

Article

# The Weak Convergence Rate of Two Semi-Exact Discretization Schemes for the Heston Model

Annalena Mickel <sup>1,2</sup> and Andreas Neuenkirch <sup>2,\*</sup>

<sup>1</sup> DFG Research Training Group 1953, University of Mannheim, B6, 26, D-68131 Mannheim, Germany; amickel@mail.uni-mannheim.de

<sup>2</sup> Mathematical Institute, University of Mannheim, B6, 26, D-68131 Mannheim, Germany

\* Correspondence: aneuenki@mail.uni-mannheim.de

**Abstract:** Inspired by the article *Weak Convergence Rate of a Time-Discrete Scheme for the Heston Stochastic Volatility Model*, Chao Zheng, *SIAM Journal on Numerical Analysis* 2017, 55:3, 1243–1263, we studied the weak error of discretization schemes for the Heston model, which are based on exact simulation of the underlying volatility process. Both for an Euler- and a trapezoidal-type scheme for the log-asset price, we established weak order one for smooth payoffs without any assumptions on the Feller index of the volatility process. In our analysis, we also observed the usual trade off between the smoothness assumption on the payoff and the restriction on the Feller index. Moreover, we provided error expansions, which could be used to construct second order schemes via extrapolation. In this paper, we illustrate our theoretical findings by several numerical examples.

**Keywords:** Heston model; discretization schemes for SDEs; exact simulation of the CIR process; Kolmogorov PDE; Malliavin calculus

**MSC:** 60H07; 60H35; 65C05; 91G60



**Citation:** Mickel, Annalena, and Andreas Neuenkirch. 2021. The Weak Convergence Rate of two Semi-Exact Discretization Schemes for the Heston Model. *Risks* 9: 23. <https://doi.org/10.3390/risks9010023>

Received: 15 October 2020  
Accepted: 28 December 2020  
Published: 12 January 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction and Main Results

The Heston Model [Heston \(1993\)](#) is a widely used stochastic volatility model to price financial options. It consists of two stochastic differential equations (SDEs) for an asset price process  $S$  and its volatility  $V$ :

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{V_t} S_t \left( \rho dW_t + \sqrt{1 - \rho^2} dB_t \right), \\ dV_t &= \kappa(\theta - V_t) dt + \sigma \sqrt{V_t} dW_t, \end{aligned} \quad (1)$$

with  $S_0, V_0, \kappa, \theta, \sigma > 0$ ,  $\mu \in \mathbb{R}$ ,  $\rho \in [-1, 1]$ ,  $T > 0$  and independent Brownian motions  $W = (W_t)_{t \in [0, T]}$ ,  $B = (B_t)_{t \in [0, T]}$ , which are defined on a filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \in [0, T]}, \mathbb{P})$ , where the filtration satisfies the usual conditions. It is a simple and popular extension of the Black–Scholes model where the volatility of the asset was assumed to be constant. 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 volatility is given by the so-called Cox–Ingersoll–Ross process (CIR). Its Feller index  $\nu = \frac{2\kappa\theta}{\sigma^2}$  will be an important parameter for our results. Throughout this article, the initial values  $S_0, V_0$  are assumed to be deterministic.

To price options with maturity at time  $T$ , one is interested in the value of

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

where  $g : [0, \infty) \rightarrow \mathbb{R}$  is the payoff function. Closed formulae for  $\mathbb{E}[g(S_T)]$  are rarely known and often Monte Carlo methods are applied, for which in turn the simulation of  $S_T$

is required. Usually, the log-Heston model instead of the Heston model is considered in numerical practice. This yields the SDE

$$\begin{aligned} d(\log(S_t)) &= \left(\mu - \frac{1}{2}V_t\right)dt + \sqrt{V_t}d\left(\rho W_t + \sqrt{1-\rho^2}B_t\right), \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t, \end{aligned} \quad (2)$$

and the exponential is then incorporated in the payoff, i.e.,  $g$  is replaced by  $f: \mathbb{R} \rightarrow \mathbb{R}$  with  $f(x) = g(\exp(x))$ .

While exact simulation schemes and their refinements are known (see, e.g., Broadie and Kaya (2006); Glasserman and Kim (2011); Malham and Wiese (2013); Smith (2007)), discretization schemes as, e.g., Altmayer and Neuenkirch (2017); Andersen (2008); Kahl and Jäckel (2006); Lord et al. (2009), are very popular for the Heston model. The latter discretization schemes can be easily extended to the multi-dimensional case and avoid computational bottlenecks of the exact schemes. In particular, Euler-type methods, such as the fully truncated Euler scheme, seem to be very efficient (see, e.g., Coskun and Korn (2018); Lord et al. (2009)), but no weak error analysis is available for them, up to the best of our knowledge.

A second order discretization scheme for the log-Heston model has been introduced in Andersen (2008) and analyzed in Zheng (2017). The so-called Broadie-Kaya trick and a removal of the drift, detailed in Section 3.1, reduce the simulation of the log-Heston model to the joint simulation of

$$\begin{aligned} dX_t &= \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)V_tdt + \sqrt{1-\rho^2}\sqrt{V_t}dB_t, \\ dV_t &= \kappa(\theta - V_t)dt + \sigma\sqrt{V_t}dW_t. \end{aligned} \quad (3)$$

Moreover, since the transition density of the CIR process  $V = (V_t)_{t \in [0, T]}$  follows a non-central chi-square distribution, it can be simulated exactly. Trapezoidal discretizations of the first component  $X = (X_t)_{t \in [0, T]}$  lead to the trapezoidal scheme

$$\begin{aligned} x_{k+1} &= x_k + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)\frac{v_{k+1} + v_k}{2}(t_{k+1} - t_k) \\ &\quad + \sqrt{1-\rho^2}\sqrt{\frac{v_{k+1} + v_k}{2}}\Delta_k B, \quad k = 0, \dots, N-1, \end{aligned}$$

where  $0 = t_0 < \dots < t_k < \dots < t_N = T$ ,  $v_k = V_{t_k}$  and  $\Delta_k B = B_{t_{k+1}} - B_{t_k}$ . This discretization avoids in particular the cumbersome exact simulation of the integrated volatility. Zheng (2017) establishes weak order two for polynomial test functions by transferring the error analysis to that of a trapezoidal rule for multidimensional deterministic integrals. Our original intention was to extend this result to a larger class of test functions  $f$  by using the Kolmogorov PDE approach. However, the required Itô-Taylor expansions turned out to be not feasible. So, instead, we analyzed the following two semi-exact discretization schemes: the Euler-type scheme

$$x_{k+1} = x_k + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)v_k(t_{k+1} - t_k) + \sqrt{1-\rho^2}\sqrt{v_k}\Delta_k B \quad (4)$$

and the semi-trapezoidal scheme

$$x_{k+1} = x_k + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)\frac{v_{k+1} + v_k}{2}(t_{k+1} - t_k) + \sqrt{1-\rho^2}\sqrt{v_k}\Delta_k B. \quad (5)$$

In both schemes, the CIR process is simulated exactly. In our opinion, the analysis of these schemes gives valuable insights in the weak error analysis of discretization schemes

for the log-Heston model and is also a good starting point for the analysis of full Euler-type discretization schemes.

Our error analysis relies on two regularity results for the Heston PDE (Briani et al. (2018); Feehan and Pop (2013)), the Kolmogorov PDE approach for the weak error analysis from Talay and Tubaro (1990), and Malliavin calculus. We also observe the usual trade off between the smoothness assumption on the payoff and the restriction on the Feller index. For payoffs of lower smoothness, a restriction on the Feller index  $\nu = 2\kappa\theta/\sigma^2$  is required, which arises from the use of Malliavin calculus tools.

In the following, we use the notation

$$\Delta t = \max_{k=1,\dots,N} |t_k - t_{k-1}|$$

for the maximal step size and the usual notations for the spaces of differentiable functions. In particular, the subscript  $c$  denotes compact support and  $pol$  denotes polynomial growth. In addition, see Section 3.1. The results of Feehan and Pop (2013) require compact support of the test functions  $f$ , while the results of Briani et al. (2018) allow polynomial growth but require higher smoothness for  $f$ .

**Theorem 1.** Let  $\varepsilon > 0$ . (i) If  $f \in C_c^{2+\varepsilon}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$  and  $\frac{2\kappa\theta}{\sigma^2} > \frac{3}{2}$ , then both schemes satisfy

$$\mathbb{E}[f(x_N, v_N)] - \mathbb{E}[f(X_T, V_T)] = O(\Delta t).$$

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

$$\mathbb{E}[f(x_N, v_N)] - \mathbb{E}[f(X_T, V_T)] = O(\Delta t).$$

Assuming more smoothness of  $f$ , we obtain more detailed results:

**Theorem 2.** Suppose that  $f \in C_{pol}^8(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ . (i) Then, the Euler scheme (4) satisfies

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

where

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

and

$$\hat{x}_t = x_n + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)v_n(t - t_n) + \sqrt{1 - \rho^2}\sqrt{v_n}(B_t - B_{t_n}), \quad t \in [t_n, t_{n+1}],$$

for  $n = 0, \dots, N - 1$ .

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

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

Here,  $u$  denotes the solution of the associated Kolmogorov PDE; see Equation (7).

(ii) For the semi-trapezoidal scheme (5), we have

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

where

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

and

$$\hat{x}_t = x_n + \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{V_t + v_n}{2} (t - t_n) + \sqrt{1 - \rho^2} \sqrt{v_n} (B_t - B_{t_n}), \quad t \in [t_n, t_{n+1}],$$

for  $n = 0, \dots, N - 1$ .

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

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

Here,  $u$  denotes again the solution of the associated Kolmogorov PDE; see Equation (7).

Thus, the semi-trapezoidal rule eliminates the first two terms of the error expansion of the Euler scheme.

Remarks

**Remark 1.** We expect that the error expansions for an equidistant discretization for both schemes satisfy

$$\mathbb{E}[f(x_N, v_N)] - \mathbb{E}[f(X_T, V_T)] = \left( \frac{T}{2} \int_0^T \mathbb{E}[\mathcal{H}(t, t, x_t, x_t, v_t, v_t)] dt \right) \cdot N^{-1} + O(N^{-2}). \quad (6)$$

However, to establish this, we would require error estimates for functionals of the type  $\mathbb{E}[f(\lambda, X_T, V_T)]$  with  $\lambda \in [0, T]$ , which are uniform in  $\lambda$ . (Compare, e.g., Proposition 2 in Talay and Tubaro (1990).) This, in turn, would require uniform regularity estimates for the Heston PDE, which are not available at the moment.

**Remark 2.** Property (6) allows to construct a second order scheme via extrapolation: If (6) holds, then

$$Y_N = 2f(x_{2N}, v_{2N}) - f(x_N, v_N)$$

satisfies

$$\mathbb{E}Y_N = \mathbb{E}f(X_T, V_T) + O((\Delta t)^2),$$

where  $(x_{2N}, v_{2N})$  uses the stepsize  $T/(2N)$  and  $(x_N, v_N)$  the stepsize  $T/N$ .

**Remark 3.** We require smoothness assumptions for  $f$  that are not met by the payoffs in practice, which are at most Lipschitz continuous or even discontinuous. However, this is a typical problem for weak approximation of SDEs as the Heston SDE, which do not satisfy the so-called standard assumptions on the coefficients. In Bally and Talay (1996), only bounded and measurable test functions  $f$  are treated assuming uniform hypoellipticity of the coefficients of the SDE. However, the Heston model does not satisfy this property. An adaptation of the strategy of Bally and Talay (1996) to the Heston model yields strong assumption on the Feller index (see Altmayer (2015)), which we want to avoid here.

**Remark 4.** Schemes built on the Broadie-Kaya trick, i.e., Equation (3), have a different structure than schemes which arise by a direct discretization of the log-Heston model as, e.g., the schemes

studied in [Altmayer and Neuenkirch \(2017\)](#); [Lord et al. \(2009\)](#). For example, the so-called absorbed Euler discretization reads as

$$z_{k+1} = z_k - \frac{1}{2}v_k(t_{k+1} - t_k) + \sqrt{v_k} \left( \rho(W_{t_{k+1}} - W_{t_k}) + \sqrt{1 - \rho^2}(B_{t_{k+1}} - B_{t_k}) \right),$$

$$v_{k+1} = (v_k + \kappa(\theta - v_k)(t_{k+1} - t_k) + \sigma\sqrt{v_k}(W_{t_{k+1}} - W_{t_k}))^+.$$

Here, the volatility  $V = (V_t)_{t \in [0, T]}$  is discretized by an Euler scheme, a fix for retaining the positivity is introduced by using the positive part, and the equation for the log-Heston price  $Z = (\log(S_t))_{t \in [0, T]}$  is discretized instead of the one for  $X = (X_t)_{t \in [0, T]}$ .

**Remark 5.** The Broadie-Kaya trick is a particular case of a more general transformation procedure, which has been introduced in [Cui et al. \(2018\)](#) for a general class of stochastic volatility models. In addition, in [Cui et al. \(2018\)](#), the weak convergence of a Markov chain approximation for these equations is established, which had been introduced in [Cui et al. \(2020\)](#). Markov chain approximations have been also studied in [Briani et al. \(2018\)](#) for the Heston and Bates model and are an alternative to a classical discretization of stochastic differential equations. In particular, for pricing American options, they can be beneficial.

## 2. Numerical Results

In this section, we will test numerically whether the convergence rates for the Euler Scheme (4) and the Semi-Trapezoidal scheme (5) are attained even under milder assumptions than those from Theorems 1 and 2. We use the following model parameters:

Model 1:  $S_0 = 100, V_0 = 0.010201, K = 100, \kappa = 6.21, \theta = 0.019, \sigma = 0.61, \rho = -0.7, T = 1, r = 0.0319$ ;

Model 2:  $S_0 = 100, V_0 = 0.09, K = 100, \kappa = 2, \theta = 0.09, \sigma = 1, \rho = -0.3, T = 5, r = 0.05$ ;

Model 3:  $S_0 = 100, V_0 = 0.0457, K = 100, \kappa = 5.07, \theta = 0.0457, \sigma = 0.48, \rho = -0.767, T = 2, r = 0.00$ .

The Feller index is  $\nu = \frac{2\kappa\theta}{\sigma^2} \approx 0.63$  in Model 1,  $\nu \approx 0.36$  in Model 2, and  $\nu \approx 2.01$  in Model 3. 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. Indicator:  $g_3(S_T) = e^{-rT} \mathbf{1}_{[0, K]}(S_T)$ .

Note that none of these payoffs satisfies the assumptions of our Theorems. Thus, the presented numerical experiments explore whether the Theorems are valid under milder assumptions. In order to measure the weak error rate, we simulated  $M = 2 \cdot 10^7$  independent copies  $g_i(s_N^{(j)})$ ,  $j = 1, \dots, M$ , of  $g_i(s_N)$  to estimate

$$\mathbb{E}(g_i(s_N))$$

by

$$p_{M,N} = \frac{1}{M} \sum_{j=1}^M g_i(s_N^{(j)})$$

for each combination of model parameters, functional and number of steps  $N \in \{2^1, \dots, 2^6\}$  where  $\Delta t = \frac{T}{N}$ . The number of Monte Carlo samples is chosen in such a way that the Monte Carlo error is sufficiently small enough, i.e., does not dominate the theoretically expected convergence rates. The Monte Carlo mean of these samples was then compared to a reference solution  $p_{\text{ref}}$ , i.e.,

$$e(N) = |p_{\text{ref}} - p_{M,N}|,$$

and the error  $e(N)$  is plotted in Figures 1–18. We then measured the rate of convergence, i.e., the decay rate of  $e(N)$ , by the slope of a least-squares fit in logarithmic coordinates. 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 Heston (1993) via the put-call-parity. The price of the digital option can be computed from the probability  $P_2$  given in Heston (1993); it equals  $e^{rT}(1 - P_2)$ . Additionally to the Euler and Semi-Trapezoidal scheme, we simulated the Trapezoidal scheme as in Zheng (2017) and the two extrapolation schemes from Remark 2. Moreover, to present a broader picture we estimated the weak error order of two Euler-type discretizations of the full Heston Model, the Full Truncation Euler (FTE) as in Lord et al. (2009), and the Symmetrized Euler as in Bossy and Diop (2015). To clarify things, we show two plots for each combination of model parameters and functional: one with the suspected order one schemes (Euler, Semi-Trapezoidal, FTE, and Symmetrized Euler) and one with the suspected order two schemes (Trapezoidal, Extrapolated Euler, and Extrapolated Semi-Trapezoidal).

