Stochastic Skew in Interest Rate Cap and Currency Option Markets

NG, Hon Yip

A Thesis Submitted in Partial Fulfilment

of the Requirements for the Degree of

Doctor of Philosophy

in

Systems Engineering and Engineering Management

The Chinese University of Hong Kong

September 2011

Thesis/Assessment Committee

Prof. Chen, Nan (Chair)

Prof. Leung, Kwai-Sun (Thesis Supervisor)

Prof. Li, Duan (Co-Supervisor)

Prof. Wong, Hoi-Ying (Committee Member)

Prof. Wu, Lixin (External Examiner)

ABSTRACT

This thesis considers the effect of stochastic skew in the interest rate cap and currency option markets, where we observe obvious stochastic variation of skew of implied volatility curve over time. To develop option pricing models consistent with empirical evidence, we adopt the Wishart process to model both stochastic volatility and stochastic skew of the asset return and to price options in both markets. As an affine model, the model is analytically tractable. Some distributional properties of the models are studied. The key feature of our model is that, when compared with the multi-factor Heston model, which generates stochastic skew through its volatility processes, the Wishart process contains not only volatility processes, but also volatility-unrelated processes which provide extra freedom to model the variation of skew that is not captured by the volatility processes. Numerical experiments demonstrate that the Wishart model has greater flexibility to model stochastic skew than the multi-factor Heston model in both the interest rate cap market and currency option market. Finally, results of calibration to market data and model estimation demonstrate the superiority of the Wishart model to the multi-factor Heston model in the interest rate cap market.

摘要

本文考慮在利率上限市場和貨幣期權市場裡的隨機偏度Stochastic skew現象，在這兩個市場裡我們觀察到引伸波動曲線的偏度表現出明顯的隨機時態變化。為了開發與實證研究一致的期權定價模型，我們採用Wishart過程模擬資產收益中的隨機波動和隨機偏度的效果，並為這兩個市場裡的期權定價。作為一個Affine模型，該模型是解析上易駕馭的。一些有關該模型的分佈特性會考察。該模型的主要特點是，當與那只能通過波動過程產生隨機偏度的多因Heston模型比較，Wishart過程不僅包含波動過程，也包含與波動無關的、提供額外自由度來捕捉單靠波動過程捕捉不到的隨機偏度的過程。數值實驗表明，Wishart模型比多因Heston模型在利率上限市場和貨幣期權市場裡有更大的彈性去捕捉隨機偏度。最後，校準市場數據和模型估計的結果顯示出Wishart模型比較多因Heston模型的優越性。

THE CHINESE UNIVERSITY OF HONG KONG

GRADUATE SCHOOL

The undersigned certify that we have read a thesis, entitled "Stochastic Skew in Interest Rate Cap and Currency Option Markets " submitted to the Graduate School by NG, Hon Yip ( ) in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Systems Engineering and Engineering Management. We recommend that it be accepted.

Prof. Leung, Kwai-Sun
Supervisor

Prof. Li, Duan
Co-Supervisor

Prof. Chen, Nan

Prof. Wong, Hoi-Ying

Prof. Wu, Lixin
External Examiner

DECLARATION

No portion of the work referred to in this thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institution of learning.

ACKNOWLEDGEMENT

I thank the Lord Jesus Christ, who carries me through all difficulties and keeps me from distresses. Without His guidance, the thesis cannot be completed.

To my supervisor, Prof. Leung, Kwai-Sun, I would like to express my deepest gratitude for his patience, guidance, support and advice to me throughout the period of PhD program. I would like to thank Prof. Li, Duan for his co-supervision, continuous encouragement and vision to me in the process of doing research. Special thanks should go to Prof. Chen, Nan, Prof. Wong, Hoi-Ying and Prof. Wu, Lixin for their commitment as thesis committee members and precious comments on my research. I would also like to thank all the staff from the Department of Systems Engineering and Engineering Management for their assistance.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

The main theme of the thesis is the development of option pricing models capable of modeling "stochastic skew" present in the interest rate cap market and currency option market, respectively.

In financial markets, a well-known stylized fact is that the implied volatility, which is the number placed in the constant volatility term of the Black and Scholes [12] option pricing formula in order to obtain the quoted option price, exhibits a shape of smile or smirk as a function of moneyness level or strike price; as a consequence, the assumption of the Black-Scholes [12] model that asset price is log-normally distributed is no longer valid. To build an option pricing model that is consistent with the observed volatility smile/smirk, the notion of stochastic volatility is introduced[1] and, particularly within the affine stochas-

[1] In the thesis, we restrict ourselves within the framework of stochastic volatility and exclude the local volatility models and jump diffusion models.

tic volatility framework (for general definition, see Dai and Singleton [31], Duffie and Kan [32], Duffie *et al.* [33]), the steepness (skew) of the model implied volatility smile/smirk is controlled by the correlation between the shocks to the asset return and to the variance dynamics so that positive (resp. negative) correlation results in upward (resp. downward) sloping smile/smirk. It is known as the leverage effect. At the beginning of development, models with single stochastic volatility, for example, Heston [48], are developed in order to model the time-varying nature of volatility. Later, models with multiple stochastic volatilities, for example, Bates [7] and Christoffersen *et al.* [25], are developed to capture multiple risk factors and better model the term structure of volatility in financial markets.

Recently, "stochastic skew", documented by Carr and Wu [21] in currency option market and by Christoffersen *et al.* [25] in index option market, refers to the fact that the slope of the volatility smile/smirk varies stochastically over time, which poses a new challenge to building option pricing models consistent with market phenomena. Within the affine stochastic volatility framework, it means that models with *stochastic* instantaneous correlation between the shocks to the asset return and to the variance are necessary. However, in the Heston [48] model, this correlation is fixed so it cannot generate stochastic skew. In the Christoffersen *et al.* [25] model, this correlation, although stochastic, given a fixed set of model parameters, is completely determined by the variance factors, thereby limiting the power to model stochastic skew.

Different from, and extending, the above models, a new class of option pricing models has been developed by incorporating the *Wishart process*, which not only models stochastic volatility, but also provides extra flexibility to model stochastic skew. The Wishart

process, closed-related to the well-known and important Wishart distribution in multivariate statistics (please refer to Appendix A), as a matrix extension of Heston [48] volatility model, is proposed by Bru [18] and is introduced into finance by Gourieroux [38]. Ever since then, the Wishart process is used as a tool to model multivariate risk (for example, Gourieroux *et al.* [39], and Gourieroux and Sufana [40]), to price options on multiple assets (for example, Branger and Muck [16], and Da Fonseca *et al.* [29]), to model stochastic skew in equity/index option markets (for example, Da Fonseca *et al.* [30], Da Fonseca and Grasselli [28], and Gruber et al. [43]) and, to investigate the choice of optimal portfolio when the correlations across assets are stochastic (for example, Buraschi *et al.* [19]). Asymptotic approximation of vanilla option prices and implied volatilities with Wishart stochastic volatility are derived in Benabid *et al.* [9] and easy-to-implement Monte Carlo (MC) procedures are provided by Gauthier and Possamaï [37].

That the Wishart process is a *symmetric positive-definite* matrix *affine* process makes it suitable to act as a source of stochastic volatility for option pricing. In the multi-asset framework, its symmetry and positivity make it resemble a stochastic covariance matrix of asset prices so that options on multiple assets can be priced (for example, Da Fonseca *et al.* [29], Gourieroux [38], and, Gourieroux and Sufana [40]); on the other hand, in the single-asset framework, when the conditional variance of asset price equals the trace of Wishart process, the off-diagonal processes of Wishart process give the freedom to model stochastic skew (for example, Da Fonseca *et al.* [30], Da Fonseca and Grasselli [28] and, Gruber et al. [43]). Finally, the affine nature of the Wishart process makes it straightforward to derive closed-form formula for vanilla option prices, in terms of the characteristic function, so that these options can be accurately and efficiently priced through Fourier

transform.

So far, the Wishart process is successfully implemented in the equity option and index option markets; however, the possibility of applying Wishart process in other financial markets is largely unexplored. Therefore, in the thesis, we study stochastic skew in other markets, namely, the interest rate cap and currency option markets, and develop consistent option pricing models through the use of Wishart process. The related background and our contribution for each case are provided below.

1.2 Stochastic Skew in Interest Rate Cap Market

In the interest rate derivative markets, the LIBOR market model, separately developed by Brace *et al.* [14], Jamshidian [50], and Miltersen *et al.* [62], is the benchmark model for pricing interest rate derivatives, where, in an arbitrage-free term structure framework, the simply compounding forward rates are log-normally distributed and that the market conventional Black's [11] formula is justified for the pricing of caplets. As a consequence, it is in contrast to the Heath, Jarrow and Morton (HJM) [46] framework, under which if the continuously compounded instantaneous forward rate being modeled is log-normally distributed, it explodes with positive probability so the bond price is zero, and, it is very difficult, if not impossible, to obtain closed-form solutions of those liquidly-traded vanilla derivatives, for example, caps and swaptions, to allow for rapid pricing and calibration. Apart from theoretical reasons, the prevalence of the LIBOR market model also lies on the fact that, when the volatility and correlation structure of the observed simple forward

rates being modeled are extracted from market prices of caps and swaptions, most, if not all, of the exotic derivatives can be priced using MC simulation or analytical pricing formulae. For more details, please refer to Bachert *et al.* [6] or Rebonato [69].

The assumption that simply forward rates are log-normally distributed implies that implied caplet volatility curves should be flat. However, this violates the observation that the implied volatility curves often exhibit a shape of smile or smirk, thereby implying that the distribution of log-forward rates has fatter left tail than when compared with the normal distribution. As such, it motivates the development of more sophisticated LIBOR market models. In particular, the stochastic volatility model is such a class of models in which the term structure is multiplicative-perturbed by positive-valued stochastic processes whose shocks are possibly correlated with the shocks to the term structure. The widely adopted approach is the Heston [48] model, for instance, Andersen and Brotherton-Ratcliffe [3], and Wu and Zhang [80], in which the variance process follows a mean-reverting square-root process. Particularly, in the Wu and Zhang [80] model, the most prominent feature is that the slope of implied volatility curve is controlled by the correlation between the shocks to the forward rate and to the variance dynamics, and the caplet price is expressed in terms of the characteristic function of log-forward rate and rapidly computed through fast Fourier transform.

Notwithstanding the above, recently, the discovery of the "unspanned stochastic volatility" (hereafter, USV) by Collin-Dufresne and Goldstein [26], where the prices of caps and swaptions are driven by stochastic volatility factors that are not spanned by the underlying forward rates or forward swap rates, shows that stochastic volatility is necessary

in term structure modelling. In fact, they show that all HJM models with stochastic volatility can exhibit USV. More importantly, based on this finding, Li and Zhao [58] demonstrate that traditional term structure models have difficulty hedging interest rate derivatives and show that multiple stochastic volatility factors are necessary to capture USV. Thus, these results invoke the application of multi-factor stochastic volatility models to price caps and swaptions,[2] for example, Han [45], Jarrow *et al.* [51], and Schwartz and Trolle [72]. All these models use at least three Heston volatilities to capture USV and, thus, are consistently shown to have better performance than do models with one or two Heston volatilities. Recently, Belomestnyl *et al.* [8] and Ma [61] also multi-factor stochastic volatility LIBOR market models.

With regard to this thesis, our objective is to develop a practically useful term structure model which is able to capture not only USV, but also, more importantly, "stochastic skew" observed in the interest rate cap market. Meanwhile, we also note that Gourieroux and Sufana [41] also develop term structure models based on the Wishart process; however, they develop an extension of the Quadratic term structure model and derive the condition of positive bond yield for all maturities, but do not provide closed-form pricing formula for any interest rate options. Different from theirs, our paper aims at developing an analytically tractable term structure model for pricing interest rate derivatives with extra flexibility to capture stochastic skew.

[2] As a remark, evidence of the presence of USV in other financial markets is provided by Andersen and Benzoni [4], Bikbov and Chernov [10], Collin-Dufresne et al. [27], Heidari and Wu [47], Schwartz and Trolle [73], and others.

First, we define and document stochastic skew in a set of U.S. interest rate cap data from July-2002 to December-2009. As a measure of the slope of volatility curve, the skew of the implied caplet volatility curve for a particular maturity is defined as the difference between the Black's implied caplet volatilities at two moneyness levels[3]. Throughout the sample of data, we observe that the skew of individual maturities exhibits stochastic variation over time.

Second, to price caps/floors consistently with such empirical phenomenon, we incorporate the Wishart process into the LIBOR market model in order to provide extra flexibility to model stochastic skew. We derive closed-form formula for caplet prices based on freezing approximation and Fourier transform. MC simulation is carried out to demonstrate the accuracy and efficiency of the pricing formula. Obviously, analytical tractability is an advantage of our model over multi-factor Heston HJM models, for example, Schwartz and Trolle [72], and Quadratic term structure models, for example, Gourieroux and Sufana [41].

Third, since we claim that the volatility-unrelated processes in our model provide an additional degree of freedom to model stochastic skew, we show that, through a numerical experiment, given a set of parameters, by varying the volatility-unrelated processes, our model can alter the volatility skew without changing the volatility level. This property obviously is an potential advantage over the multi-factor Heston LIBOR market model.

Fourthly, a calibration exercise is carried out demonstrate the superiority of the two-

[3]Moneyness is the ratio of the strike price to the forward rate of the caplet.

dimensional Wishart model to the two-factor Heston model in fitting the implied caplet volatility surface in one day.

Finally, we estimate the two-dimensional Wishart model and the two-factor Heston model using market data of 3 years (January-2004 to December-2006) and perform an out-of-sample test using market data of the following 3 years (January-2007 to December-2009). The essence of the estimation exercise is to show that, given a fixed set of model parameters, when compared with the two-factor Heston model, the volatility-unrelated processes in our model provide an additional degree of freedom to better capture the time variation of stochastic skew, which can improve the pricing performance of our model. Estimation results show that the better pricing performance of the Wishart model over the two-factor Heston model can be attributed to the better modeling of stochastic skew by the Wishart model.

It is worthwhile to emphasize that the main difference between our research and the previously-mentioned term structure literature is that our emphasis is that the enhancement in pricing is achieved by including a volatility-unrelated process which gives extra freedom to model stochastic skew, while the emphasis of Han [45], Jarrow *et al.* [51], Schwartz and Trolle [72], and Ma [61] is that the enhancement in pricing is achieved by the inclusion of multiple stochastic volatilities which are solely responsible for the skew in their models; as a result, their models have some restriction on modeling stochastic skew.

1.3 Stochastic Skew in Currency Option Market

Similar to the interest rate cap market, currency option market also exhibits stochastic skew, which is first documented by Carr and Wu [21]. Moreover, it is well-known that currency values exhibit mean-reversion (for example, Jorion and Sweeney [52], and Sweeney [74]) and stochastic volatility (for example, Bates [7], Heston [48], and Palm and Vlaar [65]). To be consistent with the latter two properties, Wong and Lo [76] assume that the log-currency value follows a mean-reverting process and the volatility follows the Heston model. To further allow for multi-scale stochastic volatility, Wong and Zhao [77] assume that the volatility dynamics is driven by two Heston stochastic volatilities; however, their model has limited degree of freedom to model stochastic skew.

To take into account the above three essential features - mean-reversion, stochastic volatility and stochastic skew of foreign exchange rate, we generalize the two-factor Heston stochastic volatility model of Wong and Zhao [77] through the use of the Wishart process and develop an accurate and efficient option pricing model which possesses extra flexibility to capture "stochastic skew" in the currency option market.

First, we use risk reversal[4] as a proxy for volatility skew and document stochastic skew in the EUR/USD currency option market. Throughout the sample, we find that the skew of individual maturities exhibits stochastic time-varying nature.

[4]The 10-delta risk reversal is defined as the difference between the implied volatilities of a 10-delta call and a 10-delta put. Risk reversal of other deltas is defined similarly. For more information on the terminologies used in currency option markets, please refer to Reiswich and Wystup [70]

Second, we generalize the Wong and Zhao [77] model by making use of the Wishart process and derive pricing formula for vanilla currency options. In addition, we extends our model to include Poisson shocks with arbitrary jump size distribution, generalizing the single-volatility framework of O'Hara and Pillay [64]. MC simulation is carried out to justify the accuracy and efficiency of the model.

Finally, using a term structure of currency futures prices, we carry out a simple numerical experiment to show that, when compared with the multi-factor Heston model, the Wishart model has an additional degree freedom to model stochastic skew once the variance factors are fitted to short-term and long-term volatilities.

1.4 Outline of Thesis

This thesis is organized as follows: Chapter 2 demonstrates the presence of stochastic skew in the interest rate cap market, incorporates the Wishart process into the LIBOR market model, with closed-form solution for caplet prices derived. Chapter 3 demonstrates the presence of stochastic skew in the currency option market and develop a currency option pricing model through the use of Wishart process, which accommodates for mean-reversion, stochastic volatility and stochastic skew, and derive closed-form formula for pricing currency options. For both chapters, accurate and efficient closed-form pricing formulas for vanilla options are demonstrated by using MC simulation, and numerical examples are provided to illustrate the crucial properties of our models. Furthermore, for the interest rate cap market, we provide calibration and estimation results of the Wishart

model and the Heston model. Chapter 4 concludes the thesis. Mathematical derivations and numerical methods applied are found in Appendices.

Chapter 2

Stochastic Skew in Interest Rate Cap Market

This chapter provides the empirical evidence of stochastic skew in the U.S. interest rate cap market and develop a practically useful, consistent term structure model by incorporating the Wishart process into the LIBOR market model. Closed-form formula of caplet prices is derived using freezing approximation and Fourier transform, whose accuracy and efficiency is demonstrated by MC simulation. Numerical example shows that, when compared with the multi-factor Heston model, the Wishart model has an additional degree of freedom to model stochastic skew once the variance factors are fitted to short-term and long-term volatilities, thereby highlighting the strength of the Wishart model. A calibration exercise is carried out to show that the two-dimensional Wishart model performs better than the two-factor Heston model in fitting the implied volatility surface in one day. Finally, we estimate the two-dimensional Wishart model and the two-factor Heston model with a time series of implied volatility surfaces, and results shows that the

two-dimensional Wishart model outperforms the two-factor Heston model.

2.1 Empirical Evidence

The sample consists of weekly data (on each Friday) of U.S. 3-month forward LIBORs and mid-quotes of U.S. Black's cap volatilities for various strike prices (1%, 2%, 3%, 4%, 5%, 6%, 7%, 8% and 9%) and maturities (1, 2, 3, 4, 5, 6, 7, 8, 9 and 10 years) from July 2002 to December 2009. On each day, we strip the Black's implied caplet volatility surface from the cap volatility surface, and, for each maturity, we linearly interpolate caplet volatility curves with respect to the strike price at fixed moneyness levels (0.7, 0.8, 0.9, 1.0, 1.1, 1.2 and 1.3). Now, the implied caplet volatility skew for a particular maturity, as a measure of the slope of volatility curve, is defined as the difference between the implied caplet volatilities at moneyness levels 1.3 and 0.7. Then, we compute the skew over all maturities for the whole sample and display in Figures 2.1 and 2.2. Figure 2.1 reveals that the implied volatility skew for all maturities fluctuates randomly over the whole sample period. The clustering effect is also observed and is more pronounced in short maturities (1 year to 4 years) than in long maturities (5 years to 10 years). On the other hand, Figure 2.2 evidences the stochastic nature of the term structure of the implied volatility skew, which exhibits different patterns like hump-shaded, decreasing or increasing pattern. Therefore, to be consistent with this market phenomenon, a term structure model should grant sufficient flexibility to capture the variation of implied volatility skew.

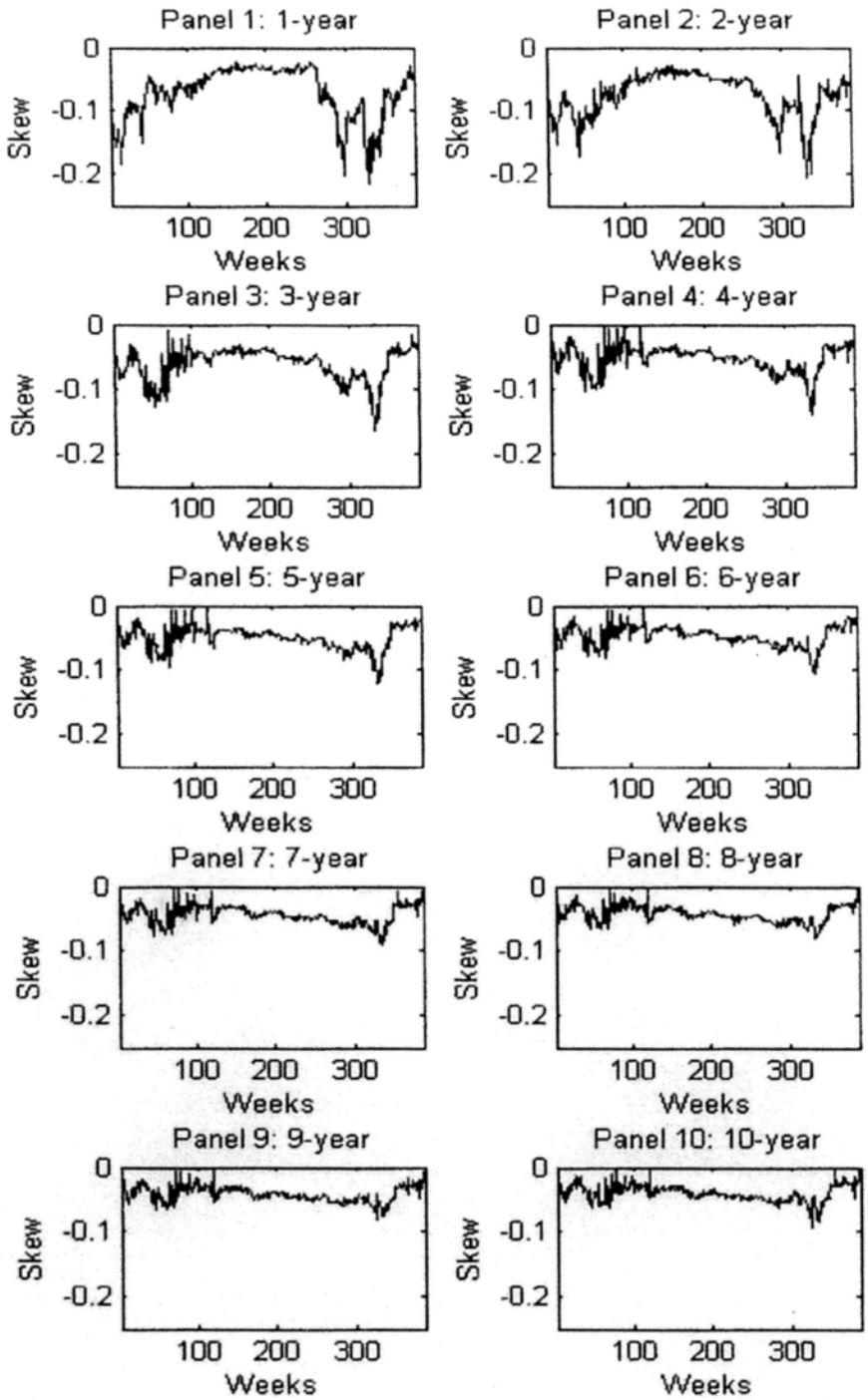

Figure 2.1: The time series of implied caplet volatility skew for various maturities from July 2002 to December 2009.

