WARWICK
THE UNIVERSITY OF WARWICK

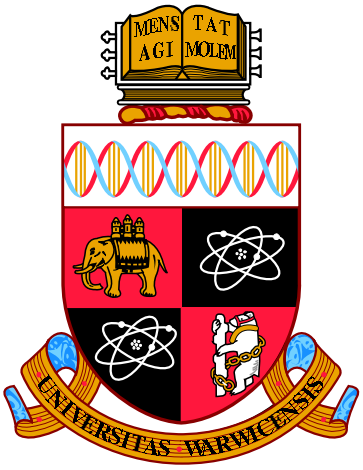
A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:
http://wrap.warwick.ac.uk/91418

Copyright and reuse:
This thesis is made available online and is protected by original copyright.
Please scroll down to view the document itself.
Please refer to the repository record for this item for information to help you to cite it.
Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

warwick.ac.uk/lib-publications



The Stochastic Volatility Markov-functional Model

by

Chuan Guo

Thesis

Submitted to the University of Warwick

for the degree of

Doctor of Philosophy

Depatment of Statistics

September 2016

THE UNIVERSITY OF
WARWICK

Contents

List of Tables **iv**

List of Figures **vi**

Acknowledgments **vii**

Declarations **viii**

Abstract **ix**

Chapter 1 Introduction **1**

Chapter 2 Interest rate markets and models **7**

2.1 Interest rate markets 7

2.1.1 Basic instruments 7

2.1.2 Interest rate options 9

2.2 Interest rate models 13

2.2.1 The role of models 14

2.2.2 Short rate models 14

2.2.3 Heath-Jarrow-Morton models 15

2.2.4 Market models 16

2.2.5 Markov-functional models 17

Chapter 3 Stochastic volatility Markov-functional models **19**

3.1 Introduction 19

3.2 Markov-functional models 21

3.2.1 One-dimensional Markov-functional models: LIBOR version . 22

3.2.2 One-dimensional Markov-functional models: Swap rate version 26

3.2.3 Multi-dimensional Markov-functional models 28

3.3 Stochastic volatility Markov-functional models 30

3.3.1 Numerical implementation 30
3.3.2 Specification of the driving process 36
3.4 Conclusion 44
3.A Appendix: Equivalence of Markov-functional models 45
3.B Appendix: Basis functions 49
3.C Appendix: Proof of Proposition 1 53
3.D Appendix: Proof of Proposition 2 56

Chapter 4 Stochastic volatility Markov-functional models: calibration 61
4.1 Introduction 61
4.2 Calibration problem description 63
4.3 Choice of marginals 65
4.3.1 Separable LIBOR market models with local volatility 65
4.3.2 Separable LIBOR market models with stochastic volatility . . 67
4.4 Calibration to caplets 69
4.4.1 SABR model 69
4.4.2 Calibration routine 70
4.4.3 Numerical study 74
4.5 Calibration to swaptions 76
4.5.1 Calibration routine 77
4.5.2 Numerical study 82
4.6 Calibration to market correlations 83
4.6.1 Calibration routine 84
4.6.2 Numerical study 88
4.7 Conclusion 89
4.A Appendix: Proof of Proposition 4 91
4.B Appendix: Swap rate based stochastic volatility Markov-functional models: calibration 94
4.B.1 Choice of marginals 95
4.B.2 Calibration to swaptions 95
4.B.3 Calibration to market correlations 98

Chapter 5 Comparison of stochastic volatility Markov-functional model and one-factor Markov-functional models 102
5.1 Introduction 102
5.2 Methodology 105
5.2.1 Bermudan swaptions 105

5.2.2 New Bermudan product . . . 107
5.2.3 One-dimensional swap Markov-functional models: driver specification and calibration . . . 111
5.2.4 Stochastic volatility Markov-functional model . . . 117
5.2.5 Expressions for vegas in Markov-functional models . . . 118
5.3 Market data . . . 121
5.4 Bermudan swaption comparison . . . 122
5.4.1 Pricing comparison results . . . 122
5.4.2 Vega comparison results . . . 127
5.5 New Bermudan product comparison . . . 130
5.5.1 Pricing comparison results . . . 130
5.5.2 Vega comparison results . . . 131
5.6 Conclusion . . . 133

Chapter 6 Quasi-Gaussian models and Markov-functional models 136
6.1 Introduction . . . 136
6.2 Quasi-Gaussian model . . . 138
6.2.1 General one-factor Quasi-Gaussian model . . . 138
6.2.2 Displaced diffusion type local volatility . . . 140
6.2.3 Calibration to swaptions . . . 141
6.2.4 Calibration to the market correlation structure . . . 144
6.2.5 Stochastic volatility Quasi-Gaussian model . . . 145
6.3 Comparison between Quasi-Gaussian models and Markov-functional models . . . 148
6.3.1 Review of Quasi-Gaussian model under Markov-functional model framework . . . 148
6.3.2 Calibration comparison . . . 151
6.A Appendix: Markovian projection . . . 153

List of Tables

2.1 Examples of short rate models 15

4.1 Calibration performance for implied volatilities (%) of the ATM caplets 99
4.2 Calibration performance for implied volatilities (%) of the ATM co-terminal swaptions 100
4.3 Scenarios for the SV MFM. 101
4.4 Calibration performance for one-step correlation of LIBORs $Corr(ln(L^i_{T_i} + \theta_i), ln(L^{i+1}_{T_{i+1}} + \theta_{i+1}))$ under SV MFMs with different scenarios 101

5.1 ATM implied volatilities in the swaption matrix linked to the Hull-White driver. 114
5.2 ATM implied volatilities in the swaption matrix linked to the one step covariance driver and the SABR driver. 115
5.3 The expressions for vegas $v_{i,k}$ for a particular i of a Bermudan swaption under swap MFMs where $\theta_i := \frac{d\tilde{\sigma}^{Berm}_{i,n+1-i}}{d\tilde{\sigma}^{ATM}_{i,n+1-i}}$. 120
5.4 Initial zero-coupon bonds on 11 March 2015. 122
5.5 Implied volatilities (%) of the co-terminal (31Y) swaptions against expiry and relative strike (bp) to the ATM on 11 March 2015 123
5.6 Implied volatilities (%) of the caplets against expiry and relative strike (bp) to the ATM on 11 March 2015 124
5.7 Implied volatilities (%) of the ATM swaptions against expiry and tenor (1Y - 10Y) on 11 March 2015 125
5.8 Implied volatilities (%) of the ATM swaptions against expiry and tenor (11Y - 20Y) on 11 March 2015 126
5.9 Implied volatilities (%) of the ATM swaptions against expiry and tenor (21Y - 30Y) on 11 March 2015 127
5.10 Scenarios for the SV MFM. 127

5.11 The prices of the Bermudan swaption under different MFMs with notional $N = 100$ million. 128
5.12 The difference between the price for the Bermudan swaption produced by different MFMs and the SV MFM (I) measured by total vega. . . 129
5.13 The prices of the new Bermudan product under different MFMs with notional $N = 100$ million. 132
5.14 The difference between the price for the new Bermudan product produced by different MFMs and the SV MFM (I) measured by total vega. 133

List of Figures

5.1 The sum of the vegas $\sum_{k=1}^{n+1-i} v_{i,k}$ for $i = 1, ..., 30$ of the Bermudan swaption under the different MFMs. 135
5.2 The sum of the vegas $\sum_{k=1}^{n+1-i} v_{i,k}$ for $i = 1, ..., 30$ of the new Bermudan product under the different MFMs. 135

Acknowledgments

First and foremost, I would like to thank my supervisor Dr Jo Kennedy for her support and patience. Without her help and suggestions throughout the last four years, this thesis would not have been possible. Thank you so much Jo.

Second, I thank my review panel, Dr Anastasia Papavasiliou and Professor Gareth Roberts, for your valuable advice and feedback of my presentation and reports.

Third, I would like to thank Duy Pham and Jaka Gogala for sharing with me your C++ code and all discussions we have. Many thanks go to the other members in the Statistics department. I have really enjoyed this friendly environment in the department.

Fourth, I am grateful to the scholarship I have received from the university of Warwick. I cannot focus all my mind on this thesis without the financial support. Thanks for providing me this valuable opportunity to study here.

I am also grateful to my friends who accompany me throughout the PhD period. Jian and Bing, I will never forget the basketball time every weekend and your brain teasers. I also want to thank my roommates Dave and Matt. Thank you for improving my English and all the joy we had together.

Finally, I would like to thank my parents and my wife for their support and love.

Declarations

I declare that this thesis is based on my own work. The thesis is original except where specified by the references. The thesis has not been submitted for any other degree at any other university. The material in this thesis has not been submitted to any other publication.

Abstract

In this thesis we study low-dimensional stochastic volatility interest rate models for pricing and hedging exotic derivatives. In particular we develop a stochastic volatility Markov-functional model. In order to implement the model numerically, we further propose a general algorithm by working with basis functions and conditional moments of the driving Markov process. Motivated by a data driven study, we choose a SABR type model as a driving process. With this choice we specify a pre-model and develop an approximation to evaluate conditional moments of the SABR driver which serve as building blocks for the practical algorithm.

Having discussed how to set up a stochastic volatility Markov-functional model next we study the calibration of a LIBOR based version of the model with the SABR type driving process. We consider a link between separable SABR LIBOR market models and stochastic volatility LIBOR Markov-functional models. Based on the link we propose a calibration routine to feed in SABR marginals by calibrating to the market vanilla options. Moreover we choose the parameters of the SABR driver by fitting to the market correlation structure.

We compare the stochastic volatility Markov-functional model developed in the thesis with one-dimensional (non-stochastic-volatility) swap Markov-functional models in terms of pricing and hedging Bermudan type products. By doing so we investigate effects of correlation structure, implied volatility smiles and the introduction of stochastic volatility on Bermudan type products.

Finally we compare Quasi-Gaussian models with Markov-functional models in terms of specification and calibration. In particular we study Quasi-Gaussian models formulated in the Markov-functional model framework to make clear the relationship between the two models.

Chapter 1

Introduction

In this thesis we work in the area of interest rate models. There are three popular categories of term structure models for pricing and hedging interest rate derivatives in practice. They are short rate models, market models and Markov-functional models. Among them, short rate models (1970s) are the earliest term structure models. They were very popular and are still used nowadays in practice because of their tractability. In a short rate model the dynamics of the short rate, which is an instantaneous spot rate and not observable in the market, is specified. So they are low-dimensional models. In LIBOR (swap) market models (1997), the dynamics of a set of contiguous market observable forward rates - LIBORs (swap rates) - are specified. Market models are high-dimensional models even when driven by a single Brownian motion. An advantage for market models is that they allow for calibration to the well-known Black's formula for vanilla options. A Markov-functional model (2000) is specified via the dynamics of a Markov process. In a Markov-functional model the zero coupon bonds, which are underlying assets in the interest rate markets, are a function of the Markov process. Besides the above interest rate models, Heath-Jarrow-Morton framework (1992) models the entire forward rate curve directly and, without appropriate conditions being imposed on the volatility of the instantaneous forward rates, this framework can be non-Markovian and infinite dimensional.

In this thesis we focus on Markov-functional models. Markov-functional models were introduced by Hunt et al. [37]. Usually in a Markov-functional model, the economy is driven by a low-dimensional Markov process which is referred to as the driving process, and the zero coupon bonds are a function of the driving process. In that sense Markov-functional models are similar to short rate models. However in a short rate model, the pricing formula for zero-coupon bonds is given to us. In con-

trast, in a Markov-functional model, the functional forms for zero-coupon bonds can be determined numerically by feeding in marginal distributions of a set of contiguous LIBORs or swap rates. In principle given a driving process, any marginal distributions can be fed into the model. As a result Markov-functional models have more calibration advantage than market models since in addition to the Black's formula Markov-functional models allow for calibration to implied volatility smiles/skews of vanilla options by feeding in appropriate marginal distributions of LIBORs (swap rates). In fact there is a link between Markov-functional models and market models. Bennett and Kennedy [7] compared a one-factor separable LIBOR Market model to a one-dimensional Markov-functional model driven by a Gaussian process calibrated to the Black's formula for caplets, and the two models turned out to be very similar numerically.

The vast majority of the research on all these term structure models is in the non-stochastic volatility setting, and this is especially true for Markov-functional models where we have not found any article in the literature on the stochastic volatility extensions. One main contribution we make in this thesis is the development of a stochastic volatility Markov-functional model.

Typically in a valuation model, the evolution of the underlying asset prices is described via a diffusion process. In a non-stochastic volatility model the volatility of the underlying asset, i.e. the coefficient of the diffusion term, is a deterministic function of the current asset level and time. For example the volatility function of the Black-Scholes [8] model is a constant. In contrast, in a stochastic volatility model, the volatility function is driven by its own SDE and at least one extra (correlated) Brownian motion. There are a number of empirical studies showing that volatility is stochastic in reality in the interest rate markets; see [15], [49] and references therein. But does this justify adding the extra complexity to a model for derivative pricing and hedging? Without stochastic volatility the set of distributions we can get for the underlying asset is constrained and does not match reality. In particular it does not have heavy enough tails. Dupire [19] introduced a local volatility model and pointed out that a suitable choice of the local volatility function allows one to fit the marginal distribution of the underlying asset derived from the market prices of European options. However as Hagan et al. [31] pointed out, without stochastic volatility the joint distributions of the underlying asset are not a good reflection of reality. The joint distributions of the underlying assets are very important for pricing and hedging path dependent derivatives such as Bermudan swaptions in practice.

In the literature, we have found many short rate models with stochastic

volatility. See, for example, [24], [49], [18], [12], [5] and references therein. After that, over the last two decades, many studies about stochastic volatility extensions to LIBOR market models have been developed. See, for example, [65], [55], [61], [23] and references therein. But market models are high dimensional so that it is difficult to implement their stochastic volatility version. These articles about stochastic volatility extensions focused on the development and calibration of the models.

In this thesis we study stochastic volatility Markov-functional models. We provide a background introduction in Chapter 2. The basic interest rate products and derivatives are introduced briefly. We also give an outline of the popular term structure models appearing in the literature.

In Chapter 3, we begin with a review of Markov-functional models with a Gaussian driving process introduced by Hunt, Kennedy and Pelsser [37]. We introduce an algorithm for implementation of Markov-functional models with a Gaussian driver. We then develop a stochastic volatility Markov-functional model. Kaisajuntti and Kennedy [44] identified a two-dimensional SABR type model as an appropriate choice for the level of rates by investigating market data. This motivates us to choose a SABR type model as the driving process. The two-dimensional algorithm for the specification of a Markov-functional model with a Gaussian driver and additive pre-model discussed in [38] relies heavily on the Gaussian assumption. This algorithm can not be used for the stochastic volatility Markov-functional model since we do not have explicit knowledge of the transition density function of the SABR driving process. To implement the model numerically, we propose an algorithm which works with conditional moments of the driver distribution based on an approximation introduced by Kennedy et al. [47]. By working with basis functions and conditional moments of the driving Markov process, we calculate and store three (conditional) expectations which can be seen as building blocks. The model can be implemented based on these building blocks. The algorithm we develop is not specific to a Gaussian driving process and could be modified to apply to all one- and multi-dimensional Markov-functional models.

In Chapter 4, we study the calibration of the stochastic volatility LIBOR Markov-functional model developed in Chapter 3. As we discussed earlier, the specification of a driving process is separated from the marginal distribution of LIBORs at their setting dates in the associated forward measure, which can be seen as an unusual feature for interest rate models. Thus the calibration issue for the stochastic volatility LIBOR Markov-functional model involves choosing the parameters for the SABR driving process and the marginals of LIBORs implied by

the input prices of the set of digital caplets which will be fed into the model.

The separation of a driver and marginals provides more flexibility for a Markov-functional model. But this may also cause an issue. A mismatch of a driving process and marginals could potentially lead to nontransparent dynamics of forward LIBORs and could result in an unstable evolution of the implied volatility surface. This potential issue has been pointed out by Andersen and Piterbarg [5] who argue that a non-parametric formulation of the marginal distribution for LIBORs may result in unrealistic evolution of the volatility smile through time. On the other hand, Bennett and Kennedy [7] showed that a LIBOR Markov-functional model with a Gaussian driver together with the Black's formula for (digital) caplets is numerically similar to the one-factor separable LIBOR market model. Gogala and Kennedy [29] extended the above link to a more general local-volatility case. Based on this link, the authors propose an approach for choosing an appropriate combination of a driving process and (digital) caplet prices, and such a combination leads to desirable dynamics of future implied volatilities. In Chapter 4 we consider a separable SABR-LIBOR market model and expect that it is similar to the stochastic volatility Markov-functional model with a SABR driver together with a SABR marginals. Based on this link the intuition behind the SDEs of the separable SABR-LIBOR market model can be applied to the corresponding stochastic volatility LIBOR Markov-functional model. This gives us an appropriate combination of the driver and marginals. Based on this link we develop a calibration routine to feed in SABR marginals by calibrating to market prices of caplets or swaptions. Moreover the parameters for the SABR driving process can be chosen by calibrating to the market implied correlations. A numerical investigation of the calibration performance is also given.

In Chapter 5, we compare the stochastic volatility Markov-functional model developed in Chapter 3 with one-dimensional swap Markov-functional models in terms of pricing and hedging a Bermudan swaption and a new Bermudan product, which has similar features but simpler payouts than callable range accruals. We consider different combinations of the specifications of driving process and marginals. By comparing their Bermudan prices and vega profiles, we investigate impacts of smiles, correlation structure and stochastic volatility on Bermudan products.

The numerical results show that the mean reversion parameter of the mean reversion and Hull-White drivers, which determines the auto-correlations of the driver, has a large effect on prices of Bermudan products. By comparing the Hull-White Markov-functional model together with log-Normal marginals to the local volatility Markov-functional model with Hull-White driver together with SABR

marginals, it turns out that the smile impact on the price of Bermudan products is very small. In order to study the impact of stochastic volatility on the prices of Bermudan products, we compare the one step covariance swap (non-stochastic-volatility) Markov-functional model with the stochastic volatility swap Markov-functional model. The results show that the introduction of stochastic volatility has a small influence on a Bermudan swaption but has a significant impact on the new Bermudan product. This suggests using a stochastic volatility model for pricing the new Bermudan product. No other papers we have found have identified a product for which stochastic volatility should be added to a term structure model.

The vega profiles of Bermudan products indicate a fundamental difference between the mean reversion driver, which is "parameterized by expiry", and the other driving processes "parameterized by time". By "parametrization by expiry" we mean that the auto-correlations of the driver are fully determined by input parameters. This implies that once input parameters have been fixed, any change in market implied volatility has no effect on the auto-correlations of the driver and therefore the swap rates at their setting dates, which is inconsistent with what we observed in the market. In contrast, for parametrization by time, the correlations of the driver are sensitive to market implied volatilities. Thus any change in market implied volatilities will result in a change in the correlations of the driver. This leads to a fundamental difference in the vega profiles between parametrizations by expiry and by time. For non-stochastic volatility models this behaviour is well-known to practitioners and is analysed in [48] for one dimensional Markov-functional models with the Black's formula. We find by introducing stochastic volatility, the row sum of vegas is decreased in comparison to the one step covariance Markov-functional model but vega profiles of the two models are still very similar. This means that the introduction of stochastic volatility does not materially alter the hedging behaviour. This finding is significant for practitioners wanting to use stochastic volatility models.

Chapter 6 compares Quasi-Gaussian models and Markov-functional models in terms of specification and calibration. These classes of models are two of the most popular low-dimensional term structure models in practical use but there are no studies in the literature which compare them theoretically or numerically. A Quasi-Gaussian model can be seen as a separable Heath-Jarrow-Morton model while a Markov-functional model with a Gaussian driver and log-Normal marginals is found to be numerically similar to the separable LIBOR market model. In order to gain insight into the relationship between these models we study a Quasi-Gaussian model in the Markov functional model framework. By doing so we can see that the

essential difference between the non-stochastic volatility versions of the models lies in the copula of the driving processes and that the Markov-functional framework offers much flexibility in matching the choice of copula for the driver to reality. These two models both allow for stochastic volatility versions and we also consider the relationship between these versions.

Chapter 2

Interest rate markets and models

In this chapter we provide an overview of the interest rate markets and models. We will introduce notation and set up a foundation for the thesis. In Section 2.1 we introduce the main interest rate products. In Section 2.2 we give a brief outline of the popular term structure models. We will assume that the reader has knowledge of pricing via an equivalent martingale measure and application of the fundamental pricing formula for various choices of numeraires. This background can be found in [38] and [46].

2.1 Interest rate markets

In this section we introduce briefly the interest rate products in the market. We first introduce some basic instruments that are liquid in the interest rate markets. We then proceed to some common interest rate options which are relevant to the thesis. The material in this section is from [38].

2.1.1 Basic instruments

In this subsection we will introduce zero-coupon bonds (ZCBs), forward rate agreements (FRAs) and interest rate swaps. Throughout this thesis we assume that the tenor structure is given by

$$0 = T_0 < T_1 < ... < T_{n+1} \tag{2.1}$$

where $\alpha_i = T_{i+1} - T_i$, $i = 0, ..., n$, are the accrual factors.

We start with underlying assets in the interest rate markets - zero coupon bonds. A zero-coupon bond with maturity time T is a contract that guarantees its holder to be paid one unit of currency at time T. The contract value at time $t \leq T$ is denoted by D_{tT}. The dependence of D_{tT} on the maturity date T is known as the term structure of ZCBs at time t.

LIBOR stands for London Interbank Offered Rate. We can define a forward LIBOR through a forward rate agreement. A forward rate agreement is an agreement between two counterparties to exchange cash payments at some specified date in the future. At the maturity T_{i+1}, $i = 1, ..., n$, a fixed payment $N\alpha_i K$, where K is the fixed rate and N is notional amount, is exchanged against a floating payment $N\alpha_i L^i_{T_i}$, where $L^i_{T_i}$ is the spot LIBOR with expiration date T_i and maturity date T_{i+1} defined by

$$L^i_{T_i} = \frac{1 - D_{T_i T_{i+1}}}{\alpha_i D_{T_i T_{i+1}}}.$$

Following a replicating portfolio argument, we have that the time-t value of this FRA is given by

$$V_t = N D_{tT_i} - N(1 + \alpha_i K) D_{tT_{i+1}}.$$

The forward LIBOR L^i_t, at time t, is defined as the fixed rate K such that the time-t value of the FRA is zero. By letting $V_t = 0$, the forward LIBOR L^i_t at time t that expires at T_i and matures at T_{i+1} is given by

$$L^i_t = \frac{D_{tT_i} - D_{tT_{i+1}}}{\alpha_i D_{tT_{i+1}}}, \qquad t \leq T_i, \tag{2.2}$$

for $i = 1, ..., n$.

Remark 1. *The value of the forward LIBOR L^i_0, $i = 1, ..., n$, seen today is assumed to be given in the market. But the value of the forward LIBOR L^i_t, $i = 1, ..., n$, at some future time t is a random variable. Later we will introduce interest rate models for the forward LIBOR processes L^i for $i = 1, ..., n$. The same remark also applies to the forward swap rate below.*

An interest rate swap, or swap for short, is an agreement between two counterparties to exchange a series of cashflows on pre-agreed dates in the future. There are two kinds of swaps: payers swap and receivers swap. Let us consider a payers interest rate swap with strike K, exercise dates $T_i, T_{i+1}, ..., T_{i+j-1}$ and payment dates $T_{i+1}, T_{i+2}, ..., T_{i+j}$. The fixed rate payer pays the fixed leg $N\alpha_s K$ in return for the

floating leg $N\alpha_s L^s_{T_s}$ at each payment date for $s = i, ..., i+j-1$. A receivers interest rate swap has a reversed cashflows. The time-t value of the above fixed leg is given by

$$V^{i,j}_{FXD}(t) = NKP^{i,j}_t, \tag{2.3}$$

where

$$P^{i,j}_t := \sum_{k=i}^{i+j-1} \alpha_k D_{tT_{k+1}}.$$

The expression $P^{i,j}$ is referred to as the *present value of a basis point* or PVBP for short. Following a replicating portfolio argument, the time-t value of the above floating leg is given by

$$V^{i,j}_{FLT}(t) = N(D_{tT_i} - D_{tT_{i+j}}).$$

It follows that the time-t value of the above interest rate swap is given by

$$\begin{aligned} V^{i,j}_{Swap}(t) &= \tau[V^{i,j}_{FLT}(t) - V^{i,j}_{FXD}(t)] \\ &= \tau N[D_{tT_i} - D_{tT_{i+j}} - KP^{i,j}_t], \end{aligned} \tag{2.4}$$

where $\tau = -1$ leads to a receivers swap and $\tau = 1$ leads to a payers swap. The forward swap rate $y^{i,j}_t$ is defined as the value of K such that the time-t value $V^{i,j}_{Swap}(t)$ of the swap is zero. Thus we have that

$$y^{i,j}_t = \frac{D_{tT_i} - D_{tT_{i+j}}}{P^{i,j}_t}, \qquad t \leq T_i. \tag{2.5}$$

If we substitute equation (2.5) back into (2.4), we have the more usual expression for the value of an interest rate swap:

$$V^{i,j}_{Swap}(t) = \tau N[P^{i,j}_t(y^{i,j}_t - K)].$$

2.1.2 Interest rate options

In this subsection we introduce some interest rate options that are relevant to the thesis. These options are caplets (floorlets), vanilla swaptions[1], digital caplets (floorlets) and PVBP-digital swaptions.

[1] Also known as European swaptions

A caplet and floorlet can be seen as an option on an FRA with the payoffs

$$V^i_{caplet}(T_{i+1};K) = N\alpha_i \max(L^i_{T_i} - K, 0)$$

and

$$V^i_{floorlet}(T_{i+1};K) = N\alpha_i \max(K - L^i_{T_i}, 0)$$

respectively at time T_{i+1} where N is the notional and K is the strike. In order to price such an option in a complete arbitrage-free model, we apply the fundamental pricing formula. In particular given some equivalent martingale measure $\mathbb{M}$ corresponding to the numeraire M, today's value of the caplet with strike K is given by

$$V^i_{caplet}(0;K) = M_0 \mathbb{E}_{\mathbb{M}}[\frac{N\alpha_i(L^i_{T_i} - K)^+}{M_{T_{i+1}}}].$$

In order to calculate this price we have to choose a numeraire M and a model for the LIBOR L^i in the measure $\mathbb{M}$. Suppose we work with the forward measure $\mathbb{F}^{i+1}$ associated with the numeraire $D_{.,T_{i+1}}$. It follows from equation (2.2) that the LIBOR L^i is a martingale in the forward measure $\mathbb{F}^{i+1}$. Suppose we model L^i by a driftless log-Normal process under the forward measure $\mathbb{F}^{i+1}$

$$dL^i_t = \sigma^i_t L^i_t dW^{i+1}_t,$$

for some deterministic function σ^i and a Brownian motion W^{i+1} under the forward measure. This yields the following well-known Black's formula introduced by Black [9]

$$V^i_{caplet}(0;K) = D_{0T_{i+1}} \mathbb{E}_{\mathbb{F}^{i+1}}[N\alpha_i(L^i_{T_i} - K)^+] \tag{2.6}$$

$$= \alpha_i N D_{0T_{i+1}}(L^i_0 \Phi(d_1) - K\Phi(d_2)) \tag{2.7}$$

where

$$d_1 = \frac{ln(L^i_0/K)}{\tilde{\sigma}\sqrt{T_i}} + \frac{1}{2}\tilde{\sigma}\sqrt{T_i}, \tag{2.8}$$

$$d_2 = d_1 - \tilde{\sigma}\sqrt{T_i} \tag{2.9}$$

$$\tilde{\sigma}^2 = \frac{1}{T_i}\int_0^{T_i} (\sigma^i_u)^2 du$$

and $\Phi(\cdot)$ is the cumulative normal distribution function. Similarly, we have the following Black's formula for a floorlet with strike K:

$$V^i_{floorlet}(0;K) = \alpha_i N D_{0T_{i+1}}(K\Phi(-d_2) - L^i_0\Phi(-d_1)).$$

We now define the implied volatility of a caplet. Given the market price $\tilde{V}^i_{caplet}(0;K)$ of the caplet struck at the strike K, the implied volatility of this caplet is defined as the volatility $\tilde{\sigma}$ such that

$$\tilde{V}^i_{caplet}(0;K) = \alpha_i N D_{0T_{i+1}}(L^i_0\Phi(d_1) - K\Phi(d_2))$$

where d_1 and d_2 are given by equations (2.8) and (2.9) respectively. In financial markets the market prices of vanilla options are quoted in terms of implied volatilities. In the interest rate markets implied volatilities are commonly observed to represent a shape of skew or smile as a function of strike. However the above log-Normal assumption of L^i implies that the implied volatilities show a flat line with respect to strike i.e. the function $\tilde{\sigma}(K)$ is always a constant. In order to capture implied volatility skews or smiles, a number of models have been developed. We will return to this topic in the later chapters.

We now consider vanilla swaptions. A vanilla swaption is an option on an interest rate swap. It gives its holder the right, without any obligation, to enter into an interest rate swap. According to the underlying interest rate swap there are two types of swaptions: receivers swaption and payers swaption. For the receivers swaptions, upon exercise the option holder enters a swap in which he receives a fixed strike rate K and pays the floating rate; a payers is the reverse. So the payoff of a payers swaption with strike K on an interest rate swap with expiry dates $T_i, T_{i+1}, ..., T_{i+j-1}$ and maturity dates $T_{i+1}, T_{i+2}, ..., T_{i+j}$ is given by

$$V^{i,j}_{sption}(T_i;K) = \max\{NP^{i,j}_{T_i}(y^{i,j}_{T_i} - K), 0\},$$

where N is the notional and K is the strike.

Following a similar explanation and assuming a driftless log-Normal process for $y^{i,j}$ under the swaption measure $\mathbb{S}^{i,j}$ associated with the numeraire $P^{i,j}$, we can obtain the following Black's formula for a vanilla swaption

$$\begin{aligned} V^{i,j}_{sption}(0;K) &= P^{i,j}_0\mathbb{E}_{\mathbb{S}^{i,j}}[N(y^{i,j}_{T_i} - K)^+] \\ &= NP^{i,j}_0(y^{i,j}_0\Phi(d_1) - K\Phi(d_2)) \end{aligned} \tag{2.10}$$

where

$$
\begin{aligned}
d_1 &= \frac{ln(y_0^{i,j}/K)}{\tilde{\sigma}\sqrt{T_i}} + \frac{1}{2}\tilde{\sigma}\sqrt{T_i}, \\
d_2 &= d_1 - \tilde{\sigma}\sqrt{T_i}, \\
\tilde{\sigma}^2 &= \frac{1}{T_i}\int_0^{T_i} (\sigma_u^i)^2 du.
\end{aligned} \tag{2.11}
$$

The explanation for implied volatilities of caplets also applies to vanilla swaptions.

The caplets (floorlets) and vanilla swaptions we introduced here are liquidly traded in the interest rate markets. These options are referred to as vanilla options. The prices of vanilla options are observable in the market. In practice a valuation model for some complex derivatives needs to be calibrated to their relevant underlying vanilla options.

In what follows, we introduce some less liquid interest rate options traded in the market - digital options. In particular, we introduce two examples that are most relevant to the thesis: digital caplets (floorlets) and PVBP-digital swaptions.

A digital caplet on the LIBOR L^i is an option paying a unit amount at time T_{i+1} if the LIBOR $L^i_{T_i}$ is above some strike level K. The payoff at time T_{i+1} is given by

$$V^i_{digcap}(T_{i+1}; K) = N\mathbf{I}_{\{L^i_{T_i} > K\}},$$

where $\mathbf{I}$ is indicator function and N is the notional. Taking the T_{i+1}-maturity ZCB $D_{\cdot,T_{i+1}}$ as numeraire and using the same log-Normal model, as we did for caplet, yields

$$V^i_{digcap}(0; K) = N D_{0T_{i+1}}\Phi(d_2),$$

where d_2 is as defined in (2.9).

Remark 2. *The payoffs of a caplet and digital caplet are related via the following equation*

$$\frac{d}{dK}(x - K)^+ = -\mathbf{I}_{\{x>K\}}.$$

In particular differentiate both sides of (2.6) with respect to the strike K, and we have that

$$\frac{dV^i_{caplet}(0; K)}{dK} = -\alpha_i V^i_{digcap}(0; K).$$

The implication of the relationship is that knowing the prices of caplet $V^i_{caplet}(0;K)$ for all strike K is equivalent to knowing the prices of digital caplet $V^i_{digcap}(0;K)$ for all K. This remark also applies to vanilla swaptions and the corresponding PVBP-digital swaptions below.

Remark 3. *Given the prices $V^i_{digcap}(0;K)$ of a digital caplet as a function of the strike $K \geq 0$, from the fundamental pricing formula we have that*

$$\begin{aligned} V^i_{digcap}(0;K) &= D_{0,T_{i+1}} \mathbb{E}_{\mathbb{F}^{i+1}}[N\mathbf{I}_{\{L^i_{T_i}>K\}}] \\ &= D_{0,T_{i+1}} N\mathbb{F}^{i+1}[L^i_{T_i} > K], \end{aligned}$$

under the forward measure $\mathbb{F}^{i+1}$ associated with the numeraire $D_{\cdot,T_{i+1}}$. Therefore the prices $V^i_{digcap}(0;K)$ implies the distribution of the LIBOR $L^i_{T_i}$ in the associated forward measure.

From the above two remarks one can see that the prices $V^i_{caplet}(0;K)$ can also determine the distribution of the LIBOR $L^i_{T_i}$ in the associated forward measure $\mathbb{F}^{i+1}$. This also applies to a vanilla swaption and PVBP-digital swaption where the distribution of the swap rate at its setting date can be determined in the associated swaption measure.

A PVBP-digital swaption on the swap rate $y^{i,j}$ with strike K has the following payoff at time T_i

$$V^{i,j}_{digsption}(T_i;K) = NP^{i,j}_{T_i}\mathbf{I}_{\{y^{i,j}_{T_i}>K\}},$$

where N is the notional. By assuming a driftless log-Normal process for $y^{i,j}$ under the swaption measure $\mathbb{S}^{i,j}$ associated with the numeraire $P^{i,j}$, we can also obtain the following Black's formula for a vanilla swaption

$$\begin{aligned} V^{i,j}_{digsption}(0;K) &= P^{i,j}_0 \mathbb{E}_{\mathbb{S}^{i,j}}[N\mathbf{I}_{\{y^{i,j}_{T_i}>K\}}] \\ &= NP^{i,j}_0 \Phi(d_2), \end{aligned} \tag{2.12}$$

where d_2 is as defined in (2.11).

2.2 Interest rate models

In this section we first discuss the role of valuation models. We then give a brief overview of interest rate models.

2.2.1 The role of models

In this thesis we use interest rate models to price and hedge exotic interest rate derivatives. By an exotic derivative we mean one which is not vanilla. Exotic options are commonly traded over-the-counter so that one needs to find their prices based on interest rate models. To do so a valuation model needs to be calibrated to other underlying instruments and market information, such as vanilla options, which are relevant to the exotic option under consideration. By calibration, we mean that the valuation model can reproduce the market prices of the chosen underlying instruments. In this sense valuation models are served as a sophisticated extrapolation from underlying instruments to produce a model price of an exotic option.

2.2.2 Short rate models

Short rate models are the earliest term structure models. The short rate r_t at time t is defined by

$$r_t := -\frac{\partial ln D_{t,T}}{\partial T}|_{T=t}.$$

The short rate r is a hypothetical interest rate which is not observable in the market. Short rate models are specified by describing the evolution of the short rate

$$dr_t = \mu(r_t, t)dt + \sigma(r_t, t)dW_t^Q, \tag{2.13}$$

where W^Q is a Brownian motion in the risk-neutral measure $\mathbb{Q}$ associated with the numeraire the bank account B which satisfies

$$dB_t = r_t B_t dt.$$

The drift and diffusion functions μ and σ need to be chosen carefully to give the model particular behaviour. Some common examples in the literature are given in Table 2.1. The parameters of the short rate model (2.13) are chosen by calibrating the model to the initial term structure and the market prices of vanilla options.

The time-t value D_{tT} of a ZCB with maturity T is given by

$$D_{tT} = \mathbb{E}_{\mathbb{Q}}[e^{-\int_t^T r_s ds}|\mathcal{F}_t], \tag{2.14}$$

where $\{\mathcal{F}_t\}$ is the natural filtration generated by the Brownian motion W^Q. From equation (2.14), the prices of all ZCBs, which are underlying assets in the interest

Model	Specification
Merton [53]	$dr_t = adt + \sigma dW_t^Q$
Vasicek [64]	$dr_t = k(\theta - r_t)dt + \sigma dW_t^Q$
Dothan [20]	$dr_t = ar_t dt + \sigma r_t dW_t^Q$
CIR [17]	$dr_t = k(\theta - r_t)dt + \sigma\sqrt{r_t}dW_t^Q$
Hull-White [36]	$dr_t = k(\theta_t - r_t)dt + \sigma dW_t^Q$

Table 2.1: Examples of short rate models

rate markets, can be obtained once we have specified the dynamics of the short rate r in the risk-neutral measure. Consequently the term structure of ZCBs are specified via the dynamics of the Markov process r. This allows one to price a derivative using an efficient algorithm such as numerical integration or finite-difference method. For details, the reader is referred to [5]. Since the short rate is not observable in the market, it may result in difficulty of calibrating to the initial term structure and the market vanilla options.

2.2.3 Heath-Jarrow-Morton models

In a Heath-Jarrow-Morton (HJM) framework [30], the instantaneous forward rate $f(t,T)$ satisfies the following SDE

$$df(t,T) = \sigma(t,T) \cdot (\int_t^T \sigma(t,u)du)dt + \sigma(t,T) \cdot dW_t^Q, \tag{2.15}$$

where W^Q is an d-dimensional Brownian motion in the risk-neutral measure $\mathbb{Q}$ associated with the numeraire the bank account B. The instantaneous forward rate $f(.,T)$ and the short rate r which can be viewed as an instantaneous spot rate are related via the following equation

$$f(t,t) = r_t.$$

A short rate model models the dynamics of the short rate r while a HJM model specifies the dynamics of the term structure of instantaneous forward rates $f(.,T)$ for all maturities T.

The time-t value D_{tT} of a ZCB with maturity T is given by

$$D_{tT} = e^{-\int_t^T f(t,u)du}.$$

From (2.15), one can see that the instantaneous forward rate is specified via the volatility function $\sigma(t,T)$. With a specific choice for the volatility $\sigma(t,T)$, the HJM

model may lead to a known short rate model. We will return to this topic later in Chapter 6.

2.2.4 Market models

Market models can be divided into two versions: LIBOR market models (LMMs) and swap market models (SMMs). In a LMM the dynamics of a set of contiguous forward LIBORs L^i, $i = 1, ..., n$, are specified. LMMs were first introduced by Brace et al. [10], where the dynamics of forward LIBORs L^i, $i = 1, ..., n$, is given by

$$\begin{aligned} dL_t^i &= -\sum_{j=i+1}^{n} \left(\frac{\sigma_t^j \alpha_j L_t^j}{1+\alpha_j L_t^j}\right)\sigma_t^i \rho_{ij} L_t^i dt + \sigma_t^i L_t^i dW_i^{n+1}(t), \quad i = 1, ..., n-1 \\ dL_t^n &= \sigma_t^n L_t^n dW_n^{n+1}(t). \end{aligned} \tag{2.16}$$

where W^{n+1} is a (correlated) Brownian motion with $dW_i^{n+1}(t)dW_j^{n+1}(t) = \rho_{ij}dt$ under the terminal measure $\mathbb{F}^{n+1}$ associated with the numeraire the T_{n+1}-maturity ZCB $D_{.,Tn+1}$, and σ_t^i is a deterministic function. For more details about the specification of the volatility function σ_t^i, the reader is referred to [11]. The drift term in the SDE (2.16) is determined by maintaining the arbitrage-free property of the model.

Following the change of measure technique, we obtain the dynamics of L^i under the forward measure $\mathbb{F}^{i+1}$ corresponding to the numeraire $D_{.,Ti+1}$ which is given by

$$dL_t^i = \sigma_t^i L_t^i dW_i^{i+1}(t), \tag{2.17}$$

where W^{i+1} is a (correlated) Brownian motion under the forward measure $\mathbb{F}^{i+1}$. This is the case since we can see from the definition of LIBORs that L^i is a martingale in the measure $\mathbb{F}^{i+1}$ so that L^i should be a driftless process. Note that the SDE (2.17) yields the Black's formula (2.7) for the prices of caplets. Thus LMMs allow for calibration to the Black's formula for caplets and floorlets, which is an advantage for LMMs.

The swap market model is specified by describing the dynamics of a set of contiguous forward (co-terminal) swap rates under the terminal measure $\mathbb{F}^{n+1}$

$$\begin{aligned} dy_t^{i,n+1-i} &= \mu_t^i(y_t) y_t^{i,n+1-i} dt + \sigma_t^i y_t^{i,n+1-i} dW_i^{n+1}(t) \quad i = 1, ..., n-1, \\ dy_t^{n,1} &= \sigma_t^n y_t^{n,1} dW_n^{n+1}(t), \end{aligned} \tag{2.18}$$

where

$$\mu_t^i(y_t) = -\sum_{j=i+1}^{n} \frac{\Psi_t^{j-1}\hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1}\hat{P}_t^{i,n+1-i}}\left(\frac{\alpha_{j-1}y_t^{j,n+1-j}}{1+\alpha_{j-1}y_t^{j,n+1-j}}\right)\sigma_t^i\sigma_t^j\rho_{ij}$$

$$\Psi_t^i := \prod_{j=1}^{i}(1+\alpha_j y_t^{j+1,n-j})$$

$$\hat{P}_t^{i,n+1-i} := \frac{P_t^{i,n+1-i}}{D_{t,T_{n+1}}}.$$

By moving to the swaption measure $\mathbb{S}^{i,n+1-i}$ associated with the numeraire $P^{i,n+1-i}$, the forward swap rate $y^{i,n+1-i}$ is given by the following driftless process

$$dy_t^{i,n+1-i} = \sigma_t^i y_t^{i,n+1-i} dW_i^{i,n+1-i}(t),$$

where $W^{i,n+1-i}$ is a (correlated) Brownian motion under the measure $\mathbb{S}^{i,n+1-i}$. This leads to the Black's formula (2.10) for the prices of vanilla swaptions.

We can see that the drift terms of the SDEs (2.16) and (2.18) are dependent on forward rates. Consequently a market model, though it is Markovian with respect to all its forward rates, is a high-dimensional model even when driven by only one Brownian motion. In practice a simulation, such as Monte Carlo methods, is required for an accurate implementation which is computationally expensive. For more details about the specification, calibration, implementation and applications of market models, the reader is referred to [58] and [59].

2.2.5 Markov-functional models

Markov-functional models (MFMs) were introduced by Hunt et al. [37]. MFMs can fit any arbitrage-free formula for caplet or swaption prices which includes the Black's formula. A MFM is specified via the SDE for a Markov process under some equivalent martingale measure which is referred to as the driving process or the driver for short. The driving process can be viewed as modelling the overall level of interest rates in the economy. The defining characteristic of MFMs is that the underlying assets in the interest rate market - ZCBs prices - are a function of a Markov process so that the term structure of ZCBs can be specified by describing the dynamics of the Markov process and the functional forms. Recall that in a short rate model, the prices of ZCBs can be computed via the formula (2.14). In a MFM the functional forms of ZCBs can be obtained numerically by calibrating to vanilla options prices. The use of a low-dimensional Markov process makes it possible to

implement the model efficiently because one only needs to track the low-dimensional Markov process. We will return to this topic later in the next chapter.

Let us conclude this section with a remark. Short rate models and market models allow for stochastic volatility extensions. See for example [24] and [62]. However there is no research in the literature about stochastic volatility extensions of MFMs. The purpose of this thesis is to study stochastic volatility interest rate models, and one main contribution we make is the development, calibration and implementation of a stochastic volatility MFM.

Chapter 3

Stochastic volatility Markov-functional models

3.1 Introduction

In this chapter we consider Markov-functional models introduced by Hunt, Kennedy and Pelsser [37]. Markov-functional models are interest rate models that allow for calibrating to any arbitrage-free formula for caplet or swaption prices. The defining characteristic of Markov-functional models is that the underlying assets - Zero Coupon Bonds - are a function of a Markov process so that the term structure of zero coupon bonds can be specified by describing the dynamics of the Markov process and the functional form. If the Markov process is chosen to be low dimensional, it is possible to implement the model efficiently because one only needs to track the low-dimensional Markov process. The functional form can be determined to fit the prices of a set of caplets or swaptions by numerical integration. Hunt and Kennedy [38] applied a low-dimensional Gaussian process as the Markov process and calibrated the Markov-functional model to the Black's formula for caplets or swaptions. They also proposed an efficient algorithm to specify the Markov-functional model. Caspers [13] considered the numerical implementation of a one-dimensional Markov-functional model with a Gaussian driver. Kaisajuntti and Kennedy [45] studied the case when the Markov process is N-dimensional. Gogala and Kennedy [29] developed a one-dimensional Markov-functional model driven by a non-Gaussian Markov process and proposed an efficient algorithm to implement the model numerically. Fries and Eckstaedt [27] studied hybrid Markov-functional models; see also [26] and [28].

In Markov-functional models the functional forms of zero coupon bonds can be obtained by numerical integration so that SDEs of forward rates or zero coupon

bonds are not given explicitly. This makes Markov-functional models less transparent in terms of the dynamics of forward rates. Bennett and Kennedy [7] partially tackled this problem. The authors compared a one-factor separable LIBOR market model with a one-dimensional Markov-functional model driven by a Gaussian process together with a log-Normal marginal distributions of LIBORs at their setting dates. The numerical results showed that under a wide range of market conditions these two models have similar dynamics and Bermudan swaption prices for short maturities (10 years). For long maturities and high volatilities the similarity begins to break down. This link gives us more understanding and intuition of a Markov-functional model from the corresponding one-factor separable LIBOR market model. Gogala and Kennedy [29] extended the above link by concluding that the one-factor separable local volatility LIBOR market model has similar dynamics to the Markov-functional model with the same local volatility type driver and pricing formula for caplets.

Over the last two decades many stochastic volatility term structure models have been proposed e.g. [4], [65], [55], [62], [61], [60], [34] and [33]. The introduction of stochastic volatility is motivated by the empirical evidence (see [4]) that the volatilities in interest rate option markets possess a random component and Hagan et al's [31] argument that the smile dynamics in stochastic volatility models behave more realistically than local volatiliy models. The stochastic volatility interest rate models mentioned above are all high-dimensional models and need to be implemented by simulation. Thus in practice it is infeasible, from the point of view of banks, to use such models for pricing and especially hedging exotic derivatives.

Andersen and Piterbarg [5] developed a stochastic volatility Quasi-Gaussian model which is a low-dimensional stochastic volatility term structure model. In this chapter we will develop another low-dimensional stochastic volatility interest rate model - the stochastic volatility Markov-functional model. As Kaisajuntti and Kennedy [44] pointed out, a one-factor stochastic process is not enough to capture the overall level of interest rates. The authors identified a SABR type model as an appropriate choice for the level of rates by investigating market data. This finding motivates us to develop a stochastic volatility Markov-functional model by taking a SABR model as the driving process to capture the overall level of interest rates and drive the whole economy. The main challenge for the specification of the model is that it is hard to obtain the transition density function of the driver. To implement the model numerically, we propose an algorithm which works with conditional moments of the driver distribution based on an approximation introduced by Kennedy et al. [47]. Unlike the two-dimensional algorithm that appears in [38]

for the additive pre-model, which relies heavily on the Gaussian assumption, the algorithm we developed is not specific and could be modified to apply to all one- and multi-dimensional Markov-functional models.

The chapter is organized as follows. In Section 3.2 we revisit the LIBOR and Swap Markov-functional models with a Gaussian driving process and introduce the numeraire approach to specify the model under the terminal measure. In Section 3.3 we develop a stochastic volatility Markov-functional model and discuss the numerical implementation by developing a general algorithm which works with basis functions and conditional moments of the driving Markov process. Moreover the specification of the driving process as well as the pre-model for the stochastic volatility Markov-functional model is given. We conclude in Section 3.4.

3.2 Markov-functional models

In this section we review Markov-functional models (MFMs) proposed by Hunt, Kennedy and Pelsser [37]. First we give a definition of a MFM from Hunt and Kennedy [38] (see also [29] for a generalized definition). We will know some general properties of a MFM from the following definition but it will not tell us the specification of a MFM in practice. We will introduce an algorithm to specify a MFM later in the section.

