Extension and Application of LIBOR Market Model

By

Fan Zhang

A Thesis Submitted to
The Hong Kong University of Science & Technology
in Partial Fulfillment of the Requirements for
the Degree of Doctor of Philosophy
in Mathematics

Hong Kong, July 17, 2003
Authorization

I hereby declare that I am the sole author of the thesis.

I authorize the University of Science and Technology to lend this thesis to other institutions or individuals for the purpose of scholarly research.

I further authorize the University of Science and Technology to reproduce the thesis by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

Fan Zhang
Extension and Application of LIBOR Market Model

By

Fan Zhang

This is to certify that I have examined the above PhD thesis and have found that it is complete and satisfactory in all respects, and that any and all revisions required by the thesis examination committee have been made.

[Signature]

Dr. Li-Xin Wu, Supervisor

[Signature]

Prof. Shiu-Yuen Cheng, Head of Department
Department of Mathematics
June, 2003
Acknowledgment

I would like to express my deep gratitude to my supervisor Dr. Li-Xin Wu. He stimulated me with challenging topics, directed me in the research, encouraged me to overcome the difficulties and provided me with unreserved help.

I would like to thank the Department of Mathematics and HKUST for providing me with the opportunity to study here.

I would like to thank Dr. Yue-Kuen Kwok and Dr. Shi-Qing Ling for their warm help and valuable suggestions. My gratitude should be also passed to Dr. Jin Zhang, Prof. Raymond Chan and Prof. Leonard Cheng for their service as examiners on my thesis.

I would like to thank Yi Xiang for his helpful discussions. I would like to thank Xiao-Ling Wang for kindly sharing Latex skills with me.

Finally I wish to thank my family and friends, especially my parents whose love and encouragement have supported me all of my life.
# Contents

Title Page i

Authorization Page ii

Signature Page iii

Acknowledgment Page iv

Table of Contents v

Abstract viii

1 Introduction 1

2 Evolution of LIBOR Market Model 4

3 Convexity Adjustment for Constant Maturity Instruments 8
   3.1 Introduction ............................................. 8
   3.2 Methodology and computation for expectation of constant maturity yield 10
   3.3 Valuation of constant maturity instruments .......................... 13
   3.3.1 Valuation of swap ....................................... 13
   3.3.2 Valuation of cap and floor ................................. 15
   3.3.3 Valuation of swaption .................................... 17
Appendix 83

A Arbitrage-free Model with Stochastic Volatility 83
   A.1 Proof of Proposition 5.2.1 83
   A.2 Proof of Theorem 5.2.1 86

B Details of Some Derivations 88
Extension and Application of LIBOR Market Model

By

Fan Zhang

Department of Mathematics
The Hong Kong University of Science & Technology

Abstract

LIBOR market model is the benchmark model for interest rate derivatives. It has been a challenge to improve the standard model to better capture market dynamics. In this thesis two extensions of standard LIBOR market model are presented, some applications of this model are also discussed.

Constant maturity instruments such as swaps, caps and floors have gained increasing popularity in recent years. A new approach which is motivated by traditional convexity adjustment and change of measure is adopted in this thesis to evaluate these constant maturity instruments, while closed-form solutions are given to apply this more accurate and computational efficient technique.

Volatility skews are reflected in many markets. The generalized CEV market model is an extension to the standard market model. The great advantage of the CEV model is its capacity to produce the volatility skew that is pronounced in the swaption prices. As
an application of the calibration of standard LIBOR CEV model by Wu (2003), we use a practical problem to illustrate how to calibrate the volatility coefficients and elasticity of the CEV market model using the matrix eigenvalue decomposition and Hessian-based algorithm. This work can also be regarded as a numerical complement to the calibration of LIBOR market model.

The phenomena of non-monotonic volatility smiles are persistent in many major markets. We extend the standard market model to cope with this feature by allowing stochastic volatility. To achieve analytical tractability, we adopt a multiplicative stochastic factor for the volatility functions of all relevant forward rates. The stochastic factor follows a squared-root process, and it can be correlated with the forward rate processes. Approximate processes for forward rates and swap rates are introduced and closed-form pricing formulae are derived. We develop a fast Fourier transform algorithm for the implementation of the formula. The approximations are well supported by pricing accuracy. By adjusting the correlation between the forward rates and the volatility in a way consistent with intuition, we can generate volatility smiles or skews of the swaption prices similar to those observed in the markets.
Chapter 1

Introduction

Much work has been done in the area of modeling the term structure of interest rate dynamics. Heath, Jarrow and Morton (1992) have established a general model for pricing interest rate related derivatives, and HJM framework is well-known as an arbitrage-free and all embracing model, but it can not be very conveniently applied.

In light of this, Brace et al (1997), Miltersen et al (1997) and Jamshidian (1997) all chose to model forward term rate and hence set up LIBOR (London Interbank Offer Rate) market model. Tractable enough to allow fast calibration to market quoted caps, this flexible model supports multiple factors and rich volatility structure, and it has since established itself as a benchmark model for interest rate options. Jamshidian (1997) developed a similar model for swap rate so that swaption pricing is no longer a difficulty. Although LIBOR market model has many advantages, it does not generate volatility skews or smiles which are pronounced in markets. This drawback motivates us to extend the standard LIBOR market model.

Owing to the importance of market model in application, we devote Chapter 2 to the introduction of the evolution of the model. We show the origin of this model and introduce some important extensions of standard model.
We show the valuation of constant maturity instruments in Chapter 3. These instruments use other treasury rates to replace 6M LIBOR rate and may assume the payments made in arrear. The difficulty for pricing those derivatives lies in how to calculate the expectation of the cash flow at cash-flow date. Although traditional convexity adjustment is easy to apply in practice, accuracy seems a concern. J. Sidenius (1999) developed the convexity adjustment approach for LIBOR-in-arrears swaps. However, this methodology can be improved by using the forward measure of the cash-flow day. We would like to present the approximate pricing formulae given the correlation between treasury and LIBOR using change of measure technique and an improved convexity adjustment formula.

Our approach can achieve any accuracy by including a sufficiently high number of terms, which greatly improves the original convexity adjustment method. The solution to swaption pricing obtained by means of summation of value of each cash flow under its corresponding forward measure is also included in this chapter.

In Chapter 4, we investigate the calibration problem for the CEV (constant elasticity of variance) model, which is an extension of the standard LIBOR market model. Researches showed that CEV model was capable of producing volatility skews and permitting analytical pricing formulae for caplets. Wu (2001) presented an efficient and robust methodology to calibrate this CEV model. However, his approach has not yet be tested with real data. In this chapter, we use a practical example to illustrate the calibration of volatility coefficients and elasticity using matrix eigenvalue decomposition and Hessian-based algorithm, which can be regarded as a complement to Wu’s work.

Although CEV model can generate volatility skews for caplets, its capacity to generate volatility skew in swaption price is limited. Hence we must consider other extension. In Chapter 5, we develop the original LIBOR market framework by allowing
stochastic volatility which has been accepted by market practitioners as the most important factor attributed for the formulation of volatility smiles/skews in the interest rate derivative market.

Specifically, we formulate the model by adopting a multiplicative stochastic factor for the volatilities of forward rates. This factor can be correlated with the Brownian motions driving the forward rates. We take a squared-root process for stochastic factor and introduce the approximate state-variable processes that carry analytical tractability. This new extension can be set in the footing of arbitrage free model. Closed-form solutions for caplets and swaptions are derived based on this model.

For the purpose of fast implementation, we show how to use Fast Fourier Transform technology. This FFT method can greatly improve the numerical evaluation of the closed-form formula.

The conclusions are presented in the last chapter.
Chapter 2

Evolution of LIBOR Market Model

Traditionally, models of interest rate have dealt with continuous compounded, instantaneous rates. Based on this instantaneous, continuously compounded spot rate, forward rates, a continuum of discount factors are constructed. Options then are evaluated using the risk neutral measure, which is martingale measure using money market account as the numeraire. Yet, in the actual market place, the rates applicable to interest rate derivatives, foremost among them LIBOR and swap derivatives, are quoted for accrual periods of at least one month, commonly three or six months, using simple rather than continuously compounding. Moreover, the market quotes liquid caps and swaptions in terms of implied Black-Scholes volatilities, implicitly assuming forward rates and swap rates follow lognormal processes with the quoted volatilities.

The traditional models take a continuum of initial instantaneous forward rates or just assume discount factors to be given, and construct a continuum of processes, making assumptions either on the dynamics of the instantaneous spot rate (possibly dependent on several state variables) or volatilities of instantaneous forward rates. In order to match market quoted prices of caps or European swaptions, practitioners need to suitably parametrize degrees of freedom in the specified dynamics of instantaneous
spot rate or volatilities of instantaneous forward rates, and then calibrate these parameters to quoted prices by a multidimensional and often highly computationally intensive, numerical root searching algorithm. The resultant processes for forward LIBOR or swap rates are analytically intractable, and generally bear no resemblance to lognormality.

By a direct hedging argument, Neuberger (1990) derived the industry standard Black-Scholes formula for European swaptions. However, a term structure of swap rates (or LIBOR rates), which is necessary for modelling of more complex derivatives, such as Bermudan swaptions, was not developed. Sandmann and Sondermann (1993, 1994) proposed a lognormal model for the effective rate and showed that it circumvented certain instabilities (particularly with Eurodollars) presented in lognormal, continuously compounded rate models. This was further developed within the framework of HJM (1992) by Goldys et al. (1994) and Musiela (1994). Continuing in this framework, Brace et al. (1997) and Miltersen et al. (1997) all shifted the emphasis from instantaneous rate to LIBOR rates. They presented an arbitrage-free interest rate model, in which forward LIBOR rates followed lognormal processes (described in details in the latter chapter), leading to the Black (1976) pricing formula for caps and floors, which had been used by market practitioners. And for this reason this model is named the "market model". Such a model is now automatically calibrated to caplet prices and can be used to evaluate more complex products like captions and callable capped floating rate notes. A similar model for swap rates and swap rate derivatives was developed by Jamshidian (1997). His so-called Swap Market Model led to the Black formula for swaptions. Andersen and Andreasen (2000) developed a very good approximation of forward swap rate dynamics by lognormal process. Such approximation led to the Black’s formula for European swaption as well and achieved consistency with caplet pricing.

There are several advantages of the market model in comparison with the tradi-
tional models, such as the instantaneous spot rate models (e.g. Vasicek (1977), Cox, Ingersoll and Ross (1985), and Hull and White (1990)) and models for instantaneous forward interest rates (Heath, Jarrow and Morton (1992) and Ritchken and Sankarasubramanian (1995)). Firstly, the use of Black formula for option prices makes calibration of market models very simple. The quoted implied Black volatilities can directly be inserted in the model, avoiding the numerical fitting procedures that are needed for the spot rate or forward rate models. Secondly, the market model are based on observable market rates, such as LIBOR rates and swap rates. Hence, one does not need the (unobserved) instantaneous short rate or instantaneous forward rates to price and hedge caps and swaptions. Given the advantages of the market models, it is not surprising these models have received a lot of attention recently.

Although market model has many advantages, the basic assumption of the LIBOR market model, lognormally distributed LIBOR rates, is, however, increasingly being found to be violated in many important cap and swaption markets. In particular, implied Black (1976) volatilities of caplet and swaption prices often tend to be decreasing functions of the strike and coupon, respectively, indicating a fat left tail of the empirical forward rate distributions relative to lognormality. This so-called volatility skew is currently most pronounced in the Japanese LIBOR market but also exists in US and German markets, among others. The presence of the volatility skew motivates the formulation of models where the diffusion coefficients of the discrete forward rates are non-linear functions of the rates themselves. For this reason Andersen and Andreasen (2000) adopted CEV (constant elasticity of variance) extension of LIBOR market model. CEV extension model has proved to be able to generate volatility skews. Andersen and Andreasen also gave the closed-form formula for pricing caplets as well as swaptions. Calibration then can be done with the implied CEV volatilities, in a way similar to the calibration of the standard market model (Wu, 2003).
Another important feature reflected by the market is volatility smile. However, since CEV model fails to cope with non-monotonic smiles, Glasserman and Kou (2000) have developed a new term structure with jump-diffusion dynamics of forward rates. Motivation for including jumps comes from several sources. First, there is evidence of the importance of jumps in capturing both returns and option prices in other markets, including equity and foreign exchange. Second, jumps in forward rates can also be used to try to reproduce the patterns in implied volatilities derived from market prices of interest rate derivatives, thus to achieve the goal of generating volatility smiles/skews by taking different mean jump sizes. Glasserman and Merener (2001) also obtained approximate closed-form solutions for caplets and swaptions.

In general, the jump-diffusion model is theoretically appealing, but market practitioners tend to believe that jump risk is not so important as stochastic volatility. Some recent research works are about extension of market model with stochastic volatility. Andersen and Brotherton-Ractliffe (2002) extended CEV model by allowing stochastic volatility. They adopted asymptotic expansion technique to deal with the market model with stochastic volatility and presented some analytical formulae.

All in all, the extension and calibration with market model has not yet been completely solved, there's still much work to do in this field.
Chapter 3

Convexity Adjustment for Constant Maturity Instruments

3.1 Introduction

A constant maturity instrument is a derivative in which we replace 6M LIBOR involved with other rates such as swap rate or treasury rate without altering the payment frequency. The constant maturity instruments are much better suited for hedging many exposures than ordinary, plain vanilla instruments. Also they can provide a flexible and market efficient access to long-dated interest rate, on the liability side, they can offer the ability to hedge long-dated position. For all these reasons it is not surprising that the constant maturity instrument market now trades in large volumes, both interbank and between corporations and financial institutions.

Because of the increasing size of constant maturity swap (CMS) market, the market has seen its margin eroding. As a consequence more accurate pricing methods become useful as market prices get more sharply defined and bid-offer spreads narrow.
Those constant maturity instruments were firstly priced by forward approximation. Until recently, most market participants improved this method by a "convexity adjustment", and this convexity adjustment turns out to be more accurate. Sidenius (1999) presented a modified convexity adjustment method to those derivatives. This method is more accurate and can deal with payment-in-arrear, but this method is based on expansion approximation when evaluating the expectation for payment-in-arrear case, so it still can be improved. Pugachevsky (2001) showed another approach, but his method was based on the traditional convexity adjustment (i.e., 2nd order of Taylor expansion), then the accuracy may become a concern. In this chapter we present a new approach to derive the more accurate pricing formulae for those instruments. One main contribution of our method is to use forward measure of cash-flow date instead of fixing date, so we don’t need to input additional correlation. Using change of measure, the present value of any constant maturity security then is given by the sum of the discount expectation values of its cash flows under the appropriate forward measures, and we derive a consistent approximation scheme where the present value is expressed as an asymptotic power series in interest rate volatility, the first few terms in this series reproduce the intuitive valuation results mentioned above, also one can obtain higher accuracy by including subsequent terms. Lastly our final results are expressed without reference to any measure.

This chapter is organized as following. In Section 2 we introduce the methodology and derive the approximate scheme for the expectation of constant maturity yield in terms of an asymptotic power series. In Section 3 we give the valuation formula for swaps, caps and floors, as well as swaptions, we also show the results if we assume LIVBOP rate process rather than zero coupon bond process. Section 4 contains numerical examples for swaps and caps. Summaries are presented in the last section of the chapter.
3.2 Methodology and computation for expectation of constant maturity yield

Recall the definition for vanilla swap, the floating side pays 6M period LIBOR. Thus 6M LIBOR is used for defining the cash flow of the floating side of the plain vanilla swap. Now suppose we replace 6M LIBOR with some other rate, for example, the yield on 10Y Treasury bonds, without altering the payment frequency or anything else. Then we have a constant maturity treasury – a 10Y CMT. In this CMT, the CMT side pays the 10Y Treasury bond yield frequently. Similarly, we can define constant maturity cap (floor, swaption) just as ordinary plain vanilla cap (floor, swaption) but with some index rate replacing LIBOR as the capped (floored, swap) rate.

Then we start from a known fact that under a forward measure, for any traded security the expectation value of its future price equals the forward price. Thus we have

\[ E[P(y)] = P(y_0). \]

Here \( y_0 \) denotes the forward value of \( y \) and \( P(y) \) is the price for that instrument at the corresponding future time.

Assume the yield follows the lognormal process, that is

\[ \frac{dy}{y} = \mu dt + \sigma dZ_t, \]

then

\[ y = F e^{\int_0^T (\mu - \frac{1}{2} \sigma^2) dt + \int_0^T \sigma dZ_t}, \]

where \( F \) is the initial value, hence

\[
E[y^i] = E[F^i e^{\int_0^T (\mu - \frac{1}{2} \sigma^2) dt + i \int_0^T \sigma dZ_t}]
= F^i e^{i \mu T - \frac{1}{2} \sigma^2 T} \cdot e^{\frac{(i-1)}{2} \sigma^2 T}
= (E[y])^i \cdot e^{\frac{(i-1)}{2} \sigma^2 T}.
\]
here \( \hat{\sigma}^2 = \frac{1}{T} \int_0^T \sigma^2 dt \). If we expand

\[
E[y] = y_0(1 + \sum_{n=1}^{\infty} r_n(\hat{\sigma} \sqrt{T})^n) = F e^{\int_0^T \mu dt},
\]

from which we can solve for \( F \), therefore

\[
y = y_0 e^{-\frac{1}{2} \int_0^T \sigma^2 dt + \int_0^T \sigma \sigma \, dz + \int_0^T \sigma \sigma \, dz}, (1 + \sum_{n=1}^{\infty} r_n(\hat{\sigma} \sqrt{T})^n),
\]

and

\[
E[y] = y_0(1 + \sum_{n=1}^{\infty} r_n(\hat{\sigma} \sqrt{T})^n).
\]

By Taylor’s expansion,

\[
P(y) = \sum_{n=0}^{\infty} P_n \cdot (y - y_0)^n,
\]

where \( P_n = \frac{1}{n!} \frac{\partial^n P(y)}{\partial y^n} \bigg|_{y=y_0} \). Taking expectation on both sides and note \( E[P(y)] = P_0 \), so

\[
E[y] = y_0 - \sum_{n=2}^{\infty} \frac{P_n}{P_1} E[(y - y_0)^n]
\]

\[
= y_0 - \sum_{n=2}^{\infty} \frac{P_n}{P_1} y_0^n E[\sum_{i=0}^{n} \left( \begin{array}{c} n \\ i \end{array} \right) e^{-\frac{1}{2} \int_0^T \sigma^2 dt + i \int_0^T \sigma \sigma \, dz + \int_0^T \sigma \sigma \, dz} (1 + \sum_{m=1}^{\infty} r_m(\hat{\sigma} \sqrt{T})^m)^i (-1)^{n-i}]\]

\[
= y_0 - \sum_{n=2}^{\infty} \frac{P_n}{P_1} y_0^n \sum_{i=0}^{n} \left( \begin{array}{c} n \\ i \end{array} \right) e^{\frac{i(i-1)}{2} \hat{\sigma}^2 T} (1 + \sum_{m=1}^{\infty} r_m(\hat{\sigma} \sqrt{T})^m)^i (-1)^{n-i}.
\]

From the above two expressions for \( E[y] \), comparing the coefficients of the same order of \( \hat{\sigma} \sqrt{T} \), we can solve for \( r_n \).

For \( n = 1 \),

\[
y_0 r_1 = -\sum_{k=2}^{\infty} \frac{P_k}{P_1} y_0^k \left[ \sum_{i=0}^{k-1} \left( \begin{array}{c} k-1 \\ i \end{array} \right) (-1)^{k-i} \right] = 0,
\]

so \( r_1 = 0 \).

For \( n = 2 \),

\[
y_0 r_2 = -\sum_{k=2}^{\infty} \frac{P_k}{P_1} y_0^k \left[ \sum_{i=0}^{k} \left( \begin{array}{c} k \\ i \end{array} \right) (-1)^{k-i} (ir_2 + \frac{i(i-1)}{2}) \right]
\]

11
\[
\begin{align*}
\frac{P_2}{P_1} y_0^2 &= - \frac{P_2}{P_1} y_0^2 \sum_{k=3}^{\infty} \frac{P_2}{P_1} y_0^k k \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^{k-i} \\
&\quad - \frac{P_2}{P_1} k(k-1) \frac{1}{2} y_0^k \sum_{i=0}^{k-2} \binom{k-2}{i} (-1)^{k-i} \\
&= \frac{P_2}{P_1} y_0^2, \\
\text{so} \quad r_2 &= \frac{P_2}{P_1} y_0.
\end{align*}
\]

Similarly, we obtain
\[
\begin{align*}
r_3 &= 0, \\
r_4 &= - \frac{1}{P_1} \left[ \left( \frac{1}{2} + 2r_2 + r_2^2 \right) P_2 y_0 + (3 + 3r_2) P_3 y_0^2 + 3 P_4 y_0^3 \right].
\end{align*}
\]

Therefore
\[
E^Q[y] = y_0 \left( 1 + r_2 \sigma^2 T + r_4 \sigma^4 T^2 \right) + o(\sigma^6 T^3). \tag{3.1}
\]

**Proposition 3.2.1** All coefficients of odd order vanish.

Proof: (By induction) on the previous discussion, we have \( r_1 = 0. \)

Suppose \( r_{2k-1} = 0 \) for all \( k \leq m, \)

then for \( k = m + 1, \) note that
\[
\begin{align*}
y_0 r_{2m+1} &= - \sum_{k=2}^{\infty} \frac{P_k}{P_1} y_0^k \sum_{i=0}^{k} \binom{k}{i} (-1)^{k-i} r_{2m+1} \\
&\quad - \sum_{k=2}^{\infty} \frac{P_k}{P_1} y_0^k r_{2m+1} \sum_{i=0}^{k-1} \binom{k-1}{i} (-1)^{k-i} \\
&= 0.
\end{align*}
\]

Therefore \( r_{2m+1} = 0, \) the proof is complete.
3.3 Valuation of constant maturity instruments

In this section we provide analytical pricing formulae for various constant maturity instruments such as constant maturity swap, cap and swaption.

