Ruprecht-Karls-Universität Heidelberg

Fakultät für Mathematik und Informatik

Master Thesis

Analysis and Computation of Black-Scholes Equation with Local Volatility

Sassan Mokhtar

Matriculation Number: 3357130

14.01.2019

Supervised by
Prof. Dr. Guido Kanschat

Acknowledgements

I would first like to thank my thesis advisor Prof. Dr. Guido Kanschat at Heidelberg University. The door to Prof. Kanschat office was always open whenever I ran into a trouble spot or had a question about my research or writing. While due to some personal problems a gap occurred in my study, his kind advice and supportive attitude enable me to overcome difficulties and pursue my research. Thank you Prof. Kanschat for everything.

I would also like to thank Mohammad Motamed, who has been acting as a consultant and advisor throughout my life. He generously shared his ideas and feedback on my progress and made a lot of contribution to developing this research.

I would also express my warmest thanks to Prof. Christoph Schnörr for the patient, encouragement, and advice that he provided throughout my time as his student. I applied what I learned throughout his lecture in this thesis. I also thank Prof. Stefan Riezler that with his lively, enthusiastic and energetic teaching style introduce me with interesting real-world applications of my theoretical knowledge.

Last but not least, I must express my very profound gratitude to my family and to my friends, especially my brother Sina, for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Sassan

Hereby I declare that I wrote this thesis myself with the help of no more than the mentioned literature and auxiliary means.

Heidelberg, 14.01.2019

. . . .

(Signature [your name])

To Mohammad Motamed

Abstract

This thesis concerns the analysis and computation of the Black-Scholes equation with local volatility. Since its derivation in 1973, the Black-Scholes equation has played a key role in the field of mathematical finance. While it is observed that the Black-Scholes equation has some limitations for capturing the real world features of option prices, it still acts as a benchmark in option pricing, and significant efforts have been made to analyze and improve this equation.

In addition to discussing the fundamental concepts in option pricing and deriving the Black-Scholes equation, we follow the book "Computational methods for option pricing" by Achdou and Pironneau to numerically approximate the Black-Scholes equation with local volatility by applying a Galerkin finite element method. This method consists of applying the Crank-Nicolson scheme to discretize the problem in time and obtain the semi-discrete problem. Then, the fully discrete problem is obtained by discretizing the problem with respect to the price variable and applying the Lagrange finite element method. In the end, the obtained linear system is solved by the conjugate gradient method.

Next, a posteriori error analysis is derived. The method consists of developing two families of error estimators, where the first family is global in price variable and local in time variable, and the second family is local with respect to both price and time variables. Then, these error estimators are applied to propose an algorithm to refine the mesh adaptively.

Finally, after a brief discussion of the shortcomings of Black-Scholes equation, we follow the paper "Computation of local volatilities from regularized Dupire equation" by Hanke and Rösler to calibrate the local volatility function from observed option prices in the market. In this method, which can be considered as a remedy for limitations of Black-Scholes equation, the observed data are smoothened by cubic splines. Then, the Dupire equation is applied to evaluate the local volatility function. However, the obtained system is underdetermined and ill-posed, and thus the first order Tikhonov regularization is used to regularize the problem.

We examine this method with the real market data of Russell 2000 index on 1st of August 2018, where the resulting volatility surface appears to be qualitatively correct.

Contents

List of Figures

1 Introduction

In 1973, the first modern option exchange, called "CBOE", was established in Chicago. Since then, options have been widely used by traders on markets and exchanges. While dealing with the problem of option pricing dates back to the work of Bachelier in 1900, the first successful attempt to derive analytical models for option valuation was the celebrated Black-Scholes equation in 1973. Black-Scholes equation is a second order parabolic partial differential equation which is considered as a benchmark in mathematical finance for pricing options. Although in the original Black-Scholes model the interest rate and volatility are assumed to be constant, in this thesis we consider the extended version of the Black-Scholes equation with the interest rate being a function of time and the volatility being a function of time and price of the underlying asset. This model is typically called "Black-Scholes equation with local volatility".

In spite of the fact that there is an analytic solution for the original Black-Scholes equation, analytical approaches are limited to constant volatilities and may not be applied to more complex options and real market data. Therefore, the numerical approximation of Black-Scholes equation needs to be considered. Among different classes of numerical methods that are applied to the Black-Scholes equation, finite difference and finite element methods are shown to have better performance. Because, beside the fact that these methods are fast and accurate, they can be used for pricing different types of options. However, one of the advantages of finite element methods are that they are supported by strong mathematical theory. Moreover, finite element methods allow a posteriori error analysis and performing local mesh refinement. Hence, we follow Achdou and Pironneau [AP05] and develop a finite element method to approximate the Black-Scholes equation with local volatility. This method consists of discretizing the problem in two steps. First, we discretize the problem in time by applying the Crank-Nicolson scheme to obtain the semi-discrete problem. Next, we apply Lagrange finite elements to discretize the problem with respect to the spatial variable and obtain the fully discrete problem. Finally, the Conjugate gradient method is used to solve the obtained linear system.

In the case of a posteriori error analysis, we consider the method proposed by Bergam, Bernardi, and Mghazli [BBM05]. The method consists of developing two families of residual-type error estimators, where the first family is global with respect to spatial variable and local in time, and the second family is local in both space and time. Then, we demonstrate the algorithm that uses these error indicators to perform adaptive mesh refinement for the Black-Scholes equation with local volatility.

On the other hand, based on empirical studies, there is a difference between the real value of the option in the market and the value that the Black-Scholes model suggests. In particular, after the US market crash in October 1987, volatility smile has appeared in the market. Volatility smile is occurred when we plot the extracted implied volatility for the same underlying asset price and the same maturity but with different strikes. According to the original Black-Scholes equation, the obtained line should be straight. Nevertheless, in practice, this process leads to a curve that is higher at the ends and lower in the center.

There are two important approaches to tackle the shortcomings of the Black-Scholes equation in capturing the real world features of option prices. The first one consists of extending the model by relaxing some restrictive assumptions on the original Black-Scholes equation (in fact this is why we consider the Black-Scholes equation with local volatility in this thesis). The second approach deals with proposing new models. In addition, while other parameters of Black-Scholes equation can be determined by information in the market easily, evaluating the volatility is a complicated task. Thus, a significant amount of work has been done by different authors to assign an appropriate value to the volatility. According to [SS91], we can categorize these efforts into two groups, deterministic volatility models and stochastic volatility models. Deterministic volatility models assume that the volatility can be determined using variables that are observed in the market. While in stochastic volatility models, it is assumed that the volatility itself is stochastic. The advantage of deterministic volatility models lies in the fact that these models can take into account some empirical regularities, for instance, time-varying volatility. Also, deterministic volatility models give us the opportunity to calibrate the volatility surface and fit the smile. However, in terms of practical use of deterministic models, better performance in pricing could be obtained at the expenses of overfitting, which limited the practical application of these models for pricing exotic options [BJ01]. On the other hand, stochastic volatility models lead to producing new parameters which make the task of evaluating the volatility more complex.

In this thesis, we follow the method which is proposed by Hanke and Rösler [HR05], to compensate for the limitation of the Black-Scholes equation. This method belongs to the class of deterministic volatility models and consists of calibrating the local volatility function from observed option prices in the market. Typically, in option markets, the information for few maturity dates and a discrete sample of strikes on each maturity is given. In this method, the observed data are smoothened by cubic splines. In the next step, the numerical differentiation should be done, and for doing this, the Dupire equation is applied. The Dupire equation determines the option value in specific price and time as a function of strike and maturity. Therefore, the advantage of applying the Dupire equation is that, unlike the Black-Scholes equation, by solving the equation just once, we obtain the option value for each pair of strikes and maturities. However, due to the nature of this problem the obtained system is underdetermined and ill-posed. Thus, we apply the first order Tikhonov regularization, to regularize the system. We apply this approach to the real data of Russell 2000 index, which can be considered as the small capitalization stock market index, on the 1st of August 2018. The obtained volatility surface appears to be qualitatively correct.

The remainder of this thesis is organized as follows: we start by reviewing the related literature in chapter 2. Chapter 3 consists of two parts. We first demonstrate the basic financial concepts related to options, and in the second part of the chapter, we derive the Black-Scholes equation. The mathematical analysis of the Black-Scholes equation with local volatility is considered in chapter 4, in which we derive the variational formulation of this equation. In the first part of chapter 5, we obtain the analytic solution to the original Black-Scholes equation. While in the second part, we apply the finite element method to approximate the Black-Scholes equation with local volatility numerically. In the end of this chapter, the obtained linear system is solved by applying the conjugate gradient method. In chapter 6, we develop a posteriori error analysis to the Black-Scholes equation with local volatility and perform a strategy to refine the mesh adaptively. Finally, in chapter 7, we first, discuss the shortcomings of the Black-Scholes equation. Then, calibration of local volatility function is done, and the method is tested for the real world data.

2 Literature Review

In this chapter, we review some literature on the numerical analysis of Black-Scholes equation. In the first section, we regard four different classes of numerical schemes for option valuation. Then, based on the fact that we are going to apply the finite element method for pricing options, we consider literature about adaptive mesh refinement of the Black-Scholes equation. Finally, we illustrate the alternative methods for calibration of local volatility surface in the last section of this chapter.

2.1 Numerical Schemes in Option Pricing

There is an explicit formula for the Black-Scholes equation of European options which is going to be discussed in chapter 4, however, it is limited to constant interest rate and volatility, and can not be applied to general market data and more complex options. Therefore, numerical schemes are widely used in option pricing. Four different classes of these numerical schemes are discussed in this section.

2.1.1 Monte Carlo Simulation

Monte Carlo method is, in fact, a numerical scheme for approximating the expected value of a random variable. On the other hand, the problem of option pricing can be stated as expectations. Therefore, it is possible to apply Monte Carlo methods for option pricing [Kwo08]. In 1977, Boyle was the first one who suggested a Monte Carlo simulation method for solving the problem of option valuation [Boy77]. After that, this method, which can be considered as an alternative to the Black-Scholes equation, has been applied widely for derivative pricing problems. Take for example Glasserman [Gla13], who discusses in detail the application of Monte Carlo methods for pricing different types of options. Moreover, the application of the Monte Carlo scheme in pricing European-style options is regarded by Seydel [Sey06], and Higham [DH96]. Meanwhile, according to [WDH93], the drawback of Monte Carlo schemes is that they are relatively slow and inflexible, compare to other numerical methods in option pricing.

2.1.2 Lattice Tree Methods

Another alternative numerical model for the problem of option valuation is Lattice tree model. As it is illustrated by Wilmott, Dewynne, and Howison [WDH93], the main idea behind the Lattice methods is to build a lattice of the possible values of the asset at maturity and probability of occurring these values, based on the current value of the underlying asset. One of the simplest Lattice methods is the binomial model, which assumes a two jump process for the asset value over each time step. In 1979, Cox, Ross and Rubinstein [CRR+79], applied the binomial method for solving the option valuation problem. Since then, various versions of this model with different assumptions have developed by different authors. For instance, pricing of European-style options by the binomial model is considered in [WHHD95] and [Cor01].

Furthermore, trinomial schemes are the extension of binomial models, in which one assumes that there is a three jump process for the price of the asset at each time step. Trinomial model for option pricing was developed by Boyle in 1986 [Boy86]. Since then, similar to the binomial scheme, many variants of this scheme has been proposed by different authors. Take for example the method which was developed by Kamrad and Ritchken in 1991 [KR91]. As it is mentioned by Kwok [Kwo08], when the number of time steps is high, both trinomial and binomial schemes satisfy the Black-Scholes equation to the first order accuracy. Nevertheless, the disadvantage of these models is their low speed.

2.1.3 Finite Difference Methods

The finite difference method has been numerously used by different authors to derive the numerical approximation of partial differential equations. Similarly, applying different schemes of finite difference method to the Black-Scholes equation has become very popular. Take for example Achdou and Pironneau [AP05], derive the fully implicit Euler scheme to the Black-Scholes equation, and obtain that this scheme is first-order accurate in time and second-order accurate in price variable. They also test the second-order accurate Crank-Nicolson scheme and observe that in some cases the order of accuracy of this scheme is less than 2.

More precise analysis of this phenomenon is done by Duffy [Duf18], which analyze the fully implicit scheme and observe that while the method is stable and do not produce false oscillations, it is only first order accurate. Moreover, by deriving the Crank-Nicolson scheme to the Black-Scholes equation he obtains that this scheme is stable and second order accurate. However, he observes that when the convective term is large or in the regions with low regularity in the payoff function (near the strike price K for European options), false oscillations are produced. Then, the Keller box scheme, which is based on the work by Keller [Kel71], is introduced. This scheme consists of reducing the second order equation to a system of first order equations, then approximate the first derivatives by averaging in a box. This scheme cures the shortcomings of Crank-Nicolson, and is second-order accurate, unconditionally stable and produce no oscillation. Furthermore, regarding the work in [Duf80], the exponentially fitted scheme is suggested by Duffy, which consists of exponentially fitted scheme in the price direction and fully implicit discretization in the time direction. This scheme is proved to be uniformly stable and oscillation-free. Also, according to [Duf13], the exponentially fitted scheme has a first-order accuracy in time and price variables, however, by applying Richardson extrapolation one can produce a second order scheme.

2.1.4 Finite Element Methods

Based on the fact that many problems in finance are one-dimensional in space, finite element seems to be unnecessarily complex. However, due to the fact that the finite element methods support local refinement and are flexible for more complex cases, they are widely used in finance [AP07].

As it is mentioned by Topper [Top05b], option pricing problems are usually of the form,

$$u_t = L[u] - f$$

where $L[u]$ is a second order differential operator. A typical approach for discretizing this type of problem is to discretize the spatial variable with finite elements, and time variable with finite differences. Nowadays, almost all finite element methods belong to the *Method of Weighted Residual* (MWR), which (in static one variable problems) consists of introducing the residual $R = L[\tilde{u}] - f$, and minimizing it. This can be done by enforcing the weighted integrals of the residual to be zero, i.e.

$$\int_{\omega} RW_j dx = 0 \quad \forall j$$

where W_j are weights. Then, in order to take into account the boundary conditions and also to determine the weighting functions, there are two popular choices, Galerkin method and Collocation method. In the case of the Galerkin method, shape function ϕ is regarded as the weighting function. One of the advantages of this method is that it produces the matrices with nice structure. In Collocation method we put $W_i(x) = \delta(x - x_i)$, where δ is a Dirac function. So that,

$$\int_{\omega} R(x, \tilde{u})\delta(x - x_i) \stackrel{!}{=} R(x_i; \tilde{u}) \quad \forall i.$$

Both of these approaches are considered by Topper [Top05a], for the valuation of European options. First of all, the Black-Scholes equation is transformed by a change of variable, then the Galerkin finite element method is derived by applying equi-distance four element model discretization and linear interpolation. Then, implementation of the Collocation finite element for European put option is done. According to his data, he concludes that this finite element approach is of first order (by applying the implicit Euler scheme) and second order (by applying the Crank-Nicolson scheme) in time, and fourth order in space.

On the other hand, in this thesis, we are going to regard the work of Achdou and Pironneau [AP05]. Their approach consists of applying Galerkin finite element method to the Black-Scholes equation with local volatility. The discretization is done in two steps. First, by applying the Crank-Nicolson scheme in time, the time-semidiscrete problem is derived. Then, in order to obtain a fully discrete problem, the Lagrange finite element is applied to discretize the domain with respect to the spatial variable.

Furthermore, discontinuous Galerkin method has been used by several authors for option pricing. Take for example Larsson [Lar13], which apply the method in the [PvP07], to semi-discretize the spatial variable by a second order not equidistant finite difference scheme. This adaptive finite difference scheme leads to the system of ordinary partial differential equations. Then, the discontinuous Galerkin method of degree r is applied to this system of equations. However, it is costly to solve the obtained linear system of the size $(r + 1)N$, where N is the number of grid points. Thus, the method in [SS00] is applied, to decouple the system, into $(r + 1)$ linear systems of size N, by choosing normalized Legendre polynomials as temporal shape functions. In the end, it is concluded that this approach can obtain the convergence rate of $2r + 1$.

2.2 A Posteriori Error Analysis of Black-Scholes Equation

For more than thirty years, a significant amount of work has done regarding a posteriori error analysis for the finite element discretization of partial differential equations [BBM05]. Due to the fact that the Black-Scholes equation is a parabolic partial differential equation, there are numerous works that take into account a posteriori error analysis of this equation.

For instance, the paper by Goll, Rannacher and Wollner [GRW15], in which the *Dual Weighted Residual* method is applied to derive a posteriori error estimator. In the case of time semi-discretization, they develop a damped Crank-Nicolson scheme. Authors argue that since irregular initial data appears naturally in the Black-Scholes model, such damping is needed to eliminate oscillations in the solution. The discretization in price variable is done by conforming finite element. Then, a posteriori error analysis is proposed by applying the *Dual Weighted Residual* method, which deals with solving the dual problem. According to the authors, possible irregularity of initial data makes damping necessary for dual equation too. Therefore, they develop a consistent discretization, which is the discretization that the dual to the discrete problem is also discrete of the dual problem. Then they follow the "goal-oriented" approach (discussed in [EEHJ96], [BR96] and [BR01]), in order to perform the error estimation. In the end, they propose the adaptive finite algorithm based on this method. According to their experiments, the convergence order of the damped Crank-Nicolson scheme is 2, while the undamped version of Crank-Nicolson has the convergence order of 1.

Another work which takes into account the adaptive finite element method for black-Scholes equation is the paper by Ern, Villeneuve, and Zanette [EVZ04]. In the case of numerical computation of Black-Scholes equation, they apply the non-conforming Galerkin method, which consists of discrete functions that are continuous in price variable and discontinuous in time variable. Next, a posteriori error analysis is derived, in which finite element residual and solution of dual problem control functional of the error. Then, an adaptive mesh refinement algorithm is derived by localizing a posteriori error estimator onto the cells of the mesh. Based on their practical experiments, while in theory, the convergence order of a posteriori error bound is 2, in practice the order of convergence fluctuates between 1.6 and 2.1. It is also worth mentioning that similar method is used by Foufas and Larsson in [FL08], for computation of Asian options.

Moreover, we can mention the paper by Jackson and Süli [JS97], that consider the generalized Black-Scholes equation and transform the equation by a change of variables. Then, in order to obtain a sharp a posteriori error bound, they propose piecewise Hermite cubic elements. Therefore, the discontinuous Galerkin method is used. In the next step, by adapting the dual problem, a residual-based a posteriori error bound is proved. Finally, regarding the bound, an adaptive strategy is developed to solve the obtained problem.

However, we follow the method which is demonstrated by Achdou and Pironneau [AP05], in order to derive a posteriori error analysis for the Black-Scholes equation. This method consists of two families of residual-type error indicators, where the first one is global with respect to price variable and local with respect to time and the second family is local with respect to both time and price.

2.3 Local Volatility Model

According to [GNC14], one important remedy for shortcomings of the Black-Scholes model in real life applications is to calibrate the local volatility function, which consists of finding the local volatilities such that the theoretical option prices match the option values in the market. Therefore, this approach leads to an optimization problem. Nevertheless, there are some difficulties with this problem. First of all, we would like to find the volatility in each grid points, thus this problem is large scale. Also, the Black-Scholes equation (and Dupire equation) is a nonlinear operator in σ. Moreover, there are only a few numbers of available data in real option markets, and this is much less than the number of parameters. Finally, small changes in option price bring about significant changes in the volatility surface. Hence, the calibration problem is a nonlinear, large scale, underdetermined and ill-posed inverse problem.

In order to overcome the large-scale nonlinear nature of the calibration problem, authors usually apply the gradient-based optimization methods. One important class of these methods consists of solving the adjoint partial differential equation of Dupire or Black-Scholes equation (e.g. [LO+97] and [BI97]). Another option is to apply the automatic differentiation. In the case of the underdetermined and ill-posed nature of the calibration problem, authors usually regularize the problem by the first order Tikhonov regularization (e.g. [BJ99]). It also should be noted that the Dupire equation determine the option value as a function of strike (K) and maturity (T), in specific price and time. Therefore, it is computationally much cheaper to consider the Dupire equation, instead of the Black-Scholes equation, in calibration problems. Because by doing this, we do not need to solve a partial differential equation for each pair of strike and maturity separately.

Among authors that take into account the calibration problem, we can mention Achdou and Pironneau [AP05], that use the Dupire equation, and regularized problem with first order Tikhonov regularization. Then, by applying the adjoint state operator, solve the problem and derive the local volatility surface. They apply their method to the S&P 500 index and plot the obtained volatility surface.

Moreover, Coleman, Li, and Verma [CLV01] apply a spline functional approach, which consists of producing the local volatility function by a spline whose values are obtained by solving an optimization problem. They do not use the Dupire equation, therefore the computational cost increase. However, the authors argue that this approach leads to a more stable method. In the end, their method is tested with real option data of the S&P 500 index. Similarly, the paper by Jackson, Süli, and Howison [JSH98] propose another method for calibrating the volatility surface without using the Dupire equation. They consider the instantaneous volatility as a spline, where the weights are obtained by solving the regularized optimization problem numerically (i.e. solve the partial differential equation with finite element methods). Finally, they construct the volatility for the straddle option prices on the FTSE-100.

Furthermore, Turinici [Tur09], develops a method that use the first order Tikhonov regularization to regularize the calibration problem. Then, he proposes an approach which can be regarded as a special case of the sequential quadratic programming and consists of Gauss-Newton style approximation of the Hessian. Also, he derives the volatility surface based on the market data of call option on S&P 500 and FOREX. We can also mention the paper by Geng, Navon, and Chen [GNC14], which apply the automatic differentiation to solve the inverse problem. However, the novelty of their approach is to use the second order Tikhonov regularization, which leads to smoother volatility surface. They verify their approach by testing both synthetic and real market data.

However, we follow the more straightforward method which is presented by Hanke and Rösler [HR05] and consists of regularizing the problem by applying the first order Tikhonov regularization. Then, data are smoothened by deriving cubic splines, and according to Dupire equation, the numerical differentiation is done.

3 Option Pricing

This chapter is an introduction to the Black-Scholes equation. In the first section of the chapter, basic financial concepts related to options are discussed. The second section deals with deriving Black-Scholes equation.

3.1 Option Basics

As illustrated in [Jia05], financial derivatives are instruments whose value depends on the price of an underlying asset, where the asset can be Bonds, Stocks, and commodities etc. Options are one of the most important financial derivatives, which frequently and in significant portion trade in financial markets. Before deriving Black-Scholes equation, it is helpful to discuss in detail basic concepts of options.

3.1.1 Brief History of Options

Whereas option markets were quite unidentified before 1973, by establishing the first modern options exchange "Chicago Bond Options Exchange (CBOE)", the number of options as well as their trading volume, increased fundamentally each year .

3.1.2 What Is an Option?

A financial option is a contract in which the writer (i.e. the seller) gives the right, but not the obligation to the holder (i.e. the purchaser), to buy or sell a certain amount of underlying asset, at the specified price (called "**Strike**") and on the specified date (called "**Maturity**"). In other words, if the option holder decides to exercise the option, the option writer has the obligation to execute the transaction. A **call option** gives the holder the right to buy the underlying asset while **put option** provides option holder the right to sell the underlying asset.

Moreover, **European options** and **American options**, are two basic types of options. While the latter can exercise by the holder at any time until maturity, the former can only exercise at maturity. During the thesis, we consider the analysis of European options in depth.

3.1.3 Moneyness

Another crucial concept in option pricing is moneyness, which indicates the relationship between the market price of the underlying asset and strike. "In the money" options are those that underlying asset price is higher than strike, whereas, in the "out of money" options, the strike is greater. If these values are the same, then the option is called "at the money". Meanwhile, it is worth noting that given definitions are for Call options, the description is reverse for Put options. [Mar01]

3.1.4 The Value of an Option

The fact of the matter is that, if the option is "in the money", then the option holder can exercise the option and obtain some profit. However, for "out of the money" situation, the option holder can simply use his or her right and refuse to exercise money and lose nothing. Thus, option trading is "no-lose situation" for option holder and "no-win situation" for option writer. The question arises here is that why someone should write an option under this circumstance. The answer is that the option writer receives some money in return for writing the option, and this amount is called **Option Value** (or **Option Premium**).

As it is mentioned in [Chr96], there is a moment in which determining the value of an option is relatively easy, and it is at expiration (maturity). Based on the fact that if the price of the underlying asset (denoted by S) is greater than the strike (denoted by K) for a call option, then the option holder can obtain $S - K$ by exercising it. On the other hand, in the case of "out of money", a rational option holder will not use his or her right to exercise the option, and the option will expire worthlessly. Consequently, a fair value of a call option in maturity is given by $max(0, S - K)$. By the same reasoning, put option value in maturity is $max(0, K - S)$. Figure 3.1 shows a graph of payoff at the expiry of European call (the left-hand side) and put (the right-hand side) options, where strike is equal to \$50.

