2011

An extensive exploration of three key quantitative approaches for pricing various financial derivatives

Wen-Ting Chen
University of Wollongong

Recommended Citation
http://ro.uow.edu.au/theses/3280
NOTE

This online version of the thesis may have different page formatting and pagination from the paper copy held in the University of Wollongong Library.

UNIVERSITY OF WOLLONGONG

COPYRIGHT WARNING

You may print or download ONE copy of this document for the purpose of your own research or study. The University does not authorise you to copy, communicate or otherwise make available electronically to any other person any copyright material contained on this site. You are reminded of the following:

Copyright owners are entitled to take legal action against persons who infringe their copyright. A reproduction of material that is protected by copyright may be a copyright infringement. A court may impose penalties and award damages in relation to offences and infringements relating to copyright material. Higher penalties may apply, and higher damages may be awarded, for offences and infringements involving the conversion of material into digital or electronic form.
An extensive exploration of three key quantitative approaches for pricing various financial derivatives

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

Doctor of Philosophy

from

University of Wollongong

by

Wen-Ting Chen, B.Sc. (Southeast University)
M.A. (Fudan University)

School of Mathematics and Applied Statistics

2011
CERTIFICATION

I, Wenting Chen, declare that this thesis, submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the School of Mathematics and Applied Statistics, the University of Wollongong, is wholly my own work unless otherwise referenced or acknowledged. The document has never been submitted for qualifications at any academic institution.

Wenting Chen
March, 2011
Acknowledgements

First and foremost, I would like to convey my sincerest gratitude and appreciation to my supervisor, Prof. Song-Ping Zhu, for his insightful supervision and continuous encouragement throughout the research. I, in particular, appreciate all his contributions of time and ideas to make my PhD studies productive and stimulating. His passion in science and rigorous attitude towards research have exceptionally inspired and enriched my growth as a student, a researcher, and a scientist to be in the future. I am indebted to him more than he knows.

My appreciation goes, in particular, to Dr. Xiao-Ping Lu for her continuous encouragement and warm care. Without her help, this thesis would never be finished so nicely. I am also thankful for the excellent example she has established as a successful female researcher, which I need to pursue in the future. I wish to thank all the fellow friends in the Center of Financial Mathematics and the School of Mathematics and Applied Statistics in the University of Wollongong, particularly Dr. Guang-Hua Lian and Ms Endah Putri, who have made my time in Wollongong so enjoyable. The financial support from the University of Wollongong with HDR tuition scholarship and University Postgraduate Research Award is also gratefully acknowledged.

Finally, I would like to thank my parents and grandparents for encouraging me to be disciplined and systematic in my work ever since I was a child, for their care and love during the past years. I am also deeply indebted to my husband, Dr. Xiang Xu, for his endless love, understanding and support.
Abstract

Options and other financial derivatives have become increasingly important in financial markets ever since Black and Scholes (1973) proposed an analytical and quantitative formula for valuing European options or other similar financial derivatives of a fixed lifetime. However, how to rationally price option derivatives efficiently and accurately is still one of the major challenges in today’s finance industry. This thesis contributes to the literature significantly by further exploring some quantitative approaches for pricing various derivatives.

Classified by the quantitative methods adopted to price option derivatives, this thesis consists three parts, with each part addressing one of the key quantitative approach. Moreover, these parts are based on ten papers published in or submitted to various top-class international journals.

The issue regarding numerically pricing option derivations, particularly, the American puts, is discussed in Part 1, which contains Chapter 2, Chapter 3 and Chapter 4. In this part, we first introduce a new numerical scheme, based on the ADI (alternating direction implicit) method, to price American put options under a stochastic volatility model. Realizing the fact that the numerical approaches designed for American puts with finite maturities are usually with low accuracy and computational inefficiency when applied to deal with the perpetual case, a new numerical scheme, based on the Legendre pseudospectral method, is then introduced to solve for the price of perpetual American puts with stochastic volatility efficiently and accurately. On the other hand, upon considering the fact that a “convergency-proved” numerical approach has never been proposed for the valuation of American options, we also introduce, in this part, an IFE (inverse finite element) approach to price American puts under the Black-Scholes model. Numerical results show that the IFE approach is quite accurate and efficient, and can be easily extended to multi-asset or stochastic
volatility pricing problems. Most remarkably, we have managed to provide a convergence analysis for the IFE approach, which ensures that our numerical solution does indeed converge to the exact one of the original nonlinear system.

In Part 2, we concentrate on deriving analytic approximations for option derivatives. Two sub-issues, regarding the asymptotic behaviour of the optimal exercise price near expiry and pricing approximation formulae for vanilla options, are discussed in details in this part. On one hand, we derive two explicit analytical expressions for the optimal exercise price near expiry under the local volatility and the stochastic volatility models, respectively, by using the method of matched asymptotic expansions. The results show that under the local volatility model, if the underlying dividend is greater than the risk-free interest rate, the behavior of the optimal exercise price is parabolic, otherwise, an extra logarithmic factor appears, which agrees with the constant volatility case. We also find that under the stochastic volatility model, the option prices are quite different from the corresponding Black-Scholes’ case, but the leading order term of the optimal exercise price remains almost the same as the constant volatility case if the spot volatility is given the same value as the constant volatility appearing in the Black-Scholes model. On the other hand, a series of approximations for the price of vanilla options are also provided in this part. In particular, by realizing that Heston’s formula is problematic, we derive a new approximation for European puts under the Heston model by using singular perturbation method. The newly-obtained formula only involves the standard normal distribution function, and is thus as fast and easy to implement as the Black-Scholes formula. Moreover, we derive three formulae for pricing perpetual American puts under the slow-mean reverting, fast-mean reverting and multi-scale stochastic volatility models, respectively. Based on the formulae, the quantitative effects of different stochastic volatility dynamics on the optimal exercise strategies of perpetual American puts are also discussed and compared. It is found that the effect of a slowly-varying volatility factor varies with respect to the spot volatility. That is, for certain values of the spot volatility, the stochastic volatility tends to add the value of the contract, but for
others, it makes the contract less valuable, whereas the fast mean-reversion factor always tends to add the price of a put option contract, had the underlying been assumed to be falling.

The last part of this thesis deals with the exact solution approach, which is extremely important in both theoretic and practical sides of option pricing. In particular, we consider the analytical pricing of Parisian-type options, i.e., Parisian and ParAsian options, which are barrier options with the knock-in or knock-out feature only activated after the underlying has spent a certain prescribed time beyond or below the barrier. By the reduction of a three-dimensional problem to a two-dimensional problem through a coordinate transform that has elegantly “absorbed” the directional derivative associated with the “barrier time” into the time derivative, we have been able to obtain two closed-form analytic formulae for prices of Parisian and ParAsian options, respectively, which can both be easily and efficiently evaluated numerically.
## Contents

1 Introduction and Background

1.1 Option derivatives ........................................... 1
  1.1.1 Vanilla options ........................................ 1
  1.1.2 Exotic options .......................................... 2

1.2 Mathematical background .................................... 3
  1.2.1 Classical numerical approaches revisited ............... 4
  1.2.2 The perturbation method ................................ 6
  1.2.3 Stochastic calculus ..................................... 8

1.3 Option pricing models ....................................... 9
  1.3.1 The Black-Scholes model ................................ 10
  1.3.2 The local volatility model .............................. 12
  1.3.3 Stochastic volatility models ............................ 13

1.4 Literature review ........................................... 17
  1.4.1 European and American options ......................... 17
  1.4.2 Parisian-type options .................................. 21

1.5 Structure of thesis .......................................... 22

2 Part I: Numerical approaches

2.1 A predictor-corrector scheme based on the ADI method for pricing

American puts with stochastic volatility 29
2.1 Introduction .................................................. 29

2.2 Pricing American options under the Heston model .......................... 32
   2.2.1 The Heston model .................................. 32
   2.2.2 The PDE for the value of American puts and the corresponding
        boundary conditions .................................. 33

2.3 Numerical method based on the ADI scheme .............................. 41
   2.3.1 Coordinate transformation ................................ 41
   2.3.2 Discretization of the PDE system ....................... 43
   2.3.3 Linearization of the nonlinear PDE system ................ 46
   2.3.4 The ADI scheme ...................................... 48

2.4 Examples and discussions ..................................... 52
   2.4.1 Valuing European put options using the ADI scheme ....... 52
   2.4.2 Discussion on convergence ............................. 53
   2.4.3 Discussion on accuracy and efficiency .................. 71

2.5 Conclusion ................................................... 72

3 A spectral-collocation method for pricing perpetual American puts
   with stochastic volatility ..................................... 74
   3.1 Introduction ............................................. 74
   3.2 Perpetual American puts under the Heston model .................... 76
   3.3 Numerical scheme based on the Legendre pseudospectral method .... 77
   3.4 Numerical results and discussions ................................ 81
      3.4.1 A test example ...................................... 82
      3.4.2 SC scheme VS PSOR method ......................... 83
      3.4.3 The impact of stochastic volatility .................. 84

3.5 Conclusion ................................................... 87

4 An inverse finite element method for pricing American options .......... 88
4.1 Introduction .................................................. 88
4.2 Mathematical formulation ................................. 90
4.3 Inverse finite element approach ......................... 93
4.4 Convergence analysis ...................................... 97
4.5 Examples and discussions ............................... 105
4.6 Conclusion .................................................. 111

Part II: Analytical approximations

5 A new analytical approximation for European puts with stochastic volatility .............................................. 114
  5.1 Introduction .............................................. 114
  5.2 Heston model and European put options with short tenor .............................................................. 116
  5.3 Numerical tests and discussions ....................... 121
  5.4 Conclusion ............................................... 124

6 Optimal exercise price of American options near expiry ................................................................. 125
  6.1 Introduction .............................................. 125
  6.2 American puts under general diffusion process .................. 127
  6.3 Matched asymptotic analysis for the optimal exercise price near expiry 128
      6.3.1 $\sigma$ is a function of S only ...................... 128
      6.3.2 $\sigma$ is a function of both $S$ and $t$ .............. 136
  6.4 Conclusion ............................................... 140

7 Asymptotic behavior of the optimal exercise price of American puts near expiry under stochastic volatility ................................................................. 142
  7.1 Introduction .............................................. 142
  7.2 American puts under the Heston model ................ 144
7.3 Matched asymptotic expansions for the optimal exercise price near expiry .................................................. 145
7.4 Numerical results .............................................................................................................................................. 159
7.5 Conclusion ...................................................................................................................................................... 161

8 Should an American option be exercised earlier or later if volatility is not assumed to be a constant? 162
8.1 Introduction ...................................................................................................................................................... 162
8.2 Perpetual American puts under the Heston model ......................................................................................... 165
8.3 Numerical tests and discussions ..................................................................................................................... 173
  8.3.1 Validation of the approximation .................................................................................................................. 173
  8.3.2 Impact of the stochastic volatility term ...................................................................................................... 176
8.4 Conclusion ...................................................................................................................................................... 182

9 Pricing perpetual American options under a stochastic volatility model with fast mean-reversion 183
9.1 Introduction ...................................................................................................................................................... 183
9.2 Perpetual American puts under fast mean-reverting volatility ...................................................................... 185
9.3 Pricing approximations ................................................................................................................................... 188
9.4 Discussions ...................................................................................................................................................... 194
9.5 Conclusion ...................................................................................................................................................... 196

10 Pricing perpetual American puts under multi-scale stochastic volatility 197
10.1 Introduction ...................................................................................................................................................... 197
10.2 Multiscale volatility models ........................................................................................................................... 200
10.3 Pricing approximations for perpetual American puts .................................................................................... 202
10.4 Discussions ...................................................................................................................................................... 216
10.5 Conclusion ................................................................. 217

Part III: Exact solution approach

11 Pricing Parisian and Par\textit{asian} options analytically ................. 219

11.1 Introduction .................................................................. 219

11.2 PDE systems for pricing Parisian and Par\textit{asian} options ............. 223

11.2.1 Parisian options ......................................................... 223

11.2.2 Par\textit{asian} options .................................................... 232

11.3 Our solution procedure .................................................... 235

11.3.1 “Moving windows” and the 2-D PDE systems ......................... 236

11.3.2 Degeneration of the 2-D systems ................................... 239

11.3.3 The analytical solution for the Parisian option price ............... 242

11.3.4 The analytical solution for the Par\textit{asian} option price ............. 249

11.4 Numerical examples and discussions ...................................... 253

11.4.1 A comparison with Wilmott \textit{et al.}\’s numerical solutions .......... 254

11.4.2 Numerical performance of our analytical solutions ............... 257

11.5 Conclusion .................................................................... 262

12 Conclusions .................................................................... 264

A The proof and the implementation details of Chapter 2 ................. 267

A.1 The consistency between the boundary condition at $v = 0$ and the
American put option pricing system with stochastic volatility ........ 267

A.2 The discretization of the ADI method .................................... 268

A.3 The matrix equation after the ADI method is applied ................ 271

B Realization details of Chapter 4 ............................................ 274

B.1 Finite element formulation for the PDE system (4.3.5) ............... 274
List of Figures

2.1 Schematic flow chart of the scheme ................................. 49

2.2 European put option prices with different variance values. ......... 54

2.3 American puts and the optimal exercise prices calculated with $\eta = 0.16$, $r = 0.1$, $\rho = 0.1$, $T - t = 0.25(\text{year})$, $K = $10.0, $\theta = \frac{1}{2}$. ............. 67

2.4 Option prices at different times to expiration. Model parameters are $\kappa = 2.5$, $\eta = 0.16$, $\sigma = 0.45$, $r = 0.1$, $\rho = 0.1$, $K = $10.0. .................. 69

3.1 Comparison of the option prices under two different models. Model parameters are $r = 0.1$, $\sigma = 0.45$, $\rho = 0.1$, $\eta = 0.2$, $\kappa = 4$, $K = $10.0. ................. 85

4.1 Comparison of optimal exercise price for the first example. ........ 107

4.2 Accuracy vs. Efficiency. ............................................. 110

4.3 Optimal exercise price for the second case. .......................... 111

5.1 European put option price with different times to expiration. ...... 122

7.1 Boundary layer structure. ............................................. 150

7.2 Comparison of the numerical results and the estimated values. Model parameters are $r = 0.1$, $\sigma = 0.2$, $\rho = 0.1$, $\eta = 0.16$, $\kappa = 1.5$, Strike price $K = $10.0. .................................................. 160

8.1 Normalized $P/K$ option price as a function of the moneyness $S/K$. . 178
8.2 The difference of optimal exercise prices with and without stochastic volatility .......................................................... 180

11.1 Pricing domain of a Parisian up-and-out option ................................. 225

11.2 Wilmott et al.’s solutions VS our results for the case $T-t \geq \bar{J}-J$.
Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$, $K = $10, $\bar{S} = $12,
$\bar{J} = 0.2\text{(year)}$ and $T-t = 1\text{(year)}$. ........................................... 256

11.3 Wilmott et al.’s solutions VS our results for the case $T-t < \bar{J}-J$.
Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$, $K = $10, $\bar{S} = $12,
$\bar{J} = 0.2\text{(year)}$ and $T-t = 0.1\text{(year)}$. ........................................... 256

11.4 The price of the Parisian up-and-out call option at different $J$ values.
Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$, $K = $10, $\bar{S} = $12,
$\bar{J} = 0.2\text{(year)}$ and $T-t = 1\text{(year)}$ ........................................... 260

11.5 The price of the Par\textit{asian} up-and-out call option at different $J$ values.
Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$, $K = $10, $\bar{S} = $12,
$\bar{J} = 0.2\text{(year)}$ and $T-t = 1\text{(year)}$ ........................................... 260

11.6 The Parisian up-and-out call option VS the corresponding Par\textit{asian}
option at different $J$ values. Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$,
$K = $10, $\bar{S} = $12, $\bar{J} = 0.2\text{(year)}$, $T-t = 1\text{(year)}$ .................................. 262
List of Tables

2.1 Comparison of the stability requirement. Model parameters are $\kappa = 2.5$, $\eta = 0.16, r = 0.1, \rho = 0.1, T-t = 0.25(\text{year}), K = 10.0$. ............................................. 65

2.2 Comparison of the computed option prices with the reference solutions. Model parameters are $\kappa = 2.5$, $\eta = 0.16$, $\sigma = 0.45$, $r = 0.1$, $\rho = 0.1$, $T-t = 0.25(\text{year}), K = 10.0$. ............................................. 66

2.3 Comparison of the computed option prices with the reference solutions. Model parameters are $\kappa = 5$, $\eta = 0.16$, $\sigma = 0.9$, $r = 0.1$, $\rho = 0.1$, $T-t = 0.25(\text{year}), K = 10.0$. ............................................. 68

2.4 EOC in the time direction, calculated with $\Delta x = \frac{X_{\text{max}}}{250}, \Delta v = \frac{V_{\text{max}}}{250}$. ............................................. 70

2.5 EOC in the $x$ direction, calculated with $\Delta \tau = \frac{T}{100000}, \Delta v = \frac{V_{\text{max}}}{400}$. ............................................. 70

2.6 EOC in the $v$ direction, calculated with $\Delta \tau = \frac{T}{100000}, \Delta x = \frac{X_{\text{max}}}{400}$. ............................................. 70

2.7 EOC with varying grid sizes, with $\Delta x_1 = \frac{X_{\text{max}}}{10}, \Delta v_1 = \frac{V_{\text{max}}}{10}$, and $\Delta \tau_1 = \frac{T_{\text{max}}}{5}$. ............................................. 71

2.8 Report on CPU time VS relative error ............................................. 71

3.1 The test example. Parameters are $\kappa = 2$, $\eta = 0.2$, $\sigma = 0.04$, $r = 0.5$, $\rho = 0.1$, $K = 10.0$. ............................................. 82

3.2 Comparison of the option prices at $v = 0.4$, calculated by the PSOR method and the SC method. Model parameters are $\kappa = 0.8$, $\eta = 0.45$, $\sigma = 0.4$, $r = 0.4$, $\rho = -0.1$, $K = 10.0$. The iteration convergence tolerance for PSOR is $10^{-10}$. ............................................. 83
3.3 “Special” v with \( \rho = 0.1, \sigma = 0.45, r = 0.1, K = $10.0 \) \hspace{1cm} 86

3.4 “Special” v with \( \kappa = 1, \sigma = 0.45, r = 0.1, K = $10.0 \) \hspace{1cm} 86

3.5 “Special” v with \( \rho = 0.1, \kappa = 1, r = 0.1, K = $10.0 \) \hspace{1cm} 86

4.1 The variations of RMSREs when the numbers of the time steps or equal-size elements are gradually increased. \( M \): the number of time intervals; \( N \): the number of elements. \hspace{1cm} 108

4.2 The variations of RMSREs with different-size elements and time steps. \( M \): the number of time intervals; \( N \): the number of elements. \hspace{1cm} 109

5.1 Relative errors of the new approximation \hspace{1cm} 123

7.1 Report on the relative error \hspace{1cm} 161

8.1 Comparison of option prices with different spot variance values. Parameters are \( \rho = 0.1, \sigma = 0.04, r = 0.1, \kappa = 0.05, K = $10.0 \) \hspace{1cm} 174

8.2 Comparison of optimal exercise price with different spot variance values. Parameters are \( \rho = 0.1, \sigma = 0.04, r = 0.1, \kappa = 0.05, K = $10.0 \) \hspace{1cm} 174

8.3 Accuracy of the approximation \hspace{1cm} 176

11.1 The price of the Parisian up-and-out call options across the barrier, i.e., \( W(t) \), with four different times to maturity. Parameters are \( \sigma = 10\% \), \( r = 5\% \), \( D = 0 \), \( K = $10 \), \( \bar{S} = $12 \), \( J = 0.2 \) (year) \hspace{1cm} 259

11.2 The price of the \textit{Parasian} up-and-out call options across the barrier, i.e., \( W(t,J) \), with four different times to maturity. Parameters are \( \sigma = 10\% \), \( r = 5\% \), \( D = 0 \), \( K = $10 \), \( \bar{S} = $12 \), \( J = 0.2 \) (year) \hspace{1cm} 259
Chapter 1

Introduction and Background

1.1 Option derivatives

The trading of option derivatives can be traced back to the early 1970’s. Before 1973, all option contracts were what is now called ‘OTC’ (over the counter). That is, they were individually negotiated by a broker on behalf of two clients, one being the buyer and the other the seller. Trading on an official exchange began in 1973 on the CBOE (Chicago Board Options Exchange). Since then, there has been a dramatic growth in option markets. Options are now traded on many exchanges throughout the world. The OTC market still exists, and options are written by institutions to meet various clients’ needs.

When an option contract is initiated, there must be two sides to the agreement. On one side of the contract is the buyer, who has the right, but not the obligation, to engage in some specific transaction on the asset. On the other side is the party who incurs the obligation to fulfill the transaction requested by the buyer. The latter is called the writer of the option.

1.1.1 Vanilla options

There are two basic types of options (colloquially called vanilla options), namely, the call options and the put options. A call option gives the holder the right to buy the
underlying by a certain date for a certain price, while a put option gives the holder the right to sell the underlying by a certain date for a certain price. The price in the contract is known as the exercise price or the strike price, and the date in the contract is named as the expiration date, exercise date, or maturity. Another classification of the option derivatives is the European and American classification, which has nothing to do with the continent of origin but refers to a technicality in the option contract. Specifically, an American option is the one that can be exercised at any time prior to expiry, while the European option is the one that can be exercised only at the expiry. The American option usually values more than its European counterpart, since the American option gives its holder greater rights than the European option, via the right of early exercise. To the mathematicians, American options are more interesting since they can be formulated as a free boundary problem [97].

1.1.2 Exotic options

Other types of options are the so-called exotic options. Broadly speaking, an exotic option is an option whose payoff is more general and different from the vanilla call or put. Nowadays, most exotic options are traded in the OTC market, and are designed by financial institutions to meet various requirements of their clients.

One of the simplest types of the exotic options is the so-called barrier option, the distinctive feature of which is that the payoff depends not only on the final price of the underlying, but also on whether or not the underlying has reached some barrier level when the option is alive. Nowadays, a number of different types of barrier options are regularly traded in the OTC market. They are attractive to some market participants because they are less expansive than the corresponding vanilla options. The barrier options can be classified as either knock-out options or knock-in options. A knock-out option is one where the option is nullified prior to expiry if the underlying touches the barrier or out strike price, while a knock-in option is an option that comes into existence only when the underlying reaches the barrier. If the knock-in and knock-out features are activated only when the underlying reaches the
barrier for a pre-specified length of time (rather than one touch of the barrier), these special barrier options are named as Parisian-type options, which are popularly used in the OTC markets.

According to different ways of measuring the time spent above or below the barrier, the Parisian-type options can be further classified into two categories, namely, the Parisian options and the Parisian options. More precisely, if the time spent above or below the underlying barrier is reset to zero each time the underlying price crosses the barrier, this type is refereed to as continuous Parisian options, or simply Parisian options, whereas if one sums the time spent below or above the barrier without resetting the counting time to zero each time the underlying touches the barrier, these options are named as cumulative Parisian options, or simply Parisian options. Hereafter, these two options are referred to as “Parisian-type options” unless otherwise stated.

At least two reasons can be employed to illustrate the popularity of the Parisian-type options. Firstly, comparing with the classical barrier options, the Parisian-type options are more robust against short-term movements of the underlying price, since a single “touch” of the underlying barrier cannot trigger the knock-in or knock-out features of these options. Secondly, the hedging problem close to the barrier, which is usually encountered in the trading of classical barrier options, can be somehow reduced, or at least “smoothed”, in the Parisian-type contracts [49].

1.2 Mathematical background

One of the major challenges in today’s financial industry is to determine the prices of financial derivatives efficiently and accurately. Such evaluations certainly require some advanced mathematics. In this section, we shall briefly review all the mathematical knowledge that is employed as a basic tool for the studies in the current thesis.
1.2.1 Classical numerical approaches revisited

With the fast development of computer science, the numerical approach has become a crucial tool for the qualitative and quantitative analysis of today’s financial market. In this subsection, three fundamental numerical approaches, namely, the finite difference scheme, the spectral method, and the finite element method, will be briefly reviewed.

The finite difference scheme

The FDM (finite difference method) is a means of finding numerical solutions to PDEs (partial differential equations) as well as linear complementarily problems. It is a very powerful and flexible technique, if applied correctly, capable of generating accurate numerical solutions to many problems arising in both the physical and financial sciences.

The underlying idea of the FDM is to replace the partial derivatives occurring in the PDEs by approximations based on Taylor series expansion of functions near that point or points of interest. There are three commonly used finite difference approximations, namely, the forward approximation, the backward approximation, and the central approximation, which are defined as

\[
\frac{f(x + \Delta x) - f(x)}{\Delta x} + \mathcal{O}(\Delta x), \quad \frac{f(x) - f(x - \Delta x)}{\Delta x} + \mathcal{O}(\Delta x), \quad \frac{f(x + \Delta x) - f(x - \Delta x)}{2\Delta x} + \mathcal{O}(\Delta x^2),
\]

respectively.

The most basic property that a FDM must have in order to be useful is that its solution approximates the solution of the corresponding PDE, and moreover, the accuracy of the approximation should improve as the grid spacings tend to zero. Such a scheme is called a convergent scheme. Proving a given scheme to be convergent is not easy in general, if attempted in a direct manner. The Lax-Richtmyer equivalence
Chapter 1.

Theorem, which is a fundamental theory of the FDM, reveals a way to check the convergence of a given FDM.

**Theorem 1** *(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 this theorem, it is clear that the more difficult result, the convergence, can be replaced by the equivalent and easily verifiable conditions of consistency and stability. Note that a numerical algorithm is said to be stable if a small error at any stage produces a smaller cumulative error.

*Spectral methods*

Spectral methods are another class of techniques often used in applied mathematics and scientific computing to numerically solve certain PDEs. Where applicable, spectral methods have excellent error properties, with the so-called “exponential convergence” being the fastest possible. Spectral methods usually include, but are not limited to, Fourier series methods for periodic geometry problems, polynomial spectral methods for finite and unbounded geometry problems, pseudo-spectral methods for highly nonlinear problems, and spectral iteration methods for fast solution of steady state problems.

The central idea of the spectral methods is to approximate the solution of a given problem as a sum of very smooth basis functions, i.e.,

\[ f(x) \approx \sum_{k=0}^{N} a_k \Phi_k(x), \]

where the \( \Phi_k(x) \) are polynomials or trigonometric functions, and the \( a_k \) are coefficients, which need to be determined during the solution process. With all the coefficients \( a_k \) available, the solution of the given problem becomes straightforward.
The finite element method

The FEM (finite element method) is also a commonly used numerical technique for finding approximate solutions of PDEs as well as of integral equations. The solution approach usually begins with the division of the whole computational domain into several subdomains, namely, the elements, which are interconnected at specified joints, called the nodes or the nodal points. The behavior of each element is then represented by an element-stiffness matrix relating the forces at the nodal points of an element to the nodal displacements. The next step is to assemble all the element-stiffness matrices and the element-loading vector to obtain a master stiffness matrix and a total applied-load vector. Finally, the unknown nodal values are found in an efficient manner. The most attractive feature of the FEM is its ability to handle complicated geometries (and boundaries) with relative ease.

It should be remarked that the FDM is a special case of the FEM. One can clearly see that if the basis functions are chosen as either piecewise constant functions or Dirac delta functions, the FEM degenerates to the FDM.

One should also notice that the spectral method and the FDM are closely related and built on with almost the same ideas. The main difference between them is that the spectral method approximates the solution as linear combination of continuous functions that are generally nonzero over the whole computational domain, while the FEM approximates the solution as a linear combination of piecewise functions that are nonzero on each subdomain, i.e., the element. Consequently, the spectral method takes on a global approach while the FDM is considered as a local approach.

1.2.2 The perturbation method

Besides the numerical approaches mentioned previously, another branch of mathematics that lends itself to option pricing related problems is perturbation theory. Perturbation theory comprises mathematical methods that are used to find an approximate solution to a problem which cannot be solved exactly, by starting from the exact solution of a related problem. Perturbation theory is applicable if the problem
at hand can be formulated by adding a “small” term to the mathematical description of an exactly solvable problem.

The perturbation method usually begins with the assumption that the desired solution $y$ can be written in terms of a formal power series, i.e., $y = \sum_{n=0}^{\infty} \epsilon^n y_n$, in which $y_0$ is usually referred to as the leading order term, while $y_1, y_2, \cdots$ represent the high order terms. By substituting such an expression into the original system, a series of simplified systems for $y_n$ can be obtained and solved. If the approximated solution $\sum_{n=0}^{N} \epsilon^n y_n$ is accurate over the entire domain, to the order of $O(\epsilon^{N+1})$, one could say that the target problem can be solved by using the regular perturbation method.

However, in a large class of physical related problems, the approximation solution produced by the regular perturbation is accurate only on some subdomain of the entire domain, leaving other subdomains consisting of one or more small areas in which the approximation is inaccurate, generally as a result of the truncated perturbation terms being not negligible there. These areas are referred to as transition layers, or boundary or interior layers depending on whether they occur at the boundary of the domain (as is the usual case in applications) or inside the domain.

To find an overall accurate solution for the problems which could not be satisfactorily solved by using the regular perturbation method, the singular perturbation method was proposed and applied. Such a method is usually implemented within the following steps. Firstly, try to construct a regular expansion in the original variable. This is possible outside the boundary layer, and the expansion is usually called the outer solution. Secondly, construct in the boundary layer(s) a local expansion in an appropriate local variable. The local expansion obtained is usually named as the inner solution or the boundary layer expansion. Thirdly, match the inner and outer solutions to obtain a formal expansion for the entire domain. This former expansion, which is valid in the whole domain, is sometimes called a uniform expansion. Lastly, prove that the formal expansions, obtained in the last three stages, represent valid asymptotic approximation of the solution of the target problem.


**1.2.3 Stochastic calculus**

In this section, some important modeling tools of studying the problem of pricing and hedging financial derivatives are revisited.

**Itô’s Lemma**

Itô’s lemma is the most important result about the manipulation of random processes. It is to functions of random processes what Taylor’s theorem is to functions of deterministic variables. In other words, Itô’s lemma relates the small change in a function of a random variable to the small change in the random variable itself.

**Theorem 2 (Itô’s Lemma).** If a random variable $G$ satisfies the following SDE (stochastic differential equation)

\[
dG = A(G, t)dW + B(G, t)dt,
\]

then any twice differentiable function $f = f(G)$ admits the stochastic dynamics given by

\[
df = A \frac{df}{dG} dW + (B \frac{df}{dG} + \frac{1}{2} A^2 \frac{d^2 f}{dG^2}) dt.
\]

**Connections between SDE and PDE**

To study the pricing problem from a PDE standpoint, just as we did in the current thesis, the Feynman-Kac theorem and Kolmogoroff (Fokker-Plank) backward equation, which bring the relationship between the SED and PDE to light, must be stressed.

**Theorem 3 (Kolmogoroff Backward Equation and Feynman-Kac theorem)**

Consider

\[
d\bar{x}_t = \mu(t, \bar{x}_t)dt + \sigma(t, \bar{x}_t)dW_t,
\]

(1.2.1)

where $\mu_i(t, x)$ and $\sigma_{ij}(t, x)$ 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 appropriate constants. Let \( f(x) : \mathbb{R}^d \to \mathbb{R} \), \( g(t, \bar{x}) : [0, T] \times \mathbb{R}^d \to \mathbb{R} \), \( K(t, \bar{x}) : [0, T] \times \mathbb{R}^d \to \mathbb{R} \) be continuous functions and satisfy

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

Suppose that \( V(t, \bar{x}) \) is continuous and belongs to \( C^{1,2}([0, T] \times \mathbb{R}^d) \), and moreover, satisfies the Cauchy problem:

\[
\begin{cases}
\frac{\partial V}{\partial t} + \mathcal{A}_t V(t, \bar{x}) - K(t, \bar{x})V + g(t, \bar{x}) = 0,
V(T, \bar{x}) = f(\bar{x}),
\end{cases}
\]

as well as the polynomial growth condition

\[
\max_{0 \leq t \leq T} |V(t, \bar{x})| \leq M(1 + \| \bar{x} \|^{2\mu}), \; \bar{x} \in \mathbb{R}^d, \; M > 0, \; \mu \geq 1,
\]

where \( \mathcal{A}_t \) is the infinitesimal generator of the stochastic process \((1.2.1)\). Then \( V(t, \bar{x}) \) admits the stochastic representation:

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

### 1.3 Option pricing models

A good pricing model should produce the price of a financial derivative close to its real market price. It has been widely acknowledged that the popular Black-Scholes model,
which relates the derivative prices to the current stock prices and the quantified risk through a constant volatility parameter, cannot capture the behavior of today’s 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 this section, we shall give an overview of different pricing models for option derivatives.

1.3.1 The 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 underlying and serves as a benchmark against which other models can be compared. In this model, the underlying is assumed to satisfy the SDE of a geometric Brownian motion:

\[ dS_t = \mu S_t dt + \sigma S_t dW, \]

where \( \mu \) is the drift, \( W \) is a standard Brownian motion, and \( \sigma \) is the constant volatility, which measures the standard deviation of the stock returns \( dS_t / S_t \). It is also assumed that there is a money market security (bank account) paying a continuously compounded annual rate \( r \). Moreover, the security markets are perfect meaning that one can trade continuously with no transaction costs and no arbitrage possibilities. Under these main assumptions, a portfolio consisting of one option and a quantity \( -\Delta \) of the underlying can now be constructed. The value of this portfolio is \( \Pi = V - \Delta S \). The change in the value of this portfolio in one time-step is \( d\Pi = dV - \Delta dS \). Then, according to the no-arbitrage assumption, \( \Pi \) must instantaneously earn the risk-free bank rate \( r \), \( d\Pi = r\Pi dt \). To eliminate the stochastic component of risk contained in the portfolio \( \Pi \), the number of the underlying must equal to \( \Delta = \frac{\partial V}{\partial S} \). Applying the Itô lemma to \( V \) and with some algebraic manipulations, the famous Black-Scholes equation

\[ \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0 \]  

(1.3.2)
Chapter 1.

can be obtained.

Three remarks about the derivation of the Black-Scholes equation should be made. Firstly, the Delta here shows the rate of change of the value of the option or portfolio of options w.r.t (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. Secondly, the linear differential operator $L_{BS}$, i.e.,

$$L_{BS} = \frac{\partial}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2}{\partial S^2} + r S \frac{\partial}{\partial S} - r I,$$

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. Thirdly, the Black-Scholes equation (1.3.2) does not contain the growth parameter $\mu$. In other words, the value of an option is independent of how rapidly or slowly the asset grows.

One of the most remarkable contributions that Black and Scholes have made to quantitative finance research is that they derived a closed-form analytic solution, known now as the Black-Scholes formula, for the price of a European call option by solving the Black-Scholes equation together with appropriate boundary conditions. The Black-Scholes formula reads:

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

where

$$d_1 = \frac{\ln(S/K) + (r + \sigma^2/2)}{\sigma \sqrt{T-t}},$$
$$d_2 = d_1 - \sigma \sqrt{T-t},$$
Chapter 1.

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

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

Whilst appealing, it has been widely recognized that there are several imperfections in the Black-Scholes model. They are evidenced by the fact that practitioners need to change the volatility parameter on a regular basis to reflect the latest information. Also, empirical evidence suggests that implied volatilities depend on the strike price (the volatility smile or smirk effect), and on the time to maturity (the volatility term structure effect). To handle such imperfections, the literature has provided several alternative ways, which will be studied in the next two subsections.

1.3.2 The local volatility model

In the literature, there have been many alternative models proposed to remedy the drawback of the constant volatility assumption lying in the heart of the Black-Scholes model. One popular class of such alternative models is the local volatility model, which treats the volatility as a function of both the underlying $S$ and the current time $t$, rather than a constant appearing in the Black-Scholes model. In particular, under the local volatility model, the underlying $S_t$, as a function of time, is assumed to follow a diffusion process:

$$dS_t = \mu S_t dt + \sigma(S, t) S_t dW,$$

where the constant $r \geq 0$ denotes the risk-free interest rate and the deterministic function $\sigma(S, t)$ represents the local volatility.

Local volatility models are useful in any option market in which the volatility of the underlying is predominantly a function of the level of the underlying, such as the interest-rate derivatives. Time-invariant local volatilities are claimed to be able to provide the best average hedge for equity index options. Moreover, local volatility models are nonetheless useful in the formulation of stochastic volatility models.
One of the attractive features of the local volatility models is that the local volatility function to match the theoretic option prices and the market prices can be deduced once the prices of call or put options are given. It is known that under the local volatility model, if no-arbitrage prices of European vanilla options are available for all strike \( K \) and expiries \( T \), \( \sigma(K, T) \) can be extracted analytically from these option prices utilizing the well known Dupire formula \([32]\), i.e.,

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

(1.3.3)

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

It should be pointed out that there are several potential problems associated with the local volatility models. Firstly, it is usually difficult to obtain the right local volatility for this model, as the prices of some financial instruments are not always enough or even available. Secondly, the local volatility calculated by using the Dupire formula does not have sufficient accuracy, because the numerator and the denominator of (1.3.3) may become quite small when the strikes are far in- or out-of the money. Lastly, 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.

### 1.3.3 Stochastic volatility models

Another 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 to make it a stochastic process, which is conceptually different from the fitting approach of the local volatility models. In fact, modeling volatility as a stochastic process is motivated \textit{a priori} by empirical studies of stock-price returns in which the estimated volatility is observed to exhibit the “random” characteristics. In addition, the effects of transaction costs show up, under many models, as uncertainty in the volatility. Moreover, the distribution of the fat-tailed returns can be simulated by stochastic
volatility, and the market “jump” phenomena are often best modeled as volatility jump process. A stochastic volatility model is, therefore, not a simple remedy for one particular assumption of the Black-Scholes framework, but rather a powerful modification that is able to describe a much more complex market.

In the literature, several alternative stochastic volatility models have been proposed, such as the Stein & Stein model [95], the Schöbel & Zhu model [89], the GARCH diffusion model [73], and etc.. Due to the great analytical tractability for European options, the model proposed by Heston [51] becomes one of the most remarkable stochastic volatility models among all the others that can be considered. In this model, Heston assumed that the underlying $S_t$, as a function of time, follows the SDE of a geometric Brownian motion:

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dw_1,$$

where $\mu$ is the drift rate, $w_1$ is a standard Brownian motion, and $\sqrt{v_t}$ is the standard deviation of the stock returns $dS_t/S_t$. Furthermore, he assumed that the variance $v_t$ (the square of the volatility) is governed by the following mean-reverting SDE:

$$dv_t = \kappa (\eta - v_t) dt + \sigma \sqrt{v_t} dw_2. \quad (1.3.4)$$

Here, $\eta$ is the long time mean of $v_t$, $\kappa$ is the rate of relaxation to this mean, and $\sigma$ is the volatility of the volatility. $w_2$ is also a standard Brownian motion, and it is related to $w_1$ with a correlation factor $\rho \in [-1, 1]$. (1.3.4) is known in financial literature as the Cox-Ingersoll-Ross (CIR) process and in mathematical statistics as the Feller process [35].

Unlike the Black-Scholes case where only one of the randomnesses, that of the underlying price, can be hedged, the present case also involves the volatility as another source of randomness. Therefore, we set up a portfolio $\Pi$ containing one option $C(S, v, t)$, a quantity $\Delta_1$ of the stock, and a quantity $\Delta_2$ of another option $U(S, v, t)$
Chapter 1.

that is used to hedge the volatility, i.e.,

$$\Pi = C - \Delta_1 S - \Delta_2 U.$$ 

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

$$d\Pi = dC - \Delta_1 dS - \Delta_2 dU.$$ \hfill (1.3.5)

Now, applying Itô’s lemma to $dC$, we obtain

$$dC = \left[ \frac{\partial C}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + \frac{1}{2} \sigma^2 v \frac{\partial^2 C}{\partial v^2} + \rho \sigma v S \frac{\partial^2 C}{\partial S \partial v} \right] dt + \frac{\partial C}{\partial S} dS + \frac{\partial C}{\partial v} dv.$$ 

Again, we apply Itô’s lemma to $dU$ and obtain

$$dU = \left[ \frac{\partial U}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 U}{\partial S^2} + \frac{1}{2} \sigma^2 v \frac{\partial^2 U}{\partial v^2} + \rho \sigma v S \frac{\partial^2 U}{\partial S \partial v} \right] dt + \frac{\partial U}{\partial S} dS + \frac{\partial U}{\partial v} dv.$$ 

Combining these two expressions $dC$ and $dU$, we can write (1.3.5) as

$$d\Pi = Adt - \Delta_2 B dt + \left( \frac{\partial C}{\partial S} - \Delta_2 \frac{\partial U}{\partial S} - \Delta_1 \right) dS + \left( \frac{\partial C}{\partial v} - \Delta_2 \frac{\partial U}{\partial v} \right) dS.$$ \hfill (1.3.6)

In order for the portfolio to be hedged against movements in the stock and against volatility, the last two terms in (1.3.6) involving $dS$ and $dv$ must be zero. This implies that the hedging parameters must be

$$\Delta_1 = \frac{\partial C}{\partial S} - \Delta_2 \frac{\partial U}{\partial S},$$

$$\Delta_2 = \left( \frac{\partial C}{\partial v} \right) / \left( \frac{\partial U}{\partial v} \right).$$

Moreover, since the portfolio must earn the risk free rate $r$, we have

$$d\Pi = r(C - \Delta_1 S - \Delta_2 U)dt$$ \hfill (1.3.7)
Thus the change in the value of the riskless portfolio now reads

\[ d\Pi = (A - \Delta_2 B)dt = r(C - \Delta_1 S - \Delta_2 U)dt, \]

which yields \( A - \Delta_2 B = r(C - \Delta_1 S - \Delta_2 U). \) Now, with the values of \( \Delta_1 \) and \( \Delta_2, \) we have

\[ A - \Delta_2 B = rC - rS \left( \frac{\partial C}{\partial S} - \Delta_2 \frac{\partial U}{\partial S} \right) - \Delta_2 rU, \]

which yields

\[ A - rC + rS \frac{\partial C}{\partial S} = \Delta_2 \left( B - rU + rS \frac{\partial U}{\partial S} \right) \]
\[ = \left( B - rU + rS \frac{\partial U}{\partial S} \right) \left( \frac{\partial C}{\partial v} \right) / \left( \frac{\partial U}{\partial v} \right). \]

Therefore, we obtain

\[ \left( \frac{1}{2} vS^2 \frac{\partial^2 C}{\partial S^2} + \rho \sigma v S \frac{\partial^2 C}{\partial S \partial v} + \frac{1}{2} \sigma^2 v \frac{\partial^2 C}{\partial v^2} + rS \frac{\partial C}{\partial S} - rC + \frac{\partial C}{\partial t} \right) \frac{\partial C}{\partial v} = \]
\[ \left( \frac{1}{2} vS^2 \frac{\partial^2 U}{\partial S^2} + \rho \sigma v S \frac{\partial^2 U}{\partial S \partial v} + \frac{1}{2} \sigma^2 v \frac{\partial^2 U}{\partial v^2} + rS \frac{\partial U}{\partial S} - rU + \frac{\partial U}{\partial t} \right) \frac{\partial U}{\partial v}. \] \hspace{1cm} (1.3.8)

Clearly, the left-hand side of (1.3.8) is a function of \( C \) only, while the right-hand side is a function of \( U \) only. The only way that (1.3.8) can be satisfied is to let both sides of (1.3.8) equal to some function depending on \( S, v, t \), i.e., \( f(S, v, t) \). The PDE governing the price of financial derivatives with stochastic volatility could thus be arrived as

\[ \frac{1}{2} vS^2 \frac{\partial^2 C}{\partial S^2} + \rho \sigma v S \frac{\partial^2 C}{\partial S \partial v} + \frac{1}{2} \sigma^2 v \frac{\partial^2 C}{\partial v^2} + rS \frac{\partial C}{\partial S} - rC + \frac{\partial C}{\partial t} = - (\alpha - \beta \lambda) \frac{\partial C}{\partial v}, \]

where, without loss of generality, we set the arbitrary function \( f(S, v, t) \) to be \(- (\alpha - \beta \lambda)\). Note that \( \lambda \) is the so-called market price of volatility risk, because it shows the quantity of the expected return of \( C \) associated with the risk (standard deviation) of \( v \) in the Capital Asset Pricing Model framework.
1.4 Literature review

1.4.1 European and American options

Since the traded options in today’s financial markets are either European or American, considerable research interests have been drawn to develop appropriate valuation approaches for both European and American options.

In the literature, the pricing of European options has been well documented. Due to its great analytic tractability, various closed-form analytic solutions have been found for European-style options. For instance, Black and Scholes [10] obtained an analytical formula for valuing European call options under the Black-Scholes framework. Their work has had a significant influence on the development of the quantitative finance research. Later, by realizing that the Black-Scholes model is inadequate to describe the asset returns and the behavior of the option markets, Heston [51] extended previous work to the stochastic volatility model, and also derived a semi-analytic formula for European style options. Whilst elegant, his formula displays some undesired characteristics. For example, his formula still requires the valuation of a logarithm with a complex argument involved in the inverse Fourier transform. The main disadvantage of a solution being left in such a form lies in the fact that the numerical calculation of these integrals needs to be handled carefully, as the integrand is multi-valued, which may sometimes give rise to numerical instability [66]. Furthermore, it is also suggested that the numerical calculation of Heston’s formula could be made extremely complicated by the fact that the integrand is typically of oscillatory nature (e.g., [18, 66]). In the literature, some authors focused entirely on the numerical implementation of Heston’s formula [18, 66].

On the other hand, for quite a long time, it has been widely acknowledged that the pricing of American options is a “much more intriguing” problem [55, 65, 75]. The challenge stems from the nonlinearity associated with the inherent characteristic that an American option can be exercised at any time during its lifespan. This additional
right of being able to exercise the option early, in comparison with a European option, turns the American option pricing problem into a free boundary problem, which is far more difficult to deal with, even under the assumption of the classical Black-Scholes model. Recently, there has been a major breakthrough in American option pricing problem as an analytical closed-form pricing formula was successfully derived by Zhu [115]. Whilst appealing, this formula is only valid under the assumption of the Black-Scholes model, and moreover, the extension of the method adopted to alternative models does not seem to be promising.

In the literature, of all the approximation methods, there are predominately two types, analytical approximations and numerical methods for the evaluation of an American option contract under the Black-Scholes model. Typical methods in the first category include the compound-option approximation method [46], the quadratic approximation method [9, 76], the randomization approach [15], the integral-equation method [17, 64, 67], and the Laplace transform method [116]. However, as pointed out by Wu and Kwok [112], the generalization of these quasi-analytical approaches to some exotic options may not be promising. Moreover, it seems quite difficult to extend all these methods mentioned above to high-dimensional problems. On the other hand, the numerical methods for the valuation of American options typically include the FDM [14, 91, 112, 125], the FEM [4], the radius basis function method [53], the binomial tree method [27], the Monte Carlo simulation technique [48], and the least square approach [75]. It should be remarked that for most of the numerical methods mentioned above, the convergence analysis has not yet been achieved. Of course, most authors demonstrated the convergence through their numerical examples. But, lack of convergence proof requires caution when these approaches are adopted on a case-by-case basis. This has motivated us to consider a “convergence-proved” approach for the valuation of American options, which should also be suitable for the practical needs of market practitioners. Such an issue forms the main content of Chapter 4.

When the pricing of American options is considered under the Heston model,
neither the analytical approximation nor the numerical approaches mentioned above can be easily extended, primarily due to the fact that, under stochastic volatility, the optimal exercise price depends, in addition to time, on the dynamics of volatility. In other words, the introduction of a second stochastic process has considerably complicated the solution process for pricing American options. In the last decade, several numerical approaches based on the FDM were introduced to solve American option prices under the Heston model. For instance, Clarke and Parrott [24], used a special version of a projected full approximation scheme with multigrid to solve the American option pricing problem. One advantage of such a multilevel method is that the number of iterations required to solve a linear complementarity problem is essentially independent of the grid size. However, their method is rather complicated because of the use of a special projected linear Gauss-Seidel smoother. Ikonen and Toivanen [60] calculated the option values by using the operator splitting method, in which an auxiliary variable is used to improve the accuracy. However, their method still requires a relatively large amount of computational storage space. Zvan et al. [127] applied the penalty method to the American option pricing problem. Their method is simpler than the one used by Clark and Parrott, but still needs a relatively large amount of computational resources to produce an accurate result. Recently, Zhu and Chen [119] proposed a predictor-corrector scheme based on the ADI (alternating direction implicit) method to deal with the valuation of American options with stochastic volatility, which forms the main contents of Chapter 2.

On the other hand, pricing a perpetual American option, which is an option that can be exercised at any time but with no expiration date, has also drawn considerable research interest. This simple contract can be viewed as an approximation for a long-dated American option. Also, analyzing such options may in principle be used as a building block in an approximation procedure for American options with finite maturities [15]. In the literature, the most remarkable step in studying perpetual American options is Merton’s closed-form solution for perpetual American puts with constant volatility [79]. Unfortunately, his approach cannot be extended directly
to stochastic volatility models, primarily due to the fact that the optimal exercise price is now an unknown curve of the volatility, rather than a constant appearing in the Black-Scholes model. To date, no documented schemes have been proposed to address this issue; efficient approximation approaches for the valuation of perpetual American options with stochastic volatility are discussed in Chapter 3, Chapter 8-Chapter 10 of this thesis.

Another important issue associated with the pricing of an American option is that its optimal exercise price is singular when the time is close to the expiry. In the literature, some analyses on the asymptotic behavior of the optimal exercise price near expiry have already been carried out for the Black-Sholes model. For instance, Barles et al. [8] derived the leading order term for the optimal exercise price of American puts on a non-dividend paying underlying by constructing a subsolution as well as a supersolution. The examples in their paper considered options with maturities less than one year, in which case, the inaccuracy of the approximation was not significant. Chen and Chadam [20] analyzed the same problem as Barles et al., and provided four approximations for the optimal exercise prices of American puts with different maturities by the method of integral equations. Evans et al. [33] considered the asymptotic behavior of the optimal exercise price near expiry on an asset with constant dividend yield rate. The approximation was derived both with the utilization of the method of integral equations and the method of matched asymptotic expansions, and was expected to be useful for a time scale in the order of days and weeks from expiry. Moreover, recently, in Chapter 6, Chen and Zhu [19] extended Evans et al.’s results to the local volatility model, in which the volatility is assumed to be dependent on both time and the underlying asset price. Their approximations degenerate to those obtained by Evans et al., when the volatility function is assumed to be a constant. Although these results are elegant and useful, they are only valid under the Black-Scholes model. Most significantly, in Chapter 7, Zhu and Chen [118] has investigated this issue under the Heston model by means of the method of matched asymptotic expansions. It turns out that the option prices
are quite different from the corresponding Black-Scholes case, while the leading order term of the optimal exercise price remains almost same as the constant volatility case if the spot volatility is given the same value as the constant volatility appearing in the Black-Scholes model.

1.4.2 Parisian-type options

With the growing demand of trading Parisian-type options, considerable research has also been done focusing on the development of different approaches that can be employed to evaluate Parisian-type options effectively. The valuation problem for Parisian-type options has been recognized as much more difficult than that of classical barrier options [70]. While a closed-form analytical solution for the latter has already been found [70], the former could only be solved approximately. Predominately two types of valuation techniques, the quasi-analytic approaches and the numerical methods are well documented in the literature. Of all the quasi-analytic methods, the most influential approach was the one proposed by Chesney et al. [21]. They used the theory of Brownian excursions and defined the value of a Parisian option in terms of an integral expressed as an inverse Laplace transform. Afterwards, their framework has been further developed by Hartly [50], Hugonnier [56], and Schröder [90] to price and hedge the Parisian-type options. Numerical methods, as another alternative, were also intensively developed recently. A typical method in this category is the PDE approach proposed by Wilmott et al. [49]. In their article, two PDE systems governing the prices of Parisian and ParAsian options are established, and then solved by using the explicit finite difference scheme. Whilst flexible and easy to implement, Zhu and Chen [120] pointed out that there are at least three major deficiencies in Wilmott et al.’s approach. Firstly, in their work [49], the singularities associated with adopting appropriate PDEs to price Parisian-type options are not explored at all; appropriately identifying these singularities and dealing with them not only make the developed PDE system (not just the PDE itself) correctly reflects what the financial clauses dictate, but also ensure any mean-
ingful numerical scheme would lead to the correct solution. Secondly, some boundary conditions set in their framework do not seem to correctly represent the corresponding financial clauses of these options. Furthermore, Zhu and Chen [120] believe that at least one boundary condition connecting the pricing domains has been totally overlooked by them, and thus their pricing systems are not properly closed. Lastly, in terms of the explicit finite difference method they adopted, it is only conditionally stable, resulting in computational inefficiency and low accuracy, especially for the 3-D problems arising from the pricing of Parisian-type options.

Based on several reasonable financial arguments, Zhu and Chen [120] have re-established the pricing systems for Parisian-type options, and derived closed-form analytic formulae for their prices. Moreover, through various numerical examples, they have shown that their solutions are accurate and efficient, and can satisfy the growing demand of trading Parisian-type options. These issues are thoroughly discussed in Chapter 11.

1.5 Structure of thesis

In this thesis, three key quantitative methods in the option pricing field, namely, the numerical schemes, the analytical approximation methods, and the exact solution approaches have been considered and extended. The thesis is organized into three parts, with each part addressing one of the issues mentioned above. Each part is further divided into several chapters, with each chapter discussing one particular topic relevant to this part.

In Part 1, we propose and test three different numerical approaches for pricing American options. This part contains three chapters, i.e., Chapter 2-Chapter 4. In Chapter 2, we introduce a new numerical scheme, based on the ADI (alternating direction implicit) method, to price American put options under a stochastic volatility model. Upon applying a front-fixing transformation to transform the unknown free boundary into a known and fixed boundary in the transformed space, a predictor-corrector finite difference scheme is then developed to solve for the optimal
exercise price and the option values simultaneously. Based on the local von Neu-
mann stability analysis, a stability requirement is theoretically obtained first and
then tested numerically. It is shown that the instability introduced by the predictor
can be damped, to some extent, by the ADI method that is used in the corrector.
The results of various numerical experiments show that this new approach is fast
and accurate, and can be easily extended to other types of financial derivatives with
American-style exercise. Another key contribution of this chapter is the proposition
of a set of appropriate boundary conditions, particularly in the volatility direction,
upon realizing that appropriate boundary conditions in the volatility direction for
stochastic volatility models appear to be controversial in the literature. A sound jus-
tification is also provided for the proposed boundary conditions mathematically as
well as financially.

In Chapter 3, based on the Legendre pseudospectral method, we propose a numer-
ical treatment for pricing perpetual American put options with stochastic volatility.
In this simple approach, a nonlinear algebraic equation system is derived first, and
then solved by the Gauss-Newton algorithm. The convergence of the current scheme
is ensured by constructing a test example similar to the original problem, and com-
paring the numerical option prices with those produced by the classical Projected
SOR (PSOR) method. Our numerical experiment suggests that the proposed scheme
is both accurate and efficient, since the spectral accuracy can be easily achieved with
a small number of iterations. Moreover, based on the numerical results, we also dis-
cuss the impact of stochastic volatility term on the prices of perpetual American
puts.

In Chapter 4, a “convergence-proved” IFE (inverse finite element) approach is
introduced to the field of financial engineering to price American options for the
first time. Without involving any linearization process at all, the current approach
deals with the nonlinearity of the pricing problem through an “inverse” approach.
Numerical results show that the IFE approach is quite accurate and efficient, and can
be easily extended to multi-asset or stochastic volatility pricing problems. The key
The contribution of this chapter to the literature is that we have managed to provide a convergence analysis for the IFE approach, which ensures that our numerical solution does indeed converge to the exact one of the original nonlinear system.

In Part 2, analytical approximation methods are further developed. This part can be further divided into six chapters, i.e., Chapter 5-Chapter 10. In Chapter 5, we apply singular perturbation techniques to price European puts with a stochastic volatility model, and derive a simple and elegant analytical formula as an approximation for the value of European put options. In contrast to the existing Heston’s semi-analytical formula, this approximation has the following unique feature: the latter only involves the standard normal distribution function, which is as fast and easy to implement as the Black-Scholes formula; whereas the former requires the evaluation of logarithm with a complex argument during the involved Fourier inverse transform, which may sometimes result in numerical instability. Various numerical experiments suggest that our new formula can achieve a high order of accuracy for a large class of option derivatives with relatively short tenor.

Chapter 6 investigates American puts on a dividend-paying underlying whose volatility is a function of both time and the underlying asset price. The asymptotic behavior of the critical price near expiry is deduced by means of singular perturbation methods. It turns out that if the underlying dividend is greater than the risk-free interest rate, the behavior of the critical price is parabolic, otherwise, an extra logarithmic factor appears, which is similar to the constant volatility case. The results of this chapter can be used as a complement to the numerical approaches to calculate the option values and the optimal exercise price in other larger times away from expiry.

Adopting the method of matched asymptotic expansions, Chapter 7 addresses the asymptotic behavior of American put options on a dividend-paying underlying with stochastic volatility near expiry. Through our analyses, we are able to show that the option price will be quite different from that evaluated under the Black-Scholes model, while the leading order term of optimal exercise price remains almost the same.
as the constant volatility case if the spot volatility is given the same value as the constant volatility in the Black-Scholes model. Results from numerical experiments also suggest that our analytical formulae derived from the asymptotic analysis are quite reasonable approximations for options with remaining time to expiry in the order of days or weeks.

In Chapter 8, we present a correction to Merton’s well-known classical case of pricing perpetual American put options by considering the same pricing problem under a stochastic volatility model with the assumption that the volatility is slowly varying. Two analytic formulae for the option price and the optimal exercise price of a perpetual American put option are derived, respectively. Upon comparing the results obtained from our analytic approximations with those calculated by a spectral collocation method, it is shown that our current approximation formulae provide fast and reasonably accurate numerical values of both the option price and the optimal exercise price of a perpetual American put option, within the validity of the assumption we have made for the asymptotic expansion. Based on the newly-derived formulae, the quantitative effect of the stochastic volatility on the optimal exercise strategy of a perpetual American put option has also been discussed. A most noticeable and interesting result is that there is a special cut-off value for the spot variance, below which a perpetual American put option priced under the Heston model should be held longer than the case of the same option priced under the traditional Black-Scholes model, when the price of the underlying is falling.

In Chapter 9, we present a “correction” to Merton’s well known classical case of pricing perpetual American puts by considering the same pricing problem under a general fast mean-reverting stochastic volatility model. By using the perturbation method, two analytic formulae are derived for the option price and the optimal exercise price, respectively. Based on the newly obtained formulae, we conduct a quantitative analysis of the impact of the stochastic volatility term on the price of a perpetual American put option as well as its early exercise strategies. It shows that the presence of a fast mean-reverting stochastic volatility tends to universally
increase the put option price and to defer the optimal time to exercise the option contract, had the underlying been assumed to be falling. It is also noted that such an effect could be quite significant when the option is near the money.

In Chapter 10, we consider the problem of pricing perpetual American put options with volatility driven by two other processes. By using a perturbation approach, we obtain approximate, but explicit, closed-form pricing formulae for the option and optimal exercise prices, respectively, under a general multi-scale stochastic volatility model. A key feature of the expansion methodology employed here is to balance the two stochastic volatility processes, while dealing with the free boundary conditions properly. It turns out that in the current formulae, the fast volatility factor does not play an explicit role, while the slow factor is quite crucial, a phenomenon that is shown to be quite reasonable through our discussions.

The last part of the thesis considers the exact solution approach for pricing option derivatives. This part contains Chapter 11. In this chapter, two exact and analytic solutions for the valuation of European-style Parisian and ParAsian options under the Black-Scholes framework are respectively presented. To the best of our knowledge, closed-form analytic formulae have never been found for the pricing of Parisian and ParAsian options, although quite a few approximate solutions and numerical approaches have been proposed. A key feature of our solution procedure is the reduction of a three-dimensional problem to a two-dimensional problem through a coordinate transform that has elegantly “absorbed” the directional derivative associated with the “barrier time” into the time derivative and thus resulted in two coupled, but simplified PDE systems. For Parisian options, the coupled PDE systems are then analytically solved by applying the Laplace transform technique in conjunction with the construction of “moving windows”, which are introduced to evaluate the option prices backwards, slide by slide, until the value of the option at a given time and trigger value is found for a given underlying price. On the other hand, due to the non-resetting mechanism of the ParAsian option, the coupled PDE systems of this type of options are much more complicated than those of their Parisian counterparts, and
the “moving window” technique fails in this case. Alternatively, the double Laplace transform technique is then applied to solve for the option prices in the Laplace space. However, our success of obtaining closed-form analytical solution hinges on overcoming the difficulty of performing Laplace inversions analytically. Finally, we have compared the results obtained from the newly-derived analytical solutions with those obtained through a numerical solution procedure. Such a comparison has not only reinforced the correctness of our newly-derived analytical solutions from numerical point of view, but also has demonstrated the efficiency of using the newly-derived analytical solutions to calculate option prices in finance practice.
Part I: Numerical approaches

Numerical methods have long been considered as a very useful and popular tool in the quantitative finance research as well as in financial markets, as they are usually fast and with acceptable accuracy. For a wide class of financial derivatives, the analytic solutions for their prices are not achievable, and the numerical approaches could be one of the best choices to determine their values accurately and efficiently. This part contributes to the literature by proposing and testing three fast and accurate numerical approaches for pricing American options. Each method forms the main content of one chapter contained in this part. More precisely, in Chapter 2, a predictor-corrector scheme based on the ADI (alternative direction implicit) method is proposed to solve the price of American puts with stochastic volatility. Most remarkably, in this chapter, we have managed to show a theoretic threshold for the conditional stability of the predictor-corrector scheme adopted to price American options. In Chapter 3, a numerical approach based on the spectral-collocation method is proposed for the valuation of perpetual American puts under the Heston model. This new method is shown to be particularly accurate and efficient, as the spectral accuracy can be easily achieved with a small number of iterations. Finally in Chapter 4, an inverse finite element method is introduced to the option pricing field for the first time. Most significantly, we have provided a convergence analysis for the inverse finite element approach, which ensures that the numerical solution does indeed converge to the exact one of the original nonlinear system.
Chapter 2

A predictor-corrector scheme based on the ADI method for pricing American puts with stochastic volatility

2.1 Introduction

It is well known that one of the most important topics in quantitative finance research is the valuation of option derivatives. Empirical evidence suggests that the Black-Scholes model, which is a breakthrough in the financial area, is inadequate to describe asset returns and the behavior of the option markets [1]. This is because the assumption on the log-normality of the value of the underlying asset has somewhat oversimplified the real process of the asset price. One possible remedy is to assume that the volatility of the asset price also follows a stochastic process [6, 38, 51, 58]. In this chapter, we shall use the stochastic model introduced by Heston for pricing American options [51].

In the last decade, several numerical approaches based on the FDM (finite difference method) are introduced to solve the free boundary problem associated with the
valuation of American options under the Heston model, as discussed in Chapter 1. However, since most of the numerical methods in the literature are either too complicated to implement or with very low computational efficiency, it is desirable to have alternative ways to deal with the valuation of American options with stochastic volatility.

In this chapter, we propose an approach based on a predictor-corrector framework, which is commonly used to numerically solve nonlinear PDEs (partial differential equations). The idea behind the predictor-corrector method is to use a suitable combination of an explicit and an implicit technique to obtain a method with better convergence characteristics. Previously, this scheme has only been applied to the pricing problem under the Black-Scholes model, such as Zhu and Zhang reported in [125]. Though their method is efficient and accurate, it can not be applied directly to the stochastic models. The purpose of this article is to introduce a new predictor-corrector scheme, which is not only suitable for the Heston model, but also for other stochastic models. In our new approach, we adopt the so-called front-fixing transformation [72] to let the unknown boundary be included in the governing equation as a nonlinear term in exchange for a fixed boundary. To tackle the nonlinear nature of American option pricing problem, which is explicitly exposed in the transformed equation, we use a predictor-corrector finite difference scheme at each time step to convert the nonlinear PDE to two linearized difference equations associated with the prediction and correction phase, respectively. The prediction phase, constructed by an explicit Euler scheme, is used to calculate the optimal exercise price, whereas the correction phase, designed by the alternating direction implicit, or ADI, method, continues to do the calculation of the option price together with the correction of the optimal exercise price. The ADI scheme used in the corrector is efficient in computing multi-dimensional problems. Moreover, it is also suggested that the good convergence property of the ADI scheme can somehow, damp the instability that might be introduced by the predictor. With the perfect combination of the explicit Euler scheme and the ADI method, the original nonlinear problem has been successfully converted
to a set of linear algebraic equations, which can be solved efficiently. In comparison with the numerical methods in the literature, the advantage of the current scheme is obvious. For example, first, our method requires almost the same storage space as a one-dimensional problem does and it will not increase even when the method is applied to option pricing problems on multi-assets. This is, however, not the case for the numerical methods proposed in [60, 61], as a substantially larger amount of the storage space is required, which will also increase as the number of the assets increases. Second, in addition to the option values, the present method captures the entire optimal exercise boundary as part of the solution procedure, whereas in [60], the optimal exercise price cannot be obtained simultaneously, and needs to be solved with some extra effort. Finally, our method requires no iterations, and can be easily extended to the valuation of American options under other models.

It is usually easy to design a numerical scheme to solve a PDE system, but much harder to provide a theoretical threshold for the stability and the convergence of the scheme. It is probably even worse to theoretically define a suitable stability criterion for the predictor-corrector method, since it is a hybrid finite difference method. For this reason, the issue of stability requirement was not even attempted in [125] for the Black-Scholes case. One could naturally imagine that with the complexity of the Heston model, it would have made a theoretical stability analysis much less achievable. Based on the local von Neumann stability analysis, combining with the "frozen" coefficient technique, which is commonly used for the stability analysis of the variable-coefficient problem [96], we have managed to not only verify that the ADI method used for the European puts under the Heston model is unconditionally stable, but also give a proper stability requirement for the predictor-corrector approach.

In the subsequent sections, we shall present this new approach together with the numerical results for American put options under the Heston model. The chapter is organized as follows: in Section 2.2, we introduce the PDE system that the price of an American put must satisfy under the Heston model, with our emphasis being placed on properly closing the system with appropriate boundary conditions, which
appear to be controversial in the literature. In Section 2.3, we present our predictor-corrector approach in detail as well as the implementation of the ADI scheme. In Section 2.4, numerical examples and some analyses are presented to demonstrate the convergence and accuracy of the current scheme. Concluding remarks are given in the last section.

2.2 Pricing American options under the Heston model

Although the Heston model has been studied by a number of authors [5, 81, 99], we still describe it in reasonable detail, in this section, for the sake of completeness of the chapter and easiness of reference for the readers. However, our emphasis will be placed on the discussion of appropriate boundary conditions in the volatility direction, which contributes greatly to the proper closeness of the PDE system to be solved numerically in the later sections.

2.2.1 The Heston model

Heston [51], assumed that the underlying $S_t$, as a function of time, follows the SDE (stochastic differential equation) of a geometric Brownian motion:

$$dS_t = \mu S_t dt + \sqrt{v_t} S_t dw_1,$$  \hspace{1cm} (2.2.1)

where $\mu$ is the drift rate, $w_1$ is a standard Brownian motion, and $\sqrt{v_t}$ is the standard deviation of the stock returns $\frac{dS_t}{S_t}$. Furthermore, he assumed that the variance $v_t$ (the square of the volatility) be governed by the following mean-reverting SDE:

$$dv_t = \kappa (\eta - v_t) dt + \sigma \sqrt{v_t} dw_2.$$  \hspace{1cm} (2.2.2)

Here, $\eta$ is the long-term mean of $v_t$, $\kappa$ is the rate of relaxation to this mean, and $\sigma$ is the volatility of the volatility. $w_2$ is also a standard Brownian motion, and it
is related to $w_1$ with a correlation factor $\rho \in [-1, 1]$. (2.2.2) is known in financial literature as the Cox-Ingersoll-Ross (CIR) process and in mathematical statistics as the Feller process [35, 38]. Various studies [5, 81, 99] suggest that it is consistent with the real market. Most remarkably, Heston found a closed-form exact solution for the price of European-style options. Unfortunately, the approach he adopted could not be easily extended to the case of American options; no analytical solution for American options under the Heston Model has yet been discovered.

### 2.2.2 The PDE for the value of American puts and the corresponding boundary conditions

Let $U(S, v, t)$ denote the value of an American put option, with $S$ being the underlying, $v$ 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 $U$ should satisfy the following bivariate PDE:

$$
\begin{align*}
\frac{1}{2}vS^2 \frac{\partial^2 U}{\partial S^2} + \rho \sigma v S \frac{\partial^2 U}{\partial S \partial v} + \frac{1}{2} \sigma^2 v \frac{\partial^2 U}{\partial v^2} + rS \frac{\partial U}{\partial S} + \left[ \kappa (\eta - v) - \lambda \sigma \sqrt{v} \right] \frac{\partial U}{\partial v} - rU + \frac{\partial U}{\partial t} &= 0, \\
\end{align*}
$$

(2.2.3)

where $\lambda$ is the market price of risk, $r$ is the risk-free interest rate. In this chapter, for simplicity, we set $\lambda$ to zero, and the extensions to the case that $\lambda$ is non-zero should be straightforward. The terminal condition to (2.2.3) is given by the payoff function. For an American put option, it reads: $U(S, v, T) = \max(K - S, 0)$, where $K$ is the strike price. Since at the expiry date, the optimal exercise price of an American put on a non-dividend underlying asset is equal to the strike price, this condition can be simplified as [115]:

$$
U(S, v, T) = 0, \text{ for } S > S_f(v, T) = K, \quad v > 0.
$$

(2.2.4)

For the valuation of American puts, a set of appropriate boundary conditions are
also needed together with the terminal condition (2.2.4) to solve (2.2.3). It is obvious that the boundary conditions along the $S$ direction are easy to justify. They are just the same as those in the Black-Scholes model. The value of a put should satisfy the far-field boundary condition,

$$\lim_{S \to \infty} U(S,v,t) = 0,$$

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 [111]:

$$U(S_f,v,t) = K - S_f,$$

$$\frac{\partial U}{\partial S}(S_f,v,t) = -1.$$ 

It should be noted that the above two conditions 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 $t$.

The boundary conditions along the $v$ direction remain unclear in the literature. Even for the European case, it is still controversial whether or not Heston’s analytical formula [51] does indeed satisfy the given boundary conditions along the $v$ direction. While some (e.g. [74]) chose the boundary conditions along the $v$ direction, for European puts, by taking the limit of the Black-Scholes formula with respect to $\sigma$, others (e.g. [114]) chose to neglect the boundary conditions along the $v$ direction altogether for American puts. From the financial point of view, there is no reason why the boundary conditions in the $v$ direction should be different when the moving
boundary for the case of American puts is associated with the $S$ direction only. There were also some authors ([24, 63]) who argued that boundary conditions along the $v$ direction are still necessary in their numerical approaches for solving the American option pricing problem under the Heston model and chose the two boundary conditions:

$$\lim_{v \to 0} U(S, v, t) = \max(K - S, 0), \quad (2.2.5)$$
$$\lim_{v \to \infty} \frac{\partial U}{\partial v}(S, v, t) = 0. \quad (2.2.6)$$

However, they did not explain exactly why the option price should be equal to the payoff function at $v = 0$. As different boundary conditions imposed will undoubtedly affect the value of an option, the controversy of what boundary conditions should be imposed in the $v$ direction has clearly jeopardized the uniqueness of the solution and thus needs to be properly investigated. In the following, we shall discuss what special price the option should be along the boundary of the $v$ direction.

Whether or not boundary conditions are needed for $v = 0$ and if they are needed, what would be their appropriate form, should be discussed from both mathematical and financial points of view.

Mathematically, $v = 0$ is a so-called “degenerate” boundary of (2.2.3), because its characteristic form vanishes as $v$ approaches zero [36]. From the mathematical point of view, boundary conditions along degenerate boundaries are not needed at all if the Fichera function is nonnegative, but should be imposed otherwise [36]. For the Heston model, it can be shown that the Fichera function along $v = 0$ equals $\kappa \eta - \frac{\sigma^2}{2}$. Therefore, if $\kappa \eta \geq \frac{\sigma^2}{2}$, the pricing system without boundary constraints at $v = 0$ is already closed, and thus there is no need to prescribe any condition along the boundary $v = 0$ at all for this case. On the other hand, if $\kappa \eta < \frac{\sigma^2}{2}$, appropriate boundary condition at $v = 0$ is still “needed” to ensure the uniqueness of the solution. However, the Fichera function does not reveal what the specific boundary conditions
should be prescribed for this case. We believe that one now has to use a financial argument to set up an appropriate boundary condition for this case.

Under the risk-neutral argument, when $v \to 0$, the leading order term of the solution of the SDE (2.1) 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 anymore 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., $[0, K) \subseteq \lim_{v \to 0} S_f(v, t)$, 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 $(K, \infty)$ belongs to the “continuous” region, i.e., $(K, \infty) \subseteq \lim_{v \to 0} S_f(v, t)$, and thus $\lim_{v \to 0} S_f(v, t) \leq K$. A combination of the above two statements leads to the conclusion that

$$\lim_{v \to 0} S_f(v, t) = K. \quad (2.2.7)$$

Having established that $S_f(v, t)$ must be equal to $K$ when $v \to 0$, it is then trivial to show

$$\lim_{v \to 0} U(S, v, t) = 0 \quad (2.2.8)$$

from the definition of $S_f$ associated with put options.

Two remarks should be made. Firstly, (2.2.8) is a simplified version of (2.2.5); it can be deduced from (2.2.5) by considering the definition of $S_f$, as well as the fact that $\lim_{v \to 0} S_f(v, t) = K$. Therefore, we believe those [24, 63] who proposed to use (2.2.5) as the boundary condition at $v \to 0$ are correct, although we argue that it is better to use the simplified version (2.2.8). Second, this boundary condition based on a financial argument is irrespective of the ratio $\frac{2\kappa \eta}{\sigma^2}$ being greater than 1 or not. Naturally, one may wonder the consistency of the two arguments for the case of $\kappa \eta \geq \frac{\sigma^2}{2}$, for which the mathematical argument shows that there is no need for any boundary condition
at \( v = 0 \), whereas the financial argument suggests that the option value vanishes there. In other words, if one also adopts the boundary condition (2.2.8) for the case \( \kappa \eta \geq \frac{\sigma^2}{2} \), will this value be consistent with the inherent value of the PDE system when \( v \to 0 \), and thus would not be an inappropriate “boundary condition” that will certainly “deteriorate” the well-posedness of the pricing problem? Note that the PDE system here refers to (2.2.16) but without boundary condition at \( v = 0 \), which is already closed for this case. Fortunately, we have managed to prove the consistency of (2.2.8) with the asymptotic behavior of the unknown function \( U(S,v,t) \) as \( v \to 0 \) when \( \kappa \eta \geq \frac{\sigma^2}{2} \). For brevity, we leave such a proof to Appendix A.1.

Combining both arguments, we can now confidently conclude that (2.2.8) should be used as the boundary condition at \( v \to 0 \) for pricing an American put option, regardless of what the ratio \( \frac{2\kappa \eta}{\sigma^2} \) is.

We now show what the boundary condition should be imposed as \( v \to \infty \). One needs to understand how the volatility impacts the option price, since \( v \) in the Heston model is nothing but the square of the volatility. Roughly speaking, the volatility of the underlying asset is a measure of the uncertainty of the future underlying price movements. As a result, if volatility increases, the probability of the underlying asset price varying within a large range would increase too, resulting in higher option prices for both puts and calls, since the holder of the option will have a chance to cash in when the movement of the underlying is in his favor and do nothing when it is not. Mathematically, this is equivalent to saying that the option price \( U(S,v,t) \) is a monotonic increasing function with respect to \( v \).

Similar to the case where the boundary conditions in the \( S \) direction for European puts and American puts are different at the lower boundary of the computational domain, those associated with European puts and American puts in the \( v \) direction are of different forms as well; they need to be discussed separately.

Firstly, we consider the case of European puts. Market observations show that when the volatility is extremely large, the option price \( U \) is independent of volatility
Chapter 2.

changes \[24\], i.e.,
\[
\lim_{v \to \infty} \frac{\partial U}{\partial v}(S, v, t) = 0.
\] (2.2.9)

In fact, this has been used by many authors as their boundary condition for the option price at the large end of \(v\) (e.g. \[24, 63\]). However, we believe that under the same argument we should be able to deduce an even simpler boundary condition that not only reflects the fact that the vega approaches zero when \(v \to \infty\), but also considerably facilitates the computation. This is achieved by realizing the fact that when \(v \to \infty\), the second-order partial derivative of \(U(S, v, t)\) with respect to \(v\) should vanish too if the vega approaches zero, i.e.,
\[
\lim_{v \to \infty} \frac{\partial^2 U}{\partial v^2}(S, v, t) = 0.
\] (2.2.10)

Then, utilizing (2.2.9) and (2.2.10), the Heston operator
\[
\mathcal{L}_H = \frac{1}{2} v S^2 \frac{\partial^2}{\partial S^2} + \rho \sigma v S \frac{\partial}{\partial S} \frac{\partial}{\partial v} + \frac{1}{2} \sigma^2 v \frac{\partial^2}{\partial v^2} + r S \frac{\partial}{\partial S} + \kappa [\eta - v] \frac{\partial}{\partial v} - r + \frac{\partial}{\partial t},
\]
degenerates to the Black-Scholes operator, with \(\sigma^2\) being substituted by \(v\),
\[
\mathcal{L}_{BS} = \frac{1}{2} v S^2 \frac{\partial^2}{\partial S^2} + r S \frac{\partial}{\partial S} - r + \frac{\partial}{\partial t}.
\]

This implies that we can use the option value calculated from the Black-Scholes formula as the boundary value of \(U(S, v, t)\) when \(\sigma \to +\infty\) under the Heston model.

