right|^2\right] &= \mathbb{E}\left[\left|\int_0^T (A_t - A_{\eta(t)}) dB_u\right|^2\right] \\ &= \int_0^T \mathbb{E}\left[\left|\int_{\eta(t)}^t a_u du\right|^2\right] dt.\end{aligned}$$

Moreover, the Cauchy-Schwartz inequality yields

$$\mathbb{E}\left[\left|\int_{\eta(t)}^t a_u du\right|^2\right] \le T^2 \sup_{t\in[0,T]} \mathbb{E}\left[|a_t|^2\right](t - \eta(t))^2 \le C \cdot (\Delta t)^2$$

and so we have

$$\mathbb{E}\left[\left|\int_0^T A_t dB_t - \sum_{i=0}^{N-1} A_{t_i}(B_{t_{i+1}} - B_{t_i})\right|^2\right] \le C \cdot N^{-2}.$$

The assertion follows now from the Lyapunov inequality. □

Since $\sum_{i=0}^{N-1} A_{t_i}(B_{t_{i+1}} - B_{t_i})$ is $\mathcal{H}_N$-measurable, we obtain that

$$\begin{aligned}
&\inf_{v \in \mathcal{V}} \mathbb{E}\left[\left| v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - \int_0^T A_t dB_t + \frac{\sigma}{2}\int_0^T B_t dW_t \right|\right] \\
&\quad = \inf_{\tilde{v} \in \mathcal{V}} \mathbb{E}\left[\left| \tilde{v}(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) + \frac{\sigma}{2}\int_0^T B_t dW_t - \int_0^T (A_t - A_{\eta(t)}) dB_t \right|\right] \\
&\quad \geq \inf_{\tilde{v} \in \mathcal{V}} \mathbb{E}\left[\left| \tilde{v}(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) + \frac{\sigma}{2}\int_0^T B_t dW_t \right|\right] - C \cdot N^{-1}
\end{aligned}$$

using that $|x| - |y| \leq |x - y|$ for all $x, y \in \mathbb{R}$. Consequently, we have reduced our initial problem to the study of

$$\inf_{v \in \mathcal{V}} \mathbb{E}\left[\left| v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - \int_0^T B_t dW_t \right|\right]. \tag{8.5}$$

8.1.4 Inserting Brownian bridges and symmetrization

For the final step let us denote the piecewise linear interpolation of B on the grid $t_0, ..., t_N$ by $\overline{B}$, i.e. $\overline{B}$ is defined as

$$\overline{B}_t = B_{t_k} + \frac{t - t_k}{t_{k+1} - t_k}(B_{t_{k+1}} - B_{t_k}), \qquad t \in [t_k, t_{k+1}],\ k = 0, \ldots, N-1.$$

Then the process B° given by

$$B_t^\circ = B_t - \overline{B}_t, \qquad t \in [0, T],$$

is a Brownian bridge on $[t_k, t_{k+1}]$ for $k = 0, \ldots, N-1$, and moreover the processes

$$(B_t^\circ)_{t \in [t_0, t_1]},\ (B_t^\circ)_{t \in [t_1, t_2]},\ \ldots,\ (B_t^\circ)_{t \in [t_{N-1}, t_N]},\ \overline{B},\ W$$

are independent. Since

$$\int_0^T \overline{B}_t dW_t = \sum_{k=0}^{N-1} B_{t_k}\left(W_{t_{k+1}} - W_{t_k}\right) + \sum_{k=0}^{N-1} \frac{B_{t_{k+1}} - B_{t_k}}{t_{k+1} - t_k} \int_{t_k}^{t_{k+1}} (t - t_k) dW_t$$

is $\mathcal{H}_N$-measurable, we have that

$$\begin{aligned}
&\inf_{v \in \mathcal{V}} \mathbb{E}\left[\left| v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - \int_0^T B_t dW_t \right|\right] \\
&\qquad = \inf_{\tilde{v} \in \mathcal{V}} \mathbb{E}\left[\left| \tilde{v}(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - \mathcal{I}(B^\circ, W) \right|\right]
\end{aligned}$$

with

$$\mathcal{I}(B^\circ, W) = \int_0^T B_t dW_t - \int_0^T \overline{B}_t dW_t.$$

Furthermore B° and $-B^\circ$ have the same law, so the independence of B° from $(W, \overline{B})$ implies that

$$\left(W, \overline{B}, B^\circ\right) \overset{d}{=} \left(W, \overline{B}, -B^\circ\right). \tag{8.6}$$

Now we will analyze $\mathcal{I}(B^\circ, W)$ in more detail.

Lemma 8.5.

(i) Let

$$\tau_{\ell,n} = \frac{\ell}{2^n}\frac{T}{N}, \qquad \ell = 0, \dots, 2^n,$$

and

$$\mathcal{I}^n(B^\circ, W) = \sum_{k=0}^{N-1}\sum_{\ell=0}^{2^n-1} B^\circ_{t_k+\tau_{\ell,n}}\left(W_{t_k+\tau_{\ell+1,n}} - W_{t_k+\tau_{\ell,n}}\right).$$

We have that

$$\mathcal{I}(B^\circ, W) = \lim_{n\to\infty} \mathcal{I}^n(B^\circ, W)$$

almost surely and in L^2.

(ii) It holds

$$\mathcal{I}(B^\circ, W) \overset{d}{=} W_1 \left(\int_0^T |B^\circ_t|^2 dt\right)^{1/2}.$$

Proof. (i) We have

$$\int_{t_k}^{t_{k+1}} B_t dW_t - \int_{t_k}^{t_{k+1}} \overline{B}_t dW_t = I_1^k - I_2^k$$

with

$$I_1^k = \int_{t_k}^{t_{k+1}} (B_t - B_{t_k}) dW_t, \qquad I_2^k = \frac{B_{t_{k+1}} - B_{t_k}}{t_{k+1} - t_k}\int_{t_k}^{t_{k+1}} (t - t_k) dW_t$$

and

$$\sum_{\ell=0}^{2^n-1} B^\circ_{t_k+\tau_{\ell,n}}\left(W_{t_k+\tau_{\ell+1,n}} - W_{t_k+\tau_{\ell,n}}\right) = I_1^{k,n} - I_2^{k,n}$$

with

$$I_1^{k,n} = \sum_{\ell=0}^{2^n-1} \left(B_{t_k+\tau_{\ell,n}} - B_{t_k}\right)\left(W_{t_k+\tau_{\ell+1,n}} - W_{t_k+\tau_{\ell,n}}\right),$$

$$I_2^{k,n} = \frac{B_{t_{k+1}} - B_{t_k}}{t_{k+1} - t_k}\sum_{\ell=0}^{2^n-1} \tau_{\ell,n}\left(W_{t_k+\tau_{\ell+1,n}} - W_{t_k+\tau_{\ell,n}}\right).$$

By straightforward calculations using the independence of B and W and the Itō isometry we have

$$\mathbb{E}\left[|I_1^k - I_1^{k,n}|^2\right] = \frac{1}{2}\left(\frac{T}{N}\right)^2 2^{-n}$$

and

$$\mathbb{E}\left[|I_2^k - I_2^{k,n}|^2\right] = \frac{1}{3}\left(\frac{T}{N}\right)^2 2^{-2n}.$$

Thus

$$\mathbb{E}\left[|\mathcal{I}(B^\circ, W) - \mathcal{I}^n(B^\circ, W)|^2\right] \leq C2^{-n},$$

which yields the L^2-convergence, and also implies

$$\sum_{n=1}^{\infty} \mathbb{E}\left[|\mathcal{I}(B^\circ, W) - \mathcal{I}^n(B^\circ, W)|\right] < \infty,$$

from which the almost sure convergence follows by an application of the Borel-Cantelli lemma.
(ii) Recall that W is independent of B°. The conditional law of $\mathcal{I}^n(B^\circ, W)$ given

$$B^\circ_{t_k+\tau_{\ell,n}} = x_{k,\ell}, \quad \ell = 0, \dots 2^n - 1, k = 0, \dots, N-1,$$

is therefore Gaussian with zero mean and variance $\sum_{k=0}^{N-1}\sum_{\ell=0}^{2^n-1}|x_{k,\ell}|^2(\tau_{\ell+1,n} - \tau_{\ell,n})$. We thus have

$$\mathcal{I}^n(B^\circ, W) \stackrel{d}{=} W_1\left(\sum_{k=0}^{N-1}\sum_{\ell=0}^{2^n-1}|B^\circ_{t_k+\tau_{\ell,n}}|^2\,(\tau_{\ell+1,n} - \tau_{\ell,n})\right)^{1/2}.$$

Since also

$$\int_0^T |B^\circ_t|^2 dt = \lim_{n\to\infty}\sum_{k=0}^{N-1}\sum_{\ell=0}^{2^n-1}|B^\circ_{t_k+\tau_{\ell,n}}|^2\,(\tau_{\ell+1,n} - \tau_{\ell,n})$$

almost surely (by continuity of almost all sample paths of B°), the assertion follows now from part (i). □

The equality of the laws in (8.6) yields that

$$\left(W, \overline{B}, \mathcal{I}^n(B^\circ, W)\right) \stackrel{d}{=} \left(W, \overline{B}, -\mathcal{I}^n(B^\circ, W)\right)$$

and (i) from previous lemma now gives

$$\left(W, \overline{B}, \mathcal{I}(B^\circ, W)\right) \stackrel{d}{=} \left(W, \overline{B}, -\mathcal{I}(B^\circ, W)\right).$$

Consequently, we have

$$\mathbb{E}\left[|v(W, B_{t_1}, B_{t_2}, \dots, B_{t_N}) - \mathcal{I}(B^\circ, W)|\right] = \mathbb{E}\left[|v(W, B_{t_1}, B_{t_2}, \dots, B_{t_N}) + \mathcal{I}(B^\circ, W)|\right]$$

and so

$$\begin{aligned}
&2\mathbb{E}\left[|\mathcal{I}(B^\circ, W)|\right] \\
&= \mathbb{E}\left[|(\mathcal{I}(B^\circ, W) - v(W, B_{t_1}, \dots, B_{t_N})) + (v(W, B_{t_1}, \dots, B_{t_N}) + \mathcal{I}(B^\circ, W))|\right] \\
&\leq 2\mathbb{E}\left[|v(W, B_{t_1}, B_{t_2}, \dots, B_{t_N}) - \mathcal{I}(B^\circ, W)|\right].
\end{aligned}$$

It follows that

$$\inf_{v\in\mathcal{V}} \mathbb{E}\left[|v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - \mathcal{I}(B^\circ, W)|\right] \geq \mathbb{E}\left[|\mathcal{I}(B^\circ, W)|\right]$$

and therefore we have

$$\begin{aligned}\inf_{v\in\mathcal{V}} \mathbb{E}\left[|v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - \mathcal{I}(B^\circ, W)|\right] &\geq \mathbb{E}\left[|W_1|\right] \mathbb{E}\left[\left(\int_0^T |B_t^\circ|^2 dt\right)^{1/2}\right] \\ &\geq \frac{1}{\sqrt{T}} \mathbb{E}\left[|W_1|\right] \int_0^T \mathbb{E}\left[|B_t^\circ|\right] dt\end{aligned}$$

by Lemma 8.5(ii) and by Jensen's inequality. Using $\mathbb{E}\left[|X|\right] = \sqrt{\frac{2}{\pi}}\sigma$ for $X \sim \mathcal{N}(0, \sigma^2)$ we obtain

$$\int_0^T \mathbb{E}\left[|B_t^\circ|\right] dt = \frac{\sqrt{2}}{\sqrt{\pi}} \int_0^T \mathbb{E}\left[|B_t^\circ|^2\right]^{1/2} dt.$$

Straightforward calculations give

$$\mathbb{E}\left[|B_t^\circ|^2\right] = \frac{(t - t_k)(t_{k+1} - t)}{t_{k+1} - t_k}, \qquad t \in [t_k, t_{k+1}],$$

which in turn yields

$$\int_0^T \mathbb{E}\left[|B_t^\circ|^2\right]^{1/2} dt = N \int_0^{T/N} \sqrt{\frac{t(T/N - t)}{T/N}} dt = \sqrt{\frac{T^3}{N}} \int_0^1 \sqrt{x(1-x)} dx.$$

Since $\int_0^1 \sqrt{x(1-x)} dx = \frac{\pi}{8}$, we have shown that

$$\sqrt{N} \inf_{v\in\mathcal{V}} \mathbb{E}\left[\left|v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - \int_0^T B_t dW_t\right|\right] \geq \frac{T}{4}. \tag{8.7}$$

Combining subsections 8.1.1–8.1.4 with Equations (8.1), (8.4), (8.5) and (8.7) concludes the proof of Theorem 8.1.

Chapter 9

Numerical Results

In this chapter, we test our results from Chapters 6, 7 and 8 by performing numerical simulations on an exemplary parameter set. We start with weak convergence simulations for the Heston model and then continue with L^1-error simulations for both the CIR process and the full Heston model.

9.1 Weak Convergence

In this section, we test numerically whether the weak convergence orders of Theorem 6.3 and Proposition 6.9 are attained even under milder assumptions on the test function f. We consider a call, a put and a digital option. These payoffs are at most Lipschitz continuous which is typical in financial applications. This lack of smoothness is in contrast to the usual assumptions on f for a weak error analysis. See also Remark 6.6. Besides the SE, AE and implicit Milstein, we also numerically test the drift-implicit Euler for the CIR process in combination with the standard Euler for the log-price process. For simplicity, we call the scheme *Drift-Implicit*. We have seen in Chapter 7 that the strong error behavior of the drift-implicit Euler for the CIR process is well analyzed and seems to be superior to the one of the Euler schemes. Weak convergence results are not available to the best of our knowledge. We would like to compare their weak convergence behaviors in the context of the full Heston model.
Our model parameters are displayed in Table 9.1.

Model	S_0	K	V_0	κ	θ	σ	ρ	T	r	ν (approx.)
1	100	100	0.04	5	0.04	0.61	-0.7	1	0.0319	1.075
2	100	100	0.04	5.5	0.04	0.55	-0.7	1	0.0319	1.45
3	100	100	0.010201	6.21	0.019	0.61	-0.7	1	0.0319	0.63
4	100	100	0.09	2	0.09	1	-0.3	5	0.05	0.36

Table 9.1: Parameters for the weak convergence test.

We have $\nu \approx 1.075$ in Model 1, $\nu \approx 1.45$ in Model 2, $\nu \approx 0.63$ in Model 3 and $\nu \approx 0.36$

in Model 4. The parameter sets for Model 3 and 4 are taken from [14]. For Model 1 and Model 2 we adjusted the parameters of Model 3 such that they have Feller indices around 1 and 1.5. With these examples we set our focus on low values of the Feller index ν since this is the most interesting parameter range. For each model, we use the following payoff functions:

1. European Call: $g_1(S_T) = e^{-rT} \max\{S_T - K, 0\}$
2. European Put: $g_2(S_T) = e^{-rT} \max\{K - S_T, 0\}$
3. Digital Option: $g_3(S_T) = e^{-rT} \mathbb{1}_{[0,K]}(S_T)$

Note that none of these payoffs satisfies the assumption of Theorem 6.3. Thus, numerical convergence orders which coincide with the orders of our Theorem indicate that the latter might be valid under milder assumptions.
In order to measure the weak error order, we simulated $M = 2 \cdot 10^7$ independent copies $g_i(s_N^{(j)})$, $j = 1, \ldots, M$, of $g_i\,(s_N)$ with $s_N = \exp(x_N)$ to estimate

$$\mathbb{E}\left[g_i(s_N)\right]$$

by

$$\hat{p}_{M,N} = \frac{1}{M} \sum_{j=1}^{M} g_i(s_N^{(j)})$$

for each combination of model parameters, payoff and number of steps $N \in \{2^3, ..., 2^8\}$ where $\Delta t = \frac{T}{N}$. To obtain a stable estimate of the convergence orders, we started with a Δt which is smaller $\frac{1}{\kappa}$ (which is required also for some auxiliary results of the proof of our main result). The Monte Carlo mean of these samples was then compared to a reference solution p_{ref}, i.e.,

$$e(N) = |p_{\mathsf{ref}} - \hat{p}_{M,N}|,$$

and the error $e(N)$ is plotted in Figures 9.1–9.12. We measure the weak error order by the slope of a least-squares fit. The reference solutions can be computed with sufficiently high accuracy from semi-explicit formulae via Fourier methods. In particular, the put price can be calculated from the call price formula given in [36] via the put-call-parity. The price of the digital option can be computed from the probability P_2 given in [36]; it equals $e^{-rT}\,(1 - P_2)$.
In Table 9.2 and Figures 9.1-9.3, we can see the measured convergence orders and the error plots for Model 1. Because of our results in Theorem 6.3 and Proposition 6.9, we would expect SE, AE and implicit Milstein to have a weak convergence order of 1 and this is indeed the case in this example. The implicit Milstein seems to have a lower convergence order for put and call but the plots 9.1–9.3 show that this might be due to the low error that this scheme produces right from the start. Also, its convergence behavior is not very regular in these cases. The Drift-Implicit scheme seems to have an overall lower convergence order than the other schemes. Only for the digital option it has a lower absolute error.

Method	Call	Put	Digital
SE	1.00	0.96	0.93
AE	1.03	1.01	0.91
Drift-Implicit	0.55	0.73	0.33
Implicit Milstein	0.73	0.72	1.04

Table 9.2: Estimated weak convergence orders Model 1

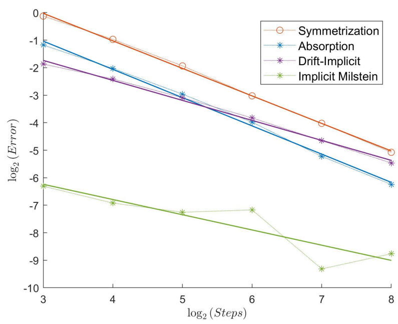


Figure 9.1: Call Model 1

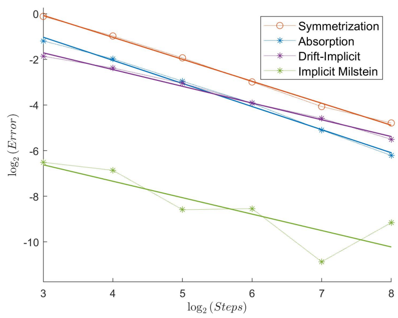


Figure 9.2: Put Model 1

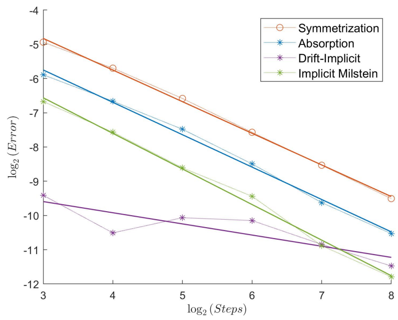


Figure 9.3: Digital Model 1

For the next model, we would again expect a convergence rate around 1. The results from Table 9.3 and plots in Figures 9.4-9.6 indicate that for particular payoffs even a higher numerical order is obtained if the Feller index is larger than 1. The orders of SE

Method	Call	Put	Digital
SE	1.34	1.27	1.18
AE	1.36	1.36	1.31
Drift-Implicit	0.74	0.79	0.65
Implicit Milstein	0.35	0.47	0.88

Table 9.3: Estimated weak convergence orders Model 2

and AE are around 1-1.3 and the convergence behavior is regular. The implicit Milstein shows a similar behavior as in the first model. Again, its error for call and put is very low but it does not seem to have a fast convergence. For the digital option, it also has a higher absolute error than the AE. The Drift-Implicit scheme performs again worse than the other schemes for call and put. It has a smaller absolute error in the digital case for our step size range but its estimated convergence order is lower.

