Delft University of Technology
Faculty Electrical Engineering, Mathematics and Computer Science
Delft Institute of Applied Mathematics

Numerical Solutions for the Stochastic Local Volatility Model

A thesis submitted to the
Delft Institute of Applied Mathematics
in partial fulfillment of the requirements

for the degree

MASTER OF SCIENCE
in
APPLIED MATHEMATICS

by

ROEL WILLEM BERTUS VAN DER WEIJST

Zurich, Switzerland

MSc thesis APPLIED MATHEMATICS

"Numerical Solutions for the Stochastic Local Volatility Model"

R.W.B. VAN DER WEIJST

Delft University of Technology

Responsible Professor

Prof. Dr. Ir. C.W. Oosterlee

Other members of the thesis committee

Dr. P. Cirillo

Dr. R.J. Fokkink

August, 2017

Zurich

Preface

Before you lies the thesis "Numerical Solutions for the Stochastic Local Volatility Model". It has been written to fulfill the graduation requirements for the degree Master of Science in Applied Mathematics at Delft University of Technology. After obtaining a bachelor in mathematics and a MSc in Economics, I wanted to expand my knowledge in financial mathematics. Due to my interest in financial markets and my job as a quantitative analyst at Derivative Partners AG, I decided to do research on hybrid LSV models.

I would like to take this opportunity to thank Prof. Dr. Ir. C.W. Oosterlee, professor at the Numerical Analysis group of Delft Institute of Applied Mathematics, and the supervisor of this thesis for his cooperation. Furthermore, I would like to thank Prof. dr. F.H.J. Redig for getting me in touch with ETH Zurich and Prof. Dr. J. Teichmann from ETH Zurich for our discussions that gave me valuable input for new ideas. Finally, I would like to thank my family for their motivational support.

R.W.B. van der Weijst

Zurich, Switzerland,

Abbreviations

CIR	Cox–Ingersoll–Ross
LV/LVM	Local Volatility Model
LSV	Local Stochastic Volatility Model
MC	Monte Carlo
MLMC	Multilevel Monte Carlo
PDE	Partial Differential Equation
r.v.	Random Variable
SDE	Stochastic Differential Equation
SV/SVM	Stochastic Volatility
TTM	Time to Maturity
QE	Quadratic Exponential

List of Symbols

Unless specified differently, these commonly used symbols and letters have the following meaning throughout this thesis.

Δ	Time step in the discretization
δ	The Greek delta $\frac{\partial C}{\partial S}$
$D(t,T)$	The zero-coupon bond price, contracted at time t with maturity T
ε	Volatility of the volatility in the Heston model
f_t	Price of the forward at time t
γ	The Greek gamma: $\frac{\partial^2 C}{\partial S^2}$
k	Strike of a Forward Start Option ($k = K + 1$)
K	Strike
κ	Rate of return in the Heston model
l	Level in the MLMC method
μ	Drift function
$\mathbb{P}$	Probability measure
$\mathbb{Q}$	Risk-neutral measure
r	Risk-free interest rate
ρ	(Greek) The Greek rho: $\frac{\partial C}{\partial r}$ (Heston) Correlation factor in the Heston model
S	Process of the underlying in the Heston model
σ	Volatility function
t	Time
θ	(Heston) Mean reversion of the volatility in the Heston model (Greek) The Greek theta: $\frac{\partial C}{\partial t}$
v	(Heston) Variance process (Greek) The Greek vega: $\frac{\partial C}{\partial \sigma}$
W	Wiener process
X	Martingale process

Abstract

This thesis is about pricing European options and forward start options under the Heston LSV model. The impact of conditionally calibrating the Heston parameters on the satisfaction of the Feller condition and thereafter correcting with a local volatility surface is investigated here. The results show that this approach is computationally time efficient and accurate. Efficient numerical approaches for this LSV model, such as the multilevel Monte Carlo method, are also investigated. Furthermore, a comparison of several discretizations schemes for the SV part have been conducted. For the calibration of the local volatility surface, the efficiency of the Particle method and the Bin method are compared. An alternative numerical approach to this problem which builds on these two methods is developed and tested.

Keywords: Local Stochastic Volatility Model (LSV), Stochastic Volatility Model (SV), Local Volatility Model (LV), Heston, Feller Condition, Discretization, Euler, Milstein, Quadratic Exponential (QE) Particle, Bins, Monte Carlo, Forward Start Options, Dupire's Model, Calibration, multilevel Monte Carlo (MLMC).

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis focuses on the pricing of European options and forward start options with local stochastic volatility models by applying numerical techniques. An option is a type of derivative: a contract that derives its value from the performance of an underlying asset. This underlying asset can be an asset that is traded at an exchange e.g., a stock. To determine the value of an option, the price process of the underlying asset price can be modelled. Several methods exist for the price determination in which the price process of the underlying is described by a stochastic differential equation. An option is the right, to buy or sell a predetermined underlying at a predetermined price (the strike) at a predetermined time (the maturity time). A commonly traded option is a European call option. This option gives the buyer the right, not the obligation, to buy the underlying asset at the strike price at the time of maturity. Therefore, the value V of this option at time of maturity is $V(S_T, T) = \max(S_T - K, 0)$, where K is the strike, T the maturity, and S_T the asset (stock) price at time T.

Research has been conducted to find the stochastic differential equation that describes the price process has been conducted. Black-Scholes [4] developed a stochastic differential equation that describes this process in 1973, the geometric Brownian motion. In the Black-Scholes model, the volatility, a statistical measure of the dispersion of returns for a given security or market index is constant. However, consider for example that the Chicago Board Options Exchange Volatility Index (known by its ticker symbol VIX), an index for the volatility, fluctuates. Market data implies that the expected volatility is usually higher for options with a strike that differs much from the price of the underlying at the time of writing the option than it is for an option with a strike close to the initial price of the underlying. As a consequence, new models were developed such as the local volatility model (LV) and the stochastic volatility model (SV) to reproduce this observed so-called volatility smile.

The local volatility model as developed by Dupire [12] and Derman & Kani [11] uses the same dynamics as the Black-Scholes model. The difference is that the volatility term is not a constant, but rather a deterministic function that depends on the time and value of the underlying. The results for pricing European options are satisfying. However, due to the simple dynamics of the Black-Scholes model, it leads to a flattened implied forward volatility when pricing forward start options [31]. This explains the inability of this model to deliver prices for complex structured products close to market prices.

Heston [22] developed a stochastic volatility model in 1993. Calibration of the Heston parameters can lead to parameters that make the model reproduce the market prices of European options quite well with various strikes and maturities. With this model, the smile that is seen in the implied volatility function can be reproduced relatively well. However, this model also has certain limitations. The accuracy of pricing short-maturity options is often unsatisfactory [13]. Also, the parameters that are calibrated using market data often make it possible for the variance process to become zero. In practice, this would mean the price of the underlying could no longer be a stochastic process anymore and would become deterministic at some moment in time. However, the process is strongly reflecting and will not stay at zero. It is time consuming to work with parameter sets in which the variance process can become zero,

as several simulations require more computational time to reach the same accuracy. Another issue arises due to the discretization of the continuous process of the underlying. The value of an option can be determined with Monte Carlo simulation (MC). A sufficient number of paths of the underlying asset are then simulated to determine the value. However, a discretization scheme has to be used to simulate the continuous paths and for this set of parameters, the volatility can even become negative. This can be prevented by taking the absolute value of the variance process, or the maximum of zero and the variance process. This of course impacts the dynamics of the underlying price process and therefore the value of the option.

Another model which became popular by Lipton [24] is the Local Stochastic Volatility Model (LSV). This model combines the local volatility model and the stochastic volatility model. Here, the volatility in the stochastic volatility model can be adjusted with a local volatility function. This keeps the dynamics of the model close to the stochastic volatility model while allowing a correction to mimic the implied volatility as observed in the market. Furthermore, this makes it easier to calibrate the model.

The dynamics of the calibrated LSV model consist of a non-linear McKean-Vlasov stochastic differential equation in which the non-linearity is due to the volatility, which depends on the marginal distribution of the process. Therefore, this stochastic differential equation is associated with a non-linear Fokker-Planck partial differential equation and is not easily solvable. Several techniques to solve this are compared, and an alternative technique is developed to see if the efficiency can be improved.

Engelmann [13] has used a finite-volume scheme to solve the Kolmogorov forward PDE. Piterbarg [29] uses a Markovian projection closed-form approximation instead of a direct solution to the LSV model. Henry-Labordère [21] also uses Markovian techniques in the particle algorithm to solve several conditional expectations that arise in the calibration process. These projection techniques preserve marginal densities. However, they do not preserve all marginal distributions of higher order.

To avoid the possibility of a negative volatility in the Heston model, the parameters can be conditionally calibrated on the non-existence of negative volatility. This is possible due the existence of a parameter set for the Heston model that does not allow the volatility to reach zero. However, this calibration does not correctly mimic the implied volatility. This thesis investigates the Heston LSV model to answer the following question: How can the implied volatility be efficiently reconstructed for a badly calibrated Heston model? Furthermore, (new) efficient methods for pricing forward start options that cannot be priced well with local volatility models are investigated.

The outline of this thesis is as follows. Chapter 1 contains the introduction. Chapter 2 discusses the preliminaries. This chapter introduces the mathematical theory of a stochastic price process. Chapter 3 begins by defining a European call option. It then states the theorems of asset pricing and explains the conditions that need to be satisfied in a pricing model to avoid arbitrage. It elaborates on the stochastic price process developed by Black-Scholes. Then it explains the Heston stochastic volatility model, Dupire's local volatility model, and finally the local stochastic volatility model.

Chapter 4 elaborates more on the Heston model and offers a comparison of global and local calibration. It explains the COS method to price options. Furthermore, it elaborates on the Monte Carlo method which can be applied to price a broad range of derivatives. Three discretization schemes are explained in Chapter 4, and a comparison between the Euler, Milstein, and Quadratic Exponential scheme is made. It explains the multilevel Monte Carlo method (MLMC) and it explains how one can reduce the computational costs for pricing options with MC simulations by discretization on a fine and coarse mesh. Finally, forward start options are introduced and methods to price these options are discussed.

Chapter 5 elaborates on the computation of Dupire's local volatility surface. It describes how one can interpolate between grid points values of the volatility to obtain a smooth surfaces. Therefore, it explains monotonic cubic splines. Furthermore, the theory of the local stochastic volatility model is discussed in more detail. Several numerical approaches to solve the local volatility function are discussed. The model of Henry-Labordère (particle) and Van der Stoep (bins) are explained in more detail. An alternative model which is in fact a model derived from the two before-mentioned models is discussed. A comparison of the alternative, the particle, and the bin models is made ,and numerical results of this comparison are

discussed. The extrapolation of the local volatility surface is also discussed.

Chapter 6 begins with the results of the comparison of efficient numerical simulation techniques. Simulations are compared for the Heston model with the Feller condition satisfied and with the Feller condition not satisfied. An experiment to determine the proportion of simulations hitting the zero variance is presented and the results are discussed. Forward start option pricing with local volatility and local stochastic volatility are compared for both the case with the Feller condition satisfied and the case where it is not satisfied. This is done by calibrating conditionally on the fact that the Feller condition is satisfied, and with synthetic data. At the end of Chapter 6, MC simulations for the LSV model for the price determination with market data of forward start options (European call option prices) are presented.

Finally, Chapter 7 presents the conclusion and offers ideas for further research.

Chapter 2

Preliminaries

This chapter contains all important theorems required to build a stochastic price process that can be used to describe the dynamics of the price of an asset over time. First, important definitions in the field of stochastic calculus and financial mathematics are given.

Definition 2.0.1 (Probability Measure)**.** *A real-valued function $\mathbb{P}$ on the collection of events of a random experiment, which assigns probabilities to events, is a probability measure if the following holds,*

- $\mathbb{P}(\mathcal{F}_i) \geq 0$*, for all events* $\mathcal{F}_i$*,*
- $\mathbb{P}(\Omega) = 1$*, with* Ω*, the sample space,*
- *It is countable additive, i.e.,* $\mathbb{P}(\cup_{i=1}^{\infty}\mathcal{F}_i) = \sum_{i=1}^{\infty}\mathbb{P}(\mathcal{F}_i)$ *holds if* $\mathcal{F}_i$*, for* i *in* I*, is a countable collection of pairwise disjoint events.*

Two important probability measures for option pricing are the real-world and the risk-neutral measure.

Definition 2.0.2 (Risk-neutral Measure)**.** *The risk-neutral measure, or martingale measure, $\mathbb{Q}$ is a probability measure under which the value of an asset is equal to its discounted expected value under this measure. It has the money market account as a numeraire.*

Besides the constructed risk-neutral measure, the real-world measure exists. In the real world, investors demand risk premia. As a consequence, assets have different expected rates of return based on the riskiness of these assets. Risk-neutral pricing is computationally powerful and convenient, as it avoids the determination of the risk premia. Therefore, this real-world measure is not used for calculations in this thesis.

Definition 2.0.3 (Filtration)**.** *A filtration $\mathcal{F}$ is a subset of an algebraic structure A. In this context, $\mathcal{F}_t$ represents all historical information up to time t, and no information about the future stochastic process. The filtration can be indexed, so $A_a \subset A_b$ holds for $t_a \leq t_b$.*

Now, a probability space can be defined.

Definition 2.0.4 (Probability Space)**.** *A probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability triple consisting of a probability measure $\mathbb{P}$, the sample space Ω, which is the non-empty set of all possible outcomes of an experiment, and the set of events $\mathcal{F}$, the filtration. Each event in this filtration is a subset of Ω.*

To define the stochastic price process to mimic the dynamics of the price process of an exchange-listed asset over time, first the definitions of an adapted process, a martingale, and a Wiener process are presented.

Definition 2.0.5 (Adapted Process). *A process X for which the random variable $X_i : \Omega \to A$ is a $(\mathcal{F}_i, \sum)$ measurable function is said to be adapted to the filtration $\mathcal{F}$. Here $(S, \sum)$ is the (measurable) state space, and $(\mathcal{F}_i)_{i \in I} = \mathcal{F}$*

Definition 2.0.6 (Discrete Martingale). *A stochastic process $X_1, X_2, ...$ in discrete-time is a Martingale with respect to filtration $\mathcal{F}$ and probability measure $\mathbb{P}$ if the following conditions are satisfied:*

- $\mathbb{E}^{\mathbb{P}}(|X_n|) < \infty$,
- $\mathbb{E}^{\mathbb{P}}(X_n | X_1, ..., X_{n-1}) = X_{n-1}$,
- *X is adapted to its filtration, $\mathcal{F}$.*

Definition 2.0.7 (Continuous Martingale). *A stochastic process $X_1, X_2, ...$ in continuous time is a Martingale with respect to filtration $\mathcal{F}$ and probability measure $\mathbb{P}$ if the following conditions are satisfied,*

- $\mathbb{E}^{\mathbb{P}}(|X_t|) < \infty$,
- $\mathbb{E}^{\mathbb{P}}(X_t | X_s) = X_s \quad \forall s < t$,
- *X is adapted to its filtration $\mathcal{F}$.*

Definition 2.0.8 (Wiener Process). *A Wiener process W is a Lévy process that satisfies,*

- $W_0 = 0$,
- *$W_t - W_s$ is independent of W_x with $x < s < t$,*
- $W_t - W_s \sim \mathcal{N}(0, (t - s))$,
- *W is continuous with probability 1.*

Itô Process

The dynamics of a price process can be defined as an Itô process, which is defined as follows:

Definition 2.0.9 (Itô Process). *An Itô process is an adapted stochastic process of the form,*

$$X_t = X_0 + \int_0^t \sigma(X_s, s) dW_s + \int_0^t \mu(X_s, s) ds, \tag{2.1}$$

with W_s a Wiener process, and $\sigma(X_s, s)$ is a W_s-integrable function. Furthermore, $\mu(X_s, s)$ is a Lebesgue integrable function and is predictable, so

$$\int_0^t (\sigma^2(X_s, s) + |\mu(X_s, s)|) ds < \infty. \tag{2.2}$$

Derivatives have a value dependent on the price process of the underlying. Itô's lemma can help to determine the value of a derivative with an underlying price process that is an Itô process.

Lemma 2.0.1 (Itô's Lemma). *For the Itô process*

$$dX_t = \mu(X_t, t) dt + \sigma(X_t, t) dW_t, \tag{2.3}$$

and any twice differentiable function $V(x, t)$, with $x, t \in \mathbb{R}$, the following equation holds

$$dV = \left(\frac{\partial V}{\partial t} + \mu(X_t, t) \frac{\partial V}{\partial x} + \frac{\sigma^2(X_t, t)}{2} \frac{\partial^2 V}{\partial x^2} \right) dt + \sigma(X_t, t) \frac{\partial V}{\partial x} dW_t. \tag{2.4}$$

This is then again an Itô process. The lemma holds as well for multidimensional processes.

Proof. For a detailed proof, see [34]. □

Itô's lemma can be used to price derivatives. The function $V(x,t)$ represents the value of the derivative, and X_t is then the price process of the underlying, and it is therefore denoted as S_t (stock).

The evolution over time of the density function of an Itô price process is described by the Fokker-Planck PDE.

Theorem 2.0.1 (Fokker-Planck Equation)**.** *For $t \in [0,T]$, the transition density function $p\colon S_t, t \mapsto \mathbb{R}$ associated with the price process*

$$\frac{dS_t}{S_t} = rdt + \sigma(S_t,t)dW_t, \tag{2.5}$$

satisfies the following equation

$$\frac{\partial}{\partial t}p(S_t,t) + r\frac{\partial}{\partial S}p(S_t,t) - \frac{1}{2}\frac{\partial^2}{\partial S^2}(\sigma^2(S_t,t)p(S_t,t)). \tag{2.6}$$

This last equation is known as the Fokker-Planck PDE.

Proof. For a detailed proof, see [34]. □

This Fokker-Planck PDE is also known as the Kolmogorov forward equation. The function p describes the probability density of the price of the underlying.

The Feynman-Kac Theorem defines the value of a derivative V with an Itô process as the underlying price process.

Theorem 2.0.2 (Feynman-Kac)**.** *Define the stochastic process S_t by*

$$\frac{dS_t}{S_t} = rdt + \sigma(S_t,t)dW_t. \tag{2.7}$$

Let $V\colon S_t, t \to \mathbb{R}$ be a sufficiently differentiable function of time t and underlying price S_t. If $V(S_t,t)$ satisfies the equations,

$$\frac{\partial V}{\partial t} + rS_t\frac{\partial V}{\partial S} + \frac{1}{2}\sigma^2(S_t,t)S_t^2\frac{\partial^2 V}{\partial S^2} - rV = 0, \tag{2.8}$$

$$V(S_T,T) = g(S_T,T), \tag{2.9}$$

with general interest rate r, volatility function $\sigma(S_t,t)$, and with given function $g(S_T,T)$, then the solution $V(S_t,t) \quad \forall t \in [0,T]$ is given by:

$$V(S_t,t) = e^{-(T-t)}\mathbb{E}[g(S_T,T)|\mathcal{F}]. \tag{2.10}$$

Here $e^{-(T-t)}$ is the discount factor.

Proof. For a detailed proof, see [10]. □

Here, V can be seen as the value of the derivative, and S is the price of the underlying e.g., stock.

Definition 2.0.10 (McKean-Vlasov SDE)**.** *A stochastic differential equation in which the volatility and or drift depend on the probability distribution $\mathbb{P}_T$ of X_t is called a McKean-Vlasov SDE. In formula this is described as follows:*

$$dX_t = \mu^{\mathbb{Q}}(X_t,t)dt + \sigma^{\mathbb{Q}}(X_t,t)dW_t. \tag{2.11}$$

Here, X_t can be an n-dimensional stochastic process, and $\mathbb{Q}$ is associated with the function μ and σ denotes the dependency of these functions on $\mathbb{Q}$.

Applying the Fokker-Planck Equation (2.6) on this McKean-Vlasov SDE gives,

$$-\partial_t p - \sum_{i=1}^{n}\partial_i(\mu^{\mathbb{Q},i}(x,t)p(x,t)) + \frac{1}{2}\sum_{i,j=1}^{n}\partial_{ij}\left[\sum_{k=1}^{d}\sigma_k^{\mathbb{Q},i}(x,t)\sigma_k^{\mathbb{Q},j}(x,t)p(x,t)\right] = 0. \tag{2.12}$$

Radon-Nikodym Derivative

Sometimes it is desirable to change the measure of the price process. For example when the underlying price process has a drift term, and a process without drift would be easier for calculations. The Radon-Nikodym derivative $\frac{d\mathbb{P}}{d\mathbb{Q}}$ can be used for this purpose.

The expectation of a continuous function $f(Z)$ on stochastic variable Z defined on the probability space $(\Omega, \mathcal{F}, \mathbb{Q})$ with probability density function p, gives the following expectation with respect to the probability measure $\mathbb{Q}$,

$$\mathbb{E}^{\mathbb{Q}}[f(Z)] = \int_{-\infty}^{\infty} f(z)p(z)dz, \tag{2.13}$$

with distribution function $\mathbb{Q}(Z \leq z) = \int_{-\infty}^{z} p(x)dx$ for $z \in \mathbb{R}$. A new probability measure can be then defined as follows:

$$\mathbb{P}(A) = \mathbb{E}^{\mathbb{Q}}\left[1_A \frac{f(Z)}{p(Z)}\right] = \mathbb{E}^{\mathbb{Q}}\left[1_A \frac{d\mathbb{P}}{d\mathbb{Q}}\right] = \int_{k\in A} \frac{f(Z(k))}{p(Z(k))} d\mathbb{Q}(k), \tag{2.14}$$

where if $f(a) = 0$, it follows directly that $p(a) = 0$. Now, it can be derived that $\mathbb{P}$ is a probability measure on $(\Omega, \mathcal{F})$. Therefore,

$$\mathbb{P}(Z \leq z) = \int_{k:Z(k)\leq z} \frac{f(Z(k))}{p(Z(k))} d\mathbb{Q}(k) = \int_{-\infty}^{z} f(x)dx, \tag{2.15}$$

such that the stochastic variable Z has a probability density function f with respect to this probability measure.

Take

$$p(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}}, \quad f(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-mt)^2}{2t}}. \tag{2.16}$$

Then

$$\frac{f(x)}{p(x)} = e^{-\frac{1}{2}m^2 t + mx}, \tag{2.17}$$

with respect to the probability measure $\mathbb{Q}$: $Z \sim \mathcal{N}(0, t)$. The function $f(x) = \frac{1}{\sqrt{2\pi t}} e^{-\frac{(x-mt)^2}{2t}}$ gives the following equation with respect to the probability measure $\mathbb{P}$: $Z \sim \mathcal{N}(mt, t)$, with m as a constant.

$$\mathbb{P}(A) = \mathbb{E}^{\mathbb{Q}}\left[1_A \frac{f(x)}{p(x)}\right] = \mathbb{E}^{\mathbb{Q}}\left[1_A \frac{d\mathbb{P}}{d\mathbb{Q}}\right]. \tag{2.18}$$

The Radon-Nikodym derivative is then defined as

$$\left.\frac{d\mathbb{P}}{d\mathbb{Q}}\right|_T = e^{-\frac{1}{2}m^2 T + mZ(T)}. \tag{2.19}$$

Chapter 3

Option Pricing Models

This chapter elaborates on models for pricing European options. First the definition and the value of a European call option are given. The Theorems of Asset Pricing are explained and conditions to avoid arbitrage in the models are stated. Then, the Black-Scholes model, the Heston stochastic volatility model, Dupire's local volatility model, and the local stochastic volatility model are discussed.

Definition 3.0.1 (Call Option)**.** *A call option is a contract that gives the buyer the right, not the obligation, to buy a predetermined asset S for the predetermined strike price K at the predetermined maturity time T, from the writer of the option. The value of such a call option C at $t = T$ is then given by:*

$$C(S_T, t, T, K) = \max(S_T - K, 0), \tag{3.1}$$

with S_T the price of the asset at time T.

The option is said to be *out of the money* if $S < K$; *at of the money* if $S = K$; or *in of the money* if $S > K$.

The price at time t is given by discounting the expected value of the option at time T by $D(t, T)$. Here, $D(t, T)$ is the zero-coupon bond price, contracted at time t with maturity T. This function is used in throughout this thesis to discount future payoffs under measure $\mathbb{Q}$.

$$C(S_t, t, T, K) = D(t, T)\mathbb{E}^{\mathbb{Q}}[\max(S - K, 0)]. \tag{3.2}$$

3.1 Arbitrage and Theorems of Asset Pricing

To price European options, certain assumptions need to be made. Therefore, the fundamental theorems of asset pricing are stated. To understand these theorems, first the concept of *arbitrage* is explained. Arbitrage is the opportunity to make money without investing money and without the risk of losing money. This can be done if a trading strategy exists in which one starts with no capital and at some later point in time is guaranteed to have not lost money but to have possibly made money. There are certain types of arbitrage. The next paragraphs describe the conditions that are needed to avoid arbitrage.

3.1.1 Call Spread Arbitrage

The following must hold to avoid call spread arbitrage:

$$C(S_t, t, T + a, K) - C(S_t, t, T, K) > 0, \tag{3.3}$$

with $a \in \mathbb{R}^+$. By letting $a \to 0$, the following equation can be obtained [27]

$$\lim_{a \to 0} \frac{1}{a} \left(C(S_t, t, T + a, K) - C(S_t, t, T, K) \right) = \frac{\partial}{\partial T} C(S_t, t, T, K). \tag{3.4}$$

3.1.2 Butterfly Arbitrage

$$C(S_t, t, T, K + a) - 2C(S_t, t, T, K) + C(S_t, t, T, K - a) > 0, \quad \forall a > 0. \tag{3.5}$$

If this equations holds, butterfly arbitrage is impossible. Otherwise, one can buy $C(S_t, t, T, K - a)$ and $C(S_t, t, T, K + a)$, and short sell $2C(S_t, t, T, K)$, and have a non-negative payoff function without the risk of a loss.

3.1.3 Put-Call Parity

The following expression holds

$$(S_T - K)^+ - (K - S_T)^+ = S_T - K, \tag{3.6}$$

which can be rewritten as

$$(S_T - K)^+ + K = (K - S_T)^+ + S_T. \tag{3.7}$$

Discounting with $D(t, T)$ gives

$$D(t, T)(S_T - K)^+ + D(t, T)K = D(t, T)(K - S_T)^+ + D(t, T)S_T. \tag{3.8}$$

Taking conditional expectations with respect to the risk-neutral measure gives

$$\begin{aligned} &D(t, T)\mathbb{E}^{\mathbb{Q}}[(S_T - K)^+ | S_t = s] + D(t, T)\mathbb{E}^{\mathbb{Q}} K | S_t = s] \\ =&D(t, T)\mathbb{E}^{\mathbb{Q}}[(K - S_T)^+ | S_t = s] + D(t, T)\mathbb{E}^{\mathbb{Q}}[S_T | S_t = s]. \end{aligned} \tag{3.9}$$

This can be simplified, as the conditional expectations of $D(t, T)K$ and $D(t, T)S_t$ are deterministic. Furthermore, the discounted value of an asset is a martingale. This implies that the two other terms are the put and call option prices. This gives [16],

$$C_t + D(t, T)K = P_t + S_t \tag{3.10}$$

This equation is the so-called put-call parity, and it should always hold for non-dividend paying stocks to avoid arbitrage.

3.1.4 Monotonicity in the Direction of the Strike

Monotonicity in the direction of the strike must hold. This means [27],

$$C(S_t, t, T, K + a) - C(S_t, t, T, K) < 0, \tag{3.11}$$

with $a \in \mathbb{R}^+$. From the put-call parity, it follows directly that

$$P(S_t, t, T, K + a) - P(S_t, t, T, K) > 0. \tag{3.12}$$

Taking the limit gives

$$\lim_{a \to 0} \frac{1}{a} \left(C(S_t, t, T, K + a) - C(S_t, t, T, K) \right) = \frac{\partial}{\partial K} C(S_t, t, T, K). \tag{3.13}$$

This can be explained by the monotonically in K decreasing payoff function $(S_t - K)^+$.

Finally, the fundamental theorems of asset pricing can be defined.