Definition 1. *An interest rate model is said to be Markov-functional if there exists some numeraire pair* $(N, \mathbb{N})$ *and some process* x *such that:*
(P.1) the process x *is a Markov process under the measure* $\mathbb{N}$*.*
(P.2) the zero coupon bond prices are of the form

$$D_{tT} = D_{tT}(x_t) \qquad 0 \leq t \leq T \leq T^*,$$

for a finite time horizon $T^* < \infty$*.*
(P.3) the numeraire N *is of the form*

$$N_t = N_t(x_t) \qquad 0 \leq t \leq T^*.$$

A MFM is said to be of d-dimension if x is a d-dimensional Markov process. We can see from the above definition that the prices of zero coupon bonds (ZCBs), which are the underlying assets of the economy, are a function of a Markov process x, which is referred to as the driving process or driver for short. This allows us to track the driving process x only in order to implement a MFM. If we choose a low-dimensional driving process, the model can be implemented efficiently. Otherwise

it needs to be implemented by simulation. Moreover the existence of a numeraire pair ensures that all the numeraire-rebased ZCBs are martingales in the equivalent martingale measure $\mathbb{N}$. In particular the prices of ZCBs can be obtained by the martingale property

$$D_{tT} = N_t E_{\mathbb{N}}[\frac{1}{N_T}|\mathcal{F}_t],$$

where $\{\mathcal{F}_t\}$ is the natural filtration generated by the Markov process x in the measure $\mathbb{N}$ and $E_{\mathbb{N}}(\cdot)$ denotes an expectation under $\mathbb{N}$. This ensures that a MFM is arbitrage-free. We will present the numeraire approach to specify a MFM under the terminal measure, which is an equivalent martingale measure, later in the section.

The above definition is very general. In fact most interest rate models are included in the category of MFMs according to the above definition. For example in a short rate model the short rate is taken as the Markov process and ZCBs are some function of it. In a LIBOR market model (LMM) however, the evolution of a set of contiguous forward LIBORs with different maturities are specified. The high-dimensional LIBORs process can be seen as the Markov process.

For efficient implementation we commonly use a low-dimensional Markov process with dimension $d \leqslant 3$. A high-dimensional MFM needs to be implemented by simulation such as Monte-Carlo methods which is much slower (see [45] for high-dimensional MFMs). In this thesis we restrict our attention to low-dimensional MFMs. Similar to Market models, there are two common versions of MFMs: LIBOR MFMs and Swap MFMs. The LIBOR version is set up by feeding in digital caplets prices whereas the swap MFM is specified by fitting PVBP-digital swaptions prices. In the next two subsections we will introduce the specification of both versions of MFMs under the terminal measure $\mathbb{F}^{n+1}$ corresponding to the numeraire $D_{.,T_{n+1}}$. MFMs can also be developed under the spot measure and details can be found in [28] and [29].

3.2.1 One-dimensional Markov-functional models: LIBOR version

Before discussing multi-dimensional MFMs, it will be helpful to review the specification of a one-dimensional LIBOR MFM under the terminal measure using the numeraire approach. We will specify the model on a grid, which is sufficient for most applications in practice and we can see it later when we use a MFM to price and hedge Bermudan swaptions. In what follows we first specify the driving Markov process x and then we determine functional forms of the numeraire $D_{T_iT_{n+1}}(x_{T_i})$ by calibrating to the input prices of a set of digital caplets.

Suppose we are given the following Gaussian process x:

$$dx_t = \sigma_t dW_t^{n+1}, \tag{3.1}$$

where W^{n+1} is a one-dimensional Brownian motion in the terminal measure $\mathbb{F}^{n+1}$ and σ_t is a deterministic function. Note that theoretically any diffusion process can be used as a driver. We choose a Gaussian process as the driving Markov process here for efficient implementation. More details about one-dimensional MFMs driven by non-Gaussian Markov processes can be found in [29].

Having specified the Markov process x, we now consider the problem of how to find the functional forms $D_{T_iT_{n+1}}(x_{T_i})$ for $i = 1, ..., n$. We proceed by backwards induction on time T_i. At time T_{n+1} by definition we have that

$$D_{T_{n+1}T_{n+1}}(x_{T_{n+1}}) = 1.$$

Assume that we have determined $D_{T_kT_{n+1}}(x_{T_k})$ for $k = i+1, ..., n+1$. At time T_i, by definition, the LIBOR $L^i_{T_i}$ is given by

$$L^i_{T_i} = \frac{1 - D_{T_iT_{i+1}}}{\alpha_i D_{T_iT_{i+1}}}. \tag{3.2}$$

It follows from equation (3.2) that the numeraire at time T_i can be expressed as

$$D_{T_iT_{n+1}}(x_{T_i}) = \frac{1}{\widehat{D}_{T_iT_{i+1}}(x_{T_i})(1 + \alpha_i L^i_{T_i}(x_{T_i}))}, \tag{3.3}$$

where the numeraire-rebased ZCB $\widehat{D}_{T_iT_{i+1}}$ is defined as

$$\widehat{D}_{T_iT_{i+1}}(x_{T_i}) := \frac{D_{T_iT_{i+1}}(x_{T_i})}{D_{T_iT_{n+1}}(x_{T_i})}.$$

It follows from the martingale property that

$$\begin{aligned} \widehat{D}_{T_iT_{i+1}}(x_{T_i}) &= E_{\mathbb{F}^{n+1}}[\frac{1}{D_{T_{i+1}T_{n+1}}(x_{T_{i+1}})}|\mathcal{F}_{T_i}] \\ &= E_{\mathbb{F}^{n+1}}[\frac{1}{D_{T_{i+1}T_{n+1}}(x_{T_{i+1}})}|x_{T_i}], \end{aligned} \tag{3.4}$$

where the last equation follows from the Markov property and the expectation conditional on x_{T_i} is a short notation for the expectation conditional on the σ-field

$\sigma(x_{T_i})$ generated by x_{T_i}. Since the Markov process

$$x_t = \int_0^t \sigma_u dW_u^{n+1}$$

is a Gaussian process, numerical integration for (conditional) expectation is calculated via

$$E_{\mathbb{F}^{n+1}}[f^j(x_{T_j})|x_{T_i}] = \int_{-\infty}^{\infty} f^j(u)\phi_{x_{T_j}|x_{T_i}}(u)du \tag{3.5}$$

for continuous function $f^j(\cdot)$. The function $\phi_{x_{T_j}|x_{T_i}}(\cdot)$ is the Gaussian density function with mean x_{T_i} and variance $\int_{T_i}^{T_j} \sigma_u^2 du$ for $0 \leq i < j \leq n+1$. Thus by approximating $\frac{1}{D_{T_{i+1}T_{n+1}}(x_{T_{i+1}})}$ in (3.4) by piecewise polynomials, the functional forms $\widehat{D}_{T_iT_{i+1}}(x_{T_i})$ can be obtained by numerical integration (3.5). Having specified the numeraire-rebased ZCBs $\widehat{D}_{T_iT_{i+1}}(x_{T_i})$, we see from equation (3.3) that if we can determine the functional form of $L^i_{T_i}(x_{T_i})$, the functional form of $D_{T_iT_{n+1}}(x_{T_i})$ is immediate.

We now fix the functional form $L^i_{T_i}(x_{T_i})$ by calibrating to the input prices of digital caplets (see Section 2.1.2) corresponding to the ith LIBOR L^i. In the market the prices $V_0^i(K)$ of digital caplets on the ith LIBOR L^i are available for only a few strikes K. However in practice a continuous function $V_0^i(K)$ w.r.t strike K can be achieved by interpolation or using e.g. the SABR model to generate the whole implied volatility smile. The technical details will be covered in the next chapter. For now we just assume that the input prices $V_0^i : [0, \infty] \to \mathbb{R}$ of digital caplets at strikes $K \geq 0$ are given for $i = 1, ..., n$. Following the fundamental pricing formula in $\mathbb{F}^{n+1}$ we have that

$$\begin{aligned} V_0^i(K) &= D_{0T_{n+1}} E_{\mathbb{F}^{n+1}}[\frac{1}{D_{T_{i+1}T_{n+1}}} \mathbf{I}_{\{L^i_{T_i} > K\}}] \\ &= D_{0T_{n+1}} E_{\mathbb{F}^{n+1}}[\widehat{D}_{T_iT_{i+1}} \mathbf{I}_{\{L^i_{T_i} > K\}}], \end{aligned} \tag{3.6}$$

where the last equation is obtained by the tower property and equation (3.4). So the input prices $V_0^i(K)$ are a decreasing function of the strike K.

In order to find the functional forms of LIBORs we still need one more assumption:

(A.1) In the model, $L^i_{T_i}$ is a monotonic increasing function of x_{T_i} for $i = 1, ..., n$.

This assumption is not restrictive since the Markov process represents the level of LIBORs.

Now we determine the functional form $L^i_{T_i}(x_{T_i})$. We choose a value x^* of x_{T_i}

and for each x^* we evaluate

$$J_0^i(x^*) = D_{0T_{n+1}} E_{\mathbb{F}^{n+1}}[\widehat{D}_{T_iT_{i+1}}(x_{T_i})\mathbf{I}_{\{x_{T_i}>x^*\}}], \tag{3.7}$$

which can be computed by means of numerical integration. Compare equation (3.7) with (3.6), and it turns out that functions J_0^i and V_0^i are the same except for the indicator function. By assumption (A.1) for any value x^*, we can always find a unique strike K such that

$$\{x_{T_i} > x^*\} = \{L_{T_i}^i > K\}. \tag{3.8}$$

Thus for any value of $J_0^i(x^*)$, we can always find a unique strike K such that

$$J_0^i(x^*) = V_0^i(K),$$

where the function $K(x^*)$ can be solved as

$$K(x^*) = (V_0^i)^{-1}(J_0^i(x^*)).$$

It follows from (3.8) that

$$L_{T_i}^i(x^*) = K(x^*).$$

We have found the functional form $L_{T_i}^i(x_{T_i})$, and therefore the functional form $D_{T_iT_{n+1}}(x_{T_i})$ is immediate by equation (3.3). Finally the functional forms of ZCBs $D_{T_iT_j}(x_{T_i})$ can be obtained by the martingale property

$$D_{T_iT_j}(x_{T_i}) = D_{T_iT_{n+1}}(x_{T_i}) E_{\mathbb{F}^{n+1}}[\frac{1}{D_{T_jT_{n+1}}(x_{T_j})}|x_{T_i}]. \tag{3.9}$$

for $i < j \leq n$.

Let us finish the specification of the model by a remark. The Markov driving process affects the model only by means of the transition density function $\phi_{x_{T_j}|x_{T_i}}(\cdot)$ with mean x_{T_i} and variance $\int_{T_i}^{T_j} \sigma_u^2 du$ for $0 \leq i < j \leq n+1$. In the model the driving process x is taken as a Gaussian process, but theoretically the Markov process can be moved away from Gaussian. Gogala and Kennedy [29] developed one-dimensional MFMs with non-Gaussian Markov processes. However using a non-Gaussian Markov process could cause challenges for the numerical implementation. In general it is more difficult to find the transition density function for a non-Gaussian driving process. In many cases we have to resort to approximation which will be discussed later in this chapter.

3.2.2 One-dimensional Markov-functional models: Swap rate version

In this subsection we consider one-dimensional swap MFMs. The specification of one-dimensional swap MFMs is similar to the LIBOR version but we find functional forms of swap rates by feeding in input prices of PVBP-digital swaptions. Suppose we are given a Gaussian process (3.1) which will be taken as the driving process.

The algorithm for swap based model specification is similar to the LIBOR version. In particular we find the functional forms $D_{T_iT_{n+1}}(x_{T_i})$ from time T_{n+1} to time T_1. At time T_{n+1} by definition we have

$$D_{T_{n+1}T_{n+1}}(x_{T_{n+1}}) = 1.$$

Assume that we have determined functional forms $D_{T_jT_k}(x_{T_j})$ for $j = i+1, ..., n+1$ and $k = j, ..., n+1$ and therefore the functional form of the PVBP $P^{i,n+1-i}_{T_{i+1}}(x_{T_{i+1}})$ is also determined. At time T_i, by definition, the swap rate $y^{i,n+1-i}_{T_i}$ is given by

$$y^{i,n+1-i}_{T_i} = \frac{1 - D_{T_iT_{n+1}}}{P^{i,n+1-i}_{T_i}}, \tag{3.10}$$

so that the numeraire at time T_i can be expressed as

$$D_{T_iT_{n+1}} = \frac{1}{\widehat{P}^{i,n+1-i}_{T_i} y^{i,n+1-i}_{T_i} + 1}. \tag{3.11}$$

where the numeraire-rebased PVBP is defined as

$$\widehat{P}^{i,n+1-i}_{T_i} := \frac{P^{i,n+1-i}_{T_i}}{D_{T_iT_{n+1}}}.$$

We can see from equation (3.11) that finding the functional form of $D_{T_iT_{n+1}}(x_{T_i})$ is equivalent to finding the functional forms of $\widehat{P}^{i,n+1-i}_{T_i}(x_{T_i})$ and $y^{i,n+1-i}_{T_i}(x_{T_i})$.

We first consider the functional form of $\widehat{P}^{i,n+1-i}_{T_i}(x_{T_i})$. It follows from the martingale property that

$$\widehat{P}^{i,n+1-i}_{T_i}(x_{T_i}) = E_{\mathbb{F}^{n+1}}[\frac{P^{i,n+1-i}_{T_{i+1}}(x_{T_{i+1}})}{D_{T_{i+1}T_{n+1}}(x_{T_{i+1}})}|x_{T_i}],$$

which can be evaluated using numerical integration (3.5).

Then we consider the functional form $y^{i,n+1-i}_{T_i}(x_{T_i})$. To do so we calibrate to the input prices of PVBP-digital swaptions. Suppose the input prices of PVBP-

digital swaptions $V_0^i(K)$ on the ith forward swap rate $y^{i,n+1-i}$,$i = 1, ..., n$, is a continuous function of the strike K i.e. $V_0^i : [0, \infty] \to \mathbb{R}$ and are given to us. The PVBP-digital swaptions on $y^{i,n+1-i}$ at strike K has the following payoff at time T_i

$$V_{T_i}^i = P_{T_i}^{i,n+1-i} \mathbf{I}_{\{y_{T_i}^{i,n+1-i} > K\}}.$$

Following the fundamental pricing formula, we have that

$$V_0^i(K) = D_{0T_{n+1}} E_{\mathbb{F}^{n+1}}[\widehat{P}_{T_i}^{i,n+1-i}(x_{T_i}) \mathbf{I}_{\{y_{T_i}^{i,n+1-i} > K\}}]. \tag{3.12}$$

We can see that the input prices $V_0^i(K)$ is a decreasing function of the strike K.

To determine the functional forms $y_{T_i}^{i,n+1-i}(x_{T_i})$ we still need one more assumption:

(A.2) In the model, $y_{T_i}^{i,n+1-i}$ is a monotonic increasing function of x_{T_i} for $i = 1, ..., n$.

We now find the functional form $y_{T_i}^{i,n+1-i}(x_{T_i})$. Choose a value x^* of x_{T_i} and for each x^* we calculate

$$J_0^i(x^*) = D_{0T_{n+1}} E_{\mathbb{F}^{n+1}}[\widehat{P}_{T_i}^{i,n+1-i}(x_{T_i}) \mathbf{I}_{\{x_{T_i} > x^*\}}], \tag{3.13}$$

which can be obtained by numerical integration. Compare equation (3.12) with (3.13), and it turns out that functions J_0^i and V_0^i are the same except for the indicator function. By assumption (A.2) for any value of x^* we can always find a unique strike K such that

$$\{x_{T_i} > x^*\} = \{y_{T_i}^{i,n+1-i} > K\}. \tag{3.14}$$

So for any value of $J_0^i(x^*)$, we can always find a unique strike K such that

$$J_0^i(x^*) = V_0^i(K), \tag{3.15}$$

where the function $K(x^*)$ can be solved as

$$K(x^*) = (V_0^i)^{-1}(J_0^i(x^*)).$$

It follows from equation (3.14) that

$$y_{T_i}^{i,n+1-i}(x^*) = K(x^*).$$

So far we have found the functional forms of $\widehat{P}_{T_i}^{i,n+1-i}(x_{T_i})$ and $y_{T_i}^{i,n+1-i}(x_{T_i})$, and the functional forms of the numeraire $D_{T_iT_{n+1}}(x_{T_i})$ is immediate by equation (3.11). Finally the functional forms of ZCBs $D_{T_iT_j}(x_{T_i})$ can be obtained by the

martingale property

$$D_{T_iT_j}(x_{T_i}) = D_{T_iT_{n+1}}(x_{T_i})E_{\mathbb{F}^{n+1}}[\frac{1}{D_{T_jT_{n+1}}(x_{T_j})}|x_{T_i}].$$

for $i < j \leq n$.

So far we have introduced the one-dimensional LIBOR and swap rate based MFMs. One may ask which version would be preferable. Bennett and Kennedy [7] investigated both LIBOR and swap one-dimensional MFMs with Gaussian driver numerically, and it turned out that the forward LIBORs may become negative in some cases for long maturities under the swap MFM. But to our knowledge there is no consensus yet as to which version is better. A rule of thumb is that the LIBOR MFM is more suitable for pricing LIBOR-based derivatives whereas the swap version is much more used to price swap-based derivatives.

3.2.3 Multi-dimensional Markov-functional models

One-dimensional MFMs are suitable for pricing and hedging level dependent derivatives such as Bermudan swaptions but not sufficient for all products. One product which needs a multidimensional model is spread options whose payoff can be seen as the difference between two different forward rates. In order to price such a product accurately the relative level of the two forward rates i.e. the skew of the forward rates curve has to be captured. So two-dimensional MFMs are needed where the two-dimensional Markov process can be viewed as capturing the level and skew of interest rates. For more details about level and skew, the reader is referred to [1] and [21].

Hunt and Kennedy [38] proposed a multi-dimensional MFM and presented an example for the two-dimensional MFM with a two-dimensional Gaussian driving process. Kaisajuntti and Kennedy [45] developed an d-dimensional MFM. Now we specify an d-dimensional LIBOR MFM briefly, and we will later adapt the ideas for stochastic volatility MFM. The swap version can also be generalized. An d-dimensional MFM is just a generalization of the one-dimensional case so that the numeraire approach can also be applied.

Suppose we are given the following d-dimensional Markov process in $\mathbb{F}^{n+1}$:

$$x = (x^1, x^2, ..., x^d).$$

Recall from the one-dimensional LIBOR MFM that the monotonic assumption (A.1) is necessary to determine the functional forms of LIBORs at their setting dates.

However the univariate and monotonicity properties are lost here due to the multi-dimensionality of x. To deal with this issue we introduce a pre-model $\widehat{L}^i_{T_i} : \mathbb{R}^d \to \mathbb{R}$ which is a function of the d-dimensional driver:

$$\widehat{L}^i_{T_i}(x_{T_i}) = f^i(x^1_{T_i}, x^2_{T_i}, ...x^d_{T_i}). \tag{3.16}$$

We will see an example for the choice of the function f^i later. Once a pre-model is chosen we make the following assumption:
(A.3): In the model, $L^i_{T_i}$ is a monotonic increasing function of the pre-model $\widehat{L}^i_{T_i}$ for $i = 1, ..., n$.
The remaining step is almost the same as the one-dimensional MFM, but all the required integrals become d-dimensional.

Hunt and Kennedy [38] considered the case where x is a two-dimensional Gaussian process. In this case we have the following two-dimensional Markov process

$$\begin{aligned} dx^1_t &= \sigma^1_t dW^1_t, \\ dx^2_t &= \sigma^2_t dW^2_t, \\ dW^1_t dW^2_t &= \rho dt, \end{aligned} \tag{3.17}$$

where σ^1 and σ^2 are deterministic functions, and W^1 and W^2 are correlated Brownian motions in $\mathbb{F}^{n+1}$. Note that the two-dimensional Markov process x in (3.17) can be viewed as representing the level as well as skew of interest rates. In this case a pre-model is commonly taken as some strictly increasing function of the linear combination of the components of the driver:

$$\widehat{L}^i_{T_i}(x_{T_i}) = g^i(\gamma^1_i x^1_{T_i} + \gamma^2_i x^2_{T_i}), \tag{3.18}$$

where γ^1_i and γ^2_i are positive constants and g^i is some strictly increasing function. Note that the pre-model can be chosen to capture some desired covariance structure in mind. More details about the choices of pre-model for the Gaussian driver (3.17) can be found in [38] and [43].

Note that the multi-dimensional driving Markov process is not forced to be Gaussian. But to our knowledge there is no paper developing multi-dimensional MFMs with non-Gaussian drivers. Gogala and Kennedy [29] developed MFMs with non-Gaussian Markov processes in the one-dimensional case. They also showed that, under some conditions, one-dimensional MFMs driven by two distinct driving processes are equivalent by using copula theory. We extend this equivalence of MFMs to the multi-dimensional case in Appendix 3.A. In the next section we

will develop a stochastic volatility MFM. We will take a two-dimensional stochastic volatility model, which is specified in a multiplicative way, as the driving process, and we will choose a pre-model different from the choice (3.18).

3.3 Stochastic volatility Markov-functional models

We have discussed the specification of MFMs without stochastic volatility in the previous section. However there is much empirical evidence supporting the stochastic volatility for interest rates; see [15], [49] and references therein. Furthermore as Hagan et al. [31] pointed out, incorporating an extra stochastic volatility factor into a model could give a more realistic evolution for the implied volatility smile. In addition Kaisajuntti and Kennedy [44] used market data to indicate that a stochastic volatility Markov process rather than a one-dimensional Markov process should be used for the level of rates. These motivate us to develop a stochastic volatility MFM by taking a stochastic volatility process as a driver.

So far our focus is on the theoretical specification of MFMs rather than the details of the numerical implementation. In Section 3.3.1 we consider the numerical implementation of a two-dimensional (stochastic volatility) LIBOR MFM by developing a practical algorithm, where we work with basis functions and conditional moments of the driving Markov process. Note that a two-dimensional algorithm that appears in [38] for the additive pre-model has been discussed earlier but this algorithm relies heavily on the Gaussian assumption. The practical algorithm proposed in this section is not specific and could be modified to apply to all one- and multi-dimensional MFMs. This general algorithm for the corresponding swap version MFM is also straightforward to develop in a similar way. In Section 3.3.2 we consider the specification of the driving process for the stochastic volatility MFM.

3.3.1 Numerical implementation

Let us develop a practical algorithm to implement a two-dimensional LIBOR MFM numerically. This algorithm is very general and does not rely on the Gaussian assumption of the driving process. Suppose we are given a two-dimensional diffusion process (y, q) and we take it as a driving Markov process. The specification of the driver for a stochastic volatility MFM will be discussed later in this section.

Grid Points

We now present the numerical implementation of the model on a grid. We construct grid points for the Markov process on the date structure T_i , $i = 1, ..., n$. we

construct m equidistant grid points $-\infty < y_{i,1} < y_{i,2} < ... < y_{i,m} < \infty$ and $-\infty < q_{i,1} < q_{i,2} < ... < q_{i,m} < \infty$ for (y_{T_i}, q_{T_i}) at time T_i for $i = 1, ..., n$. The grid points are chosen such that the intervals $[y_{i,1}, y_{i,m}]$ and $[q_{i,1}, q_{i,m}]$ cover most of probability mass of y_{T_i} and q_{T_i} respectively i.e.

$$\mathbb{F}^{n+1}(y_{i,1} \leq y_{T_i} \leq y_{i,m}) \geq 1 - \epsilon_y$$
$$\mathbb{F}^{n+1}(q_{i,1} \leq q_{T_i} \leq q_{i,m}) \geq 1 - \epsilon_q$$

where positive constants ϵ_y and ϵ_q are small enough e.g. 0.001%.

Basis Function

Since MFMs are implemented on a grid, any smooth function of the state variable (y_{T_i}, q_{T_i}) is valued and discretized on the grid points. In order to approximate and interpolate these functions we introduce the idea of basis function.

The one-dimensional basis function $b_{i,j} : \mathbb{R} \to \mathbb{R}$ for $i = 1, ..., n$ and $j = 1, ..., m$ with respect to the partition of $[y_{i,1}, y_{i,m}]$ can be written in the form of

$$b_{i,j}(y) := \sum_{j'=1}^{m-1} \sum_{d=0}^{M} b^y_{i,j,j',d} y^d \mathbf{I}_{\{y \in [y_{i,j'}, y_{i,j'+1})\}}$$

satisfying

$$b_{i,j}(y_{i,u}) = \delta_{j,u}$$

for some coefficients $b^y_{i,j,j',d} \in \mathbb{R}$, where $\delta_{j,u}$ is Kronecker delta function. The details of a basis function construction and the calculation of its coefficients can be found in Appendix 3.B.

The two-dimensional basis functions w.r.t the partition of $[y_{i,1}, y_{i,m}] \times [q_{i,1}, q_{i,m}]$ can be expressed as the product of the two corresponding one-dimensional basis functions

$$\begin{aligned} b_{i,j,k}(y, q) &:= b_{i,j}(y) \times b_{i,k}(q) \\ &= \sum_{j'k'=1}^{m-1} \sum_{d,l=0}^{M} b^y_{i,j,j',d} b^q_{i,k,k',l} y^d q^l \mathbf{I}_{\{y \in [y_{i,j'}, y_{i,j'+1})\} \cap \{q \in [q_{i,k'}, q_{i,k'+1})\}} \end{aligned} \tag{3.19}$$

satisfying

$$b_{i,j,k}(y_{i,u}, q_{i,v}) = \delta_{j,u} \times \delta_{k,v}$$

for coefficients $(b^y_{i,j,j',d}, b^q_{i,k,k',l}) \in \mathbb{R}^2$, $i = 1, ..., n$ and $j, k = 1, ..., m$.

Notice that the basis functions above are just piecewise polynominal, and they are used to approximate a smooth function. In particular, any smooth function $f : \mathbb{R} \to \mathbb{R}$ defined on the interval $[y_{i,1}, y_{i,m}]$ can be approximated by a piecewise polynomial function $\tilde{f}$ in terms of basis functions:

$$\tilde{f}(y) := \sum_{j=1}^{m} f(y_{i,j}) b_{i,j}(y). \tag{3.20}$$

Similarly, any two-dimensional smooth function $g : \mathbb{R}^2 \to \mathbb{R}$ defined on the interval $[y_{i,1}, y_{i,m}] \times [q_{i,1}, q_{i,m}]$ can be approximated by a piecewise polynomial function $\tilde{g}$ in terms of two-dimensional basis functions:

$$\tilde{g}(y, q) := \sum_{j,k=1}^{m} g(y_{i,j}, q_{i,k}) b_{i,j,k}(y, q). \tag{3.21}$$

Building Blocks

Now we define and evaluate three (conditional) expectations which will be taken as building blocks. We will show that MFMs can be implemented by using these building blocks. In practice we will store these building blocks for efficient implementation. Let us first consider the following one step conditional moments:

$$\begin{aligned} &\Theta^i_{j,k,d,l}(y_{i-1,u}, q_{i-1,v}) \\ &:= E_{\mathbb{F}^{n+1}}[y^d_{T_i} q^l_{T_i} \mathbf{I}_{\{y_{T_i} \in [y_{i,j}, y_{i,j+1})\} \cap \{q_{T_i} \in [q_{i,k}, q_{i,k+1})\}} | y_{T_{i-1}} = y_{i-1,u}, q_{T_{i-1}} = q_{i-1,v}] \end{aligned} \tag{3.22}$$

for $i = 1, ..., n$; $j, k = 1, ..., m-1$; $d, l = 0, ..., M$ and $u, v = 1, ..., m$. Note that the information on the transition density function of the driving process is given to us via the above one step conditional moments. We will present the evaluation of the above one step conditional moments later in this section after we choose and discuss the driving process. For now we assume that the functions $\Theta^i_{j,k,d,l} : \mathbb{R}^2 \to \mathbb{R}$ are given at each grid point. Suppose we have a specific driving process and pre-model in mind, which will be discussed later, and we define the following three (conditional) expectations:

$$\begin{aligned} &\Delta^i_{j,k}(y_{i-1,u}, q_{i-1,v}) := E_{\mathbb{F}^{n+1}}[b_{i,j,k}(y_{T_i}, q_{T_i}) | y_{T_{i-1}} = y_{i-1,u}, q_{T_{i-1}} = q_{i-1,v}], \\ &E_{i,j,k} := E_{\mathbb{F}^{n+1}}[b_{i,j,k}(y_{T_i}, q_{T_i})], \\ &\Gamma^i_{j,k,u} := E_{\mathbb{F}^{n+1}}[b_{i,j,k}(y_{T_i}, q_{T_i}) \mathbf{I}_{\{y_{T_i} \in [y_{i,u}, y_{i,u+1})\}}]. \end{aligned}$$

for $i = 1, ..., n$; $j, k = 1, ..., m$ and $u = 1, ..., m$. Note that the third building blocks Γ may differ for a different choice of driving process and pre-model and we will see this later.

Remark 4. *The building blocks $\Delta^i_{j,k}(y_{i-1,u}, q_{i-1,v})$ and $\Gamma^i_{j,k,u}$ will be needed to determine the functional forms of ZCBs. We use the building block $E_{i,j,k}$ in applications such as pricing options.*

Now we can show that the three (conditional) expectations can be evaluated via the function $\Theta^i_{j,k,d,l}$.

Proposition 1. *The three (conditional) expectations can be expressed as*

$$\Delta^i_{j,k}(y_{i-1,u}, q_{i-1,v}) = \sum_{j'=1}^{m-1}\sum_{d=0}^{M} b^y_{i,j,j',d} \sum_{k'=1}^{m-1}\sum_{l=0}^{M} b^q_{i,k,k',l} \Theta^i_{j',k',d,l}(y_{i-1,u}, q_{i-1,v})$$

$$E_{i,j,k} = \sum_{u,v=1}^{m} \Delta^i_{j,k}(y_{i-1,u}, q_{i-1,v}) E_{i-1,u,v}$$

$$\Gamma^i_{j,k,u} = \sum_{d=0}^{M} b^y_{i,j,u,d} \sum_{k'=1}^{m-1}\sum_{l=0}^{M} b^q_{i,k,k',l} \sum_{u',v=1}^{m} \Theta^i_{u,k',d,l}(y_{i-1,u'}, q_{i-1,v}) E_{i-1,u'v}$$

where b^y's and b^q's are coefficients for basis functions (3.19).

Proof. See Appendix 3.C. □

Having calculated the building blocks one can store and use them in the numerical implementation.

Practical Algorithm

We now present the numerical implementation of the stochastic volatility LIBOR MFM by developing a practical algorithm based on the building blocks. We determine functional forms of ZCBs by backwards induction on time T_i. At time T_{n+1} by definition we have

$$D_{T_{n+1}T_{n+1}}(y_{T_{n+1}}, q_{T_{n+1}}) = 1.$$

Assume that we have determined $D_{T_jT_k}(y_{T_j}, q_{T_j})$ for $j = i+1, ..., n+1$ and $k = j, ..., n+1$. At time T_i, by definition, the LIBOR $L^i_{T_i}$ is given by

$$L^i_{T_i} = \frac{1 - D_{T_iT_{i+1}}}{\alpha_i D_{T_iT_{i+1}}}$$

that the numeraire at time T_i can be expressed as

$$D_{T_iT_{n+1}} = \frac{1}{\widehat{D}_{T_iT_{i+1}}(1+\alpha_i L^i_{T_i})}, \tag{3.23}$$

where the numeraire-rebased ZCBs is defined as

$$\widehat{D}_{T_iT_{i+1}} := \frac{D_{T_iT_{i+1}}}{D_{T_iT_{n+1}}}.$$

We can observe from equation (3.23) that once we have determined functional forms of $\widehat{D}_{T_iT_{i+1}}(y_{T_i}, q_{T_i})$ and $L^i_{T_i}(y_{T_i}, q_{T_i})$, the functional form of the numeraire $D_{T_iT_{n+1}}(y_{T_i}, q_{T_i})$ at time T_i is immediate.

We first find functional forms of $\widehat{D}_{T_iT_{i+1}}(y_{T_i}, q_{T_i})$. It follows from the martingale property and equation (3.21) that

$$\begin{aligned}
\widehat{D}_{T_iT_{i+1}}(y_{T_i}, q_{T_i}) &= E_{\mathbb{F}^{n+1}}[\frac{1}{D_{T_{i+1}T_{n+1}}(y_{T_{i+1}}, q_{T_{i+1}})}|y_{T_i}, q_{T_i}] \\
&\approx E_{\mathbb{F}^{n+1}}[\sum_{j,k=1}^{m} \frac{1}{D_{T_{i+1}T_{n+1}}(y_{i+1,j}, q_{i+1,k})} b_{i+1,j,k}(y_{T_{i+1}}, q_{T_{i+1}})|y_{T_i}, q_{T_i}] \\
&= \sum_{j,k=1}^{m} \frac{1}{D_{T_{i+1}T_{n+1}}(y_{i+1,j}, q_{i+1,k})} \Delta^{i+1}_{j,k}(y_{T_i}, q_{T_i}),
\end{aligned}$$

where $D_{T_{i+1}T_{n+1}}(y_{i+1,j}, q_{i+1,k})$ has already been determined at time T_{i+1}.

Now we determine the functional form $L^i_{T_i}(y_{T_i}, q_{T_i})$ by feeding in the input prices of digital caplets corresponding to the ith LIBOR L^i. We assume that the input prices $V^i_0 : [0, \infty] \to \mathbb{R}$ of digital caplets at strikes $K \geq 0$ are given for $i = 1, ..., n$. Recall that following the fundamental pricing formula in $\mathbb{F}^{n+1}$ we have equation (3.6).

In order to find the functional forms of LIBORs we still need one more assumption:

(A.4): In the model, $L^i_{T_i}$ is a monotonic increasing function of y_{T_i}.

Note that one can think of the variable y_{T_i} itself as a pre-model at time T_i for $i = 1, ..., n$. For a different specification of the driving process, the form of the pre-model may differ. An explanation of this assumption will be given later after we specify the driving process (y, q).

Now we determine the functional form $L^i_{T_i}(y_{T_i})$. For each grid point $y_{i,j}$,

from $j = m$ to $j = 1$, we evaluate

$$\begin{aligned} J_0^i(y_{i,j}) &= D_{0T_{n+1}} E_{\mathbb{F}^{n+1}}[\widehat{D}_{T_iT_{i+1}}(y_{T_i}, q_{T_i})\mathbf{I}_{\{y_{T_i} > y_{i,j}\}}] \\ &\approx D_{0T_{n+1}} E_{\mathbb{F}^{n+1}}[\sum_{u,v=1}^{m} \widehat{D}_{T_iT_{i+1}}(y_{i,u}, q_{i,v}) b_{i,u,v}(y_{T_i}, q_{T_i})\mathbf{I}_{\{y_{T_i} > y_{i,j}\}}] \\ &= D_{0T_{n+1}} \sum_{u,v=1}^{m} \widehat{D}_{T_iT_{i+1}}(y_{i,u}, q_{i,v}) E_{\mathbb{F}^{n+1}}[b_{i,u,v}(y_{T_i}, q_{T_i})\mathbf{I}_{\{y_{T_i} > y_{i,j}\}}] \\ &= J_0^i(y_{i,j+1}) + D_{0T_{n+1}} \sum_{u,v=1}^{m} \widehat{D}_{T_iT_{i+1}}(y_{i,u}, q_{i,v})\Gamma^i_{u,v,j}. \end{aligned} \tag{3.24}$$

Compare equation (**??**) with (3.24), and it turns out that functions J_0^i and V_0^i are the same except for the indicator function. By assumption (A.4) for each value $y_{i,j}$ we can always find a unique strike K such that

$$\{y_{T_i} > y_{i,j}\} = \{L^i_{T_i} > K\} \tag{3.25}$$

holds. Thus for each value of $J_0^i(y_{i,j})$, we can always find a unique strike K such that

$$J_0^i(y_{i,j}) = V_0^i(K),$$

where the function $K(y_{i,j})$ can be solved as

$$K(y_{i,j}) = (V_0^i)^{-1}(J_0^i(y_{i,j})).$$

It follows from (3.25) that

$$L^i_{T_i}(y_{i,j}) = K(y_{i,j}).$$

Note that $L^i_{T_i}$ is just a function of random variable y_{T_i} so that $L^i_{T_i}(y_{T_i}, q_{T_i}) = L^i_{T_i}(y_{T_i})$.

We have found the functional form $L^i_{T_i}(y_{T_i})$, and therefore the functional form $D_{T_iT_{n+1}}(y_{T_i}, q_{T_i})$ is immediate from equation (3.23). Finally the functional forms of ZCBs $D_{T_iT_j}(y_{T_i}, q_{T_i})$, $i < j < n + 1$, can be determined by the martingale property:

$$\begin{aligned} D_{T_iT_j}(y_{T_i}, q_{T_i}) &= D_{T_iT_{n+1}}(y_{T_i}, q_{T_i}) E_{\mathbb{F}^{n+1}}[\frac{D_{T_{i+1}T_j}(y_{T_{i+1}}, q_{T_{i+1}})}{D_{T_{i+1}T_{n+1}}(y_{T_{i+1}}, q_{T_{i+1}})} | y_{T_i}, q_{T_i}] \\ &\approx D_{T_iT_{n+1}}(y_{T_i}, q_{T_i}) \sum_{u,v=1}^{m} \frac{D_{T_{i+1}T_j}(y_{i+1,u}, q_{i+1,v})}{D_{T_{i+1}T_{n+1}}(y_{i+1,u}, q_{i+1,v})} \Delta^{i+1}_{u,v}(y_{T_i}, q_{T_i}). \end{aligned}$$

In pricing applications we will need to take expectations of payoff functions. If the expectation of a smooth function f^i of the bivariate random variable (y_{T_i}, q_{T_i}) is finite, we can approximate it as

$$\begin{aligned} E_{\mathbb{F}^{n+1}}[f^i(y_{T_i}, q_{T_i})] &\approx E_{\mathbb{F}^{n+1}}[\sum_{u,v=1}^{m} f^i(y_{i,u}, q_{i,v}) b_{i,u,v}(y_{T_i}, q_{T_i})] \\ &= \sum_{u,v=1}^{m} f^i(y_{i,u}, q_{i,v}) E_{i,u,v}. \end{aligned}$$

3.3.2 Specification of the driving process

So far we have discussed the numerical implementation of a two-dimensional LIBOR MFM without giving details about the specification of the driving process (y, q). In this subsection we will focus on the specification of a diving process (y, q) by considering a SABR model. Then the evaluation of the one step conditional moments (3.22) for this particular driving process (y, q) will be given and the monotonic increasing assumption (A.4) will be discussed.

Kaisajuntti and Kennedy [44] used market data to identify a SABR type model as an appropriate choice for the level of rates. It means that a SABR model is an appropriate choice to start with. Suppose we are given the following SABR model

$$\begin{aligned} dF_t &= \sigma_t F_t^{\beta} dW_t^{n+1} \qquad \beta \in [0,1], \\ d\sigma_t &= \mu_t \sigma_t dt + v\sigma_t dB_t^{n+1} \qquad v > 0, \\ dB_t^{n+1} dW_t^{n+1} &= \rho dt \qquad \rho \in [-1,1] \end{aligned} \tag{3.26}$$

where B^{n+1} and W^{n+1} are correlated Brownian motions in $\mathbb{F}^{n+1}$, and μ_t is assumed to be piecewise constant

$$\mu_t = \sum_{j=0}^{n-1} \mu_j I_{\{t \in [T_j, T_{j+1})\}}.$$

Note that (3.26) is a modified SABR model. The volatility process σ is driftless in the original SABR model [31]. The purpose of adding the drift μ_j's to the stochastic volatility σ will be explained later in the next chapter where we will see that the drift controls the auto-correlation of the driver. One may take the above SABR process (F, σ) as the driving process of the stochastic volatility LIBOR MFM. But we note that it is hard to find the transition density of the two-dimensional process (F, σ) so

that the one step conditional moments for (F, σ) is difficult to obtain. Although Islah [40] proved an exact analytical formula for the joint density function of (F, σ) when correlation is zero, this closed-form formula involves an integral which is inefficient to calculate in practice. Thus we will take an approximation of the SABR process (F, σ) as the driving process which allows for an efficient calculation of the one step conditional moments (3.22) for such driver.

As justified by Kennedy et. al [47], the displaced diffusion (DD) SABR model can be used as a close approximation to the SABR model, and DD SABR model admits a closed-form solution. In particular consider the DD-SABR model which satisfies the following SDEs

$$\begin{aligned} dF_t &= \widehat{\sigma}_t(F_t + \theta)dW_t^{n+1} \\ d\sigma_t &= \mu_t\sigma_t dt + v\sigma_t dB_t^{n+1} \quad v > 0 \\ dB_t^{n+1}dW_t^{n+1} &= \rho dt. \end{aligned} \tag{3.27}$$

The distribution $(F_T, \sigma_T|F_s, \sigma_s)$, $s < T$, of the SABR model (3.26) and the DD-SABR model (3.27) become comparable via the following mapping

$$\widehat{\sigma}_t = \sigma_t \beta F_s^{\beta-1} \tag{3.28}$$

$$\theta = F_s \frac{1-\beta}{\beta}, \tag{3.29}$$

where $\beta \in (0, 1]$. The two models are matched exactly when $\beta = 1$. Kennedy et. al [47] justified this mapping numerically. This mapping was also dicussed by [51] and [63] in the non-stochastic volatility case.

For ease of explanation, from now on, the model (3.26) is referred to as the Normal-SABR model for $\beta = 0$ and the CEV-SABR model for $\beta \in (0, 1]$. In order to evaluate the one step conditional moments from time T_i to T_{i+1}, instead of considering the CEV-SABR model (3.26), we approximate the model (3.26) by the DD-SABR model (3.27) via the mappings (3.28) and (3.29) with $s = T_i$, as in this case the DD-SABR model gives a similar distribution of $(F_{T_{i+1}}, \sigma_{T_{i+1}}|F_{T_i}, \sigma_{T_i})$ as the CEV-SABR model. Now we reformulate the model (3.27) in a form that will help us calculate the one step conditional moments. We write down the Normal-SABR model and DD-SABR model as follows, where F satisfies (3.27), and define (y, q) to be the driver for our stochastic volatility MFM.

Normal SABR:

$$\begin{aligned} y_{T_{i+1}} &:= F_{T_{i+1}} \\ &= y_{T_i} + \frac{\rho}{v}(\sigma_{T_{i+1}} - \sigma_{T_i}) - \frac{\rho}{\nu}\int_{T_i}^{T_{i+1}} \mu_t \sigma_t dt + \sqrt{1-\rho^2}\int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1} \\ q_{T_{i+1}} &:= \ln \sigma_{T_{i+1}} \\ &= q_{T_i} + \int_{T_i}^{T_{i+1}} \mu_t dt - \frac{1}{2}v^2\alpha_i + v(B_{T_{i+1}}^{n+1} - B_{T_i}^{n+1}) \end{aligned} \tag{3.30}$$

DD-SABR:

$$\begin{aligned} y_{T_{i+1}} &:= \ln(F_{T_{i+1}} + \theta) \\ &= y_{T_i} + \omega_i \frac{\rho}{v}[(\sigma_{T_{i+1}} - \sigma_{T_i}) - \int_{T_i}^{T_{i+1}} \mu_t \sigma_t dt] \\ &- \frac{1}{2}\omega_i^2(\int_{T_i}^{T_{i+1}} \sigma_t^2 dt) + \omega_i\sqrt{1-\rho^2}\int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1} \\ q_{T_{i+1}} &:= \ln \sigma_{T_{i+1}} \\ &= q_{T_i} + \int_{T_i}^{T_{i+1}} \mu_t dt - \frac{1}{2}v^2\alpha_i + v(B_{T_{i+1}}^{n+1} - B_{T_i}^{n+1}) \end{aligned} \tag{3.31}$$

with

$$\begin{aligned} \alpha_i &= T_{i+1} - T_i \\ \omega_i &= \beta e^{(\beta-1)y_{T_i}}, \end{aligned}$$

where B^{n+1} and $\widehat{W}^{n+1}$ are independent Brownian motions in $\mathbb{F}^{n+1}$. The proof is a slight generalization of that found in the Appendix of [47].

Calculation of the one step conditional moments

From the form of the driving process (y, q) given above, we now consider the problem of how to calculate the one step conditional moments (3.22). We note that it is still hard to find the transition density function of the driver (y, q). As a result using numerical integration to solve the one step conditional moments directly seems difficult to achieve. To solve this problem we apply the tower property to the one step conditional moments (3.22), conditioning on $(y_{T_i}, q_{T_i}, q_{T_{i+1}})$, which can be rewritten

as

$$\begin{aligned}
&\Theta^{i+1}_{j,k,d,l}(y_{T_i}, q_{T_i}) \\
&:= E_{\mathbb{F}^{n+1}}[y^d_{T_{i+1}} q^l_{T_{i+1}} \mathbf{I}_{\{y_{T_{i+1}} \in [y_{i+1,j}, y_{i+1,j+1})\} \cap \{q_{T_{i+1}} \in [q_{i+1,k}, q_{i+1,k+1})\}} | y_{T_i}, q_{T_i}] \\
&= E_{\mathbb{F}^{n+1}}[q^l_{T_{i+1}} \mathbf{I}_{\{q_{T_{i+1}} \in [q_{i+1,k}, q_{i+1,k+1})\}} \Xi^{i+1}_{j,d}(y_{T_i}, q_{T_i}, q_{T_{i+1}}) | y_{T_i}, q_{T_i}], \qquad (3.32)
\end{aligned}$$

where the function $\Xi^{i+1}_{j,d}(y_{T_i}, q_{T_i}, q_{T_{i+1}})$ is defined by

$$\begin{aligned}
&\Xi^{i+1}_{j,d}(y_{T_i}, q_{T_i}, q_{T_{i+1}}) \qquad (3.33) \\
&:= E_{\mathbb{F}^{n+1}}[y^d_{T_{i+1}} \mathbf{I}_{\{y_{T_{i+1}} \in [y_{i+1,j}, y_{i+1,j+1})\}} | y_{T_i}, q_{T_i}, q_{T_{i+1}}].
\end{aligned}$$

We know from equation (3.20) that the smooth function $\Xi^{i+1}_{j,d}(y_{T_i}, q_{T_i}, q_{T_{i+1}})$ with respect to $q_{T_{i+1}}$ can be approximated by a piecewise polynomial function in terms of the basis functions:

$$\begin{aligned}
&\Xi^{i+1}_{j,d}(y_{T_i}, q_{T_i}, q_{T_{i+1}}) \\
&\approx \sum_{j'=1}^{m} \Xi^{i+1}_{j,d}(y_{T_i}, q_{T_i}, q_{i+1,j'}) b_{i+1,j'}(q_{T_{i+1}}). \qquad (3.34)
\end{aligned}$$

Inserting equation (3.34) into equation (3.32), we have that

$$\begin{aligned}
&\Theta^{i+1}_{j,k,d,l}(y_{T_i}, q_{T_i}) \qquad (3.35) \\
&\approx \sum_{j'=1}^{m} \Xi^{i+1}_{j,d}(y_{T_i}, q_{T_i}, q_{i+1,j'}) E_{\mathbb{F}^{n+1}}[q^l_{T_{i+1}} \mathbf{I}_{\{q_{T_{i+1}} \in [q_{i+1,k}, q_{i+1,k+1})\}} b_{i+1,j'}(q_{T_{i+1}}) | y_{T_i}, q_{T_i}].
\end{aligned}$$

Therefore finding the one step conditional moments $\Theta^{i+1}_{j,k,d,l}$ is equivalent to finding the function $\Xi^{i+1}_{j,d}$ and

$$E_{\mathbb{F}^{n+1}}[q^l_{T_{i+1}} \mathbf{I}_{\{q_{T_{i+1}} \in [q_{i+1,k}, q_{i+1,k+1})\}} b_{i+1,j'}(q_{T_{i+1}}) | y_{T_i}, q_{T_i}].$$

In what follows we calculate these two functions.

We first evaluate the function $\Xi^{i+1}_{j,d}$ by introducing an approximation. Note that to evaluate the function $\Xi^{i+1}_{j,d}$ we need to find the distribution $(y_{T_{i+1}} | y_{T_i}, q_{T_i}, q_{T_{i+1}})$, which is again hard to achieve. However Kennedy et. al [47] found that the conditional distribution $(y_{T_{i+1}} | y_{T_i}, q_{T_i}, q_{T_{i+1}})$ can be approximated by a Gaussian distri-

bution with the the following conditional mean and variance:

$$\mu(y_{T_i}, q_{T_i}, q_{T_{i+1}}) = E_{\mathbb{F}^{n+1}}(y_{T_{i+1}}|y_{T_i}, q_{T_i}, q_{T_{i+1}}),$$
$$\eta^2(y_{T_i}, q_{T_i}, q_{T_{i+1}}) = Var(y_{T_{i+1}}|y_{T_i}, q_{T_i}, q_{T_{i+1}}).$$

Based on this approximation we can evaluate the function $\Xi_{j,d}^{i+1}$ by numerical integration as long as we can find the conditional mean μ and variance η^2.