In other words, through this argument, we have successfully converted the Neumann boundary condition (2.2.9) into a Dirichlet boundary condition
\[
\lim_{v \to \infty} U(S, v, t) = Ke^{-r(T-t)},
\] (2.2.11)
the implementation of which would require far less computational effort. It should be remarked that (2.2.11) is a special case of (2.2.9), which means that the solution obtained by satisfying (2.2.11) would automatically satisfy (2.2.9), but not vice versa.
Now, we consider the case of American puts. To better articulate the establishment of appropriate boundary condition for this case, we form and prove the following proposition.

**Proposition 1** When \( v \) approaches infinity, the value of an American put option reaches the strike price \( K \) asymptotically, i.e.,

\[
\lim_{v \to \infty} U_A(S, v, t) = K.
\]

**Proof.** First, we consider the effect of the expiration date on the option prices. It is a well-known fact that the value of American puts is an increasing function of the time to expiry, simply because the longer the tenor of an option is, the more “right” the holder has in terms of exercising the option [57]. Mathematically, this is equivalent to saying that the option price \( U_A(S, v, t) \) is a monotonic decreasing function of \( t \), i.e.,

\[
\lim_{v \to \infty} U_A(S, v, T) \leq \lim_{v \to \infty} U_A(S, v, t) \quad \forall t \in [0, T].
\] (2.2.12)

Moreover, an American put is always worth at least as much as its European counterpart [57], i.e.,

\[
\lim_{v \to \infty} U_E(S, v, T) \leq \lim_{v \to \infty} U_A(S, v, T).
\] (2.2.13)

From (2.2.11), we have

\[
\lim_{v \to \infty} U_E(S, v, T) = Ke^{-r(T-t)}|_{t=T} = K,
\]

which, combined with the inequalities (2.2.12) and (2.2.13), yields,

\[
K \leq \lim_{v \to \infty} U_A(S, v, t).
\] (2.2.14)

On the other hand, it is also well known that no matter what the underlying value becomes, the price of an American put option can never be worth more than its strike price [57], i.e.,

\[
\lim_{v \to \infty} U_A(S, v, t) \leq K.
\] (2.2.15)
Clearly, from Eqs. (2.2.14) and (2.2.15) one must reach the conclusion

$$\lim_{v \to \infty} U_A(S, v, t) = K. \quad (2.2.16)$$

This can also be understood from the financial point of view. From the definition of a put option, its value is bounded up by the strike price $K$ [57]. Therefore, if one can show that $K$ is also the least upper bound, one must then conclude that the option price must reach $K$ when $v \to \infty$, based on another financial intuition that any option value monotonically increases with volatility. The “least upper boundness” of $K$ can be easily established by arguing that if there was another upper bound that is less than $K$, since the infinite volatility implies that the underlying may take any value within the range of $S \in [0, \infty)$ with a probability one, the option value can always be greater than this “upper bound”. A contradiction is thus reached.

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

$$\begin{cases}
\frac{1}{2} v^2 S^2 \frac{\partial^2 U}{\partial S^2} + \rho \sigma v S \frac{\partial^2 U}{\partial S \partial v} + \frac{1}{2} \sigma^2 v^2 \frac{\partial^2 U}{\partial v^2} + r S \frac{\partial U}{\partial v} + \kappa (\eta - v) \frac{\partial U}{\partial v} - r U + \frac{\partial U}{\partial t} = 0, \\
U(S, v, T) = 0, \lim_{S \to \infty} U(S, v, t) = 0, \\
U(S_f(v, t), v, t) = K - S_f(v, t), \frac{\partial U}{\partial S}(S_f(v, t), v, t) = -1, \\
\lim_{v \to 0} U(S, v, t) = 0, \lim_{v \to \infty} U(S, v, t) = K,
\end{cases}$$

for $S \in [S_f(v, t), \infty)$, $v \in [0, \infty)$, and $t \in [0, T]$. There are several remarks before we introduce an efficient and accurate numerical scheme to solve this system in the next section. First of all, the two newly-introduced boundary conditions in the $v$ direction have coincidentally well manifested the monotonicity of the option price with respect to $v$ as well as its boundedness [57]:

$$\max(K - S, 0) \leq U(S, v, t) \leq K.$$

Or, the option price is expected to monotonically increase from its lower bound
max(\(K - S, 0\)) to its upper bound \(K\) when \(v\) varies from 0 to \(\infty\) (one should not forget that when \(v = 0\), the solution domain of the above differential system is \(S \in [K, \infty)\) as we have shown before already). Second, while the last boundary condition is a special case of that adopted by some authors (e.g. [24, 63] ), the fact that this is now a Dirichlet boundary condition rather than a Neumann boundary condition as used in (e.g. [24, 63]) has considerably facilitated the numerical solution procedure to be shown in the next section. In other words, the solution satisfying the differential system (2.2.16) must also satisfy those obtained with the last boundary condition replaced by the Neumann boundary condition (2.2.9), but not vice versa. Finally, one should notice that there exists a singularity at the corner where \(S \to \infty, v \to \infty\), in addition to the well-known one along the moving boundary \(S = S_f\). However, unlike the latter, the former does not cause any computational difficulties as it is a simple and removable singularity that commonly exists in diffusion problems with different boundary values prescribed on two adjacent boundaries.

### 2.3 Numerical method based on the ADI scheme

Upon establishing a closed differential system (2.2.16) for the price of American puts under the Heston model, we propose a new predictor-corrector approach based on the ADI method, in this section, to solve this system in two phases within a time step: a prediction phase in which a rough value of the optimal exercise price \(S_f\) is calculated, and a correction phase in which the option value \(U\) as well as the final value of \(S_f\) is determined.

#### 2.3.1 Coordinate transformation

In order to solve the PDE system (2.2.16) effectively, we introduce a new variable as the time to expiration: \(\tau = T - t\). It should be noted that the backward problem (2.2.16) has been changed to an initial value problem (2.3.1) after the introduction of \(\tau\).

If we want to solve PDE system (2.2.16) directly using FDM, some kinds of
iterative methods may be adopted because of the existence of the free boundary. To avoid iterations, we first adopt the Landau transform [72], i.e., $$x = \ln\left(\frac{S}{S_f}\right)$$ to shift the moving boundary conditions to fixed boundary conditions before applying our predictor-corrector method.

After some rather simple algebraic manipulations, the PDE system (2.2.16) can be written as: for $$x \in (0, \infty), \ v \in (0, \infty), \ \tau \in (0, T]$$

$$\begin{cases}
\mathbb{L}U = 0, \\
U(x, v, 0) = 0, \ \lim_{x \to -\infty} U(x, v, \tau) = 0, \\
U(0, v, \tau) = K - S_f(v, \tau), \ \frac{\partial U}{\partial x}(0, v, \tau) = -S_f(v, \tau), \\
\lim_{v \to 0} U(x, v, \tau) = 0, \ \lim_{v \to \infty} U(x, v, \tau) = K,
\end{cases}
$$

where

$$\mathbb{L} = \left[\frac{1}{2}v + \frac{1}{2} \frac{\sigma^2 v (\partial S_f)}{S_f} \right]^2 \frac{\partial^2}{\partial x^2} + \frac{1}{2} \frac{\sigma^2 v}{S_f} \frac{\partial^2}{\partial v^2} + \left(\rho \sigma v - \frac{\sigma^2 v (\partial S_f)}{S_f} \right) \frac{\partial^2}{\partial x \partial v} + \left[-\frac{1}{2}v + \frac{1}{2} \frac{\sigma^2 v (\partial S_f)}{S_f} \right]^2 \frac{\partial^2}{\partial v^2} - \frac{1}{2} \frac{\sigma^2 v (\partial^2 S_f)}{S_f} \frac{\partial^2}{\partial v^2} + r - \kappa (\eta - v) \frac{\partial}{\partial v} + \frac{1}{S_f} \frac{\partial S_f}{\partial \tau} \frac{\partial}{\partial x} + \kappa (\eta - v) \frac{\partial}{\partial v} - r - \frac{\partial}{\partial \tau}.
$$

In order to simplify the notation of $$\mathbb{L}$$, we introduce three new notations:

$$\xi = \frac{1}{S_f} \frac{\partial S_f}{\partial v}, \ \beta = \frac{1}{S_f} \frac{\partial^2 S_f}{\partial v^2}, \ \lambda = \frac{1}{S_f} \frac{\partial S_f}{\partial \tau},$$

so that $$\mathbb{L}$$ can be written as:

$$\mathbb{L} = a(v) \frac{\partial^2}{\partial x^2} + b(v) \frac{\partial^2}{\partial v^2} + c(v) \frac{\partial^2}{\partial x \partial v} + (d(v) + \lambda) \frac{\partial}{\partial x} + e(v) \frac{\partial}{\partial v} - r - \frac{\partial}{\partial \tau},$$

where

$$a(v) = \frac{1}{2}v + \frac{1}{2} \frac{\sigma^2 v}{S_f} \xi - \rho \sigma v \xi, \ \ b(v) = \frac{1}{2} \sigma^2 v, \ \ c(v) = \rho \sigma v - \sigma^2 v \xi, \ \ d(v) = -\frac{1}{2}v + \frac{1}{2} \xi^2 \sigma^2 v - \frac{1}{2} \sigma^2 v \beta + r - \kappa (\eta - v) \xi, \ \ e(v) = \kappa (\eta - v).$$
One should notice that after the Landau transform, the nonlinear nature of the problem is explicitly exposed in the transformed equation. It consists the optimal exercise price, which is also part of the solution. Before our new predictor-corrector approach can be applied, we should discretize the highly nonlinear PDE system (2.3.1), which will be demonstrated in detail in the next subsection.

2.3.2 Discretization of the PDE system

Now the option pricing problem is defined on an unbounded domain

\[ \{(x, v, \tau)|x \geq 0, v \geq 0, \tau \in [0, T]\}. \]

In order to use finite difference approximation for spatial variables, we need to truncate the semi-infinite domain into a finite domain:

\[ \{(x, v, \tau) \in [0, x_{\text{max}}] \times [0, v_{\text{max}}] \times [0, T]\}. \tag{2.3.2} \]

Theoretically, \( x_{\text{max}} \) and \( v_{\text{max}} \) should be sufficiently large to eliminate the boundary effect. However, based on Willmott et al.’s estimate [111] that the upper bound of the asset price \( S_{\text{max}} \) is typically three or four times the strike price, it is reasonable for us to set \( x_{\text{max}} = \ln 5 \). On the other hand, the volatility value is usually very small. The highest value of the volatility that has ever been recorded on CBOE (Chicago Board Options Exchange) is only 0.85 [34]. Thus, it is quite reasonable to set \( v_{\text{max}} = 1 \).

The discretization is performed by placing a set of uniformly distributed grids in the computation domain (2.3.2). With the number of steps in the \( x, v \) and \( \tau \) directions being denoted by \( N_x, N_v \) and \( N_\tau \), respectively, the step sizes are correspondingly

\[ \Delta x = \frac{X_{\text{max}}}{N_x}, \quad \Delta v = \frac{V_{\text{max}}}{N_v}, \quad \text{and} \quad \Delta \tau = \frac{T}{N_\tau}. \]

The value of the unknown function \( U \) at a grid point is thus denoted by

\[ U_{i,j}^n \approx U(x_i, v_j, \tau_n) = U(i \Delta x, j \Delta v, n \Delta \tau). \]
where \( i = 0, \cdots, N_x; \quad j = 0, \cdots, N_v; \quad n = 0, \cdots, N_\tau. \)

The discretization of the PDE system (2.3.1) needs to be conducted both in the interior domain

\[
\Omega = \left\{(i\Delta x, j\Delta v) \middle| i = 1 \cdots N_x - 1, j = 1 \cdots N_v - 1\right\},
\]

and along the boundary \( \partial \Omega = \partial_x \Omega \cup \partial_v \Omega \cup \partial_{xv} \Omega, \) in which

\[
\partial_x \Omega = \left\{(i\Delta x, j\Delta v) \middle| i = 0, N_x, j = 1 \cdots N_v - 1\right\},
\]

\[
\partial_v \Omega = \left\{(i\Delta x, j\Delta v) \middle| i = 1 \cdots N_x - 1, j = 0, N_v\right\},
\]

\[
\partial_{xv} \Omega = \left\{(i\Delta x, j\Delta v) \middle| i = 0, N_x, j = 0, N_v\right\}.
\]

For those grid points that belong to \( \Omega, \) we use the standard second-order central difference scheme to approximate the first-order and second-order non-cross spatial derivatives. The cross-derivative term, on the other hand, is discretized as:

\[
(\delta_{xv}U)_{i,j} = \frac{(\delta_x U)_{i,j+1} - (\delta_x U)_{i,j-1}}{2\Delta v},
\]

where

\[
(\delta_x U)_{i,j+1} = \frac{U_{i+1,j+1} - U_{i-1,j+1}}{2\Delta x}, \quad (\delta_x U)_{i,j-1} = \frac{U_{i+1,j-1} - U_{i-1,j-1}}{2\Delta x}.
\]

Next, we consider the discretization of the grid points that belong to \( \partial \Omega. \) The treatment of the Dirichlet boundary conditions is quite standard. However, it is more difficult to deal with the Neumann boundary condition. In the literature, there are usually two different approaches for such kind of boundary conditions with second-order accuracy [96]. The first one is to introduce a fictitious grid point \( U^n_{0,j}, \) and approximate the Neumann boundary condition at \( x = 0 \) with the central difference,

\[
\frac{\partial U^n_{0,j}}{\partial x} = \frac{U^n_{1,j} - U^n_{1,j}}{2\Delta x}.
\]
Then, together with the assumption that the governing equation in the PDE system (2.3.1) is also satisfied at the boundary point \( U_{n,0,j} \), we obtain two different algebraic equations, from which \( U_{n,1,j} \) can be eliminated. The details of the algebraic manipulations are given in Zhu and Zhang [125]. In essence, this method is to use the grid point \( U_{1,j} \) and some extra information from the PDE, to approximate the derivative \( \frac{\partial U_{0,j}}{\partial x} \).

The second approach is to use the so-called one-sided difference, which is, in essence, a form of extrapolation that determines the value of the unknown function on the boundary in terms of its values at the interior grid points [96]. From the Taylor series, we obtain the approximation:

\[
U^n_{1,j} \approx U^n_{0,j} + \Delta x \frac{\partial U^n_{0,j}}{\partial x} + \frac{1}{2}(\Delta x)^2 \frac{\partial^2 U^n_{0,j}}{\partial x^2} + \mathcal{O}((\Delta x)^3), \tag{2.3.3}
\]

\[
U^n_{2,j} \approx U^n_{0,j} + 2\Delta x \frac{\partial U^n_{0,j}}{\partial x} + 2(\Delta x)^2 \frac{\partial^2 U^n_{0,j}}{\partial x^2} + \mathcal{O}((\Delta x)^3). \tag{2.3.4}
\]

By eliminating the second-order derivatives in (2.3.3) and (2.3.4), we obtain

\[
\frac{\partial U^n_{0,j}}{\partial x} = \frac{4U^n_{1,j} - U^n_{2,j} - 3U^n_{0,j}}{2\Delta x} + \mathcal{O}((\Delta x)^3).
\]

In short, this method is to use the interior grid values \( U^n_{1,j}, U^n_{2,j} \) and the known boundary value \( U^n_{0,j} \) to approximate the derivative \( \frac{\partial U^n_{0,j}}{\partial x} \).

In our work, we adopt the second approach to approximate the derivative \( \frac{\partial U^n_{0,j}}{\partial x} \). The specific reason will be stated in the next section. For readers’ convenience, we summarize the finite difference equations written on a grid point \((i,j,n)\) as:
Chapter 2.

\[
\begin{aligned}
\frac{\partial U^n_{i,j}}{\partial \tau} &= a_j \delta_{xx} U^n_{i,j} + b_j \delta_{vv} U^n_{i,j} + (d_j + \lambda_j) \delta_x U^n_{i,j} + c_j \delta_{xx} U^n_{i,j} + e_j \delta_{vv} U^n_{i,j} - r U^n_{i,j}, \\
U^n_{i,j} &= 0, \quad U^n_{0,j} = K - S^n_f(j), \\
4U^n_{i,j} - U^n_{2,j} - 3U^n_{0,j} &= -S^n_f(j), \\
U^n_{i,0} &= 0, \quad U^n_{i,\nu} = K,
\end{aligned}
\]

(2.3.5)

where

\[
\begin{align*}
\xi^n_j &= \frac{1}{S^n_f(j)} \frac{S^n_f(j + 1) - S^n_f(j - 1)}{2\Delta v}, \\
\beta^n_j &= \frac{1}{S^n_f(j)} \frac{S^n_f(j + 1) - 2S^n_f(j) + S^n_f(j - 1)}{(\Delta v)^2}, \\
\lambda_j &= \frac{1}{S^n_f(j)} \frac{\partial S^n_f(j)}{\partial \tau}, \\
a_j &= \frac{1}{2} v_j + \frac{1}{2} \sigma^2 v_j (\xi^n_j)^2 - \rho \sigma v_j \xi_j, \\
b_j &= \frac{1}{2} \sigma^2 v_j, \\
c_j &= \rho \sigma v_j - \sigma^2 v_j \xi^n_j, \\
d_j &= r - \frac{1}{2} v_j + \frac{1}{2} (\xi^n_j)^2 \sigma^2 v_j - \frac{1}{2} \sigma^2 v_j \beta^n_j + \kappa [\eta - v_j] \xi^n_j, \\
e_j &= \kappa [\eta - v_j].
\end{align*}
\]

One should notice that in the PDE system (2.3.5), the time derivative remains un-discretized. The discretization of the time derivative is completed in the next subsection when we design a fast and efficient numerical scheme to solve this highly nonlinear system through a linearization.

2.3.3 Linearization of the nonlinear PDE system

It can be seen from the PDE system (2.3.5) that, if the optimal exercise price \( S^n_f \) were known at the beginning of the \((n+1)\)th time step, this system would become a linear one, and the option price at the \((n+1)\)th step could be worked out directly. Based on this point, we propose a new predictor-corrector approach to solve the highly nonlinear PDE system (2.3.5) in two phases within a time step.

The first phase is to work out an estimated value of the optimal exercise price at the \((n+1)\)th time step, denoted as \( \tilde{S}^{n+1}_f \). For simplicity, we omit the subscript in the \( v \) direction. Recall that there are two ways to approximate the first-order derivative
at $x = 0$, i.e.,

$$
\frac{\partial U_{n+1}^0}{\partial x} = \frac{U_{n+1}^1 - U_{n+1}^0}{2\Delta x} + O((\Delta x)^3) = -S_f^{n+1},
$$

(2.3.6)

$$
\frac{\partial U_{n+1}^i}{\partial x} = \frac{4U_{n+1}^i - U_{n+1}^{i+1} - 3U_{n+1}^{i-1}}{2\Delta x} + O((\Delta x)^3) = -S_f^{n+1}.
$$

(2.3.7)

If (2.3.6) is adopted to approximate $\frac{\partial U_{n+1}^0}{\partial x}$, one needs to assume that the governing equation contained in (2.3.1) is also satisfied at $x = 0$, which will lead to a complicated nonlinear differential equation of $S_f(v, \tau)$ with respect to $v$, due to the complexity of the Heston operator. However, if the second approximation formula (2.3.7) is adopted, we only need to deal with a much simpler process that will lead to an explicit exposure of $S_f(v, \tau)$, while maintaining the same order of accuracy.

From the Dirichlet boundary condition at $x = 0$, we know that

$$
U_{n+1}^0 = K - S_f^{n+1}.
$$

(2.3.8)

Substituting (2.3.8) into (2.3.7), we can obtain a relation among $U_{n+1}^1, U_{n+1}^2$ and $S_f^{n+1}$ at the $(n + 1)$th time step as:

$$
S_f^{n+1} = \frac{3K + U_{n+1}^2 - 4U_{n+1}^1}{3 + 2\Delta x}.
$$

(2.3.9)

Then, by applying the explicit Euler scheme to the time derivative $\frac{\partial U_{n+1}^i}{\partial \tau}$ and the implicit Euler scheme to the time derivative $\frac{\partial S_f^{n+1}(j)}{\partial \tau}$, respectively, in the governing equation contained in (2.3.5), we obtain

$$
U_{n+1}^i = U_{n}^i + \Delta \tau \left[ a_j \delta_{xx} U_{n}^i + b_j \delta_{xv} U_{n}^i + \left( d_j + \frac{1}{S_f^n(j)} \frac{S_f^{n+1}(j) - S_f^n(j)}{\Delta \tau} \right) \delta_x U_{n}^i + c_j \delta_{xx} U_{n}^i + e_j \delta_{xv} U_{n}^i \right] - r \Delta \tau U_{n}^i, \quad i = 1, 2.
$$

(2.3.10)

Consequently, we can obtain a predicted value of the optimal exercise price at the
new time step:

\[
\hat{S}_{f}^{n+1}(j) = \frac{3K + \mathcal{D}(U_{2,j}^{n} - 4U_{1,j}^{n})}{3 + 2\Delta x - \frac{\delta x (U_{2,j}^{n} - 4U_{1,j}^{n})}{S_{f}^{n}(j)}}.
\]

Here, operator \( \mathcal{D} \) is defined as:

\[
\mathcal{D} = I + \Delta \tau \left[ a_{j} \delta_{xx} + b_{j} \delta_{vv} + \left( d_{j} - \frac{1}{\Delta x} \right) \delta_{x} + c_{j} \delta_{xv} + e_{j} \delta_{v} \right] - r \Delta \tau I.
\]

It should be noted that with the calculated \( \hat{S}_{f}^{n+1} \), we can also obtain a predicted boundary value \( \hat{U}_{0}^{n+1} \) from (2.3.8), which completes the first phase of prediction.

The second phase is to calculate \( U_{n+1} \) at all grid points by using the estimated \( \hat{S}_{f}^{n+1} \) and \( \hat{U}_{0}^{n+1} \). Then, by using the newly-obtained \( U_{1}^{n+1} \) and \( U_{2}^{n+1} \), we can obtain the corrected values of \( S_{f}^{n+1} \) and \( U_{0}^{n+1} \). Unlike the first phase in which the explicit Euler scheme is used to construct the predicted values, the whole process of calculating \( U_{n+1} \) from the linearized PDE system (2.3.5) is based on the ADI scheme, which will be described in detail in the next subsection. Then, repeat this prediction-correction process until the expiration date is reached. The schematic flow chart of the scheme is provided in Fig 2.1 to summarize what has been described above.

### 2.3.4 The ADI scheme

The ADI scheme is a very powerful tool that is especially useful for solving parabolic equations on rectangular domains. It can be also applied to equations of other types or on more general domains [96]. Generally, the ADI scheme is a way of reducing a two-dimensional problem to a succession of many one-dimensional problems. The efficiency of the ADI method lies in the fact that those reduced one-dimensional problems usually possess a good structure, that is, their final matrix is tridiagonal and can thus be efficiently dealt with. On the other hand, the ADI method requires less storage space, because it solves the two directions alternatively by fixing the variable in one direction at a time step and thus it needs almost the same storage space as that required to solve a one-dimensional problem. Since our problem now
is a two dimensional one, excluding time, it is better to use the ADI method. In the following, we shall illustrate how this method can be applied to our case.

According to what we have discussed before, our problem at hand now is to solve the PDE system (2.3.5). The time derivative of $U$ needs to be discretized before the ADI method can be used. While the details of the derivation are provided in Appendix A.2, the finite difference equation, to which the ADI method is applied, is of the form:

$$(I - \theta A_1)(I - \theta A_2)U_{n+1} = [I + A_0 + (1 - \theta)A_1 + A_2]U_n - (I - \theta A_1)\theta A_2 U_n,$$  (2.3.11)

where the definitions of linear operators $A_0, A_1$ and $A_2$ are also left in Appendix A.2.

In terms of the ADI scheme, the simplest one is the DR (Douglas-Rachford) method [30], which is of the first order in time. Other methods, such as the CS (Craig-Sneyed) method [28], the HV (Hundsdorfer and Verwier) method [59], are much more
complicated, but are of more than first-order accuracy in the time direction. We chose
the DR method to calculate $U^{n+1}$ defined in (2.3.11). One of the most important
reasons is that in the predictor, we have used the explicit Euler scheme, which is a
first-order scheme, to construct the value of $S_j^{n+1}$. Theoretically speaking, an increase
of the order of accuracy in the correction phase beyond that in the prediction phase
is futile, as far as the overall order of accuracy of the entire procedure is concerned.
Moreover, it is far more efficient to use the DR method, because there are only two
steps in the DR method, while others have at least four.

The DR method involves two steps, in which the original operator in (2.3.11)
is split into two that are applied in two spatial directions respectively. First, we
compute an intermediate value, $Y$, from

$$(I - \theta A_1)Y = [I + A_0 + (1 - \theta)A_1 + A_2]U^n,$$  \hspace{1cm} (2.3.12)$$

with the nodal values in the $v$ direction fixed. The corresponding matrix form for
calculating $Y$ can be simply written as

$$AY_j = P_j + \text{bnd}x_j,$$

with the details of $A$, $Y_j$, $P_j$ and $\text{bnd}x_j$ being defined in Appendix A.3. Note that
the matrix $A$ is tridiagonal, so the Thomas algorithm [96] can be used to accelerate
the computational speed. In short, the way to implement the first step is to use a
loop on $j$, and within each loop the Thomas algorithm is used to solve for the values
of $Y_{i,j}$, for $i = 0 \cdots N_x$.

Having computed $Y$, the second stage of computation is to compute $U^{n+1}$ from

$$(I - \theta A_2)U^{n+1} = Y - \theta A_2 U^n,$$  \hspace{1cm} (2.3.13)$$
by fixing the variable in the $x$ direction. The corresponding matrix form is:

$$BU^{n+1}_i = Q_i + \text{bnd}v_i,$$

where $B$, $U^{n+1}_i$, $Q$ and $\text{bnd}v_i$ are also defined in Appendix A.3. Similar to (2.3.12), this is a system of $N_x - 1$ tridiagonal systems of equations, one tridiagonal system for each value of $i$. Again, we use a loop on $i$, within which the Thomas algorithm is used to solve for the values of $U^{n+1}_{i,j}$, for $j = 0 \cdots N_v$.

One can easily show that solving Eqs. (2.3.12) and (2.3.13) in an alternative way is equivalent to solving the original Eq. (2.3.11).

Before presenting the results of the numerical implementation of this two-stage solution procedure in the next section, it should be remarked that boundary values of the intermediate variable $Y$ need to be produced first before the interior $Y$ values can be computed from (2.3.12). The calculation of $Y$ values on the two boundaries needs to be treated differently. The calculation on the right boundary is straightforward, since the $U$ values there are always equal to zero. That is, we can simply set

$$Y_{N_x} = 0. \quad (2.3.14)$$

On the other hand, the calculation of $Y$ values on the left boundary $x = 0$ is a bit more complicated, as $U$ is time dependent there. (2.3.13) is utilized because the estimated $U^{n+1}$ values on the boundary have already been obtained in the prediction phase, although its interior values have not yet been calculated. That is, we can calculate $Y_0$ from

$$Y_0 = (I - \theta A_2)\tilde{U}_0^{n+1} + \theta A_2U_0^n,$$

$$= (I - \theta A_2)(K - \tilde{S}_{f}^{n+1}) + \theta A_2(K - S_f^n),$$

where $\tilde{S}_{f}^{n+1}$ is the predicted optimal exercise price obtained at the beginning of the
(n+1)th time step. This completes the ADI scheme and thus the entire numerical procedure, that we have used to calculate the results presented in the next section.

2.4 Examples and discussions

Even if the ADI method used for the corrector is unconditionally stable, our predictor-corrector finite difference scheme is only conditionally stable since the explicit Euler scheme for the predictor is conditionally stable. The stability condition for the current approach remained undiscussed in the literature even when it was applied to the Black-Scholes case [125]. The complexity of the Heston model has no doubt added the degree of difficulty in verifying the conditional stability of the current approach. However, based on the local von Neumann stability analysis [96], we were able to obtain theoretically the stability requirement for the current approach applied to the Heston model.

In this section, we shall present some numerical results as well as some useful discussions, which reveal the essence of the current scheme. The section is organized into three subsections, according to three important issues that should be addressed.

In the first subsection, we calculate the European puts to test the reliability of the chosen ADI method. In the second subsection, we discuss the convergence of the current approach. The stability condition is obtained by using the local von Neumann stability analysis. In the last subsection, we study the accuracy and efficiency of the current scheme.

2.4.1 Valuing European put options using the ADI scheme

Since the corrector plays a crucial role in the current approach, it is important to test first whether the chosen DR method can be used to construct a good corrector. An efficient way to illustrate the reliability of the DR method is to calculate the value of a European put, and compare it with the existing Heston formula [51].

The option’s parameters are set as follows: reversion rate $\kappa = 5$, reversion level $\eta = 0.16$, volatility of volatility $\sigma = 0.9$, risk-free interest rate $r = 0.1$, correlation
factor $\rho = 0.1$, time to expiration $T - t = 1$ (year), strike price $K = $10.0. The computation domain is truncated as:

$$\left[0, S_{\text{max}}\right] \times \left[0, v_{\text{max}}\right] \times \left[0, T\right] = \left[0, 200\right] \times \left[0, 5\right] \times \left[0, 1\right].$$

In this numerical test, we apply the Crank-Nicolson scheme, i.e., $\theta = \frac{1}{2}$, to the time derivative. The grid numbers in the $x$ direction and $v$ direction are chosen to be 400 and 150, respectively. Furthermore, a relatively large grid size in the time direction is used, i.e., $\Delta \tau = \frac{1}{20}$. Fig 2.2 shows the comparison of the option values calculated by the DR method and those obtained by using Heston’s analytical formula. The option prices presented in Fig 2.2 are plotted against the underlying asset with different fixed variance values. The excellent agreement of the two results confirms that the DR method is accurate.

Having gained confidence on the chosen DR method as a good corrector, we then apply the predictor-corrector scheme to American put options under the Heston model, which will be demonstrated in the following subsection.

2.4.2 Discussion on convergence

In this subsection, we study the convergence of the proposed numerical scheme. As has been pointed out previously, our problem in hand is originally a nonlinear one before the linearization process is invoked. Therefore, the convergence of the linearized system and the overall convergence of the numerical solution to that of the original nonlinear PDE should both be discussed.

Stability Analysis

For the convergence of the linearized system, the Lax Equivalence Theorem [96] can be applied, which states that the convergence of a numerical scheme is equivalent to its consistency and stability. The proof of the consistency is trivial and is thus omitted here. The proof of stability, however, is not so trivial because our method is a hybrid finite difference scheme, and it is applied to a problem with variable
Figure 2.2: European put option prices with different variance values.

(a) European put option value at \( v = 0.2 \)

(b) European put option value at \( v = 0.4 \)

(c) European put option value at \( v = 0.6 \)

(d) European put option value at \( v = 1 \)
coefficients. Intuitively, even if the ADI scheme for the corrector is unconditionally stable (see Theorem 2), the current approach should be only conditionally stable since the explicit Euler scheme is used in the predictor. However, it can be anticipated that the unconditional stability in the correction phase would be weakened to some extent. This is because the main ingredient of the predictor-corrector method is the corrector, which serves as a feedback mechanism. The feedback can damp the instability that has been introduced by the predictor, and thus relax the stability requirement of the explicit Euler scheme used in the predictor. It is extremely difficult to calculate exactly “how much” the corrector can influence the predictor, and in the literature, no one has discussed this issue in detail. Based on the local von Neumann stability analysis, we have managed to show the conditional stability for the predictor-corrector scheme. In the following work, the stability requirements of the explicit Euler scheme (2.3.10) and the ADI scheme Eqs. (2.3.12)-(2.3.13) will be first discussed. Finally, we obtain the conditional stability of the predictor-corrector approach.

The von Neumann stability analysis is usually restricted to problems with constant coefficients. However, the stability conditions obtained for constant coefficient schemes can be used to give stability conditions for the same scheme applied to equations with variable coefficients, mainly due to the fact that, in essence, instability is a local phenomenon with the high frequency nodes being the most unstable ones that would result in the collapse of a numerical scheme [96]. The general procedure is to consider the so-called frozen coefficient problem, which is the corresponding problem with constant coefficients obtained by fixing the coefficients at their values attained at each grid point in the computational domain. If each frozen coefficient problem is stable, then the variable coefficient problem is also stable [96]. For simplicity, in the following analysis, we do not consider the overall effect of the boundary conditions between subdomains.

When using the frozen coefficient technique, for a typical node \((i, j, n)\), we “freeze” all the coefficients in the governing equation contained in (2.3.5) across the whole
computational domain, as if they were all constants. In short, we first consider the stability condition of the DR method applied to a two-dimensional convection-diffusion equation with constant coefficients, i.e.,

\[
\frac{\partial U}{\partial \tau} = a \frac{\partial^2 U}{\partial x^2} + b \frac{\partial^2 U}{\partial v^2} + c \frac{\partial^2 U}{\partial x \partial v} + d \frac{\partial U}{\partial x} + e \frac{\partial U}{\partial v} - rU, \tag{2.4.1}
\]

where \(a, b\) and \(r\) are of positively values. The standard procedure of von Neumann stability analysis is to express \(U_{nk,m}\) in Eqs. (2.3.12-2.3.13) by \(g_ne^{ik\varphi}e^{im\phi}\), and \(Y_{k,m}\) by \(\tilde{g}g_ne^{ik\varphi}e^{im\phi}\) [96], where \(g\) and \(\tilde{g}\) are the amplification factors of (2.3.13) and (2.3.12), respectively, with \(\varphi, \phi \in [-\pi, \pi]\). As a result, Eqs. (2.3.12)-(2.3.13) are transformed to:

\[
\begin{align*}
\tilde{g}(1 - \theta z_1) &= 1 + z_0 + (1 - \theta)z_1 + z_2; \\
g(1 - \theta z_2) &= \tilde{g} - \theta z_2,
\end{align*}
\]

which result in the amplification factor:

\[
g = 1 + \frac{z_0 + z_1 + z_2}{(1 - \theta z_1)(1 - \theta z_2)},
\]

where

\[
\begin{align*}
z_1 &= -\frac{4a\Delta \tau}{\Delta x^2} \sin^2 \frac{\varphi}{2} - \frac{r\Delta \tau}{2} + i\frac{d\Delta \tau}{\Delta x} \sin \varphi, \\
z_2 &= -\frac{4b\Delta \tau}{\Delta v^2} \sin^2 \frac{\phi}{2} - \frac{r\Delta \tau}{2} + i\frac{e\Delta \tau}{\Delta v} \sin \phi, \\
z_0 &= -\frac{c\Delta \tau}{\Delta x \Delta v} \sin \varphi \sin \phi,
\end{align*}
\]

after some algebraic manipulations.

**Proposition 2** (i) If the coefficients \(a, b, c, d, e, r\) are chosen such that

\[
|z_0| \leq 2\sqrt{\Re(z_1)\Re(z_2)}, \tag{2.4.2}
\]


then for $\theta \geq \frac{1}{2}$, the DR method (Eqs. (2.3.12)-(2.3.13)) is unconditionally stable, i.e., $|g| \leq 1$. (ii) Assuming that the coefficients satisfy $c^2 \leq 4ab$, the fully explicit scheme (Eqs. (2.3.12)-(2.3.13) with $\theta = 0$), is stable if and only if

$$\Delta \tau \leq \frac{1}{2a \Delta x^2 + 2b \Delta x^2}. \quad (2.4.3)$$

Proof. (i) First, we define two vectors as $v_j = \begin{pmatrix} \sqrt{2\Re(z_j)} \\ |1 + \theta z_j| \sqrt{2\theta} \end{pmatrix}$, where $j=1, 2$. It is trivial to show $\|v_j\|_2 = \frac{|1 - \theta z_j|}{\sqrt{2\theta}}$. Thus, we have

$$\frac{|1 - \theta z_1|}{\sqrt{2\theta}} \frac{|1 - \theta z_2|}{\sqrt{2\theta}} = \|v_1\|_2 \|v_2\|_2,$$

$$\geq v_1 \cdot v_2,$$

$$= 2\sqrt{\Re(z_1)\Re(z_2)} + \frac{|1 + \theta z_1| |1 + \theta z_2|}{\sqrt{2\theta}},$$

$$\geq |z_0| + \frac{(1 - \theta z_1)(1 - \theta z_2)}{\sqrt{2\theta}} + z_1 + z_2. \quad (2.4.4)$$

Without loss of generality, we assume that $|1 - \theta z_1| \neq |1 - \theta z_2|$. Dividing both sides of (2.4.4) with $|1 - \theta z_1|$, we have

$$\frac{1}{2\theta} \geq \frac{z_0}{(1 - \theta z_1)(1 - \theta z_2)} + \frac{1}{2\theta} + \frac{z_1 + z_2}{(1 - \theta z_1)(1 - \theta z_2)},$$

$$\geq \frac{1}{2\theta} + \frac{z_0 + z_1 + z_2}{(1 - \theta z_1)(1 - \theta z_2)}. \quad (2.4.5)$$

On the other hand, when $\theta \geq \frac{1}{2}$, we have

$$g = |1 + \frac{z_0 + z_1 + z_2}{(1 - \theta z_1)(1 - \theta z_2)}|,$$

$$= |1 - \frac{1}{2\theta} + \frac{1}{2\theta} + \frac{z_0 + z_1 + z_2}{(1 - \theta z_1)(1 - \theta z_2)}|,$$

$$\leq 1 - \frac{1}{2\theta} + \frac{1}{2\theta} + \frac{z_0 + z_1 + z_2}{(1 - \theta z_1)(1 - \theta z_2)},$$
which, combined with (2.4.5), yields $|g| \leq 1$. Therefore, if

$$|z_0| \leq 2\sqrt{R(z_1)R(z_2)},$$

the ADI scheme (for $\theta \geq \frac{1}{2}$) is unconditionally stable.

(ii) When $\theta = 0$, the DR method turns out to be a fully explicit scheme. In this case, it is efficient to first assume that $\frac{\Delta \tau}{(\Delta S)^2} = \mu_1$, $\frac{\Delta \tau}{(\Delta v)^2} = \mu_2$, where $\mu_1$ and $\mu_2$ are fixed constants. Thus, we have

$$\lambda_1 = \frac{\Delta \tau}{\Delta x} = \sqrt{\Delta \tau \mu_1} \sim O(\sqrt{\Delta \tau}),$$

$$\lambda_2 = \frac{\Delta \tau}{\Delta v} = \sqrt{\Delta \tau \mu_2} \sim O(\sqrt{\Delta \tau}),$$

which indicate that, in comparison with those $O(1)$ terms, the lower-order derivative terms only give an $O(\Delta \tau)$ contribution to $|g|^2$, and thus do not affect the stability. In fact, ignoring lower-order derivative terms when determining stability is a commonly used technique in the stability analysis [96].

Now, we need to prove that, with $\Delta \tau \leq \frac{1}{\frac{2a}{\Delta x^2} + \frac{2b}{\Delta v^2}}$

$$\left(1 - \frac{c \Delta \tau}{\Delta x \Delta v} \sin \varphi \sin \phi - \frac{4a \Delta \tau}{\Delta x^2} \sin^2 \frac{\varphi}{2} - \frac{4b \Delta \tau}{\Delta v^2} \sin^2 \frac{\phi}{2}\right) \leq 1, \quad (\varphi, \phi) \in [-\pi, \pi] \times [-\pi, \pi],$$

which is equivalent to showing that

$$0 \leq \frac{c \Delta \tau}{\Delta x \Delta v} \sin \varphi \sin \phi + \frac{4a \Delta \tau}{\Delta x^2} \sin^2 \frac{\varphi}{2} + \frac{4b \Delta \tau}{\Delta v^2} \sin^2 \frac{\phi}{2} \leq 2, \quad (\varphi, \phi) \in [-\pi, \pi] \times [-\pi, \pi],$$

and vice versa. This is achieved by using the following procedure.

Define a function

$$P(\varphi, \phi) = c \sin \varphi \sin \phi + 2 \tilde{a} \sin^2 \frac{\varphi}{2} + 2 \tilde{b} \sin^2 \frac{\phi}{2},$$
where

\[ \tilde{c} = c\sqrt{\mu_1\mu_2}, \quad \tilde{a} = 2a\mu_1, \quad \tilde{b} = 2b\mu_2. \]

Since \( P(\varphi, \phi) \) is continuous over the domain \([-\pi, \pi] \times [-\pi, \pi]\), the maximum and minimum values can be both achieved in this domain. As a result, (2.4.6) is satisfied if and only if

\[ 0 \leq P_{\min} \leq P_{\max} \leq 2. \quad \text{(2.4.7)} \]

Suppose \((\varphi, \phi)\) is an interior point of the domain of the function \( P \), and it is also a local maximum or minimum point, we have

\[ P_\varphi(\varphi, \phi) = P_\phi(\varphi, \phi) = 0, \quad \text{i.e.,} \]

\[ \begin{align*}
\tilde{c} \cos \varphi \sin \phi + \tilde{a} \sin \varphi &= 0, \\
\tilde{c} \cos \phi \sin \varphi + \tilde{b} \sin \phi &= 0.
\end{align*} \quad \text{(2.4.8)} \]

It is straightforward to show that (2.4.8) is equivalent to

\[ \begin{align*}
(\tilde{c}^2 + \tilde{a}^2)^2 \sin^2 \varphi &= \frac{\tilde{c}^4 - \tilde{a}^2\tilde{b}^2}{\tilde{c}^2}, \\
(\tilde{c}^2 + \tilde{b}^2)^2 \sin^2 \phi &= \frac{\tilde{c}^4 - \tilde{a}^2\tilde{b}^2}{\tilde{c}^2}.
\end{align*} \quad \text{(2.4.9)} \]

Clearly, the origin \((\varphi, \phi) = (0, 0)\) is the only solution of (2.4.9) because of the constraint \( c^2 \leq 4ab \), which is equivalent to \( \tilde{c}^2 \leq \tilde{a}\tilde{b} \). Thus, in \((-\pi, \pi) \times (-\pi, \pi)\), the only possible local maximum or minimum point of \( P \) is \((\varphi, \phi) = (0, 0)\). Comparing with the function values at \((0, 0)\), and along the boundary \( \varphi = \pm \pi \), or \( \phi = \pm \pi \), it is clear that

\[ P_{\max} = 4a\mu_1 + 4b\mu_2, \quad P_{\min} = 0. \]

Consequently, if \( \Delta \tau \leq \frac{1}{2a\Delta^2 + \frac{2b}{\Delta \theta}} \), then (2.4.7) is satisfied, and vice versa. Therefore, under the assumption that the coefficients of (2.4.1) satisfy \( c^2 \leq 4ab \), the sufficient and necessary condition for the stability of the explicit scheme (Eqs. (2.3.12)-(2.3.13) with \( \theta = 0 \)) is \( \Delta \tau \leq \frac{1}{2a\Delta^2 + \frac{2b}{\Delta \theta}} \). This completes the proof.

By means of Proposition 2, the following theorems can be proved.
Theorem 1 The DR method that is applied to the valuation of European puts is unconditionally stable \( (\theta \geq \frac{1}{2}) \).

Proof. In the case of European puts, \( z_0, \Re(z_1), \) and \( \Re(z_2) \) are defined as follows:

\[
\Re(z_1) = -\frac{2vx^2\Delta\tau}{\Delta x^2} \sin^2 \frac{1}{2}\varphi - \frac{r\Delta\tau}{2},
\]

\[
\Re(z_2) = -\frac{2\sigma^2 v\Delta\tau}{\Delta v^2} \sin^2 \frac{1}{2}\phi - \frac{r\Delta\tau}{2},
\]

\[
|z_0| = \left| \frac{\rho\sigma vx\Delta\tau}{\Delta x\Delta v} \sin \varphi \sin \phi \right|.
\]

where \( x \) and \( v \) are defined over the whole computational domain. We have,

\[
4\Re(z_1)\Re(z_2) - |z_0|^2 = 4\left(\frac{2vx^2\Delta\tau}{\Delta x^2} \sin^2 \frac{1}{2}\varphi + \frac{r\Delta\tau}{2}\right)\left(\frac{2\sigma^2 v\Delta\tau}{\Delta v^2} \sin^2 \frac{1}{2}\phi + \frac{r\Delta\tau}{2}\right)
\]

\[-\left(\frac{\rho\sigma vx\Delta\tau}{\Delta x\Delta v} \sin \varphi \sin \phi\right)^2,
\]

\[
\geq \left(\frac{4v\sigma x\Delta\tau}{\Delta x\Delta v} \sin \frac{1}{2}\varphi \sin \frac{1}{2}\phi\right)^2 (1 - \rho^2 \cos^2\frac{1}{2}\varphi \cos^2\frac{1}{2}\phi),
\]

\[
\geq 0,
\]

which means that for any fixed point \((x, v)\) across the whole computational domain, (2.4.2) is always satisfied. According to the first part of the proposition and the frozen coefficient technique, the ADI scheme is unconditionally stable. This completes the proof.

Theorem 2 Assuming that the optimal exercise price is known in advance, and all the variable coefficients are both bounded and sufficiently smooth, the DR method that is applied to the corrector (Eqs. (2.3.12)-(2.3.13)) is unconditionally stable \( (\theta \geq \frac{1}{2}) \).

Proof. In this case, \( z_0, \Re(z_1), \) and \( \Re(z_2) \) are defined as follows:

\[
\Re(z_1) = -\frac{(2v + 2\sigma^2 v\xi^2 - 4\rho\sigma v\xi)\Delta\tau}{\Delta x^2} \sin^2 \frac{1}{2}\varphi - \frac{r\Delta\tau}{2},
\]

\[
\Re(z_2) = -\frac{2\sigma^2 v\Delta\tau}{\Delta v^2} \sin^2 \frac{1}{2}\phi - \frac{r\Delta\tau}{2},
\]

\[
|z_0| = \left| \frac{(\rho\sigma v - \sigma^2 v\xi)\Delta\tau}{\Delta x\Delta v} \sin \varphi \sin \phi \right|.
\]
We have,

\[
4\Re(z_1)\Re(z_2) - |z_0|^2 = 4\left(\frac{2v + 2\sigma^2 v\xi^2 - 4\rho\sigma v\xi}{\Delta x^2} \sin^2 \frac{1}{2} \varphi + \frac{r\Delta \tau}{2}\left(\frac{2\sigma^2 v\Delta \tau}{\Delta v^2} \sin^2 \frac{1}{2} \phi \right)
\right.
\]

\[
+ \frac{r\Delta \tau}{2}\right) - \left(\frac{\rho\sigma v - \sigma^2 v\xi}{\Delta x}\Delta \tau \right) \sin \varphi \sin \phi)^2,
\]

\[
\geq \frac{(4\nu \sigma \Delta \tau}{\Delta x} \sin \frac{1}{2} \varphi \sin \frac{1}{2} \phi)^2 \left[(1 + \sigma^2 \xi^2 - 2\rho \sigma \xi) - (\rho - \sigma \xi)^2 \right. 
\]

\[
\left. \cos^2 \frac{1}{2} \varphi \cos^2 \frac{1}{2} \phi)\right],
\]

\[
\geq \frac{(4\nu \sigma \Delta \tau}{\Delta x} \sin \frac{1}{2} \varphi \sin \frac{1}{2} \phi)^2 \left[(1 + \sigma^2 \xi^2 - 2\rho \sigma \xi) - (\rho - \sigma \xi)^2 \right],
\]

\[
= \frac{(4\nu \sigma \Delta \tau}{\Delta x} \sin \frac{1}{2} \varphi \sin \frac{1}{2} \phi)^2 (1 - \rho^2),
\]

\[
\geq 0,
\]

which means that for any fixed point \((x, v)\) across the whole computational domain, (2.4.2) is always satisfied. According to the first part of the proposition and the frozen coefficient technique, the ADI scheme that is applied to the corrector is unconditionally stable. This completes the proof.

**Theorem 3** Assuming that the optimal exercise price is known in advance, and all the variable coefficients are both bounded and sufficiently smooth, the fully explicit scheme that is applied to the predictor (2.3.10) \((\theta = 0)\) is stable if and only if \(\Delta \tau \leq \frac{2}{a_1 + a_2}\), where

\[
a_1 = \frac{2v_{\max} - 4\rho \sigma v_{\max} \xi_{\min} + 2\sigma^2 v_{\max} \xi^2_{\min}}{(\Delta x)^2}, \quad a_2 = \frac{2\sigma^2 v_{\max}}{(\Delta v)^2}.
\]

(Here, we further assume that the correlation factor \(\rho \geq 0\)).

**Proof.** According to the second part of Proposition 2, and by using the frozen coefficient technique, we only need to show that, with

\[
\Delta \tau \leq \frac{2}{a_1 + a_2},
\]

the constraints \(c \leq 4ab\), and \(\Delta \tau \leq \frac{1}{\Delta x^2 + \Delta v^2}\) are both satisfied at each point across
the whole computational domain. In this case, $a$, $b$, $c$ are defined as:

\[
\begin{align*}
    a &= \left( \frac{1}{2} v + \frac{1}{2} \sigma^2 v \xi^2 - \rho \sigma v \xi \right), \\
    b &= \frac{1}{2} \sigma^2 v, \\
    c &= \rho \sigma v - \sigma^2 v \xi.
\end{align*}
\]

Since

\[
4ab - c^2 = (v + \sigma^2 v \xi^2 - 2 \rho \sigma v \xi) \sigma^2 v - (\rho \sigma v - \sigma^2 v \xi)^2, \\
= (v \sigma)^2 (1 - \rho^2), \\
\geq 0,
\]

all the frozen coefficient problems are stable if and only if

\[
\Delta \tau \leq \frac{1}{2a_{\text{max}}} + \frac{2b_{\text{max}}}{\Delta x^2}.
\]

Note that the optimal exercise price $S_f(v, \tau)$ is a monotonic decreasing function of $v$. On the other hand, it is assumed that all the coefficients are both bounded. Therefore, there exists a $\xi_{\text{min}}$, such that $\xi_{\text{min}} \leq \xi_m \leq 0$. As a result, $a_{\text{max}} = a_1$, $b_{\text{max}} = a_2$, this completes the proof.

Based on Theorem 3, one can easily find out that the stability condition of the fully explicit scheme is closely related to the value of $\sigma$. More steps in the time direction are needed when $\sigma$ is large.

Through previous analyses, we know that with the optimal exercise price known in advance, the corrector is unconditionally stable. Thus, if the predictor constructed by the explicit Euler scheme is stable, the whole process is stable. In other words, the amplification factor of the predictor is also the one for the entire predictor-corrector scheme. The main difference between our predictor and the single explicit Euler scheme lies in the fact that we use the corrected value, which is produced by the ADI method, to continue the prediction of the next time step. As a result, the good stable property of the ADI scheme has somehow relaxed the overall amplification
factor. It would be ideal to show the above statement numerically. This is achieved after we prove Theorem 4 as follows.

**Theorem 4** If $g_1$ is the amplification factor of the explicit Euler scheme used in the predictor, and $g_2$ is the one of the ADI scheme used in the corrector, then the predictor-corrector method is stable if and only if $|g_1 g_2| \leq 1 + M \Delta \tau$.

**Proof.**

We use tildes to denote the value that is obtained after the predictor. In the following proof, we have omitted the subscript in the $v$ direction or $x$ direction for simplicity.

Since $\tilde{S}_{j}^{n+1} = \frac{3K + \tilde{U}_2^{n+1} - 4\tilde{U}_1^{n+1}}{3 + 2\Delta x}$, we can easily deduce:

$$|\tilde{S}_{j}^{n+1} - \tilde{S}_{j}^0| \leq \frac{|\tilde{U}_2^{n+1} - \tilde{U}_2^0| + 4|\tilde{U}_1^{n+1} - \tilde{U}_1^0|}{3 + 2\Delta x}.$$ 

Thus, the predictor is stable if and only if the process of computing $\tilde{U}^{n+1}$ is stable.

From (2.3.10), we know that

$$\tilde{U}^{n+1} = (I + A_0^* + A_1^* + A_2^*)U^n$$

(2.4.10)

where star denotes the parameter $\theta = 0$. Here, $U^n$ is solved from

$$(I - \theta A_1)Y = [I + A_0 + (1 - \theta)A_1 + A_2]\tilde{U}^{n-1},$$

(2.4.11)

$$(I - \theta A_2)U^n = Y - \theta A_2\tilde{U}^{n-1}.$$ 

(2.4.12)

where $\tilde{U}^{n-1}$ is obtained after the predictor of the $n$th time step. Again, we take all the variable coefficients in these two stages to be “frozen” at constant values, so that the von Neumann analysis can be applied. By using Fourier transform on (2.4.10)
and Eqs. (2.4.11)-(2.4.12), we obtain:

\[
\tilde{U}_{n+1}^{k,m} = (1 + z^*_0 + z^*_1 + z^*_2)\check{U}_{n}^{k,m} = g_1\hat{U}_{n}^{k,m},
\]

\[
\hat{U}_{n}^{k,m} = 1 + \frac{z_0 + z_1 + z_2}{(1 - \theta z_1)(1 - \theta z_2)}\check{U}_{n-1}^{k,m} = g_2\tilde{U}_{n-1}^{k,m}.
\]

where \(\tilde{U}_{n+1}^{k,m}\) and \(\hat{U}_{n}^{k,m}\) are the Fourier transform of \(\check{U}_{n}^{k,m}\) and \(U_{n}^{k,m}\), respectively. Thus,

\[
\tilde{U}_{n+1}^{k,m} = (g_1 g_2)\tilde{U}_{n-1}^{k,m},
\]

which shows that the overall amplification factor \(g\) for the predictor-corrector approach is nothing but \(g = g_1 g_2\).

Therefore, our method is stable if and only if

\[
|g_1(x_i, v_j, \tau_n; \Delta \tau)g_2(x_i, v_j, \tau_{n-1}; \Delta \tau)| \leq 1 + M\Delta \tau,
\]

(2.4.13)

where \(M\) is a constant that is independent of \(\varphi, \phi\), as well as all the step sizes, and \(i, j, n\) are defined over the whole computational domain. This completes the proof.

A further analysis of inequality (2.4.13) shows:

\[
|g_1(x_i, v_j, \tau_n; \Delta \tau)| \leq \frac{1 + M\Delta \tau}{|g_2(x_i, v_j, \tau_{n-1}; \Delta \tau)|}.
\]

(2.4.14)

Since the corrector is unconditionally stable, i.e., \(|g_2| \leq 1\), it is straightforward that \(\Delta \tau_{\text{explicit}} \leq \Delta \tau_{\text{predictor-corrector}}\). Therefore, comparing with the explicit Euler scheme, the stability requirement of our current scheme is less restrictive. This allows us to choose a larger time step than the fully explicit Euler scheme, and thus considerably enhance the computational efficiency.

This theorem leads to the conclusion that \(\Delta \tau_{\text{explicit}} \leq \Delta \tau_{\text{predictor-corrector}}\), because the corrector is unconditionally stable, or \(g_1 < 1\).

Ideally, the above theoretical statement has been verified by the following numerical experiment (see Table 2.1). It shows the comparison of the smallest number of
time steps that each method requires to calculate a convergent solution. One should notice that in Table 2.1, the “Predictor-corrector” refers to the current method with parameter \( \theta = \frac{1}{2} \), while the “Explicit” refers to the fully explicit scheme, which is a special case of our scheme. It should be also noted that the third number in the triplets \((N_x, N_v, N_r)\) is the smallest number that each method needs to produce a convergent solution. It can be clearly seen that our current method requires a lower number of time steps than the fully explicit one. Moreover, our method is far more stable when \( \sigma \) is small. It confirms that the stability requirement of the predictor-corrector method has been influenced by the corrector. In other words, a better corrector can improve the computational efficiency of the whole scheme significantly.

Table 2.1: **Comparison of the stability requirement. Model parameters are** \( \kappa = 2.5, \eta = 0.16, r = 0.1, \rho = 0.1, T - t = 0.25 \text{(year)}, K = $10.0 \).

<table>
<thead>
<tr>
<th>( \sigma )</th>
<th>Predictor-Corrector</th>
<th>Explicit</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.01</td>
<td>(25,32,13)</td>
<td>(25,32,250)</td>
</tr>
<tr>
<td></td>
<td>(50,64,130)</td>
<td>(50,64,1000)</td>
</tr>
<tr>
<td>0.1</td>
<td>(25,32,39)</td>
<td>(25,32,250)</td>
</tr>
<tr>
<td></td>
<td>(50,64,140)</td>
<td>(50,64,1000)</td>
</tr>
<tr>
<td>0.45</td>
<td>(25,32,100)</td>
<td>(25,32,400)</td>
</tr>
<tr>
<td></td>
<td>(50,64,1100)</td>
<td>(50,64,1600)</td>
</tr>
</tbody>
</table>

Having demonstrated the convergence of the linearized system, we need to show that the numerical solution does converge to that of the original nonlinear PDE. Since there is no analytical solution for American puts under the Heston model, the only reasonable approach is to compare our numerical solution with other published results. This would be demonstrated by the following numerical examples.

**Computed option prices and optimal exercise prices**

We calculate two sets of American put options with different parameters using the current method based on the Crank-Nicolson scheme, i.e., \( \theta = \frac{1}{2} \). These prices are presented in Table 2.2 and Table 2.3 for the asset values \( S = 8, 9, 10, 11, 12 \), and for variance values \( v = 0.0625, 0.25 \). We have used different discretization grids in order to study the accuracy of the numerical solutions. The prices reported in [24, 63, 75, 78, 80, 127] are also shown in these tables for comparison. It can be seen that even with the most coarse grid, the error is only about \( 10^{-2} \). Furthermore, the prices
obtained with the finest grid are fairly close to the ones in [24, 63, 75, 78, 80, 127], and the error is about $10^{-4}$. This confirms that our numerical solution does converge to that of the original nonlinear PDE system.

Depicted in Fig 2.3-a and Fig 2.3-b are the option price $U(S, v, \tau)$ as a function of $S$ with different parameters. Clearly, the option price is a decreasing function of the asset values. Moreover, the “smooth pasty” conditions across the free boundary, which are usually difficult to implement numerically, are also satisfied well. In Fig 2.4, we show the option price $U(S, v, \tau)$ as a function of $S$ with fixed variance $v = 0.25$ at three instants: $\tau = T - t = 0.5$ (year), $\tau = T - t = 0.25$ (year) and $\tau = T - t = 0.1$ (year). Clearly, as it gets closer to the expiration of the option, i.e., $\tau = 0$, the option price becomes closer to the payoff function $\max(K - S, 0)$. Moreover, the optimal exercise price $S_f(v, \tau)$ as a function of time to expiration with different fixed variance values is shown in Fig 2. 3-c and Fig 2. 3-d. As expected, the optimal exercise price is a monotonically decreasing function with both $\tau$ and $v$.

Table 2.2: Comparison of the computed option prices with the reference solutions. 

<table>
<thead>
<tr>
<th>Volatility value</th>
<th>$(N_x, N_v, N_\tau)$</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v=0.0625$</td>
<td>(25,32,200)</td>
<td>2.0000</td>
<td>1.0682</td>
<td>0.4920</td>
<td>0.1950</td>
<td>0.0760</td>
</tr>
<tr>
<td></td>
<td>(50,64,1600)</td>
<td>2.0000</td>
<td>1.0794</td>
<td>0.4828</td>
<td>0.1851</td>
<td>0.0634</td>
</tr>
<tr>
<td></td>
<td>(100,100,6000)</td>
<td>2.0000</td>
<td>1.0774</td>
<td>0.4789</td>
<td>0.1796</td>
<td>0.0622</td>
</tr>
<tr>
<td></td>
<td>Approximation 1 [78]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Approximation 2 [78]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ref. [63]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ref. [75]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$v=0.25$</td>
<td>(25,32,200)</td>
<td>2.9877</td>
<td>1.3669</td>
<td>0.8521</td>
<td>0.4787</td>
<td>0.2913</td>
</tr>
<tr>
<td></td>
<td>(50,64,1600)</td>
<td>2.9094</td>
<td>1.3645</td>
<td>0.8410</td>
<td>0.4921</td>
<td>0.2756</td>
</tr>
<tr>
<td></td>
<td>(100,100,6000)</td>
<td>2.9093</td>
<td>1.3644</td>
<td>0.8382</td>
<td>0.4884</td>
<td>0.2685</td>
</tr>
<tr>
<td></td>
<td>Approximation 1 [78]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Approximation 2 [78]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ref. [63]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ref. [75]</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Convergence Rate**

Another important issue for a numerical scheme is its convergence rate. Theoretically speaking, we should have a first-order convergence in the $\tau$ direction and a second-order convergence in both $x$ and $v$ direction. This is because the explicit Euler scheme
(a) American put option value at $v = 0.25$ with $\kappa = 2.5$, $\sigma = 0.45$.  
(b) American put option value at $v = 0.25$ with $\kappa = 5$, $\sigma = 0.9$.  

(c) Optimal exercise prices with different volatility values. Model parameters are $\kappa = 2.5$, $\sigma$ = ity values. Model parameters are $\kappa = 5$, $\sigma = 0.9$, 0.45.  
(d) Optimal exercise prices with different volatility values.  

Figure 2.3: American puts and the optimal exercise prices calculated with $\eta = 0.16$, $r = 0.1$, $\rho = 0.1$, $T - t = 0.25$ (year), $K = 10.0$, $\theta = \frac{1}{2}$.  

Table 2.3: Comparison of the computed option prices with the reference solutions. Model parameters are $\kappa = 5$, $\eta = 0.16$, $\sigma = 0.9$, $r = 0.1$, $\rho = 0.1$, $T - t = 0.25$ (year), $K = $10.0.

<table>
<thead>
<tr>
<th>Volatility value</th>
<th>$(N_x, N_v, N_\tau)$</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v=0.0625$</td>
<td>(25,32,500)</td>
<td>2.0000</td>
<td>1.0752</td>
<td>0.5100</td>
<td>0.2200</td>
<td>0.0943</td>
</tr>
<tr>
<td></td>
<td>(50,64,5000)</td>
<td>2.0000</td>
<td>1.0908</td>
<td>0.5073</td>
<td>0.2133</td>
<td>0.0837</td>
</tr>
<tr>
<td></td>
<td>(100,100,50000)</td>
<td>2.0000</td>
<td>1.0987</td>
<td>0.5082</td>
<td>0.2106</td>
<td>0.0861</td>
</tr>
<tr>
<td>PSOR</td>
<td>2.0000</td>
<td>1.1075</td>
<td>0.5190</td>
<td>0.2129</td>
<td>0.0818</td>
<td></td>
</tr>
<tr>
<td>ref.[80]</td>
<td>2.0000</td>
<td>1.107</td>
<td>0.517</td>
<td>0.212</td>
<td>0.0815</td>
<td></td>
</tr>
<tr>
<td>ref.[24]</td>
<td>2.0000</td>
<td>1.1080</td>
<td>0.5316</td>
<td>0.2261</td>
<td>0.0907</td>
<td></td>
</tr>
<tr>
<td>ref.[127]</td>
<td>2.0000</td>
<td>1.1076</td>
<td>0.5202</td>
<td>0.2138</td>
<td>0.0821</td>
<td></td>
</tr>
<tr>
<td>$v=0.25$</td>
<td>(25,32,500)</td>
<td>2.0701</td>
<td>1.3366</td>
<td>0.8131</td>
<td>0.4654</td>
<td>0.2645</td>
</tr>
<tr>
<td></td>
<td>(50,64,5000)</td>
<td>2.0787</td>
<td>1.3335</td>
<td>0.7999</td>
<td>0.4540</td>
<td>0.2474</td>
</tr>
<tr>
<td></td>
<td>(100,100,50000)</td>
<td>2.0781</td>
<td>1.3337</td>
<td>0.7965</td>
<td>0.4496</td>
<td>0.2441</td>
</tr>
<tr>
<td>PSOR</td>
<td>2.0785</td>
<td>1.3336</td>
<td>0.7956</td>
<td>0.4481</td>
<td>0.2427</td>
<td></td>
</tr>
<tr>
<td>ref.[80]</td>
<td>2.079</td>
<td>1.334</td>
<td>0.796</td>
<td>0.449</td>
<td>0.243</td>
<td></td>
</tr>
<tr>
<td>ref.[24]</td>
<td>2.0733</td>
<td>1.3290</td>
<td>0.7992</td>
<td>0.4536</td>
<td>0.2502</td>
<td></td>
</tr>
<tr>
<td>ref.[127]</td>
<td>2.0784</td>
<td>1.3337</td>
<td>0.7961</td>
<td>0.4883</td>
<td>0.2428</td>
<td></td>
</tr>
</tbody>
</table>

is used to construct the predictor, and the central finite difference is adopted for the space variables.

It is demonstrated that the optimal exercise price is far more difficult to calculate accurately than the option price [125]. Furthermore, once $S_f$ is determined accurately, the calculation of the option price becomes straightforward. Consequently, it suffices to focus on the calculation of $S_f$ to study the order of convergence for the current scheme.

Theoretically speaking, to obtain the order of convergence in one direction, we should examine the ratios of the consecutive errors of $S_f$ with the grid spacing along this direction being successively decreased, while the grid spacings along other directions being fixed to be sufficiently small. However, as mentioned earlier, there is no analytical solution available for the American puts with stochastic volatility. One of the standard ways to demonstrate the rate of convergence is to calculate a reference solution based on very fine grids, and use it as if it were the exact solution. We have conducted such an experiment, with the reference solution being constructed with the number of grids defined as $(N_x, N_v, N_\tau) = (400, 400, 100000)$, under the parameters settings: $\kappa = 1.5$, $\eta = 0.16$, $\sigma = 0.1$, $r = 0.1$, $\rho = 0.1$, $T - t = 0.1$ (year), $K = $10.0.

To obtain the order of convergence along the $\tau$ direction, we fix the spatial grid sizes to be $\Delta x = \frac{X_{\text{max}}}{250}$, $v = \frac{V_{\text{max}}}{250}$, and vary the number of time steps from 4000 to
Chapter 2.

Figure 2.4: Option prices at different times to expiration. Model parameters are $\kappa = 2.5$, $\eta = 0.16$, $\sigma = 0.45$, $r = 0.1$, $\rho = 0.1$, $K = \$10.0$.

5000. The errors reported in the following tables are the $L_2$-norm difference between the computed numerical values and the reference solution. Moreover, the Experimental Order of Convergence (EOC) appearing in the $(i+1)$th row of Tables 2.4-2.6 is defined as,

$$\text{EOC}_{i+1} = \frac{\ln \text{error}_{i+1} - \ln \text{error}_i}{\ln N_{\tau, i} - \ln N_{\tau, i+1}}.$$  

According to Table 2.4, when the grid sizes in both $x$ and $v$ directions are fixed and kept to be quite small, the EOCs reported are close to 1, indicating that our scheme is first-order convergent in the time direction.

Similarly, when we fix the time step size to $\Delta \tau = \frac{T}{100000}$, and the grid size in the $x$ (or $v$) direction to be $\frac{X_{\text{max}}}{400}$ (or $\frac{V_{\text{max}}}{400}$), and increase the grid number in the $v$ direction (or in the $x$ direction), we find that the EOCs approach 2, as shown in Table 2.5, and Table 2.6, respectively. This indicates that a second-order convergence is achieved in both $x$ and $v$ directions.

To better investigate the convergence rate of the current scheme, we also calcu-
lated the EOCs with the time and spatial steps adjusted to each other according to the expected order of error $O(\Delta \tau) + O(\Delta x^2) + O(\Delta v^2)$. Specifically, we choose varying grid sizes, i.e., $\Delta \tau_i = h_i^2 \Delta \tau_1$, $\Delta x_i = h_i \Delta x_1$, $\Delta v_i = h_i \Delta v_1$, where $h_i$ is the rate of the grid spacings used in the $i$-th line of Table 2.7 to those appearing in the first line of the same table. We could anticipate that if the theoretical order of convergence is achieved in all directions, the EOC, which is now defined as

$$EOC_{i+1} = \frac{\ln \text{error}_{i+1} - \ln \text{error}_i}{\ln h_{i+1} - \ln h_i},$$

should approach 2. From Table 2.7, we find that the EOCs are approximately equal to 2, which confirms again that the current scheme is indeed first-order convergent in the time direction, and second-order convergent in both $x$ and $v$ directions.

Table 2.4: EOC in the time direction, calculated with $\Delta x = \frac{X_{\text{max}}}{250}$, $\Delta v = \frac{V_{\text{max}}}{250}$.

<table>
<thead>
<tr>
<th>No. of time steps</th>
<th>difference</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>4000</td>
<td>4.8699E-4</td>
<td></td>
</tr>
<tr>
<td>4500</td>
<td>4.3639E-4</td>
<td>0.9314</td>
</tr>
<tr>
<td>5000</td>
<td>3.9212E-4</td>
<td>1.0153</td>
</tr>
<tr>
<td>5500</td>
<td>3.7315E-4</td>
<td>1.0679</td>
</tr>
</tbody>
</table>

Table 2.5: EOC in the $x$ direction, calculated with $\Delta \tau = \frac{T}{100000}$, $\Delta v = \frac{V_{\text{max}}}{400}$.

<table>
<thead>
<tr>
<th>No. of steps in $x$–direction</th>
<th>difference</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.2092</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>0.0891</td>
<td>2.1052</td>
</tr>
<tr>
<td>40</td>
<td>0.0469</td>
<td>2.2305</td>
</tr>
<tr>
<td>80</td>
<td>0.0097</td>
<td>2.2711</td>
</tr>
</tbody>
</table>

Table 2.6: EOC in the $v$ direction, calculated with $\Delta \tau = \frac{T}{100000}$, $\Delta x = \frac{X_{\text{max}}}{400}$.

<table>
<thead>
<tr>
<th>No. of steps in $v$–direction</th>
<th>difference</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.0547</td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>0.0152</td>
<td>1.8475</td>
</tr>
<tr>
<td>40</td>
<td>0.0046</td>
<td>1.7244</td>
</tr>
<tr>
<td>80</td>
<td>0.0013</td>
<td>1.8231</td>
</tr>
</tbody>
</table>
Table 2.7: EOC with varying grid sizes, with $\Delta x_1 = \frac{X_{\text{max}}}{10}$, $\Delta v_1 = \frac{V_{\text{max}}}{10}$, and $\Delta \tau_1 = \frac{T_{\text{max}}}{5}$.

<table>
<thead>
<tr>
<th>h</th>
<th>difference</th>
<th>ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9265</td>
<td></td>
</tr>
<tr>
<td>0.5</td>
<td>0.2783</td>
<td>1.7352</td>
</tr>
<tr>
<td>0.25</td>
<td>0.0680</td>
<td>2.2625</td>
</tr>
<tr>
<td>0.1</td>
<td>0.0070</td>
<td>2.3077</td>
</tr>
</tbody>
</table>

2.4.3 Discussion on accuracy and efficiency

In the financial industry, an ideal numerical scheme is the one with both high efficiency and accuracy. However, generally, the computational efficiency is always inversely proportional to accuracy. Thus, whether or not one can achieve a high efficiency at the expense of losing certain degrees of accuracy should be considered as a key criterion for choosing good numerical method.

In this subsection, we study numerically the relationship between the efficiency and the accuracy of our current scheme. We shall demonstrate that with an acceptable accuracy, our scheme can achieve a relatively high speed. Once again, we shall demonstrate this issue by showing the results of the calculation of the optimal exercise price only.

The tested example is chosen with parameter values: $\kappa = 1.5$, $\eta = 0.16$, $\sigma = 0.2$, $r = 0.1$, $\rho = 0.1$, $T - t = 0.25$ (year), and $K = \$10.0$.

Unfortunately, there is no exact or analytical solution for the valuation of American put options with stochastic volatility. Thus, we computed the reference values using the predictor-corrector method with a very fine grid defined by $N_x = 200$, $N_v = 400$, and $N_\tau = 50000$. These values are used to verify the accuracy of the computed $S_f$ values based on some coarse grid.

Table 2.8: Report on CPU time VS relative error

<table>
<thead>
<tr>
<th>$(N_x,N_v,N_\tau)$</th>
<th>CPU time(s)</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>(13,25,25)</td>
<td>0.5470</td>
<td>6.67%</td>
</tr>
<tr>
<td>(13,50,25)</td>
<td>0.8440</td>
<td>6.43%</td>
</tr>
<tr>
<td>(26,50,80)</td>
<td>3.4370</td>
<td>2.26%</td>
</tr>
<tr>
<td>(52,50,250)</td>
<td>19.5420</td>
<td>1.09%</td>
</tr>
<tr>
<td>(104,100,2000)</td>
<td>529.4920</td>
<td>0.89%</td>
</tr>
</tbody>
</table>
Table 2.8 shows the relationship of the efficiency and accuracy of our scheme. Here, the computational efficiency is measured by the total CPU time in seconds consumed for each run, while the accuracy is measured by the relative error over the whole computational domain, which is defined as:

$$\text{Error} = \frac{||S_f - \tilde{S}_f||_\infty}{||S_f||_\infty}$$

where $S_f$ and $\tilde{S}_f$ denote the computed values and the reference values, respectively, and $||\cdot||_\infty$ denotes the infinite norm. All the experiments here were performed within Matlab 6.1 on an Intel Pentium 4, 3GHZ machine.

As clearly shown in Table 2.8, the accuracy is inversely varying with the efficiency. Furthermore, a better resolution in the asset values or the variance values implies more computational time. It can also be observed that our numerical results converge quickly to the reference values, and even with the most coarse grid, the accuracy is still acceptable by the industry standard. On the other hand, based on previous sections, it is suggested that the current method requires a lower number of time step intervals when $\sigma$ is small. One should note that $\sigma$ is usually very small in the real market. Therefore, our scheme is indeed a good one with a high computational efficiency and a satisfactory accuracy, and is suited for the practical use.

2.5 Conclusion

In this chapter, we first proposed appropriate boundary conditions along the $v$ direction for American puts under the Heston model, and made the valuation system closed. Then, under this closed system, a new predictor-corrector scheme based on the ADI method was demonstrated and tested. The novelty of the current scheme, in comparison with some other finite difference methods in the literature, lies in the fact that it requires no embedded iterations, and can capture the whole optimal exercise boundary as part of the solution procedure. Based on the local von Neumann analysis, combined with the “frozen” coefficient technique, a conditional stability re-
quirement was also obtained for the predictor-corrector method. It is suggested that the good convergence property of the corrector can influence the whole procedure significantly, which has also been confirmed by some of our numerical results. This approach can be easily extended to other stochastic volatility models and to non-zero choices of the market price of risk function, as long as a front-fixing transform exists. Various numerical examples suggest that the proposed scheme is both accurate and efficient, and is also suitable for practical use.
Chapter 3

A spectral-collocation method for pricing perpetual American puts with stochastic volatility

3.1 Introduction

In Chapter 2, we have proposed a predictor-corrector method to solve American options with stochastic volatility. This method is quite efficient and accurate in dealing with American options with finite maturities. However, when the lifespan of the American options becomes extremely long, the proposed method becomes less efficient and accurate. This chapter will address the numerical valuation of perpetual American options with stochastic volatility.

Perpetual American options are options that can be exercised at any time, but without a definitive expiry day. This simple contract can be viewed as an approximation for long-dated American options. Also, analyzing such options may in principle be used as a building block in an approximation procedure for American options with finite maturities [15].

The availability of a closed-form solution of perpetual American option under the BS (Black-Scholes) framework has already been achieved [79]. Empirical evidence,
however, suggests that the BS model is inadequate to describe asset returns and the behavior of the option markets [1]. The literature advocates the introduction of stochastic volatility to reproduce the implied volatility smile observed in markets. Among those SV (stochastic volatility) models (e.g. [6, 38, 51, 57]), the one proposed by Heston has received the most attention, due to its great analytical tractability for European options, and we shall use this model throughout this chapter. In fact, the extension of the current scheme to other SV models should be straightforward.

For perpetual American options under the Heston model, there is no analytical solution, primarily due to the fact that the optimal exercise price now remains unknown as a function of volatility, rather than a constant as in the BS case. In other words, the introduction of a second stochastic process has considerably complicated the solution process in pricing perpetual American options.

In the literature, several numerical approaches were introduced to solve the free boundary problem associated with the valuation of American options under the Heston model (e.g. [24, 119]), but they all concentrated on options with finite maturities. It should be pointed out that all these approaches are not suitable for pricing perpetual American options, since large time evolution is required to approximate the infinite maturity, resulting in both low accuracy and computational inefficiency. On the other hand, although the Projected SOR (PSOR) method can be adopted to solve directly the LCP (linear complementary problem) associated with the current valuation problem, the accuracy and the efficiency of this classical approach are still not satisfying. To date, no documented numerical schemes have been proposed to address this issue; an efficient numerical approach for the valuation of perpetual American options under the Heston model is the aim of this chapter.

In this chapter, a numerical approach based on the spectral-collocation (SC) method is proposed for the valuation of perpetual American puts under the Heston model. This approach consists of two steps. The first step is to derive a system of nonlinear algebraic equations by using the SC method. The second step is to transform the nonlinear system obtained in the first step into a nonlinear least-
square problem (NLSP), and solve it with the Gauss-Newton algorithm. To make sure that our numerical solution converges to the correct one, a test example similar to the original problem is constructed first, and then, the numerical results of the option price are compared with those produced by the PSOR method. Our numerical experiments show that this approach is both accurate and efficient, since the desired spectral accuracy can be easily achieved with a small number of iterations.

The chapter is organized as follows. In Section 3.2, we introduce the PDE (partial differential equation) system that the price of a perpetual American put option must satisfy under the Heston model. In Section 3.3, we introduce our numerical approach in details. In Section 3.4, numerical results and some useful discussions are presented. Concluding remarks are given in the last section.

### 3.2 Perpetual American puts under the Heston model

As pointed out previously, perpetual American option is American option with infinite expiration. Let $U(v, S)$ denote the value of a perpetual American put option, with $S$ being the underlying and $v$ being the variance. Then, under the Heston model, it can be easily shown that the valuation problem of a perpetual American put option can be formulated as a free boundary problem [97], in which the boundary location itself is part of the solution of the problem. In particular, $U(v, S)$ should satisfy:

$$
\begin{align*}
& \left\{ \begin{array}{l}
\frac{1}{2}vS^2 \frac{\partial^2 U}{\partial S^2} + \rho \sigma vS \frac{\partial^2 U}{\partial S \partial v} + \frac{1}{2} \sigma^2 v \frac{\partial^2 U}{\partial v^2} + rS \frac{\partial U}{\partial S} + \kappa (\eta - v) \frac{\partial U}{\partial v} - rU = 0, \\
\lim_{S \to -\infty} U(v, S) = 0, \\
U(v, S_f(v)) = K - S_f(v), \quad \frac{\partial U}{\partial S}(v, S_f(v)) = -1, \\
\lim_{v \to 0} U(v, S) = \max(K - S, 0), \quad \lim_{v \to -\infty} U(v, S) = K,
\end{array} \right.
\end{align*}
$$

(3.2.1)

where $S_f(v)$ denotes the optimal exercise price.