Theorem 3.1.1 (Fundamental Theorems of Asset Pricing)**.** *The two theorems are as follows:*

- *A market model with a risk-neutral probability measure does not admit arbitrage.*
- *A market model that has a risk-neutral probability measure is complete if and only if the risk-neutral measure is unique.*

Proof. For a detailed proof, see [34]. □

3.2 Black-Scholes Model

In 1973, Black and Scholes [4] developed a model for derivative pricing. This model can be used to define the evolution of a price of an underlying over time. It leads to a PDE that defines the value V, of a European option. The model assumes that the price of an underlying asset follows the geometric Brownian motion[1], defined by:

$$dS = rSdt + \sigma SdW, \tag{3.14}$$

with W a Wiener process, risk-free interest rate r, and constant volatility σ. The Black-Scholes equation is then given by:

Lemma 3.2.1 (Black-Scholes PDE)**.**

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial t^2} + rS\frac{\partial V}{\partial S} - rV = 0. \tag{3.15}$$

Proof. For a detailed proof, see [34]. □

No dividends are assumed. Eventually, this leads to an analytic equation for the call option price given by:

$$C(S_t, t, T, K) = N(d_1)S_t - N(d_2)e^{-(T-t)}K, \tag{3.16}$$

with

$$d_1 = \frac{1}{\sigma\sqrt{T-t}}\left(\log\left(\frac{S_t}{K}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)\right), \tag{3.17}$$

$$d_2 = d_1 - \sigma\sqrt{T-t}. \tag{3.18}$$

Here, T is the maturity, and N is the cumulative distribution function of the standard normal distribution

$$N(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-y^2/2}dy. \tag{3.19}$$

3.2.1 Implied Volatility

The implied volatility is the volatility function denoted by $\sigma_{imp}(K, T)$ which is 'implied' by the Black-Scholes model. This means that a spot price S_t, strike K, maturity T, and $\sigma_{imp}(K, T)$ will lead to call option value determined by the Black-Scholes Equation with these parameters, equal to the call option price in the market,

$$C(S_t, t, T, K, \sigma_{imp}) = C_{mkt}, \tag{3.20}$$

[1]The underlying asset follows a geometric Brownian motion under the risk-neutral measure. Under the real-world measure, the r term can be replaced by another drift term

with C_{mkt}, the market price of the call option, and C the price of the call option determined with the Black-Scholes model. To determine the implied volatility, a numerical approach can be used since no closed-form solution has been found. The Newton-Raphson method [2] can be used to solve this problem. This iterative method defines the following equation to find the following root

$$R(\sigma_{imp}) = C_{mkt} - C(S_t, t, T, K, \sigma_{imp}) = 0. \tag{3.21}$$

This method requires an initial guess σ^0_{imp} and requires the first derivative of the function R. If the initial guess is chosen close to the root, the method converges quadratically. This initial guess can be determined by several techniques, such as the bisection method. Now

$$\sigma^{n+1}_{imp} = \sigma^n_{imp} - \frac{R(\sigma^n_{imp})}{R'(\sigma^n_{imp})}, \tag{3.22}$$

with the superscript stating the iterative process, starting at 0. Here R' is known due to the analytic expression in (3.16). It is equal to minus vega,

$$R'(\sigma_{imp}) = -\frac{\partial C(S_t, t, T, K, \sigma_{imp}, r)}{\partial \sigma_{imp}} = -KD(t,T) f_{\mathcal{N}(0,1)}(d_2)\sqrt{T-t}, \tag{3.23}$$

here $f_{\mathcal{N}(0,1)}(d_2)$ is the standard normal probability density function and d_2 is as defined in (3.18).

3.3 Dupire's Local Volatility Model

The local volatility model of Dupire, published in 1994 [12], is based on the Black-Scholes model. However, the volatility function is different and is not constant. It depends on t and S_t. The advantage of this model is, that one can exactly create the same implied volatility for European options as is observed in the market. Call option prices $C(S_t, t, T, K)$, for different values of the strike K, and for different maturities T, have to be available given S_t to be able to calibrate the model. These call option prices then lead to the local volatility function $\sigma_{dup}(S_t, T)$, which belongs to the diffusion process

$$\frac{dS_t}{S_t} = rdt + \sigma_{dup}(S_t, t)dW_t. \tag{3.24}$$

The corresponding forward price for the delivery at the time of maturity T is given by:

$$F(t,T) = S_t e^{\int_t^T r ds}. \tag{3.25}$$

This is a martingale. Itô's lemma then gives,

$$\frac{dF_t}{F_t} = \sigma(F_t, t)dW_t, \tag{3.26}$$

with

$$\sigma(x,t) = \sigma_{dup}(xe^{-\int_t^T r ds}, t). \tag{3.27}$$

Now, assume the density function $p(S_t, t)$ of S_t exists. To calibrate Dupire's volatility surface, the existence of sufficiently many market prices for call options is required. The forward price of a call option with strike K and time to maturity T is given by:

$$C(T,K) = \int_K^\infty (x-K)p(x,T)dx, \tag{3.28}$$

where the notation $C(S_t, t, T, K)$ has been replaced by $C(T, K)$, because $t = T$ and the distribution of S_T is known by p.

Differentiation with respect to T gives

$$\frac{\partial C}{\partial T}(T,K) = \int_K^\infty \frac{\partial C}{\partial T}(T,x)dx. \tag{3.29}$$

The price at time t of this call option is given by:

$$D(t,T)C(T,K). \tag{3.30}$$

Differentiation of $C(T,K)$ with respect to K leads to the Breeden-Litzenberger formula

$$\frac{\partial C}{\partial K} = -\int_K^\infty p(x,T)dx. \tag{3.31}$$

Repeating this differentiation gives

$$\frac{\partial^2 C}{\partial K^2} = p(x,T). \tag{3.32}$$

For the derivation of Dupire's formula, an expression for the derivative of the density function of S_t, with respect to t, is required. Using Itô calculus, taking a bounded and measurable function h and defining,

$$H(x,t) = \mathbb{E}[h(F_t)|F_t = x], \tag{3.33}$$

gives

$$\mathbb{E}[h(F_T)] = \mathbb{E}[\mathbb{E}[h(F_T)|F_s]] = \int_{-\infty}^\infty H(x,t)p(x,t)dx = \int_0^\infty H(x,t)p(x,t)dx. \tag{3.34}$$

Differentiation of this equation with respect to t gives

$$0 = \int_0^\infty p\frac{\partial H}{\partial t}dx + \int_0^\infty H\frac{\partial p}{\partial t}dx. \tag{3.35}$$

From Equation (3.26) and according to Itô calculus, H satisfies

$$0 = \frac{\partial H}{\partial t} + \frac{1}{2}\sigma^2(x,t)x^2\frac{\partial^2 H}{\partial x^2}, \tag{3.36}$$

so

$$\frac{\partial H}{\partial t} = -\frac{1}{2}\sigma^2(x,t)x^2\frac{\partial^2 H}{\partial x^2}. \tag{3.37}$$

Equation (3.35) then becomes

$$\begin{aligned} 0 &= -\int_0^\infty p\frac{1}{2}\sigma^2(x,t)x^2\frac{\partial^2 H}{\partial x^2}dx + \int_0^\infty H\frac{\partial p}{\partial t}dx \\ &= \int_0^\infty \left(\frac{\partial^2}{\partial x^2}p\frac{1}{2}\sigma^2(x,t)x^2 - \frac{\partial p}{\partial t}\right)Hdx. \end{aligned} \tag{3.38}$$

Using the previous formulas this gives

$$\begin{aligned} \frac{\partial C}{\partial T}(T,K) &= \int_K^\infty (x-K)\frac{\partial p}{\partial T}(x,T)dx \\ &= \frac{1}{2}\int_K^\infty \frac{\partial^2 \sigma^2 x^2 p}{\partial x^2}(x-K)dx, \end{aligned} \tag{3.39}$$

and integrating by parts gives

$$\begin{aligned}\frac{\partial C}{\partial T}(T,K) &= -\frac{1}{2}\int_K^\infty \frac{\partial \sigma_{dup}^2 x^2 p}{\partial x} dx \\ &= \frac{1}{2}\sigma_{dup}^2(K,T)K^2 p(K,T) \\ &= \frac{1}{2}\sigma_{dup}^2(K,T)K^2\frac{\partial^2 C^2}{\partial x^2}(T,K).\end{aligned} \tag{3.40}$$

This gives Dupire's equation [12]

$$\sigma_{dup}(K,T) = \frac{1}{K}\sqrt{\frac{2\frac{\partial C}{\partial T}(T,K)}{\frac{\partial^2 C}{\partial x^2}(T,K)}}. \tag{3.41}$$

The surface of European call option prices $C(T,K)$ is not explicitly known for all T, and K. Therefore, the derivatives with respect to T and to K are not explicitly known. However, for many underlyings, enough call option prices are available to create a grid with call option prices for various strikes and maturities. Using this grid, an approximation of the above-mentioned derivatives can be made using numerical methods. Let $T_{i-1/2}$ be a time and $C_{T_i,K}$ be the call option price given strike K at time T_i. Let $T_{i+1/2}$ be another time where

$$T_{i-1/2} < T_{i+1/2}, \tag{3.42}$$

and with

$$T_{i-1/2} - T_{i+1/2} = \Delta. \tag{3.43}$$

Then the approximation of the partial derivative with respect to the maturity at $T_i = T_{i-1/2} + \frac{1}{2}\Delta$ and strike K is given by:

$$\frac{\partial C}{\partial T} \approx \frac{C_{T_{i+1/2,K}} - C_{T_{i-1/2,K}}}{\Delta}. \tag{3.44}$$

Applying the same technique for the partial derivative with respect to the strike K leads to the approximation at maturity t and strike $K_i = K_{i-1/2} + \frac{1}{2}\Delta_x$,

$$\begin{aligned}\frac{\partial C}{\partial x} &\approx \frac{C_{t,K_{i+1/2}} - C_{t,K_{i-1/2}}}{\Delta_x}, \\ \frac{\partial^2 C}{\partial x^2} &\approx \frac{\frac{\partial}{\partial x}C_{t,K_{i+1/2}} - \frac{\partial}{\partial x}C_{t,K_{i-1/2}}}{\Delta_x}, \\ \frac{\partial^2 C}{\partial x^2} &\approx \frac{C_{t,K_{i+1}} - 2C_{t,K_i} + C_{t,K_{i-1}}}{\Delta_x^2}.\end{aligned} \tag{3.45}$$

Here the following holds

$$K_{i-1} < K_{i-1/2} < K_i < K_{i+1/2} < K_{i+1}, \tag{3.46}$$

$$K_{i-1/2} - K_{i+1/2} = K_{i-1} - K_i = K_i - K_{i+1} = \Delta_x. \tag{3.47}$$

For the boundaries, one can take second order approximations by

$$\begin{aligned}\frac{\partial C}{\partial x} &\approx \frac{-\frac{3}{2}C_{t,K_i} + 2C_{t,K_{i+1}} - \frac{1}{2}C_{t,K_{i+2}}}{\Delta_x},\\ \frac{\partial^2 C}{\partial x^2} &\approx \frac{2C_{t,K_i} - 5C_{t,K_{i+1}} + 4C_{t,K_{i+2}} - C_{t,K_{i+3}}}{\Delta_x^2}.\end{aligned} \tag{3.48}$$

3.3.1 Implied Volatility

The numerical creation of Dupire's surface by Equation (3.41) can lead to significant errors. This can be explained by the numerical errors in the derivatives. In particular, errors in the second derivative can blow up due to division by this factor.

From the Black-Scholes framework, the price of a call option is given by Equation (3.16). Define the forward as in Equation (3.25). Furthermore, define the total variance by:

$$\omega(K,t) = \sigma_{imp}^2(K,t)t, \tag{3.49}$$

and the log-moneyness by:

$$x(K,t) = \log\left(\frac{K}{F(0,t)}\right). \tag{3.50}$$

Change of variables gives for the call option price,

$$C(x,\omega) = S_0 N(d_1) - S_0 e^x N(d_2), \tag{3.51}$$

with,

$$d_1 = -\frac{x}{\sqrt{\omega}} + \frac{1}{2}\sqrt{\omega}, \tag{3.52}$$

and,

$$d_2 = d_1 - \sqrt{\omega}. \tag{3.53}$$

The first- and second-order derivatives in Equation (3.41) can now be expressed as

$$\frac{\partial C}{\partial T} = -r\frac{\partial c}{\partial x} + \frac{\partial c}{\partial \omega}\frac{\partial \omega}{\partial T}, \tag{3.54}$$

$$\frac{\partial C}{\partial K} = \frac{1}{K}\frac{\partial c}{\partial x} + \frac{\partial c}{\partial \omega}\frac{\partial \omega}{\partial K}, \tag{3.55}$$

$$\frac{\partial^2 C}{\partial K^2} = 2K^2\frac{\partial c}{\partial \omega}\left(\left(1 - \frac{1}{2}\frac{Kx}{\omega}\frac{\partial \omega}{\partial K}\right)^2 + \frac{1}{2}K\left(\frac{\partial \omega}{\partial K} + K\frac{\partial^2 \omega}{\partial K^2}\right) - \frac{1}{4}K^2\left(\frac{1}{4} + \frac{1}{\omega}\right)\right). \tag{3.56}$$

Equation (3.41) then becomes

$$\sigma_{dup}^2(K,\omega) = \frac{\frac{\partial \omega}{\partial T} + rK\frac{\partial \omega}{\partial K}}{\left(1 - \frac{1}{2}\frac{Kx}{\omega}\frac{\partial \omega}{\partial K}\right)^2 + \frac{1}{2}K\left(\frac{\partial \omega}{\partial K} + K\frac{\partial^2 \omega}{\partial K^2}\right) - \frac{1}{4}K^2(\frac{\partial \omega}{\partial K})^2\left(\frac{1}{4} + \frac{1}{\omega}\right)}. \tag{3.57}$$

Using Equation (3.49), the derivatives can be expressed in terms of implied volatility σ_{imp}. The first derivative with respect to T becomes

$$\frac{\partial \omega}{\partial T} = \sigma_{imp}^2 + 2T\sigma_{imp}\frac{\partial \sigma_{imp}}{\partial T}, \tag{3.58}$$

the first derivative with respect to K becomes

$$\frac{\partial \omega}{\partial K} = 2T\sigma_{imp}\frac{\partial \sigma_{imp}}{\partial K}, \tag{3.59}$$

and the second derivative with respect to K becomes

$$\frac{\partial^2 \omega}{\partial K^2} = 2T \left(\frac{\partial \sigma_{imp}}{\partial K} \right)^2 + 2T\sigma_{imp} \frac{\partial^2 \sigma_{imp}}{\partial K^2}. \tag{3.60}$$

Now, substituting these variables in Equation (3.57) gives

$$\sigma^2_{dup}(K,T) = \frac{\sigma^2_{imp} + 2\sigma_{imp}T \left(\frac{\partial \sigma_{imp}}{\partial T} + rK \frac{\partial \sigma_{imp}}{\partial K} \right)}{\left(1 - \frac{Kx}{\sigma_{imp}} \frac{\partial \sigma_{imp}}{\partial K} \right)^2 + K\sigma_{imp}T \left(\frac{\partial \sigma_{imp}}{\partial K} - \frac{1}{4} K\sigma_{imp}T (\frac{\partial \sigma_{imp}}{\partial K})^2 + K \frac{\partial^2 \sigma_{imp}}{\partial K^2} \right)}. \tag{3.61}$$

This formula avoids the use of call option prices directly and uses the implied volatility function instead. A two dimensional grid for the volatility function described by Dupire's formula can be obtained by using Equations (3.45), (3.45), (3.48), and (3.61) and replacing C by σ_{imp}. Note that one needs to avoid arbitrage. Therefore, it can be necessary to adjust the values for Δ and Δ_x. An approximation of the entire volatility function can then be obtained by applying numerical techniques on the grid points. Monotonic cubic interpolation on the total variance makes sure that no arbitrage is possible. Formula (3.61) is used in the experiments in this thesis.

3.4 The Heston Model

In 1993, the Heston model [22] was developed, which is a stochastic volatility model. The dynamics of the volatility follows a CIR process in this model.

Definition 3.4.1 (Cox–Ingersoll–Ross (CIR) process)**.** *The following process v_t, with parameters κ, θ, and ε is a CIR process:*

$$dv_t = \kappa(\theta - v_t)dt + \varepsilon\sqrt{v_t}dW^v_t. \tag{3.62}$$

The Heston model can then be described by the following SDEs:

$$\frac{dS_t}{S_t} = \mu dt + \sqrt{v_t}dW^X_t, \tag{3.63}$$

$$dv_t = \kappa(\theta - v_t)dt + \varepsilon\sqrt{v_t}dW^v_t. \tag{3.64}$$

Here, $\kappa, \theta, \varepsilon \geq 0$. W^X and W^v are Wiener processes. It is assumed that $dW^X_v \cdot dW^v_t = \rho dt$, with correlation $\rho \in [-1, 1]$. S_t is here the price of the underlying at time t, and v_t is the variance process. Furthermore, κ is the rate at which v_t returns to the long-term mean θ, ε is the volatility of the volatility, and μ is a drift term.

Note that if $v_0 > 0$, and $2\kappa\theta \geq \varepsilon^2$, then the process v_t can never reach zero. In the case of $2\kappa\theta < \sigma^2$, the origin is attainable and strongly reflecting; this is the Feller condition [13]. If v reaches zero, the variance of the price process is zero. Calibration to market data often leads to parameters that do not satisfy the Feller condition and make it possible for this to happen. This would correspond with no uncertainty for certain simulated paths at some points in time. However, the process is strongly reflecting and it is mean-reverting to θ at rate κ, so it only takes the value zero for a very short time.

With this model, the smile that is seen in the implied volatility function can be reproduced relatively well. However, this model also has certain limitations. The accuracy of pricing short-maturity options is often unsatisfactory [13]. Furthermore, pricing forward start options is not accurate.

3.5 Calibration of the Heston Stochastic Volatility Model

Before one can use a stochastic volatility model, the parameters have to be known. These parameters are not directly clear from the market data. The price of a stock is usually known, as are European option

prices. These option prices contain information about the expectation of the price process in the future till time of maturity. Therefore, this information can be used to calibrate the parameters.

The calibration of the Heston model is based on the following insight. If the parameters are given, the option prices can be computed with the Heston model. Furthermore, the market prices are known, so a comparison between the two prices can be made. The idea is then to minimize the difference between the market prices and model prices by varying the parameters.

Under the martingale measure $\mathbb{Q}$, the price of call options with underlying asset that follows the dynamics of the Heston model is given by Equation (3.2). The maximum function in this function is not differentiable around zero. Depending on the optimization method, this can be an issue. However, the $\max(x, 0)$ function can always be approximated by the smooth function $\pi_\epsilon(x)$,

$$\pi_\epsilon(x) = \begin{cases} 0 & x < -\epsilon, \\ -\frac{1}{16\epsilon^3}x^4 + \frac{3}{8\epsilon}x^2 + \frac{1}{2}x + \frac{3\epsilon}{16} & -\epsilon \leq x \leq \epsilon, \\ x & x > \epsilon, \end{cases} \tag{3.65}$$

with ϵ a small positive real number. Then,

$$C(S_T, t, T, K) = D(t,T)\mathbb{E}^{\mathbb{Q}}[\max(S_T - K, 0)] \approx D(t,T)\mathbb{E}^{\mathbb{Q}}[\pi_\epsilon(S_T - K)]. \tag{3.66}$$

Given the call option prices in the market, a comparison with the option prices generated by the Heston model using parameters $\boldsymbol{\theta} = (\rho, v_0, \theta, \mu, \varepsilon, \kappa)$ can be made. If of I call options, the prices in the market are known, denote the ith observed call option market price by C^i_{mkt}, and the call option price determined with the Heston model by C^i, with $i \in 1, ..I$. Then, minimizing the error functional E over all I options

$$\min_{\boldsymbol{\theta}\in\Theta} E(\boldsymbol{\theta}) = \sum_{i=1}^{I} (C^i(S_t, t, T^i, K^i, \boldsymbol{\theta}) - C^i_{mkt}(S_t, t, T^i, K^i))^2, \tag{3.67}$$

over $\boldsymbol{\theta} \in \Theta$, with Θ the set of all possible parameters, gives the optimal calibration result.

3.6 Global and Local Optimization

Minimizing the error functional E, (Equation (3.67)) over $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, leads to the optimal parameters. Chen [6] shows that this objective function is not necessarily convex and it may exhibit several local minima. As a consequence, the initial guess may influence the convergence to the optimum. The obtained optimum could be a local instead of a global optimum. On the other hand, global optimization may lack the mathematical tractability, which is known for the local optimizer [7]. Another disadvantage of a global optimizer is that the computational costs are higher. The advantage of using global optimization is the continuation of the search to an optimum, even after a local optimum has been found.

This thesis uses both local and global optimization. Local optimization in this thesis has been performed with the *lsqnonlin* function, the least-squares non-linear function, of MATLAB R2017A. This function implements a trust-region reflective minimization algorithm. The global optimization in this thesis has been performed with the *asamin* function in MATLAB. This function uses stochastic methods which entails drawbacks and the convergence to its global optimum is not guaranteed [7]. The exit sequence is determined stochastically; therefore it might end early and might even lead to a solution worse than that of a local search. As a consequence, it is not a priori known which optimizer performs best. These two functions are chosen due to their relative low computation times [7].

For the calibration of the Heston model, both algorithms are used and the parameter set with the smallest error functional of the two is chosen for further calculations. For both optimizers, a bounded set of parameters has been chosen, as the values need to make sense in a financial way. If the Feller condition needs to be satisfied, the bounded set is even more limited. The volatility of volatility ε, the mean volatility θ, the volatility at $t = 0$, v_0, and the rate of return to this mean κ are assumed to be positive. Furthermore, the correlation ρ needs to be in $[-1, 1]$.

3.7 Local Stochastic Volatility Model

The local stochastic volatility model (LSV) combines the local volatility model with the stochastic volatility model. In this model, the volatility from the stochastic volatility model can be modified with a local volatility function. This keeps the dynamics of the model close to the stochastic volatility model, but allows for a correction of the volatility to mimic the implied volatility as observed in the market.

The dynamics of the calibrated LSV model is a non-linear McKean-Vlasov stochastic differential equation, in which the non-linearity is due to the volatility which depends on the marginal distribution of the process. Therefore, this stochastic differential equation is associated to a non-linear Fokker-Planck partial differential equation and is not easily solvable.

Definition 3.7.1 (Local Stochastic Volatility Model)**.** *The local stochastic volatility model is a model for the price process of an underlying S_t defined by the stochastic differential equation*

$$\frac{dS_t}{S_t} = \mu dt + \sigma(S_t, t)a_t(\rho dW_t^x + \sqrt{1-\rho^2}dW_t^v). \tag{3.68}$$

Here, the risk-neutral forward measure is $\mathbb{Q}^T$. The σ-function is the local volatility function, a_t is an Itô process (stochastic volatility), and W_t^x and W_t^v are two uncorrelated Wiener processes. The correlation coefficient is ρ.

Note that if the stochastic volatility a_t is constant and equal to one, the model simplifies to a local volatility model. If $\sigma(S_t, t) = 1$, it simplifies to a SVM. If a_t follows the CIR process, i.e., the variance process in the Heston model, then the LSV model is also called the Heston LSV model.

For the use of such a model, it is important that the parameters reflect the observed data. The computational time required to calibrate a stochastic model and the accuracy of the calibration are both of importance. To proceed, one can begin by calibrating the stochastic volatility part of this hybrid model and then proceed to calibrate the local volatility function given the parameters of the stochastic volatility model.

The dependence of $\sigma(S_t, t)$ on S_t, where S_t depends on a_t, often creates difficulties for the calibration of this model. Pierre Henry-Labordère introduced the term *effective local volatility* for the calibration of LSV [21]. This effective volatility is equal to the local volatility in Dupire's model.

Definition 3.7.2 (Effective Local Volatility)**.** *The effective local volatility σ_{loc} is defined as*

$$\sigma_{loc}(S_t, T)^2 = \sigma(S_T, T)^2 \mathbb{E}^{\mathbb{Q}^T}[a_T^2 | S_t = K]. \tag{3.69}$$

This effective local variance can be set equal to Dupire's volatility (i.e., $\sigma_{dup} = \sigma_{loc}$) to determine the local volatility function. This shows the need to know the joint density function a_t and S_t for the calibration [21]. However, this distribution is not directly known. Several researchers have found ways to determine this conditional expectation. For simple stochastic processes, it can be derived by solving the Kolmogorov forward PDE. In 2007, Piterbarg [29] published an article about the calibration that used Markovian projection closed-form approximations to determine the conditional expectation. To price European options, the one-dimensional distribution of S_T at time T needs to be known. Furthermore, if option prices for arbitrary K are known for maturity T, then this is equivalent to knowing the one-dimensional distribution of S_T at T. Piterbarg then builds his theory on this information, in combination with the following theorem [19].

Theorem 3.7.1. *Let X_t be given by:*

$$dX_t = \alpha(t)dt + \beta(t)dW_t, \tag{3.70}$$

where $\alpha(\cdot), \beta(\cdot)$ are adapted bounded-stochastic processes such that Equation (3.70) admits a unique solution. Define $a(x,t), b(x,t)$ by

$$a(x,t) = \mathbb{E}[\alpha(t)|X_t = x], \tag{3.71}$$

$$b(x,t) = \mathbb{E}[\beta^2(t)|X_t = x], \tag{3.72}$$

Then the SDE

$$dY_t = a(Y_t,t)dt + b(Y_t,t)dW_t, \tag{3.73}$$

$$Y_0 = X_0, \tag{3.74}$$

admits a weak solution Y_t. This solution then has the same one-dimensional distribution as X_t.

Proof. For a detailed proof, see [29]. □

This theorem implies that if option prices are the same for two diffusion processes, a complex diffusion process can be replaced by an easier process to obtain the calibration result.

The Markovian projection can be summarized into four steps.

- *Step 1:* Compute the quadratic variance belonging to the underlying process, driven by a single Wiener process. Fill this in, in the SDE of the price process.
- *Step 2:* Replace the diffusion and drift in the SDE of the price process by their conditional expected values. Theorem 3.7.1 then states that this does not affect option prices.
- *Step 3:* Approximate the conditional expected values from the second step. Several techniques, such as Gaussian approximations can be used for this.
- *Step 4:* Apply parameter averaging techniques to relate the time-dependent coefficients of the SDE, found in the previous step, to time-independent ones.

For more details and a general approach to solve SDEs with this projection theory, see [29].

Madan et al. [26] came up with another way to calibrate the LSV system. They solved the Kolmogorov forward PDE for the local stochastic volatility process. Theorem 3.7.1 states that the SDE of the model can be rewritten as

$$\frac{dS_t}{S_t} = \mu dt + \sigma_{dup}(S_t,t)SdW_t, \tag{3.75}$$

If σ_{dup} is computed with Dupire's model [12], determining $\mathbb{E}[a_T^2|S_T = K]$ leads to obtaining $\sigma^2(K,t)$. To solve Equation (3.69), Madan et al introduced the joint transition density function $p(x,y,T)$. The values x and y are the values that X_T, respectively Y_T, take at time T, with X, the logarithm of the price of the underlying and $Y_T = \log(a_T)$ the logarithm of the stochastic component of volatility.

The density then satisfies the Kolmogorov forward equation as described in Theorem 2.6.

$$-\frac{\partial p}{\partial t} - \frac{\partial}{\partial x}\left[\left(r - \frac{e^{2y}\sigma^2(e^x,t)}{2}\right)p\right] + \frac{\partial^2}{\partial t^2}\left[\frac{pe^{2y}\sigma^2(e^x,t)}{2}\right] - \frac{\partial}{\partial y}[\kappa(\theta - y)p] + \frac{\partial^2}{\partial y^2}\left(\frac{\varepsilon^2 p}{2}\right) = 0 \tag{3.76}$$

subject to $Y_0 = 0$ and S_0, with probability one at time $t = 0$. This leads to

$$\mathbb{E}[a_T^2|S_T = K] = \frac{\int_0^\infty e^{2y}p(e^K,y,T)dy}{\int_0^\infty p(e^K,y,T)dy} = \frac{\sum_{j=1}^{n_Y} p_j e^{2y_j}}{\sum_{j=1}^{n_Y} p_j}. \tag{3.77}$$

The next step is to solve the forward equation in p in Equation (3.76). The boundary condition $\sigma_{dup}(S_0,0)$ can be recovered from Equation (3.69) given $a_0^2 = 1$, and from Dupire's volatility function. Now, in combination with Equation (3.77), $\mathbb{E}[a_T^2|S_T = K]$ can be determined by iterating over time. Take time steps Δ and recover $\mathbb{E}[a_T^2|S_T = K]$ and $\sigma^2(K,T)$ simultaneously. Then, p can be determined

and the process can be repeated for the entire time horizon. Here p is recovered, with $p_{ij}^t = p(x_i, y_j, t)$, by finite differences as follows:

$$\begin{aligned} & p_{ij}^{1/2} + \frac{1}{2}\frac{\partial t}{\partial X}\left[\left(r - q - \frac{e^{2Y_j}\sigma^2(e^{X_{i+1}}, t)}{2}\right)p_{i+1,j}^{1/2} - \left(\mu - \frac{e^{2Y_j}\sigma^2(e^{X_{i-1}}, t)}{2}\right)p_{i-1,j}^{1/2}\right] \\ & -\frac{1}{2}\frac{\partial t}{\partial X^2}(e^{2Y_j}\sigma(e^{X_{i+1},t)}p_{i+1,j}^{1/2} - 2e^{2Y_j}\sigma(e^{X_i}, t)p_{i,j}^{1/2} + e^{2Y_j}\sigma(e^{X_{i-1},t)p_{i-1,j}^{1/2}}) = p_{i,j}^{t-\Delta} \\ = p_{i,j}^t + \frac{1}{2}\frac{\partial t}{\partial Y}[\kappa(\theta - Y_{i,j+1})p_{i,j+1}^t - \kappa(\theta - Y_{i,j-1})p_{i,j-1}^t] - \frac{\varepsilon^2}{2}\frac{\partial t}{\partial Y^2}(p_{i,j+1}^t - 2p_{ij}^t + p_{i,j-1}^t). \end{aligned} \tag{3.78}$$

Madan et al [26] state that this approach leads to a very smooth density, and that this method is stable. They solved first $p_{ij}^{1/2}$ from $p_{ij}^{t-\Delta}$ by using a tridiagonal solver. Then, afterwards p_{ij}^t can be solved from $p_{ij}^{1/2}$.