3.3.1 Valuation of swap

For the valuation of the floating leg of the swap, if the cash is paid in arrear, we know the present value is given simply as the sum of the discounted expectation values over cash-flow days.

Assume the current day is time 0, and the cash flow dates are $T_1, T_2, \ldots, T_N$, then

$$\text{Swap} = Pr \sum_{n=1}^{N} \Delta_n P_0^{T_n} E^{Q^n}[y],$$

(3.2)

where $\Delta_n = T_n - T_{n-1}$, $y$ is the index rate and $Pr$ is the principal.

Now let $t_i$ denote the fixing date, $t_i < T_i$, and we assume $y$ is lognormal, next we’ll give the formula for pricing this swap.

Since $y$ is lognormal, the following is the relationship between two Brownian motion $W(t, T_n)$ and $Z(t)$ which are under forward measure and risk neutral measure respectively.

$$dW(t, T_n) = dZ(t) - \sigma(t, T_n) dt$$

$$= dZ(t) - \sigma(t, t_n) dt + (\sigma(t, t_n) - \sigma(t, T_n)) dt$$

$$= dW(t, t_n) + \dot{\sigma}(t, t_n, T_n) dt,$$

where $\sigma(t, T_n)$ is the volatility of zero-coupon bond $P(t, T_n)$ and

$$\dot{\sigma}(t, t_n, T_n) = \sigma(t, t_n) - \sigma(t, T_n).$$
We define
\[ dB_n(t) = \frac{1}{2} \| \dot{\sigma}(t, t_n, T_n) \|^2 dt - \dot{\sigma}(t, t_n, T_n) \cdot dW(t, t_n). \]

Note that both \( W(t, t_n) \) and \( \dot{\sigma}(t, t_n, T_n) \) are vectors, then by Girsanov’s Theorem,
\[
E^{Q^{T_n}}[y] = E^{Q^{T_n}}[y \cdot e^{B_n(T_n)}] \\
= E^{Q^{T_n}}[e^{\int_0^{T_n} dB_n(t)} \cdot y] \\
= E^{Q^{T_n}}[y \cdot e^{\int_0^{T_n} dB_n(t)} \cdot e^{\int_0^{T_n} dB_n(t)}] \\
= E^{Q^{T_n}}[y \cdot e^{\int_0^{T_n} dB_n(t)}].
\]

The last equality is because of the independence and also \( e^{\int_0^{T_n} dB_n(t)} \) is a martingale.

If we assume \( y \) is lognormal under \( Q^{T_n} \),
\[
\frac{dy}{y} = \mu dt + \sigma_y \cdot dW(t, t_n)
\]  \hspace{1cm} (3.3)

where \( W(t, t_n) \) is a Brownian motion under the forward measure with delivery at \( t_n \).

From (3.3),
\[
y = E^{Q^{T_n}}[y] e^{\int_0^{T_n} \sigma_y \cdot dW(t, t_n) - \frac{1}{2} \int_0^{T_n} \| \sigma_y \|^2 dt}. 
\]

Hence
\[
E^{Q^{T_n}}[y \cdot e^{\int_0^{T_n} dB_n(t)}] = E^{Q^{T_n}}[y] E^{Q^{T_n}}[e^{\int_0^{T_n} (\sigma_y - \dot{\sigma}) \cdot dW(t, t_n) - \frac{1}{2} \int_0^{T_n} \| \sigma_y \|^2 + |\dot{\sigma}|^2 dt}] \\
= E^{Q^{T_n}}[y] e^{-\int_0^{T_n} \dot{\sigma} \cdot \sigma_y dt} \\
= E^{Q^{T_n}}[y] e^{-\int_0^{T_n} \rho |\dot{\sigma}| \| \sigma_y \| dt}, 
\]

where \( \rho \) is the correlation between \( y \) and forward rate over \([t_n, T_n]\). Thus
\[
Swap = P \sum_{n=1}^{N} \Delta_n p^{T_n} E^{Q^{T_n}}[y] e^{-\int_0^{T_n} \rho |\dot{\sigma}| \| \sigma_y \| dt}, 
\]

then by equation (3.1), the present value is obtained.

**Remark:**
If we denote the forward term rate between \( t_n \) and \( T_n \) to be \( L_n \), it can be shown that

\[
\frac{dL_n(t)}{\Delta T} = \frac{1}{P_t} \left[ \frac{P_t^n}{P_t} \sigma(t, t_n) - \sigma(t, T_n) \right] \cdot dW(t, T_n)
\]

\[
= L_n(t) \cdot \frac{1}{\Delta T} \left[ \frac{P_t^n}{P_t} - 1 + 1 \right] \left[ \sigma(t, t_n) - \sigma(t, T_n) \right] \cdot dW(t, T_n)
\]

\[
\approx L_n(t) \left( 1 + \frac{1}{\Delta T L_n(0)} \right) \left[ \sigma(t, t_n) - \sigma(t, T_n) \right] \cdot dW(t, T_n)
\]

\[
= L_n(t) \sigma_{L,n}(t) \cdot dW(t, T_n),
\]

where \( L_n(0) \) is the forward value of \( L_n \). Thus

\[
\sigma(t, t_n, T_n) = \sigma(t, t_n) - \sigma(t, T_n) = \frac{\sigma_{L,n}(t)}{\omega(L_n(0))},
\]

where \( \omega(L_n(0)) = 1 + \frac{1}{\Delta T L_n(0)} \). So in terms of \( \sigma_{L,n} \), we have

\[
E^{Q^n} [y \cdot e^{\int_{0}^{t_n} dB_n(t)}] = E^{Q^n} [y] e^{-\int_{0}^{t_n} \sigma_y dt}
\]

\[
= E^{Q^n} [y] e^{-\frac{1}{\omega(L_n(0))} \int_{0}^{t_n} \sigma_{L,n} \sigma_y dt}
\]

\[
= E^{Q^n} [y] e^{-\frac{1}{\omega(L_n(0))} \int_{0}^{t_n} \rho \sigma_{L,n} \sigma_y dt}.
\]

### 3.3.2 Valuation of cap and floor

A constant maturity cap works just like an ordinary cap except that the capped rate is some index rate, thus the value of a constant maturity cap is given by the sum of the values of individual caplets and we now proceed to the valuation of caplets, and floor may be valued using the usual parity relation.

The present value of the constant maturity cap is given by

\[
Cap = Pr \sum_{n=1}^{N} \Delta_n P_0^{T_n} E^{Q^n} [\max(y - C, 0)],
\]

(3.4)
where $C$ is the strike price.

By Black-Scholes formula and result of previous section, we can derive that

$$E^{Q^{T_n}}[\max(y - C, 0)] = E^{Q^{T_n}}[y]N(d_1) - CN(d_2)$$

$$= E^{Q^{T_n}}[y]e^{-\int_0^{T_n} \sigma_y^2 dt}N(d_1) - C \cdot N(d_2),$$

where

$$d_1 = \frac{\ln \left( \frac{E^{Q^{T_n}}[y]}{C} \right)}{\sqrt{\int_0^{T_n} \| \sigma_y(t) \|^2 dt}} + \frac{1}{2} \int_0^{T_n} \| \sigma_y(t) \|^2 dt - \int_0^{T_n} \hat{\sigma}(t, t_n, T_n) \cdot \sigma_y(t) dt \sqrt{\int_0^{T_n} \| \sigma_y(t) \|^2 dt},$$

$$d_2 = d_1 - \int_0^{T_n} \| \sigma_y(t) \|^2 dt,$$

and $N(\cdot)$ is the cumulative normal distribution function.

Since the value of cap is the sum of all those caplets, finally the price is obtained.

**Remark:**

If we use the forward term rate volatility instead,

$$E^{Q^{T_n}}[\max(y - C, 0)] = E^{Q^{T_n}}[\max(y - C, 0)e^{\int_0^{T_n} dB_n(t)}]$$

$$= E^{Q^{T_n}}[y]e^{-\int_0^{T_n} \sigma_{L,n} \cdot \sigma_y dt}N(d'_1) - C \cdot N(d'_2)$$

$$= E^{Q^{T_n}}[y]e^{-\int_0^{T_n} \rho \| \sigma_{L,n} \| \| \sigma_y \| dt}N(d'_1) - C \cdot N(d'_2),$$

where

$$d'_1 = \frac{\ln \left( \frac{E^{Q^{T_n}}[y]}{C} \right)}{\sqrt{\int_0^{T_n} \| \sigma_y(t) \|^2 dt}} + \frac{1}{2} \int_0^{T_n} \| \sigma_y(t) \|^2 dt - \frac{1}{\sqrt{\int_0^{T_n} \| \sigma_y(t) \|^2 dt}} \int_0^{T_n} \rho \| \sigma_{L,n}(t) \| \| \sigma_y(t) \| dt,$$

$$d'_2 = d'_1 - \int_0^{T_n} \| \sigma_y(t) \|^2 dt.$$
3.3.3 Valuation of swaption

In this subsection we provide the formula of swaptions. Suppose the maturity of this swaption is $T$, cash-flow dates are $T_1, T_2, \ldots, T_n$, where $T < T_1$.

Firstly we define the forward swap rate as following

$$R(t) = \frac{P(t, T) - P(t, T_n)}{B^S(t)},$$

where

$$B^S(t) = \sum_{j=1}^{n} \Delta_j P(t, T_j).$$

Now the payoff is

$$B^S(T)[\max(R(T) - K, 0)],$$

where $K$ is the strike.

Then the value of cash flow received at time $T_j$ is

$$V_j = \Delta_j P(t, T_j)(E^{Q^T_j}[R(T)]N(d_{j,1}) - KN(d_{j,2})), $$

where $R$ is lognormal distributed, $\sigma$ is its volatility and

$$d_{j,1} = \frac{\ln(E^{Q^T_j}[R]/K) + \sigma^2T/2}{\sigma\sqrt{T}},$$

$$d_{j,2} = d_{j,1} - \sigma\sqrt{T}.$$

Combined with equation (3.1), we can get the value of $V_j$. Hence the total value of the swaption is

$$V = \sum_{j=1}^{n} V_j.$$

3.4 Numerical results

In this section we present two numerical examples for constant maturity swaps and caps to show the accuracy of our method.
3.4.1 Numerical example for swaps

We present results for CMT basis swaps, i.e., CMT plus a spread against 6M LIBOR. Both sides of the swap have semiannual fixings and are fixed in advance, paid in arrears. We assume a linear swap yield curve starting at 5% at the short end and rising to a level of 7% at the long (30Y) end. The Treasury curve lies below the swap curve with a constant spread of 25 basis points on rates. And for the volatility corrections we assume

<table>
<thead>
<tr>
<th></th>
<th>Swap Maturity</th>
<th></th>
<th></th>
<th></th>
<th>Index Tenor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2Y</td>
<td>5Y</td>
<td>10Y</td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Adjustment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(0'th order)</td>
<td>-0.1687</td>
<td>-0.1688</td>
<td>-0.1689</td>
<td></td>
<td>2Y</td>
</tr>
<tr>
<td></td>
<td>-0.3613</td>
<td>-0.3611</td>
<td>-0.3608</td>
<td></td>
<td>5Y</td>
</tr>
<tr>
<td></td>
<td>-0.6556</td>
<td>-0.6541</td>
<td>-0.6518</td>
<td></td>
<td>10Y</td>
</tr>
<tr>
<td>Convexity Adjustment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2'nd order)</td>
<td>-0.1786</td>
<td>-0.2113</td>
<td>-0.3107</td>
<td></td>
<td>2Y</td>
</tr>
<tr>
<td></td>
<td>-0.3719</td>
<td>-0.4067</td>
<td>-0.5121</td>
<td></td>
<td>5Y</td>
</tr>
<tr>
<td></td>
<td>-0.6674</td>
<td>-0.7045</td>
<td>-0.8179</td>
<td></td>
<td>10Y</td>
</tr>
<tr>
<td>4'th order</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.1787</td>
<td>-0.2128</td>
<td>-0.3200</td>
<td></td>
<td>2Y</td>
</tr>
<tr>
<td></td>
<td>-0.3721</td>
<td>-0.4083</td>
<td>-0.5219</td>
<td></td>
<td>5Y</td>
</tr>
<tr>
<td></td>
<td>-0.6675</td>
<td>-0.7063</td>
<td>-0.8286</td>
<td></td>
<td>10Y</td>
</tr>
<tr>
<td>4'th order w/arrears adjustment</td>
<td>-0.1777</td>
<td>-0.2105</td>
<td>-0.3154</td>
<td>2Y</td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.3710</td>
<td>-0.4059</td>
<td>-0.5171</td>
<td></td>
<td>5Y</td>
</tr>
<tr>
<td></td>
<td>-0.6664</td>
<td>-0.7037</td>
<td>-0.8236</td>
<td></td>
<td>10Y</td>
</tr>
</tbody>
</table>

Table 3.1: Spreads for CMT basis swaps

a volatility for the Treasury yields of 15%. For the adjustment and payment in arrears, and we further assume a LIBOR volatility of 15% and a correlation between 6M LIBOR and CMT of 0.3. The results are presented as the spread on the CMT leg in units of percentage points, i.e., the CMT side pays CMT plus the spread and receives 6M LIBOR.
From these data, we'll see all the spreads are negative, which are determined by the yield curve. We also notice that as expected the effect of the volatility corrections increases with swap maturity and with index tenor. The fourth order approximation is very accurate. In fact, a higher order almost give the same results.

And we can find that in most cases convexity adjustment (2'nd order) is sufficient.

### 3.4.2 Numerical example for caps

We present results for CMT caps with semiannual fixings and payment in arrears. We assume a linear structure of swap yield starting at 5% at the short (O/N) end and rising

<table>
<thead>
<tr>
<th>Cap Tenor</th>
<th>2Y</th>
<th>5Y</th>
<th>10Y</th>
<th>Index Tenor</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Adjustment (0'th order)</td>
<td>0.0343</td>
<td>0.5170</td>
<td>2.6362</td>
<td>2Y</td>
</tr>
<tr>
<td></td>
<td>0.0532</td>
<td>0.6551</td>
<td>3.0563</td>
<td>5Y</td>
</tr>
<tr>
<td></td>
<td>0.0981</td>
<td>0.9165</td>
<td>3.7782</td>
<td>10Y</td>
</tr>
<tr>
<td>Convexity Adjustment (2'nd order)</td>
<td>0.0354</td>
<td>0.5570</td>
<td>3.0596</td>
<td>2Y</td>
</tr>
<tr>
<td></td>
<td>0.0549</td>
<td>0.7051</td>
<td>3.5447</td>
<td>5Y</td>
</tr>
<tr>
<td></td>
<td>0.1012</td>
<td>0.9846</td>
<td>4.3762</td>
<td>10Y</td>
</tr>
<tr>
<td>4'th order</td>
<td>0.0354</td>
<td>0.5586</td>
<td>3.0918</td>
<td>2Y</td>
</tr>
<tr>
<td></td>
<td>0.0550</td>
<td>0.7071</td>
<td>3.5812</td>
<td>5Y</td>
</tr>
<tr>
<td></td>
<td>0.1012</td>
<td>0.9873</td>
<td>4.4196</td>
<td>10Y</td>
</tr>
<tr>
<td>4'th order w/arrears adjustment</td>
<td>0.0353</td>
<td>0.5564</td>
<td>3.0773</td>
<td>2Y</td>
</tr>
<tr>
<td></td>
<td>0.0548</td>
<td>0.7044</td>
<td>3.5650</td>
<td>5Y</td>
</tr>
<tr>
<td></td>
<td>0.1009</td>
<td>0.9839</td>
<td>4.4007</td>
<td>10Y</td>
</tr>
</tbody>
</table>

Table 3.2: Premiums for CMT caps
to a level of 7% at the long(30Y) end. The Treasury curve lies below the swap curve with a constant spread of 25 basis points on rates. For the volatility corrections we assume a volatility for the Treasury yields of 15%. For the adjustment and payment in arrears, and we further assume a LIBOR volatility of 15% and a correlation between 6M LIBOR and CMT of 0.3. The results are presented as the premium per 100 notional of a cap with strike at 7%.

One may notice that, as expected, the effect of the volatility corrections increases with cap tenor and with index tenor. Just like swaps, it appears that convexity adjustment (2’nd order) will be sufficient in most cases.

3.5 Summary

We have shown how to value constant maturity swap (CMT and CMS), caps, floors and swaptions and how to take into account payment-in-arrears. Our results imply that the standard "convexity adjustment" is simply the first non-trivial term in a systematic expansion in volatility.

Moreover, the higher order terms are straightforward to calculate and in principle an arbitrary accuracy can be obtained by including a sufficiently high number of terms. In practice, however, as our numerical studies indicate, there does not (at present) seem to be much to gain by including more that the first few terms.
Chapter 4

Calibration of CEV Market Model

4.1 Introduction

The standard LIBOR market model is based on the assumption of lognormal process for the observed forward rates. This model has two main advantages. First, it can price caplets and swaptions in closed-form, which enables efficient calibration of the model. Second, it is a multi-factor model and thus has enough degrees of freedom to calibrate simultaneously the prices of benchmark instruments such as caplets and swaptions. Because of these features the standard LIBOR market model is now playing a predominant role in the interest rate derivative markets of various currencies. Calibration of the market model has been one of the focuses in recent research.

Although the standard market model enjoys great popularity, it still has some limitations. It’s widely agreed that the standard model can not accommodated the effect of volatility skew, and in practice the model is calibrated only to the at-the-money (ATM) caplets/floors and ATM swaptions. The volatility skew means the pattern of decreasing implied Black volatilities of caplet and swaption prices for increasing strike prices, which indicates a fat tail of empirical forward rate distributions relative to lognormal distri-
butions. CEV model can essentially capture the volatility skew, and, as an extension to the standard market model, it retains the analytical tractability and renders closed-form formulae to caplet and swaption prices.

In this chapter we will generalize the methodology of Wu for standard market model to the calibration of the CEV model. Here we utilize some theorems and theoretical results from Wu (2001), and we use numerical examples to illustrate how to use those theoretical results to calibrate the entire CEV model. The emphasis of this chapter is the application of some known results, that is, how to solve a real calibration problem under the CEV framework setting. The elasticity and implied volatility have been obtained and we also plot some figures to present the calibration results using different number of factors.

This chapter is organized as following. In section 2 we introduce the background of the LIBOR market model and CEV market model, and list the mathematical formulae for the calibration. In section 3 we review the methodology of the calibration of CEV model, and cite some important theories to be used from Wu's calibration work (2001). In section 4 we introduce the practical problem and present computational results. Finally in section 5 we conclude.

4.2 Problem formulation

4.2.1 The term structure of standard market model

Let's start from the term structure of the standard market model. The market model is based on the lognormal assumption of zero-coupon bond dynamics. Let $P(t, T)$ be the zero-coupon Treasury bond maturing at $T > t$ with par value $\$1$. Under risk-neutral
measure $P(t, T)$ is assumed to evolve according to

$$dP(t, T) = P(t, T)(r_t dt + \sigma(t, T) \cdot dZ_t),$$  \hspace{1cm} (4.1)

where $r_t$ is the risk-free rate, $\sigma(t, T)$ is the volatility vector of $P(t, T)$, and $Z_t$ is vector of independent Wiener process under the risk-neutral measure, which we denote by $\mathbb{Q}$, and "\cdot" means inner product.

With the Ito's lemma we can derive the equation for $\ln P(t, T)$,

$$d \ln P(t, T) = (r_t - \frac{\sigma^2(t, T)}{2}) dt + \sigma(t, T) \cdot dZ_t.$$  

Differentiate the above equation with respect to maturity $T$ (and multiplying by "\cdot") we arrive at equation

$$df(t, T) = \sigma_T \cdot \sigma dt - \sigma_T \cdot dZ_t$$  \hspace{1cm} (4.2)

for the instantaneous forward rate maturing at $T$, where by definition,

$$f(t, T) = -\frac{\partial \ln P(t, T)}{\partial T},$$

and

$$\sigma(t, T) = \int_t^T \sigma_T(t, s) ds.$$  \hspace{1cm} (4.3)

(4.2) is the well-known Heath-Jarrow-Morton equation (Heath et al. (1992)), which states that, under the risk neutral measure, the drift term of the forward rate is a function of its volatility. The HJM model is an all-embracing model in the sense that, by specifying specific volatility function for the forward rate, one can deduce various interest-rate models, including Ho-Lee (1977), Cox-Rose-Ingersoll (1985), Hull-White (1990) and other popular ones.

Although theoretically appealing, the HJM model is not very convenient for applications. There are two main reasons. First, it takes the instantaneous forward rate,
a non-observable quantity, as the state variable. This increases the difficulty for yield curve building and model calibration. Second, under the HJM model the pricing of most derivative securities, including the benchmark securities (caps, floors and swaptions), has to resort to Monte Carlo simulations. In the market place, meanwhile, the pricing of the benchmark securities has been done with the Black’s formula (Black, 1976), a variation of the celebrated Black-Scholes-Merton formula (Black and Scholes (1973) and Merton (1973)) for equity options. For some time such market practice had been considered inconsistent with the HJM theory.