Proposition 2. *Consider expressions (3.30) and (3.31). The conditional mean and variance of $y_{T_{i+1}}|y_{T_i}, q_{T_i}, q_{T_{i+1}}$ are given by the following closed-form expressions:*
Normal SABR:

$$\begin{aligned}
\mu(y_{T_i}, q_{T_i}, q_{T_{i+1}}) &= y_{T_i} + \frac{\rho}{v}(e^{q_{T_{i+1}}} - e^{q_{T_i}}) - \frac{\rho\mu_i}{\nu}E_{\mathbb{F}^{n+1}}[\Delta\widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}] \\
\eta^2(q_{T_i}, q_{T_{i+1}}) &= (\frac{\rho\mu_i}{\nu})^2(E_{\mathbb{F}^{n+1}}[(\Delta\widetilde{V}_{T_i})^2|q_{T_i}, q_{T_{i+1}}] - (E_{\mathbb{F}^{n+1}}[\Delta\widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}])^2) \\
&\quad + (1-\rho^2)E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}|q_{T_i}, q_{T_{i+1}}]
\end{aligned}$$

DD-SABR:

$$\begin{aligned}
\mu(y_{T_i}, q_{T_i}, q_{T_{i+1}}) &= y_{T_i} + \frac{\omega_i\rho}{v}(e^{q_{T_{i+1}}} - e^{q_{T_i}}) - \frac{\omega_i\rho\mu_i}{\nu}E_{\mathbb{F}^{n+1}}[\Delta\widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}] \\
&\quad - \frac{\omega_i^2}{2}E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}|q_{T_i}, q_{T_{i+1}}] \\
\eta^2(y_{T_i}, q_{T_i}, q_{T_{i+1}}) &= (\frac{\omega_i\rho\mu_i}{\nu})^2(E_{\mathbb{F}^{n+1}}[(\Delta\widetilde{V}_{T_i})^2|q_{T_i}, q_{T_{i+1}}] - (E_{\mathbb{F}^{n+1}}[\Delta\widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}])^2) \\
&\quad + \frac{\omega_i^4}{4}(E_{\mathbb{F}^{n+1}}[(\Delta V_{T_i})^2|q_{T_i}, q_{T_{i+1}}] - (E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}|q_{T_i}, q_{T_{i+1}}])^2) \\
&\quad + \omega_i^2(1-\rho^2)E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}|q_{T_i}, q_{T_{i+1}}] \\
&\quad + \frac{\omega_i^3\rho\mu_i}{\nu}(E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}\Delta\widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}] \\
&\quad - E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}|q_{T_i}, q_{T_{i+1}}]E_{\mathbb{F}^{n+1}}[\Delta\widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}])
\end{aligned}$$

with

$$E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}|q_{T_i}, q_{T_{i+1}}] = \frac{e^{2q_{T_i}}\sqrt{\alpha_i}}{2v}\frac{[\Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + v\sqrt{\alpha_i}) - \Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} - v\sqrt{\alpha_i})]}{\phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + v\sqrt{\alpha_i})}$$

$$\begin{aligned}E_{\mathbb{F}^{n+1}}[(\Delta V_{T_i})^2|q_{T_i}, q_{T_{i+1}}] &= -\frac{e^{4q_{T_i}}\sqrt{\alpha_i}}{4v^3}(1 + e^{2(q_{T_{i+1}}-q_{T_i})}) \\ &\times \frac{[\Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + v\sqrt{\alpha_i}) - \Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} - v\sqrt{\alpha_i})]}{\phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + v\sqrt{\alpha_i})} \\ &+ \frac{e^{4q_{T_i}}\sqrt{\alpha_i}}{4v^3}\frac{[\Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + 2v\sqrt{\alpha_i}) - \Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} - 2v\sqrt{\alpha_i})]}{\phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + 2v\sqrt{\alpha_i})}\end{aligned}$$

$$E_{\mathbb{F}^{n+1}}[\Delta \widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}] = \frac{e^{q_{T_i}}\sqrt{\alpha_i}}{\nu}\frac{[\Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + \frac{1}{2}\nu\sqrt{\alpha_i}) - \Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} - \frac{1}{2}\nu\sqrt{\alpha_i})]}{\phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + \frac{1}{2}\nu\sqrt{\alpha_i})}$$

$$\begin{aligned}E_{\mathbb{F}^{n+1}}[(\Delta \widetilde{V}_{T_i})^2|q_{T_i}, q_{T_{i+1}}] &= -\frac{2e^{2q_{T_i}}\sqrt{\alpha_i}}{\nu^3}(1 + e^{(q_{T_{i+1}}-q_{T_i})}) \\ &\times \frac{[\Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + \frac{1}{2}\nu\sqrt{\alpha_i}) - \Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} - \frac{1}{2}\nu\sqrt{\alpha_i})]}{\phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + \frac{1}{2}\nu\sqrt{\alpha_i})} \\ &+ \frac{2e^{2q_{T_i}}\sqrt{\alpha_i}}{\nu^3}\frac{[\Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + \nu\sqrt{\alpha_i}) - \Phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} - \nu\sqrt{\alpha_i})]}{\phi(\frac{q_{T_{i+1}}-q_{T_i}}{v\sqrt{\alpha_i}} + \nu\sqrt{\alpha_i})},\end{aligned}$$

and

$$\begin{aligned}E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}\Delta \widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}] &= -\frac{e^{3q_{T_i}}\sqrt{\alpha_i}}{3v^3}[e^{-3v^2\alpha_i}\frac{\Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + \frac{3}{2}\nu\sqrt{\alpha_i}) - \Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} - \frac{1}{2}\nu\sqrt{\alpha_i})}{\phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + \frac{5}{2}\nu\sqrt{\alpha_i})} \\ &- \frac{\Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + 2\nu\sqrt{\alpha_i}) - \Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} - \nu\sqrt{\alpha_i})}{\phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + \frac{3}{2}\nu\sqrt{\alpha_i})} + \frac{\Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + \nu\sqrt{\alpha_i}) - \Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}})}{\phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + \nu\sqrt{\alpha_i})} \\ &+ e^{-\frac{3}{2}v^2\alpha_i}\frac{\Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + \nu\sqrt{\alpha_i}) - \Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} - \nu\sqrt{\alpha_i})}{\phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + 2\nu\sqrt{\alpha_i})} \\ &- \frac{\Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + \frac{3}{2}\nu\sqrt{\alpha_i}) - \Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} - \frac{3}{2}\nu\sqrt{\alpha_i})}{\phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + \frac{3}{2}\nu\sqrt{\alpha_i})} + \frac{\Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + \frac{1}{2}\nu\sqrt{\alpha_i}) - \Phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} - \frac{1}{2}\nu\sqrt{\alpha_i})}{\phi(\frac{\Delta q_{T_i}}{v\sqrt{\alpha_i}} + \frac{1}{2}\nu\sqrt{\alpha_i})}],\end{aligned}$$

where $\Delta V_{T_i} := \int_{T_i}^{T_{i+1}} \sigma_t^2 dt$, $\Delta \widetilde{V}_{T_i} := \int_{T_i}^{T_{i+1}} \sigma_t dt$ *and* $\Delta q_{T_i} := q_{T_{i+1}} - q_{T_i}$. $\phi(\cdot)$ *is the standard Normal density function and* $\Phi(\cdot)$ *is the standard Normal cumulative distribution function.*

Proof. See Appendix 3.D. □

Having found the function $\Xi_{j,d}^{i+1}$, we know from equation (3.35) that finding the one step conditional moments $\Theta_{j,k,d,l}^{i+1}$ is equivalent to finding the conditional expectation

$$E_{\mathbb{F}^{n+1}}[q_{T_{i+1}}^{l}\mathbf{I}_{\{q_{T_{i+1}}\in[q_{i+1,k},q_{i+1,k+1})\}}b_{i+1,j'}(q_{T_{i+1}})|y_{T_i},q_{T_i}]. \tag{3.36}$$

To calculate it we note that the basis function $b_{i+1,j'}(\cdot)$ is a piecewise polynomial function in the form of

$$b_{i+1,j'}(q) := \sum_{u=1}^{m-1}\sum_{r=0}^{M} b_{i+1,j',u,r}^{q} q^{r}\mathbf{I}_{\{q\in[q_{i,u},q_{i,u+1})\}} \tag{3.37}$$

Then we insert the expression (3.37) of basis function $b_{i+1,j'}(q_{T_{i+1}})$ into equation (3.36), and we have that

$$\begin{aligned}
&E_{\mathbb{F}^{n+1}}[q_{T_{i+1}}^{l}\mathbf{I}_{\{q_{T_{i+1}}\in[q_{i+1,k},q_{i+1,k+1})\}}b_{i+1,j'}(q_{T_{i+1}})|y_{T_i},q_{T_i}]\\
=&\sum_{r=0}^{M} b_{i+1,j',u,r}^{q}E_{\mathbb{F}^{n+1}}[q_{T_{i+1}}^{l+r}\mathbf{I}_{\{q_{T_{i+1}}\in[q_{i+1,k},q_{i+1,k+1})\}}|y_{T_i},q_{T_i}]\\
=&\sum_{r=0}^{M} b_{i+1,j',u,r}^{q}E_{\mathbb{F}^{n+1}}[q_{T_{i+1}}^{l+r}\mathbf{I}_{\{q_{T_{i+1}}\in[q_{i+1,k},q_{i+1,k+1})\}}|q_{T_i}],
\end{aligned} \tag{3.38}$$

where the last equation follows from the observation that the increment $B_{T_{i+1}}^{n+1}-B_{T_i}^{n+1}$ is independent of the filtration $\mathcal{F}_{T_i}$ generated by the Brownian motion $(B^{n+1},\widehat{W}^{n+1})$. As a result the conditional expectation equation (3.36) can be evaluated by numerical integration since $q_{T_{i+1}}|q_{T_i}$ is a Gaussian distribution with mean $q_{T_i}+\mu_i\alpha_i-\frac{1}{2}v^2\alpha_i$ and variance $v^2\alpha_i$.

Finally we combine equation (3.38) and function $\Xi_{j,d}^{i+1}$ into the expression (3.35) for the one step conditional moments $\Theta_{j,k,d,l}^{i+1}$, and we have that

$$\begin{aligned}
&\Theta_{j,k,d,l}^{i+1}(y_{T_i},q_{T_i}) \qquad\qquad (3.39)\\
&\approx \sum_{j'=1}^{m}\Xi_{j,d}^{i+1}(y_{T_i},q_{T_i},q_{i+1,j'})\sum_{r=0}^{M} b_{i+1,j',u,r}^{q}E_{\mathbb{F}^{n+1}}[q_{T_{i+1}}^{l+r}\mathbf{I}_{\{q_{T_{i+1}}\in[q_{i+1,k},q_{i+1,k+1})\}}|q_{T_i}].
\end{aligned}$$

Note that the approximation approach we used above for the calculation of the one step conditional moments was specific to the SABR style driver. We will discuss a general driving process case later in this section.

Pre-model Specification

Having specified the stochastic volatility driving process (y, q), now we discuss the assumption (A.4). Remember that in order to specify a one-dimensional LIBOR MFM, we need to assume that the LIBORs at their setting dates are a monotonic increasing function of the one-dimensional driving process. When it comes to a multi-dimensional MFM, the univariate and monotonicity properties are lost. This is why we introduce a pre-model (3.16). The choice of a pre-model depends on the specification of the driver. In particular, as we dicussed, when a two-dimensional driving process is specified in the form of (3.17), the pre-model is chosen to be some strictly increasing function of the linear combination of the components of the driver; see (3.18). In this case the two-dimensional driver can be viewed as representing the level and skew of interest rates.

Now we return to the stochastic volatility MFM. From the specification of the stochastic volatility driving process (y, q) we can see that the driver is specified in a multiplicative way. The process y represents the level of rates while carrying the stochastic volatility process q with it. By doing so we achieve a more realistic and desirable dynamics of the model than a one-dimensional MFM where the driver is also used for capturing the level of rates but without stochastic volatility. As a result it is natural to make the assumption (A.4). In this case one can think of the variable y_{T_i} itself as a pre-model at time T_i for $i = 1, ..., n$.

We conclude this subsection with some remarks. We note that the one step conditional moments are evaluated by applying some specific approximation to the SABR model. In principle the SABR model is not the only choice for a driving Markov process. For example one could add mean reversion to the stochastic volatility process; see [35] and [25]. But the approximation discussed is specific to the SABR type model and it may not apply to other stochastic volatility models. In this case we need to find another way to evaluate the one step conditional moments. There may exist a solution or some specific approximation to the stochastic volatility model we choose so that the one step conditional moments can be solved or approximated. Otherwise we can resort to discretising the driving Markov process in time. In particular consider the following driving Markov process

$$\begin{aligned} dx_t &= \mu_x(t, x_t)dt + z_t\sigma_x(t, x_t)dW_t^{n+1} \\ dz_t &= \mu_z(t, z_t)dt + \sigma_z(t, z_t)dB_t^{n+1} \\ dB_t^{n+1}dW_t^{n+1} &= \rho dt \qquad \rho \in [-1, 1]. \end{aligned}$$

For simplicity we present the Euler scheme of the driving process

$$
\begin{aligned}
x_{t_{i+1}} &= x_{t_i} + \mu_x(t_i, x_{t_i})\Delta t + z_{t_i}\sigma_x(t_i, x_{t_i})\Delta W_{t_i}^{n+1} \\
z_{t_{i+1}} &= z_{t_i} + \mu_z(t_i, z_{t_i})\Delta t + \sigma_z(t_i, z_{t_i})\Delta B_{t_i}^{n+1} \\
E_{\mathbb{F}^{n+1}}[\Delta B_{t_i}^{n+1}\Delta W_{t_i}^{n+1}] &= \rho\Delta t \qquad \rho \in [-1, 1],
\end{aligned}
$$

where $\Delta t := t_{i+1} - t_i$, $\Delta W_{t_i}^{n+1} := W_{t_{i+1}}^{n+1} - W_{t_i}^{n+1}$ and $\Delta B_{t_i}^{n+1} := B_{t_{i+1}}^{n+1} - B_{t_i}^{n+1}$. Other discretization schemes (e.g. Predictor-corrector Scheme) could also be applied if necessary. In this case the joint distribution $(x_{t_{i+1}}, z_{t_{i+1}}|x_{t_i}, z_{t_i})$ is Gaussian so that the one step conditional moments can be evaluated. However we have to be aware that the discretization gives an inaccurate approximation to the distribution $(x_{t_{i+1}}, z_{t_{i+1}}|x_{t_i}, z_{t_i})$. In order to improve the approximation we can make the discretization smoother by using more time steps between T_i and T_{i+1} on the date structure. In particular we let $\Delta t = \frac{1}{\delta}(T_{i+1} - T_i)$. This means that we have δ time steps between each time $[T_i, T_{i+1}]$ on the date structure. Now if we take δ to be, for example, 50 the approximation should be very accurate. However the price to pay for the accuracy is the computational cost, which is the main reason why we did not use this method to approximate the one step conditional moments.

3.4 Conclusion

In this chapter we reviewed the algorithm to specify a MFM with a Gaussian driver. This algorithm relies heavily on the Gaussian assumption so that it cannot apply to a MFM with a driving process away from Gaussian. There is substantial empirical evidence supporting the introduction of a stochastic volatility model. This motivates us to develop a stochastic volatility MFM. In order to implement this model, which has a non-Gaussian driver, we developed a general algorithm by working with basis functions and conditional moments of the driving Markov process. This algorithm is not specific and could be modified to apply to all one- and multi-dimensional MFMs with various types of driving process. From a data driven study which used market data to identify a SABR type model as an appropriate choice for the level of interest rates, we chose a SABR type model as the driver of our stochastic volatility MFM. With this choice we specified a pre-model and developed an approximation to evaluate conditional moments of the SABR driver which served as building blocks for the algorithm.

3.A Appendix: Equivalence of Markov-functional models

In this appendix we will show that under some conditions two distinct driving Markov processes could lead to the same MFM. Let us first make clear what we mean by saying that two MFMs are the same. Let us consider two MFMs with two distinct driving processes $x = (x^1, x^2, ..., x^d)$ and $y = (y^1, y^2, ..., y^d)$, which are both assumed to be continuous diffusion processes, under the terminal measure on the tenor structure (2.1). Let $\{\mathcal{F}_t\}_{0\leq t\leq T^*}$ be the natural filtration generated by Brownian motions under the terminal measure. The above two MFMs are said to be the same if for any continuous and bounded function g^i we have that

$$\begin{aligned} &E_{\mathbb{F}^{n+1}}[g^i(D^x_{T_iT_{i+1}}(x_{T_i}), D^x_{T_iT_{i+2}}(x_{T_i}), ..., D^x_{T_iT_{n+1}}(x_{T_i}))|\mathcal{F}_{T_{i-1}}] \\ =\ &E_{\mathbb{F}^{n+1}}[g^i(D^y_{T_iT_{i+1}}(y_{T_i}), D^y_{T_iT_{i+2}}(y_{T_i}), ..., D^y_{T_iT_{n+1}}(y_{T_i}))|\mathcal{F}_{T_{i-1}}], \end{aligned} \tag{3.40}$$

where $D^x_{T_iT_j}(\cdot)$ and $D^y_{T_iT_j}(\cdot)$, $1 \leq i < j \leq n+1$, are functional forms of ZCBs under the two MFMs respectively.

Let us provide insight into equation (3.40). Since interest rates, such as LIBORs and swap rates, can be expressed in terms of ZCBs, equation (3.40) implies that, conditional on the σ-algebra $\mathcal{F}_{T_{i-1}}$, the distributions of interest rates at time T_i are the same for the two MFMs for $i = 1, ..., n$. Furthermore equation (3.40) also implies that the value of vanilla options, such as caplets and swaptions, are the same for the two MFMs. In what follows we show the equivalence of MFMs.

Proposition 3. *Consider two LIBOR MFMs where we take the following two distinct d-dimensional diffusion processes with continuous marginal distributions*

$$x = (x^1, x^2, ..., x^d)$$

and

$$y = (y^1, y^2, ..., y^d)$$

as driving Markov processes under the terminal measure. These two MFMs are the same if they satisfy:
1. Any component of y is a strictly increasing function of the corresponding component of x i.e. $y^i_t = f_i(t, x^i_t)$ for some strictly increasing function $f_i : [0, \infty) \times \mathbb{R} \to \mathbb{R}$ with respect to x for $t \geq 0$ and $i = 1, ..., d$.
2. The pre-model $\widehat{L}^{i,y}_{T_i}(y_{T_i})$ of y is a monotonic increasing function of the pre-model

$\widehat{L}^{i,x}_{T_i}(x_{T_i})$ *i.e.* $\widehat{L}^{i,y}_{T_i}(y_{T_i}) = h_i(T_i, \widehat{L}^{i,x}_{T_i}(x_{T_i}))$ *for some monotonic increasing function* $h_i : [0,\infty) \times \mathbb{R} \to \mathbb{R}$ *for* $i = 1, ..., n$.
3. The two MFMs are calibrated to the same input prices of digital caplets.

Proof. To prove the two MFMs are the same we need to show (3.40). The result of this proof is stronger than two MFMs being the same in the distributional sense as we have an actual functional relationship between them.

Note that for any $1 \leq i < j \leq n+1$, from the first condition we obtain that

$$\begin{aligned}\widetilde{D}^x_{T_iT_j}(x_{T_i}) &:= D^y_{T_iT_j}(y_{T_i}) \\ &= D^y_{T_iT_j}(f_1(T_i, x^1_{T_i}), ..., f_d(T_i, x^d_{T_i})).\end{aligned}$$

We will prove that

$$\widetilde{D}^x_{T_iT_j}(\cdot) = D^x_{T_iT_j}(\cdot).$$

That is the model we get from working with the process y and then writing it in terms of x, $\widetilde{D}^x_{T_iT_j}(\cdot)$ is the same model we would get if we started by setting up the MFM using x, $D^x_{T_iT_j}(\cdot)$. The proof follows the construction of the MFM, so proceeds by backwards induction on time T_i. Note that in the proof we will just focus on some key steps and a complete procedure of construction can be found in Section 3.2.3.

To begin, at time T_{n+1} by definition we have that

$$\widetilde{D}^x_{T_{n+1}T_{n+1}}(x_{T_{n+1}}) = D^x_{T_{n+1}T_{n+1}}(x_{T_{n+1}}) = 1.$$

Suppose the result is true for all time $T_{i+1}, ..., T_{n+1}$. We complete the proof by showing that the result is also true at time T_i.

At time T_i let us first consider the numeraire-rebased ZCB $\frac{D_{T_iT_{i+1}}}{D_{T_iT_{n+1}}}$. It follows

from the martingale property that

$$\begin{aligned}
\frac{\widetilde{D}^x_{T_iT_{i+1}}(x_{T_i})}{\widetilde{D}^x_{T_iT_{n+1}}(x_{T_i})} &:= \frac{D^y_{T_iT_{i+1}}(y_{T_i})}{D^y_{T_iT_{n+1}}(y_{T_i})} \\
&= E_{\mathbb{F}^{n+1}}[\frac{1}{D^y_{T_{i+1}T_{n+1}}(y_{T_{i+1}})}|y_{T_i}] \\
&= E_{\mathbb{F}^{n+1}}[\frac{1}{\widetilde{D}^x_{T_{i+1}T_{n+1}}(x_{T_{i+1}})}|y_{T_i}] \\
&= E_{\mathbb{F}^{n+1}}[\frac{1}{D^x_{T_{i+1}T_{n+1}}(x_{T_{i+1}})}|x_{T_i}] \qquad (3.41) \\
&= \frac{D^x_{T_iT_{i+1}}(x_{T_i})}{D^x_{T_iT_{n+1}}(x_{T_i})} \qquad (3.42)
\end{aligned}$$

where equation (3.41) is true since the component of y is a strictly increasing function of the corresponding component of x so that knowing y_{T_i} is equivalent to knowing x_{T_i}.

The second condition tells us that the pre-models $\widehat{L}^{i,x}_{T_i}(x_{T_i})$ and $\widehat{L}^{i,y}_{T_i}(y_{T_i})$ are chosen such that

$$\widehat{L}^{i,y}_{T_i}(y_{T_i}) = \widehat{L}^{i,y}_{T_i}(f_1(T_i, x^1_{T_i}), ..., f_d(T_i, x^d_{T_i})) = h_i(T_i, \widehat{L}^{i,x}_{T_i}(x_{T_i})) \qquad (3.43)$$

for some monotonic increasing function $h_i : [0, \infty) \times \mathbb{R} \to \mathbb{R}$. We choose values $x^* = (x^{*,1}, x^{*,2}, ..., x^{*,d})$ of x_{T_i} and the corresponding

$$y^* = (f_1(T_i, x^{*,1}), f_2(T_i, x^{*,2}), ..., f_d(T_i, x^{*,d}))$$

of y_{T_i}. From (3.42) and (3.43) we have that

$$\begin{aligned}
J^{i,y}_0(y^*) &:= D_{0T_{n+1}} E_{\mathbb{F}^{n+1}}[\frac{D^y_{T_iT_{i+1}}(y_{T_i})}{D^y_{T_iT_{n+1}}(y_{T_i})} \mathbf{I}_{\{\widehat{L}^{i,y}_{T_i}(y_{T_i}) > \widehat{L}^{i,y}_{T_i}(y^*)\}}] \\
&= D_{0T_{n+1}} E_{\mathbb{F}^{n+1}}[\frac{D^x_{T_iT_{i+1}}(x_{T_i})}{D^x_{T_iT_{n+1}}(x_{T_i})} \mathbf{I}_{\{\widehat{L}^{i,x}_{T_i}(x_{T_i}) > \widehat{L}^{i,x}_{T_i}(x^*)\}}] \\
&=: J^{i,x}_0(x^*)
\end{aligned}$$

Furthermore since we feed in the same input prices of digital caplets (see condition 3), we have that

$$L^{i,x}_{T_i}(x_{T_i}) = L^{i,y}_{T_i}(y_{T_i}),$$

and therefore

$$\widetilde{D}^x_{T_iT_{n+1}}(x_{T_i}) := D^y_{T_iT_{n+1}}(y_{T_i}) = D^x_{T_iT_{n+1}}(x_{T_i}).$$

Finally ZCBs are obtained by the martingale property

$$\begin{aligned}
\widetilde{D}^x_{T_iT_j}(x_{T_i}) &:= D^y_{T_iT_j}(y_{T_i}) \\
&= D^y_{T_iT_{n+1}}(y_{T_i}) E_{\mathbb{F}^{n+1}}[\frac{1}{D^y_{T_jT_{n+1}}(y_{T_j})}|y_{T_i}] \\
&= D^x_{T_iT_{n+1}}(x_{T_i}) E_{\mathbb{F}^{n+1}}[\frac{1}{D^x_{T_jT_{n+1}}(x_{T_j})}|x_{T_i}] \\
&= D^x_{T_iT_j}(x_{T_i})
\end{aligned}$$

for $i < j \leq n$. Once we have proved that $\widetilde{D}^x_{T_iT_j}(\cdot) = D^x_{T_iT_j}(\cdot)$ for any $1 \leq i < j \leq n+1$, equation (3.40) is immediate.

□

Note that we are not aiming for the best result possibly. There could exist some weaker assumptions to show that two MFMs with different drivers are the same. We also note that the above proposition applies to the one-dimensional case. In particular for the one-dimensional case a pre-model can be chosen to be the driver itself. As a result if two MFMs satisfy the above condition 1 they will satisfy condition 2 automatically for a one-dimensional MFM.

3.B Appendix: Basis functions

Before we discuss basis functions we introduce polynomial interpolation. Given $M+1$ distinct data points $(x_i)_{i=1}^{M+1}$ and corresponding $M+1$ values $(y_i)_{i=1}^{M+1}$, we can find a unique interpolation polynomial p up to degree M:

$$p(x) = \sum_{i=1}^{M+1} y_i l_i(x), \tag{3.44}$$

where the *Lagrange Basis Polynomials* $l_i(x)$ are defined by

$$l_i(x) := \prod_{k=1;k\neq i}^{M+1} \frac{x - x_k}{x_i - x_k}.$$

The construction of the above polynomial implies that $l_i(x_j) = \delta_{i,j}$ so that

$$p(x_i) = y_i.$$

Next we will make use of the interpolation polynomial to construct basis functions.

Let $x_1 < x_2 < ... < x_m$ be a partition of the interval $[x_1, x_m]$. The one-dimensional basis function $b_j : [x_1, x_m] \to \mathbb{R}$, $j = 1, ..., m$, is defined as a continuous function with repect to the partition of the interval $[x_1, x_m]$ such that

$$b_j(x_k) = \delta_{j,k}$$

for $k = 1, ..., m$. For the sake of efficiency, in our numerical implementation we take the basis function as a piecewise polynomial up to degree $M \geq 1$

$$b_j(x) = \sum_{k=1}^{m-1} \sum_{d=0}^{M} b_{j,k,d} x^d \mathbf{I}_{\{x \in [x_k, x_{k+1})\}}$$

satisfying

$$b_j(x_k) = \delta_{j,k}.$$

Note that the basis function can be chosen in other forms such as a Fourier basis.

Next we present the construction of the basis function and determine the coefficients $b_{j,k,d}$ of the piecewise polynomial. It is well known that given $M+1$ distinct points, there exists a unique polynomial function up to degree M and therefore we have that $m \geq M+1$. To calculate the coefficient $b_{j,k,d}$ of the basis

function b_j for the interval $[x_k, x_{k+1})$ uniquely, we make use of $M+1$ grid nodes surrounding the interval i.e.

$$x_{k-[\frac{M}{2}-1]}, ..., x_{k+[\frac{M+1}{2}]}, \tag{3.45}$$

where $[\cdot]$ is the round function. Now we are given $M+1$ distinct points

$$x_{k-[\frac{M}{2}-1]}, ..., x_{k+[\frac{M+1}{2}]}$$

and the corresponding $M+1$ values

$$b_j(x_{k-[\frac{M}{2}-1]}), ..., b_j(x_{k+[\frac{M+1}{2}]}).$$

It follows from equation (3.44) that we determine a polynomial function p up to degree M uniquely:

$$p(x) = \sum_{i=k-[\frac{M}{2}-1]}^{k+[\frac{M+1}{2}]} b_j(x_i) l_i(x) \tag{3.46}$$

where $l_i(x)$ is defined by

$$l_i(x) = \prod_{\substack{s=k-[\frac{M}{2}-1] \\ s \neq i}}^{k+[\frac{M+1}{2}]} \frac{x - x_s}{x_i - x_s}.$$

The coefficients $b_{j,k,d}$ for the interval $[x_k, x_{k+1})$ are therefore chosen such that

$$p(x) = \sum_{d=0}^{M} b_{j,k,d} x^d.$$

Following the above procedure we can fix the coefficients $b_{j,k,d}$ of the basis function b_j for $k = 1, ..., m-1$ and $d = 0, ..., M$ so that the basis function b_j can be specified.

Note that there is an issue about the construction of the basis function. We find the coefficients of polynomial for the interval $[x_k, x_{k+1})$ by using $M+1$ grid points surrounding the interval (see (3.45)). However some of these points could be out of the defined interval $[x_1, x_m]$ i.e. $k - [\frac{M}{2} - 1] < 1$ or $k + [\frac{M+1}{2}] > m$. This means that we do not have enough nodes to determine a polynomial function up to degree M. In this case we can achieve a polynomial function up to a lower degree.

In the numerical implementation it is efficient to construct the basis function.

To show this let us see equation (3.46) that $b_j(x_i)$ is non-zero only when $i = j$. Based on this observation we conclude that the coefficient $b_{j,k,d} = 0$ for interval $[x_k, x_{k+1})$ when grid point x_j is outside the interval $[x_{k-[\frac{M}{2}-1]}, x_{k+[\frac{M+1}{2}]}]$. This means that it is sufficient to fix the coefficient within the interval $[x_{j-[\frac{M+1}{2}]}, x_{j+[\frac{M}{2}]}]$ since coefficients are all zero outside. Therefore the basis function can be simplified as

$$b_j(x) = \sum_{k=j-[\frac{M+1}{2}]}^{j+[\frac{M}{2}-1]} \sum_{d=0}^{M} b_{j,k,d} x^d I_{\{x\in[x_k,x_{k+1})\}}.$$

In particular in the numerical implementation of MFMs we usually set $M = 5$ so that

$$b_j(x) = \sum_{k=j-3}^{j+2} \sum_{d=0}^{5} b_{j,k,d} x^d I_{\{x\in[x_k,x_{k+1})\}}.$$

The multi-dimensional basis function is a straightforward extension. Let $b_{j^n} : [x_1^n, x_m^n] \to \mathbb{R}$, $j^n = 1, ..., m$, $n = 1, ..., d$, be one-dimensional basis function with respect to the partition $(x_j^n)_{j=1}^m$ of the interval $[x_1^n, x_m^n]$. The d-dimensional basis function $b_{j^1,...,j^d} : [x_1^1, x_m^1] \times ... \times [x_1^d, x_m^d] \to \mathbb{R}$, $j^1, ..., j^d = 1, ..., m$, is defined as a continuous function with repect to the partition of the interval product such that

$$b_{j^1,...,j^d}(x_{k^1}^1, ..., x_{k^d}^d) = \prod_{n=1}^{d} \delta_{j^n,k^n}$$

for $k^n = 1, ..., m$ and $n = 1, ..., d$. The d-dimensional basis function can be represented as the product of the corresponding d individual one-dimensional basis functions:

$$b_{j^1,...,j^d}(x^1, ..., x^d) = \prod_{n=1}^{d} b_{j^n}(x^n).$$

In particular when $d = 2$, the two-dimensional basis function $b_{j,k} : [x_1, x_m] \times [y_1, y_m] \to \mathbb{R}$, $j, k = 1, ..., m$ w.r.t the two-dimensional partition $(x_i, y_i)_{i=1}^m$ can be expressed in the form of the product of their corresponding individual one-dimensional basis functions b_j and b_k:

$$\begin{aligned} b_{j,k}(x, y) &= b_j(x) \times b_k(y) \\ &= \sum_{j'k'=1}^{m-1} \sum_{d,l=0}^{M} b_{j,j',d}^x b_{k,k',l}^y x^d y^l \mathbf{I}_{\{x\in[x_{j'},x_{j'+1})\}\cap\{y\in[y_{k'},y_{k'+1})\}} \end{aligned}$$

satisfying

$$b_{j,k}(x_u, y_v) = \delta_{j,u} \times \delta_{k,v}.$$

3.C Appendix: Proof of Proposition 1

Let $y_{i,1} < ... < y_{i,m}$ and $q_{i,1} < ... < q_{i,m}$ be partition of the driving Markov process (y_{T_i}, q_{T_i}) at time T_i for $i = 1, ..., n$. The initial value of the process (y, q) at time T_0 is given by constants (y_0, q_0). Suppose we are given the two-dimensional basis function $b_{i,j,k}$ with respect to the above partition:

$$b_{i,j,k}(y,q) = \sum_{j'k'=1}^{m-1} \sum_{d,l=0}^{M} b^{y}_{i,j,j',d} b^{q}_{i,k,k',l} y^d q^l \mathbf{I}_{\{y \in [y_{i,j'}, y_{i,j'+1})\} \cap \{q \in [q_{i,k'}, q_{i,k'+1})\}} \tag{3.47}$$

for $j, k = 1, ..., m$. Suppose that the one step conditional moments

$$\begin{aligned}&\Theta^{i}_{j,k,d,l}(y_{i-1,u}, q_{i-1,v}) \\ &:= E_{\mathbb{F}^{n+1}}[y^d_{T_i} q^l_{T_i} \mathbf{I}_{\{y_{T_i} \in [y_{i,j}, y_{i,j+1})\} \cap \{q_{T_i} \in [q_{i,k}, q_{i,k+1})\}} | y_{T_{i-1}} = y_{i-1,u}, q_{T_{i-1}} = q_{i-1,v}]\end{aligned}$$

are given for $i = 1, ..., n$; $j, k = 1, ..., m-1$; $d, l = 0, ..., M$ and $u, v = 1, ..., m$.

We now evaluate the following three (conditional) expectations:

$$\Delta^{i}_{j,k}(y_{i-1,u}, q_{i-1,v}) := E_{\mathbb{F}^{n+1}}[b_{i,j,k}(y_{T_i}, q_{T_i}) | y_{T_{i-1}} = y_{i-1,u}, q_{T_{i-1}} = q_{i-1,v}], \tag{3.48}$$

$$E_{i,j,k} := E_{\mathbb{F}^{n+1}}[b_{i,j,k}(y_{T_i}, q_{T_i})], \tag{3.49}$$

$$\Gamma^{i}_{j,k,u} := E_{\mathbb{F}^{n+1}}[b_{i,j,k}(y_{T_i}, q_{T_i}) \mathbf{I}_{\{y_{T_i} \in [y_{i,u}, y_{i,u+1})\}}]. \tag{3.50}$$

for $i = 1, ..., n$ and $j, k, u, v = 1, ..., m$.

1. **Calculate $\boldsymbol{\Delta^{i}_{j,k}}$** : Inserting equation (3.47) into equation (3.48) leads to

$$\Delta^{i}_{j,k}(y_{i-1,u}, q_{i-1,v}) = \sum_{j'=1}^{m-1} \sum_{d=0}^{M} b^{y}_{i,j,j',d} \sum_{k'=1}^{m-1} \sum_{l=0}^{M} b^{q}_{i,k,k',l} \Theta^{i}_{j',k',d,l}(y_{i-1,u}, q_{i-1,v}),$$

   for $i = 1, ..., n$ and $j, k, u, v = 1, ..., m$.

2. **Calculate $\mathbf{E_{i,j,k}}$** : The algorithm for finding $E_{i,j,k}$ works forward iteratively from time T_1. At time T_1 we have that

$$\begin{aligned}E_{1,j,k} &:= E_{\mathbb{F}^{n+1}}[b_{1,j,k}(y_{T_1}, q_{T_1})] \\ &= E_{\mathbb{F}^{n+1}}[b_{1,j,k}(y_{T_1}, q_{T_1}) | y_0, q_0] \\ &= \Delta^{1}_{j,k}(y_0, q_0),\end{aligned}$$

   for $j, k = 1, ..., m$.

Suppose we reach time T_i now, having already found $E_{i-1,j,k}$ for $j,k = 1,...,m$. $E_{i,j,k}$ can be found by applying the tower property:

$$\begin{aligned} E_{i,j,k} &:= E_{\mathbb{F}^{n+1}}[b_{i,j,k}(y_{T_i}, q_{T_i})] \\ &= E_{\mathbb{F}^{n+1}}[E_{\mathbb{F}^{n+1}}(b_{i,j,k}(y_{T_i}, q_{T_i})|y_{T_{i-1}}, q_{T_{i-1}})] \\ &= E_{\mathbb{F}^{n+1}}[\Delta^i_{j,k}(y_{T_{i-1}}, q_{T_{i-1}})]. \end{aligned} \tag{3.51}$$

We note that the smooth function $\Delta^i_{j,k}(y_{T_{i-1}}, q_{T_{i-1}})$ can be approximated by a piecewise polynomial in terms of the basis function:

$$\Delta^i_{j,k}(y_{T_{i-1}}, q_{T_{i-1}}) \approx \sum_{u,v}^{m} \Delta^i_{j,k}(y_{i-1,u}, q_{i-1,v}) b_{i-1,u,v}(y_{T_{i-1}}, q_{T_{i-1}}). \tag{3.52}$$

Inserting equation (3.52) into equation (3.51) we have that

$$\begin{aligned} E_{i,j,k} &\approx E_{\mathbb{F}^{n+1}}[\sum_{u,v}^{m} \Delta^i_{j,k}(y_{i-1,u}, q_{i-1,v}) b_{i-1,u,v}(y_{T_{i-1}}, q_{T_{i-1}})] \\ &= \sum_{u,v}^{m} \Delta^i_{j,k}(y_{i-1,u}, q_{i-1,v}) E_{i-1,u,v}, \end{aligned}$$

for $j,k = 1,...,m$.

3. **Calculate $\mathbf{\Gamma^i_{j,k,u}}$** : It follows from equation (3.47) of basis function that

$$\begin{aligned} \Gamma^i_{j,k,u} &:= E_{\mathbb{F}^{n+1}}[b_{i,j,k}(y_{T_i}, q_{T_i}) \mathbf{I}_{\{y_{T_i} \in [y_{i,u}, y_{i,u+1})\}}] \\ &= E_{\mathbb{F}^{n+1}}[\sum_{k'=1}^{m-1} \sum_{d,l=0}^{M} b^y_{i,j,u,d} b^q_{i,k,k',l} y^d_{T_i} q^l_{T_i} \mathbf{I}_{\{y_{T_i} \in [y_{i,u}, y_{i,u+1})\} \cap \{q_{T_i} \in [q_{i,k'}, q_{i,k'+1})\}}] \\ &= \sum_{k'=1}^{m-1} \sum_{d,l=0}^{M} b^y_{i,j,u,d} b^q_{i,k,k',l} E_{\mathbb{F}^{n+1}}[y^d_{T_i} q^l_{T_i} \mathbf{I}_{\{y_{T_i} \in [y_{i,u}, y_{i,u+1})\} \cap \{q_{T_i} \in [q_{i,k'}, q_{i,k'+1})\}}]. \end{aligned} \tag{3.53}$$

Apply the tower property to the expectation in equation (3.53) and we have

that

$$
\begin{aligned}
&E_{\mathbb{F}^{n+1}}[y_{T_i}^d q_{T_i}^l \mathbf{I}_{\{y_{T_i}\in[y_{i,u},y_{i,u+1})\}\cap\{q_{T_i}\in[q_{i,k'},q_{i,k'+1})\}}] \\
&= E_{\mathbb{F}^{n+1}}[E_{\mathbb{F}^{n+1}}(y_{T_i}^d q_{T_i}^l \mathbf{I}_{\{y_{T_i}\in[y_{i,u},y_{i,u+1})\}\cap\{q_{T_i}\in[q_{i,k'},q_{i,k'+1})\}}|y_{T_{i-1}}, q_{T_{i-1}}] \\
&= E_{\mathbb{F}^{n+1}}[\Theta^i_{u,k',d,l}(y_{T_{i-1}}, q_{T_{i-1}})] \\
&= \sum_{u',v=1}^{m} \Theta^i_{u,k',d,l}(y_{i-1,u'}, q_{i-1,v}) E_{i-1,u',v}. \qquad (3.54)
\end{aligned}
$$

Insert equation (3.54) into equation (3.53) and we obtain the result

$$
\begin{aligned}
\Gamma^i_{j,k,u} &= \sum_{k'=1}^{m-1}\sum_{d,l=0}^{M} b^y_{i,j,u,d} b^q_{i,k,k',l} \sum_{u',v=1}^{m} \Theta^i_{u,k',d,l}(y_{i-1,u'}, q_{i-1,v}) E_{i-1,u',v} \\
&= \sum_{d=0}^{M} b^y_{i,j,u,d} \sum_{k'=1}^{m-1}\sum_{l=0}^{M} b^q_{i,k,k',l} \sum_{u',v=1}^{m} \Theta^i_{u,k',d,l}(y_{i-1,u'}, q_{i-1,v}) E_{i-1,u',v},
\end{aligned}
$$

for $i = 1, ..., n$ and $j, k, u = 1, ..., m$.

3.D Appendix: Proof of Proposition 2

To calculate the conditional mean and variance of $y_{T_{i+1}}|y_{T_i}, q_{T_i}, q_{T_{i+1}}$, we first find the following conditional expectations:

1. $E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}|q_{T_i}, q_{T_{i+1}}]$
2. $E_{\mathbb{F}^{n+1}}[(\Delta V_{T_i})^2|q_{T_i}, q_{T_{i+1}}]$
3. $E_{\mathbb{F}^{n+1}}[\Delta \widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}]$
4. $E_{\mathbb{F}^{n+1}}[(\Delta \widetilde{V}_{T_i})^2|q_{T_i}, q_{T_{i+1}}]$
5. $E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}\Delta \widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}]$.

The SDE of the stochastic volatility is given by

$$d\sigma_t = \mu_t\sigma_t dt + v\sigma_t dB_t^{n+1} \qquad v > 0.$$

Conditional on $B_{T_i}^{n+1}$ and $B_{T_{i+1}}^{n+1}$, the Brownian bridge $B_t^{n+1}|B_{T_i}^{n+1}, B_{T_{i+1}}^{n+1}$, $T_i \leq t \leq T_{i+1}$, is Gaussian with mean

$$B_{T_i}^{n+1} + \frac{t - T_i}{T_{i+1} - T_i}(B_{T_{i+1}}^{n+1} - B_{T_i}^{n+1})$$

and covariance between $B_t^{n+1}|B_{T_i}^{n+1}, B_{T_{i+1}}^{n+1}$ and $B_s^{n+1}|B_{T_i}^{n+1}, B_{T_{i+1}}^{n+1}$ with $s < t$

$$\frac{(T_{i+1} - t)(s - T_i)}{T_{i+1} - T_i}.$$

Moreover by Itô's lemma, we have that:

$$B_t^{n+1} = \frac{q_t - q_0 - \int_0^t \mu_s ds + \frac{1}{2}v^2 t}{v},$$

where $q_t := ln(\sigma_t)$. Thus knowing B_t^{n+1} is equivalent to knowing q_t so that

$$(B_t^{n+1}|q_{T_i}, q_{T_{i+1}}) \equiv (B_t^{n+1}|B_{T_i}^{n+1}, B_{T_{i+1}}^{n+1})$$

is a Gaussian distribution.

Next we can calculate the conditional expectations. In particular we have that

$$E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}|q_{T_i}, q_{T_{i+1}}] := E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t^2 dt|q_{T_i}, q_{T_{i+1}}]$$
$$= \sigma_{T_i}^2 \int_{T_i}^{T_{i+1}} E_{\mathbb{F}^{n+1}}[e^{(2\mu_i - v^2)(t-T_i)+2v(B_t - B_{T_i})}|q_{T_i}, q_{T_{i+1}}]dt,$$

which can be evaluated since the Gaussian distribution of $(B_t^{n+1}|q_{T_i}, q_{T_{i+1}})$, $T_i \leq t \leq T_{i+1}$, is known. After some transformation and calculation we can obtain the result. We can also obtain the conditional expectation $E_{\mathbb{F}^{n+1}}[\Delta \widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}]$ by a similar calculation.

The conditional expectation $E_{\mathbb{F}^{n+1}}[(\Delta V_{T_i})^2|q_{T_i}, q_{T_{i+1}}]$ involves more calculation. We have that

$$\begin{aligned}
&E_{\mathbb{F}^{n+1}}[(\Delta V_{T_i})^2|q_{T_i}, q_{T_{i+1}}] \\
:=&E_{\mathbb{F}^{n+1}}[(\int_{T_i}^{T_{i+1}} \sigma_t^2 dt)^2|q_{T_i}, q_{T_{i+1}}] \\
=&E_{\mathbb{F}^{n+1}}[2\int_{T_i}^{T_{i+1}} \int_{T_i}^{t} \sigma_t^2 \sigma_s^2 ds dt|q_{T_i}, q_{T_{i+1}}] \\
=&2\sigma_{T_i}^4 \int_{T_i}^{T_{i+1}} \int_{T_i}^{t} E_{\mathbb{F}^{n+1}}[e^{(2\mu_i - v^2)(t+s-2T_i)+2v(B_t^{n+1}+B_s^{n+1}-2B_{T_i}^{n+1})}|q_{T_i}, q_{T_{i+1}}]ds dt,
\end{aligned}$$

where $(B_t^{n+1} + B_s^{n+1}|q_{T_i}, q_{T_{i+1}})$, $T_i \leq s \leq t \leq T_{i+1}$, is Gaussian with mean

$$2B_{T_i}^{n+1} + \frac{t+s-2T_i}{T_{i+1}-T_i}(B_{T_{i+1}}^{n+1} - B_{T_i}^{n+1})$$

and variance

$$\frac{(T_{i+1}-t)(t-T_i) + (T_{i+1}-s)(s-T_i) + 2(T_{i+1}-t)(s-T_i)}{T_{i+1}-T_i}.$$

Thus the conditional expectation $E_{\mathbb{F}^{n+1}}[(\Delta V_{T_i})^2|q_{T_i}, q_{T_{i+1}}]$ can be evaluated by some transformation. A similar calculation also applies to the conditional expectation $E_{\mathbb{F}^{n+1}}[(\Delta \widetilde{V}_{T_i})^2|q_{T_i}, q_{T_{i+1}}]$.

Finally we calculate $E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}\Delta\widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}]$ which can be written as

$$\begin{aligned}
&E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}\Delta\widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}]\\
:=&E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}}\sigma_t^2 dt\int_{T_i}^{T_{i+1}}\sigma_t dt|q_{T_i}, q_{T_{i+1}}]\\
=&E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}}\int_{T_i}^{t}(\sigma_t^2\sigma_s+\sigma_t\sigma_s^2)dsdt|q_{T_i}, q_{T_{i+1}}]\\
=&\sigma_{T_i}^3\int_{T_i}^{T_{i+1}}\int_{T_i}^{t}E_{\mathbb{F}^{n+1}}[e^{(\mu_i-\frac{1}{2}v^2)(2t+s-3T_i)+v(2B_t^{n+1}+B_s^{n+1}-3B_{T_i}^{n+1})}|q_{T_i}, q_{T_{i+1}}]dsdt\\
+&\sigma_{T_i}^3\int_{T_i}^{T_{i+1}}\int_{T_i}^{t}E_{\mathbb{F}^{n+1}}[e^{(\mu_i-\frac{1}{2}v^2)(2s+t-3T_i)+v(2B_s^{n+1}+B_t^{n+1}-3B_{T_i}^{n+1})}|q_{T_i}, q_{T_{i+1}}]dsdt,
\end{aligned}$$

We note that $(2B_t^{n+1}+B_s^{n+1}|q_{T_i}, q_{T_{i+1}})$, $T_i\le s\le t\le T_{i+1}$, is Gaussian with mean

$$3B_{T_i}^{n+1}+\frac{2t+s-3T_i}{T_{i+1}-T_i}(B_{T_{i+1}}^{n+1}-B_{T_i}^{n+1})$$

and variance

$$\frac{4(T_{i+1}-t)(t-T_i)+(T_{i+1}-s)(s-T_i)+4(T_{i+1}-t)(s-T_i)}{T_{i+1}-T_i}.$$

Similarly $(B_t^{n+1}+2B_s^{n+1}|q_{T_i}, q_{T_{i+1}})$, $T_i\le s\le t\le T_{i+1}$, is Gaussian with mean

$$3B_{T_i}^{n+1}+\frac{t+2s-3T_i}{T_{i+1}-T_i}(B_{T_{i+1}}^{n+1}-B_{T_i}^{n+1})$$

and variance

$$\frac{(T_{i+1}-t)(t-T_i)+4(T_{i+1}-s)(s-T_i)+4(T_{i+1}-t)(s-T_i)}{T_{i+1}-T_i}.$$

Thus the conditional expectation $E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}\Delta\widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}]$ can be evaluated.

