Calibration of The Heston Model

A Thesis

submitted to
Indian Institute of Science Education and Research Pune
in partial fulfillment of the requirements for the
BS-MS Dual Degree Programme

by

Subham Kumar Samal

Indian Institute of Science Education and Research Pune
Dr. Homi Bhabha Road,
Pashan, Pune 411008, INDIA.

December, 2024

Supervisor: Michael Alexander Kouritzin

Certificate

This is to certify that this dissertation entitled Calibration of The Heston Model towards the partial fulfillment of the BS-MS dual degree program at the Indian Institute of Science Education and Research, Pune represents study/work carried out by Subham Kumar Samal at University of Alberta and Indian Institute of Science Education and Research Pune under the supervision of Michael Alexander Kouritzin, Professor, Department of Mathematics , University of Alberta, during the academic year 2023-2024.

Michael Alexander Kouritzin

Committee:

Michael Alexander Kouritzin

Anindya Goswami

This thesis is dedicated to Ace, a.k.a Aceu Kutty

Declaration

I hereby declare that the matter embodied in the report entitled Calibration of The Heston Model are the results of the work carried out by me at the Department of Mathematical and Statistical Sciences, University of Alberta and the Department of Mathematics, Indian Institute of Science Education and Research, Pune, under the supervision of Michael Alexander Kouritzin and the same has not been submitted elsewhere for any other degree. Most of the theory and calculations have been contributed by my supervisor Michael. Wherever others contribute, every effort is made to indicate this clearly, with due reference to the literature and acknowledgment of collaborative research and discussions.

Subham Kumar Samal , 20191095

Acknowledgments

This thesis would not have been possible without the guidance and support of my supervisor, Professor Michael A. Kouritzin(aka Mike). Mike has taught me so much about the science and art of Stochastic Calculus. His guidance and advice were truly invaluable. I could not have asked for a better mentor. I would also like to thank Manotosh for his help and counsel, especially during the last few months. The past year was incredibly enjoyable and transformative for me, both as a person and as a researcher. And for this, I have two special people to thank: Arya and Namasi, who inspired me with their passion for science and for this world. This section could not be complete without expressing my gratitude to my parents, who provided valuable funding in addition to their constant love and unconditional support. I would like to thank the fraternity of Hostel-1 and the entire community of IISER Pune for making my final year so memorable. Perhaps the real research is the friends we make along the way. I also want to thank you, the reader, for your interest in my work. It means the world to me.

And last, but not the least, Ace.

Abstract

The Heston Model is a stochastic volatility model heavily used in finance due to its applicability, simplicity and closed-form European call option. Calibrating the Heston Model involves estimating its parameters to match observed market prices. Traditionally, the Heston Model has been calibrated using a combination of least squares, options inference and gradient methods. These traditional calibration techniques provide relatively simple and efficient ways to estimate the parameters of the Heston Model. However, they have limitations in capturing more complex market dynamics. In this thesis, we have worked on a new calibration technique based on an explicit price solution of the Heston Model and stochastic calculus techniques. We use a novel method based on co-variation techniques to estimate the diffusion parameters and a particle filter to estimate the drift parameters. The explicit price solution and the filter used in the calibration process are found to be key for an explicit solution to the Markowitz problem for one Heston stock and one bond.

Contents

Chapter 1

Introduction

Financial models are essential as they provide a mathematical framework for modeling the uncertainty and randomness inherent in financial markets. Since the introduction of the [4], many complex models have been proposed to reflect asset price movement and their derivatives. Initially, Brownian motion [19] itself was used, followed by Geometric Brownian motion. Now, multi-factor models are usually used. Among the first publications about stochastic volatility models were [16], [34], [36] and [14]. Steven Heston introduced the Heston model in [14]. Currently, the Heston model is one of the most popular models because it is complicated enough that it can represent data in a worthwhile manner, yet simple enough that it can yield specific explicit options and stock price formulae. The model represents asset price dynamics with two stochastic differential equations. One governs the asset price, and the other governs its volatility. In the Heston model, we consider a correlation between the asset price and volatility process, as opposed to [36]. The complexity of the model, compared to constant volatility models like Black-Scholes[4], presents challenges in the calibration of its parameters to market data. Traditionally, the Heston model has been calibrated through optimization techniques such as least-squares fitting, Fourier-transform and gradient-based optimizes [9]. Most of these Monte Carlo approaches omit certain schemes and choose their parameters and discretization rather wisely when comparing schemes or showing their performance. Professor Michael Alexander Kouritzin proposed a novel calibration method that uses co-variation techniques to estimate the diffusion parameters and a branching particle filter based on an explicit solution [24] of the Heston model to estimate the drift parameters. In this thesis, we try to apply these methods in a real market. Specifically, this method is designed to improve parameter estimation in models with stochastic volatility, leading to more accurate option pricing and risk assessment. This work contributes to the existing body of research by providing an alternative calibration framework that can be effectively applied to the Heston model, offering potential improvements in both theoretical and practical applications of stochastic volatility models in finance.

1.1 Motivation

A key feature of the Heston model is stochastic volatility, making it suitable for real world applications. However, calibration of stochastic volatility models has always been challenging. Calibration is an important step to ensure accurate pricing, risk management and informed decision-making. The traditional calibration techniques, often face issues such of convergence, efficiency and robustness, especially when applied to noisy market data.

The non-linear and high-dimensional parameter space of the Heston model adds to the complexity of the calibration process. Therefore, there is a strong need for more advanced calibration methods that can improve accuracy and computational performance while remaining robust to market imperfections. The motivation behind this thesis is to address these issues by developing a more efficient and reliable calibration method using co variation techniques and the branching particle filter.

1.2 Survey of Literature

Over the past few decades, various methods have been proposed for the calibration of the Heston model. Traditionally, finite difference methods have been used like in [6], to solve the corresponding differential equations. However, they can be computationally expensive when the model has multiple factors. This led to development of Monte-Carlo based methods in [33], for which one needs simulations. The most successful simulation method for Monte-Carlo multi-factor, path-dependent option pricing is the LSM algorithm developed by [29].

Fourier transform based methods [17] allow for efficient computation but does not solve the robustness challenges associated with the calibration problem. Calibration using combination of least squares, options inference and gradient methods [1] provide relatively simple way of estimating the parameters but have limitations in capturing complex market dynamics and struggle to accurately reproduce volatility smile or skew observed in options market. Particle filtering, originally developed for sequential estimation problems, has also been used for calibration [8]. These methods allow for the estimation of latent variables in stochastic volatility models and have been applied to the Heston model, yielding improvements in estimation accuracy. In [24], it was discovered that the Heston price stochastic differential equation had an explicit solution. we work on a new calibration technique based on this explicit solution.

1.3 Original contribution

This thesis presents a novel approach to the calibration of the Heston model. We use co-variation techniques to estimate the diffusion parameters(correlation between the asset and volatility dynamics, and volatility of volatility). We will discuss this more in detail further.

We estimate the drift parameters(long term mean of the asset price, long term mean of volatility, rate of reversion of volatility) by solving a filtering problem i.e finding the signal given the observation. This mostly follows from [25]. We will build upon this further. Particle filters are being known to be used in calibration techniques. Traditional particle filters suffers from degeneracy, where the particle population collapses to a few weighted particles. We use a branching particle filter similar to the one in [26]. This mitigates the degeneracy issues by branching mechanisms, which allow a more efficient exploration of the parameter space. The original contributions of this thesis can be summarized as follows: 1. Introduction of Co-variation Techniques: By incorporating co-variation methods into the calibration process, this work provides a new approach to extracting meaningful information from market data, leading to more stable and reliable parameter estimation. 2. Application of the Branching Particle Filter: The use of the branching particle filter for calibrating the Heston model is an innovative application that improves both the accuracy and computational efficiency of the calibration process, particularly in noisy and high-dimensional environments.

Chapter 2

Theory

[1]

2.1 The Heston Model

The Heston model is a Stochastic volatility model, introduced by Stephen Heston in 1993 in [14]. It is widely used in finance to describe the volatility of an underlying asset. It assumes that the volatility follows a stochastic process, just like the asset price. This helps in capturing market phenomena such as volatility clustering[7]. The models for financial assets are generally on some real probability space $(\Omega, \mathcal{F}, \mathbb{P})$.

2.1.1 Real Model

The market-state Heston model with time-dependent drift and diffusion coefficients is given by the following stochastic differential equation as shown in [3]:

$$d\begin{pmatrix} S_t \\ V_t \end{pmatrix} = \begin{pmatrix} \mu_t S_t \\ \nu_t - \varrho_t V_t \end{pmatrix} dt + \begin{pmatrix} \left(\sqrt{1-\rho_t^2}\right) S_t V_t^{\frac{1}{2}} & \rho_t S_t V_t^{\frac{1}{2}} \\ 0 & \kappa_t V_t^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} dB_t \\ d\beta_t \end{pmatrix} \tag{2.1}$$

[1] All of the calculations and theoretical results in this chapter have been worked out by Professor Michael Alexander Kouritzin

where B, β are independent standard Brownian motions. It is usually written in the alternative form

$$d\begin{pmatrix} S_t \\ V_t \end{pmatrix} = \begin{pmatrix} \mu_t S_t \\ \nu_t - \varrho_t V_t \end{pmatrix} dt + \begin{pmatrix} S_t V_t^{\frac{1}{2}} dW_t \\ \kappa_t V_t^{\frac{1}{2}} d\beta_t \end{pmatrix} \tag{2.2}$$

but then standard Brownian motions W, β have correlation $\rho_t \in [-1, 1]$, so $dW_t = \sqrt{1 - \rho_t^2} dB_t + \rho_t d\beta_t$. We stick to the formulation where the Brownian motions are independent and emphasize that it is the real model on the real probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The Feller condition [12] : $\nu\varrho > \frac{\kappa^2}{2}$ is often imposed to ensure the volatility stays positive. The volatility usually does not hit zero in the real model but can be a concern in the option model.

2.1.2 Option Model

Options are derivatives on an underlying asset that provides the holder with the right, but not the obligation, to buy or sell the underlying asset at a predetermined *Strike price* on or before a specified expiration date. The Heston model is very useful in option pricing.
Option prices are derived under the risk-free drift setting[11]. It is under a different probability measure Q, where:

$$d\begin{pmatrix} S_t \\ V_t \end{pmatrix} = \begin{pmatrix} r_t S_t \\ \nu_t - \iota_t V_t \end{pmatrix} dt + \begin{pmatrix} \left(\sqrt{1 - \rho_t^2}\right) S_t V_t^{\frac{1}{2}} & \rho_t S_t V_t^{\frac{1}{2}} \\ 0 & \kappa_t V_t^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} dB_t^Q \\ d\beta_t^Q \end{pmatrix} \tag{2.3}$$

with $\binom{B_t^Q}{\beta_t^Q} = \binom{B_t}{\beta_t} + \int_0^t \begin{pmatrix} \frac{(\mu_s - r_s)V_s}{\sqrt{1-\rho_s^2}} + \frac{\rho_s}{\sqrt{1-\rho_s^2}} \cdot \frac{\iota_s - \varrho_s}{\kappa_s} V_s^{\frac{1}{2}} \\ \frac{\iota_s - \varrho_s}{\kappa_s} V_s^{\frac{2}{2}} \end{pmatrix} ds$ being Brownian motion with respect to some other probability Q on $(\Omega, \mathcal{F})$. Here, r_t is the risk-free interest rate, sometimes referred to as Fed funds rate, and ι_t is some risk-free volatility measure that is usually estimated from option data (since the Heston model is an incomplete market).
The term $\mu - r$ is called the risk premium. (It is the return a risk-averse agent demands, holding a unit exposure to the shocks driving the stock price) The difference ($\varrho - \iota$) V is called the *variance risk premium.* It is the return a risk-averse agent demands to hold a unit exposure to the variance innovations.
Usually, the parameters do not vary over time.
Feller becomes: $\nu\iota > \frac{\kappa^2}{2}$. Often not true, so volatility hits zero.

There are two methods for the two non-identical sets of parameters:

- Parameter estimation using (only) the stock data itself for $\rho, \kappa, \mu, \nu, \varrho$.
- Calibration, which is estimating the price using the option price data only, for $\rho, \kappa, r, \nu, \iota$.

In the latter case, we call the parameters implied parameters. Three of the parameters are the same, one can sometimes be obtained from the market i.e. from the Fed rate but ι versus ϱ generally come from these two different sources. Calibrating to the option data cedes the ability to see discrepancies and opportunities in the option pricing data. It provides less good predictions of future stock prices. Hence we will stick to the real model.

2.2 Preliminaries

Definition 2.2.1. A *càdlàg process* [38] is a stochastic process for which the paths $t \mapsto X_t$ are right-continuous with left limits everywhere, with probability one. Such processes are widely used in the theory of noncontinuous stochastic processes. For example, semi-martingales are càdlàg, and continuous-time martingales and many types of Markov processes have càdlàg modifications.
Given a càdlàg process X_t with time index t ranging over the nonnegative real numbers, its left limits are often denoted by

$$X_{t-} = \lim_{s \to t, s<t} X_s$$

for every $t > 0$. Also, the jump at time t is written as

$$\Delta X_t = X_t - X_{t-}$$

Alternative terms used to refer to a càdlàg process are *rcll* (right-continuous with left limits), $R-$ process and right-process.

Definition 2.2.2. A stochastic process X_t adapted to a filtration $\mathcal{F}_t$, is a *Martingale* if $E[X_{t+1}|\mathcal{F}_t] = X_t$. We often replace the $\mathcal{F}_t$ with the σ-algebra generated by $X_0 ... X_t$.

Definition 2.2.3. Given a (local)linear operator $\mathcal{L} : C^2(\mathbb{R}^n) \to C(\mathbb{R}^n)$, *Martingale Problem* for $\mathcal{L}$ consists in finding for each $x_0 \in \mathbb{R}^d$ a probability measure $\mathbb{P}^{x_0}$ over the space of all continuous functions $X : [0, +\infty) \to \mathbb{R}^d$ such that

$$\mathbb{P}^{x_0}\left(X(0) = x_0\right) = 1$$

and whenever

$$f \in \mathrm{C}^2\left(\ \mathrm{R}^{\mathrm{d}}\right)$$

we have that

$$f(X(t)) - f(X(0)) - \int_0^t \mathcal{L}f(X(s))ds$$

is Local Martingale under $\mathbb{P}^{x_0}$.

Theorem 2.2.4. *Suppose $\binom{B_t}{\beta_t}$ is two dimensional standard Brownian motion under P and we define a new probability measure up to time t by*

$$\frac{dQ}{dP} = \exp\left(-\int_0^t X_s^1 dB_s - \int_0^t X_s^2 d\beta_s - \frac{1}{2}\int_0^t \|X_s\|_2^2\, ds\right)$$

where $\|\|_2$ is Euclidean distance and $\int_0^t X_s^1 dB_s$ denotes Ito integration with X assumed to be predictable here. Then, $\binom{B_t^Q}{\beta_t^Q} = \binom{B_t}{\beta_t} + \int_0^t X_s ds$ is a standard Brownian motion with respect to Q.

We use this with $X_s = \begin{pmatrix} \frac{(\mu_s - r_s)V_s}{\sqrt{1-\rho_s^2}} + \frac{\rho_s}{\sqrt{1-\rho_s^2}} \cdot \frac{\iota_s - \varrho_s}{\kappa_s} V_s^{\frac{1}{2}} \\ \frac{\iota_s - \varrho_s}{\kappa_s} V_s^{\frac{1}{2}} \end{pmatrix}$ to go from the real Heston model to the option Heston model.

We consider all processes from here on to be *càdlàg*. Since martingales are constantly turning like Brownian motion or constantly jumping (called pure jump) like other stable processes (or both) standard calculus rules do not apply. Hence we use *Ito's* calculus.

Most of the definitions and facts below are taken from [21][2][18].

Definition 2.2.5. The quadratic covariation $[X, Y]$ of processes X and Y is defined as $[X, Y]_t = \lim_{\|P\| \to 0} \sum_{k=1}^n \left(X_{t_k} - X_{t_{k-1}})(Y_{t_k} - Y_{tk-1}\right)$, where P is the set of partition $\{0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t\}$ and $\|P\| = \max_k \{t_k - t_{k-1}\}$.