The works by Brace-Gatarek-Musiela (1997), Jamshidian (1997), and Miltersen, Sandmann & Sondermann (1997) reconciled practice with theory and shifted the paradigm of the interest-rate research. These researchers chose to model the forward term rates directly, while managing to maintain consistence with the dominating HJM theory. Let \( f_j(t) = f(t; T_j, T_{j+1}) \) be the arbitrage-free forward lending rate seen at time \( t \) for the period \( (T_j, T_{j+1}) \), which is an observable and tradable quantity (by forward-rate agreement, at least) in the interest-rate markets. The forward term rate relates to the zero-coupon bond price by

\[
f_j(t) = \frac{1}{T_{j+1} - T_j} \left( \frac{P(t, T_j)}{P(t, T_{j+1})} - 1 \right).
\]

As a function of two zero-coupon bonds, the dynamics of the forward term rate is determined by that of the zero-coupon bonds. Using Ito’s lemma we can derive

\[
df_j(t) = f_j(t) \gamma_j(t) \cdot [dZ_t - \sigma(t, T_{j+1})dt], \quad 1 \leq j \leq N, \tag{4.4}
\]

where \( \gamma_j(t) \) is a function of zero-coupon bond volatilities:

\[
\gamma_j(t) = \frac{1 + \Delta T_j f_j(t)}{\Delta T_j f_j(t)} [\sigma(t, T_j) - \sigma(t, T_{j+1})], \tag{4.5}
\]

which is intuitively regarded as the volatility vector for \( f_j(t) \). Under the risk neutral measure, therefore, the volatilities of the forward term rates are determined by those of
the zero-coupon bonds. Such relation can be viewed conversely. In fact, we can express the volatilities of the zero-coupon bonds in terms of the volatilities of the forward term rates:

\[
\sigma(t, T_{j+1}) = - \sum_{k=1}^{j} \frac{\Delta T_k f_k(t)}{1 + \Delta T_k f_k(t)} \gamma_k(t) + \sigma(t, T_1),
\]  

(4.6)

where \(\Delta T_j = T_{j+1} - T_j\). If one instead begins with prescribing the volatilities of the forward rates, the volatilities of the zero-coupon bonds will follow from (4.6). Given \(\gamma_j(t)\) satisfying usual regularity conditions, Brace et al. (1997) proved that \(f_j(t)\) didn’t blow up and thus put the market model on a solid mathematical foundation. They also justified that one could put \(\sigma(t, T_1) = 0\) for \(t \leq T_1\). Equations (4.4) and (4.5) constitute the so-called market model of interest rates. Roughly speaking, the stochastic evolution of the \(N\) forward rates is governed by the quantities of covariance defined by

\[
\text{Cov}^j_{ik} = \int_{T_{i-1}}^{T_i} \gamma_j(t) \cdot \gamma_k(t) dt, \quad i \leq j, k \leq N, \quad 1 \leq i \leq N.
\]  

(4.7)

Note that \(\text{Cov}^j_{ik} = 0\) for either \(j < i\) or \(k < i\) since either \(f_j\) or \(f_k\) has been reset by the time \(T_i\) (either \(\gamma_j = 0\) or \(\gamma_k = 0\)). The market model has the capacity to build in certain correlation structure between the forward rates, which has been studied extensively.

Under the market model, the use of the Black’s formula for caplets is nicely justified. Swaptions, meanwhile, can also be priced with approximate formula (see for instance Andersen and Andreasen (1998) and Sidenius (1999)) which is accurate within the bid-ask spread. However, once the implied volatilities of the caplets (of the same maturity ) are different, the market model can not simultaneously regenerate their prices. In other words, the standard market model can not be simultaneously calibrated to caplets of the same maturity but different implied Black volatilities. The reason is that in order to keep the Black’s formula valid forward rate volatilities are only taken to be functions of calendar time and forward time only, without strike. In reality the curves formed by implied Black volatilities most often take the shape of a smile or a skewed
Figure 4.1: Implied volatility of 6M maturity caplet

Figure 4.2: Implied volatility of 1Y maturity caplet

Figure 4.3: Implied volatility of 2Y maturity caplet

Figure 4.4: Implied volatility of 3Y maturity caplet

Figure 4.5: Implied volatility of 4Y maturity caplet

Figure 4.6: Implied volatility of 5Y maturity caplet

Figure 4.7: Implied volatility of 7Y maturity caplet
smile, and such phenomenon is often seen in the USD and EURO markets, especially is pronounced in the JPY market. Figures 4.1 to 4.7 show the implied volatility curves of USD caplets of various maturities, for the date July 04, 2002. One can see that smiles and downward skews coexist across maturities. For aggregated portfolio risk management, it is desirable to enlarge the capacity of the market model: to make it capable to accommodate volatility smiles or skews in the sense of regenerating market prices of the benchmark instruments.

### 4.2.2 CEV formulation

The CEV market model (Andersen and Andreasen(2000)) was based on the assumption of CEV processes for forward term rates. Let \( f_j(t) = f(t; T_j, T_{j+1}) \) be the arbitrage-free forward lending rate seen at time \( t \) for the period \( (T_j, T_{j+1}) \), then \( f_j(t) \) is assumed to follow a CEV process

\[
df_j(t) = f_j^{\alpha_j}(t) \gamma_j(t) \cdot [\sigma_{j+1}(t)dt + dZ(t)],
\]

where \( \alpha_j \) is a positive constant, \( Z(t) \) is \( n \)-dimensional independent Brownian motions for some properly chosen number \( n \), \( \gamma_j(t) \) is the vector of the instantaneous volatility coefficient, and \( \sigma_{j+1}(t) \) is the vector of instantaneous volatility coefficient of zero-coupon bond of maturity \( T_{j+1} \). If we define the percentage volatility

\[
\tilde{\gamma}_j = \frac{f_j^{\alpha_j}(t) \gamma_j(t)}{f_j(t)} = (f_j(t))^{\alpha_j-1} \gamma_j(t),
\]

then

\[
\frac{\partial \tilde{\gamma}_j}{\partial f_j} = \frac{\tilde{\gamma}_j}{f_j} = \alpha_j - 1 = \text{constant},
\]

which is the elasticity.

Consider a collection of \( N \) forward rates, \( f_j, j = 1, 2, \ldots, N \). As in the HJM model
(1992), the drifts of forward term rates in the market model are completely determined by their volatilities. The no-arbitrage condition (Brace et al., 1997) gives rise to their relation

\[ \sigma_j(t) = \sum_{k=1}^{J} \frac{\Delta T_{j-k} f_{j-k}^{\alpha_j-k}(t)}{1 + \Delta T_{j-k} f_{j-k}^{\alpha_j-k}(t)} \gamma_{j-k}(t), \]

where \( \Delta T_j = T_{j+1} - T_j \) and \( \gamma_j(t) = 0 \) for \( t \geq T_j \). As a convention we label today by \( t = T_0 = 0 \). For mathematical properties of the CEV model we refer to Andersen and Andreasen (2000).

The stochastic evolution of the \( N \) forward rates is fully described by the quantities of covariance defined by

\[ COV_{jk}^i = \int_{T_{i-1}}^{T_i} \gamma_j(t) \cdot \gamma_k(t) dt, \quad 1 \leq i \leq N. \]

Note that \( COV_{jk}^i = 0 \) for either \( j < i \) or \( k < i \) since either \( f_j \) or \( f_k \) has been reset by the time \( T_i \). The corresponding correlations are

\[ C_{jk}^i = \frac{COV_{jk}^i}{\sqrt{COV_{jj}^i \cdot COV_{kk}^i}}, \quad 1 \leq i \leq N. \]

For fixed \( i \), \( \{C_{jk}^i\} \) constitute an \( (N - i + 1) \) by \( (N - i + 1) \) non-negative symmetric matrix:

\[
C^i = \begin{pmatrix}
C_{i,i}^i & C_{i,i+1}^i & \cdots & C_{i,N}^i \\
C_{i+1,i}^i & C_{i+1,i+1}^i & \cdots & C_{i+1,N}^i \\
\vdots & \vdots & \ddots & \vdots \\
C_{N,i}^i & C_{N,i+1}^i & \cdots & C_{N,N}^i
\end{pmatrix}, \quad i = 1, 2, \ldots, N.
\]

We now introduce the pricing of swaption. The price formula for caplet follows as a special case. A swaption is an option on swap rate. Denote an annuity

\[ B^S(t) = \sum_{j=m}^{n-1} \Delta T_j P(t, T_{j+1}), \]
where \( P(t, T_{j+1}) \) is the time \( t \) price of the zero-coupon bond with maturity \( T_{j+1} \) and face value $1. The fair swap rate for the period \( (T_m, T_n) \) seen at time \( t \) is defined by

\[
R_{m,n}(t) = \frac{P(t, T_m) - P(t, T_n)}{\sum_{j=m}^{n-1} \Delta T_j P(t, T_{j+1})}.
\]

The swap rate is the fixed rate with which two parties will agree to swap fixed payments for floating payments (indexed to LIBOR) for any notional amount, at times \( T_j, j = m + 1, m + 2, \ldots, n \). The forward rates related to the zero-coupon bonds by

\[
f_j(t) = \frac{1}{\Delta T_j} \left( \frac{P(t, T_j)}{P(t, T_{j+1})} - 1 \right).
\]

By Ito’s lemma, the dynamics of \( R_{m,n}(t) \) is

\[
dR_{m,n}(t) = \sum_{j=m}^{n-1} \frac{\partial R_{m,n}(t)}{\partial f_j(t)} f_j^{\alpha_j}(t) \gamma_j(t) \cdot dW^S(t),
\]

where \( W^S(t) \) is \( n \)-dimensional independent Brownian motions under the forward swap measure induced by choosing \( B^S(t) \) as the numeraire, and

\[
\frac{\partial R_{m,n}(t)}{\partial f_j(t)} = \frac{\Delta T_j R_{m,n}(t)}{1 + \Delta T_j f_j(t)} \left[ \frac{P(t, T_n)}{P(t, T_m) - P(t, T_n)} + \sum_{k=j}^{n-1} \frac{\Delta T_k P(t, T_{k+1})}{B^S(t)} \right].
\]

The swap rate process (4.8) is apparently not a CEV process, yet following Anderson and Andreasen’s approximation, it can be approximated by “frozen coefficient”:

\[
dR_{m,n}(u) = R_{m,n}^{\alpha_{m,n}}(u) \sum_{j=m}^{n-1} \omega_j(t) \gamma_j(u) \cdot dW^S(u), \quad t \leq u < T_m,
\]

where

\[
\omega_j(t) = \frac{\partial R_{m,n}(t)}{\partial f_j(t)} \frac{f_j^{\alpha_j}(t)}{R_{m,n}^{\alpha_{m,n}}(t)},
\]

and \( \alpha_{m,n} \) is the power to be determined by least-squared fitting to the volatility skew of swaptions using formulae developed below. As we shall see that (4.9) leads to a closed-form solution for European swaption.
Theorem 4.2.1 (Andersen and Andreasen, 1998) Consider a European payer swaption on swap rate $R_{m,n}$ with strike rate $K$. Assume that the swap rate dynamics are given by the CEV specification (4.9). Define

$$
d = \frac{K^{2(1-\alpha_{m,n})}}{(1-\alpha_{m,n})^2 \xi_{m,n}^2(t)}, \quad b = \frac{1}{1-\alpha_{m,n}}, \quad f = \frac{R^{2(1-\alpha_{m,n})}}{(1-\alpha_{m,n})^2 \xi_{m,n}^2(t)},
$$

$$
g_{\pm} = \frac{\ln \frac{R_{m,n}(t)}{K} \pm \frac{1}{2} \xi_{m,n}^2(t)}{\xi_{m,n}(t)},
$$

then the swaption price at $t$ is given by

1. For $0 < \alpha_{m,n} < 1$ and an absorbing boundary at the level $R_{m,n} = 0$,

$$
PS(t, T_m, T_n) = B^S(t)[R_{m,n}(t)(1 - \chi^2(d, b + 2, f)) - K \chi^2(f, b, d)],
$$

(4.11)

where $\chi^2(\cdot, \cdot, \cdot)$ is cumulative distribution function for a non-central $\chi^2$-distributed variable.

2. For $\alpha_{m,n} = 1$,

$$
PS(t, T_m, T_n) = B^S(t)[R_{m,n}(t)N(g_+) - KN(g_-)],
$$

(4.12)

where $N(\cdot)$ is the normal accumulative function.

3. For $\alpha_{m,n} > 1$,

$$
PS(t, T_m, T_n) = B^S(t)[R_{m,n}(t)(1 - \chi^2(f, -b, d)) - K \chi^2(d, 2 - b, f)].
$$

(4.13)

Remarks: 1) The non-central $\chi^2$-distribution function can be evaluated numerically with a procedure developed by Ding (1992). 2) A caplet is a special case of swaption, corresponding to $n = m + 1$. Note that forward rates are only special cases of swap rates, i.e., $R_{m,m+1} = f_m$, we will only mention swap rates thereafter.

The first step of calibration is to fit the $\alpha_{m,n}$ and $\xi_{m,n}$ to the swaption (including
caplet) prices by least-squares fitting. In particular, if there is only one input option
price for a swap rate, we take the corresponding $\alpha_{m,n} = 1$ and solve for $\xi_{m,n}$ by root
finding. This process will translate all option prices on $R_{m,n}$ into the pairs of $\alpha_{m,n}$ and
$\xi_{m,n}$.

In this chapter, we define our calibration problem as following: given $\{\alpha_{m,n}\}$,
$\{\xi_{m,n}\}$ and $\{C^i\}$, determine the implied volatility functions $\gamma_j(t), j = 1, \ldots, N$.

We will take the non-parametric approach, looking for the volatilities in the form
of piece-wise constant functions in $t$:

$$
\gamma_j(t) = \gamma_j^i = s_j^i(a_{j,1}^i, a_{j,2}^i, \ldots, a_{j,n}^i) \equiv s_j^i a_j^i, \quad \text{for} \quad T_{i-1} \leq t \leq T_i, \ i \leq j
$$

with

$$
\|s_j^i\|_2 = \|\gamma_j^i\|_2 = 1. \quad (4.14)
$$

The total number of unknowns is proportional to $n \times N^2$, which in practice can be in
the magnitude of hundreds and far bigger than the number of input prices and elements
of correlation matrices. Hence, we are facing a middle- to large-scale under-determined
problems. Luckily, the determination of $\{s_j^i\}$ and $\{a_j^i\}$ are decoupled in the sense that,
while the former depends on both $\{C^i\}$ and $\{\xi_{m,n}\}$, the latter depends on $\{C^i\}$ only.

Suppose the rank of $\{C^i\}$ is less than or equal to $n$. Perform eigenvalue decom-
position on $C^i$:

$$
C^i = U\Lambda U^T,
$$

where $\Lambda$ is an $n$ by $n$ diagonal matrix with non-negative diagonal elements, and defined
$a_j^i$ as the $j^{th}$ row of $U\Lambda^{1/2}$, so we have

$$
C^i = \begin{pmatrix}
a_i^1 \\
\vdots \\
a_i^N
\end{pmatrix}
\begin{pmatrix}
(a_1^1)^T, & \ldots, & (a_N^1)^T
\end{pmatrix}.
\quad (4.15)
$$
By (4.14) and (4.15), the model correlation so obtained is

$$\text{Corr}(\Delta f_j(t_i), \Delta f_k(t_i)) = \frac{\Delta t_i a^j_k \cdot a^k_j}{\sqrt{\Delta t_i} ||a^j_i||_2 \cdot \sqrt{\Delta t_i} ||a^k_i||_2} = C^i_{jk},$$

where

$$\Delta f_j(t_i) = f_j^{s^i}(t_i) s^i_j a^j_i \cdot [\sigma_{j+1} \Delta t_i + \Delta Z(t_i)]$$

for some small $\Delta t_i$. Note that the columns of matrix $U\Lambda^{1/2}$ are called principle components of the matrix $C^i$.

The complication in determination of $a^i_j, j = 1, \ldots, n$ is that the rank of $C^i$ is in general much bigger than $n$. The former is typically equal to $N - i + 1$, the number of forward rate "alive". In such case the above procedure for calculating $a^i_j, j = 1, \ldots, n$ does not work. Therefore, a preprocessing is in general needed to reduce the rank of the given correlation matrix. For a given correlation matrix $C^i$, preprocessing is naturally formulated as the following minimization problem with constraints:

$$\min_{\hat{C}^i} \| C^i - \hat{C}^i \|_F,$$

s.t. $\hat{C}^i \geq 0, \ rank(\hat{C}^i) \leq n, \hat{C}^i_{kk} = 1, \ k = i, \ldots, N.$ \hspace{1cm} (4.16)

where the subindex $F$ means the Frobenius’s norm, and $\hat{C} \geq 0$ means that $\hat{C}$ is a non-negative matrix. We need to solve (4.16) for $i = 1, 2, \ldots, N$.

Once we have obtained the low-rank approximation of the correlation matrices we can proceed to the determination of the forward rate volatilities $s^i_j$, subject to the input prices. The number of $\{s^i_j\}$ to be determined is $N(N + 1)/2$. This is typically much higher than the number of the input prices. Hence, this is again an under-determined problem. For the uniqueness of the solution we must adopt an objective function, which also serves as regularization condition for smoothness and stability of the volatility surface. A rather natural candidate for the objective function is

$$\|\nabla s\|^2 + \epsilon \| s - s_0 \|^2 \equiv -(s, (\nabla \cdot \nabla) s) + \epsilon \| s - s_0 \|^2 \text{ for some } \epsilon > 0.$$ \hspace{1cm} (4.17)
Here \((\nabla \cdot \nabla)\) stands for the discrete Laplacian:

\[
(\nabla \cdot \nabla)s_j^t = s_{j-1}^t + s_{j+1}^t + s_{j-1}^{t-1} + s_{j+1}^{t-1} - 4s_j^t,
\]

hence we can rewrite right side of (4.17)

\[
\sum_{i=1}^{N} \sum_{j=1}^{N} s_j^t (-s_{j-1}^t - s_{j+1}^t - s_{j-1}^{t-1} - s_{j+1}^{t-1} + 4s_j^t) + \epsilon \sum_{i=1}^{N} \sum_{j=1}^{N} (s_j^t - s_{j,0}^t)^2.
\]

(4.19)

Note that \(s_0\) is a priori volatility surface prescribed as, for example, the volatility surface of the previous day. By minimizing (4.17), we try to achieve a balance between the smoothness and the stability of the volatility surface. In (4.19), there are some "ghost" variables whose sub- or sup-indices are out of the designated range. These "ghost" variables are eliminated by using "Neumann boundary condition":

\[
s_j^0 = s_j^1, \quad j = 1, 2, \ldots, N,
\]

\[
s_{N+1}^t = s_N^t,
\]

\[
s_j^{i+1} = s_i^t, \quad i = 1, 2, \ldots, N,
\]

\[
s_i^{i-1} = s_i^t.
\]

The constraints of prices are expressed in terms of the implied CEV volatilities:

\[
\xi_{m,n}^2 = \sum_{j,k=m,n-1} \omega_j \omega_k \sum_{i=1}^{m} s_j^i s_k^i \hat{C}_{j,k}^i \Delta T_{i-1}
\]

\[
= \sum_{i=1}^{m} \Delta T_{i-1} \sum_{j,k=m,n-1} s_j^i s_k^i (\omega_j \omega_k \hat{C}_{j,k}^i), \quad \text{for some } m \text{ and } n,
\]

(4.20)

where \(\hat{C}_{j,k}^i\) in (4.20) is an element of \(\hat{C}^i\), the low-rank approximation to correlation matrix \(C^i\). Especially, when \(n = m + 1\) we have \(\omega_j = 1, \omega_k = 0, k \neq j\), and condition (4.20) reduces to

\[
\xi_j^2 = \sum_{i=1}^{j} \Delta T_{i-1} (s_j^i)^2 \quad \text{for some } j.
\]

(4.21)
Note that all functions in (4.19)-(4.20) are quadratic functions in \( \{ s^j \} \). Such feature, as we shall see later, gives rise to a powerful numerical method.

For practical numerical implementation, we introduce matrix notations for the problem. First we line up the volatilities in a one-dimensional array

\[
X = \begin{pmatrix}
  s^1 \\
  s^2 \\
  \vdots \\
  s^N
\end{pmatrix}
\]

with

\[
s^i = \begin{pmatrix}
  s^i_1 \\
  s^i_2 \\
  \vdots \\
  s^i_N
\end{pmatrix}
\]

We then denote \( B \) the matrix corresponding to gradient part of discrete form of objective function, and associate each instrument with the following "weight" matrix

\[
W_{m,n} = \text{diag}(0, \ldots, 0, \omega_m, \ldots, \omega_{n-1}, 0, \ldots, 0),
\]

and partition diagonal "correlation matrix"

\[
G_{m,n} = \text{diag}(\Delta T_0 W_{m,n} \hat{C}^1 W_{m,n}, \ldots, \Delta T_{m-1} W_{m,n} \hat{C}^m W_{m,n}, 0, \ldots, 0).
\]

With the above matrices, the calibration to prices under the objective function (4.17) can be cast into a neat yet equivalent form

\[
\begin{align*}
\min_X & \quad X^T B X + \epsilon (X - X_0)^T (X - X_0), \\
\text{s.t.} & \quad X^T G_{m,n} X = \xi^2_{m,n}, \quad \text{for some } m \text{ and } n.
\end{align*}
\]
The objective function in (4.22) can be simplified further. Expanding the function we have

\[
X^TBX + \epsilon(X - X_0)^T(X - X_0)
= (X - \epsilon(B + \epsilon I)^{-1}X_0)^T(B + \epsilon I)(X - \epsilon(B + \epsilon I)^{-1}X_0)
+ \epsilon(X_0)^T(I - \epsilon(B + \epsilon I)^{-1})X_0,
\]

where the last term is a constant and thus can be ignored for optimization purpose. Introducing

\[
A = B + \epsilon I,
\]

which is a positive-definite matrix, we finally formulate the problem of price calibration into

\[
\begin{align*}
\min_X & (X - \tilde{X}_0)^TA(X - \tilde{X}_0), \\
\text{s.t.} & \quad X^TG_{m,n}X = \xi_{m,n}^2, \quad \text{for some } m \text{ and } n.
\end{align*}
\]

(4.23)

where \(\tilde{X}_0 = \epsilon(B + \epsilon I)^{-1}X_0\) is a constant matrix.

