TECHNISCHE
UNIVERSITÄT
WIEN

DIPLOMARBEIT

Financial market modeling based on the Kalman Filter in the setting of a Two-Factor Hull-White Model

ausgeführt am

Institut of Financial and Actuarial Mathematics
TU Wien

unter der Anleitung von

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

durch

Marlene Peck
Matrikelnummer: 01609981

Wien, am 06. März 2023

Kurzfassung

In der vorliegenden Diplomarbeit wird die Zinsstrukturkurve mittels eines Momentanzinsmodells modelliert. Als das zugrundeliegende Modell wurde das *Two-Factor Hull-White Model* gewählt, welches äquivalent zum *Gaussian Two-Factor Model* formuliert werden kann. Damit das Modell mit der initialen Zinsstrukturkurve übereinstimmt, wird das Nelson-Siegel-Svensson Verfahren angewendet. Aus täglich beobachteten Zinsdaten werden unter Anwendung des Kalman Filters die Modellparameter geschätzt. Hierfür muss zuerst das *state-space model*, welches aus der *measurement* und *transition* Gleichung besteht, für das zugrundeliegende Modell formuliert werden. Danach wird der iterative Prozess des Filters durchgeführt und die *Log-Likelihood Funktion* des Modellparametersets wird zurückgegeben. Dieser Prozess wird so oft wiederholt bis vorgegebene Grenzen des *Nelder-Mead* Optimierungsverfahren erreicht werden. Mit den geschätzten Parametern werden die zukünftigen Pfade der Nullkuponanleihe simuliert. Aus dem Durchschnitt der Pfade zu jedem Zeitpunkt werden dann mittels der Anleihenformel die Nullkuponanleihenpreise berechnet.
Um das Modell zu validieren wird zunächst überprüft, ob der Martingaltest erfüllt ist. Danach wird getestet ob das Modell konsistent ist, also ob die aus den simulierten Pfaden zukünftiger Momentanzinse erhaltene Modellparameter durch erneute Schätzung der Modellparameter, wieder gefunden werden. Schlussendlich werden die beobachteten Zinsdaten in zwei Datensätze aufgeteilt. Das erste besteht aus den Beobachtungen der ersten vier Jahre, das zweite aus der Beobachtung des letzten Jahres. Dann wird mit dem ersten Datensatz und der gewonnenen Modellparameter das nächste Jahr an Momentanzinsen geschätzt, und überprüft ob die simulierten mit den tatsächlich beobachteten übereinstimmen.

Abstract

This diploma thesis deals with modeling the interest rate term-structure using a short-rate model. The underlying model is the *Two-Factor Hull-White Model*, which is equivalent to the *Two-Factor Gaussian model*. In order to fit the model with the initial term-structure the *Nelson-Siegel-Svensson* method is applied. From daily observed short-rates we further estimate with the use of the Kalman Filter the set of model parameters. For this purpose the *state-space model*, which consists of the *measurement* and *transition* equation has to be formulated for the underlying model. Then the iterative process of the Filter can be undertaken and the return is the *log-likelihood function* of the model parameter set. This process reruns until certain thresholds of the *Nelder-Mead* optimizer are reached. With the estimated parameters the future paths of the zero-coupon bond are simulated. The mean of this paths of each time instant is used to calculate the future zero-coupon bond prices through the bond-price formula.
To validate the model it is checked whether the Martingale-test is fulfilled. Then it is tested on consistency, so if the estimated parameters of future short-rates can be re-estimated. Finally the observed short-rate is split into two sets of data. The first contains the observed short-rates of the first four years, the second the last year of observation. Then the first set is used to model the short-rates of the upcoming year with the obtained parameters and the simulated short-rates can be compared with the actual observed ones.

Danksagung

An dieser Stelle möchte ich meinen ausdrücklichen Dank an Herrn DI Michael Reichhard von der Security KAG aussprechen. Er hat mich auf dieses Thema gebracht und mich in allen Fragen und Diskussionen mit großer Geduld unterstützt.
Ebenfalls gilt mein Dank an Ao.Univ.Prof.Dipl.-Ing.Dr.techn. Peter Grandits von der TU Wien für seine Bereitschaft zur Betreuung dieser Arbeit und die ausgezeichnete Kommunikation.
Danke auch an die Kollegen DI Matthias Widman, DI Thomas Spanninger und Gregor Fleischmann, da sie sich bereit erklärt haben, diese Arbeit Korrektur zu lesen und mir dabei wertvolles Feedback gegeben haben und mir auch bei Unklarheiten weitergeholfen haben.
Last but not least vielen Dank an meine liebe Lerngruppe, die Boys (Matthias, Andi, Greg, Flex, Daniel, Knoll, Daniel und Julian) und Lavinia ihr habt großen Beitrag geleistet, dass ich meistens sehr gern zur Uni gegangen bin und mir die Studienzeit besonders schön in Erinnerung bleiben wird.

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Diplomarbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen und Hilfsmittel nicht benutzt bzw. die wörtlich oder sinngemäß entnommenen Stellen als solche kenntlich gemacht habe.

Wien, am Datum

Name des Autors

Contents

1 Introduction

Over the last decades modeling the term-structure of interest rates has been a widely researched topic. It is not only theory where this topic plays a crucial role, but also practice and especially the gap between theory and practise. Interest rates play an important role in the financial world for a wide range of matters such as: valuation, investment management, or solvency capital requirements. Especially companies with high shares of long-term liabilities such as insurances or pension-funds bear a high interest rate risk.
In the world of finance exists a variety of prices, which determine not only the current state of the economy, but are as well an indicator regarding the future development. The most important measure is the interest rate, which determines the cost of borrowing money, the return on savings and is a component of the total return on investments. Today we face the following situation: for the past few years borrowing money was cheap and people preferably took a variable loan, as the interest rate was very low. With raising interest rates many people face are major problem with paying alone the raised interest rate of the loan. So the interest rate not only determines an investors or a banks profit or loss but has direct influence on the people everyday lifes.
Even though everybody is somehow affected by the development of interest rates, only few know that a unique *interest rate* does not exist and is therefore not directly readable. There exists a variety of interest rates, marked by different maturities, liquidities, risk classes or other factors of influence. A meaningful comparison with so many different factors of influence is not possible. Therefore we try to model the term-structure, which represents interest rates of the same risk class depending on different maturities.
According to the *PRIIP-Verordnung* (PRIPP = Packaged Retail and Insurance-based Investment Products) for cash value life insurance, whose performance depends or partly not on the observed factors on the market, the insurances in Austria have to publish the so called *Basisblätter*. That is why the *AVÖ* (Aktuarvereinigung Österreich) published a guideline, on how to calculate the scenarios which are published in the *Basisblätter*. Until today Austrian insurances buy a tool, the *PIA tool*, which is provided by *IFA ULM* (Institut für Finanz und Aktuarwissenschaften). The goal of this thesis is to find a way to calibrate the model parameters for the underlying Two-Factor Gaussian model, and with them to simulate the future short-rate, and the corresponding zero-coupon bond prices.
In a first step we determine the parameter for the Gaussian Two-Factor model, which describes the dynamics of the short-rate. With this model we are able to simulate future short-rates, which can also be expressed as zero-coupon bond prices. To do so several models are applied. The basis is observed short-rates from the market, which are initially calculated into zero-coupon bond prices, and then also predicted for a longer time interval of maturities. Then the Nelson-Siegel-Svensson parameters are calculated with which the deterministic part of the Gaussian Two-Factor model can be defined. Afterwards we have to formulate the Kalman Filter equations for the underlying model. Then the procedure

starts and returns the maximum log-likelihood of the parameters. This output is then optimized through the Nelder-Mead algorithm and the Kalman Filter algorithm is repeated until a certain threshold is reached. With the estimated parameters the future paths of the zero-coupon bond are simulated. Using the mean of this paths at each time instant and applying the bond price formula, we calculate the desired future zero-coupon bond prices.
The thesis is split into eight chapters, which will be shortly introduced. The second chapter is devoted to basic definitions and concepts in the mathematical interest rate world. It starts with the Brownian motion, which is fundamental when it comes to stochastic modeling. Another important definition is the one of an Ornstein-Uhlenbeck process, as its solution will be used to formulate the equation for simulating the future paths of the zero-coupon bond. We have a closer look at term-structures and their different shapes. Furthermore, we have a look at the different kind of models, which can be classified into stochastic and statistical models. We focus on the fundamental definition of a zero-coupon bond and a bank account, which are expressed through the short-rate. With this knowledge we continue to define the different kinds of interest rates, as well as their interaction.
Chapter 3 focusses on the formulation of the Two-Factor Hull-White model and shows the equivalence with the Two-Factor Additive Gaussian model, which formulation will be used throughout this thesis. It is a no-arbitrage model and can be specified with the short rate being driven by two correlated factors and a deterministic function. For many years the fact that this model also produces negative rates has been mentioned as a major drawback. As there have actually been negative rates on the market in the past years, this property can now be mentioned as an advantage of the model. As negative rates are admissible in this model, the following section identifies the probability with which they occur. The last section of this chapter focusses on the shift of the model formulation from risk-neutral to the real-world measure. This is an extremely important step, because the real-world model formulation is used for the Kalman Filter.
Chapter 4 describes the Nelson-Siegel-Svensson technique. In order to fit the underlying model with the initial term-structure, we need the deterministic part of the short-rate equation. This again is determined through the modeled Nelson-Siegel-Svensson zero-coupon bond prices, which are gained through estimating the Nelson-Siegel-Svensson parameters from the observed short-rates. With these parameters a modeled spot-rate can be defined. Through the inverse modeled spot-rate plus one with the exponent being the respective time point, the corresponding zero-coupon prices can be calculated.
Chapter 5 introduces the formulation of the *state space* model for the Kalman Filter. The state space model again consists of the measurement and transition equation, which we formulate for the underlying model. We describe the iterative procedure of the Kalman Filter. The output of the Filter is the log-likelihood function of the set of parameters, on which a numerical approach, the *Nelder-Mead* algorithm is applied until certain thresholds are met.
Chapter 6 describes the procedure undertaken to obtain the future zero-coupon bond prices and the future short-rates. At first the observed short-rate is interpolated to enlarger the time of maturity. With the model parameters estimated through the Kalman Filter, we can simulate the future paths of zero-coupon bond prices. Applying the bond-price formula to the mean of each path at each time point, we obtain the future zero-coupon prices, which can be transformed into the short-rate.

To conclude we validate the output of the estimation with different approaches. One standard method is the well-known Martingale-test. We also compare the estimated parameters obtained through the observed market prices, with those simulated with the obtained parameters which are then reintroduced into the simulation. We confirm this data by using a different approach by splitting the data into two sets. The first contains the observed short-rates of the first four years. This is then used as the basis to simulate the short-rates for the next year using the obtained parameters as input. The results can then be compared with the second set of data, the observed short-rates from the last year of the observation period.

2 Mathematical Background

This chapter outlines the most important definitions, formulas and theories, which are fundamental for the subsequent chapters. We start this chapter by defining the properties of the Brownian motion [Definition 2.1], which is fundamental when it comes to modeling stochastic processes in the world of finance.
Besides the recapitulation of the essentials in terms of stochastic analysis, this chapter also states the notation. For the about to be specified definitions we keep in mind, that we regard a complete probability space $(\Omega, \mathcal{F}, \mathbb{F}, \mathbb{P})$. Where Ω denotes the sample space, $\mathbb{F} = (\mathcal{F}_t)_{t\in T}$ denotes the filtration, $\mathbb{P}$ denotes the probability measure and let $T \in \mathbb{R}^+$. The filtration fulfils the common properties: $\mathcal{F}_0$ contains all null-sets and $\mathbb{F}$ is right-continuous.

Definition 2.1. *(Brownian Motion)*[1] *[cf.[3]]. A $\mathbb{R}^d$-valued stochastic process $(W_t)_{t\in T}$ is called a d-dimensional Brownian motion if the following properties hold:*

1. *W_0 is $\mathcal{F}_0$-measurable*
2. *measurability of past increments, i.e., $W_s - W_r$ is $\mathcal{F}_s$-measurable $\forall\ r < s \in T$*
3. *independence of future increments, i.e., $W_t - W_s$ is independent of $\mathcal{F}_s$ $\forall\ s < t \in T$*
4. *stationarity of the distribution of the increments:*
$$\mathcal{L}(W_t - W_s) = \mathcal{L}(W_{t-s} - W_0)\ \forall\ t \in T$$
5. *normality of increments: $\mathcal{L}(W_t - W_0) = \mathcal{N}(0, tI_d)\ \forall\ t \in T$ with $I_d \in \mathbb{R}^{d\times d}$ denoting the identity matrix*
6. *W has continuous paths, i.e. for $t \in T$: $t \longmapsto W_t(\omega)$ is continuous $\forall\omega \in \Omega$*

   if in addition:

7. *starting at the origin: $\mathbb{P}[W_0 = 0] = 1$ the Brownian motion is called standard*

[1] The Brownian motion is named after the botanist Robert Brown, who studied the irregular motion of pollen using a microscope in 1827. He was the first who tried to find an explanation for this erratic behaviour. In 1900 the French mathematician Louis Bachelier used the Brownian motion in a mathematical model in order to describe the price development at the stock market. In 1905 Albert Einstein published a paper about modeling the motion of the pollen particles which are moved by individual water molecules. The direction of the force of atomic bombardment is constantly changing and at different collisions the particles are hit more on one side than the other, which leads to the seemingly random nature of the motion. The existence of Brownian motion served as convincing evidence that atoms and molecules exist. In the 1920s Norbert Wiener found proof for the existence of the Brownian motion and with this knowledge stochastic processes gained more and more in importance.[cf.[1, 2]]

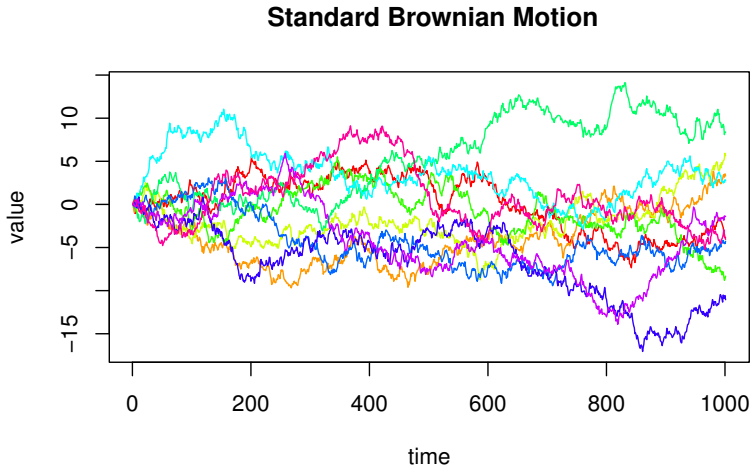


Figure 2.1: Realisation of a Standard Brownian Motion

The above given [Figure 2.1] represents a simulation of the *Standard Brownian Motion* with ten paths in the time interval $[0, 1000]$. The x-axis displays the time interval and the y-axis the corresponding values at each time point of the Brownian motion. The ten paths were drawn by simulating 1000 points for each path.

Definition 2.2. *(Brownian Motion with drift) [cf.[4]]. A process X_t is called Brownian motion with drift, if it satisfies:*

$$X_t = \mu t + \sigma W_t, \text{ with } t \geq 0, \mu \in \mathbb{R}, \sigma \in \mathbb{R}^+$$

where W_t represents the standard Brownian motion.

Definition 2.3. *(Martingale) [cf.[3, 4]]. A Martingale is an integrable, $(\mathcal{F}_t)_{t\in T}$-adapted stochastic process $(M_t)_{t\in T}$ for that holds:*

$$M_s = \mathbb{E}[M_t|\mathcal{F}_s] \text{ a.s. } \forall s \leq t \in T$$

The first condition for the process is just a technical one and the second condition says that we can actually observe the value M_t at time t. The martingale property describes that the underlying process is a fair process. To be more precise, the expectation of a future value of M, given the information available today's, equals today's observed value of M. In fact it can also be stated that a martingale has no systematic drift.

Definition 2.4. *(Semi-Martingale) [cf.[3]]. A stochastic process* $(X_t)_{t\in\mathbb{R}_0^+}$ *is called a semi-martingale, if it can be written in the form of:*

$$X_t = M_t + F_t,\ t \geq 0$$

where (M_t) *denotes a local martingale*[2] *and* F_t *an adapted process of finite variation*[3] *with* $F_0 = 0$.

Definition 2.5. *(Stochastic Differential Equation) [cf.[4]]. For the following let* $M(n,d)$ *denote the class of* $n \times d$ *matrices and we consider:*

- *a d-dimensional Brownian motion* W_t *(see [Definition 2.1])*
- *a function* μ*:* $\mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}^n$
- *a function* $\sigma : \mathbb{R}^+ \times \mathbb{R}^n \to M(n,d)$
- *a real vector* $x_0 \in \mathbb{R}^n$

The goal is to find a stochastic process X_t *which satisfies the* ***stochastic differential equation*** *(SDE):*

$$dX_t = \mu(t, X_t)dt + \sigma(t, X_t)dW_t$$
$$X_0 = x_0$$

Put in other words we want to find a process X_t *to satisfy following integral equation:*

$$X_t = x_0 + \int_0^t \mu(s, X_s)ds + \int_0^t \sigma(s, X_s)dW_s,\ \forall t \geq 0$$

Definition 2.6. *(Quadratic Variation) [cf.[6, 7]]. The quadratic variation of a stochastic process* X_t *with continuous paths* $t \to X_t(\omega)$ *is defined as:*

$$[X]_t = \lim_{n\to\infty} \sum_{i=1}^{\infty} (X_{t_i^n}(\omega) - X_{t_{i-1}^n}(\omega))^2$$

Where t_i^n represents the partition of the time index, therefore we have: $t_i^n := \frac{it}{n}$, for $i = 0, \cdots, n-1$ such that: $0 = t_0 < t_1 < \cdots < t_{n-1} < t_n = t$

[2] M is called a local martingale, if there exists a sequence $(\tau_n)_{n\in\mathbb{N}}$ of $\mathbb{R}^+$-valued $\mathcal{F}^+$-stopping times with $\tau_n \leq \tau_{n+1} \forall n \in \mathbb{N}$ and $\lim_{n\longmapsto\infty} \tau_n = \infty$ such that each τ_n reduces M. [cf.[3]]

[3] A process is called of finite variation if and only if it can be represented as the difference of two increasing processes: $\int \mu dt = \int \mu^+ dt - \int \mu^- dt$. With μ being any process such that the integral exists. [cf.[5]]

Definition 2.7. *(Stochastic Exponential)[cf.[3]]. The stochastic exponential $\mathcal{E}(X)_t$ for a continuous semi-martingale X_t is given as:*

$$\mathcal{E}(X)_t = \exp(X_t - X_0 - \frac{1}{2}[X]_t),\ t \geq 0$$

Where [.] denotes the quadratic variation of the process. The stochastic exponential with respect to the Brownian motion can be written as:

$$\mathcal{E}(X)_t = \exp(X_t - \frac{1}{2}t)$$

Definition 2.8. *(geometric Brownian Motion) [cf.[4, 8]. A geometric Brownian motion is a continuous-time stochastic process in which the logarithm of the randomly varying quantity follows a Brownian motion. In option pricing theory geometric Brownian motions are frequently used to model general asset price dynamics. Its evolution can be expressed through:*

$$dX_t = \mu X_t dt + \sigma X_t dW_t \tag{2.1}$$

with $X_0 = x_0$, μ ("the percentage drift") and σ ("the percentage volatility") are positive constants, and W_t denotes the time derivative of the Brownian motion. Applying Ito's Formula on equation (2.1) leads to following solution:

$$X_t = X_0 \cdot e^{(\mu - \frac{1}{2}\sigma^2)t + \sigma W_t}$$

Definition 2.9. *(Equivalent martingale measure) [cf.[4]]. For the following we regard the probability measure $\mathbb{Q}$ on the filtration $\mathcal{F}_T$, which is called equivalent martingale measure for the asset price vector $S_1, \cdots, S_N$ of N risky traded assets over the time interval $[0, T]$, with $S_0 = 0$, if following properties hold:*