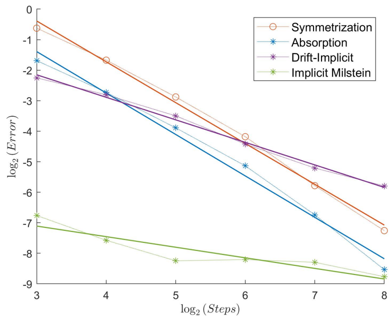


Figure 9.4: Call Model 2

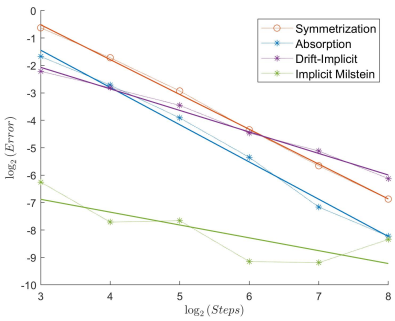


Figure 9.5: Put Model 2

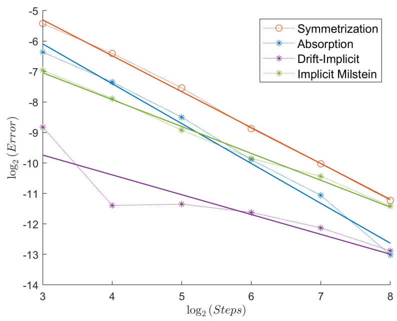


Figure 9.6: Digital Model 2

Table 9.4 shows the estimated convergence orders for Model 3. This model has a Feller index around 0.63. The simulation results indicate that this is also the convergence order for SE, AE and the Drift-Implicit scheme. The implicit Milstein has an estimated convergence order around 1. This is in line with our theoretical findings for smoother payoff functions. All plots are very regular (see Figures 9.7–9.9).

Method	Call	Put	Digital
SE	0.60	0.60	0.55
AE	0.57	0.57	0.55
Drift-Implicit	0.63	0.64	0.53
Implicit Milstein	1.05	0.94	1.30

Table 9.4: Estimated weak convergence orders Model 3

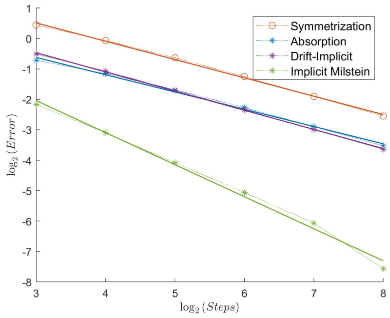


Figure 9.7: Call Model 3

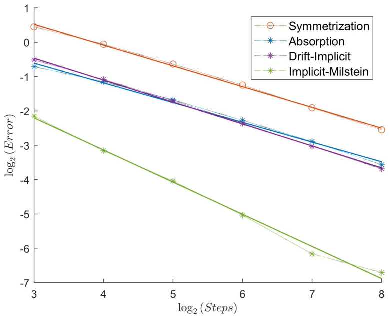


Figure 9.8: Put Model 3

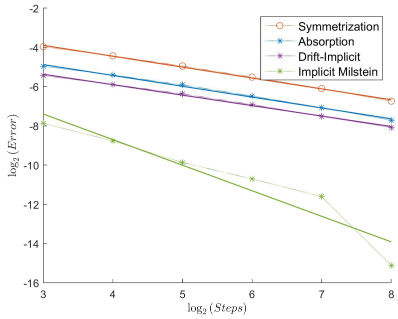


Figure 9.9: Digital Model 3

Model 4 has the lowest Feller index which is around 0.36. Again, Table 9.5 confirms this number as the numerical convergence order for SE and AE. Since $\nu < \frac{1}{2}$, we replaced the implicit Milstein by the truncated Milstein from Equation (3.11). Note that there are no weak convergence results available for this scheme to the best of our knowledge. We are not aware of a truncation of the drift-implicit Euler which was analyzed in the literature. The convergence order of truncated Milstein seems to be slightly higher and around 0.5. Our simulations confirm simulation studies in the literature that show a slow convergence for low Feller indices. Looking at Figures 9.10-9.12, the absolute values

Method	Call	Put	Digital
SE	0.47	0.47	0.40
AE	0.38	0.39	0.35
Truncated Milstein	0.53	0.52	0.45

Table 9.5: Estimated weak convergence orders Model 4

of all errors are quite high. The AE performs best up to $N = 2^8$. We again have a very regular convergence behavior.
Summarizing, we can confirm a (minimum) numerical convergence order of $\min\{\nu, 1\}$ for the symmetrized and absorbed Euler and of 1 for the implicit Milstein scheme under

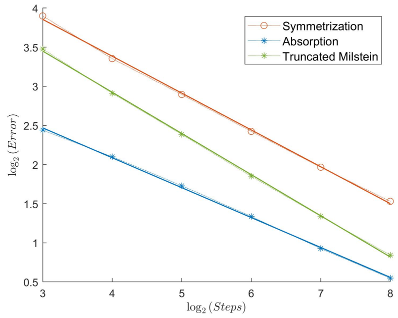


Figure 9.10: Call Model 4

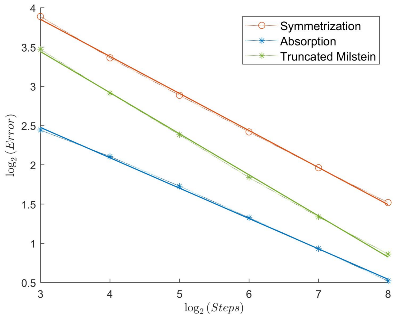


Figure 9.11: Put Model 4

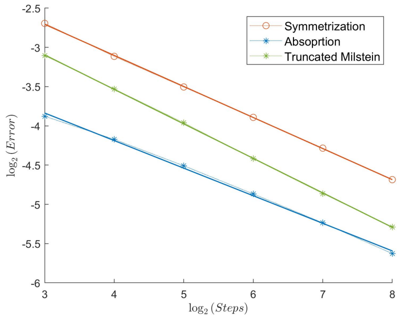


Figure 9.12: Digital Model 4

even milder assumptions on the regularity of the payoff function. We saw slightly better numerical convergence results of the Euler schemes for a higher Feller index. In most cases, the Drift-Implicit scheme performed worse than the other schemes. Furthermore, we cannot use it for low Feller indices which often occur in calibrations to real-world data. The implicit Milstein seems to have superior weak convergence properties in most of the cases than the two Euler schemes considered here. However, this effect seems to vanish for low values of ν when a truncation is needed.

9.2 Strong Convergence

In this section, we test numerically our results from Theorem 7.13, Proposition 7.14, Proposition 7.15 and Theorem 8.1. We perform the tests for all Euler schemes from Table 3.1 and, as before, for the implicit Milstein and the drift-implicit Euler.
First, we describe the design of the numerical experiments. We would like to estimate the order of the decay of the errors

$$e_v(N) = \mathbb{E}\left[\left|V_T - \hat{v}_{t_N}^{(N)}\right|\right], \qquad e_x(N) = \mathbb{E}\left[\left|X_T - \hat{x}_{t_N}^{(N)}\right|\right],$$

for the numerical scheme $\hat{v}^{(N)}, \hat{x}^{(N)}$ with step size $\Delta t = \frac{T}{N}$. Since we cannot compute these quantities exactly, we approximate their decay, see e.g. [1], by calculating

$$\texttt{err}_v(N) = \frac{1}{M}\sum_{i=1}^{M}\left|\left(\hat{v}_{t_N}^{(N)} - \hat{v}_{t_{2N}}^{(2N)}\right)^{(i)}\right|, \qquad \texttt{err}_x(N) = \frac{1}{M}\sum_{i=1}^{M}\left|\left(\hat{x}_{t_N}^{(N)} - \hat{x}_{t_{2N}}^{(2N)}\right)^{(i)}\right|,$$

where M is the number of Monte Carlo repetitions and $(\hat{v}_{t_N}^{(N)} - \hat{v}_{t_{2N}}^{(2N)})^{(i)}, i = 1, ..., M$, are iid copies of $\hat{v}_{t_N}^{(N)} - \hat{v}_{t_{2N}}^{(2N)}$. The same holds for $(\hat{x}_{t_N}^{(N)} - \hat{x}_{t_{2N}}^{(2N)})^{(i)}, i = 1, ..., M$. In our simulations, we chose $M = 10^5$ and $N \in \{2^1, ..., 2^{15}\}$. To cover a wide range of different Feller indices, we will perform numerical simulations with four different parameter sets. We always choose $T = 1$ and $S_0 = 100$. The other parameters can be found in Table 9.6. Model 1,3 and 4 are the same parameters as before (except for the time horizon) and we added Model 2 with a high Feller index. The estimates $\texttt{err}_v(N)$ and $\texttt{err}_x(N)$ for

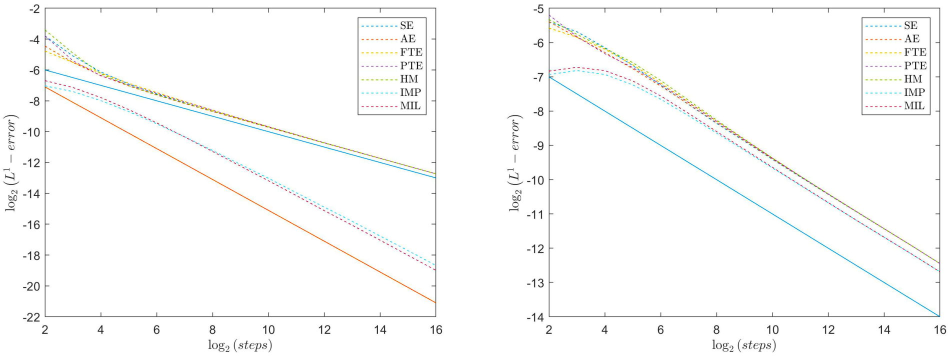
Model	V_0	κ	θ	σ	ρ	μ	ν
1	0.04	5	0.04	0.61	-0.7	0.0319	1.075
2	0.0457	5.07	0.0457	0.48	-0.767	0	2.0113
3	0.010201	6.21	0.019	0.61	-0.7	0.0319	0.63
4	0.09	2	0.09	1	-0.3	0.05	0.36

Table 9.6: Parameters for the strong convergence test.

the seven schemes are plotted in Figures 9.13–9.16 against the corresponding number of steps $2N$. For each model, we show first the convergence behavior of the error for the CIR process and then for the Heston model. Additionally, we plotted solid reference lines with suitable slopes together with the error estimates. Blue reference lines always have a slope of 0.5. We also estimated the order of convergence by the slope of a least squares fit, see Tables 9.7 – 9.10. Here, we only take errors with step sizes $N \in \{2^6, ..., 2^{15}\}$ into account to get a stable result. For all models, our simulation study shows that the numerical convergence orders of the Euler schemes do not change significantly if we extend the simulation from the CIR process to the Heston model. This indicates that the parameters of the CIR process (and especially the Feller index) solely determine the convergence behavior.

Scheme	Rate CIR	Rate Heston
SE	0.51	0.52
AE	0.51	0.52
FTE	0.52	0.53
PTE	0.52	0.52
HM	0.51	0.53
IMP	0.92	0.51
MIL	0.96	0.51

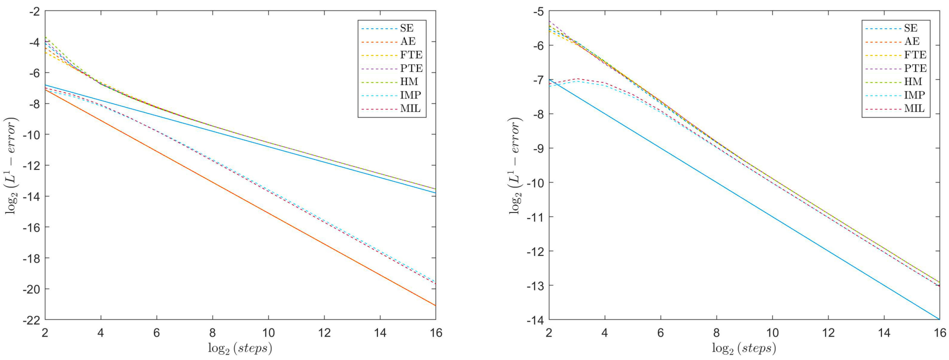
Table 9.7: Estimated strong convergence orders Model 1



(a) CIR, $\nu = 1.075$ (b) Heston, $\nu = 1.075$

Figure 9.13: Error estimates for Model 1

For the first model which has a Feller index around 1, our main result from Theorem 7.13 provides a strong convergence order of 0.5 for the Euler schemes. This can be numerically confirmed in Figure 9.13. The implicit Milstein (MIL) and the drift-implicit Euler (IMP) seem to converge with strong order 1 (which is the slope of the red solid line) which is optimal for the CIR process by the results from [34]. This indicates that our rate from Proposition 7.15 is not sharp and that the results from [2] for the latter might hold for values $\nu < 2$. However, this advantage vanishes for the full Heston model where all schemes seem to have the same strong convergence order of 0.5. This is in line with our result from Theorem 8.1.



(a) CIR, $\nu = 2.0113$ (b) Heston, $\nu = 2.0113$

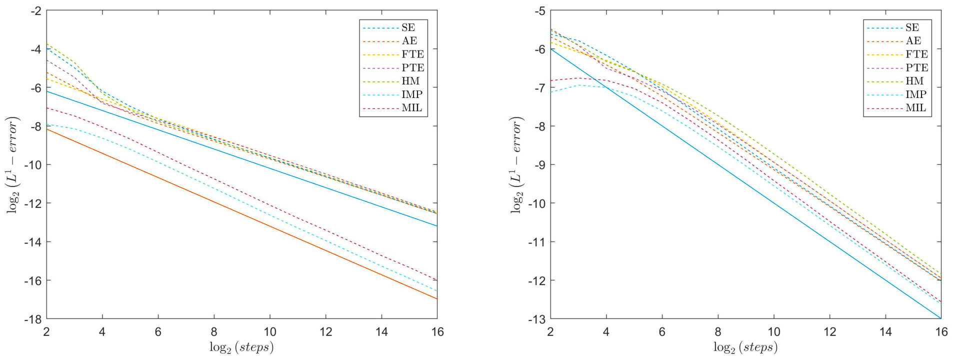
Figure 9.14: Error estimates for Model 2

For Model 2, which has a higher Feller index, Figure 9.14 and Table 9.8 confirm again the expected strong convergence order of 0.5 for the Euler case. Note that the differences

Scheme	Rate CIR	Rate Heston
SE	0.52	0.52
AE	0.52	0.52
FTE	0.52	0.52
PTE	0.52	0.52
HM	0.52	0.52
IMP	0.98	0.51
MIL	0.99	0.51

Table 9.8: Estimated strong convergence orders Model 2

between the different Euler schemes vanish for small step sizes and for high Feller indices. The Euler schemes only differ if the approximation of the CIR process becomes negative. For small step sizes and for high Feller indices this is unlikely to happen in a Monte Carlo simulation. Again, the implicit Milstein and the drift-implicit Euler seem to have a strong convergence order of 1 for the CIR case which decreases to 0.5 when applied to the full Heston model.



(a) CIR, $\nu = 0.63$ (b) Heston, $\nu = 0.63$

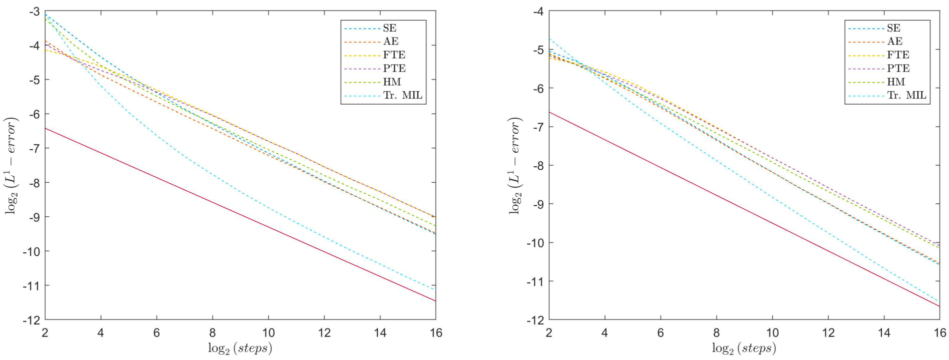
Figure 9.15: Error estimates for Model 3

Model 3 has a Feller index of 0.63 and we have shown that we can expect a strong convergence order of at least 0.315 for FTE, PTE and HM. However, looking at Table 9.9 and Figure 9.15 we can see that the rate for the Euler schemes is still around 0.5, even for SE and AE, for which we did not derive a convergence order in this case. For the IMP and MIL the convergence order dropped to a value around the Feller index itself. The red solid line in 9.15 on the left has now a slope of 0.63. This is in line with the lower bound result from [34] and indicates again that these schemes are optimal for the CIR process. For the full Heston model, all schemes seem to have a convergence order of 0.5.

