Pricing Derivatives Using the Original PDE and Simulations of the Heston Model

Seminar: Asset Prices and Financial Markets

Youssef Raad (zfw568)

20-12-2024

Abstract

This paper explores option pricing methods using the Heston model, renowned for its ability to reflect the stochastic nature of volatility and capture key market phenomena like the leverage effect and investor risk aversion. As such, we contribute to the literature by extensively deriving risk-neutral dynamics, market price of volatility risk and the pricing PDE. The model enables the pricing of volatility-sensitive derivatives and the development of volatility trading strategies, making it particularly relevant for long-dated FX options and equity markets. We compare the widely used log-Euler (Euler) method equipped with a full truncation scheme and approximated exact solution of the stock price process with the Quadratic-Exponential (QE) method and its martingale-corrected variant (QE-M) for simulating the Heston model. Using Monte Carlo simulations of European call options, we assess the bias of these methods against exact prices obtained via Fourier transform techniques. Despite the computational speed of Euler (50%–25% faster than QE-M), QE-M consistently demonstrates superior accuracy, with minimal bias even at coarse time steps. Both methods improve with increased refinement, but QE-M exhibits robust performance under less stringent conditions. These findings emphasize the advantages of advanced simulation methods like QE-M for precise and efficient option pricing.

Cooper, this is no time for caution!

Contents

Prelimiaries

- If a reference is not given proves will be supplied in Appendix B.
- We only consider European call option prices. Put options can be found by the Put-Call Parity [4, Prop. 10.2].
- We assume no market frictions.
- For readability, let χ_d^2, $\chi_d'^2(\lambda)$ denote the chi-square distribution with d degree(s) of freedom (DOF) and the non-central chi-square distribution with d DOF and non-centrality parameter (NCP) λ, respectively.

1 Financial Models

1.1 Introduction to Stochastic Volatility Models

In practice, financial markets exhibit features such as volatility clustering as already described in [21].

Stochastic volatility models assume that the variance of the underlying is itself a random process. The general form of a stochastic volatility model can be expressed by stochastic differential equations (SDE's)

$$\begin{aligned} dS_t &= a(S_t, v_t, t)dt + b(S_t, v_t, t)dW^{\mathbb{P}}_{1,t}, \quad S_0 > 0; \\ dv_t &= g(v_t, t)dt + h(v_t, t)dW^{\mathbb{P}}_{2,t}, \quad v_0 > 0, \end{aligned}$$

where $dW^{\mathbb{P}}_{1,t}$,$dW^{\mathbb{P}}_{2,t}$ are two correlated Wiener processes, $d\langle W^{\mathbb{P}}_{1,t}, W^{\mathbb{P}}_{2,t}\rangle(t) = \rho dt$ and $\rho \in (-1, 1)$ under the real world measure $\mathbb{P}$. The functions a, b, g and h are such that the model is well defined. v_t represents the instantaneous stochastic variance of the asset. By allowing stochasticity in variance, SVM's capture the empirical characteristics of asset returns [21].

Due to the additional risk in stochastic variance ($dW^{\mathbb{P}}_{1,t} \neq dW^{\mathbb{P}}_{2,t}$), the financial markets with a risk-free asset and a SVM are incomplete as changing the law of v doesn't change the discounted value of S implying the existence of an infinite amount of EMM's. However, it is possible to characterize each EMM by Girsanov's theorem [4, Thm. 12.3]

1.2 The Heston Model

Model dynamics The Heston model [15], by Steven Heston, 1993, is a SVM used in financial markets for pricing derivatives under a bivariate system. The Heston model describes the evolution of the underlying asset price S_t and its variance v_t under the measure $\mathbb{P}$ by the following system of SDE's:

$$\begin{aligned} dS_t &= \mu S_t dt + \sqrt{v_t} S_t dW^{\mathbb{P}}_{1,t}, \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dW^{\mathbb{P}}_{2,t}, \end{aligned} \tag{1.1}$$

where:

- S_t is the price of the asset at time t.
- v_t is the instantaneous variance of the asset price at time t.
- μ is the drift rate of the asset.
- θ is the long-term mean variance.
- κ is the rate at which the variance reverts to its long-term mean θ.
- σ is the volatility of the variance process.
- $W^{\mathbb{P}}_{1,t}$ and $W^{\mathbb{P}}_{2,t}$ are two Wiener processes with correlation ρ.

The process for the variance arises from the Ornstein-Uhlenbeck process for the variance $b_t = \sqrt{v_t}$ given by

$$db_t = -\beta b_t dt + \delta dW_{2,t}^{\mathbb{P}}$$

Applying Itô [4, Prop. 4.12] and the multiplication table from said Proposition yields

$$\begin{aligned}
dv_t &= 2b_t db_t + \frac{1}{2} \cdot 2 \cdot (db_t)^2 \\
&= 2b_t \left(-\beta b_t dt + \delta dW_{2,t}^{\mathbb{P}}\right) + \delta^2 dt \\
&= -2\beta b_t^2 dt + 2\delta b_t dW_{2,t}^{\mathbb{P}} + \delta^2 dt \\
&= -2\beta v_t dt + 2\delta b_t dW_{2,t}^{\mathbb{P}} + \delta^2 dt \\
&\iff \\
dv_t &= (\delta^2 - 2\beta v_t)dt + 2\delta\sqrt{v_t} dW_{2,t}^{\mathbb{P}},
\end{aligned} \tag{1.2}$$

Defining $\kappa = 2\beta$, $\theta = \delta^2/(2\beta)$ and $\sigma = 2\delta$ gives the instantaneous variance given in (1.1).

Applying Girsanov's theorem [4, Thm. 12.3] to each process of (1.1) yields

$$dS_t = rS_t dt + \sqrt{v_t} S_t dW_{1,t}^{\mathbb{Q}}, \quad W_{1,t}^{\mathbb{Q}} = \left(W_{1,t}^{\mathbb{P}} + \frac{\mu - r}{\sqrt{v_t}} t\right).$$

The risk-neutral process for the variance is obtained by introducing a function $\Lambda(S_t, v_t, t)$ into the drift of the instantaneous variance in equation (1.1), as follows

$$dv_t = (\kappa(\theta - v_t) - \Lambda(S_t, v_t, t))dt + \sigma\sqrt{v_t} dW_{2,t}^{\mathbb{Q}}, \quad W_{2,t}^{\mathbb{Q}} = \left(W_{2,t}^{\mathbb{P}} + \frac{\Lambda(S_t, v_t, t)}{\sigma\sqrt{v_t}} t\right).$$

To bear the additional risk from dv_t, a risk-averse investor would expect to obtain a *risk premium*. This means that the volatility risk premium is deducted from the drift of dv_t. As stated in [15, p. 329], the model of [5]'s consumption yields a premium proportional to the variance: $\Lambda(S_t, v_t, t) = \Lambda v_t$. Substituting for Λv_t in (1.2) the model under $\tilde{\mathbb{Q}}(\Lambda)$ can now be specified as

$$\begin{aligned}
dS(t) &= rS_t dt + \sqrt{v(t)} S_t dW_{1,t}^{\tilde{\mathbb{Q}}(\Lambda)}, \\
dv_t &= \tilde{\kappa}(\tilde{\theta} - v_t)dt + \sigma\sqrt{v_t} dW_{2,t}^{\tilde{\mathbb{Q}}(\Lambda)}, \\
\mathbb{E}_t^{\tilde{\mathbb{Q}}(\Lambda)}\left[dW_{1,t}^{\tilde{\mathbb{Q}}(\Lambda)}, dW_{2,t}^{\tilde{\mathbb{Q}}(\Lambda)}\right] &= \rho dt,
\end{aligned}$$

where

$$\tilde{\kappa} = \kappa + \Lambda \quad \text{and} \quad \tilde{\theta} = \frac{\kappa\theta}{\kappa + \Lambda}.$$

Following [13, p. 7], assume that the process where the model is calibrated to option prices precisely provides the risk-neutral process, so the price Λ for bearing additional risk due to volatility risk is set to 0. Since we assumed $\Lambda = 0$, the pricing measure can be derived from

European option prices, the choice of statistical measure is irrelevant to us which simplifies exactly to our model in (1.1).

To ease notation, we use shorthand $\tilde{\mathbb{Q}}(\Lambda) := \tilde{\mathbb{Q}}(0) := \mathbb{Q}$ unless another numeraire is specified.

The CIR process is special because of the Feller condition [12].

Theorem 1.1. *The square-root process is non-negative if Feller's condition*

$$2\kappa\theta \geq \sigma^2$$

is satisfied. Otherwise, the origin is accessible and reflective.

The condition is almost never satisfied [23, p. 4]. Furthermore, under discretization of the SDE the condition can be satisfied and still produce negative variances as nothing is ensuring that the discretization has the same property.

The European call price The time-t price of a European call on a non-dividend paying stock with spot price S_t, strike K, and time to maturity $\tau = T - t$ is

$$\begin{aligned}
Call(t) &= e^{-r\tau}\mathbb{E}_t^{\mathbb{Q}}\left[(S_T - K)^+\right] \\
&= e^{-r\tau}\mathbb{E}_t^{\mathbb{Q}}\left[(S_T - K)\mathbf{1}_{S_T \geq K}\right] \\
&= e^{-r\tau}\left(\mathbb{E}_t^{\tilde{\mathbb{Q}}}\left[S_T\mathbf{1}_{S_T \geq K}\right] - K\mathbb{E}_t^{\tilde{\mathbb{Q}}}\left[\mathbf{1}_{S_T \geq K}\right]\right) \\
&= S_t\mathbb{E}_t^{\mathbb{Q}^S}\left[\mathbf{1}_{S_T \geq K}\right] - Ke^{-r\tau}\mathbb{E}_t^{\tilde{\mathbb{Q}}}\left[\mathbf{1}_{S_T \geq K}\right] \\
&= S_t\mathbb{Q}^S(S_T \geq K) - Ke^{-r\tau}\tilde{\mathbb{Q}}(S_T \geq K) \\
&= S_tQ_1 - Ke^{-r\tau}Q_2,
\end{aligned}$$

where $\frac{d\tilde{\mathbb{Q}}}{d\mathbb{Q}^S} = \frac{B_T/B_t}{S_T/S_t}$ is the Radon-Nikodym derivative, since $\tilde{\mathbb{Q}}$ uses the bond B_t as the numeraire, while $\mathbb{Q}^S$ uses the stock price S.

The Heston PDE Assume we were to describe the evolution of a call option $U(S, v, t)$ following (1.1). By multi-dimensional Itô [4, Thm. 4.19] the process is

$$dU(S,v,t) = \frac{\partial U}{\partial t}dt + \frac{\partial U}{\partial S}dS + \frac{\partial U}{\partial v}dv + \frac{1}{2}\left(\frac{\partial^2 U}{\partial S^2}(dS)^2 + 2\frac{\partial^2 U}{\partial S \partial v}dSdv + \frac{\partial^2 U}{\partial v^2}(dv)^2\right).$$

By the multiplication table provided in [4, Prop. 4.12] we find

$$\begin{aligned}
dS_t &= rS_t\,dt + \sqrt{v_t}S_t\,dW_{1,t}^{\mathbb{P}} \\
&\iff \\
(dS_t)^2 &= \left(rS_t\,dt + \sqrt{v_t}S_t\,dW_{1,t}^{\mathbb{P}}\right)^2 \\
&= r^2S_t^2\,(dt)^2 + 2rS_t\sqrt{v_t}S_t\,dt\,dW_{1,t}^{\mathbb{P}} + v_tS_t^2\,(dW_{1,t}^{\mathbb{P}})^2 \\
&= v_tS_t^2\,dt,
\end{aligned} \tag{1.3}$$

and

$$
\begin{aligned}
dv_t &= \kappa(\theta - v_t)\,dt + \sigma\sqrt{v_t}\,dW^{\mathbb{P}}_{2,t} \\
&\iff \\
(dv_t)^2 &= \left(\kappa(\theta - v_t)\,dt + \sigma\sqrt{v_t}\,dW^{\mathbb{P}}_{2,t}\right)^2 \\
&= \kappa^2(\theta - v_t)^2\,(dt)^2 + 2\kappa(\theta - v_t)\sigma\sqrt{v_t}\,dt\,dW^{\mathbb{P}}_{2,t} + \sigma^2 v_t\,(dW^{\mathbb{P}}_{2,t})^2 \\
&= \sigma^2 v_t\,dt.
\end{aligned}
\tag{1.4}
$$

