A study of some efficient numerical techniques used in pricing options under stochastic volatility models

Xiangchen Zeng

University of Wollongong

Follow this and additional works at: https://ro.uow.edu.au/theses1

Recommended Citation

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library: research-pubs@uow.edu.au
A study of some efficient numerical techniques used in pricing options under stochastic volatility models

A thesis submitted in fulfilment of the requirements for the award of the degree

Doctor of Philosophy

from

University of Wollongong

by

Xiangchen ZENG

B. Sc. Jilin University

School of Mathematics and Applied Statistics
Certification

I, Xiangchen Zeng, declare that this thesis, submitted in partial fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Mathematics and Applied Statistics, University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has not been submitted for qualifications at any other academic institution.

________________________
Xiangchen Zeng
April 4, 2018
Abstract

Due to the development of the pricing theory, options, as one of the most important financial derivatives, have become increasingly important in financial markets. The breakthrough of the theory was made by Fischer Black and Myron Scholes in 1973, who proposed a diffusion model with a mathematical formula for pricing European call options. The formula, however, was proven to be not perfect because of the constant volatility assumption made by the authors. Hence the “stochastic volatility models” were established to provide a better fit to the random feature of volatilities. The stochastic volatility models, though, are much more complex and analytical solutions are either unavailable or very difficult to obtain. Besides, since the calculation of a large number of prices is usually required in short time, fast and accurate numerical approaches are in great need.

Among all the numerical techniques, Monte Carlo simulations, tree approaches (binomial and trinomial trees) and finite-difference methods are the three key methods that can be applied to almost all option contracts under all existing models. However, each of the three methods has its own advantage and shortcomings. The motivation of this thesis is to explore the three key methods under the more adaptive stochastic volatility model and propose faster algorithms to fit the need from both academia and industry.

The thesis is divided into two parts. The first part is composed of Chapter 3 and Chapter 4, in which the three methods for regime-switching models are investigated. Firstly, we discuss the Monte Carlo simulations and the finite difference methods for pricing European options under the regime-switching model in Chapter 3. A comparison is presented for the two methods applied to price European options on stocks having up to four regimes. The numerical performance shows that the Monte Carlo simulation outperforms the Crank-Nicolson finite difference method even though the option is written on one single underlying stock with two regimes, which is the simplest case. Even though both methods have linear growths, as the number of regimes increases, the computational time of the Crank-Nicolson finite difference scheme grows much faster than that of the Monte Carlo simulation. This verifies that the Monte Carlo simulation is a more efficient numerical approach compared to solving the partial differential equation system numerically.
In Chapter 4, we investigate tree methods under the regime-switching models. It is widely accepted that the existing trinomial tree method proposed by Yuen and Yang in [87] is the most efficient tree approach for regime-switching models. Hence in the chapter, we give a rigorous convergence analysis of their approach. According to our proof, the convergence speed of the method is of order $O(n^{-\beta})$, where $\beta = 1/2$ when the payoff function of the option contract is discontinuous (such as digital options) and $\beta = 1$ otherwise. Apart from the analytical proof, we also test our results numerically, which can be found in the chapter as well.

The second part contains Chapter 5 and Chapter 6, in which the connection between tree approaches and finite-difference methods is studied. It is well-known that under the Black-Scholes economy, the trinomial tree approach can be viewed as a special case of the explicit finite-difference method, the relation is connected by an equivalence proposed by Brennan and Schwarz in [11]. In Chapter 5, we extend this basic property to the regime-switching model. Since none of the existing trinomial tree approaches in literature can be applied to build up such a relationship under the regime-switching model, we present a new trinomial tree method for the model. The new method requires only $K + 2$ nodal values ($K$ is the total amount of regimes) involved instead of $3K$ nodal values in the Yuen and Yang’s trinomial tree method. The equivalence of the new trinomial tree approach and the explicit finite-difference method is proven analytically and verified numerically.

Then in Chapter 6, we extend the connection to Heston’s stochastic volatility model for the first time. Similar to Chapter 5, a new simple tree approach is developed in this chapter. The tree approach is simple to implement, computationally cheap and has a very clear financial interpretation. Besides, the connection between the tree and explicit finite-difference has set up. Numerical results show that the new simple tree method outperforms the existing, most efficient tree approach for the Heston model.
Acknowledgements

This PhD has been a truly life-changing experience for me and it would not have been possibly finished without the support and guidance that I received from many people.

First of all, I would like to express my sincere appreciation to my supervisor, Prof. Song-Ping Zhu for the continuous support of my entire PhD study and research. Without his help and encouragement, I could not imagine how I would get through the hardship especially at the beginning of my research life. It was his guidance and insisting that made me grow up and eventually become an independent researcher. He showed me a perfect example of what a professional researcher with rigorous attitudes should do, which has been determined as a goal that would be pursued in my future life.

Besides my supervisor, I am especially grateful for my former co-supervisor, Dr Ivan Guo, who has contributed immensely to both my personal and professional time in Wollongong. I enjoyed the stimulating discussions with him. His cleverness and dedication have inspired me numerous research ideas. Another great appreciation is given to Dr Guillaume Leduc, who provided a great opportunity of a cooperation from which I learned a lot.

My gratitude also comes to some outstanding researchers I have encountered during my PhD life, including Dr Xiaoping Lu, Dr Boda Kang, Dr Marianito Rodrigo, etc. Their friendliness and warm help encouraged me to further improve myself. In addition, I thank my fellow friends, Guiyuan Ma, Dong Yan, Ziwei Ke, Xinjiang He, Sha Lin, Michael Mampusti, Tom Pederson and many others for their care and help as well as the joy we had together, all of which are valuable and unforgettable memories of mine.

I appreciate my family for the years of support and understanding my pursuing the PhD degree. Their sincere concern about my life is always one of the major motivations that push me forward.

Last but not least, the financial support from the Chinese Scholarship Council (CSC) and the University of Wollongong (UOW) is gratefully acknowledged, without which I could not carry out such a wonderful journey in research.
Contents

1 Introduction and literature review ........................................................................... 1
  1.1 History of options ................................................................................................. 1
      1.1.1 Thales of Miletus .................................................................................. 1
      1.1.2 Tulip bulb bubble ................................................................................. 2
      1.1.3 The birth of the modern options market ............................................. 3
      1.1.4 Options pricing theory ....................................................................... 3
  1.2 Development of the options pricing theory ....................................................... 4
      1.2.1 Early approach (1900-1973) ............................................................... 4
      1.2.2 Non-arbitrage pricing theory (1973-now) ........................................... 5
      1.2.3 The main challenge ........................................................................... 6
      1.2.4 Computational approaches ............................................................... 7
  1.3 Motivation of the thesis ....................................................................................... 10
  1.4 Structure of the thesis ........................................................................................ 11

2 Mathematical background ...................................................................................... 14
  2.1 Stochastic processes .......................................................................................... 14
      2.1.1 Markov chain .................................................................................... 15
      2.1.2 Brownian motion ............................................................................. 16
      2.1.3 Geometric Brownian motion ........................................................... 17
  2.2 Black-Scholes-Merton model ........................................................................... 18
  2.3 Numerical approaches ...................................................................................... 20
      2.3.1 Monte Carlo simulations .................................................................. 21
      2.3.2 Finite-difference methods and associated schemes ....................... 22
      2.3.3 Binomial tree and trinomial tree methods ..................................... 27
      2.3.4 Historical established connection between the numerical techniques . 31
3 Pricing European options on regime-switching assets: A comparative study of Monte Carlo and Finite Difference Approaches 34

3.1 Introduction .................................................. 34
3.2 Model settings and notations ................................. 36
3.3 Monte Carlo simulation ....................................... 38
  3.3.1 Fundamental Monte Carlo .............................. 38
  3.3.2 Antithetic variates .................................. 39
  3.3.3 Control variates .................................... 40
  3.3.4 Simulating total occupation time ..................... 41
3.4 Finite difference method .................................... 42
3.5 Numerical performances and comparisons .................. 45
  3.5.1 MC vs CN .......................................... 45
  3.5.2 Comparison among simulations ....................... 47
3.6 Conclusion .................................................. 49

4 Convergence rate of regime-switching trees 51

4.1 Introduction .................................................. 51
4.2 Settings ..................................................... 56
  4.2.1 Payoff function class ................................ 56
  4.2.2 Black-Scholes discretization ......................... 57
  4.2.3 Regime-state discretizations ......................... 58
  4.2.4 Occupation time random variables ................... 59
4.3 Option price closed-form formula .......................... 60
  4.3.1 Black-Scholes model ................................ 60
  4.3.2 Regime-switching Black-Scholes model ............... 60
4.4 Outline of the chapter ...................................... 63
4.5 Partial discretization error ................................ 64
4.6 Full discretization error ................................... 69
4.7 Polynomial payoffs ......................................... 71
4.8 Yuen and Yang trinomial model ............................ 72
4.9 Numerical results .......................................... 73
4.10 Auxiliary results ................................................. 77
  4.10.1 Rate of convergence for Black-Scholes discretizations .......... 77
  4.10.2 Occupation time discretization error ........................ 79
  4.10.3 About the norm $\chi^2$ on $K^{(2)}$ ........................ 80

Part II: The connection between tree methods and finite difference
methods under stochastic volatility models 82

5 A new trinomial tree method for regime-switching models 83
  5.1 Introduction .................................................. 83
  5.2 Priori knowledge and notations ................................ 86
  5.3 New trinomial tree approach .................................. 87
  5.4 Equivalence to explicit finite difference ....................... 89
  5.5 Convergence rate ............................................. 92
  5.6 Two-state case ............................................... 92
  5.7 Numerical performance ....................................... 93
  5.8 Conclusion .................................................. 96

6 A new simple tree approach for Heston’s stochastic volatility model 97
  6.1 Introduction .................................................. 97
  6.2 Models and notations ......................................... 99
    6.2.1 Heston’s stochastic volatility model ...................... 100
    6.2.2 The regime-switching model ............................ 101
  6.3 The new simple tree approach ................................ 103
    6.3.1 Method of Lines ......................................... 103
    6.3.2 New trinomial tree for regime-switching models .......... 107
    6.3.3 New simple tree for the Heston model .................. 110
    6.3.4 The tree in the original coordinate plane ............... 112
  6.4 Equivalent to the explicit finite difference .................. 115
  6.5 Numerical performance ....................................... 118
    6.5.1 European options ......................................... 118
    6.5.2 American options ....................................... 121
    6.5.3 Equivalence test ......................................... 122
6.6 Conclusion ............................................................ 123

7 Conclusions ......................................................... 124

Bibliography ............................................................. 126

A Algorithms ............................................................ 134

B Proof of theorems .................................................. 138
  B.1 Proof of Theorem 5.5.1 ........................................ 138
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>The graphical illustration for the binomial tree model</td>
<td>27</td>
</tr>
<tr>
<td>2.2</td>
<td>The graphical illustration for the trinomial tree model</td>
<td>30</td>
</tr>
<tr>
<td>4.1</td>
<td>The convergence behaviour for pricing European put options in the ITM case where $S_0 = 90$ and initial regime $\alpha_0 = 1$</td>
<td>75</td>
</tr>
<tr>
<td>4.2</td>
<td>The convergence behaviour for pricing European put options in the ITM case where $S_0 = 90$ and initial regime $\alpha_0 = 2$</td>
<td>75</td>
</tr>
<tr>
<td>4.3</td>
<td>The convergence behaviour for pricing European put options in the ATM case where $S_0 = 100$ and initial regime $\alpha_0 = 1$</td>
<td>76</td>
</tr>
<tr>
<td>4.4</td>
<td>The convergence behaviour for pricing European put options in the ATM case where $S_0 = 100$ and initial regime $\alpha_0 = 2$</td>
<td>76</td>
</tr>
<tr>
<td>4.5</td>
<td>The convergence behaviour for pricing European put options in the OTM case where $S_0 = 110$ and initial regime $\alpha_0 = 1$</td>
<td>76</td>
</tr>
<tr>
<td>4.6</td>
<td>The convergence behaviour for pricing European put options in the OTM case where $S_0 = 110$ and initial regime $\alpha_0 = 2$</td>
<td>76</td>
</tr>
<tr>
<td>4.7</td>
<td>The convergence behaviour for pricing digital put options in the OTM case where $S_0 = 90$ and initial regime $\alpha_0 = 1$</td>
<td>76</td>
</tr>
<tr>
<td>4.8</td>
<td>The convergence behaviour for pricing digital put options in the OTM case where $S_0 = 90$ and initial regime $\alpha_0 = 2$</td>
<td>76</td>
</tr>
<tr>
<td>4.9</td>
<td>The convergence behaviour for pricing digital put options in the OTM case where $S_0 = 100$ and initial regime $\alpha_0 = 1$</td>
<td>77</td>
</tr>
<tr>
<td>4.10</td>
<td>The convergence behaviour for pricing digital put options in the OTM case where $S_0 = 100$ and initial regime $\alpha_0 = 2$</td>
<td>77</td>
</tr>
<tr>
<td>4.11</td>
<td>The convergence behaviour for pricing digital put options in the OTM case where $S_0 = 110$ and initial regime $\alpha_0 = 1$</td>
<td>78</td>
</tr>
</tbody>
</table>
4.12 The convergence behaviour for pricing digital put options in the OTM case where $S_0 = 110$ and initial regime $\alpha_0 = 2$ ................................. 78

5.1 The graphical illustration of the new trinomial tree approach applied to a two-state regime-switching model ..................................................... 93

6.1 The graphical illustration for a two-state regime-switching model .............. 109
6.2 The graphical illustration for the tree structure of the new simple tree approach ................................................................. 111
6.3 The graphical illustration for the $t-w$ plane of the new simple tree approach in the transformed coordinate plane ............................................. 111
6.4 The graphical illustration for the $t-X$ plane of the new simple tree approach in the transformed coordinate plane ........................................... 112
6.5 The graphical illustration for the $t-v$ plane of the new simple tree approach in the original coordinate plane ................................................. 115
6.6 The graphical illustration for the $t-S$ plane of the new simple tree approach in the original coordinate plane ................................................. 115
6.7 The graphical illustration for the “indirect equivalence” between the explicit finite difference and the new simple tree method under the Heston model . 118
6.8 The convergence behaviour of the new simple tree approach for pricing European call options in the case where $T = 0.25$ ............................ 120
6.9 The convergence behaviour of the new simple tree approach for pricing European call options in the case where $T = 0.5$ ............................ 120
List of Tables

3.1 The comparison between MC and CN in a low-frequency case where $\lambda_{ji} = 1$ 
(i $\neq$ j) ................................................................. 46
3.2 The comparison between MC and CN in a low-frequency case where $\lambda_{ji} = 
100$ (i $\neq$ j) ................................................................. 46
3.3 The comparison among different MC simulations for a two-state regime-
switching model in which $\lambda_{ji} = 1$ (i $\neq$ j) ......................... 48
3.4 The comparison among different MC simulations for a two-state regime-
switching model in which $\lambda_{ji} = 100$ (i $\neq$ j) ..................... 48
3.5 The comparison among different MC simulations for multi-state regime-
switching models in which $\lambda_{ji} = 1$ (i $\neq$ j) ......................... 49
3.6 The comparison among different MC simulations for multi-state regime-
switching models in which $\lambda_{ji} = 100$ (i $\neq$ j) ................. 49
4.1 The convergence behaviour for pricing European put options in the ITM 
case where $S_0 = 90$ ......................................................... 74
4.2 The convergence behaviour for pricing European put options in the ATM 
case where $S_0 = 100$ ......................................................... 74
4.3 The convergence behaviour for pricing European put options in the OTM 
case where $S_0 = 110$ ......................................................... 75
4.4 The convergence behaviour for pricing digital put options in the ITM case 
where $S_0 = 90$ ............................................................... 77
4.5 The convergence behaviour for pricing digital put options in the ATM case 
where $S_0 = 100$ ............................................................... 77
4.6 The convergence behaviour for pricing digital put options in the OTM case 
where $S_0 = 110$ ............................................................... 78
5.1 The convergence behaviour of the new trinomial tree approach for a two-state regime-switching model ................................. 94
5.2 The numerical proof of the equivalence between the new trinomial tree approach and the explicit finite difference method ......................... 95
5.3 The numerical comparison of the new trinomial tree method with Yuen and Yang’s trinomial tree approach ............................. 95
6.1 The numerical results of the new simple tree approach for pricing European call options .......................................................... 119
6.2 The convergence behaviour with respect to variable $t$ in the case where $T = 0.25$ ................................................................. 120
6.3 The convergence behaviour with respect to variable $t$ in the case where $T = 0.5$ ................................................................. 120
6.4 The convergence behaviour with respect to variable $w$ in the case where $\Delta w = 0.02$ ............................................................ 121
6.5 The convergence behaviour with respect to variable $w$ in the case where $\Delta w = 0.01$ ............................................................ 121
6.6 The numerical results of the new simple tree approach for pricing American put options .......................................................... 122
6.7 The numerical proof of the “indirect equivalence” between the explicit finite difference and the new simple tree ..................................... 123
Chapter 1

Introduction and literature review

1.1 History of options

Options, as one main category of financial derivatives, play an important role in hedging, speculation, spreading and synthetics in the financial market nowadays. It gives holders the right, but not the obligation, to buy or sell their underlying assets at a pre-determined price. The right is secured by paying what is called the price of the option, which is relatively cheaper compared to its underlying assets. This means that with the same amount of money, one can control a larger number of shares by buying options than investing in the shares themselves. For the speculators, options are ideal financial instruments because they provide a greater leverage effect while risk averters view them as a functional tool to hedge against potential loss.

Although options are very popular nowadays, it is rarely known that they were used to be notorious and had a very bad reputation. They were even banned and made illegal for decades until the late nineteenth century. Here we start with the history of options and try to find out the main reason that brings them back from the ban list and leads them to popularity.

1.1.1 Thales of Miletus

The earliest option-like contract in the history was recorded in the book, Politics, written by the famous Aristotle in the mid-fourth century. In one section, he briefly mentioned a story of another ancient Greek philosophy, Thales the Milesian, to show how useful it is to attain a monopoly in a market.
The story started with Thales predicting a vast harvest of olives in his region by studying stars. Once being harvested, olives would then need to be processed by olive presses. Knowing that the right of using olive presses was about to be sold at a higher rate, he decided to make a huge profit by what we later called as cornering the market. However, Thales was not rich enough to buy all the olive presses. Instead of owning all of them, he offered the owners a deal which secured his right of using the olive presses with the current rate during the harvest season by paying them a small amount of money each. The owners agreed to accept the deal as it was the off-season and they got a guaranteed amount of income as well. When the harvest time came around, it turned up Thales’ prediction was correct and he made a huge profit by selling his rights for a much higher price.

Even though the term “options” is not used at that time, it is still quite clear that the contracts Thales signed with the owners of the presses are basically European call options. He paid out for the right, but not obligation, to use the olive presses at a fixed rate and then exercised his options for a profit.

1.1.2 Tulip bulb bubble

Another notable occurrence of options is an event in the seventeenth century. It is widely known as the “Tulip Bulb Mania” which resulted in an adverse effect on Holland’s economy. During the 1630s, tulips in Holland began to be appreciated and further were considered to be the symbols of the Dutch aristocracy. Their popularity spread into Europe and throughout the whole world, which led to a dramatically increasing demand for tulip bulbs in the market.

By that time call and put options have been developed and primarily used for hedging purposes. For instance, farmers who grew tulip bulbs would buy put options to secure their profits against the possible decline in price. Wholesalers would buy call options in case the price of tulip bulbs goes up. However, they are not the same as the call and put options today since the market at that time was very informal and unregulated.

As the demand for tulips continued to increase, the price of the tulip bulbs increased accordingly. The raising price attracted lots of people to put their money into the secondary market which emerged to enable everyone to speculate the price of tulip bulbs. Many individuals and families invested heavily in these contracts, with some using all their
savings and some even borrowing against their properties.

However, the investors actually pushed the price of tulip bulbs to a level that was unsustainable and dangerous. Buyers started to disappear as they thought the price was too high to be acceptable. As the bubble burst, many people lost all their money as well as their properties and the economy of Holland went into a disastrous recession.

1.1.3 The birth of the modern options market

Back in the time when the market was informal and unregulated, many investors could easily walk away from the clearings of the option contracts, which resulted in a bad reputation of options. Together with the great leverage effect and liquidity issue, options were banned largely in many places throughout the world for over a hundred years.

The liquidity problem was aroused by the old, manual trading mechanism. For a long period of time, the options trading was laborious and inefficient. An option buyer usually needed to call an intermediary who would look for a seller with exactly the same price to make the deal. Since there were no correct pricing structures for options, buyers and sellers could basically set prices to be whatever they wanted, which made the trade even harder. Back in time, there were some organizations which aimed at improving the efficiency of options trading. One typical example was the “Put and Call Brokers and Dealers Association”, which tried to build a networking so that buyers and sellers could be matched. However, the complexities and inconsistent prices made it very difficult for investors to believe options were viable financial instruments. Thus options suffered from serious liquidity issues.

Then in 1973, the Chicago Board of Options Exchange (CBOE) was created and the Options Clearing Corporation was also established for the purpose of ensuring the investors to fulfil the obligation of buying options. For the first time, options were traded with proper regulations and the concerns of investors were removed. It was since that time options started to become popular financial instruments.

1.1.4 Options pricing theory

At the beginning of the CBOE opened, it only offered call options trading on 16 different stocks. Put options were not even available since they had yet to be standardized. The liquidity issue was not significantly changed. Many investors chose to wait and see because
it was still not sure whether the price of an option stood for good value in money or not.

Then still in 1973, Fisher Black and Myron Scholes presented a simple mathematical formula (the Black-Scholes formula or the BS formula) that allowed the general public to value fair prices of options by themselves. The elegant formula required only a few parameters that could be either observed in the market or calculated by historical data. And the model built by Black and Scholes was widely accepted as the start of the modern options pricing theory. The theory eventually removed the concerns of investors and made them more comfortable about trading options. Since then the daily volume of options contracts exchanged in the CBOE started to grow dramatically, from roughly 20,000 in 1977 to 350,000 today.

A general consensus is that the BS formula is the turning point of the options’ development history. It remarkably changes the fate of options, from the notorious hatred to the most popular financial instruments. Thus the options pricing theory is required to be studied to maintain a fair and healthy financial market so that it will continuously contribute to the whole world and the human race.

1.2 Development of the options pricing theory

Without the pricing theory, options can never reach the level of what it becomes today. Like any other theories, the cornerstone, the Black-Scholes model, did not suddenly appear from nowhere. Lots of previous research should not be ignored as they are very important parts of the options pricing theory. So here we proceed to review the development of the options pricing theory to discover what it was in the past, how Black and Scholes present the model and what the main challenges are nowadays.

1.2.1 Early approach (1900-1973)

At the beginning of the nineteenth century, a French mathematician Louis Bachelier first discovered the mathematical theory of Brownian motion and further present a formula for pricing options in his doctoral dissertation, *Théorie de la spéculcation* [2]. In his thesis, the author assumed the stock price to follow a time-dependent normal distribution, under which the mathematical expectation of the buyer of a call option is zero. The formula is further tested to provide accurate values for options with time to expiry being less than
one and a half months. This is widely accepted as the first attempt to employ the theory of stochastic processes into options pricing theory.

More than half a century later, Paul A. Samuelson studied Bachelier’s dissertation and found the flaw in his formula. In Samuelson’s paper [77], he pointed out that “arithmetic” Brownian motion failed to capture the property of limited liabilities for stocks and led the stock price to be negative with a fairly large probability. This resulted in the options priced by Bachelier’s formula increased in price indefinitely as the time to maturity grew, which violated the rule that option prices should always be lower than the underlying stock prices. To overcome the flaw, Samuelson assumed the return of stocks to follow a “geometric” Brownian motion instead, which prevents the stock price to fall below zero. Then by postulating that the expected return of an option was a martingale, he successfully derived a formula which was almost identical to the Black-Scholes-Merton’s formula except for two unknown parameters \( \alpha \) and \( \beta \), which were expected returns for stocks and warrants, respectively. These two parameters, which are known as the stock’s risk premium or warrant’s risk premium, were difficult to estimate from the market data and made the formula nearly impossible to use.

1.2.2 Non-arbitrage pricing theory (1973-now)

Then in 1973, a breakthrough was made by Fischer Black, Myron Scholes and Robert Merton for their celebrated Black-Scholes-Merton (BSM) model (also known as the Black-Scholes model) as well as formula. The original Black-Scholes model was presented in [5] with six assumptions:

i) The market is frictionless, which means no transaction cost is applied.

ii) The riskless interest rate is known and constant.

iii) Unlimited borrowing is allowed at the constant interest rate.

iv) No penalty fees are required for short selling.

v) Stocks follow a geometric Brownian motion with a constant volatility.

vi) No dividends or other distributions are paid by stocks.

With all six assumptions, Black and Scholes managed to create a portfolio which was composed of simultaneous and offsetting positions in an option and a stock. Being held
continuously, the hedged portfolio was further proved to have a certain return, which was equal to the riskless interest rate, by what was later known as the non-arbitrage principle. The simultaneous change of the portfolio further led to a partial differential equation (PDE) with a boundary condition (rigorously terminal condition, see expression (8) on page 643 of [5]). The solution of the PDE was derived analytically and was known as the celebrated Black-Scholes formula. This is widely accepted as the born of the modern theory of options pricing.

Then in the same year, Merton published a paper [63] that explained some mathematical concepts of the BS approach. It was in this paper that for the first time, the dynamic of stock prices was written into the famous stochastic differential equation of the geometric Brownian motion. A rigorous mathematical derivation of the PDE was given by using stochastic calculus, specifically, the Ito’s lemma. Merton had also extended the model to include dividend payments and exercise price changes.

The born of the BSM model provided a completely new perspective of options, that was, an option could be replicated by merely a long position in stocks and a short position in riskless bonds. Thus the pricing issue was linked to the idea of hedging with trading underlying assets dynamically.

In addition, the BS model also presents a deeper insight of the whole field of financial mathematics. On one side, economically, options are priced by the “non-arbitrage” approach by forming a dynamic portfolio. On the other side, mathematically, option prices are obtained by taking expectations under a properly selected “risk-neutral” probability measure. The two seemingly uncorrelated approaches reach an identical result, which is the BSM model. After the systematic study by a number of researchers, the Fundamental Theorem of Asset Pricing is then proposed (See [18, 71–73]), which can be viewed as another breakthrough of the subject of the modern mathematical finance.

1.2.3 The main challenge

While the main merit of the BSM model is its simplicity in application, it is constructed, purely theoretically, on an ideal market that satisfies all the six assumptions. The unrealistic natures of these assumptions resulted in preventing the formula from producing perfectly accurate values when tested empirically.

One of the most noted shortcomings is the assumption of constant volatilities. The
volatility is the only parameter in the BSM formula that cannot be observed from the market. It is often measured by the standard deviation from the historical data of underlying assets. Such a value is widely denoted as the “explicit” volatility. In fact, if an option is actively traded in the market, its price will be readily obtainable. Then by applying the BSM formula in a reverse way, what is called the “implied” volatility is calculated.

From the analysis of [5], equalities of estimated volatilities should hold among different option contracts written on the same underlying asset. A simple experiment can be carried out by calculating and comparing the implied volatilities for two different options, all the same, but one’s strike price is higher than the other. If the equality does not hold, then there is an evidence that the volatility is not constant.

However, it was found that the market implied the pattern of volatility was not flat with respect to all the parameters. In some markets, those patterns formed a skew curves while some others were more of a smile curve. A study could be found in [62] that showed implied volatilities changed according to strike prices (whether the option was in, at or out of the money) and time to maturity. Hence the formula turned out to somehow overprice at the money options while underpriced the deep in or out of the money options.

Hence the main challenge of the options pricing theory has been to find a more adaptive model to reflect the random feature of the market volatilities. A lot of extension models can be found in existing literature. One of the most popular extensions is the stochastic volatility (SV) model, as presented in [35, 39, 42, 79], which assumes the volatility itself to follow another diffusion process.

However, as the extra dimension of randomness being introduced, the SV models are far more complicated compared to the BSM model. As a result, analytical solutions are less likely, most time impossible, to be found. In addition to the complexity of the model, more sophisticated options, such as American options, cannot be priced by analytical solutions, either. Thus another challenge of the options pricing lies in the fact that efficient numerical methods are in great need.

1.2.4 Computational approaches

At the broadest level, there are three main methods to price options numerically: tree approaches (binomial and trinomial trees), Monte Carlo simulations and numerical partial differential equation methods (finite-difference methods). The advantages of each pricing
method largely depend on the type of contracts being priced and the model used.

Monte Carlo simulations were first introduced to options pricing by Boyle in [7] for the valuation of European options under the BS model. The principle of the method is to calculate the expected value of a quantity, which is a solution to the stochastic differential equation (SDE), by repeated random samplings. A detail introduction of Monte Carlo simulations in finance can be found in [28, 51]. It is also known as a highly flexible method, with the capability of pricing complex option contracts, as presented in [25, 29, 38, 57, 81, 82]. On the other hand, tree methods have received great popularity among practitioners and researchers since the pioneering work by Cox et. al. [18]. It was also discussed in [8, 43, 80]. Many other extensions can be found in literature, including Boyle [9], in which the method was extended to the case where the option contracts were written on two underlying stocks, Boyle [8], in which a trinomial tree method was constructed based on the binomial tree by Cox et. al. [18]. The key advantage of the method is its simplicity as well as clear economic interpretation. The tree structure approximates the distribution of the stock price so that the method can solve many complex options explicitly. Finite-difference methods are a class of numerical schemes to solve PDEs. It was first applied to the options valuation with dividend payments taking into consideration by Schwartz [78]. The method is considered as the most efficient when the dimension of the problem is low, especially one or two. However, the finite-difference method suffers greatly from the “curse of dimensionality”. The efficiency of the method declines significantly with respect to the number of underlying assets involved.

When it comes to the regime-switching model, it is quite astonishing that there has been remarkably little research into Monte Carlo simulations and finite-difference methods. Lemieux [51] proposed an algorithm of Monte Carlo simulations for pricing European options under a general multi-state regime-switching model. Based on her approach, Hieber and Schere [40] extended the method to Asian options by adopting the Brownian bridge technique. Mielkie [64, 65] studied finite-difference methods for numerically solving the PDE system from the regime-switching economy. Hence there is a need to investigate the numerical performance of the two methods when applied to the regime-switching model. This becomes the motivation of carrying out a comparative study of these two methods, which forms the main content of Chapter 3.

Tree methods, on the other hand, have been well-established during the past three
decades and became the main numerical approach for pricing options with regime-switching assets. Bollen first introduced a lattice-based method for valuing European and American-style options. His method was further modified by Liu [53] and Yuen and Yang [87]. Other tree methods can also be found in [16, 17, 47, 54–56, 86, 88]. It should be mentioned that most of these articles do not contain the part of convergence analysis. The authors do verify the convergence of their approaches numerically but the lack of convergence proof requires additional caution when these methods are applied on a case-by-case basis. Ma et. al. [59, 60] attempted to give a convergence proof for Yuen and Yang’s trinomial tree approach. However, their results were only valid for options with smooth payoff functions subject to boundedness conditions. Their issue is solved in Chapter 4, in which we present a rigorous proof of tree methods applied to a broad family of piecewise smooth payoff functions.

Another interesting problem lies in the fact that, the trinomial tree methods can be viewed as a special case of explicit finite difference methods when applied to the BSM model. The equivalence was proven by Brennan and Schwarz in their article [11]. Such a simple and elegant relationship should hold for more complex models extended from the BSM model. It is the curiosity that motivates us to study the connection between the two methods under the more adaptive stochastic volatility models, which is the main idea of Part Two of our thesis. Ma et. al. [61] presented a connection of Yuen and Yang’s tree with explicit finite-difference methods under the regime-switching frameworks. The authors established an “equivalence” except for a second-order-negligible difference term in the recursive formulas. This is, however, not quite the same as the perfect equivalence presented by Brennan and Schwarz, in which the coefficients of the recursive formulas of the two methods are identical. Such a study of the relationships under the regime-switching model of the two contents is presented in Chapter 5.

In contrast to an existing literature of the investigation of the equivalence between the two methods in the regime-switching case, the connection for Heston’s stochastic volatility model has never been brought up yet. Tree methods for the Heston model have been discussed in [3, 31, 50, 53, 83]. The main difficulty of constructing an efficient tree for the Heston model lies in the fact that, without the assumption of constant volatility, the recombining property does not hold naturally anymore. The difficulty was solved by the transformation methods, which were introduced by Beliaeva and Nawalkha [3] and
Leisen [50] to make the transformed stochastic process have a constant volatility term. Their approaches involved a “multi-jump” technique to tackle the problem aroused from the singularity appeared in the transformed drift term. Thus the tree structures are impossible to be used to build any connections with finite-difference methods. On the other hand, another approach was presented by Guan and Xiaoqi [31] and Vellekoop and Nieuwenhuis [83]. Instead of adopting transformations to ensure the recombining property, the nodes not recombined are interpolated or extrapolated to a pre-defined grid. This is clearly too far away from the grid structure of finite-difference methods. Hence we develop a new simple tree approach in Chapter 6, that allows us to further investigate the relationship between the new approach and finite-difference methods.

1.3 Motivation of the thesis

As mentioned in the previous sections, two main challenges remain in the options pricing theory: i) establishing more realistic models and ii) developing more efficient numerical techniques. This thesis deals with the second challenge. The main concentration of this thesis is to study the three key numerical methods for options pricing under the more adaptive stochastic volatility models, in particular, the regime-switching models and the Heston model.

Unlike closed-form solutions, numerical techniques are a lot more complicated in many ways. The first thing that draws lots of concerns is the efficiency, which is always tested by numerical experiments and shown tables or figures. These experiments are like “guidance” and tell researchers or practitioners how well the method is and what attention they should pay. A method without these performance tests is not reliable and not preferred. For instance, Monte Carlo simulations and finite-difference methods for regime-switching models are rarely discussed. This motivates us to implement and test the two methods for regime-switching models in Chapter 3.

On the other hand, the convergence analysis is another necessary requirement to numerical methods. In most situations, convergence analysis is very difficult and requires very advanced mathematics. Thus it is very often to see that many authors only develop a new numerical method without giving an analytical proof of its convergence. Of course in those articles, numerical experiments can always be found to verify the convergence, but
only a few cases can be included in there. The method may not even be feasible to solve
the problems in some special parameter space. The significance of the convergence analysis
is to tell users that what this method can do and what it cannot. One should always be
very careful when using the methods without a formal proof of convergence. This is why
we propose an analytical proof of convergence of tree methods for the regime-switching
model in Chapter 4.

Another interesting topic in numerical methods of options pricing is that there is a
connection between trinomial tree methods and explicit finite-difference methods. Al-
though the two methods have completely different principles and interpretations, they
have been proven to be the same under some transformations and a certain selection of
space increments. This equivalence provides great convenience in convergence analysis for
tree methods since finite-difference methods are characterized by the wealth of existing
theory. Once a connection of this type is constructed, the convergence of the tree method
can then be proven with the existing powerful tools of finite-difference methods. However,
such an equivalence holds only for the BS model now. Hence we are motivated to extend
the relationship to more adaptive models, such as regime-switching models (see Chapter
5) and Heston’s stochastic volatility model (see Chapter 6).

1.4 Structure of the thesis

Before proceeding to the difficult numerical schemes and results, we present a brief mathe-
matical background in Chapter 2, including stochastic processes and the BS model, which
are the theoretical base of the thesis. In addition, we also provide an introduction of each
of the key numerical method and how to apply them to the very basic BS model. This
preliminary knowledge helps readers have a better understanding of the main results in
the later chapters.

The rest of this thesis is divided into two parts. In the first part, which includes Chapter
3 and 4, the three numerical techniques for options pricing under the regime-switching
models are investigated. Since little research has been done for Monte Carlo simulations
and finite-difference methods with the regime-switching assets, a study is presented in
Chapter 3 for comparing the numerical efficiency of the two methods. For the Monte
Carlo simulations, we modify the algorithm from Lemieux [51] so that the new algorithm
is free from the pre-estimation of the number of random numbers to be generated, which could result in computational waste. Two antithetic variates techniques for the simulation are also tested in the chapter. As for the finite-difference method, the Crank-Nicolson scheme is applied due to the unconditional stability and second-order convergence speed. The two methods are compared in the sense that the levels of accuracy are the same while checking the computational time for each method. The one with less computing time thus is obviously the more efficient.

Even though there have been a fairly amount of articles discussing tree methods for regime-switching models, the convergence analysis for most of them is not mentioned but only verified numerically. The lack of convergence proof requires additional caution when these approaches are adopted on a case-by-case basis. Hence in Chapter 4, such a rigorous proof is presented to fill in the blank space. The main results show that the speed of convergence of tree methods vary according to the different types of payoff functions that option contracts have. The convergence analysis is verified to be solid by numerical experiments in the chapter.

The second part of the thesis focuses on the relations between the tree methods and finite-difference methods. Since the trinomial tree method and the explicit finite-difference method are equivalent under the BS model, we wonder that if such a connection still holds for the extension models, for example, the stochastic volatility models. Our first attempt comes to the regime-switching model, which forms the main content of Chapter 5. Realizing it is impossible to build a perfect equivalence with the existing tree approaches, we develop a new trinomial tree method for options pricing under the general multi-state regime-switching model. The new method is proven to be equivalent to the explicit finite-difference method and have a first-order convergence speed when it is applied to European options. The numerical performance also shows that the new method outperforms the current most popular Yuen and Yang’s trinomial tree approach.

With the equivalence holds for the regime-switching model, we proceed to the Heston model. Similar to the regime-switching case, we develop a new simple tree approach, which in each time step the recursive formula requires only four or five nodal values. The new tree method also captures the economic interpretation and is computationally efficient. More importantly, the equivalence of the tree method with finite-difference methods is established for the first time. Numerical examples show that our tree approach outperforms
the existing methods and the connection between the two methods is a perfect equivalence. The results mentioned above is given in Chapter 6, followed by the conclusions in Chapter 7.
Chapter 2

Mathematical background

In this chapter, we introduce the preliminary knowledge, such as stochastic processes, the celebrated BSM models and the basic ideas of the three numerical techniques, all of which are going to be mentioned and discussed in the future chapters.

2.1 Stochastic processes

One important assumption of the modern options pricing theory is that a stock price follows a geometric Brownian motion, which is a type of stochastic processes. In this section, we introduce the general definition of stochastic processes and further some specific ones that are often used in options pricing, including the Markov chain, Brownian motion and Geometric Brownian motion.

Consider a probability triple $(\Omega, \mathcal{F}, \mathbb{P})$, where $\Omega$ is the sample space that contains all the possible outcomes, $\mathcal{F}$ is the $\sigma$-algebra that is the collection of all events and $\mathbb{P}$ is the probability measure. The mathematical definition of a stochastic process (random process) $X_t$ is a family of random variables $\{X_t : t \in T\}$ on the common probability space $(\Omega, \mathcal{F}, \mathbb{P})$ indexed by some set $T$. All of the random variables should take values in the same mathematical space $S$. The set $T$ is called the index set of the stochastic process. There are generally two categories of set $T$: $T \in \{0, 1, 2, 3, \ldots\}$ associates with a “discrete-time” stochastic process while $T \in [0, \infty)$ corresponds to a “continuous-time” stochastic process. Here in this chapter we only consider continuous-time stochastic processes.
2.1.1 Markov chain

Let $X = \{X(t), t \in [0, \infty)\}$ be a continuous-time random process taking values from a countable state space $S$.

**Definition 2.1.1.** The stochastic process $X$ is a continuous-time Markov chain if it satisfies the Markov condition:

$$
P(X(t_n) = s_n | X(t_{n-1}) = s_{n-1}, X(t_{n-2}) = s_{n-2}, ..., X(t_1) = s_1) = \frac{P(X(t_n) = s_n | X(t_{n-1}) = s_{n-1})}{P(X(t_{n-1}) = s_{n-1})},$$

for all $s_1, s_2, ..., s_n \in S$ and any monotonic increasing time sequences $t_1 < t_2 < ... < t_n$.