### 2.1. Model 1

In Table 1, we can see the measured convergence rates for this model with a Feller index of  $\nu \approx 0.63$ . The associated plots are shown in Figures 1–6.

**Table 1.** Measured convergence rates Model 1.

Method	Call	Put	Indicator
Euler	1.5252	0.9492	1.1870
Semi-Trapezoidal	2.0174	0.2857	1.8343
FTE	1.5205	1.5205	1.2847
Symmetrized Euler	0.3693	0.3659	0.3250
Trapezoidal	2.0283	1.1119	2.4544
Euler extrap.	2.3114	2.0172	1.9719
Semi-Trapezoidal extrap.	1.8687	1.9999	0.9834

All “Order 1” schemes seem to have a very regular convergence behavior except for the Semi-Trapezoidal scheme for the Indicator, which could be explained by the low absolute error. Especially for the Call and the Indicator, both schemes from Theorem 1 seem to have very high weak convergence rates. Because of the Feller index of 0.63 in this model, this indicates that the assertion of Theorems 1 and 2 could hold under weaker assumptions. The extremely low estimated convergence rate for the Semi-Trapezoidal scheme in combination with the Put could be due to the low error. The estimated weak error order of the FTE scheme is noticeably higher than 1, whereas the Symmetrized Euler has low convergence rates. The convergence behavior of the “Order 2” schemes is a bit less regular. The Extrapolated Euler scheme seems to converge with order 2 for all payoff functions, whereas the Extrapolated Semi-Trapezoidal scheme seem to have only order 1 for the Indicator. But, again, we notice that the error for just 2 discretization steps already starts at around  $2^{-10}$ , which is extremely low.

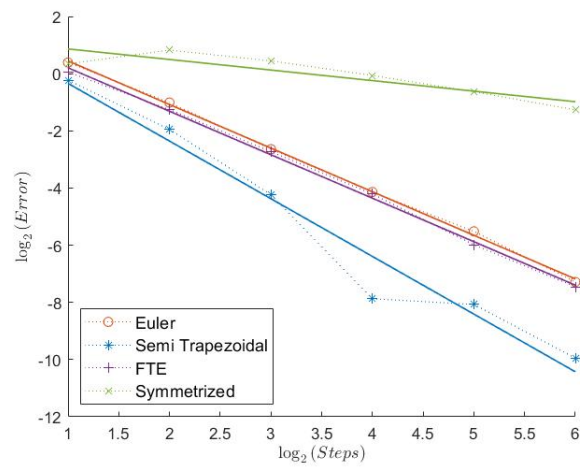


Figure 1. Call Model 1.

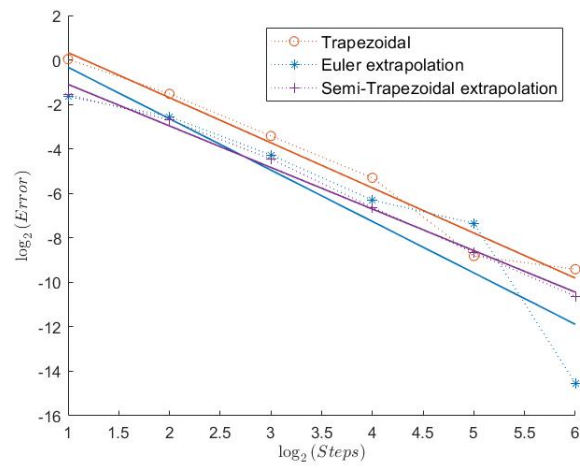


Figure 2. Call Model 1.

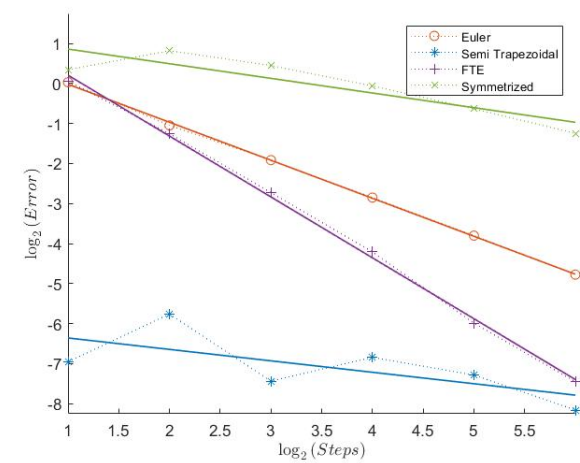


Figure 3. Put Model 1.

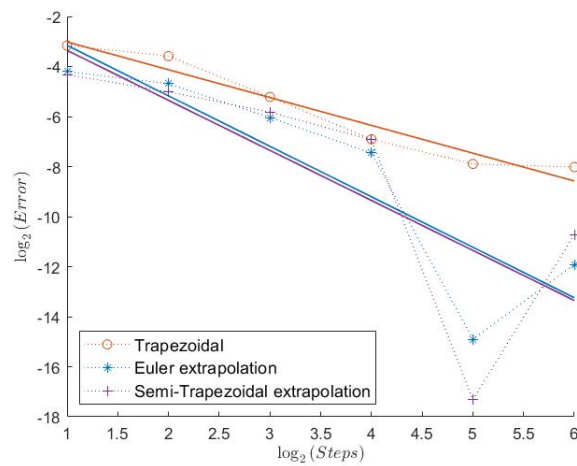


Figure 4. Put Model 1.

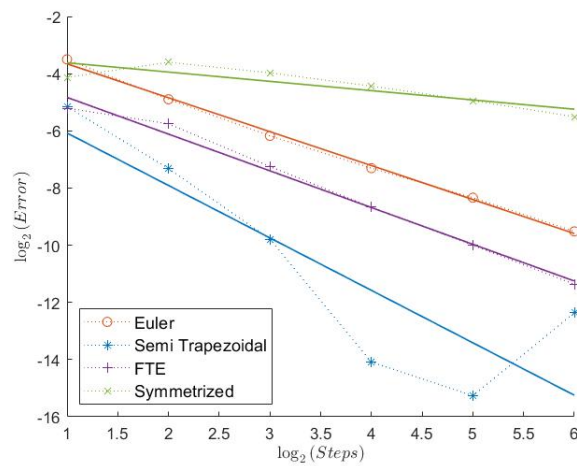


Figure 5. Indicator Model 1.

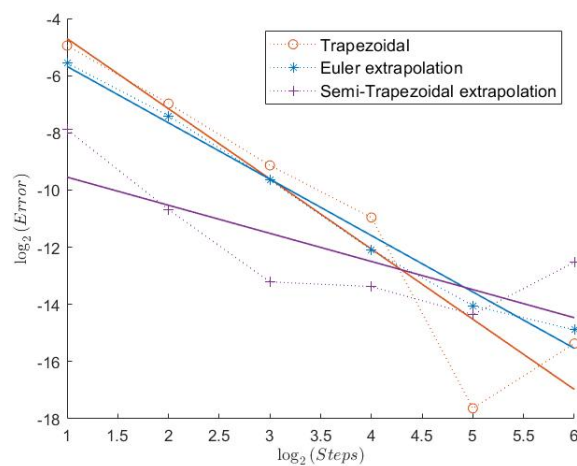


Figure 6. Indicator Model 1.

### 2.2. Model 2

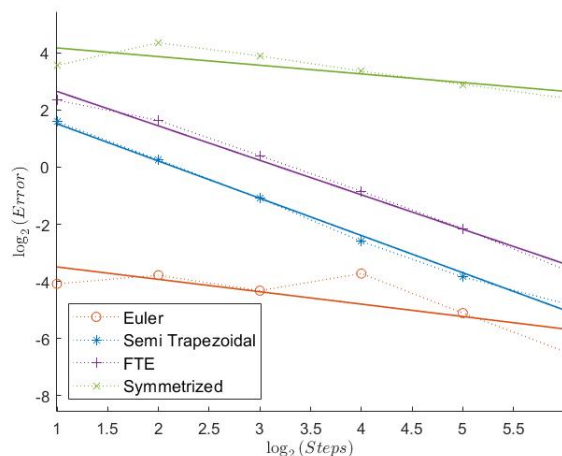
Here, we have an even lower Feller index of  $\nu \approx 0.36$ . We can see that the estimated convergence rates for all “Order 1” schemes are lower than before, see Table 2. However, the Semi-Trapezoidal scheme and the FTE scheme seem to converge with order 1. The convergence behavior is still quite regular as we can see in Figures 7, 9, and 11. In absolute



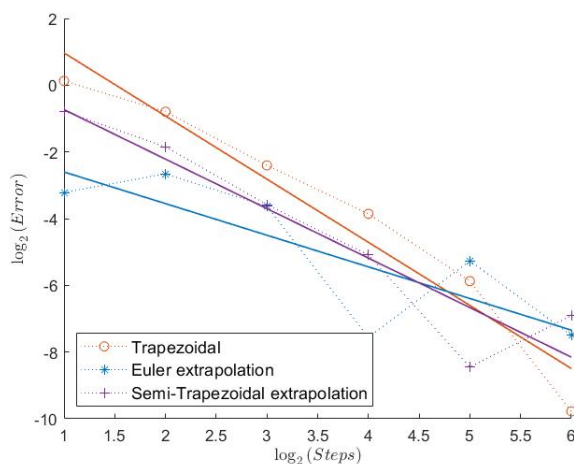
terms, the errors of the schemes from Theorem 1 are the lowest, especially for Put and Indicator. Looking at the “Order 2” schemes, the Trapezoidal discretization still shows an estimated weak convergence rate of around 2, whereas the two extrapolation schemes show a weaker performance. But, especially for the Indicator, all three schemes seem to have a very low error and a quite regular convergence behavior.

**Table 2.** Measured convergence rates Model 2.

Method	Call	Put	Indicator
Euler	0.4335	1.2898	0.8565
Semi-Trapezoidal	1.3025	0.7810	0.9518
FTE	1.2050	1.1733	1.0546
Symmetrized Euler	0.3028	0.3021	0.2421
Trapezoidal	1.8925	2.1272	1.6324
Euler extrap.	0.9483	1.4393	1.5966
Semi-Trapezoidal extrap.	1.4840	1.0481	1.2744



**Figure 7.** Call Model 2.



**Figure 8.** Call Model 2.

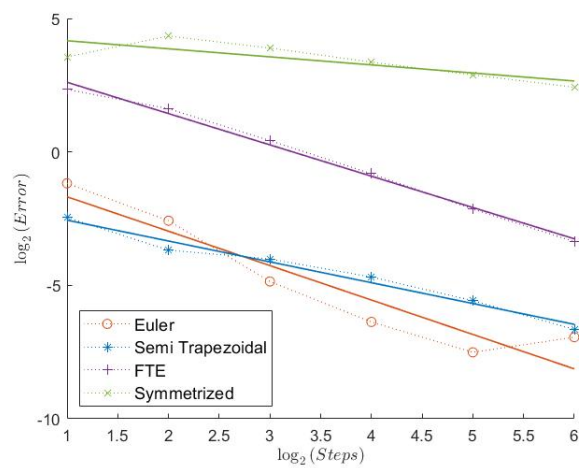


Figure 9. Put Model 2.

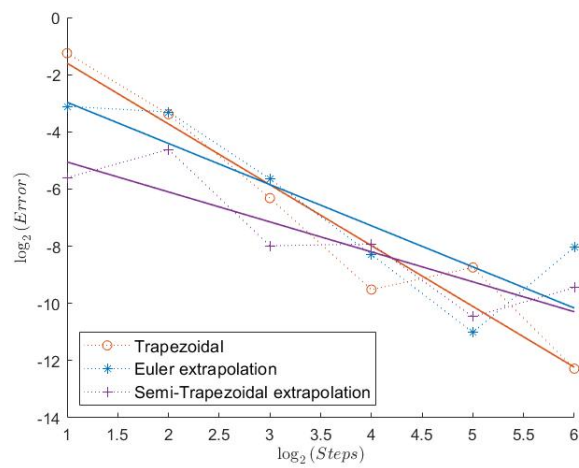


Figure 10. Put Model 2.

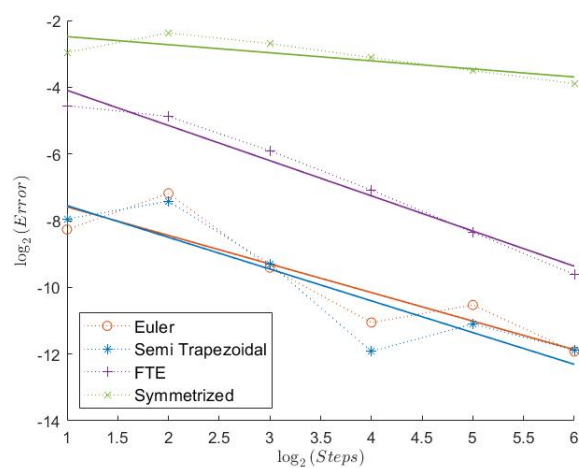


Figure 11. Indicator Model 2.

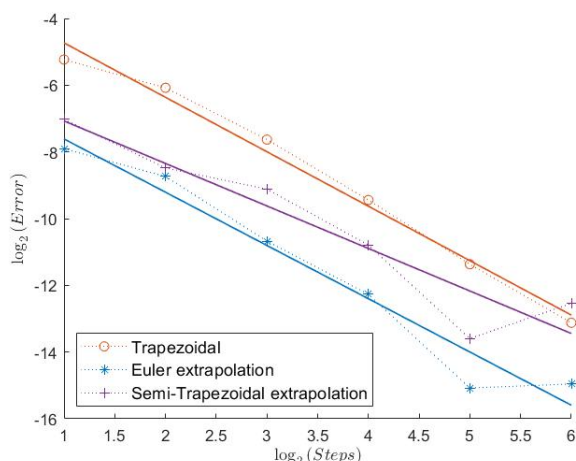


Figure 12. Indicator Model 2.

2.3. Model 3

Here, we have the highest Feller Index with  $\nu \approx 2.01$ . It is, therefore, a bit surprising that the Euler scheme seems to have a convergence rate of less than 1 in this case. In general, the errors for the “Order 1” schemes show a more irregular behavior, as can be seen from Figures 13, 15, and 17. The Semi-Trapezoidal and the FTE scheme work especially well in this scenario as we can see in Table 3. This is also the only case where the Symmetrized Euler shows an estimated convergence order of around 1. The extrapolation definitely improves the convergence rate of the Euler scheme with order 2 for the Indicator, but this is not the case for the Semi-Trapezoidal scheme.

Table 3. Measured convergence rates Model 3.