- *$\mathbb{Q}$ is equivalent to $\mathbb{P}$ on $\mathcal{F}_T$, i.e. they both have the same null-sets*
- *all price processes $S_0, \cdots, S_N$ are martingales under $\mathbb{Q}$ on the time interval $[0, T]$*

Theorem 2.1. *(Numéraire Dependent Pricing Formula) [cf.[9]]. Before having a look at the Numéraire Dependent Pricing Formula (NDPF), we recall the most important theorem in asset pricing theory: the First Fundamental Theorem of Asset Pricing (FTAP). The theorem states: The market specified by some real-world probability measure $\mathbb{P}$ is free of arbitrage, if and only if, given any numéraire N, there exists a measure $\mathbb{Q}_N$ which is equivalent to $\mathbb{P}$, and which is such that all relative price processes are $\mathbb{Q}_N$ martingales.*
The important part of the FTAP is, that it guarantees absence of arbitrage. This is relevant when it comes to pricing in financial markets, where we wish to price market-consistent if the market is arbitrage-free. To be more specific when it comes to introducing a new asset to the market, it is said to be priced market-consistent, if the market with this new asset included, still does not allow arbitrage. Let $C(t)$ denote the new asset introduced at time t, then along with the FTAP, the price is market consistent, if and only if, for at least one

equivalent martingale measure, the model allows for:

$$\frac{C(t)}{N(t)} = \mathbb{E}^{\mathbb{Q}_N}\left[\frac{C(T)}{N(T)}|\mathcal{F}_t\right]$$

Therefore we can express the price of an asset as:

$$C(t) = N(t)\mathbb{E}^{\mathbb{Q}_N}\left[\frac{C(T)}{N(T)}|\mathcal{F}_t\right]$$

Definition 2.10. *(Itô-Process) [cf.[10]]. An Itô-Process is a stochastic process $X(t)$ written in the form:*

$$\begin{aligned} X(t) &= X(0) + \int_0^t K(s)ds + \int_0^t H(s)dW(s) \qquad (2.2)\\ &= X(0) + \int_0^t K(s)ds + \sum_{j=1}^{m}\int_0^t H_j(s)dW_j(s),\ t \geq 0 \end{aligned}$$

where W_t denotes the Brownian motion and both $\{K(t)\}_{t\in T}$ and $\{H(t)\}_{t\in T}$ are progressive measurable and satisfy:

$$\int_0^t |K(s)|ds < \infty,\ \int_0^t H_i^2(s)ds < \infty,\ \forall t \geq 0,\ \ i = 1, \cdots, m$$

By writing formula (2.2) in differential notation we get following equation:

$$dX(t) = K(t)dt + H(t)dW(t),\ t \geq 0$$

Definition 2.11. *(Itô's formula for Semi-Martingales) [cf.[3]]. Let X_t be a Semi-Martingale and $f : \mathbb{R}^+ \times \mathbb{R} \to \mathbb{R}$ one time in the first, and two times continuous differentiable in the second argument. Then $f(t, X_t)$ is again a Semi-Martingale and can be written as:*

$$\begin{aligned} f(t, X_t) = f(0, X_0) + \int_0^t f_t(s, X_s)ds + \int_0^t f_x(s, X_{s-})dX_s + \frac{1}{2}\int_0^t f_{xx}(s, X_s)d[X]_s^c \\ + \sum_{0\leq s\leq t}(f(s, X_s) - f(s, X_{s-}) - f_x(s, X_{s-})\Delta X_s),\ t \geq 0 \end{aligned}$$

with $[.]_t^c = [X]_t - \sum_{0\leq s\leq t}(\Delta X_s)^2$

Written in differential notation the continuous Semi-Martingale X_t is given as:

$$df(t, X_t) = f_t(t, X_t)dt + f_x(t, X_t)dX_t + \frac{1}{2}f_{xx}(t, X_t)d[X]_t,\ t \geq 0$$

Definition 2.12. *(Ornstein-Uhlenbeck process)*[4] *[cf.[11]]. An Ornstein-Uhlenbeck process is a temporally homogeneous stationary Gauss-Markov process and is used in interest-rate modeling. Over the time it tends to drift towards its mean function, therefore it is also known as a mean-reverting process. The Ornstein-Uhlenbeck process X_t is defined through following SDE:*

$$dX_t = \theta(\mu - X_t)dt + \sigma dW_t \tag{2.3}$$
$$x_0 = a$$

In formula (2.3) we have: a, $\mu \in \mathbb{R}$, $\theta, \sigma > 0$, μ is a constant and W_t denoting the Brownian motion. Furthermore μ describes the equilibrium level, θ the stiffness and σ the diffusion level. The expected value of this process can be written as:

$$\mathbb{E}(X_t) = ae^{-\theta t} + \mu(1 - e^{-\theta t})$$

With corresponding covariance given as:

$$Cov(X_s, X_t) = \frac{\sigma^2}{2\theta}\left(e^{-\theta|t-s|} - e^{-\theta(t+s)}\right)$$

The solution of the equation (2.3), will be used in [Section 6.3] for the simulation of the future paths of zero-coupon bond prices, respectively the short-rate and is given as:

$$X_t = X_0 e^{-\theta t} + \mu(1 - e^{-\theta t}) + \sigma \int_0^t e^{-\theta(t-s)} dW_s \tag{2.4}$$

The *change of measure* is a commonly used technique when it comes to pricing theory. With this tool one can change one probability space to another probability space. Its main purpose is to change the risk-neutral model formulation to the real-world measure, so that the theory can actually be applied. The relationship between the real-world measure $\mathbb{P}$ and the risk-neutral measure $\mathbb{Q}$ can be described through the Radon-Nikodym derivative.

Definition 2.13. *(Radon-Nikodym) [cf.[4]]. Let $\mathbb{P}$ and $\mathbb{Q}$ be probability spaces or measures and θ be a random variable such that*

$$\mathbb{E}^{\mathbb{Q}} X = \mathbb{E}^{\mathbb{P}} \theta X$$

for all random variables X, then θ is called the Radon-Nikodym derivative of $\mathbb{Q}$ with respect to $\mathbb{P}$. The Radon-Nikodym process for every $0 \leq s \leq t$ is written as:

$$\mathbb{E}_s^{\mathbb{Q}} X(t) = \mathbb{E}_s^{\mathbb{P}} \frac{\theta_t}{\theta_s} X(t)$$

[4] The Ornstein-Uhlenbeck process is named after the physicians George Uhlenbeck (1900-1988) and Leonard Ornstein (1880-1941).

Definition 2.14. *(Girsanov) [cf.[4]]. Let $W(t)$ denote a Brownian motion under $\mathbb{Q}$ and $\lambda(t)$ be a k-vector process adapted to $W(t)$. If $\lambda(t)$ fulfils some mild boundedness conditions, then we can define:*

$$d\lambda(t) = -\theta(t)\lambda'(t)dW^{\mathbb{P}}(t) \ , \ \lambda(0) = 1$$

That way $\lambda(t)$ can be regarded as a Radon-Nikodym process that defines a change of measure from the real world to the risk neutral measure. The process $W^{\mathbb{Q}}(t)$ can now be written as:

$$dW^{\mathbb{Q}}(t) = \lambda(t)dt + dW^{\mathbb{P}}(t), \ W^{\mathbb{Q}}(0) = 0$$

and is a Brownian motion under $\mathbb{Q}$.

Definition 2.15. *(Log-Likelihood Function) [cf.[12]]. Let $(X_1, \cdots X_n)$ be independent and identically distributed (iid) random variables with a common probability density function (pdf) $f(x;\theta)$. For a given outcome $x = (x_1, \cdots, x_n)'$ the function*

$$l(\theta; x) = \ln L(\theta; x) = \sum_{i=1}^{n} \ln f(x_i; \theta)$$

is called the log-likelihood function.

Definition 2.16. *(Maximum Likelihood Estimation) [cf.[12]]. Given a random sample $\mathsf{x} = (x_1, \cdots, x_n)$ corresponding to the model $f(x;\theta)$ and one wants to consider the likelihood $L(\theta; x)$. Although the entire shape of the function yields valuable information about the to be estimated parameter θ, the goal is to find the parameter value $\hat{\theta} = \hat{\theta}(\mathsf{x})$ for which the likelihood is maximal, which is called its mode. Therefore this value can be understood as the best estimate of θ and is called the maximum likelihood estimate (MLE). However there is no guarantee that the MLE always exists. As the mode may not be an interior point of the parameter space Θ or the likelihood may attain its maximum value at more than one point. Furthermore it could happen that the global maximum exists but is no sensible choice, so that a local maximum is chosen (of which however there may be more than one).*

Definition 2.17. *(Splines) [cf.[13, 14]]. A Spline of degree n is an of piece-wise polynomial composed function with maximum degree of n. The points where two polynomials are connected are called knots. At these knots the Spline has to be $(n-1)$-times continuous differentiable.*
The pioneers of the spline-method are Isaac Jacob Schoenberg, Paul de Faget de Casteljau, Pierre Bézier and Carl de Boor. Jacob Schoenberg initially described with the term spline a smooth, harmonic composed curve of degree three.
Splines are mainly used for approximation and interpolation. As Splines are defined piece-wise, they are more flexible than polynomials, however easier and smoother. Let us first have a look at the basic element of Splines, the piece-wise polynomials and then define Splines:

- *For the following x is assumed to be univariate and we regard $k+1$ disjoint intervals, where each interval $(-\infty, \xi_1), [\xi_1, \xi_2), \cdots, [\xi_{k-1}, \xi_k), [\xi_k, \infty)$ defines an own polynomial function of order $\leq M$. Where the order of the polynomial describes the degree $+1$.*

By dividing the domain of x into those $k+1$ disjoint intervals a piece-wise polynomial function $f(x)$ is obtained. As already mentioned the boundaries of the interval are called knots. A piece-wise polynomial function of order $M = 2$ is called piece-wise linear, $M = 3$ quadratic, $M = 4$ cubic etc. In order to determine a piece-wise polynomial function of order M with k knots, $M(k+1)$ parameters are needed, since each of these $k+1$ polynomial consists of M coefficients.
The basic functions and constraints can be defined through:

$$h_1(x) = 1, \quad h_2(x) = x, \quad h_3(x) = (x - \xi_1)^+, \quad h_4(x) = (x - \xi_2)^+$$

Where $(.)^+$ denotes the positive part.

- *A piece-wise polynomial function of order M, which has continuous derivatives up to order $M-2$ is called Spline of order M with knots $\xi_i, i = 1, \cdots, k$. The general form of the basis function is written as:*

$$h_j(x) = x^{j-1} \ , \ j = 1, \cdots, M$$
$$h_{M+l}(x) = (x - \xi_l)_+^{M-1} \ , \ l = 1, \cdots, k$$

- *A Cubic Spline of order $M = 4$ with two knots has following basis functions:*

$$h_1(x) = 1 \ , \ h_3(x) = x^2 \ , \ h_5(x) = (x - \xi_1)_+^3$$
$$h_2(x) = x \ , \ h_4(x) = x^3 \ , \ h_6(x) = (x - \xi_2)_+^3$$

- *When it comes to Natural Splines, the basic idea is to model a non-linear relationship with piece-wise Cubic Splines. Natural Cubic Splines are chosen as they can avoid poor results from cubic splines, which occur as polynomial tend to be erratic near the lower and upper data range. That is due to having an additional constraint, which is that the function has to be linear beyond the boundary knots. Therefore the Spline function f has so satisfy: $f'' = f''' = 0$.*

Definition 2.18. *(B-Spline) [cf.[15]]. A B-Spline which is the abbreviation of Basis-Spline is referred to the basis function of Spline functions of the same order and defined over the same knots. The choice of the basis determines over rounding errors in the spline-space and therefore over the practical usage. The B-Spline is numerical stable and permits for the value calculation of the Spline-function a three-term-recursion. The B-Spline of order n is a piecewise polynomial function with degree $n-1$. It is defined using $n+1$ knots, which are in a non-descending order: $t_j \leq t_{j+1}$. However the B-Spline only contributes in the range between the first and the last knot, elsewhere it equalizes zero. Any Spline of order n on a given set of knots can be expressed through a linear combination of B-Splines:*

$$S_{n,t}(x) = \sum_i \alpha_i B_{i,n}(x)$$

The role of the basis function come from the fact the B-Splines have the same continuity properties at the knots and can be expressed as:

$$B_{i,1}(x) := \begin{cases} 1 & \text{if } t_i \leq x < t_{i+1} \\ 0 & \text{else} \end{cases}$$

$$B_{i,k}(x) := \frac{x - t_i}{t_{i+k} - t_i} B_{i,k-1}(x) + \frac{t_{i+k+1} - x}{t_{i+k+1} - t_{i+1}} B_{i+1,k-1}(x)$$

By adding the additional constraint: $\sum_i B_{i,1}(x) = 1 \ \forall x$ *between the first and last knot, the scaling factor of* $B_{i,n}(x)$ *becomes fix. The B-Splines are then the resulting in* $B_{i,n}(x)$ *Spline functions.*
The B-Spline function has a compact carrier, in other words they do not equal zero just in a small interval. A B-Spline function is a combination of flexible bands to create smooth curves and is controlled by control points. The goal is to create and manage complex shapes and surfaces using a number of points. B-Splines are used in curve fitting when no theoretical basis for choosing a fitting is given. In this case the curve is fitted through a Spline function composed of a sum of B-Splines using the least-square method. The objective function of a Spline function of degree k is given as:

$$U = \sum_x \left\{ W(x) \left[y(x) - \sum_i \alpha_i B_{i,k,t}(x) \right] \right\}^2$$

$W(x)$ *denotes the weight and* $y(x)$ *the value at* x*. The coefficients* α_i *are the to be determined parameters and the knot values are fixed.*

2.1 Term-Structure

The results for the underlying section are taken from [cf.[6, 16, 17, 18, 19]]. Generally speaking a term-structure is a function which refers a certain financial variable or parameter to its maturity. The most common examples are the term-structure of interest rates or zero-coupon bond prices. Nevertheless there also exist term-structures of option implied volatilities, credit spreads or variance swaps. The overall rate of interest that an issuer of a bond pays to the holder at maturity is called the bond's yield. The graphed relationship between yield and maturity of interest rates is called the *yield curve*. The yield curve represents a measure of the market's expectations of future interest rates, given the current market conditions.

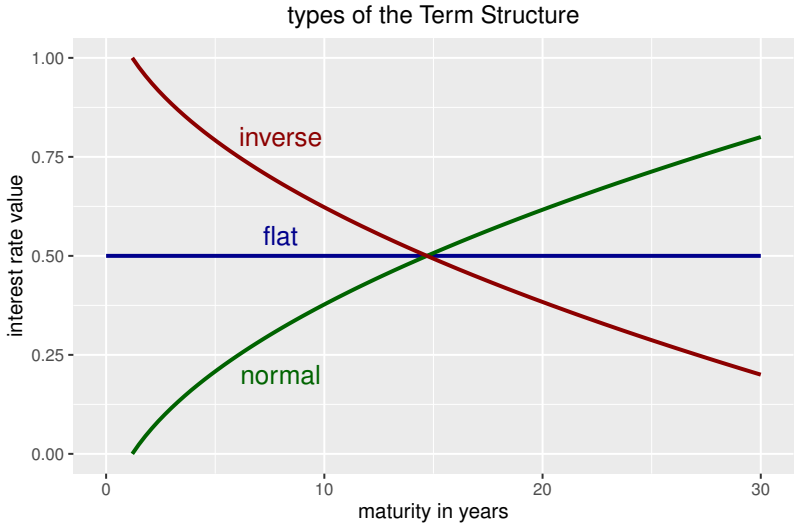


Figure 2.2: different shapes of the term-structure

As depicted in [Figure 2.2] the term-structure can distinguish between three main shapes, or even a mix of them may occur. A *normal* shape is referred to when the long-term yields are expected to be higher than the short-term ones. This is depicted in [Figure 2.2] as the green line. It is the most common shape and typically suggests a growing economy. In the case of the *flat* shape there is hardly no variation between the different maturities expected. This shape usually occurs in times of transition from the normal to the inverted shape. The economic interpretation of this transition is, that the market for long-term bonds is undesirable. In [Figure 2.2] the flat shape is represented as the blue line. The *inverse* shape is referred to when short-term yields are higher than the long-term ones. This occurs when short-term yields increase fast and investors take this as a sign of a temporary raise, but long-term yields still stay at the same level. In [Figure 2.2] we find this type shown in the red curve.

In general term-structure modeling can be split into two frameworks. The first approach is the *equilibrium framework*, which initial position is the description of the underlying economy. It works in a way that agents optimize their life-time utility function and starting from the aggregation of theses individual preference functions an equilibrium can be derived. The model is fully specified by these individual preferences, that is why interest rates, prices for derivatives and market prices of risk are endogenously determined. Moreover as these utility functions are well-specified e.g., strictly increasing and strictly concave, the model prohibits arbitrage opportunities. The pioneer of this equilibrium approach in bond- and derivative pricing was Merton in 1973. Nevertheless there are many other contributors, such as Vasicek (1977), Dothan (1978) and Cox, Ingersoll& Ross (1985) to name a few.
The second framework is defined through a *no-arbitrage approach*, which specifies that interest rates and the market price of risk do not result from individual preferences. That is why, when applying this approach one cannot determine how the model will behave due to structural changes. A major advantage of these models is that they fully represent the initial term structure, as the initial term structure is used as an input. This approach is frequently used for pricing of financial derivatives, for example: (bond)options, caps, floors and swaptions.

2.2 Interest Rates

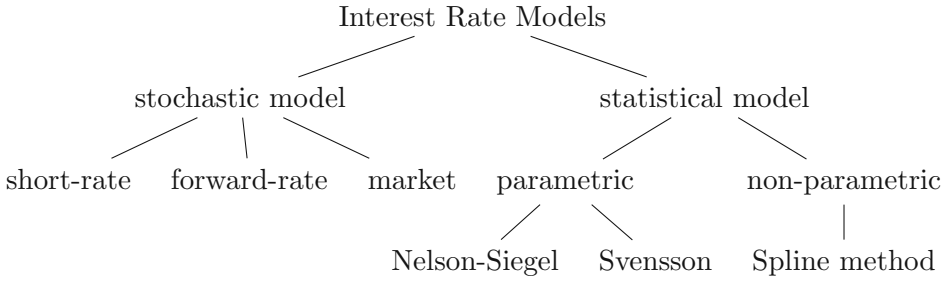
In a financial setting the term *interest rate* usually describes the rate by which money is borrowed over a certain time period. In other words it can be understood as the price one is paying for the use of a unit of another's money. That is why it has major influence on the current situation of economies and and then of course on decisions with an impact on future developments. Besides the macroeconomic importance, the interest rate is the main price in the world of finance. It is the underlying item when it comes to decisions concerning investing, portfolio composition and even for the valuation of hedging- and speculative instruments. That is why financial institutions face a major in risk in changing interest rates.
A unique *interest rate* does not exist at the market, and is therefore not directly readable. Even though there exist a variety of interest rates, which are classified by different maturities, credit ratings, liquidities and other factors of influence. However it is possible to depict interest rates of the same credit ranking dependent on the maturity through the *term-structure*. Therefore modeling the term-structure of interest rates plays an important role. This section introduces different kinds of interest rates and their link to zero-coupon bonds. The concept of interest rates belongs to our every-day life, and the basic ideas are the following. Firstly everybody expects that lending money should somehow be rewarded. Secondly receiving a given amount of money today is not equivalent to receiving the exactly same amount at any future date.
Before defining the different types of interest rates we want to distinguish between interbank rates and government rates. Interbank rates refer to rates at which deposits are exchanged between banks and at which swap transactions between banks occur. Government rates, as the name might suggest, are rates deduced by bonds and issued by governments.

