MATEMATICKO-FYZIKÁLNÍ FAKULTA
Univerzita Karlova

MASTER THESIS

Jakub Kárný

Hull-White Model: Forward-looking vs. Backward-looking Approaches

Department of Probability and Mathematical Statistics

Supervisor of the master thesis: RNDr. Daniela Flimmel, Ph.D.
Study programme: Mathematics
Study branch: Financial and Insurance Mathematics

Prague 2024

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources. It has not been used to obtain another or the same degree.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

V .... dne ....

Author's signature

I would like to thank a few people who helped me with my study and understanding of the related topics and gave me helpful advice and recommendations. My thanks go to doc. RNDr. Jan Večeř, Ph.D., who helped me with practical problems directly involved in the simulation procedure. My thanks also belong to Mgr. Jan Vávra, Ph.D. and doc. Ing. Marek Omelka, Ph.D. for their explanatory consultations of more complicated aspects of statistics theory and for their advices how to use it and, more importantly, how not to use it.

Next, I need to thank RNDr. Samuel Flimmel, Ph.D. for supervising my simulation study, for providing and explanation of input data, for advice in my first steps of using Python, and for every single patient answer to my questions. I owe my greatest thanks to my supervisor RNDr. Daniela Flimmel, Ph.D. who was patient and helpful, who finds time on every consultation I need. She helped me understand advanced theory of stochastic processes, recommended me useful sources of information, and last but not least, she led me through the whole long creation process of my thesis.

Title: Hull-White Model: Forward-looking vs. Backward-looking Approaches

Author: Jakub Kárný

Department: Department of Probability and Mathematical Statistics

Supervisor: RNDr. Daniela Flimmel, Ph.D., Department of Probability and Mathematical Statistics

Abstract: The subject of the thesis is the estimation of the parameters of the Hull-White model. The formula for the time-dependent parameter θ is derived. The constant parameters α and σ are first calibrated on the prices of the chosen interest rate derivatives in the forward-looking approach and secondly on the historical yield curves in the backward-looking approach. Calibrated Hull-White models are simulated and then the approaches used are compared in terms of available data.

Keywords: Stochastic integral, Itô lemma, Hull-White model, spot and forward interest rate, calibration, cap, swaption

Název práce: Hull-Whiteův model: přístup dívající se vpřed a vzad

Autor: Jakub Kárný

Katedra: Katedra pravděpodobnosti a matematické statistiky

Vedoucí bakalářské práce: RNDr. Daniela Flimmel, Ph.D., Katedra pravděpodobnosti a matematické statistiky

Abstrakt: Práce se zabývá odhadem parametrů Hull-Whiteova modelu. Je odvozen tvar časově závislého parametru θ. Konstantní parametry α a σ jsou nejprve odhadnuty na základě cen vybraných finančních derivátů úrokových měr v přístupu dívajícím se vpřed and poté na historických vynosových křivkách v přístupu dívajícím se vzad. Nakalibrovaný Hull-Whiteův model je simulován a poté jsou použité přístupy kalibrace porovnány pomocí dostupných dat.

Klíčová slova: Stochastický integrál, Itôovo lemma, Hull-Whiteův model, spotová a forwardová úroková míra, kalibrace, cap, swapce

Contents

Introduction **3**

1 Basic theory **5**

1.1 Stochastic processes 5

1.2 Financial Mathematics 10

2 Hull-White model **14**

2.1 General model 14

2.2 Derivation of $\theta(t)$ 17

2.3 Parameter estimation of the Hull-White model 24

2.3.1 Forward-looking approach 24

2.3.2 Backward-looking approach 26

3 Simulation study **28**

3.1 Data describtion 28

3.2 Results for the forward-looking approach 32

3.2.1 Pricing of financial derivatives 32

3.2.2 Model calibration 35

3.2.3 Parameters estimation 39

3.2.4 Simulations 41

3.3 Results for the backward-looking approach 44

3.3.1 Parameters estimation 45

3.3.2 Simulations 45

Conclusion **47**

Introduction

Interest rates of all kinds are the basic and common concept of financial theory and practice, important for understanding economical situations, expressing the time value of money, and pricing. It is very advantageous to be able to know how interest rates can behave in the future. For this purpose, there exist various stochastic processes that describe the dynamics of an interest rate. Our aim is to introduce, estimate, and simulate the Hull-White process. The calibration process is not only about estimating the constant parameters but also about the derivation of the form of a time-dependent parameter.

In the first chapter, we need to define the basic concepts of stochastic processes and financial mathematics. The most important is to set tools for expressing the Hull-White model. We need to rigorously introduce the definition of a stochastic process with emphasis on the Wiener process. It is also necessary to define stochastic integrals and recall Itô's lemma for solving a stochastic differential equation (SDE) because the Hull-White model is given by an SDE. We define the term structure, which determines the form of interest rates. We introduce the definitions of both spot and forward rates. Finally, we adopt the definitions of interest rate financial derivatives used in calibration.

In the second chapter, we show the general model of an interest rate dynamics, then introduce the Hull-White model, and briefly discuss a few related models. With the help of Itô's lemma, we derive a solution to the Hull-White model SDE. Most importantly, we derive the form of the time-dependent parameter $\theta(t)$ of the Hull-White model in detail. In the end of the chapter, we discuss several approaches on how to calibrate our model. There are mainly two possibilities: the forward-looking approach based on a calibration of market prices on implied prices of interest rate derivatives from the model and the backward-looking approach based on comparing characteristics of historical data with theoretical characteristics.

In the last chapter, we perform our simulation study on the input data from Česká spořitelna. First, we describe them and then introduce market and model pricing formulas of caplets, caps, and swaptions for the forward-looking approach. We briefly describe the theory on which optimization functions from used software are based and verify assumptions for their usage. We comment and try to interpret the results of all calibration procedures. After that we simulate the Hull-White model for each possibility of calibration to easily illustrate what such results mean. To conclude, we recommend which procedures may be considered as a preferable calibration and which may not be appropriate.

Chapter 1

Basic theory

We introduce necessary definitions from the theory of stochastic processes connected also with stochastic integrals and stochastic differential equations to work with the Hull-White model in the first chapter. In the unified notation, we also define elementary concepts from financial mathematics.

1.1 Stochastic processes

Firstly, we choose a common probability space and then define a stochastic process as a set of random variables on the respective space according to Prášková [2016].

Notation. *The common probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is as a trio of a set of elementary events* Ω*, a* σ*-algebra of subsets of* Ω *denoted by* $\mathcal{F}$ *and a measurable function* $\mathbb{P} : \mathcal{F} \to [0, 1]$ *called probability.*

Let $(\mathbb{R}, \mathcal{B})$ *be a real state space with Borel* σ*-algebra* $\mathbb{R}$ *and* $\mathcal{T} = [0, T]$*,* $T \geq 0$ *be an index set representing time. A set of random variables* $\{X_t, t \in [0, T]\}$ *with values in* $\mathbb{R}$ *is called a stochastic process.*

We denote $\mathcal{B}[0, t]$ *the Borel* σ*-algebra in the interval* $[0, t]$*,* $t \in [0, T]$*. The notation* $\mathbb{I}\{y \in (y_1, y_2)\}$ *is used as an indicator that takes on 1 if* y *belongs to the interval* (y_1, y_2) *and 0 otherwise.*

Next, we need to introduce a filtration as a system of specific σ-algebras to be able to define the measurability of a stochastic process with respect to this system. In addition, we introduce the type of measurability called progressive. These definitions and the following examples are taken from Karatzas and Shreve [1991].

Definition 1 (Filtration)**.** *A filtration on a probability space* $(\Omega, \mathcal{F}, \mathbb{P})$ *is a non-decreasing system* $\{\mathcal{F}_t,\, t \geq 0\}$ *of* σ *- algebras of* Ω *such that* $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}$
$\forall t > s \geq 0$*.*

Definition 2 (Measurability)**.** *The stochastic process* $X = \{X_t, t \in [0,T]\}$ *is called measurable if the mapping*

$$(\omega, t) \to X_t(\omega) : (\Omega \times [0,T], \mathcal{F} \otimes \mathcal{B}[0,T]) \to (\mathbb{R}, \mathcal{B})$$

is measurable.

Definition 3 (Progressive measurability)**.** *The stochastic process* $X = \{X_t, t \in [0,T]\}$ *is called* $\mathcal{F}_t$*-progressively measurable if the mapping*

$$(\omega, s) \to X_s(\omega) : (\Omega \times [0,t], \mathcal{F}_t \otimes \mathcal{B}[0,t]) \to (\mathbb{R}, \mathcal{B})$$

is measurable for every $t \in [0,T]$*.*

We show two basic examples of stochastic processes that are important for our further work. First of them is the *simple process* defined on a division of a chosen interval easily by suitable random variables and indicators that the particular variable is in the respective subinterval of the division. The paths of such a process are piecewise constant.

Example (Simple Process). Let $\mathcal{F} = \{\mathcal{F}_t, t \geq 0\}$ be a filtration and $\{0 = t_0 < \, < t_1 < \cdots < t_n = T\}$ be a division of an interval $[0,T]$. Let $(\xi_j, \ j = 0, \ldots, n-1\ )$ be a sequence of random variables such that ξ_j is a $\mathcal{F}_{t_j}$-measurable random variable for all j. The stochastic process $G = (G_t, \ t \in [0,T])$ is called a simple stochastic process in the interval $[0,T]$ if

$$G_t = \xi_0 \mathbb{I}\{t = 0\} + \sum_{j=0}^{n-1} \xi_j \mathbb{I}\{t \in (t_j, t_{j+1}]\}. \tag{1.1}$$

We denote by $\mathbb{L}_0(\mathcal{F}, [0.T])$ (shortly $\mathbb{L}_0$) the set of all simple processes in the interval $[0,T]$. The second example is the *Wiener process*. The Wiener process is important to define stochastic integrals and an SDE and is a special Gaussian process defined as follows.

Example (Wiener Process). A stochastic process $W = \{W_t, t \in [0,T]\}$ on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ satisfying the following properties

W1 $W_0 = 0$,

W2 for $0 \leq s < t \leq T$ the random variable $W_t - W_s$ is Gaussian with zero mean and variance $t - s$,

W3 for each $0 \leq t_1 \leq t_2 \leq t_3 \leq t_4$ the random variables $W_{t_2} - W_{t_1}$ and $W_{t_4} - W_{t_3}$ are independent,

W4 the paths are continuous for $\mathbb{P}$-almost all $\omega \in \Omega$,

is called a Wiener process.

Our next goal is to introduce the stochastic integral with respect to the Wiener process. It is done by taking an $\mathcal{L}_2$ limit of a sequence of integrals of simple processes. These are defined with the use of sums from the equation (1.1) with increments of the Wiener process instead of the indicators. The following part is taken from Øksendal [2003].

Definition 4 (Itô's stochastic integral of a simple process). *Let $W = \{W_t, t \in [0,T]\}$ be the Wiener process and $G = \{G_t, t \in [0,T]\}$ be the simple process. The process*

$$\int_0^t G\,dW = \sum_{j=0}^{k-1} \xi_j (W_{t_{j+1}} - W_{t_j}) + \xi_k (W_t - W_{t_k}),\ t \in [0,T], t_k \leq t < t_{k+1}$$

is called Itô's stochastic integral of a simple process G.

It is also necessary to define metrics that describe the limit behaviour of each simple process from the sequence of stochastic integrals from Definition 4. For this purpose, we need to describe a space of $\mathcal{F}_t$-progressively measurable processes where this limit behaviour works.

We set a space $\mathbb{L}_2(\mathcal{F}_t, [0.T])$ (shortly $\mathbb{L}_2$) as the set

$$\{X = (X_t, t \in [0,T]) : X \text{ is } \mathcal{F}_t \text{ progressively measurable}, \quad \mathsf{E} \int_0^t X_s^2\,ds < \infty\ \forall t \in [0,T]\}. \tag{1.2}$$

Analogously, we denote the space $\mathbb{J}_2$ for only simple processes in $\mathbb{L}_2$.

Definition 5 (Metric l). *Let $X, Y \in \mathbb{L}_2$. Then we define the metric $l(X,Y)$ by*

$$l(X,Y) = E \int_0^T (X - Y)^2\,ds.$$

It is important to note that the space $\mathbb{J}_2$ is dense in the space $\mathbb{L}_2$ which means that $\forall X \in \mathbb{L}_2$ there exists a sequence $\{G_n\}_{n\in\mathbb{N}} \subset \mathbb{J}_2$ such that $G_n \xrightarrow{l} X$. Due to this fact, we can give a meaning to the general Itô's stochastic integral of every process (not only simple) with respect to the Wiener process.

Definition 6 (Itô's stochastic integral). *Itô's stochastic integral $\int X\,dW$ of a process $X \in \mathbb{L}_2$ is the $\mathcal{L}_2$ limit of a sequence of $\{\int G_n\,dW\}_{n\in\mathbb{N}}$ where $G_n \in \mathbb{J}_2$ is such a sequence of simple processes for which $l(G_n, X) \to 0$ for $n \to \infty$.*

We introduce the general definition of an SDE and its solution. The Hull-White model and its dynamic are then described by a concrete form of the linear SDE for a process of interest rates. Then it is followed by examples of SDE such as a linear SDE and a linear SDE in the narrow sense. This part is taken from Panik [2017].

Definition 7 (SDE and its solution). *The stochastic differential equation is a stochastic differential of an unknown process X of the form*

$$dX_t = b(t, X_t)dt + \sigma(t, X_t)dW_t,\ t \in (0,T) \tag{1.3}$$

with an initial condition X_0 where X_0 is required to be a random variable independent of an increment of the Wiener process $W_t - W_0$ such as $\mathsf{E}[|X_0|^2] < \infty$. Then a solution to this SDE is an $\mathbb{R}$-valued stochastic process X satisfying

$$X_t = X_0 + \int_0^t b(s, X_s)\, ds + \int_0^t \sigma(s, X_s)\, dW_s,\ t \in [0, T],$$

where $\int_0^t b(s, X_s)\, ds$ is understood in the Lebesgue sense $\mathbb{P}$-a.s while $\int_0^t \sigma(s, X_s)\, dW_s$ is an integral from Definition 6.

We also add a theorem about the conditions under which a solution to a general SDE exists and is unique according to Panik [2017]. Later we show how it translates to linear SDE's.

Theorem 1 (Existence and uniqueness theorem for a general SDE)**.** *If the functions b and σ from Definition 7 satisfy the uniform Lipschitz condition*

$$|b(t,x) - b(t,y)| + |\sigma(t,x) - \sigma(t,y)| \leq c_1|x-y|\ \forall x, y$$

for some constant c_1 and all $t \in [0,T]$ and satisfy also the linear growth condition

$$|b(t,x)|^2 + |\sigma(t,x)|^2 \leq c_2(1+|x|^2)\ \forall x$$

for some constant c_2 and all $t \in [0,T]$, then there exists a (continuously adapted) solution X_t of the equation (1.3) *with the initial condition X_0 and uniformly bounded in the space of all real-valued processes X on the product space $\Omega \times [0,T]$. Furthermore, if X_t and $\tilde{X}_t$ are continuous and bounded solutions of the equation* (1.3) *in the space described above, then*

$$\mathbb{P}\left(\sup_{0 \leq t \leq T} |X_t - \tilde{X}_t| = 0\right) = 1$$

and thus the solution X_t is unique.

