



Milena Staneva Dimitrova

# Linear Rational Term Structure Models

MASTERARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

Masterstudium Finanz- und Versicherungsmathematik

eingereicht an der

**Technischen Universität Graz**

Betreuer/in

Ao.Univ.-Prof. Dipl.-Ing. Dr.techn.

Wolfgang Müller

Institut für Statistik

Graz, September 2018

# Abstract

In this master thesis we present the class of linear-rational term structure models developed by D. Filipović, M. Larson and A. Trolle, and described in their work [8]. This new class of term structure models has several distinct advantages. The short rate is non-negative at any time. The model admits analytical solution to swaptions. One can easily accommodate the unspanned stochastic volatility (USV) factors affecting the volatility and the risk premia. In the first part of this master thesis we briefly describe some important classes of term structure models. We discuss the concept of arbitrage free pricing in term structure models and present three different arbitrage free pricing systems, using a state price deflator (state space density), a risk-neutral probability measure and a market price of risk. We show that if one of these objects can be specified, then, under certain conditions, the other two can also be specified. We also take a closer look at affine processes. The second part of this master thesis focuses on examining linear-rational term structure models in greater detail. The parameters of the linear-rational square root models can be estimated from the market swap rates and market swaption prices by a maximum likelihood approach in conjunction with the unscented Kalman filter (UKF). In the last part of this master thesis we introduce the UKF algorithm. We also present the results from the parameter estimation of three different specifications of the linear-rational square-root diffusion models.

**Key words:** *term structure models, arbitrage free pricing, swaption pricing, MLE, UKF*

# Zusammenfassung

Diese Masterarbeit beschäftigt sich mit Linear-Rationalen Zinsstrukturmodellen, die von D. Filipović, M. Larson und A. Trolle in ihrer Arbeit [8] entwickelt wurden. Diese neue Klasse von Zinsstrukturmodellen hat einige besondere Eigenschaften. Die Short-Rate ist zu keiner Zeit negativ. Das Modell lässt analytische Bepreisung von Swaptions zu. Separate Modellierung von Zinsstrukturfaktoren und Faktoren, die die stochastische Volatilität beeinflussen, ist möglich. Im ersten Teil dieser Masterarbeit werden einige wichtige Klassen von Zinsstrukturmodellen beschrieben. Wir erklären das Konzept der arbitragefreien Bepreisung in Zinsstrukturmodellen und präsentieren drei verschiedene arbitragefreie Bepreisungsmethoden: Mit Hilfe der Zustand-Preis-dichte, des riskneutralen Wahrscheinlichkeitsmaßes und des Marktisikopreises. Unter bestimmten Bedingungen zeigen wir, dass eines dieser Objekte die anderen beiden festlegt. Der zweite Teil der Masterarbeit beschäftigt sich mit der Untersuchung von Linear-Rationalen Zinsstrukturmodellen. Die Parameter der Linear-Rationalen Wurzel-Diffusionsprozesse können durch die Market Swap Rates und Market Swaption Preise, mit Hilfe der Maximum-Likelihood Methode im Zusammenhang mit dem Unscented Kalman Filter (UKF), geschätzt werden. Im letzten Teil dieser Masterarbeit präsentieren wir den UKF Algorithmus und zeigen die Ergebnisse der Parameterschätzungen an Hand von drei verschiedenen Linear-Rationalen Wurzel-Diffusionsmodellen.

**Key words:** *Zinsstrukturmodelle, Arbitragefreiheit, analytische Bepreisung von Swaptions, ML-Schätzung, UKF*

## EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

---

Ort, Datum

---

Unterschrift

## STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

---

place, date

---

signature

# Acknowledgements

First and foremost, I would like to express my gratitude to my advisor, Ao.Univ.-Prof. Dipl.-Ing. Dr.techn Wolfgang Müller, who, with his constructive criticism and patience, has guided me through my work.

I also want to thank my colleagues who were always ready to answer my questions and give me advice when I needed it.

Lastly, I want to thank my family which has supported me and motivated me to keep going in the pursuit of my studies.

# Contents

<b>Introduction</b>	<b>1</b>
<b>1 Term Structure Models of Interest Rates</b>	<b>3</b>
1.1 Interest Rate Securities . . . . .	3
1.1.1 Bond . . . . .	4
1.1.2 Forward rate agreement . . . . .	6
1.1.3 Interest rate swap . . . . .	6
1.1.4 Swaption . . . . .	7
1.2 Term Structure of Interest Rates . . . . .	8
1.2.1 Extracting the Current Term Structure from the Market Prices	8
1.2.2 Modelling the Term Structure by Using Stochastic Variables . . . . .	9
1.3 Some Important Groups of Term Structure Models . . . . .	10
1.3.1 Parametric Models of the Current Term Structure . . . . .	10
1.3.2 Affine-Factor Diffusion Models . . . . .	10
1.3.3 Short Rate Models . . . . .	11
1.3.4 HJM Models . . . . .	11
1.3.5 Market Models . . . . .	12
1.4 Arbitrage free pricing . . . . .	12
1.5 Affine Diffusion Term Structure Models . . . . .	20
1.5.1 Affine Diffusion Processes . . . . .	20
1.5.2 Pricing in Affine Diffusion Models . . . . .	25
<b>2 Linear-Rational Term Structure Models</b>	<b>26</b>
2.1 The Linear-Rational Framework . . . . .	27
2.1.1 Term Structure Specification . . . . .	27
2.1.2 Unspanned Factors . . . . .	29
2.1.3 Unspanned stochastic volatility factors . . . . .	35
2.1.4 Swaption Price . . . . .	36
2.2 Models with Unspanned Stochastic Volatility (USV) Factors . . . . .	38

2.2.1	Linear-Rational Diffusion Models . . . . .	39
2.2.2	The Linear-Rational Square-Root Model . . . . .	41
2.2.3	Swaption Pricing in Linear Rational Square Root Model . . . . .	45
<b>3</b>	<b>Estimation of LRSQ Models</b>	<b>46</b>
3.1	Data . . . . .	46
3.2	Maximum Likelihood Estimation . . . . .	46
3.3	Unscented Kalman Filter . . . . .	48
3.3.1	Kalman Filter . . . . .	49
3.3.2	Unscented Transformation . . . . .	51
3.3.3	UKF State Estimation . . . . .	52
3.3.4	UKF Parameter Estimation . . . . .	54
3.4	Square Root Diffusion Process . . . . .	55
3.4.1	Definition . . . . .	55
3.4.2	Properties of $d$ -dimensional Square Root Diffusion Process . . . . .	57
3.4.3	Properties of one dimensional Square Root Diffusion Process (or one dimensional CIR Process) . . . . .	63
3.4.4	Euler-Maruyama Scheme for Simulation of Square Root Dif- fusion Process . . . . .	65
3.4.5	Exact Simulation of One Dimensional Square Root Diffusion Process . . . . .	66
3.5	Estimation of LRSQ(1,0) Model . . . . .	67
3.5.1	Model Specification . . . . .	67
3.5.2	Model Data . . . . .	68
3.5.3	Reconstruction of $\kappa, \theta, \sigma$ and $Z$ . . . . .	73
3.5.4	Maximum Likelihood Estimation in Conjunction with UKF . . . . .	75
3.6	Estimation of LRSQ(1,1) Model Specification . . . . .	77
3.6.1	Model Specification . . . . .	77
3.6.2	Model Data . . . . .	78
3.6.3	Maximum Likelihood Estimation in Conjunction with UKF . . . . .	83
3.7	Estimation of LRSQ(3,1) Model Specification . . . . .	86
3.7.1	Model Specification . . . . .	86
3.7.2	Model Data . . . . .	87
3.7.3	Maximum Likelihood Estimation in Conjunction with UKF . . . . .	91
	<b>Appendix A Stochastic calculus</b>	<b>102</b>
	<b>Bibliography</b>	<b>104</b>

# Introduction

In this master thesis we present the new class of linear rational term structure models developed by D. Filipović, M. Larson and A. Trolle, and described in their work [8]. The linear-rational term structure models are term structure factor models, in which the factors are modelled by a multivariate factor process with a drift that is affine in the current state and the state price density is given by an affine deterministic function of the current state. Under these assumptions the linear-rational term structure models are arbitrage free and the zero-coupon bond prices and the short rates become linear-rational functions of the current state. This class of models has several distinct advantages. The short rate is non-negative at any time. The model admits analytical solution to swaptions. One can easily accommodate the unspanned stochastic volatility (USV) factors affecting the volatility and the risk premia.

In the first part of this master thesis we introduce the most common interest rate securities and discuss their arbitrage free pricing. We briefly describe some important classes of term structure models. We discuss the concept of arbitrage free pricing of interest rate securities, a basic requirement for term structure models. Furthermore we present three different arbitrage free pricing mechanisms, using a state price deflator (state space density), a risk-neutral probability measure and a market price of risk. We show that if one of these objects can be specified, then, under certain conditions, the other two can also be specified. We also take a closer look at affine processes and give necessary and sufficient conditions for a diffusion process to be affine. Finally in this chapter we discuss the existence and the uniqueness of affine diffusion processes.

The second chapter focuses on examining linear-rational term structure models in greater detail. We define the linear-rational framework of the linear-rational term structure model. We show that the zero bond prices and the short rates are linear-rational functions of the current state as well as that the short rate in the model is non-negative at any time. We give necessary and sufficient conditions for which a  $d$ -dimensional linear-rational term structure model has exactly  $m$  term structure factors and exactly  $n = d - m$  unspanned factors. We examine linear-rational square-root models, a subclass of linear-rational term structure models, in



which the factor process is modelled by a multi-dimensional square-root diffusion process. Under this assumption the swaption prices become exponential-affine functions of the current state. We give the conditions, the parameters in LRSQ models have to satisfy in order to guarantee that a  $d$ -dimensional linear-rational square root diffusion process has exactly  $m$  term structure and exactly  $n = d - m$  unspanned stochastic volatility factors.

The parameters of the linear-rational square root models can be estimated from the market swap rates and market swaption prices by a maximum likelihood approach in conjunction with the unscented Kalman filter (UKF). The maximum likelihood approach finds the parameter values, which maximise the likelihood function, obtained by the unscented Kalman filter. For the unscented Kalman filter to be applied, the linear-rational square-root model has to be approximated by a state space model, consisting of measurement equations, describing the relationship between the observable market data and unobservable states of the factor process, and process equations describing the dynamics of the factor process. In chapter 3 we present the theory of Unscented Kalman Filter. We define the state space model corresponding to the LRSQ model. In the state space model the transition density of the factor process is approximated by a Gaussian density with identical first and second moments. We present a numerical method for calculating the first two moments of square root diffusion factor processes. Furthermore, we discuss the properties of the square root diffusion processes. We give a numerical method for calculating the first two moments and also describe the different methods for simulation of multi-dimensional square root processes. Finally, we present the results from the parameter estimation of three different specifications of the linear-rational square root diffusion model.

# Chapter 1

## Term Structure Models of Interest Rates

### 1.1 Interest Rate Securities

An **interest rate** is the price that has to be paid to the lender for using their money, usually given as a percentage of the borrowed money. The interest rates can be agreed for an immediate or for a future period of time, in this case the interest rates are called **spot rate** and **forward rate** respectively. Interest rates can also be classified by their calculation rule. Two examples of interest rates with different calculation rules are **linear (simple) interest rates** and **exponential (continuous) interest rates**. If  $R(t, T)$  is the annual linear interest rate at time  $t$ , an investment  $x$  at time  $t$  will yield  $x(1 + R(t, T)(T - t))$  at time  $T$ . If  $r(t, T)$  is the annual exponential interest rate at time  $t$ , the same investment  $x$  at time  $t$  yields  $x \exp(r(t, T)(T - t))$  at time  $T$ . [11].

**Interest rate securities** are securities that derive their value from an underlying interest rate or a set of different interest rates. They are contracts between two parties that specify conditions under which payments are to be made between them. The interest rate securities can be traded on an exchange. However trading them over-the-counter (OTC), which is a direct trade between two parties, is more common. The interest rate securities are most commonly used to hedge against the risk of an unfavourable movement of the interest rate. They are also used for speculation. The basic interest rate securities are bonds, forward rate agreements (FRA's), interest rate swaps, swaptions, interest rate caps and floors and constant maturity swaps.

### 1.1.1 Bond

A **bond** is the simplest interest rate security. It is a debt contract in which an investor (buyer of a bond) loans money to the issuer of bond for a definite period of time for predetermined scheduled payments [14]. Bonds are traded with various maturities and types of payment schedules. The typical issuers are governments, municipalities, private and public corporations, and financial institutions to fund their activities. Most of corporate and government bonds are publicly traded on exchanges, while others are traded only over-the-counter (OTC).

Dependant on the number of the scheduled payments we distinguish zero-coupon bonds and coupon bonds.

- **Zero-coupon bond**

A **zero-coupon bond** (or **discount bond**) is a bond that pays to the holder the *nominal value* on a single future date called *maturity date*. The nominal value is also known as *face value*, *par value* or *principle*. We assume that the nominal value of any bond is equal to one unit of a currency (for example 1 Euro). A zero-coupon bond is normally issued under the face value and its price is calculated by discounting the face value with the interest rate effective from the present moment till the maturity date. Therefore if a zero-coupon bond with maturity  $T$  is traded at some date  $t \leq T$  at a price  $P(t, T)$  then the **price of a zero-coupon bond** must be

$$P(t, T) = e^{-(T-t)r(t, T)} \quad (1.1)$$

where  $r(t, T)$  is the annual continuously-compounded interest rate effective for the period  $[t, T]$  (or the yield to maturity on a zero-coupon bond). Obviously  $P(T, T) = 1$  because the value of getting 1 Euro on the maturity date is 1 Euro. The zero-coupon bond with maturity date  $T$  is also denoted further in the text as  $(t, T)$  bond or  $T$ -bond.

The spot rate at time  $t$ , effective for the period  $[t, T]$ , can be given by the price of a  $(t, T)$ -bond

$$r(t, T) = -\frac{1}{T-t} \log P(t, T). \quad (1.2)$$

- **Coupon bond**

A bond, which regularly pays to its holder (semi-annual or annual) interest rates (coupons) in addition to the face value is called **coupon bond**. A coupon bond is initially traded near the price of its face value. Therefore the coupon bondholders gain from the regular payments of interest and not from the difference between the bond price and the face value they will receive at maturity. The coupon bond that pays a fixed coupon rate is called **fixed coupon bond**. The coupon bond that pays a variable coupon that is linked to a reference rate of interest, such as LIBOR or EURIBOR is called **floating coupon bond**. The coupon rate is recalculated periodically. Typically, the coupon rate, effective for the payment at the end of one period, is set at the beginning of the period at the current market interest rate for that period, e.g. to the 6-month interest rate for a floating rate bond with semi-annual payments. Any coupon bond can be seen as a portfolio of zero-coupon bonds which has the same payments as the coupon bond. Then the price of the coupon bond must be equal to the value of the portfolio, otherwise there is an arbitrage opportunity in the market. Therefore the price of any coupon bond can be expressed by the prices of the zero-coupon bonds in the portfolio and if the prices of the zero-coupon bonds are known then the price of the coupon bond is known too. [11, 14]

*Fixed coupon bond*

Let  $P_{T_1, \dots, T_n}^{C_1, \dots, C_n}$  denote the price of a fixed coupon bond, promising at time  $T_i$  the coupon payment  $C_i$  (for  $i = 1, \dots, n$ ) and additionally at time  $T_n$  the face value of 1. Since the portfolio consisting at time  $t < T_1$  of  $C_1$  units of  $T_1$ -bond,  $C_2$  units of  $T_2$ -bond, ...,  $C_{n-1}$  units of  $T_{n-1}$ -bond and  $C_n + 1$  units of  $T_n$ -bond ensures the same payments as the fixed coupon bond, the arbitrage free **price of the fixed coupon bond** at any time  $t < T_1$  must be

$$P_{T_1, \dots, T_n}^{C_1, \dots, C_n}(t) = P(t, T_n) + \sum_{i=1}^n C_i P(t, T_i) \quad (1.3)$$

*Floating coupon bond*

Let  $P_{T_0, T_1, \dots, T_n}^{float}$  denote the price of a floating coupon bond paying at time  $T_n$  the face value of 1 Euro and at time  $T_i$  the variable coupon  $C_i = R(T_{i-1}, T_i) \Delta_i$  for  $i = 1, \dots, n$ , where  $\Delta_i = T_i - T_{i-1}$  and  $R(T_{i-1}, T_i)$  is the reference annual linear rate. The reference annual rate, hence the variable coupon  $C_i$ , is known at time  $T_{i-1}$  and it is valid for the period  $[T_{i-1}, T_i]$ , therefore the time  $T_{i-1}$  is called reset date and the time  $T_i$  is called payment date. The same payments can be ensured by the following trading strategy (see [11]): *at time  $t$  buy one  $(t, T_0)$ -bond; at time  $T_0$*

the  $(t, T_0)$ -bond yields 1 Euro which is reinvested in  $\frac{1}{P(t, T_0, T_1)}$  units of  $P(T_0, T_1)$ -bond; at time  $T_i$  ( $i < n$ ) the  $(T_{i-1}, T_i)$ -bonds yields  $\frac{1}{P(T_{i-1}, T_i)} = 1 + R(T_{i-1}, T_i)\Delta_i = 1 + C_i$ , the coupon  $C_i$  is paid and 1 Euro is invested in  $\frac{1}{P(T_i, T_{i+1})}$  units of  $(T_i, T_{i+1})$ -bond; at time  $T_n$  the  $(T_{n-1}, T_n)$ -bonds yield  $1 + C_n$ . Since the investment for this trading strategy equals the price of one  $(t, T_0)$ -bond, the arbitrage free **price of the floating coupon bond** at any time  $t < T_0$  must be

$$P_{T_0, T_1, \dots, T_n}^{float}(t) = P(t, T_0). \quad (1.4)$$

### 1.1.2 Forward rate agreement

A **forward rate agreement (FRA)** is an over-the-counter contract between two parties (buyer and seller) that determines a fixed rate of interest for a future period of time which has to be paid or received. At maturity the difference between the contracted interest rate and the market interest rate is exchanged. This way the buyer of FRA hedges against the risk of rising interest rates, while the seller hedges against the risk of falling interest rates. How do Forward Rate Agreements work? Let  $f$  denote the contracted continuous forward rate at time  $t$  for the future period of time  $[T_0, T_1]$ . The buyer of FRA pays at time  $T_0$  1 Euro and receives at time  $T_1$  1 Euro and the interest  $e^{(T_1-T_0)f}$ . The same payments at time  $T_0$  and  $T_1$  can be ensured by the following trading strategy: *sell at time  $t$  one  $T_0$ -bond and buy at time  $t$   $e^{(T_1-T_0)f}$  units of  $T_1$ -bond* [11].

The value of these payments at time  $t$  and hence the arbitrage free **price of forward rate agreement** can be given by

$$P_{T_0, T_1}^{fra, f}(t) = -P(t, T_0) + e^{(T_1-T_0)f} P(t, T_1) \quad (1.5)$$

The **continuous forward rate**  $f(t, T_0, T_1)$  at time  $t$  for the future period  $[T_0, T_1]$  is the rate  $f$  for which  $P_{T_0, T_1}^{fra, f}(t) = 0$ , i.e. neither the buyer nor the seller pays for the FRA at time  $t$ ,

$$f(t, T_0, T_1) = -\frac{1}{T_1 - T_0} \log \left[ \frac{P(t, T_1)}{P(t, T_0)} \right]. \quad (1.6)$$

### 1.1.3 Interest rate swap

An **interest rate swap** is an agreement between two parties to exchange a stream of fixed interest rate payments and a stream of floating interest rate payments [11]. It is specified by a tenor structure of reset and payment dates  $T_0 < T_1 < \dots < T_n$  and a predetermined annualized rate  $K$ . At each payments date  $T_i$ , ( $i = 1, \dots, n$ ) the floating leg pays  $R(T_{i-1}, T_i)\Delta_i$  and the fixed leg pays  $K\Delta_i$ , where  $R(T_{i-1}, T_i)$  is the annualized reference rate and  $\Delta_i = T_i - T_{i-1}$ . Note that the sale of a fixed

coupon bond paying at time  $T_i$  the fixed coupon  $K\Delta_i$  and the buying at the same time of a floating coupon bond paying at time  $T_i$  the variable coupon  $R(T_{i-1}, T_i)\Delta_i$  for  $i = 1, \dots, n$  ensure the same payments as the interest swap from perspective of the fixed-rate payer. Hence, the arbitrage free **price of the interest swap** at any time  $t < T_0$  must be

$$P_{T_0, T_1, \dots, T_n}^{swap, K}(t) = P(t, T_0) - P(t, T_n) - \Delta_i K \sum_{i=1}^n P(t, T_i). \quad (1.7)$$

The rate  $K = S_t$  for which  $P_{T_0, T_1, \dots, T_n}^{swap, K}(t) = 0$  is called the **time- $t$  forward swap rate**. It is given by

$$S_t = \frac{P(t, T_0) - P(t, T_n)}{\sum_{i=1}^n \Delta_i P(t, T_i)} \quad (1.8)$$

Swaps are very similar to FRAs. Both are used for hedging against the risk of raising resp. falling interest rate. They can be used also for speculating. Unlike the swap in FRAs the payment is only made once at maturity. Interest rate swaps could be viewed as a chain of FRAs [11].

### 1.1.4 Swaption

A **swaption** is an option on a swap, it gives its holder the right, but not the obligation, to enter into an underlying swap at or before a given future date called maturity date. There are two type of swaption. A **payer swaption** gives the holder of the swaption the right but not the obligation to enter into an interest swap, paying the fixed leg at a pre-determined rate and receiving the floating leg [7]. A **receiver swaption** gives the holder the right to enter into an interest swap paying the floating leg and receiving the fixed leg. The buyer and seller of the swaption agree on the price of the swaption and expiration date, which is usually two business days prior to the start date of the underlying swap. They also agree on the parameters of the underlying swap as settlement/payment days, the fixed rate (which equals the strike of the swaption), the reference rate (for example, 3 month LIBOR paid quarterly). The holder of the swaption exercises the right to enter into the underlying interest swap if its value is positive on the maturity date. Hence, the swaption price also depends on this value and can be calculated only if the term structure of the interest rates is known. The swaption pricing includes development of the term structure models.

## 1.2 Term Structure of Interest Rates

The **current term structure of interest rates** (or **yield curve**) represents the relationship between the interest rates and different maturities. For a fixed time  $t$  the yield curve is observable at the market and can be given by the function  $T \rightarrow r(t, T)$ .

Because of the one to-one relationship between zero-coupon bond prices and spot rates,

$$r(t, T) = -\frac{1}{T-t} \log P(t, T),$$

the discount function  $T \rightarrow P(t, T)$  and the zero-coupon yield curve  $T \rightarrow r(t, T)$  give exactly the same information. If  $P(t, T)$  is defined for all maturity dates  $T$ , we can easily recover the yield (i.e. the annualized interest rate) for borrowing money for that period. The actual challenge in defining a yield curve therefore lies in the determination of the discount function  $T \rightarrow P(t, T)$ .

The entire **term structure of the interest rates** describes the relationship between the interest rate  $r(t, T)$  (or the price  $P(t, T)$ ) and the time variables  $t$  and  $T$ . It is given by  $(r(t, T))_{0 \leq t \leq T < \infty}$  or  $(P(t, T))_{0 \leq t \leq T < \infty}$ . The term structure of interest rates is not observable and has to be modelled.

### 1.2.1 Extracting the Current Term Structure from the Market Prices

If all zero-coupon bonds were traded, the term structure could be specified by their *market prices*. However, in many bond markets only very few zero-coupon bonds are issued and traded, which means that the term structure is known with certainty only for a few specific maturity dates. In markets where the coupon bonds are intensively traded the zero-coupon bonds can be constructed by forming certain portfolios of coupon bonds. The arbitrage free price of this zero-coupon bond can be derived and transformed into zero-coupon yield. This technique for constructing the yield of zero-coupon bond from the prices of a set of coupon-bearing products, e.g. bonds and swaps is called **bootstrapping**. Examples and detailed description of this technique are given in [14].

The bootstrapping technique is a simple technique, but the term structure extracted with this technique seems often unrealistic. A reason for this is that the bond prices are not only affected by interest rates but also by other factors. To determine more realistic term structure we need more sophisticated mathematical techniques, for example stochastic models.

## 1.2.2 Modelling the Term Structure by Using Stochastic Variables

The term structure of interest rates can also be modelled by a deterministic function of one or more time-dependent, non-observable, stochastic variables. These variables normally describe the factors which affect the term structure or factors which affect the price of interest rate derivatives but not the term structure. They are usually non-observable and have to be modelled.

The instantaneous spot rate and the instantaneous forward rate are two examples of factors affecting the term structure and can be used to describe the entire term structure of interest rates.

### *Instantaneous forward rate*

**Definition 1.1** (Instantaneous forward rate). *Assume that  $P(t, T)$  is differentiable with respect to  $T$ . The **instantaneous forward rate**  $f(t, T)$  is the forward rate with maturity  $T$ , arranged at time  $t$  for an infinitesimal short time interval  $[T, T+dT]$  defined by*

$$f(t, T) = \lim_{h \downarrow 0} f(t, T, T+h) = -\frac{\frac{dP}{dT}(t, T)}{P(t, T)}. \quad (1.9)$$

The zero-coupon bond prices  $P(t, T)$  for  $0 \leq t \leq T < \infty$  can be defined as solution of the differential equation (1.9) with initial condition  $P(T, T) = 1$  given by

$$P(t, T) = \exp\left(-\int_t^T f(t, u)du\right). \quad (1.10)$$

The instantaneous forward rates  $(f(t, T))_{0 \leq t \leq T < \infty}$  determine the prices  $P(t, T)$  for all  $0 \leq t \leq T < \infty$  and hence the entire term structure of interest rates. Since the instantaneous forward rates are unobservable at the market, they are usually modelled by the stochastic processes. [11]

### *Instantaneous spot rate (or short rate)*

**Definition 1.2** (Instantaneous spot rate). *Assume that  $P(t, T)$  is differentiable with respect to  $T$ . The **instantaneous spot rate (or short rate)**  $r_t$  is the spot rate for an infinitesimal short time interval  $[t, t+dt]$  defined by*

$$r_t = \lim_{h \downarrow 0} r(t, T) = f(t, t) = -\frac{dP}{dT}(t, T). \quad (1.11)$$

Unlike the instantaneous forward rate, the short rates  $(r(t))_{0 \leq t \leq T^*}$  can not determine the entire term structure alone. But if we assume that the short rate



follows an Itô process defined on the probability space  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$  with  $\mathcal{F}_0 = \{\emptyset, \Omega\}$  and choose a risk-neutral measure  $\mathbb{Q}$ , equivalent to  $\mathbb{P}$ , then the short rate process  $(r(t))_{0 \leq t \leq T^*}$  and the risk-neutral measure  $\mathbb{Q}$  together define an arbitrage free system of zero-coupon bond prices

$$P(t, T) = \mathbb{E}_{\mathbb{Q}} \left[ \exp \left( - \int_t^T r(u) du \right) \mid \mathcal{F}_t \right] \quad (1.12)$$

for  $0 \leq t \leq T \leq T^*$ , and hence, the entire term structure of interest rates.[11]

## 1.3 Some Important Groups of Term Structure Models

There is a wide variety of term structure models of interest rates, described in financial literature. We give a brief description of some important groups of term structure models [19].

### 1.3.1 Parametric Models of the Current Term Structure

In parametric models the current term structure is fitted by a smooth parametric function with a finite number of parameters, chosen in such a way that many of the typically observed shapes of the yield curve can be captured. One of the most popular models of these groups is the model of Nelson and Siegel introduced in 1987, in which the shape of the yield curve is modelled as

$$f_{NS}(x, \mathbf{z}) = z_1 + (z_2 + z_3 x) \exp(-z_4 x)$$

where  $z_1, z_2, z_3, z_4$  are parameters which have to be estimated. This family of interest rates curves was extended by Svensson as

$$f_{Sv}(x, \mathbf{z}) = z_1 + (z_2 + z_3 x) \exp(-z_4 x) + z_5 \exp(z_6 x)$$

This model is now used by central banks in Canada, Germany, France and the United Kingdom [19].

### 1.3.2 Affine-Factor Diffusion Models

The factor diffusion models assume the existence of factors which can affect the prices of all tradable securities and this factors are modelled by stochastic diffusion processes  $X_t$ . In affine-factor diffusion models the zero-coupon bond prices can be

written as an exponential-affine function of the short rate  $r(t, X_t)$ , which means there are deterministic functions  $A(t, T)$  and  $B(t, T)$  so that

$$P(t, T) = \exp(A(t, T) + r(t, X_t)B(t, T)).$$

An important result of Duffie and Kan (1996) shows that the yields are affine if and only if the drift and the square of the diffusion components of stochastic differential equations of  $X_t$  are also affine in  $r(t, X_t)$ . The functions  $A(t, T)$  and  $B(t, T)$  can be determined by solving Riccati differential equations.

### 1.3.3 Short Rate Models

In short rate models it is assumed that the short rates follow an Itô process defined on the probability space  $(\Omega, \mathcal{A}, \mathcal{F}, \mathbb{P})$  with  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ . The short rate process  $(r_t)_{0 \leq t \leq T^*}$  together with the risk-neutral pricing measure  $\mathbb{Q}$  define an arbitrage free system of zero-coupon bond prices given by (1.12), and hence they determine the entire term structure of interest rates. The first attempt in this direction was made in 1977 by O. Vasicek. He describes the short rate by an Ornstein-Uhlenbeck process

$$dr_t = \kappa(\theta - r_t)dt + \sigma dB_t.$$

The short rate in Vasicek's model is asymptotic normal distributed with mean  $\theta$  and variance  $\frac{\sigma^2}{2\kappa}$ . A disadvantage of this model is that it is theoretically possible for the interest rates to become negative. Furthermore, the model is not flexible enough to capture all shapes of the yield curve.

In the CIR model, proposed in 1985 by Cox, Ingersoll and Ross (CIR), the short rate is described by

$$dr_t = \kappa(\theta - r_t)dt + \sigma\sqrt{r_t}dB_t.$$

The short rate in the CIR model is non-centred  $\chi^2$ -square distributed and almost surely positive. Nevertheless the model is not flexible enough.

Since both the drift and  $\sigma^2(r_t)$  are affine function of  $r_t$  both Vasicek and CIR models are one factor affine models where the only factor is the short rate.

### 1.3.4 HJM Models

Heath, Jarrow and Morton (HJM) introduced in 1992 a new framework for modelling of interest rates. They model the evolution of the forward rates by a family of (an infinite number of) stochastic differential equations

$$\forall T : 0 \leq T \leq T^* \quad f(t, T) = f(0, T) + \int_0^t \alpha(t, T)dt + \int_0^t \sigma(t, T)dB_t \quad \forall t \in [0, T].$$

In this class of models  $\sigma$  is essentially chosen freely, while  $\alpha$  is determined from  $\sigma$  by the HJM drift condition in order to get an arbitrage free model.

The instantaneous forward rates  $(f(t, T))_{0 \leq t \leq T \leq T^*}$  determine the zero-coupon bond prices for all  $0 \leq t \leq T \leq T^*$  by (1.10) and hence the entire term structure of interest rates.

### 1.3.5 Market Models

The LIBOR market model does not describe the entire term structure. It models the (linear) forward rates  $L_i = F(t, T_{i-1}, T_i)$ , ( $0 \leq t \leq T_{i-1}$ ), for a fixed tenor structure  $T_0 < T_1 < \dots < T_n$  as lognormal process.

$$\frac{dL_i(t)}{L_i(t)} = \mu_i(t)dt + \sigma_i(t)dW_i, \quad i = 1, \dots, n,$$

Note that at time  $t$  the LIBOR rates  $L_i(t)$  are directly observable in the market. Its volatilities are naturally linked to traded contracts. This model class was first introduced in 1997 by Brace, Gatarek and Musiela [2] and extended by Miltersen, Sandmann and Sondermann in the same year. It is used for pricing interest rate derivatives, especially exotic derivatives like Bermudan swaptions, ratchet caps and floors, etc. [19]

## 1.4 Arbitrage free pricing

An essential requirement to term structure models is the arbitrage-free pricing of all interest rate securities. An **arbitrage** is a trading strategy that generates a risk-free profit due to the price differences at the same time at different markets [14]. Munk describes in [14] the general arbitrage free pricing mechanisms using three different objects: a *state-price deflator* (also known as *state space density*), a *risk-neutral probability measure*, and a *market price of risk*. He shows that if one of these objects can be specified, then the model is arbitrage free and a payoff stream can be specified. If one of these objects can be specified, then the other two can also be specified, under certain conditions. In this section we present these pricing mechanisms.

Consider a model of financial market in which one instantaneously risk-free asset (i.e. bank account) and one zero-coupon bond with maturity  $T$  are traded. Since the zero-coupon bond is traded up to maturity  $T$  and the prices after maturity date are not relevant, we define the model for the time period  $[0, T]$ . We assume for this model that the basic uncertainty in economy is represented by a standard Brownian motion,  $B = (B_t)_{0 \leq t \leq T}$  defined on the filtered probability space  $(\Omega, \mathcal{A}, \mathbb{P}, \mathcal{F})$  with  $\mathcal{F}_0 = \{\emptyset, \Omega\}$ .

Let  $A = (A_t)_{0 \leq t \leq T}$  denote the price process of the **bank account** continuously compounded with the instantaneously risk-free interest rate  $r_t$  for the infinitesimal interval  $[t, t + dt]$ . The price of a bank account at time  $t \leq T$  describes the value at time  $t$  of a deposit  $A_0$  made at time  $t = 0$  and therefore satisfies

$$dA_t = A_t r_t dt$$

which is equivalent to

$$A_t = e^{\int_0^t r_t dt} \quad (1.13)$$

assuming that  $A_0 = 1$  and that the process  $r = (r_t)$  is such that  $\int_0^t |r_t| dt$  is finite with probability one.

The price of a zero-coupon bond at time  $t > 0$  is random except on the maturity date  $T$ , when the nominal value is repaid. On this day it is  $P(T, T) = 1$ . We assume that the price at time  $t \leq T$  of the **zero-coupon bond with maturity day  $T$**  follows an Itô process (see Definition A.2) of the form

$$dP(t, T) = P(t, T)(\mu_t dt + \sigma_t dB_t), \quad P(T, T) = 1 \quad (1.14)$$

where  $\mu = (\mu_t)$  denotes the (relative) drift, and  $\sigma = (\sigma_t)$  reflects the relative sensitivity of the price to the exogenous shock to the economy at time  $t$ , described by the evolution of the standard Brownian motion  $B = (B_t)_{0 \leq t \leq T}$ .

For this model we will represent three equivalent arbitrage-free price systems based on three different objects: state-price deflators, risk-neutral probability measures, and markets price of risk and we show that if one of these objects is specified, any payoff stream can be priced.

## State-price deflators

**Definition 1.3.** A **state-price deflator** (or **state space density**) is a strictly positive process  $\zeta = (\zeta_t)$  with  $\zeta_0 = 1$  and the property that the product of the state-price deflator, and the price of an asset is a martingale, i.e.  $(\zeta_t P(t, T))$  and  $(\zeta_t A_t) = (\zeta_t \exp(\int_0^t r_u du))$  are  $\mathbb{P}$ -martingales. [14]

Suppose we are given a state-price deflator  $\zeta$  and hence the distribution of  $\zeta_T/\zeta_t$ . Then it follows from the martingale properties of  $(\zeta_t P(t, T))$  that the price at time  $t \leq T$  of a zero-coupon bond with maturity  $T$  is

$$P(t, T) = E_t \left[ \frac{\zeta_T}{\zeta_t} P(T, T) \right] = E_t \left[ \frac{\zeta_T}{\zeta_t} \right], \quad (1.15)$$

where  $E_t$  denote the expected value under the  $\mathbb{P}$ -probabilities conditionally on the available information at time  $t$ , i.e.  $E_t[X] = E[X | \mathcal{F}_t]$ .