Interest rate models are also known as equilibrium term-structure models, or as affine term-structure models, which are used to estimate the correct theoretical term-structure. These models are used to describe the stochastic process of the dynamics of the yield curve. Also the models identify the miss-pricing in the bond market as the estimated term-structure does not match the actual one. In this thesis we will work with a two factor short-rate model, assuming that there exist two macroeconomic variables that affect the term-structure of interest rates.

2.2.1 Types of Models

There exists a variety of interest rate model types. Following criteria can be summed up as basic properties for all of these types:

- interest rates should follow a mean reversion process
- long-term interest rates should exhibit a smaller volatility than short-term ones
- the correlation between different maturities should always be positive, however the smaller the difference in maturities, the higher the correlation
- proportionality should exist between value and volatility of interest rates



Two types of interest rate models can in general be distinguished. On the one hand there exist stochastic models, which is the focus in this thesis as the *Two-Factor Hull-White model*, or equivalently the *Two-Additive-Factor Gaussian model*, belongs to the class of the short-rate models. However on the other hand there exist static models, which again can be subdivided into parametric and non parametric models. In order to determine the parameter for the underlying short-rate model, we will use the Kalman Filter technique, which requires the Nelson-Siegel-Svensson Parameter to determine the initial term-structure with the model parameter and the model bond price. That is why we also describe the parametric model in detail in [Chapter 4].

2.2.2 Zero-Coupon Bond

The following subsection focusses on the most important tool of many financial derivatives, which are based on interest rates, the (default-free) *zero-coupon bond*. A *zero-coupon bond* is a special type of a bond without any intermediate payments. Therefore until the end of maturity no payments take place and that is why no reinvestment of the yield is possible. That again explains the exposure to a high volatility as the interest rate fluctuations have a strong impact. Zero-coupon bonds serve as a basis for the valuation of other fixed income instruments. Even though zero-coupon prices are used to directly construct the term-structure, due to the lack of market liquidity and the unavailability of all desired maturities, it is necessary to estimate the prices based on observed coupon bond prices. The foundation of the estimation is distinguished between on the one hand Interest-Rate-Swaps, which is for maturities shorter than one year. And on the other hand, Government-Bonds, which is used for maturities larger than one year. Government-Bonds generate every year a fixed coupon, or fixed interest rate in respect to the notional N. Let $K_p(t,T)$ denote a coupon-bond with fixed payments c at times $t < t_1 < \cdots < t_n < T$ and remaining maturity of the time $T - t$. Then we can write the the fixed price, which is paid at end of maturity with the notional, as

$$K_p(t,T) = c\sum_{k=1}^{n} P(t,t_k) + P(t,T)N$$

The term $P(t,t_k) = e^{-z(t,t_k)(t_k-t)}$ denotes the price of the zero-coupon bond with maturity $t_k - t$. That is why the price of a government bond is determined through the price of a zero-coupon bond. As the price of Government-bonds is observed more frequently than the one of zero-coupon bonds, the underlying combination has to be estimated. A commonly used technique is the Nelson-Siegel-Svensson method, which will be discussed in [Chapter 4]. Based on the concept *one unit currency today is worth more than one unit of the same currency tomorrow*, the time value of this unit currency is expressed through a zero-coupon bond. To be more precise it is a secured form of a loan, that guarantees its holder the payment of one unit at maturity T, without any intermediate payments. In other words one may describe a bond as a financial instrument in the market where the time value of money is traded. The time value t of a zero coupon bond with maturity T is denoted by $P(t,T)$ and the following assumptions hold:

- there exists a market for T-Bonds $\forall\ T > 0$
- $P(t,t) = 1\ \forall\ t$
- for fixed t the price $P(t,T)$ is differentiable with respect to the maturity T

It has to be mentioned that in reality not all of these assumptions fully hold. There exits a default risk, that is why $P(t,t)$ might be less than one unit of the underlying currency. Furthermore zero-coupon bonds are not in general traded for all maturities. The third assumption is a technical one and implies that the term-structure of their prices $T \to P(t,T)$ is a smooth curve, whereas $t \to P(t,T)$ is a stochastic process.

Definition 2.19. *(Zero-Coupon Yield) [cf.[4]]. Considering a zero-coupon bond with market price denoted by $P(t,T)$. If we want to determine the bond's internal rate of interest, denoted by $z(t,T)$, which can be understood as the constant short-rate of interest giving the same value to the bond as the value given by the market. We want to solve the equation:*

$$P(t,T) = e^{-z(t,T)(T-t)}$$

Which brings us to the continuously compounded zero coupon yield given by:

$$z(t,T) = -\frac{\ln P(t,T)}{T-t} \tag{2.5}$$

Therefore the yield corresponds to the spot-rate (which will be introduced in section 2.2.4) for the interval [t,T].

It has to be mentioned that throughout this thesis, we will work with observed bond-prices, which will be from now on denoted by: $P(0,t)$, as well as with modeled bond prices, which are added a superscript M, that is why these kind prices are denoted by: $P^M(0,t)$.

2.2.3 Bank Account

The bank account is specified as a risk-less investment, with profit evolving continuously at the risk-free rate. Let $B(t)$ denote the value of the bank account at time $t \geq 0$. Let us now assume $B(0) = 1$, then the evolution of the bank account can be determined through the following SDE:

$$dB(t) = r(t)B(t)d(t) \tag{2.6}$$

In equation (2.6) $r(t)$ denotes a positive function of time, and is known as the instantaneous rate, or more common *short-rate*. Investing a unit at time 0 yields to the value at time t defined in the following way:

$$B(t) = e^{\int_0^t r(s)ds} \tag{2.7}$$

In arbitrarily small time intervals $[t, t+\Delta t)$, the growth of the bank-account at each time instant t can be written as:

$$\frac{B(t+\Delta t) - B(t)}{B(t)} = r(t)\Delta t$$

When it comes to relating amount of currencies available at different times, it is recommended to look at the bank-account in terms of numeraire. The question that now arises is formulated as: What is the value at time t of one unit available at time T? Depositing A units of a currency in the bank-account at time 0, leads at time t to receiving $A \times B(t)$ units of the currency. Wanting exactly one unit at time T, e.g., $AB(T) = 1$, leads to initially investing the amount: $A = \frac{1}{B(T)}$. For simplicity it is assumed, that the interest rate process r is deterministic, hence B too and therefore the initial investment amount is known. The value at time t of the amount A invested at the initial time can be written as:

$$AB(t) = \frac{B(t)}{B(T)}$$

This brings up the definition of the stochastic discount factor $D(t,T)$ between two time instants (t,T) is the amount at time t that is *equivalent* to one unit payable at time T, e.g.:

$$D(t,T) = \frac{B(t)}{B(T)} = e^{-\int_t^T r(s)ds}$$

As the probabilistic nature of $r(t)$ affects the nature of the bank account numeraire, it is the key aspect. In various pricing application, i.e., the Black-Scholes Formula in equity markets, it is assumed that r evolves in a deterministic setting. Then again the bank-account and the stochastic discount factor at any future date are as well deterministic. Nevertheless when dealing with interest rate products, the most important parameter is the interest rate itself. That is why we drop the deterministic setting and the evolution of the short-rate is modeled as a stochastic process.

Let us first have a look at the relationship between the stochastic discount factor and a zero-coupon bond. First it is necessary to point out, that the objective of a stochastic discount factor is being an *equivalent amount of currency*, whereas of the zero-coupon bond being a *value of a contract*. In the case of the rate r being deterministic, it follows: $D(t,T) = P(t,T)$ for each pair of time instants (t,T). However as mentioned above, the deterministic assumption is dropped and we rather look at stochastic processes. The relationship can be written as: the price of a zero-coupon bond is the expectation of the random variable $D(t,T)$ under a risk-neutral probability measure.

2.2.4 Types of interest rates

There are various kinds of implied interest rates, depending on the observation date and the underlying time interval of the rate. In the following we want to define the different types of interest rates and also point-out how they correlate.

1. the simply-compounded forward rate for $[T,S]$ prevailing at t is written as:

$$F(t;T,S) = \frac{1}{S-T}\left(\frac{P(t,T)}{P(t,S)} - 1\right) \tag{2.8}$$

which is equivalent to:

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

The forward rate is characterized by having three relevant times instants: $t \leq T \leq S$. Namely at current time t: sell a T-bond and buy $\frac{P(t,T)}{P(t,S)}$ S-bonds, resulting in a zero net investment. Then at expiry time T: pay one unit of currency. At maturity S: receive $\frac{P(t,T)}{P(t,S)}$ units of currency. In other terms forward rates are simply interest rates, locked in today for an investment in an upcoming time period. Moreover they are set consistently with the current structure of discount factors.
A forward rate can be defined through a forward rate agreement (FRA). This again describes a contract, within the above mentioned time instants, which gives its holder

an interest-rate payment for the period between T and S. To be precise this works as follows: at maturity S a fixed payment based on a fixed rate K is exchanged against a floating payment based on the spot rate $F(T,S)$, which resets at time T with maturity S. In other words this contract locks in the interest rate between two time instants at a desired value K with rates in the contract that are simply compounded. Basically at time S one receives $(S-T)KN$ units of currency with paying $(S-T)F(T,S)N$ with N denoting the contract nominal value. Hence for the value of the contract in S, we have:

$$N(S-T)(K-F(T,S))$$

Using formula (2.8) leads to the formula of the value of the contract at time t:

$$FRA(t,T,S,(S-T),N,K) = N\big[P(t,S)(S-T)K - P(t,T) + P(t,S)\big]$$

2. the simple spot rate for $[t,T]$, also known as LIBOR spot rate follows the definition:

$$F(t,T) = F(t;t,T) = \frac{1}{T-t}\left(\frac{1}{P(t,T)} - 1\right)$$

3. the continuously compounded forward rate for $[T,S]$ prevailing at t is given by:

$$R(t;T,S) = -\frac{\ln P(t,S) - \ln P(t,T)}{S-T}$$

which is equivalent to

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

4. the continuously compounded spot rate, which will in the following just be called *spot rate*, for $[0,T]$ is:

$$R(t,T) = R(t;t,T) = -\frac{\ln P(t,T)}{T-t} \tag{2.9}$$

5. the instantaneous forward rate, which will in the following just be called *forward rate*, and is intuitively a forward interest rate with its maturity begin very close to its expiry:

$$f(t,T) = \lim_{S\to T^+} R(t;T,S) = -\frac{\partial \ln P(t,T)}{\partial T} \tag{2.10}$$

6. the instantaneous short-rate, which is denoted simple as *short-rate* at time t is defined as:

$$r(t) = f(t,t) = lim_{T\to t} R(t,T) \tag{2.11}$$

We can now find equivalence of formula (2.10) along with the requirement $P(T,T) = 1$ as follows:

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

with the *No-Arbitrage* condition it results in the following equation, in terms of the risk neutral measure:

$$P(t,T) = \mathbb{E}_Q\left[e^{-\int_t^T r_s ds}|\mathcal{F}_t\right]$$

3 Two-Factor Hull-White Model

The present chapter describes the Two-Factor Hull-White model and the chapter is based on [cf.[6, 20, 21]], if not stated otherwise. The Two-Factor Hull-White model is one of the most important interest-rate models being used for risk-management purposes. It is a *No-Arbitrage* short-rate model and is able to fit the current term-structure of interest rates. Moreover the formulation is equivalent to the Two-Additive-Factor Gaussian model, which formulation will be used throughout this thesis because it is easier to implement and to interpret the parameters, as the representation is the sum of two correlated factors following a mean reverting Ornstein-Uhlenbeck process see [Definition 2.12]. Furthermore a deterministic function is added in order to fit the initial term-structure. First we will have a look at the risk-neutral environment in order to determine the arbitrage free process of contingent claims. However we will also discuss the transition from the risk-neutral measure, into the real-world measure. This process is also known as *change of measure*. Usually the change of measure is applied the other way around, but we need the diffusion dynamics of the two processes in the underlying model under the real-world measure in order to apply the Kalman Filter estimation procedure.
But before we will have a closer look at the Two-Factor model, we want to point out the main characteristics of the One-Factor Hull-White model, and also why we chose the Two-Factor model over the One-Factor model.
The One-Factor Hull-White model is an extension of the Vasicek[1] model and has been introduced in their papers in 1990. The poor fitting of the initial term-structure of interest rates brought Hull and White to add a time-varying parameter, which is chosen to be a deterministic function. Matching the model and the market term-structure of rates to the current time is equivalent to solving a system with an infinite number of equations, which is one for each possible maturity. In order to solve such a system, an infinite number of parameters have to be introduced, or equivalently a deterministic function of time. One advantage of the model is, that it implies normal distribution for the short-rate process at each time step. Furthermore zero-coupon bonds and options can be priced explicitly. The short-rate process evolves under the risk-neutral measure according to:

$$dr(t) = [\vartheta(t) - ar(t)]dt + \sigma dW(t) \tag{3.1}$$

In equation (3.1) a, σ are chosen to be positive constants and $W(t)$ denotes the Brownian motion, see [Definition 2.1]. Furthermore ϑ is determined to fit the term-structure of

[1] The model describes the evolution of the short-rate process as:

$$dr(t) = k(\theta - r(t))dt + \sigma dW(t)$$

interest rates being currently observed at the market and is therefore expressed through:

$$\vartheta(t) = \frac{\partial f^M(0,t)}{\partial t} + a f^M(0,t) + \frac{\sigma^2}{2a}\left(1 - e^{-2at}\right) \tag{3.2}$$

A more detailed explanation for (3.2) can be found in [[6], p. 72 ff.] The first fraction on the rhs. of equation (3.2) denotes the partial derivative of the market instantaneous forward rate with respect to its second argument. The instantaneous forward rate again at time 0 and maturing at time T is defined as:

$$f^M(0,T) = -\frac{\partial \ln P^M(0,T)}{\partial T}$$

Even though the One-Factor model has good tractability, because of the lack of free calibration parameters, the model is not capable of reproducing a satisfactorily large volatility surface. Furthermore highly negative rates are generated too often. That is why we will from now on focus on the Two-Factor Hull-White model.

As already mentioned in [Section 2.2] the knowledge of the short-rate $r(t)$ and its distributional properties lead to the bond price formula, which is denoted by:

$$P(t,T) = \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^T r(s)ds}\right] \tag{3.3}$$

Then again with all bond prices $P(t,T)$ at a given time t one can reconstruct the whole zero-coupon interest-rate curve at the same time t. Further the evolution of the whole curve can be expressed through the evolution of one quantity, the short-rate. That is why choosing a poor model for the short-rate leads inevitably to a poor model for the evolution of the yield curve.

Having now a look at the setting of the Two-Factor Hull-White model with the short rate evolving in the risk-adjusted measure according to:

$$dr(t) = \left[\theta(t) + u(t) - \bar{a}r(t)\right]dt + \sigma_1 dZ_1(t), \quad r(0) = r_0 \tag{3.4}$$

and the stochastic mean-reversion level follows the following SDE:

$$du(t) = -\bar{b}u(t)dt + \sigma_2 dZ_2(t), \quad u(0) = 0 \tag{3.5}$$

where (Z_1, Z_2) is a two-dimensional Brownian motion and their correlation can be defined as follows:

$$dZ_1(t)dZ_2(t) = \bar{\rho}dt \quad with \ -1 \leq \bar{\rho} \leq 1$$

In equation (3.4) the deterministic function θ is chosen to fit the initial term-structure. The remaining parameters in this equation $\rho, \bar{a}, \bar{b}, \sigma_1$ and σ_2 are positive constants. Integration

of equation (3.4) and (3.5) leads to the following representation of the short-rate and corresponding mean-reversion level:

$$r(t) = r(s)e^{-\bar{a}(t-s)} + \int_s^t \theta(\nu)e^{-\bar{a}(t-\nu)}d\nu + \int_s^t u(\nu)e^{-\bar{a}(t-\nu)}d\nu + \sigma_1 \int_s^t e^{-\bar{a}(t-\nu)}dZ_1(\nu),\ s < t$$

$$u(t) = u(s)e^{-\bar{b}(t-s)} + \sigma_2 \int_s^t e^{-\bar{b}(t-\nu)}dZ_2(\nu)$$

Assuming that $\bar{a} \neq \bar{b}$ we can write:

$$\int_s^t u(\nu)e^{-\bar{a}(t-\nu)}d\nu = \int_s^t u(s)e^{-\bar{b}(\nu-s)-\bar{a}(t-\nu)}d\nu + \sigma_2 \int_s^t e^{-\bar{a}(t-\nu)} \int_s^\nu e^{-\bar{b}(\nu-x)}dZ_2(x)d\nu$$

$$= u(s)\frac{e^{-\bar{b}(t-s)} - e^{-\bar{a}(t-s)}}{\bar{a}-\bar{b}} + \sigma_2 e^{-\bar{a}t} \int_s^t e^{(\bar{a}-\bar{b})\nu} \int_s^\nu e^{\bar{b}x}dZ_2(x)d\nu$$

Following integration by parts leads to:

$$\int_s^t e^{(\bar{a}-\bar{b})\nu} \int_s^\nu e^{\bar{b}x}dZ_2(x)d\nu = \frac{1}{\bar{a}-\bar{b}} \int_s^t \left(\int_s^\nu e^{\bar{b}x}dZ_2(x) \right) d_\nu \left(e^{(\bar{a}-\bar{b})\nu} \right)$$

$$= \frac{1}{\bar{a}-\bar{b}} \left[e^{(\bar{a}-\bar{b})t} \int_s^t e^{\bar{b}x}dZ_2(x) - \int_s^t e^{(\bar{a}-\bar{b})\nu} d_\nu \left(\int_s^\nu e^{\bar{b}x}dZ_2(x) \right) \right]$$

$$= \frac{1}{\bar{a}-\bar{b}} \int_s^t \left[e^{(\bar{a}-\bar{b})t} - e^{(\bar{a}-\bar{b})\nu} \right] d_\nu \left(\int_s^\nu e^{\bar{b}x}dZ_2(x) \right)$$

$$= \frac{1}{\bar{a}-\bar{b}} \int_s^t \left[e^{\bar{a}t-\bar{b}(t-\nu)} - e^{\bar{a}\nu} \right] dZ_2(\nu)$$

Then again we can write the short-rate as:

$$r(t) = r(s)e^{-\bar{a}(t-s)} + \int_s^t \theta(\nu)e^{-\bar{a}(t-\nu)}d\nu + \sigma_1 \int_s^t e^{-\bar{a}(t-\nu)}dZ_1(\nu)$$

$$+u(s)\frac{e^{-\bar{b}(t-s)} - e^{-\bar{a}(t-s)}}{\bar{a}-\bar{b}} + \frac{\sigma_2}{\bar{a}-\bar{b}} \int_s^t \left[e^{-\bar{b}(t-\nu)} - e^{-\bar{a}(t-\nu)} \right] dZ_2(\nu)$$

This, and using $u(0) = 0$, brings us to the formulation of the short-rate as:

$$r(t) = r_0 e^{-\bar{a}t} + \int_0^t \theta(\nu) e^{-\bar{a}(t-\nu)} d\nu + \sigma_1 \int_0^t e^{-\bar{a}(t-\nu)} dZ_1(\nu)$$

$$+ \frac{\sigma_2}{\bar{a} - \bar{b}} \int_0^t \left[e^{-\bar{b}(t-\nu) - e^{-\bar{a}(t-\nu)}} \right] dZ_2(\nu)$$