We present the following special examples of SDE because the Hull-White model is described by these specific forms. This particular identification is very useful for derivation purposes in our work.

Example (Linear SDE). Let the functions b and σ from Definition 7 be linear of X on $\mathbb{R} \times [0,T]$, meaning that

$$b(t, X_t) = a(t) + A(t)X_t,\ t \in [0,T], \tag{1.4}$$
$$\sigma(t, X_t) = c(t) + C(t)X_t,\ t \in [0,T], \tag{1.5}$$

where a(t), A(t), c(t), C(t) are all mappings from $[0,T]$ to $\mathbb{R}$. Then we define a linear SDE

$$dX_t = (a(t) + A(t)X_t)dt + (c(t) + C(t)X_t)dW_t,\ t \in [0,T] \tag{1.6}$$

with an initial condition X_0.

Example (SDE linear in the narrow sense). Let us have the linear SDE with the formula (1.6) and let the function $C(t)$ be set to 0 everywhere. We call the SDE linear in the narrow sense if

$$dX_t = (a(t) + A(t)X_t)dt + c(t)dW_t,\ t \in (0, T). \tag{1.7}$$

Specifically, if the functions $a(t)$, $A(t)$, $c(t)$ and $C(t)$ from equations (1.4) and (1.5) are measurable and bounded on $[0, T]$ then the functions b and σ satisfy the conditions of Theorem 1 and the solution exists. Furthermore, if X_0 satisfies the conditions from Definition 7, then the solution is continuous and unique over our interval $[0, T]$. In the next part of this section, we can declare famous Itô lemma according to Friedman [2006] and Øksendal [2003].

Theorem 2 (Itô lemma). *Let V be a function twice continuously differentiable on $[0,T] \times \mathbb{R}$ and assume that the process $(X_t)_{t\in[0,T]}$ is given by the stochastic differential* (1.3)*, where $b(t, X_t)$ and $\sigma(t, X_t)$ are both an $\mathbb{R}$ - valued $(\mathcal{F}_t)$ - progressive process that satisfies*

$$\int_0^T |b(t, X_t)|\, dt < \infty\ \mathbb{P} - a.s., \tag{1.8}$$
$$\int_0^T |\sigma(t, X_t)|^2\, dt < \infty\ \mathbb{P} - a.s.$$

Then the process $Y_t = V(t, X_t)$ has a stochastic differential

$$dY_t = \left[\frac{\partial V}{\partial t}(t, X_t) + b(t, X_t)\frac{\partial V}{\partial x}(t, X_t) + \frac{1}{2}\sigma(t, X_t)^2\frac{\partial^2 V}{\partial x^2}(t, X_t)\right]dt + \\ + \sigma(t, X_t)\frac{\partial V}{\partial x}(t, X_t)dW_t. \tag{1.9}$$

Remark 1. We can see that the conditions of Itô's lemma on Lebesgue integrability for the linear SDE are satisfied, provided there exists solution of such an SDE under assumptions of Theorem 1. This is true because we need measurable and bounded processes b and σ according to their definition in equations (1.4) and (1.5) for the existence of the solution. Hence, they are Lebesgue integrable over a finite interval $\mathbb{P}$-a.s.

In order to differentiate a stochastic Itô integral with respect to its upper bound during later derivations, Theorem 3 can be useful. It is stated in Proposition 1.3 in Hess [2023].

Theorem 3 (Stochastic Leibniz formula). *Assume that the deterministic function $\psi : (0, \infty) \times (0, \infty) \to \mathbb{R}$ is continuous in the first variable and continuously differentiable in the second. Also suppose that for all $t \in (0, \infty)$ it holds*

$$\int_0^t \left[\psi^2(s, t) + \left(\frac{\partial \psi(s, t)}{\partial t}\right)^2\right] ds < \infty,\ t \geq s \geq 0. \tag{1.10}$$

Then for all $t \in (0, \infty)$ we have

$$d\left(\int_0^t \psi(s, t)dW_s\right) = \left(\int_0^t \frac{\partial \psi(s, t)}{\partial t}dW_s\right) + \psi(t, t)dW_t, \tag{1.11}$$

where d denotes the stochastic differential from Definition 7.

1.2 Financial Mathematics

In this section, we introduce some basic definitions from the point of view of financial mathematics. While doing so, we aim to respect the notation from the previous section. These definitions are taken from Brigo and Mercurio [2006]. We introduce the first two to explain the pricing mechanism of financial derivatives used in our simulation study. We start with a definition of a *zero-coupon bond* which we use as a numeraire (any positive non-dividend-paying asset).

Definition 8 (Zero-coupon bond)**.** *A T-maturity zero-coupon bond is a contract that guarantees its holder the payment of one unit of currency at time T, without intermediate payments. The contract value at time $t < T$ is denoted by $P(t,T)$ and, therefore, $P(T,T) = 1$ for all T.*

The basic concepts for financial mathematics are *spot* and *forward interest rates*. For creation of a yield curve in a pricing mechanism for our purpose of the estimate of the Hull-White model and for our data, we use *continuously compounded spot interest rate* and then we define yield curve through it.

Definition 9 (Continuously compounded spot interest rate)**.** *The continuously compounded spot interest rate prevailing at time t for the maturity T is denoted by $R(t,T)$ and it is the constant rate at which an investment of $P(t,T)$ units of currency at time t accrues continuously to yield a unit amount of currency at maturity T. In formulas:*

$$R(t,T) = -\frac{\log P(t,T)}{T-t}. \tag{1.12}$$

From Definition 9 it is clear that with the use of a continuously compounded spot interest rate, we can express a zero-coupon bond at time t with maturity T as

$$P(t,T) = e^{-R(t,T)(T-t)}. \tag{1.13}$$

Definition 10 (Instantaneous spot interest rate)**.** *The instantaneous spot interest rate or the short rate r_t is a limit of the spot rate*

$$r_t := \lim_{T \to t^+} R(t,T).$$

Definition 11 (Term structure of interest rates)**.** *The term structure of interest rates at time t (also called a zero-coupon curve or yield curve) is the graph of the function*

$$T \mapsto R(t,T) \text{ for } t < T.$$

The function in Definition 11 at time $t = 0$ is also called the *initial term structure of interest rates*. We can see an example of this curve constructed from our data from the Czech financial market in Figure (1.1). We present an inverted yield curve that displays a higher yield for shorter maturities than for longer.

This situation is not standart from the point of view of yield curves. Our yield curve shows that in the Czech market it is expected that short-term PRIBOR (Prague Interbank Offered Rate) based on the Czech National Bank repo rate will decrease in future years.

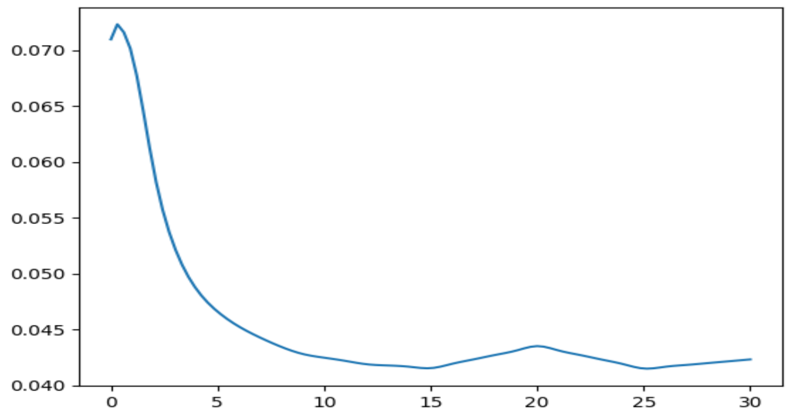

Figure 1.1: Initial term structure (yield curve at time $t = 0$) from our study data starting January 31, 2023 and ending on the last maturity date after 30 years.

All interest rate models considered in this work, mainly the Hull-White model, belong to the special group of the so-called *affine term structure models*, which possess an affine term structure of an interest rate.

Definition 12 (Affine term structure of an interest rate)**.** *We call the term structure of an interest rate affine if the countinuously compounded spot interest rate $R(t,T)$ from Definition 11 is an affine function of a short rate r_t*

$$R(t,T) = \beta(t,T) + \gamma(t,T)r_t,$$

where β and γ are deterministic functions of time and maturity.

We can use Definition 9 and the relationship between $R(t,T)$ and $P(t,T)$ in the equation (1.13) to show that this condition of the affine term structure is satisfied when

$$P(t,T) = e^{-(\beta(t,T)+\gamma(t,T)r_t)(T-t)}$$

and it equals

$$P(t,T) = B(t,T)e^{-\Gamma(t,T)r_t} \tag{1.14}$$

with

$$B(t,T) = e^{-\beta(t,T)(T-t)},$$
$$\Gamma(t,T) = \gamma(t,T)(T-t).$$

The next two definitions are dealing with forward interest rates, which are important for adopting the Hull-White model and its parameters and characteristics. Specifically, an instantaneous forward rate also gives us the possibility to consider SDE describing dynamics of the Hull-White model and formulate an exact solution to it.

Definition 13 (Simply compounded forward interest rate)**.** *The simply compounded forward interest rate prevailing at time t for the expiry $T > t$ and maturity $S > T$ is denoted by $F(t; T, S)$ and is defined by*

$$F(t;T,S) = \frac{1}{S-T}\left(\frac{P(t,T)}{P(t,S)} - 1\right). \tag{1.15}$$

Definition 14 (Instantaneous forward interest rate)**.** *The instantaneous forward interest rate prevailing at time t for the maturity $T > t$ is denoted by $f(t,T)$ and is defined as*

$$f(t,T) = \lim_{S\to T^+} F(t;T,S) = -\frac{\partial \log P(t,T)}{\partial T}. \tag{1.16}$$

For practical simulations, it is important to define some real products. The purpose is that we estimate the parameters of the Hull-White model with a usage of the calibration mechanism on specific financial derivatives. We use *caps* and *swaptions* in our case.

Definition 15 (Cap)**.** *Let $\{T_1, \ldots, T_k\}$ be a set of prespecified dates, τ_i be a tenor (year fraction between the dates T_{i-1} and T_i, $i \in \{2, \ldots, k\}$), N be the nominal value of a product, K be a fixed strike price (interest rate) and $L(T_{i-1}, T_i)$ be a floating LIBOR rate between the dates T_{i-1} and T_i, $i \in \{2, \ldots, k\}$. Then a cap is an interest rate derivative contract that pays out*

$$N\,\tau_i\,(L(T_{i-1}, T_i) - K)^+ \tag{1.17}$$

at each of the prespecified dates (positive only if floating LIBOR exceeds the strike price K).

Remark 2. An interest rate derivative at only one prespecified date is called a *caplet*. For each caplet, time T_{i-1} is called the *fixing date* and time T_i is called the *maturity date* in the equation (1.17).

Example. Caps then consist of individual caplets with the same strike price. We show an example of a payoff function based on the equation (1.17) of one caplet with nominal value 1 currency unit, tenor 0,5 year and strike price 5% as the function of an underlying interest rate z given by

$$p_{caplet}(r) = 1 * 0{,}5 * (r - 0{,}05, 0)^+. \tag{1.18}$$

We can see that caplet products serve as hedging against exceeded interest rates above the strike price. For example, if we have a caplet product with the strike price 5% and the floating rate exceeds this strike price, we get a payment from the caplet to cover the higher interest rate than we want. That means we only pay the interest rate in maximal value of the strike price.

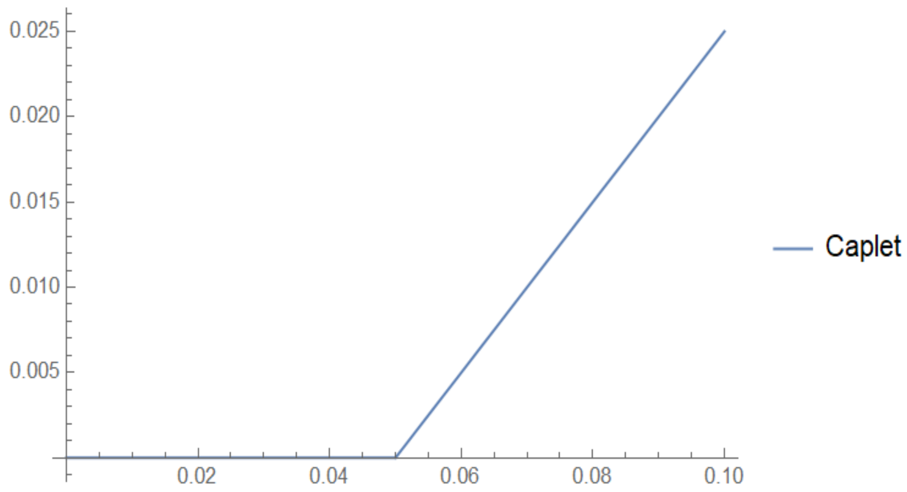

Figure 1.2: The payoff function for one caplet from the equation (1.18) for $z \in [0, 0{,}1]$.

Definition 16 (Swaption). *Let $\{T_1, \ldots, T_k\}$ be a set of the prespecified dates, T_1 is considered to be the expiry date. $T_k - T_1$ be a tenor, τ_i is a year fraction between the dates T_{i-1} and T_i, $i \in \{2, \ldots, k\}$, N be a nominal value of a product, K be a fixed strike price (interest rate) and $F(T_1; T_{i-1}, T_i)$ be forward interest rates from Definition (13), $i \in \{2, \ldots, k\}$. Then a payer swaption is an option on a payer interest rate swap which exchanges payments*

$$N\,\tau_i\,K$$

(paid) with

$$N\,\tau_i\,F(T_1; T_{i-1}, T_i)$$

(received) at each of the prespecified dates.

Remark 3. Swaption cannot be decomposed into smaller derivatives at each prespecified date, contrary to caps. Thus, we have to consider all exchange payments for discounted payoff functions of the interest rate swap at the expiry date T_1 and of swaption at time t:

$$N\,D(t, T_1)\left(\sum_{i=2}^{k} P(T_1, T_k)\,\tau_i(F(T_1; T_{i-1}, T_i) - K)\right)^+,$$

where $D(t, T_1)$ is a discount factor for swaption from the expiry date to time t and $P(T_1, T_i)$ is a discount factor for each exchange payment from time T_i to the expiry date where the interest rate swap starts. Swaption pays out only if a value of the interest rate swap is positive.

Chapter 2

Hull-White model

In this chapter, we introduce a one-factor interest rate Hull-White model. We start with the general version and then show a specific form of this model. Our goal is to estimate parameters in the backward and forward sense, and we introduce possible options to provide the final estimate.

2.1 General model

We start this chapter by introducing general dynamics for the instantaneous short-term interest rate so as to be able to show examples and describe a model which we are mainly interested in. This chapter is based on the knowledge from [Brigo and Mercurio, 2006].

Definition 17 (General model for instantaneous short-term interest rate)**.** *A general model for the dynamics of the instantaneous short-term interest rate r_t under the risk neutral measure is described by a linear SDE*

$$dr_t = (\theta(t) - \alpha(t)r_t))dt + \sigma(t)dW_t, \tag{2.1}$$

where the interest rate r_t is reverted to a long-term mean level $\frac{\theta}{\alpha}$ at a reversion rate α, parameter σ determines the overall level of volatility, and W_t is the Wiener process. Here, all parameters are allowed to be functions of time t.