The last model has the lowest Feller index. As in Section 9.1 we simulated the truncated

Scheme	Rate CIR	Rate Heston
SE	0.49	0.50
AE	0.47	0.48
FTE	0.49	0.50
PTE	0.48	0.49
HM	0.47	0.50
IMP	0.67	0.50
MIL	0.66	0.52

Table 9.9: Estimated strong convergence orders Model 3



(a) CIR, $\nu = 0.36$ (b) Heston, $\nu = 0.36$

Figure 9.16: Error estimates for Model 4

Scheme	Rate CIR	Rate Heston
SE	0.41	0.41
AE	0.38	0.40
FTE	0.37	0.39
PTE	0.37	0.38
HM	0.38	0.37
Tr. MIL	0.44	0.46

Table 9.10: Estimated strong convergence orders Model 4

Milstein alongside the Euler schemes since the Feller index is now below $\frac{1}{2}$. Again, we can see an estimated convergence order of the error for the Euler schemes that is better than expected. Here, we chose ν as the slope of the reference line in both plots from Figures 9.16. We already know that the truncated Milstein scheme is optimal in this parameter range (see Section 7.1).
The last two examples indicate that it might be possible to obtain a convergence order of $\min\{\nu, \frac{1}{2}\}$ for all Euler schemes for the CIR process and for the Heston model.
Our numerical simulations underline our result that, at least for $\nu > 1$, the L^1-convergence order of simple Euler schemes for the Heston model is already optimal. More advanced schemes can reach better error orders for the CIR process but their additional benefit in terms of the strong convergence order is not clear when applied to the full Heston model.

Chapter 10

Optimal L^2-Approximation of Stochastic Volatility Models

Now, we would like to move on to the analysis of more general models. In this chapter we study the strong approximation of the stochastic volatility model

$$\begin{aligned} dX_t &= \left(r - \frac{1}{2} f^2\left(V_t\right)\right) dt + f\left(V_t\right)\left(\rho dW_t + \sqrt{1-\rho^2} dB_t\right), \qquad && X_0 = x, \\ dV_t &= b\left(V_t\right) dt + \sigma\left(V_t\right) dW_t, && V_0 = v, \end{aligned} \tag{10.1}$$

where $V = (V_t)_{t\in[0,T]}$ takes values in an open set $D \subseteq \mathbb{R}$, $f, b, \sigma : D \to \mathbb{R}$ are appropriate functions, $\rho \in [-1, 1]$, $r \in \mathbb{R}$ and $W = (W_t)_{t\in[0,T]}$, $B = (B_t)_{t\in[0,T]}$ are independent Brownian motions. The initial values of the SDE are assumed to be deterministic and we have $x \in \mathbb{R}$, $v \in D$. The prototype example for SDE (10.1) is the generalized log-Heston model.
We analyze the minimal L^2-error for the approximation of X_T that can be obtained by arbitrary methods that use $N \in \mathbb{N}$ evaluations of each Brownian motion, that is

$$e(N) = \inf_{(s_i,t_i)_{i=1,\ldots,N}\in\Pi(N)} \inf_{u\in\mathcal{U}(N)} \left(\mathbb{E}\left[\left|u(W_{s_1}, \ldots, W_{s_N}, B_{t_1}, \ldots, B_{t_N}) - X_T\right|^2\right]\right)^{1/2}$$

where $\mathcal{U}(N)$ is the set of measurable functions $u : \mathbb{R}^{2N} \to \mathbb{R}$ and

$$\Pi(N) = \left\{(s_i, t_i)_{i=1,\ldots,N} : (s_i, t_i) \in [0, T]^2, i = 1, \ldots, N, s_N = t_N = T\right\}.$$

Our standing assumption is

Assumption 10.1. *The SDE* (10.1) *admits a unique strong solution and there exists a set $D = (l, r)$ with $-\infty \le l < r \le \infty$ and*

$$\mathbb{P}\left(V_t \in D, t \ge 0\right) = 1.$$

In the introduction of Chapter 8 we gave a brief overview of the extensive study on the optimal approximation of SDEs that has been carried out so far. In particular, if the coefficients of SDE (10.1) are Lipschitz continuous with Lipschitz continuous first derivative, then we have

$$\lim_{N\to\infty} \sqrt{N}\, e(N) \;=\; \sqrt{\frac{1-\rho^2}{4}} \int_0^T \left(\mathbb{E}\left[(f'\sigma)^2(V_t)\right]\right)^{1/2} dt,$$

from [58], where a result for more general multi-dimensional SDEs has been established. However, the coefficients of stochastic volatility models, as e.g. the log-Heston model, typically do not satisfy a global Lipschitz condition.

10.1 Lower Bound

For our first theorem, we need the following additional assumptions:

Assumption 10.2.

(a) We have $f \in C^2(D;\mathbb{R})$ and $\sigma \in C^1(D;(0,\infty))$.

(b) We have

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left|\left(f'b + \frac{1}{2}f''\sigma^2\right)(V_t)\right|^2\right] < \infty$$

and

$$\sup_{s,t\in[0,T]} \mathbb{E}\left[\frac{\left|(f'\sigma)(V_t) - (f'\sigma)(V_s)\right|^2}{|t-s|}\right] < \infty.$$

These assumptions are mainly needed to establish the Itō-Taylor expansion of the process $(f(V_t))_{t\in[0,T]}$ and to control the smoothness of the martingale part.

Theorem 10.3. *Let Assumptions 10.1 and 10.2 hold, let $\mathcal{U}(N)$ be the set of measurable functions $u : \mathbb{R}^{2N} \to \mathbb{R}$ and let*

$$e(N) = \inf_{(s_i,t_i)_{i=1,\dots,N}\in\Pi(N)} \inf_{u\in\mathcal{U}(N)} \left(\mathbb{E}\left[\left|u(W_{s_1},\dots,W_{s_N},B_{t_1},\dots,B_{t_N}) - X_T\right|^2\right]\right)^{1/2}.$$

Then, we have that

$$\liminf_{N\to\infty} \sqrt{N}\, e(N) \geq \sqrt{\frac{1-\rho^2}{6}} \int_0^T \left(\mathbb{E}\left[(f'\sigma)^2(V_t)\right]\right)^{1/2} dt.$$

10.2 Proof of Theorem 10.3

Before we start with the proof of Theorem 10.3, we need to introduce some notation. Recall that

$$\Pi(N) = \left\{(s_i,t_i)_{i=1,\ldots,N} : (s_i,t_i) \in [0,T]^2, i=1,\ldots,N,\, s_N = t_N = T\right\}.$$

We also introduce

$$\Pi_B(N) = \{(t_i)_{i=1,\ldots,N} : t_i \in [0,T], i=1,\ldots,N,\, t_N = T\}.$$

Let

$$\mathcal{G}_{\Pi(N)} = \sigma\left(W_{s_1}, W_{s_2}, \ldots, W_{s_N}, B_{t_1}, B_{t_2}, \ldots, B_{t_N}\right),$$
$$\mathcal{H}_{\Pi_B(N)} = \sigma\left(W_s, s \in [0,T], B_{t_1}, B_{t_2}, \ldots, B_{t_N}\right).$$

We also use the notation

$$\mathcal{H}_{\Pi_B(N)} = \sigma\left(W, B_{\Pi_B(N)}\right).$$

We set

$$\mathcal{Z}_{\mathcal{U}}^{\Pi(N)} = \{Z : \Omega \to \mathbb{R} : Z \text{ is } \mathcal{G}_{\Pi(N)} \text{ measurable}\},$$
$$\mathcal{Z}_{\mathcal{V}}^{\Pi_B(N)} = \{Z : \Omega \to \mathbb{R} : Z \text{ is } \mathcal{H}_{\Pi_B(N)} \text{ measurable}\}$$

and denote by $\mathcal{V}$ the set of all measurable functions $v : C([0,T];\mathbb{R}) \times \mathbb{R}^N \to \mathbb{R}$. Since we have

$$\mathcal{Z}_{\mathcal{U}}^{\Pi(N)} \subset \mathcal{Z}_{\mathcal{V}}^{\Pi_B(N)},$$

it follows that

$$\begin{aligned}
&\inf_{Z \in \mathcal{Z}_{\mathcal{U}}^{\Pi(N)}} \left(\mathbb{E}\left[|Z - X_T|^2\right]\right)^{1/2} \\
&= \inf_{u \in \mathcal{U}} \left(\mathbb{E}\left[|u(W_{s_1}, W_{s_2}, \ldots, W_{s_N}, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - X_T|^2\right]\right)^{1/2} \\
&\geq \inf_{Z \in \mathcal{Z}_{\mathcal{V}}^{\Pi_B(N)}} \left(\mathbb{E}\left[|Z - X_T|^2\right]\right)^{1/2} \\
&= \inf_{v \in \mathcal{V}} \left(\mathbb{E}\left[|v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - X_T|^2\right]\right)^{1/2}.
\end{aligned}$$

Thus, it is sufficient to analyze the quantity

$$\begin{aligned}
&\inf_{(t_i) \in \Pi_B(N)} \inf_{Z \in \mathcal{Z}_{\mathcal{V}}^{\Pi_B(N)}} \left(\mathbb{E}\left[|Z - X_T|^2\right]\right)^{1/2} \\
&\quad = \inf_{(t_i) \in \Pi_B(N)} \inf_{v \in \mathcal{V}} \left(\mathbb{E}\left[|v(W, B_{t_1}, B_{t_2}, \ldots, B_{t_N}) - X_T|^2\right]\right)^{1/2},
\end{aligned}$$

to obtain a lower bound for $e(N)$. Here and in the following we write (t_i) instead of $(t_i)_{i=1,\ldots,N}$.

10.2.1 Rewriting X_T and removing the measurable part

Recall that the SDE under consideration is given by

$$\begin{aligned} dX_t &= \left(r - \frac{1}{2} f^2 (V_t)\right) dt + f(V_t) \left(\rho dW_t + \sqrt{1-\rho^2} dB_t\right), \\ dV_t &= b(V_t)\, dt + \sigma(V_t)\, dW_t, \end{aligned} \qquad t \in [0,T].$$

Now we are going to rewrite X_T following [47].

Lemma 10.4. *Let Assumptions 10.1 and 10.2 hold. We have that*

$$X_T = U_T + \sqrt{1-\rho^2} \int_0^T f(V_t) dB_t$$

where

$$\begin{aligned} U_T = {} & x + \rho\left(F(V_T) - F(v)\right) + rT \\ & - \int_0^T \left(\frac{1}{2} f^2(V_t) + \rho \left(\frac{bf}{\sigma} + \frac{1}{2}\left(\sigma f' - \sigma' f\right)\right)(V_t)\right) dt. \end{aligned}$$

with

$$F(y) = \int_v^y \frac{f}{\sigma}(z) dz$$

for $y \in D$.

Proof. We apply Itō's formula to obtain

$$\begin{aligned} dF(V_t) &= \frac{f}{\sigma}(V_t) dV_t + \frac{1}{2}\left(\frac{\sigma f' - \sigma' f}{\sigma^2}\right)(V_t)\, d\langle V\rangle_t \\ &= \frac{fb}{\sigma}(V_t) dt + f(V_t) dW_t + \frac{1}{2}\left(\sigma f' - \sigma' f\right)(V_t)\, dt. \end{aligned}$$

Then, (X, V) solves

$$\begin{aligned} dX_t &= \rho dF(V_t) + h(V_t)\, dt + \sqrt{1-\rho^2} f(V_t)\, dB_t, \\ dV_t &= b(V_t)\, dt + \sigma(V_t)\, dW_t, \end{aligned} \qquad t \in [0,T],$$

where

$$h(y) := \left(r - \frac{1}{2} f^2(y)\right) - \rho\left(\frac{bf}{\sigma} + \frac{1}{2}\left(\sigma f' - \sigma' f\right)\right)(y).$$

Therefore, we can rewrite X_T as

$$X_T = U_T + \sqrt{1-\rho^2} \int_0^T f(V_t) dB_t$$

where

$$U_T = x + \rho\left(F(V_T) - F(v)\right) + rT - \int_0^T \left(\frac{1}{2} f^2(V_t) + \rho\left(\frac{bf}{\sigma} + \frac{1}{2}\left(\sigma f' - \sigma' f\right)\right)(V_t)\right) dt.$$

□

Expanding $f(V_t)$ using Itō's formula yields

$$f(V_t) = f(v) + \int_0^t \left(f'b + \frac{1}{2}f''\sigma^2\right)(V_s)ds + \int_0^t \left(f'\sigma\right)(V_s)dW_s, \qquad t \in [0,T].$$

Thus, we have

$$X_T = U_T + \sqrt{1-\rho^2}\left(f(v_0)B_T + \int_0^T A_t dB_t + \int_0^T Y_t dB_t\right)$$

with

$$A_t := \int_0^t \left(f'b + \frac{1}{2}f''\sigma^2\right)(V_s)ds, \quad Y_t := \int_0^t \left(f'\sigma\right)(V_s)dW_s, \qquad t \in [0,T].$$

It follows that

$$\begin{aligned}
&\inf_{Z \in \mathcal{Z}_{\mathcal{V}}^{\Pi_B(N)}} \left(\mathbb{E}\left[\left|Z - X_T\right|^2\right]\right)^{1/2} \\
&= \inf_{Z \in \mathcal{Z}_{\mathcal{V}}^{\Pi_B(N)}} \left(\mathbb{E}\left[\left|Z - U_T - \sqrt{1-\rho^2}\left(f(v)B_T + \int_0^T A_t dB_t + \int_0^T Y_t dB_t\right)\right|^2\right]\right)^{1/2} \\
&= \inf_{\tilde{Z} \in \mathcal{Z}_{\mathcal{V}}^{\Pi_B(N)}} \left(\mathbb{E}\left[\left|\tilde{Z} - \sqrt{1-\rho^2}\left(\int_0^T A_t dB_t + \int_0^T Y_t dB_t\right)\right|^2\right]\right)^{1/2}
\end{aligned}$$

and it remains to analyze

$$\inf_{(t_i) \in \Pi_B(N)} \inf_{Z \in \mathcal{Z}_{\mathcal{V}}^{\Pi_B(N)}} \left(\mathbb{E}\left[\left|Z - \int_0^T A_t dB_t - \int_0^T Y_t dB_t\right|^2\right]\right)^{1/2}.$$

10.2.2 Removing the smooth part

Since $A = (A_t)_{t\in[0,T]}$ is smooth enough, $\int_0^T A_t dB_t$ does not matter asymptotically for our approximation problem.

Lemma 10.5. *Let Assumptions 10.1 and 10.2 hold and let*

$$\Pi_B^\alpha(N) = \{\tau_0^\alpha, \tau_1^\alpha, ..., \tau_{\lceil N^\alpha \rceil}^\alpha\} = \left\{0, \frac{T}{\lceil N^\alpha \rceil}, \frac{2T}{\lceil N^\alpha \rceil}, ..., T\right\}, \qquad \alpha \in \left(\frac{1}{2}, 1\right).$$

There exists a constant $C > 0$ such that

$$\left(\mathbb{E}\left[\left|\int_0^T A_t dB_t - \sum_{k=0}^{\lceil N^\alpha \rceil - 1} A_{\tau_k^\alpha}(B_{\tau_{k+1}^\alpha} - B_{\tau_k^\alpha})\right|^2\right]\right)^{1/2} \leq C \cdot N^{-\alpha}.$$

Proof. First, we define $\eta^\alpha(t) := \max\{\tau_k^\alpha \in \Pi_B^\alpha(N) : \tau_k^\alpha \leq t\}$. Using the Itō-isometry we have

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_0^T A_t dB_t - \sum_{k=0}^{\lceil N^\alpha \rceil - 1} A_{\tau_k^\alpha}(B_{\tau_{k+1}^\alpha} - B_{\tau_k^\alpha})\right|^2\right] \\
&= \mathbb{E}\left[\left|\int_0^T (A_t - A_{\eta^\alpha(t)})dB_u\right|^2\right] \\
&= \int_0^T \mathbb{E}\left[\left|\int_{\eta^\alpha(t)}^t \left(f'b + \frac{1}{2}f''\sigma^2\right)(V_s)ds\right|^2\right] dt.
\end{aligned}$$

Moreover, the Cauchy-Schwartz inequality and Assumption 10.2 (b) yield

$$\begin{aligned}
&\mathbb{E}\left[\left|\int_{\eta^\alpha(t)}^t \left(f'b + \frac{1}{2}f''\sigma^2\right)(V_u)du\right|^2\right] \\
&\leq \left(\sup_{s\in[0,T]} \mathbb{E}\left[\left|\left(f'b + \frac{1}{2}f''\sigma^2\right)(V_s)\right|^2\right]\right)(t - \eta^\alpha(t))^2 \\
&\leq C \cdot N^{-2\alpha}
\end{aligned}$$

and so we have

$$\mathbb{E}\left[\left|\int_0^T A_t dB_t - \sum_{k=0}^{\lceil N^\alpha \rceil - 1} A_{\tau_k^\alpha}(B_{\tau_{k+1}^\alpha} - B_{\tau_k^\alpha})\right|^2\right] \leq C \cdot N^{-2\alpha}.$$

□

Next, we introduce

$$\mathcal{H}_{\Pi_B(N),\Pi_B^\alpha(N)} = \sigma\left(W, B_{\Pi_B(N)}, B_{\Pi_B^\alpha(N)}\right),$$

the set $\mathcal{V}_\alpha$ of all measurable functions $v : C([0,T];\mathbb{R}) \times \mathbb{R}^{N+\lceil N^\alpha \rceil} \to \mathbb{R}$ and

$$\mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)\cup\Pi_B^\alpha(N)} = \{Z : \Omega \to \mathbb{R} : Z \text{ is } \mathcal{H}_{\Pi_B(N),\Pi_B^\alpha(N)} \text{ measurable}\}.$$

Note that

$$\mathcal{Z}_{\mathcal{V}}^{\Pi_B(N)} \subset \mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)\cup\Pi_B^\alpha(N)}.$$

Lemma 10.6. *Let Assumptions 10.1 and 10.2 hold. We have*

$$\begin{aligned}
&\liminf_{N\to\infty} \sqrt{N}e(N) \\
&\geq \liminf_{N\to\infty} \sqrt{N}\sqrt{1-\rho^2} \inf_{(t_i)\in\Pi_B(N)} \inf_{Z\in\mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)\cup\Pi_B^\alpha(N)}} \left(\mathbb{E}\left[\left|Z - \int_0^T Y_t dB_t\right|^2\right]\right)^{1/2}.
\end{aligned}$$

Proof. Since $\sum_{k=0}^{\lceil N^\alpha \rceil -1} A_{\tau_k^\alpha}(B_{\tau_{k+1}^\alpha} - B_{\tau_k^\alpha})$ is $\mathcal{H}_{\Pi_B(N),\Pi_B^\alpha(N)}$-measurable, we have

$$\begin{aligned}
&\inf_{Z \in \mathcal{Z}_{\mathcal{V}}^{\Pi_B(N)}} \mathbb{E}\left[\left|Z - \int_0^T A_t dB_t - \int_0^T Y_t dB_t\right|^2\right]^{1/2} \\
&\geq \inf_{Z \in \mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)\cup\Pi_B^\alpha(N)}} \mathbb{E}\left[\left|Z - \int_0^T A_t dB_t - \int_0^T Y_t dB_t\right|^2\right]^{1/2} \\
&= \inf_{\tilde{Z} \in \mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)\cup\Pi_B^\alpha(N)}} \mathbb{E}\left[\left|\tilde{Z} - \int_0^T (A_t - A_{\eta^\alpha(t)}) dB_t - \int_0^T Y_t dB_t\right|^2\right]^{1/2} \\
&\geq \inf_{Z \in \mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)\cup\Pi_B^\alpha(N)}} \mathbb{E}\left[\left|Z - \int_0^T Y_t dB_t\right|^2\right]^{1/2} - \mathbb{E}\left[\left|\int_0^T (A_t - A_{\eta^\alpha(t)}) dB_t\right|^2\right]^{1/2},
\end{aligned}$$

where we applied the Minkowski inequality in the last step. Therefore, using Lemma 10.5 we have

$$\begin{aligned}
&\inf_{(t_i)\in\Pi_B(N)} \inf_{Z \in \mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)}} \mathbb{E}\left[\left|Z - \int_0^T A_t dB_t - \int_0^T Y_t dB_t\right|^2\right]^{1/2} \\
&\geq \inf_{(t_i)\in\Pi_B(N)} \inf_{Z \in \mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)\cup\Pi_B^\alpha(N)}} \mathbb{E}\left[\left|Z - \int_0^T Y_t dB_t\right|^2\right]^{1/2} - O\left(N^{-\alpha}\right).
\end{aligned}$$

□

So, we have reduced our initial problem to the study of

$$\inf_{(t_i)\in\Pi_B(N)} \inf_{Z \in \mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)\cup\Pi_B^\alpha(N)}} \mathbb{E}\left[\left|Z - \int_0^T Y_t dB_t\right|^2\right]^{1/2}. \tag{10.2}$$

In the following, we denote points from $\Pi_B(N) \cup \Pi_B^\alpha(N)$ by t_k^α and we assume without loss of generality that these points are ordered, i.e.

$$\Pi_B(N) \cup \Pi_B^\alpha(N) = \left\{0 = t_0^\alpha \leq t_1^\alpha \leq \ldots \leq t_{m(N)}^\alpha = T\right\}$$

with $m(N) = N + \lceil N^\alpha \rceil$.