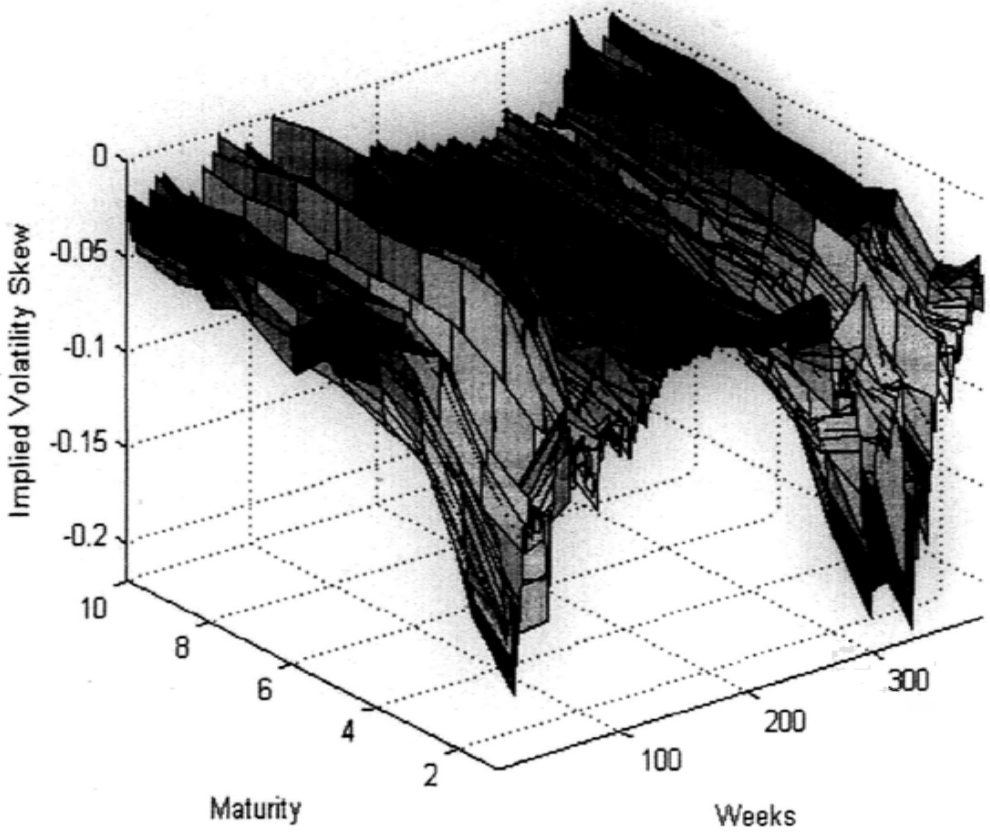

Figure 2.2: The time series of the term structure of implied caplet volatility skew from July 2002 to December 2009.

2.2 The Model

In this section, we incorporate the Wishart process into the benchmark LIBOR market model and derive the analytical formula of caplet prices through the freezing approximation and Fourier transform. We demonstrate the accuracy and efficiency of pricing formula by MC simulation.

To facilitate our discussion, we define the following notations: 1. $[A^{ij}]_{i,j=1,\ldots,n}$ (or $[A_{ij}]_{i,j=1,\ldots,n}$) as a square matrix of order n with A^{ij} (or A_{ij}) being its (i,j)-th element; 2. $[B]_{i,j}$ as the (i,j)-th element of the matrix B; 3. $\mathrm{Tr}(A)$ as the trace of the matrix A; 4. A^{T} as the transpose of the matrix A, that is, with a slight abuse of notation, $[A_{ji}]_{i,j=1,\ldots,n} = [A_{ij}]^{\mathrm{T}}_{i,j=1,\ldots,n}$; and, 5. $\mathbf{I}_n$ and $\mathbf{0}_n$ as the identity matrix and zero matrix of order n, respectively.

Let $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t\geq 0}, \mathbb{Q})$ be a filtered probability space, where the square matrix Brownian motions $W_t \triangleq [W_t^{ij}]_{i,j=1,\ldots,n}$ and $Z_t \triangleq [Z_t^{ij}]_{i,j=1,\ldots,n}$ are defined for all $t \geq 0$ and are adapted to the filtration $\{\mathcal{F}_t\}_{t\geq 0}$. Here, for $i,j = 1,\ldots,n$, W_t^{ij} (resp. Z_t^{ij}) are independent scalar Brownian motions. The probability measure $\mathbb{Q}$ is considered as risk-neutral measure.

Under the probability measure $\mathbb{Q}$, the Wishart process $\Sigma_t \triangleq [\Sigma_t^{ij}]_{i,j=1,\ldots,n}$ is an n-by-n symmetric positive-definite matrix process that follows the dynamics

$$d\Sigma_t = \left(\Omega\Omega^{\mathrm{T}} + M\Sigma_t + \Sigma_t M^{\mathrm{T}}\right)dt + \sqrt{\Sigma_t}dW_t Q + Q^{\mathrm{T}}(dW_t)^{\mathrm{T}}\sqrt{\Sigma_t}, \tag{2.1}$$

where $M \triangleq [M_{ij}]_{i,j=1,\ldots,n}$ and $Q \triangleq [Q_{ij}]_{i,j=1,\ldots,n}$ are constant real square matrices, and

$\sqrt{\Sigma_t}$, the square root of Σ_t, is defined as the unique symmetric positive definite matrix such that $\sqrt{\Sigma_t}\sqrt{\Sigma_t} = \Sigma_t$, that is, $\Sigma_t^{ij} = \sum_{l=1}^{n} \sigma_t^{il}\sigma_t^{jl}$. To guarantee mean-reversion feature and strict positivity of the Wishart process, we assume that M is negative semi-definite, $\Omega\Omega^{\mathrm{T}} = \beta Q^{\mathrm{T}}Q$ and $\beta > n-1$ (see Bru [18]). For $n = 2$, the positive definiteness condition is equivalent to $\Sigma_t^{11}\Sigma_t^{22} - (\Sigma_t^{12})^2 > 0$.

Suppose that we observe the term structure at time 0. Given a set of dates $\{T_j\}_{j=1}^{N+1}$ such that $0 \leq t < T_1 < \cdots < T_j < T_{j+1} < \cdots < T_{N+1}$, we assume that, under the probability measure $\mathbb{Q}$, the price at time t of a default-free zero coupon bond with maturity T_j at which it pays 1 unit of currency, denoted as $P_j(t)$, follows the dynamics

$$\frac{dP_j(t)}{P_j(t)} = r(t)dt + \mathrm{Tr}\Big(V_j(t)\sqrt{\Sigma_t}dZ_t\Big), \tag{2.2}$$

for $j = 1, \ldots, N+1$, where $t \in [0, T_j]$, $r(t)$ is the instantaneous risk-free rate, $V_j(t)$ is a diagonal matrix of adapted scalar processes with the i-th diagonal element denoted as $v_j^{(i)}(t)$ and $V_j(t) = \mathbf{0}_n$ for $t \geq T_j$, Σ_t is the Wishart process specified by (2.1) and Z_t is a matrix Brownian motion correlated with W_t in the way that the same correlation R_{ij} is applied to all the Brownian motions belonging to the column i of Z_t and the corresponding Brownian motions of the column j of W_t so that Z_t can be written as (see Da Fonseca *et al.* [30])

$$Z_t = W_t R^{\mathrm{T}} + B_t\sqrt{\mathbf{I}_n - RR^{\mathrm{T}}}, \tag{2.3}$$

where B_t is a matrix Brownian motion independent of W_t and $R \triangleq [R_{ij}]_{i,j=1,\ldots,n}$.

Define $L_j(t)$ as the time t value of the simple forward rate with reset date T_j and maturity date T_{j+1}, where $t \in [0, T_j]$ and $j = 1, \ldots, N$. The no-arbitrage assumption

yields the relationship

$$L_j(t) = \frac{1}{\Delta T_j}\Big(\frac{P_j(t)}{P_{j+1}(t)} - 1\Big),$$

where $\Delta T_j = T_{j+1} - T_j$. By Ito's lemma, $L_j(t)$ follows the dynamics

$$\frac{dL_j(t)}{L_j(t)} = \mathrm{Tr}\Big(U_j(t)\sqrt{\Sigma_t}\Big(dZ_t - \sqrt{\Sigma_t}V_{j+1}(t)dt\Big)\Big), \tag{2.4}$$

where $U_j(t)$ is a diagonal matrix of deterministic bounded piece-wise continuous functions with the i-th diagonal element denoted as $u_j^{(i)}(t)$ and $U_j(t) = \mathbf{0}_n$ for $t \geq T_j$, and satisfies the equation,

$$U_j(t) = \frac{1 + \Delta T_j L_j(t)}{\Delta T_j L_j(t)}\Big(V_j(t) - V_{j+1}(t)\Big).$$

Since t and T_1 are close, $P_1(t)$ behaves as a risk-free bank account so that we can set $V_1(t) \equiv \mathbf{0}_n$, which is justified by Brace *et al.* [14]. Therefore, we have

$$V_{j+1}(t) = -\sum_{k=1}^{j} \frac{\Delta T_k L_k(t)}{1 + \Delta T_k L_k(t)} U_k(t). \tag{2.5}$$

To gain the first insight of the model, consider the instantaneous variance conditional on the information up to time t,

$$\mathrm{Var}^{\mathbb{Q}}\Big(\frac{dL_j(t)}{L_j(t)}\Big|\mathcal{F}_t\Big) = \Big(\sum_{i=1}^{n} u_j^{(i)}(t)^2 \Sigma_t^{ii}\Big)dt.$$

If each $u_j^{(i)}(t)$ represents the i-th factor driving the term structure of forward rates, then the *diagonal* elements of Σ_t serve as a multiplicative, stochastic perturbation to the volatility of forward rate, thereby giving rise to volatility skew over maturities. The off-diagonal elements of Σ_t do *not* enter into the instantaneous variance but, as shown later, they give extra freedom to model the variation of implied volatility skew.

Since a cap is simply a series of separately exercisable caplets, the pricing of caps is reduced to the pricing of caplets which we will focus on. Now, a caplet with maturity T_{j+1} and strike price K is a call option on $L_j(t)$ that determines the payoff $\Delta T_j \max\{L_j(T_j) - K, 0\}$ at reset date T_j and pays at maturity date T_{j+1}. It is more convenient to consider the pricing of caplet under T_{j+1}-forward measure, $\mathbb{Q}^{T_{j+1}}$, which is equivalent to $\mathbb{Q}$, under which $L_j(t)$ is a martingale. The following proposition provides a way for deriving the dynamics of $L_j(t)$ and Σ_t under T_{j+1}-forward measure.

Proposition 2.2.1. Assume that the Novikov condition is satisfied, that is, for $t \in [0, T_{j+1}]$,

$$\mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(\frac{1}{2}\int_0^t \mathrm{Tr}(V_{j+1}(s)^2\Sigma_s)ds\Big)\Big] < \infty.$$

Then

$$dZ_t^{T_{j+1}} = dZ_t - \sqrt{\Sigma_t}V_{j+1}(t)dt \tag{2.6}$$

$$dW_t^{T_{j+1}} = dW_t - \sqrt{\Sigma_t}V_{j+1}(t)Rdt \tag{2.7}$$

are the square matrix Brownian motions for $L_j(t)$ and Σ_t under $\mathbb{Q}^{T_{j+1}}$, respectively. The dynamics of $L_j(t)$ and Σ_t under $\mathbb{Q}^{T_{j+1}}$ are, respectively,

$$\frac{dL_j(t)}{L_j(t)} = \mathrm{Tr}\Big(U_j(t)\sqrt{\Sigma_t}dZ_t^{T_{j+1}}\Big), \tag{2.8}$$

$$\begin{aligned} d\Sigma_t = {} & \Big(\beta Q^{\mathrm{T}}Q + M^{T_{j+1}}(t)\Sigma_t + \Sigma_t M^{T_{j+1}}(t)^{\mathrm{T}}\Big)dt \\ & + \sqrt{\Sigma_t}dW_t^{T_{j+1}}Q + Q^{\mathrm{T}}(dW_t^{T_{j+1}})^{\mathrm{T}}\sqrt{\Sigma_t},, \end{aligned} \tag{2.9}$$

where

$$M^{T_{j+1}}(t) = M + Q^{\mathrm{T}}R^{\mathrm{T}}V_{j+1}(t).$$

Proof: Please refer to Section B.1 of Appendix B. □

Under the probability measure $\mathbb{Q}^{T_{j+1}}$, the price of the caplet with maturity T_{j+1} and strike price K is given by

$$C(K, T_{j+1}) = \Delta T_j P_{j+1}(0) \mathbb{E}^{\mathbb{Q}^{T_{j+1}}} \Big[\max\{L_j(T_j) - K, 0\} \Big].$$

However, since the mean-reverting parameter $M^{T_{j+1}}(t)$ of the Wishart process under $\mathbb{Q}^{T_{j+1}}$ depends on forward rate processes $L_k(t)$, for $k = 1, \ldots, j$, through (2.5), it becomes non-affine in the state variables and the model loses analytical tractability. Analytical tractability can be achieved by "freezing" all forward rate processes $L_k(t)$, for $k = 1, \ldots, j$, at their initial values (see, for example, Brace *et al.* [15], Ma [61], Piterbarg [67], Schwartz and Trolle [72], and Wu and Zhang [80]):

$$\widehat{V}_{j+1}(t) = -\sum_{k=1}^{j} \frac{\Delta T_k L_k(0)}{1 + \Delta T_k L_k(0)} U_k(t). \tag{2.10}$$

Then, the mean-reverting parameter of the Wishart process becomes

$$M^{T_{j+1}}(t) = M + Q^{\mathrm{T}} R^{\mathrm{T}} \widehat{V}_{j+1}(t), \tag{2.11}$$

which is now a deterministic function of t.

Remark 2.2.1. If the matrices M, Q and R are diagonal and Σ_0 (the value of Σ_t at $t = 0$) is positive-definite and diagonal, the Wishart process (2.1) reduces to the multi-factor Heston model so that the Wishart model embraces multi-factor Heston term structure models as special cases. Interestingly, if we add a Poisson-Normal jump component[1] to the forward rate dynamics (2.8) under forward measure, we derive the multi-stochastic volatility

[1] Please refer to the working paper by Leung *et al.* [54]

LIBOR market model of Jarrow *et al.* [51] with non-zero correlation coefficients. On the other hand, if we take the limit $T_{j+1} \downarrow T_j$ and let $L_j(t)$ be continuously compounded, we derive the multi-stochastic volatility HJM model of Schwartz and Trolle [72] as a special case of the Wishart model.

Remark 2.2.2. The Wu and Zhang [80] model, which, to our knowledge, is the first correlation-based single-stochastic volatility LIBOR market model, has n factors driving the term structure and 1 additional USV factor; on the other hand, the Wishart model has n factors driving the term structure and n additional USV factors.

2.2.1 The Characteristic Function

One attractive feature of the Wishart model is that the affine nature of Wishart process enables us to compute the the characteristic function of the log-forward rate process so that the pricing of caplets can thus be achieved by fast Fourier transform.

Proposition 2.2.2. Suppose that $X_t \triangleq \ln L_j(t)$ and the characteristic function of X_{T_j} under the T_{j+1}-forward measure, conditional on the information that $\{X_t = x, \Sigma_t = \Sigma\}$, is denoted as

$$\Psi(x, \Sigma, \tau; u) \triangleq \mathbb{E}^{\mathbb{Q}^{T_{j+1}}}\Big[\exp\Big(\mathrm{i}uX_{T_j}\Big)\Big|X_t = x, \Sigma_t = \Sigma\Big], \tag{2.12}$$

where $\tau = T_j - t$, $\mathrm{i} = \sqrt{-1}$ and $u \in \mathbb{R}$. Then the characteristic function of X_{T_j} in (2.12) is given by

$$\Psi(x, \Sigma, \tau; u) = \exp\Big(\mathrm{Tr}\Big(A(\tau; u)\Sigma\Big) + \mathrm{i}ux + c(\tau; u)\Big), \tag{2.13}$$

where the functions $A(\tau;u)$ and $c(\tau;u)$ satisfy

$$\begin{aligned} A(\tau;u) &= H(\tau;u)^{-1}G(\tau;u) && (2.14)\\ c(\tau;u) &= -\frac{\beta}{2}\mathrm{Tr}\Big(\ln H(\tau;u) \\ &\quad + \int_0^\tau \Big(M^{T_{j+1}}(T_j-s)^{\mathrm{T}} + 2iuU_j(T_j-s)RQ\Big)ds\Big), && (2.15)\end{aligned}$$

where $G(\tau;u)$ and $H(\tau;u)$ are matrix functions (with $H(\tau;u)$ invertible) satisfying

$$\begin{aligned} &\frac{d}{d\tau}\begin{pmatrix} G(\tau;u) & H(\tau;u)\end{pmatrix} \\ &= \begin{pmatrix} G(\tau;u) & H(\tau;u)\end{pmatrix} && (2.16)\\ &\times \begin{pmatrix} M^{T_{j+1}}(T_j-\tau) & -2Q^{\mathrm{T}}Q \\ \frac{1}{2}iu(iu-1)U_j(T_j-\tau)^2 & -\Big(M^{T_{j+1}}(T_j-\tau)^{\mathrm{T}} + 2iuU_j(T_j-\tau)RQ\Big)\end{pmatrix}, \end{aligned}$$

with initial conditions $G(0;u)=\mathbf{0}_n$ and $H(0;u)=\mathbf{I}_n$.

Proof: Please refer to Section B.2 of Appendix B. □

Common forms of forward volatility functions $u_j^{(i)}(t)$ in $U_j(t)$ can be found in Chapter 6 of Brigo and Mercurio [17]. For parametric forms, where the forward rate volatilities are time-dependent, the differential equation (2.16) has τ-dependent coefficients (which are continuous or piecewise continuous) and, thus, has to be solved by numerical ODE (means ordinary differential equation) solvers, such as Runge-Kutta methods (please refer to Section D.1 of Appendix D for the Runge-Kutta methods). Note that we choose to apply the numerical method on the linear differential equation (2.16) but not on the original matrix differential Riccati equation in the proof because a linear system has better numerical stability than for a non-linear system.

A widely-used parametric form of forward rate volatility function is given by

$$u_j^{(i)}(t) = \Big(a_i + b_i(T_j - t)\Big) \exp(-c_i(T_j - t)) + d_i, \tag{2.17}$$

where $a_i, b_i \in \mathbb{R}$, $c_i, d_i > 0$ and $a_i + d_i > 0$ for $i = 1, \ldots, n$, whose choice allows for the humped-shaped and time-homogeneous properties of term structure of forward rate volatility often observed in the market, for which a detailed discussion can be found in Chapter 6 of Rebonato [69].

Interestingly, for non-parametric forms, where $U_j(t)$ are piecewise-constant for $0 \leq t < T_j$, the differential equation (2.16) admits an explicit solution in the form of recursive matrix exponentials (some methods of computing matrix exponentials are given in Section D.2 of Appendix D) and, therefore, the pricing of caplet can easily be implemented, as illustrated in the following corollary.

Corollary 2.2.1. Suppose that the interval $[0, T_j]$ is partitioned by $\{(\tau_i, \tau_{i+1}]\}_{i=0}^{N-1}$, where $0 = \tau_0 < \tau_1 < \cdots < \tau_{N-1} < \tau_N = T_j$. Assume that for all $k = 1, \ldots, j$, on each $(\tau_i, \tau_{i+1}]$, $U_k(T_j - \tau)$ is a constant diagonal matrix, say, $U_k^{(i)}$.[2] Then, the solution for $A(\tau; u)$ at T_j is given by

$$A(T_j; u) = H(T_j; u)^{-1} G(T_j; u), \tag{2.18}$$

where $G(\tau; u)$ and $H(\tau; u)$ are matrix functions (with H invertible) satisfying, for $i =$

[2] $U_k^{(i)}$ is a zero matrix if the associated forward rate vanishes.

$0, \ldots, N-1$,

$$\begin{aligned} &\begin{pmatrix} G(\tau_{i+1};u) & H(\tau_{i+1};u) \end{pmatrix} \\ &= \begin{pmatrix} G(\tau_i;u) & H(\tau_i;u) \end{pmatrix} \\ &\times \exp\left((\tau_{i+1}-\tau_i) \begin{pmatrix} M_i^{T_{j+1}} & -2Q^{\mathrm{T}}Q \\ \frac{1}{2}iu(iu-1)\left(U_j^{(i)}\right)^2 & -\left((M_i^{T_{j+1}})^{\mathrm{T}} + 2iuU_j^{(i)}RQ\right) \end{pmatrix} \right), \end{aligned} \tag{2.19}$$

with initial conditions $G(\tau_0;u) = \mathbf{0}$ and $H(\tau_0;u) = \mathbf{I}_n$, where

$$M_i^{T_{j+1}} \triangleq M - \Big(\sum_{l=1}^{j} \frac{\Delta T_k L_k(0)}{1+\Delta T_k L_k(0)}\Big) Q^{\mathrm{T}} R^{\mathrm{T}} U_k^{(i)}. \tag{2.20}$$

The solution for $c(\tau;u)$ at T_j is given by

$$c(T_j;u) = -\frac{\beta}{2}\mathrm{Tr}\Big(\ln H(T_j) + \sum_{i=0}^{N-1}\Big((M_i^{T_{j+1}})^{\mathrm{T}} + 2iuU_j^{(i)}RQ\Big)(\tau_{i+1}-\tau_i)\Big). \tag{2.21}$$

Proof: Please refer to Section B.3 of Appendix B. □

Corollary 2.2.1 simplifies Proposition 2.2.2 by expressing the solution of the differential equation (2.16) into a recursive matrix exponential (2.19). It should be noted that the recursive solution is also applied by Jarrow *et al.* [51], Wu and Zhang [80], and Ma [61] to solve *scalar* Riccati equations for the characteristic function.

With the characteristic function of the log-forward rate obtained in Proposition 2.2.2, the corresponding probability density function can be computed by the inverse Fourier transform of the characteristic function. Figure 2.3 shows that, when correlation coefficients R_{11} and R_{22} are taken negative (resp. positive), the distribution of the logarithm of 5-year forward rate has fatter (resp. thinner) left tail and thinner (resp. fatter) right tail (when compared with the normal distribution), which corresponds to a downward

(resp. upward) sloping implied caplet volatility curve. Thus, the Wishart model is in line with other correlation-based models, such as Wu and Zhang [80], where downward sloping implied volatility curves are generated by introducing negative correlation coefficients between the shocks to the forward rate and to the variance dynamics.