Remark 4. An interesting concept is the mean reversion effect. This principle has no general approach to how to define it. It is widely discussed by Exley et al. [2004]. We know that the mean reversion effect is given by the long-term mean level and the reversion rate, as is said in Definition 17. Alternatively, we may rewrite (2.1) as follows

$$dr_t = \alpha(t)\left(\frac{\theta(t)}{\alpha(t)} - r_t\right)dt + \sigma(t)dW_t \tag{2.2}$$

to have a better intuition for the mean reversion level $\frac{\theta(t)}{\alpha}$. This formula gives us a hint that the stochastic process for a short interest rate in this form is driftless at the mean reversion level.

In our work, we use the definition of the mean reversion recommended in Exley et al. [2004].

Definition 18 (Mean reversion effect)**.** *Let us take times* $t_1 < t_2 < t_3 < t_4$*. A process* X_t *is mean reverting if its increments over disjoint intervals* $X_{t_2} - X_{t_1}$*,* $X_{t_4} - X_{t_3}$ *are negatively correlated.*

This definition works with our intuition that after an increase of an interest rate above the long-term mean level it comes a decrease and vice versa. It is clear that when the interest rate model is constructed as driftless at the mean reversion level, the model has a negative drift above this level and it has a positive drift below this level.

We explore a few models implied from the model given by the equation (2.1) as examples, such as Vašíček, Hull-White, and Ho-Lee models. Later on, we restrict our interest only to the Hull-White model.

Example (Vašíček model). The Vašíček model is the basic model for the short-term interest rate r_t with the mean reversion effect. It is described by the *Ohrnstein-Uhlenbeck process*

$$dr_t = \alpha(\theta - r_t)dt + \sigma dW_t, \tag{2.3}$$

where all parameters are positive constants and the interpretation of parameters stays the same as in Definition 17. We can also see that this formula is given exactly the same way as in the formula (2.2). It is constructed in the sense that a constant parameter θ is the mean reversion level directly.

Example (Hull-White model). It is the direct extension of the Vašíček model from the previous example with the mean level $\theta(t)$ as a function of time t given by a process

$$dr_t = (\theta(t) - \alpha r_t))dt + \sigma dW_t. \tag{2.4}$$

Example (Ho-Lee model)*.* The Hull-White model from the equation (2.4) can also be considered as an extension of the Ho-Lee model with an added mean reversion effect. Therefore, the SDE for the Ho-Lee model is given by

$$dr_t = \theta(t)dt + \sigma dW_t \tag{2.5}$$

with the mean reversion parameter $\alpha = 0$ and $\theta(t) = \frac{\partial f(0,t)}{\partial t} + \sigma^2 t$.

Those are three particular examples of the general dynamics model. As was said earlier, we want to estimate parameters α and σ of the Hull-White model from the equation (2.4). This model is very used and an important interest rate model. Oppositely to the Vašíček model, the Hull-White and Ho-Lee models are able to provide the exact fit of the initial term structure of interest rates from Definition 11.

The exact fit provides a match of the initial term structure given by the model and observed in the market including a perfect match between volatility term structure. This is a certain advantage and leads to the creation of models such as the Hull-White and the Ho-Lee. The Hull-White model is also more advantageous in comparison to the Ho-Lee model in the sense of preserving the mean reversion effect.

Next, in our work, we prefer the Hull-White model over the general model because all parameters as functions of time hide some practical disadvantages. The easier model with fewer parameters (only one time-varying) is advantegous than the general model with three time-varying parameters with respect to our desirable feature of a possibility of the exact fit. Namely, it is important to also fit volatility term structure and there are some less liquid, not relevant volatilities for less traded maturities on a market on which the general model is sensitive. Modelled rates may then be affected.

On the other hand, the model from the equation (2.1) enables r_t to be negative through the normal distribution given by the Wiener process, and it is the general problem of this model. To ensure r_t being only nonnegative, one has to try different models, for example, the geometric Brownian motion based on lognormal distribution.

The Vašíček model (2.3) is in the form of the linear SDE according to the formula (1.7), hence we can try to express an analytical form of a solution to the equation given by (2.3) as the short-term interest rate r_t with the initial condition r_0 being a random variable r_0 satisfying the conditions in Definition 7. We apply Itô's lemma from Theorem 2 on the transformation of r_t in the form

$$V(t, r_t) = r_t e^{\alpha t}. \tag{2.6}$$

Then, we can express a unique solution to the Vašíček model after integrating the transformed differential from 0 to t, $t \geq 0$:

$$r_t = r_0 e^{-\alpha t} + \theta(1 - e^{-\alpha t}) + \sigma \int_0^t e^{-\alpha(t-u)} dW_u.$$

The Hull-White model (2.4) is also written as the linear SDE. The way we express a solution to this SDE is the same, through Itô's lemma used on the transformed differential (2.6), and with the same initial condition as in the case of the Vašíček model. Based on our main interest in this model, we try to verify assumptions of Itô's lemma rigorously.

The transforming function in the equation (2.6) fulfills the assumption of Theorem 2 about its continuous first and second derivatives on $[0, T] \times \mathbb{R}$. In the sense of Remark 1 we need to verify the measurability and boundedness of functions

$$\begin{aligned} a(t) &= \theta(t), \\ A(t) &= \alpha, \\ c(t) &= \sigma. \end{aligned}$$

We can immediately say that constant functions α and σ are measurable and bounded, hence satisfying the integrability assumptions of the Itô's lemma.

The constant σ also satisfies the other assumption of progressivity of the process $\sigma(t, X_t)$. The conditions put on a process $b(t, X_t)$ and a function $\theta(t)$ will be discussed and verified later after the derivation of an analytical form of $\theta(t)$ in the Hull-White in Section 2.2. After integrating the differential given by Itô's lemma, we get the solution

$$r_t = r_0 e^{-\alpha t} + \int_0^t \theta(u) e^{-\alpha(t-u)} du + \sigma \int_0^t e^{-\alpha(t-u)} dW_u. \tag{2.7}$$

Thanks to the Wiener process involved in the Itô integral in the solution (2.7), the interest rate r_t in the Hull-White model is normally distributed conditioned on σ-algebra $\mathcal{F}_s$ with mean and variance

$$\mathsf{E}\left[r_t|\mathcal{F}_0\right] = r_0 e^{-\alpha t} + \int_0^t \theta(u) e^{-\alpha(t-u)} du, \tag{2.8}$$

$$\operatorname{var}[r_t|\mathcal{F}_0] = \frac{\sigma^2}{2\alpha}(1 - e^{-\alpha t}). \tag{2.9}$$

The structure of $\mathcal{F}_0$ is generated by the information about the interest rate r_v at time 0, which means that $\mathcal{F}_0$ is the σ-field $\sigma\{r_0\}$.

2.2 Derivation of $\theta(t)$

The advantage of the Hull-White model is that it can fit exactly the initial term structure of interest rates. It means that this model can handle simulations of interest rates better from an unbiased origin. We achieve this by setting a specific form of the function $\theta(t)$.

Theorem 4 ($\theta(t)$ for the Hull-White model)**.** *The function $\theta(t)$ such that the Hull-White model* (2.4) *can exactly fit the initial term structure of an interest rate has the following form:*

$$\theta(t) = \frac{\partial f(0,t)}{\partial t} + \alpha f(0,t) + \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}), \tag{2.10}$$

where f is given by Definition 14.

Corollary. With the formula for $\theta(t)$ given by Theorem 4, we can express r_t as a solution of the Hull-White model from the equation (2.7):

$$r_t = r_s e^{-\alpha(t-s)} + \omega(t) - \omega(s) e^{-\alpha(t-s)} + \sigma \int_s^t e^{-\alpha(t-u)} dW_u,$$

where

$$\omega(t) = f(0,t) + \frac{\sigma^2}{2\alpha^2}(1 - e^{-\alpha t})^2.$$

Remark 5. We need to set the functions $B(t,T)$ and $\Gamma(t,T)$ in the affine term structure from the equation (1.14) for the Hull-White model because we take advantage of the affine term structure in the proof of Theorem 4. According to Brigo and Mercurio [2006], those functions are given in the Hull-White model by

$$\Gamma(t,T) = \frac{1}{\alpha}\left(1 - e^{-\alpha(T-t)}\right), \tag{2.11}$$

$$\log B(t,T) = \log P(0,T) - \log P(0,t) + \Gamma(t,T)f(0,t) - \frac{\sigma^2}{4\alpha}\left(1 - e^{-2\alpha t}\right)\Gamma(t,T)^2. \tag{2.12}$$

Proof of Theorem 4. The derivation is based on Hull [2017a] and its technical report Hull [2017b] and we comment and show each step **S1** – **S5** in a more detailed way.

S1 We derive a differential $dP(t,T)$ using Itô's lemma on transformed r_t following the Hull-White model.

S2 We express a differential $dF(t;T,S)$ in terms of the differential $dP(t,T)$.

S3 We obtain an instantaneous spot rate r_t by taking a limit of the spot rate expressed by the forward rate.

S4 We derive $\theta(t)$ by differentiating r_t with respect to t and modifying dr_t in the form of the Hull-White model from the equation (2.4).

S5 We verify retrospectively the assumptions of Itô's lemma on $\theta(t)$.

The first step is to establish an SDE for $P(t,T)$ under the risk-neutral measure with a function $u(t,T)$ such as $u(t,t) = 0$ which will be explicitly delivered later. The technical report Hull [2017b] proposes the normal choice of $dP(t,T)$ under the risk-neutral measure with a return equal to the short rate r_t:

$$dP(t,T) = r_t\, P(t,T)dt + u(t,T)\, P(t,T)dW_t, \tag{2.13}$$

where r_t follows the Hull-White model. We try to connect the form of the Hull-White model from the equation (2.4) with the term structure of an interest rate represented by the spot rate $R(t,T)$ in this derivation. Therefore, when we get a short rate r_t taking the limit of a spot rate, we can find a formula for $\theta(t)$ which links the model with the term structure of an interest rate.

The dynamics from the equation (2.13) is implied by the affine term structure in Definition 12 secured by the Hull-White model. Then we can use $P(t,T)$ in a form of the equation (1.14) and use Itô's lemma on

$$V(t,r_t) = P(t,T) = e^{\ln B(t,T) - \Gamma(t,T)r_t}. \tag{2.14}$$

It is important to realize what form the functions $b(t,r_t)$ and $\sigma(t,r_t)$ take in Itô's lemma. They are given by the stochastic process of r_t (the Hull-White model) in the equation (2.4) such that

$$b(t,r_t) = \theta(t) - \alpha r_t,$$
$$\sigma(t,r_t) = \sigma.$$

The verification of the assumptions on these two functions was already discussed earlier during the derivation of the solution of the Hull-White SDE above the equation (2.7)

We can show that the counterparts of the stochastic process generated by Itô's lemma there are given by

$$\begin{aligned}
\frac{\partial V}{\partial t}(t, r_t) &= P(t,T)\left[\frac{\partial \ln B(t,T)}{\partial t} - \frac{\partial \Gamma(t,T)}{\partial t} r_t\right], \\
b(t, r_t)\frac{\partial V}{\partial r}(t, r_t) &= -P(t,T)(\theta(t) - \alpha r_t)\Gamma(t,T), \\
\frac{1}{2}\sigma(t, r_t)^2 \frac{\partial^2 V}{\partial r^2}(t, r_t) &= \frac{1}{2} P(t,T)\, \sigma^2 \Gamma(t,T)^2, \\
\sigma(t, r_t)\frac{\partial V}{\partial r}(t, r_t) &= -P(t,T)\, \sigma\Gamma(t,T).
\end{aligned}$$

By plugging this into (1.9) we get the SDE

$$\begin{aligned}
dP(t,T) = P(t,T)\Bigg[\Bigg(\frac{\partial \ln B(t,T)}{\partial t} - \frac{\partial \Gamma(t,T)}{\partial t} r_t - (\theta(t) - \alpha r_t)\Gamma(t,T) + \\
+ \frac{1}{2}\sigma^2\Gamma(t,T)^2\Bigg) dt - \sigma\Gamma(t,T) dW_t\Bigg]. \quad (2.15)
\end{aligned}$$

The function $u(t,T)$ is then easy to derive from (2.15). From a comparison of the coefficients by dW_t in equations (2.13) and (2.15) we can see that

$$u(t,T) = -\sigma\Gamma(t,T) = -\frac{\sigma}{\alpha}\left(1 - e^{-\alpha(T-t)}\right), \quad (2.16)$$

which satisfies the condition $u(t,t) = 0$.

We are unable to see without knowledge about the form of $\theta(t)$ that a drift part of the equation (2.13) is given by the short rate r_t itself. On the other hand, we know that the Hull-White model is a generalization of the Vašíček model and $dP(t,T)$ for these two models differs only in the stochastic term according to Hull [2017a] and Hull [2017b]. The Vašíček model is also an affine term structure model, therefore we can derive $dP(t,T)$ in the same way as in the case of the Hull-White model using Itô's lemma on a transformation (2.14) with the same function $\Gamma(t,T)$ and different $\ln B(t,T)$.

The reason why $\ln B(t,T)$ is different in the Vašíček model is mainly given by the fact that this model works on a different principle with respect to an initial term structure and a mean reversion level is given by a constant parameter, not a function. In the case of constant parameters, we can get $r_t\, P(t,T)dt$ in $dP(t,T)$ with some basic algebra from the equation (2.15). From that we finish step **S1** and get for the Hull-White model

$$dP(t,T) = r_t\, P(t,T)dt - \sigma\Gamma(t,T)\, P(t,T)dW_t.$$

Our next goal is to obtain the forward rate $F(t;T,S)$, since we need to represent the spot rate $R(t,T)$ by the forward rate. It is possible to see from Definition 14 that we can work with a formula

$$F(t;T,S) = -\frac{\log P(t,S) - \log P(t,T)}{S-T}, \ S > T, \tag{2.17}$$

which is more advantageous than the formula from Definition 13. That is because we easily create a differential for the forward rate and with the usage of Itô's lemma, we substitute for a differential of the logarithm of zero-coupon bond value from the equation (2.13).

The differential for the forward rate is

$$dF(t;T,S) = \frac{d\log P(t,T)}{S-T} - \frac{d\log P(t,S)}{S-T}. \tag{2.18}$$

We obtain differentials of logarithms of the zero-coupon bond value with maturities S and T from Itô's lemma used on $V(t,P) = \log P$, where $P = P(t,T)$ follows the process from the equation (2.13). We need to verify the assumptions for the function V and

$$\begin{aligned} b(t,P) &= r_t\, P, \\ \sigma(t,P) &= -\frac{\sigma}{\alpha}\left(1 - e^{-\alpha(T-t)}\right) P(t,T). \end{aligned}$$

The transformation function V as a logarithmic function of P, $P > 0$, is twice continuously differentiable on $[0,T] \times (0,\infty)$. Then according to Remark 1, it is enough to show that the functions

$$\begin{aligned} A(t) &= r_t, \\ C(t) &= -\frac{\sigma}{\alpha}\left(1 - e^{-\alpha(T-t)}\right) \end{aligned}$$

from the linear SDE are measurable and bounded on $[0,T]$. This is true because the first one is the solution of the Hull-White model from the equation (2.7) and the second one is an exponential function.