10.2.3 Inserting Brownian bridges and symmetrization

Now we apply a symmetrization argument similar to [46, 61, 73]. To analyze (10.2) let us first denote the piecewise linear interpolation of B on the grid $t_0^\alpha, ..., t_{m(N)}^\alpha$ by $\overline{B}$, i.e. $\overline{B}$ is defined as

$$\overline{B}_t = B_{t_k^\alpha} + \frac{t - t_k^\alpha}{t_{k+1}^\alpha - t_k^\alpha}(B_{t_{k+1}^\alpha} - B_{t_k^\alpha}), \qquad t \in [t_k^\alpha, t_{k+1}^\alpha],\ k = 0, \ldots, m(N) - 1.$$

Then, the process B° given by

$$B_t^\circ = B_t - \overline{B}_t, \qquad t \in [0,T],$$

is a Brownian bridge on $[t_k^\alpha, t_{k+1}^\alpha]$ for $k = 0, \dots, m(N)-1$. Moreover, the processes

$$(B_t^\circ)_{t\in[t_0^\alpha,t_1^\alpha]},\ (B_t^\circ)_{t\in[t_1^\alpha,t_2^\alpha]},\ \dots,\ (B_t^\circ)_{t\in[t_{m(N)-1}^\alpha,t_{m(N)}^\alpha]},\ \overline{B},\ W$$

are independent. Since

$$\int_0^T Y_t d\overline{B}_t = \sum_{k=0}^{m(N)-1} \frac{1}{t_{k+1}^\alpha - t_k^\alpha} \int_{t_k^\alpha}^{t_{k+1}^\alpha} Y_t dt \, (B_{t_{k+1}^\alpha} - B_{t_k^\alpha})$$

is $\mathcal{H}_{\Pi_B(N),\Pi_B^\alpha(N)}$-measurable, we have that

$$\begin{aligned} &\inf_{Z\in\mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)\cup\Pi_B^\alpha(N)}} \left(\mathbb{E}\left[\left|Z - \int_0^T Y_t dB_t\right|^2\right]\right)^{1/2} \\ &= \inf_{\tilde{Z}\in\mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N)\cup\Pi_B^\alpha(N)}} \left(\mathbb{E}\left[\left|\tilde{Z} - \mathcal{I}_N(Y,B^\circ)\right|^2\right]\right)^{1/2} \end{aligned} \tag{10.3}$$

with

$$\mathcal{I}_N\left(Y,B^\circ\right) := \int_0^T Y_s dB_s - \int_0^T Y_s d\overline{B}_s.$$

Furthermore B° and $-B^\circ$ have the same law, so the independence of B° from $(W,\overline{B})$ implies that

$$\left(W,\overline{B},B^\circ\right) \overset{d}{=} \left(W,\overline{B},-B^\circ\right). \tag{10.4}$$

We would now like to analyze $\mathcal{I}_N\left(Y,B^\circ\right)$. Recall that $(Y_s)_{s\in[0,T]}$ is independent of $(B_s)_{s\in[0,T]}$. We define

$$\Delta_k^\alpha = t_{k+1}^\alpha - t_k^\alpha, \qquad \Delta_k^\alpha B = B_{t_{k+1}^\alpha} - B_{t_k^\alpha},$$

for $k = 0, \dots, m(N)-1$ and

$$\tau_{l,n,k}^\alpha = \frac{l}{2^n}\Delta_k^\alpha, \qquad \Delta_{k,l}^\alpha = \tau_{l+1,n,k}^\alpha - \tau_{l,n,k}^\alpha, \qquad \Delta_{k,l}^\alpha B = B_{t_k^\alpha+\tau_{l+1,n,k}^\alpha} - B_{t_k^\alpha+\tau_{l,n,k}^\alpha},$$

for $l = 0, \dots, 2^n$, $k = 0, \dots, m(N)-1$.

Lemma 10.7. *Let Assumptions 10.1 and 10.2 hold.*

(i) Let

$$\mathcal{I}_N^n(Y,B^\circ) = \sum_{k=0}^{m(N)-1} \sum_{l=0}^{2^n-1} Y_{t_k^\alpha+\tau_{l,n,k}^\alpha} \left(\Delta_{k,l}^\alpha B - \frac{\Delta_{k,l}^\alpha}{\Delta_k^\alpha}\Delta_k^\alpha B\right).$$

We have that

$$\mathcal{I}_N(Y,B^\circ) = \lim_{n\to\infty} \mathcal{I}_N^n(Y,B^\circ)$$

almost surely and in L^2.

(ii) It holds that

$$\mathcal{I}_N(Y, B^\circ) \stackrel{d}{=} Z \left(\int_0^T |\phi_s^{(t_i),\alpha}|^2 ds \right)^{1/2}$$

with

$$\phi_s^{(t_i),\alpha} := Y_s - \sum_{k=0}^{m(N)-1} \left(\frac{1}{\Delta_k^\alpha} \int_{t_k^\alpha}^{t_{k+1}^\alpha} Y_u du \right) \mathbb{1}_{[t_k^\alpha, t_{k+1}^\alpha)}(s)$$

and

$$Z \sim \mathcal{N}(0,1).$$

Proof. To simplify the notation we drop all α-superscripts in this proof. Note that Assumption 10.2 (b) implies that

$$\sup_{s,t\in[0,T]} \mathbb{E}\left[\frac{|Y_t - Y_s|^2}{|t-s|} \right] < \infty$$

and

$$\sup_{t\in[0,T]} \mathbb{E}\left[|Y_t|^2 \right] < \infty.$$

(i) We have

$$\int_{t_k}^{t_{k+1}} Y_s dB_s - \int_{t_k}^{t_{k+1}} Y_s d\overline{B}_s = I_1^k - I_2^k$$

with

$$I_1^k = \int_{t_k}^{t_{k+1}} Y_s dB_s, \qquad I_2^k = \int_{t_k}^{t_{k+1}} Y_s d\overline{B}_s$$

and

$$\sum_{l=0}^{2^n-1} Y_{t_k+\tau_{l,n,k}} \left(\Delta_{k,l} B - \frac{\Delta_{k,l}}{\Delta_k} \Delta_k B \right) = I_1^{k,n} - I_2^{k,n}$$

where

$$I_1^{k,n} = \sum_{l=0}^{2^n-1} Y_{t_k+\tau_{l,n,k}} \Delta_{k,l} B, \qquad I_2^{k,n} = \sum_{l=0}^{2^n-1} \frac{\Delta_{k,l}}{\Delta_k} Y_{t_k+\tau_{l,n,k}} \Delta_k B.$$

For brevity, we write $Y_{t_k+\tau_{l,n,k}} = Y_{k,l}$. Using the Itō-isometry, polarization and the smoothness of Y, it follows

$$\begin{aligned}
\mathbb{E}\left[\left|I_1^k - I_1^{k,n}\right|^2\right] &= \mathbb{E}\left[\left|\int_{t_k}^{t_{k+1}} Y_s dB_s - \sum_{l=0}^{2^n-1} Y_{k,l}\Delta_{k,l}B\right|^2\right] \\
&= \mathbb{E}\left[\left|\sum_{l=0}^{2^n-1}\int_{t_k+\tau_{l,n,k}}^{t_k+\tau_{l+1,n,k}} (Y_s - Y_{k,l})dB_s\right|^2\right] \\
&= \mathbb{E}\left[\sum_{l=0}^{2^n-1}\left|\int_{t_k+\tau_{l,n,k}}^{t_k+\tau_{l+1,n,k}} (Y_s - Y_{k,l})dB_s\right|^2\right] \\
&= \sum_{l=0}^{2^n-1}\int_{t_k+\tau_{l,n,k}}^{t_k+\tau_{l+1,n,k}} \mathbb{E}\left[|Y_s - Y_{k,l}|^2\right] ds \\
&\leq C\sum_{l=0}^{2^n-1}\Delta_{k,l}^2 \leq C\Delta_k^2 2^{-n}.
\end{aligned}$$

Furthermore, we have

$$\begin{aligned}
\mathbb{E}\left[\left|I_2^k - I_2^{k,n}\right|^2\right] &= \mathbb{E}\left[\left|\int_{t_k}^{t_{k+1}} Y_s d\overline{B}_s - \sum_{l=0}^{2^n-1}\frac{\Delta_{k,l}}{\Delta_k} Y_{k,l}\Delta_k B\right|^2\right] \\
&= \mathbb{E}\left[\left|\sum_{l=0}^{2^n-1}\left(\frac{1}{\Delta_k}\int_{t_k+\tau_{l,n,k}}^{t_k+\tau_{l+1,n,k}} Y_s ds\Delta_k B - \frac{\Delta_{k,l}}{\Delta_k}Y_{k,l}\Delta_k B\right)\right|^2\right] \\
&= \mathbb{E}\left[\left|\sum_{l=0}^{2^n-1}\frac{1}{\Delta_k}\int_{t_k+\tau_{l,n,k}}^{t_k+\tau_{l+1,n,k}} \left(Y_s - Y_{k,l}\right) ds\Delta_k B\right|^2\right].
\end{aligned}$$

With similar computations as before, we also obtain

$$\mathbb{E}\left[\left|I_2^k - I_2^{k,n}\right|^2\right] \leq C\Delta_k^2 2^{-n}.$$

Now,

$$\begin{aligned}
\mathbb{E}\left[\left|\mathcal{I}_N(Y,B^\circ) - \mathcal{I}_N^n(Y,B^\circ)\right|^2\right] &= \mathbb{E}\left[\left|\sum_{k=0}^{m(N)-1}\left(I_1^k - I_1^{k,n} + I_2^{k,n} - I_2^k\right)\right|^2\right] \\
&\leq 2\sum_{k=0}^{m(N)-1}\mathbb{E}\left[\left|I_1^k - I_1^{k,n}\right|^2\right] + 2\sum_{k=0}^{m(N)-1}\mathbb{E}\left[\left|I_2^k - I_2^{k,n}\right|^2\right] \\
&\leq C\sum_{k=0}^{m(N)-1}\Delta_k^2 2^{-n} \leq C2^{-n},
\end{aligned}$$

which yields the L^2-convergence, and also implies

$$\sum_{n=1}^{\infty} \mathbb{E}\left[|\mathcal{I}_N(Y,B^\circ) - \mathcal{I}_N^n(Y,B^\circ)|\right] < \infty,$$

from which the almost sure convergence follows by an application of the Borel-Cantelli lemma.
(ii) Recall that Y is independent of B°. The conditional law of $\mathcal{I}_N^n(Y,B^\circ)$ given

$$Y_{t_k+\tau_{l,n,k}} = y_{k,l}, \quad l = 0,\ldots,2^n-1, k = 0,\ldots,m(N)-1,$$

is therefore Gaussian with zero mean and variance

$$\begin{aligned}
&\mathbb{E}\left[\left(\sum_{k=0}^{m(N)-1}\sum_{l=0}^{2^n-1} y_{k,l}\left(\Delta_{k,l}B - \frac{\Delta_{k,l}}{\Delta_k}\Delta_k B\right)\right)^2\right] \\
&= \mathbb{E}\left[\sum_{k=0}^{m(N)-1}\left(\sum_{l=0}^{2^n-1} y_{k,l}\left(\Delta_{k,l}B - \frac{\Delta_{k,l}}{\Delta_k}\Delta_k B\right)\right)^2\right] \\
&\quad + 2\mathbb{E}\left[\sum_{j<k=0}^{m(N)-1}\left(\sum_{l=0}^{2^n-1} y_{k,l}\left(\Delta_{k,l}B - \frac{\Delta_{k,l}}{\Delta_k}\Delta_k B\right)\right)\left(\sum_{l=0}^{2^n-1} y_{j,l}\left(\Delta_{j,l}B - \frac{\Delta_{j,l}}{\Delta_j}\Delta_j B\right)\right)\right].
\end{aligned}$$

The Brownian increments in the second term are from disjoint intervals. Therefore, its expectation is zero. Moreover, we have

$$\begin{aligned}
&\mathbb{E}\left[\sum_{k=0}^{m(N)-1}\left(\sum_{l=0}^{2^n-1} y_{k,l}\left(\Delta_{k,l}B - \frac{\Delta_{k,l}}{\Delta_k}\Delta_k B\right)\right)^2\right] \\
&= \sum_{k=0}^{m(N)-1}\mathbb{E}\left[\sum_{l=0}^{2^n-1}\left(y_{k,l}\left(\Delta_{k,l}B - \frac{\Delta_{k,l}}{\Delta_k}\Delta_k B\right)\right)^2\right] \\
&\quad + 2\sum_{k=0}^{m(N)-1}\mathbb{E}\left[\sum_{j<l=0}^{2^n-1}\left(y_{k,l}\left(\Delta_{k,l}B - \frac{\Delta_{k,l}}{\Delta_k}\Delta_k B\right)\right)\left(y_{k,j}\left(\Delta_{k,j}B - \frac{\Delta_{k,j}}{\Delta_k}\Delta_k B\right)\right)\right].
\end{aligned}$$

Looking at the first term, we obtain

$$\begin{aligned}
&\sum_{k=0}^{m(N)-1}\mathbb{E}\left[\sum_{l=0}^{2^n-1}\left(y_{k,l}\left(\Delta_{k,l}B - \frac{\Delta_{k,l}}{\Delta_k}\Delta_k B\right)\right)^2\right] \\
&= \sum_{k=0}^{m(N)-1}\sum_{l=0}^{2^n-1} y_{k,l}^2\mathbb{E}\left[\left(\Delta_{k,l}B - \frac{\Delta_{k,l}}{\Delta_k}\Delta_k B\right)^2\right] \\
&= \sum_{k=0}^{m(N)-1}\sum_{l=0}^{2^n-1} y_{k,l}^2\left(\Delta_{k,l} - \frac{\Delta_{k,l}^2}{\Delta_k}\right).
\end{aligned}$$

The second term yields

$$
\begin{aligned}
&2\sum_{k=0}^{m(N)-1}\mathbb{E}\left[\sum_{j<l=0}^{2^n-1}\left(y_{k,l}\left(\Delta_{k,l}B-\frac{\Delta_{k,l}}{\Delta_k}\Delta_k B\right)\right)\left(y_{k,j}\left(\Delta_{k,j}B-\frac{\Delta_{k,j}}{\Delta_k}\Delta_k B\right)\right)\right]\\
&=2\sum_{k=0}^{m(N)-1}\sum_{j<l=0}^{2^n-1}-\frac{y_{k,l}y_{k,j}\Delta_{k,l}\Delta_{k,j}}{\Delta_k}.
\end{aligned}
$$

Summarizing, we have

$$
\begin{aligned}
&\mathbb{E}\left[\left(\sum_{k=0}^{m(N)-1}\sum_{l=0}^{2^n-1}y_{k,l}\left(\Delta_{k,l}B-\frac{\Delta_{k,l}}{\Delta_k}\Delta_k B\right)\right)^2\right]\\
&\quad=\sum_{k=0}^{m(N)-1}\left(\sum_{l=0}^{2^n-1}y_{k,l}^2\left(\Delta_{k,l}-\frac{\Delta_{k,l}^2}{\Delta_k}\right)-2\sum_{j<l=0}^{2^n-1}\frac{y_{k,l}y_{k,j}\Delta_{k,l}\Delta_{k,j}}{\Delta_k}\right)\\
&\quad=\sum_{k=0}^{m(N)-1}\left(\sum_{l=0}^{2^n-1}y_{k,l}^2\Delta_{k,l}\right)-\frac{1}{\Delta_k}\left(\sum_{l=0}^{2^n-1}y_{k,l}\Delta_{k,l}\right)^2.
\end{aligned}
$$

This in turns implies that

$$
\mathcal{I}_N^n(Y,B^\circ)\overset{d}{=}Z\left(\sum_{k=0}^{m(N)-1}\left(\sum_{l=0}^{2^n-1}Y_{k,l}^2\Delta_{k,l}\right)-\frac{1}{\Delta_k}\left(\sum_{l=0}^{2^n-1}Y_{k,l}\Delta_{k,l}\right)^2\right)^{1/2}
$$

with

$$
Z\sim\mathcal{N}(0,1).
$$

Now, note that

$$
\begin{aligned}
&\lim_{n\to\infty}\sum_{k=0}^{m(N)-1}\left(\sum_{l=0}^{2^n-1}Y_{k,l}^2\left(\Delta_{k,l}-\frac{\Delta_{k,l}^2}{\Delta_k}\right)-2\sum_{j<l=0}^{2^n-1}\frac{Y_{k,l}Y_{k,j}\Delta_{k,l}\Delta_{k,j}}{\Delta_k}\right)\\
&\quad=\sum_{k=0}^{m(N)-1}\left(\left(\int_{t_k}^{t_{k+1}}Y_s^2ds\right)-\frac{1}{\Delta_k}\left(\int_{t_k}^{t_{k+1}}Y_sds\right)^2\right)
\end{aligned}
$$

almost surely by continuity of almost all sample paths of Y. Defining

$$
\phi_s^{(t_i)}:=Y_s-\sum_{k=0}^{m(N)-1}\left(\frac{1}{\Delta_k}\int_{t_k}^{t_{k+1}}Y_udu\right)\mathbb{1}_{[t_k,t_{k+1})}(s),\qquad s\in[0,T],
$$

we can see that

$$\int_{t_k}^{t_{k+1}} |\phi_s^{(t_i)}|^2 ds = \int_{t_k}^{t_{k+1}} Y_s^2 ds - \frac{1}{\Delta_k}\left(\int_{t_k}^{t_{k+1}} Y_s ds\right)^2.$$

Therefore, the assertion follows now from part (i). □

Using (10.3), a symmetrization argument based on (10.4) and the Lemma 10.7, we obtain the following results.