After some integration and transformation and with the definition of:

$$\sigma_3 = \sqrt{\sigma_1^2 + \frac{\sigma_2^2}{(\bar{a} - \bar{b})^2} + 2\bar{\rho} \frac{\sigma_1 \sigma_2}{(\bar{b} - \bar{a})}}$$

$$dZ_3(t) = \frac{\sigma_1 dZ_1(t) - \frac{\sigma_2}{\bar{a} - \bar{b}} dZ_2(t)}{\sigma_3}$$

$$\sigma_4 = \frac{\sigma_2}{(\bar{a} - \bar{b})}$$

The short-rate can finally be written as:

$$r(t) = r_0 e^{-\bar{a}t} + \int_0^t \theta(\nu) e^{-\bar{a}(t-\nu)} d\nu + \int_0^t e^{-\bar{a}(t-\nu)} \left[\sigma_1 dZ_1(\nu) + \frac{\sigma_2}{\bar{b} - \bar{a}} \right] dZ_2(\nu)$$
$$+ \frac{\sigma_2}{\bar{a} - \bar{b}} \int_0^t e^{-\bar{b}(t-\nu)} dZ_2(\nu)$$

$$= r_0 e^{-\bar{a}t} + \int_0^t \theta(\nu) e^{-\bar{a}(t-\nu)} d\nu + \sigma_3 \int_0^t e^{-\bar{a}(t-\nu)} dZ_3(\nu) + \sigma_4 \int_0^t e^{-\bar{b}(t-\nu)} dZ_2(\nu)$$

In order to find the analogy with the Two-Factor Additive Gaussian Model, we want to set following parameters:

$$a = \bar{a}$$

$$b = \bar{b}$$

$$\sigma = \sigma_3$$

$$\eta = \sigma_4$$

$$\rho = \frac{\sigma_1 \bar{\rho} - \sigma_4}{\sigma_3}$$

$$\varphi(t) = r_0 e^{-\bar{a}t} + \int_0^t \theta(\nu) e^{-\bar{a}(t-\nu)} d\nu$$

The equivalence is extremely useful, as it is easier to interpret the different parameters and their influence on the price and volatility structure in the setting of the Gaussian model. That is why from now on we will regard the Gaussian model in this thesis.
The dynamics of the short rate process under the risk-neutral measure $\mathbb{Q}$ in the Gaussian model are expressed as follows:

$$r(t) = x(t) + y(t) + \varphi(t) \quad \text{with } r(0) = r_0 \tag{3.6}$$

where $\{x(t) : t \geq 0\}$ and $\{y(t) : t \geq 0\}$ are One-Factor Hull-White processes and they satisfy following SDEs:

$$\begin{aligned} dx(t) &= -ax(t)dt + \sigma dW_1(t), \quad x(0) = 0 \\ dy(t) &= -by(t)dt + \eta dW_2(t), \quad y(0) = 0 \end{aligned} \tag{3.7}$$

where (W_1, W_2) is a two-dimensional Brownian motion with correlation ρ as follows:

$$dW_1(t)dW_2(t) = \rho dt$$

The in equations (3.6) and (3.7) mentioned parameters r_0, a, b, σ, η are positive constants. Furthermore we can specify $-1 \leq \rho \leq 1$ and in particular: $\varphi(0) = r_0$. The deterministic function $\varphi(t)$ is well-defined in the time interval $[0, T^*]$. Usually the to be regarded time horizon T^* extends over the duration of $10, 30$ or 50 years. Integration of the above stated dynamics of the short rate see formula (3.7) $\forall s < t$ leads to following representation of the short rate:

$$\begin{aligned} r(t) = x(s)e^{-a(t-s)} + y(s)e^{-b(t-s)} + \sigma \int_s^t e^{-a(t-u)} dW_1(u) \\ + \eta \int_s^t e^{-b(t-u)} dW_2(u) + \varphi(t) \end{aligned}$$

The volatility of the instantaneous forward rate can be written as:

$$\sigma_f(t,T) = \sqrt{\sigma^2 e^{-2a(T-t)} + \eta^2 e^{-2b(T-t)} + 2\rho\sigma\eta e^{-(a+b)(T-t)}}$$

Now let $\mathcal{F}_t$ denote the sigma-field generated by the pair (x, y) up to time t. That is why regarding $r(t)$ conditioned on $\mathcal{F}_s$ is normally distributed and mean and variance are given as:

$$\mathbb{E}\big[r(t)|\mathcal{F}_s\big] = x(s)e^{-a(t-s)} + y(s)e^{-b(t-s)} + \varphi(t)$$

$$Var\big[r(t)|\mathcal{F}_s\big] = \frac{\sigma^2}{2a}\Big[1 - e^{-2a(t-s)}\Big] + \frac{\eta^2}{2b}\Big[1 - e^{-2b(t-s)}\Big] + 2\rho\frac{\sigma\eta}{a+b}\Big[1 - e^{-(a+b)(t-s)}\Big]$$

In particular, we can write $r(t)$ as:

$$r(t) = \sigma \int_0^t e^{-a(t-u)} dW_1(u) + \eta \int_0^t e^{-b(t-u)} dW_2(u) + \varphi(t)$$

As the Brownian motions in equation (3.7) are correlated, the two processes can also be expressed using independent Brownian motions[2]:

$$\begin{aligned} dx(t) &= -ax(t)dt + \sigma d\tilde{W}_1(t) \\ dy(t) &= -by(t)dt + \eta\rho d\tilde{W}_1(t) + \eta\sqrt{1-\rho^2} d\tilde{W}_2(t) \end{aligned} \tag{3.8}$$

Therefore it holds:

$$\begin{aligned} dW_1(t) &= d\tilde{W}_1(t) \\ dW_2(t) &= \rho d\tilde{W}_1(t) + \sqrt{1-\rho^2} d\tilde{W}_2(t) \end{aligned}$$

That is why the short rate conditioned on $\mathcal{F}_s$ can now be written as:

$$r(t)|\mathcal{F}_s = x(s)e^{-a(t-s)} + y(s)e^{-b(t-s)} + \sigma \int_s^t e^{-a(t-u)} d\tilde{W}_1(u)$$

$$+\eta\rho \int_s^t e^{-b(t-u)} d\tilde{W}_1(u) + \eta\sqrt{1-\rho^2} \int_s^t e^{-b(t-u)} d\tilde{W}_2(u) + \varphi(t)$$

The deterministic shift $\varphi(t)$ is the remaining part that needs to be solved. This will be treated in the subsequent section.

3.1 Pricing of a Zero-Coupon Bond

The present section is based on [cf.[6, 16]]. With the use of the numéraire-dependent pricing formula see [Theorem 2.1], where the money-market account denotes the numéraire. Then the price of a zero-coupon bond under the risk-neutral measure $\mathbb{Q}$ at time t with maturity T can be written as:

$$P(t,T) = \mathbb{E}_{\mathbb{Q}}\left[e^{-\int_t^T r_s ds} | \mathcal{F}_t \right] \tag{3.9}$$

In order to solve the above stated equation, we start by writing the explicit expression for the two factors $x(t)$ and $y(t)$ and then continue to find $\varphi(t)$ in formula (3.17). As a

[2] This procedure is equivalent to undertake a Cholesky decomposition on the variance-covariance matrix of the pair $(W_1(t), W_2(t))$.

first step we consider the part in the exponential. For each pair (t,T) the random variable $I(t,T)_x$ can be defined through:

$$I(t,T)_x := \int_t^T x(u)du \tag{3.10}$$

Rewriting (3.10) and considering $x(u)$ with its expectation given $\mathcal{F}_s$, leads to:

$$I(t,T)_x = \int_t^T x(t)e^{-a(u-t)}du + \sigma\int_t^T\int_u^t e^{-a(u-s)}dW_s du \tag{3.11}$$

The first term of equation (3.11) is solved by simple integration, however for the second term we have:

$$\sigma\int_t^T\int_u^t e^{-a(u-s)}dW_s du = \frac{\sigma}{a}\int_t^T \left[1-e^{-a(T-u)}\right]dW_u$$

Therefore it follows:

$$\int_t^T x(u)du = \frac{1-e^{-a(T-t)}}{a}x(t) + \frac{\sigma}{a}\int_t^T \left[1-e^{-a(T-u)}\right]dW_u$$

Analogously $I(t,T)_y$ can be expressed through:

$$I(t,T)_y := \int_t^T y(u)du = \frac{1-e^{-b(T-t)}}{b}y(t) + \frac{\eta}{b}\int_t^T \left[1-e^{-b(T-u)}\right]dW_{2u}$$

The condition expectation mean for the above stated integral is written as follows:

$$M_x(t,T) = \frac{1-e^{-a(T-t)}}{a}x(t) \tag{3.12}$$

Having now a look at the corresponding variance:

$$\begin{aligned} V_x(t,T) &= Var\left[\frac{\sigma}{a}\int_t^T \left[1-e^{-a(T-u)}\right]dW_u\right] \\ &= \frac{\sigma^2}{a^2}\int_t^T \mathbb{E}\left[1-e^{-a(T-u)}\right]^2 du \\ &= \frac{\sigma^2}{a^2}\left[T-t+\frac{2}{a}e^{-a(T-t)} - \frac{1}{2a}e^{-2a(T-t)} - \frac{3}{2a}\right] \end{aligned} \tag{3.13}$$

Following analogously formula (3.12) for $y(t)$ with replacing the corresponding terms and combining them afterwards leads to:

$$M(t,T) = M_x(t,T) + M_y(t,T) = \frac{1-e^{-a(T-t)}}{a}x(t) + \frac{1-e^{-b(T-t)}}{b}y(t) \tag{3.14}$$

Then again following formula (3.13) for the corresponding variance of $y(t)$ and combining them afterwards leads to the variance of $I(t,T) := \int_t^T [x(u) + y(u)]du$:

$$V(t,T) = V_x(t,T) + V_y(t,T) + 2\rho\sqrt{V_x(t,T)}\sqrt{V_y(t,T)} \tag{3.15}$$

$$= \frac{\sigma^2}{a^2}\left[T - t + \frac{2}{a}e^{-a(T-t)} - \frac{1}{2a}e^{-2a(T-t)} - \frac{3}{2a}\right]$$

$$+\frac{\eta^2}{b^2}\left[T - t + \frac{2}{b}e^{-b(T-t)} - \frac{1}{2b}e^{-2b(T-t)} - \frac{3}{2b}\right]$$

$$+2\rho\frac{\sigma\eta}{ab}\left[T - t + \frac{e^{-a(T-t)} - 1}{a} + \frac{e^{-b(T-t)} - 1}{b} - \frac{e^{-(a+b)(T-t)} - 1}{a+b}\right]$$

A detailed explanation with proof of (3.15) can be found in [[6], p. 169 ff.]. Along with the above given formulas, $\varphi(t)$ being a deterministic function, $x(t)$, $y(t)$ being normally distributed, the integral of the short rate $Z = \int_t^T r(s)ds$ is normally distributed with mean $\mu_z = M(t,T) + \int_t^T \varphi(u)du$ and corresponding variance $\sigma_z^2 = V(t,T)$. Then again e^Z is log-normally distributed with: $\mathbb{E}\{e^Z\} = e^{\mu_z + \frac{1}{2}\sigma_z^2}$. Therefore the price of a zero-coupon bond can be expressed as:

$$P(t,T) = exp\left\{-\int_t^T \varphi(u)du - \frac{1 - e^{-a(T-t)}}{a}x(t) - \frac{1 - e^{-b(T-t)}}{b}y(t) + \frac{1}{2}V(t,T)\right\} \tag{3.16}$$

The model is said to fit the currently-observed term-structure of discount factors for each maturity T if, the discount factor $P(0,T)$ matches the corresponding modeled discount factor $P^M(0,T)$. Therefore the following equation holds:

$$P^M(0,T) = exp\left\{-\int_0^T \varphi(u)du + \frac{1}{2}V(0,T)\right\}$$

Assuming that the current market term structure of bond prices is a sufficiently smooth function $T \longmapsto P^M(0,T)$, the modeled instantaneous forward rate, evaluated at time 0 with maturity T, implied by $P^M(0,T)$ is given as:

$$f^M(0,T) = -\frac{\partial \ln P^M(0,T)}{\partial T} = \varphi(T) - \frac{\partial}{\partial T}\frac{1}{2}V(0,T)$$

Therefore we have:

$$\varphi(T) = f^M(0,T) + \frac{\sigma^2}{2a^2}\left(1 - e^{-aT}\right)^2 + \frac{\eta^2}{2b^2}\left(1 - e^{-bT}\right)^2 + \rho\frac{\sigma\eta}{ab}\left(1 - e^{-aT}\right)\left(1 - e^{-bT}\right) \tag{3.17}$$

if, and only if

$$exp\Big\{-\int_t^T \varphi(u)du\Big\} = exp\Big\{-\int_0^T \varphi(u)du\Big\} \exp\Big\{\int_0^t \varphi(u)du\Big\}$$

$$= \frac{P^M(0,T)exp\{-\frac{1}{2}V(0,T)\}}{P^M(0,t)exp\{-\frac{1}{2}V(0,t)\}}$$

$$= \frac{P^M(0,T)}{P^M(0,t)}exp\Big\{-\frac{1}{2}[V(0,T)-V(0,t)]\Big\} \tag{3.18}$$

As a last step of this section the price of a zero-coupon bond is defined. Combining equation (3.16) with (3.18) leads to following representation of the price of a zero-coupon bond at time t with maturity T:

$$P(t,T) = \frac{P^M(0,T)}{P^M(0,t)} e^{\frac{1}{2}\Big[V(t,T)-V(0,T)+V(0,t)\Big] - \frac{1-e^{-a(T-t)}}{a}x(t) - \frac{1-e^{-b(T-t)}}{b}y(t)} \tag{3.19}$$

3.2 Probability of Negative Rates

Taking a look at the currently-observed term structure of discount factors, the expected instantaneous short-rate at time t is expressed as:

$$\mu_r(t) := \mathbb{E}[r(t)] = f^M(0,t) + \frac{\sigma^2}{2a^2}(1-e^{-at})^2 + \frac{\eta^2}{2b^2}(1-e^{-bt})^2 + \rho\frac{\sigma\eta}{ab}(1-e^{-at})(1-e^{-bt})$$

The corresponding variance $\sigma_r^2(t)$ of the instantaneous short-rate at time t is given as:

$$\sigma_r^2(t) = \mathbb{V}[r(t)] = \frac{\sigma^2}{2a}(1-e^{-2at}) + \frac{\eta^2}{2b}(1-e^{-2bt}) + 2\frac{\rho\sigma\eta}{a+b}(1-e^{-(a+b)t})$$

For a long time the possibility of attaining negative rates within this model has been mentioned as a drawback. However due to the current market environment, such as high implied volatilities and extremely low or even negative interest rates. In the risk neutral setting at time t the probability of a negative short rate can be written as:

$$Q[r(t) < 0] = \Phi\left(-\frac{\mu_r(t)}{\sigma_r(t)}\right)$$

Where Φ denotes the standard normal cumulative distribution function. Moreover the limit distribution of the short rate process is Gaussian with mean $\mu_r(\infty)$ and variance $\sigma_r^2(\infty)$ given as:

$$\mu_r(\infty) := lim_{t\to\infty}\mathbb{E}[r(t)] = f^M(0,\infty) + \frac{\sigma^2}{2a^2} + \frac{\eta^2}{2b^2} + \rho\frac{\sigma\eta}{ab}$$

$$\sigma_r^2(\infty) := lim_{t\to\infty}\mathbb{V}[r(t)] = \frac{\sigma^2}{2a} + \frac{\eta^2}{2b} + 2\frac{\rho\sigma\eta}{a+b}$$

3.3 Change of measure - model dynamics under the real world measure

The diffusion dynamics see equations in (3.7) of the underlying Two-Additive-Factor Gaussian model are formulated in the risk-neutral world, in other terms under the probability measure $\mathbb{Q}$. However when it comes to the application of the Kalman Filter the estimated bond prices are given under the real world measure $\mathbb{P}$. That is why the dynamics of the stochastic process also have to be formulated under the real world measure and therefore need to undertake a change of measure. That is why we want to change the measure from the risk-neutral to the real-world measure. Usually the change of measure is defined the other way round and follows the rule of *Girsanov's Theorem* see [Definition 2.14]. The concept for the other way round is still the same. What we aim to accomplish is to find the factor of the asset's volatility by which the drift of every asset is corrected, which can be understood as the asset's exposure to risk. Which is commonly called the *market price of risk*, as it describes the price investors are willing to pay to run a certain level of risk. However in the risk-neutral setting every asset has the same drift. In order to see how this works,we will now consider the diffusion dynamics of the factors with independent Brownian motions W_1 and W_2 under $\mathbb{Q}$, analogously to formula (3.8):

$$dx(t) = -ax(t)dt + \sigma dW_1^{\mathbb{Q}}(t) \tag{3.20}$$

$$dy(t) = -by(t)dt + \eta\rho dW_1^{\mathbb{Q}}(t) + \eta\sqrt{1-\rho^2}dW_2^{\mathbb{Q}}(t)$$

Formulating the model dynamics from equation (3.20) in general vector notation, which is written in a 2-factor SDE, we have:

$$dF(t) = \Lambda^*(F(t) - \mu^*)dt + \sum S(t)dW(t)^{\mathbb{Q}} \tag{3.21}$$

with $F(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}$, $W(t)^{\mathbb{Q}}$ denoting a 2-dimensional independent standard Brownian motion, $\sum$ is a 2×2 and possibly non-diagonal and asymmetric, and $S(t)$ a diagonal matrix, with the k-th diagonal element written as:[3]

$$[S(t)]_{kk} = \sqrt{\alpha_k}$$

[3] This thesis only considers constant volatility however if this restriction is dropped the diagonal matrix is stated as:

$$[S(t)]_{kk} = \sqrt{\alpha_k + \beta_k' F_t}$$

With the diffusion dynamics specified under $\mathbb{Q}$ and considering the constant volatility case of $S(t)$ we can plug in those formulas into the general vector notation, and get following values for the parameters:

$$\Lambda^* = \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}$$

$$\mu^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\sum = \begin{bmatrix} 1 & 0 \\ \frac{\eta\rho}{\sigma} & \sqrt{1-\rho^2} \end{bmatrix}$$

$$S(t) = S = \begin{bmatrix} \sqrt{\sigma^2} & 0 \\ 0 & \sqrt{\eta^2} \end{bmatrix} = \begin{bmatrix} \sigma & 0 \\ 0 & \eta \end{bmatrix}$$

Regarding the vector $\lambda(t)$ as a well-defined Radon-Nikodym process, then according to Girsanov's theorem see [Definition 2.14], the independent vector process $W^{\mathbb{P}}(t)$ can be defined as a 2-dimensional Brownian motion under $\mathbb{P}$. The definition of the change of measure from the risk-neutral measure $\mathbb{Q}$ to the real-world measure $\mathbb{P}$ can now be represented in the following way:

$$dW(t)^{\mathbb{P}} = -\lambda(t)dt + dW(t)^{\mathbb{Q}} \ , \ W^{\mathbb{P}}(0) = 0$$

