MSc Econometrics and Management Science

Quantitative Finance Track

Master Thesis

Interpolation of Forward-Looking and Backward-Looking Forward Rates in the Forward Market Model

Author:
Olivier Mulkin

Student ID:
503315

Academic Supervisor:
Dr. X. Xiao

Second Assessor:
Prof. Dr. M. van der Wel

Company Supervisor:
Dr. R. Pietersz

Erasmus University Rotterdam

Erasmus School of Economics

August 31, 2020

Abstract

This thesis is mainly concerned with the appropriate treatment of backward-looking forward rates, which in the face of interest rate reforms, is emerging as a plausible reference on main interest rate derivatives such as swaps, caps, floors and swaptions. These forward rates differ from traditional discretely compounded forward-looking forward rates in that the payoff value of contingent derivatives can only be determined at the end of the application period, because it is based on the daily compounding of overnight rates in that period. We analyze the Forward Market Model (FMM) presented by Lyashenko and Mercurio (2019), which exhibits the ability to handle the stochastic diffusion of these backward-looking forward rates. Introduced as an extension of the well-known Libor Market Model, it simultaneously retains the ability to diffuse forward-looking rates. In addition to providing mathematical foundations for the FMM, we also attempt to formulate an interpolation scheme capable of providing backward-looking forward rates for non-standard maturities (i.e. not modeled by the FMM). A recurrent challenge faced by interpolation methods in market models is the occurrence of a mispricing due to an artificially low implied volatility of interpolated forward rates. This thesis documents and addresses this challenge.

Table of Content

1 Introduction

The seminal paper of Brace et al. (1997) has disrupted the field of interest rate modeling by proposing the Brace-Gątarek-Musiela or BGM model. This model consists of diffusing market-observable quantities, namely the discrete forward rates, and allowed the simultaneous calibration of the parameters of all state variables through their volatility and correlation structure. The BGM came as a better alternative to existing models of theoretical values such as short rate models (for example, the Hull-White model) which only allowed the calibration of a few parameters and could not entirely replicate the yield curve after calibration. It is also sometimes referred to as the Libor Market Model (LMM) because the quantities typically modeled by the BGM are the benchmark risk-free rates (RFR) in a forward rate agreement, usually the InterBank Offer Rate or IBOR, including the infamous LIBOR (for London IBOR) which is often at the center of interest rate swaps. It is still today one of the most widely used model by financial institutions, especially for the pricing of instruments with multiple rate dependencies such as swaptions (Brigo and Mercurio (2006)), or for risk-management purposes such as the calculation of different value adjustments to the price of financial derivatives grouped under the acronym of XVA (Zeitsch (2017)). One of the main advantage of the LMM is the possiblity to simultaneously calibrate to traded products such as caps, floors and swaptions and for example, specify a desired stochastic covariance structure on the forward rates. Additionally, the fact that each rate dynamics can be specified under a single measure is a powerful feature of the LMM, since financial instruments involving multiple rates can easily be priced.

In recent years, the LIBOR and other IBOR rates have received growing criticism over scandals of market manipulation and over their failure to meet RFR standards set by central banks and government. As a consequence, the task of finding an alternative RFR to replace IBOR rates has been addressed by governments in major economies. For example, the Euro zone has introduced a new unsecured overnight rate called €STR (Euro Short-Term Rate) and has encouraged financial institutions to adopt this reference rate and abandon IBOR rates (European Central Bank (2019)). Similarly, the US has introduced the Secured Overnight Funding Rate (SOFR) and the UK selected the Sterling Overnight Index Average (SONIA) as its RFR. Naturally, transitioning to a new benchmark is a massive and incongruous challenge for banks and other financial institutions as billions worth of notional of financial instruments are linked to the IBOR rates. Not only does the discontinuation of IBOR rates affect derivatives, but they also impact other products that are based on interest rates such as loans, mortgages, etc. For financial derivatives, many such instruments often have maturities that may be very distant in time. As a consequence, the research focus in interest rate modelling has somewhat shifted to analyzing the impact of the IBOR transition (Henrard (2019)) or to the pricing of financial derivatives in this new interest rate environment, i.e. an environment where IBOR fallback options such as the SOFR or €STR are the new benchmarks as IBOR rates are discontinued (Mercurio (2018)). Such rates are called overnight because they are published daily, while IBOR rates are term rates, since they have maturities ranging from 1 month to 1 year. Overnight rates need to be converted to term rates if they are to be used as a replacement of IBOR. This also

means that, in practice, a party engaged in a swap would only know the payment at the end of the application period, when all relevant daily rates have been published. We refer to such rates as backward-looking. One could potentially simulate the daily rates and then compound them in order to estimate the payment structure. However, a paper recently published by Lyashenko and Mercurio (2019) shows that backward-looking rates can be simulated directly with the generalized Forward Market Model (FMM) which they introduce. In fact, this model is said to be hybrid in the sense that it can also simulate the forward-looking rates, using a single stochastic process. This thesis provides mathematical foundation for this model and investigates how interpolation can be applied to backward-looking rates simulated in this model.

The FMM has a very general form which allows the user to specify a great variety of adapted processes for the volatility as well as various distributions of forward rates such as the normal or the lognormal distributions. This thesis will focus on the lognormal case for multiple reasons. Firstly, we want to stay close to the LMM which is a lognormal model. Secondly, lognormal models were introduced to remedy the problem of negative interest rates occurring in other models. Given current economic conditions however, negative interest rates may be desirable, although they are only so when the magnitude of negative values produced by the model is somewhat controllable. Introducing a displacement in the LMM has made that possible, and we will see that the FMM with lognormal specification is no exception. Finally, the Black formula for options remains the standard practice when it comes to pricing financial derivatives but it is only applicable when the stochastic process has a lognormal distribution.

A market model, as Werpachowski (2010) writes, can only be considered complete when an interpolation method has been specified along the model. This practice aims to convert the model from discrete-time to continuous time in the maturity dimension. In other words, forward rates can be obtained for all desired expiry-maturity pairs, not only for the ones specified in the model. This requires the use of interpolation. Naturally, it is possible to adjust the model such that it simulates all the forward rates needed to price caplets, but books of financial institutions can easily contain financial products with a myriad of forward rates, sometimes having a maturity differing by only a few days. This presents a heavy computational burden and interpolation is consequently a better alternative, since it does not require the simulation of additional forward rates. Interpolation has so far only been applied to forward-looking forward rates as backward-looking forward rates (with payments in arrears) just recently emerged from the transition to overnight RFRs. Therefore, one of the main contributions of this thesis will be the study of interpolation methods applied to backward-looking rates in the FMM. It is far from trivial since many interpolation methods are known to produce a pricing issue when applied to the LMM. Specifically, if one plots the average implied volatility of caplets premiums created on interpolated rates as a function of maturity, the resulting line will exhibit an odd, wavy pattern, in someway resembling garlands. This directly translates into a mispricing of caplets whose expiries and maturities coincide with that of the interpolated rates. We will see that this is the case in both forward and backward-looking caplets priced in the FMM. For the forward-looking caplets, a simple solution exists and consists of changing the specification of the volatility structure in the accrual period of forward-looking rates (that is in the simulation phase). This solution bears a striking resemblance with the interpolation method developed by Werpachowski (2010) in

the context of the LMM but still differs given the differences in probability measures used to construct the FMM. For backward-looking rates, this simple solution will not be sufficient on its own. In fact, we develop an alternative interpolation method which differs from the first only by the time at which we observe (fix) the modeled forward rates which are in turn used to construct interpolated rates. With this method, the mispricing is substantially reduced although it does not completely dissipate.

In our attempt to understand the problem of lower implied volatility of "broken dates" caplets (caplet whose expiry and maturity are not those of the modeled forward rates), we derived an approximation to the instantaneous volatility of interpolated forward rates. This in turn can predict the implied volatility of the forward rates. Our primary purpose to derive such an approximation was to find out how we can manipulate our model to produce a smooth (i.e. rid of the garland shape) implied volatility function in backward-looking caplets on interpolated rates. It did not really help us in this regard, but it helped us to conclude that interpolation of backward-looking rates cannot be made optimal by the simple solution that was used in forward-looking cases. Regardless, this approximation was found to be very accurate and provide some insight into the shape of instantaneous volatility of interpolated rates. We also extend this approximation to be valid when the model contains a displacement.

Lastly, we provide a concrete application of the FMM to compute unilateral Credit Value Adjustments (CVA) of backward-looking swaps. CVA is one of the many value adjustments grouped under XVA, briefly mentioned above. It was introduced to incorporate credit risk (risk of default) of counterparties in the valuation of interest rate derivatives entered by those counterparties. As changes in daily CVA greatly affects the Profit and Loss (PnL) of financial institutions, it is important to understand how transitioning to backward-looking rates can affect CVA. Therefore, we compare CVA of forward and backward-looking swaps.

The plan of the thesis is as follows. In the next section, we define the quantities of interest in the FMM, and establish the mathematical equivalence between forward and backward-looking rates, which is the main reason for the introduction of the FMM. We then present the extended HJM framework shortly mentioned in Lyashenko and Mercurio (2019) and derive the FMM from this framework. The third section contains our most important contribution, namely a study of interpolation methods in the FMM and its consequences on caplet pricing. We apply interpolation distinctly to forward and backward-looking rates. We will also derive in that section the approximation of instantaneous volatility of interpolated forward rates. Section 4 provides a CVA application and finally, Section 5 concludes.

2 Lognormal Forward Market Model

The generalized FMM introduced by Lyashenko and Mercurio (2019) presents a very general form for the forward rate dynamics. We may choose to specify stochastic or deterministic diffusion coefficients, normal or lognormal distribution of forward rates, and additional characteristics such as a displacement, or a decaying or rising the volatility of forward rates in the accrual period. In this thesis, we will focus on forward rates following a lognormal distribution with deterministic volatilities (although we consider time-varying volatilities). As briefly explained, the purpose of lognormal models was originally the exclusion of negative interest rates produced in normal models. However, it is still possible to produce negative rates in a lognormal setting with the simple inclusion of a displacement. This technique offers more control over the magnitude of simulated negative rates as it effectively imposes a lower bound on these simulated quantities. We now provide motivations for the FMM and derive this model starting from an HJM framework.

2.1 Hybrid model for two types of rates

The main strength of the FMM when compared to the LMM is its additional feature to model backward-looking forward rates, while retaining the strength of the LMM that initially made it powerful. These strengths revolve around its ability to incorporate joint dependencies between many rates under a single probability measure. Many interest rate products often have payoffs that depend on multiple forward rates that, for instance, differ in the only maturity dimension (vanilla interest rate swaps) or the currency dimension (Cross-currency swaps). Simultaneously, if we "slice" the model, under their respective forward measure, each forward rate follows a Black model and an analytical formula to price most standard interest rate derivatives exists under this model (Black (1976)).

It can be mathematically shown that forward and backward-looking forward rates are identical before the accrual period, as defined by the period starting at the forward rate expiry and ending at the forward rate maturity. At the expiry, the forward-looking forward rates cease to exist but the backward-looking rates continues to evolve, with a different volatility structure if desired. Upon reaching their maturity, the forward rates become constant. To bring this into perspective, we first explain the notion of forward and backward-looking rates. We also define the quantities of interest and present a set of assumptions on which the FMM will be constructed. We do so by closely following the approach of Lyashenko and Mercurio (2019).

2.1.1 Terminology and assumptions

We assume the existence of a short-rate process $r(t)$ along with a so-called money-market account that has value $B(t)$ at time $t \geq 0$. The money-market account evolves according to

$$dB(t) = r(t)B(t)dt \tag{2.1}$$

with initial condition $B(0) = 1$ and solution

$$B(t) = e^{\int_0^t r(u)du} \tag{2.2}$$

The rate $r(t)$ is stochastic in nature and the money-market account is a positive asset that is part of a continuous-time economy. We start with a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where Ω is the sample space. $\mathcal{F}$ is a σ-algebra of measurable events, with filtration $\mathcal{F}_t, 0 \leq t \leq T$ representing the information available up to time t. Furthermore, information increases with time, i.e. $\mathcal{F}_s \subseteq \mathcal{F}_t$ for $s \leq t$. $\mathbb{P}$ is a probability measure under which we have a N-dimensional Brownian motion ($N > 0$) generating the filtration $\mathcal{F}_t$, with possibly correlated elements denoted by $W_k(t)$ for $k = 1, \ldots, N$. We denote by $\mathbb{E}^{\mathbb{P}}$ the expectation operator under the measure $\mathbb{P}$.

For now, we also assume the existence of an equivalent risk-neutral probability measure $\mathbb{Q}$ with associated numeraire $B(t)$. We will formally construct this measure with the use of Girsanov's theorem later in this text. This measure is risk-neutral for the numeraire $B(t)$ because asset prices expressed in terms of (discounted by) the money-market account are martingales under this measure.

We consider a set of standard tenor dates $0 = T_0 < T_1 < ... < T_M < \infty$ (the tenor structure) from which we take expiry-maturity pairs $(T_{j-1}, T_j), j = 1, \ldots, M$, to span a set of forward rates. Furthermore, denote by $\tau_j := \tau(T_{j-1}, T_j)$ the year fraction separating T_{j-1} and T_j. Unless specified otherwise, we consider the case where all τ_j are equal to γ.

A typical assumption to construct the Forward/Libor Market Model is the existence of a frictionless bond market for each maturity date.

