The Heston Model

Seminar paper

Jonas Pham

January 8, 2021

Kiel University

Contents

List of Figures

List of Tables

List of Acronyms

ITM	in-the-money
L-BFGS-B	Limited-memory Broyden-Fletcher-Goldfarm-Shanno
MAPE	mean absolute percentage error
OTM	out-of-the-money
RMSE	root-mean-squared error
SDE	stochastic differential equation

1 Introduction

Derivatives are highly liquid instruments in today's financial market. A derivative is a financial instrument whose value depends on the value of an underlying asset. Among them, European options are often used for hedging, speculation or arbitrage purposes. A European call/put option gives the holder the right to buy/sell a certain asset to an agreed on price (strike) at a certain point in time. Thus, the question arises how to price an option whose value depends on a stochastic underlying asset. Black and Scholes (1973) developed a closed-form solution for the price of European options under a risk-neutral setting. One drawback is the assumption of constant volatility which does not coincide with observations from real financial data. Computing implied volatility using the Black-Scholes formula shows variation for different strike prices. In particular, plotting the implied volatility against strike prices reveals a volatility smile structure. For stock options, the resulting volatility smile might also be skewed indicating asymmetry for increases and decreases of the stock price. Moreover, empirical stock return distributions tend to be leptokurtic and the squared or absolute returns exhibit high autocorrelation implying volatility clustering. In the Black-Scholes model however the logarithmic stock returns modelled as a generalized Wiener process follow a normal distribution. Scott (1987), Hull and White (1987) and Wiggins (1987) tackled the problem of constant volatility by modelling it as a stochastic process. A downside of these solutions is that they provide no closed-form expression and are numerically intensive in order to solve two-dimensional partial differential equations. Heston (1993) developed a stochastic volatility model that provides a closed-form solution for the price of a European option when the underlying asset is allowed to be correlated with its volatility.

In this paper we implement Heston's stochastic volatility model using real option price data and evaluate the model fit with respect to the flaws in the Black-Scholes model. Section 2 introduces Heston's stochastic volatility model and the theoretical construct of the semi-analytical solution. Furthermore, methods to calculate European option prices such as numerical integration and Monte Carlo simulation are being presented. In section 3, we show calibration procedures in order to obtain parameter estimates and note calibration difficulties that need to be taken into consideration. Section 4 discusses the empirical results and demonstrates how certain parameters have an influence on the properties of the Heston model.

The aim of this seminar paper is to give an introduction into the Heston model but also to highlight some obstacles in terms of model fit when it comes to practical implementation.

2 Theoretical framework

2.1 Stochastic Volatility Model

The stochastic process for the underlying asset, S_t, is similar to that in the Black-Scholes framework. Heston's stochastic volatility model attempts to solve the problem of non-varying volatility by modelling the variance as an additional stochastic process. Applying Girsanov's theorem onto each stochastic differential equation (SDE) results in the following processes under a risk-neutral measure $\mathbb{Q}$:

$$dS_t = rS_t dt + \sqrt{v_t} S_t dW_t^{(1)} \tag{1}$$

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t} dW_t^{(2)} \tag{2}$$

$$\rho dt = dW_t^{(1)} dW_t^{(2)} \tag{3}$$

r	risk-free interest rate
κ	mean reversion speed of the variance
θ	long run variance
σ	volatility of the variance
v_0	initial level of the variance
ρ	correlation between Brownian motion $W_t^{(1)}$ and $W_t^{(2)}$.

Equation 2 resembles a Cox-Ingersoll-Ross (CIR) process where the mean reversion rate κ ensures convergence towards the long run variance θ (Cox et al., 1985). Additionally, it allows for correlation between the underlying asset and its volatility. The CIR process is strictly positive if the Feller condition $2\kappa\theta > \sigma^2$ is fulfilled (Clark, 2011, pp. 98-103).
In option pricing, the fair price of a European option can be viewed as a function of the underlying asset and its volatility at time t

$$c_t = c(t, S_t, v_t).$$

The value of a European call option whose payoff is a function of the underlying asset at maturity T is of the form

$$c_0 = e^{-rT}\mathbb{E}_{\mathbb{Q}}[(S_T - K)^+] \tag{4}$$

$$= S_t P_1 - K e^{-rT} P_2 \tag{5}$$

where P_1 and P_2 are the probabilities of exercising under different probability measures, i.e. $Pr(S_T > K)$, conditioned on the underlying asset S_t and the volatility v_t. For further derivations it is more convenient to take the logarithm of the underlying

asset $x = ln(S)$. Applying Itô's lemma and arbitrage arguments leads to following partial differential equation[1]:

$$\frac{\partial c}{\partial t} - rc + (r - \frac{1}{2}v_t)\frac{\partial c}{\partial x} + \kappa(\theta - v_t)\frac{\partial c}{\partial v} + \frac{1}{2}v_t\frac{\partial^2 c}{\partial x^2} + \frac{1}{2}\sigma^2 v_t\frac{\partial^2 c}{\partial v^2} + \rho\sigma v_t\frac{\partial^2 c}{\partial x \partial v} = 0. \tag{6}$$

Heston attempts to find the probabilities P_1 and P_2 that satisfy equation 6 with an inverse Fourier transformation. The probabilities can be obtained using the Gil-Pelaez inversion theorem

$$P_j = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re\left[\frac{e^{-i\phi ln(K)}f_j(x, v, T; \phi)}{i\phi}\right] d\phi$$

with $j = 1, 2$ and i representing an imaginary unit. In some cases numerical integration turns out to be a challenging task. In particular, for short maturities the integrand might not be well behaved and display discontinuities or heavy oscillations. For the characteristic functions f_1 and f_2, following Heston's ansatz, Albrecher et al. (2007) made slight modifications to address these problems (Rouah, 2013, p. 31)

$$f_j(x, v_t; \phi) = exp(C_j(\tau; \phi) + D_j(\tau; \phi)v_t + i\phi x_t) \tag{7}$$

$$C_j(\tau; \phi) = ri\phi\tau + \frac{\kappa\theta}{\sigma^2}\left[b_j - \rho\sigma i\phi - d_j)\tau - 2\ ln\left(\frac{1 - c_j e^{-d_j\tau}}{1 - c_j}\right)\right] \tag{8}$$

$$D_j = \frac{b_j - \rho\sigma i\phi - d_j}{\sigma^2}\left(\frac{1 - e^{-d_j\tau}}{1 - c_j e^{-d_j\tau}}\right) \tag{9}$$

where

$$b_1 = \kappa - \rho\sigma; \quad b_2 = \kappa; \quad u_1 = \frac{1}{2}; \quad u_2 = -\frac{1}{2}$$
$$c_j = \frac{b_j - \rho\sigma i\phi - d_j}{b_j - \rho\sigma i\phi + d_j}\ ; \quad d_j = \sqrt{(\rho\sigma i\phi - b_j)^2 - \sigma^2(2u_j i\phi - \phi^2)}.$$

2.2 Option pricing

There exists a wide variety of literature when it comes to option pricing under the Heston model. Among them are numerical integration schemes that make use of inverse Fourier transformations of the characteristic functions. Furthermore, the PDE [6] can be used to apply finite differences methods. Another popular method is the Monte Carlo approximation with simulated stochastic processes. In the following, a method for numerical integration and Monte Carlo simulation will be presented for

[1] A detailed derivation of the PDEs and characteristic function can be found in the appendix

pricing of European options.

