Calibration of FX hybrid model with time-dependent parameters

Dion A.W. van Lange

in cooperation with
Rabobank

Faculty EEMCS, Delft University of Technology

TU Delft

Calibration of FX hybrid model with time-dependent parameters

by

Dion A.W. van Lange

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Wednesday December 9, 2020 at 10:00 AM.

Student number:	4462815	
Project duration:	February 12, 2020 – December 9, 2020	
Supervisor:	Dr. Ir. L. A. Grzelak,	TU Delft & Rabobank
Thesis committee:	Prof. Dr. Ir. C. W. Oosterlee	TU Delft
	Dr. Ir. L. A. Grzelak,	TU Delft & Rabobank
	Dr. R. J. Fokkink	TU Delft

This thesis is confidential and cannot be made public until December 2, 2020.

An electronic version of this thesis is available at `http://repository.tudelft.nl/`.

Cover image is from Shutterstock by Hans Engbers.

TU Delft

Abstract

This thesis captures the calibration of a FX hybrid model: The FX Black-Scholes Hull-White model. The main focus is on the calibration of the parameters in the Hull-White process: The mean reversion and the volatility parameter. The latter is commonly calibrated as a time-dependent parameter, whilst the mean reversion parameter is not. This thesis covers the calibration of the mean reversion as a time-dependent parameter. A known method from the Literature ([1], [24]) is researched, where we calibrate the mean reversion independently from the volatility parameter to the ratio of two swaptions with the same expiry but different tenor. In our research this method is extended to the negative interest rate environment by assuming that the swap rate follows a shifted lognormal distribution instead of a lognormal distribution. We show that a specific set of swaptions can be chosen, so that the calibration problem is simplified. This choice leads to sequential calibration of convex optimization problems. Numerical results of calibration to artificial and market data are presented, where we compare our method to picking the mean reversion parameter arbitrarily. The findings suggest that the choice of mean reversion parameter affects the calibration procedure. Therefore, we argue that a calibration method for the mean reversion would be appropriate.

Besides the focus on the Hull-White process, the calibration of the volatility parameter in the FX Black-Scholes process of the hybrid model is investigated. For the calibration of the FX volatility parameter ATM options on FX rates are used. Numerical results of calibration to artificial and market data are shown. A typical problem of calibration in the industry is discussed: Precision for late maturities. The results in this thesis suggest that this problem could be solved.

For research on homogeneity constraints on the time-dependent parameters, the performance of delta-hedging is investigated. The delta-hedging is performed in a simple Black-Scholes world with time-dependent volatility. We show that given a fixed amount hedges beforehand and interest rate zero, there exists an optimal distribution of time points that will lead to equal variance on the profit and loss for every volatility function that has equal implied volatility from t_0 to maturity T. Using this strategy, the homogeneity of the volatility parameter has no impact on the performance of delta-hedging. Therefore, in this thesis no homogeneity constraints are used for the calibration of the time-dependent parameters.

Keywords: Hull-White model, time-dependent parameters, Hybrid models, calibration, swaption, Black-Scholes Hull-White model, Foreign Exchange market, FX, hedging, delta-hedging.

Acknowledgements

This thesis has been submitted in partial fulfillment of the requirements for the degree of Master of Science in Applied Mathematics at Delft University of Technology. The academic supervisor of this project has been Prof. Dr. Ir. C. W. Oosterlee, of the Numerical Analysis group at the Delft Institute of Applied Mathematics. The research has been conducted under the supervision of Dr. Ir. L. A. Grzelak, partly at the Rabobank office in Utrecht, partly from Delft University of Technology and partly from home due to the COVID 19 situation. The promotor of this research has been the Financial Engineering team at Rabobank

Most of all, I would like to express my gratitude to Lech Grzelak for the opportunity of joining the Financial Engineering team at Rabobank and his invaluable supervision. His guidance has led me through the difficult world of financial engineering and I would like to thank Lech for his patience, especially during the introduction of this new world. Secondly, I would like to thanks Kees Oosterlee as academic supervisor and committee member, and Robbert Fokkink as committee member. At last, but certainly not less important, I would like to thank my friends, girlfriend and family, who brought me joy and happiness during this lonely process, whose encouragement and support allowed me to study in the most enjoyable way, who believed in me until the last moment.

Dion A.W. van Lange
Delft, December 2020

Contents

Glossary

ATM At The Money. iii, 26, 27, 36, 40

BSHW Black-Scholes Hull-White. viii, 7, 48, 75–80

CDF Cumulative Distribution Function. 5, 7, 12, 13, 48, 50, 51, 56, 58, 62

CIR Cox-Ingersoll-Ross. 30, 34, 35

CVA Credit Value Adjustment. 42

EE Expected Exposure. 42

FX Foreign Exchange. iii, vii, 1, 3, 7, 27, 47–54, 60

ITM In The Money. 4, 19, 27, 40

LGD Loss Given Default. 43

OTM Out of The Money. 4, 19, 27, 40

PDF Probability Density Function. 13

xVA x-Value Adjustment. 19

ZCB Zero Coupon Bond. 2, 10–15, 48

1 Introduction

In October 1987, a stock market crash took place on Monday 19th, which is well known as Black Monday. Although this was a terrible experience for the whole world, the computational finance engineers learned something. The constant volatility in the Black-Scholes is not correct for the pricing of deep in-the-money and out-of-the-money options [34]. From that point onward we observe an implied volatility smile in the market. Today we still use the implied volatility. Some financial products in the market are not quoted on the price of the product, but on the implied volatility, e.g. swaptions.

In 2008 the world experienced another stock market crash. Again a terrible experience, but once more an educational experience for the financial engineers. Prior to this crisis in 2008, the probability of default was considered negligible when pricing interest rate products [26]. However, in this crisis a lot of companies bankrupted, and the financial engineers learned that the probability of default cannot be neglected. Incorporating the probability of default in the pricing of financial products led to the introduction of the multi-curve framework [40] [19], which will be discussed in Section 2.3.

These two examples demonstrate that research in financial modeling is important. Perhaps one day the financial engineers are so good at modeling, that market crashes will not take place anymore. This thesis focuses mainly on the modeling of interest rate. Especially, the last five years modeling interest rate has become more interesting. Ten years ago, no one would argue that negative interest rates are possible, but today we are actually experiencing negative interest rates in the market. The Hull-White model [13] is a popular model for the modelling of interest rates. The Hull-White model had one drawback ten years ago; It allows negative interest rates. Over time the Hull-White model has become even more popular, as financial engineers slowly began to understand that negative interest rates are possible. The drawback of allowing negative interest rates turned into an advantage, and the Hull-White model got even more popular over time.

A difficult and computational expensive task is the calibration of models, e.g. Hull-White model. Calibration implies finding the best set of parameters, for which the quotes in the market correspond with the prices computed by the model. Traders use these calibrated parameters for trading purposes, e.g. hedging. Well calibrated models will improve the performance of the traders, and in turn traders will be better at providing liquidity in the markets. This is a necessity, especially in extreme down turns of financial markets as we recently experienced in March 2020. During these perilous times, liquidity seems to disappear from the market, which is horrendous for market participants (e.g. pension funds). Therefore, researching mathematical models for pricing of financial products is important. This thesis covers the calibration of Hull-White model and the FX Black-Scholes Hull-White model with time-dependent parameters.

1.1. Mathematical framework

The mathematical financial framework used in this thesis is discussed in this section. For some readers this might be basic knowledge. Nevertheless, it is important to clarify the notation used in the thesis. In order to build this financial framework, definitions will be stated one-by-one. We start with the definition of the risk-free rate.

Definition 1.1.1. (Risk-free rate) The risk-free rate $r(t)$ represents the interest rate one can obtain in the market without any risk, e.g. risk of the counterparty defaulting. The risk-free rate can be computed from from financial instruments available in the market.

In mathematical finance, we often use the risk-neutral measure for the computation of prices of financial products. In order to use this measure, we have to define the money-market account. The money-market account is closely related to the risk-free interest rate.

Definition 1.1.2. (Money-market account) The money-market account $M(t)$ represents the value of the bank account at time t. It is driven by the risk-free rate $r(t)$ in the following differential equation:

$$\mathrm{d}M(t) = r(t)M(t)\mathrm{d}t, \tag{1.1}$$

where we have that at the beginning $M(t_0) = 1$. Therefore, as solution we obtain:

$$M(t) = \mathrm{e}^{\int_{t_0}^{t} r(s)\mathrm{d}s}. \tag{1.2}$$

The money-market account is used for discounting purposes under the risk-free measure.

Besides the money-market account, zero-coupon bonds can be used for discounting under the T-forward measure. The definition of a zero-zoupon bond is given below.

Definition 1.1.3. (Zero-coupon bond) A zero-coupon bond (ZCB) is a bond that does not pay any coupons during its existence. A ZCB merely pays one at maturity time T. The value of a ZCB at time t with maturity T is denoted as $P(t,T)$, where $P(T,T) = 1$ and $T \geq t$.

Definition 1.1.4. (Annuity factor) The annuity factor is defined as a linear combination of zero coupon bonds:

$$A_{i,m}(t) = \sum_{k=i+1}^{m} \tau_k P(t,T_k), \tag{1.3}$$

where $\tau_k = T_k - T_{k-1}$. The annuity factor is used as a discount factor under the Annuity measure.

Definition 1.1.5. (Simple compounded forward libor rate) The libor rate is the London Interbank Offered Rate. This is an interest rate for which a group of banks in London lend money to one another. The simple compounded forward libor rate is used as the floating rate of financial interest rate products like swaps and swaptions. The libor rate is defined by zero coupon bonds in the following way:

$$\ell_i(t) := \ell(t,T_{i-1},T_i) = \frac{P(t,T_{i-1}) - P(t,T_i)}{\tau_i P(t,T_i)}, \tag{1.4}$$

where $\tau_i = T_i - T_{i-1}$. Equation (1.4) will be used frequently for the derivation of financial products in this thesis.

In this thesis we will cover the Hull-white model, which is part of Heath-Jarrow-Morton framework [10].This framework contains various interest rate models, e.g. Hull-White model. For details on the Heath-Jarrow-Morton framework see [10] or [25]. For this thesis we will need the instantaneous forward-rate. By using the instantaneous forward rate we can obtain $r(t_0)$.

Definition 1.1.6. (Instantaneous forward-rate) The instantaneous forward-rate is defined as:

$$f^r(0,T) := -\frac{\partial}{\partial T}\log\big(P(0,T)\big). \tag{1.5}$$

In the Heath-Jarrow-Morton framework the short rate $r(t)$ is defined in the following way:

Definition 1.1.7. (short-rate) Tthe short-rate process $r(t)$ under the Heath-Jarrow-Morton framework is given by the limit of the instantaneous forward rate [25]:

$$r(t) \equiv f^r(t,t).$$

From Definition 1.1.7, we can compute $r(t_0)$, which will be needed for the calibration of the Hull-White model. Next, the concept of change of measure will be discussed.

Measure changes

In mathematical finance, we price financial products under various probability measures. One often changes measure in order to simplify the payoff of the products, and obtain an exact pricing. For the switching between two measures Girsanov's theorem is used. Girsanov's theorem describes how the dynamics of a stochastic process change when a different measure is considered. For this method a so-called numéraire is used, which is a tradable unit of measure that the other prices are expressed in. Girsanov's theorem is very important, as it gives us the possibility to move from working under the real world measure $\mathbb{P}$ to the risk-neutral measure $\mathbb{Q}$. Besides this change of measure, Girsanov theorem allows us to perform other measure changes as well. In this thesis we frequently change measure from the risk-neutral measure to the T-forward measure, as discounting with zero coupon bonds is often more favourable than discounting with the money market account. Also the annuity measure is used in this thesis. Below the notation of the Risk-neutral, T_i-forward and Annuity measure are denoted:

- Risk-neutral measure $\mathbb{Q}$, where the money market account is the numéraire. Computing expectation under the risk-netural measure, is denoted as $\mathbb{E}^{\mathbb{Q}}[...]$
- T_i-forward measure, where the zero coupon bond $P(t,T_i)$ is the numéraire. Computing expectation under the T_i-forward measure, is denoted as $\mathbb{E}^{T_i}[...]$
- Annuity measure, where the annuity factor $A_{i,m}(t)$ is the numéraire. Computing expectation under the Annuity measure, is denoted as $\mathbb{E}^{i,m}[...]$

1.2. Interest rate products

For the calibration of FX hybrid models, one needs financial interest rate products to calibrate the model to. In this section we will discuss some of the various interest rate products that our available in the market.

Definition 1.2.1. (Interest rate swap) An interest rate swap (IRS) is a contract that allows one party to receive a fixed interest rate K and the counterparty receives a floating rate over multiple periods. The floating rate is denoted as $\ell(t)$, which commonly is the libor rate as in Definition 1.1.5. How much one would like to pay or receive is stated by the notional amount N. When the contract states that you pay the fixed rate K and receives the floating rate $\ell(t)$, we call a payer's swap. Vice versa we call it the receiver's swap. The value of the interest rate swap is given by

$$V_{swap}(t_0,T_i,T_m) = \alpha \sum_{k=i+1}^{m} \tau_k N\big(\ell_k(T_{k-1}) - K\big). \tag{1.6}$$

α denotes whether the swap is in the form of the payer or the receiver; $\alpha = 1$ for payer and $\alpha = -1$ for receiver. N is the notional amount. $\tau_k = T_k - T_{k-1}$ denotes the passed time from the previous payment. In this thesis we merely consider swaps with equidistant payments, which makes τ independent of k. As no other types of swaps will be discussed in this thesis, we will refer to interest rate swap by swap.

Definition 1.2.2. (Swap rate) The swap rate $S_{i,m}(t)$ is the strike value K that leads to a value of the swap being zero.

Theorem 1.2.1. The swap rate $S_{i,m}(t)$ is given by:

$$S_{i,m}(t) = \frac{P(t,T_i) - P(t,T_m)}{A_{i,m}(t)}. \tag{1.7}$$

Proof. We will follow the proof as in [25]. Consider a swap with life span (T_i, T_m) and payment dates $T_{i+1}, ..., T_m$. Under the risk neutral measure, we obtain for the value of a swap at time t_0:

$$\begin{aligned} V_{swap}(t_0,T_i,T_m) &= \mathbb{E}^{\mathbb{Q}}\Big[\frac{M(t_0)}{M(T_k)}\alpha \sum_{k=i+1}^{m} \tau_k N(\ell_k(T_{k-1}) - K)\Big|\mathscr{F}(t_0)\Big] \\ &= \alpha \cdot N \cdot \sum_{k=i+1}^{m} \mathbb{E}^{\mathbb{Q}}\Big[\tau_k \frac{M(t_0)}{M(T_k)}(\ell_k(T_{k-1}) - K)\Big|\mathscr{F}(t_0)\Big]. \end{aligned}$$

For now lets consider the $\alpha = -1$ case. The proof for $\alpha = 1$ is similar and can be found in [25]. We will swap from the risk-neutral measure to the T_k-forward measure. We obtain:

$$\begin{aligned} V_{swap}(t_0, T_i, T_m) &= -N \cdot \sum_{k=i+1}^{m} \mathbb{E}^{T_k}\Big[\tau_k P(t_0, T_k)(\ell_k(T_{k-1}) - K)|\mathscr{F}(t_0)\Big] \\ &= N \cdot \sum_{k=i+1}^{m} \tau_k P(t_0, T_k)K - N \cdot \sum_{k=i+1}^{m} \tau_k P(t_0, T_k)\mathbb{E}^{T_k}[\ell_k(T_{k-1})|\mathscr{F}(t_0)]. \end{aligned}$$

Note that $\ell_k(T_{k-1})$ is a martingale under the T_k-forward measure, therefore: $\mathbb{E}^{T_k}[\ell_k(T_{k-1})|\mathscr{F}(t_0)] = \ell_k(t_0)$. We can use the definition of $\ell_k(t_0)$ and the method of a telescope series to obtain:

$$\sum_{k=i+1}^{m} P(t_0, T_k)\ell_k(t_0) = \sum_{k=i+1}^{m} P(t_0, T_k)\frac{P(t_0, T_{k-1}) - P(t_0, T_k)}{P(t_0, T_k)} = \sum_{k=i+1}^{m} P(t_0, T_{k-1}) - P(t_0, T_k) = P(t_0, T_i) - P(t_0, T_m).$$

We obtain for the value of the swap:

$$V_{swap}(t_0, T_i, T_m) = N \cdot \sum_{k=i+1}^{m} \tau_k P(t_0, T_k)K - \tau_k \cdot N \cdot \big(P(t_0, T_i) - P(t_0, T_m)\big). \tag{1.8}$$

Setting the value of the swap to zero and isolating K gives the result. □

Lemma 1.2.2. The value of a swap is given by:

$$V_{swap}(t_0, T_i, T_m) = \alpha \cdot N \cdot A_{i,m}(t_0)\big(S_{i,m}(t_0) - K\big), \tag{1.9}$$

where $\alpha = 1$ represents the payer and $\alpha = -1$ represents the receiver.

Proof. By rewriting Equation (1.8) using the definition of the annuity factor and swap rate the result follows for the receiver swap. The payer swap can be obtained in a similar fashion. □

Definition 1.2.3. (Caplet / Floorlet) A caplet provides insurance may the floating rate go below a predefined level K (the cap-rate) at a future time T_i. How much one would like to pay or receive is stated by the notional amount N. The payoff of a caplet is given by:

$$V_{Caplet}(T_i) = \tau_i N \max\big(\ell_i(T_{i-1}) - K, 0\big), \tag{1.10}$$

where $\tau_i = T_i - T_{i-1}$. The floorlet is the other way around, and the payoff is given by:

$$V_{Floorlet}(T_i) = \tau_i N \max\big(K - \ell_i(T_{i-1}), 0\big). \tag{1.11}$$

Caps and floors are sums of respectively caplets and floorlets.

Definition 1.2.4. (Swaption) A Swaption is an option on a specific underlying swap; one has the possibility to buy a specific swap at maturity for a strike price K. There are two types of swaptions; the payer and the receiver. These can be thought of as a call and a put. The payoff of the swaption is given by the following:

$$V_{swpt}(T_i) = \max\big(V_{swap}(T_i), 0\big) = N \cdot \max\Big(\alpha \sum_{k=i+1}^{m} \tau_k\big(\ell_k(T_{k-1}) - K\big), 0\Big), \tag{1.12}$$

where α determines if it is a payer or receiver swap. T_i is the maturity of swaption, also known as the expiry. T_m is the maturity of the underlying swap, also known as the tenor. $\tau_k = T_k - T_{k-1}$ represents the time between two payments. N is the notional amount.

Definition 1.2.5. The strike price of a swaption with expiry T_i and tenor T_m is called at the money (ATM) if and only if

$$K = K_{ATM} := S_{i,m}(t_0) = \frac{P(t_0, T_i) - P(t_0, T_m)}{\sum_{k=i+1}^{m} \tau_k P(t_0, T_k)}. \tag{1.13}$$

We call a payer swaption out of the money (OTM) when $K > K_{ATM}$ and in the money (ITM) if $K < K_{ATM}$. For a receiver swaption it is vice versa; out of the money (OTM) when $K < K_{ATM}$, in the money (ITM) if $K > K_{ATM}$.

Theorem 1.2.3. The value of a swaption at time t_0 with maturity T_i on a swap with payment dates $T_i, ..., T_m$ is given by

$$V_{swpt}(t_0, T_i, T_m) = N \cdot A_{i,m}(t_0) \cdot \mathbb{E}^{i,m}\Big[\max\big((S_{i,m}(T_i) - K), 0\big) \Big| F(t_0) \Big]. \tag{1.14}$$

Proof. We follow the proof as in [25]. For the value of a swaption under the risk-neutral measure, we have by using Theorem 1.2.2 and Definition 1.2.4:

$$V_{swpt}(t_0, T_i, T_m) = N \cdot \mathbb{E}^{\mathbb{Q}}\Big[\frac{M(t_0)}{M(T_i)} \max\Big(A_{i,m}(T_i)\alpha(S_{i,m}(T_i) - K), 0\Big) \Big| \mathscr{F}(t_0) \Big].$$

By performing a change of measure from the risk neutral measure to the annuity measure, we will obtain the result. The corresponding Radon-Nikodym derivative is given by:

$$\lambda^{i,m}_{\mathbb{Q}}(T_i) = \frac{A_{i,m}(T_i)M(t_0)}{A_{i,m}(T_0)M(T_i)}. \tag{1.15}$$

Performing the change of measure gives:

$$\begin{aligned} V_{swpt}(t_0, T_i, T_m) &= N \cdot \mathbb{E}^{i,m}\Big[\frac{M(t_0)}{M(T_i)} A_{i,m}(T_i) \frac{M(T_i)}{M(t_0)} \frac{A_{i,m}(t_0)}{A_{i,m}(T_i)} \max\big(\alpha(S_{i,m}(T_i) - K), 0\big) \Big| \mathscr{F}(t_0) \Big] \\ &= N \cdot A_{i,m}(t_0) \cdot \mathbb{E}^{i,m}\Big[\max\big(\alpha(S_{i,m}(T_i) - K), 0\big) \Big| \mathscr{F}(t_0) \Big]. \end{aligned}$$

□

By Theorem 1.2.3, we have the value of a swaption at time t_0 expressed in an expectation. $S_{i,m}(t)$ is a martingale under the swap measure, since $S_{i,m}(t)$ is divided by the numéraire $A_{i,m}(t)$. Therefore, we have that $S_{i,m}(t)$ is driftless under the swap measure. We can use Black's formula to determine the value of a swaption. We can assume $S_{i,m}$ to be lognormal or normal , which will lead to two different pricings. Typically one chooses a lognormal distribution. However, lognormal does not allow negative interest rate, unless one shifts the distribution. Years ago, it was thought that negative interest rates were not even possible. Therefore, lognormal was considered as the optimal choice. However, today we encounter negative interest rates, which gives issues for the lognormal distribution. Today there are quotes on the implied volatility for both normal and shifted lognormal distriubtion. The market data used in this thesis consists of implied volatilities from a shifted lognormal distribution.

Theorem 1.2.4. When assuming that the swap rate is a lognormal process:

$$\mathrm{d}S_{i,m}(t) = \sigma^L_{i,m} S_{i,m} \mathrm{d}W^{i,m}(t),$$

the value of a swaption at time t_0 is given by (Black's formula [3]):

$$V^L_{Swpt}(t_0, T_i, T_m) = NA_{i,m}(t_0)\big(\alpha S_{i,m}(t_0)\Phi(\alpha d_1) - K\alpha\Phi(\alpha d_2)\big),$$

where $\alpha = 1$ for the payer swaption and $\alpha = -1$ for the receiver swaption. $\Phi(\cdot)$ is the standard lognormal CDF. With the following for d_1 and d_2:

$$d_1 = \frac{\log\left(\frac{S_{i,m}(t_0)}{K}\right) + \frac{1}{2}(\sigma^L_{i,m})^2(T_i - t_0)}{\sigma^L_{i,m}\sqrt{T_i - t_0}}, \; d_2 = d_1 - \sigma^L_{i,m}\sqrt{T_i - t_0}.$$

Alternatively one can reach a similar result when assuming a Normal distribution [5]:

$$\begin{aligned} V^N_{Swpt}(t_0, T_i, T_m) &= NA_{i,m}(t_0)\left(\alpha(S_{i,m}(t_0) - K)\Phi(\alpha d) + \frac{\sigma^N_{i,m}\sqrt{T_i - t_0}}{2\sqrt{\pi}} \mathrm{e}^{-\frac{d^2}{2}}\right) \\ d &= \frac{S_{i,m}(t_0) - K}{\sigma^N_{i,m}\sqrt{T_i - t_0}}. \end{aligned}$$

Notice the superscripts on the V and σ indicating whether it is from lognormal or normal assumption; L for lognormal and N for normal.

Proof. We will derive the proof for the payer version of the swaption assuming a lognormal distribution on the swap rate. If we now assume a lognormal distribution on the swap rate with volatility $\sigma^L_{i,m}$, i.e. $S_{i,m}$ follows a geometric brownian motion with volatility $\sigma^L_{i,m}$ under the swap measure:

$$\mathrm{d}S_{i,m}(t) = \sigma^L_{i,m} S_{i,m} \mathrm{d}W^{i,m}(t) \;\Rightarrow\; S_{i,m}(t) = S_{i,m}(t_0) \mathrm{e}^{-\frac{\left(\sigma^L_{i,m}\right)^2}{2}(t-t_0)+\sigma^L_{i,m}\left(W^{i,m}(t)-W^{i,m}(t_0)\right)}.$$

Then we can work out the expectation:

$$\begin{aligned}\mathbb{E}^{i,m}\left[\max\left((S_{i,m}(T_i)-K),0\right)\middle|\mathscr{F}(t_0)\right] &= \mathbb{E}^{i,m}\left[\left(S_{i,m}(T_i)-K\right)\mathbb{1}_{S_{i,m}(T_i)>K}\middle|\mathscr{F}(t_0)\right] \\ &= \mathbb{E}^{i,m}\left[S_{i,m}(T_i)\mathbb{1}_{S_{i,m}(T_i)>K}\middle|\mathscr{F}(t_0)\right] - \mathbb{E}^{i,m}\left[K\mathbb{1}_{S_{i,m}(T_i)>K}\middle|\mathscr{F}(t_0)\right].\end{aligned}$$

Rewriting the subscript of the indicator function gives us

$$S_{i,m}(T_i) > K \Rightarrow S_{i,m}(t_0)\mathrm{e}^{-\frac{(\sigma^L_{i,m})^2}{2}(T_i-t_0)+\sigma^L_{i,m}(W^{i,m}(T_i)-W^{i,m}(t_0))} > K \Rightarrow Z \le \frac{\log\left(\frac{S_{i,m}(t_0)}{K}\right)-\frac{(\sigma^L_{i,m})^2}{2}(T_i-t_0)}{\sigma^L_{i,m}\sqrt{T_i-t_0}} =: d_2,$$

where Z is standard normal. We obtain:

$$\begin{aligned}\mathbb{E}^{i,m}\left[\max\left((S_{i,m}(T_i)-K),0\right)\right] &= \mathbb{E}^{i,m}\left[S_{i,m}(t_0)\mathrm{e}^{\frac{(\sigma^L_{i,m})^2}{2}(T_i-t_0)-\sigma^L_{i,m}\sqrt{T_i-t_0}Z}\mathbb{1}_{Z\le d_2}\right] - \mathbb{E}^{i,m}\left[K\mathbb{1}_{Z\le d_2}\right] \\ &= S_{i,m}(t_0)\int_{-\infty}^{d_2}\mathrm{e}^{\frac{(\sigma^L_{i,m})^2}{2}(T_i-t_0)-\sigma^L_{i,m}\sqrt{T_i-t_0}s}\frac{1}{2\pi}\mathrm{e}^{-\frac{s^2}{2}}\mathrm{d}s - \int_{-\infty}^{d_2}K\mathrm{d}s \\ &= S_{i,m}(t_0)\int_{-\infty}^{d_2}\frac{1}{2\pi}\mathrm{e}^{-\frac{(s+\sigma^L_{i,m}\sqrt{T_i-t_0})^2}{2}}ds - K\Phi(d_2) \\ &= S_{i,m}(t_0)\Phi(d_2+\sigma^L_{i,m}\sqrt{T_i-t_0}) - K\Phi(d_2).\end{aligned}$$

Defining $d_1 := d_2 + \sigma^L_{i,m}\sqrt{T_i-t_0}$ gives the result. The proof of the lognormal receiver version can be derived in a similar manner. This ends the proof for the lognormal case.

Assuming a normal distribution on the swap rate, we have the following:

$$\mathrm{d}S_{i,m}(t) = \sigma^N_{i,m}\mathrm{d}W^{i,m}(t) \;\Rightarrow\; S_{i,m}(t) = S_{i,m}(t_0) + \sigma^N_{i,m}\left(W^{i,m}(t)-W^{i,m}(t_0)\right).$$

Then we can work out the expectation:

$$\begin{aligned}\mathbb{E}^{i,m}\left[\max\left((S_{i,m}(T_i)-K),0\right)\middle|\mathscr{F}(t_0)\right] &= \mathbb{E}^{i,m}\left[\left(S_{i,m}(T_i)-K\right)\mathbb{1}_{S_{i,m}(T_i)>K}\middle|\mathscr{F}(t_0)\right] \\ &= \mathbb{E}^{i,m}\left[S_{i,m}(T_i)\mathbb{1}_{S_{i,m}(T_i)>K}\middle|\mathscr{F}(t_0)\right] - \mathbb{E}^{i,m}\left[K\mathbb{1}_{S_{i,m}(T_i)>K}\middle|\mathscr{F}(t_0)\right].\end{aligned}$$

Rewriting the subscript of the indicator function gives us:

$$S_{i,m}(T_i) > K \Rightarrow S_{i,m}(t_0) + \sigma^N_{i,m}\left(W^{i,m}(t)-W^{i,m}(t_0)\right) > K \Rightarrow Z \le \frac{S_{i,m}(t_0)-K}{\sigma^N_{i,m}\sqrt{T_i-t_0}} =: d,$$

where Z is standard normal distributed random variable. We obtain:

$$\begin{aligned}\mathbb{E}^{i,m}\left[\max\left((S_{i,m}(T_i)-K),0\right)\right] &= \mathbb{E}^{i,m}\left[(S_{i,m}(t_0)+\sigma^N_{i,m}(W(t)-W(t_0))\mathbb{1}_{Z\le d}\right] - \mathbb{E}^{i,m}\left[K\mathbb{1}_{Z\le d}\right] \\ &= S_{i,m}(t_0)\int_{-\infty}^{d}\frac{\mathrm{e}^{\frac{-s^2}{2}}}{2\pi}\mathrm{d}s + \int_{-\infty}^{d}\sigma^N_{i,m}\sqrt{T_i-t_0}s\frac{\mathrm{e}^{\frac{-s^2}{2}}}{2\pi}\mathrm{d}s - K\int_{-\infty}^{d}\frac{\mathrm{e}^{\frac{-s^2}{2}}}{2\pi}\mathrm{d}s \\ &= S_{i,m}\Phi(d) - \frac{\sigma^N_{i,m}\sqrt{T_i-t_0}}{\pi}\mathrm{e}^{\frac{-d^2}{2}} - K\Phi(d) \\ &= (S_{i,m}(t_0)-K)\Phi(d) + \frac{\sigma^N_{i,m}\sqrt{T_i-t_0}}{2\sqrt{\pi}}\mathrm{e}^{-\frac{d^2}{2}}.\end{aligned}$$

□

By assuming the swap rate to follow a lognormal distribution, negative swap rates are not allowed. Consequently, negative interest rates are not allowed, which is incorrect. To dissolve this problem, one can shift the lognormal distribution by a shift parameter. In the market, prices are given for shifted lognormal distributions with various shift parameters. The following theorem gives the pricing of a swaption under shifted lognormal. A broader overview of methods for pricing in a negative interest rate environment can be found in [35].

Theorem 1.2.5. The value of a swaption at time t_0 with strike K and notional N assuming a shifted lognormal distribution on the swap rate is is given by (Black's formula):

$$V^L_{Swpt}(t_0, T_i, T_m) = NA_{i,m}(t_0)\big(\alpha \tilde{S}_{i,m}(t_0)\Phi(\alpha d_1) - \tilde{K}\alpha\Phi(\alpha d_2)\big).$$

Where $\tilde{S}_{i,m}(t) = S^s_{i,m}(t) + s$, $\tilde{K} = K + s$, $\Phi(\cdot)$ is the standard normal CDF, $\alpha = 1$ for the payer swaption and $\alpha = -1$ for the receiver swaption. With the following for d_1 and d_2:

$$d_1 = \frac{\log\left(\frac{\tilde{S}_{i,m}(t_0)}{\tilde{K}}\right) + \frac{1}{2}(\sigma^L_{i,m})^2(T_i - t_0)}{\sigma^L_{i,m}\sqrt{T_i - t_0}}, \; d_2 = d_1 - \sigma^L_{i,m}\sqrt{T_i - t_0}.$$

Proof. Assuming a shifted lognormal distribution on the swap rate with volatility $\sigma^L_{i,m}$ and shifting parameter s, gives [25]:

$$\mathrm{d}\big(S_{i,m}(t)\big) = \sigma^L_{i,m}\big(S_{i,m}(t) + s\big)\mathrm{d}W^{i,m}(t).$$

We define the shifted swap rate $\tilde{S}_{i,m}(t) = S_{i,m}(t) + s$. Note that the dynamics of the shifted swap rate are the same since s is constant, so $\mathrm{d}\tilde{S}_{i,m}(t) = \mathrm{d}S_{i,m}(t)$. Therefore, we obtain the following:

$$\mathrm{d}\big(\tilde{S}_{i,m}(t)\big) = \sigma^L_{i,m}(S_{i,m}(t) + s)\mathrm{d}W^{i,m}(t) = \sigma^L_{i,m}\tilde{S}_{i,m}(t)\mathrm{d}W^{i,m}(t).$$

We know that the value of a swaption is given by the following:

$$\begin{aligned} V_{swaption}(t_0, T_i, T_m) &= N \cdot A_{i,m}(t_0) \cdot \mathbb{E}^{i,m}\Big[\max\big(\alpha(S_{i,m}(T_i) - K), 0\big)\Big|\mathscr{F}(t_0)\Big] \\ &= N \cdot A_{i,m}(t_0) \cdot \mathbb{E}^{i,m}\Big[\max\big(\alpha(\tilde{S}_{i,m}(T_i) - s - K), 0\big)\Big|\mathscr{F}(t_0)\Big] \\ &= N \cdot A_{i,m}(t_0) \cdot \mathbb{E}^{i,m}\Big[\max\big(\alpha(\tilde{S}_{i,m}(T_i) - \tilde{K}), 0\big)\Big|\mathscr{F}(t_0)\Big], \end{aligned}$$

where $K^s = K + s$. By using Black-76 formula, we obtain the result. □

1.3. Structure of thesis

This introduction provides the reader with some knowledge on financial mathematics and interest rate products, that is needed to understand the thesis. The next two chapters will focus on interest rate, where in chapter two some light will be shed on interest rate modeling and chapter 3 dives into the calibration of the Hull-White model. Afterwards, we extend the Hull-White model to the FX Black-Scholes Hull-White model in chapter 4, where we discuss the BSHW model and its calibration. In chapter 5, we discuss the influence of time-dependent parameters on the delta-hedging process. At the end, in chapter 6, we summarize our findings and reflect on the thesis. In the Appendix, one can find the calibration algorithm and some examples of calibration. Market data has been retrieved from the Bloomberg Platform. Excel has been used for the retrieval of the data from Bloomberg and coding has been done in Python.

2

Interest rate modeling

This chapter covers aspects of interest rate modeling that we need for the calibration of the Black-Scholes Hull-White model. This chapter covers three topics. First the Hull-White model will be discussed. Afterwards the construction of the yield curve is considered. At the end, the multi-curve framework is explained.

2.1. The Hull-White model

The Hull-White model [12] is used for the modeling of interest rates. It is a no-arbitrage yield curve model in which the interest rate is driven by a mean reverting process. It comes in two variants; the one-factor and two-factor. In this thesis the one-factor is discussed, and by Hull-White model we refer to the one-factor Hull-White model. The dynamics of the Hull-White model are given by:

$$\mathrm{d}r(t) = \lambda(t)(\theta(t) - r(t))\mathrm{d}t + \eta(t)\mathrm{d}W_r(t)\ ,\ r(t_0) = r_0, \tag{2.1}$$

where $\lambda(t)$ is a time-dependent mean reversion term, $\theta(t)$ is a time-dependent drift term, $\eta(t)$ is a time-dependent volatility term and $W_r(t) := W_r^{\mathbb{Q}}(t)$ is the Brownian motion under measure $\mathbb{Q}$.

We woule like to obtain an expression for $r(t)$. Using Itô's lemma, with $y(t) = e^{\int_{t_0}^t \lambda(t)\mathrm{d}t} r(t)$ gives us:

$$\begin{aligned}\mathrm{d}y(t) &= \frac{\partial\left(\mathrm{e}^{\int_{t_0}^t \lambda(t)\mathrm{d}t} r(t)\right)}{\partial t}\mathrm{d}t + \frac{\partial\left(\mathrm{e}^{\int_{t_0}^t \lambda(t)\mathrm{d}t} r(t)\right)}{\partial r(t)}\mathrm{d}r(t) \\ &= \lambda(t)\mathrm{e}^{\int_{t_0}^t \lambda(t)\mathrm{d}t} r(t)\mathrm{d}t + \mathrm{e}^{\int_{t_0}^t \lambda(t)\mathrm{d}t}\mathrm{d}r(t).\end{aligned}$$

By substitution of the Hull-White dynamics we obtain

$$\mathrm{d}y(t) = \lambda(t)\theta(t)\mathrm{e}^{\int_{t_0}^t \lambda(t)\mathrm{d}t}\mathrm{d}t + \eta(t)\mathrm{e}^{\int_{t_0}^t \lambda(t)\mathrm{d}t}\mathrm{d}W_r(t)).$$

For ease of notation, we define the following:

$$E\big(\lambda(t), t\big) = E(t) := \mathrm{e}^{\int_{t_0}^t \lambda(t)\mathrm{d}t}. \tag{2.2}$$

Integrating gives the solution for $y(t)$:

$$y(t) = y(t_0) + \int_{t_0}^t \lambda(z)\theta(z)E(z)\mathrm{d}z + \int_{t_0}^t \eta(z)E(z)\mathrm{d}W_r(z). \tag{2.3}$$

Substitution of $y(t)$ gives the result:

$$r(t) = \frac{r_0}{E(t)} + \frac{1}{E(t)}\int_{t_0}^t \lambda(z)\theta(z)E(z)\mathrm{d}z + \frac{1}{E(t)}\int_{t_0}^t \eta(z)E(z)\mathrm{d}W_r(z). \tag{2.4}$$

From the expression of $r(t)$ in Equation (2.4) we find that $r(t)$ is normally distributed. The mean is given by:

$$\mathbb{E}\big[r(t)\big|\mathscr{F}_{t_0}\big] = \frac{r_0}{E(t)} + \frac{1}{E(t)}\int_{t_0}^t \lambda(z)\theta(z)E(z)\mathrm{d}z. \tag{2.5}$$

The variance can be obtained with help of Itô isometry:

$$v_r\big(\lambda(t),\eta(t),t\big) = v_r(t) := \mathrm{Var}\big[r(t)\big|\mathscr{F}_{t_0}\big] = \frac{1}{E^2(t)}\int_{t_0}^{t}\eta^2(z)E^2(z)\mathrm{d}z. \tag{2.6}$$

The Hull-White model belongs to the class of affine processes [25]. Therefore, the price of a zero coupon bond has the following form:

$$P(t,T) = \mathrm{e}^{A(t,T)-B(t,T)r(t)}. \tag{2.7}$$

The dynamics of the ZCB under Hull-White and risk-neutral measure are given by [24]:

$$\frac{\mathrm{d}P(t,T)}{P(t,T)} = r(t)\mathrm{d}t + \eta(t)B(t,T)\mathrm{d}W_r(t). \tag{2.8}$$

The functions $A(t,T)$ and $B(t,T)$ satisfy the so colled Ricatti ordinary diferential equations. By solving this system of Ricatti ordinary differential equations, one can find the solutions for $A(t,T)$ and $B(t,T)$. For the Hull-White model with time-dependent parameters, $A(t,T)$ and $B(t,T)$ are given [24] by:

$$B\big(\lambda(t),t,T\big) = B(t,T) := E(t)\int_t^T \frac{1}{E(u)}\mathrm{d}u, \tag{2.9}$$

$$A\big(\lambda(t),\eta(t),t,T\big) = A(t,T) := \log\Big(\frac{P(0,T)}{P(0,t)}\Big) + B(t,T)f^r(0,t) - \frac{1}{2}B^2(t,T)v_r(t,t_0), \tag{2.10}$$

with $v_r(t,t_0)$ the variance of the short rate process as in Equation (2.6) and $f^r(0,t)$ the instantaneous forward rate as in Definition (1.1.6). One of the goals in this thesis is to find functions $\lambda(t),\theta(t)$ and $\eta(t)$ such that it matches market data. The function $\theta(t)$ can be derived from the zero coupon bond curve or yield curve. In the Hull-White model $\theta(t)$ is given [24] by:

$$\theta(t) = \frac{1}{\lambda(t)}\frac{\partial}{\partial t}f^r(0,t) + f^r(0,t) + \frac{1}{2}\Big(\frac{1}{\lambda(t)}\frac{\partial^2}{\partial t^2}Y(0,t) + \frac{\partial}{\partial t}Y(0,t)\Big), \tag{2.11}$$

where $f^r(0,t)$ is the instantaneous forward rate as in Definition 1.1.6 and $Y(t,T)$ is given by:

$$Y(t,T) = \int_t^T \eta^2(u)B^2(u,T)\mathrm{d}u.$$

2.1.1. Negative mean reversion

During calibration of the time-dependent Hull-White model, we can encounter negative values for the mean reversion parameter $\lambda(t)$. In this thesis, we are modelling the Hull-White model with a time-dependent mean reversion parameter $\lambda(t)$. In recent research on calibration of the mean reversion parameter, there seems to be a general consensus that negative mean reversion parameter should be allowed [33]. However, if $\lambda(t)$ is negative for all t, we do obtain some mathematical problems. For constant $\lambda(t)$ the expectation of $r(t)$ is given by:

$$\mathbb{E}[r(t)|\mathscr{F}(t_0)] = r_0\mathrm{e}^{-\lambda t} + \lambda\int_{t_0}^{t}\theta(z)\mathrm{e}^{-\lambda(t-z)}\mathrm{d}z. \tag{2.12}$$

Note that $r_0\mathrm{e}^{-\lambda t}$ goes to ∞ or $-\infty$ if $t\to\infty$ for negative λ, depending on r_0 being positive or negative. Therefore, the dynamics of the interest rate seem to explode for negative constant λ if $\lambda\to\infty$. Also, the variance given in Equation (2.6) is unstable for $t\to\infty$, since $E(t)\to 0$ for $t\to\infty$ with negative mean reversion $\lambda(t)$. The mean reversion parameter determines how fast the dynamics will follow the long term mean $\theta(t)$. If we have a negative constant λ we will move away from the equilibrium, and by each step, moving away from the long term mean $\theta(t)$. Therefore, intuitively and the extreme case $t\to\infty$ indicate that a negative mean reversion is not suitable.