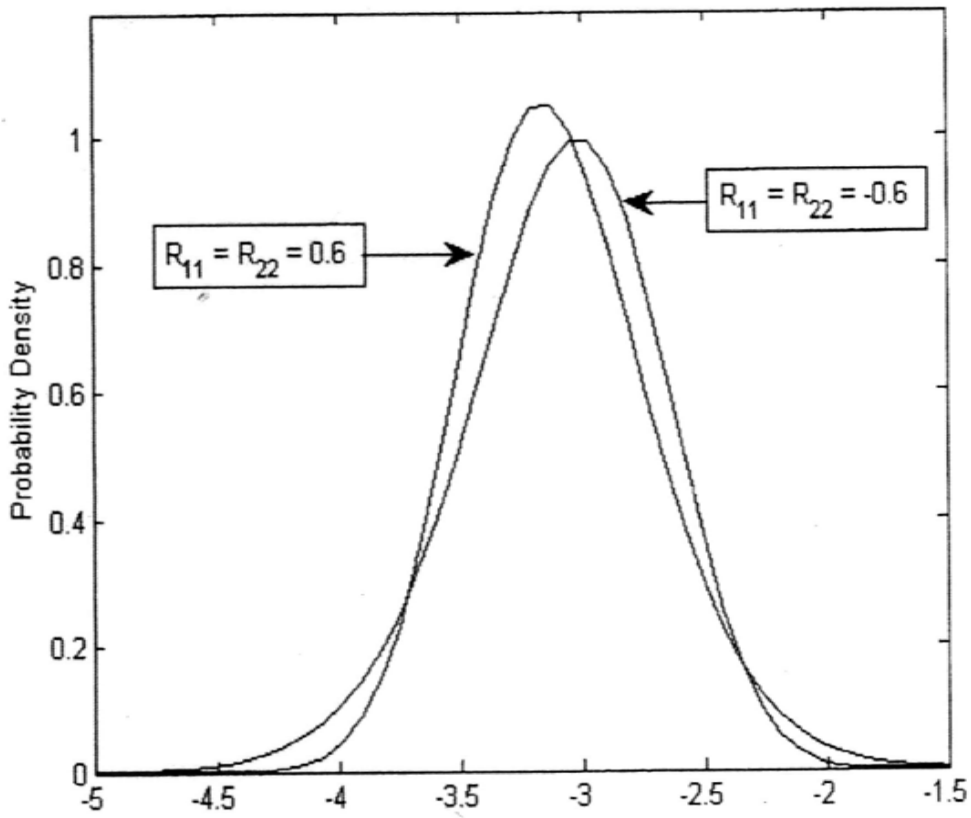

Figure 2.3: The probability density function of the logarithm of 5-year forward rate with different correlation coefficients. The parameter values are: $\beta = 5$, $M_{11} = -0.50$, $M_{22} = -0.05$, $Q_{11} = 0.40$, $Q_{22} = 0.10$, $R_{11} = R_{22} = \pm 0.60$, $\Sigma_0^{11} = \Sigma_0^{22} = 0.50$, $a_1 = a_2 = 0.01$, $b_1 = b_2 = 0.03$, $c_1 = c_2 = 0.30$ and $d_1 = d_2 = 0.13$.

2.2.2 The Pricing Formula for Caplets

Let $C(K, T_{j+1})$ be the value of caplet with maturity T_{j+1} and strike price K. We apply the and Madan citeCM99 approach to price caplets. The key is that the modified caplet price is given by

$$\begin{aligned}
\widetilde{C}(K, T_{j+1}) &= \exp(\alpha \ln K) C(K, T_{j+1}) \\
&= \Delta T_j P_{j+1}(0) \exp(\alpha \ln K) \mathbb{E}^{\mathbb{Q}^{T_{j+1}}} \left[\max\{L_j(T_j) - K, 0\} \right] \\
&= \Delta T_j P_{j+1}(0) \exp(\alpha \ln K) \int_K^\infty (e^s - K) q_j^{T_{j+1}}(s) ds,
\end{aligned}$$

where $q_j^{T_{j+1}}(s)$ is the probability density function of $\ln L_j(T_j)$ under $\mathbb{Q}^{T_{j+1}}$, whose Fourier transform is expressed in terms of the characteristic function of $\ln L_j(T_j)$ in Proposition 2.2.2. A positive parameter α is selected so that $\widetilde{C}(K, T_{j+1})$ is square-integrable in K, which guarantees the existence of its Fourier transform. The caplet price is obtained by applying inverse Fourier transform, which, after some calculations, is given below

$$C(K, T_{j+1}) = \Delta T_j P_{j+1}(0) \frac{e^{-\alpha \ln K}}{\pi} \int_0^\infty \frac{e^{-\mathrm{i}\xi \ln K} \Psi(x, \Sigma, T_j; \xi - (\alpha + 1)\mathrm{i})}{(\alpha + \mathrm{i}\xi)(\alpha + 1 + \mathrm{i}\xi)} d\xi, \qquad (2.22)$$

The above integral is approximated by Simpson's rule:

$$C(K, T_{j+1}) \approx \Delta T_j P_{j+1}(0) \frac{e^{-\alpha \ln K}}{\pi} \sum_{j=0}^{N-1} \frac{e^{-\mathrm{i}\xi_j \ln K} \Psi(x, \Sigma, T; \xi_j - (\alpha + 1)\mathrm{i})}{(\alpha + \mathrm{i}\xi_j)(\alpha + 1 + \mathrm{i}\xi_j)} w_j,$$

where N is the number of grid points for the interval $[0, T_j]$, $\eta = \frac{T_j}{N}$, $\xi_j = j\eta$ and

$$w_j = \begin{cases} \frac{1}{3}\eta & \text{for} \quad j = 0, N-1 \\ \frac{1}{3}(3 + (-1)^{j+1})\eta & \text{for} \quad j = 1, \ldots, N-2. \end{cases}$$

The rest of the implementation of the fast Fourier transform can be found in Carr and Madan [22] so it is omitted for brevity.

2.3 Simulation Study

To preserve the analytical tractability of the Wishart model, we adopt the technique of "freezing coefficients" in Equation (2.10). To investigate the accuracy of such approximation, we compare the caplet prices computed by formula (2.22) and by MC simulation for the case of $n = 2$ in the Wishart model. The procedure of the simulation is given in Section D.3 of Appendix D.

The initial term structure is the set of U.S. forward LIBORs on 12-25-2009. For the parameters of the forward rate volatility function in (2.17), for simplicity, we assume that, for $i = 1, 2$, $a_i = 0.01$, $b_i = 0.03$, $c_i = 0.30$ and $d_i = 0.13$. For the parameters of the Wishart process, we set $\beta = 5$,

$$M = \begin{pmatrix} -0.50 & 0.00 \\ 0.00 & -0.05 \end{pmatrix} \quad Q = \begin{pmatrix} 0.40 & 0.05 \\ 0.05 & 0.10, \end{pmatrix}$$

$$R = \begin{pmatrix} -0.40 & -0.20 \\ -0.20 & -0.40 \end{pmatrix} \quad \Sigma_0 = \begin{pmatrix} 0.50 & 0.20 \\ 0.20 & 0.50 \end{pmatrix}.$$

We use Simpson's rule to compute the integral in (2.22) with 128 points and grid size 0.50 and apply the sixth-order Runge-Kutta method (please refer to Section D.1 of Appendix D) with stepsize of 1/24 to solve the differential equation (2.16); on the other hand, we carry out the MC simulation with 100,000 sample paths and time step of 1/24. Table 2.1 reports the caplet prices (in basis points), CPU times and percentage pricing error produced by analytical formula (2.22) (labeled as "AF") and by MC simulation (labeled as "MC"), respectively. In the table, for all moneyeness levels and maturities, the percentage errors between AF prices and MC prices are within 1%. In addition, the CPU times for computing AF prices are much less than those for MC prices. Thus, the simulation study

demonstrates the accuracy and efficiency of the analytical formula.

Table 2.1: The caplet prices (in basis points) and CPU times produced by the analytical formula and MC simulation and their percentage pricing error.

	T=1			T=2			T=3			T=4			T=5		
Moneyness	AF	MC	% Error	AF	MC	% Error	AF	MC	% Error	AF	MC	% Error	AF	MC	% Error
0.6	10.94	10.95	-0.09	24.85	24.80	0.20	35.01	34.97	0.12	40.63	40.40	0.57	44.38	44.36	0.03
0.7	8.212	8.221	-0.11	18.95	18.90	0.28	27.27	27.23	0.17	32.26	32.05	0.66	35.81	35.80	0.03
0.8	5.561	5.569	-0.15	13.54	13.49	0.39	20.36	20.32	0.22	24.88	24.71	0.70	28.29	28.27	0.05
0.9	3.193	3.201	-0.26	8.949	8.909	0.45	14.53	14.50	0.23	18.64	18.50	0.73	21.90	21.89	0.06
1.0	1.450	1.456	-0.45	5.425	5.406	0.35	9.915	9.897	0.17	13.58	13.48	0.72	16.64	16.66	-0.08
1.1	0.494	0.497	-0.68	3.009	3.001	0.25	6.476	6.464	0.19	9.646	9.580	0.69	12.44	12.47	-0.22
1.2	0.124	0.126	-0.94	1.532	1.527	0.28	4.063	4.057	0.15	6.696	6.660	0.53	9.168	9.203	-0.38
1.3	2.36E-02	2.38E-02	-0.82	0.721	0.718	0.40	2.459	2.462	-0.11	4.555	4.546	0.20	6.671	6.711	-0.59
1.4	3.56E-03	3.55E-03	0.35	0.317	0.315	0.49	1.442	1.451	-0.60	3.046	3.059	-0.43	4.803	4.843	-0.83
Total CPU Time (sec)	1.16	277		2.07	549		3.36	888		4.67	1244		5.79	1616	

	T=6			T=7			T=8			T=9			T=10		
Moneyness	AF	MC	% Error	AF	MC	% Error	AF	MC	% Error	AF	MC	% Error	AF	MC	% Error
0.6	44.66	44.48	0.40	43.49	43.46	0.07	39.04	39.03	0.04	45.23	44.97	0.59	40.73	40.76	-0.08
0.7	36.51	36.35	0.44	35.96	35.94	0.03	32.57	32.55	0.05	38.04	37.81	0.59	34.48	34.54	-0.17
0.8	29.38	29.25	0.45	29.36	29.36	0.00	26.91	26.89	0.07	31.73	31.57	0.51	29.01	29.10	-0.30
0.9	23.28	23.21	0.39	23.71	23.72	-0.07	22.04	22.04	0.02	26.30	26.19	0.42	24.29	24.39	-0.43
1.0	18.24	18.18	0.29	18.96	18.99	-0.17	17.93	17.95	-0.08	21.69	21.61	0.35	20.24	20.35	-0.55
1.1	14.11	14.09	0.14	15.05	15.09	-0.27	14.50	14.53	-0.23	17.80	17.76	0.23	16.82	16.93	-0.66
1.2	10.81	10.82	-0.03	11.86	11.91	-0.43	11.67	11.71	-0.36	14.56	14.55	0.04	13.94	14.04	-0.74
1.3	8.218	8.228	-0.12	9.289	9.347	-0.62	9.349	9.396	-0.50	11.87	11.90	-0.19	11.53	11.63	-0.80
1.4	6.203	6.215	-0.19	7.244	7.312	-0.93	7.469	7.518	-0.66	9.665	9.714	-0.50	9.527	9.613	-0.90
Total CPU Time (sec)	7.56	2007		9.09	2411		10.32	2837		12.49	3279		14.31	3738	

2.4 Model Properties

In this section, we demonstrate the flexibility of the model on modeling the stochastic skew over the multi-factor Heston model.

2.4.1 Stochastic Correlation

Under the framework of stochastic volatility, the implied volatility skew at time t generated by the model critically depends on the instantaneous correlation between the shocks to the forward rate $L_j(t)$ and the variance dynamics $\mathrm{Tr}(\Sigma_t)$, which, in the Wishart model, is given by

$$\rho_t^{\text{Wis}} = \frac{\mathrm{Tr}\Big(U_j(t)RQ\Sigma_t\Big)}{\sqrt{\mathrm{Tr}\Big(U_j(t)^2\Sigma_t\Big)}\sqrt{\mathrm{Tr}\Big(Q^{\mathrm{T}}Q\Sigma_t\Big)}}. \tag{2.23}$$

Consider the case that $n = 2$. If we take matrices RQ and Σ_t to be diagonal in (2.23), we obtain the instantaneous correlation of the two-factor Heston model, denoted as $\rho_t^{\text{2SV-Hes}}$, which is shown to be stochastic through the variance dynamics Σ_t^{11} and Σ_t^{22}. Thus, a two-factor (also multi-factor) Heston model generates stochastic skew through its variance dynamics. For our model, if we take Σ_t to be a 2×2 symmetric positive definite matrix and R to be upper triangular in (2.23), we obtain the instantaneous correlation, denoted as $\rho_t^{\text{2D-Wis}}$, which is related with $\rho_t^{\text{2SV-Hes}}$ as follow

$$\rho_t^{\text{2D-Wis}} = \rho_t^{\text{2SV-Hes}} + \frac{u_j^{(1)}(t)R_{12}Q_{22}}{\sqrt{(u_j^{(1)}(t))^2\Sigma_t^{11} + (u_j^{(2)}(t))^2\Sigma_t^{22}}\sqrt{Q_{11}^2\Sigma_t^{11} + Q_{22}^2\Sigma_t^{22}}}\Sigma_t^{12}. \tag{2.24}$$

The above relation shows that the state variable Σ_t^{12} and parameter R_{12}, which act independently from the volatility factors Σ_t^{11} and Σ_t^{22}, offer extra flexibility to control the

term structure of skew. This flexibility is crucial when Σ_t^{11} and Σ_t^{22} have been fitted to short-term and long-term volatilities.

2.4.2 Numerical Example

To illustrate the fact that the dynamics of Σ_t^{12} grants extra freedom to control the volatility skew, we perform a numerical experiment for the case $n = 2$. The initial term structure and the parameters of the forward rate volatility functions are the same as those given in the simulation study; however, the parameters of the Wishart process are given by: $\beta = 5$,

$$M = \begin{pmatrix} -0.05 & 0.00 \\ 0.00 & -0.50 \end{pmatrix} \quad Q = \begin{pmatrix} 0.20 & 0.00 \\ 0.00 & 0.20 \end{pmatrix},$$

$$R = \begin{pmatrix} -0.40 & R_{12} \\ 0.00 & -0.40 \end{pmatrix} \quad \Sigma_0 = \begin{pmatrix} 0.50 & \Sigma_0^{12} \\ \Sigma_0^{12} & 0.50 \end{pmatrix}.$$

In order to show that the process Σ_t^{12} and parameter R_{12} offer the Wishart model extra flexibility to control the volatility skew when compared with the two-factor Heston model, we consider three different cases, where the first case is $(R_{12}, \Sigma_0^{12}) = (-0.60, 0.40)$, the second case is $(R_{12}, \Sigma_0^{12}) = (-0.60, -0.40)$ and the third case is $(R_{12}, \Sigma_0^{12}) = (0, 0)$. From Remark 2.2.1 and Equation (2.24), we refer the case that $(R_{12}, \Sigma_0^{12}) = (0, 0)$ to the two-factor Heston model.

Figure 2.4 shows the implied caplet volatility curves generated by the three cases, where the lines with *crosses*, *stars* and *dots* represent the curves generated by the Wishart model with $(R_{12}, \Sigma_0^{12}) = (-0.60, 0.40)$, $(R_{12}, \Sigma_0^{12}) = (-0.60, -0.40)$ and $(R_{12}, \Sigma_0^{12}) = (0, 0)$ (the

two-factor Heston model), respectively. Panels 1 to 5 show that when compared with the two-factor Heston model with $(R_{12}, \Sigma_0^{12}) = (0, 0)$, the implied caplet volatility skews for the Wishart model with $(R_{12}, \Sigma_0^{12}) = (-0.60, 0.40)$ (resp. $(R_{12}, \Sigma_0^{12}) = (-0.60, -0.40)$) are more (resp. less) negative, indicating that the Wishart model offers extra flexibility to control volatility skew when compared with the two-factor Heston model. Note that the turning point of the volatility curves is almost exactly on the at-the-money level, indicating that the volatility level is not affected by varying Σ_0^{12}. On the other hand, Panels 6 to 10 show that the volatility curves generated by the three cases almost coincide, indicating again that setting Σ_0^{12} non-zero has negligible effect on the volatility level, and the effect of Σ_t^{12} on volatility skew is diminishing as the Wishart process is mean-reverting. As an interpretation to the above experiment, suppose that, to estimate the models by a time series of data, for the two-factor Heston model, if the set of model parameters is fixed, it has a good fit to the volatility level by the processes Σ_t^{11} and Σ_t^{22}; however, it contains no other process to control or fit the volatility skew; therefore, changing the value of Σ_t^{11} and Σ_t^{22} must also change the volatility skew and vice versa. It makes the control of volatility level and volatility skew difficult. In contrast, for the Wishart model, once the processes Σ_t^{11} and Σ_t^{22} are fitted to volatility level, it contains another dynamics Σ_t^{12}, which does not affect the volatility level, that allows us to better fit the volatility skew. This property makes the Wishart model advantageous in pricing interest rate caps when the implied caplet volatility skews are stochastic.

Figure 2.5 shows that the term structure of skew of the Wishart model with $(R_{12}, \Sigma_0^{12}) = (-0.60, 0.40)$ (resp. $(R_{12}, \Sigma_0^{12}) = (-0.60, -0.40)$) is much steeper (resp. flatter) than that of the two-factor Heston model with $(R_{12}, \Sigma_0^{12}) = (0, 0)$. Thus, the numerical experiment

suggests that the Wishart model has extra flexibility to control volatility skew when compared with the two-factor Heston model.

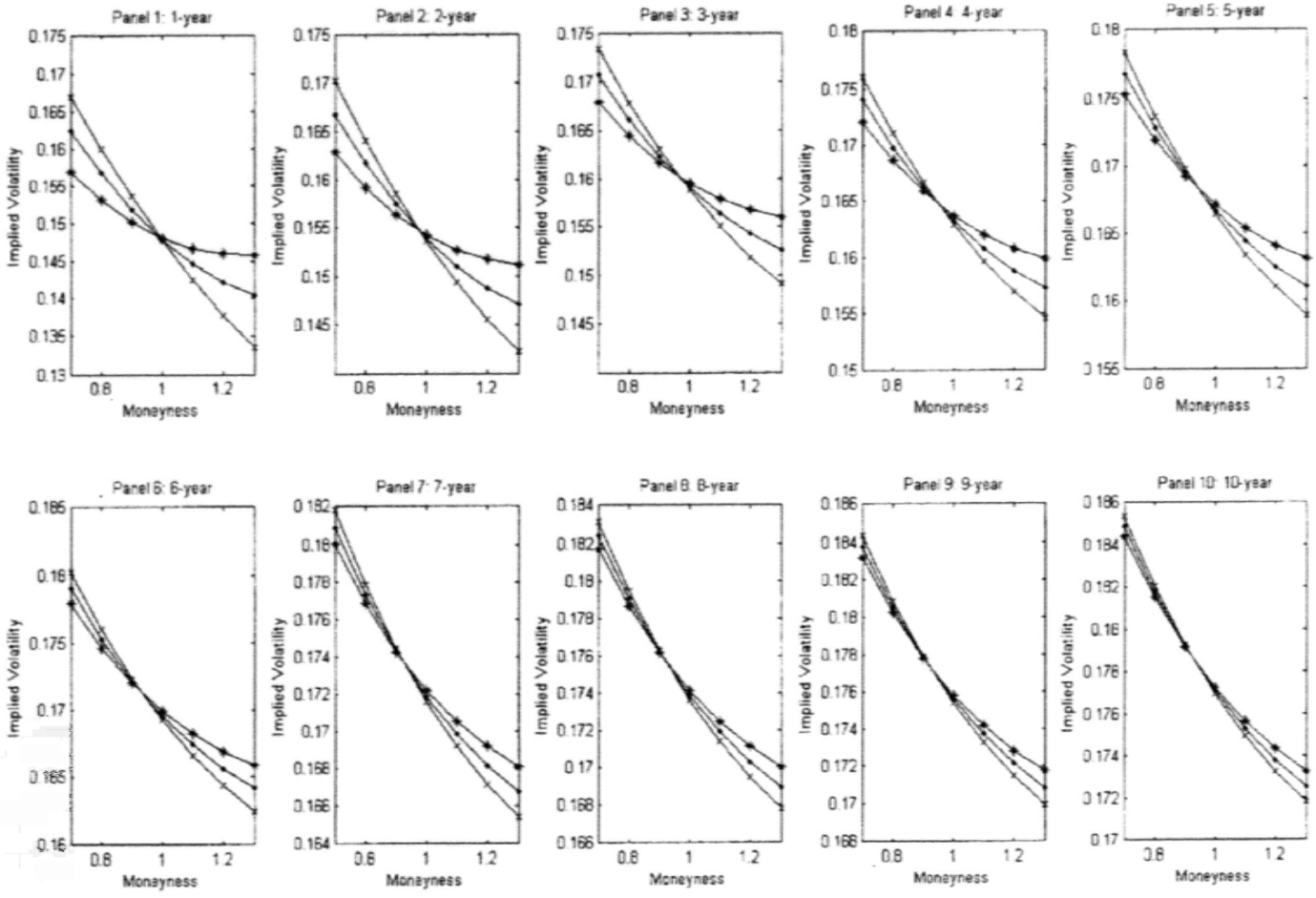

Figure 2.4: The implied caplet volatility curves generated by the Wishart model with $n = 2$ (lines with *crosses* represent the case $(R_{12}, \Sigma_0^{12}) = (-0.60, 0.40)$ and lines with *stars* represent the case $(R_{12}, \Sigma_0^{12}) = (-0.60, -0.40)$) and the two-factor Heston model (lines with *dots* represent the case $(R_{12}, \Sigma_0^{12}) = (0, 0)$).

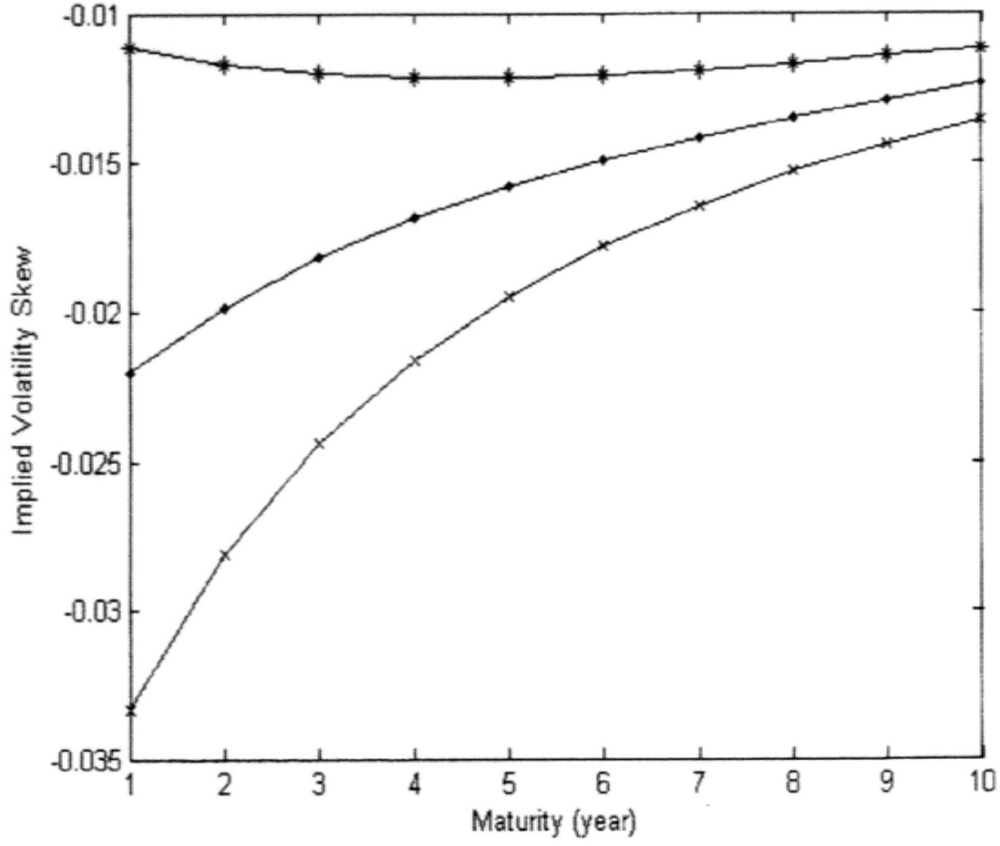

Figure 2.5: The term structure of implied caplet volatility skew generated by the Wishart model with $n = 2$ (lines with *crosses* represent the case $(R_{12}, \Sigma_0^{12}) = (-0.60, 0.40)$ and lines with *stars* represent the case $(R_{12}, \Sigma_0^{12}) = (-0.60, -0.40)$) and the two-factor Heston model (lines with *dots* represent the case $(R_{12}, \Sigma_0^{12}) = (0, 0)$).