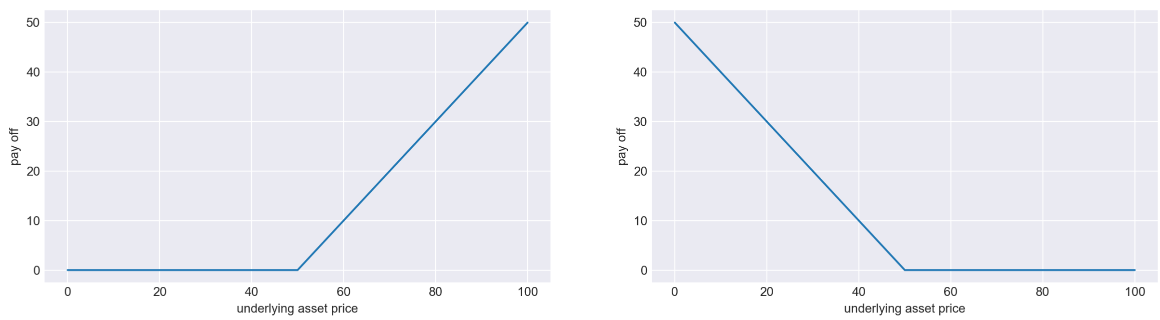

Figure 3.1: Payoffs in European call (the left-hand side) and European put (the right-hand side) options

Furthermore, the payoff is different from profit. If one would like to calculate the profit of the option holder, the premium should subtract from payoff function. Profit diagram of European call and put option is shown in figure 3.2, where strike is \$50 and the option value is equal to \$10.

It is also important to consider the fact that in order to value an option, there are some factors which should take into account. There are 6 factors that affect the value of options. These factors are consists of the strike K, maturity T, the current market price of the underlying asset S (also known as spot price), interest rate r, future volatility of the price of the underlying asset σ and dividends q for underlying asset. To put it in another way, option value is a function of these 6 factors. [Mar01]

One of the main objective in option pricing is to determine an appropriate value of options. One way to tackle this problem is deriving an equation that gives the analytic solution. The first and maybe the most important effort to obtain such analytic solution is the Black-Scholes model, which was proposed by Fischer Black and Myron Scholes in 1973. Introducing this equation is the subject of the next section.

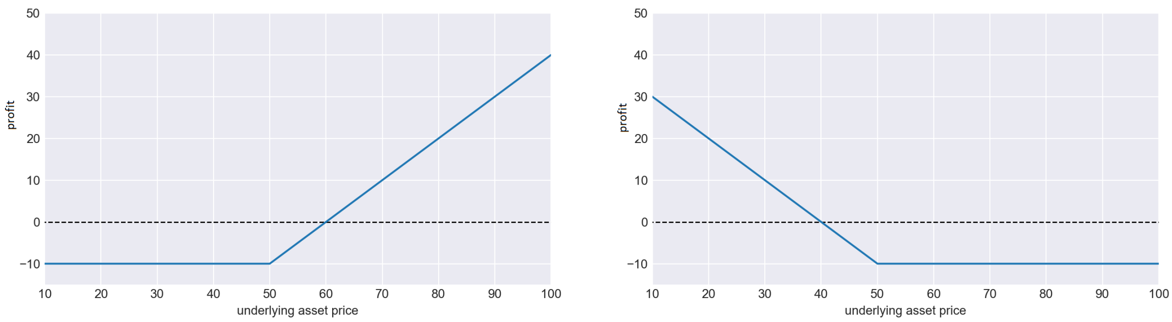

Figure 3.2: Profit from buying a European call (the left-hand side) and European put (the right-hand side) options

3.2 Black-Scholes Equation

In this section, we follow the procedure in [HB16], in order to derive the Black-Scholes equation and obtain a fair value for European options.

3.2.1 Stochastic Model

Deriving the stochastic model for capturing the price of a stock is the goal of this subsection. For doing this, illustrating some definitions and concepts are necessary.

Definition 3.2.1. Stochastic process: A *stochastic process* is a family of random variables $\{X(t)\ ,\ t \in T\}$ with parameter space T and state space, which illustrates the probabilistic evolution of the value of a random variable chronologically. [Bei06]

If T is a finite or countably infinite set, then the stochastic process is in *discrete time*, otherwise, it is categorized as *continuous time*. Similarly, stochastic processes can be classified as a *continuous variable* and *discrete variable*, where in the latter the state space is finite or countably infinite (**i.e.** the underlying variable can get finite values in a certain range) and in the former the state space in infinite.[Bei06]
It is also worth noting that, while stock prices are processes in discrete time, they will be modeled as a continuous time continuous variable stochastic processes. Because it is a proper model for deriving Black-Scholes equation. Also, continuous time processes are easier to handle analytically. [FHH04]

Definition 3.2.2. Markov process: A *Markov process* is a particular kind of stochastic process in which satisfies two conditions. First, only the present value of a variable is relevant to predict the future. Secondly, the past value of the variable and the way the current value is evolved from the past is irrelevant.

Before deriving the next definition, it is necessary to describe two concepts.
A *drift rate* of a stochastic process is the change of mean in unit time and *variance rate* is a variance per unit time.

Definition 3.2.3. Wiener process: A *Wiener process* or *Brownian process dz* is a particular type of Markov process which describes the evolution of normally distributed variable with a drift rate of zero and variance rate equal to 1.

To put it in another way, if the value of the random variable at $t = 0$ is x_0, then at the time $\bar{t}$ it is normally distributed with mean x_0 and standard deviation $\sqrt{\bar{t}}$.

Definition 3.2.4. generalized Wiener process: A *generalized Wiener process* illustrates the evolution of a variable with the normal distribution, where drift rate and variance rate per unit are a and b^2 respectively (a and b are constant). Consequently, for a variable x, generalized Wiener process can be written algebraically as

$$dx = a\,dt + b\,dz. \tag{3.1}$$

Intuitively, the adt term express the drift rate of the variable x, whereas, the term bdz is considered as the noise of the path followed by x.

Definition 3.2.5. Itô process: If a and b in the definition of generalized Wiener process be functions of the value of the underlying variable x, then the process is called Itô process. Thus, algebraically

$$dx = a(x,t)\,dt + b(x,t)\,dz. \tag{3.2}$$

Now, based on above definitions and concepts, it is possible to derive a stochastic model for stock prices.

One option is to assume that stock price follows the generalized Wiener process. However, based on the fact that expected percentage return is independent of the stock's price (**i.e.** if investor looking for 14 % annual expected return, it is not important whether the stock's price is \$10 or \$50), the constant expected drift rate is an inappropriate assumption. Instead, one should assume that the expected return is constant.

In other words, suppose that S is a stock price at time t, then one may consider the expected drift rate in S as $\mu\,S$, where a constant parameter μ is the expected rate of return on the stock. This implies that, regarding the short period of time (Δt), then the expected growth in S is $\mu\,S\,\Delta t$.

Moreover, due to the fact that investor's uncertainty of the percentage return is independent of the stock price (**i.e.** whether the price is \$ 10 or \$ 50 the risk of percentage return for the investor is the same), it makes sense to assume that in the short interval of time, Δt, the stock price should be proportional to the standard deviation of the change. Which results in the following wildly used stochastic model for describing the stock price in the real world

$$dS \;=\; \mu\,S\,dt \;+\; \sigma\,S\,dz \tag{3.3}$$

or

$$dS/S \;=\; \mu\,dt \;+\; \sigma\,dz. \tag{3.4}$$

This model is called *geometric Brownian motion*, in which μ is expected rate of return of the stock and σ is the stock price volatility. It is also worth noting that in the risk neutral world (which illustrates the situation where the investors are risk-neutral), μ is equal to the risk-free interest rate r.

3.2.2 Itô's Lemma

In order to derive a celebrated Black-Scholes partial differential equation, one needs another ingredient, called Itô's lemma [Itō51]. In fact, the role of this lemma for functions of the random variable is similar to the role of Taylor expansion for functions of deterministic variables. It relates a slight change in function of random variable x to the slight change in the random variable x itself [WDH93]. In other words, it provides us the way to take into account the behavior of functions of stochastic variables.

Regarding the previous definitions, if dz is a Wiener process, also a and b are functions of x and t. In addition, suppose that the drift rate and variance rate of variable x is a and b^2 respectively, and x follows the Itô process

$$dx = a(x,t)\, dt + b(x,t)\, dz. \tag{3.5}$$

Then Itô's lemma proves that the function G follows the process

$$dG = \Big(\frac{\partial G}{\partial x}\, a + \frac{\partial G}{\partial t} + \frac{1}{2}\,\frac{\partial^2 G}{\partial x^2}\, b^2\Big) dt + \frac{\partial G}{\partial x} b\, dz \tag{3.6}$$

where G is the function of x and t. and Wiener process dz is the same as the process in 3.5. Therefore, G follows the Itô's process either, where the drift rate is

$$\frac{\partial G}{\partial x}\, a + \frac{\partial G}{\partial t} + \frac{1}{2}\,\frac{\partial^2 G}{\partial x^2}\, b^2$$

and the variance rate is

$$(\frac{\partial G}{\partial x})^2 b^2$$

By combining the result of Itô's lemma and the obtained stochastic model of previous subsection, one can deduce that the process which follows by the function G of S and t is

$$dG = \Big(\frac{\partial G}{\partial x}\, \mu\, S + \frac{\partial G}{\partial t} + \frac{1}{2}\,\frac{\partial^2 G}{\partial x^2}\, \sigma^2\, S^2\Big) dt + \frac{\partial G}{\partial x} \sigma\, S\, dz. \tag{3.7}$$

By comparing 3.3 and 3.7, we realize the underlying source of uncertainty dz for both G and S.

3.2.3 Black-Scholes Model

Before deriving the Black-Scholes model, it is beneficial to describe very briefly the history behind the asset pricing.

History

As it is mentioned in [Din13], Louis Bachelier's is known as the founder of mathematical finance. In his Ph.D. thesis "The Theory of speculation"[Bac00], he used random walks, in order to show that stock prices follow a Brownian motion. Moreover, Bashelier introduced a method for pricing of the certain short-term options. In the mid-60s, Paul Samuelson[Sam65] observed that so as to obtain a more accurate illustration of the movement of shared prices, it is better to use geometric Brownian motion instead of simple Brownian motion. Samuelson was awarded the Nobel prize in Economic Sciences in 1970 for his contributions in economic theory.

The next major effort was done by other Noble prize winners Harry Markowitz and William Sharpe in portfolio theory. They replace the problem identifying the best stock for investment, by the problem of calculating the risk of the whole portfolio of stocks and bonds.

Although asset pricing methods developed significantly since Bachelier's thesis in 1900, obtaining an explicit formula for option pricing was missing. This revolutionary task achieved by the work of two financial economists Fischer Black and Myron Scholes, with important contributions of Robert C. Merton in 1973. Fischer Black obtained Ph.D. in applied mathematics from Harvard University. Besides that, he was working for the management consulting company "Arthur D. Little" as a financial consultant. Myron Scholes is an American-Canadian financial economist who defended his doctoral thesis at the Chicago University on the role of arbitrage in the securities market. During his first year in MIT as an assistant professor, he met Fischer Black there. On the other hand, Robert C. Merton enrolled at MIT for a graduate study in economics. Due to the fact that he was interested in dynamic portfolio selection at that time, he started to study Itô calculus, as an appropriate mathematics to model hedging strategy. While Merton was interviewed for a position at MIT, he met Scholes. Then they three start collaborating on their mutual interest.

On 1973, Black and Scholes published their famous paper [BS73], proposed two derivations of Black-Scholes formula. The first approach is based on the Capital asset pricing model, and the second one is based on Merton's idea in hedging. Merton also published an important paper [Mer73] in 1973. Scholes and Merton awarded a Noble prize in Economic Sciences in 1997, while Black did not receive an honor due to early death in 1995.

In this section, we derive the Black-Scholes equation based on the Merton's hedging approach.

Deriving the Black-Scholes differential equation

In order to begin derivation of the Black-Scholes equation, one needs to define an important concept in option pricing.

Definition 3.2.6. Delta-hedging: For a given option V, trading Δ shares of the underlying asset S in the opposite direction, such that the portfolio

$$\Pi = -V + \Delta S$$

is risk-free, is called delta-hedging.[Jia05]

Assume that the payoff of an option at maturity is denoted by $V(S,T)$, then so as to find the value of the option in different times before maturity, use earlier mentioned Itô's lemma for two variables S and t. Thus,

$$dV = \Big(\mu\, S\, \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\,\sigma^2\, S^2\, \frac{\partial^2 V}{\partial S^2}\Big) dt + \sigma\, S\, \frac{\partial V}{\partial S}\, dz. \tag{3.8}$$

Regarding the delta-hedge portfolio with $\Delta = \frac{\partial V}{\partial S}$, then the value of this portfolio is

$$\Pi = -V + \frac{\partial V}{\partial S} S$$

By considering the total profit or loss from the change in the value of the portfolio over the time period $[t, t+\Delta t]$, one obtains

$$\Delta\Pi = -\Delta V + \frac{\partial V}{\partial S}\Delta S. \tag{3.9}$$

Now, replace differential with Δ in 3.3 (in fact we consider the discretization of the stochastic model)

$$\Delta S \;=\; \mu\, S\, \Delta t \;+\; \sigma\, S\, \Delta z. \tag{3.10}$$

Following the same procedure for 3.8, leads to

$$\Delta V = \Big(\mu\, S\, \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\,\sigma^2\, S^2\, \frac{\partial^2 V}{\partial S^2}\Big)\Delta t + \sigma\, S\, \frac{\partial V}{\partial S}\, \Delta z. \tag{3.11}$$

In order to eliminate the uncertain part (Δz), plug 3.10 and 3.11 into 3.9, to obtain

$$\Delta \Pi = -\Big(\mu\, S\, \frac{\partial V}{\partial S} + \frac{\partial V}{\partial t} + \frac{1}{2}\,\sigma^2\, S^2\, \frac{\partial^2 V}{\partial S^2}\Big)\Delta t - \sigma\, S\, \frac{\partial V}{\partial S}\, \Delta z + \frac{\partial V}{\partial S}\Big(\mu\, S\, \Delta t \;+\; \sigma\, S\, \Delta z\Big)$$

And by performing introductory mathematical operations

$$\Delta \Pi = -\Big(\frac{\partial V}{\partial t} + \frac{1}{2}\,\sigma^2\, S^2\, \frac{\partial^2 V}{\partial S^2}\Big)\Delta t. \tag{3.12}$$

To continue the derivation of Black-Scholes differential equation, it is necessary to define one of the most important concepts in the theory of option pricing.

Definition 3.2.7. Arbitrage opportunity: An *Arbitrage opportunity* is an investment opportunity in which earning risk-free money is guaranteed.[Ste11]

In order to eliminate the arbitrage opportunity, the rate of return of the portfolio should be equal to the risk-free rate of return (which is interest rate r) over the time period $[t, t + \Delta t]$. Therefore,

$$\Delta \Pi = r \Pi \Delta t. \tag{3.13}$$

Plug obtained values of Π and $\Delta\Pi$ into 3.13, results in

$$\Big(-\frac{\partial V}{\partial t} - \frac{1}{2}\,\sigma^2\, S^2\, \frac{\partial^2 V}{\partial S^2}\Big)\Delta t = r\Big(-V + \frac{\partial V}{\partial S} S\Big)\Delta t. \tag{3.14}$$

Introductory mathematical simplifications lead to the famous Black-Scholes partial differential equation

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

The next step is to determine the initial and boundary conditions for this second order parabolic partial differential equation. For doing this, one needs to differentiate between call options and put options.

Based on the discussion in [WDH93] for call options, as it is mentioned in the previous section, the final value (at maturity) of an option is given by $C(S, T) = max(0, S - K)$. If $S = 0$, then the payoff will be zero too, hence the call option is worthless and $C(0, t) = 0$. When $S \to \infty$, the magnitude of the exercise price become less important and the value of the option become equal to the value of the asset

$$C(S, t) \sim S \quad as \quad S \to \infty$$

In the case of put options, the value of an option at maturity is $P(S, T) = max(0, K - S)$. Whenever S is zero, the value of the option is equal to the strike K, however, one should calculate the present value of strike. Thus, under the assumption that the interest rate r is constant, one obtains

$$P(0,t) = Ke^{-r(T-t)}$$

Nevertheless, if $S \to \infty$, then the put option is worthless, and

$$P(S,t) = 0 \quad as \quad S \to \infty$$

In this chapter, we discussed the derivation of the Black-Scholes equation based on two fundamental assumption. First, the underlying asset price follows the geometrical Brownian motion and the second assumption is that the market is arbitrage-free[Jia05]. The exact definition of Black-Scholes equation with local volatility and its mathematical analysis is given in the next chapter.

4 Mathematical Analysis of Black-Scholes Equation with Local Volatility

In this chapter, we follow the discussion in [AP05], in order to provide the required mathematical tools for analysis of the Black-Scholes equation with local volatility. The fact of the matter is that the results of this chapter are a basis for deriving the numerical solution of the Black-Scholes equation with local volatility in the next chapter. Meanwhile, we try to present the most fundamental aspects of mathematical analysis of the Black-Scholes equation with local volatility and do not include detailed mathematical discussion in this chapter.

The first part of this chapter deals with presenting the strong solution to the Black-Scholes equation with local volatility. Then, we will derive the weak formulation of the equation, which as will be discussed later, plays an important role in finite element methods. Finally, we are going to express the theorem which proves that there is a unique solution for this weak formulation.

4.1 Strong Solution

In order to derive a mathematical analysis of the Black-Scholes equation with local volatility, it is more straightforward to consider the European put option. Moreover, based on the fact that we are interested in the forward parabolic equations, the time variable t is replaced by the time to maturity $T - t$. As a consequence, in this section, we regard the forward parabolic differential equation for $S > 0$ and $t \in (0, T]$,

$$\frac{\partial}{\partial t}P(S,t) - \frac{1}{2}\sigma^2(S,t)S^2\frac{\partial^2}{\partial S^2}P(S,t) - r(t)S\frac{\partial}{\partial S}P(S,t) + r(t)P(S,t) = 0 \tag{4.1}$$

with Cauchy data

$$P(S,0) = P_0(S), \quad S \in \mathbb{R}^+ \tag{4.2}$$

where P_0 is the payoff function defined by $P_0(S) = max(0, K - S)$.

This equation is called a Cauchy problem, and as it is proved in [Fri08], if

- the function $\sigma(S,t)$ is non-negative and bounded on $\mathbb{R}^+ \times [0,T]$,
- the function $t \mapsto r(t)$ be bounded and Lipschitz continuous,
- the function $(S,t) \mapsto S\sigma(S,t)$ is Lipschitz continuous on $\mathbb{R}^+ \times [0,T]$,
- the Cauchy data P_0 satisfies $0 \leq P_0(S) \leq c(1+S)$, where c is a given constant.

Then there exist a unique function $P \in C^0(\mathbb{R}^+ \times [0,T])$, C^1-regular and C^2-regular in $\mathbb{R}^+ \times (0,T]$ with respect to t and S respectively, that is a solution to 4.1, 4.2. Also, for a given constant c', it satisfies $0 \leq P(S,t) \leq c'(1+S)$.

The function P is called the strong solution of the Cauchy problem. Nevertheless, in order to derive many numerical solutions for partial differential equations including finite elements method, one requires the concept of a weak formulation, which will be discussed in remaining of this chapter.

4.2 Weak Formulation

Before following the analysis of Black-Scholes partial differential equation, it is required to introduce some function spaces.

The Hilbert space of square integrable functions on $\mathbb{R}^+$ is denoted by $L^2(\mathbb{R}^+)$, with the norm $||v||_{L^2(\mathbb{R}^+)}= \left(\int_{\mathbb{R}^+} v(x)^2 dx\right)^{1/2}$ and the inner product $(v,w)_{L^2(\mathbb{R}^+)} = \int_{\mathbb{R}^+} v(x)w(x)dx$.

We define the function space V by

$$V = \left\{ v \in L^2(\mathbb{R}^+) : x\frac{\partial v}{\partial x} \in L^2(\mathbb{R}^+) \right\} \tag{4.3}$$

where the derivative is regarded as distributions on $L^2(\mathbb{R}^+)$. The corresponding seminorm is defined by $|v|_V= ||x\frac{\partial v}{\partial x}||_{L^2(\mathbb{R}^+)}$, and the inner product is $(v,w)_V = (v,w) + (x\frac{\partial v}{\partial x}, x\frac{\partial w}{\partial x})$.

Another ingredient in the derivation of weak formulation for the Black-Scholes equation is Poincaré's inequality.

Lemma 4.2.1. (Poincaré's inequality) If $v \in V$, then

$$||v||_{L^2(\mathbb{R}^+)} \leq 2||x\frac{\partial v}{\partial x}||_{L^2(\mathbb{R}^+)} \tag{4.4}$$

Proof. The proof of this Lemma is given in [AP05]. □

Observe that based on Poincaré's inequality

- If $|v|_V= 0$, then
$$0 = |v|_V= ||x\frac{\partial v}{\partial x}||_{L^2(\mathbb{R}^+)} \geq \frac{1}{2}||v||_{L^2(\mathbb{R}^+)}$$
thus $||v||_{L^2(\mathbb{R}^+)}= 0$, and therefore $v = 0$.

Consequently, the seminorn defined for function space V is, in fact, a norm.

Before deriving the weak solution, it is also necessary to define the dual space of V (denoted by V').

$$||w||_{V'}= \sup_{v \in V \backslash \{0\}} \frac{(w,v)}{|v|_V}. \tag{4.5}$$

After the brief discussion of function spaces, the next step is to derive the bilinear form. And for doing that, one needs to multiply the 4.1 by a smooth function ϕ on $\mathbb{R}^+$, and obtain,

$$\frac{\partial}{\partial t}P(S,t)\phi(S) - \frac{1}{2}\sigma^2(S,t)S^2\frac{\partial^2}{\partial S^2}P(S,t)\phi(S) - r(t)S\frac{\partial}{\partial S}P(S,t)\phi(S) + r(t)P(S,t)\phi(S) = 0$$

Then, integrate the obtained equation in S on $\mathbb{R}^+$,

$$\frac{\partial}{\partial t}\int_{\mathbb{R}^+} P(S,t)\phi(S)dS - \int_{\mathbb{R}^+} \frac{1}{2}\sigma^2(S,t)S^2\frac{\partial^2}{\partial S^2}P(S,t)\phi(S)dS \\ - \int_{\mathbb{R}^+} r(t)S\frac{\partial}{\partial S}P(S,t)\phi(S)dS + r(t)\int_{\mathbb{R}^+} P(S,t)\phi(S)dS = 0.$$

Next step is to apply the integration by parts to the second expression of above equation,

$$\frac{\partial}{\partial t}\int_{\mathbb{R}^+} P(S,t)\phi(S)dS + \int_{\mathbb{R}^+} \frac{S^2\sigma^2(S,t)}{2}\frac{\partial}{\partial S}P(S,t)\frac{\partial}{\partial S}\phi(S,t)dS \\ + \int_{\mathbb{R}^+}\Big(S\sigma^2(S,t) + S^2\sigma(S,t)\frac{\partial\sigma}{\partial S}(S,t)\Big)\frac{\partial}{\partial S}P(S,t)\phi(S,t)dS \\ - \int_{\mathbb{R}^+} r(t)S\frac{\partial}{\partial S}P(S,t)\phi(S)dS + r(t)\int_{\mathbb{R}^+} P(S,t)\phi(S)dS = 0.$$

Simple mathematical simplification results in,

$$\frac{\partial}{\partial t}\int_{\mathbb{R}^+} P(S,t)\phi(S)dS + \int_{\mathbb{R}^+} \frac{S^2\sigma^2(S,t)}{2}\frac{\partial}{\partial S}P(S,t)\frac{\partial}{\partial S}\phi(S,t)dS \\ + \int_{\mathbb{R}^+}\Big(-r(t) + \sigma^2(S,t) + S\sigma(S,t)\frac{\partial\sigma}{\partial S}(S,t)\Big)S\frac{\partial}{\partial S}P(S,t)\phi(S,t)dS \\ + r(t)\int_{\mathbb{R}^+} P(S,t)\phi(S)dS = 0. \quad (4.6)$$

Finally, we introduce the bilinear form of the Black-Scholes partial differential equation with local volatility as,

$$a_t(v,w) = \int_{\mathbb{R}^+} \frac{S^2\sigma^2(S,t)}{2}\frac{\partial v}{\partial S}\frac{\partial w}{\partial S}dS \\ + \int_{\mathbb{R}^+}\Big(-r(t) + \sigma^2(S,t) + S\sigma(S,t)\frac{\partial\sigma}{\partial S}(S,t)\Big)S\frac{\partial v}{\partial S}wdS + r(t)\int_{\mathbb{R}^+} vwdS. \quad (4.7)$$

Using this bilinear form, one can define the weak formulation of 4.1, 4.2.
If $C^0([0,T];L^2(\mathbb{R}^+))$ be the space of continuous functions with values in $L^2(\mathbb{R}^+)$, and $L^2(0,T;V)$ be the space of square integrable functions with values in V, then the weak formulation is defined as,

Definition 4.2.1. Weak formulation: Find $P \in C^0([0,T];L^2(\mathbb{R}^+)) \cap L^2(0,T;V)$, such that $\frac{\partial P}{\partial t} \in L^2(0,T;V')$, satisfying

$$P|_{t=0} = P_0 \quad \text{in} \quad \mathbb{R}^+ \quad \text{and for a.e.} \quad t \in (0,T) \quad (4.8)$$

$$\forall v \in V, \quad (\frac{\partial P}{\partial t}(t), v) + a_t(P(t), v) = 0. \quad (4.9)$$

In order to continue the mathematical analysis of Black-Scholes equation with local volatility and to prove that there is a unique solution for the weak formulation, one requires to show that the bilinear form is continuous and to prove Gårding's inequality. And for doing this, making some assumptions on coefficients σ and r is necessary.