Besides the mean of the Hull-White process, we can also experience problems in the computation of zero coupon bonds as in Equation (2.7). For large $\lambda(t)$, we find that $E(t)$ as in Equation (2.2) is close to zero. Consequently, $|A(t,T)-B(t,T)r(t)|$ can become very large. Therefore, we have unstable behaviour of the analytical expression of the zero coupon bond. In Figure 2.1, we can observe this unstable behaviour for a negative time-dependent mean reversion parameter.

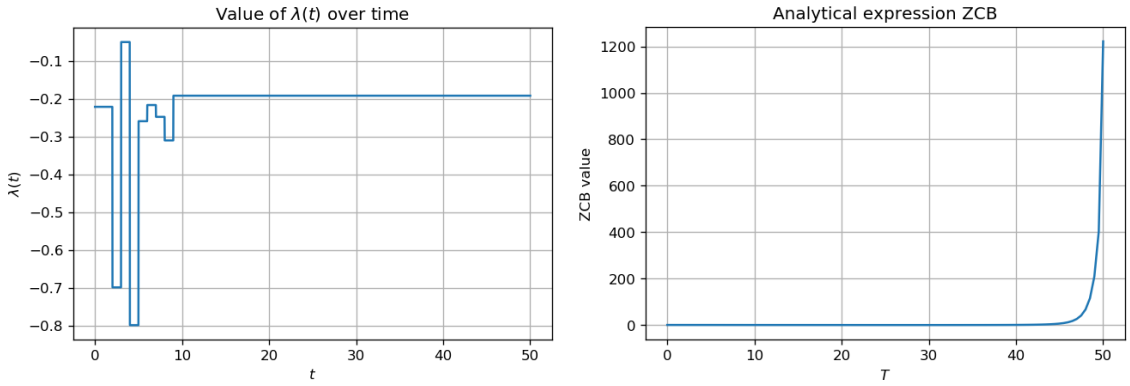

Figure 2.1: On the left we can find a calibrated $\lambda(t)$ to USD swaption data. This calibration can be found in the appendix in Figure C.3. On the right we can find the $P(0,T)$ computed by the expression in Equation (2.7). A constant volatility parameter $\eta = 0.01$ has been picked. Discount curve is observed from the market.

Although, it is against the nature of mean reversion, "past and recent developments all suggest that the use of negative mean reversion values is indispensable" [33]. Also in [24] negative mean reversion parameter is allowed. There seems to be some market consensus that the calibration benefits of allowing a negative mean reversion parameter outweighs the negative aspects of a negative mean reversion parameter in certain cases. Although, one caveat has to be made. Extrapolating the Hull-White model beyond the calibration regime can easily lead to overestimation of exposures and CVAs for longer-dated trades and portfolios [33].

Although it is theoretically hard to justify in the limit cases and for large negative mean reversion parameters we obtain problems in the analytical expression of the ZCB, for this thesis we tend to accept negative mean reversion parameters for negative $\lambda(t)$ if the negative values occur for a short period of time and do not occur frequently, since the improved calibration outweighs the negative aspects and the resulting calibration by the Hull-White model may outperform an alternative model.

2.1.2. Hull-White model under T-forward measure

In this thesis it is important to consider the Hull-White model under the T-forward measure, as working under the T-forward measure is advantageous when pricing interest rate products, e.g. swaptions. The Radon-Nikodym derivative $\lambda_{\mathbb{Q}^T}$ to change measure from risk-neutral to T-forward is given by:

$$\lambda_{\mathbb{Q}^T}(t) = \frac{P(t,T)}{P(t_0,T)}\frac{M(t_0)}{M(t)}. \tag{2.13}$$

Then by Itô's lemma, the dynamics of the zero coupon bond from Equation (2.8) and the Definition 1.1.2 of the money-savings account $M(t)$, we obtain:

$$\begin{aligned} \mathrm{d}\lambda^T_{\mathbb{Q}}(t) &= \frac{M(t_0)}{P(t_0,T)}\Big(\frac{1}{M(t)}\mathrm{d}P(t,T) - \frac{P(t,T)}{M^2(t)}\mathrm{d}M(t)\Big) \\ &= \frac{M(t_0)}{P(t_0,T)}\frac{P(t,T)}{M(t)}\eta(t)B(t,T)\mathrm{d}W^{\mathbb{Q}}(t), \\ \Rightarrow \frac{\mathrm{d}\lambda^T_{\mathbb{Q}}(t)}{\lambda^T_{\mathbb{Q}}(t)} &= \eta(t)B(t,T)\mathrm{d}W^{\mathbb{Q}}(t). \end{aligned}$$

with $B(t,T)$ as defined in Equation (2.9). We obtain for the Girsanov kernel [25]:

$$\mathrm{d}W^{\mathbb{Q}}(t) = \eta(t)B(t,T)\mathrm{d}t + \mathrm{d}W^T(t). \tag{2.14}$$

Applying this change of measure to the Hull-White dynamics gives us:

$$\mathrm{d}r(t) = \lambda(t)(\theta(t) + r(t))\mathrm{d}t + \eta(t)\big(\eta(t)B(t,T)\mathrm{d}t + \mathrm{d}W^T(t)\big) \tag{2.15}$$

$$= \lambda(t)\Big(\hat{\theta}(t) + r(t)\Big)\mathrm{d}t + \eta(t)\mathrm{d}W^T(t). \tag{2.16}$$

With $\hat{\theta}(t) = \theta(t) + \frac{\eta^2(t)B(t,T)}{\lambda(t)}$. Zero coupon bond dynamics under the T-forward measure are given by:

$$\begin{aligned}\frac{\mathrm{d}P(t,T)}{P(t,T)} &= r(t)\mathrm{d}t + \eta(t)B(t,T)\Big(\eta(t)B(t,T)\mathrm{d}t + \mathrm{d}W^T(t)\Big)\\ &= \Big(r(t) + \eta^2(t)B^2(t,T)\Big)\mathrm{d}t + \eta(t)B(t,T)\mathrm{d}W^T(t).\end{aligned}$$

2.1.3. Options on zero-coupon bonds under the Hull-White model

In this section we will derive how to price an option on a zero coupon bond. We will need this for the pricing of swaptions. Note that in this case we have two maturity times. Maturity time T_i of the option and maturity time T_m of the ZCB. We have that $T_i < T_m$. We have the following pricing of the option on a ZCB:

$$V^{ZCB}(t_0,T_i) = \mathbb{E}^{\mathbb{Q}}\Big[\frac{M(t_0)}{M(T)}\max(\alpha(P(T_i,T_m)-K),0))\Big|\mathscr{F}(t_0)\Big],$$

where $\alpha = 1$ for a call and $\alpha = -1$ for a put option.

Theorem 2.1.1. The value of a put option with strike price K and expiry T_i on a zero coupon bond with maturity T_m can be calculated by the following formula

$$\frac{V_p^{ZCB}(t_0,T_i)}{P(t_0,T_i)} = \mathrm{e}^{A(T_i,T_m)}\Big(\hat{K}\Phi(d_2) - \mathrm{e}^{B(T_i,T_m)\hat{\mu}_r(T_i)+\frac{1}{2}B^2(T_i,T_m)\hat{v}_r^2(T_i)}\Phi\big(d_1\big)\Big). \tag{2.17}$$

with $d_2 = \frac{\log(\hat{K})-B(T_i,T_m)\hat{\mu}(T_i)}{B(T_i,T_m)\hat{v}_r(T_i)}$, $d_1 = d_2 - B(T_i,T_m)\hat{v}_r(T_i)$, $\hat{\mu}(t) = \mathbb{E}^{T_i}\big[r(t)\big|\mathscr{F}_{t_0}\big]$ as in Equation (2.18) , $\hat{v}_r^2(t) = Var^{T_i}\big[r(t)\big|\mathscr{F}_{t_0}\big]$ as in Equation (2.19), and $\Phi(\cdot)$ the standard normal CDF and $\hat{K} = K\mathrm{e}^{-A(t_i,t_m)}$.

Proof. We follow the proof as in [25]. By a measure change from the risk-free measure to the T_i-forward measure, the pricing equation is given by:

$$V_p^{ZCB}(t_0,T_i) = P(t_0,T)\mathbb{E}^{T_i}\Big[\max(\bar{\alpha}(P(T_i,T_m)-K),0))\Big|\mathscr{F}(t_0)\Big].$$

When we are dealing with affine short-rate models, we can rewrite the zero coupon bond $P(t_0,T)$ to obtain the following:

$$\begin{aligned}V_p^{ZCB}(t_0,T_i) &= P(t_0,T_i)\mathbb{E}^{T_i}\Big[\max(\alpha(\mathrm{e}^{A(T_i,T_m)+B(T_i,T_m)r(T_i)}-K),0))\Big|\mathscr{F}(t_0)\Big]\\ &= P(t_0,T_i)\mathrm{e}^{A(T_i,T_m)}\mathbb{E}^{T_i}\Big[\max(\alpha(\mathrm{e}^{B(T_i,T_m)r(T_i)}-\hat{K}),0))\Big|\mathscr{F}(t_0)\Big].\end{aligned}$$

Where $A(T_i,T_m)$ and $B(T_i,T_m)$ are given in respectively Equation (2.10) and Equation (2.9). Under the T_i forward measure, we have that $r(t)$ is normally distributed with the following expectation and variance, which can be deducted from equation (2.16):

$$\hat{\mu}(t) := \mathbb{E}^{T_i}\big[r(t)\big|\mathscr{F}_{t_0}\big] = \frac{r_0}{E(t)} + \frac{1}{E(t)}\int_{t_0}^{t}\lambda(z)\hat{\theta}(z)E(z)\mathrm{d}z. \tag{2.18}$$

$$\hat{v}_r^2(t) := Var^{T_i}\big[r(t)\big|\mathscr{F}_{t_0}\big] = \frac{1}{E^2(t)}\int_{t_0}^{t}\eta^2(z)E^2(z)\mathrm{d}z. \tag{2.19}$$

Knowing that $r(t)$ is normally distributed with mean $\hat{\mu}(t)$ and variance $v_r^2(t)$

$$z(T_i) := B(T_i,T_m)r(T_i) \sim \mathscr{N}\Big(B(T_i,T_m)\hat{\mu}(T_i), B^2(T_i,T_m)\hat{v}_r^2(T_i)\Big). \tag{2.20}$$

Let $\alpha = -1$, we derive the pricing formula for a put option. We can rewrite the value of the European option on a ZCB to:

$$\begin{aligned}V_p^{ZCB}(t_0,T_i) &= P(t_0,T_i)e^{A(T_i,T_m)}\mathbb{E}^{\mathbb{Q}}\Big[\max((\hat{K}-\mathrm{e}^{z(T_i)}),0))\Big|\mathscr{F}(t_0)\Big]\\ &= P(t_0,T_i)e^{A(T_i,T_m)}\mathbb{E}^{\mathbb{Q}}\Big[(\hat{K}-\mathrm{e}^{z(T_i)})1_{\{\hat{K}-e^{z(T_i)}>0\}}\Big|\mathscr{F}(t_0)\Big].\end{aligned}$$

Let Z be a standard normal variable. Rewriting the subscript of the indicator function gives us:

$$\hat{K} - \mathrm{e}^{z(T_i)} > 0 \Longleftrightarrow B(T_i,T_m)\hat{\mu}(T_i) + B(T_i,T_m)\hat{v}_r(T_i)Z \leq \log(\hat{K}) \Longleftrightarrow Z \leq \frac{\log(\hat{K}) - B(T_i,T_m)\hat{\mu}(T_i)}{B(T_i,T_m)v_r(T_i)} =: d_2,$$

where $f_{\mathcal{N}(0,1)}(x)$ is the standard normal PDF and for the upper integral range we have $d_2 = \frac{\log\hat{K} - B(t,T)\bar{\mu}(t_0,T)}{B(t,T)v_r(T)}$. Let $\Phi(\cdot)$ be the standard normal CDF. By splitting the integral, we obtain:

$$\begin{aligned}
\frac{V_p^{ZCB}(t_0,T_i)}{P(t_0,T_i)} &= \mathrm{e}^{A(T_i,T_m)}\hat{K}\Phi(d_2) - \mathrm{e}^{A(T_i,T_m)}\int_{-\infty}^{d_2}\Big(\mathrm{e}^{B(T_i,T_m)\hat{\mu}_r(T)+B(T_i,T_m)\hat{v}_r(T)x}\Big)f_{\mathcal{N}(0,1)}(x)\mathrm{d}x \\
&= \mathrm{e}^{A(T_i,T_m)}\hat{K}\Phi(d_2) - \mathrm{e}^{A(T_i,T_m)+B(T_i,T_m)\hat{\mu}_r(T_i)+\frac{1}{2}B^2(T_i,T_m)\hat{v}_r^2(T_i)}\frac{1}{2\pi}\int_{-\infty}^{d_2}\mathrm{e}^{-\frac{\left(x-B(T_i,T_m)\hat{v}_r(T_i)\right)^2}{2}}\mathrm{d}x \\
&= \mathrm{e}^{A(T_i,T_m)}\Big(\hat{K}\Phi(d_2) - \mathrm{e}^{B(T_i,T_m)\hat{\mu}_r(T_i)+\frac{1}{2}B^2(T_i,T_m)\hat{v}_r^2(T_i)}\Phi\big(d_2 - B(T_i,T_m)\hat{v}_r(T_i)\big)\Big).
\end{aligned}$$

The pricing of a call option can be derived in a similar method. □

2.1.4. Swaptions under the Hull-White model

Under the Hull-White model we can derive a semi-analytical pricing of a swaption by Jamshidian's trick [14]. At time t_0 the value of a swaption with expiry T_i and tenor T_m is given by

$$V^{swpt}(t_0) = N\cdot\mathbb{E}^{\mathbb{Q}}\Big[\frac{M(t_0)}{M(T_i)}\max\Big(V^{Swap}(t_0,T_i,T-m),0\Big)\Big|\mathscr{F}(t_0)\Big].$$

After a change of measure from $\mathbb{Q}$ to the T_i-forward measure, we obtain for the value of the swaption

$$V^{swpt}(t_0) = N\cdot P(t_0,T_i)\cdot\mathbb{E}^{T_i}\Big[\max\Big(V^{Swap}(t_0,T_i,T_m),0\Big)\Big|\mathscr{F}(t_0)\Big].$$

Using Theorem 1.2.2, we have for the value of a payer swap

$$V^{Swap}(t,T_i,T_m) = N(P(t,T_i) - P(t,T_m)) - NK\sum_{k=i+1}^{m}\tau_k P(t,T_k),$$

and we obtain:

$$\begin{aligned}
V^{Swap}(T_i,T_i,T_m) &= 1 - P(T_i,T_m) - K\sum_{k=i+1}^{m}\tau_k P(T_i,T_k) \\
&= 1 - \sum_{k=i+1}^{m}c_k P(T_i,T_k),
\end{aligned}$$

with $c_k = K\tau_k$ for $k = i+1,\ldots,m-1$ and $c_m = 1 + K\tau_m$. By substitution, we obtain:

$$V^{swpt}(t_0) = N\cdot P(t_0,T_i)\cdot\mathbb{E}^{T_i}\Big[\max\Big(1 - \sum_{k=i+1}^{m}c_k P(T_i,T_k),0\Big)\Big|\mathscr{F}(t_0)\Big].$$

The ZCB $P(T_i,T_k)$ can be expressed as $P(T_i,T_k) = \exp(A(T_i,T_k) - B(T_i,T_k)r(T_i))$ as shown in Equation (2.7). We substitute and find:

$$V^{swpt}(t_0) = P(t_0,T_i)\cdot N\cdot\mathbb{E}^{T_i}\Big[\max\Big(1 - \sum_{k=i+1}^{m}c_k\mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r(T_i)},0\Big)\Big|\mathscr{F}(t_0)\Big]. \tag{2.21}$$

We use the property derived in [14], also known as Jamshidian's trick. we can change the maximum of the sum to the sum of the maximum. Let $\psi_k(r)$ a monotonically increasing sequence, where $\psi_k(r), \mathbb{R}\to\mathbb{R}^+$. We want to evaluate the following:

$$A = \max\Big(K - \sum_k \psi_k(r), 0\Big). \tag{2.22}$$

Note that $\sum_k \psi_k(r)$ is monotonically increasing, since $\psi_k(r)$ is monotonically increasing. Then there must be a value r^* such that

$$K = \sum_k \psi_k(r^*).$$

Now one can substitute K in Equation (2.22) and after some computations we obtain:

$$\begin{aligned} A &= \max\Big(\sum_k \psi_k(r^*) - \sum_k \psi_k(r), 0\Big) \\ &= \max\Big(\sum_k \big(\psi_k(r^*) - \psi_k(r)\big), 0\Big) \\ &= \sum_k \big(\psi_k(r^*) - \psi_k(r)\big) 1_{r>r^*} \\ &= \sum_k \max\big(\psi_k(r^*) - \psi_k(r), 0\big). \end{aligned}$$

We changed the maximum of the sum to the sum of the maximum. This trick can be applied at the maximum of sum in expression (2.21). Choose $\psi_k(r^*) = c_k \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r^*}$. Note that this is a monotone function, since $B(T_i,T_k)$ and c_k are positive for all k. We obtain:

$$\begin{aligned} \max\Big(1 - \sum_{k=i+1}^{m} c_k e^{A(T_i,T_k)-B(T_i,T_k)r(T_i)}, 0\Big) &= \max\Big(\sum_{k=i+1}^{m} c_k \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r^*} - \sum_{k=i+1}^{m} c_k \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r(T_i)}, 0\Big) \\ &= \sum_{k=i+1}^{m} c_k \max\Big(\mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r^*} - \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r(T_i)}, 0\Big) \\ &= \sum_{k=i+1}^{m} c_k \max\Big(\hat{K}_k - \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r(T_i)}, 0\Big), \end{aligned}$$

where $\hat{K}_k = \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r^*}$ and r^* can be found by solving the following equation:

$$1 - \sum_{k=i+1}^{m} c_k \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r^*} = 0.$$

We obtain as value for the swaption:

$$V^{swpt}(t_0) = P(t_0,T_i) \cdot N \cdot \sum_{k=i+1}^{m} c_k \mathbb{E}^{T_i}\Big[\max\Big(\hat{K}_k - \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r(T_i)}, 0\Big)\Big|\mathscr{F}(t_0)\Big].$$

We still have to determine the sum of the expectation. Note that it is a sum of European put options on zero-coupon bonds.

$$V_p^{ZCB}(t_0,T_i,T_k) = P(t_0,T_i)\mathbb{E}^{T_i}\Big[\max\Big(\hat{K}_k - \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r(T_i)}, 0\Big)\Big].$$

This pricing of European-type options on the ZCB under the Hull-White model gives us a closed-form solution.

Theorem 2.1.2. The value of an European swaption with expiry T_i and maturity T_m under the Hull-White model is given by

$$V^{swpt}(t_0) = N \cdot \sum_{k=i+1}^{m} c_k V_p^{ZCB}(t_0,T_i,T_k),$$

where the the strike price $\hat{K}_k; = \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r^*}$ and r^* is determined by solving the following equation

$$\sum_{k=i+1}^{m} c_k \mathrm{e}^{A(T_i,T_k)-B(T_i,T_k)r^*} = 1,$$

with $c_k = K\tau_k$ for $k = i+1, ..., m-1$ and $c_m = 1 + K\tau_m$.

The value of r^* can be obtained by solving the following optimization problem:

$$\arg\min_{r^*\in\mathbb{R}}\left(\sqrt{\left(1-\sum_{k=i+1}^{m} c_k e^{A(T_i,T_k)-B(T_i,T_k)r^*}\right)^2}\right).$$

This optimization problem is a convex optimzation problem, which can be solved by iterative methods such as Newton-Raphson. For these iterative algorithms, an initial guess of r^* is needed. In order to obtain a proper chosen initial guess, we perform a trick by interpolation on the inverse. The method is described in Algorithm 3.

Algorithm 1: Method for finding proper initial guess.

- Plug in a grid $\bar{x}$ for r^* in $\sum_{k=i+1}^{m} c_k e^{A(T_i,T_k)-B(T_i,T_k)r^*}$. We obtain the grid $\bar{y}$:

$$\bar{y} := \sum_{k=i+1}^{m} c_k e^{A(T_i,T_k)+B(T_i,T_k)\bar{x}}$$

- interpolate $\bar{y}$ and $\bar{x}$ to find an interpolated value $\hat{x} \in \bar{x}_{interpolated}$ where the corresponding interpolated $\hat{y} \in \bar{y}_{interpolated}$ is such that $\hat{y} = 1$.

- **Return:** $\hat{x}$.

2.2. Yield curve construction

Construction of the yield curve is an essential part for pricing of financial products. In the market one can find quotes for financial products, e.g. swaps, caps, etc. From these financial products one can get the zero coupon bond values. These quotes are used for the construction of the yield curve. We compute the time 0 continuously compounded risk-free rate $r_0(t)$ by the following relationship with zero coupon bonds [8]:

$$r_0(t) = -\frac{1}{t}\log\big(P(0,t)\big). \tag{2.23}$$

By the relationship between ZCB and $r_0(t)$ in Equation (2.23), we obtain a set of datapoints for $r_0(t)$. We can interpolate the $r_0(t)$ data points to obtain the curve for $r(t)$. From there, we once more use the relation in Equation (2.23) to create the zero coupon bond curve, which is needed for our computations. There are multiple ways of interpolating these points. All with different advantages and drawbacks. In [8], various methods for the construction of the yield curve are described: linear on rates, linear on the log of rates, linear on discount factors, raw interpolation, piecewise linear forward, quardrating splines, cubic splines or monotone convex.

Hagan and West [8] argue that the monotone convex method is the best method for the construction of a yield curve. However, the monotone convex method does not create a smooth curve, only continuous. This is a serious drawback of this method for our research, because the instantaneous forward rate is defined as a derivative on the zero coupon bond curve:

$$f^r(0,T) := -\frac{\partial}{\partial T}\log\big(P(0,T)\big). \tag{2.24}$$

We are not able to compute the instantaneous forward rate properly when the zero coupon bond curve is not smooth. This is poor for our applications of the zero coupon bond curve:

- For Monte Carlo simulation of Hull-White data, we compute $\theta(t)$

$$\theta(t) = \boxed{\frac{1}{\lambda(t)}\frac{\partial}{\partial t}f^r(0,t)} + f^r(0,t) + \frac{1}{2}\Big(\frac{1}{\lambda(t)}\frac{\partial^2}{\partial t^2}Y(0,t) + \frac{\partial}{\partial t}Y(0,t)\Big), \tag{2.25}$$

  Highlighted in Equation (2.25) is the derivative of the instantaneous forward rate. In order to compute this, we need the instantaneous forward rate to be smooth.

- For the calibration of the Hull-White model, we use theorem 2.1.2. In this procedure, we need the instantaneous forward rate $f^r(0,T)$. Therefore, we require the zero coupon bond curve to be smooth. We do not need the instantaneous forward rate to smooth.

Looking at the results provided by Hagan ans West [8], an optimal choice for calibration that does provide a smooth curve is Hermite cubic spline interpolation. However, under this method another problem arises when we compute by means of Monte Carlo methods. For the computation of $\theta(t)$ in Equation (2.25), we need a derivative of the instantaneous forward rate. Therefore, we need the derivative of zero coupon bond curve to be smooth. It appears that this is not the case for cubic spline interpolation.

A method that does provide us with a smooth instantaneous forward rate is interpolation by Bezier curve [9]. The Bezier curve $B(t)$ for interpolation of data points P_i is given by:

$$B(t) = \sum_{i=0}^{n} \binom{n}{i} t^i (1-t)^{(n-i)} P_i, \tag{2.26}$$

where P_i represents data point i for $i \in \{0, 1, ..., n\}$ and $0 \leq t \leq 1$. Although the Bezier interpolation method provides a smooth instantaneous forward rate, it also comes with a drawback. The Bezier curve does not pass through all data points, as can be observed in Figure 2.2. Because of this drawback, the Bezier curve is not considered as an optimal choice for the construction of the market yield curve. Using the Bezier curve or cubic spline curve gives different results for the calibration. In this thesis the cubic spline interpolation is used for the calibration, and Bezier curve is used for validation of the calibration method by means of Monte Carlo.

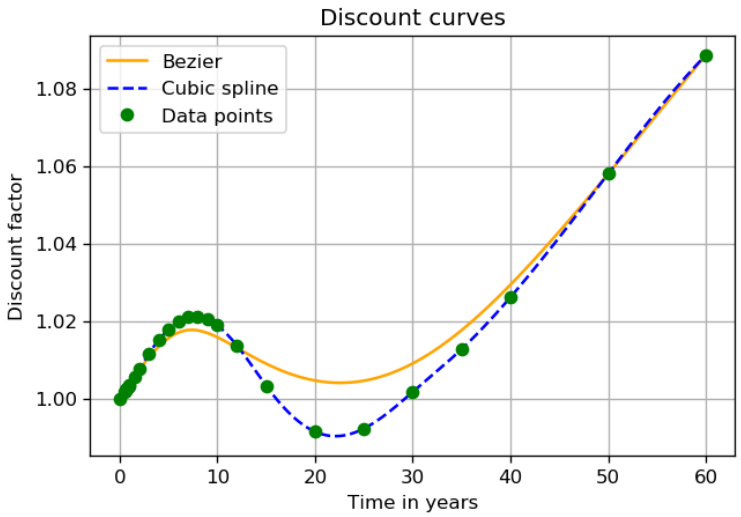

Figure 2.2: Interpolation of EUR market data by cubic spline and Bezier curve. Data is retrieved from Bloomberg July 2020.

2.3. Multi-curve framework

Before the crisis in 2008, interest rate products were priced by the single curve method. Afterwards, the multi-curve framework has been introduced. The single curve was not qualified for pricing in the post-crisis market environment. In this section the multi-curve framework will be explained.

Consider a swap with the interchanges of interest payments are set to happen on T_0, T_1, ..., T_m. The payments of swaps are done on a predetermined frequency, typically dependent on the type of product. For EUR swaptions, the underlying swap usually interchanges interest payments every 6 months, whilst for USD swaptions this is typically done every 3 months. For an EUR swaption we have $T_{i+1} - T_i = 6m$ and for USD swaption this is $T_{i+1} - T_i = 3m$. This structure of dates is called the tenor structure.

A basis swap is a swap with two float legs that have different payment frequency. Under the risk-neutral measure and neglecting risk of default, the price of a basis swap is given by [25]:

$$V_{basisswap}(t_0) = N \cdot M(t_0) \mathbb{E}^{\mathbb{Q}} \left[\sum_{k_1=i+1}^{m_1} \frac{\tau_{k_1}}{M(T_{k_1})} \ell_{k_1}(T_{k_1-1}) - \sum_{k_2=i+1}^{m_2} \frac{\tau_{k_2}}{M(T_{k_2})} \ell_{k_2}(T_{k_2-1}) \right] = 0. \tag{2.27}$$

In reality, this basis swap value (spread) is not zero. It is safer to receive payments more frequently, as the probability of default of the counterparty is never zero. In the market the basis swap spread is visible. In Figure 2.3 the development of the basis swap spread during the crisis of 2008 is shown. Before 2007 this

spread was close to zero, but during the crisis the probability of default of counterparties began to rise, and as a consequence the basis swap spread increased.

Figure 2.3: For a swap with maturity one the basis swap spread is shown for six months against three months EURIBOR rates. The figure is obtained from Bloomberg in [26].

An example is given in order to obtain a more intuitive understanding why risk of default cannot be neglected.

Example 2.3.1. Consider two contracts with the same counterparty. contract A pays two euros after two years, and contract B pays you one euro after one year, and another euro after one more year. Both contracts pay you 2 euros after 2 years. Assuming an interest rate of 0, both contracts would have the same value, when we neglect the probability of default of the counterparty. When we take the probability of default of the counterparty into consideration, contract A is cheaper than contract B. Let $t_{Default}$ be a random variable indicating the first time that a default of the counterparty occurs. We have for the value of contract A $V_A(t)$ and contract B $V_B(t)$:

$$V_A(t_0) = \mathbb{E}^{\mathbb{Q}}\left[2 \cdot 1_{t_{Default} > t_0+2}\right] = 2 \cdot P\left(t_{Default} > t_0 + 2\right),$$
$$V_B(t_0) = \mathbb{E}^{\mathbb{Q}}\left[1 \cdot 1_{t_{Default} > t_0+1} + 1 \cdot 1_{t_{Default} > t_0+2}\right] = P\left(t_{Default} > t_0 + 1\right) + P\left(t_{Default} > t_0 + 2\right).$$

Since $P\left(t_{Default} > t_0 + 1\right) \geq P\left(t_{Default} > t_0 + 2\right)$, we have that $V_A(t) \leq V_B(t)$.

The single curve framework, that was used before 2008, did not take into account counterparty risk properly. For the discounting of different tenor structures, the single curve framework uses one and the same discount curve for the discounting. When using the the single curve framework, the information contained in the Basis swap spread is neglected. Perhaps before the crisis, this was acceptable, as these spreads were very small [2]. Post-crisis, the multi-curve framework was constructed. In the multi-curve framework every swaption with different tenor structure of the underlying swap is discounted by the corresponding discount curve with the same tenor structure [18]. In Figure 2.4 the EUR 1 day discount curve and EUR 6 months discount curve is shown. Notice that difference (the spread) between the two curves is significant, we cannot neglect it.

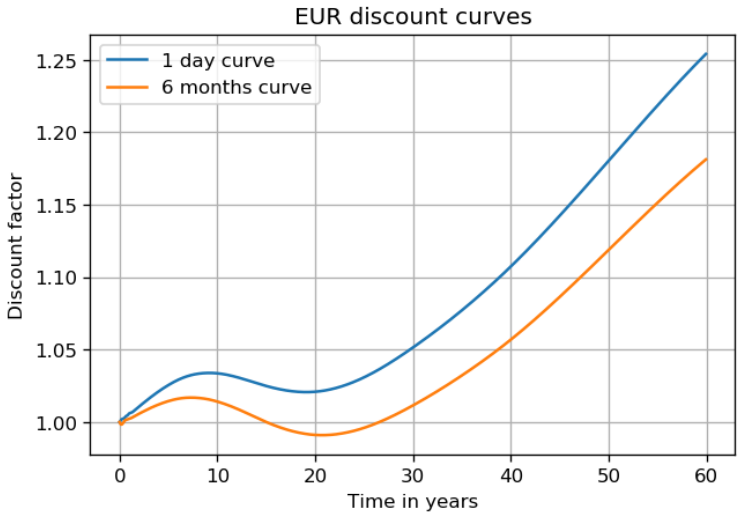

Figure 2.4: The 1 day (1d) discount curve and 6 months (6m) discount curve is shown for euros is shown. Data is extracted from Bloomberg in May 2020.

3

Calibration of the Hull-White model

By calibrating a model to market data, we are talking about finding a set of parameters such that the theoretical price reproduces the price observed in the market. Oftentimes, the market shows implied volatility instead of the prices of the financial products. For swaptions this is the case. In that case one will calibrate to the implied volatilities shown in the market. In this section the calibration of the single-factor Hull-White model is looked into. Our goal is to find optimal $\lambda(t)$, $\theta(t)$ and $\eta(t)$ for a match with market data.

In the market we observe quotes from various financial products. From these products, a yield curve is constructed. A yield curve represents today's interest rate for any maturity [25]. From the yield curve, the zero coupon bond curve can be created, in which one can observe the value of zero coupon bonds for any maturity. In the Hull-White model the time-dependent long-term average $\theta(t)$ can be expressed in terms of the mean reversion $\lambda(t)$, $\eta(t)$ and the zero coupon bond curve as in Equation (2.11). As the zero coupon bond curve can be obtained from the market, one will have the values for $\theta(t)$, once $\lambda(t)$ and $\eta(t)$ are known. Therefore, one merely has to calibrate $\lambda(t)$ and $\eta(t)$. In this chapter we discuss how to calibrate $\lambda(t)$ and $\eta(t)$ and show results. At the end we will look into the impact on xVA.

3.1. Calibration to swaptions

The first step of calibration is choosing financial products to calibrate to. Caps, floors or swaptions are generally used for the calibration of the Hull-White model. Caps, floors and swaptions are liquid products that contain the required information [36]. In fact, Swaptions are most commonly used for the calibration of the Hull-White model and can be considered as the optimal choice for calibration, because swaptions contain information about correlation between different points on the discount curve [36]. Therefore, in this thesis swaptions are used for the calibration.

As we are one step further by choosing financial products, we still need to specify the properties of the swaptions we will use. The market provides various swaptions with different properties, e.g. strike value and tenor frequency. The swaptions used for calibration in this thesis have the following properties:

1. The floating rate and fixed rate of the underlying swap of the swaption are paid at every tenor T_i .
2. The distance between tenors of the underlying swap of the swaption is the same, i.e. τ_k is the same for all k.
3. Strike price K is at the money.

The first two conditions make the modelling of swaptions easier. Besides, these swaptions are more liquid in the market than more exotic swaptions. Also, the third condition connects to this. At the money swaptions are more liquid than in the money or out of the money swaptions. Most important of all, optimization algorithms (root-finding methods) perform better when using at the money financial products [25]. For deep ITM or OTM options, the Greek vega is close to zero, which causes issues in iterative methods, e.g. Newton-Raphson. The last specification of swaptions we will be using for the calibration is the choice of expiries and tenors. This choice is more extensive and will be discussed in Section 3.4.

We state the calibration optimization problem. Our goal is to obtain the set of parameters $\Omega(t) = \big(\eta(t), \lambda(t)\big)$ such that our model gives the implied volatilities observed in the market. The optimization problem is given

by the following expression:

$$\Omega(t) = \arg\min_{\Omega(t)} \sqrt{\sum_i \sum_j \left(\sigma_{imp}^{mrkt}(t_0, T_i, T_j, K) - \sigma_{imp}^{HW}(t_0, T_i, T_j, K, \Omega(t))\right)^2} \tag{3.1}$$

The next step is choosing a method to solve the optimization problem. Before we get into this topic, a relaxation needs to made concerning the time-dependent parameters. In the next section we will go in detail in the relaxation where we assume $\lambda(t)$ and $\eta(t)$ to be piece-wise constant functions.

3.2. Piecewise constant functions

We would like to calibrate the Hull-White model using time-dependent parameters. We assume the time-dependent parameters to be piece-wise constant function. This is typically done when calibrating time-dependent functions [13]. In the Hull-White model we assume that $\eta(t)$ and $\lambda(t)$ are piece-wise constant functions. We may assume that the $T_1, ..., T_n$ for both $\eta(t)$ and $\lambda(t)$ are the same, then the piece-wise constant functions are given by:

$$\eta(t) = \begin{cases} \eta_0 & \text{if } 0 = T_0 \le t < T_1 \\ \eta_1 & \text{if } T_1 \le t < T_2 \\ \vdots & \\ \eta_{n-1} & \text{if } T_{n-1} \le t < T_n \\ \eta_n & \text{if } t \ge T_n \end{cases}, \lambda(t) = \begin{cases} \lambda_0 & \text{if } 0 = T_0 \le t < T_1 \\ \lambda_1 & \text{if } T_1 \le t < T_2 \\ \vdots & \\ \lambda_{n-1} & \text{if } T_{n-1} \le t < T_n \\ \lambda_n & \text{if } t \ge T_n \end{cases}. \tag{3.2}$$

$T_1, ..., T_n$ will correspond to the expiries or tenors of the swaptions used for the calibration, depending on the optimization problem. For this calibration, we have $2(n+1)$ parameters that we need to calibrate. One way to do this calibration is by taking at least $2(n+1)$ instruments to calibrate to. In this Thesis, we will focus on first calibrating the $n+1$ parameters of $\lambda(t)$, afterwards calibrating the $n+1$ parameters of $\eta(t)$.

In this thesis we will work with piece-wise functions only, therefore we can make some relaxations by replacing integrals with summations. Let $\Delta T_i = T_{i+1} - T_i$, we obtain:

$$E(T_k) = e^{\int_{t_0}^{T_k} \lambda(t) dt} = e^{\sum_{i=0}^{k-1} \lambda_i \Delta T_i}$$

Replacing the integral by a summation has positive effect on the computation time of the calibration. One can imagine, that during an iterative optimization process computing integrals over and over will be time consuming. Therefore, the replacement by summation is essential for an optimal computation time.

3.3. Implied volatility of swaption

In the market the swaptions are quoted on the implied volatility. By observing Theorem 1.2.4, we can find that the implied volatility of a swaption is related to the volatility of the swap rate, in the same way as an equity option to the volatility of the stock under the Black-Scholes model. Therefore, the implied volatility of the swaption is given by:

$$\sigma_{imp}(t_0, T_i) = \sqrt{\frac{1}{T_i - t_0} \int_{t_0}^{T_i} \sigma_{i,m}(t) dt},$$

where $\sigma_{i,m}$ is the volatility parameter of the swap rate dynamics. The idea is to approximate the volatility parameter of the swap rate, and look into the ratio of two approximated swaption implied volatilities with same expiry but different tenor. This ratio is independent of $\eta(t)$, hence it can be used for the calibration of $\lambda(t)$. Afterwards, we calibrate $\eta(t)$ to the implied volatilities using the calibrated $\lambda(t)$. In the first two subsections we consider the case where the swap rate follows a lognormal distribution, afterwards we will extend this to the shifted lognormal distribution in Section 3.3.3. The assumption of normally distributed swap rates is not covered in this thesis.

