2017

A further study of the inverse finite element approach for pricing American-style options

Bolujo Joseph Adegboyegun
University of Wollongong

A further study of the inverse finite element approach for pricing American-style options

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

Doctor of Philosophy

from

University of Wollongong

by

Bolujo Joseph Adegboyegun

School of Mathematics and Applied Statistics

Certification

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

Bolujo Joseph Adegboyegun

May 29, 2017

I would like to dedicate this thesis to my loving and supportive wife, Moyosola Adegboyegun, our exuberant, sweet, and kind-hearted children, Ololade and Olamide Adegboyegun.

Abstract

This thesis studies the pricing of American-style options under different formulations and frameworks. A PDE-based computational framework is adopted with focus on the inverse finite element method. The algorithms developed are based on a novel inverse isotherm finite element method which is suited to problems with phase change as known in mechanics.

The first contribution of the thesis is in Chapter 3. We investigate the feasibility of trading the roles of dependent and independent financial variables in economically and mathematically meaningful ways, thus allowing the problem to be solved in an inverse manner. The inverse algorithm is able to deal with the nonlinearity of the option pricing problems without any regularization or linearization. Moreover, we compare the computational efficiency of the direct and inverse approaches by carrying out a critical performance analysis of the two approaches against some benchmark solutions. The results demonstrate that the inverse approach is indeed more efficient than its direct counterpart, as a higher performance in terms of an acceptable computing time and accuracy is achieved.

The work presented in Chapter 4 is motivated by the tractability enjoyed by the free boundary problem of an American option. The option pricing problem is formulated as a linear complementarity problem. A new numerical algorithm based on the inverse approach is proposed. The key feature of the inverse algorithm is that the solution is limited to the yielded domain of the option, thus, the solution corresponds to the original pricing model. Furthermore, there is a reduction of the total number of unknown due to the *a priori* known option values of designated underlying. Additionally, the algorithm enjoys high efficiency as a result of using the Newton iterative scheme with its inherent quadratic convergence.

The primary contribution of this thesis is in Chapter 5, in which a new hybrid algorithm that directly prescribes dynamics of option problems under stochastic volatility model is proposed. The technique is similar in some respect to problem under classical Black-Scholes model, although with the combination of finite differences to discretize the volatility derivative terms. The implementation is based on the fact that the nodal locations along the volatility direction are fixed, while working out the motion of the nodes

along the underlying direction. The resulting non-linear system of equations is solved by the Newton iteration method to determine the optimal exercise price. We establish the convergence by reformulating the discretized problem in variational form and study the approximation of the option price.

The last part of the thesis summarizes the main results achieved and propose future research directions to extend these results.

Acknowledgements

I owe the Almighty God profound appreciation for making it possible for me to push through this research. He brought me, helped me to see and empowered me to conquer. This research would not have been possible if not for the commitment put into the reading of my drafts, and the useful guidance and suggestions offered by my supervisor, Professor Song-Ping Zhu. The same level of appreciation is extended to my co-supervisor, Dr Xiaoping Lu for generously given her time and expertise to better the work. I thank them for their contribution and their good-natured support.

I am hugely indebted to certain organisations for giving me the required funding assistance that made this research possible. These included Federal Government of Nigeria through the Tertiary Education Trust Fund (TETFUND), for the sponsorship. Many thanks to the management of Ekiti State University, Ado-Ekiti, Nigeria, for the assistance it could offer. I would like to express my deep gratitude to the Faculty of Engineering and Information Science, University of Wollongong for the travel support to attend The 12th Engineering Mathematics and Applications Conference (EMAC, 2015), University of South Australia, Adelaide, and The Australian and New Zealand Industrial and Applied Mathematics conference (ANZIAM, 2016), Canberra, where I presented aspects of my thesis.

My appreciation goes, in particular, to Professor Philip Ogunbona, who introduced me to my primary supervisor. All your initial support to get me settled in and everything subsequent to it are all appreciated. Professor E. A Ibijola and Professor J.A Kayode of Ekiti State University, Nigeria deserve my profound appreciation for the invaluable contribution they made for my sponsorship application to succeed. My postgraduate colleagues in the Financial Mathematics research group are also appreciated for every contribution they made to my work.

The decision I took to come to Australia for the PhD programme was disastrous for my wife and kids who could not be part of the journey to Australia. For them, it was a series of longings, wishes and desires. I therefore openly acknowledge the profound support of my wife, Mrs Moyosola Dorcas Adegboyegun, who bore the heavy burden of running the home, for all the pains and stress, plus the enormous pressure of certain loneliness. I also

wish to acknowledge our wonderful children, Ololade and Olamide Adegboyegun, for their naivety and belief that I can buy them anything. Every encounter with children of your ages in Australia was a painful reminder of the reality of my physical dislocation. I love you for everything you are to me.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Options

Derivative securities are financial instruments that derive their values from the performance of an underlying entity (asset, interest rate or index). Options are the most common derivative securities that frequently traded in financial markets. The theory of option pricing began in 1900 when Bachelier [7] deduced an option pricing formula based on the assumption that stock prices follow a standard Brownian motion. The main drawbacks of Bachelier's model are that the probability of negative stock prices is positive and the option prices may be greater than their respective stock prices. With the assumption that stock prices follow a geometric Brownian motion, Black and Scholes [17] derived an analytical formula for valuing European options of a fixed lifetime. Black and Scholes published their work at the time roughly coincided with the opening of the Chicago Board of Trade when trading on an official exchange began. Since then, the growth in the option markets has been quite extraordinary. As pointed out by Hunt and Kennedy [56], this dramatic growth can be attributed to two main factors.

The first, and most important, is that options can fulfil two main needs of investors: hedging and speculating. Nowadays, business organization and private enterprize with sizeable assets are exposed to financial risks connected with the movement of the world markets. For examples, manufacturers are at risk of increase in commodity prices; multinationals are exposed to unfavorable moves in exchange rates. Suitable options can help these entities reduce their exposure to adverse market moves which are beyond their control. On the other hand, options can also be used to acquire risk or to speculate about the

future. Often, investors have their own views about the movement of markets. Depending on whether their views are right or wrong, they can earn high profits or suffer great losses. By investing in options, investors may earn high profits but at a much cheaper cost, in comparison with the cost of investing directly in the underlying assets. For example, if investors believe the price of some stocks will go up for a certain period of time, then using a suitable option can produce a higher return for the investors than buying the stocks directly [55].

The second factor is the parallel development of the financial mathematics needed for financial institutions to be able to price and hedge the products demanded by their customers. Since Black and Scholes [17] showed in their celebrated paper, that the pricing of options can be formulated as a deterministic partial differential equation (PDE), the option pricing theory has experienced rapid growth and become a major area in contemporary quantitative finance research. In order to meet various needs of investors, variety of option types have been created by adding additional features to the financial contracts of plain vanilla options. As a direct consequence, the corresponding pricing problems have become much more challenging and thus there are obvious needs for more research effort on how to determine the reasonable prices of newly created options.

In the literature, there are two main groups of options, namely, Vanilla and Exotic options.

1.1.1 Vanilla options

A vanilla option is a financial contract giving its holder the right, without the obligation, to either buy (call option) or sell (put option) an underlying asset at a predetermined price, K, (the exercise or strike price) up to a specified expiration date, T. The underlying asset is typically a stock. If the holder uses his right to buy (or sell) the underlying from (or to) the issuer, he exercises the option. Options which can be exercised at any time until expiry are called American options. There are also contracts which cannot be exercised at any time $t \in [0, T]$ but only at maturity, T. Those options are called European options. Options that can be exercised at a (pre-specified) discrete set of times $\mathcal{T} = t_1, \ldots, t_m$ are called Bermudan options. American options are usually valued more than their European counterparts, since the American options give their holder more rights than the European options, via the right of early exercise. From the mathematical modelling perspective,

American options are more interesting since they can be formulated as a free boundary problem.

1.1.2 Exotic options

An exotic option differs from commonly traded "vanilla" options in terms of the underlying asset or the calculation of how or when the investor receives a certain payoff. These options are more complex than vanilla options, and they are generally traded over the counter. The complexity usually relates to determination of payoff. The payoff function of most exotic options depends on the path of the underlying asset price as well as its value at the maturity date. For example, an Asian option is a fully path dependent option. The payoff function depends on the average of the underlying asset over a specific time period. Another example of the exotic option is the barrier option. For a barrier option, the right of the exercise is either activated (an *in barrier*) or forfeited (an *out barrier*) when the underlying asset price hits a prescribed value at some time before the maturity date.

Regarding the valuation of exotic options, given their complexity, they are usually modelled using specialized simulation or lattice-based techniques, and it is relatively complicated to calculate the price or set up a hedge strategy. Despite their embedded complexities, exotic options have certain advantages over regular options, which include: more adaptable to specific risk management needs of individuals or entities and greater range of investment products to meet investors‘ portfolio needs.

In this thesis we concentrate mainly on American vanilla options, for which there are no known closed form or analytical solutions. A new numerical method to price American options under different formulations and frameworks will be discussed.

1.2 Mathematical background

Mathematical finance has been a rapidly growing area of mathematics over the past few decades. As a subsection of applied mathematics, it deals with the modelling aspects of finance. Stochastic calculus and martingale theory are the perfect mathematical tools for modelling financial derivatives. Models based on stochastic process have turned out to be highly tractable and usable in practice [56]. In this Section, we review the background mathematical tools used in this thesis.

1.2.1 Stochastic calculus

Brownian motion

We start by recalling the definition of Brownian motion, which is a fundamental example of a stochastic process. The underlying probability space $(\Omega, \mathcal{F}, \mathbb{P})$ of Brownian motion can be constructed on the space $\Omega = \mathcal{C}_0(\mathbb{R}_+)$ of continuous real-valued functions $\mathbb{R}_+$ starting from 0.

Definition 1.2.1. *The Brownian motion with drift is defined as a stochastic process $W_t \in \mathbb{R}_+$ with the following properties:*

1. *for $t > 0$ and $s > 0$, every increment $W_{t+s} - W_s$ is normally distributed with mean μt and variance $\sigma^2 t$, where μ and σ are fixed parameters,*
2. *for every $t_1 < t_2 < \cdots < t_n$, the increments $X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent random variables with distribution given in (1),*
3. *$W_0 = 0$ and the sample paths of W_t are continuous.*

It should be noted that $W_{t+s} - W_s$ is independent of history, i.e. knowing W_τ, $\tau < s$ has no effect on the probability distribution of $W_{t+s} - W_s$. For the particular case, when $\mu = 0$ and $\sigma^2 = 1$, the Brownian motion is called standard Brownian motion (or standard Wiener process) denoted as W_t, which is a continuous-time stochastic process.

Itô's Lemma

Itô's lemma is the most important result about the manipulation of random processes. It plays the role in stochastic calculus that the fundamental theorem of calculus plays in ordinary calculus. Itô's lemma is key to the study of stochastic calculus by linking the small change in a function of a random variable to the small change in the random variable itself.

Theorem 1. *(Itô's lemma). Suppose X_t is a Itô drift-diffusion process that satisfies the stochastic differential equation*

$$dX_t = \mu_t dt + \sigma_t dW_t,$$

where where W_t is a Wiener process. Then any twice differentiable function $f = f(t, x)$

admits the stochastic dynamics given by

$$df = \left(\frac{\partial f}{\partial t} + \mu_t \frac{\partial f}{\partial x} + \frac{\sigma_t^2}{2} \frac{\partial^2 f}{\partial x^2} \right) dt + \sigma_t \frac{\partial f}{\partial x} \, dW_t.$$

Connections between SDE and PDE

The Faynman-Kac theorem and Kolmogoroff (Fokker-Plank) backward equation establish the connection between the SDE and PDE. With this connection, the expectation of an Itô process can be obtained by solving the associated PDE. On the other hand, for certain PDEs, we can express its solution by an expectation of an Ito process [9, 76]. Therefore, for option pricing purposes, we stress this important result relating expectations with respect to realizations of stochastic processes to specific PDEs

Theorem 2. *(Kolmogoroff Backward Equation and Feynman-Kac theorem)*
Assume that the system state $\bar{x}(t)$ evolves according to the stochastic differential equation

$$d\bar{x}_t = \mu(t, \bar{x}_t)dt + \sigma(t, \bar{x}_t)dW_t \tag{1.1}$$

where $\mu(t, x)$ and $\sigma_{i,j}(t, x)$ are continuous, and they satisfy the Lipschiz and growth conditions:

$$\|\mu(t, x) - \mu(t, y)\| + \|\sigma(t, x) - \sigma(t, y)\| \leq C\|x - y\|,$$

$$\|\mu(t, x)\|^2 + \|\sigma(t, x)\|^2 \leq C^2(1 + \|x\|^2)$$

with C being constant. Let T be arbitrary but fixed, and let $L \geq 0, \lambda \geq 0$ be approximate constants. Let $f(x) : R^d \to R, g(t, \bar{x}_t) : [0, T] \times R^d \to R, K(t, \bar{x}_t) : [0, t] \times R^d \to R$ be continuous functions and satisfy

$$|f(\bar{x})| \leq L(1 + \|x\|^2), \text{ or } f(\bar{x}) \geq 0,$$

$$|g(t, \bar{x})| \leq \lambda(1 + \|x\|^2), \text{ or } g(t, \bar{x}) \geq 0.$$

Suppose that $V(t, \bar{x})$ is continuous and belong to $C^{1,2}([0, T] \times R^d)$, and moreover, satisfies

the Cauchy problem:

$$\begin{cases} \dfrac{\partial V}{\partial t} + A_t V(t,\bar{x}) - K(t,\bar{x})V + g(t,\bar{x}) = 0, \\ V(T,\bar{x}) = f(\bar{x}), \end{cases} \tag{1.2}$$

as well as the polynomial growth condition

$$\max_{0\le t\le T} |V(t,\bar{x})| \le M(1+\|\bar{x}\|^{2\mu}), M>0, \mu \ge 1,$$

where A_t is the infinitesimal generator of the stochastic process (1.1). Then $V(t,\bar{x})$ admits the stochastic representation:

$$V(t,\bar{x}) = E[\int_0^T exp(-\int_0^t K(u,\bar{x}_u)dug(s,\bar{x}_s)ds + exp(-\int_0^T K(\bar{x}_s(w))ds)f(\bar{x}_T)|\bar{x}_t = \bar{x}].$$

1.2.2 Basic numerical methods for partial differential equation

Here, we briefly review two numerical methods, namely, finite difference method (FDM) and finite element method (FEM), that are useful in quantitative analysis of contemporary financial markets.

Finite difference method (FDM)

Finite difference method provides a versatile tool for the numerical solution of problems described by differential equations. The method consists of a discrete grid, $\Omega_\triangle := x_j$, and a grid function, $W_\triangle := W_j$. The grid $\Omega_\triangle$ is a graph of discrete grid points $x_j \in \Omega \subset \mathbb{R}^d$ and a certain set of their neighbors, x_{jk} , $jk \in \mathcal{N}(j)$. The vectors $\{x_j - x_{jk}\}_{jk\in\mathcal{N}(j)}$ form the *difference* associated with x_j. Here, $\triangle$ abbreviates one or more discretization parameters of the underlying grid, $\Omega_\triangle$, which measure the clustering of these neighbors: the smaller $\triangle$ is, the closer x_{jk} are to x_j. Divided differences along appropriate discrete points are used to approximate the partial derivatives of the PDE. The resulting relations between the divided differences form a *finite-difference scheme*. Its solution, W_j, is sought as an approximation to the pointvalues of the unknown exact solution of the nonlinear problems, $w(x_j)$, as we refine the grid by letting $\triangle \to 0$.

The construction of finite difference scheme proceeds by replacing partial derivatives occurring in the PDEs with approximate divided differences. There are three commonly

used finite difference approximations, namely, the forward approximation, the backward approximation, and the central approximation, which are defined as

$$D_{+x}W_{jk} = \frac{W_{j+1,k} - W_{jk}}{\Delta x} + \mathcal{O}(\Delta x),\ D_{-x}W_{jk} = \frac{W_{jk} - W_{j-1,k}}{\Delta x} + \mathcal{O}(\Delta x),$$

$$D_{0x}W_{jk} = \frac{W_{j+1,k} - W_{j-1,k}}{2\Delta x} + \mathcal{O}(\Delta x),$$

respectively.

The derivation of FDM is straightforward and they are easy to implement. Moreover, they appeal to the full spectrum of linear and nonlinear PDEs. For the method to be useful, the accuracy of the approximation should improve as the grid spacings tend to zero i.e., it must be convergent. However, proving a given scheme to be convergent is not easy in general, if attempted in a direct manner. The Lax-Richtmyer equivalence theorem, which is a fundamental theory of the FDM, reveals a way to check the convergence of a given FDM.

Theorem 3. *(The Lax-Richtmyer Equivalence Theorem) A consistent finite difference scheme for a PDE for which the initial value is well-posed is convergent if and only if it is stable.*

According to the Lax-Richtmyer equivalence theorem, a consistent finite difference scheme for a well-posed linear initial value problem is convergent if and only if it is stable. Thus, when dealing with a complex problem, the convergence condition can be replaced with easily proved conditions of consistency and stability. Note that a numerical algorithm is stable if a small error at any stage produces a cumulative error of an order smaller or at most equal to that of the original one, and it is considered consistent if by reducing the mesh and time step size, the truncation error terms could be made to approach zero.

Finite element method (FEM)

Finite element method (FEM) offers greater flexibility in modeling problems with complex geometries, and, as such, they have been widely used in science and engineering as the solvers of choice for nonlinear problems. It is based on using variational methods and/or weak formulations. Instead of using the finite differences to approximate the derivatives,

the FEM converts the PDE into an integral form. The use of the integral form is advantageous since it provides a reasonable treatment of Neumann boundary conditions. These conditions are very common in quantitative finance, most especially, when estimating the behavior of an option as the underlying price goes to infinity.

To proceed with the FEM approximation, one partitions the domain of interest, $\Omega \subset \mathbb{R}^d_x$, into a set of non-overlapping polyhedrons, $\Omega_\triangle = T_j$, namely, the elements. A piecewise polynomial finite-element approximation is sought, $W(x) = \sum_k W_k \varphi_k(x)$, in terms of polynomial elements, $\varphi_k(x)$, supported on $T_\ell, \ell \in \mathcal{N}(k)$. The next step is to assemble all the element matrices to obtain a global matrix. Then, the unknown nodal values are found in an efficient manner.

It should be remarked that the FDM is a special case of the FEM for one-dimensional problem. It can be easily shown that if the basis functions are chosen as either piecewise constant functions or Dirac delta functions, the FEM degenerates to the FDM. However, unlike the finite difference algorithm that can be implemented in a straightforward manner, finite element approach is based on the variational/weak formulation. An overview of FEM implementation details is deferred to Section 2.3.

1.3 Option pricing models

Research in option pricing theory concerns, among many other issues, the computation of the value of an option during the life of the option contract. A famous equation for this is the Black-Scholes model. It represents a simple model for the values of two basic options, the so-called put option and call option. However, it has been widely acknowledged that the classical Black-Scholes model cannot capture the behaviour of modern financial markets, such as the smile or smirk effects. The natural extension of the Black-Scholes model that has been pursued in the literature and in practice is to modify the specification of the volatility. In what follows, we shall give an overview of some pricing models used for option derivatives in this thesis.

1.3.1 Black-Scholes model

The Black-Scholes model, which was introduced in 1973 by Fisher Black, Myron Scholes and Robert Merton, provides an approximate description of the behaviour of the under-

lying. The key idea behind the model is to hedge the option by buying and selling the underlying asset in just the right way and, as a consequence, to eliminate risk. This model now serves as a benchmark against which other models can be compared. The assumptions used in the Black-Scholes model are given as follows [107]:

- The risk-free interest rate, r and the volatility rate, σ are known and constant during the option life.
- The underlying asset price, S follows a geometric Brownian motion governed by:

$$dS = \mu S dt + \sigma S dZ, \tag{1.3}$$

  where μdt is a deterministic return, the volatility, σ measures the standard deviation of the returns dS/S and dZ is assumed to be a Wiener process with mean, 0 and variance, dt.
- There are no risk-free arbitrage opportunities.
- Security trading is continuous.
- There are no transaction costs or taxes in option trading.
- Short selling of securities is permitted and all securities are perfectly divisible.

Under these assumptions, we want to construct a portfolio consisting of one option and a quantity $-\triangle$ of the underlying asset so that the risk can be removed. We will denote the value of the portfolio by $\Pi = V - \triangle S$ and the change in the value of the portfolio in one time-step is $d\Pi = dV - \triangle dS$. Then, using the no-arbitrage assumption, Π must instantaneously earn the risk-free bank rate, r, $d\Pi = r\Pi dt$. In order to eliminate the stochastic component of risk contained in the portfolio, Π, the number of the underlying must equal to $\triangle = \frac{\partial V}{\partial S}$.

Applying the Itôs lemma to V and after some algebraic manipulations, the Black-Scholes equation governing the price of the option can be obtained as [107]:

$$\frac{\partial V}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0. \tag{1.4}$$

Although the original Black-Scholes model assumed no dividends, trivial extensions to the model can accommodate a continuous dividend yield factor to price a wider range of

options as will be discussed in Chapter 4.

However, it might be important to provide some financial insights behind Equation (1.4). First, with (1.4), one can perfectly hedge the option by buying and selling the underlying asset in just the right way and, consequently "eliminate risk". This hedge, in turn, implies that there is only one right price for the option, as returned by the Black-Scholes formula. Also, the Delta here shows the rate of change of the value of the option or portfolio of options with respect to the underlying, S. It is of fundamental importance in option pricing theory, and provides a measure of correlation between the movements of the option or other derivative products and those of the underlying asset.

Furthermore, the linear differential operator L_{BS}, i.e.,

$$L_{BS} = \frac{\partial}{\partial t} + \frac{\sigma^2 S^2}{2}\frac{\partial^2}{\partial S^2} + rS\frac{\partial}{\partial S} - rI,$$

where I is the identity operator, has a financial interpretation as a measure of the difference between the return on a hedged option portfolio and the return on a bank deposit. Although this difference must be identically zero for a European option, it is not so for an American option. Finally, the Black-Scholes Equation (1.4) does not contain the drift parameter μ. In other words, the value of an option is independent of how rapidly or slowly the asset grows.

One of the most remarkable contributions of Black and Scholes to quantitative finance research is the derivation of a closed-form analytic solution, known now as the Black-Scholes formula. The value of a call option for a non-dividend paying underlying stock in terms of the BlackScholes parameters is:

$$C(S,t) = N(d_1)S - N(d_2)Ke^{-r(T-t)} \tag{1.5}$$

where

$$d_1 = \frac{\ln(S/K) + (r + \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}},$$

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

and $N(d)$ is the standard normal distribution function defined as

$$N(d) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{d} e^{-x^2/2} dx.$$

Based on put-call parity, the price of a corresponding put option is:

$$P(S,t) = N(-d_2)Ke^{-r(T-t)} - N(-d_1)S \tag{1.6}$$

It should be emphasized that despite the success of the Black-Scholes model in parsimoniously describing market options prices, there are number of ways in which the Black-Scholes model has been shown to disagree with observed reality. For instance, the most questionable assumption of the model is that continuously compounded stock returns are normally distributed with constant volatility. Traders who use Black-Scholes model to hedge must continuously change the volatility assumption in order to match market prices. Also, the prices of exotic options given by models based on Black-Scholes assumptions can be wildly wrong. Dealers in such options are motivated to find models which can take the volatility smile into account when pricing it. To handle such defects, the literature has provided several alternative ways, which will be studied in the next two subsections. Nevertheless, despite the different extensions developed within the last three decades, the basic Black-Scholes-Model is still the most accepted and widely used framework in financial industry and research.

1.3.2 Stochastic volatility models

The Black-Scholes Equation (1.4) laid the foundations for modern derivatives pricing. However, the assumptions made in the Black-Scholes model are known to be overly restrictive. In particular, the Black-Scholes model assume that the underlying asset price follows a geometric Brownian motion with a constant volatility. Many derivative pricing models have been developed subsequently that use more sophisticated stochastic process for the underlying asset which result in a better match to empirically observed details [85]. That it might make sense to see volatility as a random variable should be obvious to the most casual observer of equity markets. In order to be convinced, one only needs to remember the stock market crash of October 1987. Many empirical studies and economic

arguments after the stock market crash have shown that equity return distributions exhibit skewness and kurtosis and are always negatively correlated with implied volatility. These facts conflict with the normality assumption made in the Black-Scholes model.

Using such stochastic processes is often more straightforward than relaxing the Black-Scholes assumptions to allow for discrete time trading, transactions costs and other market imperfections. Examples of more realistic stochastic processes include: jump-diffusion [82], Lévy [94], stochastic volatility (SV) [49], SV jump-diffusion [12], and also combinations of those that exhibit SV as well as jumps in both the asset price and volatility [36].

One popular class of such alternative models is the stochastic volatility model [49], which treats the volatility as a stochastic process, rather than a constant, appearing in the Black-Scholes model. There are many stochastic volatility models proposed in the literature. These include the Stein & Stein model [93], the Schöbel & Zhu model [90], and the Heston model [49]. Due to the analytical tractability of the Heston model for European options, it is widely regarded as a flexible alternative to the Black-Scholes-Merton model, and thus, serves as benchmark against which other such models are commonly judged.

The Heston stochastic volatility model is formally defined as the system of stochastic differential equations given by

$$\begin{cases} dS_t = \mu S_t dt + \sqrt{v_t} S_t dW_1 \\ dv_t = \kappa(\theta - v_t)dt + \xi\sqrt{v_t} dW_2 \\ dW_1 dW_2 = \rho dt, \end{cases} \tag{1.7}$$

where S_t denotes the spot process at time, t, v_t the volatility, μ is the drift rate, κ the mean reversion speed for the variance, θ the mean reversion level for the variance, ξ the volatility of the variance, and $W_i, i = 1, 2$, two Brownian motions with correlation $\rho \in [-1, 1]$. The model for the volatility v_t is known in financial literature as the Cox-Ingersoll-Ross (CIR) process and in mathematical statistics as the Feller process [41]. The CIR process has many uses in mathematical finance, mostly due to its non-negativity and its mean reverting properties.

The parameters ρ, ξ, and κ, in the Heston model, capture observed features of the market and produce a wide range of distributions. For instance, the parameter ρ, the correlation between the log-returns and the asset volatility, affects the skewness of the distribution and hence the shape of the implied volatility surface; the parameter ξ, the

volatility of the variance, affects the kurtosis of the distribution; the mean reversion parameter κ can be interpreted as representing the degree of volatility clustering. This phenomenon has been observed repeatedly in the market; the occurrence of large price variations makes it more likely that further large price variations will follow.

In contrast to the Black-Scholes model where there is only one source of randomness (the underlying price to be hedged), the Heston model involves the volatility as another source of randomness. This also needs to be hedged in order to form a riskless portfolio. Thus, we set up a portfolio Π containing one option $V(S, v, t)$, a quantity $\triangle_1$ of the stock, and a quantity $\triangle_2$ of another option $V_1(S, v, t)$ that is used to hedge then volatility, i.e.,

$$\Pi = V - \triangle_1 S - \triangle_2 V_1.$$

The change of this portfolio in a time dt can thus be written as

$$d\Pi = dV - \triangle_1 dS - \triangle_2 dV_1. \tag{1.8}$$

Now, applying Itô's lemma to dV, we must differentiate with respect to the variables t, S, and v. Hence

$$dV = \underbrace{\left[\frac{\partial V}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \frac{1}{2}\xi^2 v\frac{\partial^2 V}{\partial v^2} + \rho\xi vS\frac{\partial^2 V}{\partial S \partial v}\right]}_{A} dt + \frac{\partial V}{\partial S}dS + \frac{\partial V}{\partial v}dv$$

Applying Itô's Lemma to dV_1 produces the identical result, but in V_1 as

$$dV_1 = \underbrace{\left[\frac{\partial V_1}{\partial t} + \frac{1}{2}vS^2\frac{\partial^2 V_1}{\partial S^2} + \frac{1}{2}\xi^2 v\frac{\partial^2 V_1}{\partial v^2} + \rho\xi vS\frac{\partial^2 V_1}{\partial S \partial v}\right]}_{B} dt + \frac{\partial V_1}{\partial S}dS + \frac{\partial V_1}{\partial v}dv,$$

Combining these two expressions, we can write the change in portfolio value $d\Pi$ as

$$d\Pi = Adt - \triangle_2 Bdt + \left(\frac{\partial V}{\partial S} - \triangle_2\frac{\partial V_1}{\partial S} - \triangle_1\right)dS + \left(\frac{\partial V}{\partial v} - \triangle_2\frac{\partial V_1}{\partial v}\right) \tag{1.9}$$

In order for the portfolio to be hedged against movements in the stock and against volatility, the last two terms in equation (1.9) involving dS and dv must be zero. This implies

that the hedge parameters must be

$$\begin{aligned} \triangle_1 &= \frac{\partial V}{\partial S} - \triangle_2 \frac{\partial V_1}{\partial S}, \\ \triangle_2 &= \frac{\partial V}{\partial v} / \frac{\partial V_1}{\partial v} \end{aligned} \tag{1.10}$$

Moreover, the portfolio must earn the risk free rate, r. Hence $d\Pi = r\Pi dt$ which yields

$$d\Pi = r(V - \triangle_1 S - \triangle_2 V_1)dt \tag{1.11}$$

Now with the values of $\triangle_1$ and $\triangle_1$ from equation (1.10), the change in value of the riskless portfolio is

$$d\Pi = (A - \triangle_2 B)dt = r(V - \triangle_1 S - \triangle_2 V_1)dt,$$

which yields $A - \triangle_2 B = r(V - \triangle_1 S - \triangle_2 V_1)$. Substituting for $\triangle_2$ and re-arranging, produces the equality

$$\frac{A - rV + rS\frac{\partial V}{\partial S}}{\frac{\partial V}{\partial v}} = \frac{B - rV_1 + rS\frac{\partial V_1}{\partial S}}{\frac{\partial V_1}{\partial v}}. \tag{1.12}$$

Therefore, we obtain

$$\begin{aligned} &\left(\frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \rho\xi vS\frac{\partial^2 V}{\partial S \partial v} + \frac{1}{2}\xi^2 v\frac{\partial^2 V}{\partial v^2} + \frac{\partial V}{\partial S} - rV + \frac{\partial V}{\partial t}\right) / \frac{\partial V}{\partial v} = \\ &\left(\frac{1}{2}vS^2\frac{\partial^2 V_1}{\partial S^2} + \rho\xi vS\frac{\partial^2 V_1}{\partial S \partial v} + \frac{1}{2}\xi^2 v\frac{\partial^2 V_1}{\partial v^2} + \frac{\partial V_1}{\partial S} - rV_1 + \frac{\partial V_1}{\partial t}\right) / \frac{\partial V_1}{\partial v}. \end{aligned} \tag{1.13}$$

Clearly, the left-hand side of equation (1.13) is a function of V only, and the right-hand side is a function of V_1 only. This implies that both sides can be written as a function $f(S, v, t)$ of S,v, and t. Following Heston [49], specify this function as $f(S, v, t) = -\kappa(\theta - v)$ and substitute this into the left-hand side of equation (1.13), the Heston PDE governing the price of financial derivatives is obtained as

$$\frac{1}{2}vS^2\frac{\partial^2 V}{\partial S^2} + \rho\xi vS\frac{\partial^2 V}{\partial S \partial v} + \frac{1}{2}\xi^2 v\frac{\partial^2 V}{\partial v^2} + rS\frac{\partial V}{\partial S} - rV + \Big(\kappa(\theta - v)\Big)\frac{\partial V}{\partial v} + \frac{\partial V}{\partial t} = 0 \tag{1.14}$$

The original Heston PDE differed slightly from equation (1.14) in that it included an additional $-\lambda(S, v, t)\frac{\partial V}{\partial v}$ term. The $\lambda(S, v, t)$ function represents the market price of volatility

risk and in many cases it is assumed to be proportional to v, giving, $\lambda(S, v, t) = \lambda v$, for a constant λ. However, the λ parameter in the model can be scaled out of the PDE by defining,

$$\kappa^\star = \kappa + \lambda,$$

$$\theta^\star = \frac{\kappa\theta}{\kappa + \lambda}.$$

By replacing κ and θ by their starred version in equation (1.14), the original Heston PDE is recovered. Thus, equation (1.14) together with appropriate boundary conditions will be discussed using our proposed method in chapter 5.