Next we calculate the conditional mean and variance of $y_{T_{i+1}}|y_{T_i}, q_{T_i}, q_{T_{i+1}}$ for the DD-SABR model and similar calculations apply to the Normal-SABR case.

It follows from the solution of the DD-SABR model that the conditional mean is

$$\begin{aligned}\mu(y_{T_i}, q_{T_i}, q_{T_{i+1}}) = y_{T_i} &+ \frac{\omega_i\rho}{v}(e^{q_{T_{i+1}}} - e^{q_{T_i}}) - \frac{\omega_i\rho\mu_i}{\nu}E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t dt|q_{T_i}, q_{T_{i+1}}] \\ &- \frac{\omega_i^2}{2}E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t^2 dt|q_{T_i}, q_{T_{i+1}}] \\ &- \omega_i\sqrt{1-\rho^2}E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1}|q_{T_i}, q_{T_{i+1}}],\end{aligned}$$

where

$$E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1}|q_{T_i}, q_{T_{i+1}}] = 0.$$

Hence the conditional mean $\mu(y_{T_i}, q_{T_i}, q_{T_{i+1}})$ is given by

$$\begin{aligned}\mu(y_{T_i}, q_{T_i}, q_{T_{i+1}}) = y_{T_i} &+ \frac{\omega_i\rho}{v}(e^{q_{T_{i+1}}} - e^{q_{T_i}}) - \frac{\omega_i\rho\mu_i}{\nu}E_{\mathbb{F}^{n+1}}[\Delta\widetilde{V}_{T_i}|q_{T_i}, q_{T_{i+1}}] \\ &- \frac{\omega_i^2}{2}E_{\mathbb{F}^{n+1}}[\Delta V_{T_i}|q_{T_i}, q_{T_{i+1}}].\end{aligned}$$

The conditional variance $\eta^2(y_{T_i}, q_{T_i}, q_{T_{i+1}})$ is given by

$$\begin{aligned}\eta^2(y_{T_i}, q_{T_i}, q_{T_{i+1}}) = &(\frac{\omega_i\rho\mu_i}{\nu})^2 Var(\int_{T_i}^{T_{i+1}} \sigma_t dt|q_{T_i}, q_{T_{i+1}}) \\ &+ \frac{1}{4}\omega_i^4 Var(\int_{T_i}^{T_{i+1}} \sigma_t^2 dt|q_{T_i}, q_{T_{i+1}}) \\ &+ \omega_i^2(1-\rho^2)Var(\int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1}|q_{T_i}, q_{T_{i+1}}) \\ &+ \frac{\omega_i^3\rho\mu_i}{\nu}Cov(\int_{T_i}^{T_{i+1}} \sigma_t dt, \int_{T_i}^{T_{i+1}} \sigma_t^2 dt|q_{T_i}, q_{T_{i+1}}) \\ &- 2\frac{\omega_i^2\rho\mu_i}{\nu}\sqrt{1-\rho^2}Cov(\int_{T_i}^{T_{i+1}} \sigma_t dt, \int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1}|q_{T_i}, q_{T_{i+1}}) \\ &- \omega_i^3\sqrt{1-\rho^2}Cov(\int_{T_i}^{T_{i+1}} \sigma_t^2 dt, \int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1}|q_{T_i}, q_{T_{i+1}}).\end{aligned}$$

The covariance term can be written as

$$
\begin{aligned}
&Cov(\int_{T_i}^{T_{i+1}} \sigma_t dt, \int_{T_i}^{T_{i+1}} \sigma_t^2 dt | q_{T_i}, q_{T_{i+1}}) \\
=&E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t dt \int_{T_i}^{T_{i+1}} \sigma_t^2 dt | q_{T_i}, q_{T_{i+1}}] \\
-&E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t dt | q_{T_i}, q_{T_{i+1}}] E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t^2 dt | q_{T_i}, q_{T_{i+1}}] \\
=&E_{\mathbb{F}^{n+1}}[\Delta V_{T_i} \Delta \widetilde{V}_{T_i} | q_{T_i}, q_{T_{i+1}}] - E_{\mathbb{F}^{n+1}}[\Delta V_{T_i} | q_{T_i}, q_{T_{i+1}}] E_{\mathbb{F}^{n+1}}[\Delta \widetilde{V}_{T_i} | q_{T_i}, q_{T_{i+1}}],
\end{aligned}
$$

and

$$
\begin{aligned}
&Cov(\int_{T_i}^{T_{i+1}} \sigma_t dt, \int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1} | q_{T_i}, q_{T_{i+1}}) \\
=&E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t dt \int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1} | q_{T_i}, q_{T_{i+1}}] \\
-&E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t dt | q_{T_i}, q_{T_{i+1}}] E_{\mathbb{F}^{n+1}}[\int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1} | q_{T_i}, q_{T_{i+1}}] \\
=&0.
\end{aligned}
$$

Similarly,

$$
Cov(\int_{T_i}^{T_{i+1}} \sigma_t^2 dt, \int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1} | q_{T_i}, q_{T_{i+1}}) = 0.
$$

By Itô isometry the variance part $Var(\int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1} | q_{T_i}, q_{T_{i+1}})$ can be expressed as

$$
Var(\int_{T_i}^{T_{i+1}} \sigma_t d\widehat{W}_t^{n+1} | q_{T_i}, q_{T_{i+1}}) = E_{\mathbb{F}^{n+1}}[\Delta V_{T_i} | q_{T_i}, q_{T_{i+1}}].
$$

Hence the conditional variance $\eta^2(y_{T_i}, q_{T_i}, q_{T_{i+1}})$ is given by

$$
\begin{aligned}
\eta^2(y_{T_i}, q_{T_i}, q_{T_{i+1}}) = &(\frac{\omega_i \rho \mu_i}{\nu})^2 (E_{\mathbb{F}^{n+1}}[(\Delta \widetilde{V}_{T_i})^2 | q_{T_i}, q_{T_{i+1}}] - (E_{\mathbb{F}^{n+1}}[\Delta \widetilde{V}_{T_i} | q_{T_i}, q_{T_{i+1}}])^2) \\
&+ \frac{\omega_i^4}{4} (E_{\mathbb{F}^{n+1}}[(\Delta V_{T_i})^2 | q_{T_i}, q_{T_{i+1}}] - (E_{\mathbb{F}^{n+1}}[\Delta V_{T_i} | q_{T_i}, q_{T_{i+1}}])^2) \\
&+ \omega_i^2 (1 - \rho^2) E_{\mathbb{F}^{n+1}}[\Delta V_{T_i} | q_{T_i}, q_{T_{i+1}}] \\
&+ \frac{\omega_i^3 \rho \mu_i}{\nu} (E_{\mathbb{F}^{n+1}}[\Delta V_{T_i} \Delta \widetilde{V}_{T_i} | q_{T_i}, q_{T_{i+1}}] \\
&- E_{\mathbb{F}^{n+1}}[\Delta V_{T_i} | q_{T_i}, q_{T_{i+1}}] E_{\mathbb{F}^{n+1}}[\Delta \widetilde{V}_{T_i} | q_{T_i}, q_{T_{i+1}}]).
\end{aligned}
$$

Chapter 4

Stochastic volatility Markov-functional models: calibration

4.1 Introduction

In the previous chapter, we have specified a stochastic volatility LIBOR Markov-functional model by choosing a SABR driving process and proposed an algorithm to implement the model numerically. But we have not considered the problem of how to choose the parameters of the SABR driving process and the input prices of the set of digital caplets which will be fed into the model. In this chapter we will consider this by calibrating to the market correlation structure and the market prices of vanilla options.

Let us first consider the input prices of digital caplets. As we discussed in Chapter 2 the input prices of a set of digital caplets as functions of strikes determine the marginal distributions of the corresponding LIBORs at their setting dates in their associated forward measure. Feeding in the input prices of digital caplets is equivalent to feeding in marginals together with the initial term structure. From the specification of a LIBOR Markov-functional model one can see that given a driving process theoretically any marginal distribution of LIBORs at their setting dates can be fed into the model. This separates the specification of a driving process from the marginals. This is a very unusual feature for an interest rate model. In contrast to, for example, LIBOR market models and short rate models the marginal distribution of LIBORs at their setting dates are fully determined by the dynamics of the corresponding state process. This separation in a Markov-functional model pro-

vides flexibility but this may also potentially cause an unstable dynamics of forward rates. Andersen and Piterbarg [5] made a remark on Markov-functional models that a non-parametric formulation of the marginal distribution for LIBORs may result in unrealistic evolution of the volatility smile through time. To avoid these issues the authors suggested to feed in an arbitrage-free formula for digital caplet prices, where the formula can be derived from e.g. a log-Normal, displaced-diffusion or a CEV model. Bennett and Kennedy [7] showed that a LIBOR Markov-functional model with a Gaussian driver together with the Black's formula for (digital) caplets is numerically similar to the one-factor separable LIBOR market model. Recently Gogala and Kennedy [29] extended the above link to a more general local-volatility case. Based on this link, the authors proposed an approach for choosing an appropriate combination of driving process and (digital) caplet prices, and such a combination leads to desirable dynamics of future implied volatilities. In this chapter we consider the link between LIBOR Markov-functional models and separable LIBOR market models in the stochastic volatility case. We expect that a stochastic volatility Markov-functional model with a SABR driver together with a SABR marginal distribution of LIBORs should be numerically similar to the corresponding separable SABR LIBOR market model. Based on this link the intuition behind the SDEs of the separable SABR-LIBOR market model can be transferred to the corresponding stochastic volatility LIBOR Markov-functional model. This gives us a guide as to how to calibrate a stochastic volatility Markov-functional model. Based on this link we develop a calibration routine to feed in the prices of digital caplets by calibrating to the market caplets or swaptions prices. A numerical investigation of the performance of this calibration routine is also given.

In a Markov-functional model, the specification of a driving process determines the dynamics of the model. In a one-dimensional LIBOR Markov-functional model, Hunt and Kennedy [37] showed that a particular choice of Gaussian driving process leads to mean reversion of spot LIBOR. However this particular choice of the driver results in an unsatisfactory hedging performance. Furthermore Kennedy and Pham [48] showed that the specification of a Gaussian driving process has an effect on a Bermudan swaption in terms of hedging behaviour. The authors proposed a particular specification of a Gaussian driver, which is referred to as the one step covariance driver, by calibrating to the market correlation structure of swap rates in a one-dimensional swap Markov-functional model, and this choice leads to a desirable hedging behaviour. This motivates us to choose a modification of the SABR model as the driving Markov process in the stochastic volatility Markov-functional model which has been specified in the previous chapter. Recall that in

this modified SABR model, we have a drift function which is piecewise constant in the stochastic volatility process. We will show that this drift function can be chosen by calibrating to the market correlations. A numerical investigation of this calibration performance will be given. We will see in Chapter 5 that this calibration approach retains the desirable hedging behaviour which is similar to the one step covariance Markov-functional model.

The rest of the chapter is organized as follows. In Section 4.2 we describe the calibration problem. In Section 4.3 we propose an approach for choosing an appropriate combination of driving process and (digital) caplet prices by introducing a separable LIBOR market model. In Sections 4.4 and 4.5, we discuss the problem of how to calibrate a stochastic volatility LIBOR Markov-functional model to the market prices of caplets and swaptions respectively. In Section 4.6 we consider calibrating to the market correlation structure. We conclude in Section 4.7.

4.2 Calibration problem description

Consider the tenor structure (2.1). For ease of reference we present again the target SABR driving process of the stochastic volatility LIBOR Markov-functional model (MFM) discussed in the previous chapter:

$$\begin{aligned} dF_t &= \sigma_t F_t^{\beta} dW_t^{n+1} \qquad \beta \in [0,1], \\ d\sigma_t &= \mu_t \sigma_t dt + \nu \sigma_t dB_t^{n+1} \qquad \nu > 0, \\ dB_t^{n+1} dW_t^{n+1} &= \rho dt \qquad \rho \in [-1,1] \end{aligned} \tag{4.1}$$

where B^{n+1} and W^{n+1} are correlated Brownian motions under the terminal measure $\mathbb{F}^{n+1}$ associated with the numeraire the T_{n+1}-maturity zero-coupon bond (ZCB) $D_{.,Tn+1}$, and μ_t is assumed to be piecewise constant

$$\mu_t = \sum_{j=0}^{n-1} \mu_j I_{\{t \in [T_j, T_{j+1})\}}. \tag{4.2}$$

Remember that this is our target driving process, and we in fact take a transformation of it as the driving process for the sake of efficient implementation. See Section 3.3.2 for details. Recall that we set up the model by feeding in the prices $V_0^i(K)$, $V_0^i : [0,\infty] \to \mathbb{R}$, of a set of digital caplets on the ith LIBOR L^i with strike K for $i = 1, ..., n$. Thus the stochastic volatility LIBOR MFM is fully determined by

1. the input prices of digital caplets $V_0^i(K)$ as a function of the strike $K \geq 0$ for

$i = 1, ..., n$.

2. the parameters of our stochastic volatility MFM i.e. the parameters of the target driving process (4.1) which includs σ_0, β, F_0, ρ, ν and μ_j, $j = 0, ..., n-1$.

In the previous chapter, we focused on the numerical implementation of this model by assuming that the above two sets of information are given to us. In this chapter we consider the problem of how to choose them by calibrating to the market prices of vanilla options and the market correlation structure.

Remark 5. *In the original SABR model [31], the stochastic volatility process is driftless. But in the driving process (4.1), we use a modified SABR model where we have a drift function μ_j, $j = 0, ..., n-1$, for the volatility σ. We do so since we need μ_j's to capture the correlation structure. More importantly, this will lead to a desirable hedging behaviour which will be seen in Chapter 5.*

Before we proceed to the calibration of the stochastic volatility LIBOR MFM, we make an assumption. Kaisajuntti and Kennedy [44] identified a SABR style model as an appropriate choice for the level of interest rates by investigating market data. In that sense the target SABR driving process (4.1) can be determined by investigating historical market data or by traders based on their market judgement or beliefs. In particular throughout this chapter let us assume that for the target SABR driving process (4.1), the parameters β, F_0, ρ, σ_0 and ν have been determined exogenously. Consequently the remaining parameters for the target driver (4.1) are μ_j for $j = 0, ..., n-1$.

Note that the input prices of the digital caplets $V_0^i(K)$ with respect to strike K contain information on the market initial term structure and the distribution of the LIBOR $L^i_{T_i}$ under the associated forward measure $\mathbb{F}^{i+1}$ for $i = 1, ..., n$. Thus given the market initial term structure, choosing the input prices $V_0^i(K)$ is equivalent to choosing the distribution of the random variable $L^i_{T_i}$ under $\mathbb{F}^{i+1}$ for $i = 1, ..., n$. In Section 4.3 we will consider the problem of what marginal distributions of the LIBORs $\{L^i_{T_i}; i = 1, ..., n\}$ should be fed in. We will see that given the target SABR driver (4.1), the SABR type, as opposed to, for example, the CEV or log-Normal marginal distribution is an appropriate choice. In Sections 4.4 and 4.5 we will determine the input prices $V_0^i(K)$, $i = 1, ..., n$, of digital caplets at strike $K \geq 0$ by calibrating to the market prices of caplets and swaptions respectively. In Section 4.6 we will choose the parameters μ_j, $j = 0, ..., n-1$, of the target SABR driver (4.1) by calibrating to the market correlation structure of LIBORs. In this chapter we focus on the LIBOR version MFM, but all the discussion also applies to the swap

version MFM. In Appendix 4.B we will briefly discuss the calibration issue in the stochastic volatility swap MFM.

4.3 Choice of marginals

In a MFM, given a driving process, theoretically any marginal distributions of the LIBORs $\{L^i_{T_i}; i = 1, ..., n\}$ can be fed into the model. This leads to more flexibility but one needs to be careful since an inconsistent combination of driving process and marginals could potentially result in unstable dynamics of the forward rates. In this section we will find a reasonable combination by considering the link between LIBOR MFMs and separable LIBOR market models (LMMs). This link will also give us a guide as to how to perform the calibration.

In Section 4.3.1 we introduce the idea of separability and one-factor separable LMMs. We review the link between LIBOR MFMs and separable LMMs in the one-factor case. Based on this link, an appropriate combination of driver and prices of digital caplets is found. In Section 4.3.2 we consider the link between LIBOR MFMs and separable LMMs in the stochastic volatility case by considering separable SABR-LMMs. Based on this link, given the target SABR driving process (4.1) we find a consistent marginals to feed into the stochastic volatility LIBOR MFM.

4.3.1 Separable LIBOR market models with local volatility

Let us consider the following LMMs with local volatility under the terminal measure $\mathbb{F}^{n+1}$:

$$\begin{aligned} dL^i_t &= -\sum_{j=i+1}^{n} \left(\frac{\sigma^j_t \alpha_j \phi(L^i_t)}{1+\alpha_j L^j_t}\right)\sigma^i_t \phi(L^i_t)dt + \sigma^i_t \phi(L^i_t) dW^{n+1}_t, \quad i = 1, ..., n-1 \qquad (4.3) \\ dL^n_t &= \sigma^n_t \phi(L^n_t) dW^{n+1}_t, \end{aligned}$$

where $\phi(\cdot)$ is a deterministic function of the rate satisfying appropriate regularity conditions where more details can be found in [29]. One can choose the local volatility function $\phi(\cdot)$ as, for instance, $\phi(x) = x$, $\phi(x) = x^\beta$, $0 < \beta < 1$, and $\phi(x) = x + \theta$, $\theta \in \mathbb{R}$, which are corresponding to the log-Normal LMM, CEV LMM and displaced-diffusion LMM. For more choices for the local volatility function $\phi(\cdot)$, the reader is referred to [11].

We can see that the drift term of SDE (4.3) is dependent on forward LIBORs except for the drift of the nth LIBOR L^n which is zero. Consequently a LMM is a high-dimensional model and it is Markovian with respect to its all forward rates even

when it is driven by only one factor. In practice a simulation method is required for an accurate implementation of a LMM. However Pelsser et al. [56] demonstrated that a numerical approximation to a separable LMM allows for representation by a low-dimensional Markov process. Moreover they found that such an approximation is very accurate for LMMs up to 10 years. This in turn would give an efficient implementation by, for example, a finite difference method. We give the definition of separability under LMMs from [7].

Definition 2. *The one-factor LMM (4.3) is said to be separable if the volatility functions* σ^i, $i = 1, ..., n$, *satisfy*

$$\sigma_t^i = \gamma^i \sigma_t$$

for some constants γ^i, *for* $0 \leq t \leq T_i$, $i = 1, ..., n$.

Separability provides a link between Market models and MFMs. As Pelsser et al. [56] pointed out, under a one-factor separable log-Normal LMM, the drift-approximated forward LIBORs is an increasing function of a one-dimensional Markov process. It was noted by Bennett and Kennedy [7] that if we take this one-dimensional Markov process as the driving process then the corresponding MFM is numerically very similar to the one-factor separable LMM as long as we fit the Black's formula for caplets. Recently Gogala and Kennedy [29] extended the above link to a more general local-volatility case. In particular consider the following one-factor separable LMM with local volatility function

$$\begin{aligned} dL_t^i &= -\sum_{j=i+1}^{n} \left(\frac{\gamma^j \alpha_j \phi(L_t^j)}{1+\alpha_j L_t^j}\right) \gamma^i \sigma_t^2 \phi(L_t^i) dt + \gamma^i \sigma_t \phi(L_t^i) dW_t^{n+1}, \quad i = 1, ..., n-1 \quad (4.4) \\ dL_t^n &= \gamma^n \sigma_t \phi(L_t^n) dW_t^{n+1}, \end{aligned}$$

where $\phi(\cdot)$ is a deterministic function of the rate satisfying appropriate regularity conditions and W^{n+1} is a one-dimensional Brownian motion under the terminal measure. This one-factor separable local volatility LMM is found to have a similar dynamics to a LIBOR MFM with the following driving process

$$dx_t = \sigma_t \phi(x_t) dW_t^{n+1} \quad (4.5)$$

feeding in the prices $V_0^i(K)$, $i = 1, ..., n$, of digital caplets on the ith LIBOR L^i derived from the SDEs (4.4). The link between MFMs and separable LMMs implies that a high-dimensional separable LMM can be approximated by a low-dimensional arbitrage-free model which allows for a more efficient implementation.

Thanks to the above link, the intuition behind the well-understood SDE formulation of a LMM can be applied to the corresponding LIBOR MFM. This gives us a guide as to how to calibrate a MFM. In particular, given the driving process (4.5), it is reasonable to feed in the prices $V_0^i(K)$, $i = 1, ..., n$, of digital caples derived from the SDEs (4.4). This means that given a local volatility driving process, the marginal distributions of the LIBORs $\{L^i_{T_i}; i = 1, ..., n\}$ derived from the consistent local volatility style model is an appropriate choice to feed in. By doing so the resulting MFM is similar to the one-factor separable LMM (4.4). Otherwise a mismatch of a driving process and marginals of $\{L^i_{T_i}; i = 1, ..., n\}$ might potentially lead to a non-transparent dynamics of forward LIBORs and therefore result in an unstable evolution of the implied volatility surface.

4.3.2 Separable LIBOR market models with stochastic volatility

We have discussed an approach for choosing an appropriate combination of driving process and caplet prices for a one-dimensional LIBOR MFM by considering the link between one-factor separable local volatility LMMs and one-dimensional LIBOR MFMs. Now we extend the intuition to the stochastic volatility version.

Following a similar explanation for the local volatility case, in a stochastic volatility LIBOR MFM, given a SABR driving process (4.1) it would be sensible to feed in a SABR marginal distribution of the LIBORs $\{L^i_{T_i}; i = 1, ..., n\}$. To see this let us recall the link between the one-dimensional MFM and the one-factor separable LMM that we discussed. Let us take the local volatility in the form of $\phi(x) = x^\beta$ in the SDE (4.4) and this leads to the one-factor separable CEV LMM which is found to have similar dynamics to a one-dimensional LIBOR MFM with the following CEV driving process

$$dx_t = \sigma_t x_t^\beta dW_t^{n+1}$$

together with the prices $V_0^i(K)$, $i = 1, ..., n$, of digital caplet prices derived from the corresponding CEV model. We expect that the link is retained if we introduce stochastic volatility. In particular consider the following separable SABR-LMM in

the terminal measure $\mathbb{F}^{n+1}$:

$$dL_t^i = -\sum_{j=i+1}^{n} (\frac{\gamma^j \alpha_j (L_t^j)^\beta}{1+\alpha_j L_t^j}) \gamma^i \sigma_t^2 (L_t^i)^\beta dt + \gamma^i \sigma_t (L_t^i)^\beta dW_t^{n+1} \quad i = 1, ..., n-1,$$

(4.6)

$$\begin{aligned} dL_t^n &= \gamma^n \sigma_t (L_t^n)^\beta dW_t^{n+1} \qquad \beta \in [0,1], \\ d\sigma_t &= \mu_t \sigma_t dt + v \sigma_t dB_t^{n+1} \qquad v > 0, \\ dW_t^{n+1} dB_t^{n+1} &= \rho dt \qquad \rho \in [-1,1], \end{aligned}$$

where μ_t is given by (4.2). This model is expected to be numerically similar to the stochastic volatility LIBOR MFM with the target SABR driving process (4.1) together with the prices $V_0^i(K)$, $i = 1, ..., n$, for digital caplets derived from the model (4.6).

Note that the separable SABR-LMM (4.6) we discussed here is a restrictive stochastic volatility LMM. In order to obtain more flexibility, in the literature commonly we have a different stochastic volatility for each LIBOR L^i. See [33], [62] and references therein. But when we impose separability on the stochastic volatility, a common stochastic volatility σ is associated to all forward LIBORs. The cost we have to pay is the loss of flexibility, but in return the resulting separable SABR-LMM, which is a high-dimensional model, can be approximated by a low-dimensional stochastic volatility MFM which can be implemented for pricing and hedging in practice.

Based on the above link, given the target SABR driving process (4.1), one can feed in the prices $V_0^i(K)$, $i = 1, ..., n$, of digital caplets with strike $K \geq 0$ produced by the SDEs (4.6). This leads to a stochastic volatility LIBOR MFM which is expected to be similar to the separable SABR-LMM (4.6). The link also gives us a guide as to how to calibrate the stochastic volatility LIBOR MFM. We will develop a calibration routine for our stochastic volatility MFM based on the above link in the next sections. Note that we have not established the above link numerically, but we find that the calibration routine developed based on it works well which will be seen later. Thus even though the above link may not be as tight as in the non-stochastic case, it will not have an influence on the calibration routine since we just take it as a guide for calibration and test its effectiveness directly.

4.4 Calibration to caplets

Throughout this section let us assume that the parameters μ_j, $j = 0, ..., n-1$, for the target driver (4.1) are given. We will explain the details of how to determine these parameters later in this chapter. So far we have decided to feed the prices $V_0^i(K)$, $i = 1, ..., n$, of digital caplets with strike $K \geq 0$ produced by the model (4.6) into our stochastic volatility LIBOR MFM with the target SABR driver (4.1). But when we find the prices $V_0^i(K)$, we have free parameters γ^i, $i = 1, ..., n$, for the SDE (4.6) that need to be determined. In this section we will choose the parameters γ^i's by calibrating to the market prices of caplets, and then find the input prices $V_0^i(K)$ of digital caplets produced by the calibrated model (4.6). By feeding input prices $V_0^i(K)$ into our stochastic volatility LIBOR MFM, the resulting model will reproduce the market prices of caplets. To do so we develop a calibration routine in this section. The calibration routine involves the SABR formula for implied volatility smiles introduced by Hagan et al. [31]. Thus in Section 4.4.1 we review the SABR model and introduce the SABR formula. In Section 4.4.2 we develop the calibration routine to obtain the prices $V_0^i(K)$, $i = 1, ..., n$, of digital caplets.

4.4.1 SABR model

In this subsection let us step out of term structure models and focus on an interest rate model for a single forward rate. In particular we review the SABR model introduced by Hagan et al. [31]. The material in this subsection is from [31]. Let us consider the following dynamics of the forward rate

$$\begin{aligned} d\hat{F}_t &= \hat{\alpha}_t \hat{F}_t^{\beta} dW_t^1 \qquad \hat{F}_0 = f, \quad \beta \in [0,1] \\ d\hat{\alpha}_t &= \nu \hat{\alpha}_t dW_t^2 \qquad \hat{\alpha}_0 = \alpha \\ dW_t^1 dW_t^2 &= \rho dt \qquad \rho \in [-1,1] \end{aligned} \tag{4.7}$$

where W^1 and W^2 are correlated Brownian motions under some equivalent martingale measure. We note that the SABR model only treats one rate and the underlying asset $\hat{F}$ can be taken as a forward swap rate or LIBOR.

Hagan et al. [31] proposed an analytical approximation for the implied volatility of a vanilla call option as a function $\sigma_B(K, f; \alpha, \beta, \nu, \rho)$ of today's forward price f and the strike K by using singular perturbation techniques, which is

given by

$$\sigma_B(K,f;\alpha,\beta,\nu,\rho) = \frac{\sigma_0}{(fK)^{\frac{1-\beta}{2}}[1+\frac{(1-\beta)^2}{24}ln^2(\frac{f}{K})+\frac{(1-\beta)^4}{1920}ln^4(\frac{f}{K})+\cdots]}\cdot(\frac{z}{x(z)})$$
$$\cdot\{1+[\frac{(1-\beta)^2\alpha^2}{24(fK)^{1-\beta}}+\frac{\rho\beta\nu\alpha}{4(fK)^{\frac{1-\beta}{2}}}+\nu^2\frac{2-3\rho^2}{24}]T_i+\cdots\}, \quad (4.8)$$

with

$$z := \frac{\nu}{\alpha}(fK)^{\frac{1-\beta}{2}}ln(\frac{f}{K})$$

and

$$x(z) := ln(\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}).$$

The above formula for implied volatilities is also called the SABR formula, and it is widely used in practice for implied volatility smile interpolation and extrapolation. Among the parameters of the SABR formula, the parameter α mainly controls the overall level of the implied volatility curve, the parameter ρ controls the skew of the curve and ν determines the curvature of the curve or equivalently how much smile the curve can represent. The parameter β has a similar role to the parameter ρ, but it is often determined upfront from historical data. Since these parameters play separate roles to control an implied volatility curve, the calibration of the SABR formula is accurate.

Note that in financial markets vanilla options are quoted in terms of implied volatilities. The corresponding prices can be obtained by inserting the implied volatility $\sigma_B(K,f)$ (4.8) into the corresponding Black's formula.

4.4.2 Calibration routine

In this subsection we develop a calibration routine. We choose the parameters γ^i's by calibrating the separable SABR-LMM (4.6) to the market prices of a set of caplets. Then we find the prices $V_0^i(K)$, $i = 1,...,n$, of the set of digital caplets produced by the calibrated model (4.6). Finally we feed the input prices $V_0^i(K)$ into the stochastic volatility LIBOR MFM which will reproduce the market prices of caplets.

Let us first determine the parameters γ^i's in the separable SABR-LMM (4.6). To do so we consider the separable SABR-LMM (4.6) under the associated forward measure. In particular the ith forward LIBOR L^i is a (local) martingale in the

associated forward measure $\mathbb{F}^{i+1}$ corresponding to the numeraire $D_{.,T_{i+1}}$ so that the forward LIBOR process L^i should be driftless. However a change of measure also changes the drift term in the volatility because of the non-zero correlation between the two Brownian motions that drive forward rates and stochastic volatility respectively.

Remark 6. *We will assume that in addition to modelling the LIBORs there is at least one option in the economy which means that we are in a complete setting. In this case when moving to another equivalent martingale measure associated with a different numeraire the Radon-Nikodým derivative is the ratio of numeraires. This assumption also applies to the other incomplete models when we need to change measure.*

Lemma 1. *Let L^i satisfies SDEs (4.6). Under the T_{i+1}-forward measure $\mathbb{F}^{i+1}$ the dynamics of L^i satisfies*

$$\begin{aligned} dL_t^i &= \sigma_t^i (L_t^i)^\beta dW_t^{i+1} \qquad \beta \in [0,1], \\ d\sigma_t^i &= \mu^i(t,\sigma_t,L_t)\sigma_t^i dt + v\sigma_t^i dB_t^{i+1} \qquad \sigma_t^i = \gamma^i \sigma_t, \quad v > 0 \\ dW_t^{i+1} dB_t^{i+1} &= \rho dt \qquad \rho \in [-1,1], \end{aligned} \tag{4.9}$$

with

$$\mu^i(t,\sigma_t,L_t) := \mu_t + \rho v \sum_{j=i+1}^{n} \frac{\sigma_t^j \alpha_j (L_t^j)^\beta}{1+\alpha_j L_t^j}, \tag{4.10}$$

where W^{i+1} and B^{i+1} are correlated Brownian motions in the forward measure $\mathbb{F}^{i+1}$.

Proof. The Brownian motion B^{n+1} in the terminal $\mathbb{F}^{n+1}$ can be expressed as

$$dB_t^{n+1} = \rho dW_t^{n+1} + \sqrt{1-\rho^2} dZ_t^{n+1}, \tag{4.11}$$

where Z^{n+1} and W^{n+1} are independent Brownian motions. Define ς_t by

$$\varsigma_t := \left.\frac{d\mathbb{F}^{i+1}}{d\mathbb{F}^{n+1}}\right|_{\mathcal{F}_t} = \frac{D_{tT_{i+1}}/D_{0T_{i+1}}}{D_{tT_{n+1}}/D_{0T_{n+1}}},$$

where $\{\mathcal{F}_t\}$ is the filtration generated by Z^{n+1} and W^{n+1}. After calculation, we have that

$$d\varsigma_t = \varsigma_t \sum_{j=i+1}^{n} \left(\frac{\gamma^j \alpha_j (L_t^j)^\beta}{1+\alpha_j L_t^j}\right) \sigma_t dW_t^{n+1}.$$

From the Girsanov theorem we have that the processes

$$\begin{aligned} dW_t^{i+1} &:= dW_t^{n+1} - \frac{d\varsigma_t}{\varsigma_t} \cdot dW_t^{n+1} \\ &= dW_t^{n+1} - \sum_{j=i+1}^{n} (\frac{\gamma^j \alpha_j (L_t^j)^\beta}{1+\alpha_j L_t^j}) \sigma_t dt \end{aligned} \tag{4.12}$$

and

$$\begin{aligned} dZ_t^{i+1} &:= dZ_t^{n+1} - \frac{d\varsigma_t}{\varsigma_t} \cdot dZ_t^{n+1} \\ &= dZ_t^{n+1} \end{aligned} \tag{4.13}$$

are independent Brownian motions in the forward measure $\mathbb{F}^{i+1}$ corresponding to the numeraire $D_{.,T_{i+1}}$. It follows from equations (4.11), (4.12) and (4.13) that

$$\begin{aligned} dB_t^{n+1} &= \rho dW_t^{i+1} + \sqrt{1-\rho^2} dZ_t^{i+1} + \rho \sum_{j=i+1}^{n} (\frac{\gamma^j \alpha_j (L_t^j)^\beta}{1+\alpha_j L_t^j}) \sigma_t dt, \\ &= dB_t^{i+1} + \rho \sum_{j=i+1}^{n} (\frac{\gamma^j \alpha_j (L_t^j)^\beta}{1+\alpha_j L_t^j}) \sigma_t dt, \end{aligned}$$

where B^{i+1} is a Brownian motion in $\mathbb{F}^{i+1}$ defined as

$$dB_t^{i+1} = \rho dW_t^{i+1} + \sqrt{1-\rho^2} dZ_t^{i+1}.$$

The result then follows from the SDE (4.6). □

The dynamics (4.9) of the forward LIBOR L^i under the associated foward measure $\mathbb{F}^{i+1}$ is in the form similar to the SABR model (4.7) except for the drift of the volatility. The drift function in the volatility in the model (4.9) is dependent on the forward LIBORs and also a function of the common volatility σ as well as its square. This complex form of drift term prevents us from achieving the SABR formula for implied volatility of caplets. Thus we will derive an approximate dynamics in the form of (4.7). To this end we apply the following two-step approximation proposed by Morini and Mercurio [54].

1. **First step: volatility drift approximation**. We first deal with the drift term of the volatility. We note that the drift function $\mu^i(t, \sigma_t, L_t)$ depends on forward LIBORs L^j, $j \geq i+1$, as well as volatility σ itself. Our first step is to approximate it by its initial value. In particular we approximate the forward LIBORs and volatility σ_t that appear in the drift function $\mu^i(t, \sigma_t, L_t)$ by their initial

values

$$\mu^i(t,\sigma_0,L_0) := \mu_t + \rho v \sum_{j=i+1}^{n} \frac{\sigma_0^j \alpha_j (L_0^j)^\beta}{1+\alpha_j L_0^j} \approx \mu^i(t,\sigma_t,L_t). \tag{4.14}$$

Morini and Mercurio [54] has justified this freezing approximation. It found that the approximation is accurate for short maturities and the accuracy begins to be lost for long maturities (30Y). Note that there also exists some more sophisticated approximations. See, for example, predictor-corrector schemes [39] and Brownian bridge approximations [56]. In our work we simply choose constant approximation since it performs well in our numerical study which will be seen later. The freezing approximation leads to the following model in $\mathbb{F}^{i+1}$ where the stochastic volatility becomes log-Normal:

$$\begin{aligned} dL_t^i &= \sigma_t^i (L_t^i)^\beta dW_t^{i+1} \qquad \beta \in [0,1], \\ d\sigma_t^i &= \mu^i(t,\sigma_0,L_0)\sigma_t^i dt + v\sigma_t^i dB_t^{i+1} \qquad \sigma_t^i = \gamma^i \sigma_t, \quad v>0 \\ dW_t^{i+1} dB_t^{i+1} &= \rho dt \qquad \rho \in [-1,1], \end{aligned} \tag{4.15}$$

2. **Second step: SABR formula for non-zero drift**. Note that the SABR formula is derived from the SABR model of the form (4.7) where the volatility is driftless. Although we have obtained a more tractable model (4.15), there still exists a deterministic drift function. In this step we approximate the model (4.15) by the following SDEs where the volatility process is driftless

$$\begin{aligned} dL_t^i &= \tilde{\sigma}_t^i (L_t^i)^\beta dW_t^{i+1} \qquad \beta \in [0,1], \\ d\tilde{\sigma}_t^i &= v\tilde{\sigma}_t^i dB_t^{i+1} \qquad v>0 \\ dW_t^{i+1} dB_t^{i+1} &= \rho dt \qquad \rho \in [-1,1], \end{aligned} \tag{4.16}$$

such that the caplet prices given by the two models are the same. We note that the caplet price depends on the distribution of the volatility $\sigma_{T_i}^i$ at expiration time T_i, but it also depends on the path of the volatility from current time to time T_i. As justified by Morini and Mercurio [54], it is a reasonable approximation that the caplet price depends on the average volatility from now to time T_i. Thus we choose the initial value $\tilde{\sigma}_0^i$ such that

$$E_{\mathbb{F}^{i+1}}[\int_0^{T_i} \sigma_t^i dt] = E_{\mathbb{F}^{i+1}}[\int_0^{T_i} \tilde{\sigma}_t^i dt].$$

This leads to the value of $\tilde{\sigma}_0^i$ which is given by

$$\tilde{\sigma}_0^i = \frac{\gamma^i \sigma_0 \int_0^{T_i} e^{\int_0^t \mu^i(s,\sigma_0,L_0)ds} dt}{T_i}. \tag{4.17}$$

Finally, for the SABR-LMM (4.9), we obtain an approximate SABR formula

$$\sigma_B(K, L_0^i; \tilde{\sigma}_0^i, \beta, \nu, \rho)$$

given by (4.8) for the caplet on the ith LIBOR L^i with strike K, where $\tilde{\sigma}_0^i$ is given by equation (4.17).

So far we have obtained an approximate SABR formula for the SABR-LMM (4.9). We can choose the parameters γ^i's such that the approximate SABR formula $\sigma_B(K, L_0^i; \tilde{\sigma}_0^i, \beta, \nu, \rho)$ matches the market implied volatilities of the set of caplets. However we note that the only free parameters we have for the model (4.9) are parameters γ^i for $i = 1, ..., n$. Consequently we cannot match the market implied volatilities of the set of caplets for all strikes. Our choice is to match the market implied volatilities of caplets struck at some particular strike K^i for each $i = 1, ..., n$. For example one can calibrate to the ATM implied volatilities i.e. $K^i = L_0^i$ for $i = 1, ..., n$.

Suppose we are given the market implied volatilities $\sigma^{mkt}(K^i, L_0^i)$ of the caplet on the ith LIBOR L^i struck at some particular strike K^i for each $i = 1, ..., n$. We can choose γ^i's such that the SABR formula for the SABR-LMM (4.9) matches the market implied volatilities $\sigma^{mkt}(K^i, L_0^i)$:

$$\sigma_B(K^i, L_0^i; \tilde{\sigma}_0^i, \beta, \nu, \rho) = \sigma^{mkt}(K^i, L_0^i), \tag{4.18}$$

for each $i = 1, ..., n$.

Having fixed the parameters γ^i, $i = 1, ..., n$, we can find the input prices $V_0^i(K)$, $i = 1, ..., n$, of the set of digital caplets as a function of the strike K produced by the separable SABR-LMM (4.9). To do so, for this study we apply Monte Carlo methods. Finally we can feed the input prices $V_0^i(K)$, $i = 1, ..., n$, into our stochastic volatility LIBOR MFM, and the resulting model should reproduce the market implied volatilities $\sigma^{mkt}(K^i, L_0^i)$, $i = 1, ..., n$, of the set of caplets.

4.4.3 Numerical study

In this subsection, we investigate the accuracy of the calibration routine developed in Section 4.4.2. We note that the calibration routine involves some approximations

which include the one intrinsic to the SABR formula. In order to distinguish the effect of the different approximations, we give caplet prices produced at each approximation step. In particular following the calibration routine, given the market implied volatilities of the caplets, we can determine the parameters γ^i, $i = 1, ..., n$, by using the SABR formula and following (4.18). Therefore the implied volatilities of the caplets produced by the approximate SABR-LMM (4.16) using the SABR formula are exactly the same as the market implied volatilities. By comparing these implied volatilities to those produced by the approximate SABR-LMM (4.16) using Monte Carlo methods, one can investigate the accuracy of the SABR formula. We then compare the implied volatilities given by the model (4.16) to those produced by the SABR-LMM (4.9), where both models are using Monte Carlo methods, and we can distinguish the effect of the "two-step approximation". Finally we set up the stochastic volatility LIBOR MFM and produce the MFM model implied volatilities of the caplets. By comparing the MFM model implied volatilities with those given by the model (4.9), we can investigate the calibration performance of the stochastic volatility LIBOR MFM.

In our numerical study, we consider the calibration of the SV MFM with different parameters. We calibrate to the market implied volatilities $\sigma^{mkt}(L_0^i, L_0^i)$ of the ATM caplets on 17 October 2007. The numerical result is given in Table 4.1. The column "Market" is the market implied volatilities $\sigma^{mkt}(L_0^i, L_0^i)$ of the ATM caplets on 17 October 2007. The columns "Approx SABR-LMM (MC)" and "SABR-LMM (MC)" are implied volatilities of the ATM caplets produced by the approximate SABR-LMM (4.16) and the SABR-LMM (4.9) respectively using Monte Carlo (MC) methods. The last column "MFM" is the ATM implied volatilities produced by the stochastic volatility LIBOR MFM. We can see from the "Market" and "MFM" columns in Table 4.1 that the resulting MFM is able to reproduce the market implied volatilities. For all scenarios, the calibration is accurate for short maturities but the performance is getting worse for longer maturities. Let us investigate the source of the calibration error. By comparing the market implied volatilities with those produced by the approximate SABR-LMM (4.16), we can see that most calibration error is coming from the SABR formula. The SABR formula starts losing accuracy for longer maturities. Then we compare the implied volatilities produced by the models (4.16) and (4.9), and it turns out that the "two-step approximation" developed in the previous subsection performs well for short maturities but loses some precision for $T > 15$ years. We can see from the data in the last two columns that the calibration of the stochastic volatility LIBOR MFM performs quite well for all scenarios. The MFM is able to reproduce the implied volatilities produced by

the SABR-LMM (4.9). Overall the accuracy of the calibration routine developed in this section is satisfactory for short maturities and the calibration error is mainly from the SABR formula.

Remark 7. *From the numerical results, one can see that the "MFM" implied volatilities are systematically lower than the "Market" ones for longer maturities. Moreover the change in the parameter γ^i will only affect the model implied volatility of the caplet with maturity T_i for $i = 1, ..., n$. One can develop an "iterative procedure" which adjusts the value of γ^i, $i = 1, ..., n$, until achieving precise calibration of the MFM to the market. We use this technique for the calibration of the swap rate based stochastic volatility MFM in Chapter 5.*

4.5 Calibration to swaptions

In the previous section we determined the parameters γ^i, $i = 1, ..., n$, in the LIBOR based model (4.6) by calibrating to the market implied volatilities of caplets. In this section we consider the problem of how to determine γ^i's by calibrating to the market implied volatilities of co-terminal vanilla swaptions on the swap rates $\{y^{i,n+1-i}; i = 1, ..., n\}$. By "co-terminal" we mean that they have the same termination date, i.e. the same final payment date. Note that we are not forced to calibrate the parameters γ^i's to the co-terminal type swaptions. The calibration to the vanilla swaptions on the general swap rates $\{y^{i,j}; i = 1, ..., n, j = 1, ..., n+1-i\}$ is straightforward to extend. But in this chapter we focus on the co-terminal swap rates (swaptions) because they are the underlyings of the corresponding co-terminal Bermudan swaption which is the exotic option we are most interested in and will be studied in Chapter 5.

To perform the calibration we first derive an approximate formula for the vanilla swaptions prices within the model (4.6). Then we choose the parameters γ^i, $i = 1, ..., n$, in the model (4.6) by calibrating to the market prices of co-terminal vanilla swaptions using the approximate pricing formula. Finally we can find the input prices $V_0^i(K)$, $i = 1, ..., n$, of the set of digital caplets as a function of the strike K produced by the calibrated model (4.6). By feeding in the prices $V_0^i(K)$, we arrive at the stochastic volatility LIBOR MFM with the target driver (4.1) which recovers the market prices of co-terminal vanilla swaptions. Throughout this section we assume that the parameters μ_j, $j = 0, ..., n-1$, for the target driver (4.1) are given.

4.5.1 Calibration routine

Let us first derive a pricing formula for vanilla swaptions in the model (4.6). To do so we consider the dynamics of the forward swap rate $y^{i,n+1-i}$ for $i = 1, ..., n$. The forward swap rate $y^{i,n+1-i}$ can be expressed in terms of forward LIBORs:

$$y_t^{i,n+1-i} = \sum_{k=i}^{n} w_k^i(t) L_t^k$$

with the LIBOR dependent weights

$$w_k^i(t) = \frac{\alpha_k D_{tT_{k+1}}}{\sum_{s=i}^{n} \alpha_s D_{tT_{s+1}}}.$$

Notice that the weights w_k^i's depend on the forward LIBORs so that it is not straightforward to derive the dynamics of swap rates. Jackel and Rebonato [41] showed that the variation of the weights w_k^i's is insignificant compared to the variation of LIBORs. This leads to an approximation of the weights w_k^i's by their initial values and therefore we have that

$$y_t^{i,n+1-i} \approx \sum_{k=i}^{n} w_k^i(0) L_t^k. \tag{4.19}$$

with

$$w_k^i(0) = \frac{\alpha_k D_{0T_{k+1}}}{\sum_{s=i}^{n} \alpha_s D_{0T_{s+1}}}.$$

Note that this freezing approximation is familiar to practitioners and used often in the context of LMMs in the literature (see [2], [54] and [66]). From the SDE of the forward LIBORs (4.6), we apply Itô's lemma to (4.19) and derive the dynamics of the forward swap rate $y^{i,n+1-i}$ under the swaption measure $\mathbb{S}^{i,n+1-i}$ corresponding to the numeraire PVBP $P^{i,n+1-i}$.

Lemma 2. *Let L^i satisfies SDEs (4.6). In the swaption measure $\mathbb{S}^{i,n+1-i}$, we have that*

$$dy_t^{i,n+1-i} \approx \sum_{k=i}^{n} w_k^i(0) \sigma_t^k (L_t^k)^\beta dW_t^{i,n+1-i} \tag{4.20}$$

$$d\sigma_t = \tilde{\mu}^i(t, \sigma_t, y_t)\sigma_t dt + v\sigma_t dB_t^{i,n+1-i} \qquad v > 0 \tag{4.21}$$

$$dW_t^{i,n+1-i} dB_t^{i,n+1-i} = \rho dt \qquad \rho \in [-1, 1],$$

with

$$\tilde{\mu}^i(t,\sigma_t,y_t) := \mu_t + \upsilon\rho\sum_{j=i+1}^{n}\frac{\Psi_t^{j-1}\hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1}\hat{P}_t^{i,n+1-i}}(\frac{\gamma^j\alpha_{j-1}(y_t^{j,n+1-j})^\beta}{1+\alpha_{j-1}y_t^{j,n+1-j}})\sigma_t, \tag{4.22}$$

$$\Psi_t^i := \prod_{j=1}^{i}(1+\alpha_j y_t^{j+1,n-j}), \tag{4.23}$$

$$\hat{P}_t^{i,n+1-i} := \frac{P_t^{i,n+1-i}}{D_{t,T_{n+1}}}. \tag{4.24}$$

where $W^{i,n+1-i}$ and $B^{i,n+1-i}$ are correlated Brownian motions in the swaption measure $\mathbb{S}^{i,n+1-i}$.