The existence of a state-price deflator (state space density) is closely related to absence of arbitrage:

**Theorem 1.4.** *If a state-price deflator exists, the prices admit no arbitrage. [14]*

**Theorem 1.5.** *If the prices admit no arbitrage and technical conditions are satisfied, then a state-price deflator exists. [14]*

**Corollary 1.6.** *Under technical conditions, the existence of a state-price deflator is equivalent to the absence of arbitrage. [14]*

A proof of these theorems can be found in [14]. Duffie (2001) gives a precise description of the technical condition. Munk (2005) shows that if prices admit no arbitrage and technical conditions are satisfied, then the marginal rate of intertemporal substitution of some agent investing in the market, defined by  $\zeta_t = e^{-\delta t} u'(c_t)/u'(c_0)$ , is a state-price deflator. Here it is assumed that  $c = (c_t)$  is the optimal consumption process for this agent,  $u(\cdot)$  is a utility function and  $\delta$  is the time-preference rate of this agent. Since agents have different utility functions, different time preference rates, and different optimal consumption plans, there can potentially be (at least) as many state-price deflators as agents.

## Risk-neutral probability measures

Bonds are usually priced with the help of a so-called risk-neutral probability measure.

**Definition 1.7.** *A probability measure  $\mathbb{Q}$  is said to be a **riskneutral probability measure** (or **equivalent martingale measure**) if the following three conditions are satisfied:*

(i)  $\mathbb{Q}$  is equivalent to  $\mathbb{P}$ ,

(ii) the discounted price process  $\tilde{P}(t, T) = A_t^{-1} P(t, T) = P(t, T) \exp\left(-\int_0^t r_u du\right)$  is a  $\mathbb{Q}$ -martingale,

(iii) the Radon-Nikodym derivative  $d\mathbb{Q}/d\mathbb{P}$  has finite variance.

(see [14])

The marginal property (ii) implies that under the risk-neutral probability measure  $\mathbb{Q}$  the price at time  $t \leq T$  of a zero-coupon bond with maturity  $T$  is given by

$$P(t, T) = E_t^{\mathbb{Q}} \left[ e^{-\int_t^T r_u du} P(T, T) \right] = E_t^{\mathbb{Q}} \left[ e^{-\int_t^T r_u du} \right] \quad (1.16)$$

The Radon-Nikodym derivative  $d\mathbb{Q}/d\mathbb{P}$  is a random which is known at time  $T$  and usually not before time  $T$  since in our market all uncertainties are resolved at time  $T$ . Let  $d\mathbb{Q}/d\mathbb{P} = \xi_T$  and then define a density process  $\xi = (\xi_t)_{0 \leq t \leq T}$  by

$$\xi_t = E_t [d\mathbb{Q}/d\mathbb{P}] = E_t [\xi_T]$$

The process  $(\xi_t)$  is a  $\mathbb{P}$ -martingale, since for any  $t < t' \leq T$  we have

$$E_t[\xi_{t'}] = E_t[E_{t'}[\xi_T]] = E_t[\xi_T] = \xi_t$$

Then under the probability measure  $\mathbb{P}$  the price at time  $t \leq T$  of a zero-coupon bond with maturity  $T$  is given by Bayes rule

$$P(t, T) = E_t^{\mathbb{Q}} \left[ e^{-\int_t^T r_u du} \right] = \frac{E_t \left[ \xi_T e^{-\int_t^T r_u du} \right]}{E_t[\xi_T]} = E_t \left[ \frac{\xi_T}{\xi_t} e^{-\int_t^T r_u du} \right] \quad (1.17)$$

The existence of the risk-neutral measure  $\mathbb{Q}$  implies that the model, we consider, is arbitrage free.

**Theorem 1.8.** *If a risk-neutral probability measure exists, prices admit no arbitrage.*

*Proof:* (see [14], p. 76)

The next theorem shows that there is a one-to-one relation between risk-neutral probability measures and state-price deflators. Under certain technical conditions the state-price deflator can specify the risk neutral measure, and vice versa the state-space deflator can be specified by Radon-Nykodim derivative  $d\mathbb{Q}/d\mathbb{P}$ .

**Theorem 1.9.** *Given a risk-neutral probability measure  $\mathbb{Q}$ . Let  $\xi_t = E_t[d\mathbb{Q}/d\mathbb{P}]$  and define  $\zeta_t = \xi_t \exp(-\int_0^t r_u du)$ . If  $\zeta_t$  has finite variance for all  $t \leq T$ , then  $\zeta = (\zeta_t)$  is a state-price deflator. Conversely, given a state-price deflator  $\zeta$ , define  $\xi_t = \exp(\int_0^t r_u du)\zeta_t$ . If  $\xi_T$  has a finite variance, then a risk-neutral probability measure  $\mathbb{Q}$  is defined by  $d\mathbb{Q}/d\mathbb{P} = \xi_T$ . [14]*

*Proof:* Suppose that  $\mathbb{Q}$  is a risk-neutral probability measure and  $\xi_t = E_t[d\mathbb{Q}/d\mathbb{P}] = E_t[\xi_T]$  is the probability density process. We prove that  $\zeta_t = \xi_t \exp(-\int_0^t r_u du)$  is a state price density. According to Definition 1.3 of state price density we have to show that  $\zeta_0 = 1$ ,  $\zeta_T > 0$  and  $(\zeta_t A_t)$  and  $(\zeta_t P(t, T))$  are  $\mathbb{P}$ -martingals.

Obviously  $\zeta_0 = \xi_0 = \mathbb{E}[d\mathbb{Q}/d\mathbb{P}] = \mathbb{Q}(\Omega) = 1$  if  $\mathcal{F}_0 = \{\Omega, \emptyset\}$  and  $\zeta_T > 0$  since  $\xi_T > 0$ .  $(\zeta_t A_t)$  is a  $\mathbb{P}$ -martingale, because  $\zeta_t A_t = \zeta \exp(\int_0^t r_u du) = \xi_t$  and  $\xi_t = E_t[d\mathbb{Q}/d\mathbb{P}]$  is a  $\mathbb{P}$ -martingale. Further, since  $\mathbb{Q} \sim \mathbb{P}$  and  $\xi$  is the probability density process it follows that  $e^{-\int_0^t r_u du} P(t, T)$  is a  $\mathbb{Q}$ -martingale if and only if  $e^{-\int_0^t r_u du} P(t, T)\xi_t$  is a  $\mathbb{P}$ -martingale. The martingale property implies

$$e^{-\int_0^t r_u du} P(t, T)\xi_t = E_t \left[ e^{-\int_0^T r_u du} P(T, T)\xi_T \right]$$

and hence  $\zeta_t P(t, T) = E_t[\zeta_T]$  is a  $\mathbb{P}$ -martingale.

Conversely, suppose that  $(\zeta_t)_{t \leq T}$  is a state-price density, then according to Definition 1.3  $\xi_t = \zeta_t \exp(\int_0^t r_u du)$  is a  $\mathbb{P}$ -martingale. Since  $\zeta_T > 0$  and  $\mathbb{E}^{\mathbb{P}}[\xi_T] = \mathbb{E}^{\mathbb{P}}[\xi_t] = \mathbb{E}^{\mathbb{P}}[\xi_0] = 1$  (martingale property) the random variable  $\frac{d\mathbb{Q}}{d\mathbb{P}} = Z_T$  defines a probability measure  $\mathbb{Q}$ , equivalent to  $\mathbb{P}$ , with a probability density process  $\mathbb{E}^{\mathbb{P}}[Z_T | \mathcal{F}_t] = \xi_t$ . Hence,  $e^{-\int_0^t r_u du} P(t, T)$  is a  $\mathbb{Q}$ -martingale if and only if  $e^{-\int_0^t r_u du} P(t, T) \xi_t = \zeta_t P(t, T)$  is a  $\mathbb{P}$ -martingale. Since by assumption  $\zeta_t P(t, T)$  is a  $\mathbb{P}$ -martingale it follows, according to Definition 1.7, that  $\mathbb{Q}$  is an equivalent martingale measure.

□

In the previous subsection we saw that, the absence of arbitrage implies the existence of a state-price deflator under some technical conditions. It follows from Theorem 1.5 and Theorem 1.9 the absence of arbitrage also implies the existence of a risk-neutral probability measure - again under technical conditions.

**Theorem 1.10.** *If the prices admit no arbitrage and technical conditions are satisfied, then a risk neutral measure  $\mathbb{Q}$  exists.*

### Market price of risk

Suppose that  $\mathbb{Q}$  is a risk-neutral probability measure with Radon-Nikodym derivative  $d\mathbb{Q}/d\mathbb{P} = \xi_T$ . As we already discussed the process  $\xi_t = (\xi_t)$ , defined by  $\xi_t = E_t[d\mathbb{Q}/d\mathbb{P}] = E_t[\xi_T]$ , is a square integrable  $\mathbb{P}$ -martingale with  $\xi_0 = E[d\mathbb{Q}/d\mathbb{P}] = 1$  (since  $\mathbb{Q}$  is a probability measure). Then it follows from the *Martingale Representation Theorem* (see Annex), that a process  $\lambda = (\lambda_t)$  exists such that

$$d\xi_t = -\xi_t \lambda_t dB_t$$

or, equivalently (using  $\xi_0 = 1$ ),

$$\xi_t = \exp\left(-\frac{1}{2} \int_0^t \lambda_u^2 du - \int_0^t \lambda_u dB_u\right). \quad (1.18)$$

According to *Girsanov's Theorem* (see Annex), the process  $W = (W_t)$  defined by

$$dW_t = dB_t + \lambda_t dt \quad (1.19)$$

is a standard Brownian motion under the  $\mathbb{Q}$ -measure.

Let us define the process  $\lambda = (\lambda_t)$ . Since we assumed that  $\mathbb{Q}$  is a risk-neutral measure, it follows from Definition 1.7 that the discounted price process given by  $\tilde{P}(t, T) = \exp(-\int_0^t r_u du) P(t, T)$  is a  $\mathbb{Q}$ -martingale. Applying *Itô's Formula* (see

Annex) and (1.15), and then substituting  $dB_t = dW_t - \lambda_t dt$ , we obtain that the dynamics of the discounted prices are given by

$$d\tilde{P}(t, T) = \tilde{P}(t, T)(\mu_t - r_t - \sigma_t \lambda_t)dt + \sigma_t dW_t \quad (1.20)$$

Since, according to Definition 1.7, the discounted price process is a  $\mathbb{Q}$ -martingale, the drift must be zero and hence  $\lambda$  is defined as a solution of

$$\mu_t - r_t = \sigma_t \lambda_t. \quad (1.21)$$

Suppose that  $\lambda$  is a solution of (1.21). Suppose also that  $\lambda$  satisfies Novikov's condition, i.e.  $E_t \left[ \exp \left( \frac{1}{2} \int_0^T \lambda_u^2 du \right) \right] < \infty$ . Let us define the process  $\xi_t$  by (1.18) and assume that  $\xi_T$  has finite variance. Under these technical conditions the measure  $\mathbb{Q}$ , defined by  $d\mathbb{Q}/d\mathbb{P} = \xi_T$ , satisfies all conditions of risk neutral measure given in Def. 1.7, and hence the measure  $\mathbb{Q}$  is a risk-neutral probability measure. [*Proof:* (ii) Novikov's condition ensures that the process  $\xi = (\xi_t)$  is a martingale. According to Girsanov's Theorem, the process  $W$ , defined by (1.19), is a standard Brownian moving under  $\mathbb{Q}$  and the discount price process  $\tilde{P}$  is a  $\mathbb{Q}$ -martingale. (i)  $\xi_T > 0$  and  $\xi_0 = 1$  imply that  $\mathbb{Q}$  and  $\mathbb{P}$  are equivalent. (iii) satisfied by assumption]. These results are summarized in the following theorem:

**Theorem 1.11.** *If a risk-neutral probability measure exists, there must be a solution to (1.21) for all  $t$ . If a solution  $\lambda_t$  exists for all  $t$  and the process  $\lambda = (\lambda_t)$  satisfies technical conditions, then a risk-neutral probability measure exists. [14]*

**Definition 1.12.** *Any process  $\lambda = (\lambda_t)$  solving (1.21) is called a **market price of risk process**.*

The market price of risk is a measure of the extra return, or risk premium, that investors demand to bear risk.

Theorem 1.11 proves that *there is a one-to-one relation between risk-neutral probability measures and market prices of risk*. According to the result from a previous subsection we can conclude that under technical condition *the existence of a mark-price of risk is equivalent to the absence of arbitrage*.

Theorem 1.9 and Theorem 1.11 imply that up to technical condition *there is also a one-to-one relationship between state-price deflators and the market price of risk*.

Supposing that  $\lambda$  is a market price of risk and  $\xi = (\xi_t)$  is a process defined by (1.18), then under the certain technical condition, the process  $\zeta$  defined by

$$\zeta_t = \xi_t e^{-\int_0^t r_u du} = \exp \left( -\int_0^t r_u du - \frac{1}{2} \int_0^t \lambda_u^2 du - \int_0^t \lambda_u dB_u \right) \quad (1.22)$$



is a state-price deflator and  $\xi$  is a density process [14]. We apply Itô's Lemma to (1.22). Since  $\xi$  also satisfies  $d\xi_t = -\xi_t\lambda_t dB_t$ , it follows that the state-space deflator satisfies

$$d\zeta_t = -\zeta_t(r_t dt + \lambda_t dB_t). \quad (1.23)$$

We see that the relative drift of a state-price deflator equals the negative of the short-term interest rate. The sensitivity vector of a state-price deflator equals the negative of a market price of risk.

We proved that the *state-space deflator*, the *risk-neutral probability measure* and the *market price of risk* are three different but equivalent objects. If one exists, the other two can be defined. Under the certain technical condition the existence of one of them is equivalent to the absence of arbitrage. These objects give three different (but equivalent) ways to define arbitrage free price systems.

We defined three different *formulas for the arbitrage free pricing at time  $t \leq T$  of a zero-coupon bond with maturity date  $T$*  by

$$\begin{aligned} P(t, T) &= E_t \left[ \frac{\zeta_T}{\zeta_t} \right] \\ &= E_t^{\mathbb{Q}} \left[ e^{-\int_t^T r_u du} \right] \\ &= E_t \left[ \exp \left( -\int_t^T r_u du - \frac{1}{2} \int_t^T \lambda_u^2 du - \int_t^T \lambda_u dB_u \right) \right] \end{aligned}$$

where  $\zeta$  is the state-price deflator,  $\mathbb{Q}$  is the risk-neutral measure,  $\lambda$  is the market price of risk.

This arbitrage free price formulas can also be used for pricing of zero-coupon bonds with maturity date  $T' \leq T$ . Suppose that  $\mathbb{Q}$  is a risk-neutral probability measure. A zero-coupon bond with maturity date  $T' \leq T$  pays at time  $T'$  one unit of currency, e.g. 1 Euro. If we invest at time  $T'$  this payment in the bank account over the period  $[T', T]$ , then the pay off at time  $T$  is  $e^{\int_{T'}^T r_u du}$ . The price we pay at time  $t$  for these two investments is exactly the price at time  $t$  of the zero-coupon bond with maturity date  $T'$ ,  $P(t, T')$ , and it must correspond to the  $t$ -time value of the  $T$ -time pay off. Otherwise there is an arbitrage opportunity. Hence, the price at any time  $t \leq T' \leq T$  of the zero-coupon bond with maturity date  $T'$  under  $\mathbb{Q}$  must be

$$P(t, T') = E_t^{\mathbb{Q}} \left[ e^{-\int_t^{T'} r_u du} \left( P(T', T') e^{\int_{T'}^T r_u du} \right) \right] = E_t^{\mathbb{Q}} \left[ e^{-\int_t^{T'} r_u du} \right].$$

In term of state-price deflator we have

$$P(t, T') = E_t \left[ \frac{\zeta_{T'}}{\zeta_t} \right].$$

If  $\lambda$  is a market price of risk we have

$$P(t, T') = E_t \left[ \exp \left( - \int_t^{T'} r_u du - \frac{1}{2} \int_t^{T'} \lambda_u^2 du - \int_t^{T'} \lambda_u dB_u \right) \right].$$

The arbitrage free price formulas can be applied also for zero-coupon bond pricing in models with  $N$  risky assets paying a single dividend at time  $T$  [14]. In this case the assets price dynamics is given by

$$dP_t = \text{diag}(P_t)(\mu_t dt + \sigma_t dB_t)$$

Here  $P_t = (P_{1t}, \dots, P_{Nt})^T$  is the price vector of all  $N$  risky assets,  $\mu_t = (\mu_{1t}, \dots, \mu_{Nt})^T$  is a vector of expected returns,  $\sigma_t = (\sigma_{ijt})_{1 \leq i \leq N, 1 \leq j \leq d}$  is a  $N \times d$  matrix reflecting the sensitivity of the prices to the exogenous shocks and determines the instantaneous variances and covariances (and hence also the correlations) of the risky asset prices.  $B_t = (B_{1t}, \dots, B_{dt})$  is a vector of  $d$  exogenous shocks to the economy at time  $t$ . We assume that all the uncertainties that affect the investors stem from these exogenous shocks.

The market price of risk is, in this case, any process  $\lambda = (\lambda_t)$  solving the system of  $N$  equations and  $d$  unknown variables  $\lambda_{1t}, \dots, \lambda_{dt}$

$$\sigma_t \lambda_t = \mu_t - r_t \mathbf{1}_d \tag{1.24}$$

The number of solutions of this system depends on the rank of the  $N \times d$  matrix  $\sigma_t$ . Let  $k$  denote the rank of the  $\sigma_t$ . If  $k < d$ , there are several solutions to (1.24),

hence several arbitrage free price systems. Therefore the prices of the traded risky assets are not unique and they are only sensitive to  $k$  of the  $d$  exogenous shocks (risks). If  $k = d$  there is an unique solution to (1.24) given by

$$\lambda_t^* = \sigma_t^{-1}(\mu_t - r_t 1_d)$$

and therefore an unique arbitrage free price system. The prices are sensitive to all risks of the market. This is possible only in so called *complete markets* [14].

All definitions and theorems we consider in this section are valid also for models with  $N$  risky assets paying a single dividend at time  $T$ .

Finally, it is important to note that all tree arbitrage free price formulas are defined by conditional expectations and in affine term structure factor models the conditional expectations and hence the zero-coupon bond prices can be given in a closed form. Therefore the affine term structure factor models are often the preferred models. Some properties of the affine term structure factor models will be introduced in the following section.

## 1.5 Affine Diffusion Term Structure Models

The factor models assume the existence of one or several **state variables (factors)** whose current values contain all the relevant information about the economy and which can affect the short rate, market price of risk and hence the prices of all traded securities. Principally these factors are unobservable and have to be modelled.

In affine diffusion term structure models the state variables follow an affine diffusion process.

### 1.5.1 Affine Diffusion Processes

Consider a  $d$ -dimensional stochastic diffusion process  $X = (X_t)_{t \geq 0}$  with values in  $E \subseteq \mathbb{R}^d$  which solves the stochastic differential equation

$$dX_t = b(X_t, t)dt + \rho(X_t, t)dB_t, \quad X_0 = x \tag{1.25}$$

where  $x \in E$ ,  $B = (B_t)_{t \geq 0}$  is a  $d$  dimensional standard Brownian motion defined on the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ ,  $b : E \rightarrow \mathbb{R}^d$  is continuous in  $X \in E$ ,  $\rho : E \rightarrow \mathbb{R}^d$  is measurable so that the diffusion matrix  $a(X_t, t) = \rho(X_t, t)\rho(X_t, t)^\top$  is continuous in  $X \in E$ . Suppose that for any  $x \in E$  there exists a unique solution  $X = X^x$  of SDE (1.25).

**Definition 1.13 (Affine Diffusion Process).** *The stochastic diffusion process  $X$  is **affine** if  $\mathcal{F}_t$ -conditional characteristic function of  $X_T$  is exponential affine in  $X_t$ , i.e. there exist  $\mathbb{C}$ - and  $\mathbb{C}^d$ -valued functions  $\phi(t, u)$  and  $\psi(t, u)$ , respectively, with jointly continuous  $t$ -derivatives such that  $X = X^x$  satisfies*

$$\mathbb{E} \left[ e^{u^\top X_T} \mid \mathcal{F}_t \right] = e^{\phi(T-t, u) + \psi(T-t, u)^\top X_t} \quad (1.26)$$

for all  $u \in i\mathbb{R}^d, t \leq T, x \in E$ .

*Remarks:*

- 1) Since the conditional characteristic function is bounded by one, the real part of the exponent  $\phi(T-t, u) + \psi(T-t, u)^\top X_t \leq 0$  in (1.26) has to be negative [9].
- 2)  $\phi$  and  $\psi$  uniquely determined by (1.26) with  $\phi(0, u) = 0, \psi(0, u) = u$ .

The following two theorems provide necessary and sufficient conditions for the stochastic diffusion process  $X$  (1.25) to be affine [7].

**Theorem 1.14 (Necessary Condition).** *Suppose that  $X$  is affine. Then the diffusion matrix  $a(x, t)$  and drift  $b(x, t)$  are affine in  $x$ , i.e.*

$$\begin{aligned} a(x, t) &= a + \sum_{i=1}^d x_i \alpha_i \\ b(x, t) &= b + \sum_{i=1}^d x_i \beta_i = b + \mathcal{B}x \end{aligned} \quad (1.27)$$

for some  $d \times d$ -matrices  $a$  and  $\alpha_i$  and  $d$ -vectors  $b$  and  $\beta_i$ , where

$$\mathcal{B} = (\beta_1, \dots, \beta_d)$$

denotes the  $d \times d$ -matrix with  $i$ th column vector  $\beta_i, 1 \leq i \leq d$ . Moreover,  $\phi$  and  $\psi = (\psi_1, \dots, \psi_d)^\top$  solve the system of Riccati ODEs

$$\begin{aligned} \partial_t \phi(t, u) &= \frac{1}{2} \psi(t, u)^\top a \psi(t, u) + b^\top \psi(t, u) \\ \psi(0, u) &= 0 \\ \partial_t \psi_i(t, u) &= \frac{1}{2} \psi(t, u)^\top \alpha_i \psi(t, u) + \beta_i^\top \psi(t, u), \quad 1 \leq i \leq d \\ \psi(0, u) &= u. \end{aligned} \quad (1.28)$$

In particular  $\phi$  is determined by  $\psi$  via simple integration

$$\phi(t, u) = \int_0^t \left( \frac{1}{2} \psi(s, u)^\top a \psi(s, u) + b^\top \psi(s, u) \right) ds.$$

*Proof:* Suppose that  $X$  is affine. For  $T > 0$  and  $u \in i\mathbb{R}^d$  define the complex-valued Itô process

$$M(t) = \exp(\phi(T-t, u) + \psi(T-t, u)^\top X_t).$$

Applying Itô formula separate to the real and to the imaginary part of  $M$ , one obtains

$$dM(t) = I(t)dt + \psi(T-t, u)^\top \rho(X_t)dB_t, \quad t \leq T$$

with

$$\begin{aligned} I(t) = & -\partial_T \phi(T-t, u) - \partial_T \psi(T-t, u)^\top X_t \\ & + \psi(T-t, u)^\top b(X_t) + \frac{1}{2} \psi(T-t, u)^\top a(X_t) \psi(T-t, u) \end{aligned}$$

Since  $M(t)$  is a martingale, we have  $I(t) = 0$  for all  $t \leq T$  a.s. Letting  $t \rightarrow 0$ , by continuity of drift  $b(X_t)$  and diffusion matrix  $a(X_t)$  it follows that

$$\partial_T \phi(T, u) + \partial_T \psi(T, u)^\top x = \psi(T, u)^\top b(x) + \frac{1}{2} \psi(T, u)^\top a(x) \psi(T, u)$$

for all  $x \in E$ ,  $T \geq 0$ ,  $u \in i\mathbb{R}^d$ . Since  $\psi(0, u) = u$  it follows that the  $a(x)$  and  $b(x)$  are affine of the form (1.27). Plugging this into the equation above and separating first order terms in  $x$  one obtains the Ricatti SDEs (1.28). [7]

□

**Theorem 1.15 (Sufficient Condition).** *Suppose that the diffusion matrix  $a(x, t)$  and drift  $b(x, t)$  are affine of the form (1.27) and suppose there exists a solution  $(\phi, \psi)$  of the Ricatti ODEs (1.28) such that  $\text{Re}(\phi(T-t, u) + \psi(T-t, u)^\top X_t) \leq 0$  for all  $t \geq 0$ ,  $u \in i\mathbb{R}^d$  and  $x \in E$ . Then  $X$  is affine with conditional characteristic function (1.26).*

*Proof:* Suppose that  $a(x, t)$  and  $b(x, t)$  are affine of the form (1.27) and  $(\phi, \psi)$  is a solution of the Ricatti ODEs (1.28) such that  $\text{Re}(\phi(T-t, u) + \psi(T-t, u)^\top X_t) \leq 0$  for all  $t \geq 0$ ,  $u \in i\mathbb{R}^d$  and  $x \in E$ . Then  $M$ , defined as above, is uniformly bounded local martingale, and hence a martingale with  $M_T = e^{u^\top X_T}$ . Therefore  $\mathbb{E}[M_T | \mathcal{F}_t] = M_t$ , which is equivalent to (1.26). [7]

□

*Remark:* In the literature affine diffusion processes are frequently defined as stochastic diffusion process with affine drift and diffusion matrix of the form (1.27).

Up until now we assumed the existence of a unique solution  $X^x$  of SDE (1.25) on some space  $E \subset \mathbb{R}$ . Now the question arises if such a solution truly exists. Filipović and Mayerhofer prove in [9] the existence and the uniqueness of affine processes on the canonical state space  $E = \mathbb{R}_+^m \times \mathbb{R}^n$  for some integers  $m, n \geq 0$  with  $m + n = d$ .

The following theorem characterises the affine processes on the canonical state space.

**Theorem 1.16.** *The process  $X$  on the canonical state space  $\mathbb{R}_+^m \times \mathbb{R}^n$  is affine if and only if  $a(x)$  and  $b(x)$  are affine of the form (1.27) for parameters  $a, \alpha_i, b, \beta_i$  which are admissible in the following sense:*

$$\begin{aligned}
& a, \alpha_i \text{ are symmetric positive semi-definite,} \\
& a_{II} = 0 \text{ ( and thus } a_{IJ} = a_{JI}^\top = 0 \text{),} \\
& \alpha_j = 0 \text{ for all } j \in J, \\
& \alpha_{i,kl} = \alpha_{i,lk} = 0 \text{ for } k \in I \setminus \{i\}, \text{ for all } 1 \leq i, l \leq d, \\
& b \in \mathbb{R}_+^m \times \mathbb{R}^n, \\
& \mathcal{B}_{IJ} = 0, \\
& \mathcal{B}_{II} \text{ has positive off-diagonal elements.}
\end{aligned} \tag{1.29}$$

In this case, the corresponding system of Riccati equations (1.28) simplifies to

$$\begin{aligned}
\partial_t \phi(t, u) &= \frac{1}{2} \psi_J(t, u)^\top a_{JJ} \psi_J(t, u) + b^\top \psi(t, u), \\
\phi(0, u) &= 0, \\
\partial_t \psi_i(t, u) &= \frac{1}{2} \psi(t, u)^\top \alpha_i \psi(t, u) + \beta_i^\top \psi(t, u), \quad i \in I, \\
\partial_t \psi_J(t, u) &= \mathcal{B}_{JJ}^\top \psi_J(t, u), \\
\psi(0, u) &= u,
\end{aligned}$$

and there exists a unique global solution  $(\phi(\cdot, u), \psi(\cdot, u)) : \mathbb{R}_+ \rightarrow \mathbb{C}_- \times \mathbb{C}_-^m \times i\mathbb{R}^n$  for all initial values  $u \in \mathbb{C}_-^m \times i\mathbb{R}^n$ . In particular, the equation for  $\psi_J$  forms an autonomous linear system with unique global solution  $\psi_J(t, u) = e^{\mathcal{B}_{JJ}^\top t} u_J$  for all  $u_J \in \mathbb{C}^n$ . ([9], Theorem 3.2)

**Corollary 1.17 (Exponential Affine Transform Formula).** *Suppose that  $X$  is affine diffusion process. Then there exist  $\mathbb{C}$ - and  $\mathbb{C}^d$ -valued functions  $\phi(t, u, v)$  and  $\psi(t, u, v)$ , respectively, with jointly continuous  $t$ -derivatives such that  $X = X^x$  satisfies*

$$\mathbb{E} \left[ e^{u+v^\top X_T} \mid \mathcal{F}_t \right] = e^{\phi(T-t, u, v) + \psi(T-t, u, v)^\top X_t} \quad (1.30)$$

for all  $u \in \mathbb{R}$ ,  $v \in \mathbb{R}^d$ ,  $t \leq T$ ,  $x \in E$ .

- $Re(\phi(T-t, u) + \psi(T-t, u)^\top X_t) \leq 0$ ,
- $\phi$  and  $\psi$  solve the system of Riccati SDEs

$$\begin{aligned} \partial_t \phi(t, u, v) &= \frac{1}{2} \psi(t, u, v)^\top a \psi(t, u, v) + b^\top \psi(t, u, v) \\ \psi(0, u, v) &= u \\ \partial_t \psi_i(t, u, v) &= \frac{1}{2} \psi(t, u, v)^\top \alpha_i \psi(t, u, v) + \beta_i^\top \psi(t, u, v), \quad 1 \leq i \leq d \\ \psi(0, u, v) &= v. \end{aligned} \quad (1.31)$$

- $\phi$  and  $\psi$  uniquely determined by (1.30) with  $\phi(0, u, v) = u$ ,  $\psi(0, u, v) = v$ .

**Example 1.18 (Ornstein-Uhlenbeck Process).**

$$dX_t = \kappa(\theta - X_t)dt + \sigma dB_t, \quad X_t = x$$

- state space:  $\mathbb{R}^d$
- parameters:  $\kappa \in \mathbb{R}^{d \times d}$ ,  $\theta \in \mathbb{R}^d$ ,  $\sigma \in \mathbb{R}_+^{d \times d}$
- Riccati equations:

$$\begin{aligned} \partial_t \phi(t, u) &= \frac{1}{2} \psi(t, u)^\top \sigma \sigma^\top \psi(t, u) + (\kappa \theta)^\top \psi(t, u) \\ \psi(0, u) &= 0 \\ \partial_t \psi(t, u) &= -\kappa^\top \psi(t, u) \\ \psi(0, u) &= u. \end{aligned}$$

- the unique solution of Riccati equations for  $d = 1$

$$\begin{aligned} \psi(t, u) &= e^{-\kappa t} u \\ \phi(t, u) &= \theta u (1 - e^{-\kappa t}) + \frac{\sigma^2}{4\kappa} (1 - e^{-2\kappa t}). \end{aligned}$$

**Example 1.19** (CIR process).

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dB_t, \quad X_t = x$$

- *state space:*  $\mathbb{R}^d$
- *parameters:*  $\kappa \in \mathbb{R}^{d \times d}$ ,  $\theta \in \mathbb{R}^d$ ,  $\sigma = \text{Diag}(\sigma_1, \dots, \sigma_d) \in \mathbb{R}_+^{d \times d}$
- *Ricatti equations:*

$$\begin{aligned} \partial_t \phi(t, u) &= (\kappa \theta)^\top \psi(t, u) \\ \psi(0, u) &= 0 \\ \partial_t \psi(t, u) &= \frac{1}{2} \psi(t, u)^\top \sigma \sigma^\top \psi(t, u) - \kappa^\top \psi(t, u) \\ \psi(0, u) &= u. \end{aligned}$$

- *the unique solution of Ricatti equations for  $d = 1$*

$$\begin{aligned} \psi(t, u) &= \frac{ue^{-\kappa t}}{1 - \frac{\sigma^2}{2\kappa}u(1 - e^{-\kappa t})} \\ \phi(t, u) &= -\frac{2\kappa\theta}{\sigma^2} \log \left( 1 - \frac{\sigma^2}{2\kappa}u(1 - e^{-\kappa t}) \right) \end{aligned}$$

- *numeric solution for  $d > 1$ .*

## 1.5.2 Pricing in Affine Diffusion Models

Suppose that the factor process  $X$  follows an affine diffusion process of the form (1.25).

Consider a  $T$ -claim with payoff  $H_T = H(X_T, T)$  at time  $T$  so that  $\mathbb{E} \left[ \exp(-\int_0^T r(s)ds) | H(X_T, T) \right]$  is finite. Then its arbitrage free price at time  $t \leq T$  is given by

$$\pi(t) = \mathbb{E}_t^{\mathbb{Q}} \left[ e^{-\int_t^T r(s)ds} H(X_T, T) \right]$$

where  $\mathbb{Q}$  is a risk-neutral probability measure.

In affine diffusion models the **zero-coupon bond price** at time  $t \leq T$  is an exponential affine function of the short rate  $r_t$

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

where  $A(T-t)$  and  $B(T-t)$  solve the Ricatti differential equations [14].



## Chapter 2

# Linear-Rational Term Structure Models

The linear-rational term structure models are a new class of term structure models introduced by Filipović, Larsson and Trolle [8]. They assume that the state variables follow a multivariate factor process with a drift that is affine in the current state and specify the state price density as a positive affine function of the current state. Under these assumptions the zero-coupon bond prices and the short rate become linear-rational functions of the current state, and this is why the models are called linear-rational term structure models.

Specifying the state price density as a positive adaptive process ensures that the linear-rational term structure models are arbitrage-free (see Theorem 1.4) and according to (1.15) the price  $\Pi(t, T)$  at time  $t$  of any time  $T$  cash-flow  $H_T$  is given by

$$\Pi(t, T) = E_t \left[ \frac{\zeta_T}{\zeta_t} H_T \right]. \quad (2.1)$$

The linear-rational term structure models are due to the aforementioned assumptions highly tractable and compared to the affine term structure models has several distinct advantages [8]. They *i*) ensure non-negative interest rates, *ii*) easily accommodate unspanned factors affecting volatility and risk premiums, and *iii*) admit semi-analytical solutions to swaptions.

The affine term structure models can match either *i*) or *ii*), but not simultaneously, and never *iii*). In the affine term structure models non-negative interest rates are guaranteed, only if all factors are of square-root type. Easy accommodation of unspanned stochastic volatility (USV) requires at least one Gaussian factor [8].

In this chapter we define the class of linear-rational term structure models and discuss its advantages. We show also how the interest rate derivatives like swap

and swaption can be valued with this class of models. All theorems and their proofs are taken from [8].

## 2.1 The Linear-Rational Framework

### 2.1.1 Term Structure Specification

A linear-rational term structure model consists of two components:

- a *multivariate factor process*  $X_t$  whose state space is some subset  $E \subset \mathbb{R}^d$ , has a drift that is affine in the current state and is of the form

$$dX_t = \kappa(\theta - X_t)dt + dM_t \quad (2.2)$$

for some  $\kappa \in \mathbb{R}^{d \times d}$ ,  $\theta \in \mathbb{R}^d$  and some martingale  $M_t$  and

- a *state price density*  $\zeta_t$  which is assumed to be an affine function of the current state given by

$$\zeta_t = e^{-\alpha t}(\phi + \psi^\top X_t) \quad (2.3)$$

for some  $\phi \in \mathbb{R}$  and  $\psi \in \mathbb{R}^d$  such that  $\phi + \psi^\top x > 0$  for all  $x \in E$ , and some  $\alpha \in \mathbb{R}$ . Here the role of the parameter  $\alpha$  is to ensure that the short rate stays non-negative.

To specify the term structure it is necessarily the price of zero-coupon bond to be defined. Applying the price formula (2.1) one obtains the following price formula for a zero-coupon bond with a payment  $H_T = 1$  at maturity date  $T$

$$P(t, T) = e^{-\alpha(T-t)} \frac{\phi + \psi^\top E[X_T | \mathcal{F}_t]}{\phi + \psi^\top X_t}. \quad (2.4)$$

According to the lemma bellow, the conditional expectation of factor process with an affine drift is an affine function in the current state. Hence, it follows that the price of zero-coupon bond is a linear-rational function in the current state  $X_t$ .

