The Hong Kong Polytechnic University

Department of Applied Mathematics

Numerical Methods for Interest Rate Derivatives

Zhou Hongjun

A thesis submitted in partial fulfilment of the requirements for the degree of Doctor of Philosophy

June 2011

CERTIFICATE OF ORIGINALITY

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of a university or other institute of higher learning, except where due acknowledgment is made in the text.

Zhou Hongjun

Abstract

It is well known that interest rate market is an important part of the financial market, and many models have been proposed to fit the market. In this research, we study numerical methods for interest rate derivatives under several models. We consider pricing American put options on zero-coupon bonds under a single factor model of short-term rate, and valuing caps under Lognormal Forward-LIBOR Model (LFM). Monte Carlo method and a novel PDE method are illustrated for pricing caps under one-factor and two-factor LFM. Also, the performance of a short rate model (CIR model) and the one-factor LFM for pricing interest rate derivatives is compared. Calibration experiments indicate that the LFM with zero correlations is closer to the real market than the CIR model, and the LFM with nonzero correlations is even better than the LFM with zero correlations in fitting the market data.

More specifically, power penalty method is used for tackling the American put options on zero-coupon bonds for the first time. We choose the CKLS short rate model for the bond option pricing, then the option value satisfies a Linear Complementarity Problem (LCP), which is solved by the power penalty approach. Valuing caplets or caps under the one-factor LFM is usually carried out and resorted to the Black's formula. However, we think that Black's formula holds with the condition that the underlying forward rates are uncorrelated. When the underlying forward rates are correlated, Monte Carlo method is illustrated for pricing caps and European options on coupon-bearing bonds and swaptions under the one-factor LFM. Based on that, we extend the Monte Carlo method for pricing caps to the two-factor LFM. Calibration of interest rate models with market data indicates that the one-factor LFM is more practicable than the CIR model. On the other hand, we observe that caps have lower prices under the one-factor LFM than under the CIR model from numerical experiments. European options on coupon-bearing bonds under these two models are found possessing similar price behavior.

Finally, we develop a PDE approach similar to that of Heston (1993) for pricing three-period caps under the one-factor LFM, and establish numerical schemes for solving the PDEs. This PDE approach is applicable when the underlying forward rates are correlated, and can be applied to evaluate caplets and caps under the one-factor LFM with stochastic volatility.

Paper Writing

During the period of study in Department of Applied Mathematics, The Hong Kong Polytechnic University as a graduate student, I attended the conference as follows.

1. H. J. Zhou, K. F. Yiu, L. K. Li, Interest Rate Derivative Pricing under the CIR Model and LFM, *The 8th International Conference on Optimization: Techniques and Applications*, December 2010, Shanghai, China.

In addition, the following papers are based on this thesis.

1. H. J. Zhou, K. F. Yiu, L. K. Li, Evaluating American Put Options on Zero-coupon Bonds by a Penalty Method, *Journal of Computational and Applied Mathematics*, Vol. 235, No. 13, 2011, pp. 3921-3931.
2. H. J. Zhou, K. F. Yiu, L. K. Li, Comparison of the LFM and the CIR Model on Interest Rate Derivative Pricing through Monte Carlo Simulation, submitted.
3. H. J. Zhou, K. F. Yiu, L. K. Li, A Novel PDE method for Cap Pricing under the LFM, submitted.

Acknowledgments

It has been a precious experience for me to study for my PhD degree in the Department of Applied Mathematics at The Hong Kong Polytechnic University. First and foremost, I wish to express my deep gratitude to my supervisors Dr. Li Leong Kwan and Dr. Yiu Ka Fai for their constant encouragement and helpful guidance. At any time of the days, they are always willing to discuss and give insightful suggestions to my research, and those interesting and fruitful discussions with them inspired me a lot.

I wish to thank Professor Wang Song, Professor Lin Yanping, Professor Chen Xiaojun and Dr. Wong Heung for their encouragement and precious suggestions. I am grateful to Dr. Ho Ceon Hok for his help. I also wish to thank Mr. An Congpei, Miss Liu Jingzhen, Mr. Meng Kaiwen, Mr. Gan Liqing, Mr. Mao Yuan, Miss He Mu, and all my fellow students for their friendship and kindly assistance. They have made my time at The Hong Kong Polytechnic University enjoyable.

I devote my deep appreciation to my parents, my sisters and my family members living in Hubei province of China. I thank them for their unconditional support and always being by my side. I owe much to my wife, Li Xiumin. I thank her for her love, support, encouragement and tolerance.

Last but not least, I sincerely thank all the staff of the Department of Applied Mathematics and Research Office for their kindness and generosity. Thanks are also due to the Research Committee of The Hong Kong Polytechnic University for offering me financial support over the last three years. I shall always treasure the amazing environment and the excellent facilities of The Hong Kong Polytechnic University.

Contents

List of Tables

List of Figures

Notations

μ, μ_i	drift rates
$W(t), B_t, Z_i(t)$	Wiener processes
$\widetilde{r}$	risk-free interest rate
κ	speed of mean-reversion in short rate process
θ	long-term interest rate
t, τ	time variables
S	stock price
$S_f(t)$	optimal exercise stock price at time t
$S_{\alpha,\beta}(t)$	swap rate at time t
r	short rate variable
R	upper bound of variable r
$F_i, F_i(t), F_i(t; T_{i-1}, T_i)$	forward rate at t for expiry T_{i-1} and maturity T_i
$\sigma, \sigma_i(t)$	volatility constant and volatility function
$T_i (i = \alpha + 1, \ldots, \beta)$	resetting dates of forward rates
τ_i	time length between T_{i-1} and T_i
$\rho_{i,j}$	correlation of the underlying forward rates F_i, F_j
$\widetilde{N}$	nominal value of cap
T^*	maturity of zero-coupon bond
T	option maturity
K, K_i	striking prices of options
X	fixed rate of swap, cap rate of caps
$\widetilde{X}$	upper bound of stock price

$V(S,t), C(S,t), V(r,t)$	option value functions
$P(t,T)$	deterministic discount factor from T to t
$D(t,T)$	stochastic discount factor from T to t
$u(x,\tau), \tilde{u}(s,t), u(r,t)$	value functions after variable transformation
$B(r,t,T^*)$	price of the zero-coupon bond with maturity T^*
$g(S,t), h(S,t), g(r,t)$	option payoff functions
$\mathrm{Cpl}(t, T_{i-1}, T_i, \tau_i, X)$	price at time t of the caplet written on $F_i(t)$
$\mathcal{L}, \mathfrak{L}, \mathscr{L}$	operators of American put option

PDE	partial differential equation
LCP	linear complementarity problem
PSOR	projected successive over relaxation method
LFM	lognormal forward-LIBOR model
LIBOR	London inter-bank offered rate
ZBC	call option on a zero-coupon bond
ZBP	put option on a zero-coupon bond
CBO	option on a coupon-bearing bond
ATM cap	at-the-money cap

Chapter 1

Introduction

1.1 Financial Models

Options are financial contracts which give the holder right, but no obligation, to buy or sell the underlying asset at a specific price at some future time. They are categorized as different kinds of options by different criteria. Roughly speaking, there are call options and put options, European options and American options, and options on various underlying assets (e.g., stocks, bonds, interest rate and others). By definition, a European option holder can only exercise the option on the maturity date, while an American option holder can exercise at any time during its valid period. Originally, options are designed for the holders to lock their costs or income on the underlying asset, which makes them as an important tool for hedging and speculating. Option pricing has attracted much attention since the seminal paper of Black and Scholes [12]. In this thesis, we focus on numerical methods for valuing American options on a single stock, American put options on zero-coupon bonds, caps and also swaptions written on forward rates.

1.1.1 Valuing American Put Options on a Single Stock

Black and Scholes [12] studied option pricing in a risk-neutral world, which, in brief, means all investments get the same rate of return. Specifically, the risk-neutral world implies the following assumptions about the market: the market is perfectly competitive, no transaction friction (i.e., no transaction cost, no tax, no bid-ask spread, no mortgage, no short limitation), no default, and no arbitrage. Generally, the underlying stock price S_t is assumed to follow a lognormal diffusion process

$$\frac{dS_t}{S_t} = \mu dt + \sigma dB_t,$$

where μ and σ are constants, and B_t is a standard Brownian motion. By no-arbitrage principle, the European put option on this stock has the value $V(S,t)$ which satisfies the Black-Scholes equation

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \widetilde{r} S \frac{\partial V}{\partial S} - \widetilde{r} V = 0, \quad (S,t) \in (0,\infty) \times (0,T),$$

together with the terminal and boundary conditions

$$V(S,T) = (K-S)^+ := \max(K-S, 0),$$

$$\lim_{S\to 0} V(S,t) = Ke^{-\widetilde{r}(T-t)}, \quad \lim_{S\to\infty} V(S,t) = 0,$$

where $\widetilde{r}$ is the risk free interest rate, σ is the volatility of the stock price, K is the strike price of the option, and T is the option maturity.

It is well known that analytic solutions can be derived for the European options and the American call options on the above stock without dividend, if $\widetilde{r}$ and σ are constants or deterministic functions of t. However, for an American put option on the above stock, there is an unknown optimal exercise price $S_f(t)$ at time t which leads to a free boundary problem for the option pricing, thus no closed-form solution exists for the option value $V(S,t)$. The optimal exercise price $S_f(t)$ means that: once the stock price S_t is equal to or less than $S_f(t)$ at t, the option holder should exercise the option immediately and $V(S,t) = K - S_t$; otherwise, the option should be held and $V(S,t)$ satisfies the following free boundary problem (PDE Problem, see Wilmott, Howison and Dewynne [89])

$$
\begin{aligned}
&\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \widetilde{r} S \frac{\partial V}{\partial S} - \widetilde{r} V = 0, \qquad S_f(t) < S < \infty, \qquad 0 \le t < T, \\
&V(S,T) = (K-S)^+, \\
&V(S_f(t),t) = K - S_f(t), \\
&\frac{\partial V}{\partial S}(S_f(t),t) = -1, \\
&\lim_{S\to\infty} V(S,t) = 0, \qquad 0 \le t \le T.
\end{aligned}
$$

Alternatively, the American put option pricing problem can be formulated as a linear complementarity problem (LCP, see [89])

$$
\begin{aligned}
&\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \widetilde{r} S \frac{\partial V}{\partial S} - \widetilde{r} V \le 0, \\
&V(S,t) \ge g(S) := (K-S)^+, \\
&(\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + \widetilde{r} S \frac{\partial V}{\partial S} - \widetilde{r} V)(V-g) = 0.
\end{aligned}
$$

a.e. for $(S,t) \in (0,\infty) \times (0,T)$.

Here, the option value $V(S,t)$ and the optimal exercise boundary (free boundary) $S_f(t)$ are both to be determined. With the presence of the moving boundary, it was widely believed that analytical formulas for $V(S,t)$ and $S_f(t)$ did not exist until Zhu [99] claimed that he got closed-form solutions in the form of infinite series by homotopy method. Different approximation methods and numerical methods have been proposed for the above PDE problem and LCP in literature. We will review those methods and suggest a new approach for the LCP, and test their congruence by numerical examples in Chapter 2.

1.1.2 Valuing American Put Options on Zero-coupon Bonds

When dealing with options on bonds, although the underlying asset is a bond, the independent variable is the stochastic interest rate. The bond price is only used in the terminal and boundary conditions. Since various interest rate models have been proposed, it is important to choose a suitable model for the bond valuation. It is known that many term structure models for the short-term interest rate $r(t)$ can be nested within the CKLS model (Chan et al [23]) as defined by

$$dr(t) = \kappa(\theta - r(t))dt + \sigma r(t)^{\gamma} dW(t), \tag{1.1.1}$$

where $W(t)$ is a Wiener process (often called standard Brownian motion) under the risk-neutral measure, κ is the speed of mean-reversion, θ is the long-term interest rate, and σ is the volatility. Here, the risk-neutral measure Q is associated with the bank-account numerairie $Ba(t) = Ba_0 \exp\left(\int_0^t r(s)ds\right)$. Generally speaking, κ, θ, σ are constants or functions of t. For simplicity, they are assumed to be constants and $\kappa\theta \neq 0$. In addition, we need to point out that the parameter γ is critical for the model to characterize the interest rate market data, and we will take the γ estimated from real data.

Under the interest rate process (1.1.1), we can derive the pricing equations for zero-coupon bonds and options on the bonds using the traditional no-arbitrage argument. Let the price at time t of the zero-coupon bond with face value E and maturity T^* be denoted by $B(r, t, T^*)$, then it satisfies the following equation as in Li and Li [59]:

$$\frac{\partial B}{\partial t} + \frac{1}{2}\sigma^2 r^{2\gamma} \frac{\partial^2 B}{\partial r^2} + \kappa(\theta - r)\frac{\partial B}{\partial r} - rB = 0, \tag{1.1.2}$$

$$B(r, T^*, T^*) = E. \tag{1.1.3}$$

Now we consider an American put option on the zero-coupon bond with exercise price K and expiry date $T(< T^*)$. Similar to the American put options on stocks, there is an optimal exercise interest rate $r^*(t)$ at time t for this option. It is the smallest value of the interest rate at which the exercise of the put option becomes optimal. The American put option value $V(r,t)$ satisfies the free boundary problem as in Chesney, Elliott and Gibson [25]:

$$V_t + \mathcal{L}V = 0, \quad V(r,t) > g(r,t), \quad 0 < r < r^*(t), \quad 0 \le t < T, \tag{1.1.4}$$

$$V(r^*(t),t) = g(r^*(t),t), \quad V_r(r^*(t),t) = g_r(r^*(t),t), \quad 0 \le t < T, \tag{1.1.5}$$

$$V(r,t) = g(r,t), \quad r > r^*(t), \quad 0 \le t \le T, \tag{1.1.6}$$

$$V(0,t) = g(0,t), \quad 0 \le t \le T, \tag{1.1.7}$$

$$V(r,T) = g(r,T), \quad r \ge 0, \tag{1.1.8}$$

where

$$g(r,t) := \max(K - B(r,t,T^*), 0),$$
$$\mathcal{L}V := \frac{1}{2}\sigma^2 r^{2\gamma} V_{rr} + \kappa(\theta - r)V_r - rV,$$

$V_t = \frac{\partial V}{\partial t}, V_r = \frac{\partial V}{\partial r}, V_{rr} = \frac{\partial^2 V}{\partial r^2}, g_r = \frac{\partial g}{\partial r}$, and $\mathcal{L}$ denotes the differential operator. Note that K should be strictly less than $B(0,T,T^*)$; otherwise, exercising the option would never be optimal.

By imposing $r \in [0,R]$ for some sufficiently large number R, the free boundary problem (1.1.4)-(1.1.8) can be formulated as a LCP as follows

$$-V_t - \mathcal{L}V \ge 0, \tag{1.1.9}$$

$$V(r,t) \ge g(r,t), \tag{1.1.10}$$

$$(-V_t - \mathcal{L}V)[V(r,t) - g(r,t)] = 0, \tag{1.1.11}$$

a.e. in $[0,R] \times [0,T]$. The corresponding boundary and terminal conditions are

$$V(0,t) = g(0,t), \quad V(R,t) = g(R,t),$$

$$V(r,T) = g(r,T), \quad r \ge 0.$$

In order to solve the LCP (1.1.9)-(1.1.11), we need to work out the bond price in (1.1.2)-(1.1.3) first by numerical method since no analytical solution have been found. After that, we apply the power penalty method to solve the problem (1.1.9)-(1.1.11).

1.1.3 Valuing Caps and Swaptions

Amongst kinds of interest rate derivative products, the most common ones are swaps, caps (floors), swaptions and their variations. Caps and swaptions written on forward rates are derivative products of floating-rate loans where interest rate is resetted periodically (e.g. every 3 months) according to the market rates. A cap gives the holder the right to stick to a specified rate if the market rate goes higher than it. A swaption is an option with the right (and no obligation) to enter an interest rate swap (IRS) at a given future time.

"Forward-LIBOR" is the shorthand for the forward London inter-bank offered rate, which is a simply-compounded forward rate and adopted in a large range of financial contracts. Before the lognormal forward-LIBOR model (LFM, see Brace, Gatarek and Musiela [16]) was proposed, Black's formula for caps (see Brigo and Mercurio [19]) has been widely used in the market practice. It was obtained by mimicking the Black-Scholes equation for stock options under some simplified and inexact assumptions on the forward rate distribution. Brigo and Mercurio stated that the introduction of the LFM provided a new derivation of the Black's formula based on rigorous forward rate dynamics, because they believed that the correlations of the underlying forward rates would not affect the payoff of caps.

Set $F(t; T_{i-1}, T_i)$ $(i = \alpha + 1, \ldots \beta)$ as the forward rate observed at time t for a period starting at T_{i-1} and ending at T_i, and write $F_i(t)$ as its shorthand. In one-factor LFM framework, the forward rate $F_i(t)$ is assumed to follow a lognormal distribution under the forward measure Q^i (which is associated with the numeraire $P(t, T_i)$, namely, the price of the bond whose maturity coincides with the maturity of the forward rate $F_i(t)$, see Brigo and Mercurio [19]), i.e.,

$$dF_i(t) = \sigma_i(t) F_i(t) dZ_i(t), \quad t \le T_{i-1}. \tag{1.1.12}$$

In this equation, $dF_i(t)$ denotes the change in the forward rate, $F_i(t)$, in the time interval dt, Z_i is a standard Brownian motion under Q^i, and $\sigma_i(t)$ is the volatility of the forward rate $F_i(t)$. The "noises" in the dynamics of different forward rates are assumed to be instantaneously correlated according to

$$dZ_i(t)\ dZ_j(t) = \rho_{i,j} dt,$$

and $\rho_{i,j}$ is called instantaneous correlation between F_i and F_j ($\rho_{i,i} = 1$ necessarily). Note that the one-factor LFM can be extended to multi-factor LFM by setting $Z_i(t)$ as a d-dimensional Brownian motion vector, $\sigma_i(t)$ as a d-dimensional vector, and $\sigma_i(t) dZ_i(t)$ as an inner product. It can be seen that $\rho_{i,j}$ is a d-dimensional matrix in the multi-factor LFM. When no specification is given, the LFM represents the one-factor LFM in this thesis.

As mentioned above, Brigo and Mercurio deem that Black's formula is an exact pricing formula for caps under the LFM. In our opinion, Black's formula holds under the LFM with the condition that the underlying forward rates are uncorrelated (i.e., $\rho_{i,j} = 0$). When the underlying forward rates are correlated (i.e., $\rho_{i,j} \neq 0$), Black's formula does not hold any more. Numerical techniques such as Monte Carlo method and a novel PDE method will be used for evaluating caps under the LFM with $\rho_{i,j} \neq 0$ in our work. To do that, we first write the dynamics of $F_i's$ under a uniform measure, which is achieved in the following proposition (see Brigo and Mercurio [19]):

Proposition 1.1.1 *Under the assumption of lognormal distribution for each forward rate, we obtain that the dynamics of F_i under the forward measure Q^k in the* *three cases $k < i$, $k = i$ and $k > i$ are, respectively,*

$$
\begin{aligned}
&k < i \quad t \le T_k: \quad dF_i(t) = \sigma_i(t)F_i(t)\sum_{j=k+1}^{i}\tfrac{\rho_{i,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_jF_j(t)}dt + \sigma_i(t)F_i(t)dZ_i(t),\\
&k = i, \quad t \le T_{i-1}: \quad dF_i(t) = \sigma_i(t)F_i(t)dZ_i(t), \qquad (1.1.13)\\
&k > i, \quad t \le T_{i-1}: \quad dF_i(t) = -\sigma_i(t)F_i(t)\sum_{j=i+1}^{k}\tfrac{\rho_{i,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_jF_j(t)}dt + \sigma_i(t)F_i(t)dZ_i(t),
\end{aligned}
$$

where, $Z(t)$ is a Brownian motion vector under Q^k, $Z_i(t)$ is the i-th component of $Z(t)$ satisfying $dZ_i(t)dZ_j(t) = \rho_{i,j}dt$. All of the above equations admit a unique strong solution if the coefficients $\sigma_i(t)$ are bounded.

The details of our methods will be illustrated in Chapters 5 and 6.

In fact, Brigo and Mercurio have shown that a cap is equivalent to a portfolio of European put options on zero-coupon bonds. This equivalence leads to explicit formulas for cap price under those analytically tractable short-rate models. Thus, caps can be priced under the CIR model analytically. The well-known CIR model (Cox, Ingersoll and Ross [29]) for short rate $r(t)$ is given as

$$dr(t) = \kappa(\theta - r(t))dt + \sigma r(t)^{1/2}dW(t), r(0) = r_0,$$

with $r_0, \kappa, \theta, \sigma$ positive constants.

Since closed-form expressions exist for cap price under the LFM with zero correlations and the CIR model, these two models have been widely used for interest rate derivative pricing, while their performance is seldom compared. We would compare their performance in terms of pricing caps by calibrating them to market data, which will be done in Chapter 4. Moreover, we will calibrate the LFM with nonzero correlations and the LFM with zero correlations to market data, and find that the former is better than the latter in terms of fitting capability.

1.2 Literature Review

The earliest analysis on the pricing of the American option is by McKean [62]. There the problem of pricing the American option is transformed into a free boundary problem. Solving the latter, McKean writes the American option price explicitly up to knowing a certain function (the free boundary). "This work was taken further by van Moerbeke [65], who studied properties of the optimal stopping boundary. Although the American option problem was treated as an optimal stopping problem by McKean and van Moerbeke, a financial justification using hedging arguments was given only later by Bensoussan [11], Karatzas [58] and Jacka [51]" (see Myneni [68]). Here, we first survey the literature on the American options dealing with the free boundary problem (PDE problem), then we summarize those work addressing the LCP of American options.

For solving the PDE problem of the American put option on a single stock, Brennan and Schwartz [17] proposed the Brennan-Schwartz algorithm, which handles the early exercise condition by first advancing the discrete solution over a time-step without taking the constraint into account, then applying the constraint explicitly. This algorithm proves to be convergent in Jaillet, Lamberton and Lapeyre [52], nevertheless, it can only be regarded as a first order numerical approximation in time. Johnson [55] provided heuristic techniques for valuing American put options on non-dividend-paying stocks. Although these techniques are fast computationally, the accuracy of heuristic techniques is frequently sensitive to the parameter range used in the option-pricing problem. An alternative approximation method was introduced by Geske and Johnson [36] for the PDE problem. Their compound-option approximation method offers the advantages of being intuitively appealing and easily amenable to comparative-statics analysis. However, this approach is still not inexpensive since it requires the evaluation of cumulative bivariate, trivariate, and sometimes higher order multivariate normal density functions. Barone-Adesi and Whaley [10] also developed an approximation method for the PDE problem based on MacMillan's [61] quadratic approximation of the American put option valuation problem. This method is fast computationally and can be programmed on a hand-held calculator. Allegretto, Barone-Adesi and Elliott [1] modified this method to determine the critical price (the free boundary) relatively quickly. Bunch and Johnson [22] expressed the American put price as an integral involving the first-passage probability of the stock price hitting the critical stock price, by which they provided alternate derivations of expressions for put prices in the perpetual-put case. They also got a critical price expression which is exact (and approximations that can be used to evaluate it easily) for the finite-lived case.

Recently, Han and Wu [44] studied the PDE problem of the American call option on a stock with continuous dividend. They use the fundamental solution of the heat equation to derive an exact artificial boundary condition, which narrows the computation domain and improves the efficiency of their numerical method. This trick is termed as the domain decomposition technique. Resorting to the specific property of the optimal exercise boundary, they obtain the exercise boundary in the process of finding the option value. The domain decomposition method was also adopted to compute American option prices and Greeks under the CEV model in Wong and Zhao [90]. Tangman et al [80] found that the domain decomposition technique could be used for valuing American put options as well. Furthermore, they solved the moving boundary problem of $V = V_{American} - V_{European}$ using an optimal compact algorithm, and got numerical solutions with high accuracy. In addition, Tangman et al [81] solved the PDE problem with approximate boundary conditions on a truncated domain. They employ front-fixing transformation ($x = log\frac{S}{S_f(\tau)}$) and fourth-order finite difference schemes and get a second order convergence rate in terms of numerical results. Besides, Zhu [100] applied the homotopy analysis method to the PDE problem of American put options, and claimed that he got closed-form solutions for the option value and optimal exercise boundary. The solutions are constructed with Taylor's series expansion for the option value and optimal exercise boundary in terms of a parameter p, which is allowed to vary in the domain $[0, 1]$. However, each term of the series solution has a very complicated coefficient involving several double integrals hard to calculate, and about 30 terms are needed to generate convergent numerical solutions. It is computationally expensive.

On the other hand, there is numerous literature focusing on the American option pricing problem through solving the LCP, which can be equivalently regarded as a variational inequality problem (see Jaillet, Lamberton and Lapeyre [52]). Projected successive over-relaxation method (PSOR) (see Wilmott, Howison and Dewynne [89], Oosterlee [71]), linear programming method (see Dempster and Hutton [30, 31]), Lagrangian method (see Vazquez [85]), and penalty methods [34, 43, 73, 87, 101], etc, have been developed to solve the LCP.

Zvan and Forsyth [101, 34] first developed the linear (l_1) and quadratic (l_2) penalty methods. They have proved that the solution to the penalized problem converges to that to the LCP, and the convergence rates of the l_1 and l_2 penalty methods are of order $\mathcal{O}(\lambda^{-1/2})$ and $\mathcal{O}(\lambda^{-1/4})$, respectively, where λ is the penalty parameter. The convergence rates of these methods imply that large penalty parameters should be attained in order to obtain a desirable accuracy level of the solutions. However, too large penalty parameters would cause computational difficulties, which motivates the appearance of the lower order penalty method.

Wang, Yang and Teo [87] proposed the lower order penalty method for the LCP. They add a power penalty term ($[\cdot]_+^k, 0 < k < 1$) to the equivalent variational inequality and transform it into a nonlinear parabolic partial differential equation, which is then discretized by the fitted finite volume method and solved with Newton's method. They show that the solution of the penalized equation converges to that of the LCP with $\mathcal{O}(\lambda^{-1/2k})$. The convergence analysis of the numerical scheme for the penalized problem is given in Angermann and Wang [8].

When pricing bonds and American put options on zero-coupon bonds, the CKLS short rate model

$$dr(t) = \kappa(\theta - r(t))dt + \sigma r(t)^{\gamma} dW(t) \tag{1.2.14}$$

is often adopted to characterize the interest rate market. In the case $\gamma = 0$ or 0.5, the above model is the Vasicek model [84] or the CIR model [29], respectively, under both of which analytic expressions for zero-coupon bond price exist. The short-term rate r can be negative in the Vasicek model, while it is non-negative in the CIR model. Both models are pioneering works on modeling short-term rate. Chan et al [23] claimed that the CKLS model with $\gamma \geq 1$ captures the dynamics of the short-term interest rate better than the model with $\gamma < 1$ when fitting the monthly U.S. treasury bill yield from June 1964 to December 1989. Nowman and Sorwar [70] found that when fitting Canada, Hong Kong, and United States currency rates during February 1981 and December 1997, γ varied from 0.0076 to 1.2260. In particular, γ was 1.1122 and 1.2660, respectively, for 1- and 3-month United States rates. It is easy to understand that only when the interest rate model is accurate enough for fitting market rates, the pricing of derivatives written on interest rate can be trusted. Therefore, the parameter γ is critical in the pricing of such derivatives. It will be taken as 0.3912, 1.1122 and 1.2660 in this study.

On condition that $\gamma = 0.5$, zero-coupon bonds can be priced analytically, and the pricing of American put options have been studied in Allegretto, Lin and Yang [4], where both the finite volume method and the finite element method are used for setting up numerical schemes and whose stability and convergence are also analyzed. The Brennan-Schwartz algorithm is used for working out the option value there. For $\gamma = 1$, there is no analytic solution for the zero-coupon bond price, which will be used in the terminal and boundary conditions when valuing the options on the bond. It needs to be computed numerically. American put options on zero-coupon bonds under the assumption $\gamma = 1$ are tackled in Li and Li [59], also with the Brennan-Schwartz algorithm for the option value. When $\gamma \geq 1$, European call options on zero-coupon bonds are investigated in Sorwar, Barone-Adesi and Allgeretto [79]. While $\gamma = 0$ and κ and σ are deterministic functions of time, the equation (1.1.1) is Hull-White short rate model (see Hull and White [47]), under which American interest rate derivatives are evaluated by the domain decomposition method (see Wong and Zhao [91]). Basically, American put options on bonds are all priced by the PSOR method or the Brennan-Schwartz algorithm in recent researches. We propose to use power penalty approach for this bond option problem in Zhou, Yiu and Li [98].

After the seminal paper of Merton [63], the short-term interest rate models, such as the Vasicek model [84], the CIR model [29] and the Hull-White model [47], which describe dynamics for continuously-compounded instantaneous rate have been widely used in financial mathematics, because they imply analytical expressions for the price of some derivatives. Chan et al [23] reviewed several short rate models and compared their usefulness in valuing interest rate contingent claims. Calibrating to interest rate market and pricing derivatives with short rate models are also studied in Nowman and Sorwar [70], Sorwar, Barone-Adesi, and Allegretto [79], and Chen and Scott [24]. In Hull [48], the Vasicek model and the CIR model which contain time-independent drift and volatility terms are called equilibrium models. Whilst the Ho-Lee model [49], the Hull-White model [47] and the Black-Karasinski [13] model which take time-dependent drift terms are categorized as no-arbitrage models.

However, the dominant status of short rate models had changed a lot after the LFM (see Brace,Gatarek and Musiela [16], Jamshidian [53], Miltersen, Sandmann and Sondermann [64]) was introduced. The LFM aims to describe the LIBOR market. It is popular with market participants because the LIBOR is adopted in many financial contracts. In some way, it is closer to real interest rate market than the short rate models in that it gives forward rate dynamics period by period. On the contrary, short rate models usually give invariable dynamics of instantaneous rate despite of the length of the period. In the LFM, each forward rate follows a lognormal distribution (also called geometric Brownian motion) under its corresponding forward measure, and the dynamics would change if it were under another measure.

The LFM is flexible, supports multiple factors and rich volatility structures, and is tractable enough to allow fast calibration to market-quoted caps. Indeed, the LFM with zero correlations evaluates caps with Black's formula, which is the standard formula employed in the cap market. However, in its original form, the LFM does not incorporate the observable phenomenon that the implied volatilities of quoted caps are strike-dependent. It only generates flat implied volatility curves, whereas the implied volatility curves observed in the LIBOR markets often have the shape of a skew (i.e., cap volatilities decrease with strike) or smile. Here, "smile" means that both in-the-money and out-of-the-money options are traded at higher implied volatilities than at-the-money options.

There have been several extensions of the LFM that incorporate the volatility skews and smiles. Andersen and Andreasen [6] modify the assumption of lognormal diffusion to allow for local forward rate volatilities that have arbitrary dependence on the forward rate itself. They adopt constant elasticity variance (CEV) processes, which can generate monotonic implied volatility curves. Glasserman and Kou [38] and Glasserman and Merrener [39] extend the LIBOR market model to take jumps of forward rates into account. With such a jump-diffusion model, one can manipulate the slope and the curvature of a skewed smile through changing jump intensity and jump sizes. Andersen and Brotherton-Ratcliffe [7] and Wu and Zhang [93] consider the LFM with stochastic volatility process, which effectively generates different shape of implied volatility curves. This extension of the LFM is obtained in the spirit of the Heston's model [46] for equity options. In fact, empirical evidence supports the qualitative basis of the LFM with stochastic volatility, as volatilities in interest rate option market clearly possess a random component beyond that implied by the local forward rate volatility approach. In addition, we would point out that the forward rate process and the stochastic volatility process are supposed to be independent in Andersen and Brotherton-Ratcliffe [7], while they are assumed to be correlated in Wu and Zhang [93].

Under the above extensions of the LFM, i.e., the local volatility model, the jump-diffusion model and the stochastic volatility model, we have to resort to PDE method (see Andersen and Andreasen [6], Andersen and Brotherton-Ratcliffe [7], and Wu and Zhang [93]) or Monte Carlo simulation method (see Brigo and Mercurio [19] and Joshi and Rebonato [56]) to valuate caps and swaptions. Monte Carlo method is also used in the numerical tests of the LFM approximations in Brigo and Mercurio [19]. In the simulation of forward rates under the LFM and a uniform measure, note that those dynamics contain state-dependent drift terms and do not lead to distributionally known processes, we need to discretize the dynamics of forward rates. The log-Euler approach working with the logarithm of forward rates through Euler discretization is adopted in Brigo and Mercurio [19]. Recently, Denson and Joshi [32] derived a pathwise adjoint method by combining the predictor-corrector drift approximation and Monte Carlo simulation of the displaced-diffusion LIBOR market model (an extension of the LFM). The method was then used for computing Greeks (i.e., the 1st or 2nd derivatives of option price with respect to time variable or underlying asset price) of caplets and lockout Bermudan swaptions. There has been a large amount of work focusing on finding fast and accurate drift approximations to facilitate the Monte Carlo simulation under the LFM (see Glasserman and Zhao [40], Hunnter et al [50], and Joshi and Stacey [57]). However, All the literature mentioned here simulates forward rates by tackling the dynamics of the logarithm of forward rates. In contrast, we will deal with the dynamics of forward rates directly in this thesis. The details of our method will be demonstrated in Section 5.1.

As we have known, calibration of the LFM generally lays emphasis on fast fitting the volatility skew or smile of the LIBOR market, with or without taking parametric correlation matrices into account. For instance, Brigo and Mercurio [18] consider several parametric assumptions for the instantaneous correlations and provide many cases of calibration to caplets volatilities and swaption volatilities; Wu [92] develops the method for non-parametric calibration of the LFM to the prices of at-the-money caps/floors and swaptions, and to the historic correlations of the LIBOR; Basing on an extension of the stochastic volatility LFM, Piterbarg [72] derives European swaption approximation formulas that allow calibration of the model to European swaptions across different expiries, maturities and strikes. In our study, we implement the calibration of the LFM only for verifying our standpoint, i.e., the LFM with $\rho_{i,j} \neq 0$ fits the market data of caps better than the LFM with $\rho_{i,j} = 0$. We will not dig into the practical shape of the volatility curve or the efficiency of calibration algorithm.

1.3 Outline

In this thesis, we firstly review different methods for solving American vanilla options on a single stock, and testify their congruence by numerical experiments. Then we discuss numerical methods for interest rate derivatives under different interest rate models. We evaluate American put options on zero-coupon bonds under the CKLS model of short-term rate, and price forward rate derivatives under the LFM. Monte Carlo method and a novel PDE method are developed for pricing caps under the one-factor LFM. Moreover, Monte Carlo method is extended to two-factor LFM. In view of the fact that Black's formula is popular with financial practitioners, we test its effectiveness by calibrating the LFM with $\rho_{i,j} = 0$ to the market data of caps. Meanwhile, we calibrate the CIR model to the same market data, and the results indicate that the LFM with $\rho_{i,j} = 0$ is closer to the real market than the CIR model. Furthermore, we let $\rho_{i,j} \neq 0$ and set it as an exogenous matrix or a parametric correlation matrix, then implement the calibration of the LFM again. As expected, we see that the LFM with $\rho_{i,j} \neq 0$ fits the market data better than the LFM with $\rho_{i,j} = 0$. Moreover, we investigate the valuation of European options on coupon-bearing bonds and swaptions under the LFM by Monte Carlo method. Lastly, we propose a novel PDE method for pricing caps under the one-factor LFM.