Proof. Let us first derive the dynamics of the stochastic volatility σ in the swaption measure $\mathbb{S}^{i,n+1-i}$. The Brownian motion B^{n+1} in the terminal measure can be expressed as

$$dB_t^{n+1} = \rho dW_t^{n+1} + \sqrt{1-\rho^2}dZ_t^{n+1},$$

where Z^{n+1} and W^{n+1} are independent Brownian motions in the terminal measure $\mathbb{F}^{n+1}$. Define ς_t by

$$\varsigma_t := \left.\frac{d\mathbb{S}^{i,n+1-i}}{d\mathbb{F}^{n+1}}\right|_{\mathcal{F}_t} = \frac{P_t^{i,n+1-i}/P_0^{i,n+1-i}}{D_{tT_{n+1}}/D_{0T_{n+1}}},$$

where $\{\mathcal{F}_t\}$ is the filtration generated by Z^{n+1} and W^{n+1}. After calculation, we have that

$$d\varsigma_t = \varsigma_t\sum_{j=i+1}^{n}\frac{\Psi_t^{j-1}\hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1}\hat{P}_t^{i,n+1-i}}(\frac{\gamma^j\alpha_{j-1}(y_t^{j,n+1-j})^\beta}{1+\alpha_{j-1}y_t^{j,n+1-j}})\sigma_t dW_t^{n+1}.$$

From the Girsanov theorem we have that the processes

$$\begin{aligned} dW_t^{i,n+1-i} &:= dW_t^{n+1} - \frac{d\varsigma_t}{\varsigma_t}\cdot dW_t^{n+1} \\ &= dW_t^{n+1} - \sum_{j=i+1}^{n}\frac{\Psi_t^{j-1}\hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1}\hat{P}_t^{i,n+1-i}}(\frac{\gamma^j\alpha_{j-1}(y_t^{j,n+1-j})^\beta}{1+\alpha_{j-1}y_t^{j,n+1-j}})\sigma_t dt \end{aligned}$$

and

$$\begin{aligned} dZ_t^{i,n+1-i} &:= dZ_t^{n+1} - \frac{d\varsigma_t}{\varsigma_t} \cdot dZ_t^{n+1} \\ &= dZ_t^{n+1} \end{aligned}$$

are independent Brownian motions in $\mathbb{S}^{i,n+1-i}$. Thus we have that

$$\begin{aligned} dB_t^{n+1} &= \rho dW_t^{i,n+1-i} + \sqrt{1-\rho^2} dZ_t^{i,n+1-i} \qquad (4.25) \\ &+ \rho \sum_{j=i+1}^{n} \frac{\Psi_t^{j-1} \hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1} \hat{P}_t^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1} (y_t^{j,n+1-j})^\beta}{1+\alpha_{j-1} y_t^{j,n+1-j}}\right) \sigma_t dt, \\ &= dB_t^{i,n+1-i} + \rho \sum_{j=i+1}^{n} \frac{\Psi_t^{j-1} \hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1} \hat{P}_t^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1} (y_t^{j,n+1-j})^\beta}{1+\alpha_{j-1} y_t^{j,n+1-j}}\right) \sigma_t dt, \qquad (4.26) \end{aligned}$$

where $B^{i,n+1-i}$ is a Brownian motion in $\mathbb{S}^{i,n+1-i}$ defined as

$$dB_t^{i,n+1-i} = \rho dW_t^{i,n+1-i} + \sqrt{1-\rho^2} dZ_t^{i,n+1-i}.$$

Following equation (4.26), we have the dynamics of σ in the swaption measure $\mathbb{S}^{i,n+1-i}$:

$$d\sigma_t = [\mu_t + v\rho \sum_{j=i+1}^{n} \frac{\Psi_t^{j-1} \hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1} \hat{P}_t^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1} (y_t^{j,n+1-j})^\beta}{1+\alpha_{j-1} y_t^{j,n+1-j}}\right) \sigma_t] \sigma_t dt + v\sigma_t dB_t^{i,n+1-i} \quad v > 0.$$

On the other hand, since $y^{i,n+1-i}$ is a martingale in the swaption measure $\mathbb{S}^{i,n+1-i}$, by applying Itô's lemma to (4.19), the dynamics of $y^{i,n+1-i}$ approximately follows the following driftless SDE

$$dy_t^{i,n+1-i} \approx \sum_{k=i}^{n} w_k^i(0) \sigma_t^k (L_t^k)^\beta dW_t^{i,n+1-i}.$$

The result then follows. □

The SDE (4.20) shows that the volatility term of swap rates depends on forward LIBORs. We now introduce an approximation to write a SDE for the swap rate where the volatility is just a function of the swap rate itself. Following (4.20)

we have that

$$\begin{aligned} dy_t^{i,n+1-i} &\approx \sum_{k=i}^{n} w_k^i(0)\sigma_t^k (L_t^k)^\beta dW_t^{i,n+1-i} \\ &= \sum_{k=i}^{n} w_k^i(0)(\frac{L_t^k}{y_t^{i,n+1-i}})^\beta \sigma_t^k (y_t^{i,n+1-i})^\beta dW_t^{i,n+1-i} \\ &\approx \sum_{k=i}^{n} W_k^i(0)\gamma^k \sigma_t (y_t^{i,n+1-i})^\beta dW_t^{i,n+1-i} \qquad (4.27) \\ &= \xi^i \sigma_t (y_t^{i,n+1-i})^\beta dW_t^{i,n+1-i} \qquad (4.28) \end{aligned}$$

with

$$\begin{aligned} W_k^i(0) &= w_k^i(0)(\frac{L_0^k}{y_0^{i,n+1-i}})^\beta \\ \xi^i &= \sum_{k=i}^{n} W_k^i(0)\gamma^k. \qquad (4.29) \end{aligned}$$

where (4.27) is obtained by approximating the time-dependent ratio $(L_t^k/y_t^{i,n+1-i})^\beta$ by a constant ratio $(L_0^k/y_0^{i,n+1-i})^\beta$. This approximation is justified by an observation made by Andersen and Andreasen [2] that the time-dependent ratio $(L_t^k/y_t^{i,n+1-i})^\beta$ has a very low variability and is close to a constant. Thus the SDE for a swap rate can be rewritten in the form of the SABR Swap Market model (SMM).

Lemma 3. *Let L^i satisfies SDEs (4.6). Under the swaption measure $\mathbb{S}^{i,n+1-i}$ the dynamics of $y_t^{i,n+1-i}$ can be approximated by*

$$\begin{aligned} dy_t^{i,n+1-i} &\approx \sigma_t^i (y_t^{i,n+1-i})^\beta dW_t^{i,n+1-i} \qquad (4.30) \\ d\sigma_t^i &= \tilde{\mu}^i(t,\sigma_t,y_t)\sigma_t^i dt + v\sigma_t^i dB_t^{i,n+1-i} \qquad \sigma_t^i = \xi^i \sigma_t, \quad v > 0 \\ dW_t^{i,n+1-i} dB_t^{i,n+1-i} &= \rho dt \qquad \rho \in [-1,1], \end{aligned}$$

where $\tilde{\mu}^i(t,\sigma_t,y_t)$ is given by equation (4.22), and ξ^i is given by equation (4.29).

Proof. The result follows from SDEs (4.28) and (4.21) by letting $\sigma_t^i := \xi^i \sigma_t$. □

We can see from the above lemma that the SDEs (4.30) for the swap rates are in the form similar to the SABR model (4.7) except for the drift of the volatility, which is a function of the forward swap rates and the volatility itself. In order to derive an approximate SABR formula for the model (4.30), we derive an approximate dynamics in the form of (4.7). To do so we apply the following two-step

approximation which is similar to that used in Section 4.4.2.

1. **First step: volatility drift approximation**. We first apply the following freezing approximation to the drift function $\tilde{\mu}^i(t, \sigma_t, y_t)$:

$$\tilde{\mu}^i(t, \sigma_0, y_0) := \mu_t + v\rho \sum_{j=i+1}^{n} \frac{\Psi_0^{j-1} \hat{P}_0^{j,n+1-j}}{\Psi_0^{i-1} \hat{P}_0^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1} (y_0^{j,n+1-j})^\beta}{1 + \alpha_{j-1} y_0^{j,n+1-j}}\right) \sigma_0 \approx \tilde{\mu}^i(t, \sigma_t, y_t). \tag{4.31}$$

Hence we obtain the following SABR-SMM with the log-Normal dynamics of the volatility under the associated swaption measure

$$\begin{aligned} dy_t^{i,n+1-i} &\approx \sigma_t^i (y_t^{i,n+1-i})^\beta dW_t^{i,n+1-i} \\ d\sigma_t^i &\approx \tilde{\mu}^i(t, \sigma_0, y_0)\sigma_t^i dt + v\sigma_t^i dB_t^{i,n+1-i} \qquad \sigma_t^i = \xi^i \sigma_t, \quad v > 0 \\ dW_t^{i,n+1-i} dB_t^{i,n+1-i} &= \rho dt \qquad \rho \in [-1, 1]. \end{aligned} \tag{4.32}$$

2. **Second step: SABR formula for non-zero drift**. The SABR formula is derived for the SABR model of the form (4.7) where the volatility process is driftless. In this step we approximate the model (4.32) by the following model where the volatility process is driftless

$$\begin{aligned} dy_t^{i,n+1-i} &\approx \tilde{\sigma}_t^i (y_t^{i,n+1-i})^\beta dW_t^{i,n+1-i} \\ d\tilde{\sigma}_t^i &= v\tilde{\sigma}_t^i dB_t^{i,n+1-i} \qquad v > 0 \\ dW_t^{i,n+1-i} dB_t^{i,n+1-i} &= \rho dt \qquad \rho \in [-1, 1]. \end{aligned} \tag{4.33}$$

such that the co-terminal swaption prices given by the two models are approximately matched. Note that the distribution of the swap rate $y_{T_i}^{i,n+1-i}$ depends on the distribution of the volatility $\sigma_{T_i}^i$ at expiration time T_i, but it also depends on the path of the volatility from current time to time T_i. As justified by Morini and Mercurio [54], it is accurate to assume that the distribution of the swap rate $y_{T_i}^{i,n+1-i}$ depends on the average volatility from now to time T_i. Thus we choose the initial value $\tilde{\sigma}_0^i$ such that

$$E_{\mathbb{S}^{i,n+1-i}}[\int_0^{T_i} \sigma_t^i dt] = E_{\mathbb{S}^{i,n+1-i}}[\int_0^{T_i} \tilde{\sigma}_t^i dt].$$

This leads to the value of $\tilde{\sigma}_0^i$ which is given by

$$\tilde{\sigma}_0^i = \frac{\xi^i \sigma_0 \int_0^{T_i} e^{\int_0^t \tilde{\mu}^i(s, \sigma_0, y_0) ds} dt}{T_i}. \tag{4.34}$$

Finally we obtain the SABR formula $\sigma_B(K, y_0^{i,n+1-i}; \tilde{\sigma}_0^i, \beta, \nu, \rho)$ given by (4.8) for the co-terminal swaption on the swap rate $y^{i,n+1-i}$ with strike K.

Note that we only have free parameters ξ^i or equivalently $\tilde{\sigma}_0^i$ and this just allows us to calibrate to market implied volatilities with respect to one strike K^i of the swaption for $i = 1, ..., n$. As we discussed one common choice is to calibrate to the ATM implied volatilities, i.e. $K^i = y_0^{i,n+1-i}$. One may also choose the strike K^i according to the product we price. For instance, when we price a Bermudan swaption, which will be discussed in Chapter 5, with strike K^{Bem} it is natural to calibrate to the market implied volatilities of the corresponding swaptions struck at the strike K^{Bem} since this is the most relevant to the price of the Bermudan swaption. In this case we have that $K^i = K^{Bem}$ for $i = 1, ..., n$.

Suppose we are given a set of market implied volatilities $\sigma^{mkt}(K^i, y_0^{i,n+1-i})$ with respect to the strike K^i of the swaption on the swap rate $y^{i,n+1-i}$. We choose value of ξ^i such that

$$\sigma_B(K^i, y_0^{i,n+1-i}; \tilde{\sigma}_0^i, \beta, \nu, \rho) = \sigma^{mkt}(K^i, y_0^{i,n+1-i}), \tag{4.35}$$

where $\tilde{\sigma}_0^i$ is given by equation (4.34). Once we have determined parameters ξ^i's, parameters γ^i's are immediate by equation (4.29).

Having determined the parameters γ^i's, we end up with a SABR-LMM (4.6) or equivalently (4.9) that can recover the market prices of a set of European swaptions. Finally we can find the prices of digital caplets produced by the model (4.9) and feed them into the stochastic volatility LIBOR MFM. As we discussed in the previous section, we do so by using Monte Carlo methods.

4.5.2 Numerical study

In this subsection, we investigate the accuracy of the calibration routine developed in Section 4.5.1. The numerical study we perform here is similar to the one we did in Section 4.4.3. In particular given a set of market implied volatilities of the ATM swaptions on 17 October 2007, we determine the parameters ξ^i for $i = 1, ..., n$ via (4.35) by using the SABR formula. We then give the implied volatilities of the co-terminal swaptions produced by the approximate SABR-SMM (4.33), using Monte Carlo methods, and compare them with those produced by the SABR formula within the model (4.33) which is the same as the market. By doing so we can investigate the accuracy of the SABR formula. By comparing the implied volatilities given by the model (4.33) to those produced by the SABR-SMM (4.30), where we use Monte Carlo methods for both models, we can distinguish the effect of the "two-

step approximation". Finally we set up the stochastic volatility LIBOR MFM and produce the MFM model implied volatilities of the co-terminal swaptions.

In our numerical study, we consider the swaption calibration performance of the SV MFM with five different sets of parameters. We provide the numerical result in Table 4.2. The column "Market" is the market implied volatilities $\sigma^{mkt}(y_0^{i,n+1-i}, y_0^{i,n+1-i})$ of the ATM co-terminal swaption on 17 October 2007. The columns "Approx SABR-SMM (MC)" and "SABR-SMM (MC)" are implied volatilities of the ATM co-terminal swaptions produced by the approximate SABR-SMM (4.33) and the SABR-SMM (4.30) respectively using Monte Carlo methods. The last column "MFM" is the implied volatilities produced by the stochastic volatility LIBOR MFM. We draw the following conclusions from the numerical results.

1. The implied volatilities given by the MFM are close to the market implied volatilities for short maturities. The calibration routine loses accuracy for longer maturities.

2. By comparing the market implied volatilities with those produced by the approximate SABR-SMM (4.33), we can see that most calibration error is coming from the SABR formula. The SABR formula starts losing accuracy for longer maturities.

3. We compare the implied volatilities produced by the models (4.33) and (4.30), and it turns out that the "two-step approximation" developed in the previous subsection performs well for short maturities but loses some precision for $T >$ 20 years.

4. We can see from the data in the last two columns that the calibration of the stochastic volatility LIBOR MFM performs well for all scenarios.

Overall the accuracy of the calibration routine developed in this section is satisfactory for short maturities and the calibration error is mainly from the SABR formula.

4.6 Calibration to market correlations

In this section we discuss what to do with the remaining model parameters μ_j, $j = 0, ..., n-1$, of the target SABR driving process (4.1). We will choose μ_j's by calibrating to the market correlation structure of the LIBORs $\{L^i_{T_i}; i = 1, ..., n\}$. By doing so, the model will have a desirable hedging behaviour which will be discussed in Chapter 5. Note that here we focus on a LIBOR MFM, but it is straightforward to adapt the approach for the swap version of stochastic volatility MFMs where

we determine the model parameters μ_j's by calibrating to the market correlations of swap rates. In Section 4.6.1 we develop a calibration routine to determine the parameters μ_j's. In Section 4.6.2 we will investigate the accuracy of this calibration routine numerically.

4.6.1 Calibration routine

In this subsection we develop a calibration routine to choose the model parameters μ_j, $j = 0, ..., n-1$, of the target SABR driving process (4.1) by calibrating to the market correlation structure of the LIBORs $\{L^i_{T_i}; i = 1, ..., n\}$. To do so we derive an approximate formula for the correlations of the LIBORs $\{L^i_{T_i}; i = 1, ..., n\}$ in the stochastic volatility LIBOR MFM. In our approach we take the link between the separable LMMs and LIBOR MFMs as a guide. In particular we remember that the separable SABR-LMM (4.6) is expected to be numerically similar to the stochastic volatility LIBOR MFM with the SABR driving process (4.1) together with the input prices of digital caplets derived from the SDE (4.6). Thus we first derive an approximate formula for the correlation of LIBORs at their setting dates in the separable SABR-LMM (4.6), and then we borrow this formula for our stochastic volatility LIBOR MFM because of the link.

Now we derive an approximate formula for the correlations of the LIBORs $\{L^i_{T_i}; i = 1, ..., n\}$ under the separable SABR-LMM (4.6). The derivation involves some rough approximation. We will investigate the accuracy of the approximate formula numerically later in this section. We will see that the performance of this approximate formula is satisfactory.

The first approximation we adopt here is freezing the drift in the LIBORs L^i, $i = 1, ..., n-1$, of the separable SABR-LMM (4.6) under the terminal measure:

$$-\sum_{j=i+1}^{n}(\frac{\gamma^j\alpha_j(L^j_t)^\beta}{1+\alpha_j L^j_t})\gamma^i\sigma_t^2 \approx -\sum_{j=i+1}^{n}(\frac{\gamma^j\alpha_j(L^j_0)^\beta}{1+\alpha_j L^j_0})\gamma^i\sigma_0^2 =: \bar{\mu}^i_0.$$

This approximation leads to the following forward LIBORs dynamics under the terminal measure

$$\begin{aligned}
dL^i_t &= \bar{\mu}^i_0(L^i_t)^\beta dt + \gamma^i\sigma_t(L^i_t)^\beta dW^{n+1}_t \quad i = 1, ..., n-1, \\
dL^n_t &= \gamma^n\sigma_t(L^n_t)^\beta dW^{n+1}_t \qquad \beta \in [0,1], \\
d\sigma_t &= \mu_t\sigma_t dt + v\sigma_t dB^{n+1}_t \qquad v > 0, \\
dW^{n+1}_t dB^{n+1}_t &= \rho dt \qquad \rho \in [-1,1].
\end{aligned} \tag{4.36}$$

Note that the above drift approximation introduces arbitrage.

The second approximation we apply here is to approximate the model (4.36) by a stochastic volatility displaced-diffusion (DD) model. In particular applying a Taylor expansion to $(L_t^i)^\beta$:

$$\begin{aligned}(L_t^i)^\beta &\approx (L_0^i)^\beta + \beta(L_0^i)^{\beta-1}(L_t^i - L_0^i) \\ &= \delta_i(L_t^i + \theta_i),\end{aligned} \tag{4.37}$$

and inserting (4.37) into (4.36), we arrive at the following stochastic volatility displaced-diffusion model

$$\begin{aligned} dL_t^i &= \delta_i \bar{\mu}_0^i (L_t^i + \theta_i)dt + \delta_i \gamma^i \sigma_t (L_t^i + \theta_i) dW_t^{n+1} \\ dL_t^n &= \delta_n \gamma^n \sigma_t (L_t^n + \theta_n) dW_t^{n+1} \\ d\sigma_t &= \mu_t \sigma_t dt + v\sigma_t dB_t^{n+1} \qquad v > 0, \\ dW_t^{n+1} dB_t^{n+1} &= \rho dt \qquad \rho \in [-1, 1] \end{aligned} \tag{4.38}$$

with the mapping

$$\begin{aligned}\delta_i &= \beta(L_0^i)^{\beta-1} \\ \theta_i &= L_0^i \frac{1-\beta}{\beta}.\end{aligned}$$

Note that the similarity between the CEV SABR model and the corresponding DD SABR model has been justified numerically by [47]. It follows from SDE (4.38) that

$$ln(L_{T_i}^i + \theta_i) = ln(L_0^i + \theta_i) + \delta_i \bar{\mu}_0^i T_i - \frac{1}{2}\delta_i^2 (\gamma^i)^2 \int_0^{T_i} \sigma_t^2 dt + \delta_i \gamma^i \int_0^{T_i} \sigma_t dW_t^{n+1}.$$

When the variability of the term $\int_0^{T_i} \sigma_t^2 dt$ is negligible compared to the variability of the random variable $\int_0^{T_i} \sigma_t dW_t^{n+1}$, it is safe to approximate the term $\int_0^{T_i} \sigma_t^2 dt$ by its expectation so that we have the following approximation

$$ln(L_{T_i}^i + \theta_i) \approx b_{T_i}^i + \delta_i \gamma^i \int_0^{T_i} \sigma_t dW_t^{n+1}, \tag{4.39}$$

where

$$b_{T_i}^i := ln(L_0^i + \theta_i) + \delta_i \bar{\mu}_0^i T_i - \frac{1}{2}\delta_i^2 (\gamma^i)^2 E_{\mathbb{F}^{n+1}}\left(\int_0^{T_i} \sigma_t^2 dt\right)$$

is a deterministic function. Since correlation is unchanged by a linear transforma-

tion, we have the following approximate formula for the correlation of LIBORs

$$Corr(ln(L^i_{T_i}+\theta_i), ln(L^j_{T_j}+\theta_j)) \approx Corr(\int_0^{T_i}\sigma_t dW_t^{n+1}, \int_0^{T_j}\sigma_t dW_t^{n+1}). \quad (4.40)$$

Note that the derivation of this formula involves some rough approximation. But we will see later in a numerical investigation that the performance of this formula is satisfactory when the parameter Volvol ν is not too big. Furthermore we can find a formula for the correlation

$$Corr(\int_0^{T_i}\sigma_t dW_t^{n+1}, \int_0^{T_j}\sigma_t dW_t^{n+1})$$

under the separable SABR-LMM (4.6) in terms of the parameters μ_s for $s = 0, ..., n-1$.

Proposition 4. *Consider the separable SABR-LMM (4.6). The correlation between* $\int_0^{T_i}\sigma_t dW_t^{n+1}$ *and* $\int_0^{T_j}\sigma_t dW_t^{n+1}$, $1 \le i < j \le n$, *satisfies*

$$Corr(\int_0^{T_i}\sigma_t dW_t^{n+1}, \int_0^{T_j}\sigma_t dW_t^{n+1}) = \sqrt{\frac{\zeta_{T_i}}{\zeta_{T_j}}}$$

where

$$\begin{aligned}\zeta_{T_i} :=& Var(\int_0^{T_i}\sigma_t dW_t^{n+1}) \\ =& \sum_{j=0}^{i-1}\frac{\sigma_0^2}{2\mu_j+\nu^2}[\exp(2\int_0^{T_{j+1}}\mu_s ds + \nu^2 T_{j+1}) - \exp(2\int_0^{T_j}\mu_s ds + \nu^2 T_j)]. \quad (4.41)\end{aligned}$$

Proof. See Appendix 4.A. □

So far we have achieved the following approximate formula for the correlations of LIBORs at their setting dates under the separable SABR-LMM (4.6):

$$Corr(ln(L^i_{T_i}+\theta_i), ln(L^j_{T_j}+\theta_j)) \approx \sqrt{\frac{\zeta_{T_i}}{\zeta_{T_j}}}. \quad (4.42)$$

where ζ_{T_i} is given by equation (4.41). Based on the link between the separable SABR-LMM and the stochastic volatility LIBOR MFM, we borrow the approximate formula (4.42) for our stochastic volatility LIBOR MFM with the target SABR driving process (4.1). Now we can choose the remaining parameters μ_j, $j = 0, ..., n-1$, in the target SABR driving process (4.1) by calibrating to the

market correlation of LIBORs at their setting dates. We note that in a low factor model we cannot capture the whole correlation matrix. In particular the piecewise constant drift μ_j, $j = 0, ..., n-1$, can fully determine up to n correlations of the LIBORs $\{L^i_{T_i}; i = 1, ..., n\}$ so that we cannot capture all the market correlations $Corr(ln(L^i_{T_i} + \theta_i), ln(L^j_{T_j} + \theta_j))$ for $1 \leq i < j \leq n$. In our approach we calibrate to the one-step correlation of the LIBORs $Corr(ln(L^i_{T_i} + \theta_i), ln(L^{i+1}_{T_{i+1}} + \theta_{i+1}))$ for $1 \leq i \leq n-1$.

Suppose we are given a set of market one-step correlations of LIBORs

$$Corr^{mkt}(ln(L^i_{T_i} + \theta_i), ln(L^{i+1}_{T_{i+1}} + \theta_{i+1}))$$

for $i = 1, ..., n-1$. Note that the market correlations are not observable directly in the market. We have to estimate them using the swaption matrix from the market. For details the reader is referred to [48]. The parameters μ_j, $j = 0, ..., n-1$, can be chosen as follows.

1. The algorithm for fixing ζ_{T_i}'s works back iteratively from T_n. Without loss of generality, we set

$$\zeta_{T_n} = \frac{\sigma_0^2}{\nu^2}(e^{\nu^2 T_n} - 1).$$

2. Suppose we now reach T_i, having fixed ζ_{T_j} for $j = i+1, ..., n$. Given the market one-step correlation of LIBORs we can fix ζ_{T_i} by the approximate formula (4.40)

$$Corr^{mkt}(ln(L^i_{T_i} + \theta_i), ln(L^{i+1}_{T_{i+1}} + \theta_{i+1})) \approx \sqrt{\frac{\zeta_{T_i}}{\zeta_{T_{i+1}}}}$$

so that

$$\zeta_{T_i} = \zeta_{T_{i+1}}[Corr^{mkt}(ln(L^i_{T_i} + \theta_i), ln(L^{i+1}_{T_{i+1}} + \theta_{i+1}))]^2.$$

3. Once we have determined ζ_{T_i} for $i = 1, ..., n$, we can fix μ_j's by working forward iteratively from T_1. At T_1, with knowledge of ζ_{T_1} we choose μ_0 such that

$$\zeta_{T_1} = \frac{\sigma_0^2}{2\mu_0 + \nu^2}[\exp(2\mu_0 T_1 + \nu^2 T_1) - 1],$$

where μ_0 can be found by using, for instance, the bisection method.

4. Suppose we have now reached T_i, having fixed μ_j for $j = 0, ..., i-1$. We choose

μ_i such that

$$\zeta_{T_i+1} - \zeta_{T_i} = \frac{\sigma_0^2}{2\mu_i + \nu^2}[\exp(2\int_0^{T_{i+1}} \mu_s ds + \nu^2 T_{i+1}) - \exp(2\int_0^{T_i} \mu_s ds + \nu^2 T_i)].$$

Remark 8. *The correlation formula (4.42) is independent of the parameters γ^i's. This means that the correlation calibration of choosing μ_j's can precede the calibration routines developed in Sections 4.4 and 4.5. This benefit is due to the separation of the specification of driver and marginals.*

We have introduced an algorithm for calibrating to the market one-step correlation of LIBORs by choosing piecewise constants μ_j's. The calibration procedure also applies to the swap rate version of stochastic volatility MFM where we can calibrate to the market one-step correlations of swap rates. The details can be found in Appendix 4.B.

4.6.2 Numerical study

Having obtained the calibration formulas for LIBORs correlation in the SV LIBOR MFM, we investigate the calibration performance in this subsection. In our numerical study, we consider five sets of parameters which are shown in Table 4.3. Suppose we calibrate to the correlations

$$Corr^{mkt}(ln(L^i_{T_i} + \theta_i), ln(L^{i+1}_{T_{i+1}} + \theta_{i+1}))$$

where

$$\theta_i = L^i_0 \frac{1-\beta}{\beta}$$

for $i = 1, 2, ..., 29$. We generate them in the column "Market" in Table 4.4. Following the calibration routine developed in the previous subsection, we will see whether the one-step correlation of LIBORs at their setting dates produced by stochastic volatility LIBOR MFMs can match the target correlation in the column "Market". The numerical result is given in Table 4.4. The columns "MFM (I)" - "MFM (V)" are model implied correlations

$$Corr^{mod}(ln(L^i_{T_i} + \theta_i), ln(L^{i+1}_{T_{i+1}} + \theta_{i+1})).$$

produced by the stochastic volatility LIBOR MFMs with the set of parameters corresponding to the scenarios in Table 4.3.

We can see from Table 4.4 that for the first scenario where the parameter $\nu = 0.1$, the one-step correlation of LIBORs produced by the stochastic volatility LIBOR MFM is very close to the target so that the calibration performs quite well. For the second and third scenarios where the parameter ν increases to 0.3, the calibration performance is good for maturities $T < 15$ but getting worse for longer maturities. In the stress test when $\nu = 0.5$, we can see that the calibration routine loses accuracy. This is because as the parameter ν increases, the linear relationship (4.39) starts breaking down.

4.7 Conclusion

In this chapter we have addressed calibration issues of the SV LIBOR MFM developed in Chapter 3. Due to the nature of MFMs where the driver is separated from the marginals, given a SABR driver any marginal distributions of LIBORs at their setting dates can be fed into the model. The separation of the driver and marginals provides flexibility in terms of calibration.

We considered a separable SABR-LMM which combines a SABR model and a separable LMM. We expected this model could be numerically similar to the stochastic volatility LIBOR MFM. Based on this link, given the SABR driving process, we chose an appropriate SABR type marginals to feed in the MFM which could lead to a desirable dynamics of future implied volatilities.

The link between separable SABR-LMMs and SV LIBOR MFMs also gives us a guide as to how to calibrate a stochastic volatility MFM. Based on the link, we developed a calibration routine to feed in the marginals by calibrating to market vanilla options. The resulting stochastic volatility LIBOR MFM can reproduce the market vanilla options. The numerical results indicate that the calibration performance is satisfactory for short maturities. By adopting an "iterative procedure", a precise calibration can be achieved.

Finally we considered the problem of how to determine the parameters of the driving process which is a modified SABR model. By modified SABR model we mean that there exists a drift function in the stochastic volatility process. The drift function which is assumed to be piecewise constant can capture the correlation structure of LIBORs. The other parameters of the driver can be determined exogenously by investigating historical market data or by traders based on their market judgment or beliefs. Based on the link between separable SABR-LMMs and SV LIBOR MFMs, we derived an approximate formula for the correlation of LIBORs which is an analytical formula. We chose the drift parameters of the SABR driving

process by calibrating to the market one step correlation of LIBORs at their setting dates via this formula. The numerical results indicate that when the vol of volatility is small, the calibration performance is very good. As the vol of volatility increases the calibration performance gets worse. For the stress test where the vol of volatility is very large, the calibration routine loses accuracy.

4.A Appendix: Proof of Proposition 4

Consider the separable SABR-LMM (4.6). The Brownian motion W^{n+1} can be expressed as

$$W_t^{n+1} = \rho B_t^{n+1} + \sqrt{1-\rho^2}\widehat{W}_t^{n+1}$$

where B^{n+1} and $\widehat{W}^{n+1}$ are independent Brownian motions under the terminal measure $\mathbb{F}^{n+1}$.

By definition we have that

$$Corr(\int_0^{T_i} \sigma_t dW_t^{n+1}, \int_0^{T_j} \sigma_t dW_t^{n+1}) = \frac{Cov(\int_0^{T_i} \sigma_t dW_t^{n+1}, \int_0^{T_j} \sigma_t dW_t^{n+1})}{\sqrt{Var(\int_0^{T_i} \sigma_t dW_t^{n+1}) Var(\int_0^{T_j} \sigma_t dW_t^{n+1})}},$$

where

$$\begin{aligned} &Cov(\int_0^{T_i} \sigma_t dW_t^{n+1}, \int_0^{T_j} \sigma_t dW_t^{n+1}) \\ =&Var(\int_0^{T_i} \sigma_t dW_t^{n+1}) + Cov(\int_0^{T_i} \sigma_t dW_t^{n+1}, \int_{T_i}^{T_j} \sigma_t dW_t^{n+1}). \end{aligned} \tag{4.43}$$

Denote $\mathcal{F}_T = \sigma(B_u : 0 \le u \le T)$ and $V_T = \int_0^{T_i} \sigma_t^2 dt$. We have that

$$Var(\int_0^{T_i} \sigma_t dW_t^{n+1}) = E_{\mathbb{F}^{n+1}}[(\int_0^{T_i} \sigma_t dW_t^{n+1})^2] - [E_{\mathbb{F}^{n+1}}(\int_0^{T_i} \sigma_t dW_t^{n+1})]^2,$$

where by Itô isometry

$$\begin{aligned} E_{\mathbb{F}^{n+1}}[(\int_0^{T_i} \sigma_t dW_t^{n+1})^2] &= \int_0^{T_i} E_{\mathbb{F}^{n+1}}(\sigma_t^2) dt \\ &= \sigma_0^2 \int_0^{T_i} \exp(2\int_0^t \mu_s ds + \nu^2 t) dt \\ &= \sum_{j=0}^{i-1} \int_{T_j}^{T_{j+1}} \frac{\sigma_0^2}{2\mu_t + \nu^2} d\exp(2\int_0^t \mu_s ds + \nu^2 t) \\ &= \sum_{j=0}^{i-1} \frac{\sigma_0^2}{2\mu_j + \nu^2} [\exp(2\int_0^{T_{j+1}} \mu_s ds + \nu^2 T_{j+1}) \\ &\quad - \exp(2\int_0^{T_j} \mu_s ds + \nu^2 T_j)] \end{aligned}$$

and

$$E_{\mathbb{F}^{n+1}}(\int_0^{T_i}\sigma_t dW_t^{n+1})=0.$$

Thus the first term in (4.43) is given by

$$\begin{aligned}Var(\int_0^{T_i}\sigma_t dW_t^{n+1})=\sum_{j=0}^{i-1}\frac{\sigma_0^2}{2\mu_j+\nu^2}[\exp(2\int_0^{T_{j+1}}\mu_s ds+\nu^2 T_{j+1})\\ -\exp(2\int_0^{T_j}\mu_s ds+\nu^2 T_j)],\end{aligned}$$

while the second term can be obtained by

$$\begin{aligned}&Cov(\int_0^{T_i}\sigma_t dW_t^{n+1},\int_{T_i}^{T_j}\sigma_t dW_t^{n+1})\\ =&E_{\mathbb{F}^{n+1}}(\int_0^{T_i}\sigma_t dW_t^{n+1}\int_{T_i}^{T_j}\sigma_t dW_t^{n+1})-0\\ =&E_{\mathbb{F}^{n+1}}[(\rho\int_0^{T_i}\sigma_t dB_t^{n+1}+\sqrt{1-\rho^2}\int_0^{T_i}\sigma_t d\widehat{W}_t^{n+1})\\ &\times(\rho\int_{T_i}^{T_j}\sigma_t dB_t^{n+1}+\sqrt{1-\rho^2}\int_{T_i}^{T_j}\sigma_t d\widehat{W}_t^{n+1})],\end{aligned}$$

where

$$\begin{aligned}&E_{\mathbb{F}^{n+1}}[\int_0^{T_i}\sigma_t dB_t^{n+1}\int_{T_i}^{T_j}\sigma_t d\widehat{W}_t^{n+1}]\\ &=E_{\mathbb{F}^{n+1}}[\int_0^{T_i}\sigma_t dB_t^{n+1}E_{\mathbb{F}^{n+1}}(\int_{T_i}^{T_j}\sigma_t d\widehat{W}_t^{n+1}|\mathcal{F}_{T_j})]\\ &=E_{\mathbb{F}^{n+1}}[\int_0^{T_i}\sigma_t dB_t^{n+1}\times 0]\\ &=0,\end{aligned}$$

and similarly

$$E_{\mathbb{F}^{n+1}}[\int_0^{T_i}\sigma_t d\widehat{W}_t^{n+1}\int_{T_i}^{T_j}\sigma_t dB_t^{n+1}]=0,$$

and by Itô isometry

$$\begin{aligned}
&E_{\mathbb{F}^{n+1}}[\int_0^{T_i}\sigma_t dB_t^{n+1}\int_{T_i}^{T_j}\sigma_t dB_t^{n+1}]\\
&=E_{\mathbb{F}^{n+1}}[\int_0^{T_i}\sigma_t dB_t^{n+1}E_{\mathbb{F}^{n+1}}(\int_{T_i}^{T_j}\sigma_t dB_t^{n+1}|\mathcal{F}_{T_j})]\\
&=0,
\end{aligned}$$

and

$$\begin{aligned}
&E_{\mathbb{F}^{n+1}}[\int_0^{T_i}\sigma_t d\widehat{W}_t^{n+1}\int_{T_i}^{T_j}\sigma_t d\widehat{W}_t^{n+1}]\\
&=E_{\mathbb{F}^{n+1}}[E_{\mathbb{F}^{n+1}}(\int_0^{T_i}\sigma_t d\widehat{W}_t^{n+1}\int_{T_i}^{T_j}\sigma_t d\widehat{W}_t^{n+1}|\mathcal{F}_{T_j})]\\
&=E_{\mathbb{F}^{n+1}}[E_{\mathbb{F}^{n+1}}(\int_0^{T_i}\sigma_t d\widehat{W}_t^{n+1}|\mathcal{F}_{T_j})E_{\mathbb{F}^{n+1}}(\int_{T_i}^{T_j}\sigma_t d\widehat{W}_t^{n+1}|\mathcal{F}_{T_j})]\\
&=0.
\end{aligned}$$

Therefore we have that

$$\begin{aligned}
&Cov(\int_0^{T_i}\sigma_t dW_t^{n+1},\int_0^{T_j}\sigma_t dW_t^{n+1})\\
=&Var(\int_0^{T_i}\sigma_t dW_t^{n+1})\\
=&\sum_{j=0}^{i-1}\frac{\sigma_0^2}{2\mu_j+\nu^2}[\exp(2\int_0^{T_{j+1}}\mu_s ds+\nu^2 T_{j+1})-\exp(2\int_0^{T_j}\mu_s ds+\nu^2 T_j)]
\end{aligned}$$

so that

$$Corr(\int_0^{T_i}\sigma_t dW_t^{n+1},\int_0^{T_j}\sigma_t dW_t^{n+1})=\sqrt{\frac{\zeta_{T_i}}{\zeta_{T_j}}}$$

where

$$\zeta_{T_i}=\sum_{j=0}^{i-1}\frac{\sigma_0^2}{2\mu_j+\nu^2}[\exp(2\int_0^{T_{j+1}}\mu_s ds+\nu^2 T_{j+1})-\exp(2\int_0^{T_j}\mu_s ds+\nu^2 T_j)].$$

4.B Appendix: Swap rate based stochastic volatility Markov-functional models: calibration

We have discussed the calibration of the LIBOR based stochastic volatility MFM. In this appendix we discuss briefly the calibration of the swap rate based stochastic volatility MFM. We will just present some main results without too much explanation since the discussion for the LIBOR version earlier also applies to the swap version.

The target (modified) SABR driving process for the stochastic volatility swap MFM is given by (4.1). Given a driver, a swap MFM can be set up by feeding in the input prices $V_0^i(K)$, $V_0^i : [0, \infty] \to \mathbb{R}$, of a set of PVBP-digital swaptions on the ith swap rate $y^{i,n+1-i}$ with strike $K \geq 0$ for $i = 1, ..., n$. Thus the stochastic volatility swap MFM with the target driver (4.1) is fully determined by

1. the input prices of PVBP-digital swaptions $V_0^i(K)$ as a function of the strike K for $i = 1, ..., n$.

2. the parameters of the target driver (4.1) which include σ_0, F_0, β, ρ, ν and μ_j for $j = 0, ..., n-1$.

Suppose the parameters σ_0, F_0, β, ρ and ν for (4.1) have been determined exogenously by investigating historical market data or by traders based on their market judgement or beliefs. Consequently the remaining parameters we have for the target driver are μ_j for $j = 0, ..., n-1$.

In Appendix 4.B.1 we decide on the marginal distributions of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$ which will be fed into the stochastic volatility swap MFM. We do so based on a link between the stochastic volatility swap MFM and the separable SABR-swap market model (SABR-SMM). We will use this link to find the input prices $V_0^i(K)$, $i = 1, ..., n$, by calibrating to the market prices of vanilla swaptions in Appendix 4.B.2. This link also provides us a guide as to how to determine the parameters μ_j, $j = 0, ..., n-1$, by calibrating to the market correlation structure of swap rates, which will be discussed in Appendix 4.B.3.

4.B.1 Choice of marginals

Consider the following separable SABR-SMM in the terminal measure $\mathbb{F}^{n+1}$:

$$dy_t^{i,n+1-i} = -\sum_{j=i+1}^{n} \frac{\Psi_t^{j-1}\hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1}\hat{P}_t^{i,n+1-i}}\left(\frac{\gamma^j\alpha_{j-1}(y_t^{j,n+1-j})^\beta}{1+\alpha_{j-1}y_t^{j,n+1-j}}\right)\gamma^i\sigma_t^2(y_t^{i,n+1-i})^\beta dt \tag{4.44}$$

$$\begin{aligned}
&+\gamma^i\sigma_t(y_t^{i,n+1-i})^\beta dW_t^{n+1} \quad i=1,...,n-1,\\
dy_t^{n,1} &= \gamma^n\sigma_t(y_t^{n,1})^\beta dW_t^{n+1} \qquad \beta\in[0,1],\\
d\sigma_t &= \mu_t\sigma_t dt + v\sigma_t dB_t^{n+1} \qquad v>0,\\
dW_t^{n+1}dB_t^{n+1} &= \rho dt \qquad \rho\in[-1,1].
\end{aligned}$$

where μ_t, Ψ_t^i and $\hat{P}_t^{i,n+1-i}$ are given by (4.2), (4.23) and (4.24) respectively. This model is expected to be numerically similar to the stochastic volatility swap MFM with the target SABR driving process (4.1) together with the prices $V_0^i(K)$, $i=1,...,n$, for PVBP-digital swaptions produced by the model (4.44).

Following the explanations in Section 4.3.2, we feed in the prices $V_0^i(K)$, $i=1,...,n$, of PVBP-digital swaptions with strike $K\geq 0$ produced by the model (4.44). By doing so we arrive at a stochastic volatility swap MFM which is expected to be similar to the separable SABR-SMM (4.44). More importantly, it will provide us a guide as to how to calibrate the stochastic volatility swap MFM, and in what follows we will develop a calibration routine.

4.B.2 Calibration to swaptions

So far we have decided to feed the input prices $V_0^i(K)$, $i=1,...,n$, of PVBP-digital swaptions with strike $K\geq 0$ produced by the model (4.44) into our stochastic volatility swap MFM with the target SABR driver (4.1). But when we calculate the prices $V_0^i(K)$, we still have free parameters γ^i's for the SDE (4.44) that need to be determined. We will determine the parameters γ^i's by fitting co-terminal swaption volatilities. Once we have fixed the parameters γ^i's, we can find the input prices $V_0^i(K)$, $i=1,...,n$, produced by the model (4.44). Finally we feed these input prices into our stochastic volatility swap MFM, and we can obtain a MFM which will recover co-terminal swaption volatilities . Before we proceed to the calibration routine, let us assume that the parameters μ_j, $j=0,...,n-1$, for the target driver (4.1) are given. We will give the details of how to determine μ_j's later in the next subsection.

In order to choose the parameters γ^i's by calibrating to the market prices of

swaptions, let us derive an approximate formula for the swaption prices within the model (4.44). To do so we derive the SABR-SMM under the associated swaption measure.

Lemma 4. *Let $y^{i,n+1-i}$ satisfies SDEs (4.44). Under the swaption measure $\mathbb{S}^{i,n+1-i}$ the dynamics of $y^{i,n+1-i}$ satisfies*

$$\begin{aligned} dy_t^{i,n+1-i} &= \sigma_t^i (y_t^{i,n+1-i})^\beta dW_t^{i,n+1-i} \qquad \beta \in [0,1], \\ d\sigma_t^i &= \hat{\mu}^i(t,\sigma_t,y_t)\sigma_t^i dt + v\sigma_t^i dB_t^{i,n+1-i} \qquad \sigma_t^i = \gamma^i \sigma_t \\ dW_t^{i,n+1-i} dB_t^{i,n+1-i} &= \rho dt \qquad \rho \in [-1,1], \end{aligned} \tag{4.45}$$

with

$$\hat{\mu}^i(t,\sigma_t,y_t) := \mu_t + v\rho \sum_{j=i+1}^{n} \frac{\Psi_t^{j-1}\hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1}\hat{P}_t^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1}(y_t^{j,n+1-j})^\beta}{1+\alpha_{j-1}y_t^{j,n+1-j}}\right)\sigma_t \tag{4.46}$$

where $W^{i,n+1-i}$ and $B^{i,n+1-i}$ are correlated Brownian motions in the swaption measure $\mathbb{S}^{i,n+1-i}$.

Proof. The Brownian motion B^{n+1} in the terminal measure can be expressed as

$$dB_t^{n+1} = \rho dW_t^{n+1} + \sqrt{1-\rho^2} dZ_t^{n+1},$$

where Z^{n+1} and W^{n+1} are independent Brownian motions in the terminal measure $\mathbb{F}^{n+1}$. Define ς_t by

$$\varsigma_t := \left.\frac{d\mathbb{S}^{i,n+1-i}}{d\mathbb{F}^{n+1}}\right|_{\mathcal{F}_t} = \frac{P_t^{i,n+1-i}/P_0^{i,n+1-i}}{D_{tT_{n+1}}/D_{0T_{n+1}}},$$

where $\{\mathcal{F}_t\}$ is the filtration generated by Z^{n+1} and W^{n+1}. After calculation, we have that

$$d\varsigma_t = \varsigma_t \sum_{j=i+1}^{n} \frac{\Psi_t^{j-1}\hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1}\hat{P}_t^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1}(y_t^{j,n+1-j})^\beta}{1+\alpha_{j-1}y_t^{j,n+1-j}}\right)\sigma_t dW_t^{n+1}.$$

From the Girsanov theorem we have that the processes

$$\begin{aligned} dW_t^{i,n+1-i} &:= dW_t^{n+1} - \frac{d\varsigma_t}{\varsigma_t}\cdot dW_t^{n+1} \\ &= dW_t^{n+1} - \sum_{j=i+1}^{n} \frac{\Psi_t^{j-1}\hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1}\hat{P}_t^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1}(y_t^{j,n+1-j})^\beta}{1+\alpha_{j-1}y_t^{j,n+1-j}}\right)\sigma_t dt \end{aligned}$$

and

$$\begin{aligned} dZ_t^{i,n+1-i} &:= dZ_t^{n+1} - \frac{d\varsigma_t}{\varsigma_t} \cdot dZ_t^{n+1} \\ &= dZ_t^{n+1} \end{aligned}$$

are independent Brownian motions in $\mathbb{S}^{i,n+1-i}$. Thus we have that

$$\begin{aligned} dB_t^{n+1} &= \rho dW_t^{i,n+1-i} + \sqrt{1-\rho^2} dZ_t^{i,n+1-i} \\ &\quad + \rho \sum_{j=i+1}^{n} \frac{\Psi_t^{j-1} \hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1} \hat{P}_t^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1} (y_t^{j,n+1-j})^\beta}{1 + \alpha_{j-1} y_t^{j,n+1-j}}\right) \sigma_t dt, \\ &= dB_t^{i,n+1-i} + \rho \sum_{j=i+1}^{n} \frac{\Psi_t^{j-1} \hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1} \hat{P}_t^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1} (y_t^{j,n+1-j})^\beta}{1 + \alpha_{j-1} y_t^{j,n+1-j}}\right) \sigma_t dt, \end{aligned}$$

where $B^{i,n+1-i}$ is a Brownian motion in $\mathbb{S}^{i,n+1-i}$ defined as

$$dB_t^{i,n+1-i} = \rho dW_t^{i,n+1-i} + \sqrt{1-\rho^2} dZ_t^{i,n+1-i}.$$

The result then follows from the SDE (4.44). □

Following the two-step approximation introduced in Section 4.4.2, we derive an approximate SABR formula $\sigma_B(K, y_0^{i,n+1-i}; \hat{\sigma}_0^i, \beta, \nu, \rho)$ given by (4.8) for the co-terminal swaption on the ith swap rate $y^{i,n+1-i}$ with strike K based on the model (4.45) where $\tilde{\sigma}_0^i$ is given by

$$\hat{\sigma}_0^i := \frac{\gamma^i \sigma_0 \int_0^{T_i} e^{\int_0^t \hat{\mu}^i(s,\sigma_0,y_0) ds} dt}{T_i}, \tag{4.47}$$

$$\hat{\mu}^i(t, \sigma_0, y_0) := \mu_t + \upsilon \rho \sum_{j=i+1}^{n} \frac{\Psi_0^{j-1} \hat{P}_0^{j,n+1-j}}{\Psi_0^{i-1} \hat{P}_0^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1} (y_0^{j,n+1-j})^\beta}{1 + \alpha_{j-1} y_0^{j,n+1-j}}\right) \sigma_0. \tag{4.48}$$

So far we have obtained an approximate SABR formula for the SABR-SMM (4.44) or equivalently (4.45). We can choose the parameters γ^i's such that the approximate SABR formula $\sigma_B(K, y_0^{i,n+1-i}; \hat{\sigma}_0^i, \beta, \nu, \rho)$ matches the market implied volatilities of the set of vanilla swaptions. Suppose we are given the market implied volatilities $\sigma^{mkt}(K^i, y_0^{i,n+1-i})$ of the co-terminal swaption on the ith swap rate $y^{i,n+1-i}$ struck at some particular strike K^i for each $i = 1, ..., n$. We can choose γ^i such that the approximate SABR formula for the SABR-LMM (4.44) matches the

market implied volatility $\sigma^{mkt}(K^i, y_0^{i,n+1-i})$:

$$\sigma_B(K^i, y_0^{i,n+1-i}; \hat{\sigma}_0^i, \beta, \nu, \rho) = \sigma^{mkt}(K^i, y_0^{i,n+1-i}),$$

for each $i = 1, ..., n$.

Having fixed the parameters γ^i, $i = 1, ..., n$, in the separable SABR-SMM (4.45), we can find the model implied prices $V_0^i(K)$, $i = 1, ..., n$, of the set of PVBP-digital swaptions as a function of the strike K. To do so, for this study we apply Monte Carlo methods. Finally we can feed the prices $V_0^i(K)$, $i = 1, ..., n$, into our stochastic volatility swap MFM, and the resulting model can recover the market implied volatilities $\sigma^{mkt}(K^i, y_0^{i,n+1-i})$, $i = 1, ..., n$, of the set of co-terminal swaptions.