The intuition behind the Markov chain is that, given a set of previous events, the probability of the process taking a certain value depends only on the one that happens most recent. One important concept to study the continuous-time Markov chain is the transition probability, which is given in the definition below:

**Definition 2.1.2.** Probability $p_{ij}(s, t)$ is called transition probability if

$$p_{ij}(t_1, t_2) = P(X(t_2) = j | X(t_1) = i).$$

The chain is homogeneous if $p_{ij}(t_1, t_2) = p_{ij}(0, t_2 - t_1)$.

In this thesis we only consider the homogeneous chains hence we use notation $p_{ij}(t_2 - t_1)$ for short. Let $P_{ij}(t)$ be the matrix form of $(p_{ij}(t))_{i,j \in S}$. Then it is obvious that as $t \to 0$, the diagonal of $P_{ij}(t)$ converges to the identity matrix. To be more specific, suppose $X(t) = i$ at time $t$. Two types of change could take place in a small time interval $(t, t + \Delta t)$:

i) Nothing happens, with probability $p_{ii}(\Delta t) = \tilde{p}_{ii}(\Delta t) + o(\Delta t)$.

ii) The chain switches to state $j$ ($j \neq i$) with probability $p_{ij}(\Delta t) = \tilde{p}_{ij}(\Delta t) + o(\Delta t)$.

Here $\tilde{p}_{ij}$ and $\tilde{p}_{ii}$ are linear approximations of the transition probabilities $p_{ij}$ and $p_{ii}$. Hence there exists a set of constants $G = \{g_{ij} : i, j \in S\}$ such that

$$\tilde{p}_{ii}(t) = 1 + g_{ii}t,$$

$$\tilde{p}_{ij}(t) = g_{ij}t.$$
Clearly $g_{ii} \leq 0$, $g_{ij} \geq 0$ and $\sum_{i=1}^{n} g_{ij} = 0$. Such a set $G$, which is often appeared in its matrix form $G = (g_{ij})_{i,j \in S}$, is named the generator of a Markov chain.

It is worthwhile mentioning that generators can be viewed as the “first-order derivative” of transition matrices, with the following properties naturally held by its definition:

$$\lim_{t \to 0} \frac{P(t) - I}{t} = G.$$ 

Thus the transition probability can be expressed as the matrix exponential of $G$ or simply approximated by $I + Gt$ with a convergence order of $O(t)$.

### 2.1.2 Brownian motion

The continuous-time Markov chain is considered as “discrete” in the sense that it takes values in the integers or in some other countable sets. Brownian motions, however, do not belong to this category. A Brownian motion is a random process $W = \{W(t) : t \geq 0\}$, indexed by a continuous time and taking values from all real numbers $\mathbb{R}$. A general definition is given below:

**Definition 2.1.3.** A continuous-time random process $W(t)$ is a linear Brownian motion starting with $x \in \mathbb{R}$, if the following holds:

1. **Starting point:** $W(0) = x$,
2. **Independent increments:** for any given set of time $\{t_i\}_{i=1}^{n}$, the increments $\{W(t_{i+1}) - W(t_i)\}_{i=1}^{n-1}$ are independent random variables.
3. **Normal distribution:** for any $t, s \geq 0$, the random variable $W(t) - W(s)$ is normally distributed with mean 0 and variance $t - s$.
4. **Continuous path:** the mapping $t \to W(t)$ is almost surely continuous.

If $x = 0$, such a process is called the standard Brownian motion. It is obvious that all linear Brownian motions can be obtained by a linear transformation (rotation and translation) of the standard Brownian motion.

Due to (3) in Definition 2.1.3, if we select $s = 0$, together with $W(0) = 0$, it is obvious that $W(t)$ is also normally distributed with zero mean and $t$ variance. In addition, although $W(t)$ is almost surely continuous, it is almost surely nowhere differentiable.
2.1.3 Geometric Brownian motion

Using Brownian motions to model stock prices was presented in [2] and the biggest criticism was that stock prices could not become negative. Thus a Geometric Brownian motion (GBM) was proposed in [77], which is a continuous-time stochastic process that the logarithm of the randomly varying quantity follows a Brownian motion. Consider a linear Brownian motion $B(t)$ and a GBM starts with an initial value $S_0$ can be expressed by $S_t = S_0e^{B(t)}$.

A more general way of formulating the GBM is by stochastic differential equations $SDE$, which is given below

**Definition 2.1.4.** A random process $S_t$ is said to be a GBM if the following SDE holds:

$$dS_t = \mu S_t dt + \sigma S_t dW_t$$

(2.1.1)

where $W_t$ is a standard Brownian motion, $\mu$ and $\sigma$ are drift and volatility terms of the GBM.

A full detail introduction of GBM can be found in [46] and here we only mention a few important features that are going to be used in future chapters. Since it is theoretically an exponential Brownian motion, it has similar properties as Brownian motions except the increments are lognormally distributed.

**Proposition 2.1.5.** The solution to the SDE 2.1.1 with respect to an initial value $S_0$ is given by

$$S_t = S_0\exp \left( \left( \mu - \frac{1}{2}\sigma^2 \right) t + \sigma W_t \right).$$

(2.1.2)

**Proposition 2.1.6.** The local mean and variance of such a process is given by

$$E(S_t|S_0) = S_0e^{\mu t},$$

(2.1.3)

$$Var(S_t|S_0) = S_0^2e^{2\mu t} \left( e^{\sigma^2 t} - 1 \right).$$

(2.1.4)

The two propositions are of the most importance in the options pricing theory. Proposition 2.1.5 can be obtained by simple stochastic calculus and it is the core of constructing Monte Carlo simulations as we will review in Section 2.3.1. The formula (2.1.2) can also
be verified by Itô’s lemma that such a \( S_t \) satisfies the SDE (2.1.1). Proposition 2.1.6 is trivial since \( S_t \) is lognormally distributed. It is the theoretical base of the very popular tree methods as being presented in Section 2.3.3.

## 2.2 Black-Scholes-Merton model

With the six assumptions we mentioned in Chapter 1, Fischer Black, Myron Scholes and Robert Merton established the celebrated BSM model in which options pricing problem was transferred to a parabolic PDE problem. In this subsection, we introduce the BSM model, the method of deriving the governing PDE as well as presenting the BSM pricing formula.

Under the Black-Scholes-Merton economy, the price of a stock is assumed to follow the SDE (2.1.1) with drift and volatility being constant. Since the market is frictionless with unlimited short selling allowed, a portfolio that contains a long position in a call option and a short position in its underlying stock can be made with risk perfectly hedged.

Consider a European-style call option \( C(S_t, t) \) \(^1\) where \( t \) is the current time and \( S_t \) is the price of its underlying stock, which follows (2.1.1). A portfolio \( \Pi(t) \) that is compromised of buying a call and shorting \( \frac{\partial C}{\partial S} \) shares of the stock is represented as

\[
\Pi = C - \frac{\partial C}{\partial S} S. \tag{2.2.1}
\]

Then with some simple stochastic calculus (See [69] for detail), the increment of such a portfolio \( \Pi \) is

\[
d\Pi = \left( \frac{\partial C}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} \right) dt. \tag{2.2.2}
\]

It is worthwhile mentioning that (2.2.2) does not contain the Brownian motion term, which means that the increment has a zero variance or the portfolio has zero risk. In the Black-Scholes world where no transaction cost is charged, the return of such a portfolio \( \Pi \) could remain wholly deterministic if the hedging strategy is continuously held.

On the other hand, due to the non-arbitrage assumption, the rate of the increment of the portfolio \( \Pi \) can only be the riskless interest rate \( r \). Otherwise, arbitrage opportunities

\(^1\)An American-style option can lead to a BS inequality.
will raise: rational investors can either borrow money from banks with \( r \) and invest in the portfolio if the return is greater than \( r \) or short the portfolio then put money into banks if the return is less than \( r \). Either way will guarantee investors a profit with no cost, which clearly violates the assumption.

Even though it seems to be a little bit too realistic, the non-arbitrage assumption is actually reasonable because it describes an equilibrium of the market. Imagine arbitrage opportunities do exist then every investor will take action to make a profit with the price gap, which will disappear soon and an equilibrium without arbitrage is reached.

Due to the reasons above, the left-hand side of (2.2.2) can then be replaced by \( r \Pi dt \) and together with (2.2.1), the price of the call option becomes the solution of the following PDE:

\[
\frac{\partial C}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS \frac{\partial C}{\partial S} - rC = 0.
\]  

(2.2.3)

Black and Scholes suggested a transformation to the PDE (2.2.3) to turn it into a heat equation (See p643 of [5]). With the terminal condition of a European call option, \( C(S_T, T) = \max(S - E, 0) \), the analytical solution of (2.2.3) can be obtained by the convolution of the fundamental solution of the transformed heat equation and the terminal condition. A detail derivation can be found in [84].

\[
C(S, t) = SN(d_1) - Ee^{-r(T-t)}N(d_2)
\]  

(2.2.4)

where \( E \) is the exercise price, \( T \) is the maturity, \( N(\cdot) \) is the cumulative distribution function of a standard normal random variable, given by

\[
N(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}y^2} dy.
\]

Here

\[
d_1 = \frac{\log(S/E) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma \sqrt{T-t}},
\]

\[
d_2 = \frac{\log(S/E) + (r - \frac{1}{2}\sigma^2)(T-t)}{\sigma \sqrt{T-t}}.
\]

Obviously, (2.2.4) is the famous BS formula. It provides the great convenience of pricing an
option as one can simply insert the five parameters into the formula to yield the price of the European call option. However, closed-form solutions like the BS case are extremely rare. For more complicated options, e.g., American options, no efficient analytical solutions have been found yet. This motivates researchers to investigate and develop fast and simple numerical approaches to make up the blank field of the unavailability of closed-form solutions. Finite-difference methods hence are applied to numerically solve the PDE problems as one of the efficient approximations. We will review this in Section 2.3.2.

An alternative way to evaluate options is through the computation of its expected values of its discounted future cash flows. Then an expectation (given below) is taken under a properly selected “risk-neutral” probability measure $Q$ in which investors are insensitive to risk. Below we present an example of the expression of a European call option in the expectation form.

$$C(S_t, t) = e^{-r(T-t)}E_Q(\max(S_T - E, 0)|\mathcal{F}_t).$$

(2.2.5)

In other words, the return of any investments under this measure is the riskless interest rate. Then the SDE (2.1.1) can be written by

$$dS_t = rS_t dt + \sigma S_t dW_t,$$

(2.2.6)

with the drift rate $\mu$ replaced by the interest rate $r$. The theoretical framework of the existence and uniqueness of such a probability measure is provided by the first and second Fundamental Theorem of Asset Pricing ([18, 71–73]).

It should be remarked that the two approaches represent the same value of the same option so (2.2.5) leads to the PDE (2.2.3) as well. The proof can be found in [44], which is also known as the Feymann-Kac theorem. The expectation formulae (2.2.5) makes some other numerical methods possible, e.g., Monte Carlo simulations as well as tree methods, as we will review in Section 2.3.1 and 2.3.3.

2.3 Numerical approaches

In this section, we review the basic knowledge of the three key numerical methods that we concentrate on in this thesis. A brief introduction to these methods applied to the BSM
model is also presented as the base of the future extensions to stochastic volatility models in the later chapters.

2.3.1 Monte Carlo simulations

Monte Carlo simulations are a broad class of computational algorithms that adopt large amounts of repeated random samplings to approximate results. The theoretical base of the algorithms is the Law of Large Numbers (LLN) and the Central Limit Theorem (CLT) (See [30] for details). The LLN guarantees that, with a sufficiently large number of replications, the average among all the experiments converges to the expected value. On the other hand, from the CLT, the obtained result is a normal random variable with expected value being the mean and the variance being the variance of one single experiment over the total number of replications.

The main idea of constructing a Monte Carlo simulation is quite simple. We need to express the problem as a formulation with a random variable. Then the result can be obtained by repeatedly drawing samples from the distribution that the random variable follows. We have now known that under the risk-neutral measure, the solution of the SDE (2.2.6) is given by

\[ S_t = S_0\exp\left(\left(r - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right). \]

Since \( W_T \) is a normal random variable with mean 0 and variance \( T \), the terminal stock price can be further represented by

\[ S_T = S_0\exp\left(\left(r - \frac{1}{2}\sigma^2\right)T + \sigma \sqrt{T}Z\right), \tag{2.3.1} \]

where \( Z \) is a standard normal random variable. Hence for each time we draw a standard normal random number, by substituting it into (2.3.1), we will yield one sample of terminal stock price, which is further inserted into the expression to obtain one sample of option price,

\[ V(S_t, t) = e^{-r(T-t)}\mathbb{E}_Q(p(\cdot)|\mathcal{F}_t), \tag{2.3.2} \]

where \( p(\cdot) \) is payoff functions for different types of options. Path-independent options,
such as European options, require only one value on the path while path-dependent options, for example, Asian options, need a set of values depend on the specified period of time. These can all be obtained simply by the same mechanism we just mentioned. The value of the option is further approximated by averaging among all sample option prices. Pricing options with Monte Carlo simulations are generally simple, especially for some very complex path-dependent options. However, in the case of American options, which contain an optimal stopping boundary, the method becomes far more complicated, with dynamical programming involved, as presented in [12, 13, 57, 81].

The drawback of the Monte Carlo simulation is that it is normally quite efficient-less in the low-dimensional cases when other numerical approaches are also available. The convergence speed is of an order of $O(1/\sqrt{N})$, where $N$ is the number of sample paths. This is to say, with the number of paths increasing 100 times, say, from 10 to 1000, the estimator improves only by one decimal, for example, from 0.1 to 0.01. This is clearly too slow as other methods have much higher convergence rate (tree methods are of first order $O(1/N)$ and Crank-Nicolson finite-difference methods are of $O(1/N^2)$ in terms of time variable). The efficiency can be further improved by applying variance reduction techniques or using quasi-Monte Carlo approach (see [28, 51]).

In Chapter 3, we investigate Monte Carlo methods for regime-switching models and present a comparative study of Monte Carlo simulations and finite-difference methods under a regime-switching economy. The surprising results show that even with one single underlying stock, Monte Carlo simulations are still more efficient than finite-difference methods due to the hidden Markov chain feature. This indicates that the argument “Monte Carlo methods are much slower than finite-difference methods when the dimension of the problem is low” is not absolutely right.

### 2.3.2 Finite-difference methods and associated schemes

Since the options pricing problem can be viewed as a parabolic PDE problem as presented by [5], finite-difference methods are naturally capable of generating accurate numerical solutions to the BS model. The method was first applied by Eduardo Schwartz in [78] to tackle some cases, such as the one in which the stock pays discrete dividends, with no closed-form solutions available. The pioneer work transitions the option pricing problem to another stage with the wealth of existing theory of numerical PDE analysis.
The principle of finite-difference methods is to replace derivatives in PDEs by linear combinations of function values at the grid points. The theoretical base is Taylor’s expansion of functions near the points of interest. Consider a function \( f(x) \in C^\infty \) defined on the interval \([0, X]\) whose values are approximated on grid points \( x_n = n\Delta x \) with uniform mesh size \( \Delta x = X/N \). We further denote \( f(x_n) \) by \( f_n \) for simplicity. Then according to Taylor’s series expansion,

\[
\begin{align*}
    f_{n+1} &= f_n + \Delta x \left( \frac{\partial f}{\partial x} \right)_n + \frac{\Delta x^2}{2} \left( \frac{\partial^2 f}{\partial x^2} \right)_n + \frac{\Delta x^3}{6} \left( \frac{\partial^3 f}{\partial x^3} \right)_n + \frac{\Delta x^4}{24} \left( \frac{\partial^4 f}{\partial x^4} \right)_n + \cdots, \\
    f_{n-1} &= f_n - \Delta x \left( \frac{\partial f}{\partial x} \right)_n + \frac{\Delta x^2}{2} \left( \frac{\partial^2 f}{\partial x^2} \right)_n - \frac{\Delta x^3}{6} \left( \frac{\partial^3 f}{\partial x^3} \right)_n + \frac{\Delta x^4}{24} \left( \frac{\partial^4 f}{\partial x^4} \right)_n - \cdots.
\end{align*}
\]

(2.3.3) (2.3.4)

It can be easily observed that by rearranging or combining (2.3.3) and (2.3.4), the first-order derivative can be written as,

\[
\begin{align*}
    \left( \frac{\partial f}{\partial x} \right)_n &= \frac{f_{n+1} - f_n}{\Delta x} + O(\Delta x), \quad \text{forward difference,} \\
    \left( \frac{\partial f}{\partial x} \right)_n &= \frac{f_n - f_{n-1}}{\Delta x} + O(\Delta x), \quad \text{backward difference,} \\
    \left( \frac{\partial f}{\partial x} \right)_n &= \frac{f_{n+1} - f_{n-1}}{2\Delta x} + O((\Delta x)^2), \quad \text{central difference.}
\end{align*}
\]

(2.3.5) (2.3.6) (2.3.7)

The names of (2.3.5) to (2.3.7) are given with respect to the location of the point of interest compared to the others appear in the scheme. Note that the central difference is more accurate than either the forward difference or backward difference with a higher rate \( O((\Delta x)^2) \). This implies that, for instance, if we cut down the size of the increment \( \Delta x \) by a half, the error from the central-difference approximation will decrease to a quarter, while the same to forward- or backward-difference approximation only reduces a half.

When applied to the BS equation, forward- and backward-difference approximations for the term \( \partial V/\partial \tau \) lead to explicit and fully implicit schemes, respectively. Central differences of the form (2.3.7), however, are never used in practice because of their bad numerical performance. A similar form of central differences

\[
\left( \frac{\partial f}{\partial x} \right)_{n+1/2} = \frac{1}{2} \left( \frac{\partial f}{\partial x} \right)_{n+1} + \frac{1}{2} \left( \frac{\partial f}{\partial x} \right)_{n}.
\]

(2.3.8)
2.3. NUMERICAL APPROACHES

is known as the Crank-Nicolson finite-difference scheme.

The second-order partial derivative can be yielded in exactly the same way by combining (2.3.5) to (2.3.7),

\[
\left( \frac{\partial^2 f}{\partial x^2} \right)_n = \frac{f_{n+1} - 2f_n + f_{n-1}}{\Delta x^2} + O(\Delta x^2), \quad \text{central difference.}
\]

This is the most popular scheme among all the other approximations as it is symmetric and accurate to \(O(\Delta x^2)\).

Now consider a call option with price \(V(S, t)\), as a function of stock price \(S\) and time \(t\), whose exercise price is \(E\) and maturity \(T\). Then as we discussed in Section 2.2, \(V\) is the solution of equation (2.2.3) with a terminal condition \(V(S, T) = \max(S - E, 0)\). The terminal condition can be reversed to an initial condition by introducing a transformation \(\tau = T - t\).

To solve the PDE numerically, boundary conditions have to be considered. When \(S \to \infty\), the price of the call option tends to be the price of the underlying stock as option prices cannot exceed the underlying price. When \(S \to 0\), the price of the call option reaches 0 as it becomes worthless. Then we have the BS PDE with the initial condition and boundary conditions

\[
\begin{align*}
\frac{\partial V}{\partial \tau} &= \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV, \\
V(S, 0) &= \max(S - E, 0), \\
\lim_{S \to \infty} \frac{V}{S} &= 1, \\
V(0, t) &= 0.
\end{align*}
\] (2.3.9) (2.3.10) (2.3.11) (2.3.12)

To set up the mesh grids, we divide the \(S\)-axis and \(\tau\)-axis into equally spaced nodes \(\Delta x\) and \(\Delta t\) apart, respectively. This makes the original \((\tau - S)\) plane a mesh, with each point having the form \((m\Delta S, n\Delta \tau)\), with variable \(S\) truncated by \(S_{\text{max}} = M\Delta S\). For simplicity, we denote \(V(m\Delta S, n\Delta \tau)\) by \(V^n_m\).

Here we start with the explicit finite-difference approximation, which is to use the forward difference for derivatives with respect to variable \(\tau\) while central differences for
variable \( S \). Then the PDE (2.3.9) can be written as

\[
\frac{V^{m+1}_n - V^m_n}{\Delta \tau} = \mathcal{O}(\Delta \tau) + \frac{1}{2} \sigma^2 S^2_n \frac{V^{m+1}_{n+1} - 2V^m_n + V^m_{n-1}}{\Delta x^2} + \frac{rS_n}{2\Delta x} \left( \frac{V^{m+1}_{n+1} - V^m_{n-1}}{\Delta t} \right) + \mathcal{O}(\Delta x^2) - rV^m_n.
\]

(2.3.13)

It should be remarked that (2.3.13) implies the convergence order of the approximation, \( \mathcal{O}(\Delta \tau + \Delta x^2) \), which is first-order in variable \( \tau \) and second-order in variable \( S \). Ignoring the terms \( \mathcal{O}(\Delta \tau) \) and \( \mathcal{O}(\Delta x^2) \), we can rearrange (2.3.13) and this gives

\[
V^{m+1}_n = (\alpha_n + \beta_n)V^m_{n+1} + (1 - 2\alpha_n - 2\beta_n)V^m_n + (\alpha_n - \beta_n)V^m_{n-1}.
\]

(2.3.14)

where

\[
\alpha_n = \frac{1}{2} \sigma^2 n^2 \Delta \tau, \quad \beta_n = \frac{1}{2} rn \Delta \tau
\]

Note that \( V^{m+1}_n \) depends only on three known values, \( V^m_{n+1}, V^m_n \) and \( V^m_{n-1} \), from the previous time-step thus can be solved explicitly. Hence the option price can be yielded by recursively using (2.3.14) from \( m = 0 \) to \( m = M \) where \( \tau_m = M \Delta \tau \).

Explicit finite-difference schemes are the simplest among all the other finite-difference methods. However, the method has a stability problem as the rounding errors in each iteration may be magnified. Hence a stability condition should be satisfied. In this particular case, it is \( \Delta \tau < 1/(\sigma^2 N^2 + r) \), according to Von Neumann stability analysis. This could make the method rather expensive. For example, if we want to have a more accurate approximation by doubling the number of \( S \)-mesh points, then the time-step has to be roughly quartered to ensure a stable solution. Then the number of loops will become four times than what it was with the original partitions, needless to say, each iteration has the double-sized space grid involved.

On the other hand, if we replace the \( \partial V/\partial \tau \) by the backward difference, similarly, we will have

\[
-(\alpha_n + \beta_n)V^{m+1}_{n+1} + (1 + 2\alpha_n + 2\beta_n)V^m_n - (\alpha_n - \beta_n)V^{m+1}_{n-1} = V^m_n.
\]

(2.3.15)

Here in (2.3.15), \( V^{m+1}_{n+1}, V^m_n \) and \( V^{m+1}_{n-1} \) all depend on the known value \( V^m_n \). In this occasion, new nodal values cannot be explicitly solved in terms of the old, known values.
This is why it is named fully implicit finite-difference method. Therefore we can write (2.3.15) as the linear system

\[
\begin{pmatrix}
B_1 & \Gamma_1 & 0 & \ldots & 0
\vdots & \ddots & \ddots & \ddots & \vdots
0 & \ldots & 0 & \Gamma_{N-1}
0 & \ldots & A_N & B_N
\end{pmatrix}
\begin{pmatrix}
V_1^{m+1}
V_2^{m+1}
\vdots
V_N^{m+1}
\end{pmatrix}
= 
\begin{pmatrix}
0
V_2^m
\vdots
V_N^m
\Gamma_N V_N^{m+1}
\end{pmatrix}
+ 
\begin{pmatrix}
A_1 V_1^{m+1}
A_2 V_2^m
\vdots
\Gamma_N V_N^{m+1}
\end{pmatrix}
\] (2.3.16)

where

\[A_n = -(\alpha_n - \beta_n), \quad B_n = 1 + 2\alpha_n + 2\beta_n, \quad \Gamma_n = -(\alpha_n + \beta_n), \quad n = 1, 2, \ldots, N.\]

The coefficient matrix is tridiagonal and can be simply solved by the LU decomposition.

The advantage of the implicit finite-difference method is its unconditional stability. This frees us from having to take ludicrously small time-steps when we want to increase the number of the \(S\) partition.

In practice, the fully implicit finite-difference method is not preferred because another implicit scheme, the Crank-Nicolson scheme, has a second-order rate of convergence while being unconditionally stable (the fully implicit is of first-order convergence). Based on (2.3.8), the Crank-Nicolson scheme is essentially an average of explicit and fully implicit methods. It could also be obtained by averaging over (2.3.14) and (2.3.15), which gives us

\[-\frac{1}{2}(\alpha_n + \beta_n)V_{n+1}^{m+1} + (1 + \alpha_n + \beta_n)V_n^{m+1} - \frac{1}{2}(\alpha_n - \beta_n)V_{n-1}^{m+1} = \frac{1}{2}(\alpha_n + \beta_n)V_{n+1}^m + (1 - \alpha_n - \beta_n)V_n^m + \frac{1}{2}(\alpha_n - \beta_n)V_{n-1}^m. \] (2.3.17)

Solving this system is, technically, exactly the same as solving the linear system (2.3.16). Even though the right-hand side contains three terms, it could be solved explicitly. Then again, applying the LU decomposition yields the result.

The shortcoming of finite-difference methods lies in the fact that the method is much more complicated for practitioners and suffers from the “curse of dimensionality”. In the case of some stochastic volatility models, for instance, the Heston model, the governing PDE becomes a two-dimensional parabolic equation. In addition to the variable stock
price, variance-associated terms and even a cross-derivative term appear in Heston’s PDE. This makes the traditional implicit finite-difference method ridiculously expensive. A more sophisticated scheme, the Alternating Direction Implicit scheme (ADI), is applied to such problems (see [24]). However, it is widely accepted that finite-difference methods fail to solve the problems with dimensions greater than four. Researchers and practitioners instead turn to use Monte Carlo simulations.

### 2.3.3 Binomial tree and trinomial tree methods

The BSM approach was a milestone of assets pricing and the economic idea they raised, such as the delta hedging and risk-neutrality, caused a huge shock amongst the economists at that time. However, the mathematical background it included based on diffusion models appeared to be too elusive for the economists. This motivated a search for simpler modelling framework which could preserve the economic properties of the BSM model but at the same time more easily accessible.

Then in 1977, after 3 years of the born of the BSM model, John Cox, Stephen Ross and Mark Rubinstein presented a discrete-time model [18] in which a stock only attained two future prices, which was later known as the CRR model or binomial tree model. Such a simple model was proved to capture the replication property and it was also proved to converge to the BSM model as timestep went to 0.

Consider a stock whose price $S$ changes only at the discrete times $\Delta t, 2\Delta t, \cdots$, up to $N\Delta t = T$. The rate of return of the stock over each time period can have only two values: $u - 1$ with probability $p$ and $d - 1$ with probability $1 - p$. That is to say, if the stock price is $S$ at the current time $t$, its value will be either $uS$ or $dS$ at the end of the period, $t + \Delta t$. A graphical illustration is given in Figure 2.1.

![Figure 2.1: Graphical illustration for the binomial process over one period of time](image)

Let $r$ be the interest rate then we require the parameters $u$ and $d$ to satisfy $uS >$
If these inequalities do not hold, as mentioned by Cox, Ross and Rubinstein in [18], arbitrage opportunities will arise. Mathematically, the violation of the equalities will result in the probabilities $p$ or $1 - p$ being negative.

Now consider a stock with price $S^n$ at time $t = t_n = n\Delta t$. The local expectation of the random process is

$$E(S^{n+1}|S^n) = (pu + (1 - p)d)S^n.$$ 

To make the discrete-time random process converge to the continuous counterpart, we need to set the expectation of the binomial process to match the SDE (2.1.1) under the risk-neutral measure, which is to replace the drift parameter $\mu$ by the riskless interest rate $r$. Then according to Proposition (2.1.6),

$$(pu + (1 - p)d)S^n = e^{r\Delta t}S^n. \tag{2.3.18}$$

Similar to the expectation, the variance part should be matched, too. Together with relation (2.3.18), we then have

$$pu^2 + (1 - p)d^2 = e^{2r\sigma^2}. \tag{2.3.19}$$

It should be remarked that the system contains two equations (2.3.18) and (2.3.19) but three unknown variables, $u$, $d$ and $p$. Hence one more equation is required to complete the system so that the unknowns can be solved uniquely. Since (2.3.18) and (2.3.19) determine all the statistical important properties of the continuous-time random walk, the extra equation can then be selected somewhat arbitrarily. One popular extra equation is to let the tree be symmetric along the x-axis, which is

$$u = 1/d. \tag{2.3.20}$$

Putting (2.3.18), (2.3.19) and (2.3.20) together, the three variables can be solved as

$$d = A - \sqrt{A^2 - 1}, \quad u = A + \sqrt{A^2 - 1}, \quad p = \frac{e^{r\Delta t} - d}{u - d}.$$
2.3. NUMERICAL APPROACHES

where

\[ A = \frac{1}{2} \left( e^{-r\Delta t} + e^{(r+\sigma^2)\Delta t} \right). \]

With all the unknown variables settled, a tree can be formed and a backward recursive formula can be set up to approximate option prices. Here we consider that at the current time \( t = 0 \), a stock whose price \( S^0_0 \) is known. According to the binomial process, at the beginning of the next time period, the price of the stock can be either \( S^1_1 = uS \) or \( S^{1}_{-1} = dS \). Setting up the stock values repeatedly, a recombining tree can be constructed with \( n + 1 \) possible values can be taken by the stock at time \( t = t_n = n\Delta t \). Then at the final time step, \( t = T = N\Delta t \), we have \( N + 1 \) different values of the underlying stock.

From the parametrization, we can see that

\[ S^N_m = u^{N+m}d^{N-m}S^0_0 = u^mS^0_0, \quad m = N, N-2, N-4, ..., -N+2, -N. \]

Now consider a European call option and let \( V^n_m \) denote the \( m \)-th possible value of the option price at time \( t = t_n \). Then the terminal value of options can be determined by

\[ V^N_m = \max\{S^N_m - E, 0\}. \]

Starting from the final step, working backward with the recursive formula

\[ V^{n-1}_m = e^{-r\Delta t} (pV^n_{m-1} + (1 - p)V^n_{m+1}) . \]

yield the approximation of a European call option.

It should be mentioned that the parametrization we introduced in this subsection is from [84]. There are many other ways to form the parameterizations of the parametrization, such as [19, 43, 80]. However, they all lead to the same results. Some readers may find out that other researchers set \( u = 1/d = e^{\sigma\sqrt{\Delta t}} \) (see [45]) and still get the correct answer. In fact, these parameterizations are mathematically equivalent with neglecting of any terms that have a higher order than \( O(\Delta t) \).

Then in 1979, Boyle [8] first proposed a extension model named as “the trinomial tree model”. As the name suggests, the model adopts a three-jump process, which allows the
underlying stock to stay on a so-called middle (or stable) branch in addition to up and
down. The method is considered to be more accurate than the binomial tree method with
the same number of time steps. The setting of the trinomial tree model is principally
similar to the binomial trees. Consider a discrete-time stock price process at $\Delta t$, $2\Delta t$, ...
$N\Delta t$. Then instead of allowing each nodal value at the next time step to have only
two possibilities, we let three possible values attainable. Then the one unit of time of the
trinomial process is shown in the figure below. It should be remarked that due to the

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{trinomial_tree.png}
\caption{Graphical illustration for the trinomial process over one period of time}
\end{figure}

introduction of the middle branch, the number of parameters to be determined become
five, $u$, $m$, $d$, $p_u$, $p_d$, instead of three. One of the parametrization is given by Kamrad and
Ritchken [45], in which the three arbitrary equations are given as $m = 1$ and $u = 1/d =
\exp\{\lambda \sigma \sqrt{\Delta t}\}$, where $\lambda$ is a stretch parameter to adjust the risk-neutral probability. Then
by matching the first and second order moments

\begin{align*}
\lambda \sigma \sqrt{\Delta t} (p_u - p_d) &= r \Delta t, \\
\lambda^2 \sigma^2 \Delta t (p_u + p_d) &= \sigma^2 \Delta t.
\end{align*}

we can obtain the probabilities

\begin{align*}
p_u &= \frac{1}{2\lambda^2} + \frac{r \Delta t}{2\lambda \sigma}, \\
p_m &= 1 - \frac{1}{\lambda^2}, \\
p_d &= \frac{1}{2\lambda^2} - \frac{r \Delta t}{2\lambda \sigma}.
\end{align*}

Note that when $\lambda = 1$, $p_m = 0$ and the trinomial tree reduces back to a binomial tree.
Hence $\lambda$ should be strictly greater than 1 to ensure the positivity of $p_m$. Then at each
time step, the recursive formula for the trinomial tree approach can be written as

\[
V_{m-1}^n = e^{-r\Delta t} \left( p_u V_{m-1}^n + p_m V_m^n + p_d V_{m+1}^n \right).
\]

(2.3.21)

The key advantage of trinomial tree model is that the space increment can actually be independent of the time increment. The stretch parameter provides an additional dimension of flexibility, which becomes very important in forming a tree approach for the regime-switching models as we will discuss later in this thesis.

In fact, binomial tree methods are the simplest lattice method to approximate the geometric Brownian motions. One can, technically, build a model with arbitrarily many branches over one time period. The model will become more flexible as more branches are involved. However, there is no clear advantage of having more than three successor values but a growing complexity. So the binomial and trinomial tree methods are the ones research focuses on unless pricing model has some special requirements that have to be tackled by a multinomial tree model.

2.3.4 Historical established connection between the numerical techniques

Now we have introduced all three key numerical methods for options pricing problem. Based on different principles, it seems straightforward that they have no relations with each other at all. However, an interesting fact is that, under the BS economy, trinomial tree methods and explicit finite-difference methods are equivalent. The connection is established by Brennan and Schwarz [11]. In this subsection, we review the historical established relationship of the two methods, which is the preliminary knowledge of the second part of this thesis.

To derive such an equivalence, we first introduce two transforms to the BS equation. Let \( V(S,t) \) be the price of an option. The two transformations are given by

\[
x = \log(S), \quad Y = Ve^{-rt}.
\]

Then the BS PDE (2.2.3) is written as

\[
\frac{\partial Y}{\partial t} + \frac{1}{2} \sigma^2 \frac{\partial^2 Y}{\partial x^2} + r \frac{\partial Y}{\partial x} = 0.
\]

(2.3.22)
Then we replace the partial derivatives by finite-difference terms, as we introduced in Section 2.3.2, and to this end we define

\[ Y(x, t) = Y(m\Delta x, n\Delta t) = Y_m^n, \]

where \( m \) and \( n \) indicates the the discrete increments in the value of the variable \( x \) and \( t \), respectively. With the explicit finite-difference scheme, we have

\[ Y_m^n = \alpha Y_{m+1}^{n+1} + \beta Y_m^{n+1} + \gamma Y_{m-1}^{n+1}, \tag{2.3.23} \]

where

\[
\begin{align*}
\alpha &= \frac{\sigma^2 \Delta t}{2\Delta x^2} + \frac{r \Delta t}{2\Delta x}, \\
\beta &= 1 - \frac{\sigma^2 \Delta t}{\Delta x^2}, \\
\gamma &= \frac{\sigma^2 \Delta t}{2\Delta x^2} - \frac{r \Delta t}{2\Delta x}.
\end{align*}
\]

It should be remarked that at the beginning we have \( Y = e^{-rt}V \). Substituting the reverse relation, \( V = e^{rt}Y \), back in (2.3.23) yields

\[ Y_m^n = e^{-r\Delta t} \left( \alpha Y_{m+1}^{n+1} + \beta Y_m^{n+1} + \gamma Y_{m-1}^{n+1} \right). \tag{2.3.24} \]

Now we can see that (2.3.24) is of the same form of (2.3.21). The only difference is the coefficients between the two recursive formulas. In fact, if we select the space increment \( \Delta x \) to be \( \Delta x = \lambda \sigma \sqrt{\Delta t} \), where \( \lambda \) is a positive constant value, then the coefficients of the finite-difference scheme becomes

\[
\begin{align*}
\alpha &= \frac{\sigma^2 \Delta t}{2\lambda^2 \sigma^2 \Delta t} + \frac{r \Delta t}{2\lambda \sigma \sqrt{\Delta t}}, \\
\beta &= 1 - \frac{\sigma^2 \Delta t}{\lambda^2 \sigma^2 \Delta t}, \\
\gamma &= \frac{\sigma^2 \Delta t}{2\lambda^2 \sigma^2 \Delta t} - \frac{r \Delta t}{2\lambda \sigma \sqrt{\Delta t}}.
\end{align*}
\]

Hence \( \alpha \), \( \beta \) and \( \gamma \) are identical to \( p_u \), \( p_m \) and \( p_d \). This implies that the trinomial tree method can be viewed as a special case of the explicit finite-difference method (when the space increment is selected properly).
2.3. NUMERICAL APPROACHES

Part I: Numerical methods under regime-switching models

In this part, we discuss the three numerical techniques, Monte Carlo simulations, finite
difference methods and tree approaches for options pricing under regime-switching models.

The regime-switching models allow the volatility of the GBM to switch between finite
different values, with each representing a particular economic regime. The “switching”
feature is governed by an additional random process, a Markov chain, whose state space
is a set of finite numbers among which the volatility takes values. On the other hand, it
should be remarked that for each fixed state, the model is principally the BS model with
constant volatility. Hence the complexity of the regime-switching model is relatively lower
than the other stochastic volatility models in which another diffusion process is included.

Due to the PDE system that is compromised of a total of $K$-coupled PDEs ($K$ is the
total number of regimes) of the regime-switching model, the finite-difference method is sure
to be less efficient compared to the BS case, since all of the PDEs in the system has to be
solved simultaneously. On the other hand, the Monte Carlo simulation, which simulates
paths of the underlying stock, does not seem to be influenced by the additional feature.
Thus in Chapter 3, we present a comparative study of Monte Carlo simulations and Crank-
Nicolson finite-difference method for pricing options with one single underlying stock under
the regime-switching economy. In the chapter, we modify the existing algorithm of the
Monte Carlo simulation by Lemieux [51] to free the new algorithm from a pre-estimation
of the random numbers to be generated. The numerical efficiency is tested by comparing
the computational time of the two methods when they have the same level of accuracy.

In terms of tree approaches, it is widely accepted that the trinomial tree method
presented by Yuen and Yang in [87] is the most efficient. Then in 2015, Ma et. al.
[59, 60] investigated the order of convergence of the method. However, the results of the
authors are only valid for options with smooth payoff functions subject to boundedness
conditions, which excludes most of the options contracts, such as European call, put and
binary options. Then in Chapter 4, the problem is addressed and we give a rigorous proof
of the speed of convergence for option contracts with a broad family of piecewise smooth
payoff functions. The main results are further verified by numerical experiments.
Chapter 3

Pricing European options on regime-switching assets: A comparative study of Monte Carlo and Finite Difference Approaches

3.1 Introduction

Since the introduction of the Black-Scholes (BS hereafter) model [5] in 1973, analytical solutions for options pricing becomes a very popular research topic since they are fast to calculate and easy to implement. However, closed-form solutions are only available for a few specific options, which are mostly vanilla European options whose payoff is path-independent. Other options, such as many exotic options as well as American options, have no analytical formula. Thus numerical techniques play a very important role in options pricing theory. A critical criterion for a good numerical method is that the algorithm should be computationally cheap and converge to the analytical solution fast. Judging whether a new numerical algorithm is efficient or not can be done by comparing it with another existing algorithm. There has been much research on comparing various numerical methods in the BS model for both European and American options, for instance, [10], [20], etc.