Then we can use Itô's lemma and we get differentials of logarithms of zero-coupon bond value with maturities S and T in the form

$$\begin{aligned} d\log P(t,T) &= \left(r_t - \frac{u(t,T)^2}{2}\right) dt + u(t,T) dW_t, \\ d\log P(t,S) &= \left(r_t - \frac{u(t,S)^2}{2}\right) dt + u(t,S) dW_t. \end{aligned}$$

After substition, we complete Step **S2** and get

$$dF(t;T,S) = \frac{1}{S-T}\left[(r_t - \frac{u(t,T)^2}{2} - r_t + \frac{u(t,S)^2}{2})dt + (u(t,T) - u(t,S))dW_t\right],$$
$$= \left(\frac{u(t,S)^2 - u(t,T)^2}{2(S-T)}\right)dt - \left(\frac{u(t,S) - u(t,T)}{S-T}\right)dW_t.$$

Next, we define the zero-coupon bond return (yield) by the forward rate. We rewrite formula (1.12) from Definition 9 with a use of the relation between the zero-coupon bond value $P(t,T)$ and the forward rate in the equation (2.17) for $F(t,t,T)$:

$$R(t,T) = F(0;t,T) + \int_0^t dF(\tau;t,T).$$

We substitute for the differential from the equation (2.18) and we earn the following formula

$$R(t,T) = F(0;t,T) + \int_0^t \frac{u(\tau,T)^2 - u(\tau,t)^2}{2(T-t)}dt - \int_0^t \frac{u(\tau,T) - u(\tau,t)}{T-t}dW_t.$$

Now, by the definitions of r_t, $f(0,t)$ and the derivation, we take a limit

$$\lim_{T\to t^+} R(t,T) = r_t = f(0,t) + \int_0^t \frac{\partial}{\partial t}\frac{u(\tau,t)^2}{2}d\tau - \int_0^t \frac{\partial}{\partial t}u(\tau,t)dW_\tau,$$
$$= f(0,t) + \int_0^t u(\tau,t)\frac{\partial u(\tau,t)}{\partial t}d\tau - \int_0^t \frac{\partial u(\tau,t)}{\partial t}dW_\tau \qquad (2.19)$$

and get a solution to a stochastic process describing the Hull-White model. Step **S3** is thus completed.

By differentiating r_t with respect to t, we reach the stochastic process and after a substitution of $u(t,T)$ from the equation (2.16), we can find the formula for $\theta(t)$. We need to verify that the first derivative of $u(t,T)$ with respect to the second variable satisfies the assumption of Theorem 3. The function $\frac{\partial u(t,T)}{\partial T} = \psi(t,T)$ is a mapping $\psi : (0,\infty) \times (0,\infty) \to \mathbb{R}$ and is obviously continuous in the first variable and continuously differentiable in the second as we can see below:

$$\frac{\partial u(t,T)}{\partial T} = -\sigma e^{-\alpha(T-t)}, \qquad (2.20)$$
$$\frac{\partial^2 u(t,T)}{\partial T^2} = \sigma\alpha e^{-\alpha(T-t)}. \qquad (2.21)$$

Clearly, it also satisfies condition (1.10) for all $T \in (0,\infty)$ because

$$\int_0^T \left[\sigma^2 e^{-2\alpha(T-\tau)} + \sigma^2\alpha^2 e^{-2\alpha(T-\tau)}\right]d\tau =$$
$$= (\sigma^2 + \sigma^2\alpha^2)e^{-2\alpha T}\int_0^T e^{-2\alpha(T-\tau)}d\tau$$

and $e^{-2\alpha(T-\tau)}$ is integrable in the interval $[0, T]$.

Then we can use Theorem 3 to differentiate the Itô's integral in the equation (2.19). For the deterministic part, it is possible to use the deterministic Leibniz rule. Then after differentiation, dr_t looks as

$$dr_t = \left[\frac{\partial f(0,t)}{\partial t} + u(t,t)\frac{\partial u(\tau,t)}{\partial t}|_{\tau=t} + \int_0^t \left[\frac{\partial}{\partial t}(u(\tau,t)\frac{\partial u(\tau,t)}{\partial t})\right] d\tau - \right.$$
$$\left. - \int_0^t \frac{\partial^2 u(\tau,t)}{\partial t^2} dW_\tau \right] dt - \left[\frac{\partial u(\tau,t)}{\partial t}|_{\tau=t}\right] dW_t.$$

We can simplify the symbolic differentiation and use $u(t,t) = 0$, so that

$$dr_t = \left[\frac{\partial f(0,t)}{\partial t} + \int_0^t \left[u(\tau,t)\frac{\partial^2 u(\tau,t)}{\partial t^2} + \frac{\partial u(\tau,t)}{\partial t}^2\right] d\tau - \int_0^t \frac{\partial^2 u(\tau,t)}{\partial t^2} dW_\tau\right] dt -$$
$$- \left[\frac{\partial u(\tau,t)}{\partial t}|_{\tau=t}\right] dW_t. \quad (2.22)$$

The equation (2.22) is prepared for the substitution of derivatives of $u(t,T)$ in equations (2.20) and (2.21) and then from the equation (2.22) we get the following

$$dr_t = \left[\frac{\partial f(0,t)}{\partial t} + \int_0^t \left[-\sigma^2(1 - e^{-\alpha(t-\tau)})e^{-\alpha(t-\tau)} + \sigma^2 e^{-2\alpha(t-\tau)}\right] d\tau - \right.$$
$$\left. - \int_0^t \left(\sigma\alpha e^{-\alpha(t-\tau)}\right) dW_t \right] dt + \sigma dW_t.$$

After simple integration, we have almost a required form of the Hull-White model:

$$dr_t = \left[\frac{\partial f(0,t)}{\partial t} + \frac{\sigma^2}{\alpha}(e^{-\alpha t} - e^{-2\alpha t}) - \alpha\int_0^t \sigma e^{-\alpha(t-\tau)} dW_t\right] dt + \sigma dW_t. \quad (2.23)$$

We can express $\int_0^t \sigma e^{\alpha(t-\tau)} dW_t$ from the equation (2.19) after substituting for function v and its derivative meaning that

$$- \int_0^t \sigma e^{\alpha(t-\tau)} dW_t = f(0,t) + \int_0^t \left(\frac{\sigma}{\alpha}(1 - e^{-\alpha(t-\tau)})\right)\left(\sigma e^{-\alpha(t-\tau)}\right) d\tau - r_t,$$
$$= f(0,t) + \frac{\sigma^2}{\alpha^2}(1 - e^{-\alpha t}) - \frac{\sigma^2}{2\alpha^2}(1 - e^{-2\alpha t}) - r_t.$$

Using this expression in the equation (2.23) we get the following:

$$dr_t = \left[\frac{\partial f(0,t)}{\partial t} + \alpha f(0,t) + \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t}) - \alpha r_t\right] dt + \sigma dW_t.$$

From the comparison of our resulting Hull-White model and the Hull-White model in the form from the equation (2.4) we get the formula for $\theta(t)$ from the equation (2.10) and Step **S5** is completed.

Now with the derived equation of the function $\theta(t)$, we can retrospectively verify the assumptions of Itô's lemma 2 used during the search for a solution of the Hull-White SDE and also used in Step **S1**. We declared the verification of the assumption on the transformation function V and the process $\sigma(t, X_t)$, respectively, the constant function σ above the equation of the solution of the Hull-White model (2.7). It remains to verify the assumptions on the stochastic process $b(t, X_t) = \theta(t) - \alpha r_t$ and the function $\theta(t)$ itself. The process b must be $\mathbb{R}$ - valued (which holds true) and progressively measurable. The function $\theta(t)$ need to be measurable (which holds true) and bounded.

We can see from the form of $\theta(t) - \alpha r(t)$ that it is a progressively measurable stochastic process. Now, it is necessary to verify the boundedness of the function $\theta(t)$ from the equation (2.10). The boundedness of the first and second summands can be said through Definition 14 of the instantaneous forward interest rate by $P(t, T)$ and the relation (1.13). Then we can express the first and second summands as follows

$$\frac{\partial f(0,t)}{\partial t} = t\frac{\partial^2 R(0,t)}{\partial t^2} + 2\frac{\partial R(0,t)}{\partial t},$$

$$\alpha f(0,t) = \alpha t\frac{\partial R(0,t)}{\partial t} + \alpha R(0,t).$$

The yield curve R(0, T) from our data in Figure (1.1) is constructed as a natural cubic spline interpolation function. Spline interpolation is often used as a method to give data points (*knots*) such that it yields an analytical functional form. For the natural cubic spline holds that it is a bounded function on a bounded time interval $[0, T]$ because it is assumed that the cubic spline is a cubic polynomial in each interval between 2 knots and those polynomials are equal at each inner knot up to the second derivative.

This also gives us a twice continuously differentiable function, except for time 0 and T, which is bounded. Composition of bounded functions on the bounded time interval $[0, T]$ such as a linear function and, in this case, a partial derivative of the first and second order leads to the fact that those summands are bounded. We can easily see that the third summand in the equation (2.10) is also bounded on that time interval.

In total, $\theta(t)$ is a composition of bounded functions, therefore, it is a bounded function in the interval $[0, T]$ and fulfills the assumption of Itô's lemma. Now, it is shown that we can use Itô's lemma to solve the Hull-White linear SDE.

□

2.3 Parameter estimation of the Hull-White model

When we have the given formula for the Hull-White model and the derived formula for the function $\theta(t)$ which depends on two unknown parameters α and σ, we can start with a process of estimation of those parameters. After we estimate the values of α and σ, the model is determined up to the given initial term structure of the interest rates. We describe two approaches to parameter estimation and later show them and their results on our data in Chapter 3.

2.3.1 Forward-looking approach

One of approaches is the forward-looking approach which uses prices of interest rate derivatives (mostly caps/floors or swaptions) with respect to a future maturity and paying date discounted to the present. We take advantage of the market expectation about the future and estimate the parameters of the Hull-White model accordingly. On the other hand, we know nothing about what happened on the market historically. This subsection is based on the calibration methods in Russo and Torri [2019] and Gurrieri et al. [2009].

Generally, a model calibration is a non-convex optimization problem when we minimize an objective function D

$$\arg\min_{\boldsymbol{x}} D(\boldsymbol{p}(\boldsymbol{x}), \boldsymbol{p^M}) \tag{2.24}$$

dependent on the theoretical prices $\boldsymbol{p}$ and the market prices $\boldsymbol{p^M}$ of the N interest rate derivatives with respect to a vector of the parameters $\boldsymbol{x} = (\alpha, \sigma)^\top$. We try to calibrate the theoretical price given by the Hull-White model to the market expectation. It is clear that the function D should be based on some kind of a difference of the prices p and p^M.

The first of possible approaches is to calibrate both parameters simultaneously. We give examples of the objective function D. We minimize a sum of squares in principle in all 3 examples.

Example. A basic example of the function D is the sum of the distances between the theoretical and market prices of the N interest rate derivatives. We do not want to explore a sum of absolute positive distances, but only a sum of squares of distances

$$D(\boldsymbol{p}(\boldsymbol{x}), \boldsymbol{p^M}) = \sum_{i=1}^{N} \left(p_i(\boldsymbol{x}) - p_i^M \right)^2. \tag{2.25}$$

Example. A different method is to choose a relative difference of theoretical and market prices of the N interest rate derivatives, and we use it inside a sum of squares:

$$D(\boldsymbol{p}(\boldsymbol{x}), \boldsymbol{p^M}) = \sum_{i=1}^{N} \left(\frac{p_i(\boldsymbol{x}) - p_i^M}{p_i^M} \right)^2. \tag{2.26}$$

Example. We can also modify the previous example taking a square root of the equation (2.27):

$$D(\boldsymbol{p}(\boldsymbol{x}), \boldsymbol{p^M}) = \sqrt{\sum_{i=1}^{N} \left(\frac{p_i(\boldsymbol{x}) - p_i^M}{p_i^M} \right)^2}. \tag{2.27}$$

There is another possibility of how to estimate the Hull-White model using swaptions. It consists of 2 steps. First, we calibrate α independently on the volatility parameter by a minimization of difference of a market implied volatilities for swaptions and approximative swap market model implied volatilities from Definition 19. The definition is based on Brigo and Mercurio [2006]. Then in the second step, we calibrate σ by minimizing the objective function D with a known α.

Definition 19 (Swap market model)**.** *Let T_1 be expiry time, $\{T_2, \ldots, T_k\}$ be a set of the prespecified payment dates and τ_i for $i \in \{2, \ldots, k\}$ be a tenor. Then a swap market model (SMM) describing the dynamics of forward swap rates*

$$S_{1,k}(t) = \frac{P(t, T_1) - P(t, T_k)}{\sum_{i=2}^{k} \tau_i P(t, T_i)} \tag{2.28}$$

is given by an SDE

$$dS_{1,k}(t) = \sigma_{1,k}(t) S_{1,k}(t) dW_t^{1,k}, \tag{2.29}$$

where a deterministic function $\sigma_{1,k}(t)$ is an instantaneous percentage volatility of the forward swap rate and $dW_t^{1,k}$ is the Wiener process under the forward swap rate measure.

Remark 6. The *forward-swap rate measure* is a measure associated with a specific numeraire given by a denominator $\sum_{i=2}^{k} \tau_i P(t, T_i)$ in the equation (2.28). Also, it is important to note that the forward swap rates $S_{1,k}(t)$ follow the log-normal distribution.

We use SMM to approximate the Hull-White model for the purpose of deriving theoretical implied volatilities. Forward swap rates implied by the Hull-White model do not follow the log-normal distribution, but we can find that $\frac{P(t,T_i)}{P(t,T_{i+1})}$ with some fixing time T_i and paying time T_{i+1} follows log-normal distribution.

We can see it from the fact that $P(t, T)$ in the Hull-White model has the affine structure from the equation (1.14) and thus the ratio of zero-coupon bond prices is an exponential of a normal variable r_t. From that we can approximate implied volatility for swaptions in the Hull-White model with a bond ratio and it leads into formula

$$v_{sw}(T_1, T_k) = \left(\frac{P(0, T_1)}{P(0, T_1) - P(0, T_k)} \right)^2 v_p(0, T_1, T_k), \tag{2.30}$$

where $v_p(0, T_1, T_k)$ is an implied volatility for the bond ratio given by a formula

$$v_p(0, T_1, T_k) = \mathrm{var}[r_{T_1} | \mathcal{F}_0] \Gamma(T_1, T_k)^2.$$

Now we can describe what the calibration of the model looks like. To reach a variable independent of σ, we need to take a ratio of implied swaptions volatilities with the same expiry and different maturities. The reason is that the implied swaption volatility in the equation (2.30) depends only on σ in $\mathrm{var}[r_{T_1}|\mathcal{F}_0]$ which is the same for different payment times (k,1) and (k,2). Thus, the ratio is independent of σ:

$$\frac{v_{sw}(T_1, T_{k,1})}{v_{sw}(T_1, T_{k,2})} = \left[\frac{\big(P(0,T_1) - P(0,T_{k,2})\big)\Gamma(T_1, T_{k,1})}{\big(P(0,T_1) - P(0,T_{k,1})\big)\Gamma(T_1, T_{k,2})}\right]^2.$$

Then it is easy to formulate the objective function of this optimization problem with respect to the parameter α only with n_m expiry times and n_p maturities according to Example (2.3.1)

$$D(\alpha) = \sum_{1\leq i\leq n_m, 1\leq j\leq n_p} \left(\sqrt{\frac{v_{sw}(T_{1,i}, T_{k,j+1})}{v_{sw}(T_{1,i}, T_{k,j})}(\alpha)} - \frac{IV_{i,j+1}}{IV_{i,j}}\right)^2, \tag{2.31}$$

where we consider the ratio of implied swaptions volatilities as a function of the single parameter α and $IV_{i,j}$ is the implied market volatility for the swaption with expiry time $T_{1,i}$ and payment time $T_{k,j}$. After that, we minimize the objective function only with respect to the second parameter σ

$$\arg\min_{\boldsymbol{x}} D(\boldsymbol{p}(\sigma), \boldsymbol{p}^{\boldsymbol{M}}).$$

2.3.2 Backward-looking approach

A different way to estimate parameters α and σ is to use historical data from an interest rate time structure given by the spot rates $R(t,T)$. We want to minimize the sum of squares of the differences of the theoretical and observed variance of daily spot rate changes. This estimation was shown in Sivertsen [2016]. Firstly, we need to describe the dynamics of spot rates implied by the Hull-White model.