2.2.1 Numerical Integration

In Chapter 2.1 a semi-analytic solution for the pricing of European call options [5] has been shown. The probabilities of exercising the option P_1 and P_2 can be expressed as integrals using an inverse Fourier transformation. Given that the characteristic function is known, the option price can be calculated by evaluating the integrals. However, an analytic solution for the integrals is not known, since the anti-derivative of the integrand cannot be found. Therefore, the integrals must be approximated numerically. Gaussian quadratures approximate the integral over the range [a,b] by the weighted sum of the integrand evaluated at the abscissas. The Gauss-Laguerre quadrature is an extension of the Gaussian quadrature method and well suited for the Heston model, since it evaluates the integrand over the domain $(0,\infty)$

$$\int_0^\infty e^{-x} f(x) dx \approx \sum_{j=1}^{n} w_j f(x_j) \tag{10}$$

where the abscissas x_i are the roots of the Laguerre polynomial $L_n(x)$ of order n and the weights w_j are given by

$$w_j = \frac{e^{x_j}}{x_j L_n'(x_j)^2}$$
$$L_n'(x_j) = \sum_{k=1}^{n} \frac{(-1)^k}{(k-1)!} \binom{n}{k} x_j^{k-1} \quad j = 1, ..., n.$$

The advantage of Quadrature rules compared to Riemann summation methods or Newton-Cotes formulas is that the abscissas do not need to be equally spaced. Especially for not so smooth functions, Quadrature rules require less computation time due to fewer and fixed amount of abscissas. The implementation and computation is straightforward. We first calculate abscissas and weights. With the set of weights and abscissas that depend on n but not on $f(x)$, the integral for the probabilities P_1 and P_2 can be calculated by taking the sum of each weight multiplied by the integrand evaluated at each abscissa.

As mentioned before, the integrand in stochastic volatility models can cause problems in terms of numerical stability. It might display high oscillation or discontinuities and dampen very slowly in the ordinate. A remedy by Albrecher et al. (2007) for at least some of the problems was shown already. Bakshi and Madan (2000), Lewis (2001), Gatheral (2006) and Carr and Madan (1999) tackled these problems using different Fourier transform approaches. One could also use quadratures that takes care of oscillations naturally. An approach for that was presented by Le Floc'h (2018).

2.2.2 Monte Carlo Simulation

Numerical integration makes sense when a somewhat analytical solution is available. This was the case for plain vanilla options. American and path-dependent payoff structures such as Exotic options pose a challenge for numerical integration due to their complexity. A popular method for pricing of these options are Monte Carlo simulations.

Recall that the value of a European call option can be expressed as the discounted expected payoff under a risk-neutral measure $\mathbb{Q}$ [4]. In the Monte Carlo simulation, we approximate the expectation using the empirical mean as an unbiased estimator. In order to generate prices of the underlying asset, we simulate the artificial evolution of the underlying asset by discretizing the diffusion-type processes [1] and [2]. A simple discretization scheme is the Euler scheme. Given that the time steps are small, we can approximate a diffusion-type process in integral form as denoted by Kallsen (2018, p.63)

$$X(t_i) \approx X(t_{i-1}) + \int_{t_{i-1}}^{t_i} a(X(t_{i-1}), t_{i-1})dt + \int_{t_{i-1}}^{t_i} b(X(t_{i-1}), t_{i-1})dW(t). \tag{11}$$

For the stochastic process of the underlying asset and the volatility in the Heston model, applying Euler discretization we obtain

$$\begin{aligned} S_{t+dt} &= S_t + r\int_t^{t+dt} S_u du + \int_t^{t+dt} \sqrt{v}_u S_u dW_u^{(1)} \\ &= S_t + rS_t dt + \sqrt{v}_t S_t \sqrt{dt} Z_S \\ v_{t+dt} &= v_t + \int_t^{t+dt} \kappa(\theta - v_u)du + \int_t^{t+dt} \sigma\sqrt{v}_u dW_u^{(2)} \\ &= v_t + \kappa(\theta - v_u)dt + \sigma\sqrt{v}_u\sqrt{dt}Z_V \end{aligned}$$

where Z is a standard normal random variable. Simulation of stochastic processes and Monte Carlo simulations have two major drawbacks. The first drawback is that simulation of the variance process might result in negative values. If, as mentioned earlier, the Feller condition is fulfilled, then the variance process will be strictly positive. However, fulfilment of the Feller condition is not necessarily the case in practice (Clark, 2011, pp. 98-100). One way to obtain only positive values for the variance is to discretize the volatility process rather than the variance process. Applying Itô's lemma onto the volatility $\omega_t = \sqrt{v}_t$ yields following Euler discretization

$$\begin{aligned} \omega_{t+dt} &= \omega_t + \int_t^{t+dt} \frac{1}{2}\kappa\left[\left(\theta - \frac{\sigma^2}{4\kappa}\right)\omega_t^{-1} - \omega_t\right]du + \frac{1}{2}\int_t^{t+dt}\sigma dW_u^{(2)} \\ &= \omega_t + \frac{1}{2}\kappa\left[\left(\theta - \frac{\sigma^2}{4\kappa}\right)\omega_t^{-1} - \omega_t\right]dt + \frac{1}{2}\sigma\sqrt{dt}Z_V. \end{aligned}$$

After simulating the volatility process, its values will be squared in order to obtain the initial variance process. While the use of the volatility process deals with the problem of negative values for the variance and coincides with the original Heston model, the long run variance $\theta_v = (\theta - \frac{\sigma^2}{4\kappa})\omega_t^{-1}$ is now stochastic over time due to the term ω_t^{-1}. For possible negative values ω_t during the simulation, θ_v will jump between being positive or negative. Because of the strong oscillation of θ_v, the resulting simulation will not be accurate anymore. Zhu (2011) has proposed a moment-matching solution for such a transformed volatility scheme. Solving for θ_v using the moments of the volatility process[2] and rewriting the volatility process as

$$d\omega_t = \kappa_v(\theta_v^* - \omega_t)dt + \sigma_v dW_t^{(2)} \tag{12}$$

with halved mean reversion speed, volatility of the variance and moment-matching of θ_v^* improves the accuracy of the simulation significantly

$$\begin{aligned}
\kappa_v = \frac{1}{2}\kappa, \quad \sigma_v = \frac{1}{2}\sigma, \quad \theta_v^* &= \frac{\beta - \omega_t e^{-\frac{1}{2}\kappa dt}}{1 - e^{-\frac{1}{2}\kappa dt}} \\
\beta &= \sqrt{[E(v_{t+dt}) - Var(\omega_{t+dt})]^+} \\
&= (\theta + [v_t - \theta]e^{-\kappa dt} - \frac{\sigma^2}{4\kappa}(1 - e^{-\kappa dt}))^{\frac{1}{2}}.
\end{aligned}$$

The second drawback of simulation of stochastic processes is its slow convergence

$$\mathbb{E}[f(S_T)] - \frac{1}{N}\sum_{i=1}^{N} f(s_T^{M,i}) = \underbrace{\mathbb{E}[f(S_T)] - \mathbb{E}[f(S_T^M)]}_{\text{discretization error}} + \underbrace{\mathbb{E}[f(S_T^M)] - \frac{1}{N}\sum_{i=1}^{N} f(s_T^{M,i})}_{\text{Monte Carlo error}}$$

where S_T is the true underlying asset, S_T^M is the discretized underlying asset with M time steps and $s_T^{M,i}$ the N simulation paths of the underlying asset for $i = 1, ..., N$. While the empirical mean is an unbiased and consistent estimator, the Monte Carlo error crucially depends on the variance of $f(S_T^M)$ and the number of simulations. The rate of convergence for Monte Carlo simulations is $\mathcal{O}(N^{-0.5})$. In order to halve the error, four times as many simulations are needed on average. Computational cost can however be improved by applying variance reduction methods such as antithetic variables or control variates. Considering the discretization error, the Euler method has weak convergences of order 1 and thus the error of the simulation is proportional to the time step $\mathcal{O}(M)$. In general, the order of convergence can be increased by applying more suffisticated discretization schemes such as the Milstein method which expands the coefficients a and b from [11] using Itô's lemma since they are also functions of X_t. However, according to Zhu there is no Milstein scheme for the volatility process ω_t since the diffusion term is not a function of ω_t.