Method	Call	Put	Indicator
Euler	0.6977	0.5378	1.0695
Semi-Trapezoidal	1.6989	1.6551	1.6396
FTE	2.0091	1.7303	1.6008
Symmetrized Euler	1.0386	1.0426	0.9018
Trapezoidal	1.8682	1.6799	1.5219
Euler extrap.	1.1612	1.1857	2.2612
Semi-Trapezoidal extrap.	1.5660	1.0441	1.5979

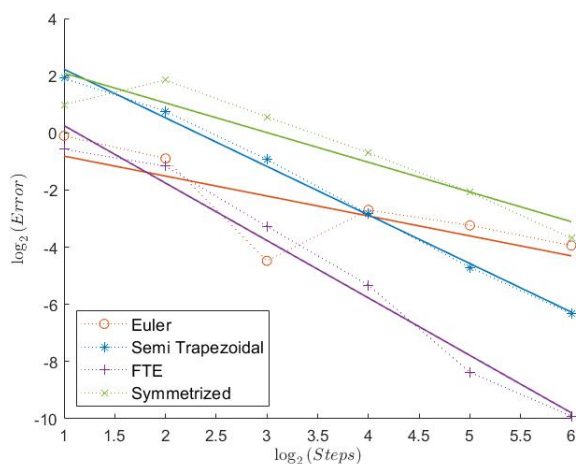


Figure 13. Call Model 3.

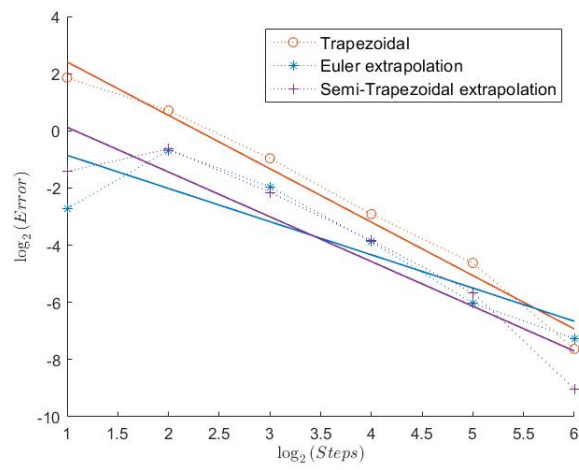


Figure 14. Call Model 3.

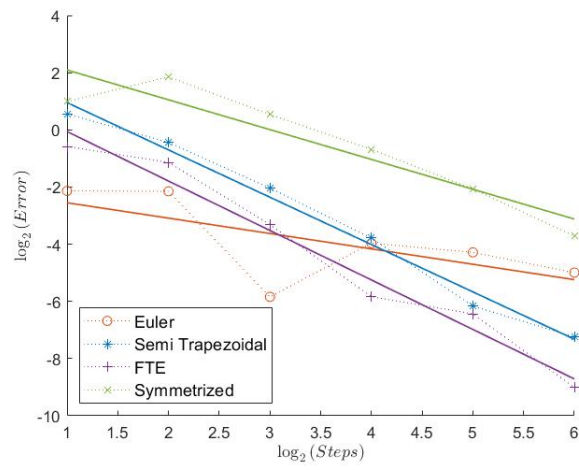


Figure 15. Put Model 3.

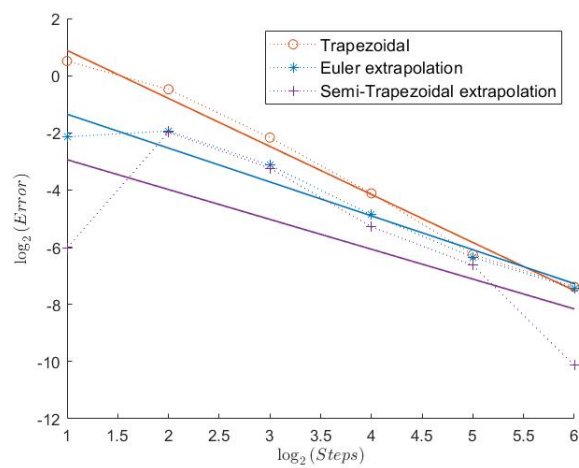


Figure 16. Put Model 3.

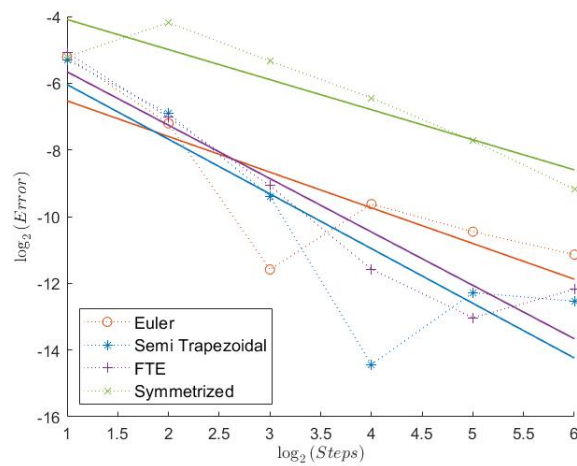


Figure 17. Indicator Model 3.

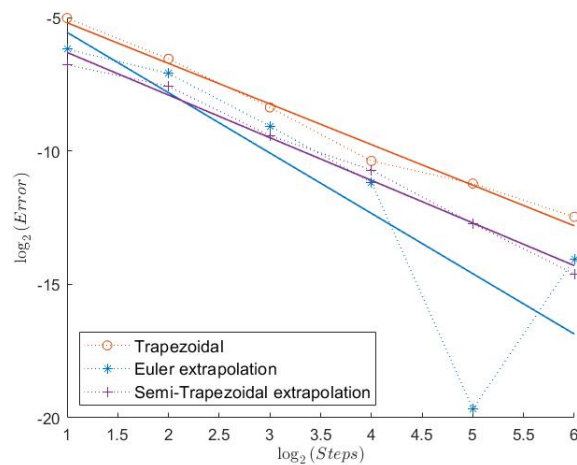


Figure 18. Indicator Model 3.

2.4. Computational Times

The computational times show the expected behavior, i.e., the simulation times for the semi-exact schemes increase as the Feller index decreases. See Tables 4 and 5. This is a well known feature of the MATLAB-generator ncx2rnd for the non-central chi-square distribution, which we used. (All simulations were carried out in MATLAB.)

Table 4. Computational times (sec.) of the semi-exact schemes for  $2^6$  time steps and  $2 \times 10^7$  paths.

Model	1	2	3
Euler	345.73	755.19	145.40
Semi-Trapezoidal	344.53	757.93	144.79
Trapezoidal	342.51	766.01	143.39
Euler extrap.	690.36	2335.94	307.62
Semi-Trapezoidal extrap.	686.55	2371.67	310.29

Table 5. Computational times (sec.) of Euler-type discretizations for  $2^6$  time steps and  $2 \times 10^7$  paths.

Model	1	2	3
FTE	142.3	138.37	141.53
Symmetrized Euler	141.64	140.98	141.67

### 2.5. Conclusions

Except for the Euler scheme for the Call in Model 3, the simulation studies support the conjecture that the convergence rates of Theorems 1 and 2 hold under weaker assumptions. For the mentioned behavior of the Euler scheme, we do not have an explanation, except the possibly pre-asymptotic step sizes. For the extrapolated schemes, which might have order two, the situation is less clear. Since the behavior of the trapezoidal scheme is regular, a too large Monte Carlo error seems an unlikely explanation. Explanations could be again the pre-asymptotic step sizes or, in fact, the non-smoothness of the considered payoffs.

### 3. Auxiliary Results

In this section, we will collect and establish, respectively, several auxiliary results for the weak error analysis.

#### 3.1. Kolmogorov PDE

Recall that the stochastic integral equations for the log-Heston model for  $0 \leq s < t \leq T$  read as

$$\begin{aligned}\log(S_t) &= \log(S_s) + \int_s^t \left( r - \frac{1}{2} V_u \right) du + \int_s^t \sqrt{V_u} d \left( \rho W_u + \sqrt{1 - \rho^2} B_u \right), \\ V_t &= V_s + \int_s^t \kappa(\theta - V_u) du + \sigma \int_s^t \sqrt{V_u} dW_u.\end{aligned}$$

Now, we apply the so-called Broadie-Kaya trick from [Broadie and Kaya \(2006\)](#). We can rearrange the second equation:

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

Then, we plug this equation into the first one:

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

Without loss of generality, we can neglect the non-integral part in  $\log(S_t) - \log(S_s)$ , since we have

$$f(\log(S_T), V_T) = f\left(X_T + \frac{\rho}{\sigma}(V_T - V_0 - \kappa\theta T + rT), V_T\right)$$

with  $X_T = X_T^{0, \log(S_0), V_0}$  given below. To get the Kolmogorov backward PDE, we look at the following integral equations:

$$\begin{aligned}V_t^{s,v} &= v + \int_s^t \kappa(\theta - V_r^{s,v}) dr + \sigma \int_s^t \sqrt{V_r^{s,v}} dW_r, \\ X_t^{s,x,v} &= x + \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \int_s^t V_r^{s,v} dr + \sqrt{1 - \rho^2} \int_s^t \sqrt{V_r^{s,v}} dB_r.\end{aligned}$$

We set

$$u(t, x, v) = \mathbb{E} \left[ f \left( X_T^{t,x,v}, V_T^{t,v} \right) \right], \quad t \in [0, T], \quad x \in \mathbb{R}, \quad v \geq 0$$

and obtain for  $f: \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$  bounded and continuous the Kolmogorov backward PDE by an application of the Feynman-Kac Theorem (see, e.g., Theorem 5.7.6 in [Karatzas and Shreve \(1991\)](#)):

$$\begin{aligned}
 u_t(t, x, v) &= - \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) v u_x(t, x, v) - \kappa(\theta - v) u_v(t, x, v) \\
 &\quad - \frac{v}{2} \left( (1 - \rho^2) u_{xx}(t, x, v) + \sigma^2 u_{vv}(t, x, v) \right), \quad t \in (0, T), x \in \mathbb{R}, v > 0, \\
 u(T, x, v) &= f(x, v), \quad x \in \mathbb{R}, v \geq 0.
 \end{aligned} \tag{7}$$

In our error analysis, we will follow the now classical approach of Talay and Tubaro (1990), which exploits the regularity of the Kolmogorov backward PDE. For the latter we will rely on the works of Feehan and Pop (2013) and Briani et al. (2018). To state these regularity results, we will need the following notation:

For a multi-index  $l = (l_1, \dots, l_d) \in \mathbb{N}^d$ , 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  $q \in \mathbb{N}$ . The set  $C^q(\mathcal{D}; \mathbb{R})$  is the set of all real-valued functions on  $\mathcal{D}$  which are  $q$ -times continuously differentiable. 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  $\varepsilon$ , and  $C_c^{q+\varepsilon}(\mathcal{D}; \mathbb{R})$  is the set of all functions from  $C^{q+\varepsilon}(\mathcal{D}; \mathbb{R})$ , who have compact support. Moreover,  $C_{pol}^q(\mathcal{D}; \mathbb{R})$  is the set of functions  $g \in C^q(\mathcal{D}; \mathbb{R})$  such that there exist  $C, a > 0$  for which

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

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

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

The work of Feehan and Pop deals with general degenerated parabolic equations and establishes a-priori regularity estimates for them. In the context of Equation (7), the main result of Feehan and Pop (2013), i.e., Theorem 1.1, reads as follows:

**Theorem 3.** *Let  $\varepsilon > 0$  and  $f \in C_c^{2+\varepsilon}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ . Then, there exists a constant  $c > 0$ , depending only on  $f, T, \rho, \kappa, \theta$  and  $\sigma$  such that the solution  $u$  of PDE (7) satisfies*

$$\begin{aligned}
 \sup_{(t,x,v) \in [0,T] \times \mathbb{R} \times [0,\infty)} (|u(t, x, v)| + |\partial_t u(t, x, v)| + |\partial_v u(t, x, v)| + |\partial_x u(t, x, v)|) &\leq c, \\
 \sup_{(t,x,v) \in [0,T] \times \mathbb{R} \times [0,1]} (|v \partial_{xx} u(t, x, v)| + |v \partial_{xv} u(t, x, v)| + |v \partial_{vv} u(t, x, v)|) &\leq c, \\
 \sup_{(t,x,v) \in [0,T] \times \mathbb{R} \times [1,\infty)} (|\partial_{xx} u(t, x, v)| + |\partial_{xv} u(t, x, v)| + |\partial_{vv} u(t, x, v)|) &\leq c.
 \end{aligned}$$

So, under the above assumptions on  $f$ , the solution  $u$  and the first order derivatives are bounded. Moreover, the second order derivatives are also bounded, if they are damped by  $v$  for  $v \in [0, 1]$ .

Assuming more smoothness on  $f$ , we can achieve more regularity for  $u$  using the above result, at least for the partial derivatives with respect to  $x$ . Set

$$\begin{aligned}
 \hat{u}(t, x, v) &:= u_x(t, x, v) = \mathbb{E} \left[ f_x(X_T^{t,x,v}, V_T^{t,v}) \right], \\
 \tilde{u}(t, x, v) &:= u_{xx}(t, x, v) = \mathbb{E} \left[ f_{xx}(X_T^{t,x,v}, V_T^{t,v}) \right].
 \end{aligned}$$

This is well defined: by continuity and boundedness of  $f_x$  and dominated convergence we have

$$\begin{aligned} u_x(t, x, v) &= \frac{\partial}{\partial x} \mathbb{E} \left[ f(X_T^{t,x,v}, V_T^{t,v}) \right] = \frac{\partial}{\partial x} \mathbb{E} \left[ f(x + Z_T^{t,v}, V_T^{t,v}) \right] \\ &= \lim_{\delta \rightarrow 0} \mathbb{E} \left[ \frac{f(x + \delta + Z_T^{t,v}, V_T^{t,v}) - f(x + Z_T^{t,v}, V_T^{t,v})}{\delta} \right] \\ &= \lim_{\delta \rightarrow 0} \int_0^1 \mathbb{E} \left[ f_x(x + \delta \lambda + Z_T^{t,v}, V_T^{t,v}) \right] d\lambda \\ &= \int_0^1 \mathbb{E} \left[ f_x(x + Z_T^{t,v}, V_T^{t,v}) \right] d\lambda \\ &= \mathbb{E} \left[ f_x(x + Z_T^{t,v}, V_T^{t,v}) \right] = \mathbb{E} \left[ f_x(X_T^{t,x,v}, V_T^{t,v}) \right] \end{aligned}$$

with

$$Z_T^{t,v} = \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \int_s^t V_r^{s,v} dr + \sqrt{1 - \rho^2} \int_s^t \sqrt{V_r^{s,v}} dB_r.$$

An analogous calculation for  $u_{xx}(t, x, v)$  shows that  $u_{xx}(t, x, v) = \mathbb{E} \left[ f_{xx}(X_T^{t,x,v}, V_T^{t,v}) \right]$ . Thus,  $u_{xx}$  is also bounded, if  $f \in C_c^{2+\varepsilon}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ . Moreover,  $\hat{u}$  fulfills the Kolmogorov backward PDE

$$\begin{aligned} \hat{u}_t(t, x, v) &= - \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) v \hat{u}_x(t, x, v) - \kappa(\theta - v) \hat{u}_v(t, x, v) \\ &\quad - \frac{v}{2} \left( (1 - \rho^2) \hat{u}_{xx}(t, x, v) + \sigma^2 \hat{u}_{vv}(t, x, v) \right), \quad t \in (0, T), x \in \mathbb{R}, v > 0, \\ \hat{u}(T, x, v) &= f_x(x, v), \quad x \in \mathbb{R}, v \geq 0, \end{aligned}$$

while  $\tilde{u}$  fulfills the same PDE with terminal condition

$$\tilde{u}(T, x, v) = f_{xx}(x, v), \quad x \in \mathbb{R}, v \geq 0.$$

Applying Theorem 3 now to  $\hat{u}$  and  $\tilde{u}$ , we obtain the following additional bounds (case (ii)) for the derivatives of  $u$ :

**Corollary 1.** (i) Let  $\varepsilon > 0$  and  $f \in C_c^{2+\varepsilon}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ . Then, there exists a constant  $c > 0$ , depending only on  $f, T, \rho, \kappa, \theta$  and  $\sigma$  such that the solution  $u$  of PDE (7) satisfies

$$\sup_{(t,x,v) \in [0,T] \times \mathbb{R} \times [0,\infty)} |\partial_{xx} u(t, x, v)| \leq c.$$

(ii) Let  $\varepsilon > 0$  and  $f \in C_c^{4+\varepsilon}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ . Then, there exists a constant  $c > 0$ , depending only on  $f, T, \rho, \kappa, \theta$  and  $\sigma$  such that the solution  $u$  of PDE (7) satisfies

$$\sup_{(t,x,v) \in [0,T] \times \mathbb{R} \times [0,\infty)} (|\partial_{xv} u(t, x, v)| + |\partial_{xx} u(t, x, v)| + |\partial_{xvv} u(t, x, v)| + |\partial_{xxx} u(t, x, v)|) \leq c.$$

The recent work of Briani et al. is a specialized approach for the log-Bates model, of which the log-Heston model is a particular case. In our setting, they obtain in Proposition 5.3 and Remark 5.4 of Briani et al. (2018) the following:

**Theorem 4.** Let  $q \in \mathbb{N}, q \geq 2$  and suppose that  $f \in C_{pol}^{2q}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ . Then, the solution  $u$  of PDE (7) satisfies  $u \in C_{pol,T}^q(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ .



In contrast to the results of Feehan and Pop, the result of Briani et al. requires more smoothness of  $f$  but allows polynomial growth instead of compact support.

### 3.2. Properties of the CIR Process

We recall here the following estimates for the CIR process, which are well known or can be found in [Hurd and Kuznetsov \(2008\)](#).

**Lemma 1.** (1) We have

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

for all  $p \geq 1$  and

$$\sup_{t \in [0, T]} \mathbb{E} V_t^p < \infty \quad \text{iff} \quad p > -\frac{2\kappa\theta}{\sigma^2}.$$

(2) For all  $p \geq 1$ , there exist constants  $c > 0$ , depending only on  $p, \kappa, \theta, \sigma, T$ , and  $V_0$ , such that

$$\mathbb{E} |V_t - V_s|^p \leq c \cdot |t - s|^{p/2}, \quad s, t \in [0, T].$$

We will need the following bound on the growth of the  $L^q$ -norm of a specific stochastic integral of the CIR process:

**Lemma 2.** For all  $q \in \left[2, \frac{4\kappa\theta}{\sigma^2}\right)$ , it holds that

$$\sup_{t \in [0, T]} t^{-q/2} \mathbb{E} \left[ \left| \int_0^t \frac{1}{\sqrt{V_u}} dB_u \right|^q \right] < \infty.$$

**Proof.** With the Burkholder-Davis-Gundy inequality and the Hölder inequality, we have

$$\begin{aligned} t^{-q/2} \mathbb{E} \left[ \left| \int_0^t \frac{1}{\sqrt{V_u}} dB_u \right|^q \right] &\leq t^{-q/2} \mathbb{E} \left[ \left| \int_0^t \frac{1}{V_u} du \right|^{q/2} \right] \\ &\leq t^{-q/2} \mathbb{E} \left[ \left( \left| \int_0^t \left( \frac{1}{V_u} \right)^{q/2} du \right|^{2/q} \left| \int_0^t dr \right|^{(q-2)/q} \right)^{q/2} \right] \\ &= t^{-q/2} \left( \int_0^t \mathbb{E} \left[ \left( \frac{1}{V_u} \right)^{q/2} \right] du \right) t^{(q-2)/2} \\ &\leq \sup_{u \in [0, t]} \mathbb{E} \left[ V_u^{-q/2} \right] \end{aligned}$$

for all  $t \in [0, T]$ . The assertion now follows from Lemma 1 (1).  $\square$

### 3.3. Malliavin Calculus

When working with low smoothness assumptions on  $f$ , we will use a Malliavin integration by parts procedure to establish weak convergence order one. As in [Altmayer and Neuenkirch \(2017\)](#), this paragraph gives a short introduction into Malliavin calculus; for more details, we refer to [Nualart \(1995\)](#).

Malliavin calculus adds a derivative operator to stochastic analysis. Basically, if  $Y$  is a random variable and  $(W_t, B_t)_{t \in [0, T]}$  a two-dimensional Brownian motion, then the Malliavin derivative measures the dependence of  $Y$  on  $(W, B)$ . The Malliavin derivative is defined by a standard extension procedure: Let  $\mathcal{S}$  be the set of smooth random variables of the form

$$S = \varphi \left( \int_0^T h_1(s) d(W_s, B_s), \dots, \int_0^T h_k(s) d(W_s, B_s) \right)$$

with  $\varphi \in C^\infty(\mathbb{R}^k; \mathbb{R})$  bounded with bounded derivatives,  $h_i \in L^2([0, T]; \mathbb{R}^2)$ ,  $i = 1, \dots, k$ , and the stochastic integrals

$$\int_0^T h_j(s) d(W_s, B_s) = \int_0^T h_j^{(1)}(s) dW_s + \int_0^T h_j^{(2)}(s) dB_s.$$

The derivative operator  $D$  of such a smooth random variable is defined as

$$DS = \sum_{i=1}^k \frac{\partial \varphi}{\partial x_i} \left( \int_0^T h_1(s) d(W_s, B_s), \dots, \int_0^T h_k(s) d(W_s, B_s) \right) h_i.$$

This operator is closable from  $L^p(\Omega)$  into  $L^p(\Omega; H)$  with  $H = L^2([0, T]; \mathbb{R}^2)$  and the Sobolev space  $\mathbb{D}^{1,p}$  denotes the closure of  $\mathcal{S}$  with respect to the norm

$$\|Y\|_{1,p} = \left( \mathbb{E}|Y|^p + \mathbb{E} \left| \int_0^T |D_s Y|^2 ds \right|^p \right)^{1/p}.$$

In particular, if  $D^W$  denotes the first component of the Malliavin derivative, i.e., the derivative with respect to  $W$ , we have

$$D_t^W Y = \begin{cases} 1_{[0,t]} & \text{if } Y = W \\ 0 & \text{if } Y = B \end{cases}$$

and vice versa for the derivative with respect to  $B$ , i.e.,

$$D_t^B Y = \begin{cases} 1_{[0,t]} & \text{if } Y = B \\ 0 & \text{if } Y = W \end{cases}$$

This, in particular, implies that, if  $Y \in \mathbb{D}^{1,2}$  is independent of  $B$ , then  $D^B Y = 0$ .

For the CIR process, we will, therefore, have that  $D^B V_t = 0$  for all  $t \in [0, T]$ .

The derivative operator follows rules similar to ordinary calculus.

**Proposition 1.** Let  $X = (X_1, \dots, X_d)$  be a random variable with components in  $\mathbb{D}^{1,p}$ . If

- (i)  $\phi : \mathbb{R}^d \rightarrow \mathbb{R}$  is in  $C^1(\mathbb{R}^d; \mathbb{R})$ ,
- (ii)  $\phi(X) \in L^p(\Omega)$ ,
- (iii)  $\partial_i \phi(X) \cdot DX_i \in L^p(\Omega; H)$  for all  $i = 1, \dots, d$ ,

then the chain rule holds:  $\phi(X) \in \mathbb{D}^{1,p}$  and

$$D\phi(X) = \sum_{i=1}^d \partial_i \phi(X) \cdot DX_i.$$

For example, for a random variable  $Y \in \mathbb{D}^{1,p}$  and  $g \in C^1(\mathbb{R}; \mathbb{R})$  with bounded derivative, the chain rule reads as

$$Dg(Y) = g'(Y) DY.$$

Another simple example for the application of this chain rule is

$$D_r^W [(W_t - W_s)^2] = 2(W_t - W_s) \mathbf{1}_{(s,t]}(r), \quad r, s, t \in [0, T], s \leq t.$$

The divergence operator  $\delta$  is the adjoint of the derivative operator. If a random variable  $u \in L^2(\Omega; L^2([0, T]; \mathbb{R}^2))$  belongs to  $\text{dom}(\delta)$ , the domain of the divergence operator, then  $\delta(u)$  is defined by the duality—also called integration by parts—relationship

$$\mathbb{E}[Y\delta(u)] = \mathbb{E} \left[ \int_0^T \langle D_s Y, u_s \rangle ds \right] \quad \text{for all } Y \in \mathbb{D}^{1,2}. \tag{8}$$

If  $u$  is adapted to the canonical filtration generated by  $(W, B)$  and satisfies  $E \int_0^T |u_t|^2 dt < \infty$ , then  $u \in \text{dom}(\delta)$  and  $\delta(u)$  coincides with the Itô integral  $\int_0^T u_1(s) dW_s + \int_0^T u_2(s) dB_s$ .

For the Malliavin regularity of the CIR process, the following is well known. See, e.g., Proposition 4.5 and Theorem 4.6 in Altmayer (2015) or Proposition 4.1 in Alos and Ewald (2008).

**Lemma 3.** Let  $t \in [0, T]$  and  $\frac{2\kappa\theta}{\sigma^2} > 1$ . Then, we have  $\sqrt{V_t} \in \mathbb{D}^{1,\infty}$  and  $V_t \in \mathbb{D}^{1,\infty}$  with

$$D_r(\sqrt{V_t}) = \frac{\sigma}{2} \exp\left(\int_r^t \left(\left(\frac{\sigma^2}{8} - \frac{\kappa\theta}{2}\right) \frac{1}{V_u} - \frac{\kappa}{2}\right) du\right) \mathbf{1}_{[0,t]}(r), \quad r \in [0, T].$$

In Altmayer and Neuenkirch (2015), this and the integration by parts formula was used to establish

$$\mathbb{E}(g(X_T)) = \frac{1}{T\sqrt{1-\rho^2}} \cdot \mathbb{E}\left(G(X_T) \cdot \int_0^T \frac{1}{\sqrt{V_t}} dB_t\right),$$

under the assumption  $\frac{2\kappa\theta}{\sigma^2} > 1$  with  $G : \mathbb{R} \rightarrow \mathbb{R}$  differentiable and  $g = G'$  bounded, see Proposition 4.1 in Altmayer and Neuenkirch (2015). Indeed, using  $u_t = 1/\sqrt{V_t}$  and  $E \int_0^T |u_t|^2 dt < \infty$ ,  $D_r^B X_t = \sqrt{1-\rho^2} \sqrt{V_t} \mathbf{1}_{[0,t]}(r)$  and the chain rule, i.e.,

$$D_r^B G(X_T) = g(X_T) D_r^B X_T,$$

we have

$$\begin{aligned} \mathbb{E}\left(G(X_T) \cdot \int_0^T \frac{1}{\sqrt{V_t}} dB_t\right) &= \mathbb{E}\left(\int_0^T g(X_T) \cdot D_t^B X_T \cdot \frac{1}{\sqrt{V_t}} dt\right) \\ &= \mathbb{E}\left(\int_0^T g(X_T) \sqrt{1-\rho^2} \sqrt{V_t} \cdot \frac{1}{\sqrt{V_t}} dt\right) = T\sqrt{1-\rho^2} \mathbb{E}[g(X_T)], \end{aligned}$$

where the first equality is due to the integration by parts formula.

In Lemmas 5 and 9, we will establish discrete counterparts for this integration by parts result, i.e., on the level of the approximation schemes. In this context, we will also need the Malliavin differentiability of  $\int_s^t \sqrt{V_u} dW_u$ . Since

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

we obtain

$$\begin{aligned} D_r^W \left( \int_s^t \sqrt{V_u} dW_u \right) &= \frac{1}{\sigma} \left( D_r^W (V_t - V_s) + \kappa \int_s^t D_r^W V_u du \right), \\ D_r^B \left( \int_s^t \sqrt{V_u} dW_u \right) &= 0, \end{aligned}$$

by exchanging the Riemann integral and the Malliavin derivative (via a standard approximation argument for the Riemann integral, Lemma 3 and Lemma 1.2.3 in Nualart (1995)) and the independence of  $(V, W)$  and  $B$ . Thus, we can conclude that

$$\int_s^t \sqrt{V_u} dW_u \in \mathbb{D}^{1,\infty}, \quad 0 \leq s < t \leq T. \tag{9}$$

### 3.4. Properties of the Euler Discretization

Recall that the Euler discretization of the price process is given by

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

with  $\Delta_k B = B_{t_{k+1}} - B_{t_k}$ . We extend this discretization in every interval  $[t_n, t_{n+1}]$  as the following Itô process:

$$\hat{x}_t = x_{n(t)} + \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \int_{\eta(t)}^t v_{n(t)} ds + \sqrt{1 - \rho^2} \int_{\eta(t)}^t \sqrt{v_{n(t)}} dB_s.$$

Here, we have set  $n(t) := \max\{n \in \{0, \dots, N\} : t_n \leq t\}$ ,  $\eta(t) := t_{n(t)}$  and  $v_k = V_{t_k}$ .

We have the following result on the Malliavin regularity of the Euler discretization:

**Lemma 4.** *Let  $t \in [0, T]$  and  $\frac{2\kappa\theta}{\sigma^2} > 1$ . Then,  $\hat{x}_t \in \mathbb{D}^{1,\infty}$ , and we have*

$$D_r^B \hat{x}_t = \sqrt{1 - \rho^2} \sqrt{v_{n(t)}} \mathbf{1}_{[0,t]}(r).$$

**Proof.** We have

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

and

$$\hat{x}_{\eta(t)} = \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \sum_{k=0}^{n(t)-1} v_k(t_{k+1} - t_k) + \sqrt{1 - \rho^2} \sum_{k=0}^{n(t)-1} \sqrt{v_k}(B_{t_{k+1}} - B_{t_k}).$$

Following the steps of the proof of Lemma 3.5 from Altmayer and Neuenkirch [Altmayer and Neuenkirch \(2017\)](#), we then have  $\hat{x}_t \in \mathbb{D}^{1,\infty}$  exploiting that  $\sqrt{V_t} \in \mathbb{D}^{1,\infty}$  and  $V_t \in \mathbb{D}^{1,\infty}$  under the assumption  $\frac{2\kappa\theta}{\sigma^2} > 1$ . The chain rule from Proposition 1 yields

$$D_r^B \hat{x}_{\eta(t)} = \sqrt{1 - \rho^2} \sum_{k=0}^{n(t)-1} \sqrt{v_k} \mathbf{1}_{(t_k, t_{k+1}]}(r)$$

and

$$D_r^B \hat{x}_t = D_r^B \hat{x}_{\eta(t)} + \sqrt{1 - \rho^2} \sqrt{v_{n(t)}} \mathbf{1}_{(t_{n(t)}, t]}(r).$$

□

Note that we write, in the following,  $v_t$  instead of  $V_t$  to unify the notation. With the above result, we can express  $\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_{xx}(t, \hat{x}_t, v_t) \right]$  without the second order derivative of  $u$ , which will be needed later on.

**Lemma 5.** *Let  $t \in [0, T]$ . Under the assumptions of Theorem 3 and  $\frac{2\kappa\theta}{\sigma^2} > 1$ , we have*

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_{xx}(t, \hat{x}_t, v_t) \right] = \frac{1}{t\sqrt{1 - \rho^2}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_t, v_t) \int_0^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \right].$$

**Proof.** To avoid stronger restrictions on the Feller index we will use a localization procedure. So, for  $\varepsilon > 0$ , let  $\psi_\varepsilon$  be a function such that

1.  $\psi_\varepsilon : \mathbb{R} \rightarrow \mathbb{R}$  is continuously differentiable with bounded derivative,
2.  $0 \leq \psi_\varepsilon(x) \leq 1$  on  $[0, \infty)$ ,
3.  $\psi_\varepsilon(x) = 1$  on  $[2\varepsilon, \infty)$ ,
4.  $\psi_\varepsilon(x) = 0$  on  $(-\infty, \varepsilon]$ .