4.B.3 Calibration to market correlations

In this subsection we choose parameters μ_j, $j = 0, ..., n-1$, for the target SABR driving process (4.1) by calibrating to the market one step correlation of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$. The calibration routine here is the same as the calibration routine developed in Section 4.6.1 with the correlation

$$Corr(ln(y_{T_i}^{i,n+1-i} + \theta_i), ln(y_{T_{i+1}}^{i+1,n-i} + \theta_{i+1}))$$

substituted for

$$Corr(ln(L_{T_i}^i + L_0^i \frac{1-\beta}{\beta}), ln(L_{T_{i+1}}^{i+1} + L_0^{i+1} \frac{1-\beta}{\beta}))$$

where

$$\theta_i := y_0^{i,n+1-i} \frac{1-\beta}{\beta}.$$

Since the calibration routines for the LIBOR based and swap rate based stochastic volatility MFMs are very similar, we will not provide the details here.

Expiry	Market	Approx SABR-LMM (MC)	SABR-LMM (MC)	MFM
$\nu = 0.1$, $\beta = 0.9$, $\rho = -0.8$, $\sigma_0 = 0.02$ and $\mu_j = 0$				
5Y	12.76	12.75	12.74	12.74
10Y	11.19	11.12	11.10	11.10
15Y	10.28	10.12	10.07	10.07
20Y	9.97	9.70	9.63	9.63
25Y	9.69	9.29	9.18	9.18
30Y	9.30	8.79	8.65	8.65
$\nu = 0.2$, $\beta = 0.7$, $\rho = -0.4$, $\sigma_0 = 0.04$ and $\mu_j = 0.01$				
5Y	12.76	12.75	12.74	12.74
10Y	11.19	11.08	11.05	11.05
15Y	10.28	10.04	9.98	9.98
20Y	9.97	9.58	9.47	9.47
25Y	9.69	9.14	9.00	9.00
30Y	9.30	8.57	8.40	8.40
$\nu = 0.3$, $\beta = 0.5$, $\rho = 0$, $\sigma_0 = 0.06$ and $\mu_j = 0.03$				
5Y	12.76	12.75	12.74	12.74
10Y	11.19	11.03	11.00	11.00
15Y	10.28	9.95	9.88	9.88
20Y	9.97	9.43	9.31	9.31
25Y	9.69	8.97	8.81	8.81
30Y	9.30	8.40	8.19	8.18
$\nu = 0.4$, $\beta = 0.3$, $\rho = 0.4$, $\sigma_0 = 0.08$ and $\mu_j = 0.01 \times j$				
5Y	12.76	12.74	12.72	12.72
10Y	11.19	10.96	10.91	10.91
15Y	10.28	9.87	9.78	9.78
20Y	9.97	9.33	9.16	9.15
25Y	9.69	8.82	8.61	8.60
30Y	9.30	8.19	7.93	7.91
$\nu = 0.5$, $\beta = 0.1$, $\rho = 0.8$, $\sigma_0 = 0.1$ and $\mu_j = 0.3 - 0.01 \times j$				
5Y	12.76	12.73	12.71	12.71
10Y	11.19	10.91	10.84	10.83
15Y	10.28	9.76	9.66	9.66
20Y	9.97	9.24	9.05	9.03
25Y	9.69	8.72	8.47	8.42
30Y	9.30	8.03	7.69	7.62

Table 4.1: Calibration performance for implied volatilities (%) of the ATM caplets

Expiry	Market	Approx SABR-SMM (MC)	SABR-SMM (MC)	MFM
$\nu = 0.1$, $\beta = 0.9$, $\rho = -0.8$, $\sigma_0 = 0.02$ and $\mu_j = 0$				
5Y	10.64	10.64	10.64	10.63
10Y	10.06	10.01	9.99	9.99
15Y	9.92	9.79	9.74	9.73
20Y	9.82	9.59	9.53	9.52
25Y	9.61	9.24	9.14	9.13
30Y	9.30	8.79	8.65	8.65
$\nu = 0.2$, $\beta = 0.7$, $\rho = -0.4$, $\sigma_0 = 0.04$ and $\mu_j = 0.01$				
5Y	10.64	10.63	10.62	10.61
10Y	10.06	9.98	9.95	9.94
15Y	9.92	9.71	9.67	9.63
20Y	9.82	9.44	9.34	9.32
25Y	9.61	9.06	8.93	8.92
30Y	9.30	8.57	8.40	8.40
$\nu = 0.3$, $\beta = 0.5$, $\rho = 0$, $\sigma_0 = 0.06$ and $\mu_j = 0.03$				
5Y	10.64	10.63	10.62	10.60
10Y	10.06	9.91	9.88	9.85
15Y	9.92	9.63	9.57	9.50
20Y	9.82	9.31	9.19	9.15
25Y	9.61	8.90	8.75	8.73
30Y	9.30	8.40	8.19	8.18
$\nu = 0.4$, $\beta = 0.3$, $\rho = 0.4$, $\sigma_0 = 0.08$ and $\mu_j = 0.01 \times j$				
5Y	10.64	10.62	10.60	10.58
10Y	10.06	9.85	9.81	9.74
15Y	9.92	9.53	9.43	9.35
20Y	9.82	9.19	9.02	8.98
25Y	9.61	8.74	8.55	8.51
30Y	9.30	8.19	7.93	7.91
$\nu = 0.5$, $\beta = 0.1$, $\rho = 0.8$, $\sigma_0 = 0.1$ and $\mu_j = 0.3 - 0.01 \times j$				
5Y	10.64	10.61	10.59	10.55
10Y	10.06	9.81	9.75	9.65
15Y	9.92	9.42	9.34	9.21
20Y	9.82	9.09	8.90	8.78
25Y	9.61	8.65	8.39	8.29
30Y	9.30	8.03	7.69	7.62

Table 4.2: Calibration performance for implied volatilities (%) of the ATM co-terminal swaptions

Scenario	ν	σ_0	β	ρ
I	0.1	0.02	0.9	-0.8
II	0.2	0.04	0.7	-0.4
III	0.3	0.06	0.5	0
IV	0.4	0.08	0.3	0.4
V	0.5	0.1	0.1	0.8

Table 4.3: Scenarios for the SV MFM.

i	Market	MFM (I)	MFM (II)	MFM (III)	MFM (IV)	MFM (V)
1	0.689	0.693	0.696	0.700	0.701	0.703
2	0.796	0.797	0.799	0.801	0.802	0.805
3	0.843	0.843	0.844	0.846	0.847	0.849
4	0.871	0.871	0.871	0.872	0.873	0.874
5	0.888	0.888	0.888	0.888	0.888	0.889
6	0.901	0.901	0.901	0.900	0.900	0.896
7	0.909	0.909	0.909	0.907	0.906	0.902
8	0.916	0.916	0.915	0.913	0.911	0.905
9	0.922	0.922	0.920	0.918	0.915	0.905
10	0.926	0.926	0.923	0.921	0.916	0.905
11	0.929	0.929	0.926	0.925	0.916	0.906
12	0.932	0.932	0.928	0.927	0.917	0.906
13	0.935	0.935	0.932	0.929	0.918	0.907
14	0.937	0.937	0.933	0.929	0.917	0.907
15	0.938	0.937	0.933	0.929	0.915	0.906
16	0.940	0.939	0.935	0.930	0.914	0.906
17	0.941	0.939	0.936	0.930	0.913	0.904
18	0.942	0.940	0.937	0.930	0.912	0.902
19	0.943	0.941	0.937	0.930	0.912	0.900
20	0.944	0.941	0.937	0.930	0.911	0.899
21	0.945	0.942	0.938	0.930	0.912	0.898
22	0.946	0.942	0.939	0.931	0.912	0.896
23	0.946	0.942	0.939	0.930	0.911	0.894
24	0.947	0.942	0.939	0.930	0.910	0.891
25	0.947	0.942	0.938	0.929	0.908	0.888
26	0.948	0.943	0.938	0.929	0.907	0.884
27	0.948	0.943	0.938	0.928	0.906	0.879
28	0.954	0.946	0.939	0.928	0.906	0.879
29	0.961	0.953	0.942	0.930	0.906	0.868

Table 4.4: Calibration performance for one-step correlation of LIBORs $Corr(ln(L^i_{T_i} + \theta_i), ln(L^{i+1}_{T_{i+1}} + \theta_{i+1}))$ under SV MFMs with different scenarios

Chapter 5

Comparison of stochastic volatility Markov-functional model and one-factor Markov-functional models

5.1 Introduction

In the interest rate markets, Bermudan swaptions are one of the most popular exotic derivatives which are traded over-the-counter. A Bermudan swaption is a derivative that allows its holder to enter into an interest rate swap on any of a set of pre-arranged exercise dates. The term "Bermudan" means that the option allows for a discrete set of exercise dates. But the option can only be exercised once. We will consider pricing and hedging Bermudan swaptions in this chapter. The other exotic option we will study in this chapter is a new Bermudan product. The new Bermudan product has more complicated payoffs, and its underlying is a structured interest rate swap, where a regular fixing leg is swapped against structured coupons. The new Bermudan product studied here is motivated by a callable range accrual which is a popular exotic option in the interest rate markets. Callable range accruals are very involved products to price whose payoff depends on the accrual of days that a LIBOR falls within a pre-arranged range. For details, the reader is referred to [32] and [6] and references therein. The new Bermudan product studied here is also traded in the markets, and has similar features but is much simpler than callable range accruals. The simplification means that we can focus on understanding the effect of stochastic volatility without introducing further approximations needed to

handle a more involved product.

There are many interest rate models that are capable of pricing and hedging the above two Bermudan products which include short rate models, LIBOR market models and Markov-functional models. We focus on Markov-functional models in this chapter. In the literature there has been substantial debate between single-factor models and multi-factor models. See [50], [3] and [22]. Pietersz and Pelsser [57] compared a one-factor Markov-functional model and multi-factor market models in terms of hedging performance for a Bermudan swaption. The numerical results showed that the performance of the two models is comparable which is consistent with the recent consensus post crisis. Thus in this chapter we focus on the low-factor Markov-functional models.

Pietersz and Pelsser [57] also studied the effect of implied volatility smiles/skews on the price of a Bermudan swaption in a one-dimensional Markov-functional model with a Gaussian driver. The authors compared the following two models: a Markov-functional model feeding in the Black's formula and a Markov-functional model feeding in the displaced diffusion formula which is able to capture implied volatility skews. The displaced diffusion Markov-functional model is calibrated to the market implied volatility smiles using the least squares method which can be seen as a "global fit" approach. In this case the numerical results indicated that the smile effect is very large in terms of pricing a Bermudan swaption. In this chapter we perform a similar numerical study. But the difference from [57] is that we adopt a "local fit" calibration approach where we fit the implied volatility at the strike of the Bermudan exactly. In this case our results show that the effect of smile is insignificant on the price of a Bermudan swaption.

Another numerical study for Markov-functional models was performed by Kennedy and Pham [48] where they investigated the effect of the specification of a Gaussian driver on a Bermudan swaption in terms of hedging performance in a one-dimensional Markov-functional model feeding in the Black's formula. Their numerical results indicated that the driver "parameterized by time" outperforms the driver "parameterized by expiry". In this chapter we make a further development. In addition to the specifications of a Gaussian driver proposed in [48], we also consider a SABR driver and take implied volatility smiles into account. Our results show that the introduction of stochastic volatility does not materially alter the hedging behaviour of parametrization by time. This finding is significant for practitioners wanting to use stochastic volatility models.

An important observation made by Pietersz and Pelsser [57] is that the price of a co-terminal Bermudan swaption is mainly determined by the joint distribution

of the underlying co-terminal swap rates at their setting dates. This motivates us to calibrate to the co-terminal vanilla swaptions as well as the market implied correlation structure of the co-terminal swap rates because they can fully determine the joint distribution of the co-terminal swap rates at their setting dates (in the case of Normal or log-Normal distribution). However due to the complex nature of Bermudan products, stochastic volatility may also influence prices and hedges of Bermudan type products. In fact empirical evidence (see [15] and [49]) supports the use of stochastic volatility for interest rates. Furthermore as Hagan et al. [31] pointed out, incorporating an extra stochastic volatility factor into a model could give a more realistic evolution for the implied volatility smile. Kaisajuntti and Kennedy [44] used market data to identify a SABR type model as an appropriate choice for the level of interest rates. These observations motivate us to investigate the stochastic volatility impact on Bermudan type products in terms of pricing and hedging. Although many stochastic volatility term structure models have been proposed in the literature, see e.g. [4], [65], [55], [62], [60], [34], [33] and [5], little is known in the literature about investigating the effect of introducing stochastic volatility on pricing and hedging Bermudan style products.

In this chapter we compare a stochastic volatility Markov-functional model with a one-dimensional Markov-functional model in terms of pricing and vega hedging performance for a Bermudan swaption and the new Bermudan product. In our numerical comparison, we calibrate a stochastic volatility *swap* Markov-functional model and one-dimensional swap Markov-functional models with different combinations of driver and marginals to the real market. Then we compare the prices and vega profiles of Bermudan products produced by the above Markov-functional models. The numerical results show a very big difference in vega profiles of Bermudan products produced by models parameterized by time and by expiry, and the result is consistent with the results of [48]. Furthermore we find that this big difference is not changed when implied volatility smiles and stochastic volatility are taken into account. The introduction of stochastic volatility has an insignificant effect on the price of a Bermudan swaption but the impact is significant on the new Bermudan product.

The rest of the chapter is organised as follows. In Section 5.2, we introduce the methodology for comparison of the Markov-functional models. In Section 5.3 we describe the market data for the comparison. In Sections 5.4 and 5.5 we compare Markov-functional models in terms of pricing and hedging a Bermudan swaption and a new Bermudan product. Section 5.6 concludes.

5.2 Methodology

In this section we describe the framework for a comparison between one-dimensional swap Markov-functional models (MFMs) and a stochastic volatility swap MFM in terms of pricing and hedging a Bermudan swaption and a new Bermudan product. In Sections 5.2.1 and 5.2.2 we introduce Bermudan swaptions and new Bermudan products, and discuss their features that need to be captured for an accurate pricing. In the numerical study we consider the following five swap MFMs for comparison:

(a) one-dimensional swap MFM with a mean reversion (MR) driver fitting the Black's formula for swaptions

(b) one-dimensional swap MFM with a Hull-White (HW) driver fitting the Black's formula for swaptions

(c) one-dimensional swap MFM with a one-step covariance driver fitting the Black's formula for swaptions

(d) one-dimensional swap MFM with a Hull-White driver feeding in SABR marginal distributions of swap rates at their setting dates

(e) stochastic volatility swap MFM with a SABR driver feeding in SABR marginal distributions of swap rates at their setting dates.

A discussion of these MFMs and an explanation of why we choose these MFMs for the comparison will be given in Sections 5.2.3 and 5.2.4. To perform a comparison we calibrate the above five MFMs to the market data. Then we compute model prices and produce vega profiles of a Bermudan swaption and a new Bermudan product within these five calibrated MFMs. By comparing the model prices and vegas, we investigate the correlation, smile and stochastic volatility impacts on the pricing and hedging performance of Bermudan products. The analytical expressions for the vegas of a Bermudan swaption in the above five MFMs are studied in Section 5.2.5. The numerical results are given in Sections 5.4 and 5.5.

5.2.1 Bermudan swaptions

In the numerical study we consider a co-terminal style Bermudan swaption where its holder has the right to enter into the remaining underlying co-terminal interest rate swap at a number of pre-arranged exercise dates. At each exercise date the holder can decide on exercising the right or waiting for the next exercise date, but only one exercising opportunity is given.

Let us first discuss how to price a co-terminal Bermudan swaption. For the tenor structure (2.1), let us consider a pay fixed co-terminal Bermudan swaption with strike K and exercise dates $T_1, T_2, ..., T_n$ on an interest rate swap with payment dates $T_2, T_3, ..., T_{n+1}$. The value $V_{Berm}(0)$ of this Bermudan swaption at time 0 can be found recursively. At exercise date T_i, $i = 1, ..., n$, the holder can decide on exercising or waiting so that the payoff $\widetilde{V}^i_{Berm}(T_i)$ is given by

$$\widetilde{V}^i_{Berm}(T_i) := \max(V^i_{Berm}(T_i), V^{i,n+1-i}_{Swap}(T_i)), \tag{5.1}$$

where $V^{i,n+1-i}_{Swap}(T_i)$ denotes the time T_i value of a pay fixed interest rate swap with setting dates $T_i, T_{i+1}, ..., T_n$ and settlement dates $T_{i+1}, T_{i+2}, ..., T_{n+1}$, which is given by

$$V^{i,n+1-i}_{Swap}(T_i) = N[P^{i,n+1-i}_{T_i}(y^{i,n+1-i}_{T_i} - K)],$$

where N is the notional amount, $P^{i,n+1-i}$ is PVBP and $y^{i,n+1-i}$ is swap rate; see Section 2.1.1. We note that at the last exercise date T_n the above Bermudan swaption is simply a vanilla swaption so that the payoff $\widetilde{V}^n_{Berm}(T_n)$ at time T_n is given by

$$\widetilde{V}^n_{Berm}(T_n) := \max(0, V^{n,1}_{Swap}(T_n)).$$

We can see that the payoff function (5.1) is a maximum between the value $V^{i,n+1-i}_{Swap}(T_i)$ of exercising the option and the value $V^i_{Berm}(T_i)$ of waiting at time T_i. In an arbitrage-free model under some equivalent martingale measure $\mathbb{Q}$ corresponding to the numeraire M, the value $V^i_{Berm}(T_i)$ of the Bermudan swaption at time T_i, $i = 0, ..., n-1$, is given by

$$V^i_{Berm}(T_i) = M_{T_i} E_{\mathbb{Q}}[\frac{\widetilde{V}^{i+1}_{Berm}(T_{i+1})}{M_{T_{i+1}}} | \mathcal{F}_{T_i}],$$

where $\widetilde{V}^{i+1}_{Berm}(T_{i+1})$ is the payoff at time T_{i+1}. In our context we price a Bermudan swaption using a MFM under the terminal measure $\mathbb{F}^{n+1}$ with the corresponding numeraire $D_{\cdot,T_{n+1}}$.

We now consider the problem of what instruments and features one needs to calibrate to in order to give an accurate Bermudan price in a low-factor model. As pointed out by Pietersz and Pelsser [57], the joint distribution of the co-terminal swap rates $\{y^{j,n+1-j}_{T_i}; j = i, ..., n, i = 1, ..., n\}$ can determine the price of a Bermudan swaption, but the main contribution is the joint distribution of the swap rates

$\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$. As a result a valuation model for the Bermudan swaption needs to be calibrated to the co-terminal vanilla swaptions so that the market implied marginal distributions of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$ in the associated swaption measure $\mathbb{S}^{i,n+1-i}$, $i = 1, ..., n$, corresponding to the numeraire $P^{i,n+1-i}$ can be captured. Note that in financial markets the prices of vanilla swaptions are given to us in the form of implied volatilities which usually are a function of strike displaying a shape of smile or skew. Ideally we would like to calibrate to the whole implied volatility smile or skew of the vanilla swaptions which is equivalent to capturing the marginal distributions. But this is not always possible. For example Black's formula for the vanilla swaption assumes that the implied volatility is a constant w.r.t strike so that it can only capture implied volatility at one strike. In this case we need to decide on which implied volatility the model should be calibrated to. A rule of thumb is to choose the implied volatility of the co-terminal vanilla swaptions struck at the strike K of the Bermudan swaption since this is the most relevant to the Bermudan swaption.

Another key feature one needs to capture is the correlation (covariance) structure of the underlying swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$. The correlation structure is not observable directly from the markets. One has to extract them from the market implied volatilities of vanilla swaptions. For details, the reader is referred to [48].

Therefore in order to price the Bermudan swaption properly, a low-factor valuation model needs to be calibrated to the co-terminal vanilla swaptions and the correlation structure of the underlying swap rates. We will discuss the problem of how to achieve this under swap MFMs later in this section.

5.2.2 New Bermudan product

The definition of a new Bermudan product is similar to a co-terminal Bermudan swaption except for the underlying interest rate swap. In particular, we recall that the underlying interest rate swap of the co-terminal Bermudan swaption, considered in Section 5.2.1, with strike K admits a fixed leg $N\alpha_i K$ and floating leg $N\alpha_i L_{T_i}^i$ at each settlement date T_{i+1} for $i = 1, ..., n$. For the new Bermudan product however the underlying is a structured interest rate swap with a fixed leg $N\alpha_i K$ and a floating leg $N\alpha_i L_{T_i}^i I_{\{l \leq L_{T_i}^i < m\}}$ at each settlement date T_{i+1}, $i = 1, ..., n$, for lower barrier l and upper barrier m.

Let us now consider a pay fixed new Bermudan product with strike K, notional amount N and exercise dates $T_1, T_2, ..., T_n$ on a structured interest rate swap with payment dates $T_2, T_3, ..., T_{n+1}$. The procedure for finding the value $V_{Prod}(0)$ of such a Bermudan product at time 0 is similar to the Bermudan swaption that is

discussed earlier except that the time T_i value $V_{Swap}^{i,n+1-i}(T_i)$ of the interest rate swap in the payoff function (5.1) is replaced by the value $V_{Strut}^{i,n+1-i}(T_i)$ of the structured interest rate swap at each exercise date T_i, $i = 1, ..., n$, which is given by

$$V_{Strut}^{i,n+1-i}(T_i) = [V_{Flt}^i(T_i) - V_{Fix}^i(T_i)],$$

where $V_{Fix}^i(T_i)$ and $V_{Flt}^i(T_i)$ denote the time T_i values of the fixed leg and the floating leg respectively. The fixed leg consists of a series of payments on dates T_{i+1},T_{i+2},...,T_{n+1} so that its value at time T_i is given by

$$\begin{aligned} V_{Fix}^i(T_i) &= NK \sum_{j=i}^{n} \alpha_j D_{T_i T_{j+1}} \\ &= NK P_{T_i}^{i,n+1-i}. \end{aligned}$$

The floating leg is more difficult to value. To do so we apply the fundamental pricing formula. In particular under some equivalent martingale measure $\mathbb{Q}$ corresponding to the numeraire M, the value $V_{Flt}^i(T_i)$ of the floating leg at time T_i is given by

$$V_{Flt}^i(T_i) = M_{T_i} N E_{\mathbb{Q}}[\sum_{j=i}^{n} \frac{\alpha_j L_{T_j}^j I_{\{l \le L_{T_j}^j < m\}}}{M_{T_{j+1}}} | \mathcal{F}_{T_i}]. \tag{5.2}$$

In the context of MFMs we choose the terminal measure $\mathbb{F}^{n+1}$ as the equivalent martingale measure corresponding to the numeraire $D_{.,T_{n+1}}$.

Having discussed the valuation of the Bermudan product we now consider the calibration problem in a low-factor valuation model. The complication of the Bermudan product makes it difficult to identify the calibrating instruments. However we notice that as the lower barrier l decreases and upper barrier m gets bigger, this Bermudan product should reduce to a Bermudan swaption. It is noted in Hagan [32] that one can hedge the new product by using a Bermudan swaption, and this requires using the same valuation model and calibration method for the new Bermudan product as would be used for Bermudan swaptions. As a result it is necessary to calibrate to the co-terminal vanilla swaptions and the correlation structure of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$. Furthermore we note that a floating coupon $N\alpha_i L_{T_i}^i I_{\{l \le L_{T_i}^i < m\}}$ at settlement dates T_{i+1}, $i = 1, ..., n$, of the underlying structured

interest rate swap can be expressed as a combination of caplets and digital caplets:

$$N\alpha_i L^i_{T_i} I_{\{l\le L^i_{T_i}<m\}} = N\alpha_i([L^i_{T_i}-l]^+ + lI_{\{L^i_{T_i}\ge l\}} - [L^i_{T_i}-m]^+ - mI_{\{L^i_{T_i}\ge m\}}), \tag{5.3}$$

so that equation (5.2) can be rewritten as

$$V^i_{Flt}(T_i) = M_{T_i}N\sum_{j=i}^{n} E_{\mathbb{Q}}[\frac{\alpha_j([L^i_{T_i}-l]^+ + lI_{\{L^i_{T_i}\ge l\}} - [L^i_{T_i}-m]^+ - mI_{\{L^i_{T_i}\ge m\}})}{M_{T_{j+1}}}|\mathcal{F}_{T_i}]. \tag{5.4}$$

As a result in order to price the value of floating leg (5.2) properly, calibrating a valuation model to the market prices of the caplets and digital caplets indicated on the right hand side of (5.4) is required.

We note that theoretically knowing the prices of caplet $V^i_{caplet}(0;K)$ for all strikes K is equivalent to knowing the price of the digital caplet $V^i_{digcap}(0;K)$ for all K. In practice however the prices of digital caplets are not always quoted in the market. But they can be replicated by a bull spread of caplets; See [52]. In particular consider the bull spread consisting of $\frac{1}{\varepsilon\alpha_i}$ caplets with strike $K-\frac{1}{2}\varepsilon$ and $-\frac{1}{\varepsilon\alpha_i}$ caplets struck at $K+\frac{1}{2}\varepsilon$, which yields the payoff

$$\begin{aligned}&\frac{1}{\varepsilon\alpha_i}[\alpha_i(L^i_{T_i}-K+\frac{1}{2}\varepsilon)^+ - \alpha_i(L^i_{T_i}-K-\frac{1}{2}\varepsilon)^+]\\ =&\begin{cases}0, & \text{if } L^i_{T_i}\le K-\frac{1}{2}\varepsilon\\ \frac{1}{\varepsilon}(L^i_{T_i}-K+\frac{1}{2}\varepsilon), & \text{if } K-\frac{1}{2}\varepsilon < L^i_{T_i} < K+\frac{1}{2}\varepsilon\\ 1, & \text{if } L^i_{T_i}\ge K+\frac{1}{2}\varepsilon.\end{cases}\end{aligned}$$

The payoff of this bullish spread of caplets reduces to the payoff of a digital caplet on the ith LIBOR L^i struck at K as $\varepsilon\to 0$. Thus a digital caplet can be replicated by the corresponding bullish spread of caplets so that

$$V^i_{digcap,ma}(0;K) = \frac{1}{\varepsilon\alpha_i}[V^i_{caplet,ma}(0;K-\frac{1}{2}\varepsilon) - V^i_{caplet,ma}(0;K+\frac{1}{2}\varepsilon)]$$

as $\varepsilon\to 0$, where the subscript "ma" stands for a market value. In practice ε is usually set to be 5bps or 10bps. In our numerical study we use this method to obtain digital caplets prices from caplets prices.

In summary, one needs to calibrate a valuation model to the market prices of a set of (digital) caplets, co-terminal vanilla swaptions and the correlation structure of swap rates. However we still have the following two calibration problems.

Effective Strikes

Firstly, we remember that in the case of non-smile valuation models for a Bermudan swaption, the implied volatility of the vanilla swaptions struck at the strike K of the Bermudan swaption, as opposed to other strikes, is the best choice to calibrate to. But the new Bermudan product has a more sophisticated underlying structured interest rate swap. In this case it is not obvious which implied volatility to calibrate to. Following standard practice we will calibrate to the implied volatility of the swaptions struck at the "effective strike". For details the reader is referred to [32]. The effective strike K_i^{eff} for each co-terminal vanilla swaption with expiry date T_i, $i = 1, ..., n$, are chosen such that

$$\frac{\text{fixed leg time-0 value of swap}(K_i^{eff})}{\text{floating leg time-0 value of swap}} = \frac{\text{fixed leg time-0 value of structured swap}(K)}{\text{floating leg time-0 value of structured swap}},$$

and therefore

$$K_i^{eff} = \frac{NK(D_{0T_i} - D_{0T_{n+1}})}{V_{Flt}^i(0)}. \tag{5.5}$$

It follows from equation (5.3) that $V_{Flt}^i(0)$ is given by

$$\begin{aligned} V_{Flt}^i(0) &= D_{0T_{n+1}} N \sum_{j=i}^{n} E_{\mathbb{F}^{n+1}}\left(\frac{\alpha_j L_{T_j}^j I_{\{l \le L_{T_j}^j < m\}}}{D_{T_{j+1}T_{n+1}}}\right) \\ &= \sum_{j=i}^{n} [V_{caplet}^j(0; l) - V_{caplet}^j(0; m) + \alpha_j l V_{digcap}^j(0; l) - \alpha_j m V_{digcap}^j(0; m)], \end{aligned}$$

where $V_{caplet}^j(0; K)$ and $V_{digcap}^j(0; K)$ are prices of a caplet and digital caplet on the jth LIBOR L^j struck at the strike K. In our numerical comparison we assume that once we have determined the effective strike via the formula (5.5), it will not change in the calculation of the vegas.

Internal Adjusters

The second calibration issue is that, in general, the market prices of co-terminal swaptions and (digital) caplets can not be reproduced by a valuation model simultaneously. To solve this problem we will introduce the idea of "internal adjusters". In particular suppose we are given a valuation model under some equivalent martingale measure $\mathbb{Q}$ corresponding to the numeraire M, and the model has been calibrated to the prices of vanilla swaptions. In general this model will not repro-

duce the market prices of the caplet and digital caplet with some strike K on the ith LIBOR L^i:

$$V^i_{digcap,mo}(0;K) = M_0 E_{\mathbb{Q}}[N I_{\{L^i_{T_i} \geq K\}} M^{-1}_{T_{i+1}}] \neq V^i_{digcap,ma}(0;K)$$
$$V^i_{caplet,mo}(0;K) = M_0 E_{\mathbb{Q}}[N\alpha_i [L^i_{T_i} - K]^+ M^{-1}_{T_{i+1}}] \neq V^i_{caplet,ma}(0;K),$$

where "mo" stands for a model value. Consequently the valuation model is incapable of pricing the floating leg (5.4) properly. To fix this problem we adjust the payoff functions for caplets and change the strike for digital caplets in the model. In particular we choose strikes $\tilde{K}_i$ and constants κ_i, $i = 1, ..., n$, such that

$$V^i_{digcap,mo}(0;\tilde{K}) = M_0 E_{\mathbb{Q}}[N I_{\{L^i_{T_i} \geq \tilde{K}_i\}} M^{-1}_{T_{i+1}}] = V^i_{digcap,ma}(0;K) \tag{5.6}$$
$$\tilde{V}^i_{caplet,mo}(0;K) = M_0 E_{\mathbb{Q}}[N\alpha_i [\kappa_i L^i_{T_i} - K]^+ M^{-1}_{T_{i+1}}] = V^i_{caplet,ma}(0;K). \tag{5.7}$$

The above strikes $\tilde{K}_i$ and κ_i, $i = 1, ..., n$, are referred to as the internal strikes and internal coefficients. When we price the floating leg (5.4) and therefore the new Bermudan product, we use internal strikes and internal coefficients rather than using the true payoff functions. Essentially we get the prices of what we're interested in right by changing their payoff functions but not changing the model.

5.2.3 One-dimensional swap Markov-functional models: driver specification and calibration

We now discuss the one-dimensional swap MFMs that will be considered in the numerical comparison. We have discussed the specification of a one-dimensional swap MFM by assuming that a Gaussian driving process is given in Section 3.2.2. In this subsection we explore the specification of the Gaussian driving process. Kennedy and Pham [48] investigated the effect of the specification of a Gaussian driver on a Bermudan swaption in terms of hedging performance. The authors studied the mean reversion (MR) driver, which is "parameterized by expiry", Hull-White (HW) and one-step covariance drivers which are "parameterized by time". Their numerical results indicated that the driver parameterized by time outperforms the driver parameterized by expiry. In this subsection we review these specifications of the Gaussian driver and explain the ideas of parametrizations by time and by expiry. We also consider the calibration issue in one-dimensional swap MFMs. The material in this subsection is from [48].

In a one-dimensional swap MFM, we consider the following driving process

$$x_t := \int_0^t \sigma_u dW_u^{n+1},$$

where W^{n+1} is a one-dimensional Brownian motion under the terminal measure $\mathbb{F}^{n+1}$ corresponding to the numeraire $D_{.,T_{n+1}}$ and σ_t is a deterministic function. Recall that we implement a MFM on a grid so that it is only necessary to specify the variance

$$\xi_{T_i} := Var(x_{T_i}) = \int_0^{T_i} \sigma_u^2 du$$

of x at time T_i for $i = 1, ..., n$. This is the case since x is a Gaussian process and the mean of it at time T_i is 0 for $i = 1, ..., n$. We will discuss different specifications of the variance ξ_{T_i}'s later. Given a driving process, in order to specify the model we need to feed in the input prices of PVBP-digital swaptions as a function of strike which implies the marginal distributions of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$ in their associated swaption measures. In our numerical study we feed in log-Normal marginal distributions and SABR marginals of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$ in their associated swaption measures.

Let us first consider the log-Normal marginals, and we will consider feeding in SABR marginals later. Suppose that the input prices of the PVBP-digital swaptions are obtained by the Black's formula (2.12) with some implied volatilities $\tilde{\sigma}_{i,n+1-i}$ for $i = 1, ..., n$. Note that the choice of $\tilde{\sigma}_{i,n+1-i}$'s depends on the product we wish to price. In our numerical comparison, it can be chosen to be implied volatilities $\tilde{\sigma}_{i,n+1-i}^{Berm}$ of the co-terminal swaptions struck at the strike of the Bermudan swaption or $\tilde{\sigma}_{i,n+1-i}^{eff}$ struck at the effective strike depending on whether we are pricing a Bermudan swaption or the new Bermudan product.

As the correlation structure of swap rates is an important feature for pricing a Bermudan product, before we proceed to the specification of the Gaussian driver we discuss the problem of how to capture the correlation structure in a swap MFM. Bennett and Kennedy [7] indicated that under the swap MFM with a Gaussian driving process x calibrated to the Black's formula with implied volatility $\tilde{\sigma}_{i,n+1-i}$ for the prices of the co-terminal swaptions, there is an approximate linear relationship:

$$\ln y_t^{i,n+1-i} \approx \eta_t^i + \gamma^i x_t, \tag{5.8}$$

for deterministic function η_t^i and constant γ^i. Thus the correlation of co-terminal

swap rates implied by the above swap MFM can be captured by the variance of x:

$$Corr(x_{T_i}, x_{T_j}) = \sqrt{\frac{\xi_{T_i}}{\xi_{T_j}}} \approx Corr^{mo}(\ln y_{T_i}^{i,n+1-i}, \ln y_{T_j}^{j,n+1-j}). \tag{5.9}$$

for $i < j$, where "mo" stands for a model value. Furthermore we have an approximation in $\mathbb{F}^{n+1}$

$$Var^{mo}(\ln y_{T_i}^{i,n+1-i}) \approx \tilde{\sigma}_{i,n+1-i}^2 T_i. \tag{5.10}$$

It follows from (5.8) and (5.10) that

$$(\gamma^i)^2 \xi_{T_i} \approx \tilde{\sigma}_{i,n+1-i}^2 T_i. \tag{5.11}$$

The approximation (5.11) will help us understand the ideas of parametrizations by time and by expiry which will be explained later. In what follows we review the three specifications of the Gaussian driving process.

Mean reversion driving process

The mean reversion driving process was first introduced by Hunt, Kennedy and Pelsser [37]. The mean reversion driving process is specified by choosing the deterministic function $\sigma_t = e^{at}$, where the constant $a > 0$ is the mean reversion parameter. The variance ξ_{T_i} of the MR driving process x at each time T_i, $i = 1, ..., n$ is given by

$$\xi_{T_i} = \int_0^{T_i} e^{2at} dt = \frac{1}{2a}(e^{2aT_i} - 1).$$

In this case the terminal correlation of swap rates in the model is determined by the mean reversion parameter a:

$$Corr^{mo}(\ln y_{T_i}^{i,n+1-i}, \ln y_{T_j}^{j,n+1-j}) \approx \sqrt{\frac{e^{2aT_i} - 1}{e^{2aT_j} - 1}} \tag{5.12}$$

for $i < j$.

We now explain the idea of parametrization by expiry. We can see that once the MR parameter a is fixed, as market implied volatilities change, the variance of x remains unchanged and therefore, from (5.12), the model implied terminal correlation of swap rates will not change. Moreover we can see from (5.11) that a change in the market implied volatilities leads to a change in expiry-dependent

parameters γ^i's. In this case, we say the mean reversion process is "parameterized by expiry" which was first introduced by Kennedy and Pham [48]. However it is observed in the market that a change in the market implied volatilities of the swaptions commonly leads to a change in the terminal correlation of swap rates. In that sense, the mean reversion process cannot reflect reality. To partially capture this market observation, we introduce a Hull-White driving process.

Hull-White driving process

The variance of the Hull-White driving process x at each time T_i, $i = 1, ..., n$, is given by

$$\xi_{T_i} = (\frac{(T_{n+1} - T_i)\tilde{\sigma}^{ATM}_{i,n+1-i}}{(1 + \alpha_i y_0^{i,n+1-i})(\psi_{T_{n+1}} - \psi_{T_i})})^2 T_i, \tag{5.13}$$

where $\psi_{T_i} = \frac{1}{a}(1 - e^{-aT_i})$, $a > 0$, and where $\tilde{\sigma}^{ATM}_{i,n+1-i}$ is the market ATM implied volatility of the co-terminal swaption. Thus the Hull-White driving process is linked to the co-terminal swaptions of the swaption matrix; see Table 5.1. For details about the derivation of the above variance, the reader is referred to [7].

Tenor		1	2	3	...	$n-2$	$n-1$	n
Expiry	1	...	...	...	...	...	...	$\tilde{\sigma}^{ATM}_{1,n}$
	2	...	...	...	...	...	$\tilde{\sigma}^{ATM}_{2,n-1}$	...
	3	...	...	...	...	$\tilde{\sigma}^{ATM}_{3,n-2}$	...	...
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
	$n-2$	...	...	$\tilde{\sigma}^{ATM}_{n-2,3}$	...	...	...	...
	$n-1$	...	$\tilde{\sigma}^{ATM}_{n-1,2}$	...	...	...	...	...
	n	$\tilde{\sigma}^{ATM}_{n,1}$	...	...	...	...	...	...

Table 5.1: ATM implied volatilities in the swaption matrix linked to the Hull-White driver.

We now explain the idea of parametrization by time. We can see from equation (5.13) that the variance ξ_{T_i} of x at time T_i depends on the market implied volatility $\tilde{\sigma}^{ATM}_{i,n+1-i}$ of the co-terminal swaption. As a result as the market implied volatility $\tilde{\sigma}^{ATM}_{i,n+1-i}$ change, the variance of x and therefore the terminal correlation of swap rates will be changed. Moreover we insert (5.13) into (5.11), and we can see that as the market implied volatilities change, the expiry-dependent parameters γ^i's will almost stay the same (exactly the same when $\tilde{\sigma}_{i,n+1-i} = \tilde{\sigma}^{ATM}_{i,n+1-i}$). Such a driving process is said to be "parameterized by time" which was first proposed by Kennedy and Pham [48].

So far we have explained parametrizations by expiry and by time. Another way to gain an insight into this idea is by considering the responses of the LIBORs to a change in the implied volatilities. For details, the reader is referred to [48]. Note that the difference in parametrizations by expiry and by time has a strong impact on the hedging behaviour of the model. For non-stochastic volatility models this behaviour is well-known to practitioners and is analysed in [48] for one dimensional Markov-functional models with the Black's formula.

We note that in the MR and HW models, we only have one free parameter a for the driving process x, which is insufficient to capture the whole correlation structure. This problem is partially solved by the introduction of the one-step covariance driving process.

One step covariance driving process

The one step covariance driving process was introduced by Kennedy and Pham [48]. For this driving process the variance ξ_{T_i} of x at time T_i, $i = 1, ..., n$, is chosen by calibrating to the market correlation structure of swap rates. Note that it is impossible to calibrate a one-factor model to the whole correlation matrix. Thus we only consider the one step covariances (correlations) between $\ln y_{T_i}^{i,n+1-i}$ and $\ln y_{T_{i+1}}^{i+1,n-i}$ for $i = 1, ..., n-1$. In particular suppose we are given the one step covariances $Cov^{ma}(\ln y_{T_i}^{i,n+1-i}, \ln y_{T_{i+1}}^{i+1,n-i})$, $i = 1, ..., n-1$, from the market. Note that the market one step covariance can be estimated by using the swaption matrix $\{\tilde{\sigma}_{i,j}^{ATM}; i = 1, ..., n, j = 1, ..., n+1-i\}$, where $\tilde{\sigma}_{i,j}^{ATM}$ is the ATM implied volatility of the vanilla swaption on the swap rate $y^{i,j}$; see Table 5.2. The estimation involves

Tenor		1	2	3	...	$n-2$	$n-1$	n
Expiry	1	$\tilde{\sigma}_{1,1}^{ATM}$	$\tilde{\sigma}_{1,2}^{ATM}$	$\tilde{\sigma}_{1,3}^{ATM}$	...	$\tilde{\sigma}_{1,n-2}^{ATM}$	$\tilde{\sigma}_{1,n-1}^{ATM}$	$\tilde{\sigma}_{1,n}^{ATM}$
	2	$\tilde{\sigma}_{2,1}^{ATM}$	$\tilde{\sigma}_{2,2}^{ATM}$	$\tilde{\sigma}_{2,3}^{ATM}$	...	$\tilde{\sigma}_{2,n-2}^{ATM}$	$\tilde{\sigma}_{2,n-1}^{ATM}$	...
	3	$\tilde{\sigma}_{3,1}^{ATM}$	$\tilde{\sigma}_{3,2}^{ATM}$	$\tilde{\sigma}_{3,3}^{ATM}$	...	$\tilde{\sigma}_{3,n-2}^{ATM}$	...	...
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	$\vdots$	$\vdots$
	$n-2$	$\tilde{\sigma}_{n-2,1}^{ATM}$	$\tilde{\sigma}_{n-2,2}^{ATM}$	$\tilde{\sigma}_{n-2,3}^{ATM}$	...	...	...	...
	$n-1$	$\tilde{\sigma}_{n-1,1}^{ATM}$	$\tilde{\sigma}_{n-1,2}^{ATM}$	...	...	...	...	...
	n	$\tilde{\sigma}_{n,1}^{ATM}$	...	...	...	...	...	...

Table 5.2: ATM implied volatilities in the swaption matrix linked to the one step covariance driver and the SABR driver.

a two-step procedure. For details the reader is referred to [48]. Now we specify the one step covariance driving process as follows.

- The algorithm for fixing the variance ξ_{T_i} of x at time T_i, $i = 1, ..., n$, works

back iteratively from T_n. At time T_n, without loss of generality, we set

$$\xi_{T_n} = (\widetilde{\sigma}^{ATM}_{n,1})^2 T_n.$$

- Suppose we have reached time T_i, having determined variance ξ_{T_j} of x for $j = i+1, ..., n$. We have that

$$\begin{aligned} Corr^{ma}(\ln y^{i,n+1-i}_{T_i}, \ln y^{i+1,n-i}_{T_{i+1}}) &= \frac{Cov^{ma}(\ln y^{i,n+1-i}_{T_i}, \ln y^{i+1,n-i}_{T_{i+1}})}{\sqrt{Var^{ma}(\ln y^{i,n+1-i}_{T_i}) Var^{ma}(\ln y^{i+1,n-i}_{T_{i+1}})}} \\ &\approx \frac{Cov^{ma}(\ln y^{i,n+1-i}_{T_i}, \ln y^{i+1,n-i}_{T_{i+1}})}{\widetilde{\sigma}^{ATM}_{i,n+1-i}\sqrt{T_i}\widetilde{\sigma}^{ATM}_{i+1,n-i}\sqrt{T_{i+1}}}, \end{aligned}$$

  where $Var^{ma}(\ln y^{i,n+1-i}_{T_i})$ is given by

$$\sqrt{Var^{ma}(\ln y^{i,n+1-i}_{T_i})} \approx \widetilde{\sigma}^{ATM}_{i,n+1-i}\sqrt{T_i}.$$

  It follows from equation (5.9) that we should get

$$\xi_{T_i} = [Corr^{ma}(\ln y^{i,n+1-i}_{T_i}, \ln y^{i+1,n-i}_{T_{i+1}})]^2 \xi_{T_{i+1}}. \tag{5.14}$$

As the specification of the one step covariance driving process depends on the swaption matrix $\{\tilde{\sigma}^{ATM}_{i,j}; i = 1, ..., n, j = 1, ..., n+1-i\}$ any changes in the market implied volatilities in the swaption matrix will influence the correlation structure. Therefore the one step covariance driver is also parameterized by time.

So far we have discussed the one-dimensional swap MFMs with three different specifications of the Gaussian driving process feeding in log-Normal marginal distributions of the swap rates $\{y^{i,n+1-i}_{T_i}; i = 1, ..., n\}$ in their associated swaption measures. Note that we choose these three specifications of the driver for comparison because they are examples of parametrizations by time and by expiry. We wish to investigate their hedging behaviour for a Bermudan swaption as well as a new Bermudan product. Moreover we wish to make clear if implied volatility smiles have an influence on their hedging behaviour. This motivates us to consider the one-dimensional swap MFM with the Gaussian driving process together with SABR marginals, which can be seen as a local volatility MFM. The SABR marginals that will be fed in are the same as the ones for the stochastic volatility swap MFM below. We will choose the Hull-White specification for the Gaussian driving process so that the local volatility MFM is also an example of parametrization by time.

Note that the reason why we choose the Hull-White specification as opposed to the mean reversion style is because the Hull-White driver is parameterized by time and it is expected to retain good hedging behaviour even though we feed in the SABR marginal distributions. The one-step covariance driver is not chosen because it is hard to calibrate to the market correlation in a local volatility MFM. By comparing the local volatility MFM to the Hull-White MFM, we can investigate the effect of implied volatility smiles on pricing and hedging a Bermudan swaption and a new Bermudan product.

5.2.4 Stochastic volatility Markov-functional model

In this subsection we discuss the stochastic volatility swap MFM that will be considered in the numerical study. We consider this model in the numerical comparison because we wish to investigate the effect of stochastic volatility on the pricing and hedging performance for Bermudan products. In what follows we focus on the calibration of the model.

We have discussed the specification and calibration of a stochastic volatility MFM in the previous chapters. In this subsection we review it briefly. Let us consider the target SABR driving process

$$\begin{aligned} dF_t &= \sigma_t F_t^{\beta} dW_t^{n+1} \qquad \beta \in [0,1] \\ d\sigma_t &= \mu_t \sigma_t dt + v\sigma_t dB_t^{n+1} \qquad v > 0, \end{aligned} \tag{5.15}$$

where B^{n+1} and W^{n+1} are correlated Brownian motions under the terminal measure such that $dB_t^{n+1} dW_t^{n+1} = \rho dt$ with $\rho \in [-1, 1]$, and the drift function μ_t is assumed to be piecewise constant

$$\mu_t = \sum_{j=0}^{n-1} \mu_j I_{\{t \in [T_j, T_{j+1})\}}. \tag{5.16}$$

In our numerical comparison we will assume that the parameters β, F_0, ρ, σ_0 and ν have been determined exogenously from historical data. The remaining parameters μ_j, $j = 0, ..., n-1$, can be chosen by calibrating to the market one step correlation (covariance) of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$. From the calibration procedure we note that the drift parameters μ_j's and the model correlation structure of swap rates will be changed when we change the market implied volatilities in the swaption matrix (see Table 5.2), which is consistent with the behaviour of a driver parameterized by time. In fact as we can see from (5.15), as the parameter v is going

to 0 the model is reduced to the one-dimensional one-step covariance MFM which is parameterized by time. This also motivates us to determine μ_i's by calibrating to the market one step correlation (covariance) of swap rates since we expect that the hedging property of the one step covariance MFM will be retained even though stochastic volatility is incorporated and we will justify this numerically later.