It is well-known that the conventional BS model with constant volatility assumption
fails to reflect the stochastic nature of financial markets. As a result, more realistic models which better reflect random market movements are required. One of these extended models is the *regime-switching model*, in which the volatility is driven by a hidden Markov chain and switches between a finite number of states. Since its introduction by Hamilton [35], growing evidence has suggested that the regime-switching model can capture the time series properties of several important financial variables. In addition, analytical solutions have also been discovered for the model with two states, which makes the model very popular. Guo [32] first presents a closed-form solution with double integrals for pricing European options whose underlying assets follow a two-state regime-switching economy. However, Fuh et al. [26] claimed that there is an error in Guo’s formula and they present a very similar, corrected version. Then Zhu et al. [91] propose another analytical solution for the same two-state regime-switching model. Their formula contains only one single integral since they manage to solve the inverse Fourier transformation analytically. Unfortunately, all the closed-form solutions for regime-switching models are only available for the specific two-state case. Formulas for the models with more than two regimes have yet to find. This means that only numerical approximations can be applied in these situations.

Research on the numerical methods for a general multi-state regime-switching models has been discussed and can be found in [6, 40, 51, 65, 87]. Here this chapter concentrates on two basic and important methods, the Monte Carlo simulation and the finite difference method, the performance of which has not yet been compared in the literature when they are applied to the regime-switching models. A comparative study of the two methods is useful and significant for helping future researchers choose the more efficient one for pricing European style options in a multi-state regime-switching model. On one hand, in terms of the Monte Carlo simulation, we present a modified version of the algorithm from Lemieux [51], which is referred to as the fundamental Monte Carlo simulation in the subsequent sections, and investigate two variance reduction techniques, antithetic variates and control variates. In addition, we also propose a new and much faster Monte Carlo algorithm, which simulates the total occupation time instead of the trajectories for the two-state case. On the other hand, the Crank-Nicolson scheme is adopted for the finite difference method because of its unconditional stability and second-order convergence rate. The numerical comparisons and analysis are given with regime-switching models with up to 4 regimes being taken into consideration.
The rest of the chapter is organized as follows. In the next section, we introduce the model settings and notations for the simplicity of discussion. In Section 3.3, we review the fundamental Monte Carlo simulation presented by Lemieux [51] and give a modified version which frees the former algorithm from pre-estimating the total amount of random numbers that require being generated. The new “simulating total occupation time” algorithm together with two variance reduction techniques can also be found in the section. This is followed by a brief outline of how to apply the finite difference with Crank-Nicolson schemes to the multi-state regime-switching PDE system in Section 3.4. The numerical examples, comparisons and analysis are shown in Section 3.5. The chapter ends with some concluding remarks in Section 3.6.

3.2 Model settings and notations

We first start with the introduction of regime-switching models. Let $S_t$ be the price of an underlying asset in the market at time $t$. We further let the market have a non-constant drift rate $\mu_t$ as well as volatility $\sigma_t$. Then in a regime-switching world where drift rates and volatilities are allowed to shift between different economic regimes, the fluctuation of an asset is assumed to follow the stochastic differential equation

$$
    dS_t = \mu(X_t)S_t dt + \sigma(X_t)S_t dW_t,
$$

where $X_t$ is a continuous-time Markov chain with a total of $K$ states and is independent of the standard Brownian motion $W_t$. Both the Markov chain and standard Brownian motion are based on the probability triplet $\{\Omega, \mathcal{F}, \mathbb{P}\}$ where $\Omega$ is the set of all possible outcomes, $\mathcal{F}$ is the set of events and $\mathbb{P}$ is the physical measure. For each state, the drift rate and the volatility are assumed to be constant and distinct, as denoted by

$$
    \mu(X_t) = \begin{cases}
    \mu_1, & X_t = 1, \\
    \mu_2, & X_t = 2, \\
    \cdots \\
    \mu_K, & X_t = K,
\end{cases}, \quad \sigma(X_t) = \begin{cases}
    \sigma_1, & X_t = 1, \\
    \sigma_2, & X_t = 2, \\
    \cdots \\
    \sigma_K, & X_t = K.
\end{cases}
$$
The generator of the Markov chain is

\[
Q = \begin{pmatrix}
\lambda_{11} & \lambda_{12} & \ldots & \lambda_{1K} \\
\lambda_{21} & \lambda_{22} & \ldots & \lambda_{2K} \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_{K1} & \lambda_{K2} & \ldots & \lambda_{KK}
\end{pmatrix}.
\]

From the Markov chain theory (see for example, Yin and Zhang [85]), the elements of the generator satisfy the following relation for each state

\[
\lambda_{jj} + \sum_{i=1}^{K} \lambda_{ji} = 0.
\]

Since another risk source \(X_t\) is introduced, the market becomes incomplete. As a result, there is no unique martingale measure. In this chapter, we select the martingale measure presented by Elliott et al. [21] and assume the interest rates are the same for all regimes. From now on we change the physical measure to the risk-neutral measure, which leads each \(\mu_j\) to be replaced by \(r\) for the rest of the chapter. Consider an option with expiration time \(T\). Under the regime-switching model (3.2.1), let \(V_j(S_t, t)\) be the price of the option at time \(t \geq 0\) when \(S_t = S\) and \(X_t = j\). Then the price of the option is given by

\[
V_j = e^{-r(T-t)}E(p(S_T)|\mathcal{F}_t, X_t = j),
\]

where function \(p(\cdot)\) is the payoff function, for example, \(\max(E - S_T, 0)\) is the payoff of a European put option. By applying Itô’s formula, \(V_j\) is associated with the following governing system of PDE (see [15] for detailed derivation)

\[
\frac{\partial V_j}{\partial t} + \frac{1}{2} \sigma_j^2 S^2 \frac{\partial^2 V_j}{\partial S^2} + rS \frac{\partial V_j}{\partial S} - rV_j = \sum_{i=1}^{K} \lambda_{ji}(V_j - V_i), \quad j = 1, 2, \ldots, K. \tag{3.2.2}
\]

It should be remarked that in contrast to the conventional BS model, the regime-switching models contain a total of \(K\) equations that have to be solved simultaneously. In the subsequent section, we will introduce how to apply numerical techniques to solve the PDE system (3.2.2)
3.3 Monte Carlo simulation

The MC simulations for regime-switching models have been discussed in the literature and can be found in \cite{40, 51}. Lemieux \cite{51} presents an algorithm to price vanilla European options in a general multi-state regime-switching model. The author applies the MC method to simulate the trajectory of the regime-switching economy. However, her algorithm requires a pre-estimation of the total amount of random numbers. Since the number of exponential random numbers in each path is uncertain, such an estimation will lead to redundant random numbers, wasting some computational resources. Hieber and Scherer \cite{40} present an efficient simulation method which is coupled with two variance reduction techniques for pricing barrier options under a two-state regime-switching model. The idea of their Brownian bridge approach can be viewed as a special case of the simulated trajectories in Lemieux \cite{51}. In this section, we present a modified version of Lemieux’s algorithm so that it successfully get rid of the pre-estimation. Our algorithm is denoted by “the fundamental MC” throughout the chapter. We also discuss two variance reduction techniques, the antithetic variates and the control variates to further improve the efficiency of the MC simulations. In addition, we propose a new algorithm to simulate the total occupation time instead of the trajectory for the two-state case. The idea is based on the path-independent property of the vanilla European options and we name it “simulating total occupation time”.

3.3.1 Fundamental Monte Carlo

The theoretical framework for simulating trajectories is based on the fact that the holding time is exponentially distributed. Further, the probability of state $j$ switching to state $i$ can be estimated by $-\lambda_{ji}/\lambda_{jj}$, given the hypothesis that a switch has taken place. Thus by keeping generating exponential random numbers with until the summation of the generated random number exceeds the life of the option, we obtain one replication with switching time determined. Then between each two switch, the conventional BS MC simulation can be applied. Repeat doing the replication many times and the option price can be obtained by averaging the results from all replications. Our modified version of the Fundamental MC method for a general multi-state regime-switching model is efficient and an example of applying our algorithm to price a European put option is given in
Algorithm 1 in Appendix A. European call options can be priced in a similar manner by replacing the put payoff function by the call payoff function.

One disadvantage of the MC simulation is that it is not efficient enough especially for financial derivatives with one single underlying asset. Thus variance reduction techniques can be adopted to improve the efficiency of MC simulations. Based on our algorithms, there are two variance reduction techniques that can be naturally applied. In the next two subsections, we will introduce them and the improvement they bring to the MC simulation is tested and presented in Section 3.5.

3.3.2 Antithetic variates

The theoretical base of an MC simulation is the Law of Large Numbers. By generating random numbers that follow the specific distribution, the result obtained from averaging all \( N \) replications is actually a random number whose mean is the true value we are looking for and variance is \( (1/\sqrt{N}) \). Thus a natural way of improving the efficiency is to reduce the variance.

To be more specific, suppose that the option price we are trying to estimate is

\[
\theta = \mathbb{E}(h(X)) = \mathbb{E}(Y).
\]

For this we have generated two samples, \( Y_1 \) and \( Y_2 \). Hence an unbiased estimate of the option price \( \theta \) is obtained by

\[
\hat{\theta} = \frac{Y_1 + Y_2}{2}.
\]

Since

\[
\text{Var}(\hat{\theta}) = \frac{\text{Var}(Y_1) + \text{Var}(Y_2) + 2\text{Cov}(Y_1, Y_2)}{4},
\]

the variance of the estimation is reduced if \( \text{Cov}(Y_1, Y_2) \) is negative. A very natural idea of constructing a pair of negatively correlated random variables is to use the symmetry of some distributions. For instance, the uniform distribution \( (A \sim U(0,1)) \) and standard normal distribution \( (B \sim N(0,1)) \) are two common ones that are frequently involved in Monte Carlo simulations. The antithetic pairs of the two distributions are \( 1-A \) and \( -B \),
respectively.

Thus, the main idea of the antithetic variates [28] is that, instead of estimating an expectation $\mu$ by averaging over $N$ independent, identical, distributed (i.i.d.) random variables, averaging over $N/2$ pairs of i.i.d. random variables, which are denoted as $Z_{ori}$, and their negatives, which are denoted as $Z_{ant}$. So the fundamental MC with the antithetic variates replaces Step 6 in Algorithm 1 with the pair of $W$ and $-W$ since the standard normal distribution is symmetric around the y-axis. Thus Line 37 to 40 is replaced by the following steps as shown in Algorithm 2.

The efficiency of the antithetic variates strongly depends on the negative correlation between $\max(E - Z_{ori}, 0)$ and $\max(E - Z_{ant}, 0)$. More negative correlation results in a better variance reduction.

### 3.3.3 Control variates

Another variance reduction technique, the control variates [28], shares a common feature with the antithetic variates, which is to use correlations to reduce variances. To apply the control variates technique, we need to find a variable $C$ whose mean is known and related to the simulated model and $C$ should also be correlated with the variable $Y$. In the option pricing problem, the price of the underlying asset (mostly stocks) actually provides a universal set of control variates. This is because under the risk-neutral measure, the stock price at time $t$ is $e^{rt}$ times the stock price today. To describe it in detail, consider a stock with price $S_t$. Thus $e^{-rt}S_t$ is a martingale under the risk-neutral measure. Further let $Y(S_t, t)$ be the discounted payoff of a financial derivative. Then the estimator of control variates is given by

$$\frac{1}{N} \sum_{n=1}^{N} (Y_n - b[S_n(T) - e^{rT}S(0)]),$$

where $S_n$ is the $n$th replication of the simulation and $b$ is the coefficient which minimizes the variance of the estimator. To adopt the control variates in the fundamental MC, several steps must be added in between Line 39 and 40 of Algorithm 1. These steps are shown in Algorithm 3.

The variance of the control variates estimator is $1 - \rho_Y^2$. Similar to the antithetic variates, a stronger correlation between $Y$ and $Z$ means the higher efficiency of the simulation with control variates technique.
3.3.4 Simulating total occupation time

Due to the Markovian property of the regime-switching model, the iterations of simulating trajectories can be avoided. This means that the generation of the exponential random numbers can be avoided. It can be accomplished by simulating the total occupation time within each regime. Theoretically, this can be achieved because the total occupation time is a random variable with its own probability density function (pdf). We find out that this is directly applicable to the two-state case since the analytical formula of the pdf of the total occupation time is available [26] for the two-state case. The pdf is shown in the following theorem.

**Theorem 3.3.1.** Assume $T_{ij}$ is the total occupation time that the Markov chain $X(t)$ spends in state $i$, given that the initial state is $j$, where $i, j \in \{1, 2\}$. Let $f_{ij}$ be the probability density function (pdf) of $T_{ij}$. Under the two-state Markov model, we have the following formula:

$$
\begin{align*}
    f_{1|1}(t, T) &= e^{-\lambda_{12} T} \delta(T - t) + e^{-\lambda_{21} (T - t)} - \lambda_{12} t^{1/2} \left( I_0(2\lambda_{12} \lambda_{21} t (T - t))^{1/2} \right) \\
    f_{2|2}(t, T) &= e^{-\lambda_{21} T} \delta(T - t) + e^{-\lambda_{12} (T - t)} - \lambda_{21} t^{1/2} \left( I_0(2\lambda_{12} \lambda_{21} t (T - t))^{1/2} \right) \\
    f_{2|1}(t, T) &= f_{1|1}(T - t, T), \quad f_{1|2}(t, T) = f_{2|2}(T - t, T),
\end{align*}
$$

where $\delta$ is a Dirac delta function and $I_0, I_1$ are modified Bessel functions such that

$$
I_\alpha(z) = \left( \frac{z}{2} \right)^\alpha \sum_{k=0}^{\infty} \frac{(z/2)^{2k}}{k! \Gamma(k + \alpha + 1)!}.
$$

Now we have got the pdf of the total occupation time and further, we can solve for the cumulative probability distribution function (cdf) by a simple numerical integration. Note that the calculation is quite straightforward everywhere apart from at $t = T$, which is a singular point and also contains a Dirac delta function. This can be interpreted as the total occupation time reaching a point mass at $t = T$, in other words, the Markov chain has remained in the same state throughout $[0, T]$. The probability of this occurring is given by $P(T_{ij} = T) = e^{-\lambda_{ij} T}$ where $j = 3 - i$, which corresponds to a discontinuity in
the cdf at $t = T$. In terms of the numerical integration, we first proceed as usual without the Dirac delta function, then add a jump of size $e^{-\lambda_{ij} T}$ to the cdf at $t = T$. A detailed algorithm of how to generate random numbers from the above pdfs is given in Algorithm 4.

Once having the simulation of total occupation time, results for each path can be obtained by simple substitution, as we have shown in Algorithm 5. It should be remarked that since this is in the two-state case, the generator matrix contains only four elements $\lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22}$. To simplify the notations, we introduce a $3 - j$ as the state other than state $j$.

The core of Algorithm 5 is to generate the random variable that follows the pdf of the total occupation time. This algorithm is much cheaper in computation than the fundamental MC because only two random numbers are generated in this algorithm, in contrast to a total of $(n + 1)$ random numbers must be generated in the fundamental MC, with $n$ being the number of state changes. This can also explain why the computational time of the fundamental MC is a monotonically increasing function of the switching intensity and the time to expiry while the computational time of this algorithm remains a constant and only depends on the number of paths.

In summary, Algorithm 5 contains two parts: one is to derive the pdf of the total occupation time and the other is to draw samples from this random variable. In theory, this algorithm can be extended to a general case with multi-state economy. Although the analytical formula of the pdf is not available in the multi-state case, a Fourier cosine expansion [23] can be applied to obtain a numerical distribution function. Drawing samples from this multivariate random number can be done by building a very complex algorithm based on the acceptance-rejection method [28]. However, the multi-state extension is no longer fast and simple to implement, which was a crucial advantage of the total occupation time approach. Thus this case will not be further explored in this chapter.

## 3.4 Finite difference method

Mielkie [64, 65] presents a detailed discussion about how to apply CN finite difference method to solve the coupled partial differential equations arising from a two-state regime-switching model. In contrast to the BS model, the governing PDE problem of a regime-
switching model contains a total of $K$ coupled PDEs, each corresponding to one of the $K$ different regimes. In this section, we extend Mielkie’s method to solving a multi-state regime-switching model. We also discuss the computational time that requires by using the finite difference method as computing time is one key factor of efficiency.

We first introduce the transformation $\tau = T - t$ and $x = \log(S/E)$. Then with the transformation, (3.2.2) becomes

$$\frac{\partial V_j}{\partial \tau} = \frac{1}{2} \sigma_j^2 \frac{\partial^2 V_j}{\partial x^2} + (r - \frac{1}{2} \sigma_j^2) \frac{\partial V_j}{\partial x} - r V_j - \sum_{i\neq j}^{K} \lambda_{ji} (V_j - V_i), \quad j = 1, 2, \ldots, K. \quad (3.4.1)$$

Here we discretize the transformed stock price region $[x_{\text{min}}, x_{\text{max}}] \times [0, T]$, into $(M + 1) \times (N + 1)$ grids, with $\Delta x = (x_{\text{max}} - x_{\text{min}})/M$ and give time increment by $\Delta \tau = T/N$. We further denote the price of the derivative at $(S_m, t_n)$ in regime $j$ by $V_{m,j}^n := V_j(x_{\text{min}} + m\Delta x, \Delta \tau)$. By applying the CN scheme, we have

$$\alpha_j V_{m-1,j}^{n+1} + \beta_j V_{m,j}^{n+1} + \gamma_j V_{m+1,j}^{n+1} + \sum_{i\neq j}^{K} \frac{\lambda_{ji} \Delta \tau}{2} V_{i,m}^{n+1} = f_{m,j}^n, \quad (3.4.2)$$

where

$$\alpha_j = - \left( \frac{\sigma_j^2 \Delta \tau}{4 \Delta x^2} - \frac{(r - \frac{1}{2} \sigma_j^2) \Delta \tau}{4 \Delta x} \right),$$

$$\beta_j = 1 + \frac{\sigma_j^2 \Delta \tau}{2 \Delta x^2} + \frac{r \Delta \tau}{2} - \frac{\lambda_{jj} \Delta \tau}{2},$$

$$\gamma_j = - \left( \frac{\sigma_j^2 \Delta \tau}{4 \Delta x^2} + \frac{(r - \frac{1}{2} \sigma_j^2) \Delta \tau}{4 \Delta x} \right),$$

$$f_{m,j}^n = -\alpha_j V_{m-1,j}^n + (2 - \beta_j) V_{m,j}^n - \gamma_j V_{m+1,j}^n - \sum_{i\neq j}^{K} \frac{\lambda_{ji} \Delta \tau}{2} V_{i,m}^n.$$ 

Writing (3.4.2) in matrix form, we obtain

$$P_j V_j^{n+1} + \sum_{i\neq j}^{K} A_{ji} V_i^{n+1} = f_j^n. \quad (3.4.3)$$
where

\[ P_j = \begin{pmatrix} \beta_j & \gamma_j \\ \alpha_j & \beta_j & \gamma_j \\ \ddots & \ddots & \ddots \\ \alpha_j & \beta_j & \gamma_j \\ \end{pmatrix} \quad \Lambda_{ji} = \begin{pmatrix} \frac{\lambda_{ji} \Delta \tau}{2} \\ \ddots \\ \frac{\lambda_{ji} \Delta \tau}{2} \end{pmatrix} \]

\[ V^n_j = \begin{pmatrix} V^n_{1,j} \\ V^n_{2,j} \\ \vdots \\ V^n_{N-2,j} \\ V^n_{N-1,j} \end{pmatrix}, \quad V^n_i = \begin{pmatrix} V^n_{1,i} \\ V^n_{2,i} \\ \vdots \\ V^n_{N-2,i} \\ V^n_{N-1,i} \end{pmatrix}, \quad f^n_j = \begin{pmatrix} f^n_{1,j} \\ f^n_{2,j} \\ \vdots \\ f^n_{N-2,j} \\ f^n_{N-1,j} - \gamma_j V^n_{N,j} \end{pmatrix}. \]

Then the whole system at time \( \tau = n \Delta \tau \) now can be written as

\[
\begin{align*}
P_1 V^n_1 + \Lambda_{12} V^n_2 + \Lambda_{13} V^n_3 + \ldots + \Lambda_{1K} V^n_K &= f^n_1, \\
P_2 V^n_2 + \Lambda_{21} V^n_1 + \Lambda_{23} V^n_3 + \ldots + \Lambda_{2K} V^n_K &= f^n_2, \\
&\vdots \\
P_K V^n_K + \Lambda_{K1} V^n_1 + \Lambda_{K2} V^n_3 + \ldots + \Lambda_{KK} V^n_{K-1} &= f^n_K.
\end{align*}
\]  

(3.4.4)

Then it becomes the system (3.4.4) needs to be solved at each time step. To have a better look at the system, we write (3.4.4) into the matrix form. Assume

\[ A = \begin{pmatrix} P_1 & \Lambda_{12} & \ldots & \Lambda_{1K} \\ \Lambda_{21} & P_2 & \ldots & \Lambda_{2K} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_{K1} & \Lambda_{K2} & \ldots & P_K \end{pmatrix}, \quad V^n = \begin{pmatrix} V^n_1 \\ V^n_2 \\ \vdots \\ V^n_K \end{pmatrix}, \quad F^n = \begin{pmatrix} f^n_1 \\ f^n_2 \\ \vdots \\ f^n_K \end{pmatrix}. \]

Then the problem is further simplified into solving \( A V^n = F^n \) at each time step and the option price can be obtained by an interpolation at the end. It should be remarked that the block matrix \( A \) is sparse since \( P_j \) is a diagonal matrix and \( \Lambda_{ij} \) is an identity matrix. Thus iterative methods are recommended to solve the linear system. In addition, it is obvious that the size of the matrix \( A \) increases linearly as the number of regimes grows. Hence the whole method grows rapidly according to the number of regimes, given the fact
that the computational complexity of matrix inversion is at maximum $O((MNK)^3)$ in this problem, even though it grows linearly. As a result, the CN tends to be far more expensive than the MC especially in the cases where $K > 2$. We present the comparison in the next section.

3.5 Numerical performances and comparisons

In this section, we present numerical examples as well as comparing the approaches we introduce in this chapter. We compare the MC simulations with CN finite difference methods under regime-switching models. We also test the fundamental MC against the two variance reduction techniques and the simulating total occupation time. The models we consider in this section have up to four regimes. The analytical solution from Zhu et al.\cite{91} is adopted as the benchmark results for the two-state regime-switching model while the trinomial tree method from Yuen and Yang \cite{87} is applied for the multi-state cases.

3.5.1 MC vs CN

We start with the comparison between the fundamental MC and CN. Since results obtained from the simulations are random numbers instead of certain values, we consider 95% confidence intervals from running simulations many times. The confidence intervals thus are advised as the accuracies of simulation methods.

Our comparisons focus on regime-switching models with two, three and four states, to show the performance of each method along with different regimes. Two scenarios are considered for each model, one is low-frequency and the other is high-frequency, which aim at comparing the two methods with different switching intensities, given the fact that the computational complexity of the fundamental MC depends on the parameter. To make the comparison fair, we adjust the number of paths for the MC simulation as well as the number of time and space steps for the finite difference so that the errors of both methods are on the same level. Then by comparing their computational time, the one with less computing time is the more efficient numerical method.

In Table 3.1 and Table 3.2, results of MC simulations are shown with an average (AVE) over multi-runs and the corresponding 95% confidence interval (CON) are given in the form of $AVE \pm CON$. The values in the parentheses in the column of CN are the
differences between results from the finite difference method and the benchmark. To obtain the benchmark results for the two-state regime-switching model, we apply the closed-form solution from Zhu et. al. [91] while for the three-state and four state models, we use the trinomial tree method from Yuen and Yang [87] with 100,000 time steps. The number of time step and space step for the finite difference method is selected as 100 and 2500, respectively. Correspondingly, MC simulations are run with 500,000 paths.

Table 3.1: MC vs CN in the low-frequency case. The parameters are \( S_0 = 36, E = 40, T = 1, \gamma = 0.1 \). The volatilities are \( (0.15, 0.25)', (0.15, 0.25, 0.35)', (0.15, 0.25, 0.35, 0.45) \) for the two-regime, three-regime, four-regimes, respectively. The jump intensities for all different models are \( \lambda_{ji} = 1, j \neq i \).

<table>
<thead>
<tr>
<th></th>
<th>MC</th>
<th>CN</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( V_1^2 )</td>
<td>( V_2^2 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2.7022 ± 0.0015</td>
<td>3.3203 ± 0.0017</td>
<td></td>
</tr>
<tr>
<td>Time(sec)</td>
<td>12.4822</td>
<td>54.6678</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( V_1^3 )</td>
<td>( V_2^3 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.3562 ± 0.0016</td>
<td>3.7653 ± 0.0018</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4.2508 ± 0.0020</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time(sec)</td>
<td>19.0468</td>
<td>426.5641</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: MC vs CN in the high-frequency case. All parameters are identical to the low-frequency case but the jump intensities now are \( \lambda_{ji} = 100, j \neq i \).

<table>
<thead>
<tr>
<th></th>
<th>MC</th>
<th>CN</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( V_1^2 )</td>
<td>( V_2^2 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.0577 ± 0.0016</td>
<td>3.0646 ± 0.0017</td>
<td></td>
</tr>
<tr>
<td>Time(sec)</td>
<td>12.4371</td>
<td>55.6734</td>
<td></td>
</tr>
<tr>
<td></td>
<td>( V_1^3 )</td>
<td>( V_2^3 )</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.8696 ± 0.0024</td>
<td>3.8733 ± 0.0024</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.8787 ± 0.0024</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Time(sec)</td>
<td>31.6948</td>
<td>431.6442</td>
<td></td>
</tr>
</tbody>
</table>

According to the two tables, each confidence interval of the MC is smaller than the error of the CN (only except \( V_1^2 \) in Table 3.1 but the two values are rather close) while the CN spends several more times in computational time than the MC. Hence it is obvious
that the MC outperforms the CN in both cases. The difference of the computation time becomes more significant as the number of the states increases. It should be noticed that the computational time of the MC in the high-frequency case is greater than the counterpart in the low-frequency case. This can be explained as follows: more iterations are involved in the simulating trajectories as the average switching times are changed from once per year to 100 times per year. In contrast to the MC simulation, computational time of the CN appears to be independent of the jump intensity. This is because the computation cost, which is mainly solving linear algebraic systems, remains the same no matter how \( \lambda \) changes. Although MC is still much cheaper in the high-frequency case we propose in this chapter, it is worth mentioning that, if we keep increasing \( \lambda \), the CN will eventually outperform MC. However, the magnitude of such a \( \lambda \) is too large to be realistic, for instance, \( \lambda = 300 \). Hence the MC simulations are still recommended here as the more efficient numerical technique for pricing European options under a multi-state regime-switching model.

### 3.5.2 Comparison among simulations

Since the simulation is cheaper than the finite difference method, we further explore the performance of the simulations coupled with variance reduction methods. For the two-state case, we also include the “simulating total occupation time” in the comparison. To better compare among the simulations, we introduce an indicator to quantify the efficiency of MC from Lemieux [51].

**Definition 3.5.1.** *The efficiency of an estimator \( \hat{\mu} \) for a quantity \( \mu \) is measured by the indicator*

\[
\text{Eff}(\hat{\mu}) = \left[ \text{MSE}(\hat{\mu}) \times C(\hat{\mu}) \right]^{-1}
\]

*where MSE(\( \hat{\mu} \)) = Var(\( \hat{\mu} \)) + B^2(\( \hat{\mu} \)) is the mean square error of \( \hat{\mu} \), B(\( \hat{\mu} \)) = E(\( \hat{\mu} \)) - \mu is the bias of \( \hat{\mu} \), and C(\( \hat{\mu} \)) is the expected computation time for \( \hat{\mu} \).*

According to the formula, the efficiency is inversely proportional to both the mean square error and the computational time. A method with a greater \( \text{Eff} \) can then be viewed as a more efficient estimator. The comparison is constructed with the same sets parameters as we used in Section 3.5.1 (the low-frequency case and the high-frequency case). We take the two-state model out from the other comparisons since the “simulating
total occupation time” is also applicable to a two-state regime-switching model.

Table 3.3: The low-frequency two-state case

<table>
<thead>
<tr>
<th></th>
<th>Fundamental MC</th>
<th>Antithetic variates</th>
<th>Control variates</th>
<th>Occupation time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_1^2$</td>
<td>2.7023 ± 0.0022</td>
<td>2.7019 ± 0.0013</td>
<td>2.7019 ± 0.0014</td>
<td>2.7032 ± 0.0011</td>
</tr>
<tr>
<td>$V_2^2$</td>
<td>3.3196 ± 0.0028</td>
<td>3.3196 ± 0.0016</td>
<td>3.3199 ± 0.0016</td>
<td>3.3197 ± 0.0010</td>
</tr>
<tr>
<td>Time(sec)</td>
<td>5.3670</td>
<td>2.9645</td>
<td>4.9100</td>
<td>0.0365</td>
</tr>
<tr>
<td>MSE</td>
<td>4.2666e−05</td>
<td>1.6323e−04</td>
<td>1.4110e−05</td>
<td>7.9885e−06</td>
</tr>
<tr>
<td>Efficiency</td>
<td>4.3670e+03</td>
<td>2.5909e+04</td>
<td>1.4434e+04</td>
<td>3.4330e+06</td>
</tr>
</tbody>
</table>

Table 3.4: The high-frequency two-state case

<table>
<thead>
<tr>
<th></th>
<th>Fundamental MC</th>
<th>Antithetic variates</th>
<th>Control variates</th>
<th>Occupation time</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_1^2$</td>
<td>3.0569 ± 0.0021</td>
<td>3.0573 ± 0.0015</td>
<td>3.0571 ± 0.0015</td>
<td>3.0574 ± 0.0012</td>
</tr>
<tr>
<td>$V_2^2$</td>
<td>3.0637 ± 0.0021</td>
<td>3.0642 ± 0.0015</td>
<td>3.0641 ± 0.0015</td>
<td>3.0636 ± 0.0011</td>
</tr>
<tr>
<td>Time(sec)</td>
<td>9.4756</td>
<td>5.0528</td>
<td>9.0112</td>
<td>0.0339</td>
</tr>
<tr>
<td>MSE</td>
<td>2.7902e−05</td>
<td>6.0305e−05</td>
<td>1.4990e−05</td>
<td>8.1818e−06</td>
</tr>
<tr>
<td>Efficiency</td>
<td>3.7823e+03</td>
<td>1.3128e+04</td>
<td>7.4030e+03</td>
<td>3.6071e+06</td>
</tr>
</tbody>
</table>

According to Table 3.3 and Table 3.4, each of the three techniques improves the simulation method in different degrees. Simulating total occupation time algorithm performs the best with a much higher efficiency. This is because the algorithm is much cheaper than others even though its mean square error (MSE) is similar to that of the others.

In addition, the computational time of the algorithm appears to be independent of the switching intensity $\lambda$ as the computational time in the low-frequency case is 0.0365 seconds against 0.0339 seconds in the high-frequency. For the two variance reduction techniques, the efficiencies of both methods decline from the low-frequency case to the high-frequency case. This results from more computations being involved as $\lambda$ increases when the MSEs remain the same for both cases. The antithetic variates estimator outperforms the control variates estimator since it only spends roughly half of the computational time of the other two methods, the fundamental MC and antithetic variates.

The rest of the multi-state models are put together, again with the same low-frequency case and high-frequency case. Performances of the two variance reduction methods are shown in Table 3.5 and Table 3.6.

According to the two tables, the conclusion of the two-state model also holds. Antithetic variates technique is the better choice as a variance reduction method with the higher efficiency in both situations no matter how many regimes are considered. In addition, it turns out that the MSEs of all of the three methods is independent of the
3.6 Conclusion

A comparative study of the MC and the CN for pricing European options with a general regime-switching model is presented in this chapter. Since the number of the PDEs within the governing system grows according to the number of regimes, solving the governing system becomes very inefficient as the number of regimes grows very large. Numerical results show that even for the two-state model, the fundamental MC is already more

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fundamental MC</th>
<th>Antithetic variates</th>
<th>Control variates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_1^3$</td>
<td>3.3555 ± 0.0022</td>
<td>3.3577 ± 0.0018</td>
<td>3.3568 ± 0.0016</td>
</tr>
<tr>
<td>$V_2^3$</td>
<td>3.7643 ± 0.0022</td>
<td>3.7664 ± 0.0019</td>
<td>3.7652 ± 0.0018</td>
</tr>
<tr>
<td>$V_3^3$</td>
<td>4.2496 ± 0.0022</td>
<td>4.2528 ± 0.0021</td>
<td>4.2518 ± 0.0019</td>
</tr>
<tr>
<td>Time(s)</td>
<td>8.1438</td>
<td>4.5162</td>
<td>7.4225</td>
</tr>
<tr>
<td>MSE</td>
<td>3.8896e-05</td>
<td>1.4884e-04</td>
<td>1.3168e-05</td>
</tr>
<tr>
<td>Efficiency</td>
<td>3.1570e+03</td>
<td>1.3373e+04</td>
<td>1.0232e+04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Fundamental MC</th>
<th>Antithetic variates</th>
<th>Control variates</th>
</tr>
</thead>
<tbody>
<tr>
<td>$V_1^4$</td>
<td>4.1038 ± 0.0029</td>
<td>4.1035 ± 0.0021</td>
<td>4.1033 ± 0.0021</td>
</tr>
<tr>
<td>$V_2^4$</td>
<td>4.3808 ± 0.0031</td>
<td>4.3791 ± 0.0020</td>
<td>4.3800 ± 0.0021</td>
</tr>
<tr>
<td>$V_3^4$</td>
<td>4.7287 ± 0.0032</td>
<td>4.7275 ± 0.0020</td>
<td>4.7270 ± 0.0023</td>
</tr>
<tr>
<td>$V_4^4$</td>
<td>5.1273 ± 0.0033</td>
<td>5.1262 ± 0.0023</td>
<td>5.1253 ± 0.0025</td>
</tr>
<tr>
<td>Time(s)</td>
<td>11.4680</td>
<td>6.7384</td>
<td>10.2069</td>
</tr>
<tr>
<td>MSE</td>
<td>3.9272e-05</td>
<td>2.2926e-04</td>
<td>1.6073e-05</td>
</tr>
<tr>
<td>Efficiency</td>
<td>1.4712e+03</td>
<td>1.0500e+04</td>
<td>6.0955e+03</td>
</tr>
</tbody>
</table>

parameter $\lambda$ and the number of the regimes. Hence we recommend to use MC simulation with the antithetic variate technique is the best choice for pricing European options under a multi-state regime-switching model.

Table 3.5: The low-frequency multi-state case

Table 3.6: The high-frequency multi-state case
efficient than the CN in both a low-frequency case a high-frequency case we consider in this chapter. The difference in efficiency becomes more severe as the number of regimes further increases. Such a finding suggests that future research on numerical techniques for regime-switching models should concentrate more on simulation-based methods.

We have also investigated two variance reduction techniques, the antithetic variates and the control variates, to improve the efficiency of the simulation methods. Numerical performance shows that the antithetic variates technique is more efficient than the control variates for regime-switching models.

Finally, we propose a much faster MC simulation algorithm, “simulating the total occupation time”, for European options in the two-state regime-switching model. Unlike the fundamental MC, the computational time of the algorithm is independent of the switching frequency. However, such an algorithm now is only applicable to the European options in a two-state regime-switching world.
Chapter 4

Convergence rate of regime-switching trees

4.1 Introduction

The acclaimed Black-Scholes model is the common language of security derivatives, and option prices are quoted using this model. In spite of this unparalleled triumph, the Black-Scholes model suffers from well-known shortcomings. One of them is that the risk-neutral rate $r$ and the volatility $\sigma$ should not be constant. The regime-switching model provides an enhancement of the Black-Scholes model which alleviates this problem. In this model, the market-related price-determining parameters $r$ and $\sigma$ of the Black-Scholes model are jointly determined by an externally driven market-related regime. While there can be several different forces acting on the price of an option, in this model one force (regime) is a dominating factor in setting the price, and the state of this regime is modelled to switch back and forth between finitely many modes. For instance, this could be the changes in preferences of the market agents [91] alternating between bullish and bearish expectations [54, 56] or, as in [14], alternating between good and bad. It can also be a business cycle [27] recurring from expansion, transition, and contraction. This price-driving force can also be determined by a hidden Markov process such as inside trading [32]. Numerous papers highlight that the regime-switching model is better than the Black-Scholes model in capturing the fat tails exhibited by empirical financial returns [17, 36, 37, 41, 52, 68]. In regime-switching models, asset prices evolve according to models determined by the state of some recurrently-switching regimes which are driven by unobserved factors resulting in
stationary regime-state changes following each other independently.

In its simplest form, the regime has two states, 1 and 2, and the risk-free rate and volatility are fully determined by this state. For simplicity, this chapter focuses on two-state regime-switching models. In an abstract form, a (two-state) regime-switching model \( \Xi := \Xi (\alpha, \xi^1, \xi^2) \) is composed of three independent components: a stochastic model \( \alpha_t \) governing the state of the regime, and two independent underlying asset stochastic models \( \xi^1_t \) and \( \xi^2_t \). Starting in a regime-state \( \alpha_0 = a \) and at a spot price \( \Xi_0 = x \), the value \( \Xi_t \) of the underlying asset matches the value of \( \xi^a_t \), that is \( \Xi_t = \xi^a_t \), until the first regime-switching occurs at time \( \tau_1 = \inf \{ t \geq 0 : \alpha_t \neq a_0 \} \). From that point on, the asset grows according to \( \xi^a_{\tau_1} \), that is \( \Xi_t = (\Xi_{\tau_1}) (\xi^a_{\tau_1} / \xi^a_{\tau_1}) \), until the regime-state changes again at time \( \tau_2 \). The process then continues as \( \Xi_t = (\Xi_{\tau_2}) (\xi^a_{\tau_2} / \xi^a_{\tau_2}) \) until the regime-state changes once more, and this scheme repeats itself forever.

In the "Black-Scholes" regime-switching model the two underlying stochastic processes, \( \xi^a \) for \( a = 1, 2 \), follow the Black-Scholes model with parameters \( r_a, \sigma_a \), while regime-state changes follow each other after waiting independent exponentially distributed times, the average waiting time being \( 1/\lambda_a \) when the regime is in state \( a \).

Let \( T \) be the maturity of some security derivatives, and let \( (T/n) \mathbb{N} \) be a discretization of the time interval. It is natural to be interested in discretizations of the Black-Scholes regime-switching model, namely piecewise constant approximations \( \Xi^{(n)}_t \) of \( \Xi_t \). These approximations include binomial and trinomial trees which are essential to price options for which a closed form solution is inexistent or computationally complicated such as in the case for American options. Taking again a high level and abstract view point, we will say that a stochastic process \( \Xi^{(n)} := \Xi^{(n)} (\alpha^n, \xi^{(1,n)}, \xi^{(2,n)}) \) is a partially discretized version of the (Black-Scholes) regime-switching model \( \Xi := \Xi (\alpha, \xi^1, \xi^2) \) if the parameters \( \alpha^n, \xi^{(1,n)}, \xi^{(2,n)} \) are either discretizations or identical versions of their corresponding parameter in \( \Xi \). A full discretization occurs when all tree parameters of \( \Xi^{(n)} \) are discretizations of their limiting \( \Xi \)-counterparts.

In the trinomial tree method for the Black-Scholes regime-switching model, the two underlying stochastic processes, \( \xi^a \) for \( a = 1, 2 \), are each approximated by a trinomial tree \( \xi^{a,n} \) for \( a = 1, 2 \). Furthermore, given that the regime is in state \( a \) at time \( t \in (T/n) \mathbb{N} \), the probability that it changes state at time \( t + T/n \) is \( 1 - \exp (-\lambda_a T/n) \). Recall that a self-similar trinomial tree \( S^{(n)} \) can be seen as a stochastic process which at every positive
time $t$ in $(T/n) \mathbb{N}$, has a probability $p_n^u$ of jumping from its current state $S_t^{(n)}$ to the state $S_t^{(n)} u_n$, a probability $p_n^d$ of jumping to the state $S_t^{(n)} d_n$, and a probability $1 - p_n^u - p_n^d$ of jumping to the state $S_t^{(n)} m_n$, for some $u_n, d_n, m_n > 0$.