Definition 2.2.6. The quadratic variation $[X]$ of X is $[X]_t = [X, X]_t$.

Definition 2.2.7. The variation $V(X)$ of X is $V(X)_t = \lim_{\|P\| \to 0} \sum_{k=1}^n \left|X_{t_k} - X_{t_{k-1}}\right|$. A process bounded variation, i.e. $V(X)_t < \infty$ for each t is said to be a of finite variation and normal (Lebesgue Stieltjes) calculus applies.

2.2.1 Some Important Facts

- $X_t = \int_0^t Z_s ds$ is of finite variation so it does not vary constantly so $[X, Y] = 0$ for any constantly varying Y. X is finite variation so $[X]_t = 0$.
- If X, Y are constantly varying but independent, then $[X, Y] = 0$
- If X or Y is finite variation, then $[X, Y] = 0$.
- If X is pure jump and Y is constantly turning, then $[X, Y] = 0$ and $[X]_t = \sum_{s \leq t} (X_s - X_{s-})^2$.
- The quadratic co-variation is bilinear: $[aX_1 + bX_2, cY_1 + dY_2] = ac\,[X_1, Y_1] + ad\,[X_1, Y_2] + bc\,[X_2, Y_1] + bd[X_2, Y_2]$.
- $\left[\int_0^t X_s dB_s, \int_0^t Y_s dB_s\right] = \int_0^t X_s Y_s ds$, when B is standard Brownian motion.
- This quadratic variation also affects normal exponential functions. $X_t = X_0 \exp(Y_t)$ is the solution to $dX_t = X_t dY_t$ if $Y_0 = 0$ and Y has finite variations e.g. $Y_t = t$.

Definition 2.2.8. [15] A stochastic exponential is of the form: $dX_t = X_{t-} dY_t$. Here the implied integral is in the Ito sense. Integrand is predictable i.e. limit from the left.
The unique solution to the stochastic exponential in the Y constant varying case is:

$$X_t = \exp\left(Y_t - \frac{1}{2}[Y]_t\right)$$

There is an extra factor if Y has jumps as well. Quadratic variation also affects the change of variables formula. The normal change of variables formula

$$f(Y_t) = f(Y_0) + \int_0^t f'(Y_s)\, dY_s$$

for finite variation processes is also affected by quadratic variation. A semi-martingale is the sum of a (local) martingale and a finite variation process.

Definition 2.2.9. [30]For semi-martingales Y and $f \in C^2(\mathbb{R}):\ f(Y_t) = f(Y_0) + \int_0^t f'(Y_s)\, dY_s + \frac{1}{2}\int_0^t f''(Y_s)\, d[Y]_s + \sum_{0<s\leq t} [f(Y_s) - f(Y_{s-}) - f'(Y_{s-})\, \Delta Y_s]$

The last term is zero if Y is continuous.

2.3 Notations and Conditions

The Heston model shown by (2.1) has five parameters. ρ_t, κ_t are the diffusion parameters. μ_t, ν_t, ϱ_t are the drift parameters.

We will show below the diffusion parameters ρ_t, κ_t in our model can be computed from the tick-by-tick data based on variation techniques. Indeed, their values can be determined instantaneously. This is not the case for the Heston drift parameters μ, ν, ϱ but we can still allow them to vary randomly in time. We will take these drift parameters to be a finite-state Markov chain independent of B, β. Let

$$Lf(\mu, \nu, \varrho) = \sum_{\mu', \nu', \varrho' \in \mathcal{A}} \lambda_{\mu,\nu,\varrho \to \mu',\nu',\varrho'} \left[f\left(\mu', \nu', \varrho'\right) - f(\mu, \nu, \varrho) \right] \tag{2.4}$$

where $\mathcal{A} = \mathcal{A}^\mu \times \mathcal{A}^\nu \times \mathcal{A}^\varrho$ is the set of values these parameters can take. Further let $D(L)$ be the continuous, bounded functions $\mathbb{R}^3 \to \mathbb{R}$.

We let $(\Omega, \mathcal{F}, P)$ be a complete probability space with null sets $\mathcal{N}$ supporting initial condition S_0, V_0, the diffusion parameters $\{(\kappa_t, \rho_t), t \geq 0\}$, the drift parameters $\{(\mu_t, \nu_t, \varrho_t), t \geq 0\}$, and the standard Brownian motions $\{(B_t, \beta_t), t \geq 0\}$ such that these four random objects are mutually independent and these four conditions are satisfied

1. $S_0, V_0 \in (0, \infty)^2$.
2. $\{(\kappa_t, \rho_t), t \geq 0\}$ is càdlàg, $\rho_t \in [-1, 1]$ for all $t \geq 0$ and there is some $c_\kappa > 0$ so that $\kappa_t \geq c_\kappa$ for all $t \geq 0$.
3. $\{(B_t, \beta_t), t \geq 0\}$ is a standard, continuous $\mathbb{R}^2$-valued Brownian motion.
4. $\{(\mu_t, \nu_t, \varrho_t), t \geq 0\}$ is the unique càdlàg solution to the martingale problem.

$$M_t^f = f\left(\mu_t, \nu_t, \varrho_t\right) - \int_0^t Lf\left(\mu_s, \nu_s, \varrho_s\right) ds \tag{2.5}$$

is a martingale for all $f \in D(L)$.

We will be using filtering techniques to determine the drift parameters in our real Heston model. The filtering techniques and the corresponding filtering problem is defined in detail in 2.8 . This martingale problem for $\{(\mu_t, \nu_t, \varrho_t), t \geq 0\}$ will form the signal of the filtering

problem. The observations will be constructed below from the price process.
The Feller condition to guarantee non-zero volatility will unfortunately not be enough for our measure change techniques. To verify the Novikov's condition below, we want actual bounds on V so we take $\epsilon > 0$ to be some very small fixed number and define the stopping time:

$$\tau_\epsilon = \left\{ t > 0 : V_t \notin \left[\epsilon, \frac{1}{\epsilon} \right] \right\} \tag{2.6}$$

We will only apply our techniques up until hitting the stopping time. However, if ϵ is very small the probability of hitting the stopping time during the time interval of interest may be very low.

Definition 2.3.1. Suppose $\{Z_t, t \geq 0\}$ is a càdlàg adapted process. Then,
$\mathcal{F}_t^Z \doteq \sigma\left(\sigma\left\{Z_s, s \leq t\right\}, \sigma\left\{\kappa_s, \rho_s, s \geq 0\right\}, \mathcal{N}\right)$, for $t \geq 0$, is the augmented filtration associated with Z.

2.4 Quadratic variation on the Heston Model

Notice that the real Heston stock price S can be written as a stochastic exponential

$$dS_t = S_t \left[\mu_t dt + \left(\sqrt{1 - \rho_t^2} \right) V_t^{\frac{1}{2}} dB_t + \rho_t V_t^{\frac{1}{2}} d\beta_t \right] \tag{2.7}$$

which from above has the solution

$$S_t = S_0 \exp\left(\int_0^t \mu_s - \frac{V_s}{2} ds + \int_0^t \left(\sqrt{1 - \rho_s^2} \right) V_s^{\frac{1}{2}} dB_s + \int_0^t \rho_s V_s^{\frac{1}{2}} d\beta_s \right) \tag{2.8}$$

or equivalently

$$S_t = S_0 \exp\left(\int_0^t \mu_s - \frac{V_s}{2} ds + \int_0^t \left(\sqrt{1 - \rho_s^2} \right) V_s^{\frac{1}{2}} dB_s + \int_0^t \rho_s V_s^{\frac{1}{2}} d\beta_s \right) \tag{2.9}$$

when written in differential form as

$$d \ln\left(\frac{S_t}{S_0} \right) = \left(\mu_t - \frac{V_t}{2} \right) dt + \left(\left(\sqrt{1 - \rho_t^2} \right) V_t^{\frac{1}{2}} \quad \rho_t V_t^{\frac{1}{2}} \right) \begin{pmatrix} dB_t \\ d\beta_t \end{pmatrix} \tag{2.10}$$

By Ito's formula, the real Heston volatility $dV_t = (\nu_t - \varrho_t V_t)\,dt + \left(\kappa_t V_t^{\frac{1}{2}} d\beta_t\right)$ satisfies

$$d\sqrt{V_t} = \frac{\nu_t - \varrho_t V_t}{2\sqrt{V_t}}dt + \frac{\kappa_t}{2}d\beta_t - \frac{\kappa_t^2}{8\sqrt{V_t}}dt \tag{2.11}$$

2.5 Observations

Estimating the drift parameters is interpreted as a filtering problem. We have to find the signal $(\mu_t, \nu_t, \varrho_t)$ given observations. We need to set appropriate observations and then solve the filtering problem.

The observations must be related to the prices because that is all we have access to. However, we do not use our prior stochastic exponential price solution. Instead, we start with our explicit price solution (from [24]):

$$S_t = S_0 \exp\left(\int_0^t \sqrt{1-\rho_s^2}V_s^{\frac{1}{2}}dB_s + \int_0^t \mu_s - \frac{\nu_s\rho_s}{\kappa_s} + \left(\frac{\rho_s\varrho_s}{\kappa_s} - \frac{1}{2}\right)V_s ds + \int_0^t \frac{\rho_s}{\kappa_s}dV_s\right) \tag{2.12}$$

if $\gamma_t = \underbrace{\sqrt{1-\rho_t^2}V_t^{\frac{1}{2}}}_{>0}$, then

$$d\ln(S_t) = \gamma_t dB_t + \left(\mu_t - \frac{\nu_t\rho_t}{\kappa_t} + \left(\frac{\rho\varrho_t}{\kappa_t} - \frac{1}{2}\right)V_t\right)dt + \frac{\rho_t}{\kappa_t}dV_t$$

Under the Feller condition, $V_t \neq 0$ on $t \in [0,\infty)$ even when not stopped at τ_ϵ, which we require now. Further, letting $Y_t = \int_0^t \gamma_s^{-1}\left(d\ln(S_s) - \frac{\rho_s}{\kappa_s}dV_s\right)$, one has that

$$dY_t = \frac{\mu_t - \frac{\nu_t\rho_t}{\kappa_t} + \left(\frac{\rho_t\rho_t}{\kappa_t} - \frac{1}{2}\right)V_t}{\gamma_t}dt + dB_t \tag{2.13}$$

which can be put in filtering form

$$Y_t = \int_0^t h^1(\mu_s, \nu_s, \varrho_s; V_s)\,ds + B_t, \quad h^1(\mu,\nu,\varrho;V) = \frac{\mu - \frac{\nu\rho}{\kappa} + \left(\frac{\rho\varrho}{\kappa} - \frac{1}{2}\right)V}{\sqrt{1-\rho^2}V^{\frac{1}{2}}}. \tag{2.14}$$

$h^1(\mu, \nu, \varrho; V)$ is our first sensor function. A sensor function shows how the observations are related to the underlying state i.e the signal. Y_t is the first observation involving noise B. Next, if we construct $\widehat{Y}_t = \int_0^t \frac{2}{\kappa_s} d\sqrt{V_s}$, then from the above Ito's formula for $d\sqrt{V_s}$ we get:

$$d\widehat{Y}_t = \underbrace{\left(\left(\frac{\nu_t}{\kappa_t} - \frac{\kappa_t}{4}\right)\frac{1}{\sqrt{V_t}} - \frac{\varrho_t}{\kappa_t}\sqrt{V_t}\right)}_{h^2(\mu_t,\nu_t,\varrho_t;V_t)} dt + d\beta_t \tag{2.15}$$

which is the second observation solely is terms of noise β. $h^2(\mu_t, \nu_t, \varrho_t; V_t)$ is our second sensor function. We observe that the sensor function h depends upon the volatility V_t, and the volatility in turn depends upon the second observation $\widehat{Y}_t$ via $V_t = \left(\int_0^t \frac{\kappa_s}{2} d\widehat{Y}_s + \sqrt{V_0}\right)^2$. We have the following proposition:

Proposition 2.5.1. *$\mathcal{F}_t^V \subset \mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(S_0, V_0) = \mathcal{F}_t^S$, for all $t \geq 0$.*

Proof. Since $V_t = \frac{d[\ln(S)]_t}{dt}$ and $\left\{\mathcal{F}_t^S\right\}_{t\geq 0}$ is right continuous, it follows that V_t is $\mathcal{F}_t^S$ measurable for all $t \geq 0$ so $\mathcal{F}_t^V \vee \sigma(V_0) \subset \mathcal{F}_t^S$ for all $t \geq 0$. Next, since $\widehat{Y}_t = \int_0^t \frac{2}{\kappa_s} d\sqrt{V_s}$, with $\sqrt{V_t}$ being an Ito process, one has that

$$\widehat{Y}_t = \lim_{\|P\|\to 0} \sum_{k=1}^{n} \frac{2}{\kappa_{t_{k-1}}}\left(\sqrt{V_{t_k}} - \sqrt{V_{t_{k-1}}}\right) \tag{2.16}$$

where P is the set of partition $\{0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t\}$ and $\|P\| = \max_k \{t_k - t_{k-1}\}$. Hence, $\widehat{Y}_t$ is $\mathcal{F}_t^V$ measurable for all t so $\mathcal{F}_t^{\widehat{Y}} \subset \mathcal{F}_t^S$ for all $t \geq 0$. Moreover, $\gamma_t = \left(\sqrt{1-\rho_t^2}\right) V_t^{\frac{1}{2}}$ is $\mathcal{F}_t^V$-measurable for all t and $Y_t = \int_0^t \gamma_s^{-1}\left(d\ln(S_s) - \frac{\rho_s}{\kappa_s} dV_s\right)$ with $\ln(S_s)$ and V being Ito processes so

$$Y_t = \lim_{\|P\|\to 0} \sum_{k=1}^{n} \frac{1}{\gamma_{t_{k-1}}}\left[\ln\left(\frac{S_{t_k}}{S_{t_{k-1}}}\right) + \frac{\rho_{t_{k-1}}}{\kappa_{t_{k-1}}}\left(V_{t_k} - V_{t_{k-1}}\right)\right] \tag{2.17}$$

where P is the set of partition $\{0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t\}$ and $\|P\| = \max_k \{t_k - t_{k-1}\}$. Hence, Y_t is $\mathcal{F}_t^V \vee \mathcal{F}_t^S = \mathcal{F}_t^S$ measurable for all t so $\mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(V_0) \subset \mathcal{F}_t^S$ for all $t \geq 0$. Conversely,

$$\int_0^t \frac{\kappa_s}{2} d\widehat{Y}_s = \int_0^t d\sqrt{V_s} = \sqrt{V_t} - \sqrt{V_0} \tag{2.18}$$

and $\sqrt{V_t} - \sqrt{V_0}$ is $\mathcal{F}_t^{\widehat{Y}}$-measurable for all $t \geq 0$ by the same mesh type argument so $\mathcal{F}_t^V \subset \mathcal{F}_t^{\widehat{Y}} \vee \sigma(V_0)$ for all $t \geq 0$. Moreover,

$$d\ln(S_t) = \gamma_t dY_t + \frac{\rho_t}{\kappa_t} dV_t \tag{2.19}$$

and $\gamma_t = \left(\sqrt{1-\rho_t^2}\right) V_t^{\frac{1}{2}}$ is $\mathcal{F}_t^V \subset \mathcal{F}_t^{\widehat{Y}} \vee \sigma(V_0)$-measurable so

$$\frac{S_t}{S_0} = \exp\left(\int_0^t \gamma_s dY_s + \int_0^t \frac{\rho_s}{\kappa_s} dV_s\right) \tag{2.20}$$

and S_t is $\mathcal{F}_t^{Y,V}$-measurable for all $t \geq 0$ by the same mesh type argument. Since $\mathcal{F}_t^V \subset \mathcal{F}_t^{\widehat{Y}} \vee \sigma(V_0)$ for all $t \geq 0, S_t$ is $\mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(V_0)$-measurable for all $t \geq 0$ and $\mathcal{F}_t^S \subset \mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(V_0)$ for all $t \geq 0$.