Lemma 10.8. *Let Assumptions 10.1 and 10.2 hold. Then, we have*

$$\inf_{Z \in \mathcal{Z}_{\mathcal{V}_\alpha}^{\Pi_B(N) \cup \Pi^\alpha B(N)}} \left(\mathbb{E}\left[|Z - \mathcal{I}_N(Y, B^\circ)|^2\right]\right)^{1/2} \geq \left(\mathbb{E}\left[\int_0^T |\phi_s^{(t_i),\alpha}|^2 ds\right]\right)^{1/2}$$

with

$$\phi_s^{(t_i),\alpha} := Y_s - \sum_{k=0}^{m(N)-1} \left(\frac{1}{\Delta_k^\alpha}\int_{t_k^\alpha}^{t_{k+1}^\alpha} Y_u du\right) \mathbb{1}_{[t_k^\alpha, t_{k+1}^\alpha)}(s), \qquad s \in [0, T].$$

Proof. Again we drop all α-superscripts in this proof. The equality of the laws in (10.4) yields that

$$\left(W, \overline{B}, \mathcal{I}_N^n(Y, B^\circ)\right) \stackrel{d}{=} \left(W, \overline{B}, -\mathcal{I}_N^n(Y, B^\circ)\right).$$

and (i) from the Lemma 10.7 now gives

$$\left(W, \overline{B}, \mathcal{I}_N(Y, B^\circ)\right) \stackrel{d}{=} \left(W, \overline{B}, -\mathcal{I}_N(Y, B^\circ)\right).$$

Consequently, it follows that

$$\begin{aligned}
&\left(\mathbb{E}\left[\left|v(W, B_{t_0}, B_{t_1}, B_{t_2}, \ldots, B_{t_{m(N)}}) - \mathcal{I}_N(Y, B^\circ)\right|^2\right]\right)^{1/2} \\
&\qquad = \left(\mathbb{E}\left[\left|v(W, B_{t_0}, B_{t_1}, B_{t_2}, \ldots, B_{t_{m(N)}}) + \mathcal{I}_N(Y, B^\circ)\right|^2\right]\right)^{1/2}
\end{aligned}$$

and so

$$\begin{aligned}
&2\left(\mathbb{E}\left[|\mathcal{I}_N(Y, B^\circ)|^2\right]\right)^{1/2} \\
&= \left(\mathbb{E}\left[|\mathcal{I}_N(Y, B^\circ) + \mathcal{I}_N(Y, B^\circ)|^2\right]\right)^{1/2} \\
&= \left(\mathbb{E}\left[\left|\left(\mathcal{I}_N(Y, B^\circ) - v(W, B_{t_0}, B_{t_1}, \ldots, B_{t_{m(N)}})\right)\right.\right.\right. \\
&\qquad\qquad \left.\left.\left. + \left(v(W, B_{t_0}, B_{t_1}, \ldots, B_{t_{m(N)}}) + \mathcal{I}_N(Y, B^\circ)\right)\right|^2\right]\right)^{1/2} \\
&\leq 2\left(\mathbb{E}\left[\left|v(W, B_{t_0}, B_{t_1}, B_{t_2}, \ldots, B_{t_{m(N)}}) - \mathcal{I}_N(Y, B^\circ)\right|^2\right]\right)^{1/2}
\end{aligned}$$

by the Minkowski inequality for any $v \in \mathcal{V}_\alpha$. It follows that

$$\inf_{v \in \mathcal{V}_\alpha} \left(\mathbb{E}\left[\left| v(W, B_{t_0}, B_{t_1}, B_{t_2}, \dots, B_{t_{m(N)}}) - \mathcal{I}_N(Y, B^\circ) \right|^2 \right] \right)^{1/2} \\ \geq \left(\mathbb{E}\left[|\mathcal{I}_N(Y, B^\circ)|^2 \right] \right)^{1/2}.$$

Using the Lemma 10.7 we then have

$$\inf_{v \in \mathcal{V}_\alpha} \left(\mathbb{E}\left[\left| v(W, B_{t_0}, B_{t_1}, B_{t_2}, \dots, B_{t_{m(N)}}) - \mathcal{I}_N(Y, B^\circ) \right|^2 \right] \right)^{1/2} \\ \geq \left(\mathbb{E}\left[\int_0^T |\phi_s^{(t_i)}|^2 ds \right] \right)^{1/2}$$

since $\mathbb{E}\left[|Z|^2\right] = \sigma^2$ for $Z \sim \mathcal{N}(0, \sigma^2)$. The assertion now follows from minimizing over all possible discretizations. □

10.2.4 Conclusion

Combining Lemma 10.6, Equation (10.3) and Lemma 10.8 we have shown that

$$\liminf_{N \to \infty} \sqrt{N} e(N) \geq \liminf_{N \to \infty} \sqrt{N} \sqrt{1 - \rho^2} \inf_{(t_i) \in \Pi_B(N)} \left(\mathbb{E}\left[\int_0^T |\phi_s^{(t_i),\alpha}|^2 ds \right] \right)^{1/2}. \quad (10.5)$$

The last part of the proof requires the following two auxiliary results:

Lemma 10.9. *Let Assumptions 10.1 and 10.2 hold. Then, the function*

$$\varphi : [0, T] \to \mathbb{R}, \qquad \varphi(t) = \left(\mathbb{E}\left[(f'\sigma)^2(V_t) \right] \right)^{1/2}$$

satisfies

$$\sup_{s,t \in [0,T]} \frac{|\varphi^2(t) - \varphi^2(s)|}{|t - s|^{1/2}} < \infty \quad (10.6)$$

and

$$\sup_{s,t \in [0,T]} \frac{|\varphi(t) - \varphi(s)|}{|t - s|^{1/4}} < \infty. \quad (10.7)$$

Proof. We have

$$\left| \mathbb{E}\left[(f'\sigma)^2(V_t) \right] - \mathbb{E}\left[(f'\sigma)^2(V_s) \right] \right| \\ = \left| \mathbb{E}\left[\left((f'\sigma)(V_t) + (f'\sigma)(V_s) \right) \left((f'\sigma)(V_t) - (f'\sigma)(V_s) \right) \right] \right|$$

and so equation (10.6) follows from Assumption 10.2 (b) and the Hölder inequality. Equation (10.7) is a consequence of

$$|\varphi(t) - \varphi(s)| = \left| \sqrt{\varphi^2(t)} - \sqrt{\varphi^2(s)} \right| \leq \sqrt{|\varphi^2(t) - \varphi^2(s)|}.$$

□

Lemma 10.10. *Let Assumptions 10.1 and 10.2 hold and let $\alpha \in (2/3, 1)$. There exists a constant $C_\alpha > 0$, such that*

$$\sup_{(t_i)\in\Pi_B(N)} \left| \mathbb{E}\left[\int_0^T |\phi_s^{(t_i),\alpha}|^2 ds - \frac{1}{6} \sum_{k=0}^{m(N)-1} \varphi^2(t_k^\alpha)(t_{k+1}^\alpha - t_k^\alpha)^2 \right] \right| \leq C_\alpha \cdot N^{-3\alpha/2},$$

where

$$\varphi^2(t) = \mathbb{E}\left[(f'\sigma)^2(V_t) \right], \qquad t \in [0,T].$$

Proof. As before, we drop the superscript α. By the definition of $\phi^{(t_i)}$, we have

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T |\phi_s^{(t_i)}|^2 dt \right] \\
&= \mathbb{E}\left[\int_0^T \left(Y_s - \sum_{k=0}^{m(N)-1} \frac{1}{\Delta_k} \int_{t_k}^{t_{k+1}} Y_u du \mathbb{1}_{[t_k,t_{k+1})}(s) \right)^2 ds \right] \\
&= \mathbb{E}\left[\int_0^T \left(\sum_{k=0}^{m(N)-1} \frac{1}{\Delta_k} \int_{t_k}^{t_{k+1}} (Y_s - Y_u) du \mathbb{1}_{[t_k,t_{k+1})}(s) \right)^2 ds \right] \\
&= \mathbb{E}\left[\int_0^T \left(\sum_{k=0}^{m(N)-1} \frac{1}{\Delta_k} \int_{t_k}^{t_{k+1}} (Y_s - Y_{u_1})\, du_1 \mathbb{1}_{[t_k,t_{k+1})}(s) \right) \right. \\
&\qquad \left. \cdot \left(\sum_{l=0}^{m(N)-1} \frac{1}{\Delta_l} \int_{t_l}^{t_{l+1}} (Y_s - Y_{u_2})\, du_2 \mathbb{1}_{[t_l,t_{l+1})}(s) \right) ds \right] \\
&= \sum_{k=0}^{m(N)-1} \frac{1}{\Delta_k^2} \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} \mathbb{E}\left[(Y_s - Y_{u_1})(Y_s - Y_{u_2}) \right] du_1 du_2 ds.
\end{aligned}$$

Recalling the definition of Y, we have for $s, u \in [t_k, t_{k+1}]$ that

$$Y_s - Y_u = \int_0^s \left(f'\sigma \right)(V_r) dW_r - \int_0^u \left(f'\sigma \right)(V_r) dW_r.$$

To calculate the value of $\mathbb{E}\left[(Y_s - Y_{u_1})(Y_s - Y_{u_2}) \right]$, we first observe that

$$\begin{aligned}
&\int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} \mathbb{E}\left[(Y_s - Y_{u_1})(Y_s - Y_{u_2}) \right] du_1 du_2 ds \\
&\qquad = 2 \int_{t_k}^{t_{k+1}} \int_{t_k}^{t_{k+1}} \int_{t_k}^{u_2} \mathbb{E}\left[(Y_s - Y_{u_1})(Y_s - Y_{u_2}) \right] du_1 du_2 ds.
\end{aligned}$$

Now, we consider three cases.

Case 1: $s \leq u_1 \leq u_2$. Using the Itō-isometry and polarization we have

$$\begin{aligned}
\mathbb{E}\left[(Y_s - Y_{u_1})(Y_s - Y_{u_2})\right] &= \mathbb{E}\left[\left(-\int_s^{u_1} (f'\sigma)(V_r)dW_r\right)\left(-\int_s^{u_2} (f'\sigma)(V_r)dW_r\right)\right] \\
&= \mathbb{E}\left[\int_s^{u_1} (f'\sigma)^2 (V_r)dr\right] \\
&= \mathbb{E}\left[(f'\sigma)^2 (V_{t_k})\right](u_1 - s) + r_k^{(1)}(s, u_1)
\end{aligned}$$

with

$$r_k^{(1)}(s, u_1) = \mathbb{E}\left[\int_s^{u_1} \left((f'\sigma)^2 (V_r) - (f'\sigma)^2 (V_{t_k})\right) dr\right]$$

and

$$\left|r_k^{(1)}(s, u_1)\right| \leq C \cdot \Delta_k^{3/2}$$

since $s, u_1 \in [t_k, t_{k+1}]$.

Case 2: $u_1 \leq s \leq u_2$. Here it follows

$$\mathbb{E}\left[(Y_s - Y_{u_1})(Y_s - Y_{u_2})\right] = \mathbb{E}\left[\left(\int_{u_1}^{s} (f'\sigma)(V_r)dW_r\right)\left(-\int_s^{u_2} (f'\sigma)(V_r)dW_r\right)\right] = 0.$$

Case 3: $u_1 \leq u_2 \leq s$. Similar to Case 1 it follows that

$$\begin{aligned}
\mathbb{E}\left[(Y_s - Y_{u_1})(Y_s - Y_{u_2})\right] &= \mathbb{E}\left[\left(\int_{u_1}^{s} (f'\sigma)(V_r)dW_r\right)\left(\int_{u_2}^{s} (f'\sigma)(V_r)dW_r\right)\right] \\
&= \mathbb{E}\left[\int_{u_2}^{s} (f'\sigma)^2 (V_r)dr\right] \\
&= \mathbb{E}\left[(f'\sigma)^2 (V_{t_k})\right](s - u_2) + r_k^{(2)}(s, u_2)
\end{aligned}$$

with

$$r_k^{(2)}(s, u_2) = \mathbb{E}\left[\int_{u_2}^{s} \left((f'\sigma)^2 (V_r) - (f'\sigma)^2 (V_{t_k})\right) dr\right]$$

and

$$\left|r_k^{(2)}(s, u_2)\right| \leq C \cdot \Delta_k^{3/2}.$$

Summarizing the different cases we have

$$\begin{aligned}
&2\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}\int_{t_k}^{u_2} \mathbb{E}\left[(Y_s - Y_{u_1})(Y_s - Y_{u_2})\right] du_1 du_2 ds \\
&= 2\varphi^2(t_k)\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}\int_{t_k}^{u_2} \left((u_1 - s)\mathbb{1}_{\{s \leq u_1 \leq u_2\}} + (s - u_2)\mathbb{1}_{\{u_1 \leq u_2 \leq s\}}\right) du_1 du_2 ds \\
&\quad + O(\Delta_k^{9/2}).
\end{aligned}$$

Now straightforward calculations yields that

$$\int_{t_k}^{u_2}(u_1-s)\mathbb{1}_{\{s\le u_1\le u_2\}}du_1=\int_s^{u_2}(u_1-s)\mathbb{1}_{\{s\le u_2\}}du_1=\frac{1}{2}(u_2-s)^2\mathbb{1}_{\{s\le u_2\}}$$

and

$$\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}\int_{t_k}^{u_2}(u_1-s)\mathbb{1}_{\{s\le u_1\le u_2\}}du_1du_2ds=\frac{1}{24}\Delta_k^4$$

as well as

$$\int_{t_k}^{u_2}(s-u_2)\mathbb{1}_{\{u_1\le u_2\le s\}}du_1=\int_{t_k}^{u_2}(s-u_2)\mathbb{1}_{\{u_2\le s\}}du_1=(s-u_2)(u_2-t_k)\mathbb{1}_{\{u_2\le s\}}$$

and

$$\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}\int_{t_k}^{u_2}(s-u_2)\mathbb{1}_{\{u_1\le u_2\le s\}}du_1du_2ds=\frac{1}{24}\Delta_k^4.$$

Consequently, we have shown that

$$\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}\mathbb{E}\left[\left(Y_s-Y_{u_1}\right)\left(Y_s-Y_{u_2}\right)\right]du_1du_2ds=\frac{1}{6}\varphi^2(t_k)\Delta_k^4+O(\Delta_k^{9/2})$$

and that

$$\begin{aligned}
&\mathbb{E}\left[\int_0^T|\phi_s^{(t_i)}|^2dt\right]\\
&=\sum_{k=0}^{m(N)-1}\frac{1}{\Delta_k^2}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}\mathbb{E}\left[\left(Y_s-Y_{u_1}\right)\left(Y_s-Y_{u_2}\right)\right]du_1du_2ds\\
&=\frac{1}{6}\sum_{k=0}^{m(N)-1}\varphi^2(t_k)\Delta_k^2+O(\Delta_{\max}^{3/2})
\end{aligned}$$

with $\Delta_{\max}=\max_{k=0,\dots,m(N)-1}|t_{k+1}-t_k|$. Since by construction

$$\Delta_{\max}\le C_\alpha\cdot N^{-\alpha}$$

uniformly over all discretizations, the assertion follows. □

Thus, we have shown

$$\mathbb{E}\left[\int_0^T|\phi_s^{(t_i),\alpha}|^2ds\right]=\frac{1}{6}\sum_{k=0}^{m(N)-1}\varphi^2(t_k^\alpha)(\Delta_k^\alpha)^2+O(N^{-3\alpha/2}),$$

uniformly in $(t_i) \in \Pi_B(N)$. Now we can apply Jensen's inequality, to obtain

$$\begin{aligned}\sqrt{N} \cdot \left(\mathbb{E}\left[\int_0^T |\phi_s^{(t_i),\alpha}|^2 dt\right]\right)^{1/2} &= \sqrt{N}\left(\frac{1}{6}\sum_{k=0}^{m(N)-1} \varphi^2(t_k^\alpha)(\Delta_k^\alpha)^2 + O(N^{-3\alpha/2})\right)^{1/2} \\ &\geq \sqrt{N}\left(\frac{1}{6}\sum_{k=0}^{m(N)-1} \varphi^2(t_k^\alpha)(\Delta_k^\alpha)^2\right)^{1/2} - O(N^{1/2-3\alpha/4}) \\ &= \sqrt{N}\sqrt{m(N)}\left(\frac{1}{6}\frac{1}{m(N)}\sum_{k=0}^{m(N)-1} \varphi^2(t_k^\alpha)(\Delta_k^\alpha)^2\right)^{1/2} \\ &\quad - O(N^{1/2-3\alpha/4}) \\ &\geq \frac{\sqrt{N}}{\sqrt{m(N)}}\frac{1}{\sqrt{6}}\sum_{k=0}^{m(N)-1} \varphi(t_k^\alpha)\Delta_k^\alpha - O(N^{1/2-3\alpha/4}).\end{aligned}$$

Since φ is a Hölder-1/4-function due to (10.7), we have that

$$\sum_{k=0}^{m(N)-1} \varphi(t_k^\alpha)\Delta_k^\alpha = \int_0^T \varphi(t)dt + O(N^{-\alpha/4})$$

and

$$\sqrt{N} \cdot \left(\mathbb{E}\left[\int_0^T |\phi_s^{(t_i),\alpha}|^2 dt\right]\right)^{1/2} \geq \frac{\sqrt{N}}{\sqrt{m(N)}}\frac{1}{\sqrt{6}}\int_0^T \varphi(t)dt - O(N^{-\alpha/4}) - O(N^{1/2-3\alpha/4}),$$

uniformly in $(t_i) \in \Pi_B(N)$. Choosing $\alpha \in (2/3, 1)$ we finally obtain

$$\liminf_{N\to\infty} \inf_{(t_i)\in\Pi_B(N)} \sqrt{N}\sqrt{1-\rho^2}\left(\mathbb{E}\left[\int_0^T |\phi_s^{(t_i),\alpha}|^2 dt\right]\right)^{1/2} \geq \sqrt{1-\rho^2}\frac{1}{\sqrt{6}}\int_0^T \varphi(t)dt,$$

since $\liminf_{N\to\infty} \frac{\sqrt{N}}{\sqrt{m(N)}} = 1$. Together with (10.5), this finishes the proof of Theorem 10.3.