The solution to the European call option problem first appeared in Heston [49]. The solution can be written as,

$$V(x, y, t) = xP_1(s, v, t; T, K) - Ke^{-r(T-t)}P_2(s, v, t; T, K), \tag{1.15}$$

where,

$$s = ln(x),$$

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

$$f_j(s, y, t; T, \phi) = exp\Big(C_j(T-t, \phi) + vD_j(T-t, \phi) + i\phi s\Big),$$

$$C_j(\tau, \phi) = r\phi i\tau + \frac{a}{b}\left((b_j - \rho\beta\phi i + d_j)\tau - 2ln\left(\frac{1 - g_j e^{d_j\tau}}{1 - g_j}\right)\right),$$

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

$$g_j(\phi)\frac{b_j - \rho\beta\phi i + d_j}{b_j - \rho\beta\phi i - d_j},$$

$$d_j(\phi) = (\sqrt{\rho\beta\phi i - b_j)^2 - \beta^2(2u_j\phi i - \phi^2)},$$

for $j = 1, 2, u_1 = 1/2, u_2 = -1/2, a = \kappa\theta, b_1 = \kappa - \rho\beta, b_2 = \kappa$ and i denoting the imaginary unit.

1.3.3 Local volatility model

Given the computational complexity of stochastic volatility models and the difficulty of fitting parameters to the current prices of vanilla options, practitioners sought a simpler way of pricing exotic options consistent with the volatility skew. One popular class of such simpler models is the local volatility model, which treats volatility as a function of both the current asset level S_t and of time t rather than a constant appearing in the Black-Scholes model . As such, a local volatility model is a generalization of the Black-Scholes model.

Under the local volatility model, the underlying S_t, as a function of time, is assumed to follow a diffusion process:

$$dS = \mu S dt + \sigma(S, t) dW \tag{1.16}$$

where the constant $\mu \geq 0$ is the drift rate and the deterministic function $\sigma(S, t)$ represents the local volatility.

Application of Itôs Lemma together with risk neutrality to (1.16), gives rise to a partial differential equation for functions of the stock price which is a straightforward generalization of Black-Scholes. In particular, the pseudo probability densities $\varphi(K, T) = \frac{\partial^2 C}{\partial K^2}$ must satisfy the Fokker-Planck equation. This leads to the following equation for the undiscounted option price C in terms of the strike price K:

$$\frac{\partial C}{\partial T} = \frac{\sigma^2 K^2}{2} \frac{\partial^2 C}{\partial K^2} + r(T) \left(C - K \frac{\partial C}{\partial K} \right) \tag{1.17}$$

where $r(t)$ is the risk-free rate.

It is known that under the local volatility model, given a complete set of European option prices for all strikes and expirations, local volatilities $\sigma(K, T)$ can be extracted analytically from these option prices utilizing the well known Dupire formula [37] as

$$\sigma(K, T) = \sqrt{\left(\frac{\partial C}{\partial T} + rK \frac{\partial C}{\partial K} \right) / \left(\frac{K^2}{2} \frac{\partial^2 C}{\partial K^2} \right)} \tag{1.18}$$

with $C(S, t, K, T)$ being the price of a European call with strike K and expiry T.

Local volatility models have a number of attractive features. Apart from the fact that they are easy to calibrate due to only one source of randomness, they are also useful in any option market in which the volatility of the underlying is predominantly a function

of the level of the underlying. Additionally, time-invariant local volatilities are claimed to be able to provide the best average hedge for equity index options. However, local volatility models have some drawbacks. For instances, successful evaluation of volatility using Dupire's general non-parametric approach requires random selection of the input implied volatility surface. Furthermore, local volatility models fail to price accurately some financial derivatives depending specifically on the random nature of volatility, such as cliquet options or forward start options, because in the local volatility models the volatility is only a deterministic function.

1.4 Structure of Thesis

In Chapter 2, we present an overview of American option pricing problem. We first introduce American option pricing problems under different formulations: free boundary problem and linear complementarity problem. Then, we give an overview of most recent related literature and discuss the current state of research in PDE-based numerical methods with emphasis on the finite element method (FEM). In particular, we apply the finite element approach to American option problems. It is the basis for the pricing methods proposed in this thesis.

Chapter 3 presents an inverse approach for pricing American options based on a combination of the finite element method and an optimization algorithm. This approach, called the inverse finite element method (iFEM), opens up a new family of valuation methods. For their formulations, the nonlinearity of the option problem is addressed directly without linearization. The relatively simple American put option is used to explore performance of the formulation against the direct finite element method (dFEM). A critical performance analysis of the iFEM and dFEM is carried out against some benchmark solutions and experimental results on accuracy-efficiency trade-off are presented. We also provide the convergence analysis of the adopted iterative scheme. Insights on the applicability and reliability of the iFEM in the field of financial engineering are also provided.

In Chapter 4, we apply the inverse finite element method to option problems under linear complementarity formulations. The principal problem about pricing American options on one asset is shown under different mathematical formulations: differential complementarity problem and variational inequalities. We then adopt the inverse isotherm finite

element proposed by Alexandrou [3] to inversely determine the unknown free boundary by over-specifying the boundary where such value is measured. The solution accuracy is examined with respect to various element shape functions.

In Chapter 5, we adapt the hybrid method, a combination of the inverse finite element and the finite-difference approach, to the pricing of American options written on the Heston model. The method eliminates the time derivatives and the nonlinearity of the pricing problem is successfully dealt with without linearization. Hence, the solution obtained corresponds to that of the original pricing model. We establish the convergence by reformulating the discretized problem in variational form and study the approximation of the option price. Interesting numerical results are provided to illustrate the effects of time-dependent volatility on the prices of American options as well as their optimal exercise boundaries.

The thesis ends with some concluding remarks about the main results and some future research directions in Chapter 6.

Chapter 2

Classical methods for pricing American options

This chapter presents an overview of American option pricing problems. Section 2.1 introduces American option problems in Black-Scholes setting under different formulations: free boundary problem and linear complementarity problems. Section 2.2 gives an overview of most recent related literature and discusses the current state of research in approximation methods for pricing American options with emphasis on PDE-based numerical method. Section 2.3 revisits the finite element approach for pricing American options. It is the basis for the pricing methods proposed in this thesis.

2.1 The American option pricing problems

The pricing of American option is a long standing problem in mathematical finance. Unlike the European options, American options can be exercised at any time until maturity. As a consequence, the holder does not know when to exercise the right a *priori* as a function of the time. It is the additional flexibility of the early exercise right that makes American options worth more than their European counterparts. On the other hand, this important feature makes the valuation process of American options more challenging because the unknown boundary of the pricing domain varies with time. Finding this free boundary as part of the solution to the problem makes the pricing of an American option difficult.

One of the first investigations into the American put option was by McKean [78], who formulated the problem in terms of a free boundary, similar to those seen in melting ice

or dam problems. This allows the option price to be written explicitly in terms of the free boundary, equivalent to the optimal stopping boundary. However, the hedging arguments for why the American option can be formulated in such a way were not laid down until the work of Bensoussan [13] and Karatzas [66]. In the following, we shall present American option problems in Black-Scholes setting under different formulations.

2.1.1 Free boundary problem formulation

Under the Black-Scholes model (BSM), the option value depends on the underlying asset price S, the current time t, in addition to other constant parameters: the exercise price K, the volatility rate σ, the risk-free interest rate r. Let $P(S,t)$ be the option value associated with asset price S and time t. Merton [80] has shown that the price $P(S,t)$ of an American put option, satisfies the following partial differential equation (PDE):

$$\frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP = 0. \tag{2.1}$$

To understand the free boundary imposed by the American option, arbitrage arguments must be used. In Figure 2.1, the value of a typical European put is seen to be lower than that of the payoff function $(K-S)$ for some range of S. If this were to be the price of the American put option P, then $P(S,t) < K - S$ in this range. If the option is bought, one would simply exercise it making an instant risk-free profit of $K-S-P$ (since $P < K-S$). Then, by arbitrage, the option price would move such that an instantaneous profit could no longer be made. Consequently, the following constraint must hold for the American put options;

$$P(S,t) \geq \max(K-S, 0). \tag{2.2}$$

Now, assume that there exists some point in S, say S_f, below which it is optimal to exercise, but not so above. For the American put, the holder will exercise in the region $S < S_f$. In this region, the return on a bank deposit is more than if the option is held. Clearly, the solution in the region $S < S_f$ does not need to be calculated since we know that the option is exercised and therefore $P = K - S$.

In the region $S > S_f$, the BSM equation must still hold, but another condition is needed in order to close the problem: the value of S_f must be chosen so as to maximize the value

of the option. This is found by examining the gradient of P at the free boundary. We can show that $\frac{\partial P}{\partial S}(S_f(t),t) = -1$, i.e., the function P runs smoothly into the payoff function (see Figure 2.2). If $\frac{\partial P}{\partial S}(S_f(t),t) < -1$, then $P < K - S$ for some region close to S_f (we have already discussed earlier how this would not be possible). For the case $\frac{\partial P}{\partial S}(S_f(t),t) > -1$, consider the strategy adopted by the holder on when to exercise. A holder would wish to allow the asset price drop as low as possible before exercising. Therefore, if $\frac{\partial P}{\partial S}(S_f(t),t) > -1$, the value of the option near S_f can be increased by taking a smaller value of S_f. We conclude, thus, that the only possibility is that $\frac{\partial P}{\partial S}(S_f(t),t) = -1$. Assuming that the

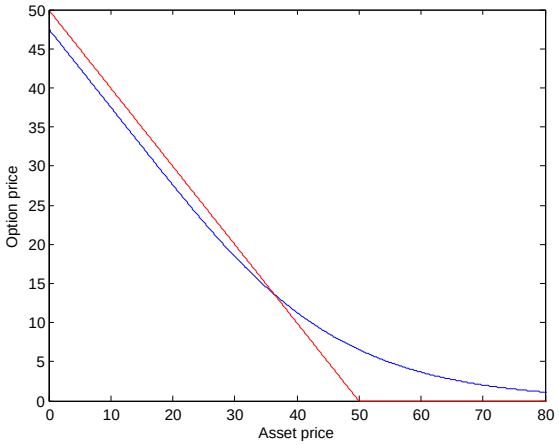

Figure 2.1: **The European put option**.
The payoff (solid red line) alongside the European put (solid blue line) with $\sigma = 0.4$, K = 50, T = 1, and r = 0.05.

option is priced under the BSM framework without continuous dividends, in the region $0 \leq S \leq S_f$ we therefore have the following

$$P = K - S, \quad \frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP < 0. \tag{2.3}$$

In the other region $S > S_f$ we have

$$P > K - S, \quad \frac{\partial P}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS\frac{\partial P}{\partial S} - rP = 0. \tag{2.4}$$

The boundary conditions at the free boundary are that P and its slope are continuous:

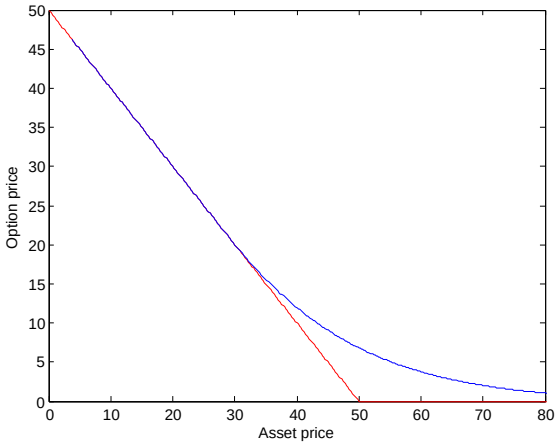

Figure 2.2: **The American put option**.
The payoff (solid red line) alongside the American put (solid blue line) with $\sigma = 0.4$, K = 50, T = 1, and r = 0.05.

$$\begin{aligned} P(S_f(t), t) &= K - S_f(t), \\ \frac{\partial P}{\partial S}(S_f(t), t) &= -1. \end{aligned} \tag{2.5}$$

It should be noted that the Dirichlet and Neumann boundary conditions in (2.5) are independent. While the Dirichlet condition is the payoff function, the Neumann condition is the result of the no-arbitrage assumption. In particular, the Neumann boundary condition indicates that the option price's derivative $\frac{\partial P}{\partial S}$ is continuous at crossing the optimal exercise boundary. This expresses the principle of American option pricing, i.e. the American option value is maximized by an exercise strategy that makes the option value and option derivative continuous. These two conditions are often referred to as the smooth pasting conditions, and ensure that the early exercise of the put option will be optimal and self financing.

The final location of the free boundary must be found by asymptotic analysis of the problem in the limit as we approach maturity. For the American put, Kim [68] found the location of the free boundary value at expiry to be

$$S_f(T) = \min\left(K, \frac{r}{\delta}K\right), \tag{2.6}$$

where δ is the continuous dividend payment on the underlying asset, and for the American put with no dividend, $S_f(T) = K$.

For both European and American put options, there is a far-field boundary condition $\lim_{S\to\infty} P(S,t) = 0$. Additionally, the fact that the value of a put option must be equal to its payoff function sets up the terminal condition

$$P(S,T) = \max(K - S, 0), \tag{2.7}$$

where T is the expiration time of the option. To summarize, the Black-Scholes-Merton (BSM) model for pricing an American put option is given as

$$\mathcal{A}\begin{cases} \dfrac{\partial P}{\partial t} + \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 P}{\partial S^2} + rS\dfrac{\partial P}{\partial S} - rP = 0, \\ P(S,T) = \max(K - S, 0), \\ P(S_f(t),t) = K - S_f(t), \\ \dfrac{\partial P}{\partial S}(S_f(t),t) = -1, \\ \lim\limits_{S\to\infty} P(S,t) = 0, \end{cases} \tag{2.8}$$

where $\mathcal{A}$ is defined on $t \in [0,T]$, $S \in [S_f(t), \infty)$. Note that for each $t \in [0,T]$, there exists a stock price S for which early exercise before final time T is advantageous. This value defines a continuous curve $S_f(t)$ and it is *a priori* unknown, and hence, defines a free boundary. The presence of $S_f(t)$ makes this type of problem nonlinear.

Similar arguments as the case of an American put option lead to the following PDE system for an American call option

$$\mathcal{B}\begin{cases} \dfrac{\partial C}{\partial t} + \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 C}{\partial S^2} + rS\dfrac{\partial C}{\partial S} - rC = 0, \\ C(S,T) = \max(S - K, 0), \\ C(S_f(t),t) = S_f(t) - K, \\ \dfrac{\partial C}{\partial S}(S_f(t),t) = 1, \\ \lim\limits_{S\to 0} C(S,t) = 0, \end{cases} \tag{2.9}$$

where $\mathcal{B}$ is defined on $t \in [0,T]$, $S \in [0, S_f(t))$. Since it is easy to show that an American call option without dividend is equivalent to a counterpart European option, we focus on pricing the former using different models.

2.1.2 Variational inequality formulation

We have already formulated the American option in the previous subsection. It is now simple to formulate this as a parabolic variational inequality, and then after discretisation, as a linear complementary problem. The idea is to reformulate the problem such that the free boundary $S_f(t)$ does not show up explicitly. Suppose will let $\mathcal{L}_{BSM}$ be the standard BSM operator defined as

$$\mathcal{L}_{BSM} = \frac{\partial}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} + rS\frac{\partial}{\partial S} - rI, \tag{2.10}$$

where I is the identity element. Then, we can write PDE system (2.8) as

$$\mathcal{C}\begin{cases} \mathcal{L}_{BSM}P(S,t).(P(S,t) - G(S,t)) = 0, \\ \mathcal{L}_{BSM}P(S,t) \leq 0, \\ P(S,t) - G(S,t) \geq 0, \\ P(S,T) = G(S,T), \end{cases} \tag{2.11}$$

such that P and $\frac{\partial P}{\partial S}$ are continuous. Note here that $G(S,t) = \max(K - S, 0)$, and consequently its derivative is discontinuous at $S = K$. However, it can be shown that $P(S = K, t) > G(S = K, t)$ for any time before expiry so the constraint will never be applied there. This is not always the case, as for example in the case of the "cash-or-nothing" American option. Here, there is a discontinuous constraint, implying that the free boundary is known a priori, since it is always optimal to exercise at the strike price. The smooth-pasting condition therefore does not apply, since $G(S_f, t)$ and $\frac{\partial G(S_f,t)}{\partial S}$ are not continuous at the strike price.

Elliot and Ockendon [38] and Friedman [43] gave detailed accounts of the existence and uniqueness proof for variational inequalities of the form (2.11). This involves knowledge of abstract functional analysis; the problem is basically reduced to minimizing a convex set of functions.

2.2 Literature review

It has been widely acknowledged that the valuation of American options is mathematically intriguing. When compared to the European option, there is a free boundary associated

with the optimal time to exercise the holder's early exercise rights. Due to the nonlinearity originating from the early exercise policy of an American option, there has been a great deal of research over the last 30 years to find an effective solution method. The Zhu [113] work derives an analytical closed-form pricing formula in the form of a Taylor series expansion for an American option under the Black-Scholes model. While the emphasis of that work is to show the existence of an analytical solution, it is however not computationally appealing, as the formula involves two infinite sums of infinite double integrals, which take a formidable amount of time to evaluate.

Even with the effort made by Zhu [113], till now, approximation methods still remain popular among market practitioners as they are usually faster with acceptable accuracy. There are predominately three types of approximation methods for the valuation of American-style derivatives in the literature. These include analytical approximation methods, stochastic simulation methods and PDE-based numerical methods. In what follows, an overview of these methods with emphasis on PDE-based numerical methods will be reviewed

2.2.1 Analytical approximation methods

This is the category of analytically tractable approximations to American option pricing problems. Such methods can provide approximations of the value of American options. MacMillan [77] and Barone-Adesi and Whaley [11] developed quadratic approximations for the option price. These methods are not convergent, and have trouble pricing long-maturity options accurately. To correct this problem, Ju and Zhong [63] developed an approximation based on the method proposed in Barone-Adesi and Whaley [11]. While this improved method prices long-maturity options more accurately, it is not convergent. Geske and Johnson [45] viewed an American option as sequence of Bermuda options and propose an approximation of the option price consisting of an infinite sequence of cumulative normal functions. Bunch and Johnson [25] proposed a modified two-point Geske-Johnson approach. Zhu [116] developed Laplace transform method based on the pseudo-steady-state approximation.

As the name suggests, analytical approximation method involves approximation in the valuation processes. For instance, in Laplace transform method [116], the author relied on the pseudo-steady-state approximation used for the classical *Stefan* problem,

and assumed that the moving boundary $S_f(t)$ is nearly a constant function during the Laplace transform. Consequently, he derived a very elegant analytical approximation formula for the optimal exercise price of American options, S_f, as well as the option price. The fundamental drive for researchers to seek analytical approximation to the problem of pricing American options is mainly to simplify computational effort. However, there is the drawback of slightly larger errors, in comparison with other approximation methods. Analytical approximation methods have a place in trading practice as long as the errors are not too big and systematic.

2.2.2 Stochastic simulation methods

Stochastic simulation methods can be divided into two sub-groups: Monte Carlo Simulation (MCS) and lattice methods (the Binomial and the Trinomial methods). Monte-Carlo methods generally simulate the dynamic of the underlying asset with arbitrary numbers in order to calculate possible price movements based on the known market figures. This method was initially employed by Boyle [18] and also applied by Tilley [97]. Some of other recent works include [46–48, 75]. While this approach can be easily applied to value a European option, it is not easily applicable to valuing American options. The difficulty in using MCS lies within the particular implementation: the path simulation requires a forward algorithm, whereas the early exercise decision is done with a backwards procedure from the maturity date. This means, when a path is traced back, to determine the exercise boundary, it depends on perfect foresight, rather than an expectation used in backward algorithms of other methods. As a consequence, it will actually overestimate the true option price. Nevertheless, its convergence is generally independent of the number of underlying variables, and thus, it can easily be adapted to extended market models such as Jump Diffusion models, multiple assets, and exotic options.

The Binomial method was initially proposed by Cox et al. [32]. Supported by the central limit theorem, this approach is fast and easy to implement for the Black-Scholes model. Since this method is not stochastic in the sense that random numbers are used, the probabilities for the up and down movements are rather a consequence of the assumed market factors. Boyle [19] extended the basic Binomial model by additional factors to a multinomial method with several possibilities for the up and down movements of the asset price. Broadie and Detemple [24] suggested a modification called BBSR, which uses the

Black-Scholes formula at the last time step before maturity, combined with Richardsons extrapolation to compensate for the oscillatory convergence of the standard Binomial Method. A major advantage of this type of simulation is that it can easily be adapted to changing market factors, such as dividend payments and change of interest rates. Also, the computation time is reasonable for an acceptable accuracy.

In general, stochastic simulation methods are more popularly used among market practitioners than any other existing methods. Apart from the ease of implementation and adaptability to other exotic options, they are more closely linked with the stochastic assumption, and thus, has a direct financial interpretation than others.

2.2.3 PDE-based numerical methods

PDE-based numerical method is the third category of existing approaches for pricing American option. Among this category are the finite difference method (FDM) [21, 26, 54, 81, 95], the finite element method (FEM) [5, 6, 50] and the finite volume method (FVM) [42, 104]. The most popular approach in this category is probably the finite difference method (FDM). Brennan et al. [81] were first to solve the Black-Scholes PDE with early exercise directly, using a numerical technique. The technique, including a simple search algorithm to find the optimal exercise boundary, is analogous to solving the variational inequality problem, explained earlier in section 2.1. Brennan and his co-worker used an implicit scheme, however, the Crank-Nicolson scheme coupled with the PSOR scheme, explained later by Wu and Kwok [109] is far superior. Some justification for the algorithm is given by Jaillet et al. [60], who noted that the scheme only works because of the specific nature of the free boundary in American option problems.

Finite element method on the Black-Scholes PDE was initially used by Wilmott et al. [106] jointly with PSOR to solve the resulting system of inequalities. Allegretto et al. [5] used the direct finite element method (dFEM) for pricing American option, which was later extended to other pricing models by Achdou and Tchou [2]. Zhang [110] studied the American options valuation through an "adaptive" finite element method using a variational inequality formulation. Zvan et al. [121] suggested a mixed Finite Element-Finite Volume approach as well as a pure Finite Volume approach in Zvan et al. [122], combined with a penalty scheme, which involves an additional term to equalize the inequality.

The PDE-based numerical method can be further categorized into two types, according

to the different ways of dealing with the nonlinearity associated with American options: those in which the moving boundary is located implicitly through the so-called Linear Complementarity (LC) formulation, and those in which the moving boundary is found explicitly in the process of solving the governing PDEs. Pricing American options under the former subcategory has been well studied in the literature. Huang and Pang [53] solved the American option pricing problem based on the LC formulation by using a second order upwind finite difference scheme. Based on the linear complementarity formulation, various numerical algorithms can be adopted to solve the final nonlinear system of equations. Koulisianis and Papatheodorou [69] compared different algorithms including Projected-SOR algorithm [1, 91], Operator splitting [58, 59], Body-Fitted Coordinated Method [105] and their own "moving index" (MI). They demonstrated that the MI method was superior to the other methods.

An alternative approach is to convert the linear complementarity problem (LCP) to a non-linear PDE by the penalty method (see [14, 15] for general theory). A penalty term is introduced in the PDE to approximate the action of early exercise constraints (see for example [79]). Penalty method and front fixing method together with the finite difference method are discussed in [83, 86, 109]. Forsyth and Vetzal [42] discretized their computational domain with the finite volume method (FVM) and implemented a nonlinear penalty iteration scheme to find the price of an American put. With a timestep "selector", they can achieve quadratic convergence. Wang et al. [103] used a power penalty function approach to solve the LC problem for the price of American options; they were able to achieve a relatively high accuracy with a small penalty parameter. Other recent studies proposed to solve penalized option problem include [50, 79, 111, 112] and the references therein. Applying the standard implicit Euler method to the discretized problem, the resulting non-linear system of algebraic equations can be solved by using iteration schemes such as the Newton iteration scheme.

The methods in the second subcategory usually adopt some sort of iterative methods to solve the discretized nonlinear system. For example, Elliott and Ockendon [38] discretized a general class of free boundary problems by the finite element method, and then solved the resulting system of nonlinear algebraic equations with an iterative approach. Topper [100], in a technical note, provides a generalized study of pricing an American option by using finite element-based residual formulation. The resulting nonlinear system is solved at each

time-step by the Newton iteration method.

Extensions of the Black Scholes Model try to make the original model fit more closely to the behaviour of certain markets, e.g. where the so-called "volatility smile" is observed. Modifications are usually applied to the continuity assumptions of the interest rate and volatility in the presumed asset price movements. PDE-based numerical methods for pricing American option with stochastic volatility models have focused on finite difference and finite element methods (see for examples [23, 29, 57, 58, 84, 121]). The alternating direction implicit (ADI) schemes are good alternative methods. For example, Hout & Foulon [51] investigate four splitting schemes of the ADI type for solving the PDE Heston equation: the Douglas scheme, the CraighSneyd scheme, the Modified CraighSneyd scheme and the HundsdorferVerwer scheme. Ikonen & Toivanen [58] propose a componentwise splitting method for pricing American options in the Heston model. In Hout & Foulon [51], the splitting method of Ikonen & Toivanen [58] is combined with ADI schemes in order to obtain more efficient numerical results.

Although, the PDE-based numerical methods are gradually becoming more popular among the market practitioners because of their computational advantages, there still exist some difficulties in using these methods. Quite often, a challenge for adopting these methods is to prove the convergency of the adopted schemes and many of the hitherto emphasized methods require some sort of linearization in addition to an intensive computation before a solution of reasonable accuracy can be reached (cf.[114]). Moreover, apart from the lack of theoretical convergence proof of the adopted schemes in some cases, the sequential nature of the algorithm makes parallelization difficult. Thus, this thesis is motivated to explore the use of a "convergence-proved" approach for the valuation of American options under different frameworks and formulations.

Among the PDE-based methods is the finite element method, which is widely documented and used, at least outside the financial community, for solving systems of partial differential equation. It is based on using variational methods and/or weak formulations. Instead of using the finite differences to approximate the derivatives, the FEM converts the PDE into an integrated form. The use of the integrated form is advantageous since it provides a reasonable treatment of Neumann boundary conditions which are very common when estimating the behavior of an option as the underlying price goes to infinity. Additionally, FEM has ability to handle complex geometries with relative ease, and a solution

for the entire domain is computed, instead of isolated nodes as in the case of FDM. However, these advantages come at the cost of a more complicated method compared to the FDM. Unlike the finite difference algorithm that can be implemented in a straightforward manner, finite element method exploit the weak formulation and replaces the underlying function space by an appropriate finite dimensional subspace. For ease of reading, details implementation of FEM is deferred to Section 2.3.

Finite element method (FEM), just like the other PDE-based methods for pricing American options, whilst appealing, suffers from the slow rate of convergence. In particular, the method requires intensive computation before a solution of reasonable accuracy can be obtained. Moreover, the solution near maturity can be of singular behavior for American options [10]. Naturally, it is difficult for the traditional FEM to calculate the option price accurately in the neighborhood of maturity. Nevertheless, FEM can be made more efficient by considering its inverse formulation (inverse finite element method), which is the main focus of this thesis.

In the next section, we present the valuation of an American option using finite element approach. The purpose is to familiarize readers with the finite element technique which is the basis for the methods discussed in this thesis.

2.3 Finite element method for pricing American options

The steps required to numerically solve the option pricing system is similar to other elliptic partial differential systems. Some projects are publicly available and the steps are solved automatically in the computer [74]. However, the most common commercial FEM packages are focused on the following steps

1. Domain discretization with finite elements (FEs),

2. Numerical integration of all FEs. In this step, a local matrix, k_e and a local load vector f_e are computed for each finite element,

3. Construction of the global sparse matrix, K from local matrices, k_e and global load vector, F from local load vectors f_e,

4. Application of boundary conditions,

5. Solution of the linear/nonlinear equation system formed previously, $KU = F$, where U represents the vector of the nodal solution.

2.3.1 Transformation and localization of domain

First, we recall the variational inequality (2.11) as the pricing problem. In view of the fact that $\mathcal{L}_{BSM}$ defined in equation (2.10) is indeed a degenerate parabolic differential operator, one usually introduces the so-called *log-prices* variable x replacing the stock price S in terms of the transformed stock variable $x = \log(S)$. This results in the disappearance of the variable coefficient S. We further change to *time-to-maturity* $\tau = T - t$, to obtain a forward parabolic problem. Also, the free boundary S_f in equation (2.11) is transformed to a function $x_f = x_f(\tau)$. Thus, setting $P(S,t) := y(x,\tau)$, problem (2.11) becomes

$$\mathcal{D}\begin{cases} \mathcal{L}y(x,\tau).(y(x,\tau) - g(x)) = 0, \\ \mathcal{L}y(x,\tau) \geq 0, \\ y(x,\tau) - g(x) \geq 0, \\ y(x,0) = g(x), \end{cases} \tag{2.12}$$

where $\mathcal{D}$ is defined on $\tau \in [0,T]$, $x \in (-\infty,\infty)$. The transformed payoff function is obtained as $g(x) = \max(K - e^x, 0)$ and the differential operator $\mathcal{L}$ defined as

$$\mathcal{L} = \frac{\partial}{\partial\tau} - \frac{1}{2}\sigma^2\frac{\partial^2}{\partial x^2} - (r - \frac{1}{2}\sigma^2)\frac{\partial}{\partial x} + rI,$$

where I is the identity operator.

The question of the existence and uniqueness of a solution to the problem (2.12) has been addressed in [14]. It should be pointed out that for each $\tau \in [0,T]$, there exists the optimal exercise price, $x_f(\tau)$ such that for all $x \leq x_f(\tau)$, the value of the American put option is the value of immediate exercise, i.e., $y(x,\tau) = g(x)$, while for $x > x_f(\tau)$ the value exceeds the immediate exercise. Note that the linear complementarity problem (2.12) do not explicitly involves the free boundary, $x_f(\tau)$.

In order to enable numerical computations, we truncate the infinite domain $\Omega = (-\infty,+\infty)$ in (2.12) to a finite domain $\Omega_k = [x^-,x^+]$, where x^+ and x^- are the maximum and minimum values of the dimensionless asset price, respectively. In practice, x^+ should be sufficiently large to eliminate the boundary effect. However, based on Wilmott

et al.'s estimate (see [106]), the upper bound of the underlying price S should be three or four times the strike price, it is reasonable to set $x^+ = \ln 5$.

2.3.2 Variational Formulation

Here, we derive an appropriate variational formulation of problem (2.12). The variational formulation of a PDE is a mathematical treatment for converting the strong formulation into a weak formulation, which permits the approximation in elements or sub-domains. In the classical setting of a PDE, one considers the solution of the desired function and its derivatives at each point of the domain. However, in the variational formulation or weak form, the solution of a PDE needs not be smooth. Rather, the solution and its derivatives only need to be square integrable. These requirements might require a modified function space when searching for a solution in the weak sense.

Let $L_2(\Omega)$ be the usual space of Lebesgue measurable and square integrable functions on Ω and denote by $H_0^1(\Omega)$ the Sobolev space of first-order weak derivatives. We define $\mathcal{K}_g \subset H_0^1(\Omega)$ as

$$\mathcal{K}_g := \{v \in H_1(\Omega) : v \geq g,\ v(x) = g(x), \forall x \in \Gamma\}, \tag{2.13}$$

where $\Gamma = \partial\Omega$ denotes the boundary, and the inequality sign means to hold pointwise $\forall$ $x \in \Omega$. Let $v \in \mathcal{K}_g$ be any test function and $y \in \mathcal{K}_g$ be a solution of problem (2.12). We multiply the second inequality in (2.12) by $v - g$ (which does not change the sign) and integrate over Ω, yielding

$$\int_\Omega \mathcal{L}y(\varphi - g)\ dx \geq 0 \tag{2.14}$$

Subtraction of the first equation in (2.12), integrated over Ω , that is $\int_\Omega \mathcal{L}y(y - g)\ dx = 0$, yields

$$\int_\Omega \mathcal{L}y(v - y)\ dx \geq 0, \tag{2.15}$$

thereby eliminating g. By the definition of the differential operator $\mathcal{L}$ and since v and y cancel out on the boundary Γ, integrating (2.15) by parts gives the formulation as

variational inequality problem

$$\begin{cases} \text{find } y \in \mathcal{K}_g, \text{ such that } \forall\ v \in \mathcal{K}_g \text{ and } 0 \leq \tau \leq T, \\ \int\limits_{\Omega} \left(\frac{\partial y}{\partial \tau}(v-y) + \frac{1}{2}\sigma^2 \frac{\partial y}{\partial x}(\frac{\partial v}{\partial x} - \frac{\partial y}{\partial x}) - \big(r - \frac{1}{2}\sigma^2\big)\frac{\partial y}{\partial x}(v-y) + ry(v-y) \right) dx \geq 0. \end{cases} \tag{2.16}$$

Conceptually, variational inequality relaxes the regularity conditions on the option price. As a consequence, one would get a *weak* solution in contrary to the *classical* or *strong* case. From treatments on variational inequalities, see e.g., [38, 70, 110], the problem (2.16) has a unique solution by a generalized theorem Lax-Milgram. More details on the variational formulation of parabolic PDEs associated with diffusion processes can be found in [28, 61, 72], with the relevant functional analytic background.