Trinomial tree methods for regime-switching models have been studied in several papers. Bollen [6] presents a lattice-based method for valuing both European and American-style options and suggests that the regime-switching option values better match implied common volatility smiles in empirical studies. A discretization of the Cox-Ross-Rubinstein type for the regime-switching model is displayed in Guo [32]. Khaliq and Liu [47] compare an implicit scheme with a tree model that generalizes the Cox-Ross-Rubinstein binomial tree model, and with an analytical approximation solution for the two-regime case described in Buffington and Elliott [14]. Liu [53] designs regime-switching recombining tree. Yuen and Yang [87] present a fast and simple tree model to price simple and exotic options in Markov regime-switching models with multiple regime-states. Yoon et al. [86] develop a lattice method for pricing lookback options in a regime-switching market environment. Fuh et al. [27] provide a closed-form formula for the arbitrage-free price of the European call option, and use a tree method, among others, for calculating prices. Liu [54] introduces a lattice tree method for pricing financial derivatives in a regime-switching mean-reverting model. In Liu and Zhao [56] a lattice approach for option pricing with two underlying assets whose prices are governed by regime-switching models is developed. Yuen et al. [88] incorporate the regime-switching effect in a discrete time binomial model for an asset’s prices via the “self-exciting” threshold principle. Costabile et al. [17] present a binomial approach for pricing contingent claims when the parameters governing the underlying asset process follows a regime-switching model. A tree approach to options pricing under a regime-switching jump diffusion model is exhibited in Liu and Nguyen [55].

These natural questions arise: at what speed do option prices converge under typical trinomial tree discretizations? How does this convergence depend on the smoothness of the payoff?

Recently, Ma and Zhu [59, 60] investigated the speed of convergence of European options under Yuen and Yang’s trinomial method [87]. The authors considered a European option with maturity $T$. Letting $a \in \{1, 2\}$ represent the state of the regime, and $S$ be any node of the trinomial tree at time $t_k = Tk/n$, they denote by $\varepsilon_a^k (S) = V (S, t_k, a) - V^k (S, a)$ the difference between the option under the regime-switching model and the
same option under the trinomial tree method when the regime-state is \(a\), the spot price is \(S\), and the time is \(t_k = T k/n\). In the main result of their paper, Ma and Zhu state that, for \(k = 1, \ldots, n - 1\) and \(a = 1, 2\), \(\|\varepsilon_a^k\|_\infty = O\left(n^{-1}\right)\), where \(\|\varepsilon_a^k\|_\infty := \max_{-n \leq j \leq n} |\varepsilon_a^k(S_j)|\) and \(S_j = u^j S_0\). Unfortunately, Ma and Zhu do not specify any conditions for the payoff function. Their main result and its proof are only valid for payoff functions which are smooth enough and subject to boundedness conditions. This excludes call options, put options, binary options, and even payoff functions such as \(f(x) = x^2\). To explain this in the simplest manner, we will assume that \(\sigma_1 = \sigma_2\) and \(r_1 = r_2\) which brings us back to the Black-Scholes model with the parameters \(r, \sigma_1\). Furthermore, as in Ma and Zhu, we set

\[
\sigma := \max(\sigma_1, \sigma_2) + \left(\sqrt{1.5} - 1\right) \frac{\sigma_1 + \sigma_2}{2} = \sqrt{1.5}\sigma_1.
\]

Define \(\Lambda_1 = \sigma/\sigma_1 > 1\), and let \(S_0 = 1\). In this special case, Yuen and Yang’s trinomial tree becomes an ordinary trinomial tree approximating the Black-Scholes model with parameters \(r, \sigma_1\). More specifically, with \(\Delta t := T/n\),

\[
\begin{align*}
    u &= e^{\Lambda_1 \sigma_1 \sqrt{\Delta t}} = e^{\sigma \sqrt{\Delta t}}, \quad m = 1, \\
    d &= e^{-\Lambda_1 \sigma_1 \sqrt{\Delta t}} = e^{-\sigma \sqrt{\Delta t}}, \\
    p_u &= \frac{e^{r_1 h} - e^{-\Lambda_1 \sigma_1 \sqrt{\Delta t}}}{e^{\Lambda_1 \sigma_1 \sqrt{\Delta t}} - e^{-\Lambda_1 \sigma_1 \sqrt{\Delta t}}} \left(1 - \frac{1}{(\Lambda_1)^2}\right) \left(1 - e^{-\Lambda_1 \sigma_1 \sqrt{\Delta t}}\right), \\
    p_d &= \frac{e^{\Lambda_1 \sigma_1 \sqrt{\Delta t}} - e^{r_1 h}}{e^{\Lambda_1 \sigma_1 \sqrt{\Delta t}} - e^{-\Lambda_1 \sigma_1 \sqrt{\Delta t}}} \left(1 - \frac{1}{(\Lambda_1)^2}\right) \left(e^{\Lambda_1 \sigma_1 \sqrt{\Delta t}} - 1\right), \\
    p_m &= 1 - \frac{1}{(\Lambda_1)^2}.
\end{align*}
\]

First, consider the smooth payoff \(f(x) = x^2\). It is easy to see that

\[
\varepsilon_1^{n-1}(x) = x^2 \varepsilon_1^{n-1}(1).
\]

As pointed out in section 4.8 below, the Yuen and Yang model satisfies equation (4.8.3) below and, with \(\gamma = 2\),

\[
\varepsilon_1^{n-1}(1) = O\left(n^{-2}\right).
\]

Therefore,

\[
\max_{-n \leq j \leq n} |\varepsilon_a^{n-1}(S_j)| = \max_{-n \leq j \leq n} |S_j^2| |\varepsilon_a^{n-1}(1)| \geq |S_n^2| O\left(n^{-2}\right) = e^{2\sigma \sqrt{T} \sqrt{n}} O\left(n^{-2}\right)
\]
and obviously
\[
\lim_{n \to \infty} e^{2\pi \sqrt{T \pi} O\left(n^{-2}\right)} = \infty.
\]
Hence \(\|\varepsilon_n^{-1}\|_\infty\) fails to be \(O\left(n^{-1}\right)\) as it is actually unbounded! Second, consider now
the case of a call option with \(E = 1 = S_0\), where \(E\) is the strike. For \(S_0 = d = \exp\left(-\Lambda_1 \sigma_1 \sqrt{\Delta t}\right) < 1\) it is clear that the price (discounted expectation) in the one-time-step trinomial tree (maturity \(\Delta t\)) is zero; on the other hand, the price \(BS(d, \Delta t)\) of the
call option in the Black-Scholes model with a maturity of \(\Delta t\) and a spot price of \(S_0 = d\) satisfies

\[
BS\left(e^{-\Lambda_1 \sigma_1 \sqrt{\Delta t}}, \Delta t\right) = \frac{\sqrt{\Delta t}}{\sqrt{\pi}} \left(2\sigma_1 e^{-\lambda_1^2} + \sqrt{\pi} \sigma_1 \Lambda_1 \text{erf} \left(\Lambda_1\right)\right) + O\left(\Delta t\right),
\]
showing again that

\[
\|\varepsilon_1^{n-1}\|_\infty \geq O\left(n^{-\frac{1}{2}}\right) , \quad \|\varepsilon_1^{n-1}\|_\infty \neq O\left(n^{-1}\right) .
\]
Finally, using the Berry-Esseen theorem, [49] proves that for digital options

\[
\sup_{x \geq 0} \left|\varepsilon_1^0(x)\right| = O\left(n^{-\frac{1}{2}}\right),
\]
which again does not match the main result in [60]. These three examples show that Ma
and Zhu’s result [59, 60] fails unless conditions are put on the payoff function.

The reason for these problems comes from the use in Ma and Zhu [60, eq. (18) and
(19)] of Taylor’s theorem to compute \(V(S_j, t_k, i) - V(S_j, t_{k+1}, i)\) and \(V(S_{j+1}, t_{k+1}, i) - V(S_j, t_{k+1}, i)\), requiring that the reminders be of order \(O\left(\Delta t^2\right)\) uniformly in \(t_k\) and \(S_j\).
Unfortunately, for this to be true, the payoff function should be sufficiently smooth and
subject to boundedness conditions. Indeed, still considering the case of the call option
(a non-differentiable payoff) when the two regimes are identical (Black-Scholes with pa-
rameters \(r_1, \sigma_1\)), and letting \(BS\left(S, t\right)\) denote the value of the option in the Black-Scholes
model at time \(t\) and spot price \(S\), it is easy to calculate that

\[
BS\left(E, 2\Delta t\right) - BS\left(E, \Delta t\right) + \left(\frac{\partial}{\partial t} BS\left(E, \Delta t\right)\right) \Delta t
= \sqrt{\Delta t} \frac{E \sigma_1}{\sqrt{\pi}} \left(2\sqrt{2} - 1\right) + O\left(\Delta t\right),
\]
which contradicts [60, equation (18)] when $t_k = (n - 2) \Delta t$.

In summary, both the result and Ma and Zhu’s proof [59, 60] can be valid only for options with smooth payoff functions subject to boundedness conditions, excluding the most interesting cases which only have a piecewise smooth payoff function such as put options, call options, and digital options.

This problem is addressed in this chapter and constitutes its main result. We show that for a broad family of piecewise smooth payoff functions (including call, put, and digital options) and for a large class of discretizations of the two-state regime-switching model, convergence occurs at a rate of $O(n^{-1})$ for continuous payoff functions, and at a rate of $O(n^{-1/2})$ when the payoff is discontinuous. These discretizations include, but are not limited to, trinomial trees and other lattice methods. In particular, Yuen and Yang’s trinomial model [87] falls under our setting.

4.2 Settings

This section describes the building blocks on which this chapter relies.

4.2.1 Payoff function class

We say that a function $h$ is piecewise $C^m$, for some integer $m \geq 0$, if there exists countably many intervals $J_\ell := [\beta_\ell, \beta_{\ell+1})$, $\beta_0 < \beta_1 < \ldots$, forming a partition of $[0, 1)$ and functions $h_\ell$ extendible to be $C^m$ on the closure of $J_\ell$, such that

$$h(x) = \sum_{\ell=0}^{\infty} h_\ell(x) 1_{[\beta_\ell, \beta_{\ell+1})}(x).$$

We use $I$ to denote the identity function, that is $I(z) := z$ for every $z$. Given an integer $k$, we set $I^k(z) := z^k$. We denote by $\mathcal{K}^{(m)}$ the class of piecewise $C^{(m)}$ functions such that $h$, $Ih$, $\ldots$, $I^m h^{(m)}$ have a limit at infinity and are of bounded variation over $[0, \infty)$. Clearly, for any $h \in \mathcal{K}^{(m)}$, functions $h$, $Ih$, $\ldots$, $I^m h^{(m)}$ are bounded and we define a norm $\zeta_m$ on $\mathcal{K}^{(m)}$ as

$$\zeta_m(h) = \sum_{k=0}^{m} \left( TV(I^k h^{(k)}) + \|I^k h^{(k)}\|_\infty \right),$$

where $TV(g)$ is the total variation of $g$ over the interval $[0, \infty)$. 

4.2.2 Black-Scholes discretization

Let $t_m := Tm/n$, for $m = 0, 1, \ldots$. In this chapter, the Black-Scholes model is discretized using geometric random walks. These are piecewise constant stochastic processes $\{\xi^{(n)}\}$ of the form

$$
\xi_t^{(n)} = \xi_0 \exp \left( \sum_{k=1}^{[nt/T]} X_n [t_{k-1}, t_k] \right),
$$

with random variables $X_n [t_{k-1}, t_k]$ independent and identically distributed as

$$
X_n \overset{\text{def}}{=} X_n [0, T/n].
$$

We consider only discrete approximations $\{\xi^{(n)}\}$ of the Black-Scholes model with parameters $r$ and $\sigma$ for which the following assumptions on $X_n$ hold:

$$
\mu_n \overset{\text{def}}{=} \mathbb{E}(X_n) = \frac{T}{n} \left( r - \frac{1}{2} \sigma^2 \right) + \mathcal{O} \left( n^{-\frac{3}{2}} \right), \quad (A1)
$$

$$
\sigma_n \overset{\text{def}}{=} \sqrt{\text{Var}(X_n)} = \sqrt{\frac{T}{n}} \left( \sigma + \mathcal{O} \left( n^{-\frac{1}{2}} \right) \right), \quad (A2)
$$

Furthermore, for every real constant $\gamma$,

$$
\mathbb{E}(\exp (\gamma X_n)) = \exp \left( \frac{1}{2} \frac{T}{n} \gamma (2r - \sigma^2 + \sigma^2 \gamma) \right) + \mathcal{O} \left( n^{-2} \right), \quad (A3)
$$

$$
\mathbb{E} \left( \exp \left( \gamma \sqrt{\frac{n}{T}} X_n \right) \right) = \mathcal{O}(1). \quad (A4)
$$

For binomial trees, it was shown in [49] that all put options, all call options, and all options with polynomial payoffs converge at a rate of $n^{-1}$ to the Black-Scholes price with risk-free rate $r$ and volatility $\sigma$ if and only if assumptions $A1$-$A4$ hold.

**Remark 1.** In a risk-neutral setting, the price of a European option coincides with the discounted expectation of the payoff at maturity. Assumption $A3$ amounts to what is called quasi risk neutrality in [49], where the discounted expectation of the asset price over an interval of time of size $T/n$ is equal to the spot price plus an error of order $1/n^2$. For mere simplicity of the presentation, we will assume in this chapter that $\xi^1$ and $\xi^2$ are risk-neutral.
4.2.3 Regime-state discretizations

Recall that $\alpha_t$, the regime-state stochastic process, is piecewise constant which values in \{1, 2\}, and it changes value after waiting independent exponentially distributed times, the average waiting time being $1/\lambda_a$ when the regime is in state $a$. In this chapter, we will approximate $\alpha_t$ by a piecewise constant process $\alpha^n_t$ which, at every time step $t_m$, changes from state $a \in \{1, 2\}$ to state $a' \neq a$ with probability $p^n_a$, and remains in state $a$ with probability $1 - p^n_a$. One possible instance of $\alpha^n_t$ is the process which changes state at time $t_m$, $m \geq 1$, if and only if $\alpha_t$ experiences at least one jump in the interval $(t_{m-1}, t_m]$. In this case $p^n_a = 1 - \exp(-\lambda a T/n)$. We will call this specific instance of $\alpha^n_t$ the default discretized regime-state space. It differs from the snapshot discretized regime-state space $\tilde{\alpha}^n_t$ defined as

$$\tilde{\alpha}^n_t = \sum_{m=0}^{\infty} \alpha_{t_m} \mathbf{1}_{(t_m,t_{m+1})}(t).$$

Note that, on any trajectory $t \mapsto \alpha_t$ where a maximum of one jumps occurs in any interval $[t_{m-1}, t_m]$, $m = 1, \ldots, n$, we have

$$\sup_{0 \leq t \leq T} |\tilde{\alpha}^n_t - \alpha^n_t| = 0.$$

Throughout this chapter we use $\lambda \overset{\text{def}}{=} \lambda_1 \lor \lambda_2$. Furthermore, for every $0 \leq s \leq t$, $N[s, t]$ denote the number of jumps of $\alpha_t$ in the interval $[s, t]$, while $N_{\lambda}[s, t]$ denote the number of jumps of a Poisson process with intensity $\lambda$ in the same interval. Recall that for any $0 \leq s \leq T$, $\Delta \alpha_s = \alpha_s - \alpha_{s-}$.

We will denote by $\mathcal{A}$ the event that, over the interval of time $[0, T]$, the trajectory $t \mapsto \alpha_t$ has a maximum of one jump in any interval $(t_{m-1}, t_m]$. That is

$$\mathcal{A} = \cap_{m=1}^{n} \mathcal{A}_m$$

(4.2.2)

where

$$\mathcal{A}_m = \left\{ \omega : \sum_{t_{m-1} < s \leq t_m} |\Delta \alpha_s| \leq 1 \right\}.$$

Consider now the probability of the complement $\mathcal{A}_m^c$ of $\mathcal{A}_m$, that is the probability that over the interval $(t_{m-1}, t_m]$ the trajectory $t \mapsto \alpha_t$ has at least two jumps. Note that $P(\mathcal{A}_m^c)$ is smaller than the probability that $t \mapsto N_{\lambda}[0, t]$ has at least two jumps over the
same interval. Hence,

\[ P (A_m^c) \leq 1 - e^{-\frac{\lambda T}{n}} \left( 1 + \left( \frac{\lambda T}{n} \right) \right) = \mathcal{O} (n^{-2}). \]

It follows that

\[ P (A) = 1 - \mathcal{O} (n^{-1}), \quad P (A^c) = \mathcal{O} (n^{-1}). \]  

(4.2.3)

This allows us to define the pseudo discretized regime-state space, \( \overline{\alpha}_t^n \), as

\[ \overline{\alpha}_t^n = 1_A \alpha_t^n + 1_{A^c} \alpha_t. \]

While this stochastic process is neither a discretization nor a Markov process, it plays an important role in this chapter because it is precisely equal to the default discretized regime-state space \( \alpha_t^n \) except on some ‘negligible’ event of probability \( \mathcal{O} (n^{-1}) \). It allows us to strictly focus our efforts only on those ‘non-negligible’ trajectories in \( A \).

#### 4.2.4 Occupation time random variables

Throughout this chapter, the maturity \( T \) is fixed. For \( a \in \{1, 2\} \), the occupation time random variable \( L_a \) is, by definition, equal to the total time spent in state \( a \) by the regime-state process \( \alpha_t \) over the interval \([0, T]\). Hence,

\[ L_a = \int_0^T \delta_a (\alpha_t) \, dt. \]

Note that \( 0 \leq L_a \leq T \), \( L_1 + L_2 = T \). We define in the same manner \( \hat{L}_a^n \), \( \tilde{L}_a^n \) and \( \overline{L}_a^n \) to be the occupation time random variables of \( \alpha_t^n \), \( \hat{\alpha}_t^n \), and \( \overline{\alpha}_t^n \). Recall \( A \) from section 4.2.3, the event that over the interval of time \([0, T]\), the trajectory \( t \mapsto \alpha_t \) has a maximum of one jump in any interval \([t_{m-1}, t_m]\). Obviously,

\[ 0 = \sup_{0 \leq t \leq T} 1_A \left| L_a^n - \hat{L}_a^n \right| = \sup_{0 \leq t \leq T} 1_A \left| L_a^n - \tilde{L}_a^n \right| = \sup_{0 \leq t \leq T} 1_{A^c} \left| \tilde{L}_a^n - L_a \right|. \]

It follows that, for any bounded measurable function \( f \),

\[ \mathbb{E} (f (L_1^n, L_2^n)) = \mathbb{E} \left( f \left( \hat{L}_1^n, \hat{L}_2^n \right) \right) + \| f \|_{\infty} \mathcal{O} (n^{-1}) \]

\[ = \mathbb{E} \left( f \left( \overline{L}_1^n, \overline{L}_2^n \right) \right) + \| f \|_{\infty} \mathcal{O} (n^{-1}). \]
4.3 Option price closed-form formula

Throughout this chapter, the price of a European option refers to the discounted expectation of the payoff at maturity. The purpose of this section is to introduce the notation related to expectations and discounted expectations of regime-switching models. Furthermore, we give a closed-form formula for the regime-switching Black-Scholes model. Here again \( \Xi := \Xi (\alpha, \xi^1, \xi^2) \) is a regime-switching Black-Scholes model, and \( \Xi(n) := \Xi(n) (\alpha^n, \xi^{1,n}, \xi^{2,n}) \) is a partial or full discretization discretization of \( \Xi \).

4.3.1 Black-Scholes model

We denote respectively by \( E_a^a \) and \( E_{a;n}^a \), the conditional expectation given that \( \xi_0^a = x \) and \( \xi_{0;n}^a = x \), where \( a \) is the regime-state, \( a = 1, 2 \). Additionally, for any payoff function \( \psi \), we denote the discounted expectations of the payoff under a Black-Scholes model, \( \xi^a \), and its discretization, \( \xi^{a,n} \), by

\[
E_t^a \psi (x) \overset{def}{=} e^{-r_a t E_x^a (\psi (\xi_t^a))},
\]

\[
E_{t;n}^a \psi (x) \overset{def}{=} e^{-r_a t E_{x;n}^a (\psi (\xi_{t;n}^a))}.
\]

4.3.2 Regime-switching Black-Scholes model

In the Black-Scholes model the price of an asset \( \xi_t^a \) satisfies

\[
d\xi_t^a = r_a \xi_t^a dt + \sigma_a \xi_t^a dW_t^a
\]

where \( W_t^a \) is a Brownian motion. Assuming that \( \xi_0^a = 1 \), this solves as

\[
\xi_t^a = \exp \left( \sigma_a W_t^a + \left( r_a - \frac{1}{2} \sigma_a^2 \right) t \right).
\]

On the other hand, in the regime-switching Black-Scholes model the price of an asset \( \Xi_t \) satisfies

\[
d\Xi_t = (r_{a_t}) \Xi_t dt + (\sigma_{a_t}) \Xi_t dW_t
\]
for some independent Brownian motion $W_t$. This solves as

$$
\Xi_t = \Xi_0 \exp \left( \int_0^t \left( r_{\alpha_s} - \frac{\sigma_{\alpha_s}^2}{2} \right) ds + \int_0^t (\sigma_{\alpha_s}) dW_s \right)
= \Xi_0 \Pi_{a=1}^2 \exp \left( \sigma_a W_{L_a}^a + \left( r_a - \frac{1}{2} \sigma_a^2 \right) L_a \right)
= \Xi_0 \xi_1^a L_1 \xi_2^a L_2
$$

where $L_a := L_a ([0, t])$ is the occupation time in state $a$ over the interval $[0, t]$, that is,

$$
L_a = \int_0^t \delta_a (\alpha_s) ds.
$$

Now for any payoff function $\psi$, let us denote by $\mathcal{E}_t \psi (a, x)$ the expectation of the discounted of the payoff $\psi (\Xi_t)$ at time $t$ given that $(\alpha_0, \Xi_0) = (a, x)$. Furthermore, we denote by $\mathbb{E}_a^{L_1}$ the expectation with respect to $L_1$ given that $\alpha_0 = a$. To avoid unnecessary complications we sometimes simply use $\mathbb{E}$ to denote the expectation given the initial value of the stochastic processes involved. Then,

$$
\mathcal{E}_t \psi (a, x) = \mathbb{E} \left( e^{-r_1 L_1} e^{-r_2 L_2} \psi (x \xi_1^1 L_1 \xi_2^2 L_2) \right). \quad (4.3.1)
$$

Moreover, because of the independence of $L_1$, $\xi_1^1$ and $\xi_2^2$,

$$
\mathcal{E}_t \psi (a, x) = \mathbb{E}_a^{L_1} \left( \mathbb{E}_1^{\xi_1} \left[ e^{-r_1 L_1} \mathbb{E}_1^{\xi_2} \left( e^{-r_2 L_2} \psi (x \xi_1^1 L_1 \xi_2^2 L_2) \right) \right] \right)
= \mathbb{E}_a^{L_1} \left( \mathbb{E}_1^{\xi_1} \left( e^{-r_1 L_1} \xi_2^2 L_2 h (x \xi_1^1) \right) \right)
= \mathbb{E}_a^{L_1} \left( \xi_1^1 \left( \xi_2^2 L_2 (x) \right) h \right).
$$

In a similar manner,

$$
\mathcal{E}_t \psi (a, x) = \mathbb{E}_a^{L_2} \left( \xi_1^1 \left( \xi_2^2 L_2 (x) \right) \right).
$$

Note that random variables $L_1$ and $L_2$ are related by $L_1 + L_2 = T$. Furthermore, $L_1$ and $L_2$ have a density. We denote by $f_1^a (t)$ and $f_2^a (t)$ the density functions of $L_1$ and $L_2$ given that $\alpha_0 = a$. A closed form formula for $f_1^a (t)$ and $f_2^a (t)$ can be found in [17, 27, 32, 66]. Then,

$$
\mathcal{E}_t \psi (a, x) = \int_0^T \mathcal{E}_t^1 \left( \xi_2^2 L_2 (x) f_1^a (t) \right) dt = \int_0^T \mathcal{E}_T^1 \left( \xi_2^2 h (x) f_1^a (t) \right) dt.
$$
It is easy to see that $E_t^1 \left( \mathcal{E}_{T-t}^2 \right) (h) (x) = E_t^* (h) (x)$, where $E_t^* (h) (x)$ is the price of an option in the Black-Scholes model with maturity $T$, spot price $x$, payoff function $h$, risk-neutral rate $r_s (t)$, and volatility $\sigma_s (t)$ where

$$r_s (t) = \frac{t}{T} r_1 + \frac{T - t}{T} r_2,$$
$$\sigma_s^2 (t) = \frac{t}{T} \sigma_1^2 + \frac{T - t}{T} \sigma_2^2.$$

This is a simple reformulation of a formula in [17, 66]. As in [91], it produces a closed form formula requiring only one integral,

$$E_t \psi (a, x) = \int_0^T E_t^* (h) (x) f_a^1 (t) \, dt. \tag{4.3.2}$$

In particular, for a European put we get

$$E_t^* (h) (x) = \Phi (-d_2^* (t)) E e^{-r_s (t) T} - \Phi (-d_1^* (t)) x,$$
$$d_1^* (t) = \frac{1}{\sigma_s (t) \sqrt{T}} \left( \ln \left( \frac{x}{E} \right) + \left( r_s (t) + \frac{\sigma_s^2 (t)}{2} \right) T \right),$$
$$d_2^* (t) = d_1^* (t) - \sigma_s (t) \sqrt{T}.$$

For the digital put option we have

$$E_t^* (h) (x) = e^{-r_s (t) T} \Phi (-d_2^* (t)).$$

**Remark 2.** Let $L$ be any measurable subset of $[0, T]$ and $L^c$ its complement. Clearly,

$$E_t \psi (a, x) = \int_0^t 1_{L} (s) E_s^1 \left( \mathcal{E}_{T-s}^2 \right) (h) (x) f_a^1 (s) \, ds$$
$$+ \int_0^t 1_{L^c} (s) E_s^1 \left( \mathcal{E}_{T-s}^2 \right) (h) (x) f_a^1 (s) \, ds$$

or, written differently,

$$E_t \psi (a, x) = \int_0^t 1_{L} (s) E_s^1 \left( \mathcal{E}_{T-s}^2 \right) (h) (x) f_a^1 (s) \, ds$$
$$+ \int_0^t 1_{L^c} (s) E_s^2 \left( \mathcal{E}_{T-s}^2 \right) (h) (x) f_a^1 (s) \, ds.$$

Analogue expressions are valid for the expectation and conditional expectation with respect
to some partial or full discretization $\Xi^{(n)}$. In particular,

$$
\mathcal{E}_t^n(a,x) = E \left( e^{-r_1 L_1^n} e^{-r_2 L_2^n} \left( x \xi_{L_1^n}^{1,n} \xi_{L_2^n}^{2,n} \right) \right)
= E_a^{L_1^n} \left( 1_{L_1^n} (L_1^n) \xi_{L_1^n}^{1,n} \xi_{L_2^n}^{2,n} \left( x \right) \right)
+ E_a^{L_1^n} \left( 1_{L_2^n} (L_1^n) \xi_{L_2^n}^{2,n} \left( x \right) \right).
$$

where $E_a^{L_1^n}$ is the expectation with respect to the random variable $L_1^n$, given that $\alpha_0 = a$.

4.4 Outline of the chapter

Given a regime-switching Black-Scholes model $\Xi := \Xi(\alpha, \xi^1, \xi^2)$, this chapter considers the following families of partial and full discretizations: $\widehat{\Xi}^{(n)} := \Xi(\alpha^n, \xi^1, \xi^2)$ and $\Xi^{(n)} := \Xi(\alpha^n, \xi^{1,n}, \xi^{2,n})$, where $\xi^{1,n}$ and $\xi^{2,n}$ are discrete approximations of Black-Scholes models $\xi^1$ and $\xi^2$ satisfying assumptions A1-A4, and where $\alpha^n$ is the default discretized regime-state space of section 4.2.3. In section 4.5, we show that for an option with payoff $h$ in $K^{(2)}$, the pricing error resulting from using model $\widehat{\Xi}^{(n)}$ instead of the regime-switching $\Xi$ is of order $O\left(n^{-\beta}\right)$, where $\beta$ depends on whether or not $h$ is continuous. More specifically, we show that, given an initial state of the regime of $a$ and a spot price of $x$, if $\widehat{\xi}_t^n h(a,x)$ and $\xi_t^n h(a,x)$ are respectively the discounted expectation of the payoff at maturity $t$ under model $\widehat{\Xi}^{(n)}$ and $\Xi$, then there exists a constant $Q$, which does not depend on $h$ or $x$, such that

$$
\left| \xi_t h(a,x) - \widehat{\xi}_t^n h(a,x) \right| \leq \chi_2(h) Q n^{-\beta},
$$

where $\beta = 1/2$ if $h$ is discontinuous and $\beta = 1$ otherwise. In section 4.6, we show that for the same option, the pricing error resulting from using model $\Xi^{(n)}$ instead of model $\widehat{\Xi}^{(n)}$ is of order $O\left(n^{-1}\right)$ when $h$ is continuous but of order $O\left(n^{-1/2}\right)$ otherwise. More specifically, we show that there exists a constant $Q$, which does not depend on $h$ or $x$, such that

$$
\left| \xi_t^n h(a,x) - \widehat{\xi}_t^n h(a,x) \right| \leq \chi_2(h) Q n^{-\beta},
$$

where $\beta = 1/2$ if $h$ is discontinuous and $\beta = 1$ otherwise. This proves that

$$
|\xi_t^n h(a,x) - \xi_t h(a,x)| \leq \chi_2(h) Q n^{-\beta},
$$
where $\beta = 1/2$ if $h$ is discontinuous and $\beta = 1$ otherwise. In section 4.7, focussing on payoffs of the form $h(x) = x^\gamma$, for any real $\gamma$, we show that there exists a constant $Q$, which does not depend on $h$ or $x$, such that for every $x > 0$,

$$|\mathcal{E}^n h(a, x) - \mathcal{E} h(a, x)| \leq Q x^\gamma n^{-1}.$$ 

In section 4.8, we explain how the trinomial method of Yuen and Yang [87] corresponds to a full discretizations of the form $(n^2) := (n^1, 1; n^2, 1)$ where $n^1$ and $n^2$ are discrete approximations of the Black-Scholes models $\xi^1$ and $\xi^2$ satisfying assumptions A1-A4, and where $\alpha^n$ is the default discretized regime-state space. Section 4.9 provides numerical results illustrating our findings. Auxiliary results are found in Section 4.10.

4.5 Partial discretization error

Here we investigate the error resulting from replacing the parameter $\alpha$ in $\Xi (\alpha, \xi^1, \xi^2)$ by $\alpha^n$. This can be seen alternatively as replacing $L_a$ by $L^n_a$ in the price formula (4.3.1).

More specifically, we show the following:

**Proposition 4.5.1** (Regime-state discretization error). Assume that properties A1-A4 hold and let $h$ belong to $\mathcal{K}^{(2)}$. Then,

$$\mathbb{E} \left( e^{-r_1 L^n_1} e^{-r_2 L^n_2} h \left( x \xi_{L^n_1}^1, \xi_{L^n_2}^2 \right) \right) = \mathbb{E} \left( e^{-r_1 L_1} e^{-r_2 L_2} h \left( x \xi_{L_1}^1, \xi_{L_2}^2 \right) \right) + \chi_2 (h) \mathcal{O} (n^{-1}),$$

where the $\mathcal{O} (n^{-1})$ term is uniform in $h$ and $x$.

**Proof.** Recall $T^n_a$ and $A$ from section 4.2. Because $h$ is bounded, because $L^n_a$ coincides with $L^n_a$ on $A$ and because $P (A^c) = \mathcal{O} (n^{-1})$, it follows that

$$\mathbb{E} \left( e^{-r_1 L^n_1} e^{-r_2 L^n_2} h \left( x \xi_{L^n_1}^1, \xi_{L^n_2}^2 \right) \right) = \mathbb{E} \left( e^{-r_1 L_1} e^{-r_2 L_2} h \left( x \xi_{L_1}^1, \xi_{L_2}^2 \right) \right) + \chi_2 (h) \mathcal{O} (n^{-1}).$$

Note that, from Lemma 1,

$$\mathbb{E} \left( |T^n_a - L_a| \right) = \mathcal{O} (n^{-1}),$$

$$P \left( \frac{|T^n_a - L_a|}{T} > \frac{T}{4} \right) = \mathcal{O} (n^{-1}).$$
Define \( L_1 \overset{def}{=} \{ L_1 \geq \frac{T}{2} \} \) and its complement \( L_2 \overset{def}{=} \{ L_2 > \frac{T}{2} \} \). Recall that

\[
\mathbb{E} \left( e^{-r_1 T_n} e^{-r_2 T_n} h \left( x \xi_{T_n}^1 \xi_{T_n}^2 \right) \right) = \mathbb{E} \left( e^{-r_2 T_n} 1_{L_1} \mathcal{E}_{T_n}^1 h \left( x \xi_{T_n}^2 \right) \right) \\
+ \mathbb{E} \left( e^{-r_1 T_n} 1_{L_2} \mathcal{E}_{T_n}^2 h \left( x \xi_{T_n}^1 \right) \right) \\
= \mathbb{E} \left( 1_{L_1} \mathcal{E}_{T_n}^2 \mathcal{E}_{T_n}^1 h \left( x \right) \right) \\
+ \mathbb{E} \left( 1_{L_2} \mathcal{E}_{T_n}^1 \mathcal{E}_{T_n}^2 h \left( x \right) \right).
\]

The above equations remain true with \( T_n \) and \( \bar{T}_n \) replaced with \( L_1 \) and \( L_2 \). Hence we only need to show that

\[
\mathbb{E} \left( e^{-r_2 T_n} 1_{L_1} \mathcal{E}_{T_n}^1 h \left( x \xi_{T_n}^2 \right) \right) = \mathbb{E} \left( e^{-r_2 L_2} 1_{L_1} \mathcal{E}_{L_1}^1 h \left( x \xi_{L_1}^2 \right) \right) \\
+ \chi_2 (h) O \left( n^{-1} \right),
\]

\[
\mathbb{E} \left( e^{-r_1 T_n} 1_{L_2} \mathcal{E}_{T_n}^2 h \left( x \xi_{T_n}^1 \right) \right) = \mathbb{E} \left( e^{-r_1 L_1} 1_{L_2} \mathcal{E}_{L_2}^2 h \left( x \xi_{L_2}^1 \right) \right) \\
+ \chi_2 (h) O \left( n^{-1} \right).
\]

The two cases are symmetric so we need to prove only the first equation. From Taylor’s theorem,

\[
1_{L_1} \mathcal{E}_{T_n}^1 h \left( x \xi_{T_n}^2 \right) = 1_{L_1} \mathcal{E}_{L_1}^1 h \left( x \xi_{L_1}^2 \right) + 1_{L_1} \int_{L_1}^{T_n} \frac{\partial}{\partial t} \mathcal{E}_{L_1}^1 h \left( x \xi_{T_n}^2 \right) dt.
\]

Note that according to Lemma 2,

\[
\sup_{t \geq \frac{T}{2}} \sup_{z \geq 0} \left\| \frac{\partial}{\partial t} \mathcal{E}_L^1 h \left( z \right) \right\|_{\infty} = O \left( 1 \right) \chi_2 (h).
\]

Hence,

\[
1_{L_1} \left\| \int_{L_1}^{T_n} \frac{\partial}{\partial t} \mathcal{E}_{L_1}^1 h \left( x \xi_{T_n}^2 \right) dt \right\| \leq O \left( 1 \right) \chi_2 (h) \left| L_1 - \bar{T}_n \right|
\]

Thus from (4.5.1),

\[
\mathbb{E} \left( 1_{L_1} e^{-r_2 T_n} \mathcal{E}_{L_1}^1 h \left( x \xi_{T_n}^2 \right) \right) = \mathbb{E} \left( 1_{L_1} e^{-r_2 L_2} \mathcal{E}_{L_1}^1 h \left( x \xi_{L_2}^2 \right) \right) + \chi_2 (h) O \left( n^{-1} \right)
\]

\[
= \mathbb{E} \left( 1_{L_1} \mathcal{E}_{L_2}^1 \mathcal{E}_{L_1}^1 h \left( x \right) \right) + \chi_2 (h) O \left( n^{-1} \right).
\]
Hence, in order to prove (4.5.3), we need only to show that

\[ E\left( L_1 E_{L_2}^2 E_{L_1} h(x) \right) = E\left( L_1 E_{L_2}^2 E_{L_1} h(x) \right) + \chi_2 (h) \mathcal{O}(n^{-1}). \]

Now let

\[ L_1' := \{ L_1 \geq \frac{T}{2} \} \cap \{ L_2 \geq T \}, \]
\[ L_1'' := \{ L_1 \geq \frac{T}{2} \} \cap \{ L_2 < T \}. \]

Because \( L_1 \) is the disjoint union of \( L_1' \) and \( L_1'' \), in order to establish (4.5.3), we only need to show that

\[ E\left( L_1' E_{L_2}^2 (E_{L_1} h) (x) \right) = E\left( L_1' E_{L_2}^2 (E_{L_1} h) (x) \right) + \chi_2 (h) \mathcal{O}(n^{-1}), \]
\[ E\left( L_1'' E_{L_2}^2 (E_{L_1} h) (x) \right) = E\left( L_1'' E_{L_2}^2 (E_{L_1} h) (x) \right) + \chi_2 (h) \mathcal{O}(n^{-1}), \]

The two cases can be treated in a completely symmetrical manner: using Taylor’s expansion around \( x \xi_{L_2}^2 \) in the first case, and around \( x \xi_{L_2}^2 \) in the second case. We will leave the second case as an exercise for the reader. Note that in order to establish (4.5.4), that is,

\[ E\left( e^{-r_2 T_2} L_1 E_{L_1} (x \xi_{L_2}^2) \right) = E\left( e^{-r_2 T_2} L_1 E_{L_1} (x \xi_{L_2}^2) \right) + \| h \|_\infty \mathcal{O}(n^{-1}), \]

we can replace \( e^{-r_2 T_2} \) by \( e^{-r_2 T_2} \) on the right hand side, that is, we only need to prove that

\[ E\left( e^{-r_2 T_2} L_1 E_{L_1} (x \xi_{L_2}^2) \right) = E\left( e^{-r_2 T_2} L_1 E_{L_1} (x \xi_{L_2}^2) \right) + \chi_2 (h) \mathcal{O}(n^{-1}) \]

Indeed, it easily follows from (4.5.1) that

\[ E\left( e^{-r_2 T_2} L_1 E_{L_1} (x \xi_{L_2}^2) \right) = E\left( e^{-r_2 T_2} L_1 E_{L_1} (x \xi_{L_2}^2) \right) + \| h \|_\infty \mathcal{O}(n^{-1}). \]
Using Taylor’s theorem, we write

\[ e^{-r_2 T_n^{L_2^n}} L_1 L_1' \mathcal{E}^1_{L_1} h \left( x \xi^2_{L_2} \right) = e^{-r_2 T_n^{L_2^n}} L_1 L_1' \mathcal{E}^1_{L_1} h \left( x \xi^2_{L_2} \right) + e^{-r_2 T_n^{L_2^n}} L_1 L_1' \frac{\partial}{\partial x} \mathcal{E}^1_{L_1} h \left( x \xi^2_{L_2} \right) \left( x \xi^2_{L_2} - x \xi^2_{L_2} \right) + 
 \]

\[ + e^{-r_2 T_n^{L_2^n}} L_1 L_1' \int_{x \xi^2_{L_2}}^{x \xi^2_{L_2}} \left( \frac{\partial^2}{\partial x^2} \mathcal{E}^1_{L_1} h \left( z \right) \right) \left( z - x \xi^2_{L_2} \right) dz. \]

To obtain (4.5.5) we will show that the expectation of the last two terms has the form \( \chi_2 \left( h \right) O \left( n^{-1} \right) \). Let us denote

\[ \Delta L_2 := L_2 - T_n^{L_2^n}, \]
\[ \bar{\xi}^2_{\Delta L_2} := \frac{\xi^2_{L_2}}{\xi^2_{L_2}}. \]

Note that, as \( \Delta L_2 \mid L_1 \mid \Delta L_2 > 0 \) on \( L_1 \). Basic properties of the geometric Brownian motion guarantee that \( \bar{\xi}^2_{\Delta L_2} \) is independent of \( \xi^2_{T_2} \) and identically distributed as \( \xi^2_{\Delta L_2} \).