We get an SDE

$$dR(t,T) = \frac{\frac{1}{2}\sigma^2\Gamma(t,T)^2 - r_t}{T-t}dt + \frac{\sigma\Gamma(t,T)}{T-t}dW_t,$$

where a function Γ is given by the equation (2.11).

We work with spot rate changes in discrete historical data, hence, it is natural to discretize the SDE into

$$\Delta R = R(t_{k+1}, T) - R(t_k, T) = \frac{\frac{1}{2}\sigma^2\Gamma(t_k,T)^2 - r_{t_k}}{T-t_k}\Delta t + \frac{\sigma\Gamma(t_k,T)}{T-t_k}\sqrt{\Delta t} w_{k+1},$$

where it is taken in discrete times $t_0, t_1, \ldots, t_p$, discrete version of dt is $\Delta\, t = t_{k+1} - t_k$ and w_{k+1} is a standard normal random variable with zero mean and unit variance.

Then it is obvious that the theoretical variance for a spot rate change with unit variance of w_{k+1} is

$$\sigma^t_{\Delta R}(t_k) = \frac{\sigma\Gamma(t_k, T)}{T - t_k}\sqrt{\Delta t} = \frac{\sigma(1 - e^{-\alpha(T-t_k)})}{\alpha(T - t_k)}\sqrt{\Delta t}. \tag{2.32}$$

We take a sample variance of daily spot rate changes from historical observations $\Delta R_1, \ldots, \Delta R_N$ for time t_k as observed variance:

$$\sigma^o_{\Delta R}(t_k) = \frac{1}{N-1}\sum_{j=1}^{N}\Big(\Delta R_j - \mathsf{E}\left[\Delta R\right]\Big)^2 \tag{2.33}$$

with a sample mean

$$\mathsf{E}\left[\Delta R\right] = \frac{1}{N}\sum_{j=1}^{N}\Delta R_j,$$

where $\Delta R_j = R(t_j, T) - R(t_{j-1}, T)$.

Now we can show that estimated parameters are a solution to a minimization problem with a recommended assumption $\Delta t = \frac{1}{251}$ for 251 trading days with equations (2.32) and (2.33):

$$\arg\min_{\alpha,\sigma}\sum_{k=1}^{p}\left(\frac{\sigma(1 - e^{-\alpha(T-t_k)})}{\alpha(T - t_k)}\frac{1}{\sqrt{12}} - \sigma^o_{\Delta R}(t_k)\right)^2,$$

where p is the number of historical yield curves with a different tenor of the spot rate.

Chapter 3

Simulation study

We described the theoretical background of the Hull-White model and its estimation. Now we show this process practically on given data from Česká spořitelna. We set prices of caplets, caps and swaptions and then calibrate the Hull-White model based on the corresponding methods in Subsection (2.3.1). Then, we use a set of historical yield curves to estimate the model based on the methods in Subsection (2.3.2). After that, simulations of estimated versions of the Hull-White model are performed.

The estimated Hull-White model is used for several needs in Česká spořitelna. One of them is to set the interest rates for a mortgage. The original purpose is to price exotic non-standardized interest rate derivatives without any analytical pricing formula. In this situation, it is useful to simulate the underlying interest rate with the help of the estimated Hull-White model and then derive the price of that specific interest rate derivative from the simulated values for a given time characteristic (maturity or expiry). It depends on the definition of the derivative.

All of the following calculations, data edits, and optimization problems solutions via minimizing objective functions are performed in Python software. Simulations of the estimated Hull-White model and their illustrations are also created in Python. Some easier graphs are made in Wolfram Mathematica software. The code with explanatory comments is provided in the attachement of this thesis.

3.1 Data describtion

Our input data given by Česká spořitelna contains a present yield curve, historical yield curves, and different volatility curves for each financial derivative. The present yield curve was displayed in Figure (1.1) as data points interpolated by a natural cubic spline. We have yields forward to 30 years and it was said that it is the inverted yield curve.

It describes 6 month interest rates for Czech currency with continuous compounding and the ACT/365 day count convention. It means that we take a real difference between two dates counted in days, and the year is supposed to have 365 days. Definition 11 also describes this curve as a zero-coupon curve. It is a reference on a zero-coupon bond which pays no coupon payments (no interest) in a lifetime of bond, and an investor receives a face value at bond's maturity.

On the same basis without interpolation, we have historical daily yields data from 31 January 2018 to 31 January 2023 for various tenors of interest rates. We show only selected yield curves because of the large amount of data that we have to historically estimate parameters of the Hull-White model. They are shown in Figure (3.1). We can see the increasing trend of yields in the last 5 years.

In 2018, yields were as expected lower for shorter interest rates and higher for longer interest rates, which resulted from added different premiums for increasing risk in a longer time period. This changed in the last two years for many economical reasons, and now we generally have higher yields.

Moreover, shorter interest rates have higher yields than longer ones. It may be partly caused by a slower reaction of shorter interest rates on what the present yield curve expects (decreasing trend) in future years. It can be seen in the whole graph that longer interest rates are more reactive to expected trends.

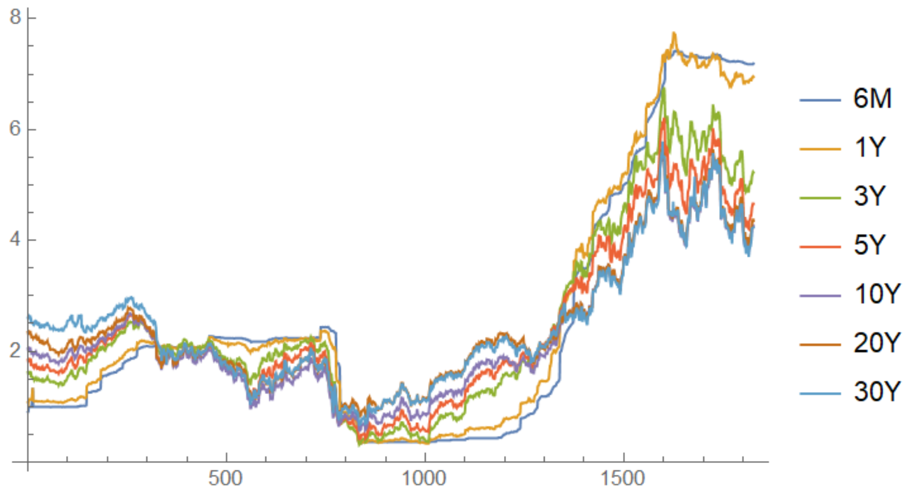

Figure 3.1: Historical yield curves starting at time $t = 0$ (in days) and ending at time $t = 1826$ (in days) for seven different tenors of the interest rates and the displays yields in %.

We present volatility curves for caplets on two separated graphs in Figure (3.2) and (3.3). The provided caplet dataset contains shifted strikes corresponding to the market valuation which will be introduced later. It is done for a clearer visualisation with a lot of data points which are connected only with lines for a data presentation. The volatility curves for given strikes have different values and shapes for the shortest maturities. The reason is that higher strikes can be far from the real value of a caplet. The value of such caplets is then zero in a majority of cases for a short period of time, and the caplets have less volatile prices.

We can see a mostly decreasing trend of volatilities for maturities up to 8 years, except for three volatility curves for the highest strikes in Figure (3.3) and then quite a consistent behaviour of all volatility curves around 0,18. With longer maturity, the distance between a strike price and a real price becomes less relevant and that is why caplets with different strikes have similarly volatile prices. Volatility curves are constructed for strikes shifted by 3 %. These volatility curves and the present yield curve are meant to be used for the forward-looking approach.

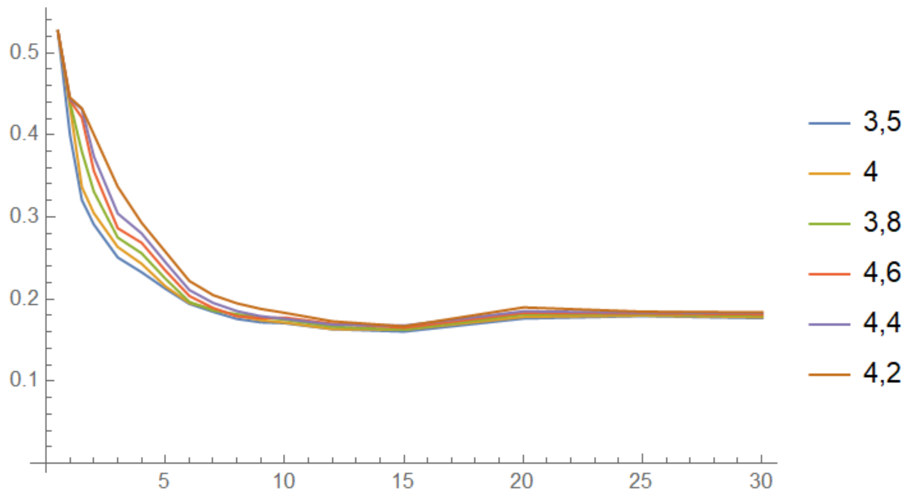

Figure 3.2: Volatility curves starting at the first maturity date 31 July 2023 at time $t = 0{,}5$ (in years) and ending at the last maturity date at time $t = 30$ (in years) for six different strikes (in %) shifted by 3 % with the same initial volatility.

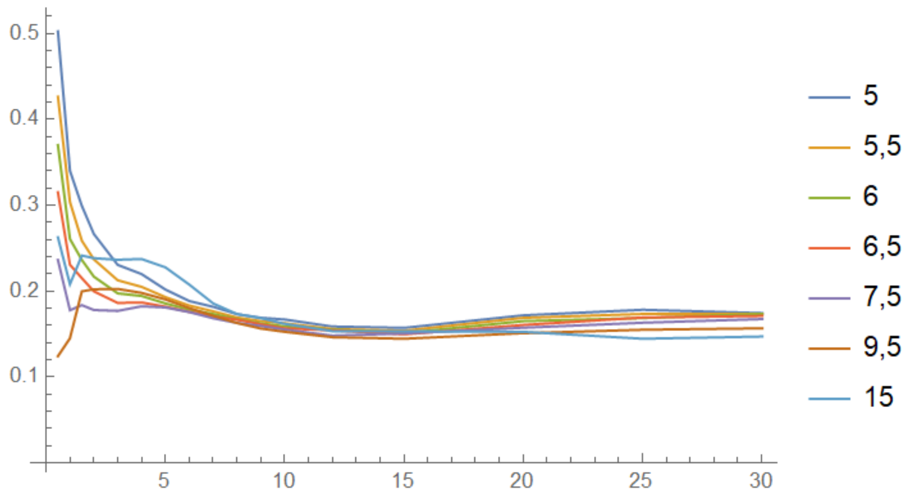

Figure 3.3: Volatility curves starting at the first maturity date 31 July 2023 at time $t = 0{,}5$ (in years) and ending at the last maturity date at time $t = 30$ (in years) for seven different strikes (in %) shifted by 3 % with the different initial volatilities than the strikes in Figure (3.2).

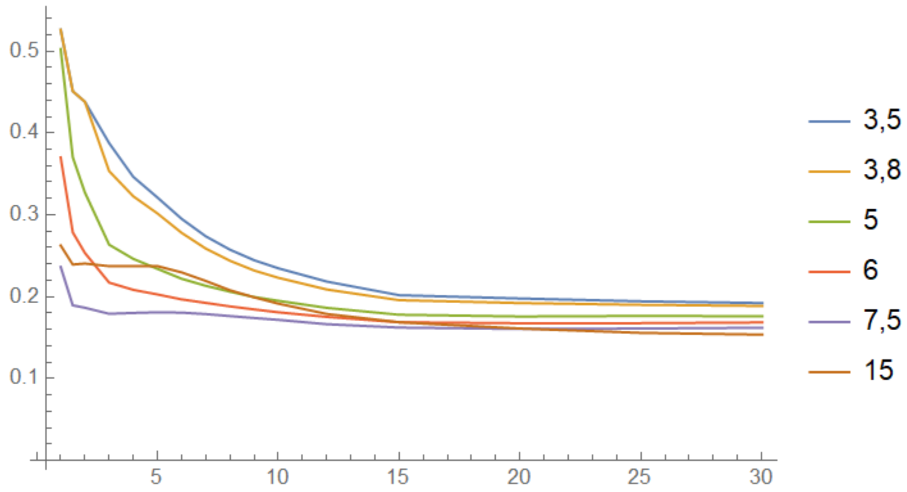

Figure 3.4: Volatility curves starting at the first maturity date 31 January 2024 at time $t = 1$ (in years) and ending at the last maturity date at time $t = 30$ (in years) for six different selected strikes (in %) shifted by 3 %.

Cap volatility curves display similar behaviour because caps consist of caplets. Only for illustration, we show volatility curves for the selected shifted strike prices in Figure (3.4), including the lowest and highest strike price. There are greater differences in volatilites for longer maturities, but it is still consistent behavior between 0,18 and 0,24. In the case of caps, it is better to see that a cap with higher strike price has less volatile price even for longer maturities. We need to point out that due to the fact that a cap consists of at least two caplets (in our case at least two half a year caplets) volatility curves start on 31 January 2024.

Swaptions with their different structure in comparison to caps have volatilities for a given expiry time and for a given tenor at date 31 January 2023. We plot the volatilities against expiry times for different tenors in Figure (3.5). There are some interesting points to comment on. Swaptions with shorter tenors have less volatile prices than those with longer tenors for expiry time up to 1 year. On the other side swaptions with longer tenors have rather less volatile prices than the ones with shorter tenors for expiry time after 10 or 15 years.

Inside the swaption dataset, there are strong dependent relations. When we look at the sample correlation coefficients between different tenors or expiries from historical volatility quotations, we can see that neighboring tenors for the same expiry or neighboring expiries for the same tenor are strongly positively correlated. This relationship weakens slowly. The sample correlation coefficients show that even relatively time-distanced tenors or expiries are still positively correlated. The shortest and the longest tenors or expiries can be said to be weakly correlated.