□

We still need to estimate $X_t = (\mu_t, \nu_t, \varrho_t)$, called the signal, given the back observation information $\mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(V_0) = \mathcal{F}_t^S$.

The goal now is to compute $\pi_s(A) = P\left(\mu_s, \nu_s, \varrho_s \in A \mid \mathcal{F}_s^S\right) = P\left(\mu_s, \nu_s, \varrho_s \in A \mid \mathcal{F}_s^{Y,\widehat{Y}} \vee \sigma(S_0, V_0)\right)$ as well as give unnormalized filter[22] to be able to judge models.

2.6 Reference probability

We use a reference probability measure Q in order to construct our model mentioned before. In particular, we let $(\Omega, \mathcal{F}, Q)$ be a complete probability space with null sets $\mathcal{N}$ supporting the initial conditions S_0, V_0, the diffusion parameters $\{(\kappa_t, \rho_t), t \geq 0\}$, the drift parameters $\{(\mu_t, \nu_t, \varrho_t), t \geq 0\}$, and standard $\mathbb{R}^2$-valued Brownian motions $\left\{\left(Y_t, \widehat{Y}_t\right), t \geq 0\right\}$ such that these four random objects are mutually independent and follow the conditions:

1. $S_0, V_0 \in (0,\infty)^2$.

2. $\{(\kappa_t, \rho_t), t \geq 0\}$ is cadlag, $\rho_t \in [-1,1]$ for all $t \geq 0$ and there is some $c_\kappa > 0$ so that $\kappa_t \geq c_\kappa$

3. $\left\{\left(Y_t, \widehat{Y}_t\right), t \geq 0\right\}$ is a standard, continuous $\mathbb{R}^2$-valued Brownian motion.

4. $\{(\mu_t, \nu_t, \varrho_t), t \geq 0\}$ is the unique cadlag solution to the martingale problem

$$M_t^f = f(\mu_t, \nu_t, \varrho_t) - \int_0^t Lf(\mu_s, \nu_s, \varrho_s)\, ds$$

is a martingale for all $f \in D(L)$.

Further, we define $h(\mu_t, \nu_t, \varrho_t; V_t) = \left(h^1(\mu_t, \nu_t, \varrho_t; V_t), h^2(\mu_t, \nu_t, \varrho_t; V_t)\right)^T$,

$$\begin{aligned}(B_t, \beta_t)^T &= \left(Y_t, \widehat{Y}_t\right)^T - \int_0^{t\wedge\tau_\epsilon} h(\mu_s, \nu_s, \varrho_s; V_s)\, ds, \\ d\begin{pmatrix} S_t \\ V_t \end{pmatrix} &= \begin{pmatrix} \mu_t S_t \\ \nu_t - \varrho_t V_t \end{pmatrix} dt + \begin{pmatrix} \left(\sqrt{1-\rho_t^2}\right) S_t V_t^{\frac{1}{2}} & \rho_t S_t V_t^{\frac{1}{2}} \\ 0 & \kappa_t V_t^{\frac{1}{2}} \end{pmatrix} \begin{pmatrix} dB_t \\ d\beta_t \end{pmatrix}\end{aligned} \tag{2.21}$$

and

$$L_t = L_t^\epsilon = \exp\left(\int_0^{t\wedge\tau_\epsilon} \left\langle h(\mu_s, \nu_s, \varrho_s; V_s), d\left(Y_s, \widehat{Y}_s\right)\right\rangle - \frac{1}{2}\int_0^{t\wedge\tau_\epsilon} |h(\mu_s, \nu_s, \varrho_s; V_s)|^2 ds\right) \tag{2.22}$$

so

$$L_t^{-1} = \exp\left(-\int_0^{t\wedge\tau_\epsilon} \langle h(\mu_s, \nu_s, \varrho_s; V_s), d(B_s, \beta_s)\rangle - \frac{1}{2}\int_0^{t\wedge\tau_\epsilon} |h(\mu_s, \nu_s, \varrho_s; V_s)|^2 ds\right)$$

Then, we define $\mathcal{Y}_t = \bigcap_{\delta>0} \sigma\left\{\mathcal{F}_{t+\delta}^{Y,\widehat{Y}}, \mathcal{F}_\infty^{\kappa,\rho,\mu,\nu,\varrho}, S_0, V_0\right\}$ and use Ito's formula to find that

$$\begin{aligned}dL_t &= 1_{t\leq\tau_\epsilon} L_t \left[\left\langle h(\mu_t, \nu_t, \varrho_t; V_t), d\left(Y_t, \widehat{Y}_t\right)\right\rangle - \frac{1}{2}|h(\mu_t, \nu_t, \varrho_t; V_t)|^2 dt\right] + \frac{L_t}{2}|h(\mu_t, \nu_t, \varrho_t; V_t)|^2 dt \\ &= 1_{t\leq\tau_\epsilon} L_t \left\langle h(\mu_t, \nu_t, \varrho_t; V_t), d\left(Y_t, \widehat{Y}_t\right)\right\rangle\end{aligned}$$

so L is at least a local $\{\mathcal{Y}_t\}$-martingale with respect to Q. However, by the Novikov's condition[32], (C1), the definition of h and our imposed bounds on V by stopping, it is in fact a Q-martingale. Now letting

$$\left.\frac{dP^\epsilon}{dQ}\right|_{\mathcal{Y}_t} = L_t^\epsilon$$

one finds by Bayes' rule [37]and Girsanov's theorem that P^ϵ gives us our desired model up to stopping time $\tau_\epsilon = \left\{t > 0 : V_t \notin \left[\epsilon, \frac{1}{\epsilon}\right]\right\}$.

Theorem 2.6.1. *(Baye's Rule) Suppose $(\Omega, \mathcal{F}, Q)$ is a probability space with filtration $\{\mathcal{Y}_t\}_{t\in[0,T]}$ and likelihood martingale $\{L_t, t \in [0,T]\}$. Let Z_t be a $\mathcal{Y}_t$-measurable random variable and $s \leq t \leq T$. Then,*

$$E^P\left[Z_t \mid \mathcal{Y}_s\right] = \frac{E^Q\left[L_t Z_t \mid \mathcal{Y}_s\right]}{L_s} \tag{2.23}$$

where $\frac{dP}{dQ} = L_T$ on $\mathcal{Y}_T$ and E^P denotes (conditional) expectation using probability measure P.

Proof. Cross-multiplying above, we need only show that

$$E^Q\left[E^P\left[Z_t \mid \mathcal{Y}_s\right] L_s 1_F\right] = E^Q\left[L_t Z_t 1_F\right]$$

for all $F \in \mathcal{Y}_s$. However, using the martingale property for L and then taking out knowns and finally total expectation

$$\begin{aligned}
E^Q\left[E^P\left[Z_t \mid \mathcal{Y}_s\right] L_s 1_F\right] &= E^Q\left[E^Q\left[L_s E^P\left[Z_t \mid \mathcal{Y}_s\right] 1_F \mid \mathcal{Y}_s\right]\right] \\
&= E^Q\left[E^Q\left[L_T \mid \mathcal{Y}_s\right] E^P\left[Z_t |\mathcal{Y}_s| 1_F\right]\right. \\
&= E^Q\left[E^Q\left[L_T E^P\left[Z_t \mid \mathcal{Y}_s\right] 1_F \mid \mathcal{Y}_s\right]\right] \\
&= E^Q\left[L_T E^P\left[Z_t \mid \mathcal{Y}_s\right] 1_F\right] \\
&= E^P\left[E^P\left[Z_t \mid \mathcal{Y}_s\right] 1_F\right] \\
&= E^P\left[E^P\left[Z_t 1_F \mid \mathcal{Y}_s\right]\right] \\
&= E^P\left[Z_t 1_F\right] \\
&= E^Q\left[L_T Z_t 1_F\right] \\
&= E^Q\left[L_t Z_t 1_F\right]
\end{aligned}$$

by the martingale property of L. This helps us establish the Girsanov's theorem that we require. □

Theorem 2.6.2. *Suppose (C0, C1, C2 '), (C 3) hold on* $(\Omega, \mathcal{F}, Q)$ *and we define* $P^\epsilon, \tau_\epsilon$ *as above. Then, (C0, C1, C2), (C3) hold on* $(\Omega, \mathcal{F}, P^\epsilon)$*. We have constructed our real Heston models on* P^ϵ *up to this stopping time.*

Proof. First consider (B, β) under the new measure. By Ito's formula

$$g(B_t, \beta_t) = g(B_0, \beta_0) + \int_0^t \frac{\Delta g(B_s, \beta_s)}{2} ds - \int_0^t \nabla g(B_s, \beta_s) h(\mu_s, \nu_s, \varrho_s; V_s) ds + \int_0^t \left\langle \nabla g(B_s, \beta_s), d\left(Y_s, \widehat{Y}_s\right)\right\rangle$$

By applying integration by parts

$$\begin{aligned} L_t g(B_t, \beta_t) &= g(B_0, \beta_0) + \int_0^t L_s \frac{\Delta g(B_s, \beta_s)}{2} ds - \int_0^t L_s \nabla g(B_s, \beta_s) h(\mu_s, \nu_s, \varrho_s; V_s) ds \\ &+ \int_0^t L_s \nabla g(B_s, \beta_s) h(\mu_s, \nu_s, \varrho_s; V_s) ds + M_t(g) \\ &= g(B_0, \beta_0) + \int_0^t L_s \frac{\Delta g(B_s, \beta_s)}{2} ds + M_t(g), \end{aligned}$$

where $t \to M_t(g)$ is a $\{\mathcal{Y}_t\}$-martingale with respect to Q. Then, by two applications of Bayes' rule and Fubini's theorem[13]

$$\begin{aligned} E^P\left[g(B_t, \beta_t) \mid \mathcal{Y}_s\right] - g(B_s, \beta_s) &= \frac{E^Q\left[L_t g(B_t, \beta_t) \mid \mathcal{Y}_s\right] - L_s g(B_s, \beta_s)}{L_s} \\ &= \frac{E^Q\left[\int_s^t L_u \frac{\Delta g(B_u, \beta_u)}{2} du + M_t(g) - M_s(g) \middle| \mathcal{Y}_s\right]}{L_s} \\ &= \int_s^t \frac{E^Q\left[L_u \frac{\Delta g(B_u, \beta_u)}{2} \middle| \mathcal{Y}_s\right]}{L_s} du \\ &= \int_s^t E^P\left[\frac{\Delta g(B_u, \beta_u)}{2} \middle| \mathcal{Y}_s\right] du \end{aligned}$$

so

$$E^P\left[g(B_t, \beta_t) - g(B_s, \beta_s) - \int_s^t \frac{\Delta g(B_u, \beta_u)}{2} du \middle| \mathcal{Y}_s\right] = 0$$

or all g and $0 \leq s \leq t$, which is equivalent to the well posed martingale problem. Hence, (B, β) is standard Brownian motion (starting at 0) with respect to $P(\cdot \mid \mathcal{Y}_0)$ and hence P.

Now, let $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t$. Then, by Bayes rule and the fact that

$$
\begin{aligned}
& E^P\left[l\left(S_0, V_0\right) \prod_{i=0}^n f_i\left(\kappa_{t_i}, \rho_{t_i}, e_{t_i}, \mu_{t_i}, \nu_{t_i}\right) \prod_{j=0}^n g_j\left(B_{t_j}, \beta_{t_j}\right)\right] \\
& = E^Q\left[L_t l\left(S_0, V_0\right) \prod_{i=0}^n f_i\left(\kappa_{t_i}, \rho_{t_i}, Q_{t_i}, \mu_{t_i}, \nu_{t_i}\right) \prod_{j=0}^n g_j\left(B_{t_j}, \beta_{t_j}\right)\right] \\
& = E^Q\left[E^Q\left[L_{t_n} \prod_{j=0}^n g_j\left(B_{t_j}, \beta_{t_j}\right) \mid \nu_0\right] l\left(S_0, V_0\right) \prod_{i=0}^n f_i\left(\kappa_{t_i}, \rho_{t_i}, \varrho_{t_i}, \mu_{t_i}, \nu_{t_i}\right)\right] \\
& = E^Q\left[E^P\left[\prod_{j=0}^n g_j\left(B_{t_j}, \beta_{t_j}\right) \mid \mu_0\right] l\left(S_0, V_0\right) \prod_{i=0}^n f_i\left(\kappa_{t_i}, \rho_{t_i}, \varrho_{t_i}, \mu_{t_i}, \nu_{t_i}\right)\right] \\
& = E^P\left[\prod_{j=0}^n g_j\left(B_{t_j}, \beta_{t_j}\right)\right] E^Q\left[l\left(S_0, V_0\right)\right] E^Q\left[\prod_{i=0}^n f_i\left(\kappa_{t_i}, \rho_{t_i}, e_{t_i}, \mu_{t_i}, \nu_{ti}\right)\right]
\end{aligned}
$$

for all bounded measurable l and f_i as well as bounded twice continuously differentiable g_j with bounded derivatives. Since this collection s.p., s.s.p. and is closed under multiplication it follows by [5] that this collection separates probability measures. Hence, the distribution of S_0, V_0 and $\kappa, \rho, \varrho, \mu, \nu$ is the same as under Q while (B, β) is a $\mathbb{R}^2$-Brownian motion independent of these other random objects. □

2.7 Unnormalized Filter and Bayes' Factor

We still need to compute the filter. Moreover, in practice, we will not know which is the correct (μ, ν, ϱ). For example, we might have several different operators L or several different initial conditions for (μ, ν, ϱ) to choose from. Hence, model selection methods like Bayes' factor are valuable. In this thesis, we have used the Wright-Fisher[35] model for the operator. We have explained the Wright-Fisher model in detail in the next chapter. We must first introduce the *unnormalized filter*

$$
\sigma_t^\epsilon(f) = E^Q\left[L_t^\epsilon f\left(\mu_t, \nu_t, \varrho_t\right) \mid \mathcal{F}_s^{Y,\widehat{Y}} \vee \sigma\left(S_0, V_0\right)\right] \tag{2.24}
$$

Proposition 2.7.1. *The following Baye's rule holds*

$$\pi_t^\epsilon(f) = \frac{\sigma_t^\epsilon(f)}{\sigma_t^\epsilon(1)} \tag{2.25}$$

where $\pi_s(A) = P\left(\mu_s, \nu_s, \varrho_s \in A \mid \mathcal{F}_s^S\right) = P\left(\mu_s, \nu_s, \varrho_s \in A \mid \mathcal{F}_s^{Y,\widehat{Y}} \vee \sigma\left(S_0, V_0\right)\right)$.