Obviously,

\[ x \xi^2_{L_2} = \left( x \xi^2_{L_2} \right) \bar{\xi}^2_{\Delta L_2}, \]
\[ x \xi^2_{L_2} - x \xi^2_{T_2} = x \xi^2_{T_2} \left( \bar{\xi}^2_{\Delta L_2} - 1 \right). \]

The independence of \( \xi^2_{T_2} \) and \( \bar{\xi}^2_{\Delta L_2} \) gives that

\[ \left| E \left( e^{-r_2 T_n^{L_2^n}} L_1 L_1' \frac{\partial}{\partial x} \mathcal{E}^1_{L_1} h \left( x \xi^2_{L_2} \right) x \xi^2_{T_2} \left( \bar{\xi}^2_{\Delta L_2} - 1 \right) \right) \right| = \left| E \left( e^{-r_2 T_n^{L_2^n}} L_1 L_1' \frac{\partial}{\partial x} \mathcal{E}^1_{L_1} h \left( x \xi^2_{L_2} \right) x \xi^2_{T_2} \left( \bar{\xi}^2_{\Delta L_2} - 1 \right) \right) \right| \]

Invoking Lemma \( 2 \) we achieve

\[ \left| E \left( e^{-r_2 T_n^{L_2^n}} L_1 L_1' \frac{\partial}{\partial x} \mathcal{E}^1_{L_1} h \left( x \xi^2_{L_2} \right) x \xi^2_{T_2} \left( \bar{\xi}^2_{\Delta L_2} - 1 \right) \right) \right| \leq \chi_2 \left( h \right) \left| E \left( \xi^2_{\Delta L_2} - 1 \right) \right| \]

\[ = \chi_2 \left( h \right) \left| E \left( x^{\Delta L_2} - 1 \right) \right| \]
\[ = \chi_2 \left( h \right) O \left( 1 \right) \left| E \left( x^{\Delta L_2} \right) \right|,
\]
and from (4.5.1) we conclude that
\[
\left| \mathbb{E} \left( e^{-r_2 T_2} \mathbb{L}_1 \left( \frac{\partial}{\partial x} \mathbb{E}^1_{L_1} h \left( x \xi_{T_2}^2 \right) x \xi_{T_2}^2 \left( \xi_{\Delta L_2}^2 - 1 \right) \right) \right) \right| \leq \chi_2 (h) \mathcal{O} (n^{-1}).
\]

Now we define \( \mathcal{L} \) as
\[
\mathcal{L} := \mathbb{E} \left( e^{-r_2 T_2} \mathbb{L}_1 \left( \int_{\xi_{T_2}^2}^{z_{\Delta L_2}} \frac{\partial^2}{\partial x^2} \mathbb{E}^1_{L_1} h \left( z \right) \left( z - x \xi_{T_2}^2 \right) dz \right) \right).
\]

To complete this proof, we need to show that \( \mathcal{L} = \chi_2 (h) \mathcal{O} (n^{-1}) \). With the change of variables \( z = x \xi_{T_2}^2 y \) we get
\[
\mathcal{L} = \mathbb{E} \left( e^{-r_2 T_2} \mathbb{L}_1 \left( \int_{\xi_{T_2}^2}^{z_{\Delta L_2}} \frac{\partial^2}{\partial x^2} \mathbb{E}^1_{L_1} \left( x \xi_{T_2}^2 y - x \xi_{T_2}^2 \right) x \xi_{T_2}^2 \right) dy \right)
= \mathbb{E} \left( e^{-r_2 T_2} \mathbb{L}_1 \left( \int_{\xi_{T_2}^2}^{z_{\Delta L_2}} \frac{\partial^2}{\partial x^2} \mathbb{E}^1_{L_1} \left( x \xi_{T_2}^2 y \right) \left( x \xi_{T_2}^2 y \right)^2 \left( \frac{y - 1}{y^2} \right) dy \right) \right).
\]

On the other hand, according to Lemma 2, there exists a constant \( Q \) depending only on the parameters \( \sigma_1, r_1, T \) such that
\[
\sup_{T \leq L_1 \leq T, z \geq 0} \left| z \frac{\partial^2}{\partial x^2} \mathbb{E}^1_{L_1} h \left( z \right) \right| \leq Q \chi_2 (h).
\]

This gives
\[
\mathbb{E} \left( \int_{\xi_{T_2}^2}^{z_{\Delta L_2}} \frac{\partial^2}{\partial x^2} \mathbb{E}^1_{L_1} \left( x \xi_{T_2}^2 y \right) \left( x \xi_{T_2}^2 y \right)^2 \left( \frac{y - 1}{y^2} \right) dy \right)
\leq Q \chi_2 (h) \mathbb{E} \left( \int_{\xi_{T_2}^2}^{z_{\Delta L_2}} \left( \frac{y - 1}{y^2} \right) dy \right).
\]

Hence,
\[
|\mathcal{L}| \leq Q \chi_2 (h) \mathbb{E} \left( \int_{\xi_{T_2}^2}^{z_{\Delta L_2}} \left( \ln \xi_{\Delta L_2}^2 - 1 + \left( \xi_{\Delta L_2}^2 \right)^{-1} \right) \right)
= Q \chi_2 (h) \mathbb{E} \left( \int_{\xi_{T_2}^2}^{z_{\Delta L_2}} \left( \Delta L_2 \left( r - \frac{1}{2} \sigma^2 \right) - 1 + \exp \left( -\Delta L_2 (r - \sigma^2) \right) \right) \right)
\leq \chi_2 (h) \mathcal{O} (1) \mathbb{E} |\Delta L_2|
= \chi_2 (h) \mathcal{O} (n^{-1}).
\]
4.6 Full discretization error

The following theorem is the main result of this chapter.

**Theorem 4.6.1** (Regime-switching discretization error). Assume that properties $A1-A4$ hold and that $h$ belongs to $K^{(2)}$. Let $\beta = 1/2$ if $h$ is discontinuous and $\beta = 1$ otherwise. Then,

$$E_t^h \psi (a, x) = E_t \psi (a, x) + \chi_2 (h) O \left( n^{-\beta} \right),$$

where the $O \left( n^{-\beta} \right)$ term is uniform in $h$ and $x$.

**Proof.** Here we want to show that

$$E \left( e^{\xi_1^n x_1^n - r_2^n x_2^n h} \xi_1^n \xi_2^n \right) = E \left( e^{\xi_1^n x_1^n - r_2^n x_2^n h} \xi_1^n \xi_2^n \right) + \chi_2 (h) O \left( n^{-\beta} \right).$$

First we write

$$E \left( e^{\xi_1^n x_1^n - r_2^n x_2^n h} \xi_1^n \xi_2^n \right) = E \left( 1_{L_1^n \geq T} \xi_1^{2,n} \xi_2^n \right) \left( \xi_1^n \xi_2^n \right) + E \left( 1_{L_2^n > T} \xi_1^{2,n} \xi_2^n \right) \left( \xi_1^n \xi_2^n \right).$$

Because $\xi^2$ and $\xi^1$ satisfy $A1-A4$, we obtain from Theorem 4.10.1 in the appendix that

$$\sup_{T \leq t \leq T} \sup_{z \geq 0} \left| E_t^h \xi^1_t h (z) - E_t^1 h (z) \right| = \chi_2 (h) O \left( n^{-\beta} \right),$$

$$\sup_{T \leq t \leq T} \sup_{z \geq 0} \left| E_t^h \xi^2_t h (z) - E_t^2 h (z) \right| = \chi_2 (h) O \left( n^{-\beta} \right).$$

As a result,

$$E \left( 1_{L_1^n \geq T} \xi_1^{2,n} \xi_2^n \right) \left( \xi_1^n \xi_2^n \right) = E \left( 1_{L_1^n \geq T} \xi_1^{2,n} \xi_2^n \right) \left( \xi_1^n \xi_2^n \right) + \chi_2 (h) O \left( n^{-\beta} \right),$$

$$E \left( 1_{L_2^n > T} \xi_1^{2,n} \xi_2^n \right) \left( \xi_1^n \xi_2^n \right) = E \left( 1_{L_2^n > T} \xi_1^{2,n} \xi_2^n \right) \left( \xi_1^n \xi_2^n \right) + \chi_2 (h) O \left( n^{-\beta} \right).$$

\[ \square \]
Note that for \( a = 1, 2 \) and for every \( t > 0 \), function \( x \mapsto \mathcal{E}_t^a h(x) \) belongs to \( C^\infty \cap K^2 \) and, furthermore, thanks to Lemma 2, there exists a constant \( Q \) such that
\[
\sup_{t \geq T} \chi_2 (\mathcal{E}_t^a h) \leq Q \chi_2 (h).
\]
Therefore, from Theorem 4.10.1 in the appendix,
\[
\sup_{0 \leq t_m < \frac{T}{2}} \sup_{z \geq 0} \left| \mathcal{E}_{t_m}^{1,n} \mathcal{E}_{T-t_m}^{2,n} h(z) - \mathcal{E}_{t_m}^{1,n} \mathcal{E}_{T-t_m}^{2,n} h(z) \right| = \chi_2 (h) \mathcal{O} (n^{-1}),
\]
\[
\sup_{0 \leq t_m < \frac{T}{2}} \sup_{z \geq 0} \left| \mathcal{E}_{t_m}^{2,n} \mathcal{E}_{T-t_m}^{1,n} h(z) - \mathcal{E}_{t_m}^{2,n} \mathcal{E}_{T-t_m}^{1,n} h(z) \right| = \chi_2 (h) \mathcal{O} (n^{-1}).
\]
As a result,
\[
\mathbb{E} \left( 1_{L_1T} e^{-r_1 L_1 h} \mathcal{E}_{L_2}^{1,n} (e^{\frac{1}{n} \mathcal{E}_{L_1}^{1,n} h}) (x) \right) = \mathbb{E} \left( 1_{L_1T} e^{-r_1 L_2 h} \mathcal{E}_{L_1}^{2,n} (e^{\frac{1}{n} \mathcal{E}_{L_1}^{2,n} h}) (x) \right) + \chi_2 (h) \mathcal{O} (n^{-\beta}),
\]
\[
\mathbb{E} \left( 1_{L_2T} e^{-r_2 L_1 h} \mathcal{E}_{L_1}^{1,n} (e^{\frac{1}{n} \mathcal{E}_{L_2}^{1,n} h}) (x) \right) = \mathbb{E} \left( 1_{L_1T} e^{-r_2 L_2 h} \mathcal{E}_{L_2}^{1,n} (e^{\frac{1}{n} \mathcal{E}_{L_2}^{2,n} h}) (x) \right) + \chi_2 (h) \mathcal{O} (n^{-\beta}).
\]

Hence,
\[
\mathbb{E} \left( e^{-r_1 L_1} e^{-r_2 L_2} h \left( x \xi_{L_1}^{1,n} \xi_{L_2}^{2,n} \right) \right) = \mathbb{E} \left( e^{-r_1 L_1} e^{-r_2 L_2} h \left( x \xi_{L_1}^{1,n} \xi_{L_2}^{2,n} \right) \right) + \chi_2 (h) \mathcal{O} (n^{-\beta}).
\]
But according to Proposition 4.5.1,
\[
\mathbb{E} \left( e^{-r_1 L_1} e^{-r_2 L_2} h \left( x \xi_{L_1}^{1,n} \xi_{L_2}^{2,n} \right) \right) = \mathbb{E} \left( e^{-r_1 L_1} e^{-r_2 L_2} h \left( x \xi_{L_1}^{1,n} \xi_{L_2}^{2,n} \right) \right) + \chi_2 (h) \mathcal{O} (n^{-1}).
\]
We have just proved that
\[
\mathbb{E} \left( e^{-r_1 L_1} e^{-r_2 L_2} h \left( x \xi_{L_1}^{1,n} \xi_{L_2}^{2,n} \right) \right) = \mathbb{E} \left( e^{-r_1 L_1} e^{-r_2 L_2} h \left( x \xi_{L_1}^{1,n} \xi_{L_2}^{2,n} \right) \right) + \chi_2 (h) \mathcal{O} (n^{-\beta}).
\]
as we wanted to.

\begin{proof}

Fix \(0 < T_1 < T_2 \leq T\), choose any arbitrary \(t \in [T_1, T_2]\), and let \(L_1, L_2, L_1^n, L_2^n\) be the usual occupation time random variables over the interval \([0, t]\). We have

\[
\mathcal{E}_t^n (I^\gamma) (a, x) = \mathbb{E} \left( e^{-r_1 L_1^n} e^{-r_2 L_2^n} \left( x^{e_1^1, n} \xi_{L_1}^{e_2^2, n} \right)^\gamma \right) \\
= x^\gamma \mathbb{E} \left( e^{-r_1 L_1^n} \left( \xi_{L_1}^{e_1^1, n} \right)^\gamma e^{-r_2 L_2^n} \left( \xi_{L_2}^{e_2^2, n} \right)^\gamma \right) \\
= x^\gamma \mathbb{E} \left( \mathcal{E}_L^{1, n} (I^\gamma) (1) \mathcal{E}_L^{2, n} (I^\gamma) (1) \right)
\]

and, invoking Theorem 4.10.1, we can continue with

\[
= x^\gamma \mathbb{E} \left( \mathcal{E}_L^{1, n} (I^\gamma) (1) \mathcal{E}_L^{2, n} (I^\gamma) (1) \right) + x^\gamma \mathcal{O} (n^{-1})
\]

where the \(\mathcal{O} (n^{-1})\) term is uniform in \(x\). Note that for any constant \(\mathcal{L} \geq 0\)

\[
\mathcal{E}_L^2 (I^\gamma) (1) = \exp (\beta_2 \mathcal{L}),
\]

where, for \(a = 1, 2, \)

\[
\beta_a = \frac{1}{2} (\gamma - 1) (\gamma \sigma_a^2 + 2r_a).
\]

Recall from (4.2.3) that \(L_a^n = L_a^n\) for \(a = 1, 2\) except on a set \(A\) with \(P (A^c) = \mathcal{O} (n^{-1})\).

Hence,

\[
\mathbb{E} \left( \mathcal{E}_L^{1, n} (I^\gamma) (1) \mathcal{E}_L^{2, n} (I^\gamma) (1) \right) = \mathbb{E} \left( \mathcal{E}_L^{1, n} (I^\gamma) (1) \mathcal{E}_L^{2, n} (I^\gamma) (1) \right) + \mathcal{O} (n^{-1}).
\]
Finally,

\[
E \left( e^{L_1^1 (I^n)} (1) e^{L_2^2 (I^n)} (1) \right) = E \left( e^{\beta_1 L_1^1 e^{\beta_2 L_2^2}} \right) \\
= E \left( e^{\beta_1 L_1 e^{\beta_2 L_2}} \right) + O(1) E \left( \left| L_1^1 - L_1 \right| \right) \\
= E \left( e^{L_1^1 (I^n)} (1) e^{L_2^2 (I^n)} (1) \right) \\
+ O(1) E \left( \left| L_1^1 - L_1 \right| \right).
\]

The result follows from Lemma 1.

4.8 Yuen and Yang trinomial model

Recall that the Yuen and Yang [87] trinomial model for regime-switching options is defined by

\[
\sigma > \max(\sigma_1, \sigma_2), \quad \Lambda_a = \frac{\sigma}{\sigma_a}, \\
u = e^{\sigma \sqrt{\Delta t}}, \quad m = 1, \quad d = e^{-\sigma \sqrt{\Delta t}}, \\
p_u^a = \frac{e^{rs_h} - e^{-\sigma \sqrt{h}} \left( 1 - \frac{1}{(\Lambda_a)^2} \right) \left( 1 - e^{-\sigma \sqrt{h}} \right)}{e^{\sigma \sqrt{h}} - e^{-\sigma \sqrt{h}}}, \\
p_d^a = \frac{e^{\sigma \sqrt{h}} - e^{rs_h} \left( 1 - \frac{1}{(\Lambda_a)^2} \right) \left( e^{\sigma \sqrt{h}} - 1 \right)}{e^{\sigma \sqrt{h}} - e^{-\sigma \sqrt{h}}}, \\
p_m^a = 1 - \frac{1}{(\Lambda_a)^2},
\]

where \( \Delta t = T/n \). Assume that at time \( t \in (\Delta t)N \), the regime-state is \( \alpha_i^n = a \) and the asset price is \( \Xi_i^{(n)} = x \). Let \( a' = 2 \) if \( a = 1 \) and \( a' = 1 \) otherwise. In the Yuen and Yang [87] model, six outcomes are possible at time \( t + \Delta t \): (a) with a probability of \( p_u^a \) \((\exp(-\lambda_a \Delta t))\), the regime remains in state \( a \) and the asset price jumps up to \( xa \); (b) with a probability of \( p_m^a \) \((\exp(-\lambda_a \Delta t))\), the regime remains in state \( a \) and the asset price also remains at level \( x \); (c) with a probability of \( p_d^a \) \((\exp(-\lambda_a \Delta t))\), the regime remains in state \( a \) and the asset price jumps down to \( xd \); (d) with a probability of \( p_u^a \) \((1 - \exp(-\lambda_a \Delta t))\), the regime switches to state \( a' \) and the asset price jumps up to \( xa \); (e) with a probability of \( p_m^a \) \((1 - \exp(-\lambda_a \Delta t))\), the regime switches to state \( a' \) and the asset price remains at level \( x \); (f) with a probability of \( p_d^a \) \((1 - \exp(-\lambda_a \Delta t))\), the regime switches to state \( a' \) and the asset price jumps down to \( xd \).
We show here that the trinomial trees $\xi^{a,n}$, $a = 1, 2$, defined by the Yuen and Yang model fall under assumptions A1-A4. To be specific, for $a = 1, 2$, let $X^a_n$ be the random variable which takes the value $\ln(u) = p\Delta t$ with probability $p^u_a$, the value $\ln(m) = 0$ with probability $p^m_a$, and the value $\ln(d) = -p\Delta t$ with probability $p^d_a$. Tedium but otherwise simple calculations that can easily be carried out by a computer algebra system give

$$E(X^a_n) = \left(r_a - \frac{1}{2}\sigma_a^2\right)\Delta t + O(\Delta t^2),$$  
(4.8.1)

$$\sqrt{Var(X^a_n)} = \sigma\sqrt{\Delta t} + O(\Delta t^{\frac{3}{2}}),$$  
(4.8.2)

and for any real constant $\gamma$,

$$E(\exp(\gamma X^a_n)) = \exp\left(\frac{1}{2}(\Delta t)\gamma(2r_a - \sigma_a^2 + \sigma_a^2\gamma)\right) + O(\Delta t^2),$$  
(4.8.3)

$$E\left(\exp\left(\gamma\sqrt{\frac{1}{\Delta t}X^a_n}\right)\right) \leq \exp(\vert\gamma\vert\sigma_a\Lambda_a) = O(1).$$  
(4.8.4)

Equations (4.8.1)-(4.8.4) precisely say that, for $a = 1, 2$, the trinomial tree $\xi^{a,n}$ approximating $\xi^a$ satisfies conditions A1-A4.

4.9 Numerical results

To illustrate the convergence behavior of security derivatives with piecewise smooth payoff functions in lattice methods for the two-state regime-switching model, we study two different kinds of options. We chose a European put option to represent the class of continuous payoff functions, and a digital put option, to represent the class of discontinuous payoff functions. The prices of these options are calculated using the Yuen and Yang trinomial model. We consider the case where the strike price is $E = 100$, and the time to maturity is $T = 1$. We choose the interest rate and the volatility to be $r_1 = 0.04, \sigma_1 = 0.25$ for regime 1, and $r_2 = 0.06, \sigma_2 = 0.35$ for regime 2. We set the jump intensity to be $\lambda_1 = \lambda_2 = 2$ for both regimes. In order to cover the three main cases, *In The Money* (ITM), *At The Money* (ATM), and *Out of The Money* (OTM), we select the value of the initial stock
price $S_0$ to be 90, 100, and 110. Numerical errors

$$\text{Err}_{\text{T}}^n h(a, x) = \mathcal{E}_{\text{T}} h(a, x) - \mathcal{E}_{\text{T}}^n h(a, x)$$

are calculated by subtracting the numerical approximations from the benchmark value obtained from our closed-form solution (4.3.2) of Section 4.3.2. Furthermore, we examine the value of $n^\beta \times \text{Err}_{\text{T}}^n h(a, x)$ to verify a convergence speed of order $O\left(n^{-\beta}\right)$. As we discussed in the previous section, European put options have a convergence of order $O\left(n^{-1}\right)$ while the convergence occurs at a speed of $O\left(n^{-1/2}\right)$ for digital options. Hence the values of $n \times \text{Err}_{\text{T}}^n h(a, x)$ and $n^{1/2} \times \text{Err}_{\text{T}}^n h(a, x)$ should be bounded for these two options.

<table>
<thead>
<tr>
<th>European put</th>
<th>Regime 1</th>
<th>Regime 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Price</td>
<td>Error</td>
</tr>
<tr>
<td>20</td>
<td>13.6292</td>
<td>-0.0383</td>
</tr>
<tr>
<td>100</td>
<td>13.6135</td>
<td>-0.0226</td>
</tr>
<tr>
<td>200</td>
<td>13.6019</td>
<td>-0.0111</td>
</tr>
<tr>
<td>500</td>
<td>13.5949</td>
<td>-0.0041</td>
</tr>
<tr>
<td>1000</td>
<td>13.5906</td>
<td>0.0003</td>
</tr>
<tr>
<td>2500</td>
<td>13.5917</td>
<td>-0.0008</td>
</tr>
<tr>
<td>5000</td>
<td>13.5911</td>
<td>-0.0002</td>
</tr>
</tbody>
</table>

Table 4.1: European put option with $S_0 = 90$

<table>
<thead>
<tr>
<th>European put</th>
<th>Regime 1</th>
<th>Regime 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Price</td>
<td>Error</td>
</tr>
<tr>
<td>20</td>
<td>9.0866</td>
<td>0.0129</td>
</tr>
<tr>
<td>100</td>
<td>9.0967</td>
<td>0.0028</td>
</tr>
<tr>
<td>200</td>
<td>9.0981</td>
<td>0.0014</td>
</tr>
<tr>
<td>500</td>
<td>9.0990</td>
<td>0.0006</td>
</tr>
<tr>
<td>1000</td>
<td>9.0993</td>
<td>0.0003</td>
</tr>
<tr>
<td>2500</td>
<td>9.0994</td>
<td>0.0001</td>
</tr>
<tr>
<td>5000</td>
<td>9.0995</td>
<td>0.0001</td>
</tr>
</tbody>
</table>

Table 4.2: European put option with $S_0 = 100$

Tables 4.1 to 4.3 collect specific values for the European put option in the ITM, ATM, and OTM cases. In each situation, we illustrate the error starting with either Regime 1 or Regime 2. We can see that when the spot price is \textit{in the money} or \textit{out of the money}, the Yuen and Yang trinomial model price oscillates around the true price. On the other hand, when the spot price is \textit{at the money}, the convergence is monotone and smooth.
4.9. NUMERICAL RESULTS

<table>
<thead>
<tr>
<th>N</th>
<th>Regime 1</th>
<th>Regime 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Error</td>
</tr>
<tr>
<td>20</td>
<td>5.9254</td>
<td>-0.0004</td>
</tr>
<tr>
<td>100</td>
<td>5.9444</td>
<td>-0.0195</td>
</tr>
<tr>
<td>200</td>
<td>5.9333</td>
<td>-0.0083</td>
</tr>
<tr>
<td>500</td>
<td>5.9265</td>
<td>-0.0016</td>
</tr>
<tr>
<td>1000</td>
<td>5.9266</td>
<td>-0.0010</td>
</tr>
<tr>
<td>2500</td>
<td>5.9259</td>
<td>-0.0002</td>
</tr>
<tr>
<td>5000</td>
<td>5.9252</td>
<td>-0.0002</td>
</tr>
</tbody>
</table>

Table 4.3: European put option with $S_0 = 110$

Figures 4.1 to 4.6 show in detail the relationship between $n \times \text{Err}^h \chi (a, x)$ and the number of time steps $n$, which varies from $n = 20$ to $n = 5,000$ with an increment of 1. The oscillations of $n \times \text{Err}^h \chi (a, x)$ in the ITM and OTM cases are unmistakable in Figures 4.1, 4.2, 4.5, and 4.6. Nonetheless, these oscillations are bounded, illustrating that the convergence is of order $O \left( n^{-1} \right)$. By contrast, while both curves in Figures 4.3 and 4.4 are also bounded, they display a smooth and monotone convergence. This numerically supports a convergence of order $O \left( n^{-1} \right)$.

![Figure 4.1: $S_0 = 90$ and $\alpha_0 = 1$](image1)

![Figure 4.2: $S_0 = 90$ and $\alpha_0 = 2$](image2)

The numerical results for the digital put option are similar to those of the European put option apart from the fact that the convergence speed is of order $O \left( n^{-1/2} \right)$. Tables 4.4 to 4.6 show the results for the digital put options when computed using the same set of parameters as for the put option, the same strike and the same maturity. From Tables 4.4 to 4.6 we can again draw the conclusion that in the ATM case the convergence is smooth and monotone while in the other two cases oscillations occur. The relationship between $\sqrt{n} \times \text{Err}^h \chi (a, x)$ and $n$ is displayed in Figures 4.7 to 4.12. Clearly, all curves in the plots
are bounded, illustrating that the speed of convergence of digital put options is of order $O(n^{-1/2})$. 

Figure 4.3: $S_0 = 100$ and $\alpha_0 = 1$

Figure 4.4: $S_0 = 100$ and $\alpha_0 = 2$

Figure 4.5: $S_0 = 110$ and $\alpha_0 = 1$

Figure 4.6: $S_0 = 110$ and $\alpha_0 = 2$
### 4.10 Auxiliary results

#### 4.10.1 Rate of convergence for Black-Scholes discretizations

The following result is from Leduc [49]. Recall that \( \text{Err}^n_h (x) := E^h (x) - E^n_h (x) \).

**Theorem 4.10.1** (Black-Scholes convergence rate for European options). Assume that

<table>
<thead>
<tr>
<th>Number of time steps</th>
<th>sqrt(N)*Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.27</td>
</tr>
<tr>
<td>500</td>
<td>0.271</td>
</tr>
<tr>
<td>1000</td>
<td>0.272</td>
</tr>
<tr>
<td>2000</td>
<td>0.273</td>
</tr>
<tr>
<td>5000</td>
<td>0.274</td>
</tr>
<tr>
<td>10000</td>
<td>0.275</td>
</tr>
<tr>
<td>25000</td>
<td>0.276</td>
</tr>
<tr>
<td>50000</td>
<td>0.277</td>
</tr>
</tbody>
</table>

---

**Table 4.4**: Digital put option with \( S_0 = 90 \)

<table>
<thead>
<tr>
<th>( \sqrt{N} \times \text{Error} )</th>
<th>( \text{Price} )</th>
<th>( \text{Error} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.6477</td>
<td>-0.0410</td>
</tr>
<tr>
<td>100</td>
<td>0.6052</td>
<td>0.0014</td>
</tr>
<tr>
<td>200</td>
<td>0.6040</td>
<td>0.0026</td>
</tr>
<tr>
<td>500</td>
<td>0.6033</td>
<td>0.0034</td>
</tr>
<tr>
<td>1000</td>
<td>0.5987</td>
<td>0.0080</td>
</tr>
<tr>
<td>2500</td>
<td>0.6054</td>
<td>0.0013</td>
</tr>
<tr>
<td>5000</td>
<td>0.6041</td>
<td>0.0026</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \sqrt{N} \times \text{Error} )</th>
<th>( \text{Price} )</th>
<th>( \text{Error} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.6355</td>
<td>-0.0389</td>
</tr>
<tr>
<td>100</td>
<td>0.5954</td>
<td>0.0012</td>
</tr>
<tr>
<td>200</td>
<td>0.5942</td>
<td>0.0024</td>
</tr>
<tr>
<td>500</td>
<td>0.5935</td>
<td>0.0031</td>
</tr>
<tr>
<td>1000</td>
<td>0.5893</td>
<td>0.0073</td>
</tr>
<tr>
<td>2500</td>
<td>0.5954</td>
<td>0.0012</td>
</tr>
<tr>
<td>5000</td>
<td>0.5942</td>
<td>0.0024</td>
</tr>
</tbody>
</table>

---

**Table 4.5**: Digital put option with \( S_0 = 100 \)

<table>
<thead>
<tr>
<th>( \sqrt{N} \times \text{Error} )</th>
<th>( \text{Price} )</th>
<th>( \text{Error} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.4093</td>
<td>0.0606</td>
</tr>
<tr>
<td>100</td>
<td>0.4423</td>
<td>0.0275</td>
</tr>
<tr>
<td>200</td>
<td>0.4503</td>
<td>0.0195</td>
</tr>
<tr>
<td>500</td>
<td>0.4575</td>
<td>0.0123</td>
</tr>
<tr>
<td>1000</td>
<td>0.4611</td>
<td>0.0087</td>
</tr>
<tr>
<td>2500</td>
<td>0.4643</td>
<td>0.0055</td>
</tr>
<tr>
<td>5000</td>
<td>0.4659</td>
<td>0.0039</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( \sqrt{N} \times \text{Error} )</th>
<th>( \text{Price} )</th>
<th>( \text{Error} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.4144</td>
<td>0.0567</td>
</tr>
<tr>
<td>100</td>
<td>0.4457</td>
<td>0.0253</td>
</tr>
<tr>
<td>200</td>
<td>0.4532</td>
<td>0.0179</td>
</tr>
<tr>
<td>500</td>
<td>0.4597</td>
<td>0.0113</td>
</tr>
<tr>
<td>1000</td>
<td>0.4631</td>
<td>0.0080</td>
</tr>
<tr>
<td>2500</td>
<td>0.4660</td>
<td>0.0051</td>
</tr>
<tr>
<td>5000</td>
<td>0.4675</td>
<td>0.0036</td>
</tr>
</tbody>
</table>

---

**Figure 4.9**: \( S_0 = 110 \) and \( \alpha_0 = 1 \)

**Figure 4.10**: \( S_0 = 110 \) and \( \alpha_0 = 2 \)
Table 4.6: Digital put option with $S_0 = 110$

<table>
<thead>
<tr>
<th>$N$</th>
<th>Regime 1</th>
<th>Regime 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Error</td>
</tr>
<tr>
<td>20</td>
<td>0.2943</td>
<td>0.0522</td>
</tr>
<tr>
<td>100</td>
<td>0.3355</td>
<td>0.0109</td>
</tr>
<tr>
<td>200</td>
<td>0.3367</td>
<td>0.0098</td>
</tr>
<tr>
<td>500</td>
<td>0.3374</td>
<td>0.0091</td>
</tr>
<tr>
<td>1000</td>
<td>0.3419</td>
<td>0.0045</td>
</tr>
<tr>
<td>2500</td>
<td>0.3456</td>
<td>0.0009</td>
</tr>
<tr>
<td>5000</td>
<td>0.3438</td>
<td>0.0026</td>
</tr>
</tbody>
</table>

Figure 4.11: $S_0 = 110$ and $\alpha_0 = 1$  

Figure 4.12: $S_0 = 110$ and $\alpha_0 = 2$

properties $A1$-$A4$ hold. Then, for every $0 < T_1 < T_2 \leq T$, and every $h$ in $K(2)$,

$$
\sup_{T_1 \leq t \leq T_2} \sup_{x \geq 0} |\text{Err}_t^h (x)| \leq \kappa_2 (h) O \left( n^{-\beta} \right),
$$

where the $O \left( n^{-\beta} \right)$ term is uniform in $h$, and where where $\beta = 1/2$ if $h$ is discontinuous and $\beta = 1$ otherwise. Furthermore, for every $x > 0$ and every real $\gamma$,

$$
\sup_{T_1 \leq t \leq T_2} |\text{Err}_t^h (I^\gamma)(x)| = x^\gamma O \left( n^{-1} \right),
$$

where the $O \left( n^{-1} \right)$ term is uniform in $x$. 
4.10.2 Occupation time discretization error

**Lemma 1.** Assume that $\mathcal{L}_a^n$ is either $L_a^n$, $\hat{L}_a^n$ or $\bar{L}_a^n$, that is $\mathcal{L}_a^n$ is either the default, snapshot or pseudo regime-state discretization defined in section 4.2.3. Then,

\[
\mathbb{E}(|L_a^n - L_a|) = \mathcal{O}(n^{-1}),
\]
\[
P\left(|L_a^n - L_a| > \frac{T}{4}\right) = \mathcal{O}(n^{-1}).
\]

**Proof.** Recall event $\mathcal{A}$ from (4.2.2) where the state process $\alpha_t$ can have at most one jump in any subinterval $(t_{m-1}, t_m]$. Then, for any $\omega \in \mathcal{A}$, $\hat{L}_a^n(\omega)$ differs from $L_a(\omega)$ only by aggregate errors of size less than or equal to $T/n$ near each jump. In other words,

\[
1_A \left| \hat{L}_a^n - L_a \right| \leq \left( \frac{T}{n} \right) N [0, T],
\]

for $a = 1, 2$, where $N [0, T]$ is the number of jumps of $\alpha_t$ in the interval $[0, T]$. Furthermore, as $\alpha_t^n$ changes state at time $t_m$ if and only if $\alpha_t$ changed stated during the interval $(t_{m-1}, t_m]$, it follows that

\[
1_A \left| L_a^n - \hat{L}_a^n \right| = 0.
\]

Thus,

\[
1_A \left| L_a^n - L_a \right| \leq \left( \frac{T}{n} \right) N [0, T].
\]

Then,

\[
\mathbb{E}(1_A |L_a^n - L_a|) \leq \left( \frac{T}{n} \right) N (0, T) \leq \left( \frac{T}{n} \right) \mathbb{E}(N(0, T)) = \mathcal{O}(n^{-1}),
\]

where $N(0, T)$ is a Poisson process with parameter $\lambda = \lambda_1 \lor \lambda_2$. Therefore,

\[
P\left(1_A |L_a^n - L_a| > \frac{T}{4}\right) \leq \frac{4\mathbb{E}(1_A |L_a^n - L_a|)}{T} = \mathcal{O}(n^{-1}).
\]

But, according to (4.2.3), $P(\mathcal{A}^c) = \mathcal{O}(n^{-1})$, and because $|L_a^n - L_a| \leq T$, it follows that

\[
\mathbb{E}(|L_a^n - L_a|) = \mathcal{O}(n^{-1}), \tag{4.10.1}
\]
\[
P\left(|L_a^n - L_a| > \frac{T}{4}\right) = \mathcal{O}(n^{-1}). \tag{4.10.2}
\]
If now $L^n_a$ is either $\hat{L}^n_a$ or $L^n_a$, then because these random variables are bounded by $T$, and because they are identical on $A$, which is a set of probability $1 - \mathcal{O}(n^{-1})$, it follows that (4.10.1) and (4.10.2) also hold when $L^n_a$ is replaced by $L^n_a$.

\[4.10.3 \quad \text{About the norm $\chi_2$ on $K^{(2)}$}\]

In the task of controlling the error of option values under approximations $\xi^n$ of geometric Brownian motions (GBM) $\xi$, the norm $\chi_2$ is quite practical because it disentangles the $\mathcal{O}$ terms from the payoff function $h$. This is particularly useful when considering options for which the payoff function is itself an option value of the form $\mathcal{E}_t h(x)$.

**Lemma 2.** Let $\xi$ be a GBM with drift $r$ and volatility $\sigma$. For every $0 \leq T_0 \leq T$ and every integer $\ell \geq 0$, there exists a constant $Q$ such that for every $h \in C^{(0)} \cap K^{(2)}$,

\[
\sup_{T_0 \leq t \leq T} \sup_{x \geq 0} \left( \left| \frac{\partial}{\partial t} \mathcal{E}_t h(x) \right| + \sum_{k=0}^{\ell} x^k \frac{\partial^k}{\partial x^k} \mathcal{E}_t h(x) \right) \leq Q \chi_2(h),
\]

\[
\sup_{T_0 \leq t \leq T} \chi_2(\mathcal{E}_t h) \leq Q \chi_2(h).
\]

**Proof.** Let $\phi$ be the density function of a standard normal random variable, and let

\[
\mathcal{E}_t^{(\ell)} h(x) = e^{-rt} \int_{-\infty}^{\infty} h(\sqrt{\sigma} z + (r - \frac{1}{2} \sigma^2) t) \phi^{(\ell)}(z) dz.
\]

According to [48],

\[
x^k \frac{\partial^k}{\partial x^k} \mathcal{E}_t h(x) = \sum_{\ell=1}^{k} \alpha_{\ell} t^{-\ell} \mathcal{E}_t^{(\ell)} h(x). \tag{4.10.3}
\]

Note that, for any given integer $k \geq 0$,

\[
\sup_{t, x \geq 0} \left| \mathcal{E}_t^{(k)} h(x) \right| \leq \|h\|_\infty \int_{-\infty}^{\infty} \left| \phi^{(k)}(z) \right| dz \leq O(1) \chi_2(h).
\]

It follows, in particular, that for any integer $k \geq 0$,

\[
\sup_{T_0 \leq t \leq T} \sup_{x \geq 0} \left| x^k \frac{\partial^k}{\partial x^k} \mathcal{E}_t h(x) \right| \leq O(1) \chi_2(h). \tag{4.10.4}
\]

The Black-Scholes equation

\[
\frac{\partial}{\partial t} \mathcal{E}_t h(x) = r \mathcal{E}_t h(x) - r x \frac{\partial}{\partial x} \mathcal{E}_t h(x) - \frac{1}{2} \sigma^2 x^2 \frac{\partial^2}{\partial x^2} \mathcal{E}_t h(x).
\]
guarantees that
\[
\sup_{T_0 \leq t \leq T} \sup_{x \geq 0} \left| \frac{\partial}{\partial t} \mathcal{E}_t h (x) \right| \leq O(1) \chi_2 (h) .
\] (4.10.5)

Now note that for every \( t \geq 0 \),
\[
TV \left( \mathcal{E}_t^{(t)} h \right) \leq TV (h) \int_{-\infty}^{\infty} \left| \varphi^{(j)} (z) \right| dz .
\]

Hence, it follows from (4.10.3) that for every integer \( k \geq 0 \),
\[
\sup_{T_0 \leq t \leq T} TV \left( x^k \frac{\partial^k}{\partial x^k} \mathcal{E}_t h (x) \right) = O(1) TV (h) \leq O(1) \chi_2 (h) .
\] (4.10.6)

Putting together (4.10.4), (4.10.5) and (4.10.6) completes the proof. \( \square \)
Part II: The connection between tree methods and finite-difference methods under stochastic volatility models

As we mentioned in Section 2.3.3, the trinomial tree methods are essentially a special case of the explicit finite-difference method under the BS model. It is reasonable to assume that the equivalence of the two methods holds for all the other models extended from the BS model. However, the equivalence of the same type as the BS case has not been discovered yet for stochastic volatility models. In this part, we extend the connection to the regime-switching model and Heston’s stochastic volatility model for the first time.