Assumption 1. The interest rate r is nonnegative and continuous and σ is continuous and sufficiently regular, such that

1. There exist nonnegative constants, $\underline{\sigma}$ and $\bar{\sigma}$, such that for all $t \in [0, T]$ and all $S \in \mathbb{R}^+$,
$$0 \leq \underline{\sigma} \leq \sigma(S,t) \leq \bar{\sigma} \tag{4.10}$$
2. There exist a positive constant C_σ, such that for all $t \in [0, T]$ and all $S \in \mathbb{R}^+$,
$$\left| S \frac{\partial \sigma}{\partial S}(S,t) \right| \leq C_\sigma. \tag{4.11}$$

Regarding the ingredients of this section, including the definitions, assumptions and bilinear form 4.7, we can derive the theorem which proves that there is a unique solution for the weak formulation, and this is the objective of the next section.

4.3 Final Result

The goal of this section is to apply the theory of Lions and Magenes [LM12], described in [AP05], in order to provide the ingredients, which are required to prove that there is a unique solution for the weak formulation. And for doing this, first of all, one needs to prove the continuity and weak ellipticity of the bilinear form 4.7, then states the final result.

Lemma 4.3.1. Considering the assumption 1, the bilinear form 4.7 is continuous on V. In other words, there is a constant $\mu \geq 0$, such that for all $v, w \in V$,

$$|a_t(v,w)| \leq \mu |v|_V |w|_V. \tag{4.12}$$

Proof. The bilinear form 4.7 consists of three expressions. The idea is to consider each expression separately, and obtain an upper bound for each one, then use the triangle inequality to obtain an upper bound for the whole bilinear form.

Therefore, using assumption 1, the definition of the function space V, and Cauchy-Schwarz inequality for the first expression of the bilinear form, results in

$$\Big| \int_{\mathbb{R}^+} \frac{S^2\sigma^2(S,t)}{2} \frac{\partial v}{\partial S} \frac{\partial w}{\partial S} dS \Big| = \Big| \int_{\mathbb{R}^+} \frac{\sigma^2(S,t)}{2} S \frac{\partial v}{\partial S} S \frac{\partial w}{\partial S} dS \Big| \leq \frac{\bar{\sigma}^2}{2} |v|_V |w|_V.$$

In the case of the second expression, applying the assumption 1, besides calling $R = \max_{t \in [0,T]} r(t)$, lead to

$$\int_{\mathbb{R}^+} \Big(-r(t) + \sigma^2(S,t) + S\sigma(S,t) \frac{\partial \sigma}{\partial S}(S,t) \Big) S \frac{\partial v}{\partial S} w dS \leq \Big(R + \bar{\sigma}^2 + C_\sigma \bar{\sigma} \Big) |v|_V ||w||_{L^2(\mathbb{R}^+)}$$

Now, applying Poincaré's inequality 4.4, yields

$$\Big(R + \bar{\sigma}^2 + C_\sigma \bar{\sigma}\Big)|v|_V||w||_{L^2(\mathbb{R}^+)} \leq 2\Big(R + \bar{\sigma}^2 + C_\sigma \bar{\sigma}\Big)|v|_V|w|_V.$$

Similarly, the upper bound for the third expression is obtained,

$$\Big|r(t)\int_{\mathbb{R}^+} vwdS\Big| \leq R||v||_{L^2(\mathbb{R}^+)}||w||_{L^2(\mathbb{R}^+)} \leq 4R|v|_V|w|_V.$$

It only remains to determine an appropriate constant μ, based on above upper bounds. Choosing $\mu = \frac{5}{2}\bar{\sigma}^2 + 2C_\sigma\bar{\sigma} + 6R$, completes the proof. □

The next Lemma is Gårding's inequality, which proves the weak ellipticity of the bilinear form. It is worth noting that unlike elliptic partial differential equations that the uniqueness of the solution proved by the Lax-Milgram theorem or Riesz theorem, and thus continuity and full ellipticity of bilinear form is required, in the case of parabolic equations (including the theory of Lions and Magenes [LM12]), the unique solution of the weak formulation is typically proved by continuity and weak ellipticity.

Lemma 4.3.2. (Gårding's inequality) Regarding assumption 1, there exist a constant $\lambda \geq 0$, such that for all $v \in V$,

$$a_t(v, v) \geq \frac{\underline{\sigma}^2}{4}|v|_V^2 - \lambda||v||^2_{L^2(\mathbb{R}^+)}. \tag{4.13}$$

Proof. We follow the approach similar to the proof of Lemma 4.3.1. As a consequence, the bound for each expression of the bilinear form 4.7 is obtained separately.

Applying the assumption 1, and the definition of the function space V to the first expression leads to

$$\Big|\int_{\mathbb{R}^+} \frac{S^2\sigma^2(S,t)}{2}(\frac{\partial v}{\partial S})^2 dS\Big| \geq \frac{\underline{\sigma}^2}{2}|v|_V^2.$$

By applying the similar reasoning to the proof of Lemma 4.3.1, for the second expression, one obtains

$$\int_{\mathbb{R}^+}\Big(-r(t) + \sigma^2(S,t) + S\sigma(S,t)\frac{\partial \sigma}{\partial S}(S,t)\Big)S\frac{\partial v}{\partial S}vdS \leq \Big(R + \bar{\sigma}^2 + C_\sigma\bar{\sigma}\Big)|v|_V||v||_{L^2(\mathbb{R}^+)}.$$

Then, using the square of a difference factorization ($(a-b)^2 = a^2 + b^2 - 2ab$), with $a = \frac{\underline{\sigma}}{2}|v|_V$ and $b = \frac{R+\bar{\sigma}^2+C_\sigma\bar{\sigma}}{\underline{\sigma}}||v||_{L^2(\mathbb{R}^+)}$, yields

$$\Big(R + \bar{\sigma}^2 + C_\sigma\bar{\sigma}\Big)|v|_V||v||_{L^2(\mathbb{R}^+)} \leq \frac{\underline{\sigma}^2}{4}|v|_V + \lambda||v||^2_{L^2(\mathbb{R}^+)}$$

where $\lambda = \frac{\big(R+\bar{\sigma}^2+C_\sigma\bar{\sigma}\big)^2}{\underline{\sigma}^2}$.
In the case of the third expression, one obtains,

$$\Big|r(t)\int_{\mathbb{R}^+} v^2 dS\Big| \leq 4R|v|_V^2.$$

Then, elementary mathematical simplification leads to the desired result, and this terminates the proof. □

Theorem 4.3.3. Considering the assumption 1, if $P_0 \in L^2(\mathbb{R}^+)$, then there is a unique solution for the weak formulation 4.2.1, and the following estimate for all t , $0 < t < T$ holds,

$$e^{-2\lambda t}||P(t)||^2_{L^2(\mathbb{R}^+)}+\frac{1}{2}\underline{\sigma}^2\int_0^t e^{-2\lambda\tau}|P(\tau)|^2_V d\tau \leq ||P_0||_{L^2(\mathbb{R}^+)}. \tag{4.14}$$

Proof. As we mentioned earlier, the derivation of this section is according to the work by Lions and Magenes [LM12]. The proof of the first part of this theorem, which is based on an abstract theory of deriving energy estimates for the approximate solutions, is given there. Moreover, Evans and Lawrence [Eva10] proposed another approach to prove the unique solution of the weak formulation using the energy estimates.

In the case of the estimate, first of all take $v = P(t)e^{-2\lambda t}$, then plug this value into the weak formulation 4.9. This yields,

$$(\frac{\partial P}{\partial t}, P(t)e^{-2\lambda t}) + a_t(P(t), P(t)e^{-2\lambda t}) = 0. \tag{4.15}$$

Now, apply Gårding's inequality 4.13, to obtain,

$$e^{-2\lambda t}a_t(P(t), P(t)) \geq e^{-2\lambda t}\Big(\frac{\underline{\sigma}^2}{4}|P(t)|^2_V - \lambda||P(t)||^2_{L^2(\mathbb{R}^+)}\Big). \tag{4.16}$$

Then, combine 4.15 and 4.16, integrate in time between 0 and t, use integration by parts and finally some mathematical simplification, result in

$$e^{-2\lambda t}||P(t)||^2_{L^2(\mathbb{R}^+)}+\frac{1}{2}\underline{\sigma}^2\int_0^t e^{-2\lambda\tau}|P(\tau)|^2_V d\tau \leq ||P_0||_{L^2(\mathbb{R}^+)}. \tag{4.17}$$

□

It should be mentioned that due to the fact that the payoff function of the call option is not in $L^2(\mathbb{R}^+)$, the above theorem does not apply to European call options. Therefore, in order to obtain the price for call options, one needs to apply the put-call parity (which will be discussed in chapter 5) or work with different Sobolev space.

By proving the unique solution of the weak formulation and obtaining a regularity estimate of Black-Scholes equation with local volatility for the European put option, this chapter terminates. In the next chapter, in addition to the derivation of the Black-Scholes formula, we apply finite element methods and derive a numerical method for approximating the Black-Scholes equation with local volatility for the European put option 4.1.

5 Solution to the Black-Scholes equation

The goal of this chapter is to present the solution to the Black-Scholes partial differential equation. In the case of the original Black-Scholes equation, it is possible to derive an exact solution, which is called a Black-Scholes formula and is discussed in the first section of this chapter. However, in the case of the Black-Scholes equation with local volatility, one requires to derive numerical methods.

There are three different class of methods for solving partial differential equations numerically, including Finite Difference Method (FDM), Finite Element Method (FEM), and Finite Volume Method (FVM). Nevertheless, due to the fact that the finite element method is more flexible and also support by strong theory for error estimation, we decide to apply Finite element method for deriving a numerical solution of Black-Scholes partial differential equation with local volatility.

5.1 Black-Scholes Formula

In this section, the Black-Scholes partial differential equation with constant coefficients is regarded, and the objective is to follow the procedure in [WDH93], to derive the explicit solution.

For doing this, consider the Black-Scholes partial differential equation with constant coefficients for call options,

$$\frac{\partial C}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + rS\frac{\partial C}{\partial S} - rC = 0 \tag{5.1}$$

with boundary conditions,

$$C(0,t) = 0, \quad C(S,t) \sim S \text{ as } S \to \infty$$

and initial condition,

$$C(S,0) = \max(S-K,0).$$

The first step is to change the following variables,

$$S = Ke^x, \qquad t = \frac{\tau}{\frac{1}{2}\sigma^2}, \qquad C = Kv(x,\tau)$$

and by taking $k_1 = \frac{r}{\frac{1}{2}\sigma^2}$, one obtains the equation,

$$\frac{\partial v}{\partial \tau} = \frac{\partial^2 v}{\partial x^2} + (k_1 - 1)\frac{\partial v}{\partial x} - k_1 v \tag{5.2}$$

with new initial condition,

$$v(x,0) = \max(e^x - 1, 0).$$

It is worth mentioning that, there is only one parameter, k_1, in the equation now. The next step is another change of variable,

$$v = e^{\alpha x + \beta \tau} u(x, \tau)$$

where α and β are constants to be found. Therefore,

$$\frac{\partial}{\partial \tau} e^{\alpha x + \beta \tau} u(x, \tau) = \frac{\partial^2}{\partial x^2} e^{\alpha x + \beta \tau} u(x, \tau) + (k_1 - 1) \frac{\partial}{\partial x} e^{\alpha x + \beta \tau} u(x, \tau) - k_1 e^{\alpha x + \beta \tau} u(x, \tau)$$

and simple differentiation yields,

$$\beta u + \frac{\partial u}{\partial \tau} = \alpha^2 u + 2\alpha \frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2} + (k_1 - 1)(\alpha u + \frac{\partial u}{\partial x}) - k_1 u.$$

Then, one requires to eliminate u and $\frac{\partial u}{\partial x}$. Taking,

$$\beta = \alpha^2 + (k_1 - 1)\alpha - k_1 \quad \text{and} \quad \alpha = \frac{1}{2}(1 - k_1)$$

leads to,

$$-\frac{1}{4}(k_1 + 1)^2 u + \frac{\partial u}{\partial \tau} = (\frac{1 - k_1}{2})^2 u + (1 - k_1) \frac{\partial u}{\partial x} + \frac{\partial^2 u}{\partial x^2} + (k_1 - 1)\Big(\frac{1 - k_1}{2} u + \frac{\partial u}{\partial x}\Big)$$

elementary mathematical simplifications result in,

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2} \qquad \text{for} \quad -\infty < x < \infty, \ \tau > 0$$

with initial condition,

$$u(x, 0) = u_0(x) = \max(e^{\frac{1}{2}(k_1+1)x} - e^{\frac{1}{2}(k_1-1)x}, 0) \tag{5.3}$$

also, recall that with this values for α and β, then,

$$v = e^{-\frac{1}{2}(k_1-1)x - \frac{1}{4}(k_1+1)^2 \tau} u(x, \tau)$$

as it is mentioned in [WDH93], the solution to this diffusion equation is given by,

$$u(x, \tau) = \frac{1}{2\sqrt{\pi \tau}} \int_{-\infty}^{\infty} u_0(s) e^{\frac{-(x-s)^2}{4\pi}} ds. \tag{5.4}$$

The only remaining step is to evaluate the integral. For doing this, using the change of variable $x' = \frac{x-s}{\sqrt{2\tau}}$, and applying the substitution rule yields,

$$u(x, \tau) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u_0(x' \sqrt{2\tau} + x) e^{-\frac{1}{2} x'^2} dx'.$$

Now, plug 5.3 into the above equation, to obtain,

$$u(x, \tau) = \\ \frac{1}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\tau}}}^{\infty} e^{\frac{1}{2}(k_1+1)(x + x'\sqrt{2\tau})} e^{-\frac{1}{2}x'^2} dx' - \frac{1}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\tau}}}^{\infty} e^{\frac{1}{2}(k_1-1)(x + x'\sqrt{2\tau})} e^{-\frac{1}{2}x'^2} dx'$$

$$= I_1 + I_2.$$

Integrating I_1, yields,

$$I_1 = \frac{1}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\tau}}}^{\infty} e^{\frac{1}{2}(k_1+1)(x+x'\sqrt{2\tau})} e^{-\frac{1}{2}x'^2} dx' = e^{\frac{1}{2}(k_1+1)x+\frac{1}{4}(k_1+1)^2\tau} N(d_1)$$

where,

$$N(d_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_1} e^{-\frac{1}{2}s^2} ds$$

and,

$$d_1 = \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k_1+1)\sqrt{2\tau}.$$

Similarly, one obtains the same result for I_2, by replacing K_1+1 by K_1-1. Therefore,

$$I_2 = \frac{1}{\sqrt{2\pi}} \int_{-\frac{x}{\sqrt{2\tau}}}^{\infty} e^{\frac{1}{2}(k_1-1)(x+x'\sqrt{2\tau})} e^{-\frac{1}{2}x'^2} dx' = e^{\frac{1}{2}(k_1-1)x+\frac{1}{4}(k_1-1)^2\tau} N(d_2)$$

where,

$$N(d_2) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d_2} e^{-\frac{1}{2}s^2} ds$$

and,

$$d_2 = \frac{x}{\sqrt{2\tau}} + \frac{1}{2}(k_1-1)\sqrt{2\tau}.$$

Thus,

$$u(x,\tau) = e^{\frac{1}{2}(k_1+1)x+\frac{1}{4}(k_1+1)^2\tau} N(d_1) - e^{\frac{1}{2}(k_1-1)x+\frac{1}{4}(k_1-1)^2\tau} N(d_2).$$

On the other hand, recall the definition of v,

$$v(x,\tau) = e^{-\frac{1}{2}(k_1-1)x-\frac{1}{4}(k_1+1)^2\tau}\Big(e^{\frac{1}{2}(k_1+1)x+\frac{1}{4}(k_1+1)^2\tau} N(d_1) - e^{\frac{1}{2}(k_1-1)x+\frac{1}{4}(k_1-1)^2\tau} N(d_2)\Big)$$
$$= e^x N(d_1) - e^{-k_1\tau} N(d_2).$$

Consequently,

$$C(x,\tau) = Kv(x,\tau) = Ke^x N(d_1) - Ke^{-k_1\tau} N(d_2).$$

Finally, in order to recover the formula based on the parameters of the Black-Scholes equation, apply the following change of variables,

$$x = \log(\frac{S}{K}), \quad \tau = \frac{\sigma^2}{2}t.$$

Then,

$$C(S,t) = SN(d_1) - Ke^{-rt} N(d_2) \tag{5.5}$$

where,

$$d_1 = \frac{\log(\frac{S}{K}) + (r + \frac{\sigma^2}{2})t}{\sigma\sqrt{t}} \quad \text{and} \quad d_2 = \frac{\log(\frac{S}{K}) + (r - \frac{\sigma^2}{2})t}{\sigma\sqrt{t}}$$

and this is called the Black-Scholes formula for European call options. In order to derive a similar formula for a European put options, the naive way is to follow the same procedure for put options. However, the smarter way is to use the obtained result for call options and apply the put-call parity,

Definition 5.1.1. Put-Call parity: Put-call parity states the relationship between the European call option and European put option of the same strike and maturity on the same stock.[Chr96]

$$C - P = S - Ke^{-rt}.$$

Therefore, the formula for the European put option is obtained by,

$$P(S,t) = -SN(-d_1) + Ke^{-rt}N(-d_2) \tag{5.6}$$

where N,d_1, and d_2 are the same as the formula for the European call option.

Graph 5.1 and 5.2 show the value for European call and European put options respectively, using the Black-Scholes formula (where $\sigma = 0.5$, $r = 0.05$, $K = 50$, and the domain $[0, 100] \times [0, 1]$).

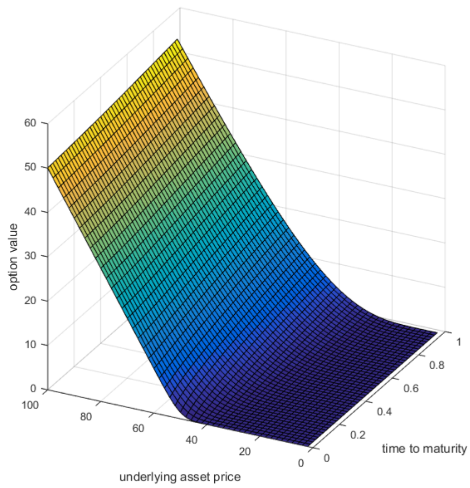

Figure 5.1: European call option using the Black-Scholes formula

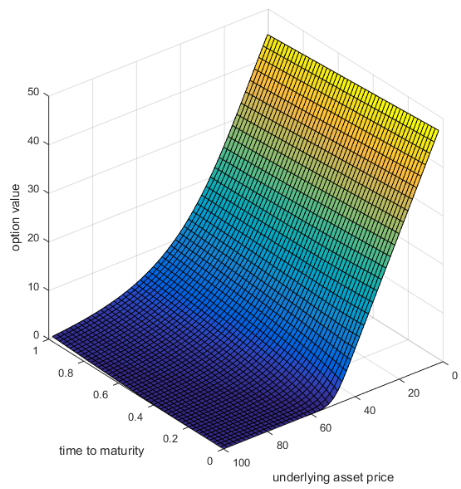

Figure 5.2: European put option using the Black-Scholes formula

5.2 Numerical Approximation

As it is mentioned earlier, the objective of this section is to derive the numerical method to approximate the Black-Scholes partial differential equation with local volatility for the European put options. We follow the discussion of Achdou and Pironneau in [AP05] and [PA09], to derive a first order finite element method based on the Crank-Nicolson scheme to the Black-Scholes parabolic differential equation with local volatility.

As it is demonstrated in [SA86], the idea behind the finite element method is to choose finite number of functions (called "shape functions"), and to approximate the exact solution by a combination of these shape functions. Shape functions are usually chosen to be piecewise polynomial, because it allows us to fit the geometry of the problem and generate polynomials. By doing this, the computer can assemble a discrete problem and solve it.

Deriving finite element method usually consists of 4 main steps. First of all, weak formulation of the partial differential equation is derived, then discretization of the weak formulation is done, the next step is to define the shape functions (the order of this shape functions is, in fact, the order of finite element method) and rewriting the obtained discrete equation in matrix form, and the final step is to solve the obtained linear system.

5.2.1 Discretization of Weak Formulation

Recall that the weak formulation of the Black-Scholes partial differential equation with local volatility for European put options was obtained in the previous chapter. Therefore, the next step is to discretize this weak formulation. In the case of parabolic partial differential equations, discretization is done in two steps. First, the semidiscrete problem is derived by discretizing the problem in one of the variables. The second step is full discretization of the problem, which is done by replacing the function space V, by a subspace of finite dimension V_h. In the case of Black-Scholes equation with local volatility, one obtains the time semidiscrete problem by introducing a partition of the interval $[0, T]$ into subintervals $[t_{m-1} - t_m]$, $1 \leq m \leq M$, such that,

$$0 < t_0 < t_1 < t_2 < ... < t_M = T.$$

The size of each interval $t_m - t_{m-1}$ is denoted by Δt_m, and define $\Delta t = \max_{m=1,...,M} \Delta t_m$. As it is mentioned earlier, our objective is to discretize the weak formulation under the Crank-Nicolson scheme. Thus, the semidiscrete problem is to find $P^m \in V$, $m = 0, ..., M$, such that $P^0 = P_0$, where it is assumed that $P_0 \in V$, and for all $P^m \in V$ and for all $v \in V$,

$$\Big(\frac{P^m - P^{m-1}}{\Delta t_m}, v\Big) + \frac{1}{2}\Big(a(P^m, v) + a(P^{m-1}, v)\Big) = 0$$

where a is the bilinear form defined earlier. It is also worth noting that the implicit Euler scheme is used to approximate derivative in t.

In order to tackle the second part of the discretization process, one requires to partition $[0, \bar{S}]$ into subintervals $\kappa_i = [S_{i-1}, S_i]$, $1 \leq i \leq N + 1$, such that,

$$0 < S_0 < S_1 < S_2 < ... < S_N < S_{N+1} = \bar{S}$$

where $\bar{S}$ is sufficiently large upper bound for underlying asset price (usually regarded as $1.5K$ or $2K$). We denote the length of the interval κ_i by h_i, and h is defined as $h =$

$\max_{i=1,..,N+1} h_i$. We also define the mesh $\mathcal{T}_h$ of the interval $[0,\bar{S}]$ as $\{\kappa_1,\kappa_2,...,\kappa_{N+1}\}$. Moreover, one requires to assume that there exists a node in $\mathcal{T}_h$ such that coincides with the strike K.

Then, we need to define the discrete space V_h,

$$V_h = \big\{v_h \in C^0([0,\bar{S}]),\ v_h(\bar{S}) = 0;\ \text{such that for all } \kappa \in \mathcal{T}_h,\ v_h|_\kappa \text{ is affine}\big\} \tag{5.7}$$

where the assumption in mesh secures that $P_0 \in V_h$.