In our discussion about forward rate derivatives, we concentrate on the original LFM, namely, the extensions of the LFM are not considered. We propose some standpoints which are contrary to the traditional view: Black's formula for caps holds under the LFM only when $\rho_{i,j} = 0$. When $\rho_{i,j} \neq 0$ (which is more practical), we have to resort to numerical methods such as Monte Carlo simulation method or a PDE method to valuate caps. In the Monte Carlo simulation under the LFM, we deal with the dynamics of forward rates directly other than the dynamics of the logarithm of forward rates. In addition, we will apply Monte Carlo method to compute cap prices under the LFM in our calibration process, and implement the novel calibration of the LFM ($\rho_{i,j} \neq 0$) to market data of cap prices. Our calibration is different from that in the literature, e.g., Brigo and Mercurio [19], where the authors use the data of swaption prices and swaption volatilities, and implement their calibration through closed-form approximate formula of "Black-like swaption volatility". While we use the data of cap prices and implement the calibration by Monte Carlo method and downhill simplex method in multidimensions (see Press [74]).

In Chapter 2, we survey the numerical methods for American vanilla options on a single stock. We have illustrated that American vanilla option pricing problem can be formulated as a free boundary PDE problem, or an equivalent LCP in Section 1.1.1. We first discuss the domain decomposition technique for solving the PDE problem. Subsequently, we review the power penalty approach (combined with fitted finite volume discretization) and the PSOR method (combined with the finite difference/finite element discretization) for solving the LCP. We propose that the PSOR method combined with the fitted finite volume discretization also works for the LCP. Next, stochastic simulation methods such as lattice method and Monte Carlo method for option valuation are summarized. At last, we examine the congruence of these mentioned methods through numerical tests.

In Chapter 3, the power penalty method is applied to value American put options on zero-coupon bonds for the first time, by which the discretized linear complementarity problem of the option value is transformed into a nonlinear PDE by adding a power penalty term. The convergence of the penalized solution to the real solution is guaranteed by the results of Wang and Yang [88]. A numerical scheme for the penalized problem is established by the finite volume method, and the corresponding stability and convergence are discussed. In our study, the parameter γ of CKLS model varies from 0.3912 to 1.2660, which is produced by fitting the model to actual market rates (see Nowman and Sorwar [70]). We take the results computed by the Brennan-Schwartz algorithm as benchmark, and find that our penalty approach provides both the option value and optimal exercise interest rate with similar accuracy. Numerical examples are also presented.

In Chapter 4, we compare the performance of the LFM and CIR model for pricing interest rate derivatives. We first introduce the essential concepts of the LFM and the Black's cap pricing formula, then explain the explicit formula for cap pricing under the CIR model. Subsequently, calibration of the LFM with $\rho_{i,j} = 0$ and the CIR model is carried out using the market data of caps. Next, We calibrate the LFM with an exogenous correlation matrix (i.e., $\rho_{i,j} = 1 - 0.025 * |i - j|$) and the LFM with a parametric correlation matrix to the market data of caps. It indicates that the LFM with $\rho_{i,j} = 1 - 0.025 * |i - j|$ is better than the LFM with $\rho_{i,j} = 0$, and the LFM with parametric correlations is even better than the LFM with $\rho_{i,j} = 1 - 0.025 * |i - j|$ in fitting the market data. Besides, we compare the prices of caps under the LFM ($\rho_{i,j} = 1$) with those under the CIR model, and find that caps have lower prices under the former model than under the latter model. We also compare the prices of European options on coupon-bearing bonds under these two models, and find similar price behavior.

In Chapter 5, Monte Carlo method is introduced for pricing interest rate derivatives under the LFM with $\rho_{i,j} \neq 0$, because no closed-form solutions exist for the price of derivatives such as caps and swaptions in this case. We first introduce forward rate path generation under the one-factor LFM, then illustrate the Monte Carlo procedures for pricing caps, European options on coupon-bearing bonds and swaptions under the one-factor LFM. The authors evaluate swaptions starting with modeling dynamics of swap rate (e.g. lognormal swap rate model), or starting with modeling forward rate dynamics to simulate swap rate in Brigo and Mercurio [19]. We will circumvent swap rate and discuss swaption pricing under the LFM, because swaptions can be defined in terms of forward rates. Lastly, we extend the Monte Carlo method to the two-factor LFM and apply it to valuate caps.

In Chapter 6, we develop a novel PDE approach for pricing caps under the LFM with $\rho_{i,j} \neq 0$. Our method is superior to the Monte Carlo method in computation time. We first write the dynamics of the underlying forward rates under a uniform measure, then introduce the PDEs governing the payoff functions of the caplets which constitute the cap. With that, we establish numerical schemes for solving the PDEs. Discounting the numerically obtained payoff functions back to the initial time, we can get the cap price. Empirical tests of three-period caps are implemented to examine the effectiveness of our method. This PDE approach can also be used for valuations of caplets and caps under the LFM with stochastic volatility.

Finally, we conclude this thesis and state some directions for future research in Chapter 7.

Chapter 2

Overview of Numerical Methods for American Vanilla Options on a Single Stock

As noted in Section 1.1, the problem of American option valuation can be formulated as a PDE problem with free boundary or an equivalent LCP. We have reviewed the literature dealing with the PDE problem. On the other hand, the PSOR method, the linear programming method, the Lagrangian method, and penalty methods, etc, have been developed to solve the LCP of American put option valuation. Since the differential operator in the LCP is degenerated at $S = 0$, variable transformation like $t = T - \frac{2\tau}{\sigma^2}, S = Ke^x, V(S,t) = Ke^{\alpha x+\beta\tau}u(x,\tau)$ is often applied to reformulate the Black-Scholes equation as the heat equation in order to facilitate numerical methods. However, this also causes a problem: the interval of the space variable, $S \in [0,\infty)$, is transformed to $x \in (-\infty,\infty)$, and truncation needs to be made on both the left boundary and the right boundary of x in contrast with only the right-hand side truncation of S before the variable transformation. It influences the accuracy of the numerical solution.

A number of numerical techniques were explored for the LCP after the above variable transformation. Wilmott, Howison and Dewynne [89] used the finite difference method to discretize the problem on a truncated domain, and solved the discrete system by the PSOR method. Allegretto, Lin and Yang [2] deduced an exact boundary condition on the left bound and applied the finite element method to discretize the problem, then they solved it through the Brennan-Schwartz algorithm. Another contribution of them was to establish the error estimate of their numerical method in Allegretto, Lin and Yang [3]. Although the Brennan-Schwartz algorithm works for solving the LCP (see Jaillet, Lamberton and Lapeyre [52]), it has a first order approximation in time variable. The PSOR method is robust for many kinds of American option pricing problems, however, its convergence rate depends on the relaxation parameter and it exhibits exponential solution-time behavior as the number of space discretization points increases. With this background, penalty methods were devised to deal with the original LCP of the American put option before the variable transformation (see [34, 43, 73, 87, 101]).

Besides those mentioned numerical methods for the PDE problem or the LCP of American put option valuation, another important class of approaches for option pricing are the stochastic simulation methods based on risk-neutral probability. The Monte Carlo simulation method [14, 15, 35, 60, 82] and the tree (lattice) method [28, 20, 21] are examples of stochastic methods. In fact, Jaillet, Lamberton, Lapeyre [52] have proved that the maximized expectation of the discounted payoff function of the American put option, i.e., $\sup\limits_{\tau\in[t,T]} \mathbb{E}^Q[e^{-\tilde{r}(\tau-t)}g(\tau, S_\tau)]$, is the solution of the LCP, where Q is the risk-neutral measure. The idea of the Monte Carlo simulation method for American option pricing is to compute the expectation in the formula through simulating price paths for the underlying asset. The main advantage of the Monte Carlo method is its ease to handle multiple state variables and path dependencies, while its main drawback is the demand for a large number of simulation trials to achieve high accuracy. Some enhanced forms of the Monte Carlo method for asset pricing have been proposed in Broadie and Glasserman [20, 21].

The lattice method was developed by Cox, Ross, and Rubinstein in [28], where the underlying asset price movement is simulated by a discrete random walk which converges to a continuous lognormal diffusion process as the time interval between successive steps tends to zero. The lattice method gives the option price independent of the expected rate of return of the asset price, which is consistent with the risk neutrality argument.

Since the American option pricing problem can be cast in the framework of a stochastic dynamic programming problem for discrete time, there are many algorithms attempting to mimic the basic backward induction algorithm of stochastic dynamic programming (see, e.g., [82, 41, 42]). Broadie and Glasserman proposed two algorithms (see [20, 21]): one is based on a multi-branch simulation tree, and the other one is based on a stochastic mesh. Their second algorithm provides an upper bound and a lower bound estimator for the option price and both estimators converge asymptotically to the true price. Longstaff and Schwartz [60] proposed an algorithm which first uses the least square method to estimate the conditional expected payoff of continuing holding the option at each time node, and then compare the holding value with the exercising value of the option. This approach is readily applicable in path-dependent and multi-factor situations where traditional finite difference techniques can not be used. The above noted literature mainly focused on the pricing problem of American options on a single asset. When it comes to multi-asset cases, they only give numerical results of the American call option on the maximum of five assets. To the best of our knowledge, there are few results about other kinds of American options on multiple assets.

In this chapter, we first review the domain decomposition technique, which shortens the solution time remarkably, for the PDE problem of American vanilla options in Section 2.1. Next, we discuss the penalty approach and the PSOR method for solving the LCP of American vanilla put options in Section 2.2. In particular, the PSOR method combined with the fitted finite volume discretization is proposed for the LCP. The stochastic simulation methods are summarized in Section 2.3. Numerical tests for examining the congruence of these mentioned methods are implemented in Section 2.4. Lastly, we present a brief summary in Section 2.5.

2.1 Domain Decomposition Technique for the PDE Problem of American Options

2.1.1 American Call Option on a Stock with Continuous Dividend

In paper [44], Han and Wu investigate the American call option on a stock with dividend paying. They apply domain decomposition technique to the PDE problem after variable transformation, i.e., separating the domain to two sets: one is infinite on which the transformed problem (which is a standard heat equation) has an analytic solution, and the other is a narrow domain with moving boundary. They find an accurate boundary condition on the separating line, which is actually an equation about the unknown value function and its partial derivatives. With this boundary condition, smaller computational domain and more accurate solution are obtained. For the moving boundary, they derive its specific property and use that to locate it. The finite difference discretization is applied to the equation over the narrow domain. Numerical results are illustrated and compared with those obtained by standard finite difference methods. The ideas and details of the domain decomposition technique are summarized as follows.

Let $C(S,t)$ denote the value of the American call option on a stock with continuous dividend, and δ denotes the dividend yield, then $C(S,t)$ is the solution of the following free boundary problem

$$\frac{\partial C}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 C}{\partial S^2} + (\widetilde{r} - \delta) S \frac{\partial C}{\partial S} - \widetilde{r} C = 0, \tag{2.1.1}$$

$$0 < S < S_f(t), \quad 0 \le t < T,$$

$$C(S,T) = h(S), \qquad 0 \le S \le S_f(T), \tag{2.1.2}$$

$$C(S_f(t),t) = h(S_f(t)), \qquad \frac{\partial C}{\partial S}(S_f(t),t) = 1, \qquad 0 \le t \le T, \tag{2.1.3}$$

$$C(S,t) = 0 \quad as \quad S \to 0, \qquad 0 \le t \le T, \tag{2.1.4}$$

where $h(S) = (S-K)^+$, $y^+ := \max(y,0)$, $K > 0$, and $S_f(T) = \max(K, \widetilde{r}K/\delta)$.

Then, we introduce the transformation of variables

$$t = T - \frac{2\tau}{\sigma^2}, \quad S = Ke^x, \quad C(S,\tau) = Ke^{\alpha x + \beta\tau} u(x,\tau),$$

where $\alpha = -\frac{\zeta-1}{2}, \beta = -\frac{(\zeta-1)^2}{4} - 2\widetilde{r}/\sigma^2, \zeta = 2(\widetilde{r}-\delta)/\sigma^2$. Thus, the problem (2.1.1)-(2.1.4) is equivalent to

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < x_f(\tau), \quad 0 \le \tau < \tau^*, \tag{2.1.5}$$

$$u(x,0) = h(x,0), \quad -\infty < x \le x_f(0), \tag{2.1.6}$$

$$u(x_f(\tau),\tau) = h(x_f(\tau),\tau), \quad 0 \le \tau \le \tau^*, \tag{2.1.7}$$

$$\alpha u(x_f(\tau),\tau) + \frac{\partial u(x_f(\tau),\tau)}{\partial x} = e^{(1-\alpha)x_f(\tau) - \beta\tau}, \quad 0 \le \tau \le \tau^*, \tag{2.1.8}$$

$$u(x,\tau) \to 0 \quad as \quad x \to -\infty, \quad 0 \le \tau \le \tau^*, \tag{2.1.9}$$

where $h(x,\tau) = e^{-\alpha x - \beta\tau}(e^x - 1)^+$, $x_f(\tau) = ln(S_f(\tau)/K)$, $\tau^* = \sigma^2 T/2$, and $x_f(\tau) > 0$ for $\tau > 0$.

Define the unknown domain $\bar{\Omega}$

$$\bar{\Omega} = \{(x,\tau) \,|\, -\infty < x < x_f(\tau), 0 < \tau \le \tau^*\}.$$

Let a be a negative number, then we have an artificial boundary $\Gamma_a = \{(x,\tau)|x = a, 0 < \tau \le \tau^*\}$, which divides the domain $\bar{\Omega}$ into two parts, i.e., the bounded part Ω_i and the unbounded part Ω_e:

$$\Omega_i = \{(x,\tau)\,|\, a \le x < x_f(\tau), 0 < \tau \le \tau^*\},$$

$$\Omega_e = \{(x,\tau)\,|\, -\infty < x < a, 0 < \tau \le \tau^*\}.$$

On Ω_e, the solution of (2.1.5)-(2.1.9), $u(x,\tau)$, satisfies

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, \quad -\infty < x < a, \quad 0 \le \tau < \tau^*, \tag{2.1.10}$$

$$u(x,0) = 0, \quad -\infty < x < a. \tag{2.1.11}$$

If the value of $u(x,\tau)$ on the boundary Γ_a is given as

$$u(a,\tau) = \phi(\tau) \tag{2.1.12}$$

with $\phi(0) = 0$, then the solution of (2.1.10)-(2.1.12) is

$$u(x,\tau) = \frac{-(x-a)}{2\sqrt{\pi}} \int_0^\tau e^{\frac{-(x-a)^2}{4(\tau-\lambda)}} \frac{\phi(\lambda)}{(\tau-\lambda)^{\frac{3}{2}}} d\lambda. \tag{2.1.13}$$

Now, introduce a new variable $\xi = \frac{x-a}{2\sqrt{\tau-\lambda}}$, we have

$$u(x,\tau) = \frac{2}{\sqrt{\pi}} \int_{-\infty}^{\frac{(x-a)}{2\sqrt{\tau}}} \phi\left(\tau - \frac{(x-a)^2}{4\xi^2}\right) e^{-\xi^2} d\xi, \tag{2.1.14}$$

$$\left.\frac{\partial u}{\partial x}\right|_{x=a} = \frac{1}{\sqrt{\pi}} \int_0^\tau \frac{\partial u(a,\lambda)}{\partial \lambda} \frac{1}{\sqrt{\tau-\lambda}} d\lambda. \tag{2.1.15}$$

Equation (2.1.15) is derived from (2.1.14). It is an exact boundary condition satisfied by $u(x,\tau)$ on the artificial boundary Γ_a. In this way, the moving boundary problem (2.1.5)-(2.1.9) in the unbounded domain $\bar{\Omega}$ is reduced to the following problem in a bounded domain Ω_i

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, \quad a \le x < x_f(\tau), \quad 0 \le \tau < \tau^*, \tag{2.1.16}$$

$$u(x,0) = h(x,0), \quad a \le x \le x_f(0), \tag{2.1.17}$$

$$u(x_f(\tau),\tau) = h(x_f(\tau),\tau), \quad 0 \le \tau \le \tau^*, \tag{2.1.18}$$

$$e^{(\alpha-1)x_f(\tau)+\beta\tau}\left[\alpha u(x_f(\tau),\tau) + \frac{\partial u(x_f(\tau),\tau)}{\partial x}\right] = 1, \quad 0 \le \tau \le \tau^*, \tag{2.1.19}$$

$$\left.\frac{\partial u}{\partial x}\right|_{x=a} = \frac{1}{\sqrt{\pi}}\int_0^\tau \frac{\partial u(a,\lambda)}{\partial \lambda}\frac{1}{\sqrt{\tau-\lambda}}d\lambda. \tag{2.1.20}$$

Hence, if $\{u^*(x,\tau), x_f^*(\tau)\}$ is the solution of the problem (2.1.16)-(2.1.20), we let

$$u(x,\tau) = \begin{cases} u^*(x,\tau), & a \le x < x_f^*(\tau), \quad 0 \le \tau < \tau^*, \\ \frac{-(x-a)}{2\sqrt{\pi}}\int_0^\tau e^{\frac{-(x-a)^2}{4(\tau-\lambda)}}\frac{u^*(a,\lambda)}{(\tau-\lambda)^{\frac{3}{2}}}d\lambda, & x < a, \quad 0 \le \tau < \tau^*, \end{cases}$$

and $x_f(\tau) = x_f^*(\tau)$, then $\{u(x,\tau), x_f(\tau)\}$ is the solution of the problem (2.1.5)-(2.1.9).

In [44], they use the Crank-Nicolson finite difference scheme to compute the solution of the problem (2.1.16)-(2.1.20). For any given τ, the boundary point $x_f(\tau)$ is the only point satisfying the equation (2.1.16) and the condition (2.1.19). Furthermore, by the strong maximum principle of the heat equation, they deduce that $u(x,\tau)$ has the following property.

Proposition 2.1.1 *For any τ, there exists $x^*(\tau)$ such that*

$$u(x,\tau) \ge h(x,\tau), \quad \textit{when} \quad x \le x_f(\tau),$$

$$u(x,\tau) < h(x,\tau), \quad \textit{when} \quad x_f(\tau) < x < x^*(\tau),$$

$$u(x,\tau) = h(x,\tau), \quad \textit{when} \quad x \ge x^*(\tau).$$

In numerical experiments, if the boundary condition $u(x,\tau) = h(x,\tau)$ is given at $x > x_f(\tau)$, $u(x,\tau) < h(x,\tau)$ will occur on the left of the boundary. Therefore, one can compute the numerical solution of the problem (2.1.16)-(2.1.20) in the interval $[a, x_f(\tau)]$, and move the right boundary $x_f(\tau)$ until the moving boundary condition (2.1.19) is satisfied. It is clear that the computational cost of solving (2.1.16)-(2.1.20) is much less than that of solving (2.1.5)-(2.1.9), because the domain of the former is $[a, x_f(\tau))$, while the domain of the latter is $(-\infty, x_f(\tau))$.

2.1.2 American Put Option on a Stock with Continuous Dividend

About the American put options on a stock with continuous dividend, Tangman et al [80] point out that they can also be valuated with the domain decomposition technique. The PDE problem of this kind of American put option, after the same variable transformation as above, can be formulated as

$$\frac{\partial u}{\partial \tau} = \frac{\partial^2 u}{\partial x^2}, \quad x_f(\tau) < x < \infty, \quad 0 \leq \tau < \frac{\sigma^2 T}{2}, \tag{2.1.21}$$

$$u(x, 0) = g(x, 0), \quad x_f(0) < x < \infty, \tag{2.1.22}$$

$$u(x_f(\tau), \tau) = g(x_f(\tau), \tau), \quad g(x, \tau) = e^{-\alpha x - \beta \tau}(1 - e^x)^+, \tag{2.1.23}$$

$$u(x, \tau) \to 0 \quad as \quad x \to \infty. \tag{2.1.24}$$

Similarly, the artificial boundary condition of the American put option problem can be obtained as

$$\frac{\partial u}{\partial x}\Big|_{x=b} = -\frac{1}{\sqrt{\pi}} \int_0^\tau \frac{\partial u(b, \lambda)}{\partial \lambda} \frac{1}{\sqrt{\tau - \lambda}} d\lambda$$

based on the fundamental solution of the heat equation. Now the problem (2.1.21)-(2.1.24) can be tackled with the domain decomposition technique.

The put-call parity formula for the American options on a stock with continuous dividend is (see Tangman et al [80]):

$$S_f(\tau, \widetilde{r} = a', \delta = b') = \frac{K^2}{S_{f_c}(\tau, \widetilde{r} = b', \delta = a')},$$

$$V(S, \tau, \widetilde{r} = a', \delta = b') = \frac{S}{K} V_c(K^2/S, \tau, \widetilde{r} = b', \delta = a'),$$

where a', b' are constants, and the terms on the left hand side are the optimal exercise boundary and option value of the American put option, while the terms on the right hand side are the counterparts of the American call option. By this formula, we have the following proposition.

Proposition 2.1.2 *(Tangman et al [80]) For any τ, there exists $x_f^{**}(\tau) < x_f(\tau)$ such that*

$$u(x, \tau) = g(x, \tau), \quad when \quad x \leq x_f^{**}(\tau),$$

$$u(x, \tau) < g(x, \tau), \quad when \quad x_f^{**}(\tau) < x < x_f(\tau),$$

$$u(x, \tau) > g(x, \tau), \quad when \quad x \geq x_f(\tau).$$

This property can be applied to locate the moving boundary for the American put option.

We give the American put option results computed by the domain decomposition technique under the parameters $\widetilde{r} = 0.1, \sigma = 0.3, \delta = 0.05, K = 100, S \in [0, 400], T \in [0, 0.5]$. We use the same numerical scheme and algorithm as in Han and Wu [44], thus the details are omitted here. We implement the numerical tests through our own C++ program. The option value, Delta and Gamma of the option value (i.e., $\partial V/\partial S$ and $\partial^2 V/\partial^2 S$), and $V(S, t) - (K - S)^+$ are presented in Figures 2.1-2.4. They are in the shape as expected.

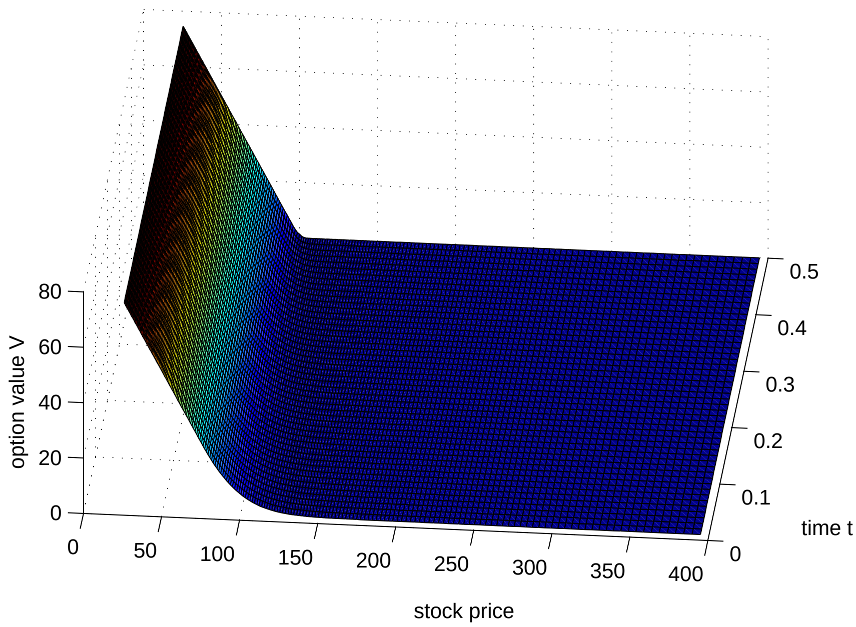

Figure 2.1: Option value

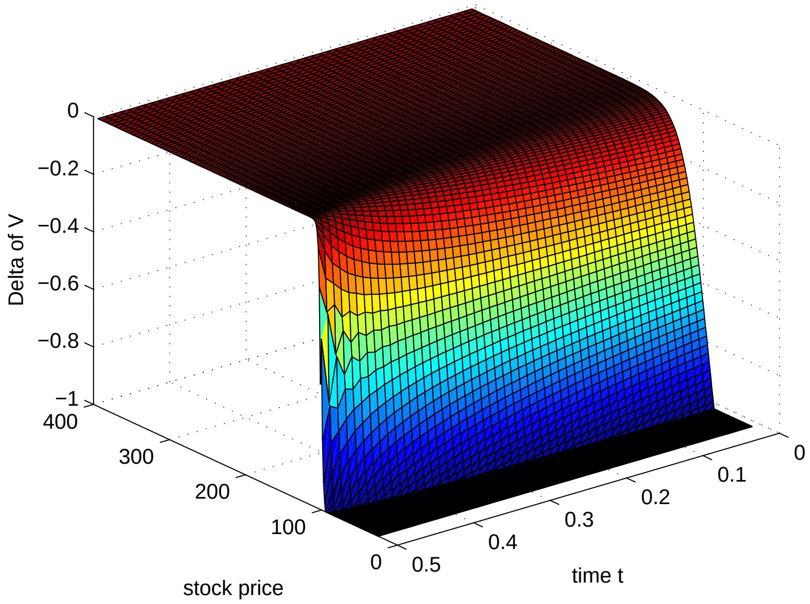

Figure 2.2: Delta of V: $\partial V/\partial S$

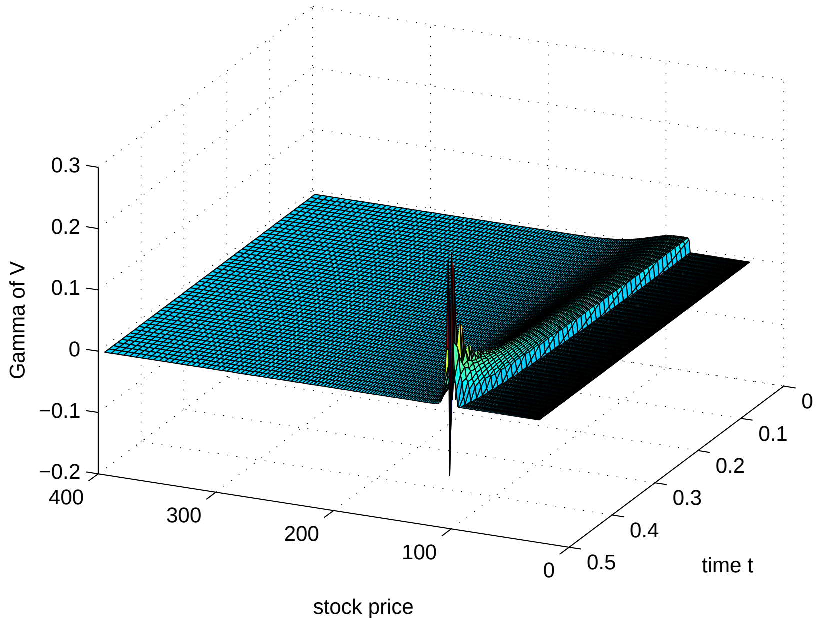

Figure 2.3: Gamma of V: $\partial^2 V/\partial^2 S$

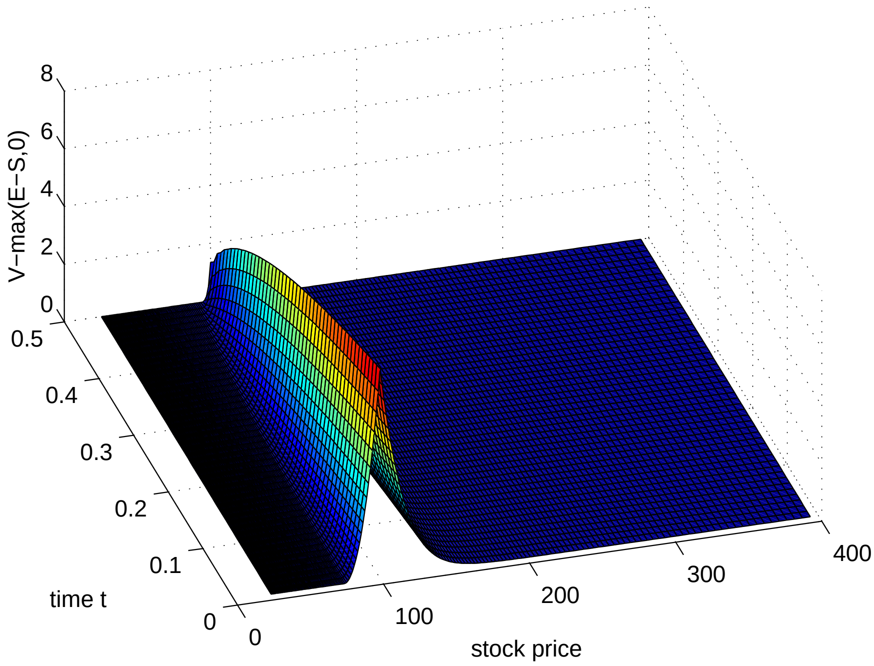

Figure 2.4: $V - (K - S)^+$

2.2 American Put Option Valuation by Solving the LCP

As known, the LCP for American put option on a stock can only be treated by numerical methods. People have developed different numerical approaches such as finite difference/element method combined with the PSOR method (see Wilmott, Howison and Dewynne [89], Zhang [97]), finite element method combined with Brennan-Schwartz algorithm (see Allegretto, Lin and Yang [2]), and finite volume method combined with penalty approach (see Wang, Yang and Teo [87], Angermann and Wang [8]), etc. In the first two methods, people usually transform the original LCP to a heat inequality problem (corresponding to the problem (2.1.21)-(2.1.24)) by the variable transformation: $t = T - \frac{2\tau}{\sigma^2}, S = Ke^x, V(S,t) = Ke^{\alpha x+\beta\tau}u(x,\tau)$, and then solve it by stable implicit schemes. By the penalty approach, the original LCP is transformed to a penalized equation and then solved numerically.

2.2.1 Power Penalty Method

The lower order power penalty method for solving the original LCP is proposed in Wang, Yang and Teo [87], where the LCP is transformed into a nonlinear parabolic partial differential equation by adding a power penalty term. It proves that the real solution of the penalized equation converges to that of the original LCP with an order depending on the penalty parameters. The penalized problem is then discretized by finite volume method and solved by Newton's method. Numerical examples are also presented to show the usefulness of the method. The convergence of the numerical scheme for the penalized problem is discussed in Angermann and Wang [8]. We describe the power penalty method in brief here.

As before, let $V(S,t)$ denote the value of the American put option with strike price K and maturity T, and S denotes the price of the underlying stock without dividend. It is known that $V(S,t)$ satisfies the following model

$$\mathfrak{L}V(S,t) := -\frac{\partial V}{\partial t} - \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - \widetilde{r} S \frac{\partial V}{\partial S} + \widetilde{r} V \geq 0, \tag{2.2.25}$$

$$V(S,t) \geq g(S) := (K-S)^+, \tag{2.2.26}$$

$$\mathfrak{L}V(S,t)[V(S,t) - g(S)] = 0, \tag{2.2.27}$$

a.e. in $\Omega := (0, \widetilde{X}) \times (0,T)$, σ and $\widetilde{r}$ are positive constants, $\widetilde{X}$ is the upper bound of S. The final condition and boundary conditions are

$$V(S,T) = g(S), \tag{2.2.28}$$

$$V(0,t) = g(0) = K, \quad V(\widetilde{X}, t) = 0. \tag{2.2.29}$$

Let $k > 0$ be a parameter and $y_+ := \max(y, 0)$, the power penalty approach is to solve a sequence of nonlinear partial differential equations of the form

$$\mathfrak{L}V_\lambda(S,t) + \lambda[g(S) - V_\lambda(S,t)]_+^{1/k} = 0, \quad (S,t) \in \Omega, \tag{2.2.30}$$

where λ is the penalty parameter, and $\lambda[g(S) - V_\lambda(S,t)]_+^{1/k}$ is the penalty term. When the constraint (2.2.26) is satisfied, (2.2.30) reduces to the Black-Scholes equation. When (2.2.26) is violated, (2.2.30) gives $[g(S) - V_\lambda]_+ = \lambda^{-k}(-\mathfrak{L}V_\lambda)^k$. In this case, if λ is sufficiently large and $\mathfrak{L}V_\lambda$ is bounded, $[g(S) - V_\lambda]_+ \approx 0$, so that (2.2.26) is satisfied within a tolerance depending on λ.

Before reformulating the LCP (2.2.25)-(2.2.27) as a variational problem, it is first transformed to an equivalent form satisfying homogeneous Dirichlet boundary conditions. Let V_0 be the linear function satisfying the boundary conditions (2.2.29):

$$V_0(S) = (1 - S/\widetilde{X})K,$$

and introduce a new variable

$$\tilde{u}(S,t) = e^{\beta' t}[V_0(S) - V(S,t)],$$

where $\beta' = \sigma^2$. Taking $\mathfrak{L}V_0$ away from both sides of (2.2.25) and transforming V in (2.2.25)-(2.2.27) into the new variable $\tilde{u}$, it is easy to show that, the LCP (2.2.25)-(2.2.27) becomes

$$\mathscr{L}\tilde{u}(S,t) := -\frac{\partial \tilde{u}}{\partial t} - \frac{\partial}{\partial S}[\tilde{a}S^2\frac{\partial \tilde{u}}{\partial S} + \tilde{b}S\tilde{u}] + \tilde{c}\tilde{u} \le f(S,t), \tag{2.2.31}$$

$$\tilde{u}(S,t) \le \tilde{u}^*(S), \tag{2.2.32}$$

$$(\mathscr{L}\tilde{u}(S,t) - f(S,t))[\tilde{u}(S,t) - \tilde{u}^*(S,t)] = 0, \tag{2.2.33}$$

where

$$f(S,t) = e^{\beta' t}\mathfrak{L}V_0(S,t), \quad \tilde{a} = \sigma^2/2, \quad \tilde{b} = \widetilde{r} - \sigma^2, \quad \tilde{c} = \widetilde{r} + \tilde{b} + \beta' = 2\widetilde{r} + \beta' - \sigma^2,$$

$$\tilde{u}^*(S,t) = e^{\beta' t}(V_0(S) - g(S)) = \begin{cases} e^{\beta' t}(1 - K/\widetilde{X})S, & 0 \le S \le K, \\ e^{\beta' t}(1 - S/\widetilde{X})K, & K < S \le \widetilde{X}. \end{cases}$$

The LCPs defined by (2.2.25)-(2.2.27) and (2.2.31)-(2.2.33) are equivalent in the sense that their solutions are related by

$$\tilde{u}(S,t) = e^{\beta' t}[V_0(S) - V(S,t)].$$

Now, we introduce a weighted L^2-norm and the space $L^2_\omega(I)$ and $H^1_{0,\omega}(I)$:

$$\|v\|_{0,\omega} := \left(\int_0^{\widetilde{X}} S^2 v^2 dS\right)^{1/2}, \quad L^2_\omega(I) := \{v : \|v\|_{0,\omega} < \infty\},$$

$$H^1_{0,\omega}(I) := \{v : v \in L^2(I), v' \in L^2_\omega(I), \quad and \quad v(\widetilde{X}) = 0\}.$$

Then

$$\|v\|_A := \left(\int_0^{\widetilde{X}} (S^2 v^2 + v^2) dS\right)^{1/2} \tag{2.2.34}$$

denotes a weighted H^1-norm on $H^1_{0,\omega}(I)$.