**Lemma 2.1.** *Assume that  $X_t$  is of the form (2.2) with integrable starting point  $X_0$ . Then for any bounded stopping time  $\rho$  and any deterministic  $\tau \geq 0$ , the random variable  $X_{\rho+\tau}$  is integrable, and we have*

$$E[X_{\rho+\tau} | \mathcal{F}_t] = \theta + e^{-\kappa\tau}(X_\rho - \theta), \quad t \leq T \quad (2.5)$$

*Proof.* I case: We prove first the lemma for  $\rho = 0$ . Itô Lemma to the process

$$Y_t = \theta + e^{-\kappa(\tau-t)}(X_t - \theta)$$

implies that

$$\begin{aligned}
dY_t &= \kappa e^{-\kappa(\tau-t)}(X_t - \theta)dt + e^{-\kappa(\tau-t)}dX_t + 0 \\
&= \kappa e^{-\kappa(\tau-t)}(X_t - \theta)dt + e^{-\kappa(\tau-t)}(\kappa(\theta - X_t)dt + dM_t) \\
&= e^{-\kappa(\tau-t)}dM_t
\end{aligned}$$

Hence  $Y_t$  is a local martingale. We show that  $Y_t$  is even a true martingale. Integration by parts yields

$$Y_t = Y_0 + e^{-\kappa(\tau-t)}M_t - \int_0^t M_s \kappa e^{-\kappa(\tau-s)} ds$$

which implies together the integrability of  $X_0$  and martingale  $M_t$  that  $Y_t$  is integrable. Then it follows from Fubini's theorem that for any  $0 \leq t \leq u$

$$\begin{aligned}
E[Y_u | \mathcal{F}_t] &= Y_0 + e^{-\kappa(\tau-u)}E[M_u | \mathcal{F}_t] - \int_0^u E[M_s | \mathcal{F}_t] \kappa e^{-\kappa(\tau-s)} ds \\
&= Y_0 + e^{-\kappa(\tau-u)}M_t - \int_0^u M_{s \wedge t} \kappa e^{-\kappa(\tau-s)} ds \\
&= Y_0 + e^{-\kappa(\tau-u)}M_t - \int_0^t M_s \kappa e^{-\kappa(\tau-s)} ds - \int_t^u M_t \kappa e^{-\kappa(\tau-s)} ds \\
&= Y_0 + e^{-\kappa(\tau-u)}M_t - \int_0^t M_s \kappa e^{-\kappa(\tau-s)} ds - (e^{-\kappa(\tau-u)} - e^{-\kappa(\tau-t)})M_t \\
&= Y_t
\end{aligned}$$

which proves that  $Y$  is a true martingale. Since  $Y_\tau = X_\tau$  it follows that

$$E[X_\tau | \mathcal{F}_0] = Y_0 = \theta + e^{-\kappa\tau}(X_0 - \theta)$$

II case: If  $\rho$  is a bounded stopping time, then the  $L^1$ -boundedness of  $Y$ , and hence of  $X$ , implies that  $X_\rho$  is integrable. Applying the  $\rho = 0$  case to the process  $(X_{\rho+s})_{s \geq 0}$  and the filtration  $(\mathcal{F}_{\rho+s})_{s \geq 0}$  we obtain

$$E[X_{\rho+\tau} | \mathcal{F}_\rho] = \theta + e^{-\kappa\tau}(X_\rho - \theta).$$

□

By substituting the conditional expectation (2.5) in the pricing formula (2.4) we get explicit formulas for zero-coupon bond prices given by  $P(t, T) = F(\tau, X_t)$ , where  $\tau = T - t$  and  $F$  is a linear-rational function

$$F(\tau, x) = \frac{(\phi + \psi^\top \theta) e^{-\alpha\tau} + \psi^\top e^{-(\alpha+\kappa)\tau}(x - \theta)}{\phi + \psi^\top x}. \quad (2.6)$$

Moreover, it follows from (1.2) that the short rate satisfies  $r_t = -\partial_T \log P(t, T) |_{T=t}$ . The simple calculations show that the short rate is given by the linear-rational function in the current state variable,

$$r_t = \alpha - \frac{\psi^\top \kappa(\theta - X_t)}{\phi + \psi^\top X_t}. \quad (2.7)$$

We see from the formula above that if the parameter  $\alpha$  is chosen large enough, the short rate is non-negative at any time  $t$ . One possible choice of  $\alpha$  is the smallest value that yields a non-negative short rate. If one defines

$$\alpha^* = \sup_{x \in E} \frac{\psi^\top \kappa(\theta - x)}{\phi + \psi^\top x} \quad \text{and} \quad \alpha_* = \inf_{x \in E} \frac{\psi^\top \kappa(\theta - x)}{\phi + \psi^\top x} \quad (2.8)$$

and sets  $\alpha = \alpha^*$  then the short rate lies in the interval

$$r_t \in [0, \alpha^* - \alpha_*] \quad (r_t \in [0, \infty] \text{ if } \alpha_* = -\infty)$$

and  $\alpha_*$  and  $\alpha^*$  depend on the model parameters, which have to be calibrated.

The parameter  $\alpha$  can also be interpreted as an infinite-maturity spot rate, since if the eigenvalues of  $\kappa$  have a non-negative real part, the equality

$$\begin{aligned} & \lim_{\tau \rightarrow \infty} -\frac{1}{\tau} \log F(\tau, x) \\ & \stackrel{(2.6)}{=} \lim_{\tau \rightarrow \infty} -\frac{1}{\tau} \log \frac{e^{-\alpha\tau} ((\phi + \psi^\top \theta) + \psi^\top e^{-\kappa\tau} (x - \theta))}{\phi + \psi^\top x} \\ & = \lim_{\tau \rightarrow \infty} \left( \underbrace{\alpha - \frac{1}{\tau} \log((\phi + \psi^\top \theta) + \underbrace{\psi^\top e^{-\kappa\tau} (x - \theta)}_{\rightarrow 0})}_{\rightarrow 0} + \underbrace{\frac{1}{\tau} \log(\phi + \psi^\top x)}_{\rightarrow 0} \right) \\ & = \alpha \end{aligned}$$

is valid for any  $x \in E$ .

## 2.1.2 Unspanned Factors

In term structure modelling, the factors affecting the term structure are known as *term structure factors*. Otherwise, they are called *unspanned factors*. Since the term structure is specified by the zero-coupon bond prices, the unspanned factors do not affect the zero-coupon bond prices, but they can affect the prices of other interest rate derivatives. The zero-coupon bond prices play a key role by determining if a model has unspanned factors or not.

The existence of unspanned factors in linear-rational term structure models can be verified easily by using only the model parameter. The authors of the linear-rational term structure model use the concept of kernel to describe the directions  $\xi \in \mathbb{R}^d$  such that the zero-coupon bond prices remain unchanged when the state vector moves along  $\xi$ . The presence of the non-zero elements in the kernel means that there are factors in the model which do not affect the term structure. Conversely, if the kernel consists only of the zero element of  $\mathbb{R}^d$ , then all factors affect the term structure, and there are no unspanned factors in the model.

In linear algebra the kernel of a linear map is the null space of the map and measures the degree of injectivity of this map. In linear-rational term structure models the term structure is specified by the linear-rational function  $F(\tau, \cdot)$ , given in (2.6), which means non-linear in  $X_t$  and therefore the usual kernel can not be used. For this reason the authors introduce the term kernel of the differentiable function and then define the term structure kernel as a kernel of the differentiable function  $F(\tau, \cdot)$  for  $\tau \geq 0$ .

**Definition 2.2 (Kernel of a differentiable function).** *The kernel of a differentiable function  $f$  on  $E \subset \mathbb{R}^d$  is the linear subspace of  $\mathbb{R}^d$  defined by*

$$\ker f = \{ \xi \in \mathbb{R}^d : \nabla f(x)^\top \xi = 0 \text{ for all } x \in E \},$$

where  $\nabla f(x)$  denote the gradient with respect to  $x$ . If  $f(x) = \nu^\top x$  is linear for some  $\nu \in \mathbb{R}^d$ , then  $\nabla f(x) = \nu$  for all  $x \in E$  and the new definition gives the usual kernel of  $f$ .

**Definition 2.3 (Term structure kernel).** *The term structure kernel is defined by*

$$\mathcal{U} = \bigcap_{\tau \geq 0} \ker F(\tau, \cdot).$$

An alternative representation of the term structure kernel in terms of the model parameters is given in the following proposition, in which the term structure kernel is presented as a kernel of a linear map.

**Proposition 2.4.** *Assume that the term structure is not trivial, i.e. the short rate is not constant. Then*

$$\mathcal{U} = \bigcap_{p=0}^{d-1} \ker \psi^\top \kappa^p. \quad (2.9)$$

*Proof:* According to Def. 2.2 the kernel is a set of elements orthogonal to the gradient  $\nabla F(\tau, x)$  for  $\tau \geq 0$ ,  $x \in E$  and hence the orthogonal complement of term structure kernel is  $\mathcal{U}^\perp = \text{span} \{ \nabla F(\tau, x) : \tau \geq 0, x \in E \}$ . To prove (2.9) we have

to prove that  $\mathcal{U}^\perp = \text{span} \{(\kappa^\top)^p \psi : p = 0, \dots, d-1\}$ . By the Cayley-Hamilton Theorem (see Horn and Johnson (1990 Theorem 2.4.3))  $(\kappa^\top)^d$  can be expressed as a linear combination of  $(\kappa^\top)^0, \dots, (\kappa^\top)^{d-1}$ . In this case we have to prove

$$\text{span} \{\nabla F(\tau, x) : \tau \geq 0, x \in E\} = \text{span} \{(\kappa^\top)^p \psi : p \geq 0\} \quad (2.10)$$

Let  $\mathcal{S}$  denote the left side. After calculation, we obtain for the gradient of  $F$

$$\nabla F(\tau, x) = \frac{e^{-\alpha\tau}}{\phi + \psi^\top x} \left[ e^{-\kappa^\top \tau} \psi - e^{-\alpha\tau} F(\tau, x) \psi \right] \quad (2.11)$$

from which  $\mathcal{S} = \text{span} \left\{ e^{-\kappa^\top \tau} \psi - e^{-\alpha\tau} F(\tau, x) \psi : \tau \geq 0, x \in E \right\}$ . We assume that there are  $x, y \in E$  and  $\tau \geq 0$  such that  $F(\tau, x) \neq F(\tau, y)$ . It follows that  $e^{-\alpha\tau} (F(\tau, x) \psi - F(\tau, y) \psi)$  and hence  $\psi$  itself, lies in  $\mathcal{S}$ . Therefore  $\mathcal{S} = \text{span} \left\{ e^{-\kappa^\top \tau} \psi : \tau \geq 0 \right\}$ .

Since  $e^{-\kappa^\top \tau} \psi = \sum_{p \geq 0} \frac{(-\tau)^p}{p!} (\kappa^\top)^p \psi$  we can conclude that  $\mathcal{S} = \text{span} \{(\kappa^\top)^p \psi : p \geq 0\}$ . □

As a corollary one obtains a necessary and sufficient condition for a linear-rational model to have only term structure factors.

**Corollary 2.5.** *Assume that  $\kappa$  is diagonalizable with real eigenvalues, i.e.  $\kappa = S^{-1} \Lambda S$  with  $\Lambda$  diagonal and real. Then  $\mathcal{U} = \{0\}$  if and only if all eigenvalues of  $\kappa$  are distinct and all components of  $S^{-\top} \psi$  are non zero.*

*Proof:* Let  $\Lambda = \text{Diag}(\lambda_1, \dots, \lambda_d)$  and consider the matrix

$$A = \begin{bmatrix} \psi & \kappa^\top \psi & \dots & (\kappa^\top)^{d-1} \psi \end{bmatrix}.$$

According to Proposition 2.4,  $\mathcal{U}$  consists of elements  $\xi \in \mathbb{R}^d$  such that  $\psi^\top \kappa^p \xi = 0$  for all  $p = 0, \dots, d-1$ , which is equivalent to  $A^\top \xi = 0$ . Then  $\mathcal{U} = \{0\}$  if and only if the determinant of the matrix  $A$  is non-zero.

Let  $\hat{\psi} = S^{-\top} \psi$ . The determinant of  $A$  is then given by

$$\begin{aligned} \det A &= \det(S^\top) \det(\hat{\psi} \quad A^\top \hat{\psi} \quad \dots \quad (\Lambda^\top)^{d-1} \hat{\psi}) \\ &= \det(S^\top) \hat{\psi}_1 \cdots \hat{\psi}_d \det \begin{pmatrix} 1 & \lambda_1 & \dots & \lambda_1^{d-1} \\ \vdots & \vdots & & \vdots \\ 1 & \lambda_d & \dots & \lambda_d^{d-1} \end{pmatrix} \\ &= \det(S^\top) \hat{\psi}_1 \cdots \hat{\psi}_d \prod_{1 \leq i < j \leq d} (\lambda_i - \lambda_j), \end{aligned}$$

The last equation follows from the formula for calculating of Vandermonde matrix. We see that the determinant of  $A$  is not zero if and only if the eigenvalues of  $\kappa$  are distinct and all components of  $\hat{\psi}$  are non-zero.

□

The term structure factors alone are not enough to describe modern finance markets. Therefore we concentrate on linear-rational models with unspanned factors. For convenience the state space of the factor process is transformed in such a way that the unspanned directions correspond to the last components of the state vector.

Assume now that the linear-rational term structure model has  $n \leq d$  unspanned factors and  $m = d - n$  term structure factors. We also assume that  $S$  is an invertible linear transformation on  $\mathbb{R}^d$ , which maps the term structure kernel into

$$S(\mathcal{U}) = \{0\} \times \mathbb{R}^n,$$

where  $n = \dim \mathcal{U}$  and  $m = d - n$ . This transformation defines a new linear-rational term structure model equivalent to the original model with

- a factor process  $d\hat{X}_t = SX_t$  whose dynamics is given by

$$d\hat{X}_t = \hat{\kappa}(\hat{\theta} - \hat{X}_t)dt + d\hat{M}_t,$$

where

$$\hat{\kappa} = SkS^{-1}, \quad \hat{\theta} = S\theta, \quad \hat{M}_t = SM_t \quad (2.12)$$

- a state price process

$$\xi_t = e^{-\alpha t}(\hat{\phi} + \hat{\psi}^\top \hat{X}_t)$$

where

$$\hat{\phi} = \phi, \quad \hat{\psi} = S^{-\top}\psi. \quad (2.13)$$

The transformed model has exactly  $n$  unspanned factors, corresponding to the last  $n$  components of the transformed factor process  $\hat{X}_t$  under certain technical conditions, given in the theorem below. In this case the transformed factor process can be decomposed into  $\hat{X}_t = (Z_t, U_t)$ , where  $Z_t$  is a  $m$ -dimensional factor process affecting the term structure and  $U_t$  is a  $n$ -dimensional factor process describing the unspanned factors.

**Theorem 2.6.** *Let  $m, n \geq 0$  be integers with  $m + n = d$ . Then*

$$S(\mathcal{U}) = \{0\} \times \mathbb{R}^n \quad (2.14)$$

*if and only if the transformed model parameters (2.12) - (2.13) satisfy:*

- (i)  $\hat{\psi} = (\hat{\psi}_Z, 0)^\top \in \mathbb{R}^m \times \mathbb{R}^n$ ;

(ii)  $\hat{\kappa}$  has block lower triangular structure,

$$\hat{\kappa} = \begin{pmatrix} \hat{\kappa}_{ZZ} & 0 \\ \hat{\kappa}_{UZ} & \hat{\kappa}_{UU} \end{pmatrix} \in \mathbb{R}^{(m+n) \times (m+n)};$$

(iii) The upper left block  $\hat{\kappa}_{ZZ}$  of  $\hat{\kappa}$  satisfies

$$\bigcap_{p=0}^{m-1} \ker \hat{\psi}_Z^\top \hat{\kappa}_{ZZ}^p = \{0\}.$$

In this case, the dimension of the term structure kernel  $\mathcal{U}$  equals  $n$ .

*Proof:* The proof of the case  $n = 0$  follows immediately from Prop. 2.4. If  $n = 0$  then  $m = d$ ,  $\hat{\psi} = \hat{\psi}_Z$ ,  $\hat{\kappa} = \hat{\kappa}_{ZZ}$  and according to Prop. 2.4  $\mathcal{U} = \bigcap_{p=0}^{m-1} \ker \hat{\psi}_Z^\top \hat{\kappa}_{ZZ}^p$ . Consider now the case  $n \geq 1$ .

$\Rightarrow$ : We denote  $\hat{\mathcal{U}} = S(\mathcal{U})$  and assume that (2.14) holds, i.e.  $\hat{\mathcal{U}} = \{0\} \times \mathbb{R}^n$ . This means that

$$\hat{\mathcal{U}} = \left\{ \hat{\xi} \in \mathbb{R}^d : \hat{\xi} = (0, \hat{\xi}_U)^\top, \text{ where } \hat{\xi}_U \in \mathbb{R}^n \right\} \subset \mathbb{R}^m \times \mathbb{R}^n \quad (2.15)$$

On the other hand, it follows from the Prop. 2.4 and Def. 2.2 that

$$\hat{\mathcal{U}} = \left\{ \hat{\xi} \in \mathbb{R}^d : \hat{\psi}^\top \hat{\kappa}^p \hat{\xi} = 0 \text{ for } p = 0, 1, \dots, d-1 \right\} \quad (2.16)$$

We partition  $\hat{\psi}$  and  $\hat{\kappa}$  so that

$$\hat{\psi} = \begin{pmatrix} \hat{\psi}_Z \\ \hat{\psi}_U \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^n, \quad \hat{\kappa} = \begin{pmatrix} \hat{\kappa}_{ZZ} & \hat{\kappa}_{ZU} \\ \hat{\kappa}_{UZ} & \hat{\kappa}_{UU} \end{pmatrix} \in \mathbb{R}^{(m+n) \times (m+n)}.$$

It follows from (2.15) and (2.16) that  $\hat{\psi}_U^\top \hat{\xi}_U = 0$  for  $\hat{\xi}_U \in \mathbb{R}^n$ . Hence  $\hat{\psi}_U^\top = 0$ , which proves (i). We show that

$$\hat{\psi}^\top \hat{\kappa}^p = \begin{pmatrix} \hat{\psi}_Z^\top \hat{\kappa}_{ZZ}^p & \hat{\psi}_Z^\top \hat{\kappa}_{ZZ}^{p-1} \hat{\kappa}_{ZU} \end{pmatrix} \quad (2.17)$$

For  $p = 1$  follows immediately from (i). Suppose that (2.17) holds for some  $p \geq 1$ . Then (2.15) and (2.16) imply

$$0 = \hat{\psi}^\top \hat{\kappa}^p \begin{pmatrix} 0 \\ \hat{\xi}_U \end{pmatrix} = \hat{\psi}_Z^\top \hat{\kappa}_{ZZ}^{p-1} \hat{\kappa}_{ZU} \hat{\xi}_U,$$

for any  $\hat{\xi}_U \in \mathbb{R}^n$  and hence

$$\hat{\psi}_Z^\top \hat{\kappa}_{ZZ}^{p-1} \hat{\kappa}_{ZU} = 0. \quad (2.18)$$



By multiplying both sides of (2.17) by  $\hat{\kappa}$  from the right we see that (2.17) holds for  $p + 1$ . It follows by induction that (2.17) and (2.18) holds for all  $p \geq 1$ .

Now we take any  $\hat{\xi}_Z$  such that  $\hat{\psi}_Z^\top \hat{\kappa}_{ZZ}^p \hat{\xi}_Z = 0$  for all  $p \geq 0$ . Then the vector  $\hat{\xi} = (\hat{\xi}_Z \quad 0)^\top \in \mathbb{R}^m \times \mathbb{R}^n$  satisfies  $\hat{\psi}^\top \hat{\kappa}^p \hat{\xi} = 0$  for all  $p \geq 0$ . Hence  $\hat{\xi} \in \hat{\mathcal{U}}$  by (2.16), and then  $\hat{\xi}_Z = 0$  by (2.15). This proves (iii).

Finally, since (2.18) holds for all  $p \geq 1$ , the range of  $\hat{\kappa}_{ZU}$  lies in the kernel of  $\hat{\psi}_Z \hat{\kappa}_{ZZ}^{p-1}$  for all  $p \geq 1$ . This implies by (iii) that the range of  $\hat{\kappa}_{ZU}$  consists of the zero vector and hence  $\hat{\kappa}_{ZU} = 0$ . This proves (ii).

$\Leftarrow$ : Conversely, we now assume that  $\hat{\psi}$  and  $\hat{\kappa}$  given in (2.12) and (2.13) satisfy the condition (i) – (iii). We show that (i) – (ii) imply  $S(\mathcal{U}) \supset \{0\} \times \mathbb{R}^n$ . Let  $\{e_1, e_2, \dots, e_d\}$  the canonical basis of  $\mathbb{R}^d$ , then we have for  $p = 0, \dots, d - 1$ ,

$$\psi^\top \kappa^p S^{-1} e_i = \psi^\top S^{-1} (S \kappa S^{-1})^p e_i = \hat{\psi}^\top \hat{\kappa}^p e_i$$

Note that  $\hat{\kappa}^p$  has the same block triangular structure as  $\hat{\kappa}$  for all  $p \geq 1$ . Thus for  $i = m + 1, \dots, d$ , the right side above is zero for all  $p \geq 0$ . Therefore  $S^{-1} e_i \in \mathcal{U}$  by (2.16) and hence  $e_i \in S(\mathcal{U})$  for  $i = m + 1, \dots, d$ . Since  $\{0\} \times \mathbb{R}^n = \text{span}(e_{m+1}, \dots, e_d)$  and according to the definition of span, is the smallest set containing the vectors  $e_{m+1}, \dots, e_d$ , it follows that  $\{0\} \times \mathbb{R}^n \subset S(\mathcal{U})$ . Suppose now that (iii) holds additionally, i.e.

$$\bigcap_{p=0}^{m-1} \ker \hat{\psi}_Z^\top \hat{\kappa}_{ZZ}^p = \left\{ \xi^* \in \mathbb{R}^m : \hat{\psi}_Z^\top \hat{\kappa}_{ZZ}^p \xi^* = 0 \text{ for all } p = 0, \dots, m - 1 \right\} = \{0\}$$

and consider a vector  $\xi \in \text{span}(S^{-1} e_1, \dots, S^{-1} e_m)$ . Then  $S\xi = (\hat{\xi}_Z, 0) \in \mathbb{R}^m \times \mathbb{R}^n$ . By (i) and (ii) we have

$$\psi^\top \kappa^p \xi = \hat{\psi}^\top \hat{\kappa}^p S\xi = \hat{\psi}_Z^\top \hat{\kappa}_{ZZ}^p \hat{\xi}_Z.$$

If  $\xi \in \mathcal{U}$ , the left side is zero for all  $p = 0, \dots, d - 1$ . Hence so is the right side, which by (iii) implies  $\hat{\xi}_Z = 0$  and therefore also  $\xi = 0$ . This means that  $\text{span}(S^{-1} e_1, \dots, S^{-1} e_m) = \{0\}$ . We deduce that

$$\mathcal{U} = \text{span}(S^{-1} e_{m+1}, \dots, S^{-1} e_d)$$

and hence

$$S(\mathcal{U}) = \text{span}(e_{m+1}, \dots, e_d) = \{0\} \times \mathbb{R}^n$$

□

The proof of Theorem (2.6) shows that the first two conditions (i) – (ii) of the theorem guarantee the existence of at least  $n$  unspanned factors in the linear rational model. The model has exactly  $n$  unspanned factors if in addition the condition (iii) is satisfied.

**Lemma 2.7.** *Let  $m, n \geq 0$  be integers with  $m + n = d$ . If the transformed model parameters (2.12)-(2.13) satisfy (i) – (ii) in Theorem 2.6, we have  $S(\mathcal{U}) \supset \{0\} \times \mathbb{R}^n$ . In this case  $\dim \mathcal{U} \geq n$ .*

Theorem 2.6 implies that if (2.14) holds then the zero-coupon bond prices depend only on the components of vector  $Z_t$  and not on the components of the vector  $U_t$  where  $\hat{X}_t = (Z_t, U_t)$ . Let  $Sx = (z, u)$  and  $\hat{\theta} = (\hat{\theta}_Z, \hat{\theta}_U)$ . According to Theorem 2.6, (2.14) holds if and only if the parameters of the transformed model satisfy (i) – (iii). Then the function

$$\begin{aligned} \hat{F}(\tau, z) = F(\tau, x) &= \frac{(\phi + \psi^\top \theta)e^{-\alpha\tau} + \psi^\top e^{-(\alpha+\kappa)\tau}(x - \theta)}{\phi + \psi^\top x} \\ &= \frac{(\phi + \psi^\top S^{-1}S\theta)e^{-\alpha\tau} + \psi^\top S^{-1}Se^{-(\alpha+\kappa)\tau}S^{-1}S(x - \theta)}{\phi + \psi^\top S^{-1}Sx} \\ &= \frac{(\hat{\phi} + \hat{\psi}_Z^\top \hat{\theta}_Z)e^{-\alpha\tau} + \hat{\psi}_Z^\top e^{-(\alpha+\hat{\kappa}_{ZZ})\tau}(z - \hat{\theta}_Z)}{\hat{\phi} + \hat{\psi}_Z^\top z} \end{aligned}$$

does not depend on  $u$  and the bond prices are given by  $P(t, T) = \hat{F}(T - t, Z_t)$ . Therefore the components of  $U_t$  do not change the zero-coupon bond prices. Just the opposite, the vector  $Z_t$  affects the zero-coupon bond prices. Its components are the term structure factors. Moreover, if the term structure function  $z \mapsto (\hat{F}(\tau, z))_{\tau \geq 0}$  is injective the vector  $Z_t$  can be reconstructed from a snapshot of the term structure at the time  $t$ . A necessary and sufficient condition for injectivity of this function is given in the following proposition.

**Proposition 2.8.** *The term structure function  $z \mapsto (\hat{F}(\tau, z))_{\tau \geq 0}$  is injective if and only if the  $\hat{\kappa}_{ZZ}$  is invertible and  $\hat{\phi} + \hat{\psi}_Z^\top \hat{\theta}_Z \neq 0$ .*

### 2.1.3 Unspanned stochastic volatility factors

Under the condition (i) – (iii) the dynamics of  $\hat{X}_t = (Z_t, U_t)$  can be decomposed into term-structure dynamics

$$dZ_t = \hat{\kappa}_{ZZ}(\hat{\theta}_Z - Z_t)dt + d\hat{M}_{Zt} \quad (2.19)$$

and unspanned factor dynamics

$$dU_t = (\hat{\kappa}_{UZ}(\hat{\theta}_Z - Z_t) + \hat{\kappa}_{UU}(\hat{\theta}_U - U_t))dt + d\hat{M}_{Ut} \quad (2.20)$$

where  $\hat{M}_t = (\hat{M}_{Zt}, \hat{M}_{Ut})$ . The state price density is affine in  $Z_t$

$$\zeta_t = e^{-\alpha t}(\hat{\phi} + \hat{\psi}_Z^\top Z_t) \quad (2.21)$$

If the unspanned factors  $U_t$  do not enter the dynamics of  $\hat{M}_{Zt}$ , then  $Z_t$  is a fully autonomous process and the unspanned factors play no role in the model. In this case the  $m$ -dimensional linear-rational term structure model (2.19) – (2.20) is equivalent to (2.2) – (2.3), has only term structure factors and is minimal. If there are some unspanned factors that enter into the dynamics of  $\hat{M}_{Zt}$ , then they can be seen as unspanned factors affecting the stochastic volatility. They are called *unspanned stochastic volatility (USV) factors*.

## 2.1.4 Swaption Price

An important advantage of the linear-rational term structure models is that they allow a semi-analytical formula for the swaption price. A *swaption* is an option on an interest rate swap that gives the right, but not the obligation to enter into the swap.

Consider a swap that pays fixed versus floating interest rates specified by a tenor structure of equidistant reset and payment dates  $T_0 < T_1 < \dots < T_n$  with  $\Delta = T_i - T_{i-1}$  and a pre-determined annualized rate  $K$ . At each date  $T_i$  the fixed leg pays  $\Delta K$  and the floating leg pays LIBOR for the preceding time period  $[T_{i-1}, T_i]$  and at the same time the floating rate for the next period is set. According to (1.7) the value of the swap at any time  $t \leq T_0$ , from the perspective of the fixed-rate payer, is given by

$$\Pi_t^{swap} = P(t, T_0) - P(t, T_n) - \Delta K \sum_{i=1}^n P(t, T_i). \quad (2.22)$$

A European payer swaption expiring at time  $T_0$  enters the swap specified above only if the swap value is positive at time  $T_0$ . Its value at expiration is

$$H_{T_0} = \left( \Pi_{T_0}^{swap} \right)^+ = \left( \sum_{i=0}^n c_i P(T_0, T_i) \right)^+ = \frac{1}{\zeta_{T_0}} \left( \sum_{i=0}^n c_i \mathbb{E}[\zeta_{T_i} | \mathcal{F}_{T_0}] \right)^+,$$

where  $c_0 = 1$ ,  $c_n = -1 - \Delta K$  and  $c_i = -\Delta K$  for  $i = 1, \dots, n-1$ . The last equality follows from (1.15).

According to Lemma 2.1 the conditional expectations  $\mathbb{E}[\zeta_{T_i} | \mathcal{F}_{T_0}]$  in a linear-rational term structure model are affine functions in  $X_{T_0}$ . Substituting the expectation in the formula above we obtain for the swaption value

$$H_{T_0} = \frac{1}{\zeta_{T_0}} \left( p_{swap}(X_{T_0}) \right)^+,$$

where  $p_{swap}$  is the explicit affine function

$$p_{swap}(x) = \sum_{i=0}^n c_i e^{-\alpha T_i} \left( \phi + \psi^\top \theta + \psi^\top e^{-\kappa(T_i - T_0)} (x - \theta) \right).$$

Applying the pricing formula (2.1) we obtain that the price of a European payer swaption at time  $t \leq T_0$  is given by

$$\Pi_t^{swpt} = \frac{1}{\zeta_t} \mathbb{E}[\zeta_{T_0} C_{T_0} | \mathcal{F}_t] = \frac{1}{\zeta_t} \mathbb{E} \left[ \left( p_{swap}(X_{T_0}) \right)^+ | \mathcal{F}_t \right]. \quad (2.23)$$

To compute the swaption price with this formula we need the conditional expectation  $E[X_{T_0} | \mathcal{F}_t]$ . If the conditional density of  $X_{T_0}$  given  $\mathcal{F}_t$  is known then the conditional expectation can be computed by numerical integration over  $\mathbb{R}^d$ . The authors of the model propose an alternative approach to evaluating the condition expectation, based on the Fourier transform methods.

**Theorem 2.9.** *Define  $\hat{q}(z) = E[\exp(z p_{swpt}(X_{T_0})) | \mathcal{F}_t]$  for every  $z \in \mathbb{C}$  such that the conditional expectation is well-defined. Pick any  $\mu > 0$  such that  $\hat{q}(\mu) < \infty$ . Then the swaption price is given by*

$$\Pi_t^{swpt} = \frac{1}{\zeta_t \pi} \int_0^\infty \operatorname{Re} \left[ \frac{\hat{q}(\mu + i\lambda)}{(\mu + i\lambda)^2} \right] d\lambda \quad (2.24)$$

*Proof:* The proof uses that the function  $f(s) = s^+$  can be presented by its inverse Fourier transform as follow

$$s^+ = \frac{1}{2\pi} \int_{\mathbb{R}} e^{(\mu+i\lambda)s} \frac{1}{(\mu+i\lambda)^2} d\lambda \quad (2.25)$$

for any  $\mu > 0$  and  $s \in \mathbb{R}$  (see Bateman and Erdélyi (1954)).

Let  $q(ds)$  denote the conditional distribution of the random variable  $p_{swap}(X_{T_0})$  given  $\mathcal{F}_t$ . Its Fourier transform is then defined as

$$\hat{q}(z) = E[\exp(z p_{swpt}(X_{T_0})) | \mathcal{F}_t] = \int_{\mathbb{R}} e^{zs} q(ds)$$

for every  $z \in \mathbb{C}$  such that the right side is well-defined and finite. Pick  $\mu > 0$  such that  $\int_{\mathbb{R}} e^{\mu s} q(ds) < \infty$ . We show now that the function  $e^{(\mu+i\lambda)s} \frac{1}{(\mu+i\lambda)^2}$  is  $d\lambda \otimes q(ds)$ -integrable and Fubini's theorem can be applied.

$$\begin{aligned} \int_{\mathbb{R}^2} \left| e^{(\mu+i\lambda)s} \frac{1}{(\mu+i\lambda)^2} \right| d\lambda \otimes q(ds) &= \int_{\mathbb{R}^2} e^{\mu s} \frac{1}{\mu^2 + \lambda^2} d\lambda \otimes q(ds) \\ &= \int_{\mathbb{R}} e^{\mu s} q(ds) \int_{\mathbb{R}} \frac{1}{\mu^2 + \lambda^2} d\lambda < \infty \end{aligned}$$

The second equality follows from Tonelli's theorem. The conditional expectation of  $p_{swap}(X_{T_0})$  given  $\mathcal{F}_t$  can be calculated as follow

$$\begin{aligned}
E [p_{swpt}(X_{T_0}) | \mathcal{F}_t] &= \int_{\mathbb{R}} s^+ q(ds) \\
&= \int_{\mathbb{R}} \left( \frac{1}{2\pi} \int_{\mathbb{R}} e^{(\mu+i\lambda)s} \frac{1}{(\mu+i\lambda)^2} d\lambda \right) q(ds) \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{1}{(\mu+i\lambda)^2} \left( \int_{\mathbb{R}} e^{(\mu+i\lambda)s} q(ds) \right) d\lambda \\
&= \frac{1}{2\pi} \int_{\mathbb{R}} \frac{\widehat{q}(\mu+i\lambda)}{(\mu+i\lambda)^2} d\lambda \\
&= \frac{1}{\pi} \int_0^\infty \operatorname{Re} \left[ \frac{\widehat{q}(\mu+i\lambda)}{(\mu+i\lambda)^2} \right] d\lambda
\end{aligned}$$

Here the third equality follows from Fubini's theorem. The last equality uses that the left, and hence the right, side is real, and that the real part of  $(\mu+i\lambda)^{-2} \widehat{q}(\mu+i\lambda)$  is an even function of  $\lambda$ . Substituting the last expression of the conditional expectation in (2.23) gives the price formula from theorem 2.9.

□

Theorem 2.9 simplifies the problem of computing an integral over  $\mathbb{R}^d$  to computing a simple line integral. The difficulty of the Fourier method for price calculation is to evaluate  $\widehat{q}(\mu+i\lambda)$  as  $\lambda$  varies through  $\mathbb{R}_+$ . If the factor process is affine, the Exponential-affine transform formula can be applied and the calculation of  $\widehat{q}(z)$  amounts to a numerical solving a system of ordinary differential equations.

## 2.2 Models with Unspanned Stochastic Volatility (USV) Factors

The **unspanned stochastic volatility (USV)** factors are unspanned factors affecting stochastic volatility of the factor process. They also affect the volatility of the zero-coupon bond price, although as unspanned factors they do not affect the price of zero-coupon bond. The unspanned factors, which affect neither the zero-coupon bond price volatility nor the covariance matrix, are called **residual factors**. Therefore, the volatility and covariance matrix of zero-bond price can be used to determine if an unspanned factor in the model is a USV factor or a residual factor.

In this section we consider linear-rational diffusion models of Filipović in which the unspanned factors are presented and give conditions under which the unspanned factors are USV factors as well. Moreover, we show which conditions the

model should satisfy, so that each unspanned factor in the model is also a USV factor. Finally, we show how a  $d$ - dimensional linear-rational square root model of Filipović et al. is specified, so that it has exactly  $m$  term structure factors and  $n = d - m$  USV factors.