3.3.1. Swap rate volatility

The swap rate and annuity factor are given by:

$$S_{i,m}(t) = \frac{P(t, T_i) - P(t, T_m)}{A_{i,m}(t)}, \qquad A_{i,m}(t) = \sum_{k=i+1}^{m} \tau_k P(t, T_k). \tag{3.3}$$

The dynamics of the swap rate are given by Itô's lemma:

$$\begin{aligned}
\mathrm{d}\big(S_{i,m}(t)\big) =& \mathrm{d}\Big(\frac{P(t,T_i)}{A_{i,m}(t)}\Big) - \mathrm{d}\Big(\frac{P(t,T_m)}{A_{i,m}(t)}\Big) \\
=& \frac{1}{A_{i,m}(t)}\Big(\mathrm{d}\big(P(t,T_i)\big) - \mathrm{d}\big(P(t,T_m)\big)\Big) - \frac{S_{i,m}(t)}{A_{i,m}(t)}\mathrm{d}\big(A_{i,m}(t)\big) \\
&+ \frac{S_{i,m}(t)}{\big(A_{i,m}(t)\big)^2}\Big(\mathrm{d}\big(A_{i,m}(t)\big)\Big)^2 - \frac{1}{\big(A_{i,m}(t)\big)^2}\Big(\mathrm{d}\big(P(t,T_i)\big) - \mathrm{d}\big(P(t,T_m)\big)\mathrm{d}\big(A_{i,m}(t)\big)\Big).
\end{aligned}$$

The dynamics of zero coupon bond and annuity rate or combinations of those are given by:

$$\begin{aligned}
\mathrm{d}\big(P(t,T_i)\big) &= P(t,T_i)\big(r(t)\mathrm{d}t - \eta(t)B(t,T_i)\mathrm{d}W^{\mathbb{Q}}(t)\big), \\
\mathrm{d}\big(A_{i,m}(t)\big) &= \mathrm{d}\Big(\sum_{k=i+1}^{m}\tau_k P(t,T_k)\Big) = \sum_{k=i+1}^{m}\tau_k \mathrm{d}\big(P(t,T_k)\big) = \sum_{k=i+1}^{m}\tau_k P(t,T_k)\big(r(t)\mathrm{d}t - \eta(t)B(t,T_k)\mathrm{d}W^{\mathbb{Q}}(t)\big), \\
\Big(\mathrm{d}\big(A_{i,m}(t)\big)\Big)^2 &= \Big(\sum_{k=i+1}^{m}\tau_k P(t,T_k)\big(r(t)\mathrm{d}t + \eta(t)B(t,T_k)\mathrm{d}W^{\mathbb{Q}}(t)\big)\Big)^2 = \eta^2(t)\Big(\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)\Big)^2\mathrm{d}t, \\
\mathrm{d}\big(P(t,T_i)\big)\mathrm{d}\big(A_{i,m}(t)\big) &= P(t,T_i)\big(r(t)\mathrm{d}t - \eta(t)B(t,T_i)\mathrm{d}W^{\mathbb{Q}}(t)\big)\cdot\sum_{k=i+1}^{m}\tau_k P(t,T_k)\big(r(t)\mathrm{d}t - \eta(t)B(t,T_k)\mathrm{d}W^{\mathbb{Q}}(t)\big) \\
&= \eta^2(t)P(t,T_i)B(t,T_i)\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)\mathrm{d}t.
\end{aligned}$$

Looking into the parts we need for the whole equation:

$$\begin{aligned}
\frac{1}{A_{i,m}(t)}\Big(\mathrm{d}\big(P(t,T_i)\big) - \mathrm{d}\big(P(t,T_m)\big)\Big) &= S_{i,m}(t)r(t)\mathrm{d}t - \frac{\eta(t)}{A_{i,m}(t)}\Big(P(t,T_i)B(t,T_i) - P(t,T_m)B(t,T_m)\Big)\mathrm{d}W^{\mathbb{Q}}(t), \\
\frac{S_{i,m}(t)}{A_{i,m}(t)}\mathrm{d}\big(A_{i,m}(t)\big) &= S_{i,m}(t)r(t)\mathrm{d}t - \eta(t)\frac{\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{A_{i,m}(t)}\mathrm{d}W^{\mathbb{Q}}(t), \\
\frac{S_{i,m}(t)}{\big(A_{i,m}(t)\big)^2}\Big(\mathrm{d}\big(A_{i,m}(t)\big)\Big)^2 &= S_{i,m}(t)\eta^2(t)\frac{\Big(\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)\Big)^2}{\big(A_{i,m}(t)\big)^2}\mathrm{d}t, \\
\frac{1}{\big(A_{i,m}(t)\big)^2}\Big(\mathrm{d}\big(P(t,T_i)\big) - \mathrm{d}\big(P(t,T_m)\big)\Big)\mathrm{d}\big(A_{i,m}(t)\big) &= \eta^2(t)\Big(P(t,T_i)B(t,T_i) - P(t,T_m)B(t,T_m)\Big)\frac{\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{\big(A_{i,m}(t)\big)^2}\mathrm{d}t.
\end{aligned}$$

Substituting these dynamics, we find for the dynamics of the swap rate:

$$\begin{aligned}
\mathrm{d}\big(S_{i,m}(t)\big) =& \eta^2(t)\Big(\frac{S_{i,m}(t)\Big(\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)\Big)^2}{\big(A_{i,m}(t)\big)^2} - \Big(P(t,T_i)B(t,T_i) - P(t,T_m)B(t,T_m)\Big)\frac{\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{\big(A_{i,m}(t)\big)^2}\Big)\mathrm{d}t \\
&- \eta(t)\Big(\frac{\big(P(t,T_i)B(t,T_i) - P(t,T_m)B(t,T_m)\big)}{A_{i,m}(t)} - \frac{\big(P(t,T_i) - P(t,T_m)\big)\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{\big(A_{i,m}(t)\big)^2}\Big)\mathrm{d}W^{\mathbb{Q}}(t). \qquad (3.4)
\end{aligned}$$

Assuming that the swap rate follows a log normal distribution, we thus have:

$$\begin{aligned}
\frac{\mathrm{d}\big(S_{i,m}(t)\big)}{S_{i,m}(t)} =& \eta^2(t)\Big(\frac{\Big(\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)\Big)^2}{\big(A_{i,m}(t)\big)^2} - \Big(P(t,T_i)B(t,T_i) - P(t,T_m)B(t,T_m)\Big)\frac{\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{S_{i,m}(t)\big(A_{i,m}(t)\big)^2}\Big)\mathrm{d}t \\
&- \eta(t)\Big(\frac{\big(P(t,T_i)B(t,T_i) - P(t,T_m)B(t,T_m)\big)}{S_{i,m}(t)A_{i,m}(t)} - \frac{\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{A_{i,m}(t)}\Big)\mathrm{d}W^{\mathbb{Q}}(t). \qquad (3.5)
\end{aligned}$$

Notice that the swap rate $S_{i,m}(t)$ is a martingale under the the annuity measure where the numéraire is $A_{i,m}(t)$ by the definition of the swap rate $S_{i,m}(t)$. A measure change does not influence the volatility term [4]. Therefore, by performing a change of measure from the risk-neutral measure to the annuity measure, we obtain:

$$\frac{\mathrm{d}\big(S_{i,m}(t)\big)}{S_{i,m}(t)} = -\eta(t)\Big(\frac{\big(P(t,T_i)B(t,T_i) - P(t,T_m)B(t,T_m)\big)}{S_{i,m}(t)A_{i,m}(t)} - \frac{\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{A_{i,m}(t)}\Big)\mathrm{d}W^{i,m}(t). \qquad (3.6)$$

We find that the implied volatility of the swap rate $S_{i,m}(t)$ from t_0 till T_i is given by:

$$\sigma_{imp}(t_0,T_i,T_m,) = \frac{1}{\sqrt{T_i-t_0}}\mathbb{E}\left[\sqrt{\int_{t_0}^{T_i}\eta^2(t)\left(\frac{\big(P(t,T_i)B(t,T_i)-P(t,T_m)B(t,T_m)\big)}{P(t,T_i)-P(t,T_m)}-\frac{\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{A_{i,m}(t)}\right)^2 \mathrm{d}t}\,\middle|\,\mathscr{F}_{t_0}\right]. \tag{3.7}$$

3.3.2. Approximation of swap rate volatility

In this section we follow the steps and results described in [24]. We have the swap rate $S_{i,m}(t)$ given by :

$$S_{i,m}(t) = \frac{P(t,T_i)-P(t,T_m)}{\sum_{k=i+1}^{m}\tau_k P(t,T_k)} = \frac{\frac{P(t_0,T_m)}{P(t,T_m)}\big(P(t,T_i)-P(t,T_m)\big)}{\frac{P(t_0,T_m)}{P(t,T_m)}\sum_{k=i+1}^{m}\tau_k P(t,T_k)}.$$

We use the following approximation of the denominator, where we freeze the at $t=t_0$

$$\frac{P(t_0,T_m)}{P(t,T_m)}\sum_{k=i+1}^{m}\tau_k P(t,T_k) \approx \mathbb{E}\Big[\frac{P(t_0,T_m)}{P(t,T_m)}\sum_{k=i+1}^{m}\tau_k P(t,T_k)\Big|\mathscr{F}(t_0)\Big] \approx \sum_{k=i+1}^{m}\tau_k P(t_0,T_k). \tag{3.8}$$

Using this approximation, we define $\tilde{S}_{i,m}$, which is an approximation of the swap rate $S_{i,m}$:

$$S_{i,m}(t) \approx \frac{\frac{P(t_0,T_m)}{P(t,T_m)}\big(P(t,T_i)-P(t,T_m)\big)}{\sum_{k=i+1}^{m}\tau_k P(t_0,T_k)} = \frac{P(t_0,T_m)}{\sum_{k=i+1}^{m}\tau_k P(t_0,T_k)}\Big(\frac{P(t,T_i)}{P(t,T_m)}-1\Big) =: \tilde{S}_{i,m}. \tag{3.9}$$

For the dynamics of $\tilde{S}_{i,m}(t)$ we have:

$$\mathrm{d}\tilde{S}_{i,m}(t) = \frac{P(t_0,T_m)}{\sum_{k=i+1}^{m}\tau_k P(t_0,T_k)}\mathrm{d}\Big(\frac{P(t,T_i)}{P(t,T_m)}\Big). \tag{3.10}$$

Note that the dynamics of $\tilde{S}_{i,m}(t)$ are easier to work with than the dynamics of $S_{i,m}(t)$. In contrast to the swap rate $S_{i,m}(t)$, the approximated swap rate $\tilde{S}_{i,m}$ is not a martingale under the annuity measure, but is a martingale under the T_m-forward measure. Therefore, by applying Itô's lemma and switching from the risk neutral measure to the T_m-forward measure, we obtain the following driftless dynamics:

$$\begin{aligned}\mathrm{d}\Big(\frac{P(t,T_i)}{P(t,T_m)}\Big) &= \frac{1}{P(t,T_m)}\mathrm{d}P(t,T_i) - \frac{P(t,T_i)}{P^2(t,T_m)}\mathrm{d}P(t,T_m)\\ &= \frac{P(t,T_i)}{P(t,t_m)}\eta(t)\big(B(t,T_m)-B(t,T_i)\big)\mathrm{d}W^{T_m}(t).\end{aligned}$$

Our goal is to obtain an approximation for the implied volatility under the annuity measure. Therefore, we need to switch to the annuity measure. Under the assumption that the swap rate is lognormal distributed, we obtain:

$$\begin{aligned}\frac{\mathrm{d}\tilde{S}_{i,m}(t)}{\tilde{S}_{i,m}(t)} &= \frac{P(t_0,T_m)P(t,T_i)}{\tilde{S}_{i,m}(t)\sum_{k=i+1}^{m}\tau_k P(t_0,T_k)P(t,T_m)}\eta(t)\big(B(t,T_m)-B(t,T_i)\big)\mathrm{d}W^{T_m}(t)\\ &= \frac{P(t,T_i)}{(P(t,T_i)-P(t,T_m))}\eta(t)\big(B(t,T_m)-B(t,T_i)\big)\mathrm{d}W^{T_m}(t)\\ &= \text{drift} + \frac{P(t,T_i)}{P(t,T_i)-P(t,T_m)}\eta(t)\big(B(t,T_m)-B(t,T_i)\big)\mathrm{d}W^{i,m}(t)\\ &\approx \text{drift} + \frac{P(t_0,T_i)}{P(t_0,T_i)-P(t_0,T_m)}\eta(t)\big(B(t,T_m)-B(t,T_i)\big)\mathrm{d}W^{i,m}(t).\end{aligned}$$

Where in the last step the following approximation by freezing technique is performed:

$$\frac{P(t,T_i)}{P(t,T_i)-P(t,T_m)} \approx \mathbb{E}\Big[\frac{P(t,T_i)}{P(t,T_i)-P(t,T_m)}\Big|\mathscr{F}(t_0)\Big] \approx \frac{P(t_0,T_i)}{P(t_0,T_i)-P(t_0,T_m)}.$$

As an approximation for the dynamics of the swap rate we obtain:

$$\frac{\mathrm{d}S_{i,m}(t)}{S_{i,m}(t)} \approx \frac{P(t_0,T_i)}{P(t_0,T_i)-P(t_0,T_m)}\eta(t)\big(B(t,T_m)-B(t,T_i)\big)\mathrm{d}W^{i,m}(t). \tag{3.11}$$

An approximation for the implied volatility of $S_{i,m}(t)$ is given by:

$$\sigma_{imp}(t_0,T_i,T_m) \approx \frac{1}{\sqrt{T_i-t_0}}\sqrt{\int_{t_0}^{T_i}\Big(\frac{P(t_0,T_i)}{P(t_0,T_i)-P(t_0,T_m)}\eta(t)\big(B(t,T_m)-B(t,T_i)\big)\Big)^2\mathrm{d}t}. \tag{3.12}$$

For the $B(t,T_m)-B(t,T_i)$ the following holds:

$$B(t,T_m)-B(t,T_i) = E(s)\int_t^{T_m}\frac{1}{E(u)}\mathrm{d}u - E(t)\int_t^{T_i}\frac{1}{E(u)}\mathrm{d}u = E(t)\int_{T_i}^{T_m}\frac{1}{E(u)}\mathrm{d}u = \frac{E(t)}{E(T_i)}B(T_i,T_m).$$

We obtain as approximation for the implied volatility of the swap rate:

$$\sigma_{imp}(T_i,T_m) \approx \Big|\frac{B(T_i,T_m)P(t_0,T_i)}{\sqrt{T_i-t_0}\big(P(t_0,T_i)-P(t_0,T_m)\big)}\Big|\sqrt{\int_{t_0}^{T_i}\eta^2(t)E^2(t)\mathrm{d}t}. \tag{3.13}$$

Only the latter part depends in Equation (3.13) depends in $\eta(t)$ and is independent of T_m. Therefore, the ratio of two implied volatilities with same T_i but different T_m is independent of $\eta(t)$.

$$\frac{\sigma_{imp}(T_i,T_{m_1})}{\sigma_{imp}(T_i,T_{m_2})} \approx \frac{B(T_i,T_{m_1})|P(t_0,T_i)-P(t_0,T_{m_2})|}{B(T_i,T_{m_2})|P(t_0,T_i)-P(t_0,T_{m_1})|}. \tag{3.14}$$

Since Equation (3.14) is independent of $\eta(t)$, we can use it for the calibration of $\lambda(t)$.

$$\operatorname{argmin}\Big(f\big(\lambda(t)\big)\Big) = \operatorname{argmin}\Bigg(\sqrt{\sum_i\Big(\frac{\sigma^{mrkt}_{swpt}(t_0,T_i,T_{m_1},K)}{\sigma^{mrkt}_{swpt}(t_0,T_i,T_{m_2},K)} - \frac{\sigma^{HW}_{swpt}(t_0,T_i,T_{m_1},K,\lambda(t))}{\sigma^{HW}_{swpt}(t_0,T_i,T_{m_2},K,\lambda(t))}\Big)^2}\Bigg).$$

3.3.3. Extension to shifted lognormal

In the shifted lognormal case we have that the dynamics of $\mathrm{d}S_{i,m}(t)$ and $\mathrm{d}S_{i,m}(t)+s$ are the same, where s is a constant shift parameter. Therefore, the following holds

$$\frac{\mathrm{d}\big(S_{i,m}(t)+s\big)}{S_{i,m}(t)+s} = \frac{S_{i,m}(t)}{S_{i,m}(t)+s}\frac{\mathrm{d}\big(S_{i,m}(t)\big)}{S_{i,m}(t)}$$

We would like to extend the method described in section 3.3.2 to the assumption of shifted lognormal. We can follow the same derivation to obtain Equation (3.11):

$$\frac{\mathrm{d}S_{i,m}(t)}{S_{i,m}(t)} \approx \text{drift} + \frac{P(t_0,T_i)}{P(t_0,T_i)-P(t_0,T_m)}\eta(t)\big(B(t,T_m)-B(t,T_i)\big)\mathrm{d}W^{i,m}(t). \tag{3.15}$$

By assuming shifted lognormal instead of lognormal distributed swap rate, we obtain:

$$\frac{\mathrm{d}\big(S_{i,m}(t)+s\big)}{S_{i,m}(t)+s} \approx \text{drift} + \frac{S_{i,m}(t_0)}{S_{i,m}(t_0)+s}\frac{P(t_0,T_i)}{P(t_0,T_i)-P(t_0,T_m)}\eta(t)\big(B(t,T_m)-B(t,T_i)\big)\mathrm{d}W^{i,m}(t) \tag{3.16}$$

$$\approx \text{drift} + \frac{P(t_0,T_i)}{(S_{i,m}(t_0)+s)A_{i,m}(t_0)}\eta(t)\big(B(t,T_m)-B(t,T_i)\big)\mathrm{d}W^{i,m}(t). \tag{3.17}$$

We obtain the following approximation:

$$\sigma_{imp}(t_0,T_i,T_m) \approx \frac{1}{\sqrt{T_i-t_0}}\sqrt{\frac{\big(P(t_0,T_i)\big)^2}{\big(S_{i,m}(t_0)+s\big)^2\big(A_{i,m}(t_0)\big)^2}\int_{t_0}^{T_i}\eta^2(t)\big(B(t,T_m)-B(t,T_i)\big)^2\mathrm{d}t} \tag{3.18}$$

$$\approx \frac{1}{\sqrt{T_i-t_0}}\Big|\frac{P(t_0,T_i)}{\big(S_{i,m}(t_0)+s\big)A_{i,m}(t_0)}\Big|\sqrt{V_r(t_0,T_i)}B(T_i,T_m). \tag{3.19}$$

For the ratio of two implied volatilities with same expiry T_i and different tenors T_{m_1} and T_{m_2} gives:

$$\frac{\sigma_{imp}(t_0,T_i,T_{m_1})}{\sigma_{imp}(t_0,T_i,T_{m_2})} = \frac{\big(S_{i,m_2}(t_0)+s\big)A_{i,m_2}(t_0)}{\big(S_{i,m_1}(t_0)+s\big)A_{i,m_1}(t_0)}\frac{B(T_i,T_{m_1})}{B(T_i,T_{m_2})}. \tag{3.20}$$

3.3.4. Alternative approximation of the implied volatility

Besides the approximation discussed in the previous two subsections, we can approximate the implied volatility in another way. Instead of applying the freezing technique in an early stage, we can apply the freezing technique in a later stage for an approximation of the implied volatility. Instead of assuming a lognormal distributed swap rate as in Equation (3.6), we assume shifted lognormal, and obtain:

$$\frac{\mathrm{d}\big(S_{i,m}(t)+s\big)}{S_{i,m}(t)+s} = -\frac{S_{i,m}(t)}{S_{i,m}(t)+s}\eta(t)\left(\frac{\big(P(t,T_i)B(t,T_i)-P(t,T_m)B(t,T_m)\big)}{S_{i,m}(t)A_{i,m}(t)} - \frac{\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{A_{i,m}(t)}\right)\mathrm{d}W^{i,m}(t) \tag{3.21}$$

$$= -\frac{\eta(t)}{(S_{i,m}(t)+s)}\left(\frac{\big(P(t,T_i)B(t,T_i)-P(t,T_m)B(t,T_m)\big)}{A_{i,m}(t)} - \frac{\big(P(t,T_i)-P(t,T_m)\big)\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{\big(A_{i,m}(t)\big)^2}\right)\mathrm{d}W^{i,m}(t). \tag{3.22}$$

We have the following implied volatility under shifted lognormal:

$$\sigma_{imp}(t_0,T_i,T_m,) = \frac{1}{\sqrt{T_i-t_0}}\mathbb{E}\left[\sqrt{\int_{t_0}^{T_i}\frac{\eta^2(t)}{\big(S_{i,m}(t)+s\big)^2}\left(\frac{\big(P(t,T_i)B(t,T_i)-P(t,T_m)B(t,T_m)\big)}{A_{i,m}(t)} - \frac{\big(P(t,T_i)-P(t,T_m)\big)\sum_{k=i+1}^{m}\tau_k P(t,T_k)B(t,T_k)}{\big(A_{i,m}(t)\big)^2}\right)^2\mathrm{d}t}\,\middle|\,\mathscr{F}_{t_0}\right]. \tag{3.23}$$

By applying a freezing technique, we obtain the following approximation:

$$\sigma_{imp}(t_0,T_i,T_m) \approx \frac{1}{\sqrt{T_i-t_0}}\sqrt{\int_{t_0}^{T_i}\frac{\eta^2(t)}{\big(S_{i,m}(t_0)+s\big)^2}\left(\frac{\big(P(t_0,T_i)B(t,T_i)-P(t_0,T_m)B(t,T_m)\big)}{A_{i,m}(t_0)} - \frac{\big(P(t_0,T_i)-P(t_0,T_m)\big)\sum_{k=i+1}^{m}\tau_k P(t_0,T_k)B(t,T_k)}{\big(A_{i,m}(t_0)\big)^2}\right)^2\mathrm{d}t}. \tag{3.24}$$

For the ratio of two implied volatilities one obtains:

$$\frac{\sigma_{imp}(t_0,T_i,T_{m_1})}{\sigma_{imp}(t_0,T_i,T_{m_2})} \approx \frac{S_{i,m_2}(t_0)+s}{S_{i,m_1}(t_0)+s}\sqrt{\frac{\int_{t_0}^{T_i}\eta^2(t)\left(\frac{\big(P(t_0,T_i)B(t,T_i)-P(t_0,T_{m_1})B(t,T_{m_1})\big)}{A_{i,m_1}(t_0)} - \frac{\big(P(t_0,T_i)-P(t_0,T_{m_1})\big)\sum_{k=i+1}^{m_1}\tau_k P(t_0,T_k)B(t,T_k)}{\big(A_{i,m_1}(t_0)\big)^2}\right)^2\mathrm{d}t}{\int_{t_0}^{T_i}\eta^2(t)\left(\frac{\big(P(t_0,T_i)B(t,T_i)-P(t_0,T_{m_2})B(t,T_{m_2})\big)}{A_{i,m_2}(t_0)} - \frac{\big(P(t_0,T_i)-P(t_0,T_{m_2})\big)\sum_{k=i+1}^{m_2}\tau_k P(t_0,T_k)B(t,T_k)}{\big(A_{i,m_2}(t_0)\big)^2}\right)^2\mathrm{d}t}}. \tag{3.25}$$

Mathematically, we cannot justify that the $\eta(t)$ would cancel out in this equation. Not even when we take into account that $\eta(t)$ is a piece-wise constant function. When $\eta(t)$ is a piece-wise constant function, we can rewrite this to.

$$\frac{\sigma_{imp}(t_0,T_i,T_{m_1})}{\sigma_{imp}(t_0,T_i,T_{m_2})} \approx \frac{S_{i,m_2}(t_0)+s}{S_{i,m_1}(t_0)+s}\sqrt{\frac{\sum_{j=0}^{n}\eta_j^2\int_{T_j}^{T_{j+1}}\left(\frac{\big(P(t_0,T_i)B(t,T_i)-P(t_0,T_{m_1})B(t,T_{m_1})\big)}{A_{i,m_1}(t_0)} - \frac{\big(P(t_0,T_i)-P(t_0,T_{m_1})\big)\sum_{k=i+1}^{m_1}\tau_k P(t_0,T_k)B(t,T_k)}{\big(A_{i,m_1}(t_0)\big)^2}\right)^2\mathrm{d}t}{\sum_{j=0}^{n}\eta_j^2\int_{T_j}^{T_{j+1}}(t)\left(\frac{\big(P(t_0,T_i)B(t,T_i)-P(t_0,T_{m_2})B(t,T_{m_2})\big)}{A_{i,m_2}(t_0)} - \frac{\big(P(t_0,T_i)-P(t_0,T_{m_2})\big)\sum_{k=i+1}^{m_2}\tau_k P(t_0,T_k)B(t,T_k)}{\big(A_{i,m_2}(t_0)\big)^2}\right)^2\mathrm{d}t}}. \tag{3.26}$$

Under the assumption of $\eta(t)$ being a constant function, we do obtain a fraction independent from $\eta(t)$:

$$\frac{\sigma_{imp}(t_0,T_i,T_{m_1})}{\sigma_{imp}(t_0,T_i,T_{m_2})} \approx \frac{S_{i,m_2}(t_0)+s}{S_{i,m_1}(t_0)+s}\sqrt{\frac{\int_{t_0}^{T_i}\left(\frac{\big(P(t_0,T_i)B(t,T_i)-P(t_0,T_{m_1})B(t,T_{m_1})\big)}{A_{i,m_1}(t_0)} - \frac{\big(P(t_0,T_i)-P(t_0,T_{m_1})\big)\sum_{k=i+1}^{m_1}\tau_k P(t_0,T_k)B(t,T_k)}{\big(A_{i,m_1}(t_0)\big)^2}\right)^2\mathrm{d}t}{\int_{t_0}^{T_i}\left(\frac{\big(P(t_0,T_i)B(t,T_i)-P(t_0,T_{m_2})B(t,T_{m_2})\big)}{A_{i,m_2}(t_0)} - \frac{\big(P(t_0,T_i)-P(t_0,T_{m_2})\big)\sum_{k=i+1}^{m_2}\tau_k P(t_0,T_k)B(t,T_k)}{\big(A_{i,m_2}(t_0)\big)^2}\right)^2\mathrm{d}t}}. \tag{3.27}$$

We will use the approximation of ratio in Equation (3.27) to compare to the approximation in Equation (3.20). For time-dependent $\eta(t)$, it is hard to justify this approximation mathematically as $\eta(t)$ only cancels out by taking it out of the integral.

3.3.5. Error on approximations

Mathematically it is difficult to quantify the errors on the approximations. For different parameters $\lambda(t)$, $\eta(t)$, T_i, T_m and zero coupon bond curve the error differs. In order to have an idea of the magnitude of the error, an example is given in Figure 3.1. A caveat has to be made, the error differs per choice of $\lambda(t)$, $\eta(t)$, T_i, T_m, shift parameter and zero coupon bond curve.

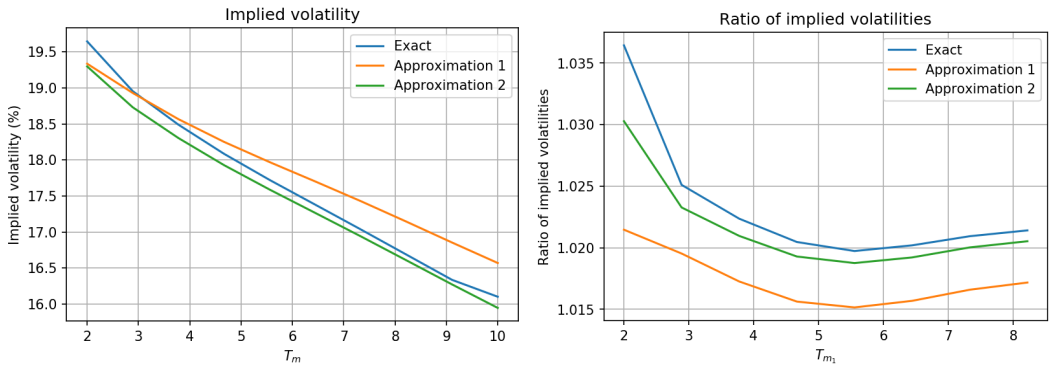

Figure 3.1: The error of the approximations can be observed in this plot. On the left the implied volatility is shown. On the right the fractions of two implied volatility ratios are shown with $T_{m_2} = T_{m_1} + 1$. Approximation 1 refers to Equation (3.20). Approximation 2 refers to Equation (3.27). In this example $\lambda(t) = 0.05 + 0.01 \cdot \sin(t)$, $\eta(t) = 0.0018 + 0.0005 \cdot \sin(t)$, $T_i = 1$, $\tau = 1$, $P(t_0, t) = e^{-0.01 \cdot t}$ and shift parameter is zero.

From the experiments conducted for this thesis, it appears that approximation 2 from Equation (3.27) is more accurate than approximation 1 from Equation (3.20). Still, we cannot justify the use of to mathematically justify the use of Equation (3.27) for time-dependent $\eta(t)$. Besides, in Section 3.7.1 the calibration results show that the computation time of approximation 2 is very large. Therefore, in this thesis we decided to focus on approximation 1.

3.4. Choice of swaption set for calibration

The choice of swaptions used for calibration is not trivial. In Table 3.1 and Figure 3.2 one can find the at the money shifted implied volatilities observed from the market for various expiries and tenors. The number of parameters in the Hull-White model is limited, therefore it cannot fit the whole swaption matrix [32]. Instead, a set of swaptions for the calibration of the Hull-White model is selected. Two common selections [32] are a co-terminal swaption set and a columnar swaption set, where co-terminal swaption set refers to the counterdiagonal of the matrix and columnar to a specific row of the swaption matrix.

		Expiries			
		1y	2y	3y	4y
Tenors	6m	0.109	0.112	0.133	0.153
	1y	0.099	0.111	0.133	0.149
	2y	0.106	0.121	0.139	0.150
	3y	0.118	0.135	0.148	0.157
	4y	0.129	0.142	0.153	0.161

Table 3.1: Shown are at the money shifted lognormal implied volatilities of swaptions observed in the market in May 2020 with shift parameter 3% and $\tau = 0.5$.

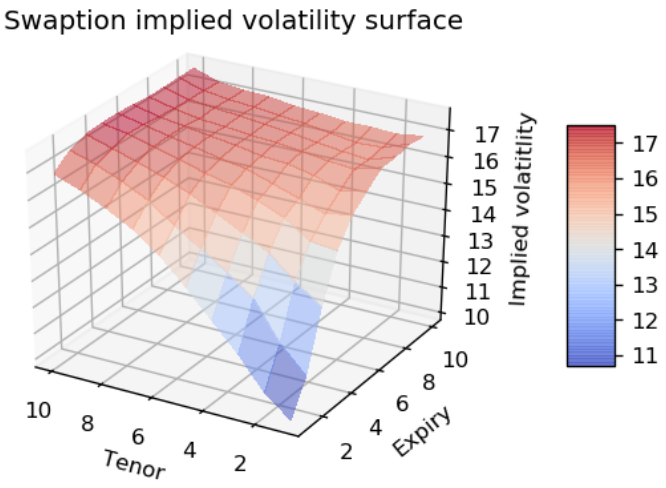

Figure 3.2: Shown are at the money shifted lognormal implied volatilities of swaptions observed in the market in May 2020 with shift parameter 3% and $\tau = 0.5$. The implied volatility surface shown in a graphical way.

A counter diagonal of the swaption matrix is known as a co-terminal basket of swaptions. The co-terminal swaption set is commonly used for the calibration of the Hull-White model, partly because of hedging reasons is argued in [24]. Also, the duration of counterparty exposure for swaptions from the co-terminal basket is the same. This could be considered as beneficial, as one would like consistency in the products with respect to counterparty risk. The choice may also depend on what intruments are in one's portfolio. In [32] is argued that for a single swap portfolio the co-terminal swaption set is ideal. Under the co-terminal swaption set, all swaptions have the same amount of time of counterparty risk. In Table 3.2 one can find the 5 year co-terminal swaption basket from the swaption matrix given above. In this thesis, the co-terminal basket is used for the calibration of $\eta(t)$. In that way we use products that have the same counterparty risk, which is beneficial for our calibration results.

5y co-terminal basket			
Expiry	Tenor	Value	Strike
1	4	0.129	ATM
2	3	0.135	ATM
3	2	0.139	ATM
4	1	0.149	ATM

Table 3.2: The 5 year co-terminal basket of table 3.1 that could be used for the calibration of $\eta(t)$. The values in this table consist of the counter diagonal of Table 3.1

For the calibration of $\lambda(t)$ we use another part of the swaption matrix. By a the subscript HW we indicate that it is under the Hull-White model. In order to calibrate $\lambda(t)$ we approximate the ratios of two implied volatilities with same expiry and different tenor by the following:

$$\frac{\sigma^{HW}_{imp}(t_0, T_i, T_{m_1}, \lambda(t))}{\sigma^{HW}_{imp}(t_0, T_i, T_{m_2}, \lambda(t))} \approx \frac{\big(S_{i,m_2}(t_0)+s\big)A_{i,m_2}(t_0)}{\big(S_{i,m_1}(t_0)+s\big)A_{i,m_1}(t_0)} \frac{B(\lambda(t), T_i, T_{m_1})}{B(\lambda(t), T_i, T_{m_2})}. \tag{3.28}$$

The last term (The fraction of B's) is the only term that depends on $\lambda(t)$. We can choose a suitable set of swaptions to create independence between the ratios. Recall that $B(T_i, T_m) = E(T_i)\int_{T_i}^{T_m} \frac{1}{E(u)} \mathrm{d}u$ with $E(T_i)$ given in Equation (2.2). Let $j = 1, \ldots, k$ denote the specific set of T_i, T_{m_1} and T_{m_2} by a superscript. We choose a specific T_i^j, $T_{m_1}^j$ and $T_{m_2}^j$ such that $T_i^j < T_{m_1}^j < T_{m_2}^j$ and for all $n, p \in \{1, \ldots, k\}, n \neq m$ that $(T_i^n, T_{m_2}^n) \bigcap (T_i^p, T_{m_2}^p) = \emptyset$. Here (A, B) denotes the open interval from A to B. We obtain

$$\frac{B(T_i,T_{m_1})}{B(T_i,T_{m_2})} = \frac{\int_{T_i}^{T_{m_1}} \frac{1}{E(u)}\mathrm{d}u}{\int_{T_i}^{T_{m_2}} \frac{1}{E(u)}\mathrm{d}u} = \frac{\int_{T_i}^{T_{m_1}} \mathrm{e}^{-\int_{t_0}^{u}\lambda(v)\mathrm{d}v}\mathrm{d}u}{\int_{T_i}^{T_{m_2}} \mathrm{e}^{-\int_{t_0}^{u}\lambda(v)\mathrm{d}v}\mathrm{d}u} = \frac{\int_{T_i}^{T_{m_1}} \mathrm{e}^{-\int_{t_0}^{T}{}_i\lambda(v)\mathrm{d}v-\int_{T_i}^{u}\lambda(v)\mathrm{d}v}\mathrm{d}u}{\int_{T_i}^{T_{m_2}} \mathrm{e}^{-\int_{t_0}^{T_i}\lambda(v)\mathrm{d}v-\int_{T_i}^{u}\lambda(v)\mathrm{d}v}\mathrm{d}u} \tag{3.29}$$

$$= \frac{\mathrm{e}^{-\int_{t_0}^{T_i}\lambda(v)\mathrm{d}v}\int_{T_i}^{T_{m_1}} \mathrm{e}^{-\int_{T_i}^{u}\lambda(v)\mathrm{d}v}\mathrm{d}u}{\mathrm{e}^{-\int_{t_0}^{t}\lambda(v)\mathrm{d}v}\int_{T_i}^{T_{m_2}} \mathrm{e}^{-\int_{T_i}^{u}\lambda(v)\mathrm{d}v}\mathrm{d}u} = \frac{\int_{T_i}^{T_{m_1}} \mathrm{e}^{-\int_{T_i}^{u}\lambda(v)\mathrm{d}v}\mathrm{d}u}{\int_{T_i}^{T_{m_2}} \mathrm{e}^{-\int_{T_i}^{u}\lambda(v)\mathrm{d}v}\mathrm{d}u}. \tag{3.30}$$

We find that $\frac{B(T_i,T_{m_1})}{B(T_i,T_{m_2})}$ merely depends on the value of $\lambda(t)$ between T_i and T_{m_2}. Because we took a set of swaptions in such a way that there is no overlap between the intervals from T_i^j T_i^j for all j, we obtain independece between the swaption ratios for the calibration procedure. Having independence between the swaption ratios gives us the opportunity to perform calibration in sequential style, instead of parallel (simultaneous) calibration. This reduces the computation time significantly, and therefore considered as a benefit. Table 3.3 shows us an example of a set of swaptions that possess the desired property.

		Expiries			
		1y	2y	3y	4y
Tenors	6m	0.109	0.112	0.133	0.153
	1y	0.099	0.111	0.133	0.149
	2y	0.106	0.121	0.139	0.150
	3y	0.118	0.135	0.148	0.157
	4y	0.129	0.142	0.153	0.161

Table 3.3: Shown are at the money shifted lognormal implied volatilities of swaptions observed in the market in May 2020 with shift parameter 3% and $\tau = 0.5$. In yellow and orange we can find an example of implied volatilities used for the calibration of $\lambda(t)$, where yellow represents the implied volatilities of T_{m_1} and orange represents the implied volatilities for T_{m_2}.

3.5. Choosing proper strike prices

At the money (ATM) implied volatilities are a common choice for the calibration purposes, since these implied volatilities are the most liquid and vega is not close to zero [25]. For deep in-the-money and out-of-the-money options vega is very small, which is problematic for some minimization algorithms, e.g. Newton Rapshon. Therefore, in this thesis at-the-money strike prices are used. One particular way to obtain ATM strike values is shown in [27]. In this article, the strike prices are chosen in the follwing way:

$$K_i(T_n) = F(0,T_n)\exp\left(0.1\cdot\sqrt{T_n}\cdot\delta_i\right)$$
$$\delta_i = -1.5,-1.0,-0.5,0.0,0.5,1.0,1.5$$

Where T_n is the expiry and $F(0,T_n)$ is the forward FX rates to T_n. $\delta_i = 0.0$ leads to ATM strike value, $\delta_i < 0$ to ITM and $\delta_i > 0$ to OTM. For swaptions ATM strike prices are defined as follows.

Definition 3.5.1. The strike price of a swaption with expiry T_i and tenor T_m is called at the money (ATM) if and only if

$$K = K_{ATM} := S_{i,m}(t_0) = \frac{P(t_0,T_i)-P(t_0,T_m)}{\sum_{k=i+1}^{m}\tau_k P(t_0,T_k)}. \tag{3.31}$$

We call a payer swaption out of the money (OTM) when $K > K_{ATM}$ and in the money (ITM) if $K < K_{ATM}$. For a receiver swaption it is vice versa; out of the money (OTM) when $K < K_{ATM}$, in the money (ITM) if $K > K_{ATM}$.

Combining the definition and the way of computing ATM strike prices in [27], Strike prices can be determined via the following formula:

$$K_j(T_i) = S_{i,m}(t_0)\exp(0.1\cdot\sqrt{T_i}\cdot\delta_j),$$
$$\delta_j = -1.5,-1.0,-0.5,0.0,0.5,1.0,1.5,$$

where $\delta_j = 0.0$ leads to ATM strike price.

3.6. Optimization methods

Choosing suitable optimization method for our problem is an extensive task, and a whole study could be devoted to the choice of the best optimization method for our problem. In this section, we will outline our thought process of choosing a good optimization method for our problem.

A decision one needs to make is whether one will use a local or global optimization method for the calibration. In our optimization problem, we can use initial guesses from the previous calibrated values for $\eta(t)$ and $\lambda(t)$, which can be considered as a first guess. Global optimization methods are considered to be more complex [22], and look for the global optimum on all set of parameters instead of neighbours. For the first calibration, a global optimization method can be used. We argue that afterwards, a local optimization method has the preference, since these will be faster and less complex.

Algorithm 2: Example of parallel calibration.

$$\eta(t) = \underset{\eta(t)}{\operatorname{argmin}}\left(\sqrt{\sum_i \left(\sigma^{mrkt}_{swpt}(t_0, T_i, T_m, K) - \sigma^{HW}_{swpt}(t_0, T_i, T_m, K, \lambda(t), \eta(t))\right)^2}\right),$$

In this way every η_i from the piece-wise function in Equation (3.2) is optimized at the same time.

Another choice that needs to be made is whether we are optimizing parallel or sequential. Optimization is called parallel if it is applied only once and sequential if is applied multiple times in a row [22]. This depends on our optimization problem. If we have dependence between the optimization problems, we cannot apply the optimization multiple times in a row (sequential calibration), as one optimization can be effected by the other. Therefore, in this thesis we decided to use parallel optimization if we have dependence between the optimization problems. As we have shown in Section 3.4, we can choose a specific set of swaptions to create independence in the optimization problem of the mean reversion parameter $\lambda(t)$. In this case sequential optimization can be performed, and is considered superior as sequential calibration is faster.

Algorithm 3: Example of sequential calibration.

for $i = 0, \ldots, n$ **do**

$$\eta_i = \underset{\eta(t)}{\operatorname{argmin}}\left(\sqrt{\left(\sigma^{mrkt}_{swpt}(t_0, T_i, T_m, K) - \sigma^{HW}_{swpt}(t_0, T_i, T_m, K, \lambda(t), \eta(t))\right)^2}\right)$$

end

In this way every η_i from the piece-wise function in Equation (3.2) is optimized separately.

For our calibration of the Hull-White model we encounter two optimization problems. Our first optimization problem is the optimization problem for $\lambda(t)$ given by:

$$\underset{\lambda(t)}{\operatorname{argmin}}\left(\sqrt{\sum_j \left(\frac{\sigma^{mrkt}_{swpt}(t_0, T_i^j, T_{m_1}^j, K)}{\sigma^{mrkt}_{swpt}(t_0, T_i^j, T_{m_2}^j, K)} - \frac{\tilde{\sigma}^{HW}_{swpt}(t_0, T_i^j, T_{m_1}^j, K, \lambda(t))}{\tilde{\sigma}^{HW}_{swpt}(t_0, T_i^j, T_{m_2}^j, K, \lambda(t))}\right)^2}\right), \tag{3.32}$$

where $\lambda(t)$ is a piece-wise constant functions with the T_i's from the swaption as change time points. $\tilde{\sigma}^{mrkt}_{swpt}(\cdot)$ is given by the market and $\sigma^{HW}_{swpt}(\cdot)$ is given by our approximation in Equation (3.19). Assuming λ to be constant, we can find out how our objective function behaves. By having a constant λ, we obtain that $E(\lambda, t) = e^{\lambda t}$ and $B(\lambda, t, T) = -\frac{1}{\lambda}(e^{-\lambda(T-t)} - 1)$. For the fraction of the implied volatilities under the HW model, we have the expression in Equation (3.20), which only depend on the $B(\lambda, T_i, T_{m_1})/B(\lambda, T_i, T_{m_2})$. We have for constant λ

$$\frac{B(\lambda, T_i, T_{m_1})}{B(\lambda, T_i, T_{m_2})} = \frac{-\frac{1}{\lambda}(e^{-\lambda(T_{m_1} - T_i)} - 1)}{-\frac{1}{\lambda}(e^{-\lambda(T_{m_2} - T_i)} - 1)} = \frac{e^{-\lambda(T_{m_1} - T_i)} - 1}{e^{-\lambda(T_{m_2} - T_i)} - 1}. \tag{3.33}$$

Notice that the expression in Equation (3.33) goes to 0 if $\lambda \to -\infty$ and goes to 1 if $\lambda \to \infty$. We have a discontinuity in $\lambda = 0$, where we have that

$$\lim_{\lambda \uparrow 0} \frac{B(\lambda, T_i, T_{m_1})}{B(\lambda, T_i, T_{m_2})} = \lim_{\lambda \downarrow 0} \frac{B(\lambda, T_i, T_{m_1})}{B(\lambda, T_i, T_{m_2})} = \frac{T_{m_1} - T_i}{T_{m_2} - T_i},$$

which can be found by looking at the Taylor expansion. We look at the derivative of the expression in Equation (3.33) to show that it is increasing, we have

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\frac{B(\lambda,T_i,T_{m_1})}{B(\lambda,T_i,T_{m_2})}=\frac{\mathrm{e}^{\lambda(T_{m_2}+T_i-T_{m_1})}\left(T_{m_1}(\mathrm{e}^{\lambda T_{m_2}}-\mathrm{e}^{\lambda T_i})-T_i(\mathrm{e}^{\lambda T_{m_2}}-\mathrm{e}^{\lambda T_{m_1}})-T_{m_2}(\mathrm{e}^{\lambda T_{m_1}}-\mathrm{e}^{\lambda T_i})\right)}{\left(\mathrm{e}^{\lambda(T_{m_1}-T_i)}-1\right)^2}. \tag{3.34}$$

Note that the expression in Equation (3.34) is positive for all $\lambda>0$ and $0<T_i<T_{m_1}<T_{m_2}$. Therefore, the expression in Equation (3.33) is increasing, we have that the problem in Equation (3.32) becomes a convex optimization problem. Since we have independence between the terms as shown in Equation (3.30), we can use sequential calibration of our optimization problem in Equation (3.32). We obtain the following optimization problem, that we have to solve for every implied volatility fraction.

$$\lambda(t) \text{ for } T_i^j\le t<T_{m_2}^j=\underset{\lambda}{\operatorname{argmin}}\left(\sqrt{\left(\frac{\sigma_{swpt}^{mrkt}(t_0,T_i^j,T_{m_1}^j,K)}{\sigma_{swpt}^{mrkt}(t_0,T_i^j,T_{m_2}^j,K)}-\frac{\tilde{\sigma}_{swpt}^{HW}(t_0,T_i^j,T_{m_1}^j,K,\lambda)}{\tilde{\sigma}_{swpt}^{HW}(t_0,T_i^j,T_{m_2}^j,K,\lambda)}\right)^2}\right), \tag{3.35}$$

We calibrate $\lambda(t)$ between T_i and T_{m_2} as a constant. Therefore, the calibration of $\lambda(t)$ can be split into subproblems where we have non-linear convex optimization problems.

The second optimization problem is for $\eta(t)$, where we use the $\lambda(t)$ found in the previous optimization.

$$\underset{\eta(t)}{\operatorname{argmin}}\left(\sqrt{\sum_i\left(\sigma_{swpt}^{mrkt}(t_0,T_i,T_m,K)-\sigma_{swpt}^{HW}(t_0,T_i,T_m,K,\lambda(t),\eta(t))\right)^2}\right), \tag{3.36}$$

where $\sigma_{swpt}^{HW}(\cdot)$ is computed by Theorem 2.1.2. This optimization problem is more difficult than the calibration of $\lambda(t)$. In the computation of $\sigma_{swpt}^{HW}(\cdot)$ we have to optimize for r^* as in Theorem 2.1.2. Due to this complex computation of the $\sigma_{swpt}^{HW}(\cdot)$, the optimization problem in Equation (3.36) is non-convex [36].