The process $\lambda(t)$ however governs the transition form the risk-neutral to the real-world measure. Let us now regard this transition in the model definition, we have that:[4]

$$\Lambda^* = \Lambda \tag{3.22}$$

$$-\Lambda\mu = \sum \Psi\alpha \tag{3.23}$$

The general vector notation of the model definition under the real-world measure following formula (3.8) and (3.21) is given as:

$$dF(t) = \Lambda(F(t) - \mu)dt + \sum S(t)dW_t^{\mathbb{P}} \tag{3.24}$$

For equation (3.24) we can specify the parameters as:

$$\Lambda = \begin{bmatrix} -a & 0 \\ 0 & -b \end{bmatrix}, -\Lambda\mu = \begin{bmatrix} \lambda_1\sigma \\ \lambda_1\eta\rho + \lambda_2\eta\sqrt{1-\rho^2} \end{bmatrix} = \begin{bmatrix} \psi_1\sigma^2 \\ \psi_1\sigma\eta\rho + \psi_2\eta^2\sqrt{1-\rho^2} \end{bmatrix}$$

[4] Again if the assumption of constant volatility is dropped, we would write:

$$\Lambda^* = \Lambda - \sum \Psi\mathbb{B}' \quad and \quad -\Lambda\mu = \sum \Psi\alpha$$

The newly introduced market price of risk for factor k which is proportional to its instantaneous standard derivation, is now chosen to be:[5]

$$\psi_k := \frac{\lambda_k}{\sqrt{\alpha_k}}$$

We follow the insurance industry standard by assuming the market price of risk to be constant. However as we will see in (3.26), the market price of risk itself depends on the local long run risk premium functions $d_x(t)$ and $d_y(t)$ which technically can also be chosen as step or linear functions. We assume the market price of risk is time-invariant and will be reformulated in (3.26). In fact modeling the market price of risk in such a way, assumes that the price investors pay for a certain risk is constant for the whole modelling horizon. To be more precise it assumes that the market price of risk is stays always the same regardless the time, or to be more precise the market situation. This does not match reality, anyhow this notation is preferred, as its intuitive and good tractable in the model.
Applying the theorem of Girsanov see [Definition 2.14] on the two correlated processes, they can be formulated under the real-world measure as:

$$dx(t) = \Big[\psi_1\sigma^2 - ax(t)\Big]dt + \sigma dW_1^{\mathbb{P}}(t)$$

$$dy(t) = \Big[\psi_1\sigma\eta\rho - \psi_2\eta^2\sqrt{1-\rho^2} - by(t)\Big]dt + \eta\rho dW_1^{\mathbb{P}}(t) + \eta\sqrt{1-\rho^2}dW_2^{\mathbb{P}}(t) \tag{3.25}$$

In order to specify ψ_1, and ψ_2 from formulas (3.25), we follow [[21], p.681, ff.] The market price of risk is there defined in a way that the two processes $x(t)$ and $y(t)$ are again formulated as Ornstein-Uhlenbeck processes.

$$\begin{bmatrix}\psi_1\\ \psi_2\end{bmatrix} = \begin{bmatrix}\frac{ad_x(t)}{\sigma^2}\\ \frac{bd_y(t)}{\eta^2\sqrt{1-\rho^2}} + \frac{\rho ad_x(t)}{\sigma\eta\sqrt{1-\rho^2}}\end{bmatrix} \tag{3.26}$$

In general three types of functions for $d_x(t)$ and $d_y(t)$ can be distinguished: the constant, the step and the linear function. These functions represent the long run risk premium for each risk factor and may depend on t. The three types are:

- constant:

$$d_x(t) = d_x$$
$$d_y(t) = d_y$$

[5] Again if constant volatility is not assumed we would consider:

$$\psi_k = \frac{\lambda_k}{\sqrt{\alpha_k + \beta_k' F_t}}$$

- step:

$$d_x(t) = \mathbb{1}_{t\leq\tau} d_x + \mathbb{1}_{t>\tau} l_x$$
$$d_y(t) = \mathbb{1}_{t\leq\tau} d_y + \mathbb{1}_{t>\tau} l_y$$

- linear:

$$d_x(t) = \mathbb{1}_{t\leq\tau}(1 - m_x t) d_x + \mathbb{1}_{t>\tau} l_x$$
$$d_y(t) = \mathbb{1}_{t\leq\tau}(1 - m_y t) d_y + \mathbb{1}_{t>\tau} l_y$$

Where d_x, d_y, l_x, l_y, m_x and m_y are real valued constants and $\mathbb{1}_A$ represents the indicator function of a subset A. As already mentioned we follow the industry standard by defining the market price of risk through a constant function. The reformulation of the dynamics see (3.25) of the processes has following representation:

$$dx(t) = a\Big(d_x - x(t)\Big)dt + \sigma dW_1^{\mathbb{P}}(t) \tag{3.27}$$
$$dy(t) = b\Big(d_y - y(t)\Big)dt + \eta dW_2^{\mathbb{P}}(t)$$

with W_1 and W_2 denoting two correlated Brownian motions under $\mathbb{P}$.
The processes $x(t)$ and $y(t)$ are still Ornstein-Uhlenbeck processes and their solution is given as:

$$x(t) = \int_0^t e^{-a(t-u)} a d_x du + \sigma \int_0^t e^{-a(t-u)} dW_1^{\mathbb{P}}(u)$$
$$y(t) = \int_0^t e^{-b(t-u)} b d_y du + \eta \int_0^t e^{-b(t-u)} dW_2^{\mathbb{P}}(u)$$

In general notation the mean reversion level of each process at time t is denoted by $d_x(t)$, respectively $d_y(t)$. We recall that the sum of $x(t)$ and $y(t)$ and the deterministic function $\varphi(t)$ under the risk neutral measure adds up to the instantaneous short rate $r(t)$. That is why changing the measure, leads also to another mean reversion level at time point t from 0 to $d_x(t)$ for the process x, and analogously to $d_y(t)$ for the process y. Hence $d_x(t) + d_y(t)$ can be interpreted as the local long run risk premium of the short rate, which is the amount added in the real world to the risk-neutral short rate, if $d_x(t) + d_y(t)$ is chosen to be constant over time. If this amount is negative, it means that future bond prices increase in expectation compared to the risk-neutral setting. Therefore a risk averse investor gets compensated for the risk he takes. This means in a market where investors are risk averse, future interest rates tend to be lower under the real world measure than in the risk-neutral. That is why $d_x(t)$ and $d_y(t)$ can be called the local long run risk premium, where the corresponding risk factor is mean reverting to at time t.

3.4 Bond price formula under the real world measure

The underlying section uses results from [21]. In the following we want to validate that the price of a zero-coupon bond under $\mathbb{P}$ is obtained through the same analytic formula as under $\mathbb{Q}$, see (3.19). The price of a zero-coupon under the real-world measure is calculated using the conditional expectation:

$$\frac{P(t,T)}{X_{P(t,T)}(t)} = \mathbf{E}^{\mathbb{P}}\left[\frac{P(T,T)}{X_{P(t,T)}(T)}|\mathcal{F}_t\right]$$

Where $X_{P(t,T)}(t)$ represents the cash-flow, that discounts the zero-coupon bond such that the discounted price-process is a martingale under $\mathbb{P}$. The dynamic of $X_{P(t,T)}(t)$ coincides with the deterministic part of the zero-coupon price dynamic and is specified by the change of measure:

$$dX_{P(t,T)}(t) = X_{P(t,T)}(t)[r(t) - B_x(t,x)ad_x - B_y(t,y)bd_y]dt \ , \ X_{P(t,T)}(0) = 1$$

With the specification of $B_x(t,x) = \frac{1-e^{-a(t,T)}}{a}$ and analogously $B_y(t,y) = \frac{1-e^{-b(t,T)}}{b}$.
The solution is given as:

$$X_{P(t,T)}(t) = e^{\int_0^t (r(u) - B_x(u,x)ad_x - B_y(u,y)bd_y)du} \tag{3.28}$$

The proof for the above stated formula, (3.28) can be found in [[21], p. 699 f.].Therefore the price of a zero-coupon bond at time t is given by:

$$P(t,T) = \mathbf{E}^{\mathbb{P}}\left[\frac{X_{P(t,T)}(t)}{X_{P(t,T)}(T)}|\mathcal{F}_t\right]$$

The ratio in the expectation amounts to:

$$\frac{X_{P(t,T)}(t)}{X_{P(t,T)}(T)} = e^{\int_t^T -(r(u) - B_x(u,x)ad_x - B_y(u,y)bd_y)du}$$

The distribution of the integral in the exponent is given as:

$$J(t,T) := \int_t^T (r(u) - B_x(u,x)ad_x - B_y(u,y)bd_y)du$$

Then $J(t,T)$ is normally distributed with mean:

$$N(t,T) = \int_t^T \varphi(u)du + \frac{1-e^{-a(T-t)}}{a}x(t) + \frac{1-e^{-b(T-t)}}{b}y(t)$$

The corresponding variance is given as:

$$V(t,T) = \frac{\sigma^2}{a^2}\left[T - t + \frac{2}{a}e^{-a(T-t)} - \frac{1}{2a}e^{-2a(T-t)} - \frac{3}{2a}\right]$$

$$+\frac{\eta^2}{b^2}\left[T - t + \frac{2}{b}e^{-b(T-t)} - \frac{1}{2b}e^{-2b(T-t)} - \frac{3}{2b}\right]$$

$$+2\rho\frac{\sigma\eta}{ab}\left[T - t + \frac{e^{-a(T-t)} - 1}{a} + \frac{e^{-b(T-t)} - 1}{b} - \frac{e^{-(a+b)(T-t)} - 1}{a+b}\right]$$

The variance is equivalent to the variance in the risk-neutral world, see (3.15). It can be shown that also the mean has the same form as in the risk-neutral world as the terms: $B_x(u,x)ad_x$ and $B_y(u,y)ad_y$ in $J(t,T)$ cancel out in calculations. The proof can be found in [[21], p. 700 ff.]. Therefore the expression $e^{-J(t,T)}$ is log-normally distributed and the zero-coupon bond price under $\mathbb{P}$ is given by the same formula as under $\mathbb{Q}$, see (3.19) with the equation of (3.18):

$$P(t,T) = \mathbf{E}^{\mathbb{P}}[e^{-\int_t^T (r(u) - B_x(u,x)ad_x - B_y(u,y)bd_y)du}|\mathcal{F}_t]$$

$$= e^{-M(t,T)+\frac{1}{2}V(t,T)}$$

$$= e^{-\int_t^T \varphi(u)du + \frac{1-e^{-a(T-t)}}{a}x(t) + \frac{1-e^{-b(T-t)}}{b}y(t) + \frac{1}{2}V(t,T)}$$

4 Nelson-Siegel-Svensson

The present chapter is based on [cf.[22, 23, 24, 25, 26, 27, 28]]. As already specified in [Section 2.2], the Nelson-Siegel and the extended Svensson model belong to the class of parametric methods. This type of method represents explicitly the function in form of a *single-piece function*. The form of the function depends on the time to maturity, but also on selected k parameters, which are chosen to approximate as good as possible. The drawback of this method is the lack of flexibility, which is due to the choice of a pre-specified function. However the advantages of this method are the easy application and the good interpretability of the model parameters.

4.1 Nelson-Siegel model

The Nelson-Siegel model is commonly used when it comes to fitting the term-structure of interest rates and was introduced in 1987. The model has little parametrization, but provides statistically and economically precise results. Initially Nelson and Siegel described a simple model containing three latent factors to fit the yield curve of the bond market. Nelson and Siegel faced the challenge of using only three variables to describe this complex curve with great performance. The described factors are also known as the level, the slope and the curvature of the yield curve. However since the introduction many researches contributed different versions to this model in order to improve the prediction. The different contributions can be classified as the *Dynamic Nelson-Siegel model*, *the Arbitrage-free Nelson-Siegel model* and the *Dynamic Generalized Nelson-Siegel model*. Moreover some researchers also extended the model to some more factors. To name the *four-factor Svensson model*, with which we will work with, or the *five-factor dynamic generalized model*.
The background of this model is the cognition of Nelson, that interest rates, such as the forward rate $f(t,T)$, can easily be modeled as differential equations of second order. Therefore the approximation function is written as a Laguerre Differential equation:

$$t \cdot L''(t) + (1-t) \cdot L'(t) + nL(t) = 0 \tag{4.1}$$

In equation (4.1) we have $n \in \mathbb{N}$ and $t \in \mathbb{R}$. The Laguerre Function is chosen to represent the approximation functions, as they represent the solution of the differential equation. Moreover the class of this functions is able to represents the majority of term-structure shapes.
The spot-rate is defined via differential equations, therefore the solutions relate to the forward-rates. The forward-rate function is represented as a Laguerre-Function of first order and a constant λ is added. In the following we will denote the forward rate $f(t,T)$

as $f(\tau)$, where τ represents the time to maturity, which is $(T-t)$.

$$f(t,T) = f(\tau) = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}' \begin{bmatrix} 1 \\ e^{-\frac{\tau}{\lambda}} \\ \frac{\tau}{\lambda} e^{-\frac{\tau}{\lambda}} \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}' \begin{bmatrix} f_0 \\ f_1 \\ f_2 \end{bmatrix} \tag{4.2}$$

With the to be estimated parameters in equation (4.2) fulfilling: $\beta_0, \beta_1, \beta_2, \lambda > 0$. The parameter $\frac{1}{\lambda}$ represents the rate of the exponential decline of the other parameter in the model. Given m observed yields with different maturities $T_1, \cdots, T_m$ we get m equations, where τ_m represent the time to maturity, which is $T_m - t$. In order to obtain these parameters we follow the subsequent strategy: fix λ and then estimate the βs using Least Square method. The parameter in the model are not assumed to be constant, but do change over time. However due to simplicity subscripts are neglected in the following.
Therefore in the Nelson-Siegel model the yield of a certain maturity can be written as the sum of different components: The model consists of a constant (f_0), which represents the interest rate level. The parameter β_0 is independent of time to maturity and can be interpreted as a constant interest rate for the various maturities. The exponential decay function, denoted by (f_1), reflects the second factor, a downward $(\beta_1 < 0)$ or upward $(\beta_1 > 0)$ slope. The parameter β_1 however is scaled with τ, resulting in dependence of the maturity. The larger the maturity, the smaller the influence of β_1 and vice-versa. Therefore β_1 influences on the short end of the curve. Furthermore a Laguerre function (f_2) of the form xe^{-x} is added to the parameter β_2. Then again β_2 is weighted with τ. The influence of β_2 works the other way round compared to β_1. The lager the maturity, the greater the influence of β_2 on the term-structure. Moreover β_2 adds the hump to the curve. As the parameter λ affects the weight function of β_1 as well as of β_2 it actually determines the position of the hump. The constraints of the parameters for equation (4.2) are set as follows:

$$\beta_0 > 0, \quad \beta_0 + \beta_1 > 0 \; and \quad \lambda > 0$$

In general the parameters can be estimated by minimising the difference between the modeled rates, denoted by: $f(\tau)^M$ and the observed rates, which are denoted by: $f(\tau)$. The optimisation problem can be stated as:

$$min_{\beta,\lambda} \sum \Big(f(\tau)^M - f(\tau) \Big)^2 \tag{4.3}$$

In order to gain the spot rates, see [Definition 2.9], the forward-rate function, see [Definition 2.10] has to be integrated and then dividing the result by the remaining time to maturity. This leads to following equation:

$$R(t,T) = R(\tau) = \frac{1}{\tau} \int_0^\tau f(u)du \tag{4.4}$$

We recall the relationship between the spot-rate $(R(t,T))$ and the short $(r(t))$ as: $r(t) = lim_{T \longmapsto t} R(t,T)$

Then again following equation (4.2) and (4.4) the spot-rates are written as:

$$R(\tau) = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}' \begin{bmatrix} 1 \\ \frac{\lambda(1-e^{-\frac{\tau}{\lambda}})}{\tau} \\ \frac{\lambda(1-e^{-\frac{\tau}{\lambda}})}{\tau} - e^{-\frac{\tau}{\lambda}} \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}' \begin{bmatrix} r_0 \\ r_1 \\ r_2 \end{bmatrix} \tag{4.5}$$

However the parameter $\beta_0, \beta_1, \beta_2$ have to be estimated from the market. The corresponding discount-function, which adjusts the data points is written as:

$$d(\tau) = e^{-\tau \cdot R(\tau)}$$

The discount-function represents the usual shapes of the function, which is either normal, inverse or with an extremum. Due to simplicity when it comes to the interpretation of the parameters, we take a look at the limes.

$$lim_{\tau \to \infty} R(\tau) = \beta_0 + lim_{\tau \to \infty} \left(\frac{\lambda\beta_1(1-e^{-\frac{\tau}{\lambda}})}{\tau} + \frac{\lambda_2\beta_2(1-e^{-\frac{\tau}{\lambda}})}{\tau} \right) = \beta_0$$

With maturity going to infinity the limes of the function is b_0, and can be interpreted as the long-term interest level. However looking at the maturity tending to zero, we get with the use of *the rule from l'Hospital*

$$lim_{\tau \to 0} R(\tau) = \beta_0 + lim_{\tau \to 0} \left(\frac{\lambda\beta_1(1-e^{-\frac{\tau}{\lambda}})}{\tau} + \frac{\lambda_2\beta_2(1-e^{-\frac{\tau}{\lambda}})}{\tau} - \beta_2 e^{-\frac{\tau}{\lambda}} \right) = \beta_0 + \beta_1$$

Therefore β_1 can be interpreted as the term, which determines the difference between the long-term and short-term interest level. The sum of β_0 and β_1 results in the short-rate.

4.2 Svensson method

The Svensson model is an extension from the just regarded Nelson-Siegel model see formula (4.2). It was introduced in 1994. Svensson added two parameter to the model with the purpose to increase flexibility and to improve the fit. It resulted in a convexity effect which improved the bond-rating for large maturities. To the function of the forward rates a second hump term is added, and can be written as:

$$f(\tau) = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}' \begin{bmatrix} 1 \\ e^{-\frac{\tau}{\lambda}} \\ \frac{\tau}{\lambda} e^{-\frac{\tau}{\lambda}} \\ e^{-\frac{\tau}{\lambda_2}} \frac{\tau}{\lambda_2} \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}' \begin{bmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \end{bmatrix} \tag{4.6}$$

The restriction for equation (4.6) are the same as for (4.2) with additional: $\lambda_2 > 0$.

The spot-rate function, which we obtain again by integrating the forward-rate function is now defined as:

$$R(\tau) = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}' \begin{bmatrix} 1 \\ \frac{\lambda(1-e^{-\frac{\tau}{\lambda}})}{\tau} \\ \frac{\lambda(1-e^{-\frac{\tau}{\lambda}})}{\tau} - e^{-\frac{\tau}{\lambda}} \\ \frac{\lambda(1-e^{-\frac{\tau}{\lambda_2}})}{\tau} - e^{-\frac{\tau}{\lambda_2}} \end{bmatrix} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix}' \begin{bmatrix} r_0 \\ r_1 \\ r_2 \\ r_3 \end{bmatrix} \tag{4.7}$$

The properties of the limes of τ of (4.7) are the same as in the Nelson-Siegel model see (4.5), and that is why the interpretation of each parameter does not change either.
However the to be estimated parameters are: $\beta_1, \beta_2, \beta_3, \beta_4, \lambda$ and λ_2. To gain the parameters one could apply the same methodology as in the Nelson-Siegel method, fix the λ, λ_2 values and then again use Least Square method to minimise the difference between the model and observed rates. However when applying this method due to collinearity problem many different parameter values give similarly-good fits.