The recast form (4.23) highlights that the objection function is in quadratic form with a positive definite matrix. Nevertheless, the Lagrange multiplier problem corresponding to (4.23) may not be well-defined in the sense that, for a set of finite multipliers, the maximum of the inner maximization problem can be infinity. For this reason, we superimpose a convex function to the objective function:

\[
\begin{align*}
\min_X U((X - \tilde{X}_0)^TA(X - \tilde{X}_0)), \\
\text{s.t.} & \quad X^TG_{m,n}X = \xi_{m,n}^2, \quad \text{for some } m \text{ and } n,
\end{align*}
\]

(4.24)

where \(U(y)\) is superlinear and monotonically increasing function for \(y \geq 0\), for examples, \(U(y) = y^2\), \(y \ln y\) or \(e^y\). Problem (4.24) then shares the same constrained minimums with (4.23). We will proceed next to solve the problem with the usual approach of descend.
4.3 Review of solution methodology

In this section we want to review some known results obtained by Wu, and show the procedure to solve the constrained minimization problems (4.16) and (4.24). Roughly speaking, this methodology is the combinations of the method of Lagrange multiplier and steepest descend. In developing the numerical methods, we have taken full advantages of the special structure of the objective function and constraints.

4.3.1 Eigen-decomposition-based rank reduction algorithm

For a given non-negative symmetric $N$ by $N$ matrix $C$, we define a low-rank approximation as the solution to the following problem

$$\min_X \|C - X\|_F, \quad \text{s.t.} \quad \text{rank}(X) \leq n < N, \quad \text{diag}(X) = \text{diag}(C).$$

(4.25)

We denote any solution to problem (4.18) by $C^*$, and the feasible set of solutions by

$$\mathcal{F} = \{X \in \mathbb{R}^{N \times N} | \text{rank}(X) \leq n, \ \text{diag}(X) = \text{diag}(C)\}.$$

For applications in the market model, $C^*$ will serve subsequently as a correlation matrix and thus is expected to be a non-negative definite matrix. It may seem that the feasible set of the optimal problem should be $\mathcal{F}^+$, the subset of $\mathcal{F}$ that consists of only positive semi-definite matrices. Inevitably, adding explicitly such constraint will increase the difficulty of the problem. Fortunately, it was proved that the solutions to (4.25) are automatically positive semi-definite, given $C$ a positive semi-definite matrix, hence the explicit imposition of the extra constraint becomes unnecessary.

Wu's approach for solving the constrained optimal approximation problem is to transform it to an equivalently min-max problem by the method of Lagrange multiplier.
Let $\mathcal{R}_n$ be the set of $N \times N$ matrices with rank less or equal to $n$. The Lagrange multiplier problem corresponding to (4.25) is defined as the following problem:

$$
\min_d \max_{X \in \mathcal{R}_n} L(X, d),
$$

(4.26)

with the Lagrange function:

$$
L(X, d) = -\|C - X\|_F^2 - 2d^T \text{diag}(C - X),
$$

(4.27)

where $d$ is the vector of the multipliers. Note that $L(X, d)$ is linear in $d$ in the following sense:

$$
L(X, td + (1 - t)\hat{d}) = tL(X, d) + (1 - t)L(X, \hat{d}).
$$

(4.28)

It also can be rigorously justified that the constrained problem (4.26) is equivalent to the problem (4.25).

Numerically the min-max problem (4.26) is treated as an minimization problem of the form

$$
\min_d V(d),
$$

(4.29)

with the objective function defined by

$$
V(d) = \max_{X \in \mathcal{R}_n} L(X, d).
$$

(4.30)

It is a matter then to look for efficient methods separately for the maximization problem (4.30) and minimization problem (4.29).

For the maximization problem (4.30), it is crucial to observe that the Lagrange function can be written into

$$
L(X, d) = -\|C + D - X\|_F^2 + \|d\|_2^2,
$$

(4.31)

where $D$ is the diagonal matrix of $d$ ($D = \text{diag}(d)$). For fixed $d$, obviously, the maximizer to (4.29) can be obtained by the eigenvalue decomposition of matrix $C + D$ (which is
symmetric). Let

\[ C + D = U \Lambda U^T \]  (4.32)

be the eigenvalue decomposition with orthogonal matrix \( U \) and eigenvalue matrix

\[ \Lambda = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_N). \]

Note that both \( U \) and \( \Lambda \) depend on the multiplier vector \( d \), hence they are also denoted by \( U(d) \) and \( \Lambda(d) \) when highlighting the dependence is necessary. We assume that the diagonal elements are put in the decreasing order in magnitude:

\[ |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_N|. \]

The solutions to the problem (4.30), the best rank-n approximations of \( C + D \), are obviously given by

\[ C(d) = C_n(d) = U_n \Lambda_n U_n^T, \]  (4.33)

where \( U_n \) is the matrix consisting of the first \( n \) columns of \( U \), and \( \Lambda_n = \text{diag}(\lambda_1, \ldots, \lambda_n) \) is the principle submatrix of \( \Lambda \) of order \( n \). Consequently we have

\[ V(d) = -\sum_{j=n+1}^{N} \lambda_j^2 + \|d\|_2^2. \]  (4.34)

Clearly, when \( |\lambda_n| > |\lambda_{n+1}| \), the solution to (4.29) is unique. In the case \( |\lambda_n| = |\lambda_{n+1}| \), the number of solutions becomes non-unique or even infinite. Throughout this chapter we limit ourselves to the case of \( |\lambda_n| > |\lambda_{n+1}| \).

Also in Zhang and Wu’s paper (2001), they proved the existence and uniqueness of the solution to that constrained problem and showed the optimal solution is differentiable in \( d \), hence the inner maximization problem (4.30) can be solved nicely by an eigenvalue decomposition. The outer minimization (4.29) then is dealt with the method of steepest descend. Below I list the algorithm.

**Algorithm:** Take \( D^{(0)} \) to be a null matrix, and repeat the following steps:

38
1. Compute the eigen-decomposition of \( C + D^{(k)}:C + D^{(k)} = U^{(k)}\Lambda^{(k)}(U^{(k)})^T \); set \( \alpha^{(k)} = 1 \) and \( \nabla V(d^{(k)}) = -2\text{diag}(C - U^{(k)}_n\Lambda^{(k)}_n(U^{(k)}_n)^T) \);

2. Define \( d^{(k+1)} = d^{(k)} - \alpha^{(k)}\nabla V(d^{(k)}) \);

3. If \( V(d^{(k+1)}) > V(d^{(k)}) - \frac{\alpha^{(k)}}{2}\|\nabla V(d^{(k)})\|^2 \), take \( \alpha^{(k)} := \alpha^{(k)}/2 \), go back to step 2;

4. If \( \|d^{(k+1)} - d^{(k)}\|_2 > \text{tol} \), go to step 1;

5. Take \( d^* = d^{(k+1)} \) and \( C^* = U^{(k)}_n\Lambda^{(k)}_n(U^{(k)}_n)^T \).

### 4.3.2 Eigenvalue problem for calibration of input prices

The calibration to input prices has been formulated in the concise form (4.24). The corresponding Lagrange multiplier problem is

\[
\min_d \max_X L(X, d),
\]

where

\[
L(X, d) = -\left((X - X_0)^TA(X - X_0)\right)^2 + 2 \sum_{i=1}^N d_i(X^TG_iX - h_i).
\]

Note that for simplicity we have used \( \{G_i, h_i\} \) in place of \( \{G_{m,n}, \ell^2_{m,n}\} \). To facilitate discussions we denote the value function for the outer minimization problem by

\[
V(d) = \max_X L(X, d),
\]

and feasible set of solutions by

\[
\mathcal{F} = \{X|X^TG_iX = h_i, i = 1, \ldots, N\}.
\]

Due to the positive-definiteness of the matrix \( A \), the maximizer of \( L(X, d) \) is finite for fixed \( d \), and \( V(d) \) therefore exists for all \( d \). It is obvious that the Lagrange function is
smooth in $X$. Any solutions to the maximization problem (4.37) hence must be critical point of the Lagrange function satisfying the following first-order condition

$$
\left((X - X_0)^T A (X - X_0)\right)(X - X_0) = \left(\sum d_i G_i\right)X = \left(\sum d_i G_i\right) (X - X_0) + \left(\sum d_i G_i\right) X_0.
$$

(4.38)

For simplicity we denote $B_d = \sum d_i G_i$ and $Y = X - X_0$. Equation (4.38) then becomes

$$
[(Y^TAY)A - B_d]Y = B_d X_0.
$$

(4.39)

The above equation can be solved through eigenvalue decomposition. Let $(\lambda_i, u_i)$ be the eigenpairs of $(B_d, A)$ such that

$$
B_d u_i = \lambda_i A u_i, \quad (4.40)
$$

$$
u_i^T A u_i = 1, \quad u_i^T B_d u_i = \lambda_i, \quad (4.41)
$$

$$
\lambda_i \geq \lambda_{i+1}, \quad i = 1, 2, \ldots, N - 1, \quad (4.42)
$$

and let $U$ denote the orthonormal eigenvector matrix of $A$

$$
U = (u_1, u_2, \ldots, u_N). \quad (4.43)
$$

To solve for $Y$, we let $\alpha = Y^T A Y$ and pre-multiply $U^T$ to equation (4.39), we then arrive at

$$
[\alpha I - \Lambda] U^{-1} Y = U^T B_d X_0, \quad (4.44)
$$

where

$$
\Lambda = diag(\lambda_1, \lambda_2, \ldots, \lambda_N)
$$

is the eigenvalue matrix. From (4.44) we obtain the solution to (4.39) as

$$
Y \equiv \tilde{Y}(\alpha) = U [\alpha I - \Lambda]^{-1} U^T B_d X_0,
$$

40
where the scalar $\alpha$ is subject to the nonlinear equation

$$\tilde{Y}^T(\alpha)AY(\alpha) = \alpha,$$

which can be solved with a few steps of iteration. When $X_0 = 0$, in particular, (4.39) becomes an eigenvalue problem and thus admits multiple solutions, and they are given by

$$Y_i = \sqrt{\max\{\lambda_i, 0\}}u_i, \quad i = 1, 2, \ldots, N.$$

To fix idea we will from now on concentrate on the special case of $X_0 = 0$, function $L(X, d)$ achieves its maximum at

$$Y = Y_1,$$  \hspace{1cm} (4.45)

and consequently

$$V(d) = \max_{j, \lambda_j \geq 0} - (Y_j^TAY_j)^2 + 2 \sum_{i=1}^{N} d_i (Y_j^T G_i Y_j - h_i)$$

$$= \max_{\lambda_j \geq 0} \lambda_j^2 - 2 \sum_{i=1}^{N} d_i h_i$$

$$= \lambda_1^2 - 2 \sum_{i=1}^{N} d_i h_i.$$

Quite often the eigenvector corresponding to the largest eigenvalue of a positive semidefinite matrix is the smoothest one amongst all eigenvectors, and it seems to be the case in our application. Very encouragingly, then result in (4.45) establishes the connection between the smoothest fit of volatility surface and the smoothest eigenvector of a generalized eigenvalue problem.

For the analytical properties of $\lambda_1(d), Y(d)(= Y_1(d))$ and $V(d)$ we have

**Theorem 4.3.1 (Wu, 2003)** If $\lambda_1(d) > \lambda_2(d),$ then

1. $(\lambda_1(d), Y(d))$ is differentiable with respect to $d$ locally;
2. $V(d)$ is differentiable in $d$ locally;

3. The gradient of $V(d)$ is

$$
\nabla_d V(d) = 2 \begin{pmatrix}
Y^T G_1 Y - h_1 \\
Y^T G_2 Y - h_2 \\
\vdots \\
Y^T G_N Y - h_N
\end{pmatrix};
$$

4. The elements of the Hessian matrix are given by

$$
H_{ij}(d) \equiv \frac{\partial^2 V}{\partial d_i \partial d_j} = 4Y^T G_i U \Phi^{-1} U^T G_j Y, 
$$

(4.46)

where

$$
\Phi = \lambda_1 I + 2 \begin{pmatrix}
\lambda_1 & 0 & \cdots & 0 \\
0 & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & 0 \\
0 & \cdots & 0 & \Lambda_i
\end{pmatrix}, 
$$

(4.47)

and the Hessian is a non-negative definite matrix.

**Remark:** In general, using this theorem, that min-max problem is nicely solved by the gradient-based algorithm, yet, when Hessian is not expensive to obtain, we should definitely use a Hessian-based algorithm for the numerical solution, and in latter section, I will use this Hessian-based algorithm to deal with the practical problem.
4.4 Practical problem and numerical results

In this section we present a practical problem. In this example, we want to calibrate the market model to the prices of a set of benchmark instruments and the correlation matrix of some forward rates. As we discuss in previous section, the goal of this calibration problem is to find out the multi-factor volatility coefficients and the elasticity for all forward rates. Here for simplicity, we just assume all forward rates and swap rate share the same elasticity, note that this assumption does not affect producing volatility skews.

In this example, all data are taken from USD market, we have a set of USD yield curve data from July 01, 1999 to July 01, 2002, and swaption and cap volatility of various maturities, for the date of July 04, 2002. The yield curves of treasury rates are plotted in Figure 4.8 while the Black-Scholes volatility of a set of caps with various strikes and at-the-money swaptions are listed in Table 4.1 and Table 4.2 (the first column and row are maturities and strikes respectively), next we need to use bootstrapping procedure to compute the implied volatility of caplets with various strikes.

<table>
<thead>
<tr>
<th>M \ S</th>
<th>1.5%</th>
<th>2%</th>
<th>3%</th>
<th>3.5%</th>
<th>4%</th>
<th>5%</th>
<th>6%</th>
<th>7%</th>
<th>8%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1y</td>
<td>51.70%</td>
<td>43.80%</td>
<td>38.00%</td>
<td>38.60%</td>
<td>39.30%</td>
<td>43.10%</td>
<td>44.90%</td>
<td>48.00%</td>
<td>52.10%</td>
</tr>
<tr>
<td>2y</td>
<td>55.60%</td>
<td>46.90%</td>
<td>39.50%</td>
<td>37.10%</td>
<td>35.00%</td>
<td>33.50%</td>
<td>32.70%</td>
<td>32.70%</td>
<td>33.00%</td>
</tr>
<tr>
<td>3y</td>
<td>52.30%</td>
<td>45.00%</td>
<td>37.30%</td>
<td>34.50%</td>
<td>32.20%</td>
<td>29.40%</td>
<td>28.30%</td>
<td>27.90%</td>
<td>27.70%</td>
</tr>
<tr>
<td>4y</td>
<td>50.60%</td>
<td>43.60%</td>
<td>35.90%</td>
<td>33.20%</td>
<td>30.90%</td>
<td>27.80%</td>
<td>26.40%</td>
<td>25.90%</td>
<td>25.70%</td>
</tr>
<tr>
<td>5y</td>
<td>48.50%</td>
<td>42.10%</td>
<td>34.60%</td>
<td>32.00%</td>
<td>29.80%</td>
<td>26.60%</td>
<td>25.00%</td>
<td>24.50%</td>
<td>24.20%</td>
</tr>
<tr>
<td>6y</td>
<td>46.90%</td>
<td>40.80%</td>
<td>33.50%</td>
<td>31.00%</td>
<td>29.90%</td>
<td>25.80%</td>
<td>24.00%</td>
<td>23.40%</td>
<td>23.00%</td>
</tr>
<tr>
<td>7y</td>
<td>45.60%</td>
<td>39.70%</td>
<td>32.60%</td>
<td>30.20%</td>
<td>28.10%</td>
<td>25.10%</td>
<td>23.10%</td>
<td>22.50%</td>
<td>22.20%</td>
</tr>
<tr>
<td>8y</td>
<td>44.30%</td>
<td>38.60%</td>
<td>31.80%</td>
<td>29.40%</td>
<td>27.40%</td>
<td>24.40%</td>
<td>22.50%</td>
<td>21.80%</td>
<td>21.40%</td>
</tr>
<tr>
<td>9y</td>
<td>43.10%</td>
<td>37.60%</td>
<td>31.00%</td>
<td>28.70%</td>
<td>26.80%</td>
<td>23.80%</td>
<td>21.80%</td>
<td>21.10%</td>
<td>20.70%</td>
</tr>
<tr>
<td>10y</td>
<td>42.00%</td>
<td>36.60%</td>
<td>30.20%</td>
<td>28.00%</td>
<td>26.10%</td>
<td>23.30%</td>
<td>21.30%</td>
<td>20.50%</td>
<td>20.10%</td>
</tr>
</tbody>
</table>

Table 4.1: Implied Black Volatilities for Caps
From Table 4.1, cap prices are quoted in terms of implied cap volatilities, let $C_{N,K}$ denote the market value of a spot starting cap with maturity $T_N$ and strike $K$. By definition, the cap volatility $\bar{\sigma}_{N,K}$ is the number that makes the following equation true:

$$C_{N,K} = \sum_{n=0}^{N-1} P(0,T_{n+1})\Delta_n BSC(f_n, K, T_{n+1}, \bar{\sigma}_{N,K}),$$

(4.48)

where $BSC(f, K, T, \sigma)$ is the Black-Scholes formula for a call with forward rate $f$, strike $K$, maturity $T$, and volatility $\sigma$. Cap volatility is the volatility implied by the cap prices.

Now we want to transfer the cap volatilities into the implied caplet volatilities. To give a precise definition of these quantities, consider a sequence of caps with same strike $K$ and different maturities $T_1, T_2, \ldots, T_N$. We assume that the prices of these caps are known. For each $n = 1, 2, \ldots, N$, we write

$$C_{n,K} = \sum_{m=0}^{n-1} Caplet_{n,m}(K),$$

(4.49)

where $Caplet_{n,m}(K)$ represents the "model value" of the caplet corresponding to $(m+1)^{st}$ cash-flow date and strike $K$. Notice that the value of the first caplet is unequivocally known at time zero, since the cash-flow to be received at time $T_1$ is known with certainty.
<table>
<thead>
<tr>
<th>Maturity xTenor</th>
<th>ATM Strike(%)</th>
<th>Black Vol(%)</th>
<th>Maturity xTenor</th>
<th>ATM Strike(%)</th>
<th>Black Vol(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25 x 1</td>
<td>2.2285</td>
<td>44.50</td>
<td>0.25 x 2</td>
<td>3.1450</td>
<td>39.20</td>
</tr>
<tr>
<td>0.25 x 3</td>
<td>3.7440</td>
<td>33.60</td>
<td>0.25 x 4</td>
<td>4.1680</td>
<td>30.70</td>
</tr>
<tr>
<td>0.25 x 5</td>
<td>4.4710</td>
<td>28.50</td>
<td>0.25 x 7</td>
<td>4.9070</td>
<td>25.90</td>
</tr>
<tr>
<td>0.25 x 10</td>
<td>5.2970</td>
<td>23.50</td>
<td>0.5 x 2</td>
<td>3.1450</td>
<td>35.70</td>
</tr>
<tr>
<td>0.5 x 3</td>
<td>3.7440</td>
<td>31.60</td>
<td>0.5 x 4</td>
<td>4.1680</td>
<td>29.20</td>
</tr>
<tr>
<td>0.5 x 5</td>
<td>4.4710</td>
<td>27.30</td>
<td>0.5 x 7</td>
<td>4.9070</td>
<td>25.00</td>
</tr>
<tr>
<td>0.5 x 10</td>
<td>5.2970</td>
<td>22.90</td>
<td>1 x 2</td>
<td>3.1450</td>
<td>30.90</td>
</tr>
<tr>
<td>1 x 3</td>
<td>3.7440</td>
<td>28.30</td>
<td>1 x 4</td>
<td>4.1680</td>
<td>26.30</td>
</tr>
<tr>
<td>1 x 6</td>
<td>4.7130</td>
<td>23.90</td>
<td>1 x 9</td>
<td>5.1920</td>
<td>22.10</td>
</tr>
<tr>
<td>2 x 2</td>
<td>3.1450</td>
<td>25.80</td>
<td>2 x 3</td>
<td>3.7440</td>
<td>24.40</td>
</tr>
<tr>
<td>2 x 5</td>
<td>4.4710</td>
<td>22.50</td>
<td>2 x 8</td>
<td>5.0630</td>
<td>21.00</td>
</tr>
<tr>
<td>3 x 2</td>
<td>3.1450</td>
<td>23.50</td>
<td>3 x 4</td>
<td>4.1680</td>
<td>21.80</td>
</tr>
<tr>
<td>3 x 7</td>
<td>4.9070</td>
<td>20.30</td>
<td>4 x 3</td>
<td>3.7440</td>
<td>21.00</td>
</tr>
<tr>
<td>4 x 6</td>
<td>4.7130</td>
<td>19.40</td>
<td>5 x 5</td>
<td>4.4710</td>
<td>18.60</td>
</tr>
</tbody>
</table>

Table 4.2: Implied Black Volatilities for Swaptions

The other caplets must be priced with the model, consistently with equation (4.49). We define the caplet volatilities \( \sigma_n(K) \), \( n = 0, 1, \ldots, N - 1 \) recursively, as follows:

\[
\sigma_0(K) = 0 = \text{implied volatility}(C_{1,K}),
\]

and

\[
\sigma_1(K) = \text{implied volatility}(C_{2,K} - C_{1,K}).
\]

The procedure is repeated each time a new cap is introduced, the caplet volatility \( \sigma_n(K) \) is then defined as

\[
\sigma_{n-1}(K) = \text{implied volatility}(C_{n,K} - C_{n-1,K}).
\]
After the bootstrapping procedure, the implied caplet volatilities are listed in Table 4.3 whose first row and column denote the maturities of the forward rates and strikes. Also the correlation matrix among forward rates is given in Table 4.4 (which are calculated with three-year USD yield curve data). The first row and column of the table show the maturities of the forward rates.