Assumption 1. *There exists a friction-less market for the zero-coupon bond with maturity T, for every maturity $T \in [0, \infty)$, and whose price at time t is denoted by $P(t, T)$. In particular, T may be a date from the standard maturity set $[T_0, T_1, ..., T_M]$. The bond price $P(t, T)$ is a strictly positive adapted process.*

We define the notion of extended bond-prices needed to construct the FMM. It has been previously used in the literature (see Andersen and Piterbarg (2010))

Definition 2.1.1 (Extended Bond Price). *The extended bond price $P(t, T)$ is defined for all time $t \geq 0$. Before maturity, it is the price of a zero-coupon bond paying 1 at maturity, i.e. $P(T, T) = 1$. The bond price is extended such that, for $t > T$, it is defined to be the value of the strategy where the initial proceed is continuously reinvested at the prevailing short rate, i.e.,*

$$P(t, T) = e^{\int_T^t r(t)dt}$$

This definition turns out to be consistent with the arbitrage-free bond price under the risk-neutral measure $\mathbb{Q}$

$$P(t, T) = \mathbb{E}^{\mathbb{Q}} \left[e^{-\int_t^T r(u)du} \,\middle|\, \mathcal{F}_t \right] \tag{2.3}$$

and therefore, for $t > T$, we have

$$P(t, T) = \mathbb{E}^{\mathbb{Q}} \left[e^{\int_T^t r(u)du} \,\middle|\, \mathcal{F}_t \right] = e^{\int_T^t r(u)du} = \frac{B(t)}{B(T)} \tag{2.4}$$

As this extended bond price process is strictly positive and is a non-dividend paying asset, the Theorem of Stochastic Representation of Assets applies and the extended bond price is a viable numeraire to be used with Girsanov's theorem (see chapter 9 in Shreve (2008)).

The forward rate is typically defined by considering the following strategy at time $t \leq T_{j-1}$: Short one unit of a T_{j-1}-maturing bond and buy $\frac{P(t,T_{j-1})}{P(t,T_j)}$ of T_j-maturing bond. The cost of setting up this strategy at t is obtained by multiplying the positions with the prices:

$$-1 \times P(t,T_{j-1}) + \frac{P(t,T_{j-1})}{P(t,T_j)} \times P(t,T_j) = 0.$$

At time T_{j-1}, the short position in the T_{j-1}-maturing bond generates a cash outflow of 1 and, at time T_j, the long position in the other bond generates an inflow of $\frac{P(t,T_{j-1})}{P(t,T_j)}$. The yield over that period is referred to as the simply compounded forward rate.

Definition 2.1.2. *The simply compounded forward rate at time t with expiry S and maturity T is given by*

$$F(t,S,T) = \frac{1}{T-S}\left(\frac{P(t,S)}{P(t,T)} - 1\right) \tag{2.5}$$

For brevity when dealing with forward rates on maturity dates of our tenor structure T_j with $j = 1,\ldots,M$, we denote $F(t,T_{j-1},T_j)$ by $F_j(t)$.

These cashflows are identical with those in a Forward Rate Agreement (FRA) initiated at par, so the fixed rate in a FRA corresponds to (2.5). The forward rate $F_j(t)$ is the time-t interest rate one can earn over the future period (T_{j-1},T_j) and these are the quantities that are modeled in the FMM. Current market conditions display negative market rates in some financial instruments. As a result, it is desirable to incorporate negative rates in our model. We therefore add a displacement to the modeled forwards. The resulting dynamics are closely related to the ones of the displaced diffusion LMM (see for instance Beveridge and Joshi (2012)).

For a given displacement parameter $\delta \in \mathbb{R}$, we define the displaced forward rate by

$$\widetilde{F}(t,S,T) := F(t,S,T) + \delta \tag{2.6}$$

To continue our set of assumptions to construct our HJM framework, we recall a well-known theorem on martingales. The concept of a martingale is essentially a condition on the expectations of a random variable. The theorem is a slightly restated version of one that can be found in Brigo and Mercurio (2006).

Theorem 2.1.3. *For an equivalent probability measure $\mathbb{P}^N$ and a filtration $\mathcal{F}_t$, a random variable Y_t is a martingale if the expectation under that measure exists and is finite, i.e. $\mathbb{E}^{\mathbb{P}^N}[Y_t] < \infty \quad \forall t$ and*

$$Y_t = \mathbb{E}^{\mathbb{P}^N}\left[Y_s \,\middle|\, \mathcal{F}_t\right] \qquad \forall t \leq s$$

Furthermore, if $N(t)$ is the numeraire associated with this probability measure, and let $X(t)$ be the price of an asset, then

$$\frac{X(t)}{N(t)} = \mathbb{E}^{\mathbb{P}^N}\left[\frac{X(s)}{N(s)} \,\middle|\, \mathcal{F}_s\right] \quad 0 \leq t \leq s.$$

From definition 2.1.2, forward rates can be rewritten as

$$F(t, T_{j-1}, T_j)P(t, T_j) = \frac{P(t, T_{j-1}) - P(t, T_j)}{T_j - T_{j-1}} \tag{2.7}$$

where the $F(t, T_{j-1}, T_j)P(t, T_j)$ can be seen as the price of a portfolio of two bonds with maturities T_{j-1} and T_j respectively. Therefore, Theorem 2.1.3 implies that, if we assume the existence of an equivalent measure $\mathbb{P}^{T_j}$ with $P(t, T_j)$ as its associated numeraire

$$\begin{aligned} \frac{F(t, T_{j-1}, T_j)P(t, T_j)}{P(t, T_j)} &= \mathbb{E}^{\mathbb{P}^{T_j}}\left[\frac{F(s, T_{j-1}, T_j)P(s, T_j)}{P(s, T_j)} \,\middle|\, \mathcal{F}_t\right] \\ \Leftrightarrow F(t, T_{j-1}, T_j) &= \mathbb{E}^{\mathbb{P}^{T_j}}\left[F(s, T_{j-1}, T_j) \,\middle|\, \mathcal{F}_t\right] \end{aligned}$$

for $0 \leq t \leq s$ and the second equation implies that the forward rates are martingales under the equivalent $\mathbb{P}^{T_j}$-measure. As a result, our model will have a diffusion process for the forward rates such that it is driftless under that measure.

2.1.2 Backward-looking forward rates

We now establish the mathematical equivalence between forward-looking and backward-looking forward rates which supports the use of a single model to diffuse both types of rate. The results in this section are similar to Lyashenko and Mercurio (2019). First, let us consider FRAs which are also called swaplets since a swap is a portfolio of FRAs.

Consider our discrete standard tenor dates, a FRA involves two counterparties which agree at a particular time t to exchange a floating interest rate, such as the LIBOR, against a fixed rate. The value of the floating rate is determined at its fixing date T_{j-1} (also called the expiry of the FRA), and the payment is done at the maturity date of the FRA T_j. The value of the contract at the payment date is given by

$$N\tau_j \left[F_j(T_{j-1}) - K\right] \tag{2.8}$$

where N is the contract notional amount, K is the fixed rate and L_j is the floating rate

$$F_j(T_{j-1}) = \frac{1}{\tau_j}\left[\frac{1}{P(T_{j-1}, T_j)} - 1\right]. \tag{2.9}$$

We denote by $L_j(t)$ the time t value of the fixed rate K such that this contract has value 0 at time t. By no-arbitrage, since we showed that $F_j(T_j)$ is a martingale under the T_j-measure, we have

$$L_j(t) := K = \mathbb{E}^{\mathbb{P}^{T_j}}\left[F_j(T_{j-1}) \,\middle|\, \mathcal{F}_t\right]$$

and $L_j(t)$ is the quantity modeled in the traditional LMM.

Now consider a backward-looking FRA, i.e. where the floating rate is instead the continuous compounding of the short rates between T_{j-1} and T_j. This is in fact an approximation to the

true payment, which would normally be the discrete compounding of the daily rates and which converges to the continuous case as the time between rate fixing shrinks to 0 (see Lyashenko and Mercurio (2019) for more details). Therefore, the floating rate is given by

$$F_j^B(T_j) = \frac{1}{\tau_j}\left[\frac{P(T_j,T_{j-1})}{P(T_j,T_j)} - 1\right] = \frac{1}{\tau_j}\left[P(T_j,T_{j-1}) - 1\right] = \frac{1}{\tau_j}\left[e^{\int_{T_{j-1}}^{T_j} r(u)du} - 1\right] \tag{2.10}$$

We use a different notation for the backward-looking rate to avoid confusion. Indeed, the theoretical forward-looking rates stops evolving and is constant in the accrual period, while the backward-looking rate continues to evolve.

The value of this backward-looking FRA at maturity is

$$N\tau_j\left[F_j^B(T_j) - K\right] \tag{2.11}$$

We denote by $R_j(t)$ the time t value of the fixed rate K such that this contract has value 0 at time t. By no-arbitrage, since $F_j^B(T_j)$ is also a martingale under the T_j-measure, we have

$$R_j(t) := K = \mathbb{E}^{\mathbb{P}^{T_j}}\left[F_j^B(T_j)\,\middle|\,\mathcal{F}_t\right] \tag{2.12}$$

Since $F_j(t)$ is a martingale, we have the following relationship between the forward and the backward-looking rates and which follows from no-arbitrage

$$F_j(T_{j-1}) = \mathbb{E}^{\mathbb{P}^{T_j}}\left[F_j^B(T_j)\,\middle|\,\mathcal{F}_{T_{j-1}}\right]$$

and therefore, for $t \le T_{j-1}$,

$$\begin{aligned} L_j(t) &= \mathbb{E}^{\mathbb{P}^{T_j}}\left[F_j(T_{j-1})\,\middle|\,\mathcal{F}_t\right] = \mathbb{E}^{\mathbb{P}^{T_j}}\left[\mathbb{E}^{\mathbb{P}^{T_j}}\left[F_j^B(T_j)\,|\,\mathcal{F}_{T_{j-1}}\right]\,\middle|\,\mathcal{F}_t\right] \\ &= \mathbb{E}^{\mathbb{P}^{T_j}}\left[F_j^B(T_j)\,\middle|\,\mathcal{F}_t\right] = R_j(t) \end{aligned} \tag{2.13}$$

where the last equality follows from the law of iterated expectations. Equation (2.13) shows that before the expiry T_{j-1}, the modeled forward rate $R_j(t)$ does not differentiate between the forward and the backward-looking rates, i.e. we can model both rates with a single quantity. Equivalently, $R_j(t)$ corresponds to the forward rates $L_j(t)$ in the LMM.

At $t = T_{j-1}$, the forward-looking rate is fixed at $R_j(T_{j-1}) = F_j(T_{j-1})$ while the backward-looking rate continues to evolve and is modeled by $R_j(t)$. Finally, at and after maturity $t \ge T_j$, the backward-looking rate is constant (but keeps existing).

2.2 Extended HJM framework

The LMM was originally derived by Brace, Gątarek, and Musiela (Brace et al. (1997)) and is based on the Heath-Jarrow-Morton or HJM (Heath et al. (1992)) framework which models the evolution of instantaneous forward rates. The FMM in Lyashenko and Mercurio (2019) is derived using a set of tools including the change-of-numeraire formula developed in Brigo and Mercurio (2006), which they also use to derive the LMM. However, the LMM can be also

constructed from the HJM framework, as is done by Shreve (2008). Lyashenko and Mercurio (2019) do postulate the existence of an underlying HJM with extended bond-prices (extended HJM) for the FMM, but do not explicitly derive it.

Therefore, in this section, we show that the FMM can be obtained from the extended HJM following the steps of Shreve (2008). Furthermore, his results consider an HJM model driven by a single Brownian motion. We instead considers a multi-factor model with nonzero instantaneous correlations, but they remain very similar.

The steps are as follows. The simply compounded forward rate is defined as a ratio of two bond prices. We first obtain the risk-neutral bond price formula under the HJM and apply the Itô-Doeblin formula to this ratio to obtain the forward rate dynamics.

2.2.1 HJM bond price dynamics

The (extended) HJM specifies a stochastic differential equation for the instantaneous forward rates. We then use the relation between instantaneous forward rates and bond prices to derive the dynamics of the bond prices under the HJM. We then change to an equivalent probability measure to exclude arbitrage.

The instantaneous forward rate $f(t,T)$ modeled in the HJM framework is the short interest rate at time T that can be locked in at $t \leq T$ and, by no-arbitrage,

$$f(t,T) = -\lim_{\delta \to 0} \frac{\log P(t,T+\delta) - \log P(t,T)}{\delta} = -\frac{\partial}{\partial T} \log P(t,T) \tag{2.14}$$

This also implies that

$$P(t,T) = \exp\left(-\int_t^T f\left(t,v\right) dv\right), \quad 0 \leq t < T. \tag{2.15}$$

The extended HJM proposes to extend equation (2.15) such that the bond price exists beyond time T. Namely, we assume

$$P(t,T) = \exp\left(\int_T^t r\left(v\right) dv\right), \quad t \geq T \tag{2.16}$$

This equation holds if $f(u,v)$ is constant and equal to $r(u)$ whenever $u \geq v$.

Assumption 2. *Under the probability measure $\mathbb{P}$, $\{\mathbf{W}^{\mathbb{P}}(t) : 0 \leq t < \infty\}$ is a correlated, N-dimensional Brownian motion driving the instantaneous forward rates. Moreover,*

$$d\mathbf{W}^{\mathbb{P}}(t) \left(d\mathbf{W}^{\mathbb{P}}(t)\right)^T = \boldsymbol{\rho}\ dt \tag{2.17}$$

where $\boldsymbol{\rho} = [\rho_{ij}]_{i,j=1}^N$ is the instantaneous correlation matrix of the Brownian motion and is constant over time.