Proof. We have to show

$$E^{P^\epsilon}\left[f\left(\mu_t, \nu_t, \varrho_t\right) \mid \mathcal{F}_t^S\right] E^Q\left[L_t^\epsilon \mid \mathcal{F}_t^S\right] = E^Q\left[L_t^\epsilon f\left(\mu_t, \nu_t, \varrho_t\right) f\left(\mu_t, \nu_t, \varrho_t\right) \mid \mathcal{F}_t^S\right]$$

Let $A \in \mathcal{F}_t^S$ and $X_t = (\mu_t, \nu_t, \varrho_t)$

$$\begin{aligned} E^Q\left[E^{P^\epsilon}\left[f\left(X_t\right) \mid \mathcal{F}_t^S\right] E^Q\left[L_t^\epsilon \mid \mathcal{F}_t^S\right] 1_A\right] &= E^Q\left[E^{P^\epsilon}\left[f\left(X_t\right) \mid \mathcal{F}_t^S\right] E^Q\left[L_t^\epsilon \mid \mathcal{F}_t^S\right] 1_A\right] \\ &= E^Q\left[E^Q\left[L_t^\epsilon E^{P^\epsilon}\left[f\left(X_t\right) \mid \mathcal{F}_t^S\right] 1_A \mid \mathcal{F}_t^S\right]\right] \\ &= E^Q\left[L_t^\epsilon E^{P^\epsilon}\left[f\left(X_t\right) 1_A \mid \mathcal{F}_t^S\right]\right] \\ &= E^{P^\epsilon}\left[E^{P^\epsilon}\left[f\left(X_t\right) 1_A \mid \mathcal{F}_t^S\right]\right] \\ &= E^{P^\epsilon}\left[f\left(X_t\right) 1_A\right] \\ &= E^Q\left[L_t f\left(X_t\right) 1_A\right]. \end{aligned}$$

Hence, we can get the probability distribution of the drift parameters from the unnormalized filter. We can also think of one P^ϵ as just being a possible model of reality corresponding to one possible Markov chain model for $(\mu_t, \nu_t, \varrho_t)$. If we are identifying static parameters, then there is just one static parameter corresponding to no movement. However, suppose we are allowing the parameters to vary over time. In that case, we have a different model for each possible way in which they move i.e. for each Markov chain generator. Suppose we are trying to select from m possible models for $(\mu_t, \nu_t, \varrho_t)$, then we have m different likelihoods, $L_t^\epsilon(1), \ldots, L_t^\epsilon(m)$, (of each model to the reference model where the observations are independent of parameters) and estimate m unnormalized filters

$$\sigma_t^{\epsilon,i}(f) = E^Q\left[L_t^\epsilon(i) f\left(\mu_t, \nu_t, \varrho_t\right) \mid \mathcal{F}_s^{Y,\widehat{Y}} \vee \sigma\left(S_0, V_0\right)\right]$$

so

$$\sigma_t^{\epsilon,i}(1) = E^Q\left[L_t^\epsilon(i) \mid \mathcal{F}_s^{Y,\widehat{Y}} \vee \sigma\left(S_0, V_0\right)\right]$$

s the estimated Likelihood of this model being true compared to the canonical model Q

where the observations are just noise. To compare the two models P_1^ϵ and P_2^ϵ, defined by two likelihoods $L_t^\epsilon(1)$ and $L_t^\epsilon(2)$, one uses the Bayes factor, which is the ratio of unnormalized filter total masses

$$\frac{E^Q\left[L_t^\epsilon(1) \mid \mathcal{F}_t^Y\right]}{E^Q\left[L_t^\epsilon(2) \mid \mathcal{F}_t^Y\right]}$$

In other words, Bayes' factor is (integrated likelihood ratio) evidence of model 1 over model 1 based upon the observed data. This method was developed in K[27], where they come up with a different unnormalized filter with explicit Bayes' factor to save steps. □

2.8 Filter Theorem

The following development largely follows [25]. They handle a more general filtering problem where the signal need not even be right continuous nor have second moments. However, they do not have the case where the sensor function h depends upon the observations (through V) here. Hence, there are significant enough differences to warrant including all the steps below.

With the goal of deriving both the unnormalized and normalized filters, we logically start with the unnormalized filter. Further, following Kouritzin and Long (2008), we will divide deriving the unnormalized filter into three steps:

1. deriving a martingale problem for it,
2. stating an appropriate martingale representation theorem and
3. combining these to come up with a stochastic differential equation for the filter.

It makes sense to start with ii) since it is actually a general result which will just be applied to filtering.

Lemma 2.8.1. *Let $\{\mathbb{M}_t, t \geq 0\}$ be a cadlag $\{\mathcal{F}_t\}_{t\geq 0}$-martingale starting at zero, $\{Y_t, t \geq 0\}$ be an $\{\mathcal{F}_t\}_{t\geq 0}$-adapted $\mathbb{R}^d$ standard Brownian motion and co-variation $[\mathbb{M}, Y]$ be differentiable with $\frac{d}{dt}[\mathbb{M}, Y]_t$ integrable almost everywhere. Then,*

$$E\left[\mathbb{M}_t \mid \mathcal{F}_t^Y\right] = \int_0^t \left\langle E\left[\frac{d}{ds}[\mathbb{M}, Y]_s \middle| \mathcal{F}_s^Y\right], dY_s\right\rangle, \forall t \in [0, T] \quad a.s. \tag{2.26}$$

Proof. $\mathcal{M}_t = E\left[\mathbb{M}_t \mid \mathcal{F}_t^Y\right]$ is the unique $\mathcal{F}_t^Y$-measurable random variable such that $E\left[\mathcal{M}_t\xi_t\right] = E\left[\mathbb{M}_t\xi_t\right]$ for all bounded $\xi_t \in \mathcal{F}_t^Y$. Without loss of generality, we can take $\xi_t = E\left[\xi_T \mid \mathcal{F}_t^Y\right]$ with $\xi_T \in \mathcal{F}_T^Y$. $\{\mathcal{M}_t, t \geq 0\}$ is verifiably a $\left\{\mathcal{F}_t^Y\right\}_{t\geq 0}$-martingale. $\left\{\mathcal{F}_t^Y\right\}_{t\geq 0}$ is continuous so $\{\mathcal{M}_t, t \geq 0\}$ and $\{\xi_t, t \geq 0\}$ have continuous modifications (see II. 2.9 of [31]). By the martingale representation theorem (Problem 3.4.16 of [20]), there are $\mathbb{R}^2$-valued $\left\{\mathcal{F}_t^Y\right\}_{t\geq 0}$-progressively measurable processes $\{\alpha_t, t \geq 0\}, \{\Phi_t, t \geq 0\}$ such that

$$\mathcal{M}_t = \int_0^t \langle \alpha_s, dY_s\rangle, \xi_t = \int_0^t \langle \Phi_s, dY_s\rangle, \forall t \in [0,T] \text{ and } \int_0^T |\alpha_s|^2 \vee |\Phi_s|^2\, ds < \infty \text{ a.s.}$$

By enlarging the class of martingales $\{\xi_t, t \geq 0\}$ under consideration, we can consider all progressively measurable $\{\Phi_t, t \geq 0\}$ such that $\int_0^{\bar{T}} |\Phi_s|^2\, ds < \infty$ a.s. (Also, by the almost sure monotonicity of conditional expectation, we can make ξ_t bounded in (ω, t).) Now, we define the stopping times

$$\kappa_N = \inf\left\{t > 0 : \left|\int_0^t \langle \alpha_s, dY_s\rangle\right| \vee \left|\int_0^t \langle \Phi_s, dY_s\rangle\right| > N\right\} \tag{2.27}$$

Then, $\int_0^t E\left[\left|\alpha_s^f\right|^2 1_{s\leq\kappa_N}\right] ds < \infty$ so that $\left|\alpha_s^f\right| 1_{s\leq\kappa_N}, |\Phi_s| 1_{s\leq\kappa_N} \in L^2(\Omega)$ almost everywhere in s. By Doob's optimal sampling theorem[10], we have that

$$\mathcal{M}_{t\wedge\kappa_N} = E\left[E\left[\mathbb{M}_t \mid \mathcal{F}_t^Y\right] \mid \mathcal{F}_{t\wedge\kappa_N}^Y\right] = E\left[\mathbb{M}_t \mid \mathcal{F}_{t\wedge\kappa_N}^Y\right] \quad \text{a.s.} \tag{2.28}$$

and it follows by Kunita-Watanabe [28] that

$$E\int_0^t \langle \alpha_s, \Phi_s\rangle 1_{s\leq\kappa_N} ds = E\left[\mathcal{M}_{t\wedge\kappa_N}\xi_{t\wedge\kappa_N}\right] = E\left[\mathbb{M}_t\xi_{t\wedge\kappa_N}\right]$$

or using Fubini and differentiating

$$E\left[\langle \alpha_t, \Phi_t\rangle 1_{t\leq\kappa_N}\right] = \frac{d}{dt}E\left[\mathcal{M}_{t\wedge\kappa_N}\xi_{t\wedge\kappa_N}\right] = \frac{d}{dt}E\left[\mathbb{M}_t\xi_{t\wedge\kappa_N}\right]$$

It follows by integration by parts that

$$\mathbb{M}_t\xi_{t\wedge\kappa_N} = \int_0^t \mathbb{M}_s\Phi_s 1_{s\leq\kappa_N} dY_s + \int_0^t \xi_{s\wedge\kappa_N} d\mathbb{M}_s + \int_0^t \Phi_s 1_{s\leq\kappa_N} d[\mathbb{M}, Y]_s$$

and by hypothesis as well as the fundamental theorem of calculus, Fubini's theorem

$$E\left[\mathbb{M}_t\xi_{t\wedge\kappa_N}\right] = E\left[\int_0^t \Phi_s 1_{s\leq\kappa_N}\frac{d}{ds}[\mathbb{M},Y]_s ds\right] = \int_0^t E\left[\Phi_s 1_{s\leq\kappa_N}\frac{d}{ds}[\mathbb{M},Y]_s\right] ds$$

Differentiating and comparing, one has that

$$E\left[\langle\alpha_t,\Phi_t\rangle\, 1_{t\leq\kappa_N}\right] = E\left[\Phi_t 1_{t\leq\kappa_N}\frac{d}{dt}[\mathbb{M},Y]_t\right]$$

and since Φ is an arbitrary progressive process one has that

$$\begin{aligned}\alpha_t &= \lim_{N\to\infty}\alpha_t 1_{t\leq\kappa_N}\\ &= \lim_{N\to\infty} E\left[1_{t\leq\kappa_N}\frac{d}{dt}[\mathbb{M},Y]_t\middle|\ \mathcal{F}_t^Y\right]\\ &= \lim_{N\to\infty} 1_{t\leq\kappa_N} E\left[\frac{d}{dt}[\mathbb{M},Y]_t\middle|\ \mathcal{F}_t^Y\right]\\ &= E\left[\frac{d}{dt}[\mathbb{M},Y]_t\middle|\ \mathcal{F}_t^Y\right]\end{aligned}$$

a.s. for almost every $t\in[0,T]$ since $1_{t\leq\kappa_N}$ is $\mathcal{F}_t^Y$-measurable. □

Now, we turn to the equations for the filter and unnormalized filter. For ease of notation, we let $X_t = (\mu_t,\nu_t,\varrho_t)$ and recall $M_t^f = f(\mu_t,\nu_t,\varrho_t) - \int_0^t Lf(\mu_s,\nu_s,\varrho_s)\,ds$ is the martingale from the signal's martingale problem for each $f\in D(L)$. Then, by our martingale problem and integration by parts, we have that

$$f(X_t)\,L_t = f(X_0)+\int_0^t L_s Lf(X_s)\,ds+\int_0^t L_s dM_s^f+\int_0^t 1_{s\leq\tau_\epsilon} f(X_s)\,L_s\left\langle h(X_s;V_s)\,,d\left(Y_s,\widehat{Y}_s\right)\right\rangle \tag{2.29}$$

Basically, we will take expectations of this to get our unnormalized filter for real Heston parameters. We need to define the innovations processes $I_t = Y_t - \int_0^t \frac{\pi_s(\mu_s)-\frac{\pi_s(\nu_s)\rho_s}{\kappa_s}+\left(\frac{\rho_s\pi_s(\varrho_s)}{\kappa_s}-\frac{1}{2}\right)V_s}{\gamma_s}ds$ and $\widehat{I}_t = \widehat{Y}_t - \int_0^t\left(\left(\frac{\pi_s(\nu_s)}{\kappa_s}-\frac{\kappa_s}{4}\right)\frac{1}{\sqrt{V_t}}-\frac{\pi_s(\varrho_s)}{\kappa_s}\sqrt{V_t}\right)ds$, where $\gamma_t = \sqrt{1-\rho_t^2}\sqrt{V_t}$, for the equation of the normalized filter. Note that the unnormalized filter equation for the classical problem is usually referred to as the *Duncan-Mortensen-Zakai* (DMZ) equation[40] after its founders and the equation for the normalized filter equation as the *Kushner-Stratonovich* or the *Fujisaki-Kallianpur-Kunita* (FKK) equation[39].

Theorem 2.8.2. *1. $\{I_t, t \geq 0\}$ and $\left\{\widehat{I}_t, t \geq 0\right\}$ are independent standard Brownian motions on $\left(\Omega, \{\mathcal{F}_t^S\}_{t\geq 0}, P\right)$,*

2. $\sigma\left\{I_v - I_u, \widehat{I}_v - \widehat{I}_u : s \leq u \leq v \leq T\right\}$ and $\mathcal{F}_s^S$ are independent,

3. the filter process π satisfies

$$\pi_t(f) = \pi_0(f) + \int_0^t \pi_s(Lf)ds + \int_0^t \pi_s\left(h_s^{V_s} f\right) - \pi_s\left(h_s^{V_s}\right)\pi_s(f)d\left\langle I_s, \widehat{I}_s\right\rangle \qquad a.s.$$

for all $f \in D(L)$, and

4. the unnormalized filter σ satisfies

$$\sigma_t(f) = \sigma_0(f) + \int_0^t \sigma_s(Lf)ds + \int_0^t \sigma_s\left(h_s^{V_s} f\right) d\left\langle Y_s, \widehat{Y}_s\right\rangle \;\; a.s.$$

for all f.

Proof. We start from the integration by parts formula above to derive the unnormalized filter equation. Noting $\sigma_t(f) = E^Q\left[L_t f(X_t) \mid \mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(S_0, V_0)\right]$ and recalling $\mathcal{F}_s^V \subset \mathcal{F}_s^{Y,\widehat{Y}} \vee \sigma(S_0, V_0)$, one has by linearity, Fubini's theorem and independent increments that

$$\begin{aligned}
\sigma_t(f) &= E^Q\left[f(X_0) \mid \mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(S_0, V_0)\right] + \int_0^t E^Q\left[L_s L f(X_s) \mid \mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(S_0, V_0)\right] ds \\
&\quad + E^Q\left[\int_0^t L_s dM_s^f + \int_0^t 1_{s\leq\tau_\epsilon} f(X_s) L_s \left\langle h(X_s; V_s), d\left(Y_s, \widehat{Y}_s\right)\right\rangle \mid \mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(S_0, V_0)\right] \\
&= \sigma_0(f) + \int_0^t \sigma_s(Lf)ds + \mathcal{M}_t^f
\end{aligned}$$

where $t \to \mathcal{M}_t = E^Q\left[\mathbb{M}_t \mid \mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(S_0, V_0)\right]$ is some $\left\{\mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma(S_0, V_0)\right\}_{t\geq 0}$-martingale, and

$$t \to \mathbb{M}_t = \int_0^t L_s dM_s^f + \int_0^t 1_{s\leq\tau_\epsilon} f(X_s) L_s \left\langle h(X_s; V_s), d\left(Y_s, \widehat{Y}_s\right)\right\rangle$$

is some $\left\{\mathcal{F}_t^X \vee \mathcal{F}_t^{Y\widehat{Y}} \vee \sigma(S_0, V_0)\right\}_{t>0}$-martingale both starting at zero. Now, we use the prior lemma to find a representation for $\mathcal{M}_t$: By independence under Q,

$$\left[\int L_s dM_s^f, (Y, \widehat{Y})\right]_t = \int_0^t L_s d\left[M^f, (Y, \widehat{Y})\right]_s = 0$$

while

$$\left[\int 1_{s\le\tau_\epsilon} f\left(X_s\right) L_s \left\langle h\left(X_s;V_s\right), d\left(Y_s,\widehat{Y}_s\right)\right\rangle, (Y,\widehat{Y})\right]_t$$
$$= \int_0^t 1_{s\le\tau_\epsilon} f\left(X_s\right) L_s h\left(X_s;V_s\right) d\left[\left(Y_s,\widehat{Y}_s\right), (Y,\widehat{Y})\right]_s$$
$$= \int_0^t 1_{s\le\tau_\epsilon} f\left(X_s\right) L_s h\left(X_s;V_s\right) ds$$

This means

$$\frac{d}{ds}[\mathbb{M},Y]_s = 1_{s\le\tau_\epsilon} f\left(X_s\right) L_s h^{V_s}\left(X_s\right)$$

and

$$\mathcal{M}_t = \int_0^t \left\langle E^Q\left[1_{s\le\tau_\epsilon} f\left(X_s\right) L_s h^{V_s}\left(X_s\right) \mid \mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma\left(S_0,V_0\right)\right], d\left(Y_s,\widehat{Y}_s\right)\right\rangle$$
$$= \int_0^t 1_{s\le\tau_\epsilon} \left\langle \sigma_s\left(fh^{V_s}\right), d\left(Y_s,\widehat{Y}_s\right)\right\rangle$$