Then let us take a look at the low-rank approximations to the market correlation matrix. We have calculated the rank-one, two, three, six approximations and the outcomes are visualized in Figure 4.9-4.13, where the first figure is the original market correlation surface. By vision we shall agree that the approximations improve with rank increasing.

<table>
<thead>
<tr>
<th>M \ S</th>
<th>1.5%</th>
<th>2%</th>
<th>3%</th>
<th>3.5%</th>
<th>4%</th>
<th>5%</th>
<th>6%</th>
<th>7%</th>
<th>8%</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25y</td>
<td>51.70%</td>
<td>43.80%</td>
<td>38.00%</td>
<td>38.60%</td>
<td>39.30%</td>
<td>43.10%</td>
<td>44.90%</td>
<td>48.00%</td>
<td>52.10%</td>
</tr>
<tr>
<td>0.5y</td>
<td>51.70%</td>
<td>43.80%</td>
<td>38.00%</td>
<td>38.60%</td>
<td>39.30%</td>
<td>43.10%</td>
<td>44.90%</td>
<td>48.00%</td>
<td>52.10%</td>
</tr>
<tr>
<td>1y</td>
<td>56.84%</td>
<td>49.02%</td>
<td>39.05%</td>
<td>36.86%</td>
<td>34.60%</td>
<td>33.25%</td>
<td>32.56%</td>
<td>32.58%</td>
<td>32.84%</td>
</tr>
<tr>
<td>2y</td>
<td>49.33%</td>
<td>42.48%</td>
<td>34.80%</td>
<td>31.94%</td>
<td>29.83%</td>
<td>26.88%</td>
<td>26.39%</td>
<td>26.42%</td>
<td>26.49%</td>
</tr>
<tr>
<td>3y</td>
<td>48.06%</td>
<td>41.02%</td>
<td>33.27%</td>
<td>30.89%</td>
<td>28.74%</td>
<td>25.61%</td>
<td>24.38%</td>
<td>24.22%</td>
<td>24.34%</td>
</tr>
<tr>
<td>4y</td>
<td>44.10%</td>
<td>38.48%</td>
<td>31.22%</td>
<td>28.93%</td>
<td>27.07%</td>
<td>23.98%</td>
<td>22.46%</td>
<td>22.37%</td>
<td>22.24%</td>
</tr>
<tr>
<td>5y</td>
<td>42.65%</td>
<td>36.83%</td>
<td>29.79%</td>
<td>27.61%</td>
<td>25.90%</td>
<td>23.37%</td>
<td>21.40%</td>
<td>20.95%</td>
<td>20.68%</td>
</tr>
<tr>
<td>7y</td>
<td>39.15%</td>
<td>33.71%</td>
<td>27.84%</td>
<td>25.32%</td>
<td>23.81%</td>
<td>20.98%</td>
<td>19.87%</td>
<td>19.08%</td>
<td>18.63%</td>
</tr>
</tbody>
</table>

Table 4.3: Implied Black Volatilities for Caplets

Before we proceed calibration procedure, we also need to compute implied CEV volatility and elasticity firstly. Suppose we have a bunch of swaptions (note caplets are special cases of swaptions), we let \( p_i \) denote the market price for \( i-th \) swaption with maturity \( T_{1,i} \), tenor \( T_{2,i} \) and strike \( K_i \), \( i = 1, \ldots, N \). using least-square fitting, we want
Table 4.4: Historical Correlation Matrix for the Forward Rates

<table>
<thead>
<tr>
<th></th>
<th>0.25</th>
<th>0.5</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.25</td>
<td>1.0000</td>
<td>0.9864</td>
<td>0.9372</td>
<td>0.8778</td>
<td>0.8352</td>
<td>0.8017</td>
<td>0.7952</td>
<td>0.7919</td>
</tr>
<tr>
<td>0.5</td>
<td>0.9864</td>
<td>1.0000</td>
<td>0.9762</td>
<td>0.9323</td>
<td>0.8988</td>
<td>0.8669</td>
<td>0.8657</td>
<td>0.8595</td>
</tr>
<tr>
<td>1</td>
<td>0.9372</td>
<td>0.9762</td>
<td>1.0000</td>
<td>0.9812</td>
<td>0.9596</td>
<td>0.9376</td>
<td>0.9308</td>
<td>0.9322</td>
</tr>
<tr>
<td>2</td>
<td>0.8778</td>
<td>0.9333</td>
<td>0.9812</td>
<td>1.0000</td>
<td>0.9891</td>
<td>0.9729</td>
<td>0.9733</td>
<td>0.9623</td>
</tr>
<tr>
<td>3</td>
<td>0.8352</td>
<td>0.8988</td>
<td>0.9506</td>
<td>0.9891</td>
<td>1.0000</td>
<td>0.9771</td>
<td>0.9793</td>
<td>0.9670</td>
</tr>
<tr>
<td>4</td>
<td>0.8017</td>
<td>0.8669</td>
<td>0.9376</td>
<td>0.9729</td>
<td>0.9771</td>
<td>1.0000</td>
<td>0.9761</td>
<td>0.9735</td>
</tr>
<tr>
<td>5</td>
<td>0.7952</td>
<td>0.8657</td>
<td>0.9398</td>
<td>0.9733</td>
<td>0.9793</td>
<td>0.9761</td>
<td>1.0000</td>
<td>0.9826</td>
</tr>
<tr>
<td>7</td>
<td>0.7919</td>
<td>0.8595</td>
<td>0.9322</td>
<td>0.9623</td>
<td>0.9670</td>
<td>0.9735</td>
<td>0.9826</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

Figure 4.9: Market correlation surface

Figure 4.10: Correlation surface of one-factor model

Figure 4.11: Correlation surface of two-factor model

Figure 4.12: Correlation surface of three-factor model
Figure 4.13: Correlation surface of six-factor model

Figure 4.14: Volatility surface of one-factor model

Figure 4.15: Volatility surface of two-factor model

Figure 4.16: Volatility surface of three-factor model

Figure 4.17: Volatility surface of six-factor model
to minimize the objective function

$$\sum_{i=1}^{N} (PS(0, T_{1,i}, T_{2,i}, K_i, \alpha, \xi^2) - p_i)^2.$$ 

Note that we assume all forward rates and swap rate share the same elasticity, and $PS(\cdot)$ is the model price for swaption which can be calculated by Theorem (4.2.1). Hence we can translate these prices into implied CEV volatilities and elasticity. For this problem, the elasticity is

$$\text{elasticity} = \alpha - 1 = -0.72.$$ 

With calculated low-rank approximation of the correlation matrix, we proceed to calculate the forward rates volatilities from the input implied CEV volatilities of caplets and swaptions. The results are plotted as volatility surfaces from Figure 4.14 to 4.17. The algorithm is implemented with a Hessian-based unconstrained minimization function in MATLAB (“fminunc” in specific).

### 4.5 Summary

We use a practical problem to illustrate how to use an efficient methodology to calibrate the CEV market model to a collection of caplet/swaption prices and historical correlation matrix, it can be regarded as a complement of Wu (2003). Numerical studies confirm the efficiency and reliability of the method. From the viewpoint of financial engineering, the whole important problem of calibrating to the volatility skews in the LIBOR derivative markets is completely solved. From mathematical point of view, this powerful new technique for general constrained minimization problem with quadratic objective function and constraints is nicely to be implemented.
Chapter 5

Market Model With Stochastic Volatility

5.1 Introduction

In this chapter we extend the standard LIBOR market model (Brace, Gatarek and Musiela (1997), Jamshidian (1997), and Miltersen, Sandermann (1997)) by taking stochastic volatilities. Over the past few years the market model has established itself as the benchmark in the market places. One of the many merits of the market model is that it renders Black’s formula for caplet and swaption prices, which is in nice agreement with the market practice. The existence of the closed-form solutions for caplets and swaptions makes calibration an amenable task, and thus has strengthened the role of the market model to be an aggregated model for the risk management of derivative portfolios.

Like other financial models, the standard market model also has its drawbacks. One of the major drawbacks is the lack of flexibility to cope with volatility smiles or skews, due to the fact that forward rate volatilities can only be functions of calendar and forward time, as otherwise the Black’s formula would be invalid. Since the phenomena of
volatility smiles/skews is persistent in major currency markets, there is a genuine need
to deal with volatility smiles/skews under the basic principles of the market model.

In Chapter 2, we introduce the main three lines of researches on the extension of
the market model. Andersen and Andreasen (2000) adopted the Constant Elasticity
Variance (CEV) extension to the lognormal process. By taking the elasticity parameter
bigger or less than one, they can generate monotonically decreasing or monotonically
increasing volatility skews. One notable merit of their work is the closed-form price
formula for options, in terms of $\chi^2$ functions, which establishes a one-to-one correspon-
dence between option prices and their implied CEV volatilities. Calibration then can
be done with the implied CEV volatilities, in a way similar to the calibration of the
standard model, and this calibration problem has been solved in Chapter 4 of this the-
sis. Using the CEV model as a "bedrock", Andersen and Brotherton-Ractliffe (2002)
later superimpose a stochastic volatility factor, which effectively adds more curvature
so that non-monotonic volatility skews can be generated. Joshi and Rebonato (2002),
meanwhile, took the Displaced Diffusion (DD) process. They expressed the forward
rate volatilities as an exponential function, and the latter is parameterized in term of
four Orstein-Uhlenbeck processes. One important motivation for their extension was to
regenerate "hockey stick" shaped volatility curves. By adding more state variables, they
were able to capture such complex shapes. Their model, however, does not yield closed-
form solution to caplets or swaptions, and one has to resort to Monte Carlo simulation to
valuate almost any derivative securities. Glasserman and Kou (2000) developed a rather
comprehensive term structure theory with the jump-diffusion dynamics of forward rates.
Based on that Glasserman and Merener (2001) obtained approximated closed-form solu-
tions for caplets and swaptions. Their model can generated volatility smiles or skews by
taking different mean jump size. In particular, it can generate sharp short-term skews.
The jump-diffusion model is theoretically appealing, but practitioners tend to believe
that jump risk may be the secondary factor for the formation of volatility smiles/skews in the interest-rate derivative markets. The primary factor is believed to be stochastic volatility.

Our generalization of LIBOR market model sets off from the same footing as that of Andersen and Andreasen (2000). We adopt a multiplicative stochastic factor for the volatilities of all involved forward rates. We, however, apart from the their approach by allowing correlation between the stochastic factor and the forward rates. The primary motivation for us to take such approach is to build a model based on the genuine mechanism behind the volatility smile and skew, which we believe are stochastic volatility and its correlation with the state variables. In fact, a stochastic volatility model with or without correlation naturally generates volatility skew or smile. A negative correlation, in particular, leads to a downward volatility skew, and such downward skew is related to the premium in option prices that has existed since the 1987 crash (Bates, 1991) for large down-side risk. Moreover, from the empirical point of view, the stochastic volatility model enjoys the desirable property of time-stationarity. Our generalization is somewhat parallel to the development of stochastic volatility for equity options (see for instance, Wiggins (1987), Hull and White (1987) and Heston (1993)). We take, in particular, a squared-root process for the stochastic factor, which enlists desirable analytical tractability and is widely accepted as a sound process for stochastic volatility. The major contribution of our work is the introduction of approximate state-variable processes that carry analytical tractability for the extended model. Under the approximate model, we can develop closed-form formulae for caplets and swaptions, using the moment generating functions of the underlying state variables. Pricing comparisons are made with Monte Carlo method to support such approximations.

Another contribution of our work is the fast implementation method for the nu-
numerical evaluation of the closed-form formula. Our approximate processes admit closed-form solution of the moment generating function for the state variables. Carr and Madan (1998) showed that the Fourier transform of an option price can be expressed in terms of the moment generating function. Having the Fourier transform analytically, the option price itself can be obtained via an inverse Fourier transform, which can be implemented through Fast Fourier Transform (FFT) technology. The FFT technology facilitates fast calibration of the model and makes it capable to mark-to-market cap and swaption prices.

The extended model is developed under the risk neutral measure. It offers a more intuitive and appealing explanation to the mechanism of volatility smile/skew, and it is not necessarily intended to address other empirical issues of the real world dynamics of the interest rates. The extended model is parsimonious compared with other existing extensions. Partly because of that it has limited capacity to accommodate more complex shapes of implied volatility curve, like, for instance, the "hockey sticks".

The remaining part of this chapter will be organized as follows. In section 2 we set up the no-arbitrage extended LIBOR market model with stochastic volatility. Section 3 is devoted to the caplet and swaption pricing, where we discuss the necessary approximations to retain the analytical tractability. In section 4 we discuss the analytical solution of the moment generating function under the assumption of piecewise constant volatility coefficients and correlation between forward rates and the stochastic volatility factor. In section 5 we introduce the FFT method for numerical option valuation. In section 6 we present numerical results obtained by our extended model and Monte Carlo method, examine the price differences, and study the relation between the shape of implied volatility curve and the correlation. Finally we conclude in last section.
5.2 The market model with stochastic volatility

In this section we need to extend the market model to cope with stochastic volatility, firstly we will construct a no-arbitrage model with stochastic volatility.

**Proposition 5.2.1** Suppose the term structure of the instantaneous forward rate under risk neutral measure and with stochastic volatility is

\[
df(t, T) = \mu(t, T) dt + \sigma_f(t, T) \sqrt{V(t)} dZ_t, \quad (5.1)
\]

\[
dV(t) = \kappa(\theta - V(t)) dt + \epsilon \sqrt{V(t)} dW_t, \quad (5.2)
\]

with the following assumptions:

1. All objects \( \mu(t, T), V(t), \sigma_f(t, T) \), are assumed to be continuously differentiable in the \( T \)-variable.

2. All processes are assumed to be regular enough to allow us to differentiate under the integral sign as well as to interchange the order of integration.

Then this model is free of arbitrage if

\[
\mu(t, T) = \frac{\partial}{\partial T} \left[ \frac{1}{2} \int_t^T \sigma_f(t, u) \sqrt{V(u)} du \right] + \phi(t, T) \sqrt{V(t)} \quad (5.3)
\]

for some predictable process \( \phi \) and \( \psi \) satisfying

\[
\int_t^T [\phi^2 + \psi^2] ds < \infty, \quad (5.4)
\]

where \( \phi \) and \( \psi \) are market prices of risk associating to Brownian motions \( Z_t \) and \( W_t \), respectively.

Then the following theorem gives the arbitrage-free model under risk neutral measure.
Theorem 5.2.1  The arbitrage-free model for forward term rate $f_n$ under risk neutral measure is

$$\frac{df_n(t)}{f_n(t)} = \gamma_n(t)\sqrt{V(t)}dZ_t + \gamma_n \sum_{k=\lambda(t)}^{n} \frac{\Delta T_k \gamma_k(t)V(t)f_k(t)}{1 + \Delta T_k f_k(t)} dt,$$

(5.5)

where $\lambda(t)$ is a right-continuous mapping function which satisfies $T_{\lambda(t)-1} < t \leq T_{\lambda(t)}$.

Remarks: Recall

$$P(t, T_n) = e^{-\int_t^{T_n} f(t,s)ds},$$

and under risk neutral measure, from (A-2) and note

$$-\int_t^T \sigma_f(t, u)\sqrt{V(t)}du = B(t, T),$$

where $B(t, T)$ is defined in Appendix A.

Then

$$\frac{dP(t, T_n)}{P(t, T_n)} = r_t dt - \left( \int_t^{T_n} \sigma_f(t, u)\sqrt{V(t)}du \right) dZ_t$$

$$= r_t dt + M_n dZ_t,$$

where

$$M_n = -\int_t^{T_n} \sigma_f(t, u)\sqrt{V(t)}du.$$

If we assume under the risk neutral measure the zero coupon bond process is

$$\frac{dP(t, T_n)}{P(t, T_n)} = r_t dt - \sigma_n \sqrt{V} dZ_t,$$

compare the above two expression, we have

$$M_n = -\sigma_n \sqrt{V},$$

which is consistent with the fact

$$\sigma_n = \int_t^{T_n} \sigma_f(t, u)du.$$
Comparing the HJM framework, one can find that in our extension we replace the volatility term in the process of zero-coupon bond by volatility itself multiplied by a stochastic factor.

Now under risk neutral measure, the no arbitrage model for forward rate has another equivalent form

\[ df_n(t) = f_n(t)\gamma_n(t)\sqrt{V(t)}[dZ_t - \sqrt{V(t)}\sigma_{n+1}(t)dt]. \]

Now we get the arbitrage-free model for forward rates with stochastic volatility. In this chapter, we adopt this natural extension to the standard market model. From the discussion above, we allow the squared forward-rate volatilities to have a multiplicative factor that follows a squared-root process:

\[
\begin{align*}
  df_j(t) &= f_j(t)\sqrt{V(t)}\gamma_j(t) \cdot [dZ_t - \sqrt{V(t)}\sigma_{j+1}(t)dt], \\
  dV(t) &= \kappa(\theta - V(t))dt + \epsilon \sqrt{V(t)}dW_t, 
\end{align*}
\]  \hspace{1cm} (5.6)

where \( \kappa, \theta \) and \( \epsilon \) are state-independent variables (\( \epsilon \) is not necessarily a small number), \( W_t \) is an additional 1-D Brownian motion under the risk-neutral measure. Note that in (5.6) \( \sqrt{V}\gamma_j \) and \( \sqrt{V}\sigma_{j+1} \) take over the roles of \( \gamma_j \) and \( \sigma_{j+1} \) in (4.4), and they also satisfy (4.5) in place of \( \gamma_j \) and \( \sigma_{j+1} \), respectively. Moreover, the forward rates and the stochastic factor can be correlated:

\[
E^Q \left[ \left( \frac{\gamma_j(t)}{\|\gamma_j(t)\|} \right) \cdot dZ_t \right] = \rho_j(t), \quad \text{with} \quad |\rho_j(t)| \leq 1. \hspace{1cm} (5.7)
\]

Here, \( (\gamma_j(t)/\|\gamma_j(t)\|) \cdot dZ_t \) is the single normalized Brownian motion for forward term rate \( f_j(t) \).

In formalism, the above extended market model appears like a special case of the extended Constant-Elasticity-Variance model (CEV) adopted by Andersen and Brotherton-Ratcliffe (2002) (for elasticity equal to constant one). But the two approaches suggest
completely different mechanism for volatility smile/skew and have very different capacity for swaption pricing. Under the extended CEV model volatility skews of caplets are generated by taking non-unitary elasticity parameter, and the stochastic factor only contributes extra convexity (or curvature) to the skews. Our lognormal extension (5.6, 5.7), on the other hand, naturally generates different patterns of volatility smiles or skews by adjusting the correlations (which are time-dependent functions) as well as the volatility of volatility factor ($\epsilon$). On swaption pricing, the extended CEV model requires a uniform elasticity parameter for all forward rate processes. Such a requirement is often not met by a CEV model calibrated to market prices of caplets, as smiles and skews, seen in Figure 4.1 and Figure 4.2 for instance, normally lead to different elasticity parameters for forward rates. The extended lognormal model, meanwhile, prices caplets and swaptions in the same framework without any restriction on the correlations. The numerical option valuation procedures for the two approaches are also drastically different. The extended CEV model uses asymptotic expansion, whenever possible, while the extended lognormal model can go beyond that and take advantage of fast Fourier transformation.

5.3 Caplet and swaption pricing

We now consider the pricing of caplets on $f_j(t)$ under the extended LIBOR model (5.6). A caplet is a call option on the forward rate. Assume the notional value of the caplet is one dollar, then its payoff is

$$\Delta T_j(f_j(T_j) - K)^+ \triangleq \Delta T_j \max\{f_j(T_j) - K, 0\},$$

where $\Delta T_j$ is the corresponding forward period for the forward term rate $f_j(t)$. To price the caplet we choose, in particular, $P(t, T_{j+1})$ to be the numeraire, and let $Q^{j+1}$ denote
the forward measure under which \( f_j(t) \) is a martingale. The next proposition states the relation between the Brownian motions between the risk-neutral and the forward measures.