By (1.3) and (1.4)

$$
\begin{aligned}
dS_t dv_t &= \left(rS_t dt + \sqrt{v_t}S_t dW^{\mathbb{P}}_{1,t}\right)\left(\kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW^{\mathbb{P}}_{2,t}\right) \\
&= rS_t\kappa(\theta - v_t)(dt)^2 + rS_t\sigma\sqrt{v_t}dtdW^{\mathbb{P}}_{2,t} + \kappa(\theta - v_t)\sqrt{v_t}S_t dtdW^{\mathbb{P}}_{1,t} + \sigma v_t S_t dW^{\mathbb{P}}_{1,t}dW^{\mathbb{P}}_{2,t} \\
&= \sigma v_t S_t \rho dt.
\end{aligned}
$$

Substituting these into the multi-dimensional Itô [4, Thm. 4.19] gives

$$
\begin{aligned}
dU(S,v,t) = \frac{\partial U}{\partial t}dt + \frac{\partial U}{\partial S}\left(rSdt + \sqrt{v}SdW^{Q}_1\right) + \frac{\partial U}{\partial v}\left(\kappa(\theta - v)dt + \sigma\sqrt{v}dW^{\mathbb{P}}_2\right) \\
+\frac{1}{2}\left[\frac{\partial^2 U}{\partial S^2}vS^2 dt + 2\frac{\partial^2 U}{\partial S\partial v}\sigma v S\rho dt + \frac{\partial^2 U}{\partial v^2}\sigma^2 v dt\right].
\end{aligned}
$$

Collecting all terms involving dt, noting that under the risk-neutral measure, $dU(S,v,t) = rU(S,v,t)dt$ as $\mathbb{E}^{\mathbb{Q}}_t[dU(S,v,t)] = rUdt$, yields

$$
\frac{\partial U}{\partial t} + rS\frac{\partial U}{\partial S} + \kappa(\theta - v)\frac{\partial U}{\partial v} + \frac{1}{2}\left(\frac{\partial^2 U}{\partial S^2}vS^2 + \frac{\partial^2 U}{\partial v^2}\sigma^2 v + 2\frac{\partial^2 U}{\partial S\partial v}\sigma v S\rho\right) = rU. \tag{1.5}
$$

The boundary conditions for the PDE in (1.5) are as follows: At $t = T$, the call option is worth its intrinsic value $U(S,v,T) = \max(0, S - K)$. When the $S_t = 0$, the call option has no value. As the stock price increases, the delta of the call option approaches one. When volatility becomes very large, the value of the call option converges to the stock price. Consequently, the boundary conditions are:

$$
U(0,v,t) = 0, \quad \frac{\partial U}{\partial S}(\infty, v, t) = 1, \quad U(S,\infty,t) = S.
$$

Now, let $x = \log S$ as it will preserve the ordering by monotonocity. By Itô [4, Prop. 4.12], the multiplication table of said Proposition, using the dynamics in (1.1), we achieve

$$
\begin{aligned}
dx_t &= \frac{1}{S_t}dS_t + \frac{1}{2}\left(-\frac{1}{S_t^2}\right)(dS_t)^2 \\
&= \frac{1}{S_t}\left(rS_t dt + \sqrt{v_t}S_t dW^{\mathbb{Q}}_{1,t}\right) + \frac{1}{2}\left(-\frac{1}{S_t^2}\right)\left(rS_t dt + \sqrt{v_t}S_t dW^{\mathbb{Q}}_{1,t}\right)^2 \\
&= \left(r - \frac{1}{2}v_t\right)dt + \sqrt{v_t}dW^{\mathbb{Q}}_{1,t}
\end{aligned}
$$

The price of the European time-t call using the dynamics of x_t can then be written as

$$Call(t) = e^x P_1 - Ke^{-r\tau} P_2 = U_1 - U_2, \tag{1.6}$$

where $U_1 = e^x P_1$ and $U_2 = Ke^{-r\tau} P_2$ are two separate options. U_1 is SP_1, which is the price of an asset that pays out $S(T)$ at maturity if $S(T) > K$, 0 otherwise. U_2 is $Ke^{-r(T-t)}P_2$, which corresponds to a binary option that pays out the present value of the exercise price if $S(T) > K$, 0 otherwise. Since both options are functions of S and v they have to fulfil the conditions of the PDE in (1.5). Rewriting the Heston PDE of (1.5) in terms of time to maturity, $\tau = T - t$ and the transformed stock price, $x = \log S$ yields

$$-\frac{\partial U}{\partial \tau} + (r - \frac{1}{2}v) + [\kappa(\theta - v)]\frac{\partial U}{\partial v} + \frac{1}{2}v\frac{\partial^2 U}{\partial x^2} + \frac{1}{2}\sigma^2 v\frac{\partial^2 U}{\partial v^2} + \rho\sigma v\frac{\partial^2 U}{\partial x \partial v} = rU, \tag{1.7}$$

U_1, U_2 have to fulfil the PDE of (1.7). We thus find the relevant derivatives:

$$\frac{\partial U_1}{\partial x} = e^x\left(P_1 + \frac{\partial P_1}{\partial x}\right), \qquad \frac{\partial^2 U_1}{\partial x^2} = e^x\left(P_1 + 2\frac{\partial P_1}{\partial x} + \frac{\partial^2 P_1}{\partial x^2}\right), \qquad \frac{\partial U_1}{\partial v} = e^x\frac{\partial P_1}{\partial v},$$
$$\frac{\partial^2 U_1}{\partial v^2} = e^x\frac{\partial^2 P_1}{\partial v^2}, \qquad \frac{\partial U_1}{\partial \tau} = e^x\frac{\partial P_1}{\partial \tau}, \qquad \frac{\partial^2 U_1}{\partial x \partial v} = e^x\left(\frac{\partial P_1}{\partial v} + \frac{\partial^2 P_1}{\partial x \partial v}\right)$$

We substitute the derivatives into the PDE in (1.7) and divide by e^x to obtain

$$\begin{aligned} -\frac{\partial P_1}{\partial \tau} + \left(r - \frac{1}{2}v\right)\left(P_1 + \frac{\partial P_1}{\partial x}\right) + \frac{1}{2}v\left(P_1 + 2\frac{\partial P_1}{\partial x} + \frac{\partial^2 P_1}{\partial x^2}\right) \\ + (\kappa(\theta - v))\frac{\partial P_1}{\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 P_1}{\partial v^2} + \rho\sigma v\left(\frac{\partial P_1}{\partial v} + \frac{\partial^2 P_1}{\partial x \partial v}\right) = rP_1 \end{aligned}$$

$$\Longleftrightarrow$$

$$-\frac{\partial P_1}{\partial \tau} + \left(r + \frac{1}{2}v\right)\frac{\partial P_1}{\partial x} + \frac{1}{2}v\frac{\partial^2 P_1}{\partial x^2} + (\kappa\theta - (\kappa - \rho\sigma)v)\frac{\partial P_1}{\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 P_1}{\partial v^2} + \rho\sigma v\frac{\partial^2 P_1}{\partial x \partial v} = 0. \tag{1.8}$$

Exactly the same is done to find the PDE that P_2 has to fulfil

$$-\frac{\partial P_2}{\partial \tau} + \left(r - \frac{1}{2}v\right)\frac{\partial P_2}{\partial x} + \frac{1}{2}v\frac{\partial^2 P_2}{\partial x^2} + (\kappa\theta - \kappa v)\frac{\partial P_2}{\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 P_2}{\partial v^2} + \rho\sigma v\frac{\partial^2 P_2}{\partial x \partial v} = 0. \tag{1.9}$$

The PDE's (1.8) and (1.9) are almost identical in form. Therefore, combine formulas for the PDE for $j = 1, 2$

$$-\frac{\partial P_j}{\partial \tau} + (r + u_j v)\frac{\partial P_j}{\partial x} + \frac{1}{2}v\frac{\partial^2 P_j}{\partial x^2} + (a - b_j v)\frac{\partial P_j}{\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 P_j}{\partial v^2} + \rho\sigma v\frac{\partial^2 P_j}{\partial x \partial v} = 0, \tag{1.10}$$

with

$$u_1 = 1/2, \quad u_2 = -1/2, \quad a = \kappa\theta, \quad b_1 = \kappa - \rho\sigma, \quad b_2 = \kappa,$$

where

$$P_j(x, v, T; \log K) = \mathbb{P}(x_T \geq \log K \mid x_t = x, v_t = v).$$

It is not possible to find a closed-form formula for the probabilities P_j, $j = 1, 2$ directly, but it can be done using their characteristic function [11]. Since the two characteristic functions are derived assets that depend on the same state variables as P_j they must satisfy the following PDE, replacing P_j with the characteristic function Ψ_j, in (1.10)

$$-\frac{\partial \Psi_j}{\partial \tau} + (r + u_j v)\frac{\partial \Psi_j}{\partial x} + \frac{1}{2}v\frac{\partial^2 \Psi_j}{\partial x^2} + (a - b_j v)\frac{\partial \Psi_j}{\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 \Psi_j}{\partial v^2} + \rho\sigma v\frac{\partial^2 \Psi_j}{\partial x \partial v} = 0, \quad (1.11)$$

with terminal condition $\Psi_j(x, v, T; u) = e^{iux}$, $u \in \mathbb{R}$.

The Heston European call price Following [15], we guess a solution for (1.11) as

$$\Psi_j(x, v, t; u) = e^{C_j(\tau;u)+D_j(\tau;u)v+iux}, \quad j = 1, 2.$$

It must hold, for the terminal conditions for the characteristic function to hold, that

$$C(0; u) = 0 \quad \text{and} \quad D(0; u) = 0.$$

The parameters a, b_j, u_j for $j = 1, 2$ are

$$u_1 = \frac{1}{2}, \quad u_2 = -\frac{1}{2}, \quad a = \kappa\theta, \quad b_1 = \kappa + \Lambda - \rho\sigma, \quad b_2 = \kappa + \Lambda,$$

and

$$\begin{aligned}
C_j(\tau; u) &= rui\tau + \frac{a}{\sigma^2}\left((b_j - \rho\sigma ui + d_j)\tau - 2\log\left(\frac{1 - g_j e^{d_j \cdot \tau}}{1 - g_j}\right)\right), \\
D_j(\tau; u) &= \frac{b_j - \rho\sigma ui + d_j}{\sigma^2} \cdot \frac{1 - e^{d_j \cdot \tau}}{1 - g_j e^{d_j \cdot \tau}}, \\
g_j &= \frac{b_j - \rho\sigma ui + d_j}{b_j - \rho\sigma ui - d_j}, \quad d_j = \sqrt{(\rho\sigma ui - b_j)^2 - \sigma^2(2u_j ui - u^2)},
\end{aligned}$$

where $C_j(\tau; u)$ and $D_j(\tau; u)$ are functions that solve a system of Riccati DE's. Appendix B has the solution accessible by a hyperlink as it is not financially relevant.

Let $k = \log K$, the conditional probability that the call option ends ITM, $Q_j = P_j(x_T \geq k \mid x_t = x, v_t = v)$, is found by a inverse Fourier transformation

$$Q_j(x, v, t; k) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty \Re\left\{\frac{e^{-iuk}\Psi_j(x, v, t; u)}{iu}\right\}du, \quad j = 1, 2$$

yielding the pricing formula

$$Call(t) = S_t Q_1(x, v, t; k) - Ke^{-r(T-t)}Q_2(x, v, t; k), \quad (1.12)$$

From [1], Heston's original formula is unstable for subsets of the parameter space. [1] solves this branch cut issue by a sign change in d in Heston's original formula which in turn yields a pricing formula that is stable over the entire domain. This is possible as the complex root, d, can partake two distinct values. We will use this representation.

[19] demonstrate that the integral converges rapidly which can be seen heuristically in Figure 1 for increasing upper integration bounds, ceteris paribus.