One should notice that, once the stochastic volatility is taken into consideration,
the valuation of the perpetual American puts is no longer as analytical achievable as the constant volatility case, because the optimal exercise price now remains unknown as a function of the volatility, while in the BS model, it is only an unknown constant. However, due to the time-independence, this pricing problem is somehow simplified, compared with the valuation of American puts with finite maturities.

### 3.3 Numerical scheme based on the Legendre pseudospectral method

Generally, it is not worth bothering with the documented numerical approaches designed for American puts with finite maturities to deal with this simplified problem, since otherwise, large time evolution is required to approximate the infinite maturity, resulting in both low accuracy and computational inefficiency. In this section, a new numerical scheme based on the Legendre pseudospectral method is introduced to solve the pricing of perpetual American puts efficiently and accurately.

In order to apply our new numerical scheme, we first adopt the Landau transform [72], i.e.,
\[
x = \ln \frac{S}{S_f}
\]
(3.3.2)
to convert the free boundary conditions to fixed boundary conditions. Furthermore, since the optimal exercise price is related to the option price by the conditions across the free boundary, we write it as
\[
S_f(v) = K - U(v, S_f).
\]
(3.3.3)

By substituting (3.3.2)-(3.3.3) into (3.2.1), we obtain

\[
\begin{cases}
    LU = 0, \\
    \lim_{x \to \infty} U(v, x) = 0, \ U(v, 0) = K + \frac{\partial U}{\partial x}(v, 0), \\
    \lim_{v \to 0} U(v, x) = \max[K - e^x(K - U(v, 0)), 0], \ \lim_{v \to \infty} \frac{\partial U}{\partial v}(v, x) = 0,
\end{cases}
\]
(3.3.4)
Chapter 3.

where

$$L = a(v) \frac{\partial^2}{\partial x^2} + b(v) \frac{\partial^2}{\partial v^2} + c(v) \frac{\partial^2}{\partial x \partial v} + d(v) \frac{\partial}{\partial x} + e(v) \frac{\partial}{\partial v} - r,$$

and

$$a(v) = \frac{1}{2} v + \frac{1}{2} \sigma^2 \xi^2 v - \rho \sigma v \xi,$$
$$b(v) = \frac{1}{2} \sigma^2 v,$$
$$c(v) = \rho \sigma v - \sigma^2 v \xi,$$
$$d(v) = -\frac{1}{2} v + \frac{1}{2} \xi^2 \sigma^2 v - \frac{1}{2} \sigma^2 v \beta + r - \kappa (\eta - v) \xi,$$
$$e(v) = \kappa (\eta - v),$$

$$\xi = \frac{1}{U(v,0) - K} \frac{\partial U}{\partial v}(v,0),$$
$$\beta = \frac{1}{U(v,0) - K} \frac{\partial^2 U}{\partial v^2}(v,0).$$

It can be clearly seen that after the Landau transform is applied, the nonlinear feature of the problem is explicitly exposed in the governing equation.

On the other hand, it can be observed that the option pricing problem is defined on an unbounded domain

$$\{ (v, x) | v \geq 0, x \geq 0 \}.$$

To implement a calculation in a computer, we truncate the semi-infinite domain into a finite domain:

$$\{ (v, x) \in [0, v_{\text{max}}] \times [0, x_{\text{max}}] \}.$$

Theoretically, $x_{\text{max}}$ and $v_{\text{max}}$ should be sufficiently large to eliminate the boundary effect. However, based on Willmott et al.’s estimate [111] that the upper bound of the asset price $S_{\text{max}}$ is typically three or four times of the strike price, it is reasonable for us to set $x_{\text{max}} = \ln 5$. On the other hand, 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 about 0.85 [34]. Thus, it is quite reasonable to set $v_{\text{max}} = 1$.

Now, with the truncated computational domain in hand, it is enough for us to introduce our new approach for solving (3.3.4). The most crucial step in the implementation of the current scheme is to derive the differential matrices (cf. [92])
for the PDE system (3.3.4). Suppose that the spectral approximation solution of (3.3.4) can be written as

\[ U(v, x) = \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} U(v_i, x_j) F_i(v) F_j(x), \]

where \( F_k(x) \) is the \( k \)-th Lagrange basis function, and \((v_i, x_j)\) are \( N + 1 \) collocation points. Based on the definition of \( F_j(v) \), it is quite straightforward to show that

\[ \frac{\partial U}{\partial x}(v_l, x_k) = \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} U(v_i, x_j) F_i(v_l) F'_j(x_k), \]

which is equivalent to

\[ \frac{\partial U}{\partial x} = U D^T_X, \] where \( \frac{\partial U}{\partial x}, U \) and \( D^T_X \) are defined as:

\[ \frac{\partial U}{\partial x} = (\frac{\partial U}{\partial x}(v_i, x_j))_{N+1,N+1}, \quad U = (U(v_i, x_j))_{N+1,N+1}, \quad D_X = (F'_j(x_i))_{N+1,N+1}. \]

Similarly, we obtain

\[ \frac{\partial U}{\partial v} = D_v U, \quad \frac{\partial^2 U}{\partial x^2} = U (D^2_X)^T, \quad \frac{\partial^2 U}{\partial v^2} = D^2_v U, \quad \frac{\partial^2 U}{\partial x \partial v} = D_v U D^T_X. \quad (3.3.6) \]

On the other hand, in order to obtain the numerical values of the differential matrices without much effort, we should choose some proper collocation points. It can be shown that if we use the so-called Legendre-Gauss-Lobatto points \( \{s_i\}_{i=1}^{N+1} \), the corresponding differential matrix \( D_s \) has the following structure [92]:

\[ (D_s)_{i,j} = \frac{L_N(s_i)}{L_N(s_j)} \frac{1}{s_i - s_j}, \quad i \neq j, \quad i, j = 2 \cdots N; \]
\[ (D_s)_{i,i} = 0, \quad i \neq 1, i \neq N + 1, \]
\[ (D_s)_{1,1} = -(D_s)_{N+1,N+1} = \frac{N(N+1)}{4}. \]

Since \( \{s_i\}_{i=1}^{N+1} \subset [-1, 1] \), the following coordinate transformation should be applied,
i.e.,
\[
v_i = \frac{(s_i + 1)v_{max}}{2}, \quad x_j = \frac{(s_j + 1)x_{max}}{2},
\]
which yields
\[
D_v = 2v_{max}D_s, \quad D_x = 2x_{max}D_s.
\]

By substituting (3.3.6) into (3.3.4), we obtain the nonlinear algebraic equations for the unknown matrix \( U \) as
\[
\begin{align*}
&\left\{ \begin{array}{l}
  a_i(UD_x^2)^T_{i,j} + b_i(D_v^2U)_{i,j} + c_i(D_vUD_x^T)_{i,j} + d_i(UD_x^T)_{i,j} + e_i(D_vU)_{i,j} \\
  -r(U)_{i,j} = 0, \quad i, j = 2 \cdots N, \\
  (U)_{1,j} = 0, \quad j = 2 \cdots N, \\
  (D_vU)_{N+1,j} = 0, \quad j = 2 \cdots N, \\
  (U)_{i,1} = K + (UD_x^T)_{i,1}, \quad i = 2 \cdots N, \\
  (U)_{i,N+1} = 0, \quad i = 2 \cdots N.
\end{array} \right.
\end{align*}
\]

The above system can be solved efficiently with the utilization of the Gauss-Newton algorithm. The iterative process is as follows:

(i) If \( U^{(k)} \) is obtained after the \( k \)-th iterative step, we can compute \( f(U^{(k)}) \), and the corresponding Jacobian \( J_f(U^{(k)}) \), where \( f(U) \) is a vector of “residuals”. Specifically,

\[
\begin{align*}
f_{(i-1)(N+1)+j} &= a_i(UD_x^2)^T_{i,j} + b_i(D_v^2U)_{i,j} + c_i(D_vUD_x^T)_{i,j} + d_i(UD_x^T)_{i,j} + e_i(D_vU)_{i,j} - r(U)_{i,j}, \quad i, j = 2 \cdots N, \\
f_j &= (U)_{1,j}, \quad j = 1 \cdots N + 1, \\
f_j &= (D_vU)_{N+1,k}, \quad j = N(N + 1) \cdots (N + 1)^2, \quad k = 1 \cdots (N + 1), \\
f_{(i-1)(N+1)+1} &= (U)_{i,1} - K - (UD_x^T)_{i,1}, \quad i = 2 \cdots N, \\
f_{(i-1)(N+1)+N+1} &= (U)_{i,N+1}, \quad i = 2 \cdots N.
\end{align*}
\]
(ii) Linearize $f$ with the current value $U^{(k)}$, i.e.,

$$f(U) \approx f(U^{(k)}) + J_f(U^{(k)})(U - U^{(k)}),$$

$$= A^{(k)}U - b^{(k)},$$

where $A^{(k)} = J_f(U^{(k)})$ and $b^{(k)} = J_f(U^{(k)})U^{(k)} - f(U^{(k)})$.

(iii) Solve the following linear-square problem:

$$\| f(U) \| = \| A^{(k)}U - b^{(k)} \|^2,$$

and obtain $U^{(k+1)} = (A^{(k)}T A^{(k)})^{-1} A^{(k)}T b^{(k)}$.

(iv) Repeat (i)-(iii) until $\| U^{(k+1)} - U^{(k)} \| < \epsilon$ is satisfied. The tolerance is set to $10^{-6}$ for all the results presented in this chapter.

One should notice that in the above Gauss-Newton method, it is quite important to choose a proper initial guess of $U$, i.e., $U^{(0)}$, since the algorithm may converge slowly or not at all if the initial guess is too far away from the final solution. For most of the practical parameter-settings, we recommend to use the analytical formula for perpetual American puts under the BS model, with the corresponding variance setting to $v$, as a good initial guess. This is because, firstly, the formula satisfies all the boundary conditions automatically, and secondly, it should be close to the final solution of (3.3.4), since with the same parameter-settings, the option prices under the two different models should not differ too much.

### 3.4 Numerical results and discussions

In this section, we shall present the numerical results as well as some useful discussions. The section is organized into three subsections, according to three important issues that should be addressed.
3.4.1 A test example

As demonstrated in Section 3.1, no analytical solution for the case of perpetual vanilla American option under the Heston model has yet been derived. Thus, in order to illustrate the reliability of the current scheme, we should conduct a test example, for which, the analytical solution can be artificially constructed. The test example is as follows:

\[
\begin{align*}
\mathbb{L}U + rK &= 0, \\
\lim_{x \to x_{\text{max}}} U(v, x) &= K + ve^{x_{\text{max}}}, \\
U(v, 0) &= K + \frac{\partial U}{\partial x}(v, 0), \\
\lim_{v \to -0} U(v, x) &= K, \\
\lim_{v \to v_{\text{max}}} U(v, x) &= K + v_{\text{max}}e^{x},
\end{align*}
\]

(3.4.7)

where \( \mathbb{L} \) is the same operator as shown in (3.3.5). This test example can be viewed as the pricing of some kind of perpetual American-style exotic options under the Heston model, and, therefore, it keeps the essential nonlinear feature of the original problem. On the other hand, the price of this option is equal to \( U_{\text{exact}} = K + ve^{x} \).

This test example has almost the same structure as our original problem. Therefore, if the error of the numerical results of this example is reasonably small, we should have confidence that the proposed scheme is quite accurate in solving (3.3.4) as well.

Table 3.1: The test example. Parameters are \( \kappa = 2, \eta = 0.2, \sigma = 0.04, \tau = 0.5, \rho = 0.1, \) \( K = 10.0 \).
Table 3.1 shows the numerical results of the test example with the initial guess $U_0 = 3 + K + v^2 x$. Here, $N$ stands for the number of collocation points along each direction, and the error is defined as the maximum point-wise error. Furthermore, since the Gauss-Newton algorithm is adopted to solve the nonlinear algebraic equation system, we have also displayed the number of iterations and the final residuals measured in the $L_2$-norm. From this table, it can be observed that the rapid convergence of our numerical solution to the exact solution can be easily achieved with a small number of iterations. Most remarkably, a desired spectral accuracy can be obtained even when very coarse grids are adopted. Therefore, the proposed scheme is both accurate and efficient in solving nonlinear problems, especially those with structures similar to the test example, such as our problem for perpetual American puts with stochastic volatility.

### 3.4.2 SC scheme VS PSOR method

As mentioned earlier, the pricing of perpetual American puts under the Heston model can also be solved by the well-known PSOR method introduced in [29]. It is quite interesting to make a comparison of the two different approaches. In Table 3.2, we compared the option prices calculated by the PSOR method and the SC scheme. Furthermore, the corresponding number of iterations and the total CPU-time cost are also displayed. All the experiments here were performed within Matlab7.5 on an Intel Pentium 4, 3GHZ machine.

<table>
<thead>
<tr>
<th>grid number $(N_v, N_x)$</th>
<th>asset values 8</th>
<th>asset values 9</th>
<th>asset values 10</th>
<th>asset values 11</th>
<th>No. of iterations</th>
<th>CPU-time(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PSOR (25,50)</td>
<td>2.2982</td>
<td>1.8159</td>
<td>1.4764</td>
<td>1.2281</td>
<td>458</td>
<td>6.4</td>
</tr>
<tr>
<td>(50,100)</td>
<td>2.3156</td>
<td>1.8368</td>
<td>1.4983</td>
<td>1.2497</td>
<td>1725</td>
<td>16.5</td>
</tr>
<tr>
<td>(100,150)</td>
<td>2.3244</td>
<td>1.8476</td>
<td>1.5095</td>
<td>1.2604</td>
<td>2484</td>
<td>246.2</td>
</tr>
<tr>
<td>(100,200)</td>
<td>2.3244</td>
<td>1.8470</td>
<td>1.5091</td>
<td>1.2602</td>
<td>6396</td>
<td>1127.8</td>
</tr>
<tr>
<td>SC (15,15)</td>
<td>2.3299</td>
<td>1.8531</td>
<td>1.5128</td>
<td>1.2796</td>
<td>6</td>
<td>2.1</td>
</tr>
<tr>
<td>(20,20)</td>
<td>2.3273</td>
<td>1.8500</td>
<td>1.5093</td>
<td>1.2560</td>
<td>7</td>
<td>4.1</td>
</tr>
<tr>
<td>(25,25)</td>
<td>2.3280</td>
<td>1.8514</td>
<td>1.5111</td>
<td>1.2581</td>
<td>7</td>
<td>9.3</td>
</tr>
<tr>
<td>(30,30)</td>
<td>2.3259</td>
<td>1.8485</td>
<td>1.5078</td>
<td>1.2546</td>
<td>8</td>
<td>20.9</td>
</tr>
</tbody>
</table>

It can be clearly seen from this table that for the finest grid, the option prices...
produced by the two numerical approaches agree well with each other, but with substantially different CPU-time consumption. Obviously, the CPU-times required by the PSOR method are significantly more than the SC approach. Furthermore, it can also be observed that the option prices produced by the SC method with the most coarse grid (the grid number \((15, 15)\)) are already very close to those computed by the PSOR method with the finest grid (the grid number \((100, 200)\)), and yet the CPU time needed for the former is remarkably less than 3 seconds, a fraction of roughly 600th of the latter!

On the other hand, by calculating the option prices with different parameter settings, it is found that the convergence rate of the PSOR method deteriorates when the parameter \(\kappa\) increases, while for the SC scheme, it seems to be less parameter-dependent. Therefore, due to the high efficiency and accuracy, it is suggested that the SC scheme, superior to the PSOR method, be adopted to value perpetual American puts under the Heston model.

Before leaving this subsection, it should be remarked that the numerical solution produced by the SC method may be adopted as a benchmark for future studies. This is because the high order convergence of the current scheme is ensured by testing the constructed perpetual American exotic option, as shown in section 3.4.1. Furthermore, the convergence of our solution to the exact one has also been numerically guaranteed by the comparison with the PSOR method.

### 3.4.3 The impact of stochastic volatility

As mentioned earlier, the option prices of perpetual American puts under the BS model can be calculated with a simple and elegant formula [79]. With the current numerical scheme, it is enough for us to make a comparison of the pricing difference for two perpetual American put option contracts being otherwise identical except the volatility terms. Such a comparison is quite interesting, since it can give us a quantitative sense on the largest effect of the stochastic volatility on the prices of American puts. This is because the impact on pricing of a stochastic volatility
usually becomes progressively larger as the lifespan of the option increases [57], and for perpetual case, the impact should undoubtedly be the most significant.

Plotted in Fig 3.1 are the comparison of the option prices under the two different models, with the variance rate under the BS model being $v$. It is interesting to notice that there is a special value of $v$, at which, the option prices under the two different models agree well with each other, as shown in Fig 3.1-a. Moreover, for $v$ less than this “special point”, the prices under the BS model are lower than the corresponding ones with stochastic volatility (Fig 3.1-b), while for $v$ larger than that point, the BS prices turn out to be higher (Fig 3.1-c).
first. Displayed in Tables 3.3-3.5 are the values of “special” \( v \) with different parameter settings. Specifically, they are calculated by fixing the parameters \((\rho, \sigma), (\kappa, \sigma)\) and \((\rho, \kappa)\), respectively. Remarkably, it is quite reasonable for us to concentrate on the variation of the “special” \( v \) related to \( \kappa, \sigma, \rho \) and \( \eta \) only, because these four parameters are introduced once the stochastic volatility is taken into consideration, and it is believed that they can provide enough information on the stochastic volatility term.

Table 3.3: “Special” \( v \) with \( \rho = 0.1, \sigma = 0.45, r = 0.1, K = \$10.0 \)

<table>
<thead>
<tr>
<th>( \kappa )</th>
<th>( \eta = 0.2 )</th>
<th>( \eta = 0.3 )</th>
<th>( \eta = 0.4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( v = 0.245 )</td>
<td>( v = 0.335 )</td>
<td>( v = 0.425 )</td>
</tr>
<tr>
<td>2</td>
<td>( v = 0.245 )</td>
<td>( v = 0.320 )</td>
<td>( v = 0.385 )</td>
</tr>
<tr>
<td>3</td>
<td>( v = 0.245 )</td>
<td>( v = 0.3 )</td>
<td>( v = 0.385 )</td>
</tr>
<tr>
<td>4</td>
<td>( v = 0.245 )</td>
<td>( v = 0.335 )</td>
<td>( v = 0.385 )</td>
</tr>
</tbody>
</table>

Table 3.4: “Special” \( v \) with \( \kappa = 1, \sigma = 0.45, r = 0.1, K = \$10.0 \)

<table>
<thead>
<tr>
<th>( \rho )</th>
<th>( \eta = 0.2 )</th>
<th>( \eta = 0.3 )</th>
<th>( \eta = 0.4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pm 0.1 )</td>
<td>( v = 0.245 )</td>
<td>( v = 0.335 )</td>
<td>( v = 0.425 )</td>
</tr>
<tr>
<td>( \pm 0.5 )</td>
<td>( v = 0.245 )</td>
<td>( v = 0.335 )</td>
<td>( v = 0.425 )</td>
</tr>
</tbody>
</table>

Table 3.5: “Special” \( v \) with \( \rho = 0.1, \kappa = 1, r = 0.1, K = \$10.0 \)

<table>
<thead>
<tr>
<th>( \sigma )</th>
<th>( \eta = 0.2 )</th>
<th>( \eta = 0.3 )</th>
<th>( \eta = 0.4 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>( v = 0.230 )</td>
<td>( v = 0.285 )</td>
<td>( v = 0.395 )</td>
</tr>
<tr>
<td>0.2</td>
<td>( v = 0.245 )</td>
<td>( v = 0.315 )</td>
<td>( v = 0.395 )</td>
</tr>
<tr>
<td>0.3</td>
<td>( v = 0.245 )</td>
<td>( v = 0.335 )</td>
<td>( v = 0.395 )</td>
</tr>
<tr>
<td>0.4</td>
<td>( v = 0.245 )</td>
<td>( v = 0.345 )</td>
<td>( v = 0.415 )</td>
</tr>
</tbody>
</table>

From Tables 3.3-3.5, it is interesting to observe that the “special” \( v \) is approximately equal to the long-term mean (\( \eta \)) of the volatility process. Once \( \eta \) is fixed, the “special” \( v \) does not change significantly with respect to the changes of other parameters. This could be probably explained as follows.

For the mean-reverting process, it is reasonable to infer that as time goes to infinity, the variance \( v \) of the volatility should approach its long-term mean \( \eta \) asymptotically. Supposing that the spot variance \( v_c < \eta (v_c > \eta) \), as time involves to infinity, it should overall increase (decrease) to \( \eta \). For the BS model, it just ignored the growing
(decreasing) tendency of the variance, resulting in option prices lower (higher) than the corresponding ones with stochastic volatility. Note that the option prices are monotonically increasing with respect to $v$. When $v_c \approx \eta$, the overall change of the variance in the long run is not significant, and thus the option prices under the two different models are almost the same. One should notice that, for the American puts with finite maturities, the above explanation is not true. This is because for finite maturity, the overall tendency of $v$ depends on several factors, such as its correlation with the asset price, its long-term mean and so on; one cannot simply determine which one is the dominant factor.

### 3.5 Conclusion

In this chapter, we have considered a spectral-collocation method for the numerical pricing of perpetual American puts when Heston’s stochastic volatility model is used. The option price can be obtained with two steps. The first step is to derive a system of nonlinear algebraic equations by means of the Legendre pseudospectral method, while the second step is to transform the above nonlinear system into a nonlinear least-square problem, which can be solved by the Gauss-Newton algorithm. Our numerical experiments suggest that the current approach is indeed fast and accurate in solving perpetual American puts with stochastic volatility. Moreover, based on the numerical results, a financially meaningful explanation for the effect of the stochastic volatility on the prices of perpetual American puts is also provided.
Chapter 4

An inverse finite element method for pricing American options

4.1 Introduction

In the previous two chapters, we have demonstrated how to price numerically American options with stochastic volatility. In this chapter, we turn back to the classical BS (Black-Scholes) model, by realizing the fact that a “convergence-proved” numerical approach has never been proposed even for the constant volatility case. We shall propose a “convergence-proved” numerical approach in solving American options with constant volatility.

For quite a long time, it has been widely acknowledged that pricing American options is a “much more intriguing” problem [55, 65, 75]. Although recently, the analytical closed-form pricing formula for the American puts under the BS model was successfully derived by Zhu [115], it is not computationally appealing, as the formula involves two infinite sums of infinite double integrals, which take a formidable amount of time to evaluate. Till now, approximation methods are still popular among those market practitioners as they are usually faster with acceptable accuracy.

In the literature, of all the approximation methods, there are predominately two types, analytical approximations and numerical methods for the valuation of an
American option contract, as demonstrated in Chapter 1. Of all the numerical methods, the multi-level FD (finite difference) approach proposed by Wu and Kwok [112] has received the most attention, as their method is quite efficient and accurate, and can be extended to the pricing of exotic options without too much additional effort. In their approach, a linearization process is firstly invoked to deal with the nonlinear nature of the American option pricing problems, and the information obtained will be used to start the second-level computation for the option prices. Whilst appealing, there was no theoretical proof for the convergence of this elegant approach, and moreover, even if the convergence could be established theoretically, it would only guarantee that the numerical results converge to the linearized system, rather than the original nonlinear one. In fact, it should be pointed out that for most of the numerical methods mentioned above, theoretical convergence analyses have not yet been provided. Of course, most authors demonstrated the convergence through their numerical examples. But, lack of convergence proof requires caution when these approaches are adopted on a case-by-case basis. This has motivated us to consider a “convergence-proved” approach for the valuation of American options. In the meantime, we also aim to have an efficient scheme that suits the practical needs of market practitioners.

In this chapter, the IFE (inverse finite element) approach [2], which was originated in solving nonlinear problems associated with phase change, is introduced to the field of option pricing for the first time. As its name suggests, this method is based on an inversely defined problem, and is implemented by using a combination of both the FE (finite element) method and the Newton iteration scheme. Essentially, the IFE approach solves the location (the nodes of the finite elements), at which the variable has a specific value. The advantage of the IFE approach mainly comes from two aspects. Computationally, this method is quite efficient since it allows the use of full Newton iteration method with its inherent quadratic convergence. Furthermore, the IFE method does not require any complicated co-ordinate transformations, in which the location of the free boundary is somehow fixed in the new co-ordinate
system, and thus has a great potential of being extended to deal with multi-asset or stochastic volatility pricing problems.

Compared with some of the widely used numerical methods for solving American option prices [112, 125], the IFE approach involves no linearization process at all, while most of the others require some sort of linearization of the PDE (partial differential equation) system. It is this rather unique feature of the IFE approach that has made it worthwhile to pursue a theoretical convergence analysis, since such an analysis can ensure that the numerical result does indeed converge to the exact solution of the original nonlinear PDE system, contrary to those involving linearization process, with which the analysis, if achieved, can only guarantee that the numerical results will converge to the linearized system rather than the original nonlinear one. On the other hand, it is usually easy to design a numerical scheme to solve a PDE system, but much harder to prove the convergence of this scheme. It is probably even more difficult to provide a theoretical proof for the convergence of the IFE approach since this method involves both the FE formulation and the Newton iteration as part of the solution procedure. For this reason, in the literature, the issue of the convergence of the IFE approach was not attempted at all. Despite those difficulties, we have managed to provide such a convergence proof to show a theoretical threshold for the convergence of the IFE scheme adopted to price American options.

The chapter is organized as follows. In Section 4.2, we introduce the PDE system that the price of an American put option must satisfy under the BS model. In Section 4.3, we present the IFE approach in details. In Section 4.4, we provide a convergence analysis of the current scheme. In Section 4.5, some numerical examples and discussions are presented to illustrate the performance of the IFE method. Concluding remarks are given in the last section.

4.2 Mathematical formulation

To demonstrate the IFE approach in a clear way, we shall adopt the simplest BS model throughout the chapter. In fact, the extensions of the current method to
higher dimensional problems, such as pricing options with multi-assets or stochastic volatilities, should be straightforward.

Let \( P(S, t) \) denote the value of an American put option, with \( S \) being the price of the underlying and \( t \) being the current time. With six main assumptions [10], Black and Scholes showed that the value of an American put option must satisfy:

\[
\begin{align*}
\frac{\partial P}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 P}{\partial S^2} + rS \frac{\partial P}{\partial S} - rP &= 0, \\
P(x, T) &= \max(K - S, 0), \\
P(S_f(t), t) &= K - S_f(t), \quad \frac{\partial P}{\partial S}(S_f(t), t) = -1, \\
\lim_{s \to \infty} P(S, t) &= 0.
\end{align*}
\]  

This PDE system is defined on \( S \in [S_f(t), \infty), \) \( t \in [0, T]. \) In (4.2.1), \( r \) and \( \sigma \) are the risk-free interest rate and the volatility of the underlying, respectively. \( K \) is the strike price and \( T \) is the expiration date. Moreover, \( S_f(t) \) is the optimal exercise price, which also needs to be determined as part of the solution of the PDE system. One should notice that although the governing differential equation itself is linear in terms of the unknown function \( P, \) it is the unknown boundary, \( S_f(t), \) that has made this type of problem highly nonlinear.

To apply the IFE approach, the range of \( P \) should be known \textit{a priori}, as shall be discussed later. However, from (4.2.1), one can easily deduce that the option price \( P \) shall fall within \([0, K - S_f(t)]\), which varies w.r.t. (with respect to) the time \( t, \) as a result of the optimal exercise price being an unknown function of \( t. \) In order to overcome this difficulty and facilitate the development of the algorithm, the following transforms are applied:

\[
x = \ln \frac{S}{K}, \quad U = \frac{P + S - K}{K}, \quad \tau = \frac{\sigma^2}{2}(T - t),
\]
which yield

\[
\begin{aligned}
\frac{\partial U}{\partial \tau} &= \frac{\partial^2 U}{\partial x^2} + (\gamma - 1) \frac{\partial U}{\partial x} - \gamma U - \gamma, \\
U(x, 0) &= \max(1 - \exp(x), 0) + \exp(x) - 1, \\
U(x_f(\tau), \tau) &= 0, \quad \frac{\partial U}{\partial x}(x_f(\tau), \tau) = 0, \\
\lim_{x \to \infty} U(x, \tau) &= \exp(x) - 1,
\end{aligned}
\]

(4.2.2)

with the only parameter \( \gamma \) being the relative interest rate, which is related to the original risk-free interest rate \( r \) by \( r = \frac{\sigma^2 \gamma}{2} \). It should be pointed out that the initial condition in (4.2.2) can be further simplified. Since the optimal exercise price is equal to the strike price at the expiration time \( T \), as shown in [115], i.e., \( x_f(0) = 0 \) in (4.2.2), we must have \( 1 - \exp(x) \leq 0 \) when \( x \) is in the range \( x_f \leq x < +\infty \). Consequently, the initial condition in (4.2.2) can be simplified as \( U(x, 0) = \exp(x) - 1 \). Moreover, it is not difficult to show that \( U \) is a strictly monotonically increasing function w.r.t. \( x \) for \( x \in (x_f, +\infty) \). This is achieved by evaluating

\[
\frac{\partial U}{\partial x} = \left( \frac{\partial P}{\partial S} + 1 \right) \frac{1}{S},
\]

which is greater than zero because the Delta of an American put option is more than \(-1\) for \( S \in (S_f, +\infty) \). Thus, \( U \) is monotonically increasing w.r.t. \( x \) for \( x \in (x_f, +\infty) \).

On the other hand, realistically to implement a calculation in a computer, we truncate the semi-infinite domain into a finite domain: \( \{ x \in [0, x_{\text{max}}] \} \). Theoretically speaking, \( x_{\text{max}} \) should be sufficiently large to eliminate the boundary effect. However, based on Willmott et al.'s estimate [111] that the upper bound of the underlying price \( S_{\text{max}} \) is typically three or four times of the strike price, it is reasonable to set \( x_{\text{max}} = \ln 5 \).

One should notice that after the transformation of variables and the truncation of the computational domain, the range of \( U \) becomes \([0, \exp(x_{\text{max}}) - 1]\), while the monotonicity is retained. This is an advantage that we can take when the IFE method is applied to solve free boundary problems, as shall be discussed in the next section.
4.3 Inverse finite element approach

The IFE approach was initially suggested by Alexandrou in solving nonlinear problems associated with phase change and in particular with solidification [2]. The essential concept of the method is to find the location (the nodes of the finite elements) at which, the original dependent variable, such as the temperature, has a specific value. In other words, the roles of “dependent” and “independent” variables are switched so that the original problem is solved in an inverse way. In financial terminology, this method uses the concept of fixing the option price while studying the motion of the different “isotherms” of the underlying. The solution of the problem through this “inverse” approach 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.

It should be pointed out that although the IFE approach is quite elegant in dealing with free boundary problems, the convergence of such an approach has never been theoretically achieved, even in [2]. Here in this chapter, we shall present the application of a “convergence-proved” IFE approach to the valuation of American put options in details.

To apply the IFE approach, the first step is to deal with the time derivative appearing in the governing equation contained in (4.2.2). This step is, in fact, necessary for most of the numerical schemes designed for solving time-dependent problems. Now, according to the essence of the IFE approach, the option price \( U \) is obtained at a selected underlying price which varies w.r.t. \( \tau \), and, therefore, we obtain

\[
\frac{dU}{d\tau} = \frac{\partial U}{\partial \tau} + \frac{\partial U}{\partial x} V_{\text{mesh}},
\]

which, combined with the fact that \( U \) is kept constant at the nodal points during the whole computation process, as shall be shown later, yields

\[
\frac{\partial U}{\partial \tau} = -\frac{\partial U}{\partial x} V_{\text{mesh}}.
\]
Here, the velocity of the mesh, $V_{\text{mesh}}$, is defined in terms of the velocity of the nodal points as $\frac{dx}{dt}$, and moreover, it is numerically approximated by using a simple first order finite difference formula, i.e.,

$$V_{\text{mesh}} \approx \tilde{V}_{\text{mesh}} = \frac{x_{t+\Delta t} - x_t}{\Delta t}. \quad (4.3.3)$$

Now, following the traditional Galerkin FE formulation, a residual equation can be constructed as

$$R = \int \left[ \frac{\partial^2 U}{\partial x^2} + (\gamma - 1 + V_{\text{mesh}}) \frac{\partial U}{\partial x} - \gamma U - \gamma \right] \varphi dx = 0, \quad (4.3.4)$$

the solution of which is identical to the one of (4.2.2) in the weak sense. Here, $\varphi$ is the trial function. The residual $R$, after using the divergence theorem, becomes

$$R = \int \left[ - \frac{\partial U}{\partial x} \frac{\partial \varphi}{\partial x} + (\gamma - 1 + V_{\text{mesh}}) \frac{\partial U}{\partial x} \varphi - \gamma U \varphi - \gamma \varphi \right] dx. \quad (4.3.5)$$

One should notice that once $V_{\text{mesh}}$ is replaced by $\tilde{V}_{\text{mesh}}$ in (4.3.5), the solution of the new residual equation, denoted as $\tilde{R} = 0$, is not identical to the one of (4.3.5), because of the truncation error brought in by the numerical approximation of $V_{\text{mesh}}$. In other words, the numerical handling of $V_{\text{mesh}}$ will affect the accuracy of the final results. However, it should be remarked that this issue is totally different from the one related to the accuracy and applicability of the IFE approach, as the final numerical results can be further improved by adopting high order approximation for $V_{\text{mesh}}$. For simplicity, in the current work, we shall adopt the first order approximation $\tilde{V}_{\text{mesh}}$, as defined in (4.3.3), and the implementation with other approximation formulae should be similar.

The next step of the traditional FE method is to evaluate the so-called element-stiffness matrices and the element-loading vector through (4.3.4), and then assemble all of the element matrices to obtain a master stiffness matrix and a total applied-load vector. Finally, after imposing the constrained boundary conditions, we obtain
the matrix form of the discretized residual  $\tilde{R}$ as

$$\tilde{R} = K^*U - Q^*. \quad (4.3.6)$$

Here $U$ is the vector of the nodal values of the entire domain, and $K^*$ and $Q^*$ are respectively, the constrained master stiffness matrix and the constrained master column matrix of the total applied generalized nodal loading [100]. The details of derivation of (4.3.6) from (4.3.5) are provided in Appendix B.1.

It should be pointed out that in the traditional “direct” FE approach, $U$ is the unknown, and can be solved from (4.3.6) as zeros of a given system. However, the “direct” method breaks down once the computational domain has moving boundaries. This is because after imposing the constrained free boundary conditions, the constrained master matrices $K^*$ and $Q^*$ are both functions of the unknown boundaries, resulting in numerical difficulty in handling the algebraic system. It should be pointed out that such a highly-nonlinear system may also be solved by using some iteration methods, but the convergence and the efficiency will be a big problem, as a result of the two types of unknowns, the free boundary and the nodal values, being coupled together.

To avoid those difficulties, the “inverse” approach is proposed based on the concept of fixing the nodal values and studying the motion of different isotherms. In other words, in the “inverse” formulation, the location of each element is unknown (in our case, the $x$ coordinate), while the nodal values $U$ are kept as known constants. Such an easy decouple of the two types of unknowns makes (4.3.6) become a nonlinear system in which only the locations of the elements need to be solved. This system can be efficiently dealt with by using any iteration scheme, such as the full Newton method. Most importantly, the convergence of such an approach is theoretically guaranteed, which is the issue of the next section.

To obtain a reasonably accurate solution, the monotonicity of $U$ is required. Otherwise, the coordinate $x$ that satisfies (4.3.6) is not unique, resulting in difficulties in
deciding which is the correct location for a fixed nodal value, even if the convergence of the adopted iteration method is guaranteed. Fortunately, in our case, $U$ is strictly monotonically increasing w.r.t. $x$, as demonstrated earlier, and thus, no such problem needs to be further considered. It should be pointed out that the monotonicity of the option price w.r.t. the underlying is an inherent characteristic dictated by the financial nature of options themselves. In this sense, the application of the IFE approach to the field of financial engineering in terms of pricing options is almost perfect.

Another important issue that should be stressed here is the proper initial guess of the unknown nodal locations, since the Newton method 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 w.r.t. $\tau$, and the time step is sufficiently small enough.

Now, with the above two crucial points in mind, the algorithm of the proposed IFE approach can be summarized as:

i) The initial input $x_0$ is a vector starting from 0, ending at $x_{\text{max}}$, with its component strictly monotonically increasing from the left to right. Mathematically, $x_0$ can be written as $x_0 = [a_1 \cdots a_N]$, where $a_1 = 0$, $a_N = x_{\text{max}}$, and $a_{i-1} < a_i < a_{i+1}$ ($2 < i < N - 1$), with $N$ being the number of the nodal points in the whole computational domain.

ii) At the $k$th ($k \geq 1$) time step, set $x_k^{(0)} = x_{k-1}^{(*)}$, where $x_{k-1}^{(*)}$ is the final solution of the $(k - 1)$th time step. Supposing $x_k^{(n)}$ is obtained after the $n$th iteration, we compute the residual $R(x_k^{(n)})$ via (4.3.6), and the corresponding Jacobian matrix $J_R(x_k^{(n)})$. 
iii) Calculate the unknown nodal locations at the \((n + 1)\)th iteration step through

\[ x^{(n+1)}_k = x^{(n)}_k - J^{-1}_R R(x^{(n)}_k). \]

iv) Repeat (ii)-(iii) until \( \| x^{(n+1)}_k - x^{(n)}_k \| < \epsilon \) is satisfied. The tolerance \( \epsilon \) is set to \( 10^{-6} \) for all the results presented in this chapter. Set the solution of the \( k \)th time step to \( x^{(n+1)}_k = x^{(n+1)}_k \).

We remark that in the above algorithm, the location of the fixed boundary should be excluded, since it is already the solution of the corresponding nodal value, and no further iteration is needed. If the location of the fixed boundary was 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 would fail.

4.4 Convergence analysis

As pointed out previously, the proposed IFE approach deals with the nonlinear PDE system directly, and involves no linearization process at all. Therefore, it is worthwhile to analyze the convergence, because if this method is convergent, the numerical solution converging to the exact one of the original nonlinear PDE system will be guaranteed.

On the other hand, one can clearly observe that there are two major parts of the IFE approach. One is the traditional FE formulation based on the residuals, and the other is the Newton iterations adopted to solve the nonlinear algebraic system resulted from the residual statement. At each time step \( k \), an initial guess is given, and the solution that satisfies the original PDE system within a given tolerance is then worked out for this particular time step. Clearly, the convergence analysis should be specified on each time step \( k \). For convenience, in the following discussions, the subscript \( k \) which implies the \( k \)th time step will be omitted. Also, our analysis is
based on the linear shape functions, and the extensions to the other shape functions should be straightforward.

Lemma 4.4.1 For every square matrix $A$ and a given norm $\| \cdot \|$, 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 details of proof can be found in [98].

Theorem 1 If $\Delta t < \frac{\Delta x_{0,\min}}{2\gamma(2\gamma - 1)}$, 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 $\Delta x_{0,\min}$ is the smallest interval of the initial input vector at $\tau = 0$, and $\| \cdot \|$ is any proper matrix norm.

Proof. The Jacobian matrix at the exact solution $x^*$ can be written as:

for $i = 2 \cdots N - 1$

$$J_R(i - 1, i) = (\frac{1}{6\Delta t} + \frac{\gamma}{3} - \frac{1}{\Delta(x_i^*)^2})\Delta U_i - \frac{\gamma}{2}U(i) - \frac{\gamma}{2},$$

$$J_R(i, i) = (\frac{1}{3\Delta t} + \frac{\gamma}{6} + \frac{1}{\Delta(x_i^*)^2})\Delta U_i + (\frac{1}{3\Delta t} + \frac{\gamma}{6} + \frac{1}{\Delta(x_{i+1}^*)^2})\Delta U_{i+1},$$

$$J_R(i + 1, i) = (\frac{1}{6\Delta t} + \frac{\gamma}{3} - \frac{1}{\Delta(x_{i+1}^*)^2})\Delta U_{i+1} + \frac{\gamma}{2}U(i) + \frac{\gamma}{2},$$

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

and

$$J_R(1, 1) = (\frac{1}{3\Delta t} + \frac{\gamma}{6} + \frac{1}{\Delta(x_1^*)^2})\Delta U_2 + \frac{\gamma}{2}U(1) + \frac{\gamma}{2},$$

$$J_R(2, 1) = (\frac{1}{6\Delta t} + \frac{\gamma}{3} - \frac{1}{\Delta(x_2^*)^2})\Delta U_2 + \frac{\gamma}{2}U(1) + \frac{\gamma}{2},$$

$$J_R(N - 1, N) = (\frac{1}{6\Delta t} + \frac{\gamma}{3} - \frac{1}{\Delta(x_N^*)^2})\Delta U_N - \frac{\gamma}{2}U(N) - \frac{\gamma}{2},$$

$$J_R(N, N) = (\frac{1}{3\Delta t} + \frac{\gamma}{6} + \frac{1}{\Delta(x_N^*)^2})\Delta U_N + (\frac{1}{3\Delta t} + \frac{\gamma}{6} + \frac{1}{\Delta(x_{N+1}^*)^2})\Delta U_{N+1},$$
where $\Delta x^*_i = x^*(i) - x^*(i-1)$, and $\Delta U_i = U(i) - U(i-1)$. The details of the
derivation are provided in Appendix B.2.

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

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

$$J_R(i,i) - \sum_{j \neq i} |J_R(j,i)| = \left(\frac{1}{2\Delta t} + \frac{\gamma}{2}\right)\Delta U_i + \left(\frac{1}{2\Delta t} + \frac{\gamma}{2}\right)\Delta U_{i+1}.$$  

Since $U$ is monotonically increasing w.r.t. the index $i$, it is straightforward to show that for any $k = 2 \cdots N + 1$, $\Delta U_k > 0$, and thus

$$J_R(i,i) > \sum_{j \neq i} |J_R(j,i)|.$$  

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

$$J_R(i,i) - \sum_{j \neq i} |J_R(j,i)| = \left(\frac{1}{6\Delta t} - \frac{\gamma}{6} + \frac{2}{(\Delta x^*_i)^2}\right)\Delta U_i + \left(\frac{1}{6\Delta t} - \frac{\gamma}{6} + \frac{2}{(\Delta x^*_{i+1})^2}\right)\Delta U_{i+1},$$

which is greater than zero if $\Delta t < \frac{\Delta x_{0\text{ min}}}{2\gamma(2e - 1)} < \frac{1}{\gamma}$.

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

$$J_R(i,i) - \sum_{j \neq i} |J_R(j,i)| = \left(\frac{1}{6\Delta t} - \frac{\gamma}{6} + \frac{2}{(\Delta x^*_i)^2}\right)\Delta U_i + \left(\frac{1}{2\Delta t} + \frac{\gamma}{2}\right)\Delta U_{i+1} + \gamma + \gamma U(i).$$

Also, it is not difficult to check that if $\Delta t < \frac{\Delta x_{0\text{ min}}}{2\gamma(2e - 1)} < \frac{1}{\gamma}$, then $J_R(i,i) > \sum_{j \neq i} |J_R(j,i)|$ is satisfied.

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

$$J_R(i,i) - \sum_{j \neq i} |J_R(j,i)| = \left(\frac{1}{2\Delta t} + \frac{\gamma}{2}\right)\Delta U_i + \left(\frac{1}{6\Delta t} - \frac{\gamma}{6} + \frac{2}{(\Delta x^*_i)^2}\right)\Delta U_{i+1} - \gamma - \gamma U(i).$$  

(4.4.7)
By applying the mean-value theorem to the right hand side of (4.4.7), we obtain

\[ J_R(i, i) - \sum_{j \neq i} | J_R(j, i) | = \exp(x_0, \xi) \Delta x_0, i \left( \frac{1}{2 \Delta t} + \frac{\gamma}{2} \right) - \gamma \exp x_0, i \]

\[ + \left( \frac{1}{6 \Delta t} - \frac{\gamma}{6} + \frac{2}{(\Delta x_{i+1}^*)^2} \right) \Delta U_{i+1}, \]

\[ = \exp(x_0, \xi) \left[ \Delta x_0, i \left( \frac{1}{2 \Delta t} + \frac{\gamma}{2} \right) - \gamma \exp(x_0, i - x_0, \xi) \right] \]

\[ + \left( \frac{1}{6 \Delta t} - \frac{\gamma}{6} + \frac{2}{(\Delta x_{i+1}^*)^2} \right) \Delta U_{i+1}, \]

\[ > \exp(x_0, \xi) \left[ \Delta x_0, i \left( \frac{1}{2 \Delta t} + \frac{\gamma}{2} \right) - \gamma \exp(\Delta x_0, i) \right], \tag{4.4.8} \]

where \( x_0(i - 1) < x_0, \xi < x_0(i) \). On the other hand, when \( \Delta t < \frac{\Delta x_{0, \min}}{2 \gamma (2e - 1)} \) it is straightforward to show that

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

Therefore,

\[ J_R(i, i) - \sum_{j \neq i} | J_R(j, i) | > \exp(x_0, \xi) \left[ \Delta x_0, i \left( \frac{\gamma (2e - 1)}{\Delta x_{0, i}} + \frac{\gamma}{2} \right) - \gamma \exp(\Delta x_0, i) \right], \]

\[ > \exp(x_0, \xi) [\gamma (2e - 1) - \gamma], \]

\[ = 2 \exp(x_0, \xi) \gamma (e - 1) > 0, \]

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

Similarly, it can be easily verified that when \( \Delta t < \frac{\Delta x_{0, \min}}{2 \gamma (2e - 1)} \), \( J_R(1, 1) > \sum_{j \neq 1} | J_R(j, 1) | \) and \( J_R(N, N) > \sum_{j \neq N} | J_R(j, N) |. \)

Therefore, based the above four cases, it can be concluded that, if \( \Delta t < \frac{\Delta x_{0, \min}}{2 \gamma (2e - 1)} \), the Jacobian matrix \( J_R(x^*) \) is strictly diagonally dominant. According to the Gershgorin circle theorem [94] that any strictly diagonally dominant matrix is non-singular, it is clear that \( J_R(x^*) \) is then invertible. By applying Lemma 4.4.1, we obtain

\[ \| J_R^{-1}(x^*) \|^{-1} = \| J_R^*(x^*) \| > 0. \]
As a result, there exists a positive constant $M$, such that $\|J_R^{-1}(x^*)\|^{-1} \geq M > 0$, which yields $\|J_R^{-1}(x^*)\| \leq \frac{1}{M}$. Therefore, the norm of the inverse of the Jacobian matrix is bounded. This completes the proof.

**Lemma 4.4.2** If $F(X) = (F_1(X), F_2(X) \cdots F_N(X))$, 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 $\| \cdot \|$ are defined as two consistent matrix and vector norms.

**Proof.** Since $F_i(X)$ is continuously differentiable on $D$, we have, for $\forall X, Y \in D,$ and $i \in \{1, 2, \cdots N\}$

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

where $M_i$ is a positive constant. The details of the proof of this statement can be found in [102].

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

$$\sum_{i=1}^{N} \|F_i(X) - F_i(Y)\|^2 \leq \sum_{i=1}^{N} M_i \|X - Y\|^2 \leq N M_{\text{max}} \|X - Y\|^2. \quad (4.4.9)$$

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

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

which, combined with (4.4.9), yields

$$\|F(X) - F(Y)\|^2_F \leq N M_{\text{max}} \|X - Y\|^2,$$

and consequently,

$$\|F(X) - F(Y)\|_F \leq \sqrt{N M_{\text{max}}} \|X - Y\|_2.$$
On the other hand, it is clear that all the matrix norms are equivalent, and so are the vector norms [98]. Thus, there exist positive constants $C_1$, $C_2$, $C_3$ and $C_4$, such that

$$C_1 \| F(X) - F(Y) \| \leq \| F(X) - F(Y) \|_{r} \leq C_2 \| F(X) - F(Y) \|;$$

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

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 $M = \frac{\sqrt{NM_{\text{max}}} C_4}{C_1}$. This completes the proof.

**Lemma 4.4.3** Let $F = (f_1, f_2 \cdots f_m) : \mathbb{R}^n \rightarrow \mathbb{R}^m$, and suppose that the partial derivatives $\frac{\partial f_i}{\partial x_j}$, $1 \leq i \leq m$, $1 \leq j \leq n$ 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 referred to [102].

In fact, Lemma 4.4.3 gives a sufficient condition for a vector function $F$ being continuously differentiable at a given point $X_0$.

**Theorem 2** 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

$$R = \frac{\min(\Delta x_{i+1}, i = 0 \cdots N - 1)}{3}.$$

**Proof.** According to the explicit form of the Jacobian matrix (see Theorem 1), it
can be shown that for \(i, j = 1 \cdots N\), \(J_R(j, i)\) 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 4.4.3, it is clear that for \(i = 1 \cdots N\), the \(i\)th 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 \(X \in B(x^*, R)\) has that property. Here, \(\overline{B(x^*, R)}\) denotes the closed form of \(B(x^*, R)\). Note that \(\overline{B(x^*, R)}\) is compact as any finite closed set is compact.

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

\[
X(i) = x^*(i) + a_i, \quad X(i - 1) = x^*(i) + b_i, \quad X(i + 1) = x^*(i + 1) + c_i, \quad i = 1 \cdots N,
\]

where \(|a_i| \leq R\), \(|b_i| \leq R\), and \(|c_i| \leq R\). Thus

\[
|X(i) - X(i - 1)| = |x^*(i) - x^*(i - 1) + a_i - b_i|,
\]

\[
\geq |x^*(i) - x^*(i - 1)| - |a_i| - |b_i|,
\]

\[
\geq R.
\]

Similarly, we obtain \(|X(i) - X(i + 1)| \geq R\). On the other hand, since each nodal value has a unique location, and moreover, the nodal value is monotonically increasing w.r.t. the index \(i\), it is straightforward to show that \(R\) is always greater than zero.

As a result, for any \(X \in \overline{B(x^*, R)}\), each of the component is different from its adjacent ones, and consequently, \(J_R(\cdot, i)\) \((i = 1 \cdots N)\) is continuously differentiable at \(X\). Therefore, for any \(X, Y \in B(x^*, R) \subset \overline{B(x^*, R)}\), by applying Lemma 4.4.2, we obtain

\[
\|J_R(X) - J_R(Y)\| \leq M \|X - Y\|,
\]

where \(M\) is a positive constant.

**Theorem 3 (Convergence for the Newton iteration method)** Let \(F : \mathbb{R}^n \to \mathbb{R}^n\) be a \(C^1\) 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 \( \| \cdot \| \) 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 \( x^* \) with

\[ \| x^{(k+1)} - x^* \| \leq CL \| x^{(k)} - x^* \|^2. \]

Proof. The details of proof can be found in [86].

Theorem 3 confirms that the Newton method is quadratically convergent only if the initial guess \( x^{(0)} \) is sufficiently close to the exact solution \( x^* \), and the Jacobian matrix is nonsingular. One should notice that in this theorem, the requirement that the Jacobian matrix is nonsingular at \( x^* \) is in fact to guarantee that the Jacobian matrix is nonsingular at any point belonging to \( B(x^*, r) \), which is the range of the solution produced after each iteration. As a result, the Newton method will not terminate until a solution with a pre-specified error bound has been found.

Based on Theorems 1-3, it is now quite clear that the Newton method adopted at each time step is locally quadratically convergent. In fact, one can easily show that for our problem, the Newton method at each time step cannot have global convergence property. For example, it can be easily checked that the Jacobian matrix \( J_R(x^{(0)}) \) is not well defined if \( x^{(0)} \) satisfies \( x^{(0)}(i + 1) = x^{(0)}(i) = x^{(0)}(i - 1) \), resulting in the Newton method failing to continue. Therefore, it is obvious that the Newton method adopted at each time step is only locally convergent.

To ensure the local convergence, the initial guess and the invertibility of the Jacobian matrix are quite crucial, as presented in Theorem 3. Fortunately, these two requirements can be both met with a proper restriction on \( \Delta t \). It is known that the initial guess at the present time step is given by the solution of the previous time step, which can be sufficiently close to the exact solution of the current time step if \( \Delta t \) is small enough, as a result of \( x \) being continuous w.r.t. \( \tau \); while for the
non-singularity of the Jacobian matrix, it is invertible if \( \Delta t < \frac{\Delta x_{0,\min}}{2\gamma(2c - 1)} \), as shown in Theorem 1. Therefore, if \( \Delta t \) is chosen such that \( x^{(0)} \in B(x^*, r) \) and the Jacobian matrix at \( x^* \) is invertible, the local-quadratic convergence of the Newton method at each time step can be guaranteed. One should notice that the convergence analysis we provided ensures that the solution at each time step will converge to the one of the discretized residual equation system, i.e., \( \tilde{R} = 0 \). Theoretically, the error between the solutions of \( \tilde{R} = 0 \) and \( R = 0 \) can be arbitrary small, if sufficiently large number of time steps and elements are adopted. Such an analysis is similar to the standard convergence analysis of the FE method, which can be found in many books (e.g. [13]), and is thus omitted here.

Several remarks should be made before we show the numerical results in the next section. First of all, the restriction on \( \Delta t \) for the convergence of the Newton iteration approach adopted at each time step is not equivalent to stability criteria required for the convergence of certain types of finite difference schemes. In fact, the discussion of the convergence of any iteration scheme has nothing to do with a stability analysis at all. Secondly, as pointed out previously, to ensure the overall convergence of the IFE approach, the solution of \( \tilde{R} = 0 \) should be sufficiently close to the one of \( R = 0 \). Such a requirement will bring in the restrictions on both \( \Delta t \) and the size of each element, the former is because of a simple first order finite difference formula being used to numerically approximate \( V_{mesh} \), while the latter is a result of the FE formulation involved in the approximating process of \( R \).

### 4.5 Examples and discussions

Since the convergence has already been shown theoretically in the last section, there is no need, in theory, to further discuss the accuracy of the IFE approach, as a convergent scheme will produce a result with arbitrary accuracy, if the length of the time step and the size of each element are sufficiently small. However, from the viewpoint that a comparison with previously published results may give readers a
sense of verification of the current approach, two numerical examples, which cover medium-term and long-term options, are still provided in this section.

To help readers who may not be used to discussing financial problems with dimensionless quantities, all results, unless otherwise stated, are now converted back to dimensional quantities in this section before they are graphed and presented.

**Example 1**
The first example is a sample case discussed in [16, 112, 115, 125]. The parameters are: strike price $K = $100, risk-free interest rate $r = 0.1$, volatility $\sigma = 0.3$, and the time to expiration $T - t = 1$ (year).

To test the reliability of the IFE approach, the best way is to calculate the option and the optimal exercise prices, and compare them with Zhu’s analytical solution [115]. However, it is better for us to focus on the comparison of the optimal exercise price only, because the optimal exercise price is far more difficult to be accurately calculated than the option price, and once the $S_f$ is determined accurately, the calculation of the option price itself is straightforward [125]. Such a comparison is shown in Fig 4.1, where the optimal exercise prices are displayed by using the IFE approach and Zhu’s analytical method [115], respectively. In this figure, two different shape functions are adopted, the linear ones and the quadratic ones, for the FE formulation involved in the current approach. Moreover, in both cases, the time step is set to $\Delta t = 0.01$, and one hundred ($N = 100$) equal-size elements along the asset direction are adopted. As can be clearly seen from this figure, the numerical results and the analytical solution appear to agree well with each other. At the expiration date, $t = T = 1$ (year), the optimal exercise price calculated by Zhu’s analytical method is $S_f(T) = $76.11, whereas they are $S_f(T) = $76.19 and $S_f(T) = $76.17, when linear and quadratic shape functions are adopted, respectively. A close examination reveals that the current approach slightly underestimates the $S_f$ values when the time is close to expiry. This is not a surprising issue due to the presence of the well-known singularity at expiry [33], which is not possible for most of the numerical schemes to deal with. However, it can be observed that the under-
estimation is considerably reduced once the quadratic shape functions are adopted. Thus, it is reasonable to infer that the proposed numerical scheme is able to well handle the singularity at expiry if higher order interpolation functions are used.

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

\[
RMSRE = \sqrt{\frac{1}{I} \sum_{i=1}^{I} \left( \frac{\tilde{a}_i - a_i}{a_i} \right)^2},
\]

where \(a_i\)s are the nodal optimal exercise prices associated with Zhu’s analytical solution, \(\tilde{a}_i\)s are the corresponding numerical values, and \(I\) is the number of sample points used in the calculation of RMSRE. In our experiment, we set \(I\) to 51 for all the following results. With RMSRE, the overall difference between the numerical solution and the exact solution can be clearly demonstrated.

Table 4.1 displays the variations of RMSREs for the optimal exercise prices when the numbers of time steps and equal-sized elements are gradually increased. In Table 4.1, \(M\) stands for the number of time intervals while \(N\) means that \(N\) equal-sized
elements are adopted. From this table, one can clearly observe that the RMSRE becomes lager, as the size of the time step, i.e., \( \frac{T}{M+1} \), becomes larger. This is because large size of time step may not only worsen the convergence condition for the Newton iteration adopted at each time step, but also result in large discretization errors between the solution obtained from the residual statement with \( \bar{R} = 0 \) and that with \( R = 0 \), which corresponds to the original residual statement without \( \tilde{V}_{\text{mesh}} \) being replaced with \( V_{\text{mesh}} \) (cf. (4.3.3)-(4.3.5)). On the other hand, large RMSRE can also be observed once few elements are used, which is simply as a result of the finite element discretization acting on the residual statement \( \bar{R} = 0 \).

Table 4.1: The variations of RMSREs when the numbers of the time steps or equal-size elements are gradually increased. \( M \): the number of time intervals; \( N \): the number of elements.

<table>
<thead>
<tr>
<th>( M )</th>
<th>( N = 10 )</th>
<th>( N = 20 )</th>
<th>( N = 40 )</th>
<th>( N = 80 )</th>
<th>( N = 100 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>7.69%</td>
<td>5.30%</td>
<td>3.69%</td>
<td>2.66%</td>
<td>2.42%</td>
</tr>
<tr>
<td>10</td>
<td>7.26%</td>
<td>4.77%</td>
<td>3.09%</td>
<td>2.01%</td>
<td>1.76%</td>
</tr>
<tr>
<td>25</td>
<td>6.98%</td>
<td>4.40%</td>
<td>2.66%</td>
<td>1.54%</td>
<td>1.28%</td>
</tr>
<tr>
<td>50</td>
<td>6.88%</td>
<td>4.27%</td>
<td>2.51%</td>
<td>1.37%</td>
<td>1.10%</td>
</tr>
</tbody>
</table>

One should notice that the numerical simulations showed above are all based on equal-size time steps and elements. In fact, the current approach is quite flexible in changing the size of each element as well as the length of the time step, an advantage that we can take to further improve the accuracy and the efficiency of the IFE approach in solving option prices. We shall discuss this issue through the following experiments.

As is now well known, there are two singularities associated with the valuation of American put option. One is near the optimal exercise price \( S_f \), at which the option price is not second order differentiable w.r.t. the underlying \( S \), although the first order derivative is continuous [58]; the other is that if the underlying dividend is less than the market interest rate, the optimal exercise price has a parabolic-logarithmic behavior when the time is close to expiry [33]. To increase the accuracy of the numerical results, a natural way is to place more elements in the neighborhood of \( S = S_f \) (x = 0), and to choose finer time steps as time approaches to the expiration
date. The variations of RMSREs of the solutions with different-size elements and
time steps are provided in Table 4.2. In this table, \( x_0^i \) \((i = 1, 2, 3)\) stands for the
initialization of the elements, and each component of \( t^s \) \((s = 1, 2)\) denotes the length
of the time step for each iteration. The specific definitions of \( x_0^i \) \((i = 1, 2, 3)\) and \( t^s \)
\((s = 1, 2)\) are provided in Appendix B.3. Moreover, a total number of \( N \) elements
and \( M + 1 \) time steps are used for the calculation.

Table 4.2: The variations of RMSREs with different-size elements and time steps. \( M \): the
number of time intervals; \( N \): the number of elements.

<table>
<thead>
<tr>
<th>( (x_0^i, t^s) )</th>
<th>( M = 10, N = 10 )</th>
<th>( M = 10, N = 25 )</th>
<th>( M = 25, N = 10 )</th>
<th>( M = 25, N = 25 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( (x_0^1, t^1) )</td>
<td>7.26%</td>
<td>4.15%</td>
<td>6.98%</td>
<td>3.76%</td>
</tr>
<tr>
<td>( (x_0^2, t^1) )</td>
<td>7.26%</td>
<td>4.15%</td>
<td>6.98%</td>
<td>3.76%</td>
</tr>
<tr>
<td>( (x_0^3, t^1) )</td>
<td>1.48%</td>
<td>1.40%</td>
<td>0.98%</td>
<td>0.9%</td>
</tr>
<tr>
<td>( (x_0^2, t^2) )</td>
<td>0.96%</td>
<td>0.8%</td>
<td>0.56%</td>
<td>0.39%</td>
</tr>
</tbody>
</table>

From this table, one can clearly observe that the accuracy of the IFE approach
has significantly increased with the RMSREs dramatically decreasing to less than
1%, once more elements are placed in the neighborhood of \( x = 0 \) to resolve the
sharp changes of the first order derivative of the option price near \( S = S_f \). However,
the changes of time-step lengths near expiry seem to have very little impact on the
accuracy of the current approach. It can also be seen that the RMSREs decrease as
the number of either the elements or the time steps increases, once the initializations
of the sizes of each element and time step are fixed.

Another important issue in the finance industry is the speed of calculation, which
is equally important, sometimes even more, than the accuracy. To clearly demon-
strate the relationship between the accuracy and efficiency of the IFE approach, in
Fig 4.2, we display the variations of RMSREs as a function of total CPU time used
in executing the code for each run. All the experiments here were performed within
Matlab7.5 on a Intel Pentium 4, 3GHZ machine. In this figure, we adopt three sets of
CPU times as a function of the RMSREs for three fixed numbers of time steps, i.e.,
with \( M \) being 10, 25, 50, respectively. Moreover, different-size elements are adopted,
and the initializations for them are provided in the vector \( x_0^3 \).
As is clearly shown in Fig 4.2, the accuracy is inversely varying with the efficiency, as expected. Furthermore, a fine resolution in the time direction implies more computational time. A great feature of the current scheme is that a high computational efficiency can still be achieved while a satisfactory accuracy can be maintained. For example, it took less than 5 seconds to produce a result with RMSRE being less than 1%. This level of accuracy and efficiency would certainly suit the practical needs of market practitioners.

**Example 2**

The second test that a good numerical scheme must pass is that the optimal exercise price should asymptotically approach to that of its corresponding perpetual counterpart, when the lifespan of the put option becomes infinite. In the literature, it was reported that some approaches lead to an oscillatory and non-monotonic optimal exercise price when the lifespan of the put option contract is extremely long [9]. To verify that the IFE approach is not restricted by the lifespan of an option in any way, a long-tenor put option is designed in this example, with the following dimensional parameters: strike price $K = \$100$, risk-free interest rate $r = 0.01$, volatility $\sigma = 0.2$, time to expiration $T - t = 50$ (year).
Figure 4.3: Optimal exercise price for the second case.

Depicted in Fig 4.3 is the optimal exercise price as a function of the time to expiration up to 50 years against the theoretical perpetual optimal exercise price calculated from the well-known Merton’s formula, i.e., \( S_f(\infty) = \frac{\gamma K}{1 + \gamma} \). From this figure, one can observe that our numerical solution exhibits a nice asymptotic approach to the theoretical optimal exercise price of the corresponding perpetual put option. No oscillation is observed at all. This confirms that our scheme is very stable and can be used for long-tenor options as well.

4.6 Conclusion

In this chapter, we have proposed and tested an IFE approach for the numerical valuation of American put options. The key feature of the current scheme, in comparison with some numerical methods in the literature [112, 125], is that our scheme does not need any linearization process at all, but still with high efficiency as a result of using the full Newton iteration method with its inherent quadratic convergence. Through a couple of numerical examples, the high accuracy and efficiency of the current scheme are demonstrated. Of course, the most important contribution of this
paper is the convergence analysis we have provided for the IFE approach for the first time, to guarantee that the numerical solution does indeed converge to the exact American put option prices.
Part II: Analytical approximations

Once analytic expressions for the option prices are not available, approximation methods, as another alternative to the numerical approaches, were also widely sought in the last decade or two. Unlike the numerical methods, various analytical approximations are made in order to reduce the intensity of the final computation and speed up the computation, while still maintain a reasonable degree of accuracy. In this part, we derive a series of analytic approximation formulae for the pricing of vanilla options, which will be particularly useful for market practitioners whose main concern in trading is the computation speed rather than the on-dot accuracy. This part is further organized into five chapters, namely, Chapter 5-Chapter 10, according to five relevant issues that need to be considered. In Chapter 5, we derive a new approximation for the valuation of European puts under the Heston model, which is as fast to implement as the Black-Scholes formula. In Chapter 6 and Chapter 7, we investigate the near-expiry asymptotic behavior of the optimal exercise prices associated with American puts under the local volatility model and under the Heston model, respectively. These approximations can be combined with the numerical approaches to calculate the option prices and the optimal exercise prices away from expiry. In Chapter 8-Chapter 10, we consider the pricing approximations of perpetual American puts under three different volatility dynamics, namely, the slow-mean reverting volatility, the fast mean-reverting volatility, and the multiscale volatility. We also discuss the effects of different volatility dynamics on the early exercise strategies of the American puts.
Chapter 5

A new analytical approximation for European puts with stochastic volatility

5.1 Introduction

It is a well-known fact that the BS (Black-Scholes) model is inadequate to describe the asset returns and the behavior of the option markets. The literature advocates the introduction of stochastic volatility to reproduce the implied volatility smile observed in markets. Among those stochastic volatility models (e.g., [6, 38, 51, 58]), the model proposed by Heston [51] has received the most attention and we shall use this model throughout the chapter. The introduction of a second stochastic factor has considerably complicated the solution process in pricing option derivatives. It leads to a bivariate PDE (partial differential equation), which is of enormous interest to financiers and mathematicians alike. Moreover, it has resulted in many diverse mathematical and computational techniques being employed in the valuation of option derivatives.

Singular perturbation theory, which is a branch of applied mathematics, has been widely used for the option pricing problems. For example, Whalley et al. [107–109]
adopted this method to study the transaction costs. Kuske and Keller [33] also used this technique to derive an explicit analytical expression for the optimal exercise boundary near expiry under the BS model. Martin et al. [110] extended this method to the valuation of a class of option derivatives under the BS model. Obviously, the use of this method has led to a significant simplification of the problems mentioned in [33, 107–110], by reducing the number of parameters. The purpose of this chapter is to derive an approximation for European puts under the Heston model by using singular perturbation theory.

Although an analytical formula has already been found for European puts under the Heston model [51], it has unfortunately, displayed some undesired characteristics. For example, the Heston formula still requires the valuation of logarithm with a complex argument involved in the Fourier inverse transform. The main disadvantage of a solution being left in such a form lies in the fact that the numerical calculation of these integrals needs to be handled very carefully, as the integrand is multi-valued, which may sometimes give rise to numerical instability [66]. Furthermore, it is also suggested that the numerical calculation of Heston’s formula could be made extremely complicated by the fact that the integrand is typically of oscillatory nature (e.g., [18, 66]). In the literature, some authors focused entirely on the numerical implementation of Heston’s formula [18, 66]. It must be pointed out that all these methods still require a certain degree of computation. The aim of making an approximation as presented in this chapter, is to reduce the intensity of the final numerical computation, while maintaining a reasonable degree of accuracy for a large class of traded options.

In this chapter, a new approximation for the European puts with stochastic volatility is presented. The formula is found through solving Heston’s PDE, with the utilization of singular perturbation method. In comparison with the complicated numerical calculation of Heston’s analytical formula, the newly-obtained approximation is rather simple and elegant, since it only involves the standard normal distribution function, which is regarded as a built-in function in most software, such as Maple
Chapter 5.

[88], Matlab [83] and so on. On the other hand, our numerical experiments suggest that the new approximation is suitable for pricing options with relatively short tenor, which is a common characteristic of most options in the traded markets. For instance, the front month option, which expires in less than four weeks, attracts the most volume [47]. Furthermore, short-tenor options, which last for only one or two months, dominate the equity option markets [26]. Therefore, our approximation has a wide range of applications in the option markets.

The chapter is organized as follows: In Section 5.2, we present the details of applying singular perturbation technique to the valuation of European puts under the Heston model. In Section 5.3, some numerical examples are presented to illustrate the accuracy and the advantage of the current approximation.

5.2 Heston model and European put options with short tenor

In standard Heston model, the price of a European put $P$ can be modeled by the following PDE system:

\[
\begin{align*}
\frac{1}{2}vS^2\frac{\partial^2 P}{\partial S^2} + \rho \sigma v S \frac{\partial P}{\partial S}\partial v + \frac{1}{2}\sigma^2 v^2 \frac{\partial^2 P}{\partial v^2} + rS \frac{\partial P}{\partial S} + \kappa (\eta - v) \frac{\partial P}{\partial v} - rP - \frac{\partial P}{\partial \tau} &= 0, \\
\text{subject to:} & \\
P(S, v, 0) &= \max(K - S, 0), \\
\lim_{S \to \infty} P(S, v, \tau) &= 0, \quad \lim_{S \to 0} P(S, v, \tau) = Ke^{-r\tau}, \\
\lim_{v \to 0} P(S, v, \tau) &= \max(Ke^{-r\tau} - S, 0), \quad \lim_{v \to \infty} \frac{\partial P}{\partial v}(S, v, \tau) = 0,
\end{align*}
\]

(5.2.1)

with $S$ being the underlying, $v$ being the variance and $\tau$ being the time to expiry. Moreover, $r$, $\kappa$, $\eta$, $\rho$, $\sigma$ and $K$ denote the risk-free interest rate, the relaxation rate of the variance process, the mean value of the variance process, the correlation factor, the volatility of the volatility and the strike price, respectively. The boundary conditions along the $v$ direction, which are based on the riskless growing argument
(as \( v \to 0 \)) and market observations (as \( v \to \infty \)), are discussed thoroughly by Zhu and Chen [119]. The original Heston boundary conditions are rather complicated, especially for the case as \( v \to 0 \), and it is still controversial whether or not Heston’s analytical formula [51] does indeed satisfy the given boundary conditions along the \( v \) direction.

One of the reasons that the Heston model is much more popular than other stochastic volatility models is that, under the Heston model, the closed-form analytical solution has already been found [51], i.e.,

\[
P(S, v, \tau) = K \exp(-r\tau) - S + SP_1(\log S, v, \tau; \log K) \\
- K \exp(-r\tau)P_2(\log S, v, \tau; \log K),
\]

(5.2.2)

where \( u_1 = 1/2, \ v_2 = -1/2, \ b_1 = \kappa - \sigma \rho, \ b_2 = \kappa, \ x = \log S, \ a = \kappa \eta, \)

\[
d_j = \sqrt{(\rho \sigma \phi_i - b_j)^2 - \sigma^2(2u_j \phi_i - \phi^2)}, \quad g_j = \frac{b_j - \rho \sigma \phi_i + d_j}{b_j - \rho \sigma \phi_i - d_j},
\]

\[
D_j = \frac{b_j - \rho \sigma \phi_i + d_j}{\sigma^2} \left( \frac{1 - e^{d_j \tau}}{1 - g_j e^{d_j \tau}} \right),
\]

\[
C_j = r \phi \sigma i + \frac{a}{\sigma^2} [(b_j - \rho \sigma \phi_i + d)\tau - 2 \log(\frac{1 - g_j e^{d_j \tau}}{1 - e^{d_j \tau}})],
\]

\[
f_j = e^{C_j + D_j v + i \phi x}, \quad P_j = \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \mathbb{R} \{ \frac{e^{-i\phi y} f_j(x, v, \tau, \phi)}{i\phi} \} d\phi.
\]

The functions \( P_j(x, v, \tau; y) \) are the cumulative distribution functions (in the variable of \( y \)) of the log-spot price after time \( \tau \), starting at \( x \) for some drift \( \mu \). It should be remarked that Heston’s solution is already in a closed form, since there is only an explicit integral left to be calculated, the same as the calculation of the cumulative distribution function required in BS’ formula. However, the sharp difference between the two is that the integrand of the latter is a well-defined and smooth real function whereas the integrand of the former is a complex-valued function, as a result of the Fourier inverse transform not being analytical performed. Consequently, the numerical implementation of Heston’s formula may not be as straightforward as numerically carrying out a quadrature, as they are multi-valued, which may cause numerical in-
stability when one decides which root is the correct one to take. In order to avoid the complicated numerical calculations, it is desirable to make an approximate solution, which is fast and easy to implement.

For small $\tau$, an approximation of the above PDE system can be derived as follows. We set $\tau = \theta T$, where $T = \mathcal{O}(1)$, and $\theta$ is a small parameter. By assuming that the solution can be expanded in powers of $\theta$, i.e., $P(S,v,T) = \sum_{n=0}^{\infty} \theta^n P_n(S,v,T)$, and substituting this into the PDE system (5.2.1), we obtain the solutions of the $\mathcal{O}(1)$ and the $\mathcal{O}(\theta)$ differential systems as $P_0(S,v,T) = \max(K - S, 0)$, and $P_1(S,v,T) = -rKT$ for $S < K - \mathcal{O}(\theta)$, $0$ for $S > K + \mathcal{O}(\theta)$, respectively. It can be observed that the first order solution $P_0$ is continuous but not differentiable. Thus, we expect that there is a corner layer, or derivative layer, at $S = K$. The difference between the corner layer and other common boundary layers which we usually encountered in mathematical problems, lies in the fact that in the so-called corner layer, the rapid changes will occur in the slope, or derivatives of the solution, and not in the value of the solution itself. In our problem, the outer solution is invalid near the strike price. This is because $\frac{\partial^2 P_0}{\partial S^2} = \delta(K - S)$, and hence the second-order derivative with respect to $S$ cannot be ignored when deriving the $\mathcal{O}(1)$ differential system as $S$ approaches $K$.

Now, we perform a local analysis in the vicinity of $S = K$ by introducing the stretched variable: $X = (S - K)/(\theta^\alpha K)$, where $\alpha$ needs to be determined such that the corresponding operator as $\theta \to 0$ contains as much information as possible. Substituting this stretched variable $X$ into the governing equation contained in (5.2.1), and letting $\tilde{Y}$ denote the solution in this region, we obtain:

$$\frac{\partial \tilde{Y}}{\partial T} = \frac{1}{2} v(\theta^n X + 1)^2 \theta^{1-2\alpha} \frac{\partial^2 \tilde{Y}}{\partial X^2} + r(\theta^n X + 1) \theta^{1-\alpha} \frac{\partial \tilde{Y}}{\partial X} - r \theta \tilde{Y} + \rho \sigma v \theta^{1-\alpha}(\theta^n X + 1) \frac{\partial^2 \tilde{Y}}{\partial X \partial v} + \frac{1}{2} \sigma^2 v \theta \frac{\partial^2 \tilde{Y}}{\partial v^2} + \kappa (\eta - v) \theta \frac{\partial \tilde{Y}}{\partial v}. \quad (5.2.3)$$

It can be observed that $\alpha = 1/2$ is a well-balanced choice. This is because the
Chapter 5.

Degenerations of the corresponding operator of (5.2.3) with other choices of $\alpha$ are contained in the degeneration obtained on choosing $\alpha = 1/2$.

With the new variable $X$, the boundary conditions, which would be concerned, all have a factor $\sqrt{\theta}K$ in common. Thus, it is quite reasonable to rescale the problem by defining $Y = \sqrt{\theta}K\hat{Y}$, and expand $Y(X, v, T)$ in terms of $\sqrt{\theta}$,

$$Y(X, v, T) = Y_0(X, v, T) + \sqrt{\theta}Y_1(X, v, T) + \mathcal{O}(\theta). \quad (5.2.4)$$

Upon collecting together those $\mathcal{O}(1)$ terms, the initial-boundary value problem for $Y_0$ is

$$\begin{cases}
\frac{\partial Y_0}{\partial T} = \frac{1}{2}v\frac{\partial^2 Y_0}{\partial X^2}, \\
Y_0(X, v, 0) = \max(-X, 0), \\
\lim_{X \to +\infty} Y_0(X, v, T) \sim 0, \quad \lim_{X \to -\infty} Y_0(X, v, T) \sim -X.
\end{cases} \quad (5.2.5)$$

The solution of (5.2.5) can be found by using similarity solution technique. It is:

$$Y_0(X, v, T) = \left(\sqrt{vT}/\sqrt{2\pi}\right) \exp\left(-X^2/(2vT)\right) - (X/2)\text{erfc}(X/\sqrt{2vT}).$$

One should notice that the conditions at $X = \pm\infty$ in (5.2.5) are consistent with the payoff function. They can also be obtained by matching with the outer solution using the Van Dyke rule [105]. This is accomplished by writing the outer solution $P_0(S, v, T)$ in terms of the inner variable $X = (S - K)/(\sqrt{\theta}K)$, and expanding $P_0(X, v, T)$ to the order of $\mathcal{O}(\sqrt{\theta})$. The coefficients in front of $\sqrt{\theta}$ are the boundary values that are needed.

