Keresztfalvi_Balint_01514136

by Balint Zoltan Keresztfalvi

Submission date: 06-Sep-2022 01:49PM (UTC+0100)
Submission ID: 185745285
File name: Keresztfalvi_Balint_01514136.pdf (1.08M)
Word count: 21610
Character count: 91879

Imperial College London

Imperial College London

Department of Mathematics

Pricing of Bermudan Swaptions in the Cheyette Model

Author: Bálint Keresztfalvi (CID: 01514136)

A thesis submitted for the degree of

MSc in Mathematics and Finance, 2021-2022

Declaration

The work contained in this thesis is my own work unless otherwise stated.

Acknowledgements

Foremost, I would like to express my gratitude to my supervisor Thomas Cass at Imperial College London for his continuous support and guidance in the writing of this thesis.

I would like to thank Darryl Copsey at TD Securities for his guidance and the opportunity to work at TD on this project. I further thank the whole QMA team for all their support and for welcoming me in their team with open arms. Special thanks to Akshay Parmar for the best lunch spots and insightful conversations, to Jeffrey Wood for his tireless efforts towards the project and all the enlightening discussions throughout. Last but not least, I would like to express my heartfelt thanks to Bartek Loskot, who was indispensable in the writing of this thesis, for his guidance and endless support.

Finally I would like thank my parents and family for all their love and relentless support throughout my studies, I would not have made it this far without you.

Abstract

This thesis gives an overview of classical interest rate modelling with a specific focus on the continuously compounded rate. We provide two different approaches to derive the Hull-White model [1] and extend it to the family of Cheyette models. In particular, we consider the Linear Cheyette model which is a Markovian two state short rate model equipped with a linear local volatility function. It is often referred to as a Quasi-Gaussian model as when one of its parameters is close to zero then it behaves like a Gaussian model. We present the ways to calibrate these two models to price Bermudan Swaptions with a PDE approach. For the Cheyette model we investigate a technique called Markovian projection that reduces the number of state variables to one so we are able to price on a similar PDE as the Hull-White model.

Contents

List of Figures

List of Tables

Introduction

Interest rates have a fundamental role in financial modelling across all asset classes as they are indispensable when talking about any form of cash flows. In this thesis, we present a standard approach to understanding interest rates through stochastic calculus by first, considering a simpler, well-established model which we then take one step further. Our final aim is to evaluate an interest rate derivative called Bermudan Swaption. However, before presenting the final results, we revisit the classical theory of interest rates that we extend with our own model of choice.

In Chapter 1 we follow the book of Brigo and Mercurio [2] to lay the foundations for the mathematical analysis later. We define different type of rates and vanilla financial products related to them, including the subject of this thesis, the Bermudan Swaption. We also give a flavour of the theory of derivative pricing as a tool to use for our models.

Chapter 2 provides a detailed overview of the two main interest rate models considered in this thesis inspired by the book of Andersen and Piterbarg [3]. Firstly, we look at the Hull-White model [1] taking the classical approach in its derivation. Then we consider the general framework of Heath-Jarrow-Morton [4], the stem of the Cheyette model showing that it is a more general case of the Hull-White model and in fact the latter can also be derived this way.

The main contributions of this thesis are in Chapters 3 and 4. In Chapter 3, we first discuss how to calibrate each model to the market of vanilla European Swaptions which is desirable for pricing of Bermudans. In the case of the Hull-White model this can be done analytically for a single Swaption but it fails to capture multiple market prices. On the contrary, the Cheyette model is able to match two Swaptions and give an overall better fit to the market. This is further demonstrated when we look at implied volatility smiles, for Hull-White adjusting parameters can only produce a parallel shift but in the Linear Cheyette case we can also adjust the slope. We provide numerical results for market data showing the differences between the two models concluding that Cheyette is superior in this sense. We implement the methods described in TD Securities's C++ library with some of the details given in Chapter 5.

In Chapter 4 we apply a general dimension reduction technique to the Cheyette model which is inherently a two state model. This enables us to treat it as a single state model like Hull-White in certain applications and we use this projected form to price Bermudan Swaptions. We investigate two different methods of projection and analyse them in a numerical setting as they will be used within numerical PDE solver. To check the accuracy of our results we run Monte Carlo simulations and use them as the benchmark for the projection.

Finally, in Chapter 6 we present numerical results for European and Bermudan Swaptions prices using a combination of methods outlined in the earlier chapters.

Chapter 1

Interest Rate Products

Interest rates are one of the most important concepts within finance. They are essential to value most financial derivatives as well as have major economic relevance for businesses and ordinary people alike. In principle, rates determine the price of borrowing or lending money which can occur when for example, someone enters a mortgage agreement, purchases a treasury bond or simply deposits money in the bank. In this chapter we are going to describe the variety of different rates used in the world of quantitative finance and some financial securities that build upon these. We will revisit the classical theory, not accounting for multiple curves, credit risk or liquidity risk as it is done in the early chapters of Brigo and Mercurio [2] and as Damiano Brigo lectured it at Imperial College London in 2021-2022 [5].

1.1 Different Rates and Vanilla Products

1.1.1 Continuously Compounded Rate

Starting from a mathematician's perspective we consider the continuously compounded rate or simply short rate. Even though it is not observable on the market it can still be viewed as one of the most fundamental rates due to its importance in modelling as we will see in Chapter 2.
Let the value of the money market account or bank account be denoted by $B(t)$ at time t, this measures the value of an investment of B_0 pounds deposited in the bank at time t after accruing interest.

Definition 1.1.1 (Short rate)**.** The rate applied to the account at time t over an infinitesimal period is the short rate, denoted by $r(t)$, with the following equations for the bank account:

$$\begin{aligned} dB(t) &= r(t)B(t)dt \\ B(t) &= B_0 \exp\left(\int_0^t r(u)du\right) \end{aligned}$$

The bank account is special as it is used to discount any future cashflows. The reason this is needed is because receiving £1 a week from now is not worth as much as receiving £1 today since in theory the pound today can be put into a bank to earn interest and turn into more than £1 a week from now.
In fact it is true that evaluating any future cashflows can be done by taking expectation of the discounted value under an appropriate probability measure. In the case of the bank account this is called the risk neutral measure, as the investment in the bank is thought to be risk free and it is denoted by $\mathbb{Q}$. In this setting $B(t)$ is referred to as the numeraire under $\mathbb{Q}$. For example, take a single payoff H_T at a future time T, then the appropriate discounting and therefore the value of the payoff at time t is given by

$$V(t) = \mathbb{E}^{\mathbb{Q}}\left[\frac{B(t)}{B(T)}H_T\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(u)du\right)H_T\middle|\mathcal{F}_t\right], \tag{1.1.1}$$

where $\mathcal{F}_t$ denotes the information available at time t as given by Brigo in [5].

1.1.2 Zero Coupon Bond

We now consider the simplest possible payoff $H_T = 1$ from the above, describing an instrument that you can buy today and the issuer agrees to pay you £1 at a future date T. This is called a zero coupon bond as it does not bear any coupon payments, only a single transaction at maturity T. Using the pricing formula (1.1.1) we have the value of the zero coupon bond purchased at time t

$$P(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(u)du\right)\middle|\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{Q}}\left[D(t,T)|\mathcal{F}_t\right], \tag{1.1.2}$$

where $D(t,T) = \frac{B(t)}{B(T)}$ is the discount factor applied to future time T at the present time t. From our earlier analogy we see that this price should be less than £1 if interest rates stay positive, however when $t = T$ we should have $P(T,T) = 1$ which is satisfied. Now with the help of $P(t,T)$ we can define rates that are not compounded continuously but instead are applied over a time period.

1.1.3 Simply Compounded Rate

The simply compounded rate or spot linear rate or often just LIBOR[1] (London InterBank Offer Rate), is one of the most common rates quoted on the market and it serves as an underlying to many interest rate derivatives. Instead of being continuously compounded like the short rate, it is only compounded once on the notional. It is defined as follows.

Definition 1.1.2 (Spot Linear Rate)**.** The Spot Linear Rate is the linearly applied constant rate such that an investment of $P(t,T)$ at time t grows to 1 by its maturity. In mathematical terms this means that

$$\begin{aligned} P(t,T)(1 + (T-t)L(t,T)) &= 1, \\ L(t,T) &= \frac{1 - P(t,T)}{(T-t)P(t,T)} \end{aligned} \tag{1.1.3}$$

This is related to the short rate in a sense that when T approaches t from above, the two become very close and eventually equal in the limit:

$$r(t) = \lim_{T \to t^+} L(t,T) \approx L(t, t+\epsilon).$$

It is also not hard to recover $L(t,T)$ from short rates by putting together equations (1.1.2) and (1.1.3).

1.1.4 Forward Rate Agreement and Forward Rates

One of the simplest derivatives that have the spot linear rate as an underlying is a forward rate agreement or commonly FRA.

Definition 1.1.3 (Forward Rate Agreement)**.** A forward rate agreement is a contract between two parties who exchange a floating and a fixed leg payment between future times $T_1 < T_2$. The payer of the floating leg agrees to pay the linear rate $L(T_1, T_2)$ for the time period and the payer of the fixed leg agrees to do the same with a previously agreed rate K .

$$\begin{array}{rccl} \text{Party A} \longrightarrow & (T_2 - T_1)K & \longrightarrow & \\ \longleftarrow & (T_2 - T_1)L(T_1, T_2) & \longleftarrow & \text{Party B} \end{array}$$

The contract is called a Receiver FRA if the holder pays the floating and receives the fixed leg, otherwise it is a Payer FRA and has the payments the other way around.

[1] In the recent years the industry have stopped using LIBOR and transitioned onto SOFR (Secured Overnight Financing Rate) but this change doesn't our derivations much as they are both just linear rates

It is shown by Brigo in [5] that the price of a receiver FRA at time t is:

$$\text{FRA}(t; T_1, T_2, K) = P(t, T_1)(T_2 - T_1)K - P(t, T_1) + P(t, T_2) \tag{1.1.4}$$

where the proof relies only on risk-neutral expectations and is otherwise model independent. Notice that for a fixed time period the price of the contract only depends on K in a linear fashion. This means that there is a choice of K that the contract is fair for both parties and has value 0 at time t. We can define:

Definition 1.1.4 (Forward Linear Rate)**.** The forward linear rate at time t for the time interval $[T_1, T_2]$ is the value K such that the price of the receiver FRA contract (1.1.4) is 0 at time t.

$$F(t; T_1, T_2) := \frac{1}{T_2 - T_1}\left(\frac{P(t, T_1)}{P(t, T_2)} - 1\right).$$

We note that it is the forward rate in a sense that when taken at time T_1 it reconciles to the spot linear rate:

$$F(T_1; T_1, T_2) = \frac{1}{T_2 - T_1}\left(\frac{P(T_1, T_1)}{P(T_2, T_2)} - 1\right) = L(T_1, T_2),$$

as $P(T_1, T_1) = 1$. Similarly to before with the relationship between the short rate and the spot linear rate, we can define an analogous instantaneous forward rate.

Definition 1.1.5 (Instantaneous Forward Rate)**.** For times $t < T$ let

$$f(t, T) = -\frac{\partial}{\partial T} \ln P(t, T) \approx F(t; T, T + \epsilon),$$

be the instantaneous forward rate.

One can invert the relationship in the above definition to get another formula for the price of the zero coupon bond

$$P(t, T) = \exp\left(-\int_t^T f(t, u)du\right), \tag{1.1.5}$$

which will be useful for us later. Similar to the short rate, this is also not quoted on the market but is used when it comes to modelling interest rates in general as we will see in the Heath-Jarrow-Morton framework in Section 2.2.

1.2 Interest Rate Products

In this section we are going to have a look at some more complicated interest rate derivatives leading up to understanding the main subject of this thesis: the Bermudan Swaption. We first define a regular swap.

Definition 1.2.1 (Interest Rate Swap)**.** An Interest Rate Swap (IRS) or simply a swap is an agreement between two parties that for a fixed tenor structure $T_0 < T_1 < \ldots < T_N$ they exchange payments at each T_n by a fixed and a floating leg:

$$\begin{array}{lll} \text{Party A} \longrightarrow & (T_n - T_{n-1})K & \longrightarrow \\ \longleftarrow & (T_n - T_{n-1})L(T_{n-1}, T_n) & \longleftarrow \text{Party B} \end{array}$$

for an agreed fixed rate K. This is called a payer IRS from perspective of Party A if they pay the fixed rate K and it is a receiver IRS if it is the other way around.

For example, if a company wishes to issue a bond with a fixed interest rate K, then they can enter into a receiver IRS with a bank to receive these fixed payments while only having to pay the current linear rate in exchange at each T_i.

Notice that each payment date in an IRS is identical to a forward rate agreement in Definition 1.1.3. This means that the above can be replicated by a bundle of FRAs and hence can be valued accordingly. We then have the price of a receiver interest rate swap

$$\text{IRS}(t, [T_0, \ldots, T_N], K) = \sum_{i=0}^{N-1} (T_{i+1} - T_i) K P(t, T_i) - P(t, T_0) + P(t, T_N) \tag{1.2.1}$$

Similarly as earlier, one can have a value for K at time t that makes the contract fair which defines the forward Swap rate.

Definition 1.2.2 (Forward Swap Rate)**.** The Forward Swap rate for a fixed tenor structure $T_0 < T_1 < \ldots < T_N$ is the fixed rate K at time t such that the value of the IRS with the same parameters in (1.2.1) is 0.

$$S_{0,N}(t) = \frac{P(t, T_0) - P(t, T_N)}{\sum_{i=0}^{N-1} \tau_i P(t, T_{i+1})}, \tag{1.2.2}$$

where $\tau_i = T_{i+1} - T_i$.

We will often just refer to this as the Swap rate. Looking at this formula, we can think of the denominator as a type of discount factor for the Swap rate. This will be important when we want to determine the price of a Swaption.

Definition 1.2.3 (Annuity)**.** For a fixed tenor structure $T_0 < T_1 < \ldots < T_N$ the annuity is given by

$$A_{0,N}(t) = \sum_{i=0}^{N-1} \tau_i P(t, T_{i+1}), \tag{1.2.3}$$

which will serve as an alternative numeraire to the money market account later.

Now we turn to Swaptions, which is short for Swap option and it has the Swap rate as its underlying quantity.
Options are common financial products throughout all asset class. In general, they give the holder the right but not the obligation to purchase the underlying asset. These products are popular because there is no additional downside once the option is purchased the worst case scenario is that the optionality is unused. However, when the option is exercised in many cases it can generate theoretically unbounded profit. In a sense this makes it similar to gambling; after placing the initial bet, the outcome can only be positive.

Definition 1.2.4 (European Swaption)**.** A European Swaption gives the holder the right but not the obligation to enter into a interest rate swap at a future date, which often coincides with the exercise date or maturity of the option. The payoff is given by

$$V(T_0) = (S_{0,N}(T_0) - K)^+ \sum_{i=0}^{N-1} \tau_i P(T_0, T_{i+1})$$

as the holder can exercise the option to enter a payer IRS with fixed rate K and at the same time enter a receiver IRS with fixed rate $S(T_0)$ to realise the profit above.

We can rewrite the above by substituting in (1.2.2) and moving the annuity from (1.2.3) inside to get

$$V(T_0) = \left(1 - P(T_0, T_N) - K \sum_{i=0}^{N-1} \tau_i P(T_0, T_{i+1})\right)^+ \tag{1.2.4}$$

The European in the name refers to the exercise type of the option. The simplest case is when there is a singular date to exercise which is the maturity. Suppose now that instead of just one, there is a series of dates when the holder can exercise the option, which are agreed at the time

the option is purchased. This is a Bermudan option. Finally, if the option can be exercised at any time[2] it is called an American option.
Intuitively it is clear that a Bermudan should be worth more than a European, and an American should be more valuable than both, given the additional flexibility to exercise when it is profitable to do so. However, this makes the valuation more difficult as even though there is more opportunity to exercise these options it is no longer clear when this should occur. For example, for a Bermudan it might be that exercising today would realise a profit, but if we have a strong view that on the next exercise date the profit is likely to be bigger, then it might be better to hold on to the option until then. This is called the exercise boundary which defines when to exercise the option.

Definition 1.2.5 (Bermudan Swaption)**.** Fix a tenor structure $T_0 < T_1 < \ldots < T_N$ and a rate K to define an IRS. A Bermudan Swaption gives the right but not the obligation to enter this IRS at each payment date T_i with exercise value

$$\begin{aligned} V_{\text{exercise}}(T_i) &= (S_{n,N}(T_n) - K)^+ \sum_{i=n}^{N-1} \tau_i P(T_n, T_{i+1}) \\ &= \left(1 - P(T_n, T_N) - K \sum_{i=n}^{N-1} \tau_n P(T_n, T_{i+1})\right)^+, \end{aligned} \tag{1.2.5}$$

and similarly it also has hold value $V_{\text{hold}}(T_i)$ which is the value of the Bermudan Swaption for tenor structure $(T_j)_{j=n+1}^N$.

Bermudan Swaptions are by far the most commonly traded exotic securities. The reason behind is the relatively high demand because of the flexibility it provides to perform hedging of large fixed income portfolios. For example, imagine a firm that wants to raise capital quickly so it issues a callable bond[3] and simultaneously enters into a swap to convert the fixed coupon payments to floating. Once the company calls back the bond, it no longer needs the swap, so it makes sense for them to buy a Bermudan Swaption when issuing the bonds, to exercise at the same time the bonds are redeemed, to close out their original swap.

1.3 Pricing of Interest Rate Products

In this section we are going to present some techniques to price the products outlined above. In general, to determine the value of a contract one needs to have a model that describes the behaviour of the underlying quantities. Most commonly this is done by considering stochastic processes and deriving a probability measure where these products depending on the processes can be valued. One of the ways this valuation is done is already presented in (1.1.1) and is derived from the following fundamental theorem given in various textbooks but we lift it from the lecture notes of Alex Tse [6]:

Theorem 1.3.1 (Fundamental Theorem of Asset Pricing)**.** *In a complete arbitrage-free market with an equivalent martingale measure $\mathbb{Q}$ and corresponding numeraire B the fair-value of a contingent claim $V(T)$ at present time t is*

$$V(t) = B(t)\mathbb{E}^{\mathbb{Q}}\left[\frac{V(T)}{B(T)}\middle|\mathcal{F}_t\right] \tag{1.3.1}$$

where $\mathcal{F}_t$ represents the information available at time t.