2.5 Model Calibration

In this section, we calibrate the two-dimensional Wishart model and the two-factor Heston model to the forward LIBORs and cap prices reported on 02-Dec-2005 by minimizing the sum of squared error of model implied and market caplet volatilities over all maturities and moneyness levels. To highlight the importance of the state variable Σ_t^{12} and parameter R_{12} on modeling stochastic skew, we take R to be an upper triangular matrix and restrict matrices M and Q to be diagonal. For both models, we adopt the parametric form in (2.17) for $i = 1, 2$ to model the term structure of caplet volatility.

Figure 2.6 shows the quality of fitting in terms of Black's implied caplet volatilities. In particular, it shows that both models match well to data. Calibrated parameters and pricing errors of both models are reported in Table 2.2, where the Wishart model produces about 27% improvement in terms of sum of squared error and about 13% in terms of average absolute percentage error[3] over the two-factor Heston model.

In short, this empirical study shows that the two-dimensional Wishart model performs better than the two-factor Heston model in fitting the implied volatility surface in one day.

[3]The average absolute percentage error is the average of the absolute value of the percentage error of model implied and market caplet volatilities over all maturities and moneyness levels.

Table 2.2: The calibrated parameters, sum of squared error and the average absolute percentage of pricing error of the two-dimensional Wishart model and the two-factor Heston model on 02-Dec-2005.

Parameters	2D-Wishart	2SV-Heston
β	1.8346	1.8512
M_{11}	-0.2522	-0.2304
M_{22}	-1.3833	-6.2181
Q_{11}	0.2004	0.3271
Q_{22}	0.1488	0.2967
R_{11}	-0.9990	-0.9558
R_{12}	0.1489	—
R_{22}	0.7527	-0.7669
Σ_0^{11}	0.0471	0.2131
Σ_0^{12}	-0.1799	—
Σ_0^{22}	0.7356	0.7087
a_1	-0.1542	-0.0064
b_1	0.8944	0.4150
c_1	0.4678	0.3771
d_1	0.1101	0.0244
a_2	-0.4300	0.1191
b_2	0.2593	-0.8780
c_2	0.4557	0.3300
d_2	0.3950	0.0668
Sum of Squared Error	1.1076×10^{-4}	1.5189×10^{-4}
Average Absolute % Error	0.4947%	0.5714%

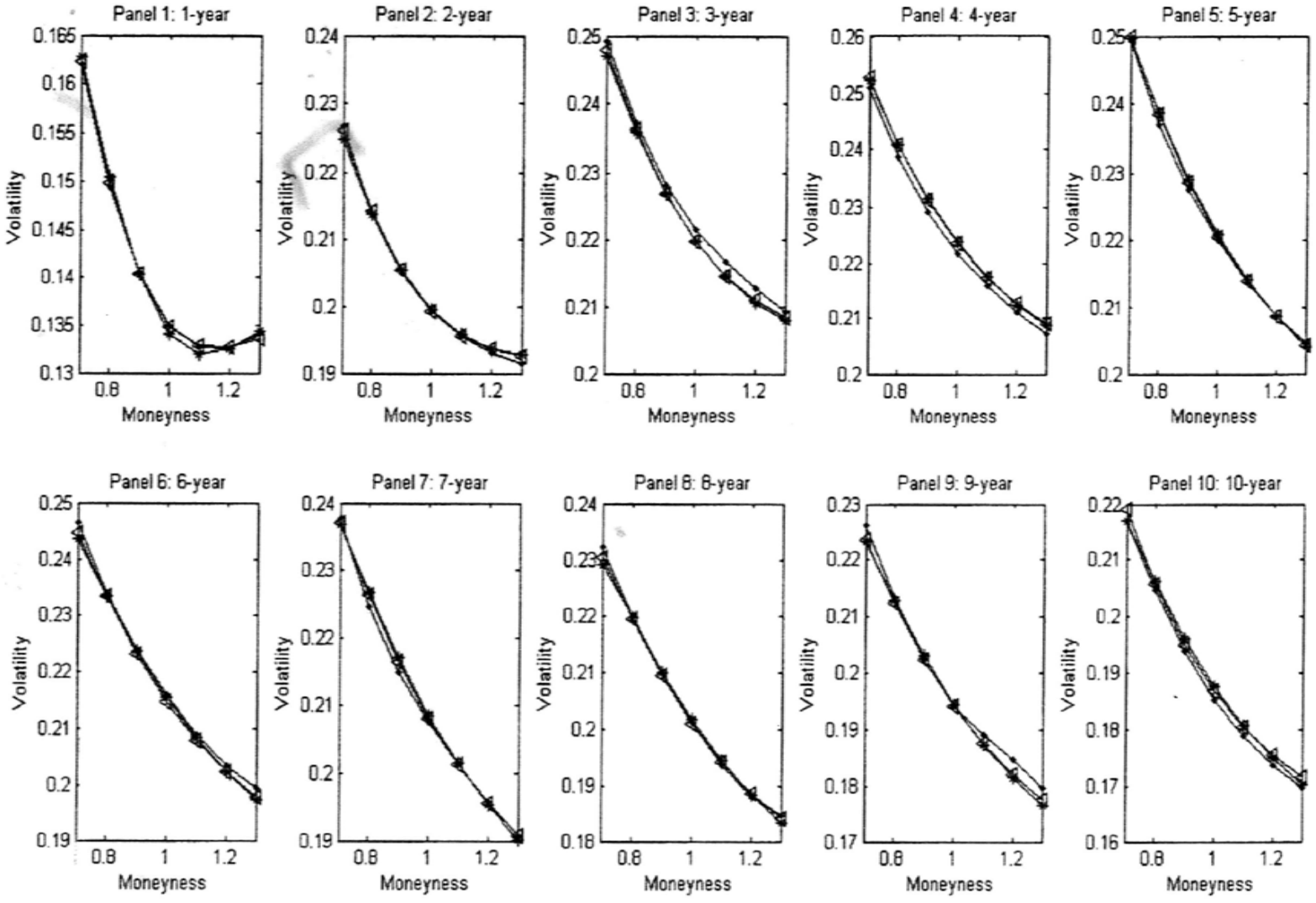

Figure 2.6: The implied caplet volatility curves for the calibrated two-dimensional Wishart model (lines with *triangles*) and the calibrated two-factor Heston model (lines with *stars*) compared with market volatility curves (lines with *dots*) on 02-Dec-2005.

2.6 Model Implementation

In Section 2.1, we observe that the implied caplet volatility skew is stochastic in nature. In the Section 2.4, we also show that the Wishart model possesses greater flexibility than the multi-factor Heston model on modeling stochastic skew. In this section, we implement the Wishart model with $n = 2$ and compare its performance with that of the two-factor Heston model.

2.6.1 Covariance Structure of Forward Rates

We estimate the covariance structure of forward rates using the approach of Han [45], Jarrow *et al.* [51] and Longstaff *et al.* [60].

Referring to Equation (2.4), under measure $\mathbb{Q}$, since, for $j = 1, \ldots, N$,

$$\frac{dL_j(t)}{L_j(t)} = (\cdots)dt + \mathrm{Tr}\Big(U_j(t)\sqrt{\Sigma_t}dZ_t\Big),$$

the covariance between the changes in $L_j(t)$ and $L_k(t)$, for $j, k = 1, \ldots, N$ is given by

$$\mathrm{Cov}^{\mathbb{Q}}\Big(\frac{dL_j(t)}{L_j(t)}, \frac{dL_k(t)}{L_k(t)}\Big|\mathcal{F}_t\Big) = \Big(\sum_{i=1}^{n} u_j^{(i)}(t)u_k^{(i)}(t)\Sigma_t^{ii}\Big)dt,$$

Therefore, the model instantaneous covariance matrix with fixed *maturities* is given by

$$\Omega_t^{\text{Maturity}} = U_t\Lambda_t U_t^{\mathrm{T}},$$

where

$$\begin{aligned} U_t &= \begin{pmatrix} u_1^{(1)}(t) & \cdots & u_1^{(n)}(t) \\ \vdots & & \vdots \\ u_N^{(1)}(t) & \cdots & u_N^{(n)}(t) \end{pmatrix} \\ \Lambda_t &= \mathbf{diag}(\Sigma_t^{11}, \ldots, \Sigma_t^{nn}). \end{aligned}$$

At time 0, suppose that the historical unconditional covariance of changes in forward rates with fixed *time-to-maturities* is approximated as $H = U\Lambda_0 U^{\mathrm{T}}$, where $\Lambda_0 \in \mathbb{R}^{n\times n}$ is a diagonal matrix whose diagonal elements are the first n largest eigenvalues in descending order, and the columns of $U \in \mathbb{R}^{N\times n}$ are the corresponding eigenvectors with unit norm.

Then, we assume that the model instantaneous covariance matrix with fixed *time-to-maturities* shares the same eigenvectors as the historical unconditional covariance matrix H, that is,

$$\Omega_t^{\text{Time-to-maturity}} = U\Lambda_t U^{\mathrm{T}}.$$

To model stochastic covariance structure, what we need is the instantaneous covariance matrix of changes in forward rates with fixed *maturities*, $\Omega_t^{\text{Maturity}}$. Therefore, at time 0, $\Omega_0^{\text{Maturity}} = \Omega_0^{\text{Time-to-maturity}}$. For $t > 0$, we obtain $\Omega_t^{\text{Maturity}}$ from $\Omega_t^{\text{Time-to-maturity}}$ through the *time-homogeneity property* (see, for example, Chapter 6 of Rebonato [69]): as time goes by, whenever a forward rate vanishes, we discard the last row of the matrix U (which represents the term structure factors of the forward rate with the longest time-to-maturity) of U, that is, for $t \in [T_k, T_{k+1})$,

$$U_t = \begin{pmatrix} U_{1,1} & \cdots & U_{1,n} \\ \vdots & & \vdots \\ U_{N-k,1} & \cdots & U_{N-k,n} \end{pmatrix}$$

For example, during the first 3-month period, the N forward rates are living and the full version of U is used. After the first 3-month period, the first forward rate L_1 vanishes at $t = T_1$, leaving only $(N-1)$ forward rates living; as a result, due to time-homogeneity, the relevant version of U_t is given by the first $(N-1)$ rows of the original U. This process is repeated until the last 3-month period, that is, the last forward rate. Implicitly, we make the assumption that the covariances are constant within adjacent reset dates, that

is, the dynamics of forward rates are given by, for $0 \leq t < T_1$ and $j = 1, \ldots, N$,

$$\frac{dL_j(t)}{L_j(t)} = (\cdots)dt + \mathrm{Tr}\Big(\mathrm{diag}(U_{j,1}, \ldots, U_{j,n})\sqrt{\Sigma_t}dZ_t\Big).$$

For $T_k \leq t < T_{k+1}$ and $j = k+1, \ldots, N$,

$$\frac{dL_j(t)}{L_j(t)} = (\cdots)dt + \mathrm{Tr}\Big(\mathrm{diag}(U_{j-k,1}, \ldots, U_{j-k,n})\sqrt{\Sigma_t}dZ_t\Big).$$

Therefore, each $U_j(t)$, for $j = 1, \ldots, N$, is a *piecewise constant* function of t.

2.6.2 Estimation Procedure

For the covariance structure of forward rates, we estimate it from the changes in forward rates from the beginning of the sample using a window of one year and roll over the procedure as time goes by. To estimate the model parameters Θ and the volatility variables Σ_t, we minimize the sum of squared *implied volatility error*, instead of the sum of squared percentage pricing error, since the quotation convention in the interest rate cap market is in terms of volatilities. Consider a sample of T weeks of forward rates and caplet volatilities, where in week t we have N_t caplet volatilities. Denote $\mathrm{IV}^{\mathrm{Market}}(t, K_i, T_i)$ as the market caplet volatility in week t of the i-th caplet with strike price K_i and maturity T_i and $\mathrm{IV}^{\mathrm{Model}}(t, K_i, T_i; \Theta, \Sigma_t)$ as the model implied caplet volatility in week t of the i-th caplet with strike price K_i and maturity T_i, given Θ and Σ_t.

In the first step, for a given set of structural parameters $\Theta = \{\beta, M, Q, R\}$, for $t = 1, \ldots, T$, Σ_t is given by

$$\Sigma_t^* = \arg\ \min \sum_{i=1}^{N_t} (\mathrm{IV}^{\mathrm{Market}}(t, K_i, T_i) - \mathrm{IV}^{\mathrm{Model}}(t, K_i, T_i; \Theta, \Sigma_t))^2$$

Since, for each day, we have 7 moneyness levels (from 0.7 to 1.3) and 10 maturities (from 1 year to 10 years), $N_t = 70$. Then, for the set of state variables $\{\Sigma_t\}_{t=1}^T$ obtained, we

estimate the model parameters Θ by minimizing the sum of squared implied volatility error over the entire sample

$$\Theta^* = \arg \min \sum_{t=1}^{T} \sum_{i=1}^{N_t} (\mathrm{IV}^{\mathrm{Market}}(t, K_i, T_i) - \mathrm{IV}^{\mathrm{Model}}(t, K_i, T_i; \Theta, \Sigma_t^*))^2.$$

Finally, we evaluate the model fit by using the implied volatility root mean squared error (IVRMSE)

$$\mathrm{IVRMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^{T} \sum_{i=1}^{N_t} (\mathrm{IV}^{\mathrm{Market}}(t, K_i, T_i) - \mathrm{IV}^{\mathrm{Model}}(t, K_i, T_i; \Theta^*, \Sigma_t^*))^2},$$

where $N = \sum_{t=1}^{T} N_t$.

2.6.3 Results

For the two-factor Heston model, it corresponds to the Wishart model with $n = 2$ and the matrices M, Q, R and Σ_t being diagonal; for the Wishart model, we restrict matrices M and Q to be diagonal and R to be upper-triangular in order to investigate the effect on pricing of the addition of the skew-related components (R_{12} and Σ_t^{12} in the Wishart model) over the two-factor Heston model. The in-sample estimation is carried out for the 3-year period, from Jan-2004 to Dec-2006, while the out-of-sample estimation is carried out for the following 3-year period, namely, from Jan-2007 to Dec-2009. The out-sample results are obtained by computing the volatility variables by using the structural parameters estimated in the in-sample period. The estimation results for both models are reported in Table 2.3. In in-sample period and out-of-sample period, the 2-D Wishart model improves on the two-factor Heston model by 8.57% and 5.88%, respectively, in terms of IVRMSE. Since the Wishart model contains one more process Σ_t^{12} and one more parameter R_{12}, which act independently from the volatility factors, than the two-factor Heston model,

Table 2.3: The estimation results for the two-dimensional Wishart model and the two-factor Heston model.

Parameters	2D-Wis	2SV-Hes
β	0.1865	0.0004
M_{11}	-0.9554	-0.6793
M_{22}	-0.0036	-0.0036
Q_{11}	1.2740	1.9998
Q_{22}	1.2289	0.2115
R_{11}	-0.1249	-0.0693
R_{12}	-0.9989	—
R_{22}	-0.2245	-0.9999
In-sample IVRMSE ($\times 10^{-2}$)	3.2140	3.5154
Out-of-sample IVRMSE ($\times 10^{-2}$)	3.5801	3.8039

better pricing performance of the Wishart model over the two-factor Heston model can be attributed to the fact that, given a fixed set of model parameters, the time variation of the volatility-unrelated process in the Wishart model provides a degree of freedom to better capture stochastic skew over time.

2.7 Summary

We propose a multi-factor term structure model to capture an essential feature observed in interest rate cap market, namely, stochastic skew, and extend the benchmark LIBOR market model by incorporating the Wishart process. Analytical pricing formula for caplets is derived and expressed in terms of characteristic functions. MC simulation shows that the Wishart model is accurate and efficient for practical uses. The essence of

the Wishart model is that while the diagonal elements of the Wishart process generate volatility smile/skew over maturities, the non-diagonal elements offers extra flexibility to control the volatility skew. In contrast, while the multi-factor Heston model has been fixed for long-term and short-term volatilities, it does not have any other state variable or parameter to control volatility skew. Through numerical example, the flexibility of modeling stochastic skew of the Wishart model is demonstrated to be superior to the multi-factor Heston model. A calibration exercise is carried out to show that the two-dimensional Wishart model performs better than the two-factor Heston model in fitting the implied volatility surface in one day. Finally, we estimate the two-dimensional Wishart model and the two-factor Heston model with a time series of implied volatility surfaces and results shows that the two-dimensional Wishart model outperforms the two-factor Heston model, where the better pricing performance of Wishart model over the Heston model can be attributed to the better modeling of stochastic skew by the Wishart model.

Chapter 3

Stochastic Skew in Currency Option Market

This chapter provides the empirical evidence of stochastic skew in the EUR/USD currency option market. To model mean-reversion, stochastic volatility and stochastic skew, we formulate the currency option pricing model with the Wishart process and derive the pricing formula of vanilla currency options, whose implementation is provided and contrast with MC simulation. Also, we extends the model to include pure jump component with Poisson arrival and jump size of arbitrary distribution, generalizing the single-volatility framework of O'Hara and Pillay [64]. Numerical example with market term structure of futures prices shows that the Wishart model provides additional freedom to better capture stochastic skew than the multi-factor Heston model.

3.1 Empirical Evidence

In this section, we provide a description of data and methodology of evidencing stochastic skew in the EUR/USD currency option market.

To show the presence of stochastic skew in currency option market, we collect quotes of the dollar price and the currency option of Euros from 1-Jan-2008 to 31-Oct-2010. The option quotes have 8 fixed time to maturities, namely, 1, 2, 3, 6, 9, 12, 18 and 24 months. We also collect quotes of 10-delta risk reversal (R10) of all maturities, which is the difference between the implied volatilities of a 10-delta call option and a 10-delta put option. Hence, the risk reversal is a measure of the slope of the implied volatility curve. Figure 3.1 shows that, for all maturities, the risk reversal (skew) is stochastically time varying, similar to the observation of Carr and Wu [21]. Hence, the empirical evidence strongly suggests that a currency option pricing model should be flexible enough to capture not only mean-reversion and stochastic volatility, but also stochastic skew.

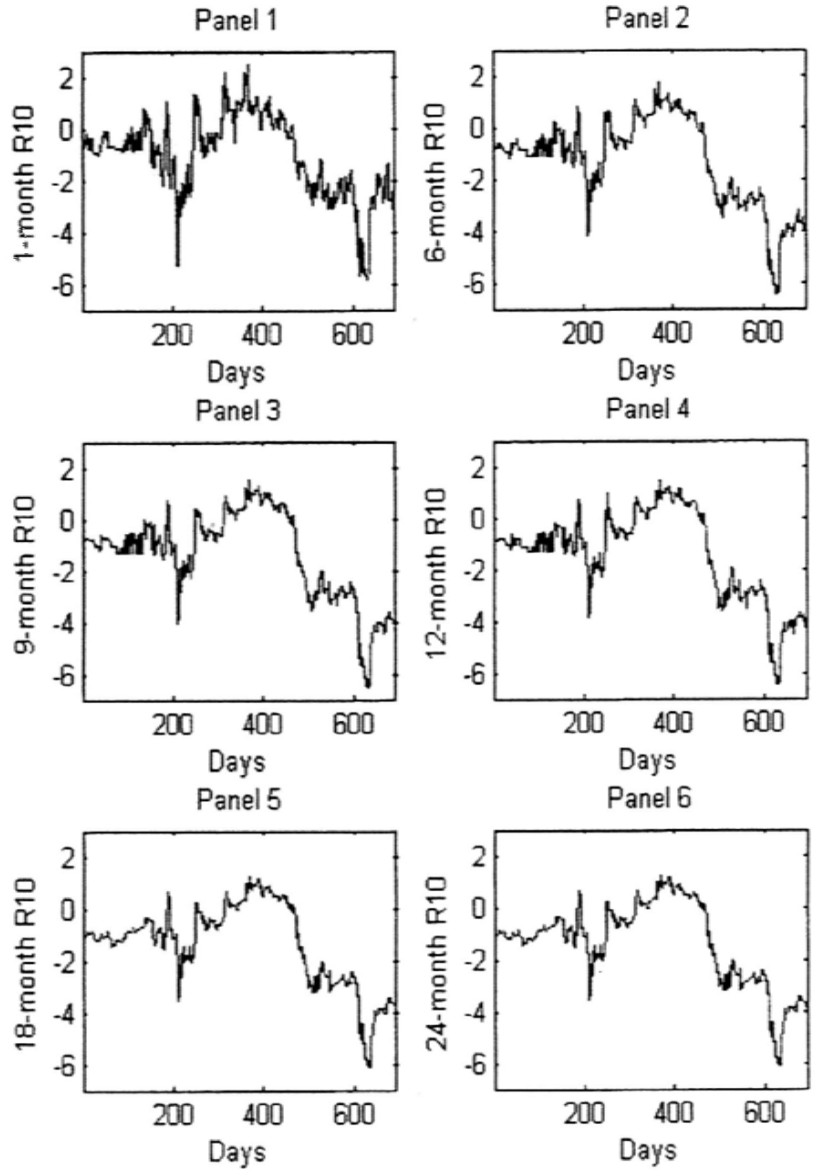

Figure 3.1: The time series of 1, 6, 9, 12, 18, 24-month 10-delta risk reversals from 1-Jan-2008 to 31-Oct-2010.

3.2 The Model

In this section, we generalize the two-factor Heston model of Wong and Zhao [77] by making use of the Wishart process. We derive the characteristic function of log-currency value with the futures-price calibrated characteristic function and provide the pricing formula for European currency call option.