The multifactor extended-HJM assumes the following dynamics

$$df(t,T) = \alpha(t,T)dt + \mathcal{I}_{\{t \leq T\}} \left(\boldsymbol{\Sigma}(t,T)\right)^T d\mathbf{W}^{\mathbb{P}}(t), \quad t \geq 0 \tag{2.18}$$

where $\mathcal{I}_{\{t\leq T\}}$ is 1 if $t \leq T$ otherwise 0, $\alpha(t,T)$ is the drift and $\mathbf{\Sigma}(t,T) = [\sigma_1(t,T), ..., \sigma_M(t,T)]^T$ is the vector of instantaneous forward volatilities. Using Leibniz's rule,

$$d\left(-\int_t^T f\left(t,v\right)dv\right) = f\left(t,t\right)dt - \int_t^T df\left(t,v\right)dv \tag{2.19}$$

For second term in (2.19), we have

$$\begin{aligned}\int_t^T df\left(t,v\right)dv &= \int_t^T \left[\alpha(t,v)dt + \mathcal{I}_{\{t\leq v\}}\left(\mathbf{\Sigma}(t,v)\right)^\top d\mathbf{W}^{\mathbb{P}}(t)\right]dv \\ &= \int_t^T \alpha(t,v)dt\ dv + \int_t^T \mathcal{I}_{\{t\leq v\}}\mathbf{\Sigma}(t,v)^\top d\mathbf{W}^{\mathbb{P}}(t)\ dv\end{aligned}$$

Now, reverse the order of integration

$$\begin{aligned}&\int_t^T \alpha(t,v)dt\ dv = \int_t^T \alpha(t,v)dv\ dt = \alpha^*(t,T)\ dt \\ &\int_t^T \mathcal{I}_{\{t\leq v\}}\left(\mathbf{\Sigma}(t,v)\right)^\top d\mathbf{W}^{\mathbb{P}}(t)\ dv = \int_t^T \mathcal{I}_{\{t\leq v\}}\mathbf{\Sigma}(t,v)^\top dv\ d\mathbf{W}^{\mathbb{P}}(t) = \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t)\end{aligned}$$

where we have defined

$$\begin{aligned}\alpha^*(t,T) &= \int_t^T \alpha(t,v)dv \\ \mathbf{\Sigma}^*(t,T) &= \int_t^T \mathcal{I}_{\{t\leq v\}}\mathbf{\Sigma}(t,v)dv \quad \Rightarrow \quad \mathbf{\Sigma}^*(t,T)^\top = \int_t^T \mathcal{I}_{\{t\leq v\}}\mathbf{\Sigma}(t,v)^\top dv\end{aligned}$$

Given that $f(t,t) = r(t)$, equation (2.19) can be rewritten as

$$d\left(-\int_t^T f\left(t,v\right)dv\right) = \left(r(t) - \alpha^*(t,T)\right)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t) \tag{2.20}$$

Using the Itô-Doeblin formula, the dynamics for the bond price is given by

$$\begin{aligned}dP(t,T) = d\left(e^{-\int_t^T f(t,v)dv}\right) &= \left(e^{-\int_t^T f(t,v)dv}\right)d\left(-\int_t^T f\left(t,v\right)dv\right) \\ &+ \frac{1}{2}\left(e^{-\int_t^T f(t,v)dv}\right)\left[d\left(-\int_t^T f\left(t,v\right)dv\right)\right]^2\end{aligned}$$

Note that $\left[\mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t)\right]$ is a scalar and therefore is symmetric (Tchuindjio (2009)), hence

$$\begin{aligned}\left[d\left(-\int_t^T f\left(t,v\right)dv\right)\right]^2 &= \left[\mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t)\right]\left[\mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t)\right]^T \\ &= \left(\int_t^T \mathcal{I}_{\{t\leq T\}}\mathbf{\Sigma}(t,T)^\top dv\right)\left[d\mathbf{W}^{\mathbb{P}}(t)\left(d\mathbf{W}^{\mathbb{P}}(t)\right)^\top\right]\left(\int_t^T \mathcal{I}_{\{t\leq T\}}\mathbf{\Sigma}(t,T)dv\right) \\ &= \left[\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\ \mathbf{\Sigma}^*(t,T)\right]dt\end{aligned}$$

which is a scalar. Plugging this is the previous expression for $dP(t,T)$ yields

$$\begin{aligned} dP(t,T) &= P(t,T)\left[(r(t)-\alpha(t,T)^*)\,dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t)\right] \\ &\quad + \frac{1}{2}P(t,T)\left[\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)\right]dt \\ \Leftrightarrow \frac{dP(t,T)}{P(t,T)} &= \left[r(t)-\alpha^*(t,T)+\frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)\right]dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t) \end{aligned} \tag{2.21}$$

In particular, when $t > T$, the equation (2.21) reduces to

$$\frac{dP(t,T)}{P(t,T)} = [r(t)-\alpha^*(t,T)]\,dt \tag{2.22}$$

To continue, recall from Shreve (2008)

Theorem 2.2.1 (First fundamental theorem of asset pricing). *If a market model has a risk-neutral probability measure, then it does not admit arbitrage.*

Furthermore, a risk-neutral probability measure is an equivalent probability measure such that the discounted asset prices under this measure are martingales. In particular, the discounted bond price

$$\frac{P(t,T)}{B(t)} = \exp\left(-\int_0^t r(u)du\right)P(t,T) \tag{2.23}$$

should be a martingale under the equivalent risk-neutral measure. Because

$$d\left(\frac{1}{B(t)}\right) = -\frac{1}{B(t)}r(t)dt$$

it follows that, using the Itô-Doeblin formula,

$$\begin{aligned} d\left(\frac{P(t,T)}{B(t)}\right) &= \frac{1}{B(t)}dP(t,T) + d\left(\frac{1}{B(t)}\right)P(t,T) \\ &= \frac{1}{B(t)}dP(t,T) - \frac{P(t,T)}{B(t)}r(t)dt \\ &= \frac{P(t,T)}{B(t)}\left[\left(r(t)-\alpha(t,T)^*+\frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)\right)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t) - r(t)dt\right] \\ &= \frac{P(t,T)}{B(t)}\left[\left(-\alpha^*(t,T)+\frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)\right)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t)\right] \end{aligned} \tag{2.24}$$

For equation (2.24) to be a martingale, we must get rid of the drift term by changing to an equivalent probability measure. Specifically, we seek an M-dimensional process $\mathbf{\Theta}(t)$ such that

$$\left[\left(-\alpha^*(t,T)+\frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)\right)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t)\right] = -\mathbf{\Sigma}^*(t,T)^\top\left[\mathbf{\Theta}(t)dt + d\mathbf{W}^{\mathbb{P}}(t)\right] \tag{2.25}$$

which is equivalent to

$$-\alpha^*(t,T)+\frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T) = -\mathbf{\Sigma}^*(t,T)^\top \mathbf{\Theta}(t) \tag{2.26}$$

We differentiate both sides of (2.26) with respect to the scalar T. Differentiating the right-hand

side and the first term on the left-hand side of (2.26) is straightforward when using Leibniz's rule of differentiation

$$\frac{\partial}{\partial T}\alpha^*(t,T) = \alpha(t,T) \tag{2.27}$$

$$\frac{\partial}{\partial T}\mathbf{\Sigma}^*(t,T)^\top \mathbf{\Theta}(t) = \mathbf{\Sigma}(t,T)^\top \mathcal{I}_{(t\leq T)}\mathbf{\Theta}(t) \tag{2.28}$$

To differentiate the remaining term, we normally would have to rely on heavily advanced algebra and define, in a Hilbert space H_n, the real inner product $\langle\cdot,\cdot\rangle : H_n \times H_n \to \mathbb{R}$ such that for two vectors $V_1, V_2 \in \mathbb{R}^2$, we have $\langle V_1, V_2\rangle = V_1^\top V_2$. Furthermore, in our case, we could use

$$\Big\langle \int_0^T V_1 ds, V_2 \Big\rangle = \int_0^T \langle V_1, V_2\rangle ds \quad ,$$

(Skorohod (1974)). However, one should check that these vectors respect some regularity and measurability conditions to go further and these technicalities are beyond the scope of this thesis. Fortunately, Tchuindjio (2009) implicitly shows that

$$\begin{aligned}\frac{\partial}{\partial T}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T) &= \frac{\partial}{\partial T}\int_t^T \Big\langle \mathbf{\Sigma}(t,v)\mathcal{I}_{(t\leq v)}, \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)\Big\rangle dv = 2\Big\langle \mathbf{\Sigma}(t,T)\mathcal{I}_{(t\leq T)}, \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)\Big\rangle \\ &= 2\, \mathbf{\Sigma}(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)\mathcal{I}_{(t\leq T)}\end{aligned}$$

This allows us to re-express equation (2.26) and the **HJM no-arbitrage condition** becomes

$$\alpha(t,T) = \mathbf{\Sigma}(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)\mathcal{I}_{(t\leq T)} - \mathbf{\Sigma}(t,T)^\top \mathbf{\Theta}(t)\mathcal{I}_{(t\leq T)} \tag{2.29}$$

Note that if $t > T, \alpha(t,T) = 0$ and $\mathbf{\Sigma}^*(t,T) = 0$ and from equation (2.21),

$$\frac{dP(t,T)}{P(t,T)} = r(t)dt \tag{2.30}$$

which was claimed in Lyashenko and Mercurio (2019) although they did not derive this result. Otherwise, provided that $\mathbf{\Sigma}(t,T)$ is not null, there is a unique $\mathbf{\Theta}(t)$ such that equation (2.29) holds. Again, one can use advanced algebra to prove this, as Tchuindjio (2009) has done by invoking the Hilbert space representation theorem. We omit these technicalities .

We can now use Girsanov's theorem to change to a probability measure $\widetilde{\mathbb{P}}$ under which

$$\mathbf{W}^{\widetilde{\mathbb{P}}}(t) = \int_0^t \mathbf{\Theta}(u)du + \mathbf{W}^{\mathbb{P}}(t) \tag{2.31}$$

$$\text{or} \quad d\mathbf{W}^{\widetilde{\mathbb{P}}}(t) = \mathbf{\Theta}(t)dt + d\mathbf{W}^{\mathbb{P}}(t) \tag{2.32}$$

is a Brownian motion. Ding (2009) shows that Girsanov's Theorem extends to correlated Brownian motion. Finally, using equation (2.26) and (2.32), we can rewrite equation (2.24) as

$$\begin{aligned} d\left(\frac{P(t,T)}{B(t)}\right) &= \frac{P(t,T)}{B(t)}\left[\left(-\alpha^*(t,T) + \frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)\right)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t)\right] \\ &= \frac{P(t,T)}{B(t)}\left[-\mathbf{\Sigma}^*(t,T)^\top \mathbf{\Theta}(t)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\mathbb{P}}(t)\right]\end{aligned}$$

$$= \frac{P(t,T)}{B(t)}(-\mathbf{\Sigma}^*(t,T))^\top \left[\mathbf{\Theta}(t)dt + d\mathbf{W}^{\mathbb{P}}(t)\right]$$
$$= \frac{P(t,T)}{B(t)}(-\mathbf{\Sigma}^*(t,T))^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t) \tag{2.33}$$

proving that the discounted bond price is a martingale under $\widetilde{\mathbb{P}}$. This also shows that the probability measure $\widetilde{\mathbb{P}}$ is associated with the numeraire $B(t)$, which we previously assumed. To derive the dynamics of the forward rates in the FMM, the dynamics of the bond price are needed. Using that $P(t,T) = B(t)\frac{1}{B(t)}P(t,T)$, the dynamics are given by

$$\begin{aligned} dP(t,T) &= r(t)B(t)\frac{1}{B(t)}P(t,T)dt + B(t)\frac{P(t,T)}{B(t)}(-\mathbf{\Sigma}^*(t,T)^\top)d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)) \\ &= P(t,T)\Big(r(t)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\Big) \end{aligned}$$

We summarize this result in the following Theorem which fully describes the extended HJM

Theorem 2.2.2. *Under the $\widetilde{\mathbb{P}}$-measure, the bond price dynamics are given by*

$$dP(t,T) = P(t,T)\Big(r(t)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\Big) \tag{2.34}$$

The solution of this stochastic differential equation is

$$\begin{aligned} P(t,T) &= P(0,T)\exp\left(\int_0^t r(u)du - \int_0^t \frac{1}{2}\mathbf{\Sigma}^*(u,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(u,T)du - \int_0^t \mathbf{\Sigma}^*(u,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(u)\right) \\ &= P(0,T)B(t)\exp\left(-\int_0^t \frac{1}{2}\mathbf{\Sigma}^*(u,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(u,T)du - \int_0^t \mathbf{\Sigma}^*(u,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(u)\right) \end{aligned} \tag{2.35}$$

Proof. The proof follows from applying the Itô-Doeblin formula and is left in Appendix. A □

2.2.2 Lognormal Forward Market Model from HJM bond prices

The extended HJM framework and Theorem 2.2.2 provides an expression for the stochastic bond prices which we can use to derive for the simply-compounded forward rates. We take

$$P(t,T_{j-1}) = P(0,T_{j-1})B(t)\exp\left(-\int_0^t \frac{1}{2}\mathbf{\Sigma}^*(u,T_{j-1})^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(u,T_{j-1})du - \int_0^t \mathbf{\Sigma}^*(u,T_{j-1})^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(u)\right)$$
$$P(t,T_j) = P(0,T_j)B(t)\exp\left(-\int_0^t \frac{1}{2}\mathbf{\Sigma}^*(u,T_j)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(u,T_j)du - \int_0^t \mathbf{\Sigma}^*(u,T_j)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(u)\right)$$

and use equations (2.5) and (2.6) to obtain

$$\begin{aligned} \widetilde{R}_j(t) + \frac{1}{\tau_j} &= \frac{1}{\tau_j}\frac{P(t,T_{j-1})}{P(t,T_j)} + \delta \\ &= \frac{P(0,T_{j-1})}{\tau_j P(0,T_j)}\exp\Bigg\{\int_0^t \frac{1}{2}\left[\mathbf{\Sigma}^*(u,T_j)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(u,T_j) - \mathbf{\Sigma}^*(u,T_{j-1})^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(u,T_{j-1})\right]du \\ &\quad + \int_0^t \left[\mathbf{\Sigma}^*(u,T_j)^\top - \mathbf{\Sigma}^*(u,T_{j-1})^\top\right]d\mathbf{W}^{\widetilde{\mathbb{P}}}(u)\Bigg\} \end{aligned}$$

We present the dynamics of $\widetilde{R}_j(t)$ in the next proposition and we leave the derivation, which is simply an application of Itô-Doeblin, in Appendix A.