The main idea behind the proof is that the discounted value of any tradeable product is a martingale, and the formula above follows from the martingale property. Furthermore, this holds true for any numeraire pair $(N, \mathbb{Q}^N)$ where N is the price process of some asset and $\mathbb{Q}^N$ is the corresponding equivalent martingale measure.
For example, in case of the Swap rate in (1.2.2) we may choose the annuity in Definition 1.2.3 as the numeraire. Then under the corresponding annuity measure denoted $\mathbb{Q}^A$ the Swap rate is a martingale. We will see this method in action later in Section 2.5.

[2] In a mathematical setting we consider this to be the case but in reality there are restrictions like business days, business hours etc.

[3] A callable bond is a bond that the issuer can terminate before maturity by settling their debt

Often in (1.3.1) the payoff $V(T)$ and the numeraire are both a function of an underlying X which we model as a stochastic process. The most general type of process we are going to deal with under the risk neutral measure is the following

Definition 1.3.2 (General stochastic differential equation)**.** Let $(X_t)_{t\geq 0}$ be a stochastic process governed by the following equation

$$dX(t) = \mu(t,\omega)dt + \sigma(t,\omega)dW_t^{\mathbb{Q}}, \tag{1.3.2}$$

where $W^{\mathbb{Q}}$ is a Wiener-process generating a filtration $\mathcal{F}_t$ with $\mu(t,\omega)$ and $\sigma(t,\omega)$ are referred to as the drift and diffusion coefficients respectively which are both $\mathcal{F}_t$ adapted denoted by the ω dependence.

The final statement means that while the values $\mu(t)$ or $\sigma(t)$ might not be known at time 0, once the process W gets to time t, based on its trajectory, these two values are determined. A simplification of the above is when both the drift and diffusion term only depend on the current value of the process X instead of any past trajectory, so we get the Markovian dynamics

$$dX(t) = \mu(t,X(t))dt + \sigma(t,X(t))dW_t^{\mathbb{Q}}. \tag{1.3.3}$$

Assuming these dynamics for the underlying we can derive a different approach to evaluate the expectation in (1.3.1).

Theorem 1.3.3 (Feynmann-Kac)**.** *Given a stochastic process of the form in (1.3.3) then*

$$V(t,x) = \mathbb{E}^{\mathbb{Q}}\left[e^{-\int_t^T r(u)du} f(X(T))\middle| X(t) = x\right]$$

is given by the solution to the partial differential equation

$$\begin{aligned} \frac{\partial V}{\partial t}(t,x) + \mu(t,x)\frac{\partial V}{\partial x}(t,x) + \frac{1}{2}\sigma^2(t,x)\frac{\partial^2 V}{\partial x^2}(t,x) &= r(t)V(t,x), \\ V(T,x) &= f(x). \end{aligned} \tag{1.3.4}$$

The above PDE can sometimes be solved explicitly, for example getting the famous Black-Scholes formula for a European Call option but it is also useful for more complex payoffs when there is no analytical solution but a numerical method can be applied. We will later use this to applied to Bermudan Swaptions, solving the PDE via an implicit finite difference scheme which is effectively applying a backwards recursion on a discretized grid for the payoff. The main idea is that for each backward step we determine if the hold value V_{hold} or the exercise value V_{exercise} of the option given in Definition 1.2.5 is larger and then we continue the recursion with the larger one.

Chapter 2

Interest Rate Modelling

In this chapter, we are going to talk about modelling interest rates as they evolve in time. Generally, all of the models described below will treat the short rate or the instantaneous forward rate described in Section 1.1 as a stochastic process over the risk neutral measure $\mathbb{Q}$ given in (1.3.2). This of course makes quite a few assumptions that might not be true, but in return we can get analytical prices for some of the products described earlier. Crucially, we will have analytical formulas for the zero coupon bond $P(t,T)$ which is a building block for many interest rate derivatives and being able to describe its dynamics gives more tractability to the models themselves. We first start by introducing one of the classical models of the short rate.

2.1 Hull-White Model

The Hull-White model was first introduced by John Hull and Alan White in [1] to describe the dynamics of the short rate based on an Ornstein–Uhlenbeck process. It is one of the most popular short rate models to this day due to it is simplicity, analytical tractability and numerical capabilities. Compared to preceding alternatives, it was able to fully capture the initial term structure while retaining easy calibration. As any model, it has its drawbacks, for example when it comes to describing the behaviour of the term structure over time or looking at volatility smiles observed on the swaption market but overall it serves as a great benchmark for any more complicated short rate model. In this section we derive the model the way it was first presented. However, later we look at the more general Heath-Jarrow-Morton [4] framework which can be formed in a way that it produces the Hull-White model.

2.1.1 General Dynamics

As mentioned above, Hull-White models the short rate process $(r_t)_{t\geq 0}$ as an Ornstein-Uhlenbeck stochastic difference equation.

Definition 2.1.1 (Hull-White short rate)**.** Define the short rate to satisfy the stochastic differential equation

$$dr(t) = (\theta(t) - \kappa(t)r(t))dt + \sigma(t)dW_t^{\mathbb{Q}}, \tag{2.1.1}$$

with deterministic parameters $\theta(t)$, $\kappa(t)$ and $\sigma(t)$.

It is often desired to eliminate $\theta(t)$ as it introduces a purely deterministic shift in the short rate. Instead of dealing with the above equation (2.1.1), we follow the idea of Brigo in [2, Second edition, Section 3.3.1, page 74] and we let

$$\begin{aligned} r(t) &:= x(t) + \varphi(t), \\ dx(t) &:= -\kappa(t)x(t)dt + \sigma(t)dW_t^{\mathbb{Q}}, \qquad x(0) := 0 \qquad (2.1.2) \\ d\varphi(t) &:= (\theta(t) - \kappa(t)\varphi(t))dt, \qquad \varphi(0) := r(0) \qquad (2.1.3) \end{aligned}$$

In this setting we have an ODE for $\varphi(t)$ and a simpler SDE for our state variable $x(t)$ which determines the short rate at time t. The main advantage of this representation however comes at

calibration when we are aiming to fit to the initial term structure of $P(0,T)$ for $T \geq 0$. Solving (2.1.2) gives the following solution if $x(s)$ is given at time $0 \leq s \leq t$:

$$x(t) = x(s)e^{-\int_s^t \kappa(u)du} + \int_s^t \sigma(u)e^{-\int_u^t \kappa(v)dv} dW^{\mathbb{Q}}(u).$$

From this expression it is clear that $x(t)$ is normally distributed which also means that the short rate is normally distributed hence it is often referred to as a Gaussian model and lots of its analytical tractability stems from this property.
Similarly, solving for $\varphi(t)$ in (2.1.3) we get:

$$\varphi(t) = \varphi(s)e^{-\int_s^t \kappa(u)du} + \int_s^t \theta(u)e^{-\int_u^t \kappa(v)dv} du, \quad 0 \leq s \leq t$$

Finally for the short rate $r(t)$ given information available at time s as above, we have:

$$\begin{aligned} r(t) &= \varphi(t) + x(t) \\ &= \varphi(t) + x(s)e^{-\int_s^t \kappa(u)du} + \int_s^t \sigma(u)e^{-\int_u^t \kappa(v)dv} dW^{\mathbb{Q}}(u) \\ &= r(s)e^{-\int_s^t \kappa(u)du} + \varphi(t) - \varphi(s)e^{-\int_s^t \kappa(u)du} + \int_s^t \sigma(u)e^{-\int_u^t \kappa(v)dv} dW^{\mathbb{Q}}(u) \end{aligned} \tag{2.1.4}$$

So clearly the short rate is also normally distributed. Some more details on this derivation are given in Appendix A.1.1.

2.1.2 Bond Reconstruction Formula

From Section 1.3 we know that under the risk neutral measure the price of the conditional zero coupon bond for time interval $[t,T]$ conditioned on the short rate at $t \leq T$ is

$$P(t,T) = \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(u)du\right)\middle|\mathcal{F}_t\right]. \tag{2.1.5}$$

In order to evaluate this expectation, we need to understand the integral of the short rate. It is relatively straightforward to show that $\int_t^T r(u)du$ is normally distributed as well, the details are shown in Appendix A.1.1, we get that

$$\int_t^T r(u)du \sim \mathcal{N}\left(x(t)G(t,T) + \int_t^T \varphi(u)du, V(t,T)\right),$$

where

$$\begin{aligned} G(t,T) &= \int_t^T e^{-\int_t^u \kappa(v)dv} du \\ V(t,T) &= \int_t^T \sigma^2(u)G(u,T)^2 du \end{aligned}$$

Now to price the conditional zero coupon bond we invoke the following lemma.

Lemma 2.1.2 (Mean of log-normal random variable)**.** *Given a normally distributed random variable $X \sim \mathcal{N}(\mu, \sigma^2)$ over some probability space $(\Omega, \mathcal{F}, \mathbb{P})$ we have that the expectation of e^X over $\mathbb{P}$ is*

$$\mathbb{E}^{\mathbb{P}}\left[e^X\right] = e^{\mu + \frac{1}{2}\sigma^2}.$$

Using the above lemma on (2.1.5) we get:

$$P(t,T) = \exp\left(-\int_t^T \varphi(u)du\right)\exp\left(-x(t)G(t,T) + \frac{1}{2}V(t,T)\right)$$

Finally we want to make sure that we fully capture the initial term structure of $P(0,T)$ by choosing $\varphi(t)$ accordingly. Following the above equation by setting $t=0$ we can see that

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

as $x(0)=0$ and hence

$$\exp\left(-\int_t^T\varphi(u)du\right)=\frac{\exp\left(-\int_0^T\varphi(u)du\right)}{\exp\left(-\int_0^t\varphi(u)du\right)}=\frac{P(0,T)}{P(0,t)}\exp\left(-\frac{1}{2}[V(0,T)-V(0,t)]\right).$$

Substituting this back we get the Hull-White conditional zero coupon bond price:

$$P(t,T,x(t))=\frac{P(0,T)}{P(0,t)}\exp\left(-x(t)G(t,T)+\frac{1}{2}[V(t,T)-V(0,T)+V(0,t)]\right) \tag{2.1.6}$$

It is also possible to reverse engineer $\varphi(t)$ and even $\theta(t)$ from the bond price (2.1.6) to get the expressions

$$\varphi(t)=f^M(0,t)+\int_0^t\sigma^2(u)e^{-\int_u^t\kappa(v)dv}G(u,t)du, \tag{2.1.7}$$

$$\theta(t)=\frac{\partial f^M(0,t)}{\partial t}+\kappa(t)f^M(0,t)+\frac{1}{2}\left(\kappa(t)\frac{\partial V}{\partial t}(0,t)+\frac{\partial^2V}{\partial t^2}(0,t)\right)$$

$$\alpha(t)=\varphi(t)-f^M(0,t)=\int_0^t\sigma^2(u)e^{-\int_u^t\kappa(v)dv}G(u,t)du \tag{2.1.8}$$

where $f^M(0,t)$ is the market implied instantaneous forward rate and define $\alpha(t)$ to be the not market implied part of the deterministic shift so that the short rate $r(t)$ is given by

$$r(t)=x(t)+f^M(0,t)+\alpha(t). \tag{2.1.9}$$

The details of the derivations in (2.1.3) are given in Appendix A.1.2. This also hints at a closer relationship of fitting perfectly onto the initial term structure and the forward rates implied by the market at time 0. In the following Section 2.2 we are going to examine modelling the forward rates in general and then stepping down to short rates by the identity $f(t,t)=r(t)$.

2.1.3 Pricing European Swaption

We now turn to pricing a European Swaption in the Hull-White model as this will be the main product the model calibrates to in order to evaluate Bermudan Swaptions. The payoff in this model is

$$V(T_0)=\left(1-P(T_0,T_N,x(T_0))-K\sum_{i=0}^{N-1}\tau_iP(T_0,T_{i+1},x(T_0))\right)^+ \tag{2.1.10}$$

from (1.2.4). We can now use Jamshidian's trick given in [7] to decompose the above payoff, using that the bond price $P(t,T,x(t))$ in (2.1.6) is strictly monotone decreasing in x. One can find x^* such that

$$P(T_0,T_N,x^*)+K\sum_{i=0}^{N-1}\tau_iP(T_0,T_{i+1},x^*)=1,$$

to obtain different synthetic strikes

$$K_i=P(T_0,T_i,x^*).$$

Substituting this back into (2.1.10) yields

$$V(T_0)=\left(K_N+\sum_{i=0}^{N-1}\tau_iK_{i+1}-P(T_0,T_N,x(T_0))-K\sum_{i=0}^{N-1}\tau_iP(T_0,T_{i+1},x(T_0))\right)\mathbb{1}_{\{x(T_0)>x^*\}},$$

using monotonic property of $P(t,T,x(t))$. Then we can decompose the payoff and convert the indicator $\mathbf{1}_{\{x(T_0)>x^*\}}$ to a positive part to get

$$V(T_0) = (K_N - P(T_0,T_N,x(T_0)))^+ + K\sum_{i=0}^{N-1}\tau_i\,(K_{i+1} - P(T_0,T_{i+1},x(T_0)))^+ \tag{2.1.11}$$

We realise this as a sum of vanilla European Put Options on different maturities. In the Hull-White model we saw that bond prices are log-normally distributed so we can evaluate the price of these according to a variation of Black-Scholes formula. We quote a result inspired by Andersen and Piterbarg [8, Section 4.5, Proposition 4.5.1, page 185]

Proposition 2.1.3 (Put Option on Zero-Coupon Bond)**.** *Consider the Put option with payoff*

$$V(T_0) = (K - P(T_0,T_1))^+$$

Then in a Gaussian model we have

$$V(t) = P(t,T_0)K\Phi(-d_-) - P(t,T_1)\Phi(-d_+) \tag{2.1.12}$$

where

$$d_\pm = \frac{\ln\left(P(t,T_1)/(KP(t,T_0))\right) \pm v/2}{\sqrt{v}},$$
$$v = \int_t^T \left(\sigma_P(u,T_1) - \sigma_P(u,T_0)\right)^2 du$$

where $\sigma_P(t,T)$ is the volatility of the log-normal bond dynamics $P(t,T)$ under the risk neutral measure $\mathbb{Q}$ and $\Phi(\cdot)$ is the cumulative normal distribution.

To apply the above we only need to determine $\sigma_P(t,T)$. In order to do this we may apply Ito's lemma to (2.1.6) to calculate $d\ln P(t,T)$:

$$d\ln P(t,T) = (\ldots)dt - G(t,T)\sigma(t)dW_t^{\mathbb{Q}}$$
$$\sigma_P^2(t,T) = G^2(t,T)\sigma^2(t)$$

So finally we have

Theorem 2.1.4 (European Swaption Price in Hull-White)**.** *For a fixed term structure $T_0 < T_1 < \ldots < T_N$ and a rate K the price of the European Swaption at time $0 \le t < T_0$ is*

$$V(t) = \mathbf{ZCBP}(t,T_0,T_N,K_N) + K\sum_{i=0}^{N-1}\tau_i\mathbf{ZCBP}(t,T_0,T_{i+1},K_{i+1}) \tag{2.1.13}$$

where $\mathbf{ZCBP}(t,T,T',K)$ is the price of the Put option of the zero coupon bond from (2.1.12) with

$$v = \int_t^T \sigma^2(u)\left(G(u,T) - G(u,T')\right)^2 du.$$

Proof. The proof is straightforward by putting together (2.1.11) and Proposition 2.1.3 with the formula for $\sigma_P(t,T)$. □

2.2 Heath-Jarrow-Morton Framework

The model considered in this section was first proposed by David Heath, Robert Jarrow, and Andrew Morton in [4] uniting existing interest rate models into one. It is often referred to as the Heath-Jarrow-Morton framework or simply for short HJM framework. In its full generality it is not practical to use within the industry due to the infinite number of stochastic processes at hand and hence a lack of analytical results for pricing. However, it provides a solid foundation to many models that are derived by introducing restrictions on certain parts improving its tractability.

2.2.1 General Dynamics

Unlike the Hull-White model in Section 2.1 we aim to describe the instantaneous forward rate processes $f(t,T)$ for all times $0 \leq t \leq T \leq \tau$ where τ is just some upper bound on the time horizon we consider for trading. This will inherently mean modelling infinitely many processes $f(\cdot,T)$ for each $T \in [0,\tau]$. In the HJM framework this is done by a general stochastic differential equation shown in (1.3.2).

Definition 2.2.1 (Instantaneous Forward Rate process)**.** We define the instantaneous forward rate $f(\cdot,T)$ to follow the stochastic differential equation

$$df(t,T) = \mu_f(t,T,\omega)dt + \sigma_f(t,T,\omega)dW_t^{\mathbb{Q}}, \tag{2.2.1}$$

where σ_f referred to as forward rate volatility.

From now on we will drop ω to ease notation. As shown in [4], imposing a no-arbitrage condition to recover the initial term structure of $P(0,T)$ gives the following lemma

Lemma 2.2.2 (No Arbitrage in Heath-Jarrow-Morton)**.** *Suppose that the economy modelling forward rates by Definition 2.2.1 is arbitrage-free then the drift*

$$\mu_f(t,T) = \sigma_f(t,T)\left(\int_t^T \sigma_f(t,u)du\right)$$

must be satisfied and then the dynamics become

$$df(t,T) = \sigma_f(t,T)\left(\int_t^T \sigma_f(t,u)du\right)dt + \sigma_f(t,T)dW_t^{\mathbb{Q}}. \tag{2.2.2}$$

We note that at this point we considered a slightly less general setting than that of in the original paper with only a single Brownian motion $W^{\mathbb{Q}}$.
From this formulation it is straightforward to derive conditional zero coupon bond prices by using the identity (1.1.5) between instantaneous forward rate and bond price shown in Section 1.1 as we just need to integrate (2.2.2) twice.