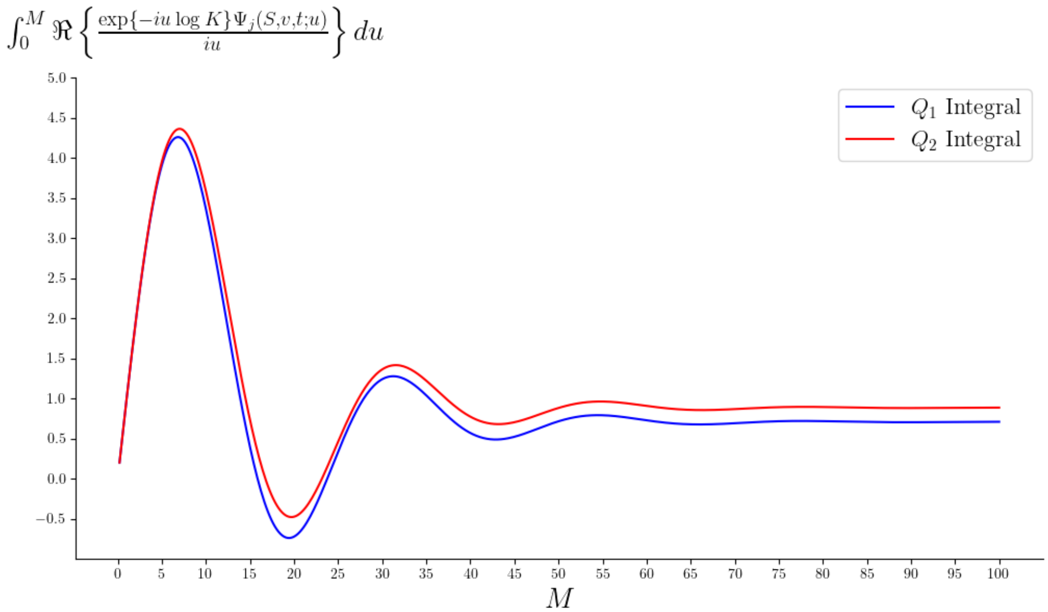

Figure 1: *Q_j convergences in increasing bounds. The fixed parameters used in the computation are set to: $\kappa = 2$, $\theta = 0.02$, $\rho = -0.5$, $\sigma = 0.3$, $\tau = 0.5$, $r = 0$, $v_0 = 0.01$, $S = 100$, and $K = 80$.*

2 Option Pricing Using Simulation

To simulate paths for (S_t, v_t) at discrete times $\mathscr{T} = \{t_i\}_{i=1}^N$, we generate random samples of $(S_{t+\Delta}, v_{t+\Delta})$ given (S_t, v_t) for any increment Δ. Repeatedly appending increments constructs the complete path $(S_t, v_t)_{t\in\mathscr{T}}$.

2.1 Log-Euler

Let $(\Omega, \mathcal{F}, \{\mathcal{F}\}_{t\geq 0}, \mathbb{P})$ be a filtered probability space. Assume X_t is driven by the SDE

$$dX_t = \underbrace{\mu(X_t, t)}_{\text{drift}} dt + \underbrace{\sigma(X_t, t)}_{\text{diffusion}} dW_t^{\mathbb{P}}, \tag{2.1}$$

where $W_t^{\mathbb{P}}$ is a Wiener process. Equally-spaced time increments are used for notational convenience, allowing us to write $t_i - t_{i-1} := \Delta$.

Integrating (2.1) from t to $t + \Delta$ yields

$$X_{t+\Delta} = X_t + \int_t^{t+\Delta} \mu(X_u, u)du + \int_t^{t+\Delta} \sigma(X_u, u)dW_u^{\mathbb{P}}. \tag{2.2}$$

At time t, $\hat{X}_t$ is known. We aim to obtain the incremented, $\hat{X}_{t+\Delta}$. Euler scheme approximates the integrals using the left end-point rule, such that the deterministic integral of (2.2) is approximated as the product of the integrand at time t and the integration range Δ

$$\int_t^{t+\Delta} \mu(X_t, u)du \approx \mu(X_t, t) \int_t^{t+\Delta} du = \mu(X_t, t)\Delta$$

Left end-points is a natural candidate as at time t the value of $\mu(X_t, t)$ is known. Now, let $Z^{\mathbb{P}} \sim \mathcal{N}(0,1)$. The stochastic integral is approximated as

$$\int_t^{t+\Delta} \sigma(X_u, u) dW_u^{\mathbb{P}} \approx \sigma(X_t, u) \int_t^{t+\Delta} dW_u^{\mathbb{P}} = \sigma(X_t, u)(W_{t+\Delta}^{\mathbb{P}} - W_t^{\mathbb{P}}) = \sigma(X_t, u)\sqrt{\Delta} Z^{\mathbb{P}},$$

because $W_{t+\Delta}^{\mathbb{P}} - W_t^{\mathbb{P}}$ and $\sqrt{\Delta} Z^{\mathbb{P}}$ are identically distributed [4, Def. 4.3]. Assembling the results yields the general form of the Euler discretization scheme:

$$\hat{X}_{t+\Delta} = \hat{X}_t + \mu(X_t, t)\Delta + \sigma(X_t, t)\sqrt{\Delta} Z^{\mathbb{P}}.$$

Discretization scheme for v_t Applying Euler discretization to dv_t in (1.1) by substituting the diffusion and drift of dv_t into $\hat{X}_{t+\Delta}$, keeping track of indices of the Wiener processes:

$$\hat{v}_{t+\Delta} = \hat{v}_t + \kappa(\theta - \hat{v}_t)\Delta + \sigma\sqrt{\hat{v}_t \Delta} Z_v^{\mathbb{P}}, \quad Z_v^{\mathbb{P}} \sim \mathcal{N}(0,1).$$

where $Z_S^{\mathbb{P}}, Z_v^{\mathbb{P}} \sim \mathcal{N}(0,1)$.

Discretization scheme for S_t We use a more robust approximation than simply applying Euler directly following [2, p. 5], [23, p. 5], [20, p. 183]. Let $x_t = \ln(S_t)$. From section 1.2

$$dx_t = \left(\mu - \frac{v_t}{2}\right) dt + \sqrt{v_t}\, dW_{1,t}^{\mathbb{P}}.$$

Integrating, the SDE for x_t

$$x_t = x_0 + \int_0^t \left(\mu - \frac{v_s}{2}\right) ds + \int_0^t \sqrt{v_s}\, dW_{1,t}^{\mathbb{P}}.$$

The solution for S_t is then

$$S_t = S_0 \exp\left(\int_0^t \left(\mu - \frac{v_s}{2}\right) ds + \int_0^t \sqrt{v_s}\, dW_{1,t}^{\mathbb{P}}\right).$$

Taking log and applying the Euler discretization we achieve

$$\log(\hat{S}_{t+\Delta}) = \log(\hat{S}_t) + \left(\mu - \frac{1}{2}\hat{v}_t\right)\Delta + \sqrt{\hat{v}_t \Delta} Z_S^{\mathbb{P}}, \quad Z_S^{\mathbb{P}} \sim \mathcal{N}(0,1).$$

Full truncation and Cholesky Decomposition To simulate, start with deterministic initial values S_0 and v_0. To generate $Z_v^{\mathbb{P}}$ and $Z_S^{\mathbb{P}}$ with correlation ρ, generate $Z_1^{\mathbb{P}}, Z_2^{\mathbb{P}} \sim \mathcal{N}(0,1)$, $Z_1^{\mathbb{P}} \perp Z_2^{\mathbb{P}}$, and set $Z_v^{\mathbb{P}} = Z_1^{\mathbb{P}}$ and $Z_S^{\mathbb{P}} = \rho Z_1^{\mathbb{P}} + \sqrt{1-\rho^2} Z_2^{\mathbb{P}}$ via a Cholesky decomposition [14, p. 43] to achieve a equivalent representation. Since $\mu = r$ under a EMM

$$\log(\hat{S}_{t+\Delta}) = \log(\hat{S}_t) + \left(\mu - \frac{1}{2}\hat{v}_t\right)\Delta + \sqrt{\hat{v}_t \Delta} Z_S^{\mathbb{Q}},$$
$$\hat{v}_{t+\Delta} = \hat{v}_t + \kappa(\theta - \hat{v}_t)\Delta + \sigma\sqrt{\hat{v}_t \Delta} Z_v^{\mathbb{Q}}.$$

In practice we will observe negative values of the variance even when Theorem 1.1 holds [2, p. 2]. When this occurs we use *full truncation* yielding the complete scheme

$$\log(\hat{S}_{t+\Delta}) = \log(\hat{S}_t) + \left(\mu - \frac{1}{2}\hat{v}_t^+\right)\Delta + \sqrt{\hat{v}_t^+\Delta} Z_S^{\mathbb{Q}},$$
$$\hat{v}_{t+\Delta} = \hat{v}_t + \kappa(\theta - \hat{v}_t^+)\Delta + \sigma\sqrt{\hat{v}_t^+\Delta} Z_v^{\mathbb{Q}}.$$

Other methods are seen in Table **1** for the form of the variance

$$\begin{aligned}\hat{v}_{t+\Delta} &= f_1(\hat{v}_t) + \kappa(\theta - f_2(\hat{v}_t))\Delta + \sigma\sqrt{f_3(\hat{v}_t)\Delta} Z_v^{\mathbb{Q}} \\ v_{t+\Delta} &= f_3(\hat{v}_{t+\Delta})\end{aligned}$$

Full truncation produces the least biased results for the variance [20].

Scheme	**Paper**	$f_1(x)$	$f_2(x)$	$f_3(x)$
Absorption	N/A	x^+	x^+	x^+
Reflection	[9], [10], [3]	$\|x\|$	$\|x\|$	$\|x\|$
Higham and Mao	[16]	x	x	$\|x\|$
Partial Truncation	[8]	x	x	x^+
Full Truncation	[20]	x	x^+	x^+

Table 1: *Euler scheme negative variance handling methods.*

2.2 Quadratic-Exponential: Martingale Corrected

Properties of v_t v_t follows a CIR process which has special distributional properties.

Proposition 2.1. *Let $F_{\chi_d'^2}(y; d, \lambda)$ be the cumulative distribution function (CDF) of the non-central chi-squared distribution with NCP λ and d DOF. Let $t < T$. $v_T \mid v_t$ has CDF*

$$\mathbb{P}\left(v_T \leq x \mid v_t\right) = F_{\chi'^2_{\frac{4\kappa\theta}{\sigma^2}}}\left(\frac{4\kappa}{\sigma^2(1 - e^{-\kappa(T-t)})}x; \frac{4\kappa\theta}{\sigma^2}, l(t,T)v_t\right),$$

where $l(t,T) := \frac{4\kappa e^{-\kappa(T-t)}}{\sigma^2\left(1 - e^{-\kappa(T-t)}\right)}$.

Next, Consider the result for the moments of $v_T \mid v_t$, $T > t$ [7, pp. 391-392].

Corollary 2.1. *Let $T > t$. Conditional on $v_T \mid v_t$ has the following first two moments:*

$$\begin{aligned}\mathbb{E}[v_T \mid v_t] &= \theta + (v_t - \theta)e^{-\kappa(T-t)}, \\ \mathbb{V}[v_T \mid v_t] &= \frac{v_t\sigma^2 e^{-\kappa(T-t)}}{\kappa}\left(1 - e^{-\kappa(T-t)}\right) + \frac{\theta\sigma^2}{2\kappa}\left(1 - e^{-\kappa(T-t)}\right)^2.\end{aligned}$$

Writing Corollary 2.1 in terms of Δ introduced in section 2, define

$$m := \theta + (\hat{v}_t - \theta)e^{-\kappa\Delta}, \tag{2.3}$$

$$s^2 := \frac{\hat{v}_t \sigma^2 e^{-\kappa\Delta}}{\kappa}\left(1 - e^{-\kappa\Delta}\right) + \frac{\theta\sigma^2}{2\kappa}\left(1 - e^{-\kappa\Delta}\right)^2, \tag{2.4}$$

$$\psi := \frac{s^2}{m^2} \tag{2.5}$$

Lastly, consider Proposition 2.1. From [18, p. 450], $\lim_{\lambda\to\infty} \chi_d'^2(\lambda)$ approaches a normal distribution. A normal r.v. with moments matched from Corollary 2.1 could be a proxy $v_{t+\Delta}$ for large v_t, while a χ_d^2-distributed r.v. proxies $v_{t+\Delta}$ for small v_t as $\lambda = l(t, T)v_t \to 0$. Recall the χ_d^2 density

$$f_{\chi_d^2}(x) = \frac{1}{2^{d/2}\Gamma(d/2)} e^{-x/2} x^{d/2-1}, \tag{2.6}$$

where $x^{d/2-1}$ dominates for small v_t. Moment-matched lognormal and Gaussian methods, shown in Figure 2, fail to fit accurately without any further optimization.

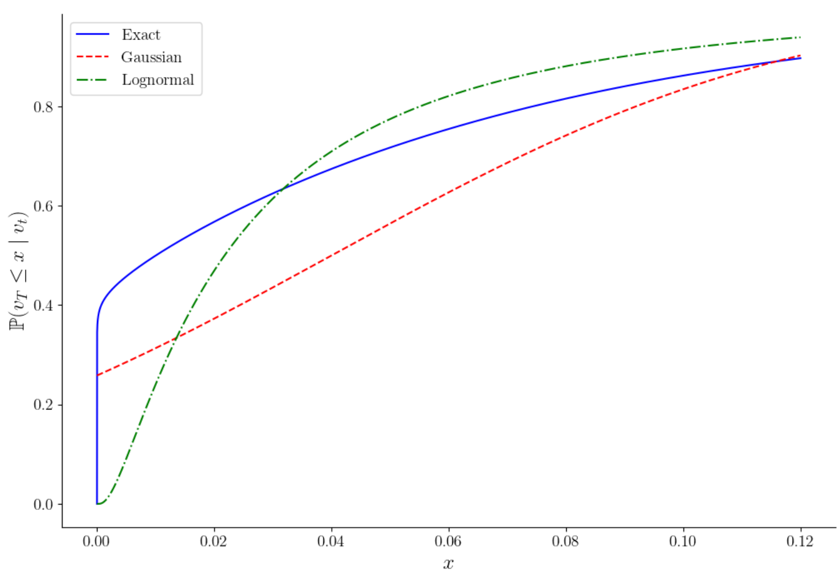

Figure 2: *CDF for $v_T \mid v_t$, with $T = 1$. The laws were parametrized by matching mean and variances to the exact distribution of v_T. Model parameters are: $\theta = 0.04$, $v_0 = 0.04$, $\kappa = 0.5$, and $\sigma = 1$.*

Discretization scheme for v_t We approximate the sampling of the true law by draws from a related moment-matched distribution. Depending on the size of λ defined in Proposition 2.1, we sample from one of the two methods:

(i): By [22, pp. 232], a $\chi_d'^2(\lambda)$-distributed r.v. with moderate/high λ-parameter can be represented well by a power-function applied to a Gaussian r.v.. By *sufficiently large* values of $\hat{v}_t$, a sample of the increment can be generated by

$$\hat{v}_{t+\Delta} = a(b + Z_v)^2, \tag{2.7}$$

where $Z_v \sim \mathcal{N}(0, 1)$. a, b are constants we determine by moment-matching.