[2] A detailed derivation is given in the appendix

3 Calibration

In order to calculate European option prices under the Heston model, the parameters $\Theta = (\kappa, \theta, \sigma, v_0, \rho)$ need to be determined. A commonly used method is to calibrate the model onto real option prices. Thus, parameters are chosen such that the model and observed option prices coincide. Since the number of parameters exceed the number of equations, a unique solution cannot be found. Hence, in practice the error between market and model values will be minimized via a loss function defined as

$$\min_{G(\Theta)} \{G(\Theta) = \frac{1}{N} \sum_{t,k} w_{tk}(X_{tk} - X_{tk}^{\Theta})^2\}$$

with a set of N_T maturities τ_t $(t = 1, ..., N_T)$, a set of N_K strike prices K_k $(k = 1, ..., N_K)$, market values $X(\tau_t, K_k) = X_{tk}$ and the corresponding model values $X(\tau_t, K_k; \Theta) = X_{tk}^{\Theta}$ calculated under the Heston model. Using identity weights w_{tk}, this loss function amounts to the MSE, the average squared distance between actual values and estimated values. The objective function $G(\Theta)$ is not necessarily convex and oscillations might make it difficult for minimization algorithms to find the global minimum. Thus, depending on the method of optimization, the accuracy might suffer. Furthermore, the type of loss function is crucial for a good fit of model prices across the moneyness spectrum and therefore the choice of loss function is important (Christoffersen and Jacobs, 2003).

3.1 Optimization

The target of minimization is a non-linear multivariate function with bounded constraints. The long run variance θ, volatility of the variance σ and the initial level of the variance v_0 should be non-negative. In order to have the mean reverting property, κ should also take a positive value. The correlation between Brownian motions ρ will be in the interval [-1,1].
For local minimization, we use the Limited-memory Broyden-Fletcher-Goldfarm-Shanno (L-BFGS-B) algorithm with box constraints from the Python package `scipy`. This quasi-Newton-type method utilizes an approximation of the Hessian in order to minimize the gradient of the objective function using a limited amount of computer memory. The advantage of a local optimization algorithm is its relatively fast computation time. However, in order to not get trapped in a local minimum, the choice of the starting values is of importance. In general, one could simply choose different initial values and check which set of parameters yield the best results.

Global optimization algorithms like the differential evolution algorithm by Storn and Price (1997) avoid the subject of initial values altogether. Again, the Python

package `scipy` is used with the differential evolution algorithm method `best1bin`. Essentially, this method first creates a random population within the specified bounds consisting of N_P parameter vectors. For each of the so called members of the population, two other different members are chosen randomly. The weighted difference of those members is added to the original member in order to create a trial candidate. Parameters from the original member and of the trial candidate are then mixed together randomly based on a crossover ratio to create a final candidate. The objective function is then evaluated with the final candidate and the original member. If the objective function output of the final candidate is smaller than the original member, it replaces the original member. This is done for each member in the population over N_G generations (iterations). In the last generation, the member of the population with the smallest objective function is chosen. From the short description, it is clear that the computation cost will be significantly more expensive than with a local optimization approach. Nonetheless, in this paper the differential evolution algorithm will be used as comparison to see whether the local optimization result and starting values are appropriate.

3.2 Loss functions

Looking at loss functions, one could minimize the squared difference between observed option prices and option prices under the Heston model. The disadvantage of using a symmetric loss function like the MSE is that large option values will have a bigger impact on the calibration overall. In this case, using the MSE or MAE as the loss function, the model fit will be better suited for long maturity in-the-money (ITM) options[3] rather than deep out-of-the-money (OTM) or short maturity options. Using the relative mean squared error will lead to the opposite result. Dividing by the option value will favour low values of options much more. A simple approach suggested by Rouah (2013) is to use ITM options only, that is, using call options for strikes less than the spot price and put options for strikes greater than the spot price. Christoffersen et al. (2009) used the reciprocal of the squared Black-Scholes Greek Vega as weights in order to give short maturity deep ITM and OTM options more relevance during the calibration process.

Another way of loss function optimization is to minimize the squared error between quoted and model implied volatilities. The range of values for implied volatilities is usually significantly smaller than option values which somewhat lessens the problem of options not contributing to the calibration. Since model fit is often evaluated by comparing implied volatilities, it also makes sense to minimize the error directly. This is however accompanied by higher computation time due to the extraction of the implied volatility after every calculation of a model option price.

[3] If an option is "in the money", $S_t > K$. Vice versa, an option is "out of the money" if $S_t < K$.

4 Empirical Analysis

For the calibration of the model parameters, we choose the two loss function types mentioned in Chapter 3.2, namely the MSE of option prices and implied volatilities. Furthermore, we will use different sets of weights to attenuate the problem of overfitting ITM options. We make use of different forms of bid-ask-spread reciprocals $\frac{1}{|Spread|}, \frac{1}{Spread^2}, \frac{1}{\sqrt{Spread}}$ denoted as $w^{(1)}, w^{(2)}$ and $w^{(3)}$ respectively. Smaller spreads will thus have a larger impact during the calibration. Following the same logic, the weights $w^{(4)}$ are the reciprocals of the squared Black-Scholes Greek Vega. These four weights are compared against the baseline weight $w^{(0)}$ of 1, i.e. having no weights, to study their influence on the model fit.

The data set consists of a DAX call option chain from November 3rd, 2020 with six maturities and 58 different strike prices for a total of 348 observations.[4] Included are options with short and long maturities as well as deep ITM and deep OTM options in order to cover a realistic range. Since we are using the DAX Performance-Index, dividends are assumed to be 0. Moreover, we assume a risk-free interest rate of 0% to reflect the current situation of negative interest rates for German treasury bonds and the EURIBOR. As an initial guess for the local optimization $\kappa = 3, \quad \theta, \sigma, v_0 = 0.5, \quad \rho = -0.5$ were chosen. A significant part of the code has its origin from Rouah (2013) who implemented the option pricing in Matlab and C#. For this paper, the code has been translated into Python and optimized for parallel computing.

The following Table 1 sums up the calibration results using local and global optimization:

Method	**Algorithm**	w	**MSE IV**	κ	θ	σ	v_0	ρ
(1) MSE Price	L-BFGS-B	$w^{(0)}$	1.331e-4	2.34303807	0.07110119	1.27077071	0.10531966	-0.71034407
(2) MSE Price	L-BFGS-B	$w^{(1)}$	1.727e-4	2.21892363	0.07269739	1.21983753	0.10190313	-0.71923353
(3) MSE Price	L-BFGS-B	$w^{(2)}$	2.152e-4	2.09977	0.07444378	1.17753398	0.09916212	-0.72522431
(4) MSE Price	L-BFGS-B	$w^{(3)}$	1.510e-4	2.28388222	0.07188725	1.24595471	0.10358197	-0.71491333
(5) MSE Price	L-BFGS-B	$w^{(4)}$	5.358e-5	2.93347265	0.07439612	1.54946749	0.11104902	-0.69409251
(6) MSE IV	L-BFGS-B	$w^{(0)}$	**5.234e-5**	2.91273117	0.07599915	1.58816529	0.1112458	-0.69618637
(1) MSE Price	DE	$w^{(0)}$	1.332e-4	2.34292578	0.07110142	1.27074574	0.10531864	-0.71034224
(2) MSE Price	DE	$w^{(1)}$	1.723e-4	2.221368	0.07268748	1.22035493	0.10192573	-0.71925155
(3) MSE Price	DE	$w^{(2)}$	2.160e-4	2.09531163	0.07446999	1.17668923	0.09912432	-0.72520626
(4) MSE Price	DE	$w^{(3)}$	1.510e-4	2.28395065	0.07188713	1.24597419	0.10358271	-0.71491259
(5) MSE Price	DE	$w^{(4)}$	5.065e-5	3.25780862	0.07332633	1.58341874	0.11266068	-0.69737511
(6) MSE IV	DE	$w^{(0)}$	**4.833e-5**	3.52661087	0.07307091	1.63110968	0.11475628	-0.70869058