Thus, the discrete problem obtained by applying the Crank-Nicolson scheme in time is: Find $(P_h^m)_{0\le m\le M}$, $P_h^m \in V_h$ satisfying,

$$P_h^0 = p_0 \tag{5.8}$$

and for all $m,\ 1 \le m \le M$, and for all $v_h \in V_h$,

$$(P_h^m - P_h^{m-1}, v) + \frac{\Delta t_m}{2}\Big(a_m(P_h^m, v_h) + a_{m-1}(P_h^{m-1}, v_h)\Big) = 0 \tag{5.9}$$

where $a_m = a_{t_m}$, and as it is defined earlier,

$$\begin{aligned} a_t(v,w) = &\int_0^{\bar{S}} \frac{S^2\sigma^2(S,t)}{2}\frac{\partial v}{\partial S}\frac{\partial w}{\partial S}dS \\ &+ \int_0^{\bar{S}} \Big(-r(t) + \sigma^2(S,t) + S\sigma(S,t)\frac{\partial \sigma}{\partial S}(S,t)\Big)S\frac{\partial v}{\partial S}wdS + r(t)\int_0^{\bar{S}} vwdS. \end{aligned} \tag{5.10}$$

5.2.2 Discrete Problem in Matrix Form

The next step is to define shape functions. We apply the first order polynomials to approximate the solution, thus the finite element method is first order. Consider the piecewise linear function w_i, where $i = 0,1,...,N+1$, that is equal to 1 at S_i, and 0 at any other nodes, defined by,

$$w_i(S) = \begin{cases} \frac{S-S_{i-1}}{h_i} & \text{if } S \in (S_{i-1}, S_i) \\ \frac{S_{i+1}-S}{h_{i+1}} & \text{if } S \in (S_i, S_{i+1}). \end{cases}$$

Note that,

$$\frac{\partial w^i}{\partial S} = \begin{cases} \frac{1}{h_i} & \text{if } S \in (S_{i-1}, S_i) \\ \frac{1}{h_{i+1}} & \text{if } S \in (S_i, S_{i+1}). \end{cases}$$

Then, $(w_i)_{i=0,...,N}$ is the nodal basis of V_h and,

$$P_h^m(S) = \sum_{i=0}^{N} P_h^m(S_i) w_i(S).$$

Furthermore, for $0 \le i,j \le N$, let $M \in \mathbb{R}^{(N+1)\times(N+1)}$ be the mass matrix, defined by $M_{i,j} = (w^i, w^j)$, and $A^m \in \mathbb{R}^{(N+1)\times(N+1)}$ denotes the stiffness matrix defined by $A_{i,j}^m = a_{t_m}(w^j, w^i)$. Also, note that $P^m = (P_h^m(S_0), P_h^m(S_1), ..., P_h^m(S_N))^T$ and $P^0 = (P_0^m(S_0), P_0^m(S_1), ..., P_0^m(S_N))^T$. Then, the matrix form of the equation 5.9 is,

$$M(P^m - P^{m-1}) + \frac{\Delta t_m}{2}(A^m P^m + A^{m-1}P^{m-1}) = 0. \tag{5.11}$$

Then, we can calculate the entries of mass matrix M and stiffness matrices A^m. Regarding the definition of shape functions and their derivatives, simple mathematical calculations lead to,

$$
\begin{aligned}
&\int_0^{\bar{S}} w^i w^{i-1} dS = \frac{h_i}{6} && \int_0^{\bar{S}} S w^i \frac{\partial w^{i-1}}{\partial S} dS = -\frac{2S_{i-1} + S_i}{6} \\
&\int_0^{\bar{S}} {w^i}^2 dS = \frac{h_i + h_{i+1}}{3} && \int_0^{\bar{S}} S w^i \frac{\partial w^i}{\partial S} dS = -\frac{h_{i+1} + h_i}{6} \quad \text{if } i > 0 \\
&\int_0^{\bar{S}} {w^0}^2 dS = \frac{h_1}{3} && \int_0^{\bar{S}} S w^0 \frac{\partial w^0}{\partial S} dS = -\frac{h_1}{6} \\
&\int_0^{\bar{S}} w^{i+1} w^i dS = \frac{h_{i+1}}{6} && \int_0^{\bar{S}} S w^i \frac{\partial w^{i+1}}{\partial S} dS = \frac{2S_{i+1} + S_i}{6}.
\end{aligned}
$$

Finally, using these calculations and applying the definition of the stiffness matrices, one obtains the entries of A^m as,

$$
\begin{aligned}
A^m_{0,0} &= \frac{r(t_m)}{2} h_1 \\
A^m_{i,i} &= \frac{S_i^2 \sigma^2(t_m, S_i)}{2} \Big(\frac{1}{h_i} + \frac{1}{h_{i+1}} \Big) + \frac{r(t_m)}{2} \Big(h_i + h_{i+1} \Big) , \quad \text{for } 1 \le i \le N \\
A^m_{i,i-1} &= -\frac{S_i^2 \sigma^2(t_m, S_i)}{2h_i} + \frac{r(t_m) S_i}{2} , \quad \text{for } 1 \le i \le N \\
A^m_{i,i+1} &= -\frac{S_i^2 \sigma^2(t_m, S_i)}{2h_{i+1}} - \frac{r(t_m) S_i}{2} , \quad \text{for } 0 \le i \le N-1.
\end{aligned}
$$

It is worth noting that, the entries of the stiffness matrices were obtained for general case, where the mesh is not necessarily uniform. However, in this chapter we regard only uniform meshes, in which $h = h_i$, for all i. Similarly, entries of the mass matrix M is obtained by,

$$
\begin{aligned}
M_{0,0} &= \frac{h_1}{3} \\
M_{i,i} &= \frac{h_i + h_{i+1}}{3} , \quad \text{for } 1 \le i \le N \\
M_{i,i-1} &= \frac{h_i}{6} , \quad \text{for } 1 \le i \le N \\
M_{i,i+1} &= \frac{h_{i+1}}{6} , \quad \text{for } 0 \le i \le N-1.
\end{aligned}
$$

Therefore, it only remains to solve the obtained linear system.

5.2.3 Conjugate Gradient Method

There are two main families of methods to solve the linear system $Ax = b$, direct methods and iterative methods. In direct methods, such as Gaussian elimination, the problem is solved by finite number of steps, and if we disregard rounding errors, then the obtained solution is exact. On the other hand, iterative methods begin by an initial guess, and produce better approximation, using the previous approximation [Str93]. Direct methods are usually more straightforward to apply, however, when the dimension of the problem is big or the matrix of the linear system is sparse, they are costly and it is more appropriate to apply iterative methods. Due to the fact that the matrix obtained by finite element method is sparse, and in general, by increasing the mesh points can become significantly big, we decide to apply an iterative method.

Conjugate gradient method is an iterative method which is widely used to solve the linear systems. As it is shown in [KS88], when applying to the technical applications, such as finite difference and finite element methods, conjugate gradient method converges very quickly. The history of the conjugate gradient method began by articles of Hestenes and Stiefel [HS52], and Lanczos [Lan50], which revolutionized the numerical solution of large linear system of equations [KNGK12].

We follow the discussion in [Pyt08], to illustrate the conjugate gradient method, and apply it to obtain the solution of the Black-Scholes partial differential equations with local volatility. The main objective of the conjugate gradient method is to choose direction vectors $\{p_i\}_{i=1}^n$, which are conjugate, i.e.

$$p_i^T A p_j = 0 \ , \quad \text{for all } i \neq j, \ i,j = 1,...,n. \tag{5.12}$$

Then, apply the expression $x_{k+1} = x_k + \alpha_k p_k$, to obtain a better approximation of the problem, where α is a constant to be found.

Regard the problem of finding $x \in \mathbb{R}^n$ satisfying the linear system,

$$Ax = b \ , \quad A \in \mathbb{R}^{n\times n}, \ b \in \mathbb{R}^n$$

where, A is real, symmetric and positive definite. The fact of the matter is that, definiteness of A eliminates division by zero, and symmetry provides a short recurrence [Str07]. This problem is equivalent to the minimization of,

$$f(x) = \frac{1}{2} x^T A x - b^T x \tag{5.13}$$

over the space $\mathbb{R}^n$. The next step is to find the constant α. For doing this, assume that the residual of the problem $Ax = b$ is denoted by $r_k = Ax_k - b = g_k$. Then, form the function,

$$\phi_k(\alpha) = f(x_k + \alpha p_k).$$

Our goal is to find α_k that minimize $\phi_k(\alpha)$, i.e. $\phi_k(\alpha_k) = 0$. Thus,

$$\phi_k(\alpha_k) = g(x_k + \alpha_k p_k)^T p_k = (A(x_k + \alpha_k p_k) - b)^T p_k = 0$$

and this yields,

$$\alpha = -\frac{r_k^T p_k}{p_k^T A p_k}. \tag{5.14}$$

It remains to calculate the direction p_k based on the approach of Hestenes and Stiefel [HS52]. They proposed the method to calculate the direction p_k, based on the previous direction and the current gradient. As a result, under the assumption that β_k is a coefficient which depends on the data from the current and previous iterations, one obtains,

$$p_k = -r_k + \beta_k p_{k-1} = -g + \beta_k p_{k-1}$$

applying 5.12 leads to,

$$p_k^T A p_{k-1} = -r_k^T A p_{k-1} + \beta_k p_{k-1}^T A p_{k-1} = 0.$$

Thus,

$$\beta_k = \frac{r_k^T A p_{k-1}}{p_{k-1}^T A p_{k-1}}. \tag{5.15}$$

It is worth mentioning that the convergence analysis of the conjugate gradient method is given in [Pyt08]. Based on the above discussion, one can derive the following algorithm for conjugate gradient method,

Algorithm 1. (The conjugate gradient algorithm)

1. Choose an initial guess $x_1 \in \mathbb{R}^n$. Set
$$k_1 = 1 \quad \text{and} \quad p_1 = -r_1 = -g_1.$$
2. Calculate,
$$\alpha_k = -\frac{r_k^T p_k}{p_k^T A p_k}.$$
3. Set $x_{k+1} = x_k + \alpha_k p_k$ and $r_{k+1} = Ax_{k+1} - b$. If $||r_{k+1}||$ is sufficiently small (in comparison to the given tolerance), then STOP. Else, calculate,
$$\beta_{k+1} = \frac{r_{k+1}^T A p_k}{p_k^T A p_k} \quad \text{and} \quad p_{k+1} = -r_{k+1} + \beta_{k+1} p_k.$$
4. Increase k by one and go to step 2.

We applied the conjugate gradient method to find the approximate solution of Black-Scholes equation for European put options, obtained by finite element method introduced in this chapter. The obtained value for the put option with $K = 50$, $r = 0.05$, $\sigma = 0.5$, 50×50 nodes, and the domain $[0, 100] \times [0, 1]$ is shown in the figure 5.3.

In addition, since we have obtained the explicit solution of Black-Scholes equation for put options, we can compare it with the approximate solution of finite element methods. Figure 5.4 shows the point-wise error produced by the Crank-Nicolson scheme and first order finite element method in the domain $[0, 200] \times [0, 1]$ with 50×50 nodes. According to our experiments, the L_2-norm of the error for 200×200 nodes is less than 0.1%, which can be regarded as an acceptable result.

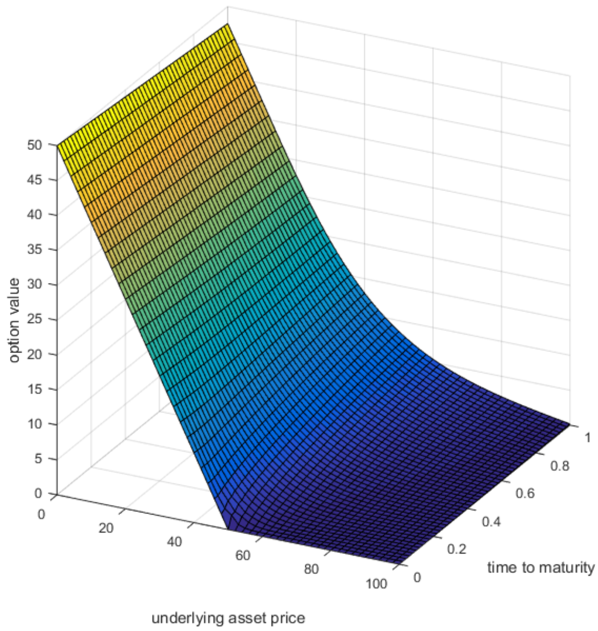

Figure 5.3: European put option using the finite element method

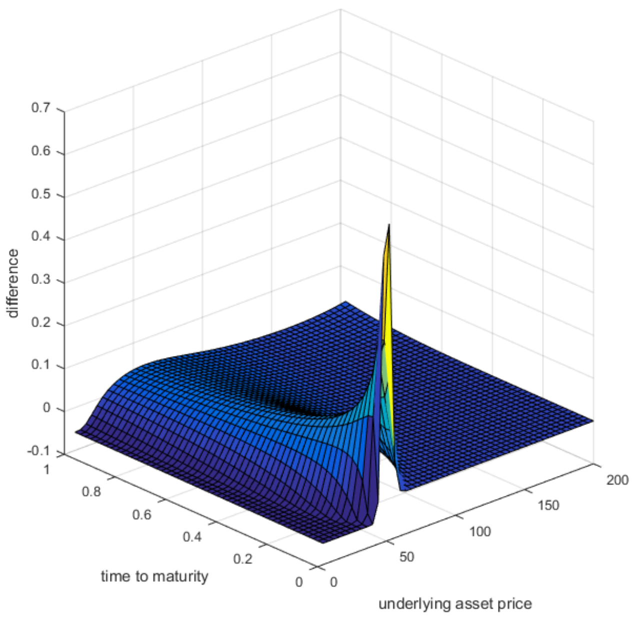

Figure 5.4: The point-wise error produced by the Crank-Nicolson scheme and first order finite element method

6 Adaptive Mesh Refinement

In this chapter, we would like to derive a posteriori error analysis to the finite element method, that applied in the previous chapter to discretize and solve the Black-Scholes partial differential equation with local volatility. We are going to apply the strategy mentioned in the paper [BBM05], which consists of obtaining two error indicators of the residual type, that one of them is global with respect to price variable S, and local with respect to time t, while the second one is local with respect to both price and time.

In order to derive the analysis, one needs some contents of previous chapters. Thus, we discuss and rewrite them in the first section of this chapter. The second section is devoted to deriving reliable and efficient error estimators. And in the third section, we propose a strategy to apply the obtained error indicators, so as to derive adaptive mesh refinement.

It is also should be noted that the contents of this chapter follow the procedure of the fifth chapter of the book by Achdou and Pironneau [AP05].

6.1 Discretization of Black-Scholes Equation with Local Volatility

As it is discussed earlier, in this section, we would like to reformulate and expand, the fundamental results of previous chapters, which is needed for the analysis of this chapter.

We regard the Black-Scholes partial differential equation with local volatility for European put options,

$$\begin{aligned}
&\frac{\partial u}{\partial t} - \frac{\sigma^2 S^2}{2}\frac{\partial^2 u}{\partial S^2} - rS\frac{\partial u}{\partial S} + ru = 0 \quad \text{in } \Omega \times (0,T] \\
&u(\bar{S},t) = 0 \quad \text{where } t \in (0,T] \\
&u(S,0) = u_0(S) = \max(S-K,0) \quad \text{in } \Omega
\end{aligned}$$

where $\Omega = (0,\bar{S})$, with the assumption that the volatility $\sigma(S,t)$ and the interest rate $r(t)$, are smooth functions. Moreover, we assume that there exist constants $\underline{\sigma}$, $\bar{\sigma}$, C_σ and R, where $0 \le \underline{\sigma} \le \bar{\sigma}$, $C_\sigma \ge 0$ and $R > 0$, such that,

$$\underline{\sigma} \le \sigma(S,t) \le \bar{\sigma} \quad \text{in} \quad \mathbb{R}^+ \times [0,T] \tag{6.1}$$

$$|S\frac{\partial \sigma}{\partial S}| \le C_\sigma \quad \text{in} \quad \mathbb{R}^+ \times [0,T] \tag{6.2}$$

$$0 \le r(t) \le R \quad \text{in} \quad [0,T]. \tag{6.3}$$

In the case of function spaces, besides L_2 and V, we require to introduce the spaces $\mathcal{D}(\mathbb{R}^+)$ and $\mathcal{D}(\bar{\Omega})$, which are the space of infinitely differentiable functions with compact support in $\mathbb{R}^+$ and Ω respectively. Then, as it is shown in [AP05], $\mathcal{D}(\bar{\Omega})$ is densely embedded in V. We define V_0 as the closure of $\mathcal{D}(\Omega)$ in V. Therefore, V_0 is a subspace of V, which contains the functions that are zero at $\bar{S}$, then its dual space V_0', is defined by,

$$||w||_{V_0'} = \sup_{v \in V_0} \frac{(w,v)}{|v|_V}. \tag{6.4}$$

Then, the bilinear form is defined by,

$$a_t(u,v) = \Big(\frac{\sigma^2}{2}S\frac{\partial u}{\partial S}, S\frac{\partial v}{\partial S}\Big) + \Big((- r + \sigma^2 + S\sigma\frac{\partial \sigma}{\partial S})S\frac{\partial u}{\partial S}, v\Big) + r(u,v). \tag{6.5}$$

As a consequence, the weak formulation is defined as,

Find $u \in C^0([0,T]; L^2(\Omega)) \cap L^2(0,T; V_0)$, such that

$$u|_{t=0} = u_0 \quad \text{in} \quad \Omega \quad \text{and for a.e.} \quad t \in (0,T) \tag{6.6}$$

$$\forall v \in V_0, \quad (\frac{\partial u}{\partial t}(t), v) + a_t(u(t), v) = 0. \tag{6.7}$$

The next step is to prove the boundedness of the bilinear form. For doing this, assume that μ is a positive constant, then for all $v, w \in V_0$,

$$|a_t(v,w)| \leq \mu |v|_V |w|_V. \tag{6.8}$$

Furthermore, assume that there exists a constant $\lambda > 0$, then based on the Gårding's inequality,

$$\forall t \in [0,T], \ \forall v \in V_0, \quad a_t(v,v) \geq \frac{\underline{\sigma}^2}{4}|v|_V^2 - \lambda ||v||^2. \tag{6.9}$$

Finally, we require to introduce the norm which is bounded with regard to the weak formulation 6.7, thus define,

$$[[v]](t) = \Big(e^{-2\lambda t}||v(t)||^2 + \frac{1}{2}\underline{\sigma}^2 \int_0^t e^{-2\lambda\tau}|v(\tau)|_V^2 d\tau\Big)^{\frac{1}{2}}. \tag{6.10}$$

Then, by plugging $u(t)e^{-2\lambda t}$ instead of v, in 6.7 and some mathematical operations, one obtains,

$$[[u]](t) \leq ||u_0||. \tag{6.11}$$

Thus, applying 6.8, integrating the expression in the norm 6.10 and mathematical simplifications yields,

$$||e^{-\lambda t}\frac{\partial u}{\partial t}||_{L^2(0,T;V_0')} \leq \sqrt{2}\frac{\mu}{\underline{\sigma}}||u_0||. \tag{6.12}$$

6.1.1 Time Semi-Discrete Problem

Recall that we partitioned the interval $[0,T]$, into subintervals $[t_{n-1}, t_n], \ 1 \leq n \leq N$, where

$$0 < t_0 < t_1 < t_2 < ... < t_N = T.$$

Then, $\Delta t_n = t_n - t_{n-1}$. In addition, we define the regularity parameter $\rho_{\Delta t}$,

$$\rho_{\Delta t} = \max_{2 \leq n \leq N} \frac{\Delta t_n}{\Delta t_{n-1}}. \tag{6.13}$$

Therefore, applying the Crank-Nicolson scheme, we obtain the following semidiscrete problem,

$$u^0 = u_0 \tag{6.14}$$

$$\forall n,\ 1 \leq n \leq N,\ \forall v \in V_0, \quad \Big(\frac{u^n - u^{n-1}}{\Delta t_m}, v\Big) + \frac{1}{2}\Big(a_{t_n}(u^n, v) + a_{t_{n-1}}(u^{n-1}, v)\Big) = 0. \tag{6.15}$$

It is also worth mentioning that we introduce the notation $u_{\Delta t}$, to show the function that is affine on each interval and satisfies $u_{\Delta t}(t_n) = u^n$.

The next step is to introduce the discrete version of the norm $[[u]]$. Therefore, regard the identity,

$$(a-b, a) = \frac{1}{2}|a|^2 + \frac{1}{2}|a-b|^2 - \frac{1}{2}|b|^2.$$

Then, replacing v by u^n in the bilinear form 6.5 and applying Gårding's inequality 6.9 yields,

$$(1 - 2\lambda\Delta t_n)||u^n||^2 + \frac{1}{2}\Delta t_n \underline{\sigma}^2 |u^n|_V^2 \leq ||u^{n-1}||^2. \tag{6.16}$$

Now, multiplying both sides of the equation by $\prod_{i=1}^{n-1}(1 - 2\lambda\Delta t_i)$ and taking the sum of equations on n, lead to,

$$\Big(\prod_{i=1}^{n}(1 - 2\lambda\Delta t_i)\Big)||u^n||^2 + \frac{1}{2}\underline{\sigma}^2 \sum_{m=1}^{n} \Delta t_m \Big(\prod_{i=1}^{m-1}(1 - 2\lambda\Delta t_i)\Big)|u^m|_V^2 \leq ||u^0||^2. \tag{6.17}$$

Hence, according to the equation 6.17, we define the discrete norm,

$$[[(v^m)]]_n = \Big(\Big(\prod_{i=1}^{n}(1 - 2\lambda\Delta t_i)\Big)||v^n||^2 + \frac{1}{2}\underline{\sigma}^2 \sum_{m=1}^{n} \Delta t_m \Big(\prod_{i=1}^{m-1}(1 - 2\lambda\Delta t_i)\Big)|v^m|_V^2\Big)^{\frac{1}{2}}. \tag{6.18}$$

As a result,

$$[[(u^m)]]_n \leq ||u^0||. \tag{6.19}$$

The next goal is to find the equivalent relation between the norms $[[(u^m)]]_n$ and $[[u_{\Delta t}]](t_n)$. This problem is addressed by the following lemma,

Lemma 6.1.1. There exists a constant $0 < \alpha \leq \frac{1}{2}$, such that for all $(v^n)_{0 \leq n \leq N} \in V_0^{N+1}$ satisfies,

$$\frac{1}{8}[[(v^m)]]_n^2 \leq [[v_{\Delta t}]]^2(t_n) \leq \max(2, 1 + \rho_{\Delta t})[[(v^m)]]_n^2 + \frac{1}{2}\underline{\sigma}^2 \Delta t_1 |v^0|_V^2 \tag{6.20}$$

where $\Delta t \leq \frac{\alpha}{\lambda}$.

Proof. Regarding the definition of the norm $[[v]]$ and $v_{\Delta t}$ lead to,

$$\frac{e^{2\lambda t_{m-1}}}{\Delta t}\int_{t_{m-1}}^{t_m} e^{-2\lambda\tau}|v_{\Delta t}|_V^2(\tau)d\tau =$$
$$\int_0^1 e^{-2\lambda\Delta t_m \tau}\Big(|v^m|_V^2 \tau^2 + |v^{m-1}|_V^2(1-\tau)^2 + 2(v^{m-1}, v^m)_V \tau(1-\tau)\Big)d\tau.$$

Now, use the inequality $ab \geq -\frac{a^2}{4} - b^2$ with $a = |v^m|_V$ and $b = |v^{m-1}|_V$ and let $\Delta t_m = 0$. Then,

$$\frac{1}{3}\Big(|v^m|_V^2 + |v^{m-1}|_V^2 + (v^m, v^{m-1})_V\Big) \geq \frac{1}{4}|v^m|_V^2.$$

In addition, observe that $e^{-2\lambda\Delta t_m \tau}$ is continuous with regard to Δt_m. Therefore, there exists a constant $0 < \alpha < \frac{1}{2}$ that,

$$\int_{t_{m-1}}^{t_m} e^{-2\lambda\tau}|v_{\Delta t}|_V^2(\tau)d\tau \geq \frac{\Delta t_m}{8} e^{-2\lambda t_{m-1}}|v^m|_V^2.$$

On the other hand, if $2\lambda\Delta t < 1$, then $\prod_{i=1}^{m-1}(1 - 2\lambda\Delta t_i) \leq e^{-2\lambda t_{m-1}}$. Two last inequalities lead to,

$$\frac{1}{2}\underline{\sigma}^2 \sum_{m=1}^{n} \Delta t_m \Big(\prod_{i=1}^{m-1}(1 - 2\lambda\Delta t_i)\Big)|v^m|_V^2 \leq 8\Big(\frac{1}{2}\underline{\sigma}^2 \int_0^{t_n} e^{-2\lambda\tau}|v_{\Delta t}|_V^2(\tau)d\tau\Big).$$

Moreover,

$$\Big(\prod_{i=1}^{n}(1 - 2\lambda\Delta t_i)\Big)||v^n||^2 \leq e^{-2\lambda t_n}|v^n|_V^2.$$

Hence, the upper estimation is obtained by combining two last expressions.