Other methods that are used to compute the conditional expectation $\mathbb{E}[a_T^2|S_T = K]$ are the Particle method, developed by Henry-Labordère [21], and the Bin method, developed by Van der Stoep [17]. These last two methods are explained into more detail in Chapter 5. Numerical experiments with these two methods, and an alternative model are discussed in Chapter 5 as well.

Chapter 4

Numerical Techniques for Stochastic Volatility Models

This chapter elaborates on the Heston stochastic volatility model. Numerical techniques for efficient computing of call option prices of European and forward start options are discussed. The Monte Carlo method is explained, and several discretizations schemes for this method are discussed and compared in this chapter. Furthermore, the multilevel Monte Carlo method is tested for these discretization schemes for the Heston model and for the CIR process, to price options computationally efficiently.

4.1 Solution of the Heston model

European call option prices can be determined by various methods. One approach is to derive a partial differential equation, and then solve this by numerical methods, see [39]. Other methods can be applied if the characteristic function of the underlying asset price is known. The characteristic functions for several price processes have been obtained in closed form. In these cases, the probability density functions do not have to be known explicitly to compute the characteristic function. Madan, Carr and Chang [25], Heston [22], and Zhu [41] have described characteristic functions of commonly used models.

Fourier analysis can be used to solve the equation by the Gil-Palaez inversion formula. However, this method is not always accurate due to the singularities in the integral. Another method ensures the existence of the Fourier transform of the call option price by the inclusion of a damping factor. This has been developed by Carr and Madan [5]. The fast Fourier transform (FFT) [14] can be used here and it saves time. This method calculates the discrete Fourier transformation, with $O(N \log N)$ operations instead of $O(N^2)$ which would be required for the same accuracy when using the direct approach. The reduction in calculation time can be seen in Figure 4.1.

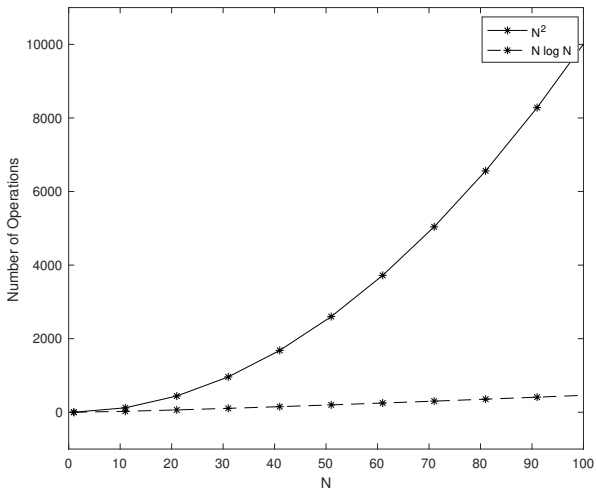
Another recently developed method that has been used in this thesis to check the results is the Cosine Series Expansion method, which is explained in more detail here.

4.1.1 Cosine Series Expansion Method

The Cosine Series Expansion Method (COS) is a method to determine the price of a derivative given the characteristic function. This section explains how this method can be applied to price options based on the article of Fang and Oosterlee [14]. The COS method can be used once the characteristic function is known, which is defined as follows:

Definition 4.1.1 (Characteristic Function)**.** *Let X be a random variable with density function $f(z)$. The*

Figure 4.1: Computational costs reduction by FFT



characteristic function *corresponding to the r.v. X is then defined by:*

$$\phi(u) = \mathbb{E}[e^{iuX}] = \int_{\mathbb{R}} e^{iuz} f(z) dz. \tag{4.1}$$

The COS method exploits the information of the density function contained in the characteristic function. The relationship between the characteristic function and the probability density function is given in lemma 4.1.1 and is referred to as the *Fourier pair*.

Lemma 4.1.1 (Inverse Fourier Transform)**.** *Let X be a random variable, with corresponding characteristic function $\phi(u)$. Then the probability density function f(z) of X is given by:*

$$f(z) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iuz} \phi(u) du \tag{4.2}$$

Proof. For a detailed proof, see [38]. □

The Fourier cosine series expansion can now be obtained for functions defined on a finite interval $[a, b] \in \mathbb{R}$, which is defined as follows:

Lemma 4.1.2 (Fourier Cosine Series Expansion)**.** *For f(z), a real-valued function which is supported on the finite interval $[a, b] \in \mathbb{R}$, the* Fourier Cosine Series Expansion *is given by:*

$$f(z) = \sum_{k=0}^{\infty} w_k A_k \cos\left(k\pi \frac{z-a}{b-a}\right), \tag{4.3}$$

with $w_0 = \frac{1}{2}$, and $w_k = 1$ for all other k, and with

$$A_k = \frac{2}{b-a} \int_a^b f(z) \cos\left(k\pi \frac{z-a}{b-a}\right) dz. \tag{4.4}$$

Proof. For a detailed proof, see [8]. □

The combination of these two lemmas gives an approach to recovering the conditional probability density function from the characteristic function.

Theorem 4.1.1 (COS Conditional Probability Density Recovery)**.** *The conditional probability density function $f(z|x)$ can be approximated by*

$$f(z|x) = \sum_{k=0}^{N-1} w_k B_k \cos\left(k\pi \frac{z-a}{b-a}\right), \tag{4.5}$$

with

$$B_k \approx \frac{2}{b-a} \Re\left(\phi\left(\frac{k\pi}{b-a} \,\middle|\, x\right) exp\left(-ik\pi \frac{a}{b-a}\right)\right) f(z) \cos\left(k\pi \frac{z-a}{b-a}\right) dz. \tag{4.6}$$

Proof. For a detailed proof, see Appendix A.1. □

Altogether, a European option, with underlying that follows a price process with known characteristic function, can now be priced with the COS pricing formula.

Theorem 4.1.2 (COS Pricing Formula)**.** *Let X_t be a stochastic price process of the underlying asset for a derivative with payoff function g(z) and with conditional characteristic function $\phi(u|x)$. The value of the European option V(x,t) at time t for $X_t = x$, can be approximated by*

$$\hat{V}(x,t) = D(t,T) \sum_{k=0}^{N-1} w_k \Re\left(\phi\left(\frac{k\pi}{b-a} \,\middle|\, x\right) exp\left(-ik\pi \frac{a}{b-a}\right)\right) \mathcal{V}_k(T), \tag{4.7}$$

with $\mathcal{V}_k(T)$, the kth Fourier cosine coefficient of the payoff at the time of expiration, T, defined by:

$$\mathcal{V}_k = \frac{2}{b-a} \int_a^b g(z) \cos\left(k\pi \frac{z-a}{a-b}\right) dz. \tag{4.8}$$

Proof. For a detailed proof, see Appendix A.2. □

Theorem 4.1.3 (COS Pricing Formula for a Lévy Process)**.** *The value of an option can be expressed as*

$$\hat{V}(x,t) = D(t,T) \sum_{k=0}^{N-1} w_k \Re\left(\phi\left(k\pi \frac{1}{b-a}\right) exp\left(ik\pi \frac{x-a}{b-a}\right)\right) \mathcal{V}_k(T), \tag{4.9}$$

where the underlying follows a Lévy process.

Proof. For a detailed proof, see Appendix A.3. □

Note that the COS method can be easily exploited by computing prices for a vector of strikes, if desired, which is important for the computational costs. Furthermore, the COS method can be applied to compute the greek $\delta = \dfrac{\partial C}{\partial S}$ by:

$$\delta = D(t,T) \frac{2}{b-a} \sum_{k=0}^{N-1} w_k \Re\left(\phi\left(\frac{k\pi}{b-a}\right) exp\left(ik\pi \frac{x-a}{b-a}\right) \frac{ik\pi}{b-a}\right) \frac{\mathcal{V}_k(T)}{S_t}, \tag{4.10}$$

and the greek $\gamma = \dfrac{\partial^2 C}{\partial S^2}$ by:

$$\gamma = D(t,T) \frac{2}{b-a} \sum_{k=0}^{N-1} w_k \Re\left(\phi\left(\frac{k\pi}{b-a}\right) exp\left(ik\pi \frac{x-a}{b-a}\right) \left(-\frac{ik\pi}{b-a} + \left(\frac{ik\pi}{b-a}\right)^2\right)\right) \frac{\mathcal{V}_k(T)}{S_t}. \tag{4.11}$$

4.1.2 Integral Solution for the European Call Option

Another method for pricing European call options involves an integral, as described by Andersen [1] in the following theorem.

Theorem 4.1.4 (Solution for a European Call Option Price Under the Heston Model)**.** *The price of a European call option, with underlying asset price following the Heston SDE, can be determined by the following formula according to the paper of Andersen [1]:*

$$\mathbb{E}[S_T - K]^+ = S_0 - \frac{K}{2\pi} \int_{-\infty}^{\infty} \frac{exp((\frac{1}{2} - ik)\log(\frac{S_0}{K}) + h_1 - (k^2 + \frac{1}{4})h_2)v_0}{k^2 + \frac{1}{4}} dk, \tag{4.12}$$

with i the complex unit and where

$$h_1 = \frac{-\kappa\theta}{\varepsilon^2}\left(d_+ T + 2\log\left(\frac{d_- + d_+ e^{\xi T}}{2\xi}\right)\right), \quad h_2 = \frac{1 - e^{-\xi T}}{d_- + d_+ e^{-\xi T}}, \quad \hat{\kappa} = \kappa - \frac{\rho\varepsilon}{2}, \tag{4.13}$$

$$d_+ = \xi - (ik\rho\varepsilon + \hat{\kappa}), \quad d_- = \xi + (ik\rho\varepsilon + \hat{\kappa}), \quad \xi = \sqrt{k^2\varepsilon^2(1-\rho^2) + 2ik\varepsilon\rho\hat{\kappa} + \hat{\kappa}^2 + \frac{\varepsilon^2}{4}}, \tag{4.14}$$

Proof. For a detailed proof, see [1]. □

4.2 Monte Carlo Method

For the Black-Scholes model, a semi-closed-form solution exists for the call option price. For the Heston model, the COS method, the FFT method, and the integral approach exist. However, for path dependent exotic options, these approaches usually cannot be used. Other asset pricing models usually cannot be used either. To determine option prices for these kinds of structured products, one can simulate many paths and determine the price based on these simulations. This can be done by the Monte Carlo method.

The Monte Carlo method (MC) is a numerical method which uses repeated random sampling to obtain an expectation. For derivative pricing it can be used by repeatedly sampling assets price paths by using the numerical discretization of a stochastic differential equation. For every realization, the value of the derivative can be computed given this path. Taking the average of all paths and discounting gives the Monte Carlo estimate.

The following formula shows the method for estimating $\mathbb{E}[V(X)]$, where X follows a stochastic price process, V is value of the derivative, with underlying price process X, and I is the total number of simulations.

$$\mathbb{E}[V(X)] = \frac{1}{I}\sum_{i=1}^{I} g(X_i), \tag{4.15}$$

with $g(x)$ the payoff function belonging to price path x. The law of large numbers states that this estimator of $\mathbb{E}[V(X)]$ is unbiased, i.e., the formula converges to $\mathbb{E}[V(X)]$, for $I \to \infty$. The central limit theorem states that the error E is then distributed as follows $E \sim \mathcal{N}\left(0, \frac{\sigma^2}{I}\right)$.

4.3 Discretization

Once the parameters of the Heston model are calibrated, European option prices can be determined by the FFT or COS method. However, certain type processes do not have a known characteristic function,

or the characteristic function is not enough to determine the value of a path dependent option. The characteristic function of a process following a local stochastic volatility model is also unknown. To price these products, Monte Carlo simulations can be used. In all the models that are discussed here, the stochastic price processes are continuous. However, these price processes can be approximated on a discrete time set. This is the so-called discretization. The density distribution of the price of the underlying is then approximated. There are different time-stepping schemes to consider, such as the Euler scheme, the Milstein scheme, and the Quadratic Exponential scheme. The advantages and disadvantages of these schemes are elaborated in the next section.

4.3.1 Euler Discretization Scheme for the Heston Model

The Euler scheme can be derived by applying Taylor expansion [1]. For a function h with a first derivative and given time discretization step Δ, this gives

$$h(t+\Delta) = h(t) + \frac{\partial h}{\partial t}(t)\Delta + \mathcal{O}(\Delta^2). \tag{4.16}$$

Applying this technique to the two equations in the Heston model

$$\begin{aligned} dS_t =& \mu S_t dt + \sqrt{v_t} dW_t^x, \\ dv_t =& \kappa(\theta - v_t)dt + \varepsilon\sqrt{v_t} W_t^v, \end{aligned} \tag{4.17}$$

with,

$$\begin{aligned} S_{t+\Delta} =& S_t + \int_t^{t+\Delta} S_u \mu(S_u, u)du + \int_t^{t+\Delta} \sigma_(S_u, u)dW_u^x, \\ v_{t+\Delta} =& v_t + \int_t^{t+\Delta} \kappa(\theta - v_u)du + \int_t^{t+\Delta} \varepsilon\sqrt{v_u} dW_u^v, \end{aligned} \tag{4.18}$$

gives,

$$\begin{aligned} v_{t+\Delta} =& v_t + \kappa(\theta - v_t)\Delta + \varepsilon\sqrt{v_t\Delta} Z_v, \\ S_{t+\Delta} =& S_t + \mu(S_t, t)S_t dt + \sqrt{v_t\Delta} S_t Z_x, \end{aligned} \tag{4.19}$$

with Z_v, and Z_x are normally distributed with mean 0 and standard deviation 1, and correlated with ρ, since $W_{t+\Delta} - W_t = \sqrt{\Delta}Z$. Note, that if the Feller condition is not satisfied, v_t can become negative. Computing the root causes a problem then. This can be avoided by replacing v_t by $v_t^+ = \max(v_t, 0)$. This can be applied to the Euler scheme, as well as to the Milstein scheme, which is explained in the next two sections. The Euler discretization scheme converges strongly with order a half. There is bias, and it has poor distribution tails [1]. Therefore, more advanced schemes are developed in what follows.

4.3.2 Milstein Discretization Scheme for the Heston Model

The Milstein scheme uses more information about the density distribution of the process than the Euler scheme. It is a general approach that can be applied to several processes. The theory of the Milstein discretization for the Heston model in this section is based on an article by Frouah [32]. Consider the price process,

$$dS_t = \mu(S_t, t)dt + \sigma(S_t, t)dW_t, \tag{4.20}$$

gives,

$$S_{t+\Delta} = S_t \int_t^{t+\Delta} \mu(S_s, s)ds + \int_t^{t+\Delta} \sigma(S_s, s)dW_s. \tag{4.21}$$

Applying Itô's lemma to μ and σ then gives

$$d\mu = \left(\frac{\partial \mu}{\partial S_t}\mu_t + \frac{1}{2}\frac{\partial^2 \mu}{\partial S_t^2}\sigma^2(S_t, t)\right)dt + \frac{\partial \mu}{\partial S_t}\sigma(S_t, t)dW_t, \tag{4.22}$$

$$d\sigma = \left(\frac{\partial \sigma}{\partial S_t} \mu_t + \frac{1}{2} \frac{\partial^2 \sigma}{\partial S_t^2} \sigma^2(S_t, t) \right) dt + \frac{\partial \sigma}{\partial S_t} \sigma(S_t, t) dW_t. \tag{4.23}$$

The derivatives to t are zero here, because it is assumed that there is no direct dependency on t. Combined this gives

$$\begin{aligned} S_{t+\Delta} = S_t &+ \int_t^{t+\Delta} \left(\mu(S_t, t) + \int_t^s \left(\frac{\partial \mu}{\partial S_u} \mu(S_u, u) + \frac{1}{2} \frac{\partial^2 \mu}{\partial S_u^2} \sigma^2(S_u, u) \right) du + \int_t^s \frac{\partial \mu(S_u, u)}{\partial S_t} \sigma(S_u, u) dW_u \right) ds \\ &+ \int_t^{t+\Delta} \left(\sigma(S_t, t) + \int_t^s \left(\frac{\partial \sigma}{\partial S_t} \mu(S_u, u) + \frac{1}{2} \frac{\partial^2 \sigma(S_u, u)}{\partial S_u^2} \sigma^2(S_u, u) \right) du + \int_t^s \frac{\partial \sigma(S_u, u)}{\partial S_t} \sigma(S_u, u) dW_u \right) dW_s \end{aligned} \tag{4.24}$$

All terms of higher order than one are ignored. The term involving $dW_u dW_s$ is retained since $dW_u dW_s = \mathcal{O}(\Delta)$ is of order one.

Now Equation (4.21) becomes

$$S_{t+\Delta} = S_t + \mu(S_t, t) \int_t^{t+\Delta} ds + \sigma(S_s, s) \int_t^{t+\Delta} dW_s + \int_t^{t+\Delta} \int_t^s \frac{\partial \sigma}{\partial S_u} \sigma(S_u, u) dW_u dW_s. \tag{4.25}$$

Furthermore, Itô's lemma gives

$$\int_t^{t+\Delta} W_s dW_s = \frac{1}{2} W_{t+\Delta}^2 - \frac{1}{2} W_t^2 - \frac{1}{2} \Delta, \tag{4.26}$$

such that

$$\begin{aligned} \int_t^{t+\Delta} \int_t^s \frac{\partial \sigma}{\partial S_u} \sigma(S_u, u) dW_u dW_s &\approx \frac{\partial \sigma}{\partial S_t} \sigma(S_t, t) \int_t^{t+\Delta} \int_t^s dW_u dW_s, \\ &= \frac{\partial \sigma}{\partial S_t} \sigma(S_t, t) \int_t^{t+\Delta} (W_s - W_t) dW_s, \\ &= \frac{\partial \sigma}{\partial S_t} \sigma(S_t, t) \left(\int_t^{t+\Delta} W_s dW_s - W_t W_{t+\Delta} + W_t^2 \right), \\ &= \frac{\partial \sigma}{\partial S_t} \sigma(S_t, t) \left(\frac{1}{2} (W_{t+\Delta} - W_t)^2 - \Delta \right). \end{aligned} \tag{4.27}$$

Equation (4.25) reads

$$S_{t+\Delta} = S_t \mu_t \Delta + \sigma(S_t, t) \sqrt{\Delta} Z + \frac{1}{2} \frac{\partial \sigma}{\partial S_t} \sigma(S_t, t) \Delta (Z^2 - 1). \tag{4.28}$$

The accuracy becomes higher because of the expanding μ and σ, in comparison with the Euler discretization.

Applying this technique to the Heston model gives

$$v_{t+\Delta} = v_t + \kappa(\theta - v_t)\Delta + \varepsilon \sqrt{v_t \Delta} Z_v + \frac{1}{4} \varepsilon^2 \Delta (Z_v^2 - 1), \tag{4.29}$$

$$S_{t+\Delta} = S_t + \mu(S_t, t) S_t \Delta + \sqrt{v_t \Delta} S_t Z_x + \frac{1}{4} S_t^2 \Delta (Z_x^2 - 1). \tag{4.30}$$

Note that the discretization is similar to the discretization following from the Euler scheme, but with an extra term to improve the accuracy. The Milstein scheme also converges strongly, but with order one, and with lower bias.

4.4 Quadratic Exponential Discretization Scheme for the Heston Model

Besides the relatively simple Euler and Milstein schemes, a numerical solution can also be determined with the Quadratic Exponential discretization scheme (QE). This section is based on the description of the QE method in an article by Andersen [1]. The QE scheme has the advantage that it determines v_t accurately while using more information of the density distribution than the above-mentioned algorithms. The approximation for v_t can be described by:

$$v_{t+\Delta} = a(b + Z_v)^2, \tag{4.31}$$

with Z_v the standard Gaussian random variable and Δ is the chosen discretization time step. Furthermore, constants a, b, and ψ are defined as

$$a = \frac{m}{1+b^2}, \quad b^2 = 2\psi^{-1} - 1 + \sqrt{2\psi^{-1}}\sqrt{2\psi^{-1}-1} \geq 0, \quad \psi = \frac{s^2}{m^2}. \tag{4.32}$$

Here s and m are defined as follows:

$$s = \sqrt{\frac{v_t\varepsilon^2 e^{-\kappa\Delta}}{\kappa}(1-e^{-\kappa\Delta}) + \frac{\theta\varepsilon^2}{2\kappa}(1-e^{-\kappa\Delta})^2}, \quad m = \theta + (v_t - \theta)e^{-\kappa\Delta}. \tag{4.33}$$

This scheme works well for larger values of v_t. However, the expression does not work well for small values of v_t. Therefore, another expression is used here that gives accurate results for small values,

$$\mathbb{P}(v_{t+\Delta} \in [x, x+dx]) \approx \left(p\delta(0) + \beta(1-p)e^{-\beta x}\right) dx, \quad x \geq 0, \tag{4.34}$$

with $x \geq 0$. In this formula δ is the Dirac deltafunction, and β and p are constants.

$$\beta = \frac{1-p}{m}, \quad p = \frac{\psi - 1}{\psi + 1} \in [0, 1), \tag{4.35}$$

and define,

$$v_{t+\Delta} = \Psi^{-1}(U_v, p, \beta) = \begin{cases} 0 & 0 \leq u \leq p, \\ \beta^{-1}\log\left(\frac{1-p}{1-u}\right) & p < u \leq 1. \end{cases} \tag{4.36}$$

In summary, the QE discretization can be described by the following steps:

- *Step 1:* Given v_t, calculate m and s according to Equation (4.33).
- *Step 2:* Calculate $\psi = s^2/m^2$.
- *Step 3:* Generate a random uniform variable U_v.
- *Step 4:* If $\psi \leq \psi_c$.
  1. Calculate a and b according to Equation (4.32).
  2. Calculate $Z_v = \phi^{-1}(U_v)$.
  3. Calculate $v_{t+\Delta} = a(b + Z_v)^2$.
- *Step 5:* When $\psi > \psi_c$ then:
  1. Calculate β and p according to Equation (4.35) and according to Equation (4.35).
  2. Calculate $v_{t+\Delta} = \Psi^{-1}(U_v, p, \beta)$.

Figure 4.2: Empirical CDF $v_0 = 0.01$

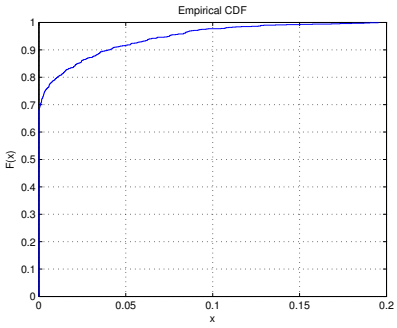
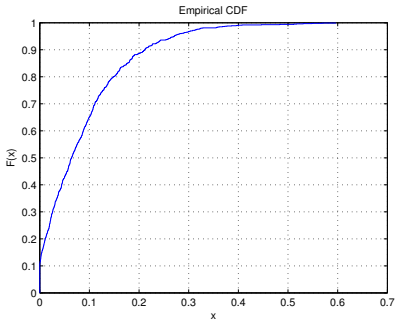


Figure 4.3: Empirical CDF $v_0 = 0.09$



Discretization of the Underlying

Since naive Euler discretization suffers from poor distribution tails [1], the following discretization is used:

$$S_{t+\Delta} = S_t exp(K_0 + K_1 v_t) exp\left(K_2 v_{t+\Delta} + \sqrt{K_3 v_t + K_4 v_{t+\Delta}} Z\right), \tag{4.37}$$

with

$$K_0 = -\frac{\rho\kappa\theta}{\varepsilon}\Delta, \quad K_1 = \gamma_1\Delta(\kappa\rho/\varepsilon - \frac{1}{2}) - \frac{\rho}{\varepsilon}, \tag{4.38}$$

$$K_2 = \gamma_2\Delta(\kappa\rho/\varepsilon - \frac{1}{2}) + \frac{\rho}{\varepsilon}, \quad K_3 = \gamma_1\Delta(1-\rho^2), \tag{4.39}$$

$$K_4 = \gamma_2\Delta(1-\rho^2). \tag{4.40}$$

Here, γ_1 and γ_2 are chosen to be both $\frac{1}{2}$, which corresponds to a central discretization. A simple Euler setting would give $\gamma_1 = 1$ and $\gamma_2 = 0$.

Now,

- *Step 1:* Given v_t generate $v_{t+\Delta}$ using the QE scheme.
- *Step 2:* Draw Z from a Gaussian distribution independent for all random numbers used for $v_{t+\Delta}$.
- *Step 3:* Compute $S_{t+\Delta}$ given S_t, v_t and $v_{t+\Delta}$.

Results of the QE Scheme

First the dependency of the distribution of v_t on v_0 is investigated. Figure 4.2 and Figure 4.3 present the cumulative distribution function of v_t,

$$F(x) = \mathbb{P}(v_t \leq x), \tag{4.41}$$

with parameters are $t = 0.1$, $\theta = 0.04$, $\kappa = 0.5$, and $\varepsilon = 1$. In Figure 4.2, $v_0 = 0.01$, and in Figure 4.3, $v_0 = 0.09$. This shows the impact that v_0 has at time t, and it indicates the importance of a good calibration of v_0.

The call option price at time, obtained by the FFT method, is denoted as C_{FFT}. The call option price obtained by the QE discretization and thereafter by using MC simulations with the QE discretization scheme is denoted as C_{MC}, and the error is defined as $C_{FFT} - C_{MC}$, to obtain insight into the bias.

Table 4.1: Pricing errors for Parameter Set 1 for different step sizes

Δ	K=70	K=100	K=140
1	-0.9677	0.0809	-1.2584
1/2	-0.3080	0.0201	-0.1446
1/4	-0.1327	-0.0213	-0.3207
1/8	0.0724	0.0204	0.1027
1/16	-0.0323	-0.0184	0.0991
1/32	0.0284	-0.0134	0.0621

The impact of the decision concerning the step size on the error is investigated. Table 4.1 includes the error following from a numerical test with 10^5 runs with parameters $T = 10$, $\theta = 0.04$, $v_0 = 0.04$, $\kappa = 0.5$, $\varepsilon = 1$, $S_0 = 100$, and $\rho = -0.9$ (Parameter Set 1).

Table 4.2 investigates the same for the parameters $T = 15$, $\theta = 0.04$, $v_0 = 0.04$, $\kappa = 0.3$, $\varepsilon = 0.9$, $S_0 = 100$, and $\rho = -0.5$ (Parameter Set 2). Here, the difference between the bias at the money, out of the money, and in the money are remarkably smaller. So it can be concluded that the bias depends on the parameter set.

Table 4.2: Pricing errors for Parameter Set 2 for different step sizes

Δ	K=70	K=100	K=140
1	0.5080	0.6469	-0.5237
1/2	-0.1378	-0.0311	-0.2433
1/4	0.1653	-0.2300	0.0596
1/8	-0.2393	-0.4323	0.2151
1/16	-0.0900	-0.1302	0.1535
1/32	0.1582	0.0823	-0.0492

The Quadratic Exponential scheme to discretize the Heston Model is a simple discretization method for simulating the volatility of the underlying. Using this volatility, a discretization for the underlying can be made. The Monte Carlo method then provides an accurate approximation of the call option price according to the results. Results rapidly become closer to the FFT solution when the steps Δ become smaller. The bias is also remarkably smaller than the bias with Euler, and it depends on the parameter set. The bias at $\Delta = 1$ for the QE scheme is similar to the bias for $\Delta = 1/16$ with the Euler scheme.

4.5 Comparison of the Discretization Schemes

For the comparison of the Euler, Milstein, and QE discretization schemes[1] used when pricing European options with underlying asset that follows the dynamics of the Heston model, a distinction between two parameter sets is made. The parameter sets that satisfy the Feller condition, and the sets which do not satisfy this condition. Define the Feller value as,

$$q = \frac{2\kappa\theta}{\varepsilon^2} - 1. \tag{4.42}$$

Then the Feller condition is satisfied if $q > 0$.

The distinction is between the scenario with this condition satisfied and not satisfied.