2.3.3 Finite element discretization

To proceed with the FEM approximation, we partition the domain of interest, $\Omega \subset \mathbb{R}^d_x$, into a set of non-overlapping polyhedrons. A partition is called admissible, if, for two arbitrary sub-domains F_i and F_j, exactly one of the following four cases is true:

1. $F_i = F_j$,
2. $\bar{F}_i \cap \bar{F}_j$ forms an entire edge of both F_i and F_j ,
3. $\bar{F}_i \cap \bar{F}_j$ is a vertex (knot) of the partition,
4. $\bar{F}_i \cap \bar{F}_j = \emptyset$,

where $F_i, i = 1, 2 \ldots$ are sets of non-overlapping polyhedrons and $\bar{F}_i, i = 1, 2 \ldots$ are sets of the complement of $F_i, i = 1, 2 \ldots$. For one dimensional problem such as option problem with single underlying asset, F_i is a discretized domain $[a, b]$, where $b > a$. In the case of higher dimensional problems, F_i can be triangular or rectangular elements.

Now, let us recall the variational inequality formulation of American options (2.16), i.e.,

$$\int\limits_{\Omega} \left(\frac{\partial y}{\partial \tau}(v-y) + \frac{1}{2}\sigma^2 \frac{\partial y}{\partial x}(\frac{\partial v}{\partial x} - \frac{\partial y}{\partial x}) + \big(\frac{1}{2}\sigma^2 - r\big)\frac{\partial y}{\partial x}(v-y) + ry(v-y) \right) dx \geq 0 \tag{2.17}$$

Since we consider one underlying asset for the option, the problem will be solved by using the finite element method on a *one*-dimensional discretized domain $\Omega_k = [x^-, x^+]$, i.e. an interval with $x^+ > x^-$. We choose equidistant mesh sizes h. However, for higher dimensional problems, the generation of a suitable grid is usually not an easy task.

By applying the Ritz-Galerkin method, a piecewise polynomial finite-element approximation is sought to approximate the variables $y = y(x,\tau)$ and $v = v(x,\tau)$ as

$$y = \sum_{i=0}^{N} y_i(\tau)\varphi_i(x),\ v = \sum_{i=0}^{N} v_i(\tau)\varphi_i(x), \tag{2.18}$$

with the finite elements $\varphi_0(x), \ldots, \varphi_N(x)$ and weights $y_i, v_i, i = 0, \ldots, N$. By abbreviating $\varphi_i' := \frac{\partial \varphi_i}{\partial x}$, the variational inequality (2.17) can now be discretized on the x-axis as follows

$$\begin{aligned}\int_\Omega \Bigg(&\sum_{i=0}^{N} \frac{\partial y_i}{\partial \tau}\varphi_i \sum_{j=0}^{N}(v_j - y_j)\varphi_j + \frac{1}{2}\sigma^2 \sum_{i=0}^{N} y_i\varphi_i' \sum_{j=0}^{N}(v_j - y_j)\varphi_j' \\ &+ (\frac{1}{2}\sigma^2 - r)\sum_{i=0}^{N} y_i\varphi_i' \sum_{j=0}^{N}(v_j - y_j)\varphi_j + r\sum_{i=0}^{N} y_i\varphi_i \sum_{j=0}^{N}(v_j - y_j)\varphi_j \Bigg)\, dx \geq 0.\end{aligned} \tag{2.19}$$

The discretized variational problem (2.19) can be written equivalently as

$$\begin{aligned}&\sum_{i,j=0}^{N} \frac{\partial y_i}{\partial \tau}(v_j - y_j) \underbrace{\int_\Omega \varphi_i\varphi_j\, dx}_{=:a_{ij}} + \\ &\sum_{i,j=0}^{N} y_i(v_j - y_j) \underbrace{\left(\int_\Omega \left(\frac{1}{2}\sigma^2\varphi_i'\varphi_j' + (\frac{1}{2}\sigma^2 - r)\varphi_i\varphi_j' + r\varphi_i\varphi_j \right) dx \right)}_{=:b_{ij}} \geq 0\end{aligned} \tag{2.20}$$

Using vector notation, we write (2.20) as

$$\begin{aligned}&\left(\frac{\partial y}{\partial \tau}\right)^T A(v - y) + y^T B(v - y) \geq 0 \quad \Leftrightarrow \\ &(v - y)^T \left[A\frac{\partial y}{\partial \tau} + By \right] \geq 0\end{aligned} \tag{2.21}$$

with

$$\begin{aligned} y &= (y_0, \ldots, y_N)^T, \quad A = (a_{ij})\ i,j = 0, ..., N, \\ v &= (v_0, \ldots, v_N)^T, \quad B = (b_{ij})\ i,j = 0, ..., N, \end{aligned}$$

and a_{ij}, b_{ij} defined as above.

So far, only the spatial axis has been discretized. A common approach, based on the ideas of Kantorovitch [65], is to separate time and spatial location. For this problem, the spatial direction x is discretized by the finite elements, while we discretize the time derivative in (2.17) with a standard finite difference scheme. We decompose the time interval into equidistant points $0 =: \tau^0 < \tau^1 \cdots < \tau^M := T$ with time step $\triangle\tau = \tau^k - \tau^{k-1}$. We approximate

$$\frac{\partial y}{\partial \tau} = \frac{y^{k+1} - y^k}{\triangle \tau},$$

yielding the familiar θ-scheme for $\theta \in [0, 1]$

$$\begin{aligned}
&(v^{k+1} - y^{k+1})^T \left(A\frac{y^{k+1} - y^k}{\triangle\tau} + \theta B y^{k+1} + (1-\theta) B y^k \right) \geq 0 \quad \Leftrightarrow \\
&(v^{k+1} - y^{k+1})^T \Bigg(\underbrace{(A + \theta B \triangle\tau)}_{=:C} y^{k+1} + \underbrace{\underbrace{(B(1-\theta)\triangle\tau - A)}_{=:D} y^k}_{=:-b} \Bigg) \geq 0 \quad \Leftrightarrow \\
&(v^{k+1} - y^{k+1})^T \big(C y^{k+1} - b \big) \geq 0
\end{aligned} \tag{2.22}$$

We recall that we would like to solve equation (2.22) forward in time. Thus, the terms indexed by k are known, while the terms with index $k+1$ are to be determined. Choosing $\theta = 1$ yields an explicit Euler scheme and $\theta = 0$ yields an implicit Euler scheme, and $\theta = \frac{1}{2}$ corresponds to the CrankNicolson scheme. Naturally, here we select the latter due to its consistency and convergence error rate $\mathcal{O}(\triangle\tau^2)$.

To complete the approximation of (2.22) as a discretized version of the variational inequality (2.17), the former equation needs to satisfy the constraints

$$v^k \geq g^k \quad \text{and} \quad y^k \geq g^k, \quad \forall\ k = 0, \ldots, M.$$

Thus, the finite element discretization of equation (2.16) is as follows

$$\begin{cases} \text{find } y = y^{k+1}, \text{ such that for all } v \geq g \\ (v - y)^T (Cy - b) \geq 0, \text{ and } y \geq g. \end{cases} \tag{2.23}$$

Proposition 2.3.1. *The problem (2.23) is equivalent to the discrete linear complementary*

problem

$$\begin{cases} \textit{find } y \in \mathbb{R} \textit{ satisfying} \\ (y-g)^T(Cy-b) = 0, \quad Cy \geq b, \quad y \geq g. \end{cases} \tag{2.24}$$

Proof. Recall the transpose property $(A \pm B)^T = A^T \pm B^T$. Thus, problem (2.23) can further be simplified by considering

$$\begin{aligned} &(v^T - y^T)(Cy-b) \geq 0 \quad \Leftrightarrow \\ &v^T(Cy-b) \geq y^T(Cy-b) \quad \Leftrightarrow \ \forall\, v \geq g. \end{aligned} \tag{2.25}$$

Clearly, $Cy-b \geq 0$. If it were negative in the i-th component, the inequality would not hold for i going to infinity-a contradiction. Using this and $y \geq g$, one obtains $(y-g)^T(Cy-b) \geq 0$. On the other hand, substituting $v = g$ in (2.23) yields $(y-g)^T(Cy-b) \leq 0$. Combining the two inequalities leads to $(y-g)^T(Cy-b) = 0$. □

The next step in our solution procedure is to evaluate the element matrices A and B, which are needed to assemble the stiffness matrices C and load vector b in (2.24). To do this, we need a suitable shape function φ. Since we are only interested in the FEM procedure rather than the solution accuracy, we use the simplest shape function-linear shape function (hat function). Details definition of a hat function procedures are provided in appendix A.1.

A variety of algorithms have been proposed for solving finite dimensional linear complementarity problems of the form (2.24) in the literature. These include matrix splitting methods such as the projected successive overrelaxation (PSOR) method, pivoting method and interior point method, see, e.g. [31, 53, 108]. The most popular method in the context of American options pricing is the PSOR algorithm [107]. In what follows, we briefly review the PSOR algorithm and point out some of its weaknesses despite its popularity.

2.3.4 The projected successive overrelaxation (PSOR) method

The finite element discretization of problem (2.16) leads to a succession of variational inequalities to be solved at each time step. Such inequalities can be expressed in the

general linear complementarity (vector) form

$$\mathcal{A}\mathcal{B} = 0, \quad \mathcal{A} \geq 0, \quad \mathcal{B} \geq 0. \tag{2.26}$$

where $\mathcal{A}\mathcal{B}$ is the product of two vectors such that at least one element in $\mathcal{A}$ and $\mathcal{B}$ must be zero at every position.

Several methods have been proposed to solve such a problem when

$$\mathcal{A} = Cy - b, \quad \text{and} \quad \mathcal{B} = y$$

where y is the n-vector to be determined, b is a known n-vector and C is an $n \times n$ matrix. Generic methods for solving such a system of equations which actually contained a linear system of inequalities are not geared toward finite element or finite difference formulation, so by adapting the successive overrelaxation (SOR) method Cryer [33] was able to exploit the sparseness of matrix C.

However, to be able to use an iterative scheme on (2.24) with $\mathcal{A} = Cy - b$ and $\mathcal{B} = y - g$, we need to make some modifications to SOR algorithm. First we rewrite (2.24) as

$$\begin{aligned} & y - C^{-1}b \geq 0, \quad y - g \geq 0, \quad (y-g)^T(Cy - b) = 0 \quad \Leftrightarrow \\ & \min_y \{y - C^{-1}b, y - g\} = 0 \qquad\qquad \Leftrightarrow \\ & y = \max\{y - C^{-1}b, g\} \end{aligned} \tag{2.27}$$

For the solution of the latter equivalence, Cryer [33] suggested the PSOR method. The idea is to use the SOR algorithm [33, 107], and to include the maximization constraint as in (2.27).

Let y^k and g denote the vectors as follows

$$y^k = \begin{pmatrix} x_1^k \\ \vdots \\ x_{N-1}^k \end{pmatrix}, \qquad g^k = \begin{pmatrix} g_1^k \\ \vdots \\ g_{N-1}^k \end{pmatrix}, \tag{2.28}$$

the PSOR algorithm for (2.24) is given as

$$\begin{cases} \text{for } i = 1, \ldots, N-1: \\ x_i^{k+1} = \dfrac{1}{c_{ii}} \left(- \sum_{i<j} c_{ji} y_i^{(k+1)} - \sum_{i>j} c_{ji} y^{(k)} + b_i \right) \\ y_i^{(k+1)} = \max\{ y^{(k)} + \omega \left(x_i^{k+1} - y_i^{(k)} \right), g \}, \end{cases} \tag{2.29}$$

where ω is the relaxation parameter. It should be noted that the values at the boundary are not included, since we do not need to iterate over them. With (2.29), the problem is reduced to maximizing the vector y^{k+1} over the solution space. This can be interpreted in finance terms as choosing the free boundary in order to maximize the value of the option.

The subsequent theorem establish the existence and uniqueness of PSOR algorithm (2.29).

Theorem 4. *(**Cryer**) Let C be symmetric and positive definite, $\omega \in (1, 2)$. Then, the PSOR method in (2.29) converges towards a unique solution.*

Proof. The details of proof can be referred to [102]. □

Next, we present some numerical results of the finite element method. We consider an American put option on non-dividend paying share with strike price $K = 100$ and expiry date $T = 1$. The risk-less interest rate is $r = 0.05$ and the volatility $\sigma = 0.25$. We compute the price of the American option at $(\tau = T)$ as well as the optimal exercise price $S_f(t)$. The results are shown in Figure 2.3 and 2.4, respectively. In addition to the American option price in Figure 2.3, we also plot the option price of the corresponding European option, which is, due to single exercise right, lower than the price of an American option.

As shown in Figure 2.4, the optimal exercise price, $S_f(t)$ decreases monotonically with time to expiry, τ. As τ approaches the expiration time 0 of the option, $S_f(t)$ rises sharply towards the strike price, $K = 100$. At $\tau = 0$, $S_f(t)$ is equivalent to the exercise price, K. This behaviour is, of course, as expected. Figure 2.4 also shows that for time near expiry, the rate of change of $S_f(t)$ is much larger than when the option contract is far from expiry. As pointed out by Zhu [113], the large rate of change of $S_f(t)$ in the neighborhood of expiration time leads to difficulties for most numerical algorithms to effectively deal with the singular behaviour of the $S_f(t)$ near $\tau = 0$.

Before we end this section, it is important to make some remarks on FEM results. The

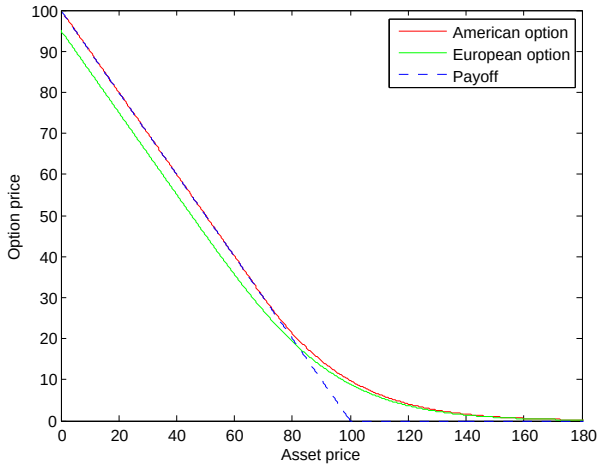

Figure 2.3: American option price

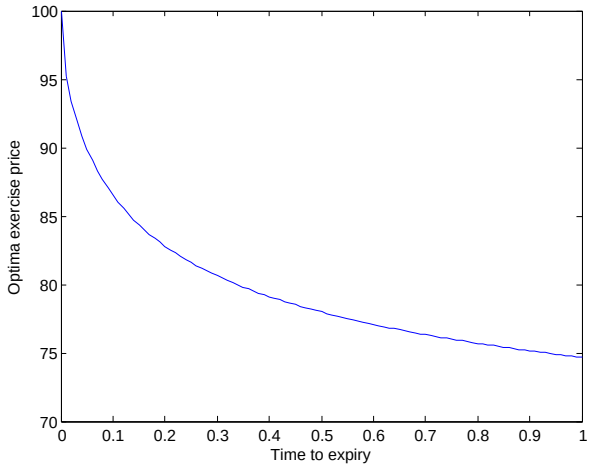

Figure 2.4: Optimal exercise price

PSOR method, has two sources of non-linear error, both from the position of the free-boundary, and from the position of the discontinuity at expiry. Of course, these errors can be reduced by modifying the PSOR algorithm using Body-Fitted Coordinated Method (BFC) [62, 105]. But, BFC method has no clear advantage over PSOR method when applying both to higher dimensional problems. Also, the algorithm suffers from a slow rate of iterative convergence, so slow in fact that it can sometimes appear not to converge at all as a result of the adopted schemes. If the volatility σ is large, and ΔS is small, then the number of iterations required to gain a sufficient level of accuracy can run into the thousands. Finally, the use of FEM in quantitative finance, whilst appealing, till now, there is no theoretical proof for the convergence of the adopted iterative scheme. Of course a number of authors have demonstrated convergence through numerical results. The lack

of proof of convergence limits wide applicability of the method except on a case-by-case basis (cf [117]).

As a consequence, we are motivated to consider a "convergence-proved" approach for the valuation of American options. Therefore, in the subsequent chapters, we shall be focusing on the efficient schemes that would suit the practical needs of market practitioners.

Chapter 3

A comparative study of the direct and the inverse finite element methods for pricing American options

3.1 Introduction

In the previous chapter, we reviewed the application of finite element method to problem in quantitative finance. However, its inverse approach to finance problems is relatively less known. Therefore, in this chapter, we will explore the use of inverse finite element method in financial industry. In particular, the computational performance of the direct and inverse finite element methods for pricing American options will be investigated.

For quite a long time, it has been widely acknowledged that pricing American options is a much more challenging problem because of the nonlinearity originating from the early exercise policy [52, 63]. Additional difficulty arises due to singular behavior of the optimal exercise boundary near the expiration of the option contract [115]. As a result, it is difficult using the conventional numerical methods to effectively and accurately determine the option price. However, with an appropriate inverse method (e.g inverse finite element), the unknown free boundary which makes the pricing problem nonlinear can be accurately determined. Once the free boundary is found, the option pricing problem becomes a fixed

boundary problem, thus, the calculation of the option price is straightforward.

The inverse finite element method (iFEM) is a numerical approach in which an optimization algorithm is coupled with finite element method in order to find optimal values for a set of target parameters which enter the finite element simulation [67]. A user defined objective function serves to measure the optimality of the parameters. Studies involving the use of inverse finite element method (iFEM) are much more limited in the literature. This approach was initially used by Alexandrou [3] in solving nonlinear problems associated with phase change, and in particular with solidification. A characteristic feature of the nonlinear problems solved by Alexandrou [3] is a demarcation line which separates two domains with different material properties. The problems are nonlinear and difficult to handle computationally. The essential concept of the iFEM is to find the location (the nodes of the finite elements) at which, the dependent variable has a predefined value. In other words, the dependent variable is fixed while solution is obtained for the independent variable without inverting the equations. This is the origin of the name of the method.

Similarly to the Alexandrou work [3], an American option is a nonlinear problem due to the presence of unknown free boundary, which separate the region where it is optimal to hold the option from where exercise is optimal. In the conventional finite element method, the question 'what is the option price at a specific location (the nodes of the elements)?' is addressed. In contrast, the iFEM addresses the question 'at what location (the nodes of the elements) does the option has a specific value?'. Essentially, this method uses the concept of fixing the option price while studying the motion of the different "isotherms" of the underlying. In a simpler form, the spatial coordinate of the nodal point becomes the dependent variable whereas the option price is treated as an independent variable. The solution of the problem through this "inverse" approach then reveals the correct location of the free boundary at each time step.

The overall purpose of the research described in this chapter is the investigation of the feasibility of trading the roles of dependent and independent financial variables in economically and mathematically meaningful ways. This allows the American option pricing problems to be solved in an inverse manner. Moreover, we compare the computational efficiency of the direct and inverse approaches. More specifically, we carry out a critical performance analysis of the two approaches against some benchmark solutions. The results of comparison of the two approaches as well as experimental results on their

accuracy-efficiency trade-off are presented.

The remainder of this chapter is organized as follows. In Section 3.2, we present the PDE system that the price of an American put option must satisfy under the Black-Scholes model. Section 3.3 details the two numerical techniques; the dFEM and the iFEM for solving free boundary problem of an American option; the differences between the two approaches are also highlighted. Numerical experiments are presented in Section 3.4. Concluding remarks are given in Section 3.5

3.2 Governing equation and boundary conditions

To compare the computational performance of the dFEM and iFEM, we adopt the Black-Scholes model for an American put option without the dividend yield. The choice of this model allows the evaluation of our results within a framework that permits objective comparison with the existing solutions.

Let $P(S,t)$ denote the value of an American put option, with S being the price of the underlying asset and t being the current time. Under the Black and Scholes [17] framework, the differential system that governs the price of an American put option can be written as

$$\mathcal{A}\begin{cases} \dfrac{\partial P}{\partial t} + \dfrac{1}{2}\sigma^2 S^2 \dfrac{\partial^2 P}{\partial S^2} + rS\dfrac{\partial P}{\partial S} - rP = 0, \\ P(S,T) = \max(K-S,0), \\ P(S_f(t),t) = K - S_f(t), \\ \dfrac{\partial P}{\partial S}(S_f(t),t) = -1, \\ \lim\limits_{S\to\infty} P(S,t) = 0, \end{cases} \tag{3.1}$$

in which r is the risk-free interest rate, σ is the volatility, K is the strike price and T is the expiration time. System $\mathcal{A}$ in (3.1) is defined on $S \in [S_f(t), +\infty), t \in [0,T]$. Here, $S_f(t)$ is the optimal exercise boundary, *a priori* unknown, which needs to be determined as part of the solution of the PDE system. At $t = T$, it has been established (see for examples [52, 68]) that $S_f(T) = K$.

Furthermore, it should be pointed out that while the governing PDE itself is linear in terms of the unknown option value, $P(S,t)$, it is the optimal exercise boundary, $S_f(t)$, that makes the system nonlinear. The nonlinearity of the system is clearly manifested once a

Landau transform is used to convert the free boundary problem to the fixed boundary equivalent. This was demonstrated by Wu and Kwok [109]; the product term of the unknown functions $\frac{1}{S_f}\frac{dS_f}{dt}\frac{\partial P}{\partial S}$, that appears in the partial differential equation, gives a good measure of the strength of the nonlinearity.

To facilitate the development of the algorithms, we shall first non-dimensionalize all variables by introducing the dimensionless variables x and τ in place of S and t, respectively, and new dependent variables $u(x,\tau)$ and $x_f(\tau)$ in place of $P(S,t)$ and $S_f(t)$ as:

$$x = \ln\frac{S}{K}, \quad x_f(\tau) = \ln\frac{S_f(t)}{K}, \quad u(x,\tau) = \frac{P+S}{K} - 1 \quad \text{and} \quad \tau = \frac{\sigma^2(T-t)}{2}.$$

System A now becomes a dimensionless system, which includes a governing differential equation together with the following corresponding initial and boundary conditions:

$$\mathcal{B}\begin{cases} \dfrac{\partial u}{\partial \tau} - \dfrac{\partial^2 u}{\partial x^2} + (1-\gamma)\dfrac{\partial u}{\partial x} + \gamma u + \gamma = 0, \\ u(x,0) = \max(e^x - 1, 0), \\ u(x_f(\tau),\tau) = 0, \\ \dfrac{\partial u}{\partial x}(x_f(\tau),\tau) = 0, \\ \lim\limits_{x\to\infty} u(x,\tau) = e^x - 1, \end{cases} \tag{3.2}$$

where $\mathcal{B}$ is defined on $x_f(\tau) \le x < +\infty, 0 \le \tau \le \frac{\sigma^2 T}{2}$. The parameter, γ is the dimensionless interest rate, and is related to the original risk-free interest rate by $\gamma = \frac{2r}{\sigma^2}$. Note that due to the introduction of the time to expiration τ as the difference between the expiration time, T and the current time, t, the terminal condition in (3.1) has become an initial condition in (3.2). Moreover, since the optimal exercise price, $S_f(t)$, is equal to the strike price, K at the expiration time, T, using the above transformed variable, we must have $x_f(0) = 0$.

For computational purposes, the common practice in the literature is to truncate the semi-infinite domain $[x_f(\tau), +\infty)$ to a finite interval $\Omega = [x_{\min}, x_{\max}]$. While for a large price of the underlying asset, the option value is negligible and is taken to be zero. Then, it is reasonable to truncate the pricing domain into a bounded domain complemented with appropriate boundary conditions. Based on Wilmott et al.'s estimate (see [106]) that the upper bound of the underlying price $S_{\max}$ is three or four times of the strike price, it is reasonable to set $x_{\max} = \ln 5$. On the other hand, since $u(x,\tau) = 0$ for $x \le x_f(\tau)$, there is

no need to show what exactly $x_{\min}$ is. However, for symmetric purposes, some published works set $x_{\min} = -x_{\max}$. This is, however, not the case for the iFEM as we only focus on the positive region. The $x_{\min}$ is set to zero. The reasons for this choice is explained in the subsequent section. In what follows, we present the implementation of the two approaches using the set up in Equation (3.2).

3.3 Formulation of the numerical techniques

This section presents the dFEM and iFEM implementation of the solution of the non-linear system of Equation (3.2). The underlying idea behind the dFEM is similar to that of conventional finite element method. For ease of reference, it is briefly outlined. On the other hand, the iFEM is relatively new development and can have applications elsewhere. In particular, this method would be extended to other pricing formulations and models in the subsequent chapters.

3.3.1 The direct finite element approach (dFEM)

Here, we present the dFEM as an approximate method that allows solving the non-linear pricing problem (3.2) directly without any linearization. As earlier pointed out the concept of dFEM presented here follows the finite element method discussed in Chapter two. However, to conform with the implementation of the iFEM, different dimensionless variable are used. Following the standard Galerkin weighted residue formulation [120], a residual equation is constructed by adopting $v(x)$ as the weighting function. More specifically, the weighted residual or equilibrium statement for the governing differential equation in (3.2) reads

$$R = \int_{\Omega} \left[\frac{\partial u}{\partial \tau} - \frac{\partial^2 u}{\partial x^2} + (1-\gamma)\frac{\partial u}{\partial x} + \gamma u + \gamma \right] v \, dx = 0, \tag{3.3}$$

where $\Omega = [x_{\min}, x_{\max}]$.

In order to reduce the regularity condition on the option price, u, we integrate (3.3) by parts via divergence theorem, to obtain

$$R = \int_{\Omega} \left[\frac{v\partial u}{\partial \tau} + \frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + (1-\gamma)\frac{v\partial u}{\partial x} + \gamma v u + \gamma v \right] dx = 0. \tag{3.4}$$

Next, we define the solution u in terms of the basis function φ_i and time-dependent coefficient $w_i(\tau)$. Similarly, the weighting function v is written in terms of φ_i and an arbitrary constant α_i:

$$u \approx \sum_{i=0}^{N} w_i(\tau)\varphi_i(x), \quad v \approx \sum_{i=0}^{N} \alpha_i \varphi_i(x), \tag{3.5}$$

where w_i are some unknown time-dependent coefficients to be determined.

To proceed with dFEM implementation, we derive the system of nonlinear ordinary differential equation which yields the semi-discrete solution u. With approximation (3.5), equation (3.4) reads

$$R = \frac{dw}{d\tau} A + w(\tau) B + F = 0, \tag{3.6}$$

where:

$$A = \int_{\Omega} \varphi_i(x)\varphi_j(x)\, dx,$$

$$B = \int_{\Omega} \Big(\varphi_i'(x)\varphi_j'(x) - (\gamma - 1)\varphi_i'(x)\varphi_j(x) + \gamma \varphi_i(x)\varphi_j(x) \Big)\, dx,$$

and

$$F = \int_{\Omega} \gamma \varphi_j(x)\, dx,$$

with the solution vector

$$w := (w_1, \ldots, w_N)^T \text{ and } \frac{dw}{d\tau} = \dot{w} := (\dot{w}_1, \ldots, \dot{w}_N)^T.$$

In order to complete the discretization of (3.6) as a fully discretized version of (3.4), we approximate the time derivative appearing in (3.6) with a standard finite difference scheme. We decompose the time interval into equidistant points $0 =: \tau^0 < \tau^1 \cdots < \tau^M := T$ with time step $\triangle\tau = \tau^k - \tau^{k-1}$. The finite difference approximation for the $\frac{dw}{d\tau}$ at time τ is

$$\frac{dw}{d\tau} = \frac{w^{(n+1)} - w^{(n)}}{\triangle\tau}. \tag{3.7}$$

By using Equation (3.7), (3.6) yields the familiar θ-scheme for $\theta \in [0, 1]$

$$R = Kw^{(n+1)} - \bar{K}w^{(n)} + F\triangle\tau = 0, \tag{3.8}$$

where $\bar{K} = A - B\bar{\theta}\triangle\tau, K = A + B\theta\triangle\tau$ and $\bar{\theta} = 1 - \theta$. Choosing $\theta = 1$ yields an explicit Euler scheme and $\theta = 0$ yields an implicit Euler scheme, and $\theta = \frac{1}{2}$ corresponds to the CrankNicolson scheme. Naturally, here we select the latter due to its consistency and convergence error rate $\mathcal{O}(\triangle\tau^2)$.

Finally, with $W = w^{(n+1)}$, $Q = \bar{K}w^{(n)} - F\triangle\tau$ and after specifying the appropriate time-dependant boundary underlying boundary conditions, we obtain a non-singular system of algebraic equations

$$R = KW - Q = 0, \tag{3.9}$$

where W is the vector of the unknown nodal values of the entire domain, K and $\bar{K}$ are the constrained master stiffness matrices and F is the master column matrix. Note that the nonlinear system of equation (3.9) is to be solved forward in time, that is, the terms indexed by n are known, while the terms with index $n + 1$ are to be determined. For numerical computation, we evaluate the element matrices using a linear basis function and then assemble all the element matrices to obtain K and Q. For ease of reading, the details of these are provided in Appendix A.1.

Next, we solve for the unknown W as the zeros of (3.9) using the popular PSOR algorithm, the review of which was presented in chapter 2. Solving problems of the form (3.9) is still a difficult task. Because, after imposing the constrained boundary conditions, the terms K and Q are both functions of the unknown boundary, making the system nonlinear. The coupling of the two types of unknowns (the optimal exercise price and the option values) makes Equation (3.9) much more computational challenging.

3.3.2 Inverse finite element approach (iFEM)

As mentioned earlier, the iFEM was used by Alexandrou [3] in solving nonlinear problems associated with phase change in mechanics. However, in quantitative finance, to the best of our knowledge, the current and only literature is a paper by Zhu and Chen [117]. Based on an algorithm proposed in [3], Zhu and Chen detailed a numerical scheme for locating

the optimal exercise boundary for American put options with no dividend yield.

Essentially, the iFEM involves the use of simulated finite elements to inversely predict desired quantities that are spatially varying with time. Any assumptions included in the finite element model and in the whole simulation of the experiment determine the quality of the inverse solution [67]. The approach can be used to deal with free boundary problems of American options. This is done by requiring the boundaries of the finite elements to remain on "isotherms" of the underlying while the option value is specified *a priori* everywhere in the domain. Therefore, the option is constant along the boundaries of unknowns locations, which are permitted to change as the adopted optimization algorithm (Newton iteration method) proceeds. In this way, the Neumann boundary condition in the PDE system (3.2) is satisfied simplicitly and is dropped from the analysis.

Furthermore, since $x_f(0) = 0$, we must have $\max(e^x - 1, 0) \geq 0$ when x is in the range $x_f \leq x \leq +\infty$. As a consequence, the initial condition in (3.2) can be simplified as $u(x, 0) = e^x - 1$. To realistically implement this approach, the range of option price, P must be known *a priori* [3]. However, in (3.1), it is not difficult to show that P would fall within $[0, K - S_f(t)]$, which varies with respect to time. Fortunately, this difficulty is overcome in system (3.2). After introducing the dimensionless variable, the transformed option price, u falls within $[0, e^x - 1]$, in which the unknown boundary is removed.

To ensure a reasonably accurate solution, we show that u increases monotonically with x for $x \in (x_f, +\infty)$. As previously shown [117], this is achieved by evaluating

$$\frac{\partial u}{\partial x} = \left(\frac{\partial P}{\partial S} + 1\right)\frac{S}{K}.$$

Here, $\frac{\partial u}{\partial x}$ is greater than zero since the delta of an American put option is greater than -1 for $S \in (S_f, +\infty)$. Therefore, u is strictly monotonically increasing with x for $x \in (x_f, +\infty)$.

Next, we proceed to the detail implementation of iFEM. The first step is to deal with the time derivative appearing in the governing partial differential equation. In contrast to the conventional FEM, where $\frac{\partial u}{\partial \tau}$ is approximated by difference scheme, here, it is decomposed into the hedge parameter, delta, i.e., $\frac{\partial u}{\partial x}$ and the velocity of the mesh, $\frac{\partial x}{\partial \tau}$. This step is, in fact, necessary for most of the indirect numerical schemes designed for solving time-dependent problems in physics.