Let $\mathcal{K} = \{v \in H^1_{0,\omega}(I) : v \leq \tilde{u}^*\}$, then the variational form corresponding to the problem (2.2.31)-(2.2.33) is as follows:
To find $\tilde{u}(t) \in \mathcal{K}$ such that

$$(-\frac{\partial \tilde{u}(t)}{\partial t}, v - \tilde{u}(t)) + A(\tilde{u}(t), v - \tilde{u}(t); t) \leq (f(t), v - \tilde{u}(t)), \quad \forall v \in \mathcal{K}, \quad (2.2.35)$$

a.e. in $(0, T)$, where $A(\cdot, \cdot; t)$ is a bilinear form defined by

$$A(\tilde{u}, v; t) := (\tilde{a}S^2\tilde{u}' + \tilde{b}S\tilde{u}, v') + (\tilde{c}\tilde{u}, v), \quad \tilde{u}, v \in H^1_{0,\omega},$$

and $\tilde{u}'$, v' are first derivatives.

The bilinear form $A(\cdot, \cdot; t)$ proves to be coercive and continuous, i.e., there exist positive constants C and M, independent of v and w, such that, for any $v, w \in H^1_{0,\omega}(I)$,

$$A(v, v; t) \geq C \|v\|_A^2 \quad \textit{and} \quad A(v, w; t) \leq M \|v\|_A \|w\|_A ,$$

for $t \in (0, T)$, where $\|v\|_A$ is the norm defined in (2.2.34).

This property of $A(\cdot, \cdot; t)$ guarantees that the variational problem (2.2.35) is uniquely solvable. In other words, the LCP (2.2.31)-(2.2.33) is guaranteed to have a unique solution in the convex set $\mathcal{K}$. Then the power penalty method is applied to solve the LCP.

The penalized equation corresponding to (2.2.31)-(2.2.33) is

$$\mathscr{L}\tilde{u}_\lambda(S, t) + \lambda[\tilde{u}_\lambda(S, t) - \tilde{u}^*(S, t)]_+^{1/k} = f(S, t), \quad (S, t) \in \Omega, \quad (2.2.36)$$

$$\tilde{u}_\lambda(0, t) = 0 = \tilde{u}_\lambda(\widetilde{X}, t), \quad \tilde{u}_\lambda(S, T) = \tilde{u}^*_\lambda(S, T), \quad (2.2.37)$$

where $\lambda > 0$ and $k > 0$ are parameters. The details of solving (2.2.36)-(2.2.37) numerically is referred to Wang, Yang and Teo [87].

2.2.2 The PSOR Method for the LCP

The PSOR method has been proposed for attacking kinds of linear complementarity problems, and its convergence is guaranteed by the positive definiteness of the coefficient matrix of the discretized system (see Elliott and Ockendon [33]). For the LCP arising from the American put option on a stock without dividend, the discrete forms of the transformed LCP generated by finite difference/element method are known to satisfy this requirement. In the meantime, the fitted finite volume discretization for the original LCP generates a nonsingular M-matrix, which is also positive definite. But the discretized system has not been solved by the PSOR method in existing literature.

We apply the PSOR method to the LCP that discretized by the fitted finite volume method (see Wang, Yang and Teo [87]) and compare its results with that produced by the power penalty method. We also include the finite difference/element discretization combined with the PSOR method in our numerical tests, and find that the results of these four methods coincide with good accuracy. In our experiments, we solve the problem on the truncated domain $[-ln\frac{\widetilde{X}}{K}, ln\frac{\widetilde{X}}{K}]$ or $[0, \widetilde{X}]$, contrasted to an exact boundary condition on the left hand side and truncated boundary condition on the right hand side in Allegretto, Lin and Yang [2]. The details of the numerical schemes and algorithms for solving the problem are omitted here, which can be referred to [87, 89, 97]. Numerical examples will be provided in Section 2.4.

2.3 Stochastic Simulation Methods for Option Valuation

The lattice (tree) method and Monte Carlo simulation method for option valuation are based on simulating sample paths for the underlying asset price. They are termed as stochastic simulation methods and have been applied to a wide range of pricing problems. The convergence rate of Monte Carlo simulation method is typically independent of the number of state variables, whereas the convergence rate of lattice method is exponentially decreasing with respect to the number of state variables. As a result, the Monte Carlo simulation method is increasingly attractive compared to the lattice method as the number of state variables grows.

2.3.1 The Binomial Tree Method

The binomial tree method is the simplest case of lattice method. It is based on no-arbitrage principle, which means all investments get the same riskless return. In the binomial tree method, the underlying stock price either goes up or down by a fixed multiplier $\mathbb{u}$ or $\mathbb{d}$ in one time-step, and the probability of up-movement and down-movement is q and $1-q$, respectively. The risk-neutral q can be picked out by the no-arbitrage principle.

We denote the option price and the underlying stock (without dividend) price at the initial time by V_0, S_0. At the end of the first time-step, corresponding to the stock price $S_{\mathbb{u}} = S_0\mathbb{u}$ and $S_{\mathbb{d}} = S_0\mathbb{d}$, the option prices are denoted by V_1, V_2, then

$$V_0 = \frac{1}{\rho}[qV_1 + (1-q)V_2].$$

Furthermore, under the condition $\mathbb{u}\mathbb{d} = 1$, we have

$$\begin{cases} q = (e^{(\tilde{r}\triangle t)} - \mathbb{d})/(\mathbb{u} - \mathbb{d}), \\ \mathbb{u} = e^{(\sigma\sqrt{\Delta}t)}, \\ \mathbb{d} = e^{(-\sigma\sqrt{\Delta}t)}, \\ \rho = e^{\tilde{r}\Delta t}. \end{cases}$$

Assume that there are N states for the stock price at T, the maturity of the option, then the European call option written on the stock has the pricing formula at time t as follows

$$e^{-\tilde{r}(T-t)} \sum_{i=0}^{N} C_i^N q^i (1-q)^{N-i} (\mathbb{u}^i \mathbb{d}^{N-i} S_0 - K)^+,$$

where $C_i^N = \frac{N!}{i!(N-i)!}$, $N!$ is a factorial and K is the strike price of the option. For European put options, they have a similar pricing formula as above except for a different payoff function.

The binomial tree method has been successful for pricing contingent claims on a stock which follows a geometric Brownian motion, because the multiplicative binomial model for stock prices includes the lognormal distribution as a limiting case when the time interval between successive steps tends to zero. It has been extended in many directions, such as the tri-nomial tree method. Although lattice method has many desirable features, its computational cost grows exponentially with regard to the number of state variables, and the rigorous proof of its convergence is difficult. The convergence of the binomial method for American options was not established until 1994 (see Amin and Khanna [5]).

2.3.2 Monte Carlo Simulation Method

We have mentioned that option price can be expressed as the expectation of its discounted payoff function under the risk-neutral measure, thus Monte Carlo method is often applied to compute the expectation. Now we describe the Monte

Carlo method for option pricing.

Assume that the underlying stock price S_t follows a lognormal diffusion process

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

where μ and σ are two positive constants and B_t is a Brownian motion. We introduce the process $\widetilde{B_t} = B_t + (\mu - \widetilde{r})t/\sigma$, which is a Brownian motion under the risk-neutral measure. Now we have

$$dS_t = S_t(\widetilde{r}dt + \sigma d\widetilde{B_t})$$

with the solution

$$S_t = S_0 e^{(\sigma \widetilde{B_t} + \widetilde{r}t - \frac{\sigma^2}{2}t)},$$

where S_0 is the price of the stock at time 0.

Let Z_i $(i = 1, \ldots, n)$ be independent samples simulating the process $\widetilde{B_t}$,

$$S_t^i = S_0 e^{(\widetilde{r} - \frac{\sigma^2}{2})t + \sigma\sqrt{t}Z_i}, \quad i = 1, \ldots, n,$$

and denote the payoff function of the option at time t by $g(t, S_t)$, then the European option price at t is

$$V(t) = \mathbb{E}^Q[e^{-\widetilde{r}(T-t)} g(T, S_T)],$$

and the American option price at t can be written as

$$\sup_{\tau \in [t,T]} \mathbb{E}^Q[e^{-\widetilde{r}(\tau - t)} g(\tau, S_\tau)],$$

where Q denotes the risk-neutral measure, and the supermum is over all stopping times $\tau \leq T$.

The Monte Carlo approach is a valuable and flexible computational tool for option pricing. It is easy to implement for European options, but not for American options, due to the estimation of the early exercise decisions in American options. We know that standard simulation programs are forward algorithms, which means that the paths of the state variable are simulated forward in time. However, pricing options with early-exercise features generally requires a backward algorithm. By working backward from the maturity date of the option via dynamic programming, the optimal exercise strategy and option value are estimated. Therefore, the problem of using simulation method to price American options lies in the inconsistency of applying an inherently forward procedure to the problem that requires a backward procedure. When the usual backward induction algorithm is applied to a sample path, the resulting future value is derived from perfect foresight, rather than based on expectation.

2.4 Numerical Examples

In this section, we take an American put option on a stock without dividend for example to test the congruence of the stochastic method and those numerical methods for the LCP. We take the results computed by the binomial tree method as benchmark, and compare them with the results generated from different methods for the LCP. We implement numerical experiments under two groups of parameters: $\widetilde{r} = 0.07, \sigma = 0.3, T = 3, S \in [0, 200], K = 45$, and $\widetilde{r} = 0.1, \sigma = 0.5, T = 1, S \in [0, 200], K = 60$. The results are shown in Figure 2.5 and Figure 2.6, where "fd" means finite difference method, "fe" means finite element method, and "fvm" means finite volume method. Numerical results indicate that these approaches, including the binomial tree method, the penalty method, and the PSOR method combined with different discretization techniques, comply with each other with good accuracy.

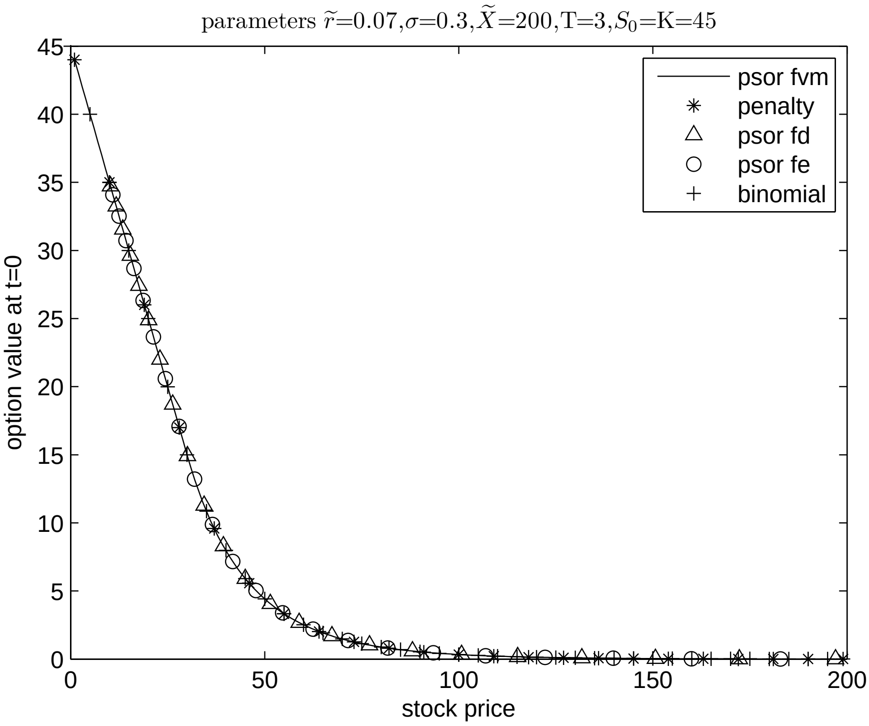

Figure 2.5: Comparison of different methods

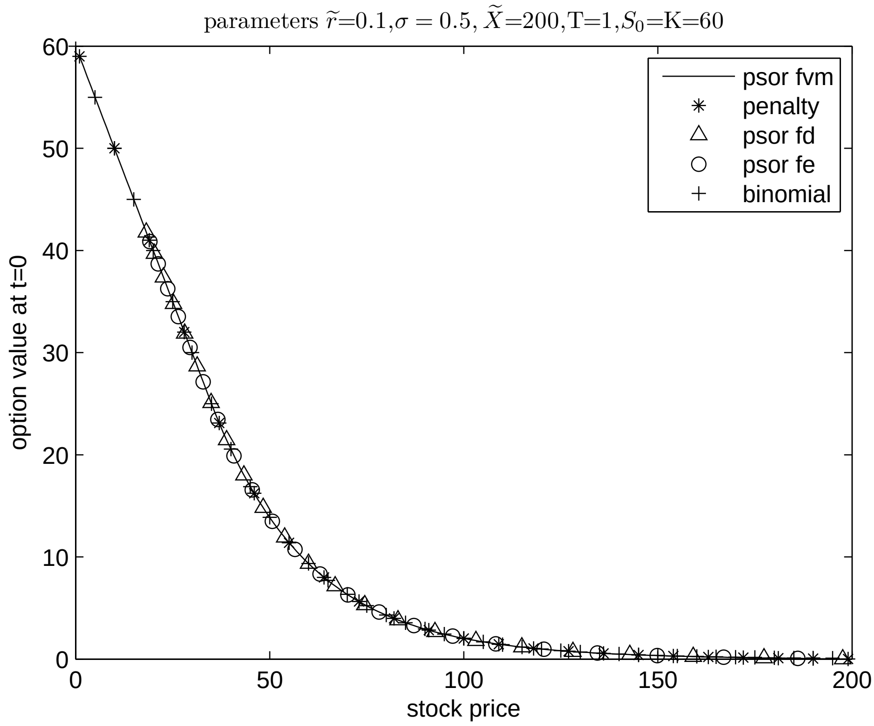

Figure 2.6: Comparison of different methods

Among these methods for option pricing, the binomial tree method and Monte Carlo method are easy to implement to get the option value, but they can hardly provide the optimal exercise boundary of American options. Different variance reduction techniques have been studied to improve the efficiency of the Monte Carlo method for option pricing, while the lattice method is hard to extend when the convergence is concerned. The penalty method and the PSOR method have been extended to option valuation under jump-diffusion models and multi-asset option valuation, since their convergence properties are relatively easier to be established.

2.5 Summary

In this chapter, we firstly reviewed the domain decomposition method for attacking the PDE problem of American options on a stock with dividend. Subsequently, the numerical techniques, such as the penalty method and the PSOR method, for solving the LCP arising from American put option valuation were summarized. We proposed that the finite volume method combined with the PSOR method could also work out the LCP. The lattice method and Monte Carlo method were then described briefly. After that, numerical tests were presented to show the congruence of these methods for American option pricing.

Chapter 3

Evaluating American Put Options on Zero-coupon Bonds by a Penalty Method

We have surveyed the numerical methods for valuing American options on a single stock in the previous chapter. Now we turn to the problem of evaluating American options on zero-coupon bonds, which can also be formulated as a LCP. We adopt finite volume method to establish a numerical scheme for the LCP, and then use the penalty method to solve the discretized LCP. This chapter is organized as follows. In Section 3.1, we introduce the pricing models of zero-coupon bonds and American put options on zero-coupon bonds under the CKLS short rate model. In Section 3.2, we establish a numerical scheme using the finite volume method and give the corresponding stability and convergence property. In Section 3.3, we illustrate the power penalty approach for the discretized LCP and its convergence. In Section 3.4, the Brennan-Schwartz algorithm is described briefly. We present our numerical results in Section 3.5 and concluding remarks in Section 3.6.

3.1 The Pricing Models of Zero-coupon Bonds and Options on the Bond

As we have mentioned in Section 1.1.2, under the CKLS model

$$dr(t) = \kappa(\theta - r(t))dt + \sigma r(t)^{\gamma} dW(t), \tag{3.1.1}$$

the price at time t of the zero-coupon bond with face value E and maturity T^*, denoted by $B(r, t, T^*)$, satisfies the following equation:

$$\frac{\partial B}{\partial t} + \frac{1}{2}\sigma^2 r^{2\gamma} \frac{\partial^2 B}{\partial r^2} + \kappa(\theta - r)\frac{\partial B}{\partial r} - rB = 0, \tag{3.1.2}$$

$$B(r, T^*, T^*) = E. \tag{3.1.3}$$

Furthermore, the price at time t of the American put option with maturity T on the above zero-coupon bond, denoted by $V(r, t)$, satisfies the following free boundary problem:

$$V_t + \mathcal{L}V = 0, \quad V(r,t) > g(r,t), \quad 0 < r < r^*(t), \quad 0 \leq t < T, \tag{3.1.4}$$

$$V(r^*(t), t) = g(r^*(t), t), \quad V_r(r^*(t), t) = g_r(r^*(t), t), \quad 0 \leq t < T, \tag{3.1.5}$$

$$V(r,t) = g(r,t), \quad r > r^*(t), \quad 0 \leq t \leq T, \tag{3.1.6}$$

$$V(0,t) = g(0,t), \quad 0 \leq t \leq T, \tag{3.1.7}$$

$$V(r,T) = g(r,T), \quad r \geq 0, \tag{3.1.8}$$

where

$$g(r,t) = \max(K - B(r, t, T^*), 0),$$

$$\mathcal{L}V := \frac{1}{2}\sigma^2 r^{2\gamma} V_{rr} + \kappa(\theta - r)V_r - rV,$$

$V_t = \frac{\partial V}{\partial t}, V_r = \frac{\partial V}{\partial r}, V_{rr} = \frac{\partial^2 V}{\partial r^2}, g_r = \frac{\partial g}{\partial r}$, and $r^*(t)$ is the optimal exercise interest rate at time t for the American put option. Note that K should be strictly less than $B(0, T, T^*)$, otherwise, exercising the option would never be optimal.

In order to solve the system defined by (3.1.4)-(3.1.8), we need to solve (3.1.2)-(3.1.3) for the bond price first by numerical method since no analytical solution can be derived. It is known that r is the short-term interest rate, we can restrict $r \in [0, R]$ for some sufficiently large number R, and define boundary conditions for (3.1.2) at $r = 0$ and $r = R$ as the bond price determined by the CIR model (which has an analytic expression). The numerical scheme for (3.1.2) will be clarified in the next section.

Actually, The free boundary problem (3.1.4)-(3.1.8) can be rewritten as a LCP as follows

$$-V_t - \mathcal{L}V \geq 0, \tag{3.1.9}$$

$$V(r,t) \geq g(r,t), \tag{3.1.10}$$

$$(-V_t - \mathcal{L}V)[V(r,t) - g(r,t)] = 0, \tag{3.1.11}$$

a.e. in $[0, R] \times [0, T]$. The boundary and terminal conditions are

$$V(0,t) = g(0,t), \quad V(R,t) = g(R,t),$$

$$V(r,T) = g(r,T), \quad r \geq 0.$$

The equivalence between the free boundary problem and the LCP for American put options is obtained in Section 7.6 of [89], thus we can solve either of them for the option value. In this paper, we discuss the power penalty method for the discrete form of the LCP (3.1.9)-(3.1.11).

3.2 Finite Volume Method for Discretization

The finite volume method was adopted for discretizing the Black-Scholes equations in [4, 86]. We apply the method to discretize (3.1.9)-(3.1.11) as in Allgrettto, Lin and Yang [4], which can also be used to discretize (3.1.2). We introduce the transformations $u(r,t) = V(r, T-t)$ and $G(r,t) = g(r, T-t)$, then the LCP (3.1.9)-(3.1.11) can be reformulated as

$$u_t - \mathcal{L}u \geq 0, \quad u \geq G, \quad 0 < r < R, \quad 0 < t \leq T, \tag{3.2.12}$$

$$(u_t - \mathcal{L}u)(u - G) = 0, \quad 0 < r < R, \quad 0 < t \leq T, \tag{3.2.13}$$

$$u(r,0) = G(r,0), \quad 0 \leq r \leq R, \tag{3.2.14}$$

$$u(0,t) = G(0,t), \quad 0 \leq t \leq T, \tag{3.2.15}$$

$$u(R,t) = G(R,t), \quad 0 \leq t \leq T. \tag{3.2.16}$$

Dividing the first inequality of (3.2.12) by $\frac{\sigma^2 r^{2\gamma}}{2}$ gives the following inequality

$$cr^{-2\gamma}u_t - u_{rr} - [ar^{-2\gamma} - br^{1-2\gamma}]u_r + cr^{1-2\gamma}u \geq 0, \tag{3.2.17}$$

where $a = \frac{2\kappa\theta}{\sigma^2}$, $b = \frac{2\kappa}{\sigma^2}$, $c = \frac{2}{\sigma^2}$. We try to simplify the expression by combining the second and the third term on the left-hand side of the above inequality. Define a function $\Psi(a, b, r, \gamma)$ (abbreviated as $\Psi(r)$) such that

$$\frac{1}{\Psi(r)}(\Psi(r)u_r)_r = u_{rr} + [ar^{-2\gamma} - br^{1-2\gamma}]u_r,$$

we can deduce that

$$\Psi(r) = \exp\left[\frac{ar^{1-2\gamma}}{1-2\gamma} - \frac{br^{2-2\gamma}}{2-2\gamma}\right].$$

Now (3.2.17) is reformulated as

$$cr^{-2\gamma}u_t - \frac{1}{\Psi(r)}(\Psi(r)u_r)_r + cr^{1-2\gamma}u \geq 0.$$

Let $w_1(r) = cr^{-2\gamma}\Psi(r)$, $w_2(r) = \Psi(r)$, $w_3(r) = cr^{1-2\gamma}\Psi(r)$, then the system

(3.2.12)-(3.2.16) is transformed into

$$w_1 u_t - (w_2 u_r)_r + w_3 u \geq 0, \quad u \geq G, \quad 0 < r < R, \quad 0 < t \leq T, \tag{3.2.18}$$

$$(w_1 u_t - (w_2 u_r)_r + w_3 u)(u - G) = 0, \quad 0 < r < R, \quad 0 < t \leq T, \tag{3.2.19}$$

$$u(r, 0) = G(r, 0), \quad 0 \leq r \leq R, \tag{3.2.20}$$

$$u(0, t) = G(0, t), \quad 0 \leq t \leq T, \tag{3.2.21}$$

$$u(R, t) = G(R, t), \quad 0 \leq t \leq T. \tag{3.2.22}$$

Similarly, the pricing equation for the zero-coupon bond (3.1.2) can be rewritten as

$$w_1 \tilde{u}_t - (w_2 \tilde{u}_r)_r + w_3 \tilde{u} = 0, \tag{3.2.23}$$

$$\tilde{u}(r, 0, T^*) = E, \tag{3.2.24}$$

where $\tilde{u}(r, t, T^*) = B(r, T^* - t, T^*)$. As we adopt the same discretization to evaluate the bond price and the option value, we skip the details of solving the bond price and go straight into the option valuation problem.

As far as the problem (3.2.18)-(3.2.22) is concerned, we use uniform mesh $h = R/N, \Delta t = T/M$, $r_j = jh, j = 0, 1, \ldots, N$, $t_m = m\Delta t, m = 0, 1, \ldots, M$, and backward difference to approximate $u_t(r, t_m)$, i.e.,

$$u_t(r, t_m) \approx \delta_t u(r, t_m) = \frac{u(r, t_m) - u(r, t_{m-1})}{\Delta t}.$$

For $j = 1, 2, \ldots, N$, integrating

$$w_1(r) u_t(r, t_m) - (w_2(r) u_r(r, t_m))_r + w_3(r) u(r, t_m)$$

in r over $[r_{j-1/2}, r_{j+1/2}] = [(r_j + r_{j-1})/2, (r_j + r_{j+1})/2]$ and using the approximations

$$u_r(r_{j\pm\frac{1}{2}}, t_m) \approx \frac{u(r_j, t_m) - u(r_{j\pm1}, t_m)}{r_j - r_{j\pm1}},$$

$$\int_{r_{j-1/2}}^{r_{j+1/2}} w_i(r)u(r,t_k)dr \approx \left(\int_{r_{j-1/2}}^{r_{j+1/2}} w_i(r)dr\right) u(r,t_k), i=1,2,3, k=m-1,m,$$

we get the discrete form of (3.2.18)-(3.2.22) as follows:

$$D(U^m - U^{m-1}) + AU^m \geq F^m, U^m \geq G^m, \tag{3.2.25}$$

$$\left[D(U^m - U^{m-1}) + AU^m - F^m\right]^T (U^m - G^m) = 0, \tag{3.2.26}$$

$$U^0 = G^0, \tag{3.2.27}$$

$$u(0,t_m) = G(0,t_m), \tag{3.2.28}$$

$$u(R,t_m) = G(R,t_m), \tag{3.2.29}$$

for $m = 1,2,\ldots,M$, where D is a diagonal matrix with entries $d_1, d_2, \ldots, d_{N-1}$, and A is a tri-diagonal matrix. In addition, u_j^m is the approximation of $u(r_j,t_m)$, and

$$U^m = \begin{pmatrix} u_1^m \\ u_2^m \\ \vdots \\ u_{N-1}^m \end{pmatrix}, F^m = \begin{pmatrix} f_1^m \\ f_2^m \\ \vdots \\ f_{N-1}^m \end{pmatrix}, G^m = \begin{pmatrix} G_1^m \\ G_2^m \\ \vdots \\ G_{N-1}^m \end{pmatrix},$$

$$A = \begin{pmatrix} a_1 & -b_1 & 0 & \cdots & 0 \\ -b_1 & a_2 & -b_2 & \ddots & \vdots \\ 0 & -b_2 & a_3 & -b_3 & 0 \\ \vdots & \ddots & \ddots & \ddots & -b_{N-2} \\ 0 & \cdots & 0 & -b_{N-2} & a_{N-1} \end{pmatrix},$$

$$a_j = b_{j-1} + b_j + c_j, \quad c_j = \int_{r_{j-1/2}}^{r_{j+1/2}} w_3(r)dr,$$

$$d_j = \frac{1}{\Delta t}\int_{r_{j-1/2}}^{r_{j+1/2}} w_1(r)dr, \quad j = 1,2,\ldots,N-1,$$

$$b_j = \frac{w_2(r_{j+1/2})}{r_{j+1} - r_j}, \quad j = 0,1,2,\ldots,N-1,$$

$$G_j^m = G(r_j,t_m), \quad j = 0,1,2,\ldots,N,$$

$$f_1^m = b_0 G_0^m, \quad f_{N-1}^m = b_{N-1} G_N^m, \quad f_j^m = 0, \quad j = 2,\ldots,N-2.$$

We construct the following $\xi-$scheme for (3.2.25)-(3.2.26):

$$D(U^m - U^{m-1}) + AU^{m-1+\xi} \geq F^{m-1+\xi}, U^m \geq G^m, \tag{3.2.30}$$

$$\left[D(U^m - U^{m-1}) + AU^{m-1+\xi} - F^{m-1+\xi}\right]^T (U^m - G^m) = 0, \tag{3.2.31}$$

for $m = 1, 2, \ldots, M$, where

$$U^{m-1+\xi} = (\xi U^m + (1-\xi)U^{m-1}), \quad F^{m-1+\xi} = (\xi F^m + (1-\xi)F^{m-1}), \quad \xi \in [0,1],$$

and the initial and boundary conditions are defined by (3.2.27)-(3.2.29). Obviously, when $\xi = 1$, 0.5 and 0, the $\xi-$scheme is the backward Euler scheme, the Crank-Nicolson scheme, and the forward Euler scheme, respectively. For the stability and convergence of the $\xi-$scheme, we state the results in Allegretto, Lin and Yang [4] as follows.

Theorem 3.2.1 *(Allegretto, Lin and Yang [4]) For any $\xi \in [0,1]$, the linear complementarity problem (3.2.30)-(3.2.31) has a unique solution U^m for $m = 1, 2, \ldots, M$. Furthermore, if*

$$\bar{\mu} = 1 - 4(1-\xi)\Lambda\frac{\Delta t}{h^2} \geq 0,$$

then

$$\max_{1\leq m\leq M} \|V^m\|^2_{0,w_1,h} + \bar{\mu} \sum_{m=1}^{M} \|V^m - V^{m-1}\|^2_{0,w_1,h}$$
$$+\Delta t \sum_{m=1}^{M} \left(\xi\|V^m\|^2_{1,w_2,h} + (1-\xi)\|V^{m-1}\|^2_{1,w_2,h}\right) \leq C,$$

where Λ is a positive constant, $V^m = (G_0^m, u_1^m, \ldots, u_{N-1}^m, G_N^m)$, C is a positive constant independent of $h, \Delta t$ and V^m, but dependent on $\|G\|_{W^{1,\infty}}$. Therefore, the Backward Euler scheme ($\xi = 1$) is unconditionally stable, and for $\xi \in [0,1)$ the scheme is conditionally stable.

Theorem 3.2.2 *(Allegretto, Lin and Yang [4]) Let $V_h^m(r)$ be the piecewise linear interpolation of V^m in r. Define $V_{h\tau}(r,t) = \sum_{m=0}^{M} V_h^m(r)\chi_\tau^m(t)$, where χ_τ^m is the characteristic function of $[t_m, t_{m+1})$. Under the conditions of Theorem 3.1, $V_{h\tau}$ has a weakly convergent subsequence in $L^2(0,T; H^1_{w^2}(0,R))$.*

3.3 Power Penalty Method for the Discretized LCP

The power penalty method was proposed to solve the continuous LCP arising from the American put options in Wang, Yang and Teo [87], where the solution of the penalized problem was proven to converge to that of the original LCP with the order of $\mathcal{O}(\lambda^{-k/2})$. It is shown in Wang and Yang [88] that when the system matrix $\mathbb{M}$ is positive definite and has positive diagonal entries together with nonpositive off-diagonal entries, the power penalty method works for the following LCP in n-dimensional real space $\mathbb{R}^n$ and the above convergence property holds:

Find $x \in \mathbb{R}^n$ such that

$$\mathbb{M}x - \mathbb{b} \leq 0, \quad x \leq 0, \tag{3.3.32}$$

$$(\mathbb{M}x - \mathbb{b})^T x = 0, \tag{3.3.33}$$

where $\mathbb{M}$ is an $n \times n$ matrix, and $\mathbb{b} \in \mathbb{R}^n$ is a vector.

The power penalty method transforms the LCP (3.3.32)-(3.3.33) into the problem as below.

Find $x_\lambda \in \mathbb{R}^n$ such that

$$\mathbb{M}x_\lambda + \lambda[x_\lambda]_+^{1/k} = \mathbb{b}, \tag{3.3.34}$$

where $\lambda > 1$ and $k > 0$ are penalty parameters, $[x]_+ = \max\{x, 0\}$ and $y^{1/k} = (y_1^{1/k}, \ldots, y_n^{1/k})^T$ for any $y = (y_1, \ldots, y_n)^T \in \mathbb{R}^n$.

It can be seen that the solution x_λ of (3.3.34) converges to that of (3.3.32)-(3.3.33) when $\lambda \to +\infty$, and the convergence rate is of the order $\mathcal{O}(\lambda^{-k/2})$ when $\mathbb{M}$ satisfies the conditions mentioned as above. As far as our problem (3.2.30)-(3.2.31) is concerned, it can be rewritten in the form of (3.3.32)-(3.3.33) with the system matrix equal to $D + \xi A$ which fulfills the properties required for $\mathbb{M}$. Therefore, the power penalty method is applicable for (3.2.30)-(3.2.31).

Adding a penalty term like $\lambda[x_\lambda]_+^{1/k}$ to (3.2.30), we get the following equation system at time t_m $(m = 1, 2, \ldots, M)$,

$$D(U^m - U^{m-1}) + AU^{m-1+\xi} + \vartheta h Q(U^{m-1+\xi}) = F^{m-1+\xi}, \tag{3.3.35}$$

where ϑ is a negative real number (penalty parameter), $Q(U^{m-1+\xi}) = \xi Q(U^m) + (1-\xi)Q(U^{m-1})$, and

$$Q(U^m) = \begin{pmatrix} Q_1(u_1^m) \\ Q_2(u_2^m) \\ \vdots \\ Q_{N-1}(u_{N-1}^m) \end{pmatrix} = \begin{pmatrix} (G_1^m - u_1^m)_+^{\frac{1}{k}} \\ (G_2^m - u_2^m)_+^{\frac{1}{k}} \\ \vdots \\ (G_{N-1}^m - u_{N-1}^m)_+^{\frac{1}{k}} \end{pmatrix}.$$

Now (3.3.35) is equivalent to

$$\begin{aligned} &D(U^m - U^{m-1}) + \xi AU^m + (1-\xi)AU^{m-1} + \vartheta h[\xi Q(U^m) + (1-\xi)Q(U^{m-1})] \\ &= \xi F^m + (1-\xi)F^{m-1}, \end{aligned} \tag{3.3.36}$$

i.e.,

$$\begin{aligned} &(D + \xi A)U^m + \vartheta h \xi Q(U^m) \\ &= \xi F^m + (1-\xi)F^{m-1} + (D - (1-\xi)A)U^{m-1} - \vartheta h(1-\xi)Q(U^{m-1}). \end{aligned} \tag{3.3.37}$$

System (3.3.37) is nonlinear in U^m. We would like to apply Newton's method to it. However, note that for $k > 1$, when $G_j^m - u \to 0$, $Q_j'(u) \to \infty$, so we have to use a smoothing technique for $Q_j(u_j^m)$ in the neighborhood of $[G_j^m - u]_+ = 0$.