Proposition 2.2.3 (Forward Rate Dynamics). *The dynamics of $\widetilde{R}_j(t)$ are given by*

$$d\widetilde{R}_j(t) = \left(R_j(t) + \frac{1}{\tau_j}\right)\left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right]^\top \left\{\boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j)dt + d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\right\}$$

Proof. See Appendix A □

Next we construct the well-known T_j-forward measure. Note that $B(t)P(t,T_{j-1})$ and $B(t)P(t,T_j)$ are both martingales under the $\widetilde{\mathbb{P}}$ measure. Then, Theorem 9.2.2 from Shreve (2008) says that $\dfrac{P(t,T_{j-1})}{P(t,T_j)}$ is a martingale under the equivalent probability measure $\mathbb{P}^{T_j}$ associated with the numeraire $P(t,T_j)$. Furthermore, this measure is defined by

$$\mathbb{P}^{T_j}(A) = \int_A \frac{B\left(T_j\right)}{P(0,T_j)} d\widetilde{\mathbb{P}} \ \forall \in \mathcal{F} \tag{2.36}$$

and, as a consequence, Girsanov's Theorem for correlated Brownian motions (Ding (2009)) implies that

$$d\mathbf{W}^{T_j}(t) = \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j)dt + d\mathbf{W}^{\widetilde{\mathbb{P}}}(t) \tag{2.37}$$

is a Brownian motion under $\mathbb{P}^{T_j}$. This allows us to express $\widetilde{R}_j(t)$ as

$$d\widetilde{R}_j(t) = \left(R_j(t) + \frac{1}{\tau_j}\right)\left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right]^\top d\mathbf{W}^{T_j}(t) \tag{2.38}$$

Since $R_j(t)$ is a martingale under the $\mathbb{P}^{T_j}$-measure, the multi-dimensional Martingale Representation Theorem (Theorem 5.4.2 in Shreve (2008)) says that there must exist an adapted, M-dimensional process $\left\{\boldsymbol{\Gamma}(t,T_j) : t \in [0,\infty)\right\}$ such that

$$d\widetilde{R}_j(t) = \widetilde{R}_j(t)\ \boldsymbol{\Gamma}(t,T_j)^\top d\mathbf{W}^{T_j}(t). \tag{2.39}$$

With this equation we can relate the discrete forward rate volatilities to the bond price volatilities

$$\begin{aligned}\boldsymbol{\Gamma}(t,T_j)\widetilde{R}_j(t) &= \frac{1}{\tau_j}\left(\tau_j R_j(t) + 1\right)\left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right] \\ \Leftrightarrow \boldsymbol{\Gamma}(t,T_j) &= \frac{\tau_j R_j(t) + 1}{\tau_j \widetilde{R}_j(t)}\left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right]\end{aligned} \tag{2.40}$$

2.2.3 Risk-neutral drift of other forwards under the $\mathbb{P}^{T_j}$-measure

Each forward rate $R_j(t)$ is a martingale under its own $\mathbb{P}^{T_j}$-measure. We now derive the risk-neutral drift of all forwards under that measure. Consider these two equivalent equations

$$\begin{aligned}d\mathbf{W}^{T_j}(t) &= \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j)dt + d\mathbf{W}^{\widetilde{\mathbb{P}}}(t) \\ d\mathbf{W}^{T_{j-1}}(t) &= \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_{j-1})dt + d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\end{aligned}$$

and subtract them to obtain

$$\begin{aligned} d\mathbf{W}^{T_j}(t) &= \Big[\boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_{j-1})\Big]dt + d\mathbf{W}^{T_{j-1}}(t) \\ &= \boldsymbol{\rho}\ \Big[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\Big]dt + d\mathbf{W}^{T_{j-1}}(t) \\ &= \ \boldsymbol{\rho}\ \boldsymbol{\Gamma}(t,T_j)\frac{\tau_j\widetilde{R}_j(t)}{\tau_j R_j(t)+1}dt + d\mathbf{W}^{T_{j-1}}(t) \end{aligned} \tag{2.41}$$

First, we derive the dynamics of $R_i(t)$ under the $\mathbb{P}^{T_j}$-measure for $i = 0, 1, ..., j-1$. Equation (2.41) provides a recursive formula for Brownian motions. Rewrite it as

$$d\mathbf{W}^{T_{j-1}}(t) = -\ \boldsymbol{\rho}\ \boldsymbol{\Gamma}(t,T_j)\frac{\tau_j R_j(t)}{\tau_j R_j(t)+1}dt + d\mathbf{W}^{T_j}(t)$$

Similarly,

$$\begin{aligned} d\mathbf{W}^{T_{j-2}}(t) &= -\ \boldsymbol{\rho}\ \boldsymbol{\Gamma}(t,T_{j-1})\frac{\tau_{j-1}\widetilde{R}_{j-1}(t)}{\tau_{j-1}R_{j-1}(t)+1}dt + d\mathbf{W}^{T_{j-1}}(t) \\ &= -\ \boldsymbol{\rho}\ \boldsymbol{\Gamma}(t,T_{j-1})\frac{\tau_{j-1}\widetilde{R}_{j-1}(t)}{\tau_{j-1}R_{j-1}(t)+1}dt - \ \boldsymbol{\rho}\ \boldsymbol{\Gamma}(t,T_j)\frac{\tau_j\widetilde{R}_j(t)}{\tau_j R_j(t)+1}dt + d\mathbf{W}^{T_j}(t) \end{aligned}$$

We can proceed till we have

$$d\mathbf{W}^{T_i}(t) = -\sum_{k=i+1}^{j}\frac{\tau_k\widetilde{R}_k(t)}{1+\tau_k R_k(t)}\boldsymbol{\rho}\ \boldsymbol{\Gamma}(t,T_k)dt + d\mathbf{W}^{T_j}(t) \tag{2.42}$$

We insert (2.42) in equation (2.39) to obtain the $\mathbb{P}^{T_j}$-dynamics for $R_i(t),\ i = 0, ..., j-1$

$$\begin{aligned} d\widetilde{R}_i(t) &= \widetilde{R}_i(t)\ \boldsymbol{\Gamma}(t,T_i)^\top d\mathbf{W}^{T_i}(t) = \widetilde{R}_i(t)\ \boldsymbol{\Gamma}(t,T_i)^\top\left[-\sum_{k=i+1}^{j}\frac{\tau_k\widetilde{R}_k(t)}{1+\tau_k R_k(t)}\boldsymbol{\rho}\ \boldsymbol{\Gamma}(t,T_k)dt + d\mathbf{W}^{T_j}(t)\right] \\ &= -\widetilde{R}_i(t)\sum_{k=i+1}^{j}\frac{\tau_i\widetilde{R}_k(t)}{1+\tau_k R_k(t)}\ \boldsymbol{\Gamma}(t,T_i)^\top\boldsymbol{\rho}\ \boldsymbol{\Gamma}(t,T_k)dt + \widetilde{R}_i(t)\ \boldsymbol{\Gamma}(t,T_i)^\top d\mathbf{W}^{T_j}(t) \end{aligned} \tag{2.43}$$

Now, we derive the same dynamics $d\widetilde{R}_i(t)$ for $i = j+1, ..., M-1$. In the same way we derived (2.41), we can obtain

$$d\mathbf{W}^{T_{j+1}}(t) = \ \boldsymbol{\rho}\ \boldsymbol{\Gamma}(t,T_{j+1})\frac{\tau_j\widetilde{R}_{j+1}(t)}{\tau_j R_{j+1}(t)+1}dt + d\mathbf{W}^{T_j}(t) \tag{2.44}$$

And again proceeding recursively,

$$d\mathbf{W}^{T_i}(t) = \sum_{k=j+1}^{i}\frac{\tau_k\widetilde{R}_k(t)}{1+\tau_k R_k(t)}\boldsymbol{\rho}\ \boldsymbol{\Gamma}(t,T_k)dt + d\mathbf{W}^{T_j}(t) \tag{2.45}$$

Inserting (2.45) into equation (2.39) gives the $\mathbb{P}^{T_j}$-dynamics for $R_i(t),\ i = j+1, ..., M-1$

$$\begin{aligned} d\widetilde{R}_i(t) &= \widetilde{R}_i(t)\ \mathbf{\Gamma}(t,T_i)^\top \left[\sum_{k=j+1}^{i} \frac{\tau_k \widetilde{R}_k(t)}{1+\tau_k R_k(t)} \boldsymbol{\rho}\ \mathbf{\Gamma}(t,T_k)dt + d\mathbf{W}^{T_j}(t) \right] \\ &= \widetilde{R}_i(t) \sum_{k=j+1}^{i} \frac{\tau_k \widetilde{R}_k(t)}{1+\tau_k R_k(t)} \mathbf{\Gamma}(t,T_i)^\top \boldsymbol{\rho}\ \mathbf{\Gamma}(t,T_k)dt + \widetilde{R}_i(t)\mathbf{\Gamma}(t,T_i)^\top d\mathbf{W}^{T_j}(t) \end{aligned} \tag{2.46}$$

This concludes the derivation of the FMM dynamics under the $\mathbb{P}^{T_j}$ forward measure. In the next section, we will also consider the dynamics under a different measure. Before that, we summarize the results in the following theorem.

Theorem 2.2.4 (Extended forward measure dynamics in the FMM). *Under the extended forward measure defined by (2.36) and the extended-HJM framework from section 2.2 driving instantaneous forward rates, the forward rate dynamics are given by*

$$\begin{aligned} i<j, \quad t \geq 0, \qquad d\widetilde{R}_i(t) &= -\widetilde{R}_i(t)\ \mathbf{\Gamma}(t,T_i)^\top \sum_{k=i+1}^{j} \frac{\tau_i \widetilde{R}_k(t)}{1+\tau_k R_k(t)}\ \boldsymbol{\rho}\ \mathbf{\Gamma}(t,T_k)dt \\ &\quad + \widetilde{R}_i(t)\ \mathbf{\Gamma}(t,T_i)^\top d\mathbf{W}^{T_j}(t) \end{aligned} \tag{2.39}$$

$$i=j, \quad t \geq 0, \qquad d\widetilde{R}_j(t) = \widetilde{R}_j(t)\ \mathbf{\Gamma}(t,T_j)^\top d\mathbf{W}^{T_j}(t) \tag{2.38}$$

$$\begin{aligned} i>j, \quad t \leq T_j, \qquad d\widetilde{R}_i(t) &= \widetilde{R}_i(t)\ \mathbf{\Gamma}(t,T_i)^\top \sum_{k=j+1}^{i} \frac{\tau_k \widetilde{R}_k(t)}{1+\tau_k R_k(t)}\ \boldsymbol{\rho}\ \mathbf{\Gamma}(t,T_k)dt \\ &\quad + \widetilde{R}_i(t)\mathbf{\Gamma}(t,T_i)^\top d\mathbf{W}^{T_j}(t) \end{aligned} \tag{2.43}$$

Equation (2.38) has a solution given by

$$\begin{aligned} \widetilde{R}_j(t) &= \widetilde{R}_j(0) \exp\left\{ -\frac{1}{2}\int_0^t \mathbf{\Gamma}(t,T_j)^\top \mathbf{\Gamma}(t,T_j)du + \int_0^t \mathbf{\Gamma}(t,T_j)^\top d\mathbf{W}^{T_j}(u) \right\} \\ &= \widetilde{R}_j(0) \exp\left\{ -\frac{1}{2}\int_0^t \|\mathbf{\Gamma}(t,T_j)\|^2 du + \int_0^t \mathbf{\Gamma}(t,T_j)^\top d\mathbf{W}^{T_j}(u) \right\} \end{aligned} \tag{2.47}$$

where $\|V\| = \sqrt{V_1^2 + ... + V_N^2}$ *for any vector* $V = [V_1, ..., V_N]^\top \in \mathbb{R}^N$.