Since  $(V, W)$  and  $B$  are independent, the chain rule from Proposition 1 implies

$$D_r^B \left( \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi_\varepsilon(v_t) u_x(t, \hat{x}_t, v_t) \right) = \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi_\varepsilon(v_t) u_{xx}(t, \hat{x}_t, v_t) D_r^B \hat{x}_t$$

with  $D_r^B \hat{x}_t = \sqrt{1 - \rho^2} \sqrt{v_{\eta(r)}} \mathbf{1}_{[0,t]}(r)$ . Recall the integration by parts formula from Equation (8), i.e.,

$$\mathbb{E} \left[ Y \left( \int_0^T u_1(s) dW_s + \int_0^T u_2(s) dB_s \right) \right] = \mathbb{E} \left[ \int_0^T \langle D_s Y, u_s \rangle ds \right],$$

where we now choose

$$D_r Y = \left( \begin{array}{c} D_r^W \left( \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi_\varepsilon(v_t) u_x(t, \hat{x}_t, v_t) \right) \\ D_r^B \left( \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi_\varepsilon(v_t) u_x(t, \hat{x}_t, v_t) \right) \end{array} \right), \quad u_r = \left( \begin{array}{c} 0 \\ \frac{1}{\sqrt{v_{\eta(r)}}} \mathbf{1}_{[0,t]}(r) \end{array} \right).$$

Before we can apply the integration by parts rule, we need to check whether

$$\begin{aligned} & \int_0^T \mathbb{E} \left[ \left| D_r^W \left( \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) \psi_\varepsilon(v_t) u_x(t, \hat{x}_t, v_t) \right|^2 \right] dr < \infty, \\ & \int_0^T \mathbb{E} \left[ \left| \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi'_\varepsilon(v_t) u_x(t, \hat{x}_t, v_t) D_r^W v_t \right|^2 \right] dr < \infty, \\ & \int_0^T \mathbb{E} \left[ \left| \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi_\varepsilon(v_t) \left( u_{xx}(t, \hat{x}_t, v_t) D_r^W \hat{x}_t + u_{xv}(t, \hat{x}_t, v_t) D_r^W v_t \right) \right|^2 \right] dr < \infty, \\ & \int_0^T \mathbb{E} \left[ \left| \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi_\varepsilon(v_t) u_{xx}(t, \hat{x}_t, v_t) D_r^B \hat{x}_t \right|^2 \right] dr < \infty, \end{aligned} \tag{10}$$

for  $t > 0$ . We deduced these terms by using again the chain rule for  $D_r Y$ . Note that the properties of the localizing function and Theorem 3 imply that

$$\psi_\varepsilon(v) u_x(t, x, v), \quad \psi'_\varepsilon(v) u_x(t, x, v), \quad \psi_\varepsilon(v) u_{xx}(t, x, v), \quad \psi_\varepsilon(v) u_{xv}(t, x, v)$$

are all uniformly bounded in  $(t, x, v)$ . So, Equation (10) holds, then, due to Lemma 1, Lemma 3, Equation (9), and Lemma 4.

Since  $\int_0^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r$  is also well-defined by Lemma 1 due to  $\frac{2x\theta}{\sigma^2} > 1$ , we obtain now

$$\begin{aligned} & \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi_\varepsilon(v_t) u_{xx}(t, \hat{x}_t, v_t) \right] \\ &= \frac{1}{t} \mathbb{E} \left[ \int_0^t \left( \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi_\varepsilon(v_t) u_{xx}(t, \hat{x}_t, v_t) \right) dr \right] \\ &= \frac{1}{t \sqrt{1 - \rho^2}} \mathbb{E} \left[ \int_0^t D_r^B \left( \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi_\varepsilon(v_t) u_x(t, \hat{x}_t, v_t) \right) \frac{1}{\sqrt{v_{\eta(r)}}} dr \right] \\ &= \frac{1}{t \sqrt{1 - \rho^2}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \psi_\varepsilon(v_t) u_x(t, \hat{x}_t, v_t) \int_0^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \right]. \end{aligned}$$

Due to Corollary 1 (i), not only  $u_x$  but also  $u_{xx}$  is bounded. Since  $\psi_\varepsilon(v_t) \rightarrow 1$  almost surely for  $\varepsilon \rightarrow 0$  and  $|\psi_\varepsilon(v_t)| \leq 1$  for all  $\varepsilon > 0$ , the assertion follows now by dominated convergence using the Itô-isometry and again Lemma 1.  $\square$

We also need the following  $L^p$ -convergence result:

**Lemma 6.** *Let  $p \geq 1$ . There exists a constant  $c > 0$ , depending only on  $p, T, \rho, \kappa, \theta, \sigma$  and  $v_0$ , such that*

$$\sup_{t \in [0, T]} \mathbb{E}|X_t - \hat{x}_t|^p \leq c \cdot (\Delta t)^{p/4}.$$

**Proof.** We have

$$X_t - \hat{x}_t = \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) \int_0^t (v_u - v_{\eta(u)}) du + \sqrt{1 - \rho^2} \int_0^t (\sqrt{v_u} - \sqrt{v_{\eta(u)}}) dB_u.$$

Assume without loss of generality that  $p \geq 2$ . Jensen’s inequality and the Burkholder-Davis-Gundy inequality now imply that there exists a constant  $c > 0$ , depending only on  $p, T$ , the parameters of the CIR process, and  $v_0$ , such that

$$\mathbb{E}|X_t - \hat{x}_t|^p \leq c \int_0^t \mathbb{E}|v_u - v_{\eta(u)}|^p du + c \int_0^t \mathbb{E}|\sqrt{v_u} - \sqrt{v_{\eta(u)}}|^p du.$$

Since  $|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|}$  for  $x, y \geq 0$ , the assertion follows from Lemma 1.  $\square$

Straightforward calculations also yield the following  $L^p$ -smoothness result for the Euler-type scheme:

**Lemma 7.** *Let  $p \geq 1$ . There exists a constant  $c > 0$ , depending only on  $p, T, \rho, \kappa, \theta, \sigma$ , and  $v_0$ , such that*

$$\mathbb{E}|\hat{x}_t - \hat{x}_s|^p \leq c \cdot |t - s|^{p/2}$$

for all  $s, t \in [0, T]$ .

### 3.5. Properties of the Semi-Trapezoidal Rule

Recall that our semi-trapezoidal rule reads as

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

Again, we write the scheme as a time-continuous process:

$$\begin{aligned} \hat{x}_t &= x_{n(t)} + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) v_{n(t)} (t - \eta(t)) + \sqrt{1 - \rho^2} \sqrt{v_{n(t)}} (B_t - B_{\eta(t)}) \\ &\quad + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) \frac{1}{2} (v_t - v_{n(t)}) (t - \eta(t)). \end{aligned}$$

Expanding the last term with Itô’s lemma, we obtain

$$\hat{x}_t = x_{n(t)} + \int_{\eta(t)}^t a_s ds + \int_{\eta(t)}^t b_s dB_s + \int_{\eta(t)}^t c_s dW_s$$

with

$$\begin{aligned}
 a_t &:= \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \left( v_{n(t)} + \frac{1}{2}(t - \eta(t))\kappa(\theta - v_t) + \frac{1}{2}(v_t - v_{n(t)}) \right), \\
 b_t &:= \sqrt{1 - \rho^2} \sqrt{v_{n(t)}}, \\
 c_t &:= \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{1}{2}(t - \eta(t))\sigma \sqrt{v_t}.
 \end{aligned}$$

Here, we have set again  $n(t) := \max\{n \in \{0, \dots, N\} : t_n \leq t\}$ ,  $\eta(t) := t_{n(t)}$ ,  $v_k = V_{t_k}$ , and we also write again  $v_t$ , instead of  $V_t$ , to unify the notation.

We need the following result on the Malliavin regularity of the semi-trapezoidal scheme:

**Lemma 8.** *Let  $t \in [0, T]$  and  $\frac{2\kappa\theta}{\sigma^2} > 1$ . Then, we have  $\hat{x}_t \in \mathbb{D}^{1,\infty}$  and*

$$D_r^B \hat{x}_t = \sqrt{1 - \rho^2} \sqrt{v_{\eta(t)}} \mathbf{1}_{[0,t]}(r).$$

**Proof.** We already know that  $v_t \in \mathbb{D}^{1,\infty}$  and  $\sqrt{v_t} \in \mathbb{D}^{1,\infty}$ . We can write  $\hat{x}_t$  as

$$\hat{x}_t = \hat{x}_{\eta(t)} + \frac{1}{2} \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) (v_t + v_{\eta(t)}) (t - \eta(t)) + \sqrt{1 - \rho^2} \int_{\eta(t)}^t \sqrt{v_{\eta(t)}} dB_s$$

with

$$\hat{x}_{\eta(t)} = \frac{1}{2} \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \sum_{k=0}^{n(t)-1} (v_{k+1} + v_k) (t_{k+1} - t_k) + \sqrt{1 - \rho^2} \sum_{k=0}^{n(t)-1} \sqrt{v_k} (B_{t_{k+1}} - B_{t_k}).$$

Following the steps of the proof of Lemma 3.5 from Altmayer and Neuenkirch (2017), we then also have  $\hat{x}_t \in \mathbb{D}^{1,\infty}$ . The chain rule from Proposition 1 yields

$$D_r^B \hat{x}_{\eta(t)} = \sqrt{1 - \rho^2} \sum_{k=0}^{n(t)-1} \sqrt{v_k} \mathbf{1}_{(t_k, t_{k+1}]}(r)$$

and

$$D_r^B \hat{x}_t = D_r^B \hat{x}_{\eta(t)} + \sqrt{1 - \rho^2} \sqrt{v_{n(t)}} \mathbf{1}_{(t_{n(t)}, t]}(r).$$

□

Note that the partial Malliavin derivative with respect to  $B$  for the Euler and the semi-trapezoidal scheme coincide. So, by analogous calculations as for the Euler scheme, we obtain the following integration by parts result:

**Lemma 9.** *Let  $t \in [0, T]$ . Under the assumptions of Theorem 3 and  $\frac{2\kappa\theta}{\sigma^2} > 1$ , we have*

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_{xx}(t, \hat{x}_t, v_t) \right] = \frac{1}{t \sqrt{1 - \rho^2}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_t, v_t) \int_0^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \right].$$

By similar calculations as for the Euler scheme, we also have:

**Lemma 10.** *Let  $p \geq 1$ . There exists a constant  $c > 0$ , depending only on  $p, T, \rho, \kappa, \theta, \sigma$ , and  $v_0$ , such that*

$$\sup_{t \in [0, T]} \mathbb{E} |X_t - \hat{x}_t|^p \leq c \cdot (\Delta t)^{p/4}.$$

**Lemma 11.** Let  $p \geq 1$ . There exists a constant  $c > 0$ , depending only on  $p, T, \rho, \kappa, \theta, \sigma$ , and  $v_0$ , such that

$$\mathbb{E}|\hat{x}_t - \hat{x}_s|^p \leq c \cdot |t - s|^{p/2}$$

for all  $s, t \in [0, T]$ .

#### 4. Proof of Theorem 1

We address both schemes and the different assumptions in separate subsections. Constants, which are, in particular, independent of the maximal stepsize

$$\Delta t = \max_{k=1, \dots, N} |t_k - t_{k-1}|,$$

and depend only  $f, T, \rho, \kappa, \theta, \sigma$ , and  $v_0, x_0$ , will be denoted by  $c$ , regardless of their value.

##### 4.1. The Euler Scheme: Expanding the Error

Since  $u(T, x_N, v_N) = \mathbb{E}f(x_N, v_N)$  and  $u(0, x_0, v_0) = \mathbb{E}f(X_T, V_T)$ , the weak error is a telescoping sum of local errors:

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

With the Itô formula and the Kolmogorov backward PDE evaluated at  $(t, \hat{x}_t, v_t)$ , we obtain

$$\begin{aligned} e_n &:= \mathbb{E}[u(t_{n+1}, x_{n+1}, v_{n+1}) - u(t_n, x_n, v_n)] \\ &= \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ u_t(t, \hat{x}_t, v_t) + \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) v_{n(t)} u_x(t, \hat{x}_t, v_t) + \kappa(\theta - v_t) u_v(t, \hat{x}_t, v_t) \right. \\ &\quad \left. + \frac{1}{2} v_{n(t)} (1 - \rho^2) u_{xx}(t, \hat{x}_t, v_t) + \frac{1}{2} v_t \sigma^2 u_{vv}(t, \hat{x}_t, v_t) \right] dt \\ &= \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) (v_{n(t)} - v_t) u_x(t, \hat{x}_t, v_t) + \frac{1}{2} (v_{n(t)} - v_t) (1 - \rho^2) u_{xx}(t, \hat{x}_t, v_t) \right] dt. \end{aligned}$$

Since

$$v_{n(t)} - v_t = - \int_{\eta(t)}^t \kappa(\theta - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s,$$

we have  $e_n = e_n^{(1)} + e_n^{(2)}$  with

$$\begin{aligned} e_n^{(1)} &:= \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \left( - \int_{\eta(t)}^t \kappa(\theta - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) u_x(t, \hat{x}_t, v_t) \right] dt, \\ e_n^{(2)} &:= \frac{1}{2} (1 - \rho^2) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \left( - \int_{\eta(t)}^t \kappa(\theta - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) u_{xx}(t, \hat{x}_t, v_t) \right] dt. \end{aligned}$$

By Theorem 3 and Corollary 1, we have that  $u_x$  and  $u_{xx}$  are bounded. So, Lemma 1 implies that

$$\int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \kappa(\theta - v_s) ds u_x(t, \hat{x}_t, v_t) \right] dt = O((\Delta t)^2)$$

and

$$\int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \kappa(\theta - v_s) ds u_{xx}(t, \hat{x}_t, v_t) \right] dt = O((\Delta t)^2).$$



Moreover, with the law of total expectation and the adaptedness of  $\hat{x}_{\eta(t)}$  and  $v_{\eta(t)}$ , we have

$$\begin{aligned} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \right] &= \mathbb{E} \left[ \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \middle| \mathcal{F}_{\eta(t)} \right] \right] \\ &= \mathbb{E} \left[ u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \middle| \mathcal{F}_{\eta(t)} \right] \right] \\ &= 0, \end{aligned}$$

due to the martingale property of the Itô integral. Therefore, we can write

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_t, v_t) \right] = \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s (u_x(t, \hat{x}_t, v_t) - u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)})) \right]$$

and obtain

$$e_n^{(1)} = O((\Delta t)^2) - \left( \rho\kappa - \frac{\sigma}{2} \right) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s (u_x(t, \hat{x}_t, v_t) - u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)})) \right] dt.$$

In the same way, we have

$$e_n^{(2)} = O((\Delta t)^2) - \frac{\sigma}{2} (1 - \rho^2) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s (u_{xx}(t, \hat{x}_t, v_t) - u_{xx}(t, \hat{x}_{\eta(t)}, v_{\eta(t)})) \right] dt.$$