Following the calibration routine proposed in Chapter 4, in order to set the stochastic volatility MFM up, we feed in SABR marginal distributions of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$ implied from the following separable SABR swap market model in the swaption measure $\mathbb{S}^{i,n+1-i}$

$$\begin{aligned} dy_t^{i,n+1-i} &= \sigma_t^i (y_t^{i,n+1-i})^\beta dW_t^{i,n+1-i} \qquad \beta \in [0,1], \\ d\sigma_t^i &= \hat{\mu}^i(t, \sigma_t, y_t)\sigma_t^i dt + v\sigma_t^i dB_t^{i,n+1-i} \qquad \sigma_t^i = \gamma^i \sigma_t \\ dW_t^{i,n+1-i} dB_t^{i,n+1-i} &= \rho dt \qquad \rho \in [-1,1], \end{aligned} \tag{5.17}$$

with

$$\begin{aligned} \hat{\mu}^i(t, \sigma_t, y_t) &:= \mu_t + v\rho \sum_{j=i+1}^{n} \frac{\Psi_t^{j-1} \hat{P}_t^{j,n+1-j}}{\Psi_t^{i-1} \hat{P}_t^{i,n+1-i}} \left(\frac{\gamma^j \alpha_{j-1} (y_t^{j,n+1-j})^\beta}{1 + \alpha_{j-1} y_t^{j,n+1-j}}\right) \sigma_t \\ \Psi_t^i &:= \prod_{j=1}^{i} (1 + \alpha_j y_t^{j+1,n-j}), \\ \hat{P}_t^{i,n+1-i} &:= \frac{P_t^{i,n+1-i}}{D_{t,T_{n+1}}}, \end{aligned}$$

where $W^{i,n+1-i}$ and $B^{i,n+1-i}$ are correlated Brownian motions in the swaption measure $\mathbb{S}^{i,n+1-i}$, and μ_t is given by (5.16). We notice that the model parameters β, ρ, σ_0, ν and μ_j's have been determined. Consequently the only free parameters we left are expiry-dependent parameters γ^i for $i = 1, ..., n$. These parameters can be chosen by calibrating to the implied volatilities $\{\tilde{\sigma}_{i,n+1-i}; i = 1, ..., n\}$ of the co-terminal vanilla swaptions struck at one particular strike. In our numerical comparison, the implied volatilities $\tilde{\sigma}_{i,n+1-i}$ can be chosen to be the implied volatility $\tilde{\sigma}_{i,n+1-i}^{Berm}$ struck at the strike of the Bermudan swaption or $\tilde{\sigma}_{i,n+1-i}^{eff}$ struck at the effective strike depending on whether we are pricing a Bermudan swaption or the new Bermudan product.

5.2.5 Expressions for vegas in Markov-functional models

Before we proceed to the numerical investigation of the vegas, let us study the analytical expressions for the vegas of a Bermudan swaption in the swap MFMs

discussed earlier. This enables us to gain more insight into the difference in vegas produced by the models parameterized by time and by expiry. In this subsection we only consider a Bermudan swaption since it is a relatively simple product while the new Bermudan product is too complicated. The simplification means that we can focus on understanding the difference between parametrizations by time and by expiry in vega profiles without handling internal adjusters of a more involved new Bermudan product.

The expressions for the vegas in the one-dimensional swap MFMs have been studied by Kennedy and Pham [48]. Here we will first review their work. Then we make further development and consider the vegas in the stochastic volatility swap MFM.

We can see from Section 5.2.3 that the one-dimensional swap MFM is fully determined by the variances ξ_{T_i}, $i = 1, ..., n$, of the driving process and the input prices of the PVBP-digital swaptions. Furthermore the input prices are obtained from the initial term structure and some implied volatilities $\widetilde{\sigma}_{i,n+1-i}$, $i = 1, ..., n$. As we discussed the choice of $\widetilde{\sigma}_{i,n+1-i}$'s depends on the derivative we price. In this context we choose the implied volatilities $\widetilde{\sigma}^{Berm}_{i,n+1-i}$ of the co-terminal swaptions struck at the strike of the Bermudan swaption. Since we focus on the vegas here, the initial yield curve will not be taken into account. Thus for the model we can express the value of a Bermudan swaption as a function $V_0^{Berm}(\xi, \widetilde{\sigma}^{Berm})$ of the variances $\xi = (\xi_{T_1}, \xi_{T_2}, ..., \xi_{T_n})$ and the (reverse diagonal) input implied volatilities $\widetilde{\sigma}^{Berm} = (\widetilde{\sigma}^{Berm}_{1,n}, \widetilde{\sigma}^{Berm}_{2,n-1}, ..., \widetilde{\sigma}^{Berm}_{n,1})$. The vega $v_{i,k}$ for $i = 1, ..., n$ and $k = 1, ..., n + 1 - i$ is defined to be the following derivative:

$$v_{i,k} := \frac{dV_0^{Berm}}{d\widetilde{\sigma}^{ATM}_{i,k}},$$

where $\widetilde{\sigma}^{ATM}_{i,k}$ is the ATM implied volailities of swaptions. In practice the vegas can be evaluated by means of the bump and revalue method: each ATM implied volatility is perturbed by a basis point shift and then the derivative is valued again.

We now consider the analytical expressions for vegas in MFMs. For a particular $i = 1, ..., n$, we give the expressions for vegas $v_{i,k}$ in Table 5.3. The notations "mr", "hw", "co", "loc" and "sv" in Table 5.3 stand for mean reversion, Hull-White, one step covariance, local volatility and stochastic volatility MFMs respectively. Note that in the local volatility MFM we adopt the HW driver.

We can see from Table 5.3 that when $k = 1, ..., n - i$, the vegas $v_{i,k}$ of MR, HW and local volatility MFMs are all zero because the variances ξ of their drivers are independent of these implied volatilities $\{\widetilde{\sigma}^{ATM}_{i,k}; k = 1, ..., n - i\}$. However we

k (Tenor)	$1, ..., n-i$	$n+1-i$
mr	0	$\frac{\partial V_0^{Berm}}{\partial \widetilde{\sigma}_{i,n+1-i}^{Berm}} \times \theta_i$
hw	0	$\frac{\partial V_0^{Berm}}{\partial \widetilde{\sigma}_{i,n+1-i}^{Berm}} \times \theta_i + \frac{\partial V_0^{Berm}}{\partial \xi_{T_i}} \times \frac{d\xi_{T_i}}{d\widetilde{\sigma}_{i,n+1-i}^{ATM}}$
co	$\sum_{s=1}^{i} \frac{\partial V_0^{Berm}}{\partial \xi_{T_s}} \times \frac{d\xi_{T_s}}{d\widetilde{\sigma}_{i,k}^{ATM}}$	$\frac{\partial V_0^{Berm}}{\partial \widetilde{\sigma}_{i,n+1-i}^{Berm}} \times \theta_i + \sum_{s=1}^{i} \frac{\partial V_0^{Berm}}{\partial \xi_{T_s}} \times \frac{d\xi_{T_s}}{d\widetilde{\sigma}_{i,n+1-i}^{ATM}}$
loc	0	$\frac{\partial V_0^{Berm}}{\partial \widetilde{\sigma}_{i,n+1-i}^{Berm}} \times \theta_i + \frac{\partial V_0^{Berm}}{\partial \xi_{T_i}} \times \frac{d\xi_{T_i}}{d\widetilde{\sigma}_{i,n+1-i}^{ATM}}$
sv	$\sum_{s=0}^{n-1} \frac{\partial \widetilde{V}_0^{Berm}}{\partial \mu_s} \times \frac{d\mu_s}{d\widetilde{\sigma}_{i,k}^{ATM}}$	$\frac{\partial \widetilde{V}_0^{Berm}}{\partial \widetilde{\sigma}_{i,n+1-i}^{Berm}} \times \theta_i + \sum_{s=0}^{n-1} \frac{\partial \widetilde{V}_0^{Berm}}{\partial \mu_s} \times \frac{d\mu_s}{d\widetilde{\sigma}_{i,n+1-i}^{ATM}}$

Table 5.3: The expressions for vegas $v_{i,k}$ for a particular i of a Bermudan swaption under swap MFMs where $\theta_i := \frac{d\widetilde{\sigma}_{i,n+1-i}^{Berm}}{d\widetilde{\sigma}_{i,n+1-i}^{ATM}}$.

can see from Table 5.2 that these implied volatilities $\{\widetilde{\sigma}_{i,k}^{ATM}; k = 1, ..., n-i\}$ are linked to the variances ξ of the one step covariance driver. In particular following the estimation procedure for the market one step covariance in [48] one can see that any change in implied volatility $\widetilde{\sigma}_{i,k}^{ATM}$ will lead to a change in ξ_{T_i} and therefore ξ_{T_j} for all $j = 1, ..., i-1$ due to equation (5.14). Therefore following the chain rule we obtain the result in Table 5.3.

When $k = n+1-i$, the vegas become more complicated. In this case all models are linked to the co-terminal ATM implied volatility $\widetilde{\sigma}_{i,n+1-i}^{ATM}$ via the input implied volatility $\widetilde{\sigma}_{i,n+1-i}^{Berm}$ of the function $V_0^{Berm}(\xi, \widetilde{\sigma}^{Berm})$. Therefore as we can see in Table 5.3, there is a term $\frac{\partial V_0^{Berm}}{\partial \widetilde{\sigma}_{i,n+1-i}^{Berm}} \times \theta_i$ where $\theta_i := \frac{d\widetilde{\sigma}_{i,n+1-i}^{Berm}}{d\widetilde{\sigma}_{i,n+1-i}^{ATM}}$ in all models. We will discuss this ratio θ_i later. In addition, we can see from (5.13) that a change in the implied volatility $\widetilde{\sigma}_{i,n+1-i}^{ATM}$ will also result in a change in ξ_{T_i} of the HW driver and therefore this gives the result in Table 5.3. Similarly it will also result in a change in ξ_{T_i} and therefore ξ_{T_j} for $j = 1, ..., i-1$ for the one step covariance driver.

We can see that the expressions for the vegas in the above models are very different. However we will see in the numerical study later that the row sums of vegas for the models parameterized by time are similar.

Remark 9. *When the implied volatility $\widetilde{\sigma}_{i,n+1-i}^{Berm}$ is chosen to be the ATM one $\widetilde{\sigma}_{i,n+1-i}^{ATM}$, it is trivial to have that $\theta_i := \frac{d\widetilde{\sigma}_{i,n+1-i}^{Berm}}{d\widetilde{\sigma}_{i,n+1-i}^{ATM}} = 1$. In practice the market implied volatilities always display a shape of smile or skew as a function of strike. The implied volatilities with different strikes are always moving simultaneously. This means that the change in the ATM implied volatility $\widetilde{\sigma}_{i,n+1-i}^{ATM}$ will result in a change in the implied volatility $\widetilde{\sigma}_{i,n+1-i}^{Berm}$ with other strike. In order to find their relationship, one can calibrate a SABR model to the market implied volatility smile/skew, and then find the derivative $\frac{d\widetilde{\sigma}_{i,n+1-i}^{Berm}}{d\widetilde{\sigma}_{i,n+1-i}^{ATM}}$ within the calibrated SABR model using the*

SABR formula since the SABR formula provides a parametric formula for $\widetilde{\sigma}^{Berm}_{i,n+1-i}$ *which depends on the ATM implied volatility* $\widetilde{\sigma}^{ATM}_{i,n+1-i}$*. Alternatively one can use a simpler approximation for the derivative* $\frac{d\widetilde{\sigma}^{Berm}_{i,n+1-i}}{d\widetilde{\sigma}^{ATM}_{i,n+1-i}}$*. There are numerical studies in the literature showing that the moving of implied volatility smiles/skews can be decomposed into three principal factors which account for 98% of the variance. Among them, the first factor which can explain around 80% of the variance can be viewed as a level effect. For more details about the study, the reader is referred to [16] and references therein. Therefore it is reasonable to assume that smiles/skews move in a parallel way. As a result we have an approximation* $\theta_i := \frac{d\widetilde{\sigma}^{Berm}_{i,n+1-i}}{d\widetilde{\sigma}^{ATM}_{i,n+1-i}} \approx 1$*.*

Having discussed the expressions for the vegas of a Bermudan swaption in a one-dimensional swap MFM, we now consider the vegas in a stochastic volatility swap MFM. Note that a stochastic volatility swap MFM is fully determined by the market prices of co-terminal vanilla swaptions and the SABR driver (5.15). As the parameters β, F_0, ρ, σ_0 and ν are assumed to be determined exogenously from historical data, the remaining piecewise constant function μ_i's can be chosen by fitting the market one step correlation (covariance). Consequently one can express the value of a Bermudan swaption as a function $\widetilde{V}^{Berm}_0(\mu, \widetilde{\sigma}^{Berm})$ of $\mu = (\mu_0, \mu_1, ..., \mu_{n-1})$ and implied volatilities $\widetilde{\sigma}^{Berm} = (\widetilde{\sigma}^{Berm}_{1,n}, \widetilde{\sigma}^{Berm}_{2,n-1}, ..., \widetilde{\sigma}^{Berm}_{n,1})$.

In line with the notation introduced in proposition 4 of Chapter 4, one can write

$$\zeta_{T_i} = g^i(\zeta_{T_{i+1}}, ..., \zeta_{T_n}; \{\widetilde{\sigma}^{ATM}_{i,k}\}_{k=i,...,n+1-i}; \{\widetilde{\sigma}^{ATM}_{s,n+1-s}\}_{s=i+1,...,n}), \tag{5.18}$$

$$\mu_j = f^j(\mu_0, ..., \mu_{j-1}; \zeta_{T_1}, ..., \zeta_{T_{j+1}}). \tag{5.19}$$

for $i = 1, ..., n$ and $j = 0, ..., n-1$ where f^j and g^i are some deterministic functions. Since the stochastic volatility MFM is calibrated to the market one step correlation (covariance), its vegas' expressions are similar to the one step covariance MFM. The only difference lies in μ. In particular from (5.18) and (5.19) we can see that as the ATM implied volatility $\widetilde{\sigma}^{ATM}_{i,k}$ changes, all μ_j for $j = 0, ..., n-1$ will be changed. This gives the expressions for vegas in Table 5.3 under the stochastic volatility MFM.

5.3 Market data

In this section we give the market data used in the numerical study. In particular we consider the market data on 11 March 2015. In our numerical study we consider 31 years annual co-terminal Bermudan products. The initial ZCBs with maturity up to 31 years are bootstrapped from the market data and displayed in Table 5.4. The

co-terminal (with maturity 31 years) swaption volatilities and caplets volatilities w.r.t expiry and strike are given in Tables 5.5 and 5.6. The implied volatilities of the ATM swaptions w.r.t expiry and tenor are displayed in Tables 5.7, 5.8 and 5.9, which will be used to extract market correlation structure of swap rates.

Maturity	1Y	2Y	3Y	4Y	5Y	6Y	7Y
Price	0.9951	0.9820	0.9650	0.9440	0.9207	0.8970	0.8711
Maturity	8Y	9Y	10Y	11Y	12Y	13Y	14Y
Price	0.8478	0.8235	0.7983	0.7788	0.7593	0.7400	0.7208
Maturity	15Y	16Y	17Y	18Y	19Y	20Y	21Y
Price	0.7018	0.6828	0.6640	0.6454	0.6269	0.6087	0.5905
Maturity	22Y	23Y	24Y	25Y	26Y	27Y	28Y
Price	0.5726	0.5549	0.5374	0.5200	0.5029	0.4860	0.4694
Maturity	29Y	30Y	31Y				
Price	0.4529	0.4367	0.4208				

Table 5.4: Initial zero-coupon bonds on 11 March 2015.

5.4 Bermudan swaption comparison

In this section we investigate the pricing and hedging performance for a Bermudan swaption under different MFMs. For the hedging comparison we will only consider the vegas, and we will explain why we do this later. The example we consider here is a 31 years pay fixed Bermudan swaption with notional $N = 100$ million. There are 30 annual exercise dates and the first exercise date of the option is in 1 year relative to today.

In Section 5.4.1 we compare the prices of the Bermudan swaption produced by different MFMs. We then compare the different MFMs in terms of the vegas in Section 5.4.2.

5.4.1 Pricing comparison results

We consider the Bermudan swaption struck at three different strikes: 2%, 3% and 4%. We display the prices in Table 5.11. In this table, the terms "MR", "HW", "1-step Cov", "Local vol" and "SV" denote the MFMs (a) - (e) that are introduced at the beginning of Section 5.2, where we use a HW driving process with $a = 0.05$ for the local volatility MFM. We set the parameter a to be 1%, 5% and 10% respectively for the MR and HW MFMs. By doing so we can investigate the impact of the correlation structure on the prices of the Bermudan swaption. In the numerical study we consider three scenarios for the SV MFM which are shown in Table 5.10.

Relative strike		-200	-150	-100	-50	0	50	100	150	200
Expiry	1Y	34.92	28.64	24.32	21.89	19.71	19.56	19.43	19.31	19.21
	2Y	33.72	28.08	23.83	21.52	19.60	19.37	19.19	19.03	18.89
	3Y	32.84	27.65	23.44	21.19	19.40	19.12	18.90	18.71	18.54
	4Y	32.22	27.28	23.06	20.86	19.17	18.85	18.60	18.38	18.20
	5Y	31.69	26.89	22.69	20.54	18.93	18.59	18.31	18.07	17.87
	6Y	31.23	26.60	22.36	20.21	18.70	18.34	18.04	17.79	17.58
	7Y	30.78	26.37	21.98	19.98	18.45	18.06	17.74	17.47	17.24
	8Y	30.27	26.18	21.61	19.73	18.17	17.75	17.41	17.12	16.88
	9Y	29.85	25.96	21.34	19.51	18.05	17.60	17.23	16.93	16.67
	10Y	29.45	25.71	21.00	19.23	17.72	17.24	16.86	16.55	16.28
	11Y	29.02	25.57	20.74	19.01	17.57	17.07	16.68	16.35	16.07
	12Y	28.62	25.35	20.51	18.98	17.51	17.00	16.59	16.25	15.97
	13Y	28.23	24.77	20.35	18.63	17.49	16.96	16.54	16.19	15.89
	14Y	27.84	24.42	20.16	18.34	17.46	16.92	16.48	16.12	15.82
	15Y	27.47	23.91	19.98	18.27	17.43	16.87	16.42	16.05	15.74
	16Y	27.10	23.71	19.80	18.12	17.37	16.80	16.34	15.96	15.64
	17Y	26.73	23.54	19.62	18.08	17.32	16.73	16.25	15.86	15.54
	18Y	26.37	23.26	19.41	18.07	17.28	16.67	16.19	15.79	15.45
	19Y	26.01	22.96	19.33	18.07	17.25	16.62	16.12	15.71	15.36
	20Y	25.70	22.69	19.32	18.05	17.23	16.58	16.06	15.64	15.28
	21Y	25.42	22.35	19.31	18.01	17.23	16.55	16.02	15.58	15.21
	22Y	25.19	22.02	19.18	17.97	17.09	16.42	15.89	15.46	15.09
	23Y	24.97	21.98	19.11	17.88	17.00	16.33	15.79	15.36	15.00
	24Y	24.77	21.75	19.05	17.76	16.86	16.19	15.65	15.22	14.86
	25Y	24.59	21.60	19.03	17.72	16.79	16.09	15.55	15.10	14.74
	26Y	24.45	21.56	19.02	17.71	16.78	16.07	15.52	15.08	14.72
	27Y	24.33	21.46	18.93	17.61	16.62	15.88	15.31	14.86	14.50
	28Y	24.23	21.42	18.87	17.58	16.51	15.64	15.01	14.51	14.12
	29Y	24.13	21.23	18.73	17.28	16.44	15.29	14.83	14.37	14.03
	30Y	24.03	21.14	18.65	17.03	16.14	15.08	14.71	14.21	13.89

Table 5.5: Implied volatilities (%) of the co-terminal (31Y) swaptions against expiry and relative strike (bp) to the ATM on 11 March 2015

In the first two scenarios (I and II) the parameters ν and σ_0 are not too big as usually expected in practice. Scenario III has a high Volvol ν and σ_0 which can be seen as an extreme case. We also produce the "total" vega for each model which stands for the difference in Bermudan price when all implied volatilities in the swaption matrix have a parallel shift of 1%. This total vega can be used to judge how significant the difference in Bermudan prices is. Table 5.12 gives the difference in Bermudan prices produced by different MFMs and the SV MFM (I) corresponding to the scenario I measured by the total vega. In particular the ratios in the table are produced by using the following formula

$$\text{ratio} = \frac{\text{SV MFM (I) price - MFM price}}{\text{SV MFM (I) total vega}}. \tag{5.20}$$

Relative strike		-200	-150	-100	-50	0	50	100	150	200
Expiry	1Y	39.27	33.72	30.12	28.42	26.96	28.26	29.58	30.91	32.26
	2Y	38.12	33.21	29.70	28.12	26.93	28.17	29.45	30.76	32.09
	3Y	37.19	32.72	29.23	27.71	26.64	27.81	29.03	30.29	31.57
	4Y	36.51	32.29	28.78	27.30	26.32	27.43	28.61	29.83	31.07
	5Y	35.83	31.72	28.21	26.75	25.82	26.86	27.96	29.11	30.29
	6Y	34.96	30.96	27.34	25.81	24.93	25.81	26.76	27.75	28.78
	7Y	34.25	30.42	26.61	25.19	24.24	25.00	25.84	26.73	27.65
	8Y	33.35	29.77	25.71	24.34	23.29	23.90	24.58	25.32	26.11
	9Y	32.68	29.27	25.12	23.76	22.77	23.26	23.84	24.48	25.17
	10Y	32.05	28.75	24.48	23.14	22.06	22.46	22.95	23.50	24.10
	11Y	31.34	28.28	23.83	22.49	21.43	21.71	22.09	22.53	23.03
	12Y	30.60	27.65	23.14	21.94	20.80	20.95	21.20	21.52	21.89
	13Y	30.00	26.83	22.70	21.28	20.43	20.49	20.65	20.89	21.19
	14Y	29.40	26.23	22.23	20.67	20.05	20.06	20.11	20.27	20.48
	15Y	28.82	25.48	21.78	20.29	19.68	19.63	19.57	19.52	19.44
	16Y	28.35	25.17	21.47	19.99	19.45	19.30	19.25	19.13	19.03
	17Y	27.88	24.88	21.15	19.80	19.23	19.02	18.93	18.72	18.68
	18Y	27.40	24.46	20.78	19.62	19.00	18.74	18.59	18.54	18.34
	19Y	26.93	24.03	20.55	19.44	18.77	18.45	18.26	18.15	18.01
	20Y	26.49	23.61	20.37	19.24	18.55	18.16	17.91	17.75	17.55
	21Y	26.08	23.12	20.19	19.00	18.33	17.87	17.56	17.34	17.19
	22Y	25.80	22.73	20.00	18.88	18.11	17.64	17.31	17.08	16.92
	23Y	25.51	22.60	19.82	18.68	17.89	17.40	17.04	16.79	16.60
	24Y	25.25	22.31	19.69	18.48	17.67	17.15	16.78	16.51	16.31
	25Y	24.98	22.07	19.56	18.31	17.45	16.89	16.47	16.17	15.93
	26Y	24.69	21.84	19.35	18.08	17.19	16.56	16.10	15.74	15.46
	27Y	24.52	21.68	19.18	17.89	16.93	16.25	15.75	15.36	15.06
	28Y	24.32	21.53	19.00	17.72	16.67	15.83	15.22	14.76	14.40
	29Y	24.28	21.34	18.81	17.37	16.51	15.37	14.93	14.48	14.15
	30Y	24.03	21.14	18.65	17.03	16.14	15.08	14.71	14.21	13.89

Table 5.6: Implied volatilities (%) of the caplets against expiry and relative strike (bp) to the ATM on 11 March 2015

From Tables 5.11 and 5.12 we draw the following conclusions.

1. The parameter a for the MR MFM and HW MFM has a significant impact on the price of the Bermudan swaption for all strikes. A 5% change in the parameter a can result in a change in value equal to a parallel implied volatilities shift of about 4%. The correlation of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$ is controlled by the parameter a in the MR and HW MFMs. In particular as the parameter a increases, the correlation decreases. It is well known that a lower correlation of swap rates gives a bigger Bermudan price. Intuitively a Bermudan swaption can be seen as a combination of the corresponding vanilla swaptions and their optionality. The "optionality value" depends on the correlation of swap rates. In particular the lower the correlations become, the higher the optionality value obtained.

Tenor		1Y	2Y	3Y	4Y	5Y	6Y	7Y	8Y	9Y	10Y
Expiry	1Y	26.96	27.01	26.95	26.84	26.77	26.22	25.66	25.14	24.64	24.10
	2Y	26.93	26.85	26.59	26.39	26.05	25.53	25.00	24.49	24.00	23.63
	3Y	26.64	26.35	26.13	25.83	25.41	24.81	24.24	23.80	23.44	23.04
	4Y	26.32	25.94	25.63	25.09	24.60	24.13	23.59	23.22	22.83	22.52
	5Y	25.82	25.33	24.94	24.43	23.91	23.47	23.01	22.61	22.36	22.03
	6Y	24.93	24.43	24.18	23.62	23.13	22.73	22.34	22.00	21.68	21.44
	7Y	24.24	23.72	23.37	22.90	22.56	22.14	21.81	21.50	21.22	21.00
	8Y	23.29	22.93	22.63	22.20	21.89	21.58	21.34	21.07	20.87	20.68
	9Y	22.77	22.32	22.00	21.65	21.34	21.02	20.86	20.67	20.50	20.33
	10Y	22.06	21.61	21.37	21.08	20.82	20.58	20.45	20.28	20.13	19.98
	11Y	21.43	21.08	20.88	20.64	20.41	20.20	20.09	19.96	19.81	19.67
	12Y	20.80	20.55	20.40	20.21	19.99	19.83	19.74	19.63	19.49	19.36
	13Y	20.43	20.22	20.09	19.92	19.72	19.57	19.49	19.39	19.25	19.12
	14Y	20.05	19.90	19.77	19.63	19.45	19.31	19.23	19.15	19.02	18.88
	15Y	19.68	19.57	19.46	19.35	19.18	19.05	18.98	18.91	18.78	18.65
	16Y	19.45	19.34	19.22	19.10	18.95	18.82	18.74	18.66	18.55	18.41
	17Y	19.23	19.11	18.97	18.86	18.72	18.59	18.50	18.42	18.31	18.17
	18Y	19.00	18.87	18.73	18.61	18.49	18.36	18.26	18.18	18.07	17.94
	19Y	18.77	18.64	18.48	18.36	18.26	18.14	18.02	17.94	17.84	17.70
	20Y	18.55	18.40	18.24	18.12	18.04	17.91	17.78	17.70	17.60	17.46
	21Y	18.33	18.19	18.05	17.92	17.83	17.68	17.55	17.46	17.35	17.23
	22Y	18.11	17.97	17.86	17.72	17.62	17.46	17.32	17.23	17.09	
	23Y	17.89	17.75	17.67	17.53	17.41	17.24	17.09	17.00		
	24Y	17.67	17.53	17.48	17.33	17.20	17.01	16.86			
	25Y	17.45	17.32	17.29	17.13	16.99	16.79				
	26Y	17.19	17.10	17.03	16.87	16.78					
	27Y	16.93	16.88	16.77	16.62						
	28Y	16.67	16.66	16.51							
	29Y	16.51	16.44								
	30Y	16.14									

Table 5.7: Implied volatilities (%) of the ATM swaptions against expiry and tenor (1Y - 10Y) on 11 March 2015

2. In order to study the implied volatility smiles effect, we compare the Bermudan price produced by the local volatility MFM, where we captured the volatility smile, with the price given by the HW (5%) MFM. We observe that the two models give very similar Bermudan prices. This implies that the impact of smile on the Bermudan price is small. Note that this conclusion depends very much on the calibration approach. In particular the local volatility MFM is set up by using the HW driving process with the parameter $a = 5\%$ together with SABR marginal distributions of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$ implied from the separable SABR swap market model (5.17) corresponding to the parameters of scenario I shown in Table 5.10. Recall that the model (5.17) only has free parameters γ^i, $i = 1, ..., n$, while the other parameters have already been determined. Commonly there are two calibration approaches to

Tenor		11Y	12Y	13Y	14Y	15Y	16Y	17Y	18Y	19Y	20Y
Expiry	1Y	23.72	23.33	23.00	22.67	22.34	22.12	21.89	21.67	21.45	21.22
	2Y	23.23	22.84	22.50	22.16	21.82	21.62	21.42	21.22	21.03	20.83
	3Y	22.69	22.35	22.05	21.76	21.46	21.26	21.05	20.85	20.64	20.44
	4Y	22.21	21.91	21.65	21.39	21.13	20.92	20.72	20.52	20.32	20.11
	5Y	21.75	21.46	21.21	20.97	20.72	20.53	20.33	20.14	19.95	19.76
	6Y	21.22	21.00	20.77	20.53	20.29	20.09	19.90	19.71	19.52	19.33
	7Y	20.80	20.60	20.42	20.23	20.05	19.84	19.62	19.41	19.20	18.98
	8Y	20.44	20.21	20.06	19.91	19.76	19.53	19.30	19.07	18.85	18.62
	9Y	20.10	19.87	19.75	19.63	19.51	19.28	19.05	18.81	18.58	18.35
	10Y	19.78	19.59	19.39	19.19	19.00	18.77	18.54	18.31	18.08	17.84
	11Y	19.48	19.28	19.09	18.89	18.70	18.47	18.25	18.02	17.79	17.57
	12Y	19.17	18.98	18.79	18.59	18.40	18.18	17.96	17.73	17.51	
	13Y	18.93	18.74	18.54	18.35	18.15	17.93	17.71	17.49		
	14Y	18.69	18.50	18.30	18.10	17.90	17.68	17.46			
	15Y	18.45	18.26	18.05	17.85	17.65	17.43				
	16Y	18.21	18.01	17.79	17.58	17.37					
	17Y	17.96	17.75	17.53	17.32						
	18Y	17.72	17.50	17.28							
	19Y	17.47	17.25								
	20Y	17.23									

Table 5.8: Implied volatilities (%) of the ATM swaptions against expiry and tenor (11Y - 20Y) on 11 March 2015

determine the parameters γ^i's. The first choice is to fix γ^i's by fitting the implied volatilities of the co-terminal vanilla swaptions struck at the strike of the Bermudan swaption. This calibration approach can ensure that the most relevant vanilla swaptions to the Bermudan swaption can be reproduced exactly by the model. The second choice however is to determine the parameters γ^i's by calibrating to the whole implied volatility smile of the co-terminal vanilla swaptions as much as we can by using the least square method. Note that the free parameters γ^i, $i = 1, ..., n$, are not sufficient to fit the whole smile. In this sense the second calibration approach is like a "global fit", where we try to match everything, while the first approach is more like a "local fit", where we just make the most important thing right. In our numerical comparison we choose the first calibration approach. So the local volatility MFM and the HW MFM are both calibrated to the implied volatilities of the vanilla swaptions struck at the strike of the Bermudan swaption, and therefore these two models give very similar Bermudan prices. Pietersz and Pelsser [57] also investigated the smile impact on the Bermudan price by comparing a mean reversion MFM to a local volatility MFM with a mean reversion driver together with displaced diffusion marginals. In the comparison the "global fit" is used and therefore this leads to a very different result which indicates that the smile impact is

Tenor		21Y	22Y	23Y	24Y	25Y	26Y	27Y	28Y	29Y	30Y
Expiry	1Y	21.06	20.91	20.75	20.59	20.44	20.29	20.15	20.00	19.85	19.71
	2Y	20.68	20.53	20.38	20.23	20.08	19.96	19.84	19.72	19.60	
	3Y	20.30	20.17	20.04	19.91	19.78	19.65	19.52	19.40		
	4Y	19.98	19.84	19.70	19.57	19.43	19.30	19.17			
	5Y	19.62	19.48	19.34	19.20	19.06	18.93				
	6Y	19.20	19.08	18.95	18.83	18.70					
	7Y	18.85	18.72	18.59	18.45						
	8Y	18.47	18.32	18.17							
	9Y	18.20	18.05								
	10Y	17.72									

Table 5.9: Implied volatilities (%) of the ATM swaptions against expiry and tenor (21Y - 30Y) on 11 March 2015

large.

3. The one step covariance MFM gives a systematically and higher Bermudan price than the SV MFM for all the three scenarios. But the difference is very small. This means that the introduction of stochastic volatility has an insignificant influence on the price of the Bermudan swaption. For most cases the one-dimensional MFM is sufficient for pricing a Bermudan swaption as long as the correlation of the underlying swap rates can be captured appropriately by the valuation model.

In summary the numerical results indicate that the correlation impact on the Bermudan price is very large. The effect of smile is very small if we fit the "right" implied volatilities of swaptions. The introduction of stochastic volatility has an insignificant influence on the Bermudan price.

Scenario	ν	σ_0	β	ρ
I	0.3	0.05	0.5	-0.7
II	0.1	0.03	-0.7	0.8
III	0.5	0.1	0	-0.2

Table 5.10: Scenarios for the SV MFM.

5.4.2 Vega comparison results

In this subsection we perform a similar numerical comparison in terms of vegas to Kennedy and Pham [48]. In particular we consider the sum of the vegas $\sum_{k=1}^{n+1-i} v_{i,k}$ for $i = 1, ..., n$ of the Bermudan swaption struck at the strike $K = 3\%$ under the different MFMs. But in addition to the one-dimensional MFMs with Gaussian drivers together with log-Normal marginals in [48], we make a further development

Strike	K=2%	K=3%	K=4%
MR (1%) MFM	21,861,477	13,440,036	8,837,719
Vega	395,023	614,416	656,073
MR (5%) MFM	24,073,433	15,814,235	10,877,533
Vega	569,952	794,238	846,667
MR (10%) MFM	26,264,916	19,031,754	13,659,641
Vega	829,426	1,050,090	1,110,018
HW (1%) MFM	21,133,705	12,755,708	8,335,399
Vega	385,009	605,845	636,633
HW (5%) MFM	23,295,892	15,017,123	10,200,357
Vega	548,153	766,518	807,385
HW (10%) MFM	25,418,062	18,020,595	12,678,291
Vega	779,854	993,464	1,039,620
1-step Cov MFM	24,070,296	15,808,558	10,870,143
Vega	520,947	733,306	760,078
Local vol MFM	23,237,083	15,001,473	10,211,473
Vega	551,233	772,472	812,312
SV MFM (I)	23,896,450	15,610,384	10,651,394
Vega	496,492	712,003	737,127
SV MFM (II)	23,953,742	15,739,932	10,791,403
Vega	509,193	720,932	749,830
SV MFM (III)	23,684,720	15,341,035	10,381,394
Vega	473,105	692,027	719,093

Table 5.11: The prices of the Bermudan swaption under different MFMs with notional $N = 100$ million.

in our numerical study and also take the local volatility MFM and the stochastic volatility MFM into account. Another difference is that we consider a 31Y annual Bermudan swaption here as opposed to 11Y in [48]. It is more challenging to price and hedge a longer maturity Bermudan swaption in practice and the difference between models is more significant.

We provide the results under the different MFMs in Figure 5.1. We set the parameter $a = 5\%$ for the mean reversion MFM and $a = 6\%$ for the Hull-White MFM and local volatility MFM so that these MFMs produce comparable Bermudan prices to the one step covariance MFM.

We draw the following conclusions from Figure 5.1.

1. We can see that the mean reversion MFM gives a very different vega profile from the other MFMs. We recall Section 5.2.5 that the different parametrization of the driving process yields a very different vega profile for a Bermudan swaption. This makes the vegas produced by the mean reversion MFM, which is parameterized by expiry, very different from the other MFMs, which are

Strike	K=2%	K=3%	K=4%
MR (1%) MFM	4.10	3.05	2.46
MR (5%) MFM	-0.36	-0.29	-0.31
MR (10%) MFM	-4.77	-4.81	-4.08
HW (1%) MFM	5.56	4.01	3.14
HW (5%) MFM	1.21	0.83	0.61
HW (10%) MFM	-3.06	-3.39	-2.75
1-step Cov MFM	-0.35	-0.28	-0.30
Local vol MFM	1.33	0.86	0.60
SV MFM (I)	0	0	0
SV MFM (II)	-0.12	-0.18	-0.19
SV MFM (III)	0.43	0.38	0.37

Table 5.12: The difference between the price for the Bermudan swaption produced by different MFMs and the SV MFM (I) measured by total vega.

examples of parametrizations by time. This result is consistent with the observation made in [48]. Note that although the stochastic volatility MFM is parameterized by time, it is not obvious this would override any effect coming from the addition of stochastic volatility. But we can see from the numerical result that although the stochastic volatility MFM produces a systematically lower row sum of vegas than the other three models of parametrizations by time, their vega profiles are still very similar.

2. The vegas produced by the Hull-White and local volatility MFMs are very close. This means that the impact of smile on the vegas of the Bermudan swaption is insignificant. The stochastic volatility MFM however gives a lower vega than the one-step covariance MFM. But the difference is insignificant in comparison to the effect of the different parametrizations.

In summary the vega profile produced by the MFM parameterized by expiry is very different from the models parameterized by time, and this is consistent with the conclusion made in [48]. The introduction of stochastic volatility makes the row sum of vegas slightly lower than the one step covariance MFM but their vega profiles are still very similar. Moreover the smile effect on the vega profile of the Bermudan swaption is insignificant.

It was noted by Kennedy and Pham [48] that the driving processes parameterized by time lead to the similar vega profiles which will control the sum of all gammas and therefore give a similar total gamma for a Bermudan swaption. Furthermore the fundamental difference in parametrizations by time and by expiry will result in the difference in vega profiles which has an effect on total gamma. The

authors concluded that the parametrization by time outperforms the driver parameterized by expiry in terms of the total gamma. In our numerical comparison we only consider the vegas. We have found that the stochastic volatility MFM has a very similar vega profile to the other MFMs parameterized by time. From the conclusions made in [48], the total gamma produced by the stochastic volatility MFM should be very similar to the other MFMs parameterized by time, and very different from the mean reversion MFM which is an example of parametrization by expiry.

So far we have compared the prices and the vegas of a Bermudan swaption under MFMs. It turns out that stochastic volatility has an insignificant effect. In the next section we will do a similar numerical comparison based on the new Bermudan product which is a more complicated Bermudan product than Bermudan swaption. For such an option, the stochastic volatility dynamics could play a more important role and we will see this in the next section.

5.5 New Bermudan product comparison

In this section we consider the above numerical comparison for the new Bermudan product. The example we are interested in here is a 31 years pay fixed new Bermudan product with notional $N = 100$ million. The lower and upper barriers for this new Bermudan product are 0 and 5%. The option has 30 annual exercise dates and the first exercise date is in 1 year relative to today.

In what follows, we first investigate the prices of the new Bermudan product based on the different MFMs. Then we compare the vegas of the new Bermudan product.

5.5.1 Pricing comparison results

In Table 5.13 we give the prices of the new Bermudan product struck at three different strikes under the different MFMs. Again we consider three scenarios for the SV MFM as shown in Table 5.10. As before in Section 5.4.1, in Table 5.14, we give the difference in the prices of the new Bermudan product produced by different MFMs and the SV MFM (I) measured in terms of the total vega by following the formula (5.20). From Tables 5.13 and 5.14 we draw the following conclusions.

1. Similar to the Bermudan swaption, the parameter a of the mean reversion MFM and Hull-White MFM has a large influence on the price of the new Bermudan product. In particular, a 5% change in the parameter a can result in a difference in price of about a parallel change of 5% of all implied volatilities in

the swaption matrix. As the parameter a becomes bigger, the price of the new Bermudan product gets higher. This numerical observation can be explained by following a similar discussion to that for the Bermudan swaption in the previous section. In particular a higher value of the parameter a gives a lower correlation structure and therefore yields a higher price of the new Bermudan product. Therefore we can conclude that the correlation structure has a large effect on the price of the new Bermudan product.

2. The price of the new Bermudan product produced by the local volatility MFM is close to the one produced by the Hull-White MFM. This means that the volatility smile has little effect on the price of the new Bermudan product.

3. To study the influence of stochastic volatility on the price of the new Bermudan product, we consider the prices given by the one-step covariance MFM and the stochastic volatility MFM. We can see that the stochastic volatility effect can be as large as a 1.55% of the total vega which can be seen as a significant effect. The effect becomes as large as a 2% of the total vega for extreme scenario III. Recall that for the Bermudan swaption the effect of the stochastic volatility is insignificant, but the impact is large here for the new Bermudan product. This is because the new Bermudan product is more complicated and its floating leg involves more information on the stochastic volatility dynamics. This could make the stochastic volatility factor more important for the new Bermudan product.

In summary the numerical results indicate that the correlation and stochastic volatility have large effects on the new Bermudan product price while the smile effect is very small. Note that stochastic volatility has a small effect on pricing a relatively simple Bermudan product such as a Bermudan swaption. But when the Bermudan product becomes more complex such as the new Bermudan product considered here, the floating leg coupon involves more information of model dynamics. In this case the stochastic volatility factor becomes more important for the product. We expect that the effect of stochastic volatility factor could be more significant for callable range accruals.

5.5.2 Vega comparison results

In this subsection we investigate the vegas of the pay fixed new Bermudan product struck at the strike $K = 3\%$ under the different MFMs. We perform a similar numerical study to Section 5.4.2. In particular we consider the sum of the vegas

Strike	K=2%	K=3%	K=4%
MR (1%) MFM	14,016,903	5,598,461	1,258,882
Vega	-1,070,703	-788,070	-558,876
MR (5%) MFM	16,230,628	7,981,468	3,272,713
Vega	-896,069	-617,335	-463,760
MR (10%) MFM	19,423,641	11,219,197	6,055,144
Vega	-709,409	-394,736	-348,397
HW (1%) MFM	13,290,844	4,915,105	1,025,339
Vega	-1,081,358	-801,991	-495,006
HW (5%) MFM	15,453,023	7,177,814	2,598,037
Vega	-918,114	-635,499	-481,800
HW (10%) MFM	18,575,322	10,191,007	5,254,696
Vega	-683,367	-424,656	-362,022
1-step Cov MFM	16,227,478	7,974,543	3,264,231
Vega	-969,372	-682,934	-533,382
Local vol MFM	15,428,372	7,150,118	2,544,429
Vega	-920,027	-634,008	-478,273
SV MFM (I)	15,248,392	7,004,218	2,310,376
Vega	-1,034,092	-746,404	-613,735
SV MFM (II)	15,800,743	7,537,331	2,829,170
Vega	-990,015	-702,701	-563,093
SV MFM (III)	14,754,820	6,601,083	1,908,403
Vega	-1,070,472	-772,038	-659,482

Table 5.13: The prices of the new Bermudan product under different MFMs with notional $N = 100$ million.

$\sum_{k=1}^{n+1-i} v_{i,k}$ for $i = 1, ..., n$ of the new Bermudan product under the different MFMs. The numerical result is given in Figure 5.2. In this numerical comparison we set the parameter $a = 5\%$ for the mean reversion MFM and $a = 6\%$ for the HW MFM and local volatility MFM so that these MFMs produce a comparable new Bermudan product prices to the one step covariance MFM.

We draw the following conclusions from the numerical results.

1. The vegas for all the MFMs become negative for maturities longer than 10 years. This is the case since we can see from the pricing formula (5.4) of the new Bermudan product that the value of the floating coupon decreases as the implied volatilities of caplets increase, and therefore decreases the value of the pay fixed new Bermudan product. The bump of the implied volatility in the swaption matrix also leads to the change in the correlation structure and the underlying swaptions, and therefore the value of the new Bermudan product. The combination of these effects leads to the result.

Strike	K=2%	K=3%	K=4%
MR (1%) MFM	-1.19	-1.88	-1.71
MR (5%) MFM	0.95	1.31	1.57
MR (10%) MFM	4.04	5.65	6.10
HW (1%) MFM	-1.89	-2.80	-2.09
HW (5%) MFM	0.20	0.23	0.47
HW (10%) MFM	3.22	4.27	4.80
1-step Cov MFM	0.95	1.30	1.55
Local vol MFM	0.17	0.20	0.38
SV MFM (I)	0	0	0
SV MFM (II)	0.53	0.71	0.85
SV MFM (III)	-0.48	-0.54	-0.65

Table 5.14: The difference between the price for the new Bermudan product produced by different MFMs and the SV MFM (I) measured by total vega.

2. The vega profile produced by the mean reversion MFM, which is an example of parametrizations by expiry, is very different from the vega profile given by MFMs parameterized by time.

3. In order to investigate the impact of smile on the vegas, we compare the vegas produced by the Hull-White MFM with the vegas produced by the local volatility MFM. The vega profiles given by these two models turn out to be very similar. Therefore the smile effect is very small.

4. By introducing stochastic volatility, the row sum of vegas decreases in comparison to the one step covariance MFM but vega profiles of the two models are still very similar. This means that the introduction of stochastic volatility does not materially alter the hedging behaviour.

In summary the numerical results show that the MFM parameterized by expiry gives a big different vega profiles from the MFMs parameterized by time. The smile impact is still very small on vega profiles. The introduction of stochastic volatility decreases the row sum of vegas in comparison to the one step covariance MFM, but the vega profile remains quite similar. This means that the introduction of stochastic volatility does not materially alter the hedging behaviour.

5.6 Conclusion

In this chapter we investigated the impact of correlation, implied volatility smiles and stochastic volatility on Bermudan type products. In particular, we compared a

stochastic volatility MFM to one-dimensional swap MFMs with different combinations of the driver and marginals in terms of pricing and hedging Bermudan type products. We focused on Bermudan swaptions and new Bermudan products. This new Bermudan product is motivated by callable range accruals. The new Bermudan product has similar features but is much simpler than callable range accruals. In order to price these two Bermudan type products accurately, we studied the features and structures which includes "effective strikes" and "internal adjusters" for the new Bermudan product. The expressions for the vegas of a Bermudan swaption in different MFMs are also given.

To perform the numerical comparison we calibrated the stochastic volatility swap MFM and one-dimensional swap MFMs with different combinations of driver and marginals to the market data. Then we compared the prices and vega profiles produced by the above MFMs.

The numerical results indicated that, for a Bermudan swaption, the correlation impact on the price and vega profiles is very large while the effects of implied volatility smiles and stochastic volatility is insignificant. For the new Bermudan product, the stochastic volatility impact on the price and vega profiles becomes larger and significant. The correlation effect is very large but the impact of implied volatility smiles is still insignificant. One important observation is that the vega profiles produced by the MFM parameterized by expiry is very different from the models parameterized by time, but this fundamental difference is not altered when the stochastic volatility is added.

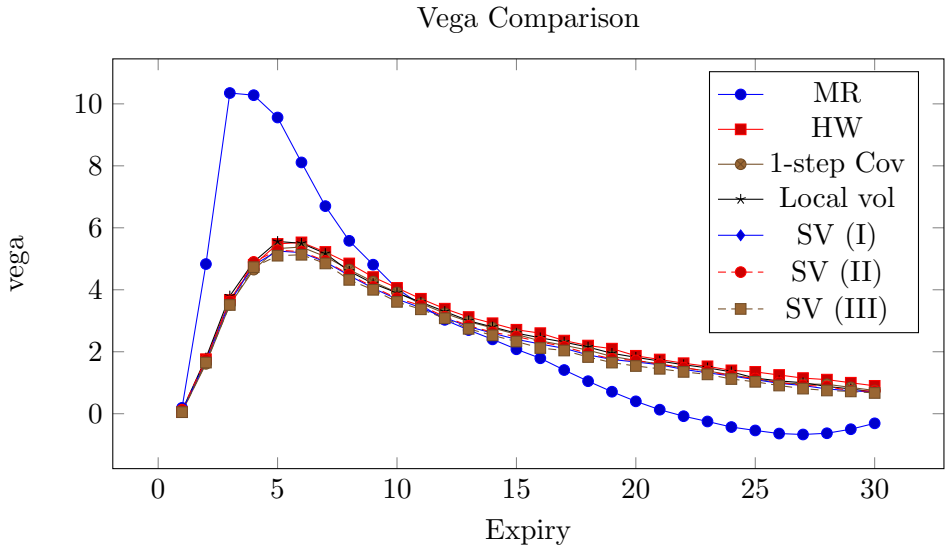


Figure 5.1: The sum of the vegas $\sum_{k=1}^{n+1-i} v_{i,k}$ for $i = 1, ..., 30$ of the Bermudan swaption under the different MFMs.

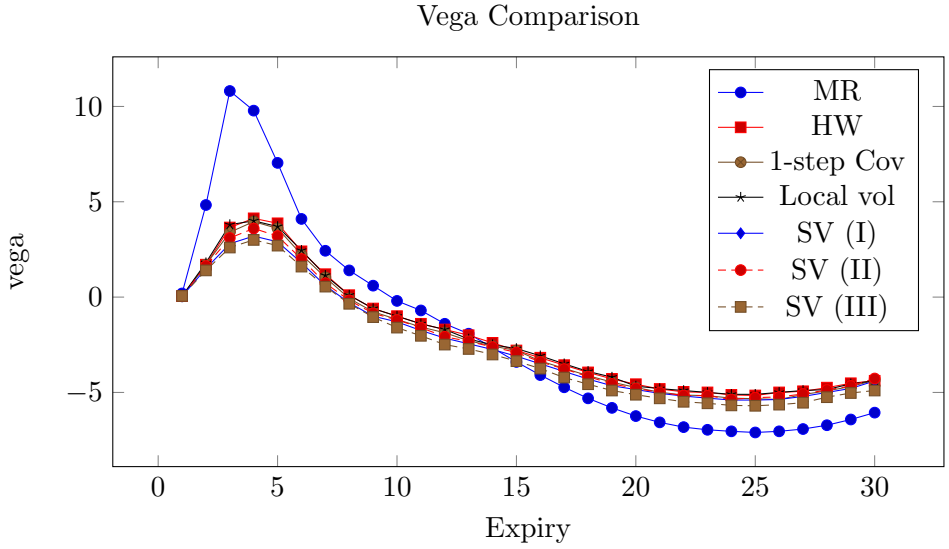


Figure 5.2: The sum of the vegas $\sum_{k=1}^{n+1-i} v_{i,k}$ for $i = 1, ..., 30$ of the new Bermudan product under the different MFMs.