(ii): For low values of $\hat{v}_t$, the moment-matching is inaccurate. Therefore, use the exponential tail in the density in (2.6) and a point-mass using a Dirac delta function, δ, at the origin to (asymptotically) approximate the density of $\hat{v}_{t+\Delta}$

$$\mathbb{P}(\hat{v}_{t+\Delta} \in [x, x+dx]) \approx \left(p\delta(0) + \beta(1-p)e^{-\beta x}\right) dx, \quad x, p, \beta \geq 0. \tag{2.8}$$

The constants a, b, p, β can locally be chosen such that the first two moments of the approximate distribution matches those of the exact distribution.

How do we sample from the above distributions?

(i): We draw $Z_v \sim \mathcal{N}(0,1)$ and apply (2.7).

(ii): To sample according to (2.8) we have to invert the CDF. The CDF is obtained by integrating the PDF

$$\begin{aligned} L(x) &:= \mathbb{P}(\hat{v}_{t+\Delta} \leq x) = p + \int_0^x \beta(1-p) \exp\{-\beta y\}\, dy \\ &= p + (1-p)\left(1 - \exp\{-\beta x\}\right), \quad x \geq 0. \end{aligned}$$

To sample, invert $L(x)$ and let $u \sim U_v \sim \mathcal{U}(0,1)$, to obtain the following inverse CDF

$$L^{-1}(u; p, \beta) = \begin{cases} 0 & \text{if } 0 \leq u \leq p, \\ \beta^{-1} \log\left(\frac{1-p}{1-u}\right) & \text{if } p < u \leq 1. \end{cases} \Rightarrow \hat{v}_{t+\Delta} = L^{-1}(U_v, p, \beta). \tag{2.9}$$

We now turn to determine the constants a, b, p and β as well as the sampling switching rule which clarifies what we meant by *sufficiently large*. To determine a and b we employ the method of moment-matching. To show the Proposition we need a Lemma.

Lemma 2.1. *Let $a, b \in \mathbb{R}$ and $Z \sim \mathcal{N}(0,1)$. If $X = a(b+Z)^2$, then $aX \sim \chi_1'^2(b^2)$.*

Proposition 2.2 (*Quadratic* scheme)**.** *Let m, s^2 and ψ be as defined in (2.3), (2.4) and (2.5), respectively. If $\psi \leq 2$, set*

$$b^2 = \frac{2}{\psi} - 1 + \sqrt{\frac{2}{\psi}}\sqrt{\frac{2}{\psi} - 1} \geq 0 \quad \textit{and} \quad a = \frac{m}{1+b^2}.$$

Let $\hat{v}_{t+\Delta}$ be as defined in (2.7) and m, s^2 and ψ be as defined in (2.3), (2.4) and (2.5), respectively. In this case $\mathbb{E}[\hat{v}_{t+\Delta}] = m$ and $\mathbb{V}[\hat{v}_{t+\Delta}] = s^2$.

Proposition 2.2 requires $\psi \leq 2$. For large ψ (small $\hat{v}_t$) we need a different scheme.

Proposition 2.3 (*Exponential* scheme)**.** *Let m, s^2, and ψ be as defined in equations (2.3), (2.4), and (2.5), respectively. Assume $\psi \geq 1$ and set*

$$p = \frac{\psi - 1}{\psi + 1} \in [0, 1) \quad \textit{and} \quad \beta = \frac{1-p}{m} = \frac{2}{m(\psi+1)} > 0.$$

Let $\hat{v}_{t+\Delta}$ be defined as the sampling of the distribution in (2.9). In this case, $\mathbb{E}[\hat{v}_{t+\Delta}] = m$ and $\mathbb{V}[\hat{v}_{t+\Delta}] = s^2$.

Next, the *switching rule*: A rule to swap between quadratic/exponential schemes. We proved that the quadratic scheme, (2.7), can only be applied for values $\psi \leq 2$ and the exponential scheme, (2.9), can only be moment-matched for $\psi \geq 1$. These domains do fortunately overlap. As such, we need a rule to determine which of schemes to apply. Choose some benchmark $\psi_c \in [1, 2]$ and define the sampling rule as

$$\hat{v}_{t+\Delta} = \begin{cases} a(b + Z_v)^2 & \text{if } \psi \leq \psi_c, \\ \begin{cases} 0 & \text{if } 0 \leq u \leq p, \\ \beta^{-1} \log\left(\frac{1-p}{1-u}\right) & \text{if } p < u \leq 1 \end{cases} & \text{if } \psi > \psi_c. \end{cases},$$

For numerical test we set $\psi_c = 1.5$. [2, p. 16] notes that $\psi_c \in [1, 2]$ means little for the test as both sampling schemes are well approximations of the exact distribution.

Now, note that

$$\psi = \frac{s^2}{m^2} = \frac{\frac{\hat{v}_t \sigma^2 e^{-\kappa\Delta}}{\kappa}\left(1 - e^{-\kappa\Delta}\right) + \frac{\theta\sigma^2}{2\kappa}\left(1 - e^{-\kappa\Delta}\right)^2}{\left(\theta + (\hat{v}_t - \theta)e^{-\kappa\Delta}\right)^2},$$

has the property for fixed $\hat{v}_t$ that $\lim_{\Delta \to 0} \psi = 0$. The limit of ψ implies that for a more refined Δ, the need to use the exponential scheme diminish. To see the accuracy of the QE scheme we recreate Figure 2 in Figure 3 with the two conditions for ψ. Notice how the QE-distribution is almost exactly equal to that of the exact distribution.

Step-by-Step Guide $\hat{v}_{t+\Delta}$

1. Given $\hat{v}_t$, compute m and s^2 from (2.3) and (2.4), respectively.
2. Compute $\psi = s^2/m^2$ from (2.5).
3. Draw $U_v \sim \mathcal{U}(0, 1)$.
4. **If $\psi \leq \psi_c$ use quadratic sampling:**
   (A) Compute a and b according to Proposition 2.2.
   (B) Compute $Z_v = \Phi^{-1}(U_v)$.
   (C) Set $\hat{v}_{t+\Delta} = a(b + Z_v)^2$.
5. **If $\psi > \psi_c$ use exponential sampling:**
   (A) Compute β and p according to Proposition 2.3.
   (B) Set
   $$\hat{v}_{t+\Delta} = \begin{cases} 0 & \text{if } 0 \leq u \leq p, \\ \beta^{-1} \log\left(\frac{1-p}{1-u}\right) & \text{if } p < u \leq 1. \end{cases}.$$

This completes the discretization of v_t.

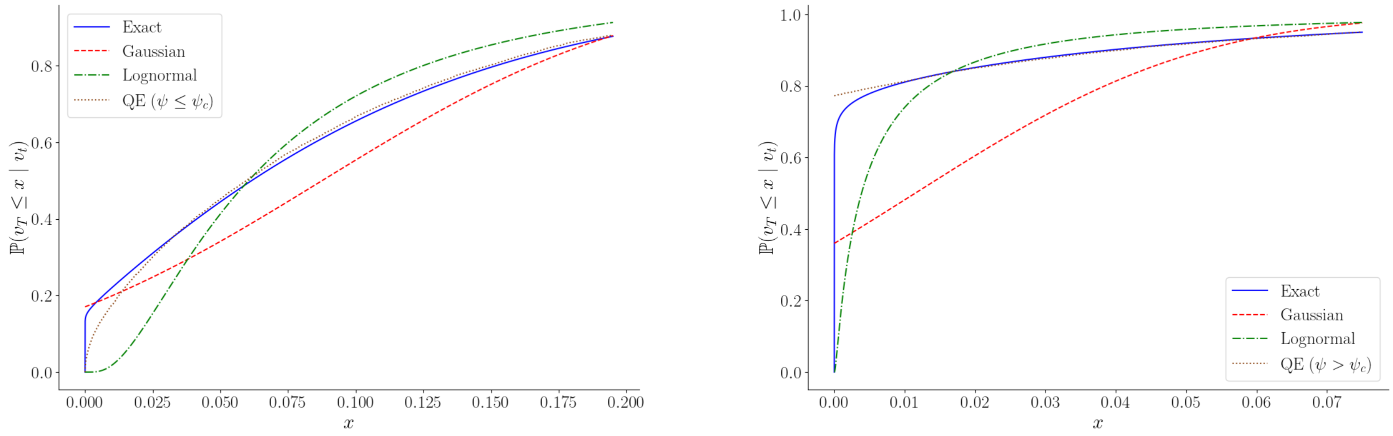

Figure 3: *CDF for $v_T \mid v_t$, with $T = 0.1$. v_0 is 0.01 in the left panel and 0.09 in the right. Laws were parametrized by moment-matching to exact law of v_T. Model parameters are: $\theta = 0.04$, $\kappa = 0.5$, and $\sigma = 1$.*

Discretization scheme for S_t The QE scheme approximates the exact solution to the (log) CIR-process found in [6, p. 219].

Proposition 2.4. *The exact solution of the log asset process in the Heston model is*

$$\log S_{t+\Delta} = \log S_t + \frac{\rho}{\sigma}(v_{t+\Delta} - v_t - \kappa\theta\Delta) + \left(\frac{\kappa\rho}{\sigma} - \frac{1}{2}\right)\int_t^{t+\Delta} v_u du + \sqrt{1-\rho^2}\int_t^{t+\Delta} \sqrt{v_u} dW_{1,u}^{\mathbb{P}}.$$

We approximate $\int_t^{t+\Delta} v_u du \approx \Delta[\gamma_1 v_t + \gamma_2 v_{t+\Delta}]$ by a central approximation $\gamma_1 = \gamma_2 = 0.5$. As $W_1^{\mathbb{P}}$ is independent of v, it follows by [4, Lem. 4.18] that conditional on v_t and $\int_t^{t+\Delta} v_u du$, that the itô integral $\int_t^{t+\Delta} \sqrt{v_u} dW_{1,u}^{\mathbb{P}} \sim \mathcal{N}\left(0, \int_t^{t+\Delta} v_u du\right)$. To include a interest rate in the scheme, consider the following Lemma.

Lemma 2.2. *Assume that the forward price process for the stock at expiry-T is given by $F_t^T = S_t \exp\left(r(T-t)\right)$. Applying a discretization with increments Δ, we derive that:*

$$\hat{S}_{t+\Delta} = \hat{S}_t \exp(r\Delta + \bullet),$$

where $\bullet$ represents the terms by the discretization, meaning, we can simply add the interest rate as $r\Delta$ linearly to the scheme terms.