Table 1: Parameter calibration using different loss functions and weights. $w^{(0)}$ is no weights, $w^{(1)}$ are the absolute spread reciprocals, $w^{(2)}$ are the squared spread reciprocals, $w^{(3)}$ are the square root spread reciprocals, $w^{(4)}$ are the Black-Scholes Vega reciprocals

[4] Due to the length of the data set and subsequent output of the analysis, comprehensive data can be found on `github.com/jnpm/heston`.

The obtained parameters from the calibration are similar for both local and global optimization methods. Furthermore, correlation ρ between the underlying asset and its volatility is strongly negative. Black (1976) noticed this inverse relationship between stock returns and its volatility. An intuitive explanation for this so-called "leverage effect" is that investors might react more erratic to negative stock returns leading to a higher volatility. Using the calibrated parameters, we then calculate model option prices via numerical integration. Given the model option prices, we can deduce the implied volatility with a simple root-finding method such as the Bisection algorithm. Unsurprisingly, minimizing the MSE of the implied volatilities yields the smallest MSE across all methods for local and global optimization.
In order to further evaluate each loss function's performance, we will use the mean absolute percentage error (MAPE) and the root-mean-squared error (RMSE) as metrics to compare the calculated option prices:

$$MAPE: \quad \frac{1}{N}\sum_{tk}\frac{|C_{tk}-C_{tk}^{\Theta}|}{C_{tk}^{\Theta}}$$

$$RMSE: \quad \sqrt{\frac{\sum_{tk}(C_{tk}-C_{tk}^{\Theta})^2}{N}}.$$

Option prices under the Heston model are calculated with the 32-point Gauss-Laguerre Quadrature and Monte Carlo simulation. For the discretization scheme, we use the transformed volatility scheme by Zhu (2011) with step size 10.000 and 500.000 simulations. Table 2 displays the error metrics for the model option prices as well as the MSE of the implied volatilities for both calculation methods.

Method			**Numerical Integration**			**Monte Carlo Simulation**		
			MSE IV	MAPE Price	RMSE Price	MSE IV	MAPE Price	RMSE Price
(1) MSE Price	L-BFGS-B	$w^{(0)}$	1.331e-4	1.866%	**9.196**	1.453e-4	2.320%	15.460
(2) MSE Price	L-BFGS-B	$w^{(1)}$	1.727e-4	1.536%	9.552	9.529e-5	1.927%	15.967
(3) MSE Price	L-BFGS-B	$w^{(2)}$	2.152e-4	**1.404%**	10.440	3.275e-4	1.682%	15.826
(4) MSE Price	L-BFGS-B	$w^{(3)}$	1.510e-4	1.670%	9.288	3.003e-4	3.064%	14.513
(5) MSE Price	L-BFGS-B	$w^{(4)}$	5.358e-5	2.026%	10.728	1.379e-4	2.685%	25.124
(6) MSE IV	L-BFGS-B	$w^{(0)}$	**5.234e-5**	1.941%	11.435	7.970e-5	3.215%	29.425
(1) MSE Price	DE	$w^{(0)}$	1.332e-4	1.866%	**9.196**	1.749e-4	2.387%	16.345
(2) MSE Price	DE	$w^{(1)}$	1.723e-4	1.536%	9.550	1.033e-4	1.901%	15.760
(3) MSE Price	DE	$w^{(2)}$	2.160e-4	**1.404%**	10.446	2.101e-4	1.702%	15.701
(4) MSE Price	DE	$w^{(3)}$	1.510e-4	1.670%	9.288	2.659e-4	2.065%	15.667
(5) MSE Price	DE	$w^{(4)}$	5.065e-5	1.945%	11.332	7.449e-5	3.271%	28.041
(6) MSE IV	DE	$w^{(0)}$	**4.833e-5**	1.696%	27.766	6.501e-5	2.842%	27.766

Table 2: Error metrics for the implied volatilities and option prices calculated using numerical integration and Monte Carlo simulation.

Across all methods used, numerical integration provides better results compared to option prices calculated via Monte Carlo simulation. Weights $w^{(2)}$ produces the smallest MAPE for option prices whereas the baseline MSE loss function has the smallest RMSE. At first glance it is not clear which loss function to choose. There was no loss function which produced the smallest errors universally during the computation. In general, a loss function should therefore be chosen based on the available data set and desired fit. For our analysis, MAPE is better suited than RMSE due to the strong heterogeneity within the same maturity.[5] The squared spread reciprocals represented by $w^{(2)}$ put more focus on deep OTM options compared to the other weights resulting in the smallest MAPE. We will thus proceed with weights $w^{(2)}$ in our subsequent analysis. For a quick comparison to the Black-Scholes model, we calibrate the parameter σ with the differential evolution algorithm. The calibrated $\sigma = 0.24306576$ has a MAPE of 15.341% and RMSE of 119.521. The Heston model performs significantly better than the Black-Scholes model for the given data set.

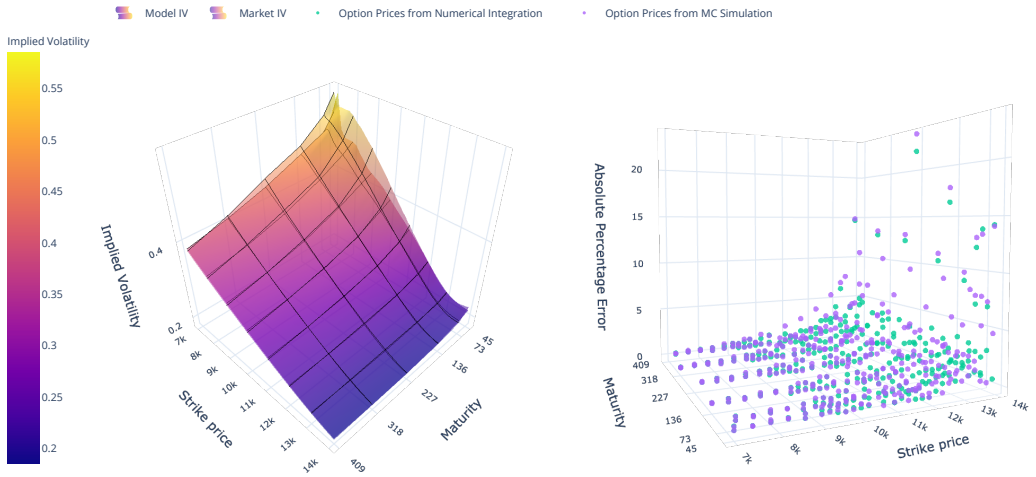

Figure 1: Implied Volatility surface on the left. The upper layer displays implied volatilities from market prices. Absolute percentage error to the market option prices on the right. Interactive graphs of all the methods can be found on `github.com/jnpm/heston`.