The connection between the Yuen and Yang’s trinomial tree method and explicit finite-difference methods has been discussed by Ma et al. [61]. However, in the so-called “equivalence” they present, a second-order error term appears as the difference of the two recursive formulas, in contrast to the perfect match in the BS case. In fact, it is impossible for the Yuen and Yang’s tree to be equivalent to the explicit finite-difference method under the regime-switching model due to too many nodal values being required for each time step. In fact, none of the existing tree methods for the regime-switching model is eligible to build the relationship with the explicit finite-difference method. Thus in Chapter 5, a new trinomial tree approach is presented. We made a different assumption about the recursive formula, which results in a significant reduction of the number of required nodal values. The new tree is proven analytically to have identical coefficients with the corresponding explicit finite-difference schemes. Numerical experiments further verify the results and show that the new trinomial tree outperforms the Yuen and Yang’s tree in pricing vanilla European options.

Then in Chapter 6, the equivalence is set up for Heston’s stochastic volatility model. Similar to the regime-switching case, a new tree approach for the Heston model is developed. Though the process is quite sophisticated, the structure of the new tree method appears to be quite simple. It also captures the key advantage of tree methods, which is the financial interpretation. The equivalence of the new approach and the explicit finite-difference has shown analytically and numerically. The efficiency of the new method is also tested in the numerical performance.
Chapter 5

A new trinomial tree method for regime-switching models

5.1 Introduction

Since its introduction by Cox et al. [19], the binomial tree method becomes one of the most popular numerical methods in options pricing. This is mainly because tree-based methods are characterized by their simplicity as well as clear financial interpretations associated with each step involved in these approaches. The essence of the method is to use discrete jumps, either up or down, to approximate the continuous diffusion process. Then the lattice is constructed as a tree with two branches from each seed node. Boyle [8] then presents a trinomial tree method with an extra branch with the meaning of the underlying asset remains the same over the time period at each seed node. Due to the presence of the additional branch, the trinomial tree has another degree of freedom since the move spacing can be set independently from the move timing [76]. Later Brennan and Schwarz [11] further investigates the lattice-based model and presents the relationship between an explicit finite difference method and a trinomial tree method under the Black-Scholes model. The authors declare that the standard explicit finite difference approximation corresponds to the three-point jump process, which is the trinomial tree approximation, while the more complex implicit finite difference approximation corresponds to a generalized jump process with infinitely many branches from each seed node. After that, Rubinstein [76] points out that the Kamrad-Ritchken trinomial tree method [45] can have an identical scheme as the standard explicit finite difference method when a logarithmic transformation is ap-
plied. Since then, it is well-known that trinomial tree method and explicit finite difference methods are equivalent under the Black-Scholes economy.

However, as broadly used in the industry, the Black-Scholes model is constantly reported to produce biased values that are far from the market prices. The flaw results from the constant volatility assumption, which fails to describe the random feature of volatilities in the practical market. Hence many practical models are presented with volatility varying along with time and price of underlying assets. Among these models, regime-switching models, as simple extensions from the BS model, have received a lot of attention since the pioneering work by Hamilton [35]. The model assumes its drift and volatility terms to depend on a continuous-time Markov chain to capture the structural changes of the market, for instance, economic expansions and recessions. Due to the random feature of Markov chains, the model can be viewed as a special case of the stochastic volatility models, while it preserves certain degrees of the simplicity of the BS model. There have been empirical studies [4, 22, 34, 70] showing the advantage of the model in pricing financial derivatives for certain markets.

Pricing options under regime-switching models usually require numerical computations. Among various numerical approaches, tree approaches for regime-switching models have been developed during the past two decades. Bollen [6] first proposes a multinomial tree method for a two-state regime-switching model. In order to make the tree branches recombine, he applies the binomial tree method [19] for each individual regime and then adds a joint middle branch shared by both trees. Thus the two binomial trees become trinomial trees which have more flexibility. By adjusting the tree parameters, the step sizes of the outer branches become twice the size of the inner branches. Then the tree recombines properly and grows only linearly as the number of time step increases. Liu [53] adopts the method and extends it to a general multi-state regime-switching model. However, the method loses its efficiency as the number of regime increases since each one more regime causes two more branches to be grown from a seed node. In the same year, Yuen and Yang [87] modifies the idea of recombining from Bollen [6] and presents a trinomial tree method for a multi-state regime-switching model. Instead of letting the outer branch be twice the length of the inner counterpart, the authors adjust them to be of the same size so that they overlap each other. Then the number of successor branches is fixed as three and becomes independent of the number of regimes. Their method is considered
as the most efficient tree method for regime-switching models.

Since a regime-switching model can be viewed as a composition of some simple GBMs, each of which has a well-established formal equivalence, the definition of which has been given in [76], between tree methods and standard explicit finite difference methods when these methods are adopted to numerically compute prices of a financial derivative, there should exist such an equivalence between tree methods and explicit finite difference methods for regime-switching models as well. The link between the two approaches is recently explored by Ma et al. [61], which presents a “second-order equivalence” of the Yuen and Yang’s trinomial tree method with an explicit finite difference method. Unfortunately, unlike the GBM case, the relation they propose is not a true equivalence because of the following two reasons. First, there is a second-order error term of $O(\Delta t^2)$ when comparing the recursive formulas of the two methods. Second, their explicit finite difference scheme has an extra perturbed term, which is apparently not standard. Thus the challenge of discovering the true equivalence remains even though their publication has presented a very close connection between the two methods. Establishing such an equivalence is the core of this chapter.

In this chapter, we present a new trinomial tree method that has a property of being formally equivalent to a standard explicit finite difference method. To achieve this, we start with an assumption that is different from Yuen and Yang’s [87], under which both methods have exactly the same coefficients at each nodal point. The evidence of our equivalence having no error terms involved and the numerical performance of our new tree approach is also shown in the chapter.

The rest of the chapter is organized as follows. Priori knowledge and notations of the model are presented in Section 5.2. Our new trinomial tree method is introduced in Section 5.3. Equivalence proof between our trinomial method and the explicit finite difference method is presented in Section 5.4. The convergence rate is reaffirmed in Section 5.5 A two-state case is demonstrated as an example in Section 5.6. Numerical performance is shown in Section 5.7, followed by conclusions in Section 5.8.
5.2 Priori knowledge and notations

In the regime-switching world where the drift rate and the volatility are shifted between different states, the fluctuation of an asset is assumed to follow the stochastic differential equation

\[ dS_t = \mu(X_t)S_t dt + \sigma(X_t)S_t dW_t, \]

where \( \mu \) and \( \sigma \) are the drift and volatility terms of the price dynamic of the underlying asset, \( X_t \) is a continuous time Markov chain with \( K \) states and is independent of the standard Brownian motion \( W_t \). They are based on the probability triple \( \{ \Omega, \mathcal{F}, \mathbb{P} \} \) where \( \mathbb{P} \) is the physical measure. For each state, the drift rate and the volatility are assumed to be constant and distinct, denoted by

\[
\begin{align*}
\mu(X_t) &= \begin{cases}
\mu_1, & X_t = 1, \\
\mu_2, & X_t = 2, \\
\vdots & \\
\mu_K, & X_t = K,
\end{cases} \\
\sigma(X_t) &= \begin{cases}
\sigma_1, & X_t = 1, \\
\sigma_2, & X_t = 2, \\
\vdots & \\
\sigma_K, & X_t = K.
\end{cases}
\end{align*}
\]

The generator of the Markov chain is

\[
Q = \begin{pmatrix}
\lambda_{11} & \lambda_{12} & \cdots & \lambda_{1K} \\
\lambda_{21} & \lambda_{22} & \cdots & \lambda_{2K} \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_{K1} & \lambda_{K2} & \cdots & \lambda_{KK}
\end{pmatrix},
\]

(5.2.1)

The sum of each row of the generator matrix is zero.

Since another risk source \( X_t \) is introduced, the market becomes incomplete. As a result, there is no unique martingale measure. Again in this chapter, we select the martingale measure presented by Elliott et al. [21] and assume the interest rates are the same for all regimes. Under the risk-neutral measure, the price of a financial derivative (which we still denote by \( V_j(S_t, t) \) as the previous chapter) with initial state \( j \) is obtained by

\[ V_j(S_t, t) = e^{-r(T-t)}\mathbb{E}(p(S_T) | \mathcal{F}_t, X_t = j), \]

(5.2.2)
where function \( p(\cdot) \) is the payoff function. For European options, substituting the payoff function into (5.2.1) and applying Itô’s formula, the governing system of partial differential equations is given by (see [15])

\[
\frac{\partial V_j}{\partial t} + \frac{1}{2} \sigma_j^2 S^2 \frac{\partial^2 V_j}{\partial S^2} + rS \frac{\partial V_j}{\partial S} - rV_j = \sum_{i=1, i \neq j}^{K} \lambda_{ji}(V_j - V_i), \quad j = 1, 2, \ldots, K. \tag{5.2.3}
\]

Since the introduction of the initial state, (5.2.3) is a PDE system with a total of \( K \) equations whose solutions have to be solved simultaneously.

### 5.3 New trinomial tree approach

Since the regime-switching model can be viewed as a composition of some simple GBMs, we first adopt Kamrad-Ritchken parameters [45] for each GBM with constant parameters. Assume \( U, M \) and \( D \) are the ratio of the stock price moving upward, stable and downward, respectively. According to [45], the three parameters are given by

\[
U = e^{\sigma \sqrt{\Delta t}}, \\
M = 1, \\
D = e^{-\sigma \sqrt{\Delta t}},
\]

where

\[
\sigma = \rho_j \sigma_j, \quad j = 1, 2, \ldots, K. \tag{5.3.1}
\]

Here \( \rho_j \) is a parameter such that the jump sizes are equal among all regimes. Thus the tree recombines and grows only linearly as the number of time steps increases. However, since the means and variances of the stock dynamic are different among regimes, the corresponding probabilities under the risk-neutral measure must not be the same. Denote the probabilities of moving up, middle and down in the regime \( j \) by \( \pi^U_j, \pi^M_j, \pi^D_j \). Then by matching the local mean and variance of the underlying distribution for each regime,
we should have

\[
\begin{align*}
\pi_j^U &= \frac{1}{2\rho_j^2} + \frac{(r - \frac{1}{2}\sigma_j^2)\sqrt{\Delta t}}{2\rho_j\sigma_j}, \\
\pi_j^M &= 1 - \frac{1}{\rho_j^2}, \\
\pi_j^D &= \frac{1}{2\rho_j^2} - \frac{(r - \frac{1}{2}\sigma_j^2)\sqrt{\Delta t}}{2\rho_j\sigma_j}.
\end{align*}
\]

Here the ratio \(\rho_j\) must be greater than 1 so that \(\pi_j^M > 0\). This means the value of \(\sigma\) must satisfy

\[
\sigma > \max_{j=1,2,...,K} \{\sigma_j\}.
\]

As suggested by Yuen and Yang [87], a good choice of \(\sigma\) is given by

\[
\sigma = \max_{j=1,2,...,K} \{\sigma_j\} + (\sqrt{1.5} - 1)\bar{\sigma}.
\]

It provides good results based on the literature of the binomial and trinomial tree models. Here the \(\bar{\sigma}\) could be either the arithmetic mean or the geometric mean of the volatilities of all regimes.

After setting up trinomial trees for all regimes, we need to combine them together via a regime-switching Markov chain. At each time step, the combined tree must account for two possible types of changes: the change of market conditions (regimes) and the change of stock prices within a regime. It’s natural to raise questions regarding the order of these changes, or whether they should occur simultaneously.

In Yuen and Yang’s tree, the authors assume that the two changes take place simultaneously. In particular, at time \(t = t_n\), a stock in regime \(j\) with price \(S = S_m\) is allowed to jump to \(S = S_{m+1}, S_m, S_{m-1}\) in all regimes at the \(t = t_{n+1}\). Thus their recursive formula requires a total of \(3K\) nodal values when computing the nodal value of the previous time step.

Our approach makes the following assumption instead. Within one time step, the stock price can either stay in the same regime and “diffuse” along the trinomial tree in the usual way, or switch regimes but without diffusing. For instance, a node \((S_m, t_n, j)\) is allowed to jump to \((S_{m+1}, t_n, j), (S_{m-1}, t_n, j)\) or \((S_m, t_n, i)\), where \(i\) means all regimes including \(j\). Hence only \(K + 2\) nodal values are required in computation at each time step. As a
result, our approach has the advantage of producing identical recursive formulas for the
two numerical methods, establishing true equivalence between them.

This can be financially interpreted as the change of market conditions occurring prior
to the change of stock prices. Under our assumption, stock prices start to randomly diffuse
only after a new regime has been switched into. Our assumption is tested and proven to
be feasible as the results provided by our tree approach is robust and accurate.

Then the main idea is as follows. Let $T$ be the time to expiry, $N$ be the total number
of time steps. Then $\Delta t = T/N$. Since the tree recombines, only $2t + 1$ nodes appear in
the time step $t$. At each time step, $K$ variables have to be solved. Here we denote option
price with initial regime $j$ at time step $t_n$ by $V_{m+1,j}^n$. Then this price can be obtained
recursively using the following equations:

$$V_{m,1}^n = e^{-r\Delta t} \left[ \pi_1 U V_{m+1,1}^{n+1} + (\pi_1^M + \lambda_{11} \Delta t)V_{m,1}^{n+1} + \pi_1 D V_{m-1,1}^{n+1} + \sum_{i=2}^{K} \lambda_{1i} \Delta t V_{m,i}^{n+1} \right],$$

$$V_{m,2}^n = e^{-r\Delta t} \left[ \pi_2 U V_{m+1,2}^{n+1} + (\pi_2^M + \lambda_{22} \Delta t)V_{m,2}^{n+1} + \pi_2 D V_{m-1,2}^{n+1} + \sum_{i=3, i \neq 2}^{K} \lambda_{2i} \Delta t V_{m,i}^{n+1} \right],$$

$$\vdots$$

$$V_{m,K}^n = e^{-r\Delta t} \left[ \pi_K U V_{m+1,K}^{n+1} + (\pi_K^M + \lambda_{KK} \Delta t)V_{m,K}^{n+1} + \pi_K D V_{m-1,K}^{n+1} + \sum_{i=1}^{K-1} \lambda_{Ki} \Delta t V_{m,i}^{n+1} \right].$$

Clearly, to compute the value of one seed node at time step $t = t_n$, we require $K + 2$ seed
nodes to be involved at time step $t = t_{n+1}$, in contrast to $3K$ seed nodes in Yuen and
Yang’s recursive equations. The evidence of our tree method being formally equivalent to
an explicit finite difference is shown in the next section.

### 5.4 Equivalence to explicit finite difference

As we mentioned earlier, the equivalence of a trinomial tree method with a standard
explicit finite difference method for a GBM has been well-established. In [76], the au-

[76] thor shows that under a logarithmic transformation, the standard explicit finite difference

[45] method has the same coefficients as the Kamrad-Ritchken trinomial tree method [45] at
each time step. In particular, the recursive formulas of these two approaches are ident-
tical at each nodal point. In this section, we show that our tree has the same recursive
formula with the standard explicit finite difference method under a general multi-state regime-switching framework.

Before starting, we first introduce the dimensionless variable transformation. Let

\[ x = \log(S), \quad Y_j = V_j e^{-rt}. \]

Then PDE (5.2.3) becomes

\[ \frac{\partial Y_j}{\partial t} + \frac{1}{2} \sigma_j^2 \frac{\partial^2 Y_j}{\partial x^2} + (r - \frac{1}{2} \sigma_j^2) \frac{\partial Y_j}{\partial x} = \sum_{i=1}^{K} \lambda_{ji} (Y_j - Y_i), \quad j = 1, 2, ..., K. \]  

(5.4.1)

Applying the forward Euler scheme to (5.4.1) gives

\[ Y_{m,j}^n = \alpha_j Y_{m+1,j}^{n+1} + (\beta_j - \sum_{i=1}^{K} \lambda_{ji} \Delta t) Y_{m,j}^{n+1} + \beta_j Y_{m-1,j}^{n+1} + \sum_{i=1}^{K} \lambda_{ji} \Delta t Y_{m,i}^{n+1}, \]  

(5.4.2)

where

\[ \alpha_j = \frac{\sigma_j^2 \Delta t}{2 \Delta x^2} + \frac{(r - \frac{1}{2} \sigma_j^2) \Delta t}{2 \Delta x}, \]

\[ \beta_j = 1 - \frac{\sigma_j^2 \Delta t}{\Delta x^2}, \]

\[ \gamma_j = \frac{\sigma_j^2 \Delta t}{2 \Delta x^2} - \frac{(r - \frac{1}{2} \sigma_j^2) \Delta t}{2 \Delta x}. \]

The recursive formula of the explicit finite difference requires a total of \( K + 2 \) nodal values to calculate the value of previous time step, which is the same as our tree structure. It’s necessary to have the same number of required nodal values in building a connection between two numerical methods. This is a key ingredient to establish a true equivalence between the two methods. Recall that Yuen and Yang’s tree requires more nodal values at each time step than the finite difference method, which leads to the presence of the error term \( O(\Delta t^2) \).

Now we have the recursive formula for the finite difference, the next step is to compare the coefficients of (5.4.2) with the formula of our tree approach. According to our parametrizations,

\[ S_{m+1}^n = US_m^n = e^{\sigma \sqrt{\Delta t}} S. \]
Thus, by substituting $x = \log(S)$, the following relation can be obtained

$$\Delta x = \sigma \sqrt{\Delta t}. \quad (5.4.3)$$

We see that the coefficients of the finite difference with transformation (5.4.3) become the same as the stock jump probabilities of the trinomial trees:

$$\alpha_j = \frac{\sigma_j^2 \Delta t}{2 \rho_j^2 \sigma^2 \Delta t} + \frac{(r - \frac{1}{2} \sigma^2) \Delta t}{2 \rho_j \sigma \sqrt{\Delta t}} = \frac{1}{2 \rho_j^2} + \frac{(r - \frac{1}{2} \sigma^2) \sqrt{\Delta t}}{2 \rho_j \sigma} = \pi^U_j,$$

$$\beta_j = 1 - \frac{\sigma_j^2 \Delta t}{\rho_j^2 \sigma^2 \Delta t} = 1 - \frac{1}{\rho_j^2} = \pi^M_j,$$

$$\gamma_j = \frac{\sigma_j^2 \Delta t}{2 \rho_j^2 \sigma^2 \Delta t} - \frac{(r - \frac{1}{2} \sigma^2) \Delta t}{2 \rho_j \sigma \sqrt{\Delta t}} = \frac{1}{2 \rho_j^2} - \frac{(r - \frac{1}{2} \sigma^2) \sqrt{\Delta t}}{2 \rho_j \sigma} = \pi^D_j,$$

Finally, substituting into (4.2) and using $V_j = Y_j e^{rt}$,

$$V^n_{m,j} = e^{-r \Delta t} \left( \alpha_j V^{n+1}_{m+1,j} + (\beta_j - \sum_{i=1 \atop i \neq j}^K \lambda_{ji} \Delta t) V^{n+1}_{m,j} + \gamma_j V^{n+1}_{m-1,j} + \sum_{i=1 \atop i \neq j}^K \lambda_{ji} \Delta t V^{n+1}_{m,i} \right)$$

$$= e^{-r \Delta t} \left( \pi^M_j V^{n+1}_{m+1,j} + (\pi^M_j - \sum_{i=1 \atop i \neq j}^K \lambda_{ji} \Delta t) V^{n+1}_{m,j} + \pi^D_j V^{n+1}_{m-1,j} + \sum_{i=1 \atop i \neq j}^K \lambda_{ji} \Delta t V^{n+1}_{m,i} \right).$$

Thus both methods have the same recursive formula and the equivalence has been proved. It should be remarked that this equivalence includes no error terms in contrast to the connection in Ma et al. [61]. Here, the coefficients of the two recursive formulas are identical. We consolidate our result by numerically testing both methods in Section 5.7.

Since explicit schemes are not conditionally stable, the stability condition is vital for the practical use of the numerical technique. In terms of our new tree approach, the following condition should be satisfied in order to maintain the stability of the method:

$$\Delta t < \min_{j=1,2,\ldots,K} \left( \frac{\pi^M_j}{\sum_{i=1 \atop i \neq j}^K \lambda_{ji}} \right).$$

This can be easily verified by the Von-Neumann’s stability analysis.
5.5 Convergence rate

Since our tree method can be viewed as a special case of a standard explicit finite difference method, it should theoretically have first-order convergence. We reaffirm this result (without relying on the equivalence) in the following theorem, which formally shows that the convergence rate of our tree method is indeed of $O(\Delta t)$.

**Theorem 5.5.1.** Let $V(S_m, t_n, j)$ denote the exact value of the option price given $S_m = m\Delta S$, $t_n = n\Delta t$ and state $j$. Further let $V^n_{m,j}$ denote the corresponding approximation from our trinomial tree method. Further denote the error of the trinomial tree method by

$$
\varepsilon^n_j(S_m) = V(S_m, t_n, j) - V^n_{m,j}, \quad j = 1, 2.
$$

(5.5.1)

and define the infinity norm at time $t_n$ by

$$
||\varepsilon^n_j||_{\infty} = \max_{-M \leq m \leq M} |\varepsilon^n_j(S_m)|, \quad j = 1, 2.
$$

Then the convergence rate can be estimated by

$$
||\varepsilon^n_j||_{\infty} = |O(\Delta t)|, \quad \text{for } n = 0, 1, ..., N - 1, \quad j = 1, 2.
$$

The proof of the theorem is quite long but straightforward. Thus it is put in the Appendix B.1.

5.6 Two-state case

In this section, we show an example of our trinomial tree method in the two-state case, in which the generator matrix is simplified to

$$
Q = \begin{pmatrix}
-\lambda_{12} & \lambda_{12} \\
\lambda_{21} & -\lambda_{21}
\end{pmatrix}.
$$

The corresponding transition probability matrix then can be approximated as

$$
P = \begin{pmatrix}
p_{11} & p_{12} \\
p_{21} & p_{22}
\end{pmatrix} = \begin{pmatrix}
1 - \lambda_{12}\Delta t + O((\Delta t)^2) & \lambda_{12}\Delta t + O((\Delta t)^2) \\
\lambda_{21}\Delta t + O((\Delta t)^2) & 1 - \lambda_{21}\Delta t + O((\Delta t)^2)
\end{pmatrix}.
$$
Since $K = 2$, the PDE problem contains only two coupled PDEs and (5.2.3) becomes

$$
\frac{\partial V_1}{\partial t} + \frac{1}{2} \sigma_1^2 S^2 \frac{\partial^2 V_1}{\partial S^2} + rS \frac{\partial V_1}{\partial S} - rV_1 = \lambda_{12} (V_1 - V_2),
$$

$$
\frac{\partial V_2}{\partial t} + \frac{1}{2} \sigma_2^2 S^2 \frac{\partial^2 V_2}{\partial S^2} + rS \frac{\partial V_2}{\partial S} - rV_2 = \lambda_{21} (V_2 - V_1).
$$

The graphical illustration of a two-state tree-based approximation is given in Figure 5.1.

![Graphical illustration for a two-state regime-switching model](image)

**Figure 5.1:** Graphical illustration for a two-state regime-switching model

From the figure, we know that there are four possibilities of the seed node moving toward next time step, moving up, middle, down and switching to the other regime. The recursive equation is simplified to

$$
V_{m+1,1}^n = e^{-r \Delta t} \left[ \pi_1^U V_{m+1,1}^{n+1} + (\pi_1^M - \lambda_{12} \Delta t) V_{m,1}^{n+1} + \pi_1^D V_{m-1,1}^{n+1} + \lambda_{12} \Delta t V_{m-2}^{n+1} \right]
$$

$$
V_{m+1,2}^n = e^{-r \Delta t} \left[ \pi_2^U V_{m+1,2}^{n+1} + (\pi_2^M - \lambda_{21} \Delta t) V_{m,2}^{n+1} + \pi_2^D V_{m-1,2}^{n+1} + \lambda_{21} \Delta t V_{m-2}^{n+1} \right]
$$

We split the probabilities of the stable branch, $\pi_1^m$ (respectively, $\pi_2^m$), into two parts, staying in the same regime, $\pi_1^m - \lambda_{12} \Delta t$ (respectively, $\pi_2^m - \lambda_{21} \Delta t$), and switching into another regime, $\lambda_{12} \Delta t$ (respectively, $\lambda_{21} \Delta t$). Thus the price of the option at each time step is obtained by a combination of four points, in contrast to the trinomial method in Yuen and Yang [87], which adopts a combination of six points. Numerical performance is provided in the next section.

## 5.7 Numerical performance

In this section, we present the numerical performance of our new trinomial tree method for options pricing under the two-state regime-switching model. Benchmark results are obtained by the closed-form solution for European options from Zhu et al. [91]. The evidence of the equivalence between our new trinomial tree method and the explicit finite
difference method is also given in this section.

Table 5.1: AS is the analytical solution from Zhu et al. [91]. Ab Err is the absolute error between the price obtained from the tree model and the analytical solution. Ratio is the ratio of the current absolute error and the previous one.

<table>
<thead>
<tr>
<th>( N )</th>
<th>Regime 1</th>
<th>Regime 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Price</td>
<td>Ab Err</td>
</tr>
<tr>
<td>20</td>
<td>0.9912</td>
<td>0.0601</td>
</tr>
<tr>
<td>40</td>
<td>1.0227</td>
<td>0.0286</td>
</tr>
<tr>
<td>80</td>
<td>1.0372</td>
<td>0.0140</td>
</tr>
<tr>
<td>160</td>
<td>1.0443</td>
<td>0.0070</td>
</tr>
<tr>
<td>320</td>
<td>1.0478</td>
<td>0.0035</td>
</tr>
<tr>
<td>640</td>
<td>1.0495</td>
<td>0.0017</td>
</tr>
<tr>
<td>1280</td>
<td>1.0504</td>
<td>0.0009</td>
</tr>
<tr>
<td>2560</td>
<td>1.0508</td>
<td>0.0004</td>
</tr>
<tr>
<td>5120</td>
<td>1.0510</td>
<td>0.0002</td>
</tr>
<tr>
<td>AS</td>
<td>1.0512</td>
<td>1.7028</td>
</tr>
</tbody>
</table>

We start with the applicability of our new tree method. European put options are calculated with a certain set of parameters with various numbers of time steps. For an applicable algorithm, the results should converge to the real value as the number of partitions goes to infinity. In Table 5.1, the initial stock price \( S_0 \) and the strike price \( E \) are set to be $40. The life of the European put option \( T \) is 3 months. The transition rates \( \lambda_{12} \) and \( \lambda_{21} \) are 1 and 0.5, respectively. The corresponding volatilities in regime 1 and regime 2 are 0.15 and 0.25. The number of the time steps are chosen to be 20, 40, 80, 160, ..., 5120. Table 5.1 shows that the errors from our trinomial tree model decrease as the number of time steps goes to zero, which implies that the method is applicable. Further, it should be pointed out that every time we double the amount of the time steps, the absolute error reduces approximately by a factor of 2 according to the ratio column. This indicates that the convergence rate of our tree model is \( O(\Delta t) \).

Then we investigate the equivalence relation between our trinomial tree and the explicit finite difference. In the implementation of the explicit finite difference, we vary the number of time steps, which in turn determines the number of space partitions. The Initial stock price is placed in the centre of the space partition. If the two methods are equivalent, the magnitude of the difference between the two methods should be negligible. The results of the numerical experiment are shown in Table 5.2. According to the table, differences between the two methods are smaller than \( 10^{-12} \) at all time, which verifies the equivalence of the two approaches. The error is due to rounding from relation (5.4.3). Results from
Table 5.2: In this table, the initial stock price and the strike price are both $100. The risk-free interest rate is set to be 0.05. The volatilities of the stock in state 1 and state 2 are 0.15 and 0.25, respectively. The life of the European put option is one year. N is the number of the time steps.

<table>
<thead>
<tr>
<th>N</th>
<th>Regime 1</th>
<th>Regime 2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Explicit</td>
<td>Tree</td>
<td>Difference</td>
</tr>
<tr>
<td>100</td>
<td>4.9673</td>
<td>4.9673</td>
<td>0.0586 x 10^{-12}</td>
</tr>
<tr>
<td>200</td>
<td>4.9899</td>
<td>4.9899</td>
<td>0.0266 x 10^{-12}</td>
</tr>
<tr>
<td>500</td>
<td>5.0034</td>
<td>5.0034</td>
<td>0.0613 x 10^{-12}</td>
</tr>
<tr>
<td>1000</td>
<td>5.0079</td>
<td>5.0079</td>
<td>0.0648 x 10^{-12}</td>
</tr>
<tr>
<td>2000</td>
<td>5.0101</td>
<td>5.0101</td>
<td>0.3118 x 10^{-12}</td>
</tr>
<tr>
<td>5000</td>
<td>5.0115</td>
<td>5.0115</td>
<td>0.2061 x 10^{-12}</td>
</tr>
<tr>
<td>10000</td>
<td>5.0119</td>
<td>5.0119</td>
<td>0.4334 x 10^{-12}</td>
</tr>
</tbody>
</table>

Table 5.3: In this table, the strike price is $40. The risk-free interest rate is set to be 0.05. The volatilities of the stock in state 1 and state 2 are 0.15 and 0.25, respectively. The life of the European put option is one year. $S_0$ is the initial stock price and is selected from $32$ to $48$.

<table>
<thead>
<tr>
<th>$S_0$</th>
<th>Regime 1</th>
<th>Regime 2</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Z-tree</td>
<td>Y-tree</td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>8.0367</td>
<td>8.0372</td>
<td>8.0369</td>
</tr>
<tr>
<td>34</td>
<td>6.4018</td>
<td>6.4024</td>
<td>6.4014</td>
</tr>
<tr>
<td>36</td>
<td>4.9562</td>
<td>4.9568</td>
<td>4.9558</td>
</tr>
<tr>
<td>38</td>
<td>3.7306</td>
<td>3.7313</td>
<td>3.7304</td>
</tr>
<tr>
<td>40</td>
<td>2.7334</td>
<td>2.7341</td>
<td>2.7352</td>
</tr>
<tr>
<td>42</td>
<td>1.9600</td>
<td>1.9607</td>
<td>1.9596</td>
</tr>
<tr>
<td>44</td>
<td>1.3770</td>
<td>1.3777</td>
<td>1.3772</td>
</tr>
<tr>
<td>46</td>
<td>0.9527</td>
<td>0.9533</td>
<td>0.9535</td>
</tr>
<tr>
<td>48</td>
<td>0.6525</td>
<td>0.6531</td>
<td>0.6532</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Z-tree</th>
<th>Y-tree</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>32</td>
<td>8.4439</td>
<td>8.4437</td>
<td>8.4440</td>
</tr>
<tr>
<td>34</td>
<td>6.9528</td>
<td>6.9526</td>
<td>6.9524</td>
</tr>
<tr>
<td>36</td>
<td>5.6333</td>
<td>5.6331</td>
<td>5.6329</td>
</tr>
<tr>
<td>38</td>
<td>4.4946</td>
<td>4.4944</td>
<td>4.4943</td>
</tr>
<tr>
<td>40</td>
<td>3.5340</td>
<td>3.5337</td>
<td>3.5352</td>
</tr>
<tr>
<td>42</td>
<td>2.7455</td>
<td>2.7452</td>
<td>2.7451</td>
</tr>
<tr>
<td>44</td>
<td>2.1071</td>
<td>2.1068</td>
<td>2.1071</td>
</tr>
<tr>
<td>46</td>
<td>1.6005</td>
<td>1.6002</td>
<td>1.6011</td>
</tr>
<tr>
<td>48</td>
<td>1.2053</td>
<td>1.2051</td>
<td>1.2059</td>
</tr>
</tbody>
</table>

the model with more regimes perform similarly and are omitted here.

Finally, we compare our trinomial method to Yuen and Yang’s tree. This is presented in Table 5.3. Both methods are calculated with 2,000 time steps. In the table, the results in the Z-tree column are obtained from our new method while the results in the Y-tree are from the Yuen and Yang’s method. It turns out that both methods provide values that are really close to the real values obtained from the formula. However, although the methods have the same convergence rate, our method is slightly better than Yuen and Yang’s method in most cases except 6 out of 18 comparisons. Our method has smaller errors except $S_0 = 40$, 46, 48 in regime 1 and $S_0 = 34$, 38, 42 in regime 2.
5.8 Conclusion

A new trinomial tree method for options pricing under a general multi-state regime-switching model is presented. Compared to the tree of Yuen and Yang [87], we make a different assumption, which would lead to an equivalence between our tree approach and a standard explicit finite difference method. The equivalence present in this paper contains no error terms, in contrast to the result of Ma et al. [61], and is numerically tested to be robust with various different numbers of time steps.
Chapter 6

A new simple tree approach for Heston’s stochastic volatility model

6.1 Introduction

Although the celebrated Black-Scholes (BS) model [5] is a breakthrough for the options pricing theory and is widely used among options traders, the results obtained from their analytical solution do not perfectly match the market data and the formula is reported sometimes to produce inaccurate values [75]. This is principally from the constant volatility assumption, in contrast to the “volatility smile” and “skew” observed in the practical world. One way to address the issue is to assume the volatility to follow another stochastic process, which is known as the stochastic volatility (SV) model.

There have been a number of different SV models developed by researchers (for example, [33, 39, 42, 79]). The main problem that this type of models brings is that either the model itself is too complicated, or the option contracts that are to be priced under these models are path-dependent, analytical solutions similar to the BS formula are rarely found. Thus numerical approximations have to be applied and efficient numerical methods are needed.

As one of the most popularly adopted numerical approaches, tree-based methods, since the pioneering work by Cox, Ross and Rubinstein [19], have been proven to be very efficient for pricing various options. Such a popularity stems from its simplicity as well as clear
financial interpretation. The essence of the tree-based method is to use a discrete-time jump process (the tree process) to approximate the stochastic differential equation (SDE) that the underlying asset follows. With the local mean and variance matched over any time periods, the tree process converges to the SDE as the length of time step goes to 0. The efficiency of the tree-based methods depends on the recombining property, which guarantees the number of nodes grows only linearly as the number of time step increases.

Tree methods for the BS model recombine naturally due to the constant volatility assumption, which results in constant jump sizes throughout the life of an option. Stochastic volatility models, on the other hand, has a volatility term that follows another random process. The recombining property is no longer trivial as the jump size varies over time. This could cause the tree branches grow exponentially and further make it computationally explosive. Thus, maintaining the recombining property of tree-based methods becomes a major difficulty for the Heston model.

To overcome this difficulty, various approaches have been discussed in the literature. One of the most efficient methods is to introduce transformations to make the new transformed variables have constant volatility terms. Leisen [50] constructs a multinomial tree approach by applying the Nelson-Ramaswamy transformation [67]. The tree has eight successor nodes at every time step, with four for each of the asset price and variance. As presented by the author, the discrete multinomial tree process weakly converges to the coupled Heston’s SDEs. However, the non-zero correlations, as presented in [3, 50], poses a major problem that unavoidable negative joint probabilities appear at each time step. Then Beliaeva and Nawalkha [3] suggests another transformation to the stock process that is comprised of three terms which orthogonalizes the two Brownian motions to make them conditionally independent. Under this transformation, the joint probabilities become the product of the marginal probabilities so that they are positive as long as the marginal probabilities are positive. The method, though, has a bad tree structure because of the multi-jump algorithm which the authors introduce to guarantee the positivity of the marginal probabilities along the variance direction (see [3]). In addition, the tree is three-dimensional and both space increments, $\Delta S$ and $\Delta v$, are dependent on time increment $\Delta t$, which results in an over-generated amount of volatility grids being taken into consideration. This significantly slows down their approach and the method is thus very inefficient.
Then Liu [53] presents another approach, in which Heston’s stochastic volatility model is approximated by a regime-switching model. The author also applies the transformation from Nelson-Ramaswamy’s [67] to make the new variables conditionally independent to each other but then discretizes the transformed volatility process by a continuous-time Markov chain. By adopting his recombining tree for the obtained regime-switching model, the price of an option can be simply calculated. Being discretized, the variance increment is independent of time increment and it has been shown in [53] that 12 variance steps are enough for the tree method to perform very well. Liu’s approach is very simple to implement and computationally efficient. His tree structure, though, is still not good enough as too many branches are needed at each time step. The challenge remains for researchers to establish a simple tree approach with least amount of branches required at each time step.

In this chapter, a modification is proposed to this approach in which we further simplify the tree-based numerical method by cutting down the number of branches at each time step. We apply the trinomial tree method presented in [90] to the approximated regime-switching model, which makes it more computationally efficient. In addition, as the well-known equivalence between the trinomial tree and the explicit finite difference under the BS model, we extend the equivalence to the Heston model, which is proven analytically and verified numerically.

The rest of the chapter is organized as follows. Section 6.2 introduces the related models and notations to be used for the simplicity of discussion. The new simple tree-based numerical method is presented in Section 6.3. The equivalence between our tree approach and explicit finite difference methods under the Heston model is proved in Section 6.4. Numerical performance and examples are shown in Section 6.5, followed by concluding remarks in Section 6.6.

6.2 Models and notations

In this section, we introduce Heston’s stochastic volatility model as well as the regime-switching model for the reason that they are much relevant to the further discussion in this chapter. We consider an underlying asset, whose price dynamic is denoted as $S_t$, to
follow a stochastic differential equation (SDE) of a geometric Brownian motion:

\[ dS_t = \mu S_t dt + \sigma_t S_t dB^1_t. \]  \hspace{1cm} (6.2.1)

Here \( \mu \) and \( \sigma_t \) are the drift term and the volatility term, respectively. \( B^1_t \) is a standard Brownian motion. The subscript \( t \) indicates the time dependence.

The Heston model and the regime-switching model are both stochastic volatility models. However, their volatilities are random in different ways. In the Heston model, the volatility is allowed to “diffuse” in the whole space controlled by a diffusion process. Thus there are two SDEs in the Heston model which are further assumed to be correlated. It is quite obvious that the range of Heston’s volatility is continuous, as Brownian motions are continuous random processes.

On the other hand, the regime-switching model considers the volatility to randomly jump among a number of different values that depend on a continuous-time Markov chain. The Markov chain is assumed to be independent of the Brownian motion of the stock price process. Unlike volatility in the Heston model, the range of the regime-switching volatility is discrete and countable. Financially the discrete range could be interpreted as a “holding time” feature of the volatility movement. This is in contrast to the Heston model, whose volatility diffuses “ceaselessly”.

### 6.2.1 Heston’s stochastic volatility model

In Heston’s stochastic volatility model, \( \sigma_t \) is selected to be the square root of a variance process \( v_t \) that follows a mean-reverting stochastic process:

\[ dv_t = \kappa(\theta - v_t)dt + \sigma \sqrt{v_t} dB^2_t. \]  \hspace{1cm} (6.2.2)

Here \( \theta \) is the long-term mean of \( v_t \), \( \kappa \) is the rate of relaxation to this mean, \( \sigma \) is the volatility term of the volatility process (vol. of vol.) and \( B^2_t \) is another standard Brownian motion. The two Brownian motions are assumed to be correlated with each other:

\[ dB^2_t = \rho dB^1_t + \sqrt{1 - \rho^2} dW_t, \]  \hspace{1cm} (6.2.3)
6.2. MODELS AND NOTATIONS

where \( W_t \) is a Brownian motion independent of \( B_t^1 \), and \( \rho \) is the correlation coefficient that satisfies \(-1 \leq \rho \leq 1\). With (6.2.1) and (6.2.2), the price of an underlying asset \( S_t \) is said to follow the Heston model, if

\[
\begin{align*}
    dS_t &= \mu S_t dt + \sqrt{v_t} S_t dB_{1t}, \\
    dv_t &= \kappa (\theta - v_t) dt + \sigma \sqrt{v_t} dB_{2t}.
\end{align*}
\]

(6.2.4)

Since \( \sigma_t = \sqrt{v_t} \), to ensure \( v_t > 0 \), the Feller condition \( 2\kappa \theta > \sigma^2 \) has to be satisfied as mentioned in [1].