By our approximations

$$\begin{aligned}
\log \hat{S}_{t+\Delta} &= \log \hat{S}_t + \frac{\rho}{\sigma}(\hat{v}_{t+\Delta} - \hat{v}_t - \kappa\theta\Delta) + \Delta\left(\frac{\kappa\rho}{\sigma} - \frac{1}{2}\right)(\gamma_1\hat{v}_t + \gamma_2\hat{v}_{t+\Delta}) + \sqrt{\Delta}\sqrt{1-\rho^2}\sqrt{\gamma_1\hat{v}_t + \gamma_2\hat{v}_{t+\Delta}} \cdot Z \\
&= \log \hat{S}_t + K_0 + K_1\hat{v}_t + K_2\hat{v}_{t+\Delta} + \sqrt{K_3\hat{v}_t + K_4\hat{v}_{t+\Delta}} \cdot Z \\
&\overset{\dagger}{\iff} \\
\log \hat{S}_{t+\Delta} &= \log \hat{S}_t + \Delta r + K_0 + K_1\hat{v}_t + K_2\hat{v}_{t+\Delta} + \sqrt{K_3\hat{v}_t + K_4\hat{v}_{t+\Delta}} \cdot Z, \qquad (2.10)
\end{aligned}$$

where $Z \sim \mathcal{N}(0,1)$ and † follows from Lemma 2.2 and defining

$$K_0 = -\frac{\rho\kappa\theta}{\sigma}\Delta, \qquad K_1 = \gamma_1\Delta\left(\frac{\kappa\rho}{\sigma} - \frac{1}{2}\right) - \frac{\rho}{\sigma},$$
$$K_2 = \gamma_2\Delta\left(\frac{\kappa\rho}{\sigma} - \frac{1}{2}\right) + \frac{\rho}{\sigma}, \quad K_3 = \gamma_1\Delta\left(1-\rho^2\right), \quad K_4 = \gamma_2\Delta\left(1-\rho^2\right).$$

Step-by-Step Guide $\hat{S}_{t+\Delta}$:

1. Given $\hat{v}_t$, generate $\hat{v}_{t+\Delta}$ using QE
2. Draw a uniform random number U_S that is independent of all random numbers used for $\hat{v}_{t+\Delta}$.
3. Set $Z = \Phi^{-1}(U_S)$.
4. Given $\log \hat{S}_t, \hat{v}_t$, and $\hat{v}_{t+\Delta}$ from step 1, compute $\log \hat{S}_{t+\Delta}$ from (2.10).

Martingale correction of the QE scheme The QE discretized form of the price, $\hat{S}_{t+\Delta}$ is not a Martingale. Taking exp, the scheme for $\hat{S}_{t+\Delta}$ in (2.10) is equivalent to

$$\hat{S}_{t+\Delta} = \hat{S}_t \exp\left\{\Delta r + K_0 + K_1\hat{v}_t\right\} \exp\left\{K_2\hat{v}_{t+\Delta} + \sqrt{K_3\hat{v}_{t+\Delta} + K_4\hat{v}_{t+\Delta}}Z\right\}. \tag{2.11}$$

We need to derive a adjustment to the discretized stock process such that $\mathbb{E}_t^{\mathbb{Q}}\left[\hat{S}_{t+\Delta}\right] = \hat{S}_t$. By the Tower Law of Conditional Expectations, ignoring the implicit conditioning on $\hat{v}_t$

$$\mathbb{E}_t^{\mathbb{Q}}\left[\hat{S}_{t+\Delta}\right] = \mathbb{E}^{\mathbb{P}}\left[\mathbb{E}_t^{\mathbb{Q}}\left[\hat{S}_{t+\Delta} \mid \hat{v}_{t+\Delta}\right]\right].$$

By (2.11), the following equation needs to be satisfied:

$$\begin{aligned}
\hat{S}_{t+\Delta} &= \hat{S}_t \exp\left\{\Delta r + K_0 + K_1\hat{v}_t\right\} \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{\sqrt{K_3\hat{v}_t + K_4\hat{v}_{t+\Delta}}Z\right\} \mid \hat{v}_{t+\Delta}\right] \\
&\overset{\dagger}{=} \hat{S}_t \exp\left\{\Delta r + K_0 + K_1\hat{v}_t\right\} \mathbb{E}_t^{\mathbb{Q}}\left[\exp\left\{K_2\hat{v}_{t+\Delta} + \frac{1}{2}(K_3\hat{v}_t + K_4\hat{v}_{t+\Delta})\right\}\right] \\
&= \hat{S}_t \exp\left\{\Delta r + K_0 + \left(K_1 + \frac{1}{2}K_3\right)\hat{v}_t\right\} \Psi_{\hat{v}_{t+\Delta}}(A),
\end{aligned} \tag{2.12}$$

with $\Psi_{\hat{v}_{t+\Delta}}(A) = \mathbb{E}\left[\exp\left\{A\hat{v}_{t+\Delta} \mid \hat{v}_t\right\}\right]$ denoting the moment generating function (MGF) of the conditioned discretized variance process evaluated at $A := K_2 + \frac{1}{2}K_4$. † follows from the expectation of the lognormal distribution. For the martingale condition to be fulfilled we need the last line of (2.12) to be $\hat{S}_t \cdot 1$:

$$\exp\left\{\Delta r + K_0 + \left(K_1 + \frac{1}{2}K_2\right)\hat{v}_t\right\} \Psi_{\hat{v}_{t+\Delta}}(A) = 1,$$

which (assuming the regularity condition $\Psi_{\hat{v}_{t+\Delta}}(A) < \infty$) is satisfied by replacing K_0 by

$$K_0^* := -\log\left(\Psi_{\hat{v}_{t+\Delta}}(A)\right) - \left(\Delta r + \left(K_1 + \frac{1}{2}K_2\right)\right)\hat{v}_t + \Delta r$$

Lastly, we characterize the MGF and find conditions for the regularity condition to hold and exist.

Proposition 2.5. *Let the QE scheme be given as in section 2.2.1 given by constants in Proposition 2.4 and 2.5. Let* $\psi_c \in [1, 2]$ *and let* m*,* s^2 *and* ψ *be defined by (2.3), (2.4) and (2.5), respectively. If* $\psi \leq \psi_c$*, then*

$$\Psi_{\hat{v}_{t+\Delta}}(A) = \frac{\exp\left\{\frac{Ab^2 a}{1-2Aa}\right\}}{\sqrt{1-2Aa}}$$

where A *has to satisfy the condition* $A < \frac{1}{2a}$*. On the contrary, if* $\psi > \psi_c$*, then*

$$\Psi_{\hat{v}_{t+\Delta}}(A) = p + \frac{\beta(1-p)}{\beta - A},$$

provided that $A < \beta$ *for the fraction to be well defined.*

From Proposition 2.5, we need two regularity conditions to be satisfied in the QE-M scheme which is not always going to be held. However, in our cases of parameters, they are. Proposition 2.5 finalizes the QE-M scheme: we simply replace K_0 with

$$K_0^* = \begin{cases} -\frac{Ab^2 a}{1-2Aa} + \frac{1}{2}\log(1-2Aa) - \left(K_1 + \frac{1}{2}K_3\right)\hat{v}_t + \Delta r, & \psi \leq \psi_c \\ -\log\left(p + \frac{\beta(1-p)}{\beta - A}\right) - \left(K_1 + \frac{1}{2}K_3\right)\hat{v}_t + \Delta r, & \psi > \psi_c \end{cases}, \tag{2.13}$$

in (2.10). Notice that we can't pre-cache K_0^* as it involves $\hat{v}_t$. We discuss this later on.

2.3 Monte Carlo Simulation: Pricing and Error

Suppose we want to estimate some parameter ρ given by

$$\rho = \mathbb{E}[g(X)],$$

where $g(X)$ is an arbitrary function, e.g. a call option. One then generate N i.i.d sample values X_1, X_2, …, X_N from the probability density function $f(x)$. The estimator of the parameter ρ could then be given by the unbiased estimator

$$\hat{\rho}_N = \frac{1}{N}\sum_{i=1}^{N} g(X_i).$$

To show that $\hat{\rho}$ is an unbiased estimator of ρ, where $\rho = \mathbb{E}[g(X_i)]$, we prove that $\mathbb{E}[\hat{\rho}_N] = \rho$. By the linearity of expectation, we have

$$\mathbb{E}[\hat{\rho}_N] = \mathbb{E}\left[\frac{1}{N}\sum_{i=1}^{N} g(X_i)\right] = \frac{1}{N}\sum_{i=1}^{N}\mathbb{E}[g(X_i)].$$

Since each X_i are i.i.d, it follows that

$$\mathbb{E}[\hat{\rho}_N] = \frac{1}{N}\sum_{i=1}^{N}\rho = \frac{1}{N}N\rho = \rho.$$

Therefore, $\hat{\rho}$ is an unbiased estimator of ρ.

We cite a (Strong) Law of Large Nubmers (SLLN) [17, p. 173].

Theorem 2.1 (SLLN)**.** *Let* $(X_N)_{N\geq 1}$ *be i.i.d. and defined on the same space. Let* $\rho = \mathbb{E}[X_j]$ *and* $\sigma^2 = \sigma^2_{X_j} < \infty$*. Then,*

$$\lim_{N\to\infty} \sum_{j=1}^{N} X_j/N = \lim_{N\to\infty} \frac{1}{N} \sum_{j=1}^{N} X_j \overset{a.s.}{=} \rho.$$

Conditions in Theorem 2.1 are satisfied such that $\hat{\rho}_N \xrightarrow{a.s.} \rho$ as $N \to \infty$. The unbiased sample variance which is an estimator for the true variance σ^2 can be written as

$$s_N^2 = \frac{1}{N-1} \sum_{i=1}^{N} \Big(g(X_i) - \hat{\rho} \Big)^2.$$

We omit showing this as it is the same story, however, note that the denominator $N-1$ instead of N causes the estimator to be unbiased. The next Theorem we need is the Central Limit Theorem (CLT) [17, p. 181].

Theorem 2.2 (CLT)**.** *Let* $(X_j)_{j\geq 1}$ *be i.i.d. with* $\mathbb{E}[X_j] = \rho$ *and* $\mathbb{V}[X_j] = \sigma^2$ *with* $0 < \sigma^2 < \infty$*. Then*

$$\frac{\sum_{j=1}^{N} X_j - N\rho}{\sigma\sqrt{N}} \xrightarrow{D} \mathcal{N}(0,1).$$

Conditions in Theorem 2.2 hold, allowing us to construct confidence intervals for our estimate $\hat{\rho}_N$. Specifically, we aim to construct an interval of the form $[\hat{\rho}_N - \varepsilon, \hat{\rho}_N + \varepsilon]$ using a method that ensures, in repeated sampling, the true parameter ρ is contained in the interval $100(1-\alpha)\%$ of the time. Here, $1-\alpha$ is the confidence level, a value close to 1. Let $Z \sim \mathcal{N}(0,1)$ and note that ρ belongs to $[\hat{\rho}_N - \varepsilon, \hat{\rho}_N + \varepsilon]$ if and only if $\hat{\rho}_N$ belongs to $[\rho - \varepsilon, \rho + \varepsilon]$. The central limit theorem states that

$$\begin{aligned}
\mathbb{P}\left(\rho - \varepsilon \leq \hat{\rho}_N \leq \rho + \varepsilon\right) &= \mathbb{P}\left(-\varepsilon \leq \hat{\rho}_N - \rho \leq \varepsilon\right) \\
&= \mathbb{P}\left(-\frac{\varepsilon\sqrt{N}}{s_N} \leq \frac{\left(\hat{\rho}_N - \rho\right)\sqrt{N}}{s_N} \leq \frac{\varepsilon\sqrt{N}}{s_N}\right) \\
&\approx \mathbb{P}\left(-\frac{\varepsilon\sqrt{N}}{s_N} \leq Z \leq \frac{\varepsilon\sqrt{N}}{s_N}\right).
\end{aligned}$$

Let z_c be the number such that $P\left(-z_c \leq Z \leq z_c\right) = 1-\alpha$. Then we have $\frac{\varepsilon\sqrt{N}}{s_N} = z_c$, implying, $\varepsilon = z_c s_N/\sqrt{N}$. Thus our confidence interval for ρ by Theorem 2.2 is

$$\hat{\rho}_N \pm \frac{z_c s_N}{\sqrt{N}} \Rightarrow \mathbb{P}\left(\hat{\rho}_N - \frac{z_c s_N}{\sqrt{N}} < \rho < \hat{\rho}_N + \frac{z_c s_N}{\sqrt{N}}\right) = 1 - \alpha.$$

Common choice for z_c in option pricing is $99\% \sim 2.576$.

3 Numerical Test

3.1 Benchmark Setup

Table **2** specifies the three different considered case parameters, Table C.1 the exact prices.

Parameter	Case I	Case II	Case III
σ	1.0	1.0	0.9
κ	0.5	1.0	0.3
ρ	-0.9	-0.3	-0.5
T (years)	10	5	15
θ (%)	4.0	9.0	4.0
r	0.00	0.05	0.00
v_0 (%)	4.0	9.0	4.0
S_0		100	
N		10^6	

Table 2: *Simulation parameters for Call-10Y, Call-5Y and Call-15Y. Each case corresponds to distinct parametrizations of the Heston model.*

The data from Case I and III broadly represent the market for long-dated FX options, such as those found in power-reverse dual contracts. These settings are difficult as we combine low mean-reversion with high volatility of the variance process. Case II can correspond to an equity setting. For all test cases, we evaluate option prices at three strike levels: $K \in [60, 100, 140]$. We would have preferred to use a larger number of sample paths, given that the biases of our new scheme is relatively low (as will be demonstrated). However, practical computing constraints make it challenging to increase the number of paths. For instance, with 32 steps per year, pricing a 15-year option using the Heston model requires generating 960 random numbers per path. With 10^6 paths, this translates to approximately 10^9 random numbers and the corresponding computations to update the stock price and variance, all needed to determine a single option price. We compute the draws outside the Monte-Carlo loop in a single operation but do not apply vectorized operations which will be discussed.