Both optimization problems in Equation (3.32) and (3.36) are non-linear optimization problems. Numerical methods for non-linear optimization are iterative. These iterative methods can be classified into two categories [37]: line search methods and trust region methods. In the following two sections we go in detail on both categories of methods.

3.6.1. Line searching algorithms

Line searching algorithms can be considered as classical methods for solving nonlinear optimization problems. Let $\{x_k\}$, $k=1,2,\ldots$ be the sequence of points that are generated by an iterative method starting with the initial guess x_0. Let $f(\cdot)$ be the target function. Line searching methods work in the following way [7]:

- Calculate the search direction p_k from x_k.
- Determine a suitable steplength $a_k>0$, such that $f(x_k+a_kp_k)<f_k$.
- Compute the new x_{k+1}: $x_{k+1}=x_k+a_kp_k$.

Examples of line searching algorithms are modified-Newton methods, quasi Newton methods and conjugate-gradient method. In this thesis, the Newton-Raphson method [38] has been used as line searching method.

3.6.2. Trust region algorithms

In contrast to Line searching algorithm, trust-region methods check whether the new step is improving the new iterative result. Let $f(\cdot)$ be the target function. The following steps are performed

- Determine the step s_k.
- Accept step s_k if the $f(x_k+s_k)$ compared to a predicted value exceeds a threshold.
- Compute the new x_{k+1}: $x_{k+1}=x_k+s_k$ if accepted, else $x_{k+1}=x_k$.

Trust region algorithms have better performance on non-convex problems compared to line searching algorithms [37]. We will give two examples of trust region methods.

The Levenberg Marquardt algorithm is a popular method used for the non-linear least-squares problems, and can be considered as a trust region method [37]. One provides a guess estimate of the parameters x_0 and

measured vector y. Let f be the cost function. We want to find the parameter vector x, such that $\hat{y} = f(x) \approx y$. Let $\epsilon = \hat{y} - y$. This is an optimization problem where we want to minimize $\epsilon^T \epsilon$. In the Levenberg Marquardt algorithm a a linear approximation of the function f in the neighborhood of x is used as a basis [17]. The Levenberg Marquardt algorithm from the scipy.optimize package in python has been used. There are three criteria of convergence that are tested every iteration. if one of the tests satisfies, a convergence is reached. Below we summarize the three convergence tests.

- $ftol$, the first criterion is the most straightforward. It determines if the change of the cost function is lower than a predetermined threshold. If this is the case, we have a convergence.
- $xtol$ computes the change of the variables and compares it to a predetermined threshold. It it is lower than the predetermined threshold, a convergence is reached.
- $gtol$, the final convergence criterion investigates the Jacobian. The value of the cosines of the angles between columns are computed. Then the maximum absolute value of these is taken and compared to a predefined threshold. Again, if it is lower than the treshold, we have a convergence.

One question that arises is how to pick an appropriate predefined for the criteria. Do we choose them the same or do we differ per convergence criterion. More details about the Levenberg Marquardt algorithm can be found in [20].

Another trust region method is Powell's Dogleg algorithm [30] [31] is a trust region method for non-linear optimization. Just like the Levenberg-Marquardt algorithm, it provides solutions for non-linear least-squares problems.

3.6.3. Choice of optimization algorithm

In [16] is shown that trust region often performs better than line search in an optical flow problem, and especially when we deal with non-linear and non-convex model. Also in [37] is argued that trust region perform better on non-linear and non-convex problems. Based on these results, we argue that trust region methods are more suitable for our optimization problem of $\eta(t)$. For the sequential calibration of $\lambda(t)$ as we showed that this problem is convex. Therefore, we argue that line search algorithms (e.g. Newton Raphson) can be used.

3.7. Calibration results

Considering the results of the calibration, how can one judge whether the results are "good"? The value of $\eta(t)$ and $\lambda(t)$ are not known in the real world, so one cannot compare these results with the real world in that way. In this thesis I used the following methods to gain insight in the performance of the calibration

- Speed. The calibration should be as fast as possible.
- Precision on market data. Considering the implied volatilities obtained by the Hull-White model with the calibrated $\eta(t)$ and $\lambda(t)$, we compare those to the implied volatilities from the market.
- Precision on artificial data. By creating data with known $\eta(t)$ and $\lambda(t)$, we can find out how good the calibration when we compare the calibrated $\lambda(t)$ and $\eta(t)$ to the artificial $\lambda(t)$ and $\eta(t)$.

3.7.1. Artificial data

Using a zero coupon bond curve obtained from the market, we can simulate data choosing $\lambda(t)$ and $\eta(t)$. By doing this, we can actually measure how well the calibration of our method is by comparing our calibrated $\lambda(t)$ and $\eta(t)$. First we will look at two different approximation methods for the fraction of two implied volatilities with same tenor but different tenor, and conclude that one is superior for calibration purposes. Afterwards, we look into stress testing under the Hull-White model. At last, we test our calibration method to data from the Cox-Ingersoll Ross (CIR) model.

Calibration of $\lambda(t)$ under different approximations

In this part we generate artificial data from the Hull-White dynamcis. We compare the calibration of $\lambda(t)$, where the fraction is approximated by Equation (3.20) and Equation (3.27). We will refer to Equation (3.20)

as approximation 1 and to Equation (3.27) as approximation 2. By looking at both approximations, one can observe that approximation 2 is more complex than approximation 1. By applying the freezing technique in a later stage, we get a more complex approximation in return, which might result in better calibration. In Figure 3.3 one can find a result of the calibration under both methods. Both methods seem to perform well under artificial data with market curve observed from the market.

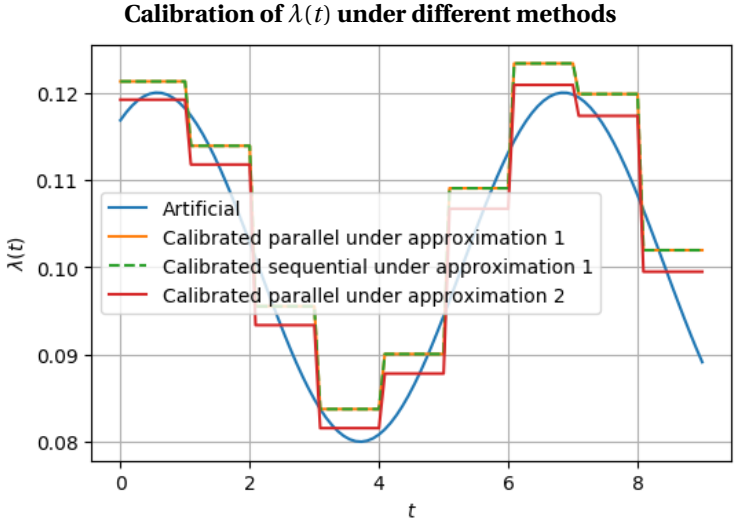

Figure 3.3: Calibration of $\lambda(t)$ under two approximation methods where approximation 1 is given by Equation (3.20) and approximation 2 is given by Equation (3.25). $\lambda(t) = 0.1 + 0.02 \cdot \sin t$ and $\eta(t) = 0.01 + 0.002 \cdot \sin t$. Calibration time for parallel under approximation 1 was 3.3 seconds. Calibration time for sequential under approximation 1 was 0.1 seconds. Calibration time for parallel under approximation 2 was 5 minutes.

In Figure 3.3 we can compare the methods of calibration of $\lambda(t)$. Notice that the calibration by a sequential method or parallel method for approximation 1 lead to the same calibrated $\lambda(t)$. This is exactly what we stated in Section 3.4 by showing that we have independence by choosing a specific set of swaptions. We argue that sequential calibration is superior over parallel calibration, since the computation time is significantly lower. Comparing the calibration under approximation 1 to the calibration under approximation 2 in Figure 3.3, it appears that the calibration is better fitted to the artificial $\lambda(t)$. By looking at the computation time, we argue that parallel calibration under approximation 2 is not practical due to the large computation time. Therefore, we decided to focus on approximation 1 for the research in this thesis. Still, we believe that for future research more examination of the second approximation would be interesting.

Hull-White process stress testing

We generate data from the Hull-White model by choosing $\lambda(t)$ and $\eta(t)$ under the zero coupon bond curve obtained from the market. By choosing different $\lambda(t)$ with corresponding $\eta(t)$, we test our calibration method. Results can be found in Figure 3.4.

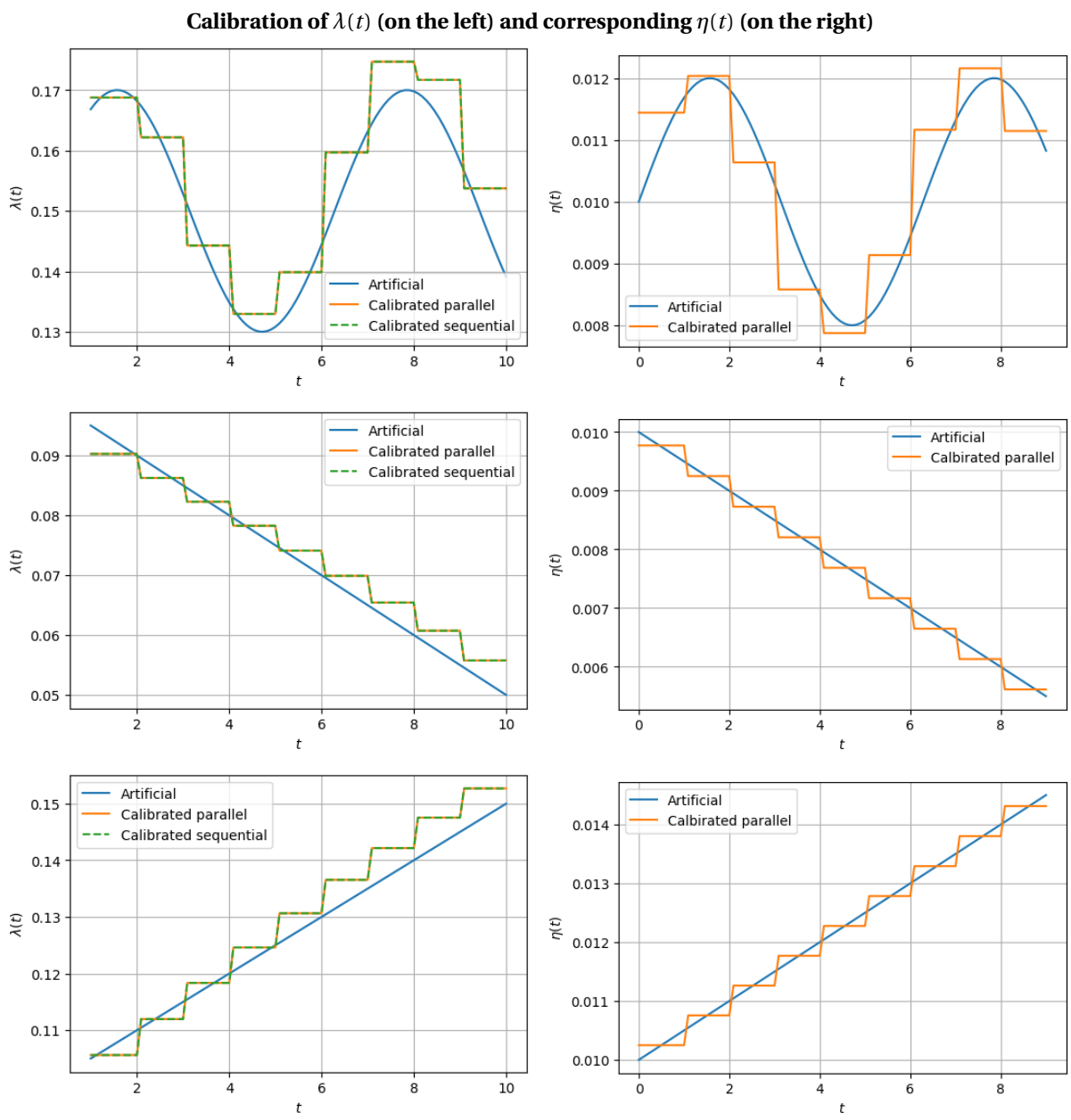

Figure 3.4: Calibration of $\lambda(t)$ by approximation of the fraction in Equation (3.20). One can observe calibration of several $\lambda(t)$ with corresponding $\eta(t)$. Computation time for sequential calibration of lambda is 0.01 seconds by Newton-Raphson. Computation time for parallel calibration of $\eta(t)$ is 13 seconds by Levenberg-Marquardt and 3 seconds under dogbox method. $\lambda(t)$ with corresponding $\eta(t)$ are chosen in such a way that the obtained implied volatilities are comparable to the implied volatilities observed in the market.

Figure 3.4 shows the calibrated $\lambda(t)$ and $\eta(t)$ look similar to the artificial, even though we use crude approximations for the calibration of $\lambda(t)$. Computation time of sequential calibration of $\lambda(t)$ is only 0.01 seconds, whilst parallel calibration time of $\eta(t)$ is 13 seconds under Levenberg-Marquardt. A method of calibration of the Hull-White model is choosing λ as a constant and calibrate only $\eta(t)$ [24]. We would like to show, that picking $\lambda(t)$ (to some extent) arbitrary can have an impact on the calibration of $\eta(t)$. We choose one big and one small constant λ. Afterwards we calibrate for the different mean reversion parameters. The results can be observed in Figure 3.5.

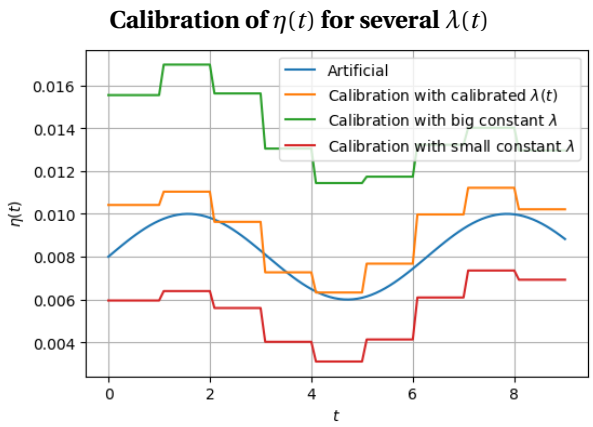

Figure 3.5: In the plot one can find the calibration of $\eta(t)$, when different mean reversion parameters are used during the calibration. Big constant λ is 0.2 and small constant λ is 0.001. The calibrated $\lambda(t)$ fluctuates around the artificial $\lambda(t) = 0.1 + 0.02 * \sin(t)$.

In Figure 3.5 we observe that the value of the mean reversion parameter $\lambda(t)$ influences the value of the volatility parameter $\eta(t)$. A bigger mean reversion $\lambda(t)$ parameters would cause the volatility parameter $\eta(t)$ to increase as well. For a lower mean reversion parameter we expect a smaller volatility parameter $\eta(t)$. This makes sense, when one thinks about this; given that the implied volatilities stay the same, a larger mean reversion would cause smaller moves in the dynamics, therefore the volatility parameter needs to compensate. This compensation of larger mean reversion by larger volatility will be discussed in the next subsection.

In conclusion, our calibration method seems to work well under artificial data with zero coupon bond curve observed from market. Important to note is that mean reversion parameter has big impact during the calibration of volatility parameter $\eta(t)$. Therefore, we argue that λ cannot be picked arbitrarily and the calibration of $\lambda(t)$ should be performed before the calibration of $\eta(t)$.

Larger λ results in larger η in the calibration

In general if we use a bigger mean reversion parameter λ for the calibration of η, this will result in a larger η. This can be observed in Figure 3.5. In this small section we will look into this mathematically. The dynamics of the Hull-White model with constant parameters are given by:

$$\mathrm{d}r(t) = \lambda\big(\theta(t) - r(t)\big)\mathrm{d}t + \eta \mathrm{d}W(t).$$

Note that if $\lambda < 0$, by each iterative step will we move away from the long term mean $\theta(t)$. For $\lambda < 0$ the long term mean works as a repeller instead of an attractor. We discussed this in Section 2.1.1. Right now we will focus on positive mean reversion. We show that the variance of process $r(t)$ becomes smaller for larger mean reversion parameter. The variance of process $r(t)$ for constant λ and η is given by

$$\mathrm{Var}[r(t)|\mathscr{F}_{t_0}] = \frac{\eta^2}{2\lambda}\big(1 - \mathrm{e}^{-2\lambda t}\big). \tag{3.37}$$

For the derivative with respect to λ we obtain:

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\mathrm{Var}\big[r(t)\big|\mathscr{F}_{t_0}\big] = -\frac{\eta^2 \mathrm{e}^{-2\lambda t}(-2\lambda t + \mathrm{e}^{2\lambda t} - 1)}{2\lambda^2}. \tag{3.38}$$

Since $-2\lambda t + e^{2\lambda t} - 1 > 0$ for all $t > 0$ and $\lambda \neq 0$, the expression in Equation (3.38) is negative for $t > 0$ and $\lambda \neq 0$. Consequently, The Equation (3.37) is a decreasing function for λ and $t > 0$. For η it is the other way around. Notice that Equation (3.37) is an increasing function for η. Therefore, the larger the η, the larger the variance of process $r(t)$. This can be observed in Figure 3.6.

Naturally, the variance of a the process $r(t)$ and the implied volatility of a swaption are positively correlated. The larger the variance of process $r(t)$, the larger the implied volatility. If we calibrate η to the implied volatility of swaptions, the larger λ we use, the larger η we obtain in return to compensate for the smaller variance caused by the λ.

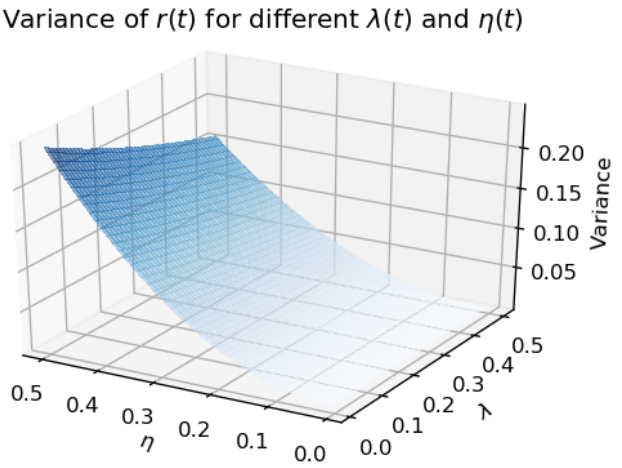

Figure 3.6: In the plot one can find the variance of stochastic process $r(t)$ for $t = 1$ several mean reversion λ and volatility η. In general, a larger mean reversion λ decreases the variance and a larger implied volatility increases the variance.

CIR process stress testing

Besides using artificial data simulated by the Hull-White model, we would like to simulate data from other short rate models to test our calibration of the Hull-White model. Another well known model is the Cox-Ingersoll-Ross (CIR) model [15]. The CIR process has the following dynamics:

$$\mathrm{d}r(t) = \lambda_{CIR}(t)\big(\theta_{CIR}(t) - r(t)\big)\mathrm{d}t + \eta_{CIR}(t)\sqrt{r(t)}\mathrm{d}W(t). \tag{3.39}$$

The CIR process is very similar to the Hull-White model. Both have a mean reversion and long-term mean parameter. The difference is in the $\sqrt{r(t)}$ in the stochastic part of the dynamics of Equation (3.39). Notice that the CIR process does not allow negative rates, therefore it is fair to assume that interest rates in the real world do not follow a CIR process. Still, one could shift the CIR process, to allow negative interest rates.

By simulating data from the CIR process, we would like to find out how our calibration of the Hull-White model performs. Because the CIR model and Hull-White model have different dynamics, we do not expect the calibrated $\lambda_{HW}(t)$ and $\eta_{HW}(t)$ to be the same as the $\lambda_{CIR}(t)$ and $\eta_{CIR}(t)$. In Figure 3.7 we can find the calibration result on artificial market data simulated by the CIR model. We find that the calibrated parameters are off from the parameters from the CIR model, however, the sinus curve remains visible in the calibration. The implied volatilities from the artificial data and generated by the calibrated Hull-White model have a good match; the difference is smaller than 10^{-7} at $T_i = 1, 2, ..., 9$.

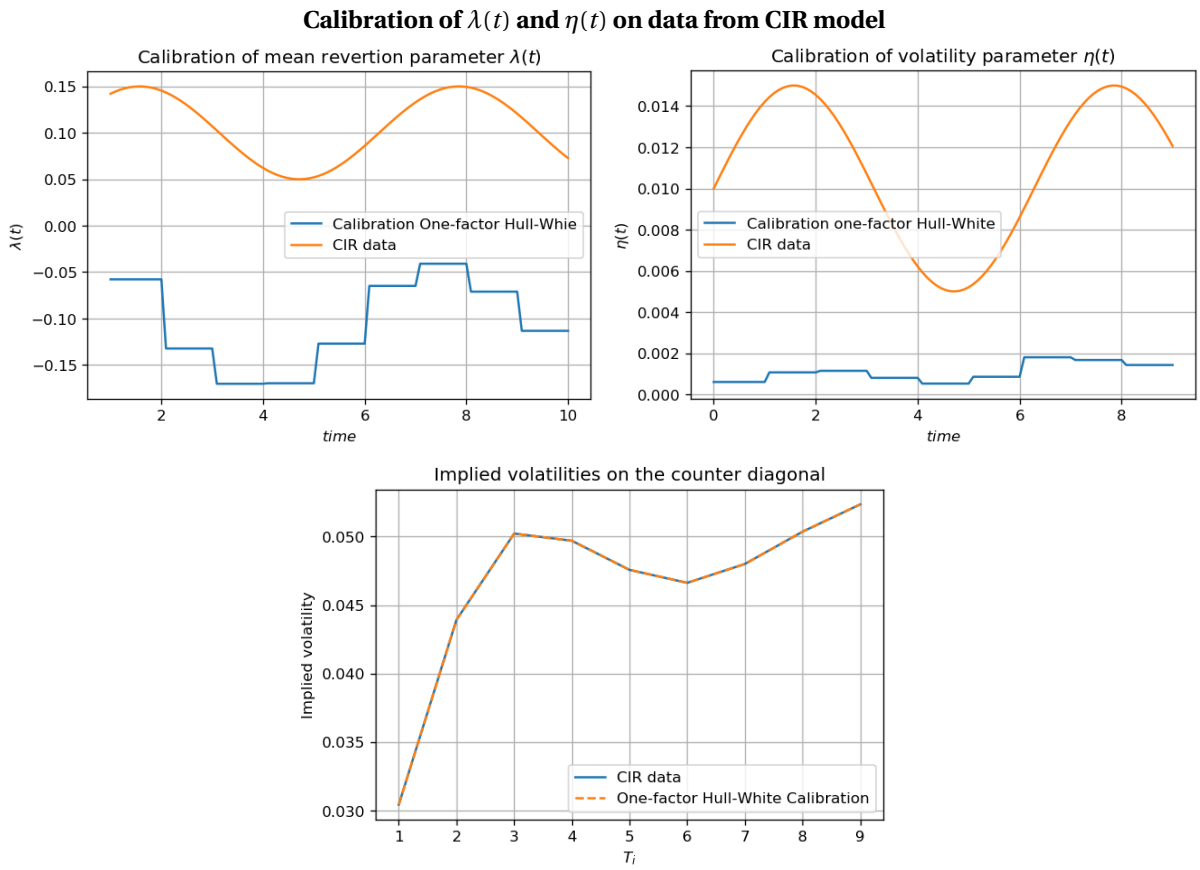

Figure 3.7: Data has been simulated under the CIR model with $\lambda_{CIR}(t) = 0.1 + 0.05 \cdot \sin(t)$ and $\eta_{CIR}(t) = 0.01 + 0.005 \cdot \sin(t)$. Zero coupon bond curve is observed from the market. $\lambda_{CIR}(t)$ and $\eta_{CIR}(t)$ are chosen in such a way that the obtained implied volatilities are comparable to the implied volatilities observed in the market.

However, we do obtain negative $\lambda(t)$ for all t. As discussed in Section 2.1.1, this is not allowed under the Hull-White model as such mean reversion $\lambda(t)$ will make the dynamics of the short rate explode. Therefore, this result gives us the insight that the Hull-White model is not suitable to capture this data, and another model would be more appropriate to use for calibration. This finding makes sense, since the dynamics of these data are indeed not from the Hull-White model, but from the CIR process. One can argue whether this finding is a positive or negative. One could argue that the calibration process should work for all possible data sets given by the market, also if the dynamics are from another process. On the other hand, one can argue that spotting that the data set is not from Hull-White dynamics is better.

3.7.2. Market data

Besides calibration of artificial data, one wants to perform calibration on real market data. Market data is more challenging. When we use artificial data, we assumed that the data comes from a Hull-White model, and specifically generated data from the Hull-White model. For market data, the data is not generated by a specific Hull-White model. The market data is determined by supply and demand, what are traders willing to pay for specific financial products. However, as mentioned before, it is more difficult to assess the calibration of market data as there is not a $\lambda(t)$ and an $\eta(t)$ in the market data. Luckily, there is one alternative calibration provided in the data, which is useful as comparison. The calibration is performed to the 30 year co-terminal basket of swaptions shown in Table 3.4. Two calibration strategies will be discussed.

- In the first strategy $\lambda(t)$ is chosen as a constant by the user, in our case $\lambda = 0.00001$. Afterwards, $\eta(t)$ will be calbirated as time-dependent piecewise-constant function.

- The second strategy does not allow the user to choose a parameter beforehand. Instead, $\lambda(t)$ will be calibrated as a time-dependent piecewise-constant function to ratio of implied volatilities of swaptions as mentioned in the previous section, afterwards $\eta(t)$ is calibrated to the implied volatilities.

30y co-terminal basket			
Expiry	Tenor	IV	Strike
1	29	0.209	ATM
2	28	0.202	ATM
3	27	0.197	ATM
4	26	0.192	ATM
5	25	0.188	ATM
6	24	0.186	ATM
7	23	0.185	ATM
8	22	0.184	ATM
9	21	0.183	ATM
10	20	0.183	ATM

Table 3.4: Swaption implied volatilities of the 30 years co-terminal basket is used for the calibration. Strike prices are at-the-money.

Choosing λ constant

A method of calibration in literature is choosing the mean reversion parameter λ as a constant [24], and with this constant λ one calibrates piece-wise constant $\eta(t)$. We choose $\lambda = 0.00001$ and use 30 years co-terminal basket for calibration of $\eta(t)$. We obtain the following optimization problem

$$\eta(t) = \arg\min_{\eta(t)} \sqrt{\sum_i \left(\sigma_{imp}^{mrkt}(t_0, T_i, T_{30}, K) - \sigma_{imp}^{HW}(t_0, T_i, T_{30}, K, \lambda, \eta(t))\right)^2},$$

for $i = 1, 2, \ldots, 10$ with $T_1 = 1, T_2 = 2, \ldots$. Besides the data from Bloomberg that has been available for our research, a calibration from the industry on the data was available as well. We will refer to this calibration from the industry as alternative calibration. This alternative calibration is only available for $\lambda = 0.00001$. This alternative calibration is useful for comparison, and will be included in plots when it is appropriate. We obtain the calibration shown in Figure 3.8.

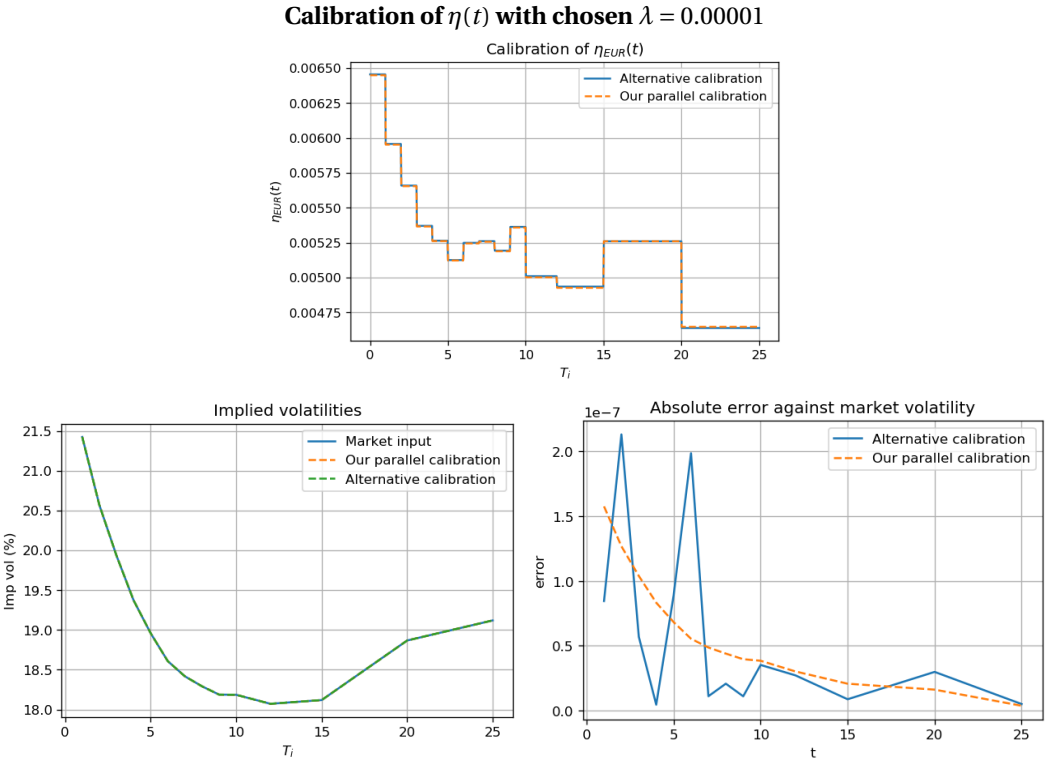

Figure 3.8: Parallel calibration of $\eta(t)$ with constant chosen $\lambda = 0.00001$ compared to an alternative calibration. Zero coupon bond curve is observed from the market.

As one can observe from Figure 3.8, the calibration gives a good match with the alternative calibration and the implied volatilities observed from the market. However, we argue that this is not the best method for calibration, as we showed in the previous section that choosing λ arbitrary can impact the calibration. By calibrating $\lambda(t)$ first, we find that the calibration of $\eta(t)$ changes significantly (see Figure 3.5).

Including calibration of $\lambda(t)$

We calibrate our $\lambda(t)$ to the set consisting of k fractions of swaptions with $T_i = 1, 2, \dots$, $T_{m_1} = T_i + 0.5$ and $T_{m_2} = T_i + 1$. We use the method as discussed in Section 3.4. Note that for this set of swaptions we indeed have that $T_i < T_{m_1} < T_{m_2}$ and for all $n, p \in \{1, \dots, k\}, n \neq m$ that $(T_i^n, T_{m_2}^n) \bigcap (T_i^p, T_{m_2}^p) = \emptyset$. Here (A, B) denotes the open interval from A to B. Therefore, we can use sequential calibration as discussed in section 3.6. First we calibrate $\lambda(t)$. Besides calibrating $\lambda(t)$ as a piece-wise constant time-dependent parameter, we calibrate the mean reversion parameter as a constant as well. We compare the results which can be found in Figure 3.9.

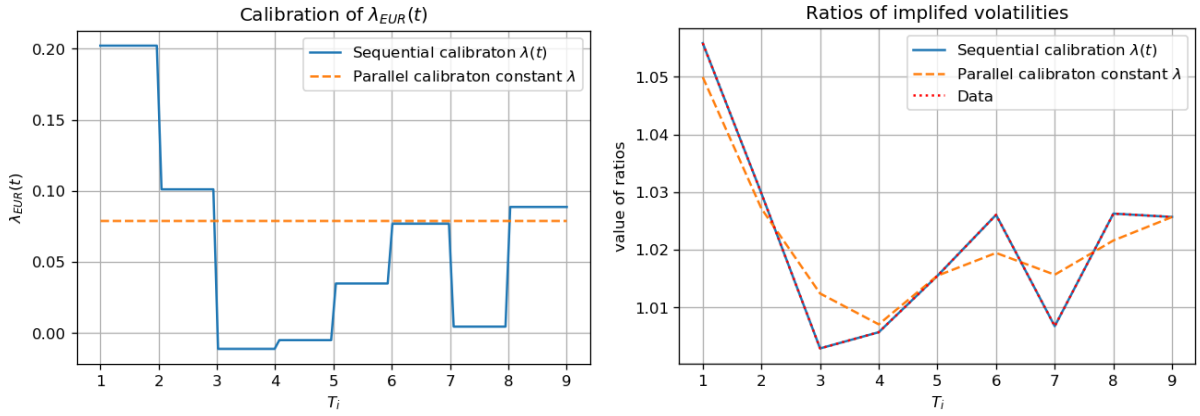

Figure 3.9: In the plot on the left one can observe the calibration of $\lambda(t)$ and constant λ to market data. On the right we observe the corresponding implied volatility for our calibration and constant λ. The ratios of implied volatilities and calibration have a good match, whilst the constant λ is slightly off. Sequential calibration computation time for time-dependent $\lambda(t)$ is 0.13 seconds. Parallel calibration computation time for constant λ is 0.17 seconds.

In Figure 3.9, we find that time-dependent $\lambda(t)$ gives perfect match with data, whilst calibrating a constant λ results in a worse match. We also observe that time-dependent calibration has big swings in the resulted calibrated $\lambda(t)$. For a small moment in time the $\lambda(t)$ is even negative. We argued that in Section 2.1.1 that this is allowed for short period of times, just like this case. In order to judge whether a time-dependent $\lambda(t)$ or constant λ is better for calibration, one has to determine what is more important: jumps in the mean reversion parameter or a better fit to the ratios of implied volatilities. In this thesis, we will not argue what is the best method, but we do argue that calibration is better than picking constant λ arbitrary.

Next, we continue with the calibration of $\eta(t)$. We include $\lambda = 0.00001$ in the plots for comparison. The results can be found in Figure 3.10

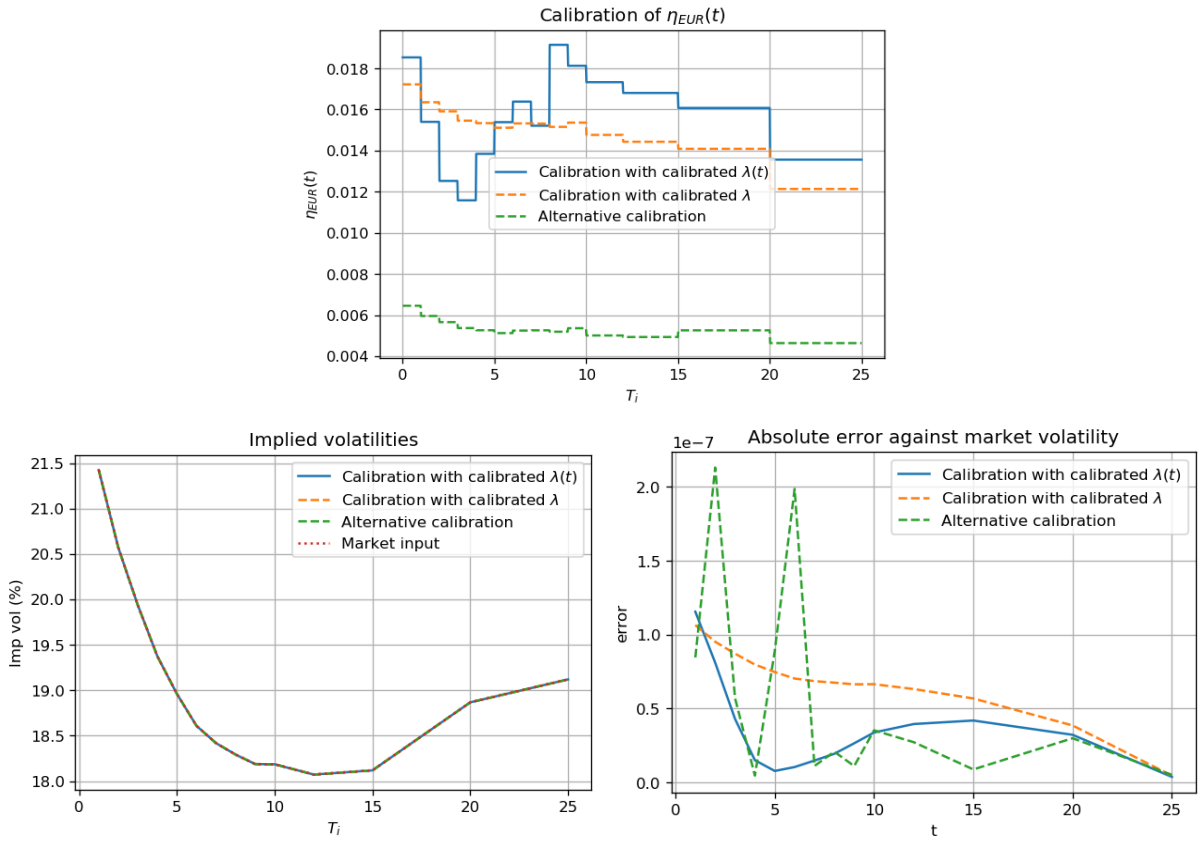

Figure 3.10: In the plot on the top we find the calibration of piece-wise constant $\eta(t)$ to the market data. On the bottom left we can observe the implied volatilities obtained by each calibration and on the right we find the error of the implied volatility compared to the market. Total calibration time is 15 seconds by Levenberg-Marquardt and 5 seconds by dogbox method.

Figure 3.10 shows us that there is a big difference in calibration of $\eta(t)$, when we use a calibrated mean reversion parameter compared to an arbitrarily chosen constant $\lambda = 0.00001$. As shown in the previous Section 3.7.1, an arbitrary constant λ can cause the calibration of $\eta(t)$ to be far off the real $\eta(t)$ used for the creation of the data. Therefore, we argue that calibrating $\lambda(t)$ is essential. Especially in Figure 3.10, where we find that arbitrarily chosen λ and calibrated $\lambda(t)$ have a big impact on the calibrated volatility parameter $\eta(t)$. Whether we prefer to use a time-dependent $\lambda(t)$ for calibration or a constant λ, remains to be seen. For the industry, calibrating λ as a constant might be preferable as this code for calibration of $\eta(t)$ can stay the same. Also, constant λ significantly lowers the jumps in the calibration of $\eta(t)$, which can be considered as a benefit. One could apply constraints to the calibration of $\eta(t)$, e.g. for all t_i $\eta(t_{i+1}) \geq \frac{1}{2}\eta(t_i)$ and $\eta(t_{i+1}) \leq \frac{3}{2}\eta(t_i)$. In Section 5 we argue that having big jumps in time dependent $\eta(t)$ is not a problem for delta hedging if one adjusts his/her hedging strategy. Therefore, using such constraints is not a necessity. Besides, one can argue that big swings in interest rate do happen in the real world. As a recent example, we have the Covid 19 pandemic in March 2020. In those situations, the volatility parameter $\eta(t)$ can increase a lot, causing a huge jump in the piece-wise constant function. This could happen from February to March in the Covid 19 for example, under the assumption that we calibrate the volatility parameter $\eta(t)$ per month. Putting constraints on $\eta(t)$ might be incorrect for that reason.

3.7.3. Performance on data not used for calibration

Over the years, there is argued that calibration of the Hull-White model cannot capture the whole swaption volatility surface [32]. In this section, we look into the performance of our calibration on implied volatilities of swaptions that are not used for the calibration.

Fit to ATM implied volatilities not used for calibration

Although we are not able to fit the whole swaption matrix, looking at the whole swaption matrix for different mean reversion parameters with corresponding calibrated volatility parameter $\eta(t)$ is meaningful. Preferably,

we would like our (calibrated) mean reversion parameter $\lambda(t)$ and calibrated volatility parameter $\eta(t)$ to be as good as possible at fitting the whole swaption matrix. Still, our first priority remains to get a perfect fit for the co-terminal basket used for the calibration of $\eta(t)$. In Figure 3.11, we look into the results on implied volatility by different calibrations of constant λ.

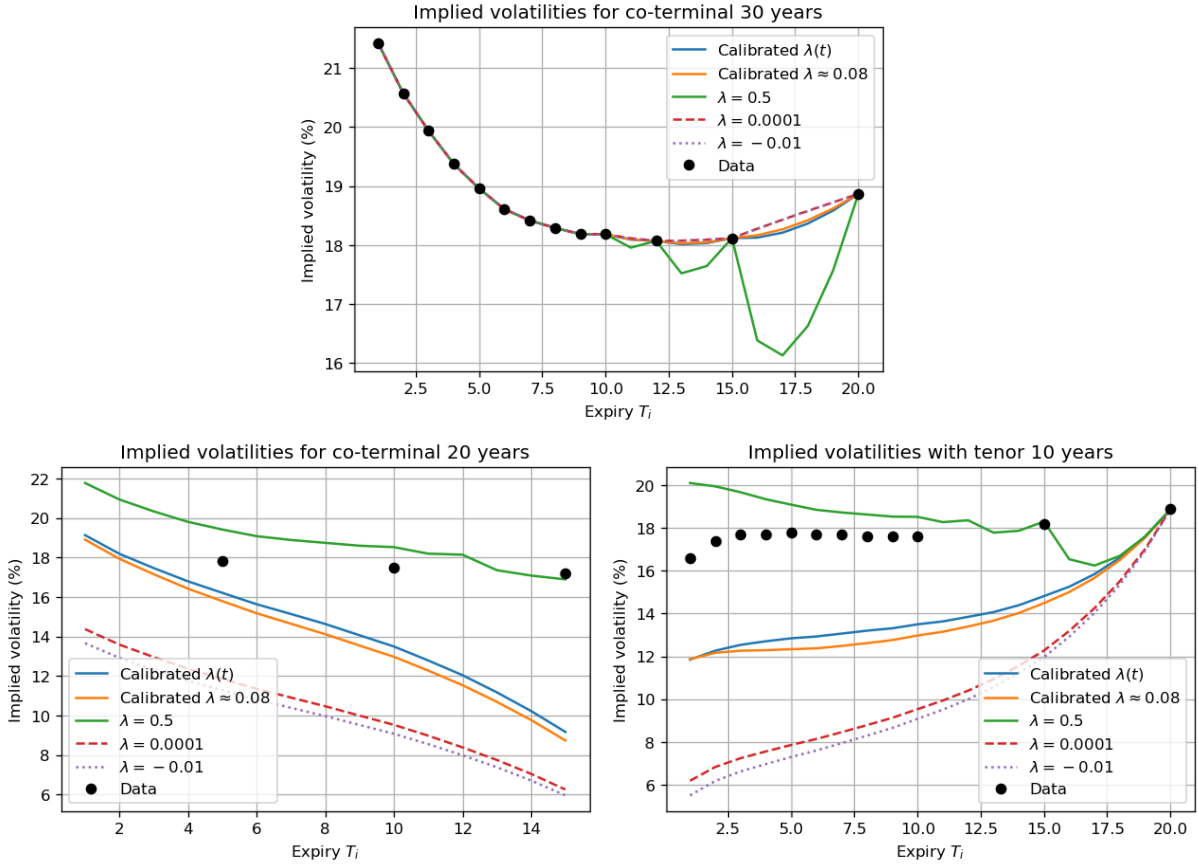

Figure 3.11: In the plot on the top we find the implied volatilities obtained from our calibrated parameter $\eta(t)$ and constant λ for the 30 year co-terminal basket which was used for the calibration. On the bottom left we can observe the implied volatilities for the co-terminal 20 years basket. On the right the implied volatilities of swaptions with tenor of 10 years is shown.