In the case of the lower estimate, again starting with the first equation of the proof, and applying mathematical simplifications result in,

$$\frac{e^{2\lambda t_{m-1}}}{\Delta t}\int_{t_{m-1}}^{t_m} e^{-2\lambda\tau}|v_{\Delta t}|_V^2(\tau)d\tau \leq$$
$$|v^m|_V^2 \int_0^1 e^{-2\lambda\Delta t_m \tau}\tau d\tau + |v^{m-1}|_V^2 \int_0^1 e^{-2\lambda\Delta t_m \tau}(1-\tau)d\tau \leq \frac{1}{2}\Big(|v^{m-1}|_V^2 + |v^m|_V^2\Big).$$

Taking sum and further simplification leads to,

$$\int_{t_{m-1}}^{t_m} e^{-2\lambda\tau}|v_{\Delta t}|_V^2(\tau)d\tau \leq \frac{1}{2}\sum_{m=1}^{n} \Delta t_m e^{-2\lambda t_{m-1}}\Big(|v^{m-1}|_V^2 + |v^m|_V^2\Big).$$

Based on the fact that there exists a constant $\alpha_2 < \frac{1}{2}$, that if $\lambda\Delta t \leq \alpha_2$ then,

$$e^{-2\lambda t_{m-1}} \leq 2\prod_{i=1}^{m-1}(1 - 2\lambda\Delta t_i)$$

one obtains,

$$\int_0^{t_n} e^{-2\lambda\tau}|v_{\Delta t}|_V^2(\tau)d\tau \leq \sum_{m=1}^{n} \Delta t_m \prod_{i=1}^{m-1}(1 - 2\lambda\Delta t_i)\Big(|v^{m-1}|_V^2 + |v^m|_V^2\Big).$$

Applying the definition of $\rho_{\Delta t}$ in 6.13, results in,

$$\int_0^{t_n} e^{-2\lambda\tau}|v_{\Delta t}|_V^2(\tau)d\tau \leq (1 + \rho_{\Delta t})\sum_{m=1}^{n} \Delta t_m \prod_{i=1}^{m-1}(1 - 2\lambda\Delta t_i)|v^m|_V^2 + \Delta t_1|v^0|_V^2.$$

Furthermore, recall the assumption $\lambda\Delta t \leq \alpha_2$. Thus, $e^{-2\lambda t_n} \leq 2\Big(\prod_{i=1}^{n}(1-2\lambda\Delta t_i)\Big)$, which leads to inequality,

$$||v^n||^2 e^{-2\lambda t_n} \leq 2||v^n||^2\Big(\prod_{i=1}^{n}(1-2\lambda\Delta t_i)\Big)$$

choosing α as a minimum of two constants α_1 and α_2 terminates the proof. □

Combining the 6.19 and 6.20 yields,

$$\forall n,\ 1 \leq n \leq N, \quad [[u_{\Delta t}]](t_n) \leq \Big(\max(2, 1+\rho_{\delta t})||u_0||^2 + \frac{1}{2}\underline{\sigma}^2\Delta t_1|u_0|_V^2\Big)^{\frac{1}{2}}. \tag{6.21}$$

6.1.2 Fully Discrete Problem

In order to illustrate the full discretization of the Black-Scholes equation for the European put options, consider the $\mathcal{T}_{nh}$ as a family of grids of Ω, for all n, $0 \leq n \leq N$. Then, the diameter of the element $\omega \in \mathcal{T}_{nh}$ is denoted by h_ω. Moreover, minimum and maximum price of each element ω is shown by $S_{min}(\omega)$ and $S_{max}(\omega)$ respectively.

We also need to make the assumption that, there exists a constant ρ_h that satisfies,

$$h_\omega \leq \rho_h h_{\omega'} \tag{6.22}$$

where ω and ω' are adjacent elements of $\mathcal{T}_{nh}$.

Then, for every h, we define the discrete spaces,

$$V_{nh} = \{v_h \in V, \forall\omega \in \mathcal{T}_{nh}, v_{h|_\omega} \in \mathcal{P}_1\}, \quad V_{nh}^0 = v_{nh} \cap V^0. \tag{6.23}$$

It remains to derive a fully discrete problem. Suppose that $u_0 \in V_{0h}$. Thus,

Find $(u_h^n)_{0\leq n\leq N}$, $u_h^n \in V_{nh}^0$, such that,

$$\forall n,\ 1 \leq n \leq N,\ \forall v_h \in V_{nh}^0, \quad \Big(\frac{u_h^n - u_h^{n-1}}{\Delta t_n}, v_h\Big) + \frac{1}{2}\Big(a_{t_n}(u_h^n, v_h) + a_{t_{n-1}}(u_h^{n-1}, v_h)\Big) = 0. \tag{6.24}$$

Hence, stability estimate leads to,

$$[[u_h^m]]_n \leq ||u^0||. \tag{6.25}$$

Note that, $u_{h,\Delta t}$ is denotes the function that is affine on each time interval and satisfies $u_{h,\Delta t} = u_h^n$

6.2 Error Indicators

In this section, we are going to derive error indicators and proving their efficiency and reliability. As it is discussed earlier, we apply two families of error indicators, the first one, which is local in t and global in S, is denoted by η_n, and the second one, which is local in both t and S, is denoted by $\eta_{n\omega}$. They can be computed explicitly, and act as an estimation for the error in each time interval (η_n) and each element ($\eta_{n\omega}$)

Besides deriving the formula for η_n and $\eta_{n\omega}$, we would like to show that these error indicators are reliable, which means that they are the upper bound for the error (subsection

6.2.1). Our next goal is to prove their efficiency, which means, the error is the upper bound for them (subsection 6.2.2).

By doing this, we obtain the formulas that estimate the error explicitly, in a reliable and efficient way. Then, we can apply these formulas to refine the mesh adaptively.

It is also worth noting that due to the fact that the proofs of this section are quite long and to make following the procedure of deriving error estimators easier, proofs are given in the appendix.

6.2.1 An Upper Bound For the Error

As it is illustrated earlier, our objective in this subsection is to bound the error $[[u - u_{h,\Delta t}]](t_n)$, by error indicators. For doing this, using the triangle inequality, the following inequality is applied,

$$[[u - u_{h,\Delta t}]](t_n) \le [[u - u_{\Delta t}]](t_n) + [[u_{\Delta t} - u_{h,\Delta t}]](t_n).$$

However, before evaluating these estimates, we require to make further assumptions.

Assume that σ and $S\frac{\partial \sigma}{\partial S}$ are Lipschitz continuous with regard to S and t. In addition, suppose that the interest rate r is Lipschitz continuous with respect to t.

Therefore, there exist constants $L_1,\ L_2$ and L_3 that for all $t, t' \in [0,T]$ satisfy,

$$\begin{aligned}
&||\sigma^2(.,t) - \sigma^2(.,t')||_{L^\infty(0,\bar{S})} \le L_1|t' - t|,\\
&||-r(t) + r(t') + \frac{\sigma^2(.,t) - \sigma^2(.,t')}{2} + S\Big(\sigma(.,t)\frac{\partial \sigma}{\partial S}(.,t) - \sigma(.,t')\frac{\partial \sigma}{\partial S}(.,t')\Big)||_{L^\infty(0,\bar{S})}\\
&\le L_2|t' - t|,\\
&|r(t) - r(t') \le L_3|t' - t|.
\end{aligned}$$

Now, we derive two proposition for evaluating mentioned estimates, then we express the theorem at the end of this subsection to illustrate the final result.

Proposition 1. Assume that $u_0 \in V_{1h}$, and there exists a constant $\alpha < \frac{1}{2}$, which under the condition that $\lambda \Delta t \le \alpha$, it holds,

$$[[u - u_{\Delta t}]](t_n) \le c\Big(\frac{L}{\underline{\sigma}^2}c(u_0)\Delta t + \frac{\mu}{\underline{\sigma}^2}(1 + \rho_{\Delta t})[[u_{\Delta t} - u_{h,\Delta t}]](t_n) + \frac{\mu}{\underline{\sigma}^2}(\sum_{m=1}^{n} \eta_m^2)^{\frac{1}{2}}\Big) \tag{6.26}$$

where,

$$\eta_m^2 = \Delta t_m e^{-2\lambda t_{m-1}} \frac{\underline{\sigma}^2}{2} |u_h^m - u_h^{m-1}|_V^2. \tag{6.27}$$

Also, $L = 4L_1 + 2L_2 + L_3$, $c(u_0) = \Big(\max(2, 1+\rho_{\delta t})||u_0||^2 + \frac{1}{2}\underline{\sigma}^2 \Delta t_1 |u_0|_V^2\Big)^{\frac{1}{2}}$, and c is a positive constant.

Moreover, the second estimate is given by the following proposition,

Proposition 2. Assume that $u_0 \in V_{1h}$, then there exists a constant c, that for each $t_n, (1 \le n \le N)$ satisfies,

$$[[u_{\Delta t} - u_{h,\Delta t}]]^2(t_n) \le \frac{c}{\underline{\sigma}^2} \max(2, 1 + \rho_{\Delta t}) \sum_{m=1}^{n} \Delta t_m \prod_{i=1}^{m-1} (1 - 2\lambda \Delta t_i) \sum_{\omega \in \mathcal{T}_{mh}} \eta_{m,\omega}^2 \tag{6.28}$$

where,

$$\eta_{m,\omega} = \frac{h_\omega}{S_{max}(\omega)} \Big|\Big| \frac{u_h^m - u_h^{m-1}}{\Delta t_m} - rS\frac{\partial u_h^m}{\partial S} + ru_h^m \Big|\Big|_{L^2(\omega)}. \tag{6.29}$$

Now, it is time to apply these two propositions to derive the theorem for a posteriori error estimate.

Theorem 6.2.1. Assume that $u_0 \in V_{1h}$ and $\lambda\Delta t \leq \alpha$, then there exists a constant c, that for each $t_n, (1 \leq n \leq N)$ satisfies,

$$[[u - u_{h,\Delta t}]](t_n) \leq c\Big(\frac{L}{\underline{\sigma}^2}c(u_0)\Delta t + \frac{\mu}{\underline{\sigma}^2}\Big(\sum_{m=1}^{n}\eta_m^2 + \frac{\Delta t_m}{\underline{\sigma}^2}(1+\rho_{\Delta t})^2\max(2, 1+\rho_{\Delta t}) \prod_{i=1}^{m-1}(1-2\lambda\Delta t_i)\sum_{\omega\in\mathcal{T}_{mh}}\eta_{m,\omega}^2\Big)^{\frac{1}{2}}\Big) \tag{6.30}$$

where, L and $c(u_0)$ are as given before, and η_m and $\eta_{m,\omega}$ are evaluated according to 6.27 and 6.29, respectively.

Therefore, we bound the error by error indicators η_m and $\eta_{m,\omega}$. The next step is to show that these error indicators are bounded by the error.

6.2.2 An Upper Bound For the Error Indicators

As illustrated before, the objective of this subsection is to derive two propositions, in order to bound η_m and $\eta_{m,\omega}$ separately.

Meanwhile, for $(v_n)_{1\leq n\leq N},\ v_n \in V_0$, we require to introduce the notation $[[v^n]]$ as,

$$[[v^n]]^2 = \frac{\underline{\sigma}^2}{2}\Delta t_n \prod_{i=1}^{n-1}(1-2\lambda\Delta t_i)|v^n|_V^2.$$

Now, we can state the proposition to bound η_m,

Proposition 3. Assume that $u_0 \in V_{1h}$ and $\lambda\Delta t \leq \alpha$, then there exists a constant c, such that for $2 \leq n \leq N$,

$$\eta_n \leq c\Bigg([[u^n - u_h^n]] + \sqrt{\rho_{\Delta t}}[[u^{n-1} - u_h^{n-1}]] + \frac{e^{-\lambda t_{n-1}}}{\underline{\sigma}}\Big(\Big|\Big|\frac{\partial(u-u_{\Delta t})}{\partial t}\Big|\Big|_{L^2(t_{n-1},t_n;V_0')} + ||u - u_{\Delta t}||_{L^2(t_{n-1},t_n;V_0)}\Big) + \Big(\frac{L}{\underline{\sigma}^2}(\max(1,\rho_{\Delta t}))^{\frac{1}{2}} + \frac{\lambda\mu}{\underline{\sigma}^2}\Big)\Delta t_n||u^0||\Bigg) \tag{6.31}$$

also,

$$\eta_1 \leq c\Bigg([[u^1 - u_h^1]] + \frac{1}{\underline{\sigma}}\Big(\Big|\Big|\frac{\partial(u-u_{\Delta t})}{\partial t}\Big|\Big|_{L^2(t_0,t_1;V_0')} + ||u - u_{\Delta t}||_{L^2(0,t_1;V_0)}\Big) + \frac{L+\lambda\mu}{\underline{\sigma}^2}\Delta t_1||u^0|| + \frac{L}{\underline{\sigma}}(\Delta t_1)^{\frac{3}{2}}|u^0|_V\Bigg). \tag{6.32}$$

Nevertheless, before deriving the corresponding expression for $\eta_{m,\omega}$, we require to make further assumptions.

Assume that, K_ω contains all elements in $\mathcal{T}_{nh}$ that share a node with ω. Then, we define,

$$V(K_\omega) = \{v \in L^2(K_\omega);\ S\frac{\partial v}{\partial S} \in L^2(K_\omega)\}$$

with the norm $||v||_{V(K_\omega)} = \Big(\int_{K_\omega} v^2(S) + S^2(\frac{\partial v}{\partial S}(S))^2\Big)^{\frac{1}{2}}$. In addition, assume that $V_0(K_\omega)$ is the closure of $\mathcal{D}(K_\omega)$ in $V(K_\omega)$, with the norm defined by $||v||_{V_0(K_\omega)} = \Big(\int_{K_\omega} S^2(\frac{\partial v}{\partial S}(S))^2\Big)^{\frac{1}{2}}$. Finally, we denote the dual space of $V_0(K_\omega)$, by $V_0'(K_\omega)$ with the dual norm.

Now, we can derive the next proposition.

Proposition 4. For each $\omega \in \mathcal{T}_{nh}$ where, $1 \le n \le N$, it holds,

$$\eta_{n,\omega} \le c\Bigg(\Big|\Big|\frac{u^{n-1} - u_h^{n-1} - (u^n - u_h^n)}{\Delta t_n}\Big|\Big|_{V_0'(K_\omega)} + \mu\Big|\Big|S\frac{\partial(u^n - u_h^n)}{\partial S}\Big|\Big|_{L^2(K_\omega)}\Bigg). \tag{6.33}$$

To conclude, by deriving a posteriori error analysis, we obtained the formula for error indicators η_m (local in t and global in S) and $\eta_{m,\omega}$ (local in t and S), and proved that the error caused by applying finite element method is bounded from below and above by these error indicators.

It remains to apply η_m and $\eta_{m,\omega}$, so as to refine the mesh adaptively. And this is the objective of the next section.

6.3 Refinement Strategy

In this section, we are going to derive and implement the algorithm, which uses η_m and $\eta_{m,\omega}$, defined by 6.27 and 6.29, respectively.

The algorithm begins with applying the uniform finite element method described in chapter 5, to the Black-Scholes partial differential equation with local volatility for European put options. Then, use 6.27 and 6.29 to compute error indicators η_m and $\eta_{m,\omega}$, respectively. Then, according to Theorem 6.2.1, the global error indicator of S-discretization in the time interval $[t_{m-1}, t_m]$ can be computed by $\frac{\Delta t_m}{\underline{\sigma}^2}\sum_{\omega\in\mathcal{T}_{nh}} \eta_{m,\omega}^2$. The corresponding expression for t-discretization in the time interval $[t_{m-1}, t_m]$, is given by η_m^2. Using these two estimates, we can also compute the global error.

If the global error is less than the given tolerance δ, then it shows that the algorithm obtains the solution with desired precision, thus it should terminate. Otherwise, we need to perform more refinements.

In order to decide which time interval to refine, form the expressions,

$$\bar{\zeta} = \max_m \eta_m \quad \text{and} \quad \underline{\zeta} = \min_m \eta_m.$$

Then, for each m, that satisfies $\eta_m > \frac{\bar{\zeta}+\underline{\zeta}}{2}$, refinement should be done. On the other hand, after determining candidates (time intervals) for refinement, we have to decide whether to refine in t or S. And for this purpose, we compare,

$$\eta_m^2 \quad \text{and} \quad \frac{\Delta t_m}{\underline{\sigma}^2}\sum_{\omega\in\mathcal{T}_{nh}} \eta_{m,\omega}^2.$$

If η_m^2 is greater or equal, it shows that the error in t is higher than S, so the refinement should be performed in t, therefore intersect the time interval $[t_{m-1}, t_m]$. Otherwise, the refinement needs to be done in the S variable. Hence, compute the following expressions,

$$\bar{\xi} = \max_{\omega} \eta_{m,\omega} \quad \text{and} \quad \underline{\xi} = \min_{\omega} \eta_{m,\omega}$$

and intersect those price intervals which satisfy $\eta_{m,\omega} > \frac{\bar{\xi}+\underline{\xi}}{2}$. It is worth noting that refining in S leads to producing hanging nodes, which is solved by applying more restrictions to the problem. Due to the fact that we applied the conforming finite element method, the constraints should simply be chosen in a way to ensure that the discrete solution space remains a subspace of the continuous space.[BKSY17]

The next step is to update variables and error indicators after refinement

In other words, the algorithm is as follows,

Algorithm 2. (Adaptive mesh refinement)

1. Calculate the discrete solution in uniform mesh

2. Compute error indicators η_m and $\eta_{m,\omega}$

3. For each time interval $[t_{m-1}, t_m]$, compute

$$\eta_m^2 \quad \text{and} \quad \frac{\Delta t_m}{\underline{\sigma}^2} \sum_{\omega \in \mathcal{T}_{nh}} \eta_{m,\omega}^2$$

4. If global error is less than the tolerance, then algorithm terminates. Otherwise, compute

$$\bar{\zeta} = \max_{m} \eta_m \quad \text{and} \quad \underline{\zeta} = \min_{m} \eta_m.$$

5. For intervals (Δt_m), that $\eta_m > \frac{\bar{\zeta}+\underline{\zeta}}{2}$, compare

$$\eta_m^2 \quad \text{and} \quad \frac{\Delta t_m}{\underline{\sigma}^2} \sum_{\omega \in \mathcal{T}_{nh}} \eta_{m,\omega}^2.$$

   If η_m^2 is greater or equal, then intersect the interval $[t_{n-1}, t_n]$. Otherwise,

6. Compute

$$\bar{\xi} = \max_{\omega} \eta_{m,\omega} \quad \text{and} \quad \underline{\xi} = \min_{\omega} \eta_{m,\omega}.$$

   And, intersect elements in the interval Δt_m that satisfy, $\eta_{m,\omega} > \frac{\bar{\xi}+\underline{\xi}}{2}$.

7. Update variables and error indicators, and start from step 3.

In the case of numerical experiments, figures 6.1 shows the value of the error estimator η_m for the put option with $K = 0$, $r = 0.05$, $\sigma = 0.5$, the domain $[0, 100] \times [0, 1]$, and 50×50 nodes.

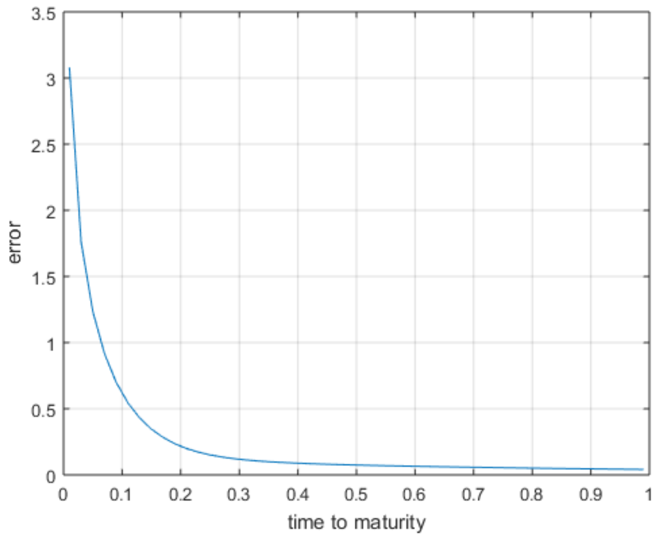

Figure 6.1: Error indicator η_m

The figure for the error estimator $\eta_{m\omega}$ with the same data is given in 6.2 . According to these figures, which totally match the actual error that we obtained in the chapter 5, the error is higher near the strike and the maturity.

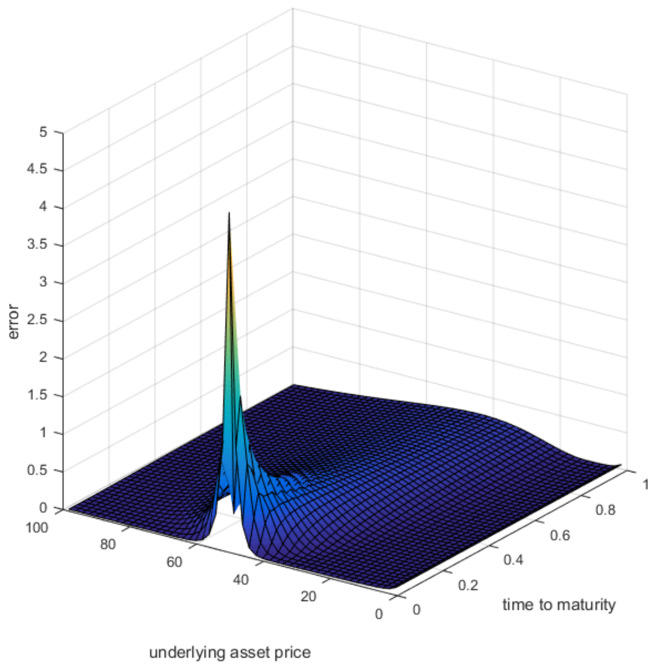

Figure 6.2: Error indicator $\eta_{m,\omega}$

Moreover, we implement this algorithm to refine the mesh adaptively. According to the shapes of the error estimators and also the results of chapter 5, we expect the mesh to be refined around maturity and strike more than other areas in the domain. Figures 6.3, 6.4, 6.5, and 6.6 show the obtained mesh grid after 0, 20, 50, and 100 refinement respectively, which fulfill our expectations.

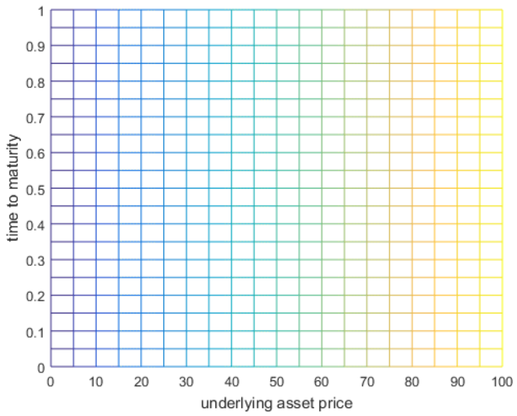

Figure 6.3: Uniform mesh

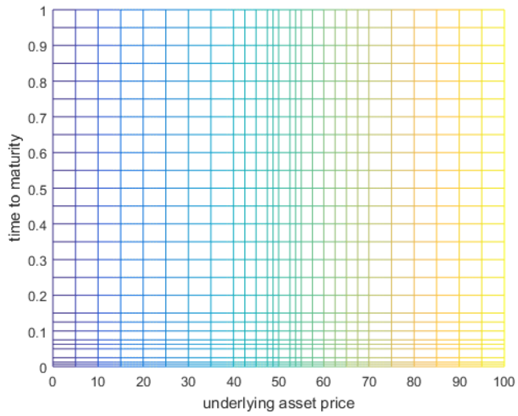

Figure 6.4: Mesh after 20 refinement

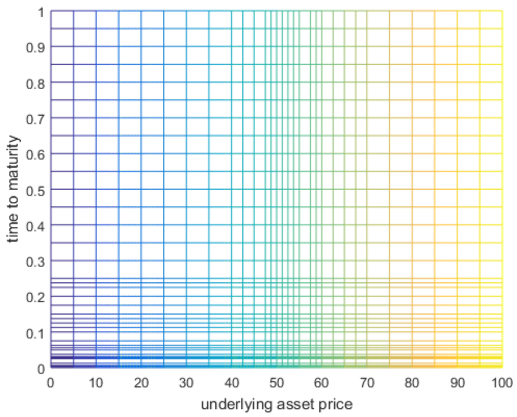

Figure 6.5: Mesh after 50 refinement

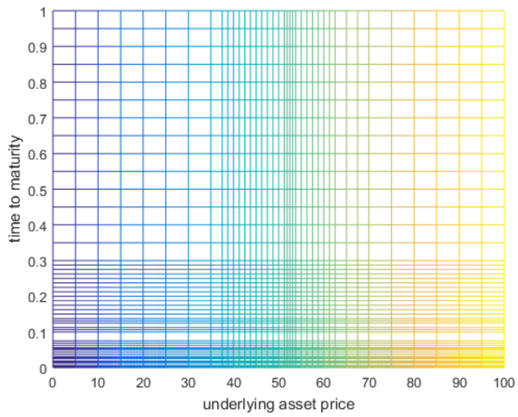

Figure 6.6: Mesh after 100 refinement

In addition, we perform the algorithm for the put option with $\sigma = 0.5$, $r = 0.05$, $K = 50$, $T = 1$, on the domain $[0, 200] \times [0, 1]$, and with 80×20 nodes. Then, we evaluate the t-discretization and S-discretization error of each time step. According to our results, the maximum error of these time steps reduce to around 10^{-2} after 50 refinement.

It also should be noted that similar to other cases in this thesis, we use MATLAB [Mat14], for performing the algorithm and obtaining the results.

7 Calibration of Local Volatility

After regarding the Black-Scholes equation in mathematical and numerical points of view in previous chapters, our objective in this chapter is to discuss some limitations of this model in real life applications. Then in the second section, we aim to follow the method mentioned in the paper [HR05], in order to tackle some of these shortcomings. In the last section of this chapter, we apply European call options available on the market, to examine this method.

7.1 Shortcomings of Black-Scholes Equation

In this section, we follow the paper [Ten11], in order to illustrate some shortcomings of Black-Scholes equation.