So,

$$\sigma_t(f) = \sigma_0(f) + \int_0^t \sigma_s(Lf)ds + \int_0^t 1_{s\le\tau_\epsilon} \left\langle \sigma_s\left(fh^{V_s}\right), d\left(Y_s,\widehat{Y}_s\right)\right\rangle, \forall t\in[0,T] \text{ a.s.}$$

for all $f\in D(L)$. By linearity, we have

$$\begin{pmatrix} I_t \\ \hat{I}_t \end{pmatrix} = \begin{pmatrix} Y_t \\ \widehat{Y}_t \end{pmatrix} - \int_0^t E^P\left[h^{V_s}\left(X_s\right) \mid \mathcal{F}_s^{Y,\widehat{Y}} \vee \sigma\left(S_0,V_0\right)\right] ds$$
$$= \begin{pmatrix} B_t \\ \beta_t \end{pmatrix} + \int_0^t h^{V_s}\left(X_s\right) - E^P\left[h^{V_s}\left(X_s\right) \mid \mathcal{F}_s^{Y,\widehat{Y}} \vee \sigma\left(S_0,V_0\right)\right] ds$$

Now, since B,β are Brownian motions with respect to large filtration $\{\mathcal{F}_t\}$ we have that

$$E^P\left[E^P\left[\begin{pmatrix} B_t \\ \beta_t \end{pmatrix}\middle| \mathcal{F}_u\right]\middle| \mathcal{F}_u^{Y,\widehat{Y}} \vee \sigma\left(S_0,V_0\right)\right] = E^P\left[\begin{pmatrix} B_u \\ \beta_u \end{pmatrix}\middle| \mathcal{F}_u^{Y,\widehat{Y}} \vee \sigma\left(S_0,V_0\right)\right]$$

which implies

$$
\begin{aligned}
E^P\left[\left.\begin{pmatrix} I_t \\ \hat{I}_t \end{pmatrix}\right| \mathcal{F}_u^{Y,\widehat{Y}} \vee \sigma\left(S_0, V_0\right)\right] &= E^P\left[\left.\begin{pmatrix} B_u \\ \beta_u \end{pmatrix}\right| F_u^{\gamma,\hat{Y}} \vee \sigma\left(S_0, V_0\right)\right] \\
&+ \int_0^u E^p\left[h^{V_s}\left(X_s\right) \mid \mathcal{F}_u^{\gamma,\hat{Y}} \vee \sigma\left(S_0, V_0\right)\right] \\
&- E^p\left[h^{V_s}\left(X_s\right) \mid \mathcal{F}_s^{\gamma,\hat{Y}} \vee \sigma\left(S_0, V_0\right)\right] ds \\
&= \begin{pmatrix} Y_u \\ \widehat{Y}_u \end{pmatrix} - \int_0^u E^P\left[h^{V_s}\left(X_s\right) \mid \mathcal{F}_s^{Y,\widehat{Y}} \vee \sigma\left(S_0, V_0\right)\right] ds \\
&= \begin{pmatrix} I_u \\ I_u \end{pmatrix}
\end{aligned}
$$

so $\left\{\binom{I_t}{\hat{I}_t}, t \geq 0\right\}$ is a $\left\{\mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma\left(S_0, V_0\right)\right\}$-martingale.

Its quadratic variation and co-variation are: $[I, I]_t = [B, B]_t = t, [\widehat{I}, \widehat{I}]_t = [\beta, \beta]_t = t$ and $[I, \widehat{I}]_t = [B, \beta]_t = 0$ by the finite variations of the differences and independence.

Hence, $\left\{\binom{I_t}{\widehat{I}_t}, t \geq 0\right\}$ is an $\mathbb{R}^2$ Brownian motion with respect to $\left\{\mathcal{F}_t^{Y,\widehat{Y}} \vee \sigma\left(S_0, V_0\right)\right\}_{t\geq 0}$.

Finally, we turn to the FKK filtering equation. From our Bayes' rule above, we have that

$$
\pi_t^\epsilon(f) = \frac{\sigma_t^\epsilon(f)}{\sigma_t^\epsilon(1)}
$$

But, $\sigma_t(1) = \sigma_0(1) + \int_0^t 1_{s\leq\tau_\epsilon}\left\langle \sigma_s\left(h^{V_s}\right), d\left(Y_s, \widehat{Y}_s\right)\right\rangle$ since L must annihilate 1 and we have by Ito's formula that

$$
\frac{1}{\sigma_t(1)} = \frac{1}{\sigma_0(1)} - \int_0^t \frac{1_{s\leq\tau_\epsilon}\left\langle \sigma_s\left(h^{V_s}\right), d\left(Y_s, \widehat{Y}_s\right)\right\rangle}{\sigma_s^2(1)} + \int_0^t \frac{1_{s\leq\tau_\epsilon}^2\left(h^{V_s}\right)}{\sigma_s^3(1)} ds
$$

and by integration by parts

$$
\begin{aligned}
\pi^\epsilon_t(f) &= \pi^\epsilon_0(f) - \int_0^t \frac{1_{s\le\tau_\epsilon}\sigma_s(f)\left\langle \sigma_s\left(h^{V_s}\right), d\left(Y_s, \widehat{Y}_s\right)\right\rangle}{\sigma_s^2(1)} + \int_0^t \frac{1_{s\le\tau_\epsilon}\sigma_s(f)\sigma_s^2\left(h^{V_s}\right)}{\sigma_s^3(1)}ds \\
&+ \int_0^t \frac{\sigma_s(Lf)}{\sigma_s(1)}ds + \int_0^t \frac{1_{s\le\tau_\epsilon}}{\sigma_s(1)}\left\langle \sigma_s\left(fh^{V_s}\right), d\left(Y_s, \widehat{Y}_s\right)\right\rangle - \int_0^t \frac{1_{s\le\tau_\epsilon}\left\langle \sigma_s\left(h^{V_s}\right), \sigma_s\left(fh^{V_s}\right)\right\rangle}{\sigma_s^3(1)}ds \\
&= \pi^\epsilon_0(f) - \int_0^t 1_{s\le\tau_\epsilon}\pi_s(f)\left\langle \pi_s\left(h^{V_s}\right), d\left(Y_s, \widehat{Y}_s\right)\right\rangle + \int_0^t 1_{s\le\tau_\epsilon}\pi_s(f)\pi_s^2\left(h^{V_s}\right)ds \\
&+ \int_0^t \pi_s(Lf)ds + \int_0^t 1_{s\le\tau_\epsilon}\left\langle \pi_s\left(fh^{V_s}\right), d\left(Y_s, \widehat{Y}_s\right)\right\rangle - \int_0^t 1_{s\le\tau_\epsilon}\left\langle \pi_s\left(h^{V_s}\right), \pi_s\left(fh^{V_s}\right)\right\rangle ds \\
&= \pi^\epsilon_0(f) + \int_0^t \pi_s(Lf)ds + \int_0^t 1_{s\le\tau_\epsilon}\left\langle \pi_s\left(fh^{V_s}\right) - \pi_s(f)\pi_s\left(h^{V_s}\right), d\left(I_s, \widehat{I}_s\right)\right\rangle a.s.
\end{aligned}
$$

□

2.9 Particle Filter

One can discretize the filtering equations and produce computer workable solutions. However, it is often more effective to go back to the signal likelihood representation and use particles instead. In particular, we work on the reference probability space $(\Omega, \mathcal{F}, Q)$ and create càdlàg particles $\{(\mu_t^i, \nu_t^i, \varrho_t^i), t \geq 0\}_{i=1}^{N_t}$ that evolve independently of each other and the observations according to the signal's martingale problem

$$M_t^{f,i} = f\left(\mu_t^i, \nu_t^i, \varrho_t^i\right) - \int_0^t Lf\left(\mu_s^i, \nu_s^i, \varrho_s^i\right) ds$$

is a martingale for all $f \in D(L)$, where

$$Lf(\mu, \nu, \varrho) = \sum_{\mu', \nu', \varrho' \in \mathcal{A}} \lambda_{\mu,\nu,\varrho \to \mu',\nu',\varrho'} \left[f\left(\mu', \nu', \varrho'\right) - f(\mu, \nu, \varrho)\right]$$

Then, we weight the particles by the likelihoods replacing the signal with the i^{th} particle

$$L_t^i = \exp\left(\int_0^{t\wedge\tau_\epsilon} \left\langle h\left(\mu_s^i, \nu_s^i, \varrho_s^i; V_s\right), d\left(Y_s, \widehat{Y}_s\right)\right\rangle - \frac{1}{2}\int_0^{t\wedge\tau_\epsilon} \left|h\left(\mu_s^i, \nu_s^i, \varrho_s^i; V_s\right)\right|^2 ds\right)$$

and find by the law of large numbers that

$$\frac{1}{N}\sum_{i=1}^{N} L_t^i \delta_{\left(\mu_t^i, \nu_t^i, e_t^i\right)}(\cdot) \Rightarrow \sigma_t(\cdot) \quad \text{a.s.}$$

However, it is a well known problem in particle filtering that most particles will drift away from the signal causing the weights $\{L_t^i\}_{i=1}^N$ to be very unequal and the effective average on the left of the above equation to be over a small portion of the particles. Consequently, for good approximation with a reasonable number of particles we should branch or interact the particles in an unbiased way. [23] provides an effective method for doing this. We will discuss the exact algorithm in the next chapter.

Chapter 3

Calibration and Simulation of the Heston model

In this chapter, first, we will discuss the algorithms used for estimating the diffusion and drift parameters. Then, we apply those algorithms on

1. Simulated Heston model data
2. Real market data(Apple and Microsoft tick-by-tick stock prices)

3.1 Estimating Diffusion parameters

By applying quadratic co-variation techniques on (2.10) and (2.11) we get,

- $d[\ln(S)]_t = V_t dt$
- $d[V, \ln(S)]_t = \rho_t \kappa_t V_t dt$
- $d[\sqrt{V}]_t = \frac{\kappa_t^2}{4} dt$
- $\frac{d[V, \ln(S)]_t}{d[\ln(S)]_t} = \rho_t \kappa_t$

We can solve for ρ_t and κ_t by computing $[\ln(S)]_t, [V, \ln(S)]_t$, and $[\sqrt{V}]_t$ from tick-by-tick data. We can compute these by applying the quadratic co-variation definitions discussed in 2.2. In all of the steps with limit, we use a small mesh instead. In step 3,9 and 15, P is the

Algorithm 1 Estimating the diffusion parameters ρ and κ using quadratic co-variation.

1: Input values of S_t from $t \in (0, N)$
2: **for** $t = 1$ to N **do**
3: Calculate the quadratic variation of the logarithm of the asset price $[\ln(S)]_t = \lim_{\|P\|\to 0} \sum_{k=1}^{n} \left(\ln(S)_{t_k} - \ln(S)_{t_{k-1}}\right)^2$
4: **end for**
5: **for** $t = 1$ to N **do**
6: Compute the volatility by just taking the rate of change of $[\ln(S)]_t$. $V_t = \frac{d[\ln(S)]_t}{dt}$
7: **end for**
8: **for** $t = 1$ to N **do**
9: Compute the quadratic variation of $\sqrt{V}$. $[\sqrt{V}]_t = \lim_{\|P\|\to 0} \sum_{k=1}^{n} \left(\sqrt{V_{t_k}} - \sqrt{V_{t_{k-1}}}\right)^2$
10: **end for**
11: **for** $t = 1$ to N **do**
12: Calculate κ_t by $\kappa_t = 2\sqrt{\frac{d[\sqrt{V}]_t}{dt}}$.
13: **end for**
14: **for** $t = 1$ to N **do**
15: Calculate the quadratic co-variation of V and $\ln S_t$. $[V, \ln(S)]_t = \lim_{\|P\|\to 0} \sum_{k=1}^{n} \left(V_{t_k} - V_{t_{k-1}}\right)\left(\ln(S)_{t_k} - \ln(S)_{t_{k-1}}\right)$.
16: **end for**
17: **for** $t = 1$ to N **do**
18: **if** $V_t = 0$ **then**
19: Skip
20: **else**
21: Compute ρ by $\rho_t = \frac{1}{\kappa_t}\frac{d[V,\ln(S)]_t}{d[\ln(S)]_t} = \frac{1}{\kappa_t V_t}\frac{d[V,\ln(S)]_t}{dt}$.
22: **end if**
23: **end for**
24: **Output** κ_t and ρ_t

set of partition $\{0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t\}$ and $\|P\| = \max_k \{t_k - t_{k-1}\}$. Notice, we can get V from the quadratic variation $[\ln(S)]_t$ so we can observe variation and the stopping time τ_ϵ defined above. Clearly, we had to stop 21 at τ_ϵ or at least before V reached zero. There will be more critical use of τ_ϵ below.

We encountered a few problems while executing the algorithm on real market data. Some of them were, feller condition not satisfying, and getting extreme values for κ_t and ρ_t. We discuss these issues and ways to get around them in detail in later sections.

3.2 Estimating Drift parameters

We use the Branching particle filter introduced in [23] and extended in [26]. Estimating the drift parameters is finding the signal given the observation.

Let the signal be $X^i = (\mu_t^i, \nu_t^i, \varrho_t^i)$. Let $\{t_n\}_{n=0}^{\infty}$ be the chosen branch times with $t_0 = 0$, set $\Delta t_n = t_n - t_{n-1}$. Let u_n be the mark at which this n^{th} jump of Y occurs. The following branching Markov process $\left\{\mathbb{S}_n^N, n = 0, 1, \ldots\right\}$ approximates the unnormalized filter $\{\sigma_{t_n}, n = 0, 1, 2, \ldots\}$ in terms of the observations. $\mathbb{N}_n$ will denote the number of particles before branching at (discrete) time t_n.
We get the observations Y_t and $\widehat{Y}_t$ from, $Y_t = \int_0^t \gamma_s^{-1} \left(d\ln(S_s) - \frac{\rho_s}{\kappa_s} dV_s \right)$ and $\widehat{Y}_t = \int_0^t \frac{2}{\kappa_s} d\sqrt{V_s}$

We use stochastic approximation techniques such as *Euler–Maruyama* method to approximate the values of the observations. Once, we have the observations, we can proceed with the particle filter.

3.2.1 Wright-Fisher Model

In line 3 of **Algorithm 2**, we need to independently evolve the particles. We use the Wright-Fischer model for that. The Wright-Fischer model is a discrete-time Markov chain that describes the evolution of the particles. The state space of this Markov chain is the set of possible values of $X^i = (\mu_t^i, \nu_t^i, \varrho_t^i)$. Each generation, a collection of particles are sampled, with replacement from the current pool of particles at generation t to form a new pool of particles at generation $t+1$. This process describes the binomial sampling of the particles in each generation. If $X_t \in \{-n, \ldots, n\}$, then the probability transition matrix for the Markov chain is:

$$\mathbb{P}(X_{t+1} = k | X_t = j) = \binom{n}{k} p^k (1-p)^{n-k}$$

where $p = \frac{X_t + n}{2n}$.

Algorithm 2 Estimating the drift parameters using a branching particle filter

Initialize $\{X_0^i\}_{i=1}^N$ are independent samples of $X_0, \mathbb{N}_0 \doteq N, \mathbb{N}_n \doteq 0$ for $n = 1, 2, \ldots$ and $\mathbb{L}_0^i \doteq 1$ for $i \doteq 1, \ldots, N$.

for $n \doteq 1, 2, \ldots$ **do**

Evolve Independently:

$$Q\left(\widehat{X}^1_{[t_{n-1},t_n]} \in \Gamma^1, \ldots, \widehat{X}^{\mathbb{N}_{n-1}}_{[t_{n-1},t_n]} \in \Gamma^{\mathbb{N}_{n-1}} \mid \mathcal{F}^X_{t_n}\right) = \prod_{l=1}^{\mathbb{N}_{n-1}} K^X_{[0,\delta t_n]}\left(X^l_{t_n}, \Gamma^l\right) a.s.$$

for all $\Gamma^l \in \sigma\left(D_{E^k}\left[0, \delta t_n\right]\right)$, where K^X denotes the Markov kernel generated by the (well posed) martingale problem. In English, we evolve the particles independently according to the martingale problem over the interval.