The upper picture in Figure 3.11 shows us the calibration on the co-terminal basket used for calibration of $\eta(t)$. As expected, we see a perfect fit on the data points. We observe that the choice of a large λ, gives a perfect fit on the data used for calibration, but behaves badly for the computed swaption implied volatilities in between the calibrated points. The curves from seem odd at first, but for large mean reversion parameter in combination with piece-wise constant volatility parameter you would actually expect these kind of shapes. These shapes are explained in Appendix B. Because these shapes arise with large mean reversion parameter, one could decide to include a constrain on the maximum of the mean reversion parameter for the calibration.

Remark 3.7.1. Large mean reversion parameter appears to be inappropriate for the calibration of the Hull-White model in combination with piece-wise constant volatility parameter. In Appendix B one can find how large mean reversion influences the value of swaptions. In Figure B.1 one observes that the swaption values under large mean reversion look very odd, and it looks like arbitrage possibilities are present. In numerical results in literature, the mean reversion often appears to be in the range of -0.03 and 0.10 ([28],[24],[39], [13]). Although a specific reasoning and quantificiation for a maximum and minimum seems to be missing. Quantification for the maximum and minimum of the mean reversion parameter for calibration would be an interesting topic for future research.

For small and negative λ, the values of the swaption between the calibrated points are more linear compared to the calibrated λ in the upper picture of Figure 3.11. Looking at the bottom two pictures, we observe that the fit of the our calibration with the implied volatilities from the data is poor. In both cases, we find that

the implied volatilities obtained with the calibrated mean reversion is closer to the data. From the results in Figure 3.11 one could argue that the calibrated mean reversion outperforms the other choices of λ.

Fit to ITM and OTM swaptions

In our calibration procedure we use at the money swaptions as we argued that this is a proper choice in Section 3.5. We wonder how our calibrated mean reversion parameter $\lambda(t)$ and mean reversion parameter $\eta(t)$ perform when computing the implied volatility of swaption that are not at the money. Figure 3.12 shows us the implied volatility of the swaption plotted against the strike price. In order to determine a suitable set of strike prices, the method described in section 3.5 for the computation of the strike prices was used. unfortunately, due to the fact that $S_{i,m}(t_0) \approx -4 \cdot 10^{-4}$ is close to zero, we obtain a very small set of strike prices. This felt inappropriate, and a range from -0.02 to 0.03 has been chosen as range for the strike prices. We can observe that the difference between the calibration is small, especially compared to the differences in the previous section, where the implied volatilities from the ATM swaption surface have been studied.

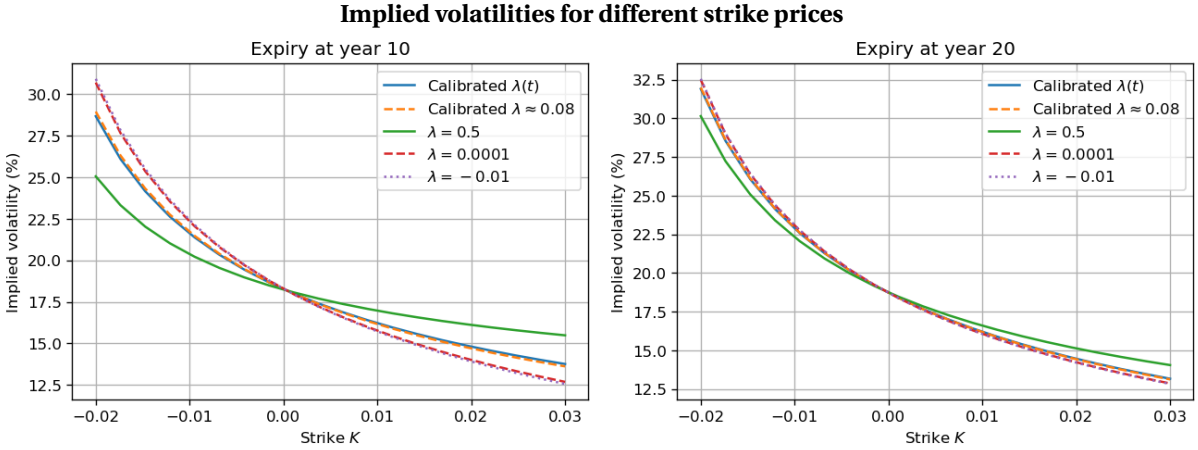

Figure 3.12: Shown are the implied volatilities for different strike prices under various mean reversion parameters, where we looked at two swaption from the 30 year co-terminal set used for the calibration of the volatility parameter $\eta(t)$. In the plot on the left, the expiry of the swaption is 10. In the plot on the right, the expiry of the swaption is 20. In both cases T_m is 30, $\tau = 0.5$ and discount curve is observed from the market.

3.7.4. Calibration on whole swaption matrix

By calibrating $\eta(t)$ on the ATM co-terminal swaption basket, the fit with implied volatilities outside of the co-terminal swaption basket is poor, as one can observe in Figure 3.11. A potential solution for this problem is calibration to the whole swaption matrix. We hope for better overall fit of the swaption matrix when we perform calibration to the whole matrix. In order to calibrate to the whole matrix, we will have to calibrate $\lambda(t)$ and $\eta(t)$ simultaneously. The optimization problem is given by:

$$\underset{\eta(t),\lambda(t)}{\operatorname{argmin}}\left(\sqrt{\sum_{i,j}\left(\sigma^{mrkt}_{swpt}(t_0,T_i,T_j,K)-\sigma^{HW}_{swpt}(t_0,T_i,T_j,K,\lambda(t),\eta(t))\right)^2}\right). \tag{3.40}$$

Results of this calibration procedure can be found in Figure 3.13. We find that calibrating to the whole swaption matrix is computational expensive. Therefore, perhaps not applicable in the industry. Also, our calibration is poor, although in non-negative interest rate environments is argued that the Hull-White model is able to capture the whole swaption matrix till some extent [24]. Others argue that the Hull-White model is not capable of calibrating the whole matrix [32]. When calibrating to the whole swaption matrix, one loses control over which implied volatilities from the swaption matrix are fitted well to the data, and which are not. Therefore, more important implied volatilities might not be fitted well to the data. One could solve this problem by introducing weights to the swaptions implied volatilities. Giving the more important swaption implied volatility a bigger weight than the other swaption implied volatilities.

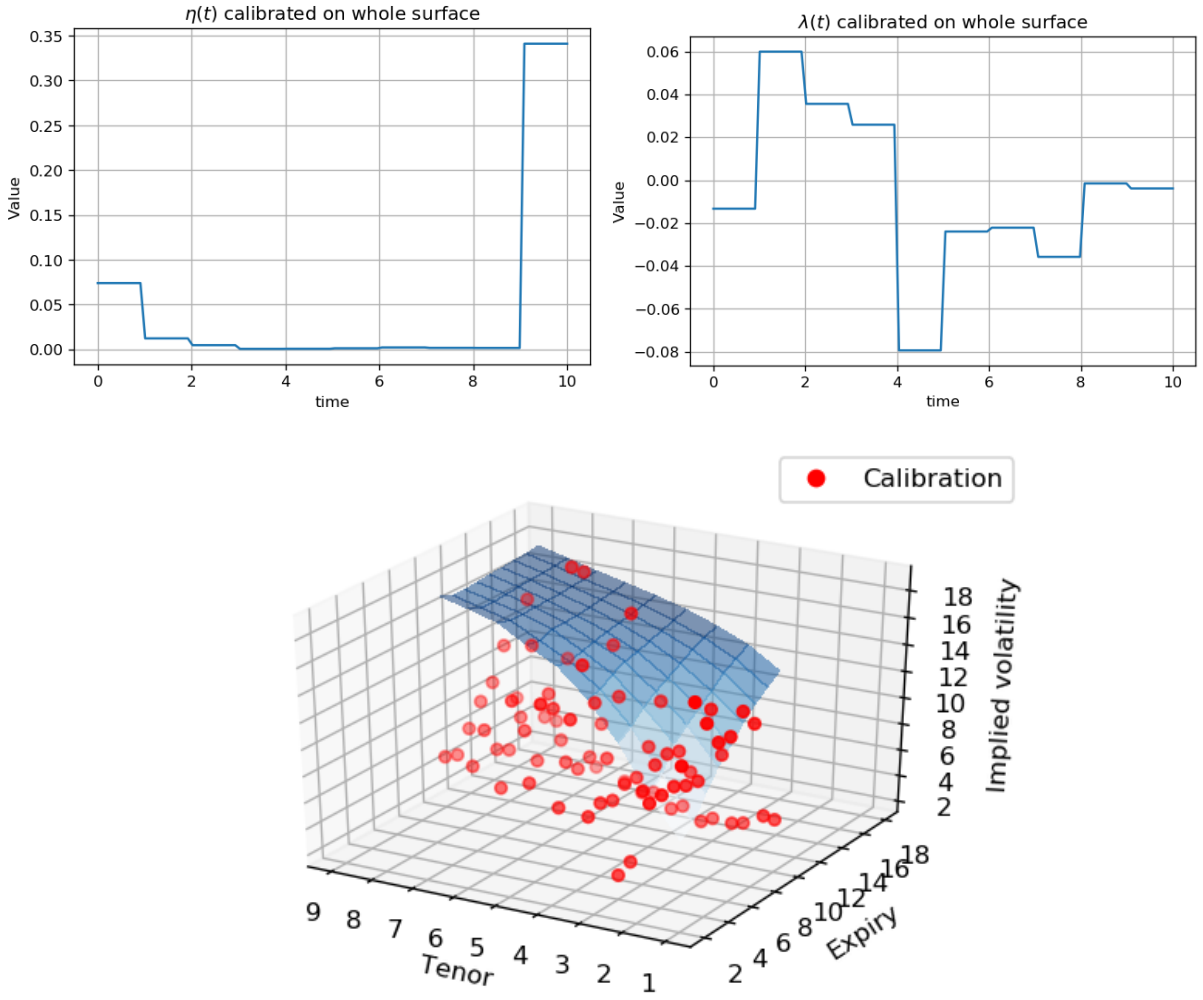

Figure 3.13: In the plots we find the calibration to the whole swaption surface. We find that the calibration is poor. Levenberg Marquardt has been used as optimization algorithm, takes 10+ minutes to calibrate.

3.8. Effect of calibration on exposure

Since the 2008 financial crash, a set of rules has been built to make sure banks do not take too much risk regarding exposure to a default. In the industry, new teams have been installed to manage these risks, and validate the models that are used for pricing. In order to deal with the risk of default, the credit valuation adjustment (CVA) is computed. The CVA is the difference between the risk-free portfolio and the portfolio that does consider risk of default. One can also interpret the CVA as the price of the counterparty credit risk. In this section the effect of different calibrations on the CVA will be discussed.

Before we dive into the CVA, we will need to define some exposures. Credit exposure is defined as the potential future losses without any recovery. Mathematically, this is defined as [25]:

$$E(t) := \max\big(V(t), 0\big), \qquad V(t) := \mathbb{E}^{\mathbb{Q}}\Big[\frac{M(t)}{M(T)} H(T,\cdot)\Big],$$

where $H(t,\cdot)$ represents the payoff of the contract at time T. Exposure changes over time, as prices of financial products change in the market. It often cannot be expressed in close-form. To quantify the exposure, we use various metrics. One of them being the expected (positive) exposure (EE), defined as:

$$EE(t_0, t) = \mathbb{E}^{\mathbb{Q}}\Big[\frac{M(t_0)}{M(t)} E(t) | \mathscr{F}(t_0)\Big]. \tag{3.41}$$

The exposure profile needs to be computed for a large number of scenarios. Therefore we need numerical methods, by means of Monte Carlo simulation. This is one of the extensive parts in the computation of the Credit Value Adjustment (CVA). These numerical methods contain essentially three stages: a forward stage for generating future asset path scenarios, a second stage to value the financial derivatives and a final stage to calculate the exposure along the generated future asset paths. An example of the expected exposure for a single swap portfolio can be found in Figure 3.14.

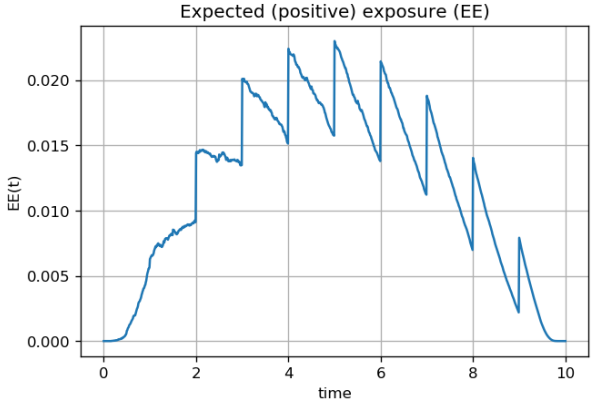

Figure 3.14: The expected exposure of the swap with $K = 0.03$ and payments are done yearly. One can observe that for every payment there is a big shift in value of the expected exposure.

CVA represents the price of counterparty credit risk. For the computation of the CVA, we need three components: The loss given default, the expected exposure and the probability of default. The loss given default can be expressed as one minus the recovery rate R. The probability of default can expressed as a survival probability in the following way: Let X denote a random variable which denotes a lifetime, for example $\mathbb{P}[X < x]$ indicates the probability of not reaching the age x in a population. The survival probability is defined as [25]:

$$\mathbb{P}[X > x] = 1 - \mathbb{P}[X \leq x] = 1 - F_X(x).$$

Using the expression for the LGD and the probability of default, The formula for computation of the CVA is given by:

$$CVA(t,T) \approx LGD \cdot \sum_{i=1}^{M} EE(t,T_i) \cdot PD = (1-R) \sum_{i=1}^{M} EE(t,T_i) \cdot \big(F_{t_D}(T_i) - F_{t_D}(T_{i-1})\big),$$

where LGD is assumed to be a fixed ratio based on market information, and the probability of default can often be obtained via the survival prbability curve in the Credit Default Swap market. This thesis captures the effect of calibration on the expected exposure. The loss given default and probability default are not in the scope of this work.

From the expression in Equation 3.41, we obtain for the expected exposure of a swap:

$$EE(t) = \mathbb{E}^{\mathbb{Q}}\Big[\frac{M(t_0)}{M(t)}\max\big(V_{Swap}(t),0\big)\Big]. \tag{3.42}$$

Notice that the expression in Equation (3.42) is similar to value of a swaption on the specific swap with strike 0 and expiry t, but this merely works for swaps with $K = 0$. A numerical example of this shown in Figure 3.15,

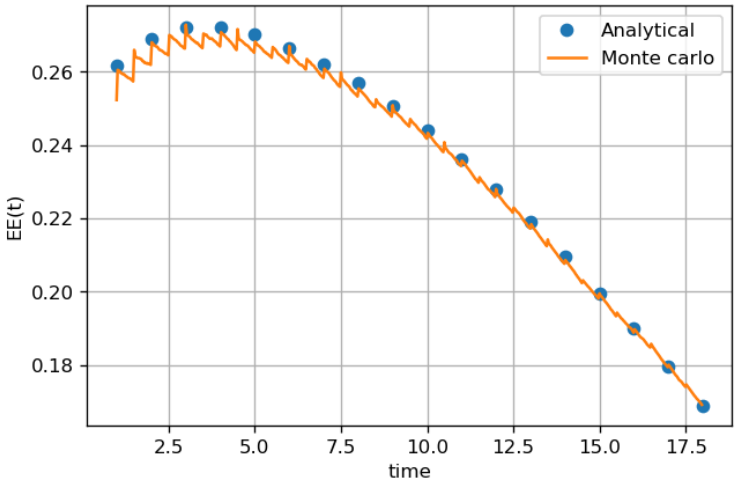

Figure 3.15: Numerical example where we show that the expected exposure can be computed by means of monte carlo methods and analytical. Monte carlo method is computed by simulating paths for the swaps, afterward computing expression in Equation (3.42). Analytical is computed by the analytical expression of the value of a swaption from Theorem 2.1.2.

For our research, we would like to investigate how different mean reversion parameter affect the expected exposure. Therefore, we pick several mean reversion parameters λ, and calibrate the volatility parmeter $\eta(t)$ for every mean reversion λ. We end up with pairs of $(\lambda, \eta(t))$ that match the market data. Then, for every pair, we look into the expected exposure of the swap that matches the co-terminal basket used for calibration. Results on the market data can be found in Figure 3.16.

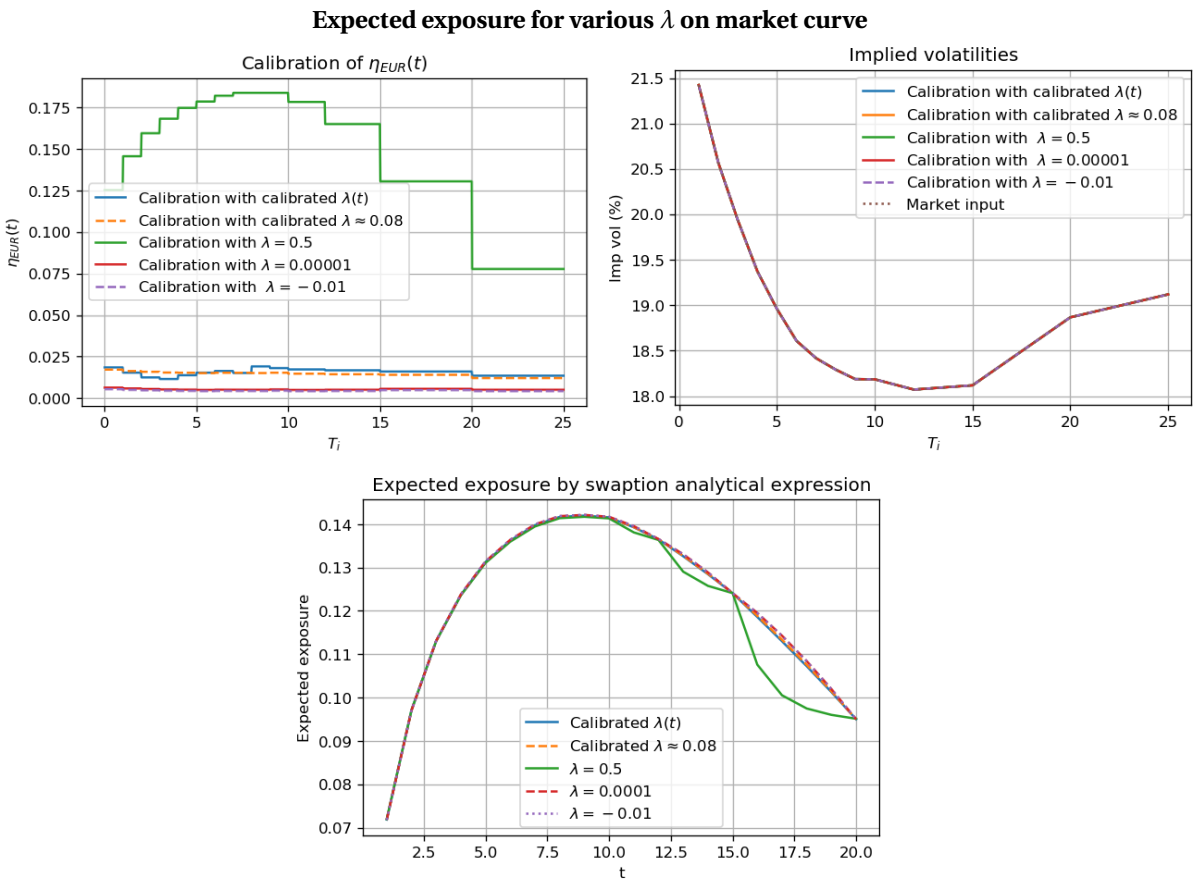

Figure 3.16: Shown in the plots is the expected exposure of a swap for various calibrations. The maturity of the swap is 30 years. The notional level is 1. On the top one can observe the calibration of $\eta(t)$ with corresponding volatilities. On the bottom, the expected exposure is computed by means of the value of the swaption with strike 0 has been computed at integer timepoints. We used the zero coupon bond curve observed from the market. For each λ, $\eta(t)$ is calibrated to the 30 year co-terminal basket.

Figure 3.16 shows that there is little difference in the expected exposure for various mean reversion parameters λ used for calibration of volatility parameter $\eta(t)$. Only for large λ, we have the same problem as encountered in the previous section. A reason why there is so little difference in the expected exposure could be the near zero interest-rate environment in the market data. We will explain this by looking at the ATM strike prices. The ATM strike prices are defined as the swap rate as in Definition 3.5.1:

$$K_{ATM} = S_{i,m}(t_0) = \frac{P(t_0, T_i) - P(t_0, T_m)}{\sum_{k=i+1}^{m} \tau_k P(t_0, T_k)}$$

In a near zero interest rate environment, we have that $P(t_0, T_i) \approx 1$ and $P(t_0, T_m) \approx 1$. Therefore, the numerator in the expression for the ATM strike price is very close to zero. For the market data the ATM strike prices are in the order of 10^{-4}. When the ATM strike price is close to zero, the data for calibration is very similar to the expression in Equation (3.42), where strike is zero. In interest rate environments that cause the ATM strike price to be not close to zero, we might obtain bigger differences in the expected exposures for various mean reversion parameters λ. To show this, we used an artificial zero coupon bond curve $P(0,t) = e^{-0.01 \cdot t}$, and the results can be found in Figure 3.17.

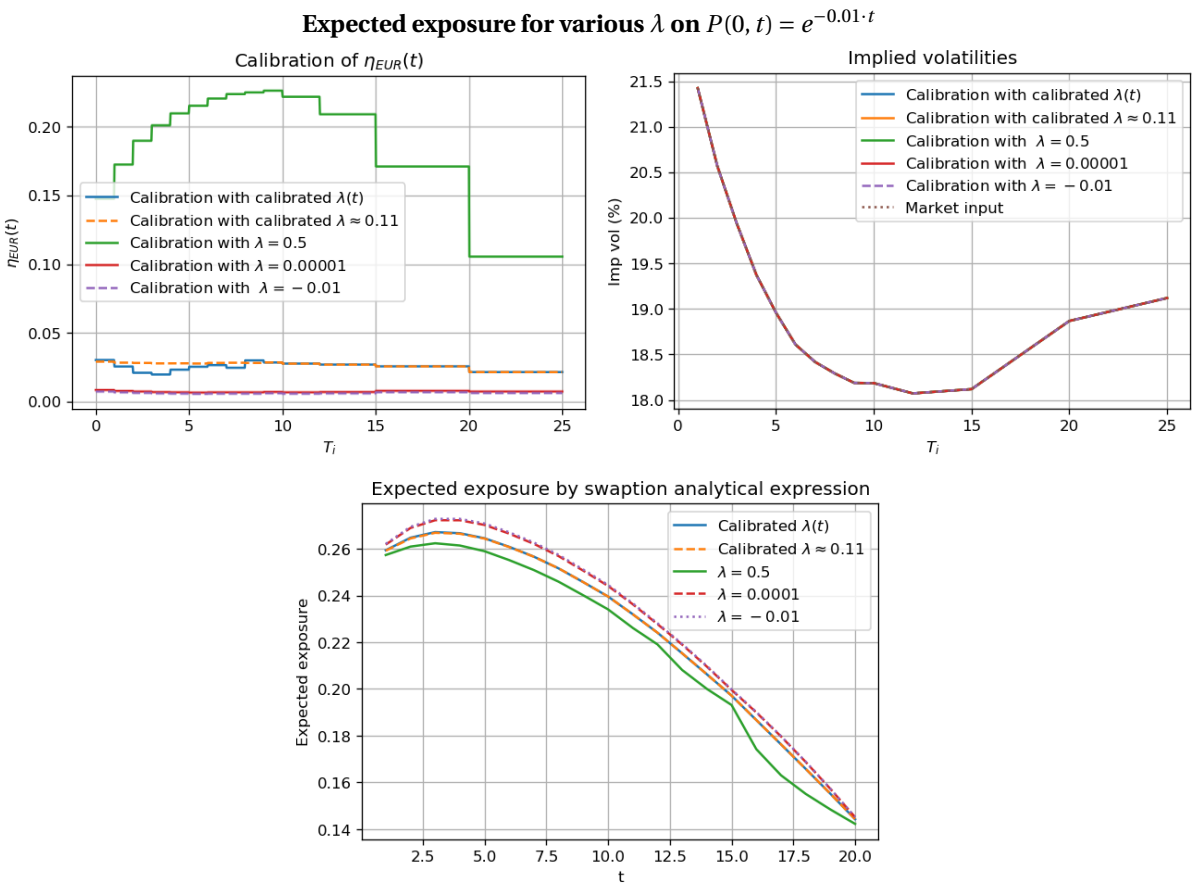

Figure 3.17: Shown in the plots is the expected exposure of a swap for various calibrations. The maturity of the swap is 30 years. The notional level is 1. On the top one can observe the calibration of $\eta(t)$ with corresponding volatilities. On the bottom one can find the expected exposure computed at integer timepoints by the swaption analytical expression. We used the zero coupon bond curve $P(0,t) = \mathrm{e}^{-0.01 \cdot t}$. For each λ, $\eta(t)$ is calibrated to the 30 year co-terminal basket.

In conclusion, we argue that the choice of mean reversion parameter λ before calibration of volatility parameter $\eta(t)$ can have an effect on the expected exposure, therefore has an effect on the credit valuation adjustment.

4

FX Black-Scholes Hull-White model

This chapter covers the calibration of the FX Black-Scholes Hull-White model. The FX Black-Scholes Hull-White model is an example of a hybrid model. In a hybrid model various asset classes are modelled. Typically a hybrid model is represented as a system of stochastic differential equations (SDEs) for these different asset classes, e.g. SDEs for stock price and interest rate. In this thesis we cover the FX Black-Scholes Hull-White model, which covers three asset classes: The FX exchange rate, the domestic interest rate and the foreign interest rate. In the first section information will be given about the FX market, afterwards we continue with FX option pricing with the Black-Scholes Hull-White model. At the end, calibration results on the Black-Scholes Hull-White model are shown.

4.1. FX market

The FX market is an abbreviation for the foreign exchange market. In this market currencies are traded, for example the dollar versus the euro. We define $y_d^f(t)$ to be the exchange rate from the domestic currency d to a foreign currency f. For notation, we write:

$$y(t) := y_d^f(t).$$

For example, if one want to trade euros for dollars, the euro is perceived as the domestic currency and the dollar as the foreign currency. For on euro, one will receive $y(t)$ dollars.

Under the real world measure $\mathbb{P}$ the exchange rate $y(t)$ follows the following stochastic process [25]:

$$\mathrm{d}y(t) = \mu y(t)\mathrm{d}t + \sigma y(t)\mathrm{d}W^{\mathbb{P}}(t). \tag{4.1}$$

We will need to switch from the real world measure $\mathbb{P}$ to the risk-neutral measure. Note that at this point there are two different risk neutral measures. We have the domestic risk-neutral measure, but also the foreign risk-neutral measure. We have two different money market account $M_f(t)$ and $M_d(t)$, one in the domestic currency and one in the foreign currency, with corresponding interest r_f and r_d. In order to obtain a closed form solution for FX options, We switch from the real-world measure to the domestic risk-neutral measure. We define the process $X(t) = y(t)\frac{M_f(t)}{M_d(t)} = y(t)e^{(r_f - r_d)t}$, which should be a martingale under the domestic risk-neutral measure. Using Itô's lemma, we obtain:

$$\mathrm{d}X(t) = (r_f - r_d)y(t)\frac{M_f(t)}{M_d(t)} + \frac{M_f(t)}{M_d(t)}\mathrm{d}y(t).$$

By plugging in Equation (4.1) for $y(t)$, we obtain:

$$\frac{\mathrm{d}X(t)}{X(t)} = (r_f - r_d + \mu) + \sigma\mathrm{d}W^{\mathbb{P}}(t).$$

Since $X(t)$ should be a martingale under the domestic risk-neutral measure, we know that the dynamics should not contain a drift term. Therefore we find:

$$\mathrm{d}W^{\mathbb{P}}(t) = \frac{r_f - r_d + \mu}{\sigma}\mathrm{d}W^{\mathbb{P}}(t).$$

Plugging this finding into Equation (4.1) gives us the dynamics of $y(t)$ under the domestic risk-neutral measure:

$$\mathrm{d}y(t) = (r_d - r_f)y(t)\mathrm{d}t + \sigma y(t)\mathrm{d}W_d^{\mathbb{Q}}(t). \tag{4.2}$$

4.1.1. Options on FX rates

For the calibration of the FX hybrid model we will use FX options. We consider European options on the exchange rate $y(t)$. Note that $y(t)$ follows lognormal distribution, therefore we can obtain the value of an option in a Black-Scholes fashioned manner. In the Black-Scholes case, with strike price K, maturity T, constant foreign exchange rate r_f and domestic exchange rate r_d we have for a call option [6]:

$$V_c^{FX}(t_0, T, K) = y(t_0)\exp(-r_f(T-t_0))\Phi(d_1) - K\exp(-r_d(T-t_0))\Phi(d_2),$$

where Φ is the standard normal CDF. d_1 and d_2 are given by:

$$d_1 = \frac{\log(\frac{y(t_0)}{K}) + (r_d - r_f + \frac{1}{2}\sigma^2)(T-t_0)}{\sigma\sqrt{T-t_0}}, \qquad d_2 = d_1 - \sigma\sqrt{T-t_0}.$$

For a put option option we have:

$$V_p^{FX}(t_0) = Ke^{-r_d(T-t_0)}\Phi(-d_2) - y(t_0)e^{-r_f(T-t_0)}\Phi(-d_1).$$

4.2. FX options under FX Black-Scholes Hull-White model

We will extent Equation (4.2) to the Black-Scholes Hull-White model by modeling interest rates as time dependent parameters following the Hull-White model and modelling the volatility parameter $\sigma(t)$ as time-dependent:

$$\begin{cases} \mathrm{d}y(t) = (r_d(t) - r_f(t))y(t)\mathrm{d}t + \sigma(t)y(t)\mathrm{d}W_y^{\mathbb{Q}}(t), \\ \mathrm{d}r_d(t) = \lambda_d(t)\big(\theta_d(t) - r_d(t)\big)\mathrm{d}t + \eta_d(t)\mathrm{d}W_d^{\mathbb{Q}}(t), \\ \mathrm{d}r_f(t) = \lambda_f(t)\big(\theta_f(t) - r_f(t)\big)\mathrm{d}t + \eta_f(t)\mathrm{d}W_f^{\mathbb{Z}}(t), \end{cases} \tag{4.3}$$

where $\mathbb{Q}$ represents the risk-neutral measure for the domestic currency, and $\mathbb{Z}$ represents the risk-neutral measure under the for the foreign currency. All the Brownian motions are correlated with one another with correlations $\rho_{y,d}$, $\rho_{y,f}$ and $\rho_{d,f}$. We denote this in a full matrix of correlations:

$$\mathrm{d}\mathbf{W}(t)(\mathrm{d}\mathbf{W}(t))^T = \begin{pmatrix} 1 & \rho_{y,d} & \rho_{y,f} \\ \rho_{y,d} & 1 & \rho_{d,f} \\ \rho_{y,f} & \rho_{d,f} & 1 \end{pmatrix}. \tag{4.4}$$

In order to obtain an analytical expression for the value of an FX option under the Hull-White Black-Scholes model, we would like all stochastic processes to be driven by the risk-neutral measure of the domestic currency. Under the risk-neutral measure of the domestic currency we obtain we obtain an extra drift term for the dynamics of the foreign short rate [25]:

$$\begin{cases} \mathrm{d}y(t) = (r_d(t) - r_f(t))y(t)\mathrm{d}t + \sigma(t)y(t)\mathrm{d}W_y^{\mathbb{Q}}(t), \\ \mathrm{d}r_d(t) = \lambda_d(t)\big(\theta_d(t) - r_d(t)\big)\mathrm{d}t + \eta_d(t)\mathrm{d}W_d^{\mathbb{Q}}(t), \\ \mathrm{d}r_f(t) = \lambda_f(t)\big(\theta_f(t) - r_f(t) - \eta_f(t)\rho_{y,f}\sigma(t)\big)\mathrm{d}t + \eta_f(t)\mathrm{d}W_f^{\mathbb{Q}}(t). \end{cases} \tag{4.5}$$

Calibration of the BSHW model is done by first calibrating the domestic short rate paramters $\lambda_d(t)$ and $\eta_d(t)$ to the domestic zero coupon bond curve and swaptions. Afterwards the foreign short rate parameters, $\lambda_f(t)$ and $\eta_f(t)$ to the foreign zero coupon bond curve and swaptions. The correlations $\rho_{y,d}$, $\rho_{y,f}$ and $\rho_{d,f}$ can be observed in the market. At the end we only need to calibrate one more parameters, which is $\sigma(t)$. We can calibrate $\sigma(t)$ to options in the FX market. The domestic ZCB and foreign ZCB are respectively denoted by $P_d(t,T)$ and $P_f(t,T)$. Under the domestic risk-neutral measure $\mathbb{Q}$, the dynamics of the zero coupon bonds are given by [25]:

$$\begin{cases} \frac{\mathrm{d}P_d(t,T)}{P_d(t,T)} = r_d(t)\mathrm{d}t + \eta_d B_d(t,T)\mathrm{d}W_d^{\mathbb{Q}}(t), \\ \frac{\mathrm{d}P_f(t,T)}{P_f(t,T)} = \big(r_f(t) - \rho_{y,f}\eta_f B_f(t,T)\big)\mathrm{d}t + \eta_f B_f(t,T)\mathrm{d}W_f^{\mathbb{Q}}(t). \end{cases} \tag{4.6}$$

In order to price options in an exact manner under the Black-Scholes Hull-White model, we perform a change of measure from the domestic risk-neutral measure to the domestic T-forward measure. Therefore, we define the forward exchange rate under the domestic T-forward measure [23]:

$$y_F(t,T) := y(t)\frac{P_f(t,T)}{P_d(t,T)}. \tag{4.7}$$

Applying Itô's lemma gives us the dynamics of the exchange rate under the domestic T-forward measure

$$\begin{aligned}\mathrm{d}y_F(t,T) =& \frac{P_f(t,T)}{P_d(t,T)}\mathrm{d}y(t) + \frac{y(t)}{P_d(t,T)}\mathrm{d}P_f(t,T) - y(t)\frac{P_f(t,T)}{\big(P_d(t,T)\big)^2}\mathrm{d}P_d(t,T) + \frac{1}{P_d(t,T)}\big(\mathrm{d}y(t)\mathrm{d}P_f(t,T)\big)\\ &- \frac{P_f(t,T)}{\big(P_d(t,T)\big)^2}(\mathrm{d}y(t)\mathrm{d}P_d(t,T)) - \frac{y(t)}{\big(P_d(t,T)\big)^2}(\mathrm{d}P_d(t,T)\mathrm{d}P_f(t,T)) + \frac{y(t)P_f(t,T)}{\big(P_d(t,T)\big)^2}(\mathrm{d}P_d(t,T))^2.\end{aligned}$$

By substitution of $y(t)$, $P_d(t,T)$ and $P_f(t,T)$ from Equation (4.5) and Equation (4.6), we obtain:

$$\frac{\mathrm{d}y_F(t,T)}{y_F(t,T)} = \eta_d B_d(t,T)\Big(\rho_{y,d}\sigma(t) + \eta_d(t)B_d(t,T) - \rho_{d,f}\eta_f(t)B_f(t,T)\Big)\mathrm{d}t \tag{4.8}$$

$$+ \sigma(t)\mathrm{d}W_y^{\mathbb{Q}}(t) + \eta_d(t)B_d(t,T)W_d^{\mathbb{Q}} - \eta_f(t)B_f(t,T)\mathrm{d}W_f^{\mathbb{Q}}(t). \tag{4.9}$$

Since $y_F(t,T)$ is a martingale under the domestic T-forward measure and change of measures only effect the drift terms [4], we obtain:

$$\frac{\mathrm{d}y_F(t,T)}{y_F(t,T)} = \sigma(t)\mathrm{d}W_y^T(t) + \eta_d(t)B_d(t,T)W_d^T(t) - \eta_f(t)B_f(t,T)\mathrm{d}W_f^T(t), \tag{4.10}$$

where the T in the superscript represents the domestic T-forward measure. For the sum of three correlated Brownian motions W_1, W_2 and W_3 with correlation $\rho_{1,2}$, $\rho_{1,3}$ and $\rho_{2,3}$ we have the following [25]:

$$aW_1(t) + bW_2(t) + cW_3(t) \overset{\mathrm{d}}{=} \sqrt{a^2+b^2+c^2+2ab\rho_{1,2}+2ac\rho_{1,3}+2ab\rho_{2,3}}W_3(t), \tag{4.11}$$

where $\overset{\mathrm{d}}{=}$ indicates equality in distribution. Applying Equation (4.11) on equation (4.10) gives the following result:

$$\frac{\mathrm{d}y_F(t,T)}{y_F(t,T)} \overset{d}{=} \bar{\sigma}_F(t,T)\mathrm{d}W_{y_F}^T, \tag{4.12}$$

where $\bar{\sigma}_F(t,T)$ is given by

$$\bar{\sigma}_F(t,T) = \sqrt{\sigma^2 + \eta_d^2B_d^2 + \eta_f^2B_f^2 + 2\sigma\eta_dB_d\rho_{y,d} - 2\sigma\eta_fB_f\rho_{y,f} - 2\eta_dB_d\eta_fB_f\rho_{d,f}}. \tag{4.13}$$

For ease of notation, all time dependent parameters have the (t) omitted in Equation (4.13). Because $y(T) = y_F(t,T)$, we obtain the following when we price an FX option:

$$\begin{aligned}V^{FX}(t_0) &= \mathbb{E}^{\mathbb{Q}}\Big[\frac{M_f(T)}{M_d(T)}\max\big(\alpha(y(T)-K),0\big)\Big|\mathscr{F}(t_0)\Big]\\ &= \frac{P_d(t_0,T)}{P_f(t_0,T)}\mathbb{E}^T\big[\max\big(\alpha(y(T)-K),0\big)\big|\mathscr{F}(t_0)\big]\\ &= \frac{P_d(t_0,T)}{P_f(t_0,T)}\mathbb{E}^T\big[\max\big(\alpha(y_F(t,T)-K),0\big)\big|\mathscr{F}(t_0)\big].\end{aligned} \tag{4.14}$$

We can use Black76 formula to obtain the result for a call option.