The left side of Figure 1 displays the implied volatility surface of the market option prices and the implied volatilities obtained from our model option prices using method 3 and numerical integration. For at-the-money (ATM) and OTM options, the implied volatilities are close. However, for deep ITM options and especially for short maturities, the model fit is considerably different. Since the spread was decreasing for an increasing strike price or maturity, we gave more weighting towards ATM and OTM options by defining the weights as the reciprocals of the spread. Moreover, short maturity deep ITM options are less sensitive to volatility. A slightly

[5] e.g. for the shortest maturity of 45 days, the largest option price in the data set was 4996 whereas the lowest value was 11.

different implied volatility for these options will thus not necessarily mean a significantly different option price. The right side of Figure 1 confirms this as it shows the relative absolute error in percent of the model option prices to the market prices. Option prices have been calculated using numerical integration (green) and Monte Carlo simulation (purple). Here, the model parameters seem to fit ITM options particularly well whereas OTM option fit is worse.

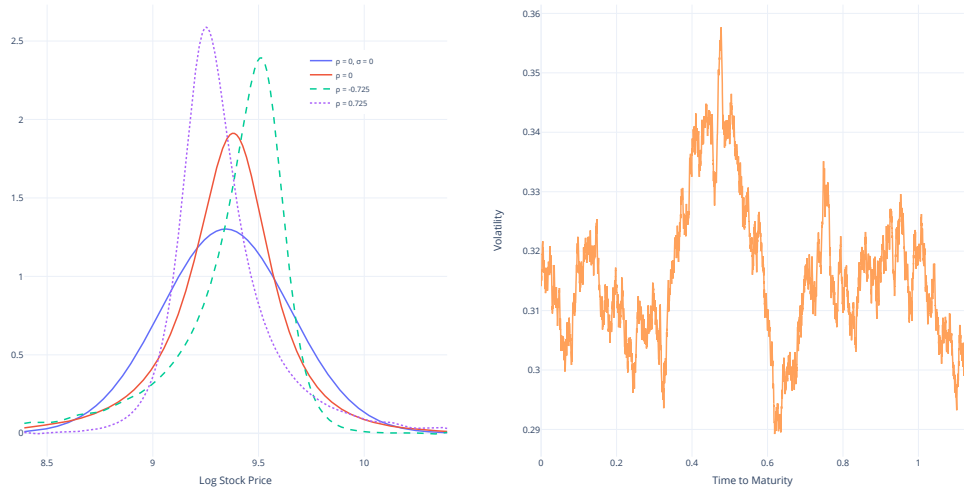

Figure 2: Influence of correlation and volatility of the variance on skewness and kurtosis. Conditional probability density function of the log stock price on the left side and a simulated volatility process over 406 days on the right side

Figure 2 shows how ρ and σ influence the skewness and kurtosis. When σ is 0, the variance process loses its stochastic part and the log stock price has a normal distribution. As σ increases, the probability that the stock price reaches extreme values is more likely due to higher fluctuation and the distribution becomes leptokurtic. The red line shows the conditional probability density function with σ equal to the calibrated parameter with weights $w^{(2)}$. The correlation between volatility and underlying asset has an impact on the skewness of the distribution. The dashed green line uses parameters from the calibration. Due to the negative skewness, more mass is on the left tail. Heston (1993) argues that since lower strikes lie in the left tail, ITM options are priced higher and OTM options lower compared to options from the Black-Scholes formula. This was indeed the case in the empirical analysis. The right graph shows a simulated mean reverting volatility process with noticeable clusters. Consequently, the Heston model captures important properties of financial data, namely heavy-tails in the distribution and volatility clustering. Ultimately, this provides a better fit to the option prices observed on the market compared to Black-Scholes pricing.

5 Conclusion

The Black-Scholes framework is an arbitrage-based pricing model with the assumption of a Geometric Brownian motion for the stock price. However, it does not take into account important properties of financial data such as returns that are characterized by leptokurtosis and volatility clustering. In this seminar paper, we introduce the Heston model which replaces the constant volatility in the Black-Scholes model by a mean-reverting stochastic volatility process. This also allows for correlation between the underlying asset and its volatility.

Calibration of multivariate objective functions can be difficult for non-convex functions. In our analysis, both local and global optimization performed on par, although for calibration of shorter maturities a global optimization might outperform a local search. Additionally, we outlined the importance of loss functions and their impact on calibration. Between the methods tested, global minimization using implied volatilities resulted in the smallest MSE of implied volatilities whereas minimization using option prices as the loss function with reciprocals of the squared spread as weights provided the smallest MAPE. Given the analysis, there exists no loss function which universally provides the best fit in all metrics. In general, one should choose a loss function based on their data set and desired fit. In terms of calculation of option prices, numerical integration yields better results compared to Monte Carlo simulation. The advantage of Monte Carlo simulations, however, lies in its simplicity and wide use cases which makes it viable for more complex options.

Based on our data, we examined how certain parameters influence the properties of the model. An increase in the volatility of the variance σ produces heavier tails while the correlation ρ has an impact on the skewness of the distribution. Overall, this flexibility leads to the Heston model creating far better option price estimates compared to the Black-Scholes model.

Further research can be done with different data sets, ranges of maturities and strikes in order to evaluate the performance of loss functions under different scenarios. Moreover, the Heston model was not able to fit the implied volatility model very well for short maturities. One could improve flexibility by introducing time dependent parameters or increase the number of parameters. Christoffersen et al. (2009) included a second stochastic variance process in their Double Heston model. Moreover, the Heston model does not take jumps into account. However, due to the recent pandemic, the appearance of jumps influencing our data can not be excluded. To provide a better fit for short maturities, stochastic volatility jump diffusion models such as Bates (1996), Scott (1987) and Pan (2002) should therefore be considered.

References

Albrecher, Hansjörg et al. (2007). "The Little Heston Trap". In: *Wilmott*, pp. 83–92.

Bakshi, Gurdip and Dilip Madan (2000). "Spanning and Derivative-Security Valuation". In: *Journal of Financial Economics* 55, pp. 205–238.

Bates, David S. (1996). "Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutsche Mark Options". In: *Review of Financial Studies* 9, pp. 69–107.

Black, Fischer (1976). "Studies of Stock Price Volatility Changes". In: *Proceedings of the 1976 Meeting of the Business and Economic Statistics Section*, pp. 177–181.

Black, Fischer and Myron Scholes (1973). "The Pricing of Options and Corporate Liabilities". In: *Journal of Political Economy* 81.3, pp. 637–654.

Carr, Peter and Dilip Madan (1999). "Option Valuation Using the Fast Fourier Transform". In: *The Journal of Computational Finance* 2, pp. 61–73.

Christoffersen, Peter, Steven Heston, and Kris Jacobs (2009). "The Shape and Term Structure of the Index Option Smirk: Why Multifactor Stochastic Volatility Models Work So Well". In: *Management Science* 55, pp. 1914–1932.

Christoffersen, Peter F. and Kris Jacobs (2003). "The Importance of the Loss Function in Option Valuation". In: *SSRN Electronic Journal*.

Clark, Iain J. (2011). *Foreign Exchange Option Pricing: A Practitioner's Guide*. Wiley Finance Series. Chichester, West Sussex, U.K: Wiley, pp. 98–103. 280 pp.

Cox, John C., Jonathan E. Ingersoll, and Stephen A. Ross (1985). "A Theory of the Term Structure of Interest Rates". In: *Econometrica* 53.2, p. 385.

Gatheral, Jim (2006). *The Volatility Surface: A Practitioner's Guide*. Wiley Finance Series. Hoboken, N.J: John Wiley & Sons. 179 pp.

Heston, Steven L. (1993). "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options". In: *Review of Financial Studies* 6, pp. 327–343.