Now, according to the concept of the iFEM, the option price, u is obtained at selected underlying price which varies with respect to time, and therefore, we obtain

$$\frac{du}{d\tau} = \frac{\partial u}{\partial \tau} + \frac{\partial u}{\partial x}\frac{\partial x}{\partial \tau}, \tag{3.10}$$

where $\frac{du}{d\tau}$ is a total derivative, i.e., is the rate of change of the option price at a node. Recall that the option price is distributed and kept constant at all times at the computational nodes. Hence, $\frac{du}{d\tau} = 0$. Moreover, it is obvious that in this case the mesh is not fixed but moves with velocity $V_{mesh} = \frac{dx}{d\tau}$. Therefore,

$$\frac{\partial u}{\partial \tau} = -\frac{\partial u}{\partial x}V_{mesh}. \tag{3.11}$$

Using Equation (3.3) and following the conventional finite element formulation, residual equation can be constructed for the governing PDE in Equation (3.2) as

$$R(x) = \int_0^{x_{\max}} \left[\frac{\partial^2 u}{\partial x^2} + (\gamma - 1 + V_{mesh})\frac{\partial u}{\partial x} - \gamma u - \gamma)\right] v \; dx = 0 \tag{3.12}$$

With the velocity of the mesh being approximated by the first order finite difference $V_{mesh} \approx P_x = \frac{x_{\tau+\Delta\tau} - x_\tau}{\Delta\tau}$ and by integrating by parts, Equation (3.12) reduces to

$$\bar{R}(x) = \int_0^{x_{\max}} \left[\frac{\partial u}{\partial x}\frac{\partial v}{\partial x} + (1 - \gamma - P_x)v\frac{\partial u}{\partial x} + \gamma v u + \gamma v\right] dx \tag{3.13}$$

where $x_{\max}$ is the location of the last node (the limit of the yielded domain). This limit (which is indeed a key parameter) is obtained as a function of time automatically with the solution. The iFEM is implemented by considering that the option price in the yielded part varies from $u = 0$ at the rotating surface $x = 0$ (i.e location of the free boundary at expiry) to $u = e^x_{\max} - 1$ at $x = x_{\max}$.

At this stage, three advantages of iFEM are obvious:

1. The solution is limited to the yielded part of the option, the singularity is automatically removed, and hence the solution corresponds to the ideal constitutive model without any regularization.

2. The boundary conditions are applied and satisfied exactly.

3. There is a reduction of the total number of unknowns due to *a priori* known option values of designated underlying asset.

Before we proceed to the next stage of our implementation procedure, it would be important to make the following remarks. V_{mesh} is approximated by P_x, and thus, the original equilibrium statement $R(x)$ in (3.12) is replaced by $\bar{R}(x)$ in (3.13). This is because of the truncation error brought in by the numerical approximation of V_{mesh} by first order finite difference. This error can be reduced by adopting higher-order approximation method. For simplicity, in the current work, we have adopted the first order approximation and, the implementation with higher order approximation should be similar.

The remaining part of the iFEM formulation involves the selection of suitable shape functions, the computation of the element matrices and assembling of the finite element contributions all follow the dFEM procedures. Finally, after specifying the appropriate time-dependant underlying boundary conditions, we obtain a non-singular system of algebraic equations of the form

$$\bar{R} = K^* W^* - Q^* \tag{3.14}$$

where W^* are the nodal values of the entire domain, K^* and Q^* are respectively, the constrained master stiffness matrix and the constrained master column matrix and they are given by:

$$K^*_{i,j} = \int_\Omega \Big(\varphi_i'(x)\varphi_j'(x) + (1-\gamma-P_x)\varphi_i'(x)\varphi_j(x) + \gamma\varphi_i(x)\varphi_j(x) \Big) dx \text{ and}$$

$$Q^*_j = \int_\Omega \gamma\varphi_j(x)dx.$$

The resulting non-linear system of equations (3.14) is solved using a Newton-Raphson scheme with its quadratic convergence characteristic. We remark that the structure of the nonlinear equation (3.14) is different from that of dFEM (3.9). While W in (3.9) are unknown nodal values, W^* in (3.14) are kept as known constant values along the nodes of unknowns location. The (3.14) formulation results to the elimination of the requirement to specify the spacing of the "isotherms" of the underlying along the moving boundary, which is replaced by the specification of the option value [3].

To obtain a reasonably accurate solution, the monotonicity of W^* is required. Otherwise, the coordinate x that satisfies (3.14) is not unique. This will lead to difficulties in deciding the correct location for a fixed nodal value, even if the convergence of the adopted iteration method is guaranteed. Fortunately, in our case, W^* are strictly monotonically increasing with respect to x, as demonstrated earlier. Therefore, no such problem needs to be further considered.

For numerical computation, the Jacobian of the Newton-Raphson procedure is saved using an element-by-element storage and solved by an iterative method based on a modification of the biconjugate gradient stabilized method [3, 16]. The Jacobi preconditioning was used to speed up convergence and the derivatives of the residual equations $\bar{R}(x)$ are obtained with respect to the unknown nodal locations x. For converged results, usually two to three iterations in the Newton-Raphson procedure are necessary at each time step and the solution advances to the next time step when all unknowns converge to the stopping criterion set to a relative error of 10^{-7}. An algorithm that guarantees the convergence of iFEM is proposed by Zhu and Chen [117] and is summarized as follows:

- At the zeroth time step, the nodal location x_0 is initialized as $x_0 = [a_1 \dots a_{N+1}$, where $a_1 = x_{\min}$, $a_{N+1} = x_{\max}$, and $a_{i-1} < a_i < a_{i+1}$ $(2 < i < N)$, with N being the number of elements in the whole computational domain.

- The Newton iteration scheme is then adopted to find the exact nodal location of the kth $(k > 1)$ time step, i.e., x_k^*. The initial guess of the solution is set as the final solution of the $(k-1)$th time step, i.e., $x_k^0 = x_{k-1}^*$.

The specific implementation of the Newton iteration for this time step is a follows:

1. Suppose that x_k^n is obtained after nth iteration $(n \geq 0)$, we compute the residual $\bar{R}(x_k^n)$ through (3.14), and the corresponding Jacobian matrix $J_{\bar{R}}(x_k^n)$

2. Calculate the unknown nodal locations at the $(n+1)$th iteration step through

   $x_k^{n+1} = x_k^n - J_{\bar{R}}^{-1}(x_k^n)$

3. Repeat steps 1 and 2 until $\| x_k^{n+1} - x_k^n \| < \epsilon$ is satisfied. Set the solution of the kth time step to $x_k^* = x_{k-1}^{n+1}$, which completes the Newton iteration for the kth time step.

In the above algorithm, the location of the fixed boundary is excluded, since it is already the solution of the corresponding nodal value, and no iteration is further needed. If the

location of the fixed boundary were still taken into consideration, the residual associated with this particular point would be zero, resulting in the corresponding row of the Jacobian matrix being zero. Consequently, the Jacobian matrix would be highly singular, and the Newton iteration fails.

In what follows, we discuss the convergence of Newton scheme.

Lemma 3.3.1. *For every square matrix A and given norm $\|.\|$, define*

$$\|A^*\| = \min_{\|x\|=1} \|Ax\|$$

Then A is invertible if and only if $\|A^*\| > 0$. If A is invertible, then $\|A^{-1}\| = \frac{1}{\|A^*\|}$.

Proof. The detail of proof can be found in [96] □

Theorem 3.3.2. *Let A be a square matrix; then*

$$\lim_{k\to 0} A^k = 0 \Leftrightarrow \rho(A) < 1,$$

where $\rho(A)$ is the spectral radius of matrix A.

Moreover, the geometric series $\sum_{i=0}^{\infty} A^k$ is convergent iff $\rho(A) < 1$. such a case

$$\sum_{i=0}^{\infty} A^k = (I - A)^{-1}.$$

As a result, if $\rho(A) < 1$ the matrix $I - A$ is invertible and the following inequalities hold

$$\frac{1}{1 + \|A\|} \leq \|(I - A)^{-1}\| \leq \frac{1}{1 - \|A\|},$$

where $\|\cdot\|$ is an induced matrix norm such that $\|A\| < 1$ and I being the identity operator.

Proof. The detail of proof can be found in [88]. □

Theorem 3.3.3. *If $\triangle\tau < \triangle x_{0,min}/2\gamma(2e-1)$, the Jacobian matrix J_R is invertible at the exact solution x^*, and moreover, there exists a positive constant C, such that $\|J_R^{-1}(x^*)\| \leq C$, where $\triangle x_{0,min}$ is the smallest interval of the initial input vector at $\tau = 0$, and $\|\cdot\|$ is any proper matrix norm.*

The details of this proof can be found in [88, 117]. It is, however, reproduced here for convenience of the reader.

Proof. The Jacobian matrix at the exact solution x^* can be written as: for $i = 2 \dots N-1$

$$J_R(i-1,i) = \left(\frac{1}{6\triangle t} + \frac{\gamma}{3} - \frac{1}{(\triangle x_i^*)^2}\right)\triangle W_i - \frac{\gamma}{2}W(i) - \frac{\gamma}{2},$$

$$J_R(i,i) = \left(\frac{1}{3\triangle t} + \frac{\gamma}{6} + \frac{1}{(\triangle x_i^*)^2}\right)\triangle W_i + \left(\frac{1}{3\triangle t} + \frac{\gamma}{6} + \frac{1}{(\triangle x_{i+1}^*)^2}\right)\triangle W_{i+1},$$

$$J_R(i+1,i) = \left(\frac{1}{6\triangle t} + \frac{\gamma}{3} - \frac{1}{(\triangle x_{i+1}^*)^2}\right)\triangle W_{i+1} + \frac{\gamma}{2}W(i) + \frac{\gamma}{2},$$

$$J_R(2 \le j < i-1, i) = J_R(i+1 < j \le N-1, i) = 0,$$

and

$$J_R(1,1) = \left(\frac{1}{3\triangle t} + \frac{\gamma}{6} + \frac{1}{(\triangle x_2^*)^2}\right)\triangle W_2 + \frac{\gamma}{2}W(1) + \frac{\gamma}{2},$$

$$J_R(2,1) = \left(\frac{1}{6\triangle t} + \frac{\gamma}{3} - \frac{1}{(\triangle x_2^*)^2}\right)\triangle W_2 + \frac{\gamma}{2}W(1) + \frac{\gamma}{2},$$

$$J_R(N-1,N) = \left(\frac{1}{6\triangle t} + \frac{\gamma}{3} - \frac{1}{(\triangle x_N^*)^2}\right)\triangle W_N - \frac{\gamma}{2}W(N) - \frac{\gamma}{2},$$

$$J_R(N,N) = \left(\frac{1}{3\triangle t} + \frac{\gamma}{6} + \frac{1}{(\triangle x_N^*)^2}\right)\triangle W_N + \left(\frac{1}{3\triangle t} + \frac{\gamma}{6} + \frac{1}{(\triangle x_{N+1}^*)^2}\right)\triangle W_{N+1},$$

where $\triangle x_i^* = x^*(i) - x^*(i-1)$, and $\triangle W_i = W(i) - W(i-1)$.

With the specific structure of $J_R(x^*)$, it suffices to show that when$\triangle t < \frac{\triangle x_{0,min}}{2\gamma(2e-1)}$, $J_R(x^*)$ is strictly diagonally dominant. This is achieved in the discussion of the following four.

Case 1. For $i = 2 \dots N-1$, if $J_R(i-1,i) < 0$, we have

$$J_R(i,i) - \sum_{j \ne i} \mid J_R(j,i) \mid = \left(\frac{1}{2\triangle t} + \frac{\gamma}{2}\right)\triangle W_i + \left(\frac{1}{2\triangle t} + \frac{\gamma}{2}\right)\triangle W_{i+1}.$$

Since W is monotonically increasing with respect to the index i , it is straightforward to show that for any $k = 2 \dots N+1, \triangle W_k > 0$, and thus

$$J_R(i,i) > \sum_{j \ne i} \mid J_R(j,i) \mid .$$

Case 2. For $i = 2 \dots N-1$, if $J_R(i-1,i) > 0$, we have

$$J_R(i,i) - \sum_{j \ne i} \mid J_R(j,i) \mid = \left(\frac{1}{6\triangle t} - \frac{\gamma}{6} + \frac{2}{(\triangle x_i^*)^2}\right)\triangle W_i + \left(\frac{1}{6\triangle t} - \frac{\gamma}{6} + \frac{2}{(\triangle x_{i+1}^*)^2}\right)\triangle W_{i+1},$$

which is greater than zero if $\triangle t < \triangle x_{0,min}/2\gamma(2e-1) < 1/\gamma$.

Case 3. For $i = 2 \ldots N-1$, if $J_R(i-1,i) > 0$ and $J_R(i+1,i) < 0$, we have

$$J_R(i,i) - \sum_{j \neq i} \mid J_R(j,i) \mid = \left(\frac{1}{6\triangle t} - \frac{\gamma}{6} + \frac{2}{(\triangle x_i^*)^2}\right)\triangle W_i + \left(\frac{1}{2\triangle t} + \frac{\gamma}{2}\right)\triangle W_{i+1} + \gamma + \gamma W(i).$$

Also, it is not difficult to check that if

$\triangle t < \triangle x_{0,min}/2\gamma(2e-1) < 1/\gamma, J_R(i,i) > \sum_{j \neq i} \mid J_R(j,i) \mid$.

Case 4. For $i = 2 \ldots N-1$, if $J_R(i-1,i) < 0$ and $J_R(i+1,i) > 0$, we obtain

$$J_R(i,i) - \sum_{j \neq i} \mid J_R(j,i) \mid = \left(\frac{1}{2\triangle t} + \frac{\gamma}{2}\right)\triangle W_i + \left(\frac{1}{6\triangle t} - \frac{\gamma}{6} + \frac{2}{(\triangle x_{i+1}^*)^2}\right)\triangle W_{i+1} - \gamma - \gamma W(i) \tag{3.15}$$

By applying the mean-value theorem to the right hand side of (3.15), we obtain

$$\begin{aligned} J_R(i,i) - \sum_{j \neq i} \mid J_R(j,i) \mid =& \exp(x_{0,\xi})\triangle x_{0,i}\left(\frac{1}{2\triangle t} + \frac{\gamma}{2}\right) - \gamma\exp(x_{0,i}) \\ &+ \left(\frac{1}{6\triangle t} - \frac{\gamma}{6} + \frac{2}{(\triangle x_{i+1}^*)^2}\right)\triangle W_{i+1} \\ &= \exp(x_{0,\xi})\Bigg[\triangle x_{0,i}\left(\frac{1}{2\triangle t} + \frac{\gamma}{2}\right) \\ &- \gamma\exp(x_{0,i} - x_{0,\xi})\Bigg] + \left(\frac{1}{6\triangle t} - \frac{\gamma}{6} + \frac{2}{(\triangle x_{i+1}^*)^2}\right)\triangle W_{i+1} \\ &> \exp(x_{0,\xi})\left[\triangle x_{0,i}\left(\frac{1}{2\triangle t} + \frac{\gamma}{2}\right) - \gamma\exp(\triangle x_{0,i})\right], \end{aligned}$$

where $x_0(i-1) < x_{0,\xi} < x_0(i)$.

On the other hand, when $\triangle t < \triangle x_{0,\min}/2\gamma(2e-1)$, it is straightforward to show that

$$\frac{1}{\triangle t} > \frac{2\gamma(2e-1)}{\triangle x_{0,\min}} \geq \frac{2\gamma(2e-1)}{\triangle x_{0,i}}.$$

Therefore,

$$\begin{aligned} J_R(i,i) - \sum_{j \neq i} \mid J_R(j,i) \mid >& \exp(x_{0,\xi})\left[\triangle x_{0,i}\left(\frac{1}{2\triangle t} + \frac{\gamma}{2}\right) - \gamma\exp(\triangle x_{0,i})\right] \\ &> \exp(x_{0,\xi})[\gamma(2e-1) - \gamma] = 2\exp(x_{0,\xi})\gamma(e-1) > 0, \end{aligned}$$

and thus $J_R(i,i) > \sum_{j \neq i} \mid J_R(j,i) \mid$.

Similarly, it can be easily verified that when

$\triangle t < \triangle x_{0,min}/2\gamma(2e-1), J_R(1,1) > \sum_{j\neq 1} \mid J_R(j,1) \mid$ and $J_R(N,N) > \sum_{j\neq N} \mid J_R(j,N) \mid$.

Therefore, based on the above four cases, it can be concluded that if $\triangle t < \triangle x_{0,min}/2\gamma(2e-1)$, the Jacobian matrix $J_R(x^*)$ is strictly diagonally dominant. According to the Gershgorin circle theorem [92] that any strictly diagonally dominant matrix is non-singular, it is clear that $J_R(x^*)$ is then invertible. By applying Lemma 3.3.1, we obtain

$$\|J_R^{-1}(x^*)\|^{-1} = \|J_R^*(x^*)\| > 0$$

As a result, there exists a positive constant M, such that $M\|J_R^{-1}(x^*)\|^{-1} > 0$, which yields $\|J_R^{-1}(x^*)\| < 1/M$. Therefore, the norm of the inverse of the Jacobian matrix is bounded. □

Lemma 3.3.4. *Let $F = (f_1, f_2 \dots f_m) : R^n \mapsto R^m$, and suppose that the partial derivatives $\partial f_i/\partial x_j$, $1 \leq i \leq m, 1 \leq j \leq n$ exist on a neighborhood of X_0, and are continuously differentiable at X_0. Then F is continuously differentiable at X_0.*

Proof. The details of proof can be found in Trench [101]. □

Lemma 3.3.4 gives a sufficient condition for a vector function F to be continuously differentiable at a given point X_0.

Lemma 3.3.5. *If $F(X) = (F_1(X), F_2(X) \dots F_N(X))$ and each $F_i(X)$: $R^N \mapsto R^N$, is continuously differentiable on an open set containing a compact set D, then there exists a positive constant M, such that $\|F(X) - F(Y)\| \leq M\|X - Y\|$, if $X, Y \in D$, where the same symbol $\|\cdot\|$ is defined as two consistent matrix and vector norms.*

Proof. Since $F_i(X)$ is continuously differentiable on D , we have, for all $X, Y \in D$, and $i \in 1, 2, \dots, N$

$$\|F_i(X) - F_i(Y)\|_2^2 \leq M_i\|X - Y\|_2^2,$$

where M_i is a positive constant. The details of the proof of this statement can be found in Trench [101].

Now, summing from $i = 1$ to $i = N$ yields

$$\sum_{i=1}^{N} \|F_i(X) - F_i(Y)\|_2^2 \leq \sum_{i=1}^{N} M_i\|X - Y\|_2^2 \leq NM_{\max}\|X - Y\|_2^2. \tag{3.16}$$

According to the definition of the F-norm, we have

$$\sum_{i=1}^{N} \|F_i(X) - F_i(Y)\|_2^2 = \|F(X) - F(Y)\|_F^2,$$

which, combined with (3.16), yields

$$\|F(X) - F(Y)\|_F^2 \leq NM_{\max}\|X - Y\|_2^2,$$

and consequently,

$$\|F(X) - F(Y)\|_F \leq \sqrt{NM_{\max}}\|X - Y\|_2.$$

On the other hand, it is clear that all the matrix norms are equivalent, and so are the vector norms (Tavernini [96]). Thus, there exist positive constants C_1, C_2, C_3 and C_4, such that

$$C_1\|F(X) - F(Y)\| \leq \|F(X) - F(Y)\|_F \leq C_2\|F(X) - F(Y)\|,$$

$$C_3\|X - Y\| \leq \|X - Y\|_2 \leq C_4\|X - Y\|,$$

where the same symbol $\|\cdot\|$ are defined as two consistent matrix and vector norms. Therefore, if $X, Y \in D$,

$$\|F(X) - F(Y)\| \leq M\|X - Y\|,$$

where $M = \sqrt{NM_{\max}}C_4/C_1$ □

The details of the above prove can be found in [88]

Theorem 3.3.6. *For any $X, Y \in B(x^*, R)$, the Jacobian matrix satisfies $\|J_R(X) - J_R(Y)\| \leq M\|X - Y\|$, where x^* is the exact solution, M is a positive constant, and $B(x^*, R)$ is an open ball centering at x^*, with radius R being defined as*

$$R = \frac{min(\triangle x_{i+1}^*, i = 1 \ldots N)}{3}.$$

Proof. The proof of the theorem is deferred to Appendix *A*.2. □

Theorem 3.3.7. *Let F: $\mathbb{R} \to \mathbb{R}$ be a $\mathcal{C}^1$ function in a convex open set D of $\mathbb{R}^n$ that contains x^*. Suppose that $J_F^{-1}(x^*)$ exists and that there exist positive constants R, C and*

L, such that $\|J_F^{-1}(x^*)\| \leq C$ *and*

$$\|J_F(x) - J_F(y)\| \leq L\|x - y\| \quad \forall\, x, y \in B(x^*; R),$$

having denoted by the same symbol $\|.\|$ *two consistent vector and matrix norms. Then, there exists* $r > 0$ *such that, for any* $x^{(0)} \in B(x^*; r)$*, the Newton iterations constructed for* $x^{(k)}$ *converges to* x^**.*

$$\|x^{(k+1)} - x^*\| \leq CL\|x^{(k)} - x^*\|. \tag{3.17}$$

Although the details of proof can be found in [88, 117], it is reproduced here for ease of reading.

Proof. Proceeding by induction on k, let us check (3.17) and, moreover, that $x^{(k+1)} \in B(x^*; r)$, where $r = \min(R, 1/(2CL))$. First, we prove that for any $x^{(0)} \in B(x^*; r)$, the inverse matrix $J_F^{-1}(x^{(0)})$ exists. Indeed

$$\|J_F^{-1}(x^*)[J_F(x^{(0)}) - J_F(x^*)]\| \leq \|J_F^{-1}(x^*)\|\|J_F(x^{(0)}) - J_F(x^*)\| \leq CLr \leq \frac{1}{2},$$

and thus, thanks to Lemma 3.3.1, we can conclude that $J_F^{-1}(x^{(0)})$ exists, since

$$\|J_F^{-1}(x^{(0)})\| \leq \frac{\|J_F^{-1}(x^*)\|}{1 - \|J_F^{-1}(x^*)[J_F(x^{(0)}) - J_F(x^*)]\|} \leq 2\|J_F^{-1}(x^*)\| \leq 2C.$$

As a consequence, $x^{(1)}$ is well defined and

$$x^{(1)} - x^* = x^{(0)} - x^* - J_F^{-1}(x^{(0)})[F(x^{(0)}) - F(x^*)].$$

Factoring out $J_F^{-1}(x^{(0)})$ on the right hand side and passing to the norms, we get

$$\|x^{(1)} - x^*\| \leq \|J_F^{-1}(x^{(0)})\|\|F(x^*) - F(x^{(0)}) - J_F(x^{(0)})[x^* - x^{(0)}]\| \leq 2C\frac{L}{2}\|x^* - x^{(0)}\|^2$$

where the remainder of Taylor's series of F has been used. The previous relation proves (3.17) in the case $k = 0$. Moreover, since $x^{(0)} \in B(x^*; r)$ we have $\|x^* - x^{(0)}\| \leq \frac{1}{2CL}$, from which $\|x^{(1)} - x^*\| \leq \frac{1}{2}\|x^* - x^{(0)}\|$. This ensures that $x^{(1)} \in B(x^*; r)$.

By a similar proof, one can check that, should (3.17) be true for a certain k, then the

same inequality would follow also for $k+1$ in place of k. □

Theorem 3.3.7 confirms that the Newton's method is quadratically convergent only if the initial guess $x^{(0)}$ is sufficiently close to the exact solution x^*, and if the Jacobian matrix is nonsingular. Thus, the iFEM, whilst appealing for solving free boundary problems, the solutions may converge slowly or may not even converge at all, if the initial guess of the adopted scheme is not properly chosen [3]. In the problem defined above, we follow closely the algorithm developed by [117] wherein nodal positions of the present time step are chosen as the initial guess of locations of the elements at the next step. For a small time interval, the nodal positions of the two adjoining time steps should not differ too much since x is continuous with respect to τ and the time step is sufficiently small (cf. [117]).

3.4 Numerical experiments

In this Section, we report the results of numerical experiments and some detailed comparisons are made between the dFEM and iFEM for pricing American options. To provide a fair and meaningful comparison, linear basis function is used for each of the discretization. Evaluation of the results is conducted with Zhu's analytical solution [113] as a benchmark. Using this equivalent set up, the goal is to compare the two methods in terms of computational performance.

In order to facilitate objective comparison, we conduct the experiments on the examples presented in [113, 117]. The parameters used are: the strike price $K = 100$, the risk-free interest rate $r = 10\%$, the volatility of the underlying asset $\sigma = 30\%$ and the tenor of the contract being $T = 1$ year.

To compare the results of the two approaches, we focus mainly on the comparisons based on the optimal exercise prices, $S_f(\tau)$ instead of the option value, P since $S_f(\tau)$ is more difficult to be accurately calculated than the option price. In fact, once $S_f(\tau)$ is accurately determined, the pricing problem becomes a fixed boundary problem and the calculation of the option price is straightforward.

3.4.1 Comparison in terms of accuracy

For us to compare the solution accuracy of dFEM and iFEM with respect to various levels of discretization and the number of time intervals, we use the $RMSRE$ (root mean square

relative error), which is defined as

$$RMSRE = \sqrt{\frac{1}{I}\sum_{i=1}^{I}\left(\frac{(a_i - \bar{a}_i)}{a_i}\right)^2}$$

where $\bar{a}_i$'s are the nodal values of the S_f associated with dFEM and iFEM, a_i's are the S_f obtained from the Zhu's analytical result and I is the number of sample points used in the $RMSRE$. With the $RMSRE$, comparison of the overall difference of the computed numerical results and the exact solution based on Zhu's analytical result can be clearly demonstrated. In our numerical experiments, I was set to be 50 in all the results presented. In order to have a good comparison of the error associated with each method, the $RMSRE$

Table 3.1: The variation of RMSRE when the grid sizes are gradually increased. M: the number of time intervals; N: the number of elements

	M=10		M=20		M=40		M=80	
	dFEM	iFEM	dFEM	iFEM	dFEM	iFEM	dFEM	iFEM
N=5	0.1045	0.0818	0.0741	0.0816	0.0486	0.0424	0.0306	0.0371
N=10	0.0917	0.0795	0.0592	0.0480	0.0402	0.0291	0.0287	0.0083
N=25	0.0872	0.0550	0.0574	0.0333	0.0358	0.0239	0.0214	0.0058
N=50	0.0748	0.0477	0.0531	0.0290	0.0322	0.0241	0.0193	0.0026

when the number of steps in both spacial and temporal directions are gradually increased are tabulated respectively in Table 3.1. From this table, one can clearly see that the dFEM produces consistently larger error than the iFEM when equal size element and number of time interval are used. The results suggest that iFEM yields a more accurate result than dFEM. This clear difference may be connected with the different convergence schedules; the adopted PSOR scheme in dFEM has slow convergence, whereas, the full Newton iterative scheme adopted in iFEM has a quadratic convergence rate. Moreover, a careful observation of the table shows that the differences between the $RMSRE$ of the two methods on a coarse grid resolution, say $N = 5$ and $M = 10$ is quite substantial, but the difference is not well pronounced on a relatively fine resolution of $N = 50$ and $M = 80$. A reason adduced to this observation is that fine grid resolution produces better results, and hence, less difference in $RMSRE$.

One should also notice from Table 3.1 that when the time interval parameter M, reduces (corresponding to an increase in the size of the time step), the $RMSREs$ for both methods become larger. Another important observation is that a reduced time interval

worsened the convergence conditions of the adopted PSOR and Newton iterative schemes for both dFEM and iFEM, respectively. This is as a result of large discretization errors when dealing with the time derivative, $\frac{dw}{d\tau}$, and the velocity of the mesh, V_{mesh} associated with the dFEM and iFEM schemes, respectively.

A reduced number of elements also produces large $RMSREs$ for the two methods. The error in this case, however, is connected with the finite element discretization acting on the residual equation. Note that in all the computations, linear basis function is used. Numerical solutions based on quadratic shape function would have a smaller $RMSREs$ than those computed on linear function.

Having compared the dFEM and iFEM based on the variation of $RMSRE$ when the grid sizes are gradually increased, it is also important to compare the accuracy of both approaches based on a fine resolution of a grid size. Such a comparison is presented in Figure 3.1 with grid resolutions of $N = 30$ and $M = 100$ for the iFEM and $N = M = 25$ for the dFEM. As shown in Figure 3.1, it can be easily seen that although both numerical results show a good convergence to the Zhu's analytical results, the results based on iFEM better approximate the benchmark solution.

Furthermore, a close examination of Figure 3.1 reveals that the iFEM almost coincide with the benchmark solution at the expiration date, $t = T = 1$ (year), the optimal exercise price calculated by Zhu‘s analytical method is $S_f(T) = 76.113$, whereas, they are $S_f(T) = 76.2478$ and $S_f(T) = 76.3085$. In addition, Figure 3.1 reveals that both FEM methods have their curves above the analytical curve when the time close to expiry. This might due to the presence of singularity at expiry, which is not possible for most of the numerical algorithms to deal with. However, at $\tau = 0$, the iFEM algorithm is designed such that the location of the optimal exercise price at expiry is known *a priori* and is already included in the algorithm as the strike price.

3.4.2 Comparison in terms of efficiency

There have been two thrusts in the development of algorithms as far as real world tasks are concerned [73]. One has emphasized higher accuracy; the other faster implementation. These two thrusts, however, have been independently pursued, without addressing the accuracy versus efficiency trade-offs. The importance of accuracy of an algorithm diminishes when response time is slow for a given task. The converse is also true; importance of a fast

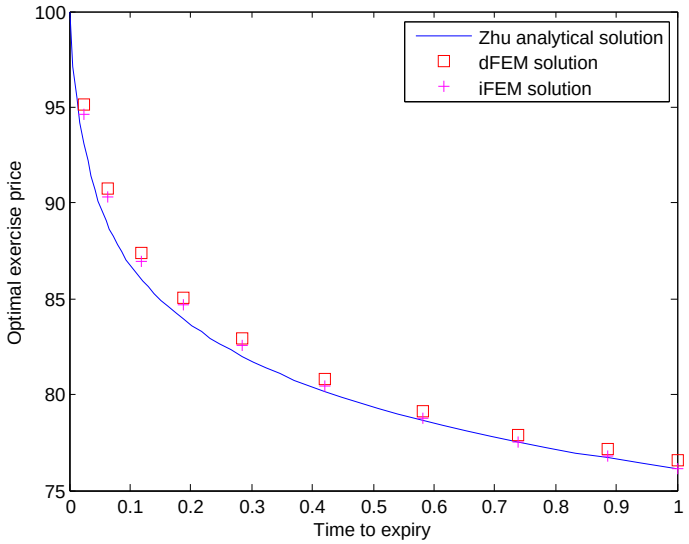

Figure 3.1: Comparison of S_f for dFEM and iFEM

algorithm diminishes if the accuracy and precision are insufficient for subsequent financial interpretations. With this in mind, comparing the dFEM and iFEM in terms of computational performance is not an easy task. Although the accuracy-efficiency characteristic is algorithm-dependant, an understanding of a general pattern is crucial in evaluating algorithm performance as far as real world tasks are concerned.

As expected, in our numerical experiments, the computing time for both methods increases with the grid size. However, the dFEM incur less computational cost than the iFEM under the same grid resolutions. In fact, for the grid resolution $N = M = 25$, the computational cost for the dFEM is just 1 second, whereas, it takes iFEM about 200 seconds to produce related result. But again, this nice feature does not make the dFEM more efficient than iFEM, because efficiency of an algorithm does not depend only on the speed of calculation, but also on the accuracy. The task of establishing a "trade-off" between accuracy and efficiency shall be our goal subsequently.

3.4.3 Accuracy versus efficiency

In our discussions pertaining to accuracy and speed, all the illustrative results are based on linear basis function. A 2-D curve characterizing the accuracy-efficiency (AE) trade-off is used to evaluate the performance of the methods. On the curve, accuracy (abscissae) is measured by the $RMSRE$, (calculated using the Zhu's analytical result [113] as the base value), and computational efficiency (ordinate) is measured by the total CPU time

consumed at each run. These curves are generated by setting parameter of the algorithms in the temporal direction to a fixed value, $M = 25$, while varying resolution in the spatial direction. Note that similar curve is obtained when grid resolution in spatial direction is fixed and the resolution in temporal direction is varied. In the AE family of performance

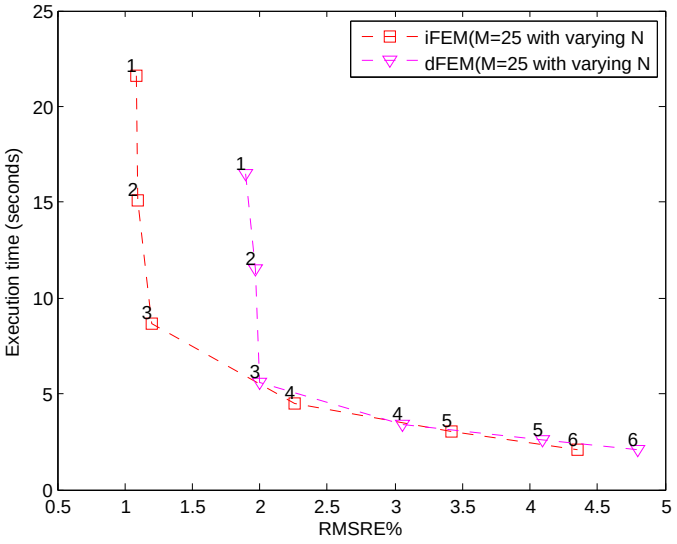

Figure 3.2: Accuracy versus efficiency

curves depicted in Figure 3.2, six different resolutions were used in the computation. Each curve corresponds to the iFEM and dFEM algorithm, respectively. A point on the performance curve denotes a certain parameter setting (grid resolution). As clearly shown, the accuracy is inversely varying with the speed of calculation for the two methods (curves); an expected result. A higher accuracy usually implies a slower run time and verse-versa for any resolution. It can also be easily observed that the dFEM curve shows a greater speed of calculation but with larger error, whereas iFEM has significantly reduced error with higher computing time under equivalent grid resolution as the dFEM.