Theorem 4.2.1. The value of a FX call option with strike K is given by:

$$V_c^{FX}(t_0) = \frac{P_d(t_0,T)}{P_f(t_0,T)}y_{F,t_0}\Phi(d_1) - K\frac{P_d(t_0,T)}{P_f(t_0,T)}\Phi(d_2), \tag{4.15}$$

where $y_F(t_0,T) := \frac{P_f(t,T)}{P_d(t,T)} y(t_0)$, $\Phi(\cdot)$ is the standard normal CDF and d_1, d_2 are given by:

$$d_1 = \frac{\log\left(\frac{y_F(t_0,T)}{K}\right) + \frac{1}{2}\sigma_c^2(T-t_0)}{\sigma_c\sqrt{T-t_0}}, \qquad d2 = \frac{\log\left(\frac{y_F(t_0,T)}{K}\right) - \frac{1}{2}\sigma_c^2(T-t_0)}{\sigma_c\sqrt{T-t_0}},$$

with σ_c is given by

$$\sigma_c^2 = \frac{1}{T-t_0}\int_{t_0}^{T} \bar{\sigma}_F^2(z)\mathrm{d}z.$$

Proof. By using Itô's lemma and integration we obtain for $y_F(t,T)$:

$$\begin{aligned}
\mathrm{d}\big(\log(y_F(t,T))\big) &= -\frac{1}{2}\bar{\sigma}_F^2(t,T)\mathrm{d}t + \bar{\sigma}_F(t,T)\mathrm{d}y_F(t),\\
\Rightarrow \log\Big(\frac{y_F(t,T)}{y_F(t_0,T)}\Big) &= -\frac{1}{2}\int_{t_0}^{t}\bar{\sigma}_F^2(s,T)\mathrm{d}s + \int_{t_0}^{t}\bar{\sigma}_F(s,T)\mathrm{d}W_F(s),\\
\Rightarrow y_F(t,T) &= y_F(t_0,T)e^{-\frac{1}{2}\int_{t_0}^{t}\bar{\sigma}_F^2(s,T)\mathrm{d}s + \int_{t_0}^{t}\bar{\sigma}_F(s,T)\mathrm{d}W_F(s)}.
\end{aligned}$$

Note that $\int_{t_0}^{t}\bar{\sigma}_F(s,T)\mathrm{d}W_F(s)$ is stochastic. $\int_{t_0}^{t}\bar{\sigma}_F(s,T)\mathrm{d}W_F(s)$ is normally distributed and has mean of zero since we take the integral with respect to a Brownian motion (the increments of Brownian motions are normally distributed with mean 0). We obtain for the variance:

$$\begin{aligned}
Var\Big(\int_{t_0}^{t}\bar{\sigma}_F(s,T)\mathrm{d}W_F(s)\Big) &= \mathbb{E}\Big[\Big(\int_{t_0}^{t}\bar{\sigma}_F(s,T)\mathrm{d}W_F(s)\Big)\Big)^2\Big] - \mathbb{E}\Big[\int_{t_0}^{t}\bar{\sigma}_F(s)\mathrm{d}W_F(s)\Big]^2\\
&= \mathbb{E}\Big[\Big(\int_{t_0}^{t}\bar{\sigma}_F(s)\mathrm{d}W_F(s)\Big)^2\Big] - 0^2.
\end{aligned}$$

By Itô isometry we obtain:

$$Var\Big(\int_{t_0}^{t}\bar{\sigma}_F(s,T)\mathrm{d}W_F(s)\Big) = \mathbb{E}\Big[\Big(\int_{t_0}^{t}\bar{\sigma}_F(s,T)\mathrm{d}W_F(s)\Big)\Big)^2\Big] = \int_{t_0}^{t}\bar{\sigma}_F^2(s,T)\mathrm{d}s.$$

$\int_{t_0}^{t}\bar{\sigma}_F(s)\mathrm{d}W_F(s)$ is normally distributed with mean 0 and variance as given above. Therefore, we have $\int_{t_0}^{t}\bar{\sigma}_F(s)\mathrm{d}W_F(s) \overset{\mathrm{d}}{=} Z\sqrt{\int_{t_0}^{t}\bar{\sigma}_F^2(s)\mathrm{d}s}$, where Z is standard normal. We obtain for $y_F(t,T)$:

$$y_F(t_0,T) \overset{\mathrm{d}}{=} y_F(t_0,T)e^{-\frac{1}{2}\int_{t_0}^{t}\bar{\sigma}_F^2(t)\mathrm{d}t + Z\sqrt{\int_{t_0}^{t}\bar{\sigma}_F^2(s)\mathrm{d}s}}. \tag{4.16}$$

By substitution of Equation (4.16), we can obtain a pricing equation that is independent of both the domestic and foreign interest rate. We will prove for the call option. The put option can be proved in a similar manner. By Equation (4.14) we have that:

$$\begin{aligned}
V_c^{fx}(t_0,S) &= \frac{P_d(t_0,T)}{P_f(t_0,T)}\mathbb{E}^T[\max(y(T)-K,0)|\mathscr{F}(t_0)]\\
&= \frac{P_d(t_0,T)}{P_f(t_0,T)}\mathbb{E}^T[\max(y_F(T,T)-K,0)|\mathscr{F}(t_0)]\\
&= \frac{P_d(t_0,T)}{P_f(t_0,T)}\Big(\mathbb{E}^T[y_F(T,T)1_{\{y_F(T,T)>K\}}|\mathscr{F}(t_0)] - \mathbb{E}^T[K1_{\{y_F(T,T)>K\}}|\mathscr{F}(t_0)]\Big).
\end{aligned}$$

Let Z be standard normal distributed. For the subscript of the indicator functions holds:

$$\begin{aligned}
y_F(t_0,T) > K &\Longleftrightarrow y_F(t_0,T)e^{-\frac{1}{2}\int_{t_0}^{t}\bar{\sigma}_F^2(s)\mathrm{d}s + Z\sqrt{\int_{t_0}^{T}\bar{\sigma}_F^2(s)\mathrm{d}s}} > K\\
&\Longleftrightarrow -\frac{1}{2}\int_{t_0}^{T}\bar{\sigma}_F^2(t)\mathrm{d}t + Z\sqrt{\int_{t_0}^{T}\bar{\sigma}_F^2(s)\mathrm{d}s} > \log\Big(\frac{K}{y_F(t_0,T)}\Big)\\
&\Longleftrightarrow Z \le \frac{\log(\frac{y_F(T,T)}{K}) - \frac{1}{2}\int_{t_0}^{T}\bar{\sigma}_F^2(t)\mathrm{d}t}{\sqrt{\int_{t_0}^{T}\bar{\sigma}_F^2(s)\mathrm{d}s}} =: d_2,
\end{aligned}$$

where we used in the last step that $Z \overset{\mathrm{d}}{=} -Z$, since Z is symmetric with mean zero. We obtain for $V_c^{fx}(t_0,S)$:

$$\begin{aligned}
V_c^{fx}(t_0,S) &= \frac{P_d(t_0,T)}{P_f(t_0,T)}\Big(\mathbb{E}^T[y_F(t_0,T)e^{-\frac{1}{2}\int_{t_0}^T \bar{\sigma}_F^2(s)\mathrm{d}s + Z\sqrt{\int_{t_0}^T \bar{\sigma}_F^2(s)\mathrm{d}s}}1_{\{Z\leq d_2\}}] - \mathbb{E}^T[K1_{\{Z\leq d_2\}}]\Big) \\
&= \frac{P_d(t_0,T)}{P_f(t_0,T)} y_F(t_0,T)\frac{1}{2\pi}\int_{d_2}^{\infty} e^{-\frac{1}{2}\left(z+\sqrt{\int_{t_0}^T \bar{\sigma}_F^2(s)\mathrm{d}s}\right)^2} dz - \frac{P_d(t_0,T)}{P_f(t_0,T)} K\Phi(d_2) \\
&= \frac{P_d(t_0,T)}{P_f(t_0,T)} y_F(t_0,T)\Phi\left(d_2 + \sqrt{\int_{t_0}^T \bar{\sigma}_F^2(s)\mathrm{d}s}\right) - \frac{P_d(t_0,T)}{P_f(t_0,T)} K\Phi(d_2),
\end{aligned}$$

with $\Phi(\cdot)$ is the standard normal CDF. We define $d_1 := d_2 + \sqrt{\int_{t_0}^T \bar{\sigma}_F^2(s)\mathrm{d}s}$, and obtain the following solution:

$$V_c^{fx}(t_0,T) = \frac{P_d(t_0,T)}{P_f(t_0,T)} y_F(t_0,T)\Phi(d_1) - \frac{P_d(t_0,T)}{P_f(t_0,T)} K\Phi(d_2).$$

We can rewrite this to the form of Black76 formula:

$$V_c^{fx}(t_0,S) = y_F(t_0,T)P(t_0,T)\Phi(d_1) - KP(t_0,T)\Phi(d_2), \tag{4.17}$$

with

$$d_1 = \frac{\log\left(\frac{y_F(t_0,T)}{K}\right) + \frac{1}{2}\sigma_c^2(T-t_0)}{\sigma_c\sqrt{T-t_0}}, \qquad d2 = \frac{\log\left(\frac{y_F(t_0,T)}{K}\right) - \frac{1}{2}\sigma_c^2(T-t_0)}{\sigma_c\sqrt{T-t_0}},$$

and σ_c is given by:

$$\sigma_c^2 = \frac{1}{T-t_0}\int_{t_0}^T \bar{\sigma}_F^2(z)dz.$$

□

4.3. Monte Carlo pricing for Black-Scholes Hull-White

In order to verify that the exact solution of an FX option under Black-Scholes Hull-White is correct, and also correctly implemented, one can compute value by means of Monte Carlo [25]. We can compute the value of an FX option by Monte Carlo is given by the following:

$$V^{FX}(t_0) = \mathbb{E}^{\mathbb{Q}}\Big[\frac{M_f(T)}{M_d(T)}\max\big(\alpha(y(T)-K),0\big)\Big|\mathscr{F}(t_0)\Big], \tag{4.18}$$

where $\alpha = 1$ for a call and $\alpha = -1$ for a put option. When one wants to use pricing formula in Equation (4.18), one needs to generate paths for $r_d(t)$, $r_f(t)$ and $y(t)$. In order to use the Monte Carlo for validation purposes, the Choleski decomposition [29] is used to obtain independent Browian Motions. Let $\mathbf{W}(t) = \left[W_d^{\mathbb{Q}}(t), W_f^{\mathbb{Q}}(t), W_y^{\mathbb{Q}}(t)\right]$, one obtains

$$\mathbf{W}(t) = \begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix}\tilde{\mathbf{W}}(t),$$

where $\tilde{W}_d$, $\tilde{W}_f$ and $\tilde{W}_y$ are independent normal distributed with mean zero and variance dt. The following relation holds with the correlation matrix.

$$\mathrm{d}\mathbf{W}(t)\big(\mathrm{d}\mathbf{W}(t)\big)^T = \begin{pmatrix} 1 & \rho_{d,f} & \rho_{y,d} \\ \rho_{d,f} & 1 & \rho_{y,f} \\ \rho_{y,d} & \rho_{y,f} & 1 \end{pmatrix} = \begin{pmatrix} a & 0 & 0 \\ b & c & 0 \\ d & e & f \end{pmatrix}\begin{pmatrix} a & b & d \\ 0 & c & e \\ 0 & 0 & f \end{pmatrix}.$$

Solving the system of equations gives us the following values

$$a = 1, \quad b = \rho_{d,f}, \quad c = \sqrt{1-\rho_{d,f}^2}, \quad d = \rho_{y,d}, \quad e = \frac{\rho_{y,f} - b\cdot d}{c}, \quad f = \sqrt{1-d^2-e^2}.$$

Therefore, we obtain the following composition of independent Brownian motions $\tilde{W}_d$, $\tilde{W}_f$ and $\tilde{W}_y$

$$\begin{aligned}
W_d &= \tilde{W}_d, \\
W_f &= \rho_{d,f}\tilde{W}_d + \sqrt{1-\rho_{d,f}^2}\tilde{W}_f, \\
W_y &= \rho_{y,d}\tilde{W}_d + \frac{\rho_{y,f}-\rho_{d,f}\rho_{y,d}}{\sqrt{1-\rho_{d,f}^2}}\tilde{W}_f + \sqrt{1-\rho_{y,d}^2-\left(\frac{\rho_{y,f}-\rho_{d,f}\rho_{y,d}}{\sqrt{1-\rho_{d,f}^2}}\right)^2}\tilde{W}_y.
\end{aligned}$$

Using this composition, one is able to perform Monte Carlo pricing. Monte Carlo pricing is relatively slow. One has to generate a number of paths, which is time expensive. For calibration the computation time is important. Therefore, Monte Carlo will not be used in the calibration procedure in this thesis. We merely use Monte Carlo for validation of our calibration.

4.4. Calibration of FX Black-Scholes Hull-White model

The dynamics of the Black-Scholes Hull-White model are given by:

$$\begin{cases}
\mathrm{d}y(t) = (r_d(t) - r_f(t))y(t)\mathrm{d}t + \boxed{\sigma(t)}\, y(t)\mathrm{d}W_d^{\mathbb{Q}}(t), \\
\mathrm{d}r_d(t) = \boxed{\lambda_d(t)}\big(\theta_d(t) - r_d(t)\big)\mathrm{d}t + \boxed{\eta_d(t)}\,\mathrm{d}W_d^{\mathbb{Q}}(t), \\
\mathrm{d}r_f(t) = \boxed{\lambda_f(t)}\big(\theta_f(t) - r_f(t) - \boxed{\eta_f(t)}\,\rho_{y,f}\,\boxed{\sigma(t)}\big)\mathrm{d}t + \boxed{\eta_f(t)}\,\mathrm{d}W_f^{\mathbb{Q}}.
\end{cases} \tag{4.19}$$

In Equation (4.19) the parameters that we want to calibrate are highlighted. We have five time-dependent parameters we want to calibrate: $\lambda_d(t)$, $\eta_d(t)$, $\lambda_f(t)$, $\eta_f(t)$ and $\sigma(t)$. Just as for the calibration as for the Hull-White model in Section 3.2, we model $\lambda_d(t)$, $\eta_d(t)$, $\lambda_f(t)$, $\eta_f(t)$ and $\sigma(t)$ as piece-wise constant parameters. We use the following calibration procedure:

1. For the calibration of $\lambda_d(t)$ and $\eta_d(t)$ we use swaptions on the domestic currency. First, we calibrate $\lambda_d(t)$, afterwards calibrating $\eta_d(t)$ by the method described in Section 3.4.
2. For the calibration of $\lambda_f(t)$ and $\eta_f(t)$ we use swaptions on the foreign currency. First, we calibrate $\lambda_f(t)$, afterwards calibrating $\eta_f(t)$ to the method described in Section 3.4.
3. At last we calibrate $\sigma(t)$ to FX options. In order to do this we use Theorem 4.2.1 with the calibrated $\lambda_d(t)$, $\eta_d(t)$, $\lambda_f(t)$ and $\eta_f(t)$ from the previous steps. We have the following optimization problem:

$$\underset{\sigma(t)}{\operatorname{argmin}}\left(\sqrt{\sum_i\left(\sigma_{FX}^{mrkt}(t_0,T_i,K) - \sigma_{FX}^{BSHW}(t_0,T_i,K,\lambda_d(t),\eta_d(t),\lambda_f(t),\eta_f(t),\rho_{y,d},\rho_{y,f},\rho_{d,f},\sigma(t))\right)^2}\right). \tag{4.20}$$

We can obtain the correlations $\rho_{y,d},\rho_{y,f},\rho_{d,f}$ from the market. Strike price K is chosen ATM for the calibration by the same reasoning as in Section 3.5: Atm strike prices are liquid and for deep ITM and OTM vega is close to zero. In Equation (4.20) $\sigma_{FX}^{BSHW}(\cdot)$ represents the implied volatility of the price given by Theorem 4.2.1. This optimization problem is a non-linear non-convex optimization problem. Therefore, by the same line of reasoning as in Section 3.6, we use trust region methods to solve this optimization problem. As there is no independence between the implied volatilities of the FX options, sequential calibration cannot be used.

4.4.1. Artificial data

By using artificial data to test our calibration method, we can pick a FX volatility parameter $\sigma(t)$, and find out whether our calibration methods returns a similar calibrated $\sigma(t)$. By using market data, we never know what the "real" underlying $\sigma(t)$ is. Therefore, it is harder to measure the performance of our calibration performance. We believe testing the calibration to artificial data is indispensable. Still one can argue that using artificial data is not realistic. How can one know that the chosen artificial data resembles data obtained from the market? We try to generate realistic artificial data by first calibrating $\lambda_d(t)$, $\lambda_f(t)$, $\eta_d(t)$ and $\eta_f(t)$ to market data. Afterwards, we choose a $\sigma(t)$ such that the obtained implied volatilities of ATM FX options are comparable to the implied volatilities of FX options in the market, i.e. approximately 8.5%. By generating data in this way, we argue that the generated data is comparable to market data to some extent.

The calibration for artificial data has been performed by the following steps:

1. Calibrate $\lambda_d(t)$ and $\eta_d(t)$ to market data as in the Section 3.4.
2. Calibrate $\lambda_f(t)$ and $\eta_f(t)$ to market data as in the Section 3.4.
3. Choose $\sigma(t)$ such that the implied volatilities for ATM FX options are comparable to those in the market.
4. calibrate $\sigma(t)$ to the artificial data.

In this research we investigate the effect of the (chosen) mean reversion parameters on the calibration of FX volatility parameter $\sigma(t)$. In order to do investigate this, we pick mean reversion parameters, besides calibrating mean reversion parameters. Afterwards we calibrate the IR volatility parameters $\eta(t)$. This results in slight adjustments in step 1 and 2, as we are not calibrating the $\lambda_d(t)$ and $\lambda_f(t)$. Results of this research can found in Figure 4.1.

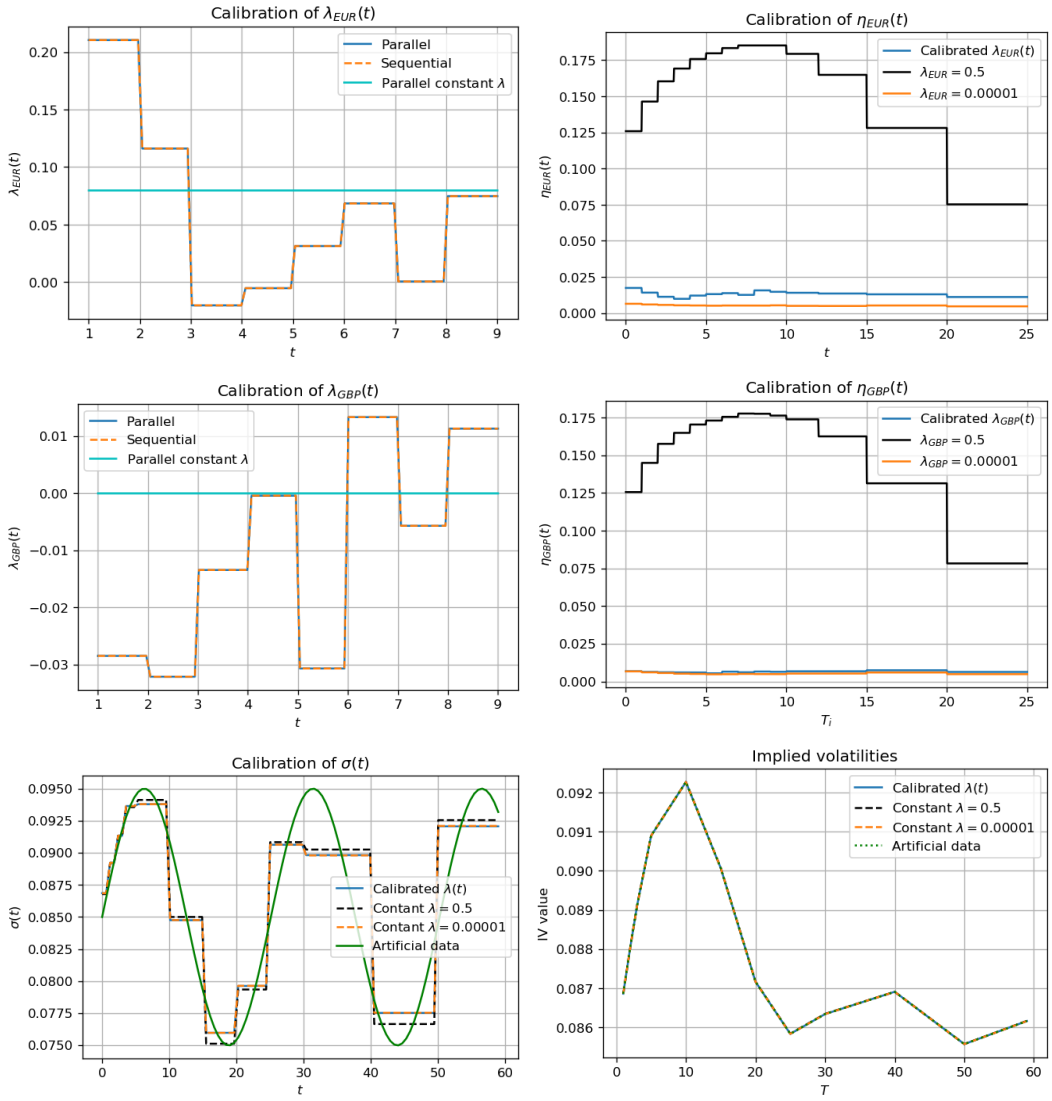

Figure 4.1: The calibration of the Black-Scholes Hull-White model to artificial data is shown. Calibration of $\lambda_{EUR}(t)$ and $\lambda_{GBP}(t)$ as time-dependent has been used in the blue line for the calibration of $\eta_{EUR}(t)$ and $\eta_{GBP}(t)$. For the calibration of $\eta_{EUR}(t)$ and $\eta_{GBP}(t)$ we used several λ, creating pairs of calibrated $\lambda(t)$ and $\eta(t)$. We calibrated $\sigma(t)$ for these different pairs. The results can be found in the two pictures on the bottom. The domestic currency is the British pound sterling (GBP) and foreign is the euro (EUR). Market data has been retrieved from Bloomberg on 16th July 2020.

Figure 4.1 shows us that the impact of different mean reversion parameters of the Hull-White processes for calibration of $\sigma(t)$ is very minimal in this experiment. Besides, we find that the implied volatilities obtained for the calibration of $\sigma(t)$ fit the artificial data. Moreover, the obtained piece-wise constant functions fit the artificial $\sigma(t)$ well.

4.4.2. Market data

In the previous section we looked into the performance of the calibration on artificial data. In this section we will look into the calibration to market data. The steps 1 to 3 described in Section 4.4 are used for the calibration, which can also be found in Algorithm 5 in Appendix A. Just as in the calibration of artificial data, we would like to investigate the effect of the mean reversion parameters. Besides calibrating the mean reversion parameter, we choose mean reversion parameters. $\eta(t)$ is calibrated accordingly to the calibrated or chosen mean reversion parameter. For this calibration an alternative calibration from the industry is available. In this alternative calibration from the industry we observe a typical problem for calibration; Bad quality of fit for late maturities. Results of this research can be found in Figure 4.2. More results on the calibration of the Black-Sholes Hull-White model with different data sets can be found in the Appendix C.1 and Appendix C.2.

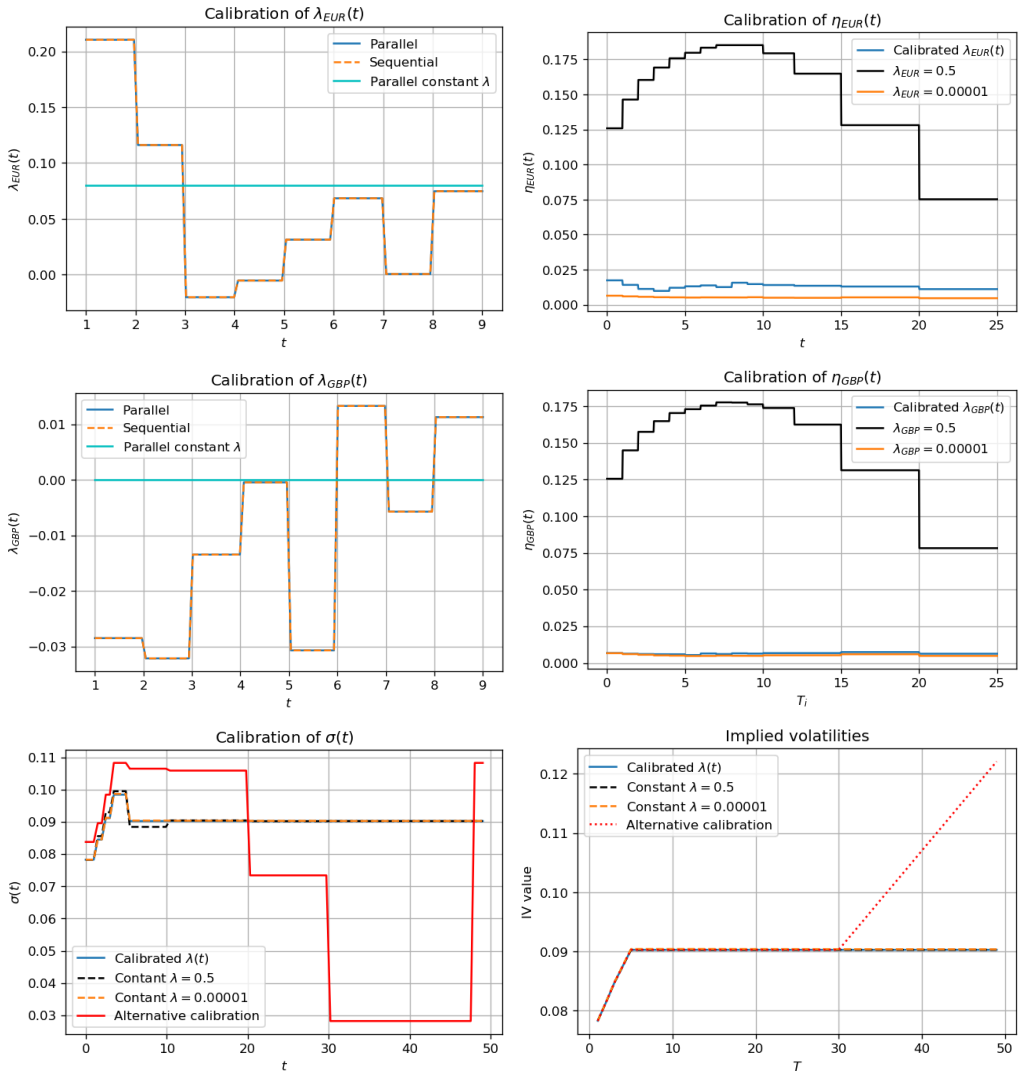

Figure 4.2: The calibration of the Black-Scholes Hull-White model to market data is shown. Calibration of $\lambda_{EUR}(t)$ and $\lambda_{GBP}(t)$ as time-dependent has been used for the calibration of $\eta_{EUR}(t)$ and $\eta_{GBP}(t)$. For the calibration of $\eta_{EUR}(t)$ and $\eta_{GBP}(t)$ we used several λ, creating pairs of calibrated $\lambda(t)$ and $\eta(t)$. We calibrated $\sigma(t)$ for these different pairs. The results can be found in the two pictures on the bottom. The domestic currency is the British pound sterling (GBP) and foreign is the euro (EUR). Market data has been retrieved from Bloomberg on 16th July 2020.

The results in Figure 4.2 show that the calibration of FX volatility parameter $\sigma(t)$ has a good quality of fit on the implied volatilities of FX options for all mean reversion parameters. For the calibrations in Appendix C.1 and Appendix C.2 we observe similar results for the calibration. The mean reversion parameter seems to have minimal influence on the calibration of $\sigma(t)$. Besides, the results suggest that the problem of bad quality of fit for late maturities can be solved.

5

Hedging

As a trader hedging your portfolio is one of the most important tasks for the mitigation of risks. Traders hedge their portfolio once they have done one or more trades in order to make the portfolio less dependent to external forces, e.g. movements in the market. The frequency of hedging is dependent on the hedging costs and the risk a trader is exposed to. The profit and loss (P&L) keeps track of the total earnings/losses of the specific trader. These risks to external forces are known as the Greeks. One of the Greeks is the Δ, which indicates the change of value of the option when there is movement in the underlying asset. This can be explained by the replicating portfolio concept in which one replicates the portfolio in the following way [25]:

$$\Pi(t,S) = V(t,S) - \Delta S.$$

We would like to delta hedge our portfolio, i.e. the value of our portfolio should not change when the underlying asset S moves. Therefore, we obtain the following expression for Δ:

$$\frac{\mathrm{d}\Pi(t,S)}{\mathrm{d}S} = \frac{\mathrm{d}V(t,S)}{\mathrm{d}S} - \Delta = 0 \Rightarrow \Delta = \frac{\mathrm{d}V(t,S)}{\mathrm{d}S}.$$

This implies that we sell Δ of the underlying asset in order to hedge our portfolio. In our case we are using swaptions for calibration purposes. Swaptions have swaps as underlying assets. Therefore, one should sell Δ of the underlying swaps when buying a swaption, where Δ corresponds to the delta of the specific swaption. Besides movement in the underlying asset, there are more external forces that can affect the value of the option. Just as delta, the movement of the option with respect to a specific external force are denoted as greek letters. Some of the other greeks are the following:

- Vega: $= \frac{\partial V}{\partial \sigma}$,
- Theta: $\Theta = -\frac{\partial V}{\partial T}$,
- Gamma: $\Gamma = \frac{\partial \Delta}{\partial S} = \frac{\partial^2 V}{\partial S^2}$.

For this thesis, we will focus on Δ hedging. We would like to know what is the impact on the hedging when we consider time-dependent parameters instead of constant parameters. We will look at delta hedging in the Black-Scholes equity world, where the stock does not pay dividend. The volatility parameter in the equity Black-Scholes model is equivalent to the volatility parameter of the FX Black-Scholes Hull-White model with the stochastic quantity $r_f - r_d(t) \equiv r(t)$. For the Hull-White process the same holds with $\lambda(t)\big(\theta(t) - r_{HW}(t)\big) \equiv r(t)$. Therefore, we argue that the findings in this chapter on the volatility parameter from the equity Black-Scholes are comparable for the volatility parameter of the $\eta_d(t)$, $\eta_f(t)$ and $\sigma(t)$ from the FX Black-Scholes Hull-White model.

Our goal in this chapter is researching how piece-wise constant volatility parameter affects the delta hedging process. Especially, how big jumps in the volatility parameter affect the delta hedging process. We will refer to the amount of big jumps in piece-wise constant parameters as the homogeneity of the parameter. The bigger the jumps in the piece-wise constant parameter, the worse the homogeneity.

5.1. Delta hedging with constant parameters

In the Black-Scholes case with r and σ constant the value of a call option is given by:

$$V_{Call}(t,S) = \Phi(d_1)S(t) + \mathrm{e}^{r(T-t)}K\Phi(d_2), \tag{5.1}$$

where Φ is the CDF of a standard normal, T is maturity and K is the strike price. d_1 and d_2 are given by:

$$d_1(t,S) = \frac{\frac{\log(S_t)}{K}(r+\frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}}, \qquad d_2(t,S) = d_1 - \sigma\sqrt{T-t}. \tag{5.2}$$

In this Black Scholes world, the $\Delta(t,S)$ is given by

$$\Delta(t,S) = \frac{\partial V_{Call}(t,S)}{\partial S(t)} = \Phi(d_1(t,S)). \tag{5.3}$$

We can perform theoretically delta hedging to research how time dependent parameters influence the hedging process, and especially how large fluctuations in value of the time-dependent parameter influence the delta hedging process. In order to perform theoretical, delta hedging we create a grid of m points and maturity T [25]. Every point in time can be expressed as $t_i = \frac{i}{m}T$. For the delta hedging of a call option $V_{call}(t,S)$ on a stock S with maturity T and strike K, the following initial computation of the P&L (t) is done:

$$\text{P\&L}(t_0) = V_{call}\big(t_0,S\big) - \Delta(t_0,S)S(t_0). \tag{5.4}$$

Afterwards, we compute the $P\&L(t)$ at every point in time in an iterative way:

$$\text{P\&L}(t_{i+1}) = \text{P\&L}(t_i)e^{r(t_{i+1}-t_i)} - \big(\Delta(t_{i+1},S) - \Delta(t_i,S)\big)S(t_{i+1}). \tag{5.5}$$

At maturity time $(t=T)$, you will exercise the call if it is in the money and re-balance your portfolio by making sure your position in the stock (your Δ) is zero. In Figure 5.1, one can see how the delta of the option and $P\&L$ changes over time. For an in the money call option one can observe a big jump in the $P\&L$ at the end due to the settling of the call option contract.

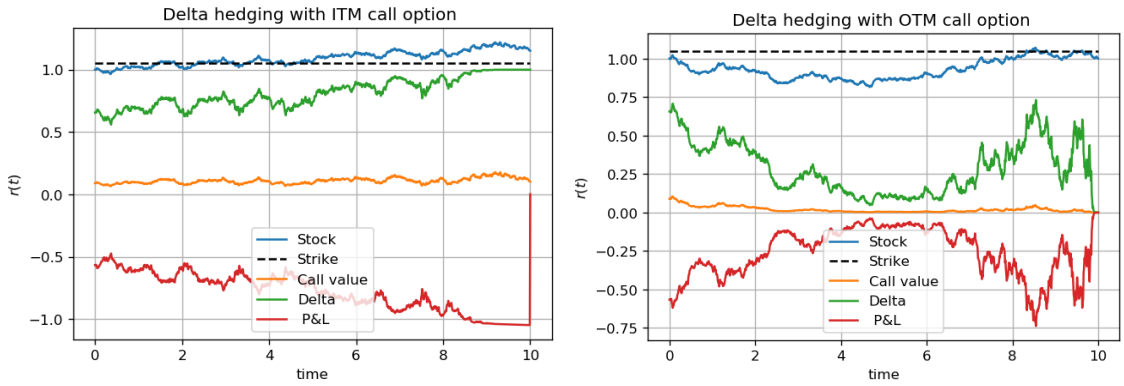

Figure 5.1: Example of how delta hedging for an in the money option and out of the money option. The plot on the left shows an in the money call option and the plot on the right an out of the money call option. Following parameters are chosen; $r = 0.01$, $\sigma = 0.05$, $T = 10$, $S_0 = 1.0$ and $K = 1.05$. In total 1000 trades for hedging are performed. Hedges are performed with equidistances between t_0 and T. This figure is similar to an example in [25].

A trader prefers to mitigate the risks as cheap as possible and with as few risks as possible. Therefore, for delta-hedging to be successful, we want the P&L to be (very close to) 0 at the maturity of the option. In other words, at time t_0 we want the expectation of P&L (T) to be zero and the variance of P&L (T) to be very small. The expectation of P&L (T) is always (very close to) zero. We measure the performance of our delta hedging process by the variance of P&L (T). The smaller the variance, the better. Parameters that influence the performance of the P&L (T) are the volatility of the underlying asset and the number of hedges used for delta hedging. In Figure 5.2 one can see that number of hedges and the chosen constant volatility have an impact on the performance of delta hedging.

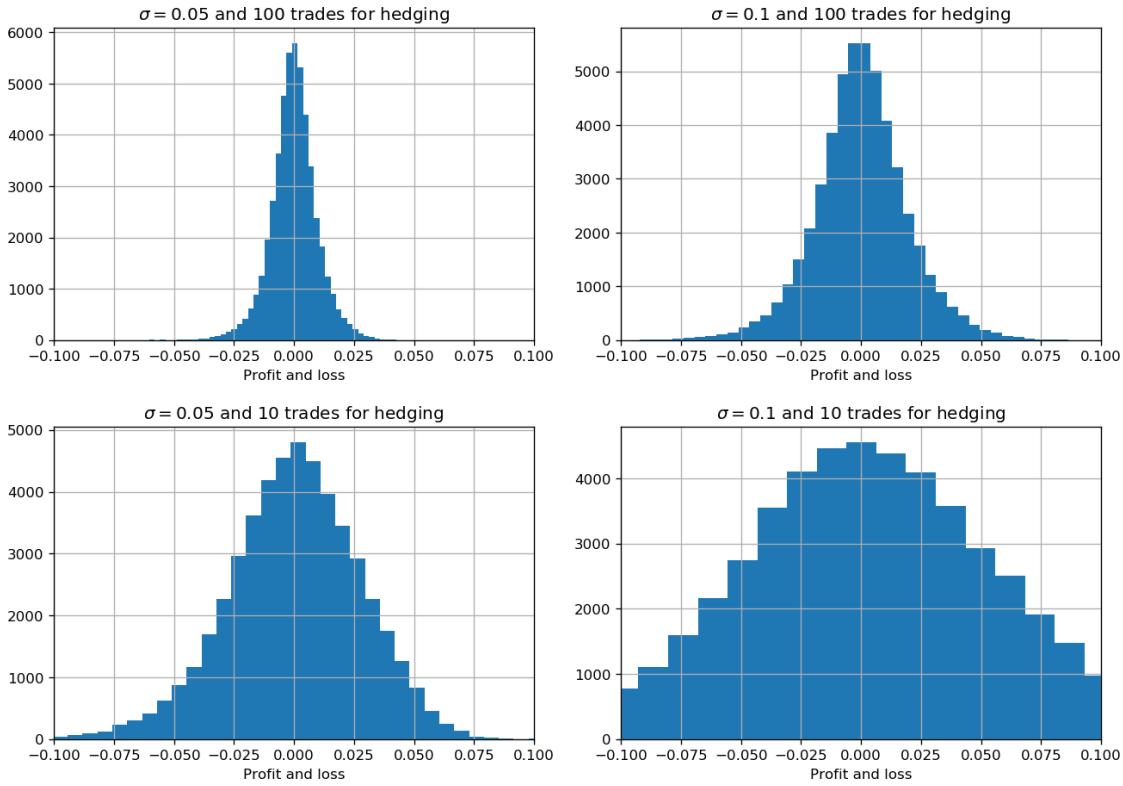

Figure 5.2: 50000 delta hedging simulations are performed with constant $r = 0.01$, $T = 10$, $S_0 = 1.0$ and $K = 1.05$. The volatility parameter σ and the number of trades for hedging are varied. One can see that a higher σ results in worse hedging performance, also a lower number of trades for hedging results in worse hedging results. Hedges are performed with equidistances between t_0 and T.

In our theoretically hedging scenario we can conclude from Figure 5.2 that a large amount of trades for hedging would be preferable. we would like to re-balance our portfolio as frequent as possible. This is not the case in reality though. There are two important issues in the real world that are not taken into account in our theoretical simulation:

- In reality there are transaction costs. Those are not taken into account in this hedging setting. Therefore, performing a trade for hedging costs a little bit of money.

- Besides the transaction costs, there is a bid-ask spread around the market price of the stock. Therefore, one is not able to buy or sell at the market price of the stock, but one has to pay a (little) premium in order to buy or sell at that moment.

These two facts show us why we do not want to re-balance our portfolio as frequent as possible. The first point shows that a trader has to make a decision on the tradeoff between the risk of his/her position in Δ and the transaction cost. The second point indicates that a trader could decide to wait for a favorable point in time to hedge his/her delta when the bid-ask spread is low.

5.2. Delta hedging with time dependent parameters

In the introduction of the this section we stated that the goal of this section is looking into the homogeneity of time-dependent parameters. In order to do this, we will from now on assume $r(t)$ and $\sigma(t)$ to be time-dependent parameters. Using time-dependent parameters will change the value of the option and the corresponding delta shown in the previous section. For time-dependent $r(t)$ and $\sigma(t)$ the value of a call option is given by:

$$V_{Call}(t,S) = \Phi(d_1)S_t + \mathrm{e}^{-\int_t^T r(t)\mathrm{d}t} K\Phi(d_2), \tag{5.6}$$

where Φ is the CDF of a standard normal, T is maturity and K is the strike price. d_1 and d_2 are given by:

$$d_1(t,S) = \frac{\frac{\log(S_t)}{K} \int_t^T \left(r(t) + \frac{1}{2}\sigma^2(t)\right) \mathrm{d}t}{\sqrt{\int_t^T \sigma^2(t)\mathrm{d}t}}, \qquad d_2(t,S) = d_1(t,S) - \sqrt{\int_t^T \sigma^2(t)\mathrm{d}t}. \tag{5.7}$$

$\Delta(t,S)$ is given by:

$$\Delta(t,S) = \frac{\partial V_{Call}(t,S)}{\partial S(t)} = \Phi\big(d_1(t,S)\big), \tag{5.8}$$

The iterative method for delta-hedging from the previous section remains the same under time-dependent parameters. We create a grid and use Equation (5.4) and Equation (5.5). In the previous section we showed that the volatility has an impact on the performance of the delta hedging process. When we compare delta hedging with a constant σ to a time dependent $\sigma(t)$, we make sure that the implied volatility from t_0 to T is the same. In that way, the total volatility from t_0 to T for time-dependent $\sigma(t)$ is equal to that of σ. When $\sigma(t)$ is known, the implied volatility from t_0 till T is given by:

$$\sigma_{imp} = \sqrt{\frac{1}{T}\int_{t_0}^T \sigma^2(t)\mathrm{d}t}. \tag{5.9}$$

Notice that if we would hedge continuously, our hedge will be perfect in our modelling. This implies that if $\Delta t \to 0$ we will obtain an expectation of P&L(T) of zero and a variance of P&L(T) of zero. In this case, arbitrary $\sigma(t)$ with equal implied volatility have the same delta hedging performance. In reality however, hedging cannot be performed at all time due to transaction costs. If we perform trades for hedges only a limited amount of times, Example 5.2.1 shows us that variance of profit and loss can be different for various $\sigma(t)$ with the same implied volatility from t_0 till T.

Example 5.2.1. The goal of this example is to illustrate why a big jump in volatility at one time point will cause more variance in the delta hedging process. Let interest rate $r = 0$, $S_0 = 1.0$ and strike price is at the money $K = 1.0$. We consider call options with maturity time $T = 2$. We will look at $\sigma_1(t)$ and $\sigma_2(t)$, where $\sigma_1(t)$ will have a big jump in volatility, and $\sigma_2(t)$ will be constant. $\sigma_1(t)$ is defined as

$$\sigma_1(t) = \begin{cases} 0.5 & \text{if } t < 1, \\ 0.05 & \text{if } t \geq 1. \end{cases} \tag{5.10}$$

From $\sigma_1(t)$ we obtain an implied volatility of approximately 0.355 from $t_0 = 0$ till $T = 2$. We define $\sigma_2(t)$ as

$$\sigma_2(t) = \sqrt{\frac{1}{2}\int_0^2 \sigma_1^2(t)\mathrm{d}t} \approx 0.355. \tag{5.11}$$

We will write a call option and delta hedge at one time point $t = 1$. We consider the scenario where $S(T) = 2.0$. Paths of S under this scenario can be observed in Figure 5.3.

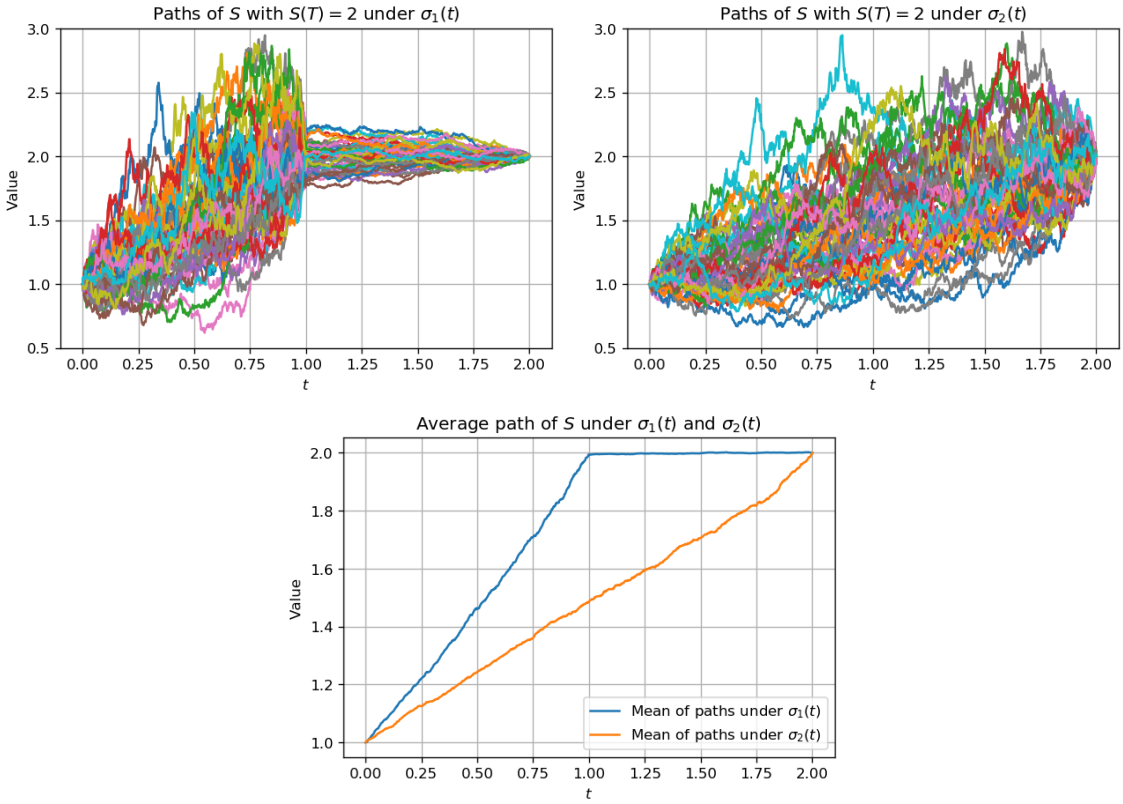

Figure 5.3: In the upper two pictures, paths under the scenario that $S(T) = 2.0$ are shown for $\sigma_1(t)$ and $\sigma_2(t)$. In the picture on the bottom we can find the mean of the paths for σ_1 and σ_2. Chosen parameters: $t_0 = 0$, $T = 2$, $S_0 = 0$ and $r = 0$.