Hull, John and Alan White (1987). "The Pricing of Options on Assets with Stochastic Volatilities". In: *The Journal of Finance* 42, pp. 281–300.

Kallsen, Jan (2018). *Lecture notes in Computational Finance*. URL: `https://www.math.uni-kiel.de/numerik/kallsen/personen/kallsen/pub/skripte/compfin.pdf`.

Le Floc'h, Fabien (2018). "An Adaptive Filon Quadrature for Stochastic Volatility Models". In: *Journal of Computational Finance* 22, pp. 65–88.

Lewis, Alan L. (2001). "A Simple Option Formula for General Jump-Diffusion and Other Exponential Levy Processes". In: *SSRN Electronic Journal*.

Pan, Jun (2002). "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time-Series Study". In: *Journal of Financial Economics* 63, pp. 3–50.

Rouah, Fabrice Douglas (2013). *The Heston Model and Its Extensions in Matlab and C#*. Hoboken, NJ, USA: John Wiley & Sons, Inc.

Scott, Louis O. (1987). "Option Pricing When the Variance Changes Randomly: Theory, Estimation, and an Application". In: *The Journal of Financial and Quantitative Analysis* 22.4, p. 419.

Storn, Rainer and Kenneth Price (1997). "Differential Evolution – A Simple and Efficient Heuristic for Global Optimization over Continuous Spaces". In: *Journal of Global Optimization* 11, pp. 341–359.

Wiggins, James B. (1987). "Option Values under Stochastic Volatility: Theory and Empirical Estimates". In: *Journal of Financial Economics* 19, pp. 351–372.

Zhu, Jianwei (2011). "A Simple and Accurate Simulation Approach to the Heston Model". In: *The Journal of Derivatives* 18, pp. 26–36.

A Derivations

A.1 European call option PDE

We start with the risk-neutral SDEs where x = $ln(S_t)$

$$\begin{aligned} dx_t &= (r - \frac{1}{2}v_t)dt + \sqrt{v_t}dW_t^{(1)} \\ dv_t &= \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^{(2)} \\ \rho dt &= dW_t^{(1)}dW_t^{(2)} \end{aligned}$$

Let $C = c(t, x_t, v_t)$ be the function of the European call option with discounted option price $\hat{C} = C \cdot e^{-rt}$. Applying Itô's lemma and simplifying leads to:

$$\begin{aligned} d\hat{C} &= \frac{\partial \hat{c}}{\partial t}dt + \frac{\partial \hat{c}}{\partial x}dx + \frac{\partial \hat{c}}{\partial v}dv + \frac{1}{2}\frac{\partial^2 \hat{c}}{\partial x^2}dx^2 + \frac{1}{2}\frac{\partial^2 \hat{C}}{\partial v^2}dv^2 + \frac{\partial^2 \hat{c}}{\partial x \partial v}dxdv \\ &= \Bigg[\frac{\partial \hat{c}}{\partial t} + (r - \frac{1}{2}v_t)\frac{\partial \hat{c}}{\partial x} + \kappa(\theta - v_t)\frac{\partial \hat{c}}{\partial v} + \frac{1}{2}v_t\frac{\partial^2 \hat{c}}{\partial x^2} + \frac{1}{2}\sigma^2 v_t\frac{\partial^2 \hat{c}}{\partial v^2} \\ &+ \rho\sigma v_t\frac{\partial^2 \hat{c}}{\partial x \partial v}\Bigg]dt + \frac{\partial \hat{c}}{\partial x}\sqrt{v_t}dW_t^{(1)} + \frac{\partial \hat{c}}{\partial v}\sigma\sqrt{v_t}dW_t^{(2)} \end{aligned}$$

The discounted fair option price $\hat{C}$ is a martingale relative to the unique equivalent martingale measure (EMM) $\mathbb{Q}$. Thus, the drift term has to disappear

$$\frac{\partial \hat{c}}{\partial t} + (r - \frac{1}{2}v_t)\frac{\partial \hat{c}}{\partial x} + \kappa(\theta - v_t)\frac{\partial \hat{c}}{\partial v} + \frac{1}{2}v_t\frac{\partial^2 \hat{c}}{\partial x^2} + \frac{1}{2}\sigma^2 v_t\frac{\partial^2 \hat{c}}{\partial v^2} + \rho\sigma v_t\frac{\partial^2 \hat{c}}{\partial x \partial v} = 0$$

or undiscounted

$$\frac{\partial c}{\partial t} - rc + (r - \frac{1}{2}v_t)\frac{\partial c}{\partial x} + \kappa(\theta - v_t)\frac{\partial c}{\partial v} + \frac{1}{2}v_t\frac{\partial^2 c}{\partial x^2} + \frac{1}{2}\sigma^2 v_t\frac{\partial^2 c}{\partial v^2} + \rho\sigma v_t\frac{\partial^2 c}{\partial x \partial v} = 0 \tag{13}$$

which is identical to [6] on page 3.

□

A.2 Characteristic functions

The following derivation has been reproduced from Rouah (2013, pp. 8-14) for the purpose of providing clarity to the reader.
The function of the European call option [5] can be written in terms of $x = ln(S)$ and $\tau = T - t$ as

$$C = e^x P_1 - Ke^{r\tau}P_2$$

This function follows the PDE [13]. We now find the required derivatives in terms of P_1 and P_2.

$$\frac{\partial c}{\partial t} = e^x\left(P_1 + \frac{\partial c}{\partial x}\right), \quad \frac{\partial c}{\partial x} = e^x\left(P_1 + \frac{\partial P_1}{\partial x}\right) - Ke^{-r\tau}\frac{\partial P_2}{\partial x}$$
$$\frac{\partial^2 c}{\partial x^2} = e^x\left(P_1 + 2\frac{\partial P_1}{\partial x} + \frac{\partial^2 P_1}{\partial x^2}\right) - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial x^2}$$
$$\frac{\partial c}{\partial v} = e^x\frac{\partial P_1}{\partial v} - Ke^{-r\tau}\frac{\partial P_2}{\partial v}, \quad \frac{\partial^2 c}{\partial v^2} = e^x\frac{\partial^2 P_1}{\partial v^2} - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial v^2}$$
$$\frac{\partial^2 c}{\partial x \partial v} = e^x\left(\frac{\partial P_1}{\partial v} + \frac{\partial^2 P_1}{\partial x \partial v}\right) - Ke^{-r\tau}\frac{\partial^2 P_2}{\partial x \partial v}$$

Heston argues that the PDE [13] holds for any strike $K \geq 0$, any value of $S \geq 0$ and any $r \geq 0$. Setting $S = 1, K = 0$ produces an option whose price is P_1. Putting the derivatives into [13] and simplifying, yields the PDE in terms of P_1

$$\frac{\partial P_1}{\partial t} + \left(r + \frac{1}{2}v_t\right)\frac{\partial P_1}{\partial x} + \frac{1}{2}v_t\frac{\partial^2 P_1}{\partial x^2} + \rho\sigma v_t\frac{\partial^2 P_1}{\partial x \partial v}$$
$$+(\rho\sigma v_t + \kappa(\theta - v_t))\frac{\partial P_1}{\partial v} + \frac{1}{2}\sigma^2 v_t\frac{\partial^2 P_1}{\partial v^2} = 0 \tag{14}$$

Conversely, setting $S = r = 0$ and $K = 1$ produces the PDE in terms of P_2

$$\frac{\partial P_2}{\partial t} + \left(r - \frac{1}{2}v_t\right)\frac{\partial P_2}{\partial x} + \frac{1}{2}v_t\frac{\partial^2 P_2}{\partial x^2} + \rho\sigma v_t\frac{\partial^2 P_2}{\partial x \partial v}$$
$$+\kappa(\theta - v_t)\frac{\partial P_2}{\partial v} + \frac{1}{2}\sigma^2 v_t\frac{\partial^2 P_2}{\partial v^2} = 0 \tag{15}$$