**Proposition 5.3.1** Let \( Z_i \) and \( W_i \) be Brownian motions under \( Q \), then \( Z_i^{j+1} \) and \( W_i^{j+1} \), defined by

\[
\begin{align*}
\frac{dZ_i^{j+1}}{dZ_i} &= \sqrt{V(t)} \sigma_{j+1}(t) dt, \\
\frac{dW_i^{j+1}}{dW_i} &= \xi_j(t) \sqrt{V(t)} dt,
\end{align*}
\]

are Brownian motions under \( Q^{j+1} \), where

\[
\xi_j(t) = \sum_{k=1}^{j} \frac{\Delta T_k f_k(t) \rho_k(t) \|\gamma_k(t)\|}{1 + \Delta T_k f_k}.
\]

Now in terms of \( Z_i^{j+1} \) and \( W_i^{j+1} \), our extended market model becomes

\[
\begin{align*}
df_j(t) &= f_j(t) \sqrt{V(t)} \gamma_j(t) \cdot dZ_i^{j+1}, \\
dV(t) &= [\kappa \theta - (\kappa + \epsilon \xi_j(t)) V(t)] dt + \epsilon \sqrt{V(t)} dW_i^{j+1}.
\end{align*}
\]

Note that \( \xi_j(t) \) depends on the forward rates, which is prohibitive for analytical tractability. To remove the dependence, we propose the approximation of \( \xi_j(t) \) by "frozen coefficients":

\[
\xi_j(t) \approx \sum_{k=1}^{j} \frac{\Delta T_k f_k(0) \rho_k(t) \|\gamma_k(t)\|}{1 + \Delta T_k f_k(0)},
\]

which can be regarded as the leading non-stochastic term of \( \xi_j(t) \). For clarity we further denote

\[
\tilde{\xi}_j(t) = 1 + \frac{\epsilon}{\kappa} \xi_j(t),
\]

and thus retain a neat equation for the process of \( V(t) \):

\[
dV(t) = \kappa [\theta - \tilde{\xi}_j(t) V(t)] dt + \epsilon \sqrt{V(t)} dW_i^{j+1}.
\]
Under forward measure $Q^{i+1}$ the forward price of the caplet must be a martingale and we thus have the following expression of caplet price

\[
Caplet(0) = P(0,T_{j+1})\Delta T_j E_0^{Q^{i+1}}[(f_j(T_j) - K)^+]
\]

\[
= P(0,T_{j+1})\Delta T_j f_j(0)G(0, f_j(0), V(0), K),
\]

where

\[
G(0, f_j(0), V(0), K) \triangleq E_0^{Q^{i+1}}[(f_j(T_j)/f_j(0) - K/f_j(0))^+]
\]

\[
= E_0^{Q^{i+1}}[e^{\ln(f_j(T_j)/f_j(0))}1_{f_j(T_j) > K}] - \frac{K}{f_j(0)} E_0^{Q^{i+1}}[1_{f_j(T_j) > K}].
\]

(5.9)

The two expectations above can be evaluated by moment generating function of the forward rate. The moment generating function is defined by

\[
\Phi(X(t), V(t), t; z) \triangleq E[\exp(zX(T_j)) | \mathcal{F}_t], \quad z \in C,
\]

where $X(t) = \ln (f_j(t)/f_j(0))$. We let $\Phi_T(z) \triangleq \Phi(0, V(0), 0; z)$ for simplicity. When $z$ is an imaginary number, $\Phi_T(z)$ becomes the characteristic function of $X(T_j)$. From the definition of a characteristic function, one can derive that (Kendall (1994))

\[
E_0^{Q^{i+1}}[1_{f_j(T_j) > K}] = \frac{\Phi_T(0)}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\text{Im}\{e^{-iu\ln(K/f_j(0))}\Phi_T(1+iu)\}}{u} du,
\]

\[
E_0^{Q^{i+1}}[e^{X(T_j)}1_{f_j(T_j) > K}] = \frac{\Phi_T(1)}{2} + \frac{1}{\pi} \int_0^{\infty} \frac{\text{Im}\{e^{-iu\ln(K/f_j(0))}\Phi_T(1+iu)\}}{u} du.
\]

Literally, what we need to do then is to get function $\Phi_T(z)$ and then perform numerical integrations. As Heston (1993) might be the first to price equity options by numerically evaluating these integrals, we call such option valuation procedure the Heston’s method for later references. When the Brownian motions $Z_t^{j+1}$ and $W_t^{i+1}$ are independent, we can work out the moment generating function directly. In general, $\Phi(x, V, t; z)$ satisfies the Komogrov backward equation corresponding to the joint forward rate and stochastic
factor process:
\[
\frac{\partial \Phi}{\partial t} + \kappa(\theta - \xi_j V) \frac{\partial \Phi}{\partial V} - \frac{1}{2} \|\gamma_j(t)\|^2 V \frac{\partial \Phi}{\partial x} \\
+ \frac{1}{2} \epsilon^2 V \frac{\partial^2 \Phi}{\partial V^2} + \epsilon \rho_j V \|\gamma_j(t)\| \frac{\partial^2 \Phi}{\partial V \partial x} + \frac{1}{2} \|\gamma_j(t)\|^2 V \frac{\partial^2 \Phi}{\partial x^2} = 0,
\]
(5.10)
with terminal condition
\[
\Phi(x, V, T_j; z) = e^{xz}.
\]
(5.11)
As we shall show, the above equation admits analytical solution provided that the coefficients are piecewise constants of \(t\). Otherwise we can at least solve \(\Phi(x, V, t; z)\) numerically. We defer solving for \(\Phi(x, V, t; z)\) until we have finished the discussion of swaption pricing, for there will be much in common in solving for the moment generating functions of forward rate and forward swap rate distributions.

Next we’ll discuss how to price swaptions. The equilibrium swap rate for the period \((T_m, T_n)\) is defined by
\[
R_{m,n}(t) = \frac{P(t, T_m) - P(t, T_n)}{B^S(t)},
\]
where
\[
B^S(t) = \sum_{j=m}^{n-1} \Delta T_j P(t, T_{j+1})
\]
is an annuity. The payoff of a swaption at \(T_m\) can be expressed as
\[
B^S(T_m) \cdot \max(R_{m,n}(T_m) - K, 0),
\]
where \(K\) is the strike rate.

The swap rate can be regarded as the price of a traded portfolio (consists of a long \(T_m\)-maturity and short \(T_n\)-maturity zero-coupon bonds) measured by the annuity \(B^S(t)\). According to the No Arbitrage Pricing theory (Harrison and Krep (1979)), the swap rate must be a martingale under the measure corresponding to the numeraire \(B^S(t)\). This measure is called the forward swap measure (Jamshidian (1997)) and is denoted by \(Q^S\)
in this chapter. Similar to the case of forward measure, we need to characterize the Brownian motions under the forward swap measure.

**Proposition 5.3.2** Let $Z_t$ and $W_t$ be Brownian motions under $\mathcal{Q}$, then $Z_t^S$ and $W_t^S$, defined by

$$
\begin{align*}
    dZ_t^S &= dZ_t - \sqrt{V(t)}\sigma^S(t)dt, \\
    dW_t^S &= dW_t + \sqrt{V(t)}\xi^S(t)dt,
\end{align*}
$$

are Brownian motions under forward swap measure $\mathcal{Q}^S$, where

$$
\begin{align*}
    \sigma^S(t) &= \sum_{j=m}^{n-1} \alpha_j \sigma(t, T_{j+1}), & \xi^S(t) &= \sum_{j=m}^{n-1} \alpha_j \xi_j, \\
\end{align*}
$$

with weights

$$
\alpha_j = \alpha_j(t) = \frac{\Delta T_j P(t, T_{j+1})}{B^S(t)}.
$$

It’s easy to show that, using Ito’s lemma, the swap rate follows

$$
\begin{align*}
    dR_{m,n}(t) &= \sqrt{V(t)} \sum_{j=m}^{n-1} \frac{\partial R_{m,n}(t)}{\partial f_j(t)} f_j(t) \tau_j(t) \cdot dZ^S(t), \\
    dV(t) &= \kappa [\theta - \bar{\xi}^S(t)V(t)]dt + \epsilon \sqrt{V(t)}dW^S(t),
\end{align*}
$$

where

$$
\bar{\xi}^S(t) = 1 + \frac{\epsilon}{\kappa} \xi^S(t).
$$

For the partial derivatives of the swap rate with respect to forward rates, we have

**Proposition 5.3.3** By the definition of forward swap rate,

$$
R_{m,n} = \frac{P(t, T_m) - P(t, T_n)}{B^S(t)},
$$

then we have

$$
\frac{\partial R_{m,n}(t)}{\partial f_j(t)} = \frac{\Delta T_j R_{m,n}}{1 + \Delta T_j f_j} \left[ \frac{P(t, T_n)}{P(t, T_m) - P(t, T_n)} + \sum_{k=j}^{n-1} \Delta T_k P(t, T_{k+1}) \right],
$$

where $m \leq j \leq n - 1$. 

61
Now under the forward swap measure we have the following expression for the swaption price

\[ PS(0) = B^S(0)E_0^S[(R_{m,n}(T_m) - K)^+] \]
\[ = B^S(0)(E_0^S[1_{R_{m,n}(T_m) > K}] - K E_0^S[1_{R_{m,n}(T_m) > K}]) \]
\[ = B^S(0)R_{m,n}(0)E_0^S[e^{\ln(R_{m,n}(T_m)/R_{m,n}(0))}1_{R_{m,n}(T_m) > K}] \]
\[ - B^S(0)KE_0^S[1_{R_{m,n}(T_m) > K}], \quad (5.15) \]

where \( E_0^S[\cdot] \) stands for the expectation under the forward swap measure conditioned on the filtration at time \( t = 0 \). Literally, \( R_{m,n}(t) \) is not lognormal distributed and an exact valuation of (5.15) is impossible. Hence, we consider the following lognormal approximations of swap-rate process instead

\[ dR_{m,n}(t) = R_{m,n}(t)\sqrt{V(t)} \sum_{j=m}^{n-1} \omega_j(0)\gamma_j(t) \cdot dZ^S(t), \quad 0 \leq t < T_m, \]
\[ dV(t) = \kappa[\theta - \xi^S_0(t)V(t)]dt + \epsilon \sqrt{V(t)}dW^S(t), \quad (5.16) \]

where

\[ \omega_j(0) = \frac{\partial R_{m,n}(0)}{\partial f_j(0)} \frac{f_j(0)}{R_{m,n}(0)}, \]
\[ \xi^S_0(t) = 1 + \frac{\epsilon}{\kappa} \xi^S(t) = 1 + \frac{\epsilon}{\kappa} \sum_{j=m}^{n-1} \alpha_j(0)\xi_j(t). \]

Coefficients in (5.16) are now non-stochastic. Equations (5.16) can be regarded as the lognormal approximation of swap-rate dynamics under the model of stochastic volatility. Such approximation was taken by Anderson and Andreasen (1998) for the deterministic volatility model. The approximations are supported by the fact that the variations of \( \omega_j(t) \) and \( \alpha_j(t) \) over time are negligible compared with those of the forward rates. With the lognormal approximation, the "closed-form" solution of swaption follows analogously to the formula for caplet. Again we let

\[ \Phi(X(t), V(t), t; z) = E^S[e^{zX(T)}|\mathcal{F}_t], \quad z \in C, \]
be the moment generating function for \( X(t) = \ln \left( \frac{R_{m,n}(t)}{R_{m,n}(0)} \right) \), and define \( \Phi_T(z) = \Phi(0, V(0), 0; z) \), we then have

\[
E_0^S[\mathbf{1}_{R_{m,n}(T_m) > K}] = \frac{\Phi_T(0)}{2} + \frac{1}{\pi} \int_0^\infty \frac{\text{Im} \left\{ e^{-i u \ln \left( \frac{K}{R_{m,n}(0)} \right)} \Phi_T(iu) \right\}}{u} \, du,
\]

\[
E_0^S[e^{X(T_m)} \mathbf{1}_{R_{m,n}(T_m) > K}] = \frac{\Phi_T(1)}{2} + \frac{1}{\pi} \int_0^\infty \frac{\text{Im} \left\{ e^{-i u \ln \left( \frac{K}{R_{m,n}(0)} \right)} \Phi_T(1 + iu) \right\}}{u} \, du.
\]

Function \( \Phi(x, V, t; z) \) satisfies the equation and terminal condition similar to those for forward rates:

\[
\frac{\partial \Phi}{\partial t} + \kappa(\theta - \frac{\zeta V}{\sigma}) \frac{\partial \Phi}{\partial V} - \frac{1}{2} \gamma_{m,n}(t) \| \frac{\partial^2 \Phi}{\partial x^2} \| V \frac{\partial \Phi}{\partial x} + \frac{1}{2} \gamma_{m,n}(t) \| \frac{\partial^2 \Phi}{\partial x^2} + \frac{1}{2} \gamma_{m,n}(t) \| V \frac{\partial^2 \Phi}{\partial x^2} = 0, \quad (5.17)
\]

with terminal condition

\[
\Phi(x, V, T_m; z) = e^{x z},
\]

where

\[
\gamma_{m,n}(t) = \sum_{j=m}^{n-1} \omega_j(0) \gamma_j(t)
\]

is the volatility of the swap rate, and

\[
\rho^S(t) = \frac{\sum_{j=m}^{n-1} \omega_j(0) \| \gamma_j(t) \| \rho_j(t)}{\| \gamma_{m,n}(t) \|}
\]

is the correlation between the swap rate and the stochastic factor. Notice that \( \{\gamma_j\} \)'s appear in the expression of \( \rho^S(t) \), which is undesirable for calibration purpose. the function \( \rho^S(t) \) is an average of correlations \( \rho_j(t) \) with weights \( \{\omega_j(0) \| \gamma_j(t) \| / \| \gamma_{m,n}(t) \|\} \).

If we assume that the effect of averaging is about the same with that of \( \{\omega_j(0)\} \), then we can have a much simpler relation:

\[
\rho^S \approx \sum_{j=m}^{n-1} \omega_j \rho_j,
\]

i.e., we approximate the correlation between the swap rate and the stochastic factor by a weighted average of the correlations between forward rates and the stochastic factor.
Such approximation, as we will see, will be supported by our test examples. We want to emphasize here that when $n = m + 1, R_{m,m+1} = f_m(t)$, the corresponding swaption reduces to the caplet on $f_m(t)$, and all equations reduce to the corresponding ones in the earlier part of this section.

Pricing swaptions through evaluating the two integrals involving the swaption price formula may not be fast enough for production purpose. The need for frequent calibration of the model (to around-the-money caps and at-the-money swaptions of various maturities) requires faster valuation method. In section 5 we introduce an alternative formula to (5.15), which is also based on the moment generating function but can be implemented tremendously faster.

### 5.4 Solving for the moment generating functions

As was seen that the moment generating functions $\Phi(x, V, t; z)$ for both forward rates and forward swap rates satisfy the partial different equation of the form

$$
\frac{\partial \Phi}{\partial t} + \kappa(\theta - \xi V) \frac{\partial \Phi}{\partial V} - \frac{1}{2} \lambda^2 V \frac{\partial^2 \Phi}{\partial x^2} + \frac{1}{2} \epsilon^2 V \frac{\partial^2 \Phi}{\partial V^2} + \epsilon \rho \lambda V \frac{\partial^2 \Phi}{\partial V \partial x} + \frac{1}{2} \lambda^2 V \frac{\partial^2 \Phi}{\partial x^2} = 0, \quad (5.18)
$$

with the terminal condition

$$
\Phi(x, V, T; z) = e^{xz}. \quad (5.19)
$$

Here, $\rho^c$ and $\lambda$ takes different functions for caplet and swaption. For caplet on $f_j$,

$$
\rho^c = \rho_j, \quad \text{and} \quad \lambda = \|\gamma_j(t)\|.
$$

For swaption on $R_{m,n},$

$$
\rho^c = \rho^S; \quad \text{and} \quad \lambda = \|\gamma_{m,n}(t)\|.
$$

64
Notice that all coefficients are either zero order or first order in $V$, following Heston (1993), we consider solution of the form

$$
\tilde{\Phi}(x, V, \tau; z) = e^{A(\tau, z) + B(\tau, z)V + \frac{\kappa}{2} z^2} = \Phi(x, V, t; z),
$$

where $\tau = T - t$ is the time to maturity. Substituting the above form solution to (5.18) and (5.19), we obtain the equations for the undetermined coefficients:

$$
\begin{align*}
\frac{dA}{d\tau} &= \kappa \theta B, \\
\frac{dB}{d\tau} &= \frac{1}{2} \kappa \theta B^2 + (\rho^2 \epsilon \lambda z - \kappa \xi) B + \frac{1}{2} \lambda^2 (z^2 - z),
\end{align*}
$$

subject to initial conditions

$$
A(0, z) = 0, \quad B(0, z) = 0. \tag{5.21}
$$

The equation for $B$ is a Riccati equation, which is known to have no analytical solution for general coefficients. In our caplet/swaption pricing model, we will consider piece-wise constant volatility and correlation functions, which offers abundant capacity for calibrating to cross sectional prices and to correlation structure among the forward rates.

The next Proposition states the analytical solution for equation (5.20) and (5.21).

**Proposition 5.4.1** For piece-wise constant coefficients, equations (5.20) and (5.21) admit solutions of the form

$$
\begin{align*}
A(\tau, z) &= A(\tau_j, z) + \kappa \theta \left\{ (a + d)(\tau - \tau_j) - 2 \ln \left[ \frac{1 - g_j e^{\mu(\tau - \tau_j)}}{1 - g_j} \right] \right\}, \\
B(\tau, z) &= B(\tau_j, z) + \frac{(a + d - \epsilon^2 B(\tau_j, z))(1 - e^{\mu(\tau - \tau_j)})}{e^{\mu(\tau_j)} - 1} e^{-\mu(\tau - \tau_j)},
\end{align*}
$$

for $\tau_j \leq \tau < \tau_{j+1}, \quad j = 0, 1, \ldots, m - 1,\tag{5.22}
$$

where

$$
a = \kappa \xi - \rho^2 \epsilon \lambda z, \quad d = \sqrt{a^2 - \lambda^2 \epsilon^2 (z^2 - z)}, \quad g_j = \frac{a + d - \epsilon^2 B(\tau_j, z)}{a - d - \epsilon^2 B(\tau_j, z)}.
$$

After we have $A$ and $B$ in analytical form, we can obtain the moment generating function $\Phi_T(z) = \Phi(0, V, 0; z)$ in closed-form, hence we can proceed to evaluate the integrals included in the pricing formulae for caplet and swaption.
5.5 Numerical integration via fast fourier transform

The application of Fast Fourier Transform in option pricing was pioneered by Carr and Madan (1998), who discovered that the Fourier transform of an option price can be expressed explicitly in terms of the characteristic function of the underlying state variables. The option price, therefore, can be obtained by performing an inverse Fourier transform, which numerically can be achieved by Fast Fourier transform. The motivation of Carr and Madan was to evaluate equity options under the Variance Gamma (VG) process. In this section we generalize their ideas to the LIBOR model with stochastic volatility.

5.5.1 Fourier method for dampened option value

Let us illustrate the idea with the valuation of a caplet. Recall definition (5.9), we now regard the forward price of the option as a function of strike:

\[ G_T(k) \triangleq G(0, f_j(0), V(0), K) = E_0^{Q^{j+1}}[((f_j(T))/f_j(0) - K/f_j(0))^+], \tag{5.23} \]

where \( k = \ln (K/f_j(0)) \).

Let \( q_T(s) \) denote the density function of \( X(T) = \ln (f_j(T)/f_j(0)) \), then,

\[ G_T(k) = \int_k^\infty (e^s - e^k)q_T(s)ds. \tag{5.24} \]

Note that \( G_T(k) \) tends to 1 when \( k \) tends to \(-\infty \) and hence \( G_T(k) \) is not square integrable over \((-\infty, \infty)\). For this reason we consider the dampened value of the caplet defined by

\[ g_T(k) = e^{ak}G_T(k), \quad \text{for some constant } a > 0. \tag{5.25} \]

The dampened value then is square integrable over the entire real line. The Fourier transform with the dampened value is given by

\[ \psi_T(u) = \int_{-\infty}^{\infty} e^{iku}g_T(k)dk \]
\begin{align*}
&= \int_{-\infty}^{\infty} e^{iuk} \int_{-\infty}^{\infty} e^{sk}(e^s - e^k) q_T(s) ds dk \\
&= \int_{-\infty}^{\infty} q_T(s) \int_{-\infty}^{\infty} (e^{s+ak} - e^{(1+a)k}) e^{iuk} dk ds \\
&= \int_{-\infty}^{\infty} q_T(s) \left[ \frac{e^{(a+1+iu)s}}{a + iu} - \frac{e^{(a+1+iu)s}}{a + 1 + iu} \right] ds \\
&= \frac{\Phi_T(1 + a + iu)}{(a + iu)(1 + a + iu)}.
\end{align*}

The caplet price follows from the inverse Fourier transform

\[ G_T(k) = e^{-ak} g_T(k) = \frac{e^{-ak}}{\pi} \int_0^\infty e^{-iuk} \tilde{\psi}(u) du. \] (5.27)

### 5.5.2 Fourier method for intrinsic value of call/put options

Instead of the dampened option values, we can also consider the time value of the call option

\[ z_T(k) \triangleq G_T(k) - (1 - K/f_i(0))^+. \]

It’s not difficult to derive

\begin{align*}
    z_T(k) &= \int_{-\infty}^{\infty} (e^s - e^k) 1_{s>k} q_T(s) ds - \int_{-\infty}^{\infty} (1 - e^k) 1_{k<0} q_T(s) ds \\
    &= \int_{-\infty}^{\infty} (e^s - e^k)(1 - 1_{s<k}) 1_{k<0} q_T(s) ds \\
    &\quad + \int_{-\infty}^{\infty} (e^s - e^k) 1_{s>k} 1_{k>0} q_T(s) ds - \int_{-\infty}^{\infty} (1 - e^k) 1_{k<0} q_T(s) ds \\
    &= \int_{-\infty}^{\infty} [(e^k - e^s) 1_{s<k<0} + (e^s - e^k) 1_{s>k>0}] q_T(s) ds.
\end{align*}

Note the last equality holds for the sake of martingale property under forward measure.