Since the implied volatility $\sigma_1(t)$ and $\sigma_2(t)$ are the same value, the mean and variance of $S(T)$ are the same under both σ_1 and σ_2. Therefore, the probability of reaching 2.0 at $T = 2$ is equal for both processes. Assuming S reaches 2.0 at T, a stock path under σ_1 has a big jump in value from $t = 0$ to $t = 1$ and a small jump from $t = 1$ to $t = 2$, whilst for $\sigma_2(t)$ this is more balanced. Notice that if we delta hedge only once at $t = 1$, the path from $S(t_0)$ till $S(1)$ is irrelevant, but the value of the stock at $S(1)$ is relevant. In this example, we show that losses under delta hedging are more extreme under σ_1 compared to σ_2. We pick the following paths comparable to the means in Figure 5.3; $S_{\sigma_1}(1) = 1.9$ and $S_{\sigma_2}(1) = 1.5$. A schematic picture can be observed in Figure 5.4, where the performance of delta hedging at $t = 1$ is shown.

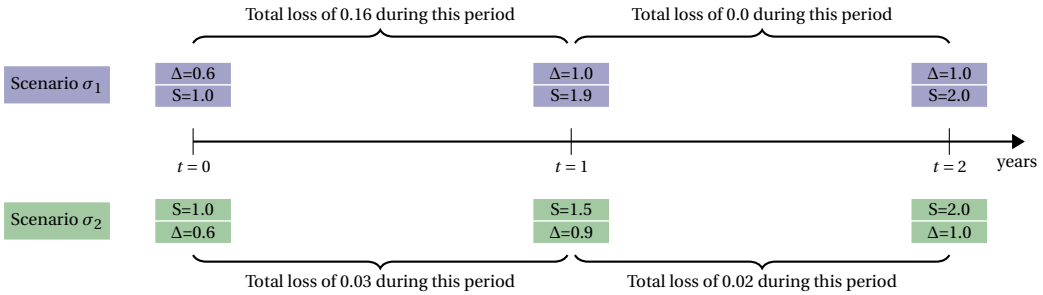

Figure 5.4: Two paths are sketched for different volatilities, where both scenarios have $S(T = 2) = 2.0$. The upper scenario in blue under $\sigma_1(t)$ and the bottom in green under $\sigma_2(t)$. Delta hedging is performed at $t = 1$. We can observe that the upper scenario in blue results in a higher loss than the scenario in green. Interest rate is considered zero in this example. $\Delta(t)$ is computed under Black Scholes by Equation (5.8).

Figure 5.4 shows that the path under $\sigma_1(t)$ results in a bigger loss than the path under $\sigma_2(t)$. Total loss under $\sigma_1(t)$ is 0.16, which all occur in the period $t = 0$ till $t = 1$. Total loss under $\sigma_2(t)$ 0.05, that is split over both periods.

Example 5.2.1 shows us a situation where paths under $\sigma_1(t)$ result in a bigger loss than paths under $\sigma_2(t)$ by assuming $S(T)$ is far in the money. The same can be shown for far out of the money values of $S(T)$; paths under $\sigma_1(t)$ will results in a larger profit than paths under $\sigma_2(t)$. In general, the variance of the profit and loss is higher for paths under $\sigma_1(t)$ than for paths under $\sigma_2(t)$.

Now that we know that a time-dependent $\sigma(t)$ can have an impact on the delta hedging process, we would like to look into comparable parameters as we obtain from real market data. Calibration of market data retrieved from Bloomberg on 21st of July 2020 shows us that the volatility parameter of FX options in the Black-Scholes Hull-White model swings around 0.10.

Example 5.2.2. We define the following $\sigma(t)$,

$$\sigma(t) = \begin{cases} 0.15 & \text{if } t < 1, \\ 0.05 & \text{if } t \geq 1, \end{cases} \tag{5.12}$$

where $\sigma(t) = 0.15$ for $0 \leq t < 1$ and $\sigma(t) = 0.05$ for $t \geq 1$, gives an implied volatility of $\sigma_{imp} \approx 0.067$. Using this time dependent $\sigma(t)$ and the constant σ_{imp} for delta hedging gives the result in Figure 5.5.

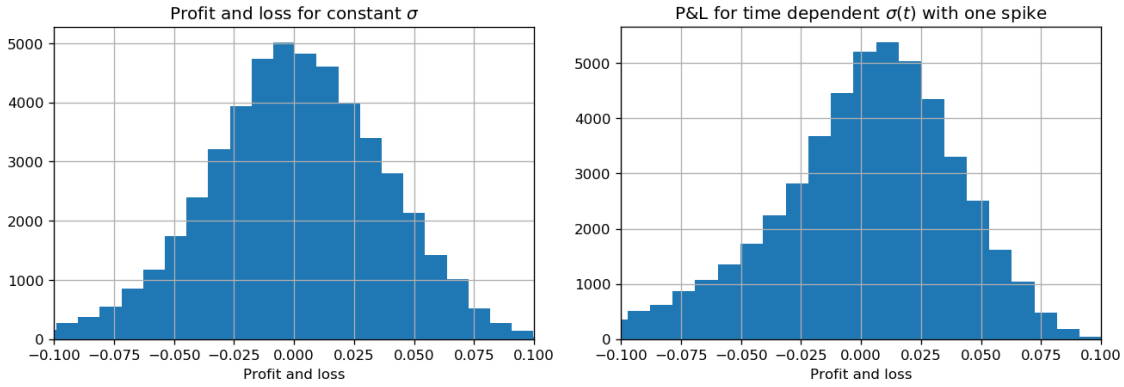

Figure 5.5: 50000 delta hedging simulations are performed with constant $r = 0.01$. Hedges are performed with equidistances between t_0 and T. Time dependent sigma is defined in Equation (5.12). The standard deviation of the histogram on the left is 0.038 and of the histogram on the right 0.042. In both cases the mean is close to 0, $2 \cdot 10^{-5}$ and $-3 \cdot 10^{-4}$ respectively.

Looking at the results in Figure 5.5, we find that the standard deviation is approximately 10% higher when we use the time-dependent volatility function in Equation (5.12) compared to the constant implied volatility function. One can also notice that the distribution seems to change slightly; the tail on the left becomes bigger and the top of the histogram moves to the right.

Example 5.2.3. In this example we show that concentrating the hedges in more volatile periods improves the variance of the profit and loss. We define the following $\sigma(t)$,

$$\sigma(t) = \begin{cases} 0.5 & \text{if } t < 1, \\ 0.05 & \text{if } t \geq 1. \end{cases} \tag{5.13}$$

The $\sigma(t)$ in Equation (5.13) gives an implied volatility of $\sigma_{imp} \approx 0.165$. We hedge with a fixed number $k = 100$. Using this time-dependent $\sigma(t)$ and the constant σ_{imp} for delta hedging gives the result in Figure 5.6, where is shown that performing more hedges in the specific volatile period improves the variance of profit and loss during the hedging process.

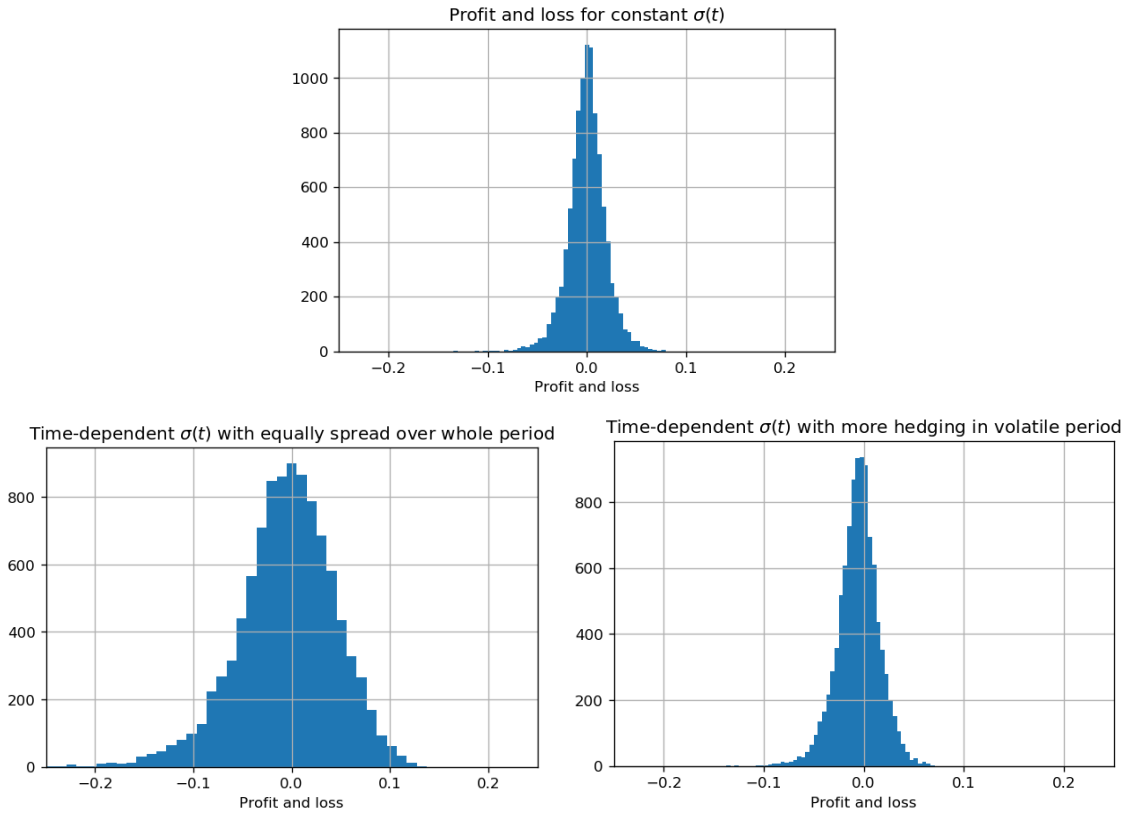

Figure 5.6: In the upper picture the profit and loss for a constant (implied) volatility function ($\sigma = 0.165$) can be observed. Hedges are performed equally spread over the whole time period from $t_0 = 0$ to $T = 10$. On the bottom we find two pictures where time-dependent volatility function $\sigma(t)$ from Equation (5.13) is used. On the left hedges are performed equally spread over the whole time period from $t_0 = 0$ to $T = 10$. On the right 70% of the hedges are performed in the volatile period from $t_0 = 0$ to $T = 1$. 10000 delta hedging simulations are performed with constant $r = 0.01$ and 100 hedges.

Considering Figure 5.6, the picture on the top and bottom right are very similar, whilst the picture on the bottom left has significantly increased variance on the profit and loss. Concentrating the rebalancing of portfolio during volatile periods has a positive impact on the variance of the profit and loss.

Theorem 5.2.1. Consider a Black Scholes world and an option at time t_0 with maturity T and strike K. Let the interest rate r be zero. Let $\sigma_1(t)$ and $\sigma_2(t)$ be two arbitrary volatility functions with equal implied volatility from t_0 to T, i.e. $\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t = \int_{t_0}^{T} \sigma_2^2(t)\mathrm{d}t$. Given that we start with a delta neutral portfolio and a fixed number of hedges for delta hedging. There exists a hedging strategy such that P&L(T) has the same distribution for $\sigma_1(t)$ and $\sigma_2(t)$.

Proof. Let k denote the number of hedges that will be performed. Both $\sigma_1(t)$ and $\sigma_2(t)$ have the same volatility from t_0 to T, which is $\sqrt{\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t}$. The idea is to choose the interval between hedges in such a way that the volatility over all periods between the hedges is the same. This can be done by choosing $t_i \in \{t_1, \ldots, t_k\}$ such that $\sqrt{\int_{t_i}^{T} \sigma_1^2(t)\mathrm{d}t} = \sqrt{\frac{k+1-i}{k+1}\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t}$. We can rewrite this to

$$\sqrt{\frac{k+1-i}{k+1}\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t} = \sqrt{\int_{t_i}^{T} \sigma_1^2(t)\mathrm{d}t} = \sqrt{\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t - \int_{t_0}^{t_i} \sigma_1^2(t)\mathrm{d}t} \tag{5.14}$$

$$\Rightarrow \int_{t_0}^{t_i} \sigma_1^2(t)\mathrm{d}t = \frac{i}{k+1}\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t \tag{5.15}$$

$$\Rightarrow \int_{t_i}^{t_{i+1}} \sigma_1^2(t)\mathrm{d}t = \frac{1}{k+1}\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t. \tag{5.16}$$

We will show that such a partition exists. Note that $\int_{t_0}^{t_i} \sigma_1^2(t)\mathrm{d}t$ is an increasing function for t_i. Therefore, one can find t_i such that $\int_{t_0}^{t_i} \sigma_1^2(t)\mathrm{d}t = \frac{1}{k+1}\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t$, where k is the number of hedges. Afterwards one can

iteratively find next the next value t_{i+1} such that $\int_{t_i}^{t_{i+1}} \sigma_1^2(t)\mathrm{d}t = \frac{1}{k+1}\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t$. One can continue to do so until $t_{i+1} = T_k$. Then we have k periods such that $\int_{t_i}^{t_{i+1}} \sigma_1^2(t)\mathrm{d}t = \frac{1}{k+1}\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t$ for $i = 0,\ldots,k-1$. This gives $\int_{t_0}^{t_k} \sigma_1^2(t)\mathrm{d}t = \frac{k}{k+1}\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t$. Therefore, we obtain for the last period:

$$\int_{t_k}^{T} \sigma_1^2(t)\mathrm{d}t = \int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t - \int_{t_0}^{t_k} \sigma_1^2(t)\mathrm{d}t = \frac{1}{k+1}\int_{t_0}^{T} \sigma_1^2(t)\mathrm{d}t, \tag{5.17}$$

which completes the partition.

Now we know a such a partition exists, we will prove that it will result in the same profit and loss at maturity time T. At every time point t_i we have the following

$$\mathrm{P\&L}(t_{i+1}) = \mathrm{P\&L}(t_i)\mathrm{e}^{\int_{t_i}^{t_{i+1}} r(t)\mathrm{d}t} - \big(\Delta(t_{i+1}) - \Delta(t_i)\big)S(t_{i+1}) \tag{5.18}$$

Given that we are at time point t_i, for arbitrary $\sigma(t)$ we have that:

$$S\big(t_{i+1}\big) = S_t\mathrm{e}^{\int_{t_i}^{t_{i+1}} r(t) - \frac{1}{2}\sigma^2(t)\mathrm{d}t + \int_{t_i}^{t_{i+1}} \sigma(t)\mathrm{d}W(t)} \tag{5.19}$$

The expectation of Equation (5.19) is given by:

$$\begin{aligned}
\mathbb{E}\big[S(t_{i+1})|\mathscr{F}(t_i)\big] &= \mathbb{E}\Big[S(t_i)\mathrm{e}^{\int_{t_i}^{t_{i+1}} r(t) - \frac{1}{2}\sigma^2(t)\mathrm{d}t + \int_{t_i}^{t_{i+1}} \sigma(t)\mathrm{d}W(t)}\Big|\mathscr{F}(t_i)\Big] \\
&= S(t_i)\mathrm{e}^{\int_{t_i}^{t_{i+1}} r(t) - \frac{1}{2}\sigma^2(t)\mathrm{d}t}\mathbb{E}\Big[\mathrm{e}^{\int_{t_i}^{t_{i+1}} \sigma(t)\mathrm{d}W(t)}\Big|\mathscr{F}(t_i)\Big] \\
&= S(t_i)\mathrm{e}^{\int_{t_i}^{t_{i+1}} r(t) - \frac{1}{2}\sigma^2(t)\mathrm{d}t}\mathbb{E}\Big[\mathrm{e}^{\frac{1}{2}\left(\int_{t_i}^{t_{i+1}} \sigma(t)\mathrm{d}W(t)\right)^2}\Big|\mathscr{F}(t_i)\Big] \\
&= S(t_i)\mathrm{e}^{\int_{t_i}^{t_{i+1}} r(t) - \frac{1}{2}\sigma^2(t)\mathrm{d}t}\mathbb{E}\Big[\mathrm{e}^{\frac{1}{2}\int_{t_i}^{t_{i+1}} \sigma^2(t)\mathrm{d}t}\Big|\mathscr{F}(t_i)\Big] \\
&= S(t_i)\mathrm{e}^{\int_{t_i}^{t_{i+1}} r(t)\mathrm{d}t},
\end{aligned}$$

where we used Ito isometry and the moment generating function. In order to obtain an expression for the variance, we compute $\mathbb{E}\big[\big(S(t_{i+1})\big)^2|\mathscr{F}(t_i)\big]$:

$$\begin{aligned}
\mathbb{E}\big[\big(S(t_{i+1})\big)^2|\mathscr{F}(t_i)\big] &= \mathbb{E}\Big[(S(t_i))^2 e^{2\int_{t_i}^{t_{i+1}} r(t) - \frac{1}{2}\sigma^2(t)\mathrm{d}t + 2\int_{t_i}^{t_{i+1}} \sigma(t)\mathrm{d}W(t)}\Big|\mathscr{F}(t_i)\Big] \\
&= S^2(t_i)\mathrm{e}^{2\int_{t_i}^{t_{i+1}} r(t) - \frac{1}{2}\sigma^2(t)\mathrm{d}t}\mathbb{E}\Big[\mathrm{e}^{2\int_{t_i}^{t_{i+1}} \sigma(t)\mathrm{d}W(t)}\Big|\mathscr{F}(t_i)\Big] \\
&= S^2(t_i)\mathrm{e}^{2\int_{t_i}^{t_{i+1}} r(t) - \frac{1}{2}\sigma^2(t)\mathrm{d}t}\mathbb{E}\Big[\mathrm{e}^{4(\int_{t_i}^{t_{i+1}} \sigma(t)\mathrm{d}W(t))^2}\Big|\mathscr{F}(t_i)\Big] \\
&= S^2(t_i)\mathrm{e}^{2\int_{t_i}^{t_{i+1}} r(t) - \frac{1}{2}\sigma^2(t)\mathrm{d}t}\mathbb{E}\Big[\mathrm{e}^{2\int_{t_i}^{t_{i+1}} \sigma^2(t)\mathrm{d}t}\Big|\mathscr{F}(t_i)\Big] \\
&= S^2(t_i)e^{\int_{t_i}^{t_{i+1}} 2r(t) + \sigma^2(t)\mathrm{d}t}.
\end{aligned}$$

Therefore, we obtain for the variance

$$Var\big[S(t_{i+1})|\mathscr{F}(t_i)\big] = \mathbb{E}\big[\big(S(t_{i+1})\big)^2|\mathscr{F}(t_i)\big] - \mathbb{E}\big[S(t_{i+1})|\mathscr{F}(t_i)\big]^2 = S^2(t_i)e^{2\int_{t_i}^{t_{i+1}} r(t)\mathrm{d}t}\Big(e^{\int_{t_i}^{t_{i+1}} \sigma^2(t)\mathrm{d}t} - 1\Big). \tag{5.20}$$

Note that the $\mathbb{E}\big[S(t_{i+1})|\mathscr{F}(t_i)\big]$ and $Var\big[S(t_{i+1})|\mathscr{F}(t_i)\big]$ is the same under $\sigma_1(t)$ and $\sigma_2(t)$, since both have equal $\int_{t_i}^{t_{i+1}} \sigma^2(t)\mathrm{d}t$ and $\int_{t_i}^{t_{i+1}} r(t)\mathrm{d}t = 0$.

Remember that $\Delta(t_i, S_{t_i})$ is defined as $\Phi\big(d_1(t_i, S_{t_i})\big)$, where $\Phi(\cdot)$ is the standard normal CDF and

$$d_1(t_i, S_{t_i}) = \frac{\frac{\log(S_{t_i})}{K} \int_{t_i}^{T} \big(\frac{1}{2}\sigma^2(t)\big)\mathrm{d}t}{\sqrt{\int_{t_i}^{T} \sigma^2(t)\mathrm{d}t}} \tag{5.21}$$

Note that $d_1(t_i, S_{t_i})$ and $d_1(t_{i+1}, S_{t_i})$ is the same under $\sigma_1(t)$ and $\sigma_2(t)$, since $\int_{t_{i,\sigma_1}}^{T} \sigma_1^2(t)\mathrm{d}t = \int_{t_{i,\sigma_2}}^{T} \sigma_2^2(t)\mathrm{d}t$, $r = 0$ and we showed S_{t_i} and $S_{t_{i+1}}$ have the same distribution under $\sigma_1(t)$ and $\sigma_2(t)$. Therefore, we have that $\Delta(t_i)$

and $\Delta(t_{i+1})$ are distributed the same under $\sigma_1(t)$ and $\sigma_2(t)$. For sake of notation we denote t_i under σ_1 as $t_i^{\sigma_1}$, whilst we normally denote under σ_1 as a subscript. Note that $t_0^{\sigma_1} = t_0^{\sigma_2} = t_0$.We know that at t_0 we have

$$\text{P\&L}_{\sigma_1}(t_0^{\sigma_1}) = V_{\sigma_1}(t_0^{\sigma_1}) - \Delta_{\sigma_1}(t_0^{\sigma_1}, S) = V_{\sigma_2}(t_0^{\sigma_2}) - \Delta_{\sigma_2}(t_0^{\sigma_2}, S) = \text{P\&L}_{\sigma_2}(t_0^{\sigma_2}). \tag{5.22}$$

Then by induction we have that:

$$\text{P\&L}_{\sigma_1}(t_{i+1}^{\sigma_1}) = \text{P\&L}_{\sigma_1}(t_i^{\sigma_1}) - (\Delta_{\sigma_1}(t_{i+1}^{\sigma_1}) - \Delta_{\sigma_1}(t_i^{\sigma_1}))S_{\sigma_1}(t_{i+1}^{\sigma_1}) \tag{5.23}$$

$$\stackrel{d}{=} \text{P\&L}_{\sigma_2}(t_i^{\sigma_2}) - (\Delta_{\sigma_2}(t_{i+1}^{\sigma_2}) - \Delta_{\sigma_2}(t_i^{\sigma_2}))S_{\sigma_2}(t_{i+1}^{\sigma_2}) = \text{P\&L}_{\sigma_2}(t_{i+1}^{\sigma_2}). \tag{5.24}$$

So, before the final transaction we have that $P\&L_{\sigma_1}(T) = P\&L_{\sigma_2}(T)$. Then the final transaction takes place

$$\text{P\&L}_{\sigma_1}(T) = -V_{\sigma_1}(T) + \Delta_{\sigma_1}(T, S) \stackrel{d}{=} -V_{\sigma_2}(T) + \Delta_{\sigma_2}(T, S) = \text{P\&L}_{\sigma_2}(T). \tag{5.25}$$

We conclude that $\text{P\&L}_{\sigma_1}(T) \stackrel{d}{=} \text{P\&L}_{\sigma_2}(T)$, which ends the proof. □

Note that the strategy used in Theorem 5.2.1 can be interpreted as a strategy where we spread the total volatility equally over all the intervals by finding such t_i's. As Example 5.2.1 shows, one would prefer to have equal volatility over the periods instead of lots of volatility focused in a specific period between to hedges. Therefore, given that one specifies the number of hedges beforehand, this way of choosing the time points for hedging can be considered as optimal for delta hedging. Some more examples will be given which illustrate Theorem 5.2.1.

In Theorem 5.2.1 we assume the interest rate to be zero, but what if $r(t)$ is not zero? Does the same proof hold? No, the same proof does not hold, since the $\mathbb{E}\big[S(t_{i+1})|\mathscr{F}(t_i)\big]$ and $Var\big[S(t_{i+1})|\mathscr{F}(t_i)\big]$ do not have to be the same for $\sigma_1(t)$ and $\sigma_2(t)$ by the proposed partition in the proof. Perhaps there exists another more advanced way to compute a partition that would incorporate the non-zero interest rate. Finding such a partition is challenging, as interest rates could be negative as well, causing $\int_{t_0}^{T} r(t)\mathrm{d}t$ to be non-monotone. During the research of this thesis, such a partition has not been found.

Example 5.2.4. We show an example of Theorem 5.2.1. $t_0 = 0$, $T = 10$, $r = 0$, $S_0 = 1.0$ and $K = 1.0$. Let $\sigma_1(t)$ and $\sigma_2(t)$ be defined by

$$\sigma_1(t) = \begin{cases} 0.5 & \text{if } t < 1 \\ 0.05 & \text{if } t \geq 1, \end{cases}, \qquad \sigma_2(t) \approx 0.165 + 0 \cdot t. \tag{5.26}$$

Both $\sigma_1(t)$ and $\sigma_2(t)$ have the same implied volatility $\sigma_{imp} \approx 0.165$. We start with a delta neutral portfolio and define beforehand that we will hedge 10 times. Using our method in Theorem 5.2.1, we obtain the following sets for $t_i^{\sigma_1}$ and $t_i^{\sigma_2}$

$$t_i^{\sigma_1} \in \{0.099, 0.198, 0.297, 0.396, 0.495, 0.595, 0.0694, 0.793, 0.892, 0.991\} \tag{5.27}$$

$$t_i^{\sigma_2} \in \left\{\frac{10}{11}, \frac{20}{11}, \frac{30}{11}, \frac{40}{11}, \frac{50}{11}, \frac{60}{11}, \frac{70}{11}, \frac{80}{11}, \frac{90}{11}, \frac{100}{11}\right\} \tag{5.28}$$

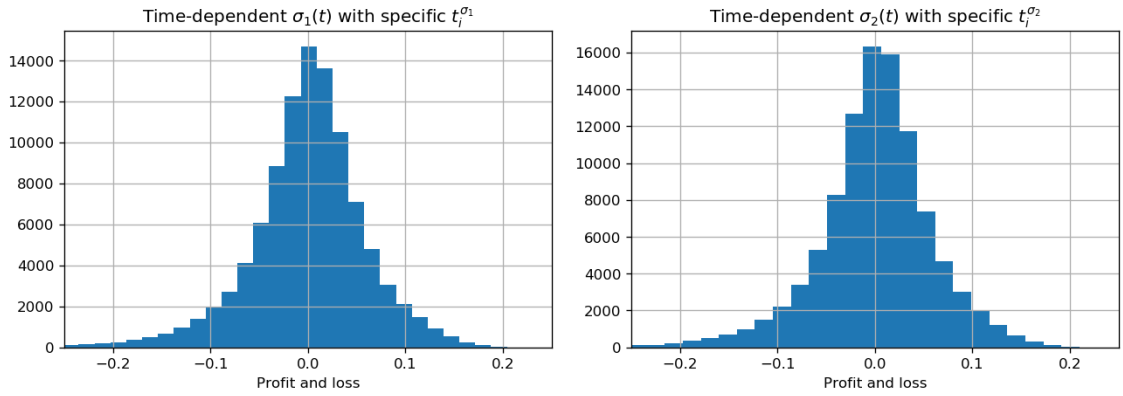

Figure 5.7: 100000 simulations have been performed. The histogram on the left has a mean $8 \cdot 10^{-5}$ and standard deviation of 0.060. The histogram on the right has a mean $-2 \cdot 10^{-4}$ and standard deviation of 0.060. We find that the profit and loss at time T of both hedging strategies are distributed with approximately mean 0 and standard deviation of 0.060.

In Figure 5.7, one can find the profit and loss at maturity time for both volatility functions σ_1 and σ_2 with corresponding hedging time points $t_i^{\sigma_1}$ and $t_i^{\sigma_2}$. We observe that both have the same distribution. If we would use the hedging time points $t_i^{\sigma_2}$ for σ_1, we would obtain the picture visible in Figure 5.8. We see that using hedging time points $t_i^{\sigma_2}$ for delta hedging with σ_1, results in a worse performance compared to using hedging time points $t_i^{\sigma_1}$.

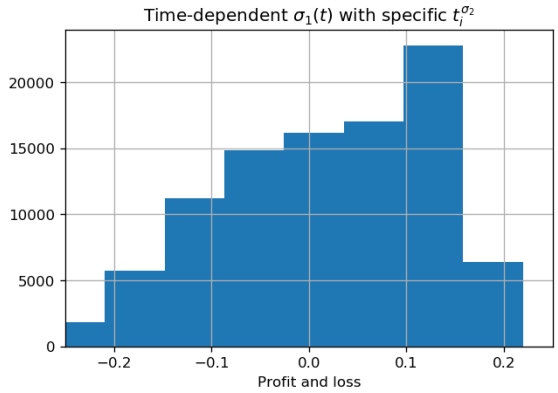

Figure 5.8: 100000 simulations have been performed. The histogram on the left has a mean $8 \cdot 10^{-5}$ and standard deviation of 0.060. The histogram has standard deviation of 0.16.

Example 5.2.5. We show an example of Theorem 5.2.1. $t_0 = 0$, $T = 10$, $r = 0$, $S_0 = 1.0$ and $K = 1.0$. Let $\sigma_1(t)$ and $\sigma_2(t)$ be defined by

$$\sigma_1(t) = \begin{cases} 0.5 & \text{if } t < 1 \\ 0.05 & \text{if } t \geq 1, \end{cases} \qquad \sigma_2(t) \approx 0.165 + 0 \cdot t. \tag{5.29}$$

Both $\sigma_1(t)$ and $\sigma_2(t)$ have the same implied volatility $\sigma_{imp} \approx 0.165$. Using the strategy from Theorem 5.2.1, we perform delta hedging using several numbers of hedges. Afterwards, we compare the standard deviation on the profit and loss at maturity time. The results can be found in Figure 5.9. The figure confirms that the standard deviation of time dependent $\sigma(t)$ in Equation (5.29) matches the standard deviation from the constant σ for any number of hedges using the strategy as in Theorem 5.2.1.

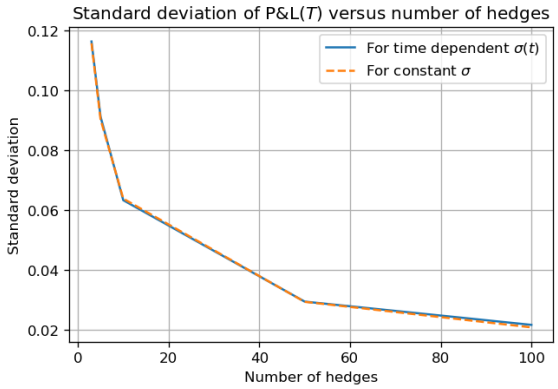

Figure 5.9: Shown in the standard deviation obtained by the delta hedging strategy from Theorem 5.2.1 for several number of hedges.

5.3. Effect of time-dependent parameters on delta hedging

If one continuously hedges, the profit and loss for arbitrary $\sigma(t)$ is perfect. In reality, this cannot be done due to transaction costs. Example 5.2.1 shows that time-dependent volatility parameter with a finite amount of trades for hedging can have a negative impact on the variance of the profit and loss. In Theorem 5.2.1 we show that given a fixed number of hedges with interest rate equal zero there exists a choice of time points for delta hedging, which will results in the same distribution of profit and loss at capital T for any $\sigma(t)$, that has

same implied volatility from t_0 to t. We argue that the method in Theorem 5.2.1 is the best way to choose time points for hedging, when one chooses a fixed number of hedges beforehand. In the experiments in this thesis, hedges are performed at certain times decided beforehand, which can be considered as unrealistic. In reality, traders could perform hedges at certain moments when the delta of the portfolio exceeds a specific threshold. This alternative strategy causes more trades for hedging to happen during the volatile periods, which corresponds to the strategy in Theorem 5.2.1. In conclusion, we argue that the effect of time-dependent volatility parameter on the delta hedging process is minimal, and does not outweigh the improved calibration by using time-dependent volatility parameter instead of a constant volatility parameter. Also, we see no reason to include homogeneity constrains to the volatility parameter with respect to the performance of delta-hedging.

6

Conclusion and outlook

6.1. Summary

This thesis captures an extension of a known method in literature [24] for the calibration of the Hull-White model. In this method we first calibrated the mean reversion parameter independently from the volatility parameter to the ratio of two swaptions with same expiry but different tenor. Afterward, we calibrate the volatility parameter, to a suitable set of swaptions, e.g. a co-terminal basket. We extend the existing method to the negative interest rate environment, by modeling the swap rate as shifted lognormal instead of lognormal. Besides this extension, we show that a smart choice can be made for the set of implied volatility fractions used for the calibration of the mean reversion parameter, causing the optimization problem to be independent of one another. Therefore, we can use sequential optimization for the calibration of the mean reversion parameter as a piece-wise constant function, which improves the speed. Besides, we show that under this specific choice of implied volatility ratios, the sequential optimization problems of the mean reversion parameter are convex. After the calibration of the mean reversion parameter, the volatility parameter is calibrated to a co-terminal basket of swaptions. On market data the calibration is well fitted. In this work the importance of a proper chosen or calibrated mean reversion parameter is emphasised. The effect of mean reversion on the calibration of the volatility parameter and expected exposure is researched by using various mean reversion parameters for calibration. In general, larger mean reversion parameters appear to result in larger volatility parameters. The results show that different mean reversion parameters result in different quality of the shape of the expected exposure.

A second novelty in this thesis lies within the calibration of an FX hybrid model; the Black-Scholes Hull-White model. Numerical results on the calibration of FX hybrid models are scarce in literature. This thesis provides more results of FX hybrid models to the literature. Besides, we show how different mean reversion parameters of the Hull-White model affect the calibration of the volatility parameter in the Black-Scholes Hull-White model. In our results, the effect of different mean reversion parameters appears to influence the calibration of the volatility parameter $\sigma(t)$ in the Black-Scholes Hull-White model, although the effect appears to be negligible.

In this thesis we looked into delta-hedging in a Black-Scholes world to investigate how the homogeneity of time-dependent parameters influence the hedging performance. Eventually, this lead to the proof of Theorem 5.2.1 in Section 5. In Theorem 5.2.1 we prove that given a fixed amount of hedges beforehand and zero interest rate, for all $\sigma(t)$ such that the implied volatility from t_0 to T is equal, there exists an optimal distribution of hedging time points such that the variance on the profit and loss on maturity is minimal. Due to this theorem, we argue that big differences in the calibration of piece-wise constant $\sigma(t)$ do not have that big of an impact on the hedging process. Therefore, it feels unnecessary to incorporate specific rules to ensure the homogeneity of the calibrated parameters. The time-dependent parameters should be able to capture big moves in extreme events, e.g. Covid 19 pandemic.

6.2. Limitations

As we propose a new method for calibration, it is important to reflect critically on the assumptions and results. This critical reflection results in some limitations for this research. These limitations are listed below.

- The novel method for the calibration of the mean reversion parameters in the Hull-White model has a limitation considering the special set of swaption we use for calibration. We need this specific set of swaptions to be available in the market for calibration. In some interest rate swaption markets that are not very liquid, these might not be available. In that case, one can still calibrate to ratio of swaptions, but we lose the independence and convex property. This causes the optimization problem to be more complex, therefore increasing the computation time.

- Another limitation is the comparison from delta-hedging in our simple Black-Scholes world to delta-hedging in the real world. Our Black-Scholes world is simple, e.g. we don't incorporate transaction costs. One can argue that this is not very realistic for that reason.

- Homogeneity constraints are not applied in the calibration of the parameters due to our findings in the chapter on delta-hedging. This can be considered as a misstep, because one would not want a piece-wise constant function that is not fluctuating so much. You would not expect these parameters to fluctuate much in reality. However, one can also argue these parameters should have the freedom to fluctuate in order to cope with extreme events in the market.

- In this thesis the Bezier curve is used for the computation of $\theta(t)$ in order to allow numerical computation of the derivative of the instantaneous forward rate. As one can see in Figure 2.2, the Bezier curve does not fit all the data points. This can be considered as a drawback. Still, the Bezier curve is merely used for Monte Carlo validation purposes.

- In this thesis we chose to use at the money products for the calibration, since these are most liquid in the market. However, one can argue that out of the money options should be used for the calibration, because most instruments available in the market are out of the money. One would like to buy a swaption or option because of the optionality. Therefore, it seems more natural that a portfolio would consist of mostly out of the money swaptions and options. This is a valid line of reasoning. The method presented in this thesis should work on out of the money market data as well, although, these calibration results are not presented in this thesis.

6.3. Future research

During our research, we encountered extensions of our research topic that would be interesting fields for research. As time is always limited for the thesis, we would like to state some of these ideas for future research. The potential of these research topics and clearer directions can be determined by need from the industry.

- The delta-hedging method is very basic. A next step would be adding transaction costs. A second element that could be added is looking in more risks than delta hedging only, which only includes risks to the movement of the asset. For example, vega-hedging could be interesting, which examines the risk of movements in the asset with respect to the volatility. Besides the strategies that we looked into, different delta hedging strategies can be examined, e.g. looking into delta-hedging only when specific thresholds are exceeded. Also, the effect of the volatility parameter on the delta-hedging procedure is researched, and the effect of a mean reversion parameter is not examined. In order to do this, one would have to assume another model instead of a Black-Scholes model, that includes a mean reversion parameter. For next research, these are things one could look into to make the modeling more realistic.

- Machine learning is a fairly new approach for calibration of financial models, e.g. Hull-White model. Recent developments show the potential of machine learning in this area. For example, in [11] and [21] calibration on the Hull-White model with constant parameters is performed by machine learning and the results show some potential. Even in [21] time-dependent parameters of the Hull-White model are assumed. Since machine learning is often considered as a "black box", it is uncertain to what extent the industry would be willing to use these methods.

- Besides the FX Black-scholes Hull-White, one could look into more FX hybrid models, and compare calibration results. One example is the Heston-Hull-White model, where the volatility on the FX rate is assumed to be stochastic and is modeled by a mean reversion process.

- The effect of negative mean reversion parameter in the Hull-White model is an excellent topic for research. In the literature there appears to be no clear agreement to what extent negative mean reversion

parameters should be allowed for calibration purposes. We argued in Section 2.1.1 that we allow negative mean reversion although it is against the nature of the mean reversion parameter. However, we have not been able to quantify to what extent negative mean reversion is permitted. Researching a method to quantify the impact of negative mean reversion on the calibration would be very interesting, and also fruitful for the industry.

- The alternative approximation in Section 3.3.4 could be researched further. If one would be able to decrease the computation time, it could be a proper approximation to include in a method for calibration. Also, more research on the error of both approximations in Section 3.3.5 would be insightful.

6.4. Conclusion

Although the Hull-White model exists for nearly 20 years [13], there is no general consensus in the market on the best method for calibration to market data. Also, it is not clear if or to what extent negative mean reversion is acceptable. This thesis proposes a novel and fast method for calibration of the mean reversion parameter, which is an extension of an existing method in the literature [24]. The calibration results on artificial data look promising. The calibration to market data obtains a perfect fit with the market swaption implied volatilities. The results in this thesis emphasise that the mean reversion parameter should be calibrated carefully. If mean reversion parameter is chosen arbitrarily, one can obtain different calibration of $\eta(t)$ and the expected exposure could behave bad for not calibrated implied volatilities in the same co-terminal basket as can be observed in Figure 3.11.

In the calibration of the Black-Scholes Hull-White model there is a general problem in the market in the calibration for late maturities. This thesis provides examples of calibration, where we do not experience this problem. Therefore, it appears that problem of poor calibration for late maturities can be solved.

Although this thesis provides a novel method for the calibration of the Hull-White model, we would like to stress that more testing and research is necessary before the method could used in the industry. The precise effects of the negative mean reversion parameter on the calibration processes should be quantified in greater detail before one allows negative mean reversion. Moreover, the natural environment often has proven to be more unpredictable than many experts or models anticipate, which is another reason for recommending precision and patience. At the same time, we conclude by sharing our hope that some of the findings of this thesis may well be applicable in some form in the near or more distal future.

A

Calibration algorithms

Algorithm 4: Calibration algorithm for Hull-White model.

- Choose desired[a] set of ATM implied volatility swaption ratios from the market data, such that $T_i^j < T_{m_1}^j < T_{m_2}^j$ and for all $n, p \in \{1, \ldots, k\}, n \neq m$ that $(T_i^n, T_{m_2}^n) \bigcap (T_i^p, T_{m_2}^p) = \emptyset$.[b]
- Calibrate $\lambda(t)$ to the implied volatility ratios in a sequential manner:

for *j in n* **do**

$$\lambda(t) \text{ for } T_i^j \leq t < T_{m_2}^j = \underset{\lambda(t)}{\operatorname{argmin}} \left(\sqrt{\sum_j \left(\frac{\sigma_{swpt}^{mrkt}(t_0, T_i^j, T_{m_1}^j, K)}{\sigma_{swpt}^{mrkt}(t_0, T_i^j, T_{m_2}^j, K)} - \frac{\sigma_{swpt}^{HW}(t_0, T_i^j, T_{m_1}^j, K, \lambda(t))}{\sigma_{swpt}^{HW}(t_0, T_i^j, T_{m_2}^j, K, \lambda(t))} \right)^2} \right)$$

end

- Pick the desired[c] co-terminal set of ATM swaptions.