2.3 Separable Volatility in HJM

We consider a restriction on the general framework above to gain some analytical tractability, mainly regarding the zero coupon bond price $P(t,T)$ which we saw derived in the Hull-White model in (2.1.6). This will give us a Separable HJM model which is often also referred to as the general Cheyette model or Quasi-Gaussian model. Originally introduced by Oren Cheyette in [9] by a specific volatility decomposition and it is revisited in [3, Chapter 13] by Andersen and Piterbarg. We will follow the methodology of the latter in our derivations showing also why this is often considered an, in a sense natural, extension of the Hull-White model.

2.3.1 General Dynamics

As advertised above we start with the Heath-Jarrow-Morton framework and impose the following restriction on the general volatility function.

Definition 2.3.1 (Separable Volatility)**.** Suppose the forward rate volatility in the Heath-Jarrow-Morton framework is separable as

$$\sigma_f(t,T,\omega) = g(t,\omega)h(T), \tag{2.3.1}$$

for some deterministic function h and $g(t,\omega)$ which is $\mathcal{F}_t$-adapted.

Again we will omit ω from the formulas to ease notation but later in Section 2.5 we will consider local volatility, which means that ω dependence will be realised as dependence on the current state of $f(t,T)$.

With this restriction we can write the instantaneous forward rate as

$$f(t,T) = f(0,T) + h(T)\int_0^t g^2(s)\left(\int_s^T h(u)du\right)ds + h(T)\int_0^t g(s)dW_s^{\mathbb{Q}}. \tag{2.3.2}$$

Now we make the following simplifications. Let $t = T$ so that $f(t,t) = r(t)$ will give us the short rate dynamics.

Definition 2.3.2 (Separable HJM decomposition)**.** Define state variable x governing the uncertain part of the short rate as follows

$$x(t) = r(t) - f(0,t), \quad x(0) = 0, \tag{2.3.3}$$

Later we will insist on $f(0,t) = f^M(0,t)$ in order to capture the initial term structure. This way one gets a stochastic differential equation for $x(t)$:

$$dx(t) = \left(\frac{h'(t)}{h(t)}x(t) + h^2(t)\int_0^t g^2(u)du\right)dt + h(t)g(t)dW_t^{\mathbb{Q}}, \quad x(0) = 0. \tag{2.3.4}$$

To simplify notation a bit we introduce another state variable $y(t)$ as follows

$$y(t) = h^2(t)\int_0^t g^2(u)du = \int_0^t \sigma_r^2(u)\frac{h^2(u)}{h^2(t)}du \tag{2.3.5}$$

which is stochastic but is locally deterministic and where $\sigma_r(t) = \sigma_f(t,t) = g(t)h(t)$ is the short rate volatility. We finally arrive at

Theorem 2.3.3 (Separable HJM or Cheyette Dynamics)**.** *Defining two state variables x as in (2.3.3) and y in (2.3.5) we get dynamics*

$$\begin{aligned} dx(t) &= (y(t) - \kappa(t)x(t))dt + \sigma_r(t)dW_t^{\mathbb{Q}} \\ dy(t) &= \left(\sigma_r^2(t) - 2\kappa(t)y(t)\right)dt \\ x(0) &= y(0) = 0, \end{aligned} \tag{2.3.6}$$

where $\kappa(t)$ is defined as follows

$$\begin{aligned} \kappa(t) &= -\frac{h'(t)}{h(t)}, \\ h(t) &= \exp\left(-\int_0^t \kappa(u)du\right) \end{aligned}$$

Proof. All the details of the proof are found in Appendix A.2. □

The formulation in (2.3.6) is why the Cheyette model is said to be a one and a half factor model as even though it is governed by two state variables, y is an auxiliary process which is locally deterministic.
It is also important to note at this point that choosing a fully deterministic volatility function $\sigma_r(t)$ as in Section 2.1 will reduce the Cheyette model into a Hull-White model also showing that that is also a type covered in the HJM framework with separable deterministic volatilities.

Proposition 2.3.4. *Define the separable volatility function in (2.3.1) with*

$$g(t) = \sigma(t)\exp\left(\int_0^t \kappa(u)du\right).$$

Then the dynamics in (2.3.6) give exactly the same short rate dynamics as the model in (2.1.1).

Proof. The proof is given in Appendix A.3. □

2.3.2 Bond Reconstruction Formula

Continuing with (2.3.2) in order to calculate the price of the conditional zero coupon bond, we need to express the instantaneous forward rate in terms of our state variables x and y

$$f(t,T) = f(0,T) + \frac{h(T)}{h(t)}\left(x(t) + y(t)G(t,T)\right) \tag{2.3.7}$$

where we use notation of Piterbarg [3] for $G(t,T)$:

$$G(t,T) = \frac{\int_t^T h(u)du}{h(t)}$$

Then integrating (2.3.7) in accordance with the identity (1.1.5) gives the following conditional zero coupon bond price in the Cheyette model conditioning on state variables $x(t)$ and $y(t)$ at $t \leq T$:

$$P(t,T,x,y) = \frac{P(0,T)}{P(0,t)} \exp\left(-G(t,T)x(t) - \frac{1}{2}y(t)G^2(t,T)\right) \tag{2.3.8}$$

The details of the calculations are again shown in Appendix A.2. Note that this formula is very similar to the bond price derived in (2.1.6) for the Hull-White model with the $y(t)$ term replaced. However, it is crucial that the formulation for Hull-White we had in (2.1.2) is **not** the same x as in the general Cheyette case, simply because of how the short rate is represented by the state variables:

$$\begin{aligned} &\text{Hull-White} \quad r(t) := x(t) + f^M(0,t) + \alpha(t) && \text{from (2.1.9)},\\ &\text{Cheyette} \quad r(t) := x(t) + f^M(0,t) && \text{from (2.3.3)}. \end{aligned}$$

However, it is true that when giving the same volatility structure (ie. separable deterministic) to both Cheyette and Hull-White models then the auxiliary variable y is fully determined by calculating (2.3.5) directly and Cheyette gives the same zero coupon bond prices as Hull-White when conditioning on the same $r(t)$ as shown in the proof of Proposition 2.3.4. This corresponds to conditioning $x(t) = x$ for Cheyette and having $x(t) = x - \alpha(t)$ for Hull-White model.

2.4 Local Volatility

The dynamics in (2.3.6) along with (2.3.8), hold in general for any separable volatility function that is adapted to the filtration of the governing Brownian motion. In the following we will examine the case when $\sigma_r(t,\cdot)$ depends on the current values of the state variables, which is called local volatility as it does not depend on the particular path of x and y but only on values at t. One could also instead consider a stochastic volatility that depends on the path $W^{\mathbb{Q}}$ takes as discussed in Piterbarg [3] but this is beyond this thesis and we focus on strictly local volatilities.
We impose the following restriction on g and hence σ:

$$\begin{aligned} g(t) &= g(t,x(t),y(t)),\\ \sigma_r(t) &= \sigma(t,x(t),y(t)). \end{aligned}$$

The reason this choice is desirable is that it introduces Markovian dynamics for both state variables. This comes with advantages at evaluating expectations depending on the states or simulating dynamics as well as deploying the Feynmann-Kac formula in Theorem 1.3.3 to solve a PDE. Our dynamics become

$$\begin{aligned} dx(t) &= (y(t) - \kappa(t)x(t))dt + \sigma_r(t,x(t),y(t))dW_t^{\mathbb{Q}}\\ dy(t) &= \left(\sigma_r^2(t,x(t),y(t)) - 2\kappa(t)y(t)\right)dt, \end{aligned} \tag{2.4.1}$$

with the zero coupon bond prices unchanged.

2.4.1 Approximations for Swap Rates

To deal with the calibration of the Cheyette model we will need to establish how to price classic European Swaptions with our end goal being to look at Bermudans. Unlike the Hull-White model, Swaption prices are not analytical in general, due to the added complexity on the dynamics. The best that can be done is to make some approximations on Swap rate dynamics and arrive at a price that way. We will follow the methodology in Andersen and Piterbarg [3, Chapter 13.1.4, page 543] to illustrate some of these approximations that we will later use for calibration.
We will consider the Swap rate dynamics under the annuity measure as it is the common practice since under this specific measure we have a martingale. For a fixed tenor structure, as it is discussed in Section 1.2 we have the Swap rate as

$$S(t,x,y) = \frac{P(t,T_0,x,y) - P(t,T_N,x,y)}{\sum_{n=0}^{N-1}(T_{n+1}-T_n)P(t,T_{n+1})}$$

where the denominator is the annuity $A(t)$ from (1.2.3). In the Cheyette model, conditional zero coupon bond prices additionally depend on the value of the state variables x and y, hence so does the Swap rate giving the following dynamics

$$dS(t,x,y) = \left(\frac{\partial S}{\partial x}(t,x(t),y(t))\right)\sigma_r(t,x(t),y(t))dW_t^A, \tag{2.4.2}$$

using Ito's lemma and the fact that $S(t)$ is a martingale under the annuity measure $\mathbb{Q}^A$ as it is discussed in Section 1.3 and given by Piterbarg [3]. Now the main idea from their book is to say that the Swap rate essentially only depends on the short rate which is fully governed by the state variable x, hence for the above instead of taking the two factor formulation we assume a local volatility in terms of the Swap rate as such with the following lemma in [3, Chapter 13.1.4, Lemma 13.1.3, page 543].

Lemma 2.4.1 (Swap Rate Dynamics)**.** *The value of all European options on the Swap rate $S(t)$ in the Local Volatility Cheyette model are identical to values computed in a vanilla model with time-dependent local volatility function:*

$$dS(t) = \varphi(t,S(t))dW_t^A, \tag{2.4.3}$$

where

$$\varphi(t,s)^2 = \mathbb{E}^A\left[\left(\frac{\partial S}{\partial x}(t,x(t),y(t))\sigma_r(t,x(t),y(t))\right)^2\middle| S(t)=s\right]. \tag{2.4.4}$$

This lemma simplifies what we had in (2.4.2) and provides a more tractable method to determine Swaption prices as well as helping with calibration to Swaption strips. Later, in the linear local volatility case, we will see that the above can be approximated by a displaced log-normal variable and can produce an option price that way.
First, we outline a general plan to calculate $\varphi(t,s)$ without assuming further properties of the volatility. Inspired by the local deterministic property of y, we want to approximate it with a fully deterministic function $\overline{y}(t)$. If this is achieved then S is a deterministic function of just t and x so we could consider an inverse function $x(t) = X(t,S(t))$. This way evaluating the expectation in (2.4.4) becomes straightforward and we get an approximate value for $\varphi(t,s)$

$$\varphi(t,s) \approx \frac{\partial S}{\partial x}(t,X(t,s),\overline{y}(t))\sigma_r(t,X(t,s),\overline{y}(t)).$$

Proposition 2.4.2 (Approximation of $y(t)$)**.** *We approximate $y(t)$ in (2.4.1) under the annuity measure by taking*

$$\overline{y}(t) = \mathbb{E}^A[y(t)] \approx h^2(t)\int_0^t \sigma_r^0(s)^2 h^{-2}(s)ds \tag{2.4.5}$$

Proof. The proof is given by Andersen and Piterbarg [3, Proposition 13.1.4, page 544]. □

We now look to approximate the inverse function giving x which can be done by simply Taylor expanding $S(t,x,y)$ around 0 replacing y by the above:

$$S(t,x,\overline{y}(t)) \approx S(t,0,\overline{y}(t)) + \frac{\partial S}{\partial x}(t,0,\overline{y}(t))x(t)$$

hence we define

$$\xi(t,s) = \frac{s - S(t,0,\overline{y}(t))}{\frac{\partial S}{\partial x}(t,0,\overline{y}(t))} \tag{2.4.6}$$

where $X(t,s) \approx \xi(t,s)$. From (2.3.3) we know that $x(t)$ is just a shifted short rate so 0 serves as a decent expansion point. Putting all of this together we get that the approximate Swap Rate dynamics

Proposition 2.4.3 (Approximate Swap Rate dynamics in the Cheyette model)**.** *Under the annuity measure $\mathbb{Q}^A$ using Lemma 2.4.1 and the above approximations we have*

$$\begin{aligned} dS(t) &\approx \widetilde{\varphi}(t,s)dW_t^A \\ \widetilde{\varphi}(t,s) &= \frac{\partial S}{\partial x}(t,\xi(t,s),\overline{y}(t))\sigma_r(t,\xi(t,s),\overline{y}(t)) \end{aligned} \tag{2.4.7}$$

where $\xi(t,s)$ is as in (2.4.6).

2.5 Linear Local Volatility in the Cheyette Model

We finally turn to the main subject of this thesis, the Linear Cheyette model. As the name suggests we consider one of the simplest local volatility functions as done in [3, Chapter 13.1.5, page 547]

$$\sigma_r(t,x(t),y(t)) = \lambda_r(t)(\alpha_r(t) + b_r(t)x(t)) \tag{2.5.1}$$

where parameters $\lambda_r(t)$, $\alpha_r(t)$ and $b_r(t)$ are all deterministic. It will become more clear at a later stage why we introduce an extra "dummy" parameter $\lambda_r(t)$ but the gist of it is to have a rough understanding of the calibrated model parameters later.

Definition 2.5.1 (Linear Cheyette dynamics)**.** The Linear Cheyette model with 2 factors and volatility function (2.5.1) is governed by the dynamics

$$\begin{aligned} dx(t) &= (y(t) - \kappa(t)x(t))dt + \lambda_r(t)(\alpha_r(t) + b_r(t)x(t))dW_t^{\mathbb{Q}} \\ dy(t) &= \left(\sigma_r^2(t) - 2\kappa(t)y(t)\right)dt \\ x(0) &= y(0) = 0. \end{aligned} \tag{2.5.2}$$

The analysis from Section 2.3 carries over to the linear volatility, zero coupon bond prices and Swap rates staying the same with simplifications to the Swap rate dynamics we discuss next.

2.5.1 Swap Rate Dynamics

Given the linearity of the local volatility function for the short rate, it is reasonable to assume the same for $\varphi(t,S(t))$ and approximate it by choosing a suitable expansion point. The most straightforward approach is to have the initial Swap rate $S(0)$ and we Taylor expand (2.4.7) around $S(0)$ to get

$$\widetilde{\varphi}(t,S(t)) = \widetilde{\varphi}(t,S(0)) + \frac{\partial \widetilde{\varphi}}{\partial s}(t,S(0))(S(t) - S(0)),$$

where

$$\begin{aligned} \widetilde{\varphi}(t,S(0)) &= \lambda_r(t)\frac{\partial S}{\partial x}(t,\xi(t,S(0)),\overline{y}(t))(\alpha_r(t) + b_r(t)\xi(t,S(0))), \\ \frac{\partial \widetilde{\varphi}}{\partial s}(t,S(0)) &= \lambda_r(t)\frac{\partial S}{\partial x}(t,\xi(t,S(0)),\overline{y}(t))\frac{\partial \xi}{\partial s}(t,S(0))) \\ &\quad \times \left[\frac{\partial^2 S(t,\xi(t,S(0)),\overline{y}(t))/\partial x^2}{\partial S(t,\xi(t,S(0)),\overline{y}(t))/\partial x}(\alpha_r(t) + b_r(t)\xi(t,S(0)) + b_r(t)\right]. \end{aligned}$$

by applying the chain rule. We can simplify this a bit further by denoting $\overline{x}(t) = \xi(t, S(0))$ and then

$$\frac{\partial \xi}{\partial s}(t, S(0)) \approx \frac{1}{\partial S(t, \overline{x}(t), \overline{y}(t))/\partial x}.$$

We then have the following

$$\widetilde{\varphi}(t, S(0)) = \lambda_r(t) \frac{\partial S}{\partial x}(t, \overline{x}(t), \overline{y}(t))(\alpha_r(t) + b_r(t)\overline{x}(t)), \tag{2.5.3}$$

$$\frac{\partial \widetilde{\varphi}}{\partial s}(t, S(0)) = \lambda_r(t) \left[\frac{\partial^2 S(t, \overline{x}(t), \overline{y}(t))/\partial x^2}{\partial S(t, \overline{x}(t), \overline{y}(t))/\partial x}(\alpha_r(t) + b_r(t)\overline{x}(t)) + b_r(t) \right]. \tag{2.5.4}$$

This will allow us to arrive at a displaced log-normal dynamics through the following proposition given by Andersen and Piterbarg [3, Corollary 13.1.9, page 548].

Proposition 2.5.2 (Swap Rate Dynamincs in the Linear Cheyette model)**.** *In the model with dynamics (2.5.2) we can approximate the Swap rate process by*

$$dS(t) \approx \lambda_S(t)(b_S(t)S(t) + (1 - b_S(t))S(0)dW_t^A \tag{2.5.5}$$

where

$$\lambda_S(t) = \frac{\widetilde{\varphi}(t, S(0))}{S(0)}, \qquad b_S(t) = S(0)\frac{\partial \widetilde{\varphi}(t, S(0))/\partial s}{\widetilde{\varphi}(t, S(0))}$$

from the expressions given above in (2.5.3) and (2.5.4).