Similarly, the initial-boundary value problem for $Y_1$ is:

$$\begin{cases}
\frac{\partial Y_1}{\partial T} = \frac{1}{2}v\frac{\partial^2 Y_1}{\partial X^2} + Xv\frac{\partial^2 Y_0}{\partial X^2} + r\frac{\partial Y_0}{\partial X} + \rho\sigma v\frac{\partial^2 Y_0}{\partial X \partial v}, \\
Y_1(X, v, 0) = 0, \\
\lim_{X \to +\infty} Y_1(X, v, T) \sim 0, \quad \lim_{X \to -\infty} Y_1(X, v, T) \sim -rT.
\end{cases} \quad (5.2.6)$$
It is not hard to find the solution of (5.2.6) by using the following lemma [54]:

**Lemma 5.2.1**  
(i) If $u - \frac{1}{2}u_{xx} = 0$ and $v - \frac{1}{2}v_{xx} = u$, then a particular solution is $v = \tau u$.  
(ii) If $u - \frac{1}{2}u_{xx} = 0$ and $v - \frac{1}{2}v_{xx} = xu$, then a particular solution is $v = x\tau u + \frac{1}{2}\tau^2 u_x$.

The proof is rather trivial, and is thus omitted here. Hence,

\[
Y_1 = \frac{(\sqrt{2vT}X)}{(4\sqrt{\pi})}\exp(-X^2/(2vT)) - (rT/2)erfc(X/\sqrt{2vT}) \\
- (\sqrt{2T}\rho\sigma X)/(4\sqrt{\pi v})\exp(-X^2/(2vT)).
\]

Theoretically speaking, in the analysis given above, we must exclude a neighborhood of the boundary points $(S, v) = (K, 0)$ and $(S, v) = (K, \infty)$, where a transition between different kinds of layers may occur. This is the reason why $v$ becomes a parameter rather than a variable in the initial-boundary value problems (5.2.5-5.2.6), and the boundary conditions along the $v$ direction are no longer needed. On the other hand, it is straightforward to show that $Y_0$ and $Y_1$ satisfy the omitted boundary conditions along the $v$ direction.

It must be pointed out here that (5.2.4) is valid in the inner region, where $S \in \mathcal{O}(K, \sqrt{\theta})$, while in the outer region, we have the outer expansion found earlier. We will now combine the two pieces of solutions to form a composite expansion. This is done by adding the expansions and then subtracting the part that is common to both. Note that by writing (5.2.4) in terms of the outer variable $S$, it equals to the outer expansion as $\theta \to 0$. In other words, the outer expansion and the common part coincide. Therefore, the inner expansion is in fact uniformly valid. Written in terms of the original variables, the new approximation formula that can be used for all $S$ and small $\tau$ is:

\[
P(S, v, \tau) = \left[ S\left(\frac{\sqrt{2vT}}{\sqrt{\pi}} - \frac{\sqrt{2\tau\rho\sigma}}{4\sqrt{\pi v}}\right) + K\left(\frac{\sqrt{2vT}}{\sqrt{\pi}} + \frac{\sqrt{2\tau\rho\sigma}}{4\sqrt{\pi v}}\right) \right] e^{-\frac{(S-K)^2}{2v\tau K^2}} \\
- (S + (r\tau - 1)K)N\left(\frac{K-S}{K\sqrt{v\tau}}\right) + o(\tau),
\]

(5.2.7)
where \( N(\cdot) \) is the standard normal distribution function.

Obviously, our newly-obtained formula is much simpler than the original Heston’s analytical solution, that is, it can be written in terms of the standard normal distribution function, and thus is as fast and easy to implement as the BS formula, whereas the latter still requires rather complicated numerical calculation. However, just as pointed out previously, the fundamental assumption of the derivation of (5.2.7) is that, the time to expiration \( \tau \) should be very small. Therefore, in the next section, we shall show numerically how our new formula performs better than the existing Heston analytical formula, by finding the range of \( \tau \) for which (5.2.7) is valid.

5.3 Numerical tests and discussions

One may argue that there is no need to find a new approximation since Heston’s formula is already in a closed form. However, several studies (e.g., [18, 66]) suggest that the numerical implementation of Heston’s formula may not be as straightforward as numerically carrying out a quadrature because the integrand involved is not only typically of oscillatory nature, but also a complex-valued function, as a result of the Fourier inverse transform not being analytically performed. In comparison with the complicated integral calculations, the advantage of our new formula is obvious, that is, it only requires to calculate the standard normal distribution function, which is regarded as a built-in function in most software, such as Maple [88], Matlab [83] and so on. In this section, we shall present some numerical results, which not only manifest the advantage of the new approximation, but also give a guideline on the application of this new formula to the option markets.

The first experiment is the same as used in [104]. The parameters used by them are: \( \kappa = 2.5, \eta = 0.16, \sigma = 0.45, \ r = 0.1, \ \rho = 0.1, \ K = \$10.0 \). We calculate the option price through the new approximation, and compare it with the analytical solution. Depicted in Fig 5.1 are the option prices as a function of the underlying \( S \) with different values of \( \tau \). It can be clearly seen that when \( \tau \) is reasonably small, the two results agree well with each other, and when \( \tau \) becomes larger, the accuracy
deteriorates gradually. However, it must be remarked that when $\tau$ is very small, i.e., $\tau = 0.0001$ (year), $\tau = 0.001$ (year), $\tau = 0.01$ (year), the Heston analytical formula cannot produce an accurate solution. This is because for very short maturities, the put value approaches its non-smooth intrinsic value causing the integrand in the Fourier inversion to be highly oscillatory [18].

![Figure 5.1: European put option price with different times to expiration.](image)

On the other hand, it should be remarked again that the fundamental small parameter in the present work is related to the time to expiration $\tau$. As a result, the higher accuracy of the new formula can be achieved with restriction on small value of $\tau$. Fortunately, empirical evidence suggests that most options in the traded markets have this characteristic. For example, the front month options, which expire in less than four weeks, attract the most volume [47]. Almost all stock and cash-settled index options expire after the close on the third Friday of the month, and thus belong to this category [47]. Furthermore, short-tenor options, which expire in one or two months, dominate the equity option markets [26]. Therefore, it is quite reasonable to expect $\tau$ to be small. At this stage, however, it is still ambiguous to answer the question of how “small” $\tau$ should be, and how our new approximation can be applied to the option markets. We shall discuss these issues through the next experiment.

Table 5.1 shows the relationship between the time to expiry $\tau$ and the accuracy of the new formula with different ranges of $r$ and $\sigma$. Note that the accuracy here is measured by the relative error, which is defined as: $Error = \frac{\| P - \tilde{P} \|_\infty}{\| \tilde{P} \|_\infty}$, where
\( P \) and \( \tilde{P} \) denote the values that are computed from (5.2.7) and (5.2.2), respectively, and \( \| \cdot \|_\infty \) denotes the infinity norm. On the other hand, for other parameters that are needed in the numerical implementation, we chose the values that were estimated daily in the June 1988 to May 1991 period [5], i.e., \( \kappa = 1.15(0.03) \), \( \eta = 0.0348(0.00) \), \( \rho = -0.64(0.01) \), where the daily averages of the estimated parameters are reported first, followed by the standard errors in parentheses.

<table>
<thead>
<tr>
<th>Table 5.1: Relative errors of the new approximation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>The risk-free interest rate ( r )</strong></td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>( \tau = 0.08 ) year (one month)</td>
</tr>
<tr>
<td>( r = 1% )</td>
</tr>
<tr>
<td>( r = 5% )</td>
</tr>
<tr>
<td>( r = 10% )</td>
</tr>
<tr>
<td>( r = 20% )</td>
</tr>
<tr>
<td>( r = 30% )</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>( \tau = 0.17 ) year (two months)</td>
</tr>
<tr>
<td>( r = 1% )</td>
</tr>
<tr>
<td>( r = 5% )</td>
</tr>
<tr>
<td>( r = 10% )</td>
</tr>
<tr>
<td>( r = 20% )</td>
</tr>
<tr>
<td>( r = 30% )</td>
</tr>
<tr>
<td>-----------------------------------------------</td>
</tr>
<tr>
<td>( \tau = 0.25 ) year (three months)</td>
</tr>
<tr>
<td>( r = 1% )</td>
</tr>
<tr>
<td>( r = 5% )</td>
</tr>
<tr>
<td>( r = 10% )</td>
</tr>
<tr>
<td>( r = 20% )</td>
</tr>
<tr>
<td>( r = 30% )</td>
</tr>
</tbody>
</table>

Furthermore, Table 5.1 also reveals some important “rules of thumb” for our new formula to be used in finance practice. First of all, the new approximation formula is insensitive to the change of \( \sigma \). We have deliberately tabulated two columns of \( \sigma \) values, demonstrating that when the interest rate as well as other parameters are held constant, the relative errors of adopting the new formula do not change very much when \( \sigma \) belongs to two very distinct ranges of \([0.1\%, 45\%]\) and \([45\%, 80\%]\), respectively. In fact, the relative error has only changed by about 1% when \( \sigma \) is varied from 0.1% to 80%, which is extremely large already as even during the recent world financial crisis, the estimated \( \sigma \) is no more than 70% (cf. [124]). This implies that one does not need to worry too much about the \( \sigma \) values when one uses this formula for the calculation of option price. On the other hand, one can clearly observe that the accuracy of the current formula inversely varies with respect to the risk-free interest rate \( r \). This means that one does need to be careful with the new formula when the
interest rate is high. However, with a quite high interest rate of 10% (e.g., at the moment, Iceland has the highest interest rate of 9.5% among developed countries.), the relative error would be only slightly higher than 1%, if the formula is used to price an option expiring in one month. For a 5% interest rate, the relative error falls below 1% for option contacts of a lifetime of one month and is slightly above 2% for option contacts of a lifetime of two months. Of course, if $r$ is very small, say around 1%, which is roughly the current interest rate in some largest economies such as the United States, Japan, Canada and most of the European countries [101], our approximation can be adopted for options expiring in even more than three months. In short, the present approximation is a good tool for traders to use in pricing short-tenor option derivatives, such as all of the front month options, most of the equity options and the index options.

5.4 Conclusion

In this chapter, we have derived a new approximation for European puts under the Heston model by using singular perturbation method. The formula only involves the standard normal distribution function, and is thus as fast and easy to implement as the BS formula, whereas the Heston analytical solution needs to be handled very carefully, as the integrand involved in his formula is multi-valued, which may result in numerical instability when one decides which root is the correct one to take. On the other hand, numerical experiments suggest that, with an acceptable degree of accuracy, this new approximation is sufficient in pricing a large class of short-tenor option derivatives with maturities of no more than three months, such as all of the front month options, most of the equity options and the index options. Therefore, our new approximation is indeed a remarkably good one for practical purposes.
Chapter 6

Optimal exercise price of
American options near expiry

6.1 Introduction

In Chapter 5, we investigated the pricing approximation for the European style options, which does indeed speed up the calculation of the option prices. From the current chapter to the end of this section, the analytical approximations for American option contracts will be considered. To start with, in this chapter, we shall work on the near-expiry behavior of the optimal exercise price associated with the American puts under the local volatility model.

In the past two decades, many researchers attempted to use different methods to tackle the non-linear problem of pricing American options. Although the analytical solutions for American puts were found by Zhu [115], under the Black-Scholes framework with non-dividend yield, and by Zhao & Wong [113], in the local volatility models, numerical methods are still preferred by market practitioners as they are usually faster with acceptable accuracy. However, due to the fact that the critical price is singular at expiry, as is the case in a similar Stefan problem [69], it is difficult to maintain the same level of accuracy in approximating the optimal exercise price at the time near expiry by using numerical methods. For example, when using both
lattice methods and the projected SOR method for the PDE (partial differential equation) system, a fine discretization must be used near expiry, which is both expensive and limited in accuracy [33]. Therefore, it is quite helpful to determine the asymptotic behavior of the critical price near expiry, and this asymptotic solution can be used as a complement to the numerical approaches to calculate the option values and the optimal exercise price in other larger times away from expiry.

In the literature, some analyses on the asymptotic behavior of the critical price near expiry have already been carried out, as presented in Chapter 1. However, these results are only valid under the Black-Scholes model with constant volatility.

Empirical evidence, however, shows that when using the observed option prices to determine the implied volatility, for which the theoretical prices fit with the observed prices, the options with different strikes have different implied volatility, which violates the Black-Scholes assumption that the volatility is a constant. One possible remedy to cope with the empirical facts is to use the local volatility model, where the volatility $\sigma$ is a deterministic function of both the underlying and the time, such as the CEV (constant elasticity of variance) model, which nests the Brownian motion, the Ornstein-Uhlenbeck process as special cases, the hyperbolic sine model, and so on. The asymptotic behavior of the critical price near expiry in the local volatility model remains unclear. Recent progress was made by Chevalier [22] in extending the previous results under the constant volatility framework to a stock-price-dependent volatility. However, it must be pointed out that the results in [22] cannot be a representative, as the stock-price-dependent volatility is only a special case of the local volatility model.

In this chapter, an explicit analytical expression for the critical price near expiry under the local volatility model is presented. The expression was found by means of the method of matched asymptotic expansions, which generates a sequence of problems for the local behavior near expiry. The results show that if the underlying dividend is greater than the market interest rate, the behavior of the critical price
is parabolic, otherwise, an extra logarithmic factor appears, which agrees with the constant volatility case.

The chapter is organized as follows. In Section 6.2, we introduce the PDE system that the price of an American put option must satisfy in the local volatility model. In Section 6.3, we first deduce the asymptotic behavior of the critical price near expiry under the assumption that the volatility is stock-price-dependent, and then extend the analysis to the general case, where $\sigma$ depends on both the underlying and the time. Concluding remarks are given in the last section.

6.2 American puts under general diffusion process

This chapter considers a general diffusion process for the underlying under the risk-neutral measure. Specifically, the underlying $S_t$, as a function of time, is assumed to follow a diffusion process:

$$\text{d}S_t = (r - D) \text{d}t + \sigma_F \text{d}W_t,$$

where the constants $r \geq 0$ and $D \geq 0$ denote for the risk-free interest rate and the dividend yield, respectively, and the deterministic function $\sigma_F$ represents the local volatility. In this chapter, two cases related to the different forms of $\sigma_F$ are discussed separately: the first one in which $\sigma_F$ is a function of $S_t$ only, i.e., $\sigma_F = \sigma_F(S_t)$, whereas in the second case, $\sigma_F$ is a function of both $S_t$ and $t$, i.e., $\sigma_F = \sigma_F(S_t, t)$.

For each case, the assumption differs as:

- for $\sigma_F = \sigma_F(S_t)$, it is assumed that $\sigma_F(S_t)$ is at least second-order differentiable at the vicinity of $S_t = K$;

- for $\sigma_F = \sigma_F(S_t, t)$, it is assumed that $\sigma_F(S_t, t)$ is at least second-order differentiable at the vicinity of $S_t = K$, $t = T_E$.

In fact, these assumptions are in line with almost all the commonly used local volatility models, such as the CEV model, the hyperbolic sine model, and so on.
Let $P_A(S, t)$ be the price of an American put option, with $S$ being the underlying and $t$ being the time. Then, under the proposed diffusion process (6.2.1), it can be easily shown that the valuation of an American put option can be formulated as a free boundary problem, with $P_A(S, t)$ satisfying

$$
\begin{align*}
\frac{\partial P_A}{\partial t} + \frac{1}{2} \sigma^2 S \frac{\partial^2 P_A}{\partial S^2} + (r - D) S \frac{\partial P_A}{\partial S} - r P_A &= 0, \\
P_A(x, T_E) &= \max(K - S, 0), \\
P_A(S_f(t), t) &= K - S_f(t), \quad \frac{\partial P_A}{\partial S}(S_f(t), t) = -1, \\
\lim_{S \to \infty} P_A(S, t) &= 0.
\end{align*}
$$

This PDE system is defined on $S \in [S_f(t), +\infty]$ and $t \in [0, T_E]$. Moreover, the critical price $S_f$ at expiry $T_E$ in the local volatility model is found out by Zhao [113], i.e.,

$$
S_f(T_E) = \min\left(\frac{r}{D}K, K\right).
$$

6.3 Matched asymptotic analysis for the optimal exercise price near expiry

6.3.1 $\sigma$ is a function of $S$ only

To make the analysis convenient, we shall first non-dimensionalize all the variables by using the new variables:

$$
S = Ke^x, \quad P = \frac{P_A e^{\rho \tau}}{K} + e^{\rho \tau}(e^x - 1), \quad S_f = Ke^{x_f}, \\
\tau = \frac{\sigma^2 F(K)}{2} (T_E - t), \quad \sigma(x) = \sigma_F(Ke^x), \quad a(x) = \frac{\sigma^2(x)}{\sigma^2(0)}.
$$

The parameters $\rho$ and $v$ are defined as

$$
\rho = \frac{2r}{\sigma^2(0)}, \quad v = \frac{2D}{\sigma^2(0)}.
$$
respectively. Then, (6.2.2) can be written in the dimensionless form
\[
\begin{align*}
\frac{\partial P}{\partial \tau} &= a(x) \frac{\partial^2 P}{\partial x^2} + (\rho - v - a(x)) \frac{\partial P}{\partial x} + e^\rho (ve^x - \rho), \\
P(x, 0) &= \max(e^x - 1, 0), \\
P(x_f, \tau) &= 0, \quad \frac{\partial P}{\partial x}(x_f, \tau) = 0, \\
\lim_{x \to \infty} P(x, \tau) &= e^\rho (e^x - 1),
\end{align*}
\]

(6.3.1)

and
\[
x_f(0) = \begin{cases} 
0, & v \leq \rho, \\
- \log\left(\frac{v}{\rho}\right), & v > \rho.
\end{cases}
\]

Now, we shall use the matched asymptotic analysis to construct the small \(\tau\) behavior of \(x_f(\tau)\) for PDE system (6.3.1).

Firstly, we consider the case in which \(D \leq r\), i.e., \(v \leq \rho\). By setting \(\tau = \epsilon T\), where \(T = O(1)\), and \(\epsilon\) is an artificial small parameter, we obtain the PDE system for \(P(x, T)\):
\[
\begin{align*}
\frac{\partial P}{\partial T} &= \epsilon (a(x) \frac{\partial^2 P}{\partial x^2} + (\rho - v - a(x)) \frac{\partial P}{\partial x} + e^{\epsilon \rho T} (ve^x - \rho)), \\
P(x, 0) &= \max(e^x - 1, 0), \\
P(x_f, T) &= 0, \quad \frac{\partial P}{\partial x}(x_f, T) = 0, \\
\lim_{x \to \infty} P(x, T) &= e^{\epsilon \rho T} (e^x - 1).
\end{align*}
\]

(6.3.2)

By assuming that the solution can be expanded in powers of \(\epsilon\), we obtain the outer solution, which is only valid for \(x > 0\) as:
\[
P(x, T) = e^x - 1 + \rho T \epsilon (e^x - 1) + O(\epsilon^2).
\]

Since the outer expansion breaks down at \(x_f(0) = 0\), we need to perform a local
analysis in the vicinity of \( x = 0 \). By using the stretched variable

\[ X = \frac{x}{\sqrt{\epsilon}}, \quad (6.3.3) \]

and substituting (6.3.3) into (6.3.2), we have

\[ \frac{\partial P}{\partial T} = a(\sqrt{\epsilon}X) \frac{\partial^2 P}{\partial X^2} + (\rho - v - a(\sqrt{\epsilon}X)) \frac{\partial P}{\partial X} + \epsilon e^{\epsilon T}(ve^x - \rho). \quad (6.3.4) \]

Since the boundary conditions, which need to be concerned, all have a factor \( \sqrt{\epsilon} \) in common, we rescale the problem by defining

\[ P = \sqrt{\epsilon}p. \quad (6.3.5) \]

On the other hand, assuming that the coefficients have the Taylor expansions to second order at \( x = 0 \), we can expand \( a \), written in the local variable \( X \), as:

\[ a(\sqrt{\epsilon}X) = a(0) + \sqrt{\epsilon}Xa'(0) + \frac{\epsilon X^2}{2}a''(0) + \mathcal{O}(\epsilon^2). \quad (6.3.6) \]

Substituting (6.3.5-6.3.6) into (6.3.4), we obtain the leading order PDE system as

\[ \begin{cases} \frac{\partial p_0}{\partial T} = \frac{\partial^2 p_0}{\partial X^2}, \\ p_0(X, 0) = \max(X, 0), \\ \lim_{X \to \infty} p_0(X, T) = X. \end{cases} \]

The solution of this PDE system can be easily found by using similarity solution techniques. It is

\[ p_0(X, T) = \frac{\sqrt{T}}{\sqrt{\pi}} e^{-\frac{X^2}{4T}} + \frac{X}{2} erf\left( -\frac{X}{2T} \right). \]

The following lemma states that the location of the free boundary \( x_f(\tau) \) is outside the layer near \( x = 0 \), which on the other hand, implies that another layer exists near
Note that in the following work, the notation $U(a, \delta)$ denotes for the neighborhood of a point $a$, i.e., $U(a, \delta) = \{ x \mid 0 \leq |x - a| < \delta \}$.

**Lemma 6.3.1** When $v \leq \rho$, we have $x_f(\tau) \notin U(0, \sqrt{\epsilon})$, where $\tau = \epsilon T$, and $T = O(1)$.

**Proof.** We shall use the method of reductio ad absurdum to complete proving this lemma. Assuming that $x_f(\tau) \in U(0, \sqrt{\epsilon})$, we have $\lim_{\epsilon \to 0} \frac{x_f(\tau)}{\sqrt{\epsilon}} = X_0$, where $X_0$ has finite values for $T = O(1)$. Therefore, we can rescale $x_f(\tau)$ and expand it in terms of $\sqrt{\epsilon}$, i.e., $X_f = \frac{x_f}{\sqrt{\epsilon}} = X_0 + \sqrt{\epsilon} X_1 + O(\epsilon)$. In order to satisfy the moving boundary conditions, the leading order term should satisfy $p_0(X_0) = \frac{\partial p_0}{\partial X}(X_0, T) = 0$, which yields

\[
\frac{\sqrt{\epsilon} T}{\sqrt{\pi}} e^{-\frac{x_0^2}{4\epsilon}} + \frac{X_0}{2} \text{erfc}\left(-\frac{X_0}{2\epsilon T}\right) = 0,
\]

\[
\text{erfc}\left(-\frac{X_0}{2\epsilon T}\right) = 0. \tag{6.3.7}
\]

By solving (6.3.7), we obtain $X_0 = \infty$, and this is in contrast with our assumption that $x_f(\tau) \in U(0, \sqrt{\epsilon})$. Therefore, the location of the free boundary should be outside the $O(\sqrt{\epsilon})$ layer near $x = 0$, and thus $\lim_{\epsilon \to 0} \frac{x_f(\tau)}{\sqrt{\epsilon}} = -\infty$. This completes the proof.

On the other hand, in order to satisfy the free boundary conditions, we use the stretched variable:

\[
z = \frac{x - x_f}{\epsilon}, \tag{6.3.8}
\]

where $z = O(1)$. Substituting (6.3.8) into the governing equation contained in (6.3.2), we have

\[
\begin{cases}
\frac{\partial P}{\partial T} - \frac{x_f}{\epsilon} \frac{\partial P}{\partial T} - \frac{x_f}{\epsilon} \frac{\partial^2 P}{\partial z^2} = a(\epsilon z + x_f) \frac{\partial^2 P}{\partial z^2} + \epsilon(\rho - v - a(\epsilon z + x_f)) \frac{\partial P}{\partial z} + \epsilon^2 e^{\epsilon \rho T}(v \epsilon^{z} x_f - \rho)), \\
P(0, T) = 0, \quad \frac{\partial P}{\partial z}(0, T) = 0.
\end{cases} \tag{6.3.9}
\]

Again, an expansion in regular powers of $\epsilon$ gives the solution of (6.3.9) as

\[
P = O(\epsilon^2). \tag{6.3.10}
\]
In order to match with the solution near \( x_f \), we need to solve the solution in the \( \mathcal{O}(\sqrt{\epsilon}) \) layer near \( x = 0 \). Assuming that

\[
p = p_0 + \sqrt{\epsilon}p_1 + \epsilon^2 p_2 + \mathcal{O}(\epsilon^3),
\]

we obtain the following sequence of PDE systems:

\[
\begin{align*}
\frac{\partial p_0}{\partial T} &= \frac{\partial^2 p_0}{\partial X^2}, \\
p_0(X, 0) &= \max(X, 0), \\
\lim_{X \to -\infty} p_0(X, T) &= X, \quad \lim_{X \to \infty} p_0(X, T) = 0,
\end{align*}
\]

\[
\begin{align*}
\frac{\partial p_1}{\partial T} &= \frac{\partial^2 p_1}{\partial X^2} + a'(0)X \frac{\partial^2 p_0}{\partial X^2} + (\rho - v - 1) \frac{\partial p_0}{\partial X} + v - \rho, \\
p_1(X, 0) &= \max(\frac{1}{2} X^2, 0), \\
\lim_{X \to -\infty} p_1(X, T) &= \frac{1}{2} X^2, \quad \lim_{X \to -\infty} \frac{\partial p_1}{\partial X}(X, T) = 0,
\end{align*}
\]

\[
\begin{align*}
\frac{\partial p_2}{\partial T} &= \frac{\partial^2 p_2}{\partial X^2} + a'(0)X \frac{\partial^2 p_1}{\partial X^2} + \frac{1}{2} a''(0)X^2 \frac{\partial^2 p_0}{\partial X^2} \\
&\quad + (\rho - v - 1) \frac{\partial p_1}{\partial X} - a'(0)X \frac{\partial p_0}{\partial X} + vX, \\
p_2(X, 0) &= \max(\frac{1}{6} X^3, 0), \\
\lim_{X \to -\infty} p_2(X, T) &= \frac{1}{6} X^3 + \rho XT, \quad \lim_{X \to -\infty} \frac{\partial^2 p_2}{\partial X^2}(X, T) = 0.
\end{align*}
\]

One should notice that, in the above PDE systems, the boundary conditions as \( X \to +\infty \) are obtained by matching with the outer expansion; whereas the ones as \( X \to -\infty \) are required to properly close those PDE systems.

The solutions of the above PDE systems can be found by using the similarity solution techniques. The details of the derivation are provided in Appendix C.1. The asymptotic behaviors for \( h_0(\xi) \), \( h_1(\xi) \) and \( h_2(\xi) \) as \( \xi \to -\infty \) can be derived as:

\[
h_0(\xi) = \frac{1}{2\sqrt{\pi}} \frac{e^{-\xi^2}}{\xi^2} + \mathcal{O}(\frac{e^{-\xi^2}}{\xi^4}),
\]

\[
h_1(\xi) = v - \rho + \mathcal{O}(\xi e^{-\xi^2}),
\]

\[
h_2(\xi) = 2v \xi + \mathcal{O}(\xi^3 e^{-\xi^2}).
\]
Now, we match the values of $P_\tau$ in the two different regions, as suggested by Keller [33], to complete the analysis. This is accomplished by taking the limit of $X \to -\infty$ ($\xi \to -\infty$ or $x \to x_f$) of $P_\tau$ given by (6.3.11) and (6.3.10). The leading order term forms the following transcendental equations:

\[
\frac{1}{2\sqrt{\pi \tau}} e^{-\frac{x_f^2}{\tau}} + v - \rho = 0, \quad v < \rho,
\]

\[
\frac{1}{2\sqrt{\pi \tau}} e^{-\frac{x_f^2}{\tau}} + vx_f = 0, \quad v = \rho,
\]

which have the solutions as

\[
x_f(\tau) = -2\sqrt{\tau} \left[ \ln \frac{1}{2(\rho - v)\sqrt{\pi \tau}} \right]^\frac{1}{2}, \quad v < \rho,
\]

\[
x_f(\tau) = -2\sqrt{\tau} \left[ \ln \frac{1}{4\sqrt{\pi v \tau}} \right]^\frac{1}{2}, \quad v = \rho,
\]

respectively. Therefore, for $D < r$,

\[
S_f(t) = Ke^{x_f(t)}
\]

\[
= K - K\sigma(K) \sqrt{(T_E - t) \ln \frac{\sigma^2(K)}{8\pi(T_E - t)(r - D)^2}} + o(\sqrt{(T_E - t) \ln \frac{1}{\sqrt{T_E - t}}}),
\]

and for $D = r$

\[
S_f(t) = Ke^{x_f(t)}
\]

\[
= K - K\sigma(K) \sqrt{2(T_E - t) \ln \frac{1}{4\sqrt{\pi D(T_E - t)}}} + o(\sqrt{(T_E - t) \ln \frac{1}{T_E - t}}).
\]

Now, we consider the case in which $D > r$, i.e., $v > \rho$. Here, we assume that $x_0 = -\log \frac{v}{\rho} << -\sqrt{\epsilon}$. The procedure in deriving the outer expansion is quite similar to the case when $D \leq r$, and the outer solution is

\[
P(x, T) = \begin{cases} 
  e^x - 1 + \rho(e^x - 1)T \epsilon + \mathcal{O}(\epsilon^2) & x \gg \sqrt{\epsilon}, \\
  (ve^x - \rho)T \epsilon + \mathcal{O}(\epsilon^2) & x \ll -\sqrt{\epsilon}.
\end{cases}
\]
One should notice that the leading order solution \( P_0(x, T) = \max(e^x - 1, 0) \) is continuous but not differentiable at \( x = 0 \). Thus, we expect that there is a corner layer at \( x = 0 \), the thickness of which is \( O(\sqrt{\epsilon}) \). The solution in this layer is derived in Appendix C.2 for interested readers. However, based on the assumption that \( x_0 = -\log\frac{v}{\rho} \ll -\sqrt{\epsilon} \), the free boundary is expected to be located outside the corner layer. Therefore, for future analysis, we only need the outer solution which is valid for \( x \ll -\sqrt{\epsilon} \).

Since the outer expansion fails to satisfy the free boundary conditions, we perform a local analysis in the vicinity of \( x_0 \) by using the following rescales:

\[
X = \frac{x - x_0}{\sqrt{\epsilon}}, \quad p = \frac{P}{\epsilon^2}, \quad X_f = \frac{x_f - x_0}{\sqrt{\epsilon}}. \tag{6.3.15}
\]

Substituting (6.3.15) into (6.3.2), the governing equation becomes

\[
\frac{\partial p}{\partial T} = a(\sqrt{\epsilon}X + x_0) \frac{\partial^2 p}{\partial X^2} + (\rho - v - a(\sqrt{\epsilon}X + x_0))\sqrt{\epsilon} \frac{\partial p}{\partial X} + e^{\rho T} \rho \frac{e^{\sqrt{\epsilon}X} - 1}{\sqrt{\epsilon}}. \tag{6.3.16}
\]

On the other hand, assume that \( a(x) \) has Taylor expansions to first order at \( x = x_0 \), i.e.,

\[
a(\sqrt{\epsilon}X + x_0) = a(x_0) + a'(x_0)\sqrt{\epsilon}X + O(\epsilon). \tag{6.3.17}
\]

Substitute (6.3.17) into (6.3.16), the leading order PDE system can be found as

\[
\left\{
\begin{align*}
\frac{\partial p_0}{\partial T} &= a(x_0) \frac{\partial^2 p_0}{\partial X^2} + \rho X, \\
p_0(X, 0) &= 0, \quad \lim_{X \to -\infty} p_0(X, T) = \rho XT.
\end{align*}
\right. \tag{6.3.18}
\]

One should notice that the boundary condition as \( X \to +\infty \) is obtained by matching with one branch of the outer expansion \( (x \ll -\sqrt{\epsilon}) \). It is straightforward to derive the solution of (6.3.18) by using the similarity solution techniques. It has the following
structure \( p_0(X, T) = T^{\frac{3}{2}}h(\xi) \), where

\[
\begin{align*}
\xi &= \frac{X}{2\sqrt{a(x_0)T}}, \\
h(\xi) &= 2\rho\xi + C[(\xi^2 + 1)e^{-\xi^2} - (2\xi^3 + 3\xi) \int_{-\infty}^{+\infty} e^{-t^2} dt],
\end{align*}
\]

with \( C \) being a constant. Now, assuming that the free boundary is located inside the layer near \( x_0 \), just as we did in analyzing the previous case, the rescaled free boundary can be thus expanded in powers of \( \sqrt{\tau} \), i.e., \( X_f = X_1 + \sqrt{\tau}X_2 + O(\epsilon) \). It is clear that \( p_0(X, T) \) should also satisfy

\[
p_0(X_1, T) = \frac{\partial p_0}{\partial X}(X_1, T) = 0, \tag{6.3.19}
\]

which is equivalent to \( h(\xi_1) = h'(\xi_1) = 0 \), where \( \xi_1 = \frac{X_1}{2\sqrt{a(x_0)T}} \). Consequently, we obtain

\[
\begin{align*}
2\rho\xi_1 + C[(\xi_1^2 + 1)e^{-\xi_1^2} - (2\xi_1^3 + 3\xi_1) \int_{-\xi_1}^{+\infty} e^{-t^2} dt] &= 0, \\
2\rho + C[3\xi_1 e^{-\xi_1^2} - (6\xi_1^2 + 3) \int_{-\xi_1}^{+\infty} e^{-t^2} dt] &= 0,
\end{align*}
\]

from which, the transcendental equation for \( \xi_1 \) can be derived as

\[
-\xi_1^3 e^{\xi_1^2} \int_{-\xi_1}^{+\infty} e^{-t^2} dt = \frac{1}{4}(1 - 2\xi_1^2). \tag{6.3.20}
\]

The solution of (6.3.20) is \( \xi_1 = 0.4517 \). Therefore

\[
\begin{align*}
x_f(\tau) &= x_0 - 2\xi_1 \sqrt{a(x_0)T} + O(\tau), \\
&= x_0 - \sqrt{2}\xi_1 \sigma \left( \frac{r}{D}K \right) \sqrt{T_E - t} + O(T_E - t),
\end{align*}
\]

and

\[
S_f(t) = Ke^{z_f} = \frac{r}{D}K \left[ 1 - \sigma \left( \frac{r}{D}K \right) \xi_1 \sqrt{2(T_E - t)} \right] + O(T_E - t).
\]
Remarkably, the leading order terms of the critical price derived in this section are believed to be reasonable, since they agree with those derived in [22], and moreover, they degenerate to Evans et al’s results when $\sigma(S)$ is independent of $S$.

### 6.3.2 $\sigma$ is a function of both $S$ and $t$

Here, we apply singular perturbation techniques to derive the explicit analytical expression for the optimal exercise price near expiry in the local volatility model where $\sigma$ is a function of both $S$ and $t$. For convenience, we shall also non-dimensionalize PDE system (6.2.2). This is achieved by adopting the new variables:

\[
S = Ke^{x}, \quad P = \frac{P_Ae^{\rho\tau}}{K} + e^{\rho\tau}(e^{x} - 1), \quad S_f = Ke^{x_f}, \quad a(x, \tau) = \frac{\sigma^2(x, \tau)}{\sigma^2(0, 0)}
\]

\[
\tau = \frac{\sigma^2_F(K, T_E)}{2}(T_E - t), \quad \sigma(x, \tau) = \sigma_F(Ke^{x}, T_E - \frac{2}{\sigma^2_F(K, T_E)}\tau).
\]

The parameters $\rho$ and $v$ are defined as $\rho = \frac{2r}{\sigma^2(0, 0)}$, $v = \frac{2D}{\sigma^2(0, 0)}$, respectively. Then, (6.2.2) can be written in the dimensionless form

\[
\begin{cases}
\frac{\partial P}{\partial \tau} = a(x, \tau)\frac{\partial^2 P}{\partial x^2} + (\rho - v - a(x, \tau))\frac{\partial P}{\partial x} + e^{\rho\tau}(ve^x - \rho), \\
P(x, 0) = \max(e^x - 1, 0), \\
P(x_f, \tau) = 0, \quad \frac{\partial P}{\partial x}(x_f, \tau) = 0, \\
\lim_{x \to \infty} P(x, \tau) = e^{\rho\tau}(e^x - 1),
\end{cases}
\]

and

\[
x_f(0) = \begin{cases} 
0, & v \leq \rho, \\
-\log\left(\frac{v}{\rho}\right), & v > \rho.
\end{cases}
\]

When $D \leq r$, i.e., $v \leq \rho$, the construction of the asymptotic expansions uses an $O(\sqrt{\epsilon})$ layer at $x = 0$, and the free boundary, in which $P = O(\tau^2)$, is located outside
this $O(\sqrt{\epsilon})$ interior layer. The analysis follows a similar way as in the previous case in which $\sigma$ is a function of $S$. Thus, we shall only describe the results.

For $x \geq \sqrt{\epsilon}$, $P(x, T)$ has the outer expansion as

$$P(x, T) = (e^x - 1) + \rho T \epsilon (e^x - 1) + O(\epsilon^2).$$

For $x = O(\sqrt{\epsilon})$, substituting the rescales $X = \frac{x}{\sqrt{\epsilon}}$ and $p = \frac{P}{\sqrt{\epsilon}}$ into PDE system (6.3.21), we obtain

$$\frac{\partial p}{\partial T} = a(\sqrt{\epsilon} X, \epsilon T) \frac{\partial^2 p}{\partial X^2} + \frac{\rho - v}{\sqrt{\epsilon}} \frac{\partial p}{\partial X} + \sqrt{\epsilon} \epsilon^2 (\epsilon^x - \rho).$$

By assuming that $a(x, \tau)$ has Taylor expansions to second order, i.e.,

$$a(\sqrt{\epsilon} X, \epsilon T) = a(0, 0) + a_x(0, 0) \sqrt{\epsilon} X + a_{\tau}(0, 0) \epsilon T + \frac{1}{2} a_{xx}(0, 0) \epsilon X^2 + O(\epsilon^{3/2}),$$

and $p$ can be expanded in powers of $\sqrt{\epsilon}$, i.e., $p(X, T) = p_0(X, T) + \sqrt{\epsilon} p_1(X, T) + \epsilon p_2(X, T) + O(\epsilon^{3/2})$, we obtain the following sequence of PDE systems:

\[
\begin{align*}
\frac{\partial p_0}{\partial T} &= \frac{\partial^2 p_0}{\partial X^2}, \\
p_0(X, 0) &= \max(X, 0), \\
\lim_{X \to -\infty} p_0(X, T) &= X, \quad \lim_{X \to \infty} p_0(X, T) = 0, \\
\frac{\partial p_1}{\partial T} &= \frac{\partial^2 p_1}{\partial X^2} + a_x(0, 0) X \frac{\partial^2 p_0}{\partial X^2} + (\rho - v - 1) \frac{\partial p_1}{\partial X} + v - \rho, \\
p_1(X, 0) &= \max(\frac{1}{2} X^2, 0), \\
\lim_{X \to -\infty} p_1(X, T) &= \frac{1}{2} X^2, \quad \lim_{X \to \infty} \frac{\partial p_1}{\partial X}(X, T) = 0, \\
\frac{\partial p_2}{\partial T} &= \frac{\partial^2 p_2}{\partial X^2} + a_x(0, 0) X \frac{\partial^2 p_1}{\partial X^2} + (a_x(0, 0) T + \frac{a_{xx}(0, 0)}{2} \epsilon X^2) \frac{\partial^2 p_0}{\partial X^2} \\
&\quad + (\rho - v - 1) \frac{\partial p_1}{\partial X} - a_x(0, 0) X \frac{\partial p_0}{\partial X} + v X, \\
p_2(X, 0) &= \max(\frac{1}{6} X^3, 0), \\
\lim_{X \to -\infty} p_2(X, T) &= \frac{1}{6} X^3 + \rho XT, \quad \lim_{X \to \infty} \frac{\partial^2 p_2}{\partial X^2}(X, T) = 0.
\end{align*}
\]
The solutions of the above PDE systems are derived in Appendix C.3.

One should notice that though the option prices \( p_0, p_1 \) and \( p_2 \) are much more complicated than the corresponding ones presented in the last subsection, fortunately, they have the same asymptotic behaviors as \( \xi \to -\infty \). Next, with the utilization of the same matching procedures as adopted in the last case, we obtain the following transcendental equations:

\[
\frac{1}{2\sqrt{\pi \tau}} e^{-\frac{x^2}{2\tau}} + v - \rho = 0, \quad v < \rho,
\]
\[
\frac{1}{2\sqrt{\pi \tau}} e^{-\frac{x^2}{2\tau}} + vx_f = 0, \quad v = \rho,
\]

from which, the asymptotic behavior of the optimal exercise price near expiry can be derived as

\[
S_f(t) = Ke^{\tau_f(t)}
\]
\[
= K - K\sigma(K, T_E) \sqrt{(T_E - t) \ln \frac{\sigma^2(K, T_E)}{8\pi(T_E - t)(r - D)^2}}
\]
\[
+ o(\sqrt{(T_E - t) \ln \frac{1}{\sqrt{T_E - t}}}), \quad D < r,
\]
\[
S_f(t) = Ke^{\tau_f(t)}
\]
\[
= K - K\sigma(K, T_E) \sqrt{2(T_E - t) \ln \frac{1}{4\sqrt{\pi}D(T_E - t)}}
\]
\[
+ o(\sqrt{(T_E - t) \ln \frac{1}{T_E - t}}), \quad D = r.
\]

When \( D > r \), i.e., \( v > \rho \), we assume that \( x_0 = -\log \frac{v}{\rho} \ll -\sqrt{\epsilon} \). The construction of the asymptotic expansion uses one \( O(\sqrt{\epsilon}) \) corner layer at \( x = 0 \), and another \( O(\sqrt{\epsilon}) \) inner layer at \( x_0 \). The free boundary is located inside the inner layer. The analysis is also similar to that of the last case. For simplicity, we shall briefly describe the difference and list the results.
The outer expansion is valid for $x \gg \sqrt{\epsilon}$ and $x \ll -\sqrt{\epsilon}$, i.e.,

$$P(x, T) = \begin{cases} 
  e^x - 1 + \rho(e^x - 1)T\epsilon + \mathcal{O}(\epsilon^2), & x \gg \sqrt{\epsilon}, \\
  (ve^x - \rho)T\epsilon + \mathcal{O}(\epsilon^2), & x \ll -\sqrt{\epsilon}.
\end{cases}$$

As mentioned in the previous section, for the later matching process, we only need one branch of the outer solution which is valid for $x \ll -\sqrt{\epsilon}$, and thus, we omit the derivation of the solutions in the corner layer. For $x \in U(x_0, \sqrt{\epsilon})$, a local analysis is performed by using the following rescales:

$$X = \frac{x - x_0}{\sqrt{\epsilon}}, \quad p = \frac{P}{\epsilon^3}, \quad X_f = \frac{x_f - x_0}{\sqrt{\epsilon}}. \quad (6.3.25)$$

Again, assume that $a(x, \tau)$ has Taylor expansions to first order at $x = x_0$ and $\tau = 0$, i.e.,

$$a(\sqrt{\epsilon}X + x_0, \epsilon T) = a(x_0, 0) + a_x(x_0, 0)\sqrt{\epsilon}X + a_\tau(x_0, 0)\epsilon T + \mathcal{O}(\epsilon). \quad (6.3.26)$$

Substituting (6.3.25-6.3.26) into (6.3.21), the PDE system governing the leading order term can be found as

$$\begin{align*}
  \frac{\partial p_0}{\partial T} &= a(x_0, 0)\frac{\partial^2 p_0}{\partial X^2} + \rho X, \\
  p_0(X, 0) &= 0, \\
  \lim_{X \to \infty} p_0(X, T) &= \rho XT,
\end{align*}$$

the solution of which is $p_0(X, T) = T^{\frac{3}{2}}h(\xi)$, where

$$\begin{align*}
  \xi &= \frac{X}{2\sqrt{a(x_0, 0)T}}, \\
  h(\xi) &= 2\rho\xi + C[(\xi^2 + 1)e^{-\xi^2} - (2\xi^3 + 3\xi)\int_\xi^{+\infty} e^{-t^2}dt],
\end{align*}$$
with $C$ being a constant. Then, by using (6.3.19) on the free boundary, we obtain

$$2\rho \xi_1 + C[(\xi_1^2 + 1)e^{-\xi_1^2} - (2\xi_1^3 + 3\xi_1)\int_{\xi_1}^{+\infty} e^{-t^2} dt] = 0,$$

and

$$2\rho + C[3\xi_1 e^{-\xi_1^2} - (6\xi_1^2 + 3)\int_{\xi_1}^{+\infty} e^{-t^2} dt] = 0,$$

from which, the transcendental equation for $\xi_1$ can be derived as

$$-\xi_1^3 e^{\xi_1^2} \int_{\xi_1}^{+\infty} e^{-t^2} dt = \frac{1}{4}(1 - 2\xi_1^2). \quad (6.3.27)$$

Here, $\xi_1 = \frac{X_1}{2\sqrt{a(x_0,0)T}}$, and $X_1$ is the leading order term of $X_f$. The solution of (6.3.27) is $\xi_1 = 0.4517$. Therefore

$$x_f(\tau) = x_0 - 2\xi_1 \sqrt{a(x_0,0)T} + O(\tau),$$

$$= x_0 - \sqrt{2}\xi_1 \sigma(DK, TE) \sqrt{TE - t} + O(TE - t),$$

and thus

$$S_f(t) = Ke^{x_f} = \frac{r}{D} K[1 - \sigma(DK, TE)\xi_1 \sqrt{2(TE - t)}] + O(TE - t).$$

It can be clearly seen that if $\sigma(S, t)$ is independent of both $S$ and $t$, our results again degenerate to the ones derived in [33].

### 6.4 Conclusion

In this chapter, the asymptotic behavior of the optimal exercise price for an American put option is investigated in the local volatility model. Based on the singular perturbation methods, the leading order term of the optimal exercise price is derived, which is expected to be complementary to the numerical methods. The result derived in this chapter is believed to be quite reasonable, since the leading order term of the optimal exercise price in the stock-price-dependent volatility model agrees with those
in the literature, and moreover, it degenerates to Evans et al’s result if the volatility function is assumed to be a constant. As the singular perturbation method is not limited to one-dimensional problems, a further task will be to consider its application to American options on underlying described by a multi-factor model, as will be shown in the next chapter.
Chapter 7

Asymptotic behavior of the optimal exercise price of American puts near expiry under stochastic volatility

7.1 Introduction

In the previous chapter, we consider the asymptotic behaviour of the optimal exercise price near expiry under the local volatility model. In this chapter, we further extend the approach presented in Chapter 6 to Heston’s stochastic volatility framework.

As has been discussed in previous chapters, the main difficulty for pricing American options stems from the inherent nonlinear nature of an American option contract itself, i.e., the additional right written in the contract for the holder to exercise the option at any time prior to expiry date. In the context of PDE (partial differential equation) approaches, this is reflected in the fact that the corresponding PDE is associated with an unknown moving boundary and the problem thus becomes a moving boundary problem. As a result, no useful analytical methods are hitherto available for pricing American options under the Heston model, and thus, numerical methods
are preferred by market practitioners. However, it is usually difficult to retain accuracy in approximating the optimal exercise price by means of numerical methods, and the inaccuracy becomes more intolerable when the time is closer to expiry. This is because within the short tenor, the velocity, with which the optimal exercise price reaches its final value, is extremely fast, and is thus, difficult to be approximated by numerical methods. For example, when using the predictor-corrector finite difference scheme [119], a very fine discretization must be adopted near expiry, which is both expensive and limited in accuracy. Therefore, it is quite reasonable to infer that, under the Heston model, the optimal exercise price is also singular at expiry, which is true under the Black-Scholes framework [33]. On the other hand, it is quite useful to determine the asymptotic behavior of the optimal exercise price near expiry, since this asymptotic solution can be used as a complement to the numerical approaches to calculate the option values and the optimal exercise price for other larger times away from the expiry.

In the literature, some analyses on the asymptotic behavior of the optimal exercise price near expiry have already been carried out for the Black-Sholes model, as shown in Chapter 1. Although these results are elegant and useful, they are only valid under the Black-Scholes model. On the other hand, it is a non-trivial task, as will be shown in this chapter, to determine the asymptotic behavior of the optimal exercise price near expiry under the Heston model, since in this case, the optimal exercise price depends, in addition to time, on the dynamics of volatility as well. In other words, the introduction of a second stochastic process has produced a number of new phenomena, which have in turn made the problem much more complicated and totally different from the case with a single stochastic process being used to describe the behavior of the underlying asset only while the volatility is assumed to be a constant.

The aim of this chapter is to present an explicit analytical expression for the optimal exercise price near expiry under the Heston model by means of the method of matched asymptotic expansions, which generates a sequence of simplified PDE
systems for the local behavior near expiry. It turns out that, even with stochastic volatility being taken into consideration, the convergence rate for the calculation of the optimal exercise price is almost the same as that for the constant volatility case.

The chapter is organized as follows. In Section 7.2, we introduce the PDE system that the price of an American put option must satisfy under the Heston model. In Section 7.3, we deduce the asymptotic behavior of the optimal exercise price near expiry by using singular perturbation method. In Section 7.4, we compare our approximation with the numerical results calculated by the predictor-corrector finite difference method [119], to illustrate the reliability of our asymptotic solution. Concluding remarks are given in the last section.

7.2 American puts under the Heston model

Let $P_A(S,v,t)$ 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. Then, under the Heston model, it is easily shown that the valuation of an American put option can be formulated as a free boundary problem [97], in which the boundary location itself is part of the solution of the problem. Specifically, $P_A$ satisfies

$$
\begin{align*}
\frac{1}{2}vS^2\frac{\partial^2 P_A}{\partial S^2} + \rho\sigma vS\frac{\partial^2 P_A}{\partial S\partial v} + \frac{1}{2}\sigma^2 v\frac{\partial^2 P_A}{\partial v^2} + (r - D)S\frac{\partial P_A}{\partial S} + \kappa(\eta - v)\frac{\partial P_A}{\partial v} - rP_A + \frac{\partial P_A}{\partial t} &= 0, \\
P_A(S,v,T_E) &= \max(K - S, 0), \quad \lim_{S \to -\infty} P_A(S,v,t) = 0, \\
P_A(S_f(v,t),v,t) &= K - S_f(v,t), \quad \frac{\partial P_A}{\partial S}(S_f(v,t),v,t) = -1, \\
\lim_{v \to 0} P_A(S,v,t) &= \max(K - S, 0), \quad \lim_{v \to -\infty} P_A(S,v,t) = K.
\end{align*}
$$

This PDE system is defined on $S \in [S_f(v,t), +\infty], v \in [0, +\infty]$ and $t \in [0, T_E]$. Unlike the Black-Scholes case, the unknown optimal exercise price $S_f$ now depends on both the time and the volatility.

For simplicity, in our work, we assume that the risk-free interest rate $r$ is greater
than the constant dividend yield $D$. It should be remarked that, when $r > D$, the riskless growing argument proposed in Zhu and Chen [119] also holds, and thus the boundary conditions established therein can still be used here. Moreover, it is also straightforward to show that [119]:

$$S_f(v, T_E) = K, \quad S_f(0, t) = K,$$

which financially simply state that at the expiration date or when the spot volatility is zero, the optimal exercise price is equal to the strike price. However, just as a similar case to the Stephan problem [97], the “velocity”, with which the optimal exercise price reaches its final value, is extremely fast, and is difficult to be approximated by numerical approaches. Therefore, in the next section, we shall construct the asymptotic behavior of the optimal exercise price near expiry by using the method of matched asymptotic expansions.

### 7.3 Matched asymptotic expansions for the optimal exercise price near expiry

To make the analysis convenient, we shall first non-dimensionalize all variables by using the new variables

$$S = Ke^x, \quad P = \frac{P_Ae^{\sigma \tau}}{K} + e^{\sigma \tau}(e^x - 1), \quad S_f = Ke^{x_f}, \quad \tau = \frac{\sigma}{2}(T_E - t),$$

where the parameters $q$ and $d$ are defined as

$$q = \frac{2r}{\sigma^2}, \quad d = \frac{2D}{\sigma^2},$$
respectively. Then, (7.2.1) can be written in the dimensionless form

\[
\begin{align*}
\frac{\partial P}{\partial \tau} &= \frac{v}{\sigma^2} \frac{\partial^2 P}{\partial x^2} + \frac{(q - v - \frac{v}{\sigma^2})}{2\rho \nu} \frac{\partial^2 P}{\partial v^2} + \frac{2k}{\sigma^2} (\eta - v) \frac{\partial P}{\partial v} \\
&\quad + \frac{e^{\eta \tau} (de^x - q)}{\sigma} \\
&\quad + \frac{2 \rho v}{\sigma} \frac{\partial^2 P}{\partial x \partial v} + e^{\eta \tau} (de^x - q),
\end{align*}
\]

\[
P(x, v, 0) = \max(e^x - 1, 0), \quad \lim_{x \to -\infty} P(x, v, \tau) = e^{\eta \tau} (e^x - 1),
\]

\[
P(x_f, v, \tau) = 0, \quad \frac{\partial P}{\partial x}(x_f, v, \tau) = 0,
\]

\[
\lim_{v \to 0} P(x, v, \tau) = e^{\eta \tau} \max(1 - e^x, 0) + e^{\eta \tau} (e^x - 1), \quad \lim_{v \to \infty} P(x, v, \tau) = e^{\eta \tau + x},
\]

(7.3.1)

together with two more conditions for the optimal exercise price

\[
x_f(v, 0) = x_f(0, \tau) = 0.
\]

(7.3.2)

One should notice that, although the governing differential equation itself in (7.3.1) is linear in terms of the unknown function \( P \), it is the unknown boundary that has made this PDE system highly nonlinear. The nonlinearity of the problem will be clearly manifested once a Landau transform is used to convert the moving boundary problem into a fixed boundary problem, as demonstrated by Zhu and Chen [119]. On the other hand, the high nonlinearity as well as the introduction of another new variable \( v \), has resulted in the analytical methods less achievable than the Black-Scholes case [33]. Consequently, we shall use the method of matched asymptotic expansions, which is an ideal tool to deal with the nonlinear problems, to construct an approximation of \( x_f(v, \tau) \) for the PDE system (7.3.1). Hereafter, we only consider the options with short tenor, i.e.,

\[
\tau = \epsilon T,
\]

(7.3.3)

where \( T = \mathcal{O}(1) \), and \( \epsilon \) is a small positive parameter. By substituting (7.3.3) into
(7.3.1), we obtain the PDE system for $P(x, v, T)$:

$$
\begin{align*}
\frac{\partial P}{\partial T} &= \epsilon \left[ \frac{v}{\sigma^2} \frac{\partial^2 P}{\partial x^2} + \left( q - d - \frac{v}{\sigma^2} \right) \frac{\partial P}{\partial x} + \frac{v}{\sigma^2} \frac{\partial^2 P}{\partial v^2} 
+ \frac{2\kappa}{\sigma^2} (\eta - v) \frac{\partial P}{\partial v} + \frac{2\rho v}{\sigma} \frac{\partial^2 P}{\partial x \partial v} \right] + \epsilon e^{\epsilon qT} (de^{x} - q), \\
\frac{\partial P}{\partial x} (x_f, v, T) &= 0, \\
\frac{\partial P}{\partial x} (x, v, T) &= 0, \\
\lim_{v \to \infty} P(x, v, T) &= e^{\epsilon qT} (e^x - 1), \\
\lim_{v \to 0} P(x, v, T) &= e^{\epsilon qT} \max(1 - e^x, 0) + e^{\epsilon qT} (e^x - 1), \\
\lim_{v \to \infty} P(x, v, T) &= e^{\epsilon qT T + x}.
\end{align*}
$$

(7.3.4)

Unlike the constant volatility case discussed in Evans et al. [33], the PDE system (7.3.4) needs to be dealt with care; there are several different regions, or the so-called “boundary layers”, in which either the unknown function $P$ or its partial derivatives change rapidly. The increased complexity, as will be shown later, is a result of the introduction of stochastic volatility and, consequently, a new dimension of the PDE system. “Boundary layer” is a phrase commonly used in physics and fluid mechanics to describe the layer of fluid in the immediate vicinity of a boundary surface [68], and it has been adopted in most of singular perturbation analyses as well. Therefore, we shall also use it in this chapter to derive the specific boundary layer structure for the PDE system (7.3.4).

The procedure usually begins with the assumption that the solution can be expanded in powers of $\epsilon$, i.e.,

$$
P_{outer} = P_0 + \epsilon P_1 + \mathcal{O}(\epsilon^2), \tag{7.3.5}
$$

in which the subscript stands for the outer solution in the region outside of boundary layer where a singular expansion is required to deal with rapid changes of the value of $P$ or its derivatives. Of course, in this “outer region”, only a regular expansion in the original variables $x$ and $v$ is needed, as shown in (7.3.5) already. After some
simple calculations, we obtain the solution as

\[ P_{\text{outer}}(x, T) = e^x - 1 + qT \epsilon (e^x - 1) + \mathcal{O}(\epsilon^2). \] (7.3.6)

It should be noted that the above solution is valid for the domain \( x > 0, 0 \leq v < \infty \), and fails to satisfy the boundary condition at \( v = \infty \). Therefore, there is a boundary layer near \( v = \infty \), with the layer thickness in the order of \( \epsilon \). It can be inferred that, in order to satisfy the boundary condition at \( v = \infty \), the exact solution of (7.3.4), compared with \( P_{\text{outer}} \), changes rapidly only within the \( \mathcal{O}(\epsilon) \) neighborhood of \( v = \infty \). In addition, in the traded market, the volatility value is usually very small, and the highest value of the volatility that has ever been recorded on Chicago Board Options Exchange (CBOE) is only 0.85 [34]. Therefore, with these two points in mind, it is perfectly justifiable not analyzing this boundary layer at all; our solution will be a reasonably good approximation for any large enough but finite \( v \) values.

On the other hand, it is clear that as \( x \to 0^+ \), too many terms on the right hand side of the governing equation contained in (7.3.4) have been dropped, and thus, we need to perform an asymptotic analysis in the vicinity of \( x = 0 \). By using the stretched variable

\[ X = \frac{x}{\epsilon^\alpha}, \] (7.3.7)

and substituting (7.3.7) into (7.3.4), we obtain

\[ \frac{\partial P}{\partial T} = \epsilon^{1-2\alpha} \frac{v}{\sigma^2} \frac{\partial^2 P}{\partial X^2} + \epsilon^{1-\alpha} (\rho - v - \frac{v}{\sigma^2}) \frac{\partial P}{\partial X} + \epsilon v \frac{\partial^2 P}{\partial v^2} + \epsilon \frac{2\kappa}{\sigma^2} (\eta - v) \frac{\partial P}{\partial v} \\
+ \epsilon^{1-\alpha} \frac{2\rho v}{\sigma} \frac{\partial^2 P}{\partial X \partial v} + \epsilon \epsilon q T (d e^{\alpha X} - q). \]

The significant degeneration of the above operator arises if \( \alpha = \frac{1}{2} \), and thus, the boundary layer near \( x = 0 \) is with the thickness of \( \mathcal{O}(\sqrt{\epsilon}) \). Assuming that \( P(X, v, T) \) has a regular expansion in this region, we write

\[ P = \sqrt{\epsilon} P_0 + \epsilon P_1 + \epsilon^2 P_2 + \mathcal{O}(\epsilon^3), \] (7.3.8)
where the explicit analytical expressions of \( P_0, P_1 \) and \( P_2 \) are derived in Appendix D.1. The solution derived in this boundary layer can be referred to as the inner solution with respect to (w.r.t. hereafter) \( P_{\text{outer}} \).

It can be easily shown that the first order derivative of \( P_0 \) w.r.t. \( v \) becomes unbounded at the corner \((X, v) = (0, 0)\). Consequently, the derivatives of \( P_0 \) w.r.t. \( v \) cannot be ignored when deriving the PDE system for \( P_0 \) around that corner. Another local analysis is again needed. By setting \( V = \frac{v}{\epsilon^\beta} \) and investigating the significant degeneration of the corresponding operator, it is obvious that \( \beta = 1 \) is a well-balanced choice. Assuming that \( Z_1(X, V, T) \) (the solution at the corner) can be expanded in powers of \( \epsilon \), we obtain

\[
Z_1 = \sqrt{\epsilon} Z_{11} + \epsilon Z_{12} + \epsilon^2 Z_{13} + \mathcal{O}(\epsilon^3),
\]

where the explicit analytical expressions of \( Z_{11}, Z_{12} \) and \( Z_{13} \) are derived in Appendix D.2. Furthermore, it can be easily shown that \( Z_{11} \) is continuous but not differentiable w.r.t. \( X \) at the corner \((X, v) = (0, 0)\). Thus, the following stretched variables are needed:

\[
X_1 = \frac{x}{\epsilon^\mu}, \quad V_1 = \frac{v}{\epsilon^\nu},
\]

where \( \mu = \nu = 1 \) are determined by investigating the significant degeneration of the corresponding operator again.

The above analysis has clearly demonstrated the boundary layer structure of our problem, as shown in Fig 7.1, where the Roman numbers I, II, III and IV stand for four different regions in which local analysis needs to be carried out consecutively. In particular, Region I represents the valid domain of \( P_{\text{outer}} \), while Region II shows the \( \mathcal{O}(\sqrt{\epsilon}) \) boundary layer near \( x = 0 \). Region III and IV denote for the corner boundary layers, and they are defined as

\[
(x, v) \in \left[-\sqrt{\epsilon}, \sqrt{\epsilon}\right] \times [0, \epsilon],
\]

\[
(x, v) \in \left[-\epsilon, \epsilon\right] \times [0, \epsilon],
\]
respectively. Moreover, near $x_f$, there might be another boundary layer, which will be discussed later. Note again that the $O(\epsilon)$ boundary layer near $v = \infty$ has been ignored.

It should be remarked here that the introduction of a second stochastic process for $v$ has indeed made the analysis much more complicated and totally different from that of the Black-Scholes case as shown in Evans et al. [33]. The complexity as well as the difference has been clearly manifested when analyzing the solution within the domain $x \in U(0, \sqrt{\epsilon})$, $0 \leq v < \infty$. Here, the notation $U(a, \delta)$ denotes for the neighborhood of a point $a$ with radius $\delta > 0$, i.e.,

$$U(a, \delta) = \{ x \mid 0 \leq |x - a| < \delta \}.$$  

In the classical Black-Scholes’ case, the analysis usually stops once the inner solution within the $O(\sqrt{\epsilon})$ boundary layer near $x = 0$ has been found. Under the Heston model, however, the “inner” region $x \in U(0, \sqrt{\epsilon})$ (relative to the “outer” Region I)
Chapter 7.

151

needs to be further divided into another set of "inner" and "outer" regions because there exists another boundary layer in the $v$ direction as a result of Region $II$ still containing the singular point at the origin. This new inner region is denoted as Region $III$. Similarly, the same process needs to be repeated, which yields Region $IV$ as a boundary layer in the $x$ again. It is envisaged that this "cascade" phenomenon of sub-dividing regions is a particular feature associated with the Heston model when it is used to price American options. However, since we are only interested in an approximate solution, based on the method of matched asymptotic expansions, of $O(\epsilon)$, it suffices to stop the sub-division at Region $IV$, as shall be discussed later.

With the division of the original solution domain into the above four regions, we can now prove the following lemma for the asymptotic behavior of $x_f(v, \tau)$ when $v \not\in U(0, \epsilon)$. Note that hereafter, we shall use a new definition $O_s(\cdot)$, i.e., $f(\epsilon) = O_s(g(\epsilon))$, if $f(\epsilon) = O(g(\epsilon))$, and $f(\epsilon) \neq o(g(\epsilon))$ for $\epsilon \to 0$. As shall be demonstrated, this new definition enables us to give a sharp estimate of the target functions.

**Lemma 7.3.1** (i). For $v = O_s(1)$, we have $x_f(v, \tau) \not\in U(0, \sqrt{\epsilon})$. (ii). For $v = O_s(\epsilon^\beta)$, where $0 < \beta < 1$, we have $x_f(v, \tau) \in U(0, \sqrt{\epsilon})$, and the leading order term of $x_f$, say $x_0$, should be bounded as $v \to 0$, and satisfy $\lim_{v \to 0} \frac{x_0(v, \tau)}{2\sqrt{a\tau}} = -\infty$, where $a = \frac{v}{\sigma^2}$.

**Proof.** (i) The first part of the lemma is proved by means of the method of proof by contradiction. Assuming that $x_f(v, \tau) \in U(0, \sqrt{\epsilon})$, we have $\lim_{\epsilon \to 0} \frac{x_f(v, \tau)}{\sqrt{\epsilon}} = X_0$, where $X_0$ should have finite values for any $v = O_s(1)$ and $T = O(1)$. Therefore, we can rescale $x_f(v, \tau)$ and expand it in terms of $\sqrt{\epsilon}$, i.e.,

$$X_f = \frac{x_f}{\sqrt{\epsilon}} = X_0 + \sqrt{\epsilon}X_1 + O(\epsilon).$$

In order to satisfy the moving boundary conditions, the leading order term should at least satisfy

$$P_0(X_0, v, T) = \frac{\partial P_0}{\partial X}(X_0, v, T) = 0,$$
which yields

\[ \frac{vT}{\sigma \sqrt{\pi}} e^{-\frac{x_0^2}{4vT}} + \frac{X_0}{2} \text{erfc}(\frac{-\sigma X_0}{2\sqrt{vT}}) = 0, \quad (7.3.9) \]

\[ \text{erfc}(\frac{-\sigma X_0}{2\sqrt{vT}}) = 0. \quad (7.3.10) \]

However, when \( v = \mathcal{O}_s(1) \), the only solution of (7.3.9)-(7.3.10) is \( X_0 = -\infty \), and this is in contrast with our assumption that \( x_f(v, \tau) \in U(0, \sqrt{\epsilon}) \). Therefore, when \( v = \mathcal{O}_s(1) \), the location of the free boundary should be outside the \( \mathcal{O}(\sqrt{\epsilon}) \) layer near \( x = 0 \), and thus \( \lim_{\epsilon \to 0} \frac{x_f(v, \tau)}{\sqrt{\epsilon}} = -\infty \). This completes the proof.

(ii) It is straightforward to show that if \( \lim_{v \to 0} \frac{x_0(v, \tau)}{2\sqrt{a\tau}} = -\infty \), (7.3.9)-(7.3.10) are always satisfied. On the other hand, when \( v = \mathcal{O}_s(\epsilon^\beta) \), the leading order term of \( x_f(v, \tau) \), i.e., \( x_0 \) can be bounded, and satisfy \( \lim_{v \to 0} \frac{x_0(v, \tau)}{2\sqrt{a\tau}} = -\infty \) at the same time. Therefore, when \( v = \mathcal{O}_s(\epsilon^\beta) \), where \( 0 < \beta < 1 \), \( x_f \) should be located inside \( U(0, \sqrt{\epsilon}) \). This completes the proof.

Since the location of the free boundary \( x_f(v, \tau) \) differs w.r.t. \( v \), it is much more convenient to discuss the asymptotic behavior of \( x_f(v, \tau) \) in different ranges of \( v \) separately.

**Case I.** \( v = \mathcal{O}_s(1) \)

In this case, \( x_f(v, \tau) \) is located outside \( U(0, \sqrt{\epsilon}) \), which indicates that there is another boundary layer near \( x_f(v, \tau) \). Now, we perform the local analysis in the vicinity of \( x_f \) by using the stretched variable:

\[ Z = \frac{x - x_f(v, \tau)}{\epsilon}, \quad (7.3.11) \]

where \( Z = \mathcal{O}(1) \). Substituting (7.3.11) into the governing equation contained in
(7.3.4), we obtain

\[
\begin{align*}
\frac{\partial P}{\partial T} - \frac{x_f \partial P}{\partial Z} &= v \frac{\partial^2 P}{\partial Z^2} + \epsilon(q - d - \frac{v}{\sigma^2}) \frac{\partial P}{\partial Z} + v \epsilon^2 \frac{\partial^2 P}{\partial v^2} - 2\epsilon \frac{\partial^2 P}{\partial v \partial Z} \frac{\partial x_f}{\partial v} \\
&\quad + 2\epsilon \frac{\partial P}{\partial v} \frac{\partial x_f}{\partial v} + \epsilon^2 \frac{\partial^2 P}{\partial v \partial Z} (\frac{\partial x_f}{\partial v})^2 - \epsilon \frac{\partial P}{\partial v} \frac{\partial^2 x_f}{\partial v^2} + \frac{2\kappa \sigma}{\sigma^2} (\eta - v) (\epsilon^2 \frac{\partial P}{\partial v} - \epsilon \frac{\partial P}{\partial Z} \frac{\partial x_f}{\partial v}) \\
&\quad + 2\frac{\partial v}{\sigma} \left( \epsilon \frac{\partial^2 P}{\partial v \partial Z} - \frac{\partial^2 P}{\partial Z^2} \frac{\partial x_f}{\partial v} \right) + \epsilon^2 e^{x_f T} (de^{xz_f} - q),
\end{align*}
\]

\(P(0, v, T) = 0, \quad \frac{\partial P}{\partial Z}(0, v, T) = 0.\)  

(7.3.12)

Again, an expansion in regular powers of \(\epsilon\) gives the solution of (7.3.12) as

\[P = \mathcal{O}(\epsilon^3).\]  

(7.3.13)

In order to match the solution in Region II, with the solution near \(x_f\), we found the asymptotic behaviors of \(h_0, h_1\) and \(h_2\) as \(\xi \to -\infty\):

\[
h_0(\xi) = \frac{\sqrt{a}}{2\sqrt{\pi}} e^{-\xi_1^2} + \mathcal{O}\left(\frac{e^{-\xi_1^2}}{\xi_1^4}\right),
\]

\[
h_1(\xi) = d - q + \mathcal{O}(\xi_1 e^{-\xi_1^2}),
\]

\[
h_2(\xi) = 2d\sqrt{a} \xi_1 + \mathcal{O}(\xi_1^4 e^{-\xi_1^2}),
\]

where \(\xi_1 = \frac{\xi}{\sqrt{a}}\).

Now, if we follow Evans et al. [33], and match the \(P_\tau\) values of (7.3.8) and (7.3.13) by taking the limit \(x \to x_f\), we obtain the following transcendental equations:

\[
\frac{\sqrt{a}}{2\sqrt{\pi \tau}} e^{-\frac{x_f^2}{4\pi \tau}} + d - q = 0, \quad d < q,
\]

(7.3.14)

\[
\frac{\sqrt{a}}{2\sqrt{\pi \tau}} e^{-\frac{x_f^2}{4\pi \tau}} + d\sqrt{ax_f} = 0, \quad d = q,
\]

(7.3.15)

which lead to the solutions of the form

\[
x_f(v, \tau) = -2\sqrt{a\tau} \sqrt{\left| \ln \frac{(q - d)2\sqrt{\pi \tau}}{\sqrt{a}} \right|}, \quad d < q,
\]

(7.3.16)

\[
x_f(v, \tau) = -2\sqrt{a\tau} \sqrt{\left| \ln \frac{1}{4\sqrt{\pi \tau} d\tau} \right|}, \quad d = q,
\]

(7.3.17)
respectively, after higher-order terms are ignored.

It should be remarked here that matching $P_\tau$ values is not the unique choice. Evans et al. [33] did in this way without any detailed explanation. In fact, this approach may be at odds with the conventional method of matched asymptotic expansions in which $P$, instead of $P_\tau$, values should be matched in the “inner” and “outer” regions. Clearly, if a solution is obtained with $P$ values being matched, so should the $P_\tau$ values, provided that $P$ function is of sufficient smoothness. However the converse is not true. One naturally wonders whether or not the two different approaches would lead to the same conclusion, once all the higher-order terms are ignored. Without a great deal of additional effort, it can be shown that matching the $P$ values of (7.3.8) and (7.3.13) by taking the limit $x \to x_f$, leads to

$$2a^{3/2} \tau^{3/2} e^{-x_f^2/4\tau} + (d - q)\tau + d\tau x_f^* = 0. \quad (7.3.18)$$

In comparison with (7.3.14)-(7.3.15), it is much more difficult to go through some additional order analysis to simplify (7.3.18), in order to analytically obtain the asymptotic behavior of $x_f^*$ like those presented in (7.3.16) and (7.3.17). However, it can still be shown that $x_f^*$ lies between $x_f$ and $-2\sqrt{a\tau}$ (see Lemma 7.3.2 below), where $x_f$ is solved from matching with $P_\tau$. Hence, as $\tau$ approaches zero, the difference between $x_f$ and $x_f^*$ will become smaller. It should be noted that we only need to consider the negative root of (7.3.18) because of the physical restriction that $x_f \leq 0$.

**Lemma 7.3.2** For small $\tau$, (7.3.18) has only one negative solution $x_f^*$, and moreover, it satisfies

$$x_f < x_f^* < -2\sqrt{a\tau}, \quad (7.3.19)$$

where

$$x_f = -2\sqrt{a\tau} \sqrt{\left| \ln \frac{(q - d)2\sqrt{\pi \tau}}{\sqrt{a}} \right|}, \quad d < q \quad (7.3.20)$$

$$x_f = -2\sqrt{a\tau} \sqrt{\left| \ln \frac{1}{4\sqrt{\pi d\tau}} \right|}, \quad d = q. \quad (7.3.21)$$
Proof. Denote
\[ f(x) = \frac{\sqrt{a\tau}e^{-x^2}}{2\sqrt{\pi}x^2} + (d - q)\tau + 2d\sqrt{a\tau}^2x. \quad (7.3.22) \]
Comparing (7.3.22) with (7.3.18), it is obvious that \( x = \frac{x^*}{2\sqrt{a\tau}} \). Then, by taking the first order derivative of \( f(x) \) w.r.t. \( x \), we obtain:
\[ f'(x) = -e^{-x^2}\frac{\sqrt{a\tau}}{\sqrt{\pi}}(x^2 + 1) + 2d\sqrt{a\tau}^2. \quad (7.3.23) \]
Since the parameters \( a, \tau \) and \( d \) are all greater than zero, we have \( f'(x) > 0 \) for any \( x < 0 \). Therefore, \( f(x) \) is monotonically increasing for negative \( x \). On the other hand, it is straightforward to show that when \( d < q \),
\[ f(-1) = \frac{\sqrt{a\tau}e^{-1}}{2\sqrt{\pi}} + (d - q)\tau - 2d\sqrt{a\tau}^2, \]
\[ f\left(-\sqrt{\frac{\ln |(q-d)2\sqrt{\pi\tau}|}{\ln |(q-d)2\sqrt{\pi\tau}|}}\right) = \tau(q-d)\left(\frac{1}{\ln |(q-d)2\sqrt{\pi\tau}|} - 1\right) \]
\[ - 2d\sqrt{a\tau}^2 \sqrt{\frac{|\ln (q-d)2\sqrt{\pi\tau}|}{\sqrt{a}}}. \]
Furthermore, for reasonably small \( \tau \), we have
\[ \frac{\sqrt{ae^{-1}}}{2\sqrt{\pi}} > (q-d)\sqrt{\tau} + 2d\sqrt{a\tau}, \quad |\ln \frac{(q-d)2\sqrt{\pi\tau}}{\sqrt{a}}| > 1. \]
Therefore, \( f(-1) > 0 \) and \( f\left(-\sqrt{\frac{\ln |(q-d)2\sqrt{\pi\tau}|}{\ln |(q-d)2\sqrt{\pi\tau}|}}\right) < 0 \). Based on the monotonicity of \( f(x) \), the only negative root of \( f \) should lie between \( -\sqrt{\frac{\ln |(q-d)2\sqrt{\pi\tau}|}{\ln |(q-d)2\sqrt{\pi\tau}|}} \) and \( -1 \).
Similarly, we can show that when \( q = d \), it lies between \( -\sqrt{\frac{1}{4\sqrt{\pi}d\tau}} \) and \( -1 \).
Therefore, for small \( \tau \), (7.3.18) has only one negative solution \( x_f^* \), and moreover, it satisfies
\[ x_f < x_f^* < -2\sqrt{a\tau}. \quad (7.3.24) \]
This completes the proof.

On the other hand, Eq. (7.3.18) can also be numerically solved and numerical
evidence indeed suggests that the difference between $x_f^*$ and $x_f$ be negligible. This probably explains why Evans et al.
[33] chose to match the $P_*$ values in their analysis.

By substituting (7.3.16)-(7.3.17) into (7.3.8), we find that the free boundary conditions can be satisfied in the following
asymptotic sense, i.e.,

$$P(x_f, v, \tau) - P_{exact}(\tilde{x}_f, v, \tau) = O(\tau), \quad d < q,$$

(7.3.25)

$$P(x_f, v, \tau) - P_{exact}(\tilde{x}_f, v, \tau) = O(\tau^{\frac{3}{2}} \sqrt{\ln \frac{1}{\tau}}), \quad d = q,$$

(7.3.26)

where $\tilde{x}_f$ stands for the exact free boundary.

**Case II.** $v = O(\epsilon^{\beta})$, $0 < \beta < 1$

According to the second part of Lemma 7.3.1, $x_f(v, \tau)$ should be located inside $U(0, \sqrt{\epsilon})$. Now, we assume that (7.3.8) can satisfy the conditions across the free boundary. By taking the limit $\epsilon \to 0$, we deduce the asymptotic behavior of $h_0$, $h_1$ and $h_2$, respectively, as:

$$h_0(\xi) = \frac{\sqrt{a}}{2\sqrt{\pi}} e^{-\xi_1^2} + O(\sqrt{a} e^{-\xi_1^2}),$$

(7.3.27)

$$h_1(\xi) = d - q + O(\xi_1 e^{-\xi_1^2}),$$

(7.3.28)

$$h_2(\xi) = 2d\sqrt{a} \xi_1 + O(\xi_1^4 e^{-\xi_1^2}),$$

(7.3.29)

where $\xi_1 = \frac{\xi}{\sqrt{a}}$. It is not strange that the asymptotic behaviors of $h_0$, $h_1$ and $h_2$ as $\epsilon \to 0$ are quite similar to those derived in case I, since when taking the limit $\epsilon \to 0$, it is equivalent to $v \to 0$, which is also equal to $\xi_1 \to -\infty$ and $a \to 0$.