Redefine $Q_j(u_j^m)$ as

$$Q_j(u_j^m) = \begin{cases} (G_j^m - u_j^m)^{1/k}, & G_j^m - u_j^m \geq \varepsilon, \\ S([G_j^m - u_j^m]_+), & G_j^m - u_j^m < \varepsilon, \end{cases} \tag{3.3.38}$$

for $k > 0$, where $0 < \varepsilon \ll 1$ is a transition parameter and $S(z)$ is the function which smooths out the original $Q_j(z)$ around $z = 0$. We choose

$$S(z) = e_1 + e_2 z + \; + e_n z^{n-1} + e_{n+1} z^n, \quad for \quad n \geq 3,$$

such that $Q_j(\cdot)$ is smooth, which requires that $S(z)$ satisfy

$$S(0) = S'(0) = 0, \quad S(\varepsilon) = \varepsilon^{1/k}, \quad S'(\varepsilon) = (1/k)\varepsilon^{1/k-1}.$$

Under this assumption, the function defined in (3.3.38) is globally smooth. Using these four conditions and setting $e_3 = \cdots = e_{n-1} = 0$, we find that

$$e_1 = e_2 = 0, \quad e_n = (n - 1/k)\varepsilon^{1/k-n+1}, \quad e_{n+1} = (1/k - n + 1)\varepsilon^{1/k-n}.$$

Note that $S(z)$ is strictly increasing on $[0, \varepsilon]$ if $k \geq 1/n$ (this is true because $S'(z) > 0$ on $[0, \varepsilon]$), we can see that the nonlinear function $Q_j(u_j^m)$ defined in (3.3.38) is smooth and monotone on $(-\infty, \infty)$ for $k \geq 1/n$ with any integer $n \geq 3$. Applying Newton's method to (3.3.37) gives

$$\begin{aligned}
&[D + \xi A + \vartheta \xi J_Q(w^{l-1})](\delta w^l) = \xi F^m + (1 - \xi)F^{m-1} \\
&+(D - (1 - \xi)A)U^{m-1} - \vartheta(1 - \xi)Q(U^{m-1}) \\
&-(D + \xi A)w^{l-1} - \vartheta \xi Q(w^{l-1}),
\end{aligned} \tag{3.3.39}$$

$$w^l = w^{l-1} + \beta(\delta w^l), \tag{3.3.40}$$

for $l = 1, 2, \ldots$, with w^0 a given initial guess, where $J_Q(w)$ denotes the Jacobian matrix of the column vector $Q(w)$ and $\beta \in (0, 1]$ is a damping parameter, and δw^l is the variation of w at each iteration. Then, choosing

$$U^m = \lim_{l \to \infty} w^l,$$

we know that w^l converges to U^m quadratically if w^0 is sufficiently close to U^m.

Observing the entries of matrices D, A, and J_Q, we know that the coefficient matrix of (3.3.39) is diagonally dominant with respect to its columns, and its diagonal elements are positive and off-diagonal elements are nonpositive, namely, it is an M-matrix. This implies that the numerical solution to (3.3.39) is convergent. Combining with Theorem 3.1, we know that when $\xi = 1$, our penalty approach gives convergent solution of the LCP.

3.4 Brennan-Schwartz Algorithm

The American put option pricing problem can be formulated as a free boundary problem or a LCP, and their equivalence is obtained in Section 7.6 of Wilmott, Howison and Dewynne [89]. The Brennan-Schwartz algorithm is first proposed to deal with the free boundary problem in Brennan and Schwartz [17], where it solves the equation with a free boundary, then it proves to be applicable to the corresponding LCP in Jaillet, Lamberton and Lapeyre [52]. Actually, the Brennan-Schwartz algorithm is adopted to tackle the American put options on zero-coupon bonds and proves to be convergent in Allegretto, Lin and Yang [4]. In this section, we apply it to solve the free boundary problem (3.1.4)-(3.1.8). In discrete form, we solve the following equation:

$$(D + A)U^m - DU^{m-1} = F^m, \quad m = 1, \dots, M, \tag{3.4.41}$$

with the constraint $u_j^m \geq G_j^m, j = 1, \dots, N-1$. It is actually equivalent to solving the LCP (3.2.25)-(3.2.26).

Transforming the tri-diagonal coefficient matrix of (3.4.41) to an upper triangular and bi-diagonal one, then the equation system is in the form:

$$l_j u_j^m + l_{j+1} u_{j+1}^m = \phi_j, \quad for \quad j = 1, \dots, N-1. \tag{3.4.42}$$

Since $u_N^m = G_N^m$, we solve u_{N-1}^m by (3.4.42); for $j = N-1, \dots, 2$,

$$if \quad u_j^m < G_j^m, u_j^m := G_j^m \to u_{j-1}^m; \quad else \quad u_j^m \to u_{j-1}^m.$$

In this way we get a complete set of u_j^m($j = N-1, \dots, 1$).

The optimal exercise boundary $r^*(t_m)$ can be located by the following property:

$$u(r, t_m) > G(r, t_m), \quad when \quad 0 \le r < r^*(t_m),$$

$$u(r, t_m) = G(r, t_m), \quad when \quad r^*(t_m) \le r \le R.$$

Since the Brennan-Schwartz algorithm works for our problem, the numerical results computed by it will be taken as the benchmark and compared with the results obtained by our penalty method.

3.5 Numerical Results

Let us illustrate the accuracy of our penalty method and compare it with the Brennan-Schwartz algorithm using some numerical examples. We use these two methods to solve the fully implicit scheme of the discrete form of the option valuation problem, i.e., let $\xi = 1$ in (3.2.30)-(3.2.31). In particular, we take $M = 1000, N = 500$ for the bond pricing equation and $M = 100, N = 500$ for the LCP, and compute the option value and the optimal exercise interest rate under three groups of parameters. These parameters of the CKLS model are estimated from market data (see Nowman and Sorwar [70]). Our software program is written in C++ and run on a personal computer with a Pentium IV 3.2 GHz processor.

Example 3.5.1: We consider a 5-year American put option on a 10-year zero-coupon bond with face value 100. The exercise price of the option is 65, i.e., $T = 5, T^* = 10, E = 100, K = 65$. Let other parameters be $R = 0.5, \kappa = 0.2436, \theta = 0.0542, \sigma = 1.4688, \gamma = 1.2660$. The penalty parameters are taken as $\vartheta = -10^6, k = 3$. We find that the results, i.e., the option value and the optimal exercise interest rate, obtained by these two methods are almost the same. Since the results computed by the Brennan-Schwartz algorithm are our benchmark, it indicates that our penalty approach provides accurate solution for the option value and optimal exercise interest rate.

The option value $V(r,t)$ and the difference between the option value and its payoff function, i.e., $V(r,t) - (K - B(r,t))_+$, computed by these two methods are presented in Figures 3.1-3.4. The discrepancy of the option value obtained by the two methods, i.e., the results by the penalty method minus the results by the Brennan-Schwartz algorithm, is shown in Figure 3.5. The optimal exercise interest rates from these two methods are compared in Figure 3.6. From Figure 3.1, we note that the option value has a concave shape for all t, which is the result of the fact that the bond price has a convex shape for all t. From Figure 3.5, we observe that the maximum discrepancy of the option value produced by these two methods is $1.5 * 10^{-4}$, namely, our penalty approach provides accurate option value. From Figure 3.6, we can see that the optimal exercise interest rate is accurate. It is in a concave shape, which is similar to the case under the CIR model (see Allegretto, Lin and Yang [4]).

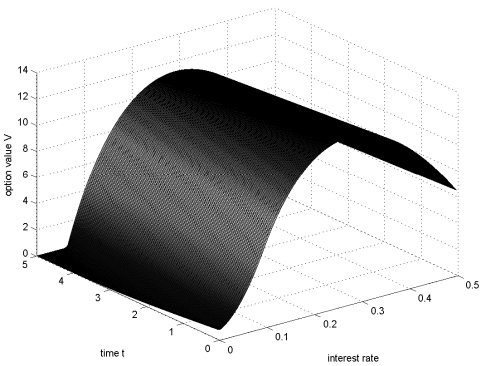

Figure 3.1: $V(r,t)$ from penalty approach

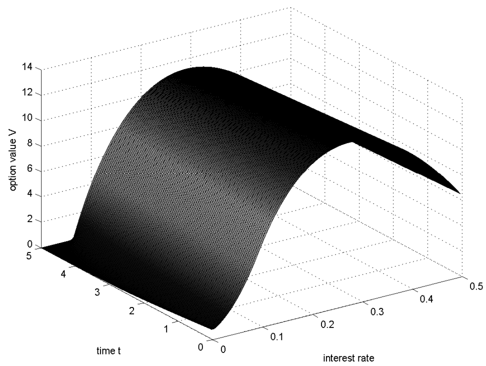

Figure 3.2: $V(r,t)$ from $B - S$ algorithm

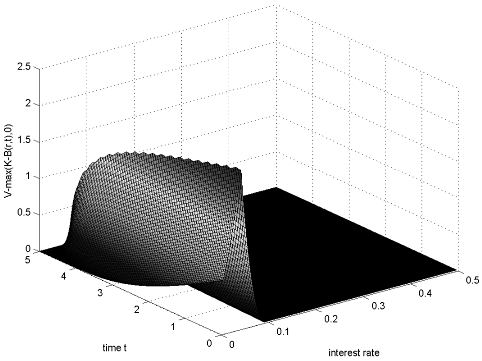

Figure 3.3: $[V(r,t) - (K - B(r,t))_+]_{penalty}$

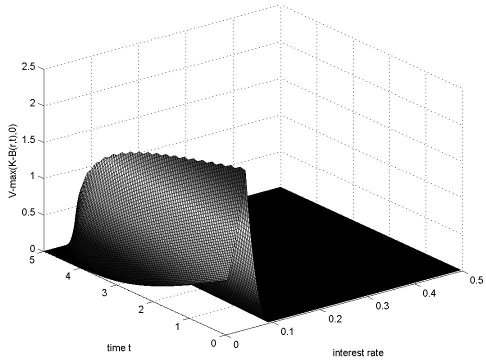

Figure 3.4: $[V(r,t) - (K - B(r,t))_+]_{B-S}$

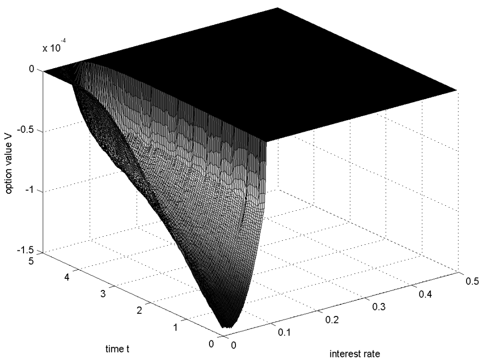

Figure 3.5: Difference of option value

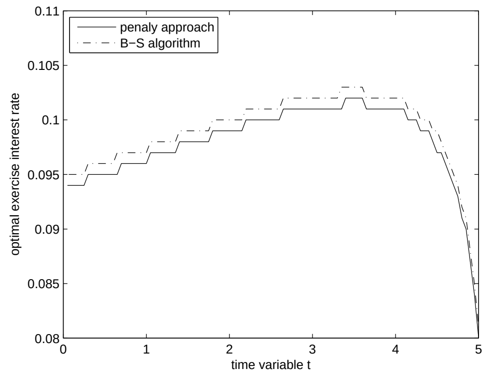

Figure 3.6: Optimal exercise interest rate

Example 3.5.2: We change the parameters κ, σ, γ of the American put option in Example 1, and let $\kappa = 0.3096, \sigma = 1.1124, \gamma = 1.1122$. We get similar results as in Example 3.5.1 (Figures are omitted).

Example 3.5.3: To demonstrate that our method is robust, we let $T = 1.0, T^* = 5, E = 100, R = 0.5, K = 60, \theta = 0.0625, \sigma = 0.2160, \kappa = 0.2880, \gamma = 0.3912$ in this example. The penalty parameters are still taken as $\vartheta = -10^6, k = 3$, and the numerical results from these two methods comply with each other with high accuracy.

The option value $V(r,t)$ and the difference between the option value and its payoff function, i.e., $V(r,t) - (K - B(r,t))_+$, computed by the penalty approach and the Brennan-Schwartz algorithm are presented in Figures 3.7-3.10. The discrepancy of the option value found by these two methods is shown in Figure 3.11. The optimal exercise interest rates from these two methods are compared in Figure 3.12. Figure 3.7 shows that the option value in this case is monotone for all t, which is due to the fact that the bond price is monotone for all t. Figure 3.11 shows that the maximum discrepancy of option value is $1.5 * 10^{-4}$, i.e., our penalty approach provides accurate option value. Figure 3.12 indicates that the optimal exercise interest rate in this case is a monotone function of t.

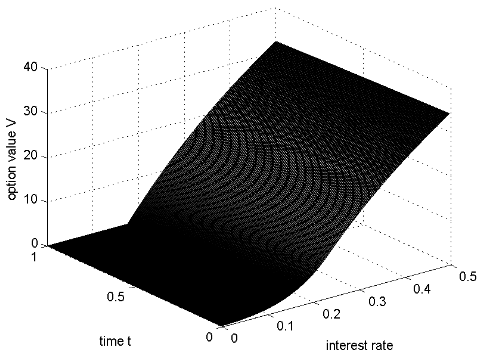

Figure 3.7: $V(r,t)$ from penalty approach

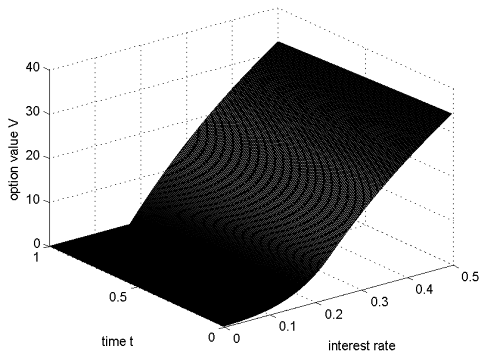

Figure 3.8: $V(r,t)$ from $B-S$ algorithm

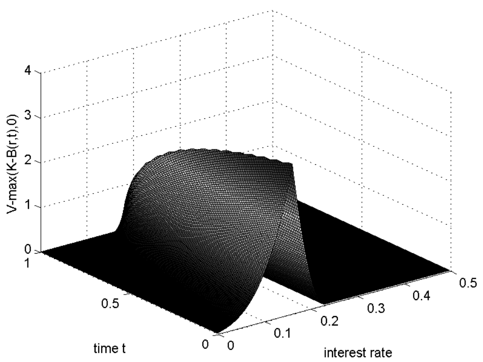

Figure 3.9: $[V(r,t) - (K - B(r,t))_+]_{penalty}$

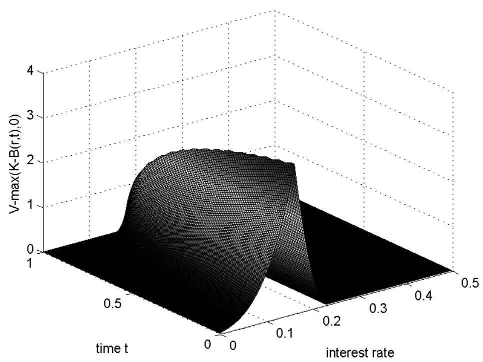

Figure 3.10: $[V(r,t) - (K - B(r,t))_+]_{B-S}$

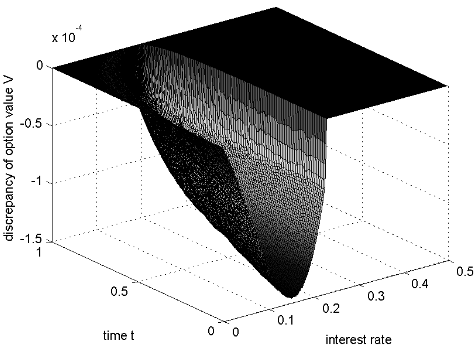

Figure 3.11: Difference of option value

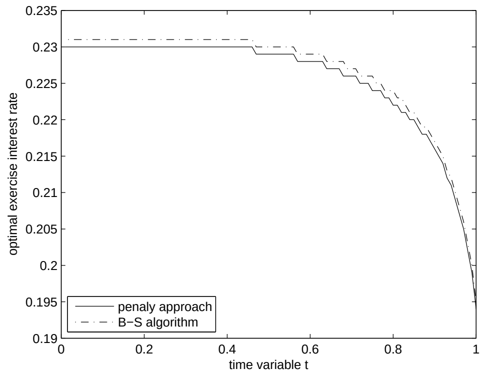

Figure 3.12: Optimal exercise interest rate

3.6 Summary

In this chapter, we have studied the American put options on zero-coupon bonds under the CKLS interest rate model. We used the practical parameters, which were estimated from the market data during February 1981-December 1997 for the CKLS model, to evaluate bonds and options. A novel power penalty method was proposed to solve the LCP for the option value and the optimal exercise boundary. The finite volume method was used to establish the numerical scheme, and the corresponding stability and convergence were given. Numerical examples have shown that our method is efficient and robust.

Chapter 4

Comparison of the LFM and the CIR Model

It is well known that caps and swaptions written on forward rates are largely traded contracts in the financial market. Black's formula for pricing caps is widely accepted by market participants. However, it provides exact cap price under the LFM with the precondition that the underlying forward rates are uncorrelated and martingales under their corresponding measures. When the underlying forward rates are correlated, Black's formula is not applicable any more.

In this chapter, we aim to compare the performance of the LFM and the CIR model in terms of pricing caps. These two models are the representatives of forward rate models and short rate models, and they are widely adopted in derivative pricing. Our "pricing under the LFM" means that we price derivatives starting from forward rate dynamics, and "pricing under the CIR model" means that we start from the CIR model to define forward rates and valuate derivatives. This chapter is organized as follows: Section 4.1 introduces the essential concepts of the LFM and Black's formula for caps. In Section 4.2, cap pricing formula under the CIR model derived based on European put options on zero-coupon bonds is reviewed. In Section 4.3, calibration of the CIR model and the LFM with $\rho_{i,j} = 0$ $(i \neq j)$ is carried out with market data of caps. The results indicate that the LFM with $\rho_{i,j} = 0$ is closer to in-the-sample market data than the CIR model. However, when it comes to predict out-of-the-sample cap prices, these two models have their own edges on different markets. In Section 4.4, we calibrate the LFM with an exogenous correlation matrix ($\rho_{i,j} = 1 - 0.025 * |i - j|$) to the market data of caps, and our results show that the LFM with $\rho_{i,j} = 1 - 0.025 * |i - j|$ is better in fitting the market data than the LFM with $\rho_{i,j} = 0$. We also calibrate the LFM with a parametric correlation matrix to the same market data. It manifests that the LFM with the parametric correlation matrix fits the data even better than the LFM with $\rho_{i,j} = 1 - 0.025 * |i - j|$. The calibration in this section is carried out by the downhill simplex method in multidimensions and Monte Carlo method. In Sections 4.5 and 4.6, we compare the prices of interest rate derivatives under the LFM ($\rho_{i,j} = 1$) with those under the CIR model using Monte Carlo method and artificial parameters. It is observed that they have underestimated prices under the former model than under the latter model. Section 4.7 summarizes this chapter.

4.1 The LFM and Black's Formula for Cap Pricing

Since the LIBOR is simply-compounded rate, we first introduce the concepts of simply-compounded spot rate and simply-compounded forward rate. Subsequently, we clarify the definitions of cap and caplet, the basic setup of LFM and Black's formula for caps (see Brigo and Mercurio [19]).

Definition 4.1.1 *The simply-compounded spot interest rate prevailing at time t*

for the maturity T is denoted by $L(t, T)$ and is the constant rate at which an investment has to be made to produce an amount of one unit of currency at maturity, starting from $P(t, T)$ units of currency at time t, when accruing occurs proportionally to the investment time. In formulas:

$$L(t, T) := \frac{1 - P(t, T)}{(T - t)P(t, T)}. \tag{4.1.1}$$

Definition 4.1.2 *The simply-compounded forward interest rate prevailing at time t for the expiry $T > t$ and maturity $T' > T$ is denoted by $F(t; T, T')$, and it is defined by*

$$F(t; T, T') := \frac{1}{T' - T}\left(\frac{P(t, T)}{P(t, T')} - 1\right), \tag{4.1.2}$$

where $T' - T$ is called the tenor of $F(t; T, T')$.

From the above definitions, it is easy to see $F(t; t, T) = L(t, T)$. Notice that $P(t, T)$ is the price of the zero-coupon bond at time t that guarantees its holder the payment of one unit of currency at maturity T. It is often used as a basic auxiliary quantity from which all interest rates can be recovered, and also appears in the definitions of $L(t, T)$ and $F(t; T, T')$. However, in the process of pricing interest rate derivatives under the LFM, $P(t, T)$ will be substituted by $\frac{1}{1+(T-t)L(t,T)}$ or $\frac{1}{1+(T-t)F(t;t,T)}$, because we do the pricing starting from forward rate dynamics. While pricing derivatives under a short rate model, we start from the short rate dynamics, and $P(t, T)$ is expressed as an analytical function of t, T and short rate r, then $P(t, T)$ is used for defining forward rates and pricing derivatives such as caps and options.

Definition 4.1.3 *A Payer (Forward-start) Interest Rate Swap (IRS) is a contract that exchanges payments between two parties, starting from a future time instant. At every instant T_i in a prespecified set of dates $\{T_{\alpha+1}, \ldots, T_\beta\}$ one party of the contract pays out the amount*

$$\widetilde{N}\tau_i X,$$

corresponding to a fixed interest rate X, a nominal value $\widetilde{N}$ and a year fraction τ_i between T_{i-1} and T_i, whereas the other party pays the amount

$$\widetilde{N}\tau_i L(T_{i-1}, T_i),$$

corresponding to the interest rate $L(T_{i-1}, T_i)$ (=$F(T_{i-1}; T_{i-1}, T_i)$) resetting at the previous instant T_{i-1} for the maturity T_i.

The discounted payoff at time $t < T_\alpha$ of the Payer IRS (abbreviated as PFS) can be expressed as

$$\widetilde{N} \sum_{i=\alpha+1}^{\beta} D(t, T_i)\tau_i(L(T_{i-1}, T_i) - X),$$

where $D(t, T_i)$ is the stochastic discount factor between the time instants t and T_i, which is taken as $P(t, T_i)$ for determinacy sometimes.

As a counterpart of the above PFS, there is a Receiver IRS (RFS) whose discounted payoff at time $t < T_\alpha$ is expressed as

$$\widetilde{N} \sum_{i=\alpha+1}^{\beta} D(t, T_i)\tau_i(X - L(T_{i-1}, T_i)).$$

Definition 4.1.4 *The forward swap rate $S_{\alpha,\beta}(t)$ at time t for the sets of times $\Gamma := \{T_\alpha, \ldots, T_\beta\}$ is the rate in the fixed leg of the IRS that makes the IRS a fair contract at the present time, i.e., it is the fixed rate X for which $RFS(t, \Gamma, \widetilde{N}, X) = 0$. Then, we obtain*

$$S_{\alpha,\beta}(t) = \frac{P(t, T_\alpha) - P(t, T_\beta)}{\sum\limits_{i=\alpha+1}^{\beta} \tau_i P(t, T_i)}.$$

Definition 4.1.5 *A cap is a contract that can be viewed as a payer IRS where each exchange payment is executed only if it has positive value. The discounted payoff of the cap at time $t < T_\alpha$ is given by*

$$\widetilde{N} \sum_{i=\alpha+1}^{\beta} D(t, T_i)\tau_i(L(T_{i-1}, T_i) - X)^+. \tag{4.1.3}$$

Definition 4.1.6 *A cap contract can be decomposed additively. Its discounted payoff is the sum of terms such as*

$$\widetilde{N} D(t, T_i)\tau_i(L(T_{i-1}, T_i) - X)^+, \tag{4.1.4}$$

each such term defines a contract that is termed as caplet.

Now we present the basic setup of the LFM. We denote $F(t; T_{i-1}, T_i)$ ($i = \alpha+1, \ldots, \beta$) as the forward rate observed at time t for a period starting at T_{i-1} and ending at T_i, and write $F_i(t)$ as the shorthand for $F(t; T_{i-1}, T_i)$. In the LFM framework, the forward rate $F_i(t)$ is assumed to follow a lognormal distribution under the forward measure Q^i (which is associated with the numeraire $P(\cdot, T_i)$, i.e., the price of the bond whose maturity coincides with the maturity of the forward rate $F_i(t)$, see Brace, Gatarek and Musiela [16]):

$$dF_i(t) = \sigma_i(t) F_i(t) dZ_i(t), \quad t \leq T_{i-1}. \tag{4.1.5}$$

In this equation, $dF_i(t)$ is the change in the forward rate, $F_i(t)$, in the time interval dt, Z_i is a standard Brownian motion under Q^i, and $\sigma_i(t)$ is the volatility coefficient for the forward rate $F_i(t)$. The "noises" in the dynamics of different forward rates are assumed to be instantaneously correlated according to

$$dZ_i(t)\ dZ_j(t) = \rho_{i,j} dt,$$

and $\rho_{i,j}$ is called the instantaneous correlation between $F_i(t)$ and $F_j(t)$ ($\rho_{i,i} = 1$ necessarily).

After setting the dynamics of all the $F_i(t)$, one can evaluate kinds of interest rate derivatives such as caps, swaptions and other options under the LFM. So far, people mainly use LFM with $Z_i(t)$ being a standard Brownian motion other than a Brownian motion vector. We call that one-factor LFM, which is adopted in this thesis except Section 5.5. When no specification is given, the LFM means the one-factor LFM in this thesis.

According to the definition, a cap can be evaluated by summing the prices of caplets under one uniform measure. It is a market practice to price caps by the following Black's formula (take the price of a cap at time zero for example)

$$\begin{aligned}
\widetilde{N} \sum_{i=\alpha+1}^{\beta} \mathrm{Cpl}(0, T_{i-1}, T_i, \tau_i, X) &= \widetilde{N} \sum_{i=\alpha+1}^{\beta} \tau_i \mathbb{E}\left\{D(0, T_i)(F(T_{i-1}; T_{i-1}, T_i) - X)^+\right\} \\
&= \widetilde{N} \sum_{i=\alpha+1}^{\beta} P(0, T_i)\tau_i \mathbb{E}(F(T_{i-1}; T_{i-1}, T_i) - X)^+ \\
&= \widetilde{N} \sum_{i=\alpha+1}^{\beta} P(0, T_i)\tau_i \mathrm{Bl}(X, F_i(0), v_i, 1) \\
&= \mathrm{Cap}^{\mathrm{Black}}(0, \Gamma, \widetilde{N}, X)
\end{aligned} \tag{4.1.6}$$

where $\Gamma = \{T_\alpha, \ldots, T_\beta\}$, $\mathrm{Cpl}(0, T_{i-1}, T_i, \tau_i, X)$ is the price at $t = 0$ of the caplet with unit principal and written on $F_i(t)$, $\mathbb{E}$ is the expectation under the risk-neutral measure, and

$$\begin{aligned}
\mathrm{Bl}(X, F_i(0), v_i, \omega) = F_i(0)\omega\Phi\left(\omega d_1(X, F_i(0), v_i)\right) - X\omega\Phi\left(\omega d_2(X, F_i(0), v_i)\right), \\
d_1(X, F_i(0), v_i) = \frac{\ln(F_i(0)/X) + v_i^2/2}{v_i}, \\
d_2(X, F_i(0), v_i) = \frac{\ln(F_i(0)/X) - v_i^2/2}{v_i}, \\
v_i = \sqrt{\int_0^{T_{i-1}} \sigma_i^2(t)dt}.
\end{aligned} \tag{4.1.7}$$

Here, $\mathrm{Bl}(X, F_i(0), v_i, \omega)$ with $\omega = 1$ corresponds to the caplet payoff at T_i, and Φ is the standard Gaussian cumulative distribution function.

The formula (4.1.6) and (4.1.7) are originally obtained by mimicking the Black-Scholes model for stock options and assuming that all forward rates in the cap follow driftless lognormal distributions under the risk-neutral measure. Because the forward rate $F_i(t)$ $(i = \alpha + 1, \ldots \beta)$ follows a driftless lognormal distribution under its corresponding forward measure Q^i in the LFM framework, $\mathbb{E}^i(F_i(T_{i-1}) - X)^+ = \mathrm{Bl}(X, F_i(0), v_i, 1)$ still holds, where $\mathbb{E}^i$ is the expectation under Q^i. Hence, Brigo and Mercurio [19] concluded that Black's formula provided an analytical solution for cap price under the LFM, i.e.,

$$\begin{aligned}
\mathrm{Cap}^{\mathrm{LFM}}(0,\Gamma,\widetilde{N},X) &= \widetilde{N}\sum_{i=\alpha+1}^{\beta} P(0,T_i)\tau_i\mathbb{E}^i(F_i(T_{i-1})-X)^+ \\
&= \widetilde{N}\sum_{i=\alpha+1}^{\beta} P(0,T_i)\tau_i \mathrm{Bl}(X,F_i(0),v_i,1) \\
&= \mathrm{Cap}^{\mathrm{Black}}(0,\Gamma,\widetilde{N},X).
\end{aligned}$$

However, we think that the first line of the above identity does not hold definitely. It is conditional on that the underlying forward rates are uncorrelated (i.e., $\rho_{i,j}=0$ for $i\neq j$). When the underlying forward rates are correlated (i.e., $\rho_{i,j}\neq 0$), Black's formula does not hold under the LFM any more. Because the caplet prices under different measures can not be summed directly to get the cap price now and "forward rates can not all be simultaneously log-normally distributed under a given pricing measure" (see Rebonato [77]). Therefore, we need to adopt Monte Carlo method or other methods to price caps under the LFM with $\rho_{i,j}\neq 0$.

As is well known, the correlation among forward rates is an important factor in pricing derivatives written on forward rates. Besides the above mentioned instantaneous correlation $\rho_{i,j}$, there is another concept "terminal correlation", which means correlation between forward rates at one payment date, e.g., correlation between F_i and F_{i+1} at time T_{i-1}. The relationship between terminal correlation and instantaneous correlation can be referred to Section 6.6 of Brigo and Mercurio [19]. In general, manipulating instantaneous correlation leads to manipulation of terminal correlation, although terminal correlation is also influenced by the way in which average volatility is distributed among instantaneous volatilities. There is an approximated formula for computing the terminal correlation of forward rates from instantaneous correlation and instantaneous volatilities (see Section 6.16 of Brigo and Mercurio [19]). Furthermore, we have that terminal correlations are, in absolute value, always less than or equal to instantaneous correlations according to the approximated formula.

When dealing with several caplets involving different forward rates, different structure of instantaneous volatilities can be employed. One can assume each forward rate has a constant instantaneous volatility or a deterministic volatility function of the time variable. Also, one can assume stochastic volatility for each forward rate. Some parametric forms for the instantaneous volatility are reviewed in Section 6.3.1 of Brigo and Mercurio [19]. The assumption of deterministic instantaneous volatility is a relatively simple case, which is adopted in Black's formula. We suppose $\sigma_i(t)$ to be constant σ_i in our study.

4.2 Cap Pricing Formula under the CIR Model

In this section, we first give the definition of short rate (i.e., continuously-compounded instantaneous rate), then we review cap pricing formula under the CIR model.

Definition 4.2.1 *We assume $Ba(t)$ is the value of a bank account at time $t \geq 0$ and it evolves according to the following differential equation*

$$dBa(t) = r(t)Ba(t)dt, \quad Ba(0) = 1.$$

As a consequence,

$$Ba(t) = \exp\left(\int_0^t r(s)ds\right),$$

where $r(t)$ is the instantaneous rate at which the bank account accrues. It is usually referred to as instantaneous spot rate, or briefly as short rate.

Based on this definition, a variety of short rate models have been proposed and used for pricing interest rate derivatives. The short rate model used here is the CIR model which guarantees the positivity and mean reversion of short rate, and it is analytically tractable (see Cox, Ingersoll and Ross [29]). In the model,

$r(t)$ satisfies the following stochastic differential equation under the risk-neutral measure

$$dr(t) = \kappa(\theta - r(t))dt + \sigma r(t)^{1/2} dW(t), r(0) = r_0,$$

where $r_0, \kappa, \theta, \sigma$ are positive constants, and $W(t)$ is a Wiener process. The condition

$$2\kappa\theta > \sigma^2$$

has to be imposed to ensure that the origin is inaccessible to the CIR process. The process $r(t)$ features a noncentral chi-squared distribution, whose density function and distribution function can be written out explicitly. Accordingly, the price at time t of the zero-coupon bond with maturity T and unit-principal is

$$P(t,T) = A(t,T)e^{-B(t,T)r(t)}, \tag{4.2.8}$$

where

$$A(t,T) = \left[\frac{2h \exp\{(\kappa + h)(T-t)/2\}}{2h + (\kappa + h)(\exp\{(T-t)h\} - 1)}\right]^{2\kappa\theta/\sigma^2}, \tag{4.2.9}$$

$$B(t,T) = \frac{2(\exp\{(T-t)h\} - 1)}{2h + (\kappa + h)(\exp\{(T-t)h\} - 1)}, \tag{4.2.10}$$

and $h = \sqrt{\kappa^2 + 2\sigma^2}$.

Remark 4.2.1 *It should be emphasized that $P(t,T)$ has a different formula under the LFM setting:*

$$P(t,T) = \frac{1}{1 + (T-t)F(t;t,T)}. \tag{4.2.11}$$

We need to use (4.2.8) or (4.2.11) at different place. That depends on the context.

Brigo and Mercurio showed that a cap (floor) is actually equivalent to a portfolio of European put (call) options on zero-coupon bonds, and this equivalence leads to explicit formulas for cap (floor) prices under those analytically tractable short rate models. Let ZBC($\cdot$) denote the price of the European call option with maturity T and strike K_i and written on a zero-coupon bond maturing at $s_i > T$. Then ZBC$(\cdot)$ has an explicit form under the CIR model according to Cox, Ingersoll and Ross [29]:

$$\begin{aligned}\mathrm{ZBC}(t,T,s_i,K_i) = P(t,s_i)\chi^2\left(2\bar{r}[\varrho+\psi+B(T,s_i)];\tfrac{4\kappa\theta}{\sigma^2},\tfrac{2\varrho^2 r(t)\exp\{(T-t)h\}}{\varrho+\psi+B(T,s_i)}\right)\\ -K_iP(t,T)\chi^2\left(2\bar{r}[\varrho+\psi];\tfrac{4\kappa\theta}{\sigma^2},\tfrac{2\varrho^2 r(t)\exp\{(T-t)h\}}{\varrho+\psi}\right),\end{aligned} \tag{4.2.12}$$

where

$$\begin{aligned}\varrho &:= \frac{2h}{\sigma^2(\exp\{(T-t)h\}-1)},\\ \bar{r} &:= \frac{\ln(A(T,s_i)/K_i)}{B(T,s_i)},\\ \psi &:= \frac{\kappa+h}{\sigma^2},\end{aligned}$$

and $\chi^2(x;k,\lambda)$ is the noncentral chi-square cumulative distribution function with degrees of freedom k and noncentrality parameter λ.

By the put-call parity formula for European options, the European put option on the same zero-coupon bond (denoting it by ZBP(t,T,s_i,K_i)) corresponding to the above European call option has price

$$\mathrm{ZBP}(t,T,s_i,K_i) = \mathrm{ZBC}(t,T,s_i,K_i) - P(t,s_i) + K_iP(t,T). \tag{4.2.13}$$

The ZBP$(\cdot)$ is an important part in the cap pricing formula under the CIR model.