Proof. Substituting in the formulas for $\lambda_S(t)$ and $b_S(t)$ gives the result. □

We recognise the dynamics of a displaced log-normal variable in (2.5.5) which is the first order approximation to any stochastic process with local volatility. The two parameters λ and b correspond to the at-the-money volatility and slope of the smile respectively. We will see later at calibration that this model will generally produce a linear smile with some room to adjust the slope at level. Nevertheless, Proposition 2.5.2 allows us to get a swaption price from these dynamics so we quote the result of Andersen [3, Proposition 13.1.10, page 549]

Proposition 2.5.3 (European Swaption Price in the Linear Cheyette model)**.** *For a fixed term structure $T_0 < T_1 < \ldots < T_N$ and a rate K the price of the payer European Swaption at the present time is approximated by*

$$V(0) \approx A(0) \left[\left(S(0) + \frac{S(0)(1 - \overline{b}_S)}{\overline{b}_S} \right) \Phi(d_+) - \left(K + \frac{S(0)(1 - \overline{b}_S)}{\overline{b}_S} \right) \Phi(d_-) \right]$$

$$d_\pm = \frac{\ln\left(\frac{S(0) + S(0)(1-\overline{b}_S)/\overline{b}_S}{K + S(0)(1-\overline{b}_S)/\overline{b}_S} \right) \pm \frac{1}{2}\overline{b}_S^2 \overline{\lambda}_S^2 T_0}{\overline{b}_S \overline{\lambda}_S \sqrt{T_0}},$$

where

$$\overline{\lambda}_S = \sqrt{\frac{1}{T_0}\int_0^{T_0} \lambda_S^2(t)dt}$$

$$\overline{b}_S = \int_0^{T_0} b_S(t) w_S(t) dt$$

$$w_S(t) = \frac{\lambda_S^2(t) \int_0^t \lambda_S^2(u)du}{\int_0^{T_0} (\lambda_S^2(u) \int_0^u \lambda_S^2(v)dv)du}$$

are time averages of the parameters from (2.5.5).

This will allow us to price European Swaptions when calibrating the model as we will discuss in the following chapter. We do not provide a proof but the idea is simply to transform the general displaced log-normal dynamics from Proposition 2.5.2 into a call option price by some averaging technique on the at-the-money(ATM) volatility and shift parameters.

Chapter 3

Calibration

In this Chapter we are going to discuss the calibration method for the Hull-White and the Linear Cheyette models defined in Chapter 2. The goal of calibration is to determine values of the model parameters such that the model, loosely speaking, fits the market as accurately as possible. Here we will be mainly concerned with the calibration of volatility parameters and we set κ, the mean reversion parameter, exogenously. First we define the specific instruments we consider for calibration and how the two models approach this. Then we will show numerical results of the calibration of both models and analyze the differences.

Given our main objective is to price Bermudan Swaptions, we look to capture the prices of a set of European Swaptions related to a given Bermudan Swaption called co-terminal Swaptions.

Definition 3.0.1 (Co-terminal Swaptions)**.** Given a tenor structure $T_0 < T_1 < \ldots < T_N$ define the series of European Swaptions $(C_n)_{0\leq n\leq N-1}$ to be co-terminal Swaptions if C_n has maturity T_n and underlying IRS with tenor structure $(T_i)_{n\leq i\leq N}$ for all $0 \leq n \leq N-1$.

Note that we do not specify the strikes of these Swaptions and they can be both payer and receiver too.
Now given a Bermudan Swaption, we can take its tenor structure and construct co-terminal Swaptions with selected strikes. If we can observe these products in the market then we can take their quoted values straight away as our calibration targets which the model produced prices should be closest to. There is often a tradeoff between the number of calibration targets chosen and the quality of the fit a model can provide. Usually the simpler the model, the less it can fit to a larger set of products and we will see this with Hull-White and Cheyette.

3.1 Hull-White Calibration

The Hull-White model has two parameters κ and σ that need to be calibrated. In our implementation we take these to be piecewise constant over a tenor structure $T_0 < T_1 < \ldots < T_N$

$$\kappa(t) = \sum_{i=1}^{N} \kappa_i \mathbb{1}_{\{t\in(T_{i-1},T_i]\}}$$
$$\sigma(t) = \sum_{i=1}^{N} \sigma_i \mathbb{1}_{\{t\in(T_{i-1},T_i]\}}$$

As mentioned before, we set the mean reversion prior to calibration to $\kappa(t) \equiv \kappa$ as recommended by Piterbarg in [3] arguing that rather than performing a global optimisation over both parameters it is better to take a rough estimate for the mean reversion and more precisely determine volatility values.
We take co-terminal Swaptions that are at-the-money, i.e. the strikes are given by the current forward Swap rate for maturity of the option, because we are only calibrating a single time dependent parameter σ_i for each call date T_i. Sometimes if the Bermudan has a different strike, we may take that strike for each of the co-terminal Swaptions to achieve a better fit for the particular Bermudan.

Once all the calibration products are determined we try to fit their price exactly by using the pricing formula in Theorem 2.1.4. Observing equation (2.1.13) we see that the price depends on the volatility parameter $\sigma(t)$ through

$$v = \int_t^T \sigma^2(u) \left(G(u,T) - G(u,T')\right)^2 du.$$

In order to hit the prices in the Hull-White model we use the Levenberg–Marquardt global minimizer [10] that achieves the value v above implied by the market price of each Swaption. We do not go into details of how this is implemented, instead we show some numerical results on market data in Section 3.3.

3.2 Linear Cheyette Calibration

For the Linear Cheyette model we take a slightly different approach to calibration. As before, we set the mean reversion parameter $\kappa(t) \equiv \kappa$ exogenously but given the additional volatility parameters, we choose to select more than one strike for each maturity date of our co-terminal Swaptions. This means that our model in principle can fit perfectly to two different prices for one maturity due to the additional degrees of freedom, giving better calibration results than Hull-White. We again consider piece-wise constant values

$$\begin{aligned}
\lambda_r(t) &= \sum_{i=1}^N \lambda_i \mathbb{1}_{\{t \in (T_{i-1}, T_i]\}}, \\
\alpha_r(t) &= \sum_{i=1}^N S_{i,N}(0) \mathbb{1}_{\{t \in (T_{i-1}, T_i]\}}, \\
b_r(t) &= \sum_{i=1}^N b_i D_i \mathbf{1}_{\{t \in (T_{i-1}, T_i]\}},
\end{aligned}$$

where we define a scaling on b_i by

$$D_i = \frac{\partial S_{i,N}}{\partial x}(t, 0, 0).$$

The reason for these choices is taken from Andersen [3] and it is to achieve that the model parameters $\lambda_r(t)$ and $b_r(t)$ are roughly the same as the displaced log-normal parameters $\lambda_S(t)$ and $b_S(t)$ defined in Proposition 2.5.2. This way we get the linear local volatility to be

$$\sigma_r(t,x) = \sum_{i=1}^N \lambda_i (S_{i,N}(0) + b_i D_i x) \mathbb{1}_{\{t \in (T_{i-1}, T_i]\}}.$$

Unlike in the Hull-White calibration above, we do not choose prices of the co-terminal Swaptions as targets. Instead, for each maturity we take a set of Swaptions and fit the displaced log-normal parameters $\overline{\lambda}_S$ and $\overline{b}_S$ to be set as targets for the model to fit directly from model parameters. In order to do this, we deploy an optimiser that takes the implied volatilities of the set of Swaptions selected as targets and comes up with values for $\overline{\lambda}_S$ and $\overline{b}_S$ such that the volatilities implied from these parameters fits the market the best. Once we have these targets we can chain together equations from Proposition 2.5.3 along with expressions

$$\lambda_S(t) = \lambda_r(t) \frac{1}{S(0)} \frac{\partial S}{\partial x}(t, \overline{x}(t), \overline{y}(t))(\alpha_r(t) + b_r(t)\overline{x}(t)), \tag{3.2.1}$$

$$b_S(t) = \frac{S(0)}{\alpha_r(t) + b_r(t)\overline{x}(t)} \frac{b_r(t)}{\frac{\partial S}{\partial x}(t, \overline{x}(t), \overline{y}(t))} + \frac{S(0) \frac{\partial^2 S}{\partial x^2}(t, \overline{x}(t), \overline{y}(t))}{\left(\frac{\partial S}{\partial x}(t, \overline{x}(t), \overline{y}(t))\right)^2}, \tag{3.2.2}$$

to find direct relationship between model parameters λ_r and b_r and the displaced log-normal. Finally, we again apply a Levenberg–Marquardt global optimiser to find values $(\lambda_i)_{0 \le i \le N}$ and $(b_i)_{0 \le i \le N}$. Alternatively one can also take a bootstrapping approach as it is described by Andersen in [3] but we stick with the global optimiser as it fits the design pattern of the general framework the implementation takes place in later discussed in Chapter 5.

Remark 3.2.1. Note that this calibration can be done in the same fashion with setting $b_r(t) \equiv 0$ to get a standard Hull-White model and instead of using the analytical formula from Theorem 2.1.4 consider the Swap rate approximation with a deterministic volatility function. In our findings the two methods display very similar results.

3.3 Numerical Results

In this section we look at calibration results of the methods described above. We take market data from TD Securities's testing framework along with some helper models. In particular, a yield curve model that gives the initial term structure $P(0,t)$ and $f^M(0,t)$ as discussed in Chapter 2 and a volatility model to get implied volatilities of Swaptions. Furthermore, as mentioned above we take the mean reversion in what follows to be $\kappa(t) \equiv 0.01$.

We first discuss the type of smiles the Linear Cheyette model can produce by examining different parameter values. In Section 2.5.1 we explained that disregarding the dummy parameter $\alpha_r(t)$ and simply setting it to the initial Swap rates observed in the market, the linear model is characterised by $\lambda_r(t)$ and $b_r(t)$ which determine the at-the-money level and skewness of the smile respectively.

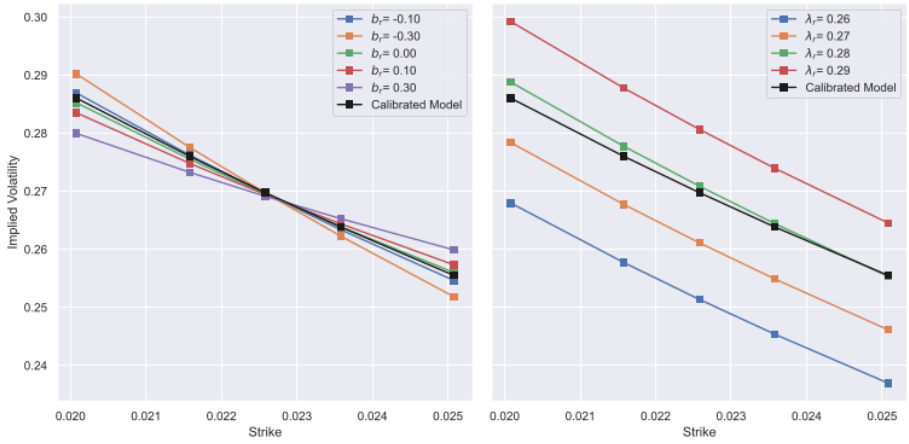

Figure 3.1: European Swaption implied volatility smiles for different flat parameters in the Linear Cheyette model for a fixed maturity $t = 1.493$

In Figure 3.1 we exhibit the shapes the smiles can have when changing parameters of a calibrated model which has initial values given by Table 3.1[1]. The original model was calibrated to a set of co-terminal Swaptions with maturities listed in the table, with implied log-normal volatilities calculated from a SABR-model which is a common volatility model in the finance industry developed by Hagan et al in [11]. We will look at this specific calibration later in more detail but for now we focus on the capabilities the two parameters give.

t	1.000	1.247	1.493	1.740	1.986	2.233	2.479	2.726	2.973	3.219	3.466	3.712
$\lambda_r(t)$	0.268	0.301	0.295	0.307	0.315	0.311	0.296	0.303	0.301	0.293	0.281	0.287
$\alpha_r(t)$	0.023	0.023	0.023	0.023	0.024	0.024	0.024	0.024	0.024	0.024	0.025	0.025
$b_r(t)$	-0.052	-0.218	-0.185	-0.200	-0.218	-0.187	-0.104	-0.077	-0.051	-0.051	-0.115	-0.097

Table 3.1: Linear Cheyette parameters calibrated to 5 Swaptions

We reset the rows in Table 3.1 with constant values for $b_r(t) \equiv b_r$ and $\lambda_r(t) = \lambda_r$. In the right hand plot of Figure 3.1 we adjust parameter λ_r which affects the at-the-money level of the smile. This is presented well as we can see a parallel shift of the original calibrated implied volatilities not

[1]Time values in the Table are given in year fractions so $t = 1.0$ equates to one year

affecting the slope or skewness. On the left of Figure 3.1 we change the skewness parameter. As shown in the plot, this changes the slope of the volatility smile with larger negative b_r value creating a steeper smile and larger positive values flattening it. It is important to note that adjusting this parameter does not affect the overall level of the implied volatilities showing that it is indeed the case that $\lambda_r(t)$ and $b_r(t)$ are in a sense "orthogonal" to each other. This gives the Linear Cheyette model a clear advantage over the general Hull-White model as we will see later when comparing calibrated values for a fixed maturity.

Remark 3.3.1. There are of course some limitations to the capabilities described above. The main issue is with the range of values $b_r(t)$ can take. Looking at the linear local volatility function from (2.5.1)

$$\sigma_r(t, x(t)) = \lambda_r(t)(\alpha_r(t) + b_r(t)x(t))$$

in order to not have zero or negative volatilities we need $\lambda_r(t) > 0$ and furthermore

$$x(t) > -\frac{\alpha_r(t)}{b_r(t)}.$$

A zero volatility would be a problem in (3.2.2) amongst making the model difficult to explain. Based on some Monte-Carlo simulations that we visit later this becomes a real possibility when $b_r(t)$ gets highly negative affecting around -1 which gives some negative volatilities in the simulation.

Remark 3.3.2. It is also worth noting again that $b_r(t) \equiv 0$ yields the Hull-White model. Unlike in the famous Black-Scholes model with a similar deterministic volatility structure, it does not produce flat Swaption smiles as we can see in Figure 3.1. This is due to the fact that Hull-White would produce flat smiles in *normal* volatility space and we are looking at *log-normal* volatilities due to the shifted log-normal approximation from Proposition 2.5.2.

We now turn to the actual calibration of Hull-White and Linear Cheyette model to a volatility surface from a calibrated SABR-model. This is a natural thing to do when it comes to implementation, since it is good to have the same volatility surface for interest rates along potentially different models and products. So to keep it consistent we use the calibrated SABR model to calculate implied volatilities for co-terminal Swaptions. For the Hull-White case, we only take the at-the-money Swaption for each maturity but for Linear Cheyette as outlined in Section 3.2 we need more than one Swaption to fit displaced log-normal parameters $\overline{\lambda}_S$ and $\overline{b}_S$. We elect to take 5 swaptions for each maturity, one which is at-the-money and the others with strikes that are ± 0.001 and ± 0.0025 away from ATM.

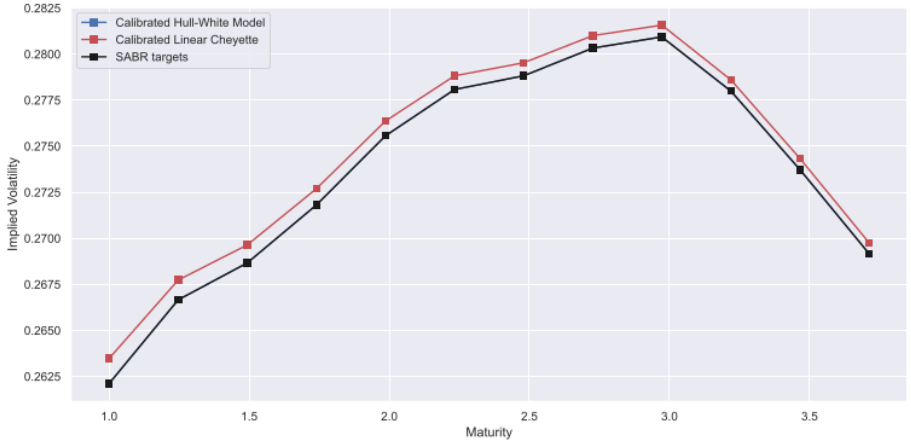

Figure 3.2: Fitted Hull-White and Linear Cheyette model implied volatilities for at-the-money Swaptions across different maturities

We first look at the quality of calibration along different maturities by considering the implied volatilities of at-the-money Swaptions across the term structure. In the market this is often reflected by a volatility "hump" which gets captured by both models. In Figure 3.2 we show this behaviour for both models, keeping in mind that the calibrated Hull-White model will hit all ATM implied volatilities unlike the linear Cheyette model because its targets are displaced log-normal parameters for each maturity. The calibrated parameters of the former are given in Table 3.2.

Remark 3.3.3. We note that we could decrease the number of Swaptions our Linear Cheyette model calibrates to for each maturity, say to just 2, in which case the displaced log-normal parameters can be fit such that the two implied volatilities are perfectly hit. We exhibit this later.

t	1.000	1.247	1.493	1.740	1.986	2.233	2.479	2.726	2.973	3.219	3.466	3.712
$\sigma(t)$	6.0e-3	6.9e-3	6.8e-3	7.2e-3	7.4e-3	7.3e-3	7.0e-3	7.2e-3	7.2e-3	7.1e-3	6.9e-3	7.2e-3

Table 3.2: Hull-White parameters calibrated to ATM Swaptions

Now we consider Swaptions along a single maturity date as in Figure 3.1 for the calibrated models. We choose two maturity dates at the extremes of our term structure $t_1 = 1.00$ and $t_2 = 3.712$ to display implied volatilities for European Swaptions with the 5 different strikes considered at the calibration of the Linear Cheyette model. The results are displayed in Figure 3.3 where the parameters of the calibrated models are given in the above two Tables 3.1 and 3.2.