Try:

Calibrate $\eta(t)$ with calibrated $\lambda(t)$ to the co-terminal set of ATM swaptions.

$$\eta(t) = \underset{\eta(t)}{\operatorname{argmin}} \left(\sqrt{\sum_i \left(\sigma_{swpt}^{mrkt}(t_0, T_i, T_m, K) - \sigma_{swpt}^{HW}(t_0, T_i, T_m, K, \lambda(t), \eta(t)) \right)^2} \right)$$

Except:

- Calibrate constant λ to the implied volatility ratio under constrain that λ is non-negative.

$$\lambda = \underset{\lambda}{\operatorname{argmin}} \left(\sqrt{\sum_j \left(\frac{\sigma_{swpt}^{mrkt}(t_0, T_i^j, T_{m_1}^j, K)}{\sigma_{swpt}^{mrkt}(t_0, T_i^j, T_{m_2}^j, K)} - \frac{\sigma_{swpt}^{HW}(t_0, T_i^j, T_{m_1}^j, K, \lambda)}{\sigma_{swpt}^{HW}(t_0, T_i^j, T_{m_2}^j, K, \lambda)} \right)^2} \right)$$

- calibrate $\eta(t)$ with (non-negative) λ to the co-terminal set of ATM swaptions.

$$\eta(t) = \underset{\eta(t)}{\operatorname{argmin}} \left(\sqrt{\sum_i \left(\sigma_{swpt}^{mrkt}(t_0, T_i, T_m, K) - \sigma_{swpt}^{HW}(t_0, T_i, T_m, K, \lambda, \eta(t)) \right)^2} \right)$$

Return: $\lambda(t)$ and $\eta(t)$.

[a] The desired set depends on the frequency of jumps in the piece-wise constant function.

[b] (A, B) denotes the open interval from A to B, $\bigcap$ denotes the intersection and $\emptyset$ denotes the empty set.

[c] The desired set depends on the portfolio of the trader and the wish of how the piece-wise constant parameter $\eta(t)$ looks like.

Algorithm 5: Calibration algorithm for Black-Scholes Hull-White model.

- Calibrate the mean reversion parameter $\lambda_d(t)$ and volatility parameter $\eta_d(t)$ of the domestic currency using Algorithm 4.
- Calibrate the mean reversion parameter $\lambda_f(t)$ and volatility parameter $\eta_f(t)$ of the foreign currency using Algorithm 4.
- Calibrate $\sigma(t)$ to the ATM options on the FX rate available in the market with corresponding correlations observable in the market:

$$\underset{\sigma(t)}{\operatorname{argmin}} \left(\sqrt{\sum_i \left(\sigma_{FX}^{mrkt}(t_0, T_i, K) - \sigma_{FX}^{BSHW}\left(t_0, T_i, K, \lambda_d(t), \eta_d(t), \lambda_f(t), \eta_f(t), \rho_{y,d}, \rho_{y,f}, \rho_{d,f}, \sigma(t)\right) \right)^2} \right)$$

Return: $\lambda_d(t)$, $\eta_d(t)$, $\lambda_f(t)$, $\eta_f(t)$ and $\sigma(t)$.

B

Shape of swaption implied volatility for large mean reversion

In order to understand this Appendix, we repeat some definitions from Section 2.1. The dynamics of the Hull-White model are given by:

$$\mathrm{d}r(t) = \lambda(t)\big(\theta(t) - r(t)\big)\mathrm{d}t + \eta(t)\mathrm{d}W_r(t).$$

The variance of process $r(t)$ is given by:

$$v_r\big(\lambda(t), \eta(t), t\big) = v_r(t) := \mathrm{Var}\big[r(t)\big|\mathscr{F}_{t_0}\big] = \frac{1}{E^2(t)}\int_{t_0}^{t} \eta^2(z)E^2(z)\mathrm{d}z,$$

where $E(\cdot)$ is defined as:

$$E\big(\lambda(t), t\big) = E(t) := \mathrm{e}^{\int_{t_0}^{t}\lambda(t)\mathrm{d}t}.$$

In this Appendix we show that piece-wise constant volatility parameter can create "unexpected" curves in the price of the swaption for large mean reversion parameter. This phenomena is caused by the fact that large mean reversion leads to little impact of past volatility on the variance of process $r(t)$. In order to illustrate this phenomena, we use constant mean reversion λ. For example, if we take $\lambda = 1$ in the dynamics of $r(t)$, process $r(t)$ will go directly to $\theta(t)$ by the $\mathrm{d}t$ term in the dynamics. Therefore, the volatility that happened before has little-to-no influence on the price, because the volatility of process r(t) that happened earlier is killed by the mean reversion. We illustrate this by an example.

Example. In this example we pick two volatility functions. We show that the variance at time 15 is similar for a large constant mean reversion, although the volatility is different in the beginning. We consider the following two volatility functions:

$$\eta_1(t) = \begin{cases} 0.12 & \text{if } t < 10 \\ 0.11 & \text{if } t \geq 10, \end{cases}$$
$$\eta_2 = 0.11.$$

For constant mean reversion $\lambda = 0.5$, $t_0 = 0$ and $t = 15$ we obtain for the variance of $\eta_1(t)$ and η_2:

$$v_r^{\eta_1(t)}(t) = \frac{0.12}{E^2(t)}\int_{t_0}^{10} E^2(z)\mathrm{d}z + \frac{0.11}{E^2(t)}\int_{10}^{t} E^2(z)\mathrm{d}z \approx 0.00081 + 0.10926 \approx 0.1101,$$
$$v_r^{\eta_2}(t) = \frac{0.11}{E^2(t)}\int_{t_0}^{10} E^2(z)\mathrm{d}z + \frac{0.11}{E^2(t)}\int_{10}^{t} E^2(z)\mathrm{d}z \approx 0.00074 + 0.10926 \approx 0.1100.$$

We find that the volatility from 0 till 10 has very little effect on the variance at time 15.

The $\frac{1}{E^2(t)}$ term in the variance of $r(t)$ causes the volatility at the beginning of process $r(t)$ to have little-to-no influence on the variance for large mean reversion parameter. The value of a swaption in Theorem 2.1.2 depends on the volatility parameter only by the value of options on zero coupon bonds and $A(\cdot,\cdot)$. The value of an option on zero coupon bonds in Theorem 2.1.1 and $A(\cdot,\cdot)$ in Equation (2.10) only depend on the volatility term parameter by the variance of $r(t)$. Therefore, for the value of swaption we experience the same property as for the variance of $r(t)$: The value of a swaption is effected very little by volatility from the beginning for large mean reversion. So, by having a larger mean reversion λ the volatility that happened before has less impact on the pricing. Therefore, if you have a big jump in the volatility parameter and a large λ, you will see a fast jump in the pricing. This is illustrated in Figure B.1.

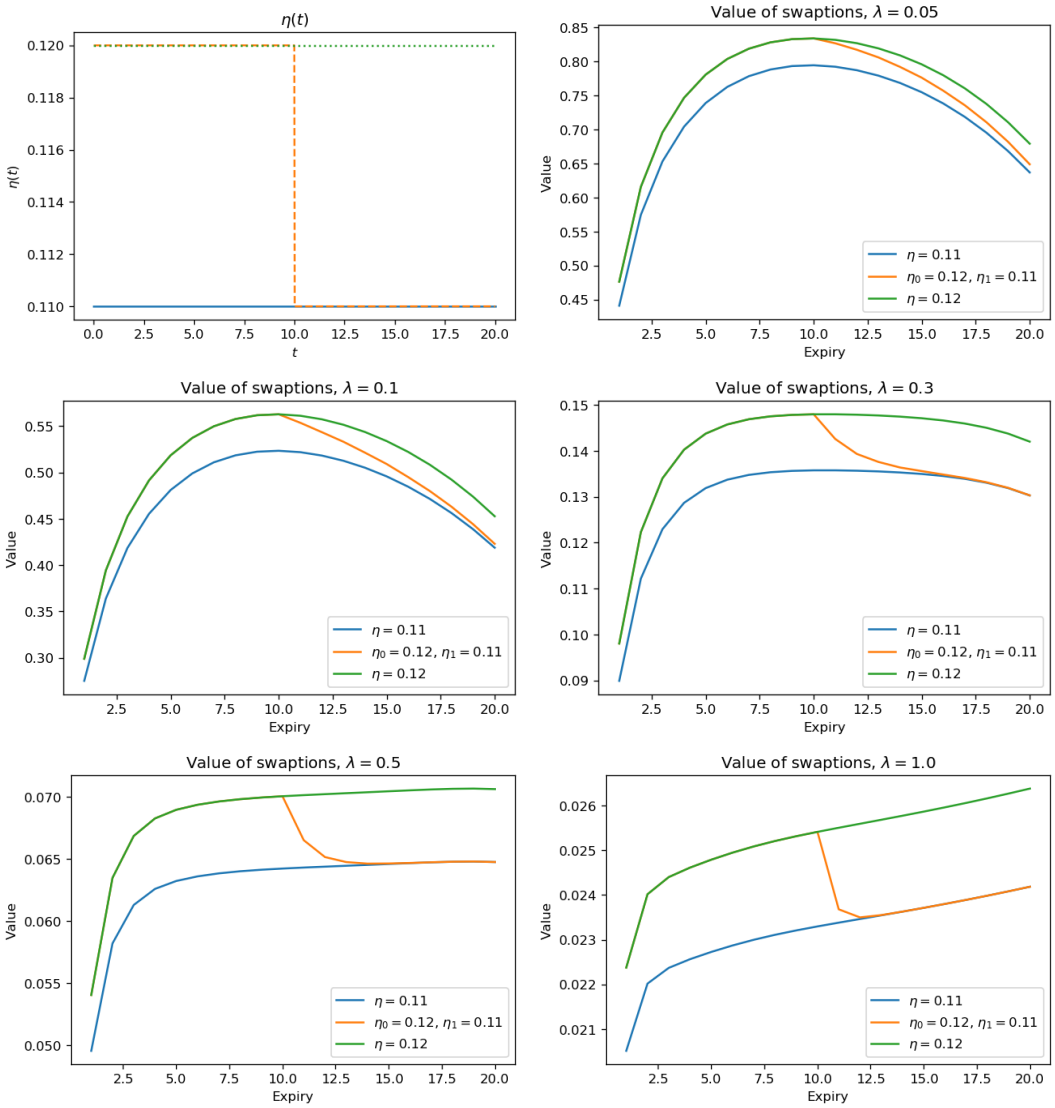

Figure B.1: Swaption values are computed for several mean reversion parameters under three different volatility functions, two constant volatility functions and one piece-wise constant function that changes at $t = 10$. Discount curve is observed from the market, $\tau = 0.5$ and $T_m = 30$. The Example in this Appendix connects to the plot on the left bottom in this Figure for $T_i = 15$.

C

BSHW calibration examples

In this Appendix two examples of calibration of the Black-Scholes Hull-White model are given. The first example is on USD-EUR data. The second example is on GBP-EUR data.

C.1. BSHW calibration example USD-EUR

We give an example of the calibration of the Black-Scholes Hull-White model to market data. The domestic currency is the dollar and foreign currency is the euro. In Figure C.1 we can find the calibration of $\sigma(t)$. $\sigma(t)$ is calibrated to the ATM FX options available in the market. An alternative calibration is available. In this calibration we can find the general difficulty of obtaining good fit for late maturity. In our calibration we do obtain a good fit for late maturity, showing that it is possible. Algorithm 5 has been used for the calibration of $\sigma(t)$.

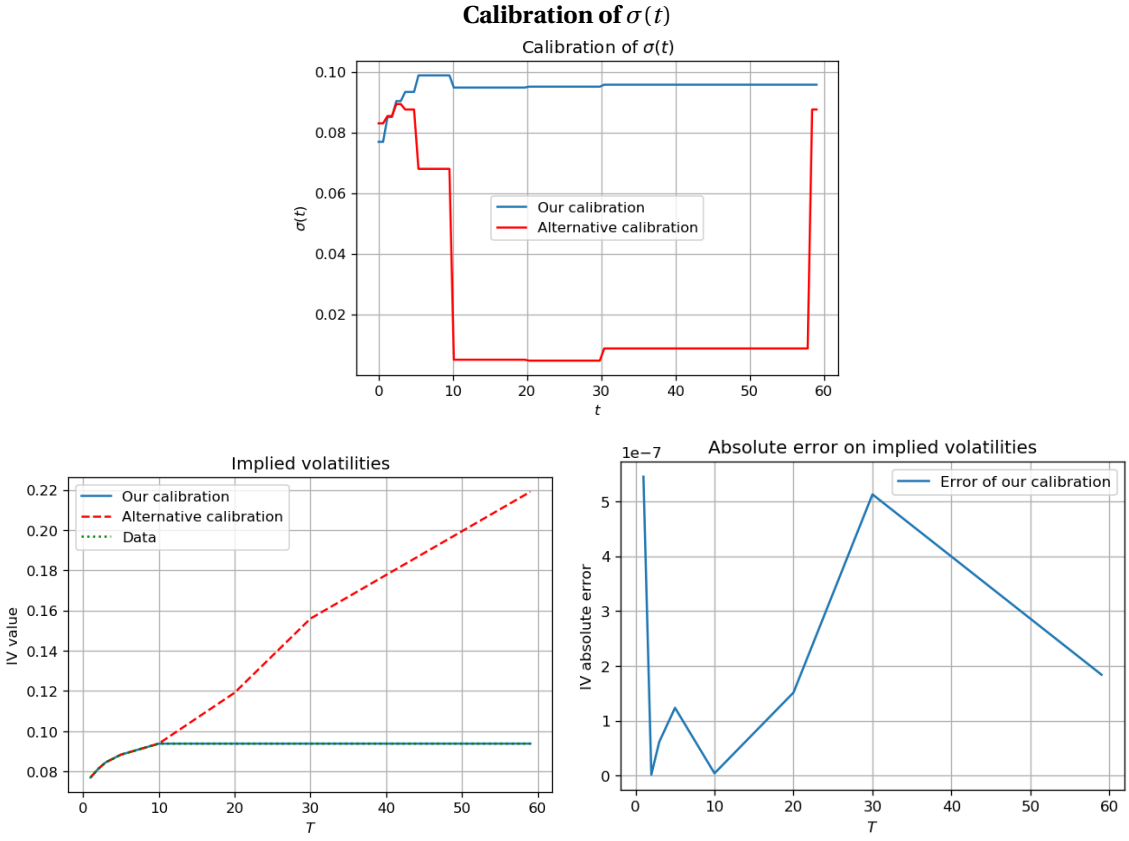

Figure C.1: The calibration of the FX volatility parameter $\sigma(t)$ can be observed in the plots.

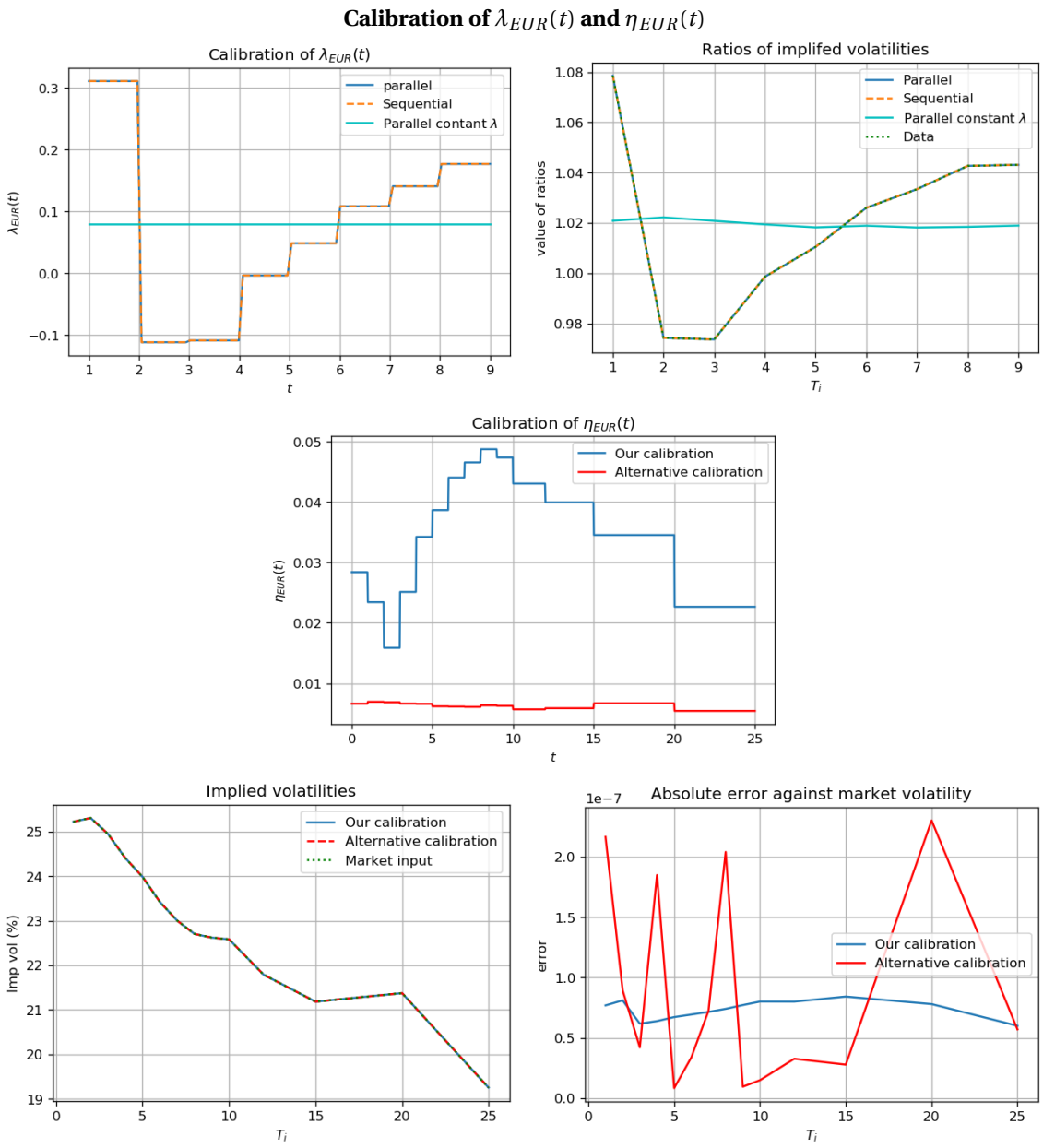

Figure C.2: The calibration of the mean reversion parameter $\lambda_{EUR}(t)$ and volatility parameter $\eta_{EUR}(t)$ can be observed in the plots.

Figure C.2 shows us the results of the calibration of the euro Hull-White process. The calibration on euro swaptions has been performed by first calibrating the mean reversion parameter $\lambda_{EUR}(t)$, afterward calibrating volatility parameter $\eta_{EUR}(t)$. The calibration procedure as shown in algorithm 4 is used. The obtained calibrated $\lambda_{EUR}(t)$ is suitable for calibration of $\eta_{EUR}(t)$. Therefore, we use time-dependent $\lambda_{EUR}(t)$ for the calibration of $\eta_{EUR}(t)$.

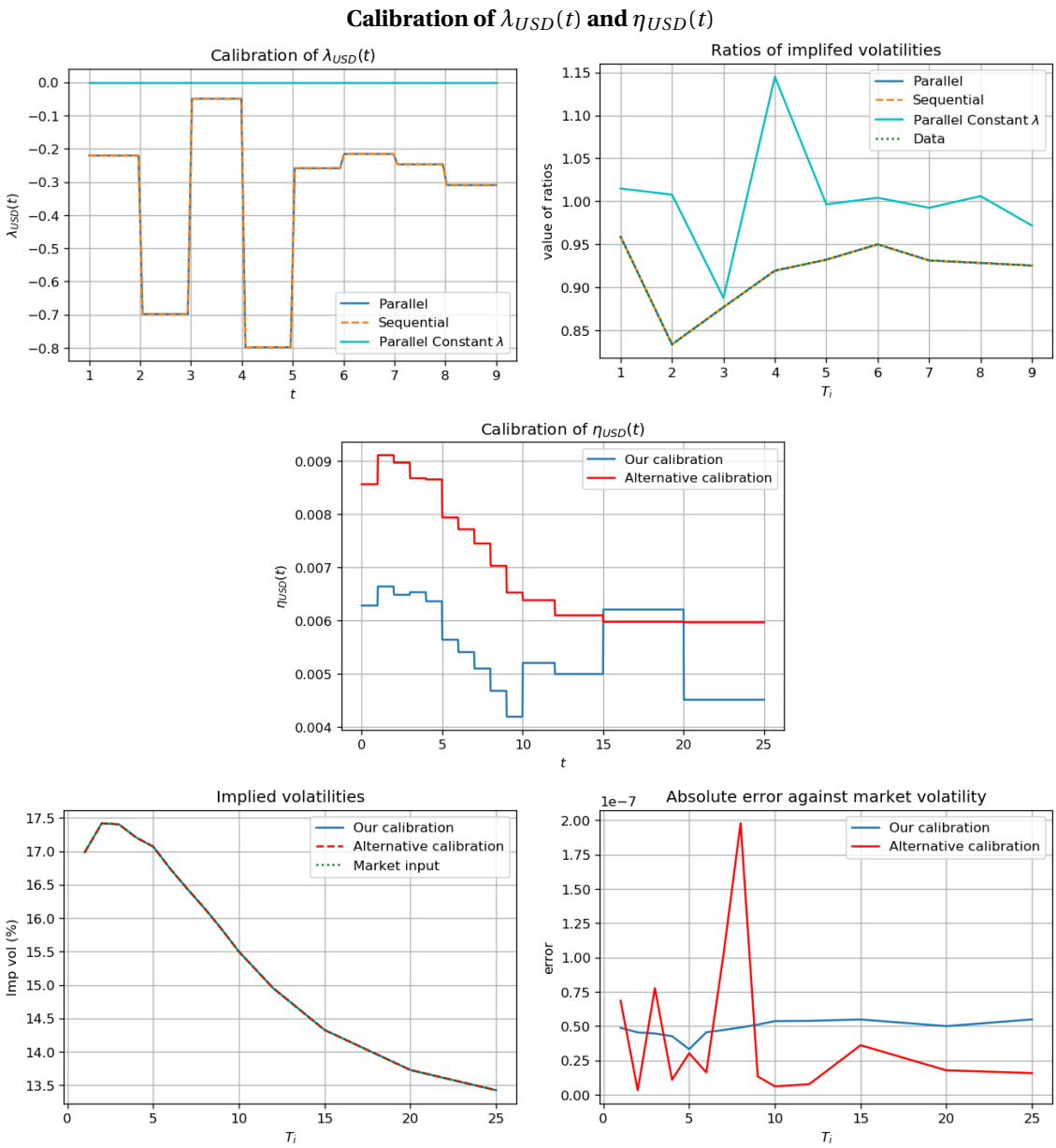

Figure C.3: The calibration of the mean reversion parameter $\lambda_{USD}(t)$ and volatility parameter $\eta_{USD}(t)$ can be observed in the plots.

Figure C.3 shows us the results of the calibration of the USD Hull-White process. The calibration on dollar swaptions has been performed by first calibrating the mean reversion parameter $\lambda_{USD}(t)$, afterward calibrating volatility parameter $\eta_{USD}(t)$. The calibration procedure as shown in algorithm 4 is used. We obtain a time-dependent $\lambda_{USD}(t)$, that is unable to calibrate $\eta(t)$. Therefore, we choose the calibrated positive constant λ_{USD} very close to zero. For this constant λ_{USD}, we are able to calibrate the volatility parameter $\eta_{USD}(t)$ to the co-terminal swaption implied volatilties.

C.2. BSHW calibration example GBP-EUR

An example of the calibration of the Black-Scholes Hull-White model to market data. The domestic currency is Pound sterling and foreign currency is the euro. In Figure C.4 we can find the calibration of $\sigma(t)$. $\sigma(t)$ is calibrated to the ATM FX options available in the market. An alternative calibration is available. In this calibration we can find the general difficulty of obtaining good fit for late maturity. In our calibration we do obtain a good fit for late maturity, showing that it is possible. Algorithm 5 has been used for the calibration of $\sigma(t)$.

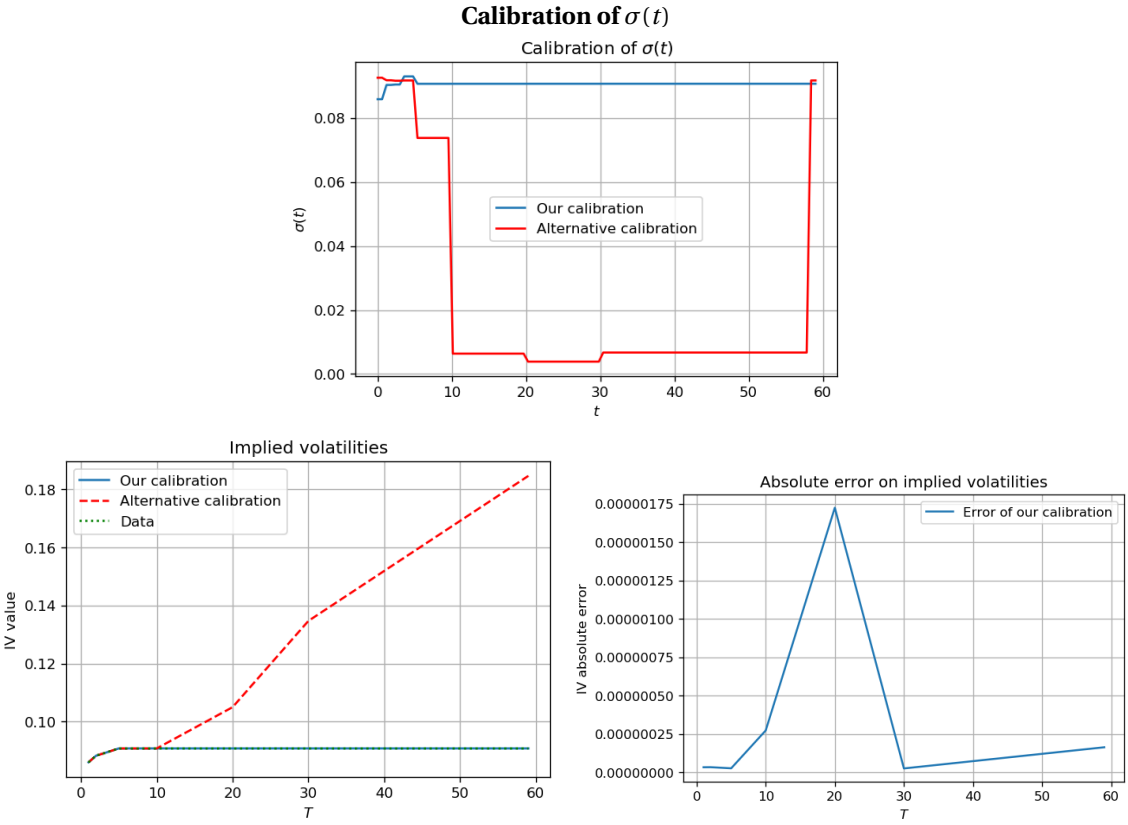

Figure C.4: The calibration of the FX volatility parameter $\sigma(t)$ can be observed in the plots.

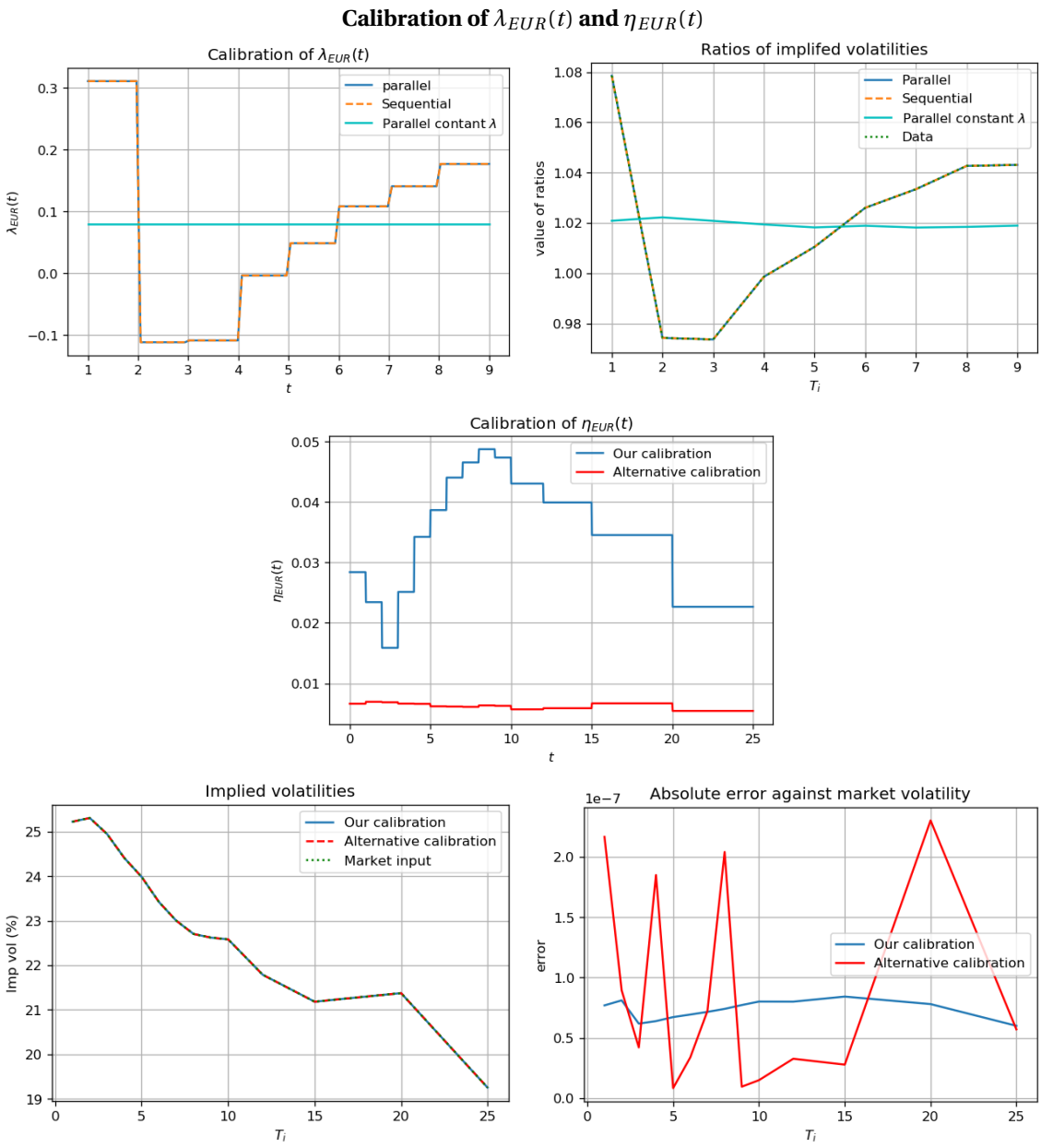

Figure C.5: The calibration of the mean reversion parameter $\lambda_{EUR}(t)$ and volatility parameter $\eta_{EUR}(t)$ can be observed in the plots.

Figure C.5 shows us the results of the calibration of the euro Hull-White process. The calibration on euro swaptions has been performed by first calibrating the mean reversion parameter $\lambda_{EUR}(t)$, afterward calibrating volatility parameter $\eta_{EUR}(t)$. The calibration procedure as shown in algorithm 4 is used. The obtained calibrated $\lambda_{EUR}(t)$ is suitable for calibration of $\eta_{EUR}(t)$. Therefore, we use time-dependent $\lambda_{EUR}(t)$ for the calibration of $\eta_{EUR}(t)$.

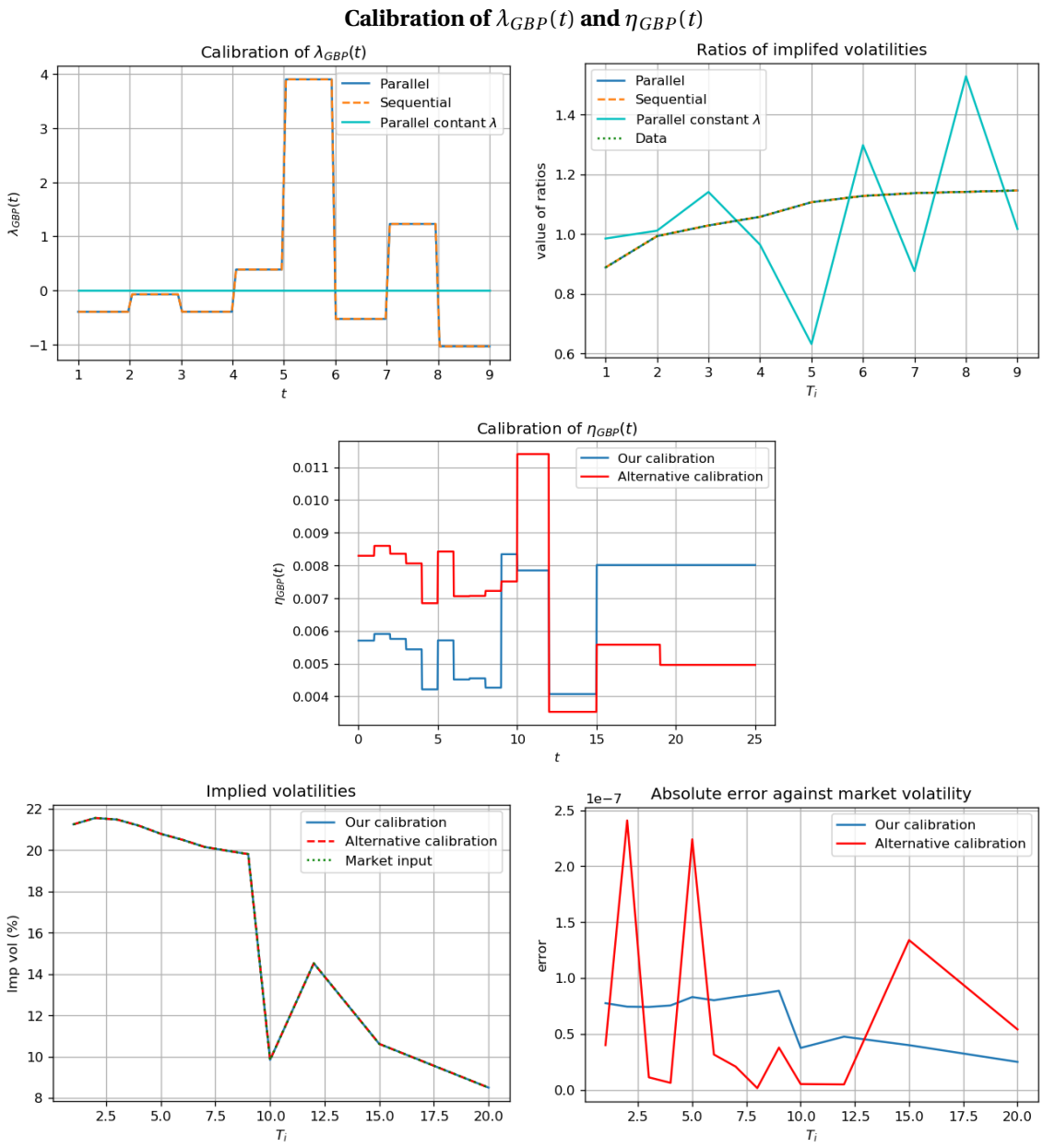

Figure C.6: The calibration of the mean reversion parameter $\lambda_{GBP}(t)$ and volatility parameter $\eta_{GBP}(t)$ can be observed in the plots.

Figure C.6 shows us the results of the calibration of the Pound sterling Hull-White process. The calibration on Pound sterling swaptions has been performed by first calibrating the mean reversion parameter $\lambda_{GBP}(t)$, afterward calibrating volatility parameter $\eta_{GBP}(t)$. The calibration procedure as shown in algorithm 4 is used. We obtain a time-dependent $\lambda_{GBP}(t)$, that is not able to calibrate. Therefore, we use the calibrated positive constant λ_{GBP}. For this calibrated constant λ_{GBP}, we are able to calibrate the volatility parameter $\eta_{GBP}(t)$ to the co-terminal swaption implied volatilities.

Bibliography

[1] Leif B.G. Andersen and Vladimir V. Piterbarg. *Interest Rate Modeling Volume 2 Term structure models*. Atlantic Financial Press, 2010.

[2] Marco Bianchetti. Two curves, one price: Pricing & hedging interest rate derivatives decoupling forwarding and discounting yield curves. *One Price: Pricing & Hedging Interest Rate Derivatives Decoupling Forwarding and Discounting Yield Curves (November 14, 2008)*, 2008.

[3] Fischer Black. The pricing of commodity contracts. *Journal of financial economics*, 3(1-2):167–179, 1976.

[4] Damiano Brigo and Fabio Mercurio. *Interest rate models-theory and practice: with smile, inflation and credit*. Springer Science & Business Media, 2007.

[5] Emanuele Casamassima. Pricing and hedging of a mortgage portfolio.

[6] Mark B. Garman and Steven W. Kohlhagen. Foreign currency option values. *Journal of international Money and Finance*, 2(3):231–237, 1983.

[7] Nicholas I.M. Gould and Sven Leyffer. An introduction to algorithms for nonlinear optimization. In *Frontiers in numerical analysis*, pages 109–197. Springer, 2003.

[8] Patrick S. Hagan and Graeme West. Methods for constructing a yield curve. *Wilmott Magazine, May*, pages 70–81, 2008.

[9] Koichi Harada and Eihachiro Nakamae. Application of the Bezier curve to data interpolation. *Computer-Aided Design*, 14(1):55–59, 1982.

[10] David Heath, Robert Jarrow, and Andrew Morton. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica: Journal of the Econometric Society*, pages 77–105, 1992.

[11] Andres Hernandez. Model calibration with neural networks. *Available at SSRN 2812140*, 2016.

[12] John Hull and Alan White. Pricing interest-rate-derivative securities. *The review of financial studies*, 3 (4):573–592, 1990.

[13] John Hull and Alan White. The general Hull–White model and supercalibration. *Financial Analysts Journal*, 57(6):34–43, 2001.

[14] Farshid Jamshidian. An exact bond option formula. 1989.

[15] John C. Cox, Jonathan E. Ingersoll Jr and Stephen A. Ross. A theory of the term structure of interest rates. In *Theory of valuation*, pages 129–164. World Scientific, 2005.

[16] El Mostafa Kalmoun and Luis Garrido. Trust region versus line search for computing the optical flow. *Multiscale Modeling & Simulation*, 11(3):890–906, 2013.

[17] Manolis I.A. Lourakis et al. A brief description of the Levenberg-Marquardt algorithm implemented by levmar. *Foundation of Research and Technology*, 4(1):1–6, 2005.

[18] Fabio Mercurio. Interest rates and the credit crunch: new formulas and market models. *Bloomberg portfolio research paper*, (2010-01), 2009.

[19] Fabio Mercurio. LIBOR market models with stochastic basis. *Bloomberg education and quantitative research paper*, (2010-05), 2010.

[20] Jorge J. Moré. The Levenberg-Marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer, 1978.

[21] Georgios Moysiadis. Calibration of the time-dependent mean reversion parameter in the Hull-White model using neural networks.

[22] Gereon Müller. Local vs. global optimization in syntax: A case study. 2004.

[23] Marek Musiela and Marek Rutkowski. Continuous-time term structure models: Forward measure approach. *Finance and Stochastics*, 1(4):261–291, 1997.

[24] Sebastien Gurrieri, Masaki Nakahayashi and Tony Wong. Calibration methods of Hull-White model. 2009.

[25] Cornelis W. Oosterlee and Lech A. Grzelak. *Mathematical Modeling and Computation in Finance*. 2020.

[26] Andrea Pallavicini and Marco Tarenghi. Interest-rate modeling with multiple yield curves. *arXiv preprint arXiv:1006.4767*, 2010.

[27] Vladimir Piterbarg. A multi-currency model with FX volatility skew. *Available at SSRN 685084*, 2005.

[28] Vladimir V Piterbarg. Cross-currency exotics: Smiling hybrids. *RISK-LONDON-RISK MAGAZINE LIMITED-*, 19(5):66, 2006.

[29] Mohsen Pourahmadi. Cholesky decompositions and estimation of a covariance matrix: orthogonality of variance–correlation parameters. *Biometrika*, 94(4):1006–1013, 2007.

[30] Michael J. D. Powell. A new algorithm for unconstrained optimization. In *Nonlinear programming*, pages 31–65. Elsevier, 1970.

[31] Michael J.D. Powell. A hybrid method for nonlinear equations. *Numerical methods for nonlinear algebraic equations*, 1970.

[32] Christoph M. Puetter and Stefano Renzitti. One-factor Hull-White model calibration for CVA-part i: Instrument selection with a kink. *Available at SSRN 3659430*, 2020.

[33] Christoph M. Puetter and Stefano Renzitti. One-factor Hull-White model calibration for CVA-part ii: Optimizing the mean reversion parameter. *Available at SSRN 3659443*, 2020.

[34] Ignacio Pena, Gonzalo Rubio and Gregorio Serna. Why do we smile? on the determinants of the implied volatility function. *Journal of Banking & Finance*, 23(8):1151–1179, 1999.

[35] Vincenzo Russo and Frank J Fabozzi. Calibrating short interest rate models in negative rate environments. *The Journal of Derivatives*, 24(4):80–92, 2017.

[36] Vincenzo Russo and Gabriele Torri. Calibration of one-factor and two-factor Hull–White models using swaptions. *Computational Management Science*, 16(1-2):275–295, 2019.

[37] Ya xiang Yuan. A review of trust region algorithms for optimization. In *Iciam*, volume 99, pages 271–282. Citeseer, 2000.

[38] Tjalling J. Ypma. Historical development of the Newton–Raphson method. *SIAM review*, 37(4):531–551, 1995.

[39] Fangfang Zhao, Zuoliang Xu, and Changjing Li. Calibration of implied volatility in generalized hull-white model. *J. Finance Account.*, 4(2):25–32, 2016.

[40] Jun Zhu. Multiple-curve valuation with one-factor Hull-White model. *Available at SSRN 2049507*, 2012.