Proof. We only prove that (2.47) is a solution to (2.38) since the rest was established before introducing Theorem 2.2.4. Applying Itô-Doeblin to (2.47), we have

$$\begin{aligned} d\widetilde{R}_j(t) &= \widetilde{R}_j(0) \exp\left\{ -\frac{1}{2}\int_0^t \mathbf{\Gamma}(t,T_j)^\top \mathbf{\Gamma}(t,T_j)du + \int_0^t \mathbf{\Gamma}(t,T_j)^\top d\mathbf{W}^{T_j}(u) \right\} \\ &\quad \times \left(-\frac{1}{2}\mathbf{\Gamma}(t,T_j)^\top \mathbf{\Gamma}(t,T_j)dt + \mathbf{\Gamma}(t,T_j)^\top d\mathbf{W}^{T_j}(t) + \frac{1}{2}\mathbf{\Gamma}(t,T_j)^\top \mathbf{\Gamma}(t,T_j)dt \right) \\ &= \widetilde{R}_j(t)\ \mathbf{\Gamma}(t,T_j)^\top d\mathbf{W}^{T_j}(t) \end{aligned}$$

□

An assumption on the structure of $\mathbf{\Gamma}(t, T_j)$ is usually adopted in the literature and in practice, and greatly simplifies the FMM derived so far. It is implicitly made in Lyashenko and Mercurio (2019) for the FMM derivation and explicitly in Brigo and Mercurio (2006) for the LMM derivation. For the rest of this thesis, we will also make this assumption and derive further results accordingly.

Assumption 3. *The forward rate volatility vector* $\mathbf{\Gamma}(t, T_j)$, $j = 1, \ldots, M$ *only has non-zero entries except for the* j^{th} *element. That is,*

$$\mathbf{\Gamma}(t, T_j) = \begin{bmatrix} 0 & 0 & \ldots & \sigma_j^*(t) & \ldots & 0 \end{bmatrix}^\top \tag{2.48}$$

where

$$\sigma_j^*(t) = \sigma_j(t) g_j(t) \tag{2.49}$$

and $\sigma_j(t)$ *is an adapted process.*

We will in fact consider two forms for $\sigma_j(t)$ for our numerical examples. The simplest form is a constant instantaneous volatility over the entire life of the forward rate (until it becomes fixed at maturity), or equivalently $\sigma_j(t) = \sigma_j$, for all $j = 1, \ldots, M$. The alternative specification of the instantaneous forward rate volatilities is a parametric form suggested by Brigo and Mercurio (2006) and consists of an exponentially decreasing volatility, namely,

$$\sigma_j(t) = \mathbf{\Phi}_j \left([a(T^* - t) + c]\, e^{-b(T^* - t)} + d \right) \tag{2.50}$$

and a, b, c and d are typically determined in a calibration routine. Furthermore, T^* is set to T_{i-1}, the expiry, in the LMM so we extend this parametric form in the FMM by setting $T^* = T_i$, regardless of whether the instrument is forward-looking or backward-looking.

The function $g_j(t)$ in equation (2.49) is introduced to allow for different behaviors of the volatility, especially in the accrual period. Lyashenko and Mercurio (2019) observe an empirical market instance of a decaying volatility in the accrual period of backward-looking rates and suggests to use

$$g_j(t) = \min \left\{ \frac{(T_j - at)^+}{T_j - bT_{j-1}}, 1 \right\} = \begin{cases} 1 & t \le T_{j-1} \\ \dfrac{T_j - at}{T_j - bT_{j-1}} & T_{j-1} < t \le T_j \\ 0 & t > T_j \end{cases} \tag{2.51}$$

with $a = b = 1$, which creates a linearly decaying volatility in the accrual period. After maturity, that is when $t > T_{j-1}$, we have $g_j(t) = 0$ which implies that the forward rate stops evolving. Unless specified otherwise, we will adopt this scaling function. (The coefficients a and b are introduced in this thesis to experiment with different functions $g_j(t)$ and will become clearer in later sections.)

2.2.4 Risk-neutral drift under the $\mathbb{Q}$-measure

Changing the probability measure only changes the drift while the diffusion coefficient stays the same (Baxter and Rennie (1996)). Denote the drift of the process $R_j(t)$ under the numeraire $P(t,T_j)$ by $\mu^{P_j}(R_j,t)$. As is clear from equation (2.38) in Theorem 2.2.4, $\mu^{P_j}(R_j,t)=0$ since $R_j(t)$ is a martingale under the $\mathbb{P}^{T_j}$-measure (it is driftless). Similarly denote the drift of $R_j(t)$ under the numeraire corresponding to the risk-neutral measure, $B(t)$, by $\mu^B(R_j,t)$. Brigo and Mercurio (2006) present a set of tools to conveniently derive the drift of a stochastic process when applying a change of numeraire (or equivalently, a change of probability measure). Specifically, they show that the two drifts can be related by the following formula

$$\begin{aligned}\mu^B(R_j,t)dt &= \mu^{P_j}(R_j,t)dt - d\ln R_j d\ln(P_j/B)\\ \Leftrightarrow \mu^B(R_j,t) &= \mu^{P_j}(R_j,t) - \frac{d\ln R_j d\ln(P_j/B)}{dt}\end{aligned} \tag{2.52}$$

The derivation in the FMM is presented in Lyashenko and Mercurio (2019) but since it is brief, we repeat it here somewhat more detailed. It will be useful to know $d\ln F_j(t)$. Using the Itô-Doeblin formula, we have

$$d\ln R_j(t) = \frac{dR_j(t)}{R_j(t)} - \frac{1}{2}\frac{1}{R_j^2(t)}(dR_j(t))^2 = -\frac{1}{2}g_j^2(t)\sigma_j^2(t)dt + g_j(t)\sigma_j(t)dW_j^{T_j}(t) \tag{2.53}$$

First note that

$$\begin{aligned}\ln\left[\frac{P(t,T_j)}{B(t)}\right] &= -\ln\left[\frac{B(t)}{P(t,T_j)}\right] = -\ln\left[\frac{P(t,0)}{P(t,T_j)}\right]\\ &= -\ln\left[\prod_{i=1}^{j}\frac{P(t,T_{i-1})}{P(t,T_i)}\right] = -\ln\left[\prod_{i=1}^{j}1+\tau_i R_i(t)\right]\\ &= -\sum_{i=1}^{j}\ln\left[1+\tau_i R_i(t)\right]\end{aligned} \tag{2.54}$$

Inserting this equation into (2.52) yields

$$\begin{aligned}\mu^B(R_j,t) &= \frac{d\ln R_j(t) d\left(\sum_{i=1}^{j}\ln\left[1+\tau_i R_i(t)\right]\right)}{dt}\\ &= \sum_{i=1}^{j}\frac{d\ln R_j(t)\ d\ln\left[1+\tau_i R_i(t)\right]}{dt}\\ &= \sum_{i=1}^{j}\frac{\tau_i}{1+\tau_i R_i(t)}\frac{d\langle\ln R_j(t), R_i(t)\rangle}{dt}\end{aligned} \tag{2.55}$$

Using the earlier expression for $d\ln R_j(t)$ and the fact that $dtd\widetilde{W}=0$,

$$\frac{d\langle\ln R_j(t), R_i(t)\rangle}{dt} = \frac{\left(-\frac{1}{2}g_j^2(t)\sigma_j^2(t)dt + g_j(t)\sigma_j(t)dW_j^{T_j}(t)\right)\left(g_i(t)\sigma_j(t)R_i(t)dW_i^{T_i}(t)\right)}{dt}$$

$$= \frac{g_j(t)g_i(t)\rho_{ij}\sigma_j(t)\sigma_i(t)R_i(t)dt}{dt} = g_j(t)g_i(t)\rho_{ij}\sigma_j(t)\sigma_i(t)R_i(t) \tag{2.56}$$

Plug (2.56) into (2.55) results in the drift of the forward rate under the risk-neutral measure

$$\begin{aligned}\mu^B(R_j,t) &= \sum_{i=1}^{j} \frac{\tau_i}{1+\tau_i R_i(t)} g_j(t)g_i(t)\rho_{ij}\sigma_j(t)\sigma_i(t)R_i(t) \\ &= g_j(t)\sigma_j(t)\sum_{i=1}^{j}\rho_{ij}\frac{g_i(t)\sigma_i(t)R_i(t)}{1+\tau_i R_i(t)}\end{aligned} \tag{2.57}$$

We summarize this result in the following theorem

Theorem 2.2.5. *Under the risk-neutral measure $\mathbb{Q}$ with associated numeraire $B(t)$ and Brownian motion $\widetilde{W}^{\mathbb{Q}}(t)$, the dynamics of the forward rates $\widetilde{R}_j(t)$ are given by*

$$d\widetilde{R}_j(t) = g_j(t)\sigma_j(t)\widetilde{R}_j(t)\sum_{i=1}^{j}\rho_{ij}\frac{g_i(t)\sigma_i(t)\widetilde{R}_i(t)}{1+\tau_i R_i(t)}dt + g_j(t)\sigma_j(t)\widetilde{R}_j(t)d\widetilde{W}^{\mathbb{Q}}(t) \tag{2.58}$$

Remark 2.2.6. *We introduce the index function $\eta(t) = \{j : T_{j-1} < t \leq T_j\}$. It corresponds to the closest next maturity after t in the tenor structure. When $t > T_j$, $g_j(t) = 0$ and therefore every forward rate whose maturity is before T_j have a zero-drift and (volatility) at time t. Therefore, equation (2.58) can be rewritten as*

$$d\widetilde{R}_j(t) = g_j(t)\sigma_j(t)\widetilde{R}_j(t)\sum_{i=\eta(t)}^{j}\rho_{ij}\frac{g_i(t)\sigma_i(t)\widetilde{R}_i(t)}{1+\tau_i R_i(t)}dt + g_j(t)\sigma_j(t)\widetilde{R}_j(t)d\widetilde{W}^{\mathbb{Q}}(t)$$

Numeraire in the FMM

The numeraire typically used in the LMM is the one associated with the spot measure, and it is the discrete equivalent of $B(t)$. It consists of a bank account that is rebalanced only at the tenor structure dates. It is given by

$$B_d(t) = \frac{P(t,T_{\eta(t)-1})}{\prod_{j=0}^{\eta(t)-1}P(T_{j-1},T_j)} \tag{2.59}$$

(see Brigo and Mercurio (2006)). It was first introduce by Jamishidian (1997) because the dynamics of the LMM under the risk-neutral dynamics presented an awkward form that made the computation of the drift a difficult task. This spot-measure is no longer needed in the FMM as Lyashenko and Mercurio (2019) show and the risk-neutral dynamics can be used as we just showed. Although, $B(t)$ is not directly accessible in the LMM, we can now retrieve it in the FMM. First note that from (2.12), we have

$$\begin{aligned} R_j(T_j) = \mathbb{E}^{\mathbb{P}^{T_j}}\left[F^B(T_j)\,\middle|\,\mathcal{F}_{T_j}\right] &= \mathbb{E}^{\mathbb{P}^{T_j}}\left[\frac{1}{\tau_j}\left[e^{\int_{T_{j-1}}^{T_j} r(u)du} - 1\right]\,\middle|\,\mathcal{F}_{T_j}\right] \\ \Leftrightarrow 1+\tau_j R_j(T_j) &= e^{\int_{T_{j-1}}^{T_j} r(u)du}\end{aligned} \tag{2.60}$$

Therefore, we have

$$\begin{aligned} B(t) &= e^{\int_0^t r(u)du} = e^{\int_0^{T_1} r(u)du} e^{\int_{T_1}^{T_2} r(u)du} \cdots e^{\int_{T_{\eta(t)-2}}^{T_{\eta(t)-1}} r(u)du} e^{\int_{T_{\eta(t)-1}}^{t} r(u)du} \\ &= e^{\int_{T_{\eta(t)-1}}^{t} r(u)du} \prod_{i=1}^{\eta(t)-1} e^{\int_{T_{i-1}}^{T_i} r(u)du} \\ &= e^{\int_{T_{\eta(t)-1}}^{t} r(u)du} \left[\prod_{i=1}^{\eta(t)-1} 1 + R_j(T_j) \right] = B(T_{\eta(t)-1})\ e^{\int_{T_{\eta(t)-1}}^{t} r(u)du} \end{aligned} \tag{2.61}$$

Thus, $B(t)$ is known at the standard tenor dates T_j, $j = 1, \ldots, M$ and is given by

$$B(T_j) = \prod_{i=1}^{j} 1 + R_j(T_j) \tag{2.62}$$

2.3 The Black formula for caplet pricing

Valuation of caps and floors in the FMM can be done by using the standard Black formula and is similar to the situation in the LMM. As an example, we briefly dis Equation (2.38) gives rise to a lognormal distribution of the forward rate $R_j(t)$ under its respective $\mathbb{P}^{T_j}$ forward measure. Under this setting, the value of caplet has an analytical expression given in the following Theorem.