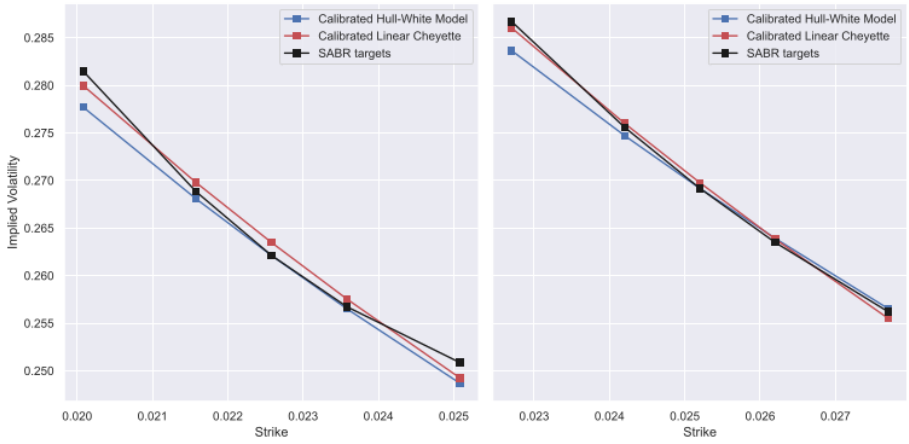

Figure 3.3: Comparing volatility smiles of Hull-White calibrated to ATM Swaptions and Linear Cheyette model calibrated to 5 Swaptions for two maturities

On the left hand graph of Figure 3.3 we have the one year maturity Swaptions which produce a more curved smile than longer maturities. Both models are poor at capturing this, which is not surprising as they can both be approximated by a displaced log-normal that produces linear smiles. The key difference between the two is that in the Linear Cheyette case, parameter $b_r(t)$ allows to adjust the slope as shown earlier, hence the model can provide a better overall fit. In Figure 3.3 the right hand plot shows the longest dated Swaptions, for which the slope of the smile is steeper. The Cheyette calibration adjusts to this change but with the Hull-White model the implied volatilities are just a parallel shift of the one-year graph.
Finally we address Remark 3.3.3 and calibrate the Linear Cheyette model to only 2 Swaptions for each maturity date. We pick at-the-money and an in-the-money[2] Swaption with strike -0.0025 or -25bps away from the spot Swap rate. For the Hull-White model we keep the calibration the same to the ATM implied volatility. As discussed before the Linear Cheyette now should hit both of these implied volatilities perfectly for each of the maturities. This is exactly what we see in Figure 3.4 considering the same two maturities as in Figure 3.3.

[2]An option is in-the-money (ITM) if it would generate a profit if the maturity was today. Similarly an option is

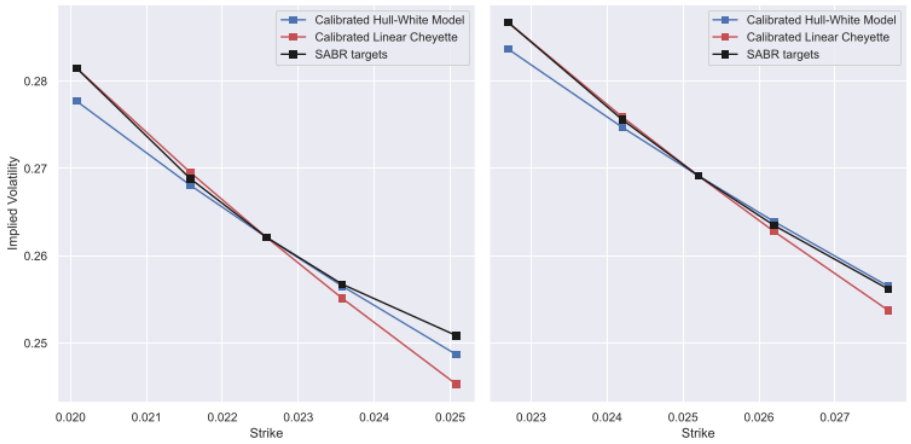

Figure 3.4: Comparing volatility smiles of Hull-White calibrated to ATM Swaptions and Linear Cheyette model calibrated to 2 ITM Swaptions for two maturities

This calibration method is useful when we want to capture specific parts of the volatility smile, for example having a Bermudan Swaption that is not at-the-money. The downside is of course that our model will be poor at capturing the behaviour for strikes away from the two targets as we see for deeply out-of-the-money Swaptions in Figure 3.4. The calibrated parameters are given below in Table 3.3

t	1.000	1.247	1.493	1.740	1.986	2.233	2.479	2.726	2.973	3.219	3.466	3.712
$\lambda_r(t)$	0.267	0.301	0.294	0.307	0.314	0.310	0.296	0.302	0.300	0.293	0.280	0.287
$\alpha_r(t)$	0.023	0.023	0.023	0.023	0.024	0.024	0.024	0.024	0.024	0.024	0.025	0.025
$b_r(t)$	-0.240	-0.306	-0.294	-0.292	-0.296	-0.262	-0.182	-0.148	-0.115	-0.120	-0.205	-0.184

Table 3.3: Linear Cheyette parameters calibrated to 2 ITM Swaptions

We can see from the parameters in the table above that the steeper smiles in Figure 3.4 correspond to larger negative values of b_r than what is given in Table 3.1 confirming the discussion earlier. We also note that $\lambda_r(t)$ only changes slightly in this calibration too, given we have the same ATM level as a target. In a similar fashion, one can calibrate to out-of-the-money Swaptions to achieve a better fit on the other end of the smile, exhibiting larger and often positive values for b_r indicating a flatter implied volatility structure. The results for this are shown below in Figure 3.5 for the same maturities and set of Swaptions as earlier with calibrated parameter values detailed in Table 3.4.

t	1.000	1.247	1.493	1.740	1.986	2.233	2.479	2.726	2.973	3.219	3.466	3.712
$\lambda_r(t)$	0.267	0.301	0.295	0.307	0.315	0.311	0.296	0.303	0.300	0.293	0.281	0.287
$\alpha_r(t)$	0.023	0.023	0.023	0.023	0.024	0.024	0.024	0.024	0.024	0.024	0.025	0.025
$b_r(t)$	0.155	-0.123	-0.068	-0.101	-0.134	-0.108	-0.019	-0.001	0.019	0.024	-0.018	-0.004

Table 3.4: Linear Cheyette parameters calibrated to 2 OTM Swaptions

Remark 3.3.4. In our data set the SABR implied smiles are all relatively close to what a Hull-White model produces. In these cases one would likely prefer the simpler model to price products as it is generally a quicker and simpler alternative. However, in an environment where the target smiles are much steeper or much flatter than what we have here, the Linear Cheyette model is comfortably a better choice.

---

out-of-the-money (OTM) if it would *not* generate a profit, i.e. would not be exercised today.

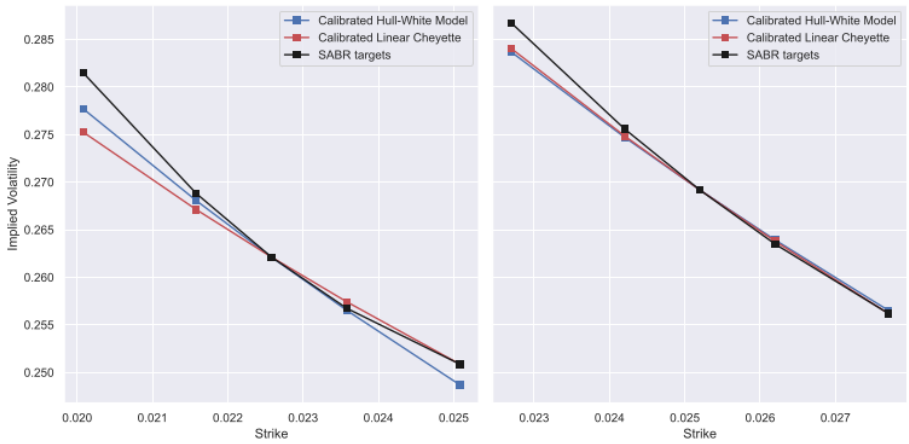

Figure 3.5: Comparing volatility smiles of Hull-White calibrated to ATM Swaptions and Linear Cheyette model calibrated to 2 OTM Swaptions for two maturities

Chapter 4

Markovian Projection

In this chapter we are going to consider a technique called Markovian projection outlined in Andersen Piterbarg [3, Chapter 13.1.9.4, page 565] which reduces the two factor model described in Section 2.5 to a one factor model. The reason one might consider reducing dimensionality is to improve the speed of numerical calculations, in particular when pricing by solving a PDE, while still retaining desirable features of the original model for calibration.

4.1 Underlying Theory

The main theorem justifying Markovian projection belongs to Gyöngy [12] and goes as follows.

Theorem 4.1.1 (Markovian projection - Gyöngy)**.** *Let $X(t)$ be defined by the stochastic differential equation*

$$dx(t) = \mu(t)dt + \sigma(t)dW_t^{\mathbb{Q}},$$

where $(W_t^{\mathbb{Q}}, \mathcal{F}_t)_{t\geq 0}$ is a Wiener process under some probability measure $\mathbb{Q}$, furthermore $\mu(t)$ and $\sigma(t)$ are bounded and $\mathcal{F}_t$-nonanticipative processes. Then the following stochastic differential equation admits a weak solution $\widetilde{x}(t)$

$$d\widetilde{x}(t) = \widetilde{\mu}(t, \widetilde{x}(t))dt + \widetilde{\sigma}(t, \widetilde{x}(t))dW_t,$$

where the drift and volatility coefficients are given by conditional expectations over $\mathbb{Q}$:

$$\widetilde{\mu}(t,x) = \mathbb{E}^{\mathbb{Q}}\left[\mu(t) \mid x(t) = x\right], \qquad (4.1.1) \qquad \widetilde{\sigma}^2(t,x) = \mathbb{E}^{\mathbb{Q}}\left[\sigma^2(t) \mid x(t) = x\right]. \qquad (4.1.2)$$

Furthermore, $\widetilde{x}(t)$ has the same one dimensional distribution as $x(t)$ for every t.

The crucial property of the projection is given by the last statement of this theorem, namely that the distribution of the projected variable $\widetilde{x}$ agrees with the original and this allows for taking expectations of *undiscounted* payoffs and have

$$\mathbb{E}^{\mathbb{Q}}\left[V(x(T))\right] = \mathbb{E}^{\mathbb{Q}}\left[V(\widetilde{x}(T))\right],$$

where the expectation is again taken with respect to the $\mathbb{Q}$-measure.
We can apply Theorem 4.1.1 to the dynamics of a general Cheyette model described in (2.3.6) in order to project $y(t)$ onto x by taking conditional expectation. Keeping the notation from earlier we get the following dynamics

$$d\widetilde{x}(t) = \left(\widetilde{y}(t, \widetilde{x}(t)) - \kappa(t)\widetilde{x}(t)\right)dt + \widetilde{\sigma}_r(t, \widetilde{x}(t))dW_t^{\mathbb{Q}}. \qquad (4.1.3)$$

The main objective is to accurately calculate $\widetilde{y}(t,x)$ governed by (4.1.1)

$$\widetilde{y}(t,x) = \mathbb{E}^{\mathbb{Q}}\left[y(t) \mid x(t) = x\right] \qquad (4.1.4)$$

as often times and also in our application of the Linear Cheyette model, the volatility does not depend on y and hence taking the conditional expectation in (4.1.2) has no effect. In the following we will assume this to be the case.

4.1.1 Recursive Method for Markovian Projection

One way to calculate $\widetilde{y}(t,x)$ as it is described by Piterbarg [3, page 568] is to discretize the process for timesteps $[t_n, t_{n+1}]$. Then using shorthand notation x_n and y_n for $x(t_n)$ and $y(t_n)$ respectively we get the following difference equation

$$y_{n+1} = y_n + \left(\sigma_r^2(t_{n+1}, x_{n+1}) - 2\kappa(t)y_n\right)\Delta_n, \quad \Delta_n = t_{n+1} - t_n,$$

noting that σ_r is evaluated at the right side of the interval as this way taking expectation conditioning on x_{n+1} and x_n is easy to handle. We get

$$\mathbb{E}^{\mathbb{Q}}\left[y_{n+1} \mid x_n, x_{n+1}\right] = \mathbb{E}^{\mathbb{Q}}\left[y_n \mid x_n, x_{n+1}\right] + \left(\sigma_r^2(t_{n+1}, x_{n+1}) - 2\kappa(t))\mathbb{E}^{\mathbb{Q}}\left[y_n \mid x_n, x_{n+1}\right]\right)\Delta_n. \quad (4.1.5)$$

Notice that by the Markov property $\mathbb{E}\left[y_n \mid x_n, x_{n+1}\right] = \mathbb{E}\left[y_n \mid x_n\right]$. Having the above we can calculate expectation of y_{n+1} conditional on x_{n+1} by integrating (4.1.5) over values of x_n. After rearranging some terms this yields

$$\mathbb{E}^{\mathbb{Q}}\left[y_{n+1} \mid x_{n+1}\right] = \int \mathbb{E}^{\mathbb{Q}}\left[y_n \mid x_n = x, x_{n+1}\right]\mathbb{Q}(x_n \in dx \mid x_{n+1}) =$$

$$\sigma_r^2(t_{n+1}, x_{n+1})\Delta_n + (1 - 2\kappa(t)\Delta_n)\int \mathbb{E}^{\mathbb{Q}}\left[y_n \mid x_n = x\right]\mathbb{Q}(x_n \in dx \mid x_{n+1}). \quad (4.1.6)$$

The recursive relationship between $\widetilde{y}_n$ and $\widetilde{y}_{n+1}$ makes this method attractive to use in numerical applications.
Suppose we are given a grid of (t,x) values with n_t points in the time direction and n_x in the x direction. From the initial dynamics in equation (2.3.6) we see that y_0 can be set 0 for all points $(0,\cdot)$ on the grid. We can now calculate in a forward propagating fashion for y_{n+1} once all y_n values are given on grid points $(t_n,\cdot)$ by applying (4.1.6). If the conditional distribution has small variance then the integral above can be approximated by just a couple of terms which gives overall complexity of this projection is $\mathcal{O}(n_t n_x)$.
The only thing left to do is to calculate the conditional probabilities along values of dx conditioned on x_{n+1}. In the next section we will do this for a Linear Cheyette model by using a Gaussian approximation which works well in general when Δ_n is small and allows us to realise the complexity given above.

4.2 Application to Linear Cheyette

As mentioned above, our approach is going to be to approximate the random variable $x_n \mid x_{n+1}$ by a Gaussian random variable. For this we need to determine the first two moments of the conditioned random variable. However one could also assume that x_n and x_{n+1} are jointly normal variables and then that implies that the conditioned random variable is also normal with mean and variance linked with mean, variance and covariance of x_n and x_{n+1}. These are the values that we are going to calculate to then eventually arrive at $\mathbb{Q}(x_n \in dx \mid x_{n+1})$. The following Lemma justifies this approach.

Lemma 4.2.1 (Conditional bivariate normal distribution)**.** *Suppose that $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$ are jointly normal random variables with correlation coefficient ρ. Then*

$$X_1 \mid X_2 = x \sim \mathcal{N}\left(\mu_1 + \frac{\sigma_1}{\sigma_2}\rho(x - \mu_2), (1 - \rho^2)\sigma_1^2\right),$$

is also normally distributed.

Recall that in Section 2.5 we defined the Linear Cheyette model with local volatility function

$$\sigma_r(t, x(t)) = \lambda(t)(a(t) + b(t)x(t))$$

and corresponding dynamics for $x(t)$ are

$$dx(t) = (y(t) - \kappa(t)x(t))dt + \lambda(t)(a(t) + b(t)x(t))dW_t$$

Then using Lemma 4.2.1 with these dynamics we get the result below.

Lemma 4.2.2 (Distribution of x_n given x_{n+1})**.** *Denoting the unconditional mean and variance of $x(t)$ as $m(t)$ and $S(t)$, following the convention that $m_n = m(t_n)$ and $S(t_n) = S_n$ then we have the approximation*

$$x_n \mid x_{n+1} = x \sim \mathcal{N}\left(m_n + \frac{\text{Cov}[x_n, x_{n+1}]}{S_{n+1}}(x - m_{n+1}), S_n - \frac{\text{Cov}^2[x_n, x_{n+1}]}{S_{n+1}}\right),$$

where

$$\begin{aligned} m(t) &= \int_0^t \overline{y}(s) e^{-\int_s^t \kappa(u)du} ds, \\ S(t) &= \int_0^t \sigma^2(s, m(s)) e^{\int_s^t \lambda^2(u)b^2(u) - 2\kappa(u)du} ds, \\ \text{Cov}[x, x + dx] &= S(t) - \kappa(t)S(t)dt. \end{aligned}$$

Noting that $\overline{y}(t)$ is defined in (2.4.5).

Proof. The proof is given in Appendix B.1. □

4.2.1 Integral Projection

Before we analyze the quality of the recursive method using the lemma above, we introduce a different method to compare. Andersen and Piterbarg [3] argue that one can directly approximate $\widetilde{y}(t, x)$ in a similar fashion to (2.4.5) but instead writing

$$\widetilde{y}(t, x) \approx h^2(t) \int_0^t \mathbb{E}^{\mathbb{Q}}\left[\sigma_r^2(u, x(u)) \middle| x(t) = x\right] h^{-2}(u) du,$$

given again that the local volatility does not explicitly depend on y. Now further approximating $\sigma_r^2(\cdot)$ by a linear function we can move the conditional expectation inside and have

$$\mathbb{E}^{\mathbb{Q}}[x(u)|x(t) = x] \approx \frac{\text{Var}(x(u))}{\text{Var}(x(t))} = \frac{\overline{y}(u)}{\overline{y}(t)} x$$

which yields

$$\widetilde{y}(t, x) \approx h^2(t) \int_0^t \sigma_r^2\left(u, \frac{\overline{y}(u)}{\overline{y}(t)} x\right) h^{-2}(u) du. \tag{4.2.1}$$

This is quite straightforward to calculate, however in application there are concerns about efficiency of evaluating this integral. Suppose we have a grid of points in t and x dimension with number of points n_t and n_x respectively (like we would in a PDE setting) then calculating this accurately for every point on the grid would require dissecting the interval $[0, t_i]$ into n_{t_i} parts and then doing the same calculation $n_t n_x$ times roughly giving complexity $\mathcal{O}(n_t^2 n_x)$ which is more costly than solving the PDE itself. Nevertheless, we use this method as a benchmark to the recursive projection which is far more efficient for finer time-grids.