With the same notations for matrix operations as in the previous chapter and the Wishart process given by the dynamics in (2.1), let S_t be the underlying currency for which the risk-neutral process is postulated as

$$\begin{aligned} S_t &= \exp\Big(X_t\Big) \\ dX_t &= \Big(\theta(t) - \kappa X_t - \frac{1}{2}\mathrm{Tr}\Big(\Sigma_t\Big)\Big)dt + \mathrm{Tr}\Big(\sqrt{\Sigma_t}dZ_t\Big) \\ d\Sigma_t &= \Big(\beta Q^{\mathrm{T}}Q + M\Sigma_t + \Sigma_t M^{\mathrm{T}}\Big)dt + \sqrt{\Sigma_t}dW_t Q + Q^{\mathrm{T}}(dW_t)^{\mathrm{T}}\sqrt{\Sigma_t}, \end{aligned} \tag{3.1}$$

where the constant κ is the mean-reversion speed for the log-currency value, the deterministic function $\theta(t)$ represents the equilibrium mean level of the log-currency-value against time and the correlation structure between W_t and Z_t is given in (2.3) in the previous chapter

$$Z_t = W_t R^{\mathrm{T}} + B_t\sqrt{\mathbf{I}_n - RR^{\mathrm{T}}},$$

where B_t is a matrix Brownian motion independent of W_t.

Remark: The proposed model embraces many existing models as special cases. When $n = 1$, it is reduced to the model of Wong and Lo [76]; when $n = 2$ and the matrices M, Q and R are diagonal, it is reduced to the model of Wong and Zhao [77]; further, when

$\theta(t) \equiv r$, the risk-free rate, and $\kappa = 0$, the proposed model is reduced to the two-factor Heston model of Christoffersen *et al.* [25].

3.2.1 The Characteristic Function

As in the previous chapter, we exploit the affine nature of the Wishart process to compute the characteristic function of the log-currency value to price currency options. Denote the characteristic function of the log-currency value X_T as

$$\Psi(x, \Sigma, \tau; u) \triangleq \mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(\mathrm{i}uX_T\Big)\Big|X_t = x, \Sigma_t = \Sigma\Big], \tag{3.2}$$

where $\mathbb{E}^{\mathbb{Q}}\Big[\cdot\Big]$ denotes the expectation under the probability measure $\mathbb{Q}$, $\tau = T - t \geq 0$ and $\mathrm{i} = \sqrt{-1}$. The following proposition holds.

Proposition 3.2.1. If X_t follows the dynamics in (3.1), then the characteristic function for X_T in (3.2) is given by

$$\Psi(x, \Sigma, \tau; u) = \exp\Big(\mathrm{Tr}\Big(A(\tau; u)\Sigma\Big) + b(\tau; u)x + c(\tau; u)\Big), \tag{3.3}$$

where the functions $A(\tau; u)$, $b(\tau; u)$ and $c(\tau; u)$ satisfy

$$A(\tau; u) = H(\tau; u)^{-1}G(\tau; u) \tag{3.4}$$

$$b(\tau; u) = \mathrm{i}ue^{-\kappa\tau} \tag{3.5}$$

$$\begin{aligned} c(\tau; u) &= \mathrm{i}u\int_0^\tau \theta(T-s)e^{-\kappa s}ds \\ &\quad -\frac{\beta}{2}\mathrm{Tr}\Big(\ln H(\tau; u) + M^{\mathrm{T}}\tau + \frac{2\mathrm{i}u}{\kappa}(1 - e^{-\kappa\tau})RQ\Big), \end{aligned} \tag{3.6}$$

where $G(\tau;u)$ and $H(\tau;u)$ are matrix functions (with $H(\tau;u)$ invertible) satisfying

$$\begin{aligned} &\frac{d}{d\tau}\Big(\begin{array}{cc} G(\tau;u) & H(\tau;u) \end{array}\Big) \\ &= \Big(\begin{array}{cc} G(\tau;u) & H(\tau;u) \end{array}\Big)\begin{pmatrix} M & -2Q^{\mathrm{T}}Q \\ \frac{1}{2}iu(iue^{-2\kappa\tau}-e^{-\kappa\tau})\mathbf{I}_n & -\Big(M^{\mathrm{T}}+2iue^{-\kappa\tau}RQ\Big) \end{pmatrix}, \end{aligned} \tag{3.7}$$

with initial conditions $H(0;u)=\mathbf{I}_n$ and $G(0;u)=\mathbf{0}_n$.

Proof: Please refer to Section C.1 of Appendix C. □

Since the matrix differential equation (3.7) involves $\exp(-\kappa\tau)$ and $\exp(-2\kappa\tau)$ as the only τ-dependent functions, we can use the classical fourth-order Runge-Kutta method to solve it accurately and efficiently (please refer to Section D.1 of Appendix D).

Now, we extend the dynamic of X_t in (3.1) by including a stochastic jump component as follows

$$dX_t^{(J)} = dX_t - m\lambda dt + d\Big(\sum_{i=1}^{N_t} J_i\Big) \tag{3.8}$$

where $\sum_{i=1}^{N_t} J_i$ is a compound Poisson process, which is assumed to be independent of W_t and Z_t, in which N_t is a Poisson process with constant intensity λ and J_i, $i=1,2,\cdots$ are i.i.d. random variables with probability density function $f(x)$. Denote $m=\int_{\mathbb{R}}(e^x-1)f(dx)$.

Corollary 3.2.1. The characteristic function of $X_T^{(J)}$ in (3.8), conditional on $X_t^{(J)}=x$ and $\Sigma_t=\Sigma$, is given by

$$\Psi^{(J)}(x,\Sigma,\tau;u)=\exp\Big(\mathrm{Tr}\Big(A(\tau;u)\Sigma\Big)+b(\tau;u)x+\widetilde{c}(\tau;u)\Big),$$

where $A(\tau;u)$ and $b(\tau;u)$ are defined in Proposition 3.2.1, and,

$$\begin{aligned}\widetilde{c}(\tau;u) &= iu\int_0^\tau \theta(T-s)e^{-\kappa s}ds - iu\lambda m(1-e^{-\kappa\tau}) \\ &\quad -\frac{\beta}{2}\mathrm{Tr}\Big(\ln II(\tau;u) + M^{\mathrm{T}}\tau + \frac{2iu}{\kappa}(1-e^{-\kappa\tau})RQ\Big) + \Lambda(u)\tau,\end{aligned}$$

where $\Lambda(u) = \lambda\int_{\mathbb{R}}(e^{iux}-1)f(dx)$.

Proof: Please refer to Appendix Section C.2 of Appendix C. □

With Corollary 3.2.1, our model can be considered as a generalization of the model proposed by O'Hara and Pillay [64] in which M, Q and R are *scalars*, the jump size is i.d.d. normally distributed with mean μ and variance γ^2.

With the characteristic function of the log-currency value obtained in Proposition 3.2.1, the probability density function of the log- currency value can be then computed by the inverse Fourier transform of the corresponding characteristic function. Figure 3.2 shows that, when correlation coefficients R_{11} and R_{22} are taken negative (resp. positive), the distribution of the logarithm of 1-year currency value has fatter (resp. thinner) left tail and thinner (resp. fatter) right tail, which corresponds to a downward (resp. upward) sloping implied volatility curve. On the other hand, Figure 3.3 shows that, when the mean-reversion coefficient κ increases, the distribution of the logarithm of 1-year currency value becomes less dispersed, which corresponds to the fact that the currency value reaches it equilibrium value faster and experiences less variation.

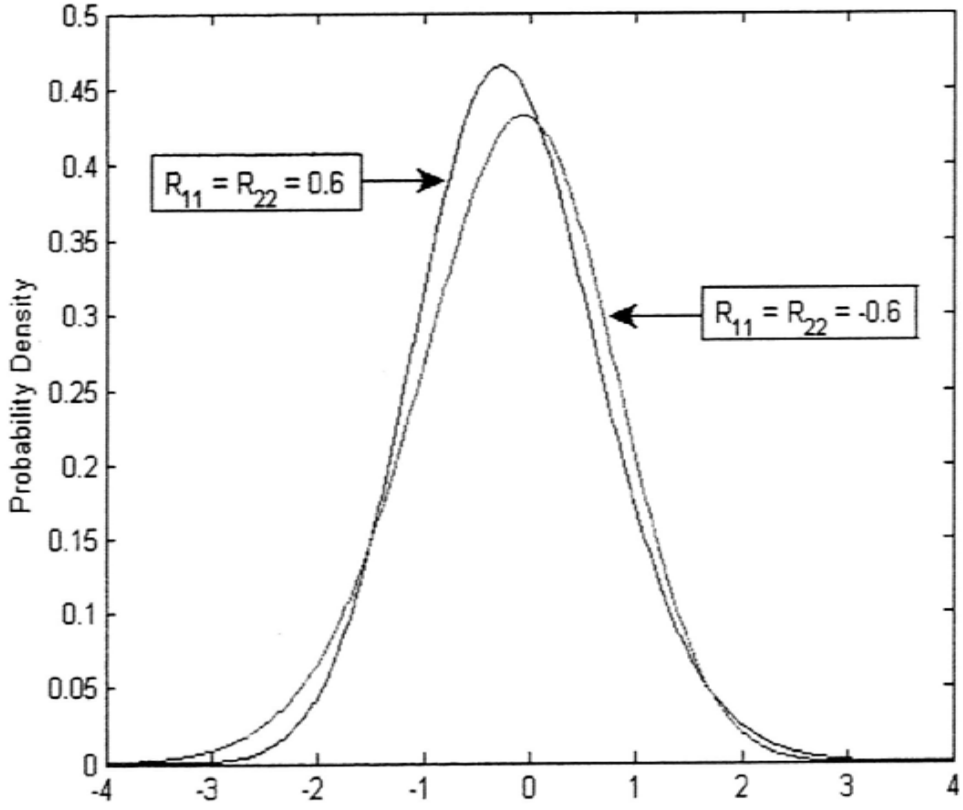

Figure 3.2: The probability density function of the logarithm of 1-year currency value with different correlation coefficients. The parameter values are: $\beta = 5$, $M_{11} = -0.50$, $M_{22} = -0.05$, $Q_{11} = 0.20$, $Q_{22} = 0.10$, $R_{11} = R_{22} = \pm 0.60$, $\Sigma_0^{11} = \Sigma_0^{22} = 0.50$, $\theta = 0.25$, $\kappa = 0.1$.

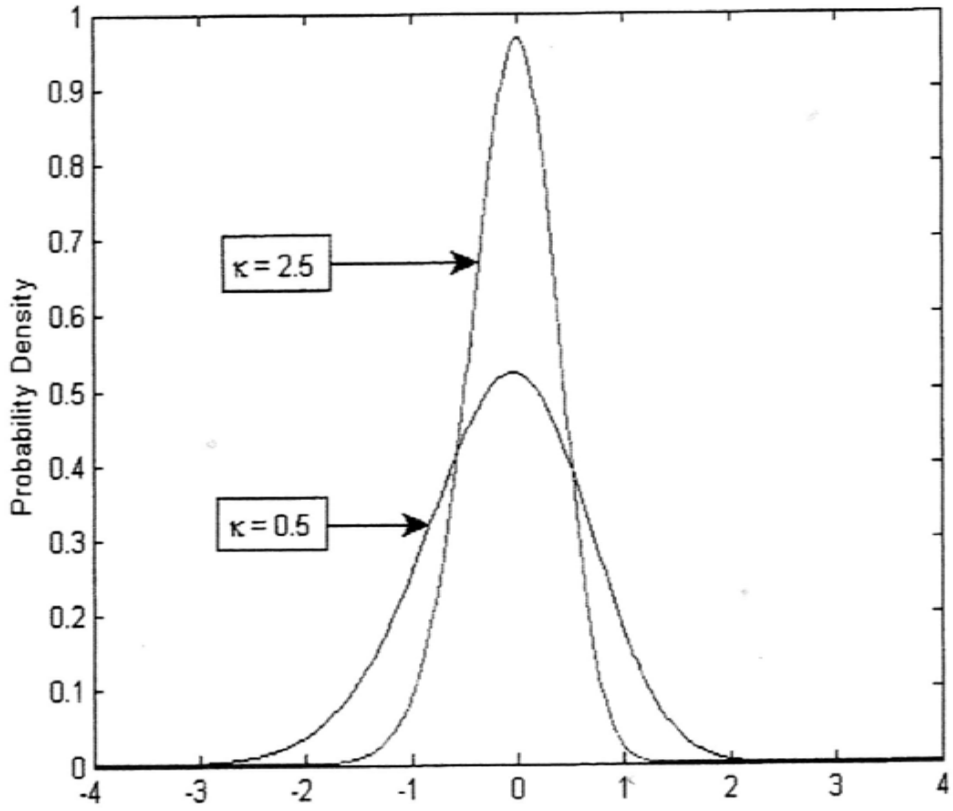

Figure 3.3: The probability density function of the logarithm of 1-year currency value with different mean reversion coefficient. The parameter values are: $\beta = 5$, $M_{11} = -0.50$, $M_{22} = -0.05$, $Q_{11} = 0.20$, $Q_{22} = 0.10$, $R_{11} = R_{22} = -0.40$, $\Sigma_0^{11} = \Sigma_0^{22} = 0.50$, $\theta = 0.25$, $\kappa = 0.5$ or 2.5.

3.2.2 Super-calibration to Currency Futures Prices

Since currency futures are actively traded, it is important to ensure that the currency option prices derived are consistent with the currency futures prices. First of all, under risk-neutral measure, the relationship between spot price and current futures price with maturity T is given by

$$\begin{aligned} F_T(t) &= \mathbb{E}^{\mathbb{Q}}\Big[S_T\Big|X_t = x, \Sigma_t = \Sigma\Big] \\ &= \Psi(x, \Sigma, \tau; -\mathrm{i}) \\ &= \exp\Big(\mathrm{Tr}\Big(A(\tau; -\mathrm{i})\Sigma\Big) + b(\tau; -\mathrm{i})x + c(\tau; -\mathrm{i})\Big). \end{aligned}$$

It is interesting to note that the first term of the function $c(\tau; -\mathrm{i})$ in (3.6), which is an integral with the time-dependent mean reversion level $\theta(t)$, can be absorbed into the term structure of futures prices so that the characteristic function can be re-written without the knowledge of the functional form of $\theta(t)$.

Proposition 3.2.2. If X_t follows the dynamics in (3.1), then the characteristic function for X_T calibrated to the current futures price $F_T(t)$ is given by

$$\Psi(x, \Sigma, \tau; u, F_T(t)) = \exp\Big(\mathrm{i}u \ln F_T(t) + \mathrm{Tr}\Big(\Delta A(\tau; u)\Sigma\Big) + \Delta c(\tau; u)\Big),$$

where

$$\begin{aligned} \Delta A(\tau; u) &= A(\tau; u) - \mathrm{i}uA(\tau; -\mathrm{i}) \\ \Delta c(\tau; u) &= -\frac{1}{2}\beta\mathrm{Tr}\Big(\ln H(\tau; u) - \mathrm{i}u \ln H(\tau; -\mathrm{i}) + M^{\mathrm{T}}\tau(1 - \mathrm{i}u)\Big). \end{aligned}$$

with $A(\tau; u)$ and $H(\tau; u)$ solved in Proposition 3.2.1.

Proof: Please refer to Appendix Section C.3 of Appendix C. □

3.2.3 Pricing Formula of Vanilla Options

A currency call option with strike price K and maturity T has the payoff

$$\max\{S_T - K, 0\}.$$

Therefore, similar to the previous chapter, by the approach of Carr and Madan [22], the price of currency call option with strike price K, maturity T and domestic risk-free rate r, denoted as $C(K,T)$, can be expressed as

$$C(K,T) = \frac{e^{-\alpha \ln K}}{\pi} \int_0^\infty \frac{e^{-rT - \mathrm{i}\xi \ln K} \Psi(x, \Sigma, T; \xi - (\alpha+1)\mathrm{i})}{(\alpha + \mathrm{i}\xi)(\alpha + 1 + \mathrm{i}\xi)} d\xi, \tag{3.9}$$

for some constant $\alpha > 0$. The above integral is approximated by Simpson's rule:

$$C(K,T) \approx \frac{e^{-\alpha \ln K}}{\pi} \sum_{j=0}^{N-1} \frac{e^{-rT - \mathrm{i}\xi_j \ln K} \Psi(x, \Sigma, T; \xi_j - (\alpha+1)\mathrm{i})}{(\alpha + \mathrm{i}\xi_j)(\alpha + 1 + \mathrm{i}\xi_j)} w_j, \tag{3.10}$$

where N is the number of grid points for the interval $[0,T]$, $\eta = \frac{T_j}{N}$, $\xi_j = j\eta$ and

$$w_j = \begin{cases} \frac{1}{3}\eta & \text{for} \quad j = 0, N-1 \\ \frac{1}{3}(3 + (-1)^{j+1})\eta & \text{for} \quad j = 1, \ldots, N-2. \end{cases}$$

The implementation of the fast Fourier transform can be found in Carr and Madan [22] so it is omitted for brevity.

3.3 Simulation Study

In this section, to investigate the accuracy and efficiency of our model in which a numerical ODE solver is used, we compare the option prices computed by the analytical pricing formula (3.9) and by Monte Carlo simulation (which involves a slight modification of Section D.3 of Appendix D). For the numerical integration (3.10), the number of grid

point is $N = 64$, the grid size is $\eta = 0.25$ and the damping coefficient is $\alpha = 3$. For the parameters of the model, $\theta = 0.1$, $\kappa = 0.25$, $S_0 = 1$, $\beta = 5$,

$$M = \begin{pmatrix} -0.50 & 0.00 \\ 0.00 & -0.05 \end{pmatrix}, Q = \begin{pmatrix} 0.10 & 0.05 \\ 0.05 & 0.05, \end{pmatrix},$$

$$R = \begin{pmatrix} -0.04 & -0.02 \\ -0.02 & -0.04 \end{pmatrix}, \Sigma_0 = \begin{pmatrix} 0.50 & 0.20 \\ 0.20 & 0.50 \end{pmatrix}.$$

In Table 3.1, the 0.5-year, 1-year, 1.5-year and 2-year call option prices are computed by analytical formula (3.9) (labeled as "AF") and MC simulation (labeled as "MC")with 100,000 sample paths and time step of 1/100, where CPU times are also reported. As shown in the table, the model prices are very close to simulated prices and the CPU times for analytical formula are much less than those for MC simulation. Thus, the simulation study demonstrates that the analytical formula is correct and efficient.

Table 3.1: The call option prices and CPU times produced by the analytical formula and MC simulation and their percentage pricing error.

	$T = 0.5$			$T = 1$		
Strike Price	AF	MC	% Error	AF	MC	% Error
0.85	0.3360	0.3364	-0.11	0.4025	0.4022	0.07
0.90	0.3119	0.3123	-0.11	0.3809	0.3806	0.07
0.95	0.2896	0.2899	-0.12	0.3606	0.3604	0.07
1.00	0.2689	0.2692	-0.12	0.3416	0.3414	0.06
1.05	0.2497	0.2500	-0.13	0.3238	0.3237	0.05
1.10	0.2320	0.2323	-0.13	0.3071	0.3070	0.04
1.15	0.2156	0.2159	-0.14	0.2915	0.2914	0.03
Total CPU Time (sec)	0.06	409		0.11	817	
	$T = 1.5$			$T = 2$		
Strike Price	AF	MC	% Error	AF	MC	% Error
0.85	0.4310	0.4316	-0.14	0.4410	0.4409	0.03
0.90	0.4107	0.4113	-0.14	0.4215	0.4214	0.02
0.95	0.3915	0.3921	-0.14	0.4031	0.4030	0.02
1.00	0.3735	0.3741	-0.14	0.3857	0.3856	0.01
1.05	0.3566	0.3571	-0.15	0.3693	0.3693	0.01
1.10	0.3406	0.3411	-0.15	0.3539	0.3538	0.01
1.15	0.3255	0.3260	-0.15	0.3393	0.3392	0.00
Total CPU Time (sec)	0.17	1223		0.22	1629	

3.4 Model Properties

In this section, we demonstrate the flexibility of the Wishart model on modeling the stochastic skew over the multi-factor Heston model in the currency option market.

3.4.1 Stochastic Correlation

The instantaneous correlation between the shocks to the log-currency and to the variance dynamics for the Wishart model, which determines the skew of implied volatility curve, is given by taking $U_j(t) \equiv \mathbf{I}_n$ in Expression 2.23. Thus, the instantaneous correlations under the Wishart and multi-factor Heston model are connected by

$$\rho_t^{\text{2D-Wis}} = \rho_t^{\text{2SV-Hes}} + \frac{R_{12}Q_{22}}{\sqrt{\Sigma_t^{11} + \Sigma_t^{22}}\sqrt{Q_{11}^2\Sigma_t^{11} + Q_{22}^2\Sigma_t^{22}}}\Sigma_t^{12}. \tag{3.11}$$

Again, the above relation shows that, when compared with the multi-factor Heston model, the Wishart model contains extra parameter R_{12} and process Σ_t^{12} that offer extra flexibility to model stochastic skew.

3.4.2 Numerical Example

On 31-Dec-2009, the spot price of the 1 EUR is 1.4405 USD and the futures prices are shown below.

Maturity	Futures Price (USD/EUR)
1-month	1.440437
2-month	1.440354
3-month	1.440260
6-month	1.439749
9-month	1.438766
12-month	1.437908
18-month	1.437360
24-month	1.438791

With Corollary 3.2.1, the form of $\theta(t)$ is not important when the term structure of futures prices is available. Consider again the case for $n = 2$. The other model parameters are given by $\kappa = 0.25$, $\beta = 5$,

$$M = \begin{pmatrix} -10.00 & 0.00 \\ 0.00 & -0.05 \end{pmatrix} \quad Q = \begin{pmatrix} 0.20 & 0.00 \\ 0.00 & 0.20 \end{pmatrix}$$

$$R = \begin{pmatrix} +0.10 & R_{12} \\ 0.00 & +0.10 \end{pmatrix} \quad \Sigma_0 = \begin{pmatrix} 0.05 & \Sigma_0^{12} \\ \Sigma_0^{12} & 0.05 \end{pmatrix}.$$

In order to show that the process Σ_t^{12} and parameter R_{12} offer our model extra flexibility to control the volatility skew when compared with the two-factor Heston model, we consider three different cases, where the first case is $(R_{12}, \Sigma_0^{12}) = (-0.40, 0.04)$, the second case is $(R_{12}, \Sigma_0^{12}) = (-0.40, -0.04)$ and the third case is $(R_{12}, \Sigma_0^{12}) = (0, 0)$. Now, we refer the case that $(R_{12}, \Sigma_0^{12}) = (0, 0)$ to the two-factor Heston model.

Figure 3.4 displays the implied volatility curves generated by the three cases, where the lines with *crosses*, *stars* and *dots* represent the curves generated by our model with $(R_{12}, \Sigma_0^{12}) = (-0.40, 0.04)$, $(R_{12}, \Sigma_0^{12}) = (-0.40, -0.04)$ and $(R_{12}, \Sigma_0^{12}) = (0, 0)$ (the two-factor Heston model), respectively. When compared with that of the two-factor Heston

model with $(R_{12}, \Sigma_0^{12}) = (0, 0)$, the implied caplet volatility skews for the Wishart model with $(R_{12}, \Sigma_0^{12}) = (-0.40, 0.04)$ (resp. $(R_{12}, \Sigma_0^{12}) = (-0.40, -0.04)$) are more negative (resp. more positive), and that the effect of twisting is diminishing as time goes by, because of the mean-reversion of Wishart process. The term structure of the implied volatility skew[1] in Figure 3.5 further confirms the conclusions which are made from Figure 3.4. In other words, the Wishart model offers extra flexibility to control the volatility skew without affecting the volatility level. The significance of the above experiment is that suppose the two-factor Heston model have a good fit to the short-term volatility by the process Σ_t^{11} and long-term volatility by the process Σ_t^{22} with suitable degree of correlations to the log-currency process, the Wishart model provides another parameter R_{12} to fit the variation of skew that is not captured by the short-term and long-term volatilities.