[1]The Kahl Jackel scheme [23], with the variance process approximated by the implicit Milstein scheme, coupled with the so-called 'IJK' discretization scheme has also been tested. However the results were disappointing with a bias of more than factor 10 larger than the bias with the Euler scheme. This is in accordance with the research of Andersen [1]. Therefore, this procedure has not been further investigated.

4.5.1 A Satisfied Feller Condition Experiment

The following parameters for the Heston model are used here $\rho = -0.9$, $v_0 = 0.04$, $\theta = 0.04$, $K = 100$, $S_0 = 100$, $r = 0$, $\varepsilon = 0.5$, $\kappa = 5$, $T = 1$. Here, $q = 0.6$, so the Feller condition is satisfied. The FFT price of a call option with these parameters is 7.5789, which is in accordance with the literature [28].

For $I = 10^5$ runs, the average price error PE,

$$PE = \frac{1}{I}\sum_{i=1}^{I} PE_i, \tag{4.43}$$

with

$$PE_i = |C^i_{MC} - C_{FFT}|, \tag{4.44}$$

is given in Table 4.3 for several step sizes. The standard deviation (Std) for 10^4 runs is given in parentheses. The standard deviation is defined as,

$$Std = \sqrt{\frac{1}{I}\sum_{i=1}^{I}(PE_i - PE)^2}. \tag{4.45}$$

Table 4.3: Pricing errors with Feller condition satisfied

Step size	Euler PE (Std)	Milstein PE (Std)	QE PE (Std)
1/16	0.1837 (0.1120)	0.1421 (0.0939)	0.0945 (0.0739)
1/32	0.1316 (0.1115)	0.1233 (0.0847)	0.0736 (0.0425)
1/64	0.1106 (0.0957)	0.1028 (0.0717)	0.0652 (0.0628)
1/128	0.0932 (0.0913)	0.0873 (0.0678)	0.0407 (0.0675)
1/256	0.0890 (0.0915)	0.0779 (0.0558)	0.0184 (0.0557)

It is clear that the pricing error is smaller by about a factor of two for the Milstein scheme relative to the Euler scheme, and by a factor of eight for the QE scheme relative to the Euler scheme. This is in accordance with the theory about the convergence and bias of the previous section. Furthermore, computational time is also important in practice. Expressed proportional to the computation time when using the Euler scheme, this is Euler=1, Milstein=1.04, and QE=1.21. Altogether, this shows that the Milstein scheme computes the price of a European option faster for a given accuracy level than the Euler scheme does, and that the QE scheme computes the price of a European option faster for a given accuracy level than the Milstein scheme does.

4.5.2 An Unsatisfied Feller Condition Experiment

The same experiment is performed with the following parameters for the Heston model $\rho = -0.9$, $v_0 = 0.04$, $\theta = 0.04$, $K = 100$, $S_0 = 100$, $r = 0$, $\varepsilon = 1$, $\kappa = 0.5$, $T = 10$. Here, $q = -0.96$; thus the Feller condition is not satisfied. The FFT price of a call option with these parameters is 13.0847, which is in accordance with the literature [28].

The price error, PE (4.43) is given in Table 4.4. The standard deviation is determined by 10 samples of 10^4. The price error PE is determined by the combination of all 10 samples, i.e., 10^5 paths.

The results are similar to the results in the previous experiment. However, the standard deviation and PE is in general larger than in the previous experiment. This can be explained by the v_t^+ factor which replaces v_t to avoid the computation of the root of a negative number. The results are again in accordance with the convergence and bias theories of the previous section. The computational time, expressed proportional to the computation time when using the Euler scheme, is here Euler=1, Milstein=1.05, and QE=1.23. The calculation time of the Euler scheme in this experiment was 1.06 which is as high as

Table 4.4: Pricing errors with Feller condition not Satisfied

Step size	Euler PE (Std)	Milstein PE (Std)	QE PE (Std)
1/16	0.2464 (0.2742)	0.1532 (0.1713)	0.0985 (0.1249)
1/32	0.2011 (0.2093)	0.1224 (0.1834)	0.0932 (0.0860)
1/64	0.1783 (0.1824)	0.1398 (0.1893)	0.0578 (0.0836)
1/128	0.1293 (0.1513)	0.1011 (0.1432)	0.0693 (0.0902)
1/256	0.1092 (0.1597)	0.0901 (0.1563)	0.0207 (0.0734)

the calculation time of the previous experiment. Altogether, this again shows that the Milstein scheme computes the price of a European option faster for a given accuracy level than the Euler scheme does, and that the QE scheme computes the price of a European option faster for a given accuracy level than the Milstein scheme does. It also shows that if the Feller condition is not satisfied, the computational time is (slightly) higher than if this condition is satisfied.

4.6 Multilevel Monte Carlo Method

For the simulation of a price process, the size of the time steps needs to be determined. Mike Giles introduced the multilevel Monte Carlo method (MLMC) in 2008 [15]. This section is based upon the theory in his article. The MLMC method uses different time steps to reduce computational costs and increase the convergence. The idea behind the model is to simulate paths with different meshes then use the differences in the estimation of the expectation in a smart way. The coarse mesh is divided into a fine mesh with factor M, while the same underlying variables (Wiener processes) are used.

The MLMC method takes different time steps Δ_l into account: $\Delta_l = \frac{T}{M^l}$, with $l = 0, 1, ..., L$ the level of the simulation, T the time of maturity, and with $M \geq 2$ the factor that decreases the time step to the next level. Let the value of a derivative be given by $V(S_T)$, with S_T following an SDE evaluated at time T. Furthermore, let S_{l,M^l} be the approximation of S_T using a discretization scheme with time step Δ_l. Now the MLMC computes the following, expectation

$$\mathbb{E}[V_L] = \mathbb{E}[V_0] + \sum_{l=1}^{L} \mathbb{E}[V_l - V_{l-1}], \tag{4.46}$$

by estimating the right-hand side. This is done in a way that minimizes the computational complexity, which gives faster convergence. Define the estimator of these expectations Y_l by,

$$Y_l = \frac{1}{N_l} \sum_{i=1}^{N_l} (V_l^i - V_{l-1}^i), \tag{4.47}$$

for $l \geq 1$, N_l the number of simulated paths at level l, and $Y_0 = \mathbb{E}[V_0]$. A crucial factor for this estimation is that the same simulated path is used for two approximations with different time steps. For a simple discretization, this can be done by first simulating the path using the fine mesh and then summing the underlying Wiener processes into groups of M. The variance is given by:

$$\mathrm{Var}[Y] = \sum_{l=0}^{L} \frac{\mathrm{Var}_l}{N_l}, \tag{4.48}$$

with $Y = \sum_{l=0}^{L} Y_l$, and with Var_l the variance of one path with discretization fineness on level l. For the Euler discretization of a Black-Scholes process, strong convergence is of order $\mathcal{O}(\sqrt{\Delta})$ (weak convergence of order $\mathcal{O}(\Delta)$). Now, for $l \to \infty$ conditional on $L \gg 1$, $\mathbb{E}[V_l - V] = \mathcal{O}(\Delta_l)$. Furthermore $\mathbb{E}[\|S_{l,M^l} - S_T\|^2] = \mathcal{O}(\Delta_l)$

For a payoff function with uniform Lipschitz bound, the following holds [15]

$$\|V(x) - V(y)\| \leq c\|x - y\| \quad \forall x, y \in \mathbb{R}^n. \tag{4.49}$$

This gives

$$\text{Var}[V_l - V] \leq \mathbb{E}[(V_l - V)^2] \leq c^2 \mathbb{E}[\|S_{l,M^l} - S_T\|^2]. \tag{4.50}$$

Therefore,

$$\text{Var}[V_l - V] = \mathcal{O}(\Delta_l) \tag{4.51}$$

Giles [15] set the number of paths per level l, N_l,

$$N_l = \mathcal{O}\left(\frac{L\Delta_l}{\epsilon^2}\right). \tag{4.52}$$

Then the variance of Y is of order $\mathcal{O}(\epsilon^2)$. Furthermore, he chooses

$$L = \left\lceil \frac{\log \frac{1}{\epsilon}}{\log M} \right\rceil. \tag{4.53}$$

Now,

$$\mathbb{E}[V_l - V] = \mathcal{O}(\Delta_l) = \mathcal{O}\left(\frac{1}{M}\right) = \mathcal{O}(\epsilon), \tag{4.54}$$

with V the exact value. Thus the mean squared error is of order $\mathcal{O}(\epsilon^2)$.

The advantage of this algorithm over the traditional MC method is the reduction in computational time for the same order error. This computational time is proportional to the number of time steps per path and the number of paths. These computational costs (CC) are expressed as

$$CC = \sum_{l=0}^{L} c \frac{N_l}{\Delta_l}, \tag{4.55}$$

with c a constant. Giles concludes in his paper that choosing N_l proportional to $\sqrt{\text{Var}_l \Delta_l}$ is optimal. For the determination of all optimal N_l, one needs to take into account the different computational costs for different paths belonging to different time steps. The computation complexity follows from the following theorem.

Theorem 4.6.1 (Complexity Theorem)**.** *Let V be a functional of the solution of $dS_t = a(S,t)dt + b(S,t)dW_t$, with $0 < t < T$ given the path of the Wiener process. Furthermore, let V_l denote the approximation of this process with discretization of stepsize $\Delta_l = \frac{T}{M}$. If positive constants $\alpha \geq \frac{1}{2}$, β, c_1, c_2, c_3, and independent estimators Y_l, exist (based on N_l MC paths) such that,*

1. *if $l = 0$, $\mathbb{E}[Y_l] = \mathbb{E}[V_0]$, and otherwise, $\mathbb{E}[Y_l] = \mathbb{E}[V_l - V_{l-1}]$*
2. *$C_l \leq c_1 \frac{N_l}{\Delta_l}$, with C_l the computational complexity of Y_l,*
3. *$\text{Var}[Y_l] \leq c_2 \frac{\Delta_l^\beta}{N_l}$,*
4. *$\mathbb{E}[V_l - V] \leq c_3 \Delta_l^\alpha$,*

then there exists a positive constant c_4 such that for any $\epsilon < e^{-1}$, there are values L and N_l for which the multilevel estimator $Y = \sum_{l=0}^{L} Y_l$ has an MSE with bound $MSE = \mathbb{E}[(Y - \mathbb{E}[V])^2] < \epsilon^2$, with a computational complexity CC with bound

$$CC \leq \begin{cases} c_4 \epsilon^{-2}, & \beta > 1, \\ c_4 (\log \epsilon)^2 \epsilon^{-2}, & \beta = 1, \\ c_4 \epsilon^{-2 - \frac{(1-\beta)}{\alpha}}, & 0 < \beta < 1. \end{cases} \tag{4.56}$$

Proof. For a detailed proof, see [15]. □

Now, from Equation (4.54) and Theorem 4.6.1, it follows that the computational complexity is $\mathcal{O}(\frac{1}{\epsilon^2})$.

Besides L, the optimal M has to be determined as well. For the $\beta = 1$ case in Theorem 4.6.1, $\text{Var}[V_l - V] \approx c_5\Delta_l$, with c_5 a positive constant[2], with a Lipschitz payoff function, and with Euler discretization. Furthermore, the following bounds for $\text{Var}[V_l - V_{l-1}]$ can be given:

$$(\sqrt{M}-1)^2 c_5\Delta_l \leq \text{Var}[V_l - V_{l-1}] \leq (\sqrt{M}+1)^2 c_5\Delta_l, \tag{4.57}$$

due to $(V_l - V) + (V_{l-1} - V) = V_l - V_{l-1}$. Now, under the assumption that $\text{Var}[V_l - V_{l-1}] \approx (M-1)c_5\Delta_l$, $c_2/c_5 = (M-1)$, and

$$N_l \approx \frac{2}{\epsilon^2}(L+1)(M-1)c_5\Delta_l. \tag{4.58}$$

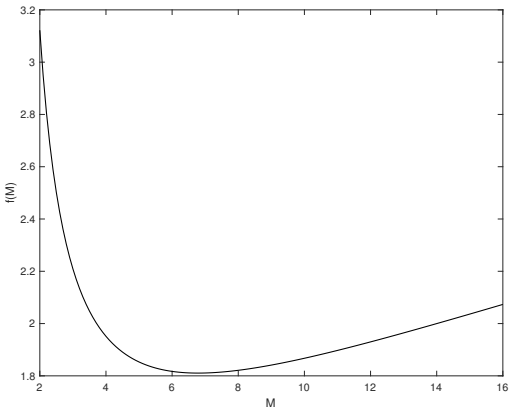
The computational costs for Y are then $N_l\left(\frac{1}{\Delta_l} + \frac{l}{\Delta_{l-1}}\right) \approx c_6\frac{2}{\epsilon^2}(L+1)(M-\frac{1}{M})$, with c_6 a constant.

Now, for $\epsilon \to 0$, asymptotically, these computational costs are

$$c_7\frac{1}{\epsilon^2}(\log\epsilon)^2\frac{M-M^{-1}}{(\log M)^2}, \tag{4.59}$$

with c_7 a positive constant. Furthermore, note that $L = \mathcal{O}\left(\frac{\log\epsilon^{-1}}{\log M}\right)$. Equation (4.59), with $c_7\frac{1}{\epsilon^2}(\log\epsilon)^2 = 1$, is given in Figure 4.6 Here, for $M = 2$, $\frac{M-M^{-1}}{(\log M)^2} = 3.1221$. For $M = 4$, $\frac{M-M^{-1}}{(\log M)^2} = 1.9513$, and for

Figure 4.4: Function $f(M) = \frac{M-M^{-1}}{(\log M)^2}$



$M = 7$, $\frac{M-M^{-1}}{(\log M)^2} = 1.8109$. The results in the numerical result section use $M = 4$ since the difference between the value that the function $\frac{M-M^{-1}}{(\log M)^2}$ takes at $M = 4$ and the minimum is small, and it reduces the number of levels in comparison with $M = 7$, where the function attains its minimum.

The bias of the estimates can be determined by $\mathbb{E}[V - V_l]$. For the Euler discretization with a Lipschitz payoff function for $l \to \infty$, the bias is approximatelty given by:

$$\mathbb{E}[V - V_l] \approx c_8\Delta_l, \tag{4.60}$$

with c_8 a constant [15]. Therefore, $\mathbb{E}[V - V_l] \approx c\Delta_l \approx \frac{\mathbb{E}[V_l - V_{l-1}]}{(M-1)}$. The MLMC method has to continue to the next level as long as the bias is larger than desired. A stopping criterion for a bias of ϵ might therefore be, stop if $|Y_l| < (M-1)\varepsilon$. This can be made stricter by including the previous level which gives the stopping criterion defined as,

$$\max\left(Y_l, \frac{Y_{l-1}}{M}\right) < (M-1)\epsilon. \tag{4.61}$$

[2]Note, $c_3 = c_5$ if $\alpha = 1$.

This information about the bias can also be used to reduce the bias (by eliminating the original leading order bias) by adding $\frac{Y_L}{(M-1)}$ to the original estimator.

From Equation (4.58), it follows that the optimal N_l is given by:

$$N_l = \left\lceil \frac{2}{\epsilon^2} \sqrt{V_l \Delta_l} \left(\sum_{l=0}^{L} \sqrt{\frac{V_l}{\Delta_l}} \right) \right\rceil . \tag{4.62}$$

For a given level of accuracy ϵ, one can choose N_l such that this is obtained with MLMC. If Equation (4.61) is scaled by $\frac{1}{\sqrt{2}}$, then the bias is smaller than $\frac{\epsilon}{\sqrt{2}}$. The choice of N_l then ensures the combined estimator variance to be smaller than $\frac{\epsilon^2}{2}$ such that the MSE is smaller than ϵ^2.

4.6.1 MLMC Results

To test the theory of the previous section in practice, this section elaborates on four experiments. Several discretization schemes are tested within the MLMC method. The Cox–Ingersoll–Ross model, see (3.62), has been implemented with the Milstein scheme, which is new to the literature. Furthermore, results for the Euler, Milstein, and partially QE discretization schemes for the Heston model are implemented and discussed.

Table 4.5: Computational costs of CIR MLMC with the Milstein scheme

eps	value	mlmc cost	std cost	savings	#levels
$5.000 \cdot 10^{-4}$	$3.008 \cdot 10^{-1}$	$2.617 \cdot 10^{8}$	$4.243 \cdot 10^{9}$	16.21	6
$1.000 \cdot 10^{-3}$	$2.997 \cdot 10^{-1}$	$6.435 \cdot 10^{7}$	$2.652 \cdot 10^{8}$	4.12	6
$2.000 \cdot 10^{-3}$	$3.001 \cdot 10^{-1}$	$1.607 \cdot 10^{7}$	$6.629 \cdot 10^{7}$	4.13	6
$5.000 \cdot 10^{-3}$	$3.046 \cdot 10^{-1}$	$2.482 \cdot 10^{6}$	$2.676 \cdot 10^{6}$	1.08	5
$1.000 \cdot 10^{-2}$	$3.098 \cdot 10^{-1}$	$6.198 \cdot 10^{5}$	$6.690 \cdot 10^{5}$	1.08	5

Option pricing with the CIR process is tested for a given accuracy level of the option value ($\kappa = 5$ $\epsilon = 0.25$, $\theta = v_0 = 100$). This gives the following results in Table 4.5. Here *eps* is the maximum error level that has to be attained and is chosen as the input. The *value*, is the value of the option, *mlmc costs* are the total number of time steps that need to be calculated (total of time steps per path times the number of paths), *std costs* are the same costs associated with the standard Monte Carlo computation for the same accuracy level, *savings* is the factor of savings in costs of the MLMC method in comparison with the standard Monte Carlo method, and #levels is the number of different levels[3] that are required to obtain the accuracy *eps*. It is clear that the MLMC substantially reduces the computational costs. The higher the desired accuracy level, the higher the savings in comparison with the standard MC method. It can also be seen that the number of levels increases as the accuracy increases and in accordance with the theory.

Table 4.6: CIR MLMC expectations with the Milstein scheme

l	ave$(V_f - V_c)$	ave(V_f)	Var$(V_f - V_c)$	Var(V_f)	kurtosis	cost
1	$-4.597 \cdot 10^{-1}$	$4.894 \cdot 10^{-1}$	1.804	$5.110 \cdot 10^{-1}$	5.11	4.00
2	$-1.625 \cdot 10^{-1}$	$3.266 \cdot 10^{-1}$	$2.053 \cdot 10^{-1}$	$2.295 \cdot 10^{-1}$	5.32	$1.60 \cdot 10^{1}$
3	$-2.056 \cdot 10^{-2}$	$3.059 \cdot 10^{-1}$	$4.849 \cdot 10^{-3}$	$2.020 \cdot 10^{-1}$	5.54	$6.40 \cdot 10^{1}$
4	$-4.564 \cdot 10^{-3}$	$3.018 \cdot 10^{-1}$	$2.516 \cdot 10^{-4}$	$1.960 \cdot 10^{-1}$	5.59	$2.56 \cdot 10^{2}$
5	$-1.107 \cdot 10^{-3}$	$3.006 \cdot 10^{-1}$	$1.506 \cdot 10^{-5}$	$1.942 \cdot 10^{-1}$	5.57	$1.02 \cdot 10^{3}$

Another experiment is performed with the CIR process with results in Table 4.6. Here, the accuracy level is fixed, and the differences per level l[4] are investigated. Here *ave$(V_f - V_c)$* is the average difference

[3] Level $l = 0$ is the coarsest level.

[4] Level $l = 0$ is not included, as no coarser mesh is available for the computations

Table 4.7: Heston MLMC computational costs with the Euler scheme

eps	value	mlmc cost	std cost	savings	#levels
$5.000 \cdot 10^{-4}$	$1.046 \cdot 10^{1}$	$3.452 \cdot 10^{10}$	$1.048 \cdot 10^{12}$	30.35	4
$1.000 \cdot 10^{-3}$	$1.046 \cdot 10^{1}$	$8.633 \cdot 10^{9}$	$2.619 \cdot 10^{11}$	30.34	4
$2.000 \cdot 10^{-3}$	$1.046 \cdot 10^{1}$	$1.686 \cdot 10^{9}$	$1.637 \cdot 10^{10}$	9.71	4
$5.000 \cdot 10^{-3}$	$1.046 \cdot 10^{1}$	$2.698 \cdot 10^{8}$	$2.619 \cdot 10^{9}$	9.71	4
$1.000 \cdot 10^{-2}$	$1.046 \cdot 10^{1}$	$5.090 \cdot 10^{7}$	$1.627 \cdot 10^{8}$	3.20	4

Table 4.8: Heston MLMC expectations with the Euler scheme

l	ave$(V_f - V_c)$	ave(V_f)	Var$(V_f - V_c)$	Var(V_f)	kurtosis	cost	
1	$2.6898 \cdot 10^{-1}$	$1.0478 \cdot 10^{1}$	$2.541 \cdot 10^{1}$	$1.812 \cdot 10^{2}$	9.44	$2.37 \cdot 10^{-1}$	4.00
2	$-1.5841 \cdot 10^{-2}$	$1.0463 \cdot 10^{1}$	$2.399 \cdot 10^{1}$	$1.879 \cdot 10^{2}$	$1.02 \cdot 10^{1}$	$1.99 \cdot 10^{-2}$	$1.60 \cdot 10^{1}$
3	$1.0665 \cdot 10^{-2}$	$1.0454 \cdot 10^{1}$	2.730	$1.907 \cdot 10^{2}$	6.96	$3.21 \cdot 10^{-1}$	$6.40 \cdot 10^{1}$
4	$2.1454 \cdot 10^{-3}$	$1.0474 \cdot 10^{1}$	$5.668 \cdot 10^{-1}$	$1.918 \cdot 10^{2}$	5.56	$2.91 \cdot 10^{-1}$	$2.56 \cdot 10^{2}$
5	$8.7639 \cdot 10^{-4}$	$1.0464 \cdot 10^{1}$	$1.366 \cdot 10^{-1}$	$1.918 \cdot 10^{2}$	5.46	$1.81 \cdot 10^{-1}$	$1.02 \cdot 10^{3}$

in the option price between two levels (V_f price fine mesh, V_c price coarse mesh) with the same underlying Wiener process (which is the estimator for $\mathbb{E}[V_l - V_{l-1}]$), *ave($V_f$)* is the average price of the option for simulations on level f, which is an estimator for $\mathbb{E}[V_l]$, *Var(V_f-V_c)*, and *Var(V_f)* denotes the corresponding variance. The *kurtosis* of $V_f - V_c$ is a measure of how a probability distribution is distributed, and is computed by,

$$\frac{\sum (V_f - V_c)^4 - 4\sum (V_f - V_c)^3 \sum (V_f - V_c) + 6\sum (V_f - V_c)^2 \sum (V_f - V_c)^2 - 3\sum (V_f - V_c)(\sum (V_f - V_c))^3}{(\sum (V_f - V_c)^2 - (\sum (V_f - V_c))^2)^2}, \tag{4.63}$$

and *cost* is the number of fine time steps per path. The difference between the computations of two subsequent levels are becoming smaller by about factor of four. The associated variance also becomes smaller with an even higher factor. Furthermore, the computation costs per path are increasing with $M = 4$ as predicted by the theory.

For the Heston model, the price of an option, with FFT price 1.046, has been determined with MLMC ($\kappa = 5$, $\sigma = 0.25$, $\rho = -0.5$, $\theta = 0.04$). The same experiments have been conducted as with the CIR model. The Euler discretization scheme gives results in Table 4.7 and Table 4.8. The Milstein discretization scheme gives results in Table 4.9 and Table 4.10. In Figure 4.5 one can see various functions for the Heston Milstein MLMC scheme. Subplots (1,1) and (1,2) show the variance and mean of V_l and $V_l - V_{l-1}$. Subplot (2,1) shows the costs per sample (linear function), and subplot (2,2) shows the kurtosis as previously defined per level. Subplot (3,1) shows the number of simulations per level per accuracy level. Subplot (3,2) shows the costs per accuracy level and makes a comparison between the standard Monte Carlo method and the MLMC. The figures for the CIR process, Euler Heston, and QE Heston can be found in Appendix D. It can be seen that the Milstein scheme substantially reduces computational costs in comparison with the Euler scheme. It can also be noted that the Euler scheme improves the computational time more in comparison with the standard MC method in comparison with the Milstein scheme.

Table 4.9: Heston MLMC computational costs with the Milstein scheme

eps	value	mlmc cost	std cost	savings	#levels
$5.000 \cdot 10^{-3}$	$1.046 \cdot 10^{1}$	$6.418 \cdot 10^{7}$	$6.545 \cdot 10^{8}$	10.20	4
$1.000 \cdot 10^{-2}$	$1.048 \cdot 10^{1}$	$1.608 \cdot 10^{7}$	$1.636 \cdot 10^{8}$	10.18	4
$2.000 \cdot 10^{-2}$	$1.045 \cdot 10^{1}$	$4.008 \cdot 10^{6}$	$4.090 \cdot 10^{7}$	10.21	4
$5.000 \cdot 10^{-2}$	$1.045 \cdot 10^{1}$	$3.203 \cdot 10^{5}$	$1.639 \cdot 10^{6}$	5.12	3
$1.000 \cdot 10^{-1}$	$1.046 \cdot 10^{1}$	$7.573 \cdot 10^{4}$	$4.097 \cdot 10^{5}$	5.41	3

Finally, the QE scheme has been tested with the same parameters as the Euler and Milstein schemes. The results are shown in Table 4.11. The QE MLMC cannot be performed in the same way as the MLMC

Figure 4.5: Heston MLMC with the Milstein scheme

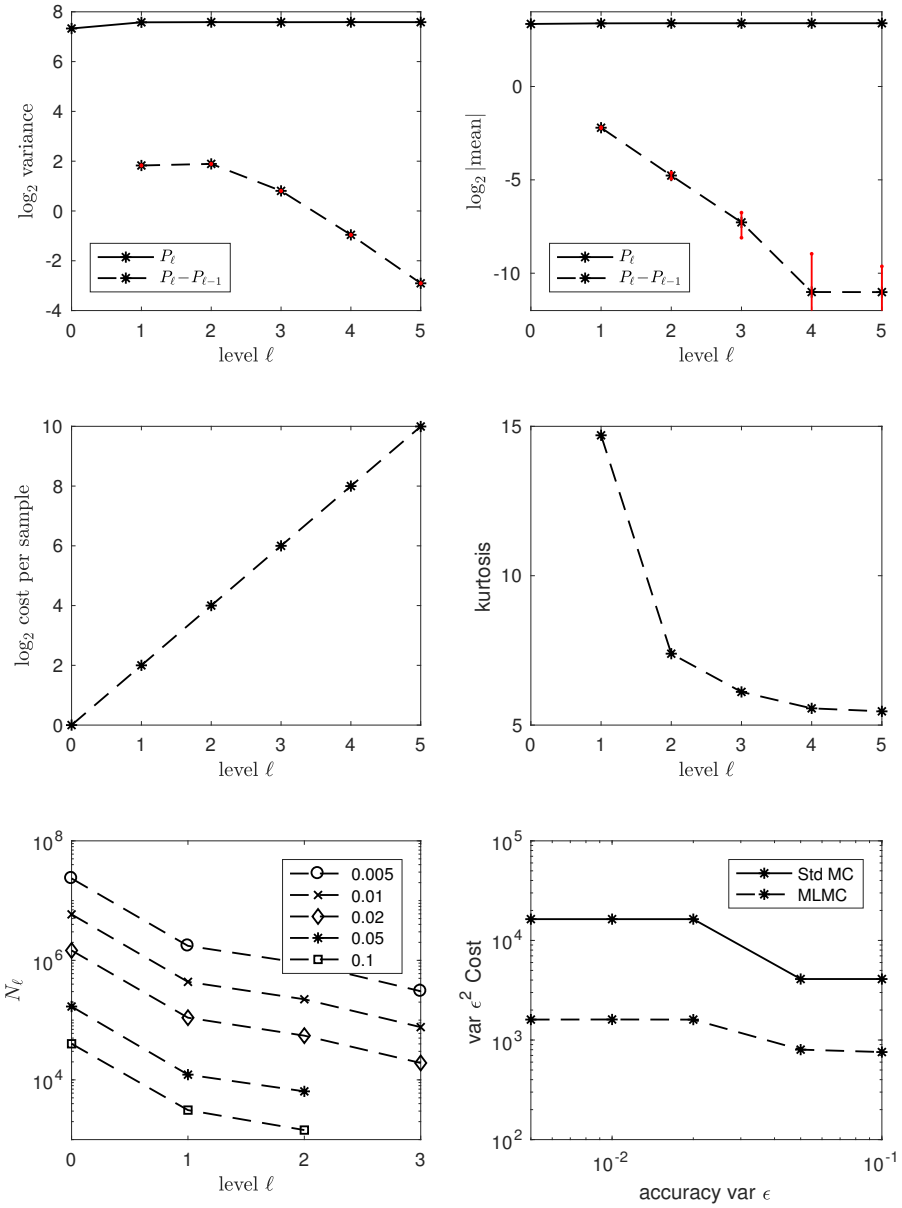


Table 4.10: Heston MLMC expectations with the Milstein scheme

l	ave$(V_f - V_c)$	ave(V_f)	Var$(V_f - V_c)$	Var(V_f)	kurtosis	cost
1	$2.164 \cdot 10^{-1}$	$1.043 \cdot 10^{1}$	3.548	$1.911 \cdot 10^{2}$	$1.47 \cdot 10^{1}$	4.00
2	$3.666 \cdot 10^{-2}$	$1.047 \cdot 10^{1}$	3.709	$1.920 \cdot 10^{2}$	7.39	$1.60 \cdot 10^{1}$
3	$6.450 \cdot 10^{-3}$	$1.046 \cdot 10^{1}$	1.750	$1.917 \cdot 10^{2}$	6.11	$6.40 \cdot 10^{1}$
4	$4.859 \cdot 10^{-4}$	$1.047 \cdot 10^{1}$	$5.147 \cdot 10^{-1}$	$1.921 \cdot 10^{2}$	5.56	$2.56 \cdot 10^{2}$
5	$4.841 \cdot 10^{-4}$	$1.046 \cdot 10^{1}$	$1.335 \cdot 10^{-1}$	$1.919 \cdot 10^{2}$	5.46	$1.02 \cdot 10^{3}$

works for the Euler and Milstein schemes. This is because the QE scheme does not allow to determine Wiener processes on the fine mesh and sum these M processes to a new Wiener process for the coarse mesh. The uniform distribution can be translated to the normal distribution as has been done in these calculations. However, the decision rule (Step 4 or Step 5 in 4.4) to determine if the variance process needs to be computed with the algorithm for small or large values is usually not the same for the M fine processes that combined form the coarse mesh. Therefore, it is not directly clear with which algorithm the coarse mesh should work. This likely causes the non-convergence of the algorithm. Here, $ave(V_f - V_c)$ does not decrease fast enough for convergence. Therefore, this method is not working theoretically. In practice it continues to finer levels and does not end. Altogether, the Milstein scheme for MLMC clearly outperforms the Euler and QE schemes with or without MLMC.