Theorem 2.3.1 (Caplet price). *A caplet (floorlet) with expiry T_{j-1} and maturity T_j is equivalent to a swaplet where the buyer of this financial product can choose to exercise the swaplet only if it has a positive value. The payoff at maturity is given by*

$$\tau_j \Big(R(T^*, T_{j-1}, T_j) - K \Big)^+ = \tau_j \Big(\widetilde{R}(T^*, T_{j-1}, T_j) - (K + \delta) \Big)^+ \tag{2.63}$$

where K is the strike price of the caplet and $T^* = T_{j-1}$ if the caplet is forward looking and $T^* = T_j$ if it is backward-looking. The price of this contract at time t can be shown to be

$$\boldsymbol{Bl}^{\,caplet}(t) = \gamma P(t, T_j) \Big(\widetilde{R}(t, T_{j-1}, T_j) \Phi(d_1) - (K + \delta) \Phi(d_2) \Big) \tag{2.64}$$

where δ is the displacement parameter and

$$d_1 = \frac{\log\left(\frac{\widetilde{R}_j(t)}{K+\delta}\right) + \frac{1}{2}\int_t^{T^*} \|\boldsymbol{\Gamma}(s, T_j)\|^2 \, ds}{\sqrt{\int_t^{T^*} \|\boldsymbol{\Gamma}(s, T_j)\|^2 \, ds}}, \qquad d_2 = d_1 - \sqrt{\int_t^{T^*} \|\boldsymbol{\Gamma}(s, T_j)\|^2 \, ds} \tag{2.65}$$

Proof. The proof is a standard result in derivative pricing and is left in Appendix A. □

We are often interested in the average volatility of the forward rate $R(t, T_{j-1}, T_j)$ up to a certain time t, which we denote $\lambda_{T_{j-1}, T_j}(t)$, as it provides an insight on the magnitude of its instantaneous volatility. Combined with Assumption (3) and the specified linearly decaying

function $g_j(t)$ of equation (2.51), in the backward looking case ($T^* = T_j$), we have

$$\lambda_{T_{j-1},T_j}(T^*) = \sqrt{\frac{1}{T^*}\int_0^{T^*} \sigma_j^2(t) g_j^2(t) ds} \tag{2.66}$$

If $\sigma_j(t)$ is constant and equal to σ_j, we have

$$\lambda_{T_{j-1},T_j}(T^*) = \sigma_i \sqrt{\frac{1}{T_j}\left(\frac{(T_j - aT_{j-1})^3 - (T_j - aT_j)^3}{3a(T_j - bT_{j-1})^2} + T_{j-1}\right)}. \tag{2.67}$$

This is easily shown. First, note that

$$\sigma_j^2 g_j^2(t) = \sigma_i^2 \mathcal{I}_{(t\in[0,T_{j-1}))} + \sigma_j^2 \left(\frac{T_j - at}{T_j - bT_{j-1}}\right)^2 \mathcal{I}_{(t\in[T_{j-1},T_j])}$$

Therefore, we can write

$$\begin{aligned}
\int_0^{T_j} \sigma_j^2(t) g_j^2(t) dt &= \int_0^{T_{j-1}} \sigma_j^2 dt + \int_{T_{j-1}}^{T_j} \sigma_j^2 \left(\frac{T_j - at}{T_j - bT_{j-1}}\right)^2 dt \\
&= \sigma_i^2 T_{i-1} + \frac{\sigma_i^2}{(T_i - bT_{i-1})^2} \int_{T_i - aT_{i-1}}^{T_i - aT_i} -\frac{1}{a} u^2 du \\
&= \sigma_i^2 T_{i-1} + \frac{\sigma_i^2}{(T_i - bT_{i-1})^2} \left(\frac{u^3}{3a}\Big|_{T_i - aT_i}^{T_i - aT_{i-1}}\right) \\
&= \sigma_i^2 T_{i-1} + \frac{\sigma_i^2 \left((T_i - aT_{i-1})^3 - (T_i - aT_i)^3\right)}{3a(T_i - bT_{i-1})^2}
\end{aligned}$$

where we used $u = T_i - at$ and $du = -adt$. Therefore the caplet average implied volatility is

$$\sqrt{\frac{1}{T_i}\left(\frac{\sigma_i^2 \left((T_i - aT_{i-1})^3 - (T_i - aT_i)^3\right)}{3a(T_i - bT_{i-1})^2} + \sigma_i^2 T_{i-1}\right)}$$

which we can rearrange to obtain (2.67).

Summary

In this chapter, we have constructed a fully fledged model, capable of diffusing a specified amount of forward rates, backward and forward-looking, with and without the presence of a displacement. We also briefly discussed pricing aspects and some numerical concerns are discussed in Appendix D since the next chapter will involve simulation results. The FMM presented here is the basis on which the remainder of this thesis rests. Currently, the model provides the evolution of forward rates $R(t, T_i, T_{i+1})$ for some consecutive dates T_i, T_{i+1} in our tenor structure. In practice, financial contracts might refer forward rates with different expiries and/or maturities. A natural solution would be to add such expiries and maturities to our tenor structure, and accordingly expand the number of modelled forward rates. This technique, however, adds to the computational burden of the implementation, and is therefore not desirable. The next chapter presents an alternative solution.

3 Interpolation in the FMM

The number of forward rates modeled is limited to M, as specified during the construction of the model. If M becomes too large, it will negatively impact the numerical speed of the simulations when the model is implemented. However, trading books of financial institutions often contain many contracts whose forward rates may not be specified by the chosen model. In this case we must find a way to derive the value of such forward rates by exploiting the ones present in the model. We can use a process called interpolation, which consists of generating the entire yield curve from modeled forward rates. There exists various application of interpolation in the context of interest rate modeling. A lot of the research has focused on deriving bond prices $P(t,T)$ by fitting econometric models to the yield curve, such as the Nelson-Siegel approach (see for instance Koopman et al. (2010)) which specifies a linear relationship between yields and three factors dependent of maturity (level, slope and curvature). After fitting a model, deriving the entire yield curve is straightforward given the relationship with maturities. Such a method however, provides only information on the bond prices themselves, while forward rates are *ratios* of bond prices and the numerator and denominator are not separately obtainable in the model.

Another application of interpolation consists of obtaining the initial bond price curve $P(0,T)$ for all T. This usually exploits live data from the markets by calibrating to and bootstrapping from traded products, and interpolation is then done using a cubic-spline or log-linear relationship between maturity and rates. Financial institutions tend to have stand-alone systems that provide them with such information and choices between types interpolation often relies on market conditions and traders' preferences (Darbyshire (2017)). However, these interpolation techniques are know to introduce arbitrage and are therefore not appropriate for the LMM or the FMM. The interpolation of the rates $R(t,S,T)$ for $t > 0$ must follow from no-arbitrage arguments.

Plenty of work has been realized to derive an interpolation scheme in the LMM. For example, Schlögl (2002) distinguishes between short-dated and long-dated bonds, and notes that at expiry, the bond prices become deterministic and proposes an interpolation method for the short-dated bonds accordingly, while long-dated bonds interpolation follows from no-arbitrage arguments. The FMM differs from the LMM in that rates continue to evolve stochastically at expiry and until the end of the accrual period. Furthermore, we are simulating forward-looking rates before expiry and backward-looking rates in the accrual period and any interpolation scheme we select must appropriately interpolate both types of rates. To the best of our knowledge, there has not been any research focusing on the interpolation of backward-looking forward rates, which we address in this section.

We concentrate our attention to a known issue occurring in the interpolation of forward rates in the LMM. When computing the implied volatility of caplets on interpolated forward rates, we observe an odd relation between the implied volatility and the maturity of the caplet. To be specific, the implied volatility of interpolated rate as a function of maturity has the shape of a "garland" (they look like the Christmas decoration garlands), which suggests that interpolated forward rates have a lower instantaneous volatility than simulated forward rates.

As we will see, backward-looking caplets also exhibit this unnatural behavior in the FMM. This presents an issue because it has mispricing implications for instruments whose payment dates or tenors differs from the ones in the specified tenor structure. The interpolation scheme we consider is a simple method sometimes used by practitioners and is equivalent to a linear interpolation of bond prices. Another interesting interpolation method to point out is the one from Werpachowski (2010), which solves the lower implied volatility issue in the LMM. He proposes to keep simulating the standard forward rates in the accrual period, which is already the case in the FMM because of the nature of the model. Note however that the numeraire implied from the spot measure in the LMM differs from the risk-neutral numeraire in the FMM as was pointed out in section 2.2.4. As a result, Werpachowski's interpolation will slightly differ from ours. In our journey to understanding and solving the problem of the "garland"-shaped implied volatility just mentioned, we also derive an approximation of the instantaneous volatility of interpolated forward rates. This approximation turns out to be extremely accurate, and applicable in both the FMM and LMM, and is equally applicable to forward-looking and backward-looking forward rates.

Finally, we have so far considered the inclusion of displacement in our model. It is of paramount importance that interpolated forward rates also reflect this specification. In fact, a small adjustment is needed to extend interpolation when the model contains a displacement. Fortunately, Bogt (2018) has derived this adjustment in the displaced-LMM. We show that his results extend to the FMM, thereby allowing us to interpolate forward and backward-looking rates with or without the presence of a displacement.

3.1 Numeraire-Relative-Bond interpolation

Bogt (2018) describes different properties that should preferably be fulfilled by the chosen interpolation scheme. Let us first briefly state the most important of these properties before we dig into the interpolation scheme.

(1) No additional rates should be simulated, as this was the motivation to find an interpolation scheme.

(2) No arbitrage is introduced when constructing interpolated rates according to this scheme. In other words, the interpolation scheme rests on no-arbitrage arguments.

(3) The interpolation scheme conserves standard forward rates. That is, when applied to the forward rates modeled by the FMM, the interpolation should leave these rates unchanged. This property is referred to as internal consistency.

(4) As described above, some interpolation methods generates interpolated forward rates with an abnormally low instantaneous volatility. This is observed through the implied volatility of caplets based on those rates, which exhibits an odd shape when presented as a function of time. This "garland" shaped implied volatility is shown in Figure 3.1. Therefore, a

desirable property of any interpolation scheme is that the implied volatility is a smooth function of maturity.

3.1.1 Description of the interpolation method

The interpolation method that we consider consists of linearly interpolating the ratio of the discount factors and the numeraire and it is sometimes used in practice. Hereafter, we shall refer to these quantities as Numeraire-Relative-Bonds (NRB)

$$NRB\,(t,T) = \frac{P(t,T)}{B(t)} \tag{3.1}$$

The reason for introducing these quantities is that they are easily obtainable in the FMM and the forward rates can be expressed in terms of the NRBs

$$R(t,S,T) = \frac{1}{\tau_{S,T}}\left[\frac{P(t,S)}{P(t,T)} - 1\right] = \frac{1}{\tau_{S,T}}\left[\frac{P(t,S)/B(t)}{P(t,T)/B(t)} - 1\right] = \frac{1}{\tau_{S,T}}\left[\frac{NRB(t,S)}{NRB(t,T)} - 1\right] \tag{3.2}$$

where $\tau_{,S,T}$ is the year fraction separating time S and time T. Therefore, if we can obtain $NRB(t,T)$ for all T, we have a complete model state (i.e., specifying all forward rates for a particular time t). The interpolation method we consider linearly interpolates $NRB(t,T)$ for any T bracketed by two tenors $T_{\eta(T)-1}$ and $T_{\eta(T)}$ of the tenor structure

$$NRB(t,T) := \alpha_T NRB(t,T_{\eta(T)-1}) + (1-\alpha_T)NRB(t,T_{\eta(T)}) \tag{3.3}$$

and it is equivalent to an interpolation of discount factors since

$$\begin{aligned}
&NRB(t,T) = \alpha_T NRB(t,T_{\eta(T)-1}) + (1-\alpha_T)NRB(t,T_{\eta(T)}) \\
\Leftrightarrow\ &\frac{P(t,T)}{B(t)} = \alpha_T \frac{P(t,T_{\eta(T)-1})}{B(t)} + (1-\alpha_T)\frac{P(t,T_{\eta(T)})}{B(t)} \\
\Leftrightarrow\ &P(t,T) = \alpha_T P(t,T_{\eta(T)-1}) + (1-\alpha_T)P(t,T_{\eta(T)})
\end{aligned}$$

where α_T is the coefficient of interpolation and is derived using no-arbitrage arguments. That is, by no-arbitrage (martingale property of discounted bond-prices under the risk-neutral measure) and linearity of the expectation operator, we have

$$\begin{aligned}
\frac{P(0,T)}{B(0)} &= \mathbb{E}^{\mathbb{Q}}\left[\frac{P(t,T)}{B(t)}\right] = \mathbb{E}^{\mathbb{Q}}\left[\alpha_T \frac{P(t,T_{\eta(T)-1})}{B(t)} + (1-\alpha_T)\frac{P(t,T_{\eta(T)})}{B(t)}\right] \\
&= \alpha_T\,\mathbb{E}^{\mathbb{Q}}\left[\frac{P(t,T_{\eta(T)-1})}{B(t)}\right] + (1-\alpha_T)\,\mathbb{E}^{\mathbb{Q}}\left[\frac{P(t,T_{\eta(T)})}{B(t)}\right] \\
&= \alpha_T \frac{P(0,T_{\eta(T)-1})}{B(0)} + (1-\alpha_T)\frac{P(0,T_{\eta(T)})}{B(0)}
\end{aligned}$$

Hence, if we have an initial curve $P(0,T), T \in [0,\infty)$, we can solve for the interpolation coefficient

$$\alpha_T = \frac{P(0,T) - P(0,T_{\eta(T)})}{P(0,T_{\eta-1}) - P(0,T_{\eta(T)})} \tag{3.4}$$

Such a curve can be obtained by observing a certain number traded products in the market and combining interpolation methods and bootstrapping to obtain the full curve (See for instance Darbyshire (2017)). Going forward, we will assume that we are provided with such an initial curve. Consequently, we are also given all the initial forward rates $L(0,S,T)$ for all $0 \leq S < T$.