10.3 Upper Bound

As shown in Lemma 10.4, a key step in our analysis is an idea of [47] to rewrite SDE (10.1) as

$$\begin{aligned} dX_t &= \rho dF(V_t) + h(V_t)dt + \sqrt{1-\rho^2}f(V_t)dB_t, \qquad && X_0 = x, \\ dV_t &= b\left(V_t\right)dt + \sigma\left(V_t\right)dW_t, && V_0 = v, \end{aligned} \tag{10.8}$$

where

$$F(y) = \int_v^y \frac{f}{\sigma}(z)dz, \qquad h(y) = r - \frac{1}{2}f^2(y) - \rho\left(\frac{bf}{\sigma} + \frac{1}{2}\left(\sigma f' - \sigma' f\right)\right)(y).$$

We now provide a matching upper bound to the lower bound from Theorem 10.3 by constructing a suitable discretization scheme for X_T. Using SDE (10.8), an approximation scheme of X_T on the discretization grid $\{0 = t_0 < t_1 < ... < t_N = T\}$ is then given by

$$\begin{aligned}\hat{x}_{t_N} &= x + \rho\left(F(\hat{v}_{t_N}) - F(v)\right) + \sum_{k=0}^{N-1} h(\hat{v}_{t_k})(t_{k+1} - t_k) \\ &\quad + \sqrt{1-\rho^2}\left(\sum_{k=0}^{N-1} \frac{1}{2}\left(f(\hat{v}_{t_k}) + f(\hat{v}_{t_{k+1}})\right)\left(B_{t_{k+1}} - B_{t_k}\right)\right),\end{aligned} \tag{10.9}$$

where $\hat{v}$ is an approximation of the volatility process V. We need the following assumptions:

Assumption 10.11. *(a) We have $f, h \in C^2(D;\mathbb{R})$ and*

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left(h'b + h''\frac{\sigma^2}{2}\right)^2(V_t)\right] < \infty, \qquad \sup_{t\in[0,T]} \mathbb{E}\left[\left(h'\sigma\right)^2(V_t)\right] < \infty,$$
$$\sup_{t\in[0,T]} \mathbb{E}\left[\left(f'b + f''\frac{\sigma^2}{2}\right)^2(V_t)\right] < \infty, \qquad \sup_{t\in[0,T]} \mathbb{E}\left[\left(f'\sigma\right)^2(V_t)\right] < \infty.$$

(b) The mapping

$$\varphi : [0,T] \to \mathbb{R}, \qquad \varphi(t) = \left(\mathbb{E}\left[(f'\sigma)^2(V_t)\right]\right)^{1/2}$$

satisfies $\varphi \in C([0,T];(0,\infty))$ and the discretization points are given by

$$t_k = \Phi^{-1}(k/N), \qquad k = 0, \ldots, N,$$

where

$$\Phi : [0,T] \to [0,1], \qquad \Phi(y) = \frac{\int_0^y \varphi(t)dt}{\int_0^T \varphi(t)dt}.$$

(c) Consider the scheme (10.9) and let $\varepsilon > 0$. We assume that there exists a constant $C_\varepsilon > 0$ such that

$$\sup_{k=0,\ldots,N} \mathbb{E}\left[|h(V_{t_k}) - h(\hat{v}_{t_k})|^2\right] \le C_\varepsilon \Delta_{\max}^{1+\varepsilon},$$
$$\sup_{k=0,\ldots,N} \mathbb{E}\left[|f(V_{t_k}) - f(\hat{v}_{t_k})|^2\right] \le C_\varepsilon \Delta_{\max}^{1+\varepsilon}$$

and

$$\mathbb{E}\left[|F(V_T) - F(\hat{v}_{t_N})|^2\right] \leq C_\varepsilon \Delta_{\max}^{1+\varepsilon},$$

where

$$\Delta_{\max} = \max_{k=1,\dots,N} |t_k - t_{k-1}|.$$

This second set of assumptions is to (a) control the Itō-Taylor expansion of $(h(V_t))_{t\in[0,T]}$ and $(f(V_t))_{t\in[0,T]}$, (b) to define the discretization points and (c) to bound the error of the approximation of the volatility.

Proposition 10.12. *Let Assumptions 10.1 and 10.11 hold. Then, the scheme* (10.9) *satisfies*

$$\limsup_{N\to\infty} \sqrt{N} \left(\mathbb{E}\left[|X_T - \hat{x}_{t_N}|^2\right]\right)^{1/2} \leq \sqrt{\frac{1-\rho^2}{4}} \int_0^T \left(\mathbb{E}\left[(f'\sigma)^2(V_t)\right]\right)^{1/2} dt.$$

10.4 Proof of Proposition 10.12

We split the proof of Proposition 10.12 into several parts and use the notation

$$\Delta_k = t_{k+1} - t_k, \qquad \Delta_k B = B_{t_{k+1}} - B_{t_k}, \qquad k = 0, 1, \dots, N-1.$$

First, we show the following lemma.

Lemma 10.13. *Let Assumptions 10.1 and 10.11 hold. For the approximation scheme* (10.9) *there exists* $\varepsilon > 0$ *such that*

$$\begin{aligned}&\left(\mathbb{E}\left[|X_T - \hat{x}_{t_N}|^2\right]\right)^{1/2} \\ &\quad \leq \sqrt{1-\rho^2}\left(\mathbb{E}\left[\left|\int_0^T f(V_t)dB_t - \sum_{k=0}^{N-1} \frac{f(V_{t_k}) + f(V_{t_k+1})}{2}\Delta_k B\right|^2\right]\right)^{1/2} + O\left(\Delta_{\max}^{\frac{1}{2}+\varepsilon}\right).\end{aligned}$$

Proof. First, we have by the Minkowski inequality

$$
\begin{aligned}
&\left(\mathbb{E}\left[|X_T - \hat{x}_{t_N}|^2\right]\right)^{1/2} \\
&= \mathbb{E}\left[\left|\rho\left(F(V_T) - F(\hat{v}_{t_N})\right) + \sum_{k=0}^{N-1}\left(\int_{t_k}^{t_{k+1}} h(V_t)dt - h(\hat{v}_{t_k})\Delta_k\right)\right.\right. \\
&\qquad \left.\left. + \sqrt{1-\rho^2}\sum_{k=0}^{N-1}\left(\int_{t_k}^{t_{k+1}} f(V_t)dB_t - \frac{1}{2}\left(f(\hat{v}_{t_k}) + f(\hat{v}_{t_{k+1}})\right)\Delta_k B\right)\right|^2\right]^{1/2} \\
&\leq \mathbb{E}\left[\left|\rho\left(F(V_T) - F(\hat{v}_{t_N})\right)\right|^2\right]^{1/2} + \mathbb{E}\left[\left|\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\left(h(V_t) - h(\hat{v}_{t_k})\right)dt\right|^2\right]^{1/2} \\
&\qquad + \mathbb{E}\left[\left|\sqrt{1-\rho^2}\sum_{k=0}^{N-1}\left(\int_{t_k}^{t_{k+1}} f(V_t)dB_t - \frac{1}{2}\left(f(\hat{v}_{t_k}) + f(\hat{v}_{t_{k+1}})\right)\Delta_k B\right)\right|^2\right]^{1/2} \\
&=: A + B + C.
\end{aligned}
$$

Then, we have by Assumption 10.11 (c) that

$$
A = |\rho|\mathbb{E}\left[|F(V_T) - F(\hat{v}_{t_N})|^2\right]^{1/2} \leq C(\Delta_{\max})^{\frac{1}{2}+\frac{\varepsilon}{2}}.
$$

For the second term, we have

$$
\begin{aligned}
B &= \mathbb{E}\left[\left|\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\left(h(V_t) - h(V_{t_k})\right)dt + \sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\left(h(V_{t_k}) - h(\hat{v}_{t_k})\right)dt\right|^2\right]^{1/2} \\
&\leq \mathbb{E}\left[\left|\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\left(h(V_t) - h(V_{t_k})\right)dt\right|^2\right]^{1/2} \\
&\qquad + \mathbb{E}\left[\left|\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\left(h(V_{t_k}) - h(\hat{v}_{t_k})\right)dt\right|^2\right]^{1/2} \\
&=: B_1 + B_2
\end{aligned}
$$

by the Minkowski inequality. By Itō's formula, integration by parts, the Itō-isometry, the Cauchy-Schwarz inequality, the Minkowski inequality and Assumption 10.11 (a), we have

$$
\begin{aligned}
B_1 &= \mathbb{E}\left[\left|\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\left(h'b+\frac{\sigma^2}{2}h''\right)(V_s)dsdt+\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t}\left(h'\sigma\right)(V_s)dW_sdt\right|^2\right]^{1/2}\\
&= \mathbb{E}\left[\left|\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}(t_{k+1}-s)\left(h'b+\frac{\sigma^2}{2}h''\right)(V_s)ds\right.\right.\\
&\qquad\left.\left.+\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}(t_{k+1}-s)\left(h'\sigma\right)(V_s)dW_s\right|^2\right]^{1/2}\\
&\le \mathbb{E}\left[\left|\int_0^T(\eta^+(t)-t)\left(h'b+\frac{\sigma^2}{2}h''\right)(V_t)dt\right|^2\right]^{1/2}\\
&\qquad+\mathbb{E}\left[\left|\int_0^T(\eta^+(t)-t)\left(h'\sigma\right)(V_t)dW_t\right|^2\right]^{1/2}\\
&\le \left(T(\Delta_{\max})^2\int_0^T\mathbb{E}\left[\left(h'b+\frac{\sigma^2}{2}h''\right)^2(V_t)\right]dt\right)^{1/2}\\
&\qquad+\left(T(\Delta_{\max})^2\int_0^T\mathbb{E}\left[\left(h'\sigma\right)^2(V_t)\right]dt\right)^{1/2}\\
&\le C\Delta_{\max},
\end{aligned}
$$

where $\eta^+(t) := \min\{t_k \in \{t_0, t_1, ..., t_N\} : t_k \ge t\}$. By the Hölder inequality and Assumption 10.11 (c), we obtain

$$
\begin{aligned}
B_2 &= \mathbb{E}\left[\left(\int_0^T\left(h(V_{\eta(t)})-h\left(\hat{v}_{\eta(t)}\right)\right)dt\right)^2\right]^{1/2}\\
&\le T^{1/2}\left(\int_0^T\mathbb{E}\left[\left|h(V_{\eta(t)})-h\left(\hat{v}_{\eta(t)}\right)\right|^2\right]dt\right)^{1/2}\\
&\le C(\Delta_{\max})^{\frac{1}{2}+\frac{\varepsilon}{2}},
\end{aligned}
$$

where $\eta(t) := \max\{t_k \in \{t_0, t_1, ..., t_N\} : t_k \le t\}$. The third term can be written as

$$
\begin{aligned}
C = \sqrt{1-\rho^2}\,\mathbb{E}&\left[\left|\int_0^T f(V_t)dB_t-\sum_{k=0}^{N-1}\frac{f(V_{t_k})+f(V_{t_{k+1}})}{2}\Delta_kB\right.\right.\\
&\qquad\left.\left.+\sum_{k=0}^{N-1}\frac{f(V_{t_k})-f(\hat{v}_{t_k})+f(V_{t_{k+1}})-f(\hat{v}_{t_{k+1}})}{2}\Delta_kB\right|^2\right]^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
&\leq \sqrt{1-\rho^2}\mathbb{E}\left[\left|\int_0^T f(V_t)dB_t - \sum_{k=0}^{N-1} \frac{f(V_{t_k})+f(V_{t_{k+1}})}{2}\Delta_k B\right|^2\right]^{1/2} \\
&\quad + \sqrt{1-\rho^2}\mathbb{E}\left[\left(\frac{1}{2}\sum_{k=0}^{N-1}\left(f(V_{t_k})-f(\hat{v}_{t_k})\right)\Delta_k B\right)^2\right]^{1/2} \\
&\quad + \sqrt{1-\rho^2}\mathbb{E}\left[\left(\frac{1}{2}\sum_{k=0}^{N-1}\left(f(V_{t_{k+1}})-f(\hat{v}_{t_{k+1}})\right)\Delta_k B\right)^2\right]^{1/2} \\
&=: C_1 + C_2 + C_3.
\end{aligned}
$$

The terms C_2 and C_3 can be bounded analogously using Assumption 10.11 (c) and the independence of W and B. That is, we have

$$
\begin{aligned}
C_2 &= \frac{\sqrt{1-\rho^2}}{2}\left(\sum_{k=0}^{N-1}\mathbb{E}\left[\left|\left(f(V_{t_k})-f(\hat{v}_{t_k})\right)\Delta_k B\right|^2\right]\right)^{1/2} \\
&= \frac{\sqrt{1-\rho^2}}{2}\left(\sum_{k=0}^{N-1}\mathbb{E}\left[\left|f(V_{t_k})-f(\hat{v}_{t_k})\right|^2\right]\mathbb{E}\left[\left|\Delta_k B\right|^2\right]\right)^{1/2} \\
&\leq \frac{\sqrt{1-\rho^2}}{2}\left(\sum_{k=0}^{N-1}\Delta_k^{2+\varepsilon}\right)^{1/2} \\
&\leq C(\Delta_{\max})^{\frac{1}{2}+\frac{\varepsilon}{2}}
\end{aligned}
$$

and

$$
C_3 \leq C(\Delta_{\max})^{\frac{1}{2}+\frac{\varepsilon}{2}}.
$$

This concludes the proof of this lemma. $\square$

The discretization points given by

$$
t_k = \Phi^{-1}(k/N), \quad k = 0,1,\ldots,N, \qquad \text{where} \qquad \Phi(y) = \frac{\int_0^y \varphi(t)dt}{\int_0^T \varphi(y)dy}, \qquad x \in [0,T],
$$

are regular, since

$$
c_\varphi := \sup_{x\in[0,T]} \frac{1}{\varphi(x)} < \infty
$$

due to Assumption 10.11 (b). More precisely, we have

$$
\Delta_{\max} = \max_{k=1,\ldots,N} |t_k - t_{k-1}| \leq c_\varphi \cdot \frac{1}{N}. \tag{10.10}
$$

Since $\Phi'(x) = \varphi(x)$, this follows from an application of the mean value theorem. We have

$$\Phi^{-1}\left(\frac{k+1}{N}\right) - \Phi^{-1}\left(\frac{k}{N}\right) = \left(\Phi^{-1}\right)'(\xi_k)\frac{1}{N} = \frac{1}{\Phi'\left(\Phi^{-1}\left(\xi_k\right)\right)}\frac{1}{N} = \frac{1}{\varphi(\tau_k)}\frac{1}{N}$$

with $\xi_k \in \left[\frac{k}{N}, \frac{k+1}{N}\right]$ and $\tau_k = \Phi^{-1}(\xi_k) \in [t_k, t_{k+1}]$. This gives

$$\Delta_{\max} = \max_{k=0,\dots,N-1}\left|\Phi^{-1}\left(\frac{k+1}{N}\right) - \Phi^{-1}\left(\frac{k}{N}\right)\right| \leq c_\varphi \cdot \frac{1}{N},$$

which is equation (10.10).
Assumption 10.11 (a) and an Itō-Taylor expansion also imply that there exists a constant $C > 0$ such that

$$\sup_{s,t\in[0,T]} \mathbb{E}\left[|f(V_t) - f(V_s)|^2\right] \leq C \cdot |t-s|. \tag{10.11}$$

Using (10.10) and (10.11) we can proceed analogously to the proof of Lemma 10.7 and obtain the following result.

Lemma 10.14. *Under Assumptions 10.1 and 10.11 we have*

$$\int_0^T f(V_t)dB_t - \sum_{k=0}^{N-1}\frac{f(V_{t_k}) + f(V_{t_k+1})}{2}\Delta_k B$$
$$\stackrel{d}{=} Z\left(\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\left(f(V_t) - \frac{1}{2}\left(f(V_{t_k}) + f(V_{t_{k+1}})\right)\right)^2 dt\right)^{1/2}$$

with

$$Z \sim \mathcal{N}(0,1).$$

Thus Lemma 10.13 gives that

$$\mathbb{E}\left[|X_T - \hat{x}_{t_N}|^2\right]^{1/2} \leq \sqrt{1-\rho^2}R_N + O\left(N^{-\frac{1}{2}-\varepsilon}\right)$$

with

$$R_N := \left(\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\mathbb{E}\left[\left(f(V_t) - \frac{1}{2}\left(f(V_{t_k}) + f(V_{t_{k+1}})\right)\right)^2\right]dt\right)^{1/2}.$$

The following result finishes the proof of Proposition 10.12.