4.3 Nelson-Siegel-Svensson parameter for latest observation

The following section explains the applied methodology in this thesis in order to gain the Nelson-Siegel-Svensson parameter, which will be used as input for the Kalman Filter procedure see [Chapter 6].
Even though the model of Nelson-Siegel-Svensson is widely used for modeling the yield curve, many authors reported *numerical difficulties* when it comes to calibrating the model. The problem is mentioned as twofold: on the one hand the optimisation problem is not convex and on the other hand it has a variety of local optima. Moreover as the Nelson-Siegel-Svensson model can also be interpreted as a factor model another problem occurs: collinearity. The β coefficients are the factor realisations. However the factor loadings are also the weight functions of these parameters. For the Nelson-Siegel-Svensson model the loadings for a maturity τ are determined through:

$$\left[1; \quad \frac{1-e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}}; \quad \frac{1-e^{-\frac{\tau}{\lambda}}}{\frac{\tau}{\lambda}} - e^{-\frac{\tau}{\lambda}}; \quad \frac{1-e^{-\frac{\tau}{\lambda_2}}}{\frac{\tau}{\lambda_2}} - e^{-\frac{\tau}{\lambda_2}}\right]'$$

By setting λ and λ_2 we impose fixed factor loadings on a specified maturity. Regarding m different maturities, we would have to solve m linear equations in order to estimate the parameters.

$$\begin{bmatrix} 1 & \frac{1-e^{-\frac{\tau_1}{\lambda}}}{\frac{\tau_1}{\lambda}} & \frac{1-e^{-\frac{\tau_1}{\lambda}}}{\frac{\tau_1}{\lambda}} - e^{-\frac{\tau_1}{\lambda}} & \frac{1-e^{-\frac{\tau_1}{\lambda_2}}}{\frac{\tau_1}{\lambda_2}} - e^{-\frac{\tau_1}{\lambda_2}} \\ & & \vdots & \\ 1 & \frac{1-e^{-\frac{\tau_m}{\lambda}}}{\frac{\tau_m}{\lambda}} & \frac{1-e^{-\frac{\tau_m}{\lambda}}}{\frac{\tau_m}{\lambda}} - e^{-\frac{\tau_m}{\lambda}} & \frac{1-e^{-\frac{\tau_m}{\lambda_2}}}{\frac{\tau_m}{\lambda_2}} - e^{-\frac{\tau_m}{\lambda_2}} \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = \begin{bmatrix} f(\tau_1) \\ \vdots \\ f(\tau_m) \end{bmatrix}$$

The system of equations is over identified, therefore the norm of the residuals would have to be minimised by applying Least Square method. Which results in many different parameter values giving a similar good fit.
That is why commonly used methods available in statistical packages, like deviation of the objective function, is not useful in order to obtain the desired parameter values. Therefore Gilli M, Große S. and Schumann E. came up with a different approach. They implemented an optimisation heuristic, to be more precise *Differential Evolution*. They also tested this methodology and were able to show that it is reliably solving the model.
The technique used is a stochastic technique for optimisation and is called *Differential Evolution* (DE) and was introduced by Storn and Prices in the 1990s. The goal is to solve the underlying optimisation problem, see formula (4.3) using *DE*. It is a useful way to find the global optimum of a real-valued function of real-valued parameters but with the function not forcedly being either continuous or differentiable. *DE* belongs to the class of genetic algorithm which use biology-inspired crossover operations, mutation, and even a selection of a population in order to minimize an objective function over various generations.
Since the invention of *DE* it has been applied in a variety of fields from computational physics to operation research. The *DE* method in general works in the following way: a problem is optimized by iteratively trying to improve a candidate solution with regard to a given measure of quality. *DE* uses floating point encoding of population members and arithmetic operations when it comes to mutation.
This method is also known as *metaheuristics* as none or only few assumptions are made and therefore the space of possible candidates for the solution is large. Furthermore it does not guarantee that an optimal solution is ever found. The problem is optimized by maintaining a population of candidate solutions and creating new candidates by combining the existing ones according to a specific scheme. It keeps whichever candidate has the best score on the problem. The procedure is repeated until a pre-specified termination criterion is satisfied.
As a first step the *true* yield curve, y is created with given parameters for the β values. The aim is to fit a smooth curve through these points. As the model is initially used to create these points, a perfect fit should be found. The objective function takes two arguments: param, which is a vector containing candidate solution, and a list data, which stores the other variables. As return we get the maximum absolute difference between a vector of the observed yields y and the model's yields y^M for the parameters from *param*. Then a model, which is here the Nelson-Siegel-Svensson is added to describe the mapping from parameters to a yield curve. Moreover possible constraints are added. One could also include a penalty function to determine whether parameters should met certain constraints. If a solution violates these constraints, a specified parameter controls how heavily it is penalised. The restriction for a valid solution is to have a penalty of zero. Afterwards the *DEOptim* function is called, with input as the objective function, the list data and also a list algo, which contains information about the population size, number of generations, step size, the probability of crossover, constraints and the penalty.
To be more precise *DEOptim* searches for minima of the specified objective function between upper and lower bounds on each of the to be optimized parameters. That is why vectors that compromise the lower and upper bounds have to be specified. And of course have

to be the same length as the parameter vector. The first argument in the function is the vector of the to be optimized function. The lower and upper bounds are specified as scalar for each parameter. The control argument stores intermediate populations from the first generations onwards. The term population refers to the set of parameter vectors, which is transformed at each generation. In order to generate an updated parameter vectors, *DE* disturbs the old one with a scaled difference of two randomly selected parameters vectors. However this pre-defined function relies on repeated evaluation of the objective function in order to move the population toward a global minimum.
Having a look at the output of the function, following elements have to be mentioned: a set of the best parameters found, the value of the function corresponding to the best values, the number of function evaluations and of the procedure iterations.

4.4 Bond price from the NSS-model

We recall that the short rate $r(t)$ from the underlying model under the real world measure follows the dynamics, see (3.27):

$$r(t) = x(t) + y(t) + \varphi(t)$$

Where we can write that two processes are solution to the SDEs, with two uncorrelated Brownian motions: W_x and W_y:

$$dx(t) = a\Big(d_x(t) - x(t)\Big)dt + \sigma dW_x(t) \ , \ x(0) = 0$$

$$dy(t) = b\Big(d_y(t) - y(t)\Big)dt + \eta\Big(\rho dW_x(t) + \sqrt{(1-\rho)^2}dW_y(t)\Big) \ , \ y(0) = 0$$

Moreover as we want to make sure, that initially the prices of zero coupon bonds of the model actually fit the obsvered bond prices from the market, which are induced through the initial term structure, the factor φ is added:

$$\varphi(t) = f^M(0,t) + \frac{\sigma^2}{2a^2}\Big(1 - e^{-at}\Big)^2 + \frac{\eta^2}{2b^2}\Big(1 - e^{-bt}\Big)^2 + \rho\frac{\sigma\eta}{ab}\big(1 - e^{-at}\big)\big(1 - e^{-bt}\big)$$

However the instantaneous forward rate, $f^M(0,t)$ results from the initial term structure, respectively the corresponding prices $P^M(0,t)$. With $P^M(0,t)$ denoting the pre-specified price at time 0 of the simulation for a zero coupon bond with maturity t, in other words the initial term structure of the model.
For the first years of the simulation is the Nelson-Siegel-Svensson method used in order to determine the initial term-structure. However after passing the first years, a flat term-structure is assumed. To be more precise after a maturity of $\hat{t}$ years a constant spot rate is applied. The spot rate $R(0,t)$, at time 0 with maturity t is modelled as follows:

$$R(0,t) = \begin{cases} \beta_0 + \beta_1(1 - e^{-\frac{t}{\tau_1}})\frac{\tau_1}{t} + \beta_2(1 - e^{-\frac{t}{\tau_1}})\frac{\tau_1}{t} + \beta_3((1 - e^{-\frac{t}{\tau_2}})\frac{\tau_2}{t} - e^{-\frac{t}{\tau_2}}) \ , \ if\, t \leq \hat{t} \\ \hat{R} \ , \ else \end{cases}$$

With the parameter $\beta_0, \beta_1, \beta_2, \beta_3, \lambda$ and λ_2 estimated as in the previous section described. We still need to determine the model bond-prices. In order to do so, we calculate them, by applying following formula:

$$P^M(0,t) = \Big(1 + R(0,t)\Big)^{-t}$$

5 Kalman Filter

The underlying chapter describes the Kalman Filter in general and we will then define the Two-Additive-Factor Gaussian model in Kalman Filter formulation. The results are used from (cf.[9, 18, 29, 30, 31, 32]).
The Kalman Filter was developed by Rudolph Kalman in 1960. The Kalman Filter is a mathematical technique commonly used in digital computers of control and navigation systems, avionics and outer-space vehicles with the purpose to extract a signal from a long sequence of noisy or incomplete measurements. When Kalman initially published his ideas he was faced scepticism. Therefore he was just given the opportunity to present his results in mechanical engineering instead of electrical or system engineering. However when he presented his ideas at the NASA Ames Research Center, they included the algorithm for the Apollo program, and then further in the NASA space shuttle, the Navy submarine and in aerospace vehicles and weapons.
For statistic and control theory the Kalman Filter is also known as a *linear quadratic estimation (LQE)*. The Kalman Filter consists of a set of mathematical equations providing an efficient recursive computational technique for optimally estimating the state of the unobservable processes. The technique is split into two phases: the prediction and the update phase. During the prediction phase the Filter produces estimates of the current state variables and their uncertainties. Followed by the measurement of the observed outcome, where these estimates are updated using a weighted average.
When it comes to estimating the term-structure of the underlying short-rate model, the Kalman Filter is applied under the real-world measure. In general the Kalman Filter is used to study the relationship between a series of possibly noisy observed measurements (here: yields) and the theoretical predictions of those measurements based on unobservable state variables. The state space model consists of the measurement and transition equations. Therefore as a first step one has to formulate the underlying model description into these equations. State space models allow to model an observed time series z_t as being explained by a vector of (possibly unobserved) state variables F_t, which are driven by a stochastic process. Moreover measurement errors are considered, in a way that an observable process is used to estimate the unobservable. The Filter uses previous data in order to predict the non-predictive process. The prediction is updated with the observed information becoming available at each time step. The updated version is then used to make a new prediction for the next time step. The observations are weighted through the variance of the measurement errors in a way that the biggest measurement errors are weighted the least.
All in all the Kalman Filter is a recursive algorithm of a set of equations which allows an estimator to be updated as soon as new information becomes available. The observation that we consider are zero-coupon yields with fixed maturities from one year up to 30 years. The state variables describe the factors. These are unobservable and require an initial set of values in order to start the recursive algorithm.

5.1 Kalman Filter specification

The Kalman Filter algorithm is an iterative procedure which consists of two stages: prediction and update. Before defining these stages in the next subsections we want to formulate the state space model. The state space model consists of the measurement and the transition equation. The measurement model describes the relationship between the state and the measurement at the current time step t as:

$$z_t = A_t + BF_t + \epsilon_t, \quad \text{where } \epsilon_t|\mathcal{F}_{t-1} \sim \mathcal{N}(0,\mathrm{H}) \tag{5.1}$$

Equation (5.1) relates an unobserved variable (F_t) to an observed variable z_t. In general there are K yields and two factors. Therefore A and z represent a vector of length K, B a $K \times 2$ matrix and F_t a vector with two entries, x_t and y_t. The term $BF_t + \epsilon_t$ is interpreted as the *disruption term*, ϵ denotes the measurement errors, which are normally distributed with mean zero and variance matrix H. The measurement error matrix H represents the accuracy of the observed yields in the market. Put in other words it describes how much noise there is in the market yields. Furthermore the measurement errors are restricted to be equal for all maturities in order to limit the number of estimated parameters to eight. Hence H is a diagonal with the time-independent vector of $(h, ..., h)$ of size $K \times 1$ on the diagonal. Let us now take a look at the other part of the state equation, which is the transition equation. The transition equation is based on a model that allows the unobserved variable to change over time:

$$F_t = C + \Phi F_{t-1} + \nu_t, \quad \text{with } \nu_t|\mathcal{F}_{t-1} \sim \mathcal{N}(0,\mathrm{Q}) \tag{5.2}$$

The term Φ represents a 2×2 matrix, and C a vector with two entries. The errors ν_t are zero mean normally distributed, with corresponding variance matrix defined in formula (5.9) and they are uncorrelated with ϵ_t.

5.1.1 Prediction Step

The measurement and transition equation are now being specified but *how works the algorithm actually?*. In the first step, is called prediction step, and this is where the Kalman Filter produces an optimal estimate of the state vector F_t based on the information till $t-1$, hence restricted on F_{t-1}. The optimal estimator for F_t is its conditional mean.The operator ˜ denotes the estimate of the variable. Hence we have following formulation for the predicted state estimate:

$$1.\ \tilde{F}_{t|t-1} = E_{t-1}[F_t] = C + \Phi\tilde{F}_{t-1}$$

The predicted state estimate results from the updated previous predicted state estimate. The corresponding predicted error covariance describes the error covariance that the Filter thinks the estimate error has. It is given as:

$$2.\ \tilde{P}_{t|t-1} = E_{t-1}[(F_t - \tilde{F}_{t|t-1})(F_t - \tilde{F}_{t|t-1})'] = \Phi\tilde{P}_{t-1}\Phi' + Q$$

5.1.2 Updating Step

In the next step, the updating step, new information of time t becomes available. In other words the yield of time t is observable. So that the predicted yield of t from $t-1$ can be compared with the actual occurred one. The measurement residual, also called the innovation represents the difference between the observed measurement and the estimated measurement as follows:

$$3.\ \xi_t = z_t - \tilde{z}_{t|t-1}, \quad \text{where } \tilde{z}_{t|t-1} = A + BF_{t|t-1}$$

The corresponding conditional variance of the prediction error is denoted by:

$$4.\ \tilde{V}_t = BP_{t|t-1}B' + H$$

With the definition of the prediction error the estimation of the state variables is updated. In a way that the so called *Kalman gain* K_t is added as a factor of the error to the initial estimate of $F_{t|t-1}$.

$$5.\ \tilde{F}_t = F_{t|t-1} + K_t\xi_t$$

Finally we come to determine the *Kalman gain*, which can be interpreted as the weight given to the additional information available at the time step between $t-1$ and t:

$$K_t = P_{t|t-1}B'V_t^{-1}$$

As the conditional expectation of the factors are updated, the conditional variance of the state system is also updated and is written as:

$$6.\ \tilde{P}_t = (I - K_tB)P_{t|t-1}$$

After repeating each of the six steps for every discrete time step in the data set, and under the assumption, that the measurement prediction errors are Gaussian, the log-likelihood function see [Definition 2.15] for the entire data set is returned and calculated as stated below:

$$l(z_1, ..., z_n; \theta) = -\frac{nK\ln(2\pi)}{2} - \frac{1}{2}\sum_{t=1}^{n}(\ln|\tilde{V}_t| + \xi_t'\tilde{V}_t^{-1}\xi_t) \tag{5.3}$$

The Kalman Filter is based on the assumption that the process and measurement models are linear and that the process and measurement noise are additive Gaussian. That is why the Kalman Filter provides optimal estimate only if the assumptions are satisfied.

5.2 Formulation of the equations for the underlying short-rate model

In the current section we want to formulate the Two-Additive-Factor Gaussian model as a state space model for the Kalman Filter. As a first step we recall that the relation between zero-coupon yields and bond prices can be written as, see (2.5):

$$z_t(\tau_K) = \frac{-\ln(P(t,T_k))}{T_k - t} \tag{5.4}$$

To improve readability the time indicator of $z(t,T)$we will from not on denote the time indicator t as a subscript. Furthermore as we denote the time to maturity as $\tau_K = T_K - t$. We recall formula (3.19) which state the price of the zero-coupon given as:

$$P(t,T) = \frac{P^M(0,T)}{P^M(0,t)} e^{\frac{1}{2}\left[V(t,T)-V(0,T)+V(0,t)\right] - \frac{1-e^{-a(T-t)}}{a}x(t) - \frac{1-e^{-b(T-t)}}{b}y(t)}$$

We now want to continue with the definition the measurement equation (5.1) for the underlying two-factor additive Gaussian model. Combining the measurement equation (5.1) with the formulation of (5.4) brings us to following definitions:

$$z_t = \begin{bmatrix} z_t(\tau_1) \\ . \\ . \\ . \\ z_t(\tau_K) \end{bmatrix}$$

$$A_t = \begin{bmatrix} -\frac{\ln \frac{P^M(0,T_1)}{P^M(0,t)}}{\tau_1} + \frac{-\frac{1}{2}(V(t,T_1)+V(0,T_1)-V(0,t))}{\tau_1} \\ . \\ . \\ . \\ -\frac{\ln \frac{P^M(0,T_K)}{P^M(0,t)}}{\tau_K} + \frac{-\frac{1}{2}(V(t,T_K)+V(0,T_K)-V(0,t))}{\tau_K} \end{bmatrix} \tag{5.5}$$

$$B = \begin{bmatrix} \frac{B'(t,T_1)}{\tau_1} \\ . \\ . \\ . \\ \frac{B'(t,T_K)}{\tau_K} \end{bmatrix} = \begin{bmatrix} \frac{B_x(t,T_1)}{\tau_1} & \frac{B_y(t,T_1)}{\tau_1} \\ . & . \\ . & . \\ . & . \\ \frac{B_x(t,T_K)}{\tau_K} & \frac{B_y(t,T_K)}{\tau_K} \end{bmatrix} = \begin{bmatrix} \frac{1-e^{-a\tau_1}}{a\tau_1} & \frac{1-e^{-b\tau_1}}{b\tau_1} \\ . & . \\ . & . \\ . & . \\ \frac{1-e^{-a\tau_K}}{a\tau_K} & \frac{1-e^{-b\tau_K}}{b\tau_K} \end{bmatrix} \tag{5.6}$$

$$F_t \equiv \begin{bmatrix} x_t \\ y_t \end{bmatrix}$$

By regarding once again the measurement equation see formula (5.1), the vector A, which is defined as in equation formula (5.5) can actually be decomposed into a time independent and a non-stochastic time-dependent part. The time dependent part arises from the fact that the variance $V(t,T_k)$ in $A(t,T_k)$ only depends on the difference $T_k - t$ and in dealing with fixed maturities, this is a vector of constants. The time-depend part reveals because we are dealing with a no-arbitrage model. The Variance from equation (5.5) follows definition of (3.15).
The measurement equation for the underlying Gaussian Two-Factor model is formulated as follows:

$$\begin{bmatrix} z_t(\tau_1) \\ . \\ . \\ . \\ z_t(\tau_K) \end{bmatrix} = \begin{bmatrix} -\frac{\ln \frac{P^M(0,T_1)}{P^M(0,t)} - \frac{1}{2}(V(t,T_1)+V(0,T_1)-V(0,t))}{\tau_1} + \\ . \\ . \\ . \\ -\frac{\ln \frac{P^M(0,T_K)}{P^M(0,t)} - \frac{1}{2}(V(t,T_K)+V(0,T_K)-V(0,t))}{\tau_K} \end{bmatrix} + \tag{5.7}$$

$$\begin{bmatrix} \frac{1-e^{-a\tau_1}}{a\tau_1} & \frac{1-e^{-b\tau_1}}{b\tau_1} \\ & . \\ & . \\ & . \\ \frac{1-e^{-a\tau_K}}{a\tau_K} & \frac{1-e^{-b\tau_K}}{b\tau_K} \end{bmatrix} \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} + \begin{bmatrix} \epsilon_1 \\ . \\ . \\ . \\ \epsilon_K \end{bmatrix}$$

In the next step we want to define the transition equation (5.2) for the underlying short rate model. We recall formula (5.2):

$$F_t = C + \Phi F_{t-1} + \nu_t, \quad \text{with } \nu_t|\mathcal{F}_{t-1} \sim \mathcal{N}(0,\mathrm{Q}) \tag{5.8}$$