Following the explanation in [34, 73], the distance from the origin to AE curve represents the overall performance (efficiency) of the algorithm. Performance point close to the origin (small error and low execution time) is indicative of better algorithm operating point. In terms of AE performance, iFEM appears more flexible and effective, because the curve is closer to the origin in about five out of the six resolutions (from the 2nd $-$ 6th). A close examination of the curves shows that at any point between the 3rd to 6th resolution, iFEM curve is closer to the origin than dFEM curve. A nice feature of the iFEM as indicated in AE curve is that using any point on the curve between the 3rd and 4th

resolutions, which appear to be the closest region on the iFEM curve to the origin, a high computational performance (efficiency) in terms of a satisfactory computing time and accuracy is achieved.

The AE curve is also useful in determining the computational cost when the same order of accuracy is maintained. For example, let us consider $RMSRE = 2.5\%$ on the curves where errors for the two methods exist. It is not difficult to see that the iFEM cost along the CPU time axis is lower than its dFEM counterpart, reaffirming the fact that when the same accuracy were to be maintained, the dFEM requires a very fine grid resolution, which lead to higher computational cost, and thus, iFEM could be the better option.

3.5 Conclusion

In this work, we have compared the direct and inverse finite element methods for pricing American put options. Based on the results of our numerical experiments, the dFEM, while appealing in terms of CPU time savings, produces larger error than its inverse counterpart for similar grid resolutions. Furthermore, by using the performance accuracy-efficiency curves to establish the trade-offs, the iFEM is indeed more flexible and efficient, as a higher performance in terms of a satisfactory computing time and accuracy can be achieved. The results presented in this work demonstrate that the iFEM deserves consideration as an alternative numerical techniques for pricing American options. In the subsequent chapters, we will exploit other advantageous features of the inverse finite element method by considering option pricing problems under different formulations and frameworks.

Chapter 4

On the inverse finite element approach for pricing American options under linear complementarity formulation

4.1 Introduction

In Chapter 3, we compared the direct and inverse finite element methods for solving the free boundary problem of American options under the Black-Scholes setting. The inverse approach was found to be more efficient than its direct counterpart. In the Black-Scholes-Merton model, as well as in more general stochastic models in finance, an American option problem can be formulated as a linear complementarity problem (LCP). An interesting question is whether or not the inverse finite element method can be used to price American options under linear complementarity formulation. Therefore, this chapter is devoted to studying the numerical performance of inverse finite element method for a linear complementarity problem arising from an American option valuation.

The LCP was introduced into mathematical field more than four decades ago, and has now developed into a very worthwhile discipline in applied mathematics. A detailed study of LCP, including a complete theory and extensive algorithms, is presented in the book by Cottle et al. [31]. Among the classic application areas of the LCP is that of optimal

stopping problem, which is the heart of all option pricing models with an American-style exercise [53]. This connection makes LCP an important technique in quantitative finance. The standard treatment of LCP for American options pricing can be found in [107], for the simple case of the Black-Scholes-Merton model and in [53], for several more complicated settings.

In this chapter, using a novel computational method (inverse finite element method) suited for the free boundary problems with singularities, we develop an algorithm which combines Newton iteration scheme with finite element method. While the inverse algorithm described in Chapter 3 does work well for the free boundary problem of an American option, the idea cannot be immediately extended to problems under linear complementarity formulation. The reason is that the unknown optimal exercise boundary is not explicitly defined under linear complementarity formulation as in the free boundary case. This formulation is beneficial for iterative solution, since the unknown free boundary can be obtained in a postprocessing step [84, 87]. To illustrate the performance of the inverse finite element method (iFEM), the solution accuracy is examined with respect to various element shape functions. The results obtained are compared with existing solutions.

The remainder of this chapter is organized as follows. In Section 4.2, we introduce the American option pricing model and its equivalent formulations: differential complementarity problem and variational inequality. Also, we discuss some implementation issues concerned with iFEM. In Section 4.3, we discuss the formulations of a system of nonlinear algebraic equations using the iFEM. Numerical experiments and comparisons of the iFEM performance with other works are presented in Section 4.5. Concluding remarks and future directions are presented in the last section.

4.2 Mathematical formulation

In this section, we present the linear complementarity problem and variational inequality of an American option which will lay the foundation for the rest of the work in this chapter. Consider an asset with price, S which satisfies the following stochastic differential equation

$$dS = \mu S dt + \sigma S dW_t,$$

where dW_t is a standard Brownian motion, μ is the drift rate and σ is the volatility of the underlying asset. We define $P(S,t)$ as the price of the option with respect to the underlying asset price, S and time, t for some function $P : (0,\infty) \times [0,T] \to \mathbb{R}$, where T is the expiry date of the contract. Under the non-arbitrage assumption, the Black-Scholes equation governing the price of an American put option can be derived as [107]:

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

where r is the risk-free interest rate and δ is the dividend yield paid by the underlying asset. Due to the early exercise possibility of American options, an additional constraint

$$P(S,t) \geq \Psi(S) \tag{4.2}$$

has to be introduced in order to avoid arbitrage possibilities. Here, $\Psi(S) = \max(K-S,0)$ is the payoff function of the option contract. The price of the option is obtained by solving the partial differential equation with the previous constraint, boundary conditions, and a final condition.

It is well-known (see for example [106]) that there is a value $S_f(t)$ for all t which divides the domain $(0,\infty)$ into two sub-domains $(0, S_f(t))$ and $(S_f(t),\infty)$ in such a way that in one of these sub-domains the price of the option equals to the payoff function while in the other one it is higher than the payoff. The price of the option satisfies the Black-Scholes Equation (4.1) in the sub-domain where it is higher than the payoff. The function $S_f(t)$ is not known beforehand and it has to be found together with the price of the option. Hence, the option pricing problem is a free boundary problem.

Wilmott et al. [106] introduced that the free boundary problem of an American option was equivalent to a parabolic linear complementarity problem:

$$\mathcal{A}\begin{cases} (\mathcal{L}_{BSM}P).(P-\Psi) = 0, \\ \mathcal{L}_{BSM}P \leq 0, \\ P - \Psi \geq 0, \end{cases} \tag{4.3}$$

with the terminal condition

$$P(S,T) = \max(K - S, 0), \tag{4.4}$$

and the boundary conditions

$$\lim_{S\to 0} P(S,t) = K, \qquad \lim_{S\to\infty} P(S,t) = 0, \tag{4.5}$$

where

$$\mathcal{L}_{BSM} = \frac{\partial}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2}{\partial S^2} + (r-\delta)S\frac{\partial}{\partial S} - rI,$$

denotes the Black-Scholes differential operator and I is the identity operator. The system $\mathcal{A}$ is defined on $t \in [0,T], S \in (0,\infty)$ such that the option price P and its first derivative $\frac{\partial P}{\partial S}$ are continuous for all $S \in (0,\infty)$. Note that the solution, $P(S,t)$ is equal to the payoff function at time, t in the domain $(0, S_f(t))$, while in the domain $(S_f(t),\infty)$, it satisfies the Black-Scholes partial differential equation.

To facilitate the development of the algorithm, we introduce the following new dimensionless variables:

$$\begin{cases} x = \ln\dfrac{S}{K}, \quad \tau = \dfrac{\sigma^2(T-t)}{2} \\ u(x,\tau) = \dfrac{1}{K}P(S,t)\exp\left(\dfrac{1}{2}(q_\delta - 1)x + (\dfrac{1}{4}(q_\delta - 1)^2 + q)\tau\right) \end{cases} \tag{4.6}$$

The parameters q and q_δ are defined as

$$q = \frac{2r}{\sigma^2}, \quad q_\delta = \frac{2(r-\delta)}{\sigma^2},$$

respectively. Under this transformation, it is easy to show that, the complementarity problem in (4.3)-(4.5) becomes, a dimensionless system, which includes a standard linear complementarity problem together with the following corresponding initial and boundary

conditions:

$$\mathcal{B}\begin{cases} u-\psi \geq 0, \\ \mathcal{L}u \geq 0, \\ \mathcal{L}u.(u-\psi)=0, \\ u(x,0)=\psi(x)=\max(\mathrm{e}^{\frac{x}{2}(q_\delta-1)}-\mathrm{e}^{\frac{x}{2}(q_\delta+1)},0), \\ u(x,\tau)=\psi(x) \quad \text{for } x\to\pm\infty, \end{cases} \tag{4.7}$$

where $\mathcal{L}$ is a partial differential operator defined as

$$\mathcal{L}=\frac{\partial}{\partial\tau}-\frac{\partial^2}{\partial x^2}.$$

Here, $\mathcal{B}$ is defined on $\tau\in[0,T\sigma^2/2]$, $x\in(-\infty,+\infty)$.

We comment that although the results of solving system $\mathcal{A}$ together with the boundary conditions (4.4) and (4.5) directly, or applying the transformation to have a dimensionless system $\mathcal{B}$ are the same, technically, it is easier to solve the latter. In fact, the differential operator $\mathcal{L}$ contains fewer terms than $\mathcal{L}_{BSM}$, which results in a simpler algorithm for the computation. Another advantage is the absence of a convection term in the latter-as a result, one obtains an (easier) symmetric system to solve, rather than a nonsymmetric one.

Because an American option is of singular behaviour near the expiry $\tau=0$, the algorithm is designed to accommodate the location of the optimal exercise price at expiry, i.e., $x_f(\tau)=0$, which is known *a priori*. In contrast to the conventional numerical methods, where grid sizes are always increased in the neighborhood of $\tau=0$, in this case, the computation is straightforward. In practice, we truncate the infinite domain $(-\infty,+\infty)$ into a finite domain $[0,x_{\max}]$. Thus, the option price shall vary from $u=0$ at the rotating surface $x_{\min}=0$ to $u=\exp(\frac{x_{\max}}{2}(q_\delta-1))-\exp(\frac{x_{\max}}{2}(q_\delta+1))$ (far-field boundary condition) at $x=x_{\max}$.

Moreover, since $\max(\exp(\frac{x}{2}(q_\delta-1))-\exp(\frac{x}{2}(q_\delta+1)),0)\geq 0$ when x is in the range $0\leq x\leq x_{\max}$, the initial condition in (4.7) can be simplified as

$$u(x,0)=\exp(\frac{x}{2}(q_\delta-1))-\exp(\frac{x}{2}(q_\delta+1)).$$

It should be remarked that the real far-field boundary condition is $u(x,\tau)=\exp(\frac{x}{2}(q_\delta-$

$1)) - \exp(\frac{x}{2}(q_\delta + 1))$ when $x \to \infty$. For computational purposes, we have adopted $u(x,\tau) = \exp(\frac{x_{\max}}{2}(q_\delta - 1)) - \exp(\frac{x_{\max}}{2}(q_\delta + 1))$. However, the truncation point $x_{\max}$ has to be sufficiently far in order to avoid excessive error due to the truncation. On the other hand, unnecessarily large value of $x_{\max}$ increases computational cost. Based on Wilmott et al.'s estimate (see [106]), we set $x_{\max} = \ln 5$. In addition, since the optimal exercise price S_f is equal to the strike price K at the expiration T, as shown by [68], we must have the transformed optimal exercise price $x_f(0) = 0$ (i.e location of the free boundary at expiry). As a result, we set $x_{\min} = 0$. This automatically remove the singularity, and hence the solution corresponds to the ideal constitutive model without any regularization [4].

In what follows, for the sake of completeness and convenience for the readers, we briefly discuss the variational inequality for system $\mathcal{B}$. Here, we derive an appropriate variational formulation of problem (4.7). As noted previously, the solution of a PDE in the variational formulation setting only requires the solution and its derivatives to be square integrable, and a modification of the function spaces.

Let $L_2(\Omega)$ be the usual space of Lebesgue measurable and square integrable functions on $\Omega = [0, x_{\max}]$ and denote by $H_0^1(\Omega)$ the Sobolev space of first-order weak derivatives. We define $\mathcal{K} \subset H_0^1(\Omega)$ as

$$\mathcal{K} := \{v \in H_1(\Omega) : v \geq \psi,\ v(\partial\Omega) = \psi(\partial\Omega)\}, \tag{4.8}$$

where the inequality sign means to hold pointwise $\forall\ x \in \Omega$. Let $v \in \mathcal{K}$ be any test function and ψ defined as in (4.7). With $u(x,\tau)$ being the solution of (4.7), the regularity requirements on $u(x,\tau)$ imply that $u \in \mathcal{K}$. For all $v \in \mathcal{K}$, we have $v - \psi \geq 0$, and in view of $\mathcal{L}u \geq 0$, it is not difficult to show that

$$\int_0^{x_{\max}} \mathcal{L}u.(v - \psi)\, dx \geq 0.$$

Also, from (4.7) we have

$$\int_0^{x_{\max}} \mathcal{L}u.(u - \psi)\, dx = 0.$$

Subtraction of the last two equations yields

$$\int_0^{x_{\max}} \mathcal{L}u.(v-u)dx \geq 0, \tag{4.9}$$

thereby eliminating ψ. By the definition of the differential operator $\mathcal{L}$ and since v and u cancel out on the boundary $\partial\Omega$, integrating (4.9) by parts gives the formulation as variational inequality problem

$$\begin{cases} \text{find } u \in \mathcal{K}, \text{ such that } \forall\ v \in \mathcal{K} \text{ and } 0 \leq \tau \leq \dfrac{\sigma^2 T}{2}, \\ \\ \displaystyle\int_\Omega \left(\frac{\partial u}{\partial \tau}(v-u) + \frac{\partial u}{\partial x}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial x} \right) \right) dx \geq 0, \end{cases} \tag{4.10}$$

in addition to the initial and boundary conditions. Theoretically, variational inequality relaxes the regularity conditions on the option price. Thus, one would get a *weak* solution in contrary to the *classical* or *strong* case. From treatments on variational inequalities, see e.g., [38, 70, 110], the problem (4.10) has a unique solution by a generalized theorem Lax-Milgram. More details on the variational formulation of parabolic PDEs associated with diffusion processes can be found in [28, 61, 72], with the relevant functional analytic background.

4.3 Inverse finite element method

The iFEM involves the use of simulated finite elements to inversely predict desired quantities that are spatially varying with time. Any assumptions included in the finite element model and in the whole simulation of the experiment determine the quality of the inverse solution [67]. To proceed with the formulation of iFEM for LCP arising from an American option contract, we require the boundaries of elements to remain on "isotherms" of the underlying such that the option value is specified *a priori* everywhere in the domain. Therefore, the option is constant along the boundaries of unknowns locations, which are permitted to change as the iteration proceeds. Since the nodal option values are fixed, the remaining unknowns are the positions of the nodes. Thus, Equation (4.10) is solved by linearization with respect to these unknowns in order to form the Jacobian of the Newton iteration. In other words, the procedure seeks to find the location of the nodes which

correspond to a predefined option price. Thus, the spatial co-ordinate of the nodal points becomes the dependent variable whereas the option price is treated as the independent variable, thereby avoiding the inversion of matrices and subsequently computational cost.

Having discussed some of the fundamental issues in Section 4.2, we now focus on the iFEM implementation. The first step is to deal with the time derivative appearing in variational problem (4.10). In contrast to the conventional pricing methods where $\frac{\partial u}{\partial \tau}$ is approximated by a finite difference scheme in most cases, here, we decompose it into the hedge parameter delta, $\frac{\partial u}{\partial x}$ and the velocity of the mesh, $\frac{\partial x}{\partial \tau}$. Using the concept of iFEM, the option price u is obtained at selected underlying price which varies with respect to time, and therefore, we obtain

$$\frac{du}{d\tau} = \frac{\partial u}{\partial \tau} + \frac{\partial u}{\partial x}\frac{\partial x}{\partial \tau}, \tag{4.11}$$

where $\frac{du}{d\tau}$ is a total derivative, i.e. is the rate of change of the option price at a node. However, since option price is distributed and kept constant at all times at the computational nodes, $\frac{du}{d\tau} = 0$. Moreover, it is obvious that in this case the mesh is not fixed but moves with velocity $V_{mesh} = \frac{dx}{d\tau}$. Therefore,

$$\frac{\partial u}{\partial \tau} = -\frac{\partial u}{\partial x} V_{mesh}. \tag{4.12}$$

With $\frac{du}{d\tau} = 0$, the integral inequality in (4.10) becomes

$$\int_{\Omega} \left(\frac{\partial u}{\partial x}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial x} \right) - V_{mesh}\frac{\partial u}{\partial x}(v-u) \right) dx \geq 0. \tag{4.13}$$

The velocity of the mesh V_{mesh} is numerically approximated by using first order finite difference, i.e.,

$$V_{mesh} \approx Q_x = \frac{x_{\tau+\Delta\tau} - x_\tau}{\Delta\tau} \tag{4.14}$$

Adopting the relation (4.14), inequality (4.13) in bilinear form becomes:

$$a(v,u) = \int_{\Omega} \left(\frac{\partial u}{\partial x}\left[\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial x} \right) - Q_x(v-u) \right] \right) dx \geq 0 \tag{4.15}$$

The discretization of the resulting inequality (4.15) follows the classical Galerkin finite element approach using the suitable element shape functions. In this work, we consider two types of element shape functions: linear and quadratic shape functions, and implementations with other higher functions should not be difficult. After applying the selected shape functions, the computation of the element matrices and assembling of the finite element contributions to obtain the global matrices are straight forward. Finally, we specify the appropriate time-dependent underlying constraint conditions to obtain a non-singular system of inequalities:

$$\begin{cases} \text{find } x = x^{(n+1)}, \text{ such that for all } v \geq \psi \\ (v-w)^T(A - Q_{x(i)}B) \geq 0, , \text{ and } w \geq \psi, \end{cases} \tag{4.16}$$

where w are the nodal values of the entire domain, A and B are the constrained global stiffness matrix and mass matrix respectively, and $Q_{x(i)}$ is the global displacement containing the location of each element to be determined. For ease of reading, the detail formulation of (4.16) from (4.15) and the definition of A and B are provided in Appendix A.3. It should be pointed out that the inequality $w \geq \psi$ is defined componentwise and the existence and uniqueness of a solution of the problem (4.16) is guaranteed by a generalized-Milgram theorem, applied to finite-dimensional spaces.

Proposition 4.3.1. *The problem (4.16) is equivalent to the discrete linear complementary problem*

$$\begin{cases} \textit{find } x = x^{(n+1)}, \textit{ such that} \\ A - Q_{x(i)}B \geq 0, \quad w \geq \psi, \quad (w-\psi)^T(A - Q_{x(i)}B) = 0, \end{cases} \tag{4.17}$$

where u is the vector of the nodal values of the entire domain.

Proof. Recall the transpose property $(P \pm Q)^T = P^T \pm Q^T$. Thus, problem (4.16) can further be simplified by considering

$$\begin{aligned} &(v^T - w^T)(A - Q_xB) \geq 0 \quad \Leftrightarrow \\ &v^T(A - Q_xB) \geq w^T(A - Q_xB) \quad \Leftrightarrow \ \forall\ v \geq \psi. \end{aligned} \tag{4.18}$$

Clearly, $A - Q_{x(i)}B \geq 0$. If it were negative in the i-th component, the inequality would not hold for i going to infinity-a contradiction. Using this and $u \geq \psi$, one obtains $(w-\psi)^T(A-$

$Q_{x(i)}) \geq 0$. On the other hand, substituting $v = \psi$ in (4.16) yields $(w-\psi)^T(A-Q_{x(i)}) \leq 0$. Combining the two inequalities leads to $(w-\psi)^T(A-Q_{x(i)}) = 0$. □

It should be pointed out that in the conventional finite element method, the unknown to be found is w. This can be solved for as zeros of a given system. However, such a system would break down along the computational region with the moving boundary. This results from the imposition of the constrained free boundary conditions which forces the constrained global matrix to be a function of the unknown free boundaries. It is numerically difficult in terms of finding the solution of the algebraic system because of the adjustment to the global matrix profile for constraint relations and matrix pivoting. However, such a nonlinear system can be solved by using some iteration methods such as projective successive overrelaxation method. The draw back of this approach could be it slow rate of convergence and efficiency issue. To avoid those draw backs, iFEM is proposed based on the concept of fixing the nodal values while studying the motion of the underlying. In other words, in the inverse formulation, while the nodal values w are kept as known constants, the location of each element need to be determined at each time step.

To complete the iFEM formulation, the resulting non-linear system of equations (4.17) together the appropriate initial and boundary conditions is solved using a Newton-Raphson scheme with its quadratic convergence characteristic. In the following section, we present the detail implementation of the solution procedures.

4.4 Numerical implementations

A key step in the iFEM applicability is the proper implementation of the solution procedures. Solving a system of equation resulting from an inverse discretization can be more computational challenging if the fundamental issue such as the monotonicity of the predefined option price is not guaranteed or the initial parameters are not properly chosen. However, in the current work, to obtain a reasonably accurate solution, the monotonicity of u is required. Otherwise, the location of each element (x coordinate) that satisfies (4.17) is not unique. This will result in difficulties in deciding the correct location for a fixed nodal value, even if the convergence of the adopted iteration scheme is guaranteed.

To establish that u is a strictly monotonically increasing function with respect to x for

$x \in [0, +\infty)$, we evaluate

$$\frac{\partial u}{\partial x} = S\frac{\partial P}{\partial S}\exp(Ax + B\tau) - uA, \tag{4.19}$$

where $A = -\frac{1}{2}(q_\delta - 1)$ and $B = -(\frac{1}{4}(q_{\delta-1})^2 + q)$ which is greater than zero because the delta of an American put option is more than -1 for $S \in (S_f, +\infty)$ and

$$q_\delta = \frac{2(r-\delta)}{\sigma^2},$$

is positive for all chosen r, δ and σ. Thus, u is strictly monotonically increasing with respect to x

The second issue that should be addressed is the proper initial guess of the unknown nodal location, since the Newton iteration scheme may converge slowly or not even converge at all, if the initial guess is far away from the real solution. For our problem, the nodal locations at the present time step are chosen as the initial guess of locations of the elements at the next time step. Clearly, for a reasonably small time interval, the nodal locations at the two adjacent time steps should not differ too much, since x is continuous with respect to τ, and the time step is sufficiently small enough.

Now, with the above crucial points in mind, the specific implementation of the Newton iteration for this time step is a follows:

1. Suppose that x_k^n is obtained after nth iteration ($n \geq 0$), we compute the residual $F(x_k^n) = A - Q_{x_k^n}B$ and the corresponding Jacobian matrix $J_F(x_k^n)$.

2. Calculate the unknown nodal locations at the $(n+1)$th iteration step through $x_k^{n+1} = x_k^n - J_F^{-1}(x_k^n)$. At each iteration, $w^n \geq \psi^n$.

3. Repeat steps 1 and 2 until $\| x_k^{n+1} - x_k^n \| < \epsilon$ is satisfied. Set the solution of the kth time step to $x_k^* = x_{k-1}^{n+1}$, which completes the Newton iteration for the kth time step.

Note that the derivatives of $F(x_k^n)$ are obtained with respect to the unknown nodal location x^n and the Jacobian of the Newton-Raphson procedure is saved using an element-by-element storage. For converged results, usually two to three iterations in the Newton-Raphson procedure are necessary at each time step and the solution advances to the next time step when all unknowns converge to the stopping criterion set to a relative error of 10^{-7}. Furthermore, in the above algorithm, the location of the fixed boundary should be

excluded from the computation, since it is already the solution of the corresponding nodal value, and no iteration is further needed.

The above algorithm is defined whenever $J_F(x_k^n)$ exists. The scheme has very attractive theoretical and practical property: if x_k^* is a solution of (4.17) at which $J_F(x_k^*)$ is nonsingular, and suppose $J_F(x_k^n)$ satisfies the Lipshitz condition

$$\| J_F(x_k^n) - J_F(x_k^*) \| \leq L \| F(x_k^n) - F(x_k^*) \|, \tag{4.20}$$

for all x_k^n close enough to x_k^*, the error at iteration $k+1$ is proportional to the square of the error at iteration k, and thus the convergence is quadratic.

In what follows, we shall consider the uniqueness of the numerical results. We shall assume that the option price $u(x,\tau)$ and the payoff function $\psi(x)$ are sufficiently smooth.

Lemma 4.4.1. *When the sizes of the time step and the elements are sufficiently small, the inequality* $\frac{\partial V_{mesh}}{\partial x} \geq 0$ *holds.*

It should be mentioned that the details proof of above lemma first appeared in [27, 117]. It is reproduced here for the completeness of this work.

Proof. According to the iFEM formulation, it is known that in each element, V_{mesh} is defined as the linear interpolation between the velocities at the nodes belonging to this element. In particular, in the i-th element,

$$V_{\text{mesh}} = \frac{x_{i+1} - x}{x_{i+1} - x_i} V_i + \frac{x - x_i}{x_{i+1} - x_i} V_{i+1},$$

and thus

$$\frac{\partial V_{\text{mesh}}}{\partial x} = \frac{V_{i+1} - V_i}{x_{i+1} - x_i} = \frac{(x_{i+1}^{\triangle\tau} - x_i^{\triangle\tau}) - (x_{i+1} - x_i)}{\triangle\tau(x_{i+1} - x_i)},$$

where x_i and $x_i^{\triangle\tau}$ denote the location of the node x_i at the kth and the $(k+1)$th time step, respectively.

Since V_{mesh} is monotonically decreasing with τ, it is clear that $x_i^{\triangle\tau} + \triangle x < x_i + \triangle x$, which, combined with the fact that u is monotonically increasing with x, yields $u(x_i^{\triangle\tau} + \triangle x, \tau) < u(x_i + \triangle x, \tau)$. On the other hand, Taylor expansion shows that

$$u(x_i^{\triangle\tau} + \triangle x, \tau + \triangle\tau) \to u(x_i^{\triangle\tau} + \triangle x, \tau), \text{ as } \triangle\tau \to 0.$$

Therefore, it is clear that as $\triangle\tau \to 0$,

$$u(x_i^{\triangle\tau} + \triangle x, \tau + \triangle\tau) \leq u(x_i + \triangle x, \tau). \tag{4.21}$$

Now, according to the essence of the iFEM, it is known that $u(x_i^{\triangle\tau}, \tau + \triangle\tau) = u(x_i, \tau)$, which, combined with (4.21), yields

$$\lim_{\triangle x \to 0} \frac{u(x_i^{\triangle\tau} + \triangle x, \tau + \triangle\tau) - u(x_i^{\triangle\tau}, \tau + \triangle\tau)}{\triangle x} \leq \lim_{\triangle x \to 0} \frac{u(x_{i+1} + \triangle x, \tau) - u(x_i, \tau)}{\triangle x}, \tag{4.22}$$

and therefore

$$\left(\frac{\partial u}{\partial x}\right)(x_i^{\triangle\tau}, \tau + \triangle\tau) \leq \left(\frac{\partial u}{\partial x}\right)(x_i, \tau).$$

Consequently, in the limit sense of the element sizes approaching zero, it is clear that

$$\frac{u(x_{i+1}^{\triangle\tau}, \tau + \triangle\tau) - u(x_i^{\triangle\tau}, \tau + \triangle\tau)}{x_{i+1}^{\triangle\tau} - x_i^{\triangle\tau}} \leq \frac{u(x_{i+1}, \tau) - u(x_i, \tau)}{x_{i+1} - x_i}. \tag{4.23}$$

On the other hand, according to the iFEM, it is known that $u(x_{i+1}^{\triangle\tau}, \tau + \triangle\tau) = u(x_{i+1}, \tau)$ and $u(x_i^{\triangle\tau}, \tau + \triangle\tau) = u(x_i, \tau)$, which, combined with (4.23), yield $x_{i+1}^{\triangle\tau} - x_i^{\triangle\tau} \geq x_{i+1} - x_i$. Therefore, when the sizes of the time step and the elements are sufficiently small, $\frac{\partial V_{\text{mesh}}}{\partial x} \geq 0$ holds. □

Theorem 4.4.2. *$a(.,.)$ in (4.15) is bounded and is a continuous $\mathcal{H}^1$-elliptic bilinear form.*

Proof. According to the definition of $a(\cdot,\cdot)$, it is clear that for all $\phi \in \mathcal{H}^1(\Omega)$,

$$a(\phi,\phi) = \int_\Omega \left(\frac{\partial \phi}{\partial x}\right)^2 d\Omega - \int_\Omega \frac{\partial V_{\text{mesh}}}{\partial x}\phi^2 \, d\Omega \geq L \int_\Omega \left[\left(\frac{\partial \phi}{\partial x}\right)^2 + \phi^2\right] d\Omega, = L\|\phi^2\|,$$

where L is a positive constant. Thanks to Lemma 4.4.1, since $\frac{\partial V_{\text{mesh}}}{\partial x} \geq 0$. Thus, the bilinear defined in (4.15) is bounded. Moreover, $\forall\ \varphi, \phi \in \mathcal{H}^1(\Omega)$,

$$a(\varphi,\phi) = \int_\Omega \frac{\partial \varphi}{\partial x}\frac{\partial \phi}{\partial x} \, d\Omega - \int_\Omega Q_x \frac{\partial \varphi}{\partial x}\phi \, d\Omega \leq \|\varphi\|_1 \|\phi\|_1 (1 + \|Q_x\|_{\Omega,\infty}).$$

Therefore, $a(\cdot,\cdot)$ is in a continuous $\mathcal{H}^1$-elliptic bilinear form, provided that Q_x is ∞-measurable on the Ω, which is the case here. □

According to Theorem 4.4.2, it is known in conjunction with the generalized Lax-Migram theorem [28, 72], the variational inequality (4.15) has a unique solution.

Finally, we remark that the convergence of the adopted iterative scheme (Newton's scheme) had been discussed in the previous chapter, hence, this aspect is left out here.

4.5 Numerical results

In this section we analyze the efficiency of the inverse finite element algorithm under linear complementarity formulation relative to the popular Zhu's analytical result [113] and the algorithm due to Zhu and Chen [117]. The example chosen for numerical tests had been used for the discussing American puts on an asset without any dividend payment [113, 117]. The relevant parameters are: the strike price $K = 100$, the risk-free interest rate $r = 10\%$, the volatility of the underlying asset $\sigma = 30\%$ and the tenure of the contract being $T =$ 1 year. First, we focus only on the zero-dividend case, i.e., we set the constant dividend yield, δ to zero. The model parameters are chosen for consistency with the referenced works. Also, a comparison with previously published results may give readers a sense of verification of the current approach. For convenience, all results presented are those associated with the original dimensional quantities before the normalization process was introduced.

In order to numerically study the performance of the iFEM presented in this work, the best way is to calculate the option price and the optimal exercise price, and compare our results with some existing works. However, since the optimal exercise price is far more difficult to be accurately calculated than the option price, we shall focus on the comparison of $S_f(\tau)$.

Such a comparison is shown in Figure 4.1, where the optimal exercise prices are displayed against the benchmark Zhu's analytical results. We compare our results with those obtained by Zhu and Chen [117] using the same iFEM but under different formulations. The key feature of Zhu and Chen formulation is that the optimal exercise price is found explicitly in the process of solving the governing PDE, in contrast to implicit location of $S_f(\tau)$ in the linear complementarity formulation in this chapter. In our approach, we have adopted the linear and quadratic shape functions for element discretization. Moreover, Zhu's analytical solutions are used as the benchmark solutions. Figure 4.1 shows the

graph of the optimal exercise prices versus time to expiry. As can be clearly seen from this figure, both schemes possess good convergence attribute and indeed converge to the exact solution. The good agreement between our current scheme and the result due to Zhu and Chen is as expected since we have assumed that the underlying asset pay no dividend by setting $\delta = 0$, the two results should naturally the same when both schemes are accurately implemented.