On the other hand, upon applying the free boundary conditions on (7.3.8), the leading order term of $x_f$ should at least satisfy

$$P_*(x_f, v, \tau) = O(\tau),$$
which is the same as what we have obtained in case I. It is now quite trivial to show that the leading order term of $x_f$ is the same as the one derived in case I, and the expansion satisfies the free boundary conditions in almost the same asymptotic sense, i.e.,

$$P(x_f, v, \tau) - P_{\text{exact}}(\tilde{x}_f, v, \tau) = \mathcal{O}(\tau) \quad d < q. \quad (7.3.30)$$

$$P(x_f, v, \tau) - P_{\text{exact}}(\tilde{x}_f, v, \tau) = \mathcal{O}(\tau^{2+\frac{d}{2}} \sqrt{\ln \frac{1}{\tau}}) \quad d = q. \quad (7.3.31)$$

$\partial P \partial x (x_f, v, \tau) - \partial P_{\text{exact}} \partial x (\tilde{x}_f, v, \tau) = \mathcal{O}(\tau^{2}) \quad d \leq q. \quad (7.3.31)$

**Case III.** $v = \mathcal{O}(\epsilon)$

In this case, if we follow the procedure described previously, we should first determine whether $x_f(v, \tau) \in U(0, \epsilon)$ or not, which is equivalent to analyzing the solution in Region $IV$ can satisfy the free boundary conditions or not. It is straightforward to show that the governing equation for the solution in Region $IV$ is

$$\frac{\partial Z_2}{\partial T} = \frac{V}{\sigma^2} \frac{\partial^2 Z_2}{\partial X_1^2} + (q - v) \frac{\partial Z_2}{\partial X_1} + \frac{2\rho V}{\sigma} \frac{\partial^2 Z_2}{\partial X_1 \partial V} + V \frac{\partial^2 Z_2}{\partial V^2} + \frac{2\kappa \eta}{\sigma^2} \frac{\partial Z_2}{\partial V} + d - q, \quad (7.3.32)$$

which contains all the derivatives the original equation has. Now, we are in an unfortunate situation of having to solve almost the full problem to be able to determine what is going on in Region $IV$. If we could solve this problem, one might wonder why it was necessary to bother with an approximation in the first place. There is no need to argue with this sentiment, and this is indeed one of the situations where the perturbation methods show some of their limitations. However, there are several remarks that should be made. Firstly, though this layer problem cannot be solved in a closed form, we may still be able to extract some useful information about the solution. Secondly, this problem can be further dealt with if we again apply the method of matched asymptotic expansions to this corner, i.e., by rescaling $T_1 = \frac{T}{\epsilon}$, and following almost the same procedure as demonstrated previously. Fortunately, there is no need to go through such a cumbersome analysis again. In order to demon-
strate the reasons in a clear way, we adopt a new notation \( \bar{x}_f \) for the actual optimal exercise boundary in this case.

As it turns out, when \( v = \mathcal{O}(\epsilon) \), the location of the optimal exercise boundary should be located either inside \( U(0, \epsilon) \) or outside. However, it is claimed that no matter what is the truth, the approximation of \( x_f \) derived before for \( d < q \) can still be used here. Firstly, if \( \bar{x}_f \) is located outside \( U(0, \epsilon) \), then (7.3.32) is defined on

\[
\Omega = \{ -\infty < X_1 < +\infty, 0 < V < +\infty, 0 < T < +\infty \}.
\]

Taking the corresponding boundary conditions into consideration, it can be identified that \( Z_2 = \epsilon Z_{21} + \mathcal{O}(\epsilon^2) \), with \( Z_{21} \) twice continuously differentiable w.r.t. \( X_1 \) or \( V \) on \( \Omega \). The stretched \( x_f \), which is derived in the previous two cases, reads

\[
X_f(V, T) = \frac{x_f}{\epsilon} = \mathcal{O}(1),
\]

which means that \( X_f \) is finite for any \( 0 < V < +\infty \) and \( 0 < T < +\infty \). Based on the continuous property of both \( Z_{21} \) and its first order derivative w.r.t. \( X_1 \), it is clear that if we adopt the approximation of \( x_f \) derived previously as the free boundary here, then the option price \( Z_2 \) satisfies the free boundary conditions in \( \mathcal{O}(\tau) \) sense, which is almost the same like the previous two cases, see (7.3.25)-(7.3.26) and (7.3.30)-(7.3.31).

Secondly, if \( \bar{x}_f \) is located inside \( U(0, \epsilon) \), it is meaningless to derived its actual form, since \( x_f = \mathcal{O}(\epsilon) \) has already been a good approximation for \( \bar{x}_f \).

Based on case I to case III, it can be concluded that when \( D < r \), (7.3.16) can be used as an approximation for small \( \tau \) and \( 0 \leq v < \infty \). Written in original variables, we obtain the leading order term of the optimal exercise price as:

\[
S_f(v, t) = K - K \sqrt{v(T_E - t)} \sqrt{\frac{v}{8\pi(T_E - t)(r - D)^2}}, \quad D < r.
\]

(7.3.33)
For $D = r$, (7.3.17) is valid for small $\tau$ and $\epsilon \leq v < \infty$. Therefore,

$$S_f(v, t) = K - K\sqrt{2v(T_E - t)}\sqrt{\ln \frac{1}{4\sqrt{\pi}(T_E - t)D}}, \quad D = r. \quad (7.3.34)$$

It is quite interesting to notice that the leading order terms of the optimal exercise price are similar to the ones with constant volatility [33], with $\sigma^2$ being substituted by $v$. One possible reason is that the moving boundary only occurs along the $S$ direction, and no critical points appear along the $v$ direction. Also, in case I and case II, the impact of $v$ is less significant, so that $v$ becomes a parameter rather than a variable, see (D.1.1)-(D.1.3). However, the option prices are much more complicated and totally different than those with constant volatility, see (D.1.5), (D.1.11) and (D.1.15).

### 7.4 Numerical results

In this section, we compare the estimated optimal exercise price with those obtained from numerical simulations based on a predictor-corrector finite difference scheme [119]. We expect our approximation to be useful for options with tenor in the order of days and weeks.

Depicted in Fig 7.2 are the optimal exercise price as a function of the time to expiry $T - t$ with different fixed values of $v$. Here, the numerical results refer to those obtained by using the predictor-corrector scheme on some fine grids, while the estimated values are obtained from (7.3.33). It is clear that, for reasonably short maturities, our approximation agrees well with those numerical results, as one would have expected. To show this in a more quantitative base, we have calculated the accuracy of our approximation, which is shown in Table 7.1, assuming that the numerical results based on some fine grids are accurate. This assumption is quite reasonable, since the convergence of the predictor-corrector scheme has already been tested by Zhu and Chen [119]. Note that the accuracy here is measured by the relative
Figure 7.2: Comparison of the numerical results and the estimated values. Model parameters are $r = 0.1$, $\sigma = 0.2$, $\rho = 0.1$, $\eta = 0.16$, $\kappa = 1.5$, Strike price $K = $10.0.
error, which is defined as
\[
\text{Error} = \frac{|S_f(v, \tau) - \tilde{S}_f(v, \tau)|}{|S_f(v, \tau)|},
\]
and
\[
\text{Overall error} = \frac{\|S_f(\cdot, \tau) - \tilde{S}_f(\cdot, \tau)\|_2}{\|S_f(\cdot, \tau)\|_2},
\]
for fixed values of \(v\) and for all \(v\) across the whole computational domain, respectively. Here, \(\| \cdot \|_2\) denotes the \(L_2\) norm, and \(S_f, \tilde{S}_f\) stand for the numerical solutions and the estimated values, respectively. From Table 7.1, it can be seen that for the maturities in the order of days and weeks, the overall errors are all very small, which demonstrates that our estimation is quite reasonable.

### Table 7.1: Report on the relative error

<table>
<thead>
<tr>
<th>Time to expiry (year)</th>
<th>Error at (v = 0.05)</th>
<th>Error at (v = 0.1)</th>
<th>Error at (v = 0.2)</th>
<th>Error at (v = 0.25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.002</td>
<td>3.445 \times 10^{-2}%</td>
<td>2.2791 \times 10^{-2}%</td>
<td>4.927 \times 10^{-2}%</td>
<td>6.1511 \times 10^{-2}%</td>
</tr>
<tr>
<td>0.005</td>
<td>0.16%</td>
<td>0.11%</td>
<td>1.727 \times 10^{-2}%</td>
<td>9.3112 \times 10^{-2}%</td>
</tr>
<tr>
<td>0.01</td>
<td>0.36%</td>
<td>0.26%</td>
<td>1.0267 \times 10^{-2}%</td>
<td>0.16%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time to expiry (year)</th>
<th>Error at (v = 0.5)</th>
<th>Error at (v = 0.6)</th>
<th>Error at (v = 0.7)</th>
<th>Overall error</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.002</td>
<td>0.18%</td>
<td>0.26%</td>
<td>0.47%</td>
<td>0.45%</td>
</tr>
<tr>
<td>0.005</td>
<td>0.54%</td>
<td>0.73%</td>
<td>0.94%</td>
<td>0.77%</td>
</tr>
<tr>
<td>0.01</td>
<td>1.05%</td>
<td>1.43%</td>
<td>1.84%</td>
<td>1.49%</td>
</tr>
</tbody>
</table>

### 7.5 Conclusion

In this chapter, we investigate the near-expiry behavior of the optimal exercise price for American put options on a dividend-paying underlying with stochastic volatility. Based on the method of matched asymptotic expansions, explicit analytical expressions for the optimal exercise price and for the values of American puts are derived. It turns out that the option prices are quite different from the corresponding Black-Scholes’ case, while the leading order term of the optimal exercise price remains almost same as the constant volatility case if the spot volatility is given the same value as the constant volatility appearing in the Black-Scholes model. Numerical experiments suggest that the current approximation is reasonably accurate for options with tenor in the order of days and weeks.
Chapter 8

Should an American option be exercised earlier or later if volatility is not assumed to be a constant?

8.1 Introduction

In the previous three chapters, we considered the asymptotic approximations for short-tenor option derivatives. In this chapter, the pricing of long-tenor options with stochastic volatility (SV), in particular, the perpetual American puts under the Heston model, will be considered.

Perpetual American options, which are options that can be exercised at any time but with no expiry, are quite important in today’s quantitative finance research. For instance, to study quantitatively the effect of a stochastic volatility term on the pricing of American options, it is better to consider those with maturities longer than one year, since the pricing impact of a stochastic volatility on options that last less than one year is fairly small in absolute terms, and becomes progressively larger as the life span of the option increases [57]. Mathematically, it is even better to assume
that the tenor of the option approaches infinity, because in this so-called “perpetual” case, the most significant effect of the stochastic volatility on the pricing of American options should be involved. In addition to this, the resulting option price can be adopted as an approximation for long-dated American options. Moreover, analyzing such kind of options may in principle be used as a building block in an approximation procedure for American options with finite maturities [15].

In the literature, the pricing of perpetual American options has been such a hot topic and is pursued by a number of authors [12, 44, 45, 79]. Most remarkably, Merton [79] derived a closed-form solution for perpetual American puts under the BS framework with constant volatility. This result has been very useful for anyone who wants to use perpetual American options to approximate long-dated options. However, his approach cannot be extended directly to the Heston model, primarily due to the fact that the optimal exercise price now becomes a function of the volatility, rather than a constant as in the BS case. As a result, no analytical method is hitherto available for the valuation of perpetual American options under the Heston model, and useful results are only obtained numerically [123]. The lack of analytical formula has made the quantitative analysis of the pricing impact of the stochastic volatility on perpetual American put options almost impossible.

In this chapter, by assuming that the volatility is slowly varying, we derive two analytical approximation formulae for perpetual American puts under the Heston model. With the utilization of the perturbation methods, which have been widely used for a number of option pricing problems [71, 117], the solution process has been greatly simplified. To check the validity of the newly-derived analytical formulae, we have compared the results obtained from these formulae with those produced by a purely numerical approach [123]. Numerical experiments show that the newly-derived analytical formulae have produced fast and reasonably accurate numerical values for both option price and the optimal exercise price of a perpetual American put option, within the validity of the assumption we have made for the asymptotic expansion. Based on the new approximations, the quantitative effect of the stochastic volatility
on the optimal exercise strategies of perpetual American options is also discussed. It turns out that for the perpetual case, if the spot variance is below a special cut-off value characterized by the given parameters, an American option priced under the Heston model should be held longer than the case of the same option priced under the traditional BS model, should the price of the underlying keep falling. A clear advantage of the analytical formulae is that the exact cut-off value has been worked here, whereas it was only observed to be approximately equal to the long-term mean of volatility in [123].

Before closing this section, we should remark that our assumption of the slowly-varying volatility is indeed quite reasonable. It is somehow related to the Born-Oppenheimer approximation which, after its introduction in 1927 for the quantum mechanical treatment of molecules, has been widely used for a great variety of physical applications [11]. In statistical mechanics, the Born-Oppenheimer approximation is especially suitable for the so-called “adiabatic elimination of the fast variables”. In finance, the slowly-varying volatility has indeed been observed and based on market observations several interesting articles have already been published [43, 85, 103]. In the so-called “slowly-varying volatility” case, the fast fluctuating property of the volatility has been eliminated, due to the fact that the changes of the volatility are not as significant as those of the underlying prices [37]. In fact, for the case of slowly-varying volatility, the newly-derived approximations can be viewed as a correction to the classical Merton formulae for perpetual American put options under the BS model with a constant volatility [79], because the assumption of “slowly-varying volatility” is essentially equivalent to treating the spot volatility as a constant plus a tiny perturbation.

We should also remark, in passing, that the validity of the newly-derived analytical approximation formulae is actually larger than what our initial assumption dictates; we shall show, at the end of Section 8.2, that the restriction on the “volatility of volatility” can be somewhat relaxed, without affecting the final appearance of the newly-derived analytical formulae. The realization of this point, prompted by
an anonymous referee’s report of [122], has made the range of applicability of our formulae wider than it was initially aimed for.

The chapter is organized as follows. In Section 8.2, we introduce the PDE (partial differential equation) system that the price of a perpetual American put option must satisfy under the Heston model, and derive an approximation solution by the perturbation method. In Section 8.3, we study the validation of the current formulae, and also compare them with the corresponding BS values. Concluding remarks are given in the last section.

8.2 Perpetual American puts under the Heston model

In the Heston model [51], the underlying $S_t$, as a function of time, is assumed to follow the SDE (stochastic differential equation) of a geometric Brownian motion:

$$dS_t = \mu S_t dt + \sqrt{\nu_t} S_t dw_1, \quad (8.2.1)$$

where $\mu$ is the drift rate, $w_1$ is a standard Brownian motion, and $\sqrt{\nu_t}$ is the standard deviation of the stock returns $\frac{dS_t}{S_t}$. Furthermore, the variance $\nu_t$ (the square of the volatility) is assumed to be governed by the following mean-reverting SDE:

$$d\nu_t = \kappa (\eta - \nu_t) dt + \sigma \sqrt{\nu_t} dw_2, \quad (8.2.2)$$

which is known in financial literature as the CIR (Cox-Ingersoll-Ross) process or the Feller process [35, 39]. For SDE (8.2.2), the following notes should be made: $\rho \in [-1, 1]$ is the correlation factor of the above two stochastic processes. $\kappa$ measures the rate with which the process $\nu_t$ reaches $\eta$, which is the long-term mean of the process $\nu_t$. In other words, $\frac{1}{\kappa}$ is the time scale of this process, meaning that it reverts to the mean value $\eta$ over times of order $\frac{1}{\kappa}$. Small values of $\kappa$ correspond to slow mean reversion, and large values of $\kappa$ correspond to fast mean reversion. $\sigma$ is the so-called
volatility of volatility. It provides the relative rate with which the spot volatility \((\sqrt{v_t})\) of the underlying moves.

We should remark here that in the following work, the volatility process is assumed to be slowly varying, with both \(\kappa\) and \(\sigma\) in the order of \(\epsilon\), where \(\epsilon\) is a small positive parameter. It should be pointed out that in the literature, \(\kappa\) is usually assumed to be very large in most of the studies (e.g. [37]). However, the slow-mean-reverting volatility was indeed found through the study of the low-frequency data, with the data necessarily ranging over times in the order of years [40]. Furthermore, some authors also assumed that the reversion-rate is quite slow, with \(\frac{1}{\kappa}\) in the order of years, and produced several interesting papers [43, 85, 103]. On the other hand, volatility of volatility is shown to be very small for an extremely long period, in the order of \(10^{-5}/\text{day}\) \((10^{-2}/\text{year})\) [23]. Therefore, it is reasonable to spend effort deriving analytical pricing formulae for the case of a slowly-varying volatility process, as is the focus of this chapter.

Let \(P_A(S, v)\) denote the value of a perpetual American put option, with \(S\) being the underlying and \(v\) being the variance. Then, under the proposed processes (8.2.1)-(8.2.2) and the additional assumption that the volatility is slowly varying, it is easily shown that the valuation of a perpetual American put option can be formulated as a free boundary problem [97], with the boundary location itself being part of the solution of the problem. In particular, as shown in [123], the PDE system is

\[
\begin{align*}
\frac{1}{2}vS^2 \frac{\partial^2 P_A}{\partial S^2} + \rho \epsilon \sigma vS \frac{\partial^2 P_A}{\partial S \partial v} + \frac{1}{2} \epsilon^2 \sigma^2 v^2 \frac{\partial^2 P_A}{\partial v^2} + rS \frac{\partial P_A}{\partial S} + \epsilon \kappa (\eta - v) \frac{\partial P_A}{\partial v} - rP_A &= 0, \\
\lim_{S \to \infty} P_A(S, v) &= 0, \\
P_A(S_f(v), v) = K - S_f(v), \quad \frac{\partial P_A}{\partial S}(S_f(v), v) &= -1, \\
\lim_{v \to 0} P_A(S, v) &= 0, \quad \lim_{v \to \infty} P_A(S, v) = K,
\end{align*}
\]

(8.2.3)

where \(\tilde{\kappa} = \frac{\kappa}{\epsilon}\) and \(\tilde{\sigma} = \frac{\sigma}{\epsilon}\) are assumed to be \(O(1)\) parameters in this chapter.

One should notice that, once the stochastic volatility is taken into consideration, the valuation of perpetual American puts is no longer as analytically achievable
as the constant volatility case under the BS framework [79], because the optimal exercise price now becomes an unknown function of the volatility, whereas under the BS model, it is only an unknown constant. However, based on the slowly-varying volatility assumption, an approximation can still be derived.

In order to facilitate the analysis, we first rewrite (8.2.3) by using the following new variables: \( x = \ln(S/K) \), \( P = P_A/K \), and obtain:

\[
\begin{align*}
\mathcal{L}_0 P + \epsilon \mathcal{L}_1 P + \epsilon^2 \mathcal{L}_2 P &= 0, \\
\lim_{x \to \infty} P(x, v) &= 0, \\
P(X_f(v), v) &= 1 - \exp(X_f(v)), \\
\frac{\partial P}{\partial x}\left(X_f(v), v\right) &= -\exp(X_f(v)), \\
\lim_{v \to 0} P(x, v) &= 0, \\
\lim_{v \to \infty} P(x, v) &= 1,
\end{align*}
\]

(8.2.4)

where the operators \( \mathcal{L}_0, \mathcal{L}_1, \) and \( \mathcal{L}_2 \) are defined as:

\[
\begin{align*}
\mathcal{L}_0 &= \frac{1}{2} v \frac{\partial^2}{\partial x^2} + (r - \frac{1}{2} v) \frac{\partial}{\partial x} - rI, \\
\mathcal{L}_1 &= \rho \tilde{\sigma} v \frac{\partial^2}{\partial x \partial v} + \tilde{\kappa}(\eta - v) \frac{\partial}{\partial v}, \\
\mathcal{L}_2 &= \frac{1}{2} \eta^2 v \frac{\partial^2}{\partial v^2}.
\end{align*}
\]

It is quite interesting to notice that the operator \( \mathcal{L}_0 \) is nothing but the perpetual BS operator with a constant volatility being equal to the spot volatility \( \sqrt{v} \).

As in the standard asymptotic analysis, we explore the solution of (8.2.4) in a series form as

\[
P(x, v) = \sum_{n=0}^{\infty} \epsilon^n P_n(x, v),
\]

(8.2.5)

\[
X_f(v) = \sum_{n=0}^{\infty} \epsilon^n X_n(v),
\]

(8.2.6)

where \( P(x, v) \) and \( X_f(v) \) are defined in (8.2.4) as the normalized option price and the logarithm of the normalized optimal exercise price, respectively. By substituting
(8.2.5) into the governing equation contained in (8.2.4), we obtain

\[ \sum_{n=0}^{\infty} \epsilon^n \mathcal{L}_0 P_n(x, v) + \sum_{n=0}^{\infty} \epsilon^{n+1} \mathcal{L}_1 P_n(x, v) + \sum_{n=0}^{\infty} \epsilon^{n+2} \mathcal{L}_2 P_n(x, v) = 0. \]  

(8.2.7)

Also, substituting (8.2.6) into the free boundary conditions yields

\[ \sum_{n=0}^{\infty} \epsilon^n P_n \bigg|_{x=\sum_{n=0}^{\infty} \epsilon^n X_n(v)} = 1 - \exp\left(\sum_{n=0}^{\infty} \epsilon^n X_n(v)\right), \]  

(8.2.8)

\[ \sum_{n=0}^{\infty} \epsilon^n \frac{\partial P_n}{\partial x} \bigg|_{x=\sum_{n=0}^{\infty} \epsilon^n X_n(v)} = - \exp\left(\sum_{n=0}^{\infty} \epsilon^n X_n(v)\right). \]  

(8.2.9)

Now, by truncating both sides of (8.2.7)-(8.2.9) to the order of \( O(\epsilon^2) \), we obtain

\[ \mathcal{L}_0 P_0 + \epsilon(\mathcal{L}_0 P_1 + \mathcal{L}_1 P_0) + \cdots = O(\epsilon^2), \]  

(8.2.10)

\[ P_0(X_0, v) + \epsilon\left(\frac{\partial P_0}{\partial x}(X_0, v) + P_1(X_0, v)\right) = 1 - \exp(X_0) - \epsilon X_1 \exp(X_0) + O(\epsilon^2), \]  

(8.2.11)

\[ \frac{\partial P_0}{\partial x}(X_0, v) + \epsilon\left(\frac{\partial^2 P_0}{\partial x^2}(X_0, v) + \frac{\partial P_1}{\partial x}(X_0, v)\right) = - \exp(X_0) - \epsilon X_1 \exp(X_0) + O(\epsilon^2). \]  

(8.2.12)

The treatment for the other boundary conditions is quite trivial, and is thus omitted here.

**Zeroth Order solution**

It is now straightforward to show that the zeroth order terms \( P_0 \) and \( X_0 \) satisfy the following PDE system:

\[
\begin{align*}
\mathcal{L}_0 P_0 &= 0, \\
\lim_{x \to \infty} P_0(x, v) &= 0, \\
P_0(X_0(v), v) &= 1 - \exp(X_0(v)), \\
\frac{\partial P_0}{\partial x}(X_0(v), v) &= - \exp(X_0(v)).
\end{align*}
\]

(8.2.13)

Clearly, the spot variance \( v \) no longer needs to be treated as a variable in (8.2.13)
because the operator $\mathcal{L}_0$ is nothing but the Black-Scholes operator, involving no partial differentiation w.r.t. $v$ at all. Consequently, the solution of (8.2.13) can be easily worked out as

$$P_0(x, v) = \frac{1}{1 + a} \left[ \frac{1 + a}{a} \exp(x) \right]^{-a},$$

$$X_0(v) = \ln \frac{a}{1 + a},$$

where $a = \frac{2r}{v}$ is the relative interest rate of the spot volatility of the underlying to the risk-free interest rate, a concept similar to $\gamma = \frac{2r}{\sigma^2}$ appearing in the classical BS model [115]. In fact, this solution is indeed identical to Merton’s [79] well-known solution under the BS model, had the spot variance $v$ be replaced by the square of volatility $\sigma^2$ used in Merton’s case. Such a degeneration back to the Black-Scholes system is expected at the zeroth order and is quite reasonable, because when volatility is assumed to slowly vary around spot volatility $\sqrt{v}$, and in the limit that the level of variation approaches to zero, $\nu_t$ should be “frozen” at its initial value $v$. Then, our solution at the next order, presented below, will be a correction to Merton’s constant volatility case, when a stochastic but slowly-varying volatility is allowed.

**First Order solution**

Continuing the process of equating terms in (8.2.10)-(8.2.12) by setting the coefficients in front of $\epsilon$ to zero, we obtain

$$\begin{align*}
\mathcal{L}_0 P_1 &= -\mathcal{L}_1 P_0, \\
\lim_{x \to \infty} P_1(x, v) &= 0, \\
P_1(X_0, v) &= -X_1 \frac{\partial P_0}{\partial x}(X_0, v) - X_1 \exp(X_0), \\
\frac{\partial P_1}{\partial x}(X_0, v) &= -X_1 \frac{\partial^2 P_0}{\partial x^2}(X_0, v) - X_1 \exp(X_0).
\end{align*}$$

(8.2.14)

Since $\frac{\partial P_0}{\partial x}(X_0, v) = -\exp(X_0)$, we have $P_1(X_0, v) = 0$. It can be easily observed that in (8.2.14), the spot variance $v$ can still be treated as if it were a constant. On the other hand, one can easily see that the problem formulated for $P_1$ is no longer a free
Chapter 8.

boundary problem, because $X_0(v)$ is known after we have found $P_0$. Furthermore, once $P_1$ is found, the first order correction of the free boundary can be explicitly obtained via

$$X_1(v) = -\frac{\partial P_0}{\partial x}(X_0, v) \frac{\partial^2 P_0}{\partial x^2}(X_0, v) + \exp(X_0),$$

(8.2.15)

once $P_1$ is found from (8.2.14) after a simple elimination of $X_1(v)$. Such an easy decouple of the unknowns in (8.2.14) and the disappearance of the “free boundary” at the first order have considerably facilitated the solution process for $P_1$, although we have to deal with an inhomogeneous ODE (ordinary differential equation) system.

The inhomogeneous ODE system that $P_1$ needs to satisfy is characterized with the same differential operator $L_0$ applied to $P_1$ while the operator that does contain differentiations w.r.t. $v$, is applied to the already-known function $P_0$, producing two known inhomogeneous terms, through which the correction to the zeroth-order optimal exercise price, $X_0(v)$, is made. This is very interesting as it implies that one only needs to deal with the moving boundary implicitly at the zeroth order and then all the corrections to the zeroth-order moving boundary can all be explicitly dealt with at subsequent higher orders. Since the solution at the zeroth order has already been found by Merton [79], the current approach preserves the elegance of Merton’s approach, while having successfully avoided using it again at higher orders where the corrections are functions of $v$ and Merton’s approach can no longer be used even if one wishes to do so.

More specifically, we first find the general solution of the governing ODE, and then determine the coefficients utilizing the boundary conditions at the first order. In the solution process at this order, the spot variance $v$ can still be treated as a parameter. After some simple algebraic manipulations, we have found $P_1(x, v)$ as

$$P_1(x, v) = -\frac{\exp(-ax)}{v(1+a)^2} \int_{X_0}^{x} 2(D + C + aD + Cy + aCy) dy.$$
where

\[ C = \left( \frac{1 + a}{a} \right)^{-a-1} \left[ \rho \tilde{\sigma} a - \frac{\tilde{\kappa} (\eta - v)}{v} \right], \]

\[ D = \left( \frac{1 + a}{a} \right)^{-a-1} \left[ \rho \tilde{\sigma} (a \ln \frac{1 + a}{a} - 1) - \frac{\tilde{\kappa} (\eta - v)}{v} \ln \frac{1 + a}{a} \right]. \]

Then, a straightforward calculation yields

\[
\frac{\partial P_1}{\partial x}(X_0, v) = \frac{2 \exp(-ax)}{v(1 + a)^2} (D + C + aD + C \ln \frac{a}{1 + a} + aC' \ln \frac{a}{1 + a})
\]

\[
= \frac{2a}{v^2(1 + a)^3} [\rho \sigma v + \kappa (\eta - v)],
\]

which can be easily substituted into (8.2.15) to produce

\[
X_1(v) = -\frac{2}{v^2(1 + a)^3} [\rho \tilde{\sigma} v + \tilde{\kappa} (\eta - v)]. \tag{8.2.16}
\]

Of course, the above solution procedure can be repeated, in theory, until any desired higher-order terms are found. However, it is quite common that in a perturbation analysis, the tediousness and cumbersomeness of algebra involved in carrying out higher-order terms progress in a way nonlinearly proportional to the number of higher-order terms to be included. It is a common practice to balance the algebra involved and the further improvement of the results obtained with higher-order terms being included. Theoretically, the approximation consisting of the zeroth-order and first-order solutions should be accurate to the order of \( \epsilon^2 \), i.e., \( 10^{-4} \) in the current work, and deriving one more terms would only contribute \( O(10^{-4}) \) to the accuracy of the approximation made here. Financially, \( O(10^{-4}) \) is very small, and can be ignored for the price of a variety of option derivatives traded in many countries. Therefore, from practical point of view, it also suffices to terminate the perturbation analysis at the order of \( \epsilon \), as long as one uses the formulae presented in this chapter with the mean-revision rate \( \kappa \) being kept roughly in \( O(10^{-2}) \).

Now, rewriting (8.2.5) and (8.2.6) in the original variables, we obtain two analytic
approximation formulae for the option price

\[ P_A(S, v) = \frac{K^{(1+2r/v)}v}{v + 2r} \left[ \frac{S(v + 2r)}{2r} \right]^{2r/v} - \frac{K^{(1+2r/v)}vS^{-2r/v}}{(v + 2r)^2} \int_{\ln 2r/(v + 2r)}^{\ln S/K} 2g(y)dy, \]

(8.2.17)

and the associated optimal exercise price

\[ S_f(v) = \frac{2rK}{v + 2r} \exp\{-\frac{2v}{(v + 2r)^3}\left[\rho \sigma v + \kappa(\eta - v)\right]\}, \]

(8.2.18)

where

\[ g(y) = \left(\frac{v + 2r}{2r}\right)^{-2r/v - 1}\left[\rho \sigma \left(\frac{2r}{v} \ln \frac{v + 2r}{2r} - 1\right) - \frac{\kappa(\eta - v)}{v} \ln \frac{v + 2r}{2r}\right] \left(\frac{v + 2r}{v}\right) \]

\[ + (1 + y + \frac{2r}{v}y)\left(\frac{v + 2r}{2r}\right)^{-2r/v - 1}\left[\frac{2\rho \sigma v - \kappa(\eta - v)}{v}\right]. \]

It should be pointed out that the newly-derived analytic formulae (8.2.17)-(8.2.18) can be viewed as a correction to the classical Merton’s formulae for perpetual American put options under the constant volatility model [79]. This is because in the current work, although the volatility follows a stochastic process, it is further assumed to be slowly varying, which is equivalent to treating the volatility as a constant with a tiny perturbation. As a result, the leading order terms are exactly the same as the corresponding ones under the BS model with the constant volatility being replaced by the spot volatility \( \sqrt{v} \), as shown in (8.2.17) and (8.2.18) by taking the limit of \( \kappa \) and \( \sigma \) to zero, while the first order terms are the ones that are responsible for the corrections needed with the tiny perturbation being taken into consideration. On the other hand, it can be clearly seen that the newly-derived formulae are explicit in terms of the variables \( S, v \) and other given parameters. Therefore, an analytical analysis of the pricing impact of the stochastic volatility on perpetual American puts can be easily achieved, which will be discussed in the next section.

Furthermore, an anonymous referee’s report of [122] made us realize that our solution is actually more general than the case restricted by the underlying assump-
tions we have initially made to restrict the order of $\kappa$ and $\sigma$ to be the same. If we relax the assumption on $\sigma$ and let it be of a larger order $\sqrt{\epsilon}$, while $\kappa$ is still kept in the order of $\epsilon$ and go through the derivation again, the results, truncated to the second order, are amazingly simple; all one needs to do is to set $\kappa$ to zero in the final expressions (8.2.17) and (8.2.18) for this case, while keeping in mind that the accuracy of the approximate solutions now becomes $O(\epsilon)$. In fact, whether $\kappa$ is set to zero or not will not affect such an order of accuracy, because $\kappa$ is in the order of $\epsilon$ and thus the terms involving $\kappa$ deem to be higher-order terms in (8.2.17) and (8.2.18) anyway for this new case. Consequently, we recommend that (8.2.17) and (8.2.18) be used for the new case as well if an approximation of an accuracy to the order of $\epsilon$ is sought. Or, in other words, formulae (8.2.17) and (8.2.18) can be used when the order of $\sigma$ varies from $O(\kappa)$ to $O(\sqrt{\kappa})$ with $\kappa$ being a sufficiently small number. The validity of this recommendation will also be confirmed by our numerical experiments shown in the next section.

8.3 Numerical tests and discussions

In this section, we shall present some numerical results and carry out some discussions on the newly-derived analytical formulae. The section contains two subsections. The first subsection is to compare the results obtained with the newly-derived analytical approximation formulae with those obtained by an SC (spectral-collocation) method, while the second subsection is to discuss the impact of adding a stochastic volatility process in pricing perpetual American options in comparison with the case where volatility is assumed to be a constant.

8.3.1 Validation of the approximation

To illustrate the reliability of the current approximation formulae, we compare the option prices and the optimal exercise price calculated from them to those calculated by an SC method, assuming that the numerical results are of high accuracy. In fact, this assumption is quite reasonable, as demonstrated in [123].
Tables 8.1-8.2 display the comparisons of the option prices and the optimal exercise price computed by the two different methods, respectively. In Table 8.1, the option prices are calculated around the money ($8 \leq S \leq 11$) at the spot variance levels $v = 0.01$, $v = 0.245$ and $v = 0.495$, while in Table 8.2, the optimal exercise prices with variance level ranging from $[0.005, 0.44]$ are shown. It can be clearly seen from these tables that, for the option prices, the maximum point-wise difference arising from the two approaches is only around 0.03, while for the optimal exercise prices, the difference is also reasonably small. Therefore, our approximation is indeed quite accurate, at least for this particular set of parameters.

Table 8.1: Comparison of option prices with different spot variance values. Parameters are $\rho = 0.1$, $\sigma = 0.04$, $r = 0.1$, $\kappa = 0.05$, $K = $10.0

<table>
<thead>
<tr>
<th>Spot variance value $v$</th>
<th>Asset prices</th>
<th>$8$</th>
<th>$9$</th>
<th>$10$</th>
<th>$11$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC method</td>
<td></td>
<td>2</td>
<td>1</td>
<td>0.2985</td>
<td>0.1304</td>
</tr>
<tr>
<td>Analytic formula</td>
<td></td>
<td>2</td>
<td>1</td>
<td>0.2905</td>
<td>0.1060</td>
</tr>
<tr>
<td>Spot variance value $v$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$v = 0.245$</td>
<td></td>
<td>8</td>
<td>9</td>
<td>3.4065</td>
<td>3.0846</td>
</tr>
<tr>
<td>SC method</td>
<td></td>
<td>3.3917</td>
<td>3.0663</td>
<td>2.8800</td>
<td>2.5781</td>
</tr>
<tr>
<td>Analytic formula</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spot variance value $v$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$v = 0.495$</td>
<td></td>
<td>8</td>
<td>9</td>
<td>4.5153</td>
<td>4.2645</td>
</tr>
<tr>
<td>SC method</td>
<td></td>
<td>4.5191</td>
<td>4.2769</td>
<td>4.0692</td>
<td>3.8884</td>
</tr>
<tr>
<td>Analytic formula</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8.2: Comparison of optimal exercise price with different spot variance values. Parameters are $\rho = 0.1$, $\sigma = 0.04$, $r = 0.1$, $\kappa = 0.05$, $K = $10.0

<table>
<thead>
<tr>
<th>$v$</th>
<th>SC method</th>
<th>Analytic formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>$v = 0.005$</td>
<td>9.6730</td>
<td>9.6461</td>
</tr>
<tr>
<td>$v = 0.01$</td>
<td>9.3860</td>
<td>9.3296</td>
</tr>
<tr>
<td>$v = 0.24$</td>
<td>4.5720</td>
<td>4.5720</td>
</tr>
<tr>
<td>$v = 0.34$</td>
<td>3.7671</td>
<td>3.7950</td>
</tr>
<tr>
<td>$v = 0.44$</td>
<td>3.2152</td>
<td>3.2343</td>
</tr>
</tbody>
</table>

On the other hand, it should be remarked again that the small parameter in the present work is related to both $\kappa$ and $\sigma$. As a result, the higher accuracy of the approximation can be only achieved with the restriction on small values of $\kappa$ and $\sigma$. Naturally, it is quite interesting to investigate how “small” $\kappa$ and $\sigma$ should be, in order to produce an acceptable approximation. We shall discuss this issue through the next experiment.
Table 8.3 shows the relationship between $\kappa$, $\sigma$, $\eta$ and the accuracy of the approximation with other parameters fixed at $\rho = 0.1$, $r = 0.1$. Note that the accuracy here is measured by the relative error, which is defined as

$$\text{Error} = \frac{\| S_f(v) - \tilde{S}_f(v) \|_2}{\| S_f(v) \|_2},$$

where $S_f(v)$ and $\tilde{S}_f(v)$ denote the optimal exercises prices computed by the SC method and our approximation, respectively, and $\| \cdot \|_2$ is the $L_2$-norm.

As pointed out previously, once the order of $\kappa$ is fixed, the assumption on $\sigma$ can be relaxed from $O(\kappa)$ to $O(\sqrt{\kappa})$. We have also tested this case for a verification. Tabulated in Table 8.3 are the results obtained with the magnitude of $\kappa$ in the order of $10^{-2}$, while the order of $\sigma$ varying from $10^{-2}$ to $10^{-1}$ (the first four rows of Panel a)). As expected, the accuracy indeed remains to be reasonable, although the approximate formulae will eventually lead to results of unacceptable accuracy if the $\sigma$ value is further increased. This has confirmed that (8.2.17) and (8.2.18) still work for $\sigma$ and $\kappa$ being in the order of $\sqrt{\epsilon}$ and $\epsilon$, respectively, as pointed out in Section 8.2.

On the other hand, compared with $\kappa$ and $\sigma$, the value of $\eta$ should be neither too big nor too small. This does make sense, since if $\eta$ is in the order higher than both $\kappa$ and $\sigma$, $\kappa \eta \frac{\partial P_0}{\partial v}$ is no longer a first order term, and should not be considered in the calculation for $P_1$. On the other hand, if $\eta$ is smaller than both $\kappa$ and $\sigma$, $\kappa \eta \frac{\partial P_0}{\partial v}$ cannot be ignored when deriving $P_0$. This leads to the conclusion that $\eta$ should be an $O(1)$ quantity before (8.2.17) and (8.2.18) can be safely used, which is verified through the data displayed in Panel b), Table 8.3. Similarly, we can examine the impact of different $\rho$ and $r$ on the accuracy of the approximation. As an overall statement, we may say that our approximation is quite accurate with the magnitude of $\kappa$ in the order of $10^{-2}$, the other parameters in the order $10^{-1}$, and $\sigma$ no more than $O(10^{-1})$.

It should be remarked that the above criteria for the safe use of (8.2.17) and
(8.2.18) are not very restrictive, especially in the case that the magnitude of \( \rho, r, \eta \) are of the order of \( 10^{-1} \), and \( \sigma \) is no more than \( \mathcal{O}(10^{-1}) \), since it is quite common to have this range of parameter-settings in financial markets [31, 106]. On the other hand, the restriction on the mean-reverting rate \( \kappa \) is in line with quite a few cases reported in the literature [43, 103]. Moreover, based on the analytical expressions of the volatility autocorrelation and the leverage functions in the Heston model [106], one can easily check that if \( \kappa \) and \( \sigma \) are in the same order (e.g., in our case \( \mathcal{O}(10^{-2}) \)), the leverage and the volatility autocorrelation have a very similar decay with characteristic times of the same order, a situation that can be found in most of the studies [40, 85]. If \( \sigma \) is in the order lower than \( \kappa \), which also makes the current formulae valid, the leverage decays faster than the volatility autocorrelation. This agrees with one of the statistical properties of the financial markets that in clear contrast with the price changes with negligible autocorrelations, the volatility autocorrelation should be still significant for a long time [85]. Therefore, it can be concluded that our restriction on the parameters does not devalue the usefulness of the newly-derived formulae for the long-dated options in financial markets.

### Table 8.3: Accuracy of the approximation

<table>
<thead>
<tr>
<th></th>
<th>( \kappa )</th>
<th>( \sigma )</th>
<th>( \eta )</th>
<th>error</th>
<th>( \kappa )</th>
<th>( \sigma )</th>
<th>( \eta )</th>
<th>error</th>
</tr>
</thead>
<tbody>
<tr>
<td>a)</td>
<td>0.05</td>
<td>0.04</td>
<td>0.2</td>
<td>1.97%</td>
<td>0.08</td>
<td>0.04</td>
<td>0.2</td>
<td>2.82%</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.08</td>
<td>0.2</td>
<td>1.90%</td>
<td>0.08</td>
<td>0.08</td>
<td>0.2</td>
<td>2.44%</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.16</td>
<td>0.2</td>
<td>2.59%</td>
<td>0.08</td>
<td>0.16</td>
<td>0.2</td>
<td>2.26%</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.32</td>
<td>0.2</td>
<td>5.03%</td>
<td>0.08</td>
<td>0.32</td>
<td>0.2</td>
<td>4.35%</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.04</td>
<td>0.2</td>
<td>3.60%</td>
<td>0.2</td>
<td>0.04</td>
<td>0.2</td>
<td>8.09%</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.08</td>
<td>0.2</td>
<td>3.19%</td>
<td>0.2</td>
<td>0.08</td>
<td>0.2</td>
<td>8.17%</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.16</td>
<td>0.2</td>
<td>2.63%</td>
<td>0.2</td>
<td>0.16</td>
<td>0.2</td>
<td>7.20%</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.32</td>
<td>0.2</td>
<td>4.24%</td>
<td>0.2</td>
<td>0.32</td>
<td>0.2</td>
<td>7.22%</td>
</tr>
<tr>
<td>b)</td>
<td>0.05</td>
<td>0.08</td>
<td>0.05</td>
<td>2.12%</td>
<td>0.08</td>
<td>0.16</td>
<td>0.05</td>
<td>3.38%</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.08</td>
<td>0.1</td>
<td>1.53%</td>
<td>0.08</td>
<td>0.16</td>
<td>0.1</td>
<td>1.97%</td>
</tr>
<tr>
<td></td>
<td>0.05</td>
<td>0.08</td>
<td>0.3</td>
<td>3.42%</td>
<td>0.08</td>
<td>0.16</td>
<td>0.3</td>
<td>5.03%</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.16</td>
<td>0.05</td>
<td>4.37%</td>
<td>0.2</td>
<td>0.16</td>
<td>0.05</td>
<td>11.36%</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.16</td>
<td>0.1</td>
<td>2.55%</td>
<td>0.2</td>
<td>0.16</td>
<td>0.1</td>
<td>8.45%</td>
</tr>
<tr>
<td></td>
<td>0.1</td>
<td>0.16</td>
<td>0.3</td>
<td>6.02%</td>
<td>0.2</td>
<td>0.16</td>
<td>0.3</td>
<td>11.78%</td>
</tr>
</tbody>
</table>

#### 8.3.2 Impact of the stochastic volatility term

With the newly-derived formulae, we can now conduct an analysis of the impact of the stochastic volatility term on the perpetual American puts quantitatively. In Fig
8.1, we show the effect of changing the parameters appearing in the Heston model on the resulting option prices. Each plot was produced with one of the four parameters $\kappa$, $\eta$, $\sigma$ and $\rho$ being slightly altered, while the other three were kept constant. From these figures, one can clearly see that, the price of a perpetual put is highly sensitive to the mean-reversion relaxation rate $\kappa$, comparing to the changes w.r.t. other parameters. This is, however, not surprising at all since for the perpetual case, the mean-reverting characteristic dominates the volatility process, as pointed out in [123]. On the other hand, one can also observe that the option prices do not differ too much w.r.t. $\sigma$ as displayed in Fig 8.1(c). This is because under the assumption that $\sigma$ is in the order of $\epsilon$, the contribution of $\sigma$ in the current approximation formulae is through the correlation of the two SDEs (8.2.1)-(8.2.2), rather than through the noise term of the variance process itself. In other words, the second-order derivative term w.r.t. $\nu$ in the PDE of (8.2.3), which represents the noise term of the variance process, does not show up in the first two orders of analysis at all; the contribution of $\sigma$ is through the cross-derivative term in the PDE, which depends on the correlation of the two stochastic processes (8.2.1)-(8.2.2). Consequently, if the underlying asset price and the variance are only slightly correlated (as in our case, $\rho = -0.1$), the option prices will not change significantly w.r.t. $\sigma$. Similar arguments can explain why the option prices are not highly sensitive w.r.t. $\eta$ and $\rho$, as shown in Fig 8.1(b) and Fig 8.1(d), respectively.

We now turn to investigate the effect of the stochastic volatility term on the early exercise strategy of a perpetual American put option. Recall that in the current work, the volatility is assumed to be slowly varying. Therefore, comparing the optimal exercise price to that of the BS model, with the volatility of the underlying being set to the spot volatility $\sqrt{\nu}$, can best reveal how the slowly-varying volatility impact upon the early exercise strategies of perpetual American puts. Of course, we should emphasize that such a comparison is only meaningful for long-term options, rather than options with short tenor, since the changes of the slowly-varying volatility
Figure 8.1: Normalized $P/K$ option price as a function of the moneyness $S/K$. 

(a) $\eta = 0.2$, $\sigma = 0.04$, $\rho = 0.1$, $r = 0.1$

(b) $\kappa = 0.02$, $\sigma = 0.04$, $\rho = -0.1$, $r = 0.1$

(c) $\kappa = 0.02$, $\eta = 0.2$, $\rho = -0.1$, $r = 0.1$

(d) $\kappa = 0.02$, $\eta = 0.2$, $\sigma = 0.04$, $r = 0.1$
Chapter 8.

should be negligible during a quite short period, but might be significant in the long run because of the “time accumulative effect”.

With the newly-derived formulae, the difference of the optimal exercise prices under the two different models can be written as:

\[
S_f^H(v) - S_{f,BS} = \frac{2rK}{v + 2r} \left( \exp \left\{ -\frac{2v}{(v + 2r)^2} [\rho \sigma v + \kappa(\eta - v)] \right\} - 1 \right). \quad (8.3.19)
\]

A straightforward calculation shows that (8.3.19) has three zeros, i.e., \( v = 0, v = \infty \), and \( v_c = \frac{\eta}{1 - \rho \sigma / \kappa} \). It is not unexpected at all to have \( v = 0 \) and \( v = \infty \) as zeros, since the boundary conditions corresponding to these two ends under the Heston model are exactly the same as the corresponding ones under the BS model in the limiting process. As a result, when the spot variance is close to any of the three zeros of (8.3.19), the difference of the optimal exercise prices under the two different models is very small, implying that the impact of the stochastic volatility on the early exercise strategies is not very significant at all. A typical functional variation of \( S_f^H - S_{f,BS} \) is displayed in Fig 8.2. Not surprisingly, from this figure, one can also observe that \( S_f^H - S_{f,BS} \) can reach two peaks at two spot variances \( v_1 \) and \( v_2 \), separating the three zeros. This in fact agrees with the Rolle’s theorem [7] that any differentiable function that attains equal value at two points, must have at least one point somewhere between them where the slope of the tangent line to the graph of the function is zero.

On the other hand, the special value \( v_c \) is a non-trivial root of \( S_f^H - S_{f,BS} = 0 \) that should be further discussed. As a combination of parameters in the parameter space of the Heston model, \( v_c \) increases with the long-term mean \( \eta \), the correlation factor \( \rho \), and volatility of volatility \( \sigma \), but decreases with the mean-reverting rate \( \kappa \). Clearly, if the spot variance \( v \) is not equal to zero and less than \( v_c \), the optimal exercise price calculated under the Heston model is lower than that calculated under the BS model, meaning that investors, who are using the Heston model to price their option would hold their option longer, assuming that the underlying price is falling.
Figure 8.2: The difference of optimal exercise prices with and without stochastic volatility

The converse is also true. To explain this from modeling point of view, we know that the optimal exercise price calculated under the Heston model is influenced by the whole dynamics of the variance process, whereas under the BS model, the optimal exercise price contains no future information of the variation of the variance at all, as a result of the constant-volatility assumption. Therefore, the special value $v_c$ can, under the current assumption of slowly-varying volatility, be viewed as some sort of average for the variance over the time. Consequently, once it is averaged out, the optimal exercise price calculated under the Heston model should be roughly the same as that calculated under the BS model with an assumed constant variance taken as $v_c$. Or mathematically, we have $S^H_f(v) \approx S^{BS}_f(v_c)$. Then, we know that under the BS framework, the optimal exercise price is monotonically decreasing with volatility, or $S^{BS}_f(v) > S^{BS}_f(v_c)$ if $v < v_c$. This then explains why the optimal exercise price calculated under the Heston model is lower than that calculated under the BS model, if the spot variance $v$ is less than $v_c$. Of course, when the “averaged” variance is lower than the spot variance now, the optimal exercise price calculated under the Heston model would exceed that calculated under the BS model, implying that those who
are using the Heston model to price their option would exercise their option earlier when the underlying price is falling.

Of course, this explanation is not in contradiction to the large-time asymptotic behavior of the variance process (8.2.2). This is because once the fluctuating property of the variance and its correlation with the underlying process are taken into consideration, it is reasonable to have a value that is only slightly different from the long-term mean $\eta$ as an average, although as time goes to infinity, the variance should approach its long-term mean asymptotically.

One should note that, for the early exercise strategies of American puts with finite maturities, the above explanation is no longer valid. That is because for finite maturity, the overall tendency of $v$ depends on several factors, such as its correlation with the asset price, its long-term mean and so on. We cannot simply determine which is the dominant factor anymore.

It should also be remarked that although the optimal exercise prices evaluated under two different models are the same at the special-cut off value $v_c$, the corresponding option prices are not. This can be easily verified by showing that

$$P^H(S, v_c) - P^{BS}(S, v_c) = -\frac{\rho \sigma K^{1+2r/v_c} S^{-2r/v_c} v_c}{v_c} \left( \ln \frac{S(2r + v_c)}{2r K} \right)^2 \left( \frac{v_c + 2r}{2r} \right)^{-2r/v_c - 1},$$

is non-zero unless $\rho = 0$. Therefore, one cannot claim that the Heston model has reduced to the BS model when $v = v_c$, although one would take the same exit strategy for this case. The only time one can do so is when both $\kappa$ and $\sigma$ approach to zero.

We can also provide a plausible reason to explain why the two option prices are different when their corresponding optimal exercise prices are the same if $v = v_c$. Under the Heston model, the option price is influenced by not only the underlying process but also the variance process as well. Suppose that the spot volatility happens to be just $\sqrt{v_c}$ at $t = 0$. It will not remain at $\sqrt{v_c}$ because of its inherent stochastic nature, and the fluctuation will certainly affect the underlying process, once the two
processes are correlated. On the other hand, under the BS model, if the constant volatility is set to $\sqrt{v_c}$, it will not change. Therefore, for the same underlying price $S$, its future dynamics depend not only on the spot volatility $\sqrt{v_c}$, but also on the correlation of the two processes (8.2.1)-(8.2.2) under the Heston model, whereas under the BS framework, it is only affected by $\sqrt{v_c}$ as a stochastic noise term. Here, we ignore the influence of the risk-free interest rate $r$, since it plays the same role in both models. Thus, it is not strange to have different option prices at the special volatility level $\sqrt{v_c}$, even when the optimal exercise prices happen to be the same. Of course, when the processes (8.2.1) and (8.2.2) are not correlated with each other at all, i.e., $\rho = 0$, then at the special cut-off value $v_c = \eta$, both the optimal exercise prices and the option prices under the two different models are the same.

8.4 Conclusion

In this chapter, we have considered the pricing of a perpetual American put option under Heston’s stochastic volatility model with the assumption that the volatility is slowly-varying. By means of the perturbation methods, two approximation formulae for the option price and the optimal exercise price are obtained, which can be viewed as corrections to Merton’s formulae for perpetual American puts under the BS model. Our numerical experiments suggest that the newly-derived formulae are quite accurate for reasonably small mean-reverting parameters. Moreover, we show that the range of applicability of our formulae is remarkably wider than it was initially aimed for, after the original assumption on the order of the “volatility of volatility” being somewhat relaxed. Based on the current approximation formulae, we have also quantitatively discussed the effect of allowing volatility to be a stochastic variable on the optimal exercise strategies of perpetual American puts, and found that if the spot volatility is below a special long-term average value, an option contract priced under stochastic volatility should be held longer than the corresponding ones with constant volatility, assuming that the price of the underlying is falling.
Chapter 9

Pricing perpetual American options under a stochastic volatility model with fast mean-reversion

9.1 Introduction

In chapter 8, we derived pricing approximations for perpetual American puts under the Heston model, with the assumption that the volatility is slowly varying. In this chapter, we shall extend the framework of Chapter 8 to a more realistic case. That is, we shall consider the pricing approximation of perpetual American puts under the fast mean-reverting volatility. The mean reversion rate of the volatility to be quite fast is one of the common features of today’s financial markets. By “mean reversion”, it refers to a linear pull back term in the drift of the volatility process itself, or in the drift of some process, of which the volatility is a function. By “fast”, it is meant that the characteristic time it takes for a process to get back to the mean level of its invariant distribution is of the order of several days. This is the so-called fast mean-reverting volatility, and has been observed in many financial markets. For example,
when analyzing the high-frequency S&P500 index data, it is found that the mean-reversion rate is typically of the order of one or two days [41]. Also, data analyses on commodity prices reveal the fast mean-reversion feature [52]. In the literature, the study of financial derivatives under a fast mean-reverting volatility has received a great deal of attention ever since the seminal work of Fouque et al. was published [42].

In the literature, the most remarkable step in studying perpetual American options is Merton’s closed-form solution for perpetual American puts with constant volatility [79]. Unfortunately, his approach cannot be extended directly to the SV models, primarily due to the fact that the optimal exercise price is now an unknown curve of the volatility, rather than just a constant as in the BS model. Recent progress was made by Zhu & Chen [122] in deriving an approximation for perpetual American puts with a slowly varying volatility. However, their assumption that the volatility is slowly varying has somehow restricted the application of their formulae. For most of the markets, where the volatility is observed to be fast mean-reverting, no useful results are hitherto available for perpetual American options. The lack of analytical formula has made the quantitative analysis of the pricing impact of the SV term on perpetual American put options less achievable.

In this chapter, the valuation of perpetual American puts is considered under a general fast mean-reverting regime of SV. Following Fouque et al.’s method [37], we obtain a sequence of simplified option pricing systems, which are so nicely constructed and can be dealt with by using the well-known Merton’s approach that was previously only available for pricing perpetual American options under the classic BS model. Based on the new approximations, the quantitative effect of a fast mean-reverting SV on the option prices as well as the optimal exercise strategies is then examined under a specific model. It turns out that the impact of a fast mean-reverting SV is quite different from that of a slowly varying volatility, as discussed in [122]. Moreover, it shows that when the option is closer to the money, the impact of a fast mean-reverting volatility becomes more significant.
Chapter 9.

The rest of this chapter is organized as follows. In Section 9.2, we introduce the PDE (partial differential equation) system that the price of a perpetual American put option must satisfy under a general fast mean-reverting volatility model. In Section 9.3, we derive an approximation solution by the perturbation method. In Section 9.4, we study the effect of the SV term, and also compare it with the corresponding result obtained when a slowly varying volatility is taken into consideration [122]. Concluding remarks are given in the last section.

9.2 Perpetual American puts under fast mean-reverting volatility

The assumption of a fast mean-reverting volatility is usually based on the OU (Ornstein-Uhlenbeck) model, which does not depend on the specific form of the volatility process. Under such a model, the underlying \( S_t \), as a function of time, is assumed to follow the SDE (stochastic differential equation) of a geometric Brownian motion:

\[
dS_t = \mu S_t dt + \sigma_t S_t dW_t,
\]

where \( \mu \) is the drift rate, \( W_t \) is a standard Brownian motion, and the volatility \( \sigma_t \) is represented by another stochastic process \( Y_t \), i.e., \( \sigma_t = f(Y_t) \), with \( f \) being a smooth positive function that is both bounded above and bounded away from zero. Here, \( Y_t \) is governed by a so-called mean-reverting process, i.e.,

\[
dY_t = \alpha(m - Y_t)dt + \sqrt{2\alpha}\nu(\rho dW_t + \sqrt{1-\rho^2}dZ_t),
\]

where \( (Z_t) \geq 0 \) is another Brownian motion independent of \( W_t \). For the second SDE, the following notes should be made, \( \rho \in [-1, 1] \) is the correlation factor of the above two stochastic processes. The parameter \( \alpha \) measures the rate with which the process \( Y_t \) reaches its long-term mean value \( m \). In other words, \( 1/\alpha \) is the time scale of this process, measuring the order of \( 1/\alpha \), with which the process \( Y_t \) reverts to its mean
value \( m \). \( \nu^2 \) is the variance of the invariant distribution of \( Y_t \), which controls the long-run level of the volatility fluctuations.

When the volatility is assumed to be fast mean reverting, it implies that the mean-reversion rate \( \alpha \) is of the order of \( 1/\epsilon \), with \( \epsilon \) being a positive small parameter. In fact, this assumption agrees well with the empirical evidence that the volatility is persistent or busty, i.e., for days at a time, it is high, and then, for a similar length of time, it is low [41]. In the literature, many authors focused on studying option derivatives in the so-called fast mean-reverting volatility scenario, and produced some interesting papers [37, 39, 41, 52].

To consider the pricing of a perpetual American option under the fast mean-reverting volatility framework, which is the main concern of the current work, the proper valuation system should be established first. Let \( P(S, y) \) denote the value of a perpetual American put option, with \( S \) being the underlying and \( y \) being the fast mean-reverting factor that drives the volatility process. Then under the risk-neutral probability \( P^\gamma \), it can be shown that the valuation of a perpetual American put option with fast mean-reverting volatility can be formulated as a free boundary problem [97], with the boundary location itself being part of the solution of the problem. In particular, the PDE system is similar to the one for the valuation of perpetual American puts with a slowly varying volatility [122], i.e.,

\[
\begin{align*}
\frac{1}{2} f^2(y) S^2 \frac{\partial^2 P}{\partial S^2} + r S \frac{\partial P}{\partial S} - r P + \frac{\sqrt{2} \nu}{\sqrt{\epsilon}} \left[ \rho f(y) S \frac{\partial^2 P}{\partial S \partial y} - \Lambda(y) \frac{\partial P}{\partial y} \right] \\
+ \frac{1}{\epsilon} (m - y) \frac{\partial P}{\partial y} + \frac{\nu^2}{\epsilon} \frac{\partial^2 P}{\partial y^2} = 0,
\end{align*}
\]

\[ (9.2.1) \]

where \( r \) is the risk-free interest rate, and the function \( \Lambda(y) \) is defined as \( \Lambda(y) = \rho (\mu - r)/f(y) + \gamma(y) \sqrt{1 - \rho^2} \), with a bounded function \( \gamma(y) \) being the market price of volatility risk.

One should notice that the boundary conditions along the \( y \) direction may also be needed for the well-posedness of (9.2.1). In case of degenerate boundary, the
corresponding boundary condition should (not) be imposed if the value of the Fichera function on the boundary (limit may need to be taken for some cases) proves to be negative (nonnegative) [36]; while for the non-degenerate case, the boundary condition should be given \textit{a priori}. For the current case, it can be shown that if the coefficient in front of $\partial^2 P/\partial y^2$, i.e., $\nu^2/\epsilon$, is not equal to zero, both $y = -\infty$ and $y = \infty$ are non-degenerate boundaries, and thus the boundary conditions should be imposed there. But, if $\nu^2/\epsilon$ vanishes, these two boundaries become degenerate, and the corresponding Fichera functions are $B_{-\infty}(y) = (m - y)/\epsilon$, $B_{\infty}(y) = (y - m)/\epsilon$.

One can then easily establish $\lim_{y \to -\infty} B_{-\infty}(y) = \infty$, $\lim_{y \to \infty} B_{\infty}(y) = \infty$, and therefore, no boundary conditions are needed along the $y$ direction for this degenerate case.

Such a dependence nature of the need (or no need) of boundary conditions along the $y$ direction on the coefficient in front of the second-order derivative of $P$ w.r.t. $y$ suggests that we shall not take the boundary conditions along the $y$ direction into consideration in our analysis here when we aim to find a general solution valid for a wide class of fast mean-reverting SV models. In other words, in the current work, we would only consider a solution that satisfies (9.2.1), but may or may not satisfy the boundary conditions along the $y$ direction, when they are required for the well-posedness of (9.2.1). Financially, the solution we try to find below is at least valid for volatility levels not being extremely high or low. It must be pointed out that the regime under which we analyze the price of the option is indeed meaningful, because in the so-called fast mean-reverting SV scenario, the volatility level fluctuates randomly around its mean level, and the epochs of high or low volatility are relatively short [42].

On the other hand, it should also be remarked that once the SV is taken into consideration, the valuation of perpetual American puts is no longer as analytically achievable as the constant-volatility case under the BS framework [79], because the optimal exercise price now remains unknown as a function of volatility, whereas under the BS model, it is only an unknown constant. However, based on the assumption that the volatility is fast mean reverting, we can still manage to derive an analytical
approximation for the price of perpetual American puts, which is shown in the next section.

9.3 Pricing approximations

In order to facilitate the analysis, we first write (9.2.1) by using the new variables $x = \ln(S/K)$, $P^\epsilon = P/K$, and obtain:

\[
\begin{aligned}
\ell^\epsilon P^\epsilon &= 0, \\
\lim_{x \to \infty} P^\epsilon(x, y) &= 0, \\
P^\epsilon(x_f^\epsilon(y), y) &= 1 - \exp(x_f^\epsilon(y)), \\
\frac{\partial P^\epsilon}{\partial x}(x_f^\epsilon(y), y) &= -\exp(x_f^\epsilon(y)),
\end{aligned}
\]

(9.3.2)

where operator $\ell^\epsilon$ is defined as: $\ell^\epsilon = \ell_0/\epsilon + \ell_1/\sqrt{\epsilon} + \ell_2$, with

\[
\begin{aligned}
\ell_0 &= (m - y) \frac{\partial}{\partial y} + \nu^2 \frac{\partial^2}{\partial y^2}, \\
\ell_1 &= \nu \sqrt{2} \rho f(y) \frac{\partial^2}{\partial x \partial y} - \Lambda(y) \frac{\partial}{\partial y}, \\
\ell_2 &= \frac{1}{2} f^2(y) \frac{\partial^2}{\partial x^2} + [r - \frac{1}{2} f^2(y)] \frac{\partial}{\partial x} - rI.
\end{aligned}
\]

It is quite interesting to notice that operator $\ell_2$ is nothing but the perpetual BS operator with a constant volatility being the spot volatility $f(y)$.

Now, following the standard asymptotic analysis, we explore the solution of (9.3.2) in the form

\[
P^\epsilon(x, y) = \sum_{n=0}^{\infty} \epsilon^{\frac{n}{2}} P_n(x),
\]

(9.3.3)

\[
x_f^\epsilon(y) = \sum_{n=0}^{\infty} \epsilon^{\frac{n}{2}} x_n(y).
\]

By substituting (9.3.3) into the governing equation contained in (9.3.2), we obtain

\[
\frac{1}{\epsilon} \ell_0 P_0 + \frac{1}{\sqrt{\epsilon}} (\ell_0 P_1 + \ell_1 P_0) + (\ell_0 P_2 + \ell_1 P_1 + \ell_2 P_0) + \sqrt{\epsilon} (\ell_0 P_3 + \ell_1 P_2 + \ell_2 P_1) = \mathcal{O}(\epsilon).
\]

(9.3.4)
Also, substituting (9.3.3) into the free boundary conditions yields

\[
P_0(x_0, y) + \sqrt{\epsilon} \left[ \frac{\partial P_0}{\partial x}(x_0, y)x_1 + P_1(x_0, y) \right] = 1 - \exp(x_0) - \sqrt{\epsilon} x_1 \exp(x_0) + O(\epsilon),
\]

\[
\frac{\partial P_0}{\partial x}(x_0, y) + \sqrt{\epsilon} \left[ \frac{\partial^2 P_0}{\partial x^2}(x_0, y)x_1 + \frac{\partial P_1}{\partial x}(x_0, y) \right] = -\exp(x_0) - \sqrt{\epsilon} x_1 \exp(x_0) + O(\epsilon).
\]

(9.3.5)

In the following, we shall first analyze (9.3.4), aiming at finding proper governing equations for \( P_0 \) and \( P_1 \). From (9.3.4), it is clear that at the lowest order, \( O(1/\epsilon) \), we have \( \mathcal{L}_0 P_0 = 0 \). At this stage, although the explicit form of \( P_0 \) is not known, we can at least deduce that \( P_0 \) does not depend on \( y \), i.e., \( P_0 = P_0(x) \). This is because operator \( \mathcal{L}_0 \) is the generator of an ergodic Markov process and acts only on the \( y \) variable, and moreover, the particular solutions that depend on \( y \) are in general ruled out by the far-field boundary condition along the \( x \) direction.

To eliminate the \( O(1/\sqrt{\epsilon}) \) term, we obtain \( \mathcal{L}_0 P_1 + \mathcal{L}_1 P_0 = 0 \), which yields \( \mathcal{L}_0 P_1 = 0 \), by using the fact that \( P_0 \) is a constant w.r.t. \( y \), and operator \( \mathcal{L}_1 \) only acts on \( y \) as well, i.e., \( \mathcal{L}_1 P_0 = 0 \). Using the same argument as that used for \( P_0 \), it is clear that \( P_1 = P_1(x) \).

At \( O(1) \), the governing equation becomes \( \mathcal{L}_2 P_0 + \mathcal{L}_1 P_1 + \mathcal{L}_0 P_2 = 0 \), and consequently,

\[
\mathcal{L}_0 P_2 + \mathcal{L}_2 P_0 = 0, \quad (9.3.6)
\]

because \( P_1 \) does not depend on \( y \). Given that \( P_0 \) is known in advance, (9.3.6) is a Poisson equation for \( P_2 \) w.r.t. operator \( \mathcal{L}_0 \) in the variable \( y \). By applying the Fredholm alternative theorem [87], it is clear that (9.3.6) has no solution unless \( \mathcal{L}_2 P_0 \) is orthogonal to the invariant distribution of the process \( Y \) whose infinitesimal generator is \( \mathcal{L}_0 \), i.e., \( \langle \mathcal{L}_2 P_0 \rangle = 0 \). Here, \( \langle \cdot \rangle \) denotes \( \int_{R_y} \cdot p_\infty dy \), with \( p_\infty \) being the invariant distribution of \( Y \), i.e., \( \mathcal{L}_0^* P_\infty = 0 \), where \( \mathcal{L}_0^* \) is the adjoint operator of \( \mathcal{L}_0 \). In our case, it is not difficult to show that \( p_\infty = 1/(\sqrt{2\pi}\nu) \exp[-(y-m)^2/(2\nu^2)] \). Now, by setting the coefficients in front of those \( O(1) \) terms in (9.3.5) to zero, we obtain
the PDE system for $P_0$ as

\[
\begin{cases}
\langle \mathcal{L}_2 P_0 \rangle = 0, \\
\lim_{x \to -\infty} P_0(x) = 0, \\
0(x_0) = 1 - \exp(x_0), \\
\partial_{x_0}(x_0) = -\exp(x_0).
\end{cases}
\] (9.3.7)

Since $P_0$ is a constant w.r.t. $y$, the governing equation contained in (9.3.7) can be further simplified as $\langle \mathcal{L}_2 \rangle P_0 = 0$. Also, from the definition of $\mathcal{L}_2$, one can deduce that $\langle \mathcal{L}_2 \rangle = \mathcal{L}_{BS}(\bar{f})$, where $\bar{f}$ is the so-called effective volatility, and is defined as the statistical average w.r.t. the invariant distribution, i.e., $\bar{f} = \langle f(y) \rangle$. It should be remarked that the effective volatility $\bar{f}$ is almost surely equal to the long-run time average of the function $f$, i.e., $\bar{f} = \lim_{t \to \infty} \frac{1}{t} \int_0^t f(Y_s) ds$, (almost surely), as a result of the ergodic theorem [37].

Now, it is quite clear that in (9.3.7), $y$ no longer needs to be treated as a variable because operator $\langle \mathcal{L}_2 \rangle$ is nothing but the perpetual BS operator, involving no partial differentiation w.r.t. $y$ at all. Consequently, the solution of (9.3.7) can be easily worked out as $P_0(x) = 1/(1 + a)[(1 + 1/a)\exp(x)]^{-a}$, $x_0 = \ln[a/(1 + a)]$, where $a = 2r/\bar{f}^2$ is the relative interest rate of the effective volatility of the underlying to the risk-free interest, a concept similar to $\gamma = 2r/\sigma^2$ appearing in the classical BS model [79]. In fact, the above solution is indeed identical to Merton’s well-known solution under the BS model, had the effective volatility $\bar{f}$ been replaced by $\sigma$ used in Merton’s case. Such a degeneration back to the BS system is expected at the zeroth-order and is a reassurance that our solution procedure is correct. From a stochastic point of view, we know that the distribution of $Y_t$ depends only on the product of the mean-reversion rate and the time, and thus when the mean-reversion rate is assumed to be very large, the distribution should be exactly the same as its large time distribution [37]. Mathematically, in the limit sense for the fast mean-reverting factor, we have $\lim_{t \to 0} f(Y_t) = \frac{1}{t} \lim_{t \to \infty} \int_0^t f(Y_s) ds$. One can now easily deduce that $\lim_{t \to 0} f(Y_t) = \bar{f}$, (almost surely), because the long-run time average of the function $f$ is almost surely equal to the effective volatility $\bar{f}$, as mentioned earlier. This also explains why the constant volatility in the BS model should be replaced by the effective volatility
in the zeroth-order solution under fast mean-reverting volatility framework. Then, our solutions at the next order, presented below, will be a correction to Merton’s constant-volatility case, when the volatility is driven by a fast mean-reverting factor.

To solve for the first order correction term $P_1$, we collect the $O(\sqrt{\epsilon})$ terms in (9.3.4), which yield $\mathcal{L}_0 P_3 + \mathcal{L}_1 P_2 + \mathcal{L}_2 P_1 = 0$. Again, by applying the Fredholm alternative theorem to this equation, we obtain $\langle \mathcal{L}_1 P_2 + \mathcal{L}_2 P_1 \rangle = 0$, which can be simplified as

$$\langle \mathcal{L}_2 \rangle P_1 = -\langle \mathcal{L}_1 P_2 \rangle,$$  

by using the fact that $P_1$ does not depend on $y$.

On the other hand, since $\mathcal{L}_2 P_0 = \mathcal{L}_2 P_0 - \langle \mathcal{L}_2 \rangle P_0$, we obtain

$$\mathcal{L}_2 P_0 = \frac{1}{2}(f^2(y) - \bar{f}^2)(\frac{\partial^2 P_0}{\partial x^2} - \frac{\partial P_0}{\partial x}),$$

which, combined with (9.3.6), yields

$$\mathcal{L}_0 P_2 = -\mathcal{L}_2 P_0 = -\frac{1}{2}(f^2(y) - \bar{f}^2)(\frac{\partial^2 P_0}{\partial x^2} - \frac{\partial P_0}{\partial x}).$$

Again using the fact that $P_0$ does not depend on $y$, we can deduce that

$$P_2 = -\frac{1}{2} \mathcal{L}_0^{-1}(f^2(y) - \bar{f}^2)(\frac{\partial^2 P_0}{\partial x^2} - \frac{\partial P_0}{\partial x}) = -\frac{1}{2}(\varphi(y) + c(x))(\frac{\partial^2 P_0}{\partial x^2} - \frac{\partial P_0}{\partial x}),$$  

(9.3.9)

where $\varphi(y)$ is the solution of $\mathcal{L}_0 \varphi(y) = f^2(y) - \bar{f}^2$, and $c(x)$ is a constant w.r.t. $y$.

Now, by substituting (9.3.9) into (9.3.8), we obtain

$$\langle \mathcal{L}_2 \rangle P_1 = \frac{1}{2} \langle \mathcal{L}_1 \varphi(y) \rangle (\frac{\partial^2 P_0}{\partial x^2} - \frac{\partial P_0}{\partial x}) = V_3 \frac{\partial^2 P_0}{\partial x^2} + (V_2 - 3V_3) \frac{\partial P_0}{\partial x} + (2V_3 - V_2) \frac{\partial^2 P_0}{\partial x^2},$$

(9.3.10)

where $V_2 = \sqrt{2}\nu(\bar{f}^2(y) - \bar{f}^2)/(2\sqrt{2}/\nu(\Lambda(y)\phi_y(y)))$ and $V_3 = \sqrt{2}\nu(\bar{f}^2(y)\phi_y(y))/2$. It should be remarked that $V_2$ and $V_3$ are nontrivial functions of the original model, but the detailed expressions of them are not necessary in the current approach. In other words, the approach adopted is able to deal with general fast mean-reverting
SV models. One should also notice that these two terms do not play a same role in (9.3.10) at all. The $V_2$ term is simply a volatility level correction, and depends on both $\rho$ and the market price of risk $\gamma$; while the $V_3$ term shows the “skew” effect due to the presence of the third-order derivative, and it only relates to $\rho$ [37].

Now, taking all the boundary conditions for $P_1$ into consideration, we find that $P_1$ satisfies

\[
\begin{cases}
\langle L_2 \rangle P_1 = D \exp(-ax), \\
\lim_{x \to \infty} P_1(x) = 0, \quad P_1(x_0) = 0, \\
\frac{\partial^2 P_0}{\partial x^2}(x_0)x_1 + \frac{\partial P_1}{\partial x}(x_0) = -\exp(x_0)x_1,
\end{cases}
\]

with $D = a(1 + 1/a)^{-a}[-(a + 2)V_3 + V_2]$.

It can be easily observed that in (9.3.11), $y$ does not appear at all as a result of taking statistical average w.r.t. the fast mean-reverting factor. On the other hand, one can see that the problem formulated above for $P_1$ is no longer a free boundary problem, because $x_0$ is known after we have found $P_0$. Furthermore, once $P_1$ is found, the first order correction of the free boundary can be explicitly calculated via $x_1 = -\frac{\partial P_1}{\partial x}(x_0)/\left[\frac{\partial^2 P_0}{\partial x^2}(x_0) + \exp(x_0)\right]$. Such an easy decoupling of unknowns in (9.3.11) and the disappearance of the “free boundary” at the first order have considerably facilitated the solution process for $P_1$, although we have to deal with an ODE (ordinary differential equation) system (9.3.11).

The inhomogeneous ODE system that $P_1$ needs to satisfy is characterized with the same differential operator $\langle L_2 \rangle$, and the source terms from the volatility level correction and the “skew” effect, through which the corrections w.r.t. the fast mean-reverting factor to the zeroth-order solutions are made. This is quite interesting as it implies that one only needs to deal with the moving boundary implicitly at the zeroth-order, and then all the corrections to the zeroth-order moving boundary w.r.t. the fast scale can be explicitly dealt with at all subsequent high orders.

More specifically, we first find the general solution of the governing ODE and then, determine the coefficients utilizing the boundary conditions at the first order.
Chapter 9.

After some simple algebraic manipulations, $P_{1,0}$ and $x_{1,0}$ are found as

$$P_1(x) = \frac{2a(1+1/a)^{-a}[-(a+2)V_3 + V_2][- \ln(1+1/a) - x] \exp(-ax)}{f^2(a+1)},$$

$$x_1 = \frac{2[-(a+2)V_3 + V_2]}{2r + f^2}.$$

It should be remarked that if the mean-reversion rate is extremely high, i.e., $\epsilon \to 0$, the above corrections are no longer needed. In other words, when $\epsilon \to 0$, the BS price with volatility being $\tilde{f}$ is already a good approximation. However, for a fast but still finite mean-reversion rate, the BS price has to be corrected due to the randomness of the volatility.

Now, by writing (9.3.3) in the original variables, we obtain the option price and the optimal exercise price of perpetual American puts under a general fast mean-reverting volatility model as

$$P = \frac{K\tilde{f}^2}{f^2 + 2r} \left[ \frac{S(f^2 + 2r)}{2r K} \right]^{-2r/f^2} - \frac{2\sqrt{2}\sqrt{\epsilon}Kr \nu}{f^2(f^2 + 2r)} \ln \frac{2r K}{S(f^2 + 2r)} \times \left[ \frac{S(f^2 + 2r)}{2r K} \right]^{-2r/f^2} \frac{2r \rho}{f^2} (f \phi_y + \langle \Lambda \phi_y \rangle),$$

$$S_f = K \exp(x_0 + \sqrt{\epsilon}x_1) \approx K \exp(x_0)(1 + \sqrt{\epsilon}x_1)$$

$$= \frac{2r K}{2r + f^2} \left[ 1 - \frac{\sqrt{2}\sqrt{\epsilon}\nu}{f^2 + 2r} \left( \frac{2r \rho}{f^2} (f \phi_y + \langle \Lambda \phi_y \rangle) \right) \right],$$

respectively.

One can easily observe that the fast mean-reverting factor $y$ does not play any explicit role in the present formulae. From a modeling point of view, this could be explained as follows. In the fast mean-reverting volatility scenario, although the volatility might fluctuate around its ergodic mean considerably over many months during the lift span of an option contract, there might also be much fluctuation of the underlying price, and the changes of the volatility are not as significant as those of the underlying price [39]. In other words, the volatility can be considered relatively constant until its next major fluctuation comes out, because the minor volatility fluctuations are insignificant comparing with the changes of the underlying price. As
a result, only the statistical average of all possible paths of $Y_t$, rather than its spot level, is involved in the current formulae.