Table 4.11: Heston MLMC expectations with the QE scheme

l	ave$(V_f - V_c)$	ave(V_f)	Var$(V_f - V_c)$	Var(V_f)	kurtosis	cost
1	-9.620	$1.415 \cdot 10^1$	$2.967 \cdot 10^2$	$2.609 \cdot 10^2$	$1.21 \cdot 10^1$	4.00
2	-7.955	$1.404 \cdot 10^1$	$2.818 \cdot 10^2$	$2.574 \cdot 10^2$	$1.54 \cdot 10^1$	$1.60 \cdot 10^1$
3	-6.765	$1.405 \cdot 10^1$	$2.724 \cdot 10^2$	$2.572 \cdot 10^2$	$1.90 \cdot 10^1$	$6.40 \cdot 10^1$
4	-6.254	$1.405 \cdot 10^1$	$2.693 \cdot 10^2$	$2.573 \cdot 10^2$	$1.82 \cdot 10^1$	$2.56 \cdot 10^2$
5	-6.069	$1.405 \cdot 10^1$	$2.682 \cdot 10^2$	$2.575 \cdot 10^2$	$1.84 \cdot 10^1$	$1.02 \cdot 10^3$

4.7 Forward Start Option Pricing

This section explains how to price forward start options for the Heston model, and is based on an article of Oosterlee [27]. First the pricing of these options with the Black-Scholes model is explained to introduce the concept. Tests have been performed with the Black-Scholes and with the Heston model to price these options. Expanded tests are however discussed in Chapter 6. First, the definition of a forward start call option is given.

Definition 4.7.1 (Forward Start Call Option)**.** *A forward start call option is a call option that starts in future date T_1. Therefore, S_{T_1}, the price of the underlying asset at T_1, is unknown. The strike K is expressed as the change of S_{T_2} relative to S_{T_1}. This gives the following payoff function at maturity T_2,*

$$C_{FS}(t_0, T_1, T_2) = \max\left(\frac{S_{T_2}}{S_{T_1}} - (K+1), 0\right). \tag{4.64}$$

This last equation is also written as

$$C_{FS}(t_0, T_1, T_2) = \max\left(\frac{S_{T_2}}{S_{T_1}} - k, 0\right), \tag{4.65}$$

with $k = K + 1$.

This derivative is the basis for more complex derivatives, such as cliquet notes. Cliquet notes consist of a series of forward start options [30]. Investors can be attracted to this product by the ability to periodically lock in profits. The value at t_0 can be expressed by

$$C_{FS}(t_0, T_1, T_2) = \mathbb{E}^{\mathbb{Q}}\left[D(t_0, T_2)\max\left(\frac{S_{T_2}}{S_{T_1}} - (K+1), 0\right) \,\middle|\, \mathcal{F}(t_0)\right]. \tag{4.66}$$

Assuming constant discounting (i.e., deterministic interest rates) gives

$$C_{FS}(t_0, T_1, T_2) = D(t_0, T_2)\mathbb{E}^{\mathbb{Q}}\left[\max\left(\frac{S_{T_2}}{S_{T_1}} - (K+1), 0\right) \,\middle|\, \mathcal{F}(t_0)\right]. \tag{4.67}$$

Define $x(T_1, T_2) = \log \frac{S_{T_2}}{S_{T_1}} = \log S_{T_2} - \log S_{T_1}$. The characteristic function is then given by:

$$\phi_x(u, x(T_1, T_2), t_0, T_2) = \mathbb{E}^{\mathbb{Q}}\left[e^{iux(T_1,T_2)} \,\middle|\, \mathcal{F}(t_0)\right]. \tag{4.68}$$

This gives

$$\phi_x(u, x(T_1, T_2), t_0, T_2) = \mathbb{E}^{\mathbb{Q}}\left[\mathbb{E}^{\mathbb{Q}}\left[e^{iux(T_1,T_2)} \;\middle|\; \mathcal{F}(T_1)\right] \;\middle|\; \mathcal{F}(t_0)\right]. \tag{4.69}$$

By discounting and assuming constant interest rates r, this equation can be rewritten, using the fact that $\log S_{T_1}$ is measurable with respect to the filtration $\mathcal{F}(T_1)$, as

$$\phi_x(u, x(T_1, T_2), t_0, T_2) = \mathbb{E}^{\mathbb{Q}}\left[e^{r(t_0-T_1)}e^{-iux(T_1)}\mathbb{E}^{\mathbb{Q}}\left[e^{r(T_1-T_2)}e^{iu\log S_{T_2}} \;\middle|\; \mathcal{F}(T_1)\right] \;\middle|\; \mathcal{F}(t_0)\right], \tag{4.70}$$

$\log(S_{T_2}) = X(T_2)$ has been filled in here. Also,

$$\mathbb{E}^{\mathbb{Q}}\left[e^{r(T_1-T_2)}e^{iu\log S_{T_2}} \;\middle|\; \mathcal{F}(T_1)\right] = \phi_X(u, X, T_1, T_2). \tag{4.71}$$

The right-hand side of (4.71) under the Black-Scholes model is given by:

$$\phi_X(u, X, T_1, T_2) = exp(r - \frac{\sigma^2}{2})iu(T_2 - T_1) - \frac{1}{2}\sigma^2u^2(T_2 - T_1) - r(T_2 - T_1) + iuX(T_1), \tag{4.72}$$

which gives

$$\phi_x(u) = \mathbb{E}^{\mathbb{Q}}\left[e^{(r-\sigma^2/2)iu(T_2-T_1)-1/2\sigma^2u^2\tau} \;\middle|\; \mathcal{F}(t_0)\right] = e^{(r-\sigma^2/)iu(T_2-T_1)-1/2\sigma^2u^2(T_2-T_1)}. \tag{4.73}$$

Note that this characteristic function is equal to the characteristic function of a normal density.

Then,

$$\begin{aligned} C_{FS}(t_0, T_1, T_2) &= e^{-rT_2}\mathbb{E}^{\mathbb{Q}}\left[\max\left(\frac{S_{T_2}}{S_{T_1}} - k, 0\right) \;\middle|\; \mathcal{F}(t_0)\right] \\ &= e^{-rT_2}\int_{-\infty}^{\infty} \max\left(e^{(r-\sigma^2/2)(T_2-T_1)+\sigma\sqrt{T_2-T_1}x} - k, 0\right) f_{\mathcal{N}(0,1)}(x)dx \\ &= e^{-rT_2}\int_{z}^{\infty} \left(e^{(r-\sigma^2/2)(T_2-T_1)+\sigma\sqrt{T_2-T_1}x} - k\right) f_{\mathcal{N}(0,1)}(x)dx \\ &= e^{-rT_2}\int_{z}^{\infty} e^{(r-\sigma^2/2)(T_2-T_1)+\sigma\sqrt{T_2-T_1}x} f_{\mathcal{N}(0,1)}(x)dx + ke^{-rT_2}(f_{\mathcal{N}(0,1)}(z) - 1), \end{aligned} \tag{4.74}$$

with $f_{\mathcal{N}(0,1)}$ the probability density function of a standard normal variable and with,

$$z = \frac{\log k - (r - \sigma^2/2)(T_2 - T_1)}{\sigma\sqrt{T_2 - T_1}}. \tag{4.75}$$

By using properties of the standard normal distribution, the integral can be rewritten as follows:

$$\begin{aligned} C_{FS}(t_0, T_1, T_2) &= e^{-rT_2}\int_{z}^{\infty} e^{(r-\sigma^2/2)(T_2-T_1)+\sigma\sqrt{T_2-T_1}x} f_{\mathcal{N}(0,1)}(x)dx + ke^{-rT_2}(F_{\mathcal{N}(0,1)}(z) - 1) \\ &= e^{-rT_2}\frac{e^{(r-\sigma^2/2)(T_2-T_1)+\sigma^2(T_2-T_1)/2}}{\sqrt{2\pi}}\int_{z}^{\infty} e^{(\sigma\sqrt{(T_2-T_1)}x-x)^2/2}dx + ke^{-rT_2}(F_{\mathcal{N}(0,1)}(z) - 1) \\ &= e^{-rT_1}(F_{\mathcal{N}(0,1)} - 1(-z + \sigma\sqrt{T_2 - T_1})) + ke^{-rT_2}(F_{\mathcal{N}(0,1)}(z) - 1), \end{aligned} \tag{4.76}$$

Therefore,

$$C_{FS}(t_0, S, T_1, T_2) = e^{-rT_2}\mathbb{E}^{\mathbb{Q}}\left[\max\left(\frac{S_{T_2}}{S_{T_1}} - k, 0\right) \;\middle|\; \mathcal{F}(t_0)\right] = e^{-rT_1}F_{\mathcal{N}(0,1)}(d_1) - ke^{-rT_2}F_{\mathcal{N}(0,1)}(d_2), \tag{4.77}$$

with,

$$d_1 = \frac{\log(1/k) + (r + \sigma^2/2)(T_2 - T_1)}{\sigma\sqrt{T_2 - T_2}}, \tag{4.78}$$

and,

$$d_2 = \frac{\log(1/k) + (r - \sigma^2/2)(T_2 - T_1)}{\sigma\sqrt{T_2 - T_2}}. \tag{4.79}$$

Under the Heston model, a similar derivation gives a way to price forward start options. However, the characteristic function is different under the Heston dynamics and is given by [27],

$$\phi_X(u, X, T_1, T_2) = e^{\bar{A}(u,T_2-T_1)+\bar{B}(u,T_2-T_1)X(T_1)+\bar{C}(u,T_2-T_1)v_{T_1}}. \tag{4.80}$$

The function $\bar{A}$, $\bar{B}$, and $\bar{C}$ are all complex-valued functions and defined by:

$$\bar{A}(\tau) = r(iu-1)\tau + \frac{\kappa\theta\tau}{\varepsilon^2}(\kappa - \varepsilon\rho iu - D_1) - \frac{2\kappa\theta}{\varepsilon^2}\log\left(\frac{1-ge^{-D_1\tau}}{1-g}\right), \tag{4.81}$$

$$\bar{B}(\tau) = iu, \tag{4.82}$$

$$\bar{C}(\tau) = \frac{1-e^{-D_1\tau}}{\varepsilon^2(1-ge^{-D_1\tau})}(\kappa - \varepsilon\rho iu - D_1), \tag{4.83}$$

with

$$D_1 = \sqrt{(\kappa - \varepsilon\rho iu)^2 + (u^2 + iu)\varepsilon^2}, \tag{4.84}$$

$$g = \frac{\kappa - \varepsilon\rho iu - D_1}{\kappa - \varepsilon\rho iu + D_1}. \tag{4.85}$$

This function directly gives

$$\phi_x(u) = \mathbb{E}^{\mathbb{Q}}\left[e^{-iuX(T_1)}e^{r(T_2-T_1)}e^{\bar{A}(u,T_2-T_1)+\bar{B}(u,T_2-T_1)X(T_1)+\bar{C}(u,T_2-T_1)v_{T_1}} \,\middle|\, \mathcal{F}(t_0)\right]. \tag{4.86}$$

Again, $\bar{B}(u, T_2 - T_1) = iu$, so

$$\phi_x(u) = e^{\bar{A}(u,T_2-T_1)+r(T_2-T_1)}\mathbb{E}^{\mathbb{Q}}\left[e^{\bar{C}(u,T_2-T_1)v_{T_1}} \,\middle|\, \mathcal{F}(t_0)\right]. \tag{4.87}$$

The determination of the conditional expectation on the right-hand side is not trivial. The following theorem about the CIR process (3.62) lead to an expression for this expectation.

Theorem 4.7.1 (Moment-generating Function for the CIR Process)**.** *The moment-generating function for the CIR process is of the closed form*

$$\mathbb{E}^{\mathbb{Q}}\left[e^{uv_t} \mid \mathcal{F}\right] = \left(\frac{1}{1-2u\bar{c}_t}\right)^{\bar{d}/2} exp\left(\frac{u\bar{c}_t\bar{\varepsilon}_t}{1-2u\bar{c}_t}\right), \tag{4.88}$$

with

$$\bar{c}_t = \frac{\varepsilon^2}{4\kappa}(1-e^{-\kappa t}), \quad \bar{d} = \frac{4\kappa\bar{v}}{\varepsilon^2}, \quad \bar{\varepsilon}_t = \frac{4\kappa v_{t_0}e^{-\kappa t}}{\varepsilon^2(1-e^{-\kappa t})}. \tag{4.89}$$

Proof. For a detailed proof, see Appendix A.4. □

Finally, by using this theorem, the characteristic function can be described by:

$$\begin{aligned}\phi_x(t) &= e^{\bar{A}(u,T_2-T_1)+r(T_2-T_1)}\mathbb{E}^{\mathbb{Q}}\left[e^{\bar{C}(u,T_2-T_1)v_{T_1}} \,\middle|\, \mathcal{F}\right] \\ &= exp\left(\bar{A}(u,T_2-T_1) + r(T_2-T_1) + \frac{\bar{C}(u,T_2-T_1)\bar{c}(T_1)\varepsilon(T_1)}{1-2\bar{C}(u,T_2-T_1)\bar{c}(T_1)}\right)\left(\frac{1}{1-2\bar{C}(u,T_2-T_1)\bar{c}(T_1)}\right)^{\bar{d}/2}.\end{aligned} \tag{4.90}$$

For an elaborated derivation, see [27]. Now, the COS method can be applied to price the forward start option since the characteristic function is known.

Chapter 5

Numerical Techniques for the Local Stochastic Volatility Model

This chapter elaborates on techniques to calibrate the local volatility surface of the LSV model. The local volatility surface of Dupire is required for these computations. First techniques to interpolate and extrapolate grid points are presented. These techniques can be applied to grid points of the local volatility surface of Dupire, as well as of the local volatility surface belonging to the LSV model. Techniques to determine the local volatility function for the LSV model are discussed, and the Particle and Bin method are explained into further detail.

5.1 Interpolation

The values at grid points of the local volatility surface can be used for interpolation to obtain a smooth local volatility surface. This can be done in several ways. It can be done by way of linear interpolation. This method creates a continuous surface which is not continuously differentiable. For the calculation of greeks, this might not be desirable.

The technique of cubic splines creates a smooth local volatility surface out of the values on the grid. In MATLAB R2017A the command *cubicinterp* in the *fit* framework involves cubic splines. It creates a smooth continuous differentiable surface. The disadvantage of this method is that the result might be counterintuitive and might lead to arbitrage opportunities. Between two grid points (a, x) and (b, y) with $a < b$ and $x < y$, values (c, z) can arise with $a < c < b$ and $z < x < y$ which is counterintuitive.

Arbitrage can occur if the total variance surface does not fulfill the criterion of non-arbitrage. Call-spread arbitrage (Equation (3.3)) needs to be avoided and monotonicity in the direction of the strike needs to be satisfied (Equation (3.11)). These criteria can be checked by determining the surface of the total variance, i.e., by checking that the total variance is increasing for all expiry times.

5.1.1 Monotonic Cubic Splines

To arrive at an arbitrage-free surface, the total variance can be determined first on all grid points. This total variance surface can then be interpolated using monotonic cubic splines, and the local volatility function can be determined by taking the inverse. In MATLAB R2017A, the function *pchip* can be used for monotonic cubic splines. All experiments in the result section imply no arbitrage opportunities. Therefore, all local volatility surfaces are converted to the total variance surfaces to make sure that no arbitrage can occur.

Figure 5.1: Monotonic cubic splines

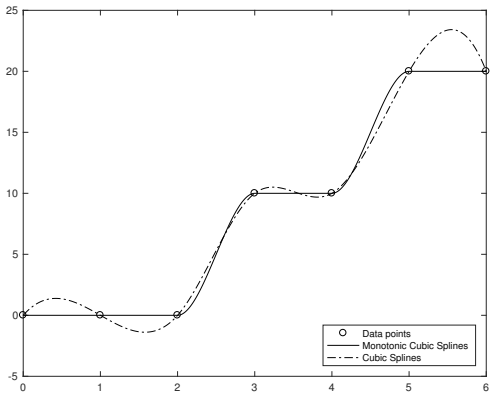


Figure 5.1 shows the difference between monotonic and standard cubic splines. MATLAB states that the monotonic cubic interpolation function makes sure that the first derivative is continuous, however the second might not be continuous. The extremum in the data is preserved, as is the monotonicity. Therefore, for all numerical experiments in this thesis, this method is used.

5.2 Extrapolation

Once a volatility surface is interpolated, the surface is defined in the t direction and the S direction. Due to the limited number of call option prices that are used for the calibration, this surface is bounded. If $\sigma(S,t)$ is given for all $t \in [t_{\min}, t_{\max}]$, and for $S \in [S_{\min}, S_{\max}]$, then simulating asset prices out of this domain is not possible with Dupire's model or the LSV model.

In Dupire's model, one can choose a flat volatility out of the given domain, and set here the value of the local volatility function equal to the average volatility in the domain. Another option is to choose the volatility function out of this domain to be equal to the expectation of the volatility in the Heston model. The variance process is then described by the CIR process:

$$dv_t = \kappa(\theta - v_t)dt + \varepsilon\sqrt{v_t}dW_t^v. \tag{5.1}$$

This gives expected variance only depending on t:

$$\mathbb{E}[v_t|v_0] = v_0 e^{-\kappa t} + \theta(1 - e^{-\kappa t}). \tag{5.2}$$

For the volatility surface of the LSV model, multiple solutions are possible as well. It is possible to set $\sigma(S,t) = 1$ for $(S,t) \notin (D_x, D_t)$, with $D_x = [S_{\min}, S_{\max}]$, and $D_t = [t_{\min}, t_{\max}]$. Out of the calibrated domain, the model then behaves like a SVM, here, as the Heston model. Another option is linear extrapolation. In all experiments this thesis, unless explicitly stated otherwise, this has been applied. The values of $\sigma(S,t)$ for $S \notin D_x$ and $S < S_{\min}$ are equal to $\sigma(S_{\min}, t)$. The values of $\sigma(S,t)$ for $S \notin D_x$ and $S > S_{\max}$ are equal to $\sigma(S_{\max}, t)$. The same holds for $t \notin D_x$, with $t_{\min}$, and $t_{\max}$.

5.3 Henry-Labordère's Approach in the Determination of the Conditional Expectation

Using the definition of the effective local volatility,

$$\sigma_{loc}(S_t,t)^2 = \sigma(S_t,t)^2 \mathbb{E}^{\mathbb{Q}^t}[a_t^2|S_t = K]. \tag{5.3}$$

Then,

$$\frac{dS_t}{S_t} = \sigma(S_t,t)a_t d\boldsymbol{W}_t, \tag{5.4}$$

with $d\boldsymbol{W}_t = (\rho dW_t^1 + \sqrt{1-\rho^2}dW_t^2)$, can be rewritten as:

$$\frac{dS_t}{S_t} = \frac{\sigma_{loc}(S_t,t)}{\sqrt{\mathbb{E}^{\mathbb{Q}}[a_t^2|S_t = K]}} a_t d\boldsymbol{W}_t. \tag{5.5}$$

Using a Malliavin representation, in the case $\sigma(S_t,t) = 1$, the effective local volatility function can be written as

$$\sigma_{loc}^{SV}(S_t,t)^2 = \frac{\mathbb{E}^{\mathbb{Q}}\left[a_t^2 \frac{e^{-\frac{U^2}{2(1-\rho^2)\int_0^t a_s^2 ds}}}{\sqrt{\int_0^t a_s^2 ds}}\right]}{\mathbb{E}^{\mathbb{Q}}\left[\frac{e^{-\frac{U^2}{2(1-\rho^2)\int_0^t a_s^2 ds}}}{\sqrt{\int_0^t a_s^2 ds}}\right]}, \tag{5.6}$$

with $U = \log\left(\frac{S}{S_0}\right) + \frac{1}{2}\int_0^t a_s^2 ds - \rho\int_0^t a_s dW_s$. Here, σ_{loc}^{SV} stands for σ_{loc} conditional on $\sigma(S_t,t) = 1$, so the process simplifies to a standard stochastic volatility process. For a detailed derivation see [20].

With the Monte Carlo method, the right-hand-side of Equation (5.6) can be computed using a discretization scheme for the process a_t.

Vladimir Piterbarg came up with a framework for mapping a LSV model to a SV model [29]. This idea can be used to find an expression for the local volatility function for the LSV model.

Let S_t be a stochastic volatility process and f_t a local stochastic volatility process given by:

$$dS_t = S_t a_t d\boldsymbol{W}_t, \tag{5.7}$$

$$df_t = f_t \sigma(f_t,t) a_t d\boldsymbol{W}_t. \tag{5.8}$$

Under the assumption that a smooth monotone mapping $S_t \mapsto f_t$ exists, this mapping $S_t = M(f_t,t)$ gives, taking $S_0 = f_0$, the following SDE:

$$dS_t = f_t \frac{\partial M(f_t,t)}{\partial f}\sigma(f_t,t)a_t d\boldsymbol{W}_t + \left(\frac{\partial M(f_t,t)}{\partial t} + \frac{1}{2}f_t^2 \frac{\partial^2 M(f_t,t)}{\partial f^2}\sigma(f_t,t)^2 a_t^2\right)dt. \tag{5.9}$$

From Equation (5.7) and the monotone mapping, it follows that Equation (5.9) is equal to $S_t a_t d\boldsymbol{W}_t$. So,

$$M(f_t,t) = f_t \frac{\partial M(f_t,t)}{\partial f}\sigma(f_t,t). \tag{5.10}$$

By definition,

$$\sigma_{loc}^{SV}(K,t)^2 = \mathbb{E}^{\mathbb{Q}^t}[a_t^2|S_t = K]. \tag{5.11}$$

Furthermore,

$$\sigma_{dup}(f,t)^2 = \sigma(f_t,t)^2\mathbb{E}^{\mathbb{Q}^t}[a_t^2|f_t = f]. \tag{5.12}$$

The monotone mapping assures that the conditional expectations in Equation (5.11) and Equation (5.12) are equal for K and f if $K = M(f,t)$. This gives,

$$\sigma(f,t) = \frac{\sigma_{dup}(f,t)}{\sigma_{loc}^{SV}(S_t,t)} = \frac{\sigma_{dup}(f,t)}{\sigma_{loc}^{SV}(M(f_t,t),t)}. \tag{5.13}$$

Equation (5.10) then gives,

$$f\frac{\partial M(f,t)}{\partial f} = \frac{\sigma_{loc}(M(f_t,t),t)}{\sigma_{dup}(f_t,t)}. \tag{5.14}$$

Integrating Equation (5.10) gives,

$$\log\frac{M(f,t)}{\Lambda(t)} = \int_{f_0}^{f}\frac{d\tilde{f}}{\tilde{f}\sigma(t,\tilde{f})}, \tag{5.15}$$

here, $\Lambda(t)$ is an integration constant. Therefore,

$$\sigma(t,f) = \frac{\sigma_{dup}(f,t)}{\sigma_{loc}\left(\Phi_t^{-1}\left(\int_{f_0}^{f}\frac{dx}{x\sigma_{dup}(x,t)}\right),t\right)}, \tag{5.16}$$

with,

$$\Phi_t^{-1} = \int_{\Lambda(t)}^{x}\frac{dz}{z\sigma_{loc}^{SV}(z,t)}. \tag{5.17}$$

$\Lambda(t)$ can be determined using the drift term in Equation (5.9),

$$\begin{aligned}&\left(\frac{\partial M(f_t,t)}{\partial t} + \frac{1}{2}f_t^2\frac{\partial^2 M(f_t,t)}{\partial f^2}\sigma(f_t,t)^2a_t^2\right)dt\\ &= M(f_t,t)\left(\frac{\partial\left(\log\Lambda(t) + \int_{f_0}^{f}\frac{dz}{z\sigma(t,z)}\right)}{\partial t} + \frac{1}{2}a_t^2(1 - \frac{\partial(f_t\sigma(f_t,t))}{\partial f})\right).\end{aligned} \tag{5.18}$$

Henry-Labordère [21], via a refferral to a result of Gyöngy, suggests substituting a_t^2 by $\mathbb{E}^{\mathbb{Q}^T}[a_t^2|f_t = f]$ and f_t by f_0. The approximation of the drift is then as follows:

$$\partial_t\log\Lambda(t) + \frac{1}{2}\sigma_{loc}^{SV}(f_0,t)^2\left(1 - \sigma(f_0,t) - f_0\frac{\partial\sigma(f_0,t)}{\partial f}\right). \tag{5.19}$$

Taking Λ such that this drift equation vanishes leads to the expression,

$$\log\frac{\Lambda(T)}{f_0} = -\frac{1}{2}\int_t^T\sigma_{loc}^{SV}(f_0,t)^2\left(1 - \sigma(f_0,t) - f_0\frac{\partial\sigma(f_0,t)}{\partial f}\right)dt, \tag{5.20}$$

which gives the optimal choice for the integration constant [29][1]. This leads to the following MC based algorithm.

[1] See [29] for a proof.

Monte Carlo Based Algortihm

To calibrate local stochastic volatility models to market smiles, Pierre Henry-Labordère [21] developed a Monte Carlo based approach which consists of three steps:

- *Step 1:* Use the formula of Dupire (Equation (3.61)) to calibrate the function $\sigma_{dup}(f,t)$ to implied volatility in the market.
- *Step 2:* Compute the function $\sigma_{loc}(f,t)$ on a space-time grid by first simulating the stochastic process a_t and then using Equation (5.6).
- *Step 3:* Set $\Lambda(T) = f_0$ and use Equation (5.16) to compute the local volatility function $\sigma(f,t)$. Use the obtained $\sigma(f_0,t)$ and $\frac{\partial\sigma(f_0,t)}{\partial f}$ to compute $\Lambda(T)$ via Equation (5.20). Repeat this process with the new $\Lambda(T)$ until convergence is reached.[2]

5.4 Particle Algorithm

The conditional expectation $\mathbb{E}^{\mathbb{Q}}[a_t^2|S_t = K]$ which arises when finding an expression for the local volatility function in Equation (3.68) can be rewritten by using the probability density function of a_t and S_t. This gives

$$\mathbb{E}^{\mathbb{Q}}[a_t^2|S_t = K] = \frac{\int a_t^2 p(t,K,a)da}{\int p(t,K,a)da}. \tag{5.21}$$

Therefore, simulating $\sigma^2(K,t) = \frac{\sigma_{dup}^2(K,t)}{\mathbb{E}^{\mathbb{Q}}[a_t^2|S_t=K]}$ requires the joint probability density function $p(t,K,a)$ of $(S_t = K, a_t)$.