4.3 Numerical Results

As it is outlined in the beginning of this Chapter we use these projections to reduce dimension of the model and to only have to solve a 1 dimensional PDE applying Feynmann-Kac Theorem 1.3.3 to the one factor model in (4.1.3). We arrive at the PDE under the risk neutral measure $\mathbb{Q}$

$$\begin{aligned} \frac{\partial V}{\partial t}(t, x) + (\widetilde{y}(t, x) - \kappa(t)x(t))\frac{\partial V}{\partial x}(t, x) + \frac{1}{2}\sigma_r^2(t, x(t))\frac{\partial^2 V}{\partial x^2}(t, x) &= r(t)V(t, x) \\ V(T, x) &= f(x) \end{aligned} \tag{4.3.1}$$

for a terminal payoff $f(x(T))$. We will use a slightly different form of this equation with an implicit finite difference scheme to determine the price of various products later but for now let us analyze

the qualitative and quantitative behaviour of the two different projection methods described above on a similar grid that would be used to solve the PDE (4.3.1).

We test the methods against a Monte Carlo simulation that aims to calculate $\widetilde{y}(T,x)$ in the Linear Cheyette model for a terminal time T. An Euler-scheme is implemented on the dynamics in (2.5.2) to simulate both state variables over N_{MC} paths. To calculate the conditional expectation (4.1.4) for $y(T)$ we sample the paths where $x(T)$ is within some tolerance of a grid point x_i and record the value of y on the same path. Finally at the end of the simulation we take expectation for each grid point along x to get proxies for $\widetilde{y}(T,x)$.
We consider two different parameter sets for the Linear Cheyette model for our simulations, one with constant parameters throughout the whole term structure and the other with varying values given in Tables 4.1 and 4.2. We take a grid with time horizon of 3 years dissected into n_t points. Similarly we define an interval $[-0.5, 0.5]$ for state variable x and dissect it into $n_x = 1000$ points.

t	0.00	0.60	1.20	1.80	2.40	3.00
$\lambda_r(t)$	1.00	1.00	1.00	1.00	1.00	1.00
$\alpha_r(t)$	0.02	0.02	0.02	0.02	0.02	0.02
$b_r(t)$	0.10	0.10	0.10	0.10	0.10	0.10

Table 4.1: Flat Linear Cheyette model parameter values

The first parameters we consider are given in Table 4.1. We run the projection with different values for n_t and the results are shown in the top part of Figure 4.1. Simultaneously we also record mean-square errors compared to the Monte Carlo simulation. The simulated values are shown as the benchmark flat 0 in the top two plots with a shaded area indicating a single standard deviation of the simulation away from the calculated means. We can see that the integral projection performs quite fell, almost immediately converging and not improving any accuracy as we increase n_t. On the other hand, the recursive method converges slower requiring a finer grid in the time direction, although this does not always improve the quality of the projection compared to the Monte Carlo simulation as we will see later. Despite the slower convergence, there is significant difference in complexity of the two algorithms which is noticeable in run-time of the experiments as well.
Overall it seems that the integral projection is better to use in this case due to more accurate results, most likely due to the fact that the integral (4.2.1) can be computed directly and this approximation works reasonably well.

Remark 4.3.1. Looking at the definition of $\widetilde{y}(t,x)$ as the conditional expectation in (4.1.4), intuitively it should be increasing in x for a fixed time t as a bigger x at the terminal time t entails a larger realised variances along the path which essentially corresponds to $y(t)$ by (2.3.5). We note that this is realised in the case of both projection methods but is not seen in Figures 4.1 or 4.2 due to slight noise in the Monte Carlo simulation.

Now we take more general parameters that change over the term structure given in Table 4.2 which are more reflective of what we saw in Chapter 3. The results are shown in a similar fashion to earlier in Figure 4.2. Unlike earlier, both methods seem to struggle more with convergence, having increased the density of the time-grid more. Furthermore, the standard deviation of the Monte Carlo simulations are also higher, indicating that it is less straightforward to determine $\widetilde{y}(T,x)$ in this setting.

The integral projection consistently undershoots compared to the simulation and seems to converge slowly as n_t increases but taking significant time to compute projections. On the contrary the recursive method again gets closer to the Monte Carlo values with relatively small $n_t = 50$ but then loses some of its accuracy.
This can be due to a number of factors but we believe the main issue is the errors on the boundaries that propagate into the recursion. The algorithm described in Section 4.1.1 struggles to compute the integral in (4.1.6) as it does not know of any values outside the grid. In our implementation we truncate at the boundaries and accumulate the rest of the probability distribution $\mathbb{Q}\left(x_n \notin [-0.5, 0.5] | x_{n+1}\right)$ with the $\widetilde{y}_n$ values coming from the boundaries. At each time-step this introduces a slight error into the projected values that over each iteration affects values further into the grid.

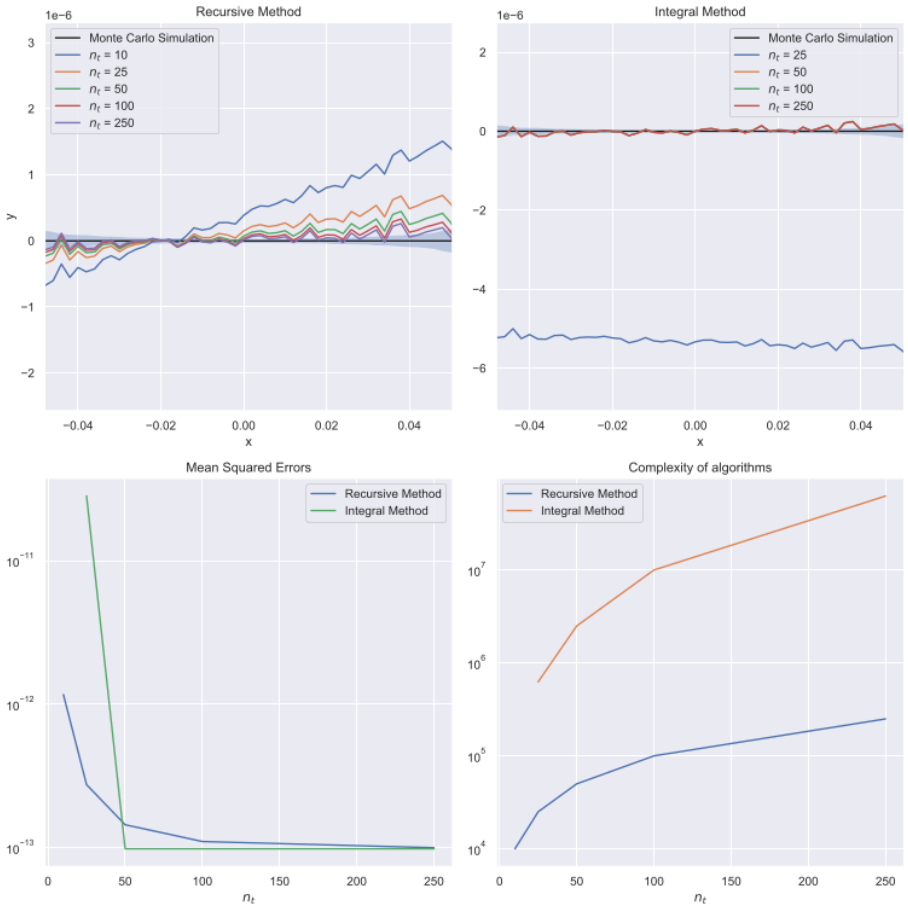

Figure 4.1: Recursive and Integral methods for Markovian projection for Linear Cheyette model with flat parameters against a Monte Carlo simulation with $N_{\text{MC}} = 100000$ paths

Remark 4.3.2. We can see that both Figures 4.1 and 4.2 show x values between much smaller intervals than the original grid [-0.5, 0.5]. This is because the Monte Carlo simulation does not often finish at the x values not shown, however the larger grid is crucial for the recursive projection in order for it to produce accurate values for the relevant x values the Monte Carlo hits.

t	0.000	0.600	1.200	1.800	2.400	3.000
$\lambda_r(t)$	1.000	1.000	1.000	1.000	1.000	1.000
$\alpha_r(t)$	0.020	0.022	0.024	0.026	0.028	0.030
$b_r(t)$	0.101	0.102	0.103	0.104	0.105	0.106

Table 4.2: Time dependent Linear Cheyette model parameter values

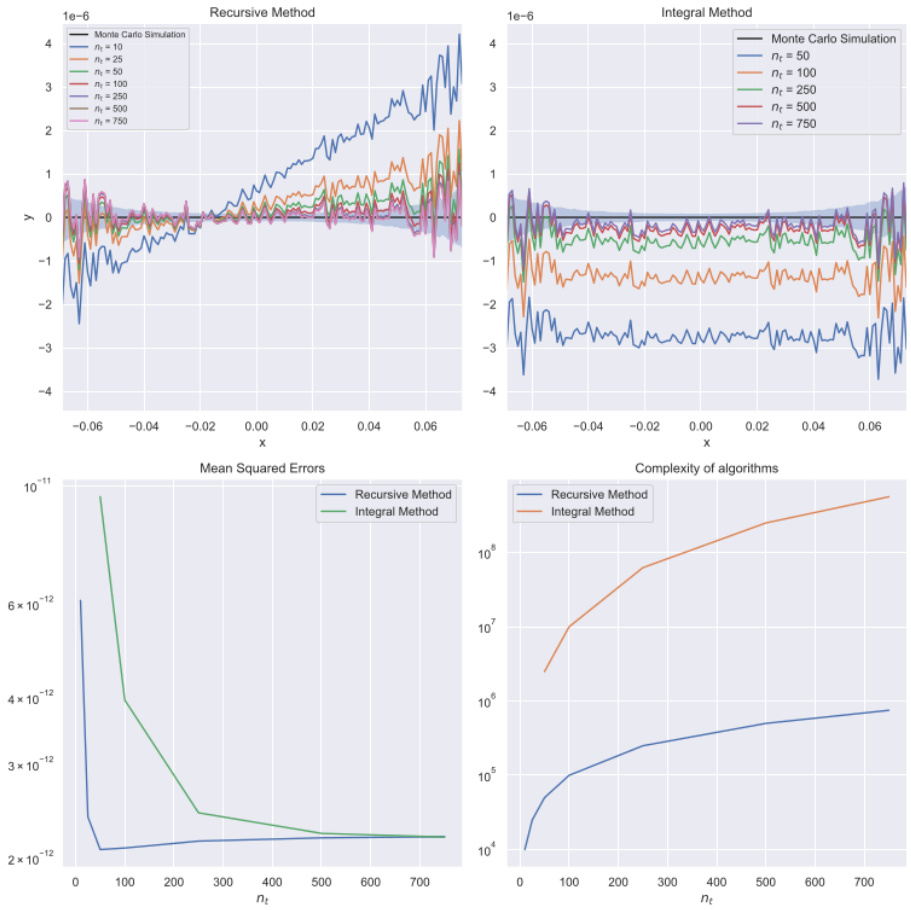

Figure 4.2: Recursive and Integral methods for Markovian projection for Linear Cheyette model with time dependent parameters against a Monte Carlo simulation with $N_{\text{MC}} = 100000$ paths

Chapter 5

Implementation

In this chapter we give a brief overview of the implementation of the Linear Cheyette model into TD Securities's C++ library UQL[1]. We provide a bird's-eye view of the design paradigm, highlighting places where special treatment was needed for our model and additional contributions.

5.1 Valuation Engines

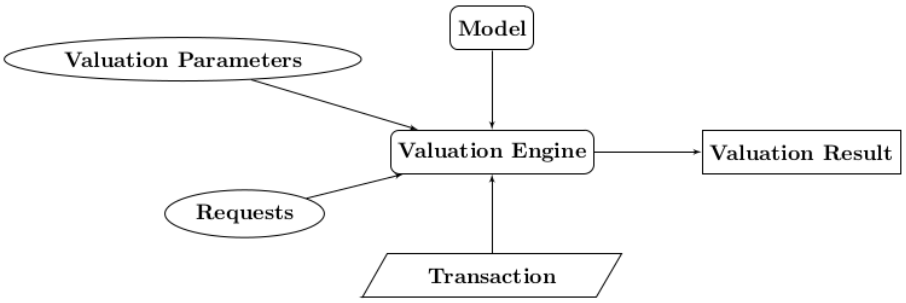

Figure 5.1: Valuation Engine design in UQL

The most common calculation UQL does is a valuation which is done by Valuation Engines. Every engine combines 4 different components to create a valuation report as shown in Figure 5.1. First and foremost it takes a transaction that it performs the calculations for, along with a request that contains information about the type of calculation. For example, this can be present value(PV) or implied volatility(IV). As we have seen earlier these kind of calculations require some sort of model that provide the relevant dynamics and formulas. This is the third component required to create a Valuation Engine. Finally, one also needs to provide valuation parameters, that are a set of instructions for the valuation engine for the specific model.

In the case of our implementation, we need two Valuation Engines. One is used during the calibration process when we need to handle European Swaptions and their implied volatilities, the second is for Bermudan Swaptions which we want to price in the end. For the former we use the approximation outlined in Section 2.5.1 to calculate PV and IV through the Linear Cheyette model. However, in the case of Bermudan Swaptions we need a different approach. The valuation is done via a partial differential equation and is set up inside the engine with the help of the model providing the coefficients, numeraire and boundary conditions. Our implementation with the help of Markovian projection supports an interface that can handle calculations related to a PDE.

It is important to note that Valuation Engines do not know of the creation of the model or the way it was calibrated and simply take dynamics from them.

[1]Unified Quantitative Library

5.2 Models

In UQL we structure models into a set of components corresponding to different discounting or funding curves. As we briefly mentioned in Chapter 1 there is not just a single simply compounded rate $L(t,T)$ but rather it exists for different tenors and currencies. To address this, within a component we pick a particular curve and use that for all discounting of cash flows. For example, this can be the USD 3 month LIBOR curve or the Federal Fund Effective Rate.
Models are created within the Model Factory as shown in Figure 5.2. There are a number of ingredients needed to build and calibrate a model from scratch. First of all, we need some market data that we wish to capture. This can vary between components so we use helpers that process this data into calibration targets for each component. This process may rely on calculations that are done through another model, called reference model. For example, when calibrating the Linear Cheyette Model in Section 3.3 we use a yield curve model to get the initial term structure and a SABR model to get implied volatilites from the market. We also provide the factory with a recipe for how to build the model called build methods. This includes determining the optimizer for the calibration but also whether the model should get implied volatilities from a reference model or quotes provided by the market data.

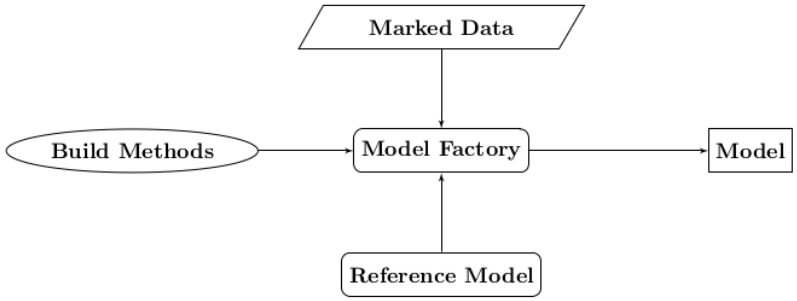

Figure 5.2: Model Factory design in UQL

The challenge for the Cheyette model was to fit the model implementation into the numerical engine for PDEs. We implemented an extra layer within the model that enabled each of the model components to be projected via methods in Chapter 4 while still keeping the model open to extension if a 2 dimensional PDE solver is added to UQL.

Chapter 6

Pricing Results for Linear Cheyette Model

In this chapter we present pricing results under a calibrated Linear Cheyette model. First, we will look at an alternative approach to Section 2.5.1 pricing European Swaptions which is used in calibration, we implement a Monte Carlo scheme following the guidance of Piterbarg in [3, Section 13.1.9.3, page 565] for this to benchmark against the Swaption approximations. Finally a PDE pricer will be used to derive Bermudan Swaption values. Let us first set up this framework.

6.1 PDE Approach for Linear Cheyette Model

Recall the Fundamental Theorem of Asset Pricing 1.3.1 from Section 1.3. As mentioned, the same theorem holds for different numeraire pairs $(N, \mathbb{Q}^N)$ and we define the T-forward measure as the equivalent martingale measure with numeraire $P(\cdot, T)$ denoted by $\mathbb{Q}^T$. Then we have the following theorem.

Theorem 6.1.1 (Pricing under the T-forward measure)**.** *In a complete arbitrage-free market with the fair-value of a contingent claim $V(T)$ at present time t under the T-forward measure is*

$$V(t) = P(t,T)\mathbb{E}^{\mathbb{Q}^T}\left[V(T)|\mathcal{F}_t\right]. \tag{6.1.1}$$

Proof. This is straightforward to see by applying the general from of Theorem 1.3.1 noticing that $P(t,T)$ is $\mathcal{F}_t$-measurable and $P(T,T) = 1$. □

We can now derive the analogous PDE to (4.3.1) using Theorem 1.3.3 in the T-forward measure on the projected model (4.1.3). First, we need to derive the projected Cheyette dynamics under $\mathbb{Q}^T$ by applying a change of measure and hence a different Brownian motion $W^{\mathbb{Q}^T}$. We use the change of numeraire toolkit from Brigo [2, Chapter 2.3, Proposition 2.3.1, page 33] to apply to T-forward measure.