## 2.2.1 Linear-Rational Diffusion Models

The linear-rational diffusion model is a linear-rational term structure model whose *factor process* follows an affine diffusion process of the form

$$dX_t = \kappa(\theta - X_t)dt + \sigma(X_t)dB_t \quad (2.26)$$

where  $\sigma : E \rightarrow \mathbb{R}^{d \times d}$  is measurable,  $B_t$  is  $d$ -dimensional Brownian motion and  $a(x) = \sigma(x)\sigma(x)^\top$  is differentiable. The *state price density* in the linear-rational diffusion model is an affine process of the form

$$\zeta_t = e^{-\alpha t}(\phi + \psi^\top X_t).$$

It follows by Itô formula that  $\zeta_t$  solves the following SDE

$$\begin{aligned} d\zeta_t &= -\alpha e^{-\alpha t}(\phi + \psi^\top X_t)dt + e^{-\alpha t}\psi^\top dX_t + \frac{1}{2} \cdot 0 \cdot (dX_t)^2 \\ &= -\alpha\zeta_t + e^{-\alpha t}\psi^\top(\kappa(\theta - X_t)dt + \sigma(X_t)dB_t) \\ &= \zeta_t \left( -\alpha dt + \frac{\psi^\top(\kappa(\theta - X_t))}{\phi + \psi^\top X_t} dt + \frac{\psi^\top \sigma(X_t)}{\phi + \psi^\top X_t} dB_t \right) \\ &= \zeta_t \left( -r_t dt + \frac{\psi^\top \sigma(X_t)}{\phi + \psi^\top X_t} dB_t \right) \\ &= \zeta_t (-r_t dt - \lambda_t dB_t), \end{aligned}$$

where  $r_t = \alpha - \frac{\psi^\top(\kappa(\theta - X_t))}{\phi + \psi^\top X_t}$  is the short rate given in (2.7) and according (1.23)  $\lambda_t = -\frac{\psi^\top \sigma(X_t)}{\phi + \psi^\top X_t}$  is the market price of risk.

In order for us to define the zero bond price volatility, we need the dynamics of the zero-coupon bond price. The zero-coupon bond price in the linear-rational diffusion model is an affine function of the factor process,  $P(t, T) = F(\tau, X_t)$ , where  $\tau = T - t$  and  $F$  is the price function (2.6). By Itô formula and (2.26) one obtains for the price dynamics

$$\begin{aligned} dP(t, T) &= -\frac{\partial F}{\partial t}(\tau, X_t)dt + \nabla F(\tau, X_t)^\top dX_t + \frac{1}{2} \sum_{i,j=0}^d \frac{\partial^2 F}{\partial x_i \partial x_j}(\tau, X_t) dX_t^i dX_t^j \\ &= P(t, T) \left( \left\{ r_t + \frac{\nabla F(\tau, x)^\top \sigma(X_t)}{F(\tau, X_t)} \lambda_t \right\} dt + \frac{\nabla F(\tau, x)^\top \sigma(X_t)}{F(\tau, X_t)} dB_t \right) \\ &= P(t, T) (\{r_t + v(t, T)^\top \lambda_t\} dt + v(t, T)^\top dB_t). \end{aligned}$$

Hence the volatility vector is given by

$$v(t, T) = \frac{\sigma(X_t)^\top \nabla F(\tau, x)}{F(\tau, X_t)}$$

and the instantaneous covariance between the returns on two bonds with maturities  $T_1$  and  $T_2$  is given by  $v(t, T_1)^\top v(t, T_2) = G(\tau_1, \tau_2, X_t)$ , where  $\tau_1 = T_1 - t, \tau_2 = T_2 - t$  and

$$G(\tau_1, \tau_2, x) = \frac{\nabla F(\tau_1, x)^\top a(x) \nabla F(\tau_2, x)}{F(\tau_1, x) F(\tau_2, x)}$$

with  $a(x) = \sigma(x)\sigma(x)^\top$ .

The variance-covariance kernel describes the directions in  $\xi \in \mathbb{R}^d$  with the property that movements of the state vector along  $\xi$  change neither the bond return volatility nor the covariation between returns on bonds with different maturities.

**Definition 2.10.** *The variance-covariance kernel, denoted by  $\mathcal{W}$ , is given by*

$$\mathcal{W} = \bigcap_{\tau_1, \tau_2 \geq 0} \ker G(\tau_1, \tau_2, \cdot).$$

That means,  $\mathcal{W}$  consists of all  $\xi \in \mathbb{R}^d$  such that  $\nabla G(\tau_1, \tau_2, x)^\top \xi = 0$  for all  $\tau_1, \tau_2 \geq 0$  and all  $x \in E$ . The linear-rational diffusion model has at least one USV factor if there is at least one element of the term structure kernel  $\mathcal{U}$  that does not lie in the variance-covariance kernel  $\mathcal{W}$ , i.e.  $\mathcal{U} \setminus \mathcal{W} \neq \emptyset$ . In the model all unspanned factors are USV factors if there is no element of the term structure kernel except the null element that lies in the variance-covariance kernel, i.e.  $\mathcal{U} \cap \mathcal{W} = \{0\}$ .

We assume now that there exists an invertible transformation  $S$  that satisfies (2.14) and additionally has the property

$$S(\mathcal{U} \cap \mathcal{W}) = \{0\} \times \{0\} \times \mathbb{R}^q,$$

where  $q = \dim \mathcal{U} \cap \mathcal{W}$ ,  $p + q = n$  and  $n = \dim \mathcal{U}$ .  $S$  transforms the state space so that the intersection  $\mathcal{U} \cap \mathcal{W}$  of the term structure kernel and variance-covariance kernel corresponds to the last components of the state vector. Therefore the unspanned factors  $U_t$  in the transformed factor process  $\hat{X}_t = (Z_t, U_t)$  decompose into  $U_t = (V_t, W_t)$ , where  $V_t$  is a vector of unspanned stochastic volatility factors and  $W_t$  is a vector of residual factors that affect neither the term structure nor the bond return volatility or covariance.

The next theorem gives sufficient condition, under which every unspanned factor in the models is a USV factor and there are no residual factors in the models. The sufficient condition is given in terms of variance-covariance matrix  $\hat{a}(z, u) = Sa(S^{-1}(z, u))S^\top$  of the transformed factor process  $\hat{X}_t = (Z_t, U_t)$ .

**Theorem 2.11.** *Assume for every  $j \in \{1, \dots, n\}$ , there exists  $i \in \{1, \dots, m\}$  such that  $\hat{a}_{ij}(z, u)$  is not constant in  $u_j$ . Then  $\mathcal{U} \cap \mathcal{W} = \{0\}$ , and therefore every unspanned factor is a USV factor. ([8], Theorem 3.4)*

The sufficient condition of Theorem 2.11 is satisfied for the square root diffusion model. Moreover, the linear-rational square root model can be specified in such a way that it has exactly  $m$  term structure and  $n$  unspanned stochastic volatility factors.

### 2.2.2 The Linear-Rational Square-Root Model

The Linear-Rational Square-Root (LRSQ) Model is a Linear-Rational Diffusion Model with state space  $E = \mathbb{R}_+^d$  in which the dynamics of the factor process is given by the multivariate square-root process of the form

$$dX_t = \kappa(\theta - X_t)dt + \text{Diag}\left(\sigma_1\sqrt{X_{1t}}, \dots, \sigma_d\sqrt{X_{dt}}\right)dB_t, \quad (2.27)$$

and the state space density is given by

$$\zeta_t = e^{-\alpha t}(\phi + \psi^\top X_t) \quad (2.28)$$

with parameters  $\phi, \psi, \kappa \in \mathbb{R}^{d \times d}, \theta \in \mathbb{R}^d$  and  $\sigma_i > 0$ .

We show how a class of LRSQ specifications with  $m$  term structure factors and  $n$  USV factors can be constructed. We first specify the state space density.

**Theorem 2.12.** *The short rate (2.7) is bounded from bellow if and only if, after coordinatewise scaling of the factor process (2.27), we have  $\zeta_t = e^{-\alpha t}(1 + \mathbf{1}^\top X_t)$ . In this case, the extremal values in (2.8) are given by  $\alpha^* = \max \mathcal{S}$  and  $\alpha_* = \min \mathcal{S}$ , where*

$$\mathcal{S} = \{\mathbf{1}^\top \kappa \theta, -\mathbf{1}^\top \kappa_1, \dots, -\mathbf{1}^\top \kappa_d\}$$

and  $\kappa_i$  denotes the  $i$ th column of  $\kappa$ . ([8], Theorem 4.1)

Hence, the state space density  $\zeta_t$  can be assumed to be of the form

$$\zeta_t = e^{-\alpha t}(1 + \mathbf{1}^\top X_t)$$

where

$$\alpha = \alpha^* = \max \mathcal{S} = \max \{\mathbf{1}^\top \kappa \theta, -\mathbf{1}^\top \kappa_1, \dots, -\mathbf{1}^\top \kappa_d\}. \quad (2.29)$$

According to Theorem 2.12, the parameters  $\phi = 1$  and  $\psi = \mathbf{1} = \mathbf{1}_d$  ensure that the short rate (2.7) in LRSQ model (2.27)-(2.28) is bounded from bellow. The choice of the parameter  $\alpha = \alpha^*$  ensures that it is non-negative for any  $t \geq 0$ .



Next we specify the factor process  $X_t$  (2.27) such that its parameters satisfy the conditions of Theorem 2.6. In this way we guarantee that the model has exactly  $m$  term structure and  $n$  unspanned factors with  $m + n = d$ . For this purpose we consider the invertible linear transformation  $S$  on  $\mathbb{R}^d$

$$S = \begin{pmatrix} Id_m & A \\ 0 & Id_n \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} Id_m & -A \\ 0 & Id_n \end{pmatrix}, \quad (2.30)$$

where  $A \in \mathbb{R}^{m \times n}$  is given by

$$A = \begin{pmatrix} Id_n \\ 0 \end{pmatrix}.$$

Let  $I = \{1, \dots, m\}$  and  $J = \{m + 1, \dots, d\}$  be the index sets and write the mean reversion matrix  $\kappa$  in block form as

$$\kappa = \begin{pmatrix} \kappa_{II} & \kappa_{IJ} \\ \kappa_{JI} & \kappa_{JJ} \end{pmatrix}$$

where  $\kappa_{IJ}$  denotes the submatrix whose rows are indexed by  $I$  and columns by  $J$ , and similarly for  $\kappa_{II}$ ,  $\kappa_{JI}$ ,  $\kappa_{JJ}$ .

Let us also decompose the mean reversion into  $\theta = (\theta_I, \theta_J) \in \mathbb{R}^m \times \mathbb{R}^n$  and fix some volatility parameters  $\sigma_i > 0, i = 1, \dots, d$ . Then the dynamics of the transformed factor process  $\hat{X}_t = S X_t = (Z_t, U_t)$  can be given by the square root diffusion process (2.27) with parameters obtained from (2.12):

$$\begin{aligned} \hat{\kappa} &= S\kappa S^{-1} = \begin{pmatrix} \kappa_{II} + A\kappa_{JI} & -\kappa_{II}A + A\kappa_{JJ} + \kappa_{IJ} - A\kappa_{JI}A \\ \kappa_{JI} & \kappa_{JJ} - \kappa_{JI}A \end{pmatrix}, \\ \hat{\psi} &= S\psi = \begin{pmatrix} \mathbf{1}_m \\ 0 \end{pmatrix}, \quad \hat{\theta} = S\theta = \begin{pmatrix} \theta_I + A\theta_J \\ \theta_J \end{pmatrix}. \end{aligned}$$

Assuming that  $\kappa_{IJ}$  satisfies the restriction

$$\kappa_{IJ} = \kappa_{II}A - A\kappa_{JJ} + A\kappa_{JI}A \quad (2.31)$$

it follows that the transformed LRSQ model satisfies the condition (i) – (ii) of Theorem 2.6 and therefore the LRSQ model (2.27)-(2.28) has at least  $n$  unspanned factors. The following result of Filipović et al. shows that if the model parameters satisfy

$$\hat{\kappa}_{ZZ} = \kappa_{II} + A\kappa_{JI} \quad (2.32)$$

and

$$\sigma_i \neq \sigma_{m+i} \quad \text{for all } i = 1, \dots, n \quad (2.33)$$

the LRSQ model (2.27)-(2.28) has exactly  $n$  USV factors.

**Proposition 2.13.** *The dimension of the term structure kernel is at least  $n$ ,  $\dim \mathcal{U} \geq n$ , with equality if  $\widehat{\kappa}_{ZZ} = \kappa_{II} + A\kappa_{JI}$  satisfies Theorem 2.6 (iii). In this case, if  $\sigma_i \neq \sigma_{m+i}$  for  $i = 1, \dots, n$ , all unspanned factors are USV factors.*

*Proof:* The first assertion about  $\dim \mathcal{U}$  follows from Theorem 2.6. The assertion that all unspanned factors are USV factors follows from Theorem 2.11. We use that the diffusion matrix of the transformed factor process  $\widehat{X}_t = (Z_t, U_t)$  is given by  $\widehat{a}(z, u) = S\sigma\sigma^\top(S^{-1}(z, u))S^\top$ . After calculation we obtain

$$\widehat{a}_{ii}(z, u) = \sigma_i^2 z_i + (\sigma_{m+i}^2 - \sigma_i^2) u_i, \quad i = 1, \dots, n,$$

which is non-constant in  $u_i$  since  $\sigma_i \neq \sigma_{m+i}$ .

The next assumptions for the parameters of LRSQ(m,n) specification

$$\kappa_{JI} = 0 \text{ and } \kappa_{JJ} = A^\top \kappa_{II} A \tag{2.34}$$

ensure the existence and the uniqueness of the square root diffusion process (2.27). Note that  $\kappa_{JJ}$  coincides with the upper left  $n \times n$  block of  $\kappa_{II}$ .

We present two specifications of LRSQ models. The first example specifies a LRSQ model with one term structure factor and one USV factor. The second example specifies a LRSQ model with three term structure factors and one USV factor.

**Example 2.14 (LRSQ(1,1)).**

In LRSQ(1,1) specification we have one term structure factor and one USV factor. According to (2.30)  $A = 1$  and hence the transformation  $S$  for this specification and its inverse are given by

$$S = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Since  $\widehat{X}_t = (Z_t, U_t) = S X_t$  the term structure factor and the unspanned factor become  $Z_t = X_{1t} + X_{2t}$  and  $U_t = X_{2t}$ . The restrictions (2.31) and (2.34) imply that the mean reversion matrix is given by

$$\kappa = \begin{pmatrix} \kappa_{11} & 0 \\ 0 & \kappa_{11} \end{pmatrix}$$

and the transformed mean reversion matrix  $\widehat{\kappa} = S\kappa S^{-1}$  coincide with  $\kappa$ ,

$$\widehat{\kappa} = \begin{pmatrix} \kappa_{11} & 0 \\ 0 & \kappa_{11} \end{pmatrix}.$$

Assuming that the volatility matrix is  $\sigma(X_t) = \text{Diag}(\sigma_1\sqrt{X_{1t}}, \sigma_2\sqrt{X_{2t}})$  then transformed volatility matrix is

$$\widehat{\sigma}(z, u) = S\sigma(x) = \begin{pmatrix} \sigma_1\sqrt{z_1 - u_1} & \sigma_2\sqrt{u_1} \\ 0 & \sigma_2\sqrt{u_1} \end{pmatrix}.$$

Therefore the transformed diffusion matrix  $\widehat{a}(z, u) = \widehat{\sigma}(z, u)\widehat{\sigma}(z, u)^\top$  satisfies  $\widehat{a}_{11}(z, u) = \sigma_1^2 z_1 + (\sigma_2^2 - \sigma_1^2)u_1$  and it is non-constant in  $u_1$  if  $\sigma_1 \neq \sigma_2$ . Then according to Theorem 2.11 the only one unspanned factor  $U_t$  in the LRSQ(1,1) specification is a USV factor if  $\sigma_1 \neq \sigma_2$ . The parameter in the LRSQ(1,1) specification which have to be calibrated are  $\alpha, \kappa_{11}, \sigma_1, \sigma_2, \theta_1, \theta_2$ . [8]

**Example 2.15 (LRSQ(3,1)).**

We consider the LRSQ(3,1) specification of a LRSQ model with three term structure factors and one unspanned factor. According to (2.30) the transformation  $S$  for this specification and its inverse are given by

$$S = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad S^{-1} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Here the matrix  $A \in \mathbb{R}^{3 \times 1}$  is  $A = (1, 0, 0, 0)^\top$ . In this case the term structure factors and the unspanned factor become

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \\ Z_{3t} \\ U_{1t} \end{pmatrix} = SX_t = \begin{pmatrix} X_{1t} + X_{4t} \\ X_{2t} \\ X_{3t} \\ X_{4t} \end{pmatrix}.$$

Under the restrictions (2.31) and (2.34) the mean reversion matrix is given by

$$\kappa = \begin{pmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & \kappa_{21} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & \kappa_{31} \\ 0 & 0 & 0 & \kappa_{11} \end{pmatrix}$$

and the transformed mean reversion matrix by

$$\widehat{\kappa} = S\kappa S^{-1} = \begin{pmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{33} & 0 \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & 0 \\ 0 & 0 & 0 & \kappa_{11} \end{pmatrix}.$$

The corresponding volatility matrix is

$$\widehat{\sigma}(z, u) = \begin{pmatrix} \sigma_1 \sqrt{z_1 - u_1} & 0 & 0 & \sigma_4 \sqrt{u_1} \\ 0 & \sigma_2 \sqrt{z_2} & 0 & 0 \\ 0 & 0 & \sigma_3 \sqrt{z_3} & 0 \\ 0 & 0 & 0 & \sigma_4 \sqrt{u_1} \end{pmatrix}.$$

Applying the Theorem 2.11 to the transformed diffusion matrix  $\widehat{a}(z, u) = \widehat{\sigma}(z, u)\widehat{\sigma}(z, u)^\top$  we see that  $\widehat{a}_{11}(z, u) = \sigma_1^2 z_1 + (\sigma_4^2 - \sigma_1^2)u_1$  is non-constant in  $u_1$  if  $\sigma_1 \neq \sigma_4$ . This means that the unspanned factor  $U_t$  in the LRSQ(3,1) specification is a USV factor if  $\sigma_1 \neq \sigma_4$ . [8]

### 2.2.3 Swaption Pricing in Linear Rational Square Root Model

The swaption price in a LRSQ model is calculated according to the price formula (2.23). Since the factor process  $X_t$  is affine, the Exponential-Affine Transform Formula can be applied (see Theorem 1.30) and the function  $\widehat{q}(z)$  in the swaption price formula can be evaluated. The evaluation of  $\widehat{q}(z)$  amounts to solving a system of ordinary differential equation given in the following theorem.

**Theorem 2.16** (Exponential-Affine Transform Formula). *Suppose that  $X_t$  is a  $d$ -dimensional affine diffusion process of the form (2.27). Then for any  $x \in \mathbb{R}^d, t \geq 0, u \in \mathbb{C}, v \in \mathbb{C}^d$  such that  $E_x [|\exp(v^\top X_t)|] < \infty$*

$$E_x \left[ e^{u+v^\top X_t} \right] = e^{\Phi(t)+x^\top \Psi(t)},$$

where  $\Phi : \mathbb{R}_+ \rightarrow \mathbb{C}, \Psi : \mathbb{R}_+ \rightarrow \mathbb{C}^d$  solve the system

$$\begin{aligned} \Phi'(\tau) &= b^\top \Psi(\tau) \\ \Psi'_i(\tau) &= \beta_i^\top \Psi(\tau) + \frac{1}{2} \sigma_i^2 \Psi_i(\tau)^2, \quad i = 1, \dots, d, \end{aligned}$$

with initial condition  $\Phi(0) = u, \Psi(0) = v$ . The solution to this system is unique.

# Chapter 3

## Estimation of LRSQ Models

### 3.1 Data

Filipović et al. estimate the model on a panel data set involving 827 weekly observations of swaps and swaptions over the time from January 29, 1997 to November 28, 2012. At each observation date, they observe rates on spot-starting swap contracts with maturities of one, two, three, five, seven and ten years, respectively as well prices on swaptions with three-month options maturities, the same six swap maturities, and strikes equal to the forward swap rates. For convenience they represent the swaption prices in term of implied volatilities. They convert the swaption prices into normal implied volatilities using the formula

$$\Pi_t^{\text{swpt}} = \sqrt{T_0 - t} \frac{1}{2\pi} \left( \sum_{i=1}^n \Delta P(t, T_i) \right) \sigma_{N,i} \quad (3.1)$$

where the zero-coupon bonds are bootstrapped from the swap curve.

### 3.2 Maximum Likelihood Estimation

For modelling of the linear-rational term structure Filipović et al. use three  $LRSQ(m, n)$  specifications in which the number of term structure factors is always set  $m = 3$  while the number of USV factors is set  $n = 1, 2$  or  $3$ , respectively. The specification  $LRSQ(3, 1)$  is used for a model in which the volatility of  $Z_{1t}$  contains one unspanned component, the specification  $LRSQ(3, 2)$  for a model in which the volatility of  $Z_{1t}$  and  $Z_{2t}$  contains unspanned components and the last specification  $LRSQ(3, 3)$  is used for a model in which the volatility of all term structure factors  $Z_{1t}, Z_{2t}, Z_{3t}$  contains unspanned components.

The model parameters are estimated by using the maximum likelihood approach in conjunction with Kalman filtering. As we will see in the next section

the Kalman filter allows simultaneous estimation of model parameters, unobservable state variables and also the likelihood function of the observable variables. For the Kalman filter to be applied, the linear-rational term structure model is approximated by a state space model with a measurement equation describing the relation between the state variables and observable swap rates and swaptions implied volatilities, and a transition equation describing the discrete-time dynamics of the state space variables.

In the linear-rational term structure models both the swap rates and the swaption implied volatilities are functions of the unobservable vector  $X_t$ . Therefore the measurement equation in state space model can be given by

$$Y_t = H(X_t, \Theta) + u_t, \quad u_t \sim N(0, \Sigma), \quad (3.2)$$

where  $Y_t$  denotes the vector consisting of observable swap rates and swaption implied volatilities,  $X_t$  denotes the vector of unobservable state variables at time  $t$ ,  $\Theta$  is the vector of model parameters,  $H$  is a vector-valued pricing function defined by (1.8) and (2.24) and  $u_t$  is a vector of i.i.d. Gaussian pricing errors with covariance matrix  $\Sigma$ . To reduce the number of parameters in  $\Sigma$  Filipović et al. assume that the pricing errors are cross-sectionally uncorrelated, i.e.  $\Sigma$  is diagonal. They also assume that all pricing errors for the swap rates and all pricing errors for the swaption implied volatilities have the same variances, denoted by  $\sigma_{rates}^2$  and  $\sigma_{swaptions}^2$  respectively.

In linear-rational term structure models the state variables follow a multidimensional CIR process. As an affine diffusion process its conditional mean and variance can be written in a closed form, however, the transition density is unknown. Filipović et al. approximate the transition density with a Gaussian density with identical first and second moments. They assume that the transition equation is of the form

$$X_t = \Phi_0 + \Phi_X X_{t-1} + \omega_t, \quad \omega_t \sim N(0, Q_t), \quad (3.3)$$

and chose the parameters  $\Phi_0$ ,  $\Phi_X$  and  $Q_t$  in such a way that the conditional moments of  $X_t$  are identical with conditional moments of the factor process in the linear-rational model. In this case the parameter  $\Phi_0$  and  $\Phi_X$  depend on the parameter  $\Theta$ . The variances  $Q_t$  are affine functions of  $X_{t-1}$  depending on the parameter  $\Theta$ .

The measurement and transition equations (3.2)–(3.3) define the state space model corresponding to the linear-rational term-structure model. Since both price functions in the measurement equations are non-linear some of the non linear Kalman Filters have to be applied for estimating the state space model. Filipović et al. use the non-linear unscented Kalman Filter. The Unscented Kalman Filter produces one-step-ahead forecasts  $\hat{Y}_{t|t-1}$  for the observation  $Y_t$  and corresponding error covariance matrices,  $\hat{F}_{t|t-1}$ , which are used to construct the log-likelihood

function

$$\mathcal{L}(\Theta) = \frac{1}{2} \sum_{t=1}^T \left( \# \log 2\pi + \log |\hat{F}_{t|t-1}| + (Y_t - \hat{Y}_{t|t-1})^\top \hat{F}_{t|t-1}^{-1} (Y_t - \hat{Y}_{t|t-1}) \right), \quad (3.4)$$

where  $T$  is the number of observation data and  $\#$  is the number of observations in  $Y_t$ . The (quasi) maximum likelihood estimator,  $\hat{\Theta}$  is then

$$\hat{\Theta} = \arg \max_{\Theta} \mathcal{L}(\Theta). \quad (3.5)$$

### 3.3 Unscented Kalman Filter

The *Unscented Kalman Filter* (UKF) is a filtering algorithm for estimating non-linear dynamic systems with measurement and process (transition) equations of the form

$$\mathbf{y}_k = H(\mathbf{x}_k, \mathbf{u}_k) \quad (3.6)$$

$$\mathbf{x}_k = F(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}) \quad (3.7)$$

where  $\mathbf{x}_k$  is a  $d$ -dimensional random vector describing the unobservable state variables of the non-linear system on time step  $k$ ,  $\mathbf{y}_k$  describes the measurement vector observed on time step  $k$  and  $\mathbf{u}_k$  and  $\mathbf{v}_k$  are uncorrelated random vectors describing the measurement and process noise, respectively with a zero-mean and diagonal covariance matrices  $\mathbf{R}_k$  and  $\mathbf{Q}_k$ , respectively. Note that it is not required that the noise sources are additive or Gaussian. The vector-valued function  $H$  describes the relation between the observations and the unobservable state variables, the vector-valued function  $F$  describes the dynamics of the state variables. It is assumed that the function  $F$  and  $H$ , as well as the initial state and the corresponding error covariance, are known. The equations (3.6) and (3.7) are called measurement and process equations, respectively.

The UKF estimation of the non-linear dynamic system includes an estimation of the unobservable states, an estimation of the unknown parameters as well as a simultaneous estimation of the states and parameters.

The Unscented Kalman Filter is a relatively new method for non-linear estimation, proposed by Julier and Uhlmann [10] as an extension of the traditional Kalman filter to the non-linear dynamic system, and at the same time as an improvement of the widely used Extended Kalman Filter for non-linear estimation.

The Unscented Kalman Filter uses an unscented transformation to approximate the joint distributions (or at least the first two moments) of the state  $\mathbf{x}_k$  and measurement  $\mathbf{y}_k$ . This way it allows the classical linear Kalman filter algorithm

to be applied for the estimation of non-linear dynamic systems. This method does not require the assumption that the noise sources are additive or Gaussian.

In order to understand better how the Unscented Kalman Filter works we first review the classical Kalman Filter algorithm. We introduce the Unscented Transformation and discuss how by means of Unscented Transformation the linear Kalman Filter algorithm can be applied for non-linear systems as well. We also show how the unobservable state variables and the unknown parameter can be estimated by UKF.

### 3.3.1 Kalman Filter

The *Kalman Filter* is a recursive algorithm for the estimation of linear time-discrete dynamic systems. On each time step the unknown state variables of the dynamic system can be estimated in a real time using only the present input measurement, the previous estimated state variables and its covariance matrix. The algorithm works in two steps:

- *Prediction (Prior estimation)*

On each time step  $k$  the Kalman Filter first predicts the current state given the measurement observed on the previous time step as well as the prior estimated state and its covariance matrix. The optimal prediction of the state,  $\hat{\mathbf{x}}_{k|k-1}$ , and its covariance matrix,  $\mathbf{P}_{k|k-1}$ , are obtained by the conditional expectations

$$\hat{\mathbf{x}}_{k|k-1} = \mathbb{E} [F(\hat{\mathbf{x}}_{k-1}, \mathbf{v}_{k-1}) \mid \mathbf{y}_{k-1}] \quad (3.8)$$

$$\mathbf{P}_{k|k-1} = \mathbb{E} [(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^\top \mid \mathbf{y}_{k-1}] \quad (3.9)$$

- *Update (Posterior estimation)*

Once the next measurement is observed, the predicted state and its covariance matrix are updated. The optimal state estimate is the expected current state  $\mathbf{x}_k$  conditional on the current observation  $\mathbf{y}_k$

$$\hat{\mathbf{x}}_k = \mathbb{E} [\mathbf{x}_k \mid \mathbf{y}_k].$$

Assuming that both the states and the measurements are Gaussian random variables, the optimal estimate of the current state is the weighted average of the predicted state and the new measurement given by the equation

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathcal{K}_k(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}), \quad (3.10)$$

its covariance is given by

$$\mathbf{P}_k = \mathbf{P}_{k|k-1} + \mathcal{K}_k \mathbf{P}_{yy} \mathcal{K}_k^\top \quad (3.11)$$



with

$$\hat{\mathbf{y}}_{k|k-1} = \mathbb{E} [H(\hat{\mathbf{x}}_{k|k-1}, \mathbf{u}_k) \mid \mathbf{y}_{k-1}] \quad (3.12)$$

$$\mathbf{P}_{yy} = \mathbb{E} [(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})^\top \mid \mathbf{y}_{k-1}] \quad (3.13)$$

$$\mathbf{P}_{xy} = \mathbb{E} [(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})^\top \mid \mathbf{y}_{k-1}] \quad (3.14)$$

$$\mathcal{K}_k = \mathbf{P}_{xy} \mathbf{P}_{yy}^{-1} \quad (3.15)$$

where  $\hat{\mathbf{y}}_{k|k-1}$  and  $\mathbf{P}_{yy}$  are the predictions of the expected observation on the time step  $k$  and its covariance matrix, respectively;  $\mathbf{P}_{xy}$  is the predicted cross correlation matrix; the weight  $\mathcal{K}_k$  is the *optimal Kalman gain* which shows how much the predicted state and its covariance matrix should be corrected and whereby the Kalman gain is chosen to minimize the mean square error of the estimate for  $\hat{\mathbf{x}}_k$ .

The Kalman Filter repeats these two step until there are no more observations or until it reaches a stable state and covariance matrix.

It is obvious that all equations in the prediction and update step are functions of the predicted value of the first two moments of the state  $\mathbf{x}_k$  and measurement  $\mathbf{y}_k$ . In linear case the Kalman Filter calculates these statistics exactly and in this way provides a closed form recursive estimate for the unobserved state variables of the linear discrete-time dynamic systems.

However, in practice the discrete-time dynamic systems are mostly non-linear. This means that both equations in the dynamic system (3.6)-(3.7) can be non-linear. In this case the traditional Kalman Filter can not be applied since the prediction of the current state and the prediction of the expected observation require calculating expectations of non-linear functions of the prior state estimate. An analytical solution of these expectations is possible only if the distribution of  $\mathbf{x}_k$  given  $\mathbf{y}_k$  is known, but even in this case the number of parameters which have to be propagated can be too great (is unlimited).

Some extensions of the Kalman Filter solve the problem with the conditional distributions and its moments, applying a transformation to the non-linear equations of the dynamic system. This way an approximation of the first two conditional moments of the state and measurement is obtained and the traditional Kalman Filter algorithm can be applied.

The *Extended Kalman Filter* (EKF), which was for many years the most common approach for non-linear estimation, first linearises the dynamic system (3.6)-(3.7) approximating the non-linear functions  $F$  and  $H$  with first order Taylor series, and then applies the Kalman Filter to the linearised dynamic system. However, the EKF has some significant drawbacks. The linearisation can introduce significant errors (in the true posterior mean and covariance). The Taylor series approximation requires the derivation of the Jacobian matrices which are not always trivial

and often lead to significant implementation difficulties. The EKF is reliable for systems which are almost linear on the time scale of the update intervals. The EKF achieves only first order (Taylor series expansion) accuracy when estimate the state [10].

The *Unscented Kalman Filter* is an improvement of the EKF. It is an extension of the Kalman Filter which uses the Unscented transformation (UT) to approximate the conditional moments of the state and measurement so that the traditional Kalman Filter can be applied. It has not the drawbacks of the EKF. The UKF does not need to calculate the Jacobian matrices as in EKF. The UKF achieves 3rd order (Taylor series expansion) estimation accuracy for any non-linearity, even though the computational complexity is the same as in EKF. [18].

### 3.3.2 Unscented Transformation

The Unscented Transformation is a method for calculating the statistics of a random variable (or vector) which undergoes a non-linear transformation [10]. Consider a random vector  $\mathbf{x}$  (dimension  $L$ ) with mean  $\bar{\mathbf{x}}$  and covariance matrix  $\mathbf{P}_{xx}$  and also the random vector  $\mathbf{y}$ , obtained after applying the non-linear vector-valued function  $f$  to the random vector  $\mathbf{x}$ , i.e.  $\mathbf{y} = f(\mathbf{x})$ . Suppose that the mean  $\bar{\mathbf{y}}$  and the covariance matrix  $\mathbf{P}_{yy}$  of the random vector  $\mathbf{y}$  have to be calculate, and the density function of  $\mathbf{y}$  can not be determined. In this case the mean and covariance of  $\mathbf{y}$  can be approximated using the unscented transformation. For this goal a set of  $2L + 1$  *sigma points*  $\mathcal{X}_i$  and corresponding weights  $W_i$  are chosen in such a way that the sample mean and sample covariance of the sigma points coincide with the mean and the covariance of  $\mathbf{x}$ :

$$\begin{aligned} \mathcal{X}_0 &= \bar{\mathbf{x}} & (3.16) \\ \mathcal{X}_i &= \bar{\mathbf{x}} + \left( \sqrt{(L + \lambda)\mathbf{P}_{xx}} \right)_i & i = 1, \dots, L \\ \mathcal{X}_i &= \bar{\mathbf{x}} - \left( \sqrt{(L + \lambda)\mathbf{P}_{xx}} \right)_{i-L} & i = L + 1, \dots, 2L \\ W_0^{(m)} &= \lambda / (L + \lambda) \\ W_0^{(c)} &= \lambda / (L + \lambda) + (1 - \alpha^2 + \beta) \\ W_i^{(m)} &= W_i^{(c)} = 1 / \{2(L + \lambda)\} & i = 1, \dots, 2L, \end{aligned}$$

where  $L$  is the dimension of  $\mathbf{x}$ ,  $\lambda = \alpha^2(L + \kappa) - L$ ,  $\alpha$ ,  $\kappa$ ,  $\beta$  are three scaling parameters, which fully determinate the scaling of the UT. The primary parameter  $\alpha$  determines the spread of the sigma points around  $\bar{\mathbf{x}}$  and varies from  $1e - 4$  to 1, the secondary parameter  $\kappa$  is usually set to 0 and the tertiary parameter  $\beta$  includes information about the prior distribution of  $\mathbf{x}$  (for Gaussian distribution

$\beta = 2$  is optimal, see [18]).  $\left(\sqrt{(L + \lambda)\mathbf{P}_{xx}}\right)_i$  is the  $i$ th row of the matrix square root. After that the sigma points are propagated through the non-linear function  $f$  to obtain a set of  $2L + 1$  *transformed sigma points*,

$$\mathcal{Y}_i = f(\mathcal{X}_i) \quad i = 1, \dots, 2L. \quad (3.17)$$

The mean and covariance of  $\mathbf{y}$  are approximated through the weighted sample mean and covariance of the transformed sigma points as follow,

$$\bar{\mathbf{y}} \approx \sum_{i=0}^{2L} W_i^{(m)} \mathcal{Y}_i \quad (3.18)$$

$$\mathbf{P}_{yy} \approx \sum_{i=0}^{2L} W_i^{(c)} (\mathcal{Y}_i - \bar{\mathbf{y}})(\mathcal{Y}_i - \bar{\mathbf{y}})^\top. \quad (3.19)$$

Note that this method resembles the Monte Carlo methods. However, there is a very important difference between them. The methods using UT use a very small number of deterministically chosen sample points to achieve an accurate approximation for the moments of the random variable  $\mathbf{y}$  while Monte Carlo methods require a huge number of sample points drawn at random [18, 10].