As in Definition 4.1.3, we denote the set of the interest payment dates by $\{T_{\alpha+1},\ldots,T_\beta\}$, and denote T_α the first reset date. The price of the caplet with unit principal and written on $F_i(t)$ is given in Brigo and Mercurio [19]:

$$\begin{aligned}\mathrm{Cpl}(t,T_{i-1},T_i,\tau_i,X) &= \mathbb{E}(e^{-\int_t^{T_i} r(s)ds}\tau_i[L(T_{i-1},T_i)-X]^+|\mathcal{F}_t)\\ &= \mathbb{E}\left(e^{-\int_t^{T_{i-1}} r(s)ds}P(T_{i-1},T_i)\tau_i[L(T_{i-1},T_i)-X]^+|\mathcal{F}_t\right),\end{aligned} \tag{4.2.14}$$

where $\mathbb{E}$ is the expectation under the risk-neutral measure, and $\mathcal{F}_t$ is the filtration generated by the Brownian motion $W(t)$ in the CIR model. By the definition of $L(T_{i-1}, T_i)$, it is easy to obtain

$$\begin{aligned}
\mathrm{Cpl}(t, T_{i-1}, T_i, \tau_i, X) &= \mathbb{E}\left(e^{-\int_t^{T_{i-1}} r(s)ds} P(T_{i-1}, T_i)\left[\tfrac{1}{P(T_{i-1},T_i)} - 1 - X\tau_i\right]^+ |\mathcal{F}_t\right) \\
&= \mathbb{E}\left(e^{-\int_t^{T_{i-1}} r(s)ds}\left[1 - (1 + X\tau_i)P(T_{i-1}, T_i)\right]^+ |\mathcal{F}_t\right) \\
&= (1 + X\tau_i)\mathbb{E}\left(e^{-\int_t^{T_{i-1}} r(s)ds}\left[\tfrac{1}{1+X\tau_i} - P(T_{i-1}, T_i)\right]^+ |\mathcal{F}_t\right).
\end{aligned} \tag{4.2.15}$$

The last expectation in the above formula is actually the price of a European put option with maturity T_{i-1} and strike $\frac{1}{1+X\tau_i}$, and written on a unit-principal zero-coupon bond maturing at T_i. Therefore,

$$\mathrm{Cpl}(t, T_{i-1}, T_i, \tau_i, X) = (1 + X\tau_i)\mathrm{ZBP}(t, T_{i-1}, T_i, X_i'),$$

where $X_i' = \frac{1}{1+X\tau_i}$.

Now the cap with nominal value $\widetilde{N}$ and strike X written on $F_i(t)$ ($i = \alpha + 1, \ldots, \beta$) has the following price at time t

$$\mathrm{Cap}(t, \Gamma, \widetilde{N}, X) = \sum_{i=\alpha+1}^{\beta} \widetilde{N}(1 + X\tau_i)\mathrm{ZBP}(t, T_{i-1}, T_i, X_i'). \tag{4.2.16}$$

It is known that the above ZBP$(\cdot)$ has an analytical expression under the CIR short rate model. Consequently, the cap can be priced analytically under the CIR model by combining (4.2.16), (4.2.13) and (4.2.12). More precisely, the cap price at time 0 is a nonlinear function of the resetting dates Γ with parameters κ, θ, σ, and r_0:

$$\begin{aligned}
&\mathrm{Cap}(0, \Gamma, \widetilde{N}, X) = \\
&\widetilde{N}(1 + X\tau_i)\sum_{i=\alpha+1}^{\beta}\Big\{P(0, T_i)\chi^2\left(2\bar{r}_i[\varrho_i + \psi + B(T_{i-1}, T_i)]; \tfrac{4\kappa\theta}{\sigma^2}, \tfrac{2\varrho_i^2 r(0)\exp\{T_{i-1}h\}}{\varrho_i+\psi+B(T_{i-1},T_i)}\right) - \\
&\tfrac{1}{1+X\tau_i}P(0, T_{i-1})\chi^2\left(2\bar{r}_i[\varrho_i + \psi]; \tfrac{4\kappa\theta}{\sigma^2}, \tfrac{2\varrho_i^2 r(0)\exp\{T_{i-1}h\}}{\varrho_i+\psi}\right) - P(0, T_i) + \tfrac{1}{1+X\tau_i}P(0, T_{i-1})\Big\},
\end{aligned} \tag{4.2.17}$$

where $\psi = (\kappa + h)/\sigma^2$, $h = \sqrt{\kappa^2 + 2\sigma^2}$, and

$$\varrho_i = \frac{2h}{\sigma^2(\exp\{(T_{i-1} - t)h\} - 1)}, \tag{4.2.18}$$

$$\bar{r}_i = \frac{\ln[(1 + X\tau_i)A(T_{i-1}, T_i)]}{B(T_{i-1}, T_i)}, \tag{4.2.19}$$

in which the functions $A(T_{i-1}, T_i)$ and $B(T_{i-1}, T_i)$ are referred to (4.2.9) and (4.2.10).

4.3 Calibrating the LFM with $\rho_{i,j} = 0$ $(i \neq j)$ and the CIR Model to Market Data of Caps

We have known that the cap price at time 0 "Cap$(0, \Gamma, \widetilde{N}, X)$" is a function of the resetting dates Γ with parameters σ_i under the LFM with $\rho_{i,j} = 0$, see (4.1.6) and (4.1.7). Meanwhile, it is a function of Γ with parameters $\kappa, \theta, \sigma, r_0$ under the CIR model, see (4.2.17)-(4.2.19). Because of these explicit formulas, these two models have been extensively used for interest rate derivative pricing, while their performances are seldom compared. We fill in the gap by calibrating them to the market data of caps.

In this section, we calibrate these two models to the market data of at-the-money (ATM) caps. The analytical expressions for cap price facilitate much the calibration of these models. We can fit their price function curves to the market data by least square curve fitting function in Matlab. Here, ATM caps are adopted because their data are easy to obtain. We first introduce the concept of ATM cap, then elaborate the calibration experiments.

Definition 4.3.1 *The cap with payment times $T_{\alpha+1}, \ldots, T_\beta$ $(T_\alpha > 0)$ and strike*

X is said to be at-the-money (ATM) if and only if

$$X = X_{ATM} := S_{\alpha,\beta}(0) = \frac{P(0,T_\alpha) - P(0,T_\beta)}{\sum\limits_{i=\alpha+1}^{\beta} \tau_i P(0,T_i)}.$$

The cap is instead said to be in-the-money (ITM) if $X < X_{ATM}$, and out-of-the-money (OTM) if $X > X_{ATM}$.

In the calibration of each model, the first step is to find the interest rate model parameters by minimizing the sum of the square percentage differences between model-giving prices and market cap prices. It is achieved by the curve fitting function in Matlab, which provides the final residual and fitted parameters as results. Therefore, the second step is to calculate the prices of market caps, including out-of-the-sample caps, with the fitted parameters. These prices are called model implied cap prices. Here, out-of-the-sample caps are those caps whose data is not used in working out the model parameters. In addition, we take the model implied cap prices into Black's formula to retrieve the implied volatility associated with each cap maturity, and compare the model implied volatilities with the quoted market volatilities.

The market typically quotes volatilities for caps with first resetting date either in three months or in six months, and progressively increasing maturities (e.g., $\alpha = 0$, T_0 equals to six months and all other T_i are equally six-months spaced). For the cap with maturity T_j, an equation is considered between the model implied price of the cap and the sum of the first j caplets prices:

$$\mathrm{Cap}^{\mathrm{Model}}(0, \{T_0, T_1, \ldots, T_j\}, 1, X) = \sum_{i=1}^{j} P(0,T_i)\tau_i \mathrm{Bl}(X, F_i(0), \sqrt{T_{i-1}}\sigma_{T_j-cap}, 1),$$

where the same constant volatility σ_{T_j-cap} has been put in all caplets up to j. The quantities σ_{T_j-cap} are called **model implied volatilities**.

The final residual of square percentage differences between model implied prices and market cap prices indicates which model is closer to the market practice. All our market data are quoted from Bloomberg system and the details of numerical experiments are presented as follows.

Calibration Example 4.3.1: We use the market data on Aug 10, 2010 of ATM (at-the-money) caps written on US forward rates, which include volatilities, cap rates (also called strike) and closing prices of caps with different maturities (see Table 4.1, where the market volatilities are actually computed by the Bloomberg system). The notional value of these caps is 1$, and the unit of cap price is $ (US dollar). The underlying interest rates of these caps are US forward rates with tenors being 3 months. Notice that the first underlying forward rate of these caps has the expire date Nov 10, 2010, and the expire dates of other forward rates are equally 3-month spaced. We need the values of these forward rates on Aug 10, 2010, see Table 4.2 (where the values are annual rates). These initial forward rates are necessary for the calibration. The numbers of forward rates and cap rates are in percentages. We take at most four digits after the decimal point as significant digit.

Table 4.1: Cap rate, volatility and price of ATM US caps on Aug 10, 2010

Maturity (year)	Cap rate (%)	Volatility	Price ($)
1	0.477	0.7993	0.0009
2	0.743	0.7594	0.0045
3	1.067	0.6856	0.0115
4	1.391	0.5764	0.0204
5	1.703	0.4915	0.0307
6	2.01	0.4255	0.0425
8	2.478	0.3591	0.0666
10	2.782	0.3215	0.0887

Table 4.2: US forward rates with tenor = 3M on Aug 10, 2010

Expire date	Forward rate (%)	Expire date	Forward rate (%)
11/10/2010	0.3703	11/10/2015	3.5936
02/10/2011	0.4798	02/10/2016	3.7307
05/10/2011	0.5553	05/10/2016	3.8644
08/10/2011	0.6661	08/10/2016	3.802
11/10/2011	0.8283	11/10/2016	3.9168
02/10/2012	1.011	02/10/2017	4.0255
05/10/2012	1.1929	05/10/2017	4.1318
08/10/2012	1.3817	08/10/2017	4.0057
11/13/2012	1.5762	11/10/2017	4.0961
02/11/2013	1.7567	02/12/2018	4.1816
05/10/2013	1.943	05/10/2018	4.2605
08/12/2013	2.1555	08/10/2018	4.1318
11/12/2013	2.3436	11/13/2018	4.2006
02/10/2014	2.5053	02/11/2019	4.2643
05/12/2014	2.6678	05/10/2019	4.3266
08/11/2014	2.8467	08/12/2019	4.2378
11/10/2014	3.0024	11/12/2019	4.2912
02/10/2015	3.1552	02/10/2020	4.3408
05/11/2015	3.3047	05/11/2020	4.3878
08/10/2015	3.4505		

By the least square curve fitting function in Matlab, the parameters of the CIR model are figured out as $\kappa = 0.0978, r_0 = 0.0004, \theta = 0.08, \sigma = 0.1176$, and the final residual is 0.3508. The CIR model implied cap prices and implied volatilities are then computed with the fitted parameters. When fitting the LFM to the data, we also use 4 parameters, but there are actually 39 unknown σ_i. Therefore, we set $\sigma_1 = \cdots = \sigma_9$, $\sigma_{10} = \cdots = \sigma_{19}$, $\sigma_{20} = \cdots = \sigma_{29}$, and $\sigma_{30} = \cdots = \sigma_{39}$, which can be seen as the piecewise constant structure for volatilities (see Brigo and Mercurio [19]). Now we have 4 volatility parameters to calibrate the LFM, and they are obtained as 0.8507, 0.3419, 0.2869, 0.2453. The corresponding final residual is 0.0011.

We use the calibrated parameters for computing the prices of in-the-sample caps (i.e., the caps on Aug 10, 2010) and out-of-the-sample caps (i.e., the caps on other dates). For in-the-sample caps, we can easily find that the LFM implied cap prices are closer to the market prices than the CIR model implied prices (see Figure 4.1). For out-of-the-sample caps, neither the LFM implied prices nor the CIR implied prices are very near to the market prices as a whole (see Figures 4.2-4.4). But we can see that the LFM implied prices for caps with short maturities (1, 2, 3 years) are accurate relative to the corresponding CIR implied prices. In addition, we put the implied volatilities of these two models together with the quoted market volatilities on Aug 10, 2010, see Figure 4.5. It is observed that the quoted market volatility is decreasing with respect to cap maturity. The CIR model implied volatility has similar property to the quoted market volatility, while the LFM implied volatility is in a humped shape with respect to cap maturity. This kind of implied volatility curve often occurs in the market practice.

Calibration Example 4.3.2: In this example, we use the market data of ATM Euro forward rate caps on Aug 10, 2010, see Table 4.3. The notional value of these caps is 1 Euro, and the unit of cap price is €(Euro). The underlying interest rates of these caps are Euro forward rates with tenor being 6 months. The first underlying forward rate of these caps has the expire date Feb 10, 2011, and the expire dates of other forward rates are equally 6-month spaced. The values of these forward rates on Aug 10, 2010 are given in Table 4.4, where the values are annual rates. The numbers of forward rates and cap rates are in percentages.

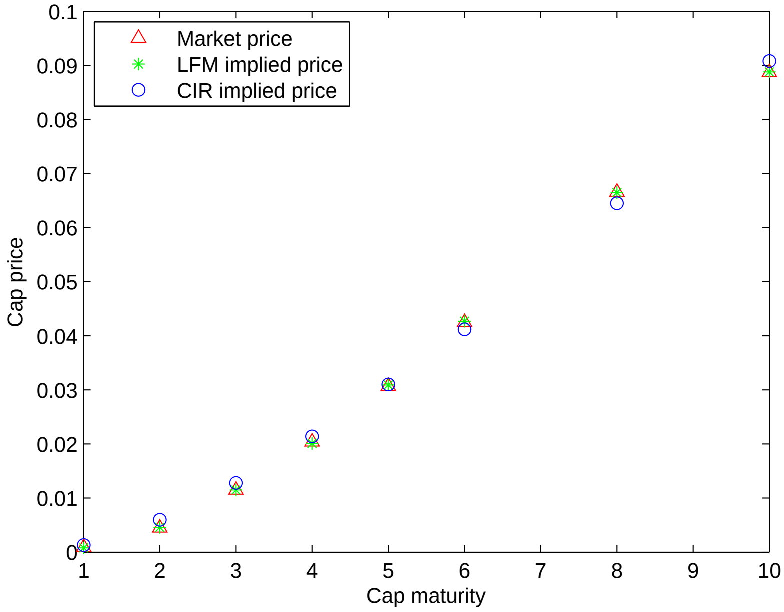

Figure 4.1: Market and model implied prices of ATM US caps on Aug 10, 2010

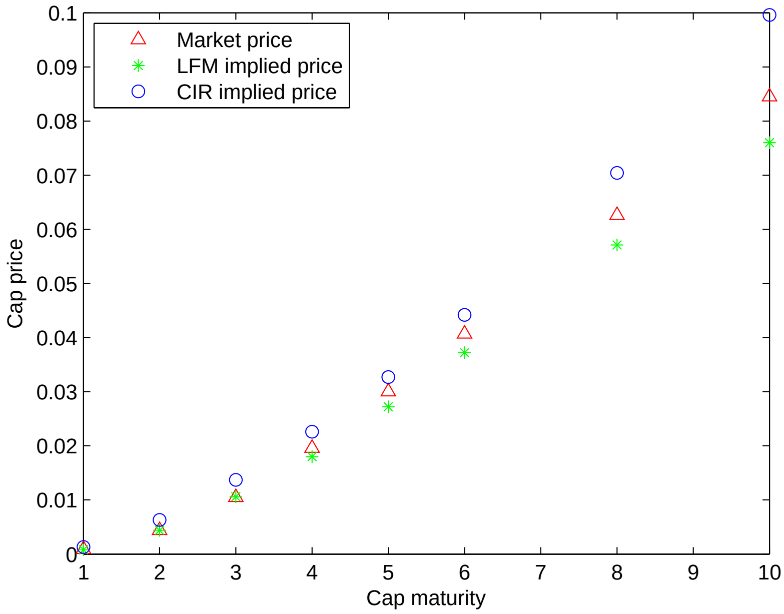

Figure 4.2: Market and model implied prices of ATM US caps on Aug 31, 2010

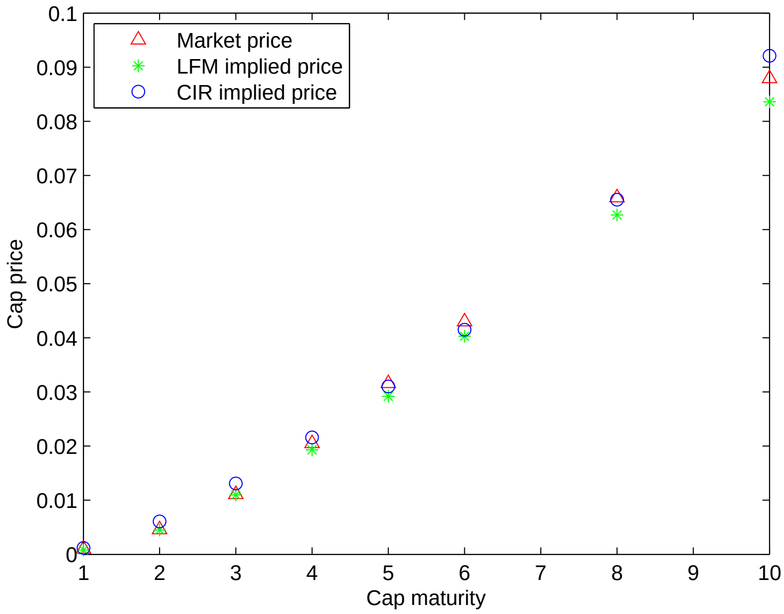

Figure 4.3: Market and model implied prices of ATM US caps on Sep 15, 2010

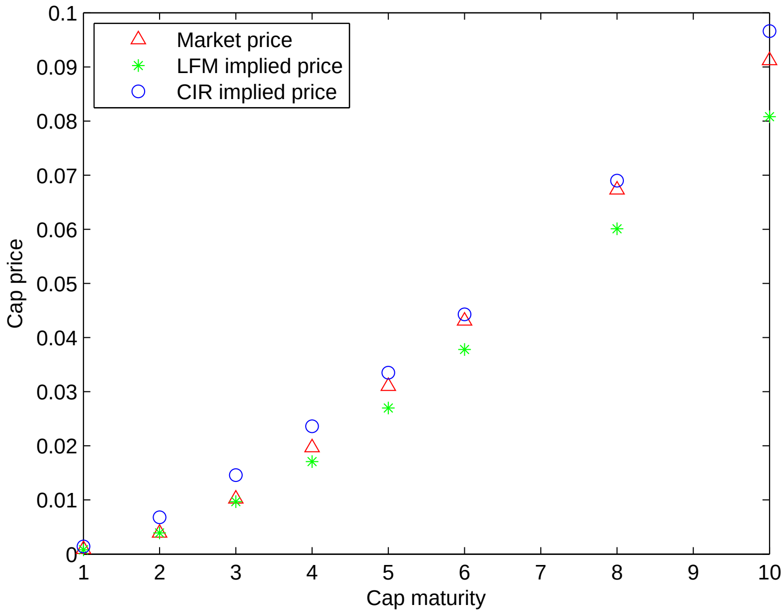

Figure 4.4: Market and model implied prices of ATM US caps on Sep 30, 2010

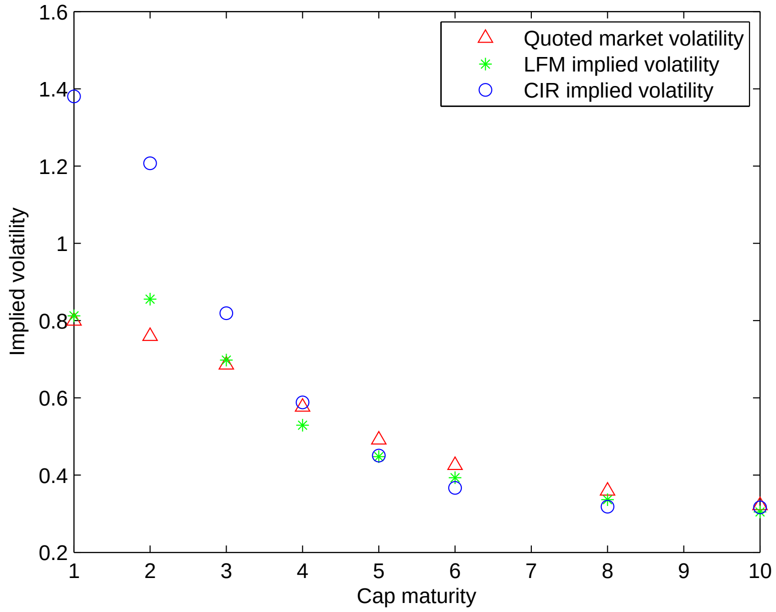

Figure 4.5: Market and model implied volatilities of ATM US caps on Aug 10, 2010

Table 4.3: Cap rate, volatility and price of ATM Euro caps on Aug 10, 2010

Maturity (year)	Cap rate (%)	Volatility	Price (€)
1	1.08	0.4797	0.0011
2	1.23	0.5421	0.0047
3	1.63	0.4684	0.0095
4	1.87	0.4358	0.0164
5	2.09	0.3918	0.0243
6	2.3	0.3477	0.0327
8	2.62	0.296	0.0496
10	2.85	0.2641	0.0658
15	3.2	0.2203	0.1025
20	3.32	0.2044	0.1322

Table 4.4: Euro forward rates with tenor = 6M on Aug 10, 2010

Expire date	Forward rate (%)	Expire date	Forward rate (%)
02/10/2011	1.299	02/10/2021	3.9532
08/10/2011	1.428	08/10/2021	4.0562
02/10/2012	1.6136	02/10/2022	4.1098
08/10/2012	1.8299	08/10/2022	4.0071
02/11/2013	2.0035	02/10/2023	4.0453
08/12/2013	2.3522	08/10/2023	4.0815
02/10/2014	2.5589	02/12/2024	4.112
08/11/2014	2.7828	08/12/2024	4.1401
02/10/2015	2.9777	02/10/2025	4.1635
08/10/2015	3.1079	08/11/2025	3.7069
02/10/2016	3.2776	02/10/2026	3.7127
08/10/2016	3.4084	08/10/2026	3.7181
02/10/2017	3.5558	02/10/2027	3.7208
08/10/2017	3.6379	08/10/2027	3.7231
02/12/2018	3.7595	02/10/2028	3.7232
08/10/2018	3.6428	08/10/2028	3.7231
02/11/2019	3.7333	02/12/2029	3.7202
08/12/2019	3.8511	08/10/2029	3.7175
02/10/2020	3.9305	02/11/2030	3.7125
08/10/2020	3.8915		

In this case, the parameters of the CIR model are fitted out as $\kappa = 0.0663, r_0 = 0.0024, \theta = 0.0597, \sigma = 0.2706$, and the final residual is 0.0910. The 4 volatility parameters of the LFM (same as that in calibration example 1) are fitted out as 0.4106, 0.2350, 0.1696, 0.1811, and the corresponding final residual is 0.0800. The calibrated parameters are then used for computing the prices of in-the-sample caps (i.e., caps on Aug 10, 2010) and out-of-the-sample caps (i.e., caps on other dates). For in-the-sample caps, the LFM implied cap prices are closer to the market prices compared with the CIR model implied prices (see Figure 4.6). For out-of-the-sample caps, the CIR model produces implied prices nearer to the market prices than the LFM (see Figures 4.7-4.9). The implied volatilities of the two models and the quoted market volatilities on Aug 10, 2010 are shown in Figure 4.10. It can be seen that the quoted market volatility of the Euro ATM cap is in a humped shape with respect to the cap maturity. Both the LFM implied volatility and the CIR implied volatility have similar shape.

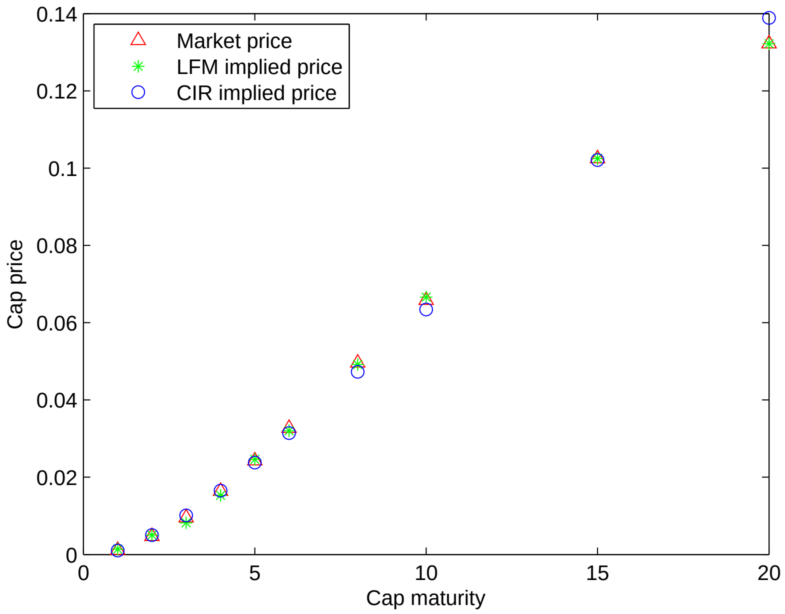

Figure 4.6: Market and model implied prices of ATM Euro caps on Aug 10, 2010

In general, these two examples show that the LFM with zero correlations fits the in-the-sample market data better than the CIR model, which affirms the significance of Black's formula for cap pricing. However, when it comes to the predicting capability for out-of-the-sample products, these two models have their own strong points and weak points on different markets. The CIR model is better

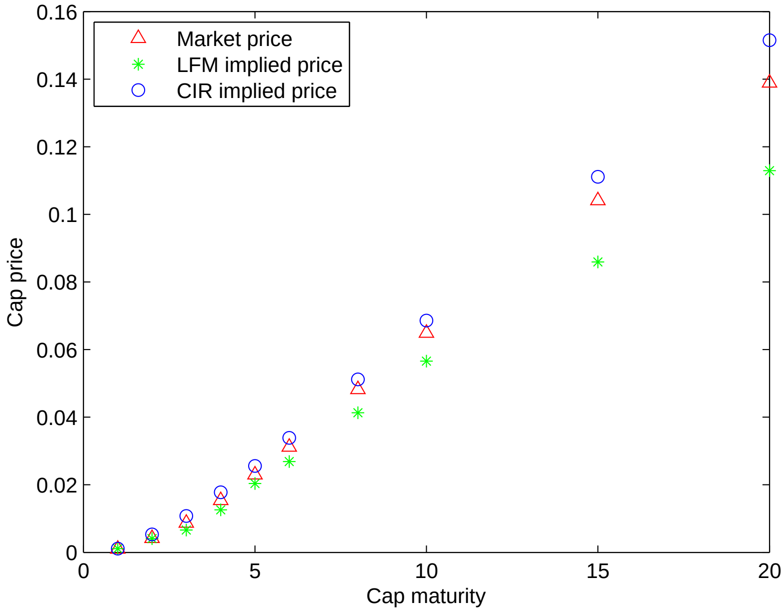

Figure 4.7: Market and model implied prices of ATM Euro caps on Aug 31, 2010

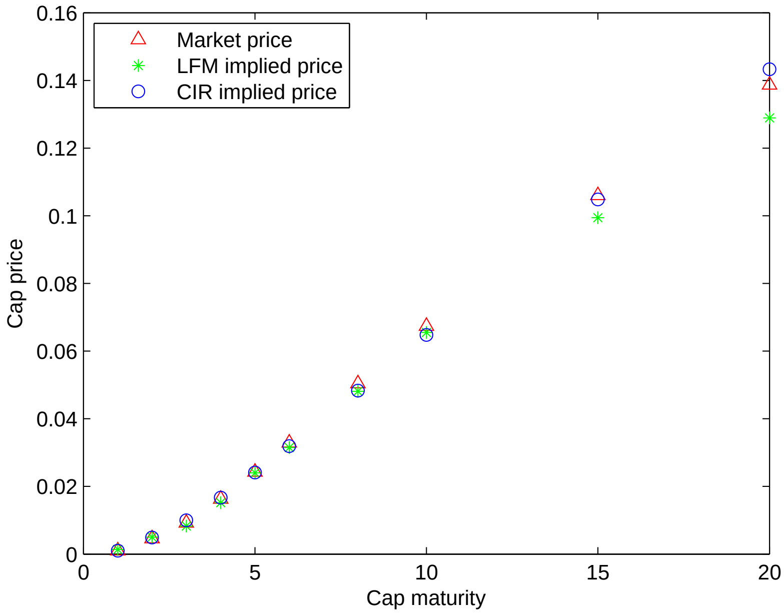

Figure 4.8: Market and model implied prices of ATM Euro caps on Sep 15, 2010

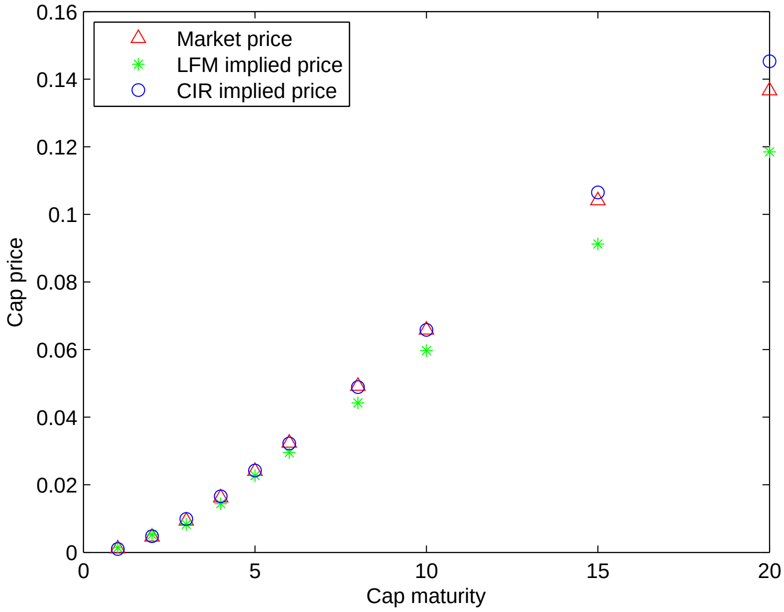

Figure 4.9: Market and model implied prices of ATM Euro caps on Sep 30, 2010

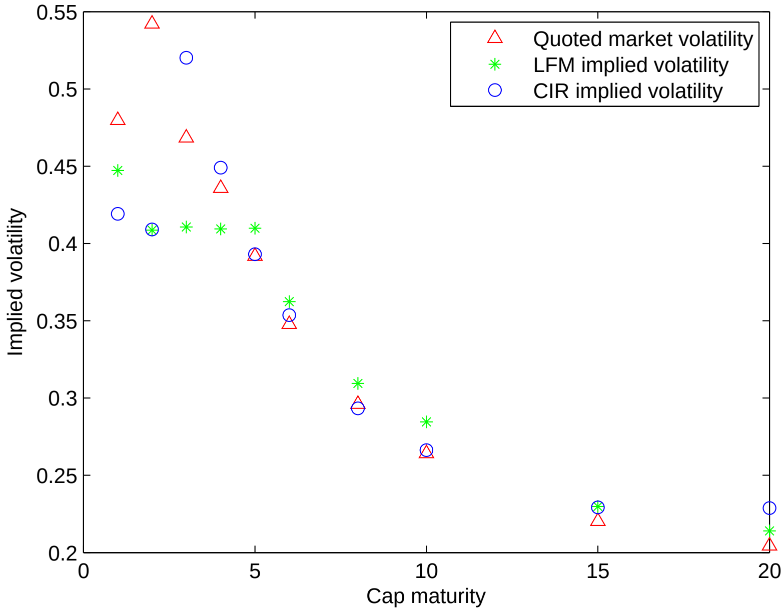

Figure 4.10: Market and model implied volatilities of ATM Euro caps on Aug 10, 2010

at predicting the ATM Euro caps, while the LFM with zero correlation is more suitable for predicting the ATM US caps. In the calibration experiments, we use 4 volatility parameters in the LFM. If we use more parameters for the calibration of the LFM, the calibration effect might be better.

4.4 Calibrating the LFM with $\rho_{i,j} \neq 0$ to Market Data of Caps

We have calibrated the LFM with $\rho_{i,j} = 0$ to market data in the last section. Black's formula is used there for computing cap prices, which brings much convenience to the calibration process. However, the assumption $\rho_{i,j} = 0$ is not reasonable. The LFM with $\rho_{i,j} \neq 0$ should be more practical than the LFM with $\rho_{i,j} = 0$. To confirm this conjecture, now we implement experiments by calibrating the LFM with $\rho_{i,j} \neq 0$ to the same market data, and compare the results with those in the last section.

The calibration of the LFM with $\rho_{i,j} \neq 0$ is carried out by Monte Carlo method here. In other words, Monte Carlo method is used for computing cap prices in the calibration (Monte Carlo procedures for pricing caps under the LFM are demonstrated in Section 5.2). In the optimization process, the downhill simplex method in multidimensions (see Nelder and Mead [69], Press et al [74]) is applied to minimize the square percentage differences between the market prices and simulated LFM prices of caps, because this method requires only function evaluations, not derivatives. The derivatives of cap prices are very computationally expensive to obtain by Monte Carlo method.

Nevertheless, we want to make clear that the calibration result of the LFM parameters in each experiment is not consistent even we use the same data and initial guess. This is because of the random simulation for forward rates and cap prices. We choose the best calibration result (i.e., the minimum final residual) from multiple trials and present it here. We use C++ program for this calibration on a PC with Intel i5 processor (2 core @ 3.33G) and 4GB memory. It is easy to understand that the calibration process is time consuming because of the Monte Carlo simulation. We present the numerical examples as below.

Example 4.4.1: In this test, we use the market data of ATM US caps, which are referred to Table 4.1 and Table 4.2. We first use the same 4 volatility parameters as in the previous section for calibrating the LFM with $\rho_{i,j} = 1 - 0.025 * |i-j|$, which is an exogenous correlation matrix satisfying the properties of an instantaneous correlation matrix (see Brigo and Mercurio [19]). Through the calibration process, the 4 volatility parameters are figured out as 0.8325, 0.3447, 0.2958, 0.2150. The model implied prices and market prices are compared in Table 4.5, where "LFM 0" represents the LFM with $\rho_{i,j} = 0$, and "LFM 1" represents the LFM with $\rho_{i,j} = 1-0.025*|i-j|$. The residual of square percentage differences between the LFM ($\rho_{i,j} = 1 - 0.025 * |i-j|$) implied cap prices and the market prices is 0.0009, while the residual of square percentage differences between the LFM ($\rho_{i,j} = 0$) implied cap prices and the market prices is 0.0011. It suggests that the LFM with $\rho_{i,j} = 1 - 0.025 * |i-j|$ is better than the LFM with $\rho_{i,j} = 0$ in fitting the market data.