## 9.4 Discussions

As mentioned earlier, the newly obtained formulae (9.3.12-9.3.13) are solutions of perpetual American puts under a general fast mean-reverting SV model. Consequently, with the formulae, the impact of the fast mean-reverting volatility on perpetual American puts can be analyzed quantitatively. To clearly address this issue, it is better to choose a particular SV model, where the parameters appearing in (9.3.12)-(9.3.13) can be explicitly worked out. In our analysis, the model used in [42] is adopted, where the volatility and the volatility risk premium are set to $\exp(y)$ and 0, respectively. Under this special SV model, the explicit forms of the parameters needed in (9.3.12)-(9.3.13) are:

\[
\begin{align*}
\bar{f}^2 &= \exp(2\nu^2 + 2m), \quad V_3 = \sqrt{2}\rho \exp(5\nu^2/2 + 3m)(1 - \exp(2\nu^2))/(2\nu), \\
V_2 &= \sqrt{2}\rho(\mu - r) \exp(\nu^2/2 + m)(\exp(2\nu^2) - 1)/(2\nu) + 2V_3.
\end{align*}
\]

To investigate how the fast mean-reverting volatility impacts upon the price of a perpetual American put option as well as its early exercise strategies, it is better to compare the current formulae with the corresponding ones under the BS model, with the volatility being set to the effective volatility. It is not difficult to show that the differences of the option and the optimal exercise prices under the two different models are respectively

\[
\begin{align*}
P^{SV} - P^{BS} &= \frac{\sqrt{2}\epsilon a[(1 + a)/a]^{-a} \ln(aK/(S + Sa))(S/K)^{-a}}{\nu(2r + f^2)} \\
&\quad (\mu + r) \exp(m + \nu^2/2)[\exp(2\nu^2) - 1], \\
S^{SV}_f - S^{BS}_f &= \frac{2rK\sqrt{2}\epsilon(\mu + r) \exp(m + \nu^2/2)[\exp(2\nu^2) - 1]}{\nu(2r + f^2)^2}.
\end{align*}
\]
Clearly, once the correlation factor $\rho$ is less than zero, a phenomenon often found through analyzing financial data [37], the option price under the SV model is higher than the one under the BS framework; whereas the optimal exercise price with SV is less than that with constant volatility. Financially, it indicates that the presence of the volatility tends to add value to a perpetual put option, and to postpone its early exercise time, had the underlying prices been assumed to be falling. Furthermore, a straightforward calculation reveals that the difference of the option prices under the two different models reaches the peak value at 

$$S = Ka \exp(1/a)/(1 + a) \approx K,$$ 

which is defined as “near the money” in financial terms, with

$$(P_{SV} - P_{BS})_{\text{max}} = -\sqrt{2} \rho (\mu + r) \exp(m + \nu^2/2 + 1)\exp(2\nu^2) - 1]/[\nu(2r + \bar{f}^2)].$$

This implies that the effect of the SV is quite significant particularly for those near-the-money options.

On the other hand, with the current formulae and the similar work for a slowly varying volatility case available [122], we can make a comparison between the two scenarios. Such a comparison should be quite interesting, as it reveals how different volatility dynamics will impact upon perpetual American put options. Firstly, one can observe that the zeroth-order terms of both approximations are perpetual BS formulae, but with volatility being replaced by the “spot” value and the “effective” value for a slowly varying volatility and a fast mean-reverting volatility, respectively. This is in fact a consistency between the two different cases, because the slowly varying volatility should be “frozen” at its initial value in the limit that the level of variations approaches to zero; whereas in the fast mean-reverting scenario, the volatility should be its “effective” level when the mean-reversion rate becomes extremely large. Secondly, it is found that the effect of a slowly varying volatility varies w.r.t. the spot volatility; that is, for some certain values of the spot volatility, the SV tends to add the value of the contract, but for others, it makes the contract less valuable; while for the current case, the SV tends to add the price of a put option contract universally. Finally, one should notice that the origins of the SV effects under the two volatility cases are totally different. For a slowly varying volatility, its effect on a perpetual American put option mainly comes from the “time accumulative effect”,
since the changes of the slowly varying volatility should be negligible during a quite short period, but might be significant in the long run [122]. On the other hand, the SV effect in the current fast mean-reverting volatility case is somehow identical to a “sum” of all the fast but finite fluctuations occurring during the time when the option is held.

9.5 Conclusion

By assuming that the instantaneous SV of the underlying is driven by a fast mean-reverting OU process, two explicit closed-form formulae are successfully derived for the valuation of perpetual American put options. Based on the newly obtained formulae, we analyze the quantitative effect of allowing a fast mean-reverting SV on the perpetual option price as well as its early exercise strategy. It turns out that the extra source of uncertainty associated with the fast mean-reverting SV can be quite significant for those near-the-money put options. Moreover, the presence of SV tends to add value to the put option price, and increase the optimal time to exercise the option contract, had the underlying prices been assumed to be falling.
Chapter 10

Pricing perpetual American puts under multi-scale stochastic volatility

10.1 Introduction

In the previous two chapters, we investigated separately the pricing of perpetual American options under a slowly-varying volatility and a fast mean-reverting volatility. In this chapter, we shall consider such a problem under the multi-scale stochastic volatility (SV), which is more general than the two volatility dynamics discussed in Chapter 8 and Chapter 9, respectively.

As is well known now, the classical Black-Scholes (BS) model, does not satisfactorily explain many properties of the financial markets, such as the volatility smile or smirk effect. One possible remedy is to assume that the volatility of the underlying also follows another stochastic process, rather than a constant assumed in the original BS framework. These two-factor models are referred to as SV (stochastic volatility) models, if the second stochastic process is introduced to describe the volatility of the first process for the underlying and they have become more popular for option pricing and hedging in the past decade or two. Recently, however, it is found out
that those SV models with a single time scale are still unable to capture some of the unique features observed in the financial markets.

One of the well known facts showing the inadequacy of using a single time scale suggested by several simple SV models [6, 37, 51, 58] is that the exponential fit with a single time scale fails to capture the long range memory characteristic of the volatility corrections, which is observed in any reasonable market models [25]. In specific, if the time scale is chosen to be small (a few days), the long time tail will be completely missed. On the contrary, if the time scale is chosen to be large, in the order of hundreds of days, the rapid initial decay will be totally lost. Another fact is related to the so-called “leverage effect”, or negative return-volatility correction, showing that the price drops are followed by an increase level of volatility [10]. If only a single time scale is adopted, both the volatility correction and the leverage functions should have approximately the same decay rate, which is in contrast to the empirical evidence that the leverage effect should decay much faster than the volatility correction [84]. As a result, it is clear that the SV models containing a single time scale are unable to illustrate the different temporal behaviors of both the volatility corrections and the leverage effect, and thus, the multiscale volatility models are introduced recently to reproduce the multi-time scale dynamics reported above. In the present chapter, we adopt a class of multiscale SV models proposed by Fouque et al. [39]. In such class of models, the volatility processes are assumed to be driven by two diffusions, one fluctuating on a fast scale, and the other on a slow time scale.

For option pricing, it is known that the valuation of perpetual American options, which can be exercised at any time but without expiration date, has received a great deal of attention [12, 44, 79, 122, 123]. This simple contract can not only be used as an approximation for the long-tenor American option but also in principle be adopted as a building block in an approximation procedure for American options with finite maturities [15]. Moreover, the price of a perpetual American option under a SV model is particularly useful for the quantitative study of the effect of the SV
term on the American option price, simply because the pricing impact of a SV is fairly small in the absolute sense, if the option lasts less than one year, but it becomes progressively larger as the life span of the option increases [57]. From a mathematical point of view, as the tenor of the option approaches to infinity, the option price should involve the most significant effect of the SV on the pricing of American options, and is thus worth being examined.

In the literature, the analytic valuation of perpetual American options under the BS framework with constant volatility has been achieved ever since Merton’s closed-form solution was produced [79]. Unfortunately, his approach cannot be extended directly to the multiscale volatility models, even to those SV models with a single time scale, primarily due to the fact that the optimal exercise price is now an unknown surface of the two volatility factors, rather than a constant as in the BS model. As a result, no useful results are hitherto available for perpetual American options with multiscale volatility. The lack of analytical formula has made the quantitative analysis of the pricing impact of the SV term on perpetual American put options less achievable.

In this chapter, we derive approximation pricing formulae for a perpetual American put option contract with multiscale volatility by using the perturbation method. Following Fouque et al.’s method for pricing European derivatives [37], we obtain a sequence of simplified option pricing systems, which are so nicely constructed and can be dealt with by using the well-known Merton’s approach that was previously only available for pricing perpetual American options under the classical BS model. The analysis not only requires balancing the two SV processes, but also needs properly dealing with the boundary conditions across the optimal exercise price. It turns out that the leading order term of the current approximation is the perpetual BS price with an effective volatility “frozen” at the initial value of the slow volatility factor. The first order correction is made up of two parts, which are derived from the fast and slow factors, respectively.

The chapter is organized as follows. In Section 10.2, we review a class of multiscale
volatility models. In Section 10.3, approximation solutions for perpetual American puts with multiscale volatility are derived by using the perturbation methods. Concluding remarks are given in the last section.

10.2 Multiscale volatility models

The concept of multiscale SV models was first introduced by Fouque et al. [39] based on various empirical studies [3, 43]. In this section, we shall describe a class of multiscale volatility models in details, for the sake of completeness of the chapter and easiness of reference for the readers. In such a multi-diffusion model, the underlying $S_t$, as a function of time, is assumed to follow the SDE (stochastic differential equation) of a geometric Brownian motion:

$$dS_t = \mu S_t dt + \sigma_t S_t dw_0,$$

(10.2.1)

where $\mu$ is the drift rate, $w_0$ is a standard Brownian motion, and $\sigma_t$ represents the SV, which is driven by two other factors $Y_t$ and $Z_t$:

$$\sigma_t = f(Y_t, Z_t),$$

with $f$ being a smooth positive function that is both bounded above and bounded away from zero.

The first factor $Y_t$ is governed by a so-called fast mean-reverting process, i.e.,

$$dY_t = \frac{1}{\epsilon} (m - Y_t) dt + \frac{\nu \sqrt{2}}{\sqrt{\epsilon}} dw_1,$$

(10.2.2)

where $\frac{1}{\epsilon}$ is the mean-reversion rate, with $\epsilon$ being a small positive parameter, and $w_1$ is also a standard Brownian motion. Moreover, $m$ is the long-term mean of $Y_t$, and $\nu^2$ is the variance of the invariant distribution of $Y_t$, which controls the long-run level of the volatility fluctuations. It can be shown that the autocorrelation of $Y_t$
decays exponentially fast on the time scale $\epsilon$ [39], and thus $Y_t$ is referred to as the fast volatility factor.

The second factor $Z_t$ is assumed to follow a slowly varying process, i.e.,

$$dZ_t = \delta c(Z_t) + \sqrt{\delta} g(Z_t) dw_2,$$

(10.2.3)

with $\delta$ being a small positive parameter that is independent of $\epsilon$, and $w_2$ being another standard Brownian motion. The functions $c(z)$ and $g(z)$ are smooth and at most linearly growing at infinity.

On the other hand, in the above multi-diffusion model (10.2.1-10.2.3), a general correlation among three standard Brownian motions is allowed, i.e.,

$$
\begin{pmatrix}
  w_0 \\
  w_1 \\
  w_2
\end{pmatrix} =
\begin{pmatrix}
  1 & 0 & 0 \\
  \rho_1 & \sqrt{1-\rho_1^2} & 0 \\
  \rho_2 & \hat{\rho}_{12} & \sqrt{1-\rho_2^2-\hat{\rho}_{12}^2}
\end{pmatrix}
W_t,
$$

(10.2.4)

where $W_t$ is a standard three-dimensional Brownian motion. Moreover, the constants $\rho_1$, $\rho_2$ and $\hat{\rho}_{12}$ satisfy $|\rho_1| < 1$ and $\rho_2^2 + \hat{\rho}_{12}^2 < 1$. It can be observed that with this parametrization, the covariation between $w_1$ and $w_2$ is given by $\hat{\rho}_{12} = \rho_1 \rho_2 + \hat{\rho}_{12} \sqrt{1-\rho_1^2}$.

Clearly, once the underlying is driven by a random non-tradable volatility, the market is no longer complete, and a family of pricing measures parameterized by the market price of volatility risk should be introduced. In our work, we adopt the same combined market prices of volatility risk $\Lambda$ and $\Gamma$ as in [39], i.e.,

$$
\Lambda(y, z) = \frac{\rho_1 (\mu - r)}{f(y, z)} + \gamma(y, z) \sqrt{1-\rho_1^2},
$$

$$
\Gamma(y, z) = \frac{\rho_2 (\mu - r)}{f(y, z)} + \gamma(y, z) \hat{\rho}_{12} + \xi(y, z) \sqrt{1-\rho_2^2-\hat{\rho}_{12}^2},
$$

with $\gamma(y, z)$ and $\xi(y, z)$ being smooth bounded functions. Then, under the risk-
neutral measure, the process (10.2.1)-(10.2.3) can be written as

\[ dS_t = rS_t dt + \sigma_t S_t dw^*_t, \quad (10.2.5) \]
\[ dY_t = \left[ \frac{1}{\epsilon}(m - Y_t) - \frac{\nu\sqrt{2}}{\sqrt{\epsilon}} \Gamma(Y_t, Z_t) \right] dt + \frac{\nu\sqrt{2}}{\sqrt{\epsilon}} dw^*_1, \quad (10.2.6) \]
\[ dZ_t = \left[ \delta c(Z_t) - \sqrt{\delta g(Z_t)} \Gamma(Y_t, Z_t) \right] + \sqrt{\delta g(Z_t)} dw^*_2. \quad (10.2.7) \]

Here, \( w^*_i \) are defined as in (10.2.4) with \( W \) being replaced by \( W^* \), with

\[ W^* = W + \int_0^t \begin{pmatrix} (\mu - r)/f(Y_s, Z_s) \\ \gamma(Y_s, Z_s) \\ \xi(Y_s, Z_s) \end{pmatrix} ds. \]

10.3 Pricing approximations for perpetual American puts

It is not difficult to show that under the proposed general multiscale SV model (10.2.5-10.2.7), the price of a perpetual American put option forms a free boundary problem analogous to the corresponding one under the BS model [115], with two additional spatial variables \( y \) and \( z \). The free boundary which needs to be determined as part of the problem, is now a surface \( S_f(y, z) \) instead of a constant under the BS model, as a result of the volatility being driven by two other stochastic factors.

Specifically, \( P(S, y, z) \) and \( S_f(y, z) \) should satisfy the PDE system

\[
\begin{aligned}
&\frac{1}{2} f^2(y, z) S^2 \frac{\partial^2 P}{\partial S^2} + r S \frac{\partial P}{\partial S} - r P + \frac{\nu^2}{\epsilon} \left[ \rho_1 f(y, z) S \frac{\partial^2 P}{\partial S \partial y} - \Lambda(y, z) \frac{\partial P}{\partial y} \right] \\
&\quad + \frac{1}{\epsilon} \left[ (m - y) \frac{\partial P}{\partial y} + \frac{\nu^2}{\epsilon} \frac{\partial^2 P}{\partial y^2} - \sqrt{\delta g(z)} \Gamma(y, z) \frac{\partial P}{\partial z} + \sqrt{\delta \rho_2 g(z)} f(y, z) \frac{\partial^2 P}{\partial S \partial z} \right] \\
&\quad + \delta c(z) \frac{\partial P}{\partial z} + \frac{\delta^2 (z)}{2} \frac{\partial^2 P}{\partial z^2} + \sqrt{\frac{\delta}{\epsilon}} \nu \rho_{12} g(z) \frac{\partial^2 P}{\partial y \partial z} = 0, \\
&\lim_{S \to \infty} P(S, y, z) = 0, \\
&P(S_f, y, z) = K - S_f, \quad \frac{\partial P}{\partial S}(S_f, y, z) = -1.
\end{aligned}
\]

(10.3.8)
To ensure the well-posedness of (10.3.8), whether or not the boundary conditions along the $y$ or $z$ directions need to be imposed remains unclear. In case of degenerate boundaries, the corresponding boundary conditions should be imposed if the Fichera function proves to be negative, but are not needed at all if the Fichera function is nonnegative [36], whereas for the non-degenerate case, the boundary conditions should be given a priori. For example, under the Heston model, where $f(y) = \sqrt{y}$, $\Lambda(y) = 0$, $\alpha = \kappa$, $m = \eta$, and $\nu = \sigma/\sqrt{2\alpha}$, it can be shown that $y = +\infty$ is non-degenerate, and thus a boundary condition at this particular point should be imposed; whereas $y = 0$ is a degenerate boundary, and the corresponding Fichera function equals to $\kappa \eta - \sigma^2/2$. Therefore, if $\kappa \eta < \sigma^2/2$, the boundary condition at $y = 0$ is mathematically needed. On the other hand, if $\kappa \eta \geq \sigma^2/2$, the boundary condition at $y = 0$ becomes mathematically redundant. For the more general case we are discussing in this chapter, it is impossible to judge whether the boundary curves along the two directions are degenerate or not, because the specific forms of the functions $c(z)$, $g(z)$, $f(y, z)$, $\Gamma(y, z)$ and $\Lambda(y, z)$ depend on the volatility process, and are not specified yet. Furthermore, even if those boundaries prove to be degenerate, it is still difficult to seek any property of the corresponding Fichera functions, because of the same reason stated before. To find a general solution valid for a wide class of multi-scale SV models, we consider a solution that satisfies (10.3.8). It may or may not satisfy the boundary conditions along $y$ or $z$ directions, if they are proposed to ensure the well-posedness of (10.3.8). Financially, the solution we present below is at least valid for volatility levels which are not being extremely high or low.

For convenience, we normalize (10.3.8) by introducing the new variables: $x = \ln \frac{S}{K}$, $P^{x,\delta} = \frac{P}{K}$, and obtain:

$$
\begin{align*}
\mathcal{L}^{x,\delta} P^{x,\delta} &= 0, \\
\lim_{x \to \infty} P^{x,\delta}(x, y, z) &= 0, \\
P^{x,\delta}(x_f^{x,\delta}(y, z), y, z) &= 1 - \exp(x_f^{x,\delta}(y, z)), \\
\frac{\partial P^{x,\delta}}{\partial x}(x_f^{x,\delta}(y, z), y, z) &= -\exp(x_f^{x,\delta}(y, z)),
\end{align*}
$$

(10.3.9)
where the operator $\mathcal{L}^{\epsilon,\delta}$ is defined as:

$$
\mathcal{L}^{\epsilon,\delta} = \frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 + \sqrt{\delta} \mathcal{M}_1 + \delta \mathcal{M}_2 + \sqrt{\frac{\delta}{\epsilon}} \mathcal{M}_3,
$$

with

$$
\begin{align*}
\mathcal{L}_0 &= (m - y) \frac{\partial}{\partial y} + \nu^2 \frac{\partial^2}{\partial y^2}, \\
\mathcal{L}_1 &= \sqrt{2} \nu [\rho_1 f(y, z) \frac{\partial^2}{\partial x \partial y} - \Lambda(y, z) \frac{\partial}{\partial y}], \\
\mathcal{L}_2 &= \frac{1}{2} f^2(y, z) \frac{\partial^2}{\partial x^2} + (r - \frac{1}{2} f^2(y, z)) \frac{\partial}{\partial x} - rI, \\
\mathcal{M}_1 &= -g(z) \Gamma(y, z) \frac{\partial}{\partial z} + \rho_2 g(z) f(y, z) \frac{\partial^2}{\partial x \partial z}, \\
\mathcal{M}_2 &= c(z) \frac{\partial}{\partial z} + \frac{g^2(z)}{2} \frac{\partial^2}{\partial z^2}, \\
\mathcal{M}_3 &= \sqrt{2} \nu \rho_1 g(z) \frac{\partial^2}{\partial y \partial z}.
\end{align*}
$$

Since there are two independent small parameters, i.e., $\epsilon$ and $\delta$, following the standard asymptotic analysis, we should explore a solution in powers of both $\epsilon$ and $\delta$. To facilitate the analysis, an alternative way is to find a series solution in powers of one of the small parameters first, while treating the other one as a normal parameter. Then, the series solution is further expanded w.r.t. the previously fixed small parameter. In our work, we first seek the solution of (10.3.9) in powers of $\sqrt{\delta}$, i.e.,

$$
P^{\epsilon,\delta}(x, y, z) = \sum_{n=0}^{\infty} \delta^{\frac{n}{2}} P_n^{\epsilon}(x, y),
$$

$$
x_f^{\epsilon,\delta}(y, z) = \sum_{n=0}^{\infty} \delta^{\frac{n}{2}} x_n^\epsilon(y).
$$

In fact, if we choose to expand the solution in powers of $\sqrt{\epsilon}$ first, the specific analysis will be quite different, but the final results should be exactly the same. For briefness, we omit the details here. One should also notice that the series solution we try to find is expanded in the powers of $\sqrt{\delta}$, instead of in the regular powers of $\delta$. This is because the operator $\mathcal{L}^{\epsilon,\delta}$ is a linear combination of the terms in the powers of $\sqrt{\delta}$,
\[ \mathcal{L}^{\epsilon, \delta} = \left( \frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 \right) (\sqrt{\delta})^0 + \left( \mathcal{M}_1 + \sqrt{\frac{1}{\epsilon} \mathcal{M}_3} \right) \sqrt{\delta} + \mathcal{M}_2 (\sqrt{\delta})^2. \]  

(10.3.12)

By substituting (10.3.10) into the governing equation contained in (10.3.9), we obtain

\[ \sum_{n=0}^{\infty} \delta^{\frac{n}{2}} \left( \frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 \right) P_n^{\epsilon}(x, y, z) + \sum_{n=0}^{\infty} \delta^{\frac{n+1}{2}} \left( \mathcal{M}_1 + \sqrt{\frac{1}{\epsilon} \mathcal{M}_3} \right) P_n^{\epsilon}(x, v) + \sum_{n=0}^{\infty} \delta^{\frac{n+1}{2}} \mathcal{M}_2 P_n^{\epsilon}(x, y, z) = 0. \]  

(10.3.13)

Also, substituting (10.3.11) into the free boundary conditions yields

\[ \sum_{n=0}^{\infty} \delta^{\frac{n}{2}} \left( \frac{1}{\epsilon^{\frac{3}{2}}} \mathcal{L}_0 + \frac{1}{\epsilon^{\frac{1}{2}}} \mathcal{L}_1 + \mathcal{L}_2 \right) \left( \sum_{n=0}^{\infty} \delta^{\frac{n}{2}} x_n^{\epsilon} \right) = 1 - \exp\left( \sum_{n=0}^{\infty} \delta^{\frac{n}{2}} x_n^{\epsilon} \right), \]  

(10.3.14)

\[ \sum_{n=0}^{\infty} \delta^{\frac{n}{2}} \frac{\partial}{\partial x} \left( \sum_{n=0}^{\infty} \delta^{\frac{n}{2}} x_n^{\epsilon} \right) = -\exp\left( \sum_{n=0}^{\infty} \delta^{\frac{n}{2}} x_n^{\epsilon} \right). \]  

(10.3.15)

Now, by truncating both sides of (10.3.13)-(10.3.15) to the order of \( O(\delta) \), we obtain

\[ \left( \frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 \right) P_0^{\epsilon} + \sqrt{\delta} \left( \frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 \right) P_0^{\epsilon} + \left( \mathcal{M}_1 + \frac{1}{\sqrt{\epsilon}} \mathcal{M}_3 \right) P_0^{\epsilon} = O(\delta), \]

\[ P_0^{\epsilon}(x_0^{\epsilon}, y, z) + \sqrt{\delta} \left( \frac{\partial}{\partial x} P_0^{\epsilon}(x_0^{\epsilon}, y, z) x_0^{\epsilon} + P_1^{\epsilon}(x_0^{\epsilon}, y, z) \right) = 1 - \exp(x_0^{\epsilon}) - \sqrt{\delta} x_0^{\epsilon} \exp(x_0^{\epsilon}) + O(\delta), \]

\[ \frac{\partial}{\partial x} P_0^{\epsilon}(x_0^{\epsilon}, y, z) + \sqrt{\delta} \left( \frac{\partial^2}{\partial x^2} P_0^{\epsilon}(x_0^{\epsilon}, y, z) x_0^{\epsilon} + \frac{\partial P_0^{\epsilon}}{\partial x}(x_0^{\epsilon}, y, z) \right) = -\exp(x_0^{\epsilon}) - \sqrt{\delta} x_0^{\epsilon} \exp(x_0^{\epsilon}) + O(\delta). \]  

(10.3.16)

The zeroth order term with respect to the slow scale

By setting the coefficients in front of those \( O(1) \) terms in (10.3.16) to zero, we obtain
Chapter 10.

the zeroth-order term w.r.t. (with respect to) the slow scale $\delta$ satisfying the PDE system

$$
\begin{cases}
\left(\frac{1}{\epsilon}\mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}}\mathcal{L}_1 + \mathcal{L}_2\right)P^\epsilon_0 = 0, \\
\lim_{x \to \infty} P^\epsilon_0(x, y, z) = 0, \\
P^\epsilon_0(x^\epsilon_0, y, z) = 1 - \exp(x^\epsilon_0), \\
\frac{\partial P^\epsilon_0}{\partial x}(x^\epsilon_0, y, z) = -\exp(x^\epsilon_0).
\end{cases}
$$

(10.3.17)

One can easily observe that $P^\epsilon_0$ is nothing but the price of a perpetual American put option for the case of fast-mean reversion in $y$. Based on Fouque et al.’s analysis for pricing European derivatives under a fast-mean reverting volatility [37], we can derive an approximation for $P^\epsilon_0$, as shall be shown in the following work.

In order to solve for $P^\epsilon_0$ and $x^\epsilon_0$, we first expand both of them in the powers of $\sqrt{\epsilon}$, i.e.,

$$
P^\epsilon_0(x, y, z) = \sum_{n=0}^{\infty} \epsilon^\frac{n}{2} P_{n,0}(x, y, z),
$$

(10.3.18)

$$
x^\epsilon_0(y, z) = \sum_{n=0}^{\infty} \epsilon^\frac{n}{2} x_{n,0}(y, z).
$$

(10.3.19)

By substituting (10.3.18) into the governing equation contained in (10.3.17), at the lowest order, $O\left(\frac{1}{\epsilon}\right)$, we obtain

$$
\mathcal{L}_0 P_{0,0} = 0.
$$

At this stage, although the explicit form of $P_{0,0}$ remains unknown, we can still deduce that $P_{0,0}$ does not depend on $y$, i.e., $P_{0,0} = P_{0,0}(x, z)$. This is because the operator $\mathcal{L}_0$ is the generator of an ergodic Markov process and acts only on the $y$ variable, and moreover, the particular solutions that depend on $y$ are in general ruled out by the far-field boundary condition along the $x$ direction.

To eliminate the $O\left(\frac{1}{\sqrt{\epsilon}}\right)$ term, we obtain

$$
\mathcal{L}_0 P_{1,0} + \mathcal{L}_1 P_{0,0} = 0.
$$

(10.3.20)
As mentioned earlier, \( P_{0,0} \) is a constant w.r.t. \( y \), and thus, we have \( \mathcal{L}_1 P_{0,0} = 0 \), which, combined with (10.3.20), yields \( \mathcal{L}_0 P_{1,0} = 0 \). Using the same argument as that for \( P_{0,0} \), it is clear that \( P_{1,0} = P_{1,0}(x, z) \).

At \( O(1) \), the governing equation becomes

\[
\mathcal{L}_2 P_{0,0} + \mathcal{L}_1 P_{1,0} + \mathcal{L}_0 P_{2,0} = 0,
\]

and consequently,

\[
\mathcal{L}_0 P_{2,0} + \mathcal{L}_2 P_{0,0} = 0,
\]

(10.3.21)

since \( P_{1,0} \) does not depend on \( y \). Given that \( P_{0,0} \) is known in advance, (10.3.21) is a Poisson equation for \( P_{2,0} \) w.r.t. the operator \( \mathcal{L}_0 \) in the variable \( y \). By applying the Fredholm alternative theorem [87], it is clear that (10.3.21) has no solution unless \( \mathcal{L}_2 P_{0,0} \) is orthogonal to the invariant distribution of the fast factor \( Y \), whose infinitesimal generator is \( \mathcal{L}_0 \), i.e.,

\[
\langle \mathcal{L}_2 P_{0,0} \rangle = 0.
\]

Here, \( \langle \cdot \rangle \) denotes \( \int_{R_y} \cdot p_\infty dy \), with \( p_\infty \) being the invariant distribution of \( Y \), i.e., \( \mathcal{L}_0^* p_\infty = 0 \), where \( \mathcal{L}_0^* \) is the adjoint operator of \( \mathcal{L}_0 \). In our case, it is not difficult to show that

\[
p_\infty = \frac{1}{\sqrt{2\pi \nu}} \exp\left(-\frac{(y - m)^2}{2\nu^2}\right).
\]

On the other hand, by using the same argument as we did when deriving the boundary conditions for \( P_0' \) and \( P_1' \), we obtain

\[
P_{0,0}(x_{0,0}, z) = 1 - \exp(x_{0,0}), \quad \frac{\partial P_{0,0}}{\partial x}(x_{0,0}, z) = -\exp(x_{0,0}).
\]
Therefore, $P_{0,0}$ is the solution of the following PDE system:

$$
\begin{align*}
\langle \mathcal{L}_2 P_{0,0} \rangle &= 0, \quad \lim_{x \to \infty} P_{0,0}(x, z) = 0, \\
P_{0,0}(x_{0,0}, z) &= 1 - \exp(x_{0,0}), \quad \frac{\partial P_{0,0}}{\partial x}(x_{0,0}, z) = -\exp(x_{0,0}).
\end{align*}
$$

(10.3.22)

Since $P_{0,0}$ is a constant w.r.t. the variable $y$, the governing equation contained in (10.3.22) can be further simplified as $\langle \mathcal{L}_2 \rangle P_{0,0} = 0$. Also, from the definition of $\mathcal{L}_2$, one can deduce that $\langle \mathcal{L}_2 \rangle = \mathcal{L}_{BS}(\bar{f}(z))$, where $\bar{f}(z)$ is the so-called effective volatility, and is defined as the statistical average w.r.t. the invariant distribution, i.e., $\bar{f}(z) = \langle f(y, z) \rangle$. Clearly, $y$ and $z$ no longer need to be treated as variables in (10.3.22) because the operator $\langle \mathcal{L}_2 \rangle$ is nothing but the perpetual BS operator, involving no partial differentiation w.r.t. $y$ or $z$ at all. Consequently, the solution of (10.3.22) can be easily found as

$$
P_{0,0}(x, z) = \frac{1}{1 + a} \left[ \frac{1 + a}{a} \exp(x) \right]^{-a},
$$

$$
x_{0,0} = \ln \frac{a}{1 + a},
$$

with $a = \frac{2r}{\bar{f}^2(z)}$ being the relative interest rate of the effective volatility of the underlying to the risk-free interest, a concept similar to $\gamma = \frac{2r}{\sigma^2}$ appearing in the classical BS model [79]. In fact, the above solution is indeed identical to Merton’s well-known solution under the BS model, had the effective volatility $\bar{f}(z)$ been replaced by $\sigma$ used in Merton’s case. Such a degeneration back to the BS system is expected at the zeroth order and is a reassurance that our solution procedure is correct. From a stochastic point of view, we know that the distribution of the fast factor $Y_t$ depends only on the product of the mean-reversion rate and the time, and thus when the mean-reversion rate is assumed to be very large, the distribution should be the same as its large time distribution [37]; while for the slow factor $Z_t$, it should be “frozen” at its initial level $z$, in the limit that the variation approaches to zero. Mathematically,
in the limit sense for both factors, we have

$$\lim_{\epsilon \to 0, \delta \to 0} f(Y_t, Z_t) = \lim_{t \to \infty} \int_0^t f(Y_s, z) ds.$$  \hfill (10.3.23)

On the other hand, according to the ergodic theorem [37], it is known that the effective volatility $\bar{f}(z)$ is almost surely equal to the long-run time average of the function $f$, i.e.,

$$\bar{f}(z) = \lim_{t \to \infty} \frac{1}{t} \int_0^t f(Y_s, z) ds \ (\text{almost surely}),$$  \hfill (10.3.24)

which, combined with (10.3.23), yields

$$\lim_{\epsilon \to 0, \delta \to 0} f(Y_t, Z_m) = \bar{f}(z) \ (\text{almost surely}).$$  \hfill (10.3.25)

This also further explains why the constant volatility in the BS model should be replaced by the effective volatility “frozen” at the spot level of the slow factor in the zeroth-order solution under multiscale volatility framework. Then, our solutions at the next order, presented below, will be corrections to Merton’s constant volatility case, when the volatility is driven by two other stochastic factors.

In order to solve for the first order correction terms w.r.t. the fast scale $\epsilon$, i.e., $P_{1,0}$, we collect the $\mathcal{O}(\sqrt{\epsilon})$ terms, which yield

$$\mathcal{L}_0 P_{3,0} + \mathcal{L}_1 P_{2,0} + \mathcal{L}_2 P_{1,0} = 0.$$  \hfill (10.3.26)

Again, by applying the Fredholm alternative theorem to (10.3.26), we obtain

$$\langle \mathcal{L}_1 P_{2,0} + \mathcal{L}_2 P_{1,0} \rangle = 0.$$  \hfill (10.3.27)

Since $P_{1,0}$ does not depend on $y$, (10.3.27) can be simplified as

$$\langle \mathcal{L}_2 \rangle P_{1,0} = -\langle \mathcal{L}_1 P_{2,0} \rangle.$$  \hfill (10.3.28)
On the other hand, since $L_2 P_{0,0} = L_2 P_{0,0} - \langle L_2 \rangle P_{0,0}$, it is clear that

$$L_2 P_{0,0} = \frac{1}{2} (f^2(y, z) - \bar{f}^2(z)) \left( \frac{\partial^2 P_{0,0}}{\partial x^2} - \frac{\partial P_{0,0}}{\partial x} \right). \quad (10.3.29)$$

Substituting (10.3.29) into (10.3.21), we obtain

$$L_0 P_{2,0} = -L_2 P_{0,0},$$

$$= -\frac{1}{2} (f^2(y, z) - \bar{f}^2(z)) \left( \frac{\partial^2 P_{0,0}}{\partial x^2} - \frac{\partial P_{0,0}}{\partial x} \right).$$

Again using the fact that $P_{0,0}$ does not depend on $y$, we can deduce that

$$P_{2,0} = -\frac{1}{2} \mathcal{L}_0^{-1} (f^2(y, z) - \bar{f}^2(z)) \left( \frac{\partial^2 P_{0,0}}{\partial x^2} - \frac{\partial P_{0,0}}{\partial x} \right),$$

$$= -\frac{1}{2} \left( \phi(y, z) + c(x, z) \right) \left( \frac{\partial^2 P_{0,0}}{\partial x^2} - \frac{\partial P_{0,0}}{\partial x} \right), \quad (10.3.30)$$

where $\phi(y, z)$ is the solution of $\mathcal{L}_0 \phi(y, z) = f^2(y, z) - \bar{f}^2(z)$, and $c(x, z)$ is a constant w.r.t. $y$.

Now, substituting (10.3.30) into (10.3.28), we obtain

$$\langle L_2 \rangle P_{1,0} = \frac{1}{2} \langle \mathcal{L}_1 \phi(y, z) \rangle \left( \frac{\partial^2 P_{0,0}}{\partial x^2} - \frac{\partial P_{0,0}}{\partial x} \right),$$

$$= V_3 \frac{\partial^3 P_{0,0}}{\partial x^3} + (V_2 - 3V_3) \frac{\partial^2 P_{0,0}}{\partial x^2} + (2V_3 - V_2) \frac{\partial P_{0,0}}{\partial x}, \quad (10.3.31)$$

where

$$V_2 = \sqrt{2} \nu \rho_1 \langle f(y, z) \phi_y(y, z) \rangle - \frac{\sqrt{2}}{2} \nu \langle \wedge(y, z) \phi_y(y, z) \rangle,$$

$$V_3 = \frac{\sqrt{2}}{2} \nu \rho_1 \langle f(y, z) \phi_y(y, z) \rangle.$$

It should be pointed out that in the current approach, the detailed expressions of $V_2$ and $V_3$ do not need to be given a priori. In other words, the approach adopted is able to deal with a wide class of multiscale volatility models. It should also be pointed out that these two terms, $V_2$ and $V_3$, do not play a same role in (10.3.31) at
all. The $V_2$ term is simply a volatility level correction, and depends on both $\rho_1$ and the market price of risk $\gamma$, whereas the $V_3$ term shows the “skew” effect due to the presence of the third order derivative, and it vanishes if $\rho_1$ is equal to zero [37].

Now, taking all the boundary conditions for $P_{1,0}$ into consideration, we find that $P_{1,0}$ satisfies

$$\begin{cases} (\mathcal{L}_2)P_{1,0} = D \exp(-ax), \\ \lim_{x \to \infty} P_{0,0}(x, z) = 0, \quad P_{1,0}(x_{0,0}, z) = 0, \\ \frac{\partial^2 P_{0,0}}{\partial x^2}(x_{0,0}, z)x_{1,0} + \frac{\partial P_{1,0}}{\partial x}(x_{0,0}, z) = -\exp(x_{0,0})x_{1,0}, \end{cases}$$

(10.3.32)

with $D = a\left(\frac{1}{a} + \frac{1}{a}\right)^{-a}[-(a + 2)V_3 + V_2]$.

One can easily observe that at this stage, the original free boundary problem has transformed to an “explicit” boundary problem, after the perturbation methods are applied. In fact, it can be shown that one only needs to deal with the moving boundary implicitly at the zeroth-order, and then all the corrections to the zeroth-order moving boundary w.r.t. the fast scale can be explicitly dealt with at subsequent high orders. Such an easy decoupling of unknowns in (10.3.32) and the disappearance of the “free boundary” have considerably facilitated the solution process for the valuation of perpetual American options. On the other hand, one should notice that the problem formulated above for $P_{1,0}$ is only an inhomogeneous ODE (ordinary differential equation) system, where $x$ is the only variable involved, while $z$ is treated as if it were a constant, and $y$ disappears as a result of taking a statistical average w.r.t. the fast factor. Moreover, this ODE system is characterized with the same differential operator $\langle \mathcal{L}_2 \rangle$, and the source terms from the volatility level correction and the “skew” effect, through which the corrections w.r.t. the fast scale to the zeroth-order solutions are made. Such an ODE system can be elegantly solved by applying an approach similar to the well-known Merton’s method, which was previously applied to solve perpetual American options with constant volatility. We first find the general solution of the governing ODE and then, determine the coefficients utilizing the
boundary conditions at the first order. After some simple algebraic manipulations, we have found $P_{1,0}$ and $x_{1,0}$ as

$$P_{1,0}(x, z) = \frac{2a(1 + 1/a)^{-a}[-(a + 2)V_3 + V_2](-\ln(1 + 1/a) - x)\exp(-ax)}{f^2(z)(a + 1)},$$  \hspace{1cm} (10.3.33)

$$x_{1,0} = \frac{2[-(a + 2)V_3 + V_2]}{2r + f^2(z)}.$$  \hspace{1cm} (10.3.34)

It should be pointed out that the BS price with volatility being $\bar{f}(z)$ is already a good approximation for the case of infinite mean-reversion rate, i.e., $\epsilon \to 0$. However, the above correction terms are needed to account for the randomness of the volatility whose mean-reversion rate is fast but finite. On the other hand, one should notice that $y$ does not appear explicitly in (10.3.33)-(10.3.34), although they are corrections to the BS price when the fast factor $y$ is taken into consideration. This is in fact quite reasonable in an average sense, because when the mean-reversion rate is very fast, the process $Y_t$ will fluctuate around its ergodic mean rapidly, resulting in the volatility oscillating around $\bar{f}(z)$.

**The first order correction with respect to the slow scale**

Clearly, in the previous derivations for $P_0^\epsilon$ and $x_0^\epsilon$, the slow factor $Z_t$ is kept constant during the whole solution process. In this sense, it is reasonable to treat $P_0^\epsilon$ and $x_0^\epsilon$ as solutions for the fast mean-reverting volatility scenario. Not surprisingly, once the slow factor is taken into consideration, there should also be corrections to account for the randomness of $Z_t$. We shall discuss this issue in the following work.

Now, continuing the process of equating terms in (10.3.16) by setting the coefficients in front of $O(\sqrt{\delta})$ to zero, we obtain

$$\begin{align*}
\left\{ \begin{array}{l}
\left( \frac{1}{\epsilon} \mathcal{L}_0 + \frac{1}{\sqrt{\epsilon}} \mathcal{L}_1 + \mathcal{L}_2 \right) P_1^\epsilon = -(\mathcal{M}_1 + \frac{1}{\sqrt{\epsilon}} \mathcal{M}_3) P_0^\epsilon, \\
\lim_{x \to \infty} P_1^\epsilon(x, y, z) = 0, \quad P_1^\epsilon(x_0^\epsilon, y, z) = 0, \\
\frac{\partial^2 P_0^\epsilon}{\partial x^2}(x_0^\epsilon, y, z) x_1^\epsilon + \frac{\partial P_1^\epsilon}{\partial x}(x_0^\epsilon, y, z) = -\exp(x_0^\epsilon)x_1^\epsilon.
\end{array} \right. \hspace{1cm} (10.3.35)
\end{align*}$$
Chapter 10.

To solve for $P_1^r$ and $x_1^r$, we expand both of them in powers of $\sqrt{\epsilon}$ as we did before, i.e.,

$$P_1^r(x, y, z) = \sum_{n=0}^{\infty} \epsilon^{\frac{n}{2}} P_{n,1}(x, y, z), \quad (10.3.36)$$

$$x_1(y, z) = \sum_{n=0}^{\infty} \epsilon^{\frac{n}{2}} x_{n,1}(y, z). \quad (10.3.37)$$

Substituting (10.3.36) into the governing equation contained in (10.3.35) yields, at the lowest order $\frac{1}{\epsilon}$, $L_0 P_{0,1} = 0$. By using the same argument as that for $P_{0,0}$, it is straightforward that $P_{0,1}$ does not depend on $y$.

By setting the coefficients in front of those $O\left(\frac{1}{\sqrt{\epsilon}}\right)$ terms to zero, we obtain

$$L_1 P_{0,1} + L_0 P_{1,1} = -M_3 P_{0,0},$$

which, combined with the fact that $L_1 P_{0,1} = 0$ and $M_3 P_{0,0} = 0$, yields $L_0 P_{1,1} = 0$, and consequently, $P_{1,1}$ is also a constant w.r.t. $y$.

Continuing the process to $O(1)$, we have

$$L_0 P_{2,1} + L_1 P_{1,1} + L_2 P_{0,1} = -M_1 P_{0,0} - M_3 P_{1,0}. \quad (10.3.38)$$

Since $L_1 P_{1,1} = 0$ and $M_3 P_{1,0} = 0$, (10.3.38) can be simplified as

$$L_0 P_{2,1} + L_2 P_{0,1} = -M_1 P_{0,0}. \quad (10.3.39)$$

Again, applying the Fredholm alternative theorem to (10.3.39), we obtain

$$\langle L_2 \rangle P_{0,1} = -\langle M_1 \rangle P_{0,0}.$$

On the other hand, expending the boundary conditions contained in (10.3.35) by
Taylor series, and collecting those leading order terms, we have

\[ P_{0,1}(x_0, z) = 0, \]
\[ \frac{\partial^2 P_{0,0}}{\partial x^2}(x_0, z)x_{0,1} + \frac{\partial P_{0,1}}{\partial x}(x_0, z) = -\exp(x_0) x_{0,1}. \]

Therefore, \( P_{0,1} \) is the solution of

\[
\begin{align*}
\langle \mathcal{L}_2 \rangle P_{0,1} &= -\langle \mathcal{M}_1 \rangle P_{0,0}, \\
\lim_{x \to \infty} P_{0,1}(x, z) &= 0, \quad P_{0,1}(x_0, z) = 0, \\
\frac{\partial P_{0,0}}{\partial x^2}(x_0, z)x_{0,1} + \frac{\partial P_{0,1}}{\partial x}(x_0, z) &= -\exp(x_0) x_{0,1}.
\end{align*}
\] (10.3.40)

It can be observed that the problem formulated for \( P_{0,1} \) is similar to the corresponding one for \( P_{1,0} \), that is, \( P_{0,1} \) is also governed by an inhomogeneous ODE system, and moreover, once \( P_{0,1} \) is found, \( x_{0,1} \) can be obtained via

\[ x_{0,1} = -\frac{\frac{\partial P_{0,1}}{\partial x}(x_0, z)}{\frac{\partial^2 P_{0,0}}{\partial x^2}(x_0, z) + \exp(x_0)} \cdot \]

Now, applying the same procedure as we did for deriving \( P_{1,0} \) to (10.3.40), we obtain

\[ P_{0,1} = -\frac{\exp(-ax)}{f^2(z)(1+a)^2} \int_{x_0}^{x} 2(D + C + aD + Cy + aCy) dy, \]

where

\[ C = \frac{4r f'(z) g(z)}{f^3(z)(1+a)} \left( \frac{1+a}{a} \right)^a [\Gamma(z) + \rho_2 \bar{f}(z) a], \]
\[ D = \frac{4r f'(z) g(z)}{f^3(z)(1+a)} \left( \frac{1+a}{a} \right)^a [\Gamma(z) \ln \frac{1+a}{a} + \rho_2 \bar{f}(z) (a \ln \frac{1+a}{a} - 1)]. \]

Therefore, \( x_{0,1} = -\frac{4f'(z)g(z)}{f^3(z)(1+a)^3} [\rho_2 \bar{f}(z) - \bar{\Gamma}(z)]. \)

It should be pointed out that \( P_{0,1} \) and \( x_{0,1} \) derived in this section result from the slow factor \( Z_t \). In fact, \( P_{0,0} + \sqrt{\epsilon} P_{0,1} \) and \( x_{0,0} + \sqrt{\epsilon} x_{0,1} \) can be viewed as solutions of
perpetual American puts under a single but slowly-varying volatility. One should also notice that for options with short tenor, there is no need to use these two first-order corrections w.r.t. the slow scale, since the change of the volatility will be negligible during a rather short period. However, for the current assumption that the options have no expiration date, the impact of the slow factor \( z \) on the option prices might be significant because of the “time accumulative effect”. The quantitative analysis for this “time accumulative effect” can be easily achieved by using our first order correction w.r.t. the slow scale, just as Zhu and Chen did in [122] with a single slowly varying stochastic process being taken into consideration.

Now, by taking both the fast and slow factors into consideration, the option price of perpetual American puts under a general multiscale volatility model can be written as (in the original variables)

\[
P = \frac{K f^2(z)}{f^2(z) + 2r} \left[ \frac{S(f^2(z) + 2r)}{2r} \right] - \frac{2\sqrt{2} \sqrt{r} K \nu}{f^2(z) (f^2(z) + 2r)} \ln \frac{2rK}{S[2r + f^2(z)]} \times
\]

\[
\left[ \frac{S(f^2(z) + 2r)}{2r} \right] - \frac{2r \rho_1 (f \phi_y)}{f^2(z)} + \langle \Lambda \phi_y \rangle
\]

\[
- \sqrt{58rK} \bar{f}'(z) \bar{f}(z) g(z) \left[ \frac{S(f^2(z) + 2r)}{2rK} \right] - \frac{2r}{f^2(z)} \int_{\ln \frac{S}{2r + f^2(z)}}^{\infty} g(y) dy,
\]

with

\[
g(y) = \left[ \left( \frac{2r}{f^2(z)} + 1 \right) \left( \ln \frac{2r + f^2(z)}{2r} + y \right) + 1 \right] \Gamma(z)
\]

\[
+ \rho_2 \bar{f}(z) \left[ \frac{2r}{f^2(z)} \left( \frac{2r}{f^2(z)} + 1 \right) \left( y + \ln \frac{2r + f^2(z)}{2r} \right) - 1 \right],
\]

and the optimal exercise price as

\[
S_f = K \exp(x_{0,0} + \sqrt{\varepsilon}x_{1,0} + \sqrt{\delta}x_{0,1}) \approx K \exp(x_{0,0})(1 + \sqrt{\varepsilon}x_{1,0} + \sqrt{\delta}x_{0,1})
\]

\[
= \frac{2rK}{2r + f^2(z)} \left\{ 1 - \frac{\sqrt{2} \nu [2r \rho_1 (f \phi_y) / f^2(z) + \langle \Lambda \phi_y \rangle]}{f^2(z) + 2r}
\right.
\]

\[
- \frac{4\sqrt{5} \bar{f}'(z) \bar{f}(z) g(z)}{(f^2(z) + 2r)^3} \left( \rho_2 \bar{f}(z) - \Gamma(z) \right) \}\}.\]
It is clear that in the approximation given above, the leading order term is the perpetual BS price with an effective volatility. The first order corrections are made up of two parts, resulting from the fast and slow factors, respectively.

10.4 Discussions

One can easily observe that the fast factor $y$ does not play any explicit role in the present formulae. From a modeling point of view, this is quite reasonable. In the fast mean-reverting volatility scenario, although the volatility might fluctuate around its ergodic mean considerably over many months during the life span of an option contract, there might also be much fluctuation of the underlying price, and the changes of the volatility are not as significant as those of the underlying price [39]. In other words, the volatility can be considered relatively constant until its next major fluctuation comes out, because the minor volatility fluctuations are insignificant comparing with the changes of the underlying price. As a result, only a statistical average of all possible paths of $Y_t$, rather than its spot level, is involved in the current formulae. On the contrary, the slow volatility factor $z$ is explicitly involved in the current approximation. This is not surprising, because for the slowly varying process $Z_t$, in the limit, it should be “frozen” at its initial value $z$.

One can also observe that the leading order term of our approximation is the perpetual BS price with the constant volatility being replaced by the effective volatility “frozen” at its initial value $z$, i.e., $\bar{f}(z)$. This can be viewed as a “sum” of the effects of two volatility dynamics on the zeroth-order solution, and is indeed quite reasonable. This is because the slowly-varying volatility should be “frozen” at its initial value in the limit when the level of variations approaches to zero; whereas in the fast mean-reverting scenario, the volatility should be its “effective” level when the mean-reversion rate becomes extremely large.

On the other hand, with the newly obtained formulae (10.3.41) and (10.3.42), the impact of the multi-scale volatility on the pricing of perpetual American puts can be analyzed quantitatively. Such an analysis is quite similar to the one with a
single volatility factor taken into consideration, just as Zhu and Chen did in [122] and [121] for a slowing-varying volatility and a fast-mean reverting volatility, respectively, and thus the details are omitted here. It is found that the effect of a slowly-varying volatility factor varies w.r.t. the spot volatility. That is, for certain values of the spot volatility, the SV tends to add the value of the contract, but for others, it makes the contract less valuable, whereas the fast mean-reversion factor always tends to add the price of a put option contract.

10.5 Conclusion

By assuming the instantaneous SV of the underlying is driven by one hidden fast-mean reverting OU process and one slowly varying process, we have successfully obtained an explicit closed-form formula for the pricing of perpetual American puts with a perturbation approach. The leading order term of the current formula is the perpetual BS price with an effective volatility “frozen” at the spot level of the slow factor, while the first order corrections are made up of two parts, which are derived respectively from the fast and slow factors of the underlying volatility. The resulting formulae are particularly useful in the quantitative study of the pricing impact of the SV.
Part III: Exact solution approach

The exact solution approach is another key quantitative method in finance. Whilst difficult, closed-form exact formulae for various option derivatives are still widely explored by a number of researchers. By “closed-form”, it is meant that the solution can be written in terms of generally accepted mathematical functions and operations. By “exact”, it is meant that no approximation is made whatsoever; the differential equations, the boundary and initial conditions of the problem can all be satisfied to any desired accuracy, and the solution is infinitely differentiable. The closed-form exact solution, once developed, will be extremely valuable for theoretical as well as practical purposes. Theoretically, a closed-form exact solution would certainly fill the gap that has been in the literature of option pricing theory for the last 30 years, ever since the Black-Scholes equation was proposed. Practically, a closed-form exact solution may also be used in trading practice as the numerical realization of such formula can be made within any desired accuracy, a feature that distinguishes this approach from the numerical methods as well as the analytical approximations. In this part, substantial progress has been made in the field of pricing Parisian-type options, by deriving closed-form analytic formulae for European style Parisian and ParAsian options. Remarkably, our analytic formulae can both be computed efficiently, and would satisfy the growing demand of trading Parisian-type options.
Chapter 11

Pricing Parisian and Par\textit{asian} options analytically

11.1 Introduction

Parisian and Par\textit{asian} options are simple extensions of classical barrier options, with a “trigger” device added on, mainly for the purpose of preventing option traders from deliberately manipulating the underlying asset price when it is close to the barrier, in order to gain advantage in the option position they hold. However, such a simple addition of a financial clause has caused considerable difficulties in quantitatively pricing these options, at least analytically. In this chapter, two closed-form analytic formulae for the prices of Parisian and Par\textit{asian} options are presented for the first time.

While both Parisian options and Par\textit{asian} options share the same feature that there is a separate “clock” set up to record the total time that the underlying has passed the barrier (either above or below, depending on the type of the barrier option), the main difference between them is how the clock is reset. If one accumulates the time spent in a row and resets it to zero each time the underlying price crosses the barrier, this type is referred to as continuous Parisian options, or simply Parisian options. On the other hand, if one adds the time spent below or above the barrier
without resetting the accumulated time to zero each time the underlying crosses the
barrier, these options are named as cumulative Parisian options, or simply Parisian
options. For simplicity of reference, we may sometimes in this chapter refer to these
two options as “Parisian-type options” when there is no need to distinguish them.

Financially, there are at least two reasons for the introduction of Parisian-type
options. Firstly, comparing with the classical barrier options, Parisian-type options
are less susceptible to short-term movements of the underlying price, since a single
touch of the barrier can no longer trigger the knock-in or knock-out feature of these
options. Secondly, the hedging problem close to the barrier, which is usually encoun-
tered in the trading of the classical barrier options, can somehow be reduced, or at
least “smoothed”, for Parisian-type contracts [49]. Of course, Parisian-type options
still possess many inherent features of classical barrier options.

The valuation problem of Parisian-type options has been recognized as much more
difficult than that of classical barrier options [70]. While a closed-form analytical
solution for the latter has already been found [70], the former could only be solved
approximately. The difficulty of pricing Parisian-type options mainly comes from the
co-existence of two different barriers specified in the option contracts: a barrier of
the underlying asset price and a barrier time, which is defined as the accumulated
time that the underlying has spent above or below the barrier. Mathematically,
the specification of the Parisian options has made a 3-D (three-dimensional) PDE
(partial differential equation) system coupled with a 2-D PDE system (or two 3-D
PDE systems coupled for the Parisian case), through the prescribed continuity of
the option price and the Delta across the barrier. Such coupled systems are hard
to handle either numerically or analytically. On one hand, although the method of
images can be used to solve for the price of classical barrier options effectively, it is the
conjunction of the two barriers that has hampered the application of this powerful
method to Parisian-type options. On the other hand, traditional numerical methods,
such as the Monte Carlo simulations, would be inefficient, as one needs to trace the
time accumulated in the “trigger” all the time. Even within Parisian-type options
themselves, the pricing of Parisian and Paraskan options, albeit being financially similar, can be quite different from mathematical point of view, as resetting or no-resetting of the amount of time already accumulated in the trigger has changed the solution approach required to produce analytical pricing formulae.

In the literature, several researchers have also addressed the pricing of Parisian-type options. Predominately two types of valuation techniques, the quasi-analytic approaches and the numerical methods, are well documented. Of all the quasi-analytic methods, the most influential approach was the one proposed by Chesney et al. [21]. They used the theory of Brownian excursions and defined the value of a Parisian option in terms of an integral expressed as an inverse Laplace transform. Afterwards, their framework has been further developed by Hartly [50], Hugonnier [56], and Schröder [90] to price and hedge the Parisian-type options. Numerical methods, as another alternative, were also intensively developed recently. A typical method in this category is the PDE approach proposed by Wilmott et al. [49]. In their article, two PDE systems governing the prices of Parisian and Paraskan options are established, and then solved by using the explicit finite difference scheme. Whilst flexible and easy to implement, there are at least three major deficiencies that should be pointed out regarding their approach. Firstly, in their work [49], the singularities associated with adopting appropriate PDEs to price Parisian-type options are not explored at all; appropriately identifying these singularities and dealing with them not only make the developed PDE system (not just the PDE itself) correctly reflects what the financial clauses dictate, but also ensure any meaningful numerical scheme would lead to the correct solution. Secondly, some boundary conditions set in their framework do not seem to correctly represent the corresponding financial clauses of these options. Furthermore, we believe that at least one boundary condition connecting the pricing domains has been totally overlooked by them, and thus their pricing systems are not properly closed. Lastly, in terms of the explicit finite difference method they adopted, it is only conditionally stable, resulting in computational inefficiency and low accuracy, especially for the current 3-D problems.
In this chapter, we present two closed-form analytic solutions for the valuation of Parisian and ParAsian options, respectively, under the Black-Scholes framework. Based on several reasonable financial arguments, two new PDE systems for the prices of Parisian and ParAsian options have been established first, with one boundary condition added on for each system to ensure its closeness. Moreover, the singularities associated with these systems are also thoroughly discussed. The newly established PDE systems are then simplified through a coordinate transform that has elegantly “absorbed” one dimension associated with the “barrier time” into the time direction. The purely analytical procedures, adopted afterwards, differ w.r.t. (with respect to) the resetting mechanisms specified in the option contracts. For a Parisian option, the resulted simplified PDE system is solved analytically by applying the Laplace transform technique together with the construction of “moving windows” to evaluate the option prices backwards, slide by slide, until the given time has been reached, whereas for a ParAsian option, its non-resetting mechanism has obstructed the application of the “moving window” technique, and we apply the double Laplace transform as an alternative to analytically solve for its option price. Finally, through Laplace inversions, two completely analytic closed-form solutions are obtained for the prices of Parisian and ParAsian options, respectively. It should be pointed out that our explicit pricing formulae for pricing Parisian-type options should be valuable in both theoretic and practical senses. Theoretically, although there are several existing methods, as mentioned above, to price Parisian-type options, the closed-form exact solutions are presented for the first time. Practically, the final form of our solutions, written in terms of a linear combination of several integrals, can be used to price Parisian-type options accurately and efficiently. With a growing demand of trading exotic options in today’s finance industry, our solution procedures may lead to the development of pricing formulae for other exotic derivatives.

The rest of the chapter is organized as follows. In Section 11.2, we introduce the PDE systems that the prices of the Parisian and ParAsian options must satisfy. In Section 11.3, we present our analytic solution procedure in detail. In Section 11.4,
some numerical examples and discussions are presented to illustrate the performance of our analytic solutions when numerical values need to be calculated from them. Our brief concluding remarks are given in the last section.

11.2 PDE systems for pricing Parisian and Par\textit{asian} options

As pointed out previously, under the Black-Scholes framework, the PDE systems for the prices of Parisian-type options have already been established in [49]. However, the complexities associated with their PDE systems have hindered the application of various analytic methods. In this section, two simplified PDE systems governing the prices of Parisian-type options are provided, which pave the way for the achievement of closed-form analytic solutions for both options. In specific, the re-establishment of the PDE system for the valuation of the Parisian options will be considered in the first subsection, while that of the Par\textit{asian} options will be provided in the second subsection.

11.2.1 Parisian options

A Parisian option is a special kind of barrier options for which the knock-in or knock-out feature is only activated if the underlying remains continually in breach of the barrier $\bar{S}$ for a pre-specified time period $\bar{J}$. Like classical barrier options, Parisian call options can have four different forms: down-and-out, up-and-out, down-and-in, up-and-in call (similarly four types of Parisian puts as well). Without loss of generality, we shall consider, here in this chapter, the pricing of a European-style Parisian up-and-out call option as an example to demonstrate our solution approach; the extensions to other cases should be rather straightforward, based on the parity relationships established in [21].

Comparing with classical barrier options, the pricing of Parisian options requires the value of a new state variable $J$, the barrier time (the time recorded in the “trigger
clock”, which dictates the “knock-in” or “knock-out” action once the trigger value $\bar{J}$ is reached), which is defined as the total time the underlying has spent continually above (for up-and-out and up-and-in Parisian options) or below (for down-and-out and down-and-in Parisian options) the barrier. For the case of an up-and-out barrier, we have

$$
\begin{align*}
J &= 0, \quad dJ = 0, \quad S \leq \bar{S}; \\
dJ &= dt, \quad S > \bar{S},
\end{align*}
$$

where $\bar{S}$ is a preset barrier of the underlying. The above expression states that when the underlying is beyond $\bar{S}$, the state variable $J$ starts to accumulate values at the same rate as the passing time $t$, and when the underlying is equal to or below $\bar{S}$, $J$ is reset to zero, and remains zero. If the barrier time is not reset to zero each time the underlying crosses the barrier $\bar{S}$, the option contract becomes Par asian type. It should be remarked that the resetting mechanism associated with a Parisian option contract has made a Parisian up-and-out call option always worth more than its Par asian counterpart, as the “out” feature of the option has been somewhat amplified by the Par asian specification, as a result of the risk being knocked out is now higher with the residue possibly left in the trigger clock.

When the “barrier” of a Parisian option takes some extreme values, one can easily obtain the option price, as in these cases, the Parisian option degenerates to either a classical barrier option or a vanilla option. For example, in the case of a Parisian up-and-out call option, when the trigger value $\bar{J}$ approaches zero, the option will be immediately knocked out once the underlying touches the barrier from the below, which is the same as the specification of a classical barrier call option with up-and-out feature. On the other hand, if $\bar{J}$ becomes infinitely large, the knock-out feature will never be activated, and thus the option remains as a European call. In terms of the barrier $\bar{S}$, when $\bar{S}$ approaches zero, as long as the time to maturity exceeds the difference between the trigger value and the accumulated barrier time, which means that $J$ can reach $\bar{J}$ for sure, the knock-out feature will certainly be activated, and thus the option price becomes zero. But, if the time to maturity is less than that
difference, the barrier time $J$ will never reach $\bar{J}$, and thus the option will never be knocked out. The price then is the same as that of a European call. Lastly, when $\bar{S}$ approaches infinity, it is clear that neither the barrier nor the trigger value could be reached, and thus the option degenerates to a European call again.

For any other non-degenerate cases, the price of a Parisian option then depends on the underlying price $S$, the current time $t$ and the barrier time $J$, in addition to other parameters such as the volatility, risk-free interest rate and the expiry time. If modeled under the Black-Scholes economy, we simply assume that the underlying asset, $S$, that attracts a continuous dividend payment at a rate $D$, follows a lognormal Brownian motion given by

$$dS = (\mu - D)Sdt + \sigma SdZ,$$  \hspace{1cm} (11.2.1)

where $Z$ is a standard Brownian motion.

Now, let $V_1(S, t)$ and $V_2(S, t, J)$ denote the option prices in the region $\mathcal{I}$ and $\mathcal{II}$, respectively, with the regions $\mathcal{I}$ and $\mathcal{II}$ referring to the plane GEOX and the

Figure 11.1: Pricing domain of a Parisian up-and-out option
cuboid ABDCGEFH, respectively, as shown in Fig 11.1. Clearly, in region \( \overline{T} \), the variable \( J \) remains unchanged, as a result of the underlying being below the barrier. By applying the Feynman-Kac theorem [93] to (11.2.1), \( V_1(S, t) \) should satisfy the classical BS (Black-Scholes) equation

\[
\frac{\partial V_1}{\partial t} + \mathbb{L} V_1 = 0,
\]

where \( \mathbb{L} = \frac{\sigma^2 S^2}{2} \frac{\partial^2}{\partial S^2} + (r - D) S \frac{\partial}{\partial S} - rI \), with \( I \) being the identity operator. The terminal condition in this region is given by the payoff function of a European call option, i.e.,

\[
V_1(S, T) = \max(S - K, 0).
\]

Besides the terminal condition, a set of boundary conditions along the \( S \) direction is also needed to solve for \( V_1 \). The fact that a call option becomes worthless when the underlying price approaches zero gives

\[
\lim_{S \to 0} V_1(S, t) = 0,
\]

whereas the continuity of the option price across the barrier \( \overline{S} \) demands

\[
\lim_{S \to \overline{S}} V_1(S, t) = \lim_{S \to \overline{S}} V_2(S, t, 0).
\]

On the other hand, in region \( \overline{\Pi} \), the barrier time \( J \) starts to accumulate. As a result, \( V_2(S, t, J) \) is governed by a modified BS equation [49]

\[
\frac{\partial V_2}{\partial t} + \frac{\partial V_2}{\partial J} + \mathbb{L} V_2 = 0,
\]

with the operator \( \mathbb{L} \) being the same as that defined earlier. Appropriate boundary conditions are also needed to close the PDE system. From the definition of the “trigger clock”, it is clear that when the variable \( J \) reaches the trigger value \( \overline{J} \), the
option becomes worthless, i.e.,

$$\lim_{J \to \bar{J}} V_2(S, t, J) = 0. \quad (11.2.7)$$

Also, due to the fact that it would take infinite amount of time for an infinitely large underlying price to fall back to the barrier $\bar{S}$, the option must be worth nothing when $S$ becomes very large, i.e.,

$$\lim_{S \to \infty} V_2(S, t, J) = 0. \quad (11.2.8)$$

Moreover, the boundary condition at $S = \bar{S}$ is specified by the so called “reset condition”, i.e.,

$$\lim_{S \to \bar{S}} V_2(S, t, J) = \lim_{S \to \bar{S}} V_1(S, t), \quad (11.2.9)$$

which indicates that $J$ is reset to zero every time the underlying $S$ falls back to the barrier $\bar{S}$ from above. It should be remarked that this “reset condition” is removed for Paraskan options, as no resetting mechanism is specified in a Paraskan option contract, a feature that distinguishes the pricing of Paraskan options from that of their Parisian counterparts.

Equations (11.2.2)-(11.2.9) constitute the differential system proposed by Wilmott et al. [49], the solution of which will give rise to the value of Parisian option at any underlying price $S$, any barrier time $J$, and any time $t$ before the expiration $T$. Remarkably, Wilmott et al.’s pricing system can be viewed as substantial work in determining the price of Parisian options from the PDE point of view. However, it should also be pointed out that there are at least two fundamental flaws in their pricing system. Firstly, the boundary condition they set at $S = \infty$ is wrong for the case $\bar{J} - J > T - t$. In this case, the knock-out feature of the Parisian option will never be activated, due to the lack of enough time for $J$ to reach the trigger value $\bar{J}$, and thus the price of a Parisian up-and-out call option equals that of a European call option, which will never become zero as $S$ approaches infinity. Secondly,
the closeness of Wilmott et al.’s PDE system is not clear. We have noticed that in their explicit finite difference approach, they must have implicitly assumed that the option price across the barrier satisfies (11.2.6). Without this implicit assumption, their approach could have not produced a unique solution. Unfortunately, (11.2.6) is not listed explicitly as a boundary condition, which we believe to be necessary to properly close the PDE system.

Furthermore, it should also be remarked that Wilmott et al.’s PDE system is rather complicated, with a 2-D and a 3-D PDE system being coupled, which greatly hinders the application of various methods to solve the price of Parisian options either numerically or analytically. In fact, by realizing the following financial arguments, the original pricing domain, i.e., the regions $\mathcal{I}$ and $\mathcal{II}$, can be reduced, and thus, Wilmott et al.’s PDE system can be further simplified. Firstly, the prism ANBFEQ (denoted by $\mathcal{III}$ hereafter) should be excluded, as in this region, the elapsed time is always less than the barrier time, i.e., $0 \leq t < J$, a case that will never happen. Secondly, in the prism LCDHMG (denoted by $\mathcal{IV}$ hereafter), the barrier time $J$ has no effect on the option price, because in this domain, there is not enough time for $J$ to reach $\bar{J}$, and thus the option will never be knocked out. Consequently, the option price in the prism $\mathcal{IV}$ does not vary w.r.t. $J$, and should be the same as that of a European call option at the time to expiry $T - t$. Now, it is quite clear that the two chunks of the original pricing domain, i.e., the prisms $\mathcal{III}$ and $\mathcal{IV}$, can be “cut off” when considering the valuation of Parisian options, leaving the actual pricing domain as the plane MEOI plus the domain ANDLMEQH, defined as

\begin{align*}
I : & \{0 \leq S \leq \bar{S}, \ 0 \leq t \leq T - \bar{J} \ J = 0\}, \\
II : & \{\bar{S} \leq S < \infty, \ J \leq t \leq J + T - \bar{J} \ 0 \leq J \leq \bar{J}\},
\end{align*}

respectively. It should be remarked that simplifying the original pricing domains has mainly resulted in two significant consequences. On one hand, by cutting the prism $\mathcal{IV}$, one no longer needs to specify the boundary condition at $S = \infty$ for
\[ \tilde{J} - J > T - t, \] as this case is excluded from the new pricing domain. On the other hand, it is this simplification that has paved the way for the development of our analytical approach, as shall be discussed later.

The simplification consists of two major modifications made to Wilmott et al.’s pricing system, in order to derive a properly closed PDE system on the simplified domain. Firstly, the original terminal condition in the region \( I \), i.e., the payoff function, should be replaced by the price of a European call option at the time to expiry \( \tilde{J} \), denoted by \( V_{BS}(S, \tilde{J}) \) as

\[
V_1(S, T - \tilde{J}) = V_{BS}(S, \tilde{J}).
\]

This is not surprising, because the value on the plane LIXC is known as a European call option price at the time to expiry \( T - t \), as discussed earlier. Secondly, we should explicitly demand that the option Delta be continuous across the barrier \( S = \tilde{S} \), i.e.,

\[
\lim_{S \to \tilde{S}} \frac{\partial V_1}{\partial S}(S, t) = \lim_{S \to \tilde{S}} \frac{\partial V_2}{\partial S}(S, t, 0).
\] (11.2.10)

Clearly, there are now three conditions that link the solutions in both regions effectively, with one being the reset condition (11.2.9), and the other two being the option price and the option Delta continuous across the barrier, i.e., (11.2.5) and (11.2.10), respectively. Hereafter, for simplicity, only (11.2.10) is referred to as the “connectivity condition”. It should be remarked that this additional connectivity condition is a necessity to make the pricing system properly closed. Of course, one may wonder why Wilmott et al. [49] could still produce a set of seemingly correct numerical results from an unclosed PDE system. This is because such a “connectivity condition” can be unintentionally satisfied when the option price has satisfied the PDE (11.2.6) itself across the barrier, a trick similar to the fictitious-point technique in dealing with Neumann boundary conditions [96]. However, this requires an implicit assumption that the option prices \( V_1 \) and \( V_2 \) be at least twice differentiable across the barrier, which is much stronger than the connectivity condition we have
imposed. From the viewpoint that a solution produced with a strong condition taken into consideration should also satisfy a weak condition, it is therefore reasonable that the pricing system with a strong condition yields the same result as that with a weak condition, as long as the latter is well-posed.

Before finally setting up the properly closed pricing system, we should carefully examine the singularities associated with the pricing of Parisian options. In fact, exploring singularities is an indispensable step in establishing any PDE systems, since otherwise, the results produced from those systems would be incorrect at least around those singularities. In the literature, no one has discussed this issue thoroughly, and we thus believe that any previously published numerical solution approach should be revisited to make sure that these singularities have been properly taken care before the obtained results can be faithfully trusted.

The singularities of the Parisian option price mainly result from the introduction of a non-cumulative trigger for the barrier $\bar{S}$, i.e., the introduction of the plane NQHD (denoted by $V_I$ hereafter) has introduced a singular line DH which is the intersection of the plane DHGC and the plane $V_I$. If the line DH is viewed to belong to the plane DHGC, the terminal condition should be imposed. On the other hand, if the line DH is viewed as part of the plane $V_I$, the knock-out condition should then be imposed. From a financially meaningful argument, we believe that the line DH should be part of the plane $V_I$, in order to ensure that any point on the plane LDHM (denoted by $V$ hereafter) may still reach $\bar{J}$ as time further increases. Such a demand is consistent with the fact that the special point L ($\infty, T - \bar{J}, J$) should lie on the plane $V$, because it would take infinite amount of time for an infinitely large underlying price to fall back to $\bar{S}$, and consequently the trigger value $\bar{J}$ would for sure be reachable when time increases by a finite amount. In fact, not only the line DH is singular, but also the entire plane $V$. There should be a jump between the option price at a point very close to the plane $V$, but in the prism $IV$, and the option value at a point on the plane $V$. Financially, one could expect that the value of the former is larger than that of the latter, because for any point on the plane
there is a risk that \( \bar{J} \) will be reached, whereas for any point in the prism \( \overline{IV} \), no matter how close it is to the plane \( \overline{V} \), it is impossible for the trigger value \( \bar{J} \) to be reached.

The existence of the plane \( \overline{V} \) also introduces another singularity along the line QH. If we classify the line QH as part of the plane \( \overline{V} \), then the option price at any point \( S = \bar{S} \) on the plane EQHM but not on the line QH is non-zero, because no matter how close this point is to the line QH, \( J \) will be reset to zero, when the underlying touches the barrier. However, the option price on the line QH is always zero, because on this line, the trigger value \( \bar{J} \) is reached, and the “knock-out” takes place.

Of course, in reality, which side the line DH or QH belongs to could be clearly defined in a contract. What one must realize is that any ambiguity left in a Parisian option contract on the belonging of these two boundary lines would ultimately lead to different views in terms of the value of the contract in the event that the underlying asset price and the barrier time reach these very special values.

Now summarizing what has been said, the PDE system for pricing a European-style Parisian up-and-out call option under the BS model can be written as:

\[
\begin{align*}
\mathcal{A}_1: & \quad \left\{ \begin{array}{l}
\frac{\partial V_1}{\partial t} + LV_1 = 0, \\
V_1(S, T - \bar{J}) = V_{BS}(S, \bar{J}), \\
\lim_{S \to 0} V_1(S, t) = 0, \\
\lim_{S \to \bar{S}} V_1(S, t) = \lim_{S \to \bar{S}} V_2(S, t, 0),
\end{array} \right. \\
& \quad \text{for } t \in [0, T - \bar{J}], \bar{J} \in [0, \bar{J}], S \in [0, \bar{S}]; \\
\mathcal{A}_2: & \quad \left\{ \begin{array}{l}
\frac{\partial V_2}{\partial t} + \frac{\partial V_2}{\partial \bar{J}} + LV_2 = 0, \\
V_2(S, t, \bar{J}) = 0, \\
\lim_{S \to \infty} V_2(S, t, J) = 0, \\
\lim_{S \to \bar{S}} V_2(S, t, J) = \lim_{S \to \bar{S}} V_1(S, t),
\end{array} \right.
\end{align*}
\]
for \( t \in [J, T - \bar{J} - J] \), \( J \in [0, \bar{J}] \), \( S \in [\bar{S}, \infty) \);

\[
\text{Connectivity condition : } \lim_{S \to \bar{S}} \frac{\partial V_1}{\partial S}(S, t) = \lim_{S \to \bar{S}} \frac{\partial V_2}{\partial S}(S, t, 0), \text{ for } t \in [0, T - \bar{J}].
\] (11.2.11)

11.2.2 **Parasian options**

Contrary to the unique “resetting” feature of a Parisian option, Parasian options are introduced with no reset of the barrier time \( J \); the knock-in or knock-out is only activated if the cumulative time spent beyond or below \( \bar{S} \) exceeds some prescribed time \( \bar{J} \). Of course, like Parisian options, Parasian options can also have eight different forms. For the purpose of illustration, we shall only consider the pricing of a European-style Parasian up-and-out call option, as it should also be straightforward to extend our work to other forms of Parasian options by using the parity relationships established in [21].

Apart from this main difference, a Parasian option behaves quite similarly to its Parisian counterpart. For extreme barrier values, it also degenerates to either a classical barrier option or a vanilla option, just the same as what the corresponding Parisian option does. For brevity, we shall not repeat the details here. For the non-degenerate cases, the pricing of Parasian options also requires three state variables, i.e., the current time \( t \), the underlying \( S \) and the barrier time \( J \). While the former two variables \( t \) and \( S \) are assumed to follow the same dynamics as those with the Parisian option, the dynamics of the latter need to be modified so that it is not reset at \( \bar{S} \) [49], i.e.,

\[
dJ = \begin{cases} 
0, & S \leq \bar{S}; \\
\ dt, & S > \bar{S}.
\end{cases}
\]

With these three state variables, Wilmott et al. [49] also established a PDE system
governing the price of the ParAsian option as

\[
\begin{aligned}
&\begin{cases}
\frac{\partial V_1}{\partial t} + L V_1 = 0, \\
V_1(S, T) = \max(S - K, 0), \\
\lim_{S \to 0} V_1(S, t; J) = 0, \\
\lim_{S \to \bar{S}} V_1(S, t; J) = \lim_{S \to \bar{S}} V_2(S, t, J),
\end{cases}
\end{aligned}
\]
for \( S \in [0, \bar{S}] \);

\[
\begin{aligned}
&\begin{cases}
\frac{\partial V_2}{\partial t} + \frac{\partial V_2}{\partial J} + L V_2 = 0, \\
V_2(S, t, \bar{J}) = 0, \\
\lim_{S \to \infty} V_2(S, t, J) = 0,
\end{cases}
\end{aligned}
\]
for \( S \in [\bar{S}, \infty) \);

\text{for } t \in [0, T], J \in [0, \bar{J}].
\tag{11.2.12}
\]

Whilst elegant and financially meaningful, it should be pointed out that the boundary condition they imposed at \( S = \infty \) is at odds with the financial clause that the ParAsian option price at the large \( S \) end would never be zero, when \( T - t < \bar{J} - J \). This is because there is not enough time left for \( J \) to reach \( \bar{J} \) in this case, no matter how large the underlying is above \( \bar{S} \), as we have already discussed for the case of Parisian options in Section 2.1. One should also notice that the PDE system (11.2.12) is not properly closed, just like the case of Parisian options; an additional connectivity condition in terms of Delta needs to be prescribed across the barrier.

Unfortunately, taking away the reset feature has also removed the simplicity we have for Parisian options. One should notice that if the barrier time \( J \) no longer needs to be reset, the trigger clock containing some residues will affect the option price when the underlying falls below \( \bar{S} \). Consequently, comparing with Parisian options, the lower cuboid EFHGXOWR (denoted by \( V IT \) hereafter), as shown in Fig 11.1, should also be included as part of the pricing domain for ParAsian options, as \( J \) serves as a parameter in this particular region as well. This additional region,
resulting from the non-resetting mechanism specified in a ParAsian option contract, has no doubt added the complexity of solving for its option prices accurately and efficiently.