### 3.3.3 UKF State Estimation

The Unscented Kalman Filter provides a recursive estimate of the state following all steps of the classical Kalman Filter (3.8)–(3.15) and applying the UT to

- predict the new state  $\hat{\mathbf{x}}_{k|k-1}$  and its covariance  $\mathbf{P}_{k|k-1}$  given by (3.8)–(3.9);
- predict the expected observation  $\hat{\mathbf{y}}_{k|k-1}$  and its covariance  $\mathbf{P}_{yy}$  given by (3.12)–(3.13);
- predict the the cross-correlation matrix  $\mathbf{P}_{xy}$  given by (3.14);

Note that in the prediction steps the UT propagates the sigma points through the non-linear functions  $H$  and  $F$  given in the dynamic system (3.6)–(3.7), which depend not only on the state vector but also on the measurement noise and the process noise, respectively. In order UT to be applied the original state is augmented with the measurement noise and the process noise and UKF is applied to the new state  $\mathbf{x}_k^a = [\mathbf{x}_k^\top \mathbf{u}_k^\top \mathbf{v}_k^\top]^\top$ . The dimension of the new state is  $\dim(\mathbf{x}_k^a) = \dim(\mathbf{x}_k) + \dim(\mathbf{u}_k) + \dim(\mathbf{v}_k)$ . The UKF can be applied with the original state only when both process and measurement noise sources are additive and Gaussian. All equations of UKF algorithm for state estimation are given in the box below.

**Initialize before executing the UKF:**

$$\begin{aligned}\hat{\mathbf{x}}_0 &= E[\mathbf{x}_0] \quad \mathbf{P}_0 = E[(\mathbf{x}_0 - \hat{\mathbf{x}}_0)(\mathbf{x}_0 - \hat{\mathbf{x}}_0)^\top] \\ \hat{\mathbf{x}}_0^a &= E[\mathbf{x}_0] = [\hat{\mathbf{x}}_0^\top \mathbf{0} \mathbf{0}]^\top \\ \mathbf{P}_0^a &= E[(\mathbf{x}_0^a - \hat{\mathbf{x}}_0^a)(\mathbf{x}_0^a - \hat{\mathbf{x}}_0^a)^\top] = \begin{bmatrix} \mathbf{P}_0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_u & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{P}_v \end{bmatrix}\end{aligned}$$

**For each**  $k = 1, 2, 3, \dots$

*Calculate the sigma points:*

$$\mathcal{X}_{k-1}^a = \left[ \hat{\mathbf{x}}_{k-1}^a \quad \hat{\mathbf{x}}_{k-1}^a + \sqrt{(L^* + \lambda)\mathbf{P}_{k-1}^a} \quad \hat{\mathbf{x}}_{k-1}^a - \sqrt{(L^* + \lambda)\mathbf{P}_{k-1}^a} \right]$$

*Prediction equation:*

$$\begin{aligned}\mathcal{X}_{k|k-1}^x &= F(\mathcal{X}_{k-1}^x, \mathcal{X}_{k-1}^u) \\ \hat{\mathbf{x}}_{k|k-1} &= \sum_{i=0}^{2L} W_i^{(m)} \mathcal{X}_{i,k|k-1}^x \\ \mathbf{P}_{k|k-1} &= \sum_{i=0}^{2L} W_i^{(c)} (\mathcal{X}_{i,k|k-1}^x - \hat{\mathbf{x}}_{k|k-1}) (\mathcal{X}_{i,k|k-1}^x - \hat{\mathbf{x}}_{k|k-1})^\top\end{aligned}$$

*Update equation*

$$\begin{aligned}\mathcal{Y}_{k|k-1} &= H(\mathcal{X}_{k-1}^x, \mathcal{X}_{k-1}^v) \\ \hat{\mathbf{y}}_{k|k-1} &= \sum_{i=0}^{2L} W_i^{(m)} \mathcal{Y}_{i,k|k-1} \\ \mathbf{P}_k^{yy} &= \sum_{i=0}^{2L} W_i^{(c)} (\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_{k|k-1}) (\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_{k|k-1})^\top \\ \mathbf{P}_k^{xy} &= \sum_{i=0}^{2L} W_i^{(c)} (\mathcal{X}_{i,k|k-1}^x - \hat{\mathbf{x}}_{k|k-1}) (\mathcal{Y}_{i,k|k-1} - \hat{\mathbf{y}}_{k|k-1})^\top \\ \mathcal{K}_k &= \mathbf{P}_k^{xy} (\mathbf{P}_k^{yy})^{-1} \\ \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_{k|k-1} + \mathcal{K}_k (\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \\ \mathbf{P}_k &= \mathbf{P}_{k|k-1} + \mathcal{K}_k \mathbf{P}_k^{yy} \mathcal{K}_k^\top\end{aligned}$$

where  $\mathbf{x}^a = [\mathbf{x}^\top \mathbf{u}^\top \mathbf{v}^\top]^\top$ ,  $\mathcal{X}^a = [(\mathcal{X}^x)^\top (\mathcal{X}^u)^\top (\mathcal{X}^v)^\top]^\top$ ,  $L$  is the dimension of the augmented state,  $\mathbf{P}_u$  is the measurement noise covariance,  $\mathbf{P}_v$  is the process noise covariance,  $\lambda$  is the scaling parameter,  $W_i^{(m)}$  and  $W_i^{(c)}$  are the weights calculated according to (3.16).

### 3.3.4 UKF Parameter Estimation

In many applications the functions  $H$  and  $F$  of the non-linear dynamic system (3.6)-(3.7) are known up to a fixed number of parameters  $\theta = (\theta_1, \dots, \theta_p)^\top$  which also have to be estimated. In this case the equation of the non-linear system can be written in the form

$$\begin{aligned}\mathbf{y}_k &= H(\mathbf{x}_k, \mathbf{u}_k, \theta) \\ \mathbf{x}_k &= F(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}, \theta).\end{aligned}$$

By the means of UKF the unknown parameter  $\theta$  can be estimated from the observations even when the state is unknown. Which estimation method is used depends on the model. We consider the most common model, in which the measurement and process noise terms are additive and Gaussian. The non-linear system of the model is given by

$$\begin{aligned}\mathbf{y}_k &= H(\mathbf{x}_k, \theta) + \mathbf{u}_k \\ \mathbf{x}_k &= F(\mathbf{x}_{k-1}, \theta) + \mathbf{v}_{k-1},\end{aligned}$$

where  $H$  and  $F$  depend no more on the noise terms and according to the assumption, the state and observation vectors are approximated through Gaussian random variables. In this case the classical methods for parameter estimation like the Maximum likelihood method can be applied in conjunction with the UKF to estimate the parameter  $\theta$  from the observation vectors  $\mathbf{y}_k$ .

As we already mentioned, the *Maximum likelihood method* finds the values of parameter  $\theta$  that maximizes the (log-)likelihood function,  $\mathcal{L}(\theta)$  and the *Maximum likelihood estimate* (MLE) for the parameter  $\theta$  is obtained by

$$\hat{\theta} = \arg \max_{\theta} \log \mathcal{L}(\theta).$$

The *likelihood function* is a function of the parameter  $\theta$ , which corresponds to the joint density function of all observation,

$$\mathcal{L}(\theta) = f(\mathbf{y}_1, \dots, \mathbf{y}_N | \theta) = \prod_{k=1}^N f(\mathbf{y}_k | \mathbf{y}_{k-1}),$$

and shows how probable the observations for a given parameter  $\theta$  are. Here  $N$  denotes the number of the observation. The likelihood function can be presented also as a product of conditional density functions  $f(\mathbf{y}_k | \mathbf{y}_{k-1})$ . The likelihood function can be calculated if all conditional density functions are known.

In models with additive Gaussian noise terms the conditional observations are Gaussian random vectors,

$$\mathbf{y}_k | \mathbf{y}_{k-1} \sim \mathcal{N} \left( \mathbb{E}[H(\mathbf{x}_k, \theta) | \mathbf{y}_{k-1}], \mathbb{E}[(\mathbf{y}_k - H(\mathbf{x}_k, \theta))(\mathbf{y}_k - H(\mathbf{x}_k, \theta))^\top | \mathbf{y}_{k-1}] \right).$$

Since the UKF produces on each time step  $k$  the one-step-ahead forecast  $\hat{\mathbf{y}}_{k|k-1}$  for  $\mathbf{y}_k$  and corresponding error covariance matrices,  $\mathbf{P}_{k|k-1}^{yy}$ , the UKF can be also used to construct the log-likelihood function

$$\log \mathcal{L}(\theta) = \frac{1}{2} \sum_{t=1}^T \left( \# \log 2\pi + \log |\mathbf{P}_k^{yy}| + (\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1})^\top \mathbf{P}_k^{yy-1} (\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \right), \quad (3.20)$$

where  $T$  is the number of observation dates and  $\#$  is the number of observations in  $\mathbf{y}_k$ .

In non additive models or models in which the likelihood function can not be constructed the Maximum likelihood approach can not be applied, and in this case other estimation approaches are applied, see [18].

## 3.4 Square Root Diffusion Process

In term structure factor models, the square root diffusion process (also known as CIR process) is usually used to model the dynamics of unknown state variables (risk factors), e.g. short rate, volatility, etc.

In CIR model, one of the most commonly used short rate models, the square root diffusion process describes the dynamics of the short rate. The CIR model was introduced in 1985 by The Cox Ingersoll and Ross, and since then it is a standard model used for evaluation of interest rate derivatives.

The square root diffusion process is also used in Heston model to model the asset price volatility as a stochastic process.

The Chen model is an example for a three factors short rate model in which the short rate, the mean of short rate and the volatility of short rate are assumed as stochastic processes and their dynamics is specified by a three dimensional square root diffusion process.

The square root diffusion process also specifies the dynamics of the factors process in the linear-rational term structure model of Filipović. For this reason we focus on the properties of square root diffusion processes.

### 3.4.1 Definition

The  $d$ -dimensional square root diffusion process  $(X_t)_{t \geq 0}$  is a stochastic process defined as a solution of the stochastic differential equation (SDE)

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dB_t, \quad X_0 = x \quad (3.21)$$

where  $d \geq 1$ ,  $x \in \mathbb{R}_+^d$ ,  $X_t = (X_t^1, \dots, X_t^d)^\top$ ,  $\sqrt{X_t} = \text{diag}(\sqrt{X_t^1}, \dots, \sqrt{X_t^d})$ ,  $B_t = (B_t^1, \dots, B_t^d)^\top$  is a  $d$ -dimensional Wiener process (modelling the random

market risk factors) defined on the filtered probability space  $(\Omega, \mathcal{A}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  with  $\mathcal{F}_0 = \{\emptyset, \Omega\}$  and  $\kappa \in \mathbb{R}^{d \times d}$ ,  $\theta \in \mathbb{R}^d$ ,  $\sigma \in \mathbb{R}^{d \times d}$ ,  $\sigma = \text{diag}(\sigma_1, \dots, \sigma_d)$  are parameters.

The square root diffusion process is well defined when the positivity of process and the existence of solution are guaranteed. According Theorem 8.1 in [9], for any  $x \in \mathbb{R}_+^d$ , there exists a unique  $\mathbb{R}_+^d$ -valued solution  $X = X^x$  of (3.21) and  $X$  satisfies  $\mathbb{E}(\sup_{t \leq T} X_t^2) < \infty$  for all  $T \geq 0$ .

The solution of the square root diffusion process can not be given in closed form. However the moments can be evaluated. To find the solution of the SDE we apply the Itô formula to the stochastic process

$$Y_t := f(X_t, t) = \theta + e^{t\kappa}(X_t - \theta),$$

where  $e^{t\kappa}$  denotes the matrix exponential on the square matrix  $\kappa$ , defined by the power series  $e^{t\kappa} = \sum_{k \geq 0} \frac{t^k}{k!} \kappa^k$  and its derivation with respect to  $t$  is given by  $\frac{d}{dt} e^{t\kappa} = \kappa e^{t\kappa} = e^{t\kappa} \kappa$ . By Itô formula, we obtain the stochastic differential equation

$$\begin{aligned} dY_t &= f_t(X_t, t)dt + f_x(X_t, t)dX_t + \frac{1}{2}f_{xx}(X_t, t)(dX_t)^2 \\ &= \kappa e^{t\kappa}(X_t - \theta)dt + e^{t\kappa} \left( \kappa(\theta - X_t)dt + \sigma \sqrt{X_t}dB_t \right) \\ &= e^{t\kappa} \sigma \sqrt{X_t}dB_t \end{aligned}$$

and its solution

$$Y_t = Y_0 + \int_0^t e^{u\kappa} \sigma \sqrt{X_u}dB_u.$$

After substitution of  $Y_t$  and  $Y_0$  in the last equation we obtain the solution of the SDE (3.21)

$$X_t - \theta = e^{-t\kappa}(X_0 - \theta) + e^{-t\kappa} \int_0^t e^{u\kappa} \sigma \sqrt{X_u}dB_u. \quad (3.22)$$

### 3.4.2 Properties of $d$ -dimensional Square Root Diffusion Process

- **Distribution**

In general, the distribution of multi-dimensional square root diffusion process is unknown, but the moments can be evaluated.

According to (3.22) the square root diffusion process can be presented as the sum of the deterministic function  $f(t, X_0) = \theta + e^{-t\kappa}(X_0 - \theta)$  and the stochastic integral  $\int_0^t e^{u\kappa} \sigma \sqrt{X_u} dB_u$  which has the properties [12]:

- (1)  $\int_0^t e^{u\kappa} \sigma \sqrt{X_u} dB_u$  is a martingale w.r. to the filtration  $\mathcal{F} = (\mathcal{F}_t)_{t \geq 0}$
- (2)  $\mathbb{E} \left[ \int_0^t e^{u\kappa} \sigma \sqrt{X_u} dB_u \right] = 0$
- (3)  $\mathbb{E} \left[ \left( \int_0^t e^{u\kappa} \sigma \sqrt{X_u} dB_u \right) \left( \int_0^t e^{u\kappa} \sigma \sqrt{X_u} dB_u \right)^\top \right] =$   
 $= \mathbb{E} \left[ \int_0^t (e^{u\kappa} \sigma \sqrt{X_u}) (e^{u\kappa} \sigma \sqrt{X_u})^\top du \right]$
- (4)  $\mathbb{E} \left[ \left( \int_s^t e^{u\kappa} \sigma \sqrt{X_u} dB_u \right) \left( \int_s^t e^{u\kappa} \sigma \sqrt{X_u} dB_u \right)^\top \mid \mathcal{F}_s \right] =$   
 $= \mathbb{E} \left[ \int_s^t (e^{u\kappa} \sigma \sqrt{X_u}) (e^{u\kappa} \sigma \sqrt{X_u})^\top du \mid \mathcal{F}_s \right]$

- **Conditional Expectation**

For  $t \geq s$  the square root diffusion process can be written in the form

$$X_t = \theta + e^{-(t-s)\kappa}(X_s - \theta) + e^{-t\kappa} \int_s^t e^{u\kappa} \sigma \sqrt{X_u} dB_u.$$

Since the stochastic integral  $\int_0^t e^{u\kappa} \sigma \sqrt{X_u} dB_u$  is a martingale (1) it follows from the martingale property  $\mathbb{E} \left[ \int_0^t e^{u\kappa} \sigma \sqrt{X_u} dB_u \mid \mathcal{F}_s \right] = \int_0^s e^{u\kappa} \sigma \sqrt{X_u} dB_u$  that the  $\mathcal{F}_s$ -**conditional expectation of the square root diffusion process** is an affine function of  $X_s$  given by

$$\mathbb{E}[X_t \mid \mathcal{F}_s] = \theta + e^{-(t-s)\kappa}(X_s - \theta) = (1 - e^{-(t-s)\kappa})\theta + e^{-(t-s)\kappa} X_s. \quad (3.23)$$



• **Moments Calculation**

We write the square root diffusion process in the form

$$dX_t = b(X_t)dt + \rho(X_t)dB_t$$

where

$$\begin{aligned} b(\mathbf{x}) &= \kappa(\theta - \mathbf{x}) \\ \rho(\mathbf{x}) &= \sigma \text{diag}(\sqrt{x_1}, \dots, \sqrt{x_d}) \sigma^\top \\ a(\mathbf{x}) &= \rho(\mathbf{x})\rho(\mathbf{x})^\top = \sigma \text{diag}(\sigma_1^2 X_1, \dots, \sigma_d^2 X_d) \sigma^\top \end{aligned}$$

Applying Itô formula to  $f \in C^2(R^d)$ , we obtain

$$\begin{aligned} df(X_t) &= \sum_{i=1}^d \frac{\partial f}{\partial x_i}(X_t) dX_t^i + \frac{1}{2} \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_i \partial x_j}(X_t) \underbrace{dX_t^i dX_t^j}_{=a_{ij}(X_t)dt} \\ &= \mathcal{A}(f)(X_t)dt + \sum_{i=1}^d \frac{\partial f}{\partial x_i}(X_t) \sum_{l=1}^d \sigma_{il}(X_t) dB_t^l \\ \text{mit } \mathcal{A}(f)(\mathbf{x}) &+ \sum_{i=1}^d b_i(\mathbf{x}) \frac{\partial f}{\partial x_i}(\mathbf{x}) + \frac{1}{2} \sum_{i,j=1}^d a_{ij}(\mathbf{x}) \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}). \end{aligned}$$

It follows for  $t \geq s$  that

$$f(X_t) = f(X_s) + \int_s^t \mathcal{A}(f)(X_u)du + \underbrace{\sum_{i,l=1}^d \int_s^t \frac{\partial(f)}{\partial x_i}(X_u) \sigma_{il}(X_u) dB_u^l}_{\text{(local) martingale}}$$

and hence

$$\mathbb{E}(f(X_t)) = f(X_0) + \int_0^t \mathbb{E}(\mathcal{A}(f)(X_u))du \quad (\text{if true martingale}).$$

Note that if  $X$  is a CIR process then

$$\mathcal{A}(f)(\mathbf{x}) = \sum_{i=1}^d e_i^\top \kappa(\theta - \mathbf{x}) \frac{\partial f}{\partial x_i}(\mathbf{x}) + \frac{1}{2} \sum_{i,j=1}^d e_i^\top \sigma \text{diag}(x_1, \dots, x_d) \sigma^\top e_j^\top \frac{\partial^2 f}{\partial x_i \partial x_j}$$

maps polynomials of degree  $\leq m$  to polynomials of degree  $\leq m$

1. moments:

Consider the function  $f(\mathbf{x}) = x_l$  with  $\frac{\partial f}{\partial x_i}(\mathbf{x}) = \delta_{i,l}$  and  $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) = 0$  for  $i = 1, \dots, d$ . Then

$$\mathcal{A}(f)(x) = e_l^\top \kappa(\theta - x)$$

and hence

$$\mathbb{E}(X_t^l) = X_0^l + e_l^\top \int_0^t \kappa(\theta - \mathbb{E}(X_u)) du$$

and

$$\Rightarrow \mathbb{E}(X_t) = X_0 + \int_0^t \kappa(\theta - \mathbb{E}(X_n)) du.$$

The vector  $y_t = (1, \mathbb{E}(X_t^1), \dots, \mathbb{E}(X_t^d))^\top$  satisfies  $y_t = y_0 + \int_0^t \underbrace{\begin{pmatrix} 0 & 0 \\ \kappa\theta & -\kappa \end{pmatrix}}_{=:A} y_u du$ , and

hence it solves the differential equation. Therefore

$$y_t = e^{tA} y_0$$

Note that  $e^{tA}$  is matrix exponential and

$$e^{tA} = \sum_{k \geq 0} \frac{t^k}{k!} \begin{pmatrix} 0 & 0 \\ \kappa\theta & -\kappa \end{pmatrix}^k = \begin{pmatrix} 0 & 0 \\ -e^{\kappa t} \theta & e^{-t\kappa} \end{pmatrix}.$$

The first moment of  $X_t$  is

$$\Rightarrow E(X_t) = -e^{-t\kappa} \theta + e^{-t\kappa} E(X_0) = -e^{-t\kappa} (\theta - E(X_0))$$

2. moments:

Consider the function  $f(\mathbf{x}) = x_l x_m$  with  $\frac{\partial f}{\partial x_l}(\mathbf{x}) = x_m$ ,  $\frac{\partial f}{\partial x_m}(\mathbf{x}) = x_l$ ,  $\frac{\partial f}{\partial x_i}(\mathbf{x}) = 0$ , for  $i \neq \{l, m\}$  and  $\frac{\partial^2 f}{\partial x_l \partial x_m}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_m \partial x_l}(\mathbf{x}) = 1$ , otherwise is 0. Then

for  $l \neq m$

$$\mathcal{A}(f)(\mathbf{x}) = e_l^\top \kappa(\theta - \mathbf{x}) x_m + e_m^\top \kappa(\theta - \mathbf{x}) x_l + \underbrace{e_l^\top \sigma \text{diag}(x_1, \dots, x_d) \sigma e_m}_{=0, \text{ if } \sigma = \text{diag}(\sigma_1, \dots, \sigma_d)}$$

for  $l = m$

$$\mathcal{A}(f)(\mathbf{x}) = 2e_l^\top \kappa(\theta - x) x_l + \underbrace{e_l^\top \sigma \text{diag}(x_1, \dots, x_d) \sigma e_l}_{=\sigma^2 x_l, \text{ if } \sigma = \text{diag}(\sigma_1, \dots, \sigma_d)}$$

and hence for  $f(x) = x_l x_m$

$$\mathcal{A}(f)(\mathbf{x}) = e_l^\top \kappa(\theta - \mathbf{x}) x_m + e_m^\top \kappa(\theta - \mathbf{x}) x_l + \delta_{l,m} \sigma_l^2 x_l$$

and

$$E(X_t^l X_t^m) = E(X_0^l X_0^m) + \int_0^t (e_l^\top \kappa(\theta E(X_u^m) - E(X_u^m X_u)) + e_m^\top \kappa(\theta E(X_u^l) - E(X_u^l X_u)) + \delta_{l,m} \sigma_l^2 E(X_u^l)) du.$$

Since  $\mathcal{A}(f)$  maps the space

$$\mathcal{L}(1, x_1, \dots, x_d, x_1^2, x_1 x_2, \dots, x_1 x_d, x_2^2, x_2 x_3, \dots, x_2 x_d, \dots, x_d^2)$$

into the same space with base elements

$$f_1 = 1, f_2 = x_1, \dots, f_{d+1} = x_d, f_{d+2} = x_1^2, f_{d+3} = x_1 x_2, \dots, f_D = x_d^2$$

where  $D = 1 + d + d + \frac{d(d-1)}{2}$ , there exists a matrix  $A \in \mathbb{R}^{D \times D}$  such that

$$\mathcal{A}(f_i) = e_i^\top A f.$$

Let  $y_t = (\mathbb{E}(f_1(X_t)), \dots, \mathbb{E}(f_D(X_t)))^\top$ , then  $y_t$  satisfies

$$y_t = y_0 + \int_0^t A y_u du, \text{ and hence also } dy_t = A y_t dt.$$

The first two moments are given by

$$y_t = e^{tA} y_0$$

**Example 1:**  $d = 2 \Rightarrow D = 1 + 2 + 2 + 1 = 6$

$$\kappa = \begin{pmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{pmatrix}, \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix},$$

$$y_t = (1, \mathbb{E}(X_t^1), \mathbb{E}(X_t^2), \mathbb{E}(X_t^1)^2, \mathbb{E}(X_t^1 X_t^2), \mathbb{E}(X_t^2)^2)$$

$$1 = 1$$

$$\mathbb{E}(X_t^1) = X_0^1 + \int_0^t (\kappa_1 \cdot \theta - \kappa_{11} \mathbb{E}(X_u^1) - \kappa_{12} \mathbb{E}(X_u^2)) du$$

$$\mathbb{E}(X_t^2) = X_0^2 + \int_0^t (\kappa_2 \cdot \theta - \kappa_{22} \mathbb{E}(X_u^2) - \kappa_{21} \mathbb{E}(X_u^1)) du$$

$$\mathbb{E}((X_t^1)^2) = (X_0^1)^2 + \int_0^t (2\kappa_1 \cdot \theta \mathbb{E}(X_u^1) - 2\kappa_{11} \cdot \left( \frac{\mathbb{E}(X_u^1)^2}{\mathbb{E}(X_u^1 X_u^2)} \right) + \sigma_1^2 \mathbb{E}(X_u^1)) du$$

$$\mathbb{E}((X_t^2)^2) = (X_0^2)^2 + \int_0^t (2\kappa_2 \cdot \theta \mathbb{E}(X_u^2) - 2\kappa_2 \cdot \left( \frac{\mathbb{E}(X_u^1 X_u^2)}{\mathbb{E}(X_u^2)} \right) + \sigma_2^2 \mathbb{E}(X_u^2)) du$$

$$\mathbb{E}(X_t^1 X_t^2) = X_0^1 X_0^2 + \int_0^t (\kappa_1 \cdot \theta \mathbb{E}(X_u^2) - \kappa_1 \cdot \left( \frac{\mathbb{E}(X_u^1 X_u^2)}{\mathbb{E}(X_u^2)^2} \right) + \kappa_2 \cdot \theta \mathbb{E}(X_u^1) - \kappa_2 \cdot \left( \frac{\mathbb{E}(X_u^1)^2}{\mathbb{E}(X_u^1 X_u^2)} \right)) du$$

$$\Rightarrow y_t = y_0 + \int_0^t A y_u du \Rightarrow dy_t = e^{tA} y_0 \Rightarrow y_t = e^{tA} y_0$$

The elements of the matrix  $A$  are given in the table below

	1	$\mathbb{E}(X_0^1)$	$\mathbb{E}(X_0^2)$	$\mathbb{E}((X_0^1)^2)$	$\mathbb{E}(X_0^1 X_0^2)$	$\mathbb{E}((X_0^2)^2)$
1	0	0	0	0	0	0
$\mathbb{E}(X_t^1)$	$\kappa_1 \cdot \theta$	$-\kappa_{11}$	$-\kappa_{12}$	0	0	0
$\mathbb{E}(X_t^2)$	$\kappa_2 \cdot \theta$	$-\kappa_{21}$	$-\kappa_{22}$	0	0	0
$\mathbb{E}((X_t^1)^2)$	0	$\sigma_1^2 + 2\kappa_1 \cdot \theta$	0	$-2\kappa_{11}$	$-2\kappa_{12}$	0
$\mathbb{E}(X_t^1 X_t^2)$	0	$\kappa_2 \cdot \theta$	$\kappa_1 \cdot \theta$	$-\kappa_{21}$	$-\kappa_{11} - \kappa_{22}$	$-\kappa_{12}$
$\mathbb{E}((X_t^2)^2)$	0	0	$\sigma_2^2 + 2\kappa_2 \cdot \theta$	0	$-2\kappa_{21}$	$-2\kappa_{22}$

**Example 2:**  $d = 4 \Rightarrow D = 1 + 4 + 4 + 6 = 15$

$$\kappa = \begin{pmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & \kappa_{14} \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & \kappa_{24} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & \kappa_{34} \\ \kappa_{41} & \kappa_{42} & \kappa_{43} & \kappa_{44} \end{pmatrix}, \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix},$$

$$y_t = e^{tA} y_0$$

The elements of the matrix  $A$  are given in the table below.

A	1	$\mathbb{E}(X_0^1)$	$\mathbb{E}(X_0^2)$	$\mathbb{E}(X_0^3)$	$\mathbb{E}(X_0^4)$	$\mathbb{E}((X_0^1)^2)$	$\mathbb{E}(X_0^1 X_0^2)$	$\mathbb{E}(X_0^1 X_0^3)$	$\mathbb{E}(X_0^1 X_0^4)$	$\mathbb{E}((X_0^2)^2)$	$\mathbb{E}(X_0^2 X_0^3)$	$\mathbb{E}(X_0^2 X_0^4)$	$\mathbb{E}((X_0^3)^2)$	$\mathbb{E}(X_0^3 X_0^4)$	$\mathbb{E}((X_0^4)^2)$
1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
$\mathbb{E}(X_t^1)$	$\kappa_1 \cdot \theta$	$-\kappa_{11}$	$-\kappa_{12}$	$-\kappa_{13}$	$-\kappa_{14}$	0	0	0	0	0	0	0	0	0	0
$\mathbb{E}(X_t^2)$	$\kappa_2 \cdot \theta$	$-\kappa_{21}$	$-\kappa_{22}$	$-\kappa_{23}$	$-\kappa_{24}$	0	0	0	0	0	0	0	0	0	0
$\mathbb{E}(X_t^3)$	$\kappa_3 \cdot \theta$	$-\kappa_{31}$	$-\kappa_{32}$	$-\kappa_{33}$	$-\kappa_{34}$	0	0	0	0	0	0	0	0	0	0
$\mathbb{E}(X_t^4)$	$\kappa_4 \cdot \theta$	$-\kappa_{41}$	$-\kappa_{42}$	$-\kappa_{43}$	$-\kappa_{44}$	0	0	0	0	0	0	0	0	0	0
$\mathbb{E}((X_t^1)^2)$	0	$\sigma_1^2 + 2\kappa_{11} \cdot \theta$	0	0	0	$-2\kappa_{11}$	$-2\kappa_{12}$	$-2\kappa_{13}$	$-2\kappa_{14}$	0	0	0	0	0	0
$\mathbb{E}(X_t^1 X_t^2)$	0	$\kappa_2 \cdot \theta$	$\kappa_1 \cdot \theta$	0	0	$-\kappa_{21}$	$-\kappa_{11} - \kappa_{22}$	$-\kappa_{23}$	$-\kappa_{24}$	$-\kappa_{12}$	$-\kappa_{13}$	$-\kappa_{14}$	0	0	0
$\mathbb{E}(X_t^1 X_t^3)$	0	$\kappa_3 \cdot \theta$	0	$\kappa_1 \cdot \theta$	0	$-\kappa_{31}$	$-\kappa_{32}$	$-\kappa_{11} - \kappa_{33}$	$-\kappa_{34}$	0	$-\kappa_{12}$	0	$-\kappa_{13}$	$-\kappa_{14}$	0
$\mathbb{E}(X_t^1 X_t^4)$	0	$\kappa_4 \cdot \theta$	0	0	$\kappa_1 \cdot \theta$	$-\kappa_{41}$	$-\kappa_{42}$	$-\kappa_{43}$	$-\kappa_{11} - \kappa_{44}$	0	0	$-\kappa_{12}$	0	$-\kappa_{13}$	$-\kappa_{14}$
$\mathbb{E}((X_t^2)^2)$	0	0	$\sigma_2^2 + 2\kappa_{22} \cdot \theta$	0	0	0	$-2\kappa_{21}$	0	0	$-2\kappa_{22}$	$-2\kappa_{23}$	$-2\kappa_{24}$	0	0	0
$\mathbb{E}(X_t^2 X_t^3)$	0	0	$\kappa_3 \cdot \theta$	$\kappa_2 \cdot \theta$	0	0	$-\kappa_{31}$	$-\kappa_{21}$	0	$-\kappa_{32}$	$-\kappa_{22} - \kappa_{33}$	$-\kappa_{34}$	$-\kappa_{23}$	$-\kappa_{24}$	0
$\mathbb{E}(X_t^2 X_t^4)$	0	0	$\kappa_4 \cdot \theta$	0	$\kappa_2 \cdot \theta$	0	$-\kappa_{41}$	0	$-\kappa_{21}$	$-\kappa_{42}$	$-\kappa_{43}$	$-\kappa_{22} - \kappa_{44}$	0	$-\kappa_{23}$	$-\kappa_{24}$
$\mathbb{E}((X_t^3)^2)$	0	0	0	$\sigma_3^2 + 2\kappa_{33} \cdot \theta$	0	0	0	$-2\kappa_{31}$	0	0	$-2\kappa_{32}$	0	$-2\kappa_{33}$	$-2\kappa_{34}$	0
$\mathbb{E}(X_t^3 X_t^4)$	0	0	0	$\kappa_4 \cdot \theta$	$\kappa_3 \cdot \theta$	0	0	$-\kappa_{41}$	$-\kappa_{31}$	0	$-\kappa_{42}$	$-\kappa_{32}$	$-\kappa_{43}$	$-\kappa_{33} - \kappa_{44}$	$-\kappa_{34}$
$\mathbb{E}((X_t^4)^2)$	0	0	0	0	$\sigma_4^2 + 2\kappa_{44} \cdot \theta$	0	0	0	$-2\kappa_{41}$	0	0	$-2\kappa_{42}$	0	$-2\kappa_{43}$	$-2\kappa_{44}$

### 3.4.3 Properties of one dimensional Square Root Diffusion Process (or one dimensional CIR Process)

- **Distribution**

The  $\mathcal{F}_s$ -conditional distribution of one dimensional CIR process ( $d = 1$ ) for any  $0 \leq s \leq t$  is, up to a scale factor, a non-central Chi-Squared distribution

$$\frac{X_t}{\tau(t-s)} \sim \chi_\nu(\mu)$$

with a degree of freedom  $\nu = \frac{4\kappa\theta}{\sigma^2}$  and non-centrality parameter  $\mu = \frac{X_s e^{-\kappa(t-s)}}{\tau(t-s)}$ , where  $\tau(t-s) = \frac{4\kappa}{\sigma^2}(1 - e^{-\kappa(t-s)})$  and  $0 \leq s \leq t$ .

*Proof:* We show that  $\frac{X_t}{\tau(t-s)}$  conditional on  $\mathcal{F}_s$  has the moment generating function of a non-central Chi-Squared distribution with parameter  $\nu$  and  $\mu$ ,

$$M_{\frac{X_t}{\tau(t-s)}}(u) = (1 - 2u)^{-\nu/2} e^{\frac{\mu u}{1-2u}}.$$

The CIR process is an affine process and therefore its SDE can be written in the form

$$dX_t = b(X_t)dt + \rho(X_t)dB_t,$$

where  $b(X_t) = b_0 + b_1 X_t$  with  $b_0 = \kappa\theta$ ,  $b_1 = -\kappa$  and  $a(X_t) = \rho(X_t)^2 = a_0 + a_1 X_t$  with  $a_0 = 0$ ,  $a_1 = \sigma^2$ .

According to the Definition for affine processes, given in [7], the  $\mathcal{F}_s$ -conditional characteristic function of  $X_t$  is exponential affine in  $X_t$ , i.e. there exist the  $\mathbb{C}$ -valued functions  $\phi(t, u)$  and  $\psi(t, u)$ , respectively, with jointly continuous  $t$ -derivatives so that the solution  $X = X^x$  satisfies

$$M_{X_t|\mathcal{F}_s}(u) = \mathbb{E} [e^{uX_t} | \mathcal{F}_s] = \exp(\phi(t-s, u) + \psi(t-s, u)X_s)$$

for all  $u \in \mathbb{C}$  and  $0 \leq s \leq t$ , and the functions  $\phi(t, u)$  and  $\psi(t, u)$  solve the system of Riccati equations

$$\begin{aligned} \partial_t \phi(t, u) &= \frac{1}{2} a_0 \psi(t, u)^2 + b_0 \psi(t, u), & \phi(0, u) &= 0 \\ \partial_t \psi(t, u) &= \frac{1}{2} a_1 \psi(t, u)^2 + b_1 \psi(t, u), & \psi(0, u) &= u \end{aligned}$$

where  $b_0 = \kappa\theta$ ,  $b_1 = -\kappa$ ,  $a_0 = 0$  and  $a_1 = \sigma^2$  (see for example [7], [11]).