Incorporate Observations through incremental weights:

$$\alpha_n\left(\widehat{X}^i\right) = \exp\left(\int_{t_{n-1}\wedge\tau_\epsilon}^{t_n\wedge\tau_\epsilon} \left\langle h\left(\mu^i_s, \nu^i_s, \varrho^i_s; V_s\right), d\left(Y_s, \widehat{Y}_s\right)\right\rangle - \frac{1}{2}\int_{t_{n-1}\wedge\tau_\epsilon}^{t_n\wedge\tau_\epsilon} \left|h\left(\mu^i_s, \nu^i_s, \varrho^i_s; V_s\right)\right|^2 ds\right)$$

(Here $V_s = \frac{d[\ln(S)]_s}{ds}$ must be supplied through the quadratic variation techniques used above.)

Weight Particles: $\widehat{\mathbb{L}}^i_n \doteq \alpha_n\left(\widehat{X}^i\right)\mathbb{L}^i_{n-1}$ for $i \doteq 1, 2, \ldots, \mathbb{N}_{n-1}$

Estimate: σ_n by: $\mathbb{S}^N_n \doteq \frac{1}{N}\sum_{i=1}^{\mathbb{N}_{n-1}} \widehat{\mathbb{L}}^i_n \delta_{\widehat{X}^i_n}$ and $\pi_n(f)$ by $\frac{\mathbb{S}^N_n(f)}{\mathbb{S}^N_n(1)}$.

Average Weight: $\mathbb{A}_n \doteq \mathbb{S}^N_n(1)$

for $i \doteq 1, 2, \ldots, \mathbb{N}_{n-1}$ **do**

if $\widehat{\mathbb{L}}^i_n \notin (a_n\mathbb{A}_n, b_n\mathbb{A}_n)$ **then**

- Offspring Number: $\mathbb{N}^i_n \doteq \left\lfloor \frac{\mathbb{L}^i_n}{\mathbb{A}_n} \right\rfloor + \rho^i_n$, with ρ^i_na $\left(\frac{\widehat{\mathbb{L}}^i_n}{\mathbb{A}_n} - \left\lfloor \frac{\widehat{L}^i_n}{\mathbb{A}_n} \right\rfloor\right)$-Bernoulli
- Resample: $\mathbb{L}^{\mathbb{N}_n+j}_n \doteq \mathbb{A}_n, X^{\mathbb{N}_n+j}_n \doteq \widehat{X}^i_n$, for $j \doteq 1, \ldots, \mathbb{N}^i_n$
- Add Offspring Number: $\mathbb{N}_n \doteq \mathbb{N}_n + \mathbb{N}^i_n$

else if $\hat{\mathbb{L}}^i_n \in (a_n\mathbb{A}_n, b_n\mathbb{A}_n)$ **then**

$$\mathbb{N}_n \doteq \mathbb{N}_n + 1, \mathbb{L}^{\mathbb{N}_n}_n \doteq \widehat{\mathbb{L}}^i_n, X^{\mathbb{N}_n}_n \doteq \widehat{X}^i_n$$

end if

end for

end for

Output μ_t, ν_t and ϱ_t

3.2.2 *Milstein* Approximation

Most of the stochastic integrals in this thesis, we have approximated using *Euler-Maruyama* methods. But the weight update equation is particularly complex and requires better approximation. Hence, we use Milstein approximation. It improves upon simpler methods like *Euler-Maruyama* by including an additional correction term. This correction term accounts for the curvature of the diffusion term, providing higher accuracy. For an SDE,

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t$$

The solution at the next time step X_{n+1} is approximated by:

$$X_{n+1} = X_n + \mu(X_n)\Delta t + \sigma(X_n)\Delta W_n + \frac{1}{2}\sigma(X_n)\frac{\partial \sigma(X_n)}{\partial X}((\Delta W_n)^2 - \Delta t)$$

- $\mu(X_n)\Delta t$ accounts for the drift over the time interval Δt.
- $\sigma(X_n)\Delta W_n$ accounts for the random diffusion.
- The correction term $\frac{1}{2}\sigma(X_n)\frac{\partial \sigma(X_n)}{\partial X}((\Delta W_n)^2 - \Delta t)$ adjusts for the non-linearity of the diffusion.

The weight update equation:

$$\alpha_n\left(\widehat{X}^i\right) = \exp\left(\int_{t_{n-1}\wedge\tau_\epsilon}^{t_n\wedge\tau_\epsilon} \left\langle h\left(\mu_s^i, \nu_s^i, \varrho_s^i; V_s\right), d\left(Y_s, \widehat{Y}_s\right)\right\rangle - \frac{1}{2}\int_{t_{n-1}\wedge\tau_\epsilon}^{t_n\wedge\tau_\epsilon} \left|h\left(\mu_s^i, \nu_s^i, \varrho_s^i; V_s\right)\right|^2 ds\right)$$

can be similarly approximated by

$$\alpha_n(\hat{X}^i) \approx \exp\left(h\left(\mu_{t_n}^i, \nu_{t_n}^i, \theta_{t_n}^i; V_{t_n}\right)\Delta Y_n + \frac{1}{2}\frac{\partial h}{\partial X}\left((\Delta Y_n)^2 - \Delta t\right) - \frac{1}{2}\left|h\left(\mu_{t_n}^i, \nu_{t_n}^i, \theta_{t_n}^i; V_{t_n}\right)\right|^2 \Delta t\right)$$

3.3 Simulated Heston

We simulate a Heston model with constant parameters.

$$\mu_t = 0.05, \nu_t = 0.04, \varrho_t = 2.0, \rho_t = -0.1, \kappa_t = 0.3$$

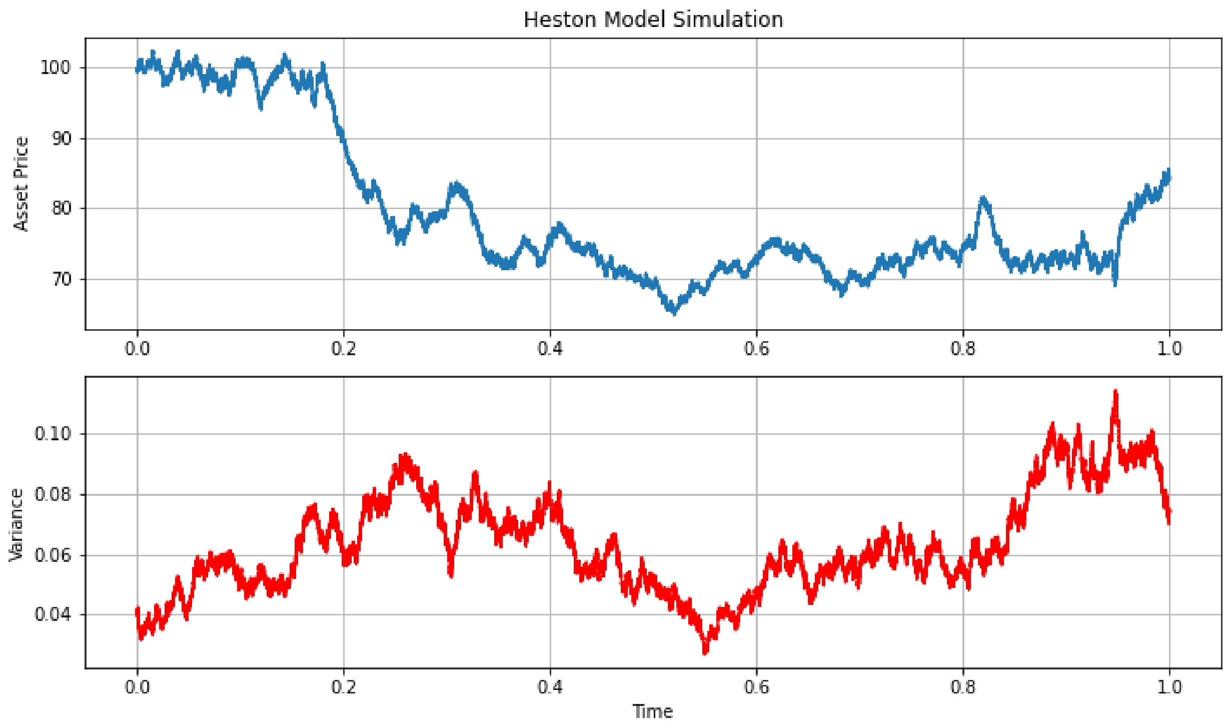

Figure 3.1: Simulated Heston Asset price and Variance

We initialize $(S_0, V_0) = (100, 0.04)$. We take the total time as one year and divide it into 1000000 timesteps. So $dt = 1/1000000$. In 3.1 we can see that the Asset price and variance

Algorithm 3 Heston Simulation

1: **Initialize** S_0 and V_0
2: **Generate** correlated Brownian motion for N steps.
3: **for** $i \in (1, N+1)$ **do**
4: Compute variance V_{t_i}
5: Compute Asset price S_{t_i}
6: **end for**
7: **Output** S_t and V_t

are showing opposite trends. This could probably be attributed to the negative correlation parameter. We apply our **Algorithm 1** and **2** on just the asset price time series and try to estimate the volatility and parameters.

We first calculate the quadratic variation of $\ln S_t$ and use it to estimate the volatility series V_t. The estimated variance is shown in figure3.2. The mean variance is 0.03662774829677189 which is around 10 percent off from the long-term mean. After we get the estimated vari-

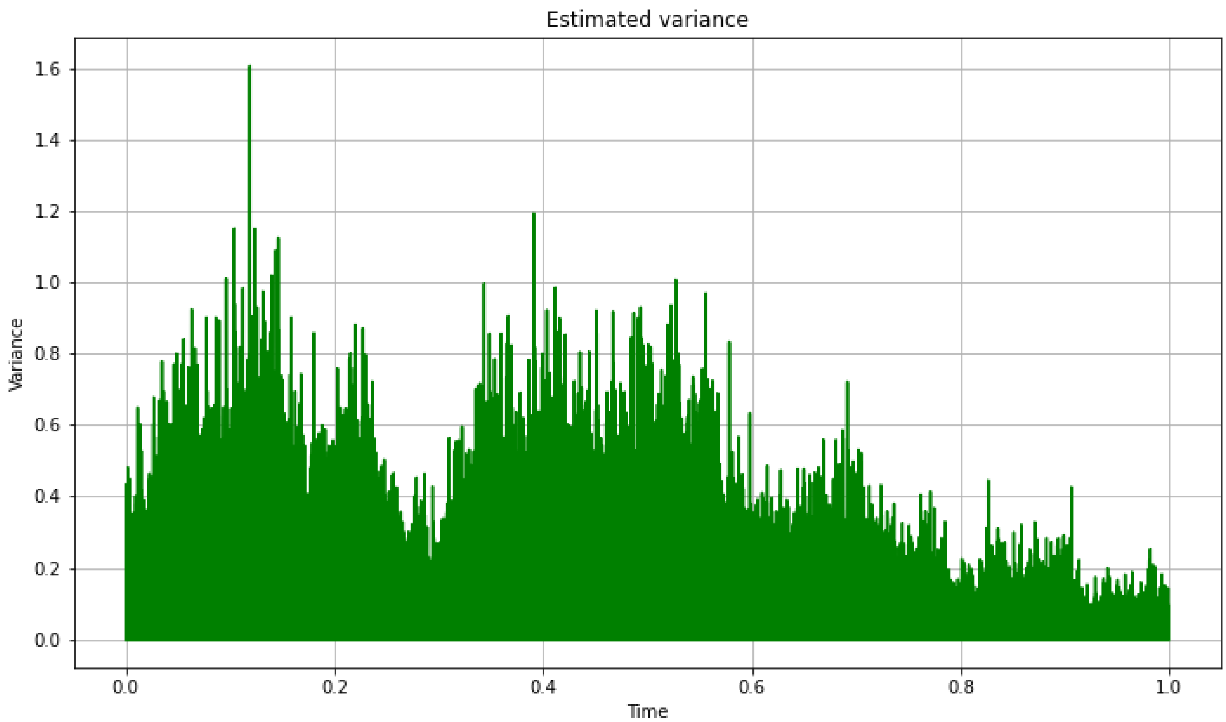

Figure 3.2: Estimated volatility of simulated Heston model

ance, we compute the quadratic variation of $\sqrt{V}]_t$. Then we get the estimated κ_t 3.3with mean 0.24586361003784957, which is approximately equal to the constant κ we used for simulation.

We the compute the quadratic co-variation of V and $\ln S$, which looks quite pretty if you see in figure3.4 From that, we compute ρ_t as seen in 3.5 with mean -0.06964 which is pretty off from the given rho. We then calculate the observations Y_t and $\widehat{Y}_t$ from $Y_t = \int_0^t \gamma_s^{-1} \left(d \ln (S_s) - \frac{\rho_s}{\kappa_s} dV_s \right)$ and $\widehat{Y}_t = \int_0^t \frac{2}{\kappa_s} d\sqrt{V_s}$ The estimated observations are shown in figure 3.6 After, we generate the observations, we create particles in the form of a grid for each of the parameters μ, ν, and ϱ. We evolve the particles using the Wright-Fischer model. After, weight updation and resampling, we get the estimated parameters μ with mean 0.04144154613681073, ν with mean 0.024106507382575217 and ϱ with mean 1.9894532915880456. 3.7 The branching particle filter is very computation heavy since there almost a 1000000 time steps, with 100000 particles and multiple for loops. This makes it very time consuming to run the filter. Hence, in the 3.7, we distribute the time into bins and reduce the number of particles significantly.