Summarizing this preliminary part, we have obtained

$$e_n = O((\Delta t)^2) + \tilde{e}_n^{(1)} + \tilde{e}_n^{(2)}, \tag{11}$$

where

$$\tilde{e}_n^{(1)} = - \left( \rho\kappa - \frac{\sigma}{2} \right) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s (u_x(t, \hat{x}_t, v_t) - u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)})) \right] dt, \tag{12}$$

$$\tilde{e}_n^{(2)} = - \frac{\sigma}{2} (1 - \rho^2) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s (u_{xx}(t, \hat{x}_t, v_t) - u_{xx}(t, \hat{x}_{\eta(t)}, v_{\eta(t)})) \right] dt. \tag{13}$$

#### 4.2. The Euler Scheme: Case (i)

So, it remains to analyze  $\tilde{e}_n^{(1)}$  and  $\tilde{e}_n^{(2)}$  under the regularity of Theorem 3 (i). We start with  $\tilde{e}_n^{(1)}$ . The mean value theorem and

$$|u_{xx}(t, x, v) + u_{xv}(t, x, v)| \leq c \left( 1 + \frac{1}{v} \right), \quad t \geq 0, x \in \mathbb{R}, v > 0$$

and give

$$\begin{aligned} &|u_x(t, \hat{x}_t, v_t) - u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)})| \\ &\leq |\hat{x}_t - \hat{x}_{\eta(t)}| \int_0^1 |u_{xx}(t, \zeta \hat{x}_t + (1 - \zeta) \hat{x}_{\eta(t)}, \zeta v_t + (1 - \zeta) v_{\eta(t)})| d\zeta \\ &\quad + |v_t - v_{\eta(t)}| \int_0^1 |u_{xv}(t, \zeta \hat{x}_t + (1 - \zeta) \hat{x}_{\eta(t)}, \zeta v_t + (1 - \zeta) v_{\eta(t)})| d\zeta \\ &\leq c (|\hat{x}_t - \hat{x}_{\eta(t)}| + |v_t - v_{\eta(t)}|) \left| \left( 1 + \frac{1}{v_t} + \frac{1}{v_{\eta(t)}} \right) \right|, \end{aligned}$$

where we used

$$\frac{1}{\xi v_1 + (1 - \xi)v_2} \leq \frac{1}{v_1} + \frac{1}{v_2}, \quad v_1, v_2 > 0.$$

With the Minkowski inequality and Lemma 1, it holds that

$$\mathbb{E} \left[ \left( \frac{1}{v_t} + \frac{1}{v_{\eta(t)}} \right)^{1+\delta} \right]^{1/(1+\delta)} \leq \mathbb{E} \left[ \left( \frac{1}{v_t} \right)^{1+\delta} \right]^{1/(1+\delta)} + \mathbb{E} \left[ \left( \frac{1}{v_{\eta(t)}} \right)^{1+\delta} \right]^{1/(1+\delta)} \leq c$$

for  $\delta \in (0, \frac{2\kappa\theta}{\sigma^2} - 1)$ , where  $c$  is in particular independent of  $t \in [0, T]$ . Finally, with the Minkowski inequality, the Burkholder-Davis-Gundy inequality, Lemma 1, and Lemma 7, we obtain for all  $p \geq 1$  that

$$\begin{aligned} & \mathbb{E} \left[ \left| \int_{\eta(t)}^t \sqrt{v_s} dW_s \right|^p (|\hat{x}_t - \hat{x}_{\eta(t)}| + |v_t - v_{\eta(t)}|)^p \right]^{1/p} \\ & \leq \mathbb{E} \left[ \left| \int_{\eta(t)}^t \sqrt{v_s} dW_s \right|^{2p} \right]^{1/2p} \mathbb{E} \left[ (|\hat{x}_t - \hat{x}_{\eta(t)}| + |v_t - v_{\eta(t)}|)^{2p} \right]^{1/2p} \\ & = \mathbb{E} \left[ \left| \int_{\eta(t)}^t v_s ds \right|^p \right]^{1/2p} (\mathbb{E} [|\hat{x}_t - \hat{x}_{\eta(t)}|^{2p}]^{1/2p} + \mathbb{E} [ |v_t - v_{\eta(t)}|^{2p} ]^{1/2p}) \\ & \leq c((t - \eta(t))^p)^{1/2p} ((t - \eta(t))^p)^{1/2p} \\ & \leq c\Delta t, \end{aligned}$$

where  $c$  is in particular independent of  $t \in [0, T]$ . The Hölder inequality then gives

$$\tilde{e}_n^{(1)} = O((\Delta t)^2).$$

For  $\tilde{e}_n^{(2)}$ , we will use the integration by parts rule to get rid of the second order derivative. Otherwise, direct estimation would only lead to weak order 1/2. First, recall that, by Lemma 5, we have

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_{xx}(t, \hat{x}_t, v_t) \right] = \frac{1}{t\sqrt{1 - \rho^2}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_t, v_t) \int_0^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \right].$$

Moreover, note that we also have

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \int_0^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \right] = 0 \tag{14}$$

and recall that

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_{xx}(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \right] = 0.$$

Thus, we can write

$$\tilde{e}_n^{(2)} = -\frac{\sigma\sqrt{1 - \rho^2}}{2t} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s (u_x(t, \hat{x}_t, v_t) - u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)})) I_t^B \right] \tag{15}$$

with  $I_t^B = \int_0^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r$ .

Before we analyze this expression further, it remains to show (14). Using the law of total expectation, the adaptedness of  $\hat{x}_{\eta(t)}$ ,  $v_{\eta(t)}$  and of the Itô integrals, we have

$$\begin{aligned} & \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \int_0^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \right] \\ &= \mathbb{E} \left[ \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \int_0^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \middle| \mathcal{F}_{\eta(t)} \right] \right] \\ &= \mathbb{E} \left[ u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_0^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \middle| \mathcal{F}_{\eta(t)} \right] \right] \\ &= \mathbb{E} \left[ u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \left( \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_0^{\eta(t)} \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \middle| \mathcal{F}_{\eta(t)} \right] \right. \right. \\ &\quad \left. \left. + \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_{\eta(t)}^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \middle| \mathcal{F}_{\eta(t)} \right] \right) \right] \\ &= \mathbb{E} \left[ u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \int_0^{\eta(t)} \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \middle| \mathcal{F}_{\eta(t)} \right] \right] \\ &\quad + \mathbb{E} \left[ u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)}) \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_{\eta(t)}^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \middle| \mathcal{F}_{\eta(t)} \right] \right]. \end{aligned}$$

Since

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \middle| \mathcal{F}_{\eta(t)} \right] = 0 = \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_{\eta(t)}^t \frac{1}{\sqrt{v_{\eta(r)}}} dB_r \middle| \mathcal{F}_{\eta(t)} \right],$$

due to the properties of the Itô integral, Equation (14) follows.

Using the mean value theorem in Equation (15), we obtain

$$|u_x(t, \hat{x}_t, v_t) - u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)})| \leq c(|\hat{x}_t - \hat{x}_{\eta(t)}| + |v_t - v_{\eta(t)}|) \left| \left( 1 + \frac{1}{v_t} + \frac{1}{v_{\eta(t)}} \right) \right|.$$

Therefore,

$$|\tilde{e}_n^{(2)}| \leq c \int_{t_n}^{t_{n+1}} \frac{1}{\sqrt{t}} \mathbb{E} \left[ \left| \int_{\eta(t)}^t \sqrt{v_s} dW_s \right| (|\hat{x}_t - \hat{x}_{\eta(t)}| + |v_t - v_{\eta(t)}|) \Theta_t \right] dt$$

with

$$\Theta_t = \frac{|I_t^B|}{\sqrt{t}} \left( 1 + \frac{1}{v_t} + \frac{1}{v_{\eta(t)}} \right).$$

By Lemmas 1 and 2, it holds that

$$\begin{aligned} & \sup_{t \in [0, T]} \mathbb{E} \left[ \left| \frac{I_t^B}{\sqrt{t}} \right|^{q-1} \right] < \infty, \\ & \sup_{t \in [0, T]} \mathbb{E} \left[ \left( 1 + \frac{1}{v_t} + \frac{1}{v_{\eta(t)}} \right)^p \right] < \infty, \end{aligned}$$

for  $q \in [2, \frac{4\kappa\theta}{\sigma^2})$  and  $p \in [0, \frac{2\kappa\theta}{\sigma^2})$ . So, the Hölder inequality leads to

$$\begin{aligned} \mathbb{E}[\Theta_t^{1+\delta}] &= \mathbb{E}\left[\left(\frac{|I_t^B|}{\sqrt{t}}\right)^{1+\delta} \left(1 + \frac{1}{v_t} + \frac{1}{v_{n(t)}}\right)^{1+\delta}\right] \\ &= \mathbb{E}\left[\left(\left(\frac{|I_t^B|}{\sqrt{t}}\right)^{1+\delta}\right)^3\right]^{1/3} \mathbb{E}\left[\left(\left(1 + \frac{1}{v_t} + \frac{1}{v_{n(t)}}\right)^{1+\delta}\right)^{3/2}\right]^{2/3} \\ &= \mathbb{E}\left[\left(\frac{|I_t^B|}{\sqrt{t}}\right)^{3(1+\delta)}\right]^{1/3} \mathbb{E}\left[\left(1 + \frac{1}{v_t} + \frac{1}{v_{n(t)}}\right)^{3(1+\delta)/2}\right]^{2/3} \leq c \end{aligned}$$

with  $\delta \in (0, \frac{4\kappa\theta}{3\sigma^2} - 1)$  and  $c$  in particular independent of  $t \in [0, T]$ .

With the Cauchy-Schwarz, Burkholder-Davis-Gundy, and Minkowski inequalities for  $p \geq 1$ , it follows that

$$\begin{aligned} &\frac{1}{\sqrt{t}} \mathbb{E}\left[\left|\int_{\eta(t)}^t \sqrt{v_s} dW_s\right|^p (|\hat{x}_t - \hat{x}_{\eta(t)}| + |v_t - v_{\eta(t)}|)^p\right]^{1/p} \\ &\leq \frac{1}{\sqrt{t}} \mathbb{E}\left[\left|\int_{\eta(t)}^t \sqrt{v_s} dW_s\right|^{2p}\right]^{1/2p} \mathbb{E}\left[(|\hat{x}_t - \hat{x}_{\eta(t)}| + |v_t - v_{\eta(t)}|)^{2p}\right]^{1/2p} \\ &\leq \frac{c}{\sqrt{t}} \mathbb{E}\left[\left|\int_{\eta(t)}^t v_s ds\right|^p\right]^{1/2p} ((t - \eta(t))^p)^{1/2p} \\ &\leq \frac{c}{\sqrt{t}} ((t - \eta(t))^p)^{1/2p} ((t - \eta(t))^p)^{1/2p} \\ &= \frac{c}{\sqrt{t}} (t - \eta(t)). \end{aligned}$$

With the Hölder inequality, we now have

$$|\tilde{e}_n^{(2)}| \leq c \int_{t_n}^{t_{n+1}} \frac{1}{\sqrt{t}} (t - \eta(t)) dt.$$

Therefore,

$$|e_n| \leq c(\Delta t)^2 + c \int_{t_n}^{t_{n+1}} \frac{1}{\sqrt{t}} (t - \eta(t)) dt,$$

and, since  $[0, T] \ni t \rightarrow \frac{1}{\sqrt{t}} \in (0, \infty)$  is Riemann-integrable, we obtain

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

which concludes the proof of this part.

#### 4.3. The Euler Scheme: Case (ii)

Starting from Equation (11) and using now the bounds of Corollary 1 for  $u_{xx}$ ,  $u_{xv}$ ,  $u_{xxx}$ , and  $u_{xxv}$ , the assertion follows from a direct application of the mean value theorem to (12) and (13), together with the Lemmata 1 and 7.

#### 4.4. Semi-Trapezoidal Rule: Expanding the Error

We look again at the telescoping sum of local errors

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

Recall that

$$\hat{x}_t = x_{\eta(t)} + \int_{\eta(t)}^t a_s ds + \int_{\eta(t)}^t b_s dB_s + \int_{\eta(t)}^t c_s dW_s$$

with

$$\begin{aligned} a_t &:= \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \left( v_{\eta(t)} + \frac{1}{2}(t - \eta(t))\kappa(\theta - v_t) + \frac{1}{2}(v_t - v_{\eta(t)}) \right), \\ b_t &:= \sqrt{1 - \rho^2} \sqrt{v_{\eta(t)}}, \\ c_t &:= \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{1}{2}(t - \eta(t))\sigma\sqrt{v_t}. \end{aligned}$$

With the Itô formula and the Kolmogorov backward PDE evaluated at  $(t, \hat{x}_t, v_t)$ , we have

$$\begin{aligned} e_n &:= \mathbb{E}[u(t_{n+1}, x_{n+1}, v_{n+1}) - u(t_n, x_n, v_n)] \\ &= \int_{t_n}^{t_{n+1}} \mathbb{E}[u_t(t, \hat{x}_t, v_t) + a_t u_x(t, \hat{x}_t, v_t) + \kappa(\theta - v_t) u_v(t, \hat{x}_t, v_t) \\ &\quad + \frac{1}{2}(b_t^2 + c_t^2) u_{xx}(t, \hat{x}_t, v_t) + c_t \sigma \sqrt{v_t} u_{xv}(t, \hat{x}_t, v_t) + \frac{1}{2} v_t \sigma^2 u_{vv}(t, \hat{x}_t, v_t)] dt \\ &= \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \left( v_{\eta(t)} + \frac{1}{2}(t - \eta(t))\kappa(\theta - v_t) + \frac{1}{2}(v_t - v_{\eta(t)}) - v_t \right) u_x(t, \hat{x}_t, v_t) \right. \\ &\quad + \frac{1}{2} \left( (1 - \rho^2) v_{\eta(t)} + \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right)^2 \frac{1}{4} (t - \eta(t))^2 \sigma^2 v_t - (1 - \rho^2) v_t \right) u_{xx}(t, \hat{x}_t, v_t) \\ &\quad \left. + \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{1}{2} (t - \eta(t)) \sigma^2 v_t u_{xv}(t, \hat{x}_t, v_t) \right] dt \\ &= \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \left( \frac{1}{2}(v_{\eta(t)} - v_t) + \frac{1}{2}(t - \eta(t))\kappa(\theta - v_t) \right) u_x(t, \hat{x}_t, v_t) \right. \\ &\quad + \frac{1}{2} \left( (1 - \rho^2)(v_{\eta(t)} - v_t) + \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right)^2 \frac{1}{4} (t - \eta(t))^2 \sigma^2 v_t \right) u_{xx}(t, \hat{x}_t, v_t) \\ &\quad \left. + \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{1}{2} (t - \eta(t)) \sigma^2 v_t u_{xv}(t, \hat{x}_t, v_t) \right] dt. \end{aligned}$$

Using again

$$v_{\eta(t)} - v_t = - \int_{\eta(t)}^t \kappa(\theta - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s,$$

we obtain

$$e_n = e_n^{(1)} + e_n^{(2)} + e_n^{(3)}$$

with

$$\begin{aligned}
 e_n^{(1)} &:= \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \frac{1}{2} \left( \int_{\eta(t)}^t \kappa(v_s - v_t) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) u_x(t, \hat{x}_t, v_t) \right] dt, \\
 e_n^{(2)} &:= \frac{1}{2}(1 - \rho^2) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \left( - \int_{\eta(t)}^t \kappa(\theta - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) u_{xx}(t, \hat{x}_t, v_t) \right] dt \\
 &\quad + \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right)^2 \frac{\sigma^2}{8} \int_{t_n}^{t_{n+1}} (t - \eta(t))^2 \mathbb{E}[v_t u_{xx}(t, \hat{x}_t, v_t)] dt, \\
 e_n^{(3)} &:= \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) \frac{\sigma^2}{2} \int_{t_n}^{t_{n+1}} (t - \eta(t)) \mathbb{E}[v_t u_{xv}(t, \hat{x}_t, v_t)] dt.
 \end{aligned}$$

Since  $vu_{xv}(t, x, v) \leq c(1 + v)$  by Theorem 3, we have

$$e_n^{(3)} = O((\Delta t)^2)$$

using Lemma 1. Moreover, since  $u_x$  and  $u_{xx}$  are bounded by Theorem 3 and Corollary 1 (i), we obtain similar to the calculations for the Euler scheme that

$$\begin{aligned}
 e_n^{(1)} &= O((\Delta t)^2) - \frac{1}{2} \left(\rho\kappa - \frac{\sigma}{2}\right) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_t, v_t) \right] dt \\
 e_n^{(2)} &= O((\Delta t)^2) - \frac{\sigma}{2} (1 - \rho^2) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_{xx}(t, \hat{x}_t, v_t) \right] dt
 \end{aligned}$$

and

$$e_n = O((\Delta t)^2) + \tilde{e}_n^{(1)} + \tilde{e}_n^{(2)}, \tag{16}$$

with

$$\tilde{e}_n^{(1)} = -\frac{1}{2} \left(\rho\kappa - \frac{\sigma}{2}\right) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s (u_x(t, \hat{x}_t, v_t) - u_x(t, \hat{x}_{\eta(t)}, v_{\eta(t)})) \right] dt, \tag{17}$$

$$\tilde{e}_n^{(2)} = -\frac{\sigma}{2} (1 - \rho^2) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s (u_{xx}(t, \hat{x}_t, v_t) - u_{xx}(t, \hat{x}_{\eta(t)}, v_{\eta(t)})) \right] dt. \tag{18}$$

#### 4.5. Semi-Trapezoidal Rule: Case (i)

Since Lemma 8 gives  $\hat{x}_t \in \mathbb{D}^{1,\infty}$  and

$$D_r^B \hat{x}_t = \sqrt{1 - \rho^2} \sqrt{v_{\eta(r)}} \mathbf{1}_{[0,t]}(r),$$

we can proceed here in the same way as for the Euler scheme by using the Lemmata 9 and 11.