The Particle algorithm is an algorithm to deal with McKean-Vlasov processes, see Definition 2.0.10, by approximating the law $\mathbb{Q}_t$. The idea behind the Particle algoritm [18] is to simulate the joint density function of (S_t, a_t). The algorithm can be generalized and applied to N processes $X^{i,N}$ which are solutions of an $\mathbb{R}^N$-dimensional linear SDE. The joint density function then can be simulated by:

$$\mathbb{Q}_t^N = \frac{1}{N}\sum_{i=1}^{N} \delta_{X_t^{i,N}}, \tag{5.22}$$

where the choice of $\delta_{X_t^{i,N}}$ is specified in Section 5.4.1, and with $i \in \{1,2,...,N\}$. Here, $X_t^{i,N}$ are solutions to

$$dX_t^{i,N} = \mu(X_t^{i,N}, t, \mathbb{Q}_t^N)dt + \sigma(X_t^{i,N}, t, \mathbb{Q}_t^N)dW_t^i. \tag{5.23}$$

Here, $\mathbb{Q}_t^N$ is a random measure on $\mathbb{R}_t^N$ and $\{W_t^i\}_{i=1,...,N}$ are Wiener processes.

$dX_t^{i,N}$ can be rewritten in the case of the McKean-Vlasov SDE where

$$\mu(x,t,\mathbb{Q}_t^N) = \int \mu(x,t,\mathbb{Q}_t^N)p(y,t)dy = \mathbb{E}[\mu(x,t,X_t)], \tag{5.24}$$

$$\sigma(x,t,\mathbb{Q}_t^N) = \int \sigma(x,t,\mathbb{Q}_t^N)p(y,t)dy = \mathbb{E}[\sigma(x,t,X_t)], \tag{5.25}$$

here, $\{X_t^{i,N}\}_{i=1,...,N}$ are Itô processes given by:

$$dX_t^{i,N} = \left(\int \mu(t,X_t^{i,N},y)d\mathbb{Q}_t^N(y)\right)dt + \left(\int \sigma(t,X_t^{i,N},y)d\mathbb{Q}_t^N(y)\right)dW_t^i. \tag{5.26}$$

[2] Henry-Labordère [21] states in his article that in his experience, no further repetition of the last step is needed to obtain convergence.

This can be discretized as,

$$dX_t^{i,N} = \frac{1}{N}\left(\sum_{j=1}^{N}\mu(t, X_t^{i,N}, X_t^{j,N})dt + \sum_{j=1}^{N}\sigma(t, X_t^{i,N}, X_t^{j,N})dW_t^i\right). \tag{5.27}$$

From the chaos propagation property of Sznitman [36], one can conclude that for a bounded continuous function V, and with $X_0^{i,N}$ independent particles, the following holds for any given $t \geq 0$,

$$\frac{1}{N}\sum_{i=1}^{N} V(X_t^{i,N}) \xrightarrow[N\to\infty]{L^1} \int_{\mathbb{R}^d} f(x)p(x,t)dx. \tag{5.28}$$

In the LSV framework, the following can be simulated:

$$\mathbb{E}^{\mathbb{Q}_t^N}[a_t^2|S_t = K] = \frac{\int \tilde{a}^2 p_N(t,K,\tilde{a})d\tilde{a}}{\int p_N(t,K,\tilde{a})d\tilde{a}} = \frac{\sum_{i=1}^{N}(a_t^{i,N})^2\delta(S_t^{i,N}-K)}{\sum_{i=1}^{N}\delta(S_t^{i,N}-K)}, \tag{5.29}$$

with δ the Dirac function. Instead of the Dirac function, Guyon et al [18] use regularizing kernels $\delta_{t,N}$ (defined in Equation 5.33) to define:

$$\sigma_N(K,t) = \sigma_{dup}(K,t)\sqrt{\frac{\sum_{i=1}^{N}\delta_{t,N}(S_t^{i,N}-K)}{\sum_{i=1}^{N}(a_t^{i,N})^2\delta_{t,N}(S_t^{i,N}-K)}}. \tag{5.30}$$

Then the following SDE can be simulated by the Monte Carlo method:

$$dS_t^{i,N} = S_t^{i,N}\sigma_N(S_t^{i,N}, t)a_t^{i,N}dW_t^i. \tag{5.31}$$

Summarized, the following steps need to be followed.

- *Step 1:* Set $\sigma_N(K,t) = \dfrac{\sigma_{dup}(K,0)}{a_0}$ for all $t \in [t_0 = 0, t_1]$. Here $\{t_k\}$ is the time discretization of $[0,T]$.
- *Step 2:* Use Equation (5.31) to simulate the $\{S_t^{i,N}, a_t^{i,N}\}_{i=1,\ldots,N}$ processes from t_{k-1} to t_k. Start with $k = 1$.
- *Step 3:* Begin by sorting the particles according to the spot values. Find the smallest index $\underline{i}(S)$ and the largest index $\bar{i}(S)$ for $S \in G_{S,t_k}$ (G_{S,t_k} the set of simulated particles) for which $\delta_{t_k,N}(S_{t_k}^{i,N} - K) > \eta$, with η a constant. Then, compute the local volatility according to,

$$\sigma_N(K,t) = \sigma_{dup}(K,t)\sqrt{\frac{\sum_{i=1}^{N}\delta_{t,N}(S_t^{i,N}-K)}{\sum_{i=1}^{N}(a_t^{i,N})^2\delta_{t,N}(S_t^{i,N}-K)}}. \tag{5.32}$$

  Then interpolate this local volatility using cubic splines and extrapolate flat outside the interval $[\min G_{S,t_k}, \max G_{S,t_k}]$. Finally, set for all $t \in [t_k, t_{k+1}]$ the volatility surface $\sigma_N(K,t) := \sigma_N(K,t_k)$.
- *Step 4:* Repeat the previous two steps with $k := k+1$ until maturity.

5.4.1 Regularizing Kernels

Regularizing kernels can be used to express $\delta_{t,N}(x)$. This leads to,

$$\delta_{t,N}(x) = \frac{1}{h_{t,N}}Ke\left(\frac{x}{h_{t,N}}\right). \tag{5.33}$$

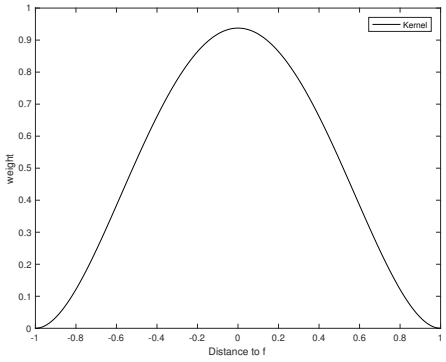
Here $h_{t,N}$ tends to zero as N goes to infinity. Furthermore, Ke is a fixed, symmetric kernel with bandwidth $h_{t,N}$. The following kernel does not require much calculation time and is therefore used for obtaining the results [18],

$$Ke(x) = \frac{15}{16}(1-x^2)^2 \mathbb{1}_{\{|x|\leq 1\}}, \tag{5.34}$$

$$h_{t,N} = \kappa f_0 \sigma_{VS,t} \sqrt{\max(t, t_{\min})} N^{-\frac{1}{5}}, \tag{5.35}$$

Here $\sigma_{VS,t}$ is the variance swap volatility at maturity. This kernel function for the Particle algorithm is shown in Figure 5.2.

Figure 5.2: Kernel function



Guyon et al [18] state that the term $N^{-\frac{1}{5}}$ comes from the minimization of the asymptotic mean integrated squared error of the Nadaraya-Watson estimator. This is the sum of the bias and the variance. They state that a smaller bandwidth leads to a smaller bias and on the other hand, a larger variance. They determined the critical bandwidth that minimizes the sum of bias and variance decreases as $N^{-\frac{1}{5}}$ for large N. This follows from Silverman's rule of thumb [35]. The factor $h_{t,N} = \kappa f_0 \sigma_{VS,t} \sqrt{\max(t, t_{\min})}$, that they take, is of order of the standard deviation of the regressor S_t. They state, that the fine-tuning of kappa is 'crucial' and then take $\kappa = 1.5$, $t_{\min} = \frac{1}{4}$.

5.4.2 Bins

There are several ways to weigh the particles instead of using the Dirac delta function in:

$$\mathbb{Q}_t^N = \frac{1}{N} \sum_{i=1}^{N} \delta_{X_t^{i,N}}, \tag{5.36}$$

by the regularizing kernels as described by Guyon et al [18] or by using so-called bins. Van der Stoep et al [17] describe in their, in 2014, published article the use of bins. They group similar realizations in bins in two different ways. Given N pairs of Monte Carlo realizations, (S_t, a_t), sort the pairs on from smallest S_t to the largest S_t and rename them $(\bar{S}_{t,j}, \bar{a}_{t,j})$, with $j = 1$ denoting the smallest value of S_t of all realizations and $j = l$ denoting the largest value of S_t of all realizations. The two different type of bins are:

1. *Equal-Size Bins* Take bins with respect to an equidistant grid specified on the domain $\bar{S}_{i,1} = B_{i,1} < B_{i,2} < ... < B_{i,l+1} = \bar{S}_{i,N}$.

   With N such that for any $u, v \in \{1, ..., l\}, u \neq v, B_{i,u+1} - B_{i,u} = B_{i,v+1} - B_{i,v}$. So,

$$B_{i,k} = \bar{S}_{i,1} + \frac{k-1}{l}(\bar{S}_{i,N} - \bar{S}_{i,1}), k = 1, ..., l+1. \tag{5.37}$$

2. *Equal Number per Bin* Take bins, such that every bin contains an equal number of Monte Carlo paths by approximation,

$$B_{i,1} = \bar{S}_{i,1}, B_{i,l+1} = \bar{S}_{i,N}, B_{i,k} = \bar{S}_{i,(k-1)N/l}, k+2, ...l. \tag{5.38}$$

If it is not stated differently, this type of bin is assumed in experiments in this thesis.

Then assign every (S_t, a_t) to its bin, take the average value of a_t of all pairs in this bin, and take this value as the expectation of a_t for value of S in this bin. These bins approximate the conditional expectation problem then by,

$$\mathbb{E}^{\mathbb{Q}_t^N}[a_t^2|S_t = K] \approx \frac{\mathbb{E}[a_{t_i}^2 \mathbb{1}_{S_{t_i} \in (b_{i,k}, b_{i,k+1}]}]}{\mathbb{Q}[S_{t_i} \in (b_{i,k}, b_{i,k+1}]]} \approx \frac{\sum_{j=1}^{N} a_{t_i}^2 \mathbb{1}_{S_{t_i} \in (b_{i,k}, b_{i,k+1}]}}{N\alpha(k)}. \tag{5.39}$$

Here $\alpha(k) = \mathbb{Q}[S_{t_i} \in (b_{i,k}, b_{i,k+1}]]$ is the probability of being in the kth bin. In the *Equal Number per Bin* scenario, this is equal to the bin size. An example is given in tables 5.1 and 5.2. The pairs (S_t^i, a_t^i) are numbered, and after the determination of the values of the pairs at $t = 1$, these pairs are assigned to a bin based on the value of S_t, and then the conditional expectation is calculated as the arithmetic average.

Table 5.1: Example pairs

i	$(S_{t=0}^i, a_{t=0}^i)$	$(S_{t=1}^i, a_{t=1}^i)$
1	(100,0.2)	(173,0.08)
2	(100,0.2)	(210,0.18)
3	(100,0.2)	(38,0.51)
4	(100,0.2)	(90,0.29)
5	(100,0.2)	(60,0.40)
6	(100,0.2)	(111,0.28)
7	(100,0.2)	(104,0.14)
8	(100,0.2)	(181,0.04)
9	(100,0.2)	(130,0.18)

Table 5.2: Example bin assignation

Bin	i	$(S_{t=1}^i, a_{t=1}^i)$	$\mathbb{E}[a_{t=1}\|S_{t=1} \in [B_j]]$
B_1	3	(38,0.51)	$(0.51 + 0.29 + 0.40)/3 = 0.4$
B_1	4	(90,0.29)	$(0.51 + 0.29 + 0.40)/3 = 0.4$
B_1	5	(60,0.40)	$(0.51 + 0.29 + 0.40)/3 = 0.4$
B_2	7	(104,0.14)	$(0.14 + 0.28 + 0.18)/3 = 0.2$
B_2	6	(111,0.28)	$(0.14 + 0.28 + 0.18)/3 = 0.2$
B_2	9	(130,0.18)	$(0.14 + 0.28 + 0.18)/3 = 0.2$
B_3	1	(173,0.08)	$(0.08 + 0.04 + 0.18)/3 = 0.1$
B_3	8	(181,0.04)	$(0.08 + 0.04 + 0.18)/3 = 0.1$
B_3	2	(210,0.18)	$(0.08 + 0.04 + 0.18)/3 = 0.1$,

With $B_1 = [0, 90], B_2 = (90, 130], and B_3 = (130, \infty)$.

The function that describes the conditional expectation will be noncontinuous at the points where one bin goes over into the next bin. It is possible to use the midpoints of the bin $(B_i - B_{i-1})/2$ and linear interpolate between this points to create a smooth function. This variant has also been investigated and is called the Smooth Bin method.

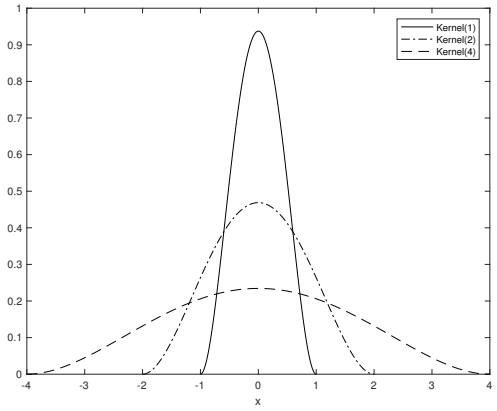
5.4.3 Particle-Bin Method

The Particle and Bin methods are quite similar, i.e., both methods simulate pairs (S_t, a_t) to define a function that describes the conditional expectation for a given t, and then interpolate these functions over time to create a local volatility surface. This makes it possible to combine both methods by taking the average of the two functions to define an alternative function to approximate the conditional expectation. Another variant that is possible, is to first assign all particles to bins, then to weigh all particles in the bin with kernels instead of with equal weights as is done in the standard Bin method. This variant and the standard Particle and standard Bin method are compared in the next section.

5.5 Comparison of the Particle, Bin, and Particle-Bin Method

5.5.1 Numerical Experiments for the Particle Method

Figure 5.3: Conditional expectation for different kernels for the Particle method



The choice of the kernel function Ke (or of the associated bandwidth) is of importance for results of the Particle method. Only pairs out of $[S_t - h_{t,N}, S_t + h_{t,N}]$ are taken into account for the determination of the conditional expectation. If there are not many pairs simulated in this range, the variance of the estimator of the conditional expectation increases. If the kernel is used as defined previously, scaling of S with $S_0 = 100$ and $S_0 = 5$ changes the estimator of the conditional expectation function if the bandwidth $h_{t,N}$ does not change. For the ease of understanding the impact of the bandwidth decision, the kernel function is here defined as follows $Ke(i) = \frac{1}{\sqrt{i}}\left(1 - \left(\frac{x}{i}\right)^2\right)^2$. In Figure 5.3, three different kernels are shown.

Figure 5.4 and Figure 5.5 show the impact of scaling S on the conditional expectation of a_t. They show the estimators of $\mathbb{E}[a_T^2|S_T = K]$ for $\kappa = 5$, $\varepsilon = 0.25$, $\rho = -0.25$, $\theta = 0.1$, $v_0 = 0.1$, $8 \cdot 10^3$ paths, $\Delta = 0.1$, and $t = 1$ (Parameter Set 3). All parameters are kept equal, besides S_0, which is 5 in Figure 5.4 and 100 in Figure 5.5. A small number of simulated paths has been chosen to show this result clearly. The two figures show that the estimator of a is more volatile for larger S_0. On the other hand, it shows that for small values of S_0, the function becomes flatter.

A comparison of the estimates of $\mathbb{E}[a_T^2|S_T = K]$ for $Ke(1)$, $Ke(4)$, and $Ke(8)$ for Parameter Set 3, with $S_0 = 50$, is shown in Figure 5.6. The figure shows that a smaller i in $Ke(i)$ leads to more variance and a more volatile function. It also shows that the differences in the estimator of a are smaller for the at-the-money value of S_t than for values (far) that are in or out the money, and that the minimum is larger for a large number of i in $Ke(i)$.

Figure 5.4: Conditional expectation for the Particle method with $Ke(1)$ and small f

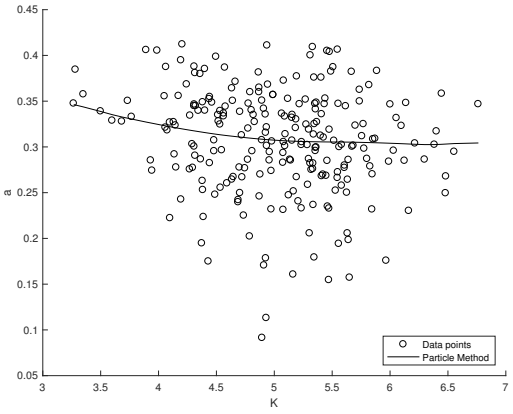
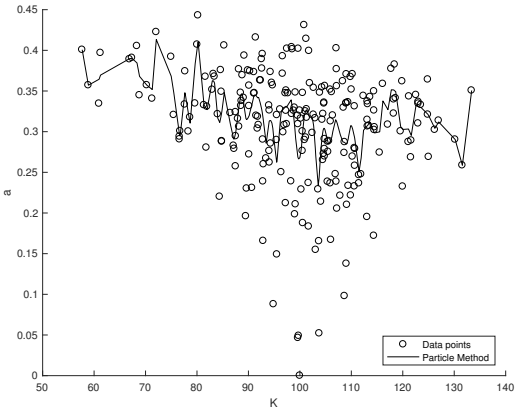
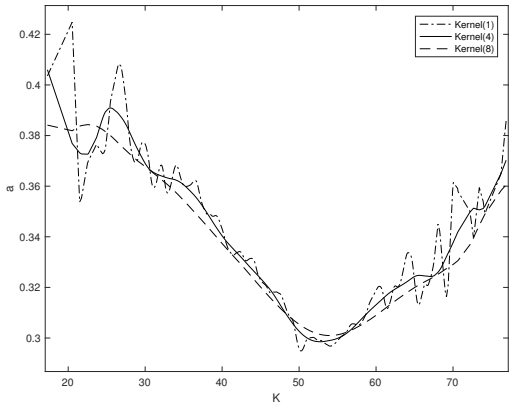


Figure 5.5: Conditional expectation for the Particle method with $Ke(1)$ and large f



To see if it is possible to use a small number i in $Ke(i)$ and then use polynomial curve fitting (*polyfit* in MATLAB R2017A) to reduce the high variability, the following experiment was performed. First two estimates of a_t were made conditional on S_t with different kernels. Then, polynomial curve fitting was applied to these results with a degree of freedom 5. The results for the estimates of $\mathbb{E}[a_T^2|S_T = K]$ are given in Figure 5.7 for Parameter Set 3. It can be seen that the polynomial curve fitting increases the minimum. Furthermore, no large differences are noted.

Figure 5.6: Conditional expectation for different kernels for the Particle method



The number of simulations has to be chosen as well. In Figure 5.8 two estimates of $\mathbb{E}[a_T^2|S_T = K]$ have been performed with one estimate which is created by 10^2 times more simulations than the other. The parameters are from Parameter Set 3. It can be seen that the larger the number of simulations, the less volatile the function becomes.

Figure 5.7: Conditional expectation for different kernels for the Particle method with polyfit

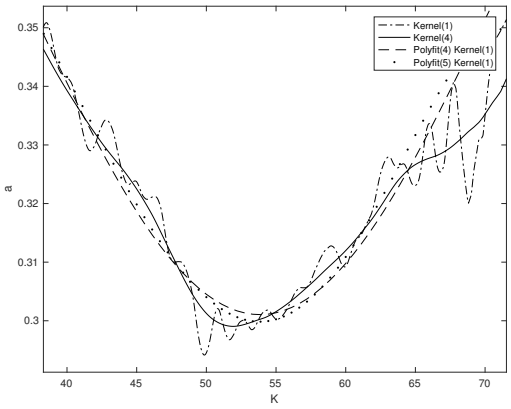
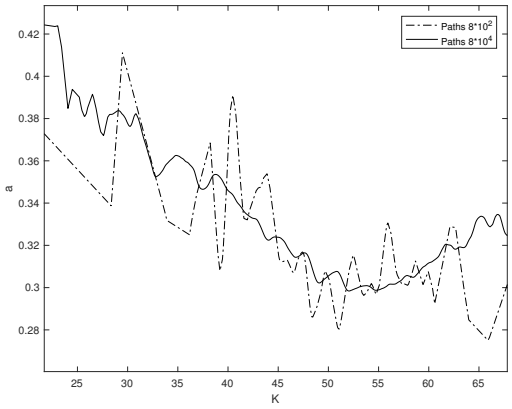


Figure 5.8: Conditional expectation for different numbers of simulations for the Particle method



5.5.2 Numerical Experiments for the Bin Method

The performance of the different types of bins has been investigated by Van der Stoep [17]. The bins, with an equal number of pairs per bin yield a smaller difference between the conditional expectation as calculated by this method and the COS method, than do bins of equal size.

The calculation time for a different numbers of bins (equal number of pairs per bin) is determined and given in Table 5.3, with parameters from Parameter Set 3. It follows that the number of bins has only a small effect on the computational time of the local volatility surface for the LSV model.

Table 5.3: Computation cost for LV Surface determination

Number of bins	2	4	8	16	32	64	128
Computation cost	1	1.003	1.007	1.011	1.021	1.043	1.087

with 'computation cost' defined, relative to the computational time for the 2 bin scenario, as $\frac{CC}{CC_{2Bins}}$ for 10^4 paths, with Computational Cost CC.

One must avoid a situation in which the basket of a_t that is used for the estimation of $\mathbb{E}[a_t^2|S_t = K]$ consists of only zeros ($a_t = 0$). It also should always consist of enough pairs. The issue of values $a_t = 0$ can be avoided by calibrating the Heston model conditional that the Feller condition is satisfied, as is done for the LSV implied volatility calculations. In practice, the Feller condition is usually not satisfied if the Heston model is calibrated without restrictions. Therefore, when simulating, cases of $a = 0$ occur at certain times. The probability that all a values are zero in a bin, is greater if there is a small number of pairs per bin. For the equal-size-of-bins method, a small number of simulated pairs occurs especially often for 'extreme values of S_t', far away from S_0 (out-of-the-money).

To make sure that every bin contains enough pairs, one can increase the number of pairs or decrease the number of bins. Figure 5.9 shows the conditional expectation $\mathbb{E}[a_T^2|S_T = K]$. Here, $8 \cdot 10^3$ paths are used and the number of bins varies. Furthermore $S_0 = 20, T = 0.5, r = 0, v_0 = 0.1, \kappa = 5, \theta = 0.1, \rho = -0.25$, and $\varepsilon = 0.25$ (Parameter Set 4). Note that there is a large difference in the expectation for values far away for S_0 (far in or out of the money). This is also observed for the Particle method.

In Figure 5.10, $8 \cdot 10^3$ paths are simulated to conduct the same experiment to determine $\mathbb{E}[a_T^2|S_T = K]$, but this time the Smooth Bin method is used with parameters from Parameter Set 4. Here, the functions are made continuous by taking the mid points of the bins and using linear interpolation between these points. It can be seen that, in all tests, a larger number of bins increases the conditional expectation in the region of the two bins that are farthest away from S_0 (far in or out the money). One can conclude

Figure 5.9: Conditional expectation for the Bin method with different numbers of bins

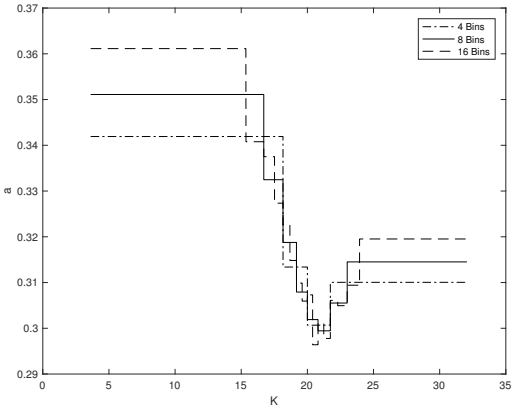
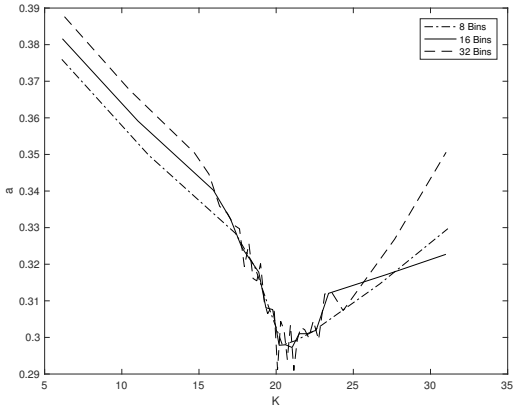


Figure 5.10: Conditional expectation for the Smooth Bin method with different numbers of bins

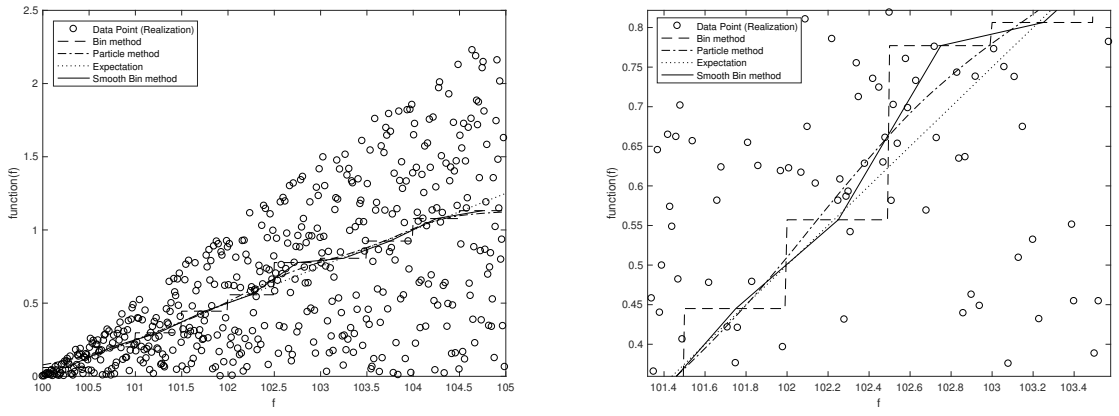


that the choice of the number of bins is especially of importance in this region, and less important in the region at the money, as the differences in the estimator of the conditional expectation of a are larger far in and out of the money.

5.5.3 Results for the Particle Versus Bin Method

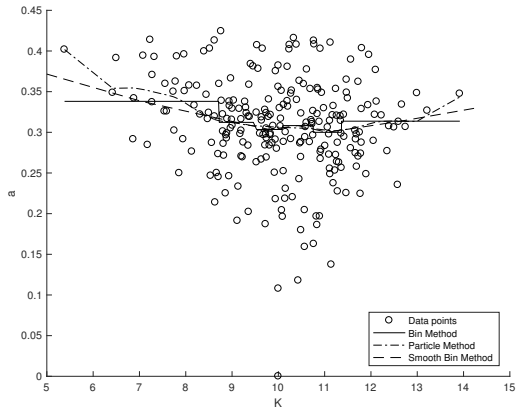
To show the differences between the Particle and the Bin methods, several experiments are conducted. Random values of a for given f are created with distribution $(0.5f - 50)U(0,1)$, where $U(0,1)$ represents the uniform distribution. The results are shown in Figure 5.11. The Particle method, with $Ke(1)$, and the Bin method (equal size of the bins and equal number of pairs per bin) are used to compute an expectation of a conditional on f. It can be seen that the Bin method is not continuous and differs on average more from the expectation than the Particle method.