Combining [14] and [15] into a single equation

$$\frac{\partial P_j}{\partial t} + \rho\sigma v_t\frac{\partial^2 P_j}{\partial v \partial x} + \frac{1}{2}v_t\frac{\partial^2 P_j}{\partial x^2} + \frac{1}{2}\sigma^2 v_t\frac{\partial^2 P_j}{\partial v^2}$$
$$+(r + u_j v_t)\frac{\partial P_j}{\partial x} + (a - b_j v_t)\frac{\partial P_j}{\partial v} = 0 \tag{16}$$

for $j = 1, 2$ and $u_1 = \frac{1}{2}, u_2 = -\frac{1}{2}, a = \kappa\theta, b_1 = \kappa - \rho\sigma, b_2 = \kappa$.
As mentioned on page 3, the probabilities can be obtained using the Gil-Pelaez inversion theorem. The characteristic function by Heston is of the form

$$f_j(x, v_t; \phi) = exp(C_j(\tau; \phi) + D_j(\tau; \phi)v_t + i\phi x_t)$$

From the Feyman-Kac theorem it can be shown that the characteristic functions follow the PDE [16].

$$-\frac{\partial f_j}{\partial \tau} + \rho\sigma v_t \frac{\partial^2 f_j}{\partial v \partial x} + \frac{1}{2} v_t \frac{\partial^2 f_j}{\partial x^2} + \frac{1}{2}\sigma^2 v_t \frac{\partial^2 f_j}{\partial v^2} + (r + u_j v_t)\frac{\partial f_j}{\partial x} + (a - b_j v_t)\frac{\partial f_j}{\partial v} = 0 \tag{17}$$

Taking the derivatives required for [17], substituting into the PDE and dropping the f_j terms, results in

$$v_t \left(-\frac{\partial D_j}{\partial \tau} + \rho\sigma i\phi D_j - \frac{1}{2}\phi^2 + \frac{1}{2}\sigma^2 D_j^2 + u_j i\phi - b_j D_j \right) - \frac{\partial C_j}{\partial \tau} + ri\phi + aD_j = 0$$

which produces the two ordinary differential equations

$$\frac{\partial D_j}{\partial \tau} = \rho\sigma i\phi D_j - \frac{1}{2}\phi^2 + \frac{1}{2}\sigma^2 D_j^2 + u_j i\phi - b_j D_j \tag{18}$$

$$\frac{\partial C_j}{\partial \tau} = ri\phi + aD_j \tag{19}$$

In order to solve the first ODE, we make use of the Riccati equation. The first-order quadratic ODE [18] is of the form of a Riccati equation with initial conditions $D_j(0, \phi) = C_j(0, \phi) = 0$.

In general, a Ricatti equation for $y(t)$ with coefficients $P(t), Q(t)$ and $R(t)$ is defined as

$$\frac{dy(t)}{dt} = P(t) + Q(t)y(t) + R(t)y(t)^2 \tag{20}$$

Considering a corresponding second-order ODE for $w(t)$

$$w'' - \left[\frac{P'}{P} + Q\right] w' + PRw = 0 \quad or \quad w'' + bw' + cw = 0 \tag{21}$$

gives the solution to [20] as $y(t) = -\frac{w'(t)}{w(t)}\frac{1}{R(t)}$. [21] can be solved with the auxiliary equation $r^2 + br + c = 0$ with the two solutions α and β being

$$\alpha = \frac{-b + \sqrt{b^2 - 4c}}{2}, \quad \beta = \frac{-b - \sqrt{b^2 - 4c}}{2}. \tag{22}$$

The solution of [21] is then $w(t) = Me^{\alpha t} + Ne^{\beta t}$ and the solution of the Riccati equation [20] is $y(t) = -\frac{M\alpha e^{\alpha t} + N\beta e^{\beta t}}{Me^{\alpha t} + Ne^{\beta t}}\frac{1}{R(t)}$.

Applying the general setting of the Riccati equation onto the Heston ODE [18] yields

$$\frac{\partial D_j}{\partial \tau} = P_j - Q_j D_j + R D_j^2$$

where $P_j = u_j i\phi - \frac{1}{2}\phi^2, \quad Q_j = b_j - \rho\sigma i\phi, \quad R = \frac{1}{2}\sigma^2$.
The corresponding second-order ODE is

$$w'' + Q_j w' + P_j R w = 0 \tag{23}$$

and the auxiliary equation to solve [23] is $r^2 + Q_j r + P_j r = 0$ with roots

$$\alpha_j = \frac{-Q_j + \sqrt{Q_j^2 - 4P_j R}}{2} = \frac{-Q_j + d_j}{2}$$
$$\beta_j = \frac{-Q_j - \sqrt{Q_j^2 - 4P_j R}}{2} = \frac{-Q_j - d_j}{2}$$

where $d_j = \alpha_j - \beta_j = \sqrt{Q_j^2 - 4P_j R} = \sqrt{(\rho\sigma i\phi - b_j)^2 - \sigma^2(2u_j i\phi - \phi^2)}$.

The solution to the Heston Riccati equation is

$$D_j = -\frac{1}{R}\frac{w'}{w} = -\frac{1}{R}\frac{M\alpha e^{\alpha\tau} + N\beta e^{\beta\tau}}{Me^{\alpha\tau} + Ne^{\beta\tau}} = -\frac{1}{R}\frac{K\alpha e^{\alpha\tau} + \beta e^{\beta\tau}}{Ke^{\alpha\tau} + e^{\beta\tau}}$$

where $K = \frac{M}{N}$. For the initial condition $D_j(0, \phi) = 0$, $K = -\frac{\beta}{\alpha}$. The solution then becomes

$$\begin{aligned} D_j &= -\frac{\beta}{R}\left(\frac{-e^{\alpha\tau} + e^{\beta\tau}}{-g_j e^{\alpha\tau} + e^{\beta\tau}}\right) = -\frac{\beta}{R}\left(\frac{1 + e^{d_j\tau}}{1 - g_j e^{d_j\tau}}\right) \\ &= \frac{Q_j + d_j}{2R}\left(\frac{1 - e^{d_j\tau}}{1 - g_j e^{d_j\tau}}\right) \end{aligned}$$

with $g_j = -K = \frac{\beta}{\alpha} = \frac{b_j - \rho\sigma i\phi + d_j}{b_j - \rho\sigma i\phi - d_j} = \frac{Q_j + d_j}{Q_j - d_j}$. Plugging everything into D_j, we get the final solution for the first ODE

$$D_j(\tau, \phi) = \frac{b_j - \rho\sigma i\phi + d_j}{\sigma^2}\left(\frac{1 - e^{d_j\tau}}{1 - g_j e^{d_j\tau}}\right) \tag{24}$$

The solution for the second ODE can be obtained by integrating [19]

$$C_j = \int_0^\tau ri\phi dy + a\left(\frac{Q_j + d_j}{\sigma^2}\right)\int_0^\tau \left(\frac{1 - e^{d_j y}}{1 - g_j e^{d_j y}}\right) dy + K_1$$

where K_1 is a constant.