#### 4.6. Semi-Trapezoidal Rule: Case (ii)

Starting from (16), the assertion of this case follows from a direct application of the mean value theorem to (17) and (18) using the regularity results from Corollary 1, together with the Lemmata 1 and 11.

### 5. Proof of Theorem 2

Now, we derive the error expansion under the regularity of Theorem 4 with  $q = 4$ , i.e., we have  $u \in C_{\rho\kappa, T}^4(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ .

### 5.1. Euler Scheme: Preliminaries

By the Lemmata 1, 4, and 7, we have that

$$\sup_{t \in [0, T]} \mathbb{E}|v_t|^p + \sup_{t \in [0, T]} \mathbb{E}|\hat{x}_t|^p < \infty$$

and

$$\mathbb{E}|v_t - v_s|^p + \mathbb{E}|\hat{x}_t - \hat{x}_s|^p \leq c \cdot |t - s|^{p/2}, \quad s, t \in [0, T],$$

for all  $p \geq 1$ . Using the Burkholder-Davis-Gundy, Hölder, and Minkowski inequalities, we also have

$$\sup_{t \in [0, T]} \mathbb{E}|X_t|^p < \infty$$

and

$$\mathbb{E}|X_t - X_s|^p \leq c \cdot |t - s|^{p/2}, \quad s, t \in [0, T],$$

for all  $p \geq 1$ . We will use this in the following at several places without explicitly mentioning it.

Recall that we obtained

$$e_n^{(1)} := \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \left( - \int_{\eta(t)}^t \kappa(\theta - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) u_x(t, \hat{x}_t, v_t) \right] dt, \quad (19)$$

$$e_n^{(2)} := \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \frac{1 - \rho^2}{2} \left( - \int_{\eta(t)}^t \kappa(\theta - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) u_{xx}(t, \hat{x}_t, v_t) \right] dt, \quad (20)$$

in Section 4.1.

If higher derivatives of  $u$  are available, then we can analyze

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_t, v_t) \right]$$

and

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_{xx}(t, \hat{x}_t, v_t) \right]$$

via another application of Itô's lemma. So, let  $k : [0, T] \times \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$  be a  $C^{1,2}$ -function that fulfills the backward PDE (7). In particular, the partial derivatives of  $u$  up to order two are such functions. Itô's formula and the Kolmogorov backward PDE (7) now give

$$\begin{aligned} k(t, \hat{x}_t, v_t) &= k(\eta(t), \hat{x}_{\eta(t)}, v_{\eta(t)}) \\ &+ \int_{\eta(t)}^t \left[ \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) k_x(s, \hat{x}_s, v_s) + \frac{(1 - \rho^2)}{2} k_{xx}(s, \hat{x}_s, v_s) \right] (v_{\eta(s)} - v_s) ds \\ &+ \int_{\eta(t)}^t k_x(s, \hat{x}_s, v_s) \sqrt{1 - \rho^2} v_{\eta(s)} dB_s + \int_{\eta(t)}^t k_v(s, \hat{x}_s, v_s) \sigma \sqrt{v_s} dW_s. \end{aligned}$$

If  $k_x$  and  $k_v$  have polynomial growth, then an application of the Itô isometry and the martingale property of the Itô integral yield

$$\begin{aligned} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s k(t, \hat{x}_t, v_t) \right] &= \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s k(\eta(t), \hat{x}_{\eta(t)}, v_{\eta(t)}) \right] \\ &+ \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_{\eta(t)}^t \mathcal{K}(s, \hat{x}_s, v_s) (v_{\eta(s)} - v_s) ds \right] \\ &+ \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_{\eta(t)}^t k_x(s, \hat{x}_s, v_s) \sqrt{1 - \rho^2 v_{\eta(s)}} dB_s \right] \\ &+ \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_{\eta(t)}^t k_v(s, \hat{x}_s, v_s) \sigma \sqrt{v_s} dW_s \right] \\ &= \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_{\eta(t)}^t \int_{\eta(s)}^s \mathcal{K}(s, \hat{x}_s, v_s) \kappa(v_u - \theta) dud s \right] \\ &- \sigma \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_{\eta(t)}^t \mathcal{K}(s, \hat{x}_s, v_s) \int_{\eta(s)}^s \sqrt{v_u} dW_u ds \right] \\ &+ \sigma \mathbb{E} \left[ \int_{\eta(t)}^t v_s k_v(s, \hat{x}_s, v_s) ds \right], \end{aligned}$$

where

$$\mathcal{K}(s, \hat{x}_s, v_s) = \left( \frac{\rho \kappa}{\sigma} - \frac{1}{2} \right) k_x(s, \hat{x}_s, v_s) + \frac{(1 - \rho^2)}{2} k_{xx}(s, \hat{x}_s, v_s).$$

If  $k_x$  and  $k_{xx}$  have polynomial growth, then an application of Hölder’s inequality and the Itô isometry yield

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s \int_{\eta(t)}^t \int_{\eta(s)}^s \mathcal{K}(s, \hat{x}_s, v_s) \kappa(v_u - \theta) dud s \right] = O((\Delta t)^{5/2}),$$

and so it follows

$$\begin{aligned} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s k(t, \hat{x}_t, v_t) \right] &= \sigma \mathbb{E} \left[ \int_{\eta(t)}^t v_s k_v(s, \hat{x}_s, v_s) ds \right] \\ &- \sigma \mathbb{E} \left[ \int_{\eta(t)}^t \mathcal{K}(s, \hat{x}_s, v_s) \int_{\eta(t)}^t \sqrt{v_u} dW_u \int_{\eta(s)}^s \sqrt{v_u} dW_u ds \right] \\ &+ O((\Delta t)^{5/2}). \end{aligned}$$

Since we have

$$\begin{aligned} &\mathbb{E} \left[ \int_{\eta(t)}^t \mathcal{K}(s, \hat{x}_s, v_s) \int_{\eta(t)}^t \sqrt{v_u} dW_u \int_{\eta(s)}^s \sqrt{v_u} dW_u ds \right] \\ &= \mathbb{E} \left[ \int_{\eta(t)}^t \mathcal{K}(s, \hat{x}_s, v_s) \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_u} dW_u \int_{\eta(s)}^s \sqrt{v_u} dW_u \middle| \mathcal{F}_s \right] ds \right] \\ &= \mathbb{E} \left[ \int_{\eta(t)}^t \mathcal{K}(s, \hat{x}_s, v_s) \left( \int_{\eta(s)}^s \sqrt{v_u} dW_u \right)^2 ds \right], \end{aligned}$$

again, by the properties of the Itô integral, we finally obtain by Hölder’s inequality and the Burkholder-Davis-Gundy inequality that

$$\mathbb{E} \left[ \int_{\eta(t)}^t \mathcal{K}(s, \hat{x}_s, v_s) \int_{\eta(t)}^t \sqrt{v_u} dW_u \int_{\eta(s)}^s \sqrt{v_u} dW_u ds \right] = O((\Delta t)^2).$$

Thus, we can conclude that

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s k(t, \hat{x}_t, v_t) \right] = \sigma \mathbb{E} \left[ \int_{\eta(t)}^t v_s k_v(s, \hat{x}_s, v_s) ds \right] + O((\Delta t)^2), \tag{21}$$

for  $k = u_x$  and  $k = u_{xx}$ , if the derivatives up to order four of  $u$  have polynomial growth.



5.2. Euler Scheme: Conclusion

Setting now  $k = u_x$  in Equation (21), we have from (19) that

$$e_n^{(1)} := -\left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) \int_{t_n}^{t_{n+1}} \int_{\eta(t)}^t \mathbb{E}[u_x(t, \hat{x}_t, v_t)\kappa(\theta - v_s)] ds dt - \sigma^2 \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) \int_{t_n}^{t_{n+1}} \int_{\eta(t)}^t \mathbb{E}[v_s u_{xv}(s, \hat{x}_s, v_s)] ds dt + O((\Delta t)^3).$$

Replacing now the function  $k$  by  $u_{xx}$  in Equation (21), we arrive from (20) at

$$e_n^{(2)} := -\frac{1}{2}(1 - \rho^2) \int_{t_n}^{t_{n+1}} \int_{\eta(t)}^t \mathbb{E}[u_{xx}(t, \hat{x}_t, v_t)\kappa(\theta - v_s)] ds dt - \frac{\sigma^2}{2}(1 - \rho^2) \int_{t_n}^{t_{n+1}} \int_{\eta(t)}^t \mathbb{E}[v_s u_{xxv}(s, \hat{x}_s, v_s)] ds dt + O((\Delta t)^3).$$

Summarizing, we have shown that

$$e_n = \left(\frac{1}{2} - \frac{\rho\kappa}{\sigma}\right) \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\int_{\eta(t)}^t \kappa(\theta - v_s) u_x(t, \hat{x}_t, v_t) + \sigma^2 v_s u_{xv}(s, \hat{x}_s, v_s) ds\right] dt - \frac{(1 - \rho^2)}{2} \int_{t_n}^{t_{n+1}} \mathbb{E}\left[\int_{\eta(t)}^t \kappa(\theta - v_s) u_{xx}(t, \hat{x}_t, v_t) + \sigma^2 v_s u_{xxv}(s, \hat{x}_s, v_s) ds\right] dt + O((\Delta t)^3)$$

and

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

where

$$\mathcal{H}(s, t, \hat{x}_s, \hat{x}_t, v_s, v_t) = \left(\frac{1}{2} - \frac{\rho\kappa}{\sigma}\right) \left(\kappa(\theta - v_s) u_x(t, \hat{x}_t, v_t) + \sigma^2 v_s u_{xv}(s, \hat{x}_s, v_s)\right) - \frac{(1 - \rho^2)}{2} \left(\kappa(\theta - v_s) u_{xx}(t, \hat{x}_t, v_t) + \sigma^2 v_s u_{xxv}(s, \hat{x}_s, v_s)\right).$$

An application of the mean value theorem, the polynomial growth of the derivatives of  $u$ , the Minkowski inequality, the Hölder inequality, and the Lemmata 1, 6 yields for  $s, t \in [t_n, t_{n+1}]$  that

$$\mathbb{E}[\mathcal{H}(s, t, \hat{x}_s, \hat{x}_t, v_s, v_t)] = \mathbb{E}[\mathcal{H}(t_n, t_n, X_{t_n}, X_{t_n}, V_{t_n}, V_{t_n})] + O((\Delta t)^{1/4}).$$

Note here that  $u \in C_{pol,T}^4(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$  implies that  $u_{tx}$  and  $u_{txx}$  are well-defined, have polynomial growth, and are continuous.

Thus, for an equidistant discretization  $t_k = kT/N, k = 0, \dots, N$ , we have

$$\mathbb{E}[f(x_N, v_N)] - \mathbb{E}[f(X_T, V_T)] = \frac{\Delta t}{2} \sum_{n=0}^{N-1} \mathbb{E}[\mathcal{H}(t_n, t_n, X_{t_n}, X_{t_n}, V_{t_n}, V_{t_n})] \Delta t + O((\Delta t)^{5/4}).$$

Since

$$\sum_{n=0}^{N-1} \mathbb{E}[\mathcal{H}(t_n, t_n, X_{t_n}, X_{t_n}, V_{t_n}, V_{t_n})] \Delta t \rightarrow \int_0^T \mathbb{E}[\mathcal{H}(t, t, X_t, X_t, V_t, V_t)] dt$$

for  $\Delta t \rightarrow 0$ , this concludes the proof of Theorem 2 (i).

### 5.3. Semi-Trapezoidal Scheme: Preliminaries

By the Lemmata 1, 8, and 11, we have that

$$\sup_{t \in [0, T]} \mathbb{E}|v_t|^p + \sup_{t \in [0, T]} \mathbb{E}|\hat{x}_t|^p < \infty$$

and

$$\mathbb{E}|v_t - v_s|^p + \mathbb{E}|\hat{x}_t - \hat{x}_s|^p \leq c \cdot |t - s|^{p/2}, \quad s, t \in [0, T],$$

for all  $p \geq 1$ . Using the Burkholder-Davis-Gundy, Hölder, and Minkowski inequalities, we also have

$$\sup_{t \in [0, T]} \mathbb{E}|X_t|^p < \infty$$

and

$$\mathbb{E}|X_t - X_s|^p \leq c \cdot |t - s|^{p/2}, \quad s, t \in [0, T],$$

for all  $p \geq 1$ .

We will use this in the following at several places without explicitly mentioning it.