Lemma 10.15. *(i) Under Assumptions 10.1 and 10.11 we have*

$$R_N = \frac{1}{2}\left(\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}(t_{k+1}-t_k)\varphi^2(t)dt\right)^{1/2} + O(N^{-3/4}).$$

(ii) Under Assumptions 10.1 and 10.11 we have

$$\limsup_{N\to\infty}\sqrt{N}\,\mathbb{E}\left[|X_T-\hat{x}_{t_N}|^2\right]^{1/2} \le \sqrt{\frac{1-\rho^2}{4}}\int_0^T\varphi(t)dt.$$

Proof. (i) An Itō-Taylor expansion yields

$$\begin{aligned}
&2f(V_t)-(f(V_{t_k})+f(V_{t_{k+1}}))\\
&\quad=(f(V_t)-f(V_{t_k}))+\big(f(V_t)-f(V_{t_{k+1}})\big)\\
&\quad=\int_{t_k}^{t}\left(f'b+f''\frac{\sigma^2}{2}\right)(V_s)ds+\int_{t_k}^{t}(f'\sigma)(V_s)dW_s\\
&\qquad-\int_{t}^{t_{k+1}}\left(f'b+f''\frac{\sigma^2}{2}\right)(V_s)ds-\int_{t}^{t_{k+1}}(f'\sigma)(V_s)dW_s\\
&\quad=\int_{t_k}^{t_{k+1}}sign(t-s)\left(f'b+f''\frac{\sigma^2}{2}\right)(V_s)ds+\int_{t_k}^{t_{k+1}}sign(t-s)(f'\sigma)(V_s)dW_s.
\end{aligned}$$

Under Assumption 10.11 (a) and using the Cauchy-Schwarz inequality, the Hölder inequality and the Itō-isometry there exists $C>0$ such that

$$\mathbb{E}\left[\left|\int_{t_k}^{t_{k+1}}sign(t-s)(f'b+f''\frac{\sigma^2}{2})(V_s)ds\right|^2\right]\le C\cdot N^{-2}$$

and

$$\mathbb{E}\left[\left|\int_{t_k}^{t_{k+1}}sign(t-s)(f'b+f''\frac{\sigma^2}{2})(V_s)ds\int_{t_k}^{t_{k+1}}sign(t-s)(f'\sigma)(V_s)dW_s\right|\right]\le C\cdot N^{-3/2}.$$

Another application of the Itō-isometry gives

$$R_n=\frac{1}{2}\left(\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\int_{t_k}^{t_{k+1}}\mathbb{E}\left[(f'\sigma)^2(V_s)\right]dsdt\right)^{1/2}+O(N^{-3/4}),$$

which concludes the proof of part (i).
(ii) With part (i), we have shown that

$$\begin{aligned}
\mathbb{E}\left[|X_T-\hat{x}_{t_N}|^2\right]^{1/2}\le\sqrt{\frac{1-\rho^2}{4}}&\left(\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}(t_{k+1}-t_k)\varphi^2(t)dt\right)^{1/2}\\
&+O(N^{-3/4})+O(N^{-1/2-\varepsilon})
\end{aligned}\tag{10.12}$$

Recall that

$$t_{k+1} - t_k = \Phi^{-1}\left(\frac{k+1}{N}\right) - \Phi^{-1}\left(\frac{k}{N}\right) = \frac{1}{\varphi(\tau_k)}\frac{1}{N}$$

with $\tau_k \in [t_k, t_{k+1}]$. Therefore, we have

$$\begin{aligned} N\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}(t_{k+1}-t_k)\varphi^2(t)dt &= \sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\frac{\varphi^2(t)}{\varphi(\tau_k)}dt \\ &= \sum_{k=0}^{N-1}\varphi(\tau_k)(t_{k+1}-t_k) + \sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}\frac{\varphi^2(t)-\varphi^2(\tau_k)}{\varphi(\tau_k)}dt. \end{aligned}$$

Since continuous functions with a compact domain of definition are uniformly continuous, we have that

$$\lim_{N\to\infty}\sup_{k=0,\ldots,N-1}\sup_{t\in[t_k,t_{k+1}]}|\varphi(t)-\varphi(\tau_k)| = 0.$$

The strict positivity and Riemann-integrability of φ imply now that

$$\lim_{N\to\infty} N\sum_{k=0}^{N-1}\int_{t_k}^{t_{k+1}}(t_{k+1}-t_k)\varphi^2(t)dt = \int_0^T \varphi(t)dt$$

and this gives the desired result together with Equation (10.12). □

10.5 Application to the Generalized Log-Heston Model

As already mentioned, the prototype example for SDE (10.1) is the generalized log-Heston model

$$\begin{aligned} dX_t &= \left(r - \frac{1}{2}V_t\right)dt + \sqrt{V_t}\left(\rho dW_t + \sqrt{1-\rho^2}dB_t\right), \qquad && X_0 = x, \\ dV_t &= \kappa(\theta - V_t)dt + \sigma V_t^\gamma dW_t, && V_0 = v, \end{aligned} \tag{10.13}$$

where $\kappa, \theta, \sigma > 0$, $\gamma \in \left[\frac{1}{2}, 1\right]$ with $\nu > 1$ if $\gamma = 1/2$. For $\gamma = \frac{1}{2}$, V is the CIR process, for $\gamma \in \left(\frac{1}{2}, 1\right)$ the volatility process is the CEV process (see [19] and Section 2.2) and for $\gamma = 1$ we are in the case of the Brennan-Schwartz model [12].
In this setup, it is well known that SDE (10.13) has a unique strong solution and that we have

$$\mathbb{P}\left(V_t \in (0,\infty), t \geq 0\right) = 1, \tag{10.14}$$

which is Assumption 10.1. See Sections 2.1 and 2.2 for $\gamma < 1$. For $\gamma = 1$, equation (10.14) follows from the explicit representation

$$\begin{aligned} V_t = v\exp&\left(-\left(\kappa + \frac{\sigma^2}{2}\right)t + \sigma W_t\right) \\ &+ \int_0^t \kappa\theta\exp\left(-\left(\kappa+\frac{\sigma^2}{2}\right)(t-s) + \sigma\left(W_t - W_s\right)\right)ds \end{aligned} \tag{10.15}$$

(see Satz 42 in [52]).

10.5.1 Verifying Assumption 10.2

We have

$$f(y) = \sqrt{y}, \qquad \sigma(y) = \sigma y^{\gamma},$$

so Assumption 10.2 (a) is also satisfied, since these maps are infinitely differentiable on $D = (0, \infty)$.

Now, it remains to verify condition 10.2 (b). Here we have

$$\left(f'b + \frac{1}{2}f''\sigma^2\right)(y) = \frac{\kappa\theta}{2\sqrt{y}} - \frac{\kappa}{2}\sqrt{y} - \frac{\sigma^2}{8}y^{2\gamma-\frac{3}{2}}$$

and

$$(f'\sigma)(y) = \frac{\sigma}{2}y^{\gamma-1/2}.$$

We will need the following Lemma:

Lemma 10.16. *We have*

$$\sup_{t\in[0,T]} \mathbb{E}\left[V_t^p\right] < \infty$$

for all $p \in \mathbb{R}$ *if* $\gamma = 1$.

Proof. For $\gamma = 1$, equation (10.15) yields

$$\frac{1}{V_t} \le \frac{1}{v}\exp\left(\left(\kappa + \frac{\sigma^2}{2}\right)t - \sigma W_t\right)$$

and the statement follows in this case from

$$\mathbb{E}\left[\exp\left(\alpha W_t\right)\right] = \exp\left(\frac{\alpha^2}{2}t\right), \qquad \alpha \in \mathbb{R},\ t \ge 0. \tag{10.16}$$

the exponential integrability of the Brownian motion.

□

Now Proposition 2.2, Proposition 2.6, Lemma 10.16 and the Minkowski inequality yield that

$$\sup_{t\in[0,T]} \mathbb{E}\left[\left|\left(f'b + \frac{1}{2}f''\sigma^2\right)(V_t)\right|^2\right] = \sup_{t\in[0,T]} \mathbb{E}\left[\left|\frac{\kappa\theta}{2\sqrt{V_t}} - \frac{\kappa}{2}\sqrt{V_t} - \frac{\sigma^2}{8}V_t^{2\gamma-\frac{3}{2}}\right|^2\right] < \infty$$

for $\gamma \in \left[\frac{1}{2}, 1\right]$ with $\nu > 1$ if $\gamma = 1/2$. Note that

$$|z^{\alpha} - y^{\alpha}| = |\alpha|\left|\int_y^z u^{\alpha-1}du\right| \le |\alpha|\,|z-y|\left(z^{\alpha-1} + y^{\alpha-1}\right)$$

for $\alpha \in \mathbb{R}, y, z > 0$. For $\gamma > 1/2$, we thus have

$$\mathbb{E}\left[\left|(f'\sigma)(V_t) - (f'\sigma)(V_s)\right|^2\right] \leq C\,\mathbb{E}\left[\left|\int_s^t \kappa(\theta - V_u)du + \sigma \int_s^t V_u^\gamma dW_u\right|^4\right]^{1/2} \leq C|t-s|$$

by Proposition 2.6, Lemma 10.16, the Minkowski, Hölder and the Burkholder-Davis-Gundy inequality. For $\gamma = 1/2$, the function $f'\sigma$ is constant. Hence, Assumption 10.2 is fulfilled and we have established the following proposition:

Proposition 10.17. *Assume that* $\gamma \in [\frac{1}{2}, 1]$ *and that* $\nu > 1$ *if* $\gamma = \frac{1}{2}$. *For SDE* (10.13) *we then have*

$$\liminf_{N\to\infty} \sqrt{N}\, e(N) \geq \frac{\sigma\sqrt{1-\rho^2}}{\sqrt{24}} \int_0^T \left(\mathbb{E}\left[V_t^{2\gamma-1}\right]\right)^{1/2} dt. \tag{10.17}$$

For the cases $\gamma \in \left\{\frac{1}{2}, 1\right\}$, we can write the right side of (10.17) in a more explicit way. For $\gamma = \frac{1}{2}$, we have

$$\mathbb{E}\left[V_t^{2\gamma-1}\right] = \mathbb{E}\left[V_t^0\right] = 1, \qquad t \in [0, T],$$

and for $\gamma = 1$ we have

$$\mathbb{E}\left[V_t^{2\gamma-1}\right] = \mathbb{E}\left[V_t\right] = ve^{-\kappa t} + \theta\left(1 - e^{-\kappa t}\right), \qquad t \in [0, T].$$

10.5.2 Verifying Assumption 10.11

For the upper bound the functions of interest in the generalized Heston model are

$$f(y) = \sqrt{y}, \qquad b(y) = \kappa\left(\theta - y\right), \qquad \sigma(y) = \sigma y^\gamma$$

and

$$F(y) = \frac{1}{\sigma\left(\frac{3}{2} - \gamma\right)}\left(y^{\frac{3}{2}-\gamma} - v^{\frac{3}{2}-\gamma}\right),$$
$$h(y) = r - \frac{1}{2}y - \rho\left(\frac{\kappa(\theta - y)}{\sigma}y^{\frac{1}{2}-\gamma} + \frac{\sigma}{2}y^{\gamma-\frac{1}{2}}\left(\frac{1}{2} - \gamma\right)\right).$$

In the following we focus on $\gamma = \frac{1}{2}$ and $\gamma = 1$. The case $\gamma \in (\frac{1}{2}, 1)$ can be analyzed similarly by extending the results from [2] and [62] to suitable non-equidistant (but non-adaptive) discretizations, but is not treated in this thesis for the sake of conciseness.

The case $\gamma = \frac{1}{2}$

Here we have

$$f(y) = y^{1/2}, \qquad b(y) = \kappa\left(\theta - y\right), \qquad \sigma(y) = \sigma y^{1/2}$$

and

$$F(y) = \frac{1}{\sigma}(y - v), \qquad h(y) = r - \frac{1}{2}y - \rho\frac{\kappa(\theta - y)}{\sigma}$$

and the functions

$$(0,\infty) \ni y \mapsto \left(h'b + h''\frac{\sigma^2}{2}\right)(y) \in \mathbb{R}, \qquad (0,\infty) \ni y \mapsto (h'\sigma)(y) \in \mathbb{R},$$
$$(0,\infty) \ni y \mapsto \left(f'b + f''\frac{\sigma^2}{2}\right)(y) \in \mathbb{R}, \qquad (0,\infty) \ni y \mapsto (f'\sigma)(y) \in \mathbb{R},$$

appearing in Assumption 10.11 (a) are bounded (in absolute value) by the function

$$(0,\infty) \ni y \mapsto C\left(1 + y + y^{-1/2}\right) \in (0,\infty),$$

for a suitable constant $C > 0$. Thus, this assumption is verified due to Proposition 2.2. In particular $(f'\sigma)(y) = \frac{\sigma}{2}$ and so Assumption 10.11 (b) is trivially satisfied. Note that the corresponding discretization points are equidistant, since $\Phi(y) = y/T$. For condition (c) we need to choose a particular approximation scheme. Here, we take the drift-implicit Euler scheme from Section 3.3.3 which approximates the process $Z = \sqrt{V}$ and is given by

$$\begin{aligned} z_{k+1} &= z_k + \left(\frac{4\kappa\theta - \sigma^2}{8}\frac{1}{z_{k+1}} - \frac{\kappa}{2}z_{k+1}\right)\Delta t + \frac{\sigma}{2}\Delta_k W, \\ v_{k+1} &= z_{k+1}^2 \end{aligned} \tag{10.18}$$

where $\Delta t = T/N$, $t_k = k\Delta t$ and $z_0 = \sqrt{v}$. In [2] it is shown that

$$\sup_{k=0,\ldots,N} \mathbb{E}\left[\left|\hat{z}_{t_k} - \sqrt{V_{t_k}}\right|^p\right] \le C_p(\Delta t)^p, \qquad \sup_{k=0,\ldots,N} \mathbb{E}\left[\left|\hat{z}_{t_k}^2 - V_{t_k}\right|^p\right] \le C_p(\Delta t)^p$$

for its time-continuous extension from Equation (7.8) if $\nu > 2$ and $1 \le p < \frac{2}{3}\nu$ and

$$\sup_{k=0,\ldots,N} \mathbb{E}\left[|\hat{z}_{t_k}|^p\right] < \infty$$

for all $p \ge 1$. For $\varepsilon \in (0, \frac{1}{3})$ and $\nu > 2$ a standard argument using Hölder's inequality yields

$$\sup_{k=0,\ldots,N} \mathbb{E}\left[\left|\hat{z}_{t_k} - \sqrt{V_{t_k}}\right|^2\right] \le C_\varepsilon(\Delta t)^{\frac{4}{3}-\varepsilon}, \qquad \sup_{k=0,\ldots,N} \mathbb{E}\left[\left|\hat{z}_{t_k}^2 - V_{t_k}\right|^2\right] \le C_\varepsilon(\Delta t)^{\frac{4}{3}-\varepsilon}.$$

Since $f(y) = \sqrt{y}$, and F, h are linear, Assumption 10.11 (c) is satisfied for

$$\hat{v}_{t_k} = \hat{z}_{t_k}^2, \qquad k = 0, \ldots, N.$$

Combining the upper and the lower bound and taking into account that $(f'\sigma)(y) = \frac{\sigma}{2}$ we obtain the following result:

Proposition 10.18. *Assume that $\gamma = \frac{1}{2}$ and $\nu > 2$. For SDE* (10.13) *and scheme* (10.9) *where $(\hat{v}_{t_k})_{k\in\{0,\dots,N\}}$ is given by the drift-implicit Euler scheme* (10.18) *with discretization points $t_k = kT/N$, $k = 0, \dots, N$, we have*

$$\begin{aligned}\frac{\sigma\sqrt{1-\rho^2}T}{\sqrt{24}} &\le \liminf_{N\to\infty} \sqrt{N}\, e(N) \\ &\le \limsup_{N\to\infty} \sqrt{N}\mathbb{E}\left(\left[|X_T - \hat{x}_{t_N}|^2\right]\right)^{1/2} \le \frac{\sigma\sqrt{1-\rho^2}T}{4}.\end{aligned}$$

The case $\gamma = 1$

In this case the coefficients of the SDE read as

$$f(y) = y^{1/2} \qquad b(y) = \kappa\,(\theta - y) \qquad \sigma(y) = \sigma y$$

and

$$F(y) = \frac{2}{\sigma}\left(y^{1/2} - v^{1/2}\right) \qquad h(y) = r - \frac{1}{2}y - \rho\left(\frac{\kappa(\theta-y)}{\sigma y^{1/2}} - \frac{\sigma}{4}y^{1/2}\right).$$

The functions

$$\begin{aligned}(0,\infty) \ni y \mapsto \left(h'b + h''\frac{\sigma^2}{2}\right)(y) \in \mathbb{R}, \qquad (0,\infty) \ni y \mapsto (h'\sigma)(y) \in \mathbb{R},\\ (0,\infty) \ni y \mapsto \left(f'b + f''\frac{\sigma^2}{2}\right)(y) \in \mathbb{R}, \qquad (0,\infty) \ni y \mapsto (f'\sigma)(y) \in \mathbb{R},\end{aligned}$$

appearing in Assumption 10.11 (a) are of the form

$$(0,\infty) \ni y \mapsto \sum_{\ell=-k}^{k} c_\ell y^{\ell/2} \in \mathbb{R}$$

for a suitable $k \in \mathbb{N}$ and $c_\ell \in \mathbb{R}$ for $\ell = -k, -k+1, \dots, k$. Lemma 10.16 implies then that Assumption 10.11 (a) is satisfied. Moreover, the function φ is given by

$$\varphi(t) = \left(\mathbb{E}\left[(f'\sigma)^2(V_t)\right]\right)^{1/2} = \frac{\sigma}{2}\left(\mathbb{E}\left[V_t\right]\right)^{1/2} = \frac{\sigma}{2}\left(v_0 e^{-\kappa t} + \theta\left(1 - e^{-\kappa t}\right)\right)^{1/2}$$

for $t \in [0,T]$ which is continuous and strictly positive. Therefore, also Assumption 10.11 (b) is satisfied. For condition (c) we need to choose again a particular approximation scheme. Here, we take the Euler-type discretization

$$\begin{aligned}\hat{v}_{t_k} &= v\exp\left(-\left(\kappa + \frac{\sigma^2}{2}\right)t_k + \sigma W_{t_k}\right) \\ &\quad + \sum_{l=0}^{k-1} \kappa\theta\exp\left(-\left(\kappa + \frac{\sigma^2}{2}\right)(t_k - t_l) + \sigma\left(W_{t_k} - W_{t_l}\right)\right)(t_{l+1} - t_l)\end{aligned} \tag{10.19}$$

for $k = 0, \dots, N$. We set

$$\Delta_{\max} := \max_{k=1,\dots,N} |t_{k+1} - t_k|.$$

For scheme (10.19), we have the following estimates:

Lemma 10.19. *Let $p \geq 1$ and $q \in \mathbb{R}$. There exist constants $C_p > 0$ such that*

$$\sup_{k=0,\dots,N} \mathbb{E}\left[|V_{t_k} - \hat{v}_{t_k}|^p\right] \leq C_p (\Delta_{\max})^p. \tag{10.20}$$

Moreover, we have

$$\sup_{k=0,\dots,N} \mathbb{E}\left[|\hat{v}_{t_k}|^q\right] < \infty. \tag{10.21}$$

Proof. Let

$$g(t,w) := \kappa\theta \exp\left(-\left(\kappa + \frac{\sigma^2}{2}\right)t + \sigma w\right), \qquad h(t,w) := \exp\left(\left(\kappa + \frac{\sigma^2}{2}\right)t - \sigma w\right).$$

Then we can write

$$V_t = g(t, W_t)\left(v + \kappa\theta \int_0^t h(s, W_s) ds\right)$$

and

$$\hat{v}_{t_k} = g(t_k, W_{t_k})\left(v + \kappa\theta \int_0^{t_k} h(\eta(s), W_{\eta(s)}) ds\right).$$