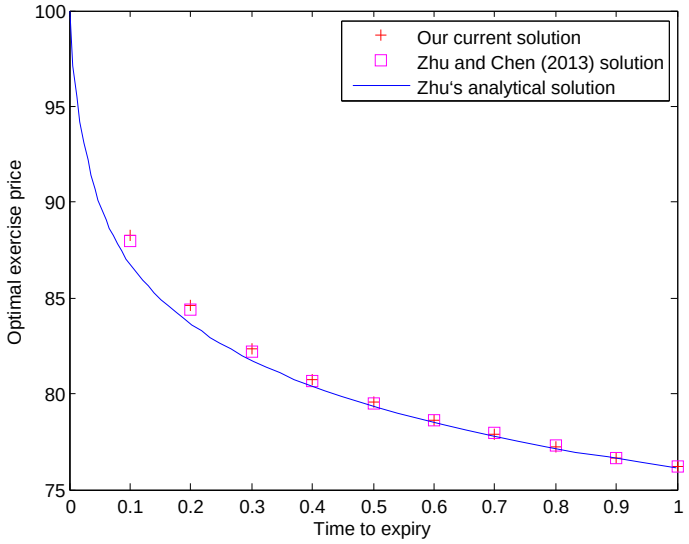

Figure 4.1: Comparison of S_f under different formulations

To better reflect the option problems traded in today's financial markets, we consider an American option problem with continuous dividend payment on the underlying asset. The relevant option parameters used in the following example are the same as those used in the zero-dividend case, except the dividend yield δ is made to be 5% and volatility of underlying asset σ are $30\%, 35\%$ and 40%, respectively. These parameters are chosen in such a way that the risk-free interest rate r is higher than the dividend yield δ. This is to avoid the parabolic-logarithm behavior associated with the optimal exercise price when the time is close to expiry [39]. Figure 4.2 demonstrates the effects of the volatility on the optimal exercise price. This graph was produced by N=45 and M=50. These numbers do not give good results although they do illustrate the trend of changing as volatility changes. The graph shows that at a given time to expiry τ, the optimal exercise price is higher for smaller values of volatility. In other words, the optimal exercise price decreases with volatility.

So far, we have presented some detailed discussions on the optimal exercise prices.

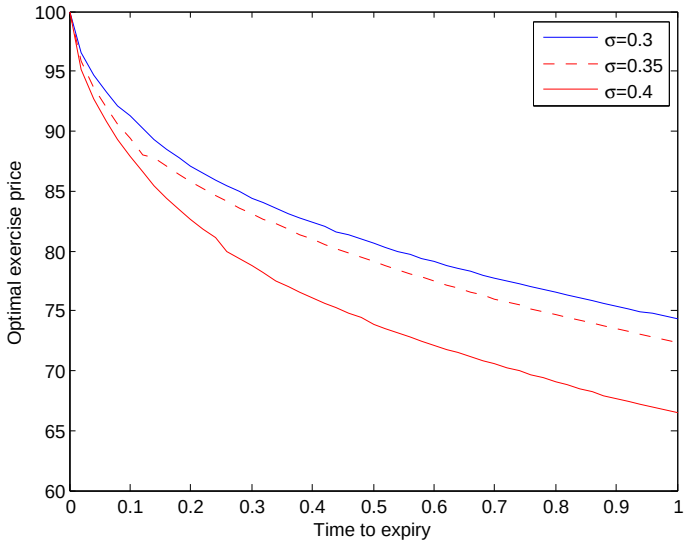

Figure 4.2: S_f for dividend yield $\delta = 0.05$ and different volatility with $N = 45$, $M = 50$

However, some readers may prefer to see how accurately the option price of an American put option can be calculated using the proposed scheme. Depicted in Figure 4.3 is the option price, $P(S, \tau)$ as a function of S with constant dividend yield $\delta = 5\%$ while all other parameters remained the same. As shown, the option price decreases with the asset values. In addition, the "smooth pasty" conditions across the optimal exercise boundary, which are usually difficult to implement numerically, are also satisfied well.

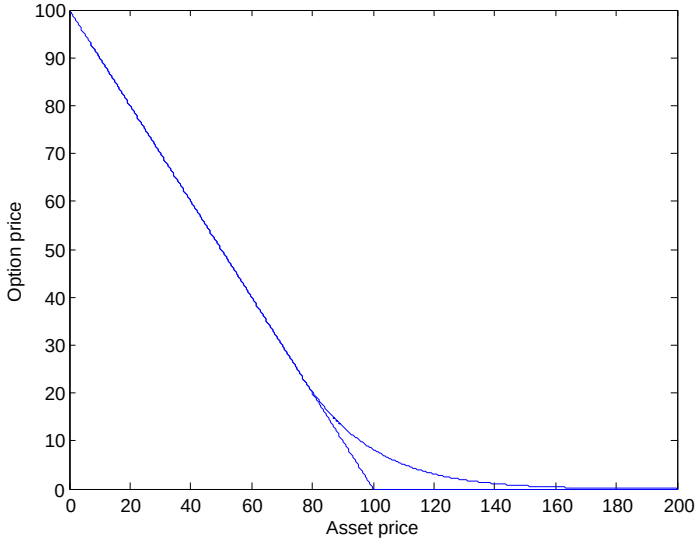

Figure 4.3: American put option value at $\delta = 5\%$

In Figure 4.4, we show the option price, $P(S, \tau)$ as a function of S with constant dividend yield, $\delta = 5\%$ and grid resolution $N = M = 50$ at three instants: $\tau = 0.5$ (year),

$\tau = 0.25$ (year) and $\tau = 0.1$ (year). Note that all other parameters remained the same as above. Clearly, as it gets closer to the expiration of the option, i.e., τ approaches 0, the option price becomes closer to the payoff function, $\max(K - S, 0)$.

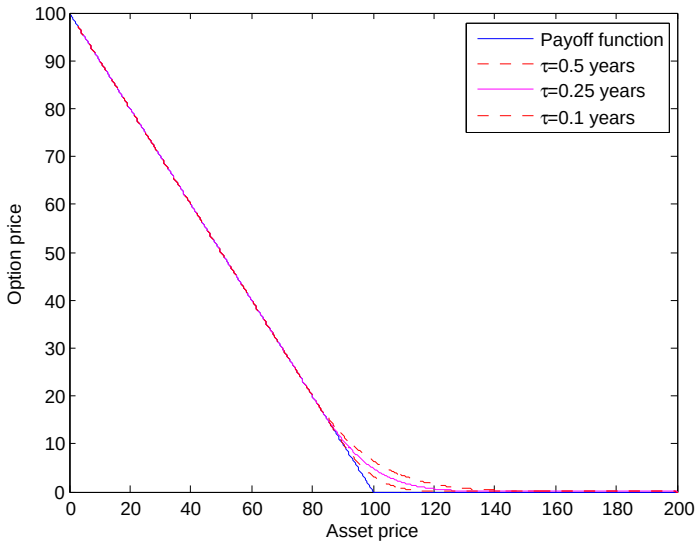

Figure 4.4: Option prices at different times to expiration.

Accuracy versus efficiency.

In quantitative finance, quite often, the computational cost is as importance as the accuracy, although it is generally a rule of thumb that efficiency is inversely proportional to accuracy [119]. In other words, when one wishes to achieve a high computational efficiency, a degree of sacrifice is suffered by the accuracy. The key question is, however, whether or not one can achieve a high efficiency with a still reasonably satisfactory accuracy. In this section, we shall demonstrate through that the relationship between accuracy and efficiency of the iFEM scheme in the current work is an inverse proportionality in nature. This is not unexpected. We shall also show that with an acceptable accuracy, say less than 2% measured in RMSRE, we can still achieve a less than 10 second execution time measured in CPU time. All the experiments in the following are performed with Matlab 7.5 on an Intel Pentium 4, 3 GHz machine.

In order to illustrate the overall performance of the iFEM, we use the $RMSRE$ (root mean square relative error), which is defined as

$$RMSRE = \sqrt{\frac{1}{I}\sum_{i=1}^{I}\left(\frac{(a_i - \bar{a}_i)}{a_i}\right)^2}$$

where $\bar{a}_i's$ are the nodal values of the S_f associated with iFEM, a_i's are the S_f obtained from the Zhu's analytical result and I is the number of sample points used in the $RMSRE$. In our numerical experiments, I was set to 50 in all the results presented. With the $RMSRE$, the overall difference between the computed numerical results and the exact solution based on Zhu's analytical result can be clearly demonstrated.

Figure 4.5 shows the variation of RMSRE as a function of computational cost for each run. The accuracy is measured by the RMSRE, which is calculated using Zhus analytical solution [113] as the base value. The computational efficiency is measured by the computational cost for each run. Three sets of computational cost as a function of RMSREs for three fixed numbers of grid points in the asset direction, i.e., with M being 40, 25 and 20 respectively, are plotted in this figure. As depicted in Figure 4.5, the accuracy is in general inversely varying with the efficiency; a higher accuracy usually implies a lower efficiency for any grid resolution. It is also generally true that a better resolution in the underlying asset values implies more computational cost. An important feature of the current scheme is that a high computational efficiency can be achieved with a satisfactory accuracy.

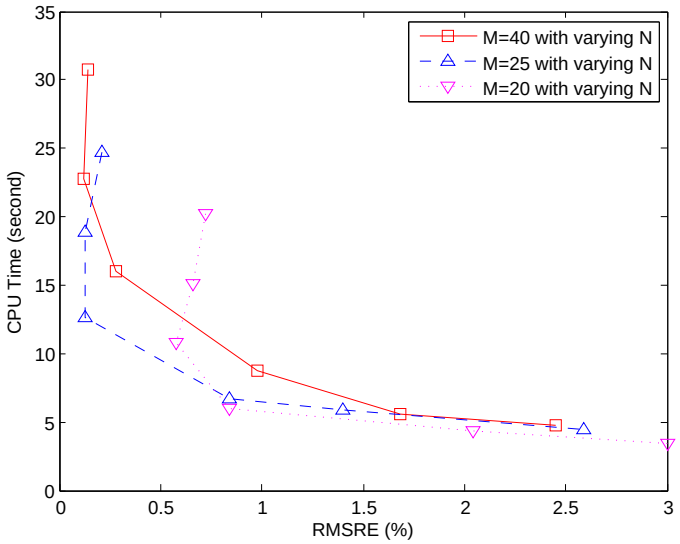

Figure 4.5: Accuracy versus efficiency

4.6 Conclusion

In this chapter we have explored the inverse finite element method for the numerical valuation of an American option under the linear complementarity formulation. The accuracy of the numerical solution has been shown to depend largely on the element shape functions adopted in the development of the algorithm, as expected. The key feature of the iFEM presented in this work, in comparison with other classical numerical methods in the literature is that the solution is limited to the yielded part of the option, and hence, the solution corresponds to the original pricing model without any co-ordinate transformation or regularization. However, it enjoys high efficiency as a result of using the Newton iterative scheme with its inherent quadratic convergence. Through a couple of numerical experiments, we have demonstrated the overall performance of the iFEM. In a subsequent chapter, the suitability of iFEM for option problem under stochastic volatility model will be examined.

Chapter 5

A hybrid approach for pricing American options under the Heston model

5.1 Introduction

In the previous chapters we studied the use of inverse finite element method to price American options written on Black-Scholes model. It is, however, well known in financial practices that the constant volatility Black-Scholes model cannot account for the 'volatility smile' which is observed in market prices for financial derivatives. One widely used method to account for the smile effect is to assume that the volatility of the underlying price is a stochastic process rather than a constant. An interesting question is whether or not the inverse finite element method (iFEM) can be extended to option problems with stochastic volatility model. In the current chapter, an attempt will be made to answer this and many more questions.

The Black-Scholes partial differential equation [17, 80] laid the foundations for contemporary derivatives pricing. However, the assumptions made in the Black-Scholes model are now known to be overly restrictive. In particular, the Black-Scholes model assume that the underlying asset price follows a geometric Brownian motion with a constant volatility. The aftermath of the 1987 global financial crisis was the empirical evidence and financial reasoning which revealed that stock return distribution manifest skewness, kutorsis and regularly possess negative relationship with implied volatility. This conflicts with the

normality assumption made in the Black-Scholes model. As a direct consequence, many derivative pricing models have been developed subsequently that use stochastic process for the underlying asset which result in a better match to empirically observed details. Examples of more realistic stochastic processes include: jump-diffusion [82], Lévy [94], stochastic volatility (SV) [49], SV jump-diffusion [12], and also combinations of those that exhibit SV as well as jumps in both the asset price and volatility [36].

The focus of this chapter is the Heston stochastic model. It has been widely regarded as the best alternative to the Black-Scholes model because of its analytical tractability for European options. We develop an inverse algorithm, a combination of the inverse finite element method and finite differences for solving option pricing differential system written on Heston's stochastic volatility model. Our technique is similar in some respects to Alexandrou [3], although with additional finite differences to discretize the volatility derivative terms. The implementation is based on the fact that the nodal locations along the volatility direction are fixed, while working out the motion of the nodes along the underlying direction. The important advantage of this formulation is that it preserves the simplicity and flexibility of conventional finite element and finite difference methods while allowing the use of full Newton iteration scheme. We establish the convergence by reformulating the discretized problem in variational form and study the approximation of the option price.

The remainder of this chapter is organized as follows. Section 5.2 reviews the Heston stochastic volatility model, the corresponding PDEs describing American option prices, and the associated boundary conditions. In Section 5.3, we present the hybrid approach in details. Section 5.4 studies the convergence property of the current scheme. In Section 5.5, numerical examples and some analyses are presented to demonstrate the efficiency of the scheme. Concluding remarks are given in the last section.

5.2 Mathematical formulation

This Section introduces the Heston stochastic volatility model and the associated boundary conditions for American put options. Although Heston's model has been studied by a number of authors [8, 71, 118], we still describe it in reasonable detail, for the sake of completeness of the Chapter and ease of reference for the readers. Furthermore, the section

contains a brief discussion of the implementation issues concerned when the nonlinear PDE system is solved using the inverse finite element method.

The Heston stochastic model is formally defined as the system of stochastic differential equations given by

$$\begin{cases} dS_t = \mu S_t dt + \sqrt{v_t} S_t dW_1 \\ dv_t = \kappa(\eta - v_t)dt + \xi\sqrt{v_t} dW_2 \\ dW_1 dW_2 = \rho dt, \end{cases} \tag{5.1}$$

where S_t denotes the spot process at time t, v_t the variance at time t, μ is the drift rate, κ the mean reversion speed for the variance, η the mean reversion level for the variance, ξ the volatility of the variance, and $W_i, i = 1, 2$, two Brownian motions with correlation $\rho \in [-1, 1]$. The model for the volatility v_t is known in financial literature as the Cox-Ingersoll-Ross (CIR) process and in mathematical statistics as the Feller process [41].

Let $P(S, v, t)$ denote the value of an American put option, with S being the price of the underlying asset, v being the variance and t being the time. For simplicity, we assume that the underlying pays no dividend. Under the Heston Model, it can be easily shown that under the risk-neutral argument, the value of a put option P should satisfy the following PDE (see e.g [44, 49]):

$$\frac{\partial P}{\partial t} + \frac{1}{2} v S^2 \frac{\partial^2 P}{\partial S^2} + \rho\xi v S \frac{\partial^2 P}{\partial S \partial v} + \frac{1}{2}\xi^2 v \frac{\partial^2 P}{\partial v^2} + rS\frac{\partial P}{\partial S} - rP + \left(\kappa(\eta - v) - \lambda\xi\sqrt{v}\right)\frac{\partial V}{\partial v} = 0, \tag{5.2}$$

where λ is the market price of risk, r is the risk-free interest rate. In this paper, for simplicity, we set λ to zero, and the extensions to the case that λ is non-zero should be straightforward. The parameters ρ, ξ, and κ, which are included in the Heston model, provide the ability to capture observed features of the market and to produce a wide range of distributions [89]. For instance, the parameter ρ, the correlation between the log-returns and the asset volatility, affects the skewness of the distribution and hence the shape of the implied volatility surface; the parameter ξ, the volatility of the variance, affects the kurtosis of the distribution; the mean reversion parameter κ can be interpreted as representing the degree of volatility clustering. This phenomenon has been observed repeatedly in the market and various studies [8, 98] have suggested that Heston's model

is consistent with the real market.

Remarkably, Heston [49] found a closed-form solution for the price of European options satisfying Equation (5.2) with associated terminal and boundary conditions. However, the approach adopted could not be easily extended to the case of American options. Therefore, in the following, we concentrate on the most interesting case from a computational point of view: American put options satisfying Equation (5.2) for which there is no closed-form solution available. For this parabolic differential equation to be solved backwards in time, we need to specify additional constraints which uniquely determine the solution. However, it should be mentioned that the main focus in this work is to provide an efficient solution method for option problems under Heston's model. This requires that the theoretical set-up of such option pricing problem admits a unique solution. In view of this, a particular choice of constraints/boundary conditions or parameters does not affect the design of the scheme.

To this end, we denote by $G(S) := \max(K - S, 0)$, the so-called payoff function which is independent of v. If at final time T, the value of the stock S is above the strike price K, one would not execute the option; the option is without value. On the other hand, if the value of S is below K, the value of the option is $K - S$. Therefore, at the final time T, the value of the option (the terminal condition) reads:

$$P(S, v, T) = \max(K - S, 0) = G(S) \tag{5.3}$$

For the valuation of American puts, a set of appropriate boundary conditions is also needed together with the terminal condition (5.3). It is obvious that the boundary conditions with respect to S are easy to justify. They are just the same as those in the Black-Scholes model. The value of an American put option should satisfy the far-field boundary condition,

$$\lim_{S \to \infty} P(S, v, t) = 0, \tag{5.4}$$

which means that when the price of the underlying becomes extremely large, a put option becomes worthless. On the other hand, just as in the Black-Scholes model, there is a critical asset price, below or equal to which it is optimal to exercise the put option. It can be shown, under the no-arbitrage argument, that the boundary conditions at the optimal

exercise boundary $S = S_f$ are [106]:

$$\begin{aligned} P(S_f, v, t) &= K - S_f, \\ \frac{\partial P}{\partial S}(S_f, v, t) &= -1. \end{aligned} \tag{5.5}$$

It should be noted that the two conditions in Equation (5.5) look very similar to the case with constant volatility. However, the main difference between the constant volatility model and the stochastic volatility model lies in the fact that in the latter case, such as in the Heston model, the optimal exercise price S_f depends, in addition to time, on the dynamics of the volatility. In other words, S_f is a function of both v and t.

Next, we discuss boundary conditions along the v direction. This is an issue which still remains unclear in the literature. Even for the European case, it is still controversial whether or not Heston's analytical formula [49] does indeed satisfy the given boundary conditions along the v direction. However, an extensive treatment from both the mathematical and the financial points of view has been provided recently in [118]. We note that through consideration of the Fichera function, the necessity for a boundary condition at $v = 0$ is only present when the Feller condition, $\kappa\eta \geq \xi^2/2$, is violated. It was argued in [30] that, when required, the appropriate boundary condition to use at this boundary is the payoff function. Moreover, it was proven in [118] that even when the Feller condition is not violated, the solution should converge to the payoff function at this boundary. Therefore, without lost of generality, in the current work, we adopt the boundary conditions along the v direction as stated in [30, 118].

For $v \to 0$, we chose the boundary condition

$$\lim_{v \to 0} P(S, v, t) = \max(K - S, 0). \tag{5.6}$$

The discussion in [30, 118] supports this choice, and even Zhu and Chen [118] proposes $\lim_{v\to 0} P(S, v, t) = 0$ as a simplified version of Equation (5.6) after they successfully established that $\lim_{v\to 0} S_f(v, t) = K$.

Lemma 5.2.1. *When v approaches 0, the value of the optimal exercise price approaches the strike price K.*

Proof. The detail of proof is deferred to Appendix A.4. □

Finally, as in [118], one expects that for $v \to \infty$ the value of an American put option reaches the strike price K asymptotically, i.e.,

$$\lim_{v\to\infty} P(S, v, t) = K. \tag{5.7}$$

In summary, the properly-closed PDE system for pricing American put options under the Heston model can be written as:

$$\begin{cases} \dfrac{\partial P}{\partial t} + \dfrac{1}{2}vS^2\dfrac{\partial^2 P}{\partial S^2} + \rho\xi vS\dfrac{\partial^2 P}{\partial S\partial v} + \dfrac{1}{2}\xi^2 v\dfrac{\partial^2 P}{\partial v^2} + rS\dfrac{\partial P}{\partial S} - rP + \Big(\kappa(\eta - v)\Big)\dfrac{\partial P}{\partial v} = 0, \\ P(S, v, T) = \max(K - S, 0), \\ P(S_f(t), v, t) = K - S_f(t), \\ \dfrac{\partial P}{\partial S}(S_f(t), t) = -1, \\ \lim\limits_{S\to\infty} P(S, v, t) = 0, \\ \lim\limits_{v\to 0} P(S, v, t) = \max(K - S, 0), \\ \lim\limits_{v\to\infty} P(S, v, t) = K. \end{cases} \tag{5.8}$$

The above PDE system is defined on $S \in [S_f(v, t), \infty)$, $v \in [0, \infty)$, and $t \in [0, T]$. Furthermore, it is obvious that for each $t \in [0, T]$, there exists a stock price S for which early exercise before final time T is advantageous. One can show that these values define a continuous curve $S_f(v, t)$. It is *a priori* unknown and therefore defines a free boundary. In fact, it is the presence of $S_f(v, t)$ that has made this type of problem highly nonlinear.

To proceed with the implementation of the hybrid iFEM and FDM, the range of the option price P should be known *a priori*, and more importantly, its monotonicity is required only along the direction where the moving boundary occurs [3]. Therefore, this method is able to handle a large class of multi-factor models, such as the stochastic volatility models, by assuming that the nodal locations along the volatility direction are fixed, while working out the motion of the nodes along the underlying direction (cf. [117]). From (5.8), one can easily deduce that the option price P shall fall within $[0, K - S_f(v, t)]$ along the S direction, which varies with respect to time t, as a result of the optimal exercise price being an unknown function of t and v. On the other hand, one can clearly see that the option price along the v direction falls within $[0, K]$.

In order to overcome the difficulty of the moving boundary in the S direction and in

view of the fact that the governing differential equation in (5.8) is indeed a degenerate parabolic differential equation, the following transforms are applied:

$$x = \ln\frac{S}{K}, \quad x_f(v,\tau) = \ln\frac{S_f(v,t)}{K}, \quad u(x,v,\tau) = \frac{P+S}{K} - 1 \quad \text{and} \quad \tau = \frac{\xi^2(T-t)}{2}.$$

The transformed value of the option $u = u(x, v, \tau)$ then satisfies the transformed Heston equation:

$$\begin{cases} \dfrac{\partial u}{\partial \tau} - vq_1\dfrac{\partial^2 u}{\partial x^2} - 2v\rho\dfrac{\partial^2 u}{\partial x \partial v} - v\dfrac{\partial^2 u}{\partial v^2} + (q_1 v - q_2)\dfrac{\partial u}{\partial x} - 2q_1\kappa(\eta - v)\dfrac{\partial u}{\partial v} + q_2 u + q_2 = 0, \\ u(x,v,0) = \max(e^x - 1, 0), \\ \lim\limits_{x\to\infty} u(x,v,\tau) = e^x - 1, \\ u(x_f(v,\tau),\tau) = 0, \quad \dfrac{\partial u}{\partial x}(x_f(v,\tau),\tau) = 0, \\ \lim\limits_{v\to 0} u(x,v,\tau) = \max(e^x - 1, 0), \\ \lim\limits_{v\to\infty} u(x,v,\tau) = e^x, \end{cases} \tag{5.9}$$

where $q_1 = 1/\xi^2$ and q_2 is the relative interest rate, which is related to the original risk-free interest rate r by $q_2\xi^2/2$. Recall that all parameters except x and v are constant real values. The transformed option pricing problem (5.9) is defined on an unbounded domain $\Omega_u^\infty := [0, T\sigma^2/2] \times [x_f, +\infty) \times [0, +\infty)$ i.e., $\Omega_u^\infty := \{(x,v,\tau) \mid x \in [x_f, \infty), v \in [0,\infty), \text{and } \tau \in [0, T\sigma^2/2]\}$.

To establish that u is a strictly monotonically increasing function along the x direction for $x \in [0, +\infty)$, we evaluate

$$\frac{\partial u}{\partial x} = \left(\frac{\partial P}{\partial S} + 1\right)\frac{S}{K},$$

which is more than zero because the delta of an American put option is greater than -1 for $S \in [S_f, +\infty)$. Hence, u is strictly monotonically increasing with respect to x for $x \in (x_f, +\infty)$.

Additionally, the two adopted boundary conditions in the v direction have coincidentally shown the monotonicity of the option price with respect to v as well as its boundedness:

$$\max(e^x - 1, 0) \le u(x,v,\tau) \le e^x.$$

The option price is monotonically increasing from its lower bound $\max(e^x - 1, 0)$ to its

upper bound e^x when v varies from 0 to ∞.

Furthermore, in the PDE system (5.9), one should notice that both the initial condition and the option price as $v \to 0$ are the same. These conditions can be further simplified using different arguments. First, the optimal exercise price is equal to the strike price at the expiration time T, as shown by Zhu [113], i.e., $x_f(v, 0) = 0$ in (5.9). Then, $\max(e^x - 1, 0) \geq 0$ when $x \in [x_f, \infty)$. As a direct consequence, the initial condition can be simplified as $u(x, v, 0) = e^x - 1$. Moreover, since the optimal exercise boundary is approaching the strike price as $v \to 0$, as shown in [118], we must have $\lim_{v\to 0} x_f(v, \tau) = 0$ in (5.9). Thus, $\forall\ v \in [0, \infty)$, $\max(e^x - 1, 0) \geq 0$, which yield $\lim_{v\to 0} u(x, v, \tau) = e^x - 1$.

For numerical computation, we need to localize the unbounded domain Ω_u^∞ by defining $\Omega_u := [0, T\sigma^2/2] \times \Omega \subset \mathbb{R}^3$ with the spatial domain $\Omega := [0, x_{\max}] \times [0, v_{\max}] \subset \mathbb{R}^2$. In practice, $x_{\max}$ and $v_{\max}$ should be large enough to remove the boundary effect. However, Willmott et al.'s [106] estimated that the upper bound of the underlying price, $S_{\max}$ is three or four times the exercise price, then, it is reasonable to set $x_{\max} = \ln 5$. On the other hand, as recently pointed out by [118], the volatility value is usually very small. The highest value of the volatility that has ever been recorded on Chicago Board Options Exchange (CBOE) is only 0.85 [40]. Thus, it is quite reasonable to set $v_{\max} = 1$, and, this has also been the case in many previous studies, (e.g. [57]).

Before we conclude this section, it should be mentioned that after the transformation of variables and the truncation of the computational domain, the range of P becomes $[0, e^{x_{\max}} - 1]$ in x direction and $[0, e^x]$ in v direction, while the monotonicity is retained. This is an advantage that we can take when the inverse finite element method is applied to solve free boundary problems, as shall be discussed in the next section.

5.3 Construction of the hybrid inverse finite element method and finite difference method (hybrid iFEM/FDM)

Upon establishing a closed differential system (5.9) for the price of American put options under the Heston model, we propose a hybrid approach, a combination of the inverse finite element method (iFEM) and finite difference method (FDM). This method is based on the fact that the nodal locations along the volatility direction is fixed, while working out the motion of the nodes along the underlying direction with the iFEM. More specifically, we

discretize the pricing system along the volatility direction by finite differences, and thus, the volatility derivative terms in the governing differential equation are elegantly removed. Finally, with the use of simulated finite elements, we predict the desired quantity (optimal exercise price) which spatially varies with time. The solution of the problem through the Newton iterations then reveals the correct location of the free boundary at each time step, and thus the advancement of the free boundary as the time marches on. In what follows, the detail solution procedures are provided.

Our implementation details of hybrid iFEM/FDM, follow closely the details described in [3]. The iFEM is a technique for solving nonlinear problems involving phase change as known in mechanics [3]. The approach has been successfully used to price American option problems written on Black-Scholes model under different formulations in the previous chapters. Essentially, iFEM is implemented by requiring the boundaries of the elements to remain on "isotherms" of the underlying such that the option value can be specified a *priori* everywhere in the domain [3]. It uses the concept of fixing the option price while studying the motion of the different "isotherms" of the underlying. In other words, the spatial coordinate of the nodal points becomes the dependent variable whereas the option price is treated as the independent variable, thereby avoiding the inversion of matrices.

To implement the hybrid iFEM/FDM, the first step is to deal with the time derivative term contained in (5.9) using the concept of fixing the option value while studying the motion of underlying asset. This step is, in fact, necessary for most of the numerical schemes designed for solving time-dependent problems. Using the essence of the inverse approach, the option price u is obtained at selected underlying price which varies with respect to τ, and therefore, we obtain

$$\frac{du}{d\tau} = \frac{\partial u}{\partial \tau} + \frac{\partial u}{\partial x}\frac{\partial x}{\partial \tau} + \frac{\partial u}{\partial v}\frac{\partial v}{\partial \tau}, \tag{5.10}$$

where $\frac{du}{d\tau}$ is a total derivative, i.e. is the rate of change of the option price at a node. However, since option price is distributed and kept constant at all times at the computational nodes, hence $\frac{du}{d\tau} = 0$.

Therefore,

$$\frac{\partial u}{\partial \tau} = -\frac{\partial u}{\partial x}\frac{\partial x}{\partial \tau} - \frac{\partial u}{\partial v}\frac{\partial v}{\partial \tau}. \tag{5.11}$$

Moreover, we stress that the mesh along the x-direction is not fixed, but moves with velocity $V_1 = \frac{dx}{d\tau}$, whereas, along the v-direction, the nodal locations are fixed, and thus, the term $\frac{dv}{d\tau} := V_2$ can be determined straightforwardly.

Next, the velocity of the mesh V_1 is numerically approximated by using first order finite difference, i.e.,

$$V_1 \approx V_{\text{mesh}} = \frac{x_{\tau+\Delta\tau} - x_\tau}{\Delta\tau} \tag{5.12}$$

Adopting the approximation (5.12), the governing differential equation (5.9) becomes

$$\begin{aligned} &vq_1\frac{\partial^2 u}{\partial x^2} + 2v\rho\frac{\partial^2 u}{\partial x \partial v} + v\frac{\partial^2 u}{\partial v^2} + (q_2 - q_1 v - V_{\text{mesh}})\frac{\partial u}{\partial x} \\ &+ 2q_1\kappa(\eta - v - V_2)\frac{\partial u}{\partial v} - q_2 u - q_2 = 0. \end{aligned} \tag{5.13}$$

It should be emphasized that the numerical treatment of V_{mesh} will affect the accuracy of the final results. However, this matter is unrelated to the issue of applicability of the current approach, as the final numerical results can be improved by adopting higher order approximation which would reduce the truncation errors due to the approximation in (5.12). For the sake of simplicity, in the current work, we confine our attention to the relatively simple case of first order difference schemes, and deferring the treatment of higher order approximation formulae to future work.

Next, we discretize differential equation (5.13) along v direction using the standard discretization method to eliminates the volatility derivative terms in the governing differential equation. The discretization is performed by placing a set of uniformly distributed grids in the computation domain $[0, v_{\max}]$. With the number of steps in the v direction being denoted by N_v, the step size is defined as $\triangle v = \frac{v_{\max}}{N_v}$. The value of the unknown function u at a grid point along v direction is thus denoted by

$$u^\tau_{x,i} \approx u(x, v_i, \tau) = u(x, i\triangle v, \tau),$$

where $i = 0, \ldots, N_v$.

The discretization along the v direction needs to be conducted in the interior domain $\Omega_v = \{i\triangle v \mid i = 1 \ldots N_v - 1\}$ to approximate the first-order and second-order spatial

derivatives using the standard forward difference schemes

$$\begin{aligned}\frac{\partial u}{\partial v} &\approx A_i = \frac{u^\tau_{x,i+1} - u^\tau_{x,i}}{\triangle v},\\ \frac{\partial^2 u}{\partial v^2} &\approx B_i = \frac{u^\tau_{x,i+1} - 2u^\tau_{x,i} + u^\tau_{x,i-1}}{(\triangle v)^2},\end{aligned} \tag{5.14}$$

At the boundary $\partial\Omega_v = \{i\triangle v \mid i = 0, N_v\}$, the boundary conditions $u^\tau_{x,0} = 0$ and $u^\tau_{x,\triangle v N_v} = e^x$ are simply incorporated into the discrete equation. With Equation (5.14), the differential equation (5.13) can be written as

$$\begin{aligned}&v_i q_1 \frac{\partial^2 u_i}{\partial x^2} + 2v_i\rho\frac{\partial A_i}{\partial x} + v_i B_i + (q_2 - q_1 v - V_{\text{mesh}})\frac{\partial u_i}{\partial x}\\ &+ 2q_1\kappa(\eta - v_i - V_2)A_i - q_2 u_i - q_2 = 0,\end{aligned} \tag{5.15}$$

which is a system of PDE to be solved. Recall that all parameters except x are constant real values in Equation (5.15) for each i. At each i, we need to numerically solve Equation (5.15) for $v = v_i = ih$ where h is the constant v step. For simplicity, we drop the subscript $i's$ hereafter.