While the Black-Scholes equation is one of the greatest achievements in financial theory and is widely used by traders, there are also some critics that challenge the assumptions of the Black-Scholes equation.

Illustration of some challenging assumptions of Black-Scholes equation is given below,

- **Volatility is constant over time:** According to volatility clustering property, "large changes tend to be followed by large changes and small changes tend to be followed by small changes" [Man67]. Moreover, it is shown that the volatility is correlated with asset price returns and trading volumes [Fre83]. Therefore, volatility cannot be constant over time, particularly, in long-term.

- **Stock price follows the random walk:** Based on this assumption, the price of the underlying asset increases and decreases with the same probability at each time. However, due to the fact that asset prices are obtained by different economic factors, their movements do not affected by the same probability in different times and prices.

- **Returns are log-normally distributed:** In contrast to stable distributions such as log-normal, according to [Cla73], asset returns usually have a semi-heavy tails with finite variance in real markets.

- **Markets are perfectly liquid:** Based on the fact that traders have financial limitations to invest, and this point that the fraction of options cannot be traded, this assumption cannot be realistic too.

- **Interest rate is constant:** On the one hand, there are no risk-free rates in real world. On the other hand, alternative rates change over time.

- **The underlying asset does not pay dividends:** In real markets, many companies pay some money to their shareholders, as a dividend.

- **There are no transaction costs:** Again, this is not a realistic assumption. Because stock brokers charge some money for trading options.

While some of these limitations can be solved by expanding the original model. Take for example the Black-Scholes with local volatility model (that was considered in this thesis), which assumes that the volatility is a function of time and underlying asset price, or including dividends in the equation. Nevertheless, there are some fundamental issues with Black-Scholes equation, which can not be addressed easily, and new models with different assumptions should be derived, so as to capture the real world features of option prices.

One important remedy for above-mentioned problems of Black-Scholes equation is to determine local volatilities, such that the obtained option values fit the real data of the market. And this method is discussed in the next section.

7.2 Calibration of Local Volatility Function

As it is mentioned before, we would like to follow the method which is proposed by Hanke and Rösler [HR05], to calibrate the volatility function from the observed underlying option prices. To this end, this method applies the Dupire equation, then in order to differentiate the data numerically, cubic splines are used. Finally, the problem is solved with Tikhonov regularization.

Consider the Black-Scholes equation, if the volatility is explicitly known, then the problem is well-posed. However, this is not the case in practice. Therefore, we need to regard an inverse problem to approximate the volatility function $\sigma(S,t)$ from observed option values at the fixed time t_0 and fixed price S_0 of the underlying asset. Nevertheless, as described by Dupire in [D$^+$94], there is a more straightforward alternative equation, that is discussed in the following subsection.

7.2.1 Dupire Equation

In this subsection, we follow [Jia05], to illustrate how the Dupire equation is derived.

Assume that the call option value $V(S,t;\sigma,K,T)$ satisfies the Black-Scholes equation at the fixed price S_0 and fixed time t_0,

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2(S,t)S^2\frac{\partial^2 V}{\partial S^2} + (r-q)S\frac{\partial V}{\partial S} - rV = 0 \tag{7.1}$$

$$V(S,T) = \max(0, S-K) \tag{7.2}$$

where $0 < S < \infty$ and $0 \leq t < T$ and q is the dividend which pays to the shareholders. Also, define G as the second derivative of V with respect to strike, i.e. $\frac{\partial^2 V}{\partial K^2} = G(S,t;K,T)$. Then G satisfies,

$$\frac{\partial G}{\partial t} + \frac{1}{2}\sigma^2(S,t)S^2\frac{\partial^2 G}{\partial S^2} + (r-q)S\frac{\partial G}{\partial S} - rG = 0 \tag{7.3}$$

$$G(S,t) = \delta(K-S) \tag{7.4}$$

where δ is a Dirac function.
As it is proved in [Jia05] in detail, regarding the fact that $\delta(x) = \delta(-x)$, then G is the fundamental solution of the 7.3, 7.4. Hence, it is also the fundamental solution to the adjoint problem of 7.3, 7.4. Thus,

$$-\frac{\partial G}{\partial T} + \frac{1}{2}\frac{\partial^2}{\partial K^2}(\sigma^2(K,T)K^2 G) - (r-q)\frac{\partial (KG)}{\partial K} - rG = 0 \tag{7.5}$$

$$G(S,t;K,T) = \delta(K-S) \tag{7.6}$$

where $0 \le K \le \infty$ and $t \le T$.
Then, substituting $\frac{\partial^2 V}{\partial K^2}$ instead of G, and integrating two times with respect to K in $[K,\infty]$, results in the Dupire equation for European call options,

$$-\frac{\partial V}{\partial T} + \frac{1}{2}\sigma^2(K,T)K^2\frac{\partial^2 V}{\partial K^2} - (r-q)K\frac{\partial V}{\partial K} - qV = 0. \tag{7.7}$$

Notice that unlike the Black-Scholes equation, that the option value is the function of underlying asset price S and time t, in Dupire equation, the option value is the function of strike K and maturity T. Consequently, by solving this equation just for one time, we evaluate the option value for all strikes and maturities, whereas, in the case of Black-Scholes equation, we have to evaluate the option value for each strike and maturity separately.

Finally, solving the equation for $\sigma(K,T)$, one obtains,

$$\sigma(K,T) = \left(\frac{2\left(\frac{\partial V}{\partial T} + (r-q)K\frac{\partial V}{\partial K} + qV\right)}{K^2\frac{\partial^2 V}{\partial K^2}}\right)^{\frac{1}{2}} \tag{7.8}$$

where all components of the right-hand side can be evaluated at $(t_0, S_0; T, K)$ from the given data in theory. However, in practice, there are some issues that cause this problem to be ill-posed. First of all, in real markets, there are just a few maturities, and in each maturity only finite number of strikes are available for traders. Secondly, the real call option values are usually rising in K and declining in T. Thus, while the positivity of the denominator is secured, numerator can be negative, and this can lead to the negativity of the whole fraction, which is clearly banned, due to the square root. Besides that, it might lead to severe oscillation of the local volatility function.

In fact, our objective for the rest of this chapter is to improve the stability of this approach by using the techniques described in [HS98], which consists of applying the smooth cubic spline interpolation for numerical differentiation of discrete data. And, using Tikhonov regularization to solve equation 7.8.

7.2.2 Differentiation of the Discrete Data

According to Dupire equation, we require to differentiate the discrete data two times with respect to strike and one time with respect to maturity. Therefore, in order to perform numerical differentiation of the discrete data with respect to strike K, we apply the Reinsch algorithm to derive smooth cubic spline. Then, in the case of taking derivative with respect to maturity T, since there is a data only for a few number of maturities and also the fact that typically there is a large gap among these maturities, numerical differentiation of the data may not be accurate. On the other hand, it also implies that the stability is not an important problem, thus we simply apply difference quotient.

Cubic Spline

In this part, we consider [GS93], to illustrate cubic spline, and derive Reinsch algorithm for computing cubic splines. First of all, we need some definitions,

Definition 7.2.1. Cubic spline: Assume that the following data is given,

$$a < K_1 < K_2 < ... < K_n < b$$

then the function g is cubic spline if g is a cubic polynomial in each interval

$$(a, K_1), (K_1, K_2), ..., (K_n, b)$$

and pieces of the polynomial fit at knots K_i, such that g and its first and second derivatives are continuous at each K_i.

Definition 7.2.2. Natural cubic spline: A cubic spline that its second and third derivatives at the endpoints are zero.

When g is a natural cubic spline, we use the notation,

$$g(i) = g(K_i) \quad \text{and} \quad \gamma_i = g''(K_i), \text{ for } i = 1, .., n.$$

Notice that for a natural cubic spline, according to its definition, $\gamma_1 = \gamma_n = 0$. Also, let $\mathbf{g} = (g_1, .., g_n)^T$ and $\gamma = (\gamma_2, .., \gamma_{n-1})^T$. For further analysis of the natural cubic spline, we require to introduce the matrices Q and R.
If Q is a $n \times (n-2)$ matrix, then its entries q_{ij}, where $i = 1, ..., n-1$ and $j = 2, ..., n-1$, are given by,

$$q_{j-1,j} = \frac{1}{h_{j-1}}, \; q_{j,j} = -\frac{1}{h_{j-1}}, \; q_{j+1,j} = \frac{1}{h_j} \quad \text{for} \quad j = 2, ..., n-1$$
$$q_{i,j} = 0 \quad \text{if} \quad |i-j| \geq 2$$

where $h_i = t_{i+1} - t_i$. Also, the $(n-2) \times (n-2)$ symmetric matrix R with entries r_{ij}, where $i, j = 2, ..., n-1$, is given by,

$$r_{i,i} = \frac{h_{i-1} + h_i}{3}, \; r_{i,i+1} = r_{i+1,i} = \frac{h_i}{6}, \quad \text{for} \quad i = 2, ..., n-1$$
$$r_{i,j} = 0 \quad \text{if} \quad |i-j| \geq 2.$$

We also define the matrix $S = QR^{-1}Q^T$. Then, we have the following theorem,

Theorem 7.2.1. Vectors $\mathbf{g}$ and γ satisfy $Q^T\mathbf{g} = R\gamma$, if and only if, g is a natural cubic spline. If this is the case, then $\int_a^b g''(t)^2 dt = \gamma^T R \gamma = \mathbf{g}^T S \mathbf{g}$

Proof. Proof is given in [GS93]. □

Theorem 7.2.2. For $K_1 < ... < K_n$ where $n \geq 2$, and values $z_1, ..., z_n$, there exists a unique natural cubic spline g with knots K_i, such that for $(i = 1, ..., n)$,

$$g(K_i) = z_i.$$

Proof. Applying Theorem 7.2.1, and the fact that the matrix R is positive definite, leads to uniqueness of γ, which is the desired condition. □

Furthermore, assume that $Y_1, ..., Y_n$ are the observations. Then, our objective is to calculate and analyze properties of the curve $\hat{g}$, that minimizes,

$$\sum_{i=1}^{n}\{Y_i - g(K_i)\}^2 + \alpha \int_a^b \{g''(x)\}^2 dx \tag{7.9}$$

where α is a smoothing parameter, and $g \in C^1([a,b])$.

It is shown in the following theorem that $\hat{g}$ have to be a natural cubic spline. Also, a formula for calculating $\hat{g}$ is presented.

Theorem 7.2.3. Assume that $n \geq 3$ and $a < K_1 < K_2 < ... < K_n < b$. Also, suppose that $Y_1, ..., Y_n$ and $\alpha > 0$ are given. If $\hat{g}$ is a natural cubic spline with knots $K_1, K_2 < ..., K_n$, such that

$$\mathbf{g} = (I + \alpha K)^{-1}\mathbf{Y} \tag{7.10}$$

where $\mathbf{Y} = (Y1, ..., Y_n)^T$. Then, for each $g \in C^1([a,b])$,

$$\sum_{i=1}^{n}\{Yi - \hat{g}(K_i)\}^2 + \alpha \int_a^b \{\hat{g}''(x)\}^2 dx \leq \sum_{i=1}^{n}\{Y_i - g(K_i)\}^2 + \alpha \int_a^b \{g''(x)\}^2 dx \tag{7.11}$$

where the equality holds, only if $\hat{g} = g$.

Proof. First of all, we show that the minimizer of the 7.9, have to be a natural cubic spline. For doing this, assume that g is any cubic spline with knots Y_i, $i = 1, ..., n$, and $\bar{g}$ is a natural cubic spline with the same knots. Then, due to the fact that for each i, we have $\bar{g}(Y_i) = g(Y_i)$, therefore $\sum\{Y_i - \bar{g}(K_i)\}^2 = \sum\{Y_i - g(K_i)\}^2$. Considering the optimality property of natural cubic spline leads to $\int \bar{g}''^2 < \int g''^2$. Finally, since α assumed to be a positive constant, we obtain the desired result.

Now for the second part of the proof, assume that g is a natural cubic spline with vectors $\mathbf{g}$ and γ, and matrices Q and R (as defined earlier). Our objective is to express 7.9, in terms of above matrices and vectors. Then not only we can prove the existence and uniqueness of the minimizer $\hat{g}$, but also it provides us with the algorithm to calculate this minimizer. Therefore, because $\mathbf{g} = (g_1, ..., g_n)^T$, then we have $\sum\{Y_i - g(K_i)\}^2 = (\mathbf{Y} - \mathbf{g})^T(\mathbf{Y} - \mathbf{g})$. In the case of the penalty term, according to theorem 7.2.1, $\int_a^b g''(t)^2 dt = \mathbf{g}^T S \mathbf{g}$. As a consequence, we can re-express the sum 7.9 as, $(\mathbf{Y} - \mathbf{g})^T(\mathbf{Y} - \mathbf{g}) + \mathbf{g}^T S \mathbf{g}$, and this is equal to,

$$\mathbf{g}^T(I + \alpha S)\mathbf{g} - 2\mathbf{Y}^T\mathbf{g} + \mathbf{Y}^T\mathbf{Y}. \tag{7.12}$$

Notice that the matrix $I + \alpha S$ is positive definite (because αS is positive semi-definite), which cause 7.12, to have a unique minimum. This minimum can be obtained by $\mathbf{g} = (I + \alpha S)^{-1}\mathbf{Y}$. □

Combining results of theorem 7.2.2 and 7.2.3, we obtain that the vector $\mathbf{g}$ defines a unique natural cubic spline g, and among all natural cubic splines with knots $K_i, (i = 1, ..., n)$, the minimum of 7.9 is given uniquely by $(I + \alpha S)^{-1}\mathbf{Y}$.

Therefore, given a set of knots with their corresponding values (observations), we can apply this formula to obtain a unique spline that fits our observations. However, it is not efficient to apply 7.10 directly and evaluate the vector $\mathbf{g}$ and thus the curve $\hat{g}$. Hence, we describe an algorithm, introduced by Christian Reinsch [Rei67], for finding the smoothing spline. The idea behind the Reinsch algorithm is to derive a non-singular linear system of equations for evaluating γ_i at knots K_i, then using explicit formulas to obtain g_i, from γ_i and Y_i.

Combining 7.10 and the definition $S = QR^{-1}Q^T$, results in,

$$(I + \alpha QR^{-1}Q^T)\mathbf{g} = \mathbf{Y}. \tag{7.13}$$

Thus,

$$\mathbf{g} = \mathbf{Y} - \alpha QR^{-1}Q^T\mathbf{g}.$$

Then, substituting $Q^T\mathbf{g} = R\gamma$, leads to,

$$\mathbf{g} = \mathbf{Y} - \alpha Q\gamma.$$

Now, multiply both sides of the equation by Q^T, and apply $Q^T\mathbf{g} = R\gamma$, to obtain, $Q^T\mathbf{Y} - \alpha Q^TQ\gamma = R\gamma$. Hence,

$$(R + \alpha Q^TQ)\gamma = Q^T\mathbf{Y}. \tag{7.14}$$

Equation 7.14 plays a key role in Reinsch algorithm. Its importance comes from the fact that the matrix $R+\alpha Q^TQ$ is a Band matrix (all non-zero elements are gathered in few diagonals) with 5 bandwidth (number of non-zero diagonals is 5), which reduces the computation costs of γ significantly. Whereas, it was not the case for the equation 7.13. Besides that, $R+\alpha Q^TQ$ is symmetric and positive definite. Therefore, there is a Cholesky decomposition,

$$R + \alpha Q^TQ = LDL^T$$

where L is the lower triangular matrix with $L_{ii} = 1$, $\forall i$, and D is a diagonal matrix with positive elements. This property makes the algorithm even cheaper. Now, we can express the Reinsch algorithm as follow,

Algorithm 7.1**. (Reinsch algorithm)**

1. Calculate the vector $Q^T\mathbf{Y}$
2. Find the Cholesky decomposition factors L and D, of the matrix $R + \alpha Q^TQ$
3. Form the system $LDL^T\gamma = Q^T\mathbf{Y}$, and calculate γ
4. Form $\mathbf{g} = \mathbf{Y} - \alpha Q\gamma$, and calculate $\mathbf{g}$.

Finally, we would like to derive the formula for g, based on the vectors $\mathbf{g}$ and γ, that obtained by performing the Reinsch algorithm.

For each interval $[K_i, K_{i+1}]$ where $i = 1, ..., n-1$, we define $h_i = K_{i+1} - K_i$. Then, as it is given in [GS93], the formula for g is,

$$\begin{aligned} g(K) = {} & \frac{(K - K_i)g_{i+1} + (K_{i+1} - K)g_i}{h_i} \\ & - \frac{1}{6}(K - K_i)(K_{i+1} - K)\left(\left(1 + \frac{K - K_i}{h_i}\right)\gamma_{i+1} + \left(1 + \frac{K_{i+1} - K}{h_i}\right)\gamma_i\right). \end{aligned} \tag{7.15}$$

In the case of $K \leq K_1$ and $K_n \leq K$, we have,

$$g(K) = g_1 - (K_1 - K)g'(K_1) \quad \text{for } K \leq K_1 \tag{7.16}$$

$$g(K) = g_n - (K - K_n)g'(K_n) \quad \text{for } K_n \leq K. \tag{7.17}$$

It only remains to consider the evaluation of the smoothing parameter α. For doing this, assume that the cubic spline, obtained for maturity T_j, where $j = 1, ..., m$ is denoted by u_j, with second derivative u_j''. Also, we show the observed data in fixed price S_0 at fixed time t_0, for maturity T_j and strike K_i, by $u(t_0, S_0; T_j, K_i)$. Recall that we have obtained the formula for computing cubic spline for each maturity, such that fits the observed data. In other words, for each maturity T_j, cubic spline u_j, minimize the following sum,

$$\sum_{i=1}^{n} |u_j(K_i) - u(t_0, S_0; T_j, K_i)|^2 + \alpha ||u_j''(K_i)||_{L^2}^2. \tag{7.18}$$

There are two fundamental strategy for determining α. The first one, which is described and analyzed in [HS01], is called "discrepancy principle", and consists of choosing the parameter α, in a way that the data uncertainty times the number of data samples equal to the sum 7.18. The second approach, which is used by Hanke and Rösler in [HR05], and is described in [EH96], is called "L-curve criterion". The idea is to evaluate the sum 7.18, at each maturity T_j for a considerable number of regularization parameter α. Then, compute the norms,

$$\rho(\alpha) = \sum_{i=1}^{n} |u_{j,\alpha}(K_i) - u(t_0, S_0; T_j, K_i)|^2, \quad \nu(\alpha) = ||u_{j,\alpha}''(K_i)||_{L^2}^2$$

and plot the curve $(\rho(\alpha), \nu(\alpha))$ in doubly logarithmic scale. This curve normally displays an L-shaped corner, where the regularization parameters relevant to the points just right of the corner are proved to be a proper choice.

Meanwhile, since g is a polynomial, we can easily compute its first and second derivative analytically. As a consequence, we have all the required ingredients for differentiating the discrete data numerically with respect to strike. The next step is to differentiate the available data with respect to maturity.

Central Difference Quotient

As it is mentioned earlier, we would like to apply the difference quotient to perform the differentiation with respect to maturity T. Based on the fact that the given maturities are typically not equidistant, we use the following centered difference quotient for inner maturities $T_2, ..., T_{m-1}$,

$$u_T(K_i, T_j) \approx \frac{1}{\tau_j + \tau_{j+1}} \left(\frac{\tau_j}{\tau_{j+1}} u_{j+1}(K_i) + \left(\frac{\tau_{j+1}}{\tau_j} - \frac{\tau_j}{\tau_{j+1}} \right) u_j(K_i) - \frac{\tau_{j+1}}{\tau_j} u_{j-1}(K_i) \right) \tag{7.19}$$

where $\tau_j = T_j - T_{j-1}$ for $j = 2, ..., m-1$.
In the case of maturities T_1 and T_m, we apply the forward and backward difference quotient respectively.

Heretofore, we have applied methods to compute u_K, u_{KK} and u_T. The next step is to use these values to solve equation 7.8 with two different approach, which will be discussed in the next subsection.

7.2.3 Solving For the Volatility Function

In order to evaluate the volatility function, we need further assumptions. First of all, if we define $\underline{\mathrm{K}} = \min_i K_i$ and $\bar{K} = \max_i K_i$, then for simplicity, assume that $\underline{\mathrm{K}}$ and $\bar{K}$ are the same for each maturity. Also, suppose that Δ is a rectangular equidistant grid with points, $\underline{\mathrm{K}} = K_1 < K_2 < ... < K_n = \bar{K}$, in x-direction, and $T_0 < T_1 < ... < T_m$ in y-direction, where $T_0 = t_0$. Then, we define,

$$z_{ij} = \sigma^2(T_i, K_j) \quad \text{where } (T_i, K_j) \in \Delta.$$

Now, if we denote the numerator and denominator of z by b and d, respectively. Then, according to equation 7.8, we have,

$$b = 2\Big(V_T + (r-q)KV_K + qV\Big) \qquad\qquad d = K^2 V_{KK}$$

where we set $b_{ij} = 0$ for $i = 0$. The next step is to apply standard lexicographic ordering, to convert the matrices z,b, and d, into one-dimensional vectors $\mathbf{z}, \mathbf{b}$ and $\mathbf{d}$. Then, we define the diagonal matrix D, by applying elements of vector $\mathbf{d}$. Therefore, our objective is to solve the system,

$$D\mathbf{z} \approx \mathbf{b}. \tag{7.20}$$

Nevertheless, due to the fact that we use natural cubic splines for the interpolation, the values of the matrix d disappear in $\underline{\mathrm{K}}$ and $\bar{K}$, therefore some elements of the matrix D are insignificant. Moreover, system 7.20 may have a negative solution, because while by definition D is always non-negative, the vector $\mathbf{b}$ can change sign. Hence, this system is ill-conditioned and some regularization is required.

Tikhonov Regularization

According to [GNC14], one of the widely used regularization methods to address ill-posed inverse problems is Tikhonov regularization. Its aim is to minimize the particular measure of the solution, such as size of the solution or norm of the first and second derivative. We apply Tikhonov regularization to minimize the following expression,

$$||D\mathbf{z} - \mathbf{b}||_2^2 + \lambda ||L\mathbf{z}||_2^2 \tag{7.21}$$

where λ is a positive regularization parameter, and L is an operator on $\mathbf{z}$.

An important issue is to define the operator L. The most obvious option is to choose identity matrix $L = I$, then it is called the zeroth order Tikhonov regularization. However, since the identity matrix can not secure non-negativity of the solution $\mathbf{z}$, it is not an appropriate choice for our problem. Because minimizing 7.21, is equivalent to solve the following linear system,

$$(D^2 + \alpha L^T L)\mathbf{z} = D\mathbf{z} \tag{7.22}$$

and choosing $L = I$, leads to a diagonal system matrix with positive entries, while the right-hand side, since $\mathbf{b}$ can change signs, may not be non-negative in general.

Most papers which address the calibration of local volatility surface, apply the first order Tikhonov regularization. It consists of defining the operator L that approximates the first derivative of the volatility function. Following the paper [HR05], we apply finite difference approximation, so as to define the discrete gradient operator L. Therefore, assume that the matrices $L_K \in \mathbb{R}^{n-1,n}$ and $L_T \in \mathbb{R}^{m,m+1}$, define by,

$$L_K = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix}, \quad L_T = \begin{bmatrix} -\frac{1}{\tau_1} & \frac{1}{\tau_1} & & & \\ & -\frac{1}{\tau_2} & \frac{1}{\tau_2} & & \\ & & \ddots & \ddots & \\ & & & -\frac{1}{\tau_m} & \frac{1}{\tau_m} \end{bmatrix}$$

Then, L is chosen as,

$$L = \begin{bmatrix} I_{m+1} \otimes L_K \\ L_T \otimes I_n \end{bmatrix} \tag{7.23}$$

where $\otimes$ denotes the Kronecker product of two matrices.

One of the benefits of this choice for the operator L is a coupling of the individual values of $\sigma^2(T_i, K_j)$ in z. Moreover, it can also be shown that with sufficiently large α, this choice of L, implies the positive solution $\mathbf{z}$. And this is proved in the following theorem.

Theorem 7.2.4. Consider the given definition of the vectors $\mathbf{b}$ and $\mathbf{d}$, and suppose that $\sum b_{ij} d_{ij} > 0$. If $\mathbf{z}_\alpha$ be the solution of the equation 7.22, and $\alpha \to \infty$, then we have,

$$\mathbf{z}_\alpha \to \frac{\sum b_{ij} d_{ij}}{\sum d_{ij}^2} \mathbf{1}$$

where $\mathbf{1}$ is the vector in $\mathbb{R}^{(m+1)n}$ with all entries equal to one.