Figure 5.11: Conditional expectation for the Particle, Standard Bin and Smooth Bin method



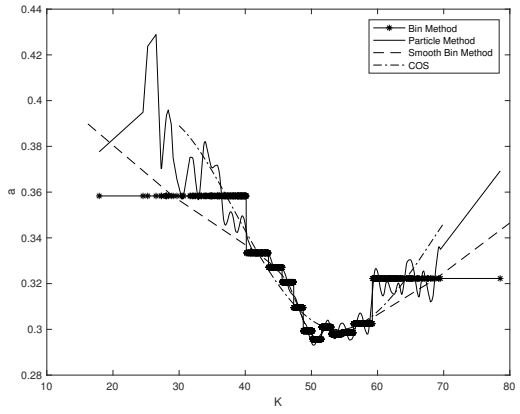
In Figure 5.12, another experiment is shown to determine $\mathbb{E}[a_T^2|S_T = K]$. Here, pairs (S_t, a_t) are computed where the underlying processes are from the Heston model with parameters, $S_0 = 10$, $\kappa = 5$, $\sigma = 0.25$, $\rho = -0.25$, $\theta = 0.1$, $v_0 = 0.1$, $8 \cdot 10^3$ paths, $\Delta = 0.1$, and with $Ke(1)$ (Parameter Set 5). At $t = 1$, the distribution of the pairs is given in this figure, together with the estimation of the conditional expectation of a given S. It can be seen that the Particle and Smooth Bin method both have very similar expectations in roughly $(S_0 - 2, S_0 + 2)$. The Bin method has a lower conditional expectation of a for large values x in $S \pm x$. To compare these two methods with the distribution as according to the COS method,

Figure 5.12: Conditional expectation for the Particle and Bin method with pairs



the same experiment was conducted to determine $\mathbb{E}[a_T^2|S_T = K]$ with Parameter Set 5: $Ke(1)$, $S_0 = 50$, $\kappa = 5$, $\sigma = 0.25$, $\rho = -0.25$, $\theta = 0.1$, $v_0 = 0.1$, $8 \cdot 10^3$ paths, $\Delta = 0.1$, $t = 1$, and 12 bins. The results are shown in Figure 5.13. This result together with those of the previous experiments clearly shows the minimal differences for values of a given S near S_0. The differences of the conditional expectation of a for large values x in $S \pm x$ are larger. Here, the Particle method is closer to the COS estimator. However, note that only a small number of pairs end up in this area. Therefore, when simulating paths, only a small number of paths are affected by the large difference between the Particle and Smooth Bin method.

Figure 5.13: Conditional expectation with the Particle, Standard Bin, Smooth Bin, and COS method

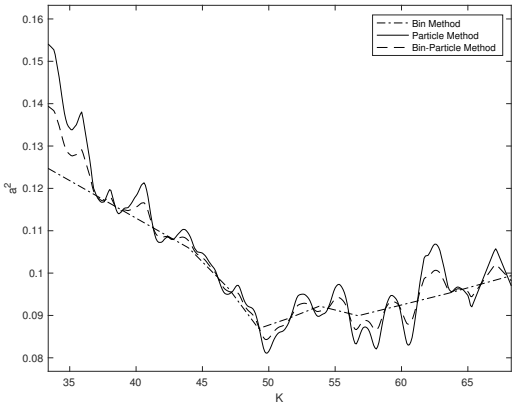


In Figure 5.14, the results are shown for the same experiment to determine $\mathbb{E}[a_T^2|S_T = K]$ as was conducted for the results in Figure 5.13 (with Parameter Set 5). Furthermore, the average of the Particle and Bin method is used for the determination of the conditional expectation, this is the Particle-Bin method. The Particle-Bin method gives a function that is smoother than the function created by the Particle method, but more volatile than the function created by the Smooth Bin method. Furthermore, no large differences are noted.

The Particle method and Bin method both give results which are close to the COS method. However, for a small number of pairs (far away from S_0), the Particle method seems to give results which are closer to the value obtained by the COS method. However, due to the small number of pairs there, the high variance, and the important impact of the number of bins (see Figure 5.9 and Figure 5.10) in the region far away from S_0, these differences are not significant (95%, normal distribution) in all experiments as can be seen in Figures 5.12, 5.13, and 5.14.

If a very small or large number of bins is chosen, or if the kernels are very large or small, a significant

Figure 5.14: The Particle-Bin model conditional expectation



(95 % confidence, normal distribution) difference between the Particle and Bin methods can be found in the experiments. Only under these circumstances, the Particle-Bin method gives a smaller difference with the results of the COS method than both the Particle and Bin method. However, it has to be noted that the differences are small and that an increase in the number of simulations (or bins) decrease the difference with the COS method drastically. This difference is measured by taking values on the conditional expectation function created by the methods, distracting the COS value at these points, and taking the absolute value. The variance of these results is determined by running these simulations and defining this difference 10 times.

Another Particle-Bin experiment with the weighting of elements in the bin is not significantly different from the Bin or Particle method. The same holds for the standard Particle-Bin method, where the average between the bin and Particle method is used for the determination of the conditional expectation function. Therefore, no further results of these experiments are presented.

Chapter 6

Numerical Results of the LSV Model

6.1 Experiment with the Feller Condition Not Satisfied

The computational times for the calibration of the SV Heston model are determined for two different data sets. These data sets consist of European call option prices with strikes $K = 50+10x$, with $x = 0, 1, ...10$, and with time to maturity $T = 1$. These European call option prices are determined with the COS method for Parameter Set 6 (Feller condition not satisfied): $\rho = -0.5, v_0 = 0.05, \theta = 0.05, S_0 = 100, r = 0, \varepsilon = 0.8, \kappa = 1$, and Parameter Set 7 (Feller condition satisfied): $\rho = -0.5, v_0 = 0.1, \theta = 0.1, S_0 = 100, r = 0, \varepsilon = 0.25, \kappa = 5$.

The calibration times for the local calibration are determined for the calibration with no requirements regarding the Feller condition and with the requirement that the Feller condition is satisfied. The calibration time conditional that the Feller condition is satisfied for Parameters set 6 is 347 seconds. The unconditional calibration time is 494 seconds and it gives the exact solution, i.e., the input parameters are the same as the parameters obtained with the calibration. For Parameter Set 7, the calibration time is 188 seconds. This shows that calibration is computationally faster if the Feller condition is satisfied in the original data. Furthermore, it shows that calibrating without any restrictions on the calibrated parameters is computationally more time consuming, than conditionally calibrating that the Feller condition has to be satisfied.

Another experiment is conducted to determine the proportion of variance processes v that attain zero. Simulations conducted with the calibrated parameters conditional on satisfaction of the Feller condition do not reach $v = 0$. With input from Parameter Set 6, with $\Delta = 1/64$, and with $T = 1$, the proportion of v processes that reach zero before or at $T = 1$ are given in Table 6.1.

For the following parameter set, $\rho = -0.5, v_0 = 0.05, \theta = 0.05, S_0 = 100, r = 0, \varepsilon = 0.5, \kappa = 1$ (Parameter Set 8), the same calculations give the results in Table 6.2. The results show that the discretization schemes do not have a large impact on this proportion. However, the Euler scheme gives different results than the other two schemes. This can be explained by a different probability density function of v_t around zero. However, a change of the ε parameter (from 0.8 to 0.5) drops the proportion significantly as seen in Table 6.1 and Table 6.2.

Table 6.1: Proportion of volatility paths that reach zero

Discretization	Proportion
Euler	0.856
Milstein	0.870
QE	0.864

Table 6.2: Proportion of volatility paths that reach zero

Discretization	Proportion
Euler	0.562
Milstein	0.522
QE	0.525

Another experiment is conducted to gain insight into the computational costs for the different discretization schemes. The computational costs are defined as the average computation time of 10^6 paths with $t = 0$ and $T = 1$. The relative computation costs (RCC) in comparison with the calibration conditional on satisfying the Feller condition (Parameter Set 6), are given in Table 6.3. The computational costs per path are therefore higher if one calibrates unconditional on the Feller condition. The results show a smaller impact of this condition for the QE scheme than for the Euler and Milstein scheme.

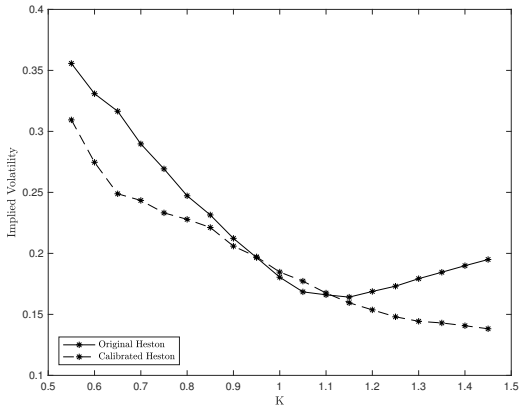
Table 6.3: Computation costs

Discretization	RCC
Euler	1.1091
Milstein	1.0920
QE	1.0041

6.2 Forward Start Options with Synthetic Data

This section describes the results of an experiment performed to determine call option and forward start option prices. First, European call option prices are generated with the COS method where the dynamics of the underlying asset price follows the SDE of the Heston model. Here, the Feller condition is not satisfied in the parameter set, $\theta = 0.05, v_0 = 0.05, \varepsilon = 0.8, \rho = -0.5, \kappa = 1$ (Parameter Set 9). The European call option prices are on a mesh, with strike varying from 10 to 190, with differences of 5, and the time to maturity varying from 0.1 to 2.3, with differences of 0.1. These call option prices are used to calibrate the Heston model. The calibration conducted conditional on the satisfaction of the Feller condition gives $\theta = 0.0392, v_0 = 0.04, \varepsilon = 0.3111, \rho = -0.772, \kappa = 1.2340$ (Parameter Set 10). Furthermore, the local volatility function with Dupire's model is determined after the calibration of the Heston model. Afterward, the local volatility surface for the LSV model is determined as well. Then, the experiment determines forward start call option prices with $T_1 = 1$ and $T_2 = 2$ for different strikes, and converts these values to implied volatilities to compare the results.

Figure 6.1: Implied volatility European call options

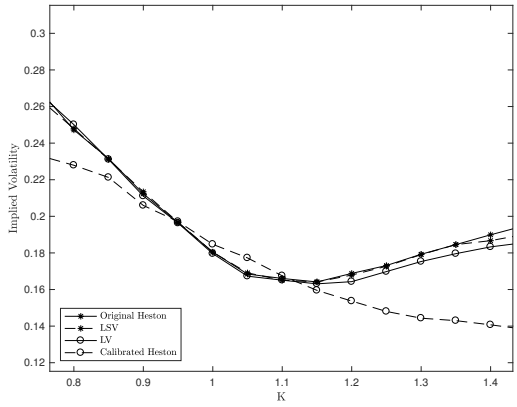


In Figure 6.1, the implied volatility is given for European call options determined by the COS function. The original Heston implied volatility uses Parameter Set 9. The calibrated Heston implied volatility is generated by using Parameter Set 10. This calibration is conditional on the satisfaction of the Feller condition. One can see that the calibrated Heston model gives a flattening of the implied volatility smile.

Figure 6.2 depicts the implied volatility for forward start option prices generated with the COS method and original parameters, as well as generated with Dupire's model (LV), the calibrated Heston model, and the LSV model are given[1]. Both the LV and the LSV model give the same implied volatility as

[1]The implied volatilities generated with this model can be found in Appendix C.1.

Figure 6.2: Implied volatility European call options



observed in the market (prices generated by the COS method) for European call option prices.

Figure 6.3 presents the implied volatility for forward start options with $T_1 = 1, T_2 = 2$, and for Parameter Set 9 which is named the original Heston implied volatility. Here, the Feller condition is not satisfied. The function describing the implied volatility corresponding to the calibrated Heston model is given as well. This calibration is conditional on satisfaction of the Feller condition and gives Parameter Set 10. Furthermore, the implied volatility corresponding to Dupire's model is given. One can see that the calibrated Heston model gives a flattening of the implied volatility smile in comparison to the original Heston model computations. The local volatility model gives the same implied volatility as observed in the market for European call option prices. However, in this figure, one can see that for the forward start options, the smile is flattened by the LV model results.

Figure 6.3: Implied volatility (market, calibrated Heston, LV)

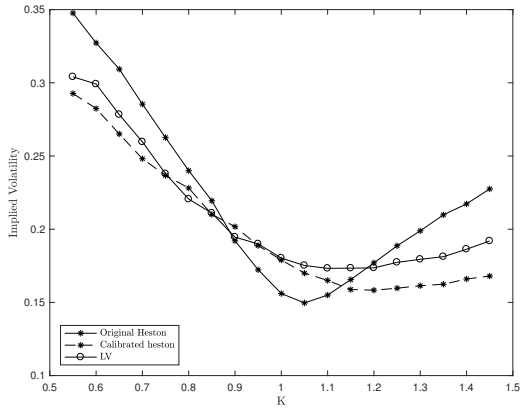


Figure 6.4 presents the LSV model results as well. The implied volatilities generated with this model can also be found in Appendix C. Here, it can be seen that the implied volatilities of the LSV model are very close to the implied volatilities of the original Heston model and improve the results of the calibrated Heston model.

For $S_0 = 100$ the implied volatility surface generated with the LSV model is given in Figure 6.5. The far out-of-the-money or in-the-money values are larger in order to increase the implied volatility there, and the values at-the-money are low to decrease the implied volatility.

Figure 6.4: Implied volatility (market, calibrated Heston, LSV)

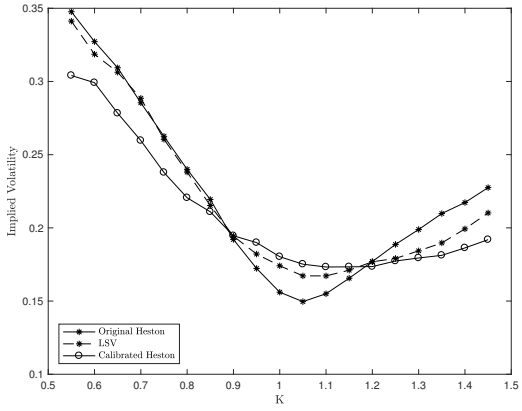
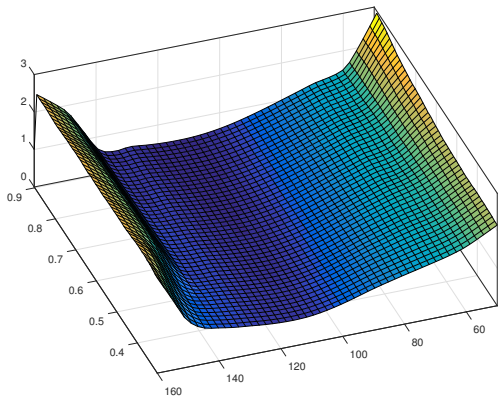


Figure 6.5: Local volatility surface for LSV



6.3 Results with Market Data

Tests with synthetic data are often desired in mathematical research since this provides an opportunity to understand the original process, such as for the Heston model. However, in practice models never describe the reality exactly. Therefore, the LSV model has been tested with real market data. The LSV model has been calibrated by first calibrating the Heston model (conditional that the Feller condition is satisfied). Then the local volatility function has been determined. The latter has been done by first numerically calculating Dupire's volatility surface on a mesh and then interpolating and applying the Particle algorithm.

The call options market data for Eurostoxx is used for the calibration of the model. The option prices, the arithmetic average between the ask and bid price, for given strike and time to maturity can be found in Appendix B.1, including their source. Appendix B.1 plots the call option prices. This surface is used for the calibration of the Heston model and for the calibration of Dupire's local volatility surface.

The calibration of Heston's model is implemented in MATLAB R2017A by minimizing the error with the *lsqnonlin* function. This is the local optimization (the global optimization gives the same result here). The parameters after calibration of the Heston model for the market data are given in Table 6.4.

For the calibration of the local volatility following the calibration of the stochastic volatility part of the LSV model, Dupire's volatility surface is first determined. The surface is given in Figure 6.6. This surface is created on a grid of 12 $(4 \cdot 3)$ pairs (K, T), which is then interpolated using cubic splines to

Table 6.4: Parameters

	Value
v_0	0.0640
κ	6.2198
θ	0.0135
σ	0.3032
ρ	0.5285

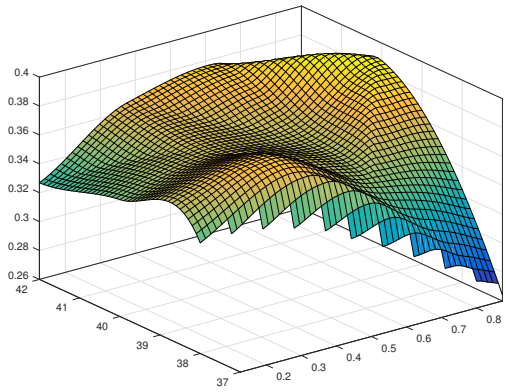


Figure 6.6: Dupire's local volatility surface

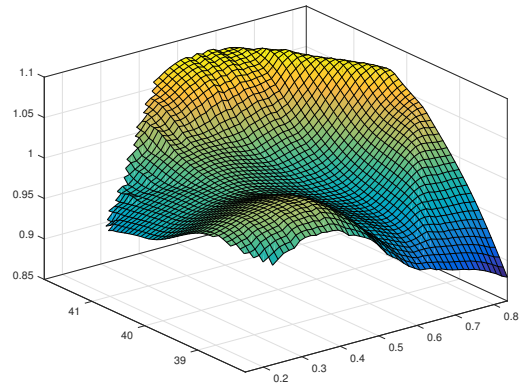


Figure 6.7: LSV local volatility surface

create a smooth surface. Given the calibration of the stochastic volatility part and Dupire's surface, the local volatility surface for the LSV model can be determined. Note that the domain of this surface cannot be larger than that of Dupire's surface. It is even smaller (in the direction of K) given a small time to maturity due to the absence of sufficient particles that end up far away from S_0 in the K direction. This surface is given in Figure 6.7. It can be seen that this surface is relatively close to 1, which is expected from the theory.

6.3.1 Extrapolation of the Local Volatility Surface with Market Data

The option prices available in the market do not cover all strikes and time to maturities. The local volatility surface, created in the experiment in Section 6.3 is only given for a limited domain. To be able to simulate paths out of this domain with the MC method, the domain needs to be expanded.

Several solutions exist, as described in Section 5.2, such as taking a constant local volatility out of the calibrated domain. Another solution is to linearly extrapolate the values of the surface at the boundaries.

Pricing call options with these different extrapolations leads to different prices. Figure 6.8 presents the percentage of MC simulated pairs (S_t, t) that are in the initial calibrated domain of the local volatility function. The x-axis represents the time t. It can be seen that the local volatility surface is not available for approximately $t < 0.1$. This is because insufficient European call option prices were available in the market to calibrate Dupire's surface with a short enough time to maturity. Furthermore, it can be seen that the function increases after about $t = 0.1$. This can be explained by the increase in values of S_t which are much smaller or larger in comparison with S_0.

In Figure 6.9, a different experiment is conducted. Here the local volatility surface is linearly extrapolated in the direction of S for approximately $0.1 < t < 0.8$. The proportion of S_t values in this region that are not in this newly expanded domain decreases to zero. This shows that this approach can be applied. However, the different pricing results for the two approaches cannot be compared with real prices, as no forward start option prices can be observed in the market.

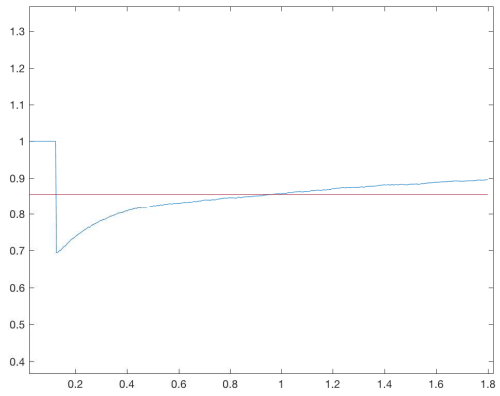


Figure 6.8: Percentage out of domain LV in Dupire's model with extrapolation in T direction

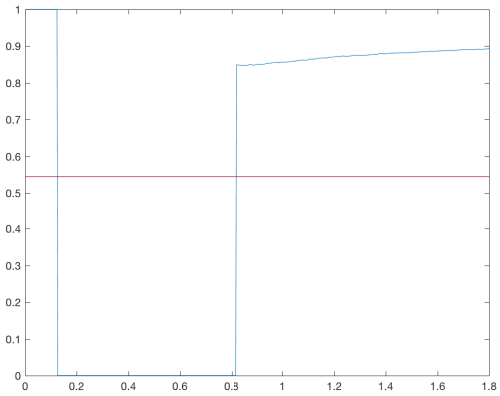


Figure 6.9: Percentage out of domain LV in Dupire's model

The generated option prices of this experiment can be found in Appendix C. Furthermore, the implied volatilities of the Heston model, the LV model, and the LSV model can be found in appendices D.1, D.2, D.3, and D.4.

Chapter 7

Conclusion and Future Research

7.1 Conclusion

Several stochastic processes have been developed to describe paths of exchange-listed asset prices to price options. The Black-Scholes model [4] does not coincide with the implied volatility of the market. The local volatility model as developed by Dupire [12] uses the same dynamics as the Black-Scholes model, but it makes it possible to precisely mimic implied volatilities for European options. In contrary to the Black-Scholes and the local volatility model, the Heston stochastic volatility model [22] allows non-deterministic volatilities. The dynamics of this model give European option prices close to the prices in the market. However, for forward start options, and thus cliquet options, the results are less precise.

The calibration of the Heston model to market prices often leads to parameters that do not satisfy the Feller condition. If this condition is not satisfied, the variance process can become zero. The calibration of these parameters is computationally less efficient and the accuracy in pricing options is lower than if the Feller condition is satisfied.

The Heston local stochastic volatility model combines the Heston model with a local volatility model. This creates a model with the dynamics of the stochastic volatility model which offers a precise match for the implied volatility. Pricing of forward start options with this model leads to prices which are close to market prices. In this model, it is possible to calibrate the Heston parameters conditionally on the satisfaction of the Feller condition, and then later to compensate the different implied volatilities with a local volatility function. The results show that this approach gives accurate prices for forward start options and is computationally efficient.

The impact of calibration on the assumption that the Feller condition is satisfied has been investigated. The results show that if the Feller condition is not satisfied, the calibration is computationally more expensive. If one simulates with the Feller condition not satisfied, this takes more time in comparison with a parameter set that satisfies the Feller condition. The Euler, Milstein, and QE discretization schemes both give a similar proportion of variance processes that reach zero given a parameter set that does not satisfy the Feller condition.

The calibration of the local volatility function in the LSV model is not straightforward [21], one ends up in the framework of Malliavin. The conditional expectation that arises for the determination of the local volatility function can be determined using the Particle method of Henry-Labordère and the bin approach described by Van der Stoep [17].

The Particle and Smooth Bin method both converge and perform well if the right kernel bandwidth (Particle method) or the right size of the bin (Bin method) is chosen and if sufficient paths are simulated. The Bin method with an equal number of pairs per bin performs better than the Bin method with equal

bin size. The larger the bin size, the more volatile the conditional expectation becomes. On the other hand, a larger bin size improves the estimation. The Particle method works well if the right bandwidth is chosen which can be done by Silverman's rule of thumb [35]. The alternative Particle-Bin model gives results similar to those of the Bin and Particle method; however, it requires more calculation time. This alternative method performs better than the Particle and Bin method if the number of bins or kernel bandwidth is chosen inaccurately. Therefore, if the right number of bins or bandwidth cannot be determined beforehand, it is better to use the Particle-Bin method for efficient option pricing. In all other cases, it is less efficient.

The multilevel Monte Carlo method can reduce the computational costs while maintaining the accuracy level. This method has been successfully implemented for the CIR process, the Milstein, and the Euler discretization for the Heston model. The large number of simulations for a coarse mesh and the small number of simulations for a fine mesh makes this method reliable and appropriate in the context of the Heston LSV model and for the pricing of (forward start) options. The MLMC Milstein discretization is the computational most efficient way to simulate paths with both the Heston model and the Heston LSV model.

Altogether the chapters in this thesis provide a detailed framework with several (alternative) techniques for both the computationally efficient pricing of forward start options and the reconstruction of the implied volatility.

7.2 Future Research

The following subjects are worth mentioning and can serve as possible topics for future research:

- The multilevel Monte Carlo method performs well for the Milstein and Euler discretization of the Heston model. The MLMC results for the quadratic exponential discretization scheme in this research are not satisfying, as they do not converge. The results show the computationally superior results of QE over the Milstein and the Euler scheme in the standard MC framework; thus this is an interesting topic. The decision rule if the variance process needs to be computed with the algorithm for small or large values is usually not the same for the M fine processes that, in combination, form the coarse mesh. Therefore, it is not directly clear which algorithm the coarse mesh should work with. This likely causes the non-convergence of the algorithm. Other approaches can therefore be studied to develop a computationally superior method using the QE scheme and the MLMC method.
- The determination of the Local Volatility surface as in Dupire's model is computationally time consuming. Computational improvements are therefore important and can be investigated.
- The global and local optimizations for the calibration of the Heston model have been tested in this thesis for only two MATLAB functions. Results show that the calibration for these two functions is time consuming. Therefore, more research on computational efficient optimization is suggested.
- The limited number of option prices available in the market makes the development of the local volatility function possible on a grid of prices. Interpolation gives a surface, but the values out of te domain of this surface are harder to estimate. Linear extrapolation has a negative influence on the pricing of (forward start) options due to a mismatch of the implied volatility. Research into the computationally most efficient and most accurate way to extrapolate (and interpolate) is therefore suggested.
- Most of the methods for the efficient calibration as well as for the discretization and pricing with the LSV model have not been tested for noncontinuous processes. The Heston model and the LSV model can be expanded by adding noncontinuous dividends.
- A framework to determine the most efficient bin size or the kernel definition based on the parameters of the models can be investigated, such that a function can be defined to chose these parameters in the most efficient way.

Bibliography

[1] Andersen, L. (2006) *Efficient Simulation of the Heston Stochastic Volatility Model*, Bank of America securities. Retrieved from `http://www.javaquant.net/papers/LeifAndersenHeston.pdf`

[2] Atkinson, K. (1989) *An Introduction to Numerical Analysis*, John Wiley & Sons Inc., ISBN 0-471-62489-6.

[3] Behvand, F. (2010) *Notes: Forward Implied Volatility*, Christ Church Oxford University. Retrieved from `http://www1.maths.ox.ac.uk/system/files/legacy/12804/StochasticImpliedVolatility.pdf`

[4] Black, F., Scholes, M. (1973) *The Pricing of Options and Corporate Liabilities*, Journal of Political Economy, Vol. 81, No. 3.

[5] Carr, P. & Madan, D.B. (1999) *Option valuation using the fast Fourier transform*, Journal of Computational Finance Vol. 3, 463-520.

[6] Chen, B. (2007) *Thesis: Calibration of the Heston Model with Application in Derivative Pricing and Hedging*. Retrieved from `http://ta.twi.tudelft.nl/mf/users/oosterle/oosterlee/chen.pdf`

[7] Crisostomo, R. (2014) *An Analysis of the Heston Stochastic Volatility Model: Implementation and Calibration Using Matlab*, CNMV Working Paper. Retrieved from `https://ssrn.com/abstract=2527818`

[8] Dawkins, P. *Notes*. Retrieved from `http://tutorial.math.lamar.edu/Classes/DE/FourierCosineSeries.aspx`

[9] Deelstra, G. & Rayée, G. (2013) *Local volatility pricing models for long-dated FX derivatives*, Applied Mathematical Finance Vol. 20 No. 4, 380–402.

[10] Del Moral, P. (2004) *Feynman-Kac Formulae, Genealogical and Interacting Particle Systems with Applications*, ISBN: 978-1-4419-1902-1.

[11] Derman, E. & Kani, I. (1998) *Stochastic implied trees: Arbitrage pricing with stochastic*, term and strike structure of volatility, International Journal of Theoretical and Applied Finance 1 (1), 61–110.

[12] Dupire, B. (1994) Pricing with a smile, Risk, Vol. 7, 18-20.

[13] Engelmann, B., Koster, F. & Oeltz, D. (2011) *Calibration of the Heston Stochastic Local Volatility Model: A Finite Volume Scheme*. Retrieved from `http://rivacon.com/sites/default/files/stoch_local_vol.pdf`

[14] Fang, F., Oosterlee, C.W. (2009) *Pricing early-exercise and discrete barrier options by Fourier-cosine series expansions*, Numerische Mathematik, Vol. 114, 27-62.

[15] Giles, M.B. (2008) Multilevel Monte Carlo path simulation, Operations Research, 607-617.

[16] Gottlieb, O. (2007) *Deriving Put-Call Parity*. Retrieved from `http://www.soarcorp.com/research/put_call_parity.pdf`

[17] Grzelak, L., Oosterlee, C.W., Van der Stoep, A. (2014) *The Heston Stochastic-Local Volatility Model: Efficient Monte Carlo Simulation*, International Journal of Theoretical and Applied Finance, Vol. 17, No. 7.

[18] Guyon, J. and Henry-Labordère, P. (2011) *The Smile Calibration Problem Solved*, Risk Magazine.

[19] Gyöngy, I. (1986) *Mimicking the one-dimensional distributions of processes having an Ito differential*, Prob. Th. Rel. Fields, Vol. 71, No. 4, 501-516.

[20] Henry-Labordère, P. (2008) *Analysis, Geometry and Modeling in Finance: Advanced Methods in Option Pricing*, Chapman & Hall/CRC, Financial Mathematics Series.

[21] Henry-Labordère, P. (2009) *Calibration of Local Stochastic Volatility Models to Market Smiles*, Risk Magazine.

[22] Heston, S.L. (1993) *A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options* The Review of Financial Studies 1993 Vol. 6, No. 2, 327-343.