3.2 Simulation Results

Following the Monte Carlo simulation theory from section 2.3, we compute an estimate of the European time-0 call price, $\widehat{Call}(0) = \mathbb{E}_0^{\mathbb{Q}}\Big((\hat{S}_T - K)^+\Big)$. Specifically, for some given discretization scheme of S_t we draw N independent samples of $\hat{S}_T^{(1)}, \hat{S}_T^{(2)}, \ldots, \hat{S}_T^{(N)}$ using an equidistant time-grid $\Big\{t_i = \frac{iT}{n},\ i = 0, \ldots, n\Big\}$ with fixed (times per year) step Δ where $n \in \mathbb{N}$. Then $\widehat{Call}(0)$ is estimated in a Monte Carlo fashion as

$$\widehat{Call}(0) \approx \frac{1}{N} \sum_{i=1}^{N} \left(\hat{S}_T^{(i)} - K\right)^+.$$

The right-hand side of this equation is a r.v. with mean $\widehat{Call}(0)$ and a standard deviation ("Monte Carlo error") of order $\mathcal{O}(N^{-1/2})$ [2, p. 24]. Using a high number of paths, we can reduce standard deviation and obtain a highly accurate estimate for $\widehat{Call}(0)$.

We compute the exact prices using (1.12). Following [23] we calculate biases using the exact time-0 call price $Call(0)$ and the estimated time-0 call price $\widehat{Call}(0)$

$$e(\Delta) = Call(0) - \widehat{Call}(0)$$

If the exact call price falls within the 99% confidence window of the Monte Carlo price, corresponding to $z_c = 2.576 = \Phi^{-1}\left(1 - \frac{0.01}{2}\right)$, meaning

$$Call(0) \in \left[\widehat{Call}(0) - \frac{z_c s_N}{\sqrt{N}}, \widehat{Call}(0) + \frac{z_c s_N}{\sqrt{N}}\right].$$

If the bias is insignificant we mark it with a star (∗). We report the confidence interval with methods developed in parenthesis.

The results can be seen in Table Table **3** or visualized in Appendix C as Figure C.1, Figure C.2 and Figure C.3. Schemes were separated in the figures as performances were extremely polarized.

		Case I				**Case II**				**Case III**			
K	Δ	Euler		QE-M		Euler		QE-M		Euler		QE-M	
		Bias ($e(\Delta)$)	Time (m)	Bias ($e(\Delta)$)	Time (m)	Bias ($e(\Delta)$)	Time (m)	Bias ($e(\Delta)$)	Time (m)	Bias ($e(\Delta)$)	Time (m)	Bias ($e(\Delta)$)	Time (m)
60	1/1	-3.158 (0.104)	0.6	-0.009* (0.064)	3.6	-1.534 (0.221)	0.49	-0.002* (0.164)	3.84	-4.056 (0.519)	0.56	-0.418 (0.133)	4.55
	1/2	-1.731 (0.084)	1.1	0.053* (0.065)	6.1	-0.773 (0.190)	0.73	-0.003* (0.163)	4.49	-2.231 (0.164)	1.21	-0.190 (0.140)	7.09
	1/4	-0.957 (0.074)	2.2	0.052* (0.065)	7.8	-0.349 (0.173)	1.30	0.047* (0.162)	8.31	-1.297 (0.164)	3.02	-0.167 (0.132)	11.70
	1/8	-0.462 (0.069)	4.3	0.008* (0.065)	13.4	-0.195 (0.166)	2.41	-0.018* (0.165)	9.91	-0.764 (0.134)	7.30	-0.146 (0.124)	20.85
	1/16	-0.268 (0.067)	9.6	-0.014* (0.064)	25.9	-0.183 (0.167)	4.56	-0.045* (0.162)	16.71	-0.518 (0.155)	13.91	-0.173 (0.128)	38.08
	1/32	-0.144 (0.066)	20.3	0.031* (0.065)	40.7	-0.086 (0.166)	9.36	-0.027* (0.169)	24.52	-0.326 (0.134)	30.14	-0.184 (0.134)	74.76
100	1/1	-6.405 (0.076)	0.6	-0.216 (0.032)	4.0	-3.427 (0.203)	0.44	0.095* (0.153)	3.12	-7.397 (0.191)	0.67	-0.039* (0.117)	4.27
	1/2	-3.741 (0.054)	1.1	-0.106 (0.034)	5.5	-1.752 (0.175)	0.70	-0.040* (0.149)	3.68	-4.463 (0.144)	2.34	-0.246 (0.124)	6.17
	1/4	-2.044 (0.044)	2.1	-0.007* (0.034)	7.6	-0.930 (0.162)	1.28	-0.076* (0.146)	4.94	-2.692 (0.130)	3.42	-0.379 (0.115)	9.81
	1/8	-1.046 (0.039)	4.2	-0.0002* (0.034)	12.5	-0.360 (0.154)	2.42	0.027* (0.157)	7.50	-1.563 (0.115)	6.85	-0.364 (0.106)	22.01
	1/16	-0.522 (0.036)	8.2	0.013* (0.034)	22.9	-0.282 (0.154)	4.92	0.070* (0.147)	12.67	-0.973 (0.117)	14.91	-0.405 (0.111)	40.54
	1/32	-0.267 (0.035)	21.7	0.003* (0.034)	43.5	-0.061 (0.151)	9.91	-0.008* (0.152)	27.35	-0.674 (0.114)	30.03	-0.420 (0.116)	78.63
140	1/1	-4.272 (0.050)	0.7	0.088* (0.006)	4.0	-4.429 (0.184)	0.44	0.501 (0.132)	4.19	-6.487 (0.206)	1.03	0.027* (0.093)	5.37
	1/2	-1.938 (0.026)	1.3	0.018* (0.007)	5.2	-2.321 (0.152)	0.68	0.286 (0.133)	5.09	-3.509 (0.136)	3.53	-0.235 (0.102)	7.76
	1/4	-0.776 (0.014)	2.3	-0.007* (0.007)	7.1	-1.094 (0.139)	1.35	0.131 (0.133)	7.43	-1.930 (0.142)	6.78	-0.429 (0.093)	10.89
	1/8	-0.266 (0.009)	4.6	0.008* (0.007)	13.2	-0.601 (0.131)	2.57	0.014* (0.134)	12.34	-1.037 (0.144)	13.12	-0.424 (0.088)	23.54
	1/16	-0.113 (0.007)	10.3	-0.015* (0.007)	23.9	-0.160 (0.133)	5.15	-0.031* (0.132)	25.62	-0.666 (0.111)	30.15	-0.413 (0.089)	42.18
	1/32	-0.067 (0.006)	20.9	0.006* (0.007)	47.3	-0.062 (0.129)	11.32	0.020* (0.134)	52.28	-0.391 (0.094)	69.02	-0.458 (0.091)	83.26

Table 3: *Bias, confidence intervals and time simulation results for the three cases presented in Table 1 with corresponding option prices in Table 2. Star denotes statistical insignificance.*

Broadly speaking, every bias is significant for the Euler scheme. Although, the bias is almost insignificant at $\Delta = 1/32$ for every Case at the 99% confidence level. Furthermore, the best performing results were are all in the equity setting, Case II, which we suspected to be the least challenging parameter subset, although not easy. The high speed of mean reversion does make the process less inconsistent even though the vol-of-variance is high. As suspected, Case III, did yield trouble with the low speed of mean reversion and high vol-of-variance. The computation times associated with the Euler method were generally lower relative to the QE-M at about 25% to 50% of that of the QE-M. The longest time was in the most challenging case, Case III, with $\Delta = 1/32$ and $K = 60$ at 29.96 minutes with the shortest at Case II with $\Delta = 1/1$ and $K = 100$, $K = 140$. From Figure C.1, Figure C.2 and Figure C.3 notice that the Euler scheme bias is always falling in refinedness of the time-step as one would expect. This does of course come at a computational cost about $2x$ when doubling the amount of yearly calculations i.e. a less refined Δ. Nevertheless, Euler scheme does converge for increasing number of calculations. Indeed, we see that the bias is over $\approx 96.4\%$ lower going from $\Delta = 1/1$ to $\Delta = 1/32$ in Case II $K = 140$ which is a impressive result. Notice, as the full truncation fix becomes deterministic with an upward drift, observe that the bias is always negative.

The QE-M scheme did perform a lot better with 30/54 insignificant biases. 2/24 significant biases were in Case I, 6/24 in Case II and the rest, 16/24, in Case III. Although the biases were low in Case III they were significant. The interesting thing to note is that the less refined times-step, $\Delta = 1/1$, did yield the only insignificant results in Case III. This is most likely due to the fact that as we stated, as $\lim_{\Delta \to 0} \psi = 0$, the need to use the exponential scheme diminish where the quadratic scheme might not be accurate at these parameter values. This could be a special case of choosing ψ_c differently yielding significant results. However, the biases in every single case are near borderline. The computational times for were shortest in Case II with $K = 100$ and $\Delta = 1/1$. The longest computation time was for Case III with $\Delta = 1/32$ and $K = 140$ at 78.99 minutes. From Figure C.1, Figure C.2 and Figure C.3, notice that the QE-M bias is in general falling in refinedness of the time-step but does not benefit as much as the Euler scheme. Inspect for example Case III; The bias is actually at some points growing when we use more yearly calculations. However, the QE-M's estimates are super accurate at every single time-step.

Both schemes did preform best for small strikes $K = 60$ and worse for large strikes $K = 140$. The reader should recall from Table C.1, high strikes yield extremely small prices, meaning the biases observed for $K = 140$ in every case are in proportion much larger than those of $K \in [60, 100]$. The bias is therefore certainly in percent decreasing in strike.

We thus see that see that there is no case dependence in terms of a scheme outperforming in a Case compared to another. However, the convergence in increasing refindedness of Δ is much more obvious for the Euler scheme than the switching scheme QE-M. The former scheme is not dynamic in the sense that a new scheme is applied depending on $\hat{v}$ and Δ where the latter absolutely is. This will, everything else equal, mean that the convergence is not going to be as obvious or predictable as the Euler scheme. As stated in [20, Thm. 4.2], the full truncation scheme converges strongly in the L^1-sense, whereas no formal convergence results are given for the QE-M scheme as such a task is not for the faint hearted.

The Euler scheme does outperform the QE-M in computation times but this is to be taken with a grain of salt. Indeed, if one were to compare the performance of the most refined time-step in the Euler scheme with the least refined time-step of the QE-M scheme one would notice that the performance is in general still better for QE-M which also yields computational times 1/8'th thoes of the Euler scheme! Take for example Case I; QE-M yields insignificant biases for $\Delta = 1/1$ and $K \in [60, 100, 140]$ where as Euler does not for $\Delta = 1/32$ and $K \in [60, 100, 140]$.

4 Discussion

This paper is a subset of a broader examination of seven numerical schemes seen in `Python` class under `HestonModel.py`, for the reader's reference. Other methods, such as the QE scheme, could also be considered. QE challenge the computational efficiency of the Euler scheme without significantly compromising performance [2, p. 29].

The practical impact of the martingale correction is often minimal, as the net drift from the martingale is typically small and can be mitigated by reducing the time step. Additionally, achieving the mean of the distribution by removing or minimizing the drift of the Wiener process does not necessarily result in more accurate option pricing. This observation could explain why, in Case III, the QE-M scheme produced insignificant results only for the coarser time steps. Moreover, as stated in Proposition 2.7, two regularity conditions must be satisfied to apply the martingale correction for the QE-M scheme. In cases where these conditions are violated, the QE scheme would be required.

In terms of computational efficiency, removing the martingale correction for the QE-M scheme to obtain the QE scheme simplifies calculations by eliminating the need to compute K_0^* from Equation (2.13). Instead, K_0 can be pre-cached, as it does not depend on $\hat{v}_{t+\Delta}$. However, as noted in the penultimate paragraph of Section 3.2, the Euler scheme is heavily reliant on the refinement of Δ. For instance, in one case, reducing Δ from 1/1 to 1/32 decreased the bias by approximately 98.6%. By contrast, the QE scheme generally benefits from finer time steps, but its performance at coarser steps is still remarkable.

This paper, while academic in nature, also has implications for practitioners. A significant disadvantage of using Monte Carlo methods for option pricing is their computational intensity. However, optimizations such as leveraging `numpy` vector operations can yield substantial improvements. For reference, the `Euler Scheme.py` file on our `GitHub` completes all cases for combinations of K and Δ in 19.53 minutes—a substantial improvement. Nevertheless, to ensure consistency across the Euler and QE-M methods, which would require a complete restructuring of the class setup, this optimization was not implemented.