Let \( U(S, v, t) \) denote the value of an option, with \( S \) being the price of the underlying asset, \( v \) being the variance of the market and \( t \) being the time. Then under the Heston model, it can be easily shown that with non-arbitrage argument, \( U \) should satisfy the following bivariate partial differential equation (PDE):

\[
\frac{\partial U}{\partial t} + \frac{1}{2} \rho^2 v S^2 \frac{\partial^2 U}{\partial S^2} + \rho \sigma S v \frac{\partial^2 U}{\partial S \partial v} + \frac{1}{2} \sigma^2 v \frac{\partial^2 U}{\partial v^2} + r S \frac{\partial U}{\partial S} + \kappa (\theta - v) \frac{\partial U}{\partial v} - r U = 0.
\]

(6.2.5)

To solve the PDE (6.2.5), the terminal condition needs to be given by the payoff function of the option. For example, for a European call option, it reads: \( U(S, v, T) = \max(S - E, 0) \), where \( E \) is the strike price.

**6.2.2 The regime-switching model**

As we mentioned, the regime-switching model belongs to another category of stochastic volatility models. Assume \( P_t \) is a continuous-time Markov chain with \( K \) states and the state space is given by \( \mathcal{P} := \{p_1, p_2, \ldots, p_K\} \). The generator of the Markov chain \( P_t \) is denoted by \( Q = (q_{ij})_{m \times n} \), which satisfies the following properties: (i) \( q_{ij} < 0 \) for all \( i = j \) and \( q_{ij} \geq 0 \) for all \( i \neq j \); (ii) \( \sum_{j=1}^{K} q_{ij} = 0 \) for all \( i = 1, 2, \ldots, K \). Then in a regime-switching model, the volatility term \( \sigma_t \) in (6.2.1) is a function of \( P_t \):

\[
\hat{\sigma}(P_t) = \begin{cases} 
    \hat{\sigma}_1, & P_t = p_1, \\
    \hat{\sigma}_2, & P_t = p_2, \\
    \vdots & \\
    \hat{\sigma}_N, & P_t = p_K.
\end{cases}
\]
Here the $\{\hat{\sigma}_n\}_{n=1}^K$ is a set of constant numbers. Thus the regime-switching SDE can be written as:

$$dS_t = \mu(P_t)S_t \, dt + \hat{\sigma}(P_t)S_t \, dB^1_t. \tag{6.2.6}$$

Here it should be remarked that the Markov chain $P_t$ is independent of the Brownian motion $B^1_t$. In addition, due to the presence of the Markov chain, the governing PDE of a regime-switching model becomes a PDE system whose solutions are required to be solved simultaneously.

To be more specific, assume an option whose underlying asset follows the regime-switching dynamic (6.2.6) to have price $\hat{U}(S, P, t)$, where $S$ and $t$ are the same variable settings as we mentioned in the Heston model while $P$ represents regime. Note that $\hat{U}(S, P, t)$ is a vector that can be expressed as $\hat{U}(S, P, t) = (U(S, p_1, t), U(S, p_2, t), \ldots, U(S, p_K, t))$. To simplify the notations, we then denote the $i$th component of the solution vector, $U(S, p_i, t)$, by $U_i(S, t)$, $i = 1, 2, \ldots, K$. Then from [21], $\hat{U}(S, P, t)$ should be the solution vector of the following PDE system

$$
\begin{align*}
\frac{\partial U_1}{\partial t} &+ \frac{1}{2} \sigma_1^2 S^2 \frac{\partial^2 U_1}{\partial S^2} + rS \frac{\partial U_1}{\partial S} - ru_1 + \sum_{j=1}^K q_{1j} U_j = 0, \\
\frac{\partial U_2}{\partial t} &+ \frac{1}{2} \sigma_2^2 S^2 \frac{\partial^2 U_2}{\partial S^2} + rS \frac{\partial U_2}{\partial S} - ru_2 + \sum_{j=1}^K q_{2j} U_j = 0, \\
&\vdots \\
\frac{\partial U_K}{\partial t} &+ \frac{1}{2} \sigma_K^2 S^2 \frac{\partial^2 U_K}{\partial S^2} + rS \frac{\partial U_K}{\partial S} - ru_K + \sum_{j=1}^K q_{Kj} U_j = 0.
\end{align*}
\tag{6.2.7}
$$

However, in practice, only one component of $\hat{U}(S, P, t)$ is normally required to be calculated since the initial regime is always known and unique. Hence it is quite inefficient to numerically solve the whole PDE system (6.2.7) just to obtain one single solution. The Monte Carlo simulation has been tested to be a more efficient numerical approach (see [89]) when the underlying asset follows a regime-switching economy.
6.3 The new simple tree approach

6.3.1 Method of Lines

As introduced in the previous section, it seems that the regime-switching model is a discrete version of Heston’s stochastic volatility model with zero-correlation. The relationship is confirmed by Liu [53], in which the author presents a “Method of Lines” (MOL) approach to discretize a decoupled Heston model and obtains a regime-switching model. So the first part of this section is to introduce his methodology, followed by our modifications.

Consider a stock whose price $S_t$ follows the Heston model’s coupled SDE (6.2.1). First, since the Markov chain of the regime-switching model is independent of the stock process, the transformation introduced by Beliaeva and Nawalkha [3] is applied to decouple the two processes of the Heston model. Let

$$
\begin{align*}
X_t & = \ln \left( \frac{S_t}{S_0} \right) - \frac{\rho}{\sigma} (v_t - v_0) - \left( r - \frac{\rho \kappa \theta}{\sigma} \right) t, \\
\omega_t & = 2 \sqrt{v_t}
\end{align*}
$$

(6.3.1)

With Ito’s lemma, we thus yield the transformed Heston model

$$
\begin{align*}
dX_t & = \frac{1}{4} \alpha \omega_t^2 dt + \frac{1}{2} \sqrt{1 - \rho^2} \omega_t dW_t, \\
d\omega_t & = \phi(w) dt + \sigma dB_t^2
\end{align*}
$$

(6.3.2)

where

$$
\alpha = \left( \frac{\rho \kappa}{\sigma} - \frac{1}{2} \right), \quad W_t = \frac{B_t^1 - \rho B_t^2}{\sqrt{1 - \rho^2}}, \quad \phi(w) = \left( 2 \kappa \theta - \frac{\sigma^2}{2} \right) \frac{1}{w} - \frac{\kappa}{2} w.
$$

The formulation of the $W_t$ indicates that the new variables $X_t$ and $\omega_t$ are uncorrelated. Hence the obstacle of the non-zero correlation has been removed. Then the next step is to construct a continuous-time Markov chain as the discrete transformed variance process.

Assume $P_t$ to be the $N$-state continuous-time Markov chain that approximates the transformed variance process $\omega_t$. Further we let the drift term and the volatility term of
the random process $X_t$ take discrete values according to $P_t$ as presented below:

$$
a(P_t) = \begin{cases} 
\frac{1}{4}w_1^2, & P_t = 1, \\
\frac{1}{4}w_2^2, & P_t = 2, \\
\ldots \\
\frac{1}{4}w_N^2, & P_t = N,
\end{cases}
$$

and

$$
b(P_t) = \begin{cases} 
\frac{1}{2}\sqrt{1 - \rho^2}w_1, & P_t = 1, \\
\frac{1}{2}\sqrt{1 - \rho^2}w_2, & P_t = 2, \\
\ldots \\
\frac{1}{2}\sqrt{1 - \rho^2}w_N, & P_t = N.
\end{cases}
$$

Here what we need is to let the two discrete functions converge to the original continuous drift and volatility terms as the number of states $N$ goes to infinity. This can be achieved by selecting $w_n = n\Delta w$, ($n = 1, 2, ..., N$), where $\Delta w$ is a given increment that will be introduced later in this section. The transformed Heston model (6.3.2) can be written as

$$
dX_t = a(P_t)dt + b(P_t)dW_t,
$$

which is a $N$-state regime-switching model.

Now the missing part is the Markov chain generator (also known as the $Q$-matrix). Liu [53] determines the $Q$-matrix by comparing the coefficients of governing PDEs of Model (6.3.2) and (6.3.3). Consider an option with expiration $T$ and here we discuss two different scenarios. In the first scenario, we assume the price and volatility of its underlying asset to follow the SDE system (6.3.2). Let $H(x, v, t)$ denote the price of the option at time $t \leq T$ when $X_t = x$ and $v_t = v$. In a risk-neutral world, $H(x, v, t)$ is the solution of the following PDE:

$$
\frac{\partial H}{\partial t} + \frac{1}{8}(1 - \rho^2)w^2 \frac{\partial^2 H}{\partial x^2} + \frac{1}{4} \alpha w^2 \frac{\partial H}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 H}{w^2} + \phi(w) \frac{\partial H}{w} - rH = 0. \tag{6.3.4}
$$

On the other hand, we consider the second scenario to be all the same but the underlying asset of the option is under a regime-switching economy (6.3.3). In this case, the price of the option is denoted by $\tilde{H}_n(x, t)$, where the subscript $n$ represents the initial regime. As we mentioned in the previous section, the general form of the governing PDE is

$$
\frac{\partial \tilde{H}_n}{\partial t} + \frac{1}{8}(1 - \rho^2)w^2 \frac{\partial^2 \tilde{H}_n}{\partial x^2} + \frac{1}{4} \alpha w^2 \frac{\partial \tilde{H}_n}{\partial x} + \sum_{k=1}^{N} q_{n,k} \tilde{H}_k - r\tilde{H}_n = 0, \quad n = 1, 2, ..., N. \tag{6.3.5}
$$
Note that it is obvious that the two equations are of high similarity. Both equations have exact same terms and coefficients except the partial derivatives with respect to the variable \( w \). Equation (6.3.4) has two partial derivative terms with respect to \( w \), one first-order and one second-order, while Equation (6.3.5) has none but a summation term with respect to initial regimes. In fact if we view \( \tilde{H}_n(x,t) \) as \( H(x,v_n,t) \), Equation (6.3.5) can be regarded as (6.3.4) with the method of line method being applied to the \( w \) variable. In particular, we discretize the variable \( w \) in Equation (6.3.5) and set up a uniform grid for the range \([0, \infty)\) with \( \Delta w \) being the space increment. Define \( H_n \) as \( H(x,n\Delta w,t) \) for \( n = 1, 2, ..., N \). Then by applying the central finite difference scheme

\[
\frac{\partial^2 H}{\partial w^2} = \frac{H_{n+1} - 2H_n + H_{n-1}}{(\Delta w)^2}, \quad \frac{\partial H}{\partial w} = \frac{H_{n+1} - H_{n-1}}{2\Delta w},
\]

the partial derivative terms with respect to \( w \) becomes

\[
\frac{1}{2} \sigma^2 \frac{\partial^2 H}{\partial w^2} + \phi(w) \frac{\partial H}{\partial w} = \pi_{n,n+1}H_{n+1} + \pi_{n,n}H_n + \pi_{n,n+1}H_{n-1}, \quad (6.3.6)
\]

where

\[
\pi_{n,n+1} = \frac{\sigma^2}{2(\Delta w)^2} + \frac{4\kappa \theta - \sigma^2}{4n(\Delta w)^2} - \frac{n\kappa}{4},
\]

\[
\pi_{n,n} = -\frac{\sigma^2}{(\Delta w)^2},
\]

\[
\pi_{n,n-1} = \frac{\sigma^2}{2(\Delta w)^2} - \frac{4\kappa \theta - \sigma^2}{4n(\Delta w)^2} + \frac{n\kappa}{4}.
\]

Then (6.3.6) can be written as

\[
\frac{1}{2} \sigma^2 \frac{\partial^2 H}{\partial w^2} + \phi(w) \frac{\partial H}{\partial w} = \sum_{k=1}^{N} \pi_{n,k}H_k,
\]

with all of the \( \{\pi_{n,k}\}_{k=1}^{N} \) being zero other than \( \pi_{n,n-1}, \pi_{n,n} \) and \( \pi_{n,n+1} \). Then the two equations become identical and \( \tilde{H}_n = H_n \). Thus we can conclude that \( \tilde{H}_n \) converges to \( H \) as \( N \) goes to infinity.

So it seems that the Markov chain generator has been determined. It actually is, however, only when \( \pi_{n,n-1} \) and \( \pi_{n,n+1} \) are both non-negative due to the property (i) of the Markov chain we have introduced in Section 6.2.2. In [53], three different classes
are considered, in which the central, backward or forward finite difference is applied to the term \( \frac{\partial H}{\partial u} \) when both \( \pi_{n,n-1} \) and \( \pi_{n,n+1} \) are greater than or equal to zero, \( \pi_{n,n+1} \) is negative or \( \pi_{n,n+1} \) is negative, respectively. Here we denote \( \beta_n^+ = q_{n,n+1}, \beta_n = q_{n,n} \) and \( \beta_n^- = q_{n,n-1} \). Then the three different cases can be written by

**Case 1:** \( \beta_n^+ = \pi_{n,n+1}, \beta_n = -(\pi_{n,n+1} + \pi_{n,n-1}), \beta_n^- = \pi_{n,n-1} \),

**Case 2:** \( \beta_n^+ = -\frac{1}{2}(\pi_{n,n+1} + \pi_{n,n-1}), \beta_n = -2\pi_{n,n-1}, \beta_n^- = -\frac{1}{2}\pi_{n,n+1} + \frac{3}{2}\pi_{n,n-1} \),

**Case 3:** \( \beta_n^+ = \frac{3}{2}\pi_{n,n+1} - \frac{1}{2}\pi_{n,n-1}, \beta_n = -2\pi_{n,n+1}, \beta_n^- = \frac{1}{2}\pi_{n,n+1} + \frac{1}{2}\pi_{n,n-1} \).

It should be mentioned that an extrapolation approach has to be adopted at the lower and upper bound of regimes, with \( 2H_n = H_{n+1} + H_{n-1} \). Then at the lower bound of the regime \( n_l \), we have

\[
\beta_{n_l}^+ = \pi_{n_l,n_l+1} - \pi_{n_l,n_l-1} = \frac{4\kappa\theta - \sigma^2}{2n_l(\Delta w)^2} - \frac{\kappa n_l}{2},
\]

\[
\beta_{n_l} = \pi_{n_l,n_l} + 2\pi_{n_l,n_l-1} = -\left( \frac{4\kappa\theta - \sigma^2}{2n_l(\Delta w)^2} - \frac{\kappa n_l}{2} \right).
\]

Here \( n_l \) is a sufficient small integer such that \( \frac{4\kappa\theta - \sigma^2}{2n_l(\Delta w)^2} - \frac{\kappa n_l}{2} \geq 0 \), given a pre-defined \( \Delta w \), to ensure the positivity of \( \beta_{n_l}^+ \). Similarly at the upper bound of the regime \( N \),

\[
\beta_N = \pi_{N,N} + 2\pi_{N,N-1} = -\frac{4\kappa\theta - \sigma^2}{2N(\Delta w)^2} + \frac{\kappa N}{2},
\]

\[
\beta_N^- = \pi_{N,N-1} - \pi_{N,N+1} = \frac{4\kappa\theta - \sigma^2}{2N(\Delta w)^2} - \frac{\kappa N}{2},
\]

where \( N \) is the total amount of space steps that should satisfy \( \frac{4\kappa\theta - \sigma^2}{2N(\Delta w)^2} - \frac{\kappa N}{2} \leq 0 \). If we denote \( f(n) = \frac{4\kappa\theta - \sigma^2}{2n(\Delta w)^2} - \frac{\kappa n}{2} \), then \( n_l \) and \( N \) are two positive integers such that \( n_l < N \), \( f(n_l) \geq 0 \) and \( f(N) \leq 0 \).

Thus the procedure is as follows. We first give a spacial increment \( \Delta w \). With the fixed \( \Delta w \), find \( \tilde{n} \) such that \( f(\tilde{n}) = 0 \) and pick two integers \( n_l \) and \( N \) which satisfy \( n_l \leq \tilde{n} \leq N \). Then range \([0, \infty)\) is truncated to \([w_{\min}, w_{\max}]\), with \( w_{\min} = n_l\Delta w \) and \( w_{\max} = N\Delta w \). Hence we obtain a set of grid \( \{w_n\}_{n=n_l}^{N} \), whose total number is \( N - n_l + 1 \).

After setting up the grid, the approximated Markov chain generator matrix \( Q \) can be
6.3. THE NEW SIMPLE TREE APPROACH

written as

\[
Q = (q_{m,n}) = \begin{pmatrix}
\beta_{n_1} & \beta_{n_1}^+ & 0 & 0 & 0 & \cdots & 0 \\
\beta_{n_1}^- & \beta_{n_1}^+ & \beta_{n_1}^+ & 0 & 0 & \cdots & 0 \\
0 & \beta_{n_1}^+ & \beta_{n_1}^+ & \beta_{n_1}^+ & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 0 & \beta_{N-2}^- & \beta_{N-2}^+ & \beta_{N-2}^+ & 0 \\
0 & \cdots & 0 & 0 & \beta_{N-1}^- & \beta_{N-1}^+ & \beta_{N-1}^+ \\
0 & \cdots & 0 & 0 & 0 & \beta_{N}^- & \beta_{N}^+
\end{pmatrix},
\]

and tree-based approximations can be constructed on this regime-switching model.

From here onward, Liu [53] starts to apply his recombining tree to the regime-switching model to approximate option prices whose underlying assets follow the Heston model. However, we find that his recombining tree approach is computationally expensive, in addition to his sophisticated regime-switching tree structure, which loses the key advantage of tree-based approximation, the financial interpretation. Thus we construct a new trinomial tree approach for the regime-switching model, which turns out to be computationally cheaper with a clear tree structure.

6.3.2 New trinomial tree for regime-switching models

In a trinomial tree model with constant drift term and volatility term, the price of the underlying stock is allowed to move upward, remain unchanged or move downward by a ratio. The ratio has to be greater than the one of a binomial tree model so that the jump probabilities are positive. Since in the CRR binomial tree model (see [19]), the ratios are assumed as \( e^{\sigma\sqrt{\Delta t}} \) and \( e^{-\sigma\sqrt{\Delta t}} \), where \( \Delta t \) is the size of time step in the model, it is natural that we assume the ratios of our trinomial tree model to be \( e^{\lambda\sigma\sqrt{\Delta t}} \), \( 1 \) and \( e^{-\lambda\sigma\sqrt{\Delta t}} \) with \( \lambda > 1 \). Although the regime-switching model allows the drift term and the volatility to change according to the Markov chain, they are constant when the state of the Markov chain is fixed. Hence for each individual state, we can propose a trinomial tree model and then make the tree branches from all states overlapped by adjusting the jump probabilities as presented in Yuen and Yang [87].

Consider a stock that under a \( N \)-state regime-switching economy follows the SDE (2.6). Let \( \phi_u^n, \phi^m_p \) and \( \phi_d^n \) be the jump probabilities corresponding to when the stock price
108 | 6.3. THE NEW SIMPLE TREE APPROACH

increases, remains the same and decreases in state \( n \), respectively. Thus the jump size for state \( n \) are \( e^{\lambda n \sigma_n \sqrt{\Delta t}} \), \( 1 \) and \( e^{-\lambda n \sigma_n \sqrt{\Delta t}} \). To make the each individual tree have the same step size, we set \( \lambda_1 \sigma_1 = \lambda_2 \sigma_2 = \ldots = \lambda_N \sigma_N = \bar{\sigma} \), which is a given value greater than all the volatilities. As suggested by Yuen and Yang [87], a popular one is selected as

\[
\bar{\sigma} = \max_{j=1,2,\ldots,K} \{\sigma_j\} + (\sqrt{1.5} - 1)\sigma, \tag{6.3.7}
\]

where \( \bar{\sigma} \) is either the arithmetic mean or geometric mean of all volatilities. Then the jump sizes and jump probabilities must match the local drift term and volatility of the stochastic process, thus

\[
\begin{align*}
\phi_n^u \lambda_n \sigma_n \sqrt{\Delta t} - \phi_n^d \lambda_n \sigma_n \sqrt{\Delta t} &= \mu_n \Delta t, \\
\phi_n^u \lambda_n^2 \sigma_n \Delta t + \phi_n^d \lambda_n^2 \sigma_n \Delta t &= \sigma_n^2 \Delta t, \\
\phi_n^u + \phi_n^m + \phi_n^d &= 1,
\end{align*}
\tag{6.3.8}
\]

where \( \mu_n \) and \( \sigma_n \) are the values that the drift and volatility terms take in regime \( n \), \( n = 1, 2, \ldots, N \). The equation system (6.3.8) can be solved with some simple algebras and we obtain

\[
\begin{align*}
\phi_n^u &= \frac{1}{2} \left( \frac{1}{\lambda_n^2} + \frac{\mu_n \sqrt{\Delta t}}{\lambda_n \bar{\sigma}} \right), \\
\phi_n^m &= 1 - \frac{1}{\sigma_n^2}, \\
\phi_n^d &= \frac{1}{2} \left( \frac{1}{\lambda_n^2} - \frac{\mu_n \sqrt{\Delta t}}{\lambda_n \bar{\sigma}} \right).
\end{align*}
\tag{6.3.9}
\]

Thus the parametrization of the trinomial tree model with constant drift term and volatility for every single state has been established. The next step is to set up the switching feature, which is to build the connection for all individual trees. At each time step, the combined tree must account for two possible types of changes: the change of market conditions (regimes) and the change of stock prices within a regime. It’s natural to raise questions regarding the order of these changes, or whether they should occur simultaneously.

Yuen and Yang [87] presents an assumption that the two changes take place simultaneously. In particular, at time \( t = t_i \), a stock in regime \( j \) with price \( S = S_m \) is allowed to
6.3. THE NEW SIMPLE TREE APPROACH

Figure 6.1: Graphical illustration of a two-state regime-switching model

jump to \( S = S_{m+1}, S_m, S_{m-1} \) in all regimes at the \( t = t_{i+1} \). Thus their recursive formula requires a total of \( 3K \) (\( K \) is the total number of regimes considered) nodal values to be involved when computing the nodal value of the previous time step.

Our approach, however, makes the following assumption instead. Within one time step, the stock price can either stay in the same regime and “diffuse” along the trinomial tree in the usual way or switch regimes but without diffusing. For instance, a node \((S_m, t_i, n)\) is allowed to jump to \((S_{m+1}, t_i, n)\), \((S_{m-1}, t_i, n)\) or \((S_m, t_i, \hat{n})\), where \( \hat{n} \) stands for all regimes including \( n \). The advantage of this assumption is that only \( K+2 \) nodal values are required in computation at each time step, which is a significant reduction when the number of regimes is quite large.

Since the weights of convex combination (recursive formula) are interpreted as the probabilities of the previous term jumping to the current term, it could be determined in a rather simple way. The up and down probabilities are not affected by the switching feature under our assumption so they remain the same as the individual trees, which are \( \phi_n^m \) and \( \pi_n^d \), respectively. The middle branch, however, contains the possibility of staying in the same regime and switching to others. Thus the summation of the coefficients of the rest \( K \) terms should be equal to \( \pi_n^m \), with each being the transition probability. We then give our recursive formula below

\[
U_{i,m,n} = e^{-r\Delta t} \left( \phi_n^m U_{m+1,n}^{i+1} + (\phi_n^m + q_{n,n} \Delta t)U_{m,n}^{i+1} + \phi_n^d U_{m-1,n}^{i+1} + \sum_{k=1}^{N} q_{n,k} \Delta t U_{m,k}^{i+1} \right). 
\]

(6.3.10)

The transition probability is approximated by \( q_{n,k} \Delta t \ (n \neq k) \), which is obtained from the definition of the Markov chain generator \( Q \). The total probability of switching out
from the current state is given by $q_{n,n} \Delta t$. Since $\sum_{k=1}^{N} q_{n,k} = 0$, the sum of all weighted coefficients is 1, which ensures the feasibility of the formulation.

Now we have established the new trinomial tree method for a general $K$-state regime-switching model and now we are ready to move on to design the simple tree method for Heston’s stochastic volatility model.

### 6.3.3 New simple tree for the Heston model

The new simple tree-based approximation is to apply the trinomial tree method (6.3.10) we introduced from the last subsection to model (6.3.3). It’s worthwhile mentioning that the total amount of the regimes in (6.3.3) is $N - n_l + 1$ instead of $N$. Such an application is quite straightforward and a brief introduction is presented here.

First of all, we need to settle the parametrization of the method. Define $a_n := a(P_t = n)$ and $b_n := b(P_t = n)$. Then the value of $\tilde{b}$ in (6.3.7) is given as

$$\tilde{b} = \max_{n \in n_l, n_l + 1, \ldots, N} \{ b_n \} + (\sqrt{1.5} - 1) \bar{b}. \quad (6.3.11)$$

where $\bar{b} = \sum_{n = n_l}^{N} b_n / (N - n_l + 1)$. Here it is remarked that the step size should be determined as $\tilde{b} \sqrt{\Delta t}$, 0, $-\tilde{b} \sqrt{\Delta t}$, since the variable $X_t$ in (6.3.1) follows a Brownian motion instead of Geometric Brownian motion. Then we proceed to determine the probabilities of the jump process. This can be simply done by replacing $\mu_n$ and $\tilde{\sigma}_n$ by $a_n$ and $b_n$ in (6.3.9). Finally, the recursive formula is obtained by substituting these parameters into (6.3.10).

By far it seems that the tree approach is merely another version of Liu [53], the method of which is for regime-switching models instead of the Heston model. However, it has another interpretation. Note that the Markov chain generator matrix $Q$ is tridiagonal, which means that a regime only switches to the two adjacent ones. Hence we can simplify (6.3.10) by

$$U_{m,n}^i = e^{-r \Delta t} \left( \phi_n^m U_{m+1,n}^{i+1} + (\phi_n^m + q_{n,n} \Delta t) U_{m,n}^{i+1} + \phi_n^d U_{m-1,n}^{i+1} + q_{n,n-1} \Delta t U_{m,n-1}^{i+1} + q_{n,n+1} \Delta t U_{m,n+1}^{i+1} \right). \quad (6.3.12)$$

Note that $\Delta t < \min_{n = n_l, n_l + 1, \ldots, N} (\phi_n^m / q_{n,n})$ has to be satisfied to insure the stability of the
method. This formula shows a five-point jump process which approximates options under the transformed Heston model (6.3.2).

\[

t_{m+1,n}^{k+1} = \begin{cases} 
U_{m,n}^k & \text{regime } n \\
U_{m,n-1}^k & \text{regime } n-1 \\
U_{m,n+1}^k & \text{regime } n+1
\end{cases}
\]

Figure 6.2: Graphical illustration for the five-point jump process

To have a closer look at our tree structure, a graphical illustration is given in Figure 6.2 over one time step. The price of the option at time \( t = t_i \) is allowed to attain one of the five future values at time \( t = t_{i+1} \). The five points are located on the \( X - w \) plane with four being vertices of a rhombus while the other one being the centre. Together with the seed node \( U_{m,n}^k \), the one unit of our tree constitutes a square pyramid. The height of the pyramid, \( U_{m,n}^k - U_{m,n}^{k+1} \), stands for both transformed variable \( X \) and \( w \) remain unchanged over the time step. The four edges represent respectively one variable change while the other one does not. The plot shows a tree structure very similar to Boyle’s tree-based method for options pricing with two-state variables as presented in [9].

Figure 6.3: Graphical illustration for the \( t - w \) plane
Note that the unit presented in Figure 6.2 is an example for the cases where $n$ satisfies $n_l + 1 \leq n \leq N - 1$. At the lower bound and the upper bound of the regimes, due to the extrapolation, only four points are involved in evaluating the price of the option from the previous time step. Therefore the rhombus goes down to a triangle at those locations. In summary, the cross-section of the tree on the $t - X$ plane is an ordinary trinomial tree process (as shown in Figure 6.4) while on the $t - w$ plane it becomes what is presented in Figure 6.3. Therefore the tree structure is compromised of a total of $N - n_l + 1$ separate trees with each one being parallel to another, which looks like a piece of cake. It’s worthwhile mentioning that our tree structure can be reduced back to the one-dimensional CRR model naturally by letting all of the regimes be the same. In other words, if we “squash” the tree along the $w$-direction, all the individual trees will overlap each other. Then it reduces back to one single tree, which is the trinomial tree presented by Kamrad and Ritchken in [45].

6.3.4 The tree in the original coordinate plane

It should be noted that the new simple tree approach is built purely on mathematically introducing variables $X$ and $w$, which have no financial meanings. On the other hand, a key advantage of tree methods, i.e., the clear economic intuition, would have been lost entirely, if we could not interpret the new tree approach financially. Hence in this subsection, we explore what the tree in the original coordinate plane, that is, stock price $S_t$ and variance $v_t$, is like, which naturally implies the financial interpretation that the tree provides.

The jump size of the tree under original coordinate is determined by the transformation (6.3.1). Upon rearranging the variables $S_t$ and $w_t$ to the left-hand side and other terms
to the right-hand side, the “inverse-transform” can be implicitly written as

\[
\begin{align*}
\log \frac{S_t}{S_0} &= X_t + \frac{\rho}{\sigma}(v_t - v_0) + (r - \frac{\rho \theta}{\sigma}) t, \\
v_t &= \frac{1}{4} w_t^2.
\end{align*}
\]

(6.3.13) \hfill (6.3.14)

Now consider the stochastic processes over a time step \( \Delta t \). The increments of \( S_t \) and \( v_t \) are then determined by

\[
\begin{align*}
\log \frac{S_{t+\Delta t}}{S_t} &= (X_{t+\Delta t} - X_t) + \frac{\rho}{\sigma}(v_{t+\Delta t} - v_t) + (r - \frac{\rho \theta}{\sigma}) \Delta t, \\
v_{t+\Delta t} - v_t &= \frac{1}{4}(w_{t+\Delta t} + w_t)(w_{t+\Delta t} - w_t).
\end{align*}
\]

(6.3.15) \hfill (6.3.16)

According to the tree settings, each of \( X_t \) and \( w_t \) is assumed to take only three future values (except the upper and lower bound of \( w_t \)). We further let \( X_t = X_m \) and \( w_t = w_n \), where \( m \) and \( n \) are indices that represent locations of the two variables at time \( t \). Then at time \( t + \Delta t \), \( X_{t+\Delta t} \) and \( w_{t+\Delta t} \) can be expressed as

\[
X_{t+\Delta t} = \begin{cases} 
X_{m+1}, & \text{up}, \\
X_m, & \text{middle}, \\
X_{m-1}, & \text{down},
\end{cases} \quad w_{t+\Delta t} = \begin{cases} 
w_{n+1}, & \text{up}, \\
w_n, & \text{middle}, \\
w_{n-1}, & \text{down},
\end{cases}
\]

with constant spatial increments \( \Delta X \) and \( \Delta w \). It should be mentioned that the “middle” here represent the branch in between “up” and “down” at each time step, which is not necessarily flat. The constant increments are the core of keeping the tree structure nice and simple. However, as we will see later, the increments of \( S \) and \( v \) are no longer constant, due to the presence of the term \( w_{t+\Delta t} + w_t \) in (6.3.16).

We start with variance \( v_t \) since it depends solely on \( w_t \). In addition, as in (6.3.15), \( v_{t+\Delta t} - v_t \) is also a factor that is included in the increment of \( S \). Since \( v_t \) is monotonically increasing with respect to \( w_t \), it is natural to define the up, middle and down branches of \( v_t \) accordingly. With relations \( w_n = w_{n+1} - \Delta w = w_{n-1} + \Delta w \), we can obtain the
three-jump process for variable \( v \):

\[
v_{t+\Delta t} = \begin{cases} 
  v_n + \frac{1}{4}(2w_n + \Delta w)\Delta w, & \text{up}, \\
  v_n, & \text{middle}, \\
  v_n + \frac{1}{4}(2w_n - \Delta w)\Delta w, & \text{down},
\end{cases} 
\]  

(6.3.17)

It is clear that the increment of \( v_t \) is not constant. Denote \( \Delta v^u_t = \frac{1}{4}(2w_n + \Delta w)\Delta w \) and \( \Delta v^d_t = \frac{1}{4}(2w_n - \Delta w)\Delta w \). Obviously, \( \Delta v^u_t > \Delta v^d_t \), which implies that the upward jump size is greater than the downward jump size.

Now we proceed to variable \( S_t \), which can be viewed as a bivariate function of \( X_t \) and \( v_t \). Similarly, we define the stock price to go up, middle and down according to both the two variables going up, middle and down at the same time. The values \( S \) can take at time \( t + \Delta t \) then can be expressed as

\[
S_{t+\Delta t} = \begin{cases} 
  S_m \exp \left( \Delta X + \frac{\rho \Delta w}{4\sigma} (2w_n + \Delta w) + (r - \frac{\rho \theta}{\sigma})\Delta t \right), & \text{up}, \\
  S_m \exp \left( (r - \frac{\rho \theta}{\sigma})\Delta t \right), & \text{middle}, \\
  S_m \exp \left( -\Delta X - \frac{\rho \Delta w}{4\sigma} (2w_n - \Delta w) + (r - \frac{\rho \theta}{\sigma})\Delta t \right), & \text{down},
\end{cases} 
\]  

(6.3.18)

Similar to \( v_t \), the jump-size of the stock price going upward (denoted by \( \Delta X^u_t \)) is different from the one of jumping downward (denoted by \( \Delta X^d_t \)). Here the sign of the correlation \( \rho \) determines which one is greater.

It is trivial to prove that the original tree whose steplength defined by (6.3.17) and (6.3.18) is recombining. Graphical illustrations are given in Figure 6.5 and Figure 6.6 as the cross-sections of the tree on \( t - v \) and \( t - S \) planes, respectively. Figure 6.5 shows that the jump size grows as the value of \( w \) increases (in the figure from bottom to top). The grid is dense at the bottom and tends to be sparse at the top. On the other hand, the middle branch of the tree in the \( t - S \) plane is not flat due to the increasing factor being \( \exp \left( (r - \frac{\rho \theta}{\sigma})\Delta t \right) \) instead of 1. Together with the two cross-sections, the tree in the original coordinate plane is a twisted and stretched pyramid.

However, a tree with such a shape is not good enough in terms of simplicity, especially when the jump sizes are non-uniform. Thus the advantage of transforming the tree to a new coordinate plane in which the tree structure is simple and has constant jump sizes
can be justified.

Figure 6.5: Graphical illustration for the $t - v$ plane

Figure 6.6: Graphical illustration for the $t - S$ plane

### 6.4 Equivalent to the explicit finite difference

It is well-known that under the BS model, the trinomial tree method can be viewed as a special case of the explicit finite difference method applied to the BS PDE, as presented in [45, 76]. As a basic property, it is reasonable to assume that the equivalence between tree methods and finite difference methods holds for all extended BS models. One
strong evidence is from [90], which provides a proof of the two methods equivalent under regime-switching models. Such as equivalence for the Heston model, however, remains undiscovered since the existing tree-based approximations are too sophisticated. Here due to the simplicity of our tree approach, we manage to build an equivalence. This is given in the theorem below.

**Theorem 6.4.1.** The tree approach (6.3.12) is equivalent to the explicit finite difference method under the condition \( \Delta x = \lambda_n b_n \sqrt{\Delta t} \).

**Proof.** The main idea of the equivalence is that both methods have exactly the same recursive formulas. So in the proof, we focus on the explicit schemes that can be written to be identical to (6.3.12). We start with introducing a dimensionless variable transformation. Let

\[ V = e^{-rt} U. \]

Then PDE (6.3.4) with this transformation becomes

\[
\frac{\partial V}{\partial t} + \frac{1}{8}(1 - \rho^2) w_n^2 \frac{\partial^2 V}{\partial x^2} + \frac{1}{4} \alpha w_n^2 \frac{\partial V}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial w^2} + \phi(w) \frac{\partial V}{\partial w} = 0. \tag{6.4.1}
\]

Then by applying the explicit Euler scheme to (6.4.1), we have

\[
V_{m,n}^{i+1} = A_n V_{m+1,n}^{i+1} + B_n V_{m,n}^{i+1} + C_n V_{m-1,n}^{i+1} + \beta_n \Delta t V_{m+1,n}^{n+1} + \beta_n^+ \Delta t V_{m-1,n}^{n+1}, \tag{6.4.2}
\]

where

\[
A_n = \frac{1}{8} \frac{\Delta t}{\Delta x^2} (1 - \rho^2) w_n^2 + \frac{1}{8} \frac{\Delta t}{\Delta x} \left( \frac{\rho_k}{\sigma} - \frac{1}{2} \right) w_n^2, \\
B_n = 1 - \frac{1}{4} \frac{\Delta t}{\Delta x^2} (1 - \rho^2) w_n^2 + \beta_n \Delta t, \\
C_n = \frac{1}{8} \frac{\Delta t}{\Delta x^2} (1 - \rho^2) w_n^2 - \frac{1}{8} \frac{\Delta t}{\Delta x} \left( \frac{\rho_k}{\sigma} - \frac{1}{2} \right) w_n^2.
\]

Here we need to mention that the coefficients \( \beta_n^-, \beta_n, \beta_n^+ \) are determined by (6.3.5), which applies one of the central, backward, forward finite difference to the first order derivative \( \frac{\partial U}{\partial w} \) so that \( \beta_n^+ \) and \( \beta_n^- \) are positive while \( \beta_n \) is negative.

Substituting \( \Delta x = \lambda_n b_n \sqrt{\Delta t} \) into \( A_n, B_n \) and \( C_n \), it is trivial to find that the following
relations hold

\[ A_n = \phi_n^u, \]
\[ B_n = \phi_n^m + q_{n,n} \Delta t, \]
\[ C_n = \phi_n^d. \]

In addition, since \( q_{n,n+1} = \beta_n^+, q_{n,n} = \beta_n \) and \( q_{n,n-1} = \beta_n^- \), as we defined in Section 6.3.1, we can rewrite (6.4.2) as

\[
V^{i+1}_{m,n} = \phi_n^u V^{i+1}_{m+1,n} + (\phi_n^m + q_{n,n}\Delta t)V^{i+1}_{m,n} + \phi_n^d V^{i+1}_{m-1,n} + q_{n,n-1}\Delta t V^{i+1}_{m,n-1} + q_{n,n+1}\Delta t V^{i+1}_{m,n+1}. \tag{6.4.3}
\]

Note that at the beginning we have \( V = e^{-rt}U \). The discrete version of the transformation is \( V^{i}_{m,n} = e^{-r \Delta t} U^{i}_{m,n} \). Substituting it back to (6.4.3) yields

\[
U^{i}_{m,n} = e^{-r \Delta t} \left( \phi_n^u U^{i+1}_{m+1,n} + (\phi_n^m + q_{n,n}\Delta t)U^{i+1}_{m,n} + \phi_n^d U^{i+1}_{m-1,n} + q_{n,n-1}\Delta t U^{i+1}_{m,n-1} + q_{n,n+1}\Delta t U^{i+1}_{m,n+1} \right),
\]

which is exactly the recursive formula (6.3.12). Thus the proof ends.

The most significant property the equivalence brings is that the convergence rate of the new simple tree approach is specified without proof. This is given in the following proposition.

**Proposition 6.4.2.** The convergence rate of the tree approach (6.3.12) is of \( O(\Delta t) + O((\Delta X)^2) + O(\Delta w) \).

It is worthwhile mentioning that the convergence rate along the \( w \) direction is only of first order. This is because of the forward and backward Euler scheme we apply to the first order derivative with respect to variable \( w \), which sacrifices one degree of convergence order. In Section 6.5, we numerically verify the convergence of the new simple tree approach.

Now the two methods are perfectly matched with the equivalence we have just presented. However, the relation is built on the transformed Heston model so the equivalence is “indirect”. To better illustrate the complex relations, we put the four main objects in a
flowchart in Figure 6.7, in which both SDEs and PDEs are connected by the transformation (6.3.1) while the transformed tree and finite difference (FD in the figure) is connected by Theorem 6.4.1. The equivalence (“IE” in the figure) is obtained by “Heston’s SDE (Tree)” $\xrightarrow{\text{Eq}(6.3.1)}$ “Transformed SDE (Tree)” $\xrightarrow{\text{Theorem 6.4.1}}$ “Transformed PDE (FD)” $\xrightarrow{\text{Eq}(6.3.1)}$ “Heston’s PDE (FD)”.