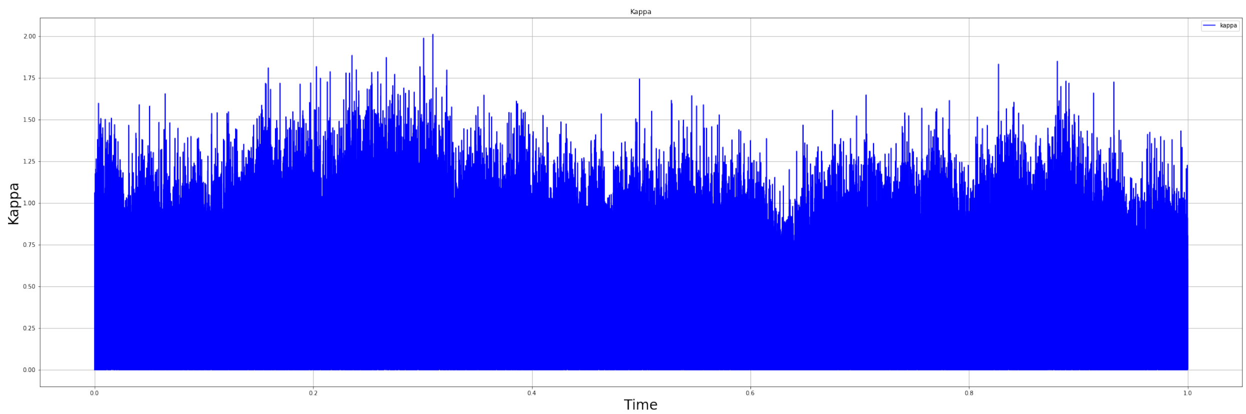

Figure 3.3: Estimated Kappa of simulated Heston

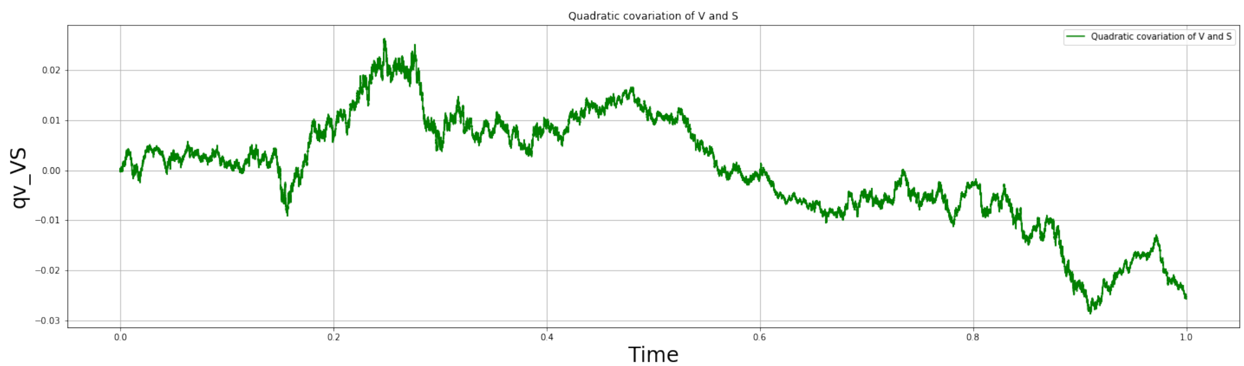

Figure 3.4: Quadratic co-variation of V and S

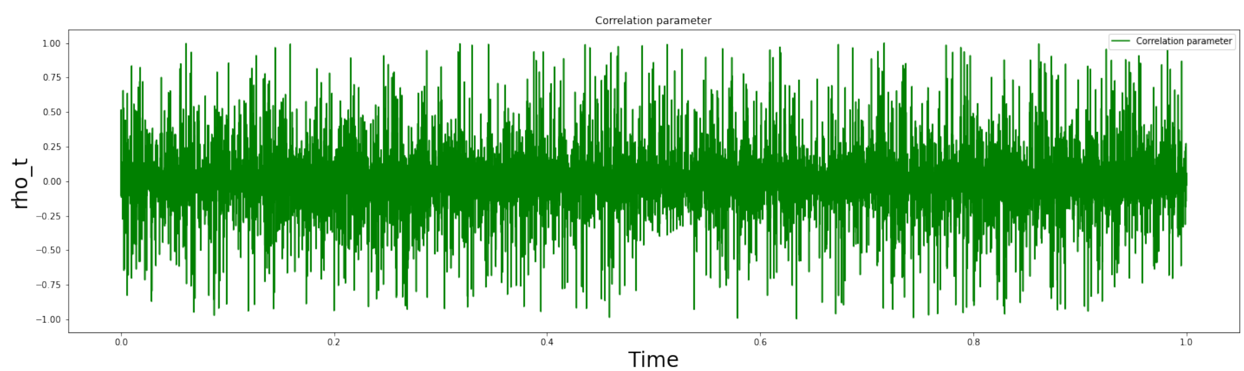

Figure 3.5: Estimated Rho

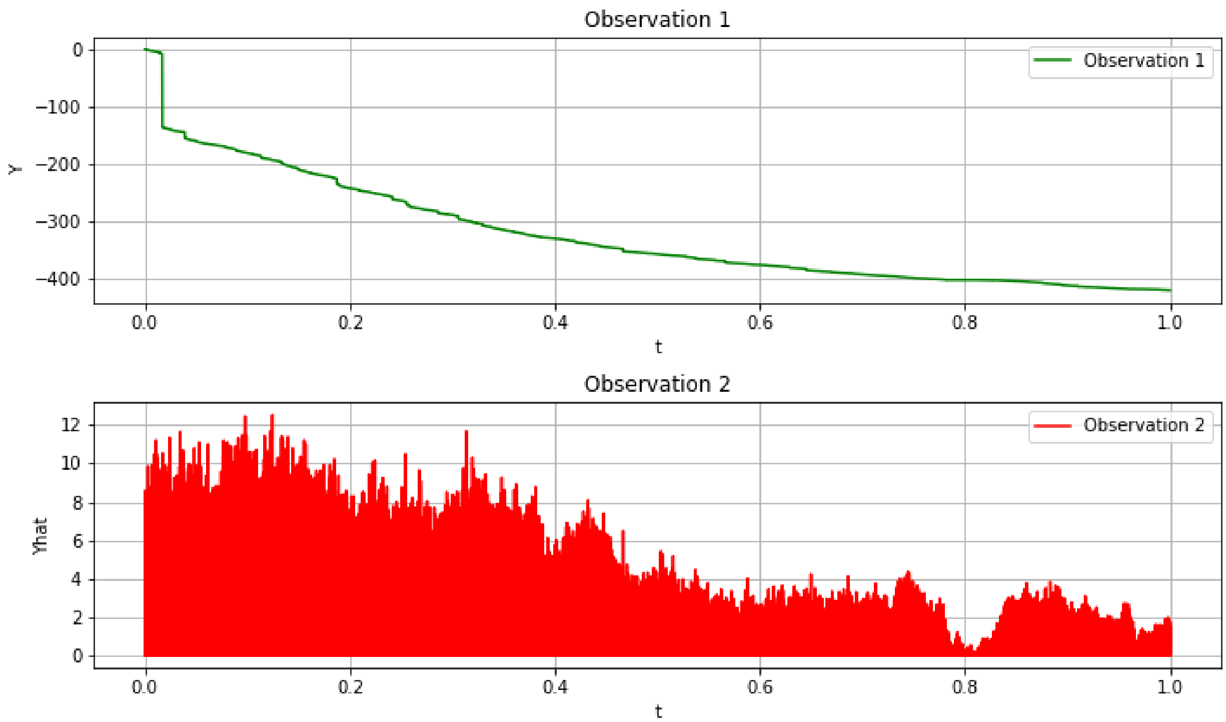

Figure 3.6: Observations constructed form estimated diffusion parameters

3.4 Real Market Data

We use tick-by-tick data of Apple quote price of one day. It contains data from thirteen different Exchanges. There are off-hour trades, so the ends need to be trimmed down. Fig(3.8) shows how the time gap at the ends are very steep while it is very gradual in the middle. Also, there are many NaN entries and certain outliers, which need to be removed. Once we are done with that, we start with the estimation of diffusion parameters.

We start with calculating the quadratic variation of $\ln S$ as usual. We take a spline fit of the quadratic variation and take it's derivative to compute the volatility. It gives a smooth volatility function as shown in , which is inherently wrong, but works as an indicator function for volatility. We then use this volatility to compute quadratic variation of $\sqrt{V_t}$. We take the derivative of that and compute the κ_t using $\kappa_t = 2\sqrt{\frac{d[\sqrt{V}]_t}{dt}}$. The estimated κ is shown in 3.10

After we have estimated κ we compute the quadratic co-variation of V and $\ln S$, which is used to estimate ρ. The estimated rho is shown in 3.11. As we can see, it is discontinued at two places, indicating volatility hitting zero.

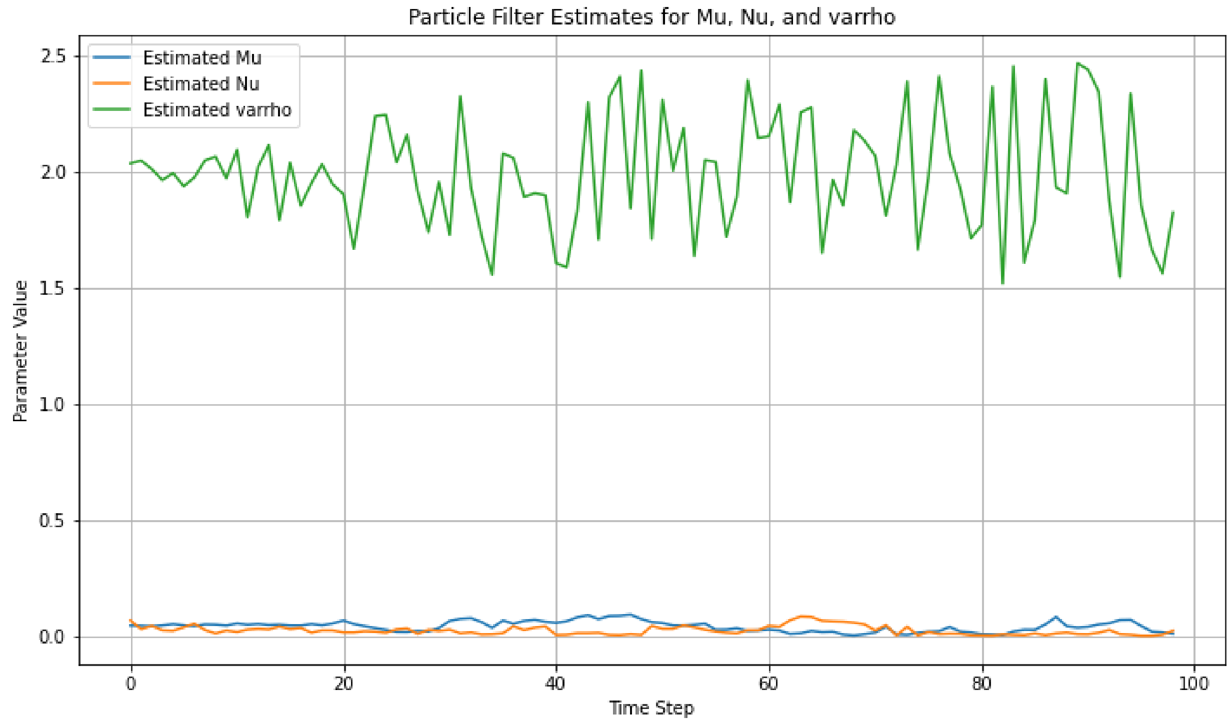

Figure 3.7: Estimated Drift parameters of the simulated Heston

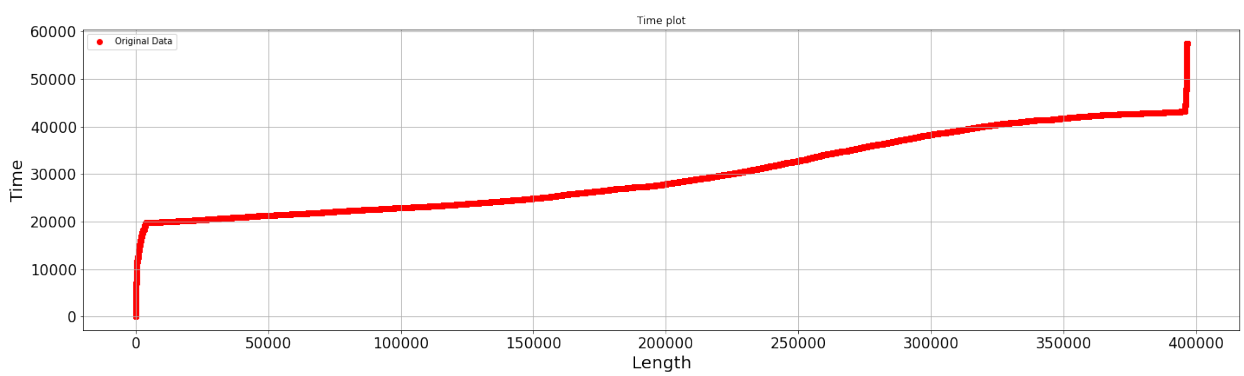

Figure 3.8: Time-plot showing how dt changes

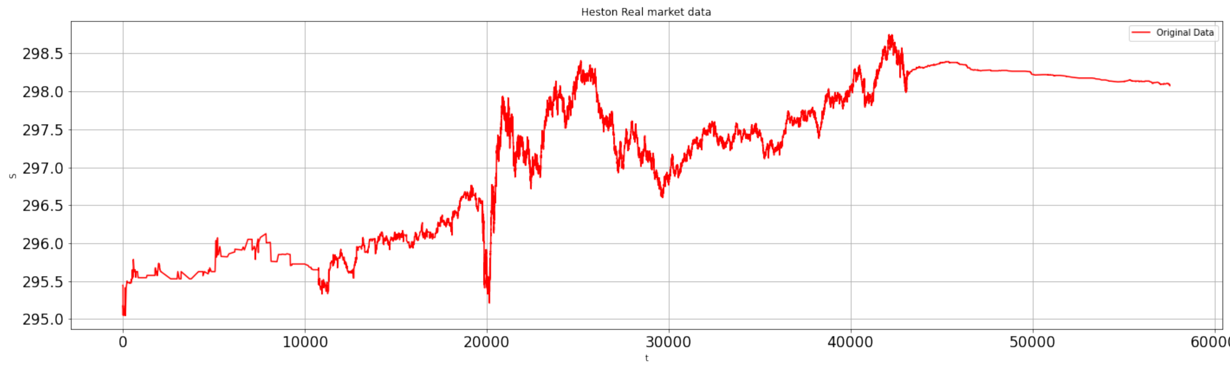

Figure 3.9: Heston market data

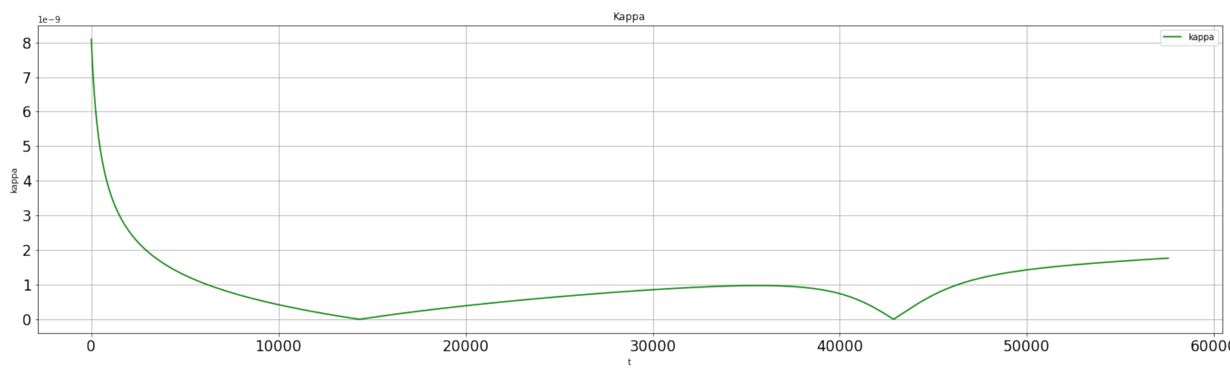

Figure 3.10: Estimated kappa of the real market data

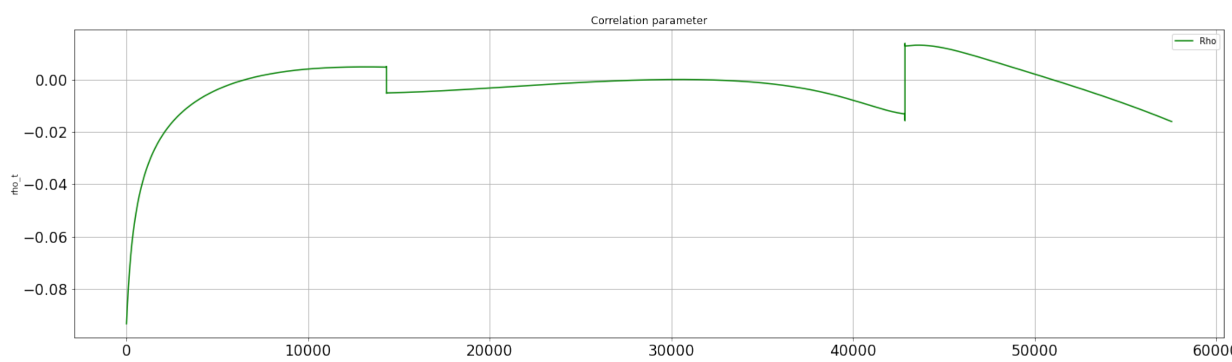

Figure 3.11: Estimated rho of real market data

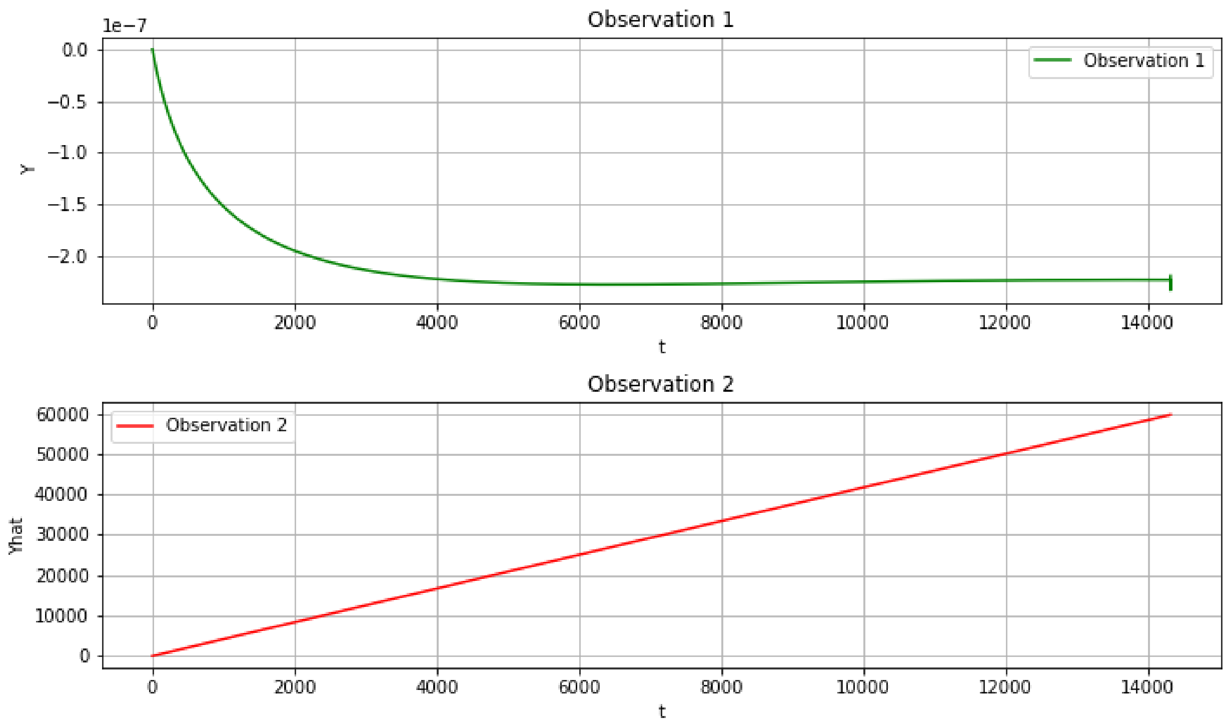

Figure 3.12: Observation for real market data

Similarly, we use the diffusion parameters to generate the observations. The 3.12 shows the generated observations.