The question of how to access the NRBs in the FMM remains. They can in fact be obtained recursively. We first compute $NRB(t, T_{\eta(T)-1})$ using (2.62), and for $t \geq T_{\eta(T)-1}$ we have

$$\begin{aligned} NRB(t, T_{\eta(T)-1}) &= \frac{P(t, T_{\eta(T)-1})}{B(t)} = \frac{P(t, T_{\eta(T)-1})}{B(T_{\eta(T)-1}) e^{\int_{T_{\eta(T)-1}}^{t} r(u)du}} = \frac{P(t, T_{\eta(T)-1})}{B(T_{\eta(T)-1}) P(t, T_{\eta(T)-1})} \\ &= \frac{1}{B(T_{\eta(T)-1})} \end{aligned} \tag{3.5}$$

where the third equality follows from the definition of extended bond price. If $t < T_{\eta(T)-1}$,

$$NRB(t, T_{\eta(T)-1}) = \frac{P(t, T_{\eta(T)-1})}{B(t)} = \frac{\mathbb{E}\left[e^{-\int_t^{T_{\eta(T)-1}} r(u)du} \,\middle|\, \mathcal{F}_t \right]}{B(T_{\eta(T)-1}) e^{-\int_t^{T_{\eta(T)-1}} r(u)du}} = \frac{1}{B(T_{\eta(T)-1})} \tag{3.6}$$

Next, from (3.2), we have

$$\begin{aligned} R_j(t) &= \frac{1}{\tau_j} \left(\frac{NRB(t, T_{j-1})}{NRB(t, T_j)} - 1 \right) \\ \Leftrightarrow NRB(t, T_j) &= NRB(t, T_{j-1}) \frac{1}{1 + \tau_j R_j(t)} \end{aligned} \tag{3.7}$$

and therefore we can obtain $NRB(t, T_{\eta(T)})$ recursively. The same arguments hold to derive $NRB(t, T_{\eta(S)-1})$ and $NRB(t, T_{\eta(S)})$.

Let us find a complete analytical expression for the interpolated forward rate

$$1 + \tau_{S,T} R(t, S, T) = \frac{NRB(t, S)}{NRB(t, T)} \tag{3.8}$$

under the method (3.3). The numerator can be written as

$$\begin{aligned} NRB(t, S) &= \alpha_S NRB(t, T_{\eta(S)-1}) + (1 - \alpha_S) NRB(t, T_{\eta(S)}) \\ &= \alpha_S \left(\frac{1}{B(T_{\eta(S)-1})} \right) + (1 - \alpha_S) \left(\frac{1}{B(T_{\eta(S)-1})} \frac{1}{1 + \tau_{\eta(S)} R_{\eta(S)}(t)} \right) \\ &= \frac{1}{B(T_{\eta(S)-1})} \left(\alpha_S + \frac{1 - \alpha_S}{1 + \tau_{\eta(S)} R_{\eta(S)}(t)} \right) \\ &= \frac{1}{B(T_{\eta(S)-1})} \left(\frac{1 + \alpha_S \tau_{\eta(S)} R_{\eta(S)}(t)}{1 + \tau_{\eta(S)} R_{\eta(S)}(t)} \right) \end{aligned} \tag{3.9}$$

Similarly, the denominator is given by

$$NRB(t, T) = \alpha_T NRB(t, T_{\eta(T)-1}) + (1 - \alpha_T) NRB(t, T_{\eta(T)})$$

$$= \alpha_T \left(\frac{1}{B(T_{\eta(S)-1})} \prod_{j=\eta(S)}^{\eta(T)-1} \frac{1}{1+\tau_j R_j(t)} \right) + (1-\alpha_T) \left(\frac{1}{B(T_{\eta(S)-1})} \prod_{j=\eta(S)}^{\eta(T)} \frac{1}{1+\tau_j R_j(t)} \right)$$

$$= \left(\frac{1}{B(T_{\eta(S)-1})} \prod_{j=\eta(S)}^{\eta(T)-1} \frac{1}{1+\tau_j R_j(t)} \right) \left(\alpha_T + \frac{1-\alpha_T}{1+\tau_{\eta(T)} R_{\eta(T)}(t)} \right)$$

$$= \frac{1}{B(T_{\eta(S)-1})} \left(\prod_{j=\eta(S)}^{\eta(T)-1} \frac{1}{1+\tau_j R_j(t)} \right) \left(\frac{1+\alpha_T \tau_{\eta(T)} R_{\eta(T)}(t)}{1+\tau_{\eta(T)} R_{\eta(T)}(t)} \right) \tag{3.10}$$

Now, we can rewrite the interpolated rates as

$$1+\tau_{S,T} R(t,S,T) = \frac{\frac{1+\alpha_S \tau_{\eta(S)} R_{\eta(S)}(t)}{1+\tau_{\eta(S)} R_{\eta(S)}(t)}}{\frac{1+\alpha_T \tau_{\eta(T)} R_{\eta(T)}(t)}{1+\tau_{\eta(T)} R_{\eta(T)}(t)} \prod_{j=\eta(S)}^{\eta(T)-1} \frac{1}{1+\tau_j R_j(t)}}$$

$$= \frac{1+\alpha_S \tau_{\eta(S)} R_{\eta(S)}(t)}{1+\alpha_T \tau_{\eta(T)} R_{\eta(T)}(t)} \frac{1+\tau_{\eta(T)} R_{\eta(T)}(t)}{1+\tau_{\eta(S)} R_{\eta(S)}(t)} \prod_{j=\eta(S)}^{\eta(T)-1} 1+\tau_j R_j(t)$$

$$\Leftrightarrow 1+\tau_{S,T} R(t,S,T) = \frac{1+\alpha_S \tau_{\eta(S)} R_{\eta(S)}(t)}{1+\alpha_T \tau_{\eta(T)} R_{\eta(T)}(t)} \prod_{j=\eta(S)+1}^{\eta(T)} 1+\tau_j R_j(t) \tag{3.11}$$

and (3.11) can provide us with both the forward-looking and the backward-looking interpolated forward rate by either setting $t = S$ (forward) or $t = T$ (backward). We refer equally to either of (3.3) or (3.11) for the same interpolation method.

Note that the interpolation scheme (3.3) is not directly applicable if the rate displacement $\delta > 0$. Therefore, in the next section, we present an extension to apply it in a displaced FMM.

3.1.2 Displacement-aware interpolation

To account for a non-zero displacement when interpolating forward rates, we compute a "displacement-aware" interpolation coefficient. This was done in Bogt (2018) in the LMM and we confirm that his results are applicable in the FMM, with a few minor changes (he in fact interpolates $\frac{1}{B_d(t)}$ but this quantity does not correspond to $\frac{1}{B(t)}$ and therefore the interpolation in practice is slightly different). We first note that α_T can be written as the ratio of two forward rates known at time 0 multiplied by their tenor length. Divide the numerator and the denominator by $P(0, T_{q(T)})$

$$\alpha_T = \frac{\frac{P(0,T)-P(0,T_{\eta(T)})}{P(0,T_{\eta(T)})}}{\frac{P(0,T_{\eta-1})-P(0,T_{\eta(T)})}{P(0,T_{\eta(T)})}} = \frac{\frac{P(0,T)}{P(0,T_{\eta(T)})}-1}{\frac{P(0,T_{\eta-1})}{P(0,T_{\eta(T)})}-1} = \frac{\tau_{T,T_{\eta(T)}}}{\tau_{\eta(T)}} \frac{R(0,T,T_{\eta(T)})}{R(0,T_{\eta(T)-1},T_{\eta(T)})} \tag{3.12}$$

where $\tau_{T,T_{\eta(T)}}$ is the year fraction separating T and $T_{\eta(T)}$. Now, it is reasonable to assume that in the displaced case, the interpolation coefficient, denoted $\widehat{\alpha}_T$, becomes

$$\widehat{\alpha}_T = \frac{\tau_{T,T_{\eta(T)}}}{\tau_{\eta(T)}} \frac{\widetilde{R}(0,T,T_{\eta(T)})}{\widetilde{R}(0,T_{\eta(T)-1},T_{\eta(T)})} = \frac{\tau_{T,T_{\eta(T)}}}{\tau_{\eta(T)}} \frac{R(0,T,T_{\eta(T)})+\delta}{R(0,T_{\eta(T)-1},T_{\eta(T)})+\delta}$$

$$
\begin{aligned}
&= \frac{\frac{P(0,T)-P(0,T_{\eta(T)})}{P(0,T_{\eta(T)})} + \delta\tau_{T,T_{\eta(T)}}}{\frac{P(0,T_{\eta-1})-P(0,T_{\eta(T)})}{P(0,T_{\eta(T)})} + \delta\tau_{\eta(T)}} \\
&= \frac{P(0,T) - P(0,T_{\eta(T)}) + \delta\tau_{T,T_{\eta(T)}}P(0,T_{\eta(T)})}{P(0,T_{\eta(T)-1}) - P(0,T_{\eta(T)}) + \delta\tau_{\eta(T)}P(0,T_{\eta(T)})}
\end{aligned} \tag{3.13}
$$

We also use an adjustment for (3.3) derived by Bogt (2018) in the LMM (It turns out to be applicable to the FMM as well), which is given by

$$
\begin{aligned}
NRB(t,T) = \alpha_T NRB(t,T_{\eta(T)-1}) &+ (1-\alpha_T)NRB(t,T_{\eta(T)}) \\
&+ \delta\left(\widehat{\alpha}_T\tau_{\eta(T)} - \tau_{T,T_{\eta(T)}}\right)NRB(t,T_{\eta(T)})
\end{aligned} \tag{3.14}
$$

Note that if $\delta = 0$, we fall back to (3.3). For the numerator of (3.8), we now have

$$
\begin{aligned}
NRB(t,S) &= \widehat{\alpha}_S NRB(t,T_{\eta(S)-1}) + (1-\widehat{\alpha}_S)NRB(t,T_{\eta(S)}) + \delta\left(\widehat{\alpha}_S\tau_{\eta(S)} - \tau_{S,T_{\eta(S)}}\right)NRB(t,T_{\eta(S)}) \\
&= \widehat{\alpha}_S\left(\frac{1}{B(T_{\eta(S)-1})}\right) + \left(1-\widehat{\alpha}_S + \delta\left(\widehat{\alpha}_S\tau_{\eta(S)} - \tau_{S,T_{\eta(S)}}\right)\right)NRB(t,T_{\eta(S)}) \\
&= \widehat{\alpha}_S\left(\frac{1}{B(T_{\eta(S)-1})}\right) + \left(1-\widehat{\alpha}_S + \delta\left(\widehat{\alpha}_S\tau_{\eta(S)} - \tau_{S,T_{\eta(S)}}\right)\right)\left(\frac{1}{B(T_{\eta(S)-1}))}\frac{1}{1+\tau_{\eta(S)}R_{\eta(S)}(t)}\right) \\
&= \frac{1}{B(T_{\eta(S)-1}))}\left(\widehat{\alpha}_S + \frac{1-\widehat{\alpha}_S + \delta\left(\widehat{\alpha}_S\tau_{\eta(S)} - \tau_{S,T_{\eta(S)}}\right)}{1+\tau_{\eta(S)}R_{\eta(S)}(t)}\right) \\
&= \frac{1}{B(T_{\eta(S)-1}))}\left(\frac{1+\widehat{\alpha}_S\tau_{\eta(S)}\left(R_{\eta(S)}(t)+\delta\right) - \delta\tau_{S,T_{\eta(S)}}}{1+\tau_{\eta(S)}R_{\eta(S)}(t)}\right) \\
&= \frac{1}{B(T_{\eta(S)-1}))}\left(\frac{1+\widehat{\alpha}_S\tau_{\eta(S)}\widetilde{R}_{\eta(S)}(t) - \delta\tau_{S,T_{\eta(S)}}}{1+\tau_{\eta(S)}R_{\eta(S)}(t)}\right)
\end{aligned} \tag{3.15}
$$

For the denominator,

$$
\begin{aligned}
NRB(t,T) &= \widehat{\alpha}_T NRB(t,T_{\eta(T)-1}) + (1-\widehat{\alpha}_T)NRB(t,T_{\eta(T)}) \\
&\quad + \delta\left(\widehat{\alpha}_T\tau_{\eta(T)} - \tau_{T,T_{\eta(T)}}\right)NRB(t,T_{\eta(T)}) \\
&= \widehat{\alpha}_T\left(\frac{1}{B(T_{\eta(S)-1})}\prod_{j=\eta(S)}^{\eta(T)-1}\frac{1}{1+\tau_j R_j(t)}\right) \\
&\quad + \left(1-\widehat{\alpha}_T + \delta\left(\widehat{\alpha}_T\tau_{\eta(T)} - \tau_{T,T_{\eta(T)}}\right)\right)\left(\frac{1}{B(T_{\eta(S)-1})}\prod_{j=\eta(S)}^{\eta(T)}\frac{1}{1+\tau_j R_j(t)}\right) \\
&= \left(\frac{1}{B(T_{\eta(S)-1})}\prod_{j=\eta(S)}^{\eta(T)-1}\frac{1}{1+\tau_j R_j(t)}\right)\left(\widehat{\alpha}_T + \frac{1-\widehat{\alpha}_T + \delta\left(\widehat{\alpha}_T\tau_{\eta(T)} - \tau_{T,T_{\eta(T)}}\right)}{1+\tau_{\eta(T)}R_{\eta(T)}(t)}\right) \\
&= \frac{1}{B(T_{\eta(S)-1})}\left(\prod_{j=\eta(S)}^{\eta(T)-1}\frac{1}{1+\tau_j R_j(t)}\right)\left(\frac{1+\widehat{\alpha}_T\tau_{\eta(T)}\widetilde{R}_{\eta(T)}(t) - \delta\tau_{T,T_{\eta(T)}}}{1+\tau_{\eta(T)}R_{\eta(T)}(t)}\right)
\end{aligned} \tag{3.16}
$$