According to Theorem A.3 in [11] the solution of the second equation is

$$\psi(t-s, u) = \frac{u e^{-\kappa(t-s)}}{1 - 2u\tau(t-s)},$$

where  $\tau(t-s) = \frac{\sigma^2}{4\kappa}(1 - e^{-\kappa(t-s)})$ . It can be also shown that the solution of the first equation is

$$\phi(t-s, z) = -\frac{2\kappa\theta}{\sigma^2} \log(1 - 2u\tau(t-s)).$$

Hence, the moment generating function of  $X_t$  conditional on  $\mathcal{F}_s$

$$M_{X_t|\mathcal{F}_s}(u) = (1 - 2u\tau(t-s))^{-\nu/2} \exp\left(\frac{ue^{-\kappa(t-s)}}{1 - 2u\tau(t-s)}X_s\right) \quad (u \in \mathbb{C})$$

and it can be rewritten as

$$M_{\frac{X_t}{\tau(t-s)}|\mathcal{F}_s}(u) = M_{X_t|\mathcal{F}_s}\left(\frac{u}{\tau(t-s)}\right) = (1 - 2u)^{-\nu/2} \exp\left(\underbrace{\frac{X_s e^{-\kappa(t-s)}}{\tau(t-s)}}_{\mu} \frac{u}{1 - 2u}\right).$$

□

### • Conditional Expectation

The mean and covariance of a non-central Chi-squared distribution with a degree of freedom  $\nu$  and a non-centrality parameter  $\mu$  are given by  $\mathbb{E}[\chi_\nu(\mu)] = \nu + \mu$  and  $\text{Var}[\chi_\nu(\mu)] = 2(\nu + 2\mu)$ .

We should obtain the same formula for **conditional expectation of square root diffusion process**  $X_t$  given  $\mathcal{F}_s$  as in (3.23)

$$\begin{aligned} \mathbb{E}[X_t | \mathcal{F}_s] &= \tau(t-s) \left( \frac{4\kappa\theta}{\sigma^2} + \frac{X_s e^{-\kappa(t-s)}}{\tau(t-s)} \right) \\ &= \frac{4\kappa\theta\tau(t-s)}{\sigma^2} + X_s e^{-\kappa(t-s)} \\ &= \theta(1 - e^{-\kappa(t-s)}) + X_s e^{-\kappa(t-s)} \end{aligned}$$

### • Conditional Variance

Unlike the multi-dimensional case the **conditional variance of one dimensional square root diffusion process** has an explicit solution,

$$\begin{aligned} \text{Var}[X_t | \mathcal{F}_s] &= 2\tau(t-s)^2 \left( \frac{4\kappa\theta}{\sigma^2} + 2\frac{X_s e^{-\kappa(t-s)}}{\tau(t-s)} \right) \\ &= \frac{8\kappa\theta\tau(t-s)^2}{\sigma^2} + 4X_s e^{-\kappa(t-s)}\tau(t-s) \\ &= \frac{\theta\sigma^2}{2\kappa} (1 - e^{-\kappa(t-s)})^2 + X_s \frac{\sigma^2}{\kappa} (e^{-\kappa(t-s)} - e^{-2\kappa(t-s)}). \end{aligned} \quad (3.24)$$

### 3.4.4 Euler-Maruyama Scheme for Simulation of Square Root Diffusion Process

The Euler-Maruyama scheme is a simple method for an approximate numerical solution of a stochastic differential equation based on the Euler method for solving ordinary deterministic differential equations. A description of the Euler-Maruyama scheme can be found in many sources.

Suppose that we have to solve the SDE of the  $d$ -dimensional square root diffusion process (3.21)

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dW_t, \quad X_0 = \mathbf{x}$$

on some interval of time  $[0, T]$  and some  $d \geq 1$ . The Euler-Maruyama scheme calculates an approximation  $\hat{X}$  of the true solution  $X_t$  using the following discretisation scheme

- The interval  $[0, T]$  is divided into  $N$  subintervals of width  $\Delta t = T/N$  choosing  $N + 1$  time-discrete points

$$0 = t_0 < t_1 < \dots < t_N = T, \quad \text{where } t_k = k\Delta t.$$

- The stochastic differentials are substituted

$$\begin{aligned} dX_k &= X_{k+1} - X_k \\ dt &= \Delta t \\ dW_k &= \Delta W_k = W_{k+1} - W_k \end{aligned}$$

Note that it follows from the properties of Wiener process that the increments  $\Delta W_k$  are independent and identically distributed normal random variables with expected value zero and variance  $\Delta t$ , i.e.  $\Delta W_k \sim N(0, \Delta t)$  for  $k = 1, \dots, N$ .

- The approximation  $\hat{X}$  of  $X$  is defined recursively for  $k = 1, \dots, N$  as follows

$$\begin{aligned} \hat{X}_0 &= X_0 \\ \hat{X}_{k+1} &= X_k + \kappa(\theta - X_k)k\Delta t + \sigma\sqrt{X_k}\Delta W_k \end{aligned} \quad (3.25)$$

Figure 3.1: Euler-Maruyama scheme for simulation of square root diffusion process  $dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dB_t$  with  $X_0 = \mathbf{x}$ .



The Euler-Maruyama approximation scheme has the advantage over other approximative methods that it is fast and simple, however it has also some drawbacks [17]. Due to the fact that the continuous time process is approximated by the discrete-time process, the Euler-Maruyama scheme can introduce a bias into the estimator used for the approximation, however the bias can be reduced to an acceptable level for a big enough  $N$ . Another serious problem is that the algorithm can break down due to the negative value and the impossibility to calculate the square root. Since the Gaussian increments are non bounded from bellow the Euler-Maruyama approximation can lose its accuracy near zero, especially, for large  $\sigma$  even when the positivity of the CIR process, that have to be approximated, is guaranteed. Many improvements to the Euler-Maruyama approximation scheme are proposed.

- Improvement of Euler-Maruyama scheme

$$\hat{X}_{k+1} = X_k + \kappa(\theta(-X_k)^+)k\Delta t + \sigma\sqrt{(X_k)^+}\Delta W_k \quad (3.26)$$

proposed by Deelstra and Delbaen.

- Improvement of Euler-Maruyama scheme

$$\hat{X}_{k+1} = \left| X_k + \kappa(\theta - X_k)k\Delta t + \sigma\sqrt{X_k}\Delta W_k \right| \quad (3.27)$$

proposed by Diop.

### 3.4.5 Exact Simulation of One Dimensional Square Root Diffusion Process

The exact simulation of a one dimensional square root diffusion process  $X$  is based on the property of simulated process that for any  $0 \leq s < t$  the distribution of  $X(t)$  given  $X(s)$  is, up to a scale factor, a non-central Chi-Squared distribution

$$X_t \sim \tau(t-s)\chi_\nu(\mu),$$

where  $\nu = \frac{4\kappa\theta}{\sigma^2}$  is a degree of freedom,  $\mu = \frac{X_s e^{-\kappa(t-s)}}{\tau(t-s)}$  is the non-centrality parameter and  $\tau(t-s) = \frac{4\kappa}{\sigma^2}(1 - e^{-\kappa(t-s)})$  is the scale factor.

The exact simulation algorithm of a one dimensional square root diffusion process

$$dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dB_t, \quad X_0 = x,$$

on a time grid  $0 = t_0 < t_1 < \dots, < t_N = T$  includes the following steps:

*Initialization*

- set time grid  $0 = t_0 < t_1 < \dots, < t_N = T$
- set degree of freedom  $\nu = \frac{4\kappa\theta}{\sigma^2}$

For each  $k = 1, 2, 3, \dots$

- calculate scale factor  $\tau(t_k - t_{k-1}) = \frac{4\kappa}{\sigma^2}(1 - e^{-\kappa(t_k - t_{k-1})})$
- calculate non-centrality parameter  $\mu = \frac{X_{t_{k-1}} e^{-\kappa(t_k - t_{k-1})}}{\tau(t_k - t_{k-1})}$
- generate random variable  $Y \sim \chi_\nu(\mu)$
- $X_{t_k} = \tau(t_k - t_{k-1})Y$

Figure 3.2: Exact simulation of one dimensional square root diffusion process  $dX_t = \kappa(\theta - X_t)dt + \sigma\sqrt{X_t}dB_t$  with  $X_0 = x$ .

## 3.5 Estimation of LRSQ(1,0) Model

### 3.5.1 Model Specification

The first model, we estimate, is the linear-rational square root model with only one term structure factor and no unspanned factors (LRSQ(1,0)). In the simplest LRSQ model the dynamics of term structure factor is given by the one-dimensional square root diffusion process

$$dZ_t = \kappa(\theta - Z_t)dt + \sigma\sqrt{Z_t}dB_t, \quad \text{with } Z_0 = x > 0 \quad (3.28)$$

and the state space density by

$$\zeta_t = e^{-\alpha t}(1 + Z_t), \quad (3.29)$$

where  $\kappa, \theta \in \mathbb{R}$  and  $\sigma \in \mathbb{R}_+$  are unknown real-valued parameters and  $(Z_t)_{t \geq 0}$  is unknown state space process with values in  $\mathbb{R}_+$ . The parameter  $\alpha$  is set

$$\alpha = \alpha^* = \max(\kappa\theta, -\kappa)$$

to guarantee that the short rate (2.7) in LRSQ(1,0) model is non-negative for any time  $t$ .

The *unknown parameters*  $\kappa, \theta$  and  $\sigma$  of the LRSQ(1,0) model as well as the *unknown state variables*  $Z_t$  for  $t \geq 0$  can be estimated from the market swap rates and the market swaption prices. As market data we use we use the model swap rates and the model swaption prices.

### 3.5.2 Model Data

We generate a panel data set including weekly data for swaps and swaptions over the time period of sixteen years. Each panel data row consists of model rates on spot-starting swap contracts with maturities of one, two, three, five, seven and ten years, respectively, and six months payment frequency as well as of models prices on swaptions with three-month options maturities, the same six swap maturities, and strikes equal to the forward swap rates.

In linear-rational term structure models the swap rates and the swaption prices are non-linear functions of the factor process  $(Z_t)_{t \geq 0}$  and depend on the parameters  $\kappa, \theta$  and  $\sigma$ . The factor process  $(Z_t)_{t \geq 0}$  is a one dimensional CIR process with the same parameters  $\kappa, \theta, \sigma$  and start value  $Z_0$ . A path of  $Z$  can be simulated.

We assume that one financial year consists of 52 weeks and the time interval between two observations is  $1/52$ . To obtain the weekly data for a period of 16 years we calculate for given parameters  $\kappa, \theta, \sigma$  and factor process  $(Z_t)_{t \geq 0}$  the aforementioned swap rates and swaption prices using the formulas (1.8) and (2.24) on the time grid

$$0 = t_0 < t_1 < \dots < t_N = 16,$$

where  $N = 16 \cdot 52$  and  $t_k = k \cdot \frac{1}{52}$  for  $k = 1, \dots, N$ .

#### *Simulation of the factor process $Z$*

Since both the swap rates and the swaption prices at time  $t$  are non-linear functions of the factor process  $Z_t$ , we first have to simulate the factor process  $Z$  on the same time grid  $0 = t_0 < t_1 < \dots < t_N = 16$ . The factor process is unobservable, but according to the assumption in the LRSQ(1,0) model it follows the one dimensional CIR process (3.28) with the parameters  $\kappa, \theta$  and  $\sigma$ . To generate the factor process  $Z$  we simulate a path of one dimensional CIR process with given values for the parameters  $\kappa, \theta$  and  $\sigma$  by using either Euler-Maruyama approach or Exact simulation.

#### *Calculation of the swap rates*

A swap with maturity  $T$  and six months payment frequency is specified by a predetermined annualized rate and a tenor structure of reset and payment dates

$$0 \leq t \leq T_0 < T_1 < \dots < T_n = T$$

where  $n$  is the number of interest rate payments and also the number of interest rate periods,  $T_i$  is the payment date for the  $i$ th period and therefore  $T_i = T_0 + i \cdot \frac{1}{2}$ . The  $t$ -time interest rate of a spot-starting swap ( $t = T_0$ ) is given, according to (1.8), by

$$S_t = \frac{1 - P(t, T_n)}{\sum_{i=1}^n \Delta_i P(t, T_i)} \quad (3.30)$$

and according to (2.6) and the assumption in LRSQ(1,0) model, the zero-coupon bond prices  $P(t, T_i)$  are given by the linear-rational function of the state  $Z_t$

$$P(t, T_i) = \frac{(\phi + \psi^\top \theta)e^{-\alpha\tau} + \psi^\top e^{-(\alpha+\kappa)\tau}(Z_t - \theta)}{\phi + \psi^\top Z_t}, \quad \text{for } i = 1, \dots, n \quad (3.31)$$

where  $\phi = \psi = 1$ ,  $\tau = T_i - t$  and  $\Delta_i = T_i - T_{i-1} = 1/2$ .

#### Calculation of the swaption prices

The price of swaption (swap option) on the same swap with three-months option maturity (i.e.  $T_0 - t = 1/4$ ) is given according to Theorem 2.9 and Theorem 2.16 by

$$\Pi_t^{swpt} = \frac{1}{\zeta_t \pi} \int_0^\infty \text{Re} \left[ \frac{\widehat{q}(\mu + i\lambda)}{(\mu + i\lambda)^2} \right] d\lambda \quad (3.32)$$

for some  $\mu > 0$  such that  $\widehat{q}(\mu) < \infty$ , where

$$\widehat{q}(\mu + i\lambda) = \mathbb{E} [\exp((\mu + i\lambda)p_{swpt}(Z_{T_0})) \mid \mathcal{F}_t] = e^{\Phi(T_0-t) + Z_t \Psi(T_0-t)}, \quad (3.33)$$

and the functions  $\Phi : \mathbb{R}_+ \rightarrow \mathbb{C}$  and  $\Psi : \mathbb{R}_+ \rightarrow \mathbb{C}$  solve the system of ordinary Ricatti differential equations

$$\begin{aligned} \Phi'(\tau) &= \kappa \theta \Psi(\tau), \\ \Psi'(\tau) &= -\kappa \Psi(\tau) + \frac{1}{2} \sigma^2 \Psi(\tau)^2 \end{aligned} \quad (3.34)$$

with initial condition

$$\begin{aligned} \Phi(0) &= u = (\mu + i\lambda) \sum_{i=0}^n c_i e^{-\alpha T_i} (\phi + \psi^\top \theta - \psi^\top e^{-\kappa(T_i - T_0)} \theta), \\ \Psi(0) &= v = (\mu + i\lambda) \sum_{i=0}^n c_i e^{-\alpha T_i} \psi^\top e^{-\kappa(T_i - T_0)}, \end{aligned}$$

where

$$\begin{aligned} c_0 &= 1, \quad c_n = -1 - \Delta K, \quad c_i = -\Delta K, \quad i = 1, \dots, n-1 \\ K &= S_{T_0}, \quad \Delta = T_i - T_{i-1} = 1/2 \quad \text{and} \\ \phi &= \psi = 1. \end{aligned}$$

We calculate the different swap rates and swaption prices with parameters  $\kappa = 0.0630$ ,  $\theta = 0.6709$ ,  $\sigma = 0.2269$  and  $\alpha = \kappa\theta$ , following the algorithm in figure 3.3 and use the calculated data as market data for the estimation of LRSQ(1,0) models.

*Initialization*

- Set the parameter  $\kappa, \theta, \sigma, \alpha = \max\{\kappa\theta, -\kappa\}$
- Set the time grid  $0 = t_0 < t_1 < \dots < t_N$

*Simulation of CIR process  $Z$  with parameters  $\kappa, \theta, \sigma$  on the time grid  $0 = t_0 < t_1 < \dots < t_N$*

- Set the start value  $Z_0$

For  $k = 1, \dots, 16 \cdot 52$

- Exact simulation of  $Z_{t_k}$

*Calculation of the swap rates on a spot-starting swap contract with maturity  $T$  and six months payment frequency on the time grid  $0 = t_0 < t_1 < \dots < t_N$*

- Set the number of payment dates:  $n = 2T$
- Set the reset/payment dates:  $T_0 = t_k; T_i = T_0 + \frac{1}{2}i, i = 1, \dots, n$

For  $k = 1, \dots, 16 \cdot 52$

- Calculate the zero-coupon bond prices  $P(t_k, T_i)$  for  $i = 1, \dots, n$
- Calculate the spot-starting swap rate  $S_{t_k}$  (for  $T_0 = t_k$ )

*Calculation of swaption prices on 3 months option*

For  $k = 1, \dots, 16 \cdot 52$

- Calculate  $\zeta_{t_k} = e^{-\alpha t_k}(1 + Z_{t_k})$
- Calculate the swap rate  $K = S_{t_k}$  ( for  $T_0 = t_k + 1/4$ )
- Calculate the swaption price  $\Pi_{t_k}^{swpt} = \frac{1}{\zeta_{t_k} \pi} \int_0^\infty \text{Re} \left[ \frac{\hat{q}(\mu+i\lambda)}{(\mu+i\lambda)^2} \right] d\lambda$

Figure 3.3: Calculation of swap rates and swaption prices given parameter values  $\kappa, \theta, \sigma$  and start value  $Z_0$

Figure 3.4 shows the simulated path of the one dimensional CIR process (with the  $\kappa = 0.0630, \theta = 0.6709, \sigma = 0.2269$  and  $\alpha = \kappa\theta$ ), used for the calculation of the swap rates and the swaption prices. In figure 3.5 we can see that the short rate is non-negative on each time point. The calculated swap rates and swaption

prices are shown in figure 3.6 and figure 3.7.

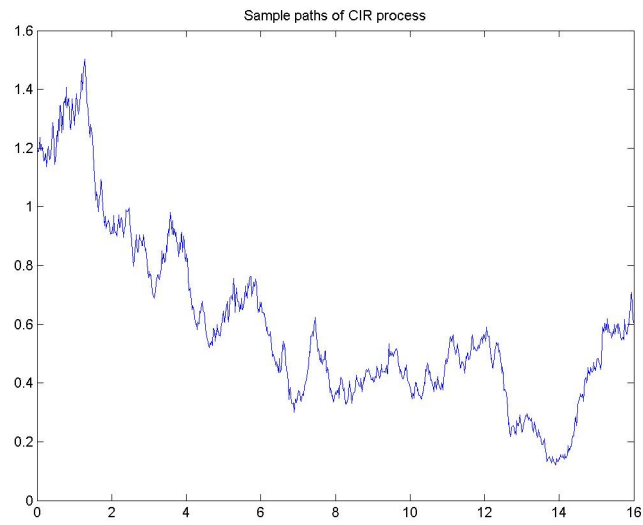


Figure 3.4: Simulated path of one dimensional CIR process  $Z$  with parameters  $\kappa = 0.0630$ ,  $\theta = 0.6709$ ,  $\sigma = 0.2269$ ,  $\alpha = \kappa\theta$  and start value  $Z_0 = 1.2$

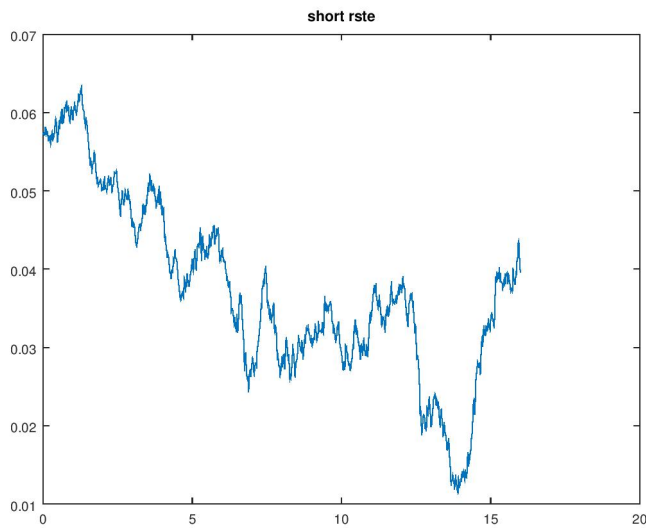


Figure 3.5: Short rate in LRSQ(1,0) model

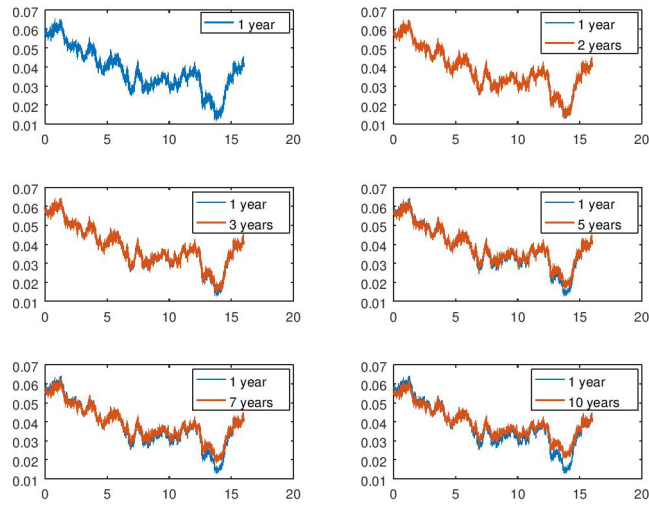


Figure 3.6: Swap rates of spot-starting swaps with 1-, 2-, 3-, 5-, 7- and 10 years maturity and 6 months payment frequency. Parameters:  $\kappa = 0.0630$ ,  $\theta = 0.6709$ ,  $\sigma = 0.2269$ ,  $\alpha = \kappa\theta$

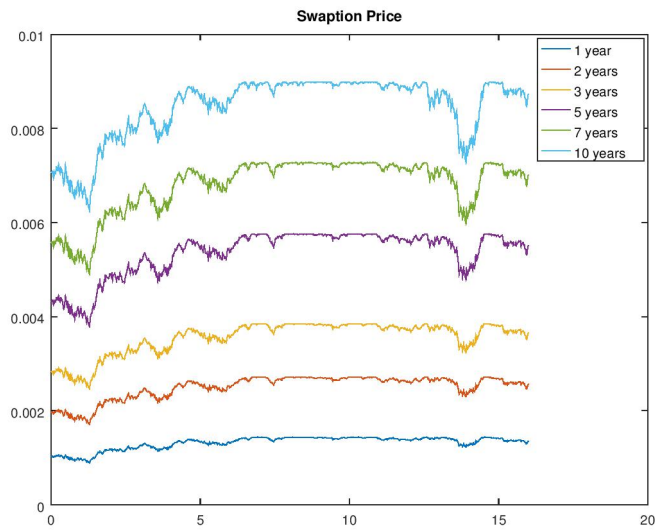


Figure 3.7: Simulated swaption prices on 3 months swap options with 1-, 2-, 3-, 5-, 7- and 10 years maturity and 6 months payment frequency. Parameters:  $\kappa = 0.0630$ ,  $\theta = 0.6709$ ,  $\sigma = 0.2269$ ,  $\alpha = \kappa\theta$

### 3.5.3 Reconstruction of $\kappa, \theta, \sigma$ and $Z$

To check our numerical calculation we reconstruct the model parameters  $\kappa, \theta, \sigma$  and the factor process  $Z$  from the swap rates and swaption prices calculated by the LRSQ(1,0) model on a time grid  $0 = t_0 < t_1 < \dots < t_k \dots < t_n = 16$  (with  $t_k = k \cdot \frac{1}{52}$ ).

- *Reconstruction of  $\kappa, \theta, \alpha$  and  $Z$*

In LRSQ(1,0) model the zero-coupon bond prices and the swap rates depend only on the parameters  $\kappa, \theta$  and the factor process  $Z$ . Only three different market swap rates (or zero-coupon bond prices) at time  $t = 0$  (for example with maturity 1, 2 and 3 years) are necessary to be determined the exact values of  $\kappa, \theta$  and  $Z_0$ . Having these values, the process  $Z$  can be exactly reconstructed from a series of swap rates (for example with maturity 1 year). Note that according to Proposition 2.8 the zero-coupon bond price function is injective and  $Z_t$  can be reconstructed from the model price at time  $t$ .

Let  $S_0^1, S_0^2, S_0^3$  denote the market/model swap rates at time  $t = 0$  of swap with maturity 1, 2, or 3 years, and 6 months payment frequency. Then  $\kappa, \theta$  and  $Z_0$  solve the following non-linear system:

$$\begin{aligned} S_0^1 &= \frac{P_0 - P_2}{(P_1 + P_2)\Delta} \\ S_0^2 &= \frac{P_0 - P_4}{(P_1 + P_2 + P_3 + P_4)\Delta} \\ S_0^3 &= \frac{P_0 - P_6}{(P_1 + P_2 + P_3 + P_4 + P_5 + P_6)\Delta} \end{aligned} \tag{3.35}$$

where for  $i = 0, \dots, 6$

$$\begin{aligned} P_i &= P(0, T_i) = \frac{(1 + \theta)e^{-\alpha\tau_i} + e^{-(\alpha+\kappa)\tau_i}(Z_0 - \theta)}{1 + Z_0} \\ \alpha &= \kappa\theta \\ t &= 0, T_0 = t + 1/4, T_i = T_0 + i \cdot 1/2 \\ \tau_i &= T_i - t, \Delta = T_i - T_{i-1} = 1/2. \end{aligned}$$

The *numerical solution* of the system (3.35) is given by

$$\begin{aligned} \hat{\kappa} &= 0.0630 \\ \hat{\theta} &= 0.6618 \\ \hat{\alpha} &= 0.0417 \\ \hat{Z}_0 &= 1.2132. \end{aligned}$$



The factor process  $(Z_{t_k})_{0=t_0 \leq t_k \leq t_N=16}$  is reconstructed (see figure 3.8) from the equation

$$S_{t_k}^1 = \frac{P(t_k, T_0) - P(t_k, T_2)}{(P(t_k, T_1) + P(t_k, T_2))\Delta},$$

where  $S_t^1$  is the market/model swap rate at time  $t = t_0, \dots, t_N$  of a swap with maturity 1 year and  $P(t, T_i)$  is the price at time  $t$  of a zero-coupon bond with maturity  $T_i$  given in (3.31).

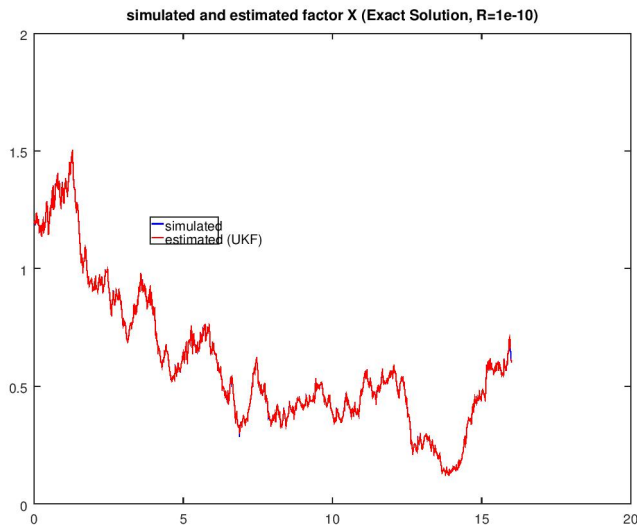


Figure 3.8: Estimated and simulated factor process  $Z$

- *Estimation of parameter  $\sigma$*

Once we have the values of the parameters  $\kappa, \theta, \alpha$  and the values of the factor process  $Z_t$  at time  $0 = t_0 < t_1 < \dots < t_k \dots < t_n = 16$  (with  $t_k = k \cdot \frac{1}{52}$ ), we need the market price of only one swaption to estimate the parameter  $\sigma$ . Let  $P_0^1$  denote the market/model price at time  $t$  of a swaption on a swap with maturity 1 year. Then  $\sigma$  solves the equation (3.32)

$$P_0^1 = \frac{1}{\zeta_t \pi} \int_0^\infty \operatorname{Re} \left[ \frac{\widehat{q}(\mu + i\lambda)}{(\mu + i\lambda)^2} \right] d\lambda.$$

The *numerical solution* of this equation is

$$\hat{\sigma} = 0.22690. \tag{3.36}$$

### 3.5.4 Maximum Likelihood Estimation in Conjunction with UKF

The *state space model* corresponding to the LRSQ(1,0) model is given by the following measurement and process equations

$$Y_{t_k} = h(Z_{t_k}, \Theta) + u_{t_k} \quad (3.37)$$

$$Z_{t_k} = \Phi_0 + \Phi_Z Z_{t_{k-1}} + w_{t_k} \quad (3.38)$$

where

- $0 \leq t_k \leq 16$  and  $t_k = k \cdot \frac{1}{52}$  for  $k = 0, 1, \dots, 16 \cdot 52$ .
- $Y_{t_k} = (Y_{t_k}^1 \ Y_{t_k}^2)^\top \in \mathbb{R}^2$  is a vector of observable market swap rates and swap-tion prices at time  $t_k$  with maturity 1, 2, 3, 5, 7 or 10 years.
- $Z_{t_k} \in \mathbb{R}$  is the unobservable state variable at time  $t_k$  and by assumption it is a Gaussian variable whose first two conditional moments are equal to the first two conditional moments of the factor process (3.28).
- $h = (h_1, h_2)^\top : \mathbb{R} \rightarrow \mathbb{R}^2$  is a vector-valued function where  $h_1$  is defined according to (3.30) and  $h_2$  according to (3.32).
- $u_{t_k}$  is the measurement noise and  $u_{t_k} \sim N(0, R)$  with  $R = \begin{pmatrix} \sigma_{swap}^2 & 0 \\ 0 & \sigma_{swpt}^2 \end{pmatrix}$ .
- $w_{t_k}$  is the process noise and  $w_{t_k} \sim N(0, Q_{t_k})$ .
- $\Theta = (\kappa, \theta, \sigma)^\top$  as well as  $\Phi_0, \Phi_X$  and  $Q_{t_k}$  are the unknown parameters, which have to be estimated from the market data.

*Estimation of parameters  $\Phi_0, \Phi_X$  and  $Q_{t_k}$*

Since, by assumption, the conditional moments of the state variables  $Z_{t_k}$  in state space model are identical to the conditional moments of the factor process (which is a CIR process), it follows that

$$\begin{aligned} \Phi_0 + \Phi_Z Z_{t_{k-1}} &= \mathbb{E} [Z_{t_k} | \mathcal{F}_{t_{k-1}}] = \theta (1 - e^{-\kappa(t_k - t_{k-1})}) + Z_{t_{k-1}} e^{-\kappa(t_k - t_{k-1})} \\ Q_{t_k} &= \mathbb{V}ar [Z_{t_k} | \mathcal{F}_{t_{k-1}}] = \frac{\theta \sigma^2}{2\kappa} (1 - e^{-\kappa(t_k - t_{k-1})})^2 + Z_{t_{k-1}}^2 \frac{\sigma^2}{\kappa} (e^{-\kappa(t_k - t_{k-1})} - e^{-2\kappa(t_k - t_{k-1})}). \end{aligned}$$

Hence, the parameters  $\Phi_0, \Phi_Z$  and  $Q_{t_k}$  can be expressed by the parameters  $\kappa, \theta, \sigma$  and the state variable  $Z_{t_{k-1}}$  as follow

$$\begin{aligned}\Phi_0 &= \theta (1 - e^{-\kappa/52}) \\ \Phi_Z &= e^{-\kappa/52} \\ Q_{t_k} &= \frac{\theta\sigma^2}{2\kappa} (1 - e^{-\kappa/52})^2 + Z_{t_{k-1}} \frac{\sigma^2}{\kappa} (e^{-\kappa/52} - e^{-2\kappa/52}).\end{aligned}$$

and therefore  $\kappa, \theta$  and  $\sigma$  are the only parameters which have to be estimated.

*Maximum likelihood estimation of parameters*

To estimate the parameters  $\Theta = (\kappa, \theta, \sigma)^\top$  we *first apply the UKF* algorithm to the state space model (3.37)-(3.38) to construct the log-likelihood function

$$\mathcal{L}(\Theta) = \frac{1}{2} \sum_{t=t_0}^{t_N} \left( 2 \log 2\pi + \log | \widehat{F}_{t|t-1} + (Y_t - \widehat{Y}_{t|t-1})^\top \widehat{F}_{t|t-1}^{-1} (Y_t - \widehat{Y}_{t|t-1}) \right),$$

where  $\widehat{Y}_{t|t-1}$  is the one-step-ahead forecast for  $Y_t$  at time  $t$  and  $\widehat{F}_{t|t-1}$  is the corresponding error covariance matrix. Then we *calculate the maximum likelihood estimation* for  $\Theta$  by

$$\widehat{\Theta} = \arg \max_{\Theta} \mathcal{L}(\Theta).$$

Applying the UKF to the model swap rates with 1 year maturity and the model swaption prices with 1 year maturity, calculated with parameters  $\kappa = 0.0630, \theta = 0.6709, \sigma = 0.2269$ , we obtain the maximum likelihood estimation for the parameters  $\kappa, \theta, \sigma$

estimated	true
$\hat{\kappa} = 0.0636$	$\kappa = 0.0630$
$\hat{\theta} = 0.6738$	$\theta = 0.6709$
$\hat{\sigma} = 0.2239$	$\sigma = 0.2269$

as well as the UKF estimate for the factor process  $Z$ . The simulates and estimated factor process are given in figure 3.9.

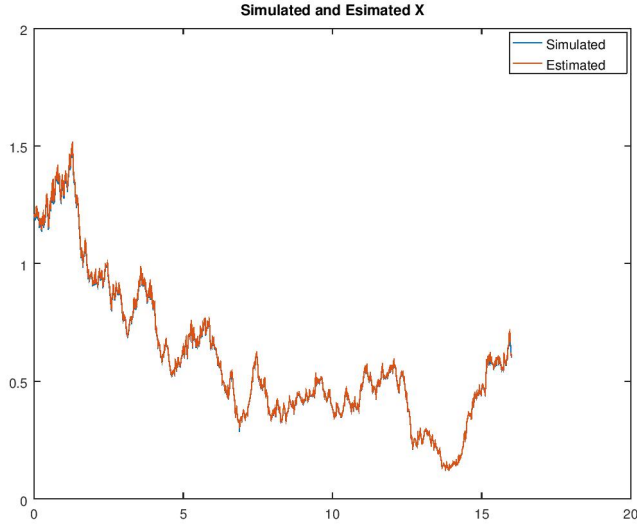


Figure 3.9: Estimated and simulated factor process  $Z$  (UKF)

## 3.6 Estimation of LRSQ(1,1) Model Specification

### 3.6.1 Model Specification

The second model, we estimate, is a two-dimensional linear-rational square root model LRSQ(1,1), which has exactly one term structure factor and one unspanned stochastic volatility (USV) factor. According to the specification of LRSQ(1,1) model (see Example 2.14) the factor process  $X$  follows the two-dimensional square root diffusion process (CIR process)

$$dX_t = \kappa(\theta - X_t)dt + \text{Diag}(\sigma_1\sqrt{X_{1t}}, \sigma_2\sqrt{X_{2t}})dB_t, \quad X_0 = x \in \mathbb{R}_+^2 \quad (3.39)$$

and the state space density  $\zeta$  is given by

$$\zeta_t = e^{-\alpha t}(1 + \mathbf{1}_2^\top X_t), \quad (3.40)$$

where

- $\kappa = \begin{pmatrix} \kappa_{11} & 0 \\ 0 & \kappa_{11} \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ ,  $\theta = \begin{pmatrix} \theta_1 \\ \theta_2 \end{pmatrix} \in \mathbb{R}^2$ ,  $\sigma_1, \sigma_2 \in \mathbb{R}_+$  with  $\sigma_1 \neq \sigma_2$  are unknown real-valued parameters;
- $X = (X_{1t}, X_{2t})_{t \geq 0}^\top$  denotes a vector of unobservable state space variables with values in  $\mathbb{R}_+^2$ ;

- $B = (B_{1t}, B_{2t})_{t \geq 0}^\top$  is a two-dimensional Browning motion;
- $\alpha$  is a real-valued parameter which has to be chosen large enough to guarantee that the short rate is non-negative for any time  $t$ . We set

$$\alpha = \alpha^* = \max \{ \mathbf{1}_2^\top \kappa \theta, -\mathbf{1}_2^\top \kappa_{\cdot 1}, -\mathbf{1}_2^\top \kappa_{\cdot 2} \} = \max \{ \kappa_{11} \theta_1, \kappa_{11} \theta_2, -\kappa_{11} \}$$

according to Theorem 2.12.