[1]The skew of the implied volatility curve for a particular maturity is defined as the difference between the Black's implied volatilities at the right-end point and the left-end point of that volatility curve.

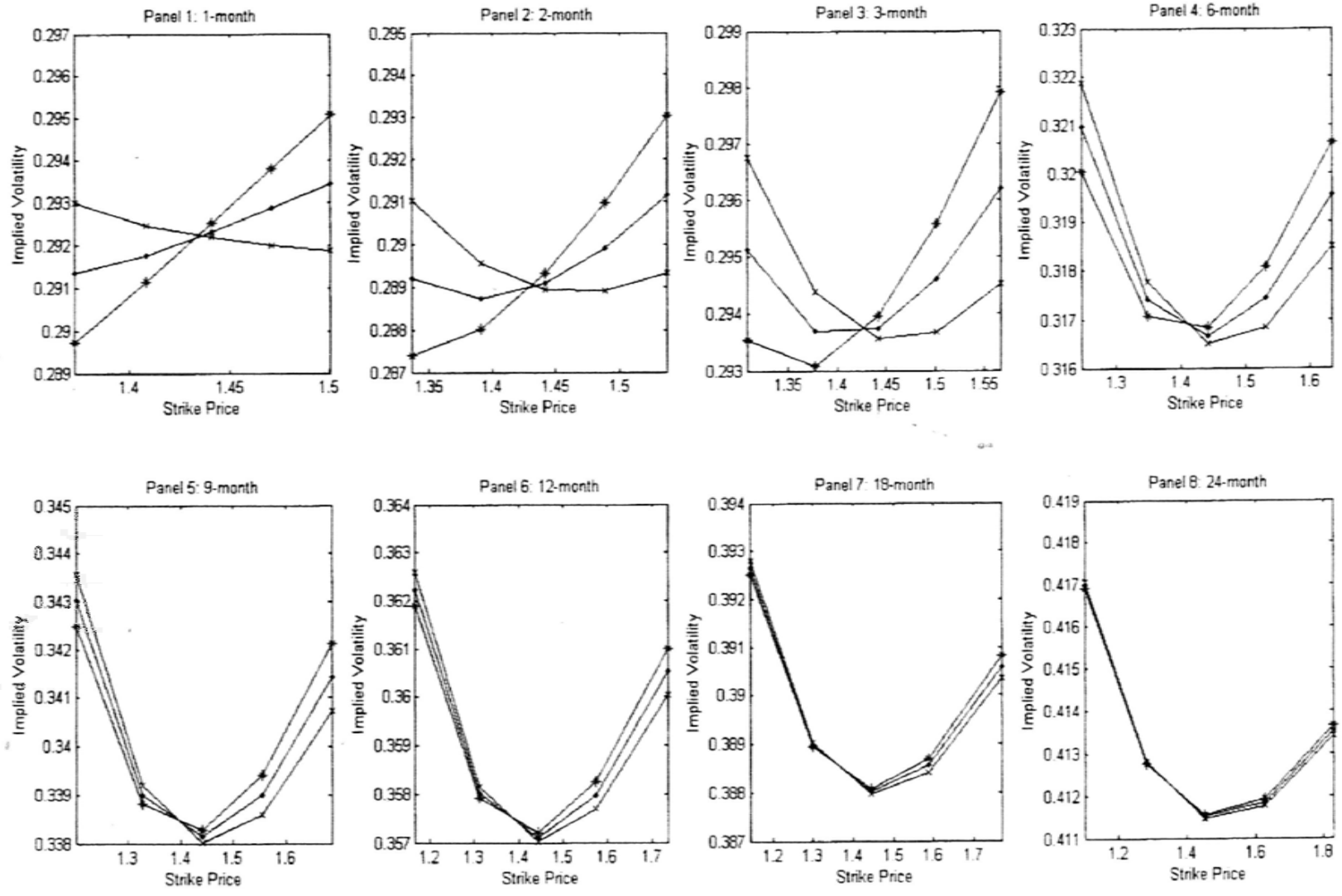

Figure 3.4: The implied volatility curves generated by the Wishart model (lines with *crosses* represent the case of $(R_{12}, \Sigma_0^{12}) = (-0.40, 0.04)$, lines with *stars* represent the case of $(R_{12}, \Sigma_0^{12}) = (-0.40, -0.04)$) and two-factor Heston model (lines with *dots* represent the case of $(R_{12}, \Sigma_0^{12}) = (0, 0)$).

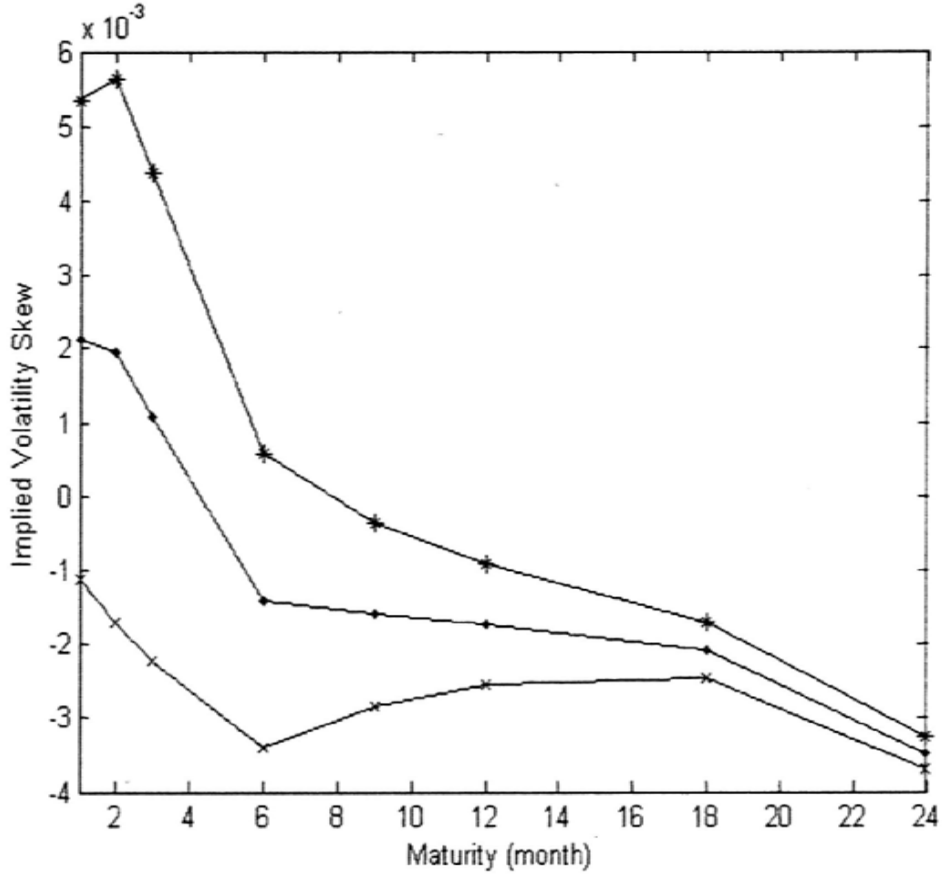

Figure 3.5: The term structure of implied volatility skew generated by the Wishart model (lines with *crosses* represent the case of $(R_{12}, \Sigma_0^{12}) = (-0.40, 0.04)$, lines with *stars* represent the case of $(R_{12}, \Sigma_0^{12}) = (-0.40, -0.04)$) and two-factor Heston model (lines with *dots* represent the case of $(R_{12}, \Sigma_0^{12}) = (0, 0)$).

3.5 Summary

We propose a currency option pricing model to simultaneously capture the three essential features observed in the currency option markets: mean-reversion, stochastic volatility and stochastic skew. Using the non-diagonal elements of matrices in the Wishart process, our model offers extra controls on the skew as compared with the multi-factor Heston model. Analytical solutions are derived for the characteristic functions and vanilla European options, which enable our framework to be implemented accurately and efficiently for practical use, as shown by MC simulation. Through a numerical example with term structure of futures prices, the flexibility of modeling stochastic skew of the Wishart model is demonstrated to be superior to the multi-factor Heston model.

Chapter 4

Conclusion and Discussion

This thesis investigates the application of the Wishart process in interest rate cap pricing and currency option pricing problems in order to capture stochastic skew present in both markets.

First, to price interest rate caps consistent with stochastic skew documented in a set of cap volatility data, we incorporate the Wishart process into the LIBOR market model and derive accurate and efficient pricing formula for caplets based on freezing approximation and transform techniques so the LIBOR market model with Wishart process can be implemented easily in practice. Most importantly, we show that, when compared with the multi-factor Heston model, the Wishart model has an additional degree of freedom to model stochastic skew after the volatility factors are fixed. Finally, calibration and estimation results show that the better pricing performance of Wishart model over the two-factor Heston model can be attributed to the better modeling of stochastic skew by the Wishart model.

Second, we incorporate the Wishart process in a currency option pricing model to price currency options consistent with stochastic skew documented in a sample of risk reversals, together with mean-reversion and stochastic volatility. Closed-form solutions for characteristic function and vanilla option price are derived. In addition, our model can be extended to include Poisson shocks, thereby giving it freedom to incorporate jumps in the currency dynamics. Using a term structure of currency futures prices, the flexibility of modeling stochastic skew of the Wishart model is demonstrated to be superior to the multi-factor Heston model.

As a final remark, given a fixed set of model parameters, that the non-diagonal elements of the Wishart process affect only the skew and do not affect the level of volatility curve makes the Wishart process able to separately fit the skew and level; in contrast, in the multi-factor Heston model, any change in level of volatility curve eventually affects the skew, and vice versa. This is the main point we emphasize in this thesis. One may argue that, for the calibration to one day's market data, the multi-factor Heston model has already performed very well and the Wishart model only performs slightly better; however, it does not mean that the Wishart model does not have any advantage but simply means that the Wishart model performs as good as multi-factor Heston model in the calibration of one day's vanilla option data. To highlight the strength of the Wishart model over the multi-factor Heston model, we investigate the model performance by in-sample and out-of-sample estimation using time series of interest rate cap data and results show that the flexibility of modeling stochastic skew indeed make the Wishart model perform better than the multi-factor Heston model.

The advantage of the Wishart model can be made clear in pricing exotic options (see Da Fonseca and Grasselli [28] and Da Fonseca *et al.* [29, 30] for a related discussion) consistently with a model capable of modeling stochastic volatility *as well as* stochastic skew. In other words, even though sometimes the Wishart model performs similarly with the multi-factor Heston model in fitting vanilla option data of one day or many days, the ability of Wishart model to capture some "stylized facts" in the market, which are determinants of exotic option prices but cannot be reproduced by the multi-factor Heston model, grants the Wishart model the strength in pricing exotic options. Therefore, we suggest that future research can compare the pricing performance of some exotic options by the Wishart model and the multi-factor Heston model.

Appendix A

The Wishart Distribution and The Wishart Process

The name "Wishart" is used in honor of John Wishart [75], who, in 1928, first formulated the Wishart distribution. In statistics, the Wishart distribution, which is defined over symmetric, nonnegative-definite matrix-valued random variables, is a generalization to multi-dimension of the chi-square distribution, or, in the case of non-integer degrees of freedom, of the gamma distribution. The Wishart distribution is the sampling distribution of the maximum-likelihood estimator (MLE) of the covariance matrix of a multivariate normal distribution with zero means and, in Bayesian inference, is the conjugate prior of the inverse of the covariance matrix (the precision matrix) of a multivariate normal distribution. Indeed, the Wishart process (2.1) considered in this thesis is closely related to the (non-central) Wishart distribution so that the Wishart process can be considered as a model for stochastic covariance dynamics.

Suppose that $\{X_1, \ldots, X_k\}$ is a set of independent random vectors in $\mathbb{R}^n$, where $n \leq k$, with multivariate normal distribution $\mathcal{N}(0, \Sigma)$. Then, the distribution of the random variable defined by

$$W = \sum_{i=1}^{k} X_i X_i^{\mathrm{T}}$$

is a Wishart distribution with k degrees of freedom, denoted by $\mathcal{W}_k(0, \Sigma)$. Moreover, the distribution of the random variable defined by

$$W = \sum_{i=1}^{k} (X_i + \mu_i)(X_i + \mu_i)^{\mathrm{T}},$$

where $\mu_i \in \mathbb{R}^n$, is a non-central Wishart distribution with k degrees of freedom, denoted by $\mathcal{W}_k(\mu, \Sigma)$ with $\mu = \sum_{i=1}^{k} \mu_i \mu_i^{\mathrm{T}}$, whose probability density function and Laplace transform (or characteristic function) can be found in Anderson [5]. The above result can be extended to any real $k > n - 1$.

Now, we state the relationship between the Wishart process and the Wishart distribution. If the random matrix process $\{\Sigma_t\}_{t \geq 0}$ has the dynamics in (2.1), then, conditional on the random matrix Σ_t, for $h \geq 0$, Σ_{t+h} has a non-central Wishart distribution $\mathcal{W}_\beta(F(h)\Sigma_t F(h)^{\mathrm{T}}, G(h))$, where

$$\begin{aligned} F(h) &= \exp\Big(hM\Big), \\ G(h) &= \int_0^h \exp\Big(sM\Big) Q^{\mathrm{T}} Q \exp\Big(sM^{\mathrm{T}}\Big) ds. \end{aligned}$$

Appendix B

Mathematical Derivations for Chapter 2

B.1 Proof of Proposition 2.2.1

Given the money market account

$$B(t) = \exp\Big(\int_0^t r(s)ds\Big),$$

the Radon-Nikodym derivative of $\mathbb{Q}^{T_{j+1}}$ with respect to $\mathbb{Q}$ is, for $t \in [0, T_j]$,

$$\begin{aligned} m_{T_{j+1}}(t) &\triangleq \frac{d\mathbb{Q}^{T_{j+1}}}{d\mathbb{Q}}\Big|_{\mathcal{F}_t} \\ &= \frac{P_{j+1}(t)/P_{j+1}(0)}{B(t)/B(0)} \\ &= \exp\Big(\int_0^t \mathrm{Tr}\Big(V_{j+1}(s)\sqrt{\Sigma_s}dZ_s\Big) - \frac{1}{2}\int_0^t \mathrm{Tr}\Big(V_{j+1}(s)^2\Sigma_s\Big)ds\Big) \end{aligned}$$

so that

$$\frac{dm_{T_{j+1}}(t)}{m_{T_{j+1}}(t)} = \mathrm{Tr}\Big(V_{j+1}(t)\sqrt{\Sigma_t}dZ_t\Big).$$

In the derivation below, we apply the two identities

$$\begin{aligned} \mathrm{Tr}\Big(AB\Big) &= \mathrm{Tr}\Big(BA\Big) \\ \mathrm{Tr}\Big(A^{\mathrm{T}}B\Big) &= \sum_{i,j=1}^{n} A_{ij}B_{ij} \end{aligned}$$

and the fact that $dW_t^{ij}dW_t^{kl} = dt$ if and only if $i = k$ and $j = l$, and equals zero otherwise.

Let $\Big\langle \cdot, \cdot \Big\rangle$ denote covariance. By Girsanov's Theorem,

$$\begin{aligned} dZ_t^{T_{j+1}} &= dZ_t - \Big\langle \frac{dm_{T_{j+1}}(t)}{m_{T_{j+1}}(t)}, dZ_t \Big\rangle \\ &= dZ_t - \Big\langle \mathrm{Tr}\Big(V_{j+1}(t)\sqrt{\Sigma_t}dZ_t\Big), dZ_t \Big\rangle \\ &= dZ_t - \Big\langle \mathrm{Tr}\Big(\Big(\sqrt{\Sigma_t}V_{j+1}(t)\Big)^{\mathrm{T}}dZ_t\Big), dZ_t \Big\rangle \\ &= dZ_t - \Big[\Big\langle \sum_{k,l=1}^{n} [\sqrt{\Sigma_t}V_{j+1}(t)]_{kl}dZ_t^{kl}, dZ_t^{kl} \Big\rangle\Big]_{k,l=1,\ldots,n} \\ &= dZ_t - \Big[[\sqrt{\Sigma_t}V_{j+1}(t)]_{kl}dt\Big]_{k,l=1,\ldots,n} \\ &= dZ_t - \sqrt{\Sigma_t}V_{j+1}(t)dt, \end{aligned}$$

Similarly,

$$
\begin{aligned}
dW_t^{T_{j+1}} &= dW_t - \Big\langle \frac{dm_{T_{j+1}}(t)}{m_{T_{j+1}}(t)}, dW_t \Big\rangle \\
&= dW_t - \Big\langle \mathrm{Tr}\Big(V_{j+1}(t)\sqrt{\Sigma_t}dZ_t\Big), dW_t \Big\rangle \\
&= dW_t - \Big\langle \mathrm{Tr}\Big(V_{j+1}(t)\sqrt{\Sigma_t}\Big(dW_t R^{\mathrm{T}} + dB_t\sqrt{\mathbf{I}_n - RR^{\mathrm{T}}}\Big)\Big), dW_t \Big\rangle \\
&= dW_t - \Big\langle \mathrm{Tr}\Big(V_{j+1}(t)\sqrt{\Sigma_t}dW_t R^{\mathrm{T}}\Big), dW_t \Big\rangle \quad \text{(since } B \text{ and } W \text{ are independent)} \\
&= dW_t - \Big\langle \mathrm{Tr}\Big(\Big(\sqrt{\Sigma_t}V_{j+1}(t)R\Big)^{\mathrm{T}} dW_t\Big), dW_t \Big\rangle \\
&= dW_t - \Big[\Big\langle \sum_{k,l=1}^{n} [\sqrt{\Sigma_t}V_{j+1}(t)R]_{kl} dW_t^{kl}, dW_t^{kl} \Big\rangle\Big]_{k,l=1,\ldots,n} \\
&= dW_t - \Big[[\sqrt{\Sigma_t}V_{j+1}(t)R]_{kl}dt\Big]_{k,l=1,\ldots,n} \\
&= dW_t - \sqrt{\Sigma_t}V_{j+1}(t)Rdt.
\end{aligned}
$$

The dynamics of forward rate under $\mathbb{Q}^{T_{j+1}}$ is obtained by putting (2.6) into (2.4) and the dynamics of the Wishart process under $\mathbb{Q}^{T_{j+1}}$ is obtained by putting putting (2.7) into (2.1).

B.2 Proof of Proposition 2.2.2

Since Wishart process is a matrix affine process, the characteristic function of X_{T_j} is exponentially affine in the state variables (see Duffie and Kan [32] and Grasselli and Tebaldi [42]),

$$
\begin{aligned}
\Psi(x, \Sigma, \tau; u) &= \exp\Big(\mathrm{Tr}\Big(A(\tau; u)\Sigma\Big) + b(\tau; u)x + c(\tau; u)\Big) \\
\Psi(x, \Sigma, 0; u) &= \exp\Big(\mathrm{i}ux\Big),
\end{aligned}
$$

so that $A(0;u) = \mathbf{0}_n$, $b(0;u) = iu$ and $c(0;u) = 0$. Our strategy is to apply the Feynman-Kac argument to obtain a set of ordinary differential equations for each of the functions.

The joint infinitesimal generator of (X, Σ) can be expressed in the form of

$$\begin{aligned}\mathcal{L}_{X,\Sigma} \triangleq\ & -\frac{1}{2}\mathrm{Tr}\Big(U_j(t)^2\Sigma\Big)\frac{\partial}{\partial x} + \frac{1}{2}\mathrm{Tr}\Big(U_j(t)^2\Sigma\Big)\frac{\partial^2}{\partial x^2} \\ & + \mathrm{Tr}\Big(\Big(\beta Q^{\mathrm{T}}Q + M^{T_{j+1}}(t)\Sigma + \Sigma M^{T_{j+1}}(t)^{\mathrm{T}}\Big)D + 2\Sigma D Q^{\mathrm{T}}Q D\Big) \\ & + \sum_{i,j=1}^{n} \varrho_{ij}\frac{\partial^2}{\partial\Sigma^{ij}\partial x} \end{aligned} \tag{B.1}$$

where $D = [\frac{\partial}{\partial\Sigma^{ij}}]_{i,j=1,\ldots,n}$ is a matrix differential operator. The first line and second line are the infinitesimal generators of the log-forward LIBOR process and the Wishart process (see Bru [18]) under $\mathbb{Q}^{T_{j+1}}$, respectively. ϱ_{ij} in the last line of Equation (B.1) is derived as follow. In fact, $\varrho_{ij}dt$ is the (i,j)-th element of

$$\begin{aligned}
& \mathbb{E}^{\mathbb{Q}^{T_{j+1}}}\Big[\mathrm{Tr}\Big(U_j(t)\sqrt{\Sigma_t}dW_t^{T_{j+1}}R^{\mathrm{T}}\Big)\Big(\sqrt{\Sigma_t}dW_t^{T_{j+1}}Q + Q^{\mathrm{T}}(dW_t^{T_{j+1}})^{\mathrm{T}}\sqrt{\Sigma_t}\Big)\Big|X_t,\Sigma_t\Big] \\
= & \ \mathbb{E}^{\mathbb{Q}^{T_{j+1}}}\Big[\Big(\sum_{k,l=1}^{n}[R^{\mathrm{T}}U_j(t)\sqrt{\Sigma_t}]_{lk}dW_t^{T_{j+1},kl}\Big)\sqrt{\Sigma_t}[dW_t^{T_{j+1},kl}]_{k,l=1,\ldots,n}Q\Big|X_t,\Sigma_t\Big] \\
& + \mathbb{E}^{\mathbb{Q}^{T_{j+1}}}\Big[\Big(\sum_{k,l=1}^{n}[R^{\mathrm{T}}U_j(t)\sqrt{\Sigma_t}]_{lk}dW_t^{T_{j+1},kl}\Big)Q^{\mathrm{T}}[dW_t^{T_{j+1},lk}]_{k,l=1,\ldots,n}\sqrt{\Sigma_t}\Big|X_t,\Sigma_t\Big] \\
= & \ \mathbb{E}^{\mathbb{Q}^{T_{j+1}}}\Big[\Big(\sum_{k,l=1}^{n}[\sqrt{\Sigma_t}U_j(t)R]_{kl}dW_t^{T_{j+1},kl}\Big)\sqrt{\Sigma_t}[dW_t^{T_{j+1},kl}]_{k,l=1,\ldots,n}Q\Big|X_t,\Sigma_t\Big] \\
& + \mathbb{E}^{\mathbb{Q}^{T_{j+1}}}\Big[\Big(\sum_{k,l=1}^{n}[R^{\mathrm{T}}U_j(t)\sqrt{\Sigma_t}]_{kl}dW_t^{T_{j+1},lk}\Big)Q^{\mathrm{T}}[dW_t^{T_{j+1},lk}]_{k,l=1,\ldots,n}\sqrt{\Sigma_t}\Big|X_t,\Sigma_t\Big] \\
= & \ \sqrt{\Sigma_t}\Big[[\sqrt{\Sigma_t}U_j(t)R]_{kl}\Big]_{k,l=1,\ldots,n}Qdt + Q^{\mathrm{T}}\Big[[R^{\mathrm{T}}U_j(t)\sqrt{\Sigma_t}]_{kl}\Big]_{k,l=1,\ldots,n}\sqrt{\Sigma_t}dt \\
= & \ \Sigma_t U_j(t)RQdt + Q^{\mathrm{T}}R^{\mathrm{T}}U_j(t)\Sigma_t dt.
\end{aligned}$$