Due to similar issues as mentioned in the last section, the estimated drift parameters are not rigorously computed.

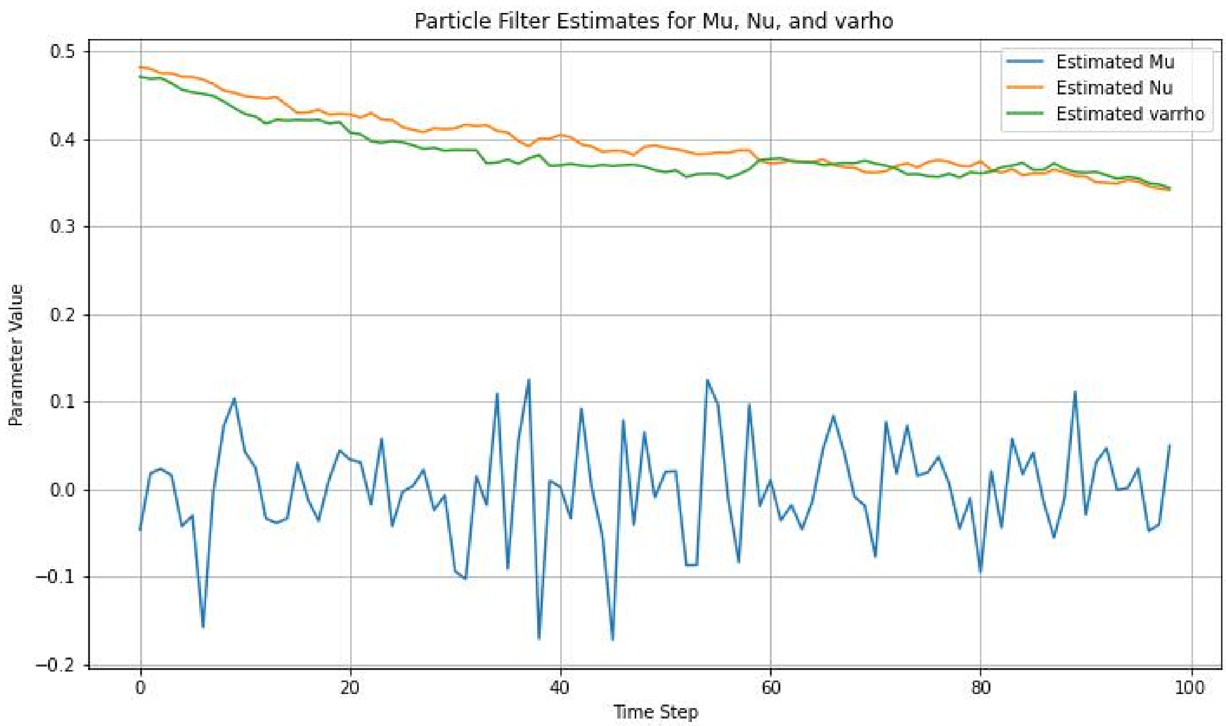

Figure 3.13: Estimated drift parameters of real market data

Chapter 4

Conclusion

In this thesis,we present a novel framework for calibrating the Heston model, based on co-variation techniques and the branching particle filter to estimate the diffusion and drift parameters in the stochastic volatility models. The Heston model is widely used for option pricing and risk assessment. This is because of the stochastic nature of volatility in the Heston Model. This a powerful alternative to constant-volatility models. However, its calibration, which requires aligning model parameters to observed market data, has remained a computationally demanding problem, particularly using high-frequency trading conditions where market dynamics are complex and data is very noisy.

To tackle these issues, we introduced an innovative calibration method built on co-variation techniques and particle filtering. We believe this approach will help in increasing the accuracy and efficiency of the calibration process . By integrating co-variation methods, which are based on the direct use of quadratic variations, we obtained estimates for the diffusion parameters (correlation and volatility of volatility).

In particular, the co-variation approach allows us to determine the diffusion parameters instantaneously using high-frequency data, thereby reducing the computational load and improving reactivity to changing market conditions. Traditional calibration methods, such as least squares fitting and Fourier transforms, often struggle to adapt to such high-frequency datasets, where complex price dynamics can interfere in the underlying parameter values. In contrast, our method's ability to capture these diffusion parameters in real time represents a meaningful advance, aligning model estimates more closely with actual market conditions and allowing for more accurate option pricing and risk assessment.

The drift parameters, which include the long-term mean of the asset price, the long-term mean of volatility, and the rate of volatility reversion, pose additional challenges for the calibration process. Traditional particle filters, while useful in estimating these latent variables, often suffer from degeneracy, where particles with low weights are prematurely discarded, leading to a loss of diversity in the particle population. This can significantly impact the accuracy of parameter estimates, especially in high-dimensional models like the Heston model. To mitigate this issue, we employed a branching particle filter, a variation that introduces a branching mechanism to regenerate particles, thereby eliminating the issues of degeneracy. This branching mechanism allows the filter to maintain a healthy number of particles.

The theoretical contributions of this thesis extend to the filtering framework itself. By constructing the model under a carefully designed reference probability, we were able to apply Bayes' factor techniques and a martingale problem approach to derive the filter's behavior under various market scenarios. This formulation enabled us to devise both unnormalized and normalized filters for the drift parameters, allowing for a clear separation of signal and noise in the estimation process. Additionally, the use of co-variation in the diffusion parameter estimation provides a novel approach that could be applied to other stochastic volatility models, potentially broadening the applicability of this work.

Our method's practical implications are equally significant. Accurate calibration of the Heston model enables financial practitioners to price options more reliably and to assess risk more accurately. This is particularly valuable in the context of algorithmic and high-frequency trading, where rapid changes in market volatility require models that can adapt in near real-time. By providing a more efficient and stable calibration framework, this work contributes to the field of financial modeling by offering a method that is both theoretically sound and practically feasible. The proposed approach can be implemented on modern computing platforms, making it accessible for real-world financial applications.

In conclusion, this thesis has made several original contributions to the field of stochastic volatility modeling:

1. **Introduction of Co-variation Techniques for Diffusion Parameter Estimation**: By using quadratic variations, we developed a method that allows for immediate estimation of diffusion parameters from high-frequency market data, thereby enhancing

the calibration process.

2. **Application of the Branching Particle Filter for Drift Parameter Estimation**: The branching mechanism within the particle filter mitigates degeneracy, improving accuracy in high-dimensional models and ensuring that the calibration process remains indifferent to market noise.

3. **Theoretical Advancements in Filtering Methods**: We formulated the model under a reference probability and derived both normalized and unnormalized filters for the drift parameters, thereby providing a clear separation of signal and noise and improving the precision of parameter estimation in stochastic volatility models.

4.0.1 Future Direction

The research presented in this thesis opens up several avenues for future investigation. First, we can use multiple different operators in the evolution of particles in the branching particle filter. We need to apply Bayes' theorem to distinguish between operators and find out which performs better. We also need to compare the model with other calibration techniques to get the actual idea of where it stands.

Secondly, while the current model has focused on a single-asset , future studies could extend this framework to multi-asset models, examining the co-variation structure between multiple assets and exploring how joint calibration of multi-asset systems could further enhance risk assessment and hedging strategies. Extending the co-variation and branching particle filter approach to multi-asset scenarios may present new computational challenges but could provide significant insights into cross-asset volatility dynamics.

Lastly, further optimization of the branching particle filter, particularly in terms of computational efficiency, could make the method even more suitable for real-time applications. Techniques such as parallel computing or GPU acceleration could be explored to speed up the filtering process, making it feasible for high-frequency trading applications where split-second decisions are critical.

Overall, this thesis has provided a foundation for a more adaptive and accurate calibration framework for stochastic volatility models, addressing key challenges in both parameter estimation and computational feasibility. By advancing the methods used to calibrate the

Heston model, this work contributes to the broader goal of developing robust financial models that can adapt to the complexities of modern markets.

Bibliography

[1] Yacine Aït-Sahalia and Robert Kimmel. Maximum likelihood estimation of stochastic volatility models. *Journal of Financial Economics*, 83(2):413–452, 2007.

[2] Ole E Barndorff-Nielsen and Neil Shephard. Estimating quadratic variation using realized variance. *Journal of Applied econometrics*, 17(5):457–477, 2002.

[3] Eric Benhamou, Emmanuel Gobet, and Mohammed Miri. Time dependent Heston model. *SIAM Journal on Financial Mathematics*, 1(1):289–325, April 2010.

[4] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81(3):637–654, 1973.

[5] Douglas Blount and Michael A. Kouritzin. On convergence determining and separating classes of functions. *Stochastic Processes and their Applications*, 120(10):1898–1907, September 2010.

[6] Michael J. Brennan and Eduardo S. Schwartz. Finite difference methods and jump processes arising in the pricing of contingent claims: A synthesis. *Journal of Financial and Quantitative Analysis*, 13(3):461–474, 1978.

[7] Rama Cont. Volatility clustering in financial markets: empirical facts and agent-based models. In *Long memory in economics*, pages 289–309. Springer, 2007.

[8] D. Crisan and A. Doucet. A survey of convergence results on particle filtering methods for practitioners. *IEEE Transactions on Signal Processing*, 50(3):736–746, 2002.

[9] Yiran Cui, Sebastian del Baño Rollin, and Guido Germano. Full and fast calibration of the heston stochastic volatility model. *European Journal of Operational Research*, 263(2):625–638, 2017.

[10] JL Doob. Conditional expectation; martingale theory. In *Measure Theory*, pages 179–204. Springer, 1994.

[11] Robin Dunn, Paloma Hauser, Tom Seibold, and Hugh Gong. Estimating option prices with heston ' s stochastic volatility model. 2014.

[12] William Feller. Two singular diffusion problems. *Annals of mathematics*, 54(1):173–182, 1951.

[13] Paolo Ghirardato. On independence for non-additive measures, with a fubini theorem. *journal of economic theory*, 73(2):261–291, 1997.

[14] Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993.

[15] Wenqing Hu. Itô's formula, the stochastic exponential, and change of measure on general time scales. *Abstract and Applied Analysis*, 2017(1):9140138, 2017.

[16] John Hull and Alan White. The pricing of options on assets with stochastic volatilities. *The journal of finance*, 42(2):281–300, 1987.

[17] Martino Grasselli JosÉ Da Fonseca and Claudio Tebaldi. A multifactor volatility heston model. *Quantitative Finance*, 8(6):591–604, 2008.

[18] Rajeeva L Karandikar and BV Rao. On quadratic variation of martingales. *Proceedings-Mathematical Sciences*, 124:457–469, 2014.

[19] Ioannis Karatzas and Steven Shreve. Methods of mathematical finance / i. karatzas, s.e. shreve. *Journal of the American Statistical Association*, 95, 06 2000.

[20] Ioannis Karatzas, Steven E Shreve, Ioannis Karatzas, and Steven E Shreve. Martingales, stopping times, and filtrations. *Brownian Motion and Stochastic Calculus*, pages 1–46, 1988.

[21] Ruben Klein and Evarist Giné. On quadratic variation of processes with gaussian increments. *The Annals of Probability*, pages 716–721, 1975.

[22] Wolfgang H. Kliemann, Giorgio Koch, and Federico Marchetti. On the unnormalized solution of the filtering problem with counting process observations. *IEEE Trans. Inf. Theory*, 36:1415–1425, 1990.

[23] Michael A Kouritzin. Convergence rates for residual branching particle filters. *Journal of Mathematical Analysis and Applications*, 449(2):1053–1093, 2017.

[24] Michael A. Kouritzin. Explicit heston solutions and stochastic approximation for path-dependent option pricing. *International Journal of Theoretical and Applied Finance*, 21(01):1850006, February 2018.

[25] Michael A. Kouritzin and Hongwei Long. On extending classical filtering equations. *Statistics Probability Letters*, 78(18):3195–3202, 2008.

[26] Michael A. Kouritzin and Anne MacKay. Branching particle pricers with heston examples, 2019.

[27] Michael A Kouritzin and Yong Zeng. Bayesian model selection via filtering for a class of micro-movement models of asset price. *International Journal of Theoretical and Applied Finance*, 8(01):97–121, 2005.

[28] Hiroshi Kunita and Shinzo Watanabe. On square integrable martingales. *Nagoya Mathematical Journal*, 30:209–245, 1967.

[29] Francis Longstaff and Eduardo Schwartz. Valuing american options by simulation: A simple least-squares approach. *Review of Financial Studies*, 14:113–47, 02 2001.

[30] Andrea Pascucci. *Itô calculus*, pages 167–201. Springer Milan, Milano, 2011.

[31] Daniel Revuz, Marc Yor, Daniel Revuz, and Marc Yor. Bessel processes and ray-knight theorems. *Continuous Martingales and Brownian Motion*, pages 409–434, 1991.

[32] Johannes Ruf. A new proof for the conditions of novikov and kazamaki. *Stochastic Processes and their Applications*, 123(2):404–421, February 2013.

[33] Gleb Sandmann and Siem Jan Koopman. Estimation of stochastic volatility models via monte carlo maximum likelihood. *Journal of Econometrics*, 87(2):271–301, 1998.

[34] Louis O Scott. Option pricing when the variance changes randomly: Theory, estimation, and an application. *Journal of Financial and Quantitative analysis*, 22(4):419–438, 1987.

[35] I.S. Stamatiou. A boundary preserving numerical scheme for the wright–fisher model. *Journal of Computational and Applied Mathematics*, 328:132–150, 2018.

[36] Elias M Stein and Jeremy C Stein. Stock price distributions with stochastic volatility: an analytic approach. *The review of financial studies*, 4(4):727–752, 1991.

[37] James V Stone. Bayes' rule: a tutorial introduction to bayesian analysis. 2013.

[38] JOSEF TEICHMANN. Foundations of martingale theory and stochastic calculus from a finance perspective. *Lecture Notes (ETH Zurich)*, 2017.

[39] Jie Xiong. *An introduction to stochastic filtering theory*, volume 18. OUP Oxford, 2008.

[40] Shing-Tung Yau and S.S.-T. Yau. Existence and uniqueness of solutions for duncan-mortensen-zakai equations. In *Proceedings of the 44th IEEE Conference on Decision and Control*, pages 536–541, 2005.