Since the time value is square integrable, so we can take Fourier transform on \( z_T(k) \),

\[ \eta_T(u) = \int_{-\infty}^{\infty} e^{iuk} z_T(k) dk. \]

\[ = \int_0^\infty dke^{iuk} \int_{-\infty}^{\infty} (e^k - e^s) q_T(s) ds + \int_{0}^{\infty} dke^{iuk} \int_{k}^{\infty} (e^s - e^k) q_T(s) ds. \]
\[
= \int_{-\infty}^{0} q_T(s) ds \left[ \int_{s}^{0} (e^{(1+iu)k} - e^{iu_k+s}) dk + \int_{0}^{s} (e^{s+iu_k} - e^{(1+iu)k}) dk \right] \\
= \frac{1 - \Phi_T(1 + iu)}{u^2 - iu}.
\]

Note that from the martingale property of \( X(T) \) we know that \( \Phi_T(0) = 1 \) and \( u = 0 \) is a removable singularity of \( \eta_T(u) \). Having \( \eta_T(u) \) in explicit form, we perform inverse Fourier transform and thus obtain the option value

\[
G_T(k) = (1 - K/f_j(0))^+ + \frac{1}{\pi} \int_{0}^{\infty} e^{-iuk} \eta_T(u) du.
\]

The above integral will be evaluated numerically, and for that sake we must truncate the domain for integration. Again by the martingale property of \( X(t) \) we have

\[
|\Phi_T(1 + iu)| = \left| E^Q_{0} \left[ e^{(1+i)X(T_j)} \right] \right| \\
\leq E^Q_{0} \left[ \left| e^{X(T_j)} e^{iuX(T_j)} \right| \right] \\
= E^Q_{0} \left[ e^{X(T_j)} \right] = 1.
\] (5.28)

It follows that

\[
|\eta_T(u)| = \left| \frac{1 - \Phi_T(1 + iu)}{u^2 - iu} \right| \\
\leq \left| \frac{2}{\sqrt{u^4 + u^2}} \right| < \frac{2}{u^2},
\] (5.29)

and

\[
\left| \int_{A}^{\infty} e^{-iuk} \eta_T(u) du \right| \leq \int_{A}^{\infty} \frac{2}{u^2} du = \frac{2}{A}.
\] (5.30)

Therefore, to numerically evaluate the integral we can truncate the integral at \( A = 10^4 \), which will ensure the accuracy of option prices to one basis point. We will see that such large truncation is too conservative, and in practice we can take much smaller \( A \).

There are two ways to further reduce that range \( A \) for numerical integration. The first one is again introduced by Carr and Madan (1998). Notice that \( z_T(k) \) is

68
not differentiable at $k = 0$, and such non-smoothness will generate high modes in the frequency space. This can be visualized in Figure 5.1, where the solid line stands for the Fourier transform of $z_T(k)$, which dies down very slowly away from origin. As a result we need to take sizable $A$ for numerical integration, which then undermine the efficiency of the model implementation. The remedy to such problem is to smooth our time value at, in particular, $k = 0$. For this purpose we take a sinusoidal function as the multiplicative factor: $\sinh(\alpha k)z_T(k)$, which has first order derivative with respect to $k$ everywhere and thus is smoother than $z_T(k)$. The Fourier transform of the smoothed out function is

$$\zeta_T(u) = \int_{-\infty}^{\infty} e^{iku} \sinh(\alpha k)z_T(k)dk = \frac{\eta_T(u - i\alpha) - \eta_T(u + i\alpha)}{2}.$$

It is represented by the dash line in Figure 5.1. From this figure we can see that the high modes dies down much more quickly than that of $\eta_T(u)$. Using $\psi_T(u)$ we have the following evaluation formula for the caplet

$$G_T(k) = (1 - K/f_j(0))^+ + \frac{1}{\sinh(\alpha k)} \frac{1}{\pi} \int_0^\infty e^{-iku} \zeta_T(u)du. \quad (5.31)$$

There is however an apparent disadvantage to the above treatment: the expression (5.31) is invalid for $k = 0$, which corresponds to at-the-money options. Interpolation may be applied to estimate the values of ATM options, but accuracy then becomes a concern.

Note that no approximation is ever made in the treatments above. The second approach, however, will involve approximation with error under control. For instance, take the Gaussian function

$$g(k) = \frac{1}{\alpha \sqrt{2\pi}} e^{-\frac{k^2}{2\alpha^2}}, \quad \alpha > 0, \quad (5.32)$$

and consider the convolution

$$z_g(k) = \int_{-\infty}^{\infty} g(k - y)z_T(y)dy. \quad (5.33)$$
$z_g(k)$ is an infinitely differentiable function and it converges to $z_T(k)$ when $\alpha \to 0^+$. For $\alpha \ll 1$, $z_g(k)$ and $z_T(k)$ are almost indistinguishable except at $k = 0$, where $z_T(k)$ is non-differentiable. We intend to replace $z_T(k)$ by $z_g(k)$. The consequence is that we may sacrifice some accuracy at $k = 0$. Yet when $\alpha \to 0$, the accuracy will improve. The Fourier transform of $z_g(k)$ relates to that of $z_T(k)$ by

$$\eta_g(u) = e^{-\frac{1}{2}u^2} \eta_T(u). \quad (5.34)$$

Apparently, for $\alpha > 0$, $\eta_g(u)$ reduces exponentially fast away from $u = 0$. Since $\eta_T(u)$ is bounded by $\frac{4}{u^2}$, we thus have

$$\left| \frac{1}{\pi} \int_A^\infty e^{-iuk} \eta_g(u) du \right| \leq \frac{2}{\pi} \int_A^\infty \frac{e^{-\frac{1}{2}u^2} u}{u^2} du \leq \frac{2 e^{-\frac{1}{2}a^2A^2}}{\pi A}. \quad (5.35)$$
Given accuracy level and \( \alpha \), we can determine \( A \) from the above bound. Suppose we take \( \alpha = 0.005 \), in order to achieve one basis point accuracy, we can take \( A = 500 \).

The same treatments apply to the valuation of swaption as well.

Finally we consider the valuation of the inverse Fourier transform in (5.31). For this purpose, we truncate that interval for an appropriately chosen number \( A \), and consider that composite trapezoidal rule for the numerical quadrature of the integral:

\[
H(k) = \frac{1}{\pi} \left( \frac{\zeta_T(0)}{2} + \sum_{m=1}^{N-1} e^{-iu_m^N k} \zeta_T(u_m) + \frac{e^{-iu_N^N k} \zeta_T(u_N)}{2} \right) \Delta u, \tag{5.36}
\]

where \( u_m = m \Delta u \) and \( \Delta u = A/N \). The composite trapezoidal rule has the order of accuracy of \( O(\Delta u^2) \). Since we are mainly interested in the around-the-money options, we take \( k \)'s around zero:

\[
k_n = -b + n \Delta k, \quad \text{for some } b > 0 \text{ and } n = 0, 1, \ldots, N - 1,
\]

with

\[
\Delta k = \frac{2b}{N}.
\]

Hence

\[
H(k_n) = \frac{\Delta u}{\pi} \left( \frac{\zeta_T(0)}{2} + \sum_{m=1}^{N-1} e^{-i \Delta u \Delta k N n} e^{i b u_m} \zeta_T(u_m) + \frac{e^{-i \Delta u \Delta k N} e^{i b u_N} \zeta_T(u_N)}{2} \right).
\]

We now choose in particular

\[
\Delta u \Delta k = \frac{2\pi}{N}, \quad \text{or} \quad b = \frac{\pi N}{A},
\]

we then end up with

\[
H(k_n) = \frac{1}{\pi} \left( \frac{\zeta_T(0) + e^{i b u_N} \zeta_T(u_N)}{2} + \sum_{m=1}^{N-1} e^{-i \Delta k N n} e^{i b u_m} \zeta_T(u_m) \right) \Delta u,
\]

for \( n = 0, 1, \ldots, N - 1 \).
The above summation fits the definition of discrete Fourier transform and can be implemented with Fast Fourier Transform technology (FFT). For convenience, we will hereafter call the method developed in this thesis the FFT method.

One can also use more accurate quadrature formula like composite Simpson’s formula, but we have found the composite trapezoidal rule is efficient enough for practical use.

5.6 Numerical results

In this section we will show the performance of the FFT method for option valuation under the LIBOR model with stochastic volatilities. In specific, we take the example from Andersen and Andreasen (2000), and compute caplet and swaption prices using FFT method, the Heston’s method, and Monte Carlo simulation method. Since there is no exact or analytical solution available to our problem, the prices produced by the Monte Carlo simulation with a big number of paths will serve as ”exact solutions”. Due to the approximations made on the forward rate and particular forward swap rate processes after their respective change of measures, deviation of prices is expected, and accuracy (or extent of deviation) actually determines the viability of the approximate models. Note that theoretically, given identical grid points for numerical quadrature, the Fourier method and Heston’s method should generate identical solutions, because they are alternative formulae based on the same characteristic function. We will also examine how the shape of volatility smiles/skews changes in response to the change of correlation between forward rates and the stochastic volatility. For brevity we omit the results on caplets, which are special yet simpler cases of swaptions, and can be priced with similar accuracy.

Example(Andersen and Andreasen (2000)): The initial term structures of the
interest rates and stochastic volatility, and the volatility term-structure of the forward rates are given below.

Forward curve: \( \Delta T_j = 0.5, \ f_j(0) = 0.04 + 0.00075j, \) for all \( j \)

Stochastic volatility dynamics: \( V(0) = \theta = \kappa = 1 \) and \( \epsilon = 1.5 \)

Volatility term structure of a two-factor model

\[
\gamma_j(T_k) = \frac{1}{\sqrt{0.04 + 0.00075j}}(0.015 + 0.025e^{-0.05(j-k)}, 0.01 - 0.05e^{-0.1(j-k)}).
\]

Table 5.1 and Table 5.2 list the values of the at-the-money swaptions (for various maturities and tenors) obtained using Monte Carlo simulations (MC), the FFT method, and the Heston’s method. In specific, Table 5.1 lists the prices by the three methods for the non-correlation case (\( \rho = 0 \)), while Table 5.2 lists the prices for the negative correlation case (\( \rho = -0.5 \)). The swaption prices are given in basis points (bps), corresponding to cents for the notional value of 100 dollars. The percentages in parentheses are the relative difference from the corresponding MC prices. Note that we have omitted the prices by the Heston’s method in Table 5.2, because, as we can see from Table 5.1, the FFT and Heston’s methods produce identical results. In Table 5.3, we use intrinsic value approach combined with convolution approximation to work out the ATM swaption prices, we choose \( \alpha = 0.001 \) and \( A = 500 \), the errors look a little bit larger than those we get in Table 5.2, which is because we may sacrifice some accuracy for around the money swaption price when \( \alpha \) is not small enough, as \( \alpha \to 0 \), the errors will be improved.

The Monte Carlo simulations are performed under spot measure (the risk neutral measure) with time-step size \( \Delta t = 1/8 \) and number of paths \( NP = 50000 \). To build in the correlation between the forward rates and the stochastic factor we recast the equation for the forward rates into

\[
\frac{df_j(t)}{f_j(t)} = -V(t)\gamma_j(t) \cdot \sigma(t, T_{j+1})dt + \sqrt{V(t)}(\sqrt{1 - \rho^2}\gamma_j(t) \cdot d\tilde{Z}_t + \rho\gamma_j(t)) \cdot dW_t, \quad (5.37)
\]
where \((\tilde{Z}_t, W_t)\) is a vector of independent Brownian motions. Treated like a lognormal process, \(f_j(t)\) is advanced by the so-called log-Euler scheme:

\[
f_j(t + \Delta t) = f_j(t) e^{-\lambda_j(t, \sigma_j) \Delta t + \frac{1}{2} \lambda_j(t, \sigma_j)^2 \Delta t} + \sqrt{\lambda_j(t, \sigma_j)^2 \Delta t + \lambda_j(t, \sigma_j)^2 \Delta W_t},
\]

**Proposition 5.6.1** The evolution of volatility in above example can be approximated by a step-wise moment-matched lognormal process:

\[
V(t + \Delta t) = \mathbb{E}^Q[V(t + \Delta t)] e^{-\frac{1}{2} \Gamma_t^2 \Delta t + \Gamma_t \Delta W_t},
\]

where

\[
\Gamma_t^2 = \frac{1}{\Delta t} \ln \frac{\mathbb{E}^Q[V^2(t + \Delta t)]}{(\mathbb{E}^Q[V(t + \Delta t)])^2},
\]

with

\[
\begin{align*}
\mathbb{E}^Q[V(t + \Delta t)] &= \theta + (V(t) - \theta)e^{-\kappa \Delta t}, \\
\mathbb{E}^Q[V^2(t + \Delta t)] &= (1 + \frac{e^2}{2\kappa \theta})(\mathbb{E}^Q[V(t + \Delta t)])^2 - \frac{e^2}{2\kappa \theta} V^2(t) e^{-2\kappa \Delta t}.
\end{align*}
\]

The derivation of (5.40) is given in Appendix B.

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Tenor</th>
<th>ATM Swap rate</th>
<th>MC (bps)</th>
<th>FFT (diff.)</th>
<th>Heston (diff.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4.19%</td>
<td>35.65</td>
<td>35.36 (0.81%)</td>
<td>35.36 (0.81%)</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>4.47%</td>
<td>134.16</td>
<td>133.01 (0.86%)</td>
<td>133.01 (0.86%)</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>4.80%</td>
<td>204.60</td>
<td>202.75 (0.90%)</td>
<td>202.75 (0.90%)</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>4.79%</td>
<td>57.62</td>
<td>57.29 (0.57%)</td>
<td>57.29 (0.57%)</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5.07%</td>
<td>223.59</td>
<td>222.86 (0.33%)</td>
<td>222.86 (0.33%)</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>5.40%</td>
<td>350.12</td>
<td>349.96 (0.05%)</td>
<td>349.96 (0.05%)</td>
</tr>
</tbody>
</table>

Table 5.1: Swaption prices for the no-correlation case, \(\rho = 0\)

74
<table>
<thead>
<tr>
<th>Maturity</th>
<th>Tenor</th>
<th>ATM Swap rate</th>
<th>MC (bps)</th>
<th>FFT (diff.) $N = 50$</th>
<th>FFT (diff.) $N = 100$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4.19%</td>
<td>34.96</td>
<td>34.73 (0.66%)</td>
<td>34.75 (0.60%)</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>4.47%</td>
<td>132.50</td>
<td>131.49 (0.76%)</td>
<td>131.49 (0.76%)</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>4.80%</td>
<td>203.14</td>
<td>201.34 (0.89%)</td>
<td>201.36 (0.88%)</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>4.79%</td>
<td>56.12</td>
<td>55.95 (0.30%)</td>
<td>55.94 (0.32%)</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5.07%</td>
<td>220.92</td>
<td>219.59 (0.60%)</td>
<td>219.59 (0.60%)</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>4.19%</td>
<td>348.97</td>
<td>347.37 (0.46%)</td>
<td>347.35 (0.46%)</td>
</tr>
</tbody>
</table>

Table 5.2: Swaption prices for the negative correlation case, $\rho = -0.5$

<table>
<thead>
<tr>
<th>Maturity</th>
<th>Tenor</th>
<th>ATM Swap rate</th>
<th>MC (bps)</th>
<th>Intrinsic Convolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4.19%</td>
<td>34.96</td>
<td>34.49 (1.34%)</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>4.47%</td>
<td>132.50</td>
<td>130.11 (1.80%)</td>
</tr>
<tr>
<td>1</td>
<td>10</td>
<td>4.80%</td>
<td>203.14</td>
<td>198.69 (2.19%)</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>4.79%</td>
<td>56.12</td>
<td>56.13 (0.02%)</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>5.07%</td>
<td>220.92</td>
<td>218.93 (0.90%)</td>
</tr>
<tr>
<td>5</td>
<td>10</td>
<td>4.19%</td>
<td>348.97</td>
<td>345.30 (1.05%)</td>
</tr>
</tbody>
</table>

Table 5.3: Swaption prices for the negative correlation case, $\rho = -0.5$

For the integration methods, we have taken $A=50$, $N=50$ and $N=100$. It can be seen that for various maturities and tenors, the price differences between the Monte Carlo method and FFT method are within one percent, while the prices from FFT and the Heston’s method are identical for almost all cases. Such accuracy strongly support the approximations taken in this chapter for the swap rate processes. There is however, some slight downward bias by the integration methods. Also notice that the price differ-
ences between $\rho = 0$ and $\rho = -0.5$ are not significant, which is perfectly acceptable as has been well known that a stochastic volatility model "tilts" the volatility curve around the at-the-money strike, which can be seen in Figure 4.4 and 4.5. The choice of $A = 50$ for truncating the integrals is suggested by the pattern of decay of $\eta(u)$ shown in Figure 5.2 and 5.3 for the in-1-to-1 swaption with notional value equal to one dollar, which show that $\eta(u)$ is a smooth function whose magnitude is negligible beyond $u = \pm 50$. The Fourier transforms of other swaptions look similar.

![Real Part of \( \eta(u) \) for the in-1-to-1 swaption](image)

**Figure 5.2**: Real part of $\eta(u)$ for the in-1-to-1 swaption
(Data taken from the example)

It is also interesting to get a qualitative idea about the relation between prices produced by the stochastic and deterministic volatility models under the same variance for the state variables. If we take

$$RMV = \sqrt{\frac{VAR(X(T_m))}{T_m}} = \sqrt{1 \over T} [\Phi_r^2(0) - (\Phi_r^2(0))^2]$$

to be the volatility for the deterministic lognormal swap-rate process for $R_{m,n}(t)$, then variances of $R_{m,n}(T_m)$ under the two model are the same. Figures 5.4 and 5.5 show the
implied volatility curve of swaption prices across strikes, in contrast to the flat curve for \( RMV \). In this calculation the correlation for the stochastic volatility model is \( \rho = -0.5 \). It can be seen that the prices by the stochastic volatility model stay around the Black's price, while having a downward skew. Such pattern of relative price was already known for equity options.

Finally we study the role of correlation in the formation of volatility smiles or skews, by examining the variation of volatility smiles/skews with respect to correlation between the forward rates and the stochastic factor. Figure 5.6 shows the volatility smile/skew for caplet, where the downward sloping skew corresponds to a negative correlation \( \rho = -0.5 \), the upward sloping skew corresponds to a positive correlation \( \rho = 0.5 \), the nearly symmetric smile corresponds to no correlation \( \rho = 0 \). Similar correspondence exists in swaptions, which is depicted in Figure 5.7.

Note that negative correlation is plausible to many market practitioners. Not only it explains the downward skew that is most often seen in interest rate derivative
markets, but it also accounts for the mechanism of adverse movements of index and implied volatilities. That is, when interest rate goes up, the implied volatility goes down, and when interest rate goes down, the implied volatility goes up. These figures strongly suggest the role of stochastic volatility in the formation of volatility smiles and
Figure 5.6: Volatility smile and skews for one-year caplets
(Data taken from the example)

Figure 5.7: Volatility smile and skews for one-year swaptions
(Data taken from the example)

skews.
5.7 Summary

In this chapter we extend the market model by taking a multiplicative stochastic factor for forward rate volatilities. We take an approach alternative to Andersen and Andreasen (2002) by allowing correlation between the forward rates and the stochastic factor. By making appropriate approximations, we obtain the moment generating function of the joint forward rate and stochastic factor in closed-form. We then express the Fourier transform of option prices in terms of the moment generating function, and apply a FFT method to the numerical valuation of caplet and swaption prices. The correlation presented in our model largely explains the phenomena of volatility smiles and skews that are pronounced in the cap and swaption markets. The FFT method will facilitate the efficient calibration of the model to benchmark instruments, and this is a topic of further studies.
Chapter 6

Conclusion

In this thesis, we review the standard LIBOR market model which is based on the log-normal processes for forward rates and illustrate its merits and shortcomings. As we point out, in order to produce volatility skews and smiles, we have to extend this standard model by incorporating other features. Two different kind of extensions of LIBOR market model are presented in this thesis. Apart from stochastic volatility extension of the standard market model, we also show the calibration of CEV model and how to evaluate constant maturity instruments.

By applying change of measure and modified convexity adjustment techniques in Chapter 3, closed-form pricing formulae for various constant maturity instruments are obtained. According to the numerical examples, we can claim that with our method an arbitrary accuracy can be achieved, and in practice first few terms are already sufficient to obtain satisfactory results. Note that we assume constant maturity yield follows log-normal process, in future research, we may improve this work by allowing the yield to follow other processes (for example, CEV process).

As for the extension of standard model, from the implementation point of view, we use a practical problem to investigate the methodology to calibrate the CEV model,
which is capable of generating volatility skews. Follow a similar approach of the calibration of standard model, we first transform the market data into the implied CEV volatilities. The entire calibration problem has been solved by a powerful technique for general constrained optimization problem with quadratic objective function and constraints.