For a hedge fund pricing 10.000 options to hedge risk, this runtime translates to approximately 61.7 hours for a computer to complete calculations for 54 option combinations (Δ-values $\times$ K-values $\times$ Cases $= 6 \cdot 3 \cdot 3 = 54$). This is clearly impractical for day-trading purposes. Furthermore, the computation times for the QE-M scheme were approximately 2–4 times those of the Euler scheme. However, trade-offs in terms of bias and time make the QE-M scheme extremely competitive when convergence rates are compared.

Another limitation of Monte Carlo simulations is their dependence on the random seed or quasi-random number generator (RNG). While this seed dependency diminishes with an increasing number of simulations, it consequently increases computational time.

Finally, our implementation utilized the log-Euler method, which, in a manner consistent with the Euler scheme, estimates integrals of the exact solution. Applying the Euler scheme directly to dS_t introduces a discretization error in the stock price process. From Table **1**, we observed that the least biased approach was to apply the negative variance fix. Altering these two *fixes* would result in inferior performances for the Euler scheme.

5 Conclusion

We derived risk neutral pricing under the Heston model showing that the investor would expect a compensation for the extra risk proportional to the size of the variance. This process is omitted in every published paper where they assume that $\Lambda = 0$. Our implementation not only works for $\Lambda \neq 0$ but we thoroughly explained its derivation. Furthermore, we derived a stable form of the Heston pricing formula via a classic argument leading to the Heston PDE.

The simulation analysis highlights the trade-offs between some computational efficiency but extreme accuracy in the two methods when parameter constraints are demanding but realistic. We considered two methods, the QE-M and log-Euler scheme with a full truncation. The former, uses moment-matching techniques to approximate the variance process depending on the size of the variance. The latter uses a simple left-point rule to approximate the integrals but a Eulerly approximation to the exact stock price solution. The QE-M method is preferable for high-precision applications despite its computational demands being $2 - 4$x those of the Euler scheme. However, simply using a rough time-step for the QE-M did yield significant results for the QE-M that were still 1/8'th the time of the Euler scheme. Conversely, for Cases where speed is prioritized over precision, especially for lower strike prices or simpler derivatives, the Euler method may be adequate if the parameter constraints are not as demanding as those given in our Cases. However, given our analysis, we would never recommend using the Euler method unless it is for discretization of the CIR process to then use for applications such as MLE estimations.

Further researched could be comparison of Poisson estimation methods, Gamma estimations or using Trees. A interesting topic of consideration is comparing Table 1 fixes and discretization methods of the stock price as we did simply use the result of [20] to give the Euler scheme a fair shot against the QE-M.

References

[1] Hansjörg Albrecher et al. "The Little Heston Trap." In: *Wilmott* (Jan. 2006), pp. 83–92.

[2] Leif BG Andersen. "Efficient simulation of the Heston stochastic volatility model." In: *Available at SSRN 946405* (2007).

[3] Abdel Berkaoui, Mireille Bossy, and Awa Diop. "Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence." In: *ESAIM: Probability and Statistics* 12 (2008), pp. 1–11.

[4] Tomas Björk. *Arbitrage theory in continuous time*. 4th ed. Oxford university press, 2020.

[5] Douglas T Breeden. "An intertemporal asset pricing model with stochastic consumption and investment opportunities." In: *Journal of financial Economics* 7.3 (1979), pp. 265–296.

[6] Mark Broadie and Özgür Kaya. "Exact simulation of stochastic volatility and other affine jump diffusion processes." In: *Operations research* 54.2 (2006), pp. 217–231.

[7] John C. Cox, Jonathan E. Ingersoll, and Stephen A. Ross. "A theory of the term structure of interest rates." In: *Econometrica* 53.2 (1985), pp. 385–407.

[8] Griselda Deelstra, Freddy Delbaen, et al. "Convergence of discretized stochastic (interest rate) processes with stochastic drift term." In: *Applied stochastic models and data analysis* 14.1 (1998), pp. 77–84.

[9] Awa Diop. "Sur la discrétisation et le comportement à petit bruit d'EDS unidimensionnelles dont les coefficients sont à dérivées singulières." PhD thesis. Nice, 2003.

[10] Mireille Bossy—Awa Diop. "An efficient discretisation scheme for one dimensional SDEs with a diffusion coefficient function of the form| x| α, α [1/2, 1." In: (2004).

[11] William Feller. *Introduction to Probality Theory and Its Applications*. Wiley Eastern, 1966.

[12] William Feller. "Two singular diffusion problems." In: *Annals of mathematics* 54.1 (1951), pp. 173–182.

[13] Jim Gatheral. *The volatility surface: a practitioner's guide*. John Wiley & Sons, 2011.

[14] Yevhen Havrylenko. *Continuous-Time Finance 2: Stochastic Volatility Models and Fourier Methods in Option Pricing*. Lecture notes. Mar. 2024. URL: `https://github.com/YoussefRaad-mathecon/Seminar-Asset-Pricing-and-Financial-Markets/blob/main/CTF_Lectures_on_Stochastic_Volatility_Models_and_Carr_Madan_2024_04_04%20(3).pdf`.

[15] Steven L Heston. "A closed-form solution for options with stochastic volatility with applications to bond and currency options." In: *The review of financial studies* 6.2 (1993), pp. 327–343.

[16] Desmond J Higham and Xuerong Mao. "Convergence of Monte Carlo simulations involving the mean-reverting square root process." In: *Journal of Computational Finance* 8.3 (2005), pp. 35–61.

[17] Jean Jacod and Philip Protter. *Probability essentials*. 2nd. Springer Science & Business Media, 2002.

[18] Norman L Johnson, Samuel Kotz, and Narayanaswamy Balakrishnan. *Continuous univariate distributions, volume 2*. Vol. 289. John wiley & sons, 1995.

[19] M.G. Kendall and A. Stuart. *The Advanced Theory of Statistics*. Vol. 1. New York: Macmillan, 1977, p. 168.

[20] Roger Lord, Remmert Koekkoek, and Dick Van Dijk. "A comparison of biased simulation schemes for stochastic volatility models." In: *Quantitative Finance* 10.2 (2010), pp. 177–194.

[21] Benoît B. Mandelbrot. "The Variation of Certain Speculative Prices." In: *The Journal of Business* 36.4 (1963), pp. 394–419.

[22] PB Patnaik. "The non-central χ 2-and F-distribution and their applications." In: *Biometrika* 36.1/2 (1949), pp. 202–232.

[23] Alexander Van Haastrecht and Antoon Pelsser. "Efficient, almost exact simulation of the Heston stochastic volatility model." In: *International Journal of Theoretical and Applied Finance* 13.01 (2010), pp. 1–43.

A Appendix: Code

GitHub repository with code.

B Appendix: Supplementary Material - Proofs

S.B.1: Deriving the PDE's of the Characteristic Functions. Click.

S.B.2: Proof Lemma 2.1. Consider $Y = (b+Z)^2$. Since $Z \sim \mathcal{N}(0,1)$, it follows that $b+Z \sim \mathcal{N}(b,1)$. By the definition of the non-central chi-square distribution, Y is distributed as a non-central chi-square r.v. with 1 DOF and NCP b^2:

$$Y \sim \chi_1'^2(b^2).$$

Now, since $X = a(b+Z)$, we have:

$$X^2 = a^2(b+Z)^2 = a^2 Y.$$

Therefore, multiplying by a^2 scales the distribution, and thus X^2 is distributed as a non-central chi-square r.v. scaled by a^2, i.e.:

$$X^2 \sim a^2 \chi_1'^2(b^2).$$

Hence, $aX \sim \chi_1'^2(b^2)$, which completes the proof.

S.B.3: Proof Corollary 2.1. Let Y be a non-central chi-squared distributed r.v. with d DOF and NCF λ, $\chi_d'^2(\lambda)$. Then $\mathbb{E}[Y] = d+\lambda$ and $\mathbb{V} = 2(d+2\lambda)$.The conditional mean and variance is then simply by Proposition 2.1 given as

$$\begin{aligned}
\mathbb{E}\left[v_T \mid v_t\right] &= \frac{\sigma^2\left(1-e^{-\kappa(T-t)}\right)}{4\kappa}\left(\frac{4\kappa\theta}{\sigma^2} + \frac{4\kappa e^{-\kappa(T-t)}}{\sigma^2\left(1-e^{-\kappa(T-t)}\right)} v_t\right) \\
&= \theta\left(1-e^{-\kappa(T-t)}\right) + v_t e^{-\kappa(T-t)},
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{V}\left(v_T \mid v_t\right) &= \frac{\sigma^4\left(1-e^{-\kappa(T-t)}\right)^2}{8\kappa^2}\left(\frac{4\kappa\theta}{\sigma^2} + \frac{8\kappa e^{-\kappa(T-t)}}{\sigma^2\left(1-e^{-\kappa(T-t)}\right)} v_t\right) \\
&= \frac{\theta\sigma^2\left(1-e^{-\kappa(T-t)}\right)^2}{2\kappa} + \frac{v_t\sigma^2 e^{-\kappa(T-t)}}{\kappa}\left(1-e^{-\kappa(T-t)}\right).
\end{aligned}$$

S.B.4: Proof Proposition 2.2. From Lemma 2.1, (2.7) gives $\hat{v}_{t+\Delta}$ as a scaled non-central chi-square variable with 1 DOF and NCP b^2. Using the known moments for this distribution, we have:

$$\mathbb{E}[\hat{v}_{t+\Delta}] = a(1+b^2), \quad \mathbb{V}[\hat{v}_{t+\Delta}] = 2a^2(1+b^2).$$

Equating these to the exact moments m and s^2, we obtain:

$$a(1+b^2) = m, \quad 2a^2(1+2b^2) = s^2.$$

Set $x = b^2$ and $\psi = \frac{s^2}{m^2}$. Solving for a from the first equation and substituting into the second gives:

$$a = \frac{m}{1+x}, \quad s^2 = 2\left(\frac{m}{1+x}\right)^2 (1+2x).$$

This simplifies to:

$$2(1+2x) = (1+x)^2\psi,$$

leading to the quadratic equation:

$$\psi x^2 + (2\psi - 4)x + (\psi - 2) = 0.$$

Solving this yields:

$$x = \frac{2(\psi-2) \pm \sqrt{16-8\psi}}{2\psi},$$

and thus,

$$x = 2\psi^{-1} - 1 + \sqrt{2\psi^{-1}}\sqrt{2\psi^{-1}-1} = b^2.$$

The solution exists if and only if $\psi \leq 2$, and we take the positive root, as required.

S.B.5: Proof Proposition 2.3. Integrating the density in (2.9), we find

$$\mathbb{E}[\hat{v}_{t+\Delta}] = \int_0^\infty x\left(p\delta(0) + \beta(1-p)e^{-\beta x}\right)dx = \frac{1-p}{\beta}.$$

For $\mathbb{E}[\hat{v}^2_{t+\Delta}]$:

$$\mathbb{E}[\hat{v}^2_{t+\Delta}] = \int_0^\infty x^2\left(p\delta(0) + \beta(1-p)e^{-\beta x}\right)dx = \frac{2(1-p)}{\beta^2}.$$

Thus, the variance is

$$\mathbb{V}[\hat{v}_{t+\Delta}] = \frac{2(1-p)}{\beta^2} - \frac{(1-p)^2}{\beta^2} = \frac{1-p^2}{\beta^2}.$$

Enforcing moment-matching conditions yields:

$$\frac{1-p}{\beta} = m, \quad \frac{1-p^2}{\beta^2} = s^2.$$

Solving for β and p, we have

$$\beta = \frac{1-p}{m}, \quad \psi = \frac{s^2}{m^2} = \frac{1-p^2}{(1-p)^2}.$$

This implies

$$p = \frac{\psi - 1}{\psi + 1}, \quad \beta = \frac{2}{m(\psi + 1)}.$$

For $p \in [0, 1)$, we require $\psi \geq 1$, as desired.