Proof. According to generalized singular value decomposition [G+96], for matrices L and D, we have,

$$L = UCX^{-1}, \quad D = WSX \tag{7.24}$$

where all matrices belong to $\mathbb{R}(m+1)n \times (m+1)n$, and C and S are diagonal matrices with non-negative elements. U and W are orthogonal matrices, and X is non-singular matrix. Suppose that the matrices W and X, are consist of vectors $w_1, ..., w_{(m+1)n}$ and $x_1, ..., x_{(m+1)n}$, respectively. Likewise, assume that the diagonal elements of the matrices C and S are given by, $c_1, ..., c_{(m+1)n}$ and $s_1, ..., s_{(m+1)n}$, respectively. If c_N be the diagonal element of C with the value equal to zero (because rank of L and C are equal, exactly one c_i is zero), then we can choose $x_{(m+1)n} = \mathbf{1}$, which leads to,

$$\mathbf{d} = D\mathbf{1} = Dx_{(m+1)n} = s_{(m+1)n} w_{(m+1)n} \quad \text{and} \quad w_{(m+1)n}^T \mathbf{b} = \frac{\mathbf{d}^T \mathbf{b}}{s_{(m+1)n}}. \tag{7.25}$$

Notice that $||w_{(m+1)n}||_2^2 = 1$, therefore,

$$s_{(m+1)n}^2 = ||\mathbf{d}||_2^2. \tag{7.26}$$

Now combining 7.22 and 7.24, and regarding the fact that the matrix D is diagonal, and thus $D^T = D$, results in,

$$\mathbf{z}_\alpha = (D^2 + \alpha L^T L)^{-1} D\mathbf{b} = X(S^2 + \alpha C^2)^{-1} X^T X^{-T} S W^T \mathbf{b} = \sum_{i=1}^{(m+1)n} w_i^T \mathbf{b} \frac{s_i}{s_i^2 + \alpha c_i^2} x_i.$$

Therefore,

$$\lim_{\alpha\to\infty} \mathbf{z}_\alpha = \frac{w_{(m+1)n}^T \mathbf{b}}{s_{(m+1)n}} \mathbf{1}.$$

Combining this with 7.25 and 7.26, lead to our desired result and terminate the proof. □

It should be noted that the asymptotic value of $\mathbf{z}_\alpha$, (z), minimize the following least square fit, corresponding to the Dupire equation,

$$||\mathbf{b} - z\mathbf{d}||_2.$$

It is also possible to generate another approach for evaluating volatility surface. It consists of rewriting the Dupire equation 7.8, as a linear system for multiplicative inverse of z_{ij}, $(\mathbf{z}^{-1})$,

$$B\mathbf{z}^{-1} \approx \mathbf{d}$$

where similar to D, the matrix B is defined as the diagonal matrix with values of vector $\mathbf{b}$ on its diagonal. Again, this problem is ill-posed and similar to the first approach, by applying the same matrix L, and sufficiently large regularization parameter β, we can obtain the following regularization problem for evaluating $\mathbf{z}^{-1}$,

$$(B^2 + \beta L^T L)\mathbf{z}^{-1} = B\mathbf{d}. \tag{7.27}$$

Moreover, by applying theorem 7.2.4, with this approach, we obtain that there exists a volatility $\frac{1}{z}$, that minimizes the following least square fit, where the regularization parameter β is sufficiently large.

$$||\mathbf{d} - \frac{1}{z}\mathbf{b}||_2.$$

7.3 Numerical Results

Our objective in this section, is to apply the observed real data of the market to the approaches 7.22 and 7.27, in order to evaluate the volatility function. For doing this, we obtain the data of Russell 2000 index on August 1, 2018 from the "Yahoo Finance" website [Fin18]. The Russell 2000 index is obtained from the Russell 3000 index, which can be regarded as a benchmark for the whole US stock market. Russell 2000 index consider the lower 2000 stocks in the Russell 3000 index, and can be regarded as a small capitalization of stock market index. We show the available data for the Russell 2000 index from August 1, 2018, in the figure 7.1.

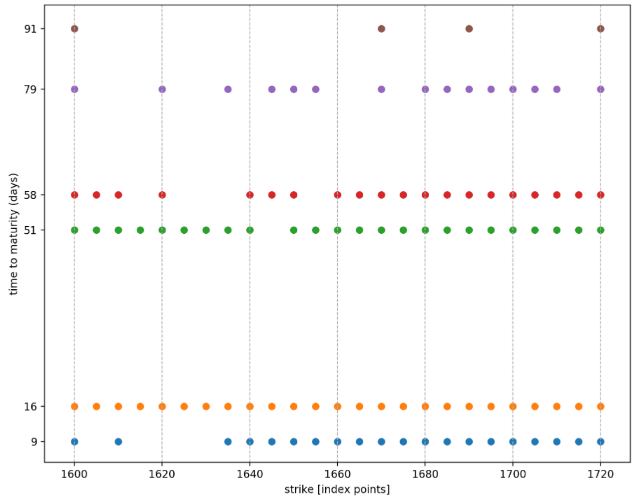

Figure 7.1: Available data for the Russell 2000 index from August 1, 2018

In the case of smooth interpolation of the data, according to our results, the standard parameter choice criterions, such as "L-curve criterion", are not useful for our data. Therefore, we assign the regularization parameter to cubic splines that both leads to the smooth interpolation, and secures the positiveness of the obtained interpolation and its second derivative. Figure 7.2 illustrates the resulting smooth interpolations.

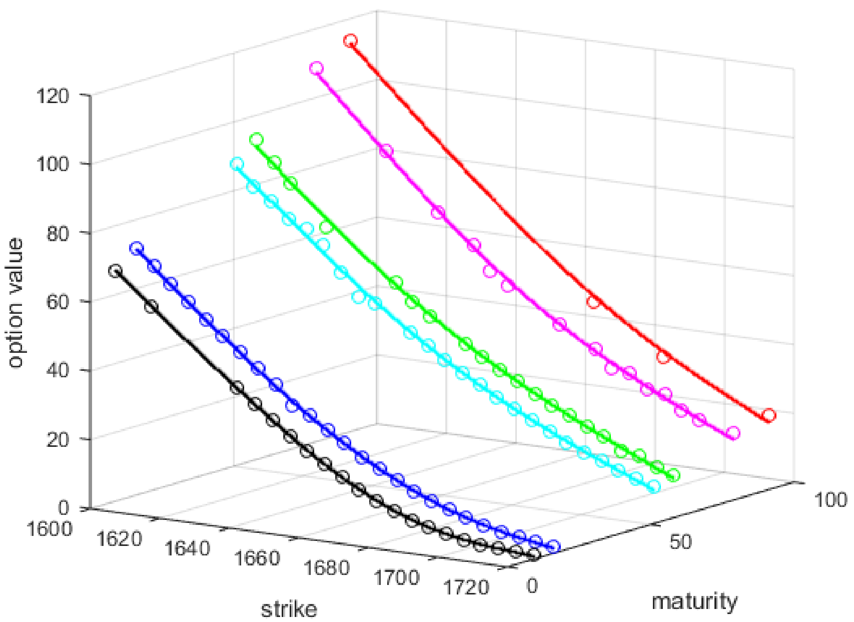

Figure 7.2: Smooth interpolation of the Russell 2000 index data

Although, as we mentioned earlier, there is no number in the market as a risk-free rate, the 4 week treasury bill rates are typically considered by authors as a risk-free interest rate. According to [otT18], the 4 week treasury bill rates in August 1, 2018 is equal to 0.019. Thus, we use this number in our calculations. Again, we choose the regularization parameter such that the obtained volatility surface becomes smooth. Using MATLAB [Mat14], the obtained volatility surfaces of our market data applying 7.22 and 7.27 are shown in figure 7.3 and 7.4, respectively. According to our results, both approaches roughly capture the volatility smile. While the second approach 7.27 gives bigger volatilities, the graph of the first approach 7.22, appears to be, qualitatively, more correct.

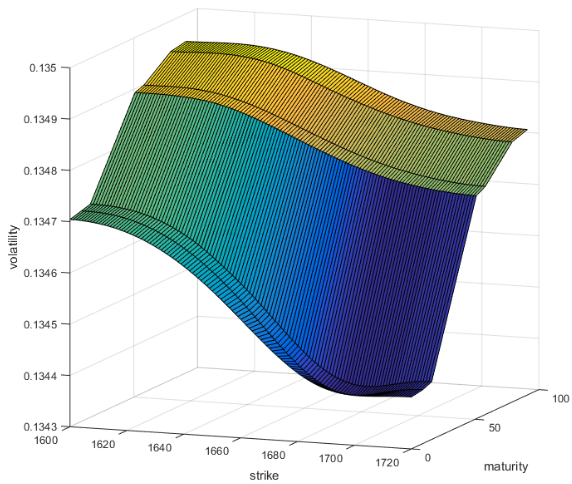

Figure 7.3: Volatility surface using 7.22

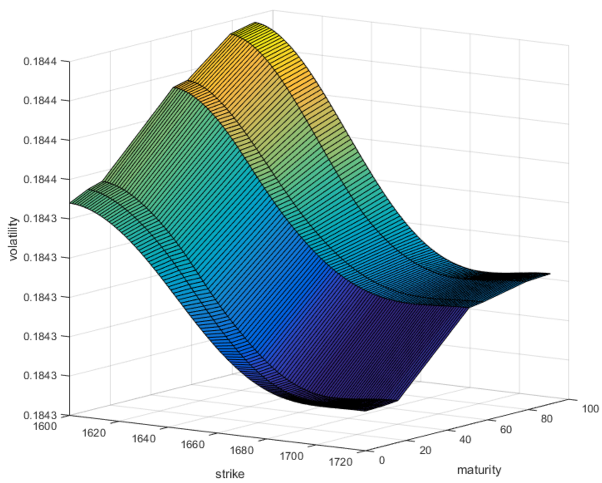

Figure 7.4: Volatility surface using 7.27

8 Conclusion

The aim of this thesis was to analyze and compute the Black-Scholes equation, which is a second order parabolic partial differential equation, to evaluate the option contracts. Throughout the thesis, we focused only on the European-type options. Although the exact solution is known for European options, numerical approximation of the Black-Scholes equation is vital for pricing more complex options, such as American options. In addition, according to empirical studies the original Black-Scholes equation fails to capture the real world features of option prices. In order to cope with this issue, it is necessary to extend the original Black-Scholes equation. Therefore, while the original Black-Scholes equation assume that the interest rate and volatility are constant, we regarded the extended version that assumes the interest rate is a function of time and volatility is a function of time and underlying asset price.

In order to develop the numerical solution of the Black-Scholes equation with local volatility, we followed the method which is mentioned in the book by Achdou and Pironneau [AP05]. This finite element method consists of obtaining the semi-discrete problem by discretizing the problem in time variable and applying the Crank-Nicolson scheme. Then, the fully discrete problem is derived by applying the Lagrange finite element and discretizing the problem in price variable. On the other hand, there is an analytic solution for the original Black-Scholes equation. Therefore, we assigned constant interest rate and volatility in our finite element method, and compared the results of two approaches together. Based on our experiments, the error around the strike and maturity is, by far, greater than other areas.

In the case of a posteriori error analysis Bergam, Bernardi and Mghazli [BBM05], regard the parabolic partial differential equations that are discretized with respect to time and space variables by an implicit Euler scheme and finite element respectively. Their main idea is to uncouple the time and space errors as much as possible. Therefore, they develop two families of error indicators, where the first one is global in space variable and local in time variable and the second family is local in both time and spatial variables. Achdou and Pironneau [AP05], apply this approach to develop a strategy to perform the adaptive mesh refinement for the Black-Scholes equation with local volatility. We followed the same method for adaptive mesh refinement where the Black-Scholes equation with local volatility is discretized in time by Crank-Nicolson scheme. Therefore, we evaluated two error indicators and show that these obtained error indicators agree with the actual error. By performing the adaptive mesh refinement algorithm, we obtain that the mesh is refined around the strike and maturity more than other areas of domain, which is according to our expectations.

To deal with shortcomings of the Black-Scholes equation in real-world applications, we followed the approach of Hanke and Rösler [HR05], which consists of calibrating the local volatility function from observed option prices in the market. First of all, they smoothening the observed data by cubic splines. Then, by applying the Dupire equation, the local volatility function is obtained by numerical differentiation. However, due to the fact that the obtained system is underdetermined and ill-posed, before solving the system, first order Tikhonov regularization is applied.

We expand the method by including the dividend in the calculations, in order to adjust the method to the real-world features of options. Then, we test the method with the real world data of Russell 2000 index on 1st of August 2018. According to our results, it seems that the method is sensitive to the initial data. In the theoretical point of view, the option prices are monotonically decreasing with respect to strike and are monotonically increasing with respect to maturity. However, this is not the case in real-world data, where some noises can be seen in the data. This phenomenon can lead to negative result, which is unacceptable. Therefore, excluding the noises are needed, which makes the underdetermined and ill-posed nature of the problem even worse. On the other hand, this method is fast and straightforward and appears to be qualitatively correct.

Bibliography

[AP05] Achdou, Yves and Olivier Pironneau: *Computational methods for option pricing*, volume 30. Siam, 2005.

[AP07] Achdou, Yves and Olivier Pironneau: *Finite element methods for option pricing*. Université Pierre et Marie Curie, 2007.

[Bac00] Bachelier, Louis: *Theory of speculation*. Dimson, E. and M. Mussavian (1998), A brief history of market efficiency, European Financial Management, 4(1):91–193, 1900.

[BBM05] Bergam, Amal, Christine Bernardi and Zoubida Mghazli: *A posteriori analysis of the finite element discretization of some parabolic equations*. Mathematics of computation, 74(251):1117–1138, 2005.

[Bei06] Beichelt, Frank: *Stochastic processes in science, engineering and finance*. Chapman and Hall/CRC, 2006.

[BI97] Bouchouev, Ilia and Victor Isakov: *The inverse problem of option pricing*. Inverse Problems, 13(5):L11, 1997.

[BJ99] Bodurtha, James N and Martin Jermakyan: *Nonparametric estimation of an implied volatility surface*. J. Comput. Finance, 2(0), 1999.

[BJ01] Buraschi, Andrea and Jens Jackwerth: *The price of a smile: Hedging and spanning in option markets*. The Review of Financial Studies, 14(2):495–527, 2001.

[BKSY17] Bangerth, Wolfgang, Imbunm Kim, Dongwoo Sheen and Jaeryun Yim: *On Hanging Node Constraints for Nonconforming Finite Elements using the Douglas–Santos–Sheen–Ye Element as an Example*. SIAM Journal on Numerical Analysis, 55(4):1719–1739, 2017.

[Boy77] Boyle, Phelim P: *Options: A monte carlo approach*. Journal of financial economics, 4(3):323–338, 1977.

[Boy86] Boyle, Phelim P: *Option valuation using a tree-jump process*. International Options Journal, 3:7–12, 1986.

[BR96] Becker, Roland and Rolf Rannacher: *A feed-back approach to error control in finite element methods: basic analysis and examples*. IWR, 1996.

[BR01] Becker, Roland and Rolf Rannacher: *An optimal control approach to a posteriori error estimation in finite element methods*. Acta numerica, 10:1–102, 2001.

[BS73] Black, Fischer and Myron Scholes: *The pricing of options and corporate liabilities.* Journal of political economy, 81(3):637–654, 1973.

[Chr96] Chriss, Neil: *Black Scholes and beyond: option pricing models.* McGraw-Hill, 1996.

[Cla73] Clark, Peter K: *A subordinated stochastic process model with finite variance for speculative prices.* Econometrica: journal of the Econometric Society, pages 135–155, 1973.

[CLV01] Coleman, Thomas F, Yuying Li and Arun Verma: *Reconstructing the unknown local volatility function.* In *Quantitative Analysis In Financial Markets: Collected Papers of the New York University Mathematical Finance Seminar (Volume II)*, pages 192–215. World Scientific, 2001.

[Cor01] Cortazar, Gonzalo: *Simulation and numerical methods in real options valuation.* 2001.

[CRR+79] Cox, John C, Stephen A Ross, Mark Rubinstein et al.: *Option pricing: A simplified approach.* Journal of financial Economics, 7(3):229–263, 1979.

[D+94] Dupire, Bruno et al.: *Pricing with a smile.* Risk, 7(1):18–20, 1994.

[DH96] Desmond, J and M Higham: *An Introduction to Financial Option Valuation*, 1996.

[Din13] Dineen, Seán: *Probability theory in finance: a mathematical guide to the Black-Scholes formula*, volume 70. American Mathematical Soc., 2013.

[Duf80] Duffy, Daniel J: *Uniformly convergent difference schemes for problems with a small parameter in the leading derivative.* PhD thesis, Trinity College Dublin, 1980.

[Duf13] Duffy, Daniel J: *Finite Difference methods in financial engineering: a Partial Differential Equation approach.* John Wiley & Sons, 2013.

[Duf18] Duffy, Daniel J: *Financial instrument pricing using C++.* John Wiley & Sons, 2018.

[EEHJ96] Eriksson, Kenneth, Don Estep, Peter Hansbo and Claes Johnson: *Computational differential equations*, volume 1. Cambridge University Press, 1996.

[EH96] Engl, Heinz W and Martin Hanke: *Neubauer 1996, Regularization of Inverse Problems*, 1996.

[Eva10] Evans, Lawrence C: *Partial differential equations.* 2010.

[EVZ04] Ern, Alexandre, Stéphane Villeneuve and Antonino Zanette: *Adaptive finite element methods for local volatility European option pricing.* International Journal of Theoretical and Applied Finance, 7(06):659–684, 2004.

[FHH04] FRANKE, JÜRGEN, WOLFGANG KARL HÄRDLE and CHRISTIAN M HAFNER: *Statistics of financial markets*, volume 2. Springer, 2004.

[Fin18] FINANCE, YAHOO: *Russell 2000*, 2018. [Online; accessed 1-August-2018].

[FL08] FOUFAS, GEORGIOS and MATS G LARSON: *Valuing Asian options using the finite element method and duality techniques.* Journal of computational and applied mathematics, 222(1):144–158, 2008.

[Fre83] FREEDMAN, DAVID: *Brownian Motion.* In *Brownian Motion and Diffusion*, pages 1–101. Springer, 1983.

[Fri08] FRIEDMAN, AVNER: *Partial differential equations of parabolic type.* Courier Dover Publications, 2008.

[G+96] GOLUB, GENE H et al.: *CF Van loan, Matrix computations.* The Johns Hopkins, 1996.

[Gla13] GLASSERMAN, PAUL: *Monte Carlo methods in financial engineering*, volume 53. Springer Science & Business Media, 2013.

[GNC14] GENG, JIAN, I MICHAEL NAVON and XIAO CHEN: *Non-parametric calibration of the local volatility surface for European options using a second-order Tikhonov regularization.* Quantitative Finance, 14(1):73–85, 2014.

[GRW15] GOLL, CHRISTIAN, ROLF RANNACHER and WINNIFRIED WOLLNER: *The damped Crank–Nicolson time-marching scheme for the adaptive solution of the Black–Scholes equation.* 2015.

[GS93] GREEN, PETER J and BERNARD W SILVERMAN: *Nonparametric regression and generalized linear models: a roughness penalty approach.* CRC Press, 1993.

[HB16] HULL, JOHN C and SANKARSHAN BASU: *Options, futures, and other derivatives.* Pearson Education India, 2016.

[HR05] HANKE, MARTIN and ELISABETH RÖSLER: *Computation of local volatilities from regularized Dupire equations.* International Journal of Theoretical and Applied Finance, 8(02):207–221, 2005.

[HS52] HESTENES, MAGNUS RUDOLPH and EDUARD STIEFEL: *Methods of conjugate gradients for solving linear systems*, volume 49. NBS Washington, DC, 1952.

[HS98] HANKE, MARTIN and OTMAR SCHERZER: *Error analysis of an equation error method for the identification of the diffusion coefficient in a quasi-linear parabolic differential equation.* SIAM Journal on Applied Mathematics, 59(3):1012–1027, 1998.

[HS01] HANKE, MARTIN and OTMAR SCHERZER: *Inverse problems light: numerical differentiation.* The American Mathematical Monthly, 108(6):512–521, 2001.

[Itō51] ITŌ, KIYOSI: *On stochastic differential equations*, volume 4. American Mathematical Soc., 1951.

[Jia05] Jiang, Lishang: *Mathematical modeling and methods of option pricing.* World Scientific Publishing Company, 2005.

[JS97] Jackson, Nicholas and Endre Suli: *Adaptive Finite Element solution of 1D European option pricing problems.* Technical Report, Unspecified, 1997.

[JSH98] Jackson, Nicholas, Endre Suli and Sam Howison: *Computation of deterministic volatility surfaces.* Technical Report, Unspecified, 1998.

[Kel71] Keller, Herbert B: *A new difference scheme for parabolic problems.* In *Numerical Solution of Partial Differential Equations–II*, pages 327–350. Elsevier, 1971.

[KNGK12] Krizek, Michal, Pekka Neittaanmäki, Roland Glowinski and Sergey Korotov: *Conjugate gradient algorithms and finite element methods.* Springer Science & Business Media, 2012.

[KR91] Kamrad, Bardia and Peter Ritchken: *Multinomial approximating models for options with k state variables.* Management science, 37(12):1640–1652, 1991.

[KS88] Knyazev, AV and AL Skorokhodov: *The rate of convergence of the method of steepest descent in a Euclidean norm.* USSR Computational Mathematics and Mathematical Physics, 28(5):195–196, 1988.

[Kwo08] Kwok, Yue-Kuen: *Mathematical models of financial derivatives.* Springer, 2008.

[Lan50] Lanczos, Cornelius: *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators.* United States Governm. Press Office Los Angeles, CA, 1950.

[Lar13] Larsson, Emil: *Option pricing using the discontinuous Galerkin method for time integration.* Uppsala University, 2013.

[LM12] Lions, Jacques Louis and Enrico Magenes: *Non-homogeneous boundary value problems and applications*, volume 1,2. Springer Science & Business Media, 2012.

[LO+97] Lagnado, Ronald, Stanley Osher et al.: *A technique for calibrating derivative security pricing models: numerical solution of an inverse problem.* Journal of computational finance, 1(1):13–25, 1997.

[Man67] Mandelbrot, Benoit: *The variation of some other speculative prices.* The Journal of Business, 40(4):393–413, 1967.

[Mar01] Marshall, John F: *Dictionary of financial engineering*, volume 74. John Wiley & Sons, 2001.

[Mat14] Mathworks: *MATLAB*, 2014. [Online; accessed 2017].

[Mer73] Merton, Robert C: *Theory of rational option pricing.* The Bell Journal of economics and management science, pages 141–183, 1973.

[otT18] Treasury, U.S. Department of the: *Daily Treasury Bill Rates Data*, 2018. [Online; accessed 1-August-2018].

[PA09] Pironneau, Olivier and Yves Achdou: *Partial differential equations for option pricing.* Handbook of numerical analysis, 15:369–495, 2009.

[PvP07] Persson, Jonas and Lina von Persson: *Pricing European multi-asset options using a space-time adaptive FD-method.* Computing and Visualization in Science, 10(4):173–183, 2007.

[Pyt08] Pytlak, Radoslaw: *Conjugate gradient algorithms in nonconvex optimization*, volume 89. Springer Science & Business Media, 2008.

[Rei67] Reinsch, Christian H: *Smoothing by spline functions.* Numerische mathematik, 10(3):177–183, 1967.

[SA86] Strang, Gilbert and Kaija Aarikka: *Introduction to applied mathematics*, volume 16. Wellesley-Cambridge Press Wellesley, MA, 1986.

[Sam65] Samuelson, Paul A: *Rational theory of warrant pricing.* IMR; Industrial Management Review (pre-1986), 6(2):13, 1965.

[Sey06] Seydel, Rüdiger: *Tools for computational finance*, volume 3. Springer, 2006.

[SS91] Stein, Elias M and Jeremy C Stein: *Stock price distributions with stochastic volatility: an analytic approach.* The review of financial studies, 4(4):727–752, 1991.

[SS00] Schötzau, Dominik and Christoph Schwab: *Time discretization of parabolic problems by the hp-version of the discontinuous Galerkin finite element method.* SIAM Journal on Numerical Analysis, 38(3):837–875, 2000.

[Ste11] Stefanica, Dan: *A primer for the Mathematics of Financial Engineering.* Fe Press, 2011.

[Str93] Strang, Gilbert: *Introduction to linear algebra*, volume 3. Wellesley-Cambridge Press Wellesley, MA, 1993.

[Str07] Strang, Gilbert: *Computational science and engineering*, volume 791. Wellesley-Cambridge Press Wellesley, 2007.

[Ten11] Teneng, Dean: *Limitations of the Black-Scholes model.* Collection of Papers 2011–Volume I, page 143, 2011.

[Top05a] Topper, Jürgen: *Financial engineering with finite elements*, volume 319. John Wiley & Sons, 2005.

[Top05b] Topper, Jürgen: *Option pricing with finite elements.* Wilmott Magazine, pages 84–90, 2005.

[Tur09] Turinici, Gabriel: *Calibration of local volatility using the local and implied instantaneous variance.* Journal of Computational Finance, 13(2):1–18, 2009.

[WDH93] Wilmott, Paul, Jeff Dewynne and Sam Howison: *Option pricing: mathematical models and computation.* Oxford financial press, 1993.