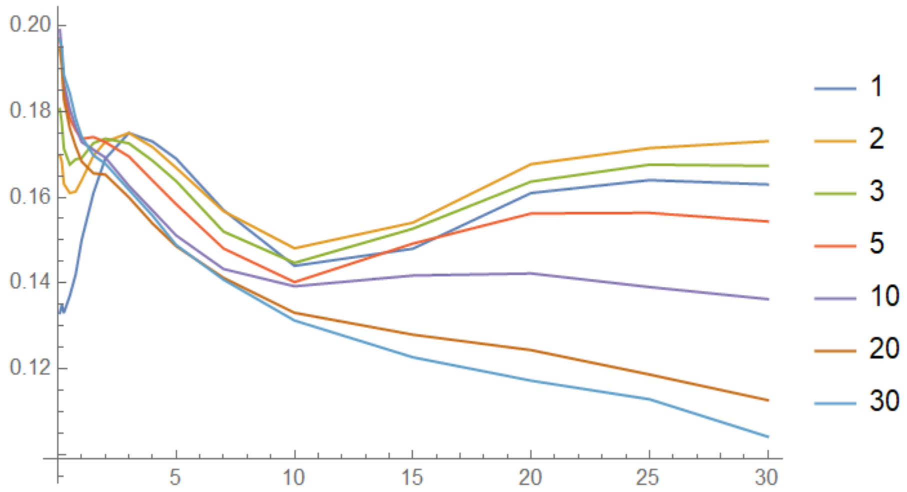

Figure 3.5: Volatilities on 31 July 2023 plotted against expiry times from 1 month to 30 years (in years) for seven selected different tenors (in years) including the shortest and the longest tenor.

3.2 Results for the forward-looking approach

In this section, we perform step by step an estimate of the Hull-White model according to the forward-looking approach of an estimate of the model. We set the prices of the respective financial derivatives for a model calibration in a market view and in the Hull-White model. Then we can use different objective functions (see examples in Section 2.3.1) for the calibration process and try to compare the estimated parameters. Finally, we show the estimated simulated Hull-White models.

3.2.1 Pricing of financial derivatives

The principle of model calibration is to compare and minimize the difference between a market price and a price given by the Hull-White model. For this purpose, we suppose that pricing with the Black-Scholes model is consistent with the financial market. Consequently to the input data with volatilities modelled for shifted strikes, we use the shifted Black-Scholes model where we shift forward rates and strike prices and Φ as a distribution function of normal distribution with zero mean and unit variance is used.

Firstly, we show the pricing formula for the shifted Black-Scholes model for caplets discounted to time 0 according to Brigo and Mercurio [2006]:

$$p^{BS}_{caplet}(0,N,T_1,T_2,K,\sigma_{BS},s) = N\,P(0,T_2)\,\tau[(F(0;T_1,T_2)+s)\Phi(d_1)-(K+s)\Phi(d_2)], \tag{3.1}$$

where N is the nominal value taken for our work as $N=1$, T_1 and T_2 are the fixing and maturity dates, $P(0,T_2)$ a zero-coupon bond price is taken as a continuous discounting factor from the maturity date to beginning at 0 given by the equation (1.13), τ is a tenor as a difference $T_2 - T_1$, σ_{BS} is a volatility from the volatility curve for a given strike price K as the quoted volatility in the market, s is a shift for the shifted Black-Scholes model taken as $s = 3\%$, $F(0; T_1, T_2)$ is a forward rate given by the equation (1.15), and d_1 and d_2 are given by the formulas

$$d_1 = \frac{\log\left(\frac{F(0;T_1,T_2)+s}{K+s}\right) + \frac{\sigma_{BS}^2 T_1}{2}}{\sigma_{BS}\sqrt{T_1}},$$
$$d_2 = d_1 - \sigma_{BS}\sqrt{T_1}.$$

Then it is easy to show how to use the price of a caplet to set a price of a cap according to Definition 15. Caps consists of caplets with the same strike price with different fixing and maturity dates, hence we can price caps in the following way:

$$p_{cap}^{BS} = \sum_{i=2}^{k} p_{caplet}^{BS}(0, 1, T_{i-1}, T_i, K, \sigma_{BS}, 3\%), \tag{3.2}$$

where the indices of the first and the last summands are consistent with the set of prescribed dates in Definition 15.

We set the tenor for each caplet as $\tau = 0{,}5$ which means that a lifetime of a caplet from a fixing date to a maturity date is half a year. Our first prescribed date T_1 is taken as 0 according to our yield curve from Figure (1.1) and, analogously, we take $T_k = 30$ as the last maturity date. Hence we have the set of prescribed dates $\{0, 0{,}5, 1, 1{,}5, \dots, 29{,}5, 30\}$.

Next, we introduce similar formulas for caplets and caps in the sense of the Hull-White model. We express prices also according to Brigo and Mercurio [2006] where the fact that a caplet is equivalent to a European put option on zero-coupon bond with modified nominal value and strike price is used. Then, a fixing date of the caplet is a maturity date of the put option, and a maturity date of the caplet is a maturity date of the zero-coupon bond. We start to show the pricing formula for caplet discounted to time 0:

$$p_{caplet}^{HW}(0, N, T_1, T_2, K) = N\left[P(0,T_1)\Phi(-h+\sigma_p) - (1+K\,\tau)P(0,T_2)\Phi(-h)\right], \tag{3.3}$$

where a nominal value N is consistent with the market setup $N = 1$, T_1 and T_2 are the fixing and maturity dates, $P(0, T_1)$ a zero-coupon bond price is taken as a continuous discounting factor from the maturity date of the put option (from the fixing date of the caplet) to the starting date at 0 given by the equation (1.13), $P(0, T_2)$ a zero-coupon bond price determining the put option, thus determining the caplet. Next, τ is a tenor as a difference $T_2 - T_1$, also taken as $\tau = 0{,}5$, K is the strike price of the caplet, $1 + K\,\tau$ is the modified strike price of the put option. It remains to set h and σ_p which are given by the following formulas:

$$h = \frac{1}{\sigma_p}\log\left(\frac{P(0,T_2)(1+K\,\tau)}{P(0,T_1)}\right) + \frac{\sigma_p}{2},$$
$$\sigma_p = \frac{\sigma}{\alpha}\left(1 - e^{-\alpha(T_2 - T_1)}\right)\sqrt{\frac{1 - e^{-2\alpha T_1}}{2\alpha}},$$

where α and σ are parameters of the Hull-White model.

After that we can write down a pricing formula for caps which (as in case of the Black-Scholes model) consists of a sum of caplet prices for different fixing and maturity dates but with the same strike price:

$$p_{cap}^{HW} = \sum_{i=2}^{k} p_{caplet}^{HW}(0, 1, T_{i-1}, T_i, K), \tag{3.4}$$

where the indices of the first and the last summands are consistent with the set of prescribed dates in Definition 15.

It is important to describe the situation when there is a possibility of no mean reversion effect according to the model calibration process. This is expressed by the Ho-Lee model from the example model in the formula (2.5). This situation is given by $\alpha = 0$. Only σ_p is dependent on the parameter α in pricing caplets or caps, thus, we take a limit

$$\lim_{\alpha \to 0} \sigma_p(\alpha, \sigma, T_1, T_2) = \sigma(T_2 - T_1)\sqrt{T_1} := \sigma_p^{HL}.$$

Last, we set prices for swaptions. Setting a swaption price is not as straightforward as in the case of caps because we do not have individual financial derivative at each prescribed date. According to the Black-Scholes model, we set market price of a payer swaption discounted to time 0 from Brigo and Mercurio [2006]:

$$p_{swaption}^{BS}(0, N, \mathcal{T}, K, \sigma_{1,k}) = N\left(S_{1,k}(0)\,\Phi(d_1) - K\,\Phi(d_2)\right)\tau \sum_{i=2}^{k} P(0, T_i), \tag{3.5}$$

where a nominal value $N = 1$, $\mathcal{T}$ is a set of an expiry date T_1 and exchange payment dates $\{T_2, \ldots, T_k\}$, $\tau = 0{,}5$ is a fraction of a year between the exchange payment dates, a strike K is taken as a forward swap rate (at-the money) at time 0 $S_{1,k}(0)$ for a swaption with an expiry date T_1, a tenor $T_k - T_1$, and with a quoted swaption volatility $\sigma_{1,k}$ given by the formula

$$S_{1,k}(0) = \frac{P(0, T_1) - P(0, T_k)}{\sum_{i=2}^{k} P(0, T_i)}.$$

A zero-coupon bond price $P(0, T_i)$, $i \in \{1, \ldots, k\}$ is taken as a continuous discounting factor from time T_i to the valuation time 0 given by the equation (1.13). We need to take the sum of zero-coupon bond prices because of our decision about the option on a whole swap made at an expiry date. Numbers d_1 and d_2 are given by formulas

$$d_1 = \frac{\log\left(\frac{S_{1,k}(0)}{K}\right) + \frac{\sigma_{1,k}^2 T_1}{2}}{\sigma_{1,k}\sqrt{T_1}},$$
$$d_2 = d_1 - \sigma_{1,k}\sqrt{T_1}.$$

Remark 7. With respect to our choice of the strike K, our formula for pricing swaptions in the Black-Scholes model gets easier:

$$p^{BS}_{swaption}(0,1,\mathcal{T},S_{1,k}(0),\sigma_{1,k}) = (P(0,T_1) - P(0,T_k))\,(\Phi(d_1) - \Phi(d_2)),$$
$$d_1 = \frac{\sigma_{1,k}\sqrt{T_1}}{2},$$
$$d_2 = -d_1.$$

It remains to set the price of a swaption in the Hull-White model according to Brigo and Mercurio [2006]:

$$p^{HW}_{swaption}(0,N,\mathcal{T},K,\mathcal{X}) = N\sum_{i=2}^{k}\left[c_i\left(K_i\,P(0,T_1)\,\Phi(-h+\sigma_p) - P(0,T_i)\,\Phi(-h)\right)\right], \tag{3.6}$$

where a nominal value $N = 1$, $\mathcal{T}$ is a set of only exchange payment dates, a strike K and a zero-coupon bond price $P(0,T_i)$ are taken in the same way as in the equation (3.5), $c_i = K\,\tau$, $i = 2,\dots,k-1$, $c_k = 1 + K\,\tau$, in both cases $\tau = 0{,}5$ is a fraction of a year between exchange payment dates, $\mathcal{X}$ is a set of bond prices (discounting factors to expiry date T_1) with maturity at some exchange payment date T_i with given a interest rate r^* such that

$$\sum_{i=2}^{k} c_i\,B(T_1,T_i)\,e^{-\Gamma(T_1,T_i)\,r^*} = 1.$$

We can see that the discounting factors X_i are implied by the affine structure:

$$X_i = B(T_1,T_i)\,e^{-\Gamma(T_1,T_i)\,r^*},\ \ i = 2,\dots,k.$$

There are still variables h and σ_p given by the formulas

$$\sigma_p = \sigma\sqrt{\frac{1-e^{-2\alpha\,T_1}}{2\alpha}}\,\Gamma(T_1,T_i),$$
$$h = \frac{1}{\sigma_p}\log\frac{P(0,T_i)}{P(0,T_1)} + \frac{\sigma_p}{2}.$$

3.2.2 Model calibration

In this subsection, we apply and describe software methods to solve a non-convex optimization problem as a model calibration by the formula (2.24) where $\boldsymbol{p}^M = \boldsymbol{p}^{BS}$ and $\boldsymbol{p}(\boldsymbol{x}) = \boldsymbol{p}^{HW}((\alpha,\sigma)^\top)$ with the i-th price for financial derivative with a set of prescribed dates $\mathcal{T}_i$, a strike K_i and for market prices with volatility σ_i, $i = 1,\dots,I$, where I is a number of caplets, caps, or swaptions.

We use the objective function from the equation (2.25) consistent with the methodology in Česká spořitelna. Estimation of α in swaption calibration is done by minimizing the objective function from the equation (2.31). In all procedures, we set bounds for estimates. We minimize the objective function on the space $(0.01,5)\times(0.0001,0.5)$ (the first interval is for α, the second for σ).

We can implement different tactics according to the functions the software provides us. We have two functions in Python, *differential_evolution* and *curve_fit*, which offer different optimization procedures. The advantage of *differential_evolution* is that we can set a required objective function and it does not put any requirements on our optimization problem but it can be a less intuitive way.

On the other side, the function *curve_fit* only works with predefined objective function from the equation (2.25) but we can set initial parameter guesses with estimates from *differential_evolution* and thanks to the nonlinear least squares estimator the estimation can be more straightforward. The main disadvantage of *curve_fit* lies in the following important assumption.

According to Jennrich [1969], the nonlinear least squares estimator works under the assumption that a vector of market prices $\boldsymbol{p}^{BS}$ has the structure

$$\boldsymbol{p}^{BS} = \boldsymbol{p}^{HW}((\alpha, \sigma)^\top) + \boldsymbol{\epsilon}, \tag{3.7}$$

where $\boldsymbol{p}^{HW}$ is the vector of known continuous nonlinear functions of the parameters α and σ and $\boldsymbol{\epsilon}$ is the vector of independent identically distributed errors (residuals) with zero mean and finite positive constant variance.

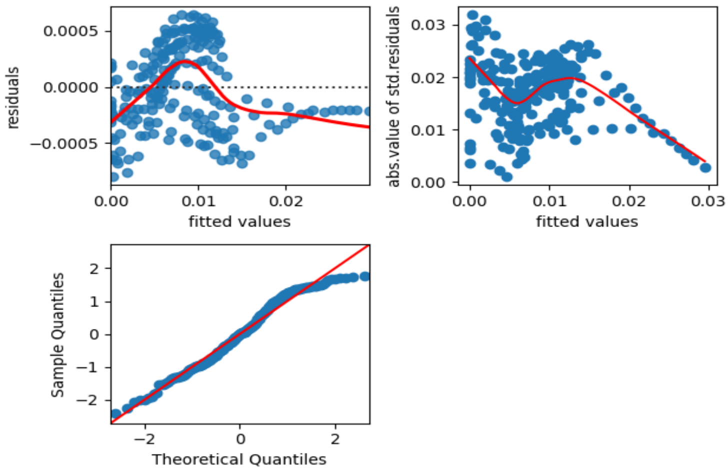

Figure 3.6: Diagnostical plots to verify the assumption of iid residuals with zero mean and constant finite variance from the vector $\boldsymbol{\epsilon}$ for caplets. There are plotted residuals and absolute value of standartized residuals against fitted values and QQ-plot of sample quantiles against theoretical quantiles of standartized normal distribution.

We can look at the diagnostic plots for caplets in Figure (3.6). We plot residuals against fitted values, and we want to analyse the red smoother. If the residuals have zero mean, the red curve should be close to zero. When we omit spare right-tail data, the assumption can be taken as fulfilled. Then we plot absolute values of standartized residuals against fitted value which can be diagnostic plot for constant variance when the red smoother is constant. Even with the omitted spare right-tail data, there is some uncertainty.

But the last diagnostic QQ-plot shows that sample quantiles of residuals correspond (except for right tail) to the theoretical quantiles of the standardized normal distribution. The result is that assumptions on residuals for the nonlinear least square estimator can hold true for caplets.

However, the diagnostic plots for the caps in Figure (3.7) lead to different results. The first plot indicates that the mean value can be shifted up from 0, the second one mainly shows that there can be heteroskedasticity in the residuals and the assumptions on the residuals are violated. In the case of swaptions, the assumptions are also violated. That means that the calibration tactic with the initial guesses from *differential_evolution* for *curve_fit* minimization can be used only for caplets.