Following the traditional finite element method, a residual equation can be constructed as

$$\begin{aligned}R(x) = \int_0^{x_{\max}} \Big[&v q_1 \frac{\partial^2 u}{\partial x^2} + 2v\rho\frac{\partial A}{\partial x} + vB + (q_2 - q_1 v - V_{\text{mesh}})\frac{\partial u}{\partial x}\\ &+ 2q_1\kappa(\eta - v - V_2)A - q_2 u - q_2\Big]\varphi\, dx = 0,\end{aligned} \tag{5.16}$$

the solution of which is identical to the one of (5.15) in the weak sense. Here, φ is the trial function. The residual $R(x)$, after using the divergence theorem, becomes

$$\begin{aligned}R(x) = \int_0^{x_{\max}} \Big[&- v q_1 \frac{\partial u}{\partial x}\frac{\partial \varphi}{\partial x} + 2v\rho A\varphi - 2v\rho A\frac{\partial \varphi}{\partial x} + vB\varphi + (q_2 - q_1 v - V_{\text{mesh}})\varphi\frac{\partial u}{\partial x}\\ &+ 2q_1\kappa(\eta - v - V_2)\varphi A - q_2\varphi u - q_2\varphi\Big]\, d\Omega.\end{aligned} \tag{5.17}$$

We now proceed to discretize residual equation (5.17) with respect to the space variable x in terms of linear finite elements (see, e.g.[20]). In particular, we have considered only linear shape function, while implementation with higher order shape functions should be similar. For this purpose, we assume that the computational domain $[0, x_{\max}]$ is discretized

into N_x line elements, and each of them is mapped isoparametrically into a basic line element with limits $-1 \leq \xi \leq 1$.

Now, following a traditional "direct" FE approach, we express u in terms of a finite element basis function as

$$u = \sum_{i=1}^{p} w_i(\tau)\varphi_i(\xi) = (\varphi_i, \ldots, \varphi_p)W^{(n)}, \tag{5.18}$$

where $W^{(n)}$ is the vector of the nodal values associated with the nth element, i.e., $W^{(n)} = (W_1 \ldots W_p)^{'}$ with the subscripts being the local numbers, and p is the total number of the nodal values of this element.

By substituting (5.18) into (5.17), we obtain the matrix form for the residual of the nth element as

$$R^{(n)}(x) = k^{(n)}W^{(n)} - q^{(n)}, \tag{5.19}$$

where

$$k^{(n)}(i,j) = \int_{-1}^{1} \left(vq_1 \frac{\varphi_i^{'}\varphi_j^{'}}{J_b} + (q_2 - q_1 v - V_{\text{mesh}})\varphi_i^{'}\varphi_j - q_2\varphi_i\varphi_j J_b \right) d\xi, \quad i,j = 1 \ldots p$$

,

$$q^{(n)}(i) = \int_{-1}^{1} \left(M\varphi_i J_b + \varphi_i^{'} \right) d\xi, \quad i = 1 \ldots p$$

and $M = 2v\rho A + vB - q_2 + 2q_1\kappa(\eta - v - V_2)A$. The term J_b is the stretch factor between x and ξ coordinates, it becomes the determinant of the Jacobian matrix of the mapping between the coordinate systems. For the linear shape function, $J_b = 1/2$. Using the solid mechanics terminology, $k^{(n)}$ is the so-called element-stiffness matrix, which characterizes the behavior of the element, whereas $q^{(n)}$ is the applied (or external) element generalized-load vector, defined from the element potential energy [99].

By substituting (5.19) into (5.17), we obtain

$$R^{(n)}(x) = \sum_{n=1}^{p} \left(k^{(n)}W^{(n)} - q^{(n)} \right), \tag{5.20}$$

which, after some algebraic manipulations, can be written as

$$R = KW - Q, \tag{5.21}$$

where W is the vector of the known nodal values of the entire domain, K is the constrained master stiffness matrix involving the unknown locations of the underlying asset, and Q represent the vector of forcing term. The detail assembling of the global matrices in (5.21) from the element matrices in (5.18) are provided in Appendix B.

To find the location of the underlying asset at each time step, we modify the Newton's algorithm discussed in chapter 4. More specifically, additional loop is created to handle the finite difference discretization of the volatility derivative terms. Moreover, in other to obtain a reasonably accurate solution, the initial guess of the unknown nodal locations must be properly chosen. This is because the adopted Newton scheme may converge slowly or not even converge at all, if the initial guess is far away from the real solution. For this work, the nodal positions at the present time step are chosen as the initial locations of the elements at the next time step. clearly, for a reasonably small time interval, the nodal positions at the two adjoining time steps should not differ too much, since x is continuous with respect to τ, and the time step is sufficiently small enough [117].

5.4 The convergence of the algorithm

In this section, we present the convergence analysis of the current approach. This is achieved by providing the error estimate for the nth time steps of resulting discrete finite element solution. After implementing the finite differences which remove the volatility-dependant terms in the governing differential equation, we reformulate the resulting system as a linear complementarity problem and establish the weak convergence to the Heston model in the underlying space.

At each time step n, it is not difficult to show that the option pricing problem (5.8) after eliminating the volatility derivative terms is equivalent to the following linear com-

plementarity problem

$$\begin{cases} (\mathcal{L}u - \Lambda).(u - g) = 0, \\ u - g \geq 0, \\ \mathcal{L}u - \Lambda \geq 0, \\ u = \psi \text{ on the boundary}, \end{cases} \tag{5.22}$$

where $g = \max(e^x - 1, 0)$, $\mathcal{L}$ is a partial differential operator defined as

$$\mathcal{L} = vq_1 \frac{\partial^2}{\partial x^2} + (q_2 - q_1 v - V_{\text{mesh}}) \frac{\partial}{\partial x} - q_2 I,$$

and

$$\Lambda = q_2 - vB - 2q_1\kappa(\eta - v - V_2)A - 2v\rho\frac{\partial A}{\partial x}.$$

The problem (5.22) is defined on an infinite domain $\Omega := (\infty, +\infty)$. In practice, we cannot solve the LCP over the whole real line. So, we truncate the infinite domain into the finite domain, i.e., $\Omega_k := [x_{\min}, x_{\max}]$. To ensure consistency with the truncation described in the previous section, we set $x_{\max} = \ln 5$. As pointed out in [64], this truncation of domain will only bring in negligible error in the pricing of American options. On the other hand, $x_{\min}$ here is set to be sufficiently small to eliminate the boundary effect. It is expected that if the desired error estimate of (5.9) is finally derived through that of (5.22), the value of $x_{\min}$ will not affect the former, as a result of the domain of (5.9) being a subset of Ω_k containing $x_{\min}$. However, for symmetric purposes, some published works set $x_{\min} = -x_{\max}$, and we have adopted same in this case.

With the truncated domain Ω_k, first, we derive the equivalent variational form of (5.22) in order to obtain the desired error estimate.

Lemma 5.4.1. *(**Variational inequality**) At each time step, the linear complementary problem (5.22) is equivalent, in the weak sense, to solving for a $w \in \mathcal{K}$, such that for all $\phi \in \mathcal{K}$, the inequality $a(w, \phi - w) \geq (g, \phi - w)$ holds, where*

$$\mathcal{K} := \{\phi \in H_1(\Omega) : \phi \geq 0,\ \phi(\partial\Omega) = 0\}, \quad \Pi = q_2 - q_1 v - V_{mesh},$$

$$a(w,\phi-w) = vq_1 \int_{\Omega_k} \frac{\partial w}{\partial x}\frac{\partial(\phi-w)}{\partial x}\, d\Omega_k - \frac{1}{2}\int_{\Omega_k} \Pi\left[\frac{\partial w}{\partial x}(\phi-w) - w\frac{\partial(\phi-w)}{\partial x}\right] dx + \int_{\Omega_k}\left(q_2 + \frac{1}{2}\frac{\partial V_{mesh}}{\partial x}\right)(\phi-w)w\, dx$$

and

$$(g,\phi-w) = -vq_1\int_{\Omega_k}\frac{\partial\psi}{\partial\psi}\frac{\partial(\phi-w)}{\partial x}\, d\Omega_k + \int_{\Omega_k}\Pi\frac{\partial\psi}{\partial x}(\phi-w)\, dx - \int_{\Omega_k} q_2\psi(\phi-w)\, dx - \int_{\Omega_k}\left[q_2 - vB - 2q_1\kappa(\eta-v-V_2)A\right](\phi-w)\, dx + \int_{\Omega_k} 2v\rho A\frac{\partial(\phi-w)}{\partial x}\, dx$$

Proof. Let $L_2(\Omega_k)$ be the usual space of Lebesgue measurable and square integrable functions on $\Omega = [x_{\min}, x_{\max}]$ and denote by $H_0^1(\Omega_k)$ the Sobolev space of first-order weak derivatives. We define $\bar{\mathcal{K}} \subset H_0^1(\Omega_k)$ as

$$\bar{\mathcal{K}} := \{\phi \in H_1(\Omega) : \phi \geq g,\ \phi(x) = g(x), \forall x \in \partial\Omega\}, \tag{5.23}$$

where the inequality sign means to hold pointwise $\forall\ x \in \Omega_k$. Let $\phi \in \bar{\mathcal{K}}$ be any test function and $u \in \bar{\mathcal{K}}$ be a solution of problem (5.22). For all $\phi \in \mathcal{K}$, we have $\phi - g \geq 0$. We multiply the third equation in (5.22) by $\phi - g$ (which does not change in sign) and integrate over Ω_k, yielding $\int_{\Omega_k} \mathcal{L}u - \Lambda).(\phi - g)\, d\Omega_k \geq 0$. Subtraction of the first equation in (5.22), integrated over Ω_k, that is, $\int_{\Omega_k} (\mathcal{L}u - \Lambda).(u - g)\, d\Omega_k = 0$, yields

$$\int_{\Omega_k} (\mathcal{L}u - \Lambda).(\phi - u)dx \geq 0,$$

thereby eliminating ψ. Moreover, we apply a further transformation $w = u - g$ in order to achieve zero boundary conditions and inequality. For this, we need to assume for the moment that g is sufficiently smooth. According to this transformation, we define a new constraint space as

$$\mathcal{K} := \{\phi \in H_1(\Omega) : \phi \geq 0,\ \phi(x) = 0\} \tag{5.24}$$

Therefore, the linear complementary problem (5.22) is equivalent to finding $w \in \mathcal{K}$ with

$$\int_{\Omega_k} [\mathcal{L}(w+g) - \Lambda].(\phi - w)dx \geq 0.$$

By applying integrating by parts technique, it is now clear that the weak solution of (5.22) is the solution of the following problem. Finding $w \in \mathcal{K}$, such that for all $\phi \in \mathcal{K}$

$$\begin{aligned} &vq_1 \int_{\Omega_k} \frac{\partial w}{\partial x}\frac{\partial(\phi - w)}{\partial x}\, dx - \int_{\Omega_k} \Pi \frac{\partial w}{\partial x}(\phi - w)\, d\Omega_k + q_2 \int_{\Omega_k} w(\phi - w)\, dx \geq \\ &- vq_1 \int_{\Omega_k} \frac{\partial \psi}{\partial \psi}\frac{\partial(\phi - w)}{\partial x}\, dx + \int_{\Omega_k} \Pi \frac{\partial \psi}{\partial x}(\phi - w)\, d\Omega_k - \int_{\Omega_k} q_2 \psi(\phi - w)\, dx \\ &- \int_{\Omega_k} \Big[q_2 - vB - 2q_1\kappa(\eta - v - V_2)A\Big](\phi - w)\, d\Omega_k + \int_{\Omega_k} 2v\rho A \frac{\partial(\phi - w)}{\partial x}\, dx \end{aligned} \tag{5.25}$$

Using the fact that

$$\begin{aligned} \frac{1}{2}\int_{\Omega_k} \Pi \frac{\partial w}{\partial x}(\phi - w)\, dx &= \frac{1}{2}\Pi(\phi - w)w\,|_{\partial\Omega_k} - \frac{1}{2}\int_{\Omega_k} \Pi \frac{\partial \Pi}{\partial x}(\phi - w)w\, dx - \frac{1}{2}\int_{\Omega_k} w\Pi \frac{\partial(\phi - w)}{\partial x}\, dx \\ &= -\frac{1}{2}\int_{\Omega_k} \Pi \frac{\partial \Pi}{\partial x}(\phi - w)w\, dx - \frac{1}{2}\int_{\Omega_k} w\Pi \frac{\partial(\phi - w)}{\partial x}\, dx, \end{aligned} \tag{5.26}$$

the left hand side of (5.26) can be written as

$$vq_1 \int_{\Omega_k} \frac{\partial w}{\partial x}\frac{\partial(\phi - w)}{\partial x}\, d\Omega_k - \frac{1}{2}\int_{\Omega_k} \Pi\left[\frac{\partial w}{\partial x}(\phi - w) - w\frac{\partial(\phi - w)}{\partial x}\right] d\Omega_k + \int_{\Omega_k} \left(q_2 + \frac{1}{2}\frac{\partial Q_x}{\partial x}\right)(\phi - w)w\, d\Omega_k$$

Therefore, at the nth time step, the linear complementary problem (5.22) becomes: solve for $w \in \mathcal{K}$, such that for all $\phi \in \mathcal{K}$, the inequality $a(w, \phi - w) \geq (g, \phi - w)$ is always satisfied. □

It should be mentioned that related proof of the above theorem can be found in [117]

Lemma 5.4.2. *When the sizes of the time step and the elements are sufficiently small, the inequality* $\frac{\partial V_{mesh}}{\partial x} \geq 0$ *holds.*

Proof. The proof is similar to that of Lemma (4.4.1). The details proof recently appeared in [117]. □

Following the reasoning by [117], Lemma 5.4.2 is proved with the condition that both the temporal and spacial step sizes are approaching zero. Thus, it is numerical difficult to identify the δ_1-neighborhood and δ_2-neighborhood of zero for Δt and Δx, respectively. Note that δ_1 and δ_2 are two small quantities, in order to validate the inequality that $\frac{\partial V_{\text{mesh}}}{\partial x} \geq 0$. Using numerical experiments, one can show that $\frac{\partial V_{\text{mesh}}}{\partial x} \geq 0$ is satisfied at least for the initialization of the time step and element sizes we have chosen. As a result of Lemma 5.4.2, the ellipticity of the bilinear form $a(\cdot,\cdot)$ and the boundness of both $a(\cdot,\cdot)$ and (ψ,ϕ) can be achieved, as shown in the subsequent theorem. The ellipticity and the boundness associated with $a(\cdot,\cdot)$ and (ψ,ϕ), respectively, are of essential importance for establishing the error estimate of the finite element solution of problem (5.22) (cf. [117]).

Theorem 5.4.3. *$a(\cdot,\cdot)$ is a continuous $\mathcal{H}^1$-elliptic bilinear form and (g,ϕ) is bounded.*

It should be mentioned that similar proof first appeared in [88] and recently in [117]. It is reproduced here for the completeness of this work.

Proof. According to the definition of $a(\cdot,\cdot)$, it is clear that for all $\phi \in \mathcal{H}^1(\Omega)$,

$$a(\phi,\phi) = v q_1 \int_\Omega \left(\frac{\partial \phi}{\partial x}\right)^2 dx + \int_\Omega \left(q_2 + \frac{1}{2}\frac{\partial V_{\text{mesh}}}{\partial x}\right)\phi^2 \, dx \geq L \int_\Omega \left[v q_1 \left(\frac{\partial \phi}{\partial x}\right)^2 + \phi^2\right] dx, =$$
$$L\|\phi^2\|, \left(\text{since } q_2 \text{ and } \frac{\partial V_{\text{mesh}}}{\partial x} \geq 0\right),$$

where L is a positive constant.
Moreover, $\forall\ \varphi, \phi \in \mathcal{H}^1(\Omega)$,

$$a(\varphi,\phi) = \int_\Omega \frac{\partial \varphi}{\partial x}\frac{\partial \phi}{\partial x}\, d\Omega - \int_\Omega M \frac{\partial \varphi}{\partial x}\phi \, dx + \int_\Omega q_2 \varphi \phi \, dx \geq \|\varphi\|_1 \|\phi\|_1 (1 + \|M\|_{\Omega,\infty} + q_2).$$

Therefore, $a(\cdot,\cdot)$ is in a continuous $\mathcal{H}^1$-elliptic bilinear form, provided that M is ∞-measurable on the Ω, which is the case here. On the other hand,

$$(g,\phi) = -\int_\Omega q_2 \phi \, dx + \int_\Omega M\phi \, dx - \int_\Omega \frac{\partial g}{\partial x}\frac{\partial \phi}{\partial x} \, dx +$$
$$\int_\Omega \Pi \frac{\partial g}{\partial x}\phi \, dx - \int_\Omega q_2 g \phi \, dx,$$

□

where $N = vB + 2Av\rho 2 + 2q1k[\eta - v - V2]A$ and $\Pi = q_2 - q_1 v - V_{\text{mesh}}$.

According to Theorem 5.4.3, it is known, in conjunction with the generalized LaxMigram theorem [28, 72], that the linear complementary problem (5.22) has a unique solution.

Next, with the linear elements we have adopted to discretize (5.22) with respect to the space variable x, Ω has been decomposed into uniform line segment with length proportional to a parameter h. We define $V_h \in \mathcal{H}^1(\Omega)$ as the finite element space spanned by the one-dimensional linear basis functions and has vanishing boundary values at the boundary.

Moreover, let

$$\bar{\mathcal{K}}_{\langle} := \{\phi_h \in V_h, \phi_h \geq 0, \text{and } \phi_h = 0 \text{ on } \partial\Omega\}.$$

It is not difficult to show that the discrete version of (5.22) is to find $w_h \in \bar{\mathcal{K}}_{\langle}$, such that $\forall \phi \in \bar{\mathcal{K}}_{\langle}$, $a(w_h, \phi_h - w_h) \geq (g, \phi_h - w_h)$ holds pointwise. Again, the existence and uniqueness of the discrete solution is guaranteed by the generalized LaxMilgram theorem, applied to finite dimensional spaces (see Ciarlet [28]).

On the other hand, the error analysis for the discrete solution of (5.22) is $\| w - w_h \|_{\mathcal{H}^1(\Omega)} = \mathcal{O}(h)$ as long as $g \in L^2(\Omega)$. This reasoning follows that of Brezzi et al. [22], where similar conditions exist. Thus, with respect to the $L^2(\Omega)$ norm, and using the explicit Euler algorithm, the error analysis for the discrete solution of the problem (5.22):

$$\| w - w_h \|_{\mathrm{L}^2[\Omega\times(0,T))]} = \mathcal{O}(h^2 + \triangle t) \tag{5.27}$$

5.5 Numerical results and discussion

This section presents numerical results in order to assess the efficiency of the hybrid iFEM/FDM for pricing American options written on Heston's model. First, we investigate whether the computed solution can be validated. Since there is no exact solution of Problem 5.8 in closed form, we compare the option price from the current hybrid approach with some of the existing numerical solutions. In addition, the section presents some graphical results to illustrate the effects of time-dependent volatility on the optimal exercise boundary of American options.

First, we compare the performance of the hybrid iFEM/FDM with various finite difference methods' pricing results given in Clark & Parrott [30], Zvan et al. [123], Oosterlee [84]

and Zhu and Chen [118]. The model parameters are presented in Table 5.1. Note that the parameters are chosen in order to permit direct comparison with the reference values provided in [30, 84, 118, 123].

Table 5.1: Model parameters

Strike price K (\$)	10
Interest rate r	0.1
Correlation parameter ρ	0.1
Mean reversion level η	0.16
Expiry time T	0.25
Mean reversion rate k	5
Initial stock prices S_0 (\$)	8,. . . ,12

In the current hybrid approach, we fix the number of points in the v coordinate as $N_v = 50$, with varying number of time and space steps: $N_\tau = N_x = 50, 75, 100$. We calculate two sets of American put options with different parameters using the hybrid iFEM and FDM. These prices are presented in Tables 5.2 and 5.3 at five stock prices $S = 8, 9, 10, 11, 12$, and for variance values $v_0 = 0.25$ and $v_0 = 0.0625$. We have used different discretization grids in order to study the accuracy of the numerical solutions. The prices reported in [30, 84, 118] are also shown in these tables for comparison. It can be seen that even with the most coarse grid $N_\tau = N_x = 50$, the error is only about 10^{-2}. Furthermore, the prices obtained with the finest grid are reasonably close to the ones in [30, 84, 118], and the error is about 10^{-3}. This confirms that our numerical solution does converge to that of the original nonlinear PDE.

Table 5.2: Comparison of the computed option prices with the reference solutions at $v_0 = 0.25$

N_x, N_τ	$S_0 = 8$	$S_0 = 9$	$S_0 = 10$	$S_0 = 11$	$S_0 = 12$
50, 50	2.0968	1.4581	0.8918	0.5493	0.2932
75, 50	2.0444	1.3326	0.7941	0.4470	0.2423
75, 75	2.0780	1.3329	0.7952	0.4477	0.2426
100, 75	2.0781	1.3333	0.7956	0.4480	0.2427
100, 100	2.0783	1.3347	0.7958	0.4481	0.2428
Ref. [118]	2.0781	1.3337	0.7965	0.4496	0.2441
Ref. [30]	2.0733	1.3290	0.7992	0.4536	0.2502
Ref. [84]	2.0790	1.3340	0.7960	0.4490	0.2430
Ref. [123]	2.0784	1.3337	0.7961	0.4483	0.2428

Next, with the hybrid iFEM/FDM solution, we graph the option value versus the stock

Table 5.3: Comparison of the computed option prices with the reference solutions at $v_0 = 0.0625$

N_x, N_τ	$S_0 = 8$	$S_0 = 9$	$S_0 = 10$	$S_0 = 11$	$S_0 = 12$
$50, 50$	1.9829	1.1067	0.5190	0.2136	0.0818
$75, 50$	1.9836	1.1075	0.5193	0.2136	0.0821
$75, 75$	1.9979	1.1075	0.5199	0.2136	0.0820
$100, 75$	2.0000	1.1076	0.5199	0.2136	0.0821
$100, 100$	2.0000	1.1076	0.5200	0.2137	0.0821
Ref. [118]	2.0000	1.0987	0.5082	0.2106	0.0861
Ref. [30]	2.0000	1.1080	0.5316	0.2261	0.0907
Ref. [84]	2.0000	1.1070	0.5170	0.2120	0.0815
Ref. [123]	2.0000	1.1076	0.5202	0.2138	0.0821

price at different time to expiry as some readers may wish to see the result in graphical forms. Depicted in Figure 1 is the option price $P(S, v, \tau)$ as a function of S with fixed variance $v = 0.25$ at four instants, $\tau = 0$, $\tau = 0.1$ (Year), $\tau = 0.5$ (Year) and $\tau = 1$ (Year). Clearly, the option price is a decreasing function of stock price. As it gets closer to the expiration of the option, i.e. $\tau = 0$ or $t = 1$ (year), the option price becomes closer to the payoff function $\max(K - S, 0)$. In fact, when $\tau = 0$, the option price is just the S axis starting from $S = \$10$, since $S_f(v, 0) = \$10$ implies that $P(S, v, \tau) = 0$ for all $S \geq \$10$.

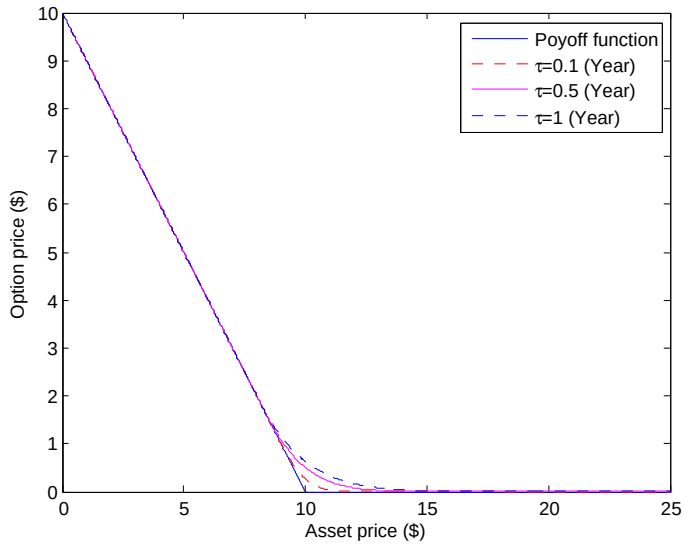

Figure 5.1: Option prices at different times to expiration. Model parameters are $k = 5$, $\eta = 0.16$, $\rho = 0.1$, $r = 0.1$, $T = 1$, $K = 10$, $\xi = 0.9$, $v_0 = 0.25$

So far, we have only presented some detailed results on the option value. For the pricing of American options, it is far more crucial to determine the optimal exercise price than the option price itself. In fact, once the optimal exercise price is accurately determined,

the problem becomes a fixed boundary problem and the determination of the option price is straightforward. In the following, we present some graphs to illustrate the effects of time-dependent volatility on the optimal exercise price.

The optimal exercise price $S_f(v,\tau)$ with different fixed variance values is shown in Figure 5.2. As clearly shown, the optimal exercise price is a monotonically decreasing function of $T-t$ or a monotonically increasing function of t. When the time approaches the expiration time T of the option, the optimal exercise price rises sharply towards the strike price $K = \$10$. At $t = T$ or $\tau = 0$, $S_f(v,\tau) = K$ as we expected. Figure 5.2 also shows that the rate of change of $S_f(v,\tau)$ is much larger near the expiration time than when the option contract is far from expiration. In fact, it is because of this large rate of change of the optimal exercise price near the expiration time that most numerical algorithms have difficulties dealing with the singular behaviour of the optimal exercise price near $t = T$ or $\tau = 0$. However, in this case, the algorithm is designed to deal with the singularity as the location of the optimal exercise price at expiry is known *a priori* and is already included in the algorithm.

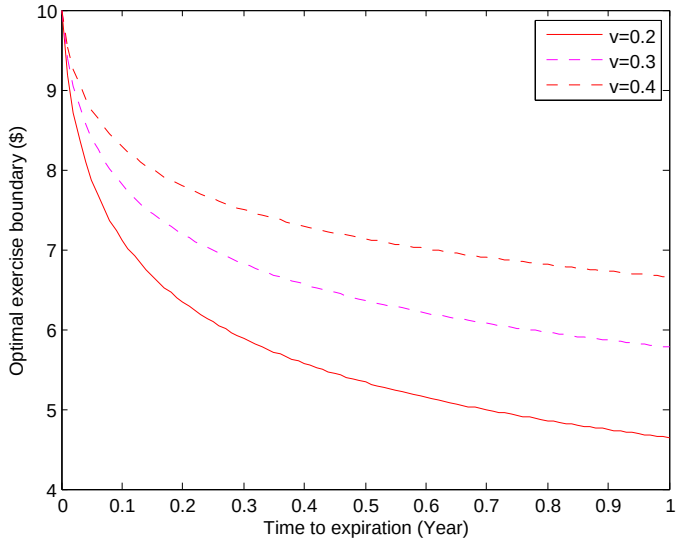

Figure 5.2: Optimal exercise prices with different volatility values. Model parameters are $k = 5$, $\eta = 0.16$, $\rho = 0.1$, $r = 0.1$, $T = 1$, $K = 10$, $\xi = 0.3$, $N = M = 100$

Depicted in Figure 5.3 is a graph of $S_f(v,\tau)$ with different fixed time to expiration τ. As clearly shown in Figure 5.3, the optimal exercise price is a monotonically decreasing function of v and $S_f(v,\tau)$ approaches the strike price as v approaches zero.

Next, we present the graph of optimal exercise price when the Feller condition $\kappa\eta \geq$

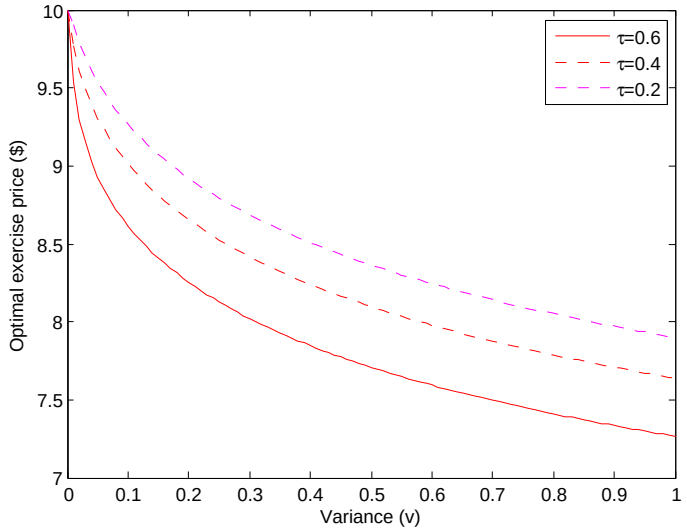

Figure 5.3: Optimal exercise prices with different time to expiration. Model parameters are $k = 5$, $\eta = 0.16$, $\rho = 0.1$, $r = 0.1$, $T = 1$, $K = 10$, $\xi = 0.9$, $N = M = 100$

$\xi^2/2$ is not satisfied. It was shown [30, 118] that whether or not the Feller condition is violated, the solution converges to the same value, at the boundary where $v = 0$. Therefore, considering the two scenarios, we graph the optimal exercise prices using the grid sizes $N \times M = 10 \times 40$ for each case. The model parameters are the same with the parameters in Figure 5.2 except using $\xi = 0.3$ and $\xi = 2$ when the Feller condition is satisfied and violated respectively. The graph shows that even though the Feller condition is violated, the results converge to the same value as $v \to 0$.

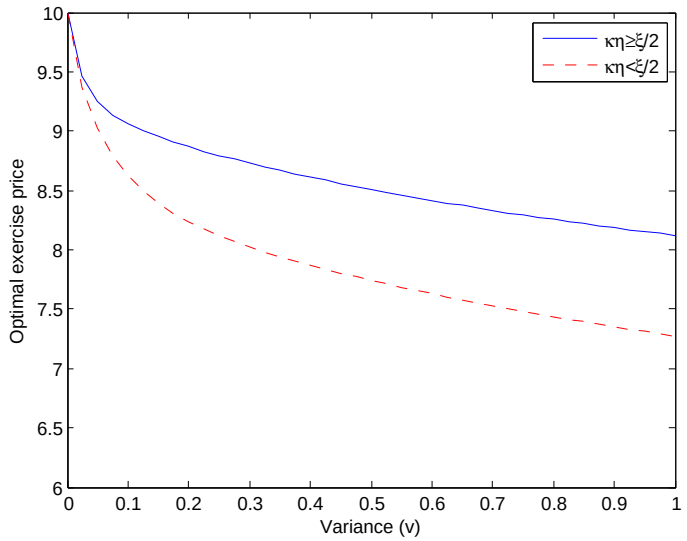

Figure 5.4: Optimal exercise prices for different conditions. Model parameters are $k = 5$, $\eta = 0.16$, $\rho = 0.1$, $r = 0.1$, $T = 1$, $K = 10$

Discussion on efficiency of the hybrid method.

Attempt is made here to compare the computational cost of the hybrid method with a known method, predictor corrector method [118], in the literature. The tested example is chosen with the parameter values $\kappa = 1.5$, $\eta = 0.16$, $\rho = 0.1$, $r = 0.1$, $K = 10$. These values are consistent with the model parameters in [118] in order to ensure a fair comparison.

Table 5.4 compares the computational cost of the current hybrid method with the predictor-corrector method of Zhu and Chen [118] at different grid resolutions. As shown in Table 5.4, the major set back of the current hybrid scheme is its high computational cost. The time required to evaluate the result is quite high and the CPU memory used is large. However, much of the implementation of the proposed scheme has been geared toward proof of concept and no optimization of the code has been attempted.

Table 5.4: Comparison of the computational cost of the hybrid method and predector corrector method

N_x	N_τ	N_v	Ref. [118]	Hybrid method
13	25	50	0.8440	418.3720
26	80	50	3.4370	7004.6543
52	250	50	19.5420	423513.7456

Next, we briefly investigate the convergence of our numerical result. We are particularly interested in the convergence properties of the algorithm as the grid is refined. To achieve this, we consider the convergence ratio proposed in D'Halluin et al. [35].

Let $\triangle\tau = h$ and $\triangle x = h$. Note that for all our tests, we simply use constant step-size h for both temporal and spatial directions. If we then carry out a convergence study by letting $h \to 0$, then we can assume that the error in the solution (at a given node) is obtained from $P_M(h) = P_{\text{exact}} + h^\omega$, and the convergence ratio is defined as

$$ratio = \frac{P_{M/2} - P_M}{P_{M/4} - P_{M/2}}, \tag{5.28}$$

where P_M denotes here the approximated price obtained with $M = M_x$ number of finite elements and P_{exact} is the assumed exact value. We fixed the number of v steps at $M_v = 50$. For each test, as we double the number of grid points by reducing the timestep and element sizes ($\triangle x = \triangle\tau = 0.02$ on the coarsest grid) in half.