One can also take a parametric correlation matrix in the calibration to avoid the perturbation from an exogenous correlation matrix. Here, we apply the classical two-parameter correlation matrix (see Rebonato [75]) in our calibration:

$$\rho_{i,j} = \rho_\infty + (1-\rho_\infty)\exp[-\tilde{\beta}|i-j|\ ], \quad 0 \le \rho_\infty \le 1, \quad \tilde{\beta} \ge 0,$$

where $\rho_{i,j}$ denotes the instantaneous correlation between the forward rates $F_i(t)$ and $F_j(t)$, ρ_∞ and $\tilde{\beta}$ are parameters. Including the above 4 volatility parameters, we have 6 parameters now. Calibrating to the ATM US caps, we get the 4 volatility parameters as 0.8432, 0.3517, 0.2550, 0.2536, and $\rho_\infty = 0.5708$, $\tilde{\beta} = 0.7159$. The LFM (with parametric correlations) implied prices are also shown in Table 4.5. They are put in the column denoted "LFM 2". In this case, the residual of square percentage differences between the LFM (with parametric correlations) implied cap prices and the market prices is 0.0005. This tells us that the LFM with parametric correlations is even better than the LFM with $\rho_{i,j} = 1 - 0.025 * |i - j|$ in terms of fitting capability.

Table 4.5: Market and LFM implied prices of ATM US caps on Aug 10, 2010 ($)

Maturity (year)	LFM 1	LFM 2	Market	LFM 0
1	0.0008	0.0008	0.0009	0.0008
2	0.0046	0.0045	0.0045	0.0046
3	0.0115	0.0114	0.0115	0.0115
4	0.0200	0.0202	0.0204	0.0201
5	0.0309	0.0313	0.0307	0.0309
6	0.0427	0.0428	0.0425	0.0427
8	0.0668	0.0663	0.0666	0.0665
10	0.0877	0.0893	0.0887	0.0888

Example 4.4.2: Now we employ the market data of ATM Euro caps, which are referred to Table 4.3 and Table 4.4. As in Example 5.1, we first calibrate the LFM with $\rho_{i,j} = 1-0.025*|i-j|$. Then we calibrate the LFM with the parametric correlation matrix, which leads to 6 parameters in the fitting process. The results are shown in Table 4.6. The 4 parameters in "LFM 1" are figured out as 0.4074, 0.2492, 0.1373, 0.2155. The residual of square percentage differences between the LFM ($\rho_{i,j} = 1 - 0.025 * |i - j|$) implied cap prices and the market prices is 0.0730, while the residual of square percentage differences between the LFM ($\rho_{i,j} = 0$) implied cap prices and the market prices is 0.0800. In the calibration of

"LFM 2", we get the 4 volatility parameters as 0.4211, 0.2099, 0.2198, 0.1424, and $\rho_\infty = 0.2554$, $\tilde{\beta} = 0.2692$. The residual of square percentage differences between the LFM (with parametric correlations) implied cap prices and the market prices is 0.0719. These results indicate that the LFM with $\rho_{i,j} = 1 - 0.025 * |i - j|$ is better than the LFM with $\rho_{i,j} = 0$, and the LFM with parametric correlations is even better than the LFM with $\rho_{i,j} = 1 - 0.025 * |i - j|$ in fitting the market data.

Table 4.6: Market and LFM implied prices of ATM Euro caps on Aug 10, 2010 (€)

Maturity (year)	LFM 1	LFM 2	Market	LFM 0
1	0.0013	0.0013	0.0011	0.0014
2	0.0050	0.0050	0.0047	0.0050
3	0.0082	0.0084	0.0095	0.0082
4	0.0152	0.0157	0.0164	0.0153
5	0.0247	0.0250	0.0243	0.0246
6	0.0325	0.0320	0.0327	0.0320
8	0.0508	0.0481	0.0496	0.0492
10	0.0695	0.0640	0.0658	0.0666
15	0.1009	0.1071	0.1025	0.1025
20	0.1349	0.1314	0.1322	0.1322

In a word, the calibration tests in this section manifest that the LFM with nonzero correlations outperforms the LFM with zero correlations, and the LFM with parametric correlations outperforms the LFM with exogenous correlations in fitting the market data. Moreover, looking into the calibration results of "LFM 0" and "LFM 2" in Table 4.5 and Table 4.6, we can see that "LFM 2" does have its advantage in fitting the prices of caps with maturities less than 5 years. For the caps with longer maturities, taking the correlations among forward rates into account does not improve the calibration results. This can be explained by the fact that joint movements of far away rates are less correlated than movements of rates with close maturities, and correlations have weak effect on prices of the caps with long maturities.

4.5 Comparison of the CIR Model and the LFM for Cap Pricing

We have known that cap pricing under the LFM with $\rho_{i,j} \neq 0$ often resorts to Monte Carlo method, by which one can compute the expectation of the discounted payoff function of the cap. The details of the Monte Carlo method for valuing interest rate derivatives under the LFM will be introduced in the next chapter.

In this section, we compare the prices of caps under the LFM ($\rho_{i,j} = 1$) with those under the CIR model. In the comparison, we adopt $\rho_{i,j} = 1$ in the LFM because the CIR model implies perfectly correlated spot rates with different tenors, which can also be seen as forward rates at expiry dates. We know that the LFM and the CIR model have different parameters and initial values. For the comparability, we assume artificial parameters for the CIR model, and use them to produce the initial forward rates and volatilities for the LFM since the CIR model implies a forward rate model (4.5.21). Then these initial values are adopted in the LFM. We can conjecture that pricing caps under the CIR model or (4.5.21) should have the same results. Therefore, comparing the CIR model with the LFM turns into comparing (4.5.21) with the LFM. In the numerical experiment, we apply the analytical formulas (4.2.17)-(4.2.19) to evaluate caps under the CIR model, and employ Monte Carlo method to compute cap price under the LFM.

The parameters of the CIR model

$$dr(t) = \kappa(\theta - r(t))dt + \sigma r(t)^{1/2} dW(t), r(0) = r_0,$$

are taken as: $\kappa = 0.2, \theta = 0.1, \sigma = 0.06, r_0 = 0.1$.

From the short rate $r(t)$ to the forward rate $F(t; T, T')$, we have the following formula:

$$F(t; T, T') = \frac{1}{T' - T} \left(\frac{P(t, T)}{P(t, T')} - 1 \right). \tag{4.5.20}$$

Since $P(t, T)$ has the explicit form (4.2.8) under the CIR model, we can get the initial forward rate $F(0; T, T')$ by the formula (4.5.20). The forward rate dynamics implied by the CIR short rate dynamics can also be derived (see Brigo and Mercurio [19]):

$$\begin{aligned} dF(t; T, T') &= \sigma F(t; T, T') \left(1 + \tfrac{1}{(T'-T)F(t;T,T')}\right) \\ &\cdot \sqrt{(B(t, T') - B(t, T)) \ln \left[((T' - T)F(t; T, T') + 1) \tfrac{A(t,T')}{A(t,T)} \right]} dW^{T'}(t), \end{aligned} \tag{4.5.21}$$

where $W^{T'}$ is the Brownian motion under the forward measure $Q^{T'}$, the functions $A(t, T)$ and $B(t, T)$ are given in (4.2.8). This is different from the lognormal dynamics assumed for $F(t; T, T')$ under the LFM, where typically

$$dF(t; T, T') = \sigma(t) F(t; T, T') dW^{T'}(t).$$

This difference explains the distinct performance of the CIR model and the LFM for pricing interest rate derivatives. We illustrate the difference manifested in cap pricing by numerical examples.

We consider the caps with maturities being $1, 2, \ldots, 5$ years. Their nominal value is \$100, and the cap rates range from 0.08 to 0.10. The interest payments are made every 6 months. By (4.5.20) and (4.5.21), the corresponding initial forward rates $F_i(0; T_{i-1}, T_i)$ ($T_i = (i + 1) * 0.5, i = 1, 2, \ldots, 9$) are 0.1024, 0.1023, 0.1021, 0.1019, 0.1017, 0.1015, 0.1012, 0.1010, 0.1008, and the initial volatilities are 0.1674, 0.1512, 0.1365, 0.1231, 0.1110, 0.1000, 0.0901, 0.0811, 0.0730. In addition, $T_0 = 0.5$. We compute the cap prices under the LFM using the Monte Carlo simulation by C++ coding, and get the cap prices under the CIR model by Matlab programming. C++ is used for the Monte Carlo simulation because of its speed advantage.

Calculating a cap price from each forward rate path and simulating 10^5 paths (we set $dt = 0.005$ in the simulation), we get the cap prices under the LFM with the above varying volatilities. We also test the LFM with fixed initial volatility 0.06, and get cap prices far different from those under the CIR model (see Table 4.7, where the unit is \$). In Table 4.7, "LFM(fv)" means the LFM with fixed volatility 0.06 in all $F_i(t)$, and "LFM(vv)" means the LFM with the above varying volatilities for $F_i(t)$. It is observed that the cap prices under the LFM(vv) are higher than those under the LFM(fv), which is due to the term $[L(T_{i-1}, T_i) - X]^+$ in (4.2.14) plays a dominant role in this comparison. The varying volatilities are larger than the fixed volatility, and thus generate more fluctuating rates and higher cap prices under the LFM(vv). It is also observed that the cap prices under the CIR model are higher than those under the LFM(vv), which can be explained by the second line of (4.2.15). We deem that the discount factor $e^{-\int_t^{T_{i-1}} r(s)ds}$ plays a dominant role in this comparison. The discount factor under the CIR model is larger than that under the LFM(vv) because the integrand $r(s)$ under the CIR model is less than that under the LFM(vv). In addition, we can see that the discrepancy broadens when the life of the cap increases because of the Monte Carlo simulation of more forward rates.

In brief, these results indicate that caps have lower prices under the LFM ($\rho_{i,j} = 1$) than under the CIR model.

Table 4.7: Prices of caps with nominal value \$100

Life of cap (year)	Model	Cap rate (% per annum)		
		8.0	9.0	10.0
1.0	CIR	1.02	0.61	0.29
	LFM(fv)	1.01	0.56	0.14
	LFM(vv)	1.01	0.59	0.27
2.0	CIR	2.96	1.86	1.00
	LFM(fv)	2.87	1.59	0.48
	LFM(vv)	2.91	1.78	0.91
3.0	CIR	4.74	3.07	1.77
	LFM(fv)	4.53	2.51	0.82
	LFM(vv)	4.63	2.87	1.53
4.0	CIR	6.37	4.20	2.50
	LFM(fv)	6.00	3.34	1.18
	LFM(vv)	6.16	3.83	2.07
5.0	CIR	7.85	5.23	3.19
	LFM(fv)	7.31	4.10	1.52
	LFM(vv)	7.51	4.67	2.51

4.6 Comparison of the CIR Model and the LFM for Valuing European Options on Coupon-bearing Bonds

As Brigo and Mercurio [19] mentioned, under the CIR model and the Hull-White extended Vasicek short rate model, European options on coupon-bearing bonds can be explicitly priced by means of Jamshidian's decomposition (see Jamshidian

[54]), which transforms the problem into evaluating a series of European options on zero-coupon bonds.

We consider a European call option with strike price K and maturity T written on a coupon-bearing bond which pays coupon c_i at time $s_i > T$ $(i = 1, 2, \ldots, M)$. Denote by $\hat{r}$ the short rate at time T such that the coupon-bearing bond price equals the strike price, i.e.,

$$\sum_{i=1}^{M} c_i P(\hat{r}, T, s_i) = K,$$

where $P(r(t), t, s_i)$ is given in (4.2.8), then K can be decomposed as $K = \sum_{i=1}^{M} c_i K_i$, with $K_i = P(\hat{r}, T, s_i)$. Now the payoff function of the above European option

$$\left[\sum_{i=1}^{M} c_i P(r(T), T, s_i) - K\right]^+ \tag{4.6.22}$$

can be written as

$$\sum_{i=1}^{M} c_i [P(r(T), T, s_i) - K_i]^+,$$

because $P(r, T, s_i)$ $(i = 1, 2, \ldots, M)$ are decreasing functions of r.

Discounting this payoff function from T to t and taking expectation, we have

$$\mathrm{CBO}(t, r(t), T, K) = \sum_{i=1}^{M} c_i \mathbb{E}\left(e^{-\int_t^T r(s)ds}[P(r(T), T, s_i) - K_i]^+\right), \tag{4.6.23}$$

where $\mathrm{CBO}(\cdot)$ denotes the price of the above mentioned European call option on the coupon-bearing bond. It is known that

$$\mathbb{E}\left(e^{-\int_t^T r(s)ds}[P(r(T), T, s_i) - K_i]^+\right) = \mathrm{ZBC}(t, T, s_i, K_i). \tag{4.6.24}$$

As in Section 4.2, $\mathrm{ZBC}(\cdot)$ denotes the price of European call option with maturity T and strike K_i written on a zero-coupon bond maturing at $s_i > T$. It has other analytical forms under the CIR model and the Hull-White extended Vasicek short rate model (see Brigo and Mercurio [19]), thus the $\mathrm{CBO}(\cdot)$ has analytical solutions under these models.

We consider extending the equation (4.6.23) to the LFM context, therefore, we need to express every ZBC($\cdot$) in (4.6.23) analytically under the LFM. However, for multiple forward rates under one measure, those related ZBC($\cdot$) do not have explicit forms simultaneously, thus (4.6.23) is not analytical under the LFM. We compute CBO($\cdot$) under the LFM by discounting its payoff function, i.e., (5.3.8), using Monte Carlo method (see Section 5.3). For the comparability of the option price under the LFM with $\rho_{i,j} = 1$ with that under the CIR model, we still use the CIR model to produce the initial forward rates and volatilities for the LFM as in the last section. Then we generate forward rate paths and compute CBO($\cdot$) under the LFM. As far as the price of the same option under the CIR model is concerned, we cite it from Hull and White [47].

Consider European call options with different maturities on a 5-year bond that has a par \$100 and pays a coupon of 10% per annum semiannually. The parameters for the CIR model are $\kappa = 0.2, \theta = 0.1, \sigma = 0.06, r_0 = 0.1$. We have the initial forward rates $F_i(0; T_{i-1}, T_i)$ ($T_i = (i+1) * 0.5, i = 1, 2, \ldots, 17$) and their corresponding initial volatilities σ_i for the LFM by (4.5.20) and (4.5.21). In addition, $T_0 = 0.5$. The values of $F_i(0; T_{i-1}, T_i)$ are 0.1024, 0.1023, 0.1021, 0.1019, 0.1017, 0.1015, 0.1012, 0.1010, 0.1008, 0.1006, 0.1004, 0.1002, 0.1000, 0.0999, 0.0997, 0.0996, 0.0994. The initial values of σ_i are 0.1674, 0.1512, 0.1365, 0.1231, 0.1110, 0.1000, 0.0901, 0.0811, 0.0730, 0.0657, 0.0591, 0.0532, 0.0478, 0.0430, 0.0386, 0.0347, 0.0312. Now we simulate $F(t; T_{i-1}, T_i)$ under Q^1 using Monte Carlo method and get $F(T_{i-1}; T_{i-1}, T_i)$ for $i = 1, 2, \ldots, 17$ (we simulate 10^5 paths and set $dt = 0.005$ in the simulation). Then the prices of the options with different strike prices and maturities can be calculated. The prices of these options under the CIR model are cited from Hull and White [47]. The results are shown in Table 4.8 (where the unit is \$), where "LFM(fv)" means the LFM with fixed volatility 0.06 for all $F_i(t)$, and "LFM(vv)" means the LFM with varying volatilities for $F_i(t)$.

It is observed that the option prices under the CIR model are higher than those under the LFM (both LFM(fv) and LFM(vv)), which can be ascribed to that the discount factor $e^{-\int_t^T r(s)ds}$ and the term $[P(r(T), T, s_i) - K_i]^+$ in (4.6.23) are both decreasing with respect to r. These two terms under the CIR model are both larger than those under the LFM.

Table 4.8: Prices of call options on a 5-year bond ($)

Maturity (year)	Model	Exercise price		
		95.0	97.5	100.0
0.5	CIR	4.30	2.32	0.94
	LFM(fv)	3.14	0.94	0.04
	LFM(vv)	3.12	1.18	0.21
1.0	CIR	4.32	2.54	1.24
	LFM(fv)	3.05	0.97	0.06
	LFM(vv)	3.01	1.17	0.22
2.0	CIR	4.12	2.52	1.31
	LFM(fv)	2.84	1.00	0.10
	LFM(vv)	2.79	1.10	0.21
3.0	CIR	3.73	2.21	1.05
	LFM(fv)	2.66	1.01	0.14
	LFM(vv)	2.59	1.02	0.18
4.0	CIR	3.32	1.77	0.60
	LFM(fv)	2.45	0.99	0.18
	LFM(vv)	2.43	0.94	0.14

4.7 Summary

In this chapter, we introduced the forward rate model LFM and the CIR short rate model, and derived the explicit cap pricing formulas under these two models. We calibrated the LFM with zero correlations and the CIR model to the market data of caps, and found that the LFM with zero correlations fitted in-the-sample market data better than the CIR model. However, these two models have their own strengths on predicting out-of-the-sample cap prices in different markets. The CIR model is better at predicting the ATM Euro caps, while the LFM with zero correlation is more suitable for predicting the ATM US caps with short maturities. In our opinion, the assumption of $\rho_{i,j} = 0$ hidden in Black's formula is not reasonable. Thus we calibrated the LFM with an exogenous correlation matrix ($\rho_{i,j} = 1 - 0.025 * |i - j|$) and the LFM with a parametric correlation matrix to the market data. This calibration was carried out by the downhill simplex method in multidimensions and Monte Carlo method. The results show that the LFM with $\rho_{i,j} = 1 - 0.025 * |i - j|$ outperforms the LFM with $\rho_{i,j} = 0$, and the LFM with parametric correlations outperforms the LFM with $\rho_{i,j} = 1 - 0.025 * |i - j|$ in terms of fitting capability. In addition, we compared the prices of caps under the one-factor LFM ($\rho_{i,j} = 1$) with those under the CIR model. We observed that caps had lower prices under the former model than under the latter model. We also compared the prices of European options on coupon-bearing bonds under these two models, and found similar price behavior. For the comparability, the initial forward rates and volatilities used in the LFM were derived from the CIR model.

Chapter 5

Monte Carlo Method for Pricing Interest Rate Derivatives under the LFM

We have introduced the widely used forward rate model, i.e., the LFM, and Black's formula for pricing caps in the previous chapter. Also, we proposed that Black's formula holds under the LFM if and only if $\rho_{i,j} = 0$ $(i \neq j)$. This viewpoint will be verified by numerical results later. For pricing caps under the LFM with $\rho_{i,j} \neq 0$, we need to resort to Monte Carlo method or other methods, which constitute the main contents of this and the next chapter.

Within the LFM framework, no matter $\rho_{i,j} = 0$ or $\rho_{i,j} \neq 0$, the prices of most interest rate derivatives can be computed by Monte Carlo simulation method, the details of which will be described later. The main advantage of Monte Carlo method is its applicability to price a large range of derivatives, and straightforward implementation directly from the stochastic model rather than requiring further derivation (as for PDE method). However, it is computationally expensive because a large number of trials need to be generated. For derivative pricing problems in this chapter, after pricing tests with a quantity of paths, we find that the results using 10^5 paths are very close to those using 3×10^4 paths. For instance, the average price discrepancy of caps in Section 5.2 is within 1.0×10^{-4}. In other words, 10^5 paths are enough for the results to converge. It is acceptable in our problems from the perspective of computation speed.

The entire Monte Carlo simulation procedures can be divided into three stages, namely, simulation initialization, forward rate path generation and post processing for determining derivative price. We need to complete these three steps for pricing all kinds of interest rate derivatives. In this chapter, we first illustrate the Monte Carlo method for pricing caps, European options on coupon-bearing bonds and swaptions under the one-factor LFM, then extend it to the two-factor LFM. In Section 5.1, we introduce forward rate path generation under the one-factor LFM. The generated paths are then used for pricing interest rate derivatives in Section 5.2-5.4. In Section 5.2, caps are evaluated, and the cap prices obtained by the Monte Carlo method when $\rho_{i,j} = 0$ $(i \neq j)$ totally coincide with that computed by Black's formula. In Section 5.3, the forward rates are used to valuate European options on coupon-bearing bonds. In Section 5.4, swaptions are valuated under the one-factor LFM, and the use of swap rates is avoided. In Section 5.5, we extend the Monte Carlo simulation of forward rates to the two-factor LFM, and investigate cap pricing again. Section 5.6 is the summary of this chapter.

5.1 Monte Carlo Simulation of Forward Rates under the One-factor LFM

We have introduced the basic setup of one-factor LFM in Section 4.1. Now we give a general form of the dynamics of those forward rates, in which each forward rate is described by a Brownian motion vector. We call it multi-factor LFM.

Suppose the resetting dates of the underlying forward rates of caps are: $T_0, T_1, \ldots, T_N$ ($T_0 > 0$). In the multi-factor LFM framework, the dynamics of $F_n(t)$ under the forward measure Q^n is

$$\frac{dF_n(t)}{F_n(t)} = \vec{\sigma}_n(t) \cdot d\overrightarrow{Z_n}(t), \quad n = 1, \ldots, N, \tag{5.1.1}$$

where dF_n is the change of forward rate F_n in the time interval dt, $\vec{\sigma}_n(t)$ is the d-dimensional constant volatility vector, and $d\overrightarrow{Z_n}(t)$ is the differential of a d-dimensional Brownian motion vector $\overrightarrow{Z_n}(t)$.

To facilitate Monte Carlo evaluation of interest rate derivatives, we need to write the dynamics of different forward rates uniformly under one measure. That is done in Proposition 6.1.1 for the one-factor LFM. We do the similar transformation for the multi-factor LFM here. Hence, we write the dynamics of $F_n(t)$ under the forward measure Q^1:

$$\frac{dF_n(t)}{F_n(t)} = (\vec{\sigma}_n(t) \cdot \vec{\mu}_n(t))dt + \vec{\sigma}_n(t) \cdot d\overrightarrow{Z_n}(t), \quad n = 1, \ldots, N. \tag{5.1.2}$$

Here, the drift coefficient $\vec{\mu}_n$ is a d-dimensional vector given by

$$\begin{cases} \vec{\mu}_1(t) = \vec{0}, \\ \vec{\mu}_n(t) = \sum\limits_{i=2}^{n} \frac{\tau_i F_i(t) \rho_{n,i} \vec{\sigma}_i(t)}{1+\tau_i F_i(t)}, \quad t \le T_0, \quad n = 2, \ldots, N, \\ \vec{\mu}_n(t) = \sum\limits_{i=j+2}^{n} \frac{\tau_i F_i(t) \rho_{n,i} \vec{\sigma}_i(t)}{1+\tau_i F_i(t)}, \quad t \in (T_j, T_{j+1}], \quad j = 0, \ldots, n-2, \quad n = 2, \ldots, N, \end{cases} \tag{5.1.3}$$

where $\tau_i = T_i - T_{i-1}$, and $\rho_{n,i}$ is a $d \times d$ matrix which describes the instantaneous correlation between $\overrightarrow{Z_n}(t)$ and $\overrightarrow{Z_i}(t)$. Notice that $\vec{\mu}_n(t)$ ($n = 2, \ldots, N$) are piecewise functions, which should be tackled with care in the simulation.

It is easy to see that the drift terms in (5.1.3) are the vector forms of their counterparts in Proposition 6.1.1. Each component of $d\overrightarrow{Z_n}(t)$ can be written as

$$[d\overrightarrow{Z_n}(t)]_k = [\vec{\epsilon}_n]_k \sqrt{dt}, \quad k = 1, \ldots, d, \tag{5.1.4}$$

where $[\vec{\epsilon}_n]_k$ is a Gaussian random number drawn from the standard normal distribution, i.e., $[\vec{\epsilon}_n]_k \sim \Phi(0, 1.0)$. Therefore, a Gaussian random number generator is required to simulate the Brownian motion.

As we have mentioned, the Monte Carlo simulation procedures for evaluating interest rate derivatives can be divided into three stages. First, we initialize $F_n(0)$ $(n = 1, \ldots, N)$, the volatility vector $\vec{\sigma}_n$ and the instantaneous correlations $\rho_{n,i}$ in the initialization stage of the simulation.

In the second stage, the LFM paths are generated according to (5.1.2). The pseudo-code for the LFM model can be described as

Step 1: for $n = CurrPeriod + 1$ to N

Step 2: $factor = \tau_n F_n \rho_{n,i}/(1.0 + \tau_n F_n)$

Step 3: $\vec{\mu}_n = factor \times \vec{\sigma}_n$

Step 4: $\vec{\mu}_n = \vec{\mu}_n + \vec{\mu}_{n-1}$

Step 5: $\varsigma = (\vec{\mu}_n \cdot \vec{\sigma}_n)dt + (d\overrightarrow{Z_n}(t) \cdot \vec{\sigma}_n)$

Step 6: $dF_n = \varsigma \times F_n$

Step 7: $F_n = F_n + dF_n$

where $CurrPeriod$ is the index of the current period and N is the number of forward rates. From this process, we can see that we deal with the dynamics of forward rates instead of the dynamics of the logarithm of forward rates, while the latter way is adopted in the existing literature. After forward rate paths are generated, we can evaluate the derivative products.

In order to simplify the simulation in the second stage, we set $d = 1$, by which the vectors $\vec{\sigma}_n(t)$, $\vec{\mu}_n(t)$, $\overrightarrow{Z_n}(t)$ and $\vec{\epsilon}_n$ become scalers $\sigma_n(t)$, $\mu_n(t)$, $Z_n(t)$ and ϵ_n. Note that $\rho_{n,i} = 1$ (n=2, ..., N, i=2, ..., n) is the simplest case for generating forward rate paths, so we firstly let $\rho_{n,i} = 1$ in simulating forward rates and valuing derivatives in numerical experiments. Other two groups of values are also taken for $\rho_{n,i}$ for comparison.

Remark 5.1.1 *In case different values are used for $\rho_{n,i}$, ϵ_n should be generated accordingly. Suppose,*

(i) $\rho_{n,i} = 1$, which means that $Z_n(t)$ $(n = 1, \ldots, N)$ can be seen as one Brownian motion, and ϵ_n $(n = 1, \ldots, N)$ can be taken as one random variable which follows the standard normal distribution.

(ii) $0 < \rho_{n,i} < 1$ $(n \neq i)$, which means that $Z_n(t)$ $(n = 1, \ldots, N)$ are correlated Brownian motions in a Brownian motion vector, and ϵ_n $(n = 1, \ldots, N)$ should be the components of multivariate normal deviate with covariance $(\rho_{i,j})_{N\times N}$ and mean $\vec{0}$. There is a general way to construct a vector deviate x with a specified covariance $\sum$ and mean μ, starting with a vector y of independent random deviates that are of zero mean and unit variance. First, we use Cholesky decomposition to factor $\sum$ into a left triangular matrix L times its transpose, i.e., $\sum = LL^T$. Next, construct $x = Ly + \mu$.

(iii) $\rho_{n,i} = 0$ $(n \neq i)$, which means that $Z_n(t)$ $(n = 1, \ldots, N)$ are uncorrelated Brownian motions, and ϵ_n $(n = 1, \ldots, N)$ should be independent random deviates following the standard normal distribution.

5.2 Valuations of Caps

Recall that a cap consists of a series of caplets in each of which the payoff between the floating rate $F_n(T_{n-1})$ and the cap rate X in period $[T_{n-1}, T_n]$ is settled at time T_n $(n = 1, \ldots, N)$. For pricing the cap via Monte Carlo method, a large number of interest rate paths need to be generated using pseudo-random numbers. In each path, the forward rate $F_n(T_{n-1})$ $(n = 1, \ldots, N)$ is realized which enables the caplet payoff to be calculated as

$$\text{payoff}_n = \text{principal} \times \tau_n \times \max(F_n(T_{n-1}) - X, 0). \tag{5.2.5}$$

As the amount payoff_n is to be received at T_n, its value at time zero is the amount that would grow to payoff_n with interest rates from 0 to T_n. To find the value of payoff_n at 0, the discount factor for discounting payoff_n at T_n back to 0 is

$$D(0,T_n) = D(0,T_0) \cdot D(T_0,T_n) = D(0,T_0) \cdot \prod_{i=1}^{n} \frac{1}{1+\tau_i F_i(T_{i-1})}. \tag{5.2.6}$$

Now the payoff of each caplet can be backward discounted to time zero and summed up to form the price of the cap under one Monte Carlo trial. The average value of cap prices in all the Monte Carlo trials is what we want. Therefore, we have

$$\text{Cap}^{\text{MC}} = \mathbb{E}^1 \left(\sum_{n=1}^{N} D(0,T_n) \cdot \ \text{payoff}_n \right), \tag{5.2.7}$$

where $\mathbb{E}^1$ means expectation of sample cap prices under Q^1. To verify that our program of Monte Carlo method is correct, the cap prices obtained by Monte Carlo method when $\rho_{i,j} = 0$ $(i \neq j)$ will be compared with that computed by Black's formula. The expression of Black's formula is referred to Section 4.1. Approximately, we can take $D(0,T_n)$ in Black's formula as

$$D(0,T_n) = \frac{1}{1+T_0 F_1(0)} \prod_{i=1}^{n} \frac{1}{1+\tau_i F_i(0)}.$$

We compute the prices of caps by Monte Carlo method when the correlations $\rho_{n,i}$ are taken as different values in numerical experiments. For example, we assume that cap rate $X = 0.1$ and cap maturities vary from 1 to 5 years. Suppose the principal of these caps is \$1, and the time to maturity is segmented by $T_0 = 0.5, T_1 = 1, \ldots, T_9 = 5$, and $\tau_n = 0.5$ $(n = 1, \ldots, 9)$. The initial forward rates are taken as $F_n(0) = 0.1$ $(n = 1, \ldots, 9)$. The constant volatility parameter $(\sigma_1 = \ldots = \sigma_9)$ varies from 0.06 to 0.41. The correlations are taken as: (i) $\rho_{n,i} = 1$, (ii)$\rho_{n,i} = 1 - 0.05 * |n - i|$, (iii) $\rho_{i,i} = 1$, and $\rho_{n,i} = 0$ $(n \neq i)$. We set $dt = 0.005$ in the Monte Carlo simulation and simulate 10^5 paths for the forward rates, by which the prices of these caps have converged. In Black's formula, we let $D(0, T_n) = 1/(1 + 0.5 * 0.1)^{n+1}$. The cap prices at time 0 computed by the two methods are shown in Table 5.1, where they are reported in \$1/10000. In the table, "MC (i)", "MC(ii)" and "MC (iii)" mean that cap prices are obtained by Monte Carlo method with correlations (i), (ii) and (iii), respectively. We need to point out that the 1 year cap here is written on forward rate $F(t; 0.5, 1)$, and the 2 year cap is written on forward rate $F(t; 0.5, 1)$, $F(t; 1, 1.5)$, and $F(t; 1.5, 2)$. Other caps are defined in the similar way.

Table 5.1: Prices of caps at time 0 ($1/10000)

Life of cap (year)	Method	Volatility parameter			
		0.06	0.1	0.2	0.41
1.0	MC (i)	7.68	12.74	25.59	52.53
	MC (ii)	7.65	12.83	25.50	51.87
	MC (iii)	7.65	12.89	25.60	52.29
	Black's	7.78	12.84	25.49	52.11
2.0	MC (i)	30.12	50.27	101.22	208.54
	MC (ii)	30.14	50.54	100.50	207.06
	MC (iii)	30.01	50.09	100.04	205.01
	Black's	30.09	50.24	100.00	203.94
3.0	MC (i)	57.92	96.81	195.65	407.39
	MC (ii)	58.02	96.97	194.39	406.99
	MC (iii)	57.44	95.56	190.76	389.47
	Black's	57.51	95.90	191.07	388.12
4.0	MC (i)	88.49	148.10	302.10	640.66
	MC (ii)	88.64	148.21	299.62	638.88
	MC (iii)	87.22	145.70	290.35	588.62
	Black's	87.35	145.60	290.35	587.90
5.0	MC (i)	120.51	202.19	416.21	898.79
	MC (ii)	120.52	202.00	411.91	893.30
	MC (iii)	118.21	197.45	393.12	795.27
	Black's	118.31	197.31	393.10	793.62

In these numerical results, the errors between the Monte Carlo method and Black's formula are within $\$1.65\times10^{-4}$ when the forward rates are instantaneously uncorrelated. However, the maximum error will go to $\$105.17\times10^{-4}$ as the forward rates are perfectly instantaneously correlated. From the results of the last chapter, we know that the implied volatility of the 5 year Euro-forward-rate cap and 5 year US-forward-rate cap is 0.4099 and 0.4479, respectively. Therefore, it is reasonable that we take the volatility parameter as 0.41 to examine the discrepancy between Monte Carlo method and Black's formula.

The results in Table 5.1 show that the cap prices are not independent of the correlations under the LFM, and verified that Black's formula holds under the LFM with precondition that the underlying forward rates are instantaneously uncorrelated.

5.3 Valuing European Options on Coupon-bearing Bonds

Now we study the European call option of a coupon-bearing bond which pays coupon c_j at time $s_j > T$ $(j = 1, 2, \ldots M)$. Since there is no closed-form solution for this option price under the LFM, we use the Monte Carlo method.

Naturally, thousands of interest rate paths need to be generated according to Eq. (5.1.2). In each path, the payoff of the option at time T is calculated by

$$\text{payoff}_M = \max\left[\sum_{j=1}^{M} c_j D(T, s_j) - K, 0\right], \tag{5.3.8}$$

where the discount factor is

$$D(T, s_j) = \prod_{i=1}^{j} \frac{1}{1 + \tau_i F_i(T_{i-1})},$$

and $T = T_0, s_j = T_j$. In general, $\{T, s_j\ (j = 1, 2, \ldots M)\}$ should be a subset of the forward rates resetting dates $\{T_n | n = 0, 1, \ldots, N\}$. The payoff_M is then discounted backward to time zero by multiplying $D(0, T)$, which leads to the price of the option under one Monte Carlo trial. The average of the option prices in all the Monte Carlo trials is the option price we want to work out.

In our experiment, the underlying asset of the option is the 5-year bond which has the par \$100 and pays a coupon of 10% per annum semiannually, i.e., $c_j = \$5$, $s_j - s_{j-1} = 0.5$ for $2 \leq j \leq M$, and $s_1 - T = 0.5$. In the simulation of forward rates, we let $\tau_n = 0.5$, $\sigma_n = 0.06$, and the initial forward rates be $F_n(0) = 0.1$. The correlations $\rho_{n,i}$ are taken as 1. We set $dt = 0.005$ and simulate 10^5 paths for the forward rates. The option maturity T is taken from the set $\{0.5, 1, 2, 3\}$ (the unit is year), while K is taken from $\{95, 97.5, 100\}$. The prices of these options at time 0 computed by the Monte Carlo method are shown in Table 5.2, where they are reported in \$. Besides pricing caps and the above European options, Monte Carlo method can also be used for pricing swaptions in the LFM framework.