Proposition 6.1.2 (Change to T-forward measure)**.** *Let X be a stochastic process under the risk neutral measure $\mathbb{Q}$ with numeraire*

$$dX(t) = \mu(t, X(t))dt + \sigma(t, X(t))dW_t^{\mathbb{Q}},$$

where both $\mu_t^{\mathbb{Q}}$ and σ_t are adapted to the filtration generated by 1 dimensional Brownian motion $W^{\mathbb{Q}}$. Let us assume the two numeraires B and $P(\cdot, T)$ that evolve under the T-forward measure $\mathbb{Q}^T$ as

$$dB(t) = (\ldots)dt + \sigma_t^B dW_t^{\mathbb{Q}^T}$$
$$dP(t,T) = (\ldots)dt + \sigma_t^P dW_t^{\mathbb{Q}^T}$$

where both σ_t^B and σ_t^P are adapted as well. Then the X can be written under the T-forward measure as

$$dX(t) = \mu^T(t, X(t))dt + \sigma(t, X(t))dW_t^{\mathbb{Q}^T}$$

where the drift term is given by

$$\mu^T(t, X(t)) = \mu(t, X(t)) - \sigma(t, X(t)) \left(\frac{\sigma_t^B}{B(t)} - \frac{\sigma^P(t)}{P(t,T)} \right) \tag{6.1.2}$$

Proof. This is simply a special case of the outlined Proposition 2.3.1 in [2]. □

Using the above we just need to determine the diffusion coefficient $\sigma^P(t)$ to have the T-forward dynamics. We use Ito's lemma on the derived zero coupon bond price in (2.3.8) to get

$$dP(t,T) = (\ldots)dt - P(t,T)G(t,T)\sigma_r(t, x(t))dW_t^{\mathbb{Q}^T},$$

bearing in mind that the volatility of the process stays the same under different equivalent martingale measures and that we substitute $\widetilde{y}(t,x)$ during the calculation. Hence we have that the new drift of state variable x using (6.1.2) is

$$\mu^T(t, x(t)) = \widetilde{y}(t,x) - \kappa(t)x(t) - G(t,T)\sigma_r^2(t, x(t))$$

Hence the dynamics of the projected Linear Cheyette model under the T-forward measure are

$$dx(t) = \left(\widetilde{y}(t,x) - \kappa(t)x(t) - G(t,T)\sigma_r^2(t, x(t))\right) dt + \sigma_r(t, x(t))dW_t^{\mathbb{Q}^T} \tag{6.1.3}$$

So now when pricing in the T-forward measure with the above dynamics, we can apply Feynmann-Kac 1.3.3 to get

Proposition 6.1.3 (Projected Linear Cheyette PDE under T-forward measure)**.** *In a complete arbitrage-free market the fair-value $V(t,x)$ at time t and state $X(t) = x$ of a contingent claim depending on the terminal state of X by $f(X(T))$ under the T-forward measure is given by the solution to the partial differential equation*

$$\begin{aligned} \frac{\partial \widetilde{V}}{\partial t}(t,x) + (\widetilde{y}(t,x) - \kappa(t)x(t) - G(t,T)\sigma_r^2(t,x(t)))\frac{\partial \widetilde{V}}{\partial x}(t,x) + \frac{1}{2}\sigma_r^2(t,x(t))\frac{\partial^2 \widetilde{V}}{\partial x^2}(t,x) = 0 \quad (6.1.4) \\ \widetilde{V}(T,x) = f(x) \end{aligned}$$

where $\widetilde{V}(t,x) = \frac{V(t,x)}{P(t,T)}$ is discounted value with respect to the bond.

Proof. Simply apply Feynmann-Kac to the conditional expectation in (6.1.1) with $r(u) \equiv 0$ and move $P(t,T)$ to the other side to get the result. □

This PDE will be our main tool to use when pricing in the Projected Linear Cheyette model and we will exclusively use this for Bermudan Swaptions.

Remark 6.1.4. The same PDE will hold for the Hull-White model where there is no need for projection as y is deterministic and we can straight away apply the change of numeraire and Feynmann-Kac to the model itself. We will use this to compare our results for the Linear Cheyette later.

6.2 European Swaption Results

In this section we check the accuracy of the Swap rate and Swaption approximations as described in Section 2.5.1 used for the calibration of the full Linear Cheyette model before any projection takes place. We benchmark the prices produced by the model to a Monte Carlo simulation run on the same parameters. Our approach is the same as outlined by Piterbarg [3, Section 13.1.3.9, page 565], using the trick that instead of simulating the auxiliary state y directly we define

$$u(t) = y(t) - \overline{y}(t),$$

which has a simpler drift term easing computations. Otherwise, we implement a simple Euler-scheme on the discretized dynamics of (2.3.6) and deploy an antithetic scheme[1]. This helps to somewhat reduce the final standard deviation of the simulations.

[1]For each Brownian path generated we take the mirrored path into account at the simulation too

We consider the model fitted to the general smile from Section 3.3 with parameter values given in Table 3.1 and we value the same Swaptions given in the calibration with a standard notional of 100.0. The two maturities we take are the same as in Figure 3.3 to test the accuracy of the Swaption approximation on both ends of time scale used in calibration and to examine if the accuracy depends on the curvature of the smile.
The results for the shorter maturity $t = 1.0$ are presented in Table 6.1 and in Table 6.2 for the longer dated one. N_{paths} denotes the number of paths simulated, where we kept the number of timesteps within each path constant at $N_{\text{step}} = 1000$.

	Strikes	0.02008	0.02158	0.02258	0.02358	0.02508
	Model Prices	1.77879	1.3484	1.09955	0.881843	0.612737
		Monte Carlo				
$N_{\text{paths}} = 10000$	Price	1.77171	1.34109	1.09163	0.874023	0.60543
	Std.	0.01913	0.01697	0.01544	0.013864	0.011498
	Error	0.00708	0.00731	0.00792	0.00782	0.007307
$N_{\text{paths}} = 100000$	Price	1.78216	1.35246	1.10415	0.886151	0.61632
	Std.	0.00612	0.00544	0.00495	0.004458	0.003717
	Error	0.00337	0.00406	0.0046	0.004308	0.003583
$N_{\text{paths}} = 1000000$	Price	1.77909	1.3487	1.09977	0.882138	0.613082
	Std.	0.00193	0.00172	0.00156	0.001409	0.001174
	Error	0.0003	0.0003	0.00022	0.000295	0.000345
$N_{\text{paths}} = 10000000$	Price	1.7795	1.34919	1.10035	0.882647	0.613488
	Std.	0.00061	0.00054	0.0005	0.000445	0.000372
	Error	0.00071	0.00079	0.0008	0.000804	0.000751

Table 6.1: Linear Cheyette model produced European Swaption prices across strikes for fixed maturity $t = 1.0$ by Swaption approximation and Monte Carlo simulations

From the results we can see that the Monte Carlo simulation slowly converges gaining about a factor of 10 in accuracy for an increase of 100 times the number of simulations. The number of paths is also proportional to the time it takes to run the simulations of course, hence this is just a benchmark to compare the Swaption approximations to. Overall in both tables we can see that the model provided price is within one standard deviation of the simulation. The only place where this starts to get violated is when the number of paths is 10 million for the one year maturity with the error being consistently around 0.0008 which is not very significant. The reason it might be roughly the same along each row is that the Swaptions are valued simultaneously at the end of each simulated path introducing this bias. For the longer maturity all errors stay within one standard deviation but they still exhibit high correlation, again due to the same bias.

	Strikes	0.0227	0.0242	0.0252	0.0262	0.0277
	Model Prices	1.35035	1.17549	1.06677	0.964386	0.822772
		Monte Carlo				
$N_{\text{paths}} = 10000$	Price	1.34496	1.16912	1.06003	0.957508	0.816356
	Std.	0.01605	0.01507	0.01441	0.013729	0.012702
	Error	0.00539	0.00637	0.00674	0.006878	0.006416
$N_{\text{paths}} = 100000$	Price	1.35307	1.1783	1.06945	0.967031	0.825573
	Std.	0.0051	0.0048	0.00459	0.00437	0.004044
	Error	0.00272	0.00281	0.00268	0.002645	0.002801
$N_{\text{paths}} = 1000000$	Price	1.35116	1.17629	1.06761	0.965276	0.823622
	Std.	0.00161	0.00151	0.00144	0.001378	0.001276
	Error	0.00081	0.0008	0.00084	0.00089	0.00085
$N_{\text{paths}} = 10000000$	Price	1.35052	1.17564	1.06691	0.964533	0.822953
	Std.	0.00051	0.00048	0.00046	0.000435	0.000403
	Error	0.00017	0.00015	0.00014	0.000147	0.000181

Table 6.2: Linear Cheyette model produced European Swaption prices across strikes for fixed maturity $t = 3.712$ by Swaption approximation and Monte Carlo simulations

Remark 6.2.1. It is important to note that even though the Swaption approximation is heavily used when calibrating the model, it does not mean that the Monte Carlo simulation ran on the calibrated model should reproduce these prices. It should give the "true" European Swaption prices implied by the dynamics with the given parameters. Our findings in fact show that the approximations made in Section 2.5.1 are completely justified as the prices they produce lie very close to "true" price within the model.

6.3 Bermudan Swaption Results

Finally we turn to pricing Bermudan Swaptions. We are going to use the calibrated Linear Cheyette model from Section 3.3 with parameters in Table 3.1 for the evaluation and compare them with the Hull-White model from the same section. We pick Bermudans that have tenor structure aligning with the co-terminal Swaptions used for calibration and are all at-the-money. We consider different exercise structures starting from the most frequent. Our expectations are the following.

- The less frequent the exercise dates are, the less dollar amount the Bermudan is worth given the reduced optionality overall.
- Similarly Bermudan Swaptions should be worth more in dollar amount than any of the co-terminal European Swaptions but not more than the sum of Europeans for each exercise date of the Bermudan. Otherwise we could create an arbitrage by selling the Bermudan and buying a bundle of European Swaptions with the same exercise dates.

We present our results in Table 6.3 where the call dates of the Bermudans have the same term structure as our co-terminal Swaptions in Section 3.3. We display the first call at the maturities in the table and then all further call dates are the remaining maturities for the co-terminal Swaptions.

First Call Date	1.000	1.247	1.493	1.740	1.986	2.233	2.479
Linear Cheyette Price	9.2994	8.8371	8.3521	7.8620	7.3459	6.8408	6.3932
Hull-White Price	9.3008	8.8384	8.3534	7.8633	7.3472	6.8420	6.3943
Difference	-0.0014	-0.0013	-0.0014	-0.0013	-0.0013	-0.0012	-0.0011

Table 6.3: ATM Bermudan Swaption prices given by the Linear Cheyette and Hull-White models calibrated to the same set of co-terminal Swaptions

We can see that as expected the prices are decreasing consistently as we remove available call dates and the prices are less than the cost of the individual Europeans making up the Bermudan. We observe the Hull-White prices to be consistently over the Linear Cheyette by a roughly 0.0013 for all the different exercise structures. We compare these results with the two different calibrations of the Linear Cheyette model we saw in Section 3.3, one for ITM and other for OTM European Swaptions. The results are presented in Table 6.4

First Call Date		1.000	1.247	1.493	1.740	1.986	2.233	2.479
OTM Cheyette	Price	9.3010	8.8388	8.3539	7.8636	7.3475	6.8422	6.3945
	Diff	-0.0016	-0.0017	-0.0018	-0.0017	-0.0015	-0.0014	-0.0013
ITM Cheyette	Price	9.2994	8.8369	8.3517	7.8616	7.3456	6.8404	6.3928
	Diff	0.0001	0.0003	0.0004	0.0004	0.0004	0.0004	0.0004

Table 6.4: ATM Bermudan Swaption prices given by the Linear Cheyette models calibrated to 2-2 ITM and OTM European Swaptions compared to the prices of the model calibrated to the whole smile

We can see that again both models mostly agree with the prices of the Linear Cheyette calibrated to the whole smile. The main difference we see in these tables is how the Bermudan price is affected by the slope of the volatility smile. For a steeper smile at the same ATM level we can lower prices whereas for the flatter ones the price increases giving the highest values in case of the OTM calibrated Cheyette. We note that the prices do not seem to differ too much for the Hull-White model which is not too surprising given it is realised when Cheyette has $b_r(t) \equiv 0$ which was approximately the case in these applications. This further shows that the projection works well as the prices produced close to when $y(t)$ is deterministic agree with the Hull-White model.

Conclusion

Throughout this thesis we have been investigating the differences between the Hull-White and Linear Cheyette models in the setting of pricing different type of Swaptions. We can safely conclude that the latter is an extension of the former both in the mathematical and the model capabilities sense. In terms of calibration, we showed that Hull-White is only able to produce one type of volatility smile for European Swaptions and tweaking the parameters can only apply a parallel shift. For the Linear Cheyette model the additional parameter in the local volatility function gives the extra degree of freedom to adjust the slope of the smile as well. However, there are still limitations on the slopes we can achieve without the risk of getting negative volatilities and failing the Swaption approximations applied on the way. A possible further area of research is to replace the linear local volatility function with a different function that is non-negative for any state. One such function with similar properties could be a truncated linear function that does not go below a small $\epsilon > 0$ or the sigmoid function which also introduces a volatility cap that could be desired in some applications.

We further believe that the Swap rate approximation given in Section 2.5.1 while providing great results against the Monte Carlo simulation for the Linear Cheyette model might be insufficient for more general volatility functions with more than 2 parameters. The reason behind this is that a displaced log-normal variable itself only has two parameters and can only provide the linear smiles seen in Section 3.3. If we take a more general, say quadratic volatility function, we would want to fit to more than two parameters at each expiry. For this we propose that the displaced log-normal to be replaced by a higher approximate of the general local volatility, for example, for a quadratic local volatility function with a second order approximation of the Swap rate dynamics.

Finally, we think the capabilities of Markovian projection are overall positive. In particular, the recursive approach is very efficient numerically and allows for a quicker PDE application. However, we do propose further research on its performance against a two dimensional solver in terms of both complexity and numerical accuracy benchmarked to market prices.

Appendix A

Interest Rate Model Derivations

A.1 Hull-White Dynamics

A.1.1 General Derivation and Zero Coupon Bond Price

Here we provide details of the derivation of the Hull-White model dynamics. Start with the Ornstein-Uhlenbeck process

$$dr(t) = (\theta(t) - \kappa(t)r(t))dt + \sigma(t)dW_t^{\mathbb{Q}}$$

Then let $r(t) = x(t) + \varphi(t)$ so that

$$\begin{aligned} dx(t) &= -\kappa(t)x(t)dt + \sigma(t)dW_t^{\mathbb{Q}}, & x(0) &= 0, \\ d\varphi(t) &= (\theta(t) - \kappa(t)\varphi(t))dt, & \varphi(0) &= r(0). \end{aligned}$$

Now we solve for $x(t)$. Using an integrating factor $e^{\int_0^t \kappa(u)du}$ we have that

$$\begin{aligned} z(t) &= x(t)e^{\int_0^t \kappa(u)du} \\ dz(t) &= \kappa(t)x(t)e^{\int_0^t \kappa(u)du}dt + e^{\int_0^t \kappa(u)du}dx(t) \\ &= \kappa(t)z(t) - \kappa(t)z(t) + e^{\int_0^t \kappa(u)du}\sigma(t)dW_t^{\mathbb{Q}} \\ &= \sigma(t)e^{\int_0^t \kappa(u)du}dW_t^{\mathbb{Q}}. \end{aligned}$$

Hence we get

$$\begin{aligned} x(t)e^{\int_0^t \kappa(u)du} &= x(s)e^{\int_0^s \kappa(u)du} + \int_s^t \sigma(u)e^{\int_0^u \kappa(v)dv}dW_u^{\mathbb{Q}} \\ x(t) &= x(s)e^{-\int_s^t \kappa(u)du} + \int_s^t \sigma(u)e^{-\int_u^t \kappa(v)dv}dW_u^{\mathbb{Q}} \end{aligned}$$

For $\varphi(t)$ we can perform the same calculation but instead of a diffusion term there is a drift term independent of $\varphi(t)$. So we get

$$\varphi(t) = \varphi(s)e^{-\int_s^t \kappa(u)du} + \int_s^t \theta(u)e^{-\int_u^t \kappa(v)dv}du.$$

Putting the two together gives the derived short rate (2.1.4):

$$r(t) = r(s)e^{-\int_s^t \kappa(u)du} + \varphi(t) - \varphi(s)e^{-\int_s^t \kappa(u)du} + \int_s^t \sigma(u)e^{-\int_u^t \kappa(v)dv}dW_u^{\mathbb{Q}}$$

In order to determine bond prices instead of directly integrating the above we integrate $x(t)$:

$$\int_t^T x(u)du = \int_t^T x(t)e^{-\int_t^u \kappa(v)dv}du \int_t^T \int_t^u \sigma(v)e^{-\int_t^v \kappa(w)dw}dW_v^{\mathbb{Q}}du$$

Let

$$G(t,T) = \int_t^T e^{-\int_t^u \kappa(v)dv} du.$$

Then using Fubini with the appropriate limit changes

$$\begin{aligned}\int_t^T x(u)du &= x(t)G(t,T) + \int_t^T \int_v^T \sigma(v) e^{-\int_t^v \kappa(w)dw} du dW_v^{\mathbb{Q}} \\ &= x(t)G(t,T) + \int_t^T \sigma(u)G(u,T)dW_u^{\mathbb{Q}}\end{aligned}$$

So then simply for the integral of the short rate

$$\int_t^T r(u)du = x(t)G(t,T) + \int_t^T \varphi(u)du + \int_t^T \sigma(u)G(u,T)dW_u^{\mathbb{Q}}$$

Hence it is easy to see that the above is normally distributed $\mathcal{N}(\mu, \sigma^2)$ with

$$\begin{aligned}\mu &= x(t)G(t,T) + \int_t^T \varphi(u)du, \\ \sigma^2 &= \int_t^T \sigma^2(u)G(u,T)^2 du = V(t,T).\end{aligned}$$

A.1.2 Reconstructing Parameters

In this section we will reverse engineer the values for $\varphi(t)$ and $\theta(t)$ from the derivation of the bond formula in (2.1.6). We start with

$$-\ln P(0,T) = \int_0^T \varphi(u)du - \frac{1}{2}V(0,T),$$

then we differentiate both sides to get

$$\begin{aligned}f^M(0,T) &= -\frac{\partial \ln P(0,T)}{\partial T} \\ &= \frac{\partial}{\partial T}\int_0^T \varphi(u)du - \frac{1}{2}\frac{\partial}{\partial T}\left(\int_0^T \sigma^2(u)G^2(u,T)du\right) \\ &= \varphi(T) - \frac{1}{2}\left(\sigma^2(T)G^2(T,T) + \int_0^T \sigma^2(u)\frac{\partial}{\partial T}G^2(u,T)du\right) \\ &= \varphi(T) - \frac{1}{2}\int_0^T \sigma^2(u)2G(u,T)\frac{\partial}{\partial T}G(u,T)du\end{aligned}$$

Rearranging gives

$$\varphi(t) = f^M(0,t) + \int_0^t \sigma^2(u)e^{-\int_u^t \kappa(v)dv}G(u,t)du$$