The *estimation of LRSQ(1,1)* specification includes the estimation of *unknown parameters*  $\Theta = (\kappa_{11}, \theta_1, \theta_2, \sigma_1, \sigma_2)$  as well as the estimation of *unobservable state variables*  $X_t = (X_{1t}, X_{2t})^\top$  on specific time points from the market swap rates and the market swaption prices. As market data we use the model swap rates and the model swaption prices. We estimate the unknown parameters and the unknown state variables by *Maximum likelihood approach in conjunction with Unscented Kalman Filter*.

### 3.6.2 Model Data

Similar to LRSQ(1,0) model, we generate a panel data set including weekly data for swaps and swaptions over the time period of sixteen years. Each panel data row consists of model rates on spot-starting swap contracts with maturities of one, two, three, five, seven and ten years, respectively, and six months payment frequency as well as of models prices on swaptions with three-month options maturities, the same six swap maturities, and strikes equal to the forward swap rates. The model swap rates and swaption prices in the LRSQ(1,1) model are functions of the two dimensional factor process  $(X_t)_{t \geq 0}$  and depend on the parameters  $\kappa_{11}, \theta_1, \theta_2, \sigma_1, \sigma_2$ . To obtain the model data we first have to simulate the factor process  $(X_t)_{t \geq 0}$  for given parameters  $\kappa_{11}, \theta_1, \theta_2, \sigma_1, \sigma_2$  and a given start value  $X_0$  and then calculate the model swap rates and the model swaption prices according to (1.8) and (2.24).

*Calculation of model swap rates and model swaption prices*

We calculate the *model swap rates* and *model swaption prices* on the time grid

$$0 = t_0 < t_1 < \dots < t_N = 16, \quad t_k = k \cdot 1/52 \text{ for } k = 1, \dots, N$$

for the parameters

$$\Theta = (\kappa_{11}, \theta_1, \theta_2, \sigma_1, \sigma_2)^\top = (0.0630, 0.6709, 0.2903, 0.2269, 0.1688)^\top$$

and the start value

$$X_0 = (1.2, 0.5)^\top.$$

Following the algorithm in figure 3.3 we

1. set the time grid

$$0 = t_0 < t_1 < \dots < t_N = 16,$$

where  $N + 1 = 16 \cdot 52 + 1$  is the number of the simulated weekly observations and  $t_k = k \cdot \frac{1}{52}$  for  $k = 1, \dots, N$ .

2. simulate a path of the two-dimensional CIR process (3.39) with start value  $X_0 = (1.2, 0.5)^\top$  on the time grid  $0 = t_0 < t_1 < \dots < t_N = 16$  using Euler-Maruyama approach.
3. calculate for each  $t_k$  the swap rates of spot-starting swaps with maturity  $T = 1, 2, 3, 5, 7, 10$  and payment dates  $t_k = T_0 < T_1 < \dots < T_n = T$  according to

$$S_{t_k} = \frac{P(t_k, T_0) - P(t_k, T_n)}{\sum_{i=1}^n \Delta_i P(t_k, T_i)} \quad (3.41)$$

where

$$P(t_k, T_i) = \frac{(\phi + \psi^\top \theta) e^{-\alpha \tau} + \psi^\top e^{-(\alpha + \kappa) \tau} (X_{t_k} - \theta)}{\phi + \psi^\top X_{t_k}}, \quad (3.42)$$

$$\phi = 1, \quad \psi = (1 \ 1)^\top, \quad \tau = T_i - t, \quad \Delta_i = T_i - T_{i-1} = 1/2,$$

$$n = 2T = 2, 4, 6, 10, 14 \text{ or } 20.$$

4. calculate for each  $t_k$  the prices of swaptions on three-month swap options with maturity  $T = 1, 2, 3, 5, 7, 10$ , payment dates  $t_k = T_0 < T_1 < \dots < T_n = T$  and strikes equal to the forward swap rates according to

$$\Pi_{t_k}^{swpt} = \frac{1}{\zeta_{t_k} \pi} \int_0^\infty \operatorname{Re} \left[ \frac{\widehat{q}(\mu + i\lambda)}{(\mu + 1\lambda)^2} \right] d\lambda \quad (3.43)$$

for some  $\mu > 0$  such that  $\widehat{q}(\mu) < \infty$ , where

$$\widehat{q}(\mu + i\lambda) = \mathbb{E} \left[ e^{(\mu+i\lambda)p_{\text{swpt}}(X_{T_0})} \mid \mathcal{F}_t \right] = e^{\Phi(T_0-t) + X_t^\top \Psi(T_0-t)}, \quad (3.44)$$

and  $\Phi : \mathbb{R}_+ \rightarrow \mathbb{C}, \Psi : \mathbb{R}_+ \rightarrow \mathbb{C}^2$  solve the system of Ricatti ordinary differential equations

$$\begin{aligned} \Phi'(\tau) &= \kappa \theta^\top \Psi(\tau) \\ \Psi'(\tau) &= -\kappa^\top \Psi(\tau) + \frac{1}{2} \Psi(\tau)^\top \text{Diag}(\sigma_1^2, \sigma_2^2) \Psi(\tau) \end{aligned} \quad (3.45)$$

with initial condition

$$\begin{aligned} \Phi(0) = u &= (\mu + i\lambda) \sum_{i=0}^n c_i e^{-\alpha T_i} (\phi + \psi^\top \theta - \psi^\top e^{-\kappa(T_i-T_0)} \theta) \\ \Psi(0) = v &= (\mu + i\lambda) \sum_{i=0}^n c_i e^{-\alpha T_i} \psi^\top e^{-\kappa(T_i-T_0)} \end{aligned}$$

where

$$\begin{aligned} c_0 &= 1, c_n = -1 - \Delta K, c_i = -\Delta K, i = 1, \dots, n-1 \\ K &= S_{T_0}, \tau = T_0 - t_k = 1/4, \Delta = T_i - T_{i-1} = 1/2 \text{ and} \\ \phi &= 1, \psi = (1, 1)^\top. \end{aligned}$$

In figure 3.10, figure 3.12 and figure 3.13 can be seen the simulated CIR process  $X$ , the model swap rates and the model swaption prices with parameters  $\kappa_{11} = 0.0630, \theta_1 = 0.6709, \theta_2 = 0.2903, \sigma_1 = 0.2269, \sigma_2 = 0.1688, \alpha = 0.0636$ .

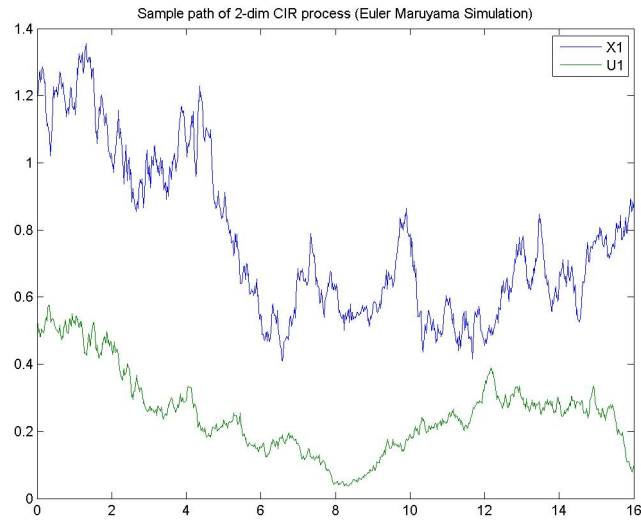


Figure 3.10: Simulated path of two dimensional CIR process  $X = (X_{1t} \ X_{2t})^\top$  with parameters  $\kappa_{11} = 0.0630$ ,  $\theta_1 = 0.6709$ ,  $\theta_2 = 0.2903$ ,  $\sigma_1 = 0.2269$ ,  $\sigma_2 = 0.1688$ ,  $\alpha = 0.0636$  and start value  $X_0 = (1.2, 0.5)^\top$

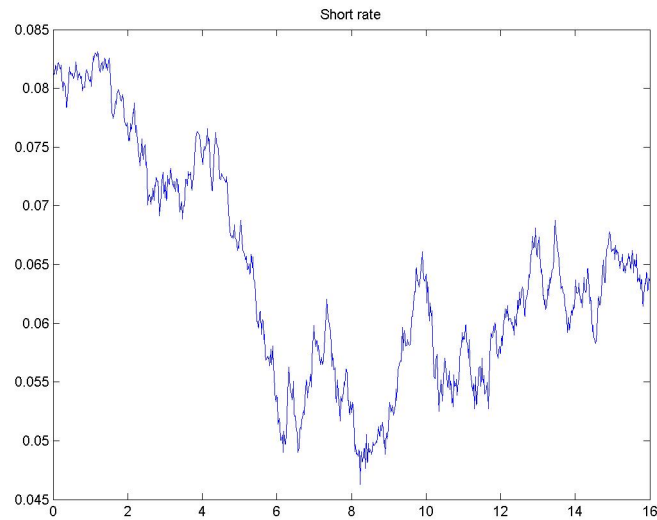


Figure 3.11: Short rate in LRSQ(1,1) model



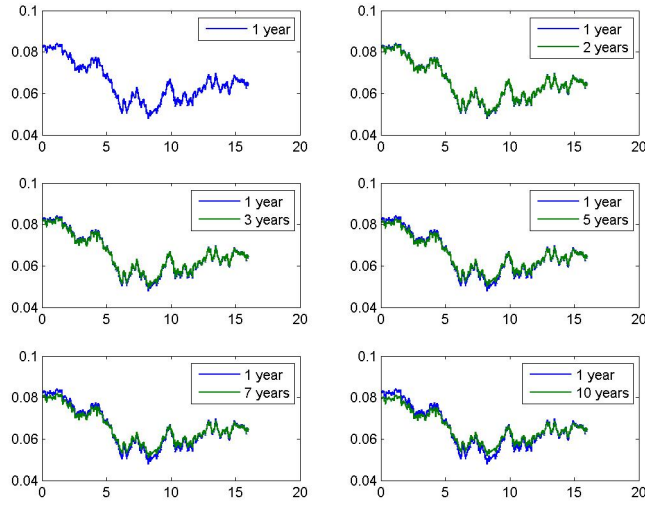


Figure 3.12: Swap rates of spot-starting swaps with 1-, 2-, 3-, 5-, 7- and 10 years maturity and 6 months payment frequency. Parameters:  $\kappa_{11} = 0.0630$ ,  $\theta_1 = 0.6709$ ,  $\theta_2 = 0.2903$ ,  $\sigma_1 = 0.2269$ ,  $\sigma_2 = 0.1688$ ,  $\alpha = 0.0636$  and start value  $X_0 = (1.2, 0.5)^\top$

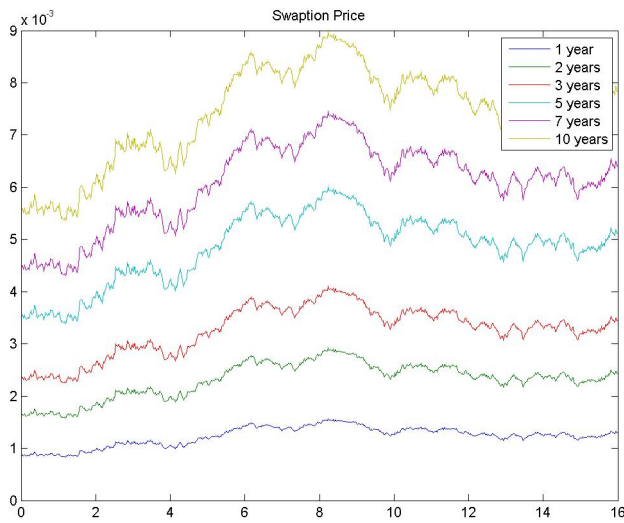


Figure 3.13: Swaption prices on 3 months swap options with 1-, 2-, 3-, 5-, 7- and 10 years maturity and 6 months payment frequency. Parameters:  $\kappa_{11} = 0.0630$ ,  $\theta_1 = 0.6709$ ,  $\theta_2 = 0.2903$ ,  $\sigma_1 = 0.2269$ ,  $\sigma_2 = 0.1688$ ,  $\alpha = 0.0636$  and start value  $X_0 = (1.2, 0.5)^\top$

### 3.6.3 Maximum Likelihood Estimation in Conjunction with UKF

The *state space model* corresponding to the LRSQ(1,1) model is given by the following measurement and process equations

$$Y_{t_k} = h(X_{t_k}, \Theta) + u_{t_k} \quad (3.46)$$

$$Z_{t_k} = \Phi_0 + \Phi_X X_{t_{k-1}} + w_{t_k} \quad (3.47)$$

where

- $0 \leq t_k \leq 16$  and  $t_k = k \cdot \frac{1}{52}$  for  $k = 0, 1, \dots, 16 \cdot 52$ .
- $Y_{t_k} = (Y_{t_k}^1 \ Y_{t_k}^2)^\top \in \mathbb{R}^{2 \times 6}$  is a vector of observable market swap rates and swaption prices at time  $t_k$  with maturity 1, 2, 3, 5, 7, 10 years, respectively.
- $X_{t_k} \in \mathbb{R}^2$  is the unobservable vector of state variables at time  $t_k$  and by assumption it is a Gaussian random vector, whose first two conditional moments are equal to the first two conditional moments of the factor process (3.39).
- $h = (h_1, h_2)^\top : \mathbb{R} \rightarrow \mathbb{R}^2$  is a vector-valued function where  $h_{1j}$  are defined according to (3.41) and  $h_{2j}$  according to (3.43).
- $u_{t_k}$  is the measurement noise and  $u_{t_k} \sim N(0, R)$  with  $R = \begin{pmatrix} \sigma_{swap}^2 & 0 \\ 0 & \sigma_{swpt}^2 \end{pmatrix}$ .
- $w_{t_k}$  is process noise and  $w_{t_k} \sim N(0, Q_{t_k})$ .
- $\Theta = (\kappa_{11}, \theta_1, \theta_2, \sigma_1, \sigma_2)^\top$  as well as  $\Phi_0, \Phi_X$  and  $Q_{t_k}$  are the unknown parameters which have to be estimated from the market data.

*Estimation of parameters  $\Phi_0, \Phi_X$  and  $Q_{t_k}$*

The assumption, that the first two conditional moments of the state variables  $X_{t_k}$  in the state space model and conditional moments of the factor process (which follow a CIR process) are identical, implies

$$\begin{aligned} \Phi_0 + \Phi_Z Z_{t_{k-1}} &= \mathbb{E} [Z_{t_k} \mid \mathcal{F}_{t_{k-1}}] = \theta (1 - e^{-\kappa(t_k - t_{k-1})}) + X_{t_{k-1}} e^{-\kappa(t_k - t_{k-1})} \\ Q_{t_k} &= \text{Var} [Z_{t_k} \mid \mathcal{F}_{t_{k-1}}]. \end{aligned}$$

From the first equality we obtain that the parameters  $\Phi_0, \Phi_X$  depend only on the parameters  $\kappa_{11}, \theta_1$  and  $\theta_2$ , and can be given in closed form by

$$\begin{aligned} \Phi_0 &= \theta (1 - e^{-\kappa/52}) \\ \Phi_X &= e^{-\kappa/52}. \end{aligned}$$

From the second equality, it follows that the variance of the process noise is equal to the conditional variance of the two dimensional CIR process (3.39), which can not be given in a closed form but it can be calculated numerically if the parameters  $\kappa_{11}, \theta_1, \theta_2, \sigma_1, \sigma_2$  and the previous state are known.

Therefore the *unknown parameters of the state space model* are only  $\kappa_{11}, \theta_1, \theta_2, \sigma_1, \sigma_2$ .

*Maximum likelihood estimation of parameters  $\Theta = (\kappa_{11} \theta_1 \theta_2 \sigma_1 \sigma_2)^\top$  and state variables  $X_t$*

To estimate the unknown parameters  $\Theta = (\kappa_{11}, \theta_1, \theta_2, \sigma_1, \sigma_2)^\top$ , we *apply the UKF* algorithm to the state space model (3.46)-(3.47). The Unscented Kalman Filter produces a recursive estimation for of the state variables  $X_{t_k}$  and at the same time constructs the log-likelihood function

$$\mathcal{L}(\Theta) = \frac{1}{2} \sum_{t=t_0}^{t_N} \left( 2 \log 2\pi + \log | \widehat{F}_{t|t-1} + (Y_t - \widehat{Y}_{t|t-1})^\top \widehat{F}_{t|t-1}^{-1} (Y_t - \widehat{Y}_{t|t-1}) \right),$$

where  $\widehat{Y}_{t|t-1}$  is the one-step-ahead forecast for  $Y_t$  at time  $t$  and  $\widehat{F}_{t|t-1}$  is the corresponding error covariance matrix. The *maximum likelihood estimation* for  $\Theta$  is obtained by

$$\widehat{\Theta} = \arg \max_{\Theta} \mathcal{L}(\Theta).$$

Applying the UKF to the model swap rates with 1 year maturity and the model swaption prices with 1 year maturity, calculated with the parameters  $\kappa_{11} = 0.0630, \theta_1 = 0.6709, \theta_2 = 0.2203, \sigma_1 = 0.2269, \sigma_2 = 0.1688$ , we obtain the maximum likelihood estimation for the parameters  $\kappa, \theta, \sigma$

estimated	true
$\widehat{\kappa}_{11} = 0.0640$	$\kappa_{11} = 0.0630$
$\widehat{\theta}_1 = 0.7021$	$\theta_1 = 0.6709$
$\widehat{\theta}_2 = 0.2204$	$\theta_2 = 0.2203$
$\widehat{\sigma}_1 = 0.2250$	$\sigma_1 = 0.2269$
$\widehat{\sigma}_2 = 0.1629$	$\sigma_2 = 0.1688$

as well as the UKF estimate for the factor process  $X$ . The simulates and estimated factor process are given in figure figure 3.14 and figure 3.15.

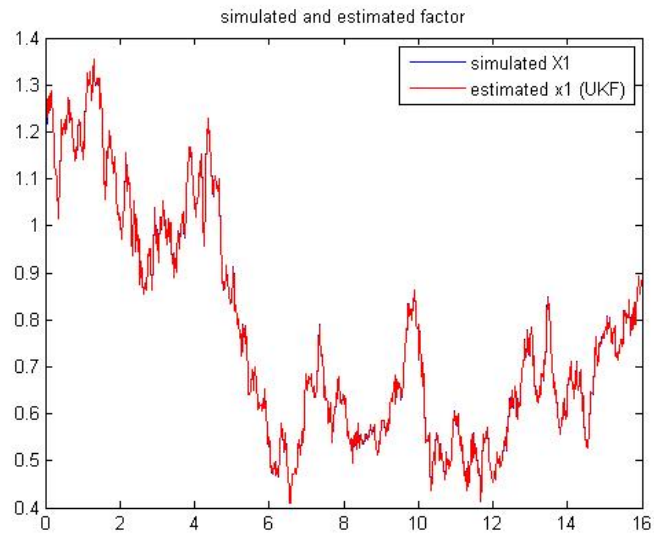


Figure 3.14: Estimated and simulated factor process  $X$  (UKF)

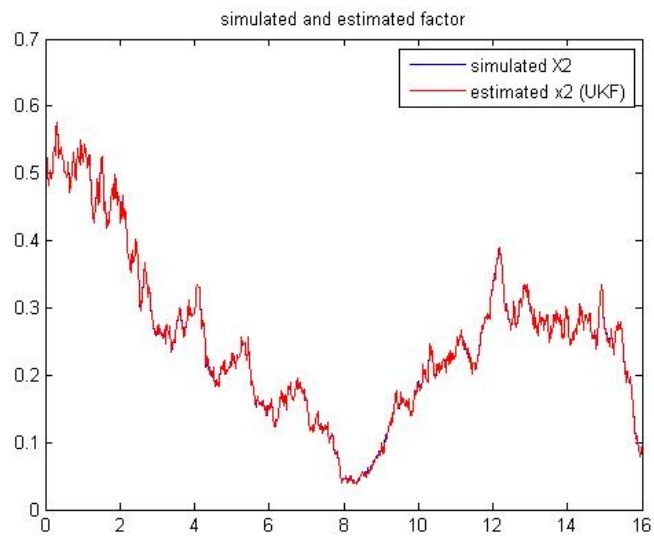


Figure 3.15: Estimated and simulated factor process  $X$  (UKF)

## 3.7 Estimation of LRSQ(3,1) Model Specification

### 3.7.1 Model Specification

The LRSQ(3,1) model is a four-dimensional linear-rational square root model, which has exactly three term structure factors and one unspanned stochastic volatility factor. According to LRSQ(3,1) specification (see Example 2.15), the factor process  $X$  follows the four-dimensional square root diffusion process (CIR process)

$$\begin{aligned} dX_t &= \kappa(\theta - X_t)dt + \text{Diag}(\sigma_1\sqrt{X_{1t}}, \sigma_2\sqrt{X_{2t}}, \sigma_3\sqrt{X_{3t}}, \sigma_4\sqrt{X_{4t}})dB_t \\ X_0 &= x \in \mathbb{R}^4 \end{aligned} \quad (3.48)$$

and the state space density is given by

$$\zeta_t = e^{-\alpha t}(1 + \mathbf{1}_4^\top X_t), \quad (3.49)$$

where

$$\bullet \kappa = \begin{pmatrix} \kappa_{11} & \kappa_{12} & \kappa_{13} & 0 \\ \kappa_{21} & \kappa_{22} & \kappa_{23} & \kappa_{21} \\ \kappa_{31} & \kappa_{32} & \kappa_{33} & \kappa_{31} \\ 0 & 0 & 0 & \kappa_{11} \end{pmatrix} \in \mathbb{R}^{4 \times 4}, \quad \theta = \begin{pmatrix} \theta_1 \\ \theta_2 \\ \theta_3 \\ \theta_4 \end{pmatrix} \in \mathbb{R}^4, \text{ and}$$

$\sigma_1, \sigma_2, \sigma_3, \sigma_4 \in \mathbb{R}_+$  with  $\sigma_1 \neq \sigma_4$  are unknown real-valued parameters;

- $X = (X_{1t}, X_{2t}, X_{3t}, X_{4t})_{t \geq 0}^\top$  denotes the vector of unobservable state space variables with values in  $\mathbb{R}_+^4$ ;
- $B = (B_{1t}, B_{2t}, B_{3t}, B_{4t})_{t \geq 0}^\top$  is a four-dimensional Brownian motion;
- $\alpha = \alpha^* = \max \{ \mathbf{1}_4^\top \kappa \theta, -\mathbf{1}_4^\top \kappa_{\cdot 1}, -\mathbf{1}_4^\top \kappa_{\cdot 2}, -\mathbf{1}_4^\top \kappa_{\cdot 3}, -\mathbf{1}_4^\top \kappa_{\cdot 4} \}$   
ensures that the short rate in the LRSQ(3,1) model stays non-negative for any time  $t$  (see Theorem 2.12).

The *estimation of LRSQ(3,1)* specification includes an estimation of the *unknown parameters*

$$\Theta = (\kappa_{11}, \kappa_{12}, \kappa_{13}, \kappa_{21}, \kappa_{22}, \kappa_{23}, \kappa_{31}, \kappa_{32}, \kappa_{33}, \theta_1, \theta_2, \theta_3, \theta_4, \sigma_1, \sigma_2, \sigma_3, \sigma_4)$$

as well as an estimation of the *unobservable state variables*  $X_t = (X_{1t}, X_{2t}, X_{3t}, X_{4t})^\top$  on specific time points from the market swap rates and the market swaption prices. As market data we use the model swap rates and the model swaption prices with added white noise. We estimate the unknown parameters by Maximum likelihood approach in conjunction with Unscented Kalman Filter.

### 3.7.2 Model Data

We generate a panel data set for the LRSQ(3,1) model including weekly data for swaps and swaptions over the time period of sixteen years. Each panel data row consists of model rates on spot-starting swap contracts with maturities of one, two, three, five, seven and ten years, respectively, and six months payment frequency as well as of models prices on swaptions with three-month options maturities, the same six swap maturities, and strikes equal to the forward swap rates. The model swap rates and swaption prices in the LRSQ(3,1) model are functions of the four dimensional factor process  $(X_t)_{t \geq 0}$  and depend on the parameter vector  $\Theta$ . To obtain the model data we first simulate the four dimensional CIR process  $(X_t)_{t \geq 0}$  for a given parameters vector  $\Theta$  and a given start value  $X_0$  and then calculate the model swap rates and the model swaption prices according to (1.8) and (2.24). To obtain more realistic data we add a white noise to the model swap rates and the model swaption prices.

*Calculation of model swap rates and model swaption prices with noise*

We calculate the *model swap rates* and *model swaption prices* on the time grid

$$0 = t_0 < t_1 < \dots < t_N = 16, \quad t_k = k \cdot 1/52 \text{ for } k = 1, \dots, N = 16 \cdot 52$$

for the parameters

$$\begin{aligned} \Theta &= (\kappa_{11}, \kappa_{12}, \kappa_{13}, \kappa_{21}, \kappa_{22}, \kappa_{23}, \kappa_{31}, \kappa_{32}, \kappa_{33}, \theta_1, \theta_2, \theta_3, \theta_4, \sigma_1, \sigma_2, \sigma_3, \sigma_4) \\ &= (0.0630, 0.0000, 0.0000, -0.1266, 0.4377, 0.0000, 0.0000, -0.5012, 0.1652, \\ &\quad 0.6709, 0.2903, 0.8810, 0.3275, \\ &\quad 0.2269, 0.1688, 0.1229, 1.8097) \end{aligned}$$

and the start value

$$X_0 = (1.2, 0.5, 0.8, 0.1)^\top.$$

We generate  $6(N+1)$  i.i.d. Gaussian variables  $Z_{t_k, T}^{swap} \sim \mathcal{N}(0, 1e - 8)$  and  $6(N+1)$  i.i.d. Gaussian variables  $Z_{t_k, T}^{swpt} \sim \mathcal{N}(0, 1e - 10)$  and add them to the model swap rates and to the model swaption prices, respectively.

*Calculation scheme*

1. set the time grid

$$0 = t_0 < t_1 < \dots < t_N = 16,$$

where  $N + 1 = 16 \cdot 52 + 1$  is the number of the simulated weekly observations and  $t_k = k \cdot \frac{1}{52}$  for  $k = 1, \dots, N$ .

2. simulate a path of the four-dimensional CIR process (3.48) with start value  $X_0 = (1.2 \ 0.5 \ 0.8 \ 0.1)^\top$  on the time grid  $0 = t_0 < t_1 < \dots < t_N = 16$  using Euler-Maruyama approach.
3. generate  $6(N + 1)$  random variables  $Z_{t_k, T}^{swap} \stackrel{iid}{\sim} \mathcal{N}(0, 1e - 8)$ .
4. generate  $6(N + 1)$  random variables  $Z_{t_k, T}^{swpt} \stackrel{iid}{\sim} \mathcal{N}(0, 1e - 10)$ .
5. calculate for each  $t_k$  the **noisy** swap rates of spot-starting swaps with maturity  $T = 1, 2, 3, 5, 7, 10$  and payment dates  $t_k = T_0 < T_1 < \dots < T_n = T$  according to

$$S_{t_k, T} = \frac{P(t_k, T_0) - P(t_k, T_n)}{\sum_{i=1}^n \Delta_i P(t_k, T_i)} + Z_{t_k, T}^{swap}, \quad (3.50)$$

where

$$P(t_k, T_i) = \frac{(\phi + \psi^\top \theta) e^{-\alpha \tau} + \psi^\top e^{-(\alpha + \kappa) \tau} (X_{t_k} - \theta)}{\phi + \psi^\top X_{t_k}}, \quad (3.51)$$

$$\phi = 1, \psi = (1 \ 1 \ 1 \ 1)^\top, \tau = T_i - t, \Delta_i = T_i - T_{i-1} = 1/2,$$

$$n = 2T = 2, 4, 6, 10, 14 \text{ or } 20.$$

6. calculate for each  $t_k$  the *noisy* prices of swaptions on three-month swap options with maturity  $T = 1, 2, 3, 5, 7, 10$ , payment dates  $t_k = T_0 < T_1 < \dots < T_n = T$  and strikes equal to the forward swap rates according to

$$\Pi_{t_k, T}^{swpt} = \frac{1}{\zeta_{t_k} \pi} \int_0^\infty \text{Re} \left[ \frac{\widehat{q}(\mu + i\lambda)}{(\mu + i\lambda)^2} \right] d\lambda + Z_{t_k, T}^{swpt} \quad (3.52)$$

for some  $\mu > 0$  such that  $\widehat{q}(\mu) < \infty$ , where

$$\widehat{q}(\mu + i\lambda) = \mathbb{E} \left[ e^{(\mu + i\lambda) p_{swpt}(X_{T_0})} \mid \mathcal{F}_t \right] = e^{\Phi(T_0 - t) + X_t^\top \Psi(T_0 - t)}, \quad (3.53)$$

given  $X_t = x$  and  $\Phi : \mathbb{R}_+ \rightarrow \mathbb{C}, \Psi : \mathbb{R}_+ \rightarrow \mathbb{C}^4$  solve the system of Ricatti differential equations

$$\begin{aligned} \Phi'(\tau) &= \kappa \theta^\top \Psi(\tau) \\ \Psi'(\tau) &= -\kappa^\top \Psi(\tau) + \frac{1}{2} \Psi(\tau)^\top \text{Diag}(\sigma_1^2, \sigma_2^2, \sigma_3^2, \sigma_4^2) \Psi(\tau) \end{aligned} \quad (3.54)$$

with initial condition

$$\Phi(0) = u = (\mu + i\lambda) \sum_{i=0}^n c_i e^{-\alpha T_i} (\phi + \psi^\top \theta - \psi^\top e^{-\kappa(T_i - T_0)} \theta)$$

$$\Psi(0) = v = (\mu + i\lambda) \sum_{i=0}^n c_i e^{-\alpha T_i} \psi^\top e^{-\kappa(T_i - T_0)}$$

where

$$c_0 = 1, c_n = -1 - \Delta K, c_i = -\Delta K, i = 1, \dots, n-1$$

$$K = S_{T_0}, \tau = T_0 - t_k = 1/4, \Delta = T_i - T_{i-1} = 1/2 \text{ and}$$

$$\phi = 1, \psi = (1 \ 1 \ 1 \ 1)^\top.$$

The figure 3.16, figure 3.17, figure 3.18 and figure 3.19 show the simulated CIR process  $X$ , the model short rate, the noisy model swap rates and the noisy model swaption prices with parameters  $\kappa_{11} = 0.0630, \kappa_{12} = 0, \kappa_{13} = 0, \kappa_{21} = -0.1266, \kappa_{22} = 0.4377, \kappa_{23} = 0, \kappa_{31} = 0, \kappa_{32} = -0.5012, \kappa_{33} = 0.1652, \theta_1 = 0.6703, \theta_2 = 0.2903, \theta_3 = 0.8810, \theta_4 = 0.3275, \sigma_1 = 0.2269, \sigma_2 = 0.1688, \sigma_3 = 0.1229, \sigma_4 = 1.8097, \alpha = 0.0636$ .

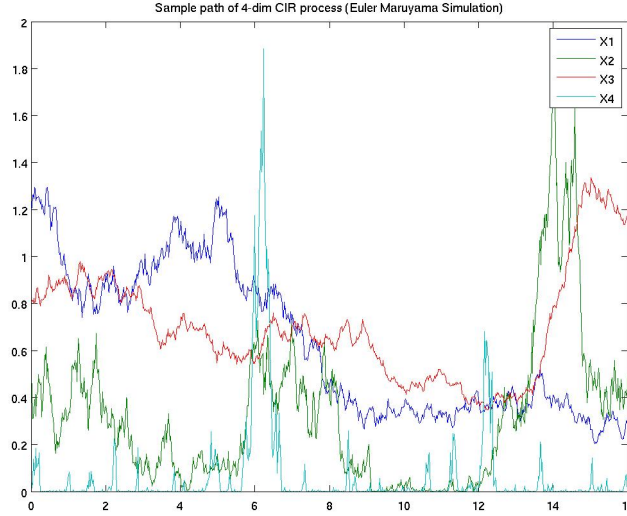


Figure 3.16: Simulated path of four dimensional CIR process  $X = (X_{1t}, X_{2t}, X_{3t}, X_{4t})^\top$  with parameters  $\kappa_{11} = 0.0630, \kappa_{12} = \kappa_{13} = 0, \kappa_{21} = -0.1266, \kappa_{22} = 0.4377, \kappa_{23} = 0, \kappa_{31} = 0, \kappa_{32} = -0.5012, \kappa_{33} = 0.1652, \theta_1 = 0.6703, \theta_2 = 0.2903, \theta_3 = 0.8810, \theta_4 = 0.3275, \sigma_1 = 0.2269, \sigma_2 = 0.1688, \sigma_3 = 0.1229, \sigma_4 = 1.8097, \alpha = 0.0636$  and start value  $X_0 = (1.2, 0.5, 0.8, 0.1)^\top$



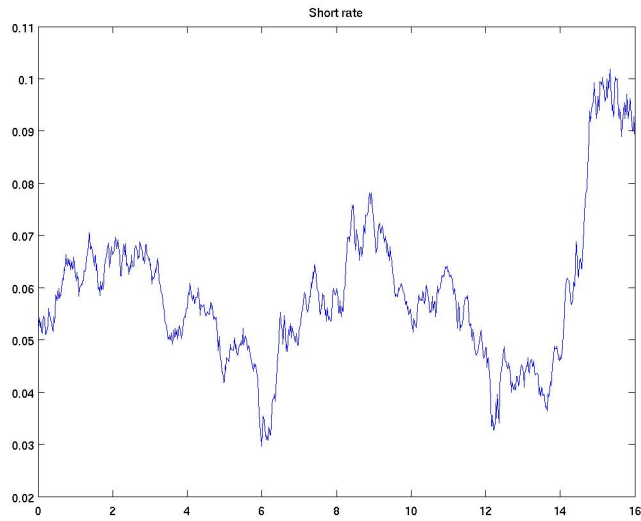


Figure 3.17: Short rate in LRSQ(3,1) model

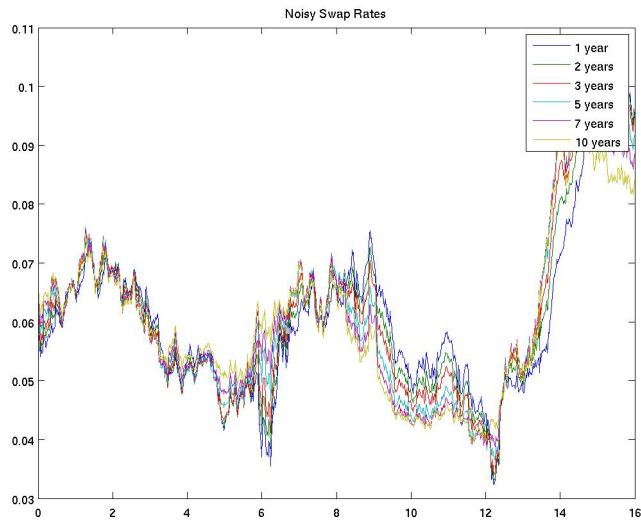


Figure 3.18: Noisy swap rates of spot-starting swaps with 1-, 2-, 3-, 5-, 7- and 10 years maturity and 6 months payment frequency.