Table 5.2: Prices of European options at time 0 (\$)

Option maturity (year)	Strike price (\$)		
	95	97.5	100
0.5	3.78	1.46	0.12
1.0	3.57	1.39	0.14
2.0	3.21	1.28	0.17
3.0	2.85	1.15	0.18

5.4 Pricing Swaptions

Recall the definition 4.1.4, the forward swap rate $S_{\alpha,\beta}(t)$ at time $t \leq T_\alpha$ for the sets of times $\{T_\alpha, T_{\alpha+1}, \ldots, T_\beta\}$ is

$$S_{\alpha,\beta}(t) = \frac{P(t,T_\alpha) - P(t,T_\beta)}{\sum\limits_{i=\alpha+1}^{\beta} \tau_i P(t,T_i)} = \frac{1 - FP(t;T_\alpha,T_\beta)}{\sum\limits_{i=\alpha+1}^{\beta} \tau_i FP(t;T_\alpha,T_i)},$$

where

$$FP(t;T_\alpha,T_i) = \frac{P(t,T_i)}{P(t,T_\alpha)} = \prod_{j=\alpha+1}^{i} \frac{1}{1+\tau_j F_j(t)}.$$

Note that the forward swap rate is actually a nonlinear function of the underlying forward LIBOR. It can be written in another form:

$$S_{\alpha,\beta}(t) = \sum_{i=\alpha+1}^{\beta} \omega_i(t) F_i(t), \tag{5.4.9}$$

$$\omega_i(t) = \frac{\tau_i FP(t; T_\alpha, T_i)}{\sum\limits_{k=\alpha+1}^{\beta} \tau_i FP(t; T_\alpha, T_k)} = \frac{\tau_i P(t, T_i)}{\sum\limits_{k=\alpha+1}^{\beta} \tau_k P(t, T_k)}.$$

This is an important formula connecting the forward LIBOR with the forward swap rate. According to Brigo and Mercurio [19], the variability of the $\omega_i's$ is less compared to the variability of the $F_i's$, thus one can approximate the $\omega_i's$ by their initial values $\omega_i(0)$ and obtain

$$S_{\alpha,\beta}(t) \approx \sum_{i=\alpha+1}^{\beta} \omega_i(0) F_i(t), \tag{5.4.10}$$

which is often used for simulating swap rates starting from the LFM.

There are two main types of swaptions, i.e., payer swaptions and receiver swaptions. For the Payer IRS (see Definition 4.1.3) with unit nominal value and X different from the swap rate $S_{\alpha,\beta}(t)$, its discounted payoff at time t can be expressed in terms of swap rates as

$$D(t, T_\alpha)(S_{\alpha,\beta}(T_\alpha) - X) \sum_{i=\alpha+1}^{\beta} \tau_i P(T_\alpha, T_i). \tag{5.4.11}$$

Definition 5.4.1 *A swaption is a contract that gives its holder the right (but not the obligation) to enter at a future time $T_\alpha > 0$ an IRS, whose first reset time usually coincides with T_α, with payments occurring at dates $T_{\alpha+1}, T_{\alpha+2}, \ldots, T_\beta$. The fixed rate X of the underlying swap is usually called the swaption strike. If we assume with unit notional amount, the payer swaption payoff at time t $(0 \le t \le T_\alpha)$ can be written as*

$$D(t, T_\alpha)(S_{\alpha,\beta}(T_\alpha) - X)^+ \sum_{i=\alpha+1}^{\beta} \tau_i P(T_\alpha, T_i). \tag{5.4.12}$$

The underlying-IRS length ($T_\beta - T_\alpha$ in our notation) is called the tenor of the swaption, and T_α is called the maturity of the swaption.

The payer swaption is said to be at-the-money (ATM) at time t if $X = S_{\alpha,\beta}(t)$, and in-the-money (ITM) if $X < S_{\alpha,\beta}(t)$, and out-of-the-money (OTM) if $X > S_{\alpha,\beta}(t)$. Clearly, moneyness for a receiver swaption is defined in the opposite way.

For computing the payer swaption price at time $t = 0$

$$\mathbb{E}\left(D(0,T_\alpha)(S_{\alpha,\beta}(T_\alpha) - X)^+ \sum_{i=\alpha+1}^{\beta} \tau_i P(T_\alpha, T_i)\right),$$

one can start with forward rate dynamics and use (5.4.10) to simulate $S_{\alpha,\beta}(T_\alpha)$ (refer to Section 6.10 of Brigo and Mercurio [19] about the details), or start by modeling the dynamics of $S_{\alpha,\beta}(t)$, and then calculate the above expectation. Actually, we can circumvent $S_{\alpha,\beta}(t)$ by defining swaptions with forward LIBOR and discuss swaption pricing under the LFM.

Proposition 5.4.1 *The discounted payoff at time t of the payer swaption in Definition 5.4.1 can be expressed in terms of forward LIBOR as*

$$D(t,T_\alpha)\left(\sum_{i=\alpha+1}^{\beta} \tau_i P(T_\alpha, T_i)(F(T_\alpha; T_{i-1}, T_i) - X)\right)^+, \tag{5.4.13}$$

which is equivalent to (5.4.12).

Proof. According to the definition of $S_{\alpha,\beta}(T_\alpha)$, we have

$$\begin{aligned}
&(S_{\alpha,\beta}(T_\alpha) - X)^+ \sum_{i=\alpha+1}^{\beta} \tau_i P(T_\alpha, T_i) \\
&= \left(P(T_\alpha, T_\alpha) - P(T_\alpha, T_\beta) - \sum_{i=\alpha+1}^{\beta} \tau_i X P(T_\alpha, T_i)\right)^+.
\end{aligned}$$

On the other hand,

$$\begin{aligned}
&\left(\sum_{i=\alpha+1}^{\beta} \tau_i P(T_\alpha, T_i)[F(T_\alpha; T_{i-1}, T_i) - X]\right)^+ \\
&= \left(\sum_{i=\alpha+1}^{\beta} P(T_\alpha, T_i)[\tfrac{P(T_\alpha, T_{i-1})}{P(T_\alpha, T_i)} - 1 - \tau_i X]\right)^+ \\
&= \left(\sum_{i=\alpha+1}^{\beta} [P(T_\alpha, T_{i-1}) - P(T_\alpha, T_i) - \tau_i X P(T_\alpha, T_i)]\right)^+ \\
&= \left(P(T_\alpha, T_\alpha) - P(T_\alpha, T_\beta) - \sum_{i=\alpha+1}^{\beta} \tau_i X P(T_\alpha, T_i)\right)^+ .
\end{aligned}$$

Multiplying $D(t, T_\alpha)$ on the left side of the above two equations, we get the equivalence of (5.4.12) and (5.4.13).□

When pricing the swaption with (5.4.13), we first simulate forward rates by Monte Carlo method. Note that we need to consider the joint action of the rates involved in the contract payoff, i.e., correlations between different forward rates at T_α (also called terminal correlations). In the simulation, we presume the instantaneous correlations of the forward rates, i.e., $\rho_{n,i}$, in the initialization stage.

Now we provide numerical examples of pricing swaptions under the one-factor LFM. We investigate swaptions with maturity $T_0 = 0.5$ year and strike $X = 0.1$. The nominal value of the swaptions is 1\$, and the tenors of the swaptions vary from 0.5 to 4.5 years. We use the same initial forward rates as in Section 5.2 for generating forward rates, i.e., $F_n(0) = 0.1$ $(n = 1, \ldots, 9)$, $T_0 = 0.5, T_1 = 1, \ldots, T_9 = 5$, $\tau_n = 0.5$ $(n = 1, \ldots, 9)$. The volatility parameter $(\sigma_1 = \sigma_2 = \cdots = \sigma_9)$ is taken from $\{0.06, 0.1, 0.2\}$. The correlations are taken as: (i) $\rho_{n,i} = 1$, (ii) $\rho_{n,i} = 1 - 0.05 * |n - i|$, (iii) $\rho_{i,i} = 1$, and $\rho_{n,i} = 0$ $(n \neq i)$. In this experiment, we still set $dt = 0.005$ in the simulation and simulate 10^5 paths for the forward rates, by which we compute the prices of the swaptions at time 0 and the results converge. The results are presented in Table 5.3 in the unit of \$1/10000.

Table 5.3: Prices of swaptions at time 0 ($1/10000)

Tenor of swaption (year)	Correlations	Volatility parameter 0.06	0.1	0.2
0.5	(i)	7.71	12.80	25.42
	(ii)	7.59	12.82	25.56
	(iii)	7.69	12.82	25.59
1.5	(i)	12.74	21.21	42.29
	(ii)	21.22	35.85	71.33
	(iii)	12.69	21.19	42.45
2.5	(i)	15.68	26.08	52.34
	(ii)	33.08	55.95	111.35
	(iii)	15.75	26.12	52.29
3.5	(i)	17.74	29.53	59.10
	(ii)	43.38	73.28	146.17
	(iii)	17.85	29.61	59.23
4.5	(i)	19.26	32.03	64.18
	(ii)	52.29	88.26	176.36
	(iii)	19.34	32.28	64.46

Through numerical examples, we demonstrate that these swaptions have different prices when different instantaneous correlations are adopted. We observed that the prices of the swaptions would not be higher than their corresponding caps (Note that the swaption with tenor being 4.5 year corresponds to the 5 year cap, since their underlying forward rates are the same).

5.5 Cap Pricing under the Two-factor LFM

As we have mentioned in Section 5.1, the stochastic term $d\overrightarrow{Z_n}(t)$ could be a d-dimensional vector, which means that each forward rate is modeled by a Brwonian motion vector. In Section 5.2-5.4, we have studied the derivative pricing when $d = 1$. Now we discuss the cap pricing when $d = 2$, which corresponds to two-factor LFM. The procedures are the same as in the one-factor LFM except some details. We first simulate forward rates under the two-factor LFM, then investigate cap pricing based on that. By comparing the pricing results of the one-factor and two-factor LFM, we hope to find some connections between them.

In case $d = 2$, $\vec{\mu}_n$, $\vec{\sigma}_n$ and $\overrightarrow{Z_n}(t)$ in (5.1.2) are 2-dimensional vectors, and $\rho_{n,i}$ is a 2×2 matrix which expresses the correlation between $(Z_{n1}, Z_{n2})^T$ and $(Z_{i1}, Z_{i2})^T$, i.e.,

$$\rho_{n,i} = \begin{pmatrix} c_{n1,i1} & c_{n1,i2} \\ c_{n2,i1} & c_{n2,i2} \end{pmatrix},$$

where $c_{k,l}$ represents the correlation between Z_k and Z_l. All the $\rho_{n,i}$ $(n, i = 1, \ldots, N)$ constitute a symmetric matrix $\sum_{2N\times 2N}$, which is the instantaneous correlation matrix of $F_i's$ $(i = 1, \ldots, N)$, and all the $\vec{\epsilon}_n$ $(n = 1, \ldots, N)$ form a $2N \times 1$ vector which has the covariance $\sum_{2N\times 2N}$ and mean $\vec{0}$. After initializing $F_n(0)$, $\vec{\sigma}_n$ and $\sum_{2N\times 2N}$, we use the pseudo-code in Section 5.1 for generating forward rates, but need to pay some attention to the vector computation of $\vec{\mu}_n$ and $\vec{\sigma}_n$. Next, we compute the cap price by formula (5.2.7). A numerical experiment is implemented as below.

We still consider the caps in Section 5.2. The cap rate is $X = 0.1$ and the maturities range from 1 to 5 years. The principal of these caps is \$1. The resetting dates are $T_0 = 0.5, T_1 = 1, \ldots, T_9 = 5$, and $\tau_n = 0.5$ $(n = 1, \ldots, 9)$. The initial forward rates are $F_n(0) = 0.1$ $(n = 1, \ldots, 9)$. The volatility parameter $(\vec{\sigma}_1 = \vec{\sigma}_2 = \cdots = \vec{\sigma}_9)$ is taken as $(0.03, 0.03)^T$ or $(0.05, 0.05)^T$ or $(0.1, 0.1)^T$. Here, the sum of the components in each vector $\vec{\sigma}_n$ is equal to σ_n in the one-factor LFM. This arrangement is to facilitate the comparison between the one-factor LFM and two-factor LFM. The instantaneous correlations in the two-factor LFM are taken as: (i) $\sum_{i,i} = 1$, and $\sum_{i,j} = 0$ $(i \neq j)$, (ii) $\sum_{i,j} = 1 - 0.05 * |i - j|$ $(i, j = 1, 2, \ldots, 18)$. We still use $dt = 0.005$ in the simulation and simulate 10^5 paths for the forward rates, by which the cap prices converge. The results are presented in Table 5.4, where the rows denoted "Black's" are the results computed by Black's formula under the one-factor LFM, and the rows denoted "MC(i/ii)" are results arising from Monte Carlo method with correlations (i/ii) under the two-factor LFM. The results in in Table 5.4 are reported in \$1/10000.

Table 5.4: Prices of caps at time 0 under the two-factor LFM (\$1/10000)

Life of cap (year)	Method	Volatility parameter $(0.03, 0.03)^T$	$(0.05, 0.05)^T$	$(0.1, 0.1)^T$
1.0	MC (i)	5.43	8.99	18.11
	MC (ii)	7.62	12.65	25.55
	Black's	7.78	12.84	25.49
2.0	MC (i)	21.27	35.45	71.02
	MC (ii)	29.88	49.77	100.70
	Black's	30.09	50.24	100.00
3.0	MC (i)	40.70	67.89	135.47
	MC (ii)	57.15	95.58	193.44
	Black's	57.51	95.90	191.07
4.0	MC (i)	61.70	103.18	205.86
	MC (ii)	87.00	145.83	295.91
	Black's	87.35	145.60	290.35
5.0	MC (i)	83.66	139.63	279.22
	MC (ii)	118.14	197.97	403.94
	Black's	118.31	197.31	393.10

From the table, we have seen that the cap prices under the two-factor LFM with correlations $\sum_{i,j} = 1 - 0.05 * |i - j|$ are close to those under the one-factor LFM. But when the correlations are taken as other values under the two-factor LFM, the cap prices are far away from that under the one-factor LFM. This phenomenon implies that if we do not adopt appropriate correlations in the two-factor LFM, it would not outperform the one-factor LFM in pricing caps.

5.6 Summary

In this chapter, we first presented the details of generating forward rate paths under the LFM by Monte Carlo method. Our simulation of forward rates is different from that in the literature. Basing on the generated forward rate paths, we computed the prices of caps, European options on coupon-bearing bonds and swaptions. Our Monte Carlo simulation program is testified to be correct, since the cap prices obtained from the simulation method under the LFM with $\rho_{n,i} = 0$ highly coincide with that calculated by Black's formula. When computing the price of swaptions by Monte Carlo method, we circumvented swap rates and dealt with it under the LFM, since swaptions could be defined in terms of forward rates. Lastly, we demonstrated how to price caps under the two-factor LFM using the Monte Carlo method. We find that it would not outperform the one-factor LFM if we do not adopt appropriate correlations in the two-factor LFM.

Chapter 6

A Novel PDE Method for Cap Pricing under the LFM

The Monte Carlo method for pricing caps under the LFM with $\rho_{i,j} \neq 0$ is illustrated in the previous chapter. Recall that in the one-factor LFM framework, the forward rate $F_i(t)$ is assumed to follow a lognormal distribution under the forward measure Q^i:

$$dF_i(t) = \sigma_i(t)F_i(t)dZ_i(t), \quad i = 1, \ldots, N, \tag{6.0.1}$$

with

$$dZ_i(t)\ dZ_j(t) = \rho_{i,j}dt.$$

In the Monte Carlo procedures , we first write the dynamics of $F_i's$ under the uniform measure Q^1 as

$$\frac{dF_i(t)}{F_i(t)} = \sigma_i(t) \cdot \mu_i(t)dt + \sigma_i(t) \cdot dZ_i(t), \quad i = 1, \ldots, N, \tag{6.0.2}$$

where

$$\begin{cases} \mu_1(t) = 0, \\ \mu_n(t) = \sum\limits_{i=2}^{n} \frac{\tau_i F_i(t)\rho_{n,i}\sigma_i(t)}{1+\tau_i F_i(t)}, \quad t \leq T_0, \quad n = 2, \ldots, N, \\ \mu_n(t) = \sum\limits_{i=j+2}^{n} \frac{\tau_i F_i(t)\rho_{n,i}\sigma_i(t)}{1+\tau_i F_i(t)}, \quad t \in (T_j, T_{j+1}], \quad j = 0, \ldots, n-2, \quad n = 2, \ldots, N. \end{cases} \tag{6.0.3}$$

Then the paths of $F_i's$ are generated according to the above dynamics, and taken into the following formula for computing the cap price

$$\text{principal } \cdot \sum_{i=1}^{N} P(0,T_i)\tau_i \mathbb{E}^1(F_i(T_{i-1}) - X)^+, \tag{6.0.4}$$

where the expectations are taken under the measure Q^1, and $P(0,T_i)$ is the discount factor:

$$P(0,T_i) \approx \frac{1}{1+T_0F_1(0)} \prod_{k=1}^{i} \frac{1}{1+\tau_k F_k(0)}.$$

Although the Monte Carlo method has the advantage of wide applicability and easy implementation, we need a large amount of forward rate paths which lead to long computation time. In this chapter, we develop a PDE approach to determine $\mathbb{E}^1(F_i(T_{i-1}) - X)^+$ in (6.0.4). It requires us to solve a PDE to get the expectation for each caplet. Then we evaluate the cap consisting of multiple caplets. This PDE approach is similar to that of Heston [46], whereas Heston dealt with a European call option on an asset with stochastic volatility. Naturally, this PDE approach can be applied to price caplets under the LFM with stochastic volatility, which is more general than the LFM with deterministic volatility.

This chapter is structured as follows: Section 6.1 introduces the governing PDEs of the expected payoff functions of caplets. After solving these PDEs, we can get the cap price by the formula (6.0.4). Because the dimension of the PDE is increasing with respect to the number of caplets, the complexity of seeking its numerical solution grows rapidly. We take a three-period cap as an example to demonstrate our method, and prove the stability of the corresponding numerical scheme in Section 6.2. Section 6.3 provides numerical examples to show the effectiveness of our method. Section 6.4 concludes this chapter.

6.1 PDEs for Expected Payoff Functions of Caplets

We know that a caplet can be viewed as a European call option, so it can be priced resorting to a PDE analogous to the Black-Scholes equation. According to Feynman-Kac theorem (see Appendix C.5 of Brigo and Mercurio [19]), the solution of the PDE is equivalent to the expectation of the payoff function of the caplet. A cap is a portfolio of caplets, thus it is impossible to solve the payoff function of the cap by only one PDE because multi-payoff and different forward rates are involved in its duration. The multi-payoff property determines that we need to derive and solve PDEs for different caplets in order to get the cap price through (6.0.4). This is what we called PDE method for pricing caps.

Before we write out the PDEs governing the expectation of payoff functions of caplets, we need to unify forward rates in different periods under one measure. It is done in the following proposition (see Brigo and Mercurio [19]).

Proposition 6.1.1 *Under the assumption of lognormal distribution for each forward rate, we obtain that the dynamics of F_i under the forward measure Q^k in the three cases $k < i$, $k = i$ and $k > i$ are, respectively,*

$$\begin{aligned}
&k < i \quad t \leq T_k : \quad dF_i(t) = \sigma_i(t)F_i(t)\sum_{j=k+1}^{i} \tfrac{\rho_{i,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_j F_j(t)}dt + \sigma_i(t)F_i(t)dZ_i(t),\\
&k = i, \quad t \leq T_{i-1} : \quad dF_i(t) = \sigma_i(t)F_i(t)dZ_i(t), \qquad\qquad (6.1.5)\\
&k > i, \quad t \leq T_{i-1} : \quad dF_i(t) = -\sigma_i(t)F_i(t)\sum_{j=i+1}^{k} \tfrac{\rho_{i,j}\tau_j\sigma_j(t)F_j(t)}{1+\tau_j F_j(t)}dt + \sigma_i(t)F_i(t)dZ_i(t),
\end{aligned}$$

where, $Z(t)$ is a Brownian motion vector under Q^k, $Z_i(t)$ is the i-th component *of $Z(t)$ satisfying $dZ_i(t)dZ_j(t) = \rho_{i,j}dt$. All of the above equations admit a unique strong solution if the coefficients $\sigma_i(t)$ are bounded.*

The proof can be found in [19]. Obviously, the equations (6.0.2) and (6.0.3) are obtained by taking Q^k as Q^1 in this proposition. We will use the dynamics (6.0.2) in the sequel.

As we have mentioned, the expectation of caplet payoff function satisfies the PDE which is analogous to the Black-Scholes equation [6, 93]. We state the result in the following proposition.

Proposition 6.1.2 *When the forward rate $F(t; T_{i-1}, T_i)$ follows*

$$dF_i(t) = \psi(F_i)dt + \sigma(F_i)dZ_i(t), \quad i = 1, \ldots, N, \tag{6.1.6}$$

the price at time t of the caplet with unit principal, whose payoff is set at T_{i-1} and paid at T_i, can be written as

$$\begin{aligned} Cpl(t, T_{i-1}, T_i, \tau_i, X) &= P(t, T_i)\tau_i \mathbb{E}[(F_i(T_{i-1}) - X)^+ | \mathcal{F}_t^{Z_i}] \\ &= P(t, T_i)\tau_i G_i(t, F_i), \end{aligned} \tag{6.1.7}$$

where $\mathcal{F}_t^{Z_i}$ is the filtration generated by the Brownian motion $Z_i(t)$. The function $G_i(t, F_i)$ (denoted by G_i for brevity) satisfies the following PDE

$$\frac{\partial G_i}{\partial t} + \psi(F_i)\frac{\partial G_i}{\partial F_i} + \frac{1}{2}\sigma(F_i)^2 \frac{\partial^2 G_i}{\partial F_i^2} = 0, \quad 0 \le t \le T_{i-1}, \tag{6.1.8}$$

with terminal condition

$$G_i(T_{i-1}, F_i) = (F_i - X)^+.$$

Remark 6.1.1 *This proposition is similar to Theorem 2 in Andersen and Andreasen [6]. It is derived from the Feynman-Kac theorem by setting riskless return rate equal to 0. Notice that $G_i(t, F_i)$ is the expectation of payoff function of the caplet written on $F_i(t)$, and the dynamics of $F_i(t)$ has drift $\psi(F_i)$ and volatility $\sigma(F_i)$.*

Going back to the dynamics (6.0.2), we have

$$\mu_1(t) = 0, \quad \mu_2(t) = \frac{\sigma_2 \tau_2 F_2(t)}{1 + \tau_2 F_2(t)},$$

and $\mu_i(t)$ $(i = 3, \ldots, N)$ is dependent on $F_2(t), \ldots, F_i(t)$. That is to say, $F_i(t)$ $(i = 3, \ldots, N)$ has different dynamics from (6.1.6), thus, the governing PDE of the corresponding $G_i(t, F_i)$ is different from (6.1.8). We could take $G_i(t, F_i)$ $(i = 3, \ldots, N)$ as the function of t, $F_2(t), \ldots, F_i(t)$, and derive its governing PDE by the means of getting multi-dimensional Black-Scholes equation.

We take a cap of three periods as an example to demonstrate our PDE method for pricing caps. This cap consists of three caplets $\mathrm{Cpl}(t, T_0, T_1, \tau_1, X)$, $\mathrm{Cpl}(t, T_1, T_2, \tau_2, X)$ and $\mathrm{Cpl}(t, T_2, T_3, \tau_3, X)$ (the nominal value of these caplets is \$1), and the underlying forward rates are $F_1(t)$, $F_2(t)$ and $F_3(t)$. Their dynamics under the measure Q^1 can be written as

$$\begin{cases} dF_1(t) = \sigma_1 F_1(t) dZ_1(t), \quad 0 \le t \le T_0, \\ dF_2(t) = \frac{\tau_2 \sigma_2^2 F_2^2(t)}{1+\tau_2 F_2(t)} dt + \sigma_2 F_2(t) dZ_2(t), \quad 0 \le t \le T_1, \\ dF_3(t) = \sigma_3 F_3(t) \left(\frac{\rho_{2,3} \sigma_2 \tau_2 F_2(t)}{1+\tau_2 F_2(t)} + \frac{\sigma_3 \tau_3 F_3(t)}{1+\tau_3 F_3(t)} \right) dt + \sigma_3 F_3(t) dZ_3(t), \quad 0 \le t \le T_1, \\ dF_3(t) = \frac{\tau_3 \sigma_3^2 F_3^2(t)}{1+\tau_3 F_3(t)} dt + \sigma_3 F_3(t) dZ_3(t), \quad t \in (T_1, T_2], \end{cases}$$

where $Z_1(t)$, $Z_2(t)$ and $Z_3(t)$ are standard Brownian motions under Q^1, σ_1, σ_2 and σ_3 are constant volatilities. In the sequel, the correlation $\rho_{2,3}$ is taken as 1 since F_2 and F_3 are close on time axis. Then G_1 is the solution to the PDE

$$\frac{\partial G_1}{\partial t} + \frac{1}{2} \sigma_1^2 F_1^2 \frac{\partial^2 G_1}{\partial F_1^2} = 0, \quad 0 \le t \le T_0, \tag{6.1.9}$$

with the terminal condition

$$G_1(T_0, F_1) = (F_1 - X)^+.$$

G_2 is the solution to the PDE

$$\frac{\partial G_2}{\partial t} + \frac{\tau_2 \sigma_2^2 F_2^2}{1 + \tau_2 F_2} \frac{\partial G_2}{\partial F_2} + \frac{1}{2} \sigma_2^2 F_2^2 \frac{\partial^2 G_2}{\partial F_2^2} = 0, \quad 0 \le t \le T_1, \tag{6.1.10}$$

with the terminal condition

$$G_2(T_1, F_2) = (F_2 - X)^+.$$

And G_3 satisfies

$$\frac{\partial G_3}{\partial t} + \frac{\tau_3\sigma_3^2 F_3^2}{1+\tau_3 F_3}\frac{\partial G_3}{\partial F_3} + \frac{1}{2}\sigma_3^2 F_3^2 \frac{\partial^2 G_3}{\partial F_3^2} = 0, \quad t \in (T_1, T_2], \tag{6.1.11}$$

with the terminal condition

$$G_3(T_2, F_2, F_3) = (F_3 - X)^+.$$

On $[0, T_1]$, we have

$$\begin{aligned} &\frac{\partial G_3}{\partial t} + \frac{\tau_2\sigma_2^2 F_2^2}{1+\tau_2 F_2}\frac{\partial G_3}{\partial F_2} + \frac{1}{2}\sigma_2^2 F_2^2 \frac{\partial^2 G_3}{\partial F_2^2} + \sigma_3 F_3\left(\frac{\sigma_2\tau_2 F_2}{1+\tau_2 F_2} + \frac{\sigma_3\tau_3 F_3}{1+\tau_3 F_3}\right)\frac{\partial G_3}{\partial F_3} \\ &\quad + \frac{1}{2}\sigma_3^2 F_3^2 \frac{\partial^2 G_3}{\partial F_3^2} + \sigma_2\sigma_3 F_2 F_3 \frac{\partial^2 G_3}{\partial F_2 \partial F_3} = 0, \quad 0 \le t \le T_1, \end{aligned} \tag{6.1.12}$$

and the terminal condition at time $t = T_1$ is solved by (6.1.11).

Now the three-period-cap consisting of $\mathrm{Cpl}(t, T_0, T_1, \tau_1, X)$, $\mathrm{Cpl}(t, T_1, T_2, \tau_2, X)$ and $\mathrm{Cpl}(t, T_2, T_3, \tau_3, X)$ has its price at $t = 0$:

$$P(0,T_1)\tau_1 G_1(0, F_1(0)) + P(0,T_2)\tau_2 G_2(0, F_2(0)) + P(0,T_3)\tau_3 G_3(0, F_3(0)). \tag{6.1.13}$$

Here we apply the unconditionally stable and convergent finite difference scheme (Crank-Nicolson scheme) to equations (6.1.9) -(6.1.12) to work out $G_j(t, F_j(0))$ ($j = 1, 2, 3$) numerically. The stability of the Crank-Nicolson scheme for parabolic PDEs with variable coefficients can be obtained by von Neumann method, which is referred to Section 5.4 of Zhu, Wu and Chern [100]. The analysis there is adaptable to our equations (6.1.9)-(6.1.11).

6.2 Stability of the Numerical Schemes

We know that the stability of a numerical scheme is concerned with the propagation of errors, because the truncation errors of approximate solutions and rounding errors of computer are introduced into numerical solutions. For a given numerical scheme, it is said to be stable when the errors are not magnified at each step in some norm.

6.2.1 Stability of the Crank-Nicolson Scheme for (6.1.9)

There are two different norms that are often used in studying stability. Suppose

$$\mathbf{x} = (x_1, x_2, \dots, x_{M-1})^T$$

is a vector with $M-1$ components. The l_∞ and l_2 norms of the vector $\mathbf{x}$ are defined as

$$\|\mathbf{x}\|_\infty = \max_{1\le m\le M-1} |x_m|,$$

and

$$\|\mathbf{x}\|_2 = \left(\frac{1}{M-1}\sum_{m=1}^{M-1} x_m^2\right)^{1/2}.$$

For the heat equation

$$\begin{cases} \frac{\partial u}{\partial \tau} = a\frac{\partial^2 u}{\partial x^2}, \quad x_l \le x \le x_r, \quad 0 \le \tau \le T, \\ u(x,0) = f(x), \quad x_l \le x \le x_r, \\ u(x_l,\tau) = f_l(\tau), \quad 0 \le \tau \le T, \\ u(x_r,\tau) = f_r(\tau), \quad 0 \le \tau \le T, \end{cases} \tag{6.2.14}$$

where a is a positive constant. Its Crank-Nicolson scheme is

$$\begin{cases} -\frac{\alpha}{2}u_{m-1}^{n+1} + (1+\alpha)u_m^{n+1} - \frac{\alpha}{2}u_{m+1}^{n+1} = \frac{\alpha}{2}u_{m-1}^n + (1-\alpha)u_m^n + \frac{\alpha}{2}u_{m+1}^n, \\ \quad m = 1,2,\dots,M-1, \quad n = 0,1,\dots,N-1, \\ u_0^n = f_l(\tau^n), \quad n = 0,1,\dots,N, \\ u_M^n = f_r(\tau^n), \quad n = 0,1,\dots,N, \\ u_m^0 = f(x_m), \quad m = 1,2,\dots,M-1, \end{cases} \tag{6.2.15}$$

where $\alpha = \frac{a\Delta\tau}{\Delta x^2}$. Given $u(x_l, \tau)$ and $u(x_r, \tau)$, the $M-1$ equations in (6.2.15) can be written as the system:

$$\mathbf{A}\mathbf{u}^{n+1} = \mathbf{B}\mathbf{u}^n + \mathbf{b}^n, \tag{6.2.16}$$

where

$$\mathbf{A} = \begin{pmatrix} 1+\alpha & -\alpha/2 & 0 & \cdots & 0 \\ -\alpha/2 & 1+\alpha & -\alpha/2 & \ddots & \vdots \\ 0 & -\alpha/2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & -\alpha/2 \\ 0 & \cdots & 0 & -\alpha/2 & 1+\alpha \end{pmatrix},$$

$$\mathbf{B} = \begin{pmatrix} 1-\alpha & \alpha/2 & 0 & \cdots & 0 \\ \alpha/2 & 1-\alpha & \alpha/2 & \ddots & \vdots \\ 0 & \alpha/2 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \alpha/2 \\ 0 & \cdots & 0 & \alpha/2 & 1-\alpha \end{pmatrix},$$

$$\mathbf{u}^n = \begin{pmatrix} u_1^n \\ u_2^n \\ \vdots \\ u_{M-2}^n \\ u_{M-1}^n \end{pmatrix}, \quad \mathbf{b}^n = \begin{pmatrix} \alpha u_0^n/2 + \alpha u_0^{n+1}/2 \\ 0 \\ \vdots \\ 0 \\ \alpha u_M^n/2 + \alpha u_M^{n+1}/2 \end{pmatrix}.$$

Let $\mathbf{e}^n$ be the error vector between $\mathbf{u}^n$ and its real value. According to (6.2.16), we have

$$\mathbf{A}\mathbf{e}^{n+1} = \mathbf{B}\mathbf{e}^n. \tag{6.2.17}$$

Assume that

$$\mathbf{e}_{\omega_k} = \begin{pmatrix} \sin\omega_k \\ \sin 2\omega_k \\ \vdots \\ \sin(M-1)\omega_k \end{pmatrix},$$

where $\omega_k = k\pi/M$, $k = 1, 2, \ldots, M-1$. It can be verified that $\{\mathbf{e}_{\omega_k}, k = 1, 2, \ldots, M-1\}$ are orthogonal, thus, any initial error $\mathbf{e}^0$ can be expressed as

$$\mathbf{e}^0 = \sum_{k=1}^{M-1} \varepsilon_{\omega_k} \mathbf{e}_{\omega_k}.$$

Suppose the eigenvalues of $\mathbf{A}^{-1}\mathbf{B}$ are $\{\lambda_{\omega_k},\ k = 1, 2, \ldots, M-1\}$, because of (6.2.17), we get

$$\mathbf{e}^n = \sum_{k=1}^{M-1} \varepsilon_{\omega_k} \lambda_{\omega_k}^n \mathbf{e}_{\omega_k}, \quad n = 0, 1, \ldots, N-1.$$

By the expressions of $\mathbf{e}^0$ and $\mathbf{e}^n$, we have

$$\|\mathbf{e}^0\|_2 = \left(\frac{1}{M-1} \sum_{k=1}^{M-1} \varepsilon_{\omega_k}^2 \|\mathbf{e}_{\omega_k}\|_2^2 \right)^{1/2}$$

and

$$\|\mathbf{e}^n\|_2 = \left(\frac{1}{M-1} \sum_{k=1}^{M-1} \varepsilon_{\omega_k}^2 \lambda_{\omega_k}^{2n} \|\mathbf{e}_{\omega_k}\|_2^2 \right)^{1/2}.$$

Consequently, we obtain

$$\|\mathbf{e}^n\|_2 \le \|\mathbf{e}^0\|_2$$

if all the eigenvalues λ_{ω_k} are in $[-1, 1]$.

On the other hand, we have

$$\lambda_{\omega_k} \mathbf{A} \mathbf{e}_{\omega_k} = \mathbf{B} \mathbf{e}_{\omega_k},$$

i.e.,

$$\begin{aligned} &\lambda_{\omega_k}[(1+\alpha)\sin m\omega_k - \tfrac{\alpha}{2}(\sin(m+1)\omega_k + \sin(m-1)\omega_k)] \\ &= [(1-\alpha)\sin m\omega_k + \tfrac{\alpha}{2}(\sin(m+1)\omega_k + \sin(m-1)\omega_k)]. \end{aligned} \tag{6.2.18}$$

Therefore,

$$\lambda_{\omega_k} = \frac{(1-\alpha)\sin m\omega_k + \alpha \sin m\omega_k \cos\omega_k}{(1+\alpha)\sin m\omega_k - \alpha \sin m\omega_k \cos\omega_k} = \frac{1 - 2\alpha \sin^2 \frac{\omega_k}{2}}{1 + 2\alpha \sin^2 \frac{\omega_k}{2}}$$

and $|\lambda_{\omega_k}| \le 1$ for any k. In other words, the Crank-Nicolson scheme for the heat equation is stable in the l_2 norm. This method for stability analysis is called von Neumann method. It is adaptable to our problem (6.1.9), and the result is that the Crank-Nicolson scheme for (6.1.9) is stable.