[WHHD95] Wilmott, Paul, Susan Howson, Sam Howison and Jeff Dewynne: *The mathematics of financial derivatives: a student introduction.* Cambridge university press, 1995.

Appendix

Proof of Proposition 1.
Regard the weak formulation 6.7, and replace $u_{\Delta t}$ by u and, add and subtract the expression $a_{t_n}(u_{\Delta t}(t), v) + a_{t_n}(u^n, v)$. Then,

$$\begin{aligned}
&\Big(\frac{\partial u_{\Delta t}}{\partial t}(t), v\Big) + a_t(u_{\Delta t}(t), v) = \\
&\Big(\frac{u^n - u^{n-1}}{\Delta t_n}, v\Big) + (a_t(u_{\Delta t}(t), v) - a_{t_n}(u_{\Delta t}(t), v)) + a_{t_n}(u_{\Delta t}(t) - u^n, v) + a_{t_n}(u^n, v) = \\
&\Big((a_t(u_{\Delta t}(t), v) - (a_{t_n}(u_{\Delta t}(t), v)\Big) + (a_{t_n}(u_{\Delta t}(t) - u^n, v)
\end{aligned}$$

for each $t \in (t_{n-1}, t_n]$ and $v \in V_0$.
Now, subtracting from the weak formulation 6.7, results in,

$$\begin{aligned}
&\Big(\frac{\partial(u - u_{\Delta t})}{\partial t}(t), v\Big) + a_{t_n}((u - u_{\Delta t})(t), v) = \\
&\qquad\qquad - (a_t(u_{\Delta t}(t), v) - a_{t_n}(u_{\Delta t}(t) - u^n, v)) - a_{t_n}(u_{\Delta t}(t), v).
\end{aligned}$$

Now, apply the change of variable, $e^{2\lambda t}v(t) = (u - u_{\Delta t})(t)$. Then, integrating, summing and mathematical simplification lead to,

$$\begin{aligned}
&[[u - u_{\Delta t}]]^2 \leq \\
&-2\sum_{m=1}^{n}\int_{t_{m-1}}^{t_m}\Big(a_\tau(u_{\Delta t}(\tau), v) - a_{t_m}(u_{\Delta t}(\tau), v)\Big)d\tau - 2\sum_{m=1}^{n}\int_{t_{m-1}}^{t_m} a_{t_m}(u_{\Delta t}(\tau) - u^m, v)d\tau.
\end{aligned}$$

As a result, we obtained the inequality with two terms in the right hand side , and our objective is to simplify each term.

In the case of first expression, applying the definition of bilinear form 6.5, previous assumptions and mathematical simplifications, lead to,

$$\begin{aligned}
&\Big|\int_{t_{m-1}}^{t_m}\Big(a_\tau(u_{\Delta t}(\tau), v) - a_{t_m}(u_{\Delta t}(\tau), v)\Big)d\tau\Big| \leq \\
&\Delta t_m \frac{2L_1 + L_2 + \frac{L_3}{2}}{\underline{\sigma}^2}\int_{t_{m-1}}^{t_m}\frac{\underline{\sigma}^2}{2}|u_{\Delta t}|_V|u - u_{\Delta t}|_V e^{-2\lambda\tau}d\tau \leq \\
&\Delta t_m \frac{2L_1 + L_2 + \frac{L_3}{2}}{\underline{\sigma}^2}\Big(\int_{t_{m-1}}^{t_m}\frac{\underline{\sigma}^2}{2}|u_{\Delta t}|_V^2 e^{-2\lambda\tau}d\tau\Big)^{\frac{1}{2}}\Big(\int_{t_{m-1}}^{t_m}\frac{\underline{\sigma}^2}{2}|u - u_{\Delta t}|_V^2 e^{-2\lambda\tau}d\tau\Big)^{\frac{1}{2}}.
\end{aligned}$$

The next step is to take the sum over m, in order to obtain the result for all time intervals.

Therefore,

$$
\begin{aligned}
&2\Big|\sum_{m=1}^{n}\int_{t_{m-1}}^{t_m}\Big(a_\tau(u_{\Delta t}(\tau),v)-a_{t_m}(u_{\Delta t}(\tau),v)\Big)d\tau\Big|\leq\\
&\frac{L}{\underline{\sigma}^2}c(u_0)\Delta t\Big(\sum_{m=1}^{n}\int_{t_{m-1}}^{t_m}\frac{\underline{\sigma}^2}{2}|u-u_{\Delta t}|_V^2e^{-2\lambda\tau}d\tau\Big)^{\frac{1}{2}}\\
&\frac{L}{\underline{\sigma}^2}c(u_0)\Delta t[[u-u_{\Delta t}]](t_n).
\end{aligned}
$$

It remains to consider the second term, applying the equation 6.8 and some mathematical operations, bring about,

$$
\begin{aligned}
&\Big|\int_{t_{m-1}}^{t_m}a_{t_m}(u_{\Delta t}(\tau)-u^m,v)d\tau\Big|\leq\\
&\mu\Big(\int_{t_{m-1}}^{t_m}|u_{\Delta t}(\tau)-u^m|_V^2e^{-2\lambda\tau}d\tau\Big)^{\frac{1}{2}}\Big(\int_{t_{m-1}}^{t_m}|v|_V^2e^{-2\lambda\tau}d\tau\Big)^{\frac{1}{2}}\leq\\
&\frac{\sqrt{2}\mu}{\underline{\sigma}}\Big(\int_{t_{m-1}}^{t_m}|u_{\Delta t}(\tau)-u^m|_V^2e^{-2\lambda\tau}d\tau\Big)^{\frac{1}{2}}\Big(\int_{t_{m-1}}^{t_m}\frac{\underline{\sigma}^2}{2}|u-u_{\Delta t}|_V^2e^{-2\lambda\tau}d\tau\Big)^{\frac{1}{2}}.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
&\Big(\int_{t_{m-1}}^{t_m}|u_{\Delta t}(\tau)-u^m|_V^2e^{-2\lambda\tau}d\tau\Big)^{\frac{1}{2}}\leq\\
&\Big(\frac{\Delta t_m}{3}\Big)^{\frac{1}{2}}e^{-\lambda t_{m-1}}|u^{m-1}-u^m|_V\leq\\
&\Bigg(\Delta t_me^{-2\lambda t_{m-1}}\Big(|u_h^{m-1}-u_h^m|_V^2+|u^{m-1}-u_h^{m-1}|_V^2+|u^m-u_h^m|_V^2\Big)\Bigg)^{\frac{1}{2}}
\end{aligned}
$$

where in the first inequality, we used the fact that $u_{\Delta t}(\tau)-u^m=\frac{t^m-\tau}{\Delta t_m}(u^{m-1}-u^m)$. Then, we add and subtract u_h^{m-1} and u_h^m to obtain the last inequality.

Then, we can bound the last term by,

$$
2(1+\rho_{\Delta t})\sum_{m=1}^{n}\Delta t_m\prod_{i=1}^{m-1}(1-2\lambda\Delta t_i)|u^m-u_h^m|_V^2\leq\frac{32}{\underline{\sigma}^2}(1+\rho_{\Delta t})[[u_{\Delta t}-u_{h,\Delta t}]]^2(t_n).
$$

Notice that we applied 6.20, and use the inequality $e^{-2\lambda t_n}\leq 2(\prod_{i=1}^{n}(1-2\lambda\Delta t_i))$, that was obtained during prove of lemma 6.1.1.

Finally,

$$
\begin{aligned}
&\Big|2\sum_{m=1}^{n}\int_{t_{m-1}}^{t_m}a_{t_m}(u_{\Delta t}(\tau)-u^m,v)d\tau\Big|\leq\\
&\frac{4\mu}{\underline{\sigma}^2}\Big(16(1+\rho_{\Delta t})[[u_{\Delta t}-u_{h,\Delta t}]]^2(t_n)+\sum_{m=1}^{n}\Delta t_me^{-2\lambda t_{m-1}}\frac{\underline{\sigma}^2}{2}|u_h^m-u_h^{m-1}|_V^2\Big)^{\frac{1}{2}}[[u-u_{\Delta t}]](t_n).
\end{aligned}
\tag{.1}
$$

Proof of Proposition 2. Assume that $v \in V_0$ and $v_h \in V_{nh}$, then,

$$(u^n - u_h^n, v) + \Delta t_n a_{t_n}(u^n - u_h^n, v) = \\ (u^n - u_h^n, v_h) + (u^n - u_h^n, v - v_h) + \Delta t_n a_{t_n}(u^n - u_h^n, v - v_h).$$

Now, suppose that we denote the minimum and maximum price of each element ω by ψ_1 and ψ_2, respectively. Then, applying integration by parts to the bilinear form terms lead to,

$$(u^n - u_h^n, v - v_h) + \Delta t_n a_{t_n}(u^n - u_h^n, v - v_h) = \\ (u^{n-1} - u_h^{n-1}, v - v_h) - (u_h^n - u_h^{n-1}, v - v_h) + \\ \Delta t_n \sum_{\omega \in \mathcal{T}_{nh}} \left(\int_\omega \Big(rS \frac{\partial u_h^n}{\partial S} - r u_h^n \Big)(v - v_h) dS - \frac{1}{4} \sum_{i=1}^{2} \sigma^2(\psi_i, t_n) \psi_i^2 \Big[\frac{\partial u_h^n}{\partial S} \Big](\psi_i)(v - v_h)(\psi_i) \right).$$

For the next step, assume that $v_v \in V_h$ are chosen in a way that at the mesh points, except for $S_i = 0$, $v_h(S_i) = v(S_i)$. Also in the case of $v_h(0)$, it is satisfied, $\int_0^{S_1}(v - v_h) = 0$. Then,

$$||v - v_h||_{L^2(\omega)} \leq \frac{h_\omega}{S_{max}} \Big|\Big| S \frac{\partial v}{\partial S} \Big|\Big|_{L^2(\omega)} \quad \text{and} \quad ||S(v - v_h)||_{L^\infty(\omega)} \leq C h_\omega^{\frac{1}{2}} \Big|\Big| S \frac{\partial v}{\partial S} \Big|\Big|_{L^2(\omega)}.$$

Hence,

$$-\frac{1}{4} \sum_{i=1}^{2} \sigma^2(\psi_i, t_n) \psi_i^2 \Big[\frac{\partial u_h^n}{\partial S} \Big](\psi_i)(v - v_h)(\psi_i) = 0.$$

Now, we can obtain the following bound,

$$\Bigg| - \int_\omega \frac{u_h^n - u_h^{n-1}}{\Delta t_n}(v - v_h) dS + \int_\omega \Big(rS \frac{\partial u_h^n}{\partial S} - r u_h^n \Big)(v - v_h) dS \\ - \frac{1}{4} \sum_{i=1}^{2} \sigma^2(\psi_i, t_n) \psi_i^2 \Big[\frac{\partial u_h^n}{\partial S} \Big](\psi_i)(v - v_h)(\psi_i) \Bigg| \leq C \eta_{n,\omega} \Big|\Big| S \frac{\partial v}{\partial S} \Big|\Big|_{L^2(\omega)}.$$

Regard the first expression of the proof, and apply the change of variable, $v = (u^n - u_h^n)$. Thus,

$$(1 - 2\lambda \Delta t_n)||u^n - u_h^n||^2 + \frac{1}{4} \Delta t_n \underline{\sigma}^2 |u^n - u_h^n|_V^2 \leq ||u^{n-1} - u_h^{n-1}||^2 + \frac{4C^2}{\underline{\sigma}^2} \Delta t_n \sum_{\omega \in \mathcal{T}_{nh}} \eta_{n,\omega}^2.$$

Notice that our objective is to bound the discrete norm $[[]]_n$, thus we would like to produce the terms of this norm on the left hand side of this inequality. As a result, multiply both sides of the inequality by $\prod_{i=1}^{n-1}(1 - 2\lambda \Delta t_i)$, and sum up over n. Therefore,

$$[[(u^m - u_h^m)]]_n^2 \leq \frac{c}{\underline{\sigma}^2} \sum_{m=1}^{n} \Delta t_m \prod_{i=1}^{m-1} (1 - 2\lambda \Delta t_i) \sum_{\omega \in \mathcal{T}_{mh}} \eta_{n,\omega}^2$$

where c is a constant.

Finally, applying the result of lemma 6.1.1, give us the following bound, which terminates the proof,

$$[[(u_{\Delta t} - u_{h,\Delta t})]]^2(t_n) \leq \frac{c}{\underline{\sigma}^2} \max(2, 1 + \rho_{\Delta t}) \sum_{m=1}^{n} \Delta t_m \prod_{i=1}^{m-1} (1 - 2\lambda \Delta t_i) \sum_{\omega \in \mathcal{T}_{mh}} \eta_{n,\omega}^2.$$

Proof of Proposition 3. In order to prove the bound for η_n, first of all apply the triangle inequality to obtain,

$$\eta_n \leq \sqrt{\Delta t_n} e^{-\lambda t_{n-1}} \frac{\sigma}{\sqrt{2}} |u^n - u^{n-1}|_V + \sqrt{\Delta t_n} e^{-\lambda t_{n-1}} \frac{\sigma}{\sqrt{2}} |u^n - u_h^n|_V + \sqrt{\Delta t_n} e^{-\lambda t_{n-1}} \frac{\sigma}{\sqrt{2}} |u^{n-1} - u_h^{n-1}|_V.$$

Our goal in this proof is to estimate each terms of the above bound separately. In the case of the second and third expressions, apply the inequality $e^{-2\lambda t_n} \leq 2\Big(\prod_{i=1}^n (1 - 2\lambda\Delta t_i)\Big)$, which mentioned in the proof of lemma 6.1.1, and the definition of regularity parameter 6.13, respectively, to obtain following results,

$$\sqrt{\Delta t_n} e^{-\lambda t_{n-1}} \frac{\sigma}{\sqrt{2}} |u^n - u_h^n|_V \leq \sqrt{2}[[u^n - u_h^n]]$$

and

$$\sqrt{\Delta t_n} e^{-\lambda t_{n-1}} \frac{\sigma}{\sqrt{2}} |u^{n-1} - u_h^{n-1}|_V \leq \sqrt{2\rho_{\Delta t}}[[u^{n-1} - u_h^{n-1}]].$$

However, obtaining the estimate for the first term is much more difficult. First of all, observe that by defining I, II, III and IV as,

$$I = 4e^{-2\lambda t^{n-1}} \int_{t_{n-1}}^{t_n} \frac{\partial(u - u_{\Delta t})}{\partial t}(\tau)(u^n - u^{n-1}) d\tau$$

$$II = 4e^{-2\lambda t^{n-1}} \int_{t_{n-1}}^{t_n} a_\tau(u - u_{\Delta t}(\tau), u^n - u^{n-1}) d\tau$$

$$III = 4e^{-2\lambda t^{n-1}} \int_{t_{n-1}}^{t_n} \Big(a_\tau(u_{\Delta t}(\tau), u^n - u^{n-1}) - a_{t_n}(u_{\Delta t}(\tau), u^n - u^{n-1})\Big) d\tau$$

$$IV = 2\lambda\Delta t_n e^{-2\lambda t_{n-1}} ||u^n - u^{n-1}||^2.$$

Then we have,

$$\sqrt{\Delta t_n} e^{-\lambda t_{n-1}} \frac{\sigma}{\sqrt{2}} |u^n - u^{n-1}|_V \leq I + II + III + IV.$$

Hence, we need to obtain the estimate for each term separately.

For the first term, simple mathematical operations lead to,

$$|I| \leq \frac{4\sqrt{2\Delta t_n}}{\underline{\sigma}} e^{-2\lambda t_{n-1}} \Big|\Big| \frac{\partial(u - u_{\Delta t})}{\partial t} \Big|\Big|_{L^2(t_{n-1}, t_n; V_0')} \frac{\sigma}{\sqrt{2}} |u^n - u^{n-1}|_V = \frac{4\sqrt{2}}{\underline{\sigma}} e^{-\lambda t_{n-1}} \Big|\Big| \frac{\partial(u - u_{\Delta t})}{\partial t} \Big|\Big|_{L^2(t_{n-1}, t_n; V_0')} \eta_n.$$

Likewise, we obtain the following estimate for the second term,

$$|II| \leq \frac{4\sqrt{2\mu\Delta t_n}}{\underline{\sigma}} e^{-2\lambda t_{n-1}} ||u - u_{\Delta t}||_{L^2(t_{n-1}, t_n; V_0)} \frac{\sigma}{\sqrt{2}} |u^n - u^{n-1}|_V = \frac{4\sqrt{2\mu}}{\underline{\sigma}} e^{-\lambda t_{n-1}} ||u - u_{\Delta t}||_{L^2(t_{n-1}, t_n; V_0)} \eta_n.$$

In the case of third term, we require to consider two different condition. Following the procedure in the proof of proposition 1, and applying previous assumptions and definitions, we obtain

- If $i = 1$,

$$|III| \leq \frac{\sqrt{2}L}{\underline{\sigma}}\Delta t_1 ||u_{\Delta t}||_{L^2(0,t_1;V_0)}\eta_1 \leq \frac{L}{\underline{\sigma}}(\Delta t_1)^{\frac{3}{2}}(|u^1|_V^2 + |u^0|_V^2)^{\frac{1}{2}}\eta_1$$
$$\leq \frac{2\sqrt{2}L}{\underline{\sigma}^2}\Delta t_1\Big([[(u^m)]]_1 + \frac{\underline{\sigma}}{\sqrt{2}}\sqrt{\Delta t_1}|u^0|_V\Big)\eta_1$$
$$\leq \frac{2\sqrt{2}L}{\underline{\sigma}^2}\Delta t_1\Big(||u^0|| + \frac{\underline{\sigma}}{\sqrt{2}}\sqrt{\Delta t_1}|u^0|_V\Big)\eta_1.$$

- If $i > 1$,

$$|III| \leq \frac{\sqrt{2}L}{\underline{\sigma}}e^{-\lambda t_{n-1}}\Delta t_n ||u_{\Delta t}||_{L^2(t_{n-1},t_n;V_0)}\eta_n$$
$$\frac{2L}{\underline{\sigma}}\Big(\prod_{i=1}^{n-1}(1-2\lambda\Delta t_i)\Big)^{\frac{1}{2}}\Delta t_n ||u_{\Delta t}||_{L^2(t_{n-1},t_n;V_0)}\eta_n$$
$$\frac{\sqrt{2}L}{\underline{\sigma}}(\max(1,\rho_{\Delta t}))^{\frac{1}{2}}\Delta t_n\Big(\Delta t_n\prod_{i=1}^{n-1}(1-2\lambda\Delta t_i)|u^n|_V^2 + \Delta t_{n-1}\prod_{i=1}^{n-2}(1-2\lambda\Delta t_i)|u^{n-1}|_V^2\Big)^{\frac{1}{2}}\eta_n$$
$$\frac{2L}{\underline{\sigma}^2}(\max(1,\rho_{\Delta t}))^{\frac{1}{2}}\Delta t_n[[(u^m)]]_n\eta_n \leq \frac{2L}{\underline{\sigma}^2}(\max(1,\rho_{\Delta t}))^{\frac{1}{2}}\Delta t_n||u^0||\eta_n.$$

It remains to obtain the estimate for the last term. We have,

$$IV = 2\lambda e^{-2\lambda t_{n-1}}\Big(\int_{t_{n-1}}^{t_n}(u^n - u^{n-1} - \Delta t_n\frac{\partial u}{\partial t})(u^n - u^{n-1})d\tau + \int_{t_{n-1}}^{t_n}\Delta t_n\frac{\partial u}{\partial t}(u^n - u^{n-1})d\tau\Big).$$

- For the first term

$$2\lambda e^{-2\lambda t_{n-1}}\Big|\int_{t_{n-1}}^{t_n}(u^n - u^{n-1} - \Delta t_n\frac{\partial u}{\partial t})(u^n - u^{n-1})d\tau\Big|$$
$$= 2\lambda e^{-2\lambda t_{n-1}}\Delta t_n\Big|\int_{t_{n-1}}^{t_n}\Big(\frac{\partial u_{\Delta t}}{\partial t} - \frac{\partial u}{\partial t}\Big)(\tau)(u^n - u^{n-1})d\tau\Big|$$
$$\leq \frac{2\sqrt{2}\alpha}{\underline{\sigma}}e^{-\lambda t_{n-1}}\Big|\Big|\frac{\partial(u-u_{\Delta t})}{\partial t}\Big|\Big|_{L^2(t_{n-1},t_n;V_0')}\eta_n.$$

- For the second term,

$$2\lambda e^{-2\lambda t_{n-1}}\Big|\int_{t_{n-1}}^{t_n}\Delta t_n\frac{\partial u}{\partial t}(u^n - u^{n-1})d\tau\Big|$$
$$\leq \frac{2\sqrt{2}\lambda}{\underline{\sigma}}\Delta t_n e^{\lambda\Delta t_n}||e^{-\lambda t}\frac{\partial u}{\partial t}||_{L^2(t_{n-1},t_n;V_0')} \leq \frac{4\mu\lambda}{\underline{\sigma}^2}\Delta t_n e^{\alpha}||u^0||\eta_n.$$

Similar to previous cases and applying 6.12,we obtain the estimate for these two terms. And this estimates finish the proof.

Proof of Proposition 6.2.2.
Recall that we obtained the following expression in proof of proposition 2,

$$(u^n - u_h^n, v - v_h) + \Delta t_n a_{t_n}(u^n - u_h^n, v - v_h) =$$
$$(u^{n-1} - u_h^{n-1}, v - v_h) - (u_h^n - u_h^{n-1}, v - v_h) +$$
$$\Delta t_n \sum_{\omega \in \mathcal{T}_{nh}} \Bigg(\int_\omega \Big(rS\frac{\partial u_h^n}{\partial S} - ru_h^n \Big)(v - v_h)dS - \frac{1}{4}\sum_{i=1}^{2} \sigma^2(\psi_i, t_n)\psi_i^2 \Big[\frac{\partial u_h^n}{\partial S}\Big](\psi_i)(v - v_h)(\psi_i) \Bigg).$$

Suppose that ψ_ω is a bubble function on ω, then, apply the change of variable $v^h = 0$. Also,

$$v = \Big(\frac{u_h^n - u_h^{n-1}}{\Delta t_n} - rS\frac{\partial u_h^n}{\partial S} + ru_h^n \Big)\psi_\omega \text{ on } \omega, \quad v = 0 \text{ elsewhere.}$$

Then,

$$||v||^2_{L^2(\omega)} = \Big|\Big| \Big(\frac{u_h^n - u_h^{n-1}}{\Delta t_n} - rS\frac{\partial u_h^n}{\partial S} + ru_h^n \Big)\psi_\omega^{\frac{1}{2}} \Big|\Big|^2_{L^2(\omega)}$$
$$= \frac{1}{\Delta t_n}(u^{n-1} - u_h^{n-1}, v) - \frac{1}{\Delta t_n}(u^n - u_h^n, v) + a_{t_n}(u_h^n - u^n, v).$$

Now, for all w that belong to the space of first order polynomials, following inverse inequalities are derived,

$$||w||_{L^2(\omega)} \le c_1 ||w\psi_\omega^{\frac{1}{2}}||_{L^2(\omega)} \quad \text{and} \quad \Big|\Big| S\frac{\partial w}{\partial S} \Big|\Big|_{L^2(\omega)} \le c_2 \frac{S_{max}(\omega)}{h_\omega} ||w||_{L^2(\omega)}.$$

Hence, by combining these results we obtain,

$$\Big|\Big| \Big(\frac{u_h^n - u_h^{n-1}}{\Delta t_n} - rS\frac{\partial u_h^n}{\partial S} + ru_h^n \Big) \Big|\Big|^2_{L^2(\omega)}$$
$$\le c\frac{S_{max}(\omega)}{h_\omega} \Bigg(\Big|\Big| \frac{u^{n-1} - u_h^{n-1} - u^n + u_h^n}{\Delta t_n} \Big|\Big|_{V_0'(\omega)} + \mu \Big|\Big| S\frac{\partial (u^n - u_h^n)}{\partial S} \Big|\Big|_{L^2(\omega)} \Bigg) ...$$
$$... \Big|\Big| \Big(\frac{u_h^n - u_h^{n-1}}{\Delta t_n} - rS\frac{\partial u_h^n}{\partial S} + ru_h^n \Big) \Big|\Big|_{L^2(\omega)}.$$

Finally, eliminating common expressions from both sides, leads us to our desired result,

$$\frac{h_\omega}{S_{max}(\omega)} \Big|\Big| \Big(\frac{u_h^n - u_h^{n-1}}{\Delta t_n} - rS\frac{\partial u_h^n}{\partial S} + ru_h^n \Big) \Big|\Big|_{L^2(\omega)}$$
$$\le \Bigg(\Big|\Big| \frac{u^{n-1} - u_h^{n-1} - u^n + u_h^n}{\Delta t_n} \Big|\Big|_{V_0'(\omega)} + \mu \Big|\Big| S\frac{\partial (u^n - u_h^n)}{\partial S} \Big|\Big|_{L^2(\omega)} \Bigg).$$