Substituting $x = e^{d_j y}$, evaluating the second integral using partial fractions and

substituting for d_j, Q_j and g_j gives the following solution for C_j under the initial condition $C_j(0,\phi) = 0$

$$C_j(\tau,\phi) = ri\phi\tau + \frac{\kappa\theta}{\sigma^2}\left[(b_j - \rho\sigma i\phi + d_j)\tau - 2\ln\left(\frac{1-g_j e^{d_j\tau}}{1-g_j}\right)\right]. \tag{25}$$

Albrecher's little Heston trap [9] on page 3 can be obtained by multiplying by $exp(-d_j\tau)$ in the numerator and denominator.

$$D_j = \frac{b_j - \rho\sigma i\phi + d_j}{g_j\sigma^2}\left(\frac{1-e^{-d_j\tau}}{1-e^{-d_j\tau}/g_j}\right) = \frac{b_j - \rho\sigma i\phi - d_j}{\sigma^2}\left(\frac{1-e^{-d_j\tau}}{1-c_j e^{-d_j\tau}}\right) \tag{26}$$

where $c_j = \frac{1}{g_j} = \frac{b_j - \rho\sigma i\phi - d_j}{b_j - \rho\sigma i\phi + d_j}$.

The logarithm in C_j can be written as

$$\begin{aligned} d_j\tau - 2\ln\left[\frac{1-g_j e^{d_j\tau}}{1-g_j}\right] &= d_j\tau - 2\ln\left[e^{d_j\tau}\left(\frac{e^{-d_j\tau}-g_j}{1-g_j}\right)\right] \\ &= -d_j\tau - 2\ln\left[\frac{1-c_j e^{-d_j\tau}}{1-c_j}\right] \end{aligned}$$

which results in C_j of the following form

$$C_j = ri\phi\tau + \frac{\kappa\theta}{\sigma^2}\left[(b_j - \rho\sigma i\phi - d_j)\tau - 2\ln\left(\frac{1-c_j e^{d_j\tau}}{1-c_j}\right)\right]. \tag{27}$$

□

A.3 Transformed Volatility scheme

We start with finding the SDE for the volatility process. The variance process has been defined as

$$dv_t = \kappa(\theta - v_t)dt + \sigma\sqrt{v_t}dW_t^{(2)}$$

Let $\omega(t, v_t) = \sqrt{v_t}$ be the volatility. Applying Itô's lemma yields

$$\begin{aligned} d\omega &= \frac{\partial\omega}{\partial t}dt + \frac{\partial\omega}{\partial v}dv + \frac{1}{2}\frac{\partial^2\omega}{\partial v^2}dv^2 \\ &= \left[\frac{\partial\omega}{\partial t} + \kappa(\theta - v_t)\frac{\partial\omega}{\partial v} + \frac{1}{2}\sigma^2 v_t\frac{\partial^2\omega}{\partial v^2}\right]dt + \sigma\sqrt{v_t}dW_t^{(2)} \end{aligned} \tag{28}$$

Finding the derivatives required

$$\frac{\partial \omega}{\partial t} = 0, \quad \frac{\partial \omega}{\partial v} = \frac{1}{2\sqrt{v_t}}, \quad \frac{\partial^2 \omega}{\partial v^2} = -\frac{1}{4v_t^{3/2}}$$

and substituting them into the PDE gives following equation

$$d\omega = \frac{1}{2}\kappa \left[(\theta - \frac{\sigma^2}{4\kappa})\omega_t^{-1} - \omega_t \right] dt + \frac{1}{2}\sigma dW_t^{(2)} \tag{29}$$

As mentioned in Section 2.2.2, the mean level $\theta_v = (\theta - \frac{\sigma^2}{4\kappa})\omega_t^{-1}$ is now stochastic over time. Zhu (2011) uses moment-matching to calculate θ_v. The transformed volatility process has the form:

$$\begin{aligned} d\omega_t &= \kappa_v(\theta_v^* - \omega_t)dt + \sigma_v dW_t^{(2)} \\ \kappa_v &= \frac{1}{2}\kappa, \quad \sigma_v = \frac{1}{2}\sigma \end{aligned}$$

Since $\mathbb{E}[\omega_t^2] = \mathbb{E}[v_t]$, we have the relationship

$$\mathbb{E}[v_{t+dt}|\omega_t] = Var[\omega_{t+dt}|\omega_t] + \mathbb{E}[\omega_{t+dt}|\omega_t]^2. \tag{30}$$

In the next step, we find the solutions $v_{t+\Delta}$ and $\omega_{t+\Delta}$ that satisfy the SDEs [2] and [12]. Note that the two SDEs are similar to an Ornstein-Uhlbeck process. Applying Itô's lemma on the ansatz $g(v,t) = v_t e^{\kappa t}$, writing in integral form and evaluating the integral yields

$$\begin{aligned} dg &= \kappa\theta e^{\kappa t}dt + \sigma\sqrt{v_t}e^{\kappa t}dW_t^{(2)} \\ v_{t+dt}e^{k(t+dt)} &= v_t e^{\kappa t} + \int_t^{t+dt} \kappa\theta e^{\kappa s}ds + \int_t^{t+dt} \sigma\sqrt{v_s}e^{\kappa s}dW_s^{(2)} \\ v_{t+dt} &= \theta + (v_t - \theta)e^{-\kappa dt} + \int_t^{t+dt} \sigma\sqrt{v_s}e^{-\kappa(t+dt-s)}dW_s^{(2)} \end{aligned}$$

The conditional expectation of the variance process is therefore $\mathbb{E}[v_{t+dt}|\omega_t] = \theta + (v_t - \theta)e^{-\kappa dt}$. Doing the same for the transformed volatility process with the ansatz $h(\omega,t) = \omega_t e^{\kappa_v t}$ yields

$$\begin{aligned} dh &= \kappa_v\theta_v^* e^{\kappa_v t}dt + \sigma_v e^{\kappa_v t}dW_t^{(2)} \\ \omega_{t+dt} &= \theta_v^* + (\omega_t - \theta_v^*)e^{-\kappa_v dt} + \int_t^{t+dt} \sigma_v e^{-\kappa_v(t+dt-s)}dW_s^{(2)}. \end{aligned}$$

Again, the conditional expectation of the transformed volatility process is $\mathbb{E}[\omega_{t+dt}|\omega_t] = \theta_v^* + (\omega_t - \theta_v^*)e^{-\frac{1}{2}\kappa dt}$.

For the variance of the transformed volatility process, we make use of Itô isometry

$$\begin{aligned}
Var[\omega_{t+dt}|\omega_t] &= \mathbb{E}[(\omega_{t+dt} - \mathbb{E}[\omega_{t+dt}|\omega_t])^2|\omega_t] \\
&= \mathbb{E}\left[\left(\int_t^{t+dt} \sigma_v e^{-\kappa_v(t+dt-s)} dW_s^{(2)}\right)^2\right] \\
&= \mathbb{E}\left[\int_t^{t+dt} \sigma_v^2 e^{-2\kappa_v(t+dt-s)} ds\right] \\
&= \frac{\sigma_v^2}{2\kappa_v}(1 - e^{-2\kappa_v dt}) = \frac{\sigma^2}{4\kappa}(1 - e^{-\kappa dt})
\end{aligned}$$

Plugging the results into equation [30] and solving for θ_v^* yields

$$\begin{aligned}
\underbrace{\sqrt{[E(v_{t+dt}) - Var(\omega_{t+dt})]^+}}_{\beta} &= \theta_v^* + (\omega_t - \theta_v^*)e^{-\frac{1}{2}\kappa dt} \\
\beta - \omega_t e^{-\frac{1}{2}\kappa dt} &= \theta_v^*(1 - e^{-\frac{1}{2}\kappa dt}) \\
\theta_v^* &= \frac{\beta - \omega_t e^{-\frac{1}{2}\kappa dt}}{1 - e^{-\frac{1}{2}\kappa dt}}
\end{aligned}$$

□