Despite a 3-D cuboid $\{III\}$ having replaced a 2-D plane $\{I\}$, the simplification of the pricing domain, which is a key step in the development of our analytic approach for finding the price of ParAsian options, can still proceed like in the case of Parisian options. In specific, both of the prisms ANBWOP (denoted by $\{III\}$ hereafter) and LCDRIX (denoted by $\{IV\}$ hereafter) can be excluded for the same reasons stated in Section 2.1. Then, following almost the same arguments demonstrated in Section 2.1, we can re-establish a properly closed PDE system governing the price of a ParAsian up-and-out call option as:

\[
A_1 : \begin{cases}
\frac{\partial V_1}{\partial t} + \mathbb{I}V_1 = 0, \\
V_1(S, T - \bar{J} + J; J) = V_{BS}(S, \bar{J} - J), \\
\lim_{S \to 0} V_1(S, t; J) = 0, \\
\lim_{S \to S} V_1(S, t; J) = \lim_{S \to S} V_2(S, t, J),
\end{cases}
\]

for $S \in [0, \bar{S}]$;

\[
A_2 : \begin{cases}
\frac{\partial V_2}{\partial t} + \frac{\partial V_2}{\partial J} + \mathbb{I}V_2 = 0, \\
V_2(S, t, \bar{J}) = 0, \\
\lim_{S \to \infty} V_2(S, t, J) = 0, \\
\lim_{S \to S} V_2(S, t, J) = \lim_{S \to S} V_1(S, t; J),
\end{cases}
\]

for $S \in [\bar{S}, \infty]$;

Connectivity condition : $\lim_{S \to \bar{S}} \frac{\partial V_1}{\partial S}(S, t; J) = \lim_{S \to \bar{S}} \frac{\partial V_2}{\partial S}(S, t, J)$, for $t \in [J, T - \bar{J} + J], J \in [0, \bar{J}]$.

Clearly, there are two major differences between the newly established PDE system (11.2.13) and the one proposed in [49] (i.e., the PDE system (11.2.12)). Firstly, the original terminal condition for $V_1$ in (11.2.12) has been replaced by a European call
option price at the time to expiry $\bar{J} - J$. Secondly, to properly close the PDE system (11.2.12), we have further assumed that the option Delta be continuous across the barrier, i.e.,

$$\lim_{S \to \bar{S}} \frac{\partial V_1}{\partial S}(S, t; J) = \lim_{S \to \bar{S}} \frac{\partial V_2}{\partial S}(S, t, J).$$  \hspace{1cm} (11.2.14)

Comparing (11.2.13) with (11.2.11), one can observe that pricing ParAsian options is much more complicated than pricing their Parisian counterparts. The complexity mainly comes from two aspects. Firstly, when $S < \bar{S}$, the ParAsian option price $V_1(S, t; J)$ varies w.r.t. the parameter $J$, whereas for a Parisian option, its price $V_1(S, t)$ is independent of $J$. Secondly, the connectivity condition of the ParAsian option should be applied across the entire barrier plane $S = \bar{S}$, while for its Parisian counterpart, this condition is only specified across the line $S = \bar{S}, J = 0$.

On the other hand, it should be pointed out that the two newly established PDE systems (11.2.11) and (11.2.13) are quite different from the corresponding ones used in [49]. One may wonder whether different pricing systems will yield the same option price. This issue will be further discussed in Section 4.1. It should also be remarked that although (11.2.11) and (11.2.13) are linear, they are both coupled systems, with the solution in one domain being the boundary condition for another domain. Whilst difficult, we still manage to develop two analytic approaches to solve (11.2.11) and (11.2.13), respectively. These approaches are discussed in the next section.

### 11.3 Our solution procedure

Although the newly established PDE systems are simpler, to some extent, than the original ones used in [49], they are still in 3-D, a difficulty that hampers the achievement of the analytic solutions. By taking advantage of the shape of the current pricing domain, however, the above 3-D PDE systems (11.2.11) and (11.2.13) can be further simplified to two 2-D PDE systems, which can then be easily solved analytically by utilizing the Laplace transform technique.

In this section, we shall demonstrate our solution procedure in detail. The trans-
formations from (11.2.11) and (11.2.13) to 2-D PDE systems will be discussed in the first subsection. Then in the second subsection, we shall explore the degenerations of the resulted 2-D PDE systems for some special cases. Finally, the details of analytically solving for the prices of the Parisian and Par\textit{asian} options from the 2-D PDE systems will be presented in the third and fourth subsections, respectively.

11.3.1 “Moving windows” and the 2-D PDE systems

Clearly, to transform (11.2.11) to a 2-D PDE system, we only need to deal with the system governing $V_2$, i.e., $A_2$, as the one that $V_1$ needs to satisfy is already in 2-D. On the other hand, to reduce dimensionality of a PDE system usually requires the application of some sorts of advanced transformation techniques, such as the Fourier transform, the Laplace transform, and so on. The main drawback of the traditional way is that either the resulted 2-D system could still be hard to solve analytically or, in the case that a solution can be found in the transformed space, the inverse of this solution could still not be found. For Parisian options, however, one additional variable in $A_2$ can still be elegantly “absorbed” without applying any transformation methods mentioned above, which will pave the way for our final analytical solution of this problem.

One can observe that the pricing domain $II$ is a parallelepipedon, and can be decomposed into infinite many cross-sections (which will be referred to as “slides” hereafter), all of which are of $45^\circ$ angles to both of the plane $t = 0$, and $J = 0$. However, for any given state point $(S, t, J)$ in $II$, there is a unique slide passing through that point. In light of this geometric characteristic of the domain $II$, it is clear that the option value $V_2$ at any given point $(S, t, J)$ can be uniquely determined as long as enough information along the very slide passing through that static point is known. In other words, the original 3-D problem can be decomposed into a set of 2-D problems defined on each slide, if viewed from a $45^\circ$ rotated coordinate system.

Mathematically, to obtain the PDE governing $V_2$ in the rotated coordinate system, we can replace the sum of the two partial derivatives appearing in $A_2$, i.e.,
\[ \frac{\partial V_2}{\partial t} + \frac{\partial V_2}{\partial J}, \text{ by the directional derivative } \sqrt{2} \frac{\partial V_2}{\partial l}, \text{ which represents the instantaneous rate of change of the function } V_2 \text{ at the point } (t, J), \text{ in the direction of } \left( \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2} \right), \]
i.e., a counter clockwise 45° plane viewed from the top. As a result, the governing equation in the new coordinate system can be written as

\[(\sqrt{2} \frac{\partial}{\partial l} + L)V_2(S, l; t) = 0. \quad (11.3.15)\]

Here \( t \) serves as a parameter, which identifies the slide passing through the point \((\bar{S}, t, 0)\), and thus the solution \( V_2(S, l; t) \) in the new coordinate system corresponds to \( V_2(S, t, J) \), with \( t \in [t, \frac{l}{\sqrt{2}} + t] \) and \( J = \frac{l}{\sqrt{2}} \) in the original 3-D coordinate system. In other words, the new solution \( V_2(S, l; t) \) equals \( V_2(S, t + \frac{l}{\sqrt{2}}, \frac{l}{\sqrt{2}}) \) in the original pricing domain \( II \).

It should be remarked that replacing the derivatives by a directional derivative along the \( l \) direction is also financially meaningful, because such a replacement corresponds to the restriction that when \( S > \bar{S} \), \( t \) and \( J \) increase at the same rate as stated in the definition of the “barrier time”, and thus the movement of the underlying can be viewed as if it were along the diagonal plane between \( t = 0 \) and \( J = 0 \). It is this restriction that has reduced the 3-D problem to a 2-D problem in a rotated coordinate system. Furthermore, the constant \( \sqrt{2} \) can be absorbed by rescaling \( l \), i.e., \( l' = \frac{l}{\sqrt{2}} \), and (11.3.15) becomes

\[ (\frac{\partial}{\partial l'} + L)V_2(S, l'; t) = 0, \quad (11.3.16)\]

which is nothing but the BS equation! Such a degeneration makes perfect sense too, because for \( S > \bar{S} \), the movements of the underlying are still the same as \( S < \bar{S} \), except that they should be viewed from a different space, and thus the governing equation in the rotated and rescaled coordinate system, i.e., along this 45° plane, should certainly be nothing but the BS equation.

The boundary conditions set for \( V_2(S, l'; t) \) can be extracted from the correspond-
ing boundary conditions that \( V_2(S, t, J) \) needs to satisfy. The fact that the trigger value \( \bar{J} \) will definitely be crossed if the underlying price is infinitely large gives

\[
\lim_{S \to \infty} V_2(S, \bar{J}; t) = 0.
\] (11.3.17)

Moreover, the knock-out condition in the rotated coordinate system now reads

\[
\lim_{l' \to \bar{J}} V_2(S, l'; t) = 0,
\] (11.3.18)

which can be viewed as the “terminal condition” along the \( l' \) direction. In addition, in the new coordinate system, the reset condition and the connectivity condition become

\[
V_2(S, l'; t) = V_1(S, t + l'), \quad \lim_{S \to 0} V_1(S, t) = 0, \quad \lim_{S \to \bar{S}} V_2(S, 0; t) = \lim_{S \to \bar{S}} \frac{\partial V_1}{\partial S}(S, t),
\]
respectively, simply because the solution \( V_2(S, l'; t) \) now corresponds to \( V_2(S, t+l', l') \) in the original pricing domain \( II \), as demonstrated earlier.

Therefore, the 2-D PDE systems that govern the price of a Parisian up-and-out call option can be now summarized as:

\[
\mathcal{A}_1(S \in [0, \bar{S}]) \begin{cases} \\
\frac{\partial V_1}{\partial t} + \mathcal{L}V_1 = 0, \\
V_1(S, T - \bar{J}) = V_{BS}(S, \bar{J}), \\
\lim_{S \to 0} V_1(S, t) = 0, \\
\lim_{S \to \bar{S}} V_1(S, t) = W(t), \\
\end{cases}
\]

\[
\mathcal{A}_2(S \in [\bar{S}, \infty)) \begin{cases} \\
\frac{\partial V_2}{\partial l'} + \mathcal{L}V_2 = 0, \\
V_2(S, \bar{J}; t) = 0, \\
\lim_{S \to \infty} V_2(S, \bar{J}; t) = 0, \\
\lim_{S \to \bar{S}} V_2(S, l'; t) = W(t + l'), \\
\end{cases}
\]

Connectivity condition: \( \lim_{S \to \bar{S}} \frac{\partial V_1}{\partial S}(S, t) = \lim_{S \to \bar{S}} \frac{\partial V_2}{\partial S}(S, 0; t), \)

for \( t \in [0, T - \bar{J}], l' \in [0, \bar{J}], \) (11.3.19)
with \( W(t) \) being defined as \( \lim_{S \to \bar{S}} V_2(S, 0; t) \), which needs to be solved as part of the solution.

Similarly, we can also deduce the 2-D systems for the valuation of a Parisian up-and-out call option as:

\[
\begin{align*}
\mathcal{A}_1(S \in [0, \bar{S}]) & \quad \begin{cases} \\
\frac{\partial V_1}{\partial t} + \mathbb{L} V_1 = 0, \\
V_1(S, T - \bar{J} + J; J) = V_{BS}(S, \bar{J} - J), \\
\lim_{S \to 0} V_1(S, t; J) = 0, \\
\lim_{S \to \bar{S}} V_1(S, t; J) = W(t, J), \\
\end{cases} \\
\mathcal{A}_2(S \in [\bar{S}, \infty)) & \quad \begin{cases} \\
\frac{\partial V_2}{\partial l'} + \mathbb{L} V_2 = 0, \\
V_2(S, \bar{J}; t) = 0, \\
\lim_{S \to \infty} V_2(S, l'; t) = 0, \\
\lim_{S \to \bar{S}} V_2(S, l'; t) = W(t + l', l'), \\
\end{cases}
\end{align*}
\]

Connectivity condition: \( \lim_{S \to \bar{S}} \frac{\partial V_1}{\partial S}(S, t; J) = \lim_{S \to \bar{S}} \frac{\partial V_2}{\partial S}(S, J; t - J) \),

for \( t \in [0, T - \bar{J}], l' \in [0, \bar{J}] \), \hspace{1cm} (11.3.20)

with \( W(t, J) \) being defined as \( \lim_{S \to \bar{S}} V_2(S, J; t - J) \).

One should notice that the coupling of \( V_1 \) and \( V_2 \) is now through a function of both \( t \) and \( J \), i.e., \( W(t, J) \), rather than a function of \( t \) only for the Parisian option. This additional variable \( J \) explicitly appearing in \( W \) is a result of the non-resetting mechanism specified in the Parisian option contract. It is also this additional variable that has made the solution procedure for the Parisian option much more complicated than and totally different from that of its Parisian counterpart.

### 11.3.2 Degeneration of the 2-D systems

An efficient way to verify the derivation of (11.3.19) from (11.2.11) and that of (11.3.20) from (11.2.13) is to investigate the degenerations of (11.3.19) and (11.3.20) with the involved parameters taking some extreme values, because whether or not
the degenerate cases agree with the financial terms set for the corresponding options is a necessary condition to verify these 2-D systems. There are four special cases that need to be examined.

The first case is associated with the zero value of the trigger time $\bar{J}$. One can verify that the terminal condition appearing in $\mathcal{A}_1$ of (11.3.19) becomes the payoff function of a European call option, by simply taking the limit process, i.e.,

$$V_1(S, T) = \lim_{\bar{J} \to 0} V_1(S, T - \bar{J}) = \lim_{\bar{J} \to 0} V_{BS}(S, T - \bar{J}) = \max(S - K, 0).$$

Moreover, it is also true that the variable $l'$ approaches zero as well in the current case, because $l'$ varies within the range $[0, \bar{J}]$. As a result, $\mathcal{A}_2$ in (11.3.19) vanishes, leaving $V_2(S, 0; t) = 0$ valid for $S \in [\bar{S}, \infty)$, $t \in [0, T]$. Combining these two points, it is clear that when $\bar{J} \to 0$, the pricing system (11.3.19) degenerates to

$$\begin{align*}
\frac{\partial V_1}{\partial t} + LV_1 &= 0, \\
V_1(S, T) &= \max(S - K, 0), \\
\lim_{S \to 0} V_1(S, t) &= 0, \\
\lim_{S \to \bar{S}} V_1(S, t) &= 0,
\end{align*}$$

for $S \in [0, \bar{S}], t \in [0, T]$.

This is indeed the same as the PDE system governing the price of a barrier call option with up-and-out feature, as expected.

Secondly, when $\bar{J} \to \infty$, it is clear that for any finite option maturity value $T$, $T - \bar{J}$ is always less than zero, which is equivalent to saying that both $\mathcal{A}_1$ and $\mathcal{A}_2$ defined in (11.3.19) vanish. Geometrically, the above statement reveals that the pricing domain in the current case is in fact part of the prism $\mathcal{IV}$. Since the value in the prism $\mathcal{IV}$ is defined as the price of a European call option, the degeneration as $\bar{J} \to \infty$ becomes obvious, and does indeed agree with the financial definition of a Parisian up-and-out call option contract.
Chapter 11.

The third case is when the barrier $\tilde{S}$ approaches zero. Since the option value for any $t \in (T - \bar{J}, T]$ is already defined as a European call option price, we only need to consider the case when $t \in [0, T - \bar{J}]$. Clearly, when $\tilde{S} \to 0$, $A_1$ in (11.3.19) vanishes, leaving $V_1(0, t) = 0$, and thus, for any $t \in [0, T - \bar{J}]$, we obtain

$$W(t) = \lim_{\tilde{S} \to 0} V_1(S, t) = \lim_{\tilde{S} \to 0} V_1(S, t) = 0.$$  

In fact, we could conclude that $W(t)$ equals zero for any $t \in [0, T]$, because in the additional region $(T - \bar{J}, T]$, $W(t)$ is defined as the value of a European call option with zero underlying price, i.e.,

$$W(t) = \lim_{\tilde{S} \to 0} V_{BS}(S, t) = 0.$$  

Now, it is clear that $W(t + \ell')$ appearing in $A_2$ equals zero, as a result of $t + \ell'$ varying within $[0, T]$. Consequently, $A_2$ becomes a homogenous system defined on $t \in [0, T - \bar{J}]$, $S \in [0, \infty]$, $J \in [0, \bar{J}]$, and thus $A_2$ only has a trivial solution. Therefore, when $\tilde{S} \to 0$, the Parisian up-and-out call option price calculated from our newly defined system (11.3.19) is indeed equal to a European call option price if $t \in [0, T - \bar{J}]$, and zero otherwise.

Finally, when $\tilde{S} \to \infty$, it is clear that $A_2$ vanishes, and moreover, $A_1$ degenerates to the following backward problem:

\[
\begin{align*}
\frac{\partial V_1}{\partial t} + LV_1 &= 0, \\
V_1(S, T - \bar{J}) &= V_{BS}(S, T - \bar{J}), \\
\text{for } S \in [0, \infty), t \in [0, T - \bar{J}],
\end{align*}
\]

the solution of which is identical to a European call option price at the time to expiry $T - t$. In other words, a European call option can be viewed as a Parisian up-and-out call option without barriers. Therefore, the result of such a degeneration mathematically indeed matches with the financial interpretation.
Although the above arguments are established for a Parisian up-and-out call option, they also hold under the framework of a corresponding ParAsian option. Therefore, we shall not repeat the details here for the case of ParAsian options.

From the above discussions, one can thus conclude that, with the parameters taken on some extreme values, the 2-D PDE systems (11.3.19) and (11.3.20) indeed degenerate to the ones that could be expected from the financial point of view. This has somehow confirmed that the way of establishing (11.3.19) and (11.3.20) is valid and reasonable. Having gained confidence on the correctness of the newly established 2-D PDE systems, two purely analytic approaches will be developed in the next two subsections to solve (11.3.19) and (11.3.20), respectively.

11.3.3 The analytical solution for the Parisian option price

To solve the newly established pricing system (11.3.19) effectively, we shall first non-dimensionalize all variables by introducing dimensionless variables

\[
x = \ln \frac{S}{K}, \bar{x} = \ln \frac{\bar{S}}{K}, V_1' = \frac{V_1}{K}, V_2' = \frac{V_2}{K}, \tau = (T - \bar{J} - t) \frac{\sigma^2}{2}, W' = \frac{W}{K},
\]

\[
\bar{J} = \frac{\sigma^2 \bar{j}}{2}, \bar{t} = \frac{\sigma^2}{2}(\bar{J} - \bar{t}), \gamma = \frac{2r}{\sigma^2}, q = \frac{2D}{\sigma^2}.
\]

With all primes and tildes dropped from now on, the dimensionless PDE system reads:

\[
\mathcal{A}_1(x \in (-\infty, \bar{x}]) \begin{cases}
\frac{\partial V_1}{\partial \tau} = LV_1, \\
V_1(x, 0) = V_{BS}(Ke^{\bar{x}}, \frac{2\bar{J}}{\sigma^2})/K, \\
\lim_{x \to -\infty} V_1(x, \tau) = 0, \\
\lim_{x \to \bar{x}} V_1(x, \tau) = W(\tau),
\end{cases}
\]
\[ A_2(x \in [\bar{x}, +\infty)) \begin{cases} \frac{\partial V_2}{\partial t} = \mathbb{L}V_2, \\
V_2(x, 0; \tau) = 0, \\
\lim_{x \to \infty} V_2(x, l; \tau) = 0, \\
\lim_{x \to \bar{x}} V_2(x, l; \tau) = W(\tau - \bar{J} + l), 
\end{cases} \]

Connectivity condition: \[ \lim_{x \to \bar{x}} \frac{\partial V_1}{\partial x}(x, \tau) = \lim_{x \to \bar{x}} \frac{\partial V_2}{\partial x}(x, \bar{J}; \tau), \]
for \( \tau \in [0, \bar{\tau}], l \in [0, \bar{J}], \) \hspace{1cm} (11.3.22)

where operator \( \mathbb{L} = \frac{\partial^2}{\partial x^2} + k \frac{\partial}{\partial x} - \gamma I, \) with \( k \) being equal to \( \gamma - q - 1. \)

From (11.3.22), it can be observed that once \( W(\tau) \) is found, \( V_1 \) and \( V_2 \) are no longer coupled, and the corresponding solutions \( V_1 \) and \( V_2 \) can be obtained straightforwardly. While it seems quite natural to treat the determination of \( W(\tau) \) as a key step of solving (11.3.22), it is, however, not an easy task.

Since the time to expiry \( \tau \) is the only variable of the unknown function \( W(\tau), \) one could expect \( W(\tau) \) to be governed by some sort of simple equations (e.g. ODE (ordinary differential equation), integral equation, integral ODE, etc.), whose solution could be more analytically achievable, rather than to solve for \( W(\tau) \) directly from the coupled system (11.3.22).

To find the governing equation for \( W(\tau) \), we shall first find the integral representations of \( V_1 \) and \( V_2, \) in terms of the unknown function \( W(\tau) \). This can be achieved by solving \( A_1 \) and \( A_2 \) in (11.3.22) separately, as if they were not coupled. Specifically, the solution of \( A_1 \) can be found by splitting the linear problem into two problems, a technique frequently used in solving linear PDEs. The solution of the first problem, which involves a homogeneous differential equation and homogeneous boundary conditions but arbitrary initial condition, can be easily worked out utilizing the Laplace transform technique as well as the Green function method for the resulted ODE, while the solution of the second problem, which involves a homogeneous differential equation and zero initial condition but non-homogeneous boundary condition at \( x = \bar{x} \) can be obtained by applying the Laplace transform technique. Without going
through the lengthy derivation procedures, the integral representation of $V_1$ can be written as

$$V_1(x, \tau) = F(x, \tau) + \int_0^\tau W(s)g_1(x, \tau - s)ds, \quad (11.3.23)$$

where

$$F(x, \tau) = \int_{-\infty}^x \frac{1}{2\sqrt{\pi \tau}} e^{-\frac{1}{2}(x-z)^2} \left( e^{-\frac{(x-z)^2}{4\tau}} - e^{-\frac{(x-z-2z)^2}{4\tau}} \right) |f(z)dz,$$

$$g_1(x, \tau) = -\frac{x-\bar{x}}{2\sqrt{\pi \tau^2}} e^{-(k^2/4+\gamma)\tau - \frac{(x-\bar{x})^2}{4\tau} - \frac{1}{2}(x-\bar{x})},$$

$$f(z) = V_{BS}(z, \bar{J}).$$

Similarly, $V_2$ can be written as

$$V_2(x, l) = \int_0^l W(\tau - \bar{J} + s)g_2(x, l - s)ds, \quad (11.3.24)$$

where $g_2(x, l) = -g_1(x, l)$. Now, applying the connectivity condition to (11.3.23) and (11.3.24), we obtain the integral equation governing $W(\tau)$ as

$$\frac{\partial F}{\partial x}(x, \tau)|_{x=\bar{x}} + \int_0^\tau W(s) \frac{\partial g_1}{\partial x}(x, \tau - s)ds|_{x=\bar{x}} = \int_0^\tau W(\tau - \bar{J} + s) \frac{\partial g_2}{\partial x}(x, \bar{J} - s)ds|_{x=\bar{x}}. \quad (11.3.25)$$

It can be observed that the LHS (left hand side) of (11.3.25) contains the information of $W(s)$ from the expiry ($\tau = 0$) to the current time to expiry, $\tau$, while its RHS (right hand side) involves the value of $W(s)$, with $s$ varying within $[\tau - \bar{J}, \tau]$, which coincides with the projection of the slide passing through $(\bar{S}, \tau, 0)$ on the plane $J = 0$. Note that for simplicity, we shall name the projection of a slide on the plane $J = 0$ as a “window” from now on. The initial window (the zeroth window) is thus the one with $\tau$ varying within $[-\bar{J}, 0]$, where $W(\tau)$ is defined as the price of a European call option at $S = \bar{S}$. Intuitively, by solving (11.3.25) once, it could at most be expected to have the unknown function $W$ determined on one particular window, rather than over the entire domain $[0, \tau]$, unless $[0, \tau]$ belongs to the first window, which corresponds to the case $0 \leq \tau \leq \bar{J}$. 
Consequently, for a state point \((S, \tau, J)\), one can evaluate \(W\) forwards, window by window, until the value at the required time \(\tau\) is found. Such a procedure is equivalent to treating the window as if it were moving, each time with a \(J\) distance, from the initial window to the place containing the given \(\tau\) value. Therefore, determining the corresponding \(W\) value on the window moving once from the initial window \([-\bar{J}, 0]\) becomes crucial, as the procedure can be repeated until the desired place is reached. In the following, we shall first solve for \(W(\tau), \ \tau \in [0, \bar{J}]\), which lies on the window that has moved once from the initial place. Hereafter, the window that moves \(n\) \((n \geq 1)\) times from the initial window is called the \(n\)th window for simplicity.

Now defining a new variable \(\xi = \tau - \bar{J} + s\), and substituting it into the RHS of (11.3.25), we obtain

\[
\int_{\bar{J}}^{0} \left. W(\tau - \bar{J} + s) \frac{\partial g_2}{\partial x}(x, \bar{J} - s) \right|_{x=\bar{x}} = \int_{\tau - \bar{J}}^{\tau} \left. W(\xi) \frac{\partial g_2}{\partial x}(x, \tau - \xi) \right|_{x=\bar{x}} \right|_{x=\bar{x}}. \tag{11.3.26}
\]

Since \(W(\tau)\) is already defined for \(\forall \tau \in [-\bar{J}, 0]\), it is reasonable to split (11.3.26) into two terms, i.e.,

\[
\int_{\tau - \bar{J}}^{\tau} \left. W(\xi) \frac{\partial g_2}{\partial x}(x, \tau - \xi) \right|_{x=\bar{x}}\right|_{x=\bar{x}} = \int_{\tau - \bar{J}}^{0} \left. W(\xi) \frac{\partial g_2}{\partial x}(x, \tau - \xi) \right|_{x=\bar{x}}\right|_{x=\bar{x}} + \int_{0}^{\tau} \left. W(\xi) \frac{\partial g_2}{\partial x}(x, \tau - \xi) \right|_{x=\bar{x}}\right|_{x=\bar{x}}. \tag{11.3.27}
\]

with the first term already known. Substituting (11.3.27) into (11.3.25), we obtain

\[
\left. \left[ \frac{\partial F}{\partial x}(x, \tau) + \int_{0}^{\tau} W(s) \frac{\partial g_1}{\partial x}(x, \tau - s) ds \right] \right|_{x=\bar{x}} = \left[ \int_{\tau - \bar{J}}^{\tau} \left. W(\xi) \frac{\partial g_2}{\partial x}(x, \tau - \xi) \right|_{x=\bar{x}}\right|_{x=\bar{x}} + \int_{0}^{\tau} \left. W(\xi) \frac{\partial g_2}{\partial x}(x, \tau - \xi) \right|_{x=\bar{x}}\right|_{x=\bar{x}}. \tag{11.3.28}
\]

Clearly, the unknown function \(W(\xi)\) \((\xi \in [0, \tau])\) is now only involved in the two convolutions that appear on both sides of (11.3.28). According to the Convolution Theorem [77], the convolution integral can be further eliminated by taking the Laplace
transform w.r.t. $\tau$, and thus we obtain
\[
\hat{H}(x, p)|_{x = \bar{x}} + \hat{W} \frac{\partial \hat{g}_1}{\partial x}(x, p)|_{x = \bar{x}} = \hat{G}(x, p)|_{x = \bar{x}} + \hat{W} \frac{\partial \hat{g}_2}{\partial x}(x, p)|_{x = \bar{x}},
\]
(11.3.29)

where
\[
\hat{H}(x, p) = \mathcal{L}\left[\frac{\partial F}{\partial x}(x, \tau)\right], \quad \hat{G}(x, p) = \mathcal{L}\left[\int_{\tau - \bar{x}}^{0} W(\xi) \frac{\partial g_2}{\partial x}(x, \tau - \xi) d\xi\right],
\]
\[
\hat{g}_1(x, p) = \mathcal{L}[g_1(x, \tau)] = e^{\lambda_1(x - \bar{x})}, \quad \hat{g}_2(x, p) = \mathcal{L}[g_2(x, \tau)] = e^{\lambda_2(x - \bar{x})},
\]
\[
\hat{W}(p) = \mathcal{L}[W(\tau)],
\]

with
\[
\lambda_1 = -\frac{k}{2} + \sqrt{\left(\frac{k^2}{4} + \gamma\right) + p}, \quad \lambda_2 = -\frac{k}{2} - \sqrt{\left(\frac{k^2}{4} + \gamma\right) + p},
\]

and $p$ being the Laplace parameter. After some simple algebraic manipulations, the unknown function $W$ in the Laplace space can be solved as
\[
\hat{W}(p) = \frac{\hat{H}(x, p)|_{x = \bar{x}} + \hat{G}(x, p)|_{x = \bar{x}}}{\lambda_2 - \lambda_1}.
\]
(11.3.30)

Although (11.3.30) is remarkably simple, it is unfortunately still in terms of the Laplace parameter $p$. In order to obtain an analytical formula for $W(\tau)$, one still needs to carry out the Laplace inversion, a formidable process that often prevents this great technique being widely used to solve PDEs. However, the significance of (11.3.30) should never be underestimated, even though it is in terms of the Laplace parameter $p$. As pointed out previously, finding the option price across the barrier, i.e., $W(\tau)$, is a key step to solve the Parisian option price. Once $W(\tau)$ is found, $V_1$ and $V_2$ are no longer coupled, and can be obtained analytically through (11.3.23)-(11.3.24).

To analytically invert (11.3.30) is quite difficult, but fortunately, achievable. With the derivation details left in Appendix E.1 and Appendix E.1, we obtain the fully
explicit analytic expression for $W(\tau)$ as

\[
W(\tau) = \mathcal{L}^{-1}\left[ \frac{\hat{H}(x,p)}{\lambda_2 - \lambda_1} \right]_{x = \tilde{x}} + \mathcal{L}^{-1}\left[ \frac{\hat{G}(x,p)}{\lambda_1 - \lambda_2} \right]_{x = \tilde{x}} = \int_{-\infty}^{\tilde{x}} e^{-\frac{1}{2}(\tilde{x}-z)^2 - (\frac{k^2}{4} + \gamma)\tau} f(z) dz + \frac{W(0)}{2} e^{-(\frac{k^2}{4} + \gamma)\tau} - \frac{e^{-(\frac{k^2}{4} + \gamma)n\bar{J}}}{2\pi \sqrt{n\bar{J}}} \int_{0}^{\tau} \frac{e^{-(\frac{k^2}{4} + \gamma)\tau - s}}{\sqrt{\tau - s}} W(s - \bar{J}) ds - \frac{1}{\pi} \int_{0}^{\tau} \frac{e^{-(\frac{k^2}{4} + \gamma)(\tau - s)}}{\sqrt{\tau - s}} \int_{\sqrt{\tau}}^{\bar{J}} e^{-(\frac{k^2}{4} + \gamma)t^2} [(\frac{k^2}{4} + \gamma)W(s - t^2) + W'(s - t^2)] dt ds,
\]

(11.3.31)

where the function $W(y), (y \in [-\tilde{J}, 0])$ involved in the RHS is already defined as the price of a European call option at the barrier $\tilde{S}$, with the time to expiry being equal to $y + \bar{J}$. Considering the complexity of the problem, this explicit and closed-form solution for the price across the barrier is remarkably simple. Moreover, from the viewpoint that even the price of a standard barrier option with rebate involves the calculation of double integrals, it is reasonable for us to believe that (11.3.31) is already in its simplest form.

It should be pointed out that although (11.3.31) is valid only on the first window, the detailed derivation can be in fact generalized to determine $W$ on the $(n + 1)$th window, assuming that the option price on the $n$th window has already been found. This can be easily achieved by realizing the following facts. Firstly, according to the semi-group property of the solution of a general heat transfer problem [82], one can solve $V_1(x, \tau) (\tau \in [n\bar{J}, (n + 1)\bar{J}])$ as if the diffusion started at $n\bar{J}$, with the initial condition now being equal to $f(x) = V_1(x, n\bar{J})$, and the time length being $\tau - n\bar{J}$.

Secondly, solving $W$ on the $(n + 1)$th window with the known option price on the $n$th window is equivalent to determining the $W$ value on the first window defined in the new coordinate system $\tilde{\tau} = \tau - n\bar{J}$. Therefore, for $\tau \in [n\bar{J}, (n + 1)\bar{J}]$, $W(\tau)$ has the same expression as (11.3.31), except that in this case, $f(z) = V_1(z, n\bar{J})$, and the function $W(y), (y \in [(n - 1)\bar{J}, n\bar{J}])$ involved in the RHS is defined as the price of a Parisian up-and-out call option at the barrier $\tilde{S}$, with the time to expiration being...
More specifically, the closed-form analytic formula for the Parisian up-and-out call option at any given \( \tau \) can be written as:

\[
W(n) = \begin{cases} 
W_0, & \tau \in [-\bar{J}, 0] \\
W_1, & \tau \in [0, \bar{J}] \\
\ldots, & \\
W_n, & \tau \in [(n-1)\bar{J}, n\bar{J}] \\
W_{n+1}, & \tau \in [n\bar{J}, (n+1)\bar{J}], 
\end{cases}
\]

where \( W_{n+1} \) can be found recursively as:

\[
W_{n+1}(\tau) = \int_{-\infty}^{\bar{J}} e^{-\frac{1}{2}(\bar{x}-z)^2 - (\frac{k^2}{4} + \gamma)(\tau-n\bar{J})} e^{-\frac{(\bar{x}-z)^2}{2}} f_n(z) dz + \frac{W_n(n\bar{J})}{2} e^{-(\frac{k^2}{4} + \gamma)(\tau-n\bar{J})} e^{-\frac{(\bar{x}-z)^2}{2}} f_n(z) dz + \\
-\frac{1}{2\pi} e^{-\frac{(k^2}{4} + \gamma)(\tau-n\bar{J})} \int_{n\bar{J}}^{\bar{J}} e^{-\frac{(k^2}{4} + \gamma)(\tau-s)} W_n(s-n\bar{J}) ds - \frac{1}{\pi} \int_{n\bar{J}}^{\bar{J}} \frac{e^{-\frac{(k^2}{4} + \gamma)(\tau-s)}}{\sqrt{\tau-s}} W_n(s-n\bar{J}) ds - \\
\int_{n\bar{J}}^{\tau} \frac{e^{-\frac{(k^2}{4} + \gamma)(\tau-s)}}{\sqrt{\tau-s}} \int_{s-n\bar{J}}^{\bar{J}} e^{-\frac{(k^2}{4} + \gamma)t^2} dt ds,
\]

for \( n = 0, 1, 2 \ldots \), with

\[
\begin{align*}
f_0(z) &= V_{BS}(z, \bar{J}), \\
f_n(x) &= F(x, n\bar{J}) + \sum_{i=1}^{n} \int_{(i-1)\bar{J}}^{i\bar{J}} W_i(s) g_1(x, n\bar{J} - s) ds, \quad n = 1, 2 \ldots \\
F(x, \tau) &= \frac{1}{2\pi} \int_{-\infty}^{\bar{x}} e^{-\frac{1}{2}(x-y)^2 - (\frac{k^2}{4} + \gamma)\tau} [e^{-\frac{(x-y)^2}{2}} - e^{-\frac{(x+y-2\bar{J})^2}{4}}] f_0(z) dz, \\
W_0(\tau) &= V_{BS}(x, \tau + \bar{J}), \quad \tau \in [-\bar{J}, 0].
\end{align*}
\]

The recursive nature of this analytic formula guarantees that \( W(\tau) \) can be evaluated explicitly and for sure analytically, window by window, until the window containing the given \( \tau \) value has been reached. Once \( W(\tau) \) is solved, the price of a Parisian up-and-out call option can be calculated straightforwardly by means of (11.3.23) and (11.3.24).
11.3.4 The analytical solution for the ParAsian option price

Having successfully found the analytic formula for the Parisian option price, we now move to solve for the price of its ParAsian counterpart. Similar to the solution procedure of the Parisian option, a key step here is to determine the ParAsian option price across the barrier, i.e., \( W(\tau, J) \). Once \( W(\tau, J) \) values are found, the calculation of the ParAsian option price becomes straightforward, too. However, one should notice that \( W(\tau, J) \) is now a function of both \( t \) and \( J \), rather than a function of \( t \) only for the case of Parisian options. With an additional variable \( J \) added on, one could except that the analytical expression of \( W(\tau, J) \) would be much more complicated than and totally different from \( W(\tau) \) of the Parisian option. Moreover, the involvement of the additional variable \( J \) in \( W \) has prevented the application of the exactly same solution procedure adopted in the case of pricing Parisian options for the current case; a different approach to solve for \( W(\tau, J) \) needs to be explored and is the focus of this subsection.

To solve (11.3.20) effectively, we shall also non-dimensionalize this system first. By using the same dimensionless variables introduced in the last subsection, and dropping all the primes and tildes afterwards, we obtain,

\[
\mathcal{A}_1(x \in (-\infty, \bar{x}]) \quad \begin{cases} 
\frac{\partial V_1}{\partial \tau} = \mathcal{L} V_1, \\
V_1(x, -J; J) = V_{BS}(K e^{x}, \frac{2(\bar{J} - J)}{\sigma^2})/K, \\
\lim_{x \to -\infty} V_1(x, \tau; J) = 0, \\
\lim_{x \to \bar{x}} V_1(x, t; J) = W(\tau, J),
\end{cases}
\]

Where \( \mathcal{L} \) is the differential operator defined in the last subsection.
\[ \mathcal{A}_2(x \in [\bar{x}, +\infty)) \begin{cases} \frac{\partial V_2}{\partial l} = LV_2, \\ V_2(x, 0; \tau) = 0, \\ \lim_{x \to -\infty} V_2(x, l; \tau) = 0, \\ \lim_{x \to \bar{x}} V_2(x, l; \tau) = W(\tau - \bar{J} + l, \bar{J} - l), \end{cases} \]

Connectivity condition: \( \lim_{x \to \bar{x}} \frac{\partial V_1}{\partial x}(x, \tau - \bar{J} + l; \bar{J} - l) = \lim_{x \to \bar{x}} \frac{\partial V_2}{\partial x}(x, l; \tau) \), for \( \tau \in [0, \bar{\tau}], l \in [0, \bar{J}] \), (11.3.32)

where the operator \( L = \frac{\partial^2}{\partial x^2} + k \frac{\partial}{\partial x} - \gamma I \), with \( k \) being equal to \( \gamma - q - 1 \).

Then, we solve for the integral representations of \( V_1 \) and \( V_2 \), in terms of the unknown function \( W(\tau, J) \) and obtain

\[ V_1(x, \tau, J) = F(x, \tau; J) + \int_0^{\tau + J} W(s - J, J)g_1(x, \tau + J - s)ds, \] (11.3.33)
\[ V_2(x, l; \tau) = \int_0^l W(\tau - \bar{J} + s, \bar{J} - l)g_2(x, l - s)ds, \] (11.3.34)

where

\[ F(x, \tau; J) = \int_{-\infty}^x \frac{1}{2\sqrt{\pi}(\tau + J)} e^{-\frac{(x-z)^2}{2(\tau + J)}} e^{-\frac{(x+z-\bar{J})^2}{4(\tau + J)-e^{-\frac{(x+z-\bar{J})^2}{4(\tau + J)}}}} f(z)dz, \]

with the functions \( g_1(x, \tau), g_2(x, l), \) and \( f(z) \) being the same as those defined in the last subsection. Now, applying the connectivity condition to (11.3.33) and (11.3.34), we obtain the integral equation governing \( W(\tau, J) \) as

\[ \frac{\partial F}{\partial x}(x, \tau - \bar{J} + l; \bar{J} - l)|_{x=\bar{x}} + \int_0^\tau W(s + l - \bar{J}, \bar{J} - l) \frac{\partial g_1}{\partial x}(x, \tau - s)ds|_{x=\bar{x}} = \int_0^l W(\tau - \bar{J} + s, \bar{J} - s) \frac{\partial g_2}{\partial x}(x, l - s)ds|_{x=\bar{x}}. \] (11.3.35)

To facilitate the analytical procedure, we denote \( W(\tau, l) = W(\tau + l - \bar{J}, \bar{J} - l) \),
and thus rewrite (11.3.35) as

\[
\frac{\partial F}{\partial x}(x, \tau - \bar{J} + l; \bar{J} - l)|_{x = \bar{x}} + \int_0^{\tau} W(s, l) \frac{\partial g_1}{\partial x}(x, \tau - s) ds |_{x = \bar{x}}
\]

\[
= \int_0^{l} W(\tau, s) \frac{\partial g_2}{\partial x}(x, l - s) ds |_{x = \bar{x}}.
\]

One can observe that this integral equation governing \(\bar{W}(\tau, l)\) involves two different convolutions w.r.t. two variables, i.e., \(\tau\) and \(l\). To eliminate the two convolutions, a double Laplace transform is applied, which yields

\[
\hat{H}(x, p_1, p_2)|_{x = \bar{x}} + \hat{W}(p_1, p_2) \lambda_1(p_1) = \hat{W}(p_1, p_2) \lambda_2(p_2),
\]

(11.3.36)

where

\[
\hat{H}(x, p_1, p_2) = \mathcal{L}_{p_1}[\mathcal{L}_{p_2}[\frac{\partial F}{\partial x}(x, \tau - \bar{J} + l; \bar{J} - l)]], \quad \hat{W}(p_1, p_2) = \mathcal{L}_{p_1}[\mathcal{L}_{p_2}[W(\tau, l)]],
\]

\[
\lambda_1(p_1) = -\frac{k}{2} + \sqrt{\left(\frac{k^2}{4} + \gamma\right) + p_1}, \quad \lambda_2(p_2) = -\frac{k}{2} - \sqrt{\left(\frac{k^2}{4} + \gamma\right) + p_2},
\]

with \(p_1, p_2\) being the Laplace parameters corresponding to the original variables \(\tau\) and \(l\), respectively, and the subscripts of the Laplace operator denote the corresponding Laplace transform.

From (11.3.36), it is clear that the unknown function \(\bar{W}\) in the Laplace space can be found as

\[
\hat{W}(p_1, p_2) = \frac{\hat{H}(x, p_1, p_2)}{\lambda_2(p_2) - \lambda_1(p_1)}|_{x = \bar{x}},
\]

(11.3.37)

and thus

\[
\bar{W}(\tau, l) = \mathcal{L}_{p_2}^{-1}[\mathcal{L}_{p_1}^{-1}\left[\frac{\hat{H}(x, p_1, p_2)}{\lambda_2(p_2) - \lambda_1(p_1)}|_{x = \bar{x}}\right]].
\]

(11.3.38)

Like performing Laplace inversion analytically for a single Laplace transform encountered for the case of Parisian options, analytically performing a Laplace inversion of (11.3.38) was initially considered as an even-more luxury and we were going to resort to a numerical inversion as did in [126]. Fortunately, we eventually managed
to have carried out this seemingly-impossible task and obtained a fully closed-form analytic expression for $\tilde{W}(\tau, l)$ as

\[
\tilde{W}(\tau, l) = \frac{\sqrt{2}}{\sqrt{\pi}} e^{-\left(\frac{l^2}{4} + \frac{\tau}{2}\right) (\tau+l)} \int_0^\infty \left\{ e^{(1+\frac{k}{2})^2 l + \frac{x^2}{4} - \sqrt{2\tau(l+1)\eta}} N\left(\frac{x - \sqrt{2\tau\eta}}{\sqrt{2l}}\right) + \frac{\sqrt{2l}}{2} \right\} d\eta \\
- e^{\frac{k^2}{4} l - \frac{\tau^2}{2} - k\eta} N\left[ \frac{\tau - \sqrt{2\tau\eta}}{\sqrt{2l}} + \frac{\sqrt{2l}k}{2} \right] d\eta \\
- \frac{\sqrt{2}}{\pi} e^{-\left(\frac{l^2}{4} + \frac{\tau}{2}\right) (\tau+l)} \int_0^\frac{\pi}{2(\tau+l)} e^{rac{x^2}{2}(\tau+l)\xi} \left( \frac{e^{\frac{x^2}{2}(\tau+l)\xi} \eta^2}{2\eta^2} - 1 \right) - N\left(\sqrt{\frac{l}{\xi}}\right) d\xi \\
+ \frac{2\sqrt{2}}{\pi} e^{-\left(\frac{l^2}{4} + \frac{\tau}{2}\right) (\tau+l)} \\
\int_{-\infty}^\frac{x}{\sqrt{\eta^2}} e^{-\frac{2^2}{2} \frac{(x-z)^2}{2\eta^2}} \int_0^\frac{\pi}{\sqrt{2\eta}} f_3(\theta, \tau - \frac{(x - z)^2}{2\eta^2}, z) d\theta d\eta dz \\
+ \frac{2\sqrt{2}}{\sqrt{\pi}} e^{-\left(\frac{l^2}{4} + \frac{\tau}{2}\right) (\tau+l)} \\
\int_{-\infty}^\frac{x}{\sqrt{\eta^2}} e^{-\frac{2^2}{2} \frac{(x-z)^2}{2\eta^2}} \int_{-\infty}^\frac{y}{\sqrt{2\eta^2}} f_4(y, \tau - \frac{(x - z)^2}{2\eta^2}, z) dy d\eta dz
\]

(11.3.39)

where

\[
\begin{align*}
 f_3(\theta, w, z) &= e^{\frac{1}{2} \frac{(w+1)^2}{2l}(\theta \cos \theta)} - \left(\frac{k}{2} + 1\right)^2 N(d_1) \sqrt{\frac{1}{l + w}} \cos \theta - \left(\frac{k}{2} + 1\right) \frac{e^{-\frac{d_1^2}{4\sqrt{\pi}}}}{2\sqrt{l}} \\
 &\quad + e^{\frac{k^2}{4} (l+w) \cos^2 \theta} \left[ \frac{k^2}{4} N(d_2) \sqrt{\frac{1}{l + w}} \cos \theta + \frac{e^{-\frac{d_2^2}{4\sqrt{\pi}}}}{4\sqrt{\pi}} \right], \\
 f_4(y, w, z) &= \frac{z}{4\sqrt{\pi}(l+w)} \left[ \frac{e^{\frac{1}{2} \frac{(w+1)^2}{2l}(\theta \cos \theta)}}{\sqrt{2l}(l+w)} \frac{1}{(y^2 + 1)} - \frac{d_3^2}{2(l+w)} \right],
\end{align*}
\]

\[
\begin{align*}
 d_1 &= \frac{z}{\sqrt{2(l+w)\cos \theta}} + \left(\frac{k}{2} + 1\right) \sqrt{\frac{2(l+w)}{\cos \theta}}, \\
 d_2 &= \frac{z}{\sqrt{2(l+w)\cos \theta}} + \frac{k}{2} \sqrt{\frac{2(l+w)}{\cos \theta}}, \\
 d_3 &= \frac{z^2(1+y^2)}{2(l+w)} + z(k + 2) + \frac{2(l+w)}{(y^2 + 1)} \left(\frac{k}{2} + 1\right)^2, \\
 d_4 &= \frac{z^2(1+y^2)}{2(l+w)} + k z + \frac{(l+w)k^2}{2(y^2 + 1)},
\end{align*}
\]

with $N(\cdot)$ being the standard cumulative distribution function. While the details of
the double Laplace inversion is left in Appendix E.3 for interested readers, it should be remarked that although the current expression (11.3.39) involves the evaluation of triple integrals, we believe that it is already in its simplest form, due to the complexity arising from the non-resetting mechanism associated with the ParAsian options. Also, as a result of a simple financial switch from a resetting “trigger” to a non-resetting one, the nice recursive nature in the evaluation of (11.3.31) for the case of Parisian options has been destroyed; the evaluation of $W(\tau, J)$ from (11.3.39) requires an additional fold of integration. It is thus anticipated that the computational efficiency of adopting (11.3.39) to calculate numerical values of a ParAsian option should be slightly worse off than that of adopting (11.3.31) to calculate numerical values of a Parisian option. Some direct comparisons of numerical efficiency are provided in the next section through some numerical examples.

11.4 Numerical examples and discussions

The derivations of (11.3.31) from (11.3.22) and (11.3.39) from (11.3.32) are carried out rigorously and deductively, and thus there is no need to discuss the “accuracy” of these closed-form solutions and present any calculated results. However, from the viewpoint that a comparison with previously published results (numerical solutions or solutions obtained through other approximation methods) may give readers a sense of verification of the newly found closed-form solutions, as well as help them understand the improvement in accuracy and efficiency with our exact formulae, several numerical examples are given in this section. The section is organized into two subsections, according to two important issues that should be addressed. The first subsection is to compare the results obtained from our newly proposed PDE systems (11.2.11) and (11.2.13) with the corresponding ones proposed in [49], while the second subsection is to test the numerical performance of our closed-form analytic solutions.

To help readers who may not be used to discussing financial problems with di-
mensionless quantities, all results, unless otherwise stated, are now converted back to dimensional quantities in this section before they are graphed and presented.

11.4.1 A comparison with Wilmott et al.’s numerical solutions

As pointed out earlier, our pricing systems (11.2.11) and (11.2.13) are primarily based on what were proposed by Wilmott et al. [49]. However, we found that there are some errors in [49]. It is thus interesting to have a direct comparison between the results produced with our new PDE systems and those produced with Wilmott et al.’s original PDE system. Of course, in addition to the financial arguments provided in Section 2, here in this subsection, we share focus on the results produced by our newly-derived analytic solution and those produced by Wilmott et al.’s explicit finite difference scheme.

There are two major differences between Wilmott et al.’s PDE system and ours presented in this chapter. The first one is that our solution has been worked out from a truncated domain; apart from the results near the boundary where different boundary conditions are imposed, we expect that the results produced with our newly-derived analytic solutions should match those produced by Wilmott et al.’s explicit finite difference scheme. This is indeed so, as shown in Fig 11.2-11.3. In Fig 11.2(a-b), the Parisian and Par asian option prices with three different values of the time left in the trigger ($\bar{J} - J$) being less than or equal to the time to expiry, i.e., $T - t \geq \bar{J} - J$, are plotted, respectively. From these two figures, one can clearly observe that for the case $T - t \geq \bar{J} - J$, both Parisian and Par asian option prices calculated from our systems and those produced from Wilmott et al.’s systems [49] agree amazingly well with each other.

The second difference is that different boundary conditions at $S = \infty$ are proposed in the prism $\mathcal{IV}$, where $J$ is impossible to reach $\bar{J}$. In this case, we should expect that the results produced by the two different PDE systems are different. A comparison of option prices for this case is shown in Figs 11.3(a-b) and Figs 11.3(c-d)
for the Parisian and ParAsian options, respectively. It can be observed that for reasonably finite values of the underlying, our solutions seem to agree well with Wilmott et al.’s, whether the inequity $T - t \geq \bar{J} - J$ is held or not, as shown in Fig 11.3(a) and Fig 11.3(b). In these two figures, one may also notice that the option prices corresponding to the case $T - t \geq \bar{J} - J$ appear to have a sharp corner around $\bar{S}$, which seems to be contradictory to the financial clause. In fact, this is only a result of the scale with which these figures were plotted; a locally “zoom-in” plot would reveal that the “sharp” corner becomes a smooth curve as a result of imposing the continuity of the Delta between the two regions. On the other hand, since different boundary conditions at $S = \infty$ are imposed in the prism $\mathcal{IV}$, one can expect that for the case $T - t < \bar{J} - J$, the large $S$ asymptotic behaviors of our solutions and Wilmott et al.’s should be different, although they agree very well with each other for reasonably finite underlying (see the option prices provided in Fig 11.3(a) and Fig 11.3(c)). To visualize the price differences at the large end of $S$ values, we have re-plotted two sets of option price data with $J = 0$ displayed in Fig 11.3(a) and Fig 11.3(c), with a larger range ($S = 50$ instead of $S = 15$) in Fig 11.3(b) and Fig 11.3(d), respectively. Clearly, at the large $S$ end, the option prices calculated from our systems increase to infinity at almost the same rate as the underlying, whereas Wilmott et al.’s option prices exhibit a dramatic decrease to zero with sufficiently large $S$ values, as shown in Fig 11.3(b) and Fig 11.3(d) for Parisian and ParAsian options, respectively. This is not surprising, as in the region where $T - t < \bar{J} - J$ (the prism $\mathcal{IV}$), the boundary condition at $S = \infty$ in our systems is set to $S$ to ensure that the option price in this domain equals a European call option price, whereas in Wilmott et al.’s systems, it is set to zero no matter what value $\tau$ is. It should be remarked that for $T - t < \bar{J} - J$, the resulting option prices differing only for sufficiently large underlying is indeed reasonable, since the boundary condition usually only has a local effect on the final solution.

Several remarks should be made before we discuss the accuracy and efficiency of the explicit-closed form analytic formulae, in terms of numerically evaluating the
(a) Price of the Parisian up-and-out call options.

(b) Price of the Parisian up-and-out call options at finite S values.

Figure 11.2: Wilmott et al.’s solutions VS our results for the case $T - t \geq \bar{J} - J$. Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$, $K = $10, $\bar{S} = $12, $\bar{J} = 0.2$ (year) and $T - t = 1$ (year).

(c) Price of the Parisian up-and-out call options at large S end.

(d) Price of the Parisian up-and-out call options at large S end.

Figure 11.3: Wilmott et al.’s solutions VS our results for the case $T - t < \bar{J} - J$. Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$, $K = $10, $\bar{S} = $12, $\bar{J} = 0.2$ (year) and $T - t = 0.1$ (year).
involved integrals, as is the main issue of the next subsection. Firstly, whether the prism \( \mathcal{P} \) (or \( \mathcal{P} \)) is cut or not has no influence on the final solution for the Parisian (Parasian) options, due to the backward property of the current problem. In fact, under our pricing systems, the prism \( \mathcal{P} \) (or \( \mathcal{P} \)) can be artificially filled in once a Parisian (Parasian) option with longer maturity is considered. In other words, in this region, our option prices should be identical to Wilmott et al.’s solutions.

Secondly, one may wonder why there is an extra condition (11.2.10) (or (11.2.14)) in our systems (11.2.11) (or (11.2.13)) whereas there is no such condition in Wilmott et al.’s PDE systems governing the price of a Parisian option (or Parasian option), yet the results from both systems agree well for reasonably small \( S \). The reason is that when applying the explicit finite difference scheme, both of the connectivity conditions involved in the pricing systems (11.2.11) and (11.2.13) become redundant, since the option prices across the barrier are implicitly assumed to satisfy the governing equation, which is a much stronger condition than the continuity of the option Delta across the barrier. In fact, handling with the connectivity conditions in this way will not result in any difference on the final results, as clearly demonstrated in Section 2.1.

\section{Numerical performance of our analytical solutions}

To test the accuracy and efficiency of our exact solutions, in terms of numerically calculating the integrals, the best way is to compare the analytical solutions with those calculated directly from Wilmott et al.’s pricing systems with the utilization of their explicit finite difference method [49]. Such comparisons are shown in Table 11.1 and Table 11.2, in which the Parisian and Parasian up-and-out call option prices across the barrier are tabulated, respectively. Furthermore, we have considered four different times to maturity, i.e., \( T-t = 0.3 \) (year), \( T-t = 0.4 \) (year), \( T-t = 0.5 \) (year), and \( T-t = 1 \) (year). All the experiments were performed within Matlab7.8 on an Intel Pentium 4, 3GHZ machine. In these tables, columns marked with “FDM” display the results obtained from Wilmott et al.’s explicit finite difference scheme with extremely
fine grids defined as $\Delta S = 0.05$, $\Delta \tau = \Delta J = 0.001$, and the CPU time means the
total time it takes to numerically evaluate our analytic formulae.

From Table 11.1, it is clear that our analytical results for the Parisian up-and-out
call options agree well with those produced by using the explicit finite difference
method, with the point-wise difference between the two being in the order of $O(10^{-3})$.
Moreover, it only takes a few seconds to numerically carry out our exact solution,
which is hundreds of folds less than what it takes to execute the code written with
the explicit finite difference method for the same case, especially when options are
of long maturity.

Provided in Table 11.2 are the corresponding results for the Parasian up-and-out
call options. From this table, it is clear that our analytic results also agree with
those calculated by using the explicit finite difference method, with the point-wise
difference now being in the order of $O(10^{-2})$. The loss of accuracy of the explicit
finite difference method adopted to price the Parasian options is indeed expected,
as in this case, the error along the $J$ direction starts to accumulate when solving for
both $V_1$ and $V_2$, whereas for a corresponding Parisian option, the error along the $J$
direction will not affect the accuracy of $V_1$. In this sense, the explicit finite difference
scheme used to price the Parasian options should be less accurate than that with the
same grid size adopted to price the corresponding Parisian options. One may also
notice from Table 11.2 that the point-wise relative errors between our exact solution
and the numerical result are larger as the time to maturity becomes longer or the
barrier time $J$ is closer to the trigger value $\bar{J}$. This is not surprising at all, as in
both cases mentioned above, the absolute value of $W(t, J)$ becomes very small, and
discussing the relative error between two numbers which are quite small in absolute
terms means very little. On the other hand, we should point out that although our
analytic solution for $W(t, J)$ involves the evaluation of triple integrals, it is still far
more efficient to implement than the explicit finite difference method, with the CPU
times for the former being at least 20 folds less than those for the latter, as shown in
Table 11.2. Therefore, from both efficiency and accuracy, points of view, our analytic
formulae have clear edge over the explicit finite difference method in pricing Parisian and ParAsian options.

Table 11.1: The price of the Parisian up-and-out call options across the barrier, i.e., $W(t)$, with four different times to maturity. Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$, $K = $10, $S = $12, $J = 0.2\text{(year)}$.

<table>
<thead>
<tr>
<th>The time to expiry</th>
<th>FDM (CPU(second))</th>
<th>exact solution (CPU(second))</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T - t = 0.3$ (year)</td>
<td>1.3660 (375s)</td>
<td>1.3883 (1s)</td>
</tr>
<tr>
<td>$T - t = 0.4$ (year)</td>
<td>1.0819 (503s)</td>
<td>1.0934 (1s)</td>
</tr>
<tr>
<td>$T - t = 0.5$ (year)</td>
<td>0.8887 (631s)</td>
<td>0.8888 (2s)</td>
</tr>
<tr>
<td>$T - t = 1.0$ (year)</td>
<td>0.4246 (1260s)</td>
<td>0.4296 (6s)</td>
</tr>
</tbody>
</table>

As mentioned earlier, once the values of $W$ are known, the option prices can be obtained straightforwardly from the integral representations (11.3.23)-(11.3.24) and (11.3.33)-(11.3.34), for the Parisian and ParAsian up-and-out call options, respectively. We can thus further test the accuracy of our analytic formulae by calculating the option prices via (11.3.23)-(11.3.24) and (11.3.33)-(11.3.34), and comparing them with those produced by using Wilmott et al.’s method, as shown in Fig 11.4 and Fig 11.5, where the option prices are displayed as a function of the underlying $S$, for different $J$ values. The excellent agreement of our analytic option prices and those calculated by the explicit finite difference scheme again provides a sense of verification of our closed-form explicit formulae.

Table 11.2: The price of the ParAsian up-and-out call options across the barrier, i.e., $W(t, J)$, with four different times to maturity. Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$, $K = $10, $S = $12, $J = 0.2\text{(year)}$.

<table>
<thead>
<tr>
<th>The time to expiry</th>
<th>FDM (CPU(second))</th>
<th>exact solution (CPU(second))</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T - t = 0.3$ (year)</td>
<td>0.9974 (439s)</td>
<td>0.9994 (23s)</td>
</tr>
<tr>
<td>$T - t = 0.4$ (year)</td>
<td>0.7055 (587s)</td>
<td>0.7312 (23s)</td>
</tr>
<tr>
<td>$T - t = 0.5$ (year)</td>
<td>0.5433 (796s)</td>
<td>0.5762 (23s)</td>
</tr>
<tr>
<td>$T - t = 1.0$ (year)</td>
<td>0.2297 (1476s)</td>
<td>0.2462 (23s)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The time to expiry</th>
<th>FDM (CPU(second))</th>
<th>exact solution (CPU(second))</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T - t = 0.3$ (year)</td>
<td>0.5427 (439s)</td>
<td>0.5339 (23s)</td>
</tr>
<tr>
<td>$T - t = 0.4$ (year)</td>
<td>0.4043 (587s)</td>
<td>0.4050 (23s)</td>
</tr>
<tr>
<td>$T - t = 0.5$ (year)</td>
<td>0.3184 (763s)</td>
<td>0.3222 (23s)</td>
</tr>
<tr>
<td>$T - t = 1.0$ (year)</td>
<td>0.1388 (1476s)</td>
<td>0.1400 (23s)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>The time to expiry</th>
<th>FDM (CPU(second))</th>
<th>exact solution (CPU(second))</th>
</tr>
</thead>
<tbody>
<tr>
<td>$T - t = 0.3$ (year)</td>
<td>0.3446 (439s)</td>
<td>0.3230 (23s)</td>
</tr>
<tr>
<td>$T - t = 0.4$ (year)</td>
<td>0.2599 (587s)</td>
<td>0.2463 (23s)</td>
</tr>
<tr>
<td>$T - t = 0.5$ (year)</td>
<td>0.2059 (763s)</td>
<td>0.1963 (23s)</td>
</tr>
<tr>
<td>$T - t = 1.0$ (year)</td>
<td>0.0904 (1476s)</td>
<td>0.0852 (23s)</td>
</tr>
</tbody>
</table>
Figure 11.4: The price of the Parisian up-and-out call option at different $J$ values. Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$, $K = $10, $\bar{S} = $12, $J = 0.2$ (year) and $T - t = 1$ (year)

Figure 11.5: The price of the Parisian up-and-out call option at different $J$ values. Parameters are $\sigma = 10\%$, $r = 5\%$, $D = 0$, $K = $10, $\bar{S} = $12, $J = 0.2$ (year) and $T - t = 1$ (year)
On the other hand, Fig 11.4 and Fig 11.5 also clearly exhibit the up-and-out feature of the Parisian-type options. It can be observed that the prices of both Parisian and Parasian up-and-out call options have peaked before the barrier $\bar{S}$ is reached. Then, as $S$ is further increased, the danger of being “out” has already been factorized in the price, which starts to decrease from the peak of the price and eventually tends to zero when the underlying price gets too high. These results make sense financially as once the underlying price gets closer to the barrier, market at some point will start to consider the danger of being “knocked out”. One should also notice that because of different resetting structures, the option price of a Parisian option is not affected at all by the residue left in the trigger (because the barrier time will be reset to zero once $S$ falls back to $\bar{S}$), whereas the changes of the Parasian option price with the residue left in the trigger are “felt” in the entire domain of the underlying, as a result of the non-resetting mechanism on the trigger. It can be further observed from these two figures that when the underlying price is near $\bar{S}$, the Parisian specification provides a smooth transition only at $J = 0$, and moreover, its option Delta increases dramatically to infinity as the barrier time $J$ becomes closer to the trigger value $\bar{J}$. This is because as long as the trigger value $\bar{J}$ is not reached, $J$ will be reset to zero when the barrier is touched, and furthermore, the option Delta being continuous is only specified at $J = 0$ for the Parisian option (see the connectivity condition (11.2.10)). In contrast, the Parasian specification has made the option prices smooth across the barrier $\bar{S}$ for all non-zero $J$ values. This is indeed a result of the option Delta of a Parasian option being assumed to be continuous across the whole barrier plane (see the connectivity condition (11.2.14)). Note that “smooth” here refers to the option Delta to be continuous.

Depicted in Fig 11.6 is the comparison of a Parisian up-and-out call option and the corresponding Parasian option. Clearly, at any given $J$, the Parisian option price is always higher than its Parasian counterpart for all underlying $S$. This is indeed reasonable, as the “out” feature of the Parasian option has been amplified by the cumulative effect on $J$, and thus a Parasian option is more vulnerable to be knocked
out than its Parisian counterpart. Consequently, its value should be lower than the corresponding Parisian option price, if all the other terms are the same.

11.5 Conclusion

In this chapter, substantial progress has been made for pricing Parisian-type options, with the successful derivation of two closed-form analytic formulae for European-style Parisian and Parasian options, respectively. A key step of two approaches presented here is to simplify the domain of the pricing systems, and consequently reduce a 3-D problem to two coupled 2-D systems. Through numerical examples, we have shown that our newly established pricing systems can indeed correctly replace the original 3-D ones that have been used in the literature. We have also compared the results produced from our analytical solutions with those calculated by using Wilmott et al.’s method, and found that our results agree with Wilmott et al.’s perfectly for reasonably finite values of the underlying.

The significance of our work has been illustrated from two aspects. Theoreti-
cally, although the literature provides quite a few approximate methods for pricing Parisian-type options, the closed-form solutions are presented for the first time. Practically, computational efficiency has been enormously enhanced through using the newly derived analytical formulae, which can assist market practitioners to price Parisian-type options much more quickly and accurately.
Chapter 12

Conclusions

In this thesis, we proposed some highly efficient approaches to price option derivatives under as well as beyond the Black-Scholes model. In particular, three key quantitative methods in the option pricing field, namely, the numerical schemes, the analytic approximation, and the exact solution approach, have been investigated and further developed.

The first issue discussed is the numerical valuation of American options with and without volatility. We propose a new predictor-corrector scheme based on the ADI method to solve the price of American puts under the Heston model. The novelty of this scheme, in comparison with some other finite difference methods in the literature, lies in the fact that it requires no embedded iterations, and can capture the whole optimal exercise boundary as part of the solution procedure. Based on the local von Neumann analysis, combining with the “frozen” coefficient technique, a conditional stability requirement is also obtained for the predictor-corrector method. By realizing that either the predictor-corrector scheme or the documented numerical schemes in the literature fail to solve the prices of perpetual American options efficiently and accurately, we then propose a spectral-collocation method for the numerical valuation of perpetual American puts when Heston’s stochastic volatility model is used. The proposed method can be easily and efficiently implemented, and moreover, the spectral accuracy can be easily achieved with a small number of iterations. On the
other hand, in terms of pricing American options under the classical Black-Scholes model, we have proposed and tested an inverse finite element approach, which does not involve any linearization process at all, but still with high efficiency as a result of using the full Newton iteration method with its inherent quadratic convergence. The key contribution of this work to the literature is that a convergence analysis for the inverse finite element approach is provided for the first time, which ensures that the numerical solution does indeed converge to the exact one of the original nonlinear system.

The derivations of various analytic approximation formulae form the second major issue of the current thesis, which is further divided into two sub-issues, i.e., the asymptotic behavior of the optimal exercise price near expiry and pricing approximation formulae for vanilla options. On one hand, we first investigate the asymptotic behavior of the optimal exercise price for an American put option in the local volatility model. The result derived is believed to be quite reasonable, since the leading order term of the optimal exercise price in the stock-price-dependent volatility model agrees with that in the literature, and moreover, it degenerates to Evans et al.'s result if the volatility function is assumed to be a constant. Then, our approach is extended to analyze the near-expiry behavior of the optimal exercise price for American put options on a dividend-paying underlying with stochastic volatility. It turns out that the option prices are quite different from the corresponding Black-Scholes case, while the leading order term of the optimal exercise price remains almost same as the constant volatility case if the spot volatility is given the same value as the constant volatility appearing in the Black-Scholes model. Numerical experiments suggest that the approximation is reasonably accurate for options with tenor in the order of days and weeks. On the other hand, we have derived a series of analytical approximation formulae for the pricing of vanilla options, which are particularly useful in today's algorithmic trading and VaR (Value at Risk)-based risk management. In specific, we have derived a new approximation for the valuation of European put options under the Heston model, which is as easy to implement as the Black-Scholes formula!
have also derived the pricing approximations for perpetual American options under
different volatility dynamics, such as slowly varying volatility, fast mean-reverting
volatility, and even multiscale volatility. Moreover, based on the newly obtained
formulae, the quantitative effect of different stochastic volatility dynamics on the
optimal exercise strategies of perpetual American puts have also been analyzed.

Finally, we discuss the exact solution approach for the pricing of some exotic-options. We present exact and analytic solutions for the valuation of Parisian and Parasion options under the Black-Scholes framework. Through a coordinate transform that has elegantly “absorbed” the directional derivative associated with the
“barrier trigger” into the time derivative, we obtain two coupled, but simplified PDE
systems. For Parisian options, the coupled PDE systems are then analytically solved
by applying the Laplace transform technique in conjunction with the construction of
“moving windows”, which are introduced to evaluate the option prices backwards,
slide by slide, until the value of the option at a given time and trigger value is found
for a given underlying price. On the other hand, due to the non-resetting mechanism
of the Parasion option, the resulted coupled PDE systems of this option are much
more complicated than those of its Parisian counterpart, and the “moving window”
technique fails in this case. Alternatively, the double Laplace transform technique is
applied to solve for the option prices in the Laplace space. Finally, by performing
the analytical Laplace inversions, the completely analytical closed-form solutions for
the Parisian and Parasion options can both be obtained. Various numerical experi-
ments show that our solutions are quite accurate, and moreover, both of them can
be efficiently carried out, although quite a few integrals are involved.

Before closing the conclusion part, we remark that the extensions of all the ap-
proaches presented above to some even more complicated stochastic volatility models
are indeed promising, and could thus be very interesting topics for the future research.
Appendix A

The proof and the implementation details of Chapter 2

A.1 The consistency between the boundary condition at \( v = 0 \) and the American put option pricing system with stochastic volatility

In this Appendix, we consider the consistency between (2.2.8) and a newly constructed PDE system, which is the same as the PDE system (2.2.16), but without the boundary constrain at \( v = 0 \). As demonstrated in Section 2.2, this new PDE system is already closed when \( \kappa \eta \geq \frac{\sigma^2}{2} \). Note that in the following theorem, “the PDE system” always refers to this newly-constructed PDE system, unless stated elsewhere.

**Theorem 1** For \( \kappa \eta \geq \frac{\sigma^2}{2} \), the proposed boundary condition (2.2.8) for \( v = 0 \), is consistent with the inherent value of the PDE system as \( v \to 0 \).

**Proof.** From a perturbation point of view, this statement is equivalent to showing that (2.2.8) will not create any singularity or boundary layer as \( v \to 0 \). Mathematically, we need to prove that the leading order term as \( v \to 0 \) deduced from the PDE system is exactly the same as (2.2.8).
The asymptotic behavior of the PDE system can be examined with the perturbation technique. First, we rewrite $v$ as $v = \epsilon V$, with $\epsilon$ being a small parameter and $V$ being an $O(1)$ variable, to reflect the fact that $v$ is an extremely small number in the limit process of $v$ approaching zero (cf. [33]). Now, substituting $v = \epsilon V$ into the governing equation of the PDE system, and assuming that its solution can be written in powers of $\epsilon$, i.e., $U(S,V,t) = \sum_{n=0}^{\infty} U_n(S,V,t)\epsilon^n$, we obtain the leading order term as

$$U_0(S,V,t) = C_1(S,t) + C_2(S,t)\frac{V^{1-\alpha}}{\alpha}$$

as $v \to 0$, where $\alpha = \frac{2\kappa \eta}{\sigma^2}$. Clearly, when $\alpha > 1$, i.e., $\kappa \eta \geq \frac{\sigma^2}{2}$, $U_0$ has an exponential growth as $V \to 0$, unless $C_2$ is set to zero. With $C_2$ being set to zero to maintain the boundedness of the American put value as the solution of the PDE system, $U_0$ is now a function of $S$ and $t$ only. Since $U_0$ also needs to satisfy the initial and boundary conditions along the $t$ and $S$ directions, respectively, one has no choice but letting $C_1(S,t) = \max(K - S, 0)$. Using (2.2.7), the boundary condition (2.2.8) can be easily deduced. Consequently, the consistency between (2.2.8) and the inherent value of the PDE system as $v \to 0$ for the case of $\kappa \eta \geq \frac{\sigma^2}{2}$ is established.

The proof of this theorem shows that for $\kappa \eta \geq \frac{\sigma^2}{2}$, the proposed boundary condition at $v = 0$, although mathematically “redundant”, does indeed not “deteriorate” the well-posedness of the pricing problem.

A.2 The discretization of the ADI method

Once the space discretization is performed, the governing equation in the PDE system (2.3.1) becomes:

$$\frac{\partial U_{i,j}^n}{\partial \tau} = a_j \delta_{xx} U_{i,j}^n + b_j \delta_{vv} U_{i,j}^n + (d_j + \lambda_j) \delta_x U_{i,j}^n + c_j \delta_{xx} U_{i,j}^n + e_j \delta_v U_{i,j}^n - r U_{i,j}^n.$$
By applying the first-order fully implicit Euler scheme to the time derivative $\frac{\partial U_{i,j}^n}{\partial \tau}$, we can obtain

$$
\frac{U_{i,j}^{n+1} - U_{i,j}^n}{\Delta \tau} = a_j \delta_{xx} U_{i,j}^{n+1} + b_j \delta_{vv} U_{i,j}^{n+1} + (d_j + \lambda_j) \delta_x U_{i,j}^{n+1} + c_j \delta_{xx} U_{i,j}^{n+1} + e_j \delta_x U_{i,j}^{n+1} - r U_{i,j}^{n+1}.
$$

In order to demonstrate the ADI method clearly, we use linear operators $A_0, A_1$ and $A_2$ to denote the mixed derivative, the spatial derivatives in the $x$ direction and the spatial derivatives in the $v$ direction, respectively, i.e.,

$$
A_0 U_{i,j}^n = \Delta \tau \cdot c_j \delta_{xx} U_{i,j}^n,
$$

$$
A_1 U_{i,j}^n = \Delta \tau \cdot (a_j \delta_{xx} U_{i,j}^n + (d_j + \lambda_j) \delta_x U_{i,j}^n - \frac{r}{2} U_{i,j}^n),
$$

$$
A_2 U_{i,j}^n = \Delta \tau \cdot (b_j \delta_{vv} U_{i,j}^n + e_j \delta_v U_{i,j}^n - \frac{r}{2} U_{i,j}^n).
$$

Thus, the governing equation in the PDE system (2.3.1) can be written shortly as:

$$
[I - (A_0 + A_1 + A_2)] U^{n+1} = U^n + \mathcal{O}((\Delta \tau)^2).
$$

Similarly, the first-order explicit Euler scheme reads:

$$
[I + (A_0 + A_1 + A_2)] U^n = U^{n+1} + \mathcal{O}((\Delta \tau)^2).
$$

Thus, the weighted average of the fully implicit scheme and explicit scheme can be written as

$$
[I - \theta(A_0 + A_1 + A_2)] U^{n+1} = [I + (1 - \theta)(A_0 + A_1 + A_2)] U^n + \mathcal{O}((\Delta \tau)^3). \quad (A.2.1)
$$

One should note that when $\theta$ is equal to zero or one, (A.2.1) goes back to fully
explicit Euler scheme or fully implicit one. When \( \theta = \frac{1}{2} \), it is equivalent to applying the Crank-Nicolson scheme to the time derivative \( \frac{\partial U^n}{\partial \tau} \).

Adding \( \theta^2 A_1 A_2 U^{n+1} \) to both sides of (A.2.1), we have:

\[
[I - \theta A_0 - \theta A_1 - \theta A_2 + \theta^2 A_1 A_2] U^{n+1} = [I + (1 - \theta)A_0 + (1 - \theta)A_1 \\
+ (1 - \theta)A_2 + \theta^2 A_1 A_2] U^n \\
+ \theta^2 A_1 A_2 (U^{n+1} - U^n) + \mathcal{O}((\Delta \tau)^3).
\]

(A.2.2)

As \( U^{n+1} - U^n \sim \mathcal{O}(\Delta \tau) \), and \( A_1 A_2 \sim \mathcal{O}((\Delta \tau)^2) \), we can merge the term \( \theta^2 A_1 A_2 (U^{n+1} - U^n) \) into the error term. Thus, (A.2.2) becomes:

\[
(I - \theta A_1)(I - \theta A_2) U^{n+1} - \theta A_0 U^{n+1} = [I + (1 - \theta)A_0 + (1 - \theta)A_1 + \\
(1 - \theta)A_2 + \theta^2 A_1 A_2] U^n + \mathcal{O}((\Delta \tau)^3).
\]

(A.2.3)

However, it is still difficult to solve (A.2.3) alternatively in two directions because of the existence of \( \theta A_0 U^{n+1} \). One possible measure is to treat the mixed derivative fully explicit at the expense of losing one order of accuracy in the time direction, i.e.,

\[
(I - \theta A_1)(I - \theta A_2) U^{n+1} = [I + A_0 + (1 - \theta)A_1 + (1 - \theta)A_2 + \theta^2 A_1 A_2] U^n \\
+ \theta A_0 (U^{n+1} - U^n) + \mathcal{O}((\Delta \tau)^3),
\]

\[
= [I + A_0 + (1 - \theta)A_1 + (1 - \theta)A_2 \\
+ \theta^2 A_1 A_2] U^n + \mathcal{O}((\Delta \tau)^2),
\]

which is equivalent to:

\[
(I - \theta A_1)(I - \theta A_2) U^{n+1} = [I + A_0 + (1 - \theta)A_1 + A_2] U^n - (I - \theta A_1) \theta A_2 U^n. \quad \text{(A.2.4)}
\]
Here, the linear operators $A_0$ and $A_1$ in (A.2.4) are time-dependent. We value them at the $(n + \theta)$th time step, i.e.,

$$A_0 = \Delta \tau c_j \delta_{xy},$$
$$A_1 = \Delta \tau [a_j \delta_{xx} + (d_j + \lambda_j) \delta_x - \frac{r}{2} U_{i,j}],$$
where

$$\xi = \delta_{yy}(\theta S^n_{f} + (1 - \theta) S^n_{f}),$$
$$\beta = \delta_{yv}(\theta S^n_{f} + (1 - \theta) S^n_{f}),$$
$$\lambda = \frac{S^n_{f} - S^n_{f}}{(\theta S^n_{f} + (1 - \theta) S^n_{f}) \Delta \tau}.$$

**A.3 The matrix equation after the ADI method is applied**

The matrix form for calculating $Y$ corresponding to (2.3.12) is:

$$AY_j = P_j + \text{bnd} x_j,$$
with

$$Y_j = (Y_{1,j}, Y_{2,j}, \cdots, Y_{N_x-1,j})^T,$$
$$P_j = (p_{1,j}, p_{2,j}, \cdots, p_{N_x-1,j})^T,$$
\[ \text{bnd}x_j = \begin{pmatrix} 
\theta \left( \frac{a_j \Delta \tau}{(\Delta x)^2} - \frac{d'_j \Delta \tau}{2 \Delta x} \right) Y_{0,j} \\
\vdots \\
\theta \left( \frac{a_j \Delta \tau}{(\Delta x)^2} + \frac{d'_j \Delta \tau}{2 \Delta x} \right) Y_{N_x,j} 
\end{pmatrix}, \]

\[ \begin{align*}
p_{ij} &= U^n_{i,j} + \frac{c_j \Delta \tau}{4 \Delta x \Delta v} (U^n_{i+1,j+1} + U^n_{i-1,j-1} - U^n_{i+1,j-1} - U^n_{i-1,j+1}) \\
&\quad + (1 - \theta) \left( \frac{a_j \Delta \tau}{(\Delta x)^2} - \frac{d'_j \Delta \tau}{2 \Delta x} \right) U^n_{i-1,j} + (1 - \theta) \left( -\frac{a_j \Delta \tau}{(\Delta x)^2} - \frac{r \Delta \tau}{2} \right) U^n_{i,j} \\
&\quad + (1 - \theta) \left( \frac{a_j \Delta \tau}{(\Delta x)^2} + \frac{d'_j \Delta \tau}{2 \Delta x} \right) U^n_{i+1,j} + \left( \frac{b_j \Delta \tau}{(\Delta v)^2} - \frac{e_j \Delta \tau}{2 \Delta v} \right) U^n_{i,j-1} \\
&\quad + \left( -\frac{b_j \Delta \tau}{(\Delta v)^2} - \frac{r \Delta \tau}{2} \right) U^n_{i,j} + \left( \frac{b_j \Delta \tau}{(\Delta v)^2} + \frac{e_j \Delta \tau}{2 \Delta v} \right) U^n_{i,j+1},
\end{align*} \]

\[ A = \begin{pmatrix} 
1 + \theta \left( \frac{2 a_j \Delta \tau}{\Delta x^2} + \frac{r \Delta \tau}{2} \right) & -\theta \left( \frac{a_j \Delta \tau}{\Delta x^2} + \frac{d'_j \Delta \tau}{2 \Delta x} \right) & \vdots \\
-\theta \left( \frac{a_j \Delta \tau}{\Delta x^2} - \frac{d'_j \Delta \tau}{2 \Delta x} \right) & 1 + \theta \left( \frac{2 a_j \Delta \tau}{\Delta x^2} + \frac{r \Delta \tau}{2} \right) & \vdots \\
\vdots & \vdots & \ddots \\
\vdots & \vdots & \ddots & 1 + \theta \left( \frac{2 a_j \Delta \tau}{\Delta x^2} + \frac{r \Delta \tau}{2} \right)
\end{pmatrix}, \]

where \( d'_j = (d_j + \lambda) \), and \( \text{bnd}x_j \) stands for the vector that contains the boundary value of the intermediate variable \( Y \) in the \( x \) direction.