Parameters:  $\kappa_{11} = 0.0630$ ,  $\kappa_{12} = \kappa_{13} = 0$ ,  $\kappa_{21} = -0.1266$ ,  $\kappa_{22} = 0.4377$ ,  $\kappa_{23} = 0$ ,  $\kappa_{31} = 0$ ,  $\kappa_{32} = -0.5012$ ,  $\kappa_{33} = 0.1652$ ,  $\theta_1 = 0.6703$ ,  $\theta_2 = 0.2903$ ,  $\theta_3 = 0.8810$ ,  $\theta_4 = 0.3275$ ,  $\sigma_1 = 0.2269$ ,  $\sigma_2 = 0.1688$ ,  $\sigma_3 = 0.1229$ ,  $\sigma_4 = 1.8097$ ,  $\alpha = 0.0636$  and start value  $X_0 = (1.2, 0.5, 0.8, 0.1)^\top$

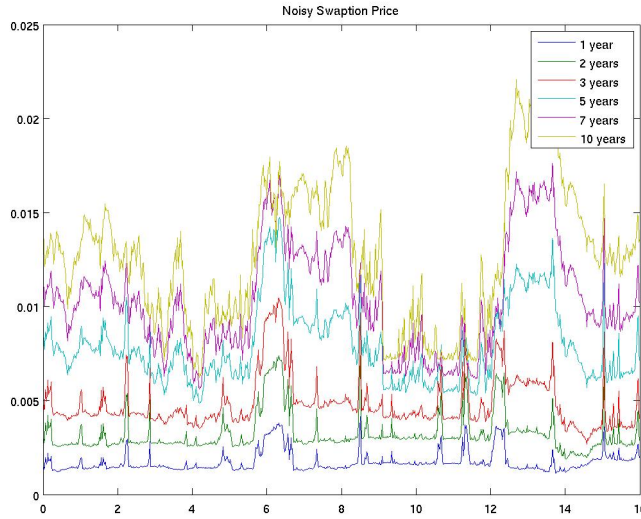


Figure 3.19: Noisy swaption prices on 3 months swap options with 1-, 2-, 3-, 5-, 7- and 10 years maturity and 6 months payment frequency. Parameters:  $\kappa_{11} = 0.0630$ ,  $\kappa_{12} = \kappa_{13} = 0$ ,  $\kappa_{21} = -0.1266$ ,  $\kappa_{22} = 0.4377$ ,  $\kappa_{23} = 0$ ,  $\kappa_{31} = 0$ ,  $\kappa_{32} = -0.5012$ ,  $\kappa_{33} = 0.1652$ ,  $\theta_1 = 0.6703$ ,  $\theta_2 = 0.2903$ ,  $\theta_3 = 0.8810$ ,  $\theta_4 = 0.3275$ ,  $\sigma_1 = 0.2269$ ,  $\sigma_2 = 0.1688$ ,  $\sigma_3 = 0.1229$ ,  $\sigma_4 = 1.8097$ ,  $\alpha = 0.0636$  and start value  $X_0 = (1.2, 0.5, 0.8, 0.1)^\top$

### 3.7.3 Maximum Likelihood Estimation in Conjunction with UKF

The *state space model*, corresponding to the LRSQ(3,1) model, is given by the following measurement and process equations

$$Y_{t_k} = h(X_{t_k}, \Theta) + u_{t_k} \quad (3.55)$$

$$Z_{t_k} = \Phi_0 + \Phi_X X_{t_{k-1}} + w_{t_k} \quad (3.56)$$

where

- $0 \leq t_k \leq 16$  and  $t_k = k \cdot \frac{1}{52}$  for  $k = 0, 1, \dots, 16 \cdot 52$ .
- $Y_{t_k} = (Y_{t_k}^1, Y_{t_k}^2, Y_{t_k}^3, Y_{t_k}^4)^\top \in \mathbb{R}^4$  is a vector of observable market swap rates and swaption prices at time  $t_k$  with maturity 1, 2, 3, 5, 7, 10 years, respectively. Note that we need at least four measurement equations to estimate four unknown variable.
- $X_{t_k} \in \mathbb{R}^4$  is a vector of 4 unobservable state variable at time  $t_k$  and by assumption it is a multidimensional Gaussian random vector, whose first

two conditional moments are equal to the first two conditional moments of the factor process (3.48).

- $h = (h_i)_{1 \leq i \leq 4} : \mathbb{R}^4 \rightarrow \mathbb{R}^4$  is a vector-valued function where  $h_1$  and  $h_2$  are defined according to (3.50) and  $h_3$  and  $h_4$  according to (3.52).
- $u_{t_k}$  is the measurement noise and  $u_{t_k} \sim N(0, R)$  with  $R = \begin{pmatrix} \sigma_{swap}^2 Id_2 & 0 \\ 0 & \sigma_{swpt}^2 Id_2 \end{pmatrix}$ .
- $w_{t_k}$  is process noise and  $w_{t_k} \sim N(0, Q_{t_k})$ .
- $\Theta = (\kappa_{11}, \kappa_{12}, \kappa_{13}, \kappa_{21}, \kappa_{22}, \kappa_{23}, \kappa_{31}, \kappa_{32}, \kappa_{33}, \theta_1, \theta_2, \theta_3, \theta_4, \sigma_1, \sigma_2, \sigma_3, \sigma_4)$  as well as  $\Phi_0, \Phi_X$  and  $Q_{t_k}$  are the unknown parameters which have to be estimated from the market data.

*Estimation of parameters  $\Phi_0, \Phi_X$  and  $Q_{t_k}$*

The assumption that the first two conditional moments of the state variables  $X_{t_k}$  in the state space model and conditional moments of the factor process (which is a CIR process) are identical, implies

$$\begin{aligned} \Phi_0 + \Phi_Z Z_{t_{k-1}} &= \mathbb{E} [Z_{t_k} | \mathcal{F}_{t_{k-1}}] = \theta (1 - e^{-\kappa(t_k - t_{k-1})}) + X_{t_{k-1}} e^{-\kappa(t_k - t_{k-1})} \\ Q_{t_k} &= \text{Var} [Z_{t_k} | \mathcal{F}_{t_{k-1}}]. \end{aligned}$$

From the first equality, we obtain that the parameters  $\Phi_0, \Phi_X$  depend on the parameters  $\kappa_{11}, \kappa_{12}, \kappa_{13}, \kappa_{21}, \kappa_{22}, \kappa_{23}, \kappa_{31}, \kappa_{32}, \kappa_{33}, \theta_1, \theta_2, \theta_3, \theta_4$  and can be given in a closed form by

$$\begin{aligned} \Phi_0 &= \theta (1 - e^{-\kappa/52}) \\ \Phi_X &= e^{-\kappa/52}. \end{aligned}$$

From the second equality, it follows that the variance of the process noise is equal to the conditional variance of the four dimensional CIR process (3.48) which can be calculated numerically, if all parameters in  $\Theta$  and the previous state are known.

Therefore the state space model parameters are

$$\Theta = (\kappa_{11}, \kappa_{12}, \kappa_{13}, \kappa_{21}, \kappa_{22}, \kappa_{23}, \kappa_{31}, \kappa_{32}, \kappa_{33}, \theta_1, \theta_2, \theta_3, \theta_4, \sigma_1, \sigma_2, \sigma_3, \sigma_4).$$

*Maximum likelihood estimation (MLE) for  $\Theta$*

To obtain the MLE for the unknown parameter  $\Theta$  we *first* apply the UKF to the state space model (3.55)-(3.56) to construct the log-likelihood function

$$\mathcal{L}(\Theta) = \frac{1}{2} \sum_{t=t_0}^{t_N} \left( 4 \log 2\pi + \log | \hat{F}_{t|t-1} + (Y_t - \hat{Y}_{t|t-1})^\top \hat{F}_{t|t-1}^{-1} (Y_t - \hat{Y}_{t|t-1}) \right),$$

where  $\hat{Y}_{t|t-1}$  is a one-step-ahead forecast for  $Y_t$  at time  $t$  and  $\hat{F}_{t|t-1}$  is the corresponding error covariance matrix. *Then* we find numerically the parameter value  $\hat{\Theta}$  which minimize the negative log-likelihood function  $\mathcal{L}(\Theta)$  by

$$\hat{\Theta} = \arg \max_{\Theta} \mathcal{L}(\Theta).$$

*UKF state estimation given parameter  $\Theta$*

We first consider a state space model in which all parameters are known and estimate by UKF the state variables in the model from the the noisy model swap rates and with maturity 1- and 5 years and the noisy model swaption prices with the same maturity.

Figure 3.20 - 3.23 show the estimated and simulated factor process with parameters  $\kappa_{11} = 0.0630$ ,  $\kappa_{12} = \kappa_{13} = 0$ ,  $\kappa_{21} = -0.1266$ ,  $\kappa_{22} = 0.4377$ ,  $\kappa_{23} = 0$ ,  $\kappa_{31} = 0$ ,  $\kappa_{32} = -0.5012$ ,  $\kappa_{33} = 0.1652$ ,  $\theta_1 = 0.6703$ ,  $\theta_2 = 0.2903$ ,  $\theta_3 = 0.8810$ ,  $\theta_4 = 0.3275$ ,  $\sigma_1 = 0.2269$ ,  $\sigma_2 = 0.1688$ ,  $\sigma_3 = 0.1229$ ,  $\sigma_4 = 1.8097$ ,  $\alpha = 0.0636$  and start value  $X_0 = (1.2, 0.5, 0.8, 0.1)^\top$ .

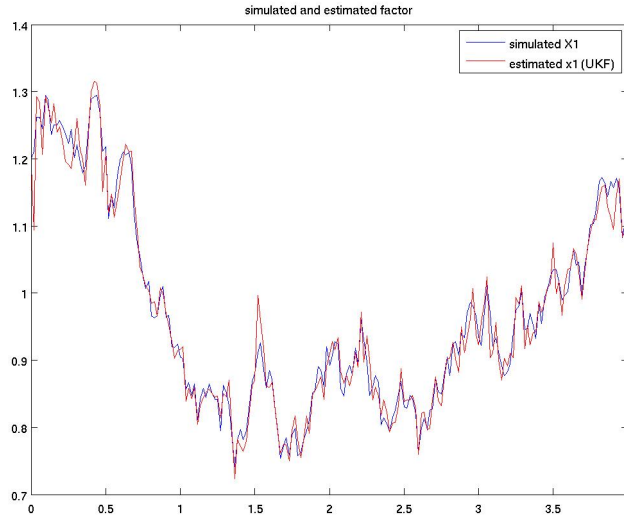


Figure 3.20: Estimated and simulated factor process  $X_1$  (UKF)

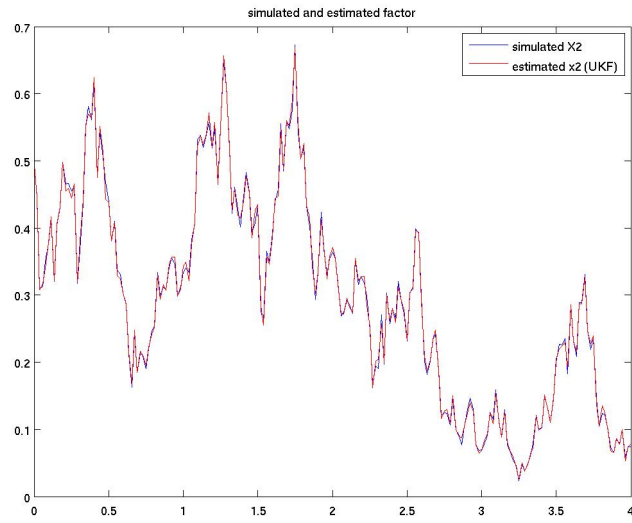


Figure 3.21: Estimated and simulated factor process  $X_2$  (UKF)

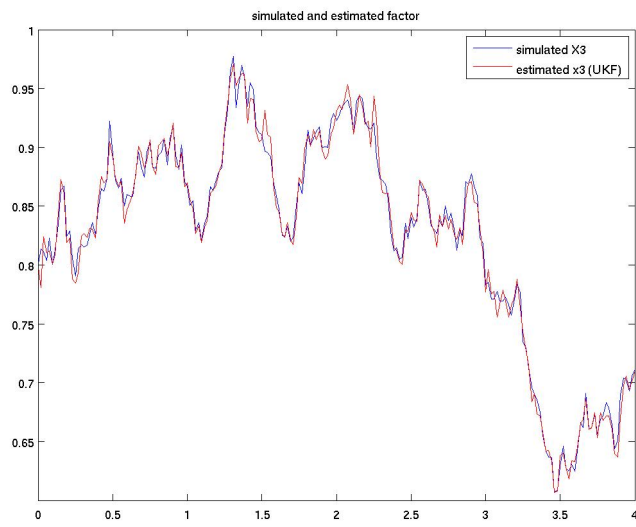


Figure 3.22: Estimated and simulated factor process  $X_3$  (UKF)

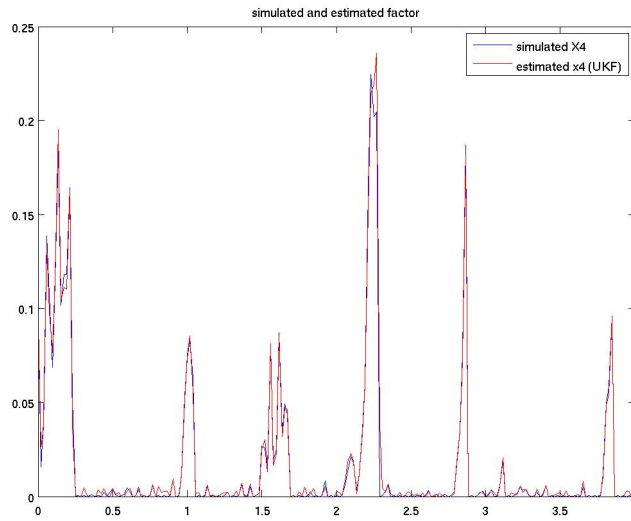


Figure 3.23: Estimated and simulated factor process  $X_4$  (UKF)

*Maximum likelihood estimation for  $\Theta$*

In practice, the model parameters are unknown and have to be estimated from the market prices.

We assume that in the state space model (3.55)-(3.56) both the parameters and the state variables are unknown. We apply the Maximum likelihood approach in conjunction with UKF to estimate the parameter  $\Theta$  from the market data. As market data we use the noisy model swap rates with maturity 1- and 5 years and the noisy model swaption prices with the same maturity.

### Example 1

$$\Theta^0 = \Theta^*$$

$$\hat{\Theta} = \text{fminsearch}(-\mathcal{L}(\Theta \mid \mathbf{x}^{209}), \Theta^0, \text{options})$$

In the table below are given the MLE  $\hat{\Theta}$ , the true model parameter  $\Theta^*$  as well as the start value  $\Theta^0$  for the numerical minimization of the likelihood function. Figure (3.24)-(3.27) show the estimated and simulated factor process.

estimate $\hat{\Theta}$	true $\Theta^*$	start $\Theta^0$
$\hat{\kappa}_{11} = 0.0640$	$\kappa_{11} = 0.0630$	$\kappa_{11}^0 = 0.0630$
$\hat{\kappa}_{12} = 0.0001$	$\kappa_{12} = 0.0000$	$\kappa_{12}^0 = 0.0000$
$\hat{\kappa}_{13} = 0.0001$	$\kappa_{13} = 0.0000$	$\kappa_{13}^0 = 0.0000$
$\hat{\kappa}_{21} = -0.1274$	$\kappa_{21} = -0.1266$	$\kappa_{21}^0 = -0.1266$
$\hat{\kappa}_{22} = 0.4393$	$\kappa_{22} = 0.4377$	$\kappa_{22}^0 = 0.4377$
$\hat{\kappa}_{23} = 0.0001$	$\kappa_{23} = 0.0000$	$\kappa_{23}^0 = 0.0000$
$\hat{\kappa}_{31} = 0.0001$	$\kappa_{31} = 0.0000$	$\kappa_{31}^0 = 0.0000$
$\hat{\kappa}_{32} = -0.5000$	$\kappa_{32} = -0.5012$	$\kappa_{32}^0 = -0.5012$
$\hat{\kappa}_{33} = 0.1626$	$\kappa_{33} = 0.1652$	$\kappa_{33}^0 = 0.1652$
$\hat{\theta}_1 = 0.6730$	$\theta_1 = 0.6709$	$\theta_1^0 = 0.6709$
$\hat{\theta}_2 = 0.2918$	$\theta_2 = 0.2903$	$\theta_2^0 = 0.2903$
$\hat{\theta}_3 = 0.8729$	$\theta_3 = 0.8810$	$\theta_3^0 = 0.8810$
$\hat{\theta}_4 = 0.3220$	$\theta_4 = 0.3275$	$\theta_4^0 = 0.3275$
$\hat{\sigma}_1 = 0.2261$	$\sigma_1 = 0.2269$	$\sigma_1^0 = 0.2269$
$\hat{\sigma}_2 = 0.6905$	$\sigma_2 = 0.6882$	$\sigma_2^0 = 0.6882$
$\hat{\sigma}_3 = 0.1216$	$\sigma_3 = 0.1229$	$\sigma_3^0 = 0.1229$
$\hat{\sigma}_4 = 1.8389$	$\sigma_4 = 1.8097$	$\sigma_4^0 = 1.8097$
$\hat{\alpha} = 0.0631$	$\alpha = 0.0630$	$\alpha^0 = 0.0630$
$\mathcal{L}(\hat{\Theta} \mid \mathbf{x}^{209}) = 5472.38$	$\mathcal{L}(\Theta^* \mid \mathbf{x}^{209}) = 5367.91$	$\mathcal{L}(\Theta^0 \mid \mathbf{x}^{209}) = 5367.91$

*Remark:* The sample period consists of 209 simulated weekly observations.

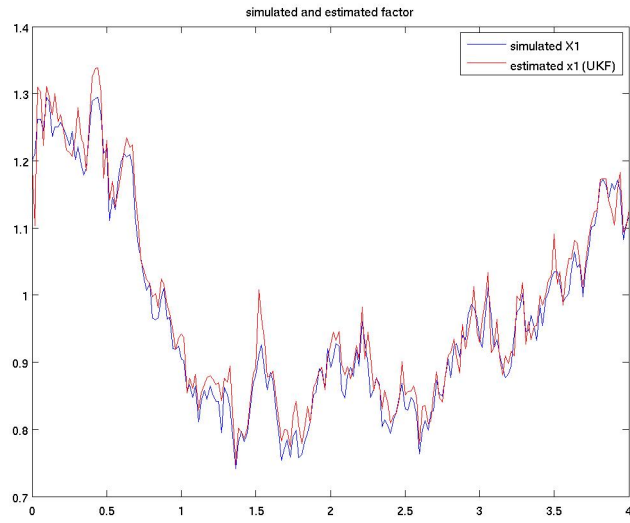


Figure 3.24: Estimated and simulated factor process  $X_1$  (UKF)

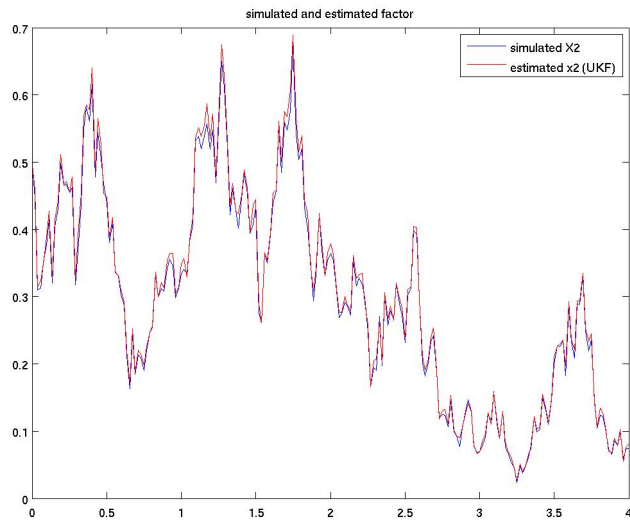


Figure 3.25: Estimated and simulated factor process  $X_2$  (UKF)



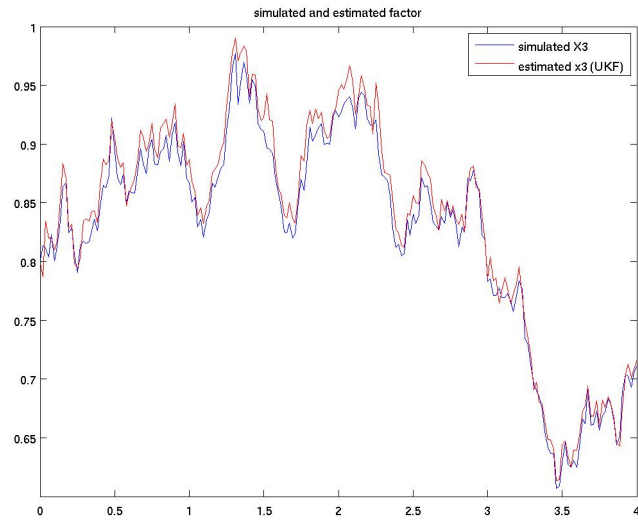


Figure 3.26: Estimated and simulated factor process  $X_3$  (UKF)

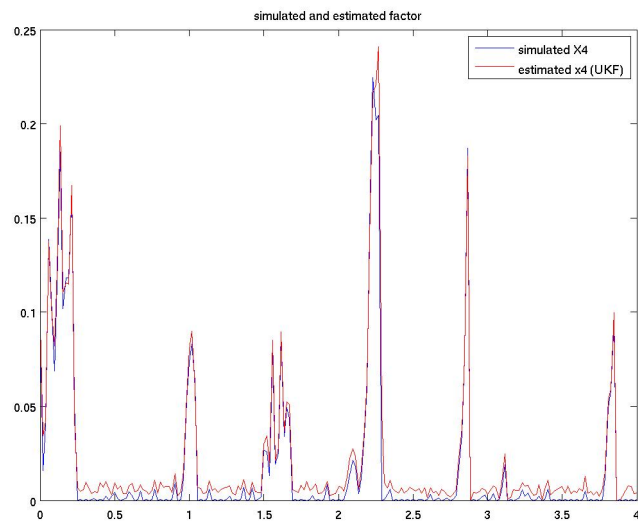


Figure 3.27: Estimated and simulated factor process  $X_4$  (UKF)

## Example 2

$$\Theta^0 \approx \Theta^*$$

$$\hat{\Theta} = \text{fminsearch}(-\mathcal{L}(\Theta \mid \mathbf{x}^{209}), \Theta^0)$$

The MLE estimation for the parameter  $\Theta$  is given in the table below. Figure (3.28)-(3.31) show the estimated and the simulated factor process.

estimate $\hat{\Theta}$	true $\Theta^*$	start $\Theta^0$
$\hat{\kappa}_{11} = 0.05925$	$\kappa_{11} = 0.0630$	$\kappa_{11}^0 = 0.0600$
$\hat{\kappa}_{12} = 0.00003$	$\kappa_{12} = 0.0000$	$\kappa_{12}^0 = 0.0000$
$\hat{\kappa}_{13} = 0.00006$	$\kappa_{13} = 0.0000$	$\kappa_{13}^0 = 0.0000$
$\hat{\kappa}_{21} = -0.12189$	$\kappa_{21} = -0.1266$	$\kappa_{21}^0 = -0.1200$
$\hat{\kappa}_{22} = 0.43090$	$\kappa_{22} = 0.4377$	$\kappa_{22}^0 = 0.4300$
$\hat{\kappa}_{23} = 0.000021$	$\kappa_{23} = 0.0000$	$\kappa_{23}^0 = 0.0000$
$\hat{\kappa}_{31} = 0.00006$	$\kappa_{31} = 0.0000$	$\kappa_{31}^0 = 0.0000$
$\hat{\kappa}_{32} = -0.49779$	$\kappa_{32} = -0.5012$	$\kappa_{32}^0 = -0.5000$
$\hat{\kappa}_{33} = 0.16166$	$\kappa_{33} = 0.1652$	$\kappa_{33}^0 = 0.1600$
$\hat{\theta}_1 = 0.6702$	$\theta_1 = 0.6709$	$\theta_1^0 = 0.6700$
$\hat{\theta}_2 = 0.2922$	$\theta_2 = 0.2903$	$\theta_2^0 = 0.2900$
$\hat{\theta}_3 = 0.8898$	$\theta_3 = 0.8810$	$\theta_3^0 = 0.8800$
$\hat{\theta}_4 = 0.3241$	$\theta_4 = 0.3275$	$\theta_4^0 = 0.3200$
$\hat{\sigma}_1 = 0.2232$	$\sigma_1 = 0.2269$	$\sigma_1^0 = 0.2200$
$\hat{\sigma}_2 = 0.6588$	$\sigma_2 = 0.6882$	$\sigma_2^0 = 0.6800$
$\hat{\sigma}_3 = 0.1213$	$\sigma_3 = 0.1229$	$\sigma_3^0 = 0.1200$
$\hat{\sigma}_4 = 1.7790$	$\sigma_4 = 1.8097$	$\sigma_4^0 = 1.8000$
$\hat{\alpha} = 0.06692$	$\alpha = 0.0630$	$\alpha^0 = 0.0700$
$\mathcal{L}(\hat{\Theta} \mid \mathbf{x}^{209}) = 5472.38$	$\mathcal{L}(\Theta^* \mid \mathbf{x}^{209}) = 5469.18$	$\mathcal{L}(\Theta^0 \mid \mathbf{x}^{209}) = 5319.79$

*Remark:* The sample period consists of 209 simulated weekly observation.

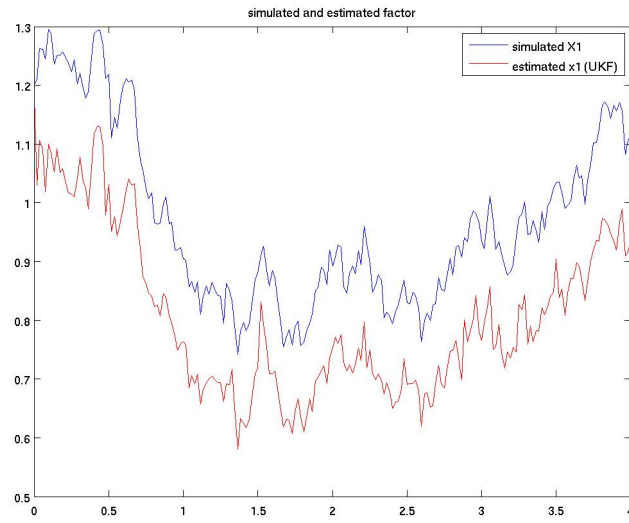


Figure 3.28: Estimated and simulated factor process  $X_1$  (UKF)

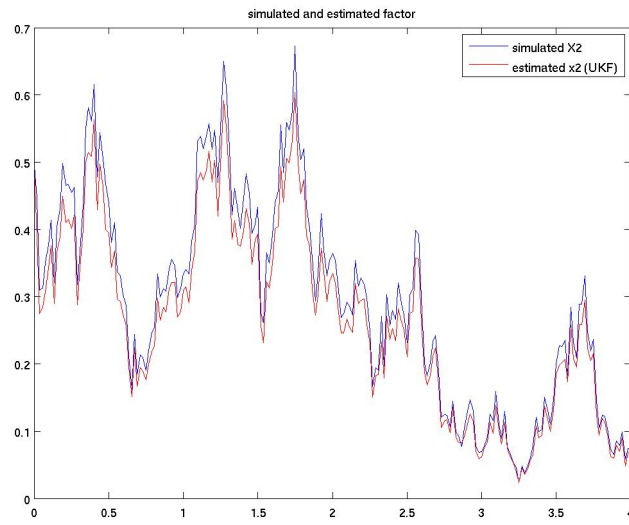


Figure 3.29: Estimated and simulated factor process  $X_2$  (UKF)

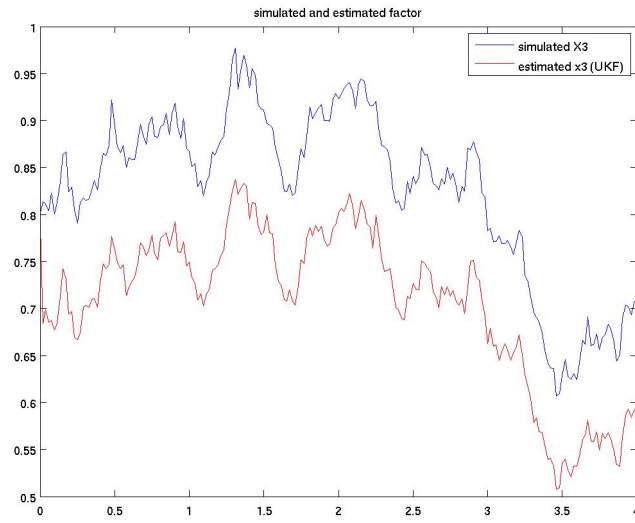


Figure 3.30: Estimated and simulated factor process  $X_3$  (UKF)

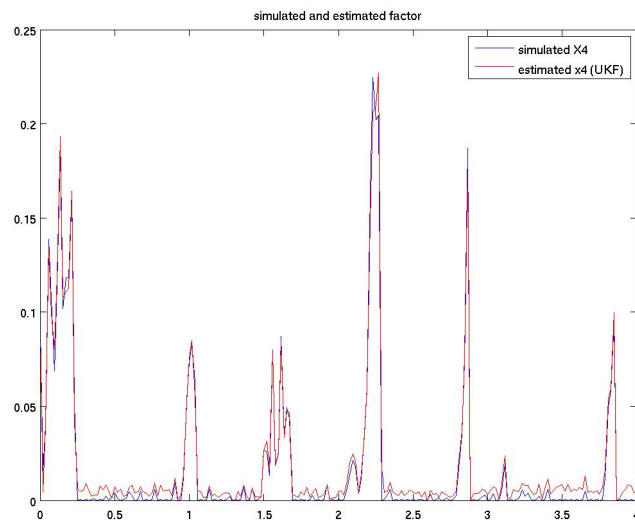


Figure 3.31: Estimated and simulated factor process  $X_4$  (UKF)

# Appendix A

## Stochastic calculus

Assume that  $X = (X_t)_{t \geq 0}$  is a stochastic process defined on the filtered probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ .

**Definition A.1 (Martingale).** A stochastic process  $X = (X_t)_{t \geq 0}$  is called **martingale** with respect to the filtration  $(\mathcal{F}_t)_{t \geq 0}$ , if  $E(|X_t|) < \infty$  for all  $t \geq 0$ ,  $X$  is adapted and

$$E(X_t | \mathcal{F}_s) = X_s \quad \mathbb{P}\text{-a.s. for all } 0 \leq s \leq t.$$

Examples: 1)  $X_t = E(Y | \mathcal{F}_t)$ , where  $Y$  is a random variable with  $E(|Y|) < \infty$ ;  
2)  $B = (B_t)_{t \geq 0}$ ;

**Definition A.2 (Itô process).** A  $d$ -dimensional stochastic process  $X$  is said to be an **Itô process**, if the local increments are on the form

$$dX_t = \mu_t dt + \sigma_t dB_t \tag{A.1}$$

where  $B = (B_t)_{t \geq 0}$  is assumed to be a  $d$ -dimensional standard Brownian motion on  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , the drift is a  $d$ -dimensional process  $\mu = (\mu_t)_{t \geq 0}$ , and the sensitivity towards the shock is a stochastic processes  $\sigma = (\sigma_t)_{t \geq 0}$  taking  $d \times d$  matrices as values. [14]

**Lemma A.3 (Itô's Formula).** Let  $X = (X_1, \dots, X_d)^\top$  be a continuous  $d$ -dimensional semimartingale, taking  $\mathbb{P}$ -a.s values in an open set  $U \subset \mathbb{R}^d$ , and let  $f : U \rightarrow \mathbb{R}$  be a twice continuously differentiable function ( $f \in C^2(U)$ ). Then the process  $f(X)$  is a continuous semimartingale and  $\mathbb{P}$ -a.s

$$f(X_t) = f(X_0) + \sum_{i=1}^d \int_0^t \frac{\partial f}{\partial x_i}(X_s) dX_s^i + \frac{1}{2} \sum_{i,j=1}^d \int_0^t \frac{\partial^2 f}{\partial x_i \partial x_j}(X_s) d[X^i, X^j]_s.$$

[see [12], Theorem 10.2]

**Theorem A.4 (Martingale Representation Theorem).** Let  $B = (B_t)_{t \geq 0}$  be a  $d$ -dimensional Brownian motion on the probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ , where  $\mathcal{F} = \sigma(\mathcal{F}_t^B \cup \mathcal{N})$  is the standard filtration. Then for each  $d$ -dimensional martingale  $X = (X_1, \dots, X_d)$  with respect to  $\mathcal{F}$  there exist a predictable processes  $Y^1, \dots, Y_d \in \mathcal{H}^2$  such that

$$X_t^i = X_0^i + \sum_{j=1}^d \int_0^t Y_s^j dB_s^j, \quad (\mathbb{P}\text{-a.s.}), \quad t \geq 0.$$

[see [12], Theorem 12.2]

*Remark:*  $\mathcal{H}^2 = \left\{ Y : Y \text{ is adapted, measurable and } \mathbb{E} \left( \int_0^\infty Y_s^2 ds \right)^{1/2} \right\} < \infty$

**Theorem A.5 (Girsanov's theorem).** Let  $B = (B_t)_{t \geq 0}$  be a Brownian motion on the probability space  $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$  and  $Y \in L^1(B)$ , and let

$$Z_s = \exp \left( \int_0^s Y_u dB_u - \frac{1}{2} \int_0^s Y_u^2 du \right), \quad \text{for } 0 \leq s \leq t,$$

be a  $\mathbb{P}$ -martingale. Then the measure  $\mathbb{Q}$  can be defined such that  $\mathbb{Q} \sim \mathbb{P}$  and  $\frac{d\mathbb{Q}_t}{d\mathbb{P}} = Z_t$ , and the process  $W = (W_t)_{s \leq t}$ , defined by

$$W_s = B_s - \int_0^s Y_u du,$$

is a Brownian motion with respect to  $\mathbb{Q}$ . [see [12], Theorem 11.8]

*Remark:* The process  $(Z_s)_{s \leq t}$  is a  $\mathbb{P}$ -martingale, if the Novikov condition,

$$\mathbb{E}^{\mathbb{P}} \left( \exp \left( \frac{1}{2} \int_0^t Y_u^2 du \right) \right) < \infty,$$

is satisfied.

# Bibliography

- [1] Bateman, H. and A. Erdélyi (1954) Tables of integral transform. New York: McGraw-Hill.
- [2] Brace, A., D. Gatarek and M. Musiela (1997): The Market Model of Interest Rate Dynamics. In: Mathematical Finance. Band 7, 1997, S. 127147
- [3] Brace, A and M. Musiela (1994). A Multifactor Gauss Markov Implementation of Heath, Jarrow, and Morton. *JMathematical Finance*, 4(3):259283, July 1994.
- [4] Dai. Q. and K. Singleton (2000). Specification analysis of affine term structure models. *Journal of finance* 55,1943-1978.
- [5] Duffie, D. (2001). *Dynamic Asset Pricing Theory* (Third ed.). Princeton University Press.
- [6] Duffie, D. and R. Kan (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance* 6 (4), 379406.
- [7] Filipović (2009). *Term Structure Models - A graduate Course*. Springer.
- [8] Filipović, D., Larson, M., Trolle, A. (2013). *Linear-Rational Term Structure Models*. Swiss Finance Institute @ EPFL
- [9] Filipović, D., Mayerhofer, E. (2009). *Affine Diffusion Processes: Theory and Applications*. ....
- [10] Julier, S.J. and Uhlmann, J.K. (1997). A New Extension of the Kalman Filter to Nonlinear Systems In *Proc. of AeroSense: The 11th Int. Symp. on Aerospace/Defense Sensing, Simulations and Controls*.
- [11] Müller, W. (2016) *Advanced Financial Mathematics, Lecture notes*
- [12] Müller, W. (2013) *Stochastic Analysis, Lecture Notes*

- [13] Munk, C. (2011). Fixed Income Modelling. Oxford University Press.
- [14] Munk, C. (2005). Fixed Income Analysis: Securities, Pricing, and Risk Management
- [15] Nelson, C., and Siegel, A. F., Parsimonious Modelling of Yield Curves, Journal of Business, Vol. 60 (October 1987), 473489.
- [16] Øksendal, B. (1998). Stochastic Differential Equations (Fifth ed.). Springer-Verlag.
- [17] Shao, A (2012) Fast and exact simulation for CIR process. A Dissertation presented to the gradual school of the University of Florida in partial Fulfilment of the requirements for the degree of Doctor of Philosophy
- [18] Wan, E.A. and van der Merwe, R. (2000) The Unscented Kalman Filter for Nonlinear Estimation. Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373).
- [19] <https://fam.tuwien.ac.at/public/posters/zinsposter.pdf>