Since it is a detour path, we name it “indirect equivalence” between the two methods.

6.5 Numerical performance

In this section, we apply our method to both European and American types of options in a number of numerical experiments. We compare our methods to Liu’s approach [53] and the results are shown in tables. We also numerically test the “indirect equivalence” that we set up between our tree approach and the explicit finite difference method.

6.5.1 European options

Under the Heston model, European options can be priced in a semi-analytical formula [39], which is compromised of two inverse Fourier transform integrals. It should be remarked that attention has to be paid to the integrands as they sometimes produce high oscillations that dampen extremely slowly along the integration axis, as reported in [74]. To fix the problem, one can apply either the equivalent formulation from Albrecher et al. [1] or the “Rotation Counts Algorithm” presented by Lord and Kahl [58].

In the first experiment, we test our method against the analytical solution as well as the tree-based approach from Liu [53]. To control all the variates while doing the experiments, we use exactly the same set of parameters from the example in [53]. The parameters are $r = 0.05$, $\rho = -0.1$, $\kappa = 3.0$, $\theta = 0.04$, $\sigma = 0.1$, $E = \$100$. There are a total of 12 different
cases are taken into consideration, including two different time to maturities $T = 0.25, 0.5,$
three initial stock prices $S_0 = 90, 100, 110$ and two initial variances $v_0 = 0.04, 0.09$.

The time increment is chosen to be $\Delta t = 10^{-3}$ while the space increment $\Delta w = 0.02$. The upper and lower bound $n_l$ and $N$ are selected as $n_l = 15, N = 40$. Hence the space range is $[0.0225; 0.16]$, with a total of 26 regimes taken into consideration. The results of Experiment 1 are presented in Table 6.1.

![Table 6.1](image)

All the results provided in Table 6.1 are European call options. The “L-Tree” column in the table means the results from the column is obtained from Liu’s tree method while “Z-Tree” stands for our tree approach. It can be verified from the table that Liu’s tree is not performing well in the In-The-Money (ITM) case, as the errors go up to 0.0045. On the other hand, our tree approach is stable throughout all cases, with errors being less than 0.0010.

In the second experiment, we test the convergence behaviour of our method. The method is discretized along both $t$-axis and $w$-axis so it requires being discussed separately. Here we start with the convergence behaviour with respect to the time discretization. To have a better idea about how the method converges as the time increment goes to zero, we adopt two sets of parameters from what we mentioned above (except for $S_0 = 100$, $v_0 = 0.04$ and $T = 0.25$ and 0.5 for each case). Note that the only difference between the two cases is the time to maturity.
According to Table 6.2 and 6.3, for each time the number of time step is doubled, the error compared to the analytical solution (Diff column in both tables) roughly reduced to a half. This verifies the convergence order of the method is of order $O(\Delta t)$ in terms of the $t$-direction. More illustration can be found in Figure 6.8 and Figure 6.9, in which numbers of time steps are allowed to vary from 120 to 1200 with a increment of 50. As we can see from the figures, both curves slide down from upper left to lower right, which further proves that the method converges to the real solution as $N_T$ goes to infinity, although the patterns have slight oscillatory.

Then we proceed to the convergence with respect to variable $w$. We first fix $\Delta w$ and check how the number of regimes influences on the convergence. Again we start with the same set of parameters ($S_0 = E = $100, $r = 0.05$, $\rho = -0.1$, $\kappa = 3.0$, $\theta = 0.04$, $\sigma = 0.1$, $T = 0.25$, $v_0 = 0.04$, $\Delta w = 0.02$). Here we choose the $N_T$ to be 1280 since it provides the best results in Table 6.2 and 6.3. It can be simply calculated that the solution of equation $f(n) = 0$ is $\bar{n} = 19.7906$, which means $n_l \leq 19$ and $N \geq 20$.

In Table 6.4, we approximate European call option prices with different numbers of regimes (NOR column in the table) from 2 to 32, with the latter being double the amount
6.5. NUMERICAL PERFORMANCE

<table>
<thead>
<tr>
<th>$n_l$</th>
<th>$N$</th>
<th>NOR</th>
<th>Range</th>
<th>Tree</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>20</td>
<td>2</td>
<td>[0.0361, 0.0400]</td>
<td>4.616048</td>
<td>0.005711</td>
</tr>
<tr>
<td>18</td>
<td>21</td>
<td>4</td>
<td>[0.0324, 0.0441]</td>
<td>4.612996</td>
<td>0.002658</td>
</tr>
<tr>
<td>16</td>
<td>23</td>
<td>8</td>
<td>[0.0256, 0.0529]</td>
<td>4.610412</td>
<td>0.000074</td>
</tr>
<tr>
<td>12</td>
<td>27</td>
<td>16</td>
<td>[0.0144, 0.0729]</td>
<td>4.610464</td>
<td>0.000127</td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td>32</td>
<td>[0.0016, 0.1225]</td>
<td>4.610139</td>
<td>0.000199</td>
</tr>
</tbody>
</table>

Table 6.4: $\Delta w = 0.02$

of the former. $n_l$ and $N$ vary with $\bar{n}$ being the centre. The range column represents the truncated volatility interval instead of variable $w$. According to Table 6.4, the method converges rapidly with respect to $w$ and start to fluctuate.

For comparison, we remain all parameters to be the same but change $\Delta w$ to 0.01. As $\Delta w$ is changed, the solution of $f(n) = 0$ becomes $\bar{n} = 39.5811$. So this time we have $n_l \leq 39$ while $N \geq 40$.

<table>
<thead>
<tr>
<th>$n_l$</th>
<th>$N$</th>
<th>NOR</th>
<th>Range</th>
<th>Tree</th>
<th>Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>39</td>
<td>40</td>
<td>2</td>
<td>[0.0380, 0.0400]</td>
<td>4.616082</td>
<td>0.005745</td>
</tr>
<tr>
<td>38</td>
<td>41</td>
<td>4</td>
<td>[0.0361, 0.0420]</td>
<td>4.615237</td>
<td>0.004900</td>
</tr>
<tr>
<td>36</td>
<td>43</td>
<td>8</td>
<td>[0.0324, 0.0462]</td>
<td>4.612440</td>
<td>0.002102</td>
</tr>
<tr>
<td>32</td>
<td>47</td>
<td>16</td>
<td>[0.0256, 0.0552]</td>
<td>4.610527</td>
<td>0.000189</td>
</tr>
<tr>
<td>24</td>
<td>55</td>
<td>32</td>
<td>[0.0144, 0.0756]</td>
<td>4.610344</td>
<td>0.000007</td>
</tr>
<tr>
<td>8</td>
<td>71</td>
<td>64</td>
<td>[0.0016, 0.1260]</td>
<td>4.610125</td>
<td>0.000213</td>
</tr>
</tbody>
</table>

Table 6.5: $\Delta w = 0.01$

In summary, not too many regimes are required to maintain good results. From our experiments, it turns out that 12 regimes are generally good enough. The option values start to fluctuates slightly after the number of regimes is sufficiently large. Practically the optimal number of regimes can be obtained by experimenting the algorithm a few times.

The biggest advantage of our method is that it is computationally cheap. The computational time of our algorithm with 1280 time steps and 8 regimes cost only 1.150 seconds, in contrast to 2.162 seconds for only $N = 200$ steps of [3] with results from our algorithm outperforming theirs.

6.5.2 American options

One main merit of tree methods is that it can be easily used to price American options. In contrast to pricing European options, at each node, the nodal value obtained from the recursive formula (3.13) has to be compared to the instant exercise value, which is the payoff for exercising at that very point. The larger one of the two values will be involved
in the next-step calculation. To be more specific, for an American put option whose price is denoted as $\bar{U}$, we have

$$\bar{U}_{i,m;n}^i = \max \{ U_{i,m;n}^i, f(U_{i,m;n}^i) \},$$

where

$$f(U_{i,m;n}^i) = \max \left\{ E S_0 \exp \left( x_m \frac{\rho}{\sigma} (v_n - v_0) + \left( r - \frac{\rho \theta}{\sigma} \right) i \Delta t \right), 0 \right\},$$

and $U_{i,m;n}^i$ is obtained from the recursive formula.

Again we compare our tree method with Liu’s tree approach. Similarly, we set the parameters to be exactly the same as Experiment 1. Benchmark results are obtained from [3] with 200 steps in each space dimension with control variate technique as the chapter mentioned. Like the case of European options, results obtained from our tree are mostly closer to the benchmark solution than Liu’s tree. The comparison is given in Table 6.6 below.

<table>
<thead>
<tr>
<th>$S_0$</th>
<th>$v_0$</th>
<th>$\Delta t$</th>
<th>L-Tree(error)</th>
<th>Z-Tree(error)</th>
<th>Benchmark</th>
</tr>
</thead>
<tbody>
<tr>
<td>90</td>
<td>0.04</td>
<td>0.25</td>
<td>10.1719(-0.0008)</td>
<td>10.1718(-0.0007)</td>
<td>10.1711</td>
</tr>
<tr>
<td>100</td>
<td>0.04</td>
<td>0.25</td>
<td>3.4753(-0.0005)</td>
<td>3.4750(-0.0002)</td>
<td>3.4748</td>
</tr>
<tr>
<td>110</td>
<td>0.04</td>
<td>0.25</td>
<td>0.7738(-0.0002)</td>
<td>0.7736(0.0000)</td>
<td>0.7736</td>
</tr>
<tr>
<td>90</td>
<td>0.09</td>
<td>0.25</td>
<td>11.0316(-0.0092)</td>
<td>11.0291(-0.0067)</td>
<td>11.0224</td>
</tr>
<tr>
<td>100</td>
<td>0.09</td>
<td>0.25</td>
<td>4.9567(-0.0115)</td>
<td>4.9543(-0.0091)</td>
<td>4.9542</td>
</tr>
<tr>
<td>110</td>
<td>0.09</td>
<td>0.25</td>
<td>1.8095(-0.0111)</td>
<td>1.8081(-0.0097)</td>
<td>1.7984</td>
</tr>
<tr>
<td>90</td>
<td>0.04</td>
<td>0.5</td>
<td>10.6501(-0.0019)</td>
<td>10.6499(-0.0017)</td>
<td>10.6482</td>
</tr>
<tr>
<td>100</td>
<td>0.04</td>
<td>0.5</td>
<td>4.6485(-0.0012)</td>
<td>4.6484(-0.0011)</td>
<td>4.6473</td>
</tr>
<tr>
<td>110</td>
<td>0.04</td>
<td>0.5</td>
<td>1.6837(-0.0005)</td>
<td>1.6837(-0.0005)</td>
<td>1.6832</td>
</tr>
<tr>
<td>90</td>
<td>0.09</td>
<td>0.5</td>
<td>11.8634(-0.0117)</td>
<td>11.8600(-0.0083)</td>
<td>11.8517</td>
</tr>
<tr>
<td>100</td>
<td>0.09</td>
<td>0.5</td>
<td>6.2629(0.0131)</td>
<td>6.2596(-0.0098)</td>
<td>6.2498</td>
</tr>
<tr>
<td>110</td>
<td>0.09</td>
<td>0.5</td>
<td>2.9851(-0.0124)</td>
<td>2.9828(-0.0101)</td>
<td>2.9727</td>
</tr>
</tbody>
</table>

Table 6.6: Comparing our results with Liu’s approach for American put options

6.5.3 Equivalence test

In the end, we present the numerical proof for the equivalence we have discussed in Section 4. To apply the explicit finite difference scheme, we need to specify the boundary conditions. Since the $t - X$ plane is an ordinary trinomial tree, the boundary condition for the
X variable is thus freed as the space increment $\Delta x$ is dependent on the time increment $\Delta t$. However, as the variable $w$ is discretized, boundary conditions for $w$ are required. A simple boundary condition can be given by the smoothness of $U$ with respect to $w$:

$$U_{m,n,t}^i = U_{m,n,t+1}^i, \quad U_{m,N}^i = U_{m,N-1}^i$$

Hence the explicit finite difference scheme can be applied and the comparison of the two methods is denoted as Experiment 3, in which a set of parameters are given as $S_0 = E = 40$, $\theta = 0.06$, $\sigma = 0.2$, $r = 0.04$, $\rho = -0.2$, $\nu_0 = 0.03$, $\kappa = 2.5$ with the time to maturity $T = 0.25$ years. We consider one European call option and one European put option with a range of different numbers of time steps to prove the equivalence is robust and stable along the time axis. The number of time steps $N_t$ is selected from the set \{20, 40, 80, 160, 320, 640, 1280, 2560\}.

<table>
<thead>
<tr>
<th>$N_t$</th>
<th>Put</th>
<th>Call</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>2.6645e-14</td>
<td>3.9080e-14</td>
</tr>
<tr>
<td>40</td>
<td>4.4409e-16</td>
<td>2.2204e-16</td>
</tr>
<tr>
<td>80</td>
<td>1.1102e-15</td>
<td>1.7764e-15</td>
</tr>
<tr>
<td>160</td>
<td>6.6613e-16</td>
<td>4.4409e-16</td>
</tr>
<tr>
<td>320</td>
<td>3.3307e-15</td>
<td>5.7732e-15</td>
</tr>
<tr>
<td>640</td>
<td>3.1086e-15</td>
<td>3.1086e-15</td>
</tr>
<tr>
<td>1280</td>
<td>1.2434e-14</td>
<td>2.3093e-14</td>
</tr>
<tr>
<td>2560</td>
<td>7.7716e-15</td>
<td>3.9968e-15</td>
</tr>
</tbody>
</table>

Table 6.7: Numerical examples for comparing the tree method with the explicit finite difference method

In Table 6.7, differences between the two methods are given for both European call and put options. From the table, errors fluctuate between $10^{-14}$ to $10^{-16}$. This can be interpreted as differences are compromised of rounding errors only.

### 6.6 Conclusion

We develop a new simple tree approach for pricing options under Heston’s stochastic volatility model. The new tree-based method is very simple to implement, computationally cheap and has a clear financial interpretation. In addition, we also find an indirect way to present an equivalence theorem of our tree method with explicit finite difference methods under the Heston model. Numerical performance has provided the evidence that our new tree approach is robust and efficient.
Chapter 7

Conclusions

In this thesis, we have explored three main numerical techniques, Monte Carlo simulations, tree approaches, and finite difference methods, when they are used in options pricing under regime switching models and Heston’s stochastic volatility models.

The thesis contains two parts. In the first part, we study the three methods under the regime-switching model while the connections between tree approaches and explicit finite-difference methods are investigated in the second part.

The first part starts with Chapter 3, in which we present a comparative study of Monte Carlo simulations and Crank-Nicolson finite-difference method under the regime-switching model. The numerical performance shows that the Monte Carlo simulation is generally more efficient than the finite-difference method even with one underlying stock having two regimes, which is the simplest case. The only exception occurs where the switching frequency (parameter $\lambda$ in the model) is sufficiently large. But such a large $\lambda$ can only be taken mathematically as it does not exist in the practical world. Even though the computational cost of the two methods grows linearly with respect to the number of regimes, the increasing rate of the finite-difference method is shown significantly greater than the Monte Carlo simulation. The chapter also shows that in terms of variation reduction techniques, the antithetic variates method has been tested to outperform the control variates in the regime-switching world.

Chapter 4 compromises the convergence analysis for tree approaches for the regime-switching model. The convergence rate of Yuen and Yang’s trinomial tree is proven to be of order $O(n^{-\beta})$, where $\beta = 1/2$ if the payoff function of the option contract is discontinuous and $\beta = 1$ elsewhere. A European put option and a digital put option are taken as
examples to verify our proof. As shown in the numerical experiments, the convergence rate for the case of the European put option is of order $O(n^{-1})$ as its payoff function $\max\{K - S, 0\}$ is continuous while the corresponding rate for digital put option is $O(n^{-1/2})$, since the payoff for a digital put options is either one or zero, which is discontinuous.

The second part start with Chapter 5, in which we first present a new trinomial tree method for regime-switching models in which only a total of $K + 2$ ($K$ is the number of regimes) nodal values are involved for each time step instead of $3K$ in Yuen and Yang’s tree. Then the new method is proven analytically to be equivalent to the corresponding explicit finite-difference method, under a lognormal transformation. The numerical results show that the new method outperforms Yuen and Yang’s tree and the equivalence is robust and stable with different numbers of time partitions.

The new trinomial tree approach is then used to become a new simple tree approach for Heston’s stochastic volatility model, as presented in Chapter 6. Based on Liu’s approach, we apply the new trinomial tree method to the regime-switching model which is obtained by transforming and discretizing the Heston model. The new simple tree method has very clear advantages in simplicity as well as economic interpretation. Again, an equivalent has been established between the new tree method and the explicit finite-difference method for the transformed Heston’s PDE. From the numerical experiment, the new simple tree method is more efficient than the original approach by Liu in pricing both European and American options. The equivalence is also verified and presented in the numerical results in the chapter.
Bibliography


[34] Markus Hahn, Sylvia Frühwirth-Schnatter, and Jörn Sass. Markov chain monte


Appendix A

Algorithms

This appendix contains all the algorithms mentioned in Chapter 3.
Algorithm 1 Fundamental Monte Carlo simulation

function FMC\((S_0, E, r, G, T, \sigma, j, N)\):

**Require:**
1. Initial stock price \(S_0\);
2. Strike price \(E\);
3. Interest rates \(r\);
4. Generator matrix of the Markov chain \(G = (\lambda_{mn}) \in \mathbb{R}^{K \times K}\);
5. Time to maturity \(T\);
6. Volatilities vector \(\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_K)\);
7. Initial regime \(j\) \((j = 1, 2, \ldots, K)\);
8. Numbers of simulation paths \(N\);

**Ensure:**
9. European put option price with respect to initial regime \(j\);
10. for \(n = 1\) to \(N\) do
11. for \(k = 1\) to \(K\) do
12. \(J[k] \leftarrow 0\);
13. end for
14. \(t \leftarrow T; I \leftarrow j; Q \leftarrow 0\);
15. while \(t > 0\) do
16. Generate \(U_1, U_2 \sim \text{uniform}(0, 1)\);
17. \(\tau \leftarrow \log(U_1)/\lambda_{II}\);
18. \(pr \leftarrow 0; m \leftarrow 1\);
19. while \(U_2 > pr\) do
20. if \(m \neq I\) then
21. \(pr \leftarrow pr - \lambda_{Im}/\lambda_{II}\);
22. else
23. \(m \leftarrow m + 1\);
24. end if
25. end while
26. if \(\tau > t\) then
27. \(J[I] \leftarrow J[I] + t\);
28. else
29. \(J[I] \leftarrow J[I] + \tau\);
30. end if
31. \(t \leftarrow t - \tau; I \leftarrow m - 1\);
32. end while
33. for \(k = 1\) to \(K\) do
34. \(Q \leftarrow Q + \sigma_k^2 J[k]\);
35. end for
36. Generate \(W \sim \text{normal}(0, 1)\);
37. \(Z[n] \leftarrow S_0\exp\{(rT - Q/2) + \sqrt{Q}W\}\);
38. \(Y[n] \leftarrow e^{-rT}\max(E - Z[n], 0)\);
39. end for
40. \(V = \frac{1}{N} \sum_{n=1}^{N} Y[n]\);
41. return \(V\);
Algorithm 2 Monte Carlo with antithetic variates

37: \( Z_{ori}[n] \leftarrow S_0 \exp\{(rT - Q/2) + \sqrt{Q}W\}; \ Z_{ant}[n] \leftarrow S_0 \exp\{(rT - Q/2) - \sqrt{Q}W\} \)
38: \( Y[n] \leftarrow \frac{1}{2} e^{-rT}(\max(E - Z_{ori}[n], 0) + \max(E - Z_{ant}[n], 0)); \)
39: \( V = \frac{1}{N} \sum_{n=1}^{N} Y[n]; \)
40: return \( V; \)

Algorithm 3 Monte Carlo with control variates

40: \( \bar{Z} = \frac{1}{N} \sum_{n=1}^{N} Z[n]; \ \bar{Y} = \frac{1}{N} \sum_{n=1}^{N} Y[n]; \)
41: \( b = \frac{\sum_{n=1}^{N} (Z[n] - \bar{Z})(Y[n] - \bar{Y})}{\sum_{n=1}^{N} (Z[n] - \bar{Z})^2}; \)
42: \( V = \frac{1}{N} \sum_{n=1}^{N} (Y[n] - b(Z[n] - e^{rT}S_0)); \)
43: return \( V; \)

Algorithm 4 Generating random numbers from the \( f_{j|i} \) function GRN \((\lambda_{ji}, \lambda_{ij}, T, M)\);

function \( \text{GRN}(\lambda_{ji}, \lambda_{ij}, T, M); \)

Require:
1: Jump intensity \( \lambda_{ji}, \lambda_{ij}; \)
2: Time to maturity \( T; \)
3: Numbers of time partitions \( M; \)
Ensure:
4: A random number following the pdf \( f_{j|i}; \)
5: \( t[0] \leftarrow 0; \)
6: \( \Delta t = T/M; \)
7: for \( m = 1 \) to \( M \) do
8: \( t[m] \leftarrow t[m-1] + \Delta t; \)
9: \( f_{j}[m-1] \leftarrow e^{-\lambda_{ji}(T-t[m-1]) - \lambda_{ij}t[m-1]} \lambda_{ji} I_0(2(\lambda_{ji} \lambda_{ij} t(T-t[m-1]))^{1/2}) + \lambda_{ij} I_1(2(\lambda_{ji} \lambda_{ij} t(T-t[m-1]))^{1/2}); \)
10: end for
11: \( f_{j}[M] \leftarrow f_{j}[M-1] + e^{-\lambda_{ji}T}; \)
12: \( B \leftarrow 0; \)
13: for \( m = 0 \) to \( M \) do
14: \( B \leftarrow B + f_{j}[m]; \)
15: \( c_j[m] \leftarrow BT/M; \)
16: end for
17: Generate \( U \sim \text{uniform}(0, 1); \)
18: \( n \leftarrow 0; \)
19: while \( U > c_j[n] \) do
20: \( n \leftarrow n + 1; \)
21: end while
22: return \( t[n]; \)
Algorithm 5 Simulating total occupation time

\textbf{function TOT} \((S_0, E, r, G, T, \sigma, j, N, M)\);

\textbf{Require:}
1: Initial stock price \(S_0\);
2: Strike price \(E\);
3: Interest rates \(r\);
4: Generator matrix of the Markov chain \(G = (\lambda_{mn}) \in \mathbb{R}^{2 \times 2}\);
5: Time to maturity \(T\);
6: Volatilities vector \(\sigma = (\sigma_1, \sigma_2)\);
7: Initial regime \(j\) (\(j = 1, 2\));
8: Numbers of simulation paths \(N\);
9: Number of time partition for generating random variable from the pdf \(f_j\) \(M\);

\textbf{Ensure:}
10: European put option price with respect to initial regime \(j\);
11: \textbf{for} \(n = 1\) to \(N\) \textbf{do}
12: \hspace{1em} \(J \leftarrow \text{GRN}(\lambda_{ji}, \lambda_{ij}, T, M)\);
13: \hspace{1em} \(I \leftarrow T - J\);
14: \hspace{1em} \(Q \leftarrow \sigma_j^2 J + \sigma_i^2 I\);
15: \hspace{1em} Generate \(W \sim \text{normal}(0, 1)\);
16: \hspace{1em} \(Z[n] = S_0 \exp\{\left(rT - U/2\right) + \sqrt{U}W\}\);
17: \hspace{1em} \(Y[n] = e^{-rT} \max(E - Z[n], 0)\);
18: \textbf{end for}
19: \(V = \frac{1}{N} \sum_{n=1}^{N} Y[n]\);
20: \textbf{return} \(V\);
Appendix B

Proof of theorems

B.1 Proof of Theorem 5.5.1

The idea of the proof is from Ma and Zhu [60] and Ma and Zhu [59].

We start by introducing a definition of local remainder

\[ R^m_j = V(S_m, t_n, j) - e^{-r\Delta t} \left[ \pi_j^U V(S_{m+1}, t_{n+1}, j) + \pi_j^M - \sum_{i \neq j}^{K} \lambda_{ji}\Delta t V(S_m, t_{n+1}, j) \right. \]

\[ \left. + \pi_j^D V(S_{m-1}, t_{n+1}, j) + \sum_{i=1}^{K} V(S_m, t_{n+1}, i) \right] , \quad (B.1) \]

where \( V(S_m, t_n, j) \) is the analytical value of the option price in regime \( j \) given \( S = m\Delta S, t = n\Delta t \). For the local remainder, we have the following proposition.

**Proposition B.1.1.** \( R^m_j \) defined by (B.1) is estimated by \( R^m_j = O \left((\Delta t)^2 \right) \).

**Proof.** Apply Taylor’s expansion to \( V(S_m, t_n, j) \) at time \( t = t_{n+1} \), we obtain

\[ V(S_m, t_n, j) = V(S_m, t_{n+1}, j) - \frac{\partial V(S_m, t_{n+1}, j)}{\partial t} \Delta t + O \left( (\Delta t)^2 \right) . \quad (B.1.2) \]

Do the same to \( V(S_{m+1}, t_{n+1}, j) \) and \( V(S_{m-1}, t_{n+1}, j) \) at the point \( S = S_m \)

\[ V(S_{m+1}, t_{n+1}, j) = V(S_m, t_{n+1}, j) + \frac{\partial V(S_m, t_{n+1}, j)}{\partial S} S_m (U - 1) \]

\[ + \frac{1}{2} \frac{\partial^2 V(S_m, t_{n+1}, j)}{\partial^2 S} S_m^2 (U - 1)^2 + \frac{1}{6} \frac{\partial^3 V(S_m, t_{n+1}, j)}{\partial^3 S} S_m^3 (U - 1)^3 + O \left( (U - 1)^4 \right) , \quad (B.1.3) \]
and

\[
V(S_{m-1}, t_{n+1}, j) = V(S_m, t_{n+1}, j) - \frac{\partial V(S_m, t_{n+1}, j)}{\partial S} S_m(D - 1) \\
+ \frac{1}{2} \frac{\partial^2 V(S_m, t_{n+1}, j)}{\partial^2 S} S_m^2 (D - 1)^2 - \frac{1}{6} \frac{\partial^3 V(S_m, t_{n+1}, j)}{\partial^3 S} S_m^3 (D - 1)^3 + O \left( (D - 1)^4 \right).
\]

(B.1.4)

Substitute (B.1.2) to (B.1.4) into the definition of the local remainder (B.1.1),

\[
R^n_j = V(S_m, t_{n+1}, j) - \frac{\partial V(S_m, t_{n+1}, j)}{\partial t} \Delta t + O \left( (\Delta t)^2 \right) \\
- e^{-r\Delta t} \left[ (1 - \sum_{i=1}^{K} \lambda_{ji} \Delta t) V(S_m, t_{n+1}, j) + \sum_{i=1}^{K} \lambda_{ji} \Delta t V(S_m, t_{n+1}, i) + A^n_j + B^n_j + C^n_j \\
+ O \left( (U - 1)^4 \right) + O \left( (D - 1)^4 \right) \right],
\]

(B.1.5)

where

\[
A^n_j = \frac{\partial V(S_m, t_{n+1}, j)}{\partial S} S_m \left[ \pi_j^U(U - 1) + \pi_j^D(D - 1) \right],
\]

(B.1.6)

\[
B^n_j = \frac{\partial^2 V(S_m, t_{n+1}, j)}{\partial^2 S} S_m^2 \left[ \pi_j^U(U - 1)^2 + \pi_j^D(D - 1)^2 \right],
\]

(B.1.7)

\[
C^n_j = \frac{\partial^3 V(S_m, t_{n+1}, j)}{\partial^3 S} S_m^3 \left[ \pi_j^U(U - 1)^3 + \pi_j^D(D - 1)^3 \right].
\]

(B.1.8)

From the Kamraud-Mitchken parametrization, the following equations hold

\[
\pi_j^U(U - 1) + \pi_j^D(D - 1) = r \Delta t + O \left( (\Delta t)^2 \right),
\]

(B.1.9)

\[
\pi_j^U(U - 1)^2 + \pi_j^D(D - 1)^2 = \sigma_i^2 \Delta t + O \left( (\Delta t)^2 \right),
\]

(B.1.10)

\[
\pi_j^U(U - 1)^3 + \pi_j^D(D - 1)^3 = O \left( (\Delta t)^2 \right).
\]

(B.1.11)

These three equations can be proved trivially by substituting the parameters in so that
we omit it here. Inserting the results from (B.1.6) to (B.1.11) back to (B.1.5) gives that

\[ R^n_j = V(S_m, t_{n+1}, j) - \frac{\partial V(S_m, t_{n+1}, j)}{\partial t} \Delta t + \mathcal{O}((\Delta t)^2) \]

\[ - e^{-r\Delta t} \left[ (1 - \sum_{i=1}^{K} \lambda_{ji} \Delta t)V(S_m, t_{n+1}, j) + \sum_{i=1}^{K} \lambda_{ji} \Delta t V(S_m, t_{n+1}, i) + rS_m \frac{\partial V(S_m, t_{n+1}, j)}{\partial S} \Delta t \right. \]

\[ + \frac{1}{2} \sigma_j^2 S_m \frac{\partial^2 V(S_m, t_{n+1}, j)}{\partial^2 S} \Delta t + + \mathcal{O}((\Delta t)^2) + \mathcal{O}\left((U - 1)^4 + (D - 1)^4\right) \]  

(B.1.12)

Because of the following relations,

\[ e^{-r\Delta t} = 1 - r\Delta t + \mathcal{O}((\Delta t)^2), \]

\[ \mathcal{O}((U - 1)^4) = \mathcal{O}((\Delta t)^2), \]

\[ \mathcal{O}((D - 1)^4) = \mathcal{O}((\Delta t)^2). \]

(B.1.12) can be further simplified by

\[ R^n_j = \mathcal{O}((\Delta t)^2) - \left[ \frac{\partial V(S_m, t_{n+1}, j)}{\partial t} + \frac{1}{2} \sigma_j^2 S_m \frac{\partial^2 V(S_m, t_{n+1}, j)}{\partial^2 S} \right. \]

\[ \left. + rS_m \frac{\partial V(S_m, t_{n+1}, j)}{\partial S} - rV(S_m, t_{n+1}, j) - \sum_{i=1, i \neq j}^{K} \lambda_{ji} \left( V(S_m, t_{n+1}, j) - V(S_m, t_{n+1}, i) \right) \right] \Delta t. \]  

(B.1.13)

As we see the terms in the bracket of (B.1.13) is exactly the governing PDE of the regime-switching diffusion process when the initial economic regime is in state j. Thus \( R^n_j = \mathcal{O}((\Delta t)^2). \)

Once the local remainder is determined, according to (B.1.1), we can have

\[ V(S_m, t_n, j) = R^n_j + e^{-r\Delta t} \left[ \pi^U_j V(S_{m+1}, t_{n+1}, j) + \left( \pi^M_j - \sum_{i=1, i \neq j}^{K} \lambda_{ji} \Delta t \right) V(S_m, t_{n+1}, j) + \right. \]

\[ \left. \pi^D_j V(S_{m-1}, t_{n+1}, j) + \sum_{i=1, i \neq j}^{K} V(S_m, t_{n+1}, i) \right], \]  

(B.1.14)
then subtracting (B.1.14) by (5.5.1) in Theorem 5.5.1 gives that

\[
\epsilon_j^n(S_m) = R_j^n + e^{-r\Delta t}\left[\pi_j^L \epsilon_j^{n+1}(S_{m+1}) + (\pi_j^M - \sum_{i=1}^{K} \lambda_{ji}\Delta t)\epsilon_j^{n+1}(S_m) + \cdots \right. \\
\pi_j^D \epsilon_j^{n+1}(S_{m-1}) + \sum_{i=1}^{K} \epsilon_i^{n+1}(S_m) \left. \right].
\]

(B.1.15)

Therefore, by the definition of \(\|\epsilon_j^n\|_\infty\) we can derive that

\[
\|\epsilon_j^n\|_\infty \leq |\mathcal{O}((\Delta t)^2)| + e^{-r\Delta t}\left(\lambda_{jj}\Delta t\|\epsilon_j^{n+1}\|_\infty + \sum_{i=1}^{K} \lambda_{ji}\Delta t\|\epsilon_i^{n+1}\|_\infty \right),
\]

(B.1.16)

given the fact that

\[
\lambda_{jj} = 1 - \sum_{i \neq j}^{K} \lambda_{ji}.
\]

This relationship holds for all \(j = 1, 2, \ldots, K\) at time \(t_n\). Thus a general matrix form of (A-16) can be written as

\[
\begin{pmatrix}
\|\epsilon_1^n\|_\infty \\
\|\epsilon_2^n\|_\infty \\
\vdots \\
\|\epsilon_K^n\|_\infty
\end{pmatrix} \leq e^{-r\Delta t} \begin{pmatrix}
\lambda_{11}\Delta t & \lambda_{12}\Delta t & \cdots & \lambda_{1K}\Delta t \\
\lambda_{21}\Delta t & \lambda_{22}\Delta t & \cdots & \lambda_{2K}\Delta t \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_{K1}\Delta t & \lambda_{K2}\Delta t & \cdots & \lambda_{KK}\Delta t
\end{pmatrix} \begin{pmatrix}
\|\epsilon_1^{n+1}\|_\infty \\
\|\epsilon_2^{n+1}\|_\infty \\
\vdots \\
\|\epsilon_K^{n+1}\|_\infty
\end{pmatrix} + \begin{pmatrix}
|\mathcal{O}((\Delta t)^2)| \\
|\mathcal{O}((\Delta t)^2)| \\
\vdots \\
|\mathcal{O}((\Delta t)^2)|
\end{pmatrix}.
\]

(B.1.17)

Denote the coefficient matrix by \(C\)

\[
C = e^{-r\Delta t} \begin{pmatrix}
\lambda_{11}\Delta t & \lambda_{12}\Delta t & \cdots & \lambda_{1K}\Delta t \\
\lambda_{21}\Delta t & \lambda_{22}\Delta t & \cdots & \lambda_{2K}\Delta t \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_{K1}\Delta t & \lambda_{K2}\Delta t & \cdots & \lambda_{KK}\Delta t
\end{pmatrix}.
\]
Then by applying Taylor’s expansion
\[ e^{-r\Delta t} = 1 - r\Delta t + \mathcal{O} ((\Delta t)^2), \]
matrix \( C \) can be approximated by
\[
C = \begin{pmatrix}
1 - (r + q_1)\Delta t & \lambda_{12}\Delta t & \cdots & \lambda_{1K}\Delta t \\
\lambda_{21}\Delta t & 1 - (r + q_2)\Delta t & \cdots & \lambda_{2K}\Delta t \\
\vdots & \vdots & \ddots & \vdots \\
\lambda_{K1}\Delta t & \lambda_{K2}\Delta t & \cdots & 1 - (r + q_K)\Delta t
\end{pmatrix},
\]
where
\[ q_j = \sum_{i=1}^{K} \lambda_{ji}. \]
Iterating inequality (B.1.17) for \( N \) times gives that
\[
\begin{pmatrix}
\|\epsilon^0_1\|_\infty \\
\|\epsilon^0_2\|_\infty \\
\vdots \\
\|\epsilon^0_K\|_\infty
\end{pmatrix} \leq C^N \begin{pmatrix}
\|\epsilon^N_1\|_\infty \\
\|\epsilon^N_2\|_\infty \\
\vdots \\
\|\epsilon^N_K\|_\infty
\end{pmatrix} + \left( I + \sum_{p=1}^{N-1} C^p \right) \begin{pmatrix}
\|\mathcal{O} ((\Delta t)^2)\|_\infty \\
\|\mathcal{O} ((\Delta t)^2)\|_\infty \\
\vdots \\
\|\mathcal{O} ((\Delta t)^2)\|_\infty
\end{pmatrix}, \tag{B.1.18}
\]
where \( I \) is \( K \times K \) identity matrix. Since \( \|\epsilon^N_j\|_\infty \) is the error for the price of the financial derivative at \( t = t_N = T \), which is the terminal condition, the first term of the right hand side (RHS) of inequality (B.1.18) becomes zero vector. Therefore, the second term determines the upper bound of \( \|\epsilon^0_j\|_\infty \). Note that
\[
I + \sum_{p=1}^{N-1} C^p = \begin{pmatrix}
N - \frac{N(N-1)}{2} (r + q_1)\Delta t & \frac{N(N-1)}{2} \lambda_{12}\Delta t & \cdots & \frac{N(N-1)}{2} \lambda_{1K}\Delta t \\
\frac{N(N-1)}{2} \lambda_{21}\Delta t & N - \frac{N(N-1)}{2} (r + q_2)\Delta t & \cdots & \frac{N(N-1)}{2} \lambda_{2K}\Delta t \\
\vdots & \vdots & \ddots & \vdots \\
\frac{N(N-1)}{2} \lambda_{K1}\Delta t & \frac{N(N-1)}{2} \lambda_{K2}\Delta t & \cdots & k - \frac{N(N-1)}{2} (r + q_K)\Delta t
\end{pmatrix}, \tag{B.1.19}
and \( \Delta t = T/N \), (B.1.18) becomes

\[
\left( \begin{array}{c}
\| \epsilon_1^0 \|_\infty \\
\| \epsilon_2^0 \|_\infty \\
\vdots \\
\| \epsilon_K^0 \|_\infty 
\end{array} \right) \leq \left( \begin{array}{cccc}
N - \frac{T(N-1)}{2} (r + q_1) & \frac{T(N-1)}{2} \lambda_1 & \cdots & \frac{T(N-1)}{2} \lambda_{1K} \\
\frac{T(N-1)}{2} \lambda_{21} & N - \frac{T(N-1)}{2} (r + q_2) & \cdots & \frac{T(N-1)}{2} \lambda_{2K} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{T(N-1)}{2} \lambda_{K1} + & \frac{T(N-1)}{2} \lambda_{K2} & \cdots & k - \frac{T(N-1)}{2} (r + q_K)
\end{array} \right) \left( \begin{array}{c}
|\mathcal{O}((\Delta t)^2)| \\
|\mathcal{O}((\Delta t)^2)| \\
|\mathcal{O}((\Delta t)^2)| \\
|\mathcal{O}((\Delta t)^2)|
\end{array} \right)
\]

\[
= \left( \begin{array}{c}
(1 - \frac{T}{2}) N \\
(1 - \frac{T}{2}) N \\
\vdots \\
(1 - \frac{T}{2}) N
\end{array} \right) \left( \begin{array}{c}
|\mathcal{O}((\Delta t)^2)| \\
|\mathcal{O}((\Delta t)^2)| \\
\vdots \\
|\mathcal{O}((\Delta t)^2)|
\end{array} \right) + \left( \begin{array}{c}
\frac{T}{2} \\
\frac{T}{2} \\
\vdots \\
\frac{T}{2}
\end{array} \right) \left( \begin{array}{c}
|\mathcal{O}((\Delta t)^2)| \\
|\mathcal{O}((\Delta t)^2)| \\
\vdots \\
|\mathcal{O}((\Delta t)^2)|
\end{array} \right)
\]

\[
= \left( \begin{array}{c}
1 - \frac{rT^2}{2} \\
1 - \frac{rT^2}{2} \\
\vdots \\
1 - \frac{rT^2}{2}
\end{array} \right) \left( \begin{array}{c}
|\mathcal{O}(\Delta t)| \\
|\mathcal{O}(\Delta t)| \\
\vdots \\
|\mathcal{O}(\Delta t)|
\end{array} \right) .
\]

Hence we can conclude that

\[ \| \epsilon_j^0 \|_\infty = |\mathcal{O}(\Delta t)|, \]

is true for all \( j = 1, 2, 3, \ldots, K \). Therefore the proof of the theorem is complete.