The corresponding matrix form for computing \( U^{n+1} \) from (2.3.13) is:

\[ BU^{n+1}_i = Q_i + \text{bnd}v_i, \]

with

\[ U^{n+1}_i = (U^{n+1}_{i,1}, U^{n+1}_{i,2}, \ldots U^{n+1}_{i,N_x-1})^T, \]

\[ Q_i = (q_{i,1}, q_{i,2}, \ldots q_{i,N_x-1})^T, \]
\[
bnd v_j = \begin{pmatrix}
\theta \left( \frac{b_j \Delta \tau}{\Delta v^2} - \frac{e_j \Delta \tau}{2 \Delta v} \right) U_{i,0}^{n+1} \\
\vdots \\
\theta \left( \frac{b_j \Delta \tau}{\Delta v^2} + \frac{e_j \Delta \tau}{2 \Delta v} \right) U_{i,N_v}^{n+1}
\end{pmatrix},
\]

\[
q_{i,j} = Y_{i,j} - \theta \left( \frac{b_j \Delta \tau}{\Delta v^2} - \frac{e_j \Delta \tau}{2 \Delta v} \right) U_{i,j-1}^n - \theta \left( -\frac{b_j \Delta \tau}{\Delta v^2} - \frac{r \Delta \tau}{2} \right) U_{i,j}^n - \theta \left( \frac{b_j \Delta \tau}{\Delta v^2} + \frac{e_j \Delta \tau}{2 \Delta v} \right) U_{i,j+1}^n,
\]

\[
B = \begin{pmatrix}
1 + \theta \left( \frac{2b_j \Delta \tau}{\Delta v^2} + \frac{r \Delta \tau}{2} \right) & -\theta \left( \frac{b_j \Delta \tau}{\Delta v^2} + \frac{e_j \Delta \tau}{2 \Delta v} \right) \\
-\theta \left( \frac{b_j \Delta \tau}{\Delta v^2} - \frac{e_j \Delta \tau}{2 \Delta v} \right) & 1 + \theta \left( \frac{2b_j \Delta \tau}{(\Delta v)^2} + \frac{r \Delta \tau}{2} \right) \\
& \ddots \\
& & \ddots \\
& & & 1 + \theta \left( \frac{2b_j \Delta \tau}{\Delta v^2} + \frac{r \Delta \tau}{2} \right)
\end{pmatrix},
\]

where \( \text{bnd} v_j \) stands for the vector that includes the boundary value of \( U^{n+1} \) in the \( v \) direction.
Appendix B

Realization details of Chapter 4

B.1 Finite element formulation for the PDE system (4.3.5)

Suppose that the computational domain \([0, \exp(x_{\text{max}}) - 1]\) is now discretized into \(N\) 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

\[
U = \sum_{i=1}^{p} \varphi_i(\xi) U_i = (\varphi_1, \cdots, \varphi_p) U^{(n)},
\]

where \(U^{(n)}\) is the vector of the nodal values associated with the \(n\)th element, i.e., \(U^{(n)} = (U_1 \cdots U_p)'\) with the subscripts being the local numbers, and \(p\) is the total number of the nodal values of this element. By substituting (B.1.1) into (4.3.5), we obtain the matrix form for the residual of the \(n\)th element as

\[
R^{(n)} = k^{(n)} U^{(n)} - q^{(n)},
\]
where

$$k^{(n)}(i, j) = \int_{-1}^{1} \left[ -\frac{\varphi_i \varphi'_j}{J_a} + (\gamma - 1 + Q_x) \varphi_i \varphi'_j - \gamma \varphi_i \varphi_j J_a \right] d\xi, \quad i, j = 1 \cdots p,$$

and

$$q^{(n)}(i) = \int_{-1}^{1} \gamma \varphi_i J_a d\xi, \quad i = 1 \cdots p.$$

In the terminology of solid mechanics, $k^{(n)}$ 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 [100].

Now, substituting (B.1.2) into (4.3.5), yields

$$R = \sum_{n=1}^{N} (k^{(n)} U^{(n)} - q^{(n)}),$$

which, after some algebraic manipulations, can be written as

$$R = KU - Q,$$

where $U$ is the vector of the nodal values of the entire domain, and $K$ and $Q$ are respectively, the master stiffness matrix and the master column matrix of the total applied generalized nodal loading [100].

In the actual evaluation of $K$ and $Q$, we assemble one element after another, and add the contribution of each element to the proper location in the master matrices, rather than determine the components $K(I, J)$ and $Q(I)$ for a fixed $I$ and $J$ by adding the contributions from each of the element at one time, as implied by (B.1.3-B.1.4). Mathematically, let the master matrices be denoted by $K^{(0)} (= 0)$ and $Q^{(0)} (= 0)$ before any contributions are added from an element, and $K^{(n)}$ and $Q^{(n)}$ after the $n$th element has been assembled, and the assembling procedure described above can be
written as

\[ K^{(n)}(i_r, i_s) = K^{(n-1)}(i_r, i_s) + k^{(n)}(r, s), \quad r, s = 1 \cdots p, \]

\[ Q^{(n)}(i_r) = Q^{(n-1)}(i_r) + q^{(n)}(r), \quad r = 1 \cdots p, \]

and the rest of \( K^{(n)}(i, j) \) are kept the same as \( K^{(n-1)}(i, j) \). Here, \( i_1 \cdots i_p \) are the global numbers, and they represent the transformation from local degree numbers to the global degree numbers.

For the case of linear shape functions, i.e., \( p = 2 \), \( \varphi_1(\xi) = \frac{1}{2}(1 - \xi) \), \( \varphi_2(\xi) = \frac{1}{2}(1 + \xi) \), we have, for \( I, J = 2 \cdots N \)

\[ K(I, J) = 0, \quad J < I - 1 \text{ or } J > I + 1, \]

\[ K(I - 1, I) = k^{(I-1)}(2, 1), \]

\[ K(I + 1, I) = k^{(I)}(1, 2), \]

\[ K(I, I) = k^{(I-1)}(2, 2) + k^{(I)}(1, 1); \]

and

\[ K(1, 1) = k^{(I-1)}(2, 1), \]

\[ K(I + 1, I) = k^{(I)}(1, 2), \]

\[ K(N + 1, N) = k^{(N)}(2, 1), \]

\[ K(N + 1, N + 1) = k^{(N)}(2, 2). \]

If the quadratic or even more complicated shape functions are adopted, the corresponding master matrices can also be obtained straightforwardly. For brevity, we omit the details here.

The next important step is to impose the constrained boundary conditions. Generally, if the \( j \)th component of \( U \) is already known, we define \( Q^* \) by multiplying the \( j \)th column of \( K \) by \( U(j) \), and subtracting it from \( Q \), except for the \( j \)th component,
which is set to $U(j)$. Then, $K^*$ is defined by setting the $j$th column and $j$th row of $K$ to zero, except for the diagonal term $K(j,j)$, which is set to 1. Specifically, we have:

for $i \leq j - 1$ and $i \geq j + 1$

$$Q^*(i) = Q(i) - K(i,j)U(j),$$
$$K^*(i,j) = K^*(j,i) = 0;$$

for $i = j$

$$Q^*(i) = U(j),$$
$$K^*(i,i) = 1.$$

For our problem, only the far field boundary condition needs to be imposed. When the linear shape functions are adopted, the constrained master matrices are as follows:

$$K^*(I,J) = K(I,J), \ 1 \leq I, J \leq N - 1,$$
$$K^*(I,J) = 0, \ I \ or \ J = N + 1,$$
$$K^*(N + 1, N + 1) = 1;$$

and

$$Q^*I = Q(I), \ 1 \leq I \leq N - 1,$$
$$Q^*(N) = Q(N) - K(N, N + 1)U(N + 1),$$
$$Q^*(N + 1) = U(N + 1).$$
B.2 The explicit form of the Jacobin matrix with linear shape functions

As mentioned earlier, the component that is related to the location of the fixed boundary should be excluded in the final form of the Jacobin matrix $J_R(x^*)$. For convenience, we first evaluate the complete Jacobian matrix of the residual $R$, and the required form can be obtained by deleting the last row and column of the complete one.

According to the definition of the Jacobin matrix, its $i$th column is defined as

$$
J_R(:, i) = \frac{\partial R}{\partial x^*(i)} = \frac{\partial K^*}{\partial x^*(i)} U - \frac{\partial Q^*}{\partial x^*(i)}.
$$

(B.2.1)

In case of linear shape functions, $\frac{\partial K^*}{\partial x^*(i)}$ and $\frac{\partial Q^*}{\partial x^*(i)}$ can be explicitly worked out as:

for $i = 2 \cdots N - 1$,

$$
\frac{\partial K^*}{\partial x^*(i)}_{N+1,N+1} = \begin{pmatrix}
0 & \cdots & \cdots & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\partial k^{(i-1)}(1, 1) & \partial k^{(i-1)}(2, 1) & 0 & \vdots & \vdots \\
\partial k^{(i-1)}(1, 1) & \partial k^{(i-1)}(2, 1) & \partial k^{(i)}(1, 1) & \partial k^{(i)}(1, 2) & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
0 & \cdots & \cdots & \cdots & 0
\end{pmatrix}
$$

(B.2.2)
\[
\left( \frac{\partial Q^*}{\partial x^*(i)} \right)_{N+1,1} = \begin{pmatrix}
0 \\
\vdots \\
\frac{\partial q^{(i-1)}(1)}{\partial x^*(i)} \\
\frac{\partial q^{(i-1)}(2)}{\partial x^*(i)} + \frac{\partial q^{(i-1)}(2)}{\partial x^*(i)} \\
\vdots \\
0
\end{pmatrix}; \quad (B.2.3)
\]

for \(i = 1\)

\[
\left( \frac{\partial K^*}{\partial x^*(1)} \right)_{N+1,N+1} = \begin{pmatrix}
\frac{\partial k^{(1)}(1,1)}{\partial x^*(1)} & \frac{\partial k^{(1)}(1,2)}{\partial x^*(1)} & \cdots & 0 \\
\frac{\partial k^{(1)}(2,1)}{\partial x^*(1)} & \frac{\partial k^{(1)}(2,2)}{\partial x^*(1)} & \cdots & \vdots \\
\vdots & \vdots & \ddots & \vdots \\
0 & \cdots & \cdots & 0
\end{pmatrix}; \quad (B.2.4)
\]

\[
\left( \frac{\partial Q^*}{\partial x^*(1)} \right)_{N+1,1} = \begin{pmatrix}
0 \\
\vdots \\
\frac{\partial q^{(1)}(1)}{\partial x^*(1)} \\
\frac{\partial q^{(1)}(2)}{\partial x^*(1)} \\
\vdots \\
0
\end{pmatrix}; \quad (B.2.5)
\]

for \(i = N\)

\[
\left( \frac{\partial K^*}{\partial x^*(N)} \right)_{N+1,N+1} = \begin{pmatrix}
0 & \cdots & \cdots & 0 \\
\vdots & \ddots & \vdots & \vdots \\
\frac{\partial k^{(N-1)}(1,1)}{\partial x^*(N)} & \cdots & \frac{\partial k^{(N-1)}(2,2)}{\partial x^*(N)} + \frac{\partial k^{(N)}(1,1)}{\partial x^*(N)} \\
\frac{\partial k^{(N-1)}(2,1)}{\partial x^*(N)} & \cdots & \vdots
\end{pmatrix}; \quad (B.2.6)
\]
\[
\left( \frac{\partial Q^*}{\partial x^*(N)} \right)_{N+1,1} = \begin{pmatrix}
0 \\
\vdots \\
\frac{\partial q^{(N-1)}(1)}{\partial x^*(N)} + \frac{\partial q^{(N)}(1)}{\partial x^*(N)} - \frac{\partial k^{(N)}(1,2)}{\partial x^*(N)} U(N + 1)
\end{pmatrix}, \quad (B.2.7)
\]

where

\[
\frac{\partial k^{(i-1)}(1,1)}{\partial x^*(i)} = -\frac{1}{6\Delta t} - \frac{\gamma}{3} + \frac{1}{(\Delta x^*_i)^2}, \quad \frac{\partial k^{(i)}(1,1)}{\partial x^*(i)} = -\frac{1}{3\Delta t} + \frac{\gamma}{3} - \frac{1}{(\Delta x^*_i)^2}, \quad (B.2.8)
\]

\[
\frac{\partial k^{(i-1)}(1,2)}{\partial x^*(i)} = \frac{1}{6\Delta t} - \frac{\gamma}{6} - \frac{1}{(\Delta x^*_i)^2}, \quad \frac{\partial k^{(i)}(1,2)}{\partial x^*(i)} = \frac{1}{3\Delta t} + \frac{\gamma}{6} + \frac{1}{(\Delta x^*_i)^2}, \quad (B.2.9)
\]

\[
\frac{\partial k^{(i-1)}(2,1)}{\partial x^*(i)} = -\frac{1}{3\Delta t} - \frac{\gamma}{6} - \frac{1}{(\Delta x^*_i)^2}, \quad \frac{\partial k^{(i)}(2,1)}{\partial x^*(i)} = -\frac{1}{6\Delta t} + \frac{\gamma}{6} - \frac{1}{(\Delta x^*_i)^2}, \quad (B.2.10)
\]

\[
\frac{\partial k^{(i-1)}(2,2)}{\partial x^*(i)} = \frac{1}{3\Delta t} - \frac{\gamma}{3} + \frac{1}{(\Delta x^*_i)^2}, \quad \frac{\partial k^{(i)}(2,2)}{\partial x^*(i)} = \frac{1}{6\Delta t} + \frac{\gamma}{3} - \frac{1}{(\Delta x^*_i)^2}. \quad (B.2.11)
\]

By substituting (B.2.2-B.2.11) into (B.2.1), and deleting the last row and column of the complete Jacobian matrix, we obtain the reduced Jacobian form at the exact solution \( x^* \) as: for \( i = 2 \cdots N - 1 \)

\[
J_R(i - 1, i) = \left( \frac{1}{6\Delta t} + \frac{\gamma}{3} - \frac{1}{\Delta(x^*_i)^2} \right) \Delta U_i - \frac{\gamma}{2} U(i) - \frac{\gamma}{2},
\]

\[
J_R(i, i) = \left( \frac{1}{3\Delta t} + \frac{\gamma}{6} + \frac{1}{\Delta(x^*_i)^2} \right) \Delta U_i + \left( \frac{1}{3\Delta t} + \frac{\gamma}{6} + \frac{1}{\Delta(x^*_i+1)^2} \right) \Delta U_{i+1},
\]

\[
J_R(i + 1, i) = \left( \frac{1}{6\Delta t} + \frac{\gamma}{3} - \frac{1}{\Delta(x^*_i+1)^2} \right) \Delta U_{i+1} + \frac{\gamma}{2} U(i) + \frac{\gamma}{2},
\]

\[
J_R(2 \leq j < i - 1, i) = J_R(i + 1 < j \leq N - 1, i) = 0;
\]
and

\[ J_R(1, 1) = \left( \frac{1}{3 \Delta t} + \frac{\gamma}{6} + \frac{1}{\Delta (x^*_2)^2} \right) \Delta U_2 + \frac{\gamma}{2} U(1) + \frac{\gamma}{2}, \]

\[ J_R(2, 1) = \left( \frac{1}{6 \Delta t} + \frac{\gamma}{3} - \frac{1}{\Delta (x^*_2)^2} \right) \Delta U_2 + \frac{\gamma}{2} U(1) + \frac{\gamma}{2}, \]

\[ J_R(N - 1, N) = \left( \frac{1}{6 \Delta t} + \frac{\gamma}{3} - \frac{1}{\Delta (x^*_N)^2} \right) \Delta U_N - \frac{\gamma}{2} U(N) - \frac{\gamma}{2}, \]

\[ J_R(N, N) = \left( \frac{1}{3 \Delta t} + \frac{\gamma}{6} + \frac{1}{\Delta (x^*_N)^2} \right) \Delta U_N + \left( \frac{1}{3 \Delta t} + \frac{\gamma}{6} + \frac{1}{\Delta (x^*_N+1)^2} \right) \Delta U_{N+1}, \]

where \( \Delta x^*_i = x^*(i) - x^*(i - 1) \), and \( \Delta U_i = U(i) - U(i - 1) \).

### B.3 Different initializations of elements and time steps

Different initializations of elements \( x^i_0 \) (i = 1, 2, 3) are defined as

\[ x^1_0 = \text{linspace}(0, \ln 5, N + 1), \]

\[ x^2_0 = [0, 0.01, 0.02, 0.04, 0.06, \text{linspace}(0.12, \ln 5, N - 4)], \]

\[ x^3_0 = [0, 0.001, 0.005, 0.01, 0.015, 0.02, 0.04, 0.08, 0.12, \text{linspace}(0.2, \ln 5, N - 8)], \]

where \( \text{linspace}(a, b, \text{Num}) \) stands for a linearly spaced vector starting from \( a \), ending at \( b \), with the step size \( \frac{b - a}{\text{Num} - 1} \).

The length of time step for each run is defined as

\[ t^1 = \text{linspace}(0, 1, M), \]

for \( M = 10 \),

\[ t^2 = [0.36, 0.36, 0.4136, 0.5527, 0.7387, 0.9872, 1.3193, 1.7632, 2.3563, 3.1491]/12*\sigma^2/2, \]
for $M = 25$, 

\[ t^2 = [0.1440, 0.1440, 0.1440, 0.1440, 0.1440, 0.1512, 0.1698, 0.1907, 0.2141, 0.2405, \\
0.2701, 0.3033, 0.3406, 0.3825, 0.4295, 0.4823, 0.5417, 0.6083, 0.6831, 0.7671, \\
0.8615, 0.9674, 1.0864, 1.2200, 1.3701]/12 \times \sigma^2/2. \]
Appendix C

The analysis for Chapter 6

C.1 The solution procedure of PDE system (6.3.12)-(6.3.14)

To find the solution of PDE system (6.3.12), we assume it can be written as

\[ p_0(X, T) = \sqrt{T}h_0(\xi), \tag{C.1.1} \]

where \( \xi = \frac{X}{2\sqrt{T}} \). By substituting (C.1.1) into (6.3.12), we obtain the ODE system for \( h_0(\xi) \) as

\[
\begin{align*}
  h_0''(\xi) + 2\xi h_0'(\xi) - 2h_0(\xi) &= 0, \\
  \lim_{\xi \to -\infty} h_0(\xi) &= 2\xi, & \lim_{\xi \to -\infty} h_0(\xi) &= 0.
\end{align*}
\]

The analytical solution of this ODE system can be readily found as

\[ h_0(\xi) = \frac{1}{\sqrt{\pi}} e^{-\xi^2} + \xi \text{erfc}(-\xi). \]

Similarly, by assuming that the solution of PDE system (6.3.13) is in the form of

\[ p_1(X, T) = Th_1(\xi), \]

we have

\[
\begin{cases}
    h''_1(\xi) + 2\xi h'_1(\xi) - 4h_1(\xi) = 2(1 - \rho + v)erf(-\xi) - \frac{4a'(0)}{\sqrt{\pi}} \xi e^{-\xi^2} + 4(\rho - v), \\
    \lim_{\xi \to \infty} h_1(\xi) = 2\xi^2, \quad \lim_{\xi \to -\infty} h'_1(\xi) = 0.
\end{cases}
\]  

(C.1.2)

Suppose that the solution \( h_1(\xi) \) has the following structure:

\[
h_1(\xi) = f(\xi)e^{-\xi^2} + g(\xi) \int_{-\infty}^{\xi} e^{-t^2} dt + m(\xi),
\]  

(C.1.3)

where \( f(\xi), g(\xi) \) and \( m(\xi) \) are polynomials in \( \xi \). By substituting (C.1.3) into (C.1.2), we obtain

\[
(f'' - 2\xi f' + 2g' - 6f)e^{-\xi^2} + (g'' + 2\xi g' - 4g) \int_{-\infty}^{\xi} e^{-t^2} dt + m'' + 2\xi m' - 4m
\]

\[
= \frac{4(1 - \rho + v)}{\sqrt{\pi}} \int_{-\infty}^{\xi} e^{-t^2} dt - \frac{4a'(0)}{\sqrt{\pi}} \xi e^{-\xi^2} + 4(\rho - v),
\]

which, combined with the boundary conditions at \( \xi = \pm \infty \), yields

\[
\begin{cases}
    m'' + 2\xi m' - 4m = 4(\rho - v), \\
    \lim_{\xi \to -\infty} m' = 0,
\end{cases}
\]

(C.1.4)

\[
\begin{cases}
    g'' + 2\xi g' - 4g = \frac{4(1 - \rho + v)}{\sqrt{\pi}}, \\
    \lim_{\xi \to -\infty} \sqrt{\pi}g + m = 2\xi^2,
\end{cases}
\]

(C.1.5)

\[
f'' - 2\xi f' + 2g' - 6f = -\frac{4a'(0)}{\sqrt{\pi}} \xi.
\]

(C.1.6)

The polynomial solutions of (C.1.4-C.1.6) can be readily found as:

\[
m(\xi) = v - \rho,
\]

\[
g(\xi) = \frac{2}{\sqrt{\pi}} \xi^2 + \frac{\rho - v}{\sqrt{\pi}},
\]

\[
f(\xi) = \frac{1}{\sqrt{\pi}} (1 + \frac{a'(0)}{2}) \xi.
\]
Therefore,

\[ h_1(\xi) = \frac{1}{\sqrt{\pi}} \left( 1 + \frac{a'(0)}{2} \right) \xi e^{-\xi^2} + \left( \frac{2}{\sqrt{\pi}} \xi^2 + \frac{\rho - v}{\sqrt{\pi}} \right) \int_{-\infty}^{\xi} e^{-t^2} dt + v - \rho. \]

With the utilization of the above solution technique, it is not hard to find the solution of (6.3.14), though it is quite complicated. By assuming that

\[ p_2(X, T) = T^{\frac{3}{2}} h_2(\xi), \]

and substituting it to (6.3.14), we obtain the ODE system for \( h_2(\xi) \) as

\[
\begin{align*}
& h_2''(\xi) + 2\xi h_2'(\xi) - 6h_2(\xi) = (C\xi^4 + A\xi^2 + B)e^{-\xi^2} + \frac{8(1 + v - \rho)}{\sqrt{\pi}} \xi \int_{-\infty}^{\xi} e^{-t^2} dt - 8v\xi, \\
& \lim_{\xi \to -\infty} h_2'(\xi) = \frac{4}{3} \xi^3 + 2\rho\xi, \quad \lim_{\xi \to -\infty} h_2''(\xi) = 0,
\end{align*}
\]

where

\[
\begin{align*}
A &= -\frac{4a''(0)}{\sqrt{\pi}} + \frac{6(a'(0))^2}{\sqrt{\pi}} - \frac{4a'(0)}{\sqrt{\pi}} - \frac{6a'(0)(v - \rho)}{\sqrt{\pi}}, \\
B &= \frac{\rho - v - 1}{\sqrt{\pi}} (-2 - a'(0) + 2v - 2\rho), \\
C &= -\frac{4(a''(0))^2}{\sqrt{\pi}}.
\end{align*}
\]

Suppose \( h_2(\xi) \) can be written as

\[ h_2(\xi) = f(\xi)e^{-\xi^2} + g(\xi) \int_{-\infty}^{\xi} e^{-t^2} dt + m(\xi), \]

where \( f(\xi), g(\xi) \) and \( m(\xi) \) are polynomials in \( \xi \). By using the same procedure as in deriving \( h_1 \), the ODE systems for \( f(\xi), g(\xi) \) and \( m(\xi) \) can be easily found as

\[
\begin{align*}
& m'' + 2\xi m' - 6m = -8v\xi, \\
& \lim_{\xi \to -\infty} m'' = 0,
\end{align*}
\]

(C.1.7)
\[
\begin{align*}
\begin{cases}
g'' + 2\xi g' - 4g &= \frac{8(1 + v - \rho)}{\sqrt{\pi}}, \\
\lim_{\xi \to \infty} \sqrt{\pi}g + m &= \frac{4\xi^2}{3} + 2\rho \xi,
\end{cases} \\
\int_{-\infty}^{\xi} e^{-t^2} dt &+ 2v \xi.
\end{align*}
\]

The polynomial solutions of (C.1.7-C.1.9) are:

\[
\begin{align*}
m(\xi) &= 2v \xi, \\
g(\xi) &= \frac{4\xi^3}{3\sqrt{\pi}} + \frac{2(\rho - v)\xi}{\sqrt{\pi}}, \\
f(\xi) &= -\frac{C}{16} \xi^4 - \left(\frac{C}{16} + \frac{\tilde{A}}{12}\right) \xi^2 - \frac{C}{64} - \frac{\tilde{A}}{48} - \frac{\tilde{B}}{8},
\end{align*}
\]

where

\[
\begin{align*}
\tilde{A} &= A - \frac{8}{\sqrt{\pi}}, \\
\tilde{B} &= B - \frac{4(\rho - v)}{\sqrt{\pi}}.
\end{align*}
\]

Therefore,

\[
h_2(\xi) = \left[ -\frac{C}{16} \xi^4 - \left(\frac{C}{16} + \frac{\tilde{A}}{12}\right) \xi^2 - \frac{C}{64} - \frac{\tilde{A}}{48} - \frac{\tilde{B}}{8} \right] e^{\xi^2} + \left(\frac{4\xi^2}{3\sqrt{\pi}} + \frac{2(\rho - v)\xi}{\sqrt{\pi}}\right) \int_{-\infty}^{\xi} e^{-t^2} dt + 2v \xi.
\]

### C.2 Corner layer analysis

In the corner layer, we adopt the following rescales:

\[
X = \frac{x}{\sqrt{\epsilon}}, \quad p = \frac{P}{\sqrt{\epsilon}}.
\] (C.2.1)

Assuming that \( p \) can be expanded in powers of \( \sqrt{\epsilon} \), i.e.,

\[
p = p_0 + \sqrt{\epsilon}p_1 + \epsilon p_2 + \mathcal{O}(\epsilon^2),
\] (C.2.2)
and substituting (C.2.1-C.2.2) into (6.3.2), we obtain the following sequence of PDE systems:

\[
\begin{align*}
\frac{\partial p_0}{\partial T} &= \frac{\partial^2 p_0}{\partial X^2}, \\
p_0(X, 0) &= \max(X, 0), \\
\lim_{X \to +\infty} p_0(X, T) &= X, \quad \lim_{X \to -\infty} p_0(X, T) = 0,
\end{align*}
\]

\[
\begin{align*}
\frac{\partial p_1}{\partial T} &= \frac{\partial^2 p_1}{\partial X^2} + a'(0)X \frac{\partial^2 p_0}{\partial X^2} + (\rho - v - 1) \frac{\partial p_0}{\partial X} + v - \rho, \\
p_1(X, 0) &= \max\left(\frac{1}{2}X^2, 0\right), \\
\lim_{X \to +\infty} p_1(X, T) &= \frac{1}{2}X^2, \quad \lim_{X \to -\infty} p_1(X, T) = (v - \rho)T,
\end{align*}
\]

\[
\begin{align*}
\frac{\partial p_2}{\partial T} &= \frac{\partial^2 p_2}{\partial X^2} + a'(0)X \frac{\partial^2 p_1}{\partial X^2} + \frac{1}{2} a''(0)X^2 \frac{\partial^2 p_0}{\partial X^2} + (\rho - v - 1) \frac{\partial p_1}{\partial X} - a'(0)X \frac{\partial p_0}{\partial X} + vX, \\
p_2(X, 0) &= \max\left(\frac{1}{6}X^3, 0\right), \\
\lim_{X \to +\infty} p_2(X, T) &= \frac{1}{6}X^3 + \rho XT, \quad \lim_{X \to -\infty} p_2(X, T) = vX T.
\end{align*}
\]

Here, the boundary conditions as \(X \to +\infty\) are obtained by matching with one branch of the outer expansion, which is valid for \(X \gg \sqrt{\epsilon}\); whereas the ones as \(X \to -\infty\) are obtained by matching with another branch (\(X \ll -\sqrt{\epsilon}\)). The solutions of these PDE systems are the same as those in Appendix C.1, i.e.,

\[
\begin{align*}
p_0(X, T) &= \sqrt{T} h_0(\xi), \\
p_1(X, T) &= T h_1(\xi), \\
p_2(X, T) &= T^3 h_2(\xi),
\end{align*}
\]

where \(h_0(\xi), h_1(\xi)\) and \(h_2(\xi)\) are defined in Appendix C.1.
C.3 The solution procedure of the PDE systems

(6.3.22)-(6.3.24)

Again, we shall use the similarity solution techniques to derive the solutions of

\( (6.3.22-6.3.24) \). Suppose

\[
p_0(X,T) = \sqrt{T}h_0(\xi), \quad p_1(X,T) = Th_1(\xi), \quad p_2(X,T) = T^3h_2(\xi),
\]

where \( \xi = \frac{X}{2\sqrt{T}} \). The ODE systems for \( h_0(\xi), h_1(\xi) \) and \( h_2(\xi) \) can be derived as

\[
\begin{cases}
    h_0''(\xi) + 2\xi h_0'(\xi) - 2h_0(\xi) = 0 \\
    \lim_{\xi \to -\infty} h_0(\xi) = 2\xi, \quad \lim_{\xi \to -\infty} h_0(\xi) = 0,
\end{cases}
\]

\[
\begin{cases}
    h_1''(\xi) + 2\xi h_1'(\xi) - 2h_1(\xi) = 2(1 - \rho + v)erfc(-\xi) - \frac{4a_x(0,0)}{\sqrt{\pi}} \xi e^{-\xi^2} + 4(\rho - v), \\
    \lim_{\xi \to -\infty} h_1(\xi) = 2\xi^2, \quad \lim_{\xi \to -\infty} h_1'(\xi) = 0,
\end{cases}
\]

\[
\begin{cases}
    h_2''(\xi) + 2\xi h_2'(\xi) - 2h_2(\xi) = (C\xi^4 + A\xi^2 + B)e^{-\xi^2} \\
    \quad + \frac{8(1 + v - \rho)}{\sqrt{\pi}} \int_{-\infty}^{\xi} e^{-t^2} dt - 8v\xi,
\end{cases}
\]

\[
\begin{cases}
    \lim_{\xi \to -\infty} h_2(\xi) = \frac{4}{3}\xi^3 + 2\rho\xi, \quad \lim_{\xi \to -\infty} h_2'(\xi) = 0,
\end{cases}
\]

where

\[
A = -\frac{2a_x(0,0)(2 - 3a_x(0,0) + 2v - 2\rho)}{\sqrt{\pi}} - \frac{2a_x(0,0)(1 + v - \rho)}{\sqrt{\pi}} - \frac{4a_x(0,0)}{\sqrt{\pi}},
\]

\[
B = -\frac{(1 + v - \rho)(-2 - a_x(0,0) + 2v - 2\rho)}{\sqrt{\pi}} - \frac{2a_x(0,0)}{\sqrt{\pi}},
\]

\[
C = -\frac{4a_x^2(0,0)}{\sqrt{\pi}}.
\]
By using the solution techniques as introduced in Appendix C.1, we obtain

\[ h_0(\xi) = \frac{1}{\sqrt{\pi}} e^{-\xi^2} + \xi \text{erfc}(-\xi), \]

\[ h_1(\xi) = \frac{1}{\sqrt{\pi}} (1 + \frac{a_x(0, 0)}{2}) \xi e^{-\xi^2} + \left( \frac{2}{\sqrt{\pi}} \xi^2 + \frac{\rho - v}{\sqrt{\pi}} \right) \int_{-\infty}^{\xi} e^{-t^2} dt + v - \rho, \]

\[ h_2(\xi) = \left[ -\frac{C_1}{16} \xi^4 - \left( \frac{C_1}{16} + \frac{\tilde{A}}{12} \right) \xi^2 - \frac{C_2}{64} - \frac{\tilde{A}}{48} + \left( \frac{4 \xi^3}{3 \sqrt{\pi}} + \frac{2(\rho - v) \xi}{\sqrt{\pi}} \right) \right] e^{-\xi^2} \int_{-\infty}^{\xi} e^{-t^2} dt + 2v \xi, \]

where

\[ \tilde{A} = A - \frac{8}{\sqrt{\pi}}, \]

\[ \tilde{B} = B - \frac{4(\rho - v)}{\sqrt{\pi}}. \]
Appendix D

The analysis for Chapter 7

D.1 The solution procedure for the solution in the inner region

The PDE systems for $P_0$, $P_1$ and $P_2$ are:

\[
\begin{align*}
\frac{\partial P_0}{\partial T} &= \frac{v}{\sigma^2} \frac{\partial^2 P_0}{\partial X^2}, \\
P_0(X, v, 0) &= \max(X, 0), \quad \lim_{X \to \infty} P_0(X, v, T) = X, \quad \lim_{X \to -\infty} P_0(X, v, T) = 0, \\
\frac{\partial P_1}{\partial T} &= \frac{v}{\sigma^2} \frac{\partial^2 P_1}{\partial X^2} + \left( q - d - \frac{v}{\sigma^2} \right) \frac{\partial P_0}{\partial X} + \frac{2 \rho v}{\sigma} \frac{\partial^2 P_0}{\partial X \partial v} + d - q, \\
P_1(X, v, 0) &= \max(\frac{1}{2} X^2, 0), \quad \lim_{X \to \infty} P_1(X, v, T) = \frac{1}{2} X^2 + qXT, \quad \lim_{X \to -\infty} \frac{\partial P_1}{\partial X}(X, v, T) = 0, \\
\frac{\partial P_2}{\partial T} &= \frac{v}{\sigma^2} \frac{\partial^2 P_2}{\partial X^2} + \left( q - d - \frac{v}{\sigma^2} \right) \frac{\partial P_1}{\partial X} + \frac{2 \rho v}{\sigma} \frac{\partial^2 P_1}{\partial X \partial v} \\
&\quad + \left( q - d - \frac{v}{\sigma^2} \right) \frac{\partial^2 P_0}{\partial v^2} + q \frac{2 \kappa \sigma^2}{\sigma^2} (\eta - v) \frac{\partial P_0}{\partial v} + dX, \\
P_2(X, v, 0) &= \max(\frac{1}{6} X^3, 0), \quad \lim_{X \to -\infty} \frac{\partial^2 P_2}{\partial X^2}(X, v, T) = 0, \\
&\lim_{X \to \infty} P_2(X, v, T) = \frac{1}{6} X^3 + qXT, \quad \lim_{X \to -\infty} \frac{\partial^2 P_2}{\partial X^2}(X, v, T) = 0.
\end{align*}
\]
To find the solution of the PDE system (D.1.1), we assume it can be written as

$$P_0(X, v, T) = \sqrt{T} h_0(\xi; a),$$

(D.1.4)

where $\xi = \frac{X}{2\sqrt{T}}$ and $a = \frac{v}{\sigma^2}$. By substituting (D.1.4) into (D.1.1), we obtain the ODE system for $h_0(\xi; a)$ as

$$\begin{cases}
  a h''_0(\xi; a) + 2 \xi h'_0(\xi; a) - 2 h_0(\xi; a) = 0, \\
  \lim_{\xi \to -\infty} h_0(\xi; a) = 2\xi, \\
  \lim_{\xi \to \infty} h_0(\xi; a) = 0.
\end{cases}$$

The analytical solution of this ODE system can be readily found as

$$h_0(\xi; a) = \frac{\sqrt{a}}{\sqrt{\pi}} e^{-\xi^2} + \xi \text{erfc}(\frac{\xi}{\sqrt{a}}).$$

(D.1.5)

Similarly, by assuming that the solution of PDE system (D.1.2) is in the form of

$$P_1(X, v, T) = Th_1(\xi; a) = Ta\tilde{h}_1(\xi_1),$$

where $\xi_1 = \frac{\xi}{\sqrt{a}}$, we have

$$\begin{cases}
  \tilde{h}''_1(\xi_1) + 2 \xi \tilde{h}'_1(\xi_1) - 4 \tilde{h}_1(\xi_1) = 2 \frac{(a - q + d)}{a} \text{erfc}(\xi_1) \\
  + \frac{4\rho}{\sigma \sqrt{\pi} a} \xi_1 e^{-\xi_1^2} + 4 \frac{(q - d)}{a}, \\
  \lim_{\xi_1 \to -\infty} \tilde{h}_1(\xi_1) = 2\xi_1^2, \\
  \lim_{\xi_1 \to +\infty} \tilde{h}'_1(\xi_1) = 0.
\end{cases}$$

(D.1.6)

Suppose that the solution $\tilde{h}_1(\xi_1)$ has the following structure:

$$h_1(\xi) = f(\xi_1)e^{-\xi_1^2} + g(\xi_1) \int_{-\infty}^{\xi_1} e^{-t^2} dt + m(\xi_1),$$

(D.1.7)
where $f(\xi_1), g(\xi_1)$ and $m(\xi_1)$ are polynomials in $\xi_1$. By substituting (D.1.7) into
(D.1.6), we obtain
\[
(f'' - 2\xi_1 f' + 2g' - 6f)e^{-\xi_1^2} + (g'' + 2\xi_1 g' - 4g)\int_{-\infty}^{\xi_1} e^{-t^2} dt + m'' + 2\xi_1 m' - 4m
= \frac{4(a - q + d)}{a\sqrt{\pi}} \int_{-\infty}^{\xi_1} e^{-t^2} dt + \frac{4\rho}{\sigma a \sqrt{\pi}} \xi_1 e^{-\xi_1^2} + 4 \frac{(q - d)}{a},
\]
which, combined with the boundary conditions at $\xi_1 = \pm \infty$, yields
\[
\begin{cases}
  m'' + 2\xi_1 m' - 4m = \frac{4(q - d)}{a}, \\
  \lim_{\xi_1 \to -\infty} m' = 0,
\end{cases}
\]
(D.1.8)
\[
\begin{cases}
  g'' + 2\xi_1 g' - 4g = \frac{4(a - q + d)}{a\sqrt{\pi}}, \\
  \lim_{\xi_1 \to -\infty} \sqrt{\pi} g + m = 2\xi_1^2,
\end{cases}
\]
(D.1.9)
\[
f'' - 2\xi_1 f' + 2g' - 6f = \frac{4\rho}{\sigma \sqrt{\pi} a} \xi_1.
\]
(D.1.10)
The polynomial solutions of (D.1.8)-(D.1.10) can be readily found as:
\[
\begin{align*}
m(\xi_1) &= \frac{d - q}{a}, \\
g(\xi_1) &= \frac{2}{\sqrt{\pi}} \xi_1^2 + \frac{q - d}{\sqrt{a\pi}}, \\
f(\xi_1) &= \frac{1}{\sqrt{\pi}} (1 - \frac{\rho}{2\sigma a}) \xi_1.
\end{align*}
\]
Therefore,
\[
h_1(\xi; a) = \frac{1}{\sqrt{\pi}} (a - \frac{\rho}{2\sigma}) \frac{\xi}{\sqrt{a}} e^{-\xi^2} + \left( \frac{2}{\sqrt{\pi}} \xi^2 + \frac{q - d}{\sqrt{a\pi}} \right) \int_{-\infty}^{\xi} e^{-t^2} dt + d - q.
\]
By means of the above solution technique, it is not hard to find the solution of
(D.1.3), though the expression is quite complicated. By assuming
\[
P_2(X, v, T) = T^{\frac{3}{2}} h_2(\xi; a) = T^{\frac{3}{2}} a^{\frac{3}{2}} \tilde{h}_2(\xi_1),
\]
and substituting it to (D.1.3), we obtain the ODE system for \( \tilde{h}_2(\xi_1) \) as

\[
\begin{align*}
\tilde{h}_2''(\xi_1) + 2\xi_1 \tilde{h}_2'(\xi_1) - 6\tilde{h}_2(\xi_1) &= (C\xi_1^4 + A\xi_1^2 + B)e^{-\xi_1^2} \\
+ \frac{8(a + d - q)}{a\sqrt{\pi}} \int_{-\infty}^{\xi_1} e^{-t^2} dt - \frac{8d}{a}\xi_1, \\
\lim_{\xi_1 \to -\infty} \tilde{h}_2(\xi_1) &= \frac{4}{3}\xi_1^3 + \frac{2q}{a}\xi_1, \\
\lim_{\xi_1 \to +\infty} \tilde{h}_2''(\xi_1) &= 0,
\end{align*}
\]

where

\[
\begin{align*}
A &= \frac{2(a - q + d)\rho}{a^2\sigma\sqrt{\pi}} - \frac{\rho}{4a\sqrt{\pi}a^2}(\frac{32\rho}{\sigma} - 16a + 16q - 16d) - \frac{2}{\sigma^2\sqrt{\pi}a^2}, \\
B &= \frac{2(a - q + d)}{a^2}\left(-\frac{\rho}{2\sigma\sqrt{\pi}} + \frac{a}{\sqrt{\pi}} + \frac{q - d}{\sqrt{\pi}}\right) - \frac{\rho}{4a\sqrt{\pi}a^2}(8a + \frac{4\rho}{\sigma} - 8q + 8d) \\
&\quad + \frac{1}{\sigma^2\sqrt{\pi}a^2} - \frac{4\kappa}{\sigma^2\sqrt{\pi}a^2}(\frac{q}{\sigma^2} - a), \\
C &= -\frac{4\rho^2}{\sigma^2\sqrt{\pi}a^2}.
\end{align*}
\]

Suppose \( \tilde{h}_2(\xi_1) \) can be written as

\[
\tilde{h}_2(\xi_1) = f(\xi_1)e^{-\xi_1^2} + g(\xi_1) \int_{-\infty}^{\xi_1} e^{-t^2} dt + m(\xi_1),
\]

where \( f(\xi_1), g(\xi_1) \) and \( m(\xi_1) \) are polynomials in \( \xi_1 \). By using the same procedure as in deriving \( h_1 \), the systems for \( f(\xi_1), g(\xi_1) \) and \( m(\xi_1) \) can be easily found as

\[
\begin{align*}
&\begin{cases}
m'' + 2\xi_1 m' - 6m = -\frac{8d}{a}\xi_1, \\
\lim_{\xi_1 \to -\infty} m'' = 0,
\end{cases} \quad \text{(D.1.12)} \\
&\begin{cases}
g'' + 2\xi_1 g' - 6g = \frac{8(a + d - q)}{a\sqrt{\pi}}\xi_1, \\
\lim_{\xi_1 \to -\infty} \sqrt{\pi}g + m = \frac{4\xi_1^2}{3} + \frac{2q}{a}\xi_1, 
\end{cases} \quad \text{(D.1.13)} \\
f'' - 2\xi_1 f' - 8f + 2g' = C\xi_1^4 + A\xi_1^2 + B. \quad \text{(D.1.14)}
\end{align*}
\]
The polynomial solutions of (D.1.12)-(D.1.13) are:

\[
\begin{align*}
m(\xi_1) &= \frac{2d}{a} \xi_1, \\
g(\xi_1) &= \frac{4\xi_1^3}{3\sqrt{\pi}} + \frac{2(q-d)\xi_1}{a\sqrt{\pi}}, \\
f(\xi_1) &= -\frac{C_16\xi_4}{16\xi_1} - \frac{C}{16} + \frac{\hat{A}}{12} \xi_1^2 - \frac{C}{64} - \frac{\hat{A}}{48} - \frac{\hat{B}}{8},
\end{align*}
\]

where

\[
\begin{align*}
\hat{A} &= A - \frac{8}{\sqrt{\pi}}, \\
\hat{B} &= B - \frac{4(q-d)}{a\sqrt{\pi}}.
\end{align*}
\]

Therefore,

\[
\begin{align*}
h_2(\xi; a) &= a^\frac{3}{2} \left[ -\frac{C\xi_4^4}{16 a^2} - \frac{C}{16} + \frac{\hat{A}}{12} \frac{\xi_2^2}{a} - \frac{C}{64} - \frac{\hat{A}}{48} - \frac{\hat{B}}{8} e^{-\xi_1^2} \right. \\
&\quad + a^\frac{3}{2} \left( \frac{4\xi_1^3}{3\sqrt{\pi}a^\frac{3}{2}} + \frac{2(q-d)\xi_1}{\sqrt{\pi}a^\frac{3}{2}} \right) \int_{-\infty}^{\xi_1} e^{-t^2} dt + 2d\xi, \tag{D.1.15}
\end{align*}
\]
D.2 The solution procedure for the solution inside the corner

The PDE systems for $Z_{11}$, $Z_{12}$ and $Z_{13}$ are:

\[
\begin{align*}
\frac{\partial Z_{11}}{\partial T} &= V \frac{\partial^2 Z_{11}}{\partial V^2} + \frac{2\kappa}{\eta} \frac{\partial Z_{11}}{\partial V}, \\
Z_{11}(X, V, 0) &= \max(X, 0), \\
\lim_{V \to \infty} Z_{11}(X, v, T) &= \max(X, 0), \lim_{V \to 0} P_0(X, v, T) = X, \quad X > 0,
\end{align*}
\]

\[
\begin{align*}
\frac{\partial Z_{12}}{\partial T} &= V \frac{\partial^2 Z_{12}}{\partial V^2} + \frac{2\kappa}{\eta} \frac{\partial Z_{12}}{\partial V} + (q - d) \frac{\partial Z_{11}}{\partial X} + \frac{2\rho V}{\sigma} \frac{\partial^2 Z_{11}}{\partial X \partial V} + d - q, \\
Z_{12}(X, v, 0) &= \max\left(\frac{1}{2}X^2, 0\right), \lim_{V \to 0} Z_{12} = \frac{1}{2}X^2, \quad X > 0,
\end{align*}
\]

\[
\begin{align*}
\lim_{V \to \infty} Z_{12}(X, v, T) &= \frac{1}{2}X^2, \quad X > 0, \\
(d - q)T, \quad X < 0,
\end{align*}
\]

\[
\begin{align*}
\frac{\partial Z_{13}}{\partial T} &= V \frac{\partial^2 Z_{13}}{\partial V^2} + \frac{2\kappa}{\eta} \frac{\partial Z_{13}}{\partial V} + (q - d) \frac{\partial Z_{12}}{\partial X} + \frac{2\rho V}{\sigma} \frac{\partial^2 Z_{12}}{\partial X \partial V} \\
&\quad + \frac{V}{\sigma^2} \frac{\partial^2 Z_{11}}{\partial X^2} - \frac{2\kappa V}{\sigma^2} \frac{\partial Z_{11}}{\partial V} + dX, \\
Z_{13}(X, v, 0) &= \max\left(\frac{1}{6}X^3, 0\right), \lim_{V \to 0} Z_{13}(X, v, T) = qTX + \frac{1}{6}X^3, \quad X > 0.
\end{align*}
\]

\[
\begin{align*}
\lim_{V \to \infty} Z_{13}(X, v, T) &= qTX + \frac{1}{6}X^3, \quad X > 0, \\
dTX, \quad X < 0,
\end{align*}
\]

It is straightforward to derive the solutions of the above PDE systems, i.e.,

\[
Z_{11}(X, V, T) = \begin{cases} 
X, & X > 0, \\
0, & X < 0;
\end{cases}
\]

\[
Z_{12}(X, V, T) = \begin{cases} 
\frac{X^2}{2}, & X > 0, \\
(d - q)T, & X < 0;
\end{cases}
\]

\[
Z_{13}(X, V, T) = \begin{cases} 
qTX + \frac{1}{6}X^3, & X > 0, \\
dTX, & X < 0.
\end{cases}
\]
\[ Z_{13}(X, V, T) = \begin{cases} 
&T X + \frac{1}{6} X^3, \quad X > 0, \\
&d T, \quad X < 0. 
\end{cases} \]
Appendix E

Laplace inversions of Chapter 11

E.1 The evaluation of \( L^{-1}\left[ \hat{H}(x,p) |_{x=x}\right] \)

Since it is already known that

\[
F(x, \tau) = \int_{-\infty}^{\bar{x}} \frac{1}{2\sqrt{\pi \tau}} e^{-\frac{1}{2}(x-z)^2 - (\frac{k^2}{4} + \gamma + p)(x-z)} f(z) dz,
\]

we obtain

\[
\hat{F}(x, p) = L(F(x, \tau)) = \left[ \int_{-\infty}^{\bar{x}} e^{\lambda_1 (x-z)} f(z) dz + \int_{-\infty}^{\bar{x}} e^{\lambda_2 (x-z)} f(z) dz + \int_{x}^{\bar{x}} e^{\lambda_1 (x-z)} f(z) dz \right],
\]

and thus

\[
\frac{\hat{H}(x, p)}{\lambda_2 - \lambda_1} |_{x=\bar{x}} = \left. \frac{\partial F}{\partial x}(x, \tau) \right|_{x=\bar{x}} = \left. \frac{\partial \hat{F}}{\partial x}(x, p) \right|_{x=\bar{x}} = \int_{-\infty}^{\bar{x}} \frac{e^{\lambda_2 (x-z)}}{\lambda_1 - \lambda_2} f(z) dz
\]

\[
= \int_{-\infty}^{\bar{x}} \frac{e^{\lambda_1 (x-z)}}{\lambda_2 - \lambda_1} f(z) dz = \int_{-\infty}^{\bar{x}} e^{\frac{k^2 (x-z)}{4} + \gamma + p} f(z) dz.
\]  
(E.1.1)
Now, applying the Laplace inversion to (E.1.1), we obtain

\[
L^{-1}\left[ \frac{\hat{H}(x,p)}{\lambda_2 - \lambda_1} \bigg|_{x=\bar{x}} \right] = L^{-1}\left[ \int_{-\infty}^{\bar{x}} \frac{e^{-\frac{1}{2}(\bar{x}-z)^2 + \gamma + p(\bar{x}-z)}}{2\sqrt{\frac{k^2}{4} + \gamma + p}} f(z) dz \right]
\]

\[
= \int_{-\infty}^{\bar{x}} \frac{e^{-\frac{1}{2}(\bar{x}-z)^2}}{2} f(z) L^{-1}\left[ \frac{e^{\sqrt{\frac{k^2}{4} + \gamma + p}(\bar{x}-z)}}{2\sqrt{\frac{k^2}{4} + \gamma + p}} \right] dz
\]

\[
= \int_{-\infty}^{\bar{x}} \frac{e^{-\frac{1}{2}(\bar{x}-z)^2}}{2} f(z) e^{-\left(\frac{k^2}{4} + \gamma\right)\tau} L^{-1}\left[ \frac{e^{-\sqrt{\frac{k^2}{4} + \gamma}(\bar{x}-z)}}{\sqrt{\tau}} \right] dz \quad \text{(shift theorem)}
\]

\[
= \int_{-\infty}^{\bar{x}} \frac{e^{-\frac{1}{2}(\bar{x}-z)^2} - \left(\frac{k^2}{4} + \gamma\right)\tau}{2\sqrt{\pi \tau}} e^{-\frac{(\bar{x}-z)^2}{4\tau}} f(z) dz.
\]

### E.2 The evaluation of \( L^{-1}\left[ \frac{\hat{G}(x,p)}{\lambda_1 - \lambda_2} \bigg|_{x=\bar{x}} \right] \)

According to the Convolution theorem [77], we obtain

\[
L^{-1}\left[ \frac{\hat{G}(x,p)}{\lambda_1 - \lambda_2} \bigg|_{x=\bar{x}} \right] = \int_{0}^{\tau} \frac{e^{-\left(\frac{k^2}{4} + \gamma\right)(\tau-s)}}{2\sqrt{\pi (\tau-s)}} G(x, s)_{|_{x=\bar{x}}} ds
\]

\[
= \int_{0}^{\tau} \frac{e^{-\left(\frac{k^2}{4} + \gamma\right)(\tau-s)}}{2\sqrt{\pi (\tau-s)}} \int_{s-j}^{0} W(\xi) \frac{\partial q_2}{\partial x} (x, s - \xi)_{|_{x=\bar{x}}} d\xi ds
\]

\[
= \int_{0}^{\tau} \frac{e^{-\left(\frac{k^2}{4} + \gamma\right)(\tau-s)}}{2\sqrt{\pi (\tau-s)}} \int_{0}^{j-s} W(s - \bar{J} + m) \frac{\partial q_2}{\partial x} (x, \bar{J} - m)_{|_{x=\bar{x}}} dmds \quad (E.2.1)
\]

On the other hand, since \( \frac{\partial g_2}{\partial x}(\bar{x}, \tau) \) is singular at \( (\bar{x}, 0) \), we evaluate the value \( \frac{\partial g_2}{\partial x}(\bar{x}, \tau) \) in the Laplace space, and then convert it back. We obtain

\[
\frac{\partial g_2}{\partial x}(\bar{x}, \tau) = L^{-1}[\lambda_2]
\]

\[
= -\frac{k}{2} \delta(\tau) - \frac{e^{-\left(\frac{k^2}{4} + \gamma\right)\tau}}{\sqrt{\pi}} \int_{0^{-}}^{\tau} \frac{\delta(s)}{\sqrt{\tau - s}} ds. \quad (E.2.2)
\]

Here, the left limit \( -' \) and the right limit \( +' \) are introduced to ensure the regularity of the integration. When the kernel function is integrable, such sided limit processes are longer needed.
By substituting (E.2.2) into (E.2.1), we obtain

\[
L^{-1}\left[ \frac{G(x, p)}{\lambda_1 - \lambda_2} \right]_{x=\bar{x}} = -\frac{k}{4\sqrt{\pi}} \left[ \int_0^\tau e^{-\left(\frac{k^2}{\tau} + \gamma\right)(\tau-s)} W(s - \bar{J} + m)\delta(\bar{J} - m) \,emd\tau \int_0^{(\bar{J}-m)^+} \frac{\delta'(\xi)}{\sqrt{\bar{J} - m - \xi}} \,d\xi \right] \\
= -\frac{1}{2\pi} \left[ \int_0^\tau e^{-\left(\frac{k^2}{\tau} + \gamma\right)(\tau-s)} W(s - \bar{J} + m) \frac{\delta'(\xi)}{\sqrt{\tau - s}} \,d\xi \int_0^{(\bar{J}-m)^+} \frac{\delta'(\xi)}{\sqrt{\bar{J} - m - \xi}} \,d\xi \right] \\
= 0.
\]

In the following, we shall evaluate the integrals \( I \) and \( II \) separately. When evaluating \( I \), we can change the order of the two integrals, and we obtain

\[
I = \left[ \int_0^{(J-\gamma)^+} \delta(\bar{J} - m) \int_0^\tau e^{-\left(\frac{k^2}{\tau} + \gamma\right)(\tau-s)} W(s - \bar{J} + m) \,ds \,dm \right] \\
+ \int_{(J-\gamma)^+}^{J^+} \delta(\bar{J} - m) \int_0^{(J-m)^+} e^{-\left(\frac{k^2}{\tau} + \gamma\right)(\tau-s)} W(s - \bar{J} + m) \,ds \,dm.
\]

It should be remarked that the above change of the order of the double integrals makes sense, as the inner integrated function \( e^{-\left(\frac{k^2}{\tau} + \gamma\right)(\tau-s)} W(s - \bar{J} + m) \) only has a removable singularity at \( \tau = s \), as long as \( W(\tau) \in C^1 \), which is certainly true from the financial point of view. Therefore, when \( \tau > 0 \),

\[
I = -\frac{k}{4\sqrt{\pi}} \left[ 0 + \lim_{m \to J} \int_0^{J-m} e^{-\left(\frac{k^2}{\tau} + \gamma\right)(\tau-s)} W(s - \bar{J} + m) \,ds \right] \\
= 0,
\]

whereas when \( \tau = 0 \),

\[
I = -\frac{k}{4\sqrt{\pi}} \left[ \left( \lim_{\tau \to 0} \lim_{m \to J} \int_0^{\tau} + \lim_{m \to J} \int_0^{(J-m)^+} \right) e^{-\left(\frac{k^2}{\tau} + \gamma\right)(\tau-s)} \frac{W(s - \bar{J} + m)}{\sqrt{\tau - s}} \right] \\
= 0.
\]
Thus, one can conclude that the integral $I$ equals to zero.

On the other hand, to evaluate $II$, we also perform the change of the order of the integrals first, and we obtain

$$II = -\frac{1}{2\pi} \left[ \int_{0}^{\tau} \delta'(\xi) \int_{0}^{J-s} K(m, s; \xi) dmdsd\xi \right]$$

$$= \frac{1}{2\pi} \lim_{\xi \to 0} \frac{\partial}{\partial \xi} \left[ \int_{0}^{\tau} \int_{0}^{J-s} K(m, s; \xi) dmds \right]$$

$$= \frac{1}{2\pi} \int_{0}^{\tau} \frac{e^{-\left(\frac{k^2}{4}\tau + \gamma\right)\left(\tau-s\right)}}{\sqrt{\tau-s}} \lim_{\xi \to 0} \frac{\partial}{\partial \xi} \left[ \int_{0}^{J-s} \frac{e^{-\left(\frac{k^2}{4}\tau + \gamma\right)(J-m)}}{\sqrt{J-m-\xi}} dm \right] ds. \quad (E.2.3)$$

Now, introducing a new variable $t = \sqrt{J-m-\xi}$, and substituting it into (E.2.3),
we obtain

\[
II = \frac{1}{2\pi} \int_0^\tau \frac{e^{-(k^2/4)(\tau-s)}}{\sqrt{\tau-s}} \lim_{\xi \to 0} \frac{\partial}{\partial \xi} \left[ \int_{\sqrt{s-\xi}}^{\sqrt{\tau-\xi}} 2e^{-(k^2/4)(\xi-t^2)} W(s-\xi-t^2) dt \right] ds
\]

\[
= \frac{1}{2\pi} \int_0^\tau \frac{e^{-(k^2/4)(\tau-s)}}{\sqrt{\tau-s}} \left\{ W(0) \frac{e^{-(k^2/4)s}}{\sqrt{s}} - W(s-J) \frac{e^{-(k^2/4)s}J}{\sqrt{J}} \right\} ds
\]

\[
-2 \int_s^J e^{-(k^2/4)t^2} \left[ (\frac{k^2}{4} + \gamma) W(s-t^2) + W'(s-t^2) \right] dt ds.
\]

Therefore,

\[
L^{-1} \left[ \frac{\hat{G}(x,p)}{\lambda_1 - \lambda_2} \big|_{x=\hat{x}} \right] = \frac{W(0)}{2} e^{-(k^2/4)x} - \frac{e^{-(k^2/4)\tau}}{2\pi \sqrt{J}} \int_0^\tau \frac{e^{-(k^2/4)(\tau-s)}}{\sqrt{\tau-s}} W(s-J) ds
\]

\[
- \frac{1}{\pi} \int_0^\tau \frac{e^{-(k^2/4)(\tau-s)}}{\sqrt{\tau-s}} \int_{\sqrt{s}}^{\sqrt{\tau}} e^{-(k^2/4)t^2} \left[ (\frac{k^2}{4} + \gamma) W(s-t^2) + W'(s-t^2) \right] dt ds.
\]

E.3 The double Laplace inversion

In this appendix, we shall apply the double Laplace inversion to (11.3.37) to solve for its explicit form in the original \((\tau, l)\) space. By applying the double Laplace inversion, we obtain

\[
\hat{W}(\tau, l) = L_{p_2}^{-1} \left[ L_{p_1}^{-1} \left[ \frac{\hat{H}(x, p_1, p_2)}{\lambda_2(p_2) - \lambda_1(p_1)} \big|_{x=\hat{x}} \right] \right]
\]

\[
= L_{p_2}^{-1} \left[ L_{p_1}^{-1} \left[ \int_{-\infty}^x \int_{-\infty}^{z} e^{\lambda_2(p_2)z} e^{\lambda_1(p_1)z} \frac{1}{2\sqrt{k^2/4 + \gamma + p_2(\sqrt{k^2/4 + \gamma + p_1 + \sqrt{k^2/4 + \gamma + p_2})}} \right] \right]
\]

\[
\max(e^s - 1) ds dz + \int_{-\infty}^x \int_{z}^{+\infty} e^{\lambda_2(p_2)(z-z)} e^{\lambda_1(p_1)(\gamma + p_2)} \frac{1}{2\sqrt{k^2/4 + \gamma + p_2(\sqrt{k^2/4 + \gamma + p_1 + \sqrt{k^2/4 + \gamma + p_2})}} \right]
\]

\[
\max(e^s - 1) ds dz\right],
\]
which can be simplified as

\[
\tilde{W}(\tau, l) = \int_{-\infty}^{\bar{x}} \int_{-\infty}^{z} \frac{1}{2} e^{-(\frac{a^2}{4} + \gamma)(\tau + l) - \frac{a}{2}(\bar{x} - s)} L_p^{-1} \left[ L_p^{1-1} \left( \frac{e^{-\sqrt{p_1}(\bar{x} - \bar{z}) - \sqrt{p_2}(\bar{z} - s)}}{\sqrt{p_2}(\sqrt{p_1} + \sqrt{p_2})} \right) \right] \max(e^s - 1) ds dz
\]

\[
+ \int_{-\infty}^{\bar{x}} \int_{z}^{\infty} \frac{1}{2} e^{-(\frac{a^2}{4} + \gamma)(\tau + l) - \frac{a}{2}(\bar{x} - s)} L_p^{-1} \left[ L_p^{1-1} \left( \frac{e^{-\sqrt{p_1}(\bar{x} - \bar{z}) + \sqrt{p_2}(\bar{z} - s)}}{\sqrt{p_2}(\sqrt{p_1} + \sqrt{p_2})} \right) \right] \max(e^s - 1) ds dz.
\]

The key step now is to evaluate

\[
L_p^{1-1} \left[ L_p^{1-1} \left( \frac{e^{-a\sqrt{p_1} - b\sqrt{p_2}}}{\sqrt{p_2}(\sqrt{p_1} + \sqrt{p_2})} \right) \right], \quad a \geq 0, \quad b \geq 0.
\]

\[
L_p^{1-1} \left[ L_p^{1-1} \left( \frac{e^{-a\sqrt{p_1} - b\sqrt{p_2}}}{\sqrt{p_2}(\sqrt{p_1} + \sqrt{p_2})} \right) \right] = L_p^{1-1} \left[ e^{-b\sqrt{p_2}} L_p^{1-1} \left( \frac{1}{p_2(1 + \sqrt{p_1} / p_2)} \right) \right]
\]

\[
= L_p^{1-1} \left[ \frac{e^{-b\sqrt{p_2}}}{\sqrt{p_2}} \int_{0}^{\bar{x}} \frac{ae^{-\frac{a^2}{4}(\tau - s)}}{\sqrt{\pi(\tau - s)^2}} \frac{1}{\sqrt{\pi s}} ds \right]
\]

\[
- L_p^{1-1} \left[ \int_{0}^{\bar{x}} \frac{ae^{-\frac{a^2}{4}(\tau - s)}}{\pi \sqrt{(\tau - s)^2}} \int_{\sqrt{p_2}}^{+\infty} e^{p_2 s - b\sqrt{p_2} - w^2} dw ds \right]
\]

\[
= \frac{1}{\pi \sqrt{\pi l}} e^{-\frac{a^2}{4\pi} \bar{x}^2} \frac{a}{\pi} \int_{0}^{\bar{x}} \frac{e^{-\frac{a^2}{4}(\tau - s)}}{\sqrt{(\tau - s)^2}} \int_{\sqrt{p_2}}^{+\infty} L_p^{1-1} \left( \frac{e^{-b\sqrt{p_2} - p_2(\xi^2 - s)}}{\sqrt{\pi}} \right) d\xi ds
\]

\[
= \frac{1}{\pi \sqrt{\pi l}} e^{-\frac{a^2}{4\pi} \bar{x}^2} \frac{a}{\pi} \int_{0}^{\bar{x}} \frac{e^{-\frac{a^2}{4}(\tau - s)}}{\sqrt{(\tau - s)^2}} \int_{\sqrt{\pi}}^{+\infty} \int_{0}^{(l + s - \xi^2)^+} \frac{\delta'(\eta)}{\sqrt{\pi(l + s - \xi^2 - \eta)}} e^{-\frac{1}{4(\xi^2 - \eta)}} d\eta d\xi ds.
\]
Therefore,

\[ \bar{W}(\tau, l) = \frac{1}{2\pi \sqrt{\tau l}} \int_{-\infty}^{\infty} \int_{0}^{+\infty} e^{-\frac{k}{2}(\bar{x}-s)}(e^{s} - 1) e^{-\frac{(\bar{x}-z)^2}{4\tau}} ds dz \]

\[ - \frac{e^{-(\frac{k^2}{2\tau} + \gamma)(\tau + l)}}{2\pi} \int_{-\infty}^{\infty} \int_{0}^{+\infty} e^{-\frac{k}{2}(\bar{x}-s)}(e^{s} - 1)(\bar{x} - z) \int_{0}^{\tau} e^{-\frac{(x-z)^2}{4(\tau - w)}} \int_{0}^{\sqrt{l+w}} \int_{-\infty}^{+\infty} \int_{0}^{+\infty} \int_{0}^{l+w-\eta} \int_{0}^{\sqrt{l+w-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \delta'(\eta) \int_{0}^{\sqrt{l+w-\eta}} \int_{0}^{\sqrt{l+w-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} \int_{0}^{\frac{(x-z)^2}{4l+w-\xi^2-\eta}} di dx dz \\

By changing variables, \( II \) can be written in the form of the normal distribution function, i.e.,

\[ II = 2e^{-\frac{k}{2}(x-z)+z+(\frac{k}{2}+1)^2(l+w-\xi^2-\eta)} N(d_1) - 2e^{-\frac{k}{2}(x-z)+\frac{k^2}{4}(l+w-\xi^2-\eta)} N(d_2), \]

where

\[ d_1 = \frac{z}{\sqrt{2(l+w-\xi^2-\eta)}} + (\frac{k}{2}+1)\sqrt{2(l+w-\xi^2-\eta)}, \]

\[ d_2 = \frac{z}{\sqrt{2(l+w-\xi^2-\eta)}} + \frac{k}{2}\sqrt{2(l+w-\xi^2-\eta)}. \]
As a result,

\[
\int_{0^{-}}^{+} \delta'(\eta) \int_{\eta}^{\sqrt{1+w-\eta}} I d\xi d\eta = \frac{1}{\sqrt{1+w}} e^{-\frac{1}{2}(x-z)}(e^{z} - 1)N(\frac{z}{\theta}) - 2e^{-\frac{1}{2}(x-z)+z} \int_{\sqrt{\eta}}^{\sqrt{1+w}} e(\frac{z}{\theta}+1)^2(l+w-\xi^2)\]

\[
\{- (\frac{k}{2} + 1)^2 N(d_{1} |_{\eta=0}) + \frac{e^{-\frac{d_1^2}{2(\eta)}}}{\sqrt{2\pi}} \left[ \frac{z}{2(\sqrt{2}/(l+w-\xi^2)^3) - \frac{\sqrt{2k}}{4\sqrt{(l+w-\xi^2)}}} \right] \}
\]

\[-2e^{-\frac{k}{2}(x-z)} \int_{\sqrt{\eta}}^{\sqrt{1+w}} e^{\frac{k^2}{l+w-\xi^2}} \left\{ \frac{k^2}{4} N(d_{2} |_{\eta=0}) - \frac{e^{-\frac{d_2^2}{2(\eta)}}}{\sqrt{2\pi}} \left[ \frac{z}{2(\sqrt{2}/(l+w-\xi^2)^3) - \frac{\sqrt{2k}}{4\sqrt{(l+w-\xi^2)}}} \right] \right\}
\]

\[
= \frac{1}{\sqrt{1+w}} e^{-\frac{1}{2}(x-z)}(e^{z} - 1)N(\frac{z}{\theta}) - 2e^{-\frac{1}{2}(x-z)} \left\{ \int_{\arcsin(\sqrt{\eta})}^{\frac{\pi}{2}} f_3(\theta, w, z) d\theta + \int_{\sqrt{\eta}}^{+\infty} f_4(y, w, z) dy \right\},
\]

where

\[
f_3(\theta, w, z) = e^{z+(\frac{k}{2}+1)^2(l+w)\cos^2\theta} \left[ - (\frac{k}{2} + 1)^2 N(\tilde{d}_1) \sqrt{1+w} \cos \theta - (\frac{k}{2} + 1) e^{-\frac{d_1^2}{2\sqrt{\pi}}} \right]
\]

\[
+ e^{\frac{k^2}{l+w} \cos^2\theta} \left\{ \frac{k^2}{4} N(\tilde{d}_2) \sqrt{1+w} \cos \theta + \frac{e^{-\frac{d_2^2}{2\sqrt{\pi}}}}{4\sqrt{\pi}} \right\},
\]

\[
f_4(y, w, z) = \frac{z}{4\sqrt{\pi}(l+w)} \left[ e^{z+(\frac{k}{2}+1)^2(l+w)\sqrt{\frac{1}{y^2+1} - \frac{k^2}{4\sqrt{\pi}}} - e^{\frac{k^2}{l+w} \sqrt{\frac{1}{y^2+1} - \frac{d_4^2}{2\sqrt{\pi}}}} \right],
\]

\[
\tilde{d}_1 = \frac{z}{\sqrt{2(l+w) \cos \theta}} + (\frac{k}{2} + 1) \sqrt{\frac{2(l+w) \cos \theta},}
\]

\[
\tilde{d}_2 = \frac{z}{\sqrt{2(l+w) \cos \theta}} + k \sqrt{\frac{2(l+w) \cos \theta},}
\]

\[
d_3 = \frac{2(l+w)}{2(l+w)} + \frac{z(k+2)}{(y^2+1)} \left( \frac{k}{2} + 1 \right)^2,
\]

\[
d_4 = \frac{2(l+w)}{2(l+w)} + \frac{(l+w)k^2}{2(y^2+1)}.
\]
Therefore

\[
\tilde{W}(\tau, l) = \frac{1}{2\pi\sqrt{\tau l}} \left( \int_{-\infty}^{\infty} e^{-\frac{1}{2}(x-s)}(e^x - 1)e^{-\frac{(\tilde{x}-z)^2}{4\tau}}dsdz \right) \\
\left. - \frac{e^{-\frac{(k^2+\gamma)(\tau+l)}{\tau}}}{2\pi} \int_{-\infty}^{\infty} \int_{0}^{\bar{x}} e^{-\frac{(\tilde{x}-z)^2}{4(\tau-w)}} \frac{(\bar{x}-z)e^{-\frac{1}{2}(x-z)}(e^x - 1)N(\frac{z}{\tau})dwdz}{\sqrt{(\tau-w)^3} \sqrt{1+w}} \right) \\
\left. + \frac{e^{-\frac{(k^2+\gamma)(\tau+l)}{\tau}}}{\pi} e^{-\frac{1}{2}(x-z)} \int_{-\infty}^{\infty} \int_{0}^{\bar{x}} \frac{e^{-\frac{(x-z)^2}{4(\tau-w)}}}{\sqrt{(\tau-w)^3}} \int_{\arcsin\sqrt{\frac{w}{\tau+w}}}^{\frac{\pi}{2}} f_3(\theta, w, z)d\theta dwdz \right) \\
\left. + \frac{e^{-\frac{(k^2+\gamma)(\tau+l)}{\tau}}}{\pi} e^{-\frac{1}{2}(x-z)} \int_{-\infty}^{\infty} \int_{0}^{\bar{x}} \frac{e^{-\frac{(x-z)^2}{4(\tau-w)}}}{\sqrt{(\tau-w)^3}} \int_{\arcsin\sqrt{\frac{w}{\tau+w}}}^{\frac{\pi}{2}} f_4(y, w, z)dydwdz \right).
\]

By changing variables, the first integral can be simplified as

\[
I = \frac{\sqrt{2}}{\sqrt{\pi}} e^{-\frac{k^2(\tau+l)}{\tau}} \int_{0}^{\infty} \left\{ e^{(1+\frac{k^2}{2})\xi^2+\bar{x}^2-\frac{\bar{x}^2}{2\tau}}} - e^{\frac{k^2\tau-w}{2}} N[\frac{\bar{x} - \sqrt{2\tau}\eta}{\sqrt{2l} + \sqrt{2l(\bar{x} - \sqrt{2\tau}\eta)}}] \right\} d\eta.
\]

Denoting \( \eta = \sqrt{\frac{\tau+w}{\tau-l}}, \ \xi = \frac{\bar{x} - z}{\sqrt{2(\tau+l)}} \) and substituting them to II, we have

\[
II = -\frac{\sqrt{2}e^{-\frac{1}{2}(\tau+l)}(\tau+l)}{\pi} \int_{0}^{\frac{\sqrt{2}(\tau+l)}{\tau}} e^{-\frac{1}{2}\sqrt{2(\tau+l)}\xi-\frac{1}{2}(\bar{x} - \sqrt{2(\tau+l)}\xi - 1)(1 - N(\sqrt{\frac{l}{\tau}}))d\xi.}
\]
Denoting $\eta = \frac{x - z}{\sqrt{2(\tau - w)}}$, and substituting it to $III - IV$, we obtain

$$III = \frac{2\sqrt{2}e^{-\left(\frac{k^2}{4} + \gamma\right)(\tau + l)}}{\pi} \int_{-\infty}^{+\infty} \int_{\bar{x} - \infty}^{+\infty} e^{-\frac{\eta^2}{2} - \frac{1}{2}(\bar{x} - z)}$$

$$\int_{\text{arcsin} \sqrt{\frac{2\eta^2(\bar{x}^2 - z^2)}{2\eta^2(\bar{x}^2 + (\bar{x} - z)^2)}}}^{\eta} f_3(\theta, \tau - \frac{(\bar{x} - z)^2}{2\eta^2}, z) d\theta d\eta dz,$$

$$IV = \frac{2\sqrt{2}e^{-\left(\frac{k^2}{4} + \gamma\right)(\tau + l)}}{\pi} \int_{-\infty}^{+\infty} \int_{\bar{x} - \infty}^{+\infty} e^{-\frac{\eta^2}{2} - \frac{1}{2}(\bar{x} - z)}$$

$$\int_{\sqrt{\frac{2\eta^2(\bar{x}^2 - z^2)}{2\eta^2}}}^{+\infty} f_4(y, \tau - \frac{(\bar{x} - z)^2}{2\eta^2}, z) dy d\eta dz.$$
Bibliography


[31] A. A. Drăgulescu and V. M. Yakovenko. Probability distribution of returns in
the Heston model with stochastic volatility. *Quantitative Finance*, 2:443–453,
2002.


[33] J. D. Evans, R. Kuske, and J. B. Keller. American options on assets with


182, 1951.


volatility. *Internal Journal of Theoretical and Applied Finance*, 3:101–142,
1998.

markets with stochastic volatility*. Cambridge University Press, Cambridge,
2000.

markets with stochastic volatility*. Cambridge University Press, Cambridge,
2000.


[106] R. Vicente, C. M. de Toledo, V. B. P. Leite, and N. Caticha. Underlying
dynamics of typical fluctuations of an emerging market price index: the Heston

of the 7th European Conference on Mathematics in Industry*, pages 427–434,
1994.

1994.

model for option pricing with transaction costs. *Mathematical Finance*, 7:307–
324, 1997.

[110] M. Widdicks, P. W. Duck, A. D. Andricopoulos, and D. P. Newton. The Black-
Scholes equation revisited: Asymptotic expansions and singular perturbation.


(to appear), 2009.

[114] Z. Zhe and K.-G. Lim. A non-lattice pricing model of American options under


Publication list of the Author


- S.-P. Zhu. and W.-T. Chen. Should an American option be exercised earlier or late if volatility is not assumed to be a constant? *International Journal of Theoretical and Applied Finance (accepted)*, 2011.


2010.

tan options analytically.
  