Chapter 6

Quasi-Gaussian models and Markov-functional models

6.1 Introduction

In this chapter we compare Markov-functional models to Quasi-Gaussian models in terms of the model specification and calibration. We try to make clear the link and difference between these two low-dimensional term structure models as well as their stochastic volatility versions.

Quasi-Gaussian models[1] were introduced by Jamshidian [42] and Cheyette [14]. They were further developed by Andersen and Piterbarg [5]. Quasi-Gaussian models belong to the class of short rate models and are obtained by imposing separability condition on Heath-Jarrow-Morton models. In a one-factor Quasi-Gaussian model, the economy is driven by a two-dimensional Markov process. This means that one can implement one-factor Quasi-Gaussian models efficiently since one only needs to keep track of the two-dimensional Markov process. By choosing an appropriate volatility function, one-factor Quasi-Gaussian models allow for analytical approximate pricing formulas for vanilla options and are capable of capturing implied volatility skews of vanilla options. They can also calibrate to the market correlation structure which is very important for pricing and hedging Bermudan type products. Furthermore Quasi-Gaussian models allow for stochastic volatility extensions.

Due to these calibration and implementation advantages, Quasi-Gaussian models have become one of the most popular low-dimensional term structure models. Another class of low-dimensional term structure models that have similar ad-

[1] Also known as Cheyette models

vantages and are widely used in the City are Markov-functional models. In this chapter we compare these two models in terms of model specification and calibration. By considering a (stochastic volatility) Quasi-Gaussian model using the Markov-functional approach, we make clear the link and difference between the two models. By the general definition of a Markov-functional model given in Section 3.2, Quasi-Gaussian models belong to the class of Markov-functional models. However the Quasi-Gaussian model is set up in a different way from the Markov-functional approach we discussed earlier. This difference stems from the way the functional forms of zero-coupon bonds are determined. In particular in a Quasi-Gaussian model the functional forms of zero-coupon bonds are given analytically while the functional forms are determined numerically by feeding in marginals for the Markov-functional approach. This difference also leads to more flexibility for the Markov-functional approach. In particular, given a driving process for a Markov-functional model, one can feed in any marginal distributions of LIBORs (swap rates). However once a driving process is given in a Quasi-Gaussian model, the marginal distributions of LIBORs and swap rates are determined via the explicit formula for zero-coupon bonds. This separation of the driver and marginals gives Markov-functional models more flexibility.

Another flexibility for a Markov-functional model is the freedom to choose its driving process. Quasi-Gaussian models are obtained from the separable Heath-Jarrow-Morton models. As a result the form of the driver is forced upon us as this is required to rule out arbitrage in the separable Heath-Jarrow-Morton framework. In that sense we do not truly choose a driving process and it is the mathematics that drives us to take it as the driving process. In a Markov-functional model, in principle, we are free to choose any diffusion process as the driving process to capture the level of rates. For example one can choose a Gaussian process as the driver for the sake of efficient implementation. Alternatively one can take a SABR process as the driver according to the data driven study [44] where the authors identified a SABR style model as an appropriate choice for the level of interest rates.

Quasi-Gaussian models and Markov-functional models are both capable of capturing correlation structure of swap rates, which is important for pricing and hedging Bermudan type products. In particular, in both models, the parameters of drivers can be chosen by calibrating to the market correlation of swap rates. In that sense they are both examples of "parametrizations by time" according to the explanation in Chapter 5.

The rest of the chapter is organized as follows. In Section 6.2 we review briefly one-factor Quasi-Gaussian models and their stochastic volatility extensions.

The calibration to vanilla options and correlation structure is also discussed. In Section 6.3 we compare Quasi-Gaussian models to Markov-functional models in terms of model specification and calibration. We specify a Quasi-Gaussian model using the Markov-functional approach.

6.2 Quasi-Gaussian model

In this section we review briefly Quasi-Gaussian (QG) models. The material in this section comes from Andersen and Piterbarg [5]. We begin with the specification of a general one-factor Quasi-Gaussian model. We then focus on a special case when the volatility is chosen to be of displaced diffusion form. With such a choice we derive approximate formulas for vanilla swaptions prices and correlations of swap rates. The stochastic volatility version is also discussed. We just provide some main results of Quasi-Gaussian models without proof. For details the reader is referred to [5].

6.2.1 General one-factor Quasi-Gaussian model

A Quasi-Gaussian model can be seen as a separable Heath-Jarrow-Morton (HJM) model. We have introduced HJM models in Section 2.2.3. For ease of reference we present here the SDE of an instantaneous forward rate $f(.,T)$ in a HJM model:

$$df(t,T) = \sigma_f(t,T)(\int_t^T \sigma_f(t,u)du)dt + \sigma_f(t,T)dW_t^Q, \quad 0 \leq t \leq T, \tag{6.1}$$

where $W^{\mathbb{Q}}$ is a Brownian motion in the risk-neutral measure $\mathbb{Q}$ associated with the numeraire the bank account. Note that the volatility function $\sigma_f(t,T)$, $0 \leq t \leq T$, in (6.1) can be stochastic, in which case we use the notation $\sigma_f(t,T,\omega)$. We now give the definition of separability under HJM models.

Definition 3. *The instantaneous volatility function $\sigma_f(t,T,\omega)$, $0 \leq t \leq T$, is separable if there exists a function g such that*

$$\sigma_f(t,T,\omega) = g(t,\omega)h(T) \tag{6.2}$$

for some deterministic function $h(\cdot)$.

A HJM model is said to be separable if separability is imposed on the volatility function. Following some transformations, one can obtain a one-factor Quasi-

Gaussian model from a one-factor separable HJM model which is given by

$$\begin{aligned} dx_t &= (q_t - \varkappa_t x_t)dt + \sigma_r(t,\omega)dW_t^{\mathbb{Q}}, \quad x_0 = 0 \\ dq_t &= (\sigma_r(t,\omega)^2 - 2\varkappa_t q_t)dt, \quad q_0 = 0 \end{aligned} \tag{6.3}$$

where

$$\begin{aligned} \varkappa(t) &= -\frac{h'(t)}{h(t)}, \\ \sigma_r(t,\omega) &= \sigma_f(t,t,\omega) = g(t,\omega)h(t). \end{aligned} \tag{6.4}$$

The instantaneous forward rates can be expressed in terms of x and q:

$$f(t,T) = f(0,T) + \frac{h(T)}{h(t)}(x_t + G(t,T)q_t),$$

where

$$G(t,T) = \frac{\int_t^T h(s)ds}{h(t)}.$$

Furthermore the short rate r is given by

$$r_t = f(t,t) = f(0,t) + x_t.$$

Note that Quasi-Gaussian models can be viewed as short rate models. Recall that in a short rate model, the pricing formula for zero-coupon bonds (ZCBs) is given to us though we may not have an analytic expression for it.

Proposition 5. *In the one-factor Quasi-Gaussian model (6.3). Zero-coupon bonds prices are deterministic functions of the processes x_t and q_t,*

$$D_{t,T} = D_{t,T}(x_t, q_t)$$

where

$$D_{t,T}(x,q) = \frac{D_{0,T}}{D_{0,t}} \exp(-G(t,T)x - \frac{1}{2}G(t,T)^2 q). \tag{6.5}$$

Proof. See Proposition 13.1.1 in [5]. □

The proposition demonstrates that the evolution of the whole interest rate curve in the model can be described by the evolution of a two-dimensional Markov

process (x, q) which is given by (6.3). Observe that in general, the finite variation process q is not deterministic.

Remark 10. *The general one-factor Quasi-Gaussian model (separable HJM model) is a low-dimensional Markov-functional model (MFM) according to the general definition of a MFM given in Chapter 3. However the Quasi-Gaussian model is not set up according to the Markov-functional approach discussed earlier in Section 3.2.3 as the functional forms of ZCBs here are given analytically in (6.5) while the functional forms need to be determined numerically for the Markov-functional approach.*

6.2.2 Displaced diffusion type local volatility

Having defined a general Quasi-Gaussian model (6.3), in this section we focus on a special case when the volatility function $\sigma_r(t, \omega)$ is taken to be of displaced diffusion type:

$$\sigma_r(t, \omega) = \lambda_r(t)(\alpha_r(t) + b_r(t)x_t), \tag{6.6}$$

where $\lambda_r(t)$ and $b_r(t)$ are assumed to be piecewise constant:

$$\lambda_r(t) = \sum_{i=0}^{n-1} \lambda_i I_{\{t \in (T_i, T_{i+1}]\}}, \tag{6.7}$$

$$b_r(t) = \sum_{i=0}^{n-1} b_i I_{\{t \in (T_i, T_{i+1}]\}}. \tag{6.8}$$

The parameter $\alpha_r(t)$ is redundant and can be chosen exogenously. Then we arrive at the following local volatility Quasi-Gaussian model

$$\begin{aligned} dx_t &= (q_t - \varkappa_t x_t)dt + \lambda_r(t)(\alpha_r(t) + b_r(t)x_t)dW_t^{\mathbb{Q}}, \quad x_0 = 0 \\ dq_t &= (\lambda_r^2(t)(\alpha_r(t) + b_r(t)x_t)^2 - 2\varkappa_t q_t)dt, \quad q_0 = 0. \end{aligned} \tag{6.9}$$

Note that we choose displaced diffusion type local volatility for tractability, and also it is capable of producing implied volatility skews of vanilla swaptions which will be seen later. Alternative choices for local volatility are, for example,

$$\sigma_r(t, \omega) = \lambda_r(t)x_t$$

and

$$\sigma_r(t, \omega) = \lambda_r(t)x_t^{\beta_r(t)}$$

which give rise to models with log-Normal dynamics and CEV dynamics respectively.

6.2.3 Calibration to swaptions

In this subsection we consider the problem of how to calibrate the displaced diffusion Quasi-Gaussian model (6.9) to the market prices of vanilla swaptions. To do so we derive an approximate pricing formula for vanilla swaptions.

Swap rate dynamics

Let us first derive the dynamics of the co-terminal swap rate $y^{i,n+1-i}$, $i = 1, ..., n$, within the model (6.9). Since the swap rate $y^{i,n+1-i}$ is a martingale in the associated swaption measure $\mathbb{S}^{i,n+1-i}$ corresponding to the numeraire $P^{i,n+1-i}$, the swap rate $y^{i,n+1-i}$ should be a driftless process. From (6.9) we have that

$$dy_t^{i,n+1-i} = (\frac{\partial y_t^{i,n+1-i}}{\partial x}(x_t, q_t))\lambda_r(t)(\alpha_r(t) + b_r(t)x_t)dW_t^{i,n+1-i} \tag{6.10}$$

where $W^{i,n+1-i}$ is a Brownian motion in the swaption measure $\mathbb{S}^{i,n+1-i}$. By the definition of swap rates, we have that

$$y_t^{i,n+1-i}(x_t, q_t) = \frac{D_{tT_i}(x_t, q_t) - D_{tT_{n+1}}(x_t, q_t)}{P_t^{i,n+1-i}(x_t, q_t)}, \qquad t \leq T_i, \tag{6.11}$$

where $D_{tT}(x_t, q_t)$ is given by (6.5). Apply Itô's lemma to (6.11) and we have that

$$\begin{aligned}\frac{\partial y_t^{i,n+1-i}}{\partial x}(x, q) = &- \frac{1}{P_t^{i,n+1-i}(x, q)}(D_{tT_i}(x, q)G(t, T_i) - D_{tT_{n+1}}(x, q)G(t, T_{n+1})) \\ &+ \frac{y_t^{i,n+1-i}(x, q)}{P_t^{i,n+1-i}(x, q)} \sum_{j=i}^{n} \alpha_j D_{tT_{j+1}}(x, q)G(t, T_{j+1}).\end{aligned} \tag{6.12}$$

Swap rate dynamics approximation

From SDE (6.10) we can see that the co-terminal swap rates have a local volatility which is a function of (x, q). This dynamics prevents us from deriving a pricing formula for co-terminal swaptions. Therefore we approximate (6.10) by replacing the diffusion term which is a function of (x, q) with a local volatility which is a function of the swap rate itself. The approximations applied here are quite involved so we will not provide complete technical details. For details, the reader is referred to [5].

We first apply the Markovian projection (see Appendix 6.A) to the dynamics of the co-terminal swap rates (6.10), and we have the following lemma.

Lemma 5. *The marginal distributions of the swap rate $y^{i,n+1-i}$ in the model (6.10) are the same as the marginal distributions of $y^{i,n+1-i}$ implied from the following dynamics in the associated swaption measure $\mathbb{S}^{i,n+1-i}$:*

$$dy_t^{i,n+1-i} = \varphi^i(t, y_t^{i,n+1-i})dW_t^{i,n+1-i}, \tag{6.13}$$

where

$$\varphi^i(t,y)^2 = \mathbb{E}_{\mathbb{S}^{i,n+1-i}}[((\frac{\partial y_t^{i,n+1-i}}{\partial x}(x_t, q_t))\lambda_r(t)(\alpha_r(t) + b_r(t)x_t))^2 | y_t^{i,n+1-i} = y]. \tag{6.14}$$

Remark 11. *Note that (6.10) and (6.13) admit the same marginal distributions of the swap rates $y^{i,n+1-i}$ for $i = 1, ..., n$ rather than the joint distributions. The same marginals implies that the co-terminal vanilla swaption prices produced by models (6.10) and (6.13) are identical.*

So far we have obtained the dynamics (6.13) of swap rates where the local volatility is a function of the swap rate itself. However it is still difficult to calculate the conditional expectation (6.14). To solve this problem we approximate q_t by a deterministic function $\bar{q}_t$. By doing so the swap rate $y_t^{i,n+1-i}$ would just be a deterministic function of x_t and time t. Similarly x_t would also be a deterministic function of the co-terminal swap rate $y_t^{i,n+1-i}$ and time t, i.e. $x_t = x(t, y_t^{i,n+1-i})$. Now the conditional expectation (6.14) is given by

$$\varphi^i(t,y) = \frac{\partial y_t^{i,n+1-i}}{\partial x}(x(t,y), \bar{q}_t)\lambda_r(t)(\alpha_r(t) + b_r(t)x(t,y)),$$

where $\partial y^{i,n+1-i}/\partial x$ is given by (6.12).

We now give a simple approximation for q_t and functions φ^i for $i = 1, ..., n$. We let $\bar{q}_t = 0$, and apply Taylor expansion to $y_t^{i,n+1-i}(x_t, \bar{q}_t)$:

$$y_t^{i,n+1-i}(x,0) \approx y_t^{i,n+1-i}(0,0) + \frac{\partial y_t^{i,n+1-i}}{\partial x}(0,0)x.$$

This gives the following approximation for $x(t, y_t^{i,n+1-i})$

$$x(t,y) \approx \frac{y - y_t^{i,n+1-i}(0,0)}{\partial y_t^{i,n+1-i}/\partial x(0,0)}.$$

Then we make the following rough approximation

$$\frac{\partial y_t^{i,n+1-i}}{\partial x}(x,0) \approx \frac{\partial y_t^{i,n+1-i}}{\partial x}(0,0),$$

and we obtain the approximation for functions φ^i, $i = 1, ..., n$,

$$\varphi^i(t,y) \approx \frac{\partial y_t^{i,n+1-i}}{\partial x}(0,0)\lambda_r(t)(\alpha_r(t) + b_r(t)\frac{y - y_t^{i,n+1-i}(0,0)}{\partial y_t^{i,n+1-i}/\partial x(0,0)}).$$

Finally we apply Taylor expansion to the function φ^i, and we have that

$$\varphi^i(t, y_t^{i,n+1-i}) \approx \varphi^i(t, y_0^{i,n+1-i}) + \frac{\partial \varphi^i}{\partial y}(t, y_0^{i,n+1-i})(y_t^{i,n+1-i} - y_0^{i,n+1-i}), \tag{6.15}$$

where

$$\varphi^i(t, y_0^{i,n+1-i}) = \frac{\partial y_t^{i,n+1-i}}{\partial x}(0,0)\lambda_r(t)(\alpha_r(t) + b_r(t)\frac{y_0^{i,n+1-i} - y_t^{i,n+1-i}(0,0)}{\partial y_t^{i,n+1-i}/\partial x(0,0)}),$$

$$\frac{\partial \varphi^i}{\partial y}(t, y_0^{i,n+1-i}) = \frac{\partial y_t^{i,n+1-i}}{\partial x}(0,0)(\frac{\lambda_r(t)b_r(t)}{\partial y_t^{i,n+1-i}/\partial x(0,0)}).$$

Note that the approximation introduced here is rather rough. A more accurate approximation can be found in [5].

Proposition 6. *In the one-factor displaced diffusion Quasi-Gaussian model (6.9), the dynamics of the swap rate $y^{i,n+1-i}$ in the associated swaption measure $\mathbb{S}^{i,n+1-i}$ can be approximated by*

$$dy_t^{i,n+1-i} \approx \lambda_y^i(t)(b_y^i(t)y_t^{i,n+1-i} + (1 - b_y^i(t))y_0^{i,n+1-i})dW_t^{i,n+1-i}, \tag{6.16}$$

where

$$\lambda_y^i(t) := \frac{\varphi^i(t, y_0^{i,n+1-i})}{y_0^{i,n+1-i}}, \tag{6.17}$$

$$b_y^i(t) := \frac{y_0^{i,n+1-i}}{\varphi^i(t, y_0^{i,n+1-i})}\frac{\partial \varphi^i}{\partial y}(t, y_0^{i,n+1-i}). \tag{6.18}$$

Proof. The result follows from (6.13) and (6.15). □

Note that (6.16) is a displaced diffusion model with time-dependent parameters $\lambda_y^i(t)$ and $b_y^i(t)$. We approximate (6.16) by replacing time-dependent parameters $\lambda_y^i(t)$ and $b_y^i(t)$ with parameters $\overline{\lambda}_y^i$ and $\overline{b}_y^i$ using the parameter averaging technique

which is introduced by Andersen and Piterbarg [5]. Therefore we arrive at the following approximate dynamics:

$$dy_t^{i,n+1-i} \approx \overline{\lambda}_y^i[\overline{b}_y^i y_t^{i,n+1-i} + (1-\overline{b}_y^i)y_0^{i,n+1-i}]dW_t^{i,n+1-i}, \tag{6.19}$$

where

$$\overline{\lambda}_y^i := (\frac{1}{T_i}\int_0^{T_i} \lambda_y^i(t)^2 dt)^{1/2}, \tag{6.20}$$

$$\overline{b}_y^i := \int_0^{T_i} b_y^i(t) w_y^i(t) dt, \tag{6.21}$$

$$w_y^i(t) := \frac{\lambda_y^i(t)^2 \int_0^t \lambda_y^i(s)^2 ds}{\int_0^{T_i}(\lambda_y^i(u)^2 \int_0^u \lambda_y^i(s)^2 ds)du}.$$

Note that by using the parameter averaging technique, the resulting model (6.19) has very similar marginal distributions for the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$ to the model (6.16). This gives the following approximate swaption pricing formula.

Proposition 7. *In the one-factor displaced diffusion Quasi-Gaussian model (6.9), an approximate pricing formula for a payer swaption with strike K and expiry T_i on the swap rate $y^{i,n+1-i}$ is given by*

$$\begin{aligned} V^{i,n+1-i}(0;K) \approx P_0^{i,n+1-i}[&(y_0^{i,n+1-i} + y_0^{i,n+1-i}(1-\overline{b}_y^i)/\overline{b}_y^i)\Phi(d_+) \\ &- (K + y_0^{i,n+1-i}(1-\overline{b}_y^i)/\overline{b}_y^i)\Phi(d_-)], \end{aligned} \tag{6.22}$$

$$d_\pm = \frac{\ln(\frac{y_0^{i,n+1-i}+y_0^{i,n+1-i}(1-\overline{b}_y^i)/\overline{b}_y^i}{K+y_0^{i,n+1-i}(1-\overline{b}_y^i)/\overline{b}_y^i}) \pm \frac{1}{2}(\overline{b}_y^i\overline{\lambda}_y^i)^2 T_i}{\overline{b}_y^i\overline{\lambda}_y^i\sqrt{T_i}},$$

where $\overline{\lambda}_y^i$ and $\overline{b}_y^i$ are given by (6.20) and (6.21).

From the pricing formula (6.22), one can see that the model (6.9) is capable of capturing implied volatility skews of vanilla swaptions. The parameters λ_i's and b_i's of (6.7) and (6.8) can be chosen by calibrating to the market implied volatility smiles/skews of swaptions via the pricing formula (6.22).

6.2.4 Calibration to the market correlation structure

We now consider the problem of how to determine the parameters $\varkappa(t)$ by calibrating to the market correlation of swap rates. Let us consider the following formula for the correlation between co-terminal swap rates (see [5]).

Proposition 8. *The correlation between the swap rates $y_{T_i}^{i,n+1-i}$ and $y_{T_j}^{j,n+1-j}$, $i<j$, in the Quasi-Gaussian model (6.9) can be approximated by*

$$Corr(y_{T_i}^{i,n+1-i}, y_{T_j}^{j,n+1-j}) \approx \int_0^{T_i} (\frac{\partial y_t^{i,n+1-i}}{\partial x}(0,0))(\frac{\partial y_t^{j,n+1-j}}{\partial x}(0,0))dt$$
$$\times (\int_0^{T_i} (\frac{\partial y_t^{i,n+1-i}}{\partial x}(0,0))^2 dt)^{-1/2} (\int_0^{T_j} (\frac{\partial y_t^{j,n+1-j}}{\partial x}(0,0))^2 dt)^{-1/2}. \tag{6.23}$$

From the above formula, we can capture the correlation structure of swap rates. However note that in a low-factor model we cannot capture the whole correlation matrix. By letting

$$\varkappa(t) = \sum_{i=0}^{n-1} \varkappa_i I_{\{t\in(T_i,T_{i+1}]\}}$$

one can choose parameters $\varkappa_i$ by calibrating to the market one step correlation of swap rates:

$$Corr(y_{T_i}^{i,n+1-i}, y_{T_{i+1}}^{i+1,n-i})$$

for $i = 1, ..., n-1$.

We conclude this subsection with a remark. Note that the formula (6.23) is independent of the volatility parameters λ_i's and b_i's. This means that we can carry out the correlation calibration before the swaption calibration discussed in Section 6.2.3. This feature is very similar to the Markov-functional approach where the market one step correlation (covariance) calibration precede the calibration to vanilla options.

6.2.5 Stochastic volatility Quasi-Gaussian model

Before we proceed to the comparison between Quasi-Gaussian models and MFMs, we review the stochastic volatility version of Quasi-Gaussian models.

The stochastic volatility version of a Quasi-Gaussian model is straightforward to extend. Consider the volatility function $g(t,\omega)$ (6.4) of the one-factor Quasi-Gaussian model (6.3). The stochastic volatility version can be obtained by adding a stochastic volatility process z into the volatility function $g(t,\omega)$. In particular we have

$$g(t,\omega) = \sqrt{z(t)}g(t,x_t,q_t),$$

and

$$\sigma_r(t,x,q) = g(t,x,q)h(t).$$

The stochastic volatility Quasi-Gaussian model is defined by the following SDEs:

$$\begin{aligned} dx_t &= (q_t - \varkappa_t x_t)dt + \sqrt{z_t}\sigma_r(t,x_t,q_t)dW_t^{\mathbb{Q}}, \quad x_0 = 0, && (6.24)\\ dq_t &= (z_t\sigma_r(t,x_t,q_t)^2 - 2\varkappa_t q_t)dt, \quad q_0 = 0, \\ dz_t &= \theta(z_0 - z_t)dt + \eta_t\sqrt{z_t}dZ_t^{\mathbb{Q}}, \quad z_0 = 1, && (6.25)\\ dZ_t^{\mathbb{Q}}dW_t^{\mathbb{Q}} &= 0, \end{aligned}$$

where $W^{\mathbb{Q}}$ and $Z^{\mathbb{Q}}$ are independent Brownian motions in the risk neutral measure and the function η_t is piecewise constant of the form

$$\eta_t = \sum_{i=0}^{n-1} \eta_i I_{\{t\in(T_i,T_{i+1}]\}}.$$

We will assume that in addition to ZCBs there is at least one option in the economy which means that the model (6.24) has been chosen and we are in a complete setting. In this case when moving to another equivalent martingale measure associated with a different numeraire the Radon-Nikodým derivative is the ratio of numeraires.

Remark 12. *Note that we make the assumption of zero correlation $dZ_t^{\mathbb{Q}}dW_t^{\mathbb{Q}} = 0$ for tractability. It is helpful when we change measure, as we will see later. This assumption is not a restriction for capturing implied volatility smiles since the correlation is not the only parameter that can control the slope of the smile. However it does cause a restriction of the dynamics and it could have an effect on the use of the model for a Bermudan type product, especially, in terms of hedging. In contrast, in the stochastic volatility MFM developed in Chapter 3, we do not have this restriction.*

The pricing formulas for ZCBs in the model (6.24) are given in Proposition 5 which are the same as the local volatility version. Thus the formulas for ZCBs are not dependent on the stochastic volatility z and this leads to an unspanned stochastic volatility model in which case the stochastic volatility risk cannot be completely hedged by ZCBs. For more details about (un)spanned stochastic volatility models, the reader is referred to [15].

In what follows we focus on a special case when the volatility function $\sigma_r(t,x_t,q_t)$ in (6.24) is of displaced diffusion type which is given by (6.6). This

leads to the following stochastic volatility Quasi-Gaussian model:

$$\begin{aligned} dx_t &= (q_t - \varkappa_t x_t)dt + \sqrt{z_t}\lambda_r(t)(\alpha_r(t) + b_r(t)x_t)dW_t^{\mathbb{Q}}, \quad x_0 = 0 \\ dq_t &= (z_t\lambda_r^2(t)(\alpha_r(t) + b_r(t)x_t)^2 - 2\varkappa_t q_t)dt, \quad q_0 = 0. \\ dz_t &= \theta(z_0 - z_t)dt + \eta_t\sqrt{z_t}dZ_t^{\mathbb{Q}}, \quad z_0 = 1, \\ dZ_t^{\mathbb{Q}}dW_t^{\mathbb{Q}} &= 0. \end{aligned} \tag{6.26}$$

We now review the calibration in the stochastic volatility Quasi-Gaussian model (6.26). Most approximations and techniques we used for the local volatility version in Section 6.2.3 can also apply to the stochastic volatility version here. So we will not provide complete details which can be found in [5].

Proposition 9. *In the stochastic volatility Quasi-Gaussian model (6.26), the dynamics of the swap rate $y^{i,n+1-i}$ in the associated swaption measure $\mathbb{S}^{i,n+1-i}$ can be approximated by*

$$\begin{aligned} dy_t^{i,n+1-i} &= \sqrt{z_t}\lambda_y^i(t)(b_y^i(t)y_t^{i,n+1-i} + (1 - b_y^i(t))y_0^{i,n+1-i})dW_t^{i,n+1-i}, \\ dz_t &= \theta(z_0 - z_t)dt + \eta_t\sqrt{z_t}dZ_t^{i,n+1-i}, \quad z_0 = 1, \\ dZ_t^{i,n+1-i}dW_t^{i,n+1-i} &= 0, \end{aligned} \tag{6.27}$$

where $\lambda_y^i(t)$ and $b_y^i(t)$ are given by (6.17) and (6.18).

Note that due to the assumption of zero correlation $dZ_t^{\mathbb{Q}}dW_t^{\mathbb{Q}} = 0$ and completeness, the drift term of the stochastic volatility process z is not altered when the measure is changed from the risk-neutral measure $\mathbb{Q}$ to the swaption measure $\mathbb{S}^{i,n+1-i}$.

Applying the parameter averaging technique, the time-dependent stochastic volatility model (6.27) can be approximated by a time-independent stochastic volatility model which allows for a closed-form pricing formula for vanilla swaptions. The parameters λ_i's, b_i's and η_i's in the stochastic volatility Quasi-Gaussian model (6.26) can be chosen by calibrating to the market implied volatility smiles/skews of vanilla swaptions. The remaining parameters $\varkappa_i$'s can be determined by calibrating to the market correlation structure. We can still use formulas introduced for the local volatility Quasi-Gaussian model in spite of effects of the stochastic volatility. But the calibration performance is not expected to be very precise.

6.3 Comparison between Quasi-Gaussian models and Markov-functional models

6.3.1 Review of Quasi-Gaussian model under Markov-functional model framework

We now consider the Quasi-Gaussian model (6.9) using the Markov-functional approach to make clear the relationship between the Quasi-Gaussian models and MFMs. In particular we specify a swap MFM under the terminal measure $\mathbb{F}^{n+1}$ corresponding to the numeraire $D_{\cdot,T_{n+1}}$. By choosing a particular combination of driving process, pre-model and marginals, the resulting swap MFM is expected to be close to the Quasi-Gaussian model (6.9).

Driving process

In order to specify a MFM, we first consider the choice of driving process. In Section 3.2, we specified a MFM by taking a Gaussian process as the driver for the sake of efficient implementation. Note that we are not forced to choose a Gaussian driving process. In order to match the Quasi-Gaussian model (6.9), we take the process (x, q) of the model (6.9) as the driver. However we will specify the MFM under the terminal measure while the process (x, q) is given under the risk-neutral measure. Thus following the change of numeraire technique we write down SDEs of (x, q) under the terminal measure.

Proposition 10. *Under the terminal measure $\mathbb{F}^{n+1}$ the process (x, q) in the Quasi-Gaussian model (6.9) satisfies*

$$\begin{aligned} dx_t &= (q_t - \varkappa_t x_t - G(t, T_{n+1})\lambda_r^2(t)(\alpha_r(t) + b_r(t)x_t)^2)dt \\ &\quad + \lambda_r(t)(\alpha_r(t) + b_r(t)x_t)dW_t^{n+1} \\ dq_t &= (\lambda_r^2(t)(\alpha_r(t) + b_r(t)x_t)^2 - 2\varkappa(t)q_t)dt, \end{aligned} \tag{6.28}$$

where the Brownian motion W^{n+1} under the terminal measure is defined by

$$dW_t^{n+1} := dW_t^{\mathbb{Q}} + G(t, T_{n+1})\lambda_r(t)(\alpha_r(t) + b_r(t)x_t)dt.$$

Proof. Define ρ_t by

$$\rho_t := \left.\frac{d\mathbb{F}^{n+1}}{d\mathbb{Q}}\right|_{\mathcal{F}_t} = \frac{D_{tT_{n+1}}/D_{0T_{n+1}}}{B_t/B_0},$$

where B is bank account and $\{\mathcal{F}_t\}$ is the filtration generated by $W^{\mathbb{Q}}$. After calcu-

lation, we have that

$$d\rho_t = -\rho_t G(t, T_{n+1})\lambda_r(t)(\alpha_r(t) + b_r(t)x_t)dW_t^{\mathbb{Q}}.$$

From the Girsanov theorem we have that the process

$$\begin{aligned} dW_t^{n+1} &:= dW_t^{\mathbb{Q}} - \frac{d\rho_t}{\rho_t} \cdot dW_t^{\mathbb{Q}} \\ &= dW_t^{\mathbb{Q}} + G(t, T_{n+1})\lambda_r(t)(\alpha_r(t) + b_r(t)x_t)dt. \end{aligned}$$

is a Brownian motion in the terminal measure $\mathbb{F}^{n+1}$ corresponding to the numeraire $D_{.,T_{n+1}}$. The result then follows from SDE (6.9). □

Usually when using the Markov-functional approach for a low-dimensional MFM, the driving process is chosen to model the level of rates. For instance for the stochastic volatility MFM we developed in Chapter 3, we choose a SABR type model as the driver because Kaisajuntti and Kennedy [44] identified a SABR style model as an appropriate choice for the level of interest rates by investigating an extensive set of market data. The process x of (6.28) can be seen as modelling the level of interest rates while carrying the process q with it.

We note that the process x itself is not Markovian but (x, q) is a two-dimensional Markov process. Thus we have to set up a two-dimensional MFM although there is only one factor (Brownian motion) driving the whole economy. From the modelling perspective, this does not look a likely choice for a driver. In particular, in the Quasi-Gaussian model the form of the driver (6.3) and the necessity to carry around the finite variation process q is forced upon us as this is required to avoid arbitrage in the separable HJM framework. In that sense we do not truly choose a driver and we are just stuck with it. Though the driver (6.3) leads to a non-Gaussian copula, there are other choices of non-Gaussian copula, e.g. CEV driver, which would not require a two-dimensional process for a one-factor model and which maybe better from a modelling point of view. Thus the Markov-functional approach allows for more flexibility in the choice of driver.

The Markov-functional approach is closer in spirit to market models. In particular the Quasi-Gaussian model is obtained from the separable HJM model while the MFM with particular choices of driver and marginals is found to be very similar to the corresponding separable market model. The HJM model models the instantaneous forward rate which is not observable in the market while market models model the market interest rate directly which is more transparent.

Pre-model

Remember that under multi-dimensional MFMs we retain the univariate and monotonicity properties by introducing the idea of pre-model. Here we apply the same technique and choose a pre-model. Recall that a pre-model $\widehat{y}_{T_i}^{i,n+1-i} : \mathbb{R}^2 \to \mathbb{R}$ is a function of (x_{T_i}, q_{T_i}):

$$\widehat{y}_{T_i}^{i,n+1-i}(x_{T_i}, q_{T_i}) = f^i(x_{T_i}, q_{T_i}).$$

In order to match the Quasi-Gaussian model (6.9), we choose the pre-model to be the function $y_{T_i}^{i,n+1-i}(x_{T_i}, q_{T_i})$ which is given in the Quasi-Gaussian model.

Since ZCBs are deterministic functions (6.5) of the process (x, q) in the Quasi-Gaussian model (6.9), the co-terminal swap rates can be expressed in the form of

$$y_t^{i,n+1-i}(x_t, q_t) = \frac{D_{tT_i}(x_t, q_t) - D_{tT_{n+1}}(x_t, q_t)}{\sum_{j=i}^{n} \alpha_j D_{tT_{j+1}}(x_t, q_t)}, \tag{6.29}$$

where the functional form $D_{tT}(x_t, q_t)$ is given by (6.5). We choose the pre-model $\widehat{y}_{T_i}^{i,n+1-i}(x_{T_i}, q_{T_i})$ as

$$\hat{y}_{T_i}^{i,n+1-i}(x_{T_i}, q_{T_i}) := y_{T_i}^{i,n+1-i}(x_{T_i}, q_{T_i}). \tag{6.30}$$

Marginal distribution

Having chosen the driving process and the pre-model for our MFM, we now consider the problem of how to choose the marginal distribution of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$. In order to specify a MFM which is similar to the Quasi-Gaussian model (6.9), we choose the marginals implied from the Quasi-Gaussian model. As we discussed in the Quasi-Gaussian model (6.9), the marginal distributions of $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$ are approximately given by the displaced diffusion model (6.19). Therefore we feed the displaced diffusion marginal distribution implied from (6.19) into our MFM.

Remark 13. *Given a driving process for a MFM, one can feed in any exact marginal distributions of swap rates. This separation of the driver and marginals gives a MFM more flexibility. In contrast, in Quasi-Gaussian models, in order to fit the marginal distributions of swap rates we applied complicated approximations (see Section 6.2.3). Once the driving process of a Quasi-Gaussian model is given, the marginal distributions of swap rates are determined via the explicit formula for ZCBs (6.5).*

In this subsection we have chosen a particular combination of the driver,

pre-model and marginal distributions and we arrive at a MFM which is expected to be similar to the Quasi-Gaussian model (6.9). However we have flexibility to choose another combination and this will lead to a MFM different from the Quasi-Gaussian model (6.9). In this sense the Quasi-Gaussian model can be viewed as a special case of the MFM class.

In similar way, the stochastic volatility Quasi-Gaussian model (6.26) can also be viewed from the Markov-functional approach perspective by choosing an appropriate driver ((x, q) process (6.28) together with stochastic volatility process z (6.25)), pre-model (6.30) and the marginals implied from the stochastic volatility model (6.27).

Remark 14. *In the stochastic volatility Quasi-Gaussian model, one has to keep track of three processes. In the stochastic volatility MFM introduced in Chapter 3 however we only need to deal with a two-dimensional Markov process.*

6.3.2 Calibration comparison

In this subsection we compare the Quasi-Gaussian model (6.9) to a swap MFM in terms of the calibration. In particular we consider the one-dimensional swap MFM with a Gaussian driving process together with a displaced diffusion marginal distributions of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$.

The Quasi-Gaussian model can be calibrated in a similar way to the swap MFM. They can be both calibrated to the displaced diffusion type implied volatility skews of the co-terminal swaptions and the market correlations of swap rates. Recall that in MFMs, we discussed the one-step covariance type driving process which can be specified by calibrating to the market one-step covariance (correlation) of swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$, and this is an example of parametrizations by time. Therefore the Quasi-Gaussian model is also parameterized by time.

When the Quasi-Gaussian model and the swap MFM are calibrated to the same market prices of swaptions and market correlation structure, we are interested in the problem of whether or not the two models produce the same prices and hedges of, for example, a Bermudan swaption. We expect that the results could be different because the two models have different copulas for the driving processes. In particular the MFM has a Gaussian copula while the Quasi-Gaussian model has a non-Gaussian copula. Although they have the same marginal distributions of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$, different copulas for the drivers result in different joint distributions of the swap rates $\{y_{T_i}^{i,n+1-i}; i = 1, ..., n\}$ which determines the prices of the corresponding Bermudan swaption. Therefore the prices and hedges

for a Bermudan swaption produced by the two models could be different. But we need to investigate if this leads to a significant difference in practice. Similar questions arise for the stochastic volatility versions of these models.

In this chapter we have studied the Quasi-Gaussian models via the Markov-functional model framework. This enables us to gain insight into how the models are similar and how they are different. The Quasi-Gaussian model without stochastic volatility can be calibrated to the one-step correlation (covariance) of swap rates and its marginal distributions are approximately that of a displaced diffusion. The main difference from the standard Markov-functional model (with displaced diffusion marginal) is the driver which is non-Gaussian and requires us to track a two-dimensional process. Similarly the stochastic volatility Quasi-Gaussian model requires a three-dimensional driver whereas the stochastic volatility Markov-functional model developed in this thesis requires a two-dimensional driver and the choice of driver is motivated by empirical data rather than tractability considerations. Further numerical work is required to study the differences between these two models in terms of pricing and hedging path dependent derivatives.

6.A Appendix: Markovian projection

We give the following Markovian projection from [5] and the proof can also be found there. Let us consider the following SDE

$$dX_t = \lambda_t dW_t, \tag{6.31}$$

where W is a one-dimensional Brownian motion under some probability measure $\mathbb{P}$ and the process λ is adapted and bounded such that (6.31) has a unique strong solution. Define $b(t,x)$ by

$$b(t,x)^2 = \mathbb{E}_{\mathbb{P}}(\lambda_t^2 | X_t = x).$$

Then the SDE

$$dY_t = b(t,Y_t)dW_t, \qquad Y_0 = X_0,$$

admits a weak solution Y that has the same marginal distributions as X.

Bibliography

[1] Alexander, C. (2001). Principal component analysis of volatility smiles and skews. In EFMA 2001 Lugano Meetings (pp. 2000-10).

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

[3] Andersen, L., & Andreasen, J. (2001). Factor dependence of bermudan swaptions: fact or fiction?. Journal of Financial Economics, 62(1), 3-37.

[4] Andersen, L. B., & Brotherton-Ratcliffe, R. (2001). Extended LIBOR market models with stochastic volatility. Available at SSRN 294853.

[5] Andersen, L., and Piterbarg, V. (2010). Interest Rate Modeling, Atlantic Financial Press.

[6] Beveridge, C., & Joshi, M. S. (2009). Interpolation schemes in the displaced-diffusion libor market model and the efficient pricing and Greeks for callable range accruals. Available at SSRN 1461285.

[7] Bennett, M. N., & Kennedy, J. E. (2005). A comparison of Markov-functional and market models: The one-dimensional case. The Journal of Derivatives, 13(2), 22-43.

[8] Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. The journal of political economy, 637-654.

[9] Black, F. (1976). The pricing of commodity contracts. Journal of financial economics, 3(1), 167-179.

[10] Brace, A., & Musiela, M. (1997). The market model of interest rate dynamics. Mathematical finance, 7(2), 127-155.

[11] Brigo, D., & Mercurio, F. (2007). Interest rate models-theory and practice: with smile, inflation and credit. Springer.

[12] Casassus, J., Collin-Dufresne, P., & Goldstein, B. (2005). Unspanned stochastic volatility and fixed income derivatives pricing. Journal of Banking & Finance, 29(11), 2723-2749.

[13] Caspers, P. (2013). Markov Functional One Factor Interest Rate Model Implementation in QuantLib. Available at SSRN 2183721.

[14] Cheyette, O. (1992). Markov representation of the Heath-Jarrow-Morton model. Available at SSRN 6073.

[15] CollinDufresne, P., & Goldstein, R. S. (2002). Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility. The Journal of Finance, 57(4), 1685-1730.

[16] Cont, R., & Da Fonseca, J. (2002). Dynamics of implied volatility surfaces. Quantitative finance, 2(1), 45-60.

[17] Cox, J. C., Ingersoll Jr, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. Econometrica: Journal of the Econometric Society, 385-407.

[18] Duffie, D., & Kan, R. (1996). A yieldfactor model of interest rates. Mathematical finance, 6(4), 379-406.

[19] Dupire, B. (1994). Pricing with a smile. Risk, 7(1), 18-20.

[20] Dothan, L. U. (1978). On the term structure of interest rates. Journal of Financial Economics, 6(1), 59-69.

[21] Driessen, J., Klaassen, P., & Melenberg, B. (2003). The performance of multi-factor term structure models for pricing and hedging caps and swaptions. Journal of Financial and Quantitative Analysis, 38(03), 635-672.

[22] Fan, R., Gupta, A., & Ritchken, P. (2003). Hedging in the possible presence of unspanned stochastic volatility: Evidence from swaption markets. The Journal of Finance, 58(5), 2219-2248.

[23] Ferreiro, A. M., Garca-Rodrguez, J. A., Lpez-Salas, J. G., & Vzquez, C. (2014). SABR/LIBOR market models: Pricing and calibration for some interest rate derivatives. Applied Mathematics and Computation, 242, 65-89.

[24] Fong, H. G., & Vasicek, O. A. (1991). Fixed-income volatility management. The Journal of Portfolio Management, 17(4), 41-46.

[25] Fouque, J. P., Papanicolaou, G., & Sircar, K. R. (2000). Mean-reverting stochastic volatility. International Journal of theoretical and applied finance, 3(01), 101-142.

[26] Fries, C. (2007). Mathematical finance: theory, modeling, implementation. John Wiley & Sons.

[27] Fries, C., & Eckstaedt, F. (2011). A hybrid Markov-functional model with simultaneous calibration to the interest rate and FX smile. Quantitative Finance, 11(4), 587-597.

[28] Fries, C. P., & Rott, M. G. (2004). Cross-currency and hybrid Markov-functional models. Available at SSRN 532122.

[29] Gogala, J., & Kennedy, J. (2015). One-Dimensional Markov-functional Models driven by Non-Gaussian Markov Processes. Statistics Department, University of Warwick.

[30] Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. Econometrica: Journal of the Econometric Society, 77-105.

[31] Hagan, P. S., Kumar, D., Lesniewski, A. S., & Woodward, D.E. (2002). Managing smile risk. The Best of Wilmott, 249.

[32] Hagan, P. S. (2004). Accrual swaps and range notes. Working paper.

[33] Hagan, P., & Lesniewski, A. (2008). LIBOR market model with SABR style stochastic volatility. JP Morgan Chase and Ellington Management Group, 32.

[34] Henry-Labordere, P. (2007). Combining the SABR and LMM models. Risk, 20, 102-107.

[35] Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. Review of financial studies, 6(2), 327-343.

[36] Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. Review of financial studies, 3(4), 573-592.

[37] Hunt, P. J., Kennedy, J. E. & Pelsser, A. A. J. [2000], 'Markov-functional interest rate models', Finance and Stochastic 4(4), 391-408.

[38] Hunt, P., & Kennedy, J. (2004). Financial derivatives in theory and practice. Wiley.

[39] Hunter, C., Jackel, P. & Joshi, M. [2001], 'Getting the drift', Risk Magazine (July).

[40] Islah, O. (2009). Solving SABR in exact form and unifying it with LIBOR market model. Exact Form and Unifying it with LIBOR Market Model (October 15, 2009).

[41] Jackel, P., & Rebonato, R. (2000). Linking caplet and swaption volatilities in a BGM/J framework: Approximate solutions. Quantitiative Research Centre, The Royal Bank of Scotland.

[42] Jamshidian, F. (1991). Bond and option evaluation in the Gaussian interest rate model. Research in Finance, 9, 131-170.

[43] Johnson, Simon. "Numerical Methods for the Markov Functional model." Wilmott Magazine, Jan (2006): p68.

[44] Kaisajuntti, L., & Kennedy, J. (2014). Stochastic volatility for interest rate derivatives, Quantitative Finance, Volume 14, Issue 3 (pp 457-480)

[45] Kaisajuntti, L., & Kennedy, J. (2013). An n-dimensional Markov-functional Interest Rate Model, Journal of Computational Finance, Volume 17, Issue 1 (pp 3-41).

[46] Karatzas, I., & Shreve, S. E. (1991). Brownian motion and stochastic calculus. Graduate Texts in Mathematics, 113.

[47] Kennedy, J. E., Mitra, S., & Pham, D. (2012). On the approximation of the SABR model: a probabilistic approach. Applied Mathematical Finance, 19(6), 553-586.

[48] Kennedy, J. E., & Pham, D. (2013). Implications for hedging of the choice of driving process for one-factor Markov-functional models. International Journal of Theoretical and Applied Finance, 16(5).

[49] Longstaff, F. A., & Schwartz, E. S. (1992). Interest rate volatility and the term structure: A twofactor general equilibrium model. The Journal of Finance, 47(4), 1259-1282.

[50] Longstaff, F. A., Santa-Clara, P., & Schwartz, E. S. (2001). Throwing away a billion dollars: The cost of suboptimal exercise strategies in the swaptions market. Journal of Financial Economics, 62(1), 39-66.

[51] Marris, D. (1999). Financial option pricing and skewed volatility. Unpublished master's thesis, University of Cambridge.

[52] McInerney, D., & Zastawniak, T. (2015). Stochastic Interest Rates. Cambridge University Press.

[53] Merton, R. C. (1973). Theory of rational option pricing. The Bell Journal of economics and management science, 141-183.

[54] Morini, M., & Mercurio, F. (2007). No-arbitrage dynamics for a tractable SABR term structure LIBOR model. Bloomberg Portfolio Research Paper, (2010-03).

[55] Piterbarg, V. (2003). A stochastic volatility forward Libor model with a term structure of volatility smiles. SSRN eLibrary.

[56] Pietersz, R., Pelsser, A., & Van Regenmortel, M. (2004). Fast drift approximated pricing in the BGM model. Journal of Computational Finance, 8(1).

[57] Pietersz, R., & Pelsser, A. (2010). A comparison of single factor Markov-functional and multi factor market models. Review of Derivatives Research, 13(3), 245-272.

[58] Rebonato, R. (2002). Modern pricing of interest-rate derivatives: The LIBOR market model and beyond. Princeton University Press.

[59] Rebonato, R. (2005). Volatility and correlation: the perfect hedger and the fox. John Wiley & Sons.

[60] Rebonato, R. (2007) A time-homogeneous, SABR-consistent extension of the LMM: Calibration and numerical results. Working Paper. Tanaka Business School.

[61] Rebonato, R., McKay, K., & White, R. (2011). The SABR/LIBOR Market Model: Pricing, calibration and hedging for complex interest-rate derivatives. John Wiley & Sons.

[62] Rebonato, R., & White, R. (2009). Linking caplets and swaptions prices in the LMM-SABR model. Journal of Computational Finance, 13(2), 19.

[63] Svoboda-Greenwood, S. (2009). Displaced diffusion as an approximation of the constant elasticity of variance. Applied Mathematical Finance, 16(3), 269-286.

[64] Vasicek, O. (1977). An equilibrium characterization of the term structure. Journal of financial economics, 5(2), 177-188.

[65] Wu, L., & Zhang, F. (2002). Libor market model: from deterministic to stochastic volatility. Working paper. Claremont Graduated University and Hong Kong University of Science and Technology.

[66] Wu, L., & Zhang, F. (2006). LIBOR market model with stochastic volatility. Journal of industrial and management optimization, 2(2), 199.