6.2.2 Stability of the Crank-Nicolson Scheme for (6.1.10)

In practice, some parabolic PDE problems with variable coefficients can not be simplified as the heat equation, e.g., our equation (6.1.10). The stability of the numerical schemes for these problems can also be analyzed by the von Neumann method.

Consider the following scheme with variable coefficients:

$$\alpha_{1,m}^n u_{m+1}^{n+1} + \alpha_{0,m}^n u_m^{n+1} + \alpha_{-1,m}^n u_{m-1}^{n+1} = \beta_{1,m}^n u_{m+1}^n + \beta_{0,m}^n u_m^n + \beta_{-1,m}^n u_{m-1}^n, \tag{6.2.19}$$

where $m = 1, 2, \ldots, M-1$, $n = 0, 1, \ldots, N-1$. Suppose

$$|f_{m+1}^n - f_m^n| < C\Delta x, \quad |f_{m+1}^n - 2f_m^n + f_{m-1}^n| < C\Delta x^2,$$

and

$$|f_m^{n+1} - f_m^n| < C\Delta\tau$$

for $f = \alpha_1, \alpha_0, \alpha_{-1}, \beta_1, \beta_0$ and β_{-1}. Assume the error between u_m^n and its true value has the form

$$e_m^n = \lambda_\theta^n \mathrm{e}^{im\theta},$$

where θ can be any real number in the interval $[0, 2\pi]$, i is the imaginary unit, and $\mathrm{e}^{im\theta}$ is an exponential function. Substituting the error terms into (6.2.19) yields

$$\lambda_\theta(x_m, \tau^n) = \frac{\beta_{1,m}^n \mathrm{e}^{i(m+1)\theta} + \beta_{0,m}^n \mathrm{e}^{im\theta} + \beta_{-1,m}^n \mathrm{e}^{i(m-1)\theta}}{\alpha_{1,m}^n \mathrm{e}^{i(m+1)\theta} + \alpha_{0,m}^n \mathrm{e}^{im\theta} + \alpha_{-1,m}^n \mathrm{e}^{i(m-1)\theta}}.$$

If the amplification factor satisfies

$$|\lambda_\theta(x_m, \tau^n)| \le 1$$

for every point, the scheme (6.2.19) is stable. Usually, the condition $|\lambda_\theta(x_m, \tau^n)| \le 1$ is written as

$$|\beta_{1,m}^n \mathrm{e}^{i\theta} + \beta_{0,m}^n + \beta_{-1,m}^n \mathrm{e}^{-i\theta}| - |\alpha_{1,m}^n \mathrm{e}^{i\theta} + \alpha_{0,m}^n + \alpha_{-1,m}^n \mathrm{e}^{-i\theta}| \le 0, \tag{6.2.20}$$

because the latter is easier to use.

Now we consider the stability of the numerical scheme for the following parabolic PDE problem which includes the equation (6.1.10):

$$\begin{cases} \frac{\partial u}{\partial \tau} = a(x,\tau)\frac{\partial^2 u}{\partial x^2} + b(x,\tau)\frac{\partial u}{\partial x} + c(x,\tau)u + d(x,\tau), \\ \qquad x_l \le x \le x_r, \quad 0 \le \tau \le T, \\ u(x,0) = f(x), \quad x_l \le x \le x_r, \\ u(x_l,\tau) = f_l(\tau), \quad 0 \le \tau \le T, \\ u(x_r,\tau) = f_r(\tau), \quad 0 \le \tau \le T, \end{cases} \tag{6.2.21}$$

where $a(x,\tau) > 0$ on the domain $[x_l, x_u] \times [0,T]$. The Crank-Nicolson scheme for the problem (6.2.21) is

$$\begin{cases} \frac{u_m^{n+1}-u_m^n}{\Delta\tau} = \frac{a_m^{n+1/2}}{2}\left(\frac{u_{m+1}^{n+1}-2u_m^{n+1}+u_{m-1}^{n+1}}{\Delta x^2} + \frac{u_{m+1}^{n}-2u_m^{n}+u_{m-1}^{n}}{\Delta x^2}\right) \\ +\frac{b_m^{n+1/2}}{2}\left(\frac{u_{m+1}^{n+1}-u_{m-1}^{n+1}}{2\Delta x} + \frac{u_{m+1}^{n}-u_{m-1}^{n}}{2\Delta x}\right) + \frac{c_m^{n+1/2}}{2}(u_m^{n+1}+u_m^n) + d_m^{n+1/2}, \\ \quad m = 1,2,\ldots,M-1, \quad n = 0,1,\ldots,N-1, \\ u_0^{n+1} = f_l(\tau^{n+1}), \quad n = 0,1,\ldots,N, \\ u_M^{n+1} = f_r(\tau^{n+1}), \quad n = 0,1,\ldots,N, \\ u_m^0 = f(x_m), \quad m = 1,2,\ldots,M-1. \end{cases} \tag{6.2.22}$$

This scheme can be reformulated in the form of (6.2.19) with

$$\alpha_{1,m}^n = -\left(\frac{a_m^{n+1/2}}{2\Delta x^2} + \frac{b_m^{n+1/2}}{4\Delta x}\right)\Delta\tau,$$

$$\alpha_{0,m}^n = 1 + \frac{a_m^{n+1/2}}{\Delta x^2}\Delta\tau,$$

$$\alpha_{-1,m}^n = -\left(\frac{a_m^{n+1/2}}{2\Delta x^2} - \frac{b_m^{n+1/2}}{4\Delta x}\right)\Delta\tau,$$

$$\beta_{1,m}^n = -\alpha_{1,m}^n,$$

$$\beta^n_{0,m} = 2 - \alpha^n_{0,m},$$

$$\beta^n_{-1,m} = -\alpha^n_{-1,m},$$

based on which (6.2.20) holds (see Zhu, Wu and Chern [100]). That is to say, the Crank-Nicolson scheme for (6.2.21) is stable.

6.2.3 Crank-Nicolson Scheme for (6.1.12)

We rewrite the equation (6.1.12) in a concise form:

$$-u_t + a_1 u_{xx} + a_2 u_{xy} + a_3 u_{yy} + a_4 u_x + a_5 u_y = 0, \tag{6.2.23}$$

with the initial condition

$$u(0, x, y) = (y - X)^+,$$

where X is the cap rate and a_i $(i = 1, \ldots, 5)$ represents the coefficient function.

The Crank-Nicolson scheme for (6.2.23) is

$$\begin{aligned}
&-\tfrac{u^{n+1}_{m,i} - u^n_{m,i}}{\Delta t} + \tfrac{a_1}{2\Delta x^2}(u^{n+1}_{m+1,i} - 2u^{n+1}_{m,i} + u^{n+1}_{m-1,i} + u^n_{m+1,i} - 2u^n_{m,i} + u^n_{m-1,i}) \\
&+\tfrac{a_2}{8\Delta x \Delta y}(u^{n+1}_{m+1,i+1} - u^{n+1}_{m+1,i-1} - u^{n+1}_{m-1,i+1} + u^{n+1}_{m-1,i-1} + u^n_{m+1,i+1} - u^n_{m+1,i-1} \\
&-u^n_{m-1,i+1} + u^n_{m-1,i-1}) + \tfrac{a_3}{2\Delta y^2}(u^{n+1}_{m,i+1} - 2u^{n+1}_{m,i} + u^{n+1}_{m,i-1} + u^n_{m,i+1} - 2u^n_{m,i} + u^n_{m,i-1}) \\
&+\tfrac{a_4}{4\Delta x}(u^{n+1}_{m+1,i} - u^{n+1}_{m-1,i} + u^n_{m+1,i} - u^n_{m-1,i}) + \tfrac{a_5}{4\Delta y}(u^{n+1}_{m,i+1} - u^{n+1}_{m,i-1} + u^n_{m,i+1} - u^n_{m,i-1}) = 0
\end{aligned} \tag{6.2.24}$$

It has been shown in Zhu, Wu and Chern [100] this scheme is stable and convergent.

Now we can solve the equations (6.1.9)-(6.1.12) for pricing the three-period caps. To examine the effectiveness of the proposed PDE method for cap pricing, we apply it to the 1 year cap on USD forward rates and the 2 year cap on Euro forward rates, which are introduced in Section 4.4. We find that our PDE method produces the cap prices close to the market data. We demonstrate the numerical results below.

6.3 Numerical Results for Three-period Caps

For the 1 year cap on 3-month USD forward rates on Aug 10, 2010, it is based on 3 forward rates. The underlying forward rates have expire dates Nov 10, 2010, Feb 10, 2011, and May 10, 2011, respectively, and the corresponding maturity dates are Feb 10, 2011, May 10, 2011, and Aug 10, 2012. We set Aug 10, 2010 as the origin of time axis, then the underlying forward rates can be written as $F(t; 0.25, 0.5)$, $F(t; 0.5, 0.75)$ and $F(t; 0.75, 1.0)$. As before, we denote them as $F_1(t)$, $F_2(t)$ and $F_3(t)$ for simplicity. From the data in Section 4.3, we know that $X = 0.477\%$, $F_1(0) = 0.3703\%$, $F_2(0) = 0.4798\%$, $F_3(0) = 0.5553\%$, $\tau_1 = \tau_2 = \tau_3 = 0.25$, and $\sigma_1 = \sigma_2 = \sigma_3 = 0.7993$. Now we solve the equations (6.1.9)-(6.1.12), and get $G_1(0, F_1(0)) = 0.0003$, $G_2(0, F_2(0)) = 0.0011$, $G_3(0, F_2(0), F_3(0)) = 0.0022$. Using the discount factors

$$P(0, T_i) \approx \frac{1}{1 + 0.25F_1(0)} \prod_{k=1}^{i} \frac{1}{1 + \tau_k F_k(0)}, \quad i = 1, 2, 3,$$

we obtain that the cap price at $t = 0$ is 0.0009 by Eq. (6.1.13), which is the same as the market data.

For the 2 year cap on 6-month Euro forward rates on Aug 10, 2010, it is also based on three forward rates. The underlying forward rates expire on Feb 10, 2011, Aug 10, 2011 and Feb 10, 2012, respectively, and mature on Aug 10, 2011, Feb 10, 2012, and Aug 10, 2012 correspondingly. We still set Aug 10, 2010 as the origin of time axis, and denote $F_1(t) = F(t; 0.5, 1.0)$, $F_2(t) = F(t; 1.0, 1.5)$ and $F_3(t) = F(t; 1.5, 2.0)$. In this case, $X = 1.23\%$, $F_1(0) = 1.299\%$, $F_2(0) = 1.428\%$, $F_3(0) = 1.6136\%$, $\tau_1 = \tau_2 = \tau_3 = 0.5$, and $\sigma_1 = \sigma_2 = \sigma_3 = 0.5421$. Solving the equations (6.1.9)-(6.1.12), we get $G_1(0, F_1(0)) = 0.0017$, $G_2(0, F_2(0)) = 0.0031$, $G_3(0, F_2(0), F_3(0)) = 0.0052$. Taking the discount factors as

$$P(0, T_i) \approx \frac{1}{1 + 0.5F_1(0)} \prod_{k=1}^{i} \frac{1}{1 + \tau_k F_k(0)}, \quad i = 1, 2, 3,$$

we obtain that the cap price at $t = 0$ is 0.0048 by Eq. (6.1.13), which is very close to the market price 0.0047.

The above two examples indicate that our PDE approach is effective and accurate for pricing three-period caps. Now we illustrate the solution time of our PDE method in the following table.

Table 6.1: Solution time for solving the expected payoff functions of caplets

Caplet payoff	Computation grid	Solution time (second)
$G_1(0, F_1(0))$	1000*1000	0.1
$G_2(0, F_2(0))$	1000*1000	0.1
$G_3(0, F_2(0), F_3(0))$	250*250*100	68.61

In our experiments, the computation of the PDE method is implemented by Matlab, while the Monte Carlo method is realized in C++. We run all the programs on a laptop with Intel i5 processor (2 core @ 2.6G) and 4GB memory. For each cap, the same cap price is obtained by these two methods. We can see that it takes 68.81 seconds to price the three-period cap by the PDE method. In contrast, it takes 94.5 seconds to price the same cap by Monte Carlo simulation method (we simulate 10^5 forward rate paths and let $dt = 0.003$). That is to say, the PDE method saves 25.69 seconds relative to the Monte Carlo method.

6.4 Summary

In this chapter, we proposed a novel PDE method for pricing caps of three periods under the LFM. The PDEs governing the expected payoff functions of caplets were obtained by Feynman-Kac theorem. We established stable and convergent numerical schemes for these PDEs, and implemented numerical experiments to examine the effectiveness of our method. This PDE approach can deal with cap pricing under the LFM with nonzero correlations. It can also be used to evaluate caplets or caps under the LFM with stochastic volatility. It is more general than Black's formula for caps and computationally faster than Monte Carlo method. However, the dimension of the PDEs is mounting up as the number of caplets (which constitute the cap) increases, and the complexity for solving the PDEs grows quickly. Therefore, the PDE approach has its edge over Monte Carlo method only for pricing caps containing several caplets.

Chapter 7

Conclusions and Suggestions for Future Studies

In this thesis, we firstly summarized different methods for valuing the American vanilla options on a single stock, and testified their congruence by numerical experiments. Then we resorted to one of those methods, i.e., the power penalty method, for attacking American put options on zero-coupon bonds. The new problem is a LCP similar to that of American options on a single stock, while the key variable of the new problem is short-term rate and the boundary conditions involve zero-coupon bond price. We chose the CKLS model to describe the short-term rate, and solved the bond price numerically in the beginning, then we figured out the option value and the optimal exercise boundary by power penalty method. The main advantages of the penalty approach are: (i) when sophisticated discretization methods other than standard central or upstream weighting methods are employed, it would reduce the computational cost and time; (ii) it may converge faster than the PSOR method for American options with early exercise constraint.

Next, we compared the performance of the one-factor LFM and the CIR model in terms of pricing interest rate derivatives. Through a number of cases of calibration to the market data in August 2010, we got the following conclusions:

1. The LFM with $\rho_{i,j} = 0$ is closer to the real market than the CIR model.
2. The LFM with an exogenous correlation matrix (i.e., $\rho_{i,j} = 1 - 0.025 * |i - j|$) outperforms the LFM with $\rho_{i,j} = 0$ $(i \neq j)$.
3. The LFM with parametric correlations outperforms the LFM with $\rho_{i,j} = 1 - 0.025 * |i - j|$.

Our calibration uses the data of cap prices and is implemented through Monte Carlo method and the downhill simplex method in multidimensions, while others use the data of swaption prices and swaption volatilities, and perform their calibration through closed-form approximate formulas. Moreover, our procedure for simulating the forward rates under the LFM differs from others. We deal with the dynamics of forward rates other than the dynamics of the logarithm of forward rates.

After that, we also compared the prices of interest rate derivatives under the LFM ($\rho_{i,j} = 1$) with those under the CIR model using Monte Carlo method and artificial parameters. It is observed that they have underestimated prices under the former model than under the latter model.

The rest of the thesis discussed pricing methods of derivatives under the LFM. We presented the Monte Carlo simulation of forward rate paths, and illustrated the simulation procedures for pricing caps, European options on coupon-bearing bonds and swaptions. When computing the prices of swaptions under the LFM, we circumvented swap rates and started from forward rates, since swaptions can be defined in terms of forward rates. Also, we demonstrated the Monte Carlo method for pricing caps under the two-factor LFM, but found that it would not outperform the one-factor LFM if we did not adopt appropriate correlations in the two-factor LFM.

At last, we developed a PDE approach for pricing caps under the LFM with $\rho_{i,j} \neq 0$. We introduced the PDEs governing the expected payoff functions of the caplets, which constitute the cap. Solving the PDEs and discounting the expectations back to the initial time, we got the price of the cap. This PDE approach can be extended to pricing caplets and caps under the one-factor LFM with stochastic volatility. It is computationally faster than the Monte Carlo method for cap pricing. However, this advantage exists only for pricing caps containing several caplets.

We shall continue our work in the following related topics.

1. Piecewise-constant instantaneous-volatility structures and parametric volatility structures for the LFM have been demonstrated in Brigo [19]. The LFM with such assumptions for the volatility may possess better fitting capability to the LIBOR market than the LFM with constant volatility. It should be interesting to include different volatility structures in the calibration of the LFM ($\rho_{i,j} \neq 0$).
2. We use Monte Carlo method and the downhill simplex method in multidimensions in the calibration of the LFM ($\rho_{i,j} \neq 0$). It takes tens of minutes to compute the function value on each vertex of the simplex even we just generate 10^5 forward rate paths. How to speed up the computation and apply the variance reduction techniques reviewed in Boyle, Broadie and Glasserman [14] to the Monte Carlo simulation is an interesting topic.
3. The extensions of the LFM, such as the local volatility forward rate model, the jump-diffusion forward rate model and the stochastic volatility forward

rate model, could be calibrated to the market data as what was done for the LFM with $\rho_{i,j} \neq 0$.

4. Our PDE method for pricing caps under the one-factor LFM have been used only for three-period caps. It will be computationally expensive but meaningful to extend it for pricing caps of more periods. This PDE method can also be applied to price caps under the one-factor LFM with stochastic volatility.

Bibliography

[1] W. Allegretto, G. Barone-Adesi, R. Elliott, Numerical evaluation of the critical price of American options, *The European Journal of Finance*, Vol. 1, 1995, pp. 69-78.

[2] W. Allegretto, Y. Lin, H. Yang, A novel approach to the valuation of American options, *Global Finance Journal*, Vol. 13, 2002, pp. 17-28.

[3] W. Allegretto, Y. Lin, H. Yang, Finite element error estimates for a nonlocal problem in American option valuation, *SIAM Journal on Numerical Analysis*, Vol. 39, No. 3, 2002, pp. 834-857.

[4] W. Allegretto, Y. Lin, H. Yang, Numerical pricing of American put options on zero-coupon bonds, *Applied Numerical Mathematics*, Vol. 23, 2003, pp. 113-134.

[5] K. Amin, A. Khanna, Convergenc of American option values from discrete- to continuous-time financial models, *Mathematical Finance*, Vol. 4, No. 4, 1994, pp. 289-304.

[6] L. Andersen, J. Andreasen, Volatility skews and extensions of the libor market model, *Applied Mathematical Finance*, Vol. 7, No. 1, 2000, pp. 1-32.

[7] L. Andersen, R. Brotherton-Ratcliffe, Extended libor market models with stochastic volatility, Available at SSRN: http://ssrn.com/abstract=294853 or doi:10.2139/ssrn.294853, 2001.

[8] L. Angermann, S. Wang, Convergence of a fitted volume finite volume method for the penalized Black-Scholes equation governing European and American option pricing, *Numerische Mathematik* , Vol. 106, 2007, pp. 1-40.

[9] G. Barone-Adesi, R. Elliott, Approximations for the values of American options, *Stochasitic Analysis and Applications*, Vol. 9, No. 2, 1991, pp. 115-131.

[10] G. Barone-Adesi, R. Whaley, Efficient analytic approximation of American option values, *The Journal of Finance*, Vol. 42, No. 2, 1987, pp. 301-320.

[11] A. Bensoussan, On the theory of option pricing, *Acta Applicandae Mathematicae* Vol. 2, No. 2, 1984, pp. 139-158.

[12] F. Black, M. Scholes, The pricing of options and corporate liabilities, *Journal of Political Economy*, Vol. 81, No.3, 1973, pp. 637-654.

[13] F. Black, P. Karasinski, Bond and option pricing when short rates are lognormal, *Financial Analysts Journal*, Vol. 47, No. 4, 1991, pp. 52-59.

[14] P. Boyle, M. Broadie, P. Glasserman, Monte Carlo methods for security pricing, *Journal of Economic Dynamics and Control*, Vol. 21, No. 8-9, 1997, pp. 1267-1321.

[15] P. Boyle, Options: a Monte Carlo approach, *Journal of Financial Economics*, Vol. 4, No. 3, 1977, pp. 323-338.

[16] A. Brace, M. Gatarek, M. Musiela, The market model of interest rate dynamics, *Mathematical Finance*, Vol. 7, No. 2, 1997, pp. 127-155.

[17] M. J. Brennan, E. S. Schwartz, The valuation of American put options, *The Journal of Finance*, Vol. 32, No. 2, 1977, pp. 449-462.

[18] D. Brigo, F. Mercurio, M. Morini, The LIBOR model dynamics: Approximations, calibration and diagnostics, *European Journal of Operational Research*, Vol. 163, No. 1, 2005, pp. 30-51.

[19] D. Brigo, F. Mercurio, Interest Rate Models - Theory and Practice, *Springer-Verlag Berlin Heidelberg*, 2006.

[20] M. Broadie, P. Glasserman, Pricing American-style securities using simulation, *Journal of Economic Dynamics and Control*, Vol. 21, No. 8-9, 1997, pp. 1323-1352.

[21] M. Broadie, P. Glasserman, A Stochastic Mesh Method for Pricing High-Dimensional American Options, 1998, working paper.

[22] D. S. Bunch, H. Johnson, The American put option and its critical stock price, *The Journal of Finance*, Vol. 55, No. 5, 2000, pp. 2333-2356.

[23] K. C. Chan, G. A. Karolyi, F. A. Longstaff, Anthony B. Sanders, An empirical comparison of alternative models of the short-term interest rate, *The Journal of Finance*, Vol. 47, No. 3, 1992, pp. 1209-1227.

[24] R. R. Chen, L. Scott, Pricing interest rate options in a two-factor Cox-Ingersoll-Ross model of the term structure, *The Review of Financial Studies*, Vol. 5, No. 4, 1992, pp. 613-636.

[25] M. Chesney, R. Elliott, R. Gibson, Analytical solutions for the pricing of American bond and yield options, *Mathematical Finance*, Vol. 3, No. 3, 1993, 277-294.

[26] N. Clarke, K. Parrott, Multigrid for American option pricing with stochastic volatility, *Applied Mathematical Finance*, Vol. 6, 1999, pp. 177-195.

[27] R. Cont, E. Voltchkova, A finite difference scheme for option pricing in jump diffusion and exponential Levy models, *SIAM Journal on Numerical Analysis*, Vol. 43, No. 4, 2005, pp. 1596-1626.

[28] J. C. Cox, S. A. Ross, M. Rubinstein, Option pricing: A simplified approach, *Journal of Financial Economics*, Vol. 7, No. 3, 1979, pp. 229-263.

[29] J. C. Cox, J. E. Ingersoll, S. A. Ross, A theory of the term structure of interest rates, *Econometrica*, Vol. 53, No. 2, 1985, pp. 385-407.

[30] M. A. H. Dempster, J. P. Hutton, D.G. Richards, LP valuation of exotic American options expoiting structure, *Journal of computational Finance*, Vol. 2, No. 1, 1998, pp. 61-84.

[31] M. A. H. Dempster, J. P. Hutton, Pricing American stock options by linear programmings, *Mathematical Finance*, Vol. 9, No. 3, 1999, pp. 1229-254.

[32] N. Denson, M. Joshi, Fast and accurate Greeks for the LIBOR market model, *The Journal of Computational Finance*, Vol. 14, No. 4, 2011, pp. 115-140.

[33] C. M. Elliott, J. R. Ockendon, Weak and Variational Methods for Moving Boundary Problems, *Pitman Publishing*, 1982.

[34] P. A. Forsyth, K. R. Veztal, Quadratic convergence for valuing American options using a penalty method, *SIAM Journal on Scientific Computing*, Vol. 23, 2002, pp. 2095-2122.

[35] M. C. Fu, S. B. Laprise, D. B. Madan, Y. Su, R. Wu, Pricing American Options: A Comparison of Monte Carlo Simulation Approaches, working paper, 2000.

[36] R. Geske, H. Johnson, The American put options valued analytically, *The Journal of Finance*, Vol. 39, No. 5, 1984, pp. 1511-1524.

[37] P. Glasserman, Monte Carlo Methods in Financial Engineering, *Springer*, 2003.

[38] P. Glasserman, S. G. Kou, The term structure of simple forward rates with jump risk, *Mathematical Finance*, Vol. 13, No. 3, 2003, pp. 383-410.

[39] P. Glasserman, N. Merrener, Numerical solution of jump-diffusion LIBOR market models, *Journal of Computational Finance*, Vol. 7, No. 1, 2003, pp. 1-27.

[40] P. Glasserman, X. Zhao, Arbitrage-free discretization of lognormal forward LIBOR and swap rate models, *Finance and Stochastics*, Vol. 4, No. 1, 2000, pp. 35-68.

[41] D. Grant, G. Vora, D. Weeks, Simulation and the Early-Exercise Option Problem, *Journal of Financial Engineering*, Vol. 5, No. 3 , 1996, pp. 211-227.

[42] D. Grant, G. Vora, D. Weeks, Path-Dependent Options: Extending the Monte Carlo Simulation Approach, *Management Science*, Vol. 43, 1997, pp. 1589-1602.

[43] Y. Halluin, P. A. Forsyth, G. Labahn, A penalty method for American options with jump diffusion process, *Numerische Mathematik*, Vol. 97, No. 2, 2004, pp. 321-352.

[44] H. Han, X. Wu, A fast numerical method for the Black-Scholes equation of American options, *SIAM Journal on Numerical Analysis*, Vol. 41, No. 6, 2003, pp. 2081-2095.

[45] D. Heath, R. Jarrow, A. Morton, Bond pricing and the term structure of the interest rates: A new methodology, *Econometrica*, Vol. 60, No. 1, 1992, pp. 77-105.

[46] S. Heston, A closed-form solution for options with stochastic volatility with applications to bond and currency options, *The Review of Financial Studies*, Vol. 6, No. 2, 1993, pp. 327-343.

[47] J. Hull, A. White, Pricing interest rate derivative securities, *The review of financial studies*, Vol. 3, No. 4, 1990, pp. 573-592.

[48] J. Hull, Options, futures and other derivatives, *Pearson Prentice Hall*, 2009.

[49] T. S. Y. Ho, S. B. Lee, Term structure movements and pricing interest rate contingent claims, *Journal of Finance*, Vol. 41, No. 5, 1986, pp. 1011-1029.

[50] C. Hunter, P. Jackel, M. Joshi, Getting the drift, *Risk*, Vol. 14, No. 7, 2001, pp. 81-84.

[51] S. D. Jacka, Optimal stopping and the American put, *Mathematical Finance*, vol. 1, No. 2, 1991, pp. 1-14.

[52] P. Jaillet, D. Lamberton, B. Lapeyre, Variational inequalities and the pricing of American options, *Acta Applicandae Mathematicae*, Vol. 21, 1990, pp. 263-289.

[53] F. Jamshidian, LIBOR and swap market models and measures, *Finance and Sthochastic*, vol. 1, No. 4, 1997, pp 293-330.

[54] F. Jamshidian, An exact bond option pricing formula, *The Journal of Finance*, Vol. 44, 1989, pp. 205-209.

[55] H. Johnson, An Analytic Approximation for the American Put Price, *Journal of Financial and Quantitative Analysis*, Vol. 18, No. 1, 1983, pp. 141-148.

[56] M. Joshi, R. Rebonato, A stochastic-volatility, displaced-diffusion extension of the LIBOR market model, *Quantitative Finance*, Vol. 3, No. 6, 2003, pp. 458-469.

[57] M. Joshi, A. Stacey, New and robust drift approximations for the LIBOR market model, *Quantitative Finance*, Vol. 8, No. 4, 2008, pp. 427-434.

[58] I. Karatzas, On the pricing of American options, *Applied Mathematics and Optimization*, Vol. 17, No. 1, 1988, pp. 37-60.

[59] S. Li, S. Li, Pricing American interest rate option on zero-coupon bond numerically, *Applied Mathematics and Computation*, Vol. 175, 2006, pp. 834-850.

[60] F. Longstaff, E. Schwartz, Valuing American options by simulation: a simple least-squares approach, *The Review of Financial Studies Spring*, Vol. 14, No. 1, 2001, pp. 113-147.

[61] L. W. MacMillan, Analytic approximation for the American put option, *Advances in Futures and Option Research*, Vol. 1, 1986, pp. 119-139.

[62] H. Mckean, Appendix: A free boundary problem for the heat equation arising from a problem in mathematical economics, *Industrial Management Review*, Vol. 6, No. 2, 1965, pp. 32-39.

[63] R. C. Merton, Theory of rational option pricing, *Bell Journal of Economics and Management Science*, Vol. 4, No. 1, 1973, pp. 141-183.

[64] K. R. Miltersen, K. Sandmann, D. Sondermann, Closed form solutions for term structure derivatives with lognormal interest rates, *The Journal of Finance*, Vol. 52, No. 1, 1997, pp. 409-430.

[65] P. Moerbeke, On optimal stopping and free boundary problems, *Archives for Rational Mechanics and Analysis*, Vol. 60, No. 2, 1976, pp. 101-148.

[66] M. Musiela, M. Rutkowshi, Continuous-time term structure models: foward measure approach, *Finance and Stochastics*, Vol. 1, No. 4, pp. 261-292.

[67] M. Musiela, M. Rutkowshi, Martingale Methods in Financial Modelling, Second edition, *Springer*, 2005.

[68] R. Myneni, The pricing of the American option, *The Annals of Applied Probability*, Vol. 2, No. 1, 1992, pp. 1-23.

[69] J. A. Nelder, R. Mead, A simplex method for function minimization, *Computer Journal*, Vol. 7, No. 4, 1965, pp. 308-313.

[70] K. B. Nowman, G. Sorwar, Derivative prices from interest rate models: results for Canada, Hong Kong, and United States, *Intern ational Review of Financial Analysis*, Vol. 14, 2005, 428-438.

[71] C. W. Oosterlee, On multigrid for linear complementarity problems with application to American-style options, *Electronic Transactions on Numerical Analysis*, Vol. 15, 2003, pp.165-185.

[72] V. Piterbarg, A stochastic volatility forward Libor model with a term structure of volatility smiles, *Applied Mathematical Finance*, Vol. 12, 2003, pp. 147-185.

[73] D. M. Pooley, P. A. Forsyth, K. R. Vetzal, Numerical convergence properties of option pricing PDEs with uncertain volatility, *IMA Journal of Numerical Analysis*, Vol. 23, 2003, pp. 241-267.

[74] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical Recipes, Third edition, *Cambridge University Press*, 2007.

[75] R. Rebonato, On the simultaneous calibration of multifactor lognormal interest rate models to Black volatilities and to the correlation matrix, *Journal of Computational Finance*, Vol. 2, No. 4, 1999, pp. 5-27.

[76] R. Rebonato, Volatility and Correlation, First edition, *Wiley*, 1999.

[77] R. Rebonato, Modern Pricing of Interest-Rate Derivatives, *Princeton University Press*, 2002.

[78] J. Schoenmakers, B. Coffey, LIBOR Rate Models, Related Derivatives and Model Calibration, *WIAS*, 1999.

[79] G. Sorwar, G. Barone-Adesi, W. Allegretto, Valuation of derivatives based on single-factor interest rate models, *Global Finance Journal*, Vol. 18, No. 2, 2007, pp. 251-269.

[80] D. Y. Tangman, A. Gopaul, M. Bhuruth, A fast high-order finite difference algorithm for pricing American options, *Journal of Computational and Applied mathematics*, 2007, doi:10.1016/j.cam.2007.10.044.

[81] D. Y. Tangman, A. Gopaul, M. Bhuruth, Numerical pricing of options using high-order compact finite difference schemes, *Journal of Computational and Applied mathematics*, Vol. 218, 2008, pp. 270-280.

[82] J. Tilley, Valuing American Options in a Path Simulation Model, *Transactions of the Society of Actuaries*, 45, 1993, pp. 83-104.

[83] I. Z. Valeri, American option pricing with transaction costs, 2003, http://129.3.20.41/eps/fin/papers/0311/0311012.pdf .

[84] O. Vasicek, An equilibrium characterization of the term structure, *Journal of Financial Economics*, Vol. 5, No. 2, 1977, pp. 177-188.

[85] C. Vazquez, An upwind numerical approach for an American and European option pricing model, *Applied Mathematics and Computation*, Vol. 97, 1998, pp. 273-286.

[86] S. Wang, A fitted finite volume method for the Black-Scholes equation governing European option pricing, *IMA journal of Numerical Analysis*, Vol. 24, 2004, pp. 699-720.

[87] S. Wang, X. Q. Yang, K.L. Teo, Power penalty method for a linear complementarity problem arising from the American option valuation, *Journal of Optimization Theory and Applications*, Vol. 129, No. 2, 2006, pp. 227-254.

[88] S. Wang, X. Q. Yang, A power penalty method for linear complementarity problems, *Operations Research Letter* Vol. 36, 2008, pp. 211-214.

[89] P. Wilmott, S. Howison, J. Dewynne, The Mathematics of Financial Derivatives, *Cambridge University Press*, 1995.

[90] H. Wong, J. Zhao, An Artificial Boundary Method for American Option Pricing under the CEV Model, *SIAM Journal on Numerical Analysis*, Vol. 46, No. 4, 2008, pp. 2183-2209.

[91] H. Wong, J. Zhao, An artificial boundary method for the Hull-White model of American interest rate derivatives, *Applied Mathematics and Computation*, Vol. 217, No. 9, 2011, pp. 4627-4643.

[92] L. Wu, Fast at-the-money calibration of LIBOR market model through Lagrange multipliers, *Journal of Computational Finance*, Vol. 6, No. 2, 2002, pp. 39-77.

[93] L. Wu, F. Zhang, Libor market model with stochastic volatility, *Journal of Industrial and Management Optimization*, Vol. 2, No. 2, 2006, pp. 199-227.

[94] K. Zhang, S. Wang, Pricing options under jump diffusion process with a fitted finite volume method, *Applied Mathematics and Computation*, Vol. 201, 2008, pp. 398-413.

[95] K. Zhang, X. Q. Yang, K. L. Teo, Convergence analysis of a monotonic penalty method for American option pricing, *Journal of Mathematical Analysis and Applications* Vol. 348, 2008, pp. 915-926.

[96] K. Zhang, S. Wang, Convergence property of an interior penalty approach to pricing American option, *Journal of Industrial and Management Optimization*, Vol. 7, No. 2, 2011, pp. 435-447.

[97] T. Zhang, Numerical Methods for the pricing problem of American put option, *Acta Mathematicae Applicatae Sinica (Chinese Series)*, Vol. 25, No. 1, 2002, pp. 113–122.

[98] H. J. Zhou, K. F. Yiu, L. K. Li, Evaluating American Put Options on Zero-coupon Bonds by a Penalty Method, *Journal of Computational and Applied Mathematics*, Vol. 235, No. 13, 2011, pp. 3921-3931.

[99] S. P. Zhu, An exact and explicit solution for the valuation of American put options, *Quantitative Finance*, Vol. 6, No. 3, 2006, pp. 229-242.

[100] Y. Zhu, X. Wu, I. L. Chern, Derivative securities and Differcence methods, *Springer*, 2004.

[101] R. Zvan, P. A. Forsyth and K. R. Veztal, Penalty methods for American options with stochastic volatility, *Journal of Computational and Applied mathematics*, Vol. 91, 1998, pp. 199-218.