We will now take also a closer look at the error of the semi-trapezoidal discretization for  $u \in C^4_{pol, T}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ . Recall that

$$\begin{aligned} e_n^{(1)} &:= \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \frac{1}{2} \left( - \int_{\eta(t)}^t \kappa(v_t - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) u_x(t, \hat{x}_t, v_t) \right], \\ e_n^{(2)} &:= \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \frac{1}{2} (1 - \rho^2) \left( - \int_{\eta(t)}^t \kappa(\theta - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) u_{xx}(t, \hat{x}_t, v_t) \right. \\ &\quad \left. + \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right)^2 \frac{\sigma^2}{8} (t - \eta(t))^2 v_s u_{xx}(t, \hat{x}_t, v_t) \right] dt, \\ e_n^{(3)} &:= \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{\sigma^2}{2} \int_{t_n}^{t_{n+1}} \mathbb{E} [(t - \eta(t)) v_t u_{xv}(t, \hat{x}_t, v_t)] dt. \end{aligned}$$

We can again use the Itô formula and the Kolmogorov backward PDE (7) evaluated at  $(s, \hat{x}_s, v_s)$  and obtain for a  $C^{1,2}$ -function  $k$ , which fulfills the PDE (7), that

$$\begin{aligned} &k(t, \hat{x}_t, v_t) - k(t_n, \hat{x}_{t_n}, v_{t_n}) \\ &= \int_{t_n}^t k_t(s, \hat{x}_s, v_s) ds + \int_{t_n}^t k_v(s, \hat{x}_s, v_s) dv_s + \int_{t_n}^t k_x(s, \hat{x}_s, v_s) d\hat{x}_s \\ &\quad + \frac{1}{2} \int_{t_n}^t k_{xx}(s, \hat{x}_s, v_s) d\langle \hat{x} \rangle_s + \int_{t_n}^t k_{xv}(s, \hat{x}_s, v_s) d\langle \hat{x}, v \rangle_s + \frac{1}{2} \int_{t_n}^t k_{vv}(s, \hat{x}_s, v_s) d\langle v \rangle_s \\ &= \int_{t_n}^t \left( a_s - \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) v_s \right) k_x(s, \hat{x}_s, v_s) ds + \frac{1}{2} \int_{t_n}^t \left( b_s^2 + c_s^2 - (1 - \rho^2) v_s \right) k_{xx}(s, \hat{x}_s, v_s) ds \\ &\quad + \int_{t_n}^t c_s \sigma \sqrt{v_s} k_{xv}(s, \hat{x}_s, v_s) ds + \int_{t_n}^t b_s k_x(s, \hat{x}_s, v_s) dB_s + \int_{t_n}^t c_s k_x(s, \hat{x}_s, v_s) dW_s \\ &\quad + \int_{t_n}^t \sigma \sqrt{v_s} k_v(s, \hat{x}_s, v_s) dW_s \end{aligned} \tag{22}$$

with

$$\begin{aligned} a_t &:= \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \left( v_{\eta(t)} + \frac{1}{2} (t - \eta(t)) \kappa(\theta - v_t) + \frac{1}{2} (v_t - v_{\eta(t)}) \right), \\ b_t &:= \sqrt{1 - \rho^2} \sqrt{v_{\eta(t)}}, \\ c_t &:= \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{1}{2} (t - \eta(t)) \sigma \sqrt{v_t}. \end{aligned}$$

Analogous calculations as for the Euler scheme yield that

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_\tau} dW_\tau k(t, \hat{x}_t, v_t) \right] = \sigma \mathbb{E} \left[ \int_{\eta(t)}^t v_\tau k_v(\tau, \hat{x}_\tau, v_\tau) d\tau \right] + O((\Delta t)^2) \tag{23}$$

for  $k = u_x$  and  $k = u_{xx}$  under the assumption  $u \in C_{pol,T}^4(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ .

5.4. Semi-Trapezoidal Rule: Calculations for  $e_n^{(1)}$ ,  $e_n^{(2)}$ , and  $e_n^{(3)}$

Rewriting the terms of  $e_n^{(1)}$  using (23) for the last term gives

$$\begin{aligned} e_n^{(1)} &= - \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{1}{2} \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \kappa(v_t - v_s) ds u_x(t, \hat{x}_t, v_t) \right] dt \\ &\quad - \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{\sigma}{2} \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_x(t, \hat{x}_t, v_t) \right] dt \\ &= - \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{1}{2} \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \kappa \left( \int_s^t \kappa(\theta - v_u) du + \sigma \int_s^t \sqrt{v_u} dW_u \right) ds u_x(t, \hat{x}_t, v_t) \right] dt \\ &\quad - \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{\sigma^2}{2} \int_{t_n}^{t_{n+1}} \int_{\eta(t)}^t \mathbb{E}[v_s u_{xv}(s, \hat{x}_s, v_s)] ds dt + O((\Delta t)^3). \end{aligned}$$

Applying, again, (23) with  $s$  instead of  $\eta(t)$  as the lower bound of the integral to the second summand of the first term, using the polynomial growth of the derivatives of  $u$  and Hölder’s inequality, we also have

$$\int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \kappa \left( \int_s^t \kappa(\theta - v_u) du + \sigma \int_s^t \sqrt{v_u} dW_u \right) ds u_x(t, \hat{x}_t, v_t) \right] dt = O((\Delta t)^3)$$

and so

$$e_n^{(1)} = - \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{\sigma^2}{2} \int_{t_n}^{t_{n+1}} \int_{\eta(t)}^t \mathbb{E}[v_s u_{xv}(s, \hat{x}_s, v_s)] ds dt + O((\Delta t)^3).$$

Adding  $e_n^{(3)}$  yields

$$e_n^{(1)} + e_n^{(3)} = O((\Delta t)^3) - \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right) \frac{\sigma^2}{2} \int_{t_n}^{t_{n+1}} \int_{t_n}^t \mathbb{E}[u_{xv}(t, \hat{x}_t, v_t) v_t - u_{xv}(s, \hat{x}_s, v_s) v_s] ds dt.$$

However, Itô’s formula gives for sufficiently smooth  $k : [0, T] \times \mathbb{R} \times [0, \infty) \rightarrow \mathbb{R}$  that

$$\begin{aligned} &k(t, \hat{x}_t, v_t) - k(s, \hat{x}_s, v_s) \\ &= \int_s^t k_t(r, \hat{x}_r, v_r) dr + \int_s^t k_v(r, \hat{x}_r, v_r) dv_r + \int_s^t k_x(r, \hat{x}_r, v_r) d\hat{x}_r \\ &\quad + \frac{1}{2} \int_s^t k_{xx}(r, \hat{x}_r, v_r) d\langle \hat{x} \rangle_r + \int_s^t k_{xv}(r, \hat{x}_r, v_r) d\langle \hat{x}, v \rangle_r + \frac{1}{2} \int_s^t k_{vv}(r, \hat{x}_r, v_r) d\langle v \rangle_r \end{aligned} \tag{24}$$

with

$$dv_t = \kappa(\theta - v_t) dt + \sigma \sqrt{v_t} dW_t, \quad d\hat{x}_t = a_t dt + b_t dB_t + c_t dW_t,$$

where

$$\begin{aligned}
 a_t &:= \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) \left(v_{\eta(t)} + \frac{1}{2}(t - \eta(t))\kappa(\theta - v_t) + \frac{1}{2}(v_t - v_{\eta(t)})\right), \\
 b_t &:= \sqrt{1 - \rho^2} \sqrt{v_{\eta(t)}}, \\
 c_t &:= \left(\frac{\rho\kappa}{\sigma} - \frac{1}{2}\right) \frac{1}{2}(t - \eta(t))\sigma \sqrt{v_t}
 \end{aligned}$$

and

$$d\langle \hat{x} \rangle_t = (b_t^2 + c_t^2)dt, \quad d\langle \hat{x}, v \rangle_t = \sigma c_t \sqrt{v_t} dt, \quad d\langle v \rangle_t = \sigma^2 v_t dt.$$

Since  $u \in C^4_{pol,T}(\mathbb{R} \times \mathbb{R}_+; \mathbb{R})$ , we can apply this to  $k(t, x, v) = u_{xv}(t, x, v)v$  and taking expectations gives then

$$\mathbb{E}[u_{xv}(t, \hat{x}_t, v_t)v_t - u_{xv}(s, \hat{x}_s, v_s)v_s] = O(|t - s|).$$

So, we end up with

$$e_n^{(1)} + e_n^{(3)} = O((\Delta t)^3).$$

Looking at  $e_n^{(2)}$ , the last term is already of third order:

$$\begin{aligned}
 e_n^{(2)} &= \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \frac{1}{2}(1 - \rho^2) \left( - \int_{\eta(t)}^t \kappa(\theta - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) u_{xx}(t, \hat{x}_t, v_t) \right. \\
 &\quad \left. + \left( \frac{\rho\kappa}{\sigma} - \frac{1}{2} \right)^2 \frac{\sigma^2}{8} (t - \eta(t))^2 v_s u_{xx}(t, \hat{x}_t, v_t) \right] dt \\
 &= \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \frac{1}{2}(1 - \rho^2) \left( - \int_{\eta(t)}^t \kappa(\theta - v_s) ds - \sigma \int_{\eta(t)}^t \sqrt{v_s} dW_s \right) u_{xx}(t, \hat{x}_t, v_t) \right] dt \\
 &\quad + O((\Delta t)^3).
 \end{aligned}$$

Since

$$\mathbb{E} \left[ \int_{\eta(t)}^t \sqrt{v_s} dW_s u_{xx}(t, \hat{x}_t, v_t) \right] = \sigma \mathbb{E} \left[ \int_{\eta(t)}^t v_s u_{xxv}(s, \hat{x}_s, v_s) ds \right] + O((\Delta t)^2),$$

by (23), it follows

$$\begin{aligned}
 e_n^{(2)} &= -\frac{1}{2}(1 - \rho^2) \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \kappa(\theta - v_s) ds u_{xx}(t, \hat{x}_t, v_t) \right] dt \\
 &\quad - \frac{1}{2}(1 - \rho^2) \sigma^2 \int_{t_n}^{t_{n+1}} \int_{\eta(t)}^t \mathbb{E}[v_s u_{xxv}(s, \hat{x}_s, v_s)] ds dt + O((\Delta t)^3).
 \end{aligned}$$

### 5.5. Semi-Trapezoidal Scheme: Conclusion

Summarizing, we have shown that

$$e_n = -\frac{(1 - \rho^2)}{2} \int_{t_n}^{t_{n+1}} \mathbb{E} \left[ \int_{\eta(t)}^t \kappa(\theta - v_s) u_{xx}(t, \hat{x}_t, v_t) + \sigma^2 v_s u_{xxv}(s, \hat{x}_s, v_s) ds \right] dt + O((\Delta t)^3)$$

and

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

where

$$\mathcal{H}(s, t, \hat{x}_s, \hat{x}_t, v_s, v_t) = -\frac{(1-\rho^2)}{2} \left( \kappa(\theta - v_s) u_{xx}(t, \hat{x}_t, v_t) + \sigma^2 v_s u_{xxv}(s, \hat{x}_s, v_s) \right).$$

An application of the mean value theorem, the polynomial growth of the derivatives of  $u$ , the Minkowski inequality, the Hölder inequality, and the Lemmata 1, 10 yields for  $s, t \in [t_n, t_{n+1}]$  that

$$\mathbb{E}[\mathcal{H}(s, t, \hat{x}_s, \hat{x}_t, v_s, v_t)] = \mathbb{E}[\mathcal{H}(t_n, t_n, X_{t_n}, X_{t_n}, V_{t_n}, V_{t_n})] + O((\Delta t)^{1/4}).$$

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

$$\mathbb{E}[f(x_N, v_N)] - \mathbb{E}[f(X_T, V_T)] = \frac{\Delta t}{2} \sum_{n=0}^{N-1} \mathbb{E} \mathcal{H}(t_n, t_n, X_{t_n}, X_{t_n}, V_{t_n}, V_{t_n}) \Delta t + O((\Delta t)^{5/4})$$

and the convergence

$$\sum_{n=0}^{N-1} \mathbb{E} \mathcal{H}(t_n, t_n, X_{t_n}, X_{t_n}, V_{t_n}, V_{t_n}) \Delta t \rightarrow \int_0^T \mathbb{E} \mathcal{H}(t, t, X_t, X_t, V_t, V_t) dt$$

for  $\Delta t \rightarrow 0$  concludes the proof of Theorem 2 (ii).

**Author Contributions:** All authors have contributed substantially to this manuscript. All authors have read and agreed to the published version of the manuscript.

**Funding:** DFG Research Training Group 1953 “Statistical Modeling of Complex Systems”.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Data sharing not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Alos, Elisa, and Christian-Oliver Ewald. 2008. Malliavin differentiability of the Heston volatility and applications to option pricing. *Advances in Applied Probability* 40: 144–62. [CrossRef]
- Altmayer, Martin. 2015. Quadrature of Discontinuous SDE Functionals Using Malliavin Integration by Parts. Ph.D. dissertation, University of Mannheim, Germany.
- Altmayer, Martin, and Andreas Neuenkirch. 2015. Multilevel Monte Carlo quadrature of discontinuous payoffs in the generalized Heston model using Malliavin integration by parts. *SIAM Journal on Financial Mathematics* 6: 22–52.
- Altmayer, Martin, and Andreas Neuenkirch. 2017. Discretising the Heston model: an analysis of the weak convergence rate. *IMA Journal of Numerical Analysis* 37: 1930–60. [CrossRef]
- Andersen, Leif. 2008. Simple and efficient simulation of the Heston stochastic volatility model. *Journal of Computational Finance* 11: 29–50.
- Bally, Vlad, and Denis Talay. 1996. The law of the Euler scheme for stochastic differential equations. I: Convergence rate of the distribution function. *Probability Theory and Related Fields* 104: 43–60. [CrossRef]
- Bossy, Mireille, and Awa Diop. 2015. Weak convergence analysis of the symmetrized Euler scheme for one dimensional SDEs with diffusion coefficient  $|x|^a, a \in [1/2, 1)$ . *arXiv arXiv:1508.04573*.
- Briani, Maya, Lucia Caramellino, and Giulia Terenzi. 2018. Convergence rate of Markov chains and hybrid numerical schemes to jump-diffusions with application to the Bates model. *arXiv arXiv:1809.10545*.
- Broadie, Mark, and Özgür Kaya. 2006. Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research* 54: 217–31. [CrossRef]
- Coskun, Sema, and Ralf Korn. 2018. Pricing barrier options in the Heston model using the Heath–Platen estimator. *Monte Carlo Methods and Applications* 24: 29–41.
- Cui, Zhenyu, Justin Kirkby, and Duy Nguyen. 2018. A general valuation framework for SABR and stochastic local volatility models. *SIAM Journal on Financial Mathematics* 9: 520–63.
- Cui, Zhenyu, Justin Kirkby, and Duy Nguyen. 2020. Efficient simulation of generalized SABR and stochastic local volatility models based on markov chain approximations. *European Journal of Operational Research*. [CrossRef]

- Feehan, Paul, and Camelia Pop. 2013. A Schauder approach to degenerate-parabolic partial differential equations with unbounded coefficients. *Journal of Differential Equations* 254: 4401–45. [[CrossRef](#)]
- Glasserman, Paul, and Kyoung-Kuk Kim. 2011. Gamma expansion of the Heston stochastic volatility model. *Finance and Stochastics* 15: 267–96. [[CrossRef](#)]
- Heston, Steven. 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies* 6: 327–43. [[CrossRef](#)]
- Hurd, Thomas, and Alexey Kuznetsov. 2008. Explicit formulas for Laplace transforms of stochastic integrals. *Markov Processes and Related Fields* 14: 277–90.
- Kahl, Christian, and Peter Jäckel. 2006. Fast strong approximation Monte-Carlo schemes for stochastic volatility models. *Quantitative Finance* 6: 513–36.
- Karatzas, Ioannis, and Steven Shreve. 1991. *Brownian Motion and Stochastic Calculus*, 2nd ed. New York: Springer.
- Malham, Simon, and Anke Wiese. 2013. Chi-square simulation of the CIR process and the Heston model. *International Journal of Theoretical and Applied Finance* 16: 1350014.
- Nualart, David. 1995. *The Malliavin Calculus and Related Topics*. New York: Springer.
- Lord, Roger, Remmert Koekoek, and Dick van Dijk. 2009. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance* 10: 177–94.
- Smith, Robert. 2007. An almost exact simulation method for the Heston model. *Journal of Computational Finance* 11: 115–25.
- Talay, Denis, and Luciano Tubaro. 1990. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Analysis and Applications* 8: 483–509.
- Zheng, Chao. 2017. Weak convergence rate of a time-discrete scheme for the Heston stochastic volatility. *SIAM Journal on Numerical Analysis* 55: 1243–63. [[CrossRef](#)]