Taking $\Sigma_t = \Sigma$ yields

$$\begin{aligned}
\sum_{k,l=1}^{n} \varrho_{ij} \frac{\partial^2}{\partial \Sigma^{kl} \partial x} &= \sum_{k,l=1}^{n} \Big([\Sigma U_j(t) RQ]_{k,l} \frac{\partial}{\partial \Sigma^{kl}} + [Q^{\mathrm{T}} R^{\mathrm{T}} U_j(t) \Sigma]_{k,l} \frac{\partial}{\partial \Sigma^{kl}} \Big) \frac{\partial}{\partial x} \\
&= \sum_{k,l=1}^{n} \Big([\Sigma U_j(t) RQ]_{k,l} \frac{\partial}{\partial \Sigma^{lk}} + [\Sigma U_j(t) RQ]_{l,k} \frac{\partial}{\partial \Sigma^{kl}} \Big) \frac{\partial}{\partial x} \\
&= 2\mathrm{Tr}\Big(\Sigma U_j(t) RQD \Big) \frac{\partial}{\partial x}.
\end{aligned}$$

Now, applying the Feynman-Kac argument on the characteristic function (2.13) by using the joint infinitesimal generator (B.1) gives the following partial differential equation for the characteristic function

$$\begin{aligned}
\frac{\partial \Psi}{\partial \tau} &= -\frac{1}{2}\mathrm{Tr}\Big(U_j(t)^2 \Sigma \Big) \frac{\partial}{\partial x} \Psi + \frac{1}{2}\mathrm{Tr}\Big(U_j(t)^2 \Sigma \Big) \frac{\partial^2}{\partial x^2} \Psi \\
&\quad + \mathrm{Tr}\Big(\Big(\beta Q^{\mathrm{T}} Q + M^{T_{j+1}}(t) \Sigma + \Sigma M^{T_{j+1}}(t)^{\mathrm{T}} \Big) D + 2\Sigma D Q^{\mathrm{T}} Q D \Big) \Psi \\
&\quad + 2\mathrm{Tr}\Big(\Sigma U_j(t) RQD \Big) \frac{\partial}{\partial x} \Psi \\
\Psi(x, \Sigma, 0; u) &= \exp\Big(iux \Big),
\end{aligned}$$

By rearranging the terms,

$$\begin{aligned}
0 &= -\mathrm{Tr}\Big(\frac{d}{d\tau} A(\tau; u) \Sigma \Big) - \frac{d}{d\tau} b(\tau; u) x - \frac{d}{d\tau} c(\tau; u) \\
&\quad + \mathrm{Tr}\Big(\Big(\frac{1}{2} b(\tau; u)(b(\tau; u) - 1) U_j(t)^2 + A(\tau; u) M^{T_{j+1}}(t) \\
&\quad + \Big(M^{T_{j+1}}(t)^{\mathrm{T}} + 2b(\tau; u) U_j(t) RQ \Big) A(\tau; u) + 2A(\tau; u) Q^{\mathrm{T}} Q A(\tau; u) \Big) \Sigma \Big) \\
&\quad + \beta \mathrm{Tr}\Big(Q^{\mathrm{T}} Q A(\tau; u) \Big),
\end{aligned}$$

and identifying the coefficients of Σ and x and the constant terms respectively on both

sides, with $t = T_j - \tau$, we obtain the following system of ODEs:

$$\begin{aligned}
\frac{d}{d\tau}A(\tau;u) &= \frac{1}{2}b(\tau;u)(b(\tau;u)-1)U_j(T_j-\tau)^2 + A(\tau;u)M^{T_{j+1}}(T_j-\tau) \\
&\quad + \Big(M^{T_{j+1}}(T_j-\tau)^{\mathrm{T}} + 2b(\tau;u)U_j(T_j-\tau)RQ\Big)A(\tau;u) \\
&\quad + 2A(\tau;u)Q^{\mathrm{T}}QA(\tau;u) \qquad \text{(B.2)} \\
\frac{d}{d\tau}b(\tau;u) &= 0 \qquad \text{(B.3)} \\
\frac{d}{d\tau}c(\tau;u) &= \beta\mathrm{Tr}\Big(Q^{\mathrm{T}}QA(\tau;u)\Big), \qquad \text{(B.4)}
\end{aligned}$$

with the initial conditions

$$\begin{aligned}
A(0;u) &= \mathbf{0}_n \qquad \text{(B.5)} \\
b(0;u) &= \mathrm{i}u \qquad \text{(B.6)} \\
c(0;u) &= 0. \qquad \text{(B.7)}
\end{aligned}$$

The solution to Equation (B.3) with initial condition (B.6) is

$$b(\tau;u) = \mathrm{i}u.$$

Consider Equation (B.2). By Radon's lemma (see Freiling [35]), let

$$G(\tau;u) = H(\tau;u)A(\tau;u),$$

with $H(\tau;u)$ invertible, where $G(0;u) = \mathbf{0}_n$ and $H(0;u) = \mathbf{I}_n$. With Equation (B.2), differentiating both sides with respect to τ yields

$$\begin{aligned}
\frac{d}{d\tau}G(\tau;u) &= \Big(\frac{d}{d\tau}H(\tau;u)\Big)A(\tau;u) + H(\tau;u)\frac{d}{d\tau}A(\tau;u) \\
&= \Big(\frac{d}{d\tau}H(\tau;u)\Big)A(\tau;u) + H(\tau;u)\Big[\frac{1}{2}\mathrm{i}u(\mathrm{i}u-1)U_j(T_j-\tau)^2 + A(\tau;u)M^{T_{j+1}}(T_j-\tau) \\
&\quad + \Big(M^{T_{j+1}}(T_j-\tau)^{\mathrm{T}} + 2\mathrm{i}uU_j(T_j-\tau)RQ\Big)A(\tau;u) + 2A(\tau;u)Q^{\mathrm{T}}QA(\tau;u)\Big] \\
&= G(\tau;u)M^{T_{j+1}}(T_j-\tau) + \frac{1}{2}\mathrm{i}u(\mathrm{i}u-1)H(\tau;u)U_j(T_j-\tau)^2 \\
&\quad + \Big(\frac{d}{d\tau}H(\tau;u) + 2G(\tau;u)Q^{\mathrm{T}}Q + H(\tau;u)\Big(M^{T_{j+1}}(T_j-\tau)^{\mathrm{T}} + 2\mathrm{i}uU_j(T_j-\tau)RQ\Big)\Big)A(\tau;u),
\end{aligned}$$

so that

$$
\begin{aligned}
0 \;=\; & \Big(-\frac{d}{d\tau}G(\tau;u) + G(\tau;u)M^{T_{j+1}}(T_j-\tau) + \frac{1}{2}\mathrm{i}u(\mathrm{i}u-1)H(\tau;u)U_j(T_j-\tau)^2\Big) \\
& + \Big(\frac{d}{d\tau}H(\tau;u) + 2G(\tau;u)Q^{\mathrm{T}}Q + H(\tau;u)\Big(M^{T_{j+1}}(T_j-\tau)^{\mathrm{T}} + 2\mathrm{i}uU_j(T_j-\tau)RQ\Big)\Big)A(\tau;u),
\end{aligned}
$$

and, then, identifying the coefficients of $A(\tau;u)$ and the constant terms respectively on both sides yield the following system of ODEs

$$
\begin{cases}
\frac{d}{d\tau}G(\tau;u) &= G(\tau;u)M^{T_{j+1}}(T_j-\tau) + \frac{1}{2}\mathrm{i}u(\mathrm{i}u-1)H(\tau;u)U_j(T_j-\tau)^2 \\
\frac{d}{d\tau}H(\tau;u) &= -2G(\tau;u)Q^{\mathrm{T}}Q - H(\tau;u)\Big(M^{T_{j+1}}(T_j-\tau)^{\mathrm{T}} + 2\mathrm{i}uU_j(T_j-\tau)RQ\Big),
\end{cases}
$$

The above system of ODEs can be re-written as follow:

$$
\begin{aligned}
& \frac{d}{d\tau}\begin{pmatrix} G(\tau;u) & H(\tau;u) \end{pmatrix} \\
= \; & \begin{pmatrix} G(\tau;u) & H(\tau;u) \end{pmatrix}\begin{pmatrix} M^{T_{j+1}}(T_j-\tau) & -2Q^{\mathrm{T}}Q \\ \frac{1}{2}\mathrm{i}u(\mathrm{i}u-1)U_j(T_j-\tau)^2 & -\Big(M^{T_{j+1}}(T_j-\tau)^{\mathrm{T}} + 2\mathrm{i}uU_j(T_j-\tau)RQ\Big) \end{pmatrix},
\end{aligned}
$$

where $G(0;u) = \mathbf{0}_n$ and $H(0;u) = \mathbf{I}_n$. Therefore, $A(\tau;u)$ is solved. Finally, consider Equation (B.4)

$$
\begin{aligned}
\frac{d}{d\tau}c(\tau;u) &= \beta\mathrm{Tr}\Big(Q^{\mathrm{T}}QA(\tau;u)\Big) \\
&= \beta\mathrm{Tr}\Big(Q^{\mathrm{T}}QH(\tau;u)^{-1}G(\tau;u)\Big) \\
&= -\frac{\beta}{2}\mathrm{Tr}\Big(Q^{\mathrm{T}}QH(\tau;u)^{-1}\Big(\frac{d}{d\tau}H(\tau;u) + H(\tau;u)\Big(M^{T_{j+1}}(T_j-\tau)^{\mathrm{T}} + 2\mathrm{i}uU_j(T_j-\tau)RQ\Big)\Big) \\
&\quad \times(Q^{\mathrm{T}}Q)^{-1}\Big) \\
&= -\frac{\beta}{2}\mathrm{Tr}\Big(H(\tau;u)^{-1}\frac{d}{d\tau}H(\tau;u) + M^{T_{j+1}}(T_j-\tau)^{\mathrm{T}} + 2\mathrm{i}uU_j(T_j-\tau)RQ\Big).
\end{aligned}
$$

The solution of $c(\tau;u)$ can be obtained by directly integrating from 0 to τ using initial condition (B.7).

B.3 Proof of Corollary 2.2.1

For $\tau \in (\tau_i, \tau_{i+1}]$, $U_k(T_j - \tau) = U_k^{(i)}$ and $M^{T_{j+1}}(T_j - \tau) = M_i^{T_{j+1}}$, so the matrix differential equation (2.16) now has *constant coefficients* and, therefore, has closed-form solution expressed as matrix exponential in (2.19). The solution for A in (2.18) is obtained in the final step of recursion and the solution for c in (2.21) is obtained by integration on each interval $(\tau_i, \tau_{i+1}]$.

Appendix C

Mathematical Derivations for Chapter 3

C.1 Proof of Proposition 3.2.1

The general procedure is largely similar to that in Section B.2 of Appendix B; however, the system of ODEs for the characteristic function now becomes:

$$\begin{aligned}
\frac{d}{d\tau}A(\tau;u) &= \frac{1}{2}b(\tau;u)(b(\tau;u)-1)\mathbf{I}_n + A(\tau;u)M + \Big(M^{\mathrm{T}} + 2b(\tau;u)RQ\Big)A(\tau;u) \\
&\quad +2A(\tau;u)Q^{\mathrm{T}}QA(\tau;u) \\
\frac{d}{d\tau}b(\tau;u) &= -\kappa b(\tau;u) \\
\frac{d}{d\tau}c(\tau;u) &= \theta(T-\tau)b(\tau;u) + \beta\mathrm{Tr}\Big(Q^{\mathrm{T}}QA(\tau;u)\Big),
\end{aligned}$$

with initial conditions

$$\begin{aligned} A(0;u) &= \mathbf{0}_n \\ b(0;u) &= iu \\ c(0;u) &= 0. \end{aligned}$$

The solution of $A(\tau;u)$, $b(\tau;u)$ and $c(\tau;u)$ can be solved using similar techniques in Section B.2 of Appendix B.

C.2 Proof of Proposition 3.2.1

Rewrite $dX_t^{(J)}$ as

$$dX_t^{(J)} = d\widetilde{X}_t + d\Big(\sum_{i=1}^{N_t} J_i\Big)$$

where $d\widetilde{X}_t = dX_t - m\lambda dt$. Now,

$$\begin{aligned} \Psi^{(J)}(x,\Sigma,\tau;u) &= \mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(iuX_T^{(J)}\Big)\Big|X_t^{(J)}=x,\Sigma_t=\Sigma\Big] \\ &= \mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(iu\Big(\widetilde{X}_T+\sum_{i=1}^{N_\tau}J_i\Big)\Big)\Big|X_t^{(J)}=x,\Sigma_t=\Sigma\Big] \\ &= \mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(iu\widetilde{X}_T\Big)\Big|X_t^{(J)}=x,\Sigma_t=\Sigma\Big]\mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(iu\Big(\sum_{i=1}^{N_\tau}J_i\Big)\Big)\Big] \end{aligned}$$

The last line is obtained by using the independence of the compound Poisson process and the process $\widetilde{X}_T$. The first expectation in the last row can be derived from Proposition 3.2.1 by simply replacing $\theta(t)$ with $\theta(t)-m\lambda$ and the second one is just the characteristic function of the compound Poisson process, which is given by

$$\mathbb{E}^{\mathbb{Q}}\Big[\exp\Big(iu\Big(\sum_{i=1}^{N_\tau}J_i\Big)\Big]=\exp\Big(\lambda\tau\int_{\mathbb{R}}(e^{iux}-1)f(dx)\Big).$$

C.3 Proof of Proposition 3.2.2

The current futures price with maturity T is given by

$$\begin{aligned} F_T(t) &= \mathbb{E}\Big[S_T\Big|X_t = x, \Sigma_t = \Sigma\Big] \\ &= \Psi(x, \Sigma, \tau; -\mathrm{i}) \\ &= \exp\Big(\mathrm{Tr}\Big(A(\tau; -\mathrm{i})\Sigma\Big) + b(\tau; -\mathrm{i})x + c(\tau; -\mathrm{i})\Big). \end{aligned}$$

where $A(\tau; -\mathrm{i})$ is solved in Proposition 3.2.1 and

$$\begin{aligned} b(\tau; -\mathrm{i}) &= e^{-\kappa\tau} \\ c(\tau; -\mathrm{i}) &= \int_0^\tau \theta(T-s)e^{-\kappa s}ds - \frac{1}{2}\beta\mathrm{Tr}\Big(\ln H(\tau; -\mathrm{i}) + M^{\mathrm{T}}\tau + \frac{2}{\kappa}(1-e^{-\kappa\tau})RQ\Big). \end{aligned}$$

Rearrangement of terms yields

$$\begin{aligned} \int_0^\tau \theta(\tau-s)e^{-\kappa s}ds &= \ln F_T(t) - \mathrm{Tr}\Big(A(\tau; -\mathrm{i})\Sigma\Big) - xe^{-\kappa\tau} + \frac{1}{2}\beta\mathrm{Tr}\Big(\ln H(\tau; -\mathrm{i}) \\ &\quad + M^{\mathrm{T}}\tau + \frac{2}{\kappa}(1-e^{-\kappa\tau})RQ\Big). \end{aligned}$$

Now, substituting the above expression into the characteristic function (3.3) yields

$$\begin{aligned} \Psi(x, \Sigma, \tau; u, F_T(t)) &= \exp\Big(\mathrm{Tr}\Big(A(\tau; u)\Sigma\Big) + \mathrm{i}ux e^{-\kappa\tau} \\ &\quad + \mathrm{i}u\Big[\ln F_T(t) - \mathrm{Tr}\Big(A(\tau; -\mathrm{i})\Sigma\Big) - xe^{-\kappa\tau} \\ &\quad + \frac{1}{2}\beta\mathrm{Tr}\Big(\ln H(\tau; -\mathrm{i}) + M^{\mathrm{T}}\tau + \frac{2}{\kappa}(1-e^{-\kappa\tau})RQ\Big)\Big] \\ &\quad - \frac{1}{2}\beta\mathrm{Tr}\Big(\ln H(\tau; u) + M^{\mathrm{T}}\tau + \frac{2\mathrm{i}u}{\kappa}(1-e^{-\kappa\tau})RQ\Big)\Big) \\ &= \exp\Big(\mathrm{i}u\ln F_T(t) + \mathrm{Tr}\Big(\Big(A(\tau; u) - \mathrm{i}uA(\tau; -\mathrm{i})\Big)\Sigma\Big) \\ &\quad - \frac{1}{2}\beta\mathrm{Tr}\Big(\ln H(\tau; u) - \mathrm{i}u\ln H(\tau; -\mathrm{i}) + M^{\mathrm{T}}\tau(1-\mathrm{i}u)\Big)\Big). \end{aligned}$$

Appendix D

Numerical Methods

D.1 The Runge-Kutta Methods

In the numerical solution of differential equations, the Taylor-series method is simple to implement but has the drawback of requiring higher order derivatives and some error analysis prior to implementation. To circumvent these problems, the family of Runge-Kutta methods imitates the Taylor-series method by means of clever combinations of the values of the first derivative through the repeating use of chain rule of differentiation.

Consider the general (non-linear) ordinary differential equation (ODE)

$$\frac{d}{d\tau}x(t) = f(t, x(t)),$$

where $x(a) \in \mathbb{C}^{n\times n}$. The interval $[a, b]$ is partitioned into N subintervals $[t_{k-1}, t_k]$, where $k = 1, \ldots, N$ with $t_0 = a$ and $t_N = b$. A numerical method for ODE approximates the value of $x(t_{k+1})$ with a time-stepping advance procedure of the form

$$x_{k+1} = \Psi(h_k, x_k, \ldots, h_0, x_0),$$

starting from $x(t_0)$, where $h_k = t_{k+1} - t_k$. Here the map Ψ depends on the specific numerical method applied.

Now, the general class of m-stage Runge-Kutta methods are characterized by the real numbers a_{ij}, b_i, for $i, j = 1, \ldots, m$, and $c_i = \sum_{j=1}^{m} a_{ij}$, as

$$Y_i = x_k + h_k \sum_{j-1}^{m} a_{ij} f(t_k + c_j h_k, Y_j),$$

for $i = 1, \ldots, m$, and,

$$x_{k+1} = x_k + h_k \sum_{i=1}^{m} b_i f(t_k + c_i h_k, Y_i),$$

where Y_i, for $i = 1, \ldots, m$, are the intermediate stages. The associated coefficients are usually displayed with the so-called Butcher tableau as follows:

$$\begin{array}{c|ccc} c_1 & a_{11} & \cdots & a_{1m} \\ \vdots & \vdots & & \vdots \\ c_m & a_{m1} & \cdots & a_{mm} \\ \hline & b_1 & \cdots & b_m \end{array}$$

If $a_{ij} = 0$, for $j \geq i$, then the intermediate stages Y_i can be evaluated recursively and the method is explicit. In that case, the zero a_{ij} coefficients (in the upper triangular part of the tableau) are omitted for clarity. With this notation, the 4-stage fourth-order method Runge-Kutta method applied in currency option pricing is given by

$$\begin{array}{c|cccc} 0 & & & & \\ \frac{1}{2} & \frac{1}{2} & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & \\ \frac{1}{2} & 0 & 0 & 1 & \\ \hline & \frac{1}{6} & \frac{2}{6} & \frac{2}{6} & \frac{1}{6} \end{array}$$

while the 7-stage sixth-order Runge-Kutta method applied in cap pricing is given by

0							
$\frac{5-\sqrt{5}}{10}$	$\frac{5-\sqrt{5}}{10}$						
$\frac{5+\sqrt{5}}{10}$	$\frac{-\sqrt{5}}{10}$	$\frac{5+2\sqrt{5}}{10}$					
$\frac{5-\sqrt{5}}{10}$	$\frac{-15+7\sqrt{5}}{20}$	$\frac{-1+\sqrt{5}}{4}$	$\frac{15-7\sqrt{5}}{10}$				
$\frac{5+\sqrt{5}}{10}$	$\frac{5-\sqrt{5}}{60}$	0	$\frac{1}{6}$	$\frac{15+7\sqrt{5}}{60}$			
$\frac{5-\sqrt{5}}{10}$	$\frac{5+\sqrt{5}}{60}$	0	$\frac{9-5\sqrt{5}}{12}$	$\frac{1}{6}$	$\frac{-5+3\sqrt{5}}{10}$		
1	$\frac{1}{6}$	0	$\frac{-55+25\sqrt{5}}{12}$	$\frac{-25-7\sqrt{5}}{12}$		$\frac{5+\sqrt{5}}{2}$	
	$\frac{1}{12}$	0	0	0	$\frac{5}{12}$	$\frac{5}{12}$	$\frac{1}{12}$

Very briefly, the m-order Runge-Kutta method is called an *explicit one-step method* and involves *local truncation error* of order m so, as a one-step method, it is *convergent*. Moreover, it is *strongly stable*. For the efficiency of implementation, we fix the stepsize; however, to impose error control, *adaptive* algorithms, for example, the Runge-Kutta-Fehlberg Method, which vary stepsize to accommodate local peculiarities in the solution should be implemented. More results can be found in Chapter 5 of Burden and Faires [20] or Chapter 7 of Allen and Isaacson [1].

D.2 Approximation of Matrix Exponentials

We quote some easy-to-implement methods from Moler and Van Loan [63] for the computation of matrix exponential. Suppose that A is a square matrix with complex entries. The exponential of matrix A, denoted as $\exp\Big(A\Big)$, is formally defined by the convergent power series

$$\exp\Big(A\Big) \triangleq \sum_{n=0}^{\infty} \frac{A^n}{n!}.$$

The most intuitive way to compute $\exp\Big(A\Big)$ is to truncate the above power series into a finite sum; however, a large number of terms have to be compute in order to assure a certain degree of accuracy and, thus, this method is computationally expensive.

Alternatively, the simplest way to compute $\exp\Big(A\Big)$ is to use the eigenvalue decomposition. Suppose that the matrix A is decomposed as $A = P^{-1}DP$, where D is the diagonal matrix of the eigenvalues of matrix A and P contains the columns of the corresponding eigenvectors of matrix A. Then, we have

$$\exp\Big(A\Big) = P^{-1}\exp\Big(D\Big)P,$$

and $\exp\Big(D\Big)$ is simply the diagonal matrix of the exponentiated diagonal elements of D. This method gives the exact answer; however, the theoretical difficulty occurs when A does not have a complete set of linearly independent eigenvectors and is thus defective. In this case, there is no invertible matrix P of eigenvectors and the algorithm breaks down. An example of defective matrix is

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Another solution is to invoke a numerical ODE solver. Since $\exp\Big(tA\Big)$ is the solution to the ODE

$$\begin{aligned} \frac{d}{dt}x(t) &= Ax(t) \\ x(0) &= \mathbf{I}_n, \end{aligned}$$

it is straightforward to consider compute the solution by a numerical ODE solver, adaptive or non-adaptive, single-step or multi-step, implicit or explicit. One example is the

Runge-Kutta methods. However, an obvious disadvantage is the cost of intensive computation.