In the case of quadratic convergence ($\omega = 2$), then ratio = 4, while for linear conver-

Table 5.5: Ratio for the price of American put options as the starting point S_0 varies

M	$S_0 = 8$	$S_0 = 8$	$S_0 = 10$	$S_0 = 11$	$S_0 = 12$
50	2.41087	2.73284	3.02986	2.96705	3.28045
100	2.21926	2.37261	2.39418	2.41290	2.43874
200	2.01274	2.11748	2.14439	2.17503	2.164229

gence ($\omega = 2$), ratio = 2. In Table 5.5 we show the convergence rate for a American put option written on Heston's model using the data in Table 5.1. The table shows that ratios of the present method are all around 2 and 3, which suggests that the convergence ratio for the current method is linear in both x and τ direction. Moreover, Table 5.5 shows that the observed ratios are very stable, and this gives an evidence of the stability of the method.

5.6 Conclusion

In this chapter, we have presented a hybrid method, a combination of the inverse finite element and the finite-difference approach, to the pricing of American options written on the Heston model. The method eliminates the time derivatives and the nonlinearity of the pricing problem is successfully dealt with without any linearization. We establish the convergence by reformulating the discretized problem in variational form and study the approximation of the option price. Various numerical experiments suggest that the current method is efficient, and can be easily extended to price American-style options under other stochastic volatility models. Based on the numerical results, we have also examined quantitatively the influence of the time-dependent volatility on the optimal exercise price.

Chapter 6

Conclusion

In this thesis, we have exploited the application of inverse finite element method (iFEM) in quantitative finance. We start with the comparative study of direct and inverse finite element methods for solving the free boundary problem of an American option under the Black-Scholes model. More specifically, we compare the computational efficiency of the direct and inverse approaches by carrying out a critical performance analysis of the two approaches against some benchmark solutions. This was done for various grid and time step refinements. The result has shown that under the same grid resolutions, the inverse method is more accurate than its direct counterpart. The results demonstrate that the inverse approach is indeed more efficient than its direct counterpart, as a higher performance in terms of an acceptable computing time and accuracy is achieved.

We further exploit the application of iFEM by formulating the option pricing problems as linear complementarity problems. The LCP formulation is beneficial for iterative solution, since the unknown boundary does not appear explicitly, and can be obtained in a postprocessing step [84, 87]. The key feature of the inverse algorithm is that the solution is limited to the yielded part of the option, thus, the solution corresponds to the original pricing model. Furthermore, the algorithm adapted is able to deal with the nonlinearity of the option pricing problems without any regularization or linearization when compared to most existing numerical methods. Additionally, it enjoys high efficiency as a result of using the Newton iterative scheme with its inherent quadratic convergence. Through a couple of numerical experiments, we demonstrated the overall performance of the proposed method.

For American option under stochastic volatility model, our technique is similar in

some respects to the case of classical Black-Scholes model, although with additional finite differences to discretize the volatility derivative terms. As in Alexandrou [3], we have assumed that the nodal locations along the volatility direction are fixed, while working out the motion of the nodes along the underlying direction. The resulting non-linear system of equation, which satisfies all the appropriate boundary conditions, is solved once and for all by Newton iteration scheme to determine the optimal exercise price as well as the option value. We establish the convergence by reformulating the discretized problem in variational form and study the approximation of the option price. The advantage of this algorithm is that it preserves the simplicity and flexibility of conventional finite element and finite difference methods while allowing the use of full Newton iteration scheme.

Before closing the conclusion part, we remark that there are several obvious avenues for future research with the inverse approach presented in this thesis. One would be a detailed analysis of pricing and hedging of American vanilla options where the underlying asset follows a jump diffusion process. Similarly, it would be interesting to explore the application of inverse finite element to exotic options and convertible bonds.

Appendix A

Appendix

A.1 Element and global matrices

A.1.1 Linear basis function

For $1 \leq i \leq N-1$, the linear basis function is defined as

$$\varphi_i(x) = \begin{cases} \frac{x-x_{i-1}}{x_i-x_{i-1}} & \text{for} \quad x_{i-1} \leq x \leq x_i \\ \frac{x_{i+1}-x}{x_{i+1}-x_i} & \text{for} \quad x_i \leq x \leq x_{i+1} \\ 0 & \text{elsewhere} \end{cases} \tag{A.1}$$

In order to find the unknowns degree of freedom in (3.9), (3.14) and (4.17), respectively, it is necessary to compute the $N-1 \times N-1$ elements for the constrained global stiffness matrices K, $\bar{K}$, K^*, A and B and $N-1$ elements for the total load vectors F. Note that for the applied load vector, we have $Q^* = F$. Since each of the function $\varphi(x)$ is defined in the same way, it is possible to compute k, $\bar{k}$ and k^*, and f for a generic element, and add the contribution of each element to the proper location in the master matrices.

Consider a generic interior element Ω_e on the interval x_a to x_b. It is more convenient to map the element domain $[x_a, x_b]$ to a standardized reference element domain $[-1, 1]$. By writing x in term of ξ, a dummy variable for x, for the dFEM, we obtain the element matrices as:

$$k_{i,j} = \int_{-1}^{1} \left[\varphi_i(\xi)\varphi_j(x)J_a + \left(\frac{\varphi_i'(\xi)\varphi_j'(\xi)}{J_a} - (\gamma-1)\varphi_i'(\xi)\varphi_j(\xi) + \gamma\varphi_i(\xi)\varphi_j(\xi)J_a \right)\theta\Delta\tau \right] d\xi$$

$$\bar{k}_{i,j} = \int_{-1}^{1} \left[\varphi_i(\xi)\varphi_j(x)J_a - \left(\frac{\varphi_i'(\xi)\varphi_j'(\xi)}{J_a} - (\gamma-1)\varphi_i'(\xi)\varphi_j(\xi) + \gamma\varphi_i(\xi)\varphi_j(\xi)J_a \right)\theta\Delta\tau \right] d\xi$$

$$k^*_{i,j} = \int_{-1}^{1} \left(\frac{\varphi_i'(\xi)\varphi_j'(\xi)}{J_a} + (1-\gamma-P_x)\varphi_i'(\xi)\varphi + \gamma\varphi_i(\xi)\varphi_j(\xi)J_a \right) d\xi,$$

$$f_j = \int_{-1}^{1} \gamma\varphi_j(\xi)J_a \, d\xi,$$

$$a_{i,j} = \int_{-1}^{1} \left(\varphi_i'(\xi)\varphi_j'(x)/J_a \right) d\xi,$$

and

$$b_{i,j} = \int_{-1}^{1} \left(\varphi_i(\xi)\varphi_j'(x)J_a \right) d\xi$$

where $r, s = 1 \ldots, p$. The term J_a is the stretch factor between x and ξ coordinates, it becomes the determinant of the Jacobian matrix of the mapping between the coordinate systems. For the linear shape function, $J_a = h/2$. It must also be pointed out that $\varphi(\xi)$ are simply the basis functions defined through the local nodes in the reference element. These are the same for all elements and independent of the element geometry, a great advantage over the original functions.

For the case of linear shape functions,i.e., $p = 1$, $\varphi_1(\xi) = \frac{1}{2}(1-\xi)$ and $\varphi_2(\xi) = \frac{1}{2}(1+\xi)$. With the above, evaluating the so-called 2×2 element-stiffness matrices, $k_{i,j}$, $\bar{k}_{i,j}$ and $k^*_{i,j}$ and 2×1 applied-load vector, f_j are straightforward. By using the same basic idea, the element matrices based on higher order shape functions can be easily implemented.

A.1.2 Assembly of elementwise computation

For the mathematical notation, we let r and s be the local number corresponding to the global node numbers $i_r = q(e, r)$ and $j_s = q(e, s)$. We denote each element matrices by $k_{r,s}$, $\bar{k}_{r,s}$, $k^*_{r,s}$, and f_r. In order to evaluate the global stiffness matrices K_{i_r,i_s}, $\bar{K}_{i_r,i_s}$ and $K^*_{i_r,i_s}$, and the total applied-load vector F_{i_r}, we assemble each element one after another, and add the contribution of each element to the proper location in the matrices, rather than determine the components $K_{I,J}$, $\bar{K}_{I,J}$, $K^*_{I,J}$ and $F(I)$ for a fixed I and J by adding the contribution from each of the element matrices at one.

For computer implementation, we initialize K_{i_r,i_s}, $\bar{K}_{i_r,i_s}$, $K^*_{i_r,i_s}$ and f_r all to zero be-

fore any contribution is added from an element matrix, and $K^q_{i_r,i_s}$, $\bar{K}^q_{i_r,i_s}$, $K^{q*}_{i_r,i_s}$ and F^q_r after qth element matrices are added. Since the implementation process is the same for all global stiffness matrices, we shall write the assembling procedures described above for one global matrix and the total applied load vector as:

$$K^q_{i_r,i_s} = K^{q-1}_{i_r,i_s} + k^q_{r,s}, \quad r,s = 1\ldots,p$$

and

$$F^q_r = F^{q-1}_r + f^q_r, \quad r = 1\ldots,p$$

The next step after assembling the global matrices is to impose the boundary conditions. For this problem, based on our governing PDE system (2.2), $u(1)$ and $u(N+1)$ are both specified. Then the first and the last rows of all the matrices need not to be included, but the second and $N-2$ rows of the applied load vector must be adjusted. However, since $u(1) = 0$, only the far field boundary condition needs to be imposed. With the linear basis functions, for $I, J = 2\ldots N$, the constrained master matrices are as follows:

$$\hat{K}_{I,J} = K_{I,J}, \quad 2 \le I, J \le N$$

$$\hat{F}_I = F_I, \quad 2 \le I, J \le N-1$$

and

$$\hat{F}_N = F_N - u(N+1)K_{N,N+1}.$$

A.2 The proof of Theorem 3.3.6

Proof. According to the explicit form of the Jacobian matrix (see Theorem 3.3.4, it can be shown that for $i, j = 1\ldots N$, $J_R(j,i)$ is continuously differentiable if and only if $x(i+1) \neq x(i)$, and $x(i-1) \neq x(i)$. Now, if each column of the Jacobian matrix is treated as a vector function, by applying Lemma 3.3.4, it is clear that for $i = 1\ldots N$, the ith column of the Jacobian matrix $J_R(\cdot,i)$ is continuously differentiable at X, if each component of X is different from its neighbors. In the following, we shall show that any $X \in B(\bar{x^*}, R)$ has that property. Here, $B(\bar{x^*}, R)$ denotes the closed form of $B(x^*, R)$. Note that $B(\bar{x^*}, R)$ is compact as any finite closed set is compact.

For all $X \in B(\bar{x^*}, R)$, its component can be written as

$$X(i) = x^*(i) + a_i, \quad X(i-1) = x^*(i-1) + b_i, \quad X(i+1) = x^*(i+1) + c_i, \quad i = 2\ldots N$$

where $\mid a_i \mid \leq R$, and $\mid b_i \mid \leq R$, and $\mid c_i \mid \leq R$. Thus

$$\mid X(i) - X(i-1) \mid = \mid x^*(i) - x^*(i-1) + a_i - b_i \mid \geq \mid x^*(i) - x^*(i-1) \mid - \mid a_i \mid - \mid b_i \mid \geq R.$$

Similarly, we obtain $\mid X(i) - X(i+1) \mid \geq R$. On the other hand, since each nodal value has a unique location, and moreover, the nodal value is monotonically increasing with respect to the index i, it is straightforward to show that R is always greater than zero. As a result, for any $X \in B(\bar{x^*}, R)$, each of the component is different from its adjacent ones, and consequently, $J_R(\cdot, i)(i = 1 \dots N)$ is continuously differentiable at X . Therefore, for any $X, Y \in B(x^*, R) \subset B(\bar{x^*}, R)$, by applying Lemma 4.4, we obtain $\|J_R(X) - J_R(Y)\| \leq M\|X - Y\|$, where M is a positive constant. □

A.3 Derivation of problem (4.16)

Recall the bilinear form:

$$a(v, u) = \int_{\Omega} \left(\frac{\partial u}{\partial x} \left[\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial x} \right) - Q_x(v - u) \right] \right) dx \geq 0 \tag{A.2}$$

Then, by applying the Ritz-Galerkin method, the variables $u = u(x, \tau)$ and $v = v(x, \tau)$ can be approximated by

$$u \approx \sum_{i=0}^{m} w_i(\tau)\varphi_i(x), \quad v \approx \sum_{i=0}^{m} v_i\varphi_i(x),$$

with the finite elements $\varphi_0(x), \dots ..., \varphi_m(x)$ and weights $w_i, v_i, i = 1, \dots, m$. With the abbreviation $\varphi_i^{'}(x) := \frac{\partial \varphi_i}{\partial x}$, the $A.2$ can now be discretized on the x-axis as follows

$$a(v, u) = \sum_{i,j=0}^{m} w_j(v_i - w_j) \underbrace{\int_{x_0}^{x_m} \varphi_i^{'}\varphi_j^{'} \, dx}_{=:a_{ij}} - \sum_{i,j=0}^{m} w_j(v_i - w_j) \underbrace{\int_{x_0}^{x_m} \varphi_i\varphi_j^{'} \, dx}_{=:b_{ij}} \geq 0.$$

Using vector notation and after some algebraic manipulations, we can write the later inequality equivalently

$$(v - w)^T(A - Q_{x(i)}B) \geq 0$$

A.4 The proof of Lemma 5.2.1

Proof. Under the risk-neutral argument, when $v \to 0$, the leading order term of the solution of the SDE

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dW_1$$

in Equation (5.1) is $S = e^{rt} S_0$, i.e., the underlying becomes virtually riskless, and its price should appreciate at a deterministic rate r when $v \to 0$. Therefore, if $S < K$, the put option should be immediately exercised as there is no reason to hold the option any more if one knows that the underlying will definitely increase for sure. In other words, the underlying price range $[0, K)$ belongs to the "exercise" region, i.e., $[0, K) \subseteq [0, \lim_{v \to 0} S_f(v, t)]$, and thus $\lim_{v \to 0} S_f(v, t) \geq K$. On the other hand, if $S > K$, it is obvious that the value of the put option becomes zero, and therefore it is better to hold the option as the option may still have some time value before its expiration date is reached. That is to say, the underlying price range (K, ∞) belongs to the "continuous" region, i.e., $(K, \infty) \subseteq (\lim_{v \to 0} S_f(v, t), \infty)$, and thus $\lim_{v \to 0} S_f(v, t) \leq K$. A combination of the above two statements leads to the conclusion that $\lim_{v \to 0} S_f(v, t) = K$ □

Bibliography

[1] Y Achdou and O Pironneau. *Computational methods for option pricing*, volume 30. Siam, 2005.

[2] Y Achdou and N Tchou. Variational analysis for the Black and Scholes equation with stochastic volatility. *ESAIM: Mathematical Modelling and Numerical Analysis*, 36(3):373–395, 2002.

[3] A N Alexandrou. An inverse finite element method for directly formulated free boundary problems. *International journal for numerical methods in engineering*, 28(10):2383–2396, 1989.

[4] G Alexandrou, A N & Georgiou. On the early breakdown of semisolid suspensions. *Journal of non-newtonian fluid mechanics*, 142(1):199–206, 2007.

[5] W Allegretto, Y Lin, and H Yang. A fast and highly accurate numerical method for the evaluation of American options. *Dynamics of Continuous Discrete and Impulsive Systems Series B*, 8:127–138, 2001.

[6] A Andalaft-Chacur, M M Ali, and J G Salazar. Real options pricing by the finite element method. *Computers & Mathe matics with Applications*, 61(9):2863–2873, 2011.

[7] L Bachelier. *Louis Bachelier's theory of speculation: the origins of modern finance*. Princeton University Press, 2011.

[8] Gurdip Bakshi, Charles Cao, and Zhiwu Chen. Empirical performance of alternative option pricing models. *The Journal of finance*, 52(5):2003–2049, 1997.

[9] E Barles, G & Lesigne. Sde, bsde and pde. Pitman Research Notes in Mathematics Series. *Pitman Research Notes in Mathematics Series*, pages 47–82, 1997.

[10] G Barles, J Burdeau, M Romano, and N Samsoen. Critical stock price near expiration. *Mathematical finance*, 5(2):77–95, 1995.

[11] G Barone-Adesi and R E Whaley. Efficient analytic approximation of american option values. *Journal of Finance*, 42(2):301–320, 1987.

[12] David S Bates. Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options. *Review of financial studies*, 9(1):69–107, 1996.

[13] A Bensoussan. On the theory of option pricing. *Acta Applicandae Mathematica*, 2(2):139–158, 1984.

[14] A Bensoussan and J L Lions. *Impulse Control And Quasi-Variational Inequalities*. Gaunthier-Villars, 1984.

[15] A Bensoussan and J L Lions. *Applications of variational inequalities in stochastic control*. Elsevier, 2011.

[16] H V D V Bi-CGSTAB. A fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems. *SIAM J. Sci. Stat. Comput*, 13:631–644, 1992.

[17] F Black and M Scholes. The pricing of options and corporate liabilities. *The journal of political economy*, pages 637–654, 1973.

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

[19] P P Boyle. A lattice framework for option pricing with two state variables. *Journal of Financial and Quantitative Analysis*, 1(1-12), 23.

[20] Dietrich Braess. *Finite elements: Theory, fast solvers, and applications in solid mechanics*. Cambridge University Press, 2001.

[21] M J Brennan and E S Schwartz. Finite difference methods and jump processes arising in the pricing of contingent claims: A synthesis. *Journal of Financial and Quantitative Analysis*, 13(03):461–474, 1978.

[22] Franco Brezzi, William W Hager, and Pierre-Arnaud Raviart. Error estimates for the finite element solution of variational inequalities. *Numerische Mathematik*, 28(4):431–443, 1977.

[23] M Briani and A Caramellino, L & Zanette. A hybrid approach for the implementation of the Heston model. *IMA Journal of Management Mathematics*, dpv032, 2015.

[24] J Broadie, M & Detemple. American option valuation: new bounds, approximations, and a comparison of existing methods. *Review of Financial Studies*, 9(4):1211–1250, 1996.

[25] H Bunch, D S & Johnson. A Simple and Numerically Efficient Valuation Method for American Puts Using a Modified Geske-Johnson Approach. *The Journal of Finance*, 47(2):809–816, 1992.

[26] M M Chawla and D J Evans. Numerical volatility in option valuation from Black-Scholes equation by finite differences. *International Journal of Computer Mathematics*, 81(8):1039–1041, 2004.

[27] Wen-Ting Chen. An extensive exploration of three key quantitative approaches for pricing various financial derivatives. 2011.

[28] P G Ciarlet. *The finite element method for elliptic problems*, volume 40. Siam, 2002.

[29] K Clarke, N & Parrott. Multigrid for American option pricing with stochastic volatility. *Applied Mathematical Finance*, 6(3):177–195, 1999.

[30] Nigel Clarke and Kevin Parrott. Multigrid for American option pricing with stochastic volatility. *Applied Mathematical Finance*, 6(3):177–195, 1999.

[31] R W Cottle and R E Pang, J S & Stone. *The linear complementarity problem*, volume 60. Siam, 2009.

[32] J C Cox, S A Ross, and M Rubinstein. Option pricing: A simplified approach. *Journal of financial Economics*, 7(3):229–263, 1979.

[33] C W Cryer. The solution of a quadratic programming problem using systematic overrelaxation. *SIAM Journal on Control*, 9(3):385–392, 1971.

[34] J Cuadrado, R Gutierrez, M A Naya, and P Morer. A comparison in terms of accuracy and efficiency between a MBS dynamic formulation with stress analysis and a non-linear FEA code. *International Journal for Numerical Methods in Engineering*, 51(9):1033–1052, 2001.

[35] Yann d'Halluin, Peter A Forsyth, and Kenneth R Vetzal. Robust numerical methods for contingent claims under jump diffusion processes. *IMA Journal of Numerical Analysis*, 25(1):87–112, 2005.

[36] Darrell Duffie, Jun Pan, and Kenneth Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–1376, 2000.

[37] B Dupire. Pricing with a smile. *Risk*, 7(1):18–20, 1994.

[38] C M Elliott and J R Ockendon. *Weak and variational methods for moving boundary problems*, volume 59. Pitman Publishing, 1982.

[39] J D Evans, R Kuske, and J B Keller. American options on assets with dividends near expiry. *Mathematical Finance*, 12(3):219–237, 2002.

[40] Chicago Board Options Exchange. The CBOE volatility index-VIX. *White Paper*, pages 1–23, 2009.

[41] W Feller. Two singular diffusion problems. nnals of mathematics, 173-182., journal = Annals of mathematics, year = 1951, pages = 173-182,.

[42] P A Forsyth and K R Vetzal. Quadratic convergence for valuing american options using a penalty method. *SIAM Journal on Scientific Computing*, 23(6):2095–2122, 2002.

[43] A Friedman. *Variational principles and free-boundary problems*. Courier Corporation, 2010.

[44] Jim Gatheral. *The volatility surface: a practitioner's guide*, volume 357. John Wiley & Sons, 2011.

[45] R Geske and H E Johnson. The American put option valued analytically. *Journal of Finance*, pages 1511–1524, 1984.

[46] D Grant and D Vora, G & Weeks. Path-dependent options: Extending the Monte Carlo simulation approach. *Management Science*, 43(11):1589–1602, 1997.

[47] D Grant and D E Vora, G & Weeks. Simulation and the early exercise option problem. *J OF FINANCIAL ENGINEERING*, J5(3), 1996.

[48] L Haugh, M B & Kogan. Pricing American options: a duality approach. *Operations Research*, 52(2):258–270, 2004.

[49] S L Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options.

[50] A D Holmes and H Yang. A front-fixing finite element method for the valuation of American options. *SIAM journal on scientific computing*, 30(4):21–58, 2008.

[51] S Hout, K J & Foulon. ADI finite difference schemes for option pricing in the Heston model with correlation. *Int. J. Numer. Anal. Model*, 7(2):303–320, 2010.

[52] H Z Huang, M G Subrahmanyam, and G G Yu. Pricing and hedging american options: A recursive integration method. *The Review of Financial Studies*, 9, 1996.

[53] J Huang and J S Pang. Option pricing and linear complementarity. *The Journal of Computational Finance*, 2(1):31–60, 1998.

[54] J Hull and A White. Valuing derivative securities using the explicit finite difference method. *Journal of Financial and Quantitative Analysis*, 25(01):87–100, 1990.

[55] J C Hull. *Options, futures, and other derivatives*. Boston, Massechusetts London Pearson, 2012.

[56] J Hunt, P & Kennedy. *Financial derivatives in theory and practice*. John Wiley & Sons, 2004.

[57] J Ikonen, S & Toivanen. Efficient numerical methods for pricing American options under stochastic volatility. *Numerical Methods for Partial Differential Equations*, 24(1):104–126, 2008.

[58] J Ikonen, S & Toivanen. Operator splitting methods for pricing American options under stochastic volatility. *Numerische Mathematik*, 113(2):299–324, 2009.

[59] J Ikonen, S andToivanen. Operator splitting methods for american option pricing. *Applied mathematics letters*, 17(7):809–814, 2004.

[60] P Jaillet, D Lamberton, and B Lapeyre. Variational inequalities and the pricing of American options. *Acta Applicandae Mathematica*, 21(3):263–289, 1990.

[61] C Johnson. *Numerical solution of partial differential equations by the finite element method*. Courier Corporation, 2012.

[62] P Johnson. *Improved Numerical Techniques for Occupation-Time Derivatives and Other Complex Financial Instruments*. PhD thesis, University of Manchester, 2008.

[63] N Ju. Pricing by American option by approximating its early exercise boundary as a multipiece exponential function. *Review of Financial Studies*, 11(3):627–646, 1998.

[64] Raul Kangro and Roy Nicolaides. Far Field Boundary Conditions for Black-Scholes Equations. *SIAM Journal on Numerical Analysis*, 38(4):1357–1368, 2000.

[65] L V Kantorovich and V I Krylov. *Approximate methods of higher analysis*. Interscience, 3rd edition, 1964.

[66] I Karatzas. On the pricing of American options. *Applied mathematics and optimization*, 17(1):37–60, 1988.

[67] M Kauer. *Inverse finite element characterization of soft tissues with aspiration experiments*. PhD thesis, ETH Zr̈ich, Swiss Federal Institute of Technology, Zr̈ich, Switzerland, 2001.

[68] I N Kim. The analytic valuation of American options. *Review of financial studies*, 3(4):547–572, 1990.

[69] M D Koulisianis and T S Papatheodorou. Pricing of american options using linear complementarity formulation: methods and their evaluation. *Neural, Parallel & Scientific Computations*, 11(4):423–444, 2003.

[70] M Krizek and P Goittaanmaki. *Finite element approximation of variational problems and applications*, volume 50. Longman Scientific & Technical, 1990.

[71] A Kunoth, C Schneider, and K Wiechers. Multiscale methods for the valuation of american options with stochastic volatility. *International Journal of Computer Mathematics*, 89(9):1145–1163, 2012.

[72] S Larsson and V Thome. *Partial differential equations with numerical methods*, volume 45. Springer Science & Business Media, 2008.

[73] H Liu, T. H Hong, M Herman, T Camus, and R Chellappa. Accuracy vs efficiency trade-offs in optical flow algorithms. *Computer vision and image understanding*, 72(3):271–286, 1998.

[74] Anders Logg, Kent-Andre Mardal, and Garth Wells. *Automated solution of differential equations by the finite element method: The FEniCS book*, volume 84. Springer Science & Business Media, 2012.

[75] F A Longstaff and E S Schwartz. Valuing American options by simulation: a simple least-squares approach. *Review of Financial studies*, 14(1):113–147, 2001.

[76] Ma, Protter, Martin, and Torres. The connection between bsde and pde and numerical methods to solve bsde. Working group on portfolio management, SAMSI, 2005.

[77] L W MacMillan. Analytic approximation for the American put option. *Advances in futures and options research*, 1(1):119–139, 1986.

[78] H P McKean. A free boundary problem for the heat equation arising from a problem in mathematical economics. *Sloan Management Review*, 6(2):32, 1965.

[79] S MEMON. Finite element method for american option pricing: A penalty approach. *International Journal of Numerical Analysis and Modelling, Series B Computing and Information*, 3(3):345–370, 2012.

[80] R C Merton. The theory of rational option pricing. *Bell Journal of Economics and Management Science*, 1:141–183, 1973.

[81] R C Merton and E S Brennan, M J & Schwartz. The valuation of American put options. *The Journal of Finance*, 32(2):449–462, 1977.

[82] Robert C Merton. Option pricing when underlying stock returns are discontinuous. *Journal of financial economics*, 3(1-2):125–144, 1976.

[83] B F Nielsen, O Skavhaug, and A Tveito. Penalty and front-fixing methods for the numerical solution of american option problems. *Journal of Computational Finance*, 5(4):69–98, 2002.

[84] C W Oosterlee. On multigrid for linear complementarity problems with application to American-style options. *Electronic Transactions on Numerical Analysis*, 15(165-185):2–7, 2003.

[85] CONALL O'SULLIVAN and STEPHEN O'SULLIVAN. Pricing European And American Options In The Heston Model With Accelerated Explicit Finite Differencing Methods. *International Journal of Theoretical and Applied Finance*, 16(03):1350015, 2013.

[86] K N Pantazopoulos, E N Houstis, and S Kortesis. Front-tracking finite difference methods for the valuation of american options. *Computational Economics*, 12(3):255–273, 1998.

[87] O Pauly. *Numerical simulation of American options*. PhD thesis, Universitt Ulm, 2004.

[88] A Quarteroni and F Sacco, R & Saleri. *Numerical mathematics*, volume 37. Springer Science & Business Media, 2010.

[89] Fabrice D Rouah. *The Heston Model and its Extensions in Matlab and C*. John Wiley & Sons, 2013.

[90] J Schbel, R & Zhu. Stochastic volatility with an OrnsteinUhlenbeck process: an extension. *European Finance Review*, 3(1):23–46, 1999.

[91] R Seydel and R Seydel. *Tools for computational finance*, volume 4. Springer, Berlin, 2002.

[92] Vladimir Nikolaevich Solov'ev. A generalization of Gershgorin's theorem. *Mathematics of the USSR-Izvestiya*, 23(3):545, 1984.

[93] J C Stein, E M & Stein. Stock price distributions with stochastic volatility: an analytic approach. *Review of financial Studies*, 4(4):727–752, 1991.

[94] Peter Tankov. *Financial modelling with jump processes*, volume 2. CRC press, 2003.

[95] D Tavella and C Randall. *Pricing financial instruments: The finite difference method*. John Wiley & Sons, USA, 2000.

[96] L Tavernini. Numerical analysis note 2, 2008.

[97] J A Tilley. Valuing American options in a path simulation model. *Transactions of the Society of Actuaries*, 45(83):104, 1993.

[98] Robert G Tompkins et al. Stock index futures markets: stochastic volatility models and smiles. *Journal of Futures Markets*, 21(1):43–78, 2001.

[99] P Tong and J N Rossettos. *Finite-element Method*. The MIT Press, 1978.

[100] J Topper. Option pricing with finite elements. In: Wilmott Magazine, January 2005. pp. 84-90.

[101] William F Trench. Introduction to real analysis. 2013.

[102] K Urban and M Tang, S & Rometsch. Numerical Finance. lecture notes, University of Ulm, 2009.

[103] S Wang, X Q Yang, and K L Teo. Power penalty method for a linear complementarity problem arising from american option valuation. *Journal of Optimization Theory and Applications*, 129(2):227–254, 2006.

[104] S Wang, S Zhang, and Z Fang. A superconvergent fitted finite volume method for BlackScholes equations governing European and American option valuation. *Numerical Methods for Partial Differential Equations*, 31(4):1190–1208, 2015.

[105] M Widdicks. *Examination, extension and creation of methods for pricing options with early exercise features*. PhD thesis, University of Manchester, 2002.

[106] P Wilmott, J Dewynne, and S Howison. *Option pricing: Mathematical models and computation*. Oxford financial press, 1993.

[107] P Wilmott, S Howison, and J Dewynne. *The mathematics of financial derivatives: a student introduction*. Cambridge University Press, 1995.

[108] S J Wright. *Primal-dual interior-point methods*. Siam, Chicago, 1997.

[109] L Wu and Y K Kwok. A front-fixing finite difference method for the valuation of American options. *Journal of Financial Engineering*, 6(4):83–97, 1997.

[110] C S Zhang. *Adaptive finite element methods for variational inequalities: Theory and applications in finance*. PhD thesis, University of Maryland, USA, 2007.

[111] K Zhang, H Song, and J Li. Front-fixing FEMs for the pricing of American options based on a PML technique. *Applicable Analysis*, 94(5):903–931, 2015.

[112] R Zhang, Q Zhang, and H Song. An efficient finite element method for pricing American multi-asset put options. *Communications in Nonlinear Science and Numerical Simulation*, 29(1):25–36, 2015.

[113] S P Zhu. An exact and explicit solution for the valuation of American put options. *Quantitative Finance*, 6(3):229–242, 2006.

[114] S P Zhu. A simple approximation formula for calculating the optimal exercise boundary of American puts. *Journal of Applied Mathematics and Computing*, 37(1-2):611–623, 2011.

[115] Song-Ping Zhu. Pricing American Options-an Important Fundamental Research in Pricing Financial Derivatives. *Wollongong, Wollongong, Australia: University of Wollongong*.

[116] Song-Ping Zhu. A new analytical approximation formula for the optimal exercise boundary of American put options. *International Journal of Theoretical and Applied Finance*, 9(07):1141–1177, 2006.

[117] Song-Ping Zhu and W T Chen. An inverse finite element method for pricing American options. *Journal of Economic Dynamics and Control*, 37(1):231–250, 2013.

[118] Song-Ping Zhu and Wen-Ting Chen. A predictor-corrector scheme based on the ADI method for pricing american puts with stochastic volatility. *Computers & Mathematics with Applications*, 62(1):1–26, 2011.

[119] Song-Ping Zhu and Jin Zhang. A new predictor-corrector scheme for valuing American puts. *Applied Mathematics and Computation*, 217(9):4439–4452, 2011.

[120] O. Zienkiewicz. *The Finite Element Method*. McGraw-Hill, London, 1989.

[121] R Zvan and K R Forsyth, P A & Vetzal. Penalty methods for American options with stochastic volatility. *Journal of Computational and Applied Mathematics*, 91(2):199–218, 1998.

[122] R Zvan and K R Forsyth, P A & Vetzal. A finite volume approach for contingent claims valuation. *IMA Journal of Numerical Analysis*, 21(3):703–731, 2001.

[123] Robert Zvan, PA Forsyth, and KR Vetzal. Penalty methods for American options with stochastic volatility. *Journal of Computational and Applied Mathematics*, 91(2):199–218, 1998.

List of my publications

[1] Bolujo Joseph Adegboyegun and Xiaoping Lu. A comparative study of the direct and the inverse finite element methods for pricing american options. *To be submitted.*

[2] Bolujo Joseph Adegboyegun and Xiaoping Lu. A hybrid approach for pricing american options under the heston model. *To be submitted.*

[3] Xiaoping Lu and Bolujo Joseph Adegboyegun. On the inverse finite element approach for pricing american options under linear complementarity formulation. *To be submitted.*