S.B.6: Proof Lemma 2.2. Consider the forward price process at time t, F_t^T, defined by:

$$F_t^T = S_t e^{r(T-t)}.$$

When applying a discretization scheme with increment Δ, the forward price process at time $t + \Delta$ is given by:

$$\begin{aligned} S_{t+\Delta} &= F_{t+\Delta}^T e^{-r(T-(t+\Delta))} \\ &= F_t^T \exp(\ \bullet\ ) e^{-r(T-(t+\Delta))} \\ &= S_t e^{r(T-t)} \exp(\ \bullet\ ) e^{-r(T-(t+\Delta))} \\ &= S_t \exp(r\Delta + \ \bullet\ ). \end{aligned}$$

S.B.7: Proof Proposition 2.4. Consider $X_t = \log S_t$. By section 1.2 we have

$$S_{t+\Delta} = S_t \exp\left\{ \int_t^{t+\Delta} \left(\mu - \frac{v_u}{2} \right) du + \int_t^{t+\Delta} \sqrt{v_u} dW_{1,u}^{\mathbb{P}} \right\}.$$

Using a Cholesky decomposition and considering log-prices instead yields

$$\log(S_{t+\Delta}) = \log(S_t) - \frac{1}{2} \int_t^{t+\Delta} v_u du + \rho \int_t^{t+\Delta} \sqrt{v_u} dW_{2,u}^{\mathbb{P}} + \sqrt{1-\rho^2} \int_t^{t+\Delta} \sqrt{v_u} dW_{1,u}^{\mathbb{P}}. \tag{*}$$

Integrating the square-root variance process of (1.1) yields

$$\begin{gathered} v_{t+\Delta} = v_t + \int_t^{t+\Delta} \kappa(\theta - v_u) du + \sigma \int_t^{t+\Delta} \sqrt{v_u} dW_{2,u}^{\mathbb{P}} \\ \iff \\ \int_t^{t+\Delta} \sqrt{v_u} dW_{2,u}^{\mathbb{P}} = \frac{1}{\sigma} \left(v_{t+\Delta} - v_t - \kappa\theta\Delta + \kappa \int_t^{t+\Delta} v_u du \right). \end{gathered} \tag{**}$$

Substituting $(**)$ into $(*)$ and reducing yields the result.

S.B.8: Proof Proposition 2.5. *case 1.* $\psi \geq \psi_c$: The MGF is computed directly, knowing that in this case QE sets $\hat{v}_{t+\Delta} = a(b + Z_v)^2$ the distribution of which is $\chi_1'(b^2)$ by Lemma 2.1

$$
\begin{aligned}
\Psi_{\hat{v}_{t+\Delta}}(A) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{(a\cdot A)\{z+b\}^2} e^{-z^2/2} \mathrm{d}z \\
&= \frac{1}{\sqrt{2\pi}} \int \exp\left\{-\left(\frac{1}{2} - a\cdot A\right) z^2 + 2b(a\cdot A)z + b^2(a\cdot A)\right\} \mathrm{d}z \\
&= \frac{1}{\sqrt{2\pi}} \int \exp\left\{-\left(\frac{1}{2} - a\cdot A\right) \{z - Q\}^2 + b^2(a\cdot A) + \frac{2b^2(a\cdot A)^2}{1-2(a\cdot A)}\right\} \mathrm{d}z, \qquad Q = \frac{2b(a\cdot A)}{1-2(a\cdot A)} \\
&= \exp\left\{b^2(a\cdot A) + \frac{2b^2(a\cdot A)^2}{1-2(a\cdot A)}\right\} \frac{1}{\sqrt{2\pi}} \int \exp\left\{-\frac{\{z-Q\}^2}{2(1-2(a\cdot A))^{-1}}\right\} \mathrm{d}z \\
&= \exp\left\{b^2(a\cdot A) + \frac{2b^2(a\cdot A)^2}{1-2(a\cdot A)}\right\} (1-2(a\cdot A))^{-1/2} \\
&\quad \times \frac{1}{\sqrt{2\pi}(1-2(a\cdot A))^{-1/2}} \int \exp\left\{-\frac{\{z-Q\}^2}{2(1-2(a\cdot A))^{-1}}\right\} \mathrm{d}z \\
&= \exp\left\{b^2(a\cdot A) + \frac{2b^2(a\cdot A)^2}{1-2(a\cdot A)}\right\} (1-2(a\cdot A))^{-1/2} \times 1 \\
&= \exp\left\{b^2(a\cdot A) + \frac{2b^2(a\cdot A)^2}{1-2(a\cdot A)}\right\} (1-2(a\cdot A))^{-1/2} \\
&= \frac{\exp\left\{\frac{Ab^2a}{1-2Aa}\right\}}{\sqrt{1-2Aa}}
\end{aligned}
$$

where we demand $A < \frac{1}{2a}$ for the expression to be well-defined.

case 2. $\psi < \psi_c$: We simply use integrate the RHS of (2.8), i.e.

$$
\begin{aligned}
\Psi_{\hat{v}_{t+\Delta}}(A) &= p + \beta(1-p) \int_0^{\infty} \exp\{(A-\beta)u\}\, du \\
&= p + \frac{\beta(1-p)}{\beta - A},
\end{aligned}
$$

where we demand $A < \beta$, otherwise the expectation does not exists as the fraction will be zero in the denominator.

C Appendix: Supplementary Material - Plots & Tables

Parameters	**Case I**			**Case II**			**Case III**		
K	Price	Error	Time (ms)	Price	Error	Time (ms)	Price	Error	Time (ms)
60	44.330	$1.03 \cdot 10^{-8}$	19	56.57508	$8.99 \cdot 10^{-9}$	5.0	45.101	$5.75 \cdot 10^{-9}$	12
100	13.085	$5.12 \cdot 10^{-9}$	10	33.597	$3.78 \cdot 10^{-10}$	5.6	16.245	$6.59 \cdot 10^{-12}$	8.1
140	0.29577	$1.43 \cdot 10^{-9}$	8.0	18.157	$5.38 \cdot 10^{-9}$	5.5	4.88854	$5.07 \cdot 10^{-9}$	2.6

Table C.1: *Option prices, computing time in milliseconds and numerical integration errors for three case parameters given in Table 1 with strikes $K \in \{60, 100, 140\}$.*

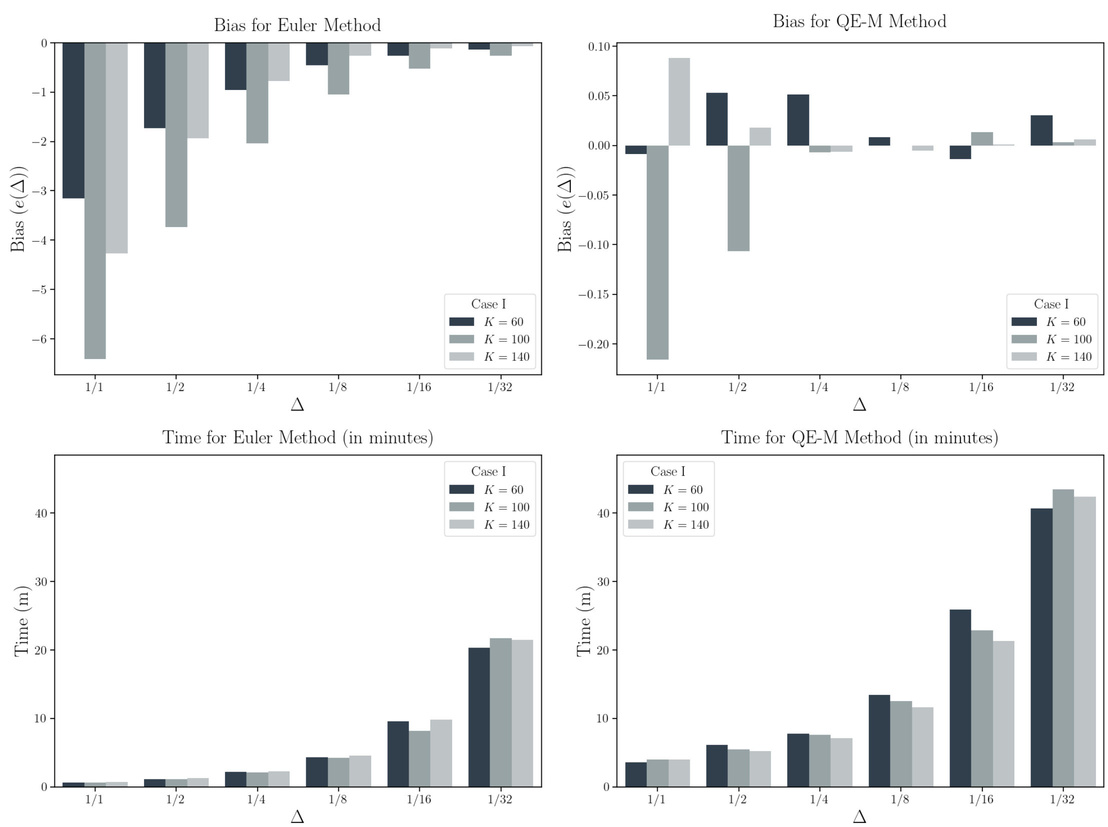

Figure C.1: *Simulation results visualized for the Euler and QE-M Case I.*

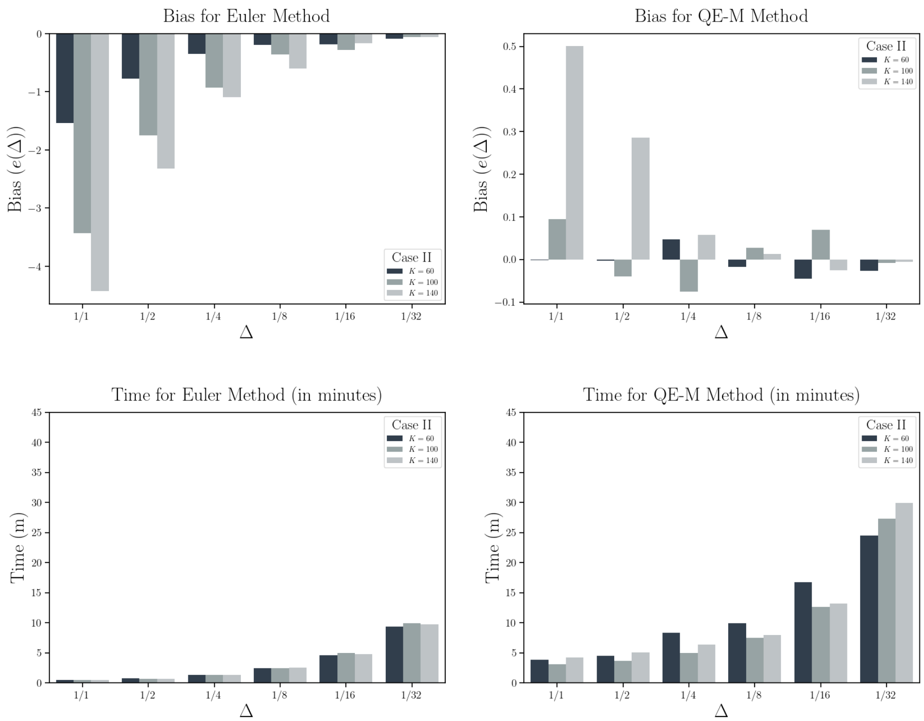

Figure C.2: *Simulation results visualized for the Euler and QE-M for Case II.*

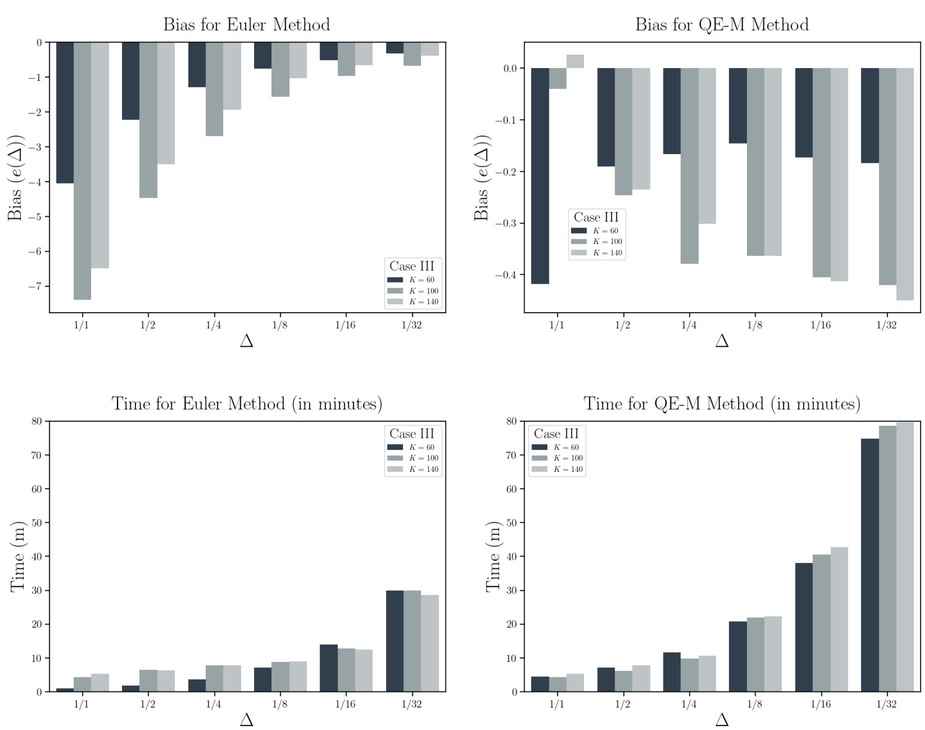

Figure C.3: *Simulation results visualized for the Euler and QE-M.*