The last method is the combination of Padé approximation and the method scaling and squaring. This method exploits the fundamental equality

$$\exp\Big(A\Big) = \exp\Big(A/m\Big)^m.$$

First, choose the smallest integer j such that

$$\|A\| \leq 2^{j-1},$$

where $\|A\| = \max_{\|x\|=1} \|Ax\|$ and $\|x\| = \Big(\sum_{i=1}^{n} |x_i|^2\Big)^{\frac{1}{2}}$. Now, let $m = 2^{j-1}$. The (p, q) **Padé approximation of the matrix** $\exp\Big(A/m\Big)$ is given by

$$R_{pq}(A/m) = [D_{pq}(A/m)]^{-1} N_{pq}(A/m),$$

where

$$\begin{aligned} D_{pq}(A/m) &= \sum_{j=0}^{q} \frac{(p+q-j)!q!}{(p+q)!j!(q-j)!}(-A/m)^j \\ N_{pq}(A/m) &= \sum_{j=0}^{p} \frac{(p+q-j)!p!}{(p+q)!j!(p-j)!}(A/m)^j. \end{aligned}$$

It is shown that

$$[R_{pq}(A/m)]^m = \exp\Big(A + E\Big),$$

where

$$\|E\| \leq \frac{1}{2^{p+q-3}} \frac{p!q!}{(p+q)!(p+q+1)!},$$

that is, the above approximation makes an error bounded by a small number. The combination of Padé approximation with scaling and squaring is implemented in MATLAB by the function expm. The code in expm1.m implements scaling and squaring and the $(6,6)$ Padé approximation.

D.3 The Simulation of Forward Rates in the Wishart model

To simulate forward rates in the Wishart model, we apply the easy-to-implement and efficient OU-discretization scheme proposed by Gauthier and Possamaï [37] with β being positive integers. With a simple change of probability measure, it can be extended to all values of $\beta \geq n+1$, but it cannot be used when $\beta \in (n-1, n+1) \setminus \{n\}$.

The Relationship between OU Process and Wishart Process

Let $\beta \in \mathbb{N}$ and let $\{X_{m,t} : t \geq 0\}_{1 \leq m \leq \beta}$ be independent OU processes in $\mathbb{R}^n$ which follow the dynamics

$$dX_{m,t} = MX_{m,t}dt + Q^{\mathrm{T}}dW_{m,t},$$

where $\{W_{m,t} : t \geq 0\}_{1 \leq m \leq \beta}$ are independent vector Brownian motions in $\mathbb{R}^n$, M and Q are some real-valued square matrices of order n. Then, the matrix process defined by

$$\Sigma_t = \sum_{m=1}^{\beta} X_{m,t}X_{m,t}^{\mathrm{T}}$$

follows the dynamics

$$d\Sigma_t = (\beta Q^{\mathrm{T}}Q + M\Sigma_t + \Sigma_t M^{\mathrm{T}})dt + \sqrt{\Sigma_t}dW_t Q + Q^{\mathrm{T}}(dW_t)^{\mathrm{T}}\sqrt{\Sigma_t},$$

where W_t is a square matrix Brownian motion. Note that

$$\sqrt{\Sigma_t}dW_t \sim \sum_{m=1}^{\beta} X_{m,t}dW_{m,t}^{\mathrm{T}}.$$

The Simulation Procedure of Forward Rates

Under the T_{j+1}-forward measure, the log-forward rate follows the dynamics

$$\begin{aligned}\ln L_j(t+\Delta t) &= \ln L_j(t) - \frac{1}{2}\int_t^{t+\Delta t} \mathrm{Tr}\Big(U_j(s)^2\Sigma_s\Big)ds + \int_t^{t+\Delta t} \mathrm{Tr}\Big(U_j(s)\sqrt{\Sigma_s}dW_s R^{\mathrm{T}}\Big) \\ &\quad + \int_t^{t+\Delta t} \mathrm{Tr}\Big(U_j(s)\sqrt{\Sigma_s}dB_s\sqrt{\mathbf{I}_n - RR^{\mathrm{T}}}\Big),\end{aligned}$$

where B_t and W_t are independent matrix Brownian motions. The stochastic integrals in the above expression can be approximated as follows:

$$\begin{aligned}\int_t^{t+\Delta t} \mathrm{Tr}\Big(U_j(s)^2\Sigma_s\Big)ds &\approx \frac{\Delta t}{2}\mathrm{Tr}\Big(U_j(t+\Delta t)^2\Sigma_{t+\Delta t} + U_j(t)^2\Sigma_t\Big) \\ \int_t^{t+\Delta t} \mathrm{Tr}\Big(U_j(s)\sqrt{\Sigma_s}dW_s R^{\mathrm{T}}\Big)ds &= \mathrm{Tr}\Big(\sum_{m=1}^{\beta}\int_t^{t+\Delta t} U_j(s)X_{m,s}dW_{m,s}^{\mathrm{T}}R^{\mathrm{T}}\Big) \\ &\approx \sqrt{\Delta t}\mathrm{Tr}\Big(U_j(t)\sum_{m=1}^{\beta} X_{m,t}\epsilon_{m,t+\Delta t}^{\mathrm{T}}R^{\mathrm{T}}\Big) \\ \int_t^{t+\Delta t} \mathrm{Tr}\Big(U_j(s)\sqrt{\Sigma_s}dB_s\sqrt{\mathbf{I}_n - RR^{\mathrm{T}}}\Big)ds &\sim \mathcal{N}\Big(0, \int_t^{t+\Delta t}\mathrm{Tr}\Big(U_j(s)\Sigma_s U_j(s)(\mathbf{I}_n - RR^{\mathrm{T}})\Big)ds\Big) \\ &\approx Z\sqrt{\begin{array}{l}\frac{\Delta t}{2}\mathrm{Tr}\Big(\Big(U_j(t+\Delta t)\Sigma_{t+\Delta t}U_j(t+\Delta t) \\ +U_j(t)\Sigma_t U_j(t)\Big)(\mathbf{I}_n - RR^{\mathrm{T}})\Big)\end{array}},\end{aligned}$$

where $\{\epsilon_{m,t+\Delta t}\}_{1\le m\le\beta}$ are vectors in $\mathbb{R}^n$ of independent standard normal random variables for the time interval $[t, t+\Delta t]$ and Z is a standard normal random variable.

Here is the simulation procedure:

Step 1: Partition the interval $[0, T_j]$ into N equal subintervals $\{t_i, t_{i+1}\}_{i=0}^{N-1}$ such that

$t_0 = 0$, $t_{i+1} = t_i + \Delta t$ and $\Delta t = T_j/N$.

Step 2: To initialize the simulation, an eigenvalue decomposition on the symmetric positive definite matrix Σ_{t_0} yields

$$\Sigma_{t_0} = \sum_{m=1}^{n} \lambda_m \phi_m \phi_m^{\mathrm{T}}$$

where λ_m are the eigenvalues and ϕ_m are the corresponding eigenvectors of Σ_{t_0}. Thus, the initial state of the vector OU process is, for $m = 1, \ldots, \beta$,

$$X_{m,t_0} = \mathbf{1}_{1 \le m \le n} \sqrt{\lambda_m} \phi_m.$$

Step 3: To generate the OU process, for $k = 1, \ldots, \beta$,

$$X_{m,t_{i+1}} = X_{m,t_i} + \Delta t M^{T_{j+1}}(t_i) X_{m,t_i} + \sqrt{\Delta t} Q^{\mathrm{T}} \epsilon_{m,t_{i+1}},$$

where $\{\epsilon_{m,t_{i+1}}\}_{1 \le m \le \beta}$ are vectors independent standard normal random variables in $\mathbb{R}^n$ for the time interval $[t_i, t_{i+1}]$.

Step 4: The Wishart process is generated by

$$\Sigma_{t_{i+1}} = \sum_{m=1}^{\beta} X_{m,t_{i+1}} X_{m,t_{i+1}}^{\mathrm{T}}.$$

Step 5: To generate log-forward rate under the Wishart model,

$$\begin{aligned} \ln_j L(t_{i+1}) \;=\; & \ln L_j(t_i) - \frac{1}{4}\Delta t \mathrm{Tr}\Big(U_j(t_{i+1})^2 \Sigma_{t_{i+1}} + U_j(t_i)^2 \Sigma_{t_i}\Big) + \sqrt{\Delta t}\mathrm{Tr}\Big(U_j(t_i) \sum_{m=1}^{\beta} X_{m,t_i} \epsilon_{m,t_{i+1}}^{\mathrm{T}} R^{\mathrm{T}}\Big) \\ & + \sqrt{\frac{\Delta t}{2}\mathrm{Tr}\Big(\Big(U_j(t_{i+1})\Sigma_{t_{i+1}} U_j(t_{i+1}) + U_j(t_i)\Sigma_{t_i} U_j(t_i)\Big)(\mathbf{I}_n - RR^{\mathrm{T}})\Big)}\, Z. \end{aligned}$$

Step 6: Repeat Steps 3 to 5 for $i = 1, \ldots, N-1$.

Bibliography

[1] Allen III, M.B., and Isaacson E.L. (1998). Numerical analysis for applied sicence, John Wiley & Sons, Inc.

[2] Andersen, L., and Andreasen, J. (2000). Volatility skews and extensions of the LIBOR market model, *Applied Mathematical Finance*, Vol. 7, 1-32.

[3] Andersen, L., and Brotherton-Ratcliffe, R. (2005). Extended LIBOR market models with stochastic volatility, *Journal of Computational Finance*, Vol. 9, No. 1, 1-40.

[4] Andersen, T.G., and Benzoni, L. (2010). Do bonds span volatility risk in the U.S. treasury market? A specification test for affine term structure models, *The Journal of Finance*, Vol. 65, No. 2, 603-653.

[5] Andersen, T.W. (1946). The non-central Wishart distribution and certain problems of multivariate statistics, *The Annals of Mathematical Statistics*, Vol. 17, No. 4, 409-431.

[6] Bachert, P., Gatarek, D., and Maksymiuk, R. (2006). The LIBOR market model in practice, *Wiley Finance Series*, John Wiley & Sons.

[7] Bates, D.S. (2000). Post-'87 crash fears in the S&P 500 futures option market, *Journal of Econometrics*, Vol. 94, 181-238.

[8] Belomestnyl, D., Mathew, S., and Schoenmakers, J. (2009). Multiple stochastic volatility extension of the LIBOR market model and its implementation, *Monte Carlo Methods and Applications*, Vol. 15, No. 4, 285-310.

[9] Benabid, A., Bensusan, H., and El Karoui, N. (2009). Wishart stochastic volatility: Asymptotic smile and numerical framework. *Working paper.*

[10] Bikbov, R., and Chernov, M. (2009). Unspanned stochastic volatility in affine models: Evidence from Eurodollar futures and options, *Management Science*, Vol. 55, No. 8, 1292-1305.

[11] Black, F. (1976). The pricing of commodity contracts, *Journal of Financial Economics*, Vol. 3, 167-179.

[12] Black, F., and Scholes, M. (1973). The pricing of options and corporate liabilities, *Journal of Political Economy*, Vol. 81, 637-653.

[13] Bleaney, M.F., Leybourne, S.J., and Mizen, P. (1999). Mean reversion of real exchange rates in high-inflation countries, *Southern Economic Journal*, Vol. 65, No. 4, 839-854.

[14] Brace, A., Gatarek, D., and Musiela, M. (1997). The market model of interest rate dynamics, *Mathematical Finance*, Vol. 7, No. 2, 127-155.

[15] Brace, A., Dunn, T., and Barton, G. (2001). Towards a central interest tate model, *Handbooks in Mathematical Finance*, Cambridge, Cambridge University Press.

[16] Branger, N., and Muck, M. (2009). Keep on smiling? Volatility surface and the pricing of quanto options when all covariances are stochastic, Working Paper.

[17] Brigo, D., and Mercurio, F. (2007). Interest rate models: Theory and practice, Third Edition, *Springer*, Berlin Heidelberg, New York.

[18] Bru, M.F. (1991). Wishart processes, *Journal of Theoretical Probability*, Vol. 4, 725-743.

[19] Buraschi, A., Porchia, P., and Trojani, F. (2010). Correlation risk and optimal portfolio choice, *The Journal of Finance*, Vol. 65, No. 1, 393-420.

[20] Burden, R.L., and Faires, J.D. (2011). Numerical analysis, Ninth Edition, Brooks/Cole, Cengage Learning.

[21] Carr, P., and Wu, L. (2007). Stochastic skew in currency options, *Journal of Financial Economics*, Vol. 86, 213-247.

[22] Carr, P., and Madan, D.B. (1999). Option valuation using the fast Fourier transform, *Journal of Computational Finance*, Vol. 2, No. 4, 61-73.

[23] Casassus, J., Collin-Dufresne, P., and Goldstein, B. (2005). Unspanned stochastic volatility and fixed income derivatives pricing, *Journal of Banking and Finance*, Vol. 29, 2723-2749.

[24] Chen, S.N., and Jeon, K. (1999). Mean reversion behavior of the returns on currency assets, *International Review of Economics and Finance*, Vol. 7, No. 2, 185-200.

[25] Christoffersen, P., Heston, S., and Jacobs, K. (2009). The shape and term structure of the index option smirk: Why multifactor stochastic volatility models work so well?, *Management Science*, Vol. 55, No. 12, 1914-1932.

[26] Collin-Dufresne, P., and Goldstein, R. (2002). Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility, *The Journal of Finance*, Vol. 57, No. 4, 1685-1730.

[27] Collin-Dufresne, P., Goldstein, R., and Jones, C. (2009). Can interest rate volatiltiy be extracted from the cross section of bond yields?, *Journal of Financial Economics*, Vol. 94, 47-66.

[28] Da Fonseca, J., and Grasselli, M. (2010). Riding on the smiles, *SSRN*.

[29] Da Fonseca, J., Grasselli, M., and Tebaldi, C. (2007). Option pricing when correlations are stochastic-an analytical framework, *Review of Derivatives Research*, Vol. 10, No. 2, 151-180.

[30] Da Fonseca, J., Grasselli, M., and Tebaldi, C. (2008). A multifactor volatility Heston model, *Quantitative Finance*, Vol. 8, No. 6, 591-604.

[31] Dai, Q., and Singleton, K.J. (2000). Specification analysis of affine term structure models, *The Journal of Finance*, Vol. 55, No. 5, 1943-1978.

[32] Duffie, D., and Kan, R. (1996). A yield-factor model of interest rates, *Mathematical Finance*, Vol. 6, No. 4, 379-406.

[33] Duffie, D., Pan, J., and Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions, *Econometrica*, Vol. 68, No. 6, 1343-1376.

[34] Ejvall, N., Jennergren, L.P., and Naslund, B. (1997), Currency option pricing with mean reversion and uncovered interest parity: A revision of the Garman-Kohlhagen model, *European Journal of Operational Research*, Vol. 100, 41-59.

[35] Freiling, G. (2002), A survey of nonsymmetric Riccati equations, *Linear Algebra and its Applications*, Vol. 351-352, 243-270.

[36] Garman, M., and Kohlhagen, S. (1983). Foreign currency option values, *Journal of International Money and Finance*, Vol 2, 231-237.

[37] Gauthier, P., and Possamaï, D. (2009). Efficient simulation of the Wishart model, *Working Paper.*

[38] Gourieroux, C. (2006). Continuous time Wishart process for stochastic risk, *Econometric Review*, Vol. 25, 177-217.

[39] Gourieroux, C., Jasiak, J., and Sufana, R. (2009). The Wishart autoregressive process of multivariate stochastic volatility, *Journal of Econometrics*, Vol. 150, 167-181.

[40] Gourieroux, C., and Sufana, R. (2010). Derivative pricing with Wishart multivariate stochastic volatility, *Journal of Business and Economics Statisitcs*, Vol. 28, No. 3, 438-451.

[41] Gourieroux, C., and Sufana, R. (2011). Discrete time Wishart term structure models, *Journal of Economic Dynamics and Control*, Vol. 35, No. 6, 815-824.

[42] Grasselli, M., and Tebaldi, C. (2008). Solvable affine term structure models, *Mathematical Finance*, Vol. 18, No. 1, 135-153.

[43] Gruber, P.H., Tebaldi, C., and Trojani, F. (2010). Three make a dynamic smile-unspanned skewness and interacting volatility components in option valuation, *SSRN*.

[44] Hagan, P.S., Kumar, D., Lesniewski, A.S., and Woodward, D.E., (2002). Managing smile risk, *Wilmott Magazine*, September, 84-108.

[45] Han, B. (2007). Stochastic volatilities and correlations of bond yields, *The Journal of Finance*, Vol. 62, No. 3, 1491-1524.

[46] Heath, D., Jarrow, R., and Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation, *Econometrica*, Vol. 60, No. 1, 77-105.

[47] Heidari, M., and Wu, L. (2003). Are interest rate derivatives spanned by the term structure of interest rates, *Journal of Fixed Income*, Vol 13, 75-86.

[48] Heston, S.L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options, *Review of Financial Studies*, Vol. 6, No. 2, 327-343.

[49] Hull, J., and White, A. (2000). Forward rate volatilities, swap rate volatilities and the implementation of the LIBOR market model, *Journal of Fixed Income*, Vol. 10, 46-62.

[50] Jamshidian, F. (1997). LIBOR and swap market models and measures, *Finance and Stochastics*, Vol. 1, 293-330.

[51] Jarrow, R., Li, H., and Zhao, F. (2007). Interest rate caps "smile" too! But can the LIBOR market models capture the smile?, *The Journal of Finance*, Vol. 62, No. 1, 345-382.

[52] Jorion, P., and Sweeney, R.J. (1996). Mean reversion in real exchange rates: Evidence and implications for forecasting, *Journal of International Money and Finance*, Vol. 15, No. 4, 535-550.

[53] Joshi, M.S., and Rebonato, R. (2003). A displaced-diffusion stochastic volatility LIBOR market model: Motivation, definition and implementation, *Quantitative Finance*, Vol. 3, No. 6, 458-469.

[54] Leung, K.S., Ng, H.Y., and Wong, H.Y. (2011). Currency option pricing with Wishart process, *Working Paper*

[55] Lewis, A. (2000). Option valuation under stochastic volatility, *Finance Press*, Newport Beach, CA.

[56] Leybourne S., Newbold, P., and Sollis, R. (2002). Tests for symmetric and asymmetric nonlinear mean reversion in real exchange rates, *Journal of Money, Credit, and Banking*, Vol. 34, No. 3, 686-700.

[57] Lindberg, H., and Söderlind, P. (1994). Intervention policy and mean reversion in exchange rate target zones: The Swedish case, *The Scandinavian Journal of Economics*, Vol. 96, No. 4, 499-513.

[58] Li, H., and Zhao, F. (2006). Unspanned stochastic volatility: Evidence from hedging interest rate derivatives, *The Journal of Finance*, Vol. 61, No. 1, 341-378.

[59] Litterman, R., and Scheinkman, J. (1991). Common factors affecting bond returns, *Journal of Fixed Income*, Vol. 1, 62-74.

[60] Longstaff, F., Santa-Clara, P. and Schwartz, E. (2001). The relative valuation of caps and swaptions: Theory and evidence, *The Journal of Finance*, Vol. 56, No. 6, 2067-2109.

[61] Ma, K. (2010). Cross-currency LIBOR market model with stochastic volatilities, *Hong Kong Unveristy of Science and Technology*, MPhil Thesis.

[62] Miltersen, K.R., Sandmann, K., and Sondermann D. (1997). Closed form solutions for term structure derivatives with log-normal interest rates, *The Journal of Finance*, Vol. 52, No. 1, 409-430.

[63] Moler, C., and Van Loan, C. (2003). Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later, *SIAM Review*, Vol. 45, No. 1, 3-49

[64] O'Hara, J.G., and Pillay, E. (2011). FFT based option pricing under a mean reverting process with stochastic volatility and jumps, *Journal of Computational and Applied Mathematics*, Vol. 235, 3378-3384.

[65] Palm, F.C., and Vlaar, P.J.G. (1993). The message in weekly exchange rates in the European monetary system: Mean reversion, conditional heteroscedasticity, and jumps, *Journal of Business and Economic Statistics*, Vol. 11, No. 3, 351-360.

[66] Peel, D.A., Sarno, L., and Taylor, M.P. (2001). Nonlinear mean-reversion in real exchange rates: Toward a solution to the purchasing power parity puzzles, *International Economic Review*, Vol. 42, No. 4, 1015-1042.

[67] Piterbarg, V. (2005). Stochastic volatility model with a time-dependent skew, *Applied Mathematical Finance*, Vol. 12, No. 2, 147-185.

[68] Rebonato, R. (1998). Interest rate option models, Second Edition, *Wiley*, Chichester.

[69] Rebonato, R. (2002). Modern pricing of interest-rate derivatives, Second Edition, *Princeton University Press*, New Jersey.

[70] Reiswich, D., and Wystup. U. (2010). A guide to FX options quoting conventions, *Journal of Derivatives*, Vol. 18, No. 2, 58-68.

[71] Sarno, L., and Taylor, M.P. (1998). Real exchange rates under the recent float: Unequivocal evidence of mean reversion, *Economics Letters*, Vol. 60, 131-137.

[72] Schwartz, E., and Trolle, A.B. (2009a). A general stochastic volatility model for the pricing of interest rate derivatives, *Review of Financial Studies*, Vol. 22, No. 5, 2007-2057.

[73] Schwartz, E., and Trolle, A.B. (2009b). Unspanned stochastic volatility and the pricing of commodity derivatives, *Review of Financial Studies*, Vol. 22, No. 11, 4423-4461.

[74] Sweeney, R.J. (2006). Mean reversion in G-10 nominal exchange rates, *Journal of Financial and Quantitative Analysis*, Vol. 41, 685-708.

[75] Wishart, J. (1928). "The Generalised Product Moment Distribution in Samples from a Normal Multivariate Population,". *Biometrika*, 20A(1/2), 32-52.

[76] Wong, H.Y., and Lo, Y.W. (2009). Option pricing with mean reversion and stochastic volatility, *European Journal of Operational Research*, Vol. 197, 179-187.

[77] Wong, H.Y., and Zhao, J. (2010). Currency option pricing: Mean reversion and multi-scale stochastic volatility, *Journal of Futures Markets*, Vol. 30, No. 10, 938-956.

[78] Wu, L. (2003). Fast at-the-money calibration of LIBOR market model through Lagrange multipliers, *Journal of Computational Finance*, Vol. 6, No. 2, 39-77.

[79] Wu, L. (2003). On the calibration of the market model with a square-root volatility process, *Working Paper*.

[80] Wu, L., and Zhang, F. (2006). LIBOR market model with stochastic volatility, *Journal of Industrial and Management Optimization*, Vol. 2, No. 2, 199-227.