The errors ν_t are zero mean normally distributed, with corresponding variance matrix defined in formula (5.9) and they are uncorrelated with ϵ_t.
In order to determine C, Φ and Q we have to take a look at the model dynamics specified under the real-world measure, $\mathbb{P}$. We recall formula (3.24) which states:

$$dF_t = \Lambda(F_t - \mu)dt + \sum S_t dW_t^{\mathbb{P}}$$

As this process belongs to the class of *Ornstein-Uhlenbeck processes*, we recall that we have already seen in [Definition 2.12], how to obtain a solution. Therefore we directly write the solution, which is:

$$F_{t+\Delta t}|\mathcal{F}_t = e^{\Lambda\Delta t}F_t + \int_t^{t+\Delta t} e^{\Lambda\Delta t}(-\Lambda\mu)ds + \int_t^{t+\Delta t} e^{\Lambda\Delta t}\sum S_t dW_s$$

The expectation for the factors at $t + \Delta t$ given $\mathcal{F}_t$ is defined through:

$$\mathbb{E}_t[\mathcal{F}_{t+\Delta t}] = e^{\Lambda \Delta t} F_t + \int_t^{t+\Delta t} e^{\Lambda \Delta t}(-\Lambda \mu) ds) =$$

$$\underbrace{\begin{bmatrix} e^{-a\Delta t} & 0 \\ 0 & e^{-b\Delta t} \end{bmatrix}}_{\Phi} F_t + \underbrace{\begin{bmatrix} \frac{\psi_1 \sigma^2}{a}(1 - e^{-a\Delta t}) \\ \frac{\psi_1 \sigma \eta \rho + \psi_2 \eta^2 \sqrt{(1-\rho^2)}}{b}(1 - e^{-b\Delta t}) \end{bmatrix}}_{C}$$

Hence we can calculate Matrix Q from the transition equation as follows:

$$Var_t(F_t) = \mathbb{E}_t[F_{t+\Delta t} F'_{t+\Delta t}] = \int_t^{t+\Delta t} e^{\Lambda(t+\Delta t-s)} \sum S_t (\sum S_t)' (e^{\Lambda(t+\Delta t-s)})' dW_s^{\mathbb{P}}$$

$$= \underbrace{\begin{bmatrix} \frac{\sigma^2}{2a}(1 - e^{-2a\Delta t}) & \frac{\rho\sigma\eta}{a+b}(1 - e^{-(a+b)\Delta t}) \\ \frac{\rho\sigma\eta}{a+b}(1 - e^{-(a+b)\Delta t}) & \frac{\eta^2}{2b}(1 - e^{-2b\Delta t}) \end{bmatrix}}_{Q} \tag{5.9}$$

The state space model is fully specified but before starting the iterative procedure, the initial values have to be determined:

$$F_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The initial value of the variance is defined through the variance of the factors:

$$Q_0 = \begin{bmatrix} \frac{\sigma^2}{2a} & \frac{\rho\sigma\eta}{a+b} \\ \frac{\rho\sigma\eta}{a+b} & \frac{\eta^2}{2b} \end{bmatrix}$$

With having now the state space model and the initial values declared, we can insert the above gained information about the specifications in each step of the Kalman Filter, which is described in the following subsections.

6 Modeling of the future bond prices

Since the previous chapters cover the basic theory for this thesis, the underlying chapter focusses on the application of the theory and describes the methodology to estimate a set of parameters for the Two-Additive-Factor Gaussian model. With which we further model the future zero-coupon prices. The methodology is implemented in the statistic-software R, which is a free software environment for statistical computing and graphics.
As already mentioned in [Chapter 5] the Kalman Filter is a recursive algorithm, therefore an initial set of parameters has to be specified. However these parameters can be randomly chosen, they just have to full fill a certain criterion of the optimizer function, which is explained in [Subsection 6.2.1].
To start with the specification of the setting: we do 10.000 path simulations, choose a maximum maturity of 30 years, the widest simulation horizon represents 40 years, and the time-delta between the maturities, as well as between the simulation horizons is one year, the observed maturities of the short-rate is also up to 30 years. However as we have daily-observed short-rates the observed time-delta is $1/254$.
For the Kalman Filter we have to specify the initial values x_0, y_0, furthermore the drifts, volatilities, correlation, and market prices of risk. We chose them to have following values:

$$\begin{bmatrix} x_0 \\ y_0 \\ a \\ b \\ \sigma \\ \eta \\ \rho \\ \psi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0.05 \\ 0.05 \\ 0.0101 \\ 0.0101 \\ -0.5 \\ 0.0009 \\ 0.0009 \end{bmatrix}$$

Despite the specification of the criterion for the optimizer this guideline for the initial set of parameters is advisable: $a, b \in [0, 2]$, the mean reversion should not be grater than 1, hence $\sigma, \eta \in [0, 1]$, the correlation should be in-between $[-1, 1]$ and the measurement error should not exceed 150 basis points, and is chosen to be 10^{-3}.
We then start the routine by loading the observed interest rates data, which is daily observed short-rates in the time period between 01.01.2015 and 31.12.2019. As week-ends and holidays are excluded we get a total of 1272 observations.
As a next step we just filter on the latest observation date and interpolate the corresponding spot curve. This is needed as we only have observed maturities for 30 years, but we want to have a predicted model with maturities of 70 years, which is the widest horizon we regard. The subsequent section describes the applied procedure.

6.1 Interpolation of the spot curve

The underlying section uses results from [cf.[13, 33, 34, 35, 36, 37]]. In a mathematical sense *interpolation* is referred to a kind of estimation, which constructs or finds new data points based on a set of given data points. We distinguish piece-wise, linear, polynomial spline and mimetic interpolation. As natural cubic splines belongs to the class of piece-wise interpolation, we will have a closer look at this method. Piece-wise interpolation fits a large number of data points with low-degree polynomials. It works in a way that, given a set of data points, which is in this case short-rates with different maturities, a different polynomial for each interval is used to interpolate several interpolants at successive points. Therefore the terms such as knots, breakpoints or control points are derived, as these are abscissas at which the interpolant changes from one polynomial to another. The goal is to find an interpolation function that is smooth and does not change too much between the node points. The natural cubic spline is defined as a piece-wise cubic polynomial that is twice continuously differentiable.
For the interpolation of the spot curve, the *lm function* from the R-package *stats*[1] *(version 3.6.2)* with a cubic spline proceeding is applied. The main purpose of the *lm* function is to fit linear models. Linear models are used to predict the value of an unknown variable, with the basis being independent variables. Therefore it should find the relationship between the given variables and a forecast. The function *lm* is usually applied when it comes to carry out regression, single stratum analysis of variance or analysis of covariance to predict the missing values of the underlying data. A typical model hast the form *response* $\sim$ *terms*, with the first argument being a (numeric) response vector and the other specifies a linear predictor for the *response*. The response term in this case is the short-rate, with the linear predictor being specified as a natural spline procedure. The function *natural spline (ns)* is from the package *Epi*[2] *(version 2.46)* and generates a B-spline basis matrix for a natural cubic spline. The *ns* function consists of following arguments: a predictor variable, which is in our case time, the degrees of freedom, which is 30 in our case. Moreover one could specify knots, which represent the breakpoints that define the spline, however we have not defined any knots. Also one could determine an intercept included in the basis, but we have not chosen to do so. Furthermore there is the possibility to set boundary knots. These are boundary points at which to impose the natural boundary conditions and anchor the B-spline bases.
Afterwards the *predict* function is applied on the data within the interpolated model. The function evaluates a pre-defined spline basis, given the input values. The *predict* function takes two arguments as input: the object, which is in our case the interpolated spot model and the new values at which evaluations are required, which is a data frame with maturities from 1 to 70 years. After applying the function on the short-rates, we have expanded the observation period of with corresponding short-rates from the initial 30 years to 70 years of maturities.
Then before we continue with the estimation procedure, the column *bond price* is added to the underlying data frame. Applying following formula on the spot rates returns, we get

[1]The package can be found, following: https://rdocumentation.org/packages/stats/versions/3.6.2
[2]The package can be found, following https://rdocumentation.org/packages/Epi/versions/2.46

an addition column with the corresponding bond prices:

$$P(0,t) = \Big(1 + r(t)\Big)^{-t} \tag{6.1}$$

So far we have a data frame containing short-rates with maturities from 1 to 70 years and corresponding bond prices.

6.2 Estimation process

When it comes to the estimation procedure, we continue with the calculation of the Nelson-Siegel-Svensson parameters for the latest yield. We recall that we obtain the latest yield by applying following formula on the obtain bond-prices, see formula (6.1):

$$z_t(\tau_K) = -\frac{\ln\Big(P_{obs}(t,T)\Big)}{\tau_K}$$

With the latest yield we can continue to calculate the Nelson-Siegel-Svensson parameter: $\beta_1, \beta_2, \beta_3, \beta_4, \tau_1, \tau_2$, see (4.7). In order to obtain the NSS parameters an optimisation heuristic method, *Differential Evolution* is implemented. However parameter identification is only possible if specific parameters are restricted to certain ranges. The exact procedure is explained in [Chapter 4].
We then start with the procedure which returns the log-likelihood function, see (5.3). The procedure has following input parameters: the initially defined model parameters, the maturities, the observed yields, the observed time steps and the Nelson-Siegel-Svensson parameters.
In a first step the variance of the function is calculated, following formula (3.15):

$$V(t,T) = \frac{\sigma^2}{a^2}\left[T - t + \frac{2}{a}e^{-a(T-t)} - \frac{1}{2a}e^{-2a(T-t)} - \frac{3}{2a}\right] + \frac{\eta^2}{b^2}\left[T - t + \frac{2}{b}e^{-b(T-t)} - \frac{1}{2b}e^{-2b(T-t)} - \frac{3}{2b}\right] +$$
$$2\rho\frac{\sigma\eta}{ab}\left[T - t + \frac{e^{-a(T-t)} - 1}{a} + \frac{e^{-b(T-t)} - 1}{b} - \frac{e^{-(a+b)(T-t)} - 1}{a+b}\right]$$

As a next step the model bond prices are calculated using the Nelson-Siegel-Svensson parameters. The corresponding modeled short-rate with maturity t is calculated, following equation (4.7):

$$R(0,t) = \beta_0 + \beta_1(1 - e^{-\frac{t}{\tau_1}})\frac{\tau_1}{t} + \beta_2(1 - e^{-\frac{t}{\tau_1}})\frac{\tau_1}{t} + \beta_3((1 - e^{-\frac{t}{\tau_2}})\frac{\tau_2}{t} - e^{-\frac{t}{\tau_2}}) \tag{6.2}$$

Then again with the knowledge of the modeled short-rate, the bond prices for the model, which are denoted by: $P^M(0,t)$ are calculated, applying formula (6.1) on the modeled spot-rates from (6.2):

$$P^M(0,t) = \Big(1 + R(0,t)\Big)^{-t}$$

To continue with the application of the Kalman Filter, we recall that we need to formulate the state equation, which consists of the measurement and transition equation. With the measurement equation once again specified as:

$$z_t = A_t + BF_t + \epsilon_t$$

With the equation for the short-rate model specified:

$$z_t(\tau_K) = \frac{-\ln(P(t,T_k))}{T_k - t} = \frac{-\ln(\mathcal{A}(t,T_k)) + B_x(t,T_k)x(t) + B_y(t,T_k)y(t)}{T_k - t}$$

The other part of the state equation, the transition equation is once again formulated:

$$F_t = C + \Phi F_{t-1} + \nu_t$$

Afterwards we switch to the model dynamics under $\mathbb{P}$ in order to determine C, Φ and Q:

$$dF_t = \lambda(F_t - \mu)dt + \sum S_t dW_t^{\mathbb{P}}$$

With having now all input parameter specified, we call the function *fkf*[3] from the same labelled package, which routine is described in [Chapter 5]. The output is the log-likelihood function of the set of parameters from the Kalman Filter, which is defined in (5.3).
As a next step, the yield from the observed bond prices are calculated by applying following formula on the observed bond prices:

$$z(t,T) = -\left(\frac{\ln P(t,T)}{(T-t)}\right)$$

The observed yields are needed as input for the optimizer for the estimated model parameters, which is described in the subsequent subsection.

6.2.1 Optimizier

The underlying section uses results from [cf.[38, 39, 40]]. In the following we regard the numerical approach for the MLE for the set of parameters. For this the R-Function *constrOptim* from the packages "stats" (version 3.6.2) is applied. It provides a method to minimise a function subject to linear inequality constraints with the use of an adaptive barrier algorithm. The here applied method is called *Nelder-Mead method*, which is also known as a downhill simplex. It is a numerical method used to find the minimum or maximum of an objective function in a multidimensional space. The method was introduced by John Nelder and Roger Mead in 1965. The algorithm is a deterministic search strategy relying on function evaluations only. Moreover it can be regarded as an optimization method for non-linear problems using a simplex of $n + 1$. The aim of the algorithm is to find a minimum by evaluating the objective function and moving the simplex until just the minimum is found. Doing so it rescales the simplex based on local behaviour of the function while three basic procedures are applied: reflection, expansion and contraction.

[3]The package can be found, following: https://rdocumentation.org/packages/FKF/versions/0.2.4

The *ConstrOptim* function takes following inputs: numeric starting values, which is in this case the initial set of parameters, the to be minimised function, which is the negative log-likelihood function from the estimation step, see (5.3), a constraint matrix u_i, a constraint vector c_i of row length from u_i, the iterations of the barrier algorithm is set to 200 and the relative convergence tolerance is set to: 10^{-6}. The starting vector, which contains the initial set of parameters has to full-fill following condition:

$$ui * \theta - ci \geq 0$$

Where $*$ represent a matrix multiplication. Furthermore θ represents a vector containing the initial set of parameters.
The output of this procedure is the set of parameters for the Two-Additive-Factor Gaussian model. In particular we obtain the beneath stated values for the parameters:

$$\begin{bmatrix} x_0 \\ y_0 \\ a \\ b \\ \sigma \\ \eta \\ \rho \\ \psi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0.9136051683 \\ 0.9989999468 \\ 0.0101236500 \\ 0.0126402162 \\ -0.9988806370 \\ 0.0009724382 \\ 0.0009983962 \end{bmatrix} \tag{6.3}$$

With the modeled parameters we can finally start the path simulation of the future zero-coupon bond prices, which is explained in the next section.

6.3 Simulation

With the estimated set of parameters we want to continue to create possible paths of the future development. As already mentioned at the beginning of this chapter, we chose to do a simulation of 10.000 paths, and the mean of these is used to calculated the future zero-coupon bond prices.
We want to mention again, that we regard two correlated Ornstein-Uhlenbeck processes which are given as:

$$dx(t) = \big(\psi_1\sigma^2 - ax(t)\big)dt + \sigma dW_1(t)$$

$$dy(t) = \big(\psi_1\sigma\eta\rho - \psi_2\eta^2\sqrt{1-\rho^2}\big)dt + \eta\rho dW_1(t) + \eta\sqrt{1-\rho^2}dW_2(t)$$

We recall that the solution of an Ornstein-Uhlenbeck process is of the form, see equation (2.4):

$$X_t = X_0 e^{-\theta t} + \mu(1 - e^{-\theta t}) + \sigma \int_0^t e^{-\theta(t-s)} dW_s$$

Which leaves us to determine the parameters for the underlying model. Which brings us to following solutions of the processes:

$$x_{t+1} = x_t e^{-a(t+1)} + \frac{\psi_1\sigma^2}{a}(1 - e^{-a(t+1)}) + \sigma \int_t^{t+1} e^{-a(t+1-s)} dW_1(s)$$

$$y_{t+1} = y_t e^{-b(t+1)} + \frac{\psi_1 \sigma\eta\rho - \psi_2 \eta^2 \sqrt{1-\rho^2}}{b}(1 - e^{-a(t+1)}) + \eta\rho \int_t^{t+1} e^{-b(t+1-s)} dW_1(s)$$
$$+\eta\sqrt{1-\rho^2} \int_t^{t+1} e^{-b(t+1-s)} dW_2(s)$$

Following the above stated formulas we get 10.000 paths for each of these processes at each time-step.

6.4 Calculate zero-coupon bond prices

With now having the simulated paths of the processes we calculate the mean of these bond prices for every time instant and every maturity. We follow the zero-coupon bond price formula, see (3.19):

$$P(t,T) = \frac{P(0,T)}{P(0,t)} e^{\frac{1}{2}(V(t,T)-V(0,T)+V(0,t)-\frac{1-e^{-a(T-t)}}{a}x(t)-\frac{1-e^{-b(T-t)}}{b}y(t)}$$

Where $V(.,.)$ represents the variance, which is determined in equation (3.15). The zero-coupon bond prices $P(0,T)$ and $P(0,t)$ are determined through the relationship between zero-coupon-bonds prices and the short-rate. Which is given by:

$$P(0,t) = \Big(1 + r(t)\Big)^{-t} \tag{6.4}$$

The above stated formula (6.4) is applied on the initial observed short-rate data from the latest observation date. The factor a and b are the estimated model parameters. With having every part of the zero-coupon bond price formula determined, the mean at each time of these prices is calculated. The output is a matrix, containing the mean zero-coupon bond prices with maturities from 1 year up to 30 years and a simulation horizon up to 40 years. These zero-coupon bond prices can then again be transformed to the corresponding short-rates, see (6.4). To model this short-rate was the goal of this thesis. That is why with this output we have reached it. However it is still left to validate the output, which is the focus of the subsequent chapter.

7 Parameter Valuation

So far we have the estimated set of model parameters and with those estimated the future zero-coupon bond prices respectively the corresponding short-rates. But is our output valid? There are various ways to check, whether the model is stable and calibrates good results. In the underlying chapter we use different approaches to validate the model and the output.

7.1 Martingale-Test

A simple and frequently used technique is the Martingale-test. The Martingale-test proves if the underlying modeled bond-prices are a martingale see [Definition 2.3]. To be more precise this means that one unit of currency invested at time t_0 at the end of maturity t_n and being discounted, should match in average one unit of the currency again. Actually discounted at every time instant t_i it should match the one unit of currency. The underlying plot [Figure 7.1] represents the output of the martingale test for the simulated bond prices.

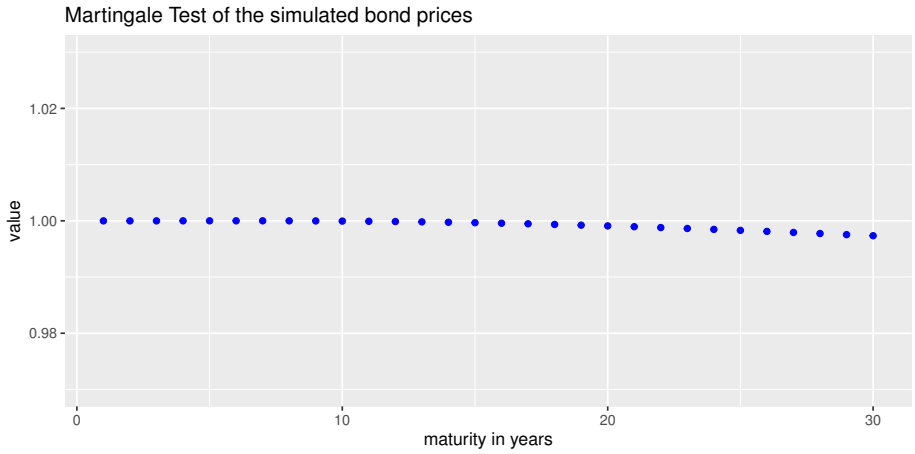


Figure 7.1: Martingale Test of the simulated bond prices

We can identify that the for the first ten years, we meet the martingale property quite perfectly. However as maturity grows, we can recognize a falling tendency. However it is important to pay a close look at the scale, as the lines of y-axis are represented with step-size of 0.01. The discounted prices stay in an interval of [0.9999998; 0.99973412]. That is why we can say that we fit the Martingale-test quiet well, as the requirement is to stay close to the value of 1. The deviations of the Martingale.test stay in between of a 1 percent range of the expected value of 1. Therefore the outcome of the modeled short-rates is satisfying.