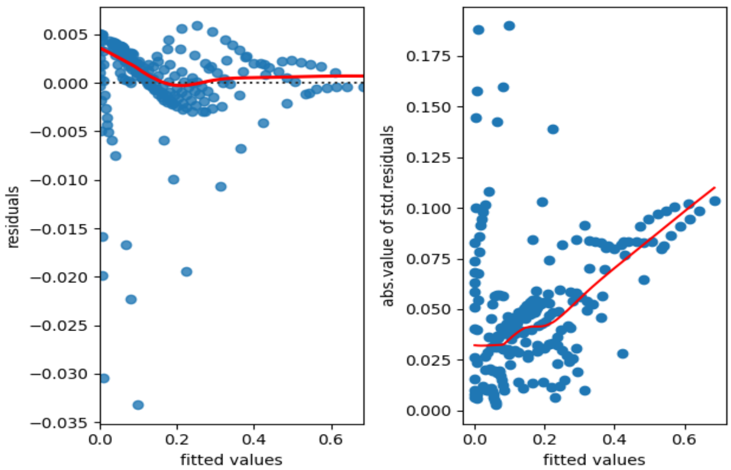

Figure 3.7: Diagnostical plots to verify the assumption of iid residuals with zero mean and constant finite variance from the vector $\boldsymbol{\epsilon}$ for caplets. There are plotted residuals and absolute value of standartized residuals against fitted values.

In this work, we perform two possible estimation procedures. The first one involves estimation of the parameters α and σ through *differential_evolution*. The second one uses estimates from *differential_evolution* as initial guesses for *curve_fit*, which enables us to estimate parameters only under assumption described below the equation (3.7). Then we can compare both ways with respect to SSE (sum of squared errors or residuals). In the case where we cannot fulfill the assumption of nonlinear least squares estimate method, we use the value from *differential_evolution* as the final estimate of the parameter.

Remark 8. In the case of *differential_evolution*, we can compare the quality of estimates through the value of an objective function (OF). We need to reach the lowest possible value. When we use *curve_fit*, the natural way to compare estimates is with respect to SSE. Again, the estimates with lower SSE are better. Specifically, for the objective function from the equation (2.25), the value of the objective function at estimated parameters from *differential_evolution* is the same as the SSE from *curve_fit*.

We briefly introduce both functions. First, we describe a principle of *dif-*

ferential_evolution according to Qiang et al. [2016]. It is an implemenation of a stochastic based evolution algorithm for global optimization. The algorithm starts with a population initialization of S solutions, which can be done by a generation from a uniform distribution on a D dimensional parameter space even without any given information about the solution of the optimization process. In general, it is performed in three steps.

Each mutation step is about to create S mutant vectors $\boldsymbol{v}_i$ from existing members $\boldsymbol{x}_i$, $i \in \{i = 1, 2, \dots, S\}$ of a current generation. Mutation is provided by the so-called *best1bin* strategy in the function *differential_evolution* which is given by the formula

$$\boldsymbol{v}_i = \boldsymbol{x}_b + M(\boldsymbol{x}_{r1} - \boldsymbol{x}_{r2}), \tag{3.8}$$

where $\boldsymbol{x}_b$ is the best solution from the current generation, $r1$ and $r2$ are random integers from the interval $[1, S]$ and are not equal to index i and $M \in [0{,}5, 1]$ is a mutation weighting factor.

The next crossover step combines each $\boldsymbol{v}_i$ and $\boldsymbol{x}_i$ into a new trial vector $\boldsymbol{U}_i = (u_{i,1}, \dots, u_{i,D})^\top$, $i \in \{i = 1, 2, \dots, S\}$ in the following way

$$\begin{aligned} u_{i,j} &= v_{i,j}, \text{ if } rand_j \leq CR \text{ or } j = mbr_j, \\ &= x_{i,j}, \text{ otherwise,} \end{aligned}$$

where CR is crossover probability and represents a benchmark for selection of mutant vectors and existing members, $rand_j$ is a randomly chosen real number from $[0, 1]$ for each $u_{i,j}$ and mbr_i is a random chosen integer from $[1, D]$ for the whole $\boldsymbol{U}_i$.

The last is the selection step. Members of the new generation are selected on the basis of a boundary check and a comparison with the member of the current generation. If the trial solution $\boldsymbol{U}_i$ is outside the boundary of the parameter space, the new trial vector is created within the parameter space. We select $\boldsymbol{U}_i$ for the new generation if it has a lower value of the objective function than $\boldsymbol{x}_i$, otherwise we stay with $\boldsymbol{x}_i$. The algorithm ends after reaching a convergence condition or after a given maximum number of iterations.

In the next part, we describe how the function *curve_fit* works for our data. According to the documentation of the SciPy package in Python software, *subspace trust region interior reflective (STIR)* method of nonlinear least squares is used to minimize our optimization problem

$$\min_{\boldsymbol{x}} f(\boldsymbol{x}) : \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u}, \tag{3.9}$$

where $\boldsymbol{x} = (\alpha, \sigma)^\top$ is the vector of model parameters and $\boldsymbol{l}$ and $\boldsymbol{u}$ are vectors of lower and upper bounds for parameters.

The STIR method generally formulated by Branch et al. [1999] is an iterative method with given steps. First, our chosen objective function $f(x_k)$ is approximated by the quadratic function

$$\kappa_k(\boldsymbol{s}) = \boldsymbol{g}_k^\top \boldsymbol{s} + \frac{1}{2}\boldsymbol{s}^\top (H_k + C_k)\boldsymbol{s},$$

where $\boldsymbol{s}$ is a step that can improve a solution of the optimization problem from the kth iteration, $\boldsymbol{g}_k$ and H_k are gradient and Hessian matrix for the objective function in the kth iteration and C_k is a diagonal scaling matrix in the kth iteration.

This quadratic function enables to define a trust region subproblem which is solved with respect to step s to determine possible improvement of an optimal solution:

$$\min_{\boldsymbol{s}} \kappa_k(\boldsymbol{s}) : ||D_k\boldsymbol{s}|| \leq \delta_k, \boldsymbol{s} \in S_k,$$

where D_k is a positive diagonal scaling matrix, δ_k is a positive scalar, $||.||$ is the Euclidean norm, and S_k is a given two-dimensional subspace in $\mathbb{R}^n$, all for the k-th iteration. We then compare the values of the objective function $f(\boldsymbol{x}_k + \boldsymbol{s}_k)$ and $f(\boldsymbol{x}_k)$. If we improve the value of the objective function in the process of minimization, we take

$$\boldsymbol{x}_{k+1} = \boldsymbol{x}_k + \boldsymbol{s}_k \in \{\boldsymbol{x} : \boldsymbol{l} \leq \boldsymbol{x} \leq \boldsymbol{u}\},$$

otherwise, we stay with $\boldsymbol{x}_k$ for the next iteration. The last step is to decide δ_{k+1} based on the comparison of the values of the objective function. This iterative algorithm in the function *curve_fit* ends after satisfying the first-order optimality condition given by a drop of $\kappa_k(\boldsymbol{s}_k)$ below a certain level.

3.2.3 Parameters estimation

In this section, we present the collected estimates forward-calibrated by the described methods on each financial derivative. First, we calibrated the Hull-White model on caplets. The results are given in Table (3.1). The important fact is that the *curve_fit* minimization function does not give any significant improvement in estimates according to no significant change in either SSE or estimate values.

While σ has a reasonable value for the standard deviation of the model, the fact that the mean reversion speed rate paramater α has a similar value can indicate that such calibrated model is relatively volatile against the low speed rate of return of the interest rate r_t to the long-term mean reversion level.

Parameter	Initial guess	OF value	Estimate	SSE
α	0.0152	0.0000340	0.0153	0.0000340
σ	0.0118		0.0118	

Table 3.1: Estimates of the parameters α and σ of the Hull-White model calibrated on caplets from the *curve_fit* function with the initial guesses from the *differential_evolution* function both rounded to 3 significant figures. The quality of the estimates from *differential_evolution* are compared according to the value of the objective function at the estimated parameters and from *curve_fit* are compared according to SSE.

It leads us to think about the possibility of a model without the mean reversion effect for caplets, thus the Ho-Lee model from the equation (2.5). There is only an estimated σ parameter since $\alpha = 0$. The estimation is provided in the same way as in the case of Hull-White and the results are in Table (3.2).

Initial guess	OF value	Estimate	SSE
0.0107	0.0000375	0.0107	0.0000375

Table 3.2: Estimates of the parameter σ of the Ho-Lee model calibrated on caplets with the *curve_fit* function with the initial guesses from the *differential_evolution* function both rounded to 3 significant figures.

Again, there is no improvement after using the *curve_fit* function and the value of σ is very similar to that from the Hull-White model. The main difference is in the interpretation of θ parameter and it gives another option to see how simulations can be seen. We comment on this in Subsection 3.2.4.

Secondly, we present estimates calibrated on caps and swaptions from *differential_evolution* function in Table (3.3). There are written values of the objective function from the equation (2.25) for comparability. It means that for swaptions it is the value of the objective function from calibration of σ with single calibrated α from the objective function in the equation (2.31).

Minimal values of the objective function are much higher than for caplets. It does not necessarily mean that caplets are the best for calibration. The better fit can be caused by the easier structure of individual caplets and may be the reason why caplets are not as good for calibration as other complex financial derivatives. Swaptions and caps capture longer time-dependent behaviour, which can lead to better calibration of the dynamics of interest rates.

	Parameter	Estimate	OF value
Caps	α	3.17	0.0277
	σ	0.155	
Swaption	α	0.384	1.37
	σ	0.0157	

Table 3.3: Estimates of the parameters α and σ of the Hull-White model calibrated on caps and swaptions with the *differential_evolution* function both rounded to 3 significant figures. The quality of the estimates from *differential_evolution* are compared according to the value of the objective function at the estimated parameters.

There are two important observations. The first is that the values of estimated α are an order of magnitude greater than the values of σ for both derivatives. It is more expected than similar values in the case of caplets, because a mean reversion speed rate higher than the effect of a standard deviation may prevent a wider spread of simulated paths of interest rates. On the other hand, if we look at the great absolute values of estimates α and σ for caps, they can lead to a less realistic model. The second observation is that the calibration procedure for caps leads to a lower value of the objective function than for the calibration procedure for swaptions.

The recommendation should be that the usage of caps for calibration procedure of the Hull-White model can be better than swaptions and caplets. Notice that with different input datasets, the result could be different and it is also not intuitive to understand the dynamics of interest rates just from estimated values. That is the reason why we want to look at simulations.

3.2.4 Simulations

Finally, we create simulations of the estimated models from Subsection 3.2.3 to display behavior of the Hull-White model and show some characteristics. We use 1000 simulations of each model. We simulate in Python software with functions from *Quantlib* library. We first introduce technically the simulation procedure.

We need to start with creating an interest rate term structure with a natural cubic spline of the zero-curve input data based on Actual/Actual day-count convention (actual number of days in months and in whole year) and the Czech calendar in the specific format for *Quantlib* library. Then we define the Hull-White model through the built-in function with the processed interest rate term structure and the estimated parameters α and σ. Thanks to generators of random sequences and process paths, we create simulated paths for 30 years in 360 discrete timespans (one for each month).

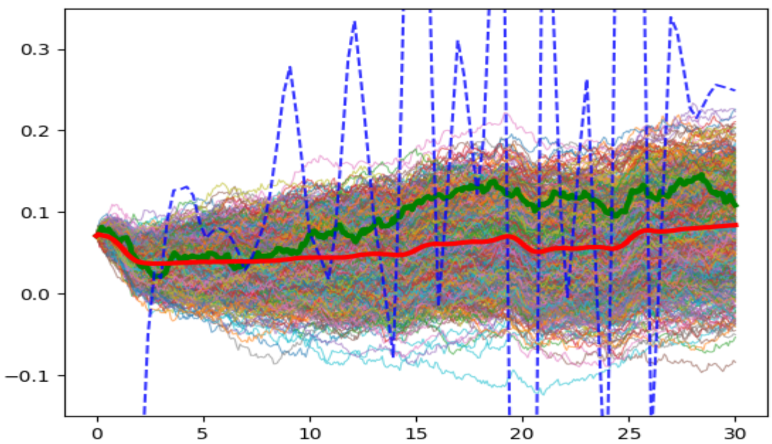

Figure 3.8: Plot of 1000 simulations of the Hull-White model calibrated on caplets with α and σ from Table (3.1) from time 0 to time 30 (in years) with highlighted the red line of means of all simulated values in given time t, the dashed blue line $\theta(t)/\alpha$ and the green line representing one example of the simulated path.

We recall that the estimates from the calibration on caplets fit market data relatively well, but values of mean reversion speed rate and of overall standard deviation are similar and the process is slowly mean reverting. Figure (3.8) shows that it creates a relatively reasonable range of simulated paths, but the sample mean of these paths and its relation to the long-term mean reversion level $\theta(t)/\alpha$ is difficult to interpret with respect to the mean reversion effect.

It is expected that the sample mean should be close to the long-term mean or at least around the long-term mean as the simulated process is mean reverting. Since the function $\theta(t)$ is divided by the small α, the mean reversion effect is logically weak with a low mean reversion speed rate and can cause a wider spread in longer time. The highlighted path displays an example of such a weak mean reversion effect not even going below the sample mean and having a confusing relation with the blue dashed line.

That is another reason why we attempted to calibrate the Ho-Lee model on our caplets data. If we look at Figure (3.9), the sample mean line and the simulated paths are not very different from Figure (3.8), but now we display only $\theta(t)$ for the Ho-Lee model, which has a different meaning here. The function $\theta(t)$ can be taken as the average future direction of interest rates, and even it can be approximated by the slope $\frac{\partial f(0,t)}{\partial t}$ from the equation (2.5) according to Hull [2017a] which corresponds to the movements of the sample mean line.

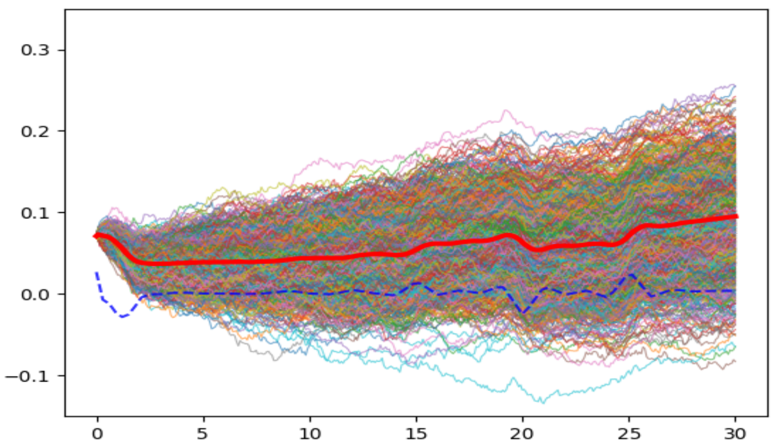

Figure 3.9: Plot of 1000 simulations of the Ho-Lee model calibrated on caplets with $\alpha = 0$ and σ from Table (3.2) from time 0 to time 30 (in years) with highlighted the red line of means of all simulated values in given time t and the dashed blue line $\theta(t)$ from the equation (2.5).