[23] Kahl, C. & P. Jäckel (2005) *Fast strong approximation Monte-Carlo schemes for stochastic volatility model*, Quantitative Finance, Vol. 6, 513–536.

[24] Lipton, A. (2002) *The vol smile problem*, Risk Magazine, Vol. 15, 61-65.

[25] Madan, D.B., Carr, P. & Chang, E. (1998) *The variance gamma process and option pricing*, European Finance Review 2, 79-105.

[26] Madan, D.B., Qian, M. & Ren, Y. (2007) *Calibrating and pricing with embedded local volatility models*, Risk Magazine Vol. 20 No. 9, 138–143.

[27] Oosterlee, C.W. (2017) *Notes: Financial Mathematics*, Chapter 3, Local Volatility Models 68-69.

[28] Oosterlee, C.W. & Fang, F. (2011) *A Fourier-based Valuation Method for Bermudan and Barrier Options under Heston's Model*, SIAM Journal on Financial Mathematics, Vol. 2.

[29] Piterbarg, V. (2006) *Markovian Projection for Volatility Calibration*, Risk Magazine.

[30] Raab, M. (2016) *Inside Structured Products*, EUSIPA - European Structured Products Association, ISBN: 978-3-033-05912-2.

[31] Rebonato, R. (1999) *Volatility and Correlation in the Pricing of Equity*, FX, and Interest-rate Options. New York: John Wiley & Sons, Ltd.

[32] Rouah, F.D. (2013) *Notes: Euler and Milstein Discretization*. Retrieved from `http://www.frouah.com/finance%20notes/Euler%20and%20Milstein%20Discretization.pdf`

[33] Rubinstein, M. and Reiner, E. (1991) *Breaking Down the Barriers*, Risk Magazine, Vol. 4, No. 8, 28-35.

[34] Shreve, S.E. (2004) *Stochastic Calculus for Finance II: Continuous-Time Models*, Springer Finance, Springer.

[35] Silverman, B. (1986) *Density Estimation for Statistics and Data Analysis*, Chapman & Hall.

[36] Sznitman, A. (1989) *Topics in propagation of chaos*, Ecole d'été de probabilités de Saint-Flour, Volume 1464 of Lecture Notes in Mathematics, Springler-Verlag.

[37] Tataru, G., Fisher, T. & Yu, J. (2012) *The Bloomberg stochastic local volatility model for FX exotics*, Bloomberg second edition.

[38] Williams, D. (1991) *Probability with Martingales*, Cambridge University Press.

[39] Wilmott, P. (1998) *The Theory and Practice of Financial Engineering*, John Wiley and Sons.

[40] Yiran Cui, Sebastian del Baño Rollin, Guido Germano (2015) *Full and fast calibration of the Heston stochastic volatility model*. Retrieved from `https://arxiv.org/abs/1511.08718`

[41] Zhu, J. (2000) *Modular Pricing of Options*. An Application of Fourier Analysis, Springer.

Appendix A

Proofs

A.1 Proof of COS Conditional Probability Density Recovery Theorem

Proof. Suppose $[a, b] \in \mathbb{R}$ chosen, such that the trunctated integral approximates the infinite counterpart. Then the characteristic function corresponding to this truncated integral is,

$$\hat{\phi}(u|x) = \int_a^b e^{izu} f(z|x)dz \approx \int_{\mathbb{R}} e^{izu} f(z|x)dz = \phi(u|x). \tag{A.1}$$

Substituting $u = \frac{k\pi}{b-a}$ and multiplication by $exp\left(-ik\pi\frac{a}{b-a}\right)$ gives

$$\hat{\phi}\left(\frac{k\pi}{b-a}|x\right) exp\left(-ik\pi\frac{a}{b-a}\right) = \int_a^b exp\left(ik\pi\frac{z}{b-a} - ik\pi\frac{a}{b-a}\right) f(z|x)dz. \tag{A.2}$$

This can be rewritten given $\Re(e^{iu}) = \cos u$,

$$\Re\left(\hat{\phi}\left(\frac{k\pi}{b-a}|x\right) exp\left(-ik\pi\frac{a}{b-a}\right)\right) = \int_a^b \cos\left(k\pi\frac{z-a}{b-a}\right) f(z|x)dz. \tag{A.3}$$

From the definition of A_k 4.4 it follows that,

$$\bar{A}_k = \frac{2}{b-a}\Re\left(\phi\left(\frac{k\pi}{b-a}|x\right) exp\left(-ik\pi\frac{a}{b-a}\right)\right) \tag{A.4}$$

Replacing A_k by $\bar{A}_k$ and the truncation of the series summation then gives,

$$f(z|x) \approx \sum_{k=0}^{N-1} w_k \bar{A}_k \cos\left(k\pi\frac{y-a}{b-a}\right) \tag{A.5}$$

□

A.2 Proof of the COS Pricing Formula

Proof.

$$\hat{V}(x,t) = D(t,T)\int_a^b g(z) \sum_{k=0}^{N-1} w_k \bar{A}_k \cos\left(k\pi\frac{z-a}{b-a}\right) dz. \tag{A.6}$$

Furthermore,

$$\bar{A}_k \cos\left(k\pi\frac{z-a}{b-a}\right) = \frac{2}{b-a}\Re\left(\phi\left(\frac{k\pi}{b-a}|x\right) exp\left(-ik\pi\frac{a}{b-a}\right)\right)\cos\left(k\pi\frac{z-a}{b-a}\right)dz. \tag{A.7}$$

Applying Fubini's Theorem gives

$$\hat{V}(x,t) = D(t,T)\sum_{k=0}^{N-1} w_k \Re\left(\phi\left(\frac{k\pi}{b-a}|x\right) exp\left(-ik\pi\frac{a}{b-a}\right)\right)\frac{2}{b-a}\int_a^b g(z)\cos\left(k\pi\frac{z-a}{a-b}\right)dz. \tag{A.8}$$

Now define $\mathcal{V}_k$ by

$$\mathcal{V}_k = \frac{2}{b-a}\int_a^b g(z)\cos\left(k\pi\frac{z-a}{a-b}\right)dz. \tag{A.9}$$

□

A.3 Proof of the COS Pricing formula for a Lévy process

Proof. The conditional characteristic function of a Lévy process is given by:

$$\phi(u|x) = \phi(u|x=0)e^{iux}, \tag{A.10}$$

Therefore,

$$\begin{aligned}\hat{V}(x,t) &= D(t,T)\sum_{k=0}^{N-1} w_k \Re\left(\phi\left(k\pi\frac{1}{b-a}\right) exp\left(-ik\pi\frac{a}{b-a}\right)\right)\mathcal{V}_k(T)\\ &= D(t,T)\sum_{k=0}^{N-1} w_k \Re\left(\phi\left(k\pi\frac{1}{b-a}\right) exp\left(ik\pi\frac{x-a}{b-a}\right)\right)\mathcal{V}_k(T),\end{aligned} \tag{A.11}$$

which gives the desired result □

A.4 Proof of the moment-generation function for CIR process theorem

Proof. From [27] it follows that the density for the r.v. v_t is defined by:

$$f_{v_t}(x) = \frac{1}{\bar{c}_t} f_{\chi^2}(\bar{d},\bar{\gamma}_t)\left(\frac{x}{\bar{c}_t}\right), \tag{A.12}$$

and the noncentral chi-squared distribution, $\chi^2(\bar{d},\bar{\gamma}_t)$ is defined by:

$$f_{\chi^2(\bar{d},\bar{\gamma}_t)}(x) = \sum_{n=0}^{\infty}\frac{1}{n!}e^{-\bar{\lambda}_t/2}\left(\frac{\bar{\lambda}_t}{2}\right)^n f_{\chi^2(\bar{d}+2n)}(x). \tag{A.13}$$

Here $\chi^2(a)$ is distributed as a chi-squared distribution with a degrees of freedom. Using

$$M_{\chi^2(\bar{d}+2n)}(u\bar{c}_t) = \int_0^{\infty} e^{u\bar{c}_t x} f_{\chi^2(\bar{d}+2n)}(x)dx = \left(\frac{1}{1-2u\bar{c}_t}\right)^{\bar{d}/2+n}, \tag{A.14}$$

and then defining the moment-generating function and changing variables gives

$$\begin{aligned}
M_{v_t}(u) &= \mathbb{E}^{\mathbb{Q}}\left[e^{uv_t} \mid \mathcal{F}(t_0)\right] \\
&= \frac{1}{\bar{c}_t}\int_0^\infty e^{ux}\sum_{n=0}^{\infty}\frac{1}{n!}e^{-\bar{\lambda}_t/2}\left(\frac{\bar{\lambda}_t}{2}\right)^2 f_{\chi^2(\bar{d}+2n)}\left(\frac{x}{\bar{c}_t}\right)dx \\
&= \sum_{n=0}^{\infty}\frac{1}{n!}e^{-\bar{\lambda}_t/2}\left(\frac{\bar{\lambda}_t}{2}\right)^n\int_0^\infty e^{u\bar{c}_t x} f_{\chi^2(\bar{d}+2n)}(x)dx \\
&= \sum_{n=0}^{\infty}\frac{1}{n!}e^{-\bar{\lambda}_t/2}\left(\frac{\bar{\lambda}_t}{2}\right)^n\left(\frac{1}{1-2u\bar{c}_t}\right)^{\bar{d}/2+n} \\
&= \left(\frac{1}{1-2u\bar{c}_t}\right)^{\bar{d}/2} exp\left(\frac{\bar{\lambda}_t}{2(1-2u\bar{c}_t)}-\frac{\bar{\lambda}_t}{2}\right)\sum_{n=0}^{\infty}\frac{1}{n!}e^{-\bar{\lambda}_t/(2(1-2u\bar{c}_t))}\left(\frac{\bar{\lambda}_t}{2(1-2u\bar{c}_t)}\right)^n \\
&= \left(\frac{1}{1-2u\bar{c}_t}\right)^{\bar{d}/2} exp\left(\frac{\bar{\lambda}_t}{2(1-2u\bar{c}_t)}-\frac{\bar{\lambda}_t}{2}\right)\sum_{n=0}^{\infty}\mathbb{P}[Z=n],
\end{aligned} \tag{A.15}$$

where the last P stand for the probability under a Poisson process, due to its definition,

$$\mathbb{P}[Z=n] = \frac{1}{n!}e^{-\alpha}\alpha^n = \frac{1}{n!}e^{-\bar{\lambda}_t/(2(1-2u\bar{c}_t))}\alpha^{\bar{\lambda}_t/(2(1-2u\bar{c}_t))}. \tag{A.16}$$

Therefore,

$$\begin{aligned}
M_{v_t}(u) &= \mathbb{E}^{\mathbb{Q}}\left[e^{uv_t} \mid \mathcal{F}(t_0)\right] \\
&= \left(\frac{1}{1-2u\bar{c}_t}\right)^{\bar{d}/2} exp\left(\frac{u\bar{c}_t\bar{\lambda}_t}{1-2u\bar{c}_t}\right).
\end{aligned} \tag{A.17}$$

□

Appendix B

Data

B.1 Market Data for Eurostoxx

Data from Bloomberg 26-11-2016, 14.30.

Strike ($\cdot 10^2$ EUR)	TTM	Market Price ($\cdot 10^2$ EUR)
36	0.0493150	3.750
38	0.0493150	2.145
40	0.0493150	1.035
42	0.0493150	0.435
44	0.0493150	0.170
36	0.1260273	4.300
38	0.1260273	2.910
40	0.1260273	1.850
42	0.1260273	1.095
44	0.1260273	0.615
36	0.3753424	5.550
38	0.3753424	4.350
40	0.3753424	3.350
42	0.3753424	2.550
44	0.3753424	1.920
36	0.6246575	6.475
38	0.6246575	5.350
40	0.6246575	4.400
42	0.6246575	3.600
44	0.6246575	2.920
35	0.8739726	7.775
37	0.8739726	6.675
40	0.8739726	5.250
42	0.8739726	4.425
45	0.8739726	3.425
35	1.8684931	10.125
37	1.8684931	9.200
40	1.8684931	7.850
42	1.8684931	7.100
45	1.8684931	6.100

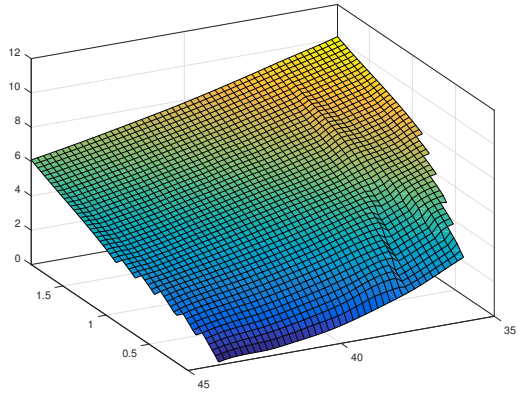


Figure B.1: Call option prices Yahoo! given time to maturity TTM (z-axis) and Strike (x-axis)

Greek	
Delta (δ)	$\frac{\partial C}{\partial S}$
Gamma (γ)	$\frac{\partial^2 C}{\partial S^2}$
Vega (v)	$\frac{\partial C}{\partial \sigma}$
Theta (θ)	$\frac{\partial C}{\partial t}$
Rho (ρ)	$\frac{\partial C}{\partial r}$

Appendix C

Numerical Results

For the described Eurostoxx options, prices observed in the market have been compared with Monte Carlo simulations. The market prices are the arithmetic average between the bid and ask, from Bloomberg

$P_{MCHeston}$	P_{MCLV}	P_{MCLVS}	P_{Mkt}
3.8046	3.8072	3.8046	3.750
2.2085	2.2120	2.2152	2.145
1.0863	1.0921	1.0832	1.035
0.4326	0.4432	0.4300	0.435
0.1390	0.1460	0.1401	0.170
4.2727	4.2666	4.2695	4.300
2.8870	2.8923	2.8824	2.910
1.8327	1.8370	1.8297	1.850
1.0771	1.0829	1.0792	1.095
0.5908	0.6169	0.5888	0.615
5.4665	5.4550	5.4950	5.550
4.2775	4.2753	4.2789	4.350
3.3011	3.2715	3.2850	3.350
2.4842	2.4831	2.4570	2.550
1.8404	1.8332	1.8102	1.920
6.3482	6.2952	6.2879	6.475
5.2389	5.2736	5.2477	5.350
4.2642	4.2678	4.2809	4.400
3.4490	3.5282	3.5200	3.600
2.7681	2.8353	2.8175	2.920
7.6635	7.6821	7.6543	7.775
6.5104	6.5205	6.5163	6.675
5.0832	5.1256	5.0468	5.250
4.2703	4.3565	4.3218	4.425
3.2645	3.4105	3.3520	3.425
9.7184	9.9142	9.8324	10.125
8.8858	8.9214	8.8255	9.200
7.5631	7.7098	7.4569	7.850
6.6933	7.0415	6.6961	7.100
5.6857	5.8947	5.7177	6.100

Option prices P, as determined by MC simulations.

Heston	LV	LSV
0.0546	0.0572	0.0546
0.0635	0.0670	0.0702
0.0513	0.0571	0.0482
-0.0024	0.0082	-0.0050
-0.0310	-0.0240	-0.0299
-0.0273	-0.0334	-0.0305
-0.0230	-0.0177	-0.0276
-0.0173	-0.0130	-0.0203
-0.0179	-0.0121	-0.0158
-0.0242	0.0019	-0.0262
-0.0835	-0.0950	-0.0550
-0.0725	-0.0747	-0.0711
-0.0489	-0.0785	-0.0650
-0.0658	-0.0669	-0.0930
-0.0796	-0.0868	-0.1098
-0.1268	-0.1798	-0.1871
-0.1111	-0.0764	-0.1023
-0.1358	-0.1322	-0.1191
-0.1510	-0.0718	-0.0800
-0.1519	-0.0847	-0.1025
-0.1115	-0.0929	-0.1207
-0.1646	-0.1545	-0.1587
-0.1668	-0.1244	-0.2032
-0.1547	-0.0685	-0.1032
-0.1605	-0.0145	-0.0730
-0.4066	-0.2108	-0.2926
-0.3142	-0.2786	-0.3745
-0.2869	-0.1402	-0.3931
-0.4067	-0.0585	-0.4039
-0.4143	-0.2053	-0.3823

Differences between the market prices and MC simulated prices.

Table C.1: Implied volatilities European call option

Strike	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
Original Heston	0.3557	0.3309	0.3165	0.2897	0.2693	0.2472	0.2315	0.2123	0.1966	0.1805
Calibrated Heston	0.3094	0.2747	0.2489	0.2434	0.2332	0.2279	0.2212	0.2059	0.1971	0.1847
LV	0.3502	0.3312	0.3183	0.2863	0.2674	0.2501	0.2311	0.2111	0.1963	0.1796
LSV	0.3552	0.3301	0.3187	0.2867	0.2643	0.2479	0.2311	0.2134	0.1968	0.1804
Strike	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45	
Original Heston	0.1685	0.1660	0.1641	0.1688	0.1731	0.1793	0.1845	0.1899	0.1950	
Calibrated Heston	0.1773	0.1675	0.1595	0.1537	0.1480	0.1443	0.1430	0.1407	0.1382	
LV	0.1673	0.1651	0.1631	0.1643	0.1698	0.1753	0.1796	0.1833	0.1857	
LSV	0.1691	0.1651	0.1641	0.1679	0.1727	0.1791	0.1845	0.1867	0.1902	

Table C.2: Implied volatilities forward start option

Strike	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95	1.00
Original Heston	0.3093	0.2854	0.2625	0.2400	0.2194	0.1921	0.1723	0.1561	0.1496	0.1550
Calibrated Heston	0.2651	0.2481	0.2368	0.2281	0.2101	0.2017	0.1888	0.1789	0.1700	0.1650
LV	0.2783	0.2597	0.2378	0.2206	0.2110	0.1946	0.1899	0.1803	0.1752	0.1733
LSV	0.3063	0.2885	0.2605	0.2381	0.2153	0.1943	0.1821	0.1741	0.1672	0.1672
Strike	1.05	1.10	1.15	1.20	1.25	1.30	1.35	1.40	1.45	
Original Heston	0.1496	0.1550	0.1656	0.1770	0.1887	0.1989	0.2098	0.2173	0.2275	
Calibrated Heston	0.1700	0.1650	0.1588	0.1583	0.1597	0.1614	0.1624	0.1660	0.1680	
LV	0.1752	0.1733	0.1734	0.1736	0.1774	0.1794	0.1812	0.1863	0.1919	
LSV	0.1672	0.1672	0.1711	0.1768	0.1791	0.1843	0.1897	0.1993	0.2102	

Appendix D

Figures

For the first 4 figures, the x-axis and y-axis represent the time to maturity (0-1) and Strike $\cdot 10^2$ EUR (35-45).

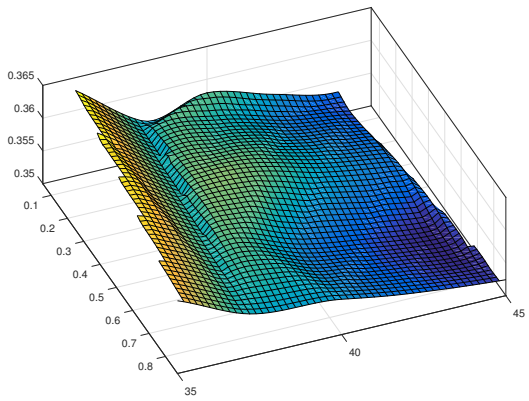


Figure D.1: Implied volatility Heston

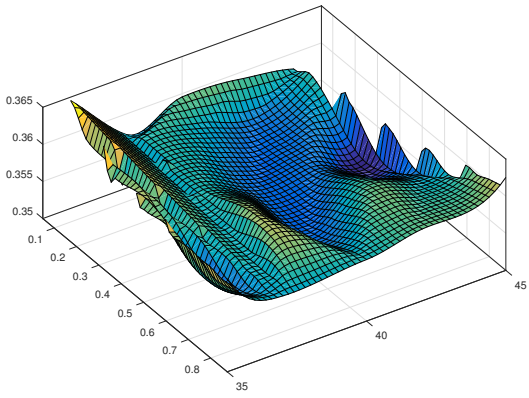


Figure D.2: Implied volatility Dupire

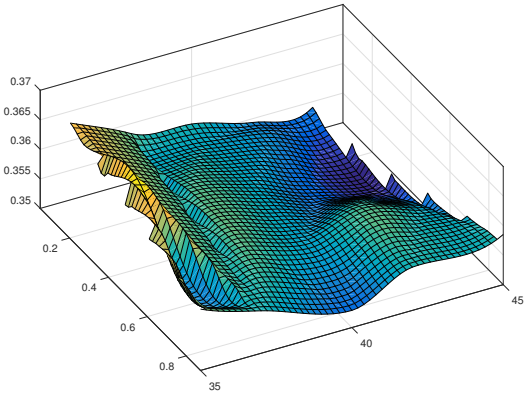


Figure D.3: Implied volatility LSV

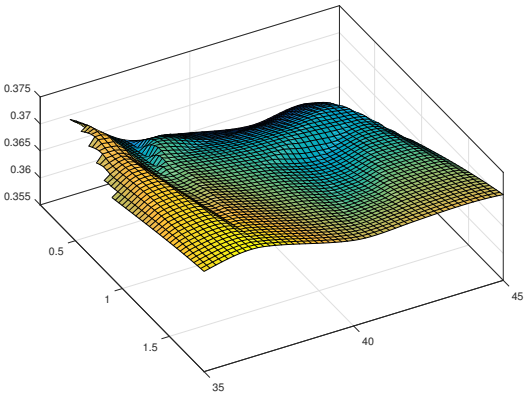


Figure D.4: Implied volatility market price

Figure D.5: Heston MLMC with the Euler scheme

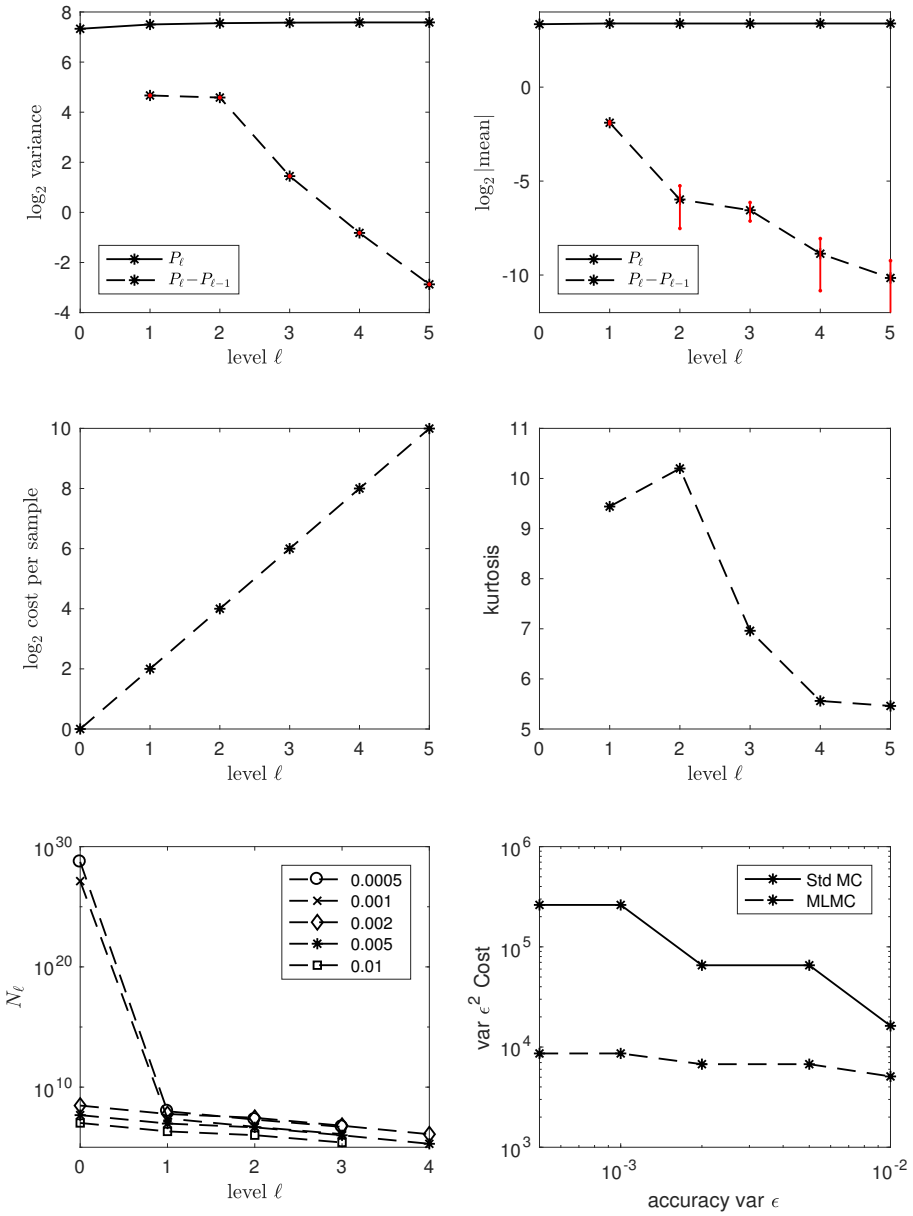


Figure D.6: CIR MLMC with the Milstein scheme

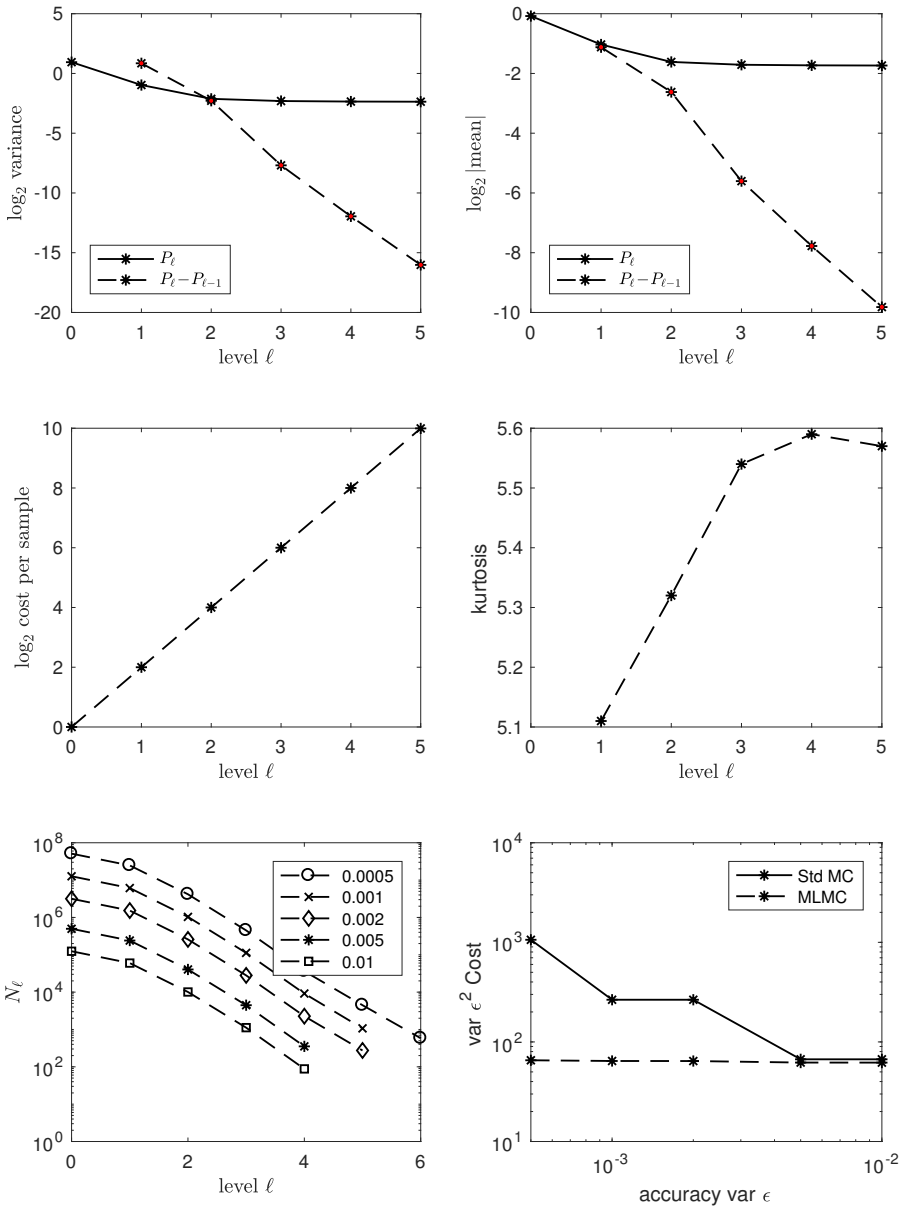


Figure D.7: Heston MLMC with the QE scheme