It is well known that

$$\max_{k=0,\dots,N} \mathbb{E}\left[\left|\int_0^{t_k} h(s, W_s) ds - \int_0^{t_k} h(\eta(s), W_{\eta(s)}) ds\right|^p\right] \leq C_p (\Delta_{\max})^p, \tag{10.22}$$

however we could not find a reference for this. The closest reference is [29], which considers SDEs with bounded coefficients instead of geometric Brownian motion. However, the above estimate can be shown using standard arguments based on an Itō-Taylor expansion, the Minkowski, Hölder and Burkholder-Davis-Gundy inequalities, see also the proof of Lemma 10.13. Thus, another application of the Hölder inequality, of equation (10.16) and of equation (10.22) yield

$$\begin{aligned}
&\sup_{k=0,\dots,N} \mathbb{E}\left[|V_{t_k} - \hat{v}_{t_k}|^p\right] \\
&\quad \leq (\kappa\theta)^p \sup_{k=0,\dots,N} \left(\mathbb{E}\left[g(t_k, W_{t_k})^{2p}\right]\right)^{1/2} \\
&\qquad \cdot \sup_{k=0,\dots,N} \left(\mathbb{E}\left[\left|\int_0^{t_k} h(s, W_s) ds - \int_0^{t_k} h(\eta(s), W_{\eta(s)}) ds\right|^{2p}\right]\right)^{1/2} \\
&\quad \leq C_p (\Delta_{\max})^p,
\end{aligned}$$

which is (10.20), i.e. the first statement.
For the second statement (10.21) note that

$$\frac{1}{\hat{v}_{t_k}} \geq \frac{1}{v} \exp\left(\left(\kappa + \frac{\sigma^2}{2}\right) t_k - \sigma W_{t_k}\right),$$

and for $q < 0$ the assertion follows again from equation (10.16). For $q \geq 1$ we can use that

$$\hat{v}_{t_k} = V_{t_k} + (\hat{v}_{t_k} - V_{t_k})$$

Lemma 10.16, the estimate (10.20) and the Minkowski inequality. Finally, for $q \in [0, 1)$ we can apply $|y|^q \leq 1 + |y|$. □

Now recall that

$$\begin{aligned} f(y) &= \sqrt{y}, \qquad F(y) = \frac{2}{\sigma}\left(\sqrt{y} - \sqrt{v}\right), \\ h(y) &= r - \frac{1}{2}y - \rho\left(\frac{\kappa\theta}{\sigma}\frac{1}{\sqrt{y}} - \left(\frac{\kappa}{\sigma} + \frac{\sigma}{4}\right)\sqrt{y}\right). \end{aligned}$$

Since

$$|\sqrt{z} - \sqrt{y}| \leq \left(\sqrt{z} + \sqrt{y}\right)|z - y|, \qquad y, z > 0,$$

and

$$\left|\frac{1}{\sqrt{z}} - \frac{1}{\sqrt{y}}\right| \leq \frac{1}{2}\left(z^{-3/2} + y^{-3/2}\right)|z - y|, \qquad y, z > 0,$$

Lemma 10.16 and 10.19 now imply that also Assumption 10.11 (c) is satisfied. Thus we obtain the following result:

Proposition 10.20. *Assume that $\gamma = 1$ and let $\varphi(t) = \frac{\sigma}{2}\left(ve^{-\kappa t} + \theta\left(1 - e^{-\kappa t}\right)\right)^{1/2}, t \in [0, T]$. For SDE (10.13) and the schemes (10.9) and (10.19) with discretization points given by $t_k = \Phi^{-1}(k/N)$, $k = 0, 1, \ldots, N$, where $\Phi(y) = \int_0^y \varphi(t)dt / \int_0^T \varphi(t)dt$, we have*

$$\begin{aligned} \sqrt{\frac{1-\rho^2}{6}} \int_0^T \varphi(t)\, dt &\leq \liminf_{N\to\infty} \sqrt{N}\, e(N) \\ &\leq \limsup_{N\to\infty} \sqrt{N}\left(\mathbb{E}\left[|X_T - \hat{x}_{t_N}|^2\right]\right)^{1/2} \leq \sqrt{\frac{1-\rho^2}{4}} \int_0^T \varphi(t)\, dt. \end{aligned}$$

Chapter 11

Conclusion

In this thesis, we analyzed several numerical schemes for the log-Heston model and the CIR process. Our main motivation was to provide weak and strong convergence results for explicit Euler-type discretization schemes which are very easy to implement. For these schemes, results from the literature were rare and came with strong restrictions on the parameter range.

Our first main result was Theorem 6.3 where we provided the first weak convergence result for an Euler discretization of the log-Heston model. We could also observe in Section 9.1 that our theoretical convergence rates are attained under even milder assumptions.

In Chapter 7 we could prove strong convergence rates for all known Euler-type schemes of the CIR process and the log-Heston model (Theorem 7.13 and Proposition 7.14). For the first group of Euler schemes, which allow negative values of the CIR approximation throughout the simulation, our results hold without any additional assumptions. For the second group, our proof is valid for $\nu > 1$. Together with Theorem 8.1 we could show that the achieved convergence order of the Euler schemes is already optimal for the log-Heston model in this parameter range. In terms of the convergence order, there is no advantage in using a more sophisticated scheme. Our simulations in Section 9.2 confirm these theoretical results. For the parameter range $\nu \leq 1$, the numerical simulations indicate that an even better convergence order than the one from Proposition 7.14 might be possible.

In Proposition 6.9, we extended existing weak convergence results for a Milstein-type scheme. In Proposition 7.15 we could also show new strong convergence results for this method. Both proofs hold for the whole parameter range where the implicit Milstein scheme is positivity preserving.

Finally, we analyzed the minimal L^2-error for general stochastic volatility models in Chapter 10. We could prove a lower bound in Theorem 10.3 and a matching upper bound (up to a factor $\sqrt{3/2}$) in Proposition 10.12.

List of Figures

List of Tables

Bibliography

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

[2] A. Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comput.*, 79(269):209–237, 2010.

[3] M. Altmayer. *Quadrature of discontinuous SDE functionals using Malliavin integration by parts.* PhD thesis, Mannheim, 2015.

[4] M. Altmayer and A. Neuenkirch. Multilevel Monte Carlo quadrature of discontinuous payoffs in the generalized Heston model using Malliavin integration by parts. *SIAM J. Finan. Math.*, 6(1):22–52, 2015.

[5] M. Altmayer and A. Neuenkirch. Discretising the Heston model: an analysis of the weak convergence rate. *IMA J. Numer. Anal.*, 37(4):1930–1960, 2017.

[6] L.B.G. Andersen. Simple and efficient simulation of the Heston stochastic volatility model. *J. Comput. Finance*, 11(3):29–50, 2008.

[7] L.B.G. Andersen and V.V. Piterbarg. Moment explosions in stochastic volatility models. *Finance Stoch.*, 11(1):29–50, 2007.

[8] V. Bally and D. Talay. The law of the Euler scheme for stochastic differential equations. I: Convergence rate of the distribution function. *Probab. Theory Relat. Fields*, 104(1):43–60, 1996.

[9] A. Berkaoui, M. Bossy, and A. Diop. Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM: PS*, 12:1–11, 2008.

[10] M. Bossy and A. Diop. An efficient discretisation scheme for one dimensional SDEs with a diffusion coefficient function of the form $|x|^a$, $a \in [1/2, 1)$. Research Report RR-5396, INRIA, 2007. Version 2.

[11] M. Bossy and H. Olivero Quinteros. Strong convergence of the symmetrized Milstein scheme for some CEV-like SDEs. *Bernoulli*, 24(3):1995–2042, 2018.

[12] M.J. Brennan and E.S. Schwartz. Analyzing convertible bonds. *The Journal of Financial and Quantitative Analysis*, 15(4):907–929, 1980.

[13] M. Briani, L. Caramellino, and G. Terenzi. Convergence rate of Markov chains and hybrid numerical schemes to jump-diffusions with application to the Bates model. *SIAM Journal on Numerical Analysis*, 59(1):477–502, 2021.

[14] M. Broadie and Ö. Kaya. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Oper. Res.*, 54(2):217–231, 2006.

[15] S. Cambanis and Y. Hu. Exact convergence rate of the Euler-Maruyama scheme, with application to sampling design. *Stochastics Stochastics Rep.*, 59(3-4):211–240, 1996.

[16] F. Castell and J. Gaines. The ordinary differential equation approach to asymptotically efficient schemes for solution of stochastic differential equations. *Ann. Inst. Henri Poincaré, Probab. Stat.*, 32(2):231–250, 1996.

[17] J.-F. Chassagneux, A. Jacquier, and I. Mihaylov. An explicit Euler scheme with strong rate of convergence for financial SDEs with non-Lipschitz coefficients. *SIAM J. Financial Math.*, 7:993–1021, 2016.

[18] J.M.C. Clark and R.J. Cameron. The maximum rate of convergence of discrete approximations for stochastic differential equations. In *Stochastic differential systems (Proc. IFIP-WG 7/1 Working Conf., Vilnius, 1978)*, pages 162–171. Springer, Berlin-New York, 1980.

[19] J.C. Cox. The constant elasticity of variance option pricing model. *Journal of Portfolio Management*, 23(5):15–17, 1996.

[20] J.C. Cox, J.E. Ingersoll, and S.A. Ross. A Theory of the Term Structure of Interest Rates. *Econometrica*, 55(2):385–407, 1985.

[21] A. Cozma and C. Reisinger. Strong order 1/2 convergence of full truncation Euler approximations to the Cox-Ingersoll-Ross process. *IMA Journal of Numerical Analysis*, 00:1–19, 2018.

[22] T. De Angelis, M. Germain, and E. Issoglio. A numerical scheme for stochastic differential equations with distributional drift. *arXiv:1906.11026v4*, 2022.

[23] G. Deelstra and F. Delbaen. Convergence of discretized stochastic (interest rate) processes with stochastic drift term. *Appl. Stochastic Models Data Anal.*, 14(1):77–84, 1998.

[24] S. Dereich, A. Neuenkirch, and L. Szpruch. An Euler-type method for the strong approximation of the Cox–Ingersoll–Ross process. *Proc. R. Soc. A*, 468(2140):1105–1115, 2012.

[25] D. Duffie and P. Glynn. Efficient Monte Carlo Simulation of Security Prices. *Ann. Appl. Probab.*, 5(4):897 – 905, 1995.

[26] P.M.N. Feehan and C.A. Pop. A Schauder approach to degenerate-parabolic partial differential equations with unbounded coefficients. *J. Differ. Equations*, 254(12):4401–4445, 2013.

[27] W. Feller. Two singular diffusion problems. *Annals of Mathematics*, 54(1):173–182, 1951.

[28] M. Giles. Multilevel monte carlo path simulation. *Operations Research*, 56(3):607–6017, 2008.

[29] E. Gobet and S. Menozzi. Discrete sampling of functionals of Itô processes. In *Séminaire de Probabilités XL*, pages 355–374. Berlin: Springer, 2007.

[30] I. Gyöngy and M. Rásonyi. A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients. *Stochastic Process. Appl.*, 121(10):2189–2200, 2011.

[31] M. Hefter and A. Herzwurm. Optimal strong approximation of the one-dimensional squared Bessel process. *Commun. Math. Sci.*, 15(8):2121–2141, 2017.

[32] M. Hefter and A. Herzwurm. Strong convergence rates for Cox–Ingersoll–Ross processes — Full parameter range. *Journal of Mathematical Analysis and Applications*, 459(2):1079–1101, 2018.

[33] M. Hefter, A. Herzwurm, and T. Müller-Gronbach. Lower error bounds for strong approximation of scalar SDEs with non-Lipschitzian coefficients. *Ann. Appl. Probab.*, 29(1):178–216, 2019.

[34] M. Hefter and A. Jentzen. On arbitrarily slow convergence rates for strong numerical approximations of Cox-Ingersoll-Ross processes and squared Bessel processes. *Finance Stoch.*, 23(1):139–172, 2019.

[35] S. Heinrich. Multilevel Monte Carlo Methods. In *Large-Scale Scientific Computing*, pages 58–67. Springer Berlin Heidelberg, 2001.

[36] S.L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financial Studies*, 6(2):327–343, 1993.

[37] D. J. Higham and X. Mao. Convergence of the Monte Carlo simulations involving the mean-reverting square root process. *J. Comput. Finance*, 8:35–62, 2005.

[38] N. Hofmann, T. Müller-Gronbach, and K Ritter. Optimal approximation of stochastic differential equations by adaptive step-size control. *Math. Comput.*, 69(231):1017–1034, 2000.

[39] N. Hofmann, T. Müller-Gronbach, and K. Ritter. Step size control for the uniform approximation of systems of stochastic differential equations with additive noise. *Ann. Appl. Probab.*, 10(2):616–633, 2000.

[40] N. Hofmann, T. Müller-Gronbach, and K. Ritter. The optimal discretization of stochastic differential equations. *J. Complexity*, 17(1):117–153, 2001.

[41] N. Hofmann, T. Müller-Gronbach, and K. Ritter. Linear vs standard information for scalar stochastic differential equations. *J. Complexity*, 18(2):394–414, 2002.

[42] T.R. Hurd and A. Kuznetsov. Explicit formulas for Laplace transforms of stochastic integrals. *Markov Process. Relat. Fields*, 14(2):277–290, 2008.

[43] M. Hutzenthaler, A. Jentzen, and M. Noll. Strong convergence rates and temporal regularity for Cox-Ingersoll-Ross processes and Bessel processes with accessible boundaries. *arXiv:1403.6385*, 2014.

[44] N. Ikeda and S. Watanabe. *Stochastic differential equations and diffusion processes.* North-Holland, 1989.

[45] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical methods for financial markets.* Springer Dordrecht Heidelberg London New York, 2009.

[46] A. Jentzen, T. Müller-Gronbach, and L. Yaroslavtseva. On stochastic differential equations with arbitrary slow convergence rates for strong approximation. *Commun. Math. Sci.*, 14(6):1477–1500, 2016.

[47] B. Jourdain and M. Sbai. High-order discretization schemes for stochastic volatility models. *J. Comput. Finance*, 17(2), 2013.

[48] C. Kahl and P. Jäckel. Fast strong approximation Monte-Carlo schemes for stochastic volatility models. *Quant. Finance*, 6(6):513–536, 2006.

[49] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus.* New York, Springer-Verlag, 2nd edition, 1991.

[50] S. Karlin and H.M. Taylor. *A second course in stochastic processes.* Academic Press New York, 1981.

[51] P. Kloeden and A. Neuenkirch. Convergence of numerical methods for stochastic differential equations in mathematical finance. In *Recent Developments in Computational Finance*, chapter 2, pages 49–80. World Sci. Publ., Hackensack, NJ, 2013.

[52] E. Korn and R. Korn. *Optionsbewertung und Portfoliooptimierung.* Vieweg, 1999.

[53] A. Mickel and A. Neuenkirch. The order barrier for the L^1-approximation of the log-Heston SDE at a single point. *arXiv:2212.07252v2; to appear in: A. Hinrichs, P. Kritzer, F. Pillichshammer (eds.). Monte Carlo and Quasi-Monte Carlo Methods 2022. Springer Verlag.*

[54] A. Mickel and A. Neuenkirch. The weak convergence rate of two semi-exact discretization schemes for the Heston model. *Risks*, 9(1), 2021.

[55] A. Mickel and A. Neuenkirch. Sharp L^1-Approximation of the log-Heston SDE by Euler-type methods. *arXiv:2206.03229*, 2022.

[56] A. Mickel and A. Neuenkirch. The weak convergence order of two Euler-type discretization schemes for the log-Heston model. *arXiv:2106.10926v2; to appear in: IMA J. Numer. Anal.*, 2022.

[57] T. Müller-Gronbach. The optimal uniform approximation of systems of stochastic differential equations. *Ann. Appl. Probab.*, 12(2):664–690, 2002.

[58] T. Müller-Gronbach. Strong approximation of systems of stochastic differential equations, Habilitation thesis, Januar 2002.

[59] T. Müller-Gronbach. Optimal pointwise approximation of SDEs based on Brownian motion at discrete points. *Ann. Appl. Probab.*, 14(4):1605–1642, 2004.

[60] T. Müller-Gronbach and K. Ritter. Minimal errors for strong and weak approximation of stochastic differential equations. In *Monte Carlo and quasi-Monte Carlo methods 2006. Selected papers based on the presentations at the 7th international conference 'Monte Carlo and quasi-Monte Carlo methods in scientific computing', Ulm, Germany, August 14–18, 2006.*, pages 53–82. Berlin: Springer, 2008.

[61] T. Müller-Gronbach and L. Yaroslavtseva. A note on strong approximation of SDEs with smooth coefficients that have at most linearly growing derivatives. *J. Math. Anal. Appl.*, 467(2):1013–1031, 2018.

[62] A. Neuenkirch and L. Szpruch. First order strong approximations of scalar SDEs defined in a domain. *Numer. Math.*, 128(1):103–136, 2014.

[63] N.J. Newton. An asymptotically efficient difference formula for solving stochastic differential equations. *Stochastics*, 19:175–206, 1986.

[64] N.J. Newton. An efficient approximation for stochastic differential equations on the partition of symmetrical first passage times. *Stochastics Stochastics Rep.*, 29(2):227–258, 1990.

[65] N.J. Newton. Asymptotically efficient Runge-Kutta methods for a class of Itô and Stratonovich equations. *SIAM J. Appl. Math.*, 51(2):542–567, 1991.

[66] B. Oksendahl. *Stochastic Differential Equations : An Introduction with Applications.* Berlin,Heidelberg, Springer-Verlag, 6th edition, 2003.

[67] R. Lord, R. Koekkoek, and D.J.C. van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quant. Finance*, 10(2):177–194, 2009.

[68] D. Revus and M.Yor. *Continuous Martingales and Brownian Motion.* Springer Berlin Heidelberg New York, 3rd edition, 1999.

[69] R.D. Smith. An almost exact simulation method for the Heston model. *J. Comput. Finance*, 11(1):115–125, 2007.

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

[71] J. F. Traub, G. W. Wasilkowski, and H. Woźniakowski. *Information-based complexity.* Computer Science and Scientific Computing. Academic Press, Inc., Boston, MA, 1988. With contributions by A. G. Werschulz and T. Boult.

[72] A. van Haastrecht and A Pelsser. Efficient, almost exact simulation of the heston stochastic volatility model. *International Journal of Theoretical and Applied Finance*, 13(1):1–43, 2010.

[73] L. Yaroslavtseva. On non-polynomial lower error bounds for adaptive strong approximation of SDEs. *J. Complexity*, 42:1–18, 2017.

[74] C. Zheng. Weak Convergence Rate of a Time-Discrete Scheme for the Heston Stochastic Volatility. *SIAM J. Numer. Anal.*, 55(3):1243–1263, 2017.