Beside the CEV approach, we also show another extension of LIBOR market model, this new extension includes stochastic volatility and is free of arbitrage, hence can produce not only skews but also non-monotonic volatility smiles. Approximate state-variable processes which carry analytical tractability are introduced, and with some appropriate approximations, closed-form solutions can be obtained, from the fast implementation viewpoint, we show that fast fourier transform technology can make our approach no longer computational expensive. The reasonableness of approximation is justified by the numerical examples. For the future study, we may construct a more complex model with stochastic volatility and jump, and the calibration of this model is another problem needed to be solved.
Appendix A

Arbitrage-free Model with Stochastic Volatility

A.1 Proof of Proposition 5.2.1

We define the Radon-Nikodym derivative
\[ \frac{dP^*}{dP} = \exp \left( -\frac{1}{2} \int_0^t (\phi^2 + \psi^2) ds - \int_0^t \phi dZ_s - \int_0^t \psi dW_s \right), \]

so under the new measure
\[ d\hat{Z}_t = dZ_t + \phi dt, \]
\[ d\hat{W}_t = dW_t + \psi dt, \]

and
\[ \int_0^t (\phi^2 + \psi^2) ds < \infty, \]

thus
\[ dV(t) = (\kappa \theta - \kappa V(t)) dt - \epsilon \sqrt{V(t)} \psi dt + \epsilon \sqrt{V(t)} d\hat{W}_t \]
\[ = \kappa (\hat{\theta} - V(t)) dt + \epsilon \sqrt{V(t)} d\hat{W}_t. \]

It’s easy to see
\[ V(t) = V(0) + \int_0^t \kappa (\hat{\theta} - V(s)) ds + \int_0^t \epsilon \sqrt{V(s)} d\hat{W}_s, \]
\[ f(t, T) = f(0, T) + \int_0^t \mu(s, T) ds + \int_0^t \sigma_f(s, T) \sqrt{V(s)} dZ_s. \]

Denote
\[ Y(t, T) = -\int_t^T f(t, u) du, \]
then
\[ Y(t, T) = - \int_t^T \left[ f(0, u) + \int_0^t \mu(s, u) ds + \int_0^t \sigma_f(s, u) \sqrt{V(s)} dZ_s \right] du. \]

For each term of above expression, after splitting up the integral and changing the order of integration, we get

\[
- \int_t^T f(0, u) du = - \int_0^T f(0, u) du + \int_0^t f(0, u) du
\]
\[= Y(0, T) + \int_0^t f(0, u) du, \]

\[
- \int_t^T \int_t^s \mu(s, u) ds du = - \int_0^t \int_t^s \mu(s, u) duds
\]
\[= - \int_0^t \int_s^T \mu(s, u) duds + \int_0^t \int_s^t \mu(s, u) duds
\]
\[= - \int_0^t \int_s^T \mu(s, u) duds + \int_0^t \int_0^u \mu(s, u) ds du, \]

\[
- \int_t^T \int_0^t \sigma_f(s, u) \sqrt{V(s)} dZ_s du
\]
\[= - \int_0^t \int_t^T \sigma_f(s, u) \sqrt{V(s)} dudZ_s
\]
\[= - \int_0^t \int_s^T \sigma_f(s, u) \sqrt{V(s)} dudZ_s + \int_0^t \int_0^t \sigma_f(s, u) \sqrt{V(s)} dudZ_s
\]
\[= - \int_0^t \int_s^T \sigma_f(s, u) \sqrt{V(s)} dudZ_s + \int_0^t \int_0^u \sigma_f(s, u) \sqrt{V(s)} dudZ_s du. \]

Consider the instantaneous interest rate
\[ r_u = f(u, u) \]
\[= f(0, u) + \int_0^u \mu(s, u) ds + \int_0^u \sigma_f(s, u) \sqrt{V(s)} dZ_s. \]

Therefore we get
\[ Y(t, T) = Y(0, T) + \int_0^t r_u du - \int_0^t \int_0^t \mu(s, u) duds
\]
\[- \int_0^t \int_s^T \sigma_f(s, u) \sqrt{V(s)} dudZ_s. \]
Hence
\[ dY(t, T) = \left(r_t + A(t, T)\right)dt + B(t, T)dZ_t, \]  
(A-1)
where
\[ A(t, T) = -\int_t^T \mu(t, u)du, \]
\[ B(t, T) = -\int_t^T \sigma_f(t, u)\sqrt{V(t)}du. \]

Since the zero coupon bond
\[ P(t, T) = e^{-\int_t^T f(t, s)ds} = e^{Y(t, T)}, \]

by Ito’s Lemma
\[ dP(t, T) = P(t, T)dY(t, T) + \frac{1}{2} P(t, T)B^2(t, T)dt, \]
thus
\[ \frac{dP(t, T)}{P(t, T)} = \left[ r_t + A(t, T) + \frac{1}{2} B^2(t, T) \right] dt + B(t, T)dZ_t. \]  
(A-2)

Note
\[ dZ_t = -\phi dt + d\hat{Z}_t, \]
then
\[ A(t, T) + \frac{1}{2} B^2(t, T) - \phi(t)B(t, T) = 0. \]

Differentiate it with respect to \( T \),
\[ \mu(t, T) = \frac{\partial}{\partial T} \left[ \frac{1}{2} \left( \int_t^T \sigma_f(t, u)\sqrt{V(t)}du \right)^2 \right] + \phi \sigma_f(t, T)\sqrt{V(t)}. \]

The proof is complete.
A.2 Proof of Theorem 5.2.1

By Proposition 5.2.1, under risk neutral measure, we have

$$\phi \equiv 0,$$

and for simplicity, we put

$$\psi = 0.$$  

So

$$\mu(t, T) = \frac{\partial}{\partial T} \left[ \frac{1}{2} \left( \int_t^T \sigma_f(t, u) \sqrt{V(t)} du \right)^2 \right],$$

and since

$$f_n = f(t, T_n, T_n + \Delta T_n) = \frac{1}{\Delta T_n} \left( e^{\int_{T_n}^{T_{n+1}} f(t, u) du} - 1 \right),$$

hence

$$df_n = \frac{1}{\Delta T_n} d\left( e^{\int_{T_n}^{T_{n+1}} f(t, u) du} \right) = \frac{1}{\Delta T_n} d\left( e^{Z(t, T_n)} \right),$$

where

$$Z(t, T_n) = \int_{T_n}^{T_{n+1}} f(t, u) du.$$  

Note

$$dZ(t, T_n) = \int_{T_n}^{T_{n+1}} df(t, u) du$$

$$= \left( \int_{T_n}^{T_{n+1}} \sigma_f(t, u) \sqrt{V(t)} du \right) dZ_t + \left( \int_{T_n}^{T_{n+1}} \frac{\partial}{\partial u} \left[ \frac{1}{2} \int_t^u \sigma_f(t, s) \sqrt{V(s)} ds \right]^2 du \right) dt$$

$$= \left( M_n - M_{n+1} \right) dZ_t + \frac{1}{2} \left( M_{n+1}^2 - M_n^2 \right) dt,$$

where

$$M_n = - \int_t^{T_n} \sigma(t, u) \sqrt{V(t)} du.$$  

So by Ito's Lemma

$$df_n = \frac{1}{\Delta T_n} e^{Z(t, T_n)} dZ(t, T_n) + \frac{1}{2\Delta T_n} e^{Z(t, T_n)} (M_n - M_{n+1})^2 dt$$

$$= \frac{1}{\Delta T_n} \left( 1 + \Delta T_n L_n \right) (M_n - M_{n+1})(dZ_t - M_{n+1} dt).$$
Set
\[ M_n - M_{n+1} = \frac{\Delta T_n \gamma_n \sqrt{V} f_n}{1 + \Delta T_n f_n}, \]
which is equivalent to
\[ \int_{T_n}^{T_{n+1}} \sigma_f(t, u) du = \frac{\Delta T_n \gamma_n f_n}{1 + \Delta T_n f_n}, \]
hence
\[ df_n = \gamma_n \sqrt{V} f_n (dZ_t - M_{n+1} dt). \]
Note
\[ -M_{n+1} = \sum_{k=\lambda(t)}^{n} (M_k - M_{k+1}) = \sum_{k=\lambda(t)}^{n} \frac{\Delta T_k \gamma_k \sqrt{V} f_k}{1 + \Delta T_k f_k} \]
as \( M_{\lambda(t)} = 0. \)
Finally we get
\[ \frac{df_n}{f_n} = \gamma_n \sqrt{V} dZ_t + \gamma_n \sum_{k=\lambda(t)}^{n} \frac{\Delta T_k \gamma_k V f_k}{1 + \Delta T_k f_k} dt. \]
Appendix B

Details of Some Derivations

Proof of Proposition 5.3.1:

The Radon-Nikodym derivative for $Q^{i+1}$ is

$$
\frac{dQ^{i+1}}{dQ} = \frac{P(t, T_{j+1})/P(0, T_{j+1})}{B(t)}
\begin{align*}
&= e^{\int_0^t \frac{1}{2}V(s)\sigma_{j+1}^2(s)ds + \sqrt{V(s)}\sigma_{j+1}dZ_s} \\
&\triangleq m_{j+1}(t), \quad t \leq T_{j+1}.
\end{align*}
$$

Clearly we have

$$
dm_{j+1}(t) = m_{j+1}(t)\sqrt{V(t)}\sigma_{j+1}(t) \cdot dZ_t.
$$

Let $<\cdot, \cdot>$ denote covariance, we have the Brownian motions under $Q^{i+1}$ to be

$$
dZ_t^{i+1} = dZ_t - <dZ_t, dm_{j+1}(t)/m_{j+1}(t)>
= dZ_t - <dZ_t, \sqrt{V(t)}\sigma_{j+1}(t) \cdot dZ_t>
= dZ_t - \sqrt{V(t)}\sigma_{j+1}(t)dt,
$$

$$
dW_t^{i+1} = dW_t - <dW_t, dm_{j+1}(t)/m_{j+1}(t)>
= dW_t - <dW_t, \sqrt{V(t)}\sigma_{j+1}(t) \cdot dZ_t>
= dW_t + \sqrt{V(t)}\sum_{k=1}^{j} \frac{\Delta T_k f_k(t)\|\gamma_k(t)\|}{1 + \Delta T_k f_k(t)} <dW_t, \frac{\gamma_k(t)}{\|\gamma_k(t)\|} \cdot dZ_t>
= dW_t + \sqrt{V(t)}\sum_{k=1}^{j} \frac{\Delta T_k f_k(t)\|\gamma_k(t)\|}{1 + \Delta T_k f_k(t)} \rho_k(t)dt
= dW_t + \xi_j\sqrt{V(t)}dt.
$$

The proof is complete.
Proof of Proposition 5.3.2:

The Radon-Nikodym derivative for forward swap measure $Q^S$ is

\[
\frac{dQ^S}{dQ} = \frac{B^S(t)/B^S(0)}{B(t)}
\]

\[
= \frac{1}{B^S(0)} \sum_{j=m}^{n-1} \Delta T_j P(0, T_{j+1}) e^{\int_0^t -\frac{1}{2} V(s) \sigma^2_{j+1}(s) ds + \sqrt{V(s)} \sigma_{j+1}(s) dZ_s}
\]

\[
\triangleq m_S(t), \quad t \leq T_m.
\]

Denote

\[
M_{j+1}(t) = e^{\int_0^t -\frac{1}{2} V(s) \sigma^2_{j+1}(s) ds + \sqrt{V(s)} \sigma_{j+1}(s) dZ_s},
\]

hence

\[
dm_S(t) = \frac{1}{B^S(0)} \sum_{j=m}^{n-1} \Delta T_j P(0, T_{j+1}) M_{j+1}(t) \sqrt{V(t)} \sigma_{j+1}(t) \cdot dZ_t
\]

\[
= \frac{1}{B^S(0) B(t)} \sum_{j=m}^{n-1} \Delta T_j P(t, T_{j+1}) \sqrt{V(t)} \sigma_{j+1}(t) \cdot dZ_t
\]

\[
= m_S(t) \sum_{j=m}^{n-1} \alpha_j \sqrt{V(t)} \sigma_{j+1}(t) \cdot dZ_t.
\]

It follows that

\[
dZ^S_i = dZ_i - <dZ_i, dm_S(t)/m_S(t)>
\]

\[
= dZ_i - <dZ_i, \sqrt{V(t)} \sum_{j=m}^{n-1} \alpha_j \sigma_{j+1}(t) \cdot dZ_t>
\]

\[
= dZ_i - \sqrt{V(t)} \sum_{j=m}^{n-1} \alpha_j \sigma_{j+1}(t) dt
\]

\[
= dZ_i - \sqrt{V(t)} \sigma_S(t) dt,
\]

\[
dW^S_i = dW_i - <dW_i, dm_S(t)/m_S(t)>
\]

\[
= dW_i - <dW_i, \sqrt{V(t)} \sum_{j=m}^{n-1} \alpha_j \sigma_{j+1}(t) \cdot dZ_t>
\]

\[
= dW_i + \sqrt{V(t)} \sum_{j=m}^{n-1} \alpha_j \sum_{k=1}^{j} \frac{\Delta T_k f_k(t) ||\gamma_k(t)||}{1 + \Delta T_k f_k(t)} \cdot dZ_t >
\]

89
\[
= dW_t + \sqrt{V(t)} \sum_{j=m}^{n-1} \alpha_j \sum_{k=1}^{j} \frac{\Delta T_k f_k(t)}{1 + \Delta T_k f_k(t)} \gamma_k(t) \rho_k(t) \, dt \\
= dW_t + \sqrt{V(t)} \sum_{j=m}^{n-1} \alpha_j \xi_j(t) \, dt \\
= dW_t + \sqrt{V(t)} \xi^S(t) \, dt.
\]

The proof is complete.

Proof of Proposition 5.3.3:

Note by definition of forward swap rate, we have

\[
R_{m,n}(t) = \frac{\beta_m - \beta_n}{\sum_{k=m}^{n-1} \Delta T_k \beta_k^{k+1}},
\]

where

\[
\beta_k = \frac{P(t, T_k)}{P(t, T_m)}.
\]

From the price-yield relation we have

\[
P(t, T_{k+1}) = \frac{P(t, T_k)}{1 + \Delta T_k f_k} = \ldots = \frac{P(t, T_m)}{\prod_{l=m}^{k} (1 + \Delta T_l f_l)},
\]

so

\[
\beta_k = \frac{1}{\prod_{l=m}^{k-1} (1 + \Delta T_l f_l)}, \quad \text{for } m + 1 \leq k \leq n.
\]

Apparently

\[
\frac{\partial \beta_k}{\partial f_j} = \begin{cases} \\
\frac{-\Delta T_j}{1 + \Delta T_j f_j} \cdot \beta_k, & k > j, \\
0, & k \leq j.
\end{cases}
\]

Therefore

\[
\frac{\partial R_{m,n}(t)}{\partial f_j(t)} = \frac{\partial}{\partial f_j} \left( \frac{\beta_m - \beta_n}{\sum_{k=m}^{n-1} \Delta T_k \beta_k^{k+1}} \right) \\
= \frac{P(t, T_m)}{B^S(t)} \left( -\frac{\partial \beta_n}{\partial f_j} \right) - \frac{R_{m,n} P(t, T_m)}{B^S(t)} \left( \sum_{k=m}^{n-1} \Delta T_k \cdot \frac{\partial \beta_k^{k+1}}{\partial f_j} \right)
\]

90
\[ \begin{align*}
&= \frac{P(t, T_m)}{B^2(t)} \cdot \left( \frac{\Delta T_j \cdot \beta_n}{1 + \Delta T_j f_j} \right) \\
&
+ \frac{R_{m,n} P(t, T_m)}{B^8(t)} \left( \sum_{k=j}^{n-1} \frac{\Delta T_j}{1 + \Delta T_j f_j} \cdot \beta_{k+1} \right) \\
&= \frac{\Delta T_j R_{m,n}}{1 + \Delta T_j f_j} \left[ \frac{P(t, T_n)}{P(t, T_m) - P(t, T_n)} + \sum_{k=j}^{n-1} \frac{\Delta T_k P(t, T_{k+1})}{B^8(t)} \right].
\end{align*} \]

The proof is complete.

Proof of Proposition 5.4.1:

Consider initial value problem with constant coefficients

\[ \begin{align*}
\frac{dA}{d\tau} &= a_0 B, \\
\frac{dB}{d\tau} &= b_2 B^2 + b_1 B + b_0,
\end{align*} \]

subject to initial conditions

\[ A(0) = A_0, \quad B(0) = B_0. \]

Since \( B \) is independent of \( A \), so it will be solved firstly. Let \( Y_1 \) be the solution to

\[ b_2 Y^2 + b_1 Y + b_0 = 0, \]

then we get

\[ Y_1 = \frac{-b_1 \pm d}{2b_2}, \quad \text{with} \quad d = \sqrt{b_1^2 - 4b_0 b_2}. \]

Without loss of generality we will take the "+" sign for \( Y_1 \). We then consider the difference between \( Y_1 \) and \( B \):

\[ Y_2 = B - Y_1. \]

Clearly \( Y_2 \) satisfies

\[ \begin{align*}
\frac{dY_2}{d\tau} &= \frac{d(Y_1 + Y_2)}{d\tau} \\
&= b_2 (Y_1 + Y_2)^2 + b_1 (Y_1 + Y_2) + b_0
\end{align*} \]
\[ = b_2 Y_2^2 + (2b_2 Y_1 + b_1) Y_2 \]
\[ = b_2 Y_2^2 + d Y_2, \]

with initial condition
\[ Y_2(0) = B_0 - Y_1. \]

Note in the last equality we have used the expression of \( Y_1 \), also the above ODE belongs to the class of Bernoulli equation which can be solved explicitly. One can verify that solution for \( Y_2 \) is
\[ Y_2 = \frac{d \cdot g e^{d \tau}}{b_2 (1 - g e^{d \tau})}, \quad \text{with} \quad g = \frac{-b_1 + d - 2B_0 b_2}{-b_1 - d - 2B_0 b_2}, \]
hence we have
\[ B(\tau) = Y_1 + Y_2 \]
\[ = \frac{-b_1 + d}{2b_2} + \frac{d}{b_2 (1 - g e^{d \tau})} \]
\[ = B_0 + \frac{(-b_1 + d - 2b_2 B_0)(1 - e^{d \tau})}{2b_2(1 - g e^{d \tau})}. \]

Having the solution of \( B \), we integrate it to get \( A \):
\[ A(\tau) = A_0 + a_0 \int_0^\tau B(s) ds \]
\[ = A_0 + a_0 B_0 \tau + \frac{a_0(-b_1 + d - 2b_2 B_0)}{2b_2} \int_0^\tau \frac{1 - e^{d \tau}}{1 - g e^{d \tau}} d\tau \]
\[ = A_0 + a_0 B_0 \tau + \frac{a_0(-b_1 + d - 2b_2 B_0)}{2b_2} \left[ \tau - \int_0^\tau \frac{(1 - g) e^{d \tau}}{1 - g e^{d \tau}} d\tau \right] \]
\[ = A_0 + \frac{a_0(-b_1 + d) \tau}{2b_2} - \frac{a_0(-b_1 + d - 2b_2 B_0)}{2b_2 d} \int_1^{e^{d \tau}} \frac{(1 - g) u}{1 - g u} du \]
\[ = A_0 + \frac{a_0(-b_1 + d) \tau}{2b_2} - \frac{a_0(-b_1 + d - 2b_2 B_0) (g - 1)}{g} \ln \left( \frac{1 - g e^{d \tau}}{1 - g} \right) \]
\[ = A_0 + \frac{a_0}{2b_2} \left[ (-b_1 + d) \tau - 2 \ln \left( \frac{1 - g e^{d \tau}}{1 - g} \right) \right]. \]
Specifying $a_0, b_0, b_1, b_2,$ and $A_0, B_0$ by

\[
\begin{align*}
  a_0 &= \kappa \theta, \\
  b_0 &= \frac{1}{2} \lambda^2 (z^2 - z), \\
  b_1 &= \rho \epsilon \lambda z - \kappa \xi, \\
  b_2 &= \frac{\epsilon^2}{2}, \\
  A_0 &= A(\tau_j, z), \\
  B_0 &= B(\tau_j, z),
\end{align*}
\]

and replacing $\tau$ by $\tau - \tau_j$, then we can derive (5.22). The proof is complete.

**Proof of Proposition 5.6.1:**

Given the process

\[dV(t) = \kappa(\theta - V(t))dt + \epsilon \sqrt{V(t)}dW_t,\]

take expectation under $Q$ on both sides and exchange with differentiation, we have

\[
\frac{dE[V]}{dt} = \kappa(\theta - E[V]).
\]

Solving this ODE we obtain

\[E^Q[V(t + \Delta t)] = \theta + (V(t) - \theta)e^{-\kappa \Delta t}.\]

For the second part of (5.40), firstly by Ito's lemma,

\[dV^2 = (\epsilon^2 V + 2\kappa(\theta V - V^2))dt + 2\epsilon V \sqrt{V}dW_t.\]

Similarly, we have

\[dE[V^2] = ((2\kappa \theta + \epsilon^2)E[V] - 2\kappa E[V^2])dt,\]

hence

\[d(e^{2\kappa t}E[V^2]) = (2\kappa \theta + \epsilon^2)e^{2\kappa t}E[V]dt.\]
On the other hand,

\[ d(e^{2\kappa t}(E[V])^2) = 2e^{\kappa t}E[V]d(e^{\kappa t}E[V]) = 2\kappa \theta e^{2\kappa t}E[V]dt. \]

It is easily seen that

\[ d \left( e^{2\kappa t}E[V^2] - \frac{2\kappa \theta + \epsilon^2}{2\kappa \theta} e^{2\kappa t}E[V]^2 \right) = 0. \]

Integrate the above equation over \((t, t + \Delta t)\), we obtain

\[ e^{2\kappa (t + \Delta t)}E[V^2(t + \Delta t)] - \frac{2\kappa \theta + \epsilon^2}{2\kappa \theta} e^{2\kappa (t + \Delta t)}E[V(t + \Delta t)]^2 = -\frac{\epsilon^2}{2\kappa \theta} e^{2\kappa t}V^2(t). \]

The second part of (5.40) then follows, after we have (5.40), then by moment-matching, we can easily derive (5.38), and the proof is complete.
Bibliography


95