We can similarly determine an expression for $\theta(t)$.

$$\begin{aligned}P(t,T) &= \mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T r(u)du\right)\middle|\mathcal{F}_t\right] \\ &= \exp\left(-r(t)G(t,T) - \int_t^T \theta(u)G(u,T)du\right)\mathbb{E}^{\mathbb{Q}}\left[\exp\left(-\int_t^T \sigma(u)G(u,T)dW_t^{\mathbb{Q}}\right)\middle|\mathcal{F}_t\right] \\ &= \exp\left(-r(t)G(t,T) - \int_t^T \theta(u)G(u,T)du + \frac{1}{2}V(t,T)\right)\end{aligned}$$

we can then write

$$\begin{aligned} f^M(0,t) &= -\frac{\partial \ln P(0,t)}{\partial t} \\ &= r(0)e^{-\int_0^t \kappa(u)du} + \int_0^t \theta(u)e^{-\int_u^t \kappa(v)dv}du - \frac{1}{2}\frac{\partial V(0,t)}{\partial t} \end{aligned}$$

we can differentiate both sides of this expression to get $\theta(t)$

$$\theta(t) = \frac{\partial f^M(0,t)}{\partial t} + \kappa(t)f^M(0,t) + \frac{1}{2}\left(\kappa(t)\frac{\partial V}{\partial t}(0,t) + \frac{\partial^2 V}{\partial t^2}(0,t)\right)$$

A.2 Cheyette Dynamics

Here we provide details of the derivation of the general Cheyette dynamics discussed in Section 2.3. We start from a general HJM framework with a separable volatility function:

$$\begin{aligned} df(t,T) &= \sigma_f(t,T)\left(\int_t^T \sigma_f(t,u)du\right)dt + \sigma_f(t,T)dW_t^{\mathbb{Q}} \qquad \text{(A.2.1)} \\ \sigma_f(t,T,\omega) &= g(t,\omega)h(T) \end{aligned}$$

From now on we omit ω from the formulas to ease notation. We directly integrate (A.2.1) to get $f(t,T)$ as in (2.3.2):

$$\begin{aligned} f(t,T) &= f(0,T) + \int_0^t df(s,T) \\ &= f(0,T) + \int_0^t \sigma_f(s,T)\left(\int_s^T \sigma_f(s,u)du\right)ds + \int_0^t \sigma_f(s,T)dW_s^{\mathbb{Q}} \\ &= f(0,T) + \int_0^t g(s)h(T)\left(\int_s^T g(s)h(u)du\right)ds + \int_0^t g(s)h(T)dW_s^{\mathbb{Q}} \\ &= f(0,T) + h(T)\int_0^t g^2(s)\left(\int_s^T h(u)du\right)ds + h(T)\int_0^t g(s)dW_s^{\mathbb{Q}} \qquad \text{(A.2.2)} \end{aligned}$$

Let $x(t) = f(t,t) - f(0,t) = r(t) - f(0,t)$, then

$$\begin{aligned}
dx(t) &= d\left(h(t)\int_0^t g^2(s)\left(\int_s^t h(u)du\right)ds + h(t)\int_0^t g(s)dW_s^{\mathbb{Q}}\right) \\
&= \left(h(t)\int_0^t g^2(s)\left(\int_s^t h(u)du\right)ds\right)'dt + \left(h(t)\int_0^t g(s)dW_s^{\mathbb{Q}}\right)'dt \\
&= h'(t)\int_0^t g^2(s)\left(\int_s^t h(u)du\right)dsdt + h(t)\left(\int_0^t g^2(s)\left(\int_s^t h(u)du\right)ds\right)'dt \\
&\quad + h'(t)\int_0^t g(s)dW_s^{\mathbb{Q}}dt + h(t)g(t)W_t^{\mathbb{Q}} \\
&= h'(t)\left(\int_0^t g^2(s)\left(\int_s^t h(u)du\right)ds\right)dt + h(t)\Bigg(g^2(t)\overbrace{\left(\int_t^t h(u)du\right)}^{0} \\
&\quad + \int_0^t g^2(s)\left(\int_s^t h(u)du\right)'ds\Bigg)dt + h'(t)\int_0^t g(s)dW_s^{\mathbb{Q}}dt + h(t)g(t)W_t^{\mathbb{Q}} \\
&= h'(t)\left(\int_0^t g^2(s)\left(\int_s^t h(u)du\right)ds\right)dt + h^2(t)\left(\int_0^t g^2(u)du\right)dt \\
&\quad + h'(t)\int_0^t g(s)dW_s^{\mathbb{Q}}dt + h(t)g(t)W_t^{\mathbb{Q}} \\
&= \left(\frac{h'(t)}{h(t)}x(t) + h^2(t)\int_0^t g^2(u)du\right)dt + h(t)g(t)dW_t^{\mathbb{Q}} \\
&= \left(y(t) - \kappa(t)x(t)\right)dt + \sigma_r(t)dW_t^{\mathbb{Q}} \qquad \text{(A.2.3)}
\end{aligned}$$

as in (2.3.4), making the substitutions

$$\begin{aligned}
\sigma_r(t) &= h(t)g(t) = \sigma_f(t,t), \\
\kappa(t) &= -\frac{h'(t)}{h(t)}, \\
y(t) &= h^2(t)\int_0^t g^2(u)du = \int_0^t \sigma_r^2(u)\exp\left(-2\int_u^t \kappa(s)ds\right)du. \qquad \text{(A.2.4)}
\end{aligned}$$

Using that we can invert relationship between $\kappa(t)$ and $h(t)$ to get

$$h(t) = \exp\left(-\int_0^t \kappa(u)du\right)$$

Notice that (A.2.4) solves the following ODE

$$dy(t) = \left(\sigma_r^2(t) - 2\kappa(t)y(t)\right)dt, \quad y(0) = 0$$

It is clear that we arrive at the Cheyette dynamics in (2.3.6) putting this together with (A.2.3). We continue with deriving the zero coupon bond price formula. From (A.2.2)

$$\begin{aligned}
f(t,T) &= f(0,T) + h(T)\int_0^t g^2(s)\left(\int_s^T h(u)du\right)ds + h(T)\int_0^t g(s)dW_s^{\mathbb{Q}} \\
&= f(0,T) + \frac{h(T)}{h(t)}\left(h(t)\int_0^t g^2(s)\left(\int_s^t h(u)du\right)ds + h(t)\int_0^t g(s)dW_s^{\mathbb{Q}}\right) \\
&\quad + h(T)\int_0^t g^2(s)\left(\int_t^T h(u)du\right)ds \\
&= f(0,T) + \frac{h(T)}{h(t)}x(t) + \frac{h(T)}{h^2(t)}y(t)\left(\int_t^T h(u)du\right) \\
&= f(0,T) + \frac{h(T)}{h(t)}\left(x(t) + \frac{y(t)}{h(t)}\int_t^T h(u)du\right) \qquad \text{(A.2.5)}
\end{aligned}$$

as in (2.3.7). Now for convenience define

$$H(t,T) = \int_t^T h(u)du$$
$$G(t,T) = \frac{\int_t^T h(u)du}{h(t)} = \frac{H(t,T)}{h(t)}$$

Then integrating (A.2.5)

$$\begin{aligned}
\int_t^T f(t,u)du &= \int_t^T f(0,u)du + \int_t^T \frac{h(u)}{h(t)}x(t)du + \int_t^T \frac{y(t)h(u)}{h^2(t)}\left(\int_t^u h(s)ds\right)du \\
&= \int_t^T f(0,u)du + G(t,T)x(t) + \frac{y(t)}{h^2(t)}\int_t^T \underbrace{h(u)}_{H'(t,u)}\underbrace{\left(\int_t^u h(s)ds\right)}_{H(t,s)}du \\
&= \int_t^T f(0,u)du + G(t,T)x(t) + \frac{y(t)}{h^2(t)}\frac{1}{2}H^2(t,T) \\
&= \int_0^T f(0,u)du - \int_0^t f(0,u)du + G(t,T)x(t) + \frac{1}{2}y(t)G^2(t,T)
\end{aligned}$$

Using the formula for $P(t,T)$ in (1.1.5) on the above we get

$$\begin{aligned}
P(t,T) &= \exp\left(-\int_0^T f(0,u)du + \int_0^t f(0,u)du - G(t,T)x(t) - \frac{1}{2}y(t)G^2(t,T)\right) \\
&= \frac{P(0,T)}{P(0,t)}\exp\left(-G(t,T)x(t) - \frac{1}{2}y(t)G^2(t,T)\right)
\end{aligned}$$

which is just as in (2.3.8).

A.3 Proof of Proposition 2.3.4

Proof. Given $g(t)$ we know that we have

$$\sigma_r(t) = g(t)h(t) = \sigma(t)$$

which makes the expression for $y(t)$ deterministic and directly computable

$$y(t) = \int_0^t \sigma_r^2(u)\exp\left(-2\int_u^t \kappa(s)ds\right)du$$

It suffices to show that the short rate has the same dynamics if the zero coupon bond prices are the same given we condition on the same short rate at time t. We have

$$r(t) = x_{\text{Cheyette}}(t) + f^M(0,t) = x_{\text{Hull-White}}(t) + \alpha(t) + f^M(0,t)$$

with bond prices

$$\begin{aligned}
P_{\text{Cheyette}}(t,T,x,y) &= \frac{P(0,T)}{P(0,t)}\exp\left(-x_{\text{Cheyette}}(t)G(t,T) - \frac{1}{2}y(t)G^2(t,T)\right) \\
P_{\text{Hull-White}}(t,T,x) &= \frac{P(0,T)}{P(0,t)}\exp\left(-x_{\text{Hull-White}}(t)G(t,T) + \frac{1}{2}[V(t,T) - V(0,T) + V(0,t)]\right)
\end{aligned}$$

Making these two equal and taking $x_{\text{Hull-White}}(t) = x_{\text{Cheyette}}(t) - \alpha(t) \equiv x$ it remains to show

$$-\alpha(t)G(t,T) - \frac{1}{2}y(t)G^2(t,T) = \frac{1}{2}[V(t,T) - V(0,T) + V(0,t)]$$

Substituting in for the left hand side we get

$$
\begin{aligned}
&- G(t,T)\int_0^t \sigma^2(u)\exp\left(-\int_u^t \kappa(v)dv\right)G(u,t)du - \frac{1}{2}G^2(t,T)\int_0^t \sigma^2(u)\exp\left(-2\int_u^t \kappa(s)ds\right)du \\
&= -G(t,T)\int_0^t \sigma^2(u)\frac{h(t)}{h(u)}G(u,t)du - \frac{1}{2}G^2(t,T)\int_0^t \sigma^2(u)\frac{h^2(t)}{h^2(u)}du
\end{aligned}
$$

For right hand side

$$
\begin{aligned}
&\frac{1}{2}\left[\int_t^T \sigma^2(u)G(u,T)^2du - \int_0^T \sigma^2(u)G(u,T)^2du + \int_0^t \sigma^2(u)G(u,t)^2du\right] \\
&\qquad = \frac{1}{2}\left[\int_0^t \sigma^2(u)G(u,t)^2du - \int_0^t \sigma^2(u)G(u,T)^2du\right]
\end{aligned}
$$

Let us expand the following

$$
\begin{aligned}
&\frac{1}{2}\int_0^t \sigma^2(u)G(u,T)^2du = \frac{1}{2}\int_0^t \sigma^2(u)\frac{\left(\int_u^T h(v)dv\right)^2}{h^2(u)}du \\
&= \frac{1}{2}\int_0^t \sigma^2(u)\frac{\left(\int_u^t h(v)dv + \int_t^T h(v)dv\right)^2}{h^2(u)}du \\
&= \frac{1}{2}\int_0^t \sigma^2(u)\frac{\left(\int_u^t h(v)dv\right)^2}{h^2(u)}du + \int_0^t \sigma^2(u)\frac{\left(\int_u^t h(v)dv\right)\left(\int_t^T h(v)dv\right)}{h^2(u)}du + \frac{1}{2}\int_0^t \sigma^2(u)\frac{\left(\int_t^T h(v)dv\right)^2}{h^2(u)}du \\
&= \frac{1}{2}\int_0^t \sigma^2(u)G^2(u,t)du + \int_0^t \sigma^2(u)G(u,t)G(t,T)\frac{h(t)}{h(u)}du + \frac{1}{2}\int_0^t \sigma^2(u)G^2(t,T)\frac{h^2(t)}{h^2(u)}du
\end{aligned}
$$

Now notice that negating both sides and rearranging gives the identity we wanted. □

Appendix B

Markovian Projection Derivations

B.1 Proof of Lemma 4.2.2

Proof. Let us first deal with the cases of $m(t)$, $S(t)$ and $\text{Cov}[x, x+dx]$, once we have these values then invoking Lemma 4.2.1 with some rearranging will yield the result. We get the following ordinary differential equation:

$$\begin{aligned} dm(t) &= \mathbb{E}[x(t) + dx(t)] - \mathbb{E}[x(t)] \\ &= m(t) + \mathbb{E}[y(t) - \kappa(t)x(t)]dt + \mathbb{E}[\lambda(t)(a(t) + b(t)x(t))dW_t] - m(t) \\ &= (\overline{y}(t) - \kappa(t)m(t))dt. \end{aligned}$$

After approximating $\mathbb{E}[y(t)]$ by (2.4.5). We have boundary condition $m(0) = 0$ from boundary conditions in 2.3.6 so the unique solution for $m(t)$ is:

$$m(t) = \int_0^t \overline{y}(s) e^{-\int_s^t \kappa(u)du} ds,$$

as required.
Similarly for the variance we can write

$$\begin{aligned} dS(t) &= \mathbb{E}\left[(x(t) - m(t) + dx(t) - dm(t))^2\right] - \mathbb{E}\left[(x(t) - m(t))^2\right] \\ &= \mathbb{E}\left[(dx(t))^2 + (dm(t))^2 + 2dx(t)(x(t) - m(t)) - 2dm(t)(x(t) - m(t))\right] \\ &= \mathbb{E}\left[\sigma_r^2(t, x(t))dt + 2\big(x(t)\overline{y}(t) - \kappa(t)x^2(t) - m(t)\overline{y}(t) + m(t)\kappa(t)x(t) \right. \\ &\quad \left. + x(t)\kappa(t)m(t) - x(t)\overline{y}(t) - \kappa(t)m^2(t) + m(t)\overline{y}(t)\big)dt\right] \\ &= \mathbb{E}\left[\sigma_r^2(t, x(t))\right] dt - 2\kappa(t)\mathbb{E}\left[(x(t) - m(t))^2\right] dt. \end{aligned}$$

Expanding on $\mathbb{E}\left[\sigma_r^2(t, x(t))\right]$ as it is a linear local volatility we have

$$\begin{aligned} \mathbb{E}\left[\sigma_r^2(t, x(t))\right] &= \mathbb{E}\left[\lambda^2(t)(a^2(t) + 2a(t)b(t)x(t) + b^2(t)x^2(t))\right] \\ &= \lambda^2(t)(a^2(t) + 2a(t)b(t)m(t) + b^2(t)(S(t) + m^2(t))) \\ &= \lambda^2(t)b^2(t)S(t) + \sigma_r^2(t, m(t)). \end{aligned}$$

Then putting it all together we get

$$\begin{aligned} \frac{dS(t)}{dt} &= (\lambda^2(t)b^2(t) - 2\kappa(t))S(t) + \sigma_r^2(t, m(t)), \\ S(0) &= 0 \end{aligned}$$

Where the boundary condition again comes from the boundary conditions of the process $x(t)$. The unique solution is given by

$$S(t) = \int_0^t \sigma^2(s, m(s)) e^{\int_s^t \lambda^2(u)b^2(u) - 2\kappa(u)du} ds,$$

as required. Finally for the covariance we can write

$$\begin{aligned}
\mathrm{Cov}[x(t), x(t)+dx(t)] &= \mathbb{E}[x(t)(x(t)+dx(t))] - \mathbb{E}[x(t)]\mathbb{E}[x(t)+dx(t)] \\
&= \mathbb{E}[x^2(t)] + \mathbb{E}[x(t)dx(t)] - \mathbb{E}[x(t)]^2 - \mathbb{E}[x(t)]\mathbb{E}[dx(t)] \\
&= S(t) + [m(t)\overline{y}(t) - \kappa(t)(S(t)+m^2(t))]dt - m(t)[\overline{y}(t) - \kappa(t)m(t)]dt \\
&= S(t) - \kappa(t)S(t)dt,
\end{aligned}$$

which yields the final result. □

Bibliography

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

[2] Damiano Brigo, Fabio Mercurio, et al. *Interest rate models: theory and practice*, volume 2. Springer, 2001.

[3] L.B.G. Andersen and V.V. Piterbarg. *Interest Rate Modeling*. Number v. 2 in Interest Rate Modeling. Atlantic Financial Press, 2010.

[4] 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.

[5] Damiano Brigo. Interest rate modelling. Lecture notes, Imperial College London, 2021-2022.

[6] Alex Tse. Topics in derivatives pricing. Lecture notes, Imperial College London, 2021-2022.

[7] Farshid Jamshidian. An exact bond option formula. *The journal of Finance*, 44(1):205–209, 1989.

[8] L.B.G. Andersen and V.V. Piterbarg. *Interest Rate Modeling*. Number v. 1 in Interest Rate Modeling. Atlantic Financial Press, 2010.

[9] O Cheyette. Markov representation of the heath-jarrow-morton model (working paper). *Berkeley: BARRA Inc*, 1994.

[10] Kenneth Levenberg. A method for the solution of certain non-linear problems in least squares. *Quarterly of applied mathematics*, 2(2):164–168, 1944.

[11] Patrick S Hagan, Deep Kumar, Andrew S Lesniewski, and Diana E Woodward. Managing smile risk. *The Best of Wilmott*, 1:249–296, 2002.

[12] I. Gyöngy. Mimicking the one-dimensional marginal distributions of processes having an ito differential. *Probability Theory and Related Fields*, 71(4):501–516, 1986.