This makes us suggest that with the estimated set of parameters, the underlying short-rate model fulfils the martingale property for the modeled short-rate.

7.2 Simulation with the estimated parameters

7.2.1 Parameter comparing

For an admissible model we expect that the model is able to reproduce its set of parameters. To be more precise with the gained set of parameters we have already simulated the future short-rates. This time the observation period, which represents the time delta, represents one year, the maturities we regard are from 1 to 30 years. The initial set of parameters for the Kalman Filter, is the obtained set of parameters from the observed short-rates, see equation (6.3). With those parameters we rerun the Kalman Filter algorithm. Then again the *Nelder-Mead* optimizer is applied and the output is the following set of parameters.

$$\begin{bmatrix} a \\ b \\ \sigma \\ \eta \\ \rho \\ \psi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} 0.94050491442 \\ 0.99899980290 \\ 0.01620579552 \\ 0.01618853236 \\ -0.99899155700 \\ 0.00099996790 \\ 0.00009498986 \end{bmatrix} \tag{7.1}$$

By comparing these two sets of parameters ((6.3) and (7.1)) we assert that the absolute difference between those sets is small. The behaviour of the terms is pretty similar. The volatility terms, σ and η are not exactly the same, however those are also influenced through the random behaviour of the Brownian motion. Therefore it can be said that in sum the parameters are quite good replicated. Which makes us suggest that the implementation of the Kalman Filter, as well as of the path simulation is right. However to be sure we repeat the methodology and estimate on basis of the parameter of (7.1) the zero-coupon bond prices, respective the future short-rates again. On basis of this short-rates we rerun the Kalman Filter algorithm with this set of parameters again. The output this time is:

$$\begin{bmatrix} a \\ b \\ \sigma \\ \eta \\ \rho \\ \psi_1 \\ \psi_2 \end{bmatrix} = \begin{bmatrix} 0.94050490000 \\ 0.99899980000 \\ 0.01622942747 \\ 0.01650635754 \\ -0.998991600 \\ 0.00099996790 \\ 0.00009498986 \end{bmatrix}$$

By comparing the three sets of parameters we come to the conclusion that model is able to reproduce its parameters, as we expect it to do. Therefore we conclude that the procedure is implemented correctly.

7.3 Model bond prices compared to the observed bond prices

The last approach to validate the model is as follows: The observed interest-rates data is split into two sets. The first set contains the first four years of observation, and the other the last year of observation. We then take the first set and start with the parameters of (6.3) a path simulation of the zero-coupon bond prices for the next year with maturities from 1 to 30 years. We then compare the actual observed bond prices from the latest observation, with the simulated one. The output is represented in the above given plot, see [Figure 7.2].

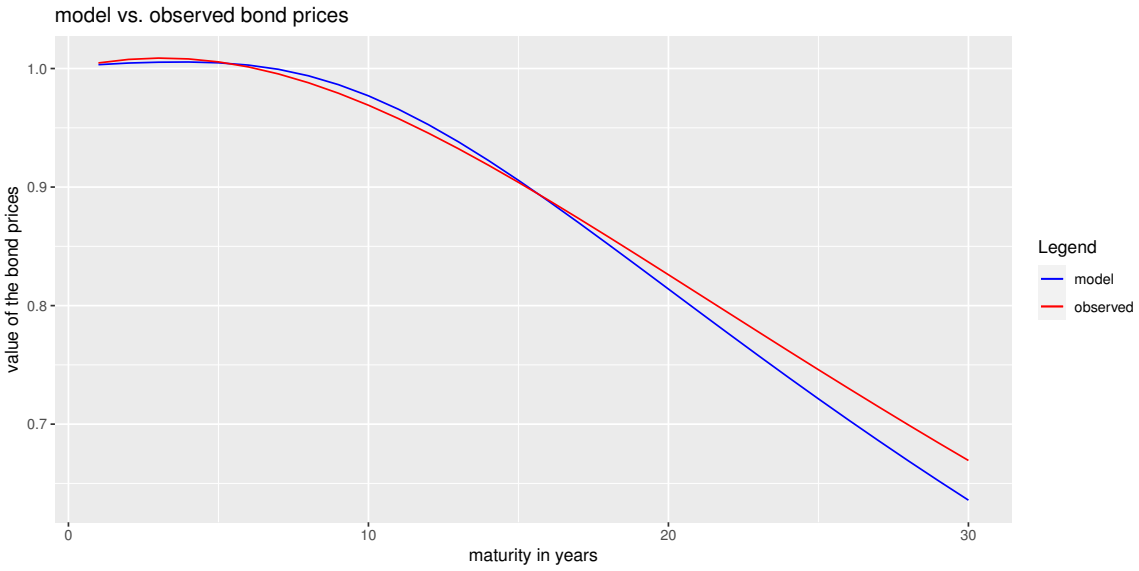


Figure 7.2: Model Bond Price and Observed Bond Price

As we can see in the above given plot, see [Figure 7.2], the bond prices match quite well for the first 17 years. Then however the modeled prices have a more downward tendency than the observed ones. Regarding the difference for a maturity of 30 years, it has the value of about 0.033. The gap for the longer maturities could be explained, through the fact that for the underlying short-rates we regard four years of observed short-rates, however the simulation here is for maturities to 30 years. That means for the first years we have a quite good fit, however as the maturities grow, the observed data for the underlying model should as well be extended. To sum up we can say that the model fits short-term maturities well, however as the maturities grow the more the values spread. Even tough the fit is not perfect, we can say that all in all we have a quiet accurate fit with the model.

8 Conclusion

We successfully demonstrated to calibrate a short-rate model. With the Brownian motion, see [Definition 2.1] and the stochastic differential equation, see [Definition 2.5] we have the basis tools for the short-rate model. Through the Ornstein-Uhlenbeck process, see [Definition 2.12] we have a solution of the process, which will be used for the path simulation for the future zero-coupon bond prices. With the definition of the zero-coupon bond in [Section 2.2.2] we are able to identify the corresponding yield. Through the representation of the different kinds of interest rates we are able to show their interaction.
When focussing on the underlying short-rate model we first take a look at the One-Factor Hull-White model and the problems, which may arise from depending on just one factor. This brings us to the conclusion to extend the model to another factor. The Two-Factor Hull-White model can equivalently be formulated as a Gaussian-Additive Two-Factor model, which is represented as the sum of two correlated Gaussian factors plus a deterministic function. With Girsanov's theorem the model formulation under the risk-neutral measure can be transformed to the real-world measure, which we will need for the Kalman Filter. The deterministic function is added to fit the initial term-structure and in order to define this function we need the Nelson-Siegel-Svensson technique. To be specific we need the Nelson-Siegel-Svensson parameters which are estimated through the latest observation of the observed short-rates. With these parameters the corresponding modeled zero-coupon bond-price can be calculated. In order to receive the set of Nelson-Siegel-Svensson parameter the optimisation heuristic *Differential Evolution* is applied.
The Kalman Filter is recursive algorithm and is used to obtain the set of model parameters for the Two-Factor Gaussian-Additive model. After formulating the state-space equation for the underlying model, we focus on the methodology: with an initial set of parameters, the equations for the underlying model and the estimated Nelson-Siegel-Svensson parameters the Kalman Filter algorithm can start. The output is the log-likelihood function of the parameters. Through the Nelder-Mead optimizer the Kalman Filter algorithm reruns until a certain threshold is reached.
With the basis theory for this thesis covered, we describe the process it takes in order to obtain the model parameters and then estimate the paths of the future zero-coupon bond prices. Those zero-coupon bond prices can also be transformed into the short-rate. To continue with the interpolation of the short-rates in order to enlarger the 30 observations to 70, as the simulation horizon corresponds to 40 years. We continue with the estimation process, which first identifies the Nelson-Siegel-Svensson parameters, and the corresponding model prices. We then start the Kalman Filter with the output being the optimized log-likelihood function of the to be estimated parameters. With the just estimated parameters we can start the simulation of the paths, of which we calculate the future mean bond-prices.
The goal of this thesis is to find a reasonable set of parameters for the short-rate model. We

wanted to confirm the validity of the output using the estimated set of parameters. For the validation of the output we used different approaches. Fulfilling the martingale property is evaluated using the Martingale-test. It states that the average future zero-coupon bond price discounted at each time point should match one unit of currency. As the discounted price range stays within one percent of one unit, we conclude that the Martingale property is fulfilled. As a second approach we want to recalibrate the set of model parameters. To do so we transform the obtained zero-coupon bond prices into short-rates. These short-rates are handled as the observed-short rates and the obtained set of model parameters is chosen to be the initial set of parameters for the Kalman Filter algorithm. The output of the model parameters of the second procedure fits quite well the output from the first procedure. We also tried to rerun the routine again: Using the data from the second procedure as input parameters we simulated paths and calculated the mean of future zero-coupon prices, which then again are transformed into short-rates and handled as the observed short-rates. This time the initial set of parameters for the Kalman Filter is the second set of parameters. The output matches again quite well with the other two sets, which makes us suggest that the implementation is correct, as the model behaves as expected. Last but not least the observed short-rates are divided into two data-sets so that one contains the observations of the first four years, and the other set of the last year. Then we take the first set as the short-rate basis and the estimated set of parameters as input for the path simulation. Then again we calculate the mean of this paths into zero-coupon bond prices. By comparing the modeled set with the actual observed one we see quite a perfect fit for the first 17 years of maturities. Until the end of maturity, which is year 30, the gap widens and the modeled short-rates are lower than the observed ones. We still regard the test as passed, as the first years fit quiet well and the observation period was short. In a next step one could widen the range of observed years to improve the fit, but this is beyond of the scope of this thesis. In summary we can say that we found a way to estimate a set of model parameter for the Gaussian-Additive Two Factor model. With those parameters future zero-coupon bond prices and the short-rate is modeled successfully.

Bibliography

[1] Behrends, Ehrard: *Markovprozesse und stochastische Differentialgleichungen: Vom Zufallsspaziergang zur Black-Scholes-Formel*. 2013. https://link.springer.com/content/pdf/10.1007/978-3-658-00988-5.pdf, visited on 13.01.2023.

[2] Wikipedia-contributors: *Brownian motion*. Wikipedia, The Free Encyclopedia, 2022. https://en.wikipedia.org/w/index.php?title=Brownian_motion&oldid=1098659132, visited on 25.07.2022.

[3] Schmock, Uwe: *Stochastic analysis for financial and acturial mathematics*. TU WIEN, 2020.

[4] Björk, Tomas: *Arbitrage Theory in Continuous Time*. Oxford University Press Inc., New York, third edition, 2009.

[5] Rheinländer, Thorsten and Sexton, Jenny: *Hedging derivatives*. World Scientific Publishing Co. Pte. Ltd, 2011.

[6] Brigo, Damiano and Mercurio, Fabio: *Interest Rate Models - Theory and Practice*. Springer Berlin, second edition (2006) edition, 2001, ISBN 3540417729.

[7] Whiteside, Benjamin: *Quadratic variation*, 2017. https://benjaminwhiteside.com/2017/01/26/quadratic-variation/, visited on 16.02.2023.

[8] Wikipedia-contributors: *Geometric brownian motion*, 2023. https://en.wikipedia.org/wiki/Geometric_Brownian_motion, visited on 26.02.2023.

[9] Hendriks: *Calibration and estimation of a correlated multi-factor term structure model*. Tilburg University, 2015.

[10] Korn, Ralf: *Moderne Finanzmathematik- Theorie und praktische Anwendung*. Springer, 2014.

[11] Wikipedia-contributors: *Ornstein-uhlenbeck process*, 2022. https://en.wikipedia.org/wiki/Ornstein\OT1\textendashUhlenbeck_process, visited on 03.07.2022.

[12] Gurker, Werner: *Introduction to Mathematical Statistics using R*. TU Wien, 2018.

[13] Filzmoser, Peter: *Advanced methods for regression and classification*. TU WIEN, 2022. https://tuwel.tuwien.ac.at/mod/resource/view.php?id=1727170, visited on 29.10.2022.

[14] Wikipedia-contributors: *Spline*. Wikipedia, The Free Encyclopedia, 2022. https://de.wikipedia.org/w/index.php?title=Spline&oldid=227186133#B-Splines, visited on 02.01.2023.

[15] Wikipedia-contributors: *B-spline*. Wikipedia, The Free Ecyclopedia, 2022. https://en.wikipedia.org/w/index.php?title=B-spline&oldid=1102440832, visited on 03.11.2022.

[16] Blanchard, Arnaud: *The two-factor hull-white model: Pricing and calibration of interest rates derivatives*. KTH Vetenskap Och Konst.

[17] Filipovic, Damir: *Term-Structure Models*. Springer Berlin, first edition, 2009.

[18] Ibhagui, Iyakhilome: *Application of the kalman filter to interest rate modelling*. Munich Personal RePEc Archive, 2010.

[19] Ferstl, Robert and Hayden, Josef: *Zero-coupon yield curve estimation with the package termstrc*. Journal of Statistical Software, 2010.

[20] Scheffer, Marcus and Zacharias, Mario: *A comparative study of the 1-factor hull white and the g2++ interest rate model*, 2018. https://www.milliman.com/-/media/milliman/importedfiles/uploadedfiles/insight/2018/a_comparative_study_of_the_1-factor_hull_white.ashx, visited on 28.01.2023.

[21] Berninger, Christoph and Pfeiffer, Julian: *The gauss2++ model: a comparison of different measure change specifications for a consisten risk neutral and real world calibtration*. European Acturial Journal, 2021.

[22] (AVÖ), Aktuarvereinigung Österreichs: *Leitfaden zum Österreichischen branchstandard für priips der kategorie 4*, 2019. https://avoe.at/wp-content/uploads/2019/03/Leitfaden-AV\unhbox\voidb@x\bgroup\accent127O\protect\penalty\@M\hskip\z@skip\egroup-Branchenstandard-PRIIP-Kategorie-4-02_2019-final.pdf, visited on 28.01.2023.

[23] Annaert, Jan, Cleas, Anouk, De Ceuster, Marc, and Zhang, Hairui: *Estimating the yield curve using the nelson-siegel model*. Universiteit Antwerpen. https://www.efmaefm.org/0EFMAMEETINGS/EFMA%20ANNUAL%20MEETINGS/2010-Aarhus/papers/Smoothed_Bootstrap_-_Nelson-Siegel_Revisited_June_2010.pdf, visited on 28.01.2023.

[24] Svensson, Lars: *Estimating and interpreting forward interest rates:sweden 1992-1994*. NATIONAL BUREAU OF ECONOMIC RESEARCH, 1994. https://www.nber.org/system/files/working_papers/w4871/w4871.pdf, visited on 28.01.2023.

[25] Huang, Zhe: *Fitting yield curve with dynamic nelson-siegel models: Evidence from sweden*, 2021. http://uu.diva-portal.org/smash/get/diva2:1593645/FULLTEXT01.pdf, visited on 28.01.2023.

[26] Weißensteiner, Felix: *Zinskurvenmodelle und die prognose mithilfe des nelson-siegel modells*, 2017. https://unipub.uni-graz.at/obvugrhs/content/titleinfo/2293481/full.pdf, visited on 28.01.2023.

[27] Gilli, Manfred, Große, Stefan, and Schumann, Enrico: *Calibrating the nelson-siegel-svensson model*, 2010. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1676747, visited on 16.10.2022.

[28] Mullen, Katharine, Ardia, David, Gil, David, Windover, Donald, and Cline, James: *Package 'deoptim'*. https://cran.r-project.org/web/packages/DEoptim/DEoptim.pdf, visited on 28.01.2023.

[29] Vo, Long H.: *Application of kalman filter on modelling interest rates*. Journal of Management Sciences.

[30] Bolder, David Jamieson: *Affine term-structure models: Theory and implementation*. Fiancial Markets Department Bank of Canada.

[31] Kim, Youngjoo and Bang, Hyochoong: *Introduction to kalman filter and its applications*, 2018. https://cdn.intechopen.com/pdfs/63164.pdf, visited on 16.12.2022.

[32] Wikipedia-contributors: *Rudolf e. kalman*. Wikipedia, The Free Encyclopedia. https://en.wikipedia.org/wiki/Rudolf_E._Kalman, visited on 13.10.2022.

[33] RDocumentation: *lm:fitting linear models*. https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/lm, visited on 22.12.2022.

[34] GeeksforGeeks: *How to use lm() function in r to fit linear models*, 2019. https://www.geeksforgeeks.org/how-to-use-lm-function-in-r-to-fit-linear-models, visited on 28.10.2022.

[35] RDocumentation: *Ns:natural splines -(cubic splines linear beyond outermost knots) with convenient specification of knots and possibility of centering, detrending and clamping*. https://www.rdocumentation.org/packages/splines/versions/3.6.2/topics/ns, visited on 28.01.2023.

[36] Leal, Lois: *Numerical interpolation Natural cubic spline*. Towards Data Science, 2018. https://towardsdatascience.com/numerical-interpolation-natural-cubic-spline-52c1157b98ac, visited on 28.01.2022.

[37] Wikipedia-contributors: *Interpolation*. Wikipedia, The Free Encyclopedia, 2022. https://en.wikipedia.org/w/index.php?title=Interpolation&oldid=1129289791, visited on 11.01.2023.

[38] RDocumentation: *Linearly constrained optimization*. https://stat.ethz.ch/R-manual/R-devel/library/stats/html/constrOptim.html, visited on 27.12.2022.

[39] Sofia, Ayouche, Ellaia, Rachid, and Aboulaich, Rajae: *A hybrid particle swarm optimization-nelder-mead algorithm (pso-nm) for nelson-siegel-svensson calibration*. World Academy of Science, 2016.

[40] Wikipedia-contributors: *Nelder-mead method*. Wikipedia, The Free Encyclopedia. `https://en.wikipedia.org/w/index.php?title=Nelder\OT1\textendashMead_method&oldid=1114246182`, visited on 04.11.2022.

[41] Schumann, Enrico: *Fitting the nelson-siegel-svensson model with differential evolution*. `https://cran.r-project.org/web/packages/NMOF/vignettes/DEnss.pdf`, visited on 27.12.2022.

[42] Petrovska, Roksolana: *Using natural splines in linear modeling*, 2019. `https://www.clayford.net/statistics/using-natural-splines-in-linear-modeling/`, visited on 28.10.2022.

[43] RDocumentation: *predict.lm: Predict method for linear model fits*. `https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/predict.lm`, visited on 18.12.2022.

[44] Johansen, A.M: *Monte carlo methods*. International Encyclopedia of Education, 2010.

[45] Muralidhar, Krishnamurty: *Monte carlo simulation*. Encyclopedia of Information Systems, 2010.

[46] Moudiki, Thierry: *Esgtoolkit, a tool for stochsatic simulation (v0.2.0)*, 2020. `https://www.researchgate.net/publication/338549100_ESGtoolkit_a_tool_for_stochastic_simulation_v020`, visited on 23.10.2022.

[47] Miller, Steven J.: *The method of least squares*. Mathematics Department Brown University. `https://web.williams.edu/Mathematics/sjmiller/public_html/BrownClasses/54/handouts/MethodLeastSquares.pdf`, visited on 02.11.2022.

[48] RDocumentation: *predict.bs:evaluate a spline basis*. DataCamp. `https://www.rdocumentation.org/packages/splines/versions/3.6.2/topics/predict.bs`, visited on 04.11.2022.

[49] Fisher, Mark: *Modeling the term structure of interest rates: An introuduction*, 2004. Federal Reserve Bank of Atlanta.