However, caplet is not the main financial derivative to examine according to results and also simulations. We move on to the caps. Unlike caplets, there is a high mean reversion speed rate and an order of magnitude lower but still relatively high standard deviation. Altogether, it creates the mean reverting process almost symmetric around long-term mean level, which collides with the sample mean line in Figure (3.10).

We can see there that the spread of the paths is consistent and relatively wide and contains highly negative interest rates. For illustration, there are added boundary lines for the 95% confidence interval around the sample mean line. The confidence interval is based on empirical quantiles of 97,5% and 2,5% of simulated values at each time t. The low boundary line is slightly above value -10 %, which is a very low value in terms of interest rates.

The quality of estimates calibrated on caps is greater than of those calibrated on swaptions according to the minimized value of the objective function in Table (3.3) but the question is whether the dynamics of the interest rates given by caps is realistic or reasonable. We might be sceptical about the caps calibration based on the listed reasons, such as high volatility of the model or the significantly negative interest rates with relatively high probability.

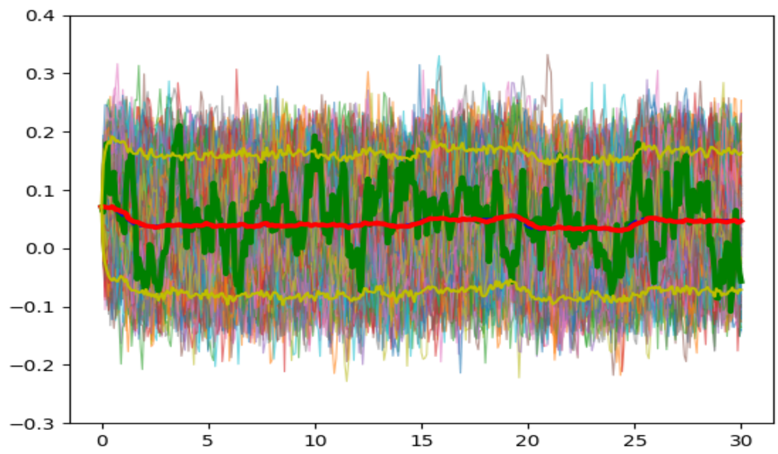

Figure 3.10: Plot of 1000 simulations of the Hull-White model calibrated on caps with α and σ from Table (3.3) from time 0 to time 30 (in years) with highlighted the red line of means of all simulated values in given time t, the yellow lines defining 95 % confidence interval around the mean line, the dashed blue line $\theta(t)/\alpha$ and the green line representing one example of the simulated path.

The next step is to look at simulations of the Hull-White model calibrated on swaptions in Figure (3.11). We need to compare it to caps, and we can see improvement in the width of spread of simulated paths and in the possibility of reaching negative interest rates. The low boundary line for 95% confidence interval drops below the zero level only between years 20 and 25 and even there not significantly. The sample mean red line and the example of the simulated path are close to the long-term mean reversion level excluding the first 3 years of initial drop of the yield curve and times $t \in \{20, 25\}$ probably caused by a low liquidity of such swaptions with the long expiry.

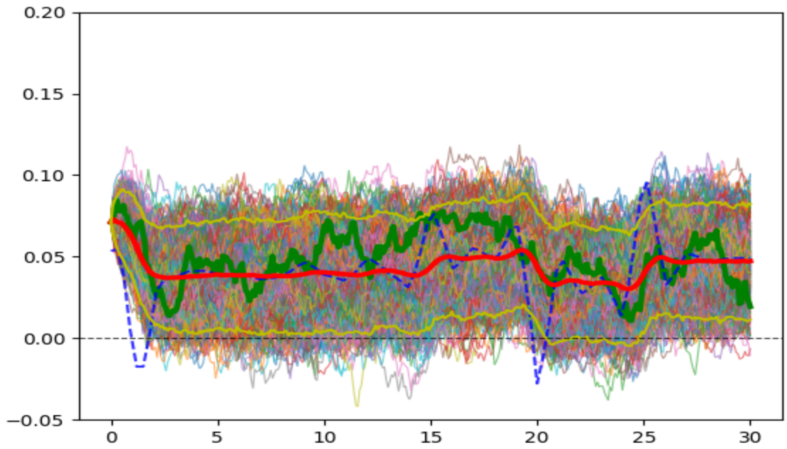

Figure 3.11: Plot of 1000 simulations of the Hull-White model calibrated on swaptions with α and σ from Table (3.3) from time 0 to time 30 (in years) with highlighted the red line of means of all simulated values in given time t, the yellow lines defining 95 % confidence interval around the mean line, the dashed blue line $\theta(t)/\alpha$ and the green line representing one example of the simulated path.

We might be willing to use swaptions rather than caps for the calibration procedure based on these conclusions. It is important to point out that a different usage of the calibrated and simulated Hull-White model can require different input datasets containing specific lengths of tenors, maturities, and expiries. There are two main conclusions based on our result. We should choose swaptions or caps for the forward-looking approach of the calibration of the Hull-White model based on our requirements. According to the overall calibration with all available tenors, maturities, and expiries for caps and swaptions, we recommend using swaptions because of their reasonable estimated value of parameters and the corresponding behaviour of the simulations.

3.3 Results for the backward-looking approach

Now we want to follow backward-looking approach methods based on historical data to estimate the Hull-White model. We described theoretically in Subsection 2.3.2 how to define the objective function to use historical data directly. There is no need to create other data such as prices of financial derivatives in the case of the forward-looking approach. We are immediately able to minimize the objective function and estimate the Hull-White model. The simulations are then performed.

3.3.1 Parameters estimation

We introduce the estimates of the Hull-White parameters from the backward-looking approach in Table (3.4). We can see that the minimal value of the objective function is even lower than the value in Table (3.1) and α is an order of magnitude larger than σ. On the other hand, the estimated value of α is at the edge of the space where we minimize the objective function.

Parameter	Estimate	OF value
α	0.0100	0.00000565
σ	0.00895	

Table 3.4: Estimates of the parameters α and σ of the Hull-White model based on the backward-looking approach from *differential_evolution* rounded to 3 significant figures. The quality of the estimate from *differential_evolution* is given by the value of the objective function at the estimated parameters.

Such a low mean reversion speed can cause worse behavior and interpretability of simulations of the Hull-White model, such as those of caplets. Furthermore, the backward-looking approach can be limiting in the sense of setting θ parameter by forward rates and may cause such small values of the parameters, especially α.

3.3.2 Simulations

We illustrate results from historical calibration on the same simulations as in the case of the forward-looking approach, therefore, we simulate 1000 paths of the Hull-White model with α and σ from Table (3.4) for 30 years in 360 timespans (one for each month). We can see in Figure (3.12) that the simulations based on historical data are similar to that based on caplet calibration because there are similar results, as low mean reversion speed but even less volatile. This leads to a narrower spread of paths.

There are also similar problems to those with the simulations calibrated on caplets because low α indicates a lack of the mean reversion effect. This observation can explain the misleading display of the long-term mean reversion level $\theta(t)/\alpha$ which obviously has no senseful relation with the simulated process and the sample mean line. Despite a very good fit of market data through the minimization of the objective function, the conclusion is that this approach may not be the best for estimating the Hull-White model. Again, it is possible to perform the Ho-Lee model, but it is not our goal here.

The explanation could be that we calibrate the model connected through $\theta(t)$ with forward rates on historical data, which can cause mismatch even if we minimize the objective function based on comparing theoretical volatilities and observed volatilities on daily yield changes not based on prices. The solution could be found in the subsequent work in comparing different characteristics or in choosing the way how to combine the forward and the backward-looking approach more properly.

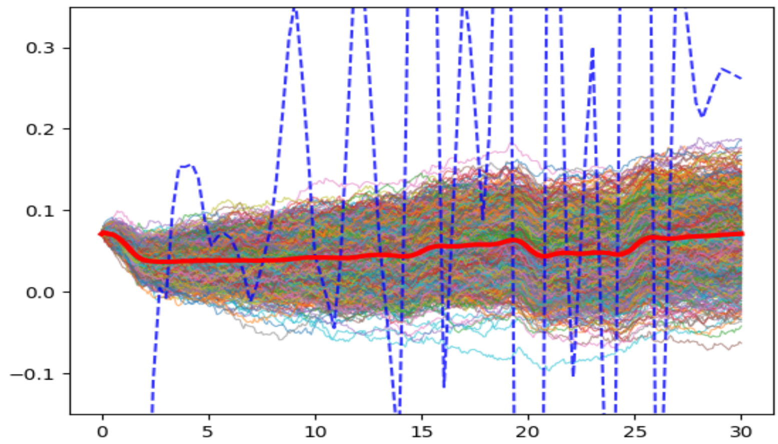

Figure 3.12: Plot of 1000 simulations of the Hull-White model estimated from the backward-looking approach with α and σ from Table (3.4) from time 0 to time 30 (in years) with the red line of means of all simulated values in given time t and the dashed blue line $\theta(t)/\alpha$.

Conclusion

The aim of the thesis was to introduce the Hull-White model and to study the corresponding stochastic process and parameters. In particular, we derived the formula for the time-dependent parameter $\theta(t)$. Next, we set different calibration procedures with the forward and backward-looking approach and the necessary tools such as interest rate derivative prices, a theoretical volatility for swaptions and theoretical volatilities of historical daily changes of a spot rate. Finally, we performed the calibration in terms of the nonlinear optimization problem in the simulation study and simulated the Hull-White model for each founded result.

We formalized the general form of the model that describes the dynamics of the instantaneous spot rate r_t for our thesis. As one of special examples, we introduced the Hull-White model as a stochastic mean reverting process with constant parameters α and σ and the time-dependent parameter $\theta(t)$ standing for the long-term mean reversion level. We solved this model as an SDE with the usage of Itô's lemma.

After that, we derived step by step the form of $\theta(t)$ implied by the Hull-White model and determined not only with the parameters of the model, but also with the instantaneous forward rate that provides the ability of good future simulations. We took the stochastic process for the zero-coupon bond prices and the relations between the forward and the spot rates to derive $\theta(t)$ in the way that it linked the model with the initial term structure of the interest rates. The derivation is based on Hull [2017b]. We performed this derivation in more detail in five structured steps in the thesis.

In the simulation study, we calibrated the Hull-White model on caplets, caps, and swaptions in the framework of the forward-looking approach and then calibrated the model on comparing theoretical and observed volatilities of daily changes of historical spot rates. For each of this approach, we simulated the estimated model and compared them based on combined results of the estimates and the simulations. There are conclusions and recommendations of the calibration procedure that should be used and why and what can be done differently for various requirements.

The historical estimate and the caplet calibration performed a good fit of theoretical and market characteristics on implied characteristics by the model, but the simulations had a lack of the mean reversion effect important for the Hull-White model, since α parameter describing the mean reversion speed was very low and therefore there was a lack of a reasonable interpretability.

The calibrations on swaptions and caps demonstrate a senseful fit in terms of the optimization procedure, and simulations worked better based on the estimated mean reversion and the standard deviation. The cap simulations unfortunately offered the wide spread of the interest rates with non-negligible probability of reaching extremely negative values, which does not seem very reasonable. The swaption simulations did not contain the disadvantages of the previously mentioned simulations, and the concluding recommendation should be that the calibration on swaptions may estimate the Hull-White model best based on our input dataset.

Bibliography

M.A. Branch, T.F. Coleman, and Y. Li. A subspace, interior and conjugate gradient method for large-scale bound-constrained minimization problem. *SIAM Journal on Scientific Computing*, 21(1):1–23, 1999. Available at https://epubs.siam.org/doi/epdf/10.1137/S1064827595289108.

D. Brigo and F. Mercurio. *Interest Rate Models – Theory and Practice.* Second Edition. Springer-Verlar, Berlin, 2006. ISBN 978-3-540-22149-4.

Jon Exley, Shyam Mehta, and Andrew Smith. Mean reversion. In *finance and investment conference*, pages 1–31, 2004. Available at https://www.researchgate.net/profile/Jon_Exley/publication/264858032_MEAN_REVERSION/links/60815d4b8ea909241e17585b/MEAN-REVERSION.pdf.

A. Friedman. *Stochastic differential equations and applications: two volumes bound as one.* Dover Publications, Mineola, N.Y., 2006. ISBN 0-486-45359-6.

S. Gurrieri, M. Nakabayashi, and T. Wong. Calibration methods of hull-white model. 2009. Available at https://dx.doi.org/10.2139/ssrn.1514192.

M. Hess. The stochastic leibniz formula for volterra integrals under enlarged filtrations. *Stochastic Models*, 39(4):823–850, 2023. Available at http://dx.doi.org/10.2139/ssrn.3308400.

J.C. Hull. *Options, futures and other derivatives.* Ninth Edition. Pearson, London, 2017a. ISBN 978-0-13-345631-8.

J.C. Hull. Technical note no. 31. http://www-2.rotman.utoronto.ca/~hull/technicalnotes/TechnicalNote31.pdf, 2017b. Technical note for Options, futures and other derivatives.

R.I. Jennrich. Asymptotic properties of non-linear least square estimators. *The Annals of Mathematical Statistics*, 40(2):633–643, 1969.

I. Karatzas and S. Shreve. *Brownian motion and stochastic calculus.* Second Edition. Springer, New York, 1991. ISBN 0-387-97655-8.

B. Øksendal. *Stochastic differential equations: an introduction with applications.* Sixth Edition. Springer-Verlag, Berlin, 2003. ISBN 3-540-04758-1.

M.J. Panik. *Stochastic differential equation.* First Edition. John Wiley & Sons, New Jersey, 2017. ISBN 978-1-11-937741-2.

Z. Prášková. *Základy náhodných procesů II.* Druhé upravené vydání. Matfyzpress, Praha, 2016. ISBN 978-80-246-3529-3.

J. Qiang, C. Mitchell, and A. Qiang. Tuning of an adaptive unified differential evolution algorithm for global optimization. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 4061–4068. IEEE, 2016. Available at `https://www.osti.gov/servlets/purl/1163659`.

V. Russo and G. Torri. Calibration of one-factor and two-factor hull–white models using swaptions. *Computational Managment Science*, 16(1):275–295, 2019. Available at `https://doi.org/10.1007/s10287-018-0323-z`.

K. Sivertsen. Interest rate models in solvency ii. Master's thesis, University of Bergen, 2016. Available at `https://bora.uib.no/bora-xmlui/bitstream/handle/1956/15421/150432241.pdf?sequence=1`.

List of Tables

3.1 Estimates of the parameters α and σ of the Hull-White model calibrated on caplets from the *curve_fit* function with the initial guesses from the *differential_evolution* function both rounded to 3 significant figures. The quality of the estimates from *differential_evolution* are compared according to the value of the objective function at the estimated parameters and from *curve_fit* are compared according to SSE. 39

3.2 Estimates of the parameter σ of the Ho-Lee model calibrated on caplets with the *curve_fit* function with the initial guesses from the *differential_evolution* function both rounded to 3 significant figures. 40

3.3 Estimates of the parameters α and σ of the Hull-White model calibrated on caps and swaptions with the *differential_evolution* function both rounded to 3 significant figures. The quality of the estimates from *differential_evolution* are compared according to the value of the objective function at the estimated parameters. . 40

3.4 Estimates of the parameters α and σ of the Hull-White model based on the backward-looking approach from *differential_evolution* rounded to 3 significant figures. The quality of the estimate from *differential_evolution* is given by the value of the objective function at the estimated parameters. 45