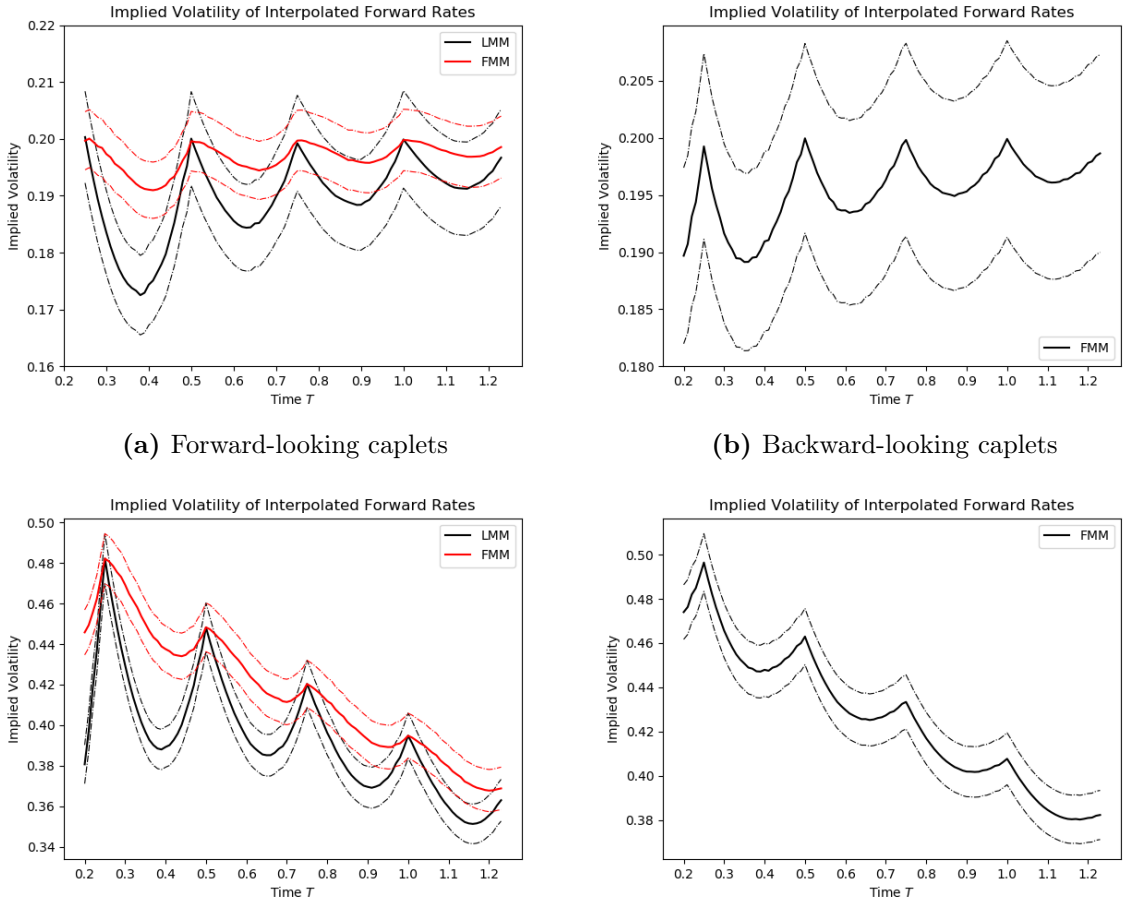
Taking the ratio, we can rewrite the interpolated rates as

$$
\begin{aligned}
1+\tau_{S,T}\widetilde{R}(t,S,T) &= \frac{\frac{1+\widehat{\alpha}_S\tau_{\eta(S)}\widetilde{R}_{\eta(S)}(t)-\delta\tau_{S,T_{\eta(S)}}}{1+\tau_{\eta(S)}R_{\eta(S)}(t)}}{\frac{1+\widehat{\alpha}_T\tau_{\eta(T)}\widetilde{R}_{\eta(T)}(t)-\delta\tau_{T,T_{\eta(T)}}}{1+\tau_{\eta(T)}R_{\eta(T)}(t)}\prod_{j=\eta(S)}^{\eta(T)-1}\frac{1}{1+\tau_jR_j(t)}} \\
&= \frac{1+\widehat{\alpha}_S\tau_{\eta(S)}\widetilde{R}_{\eta(S)}(t)-\delta\tau_{S,T_{\eta(S)}}}{1+\widehat{\alpha}_T\tau_{\eta(T)}\widetilde{R}_{\eta(T)}(t)-\delta\tau_{T,T_{\eta(T)}}}\frac{1+\tau_{\eta(T)}R_{\eta(T)}(t)}{1+\tau_{\eta(S)}R_{\eta(S)}(t)}\prod_{j=\eta(S)}^{\eta(T)-1}1+\tau_jR_j(t) \\
\Leftrightarrow 1+\tau_{S,T}\widetilde{R}(t,S,T) &= \frac{1+\widehat{\alpha}_S\tau_{\eta(S)}\widetilde{R}_{\eta(S)}(t)-\delta\tau_{S,T_{\eta(S)}}}{1+\widehat{\alpha}_T\tau_{\eta(T)}\widetilde{R}_{\eta(T)}(t)-\delta\tau_{T,T_{\eta(T)}}}\prod_{j=\eta(S)+1}^{\eta(T)}1+\tau_jR_j(t) \qquad (3.17)
\end{aligned}
$$

3.1.3 Lower instantaneous volatility of interpolated forward rates

We now demonstrate interpolation method (3.3) (and (3.14) if $\delta \neq 0$) and its implications on caplet pricing in an example. We consider an at-the-money caplet on the forward rate $R(T^*, T, T+\gamma)$, where we set $\gamma = 0.25$ and we plot the implied volatility of this caplet against $T \in [0.25, 1.25]$ in Figure 3.1, along with the 95% Monte Carlo confidence interval. Figure 3.1a shows the result for a forward-looking caplet ($T^* = T$) while Figure 3.1b concerns a backward-looking caplet ($T^* = T+\gamma$) which, we recall, can only be priced in the FMM. The forward rates were all simulated using a fixed input volatility, i.e. $\sigma_j = 20\%\ \forall j = 1, ..., M$. We consider two models, namely the LMM and the FMM. The interpolation method derived in this section is applicable to the LMM as well. The difference is that the LMM dynamics are derived under the spot measure, described in section 2.2.4. As a consequence, we replace $B(t)$ by the corresponding numeraire $B_d(t)$ in the equations above. We can see that this interpolation method exhibits the so-called volatility dips in both models. When T is one of the standard dates from our tenor structure, the implied volatilities in the graphs are equal to the model input volatility of 20% (with a certain confidence interval shown by the dotted lines). In between the tenor structure dates, the implied volatility plunges and reaches a minimum roughly halfway between the two closest standard dates. We note that the minimum of each volatility dip is more stressed for caplets whose expiry is closer to T_0. A possible explanation is that the mispricing comes mostly from the accrual period of the interpolated forward rates (this will become clear in the next section). For caplets with a short lifetime, the accrual period represents a larger proportion of the total lifetime of the instrument. If the mispricing occurs in that period, it will have a larger effect for these caplets, and the effect will smooth out as the expiry is further in time.

We also point out that this mispricing is much less exacerbated in the FMM and the minimum seems to be further away from the midpoint of two standard dates. This is likely related to the shape of the linearly decaying function $g(t)$ and the instantaneous volatility is lower towards the end of the accrual period and the mispricing is therefore more severe in that period. It is intriguing that the implied volatility increases again after reaching a minimum (rather than sharply recovering to 20% just before the next standard tenor date). As we will show, the instantaneous volatility of interpolated rate is in fact a weighted average of the volatility of the previous and the next forward rates. As the expiry of the interpolated rate approaches the next standard date, progressively more weight is put on the volatility of the next forward rate.



(a) Forward-looking caplets

(b) Backward-looking caplets

(c) Forward-looking caplets

(d) Backward-looking caplets

Figure 3.1: Implied volatility of caplets on interpolated forward rates. The dotted line shows the 95% Monte Carlo confidence intervals. A fairly accurate price was reached with 3000 simulations for constant volatility and 20000 for exponential volatility.

Setting $\sigma_j(t) = \sigma_j$ is a simplistic situation rarely used in practice since there is no calibration involved. We will now repeat the above experiment when rates are simulated in combination with an exponentially decreasing function, given by (2.50) and where we set $a = c = 0$, $b = 0.4$, $d = 1$ and $\Phi_j = 0.6$, $\forall j = 1, \ldots, M$. We show the results in Figure C.1. We observe the same behaviors as in the case of a constant volatility. Namely, for forward-looking caplets in Figure 3.1c, our interpolation method produce garlands in the implied volatility functions under both the LMM and the FMM, although this problem is again substantially less severe in the FMM. Observations previously made on the location of the largest mispricing also hold here. Again, in backward-looking caplets (Figure 3.1d), this mispricing is also present.

We also show that the displacement-aware interpolation method (3.14) agrees with the above observations. We left the resulting graphs in Appendix C due to their similarity to the non-displaced case (see Figures C.1a and C.1b for constant instantaneous volatility, and Figures C.1c and C.1d for exponential volatility). Also note that all simulations were realized with an identical seed for replicability and comparison purposes. This leads to the conclusion that the adjustments derived by Bogt (2018) (equations (3.13) and (3.14)) perfectly fit in the context of a displaced FMM.

The examples shown in this section stress the need for a better interpolation method to overcome the issue of lower implied volatility of interpolated forward rates. Before we tackle the search for a better alternative, we will first investigate whether the analytical expression of interpolated forward rates can help us understand this issue, for instance by analyzing the instantaneous volatility of these rates in the next section.

3.2 Approximation of the volatility of interpolated rates

In this section, we derive an analytical approximation of the instantaneous volatility of interpolated forward rates and we show that the "garland"-shaped implied volatility function, resulting from Monte Carlo simulation, can in fact be predicted accurately using this approximation and will be applicable to both the LMM and the FMM. To the best of our knowledge, such an approximation has not yet been derived in the LMM.

First, (3.11) shows that an interpolated rate is a linear combination of standard forward rates (just as swap rates are, see Brigo and Mercurio (2006)), which are stochastic in nature. Therefore, they should also have a stochastic differential equation which is given by

$$d(1+\tau_{S,T}R(t,S,T)) = \tau_{S,T}dR(t,S,T) = d\left(\frac{1+\alpha_S\tau_{\eta(S)}R_{\eta(S)}(S)}{1+\alpha_T\tau_{\eta(T)}R_{\eta(T)}(S)}\prod_{j=\eta(S)+1}^{\eta(T)}1+\tau_jR_j(t)\right)$$

Consider the situation with $T \in [T_{\eta(S)}, T_{\eta(S)+1}]$, and to simplify notation let $\eta(S) = j$. Thus, $\eta(T) = j+1$ and we can rewrite the above as

$$d(1+\tau_{S,T}R(t,S,T)) = d\left(\frac{1+\alpha_S\tau_jR_j(t)}{1+\alpha_T\tau_{j+1}R_{j+1}(t)}(1+\tau_{j+1}R_{j+1}(t))\right)$$

Using the product rule for stochastic processes, and ignoring the deterministic terms involving dt (since we are only interest in the diffusion coefficient, and hence the volatility), we have

$$\begin{aligned}
d(1+\tau_{S,T}R(t,S,T)) &= \left(\frac{1+\alpha_S\tau_jR_j(t)}{1+\alpha_T\tau_{j+1}R_{j+1}(t)}\right)\tau_{j+1}dR_{j+1}(t) \\
&\quad + (1+\tau_{j+1}R_{j+1}(t))d\left(\frac{1+\alpha_S\tau_jR_j(t)}{1+\alpha_T\tau_{j+1}R_{j+1}(t)}\right) \\
&= \left(\frac{1+\alpha_S\tau_jR_j(t)}{1+\alpha_T\tau_{j+1}R_{j+1}(t)}\right)\tau_{j+1}\sigma_{j+1}g_{j+1}R_{j+1}dW(t) \\
&\quad + (1+\tau_{j+1}R_{j+1}(t))d\left(\frac{1+\alpha_S\tau_jR_j(t)}{1+\alpha_T\tau_{j+1}R_{j+1}(t)}\right)
\end{aligned}$$

Applying the product rule again, for the diffusion of the term involving the ratio, we have

$$\begin{aligned}
d\left(\frac{1+\alpha_S\tau_jR_j(t)}{1+\alpha_T\tau_{j+1}R_{j+1}(t)}\right) &= \frac{\alpha_S\tau_jdR_j(t)}{1+\alpha_T\tau_{j+1}R_{j+1}(t)} - \frac{\alpha_T\tau_{j+1}dR_{j+1}(t)(1+\alpha_S\tau_jR_j(t))}{(1+\alpha_T\tau_{j+1}R_{j+1}(t))^2} \\
= \frac{\alpha_S\tau_j\sigma_j(t)g_j(t)R_j(t)dW(t)}{1+\alpha_T\tau_{j+1}R_{j+1}(t)} &- \frac{\alpha_T\tau_{j+1}\sigma_{j+1}(t)g_{j+1}(t)(1+\alpha_S\tau_jR_j(t))R_{j+1}(t)dW(t)}{(1+\alpha_T\tau_{j+1}R_{j+1}(t))^2}
\end{aligned}$$

With these equations, we can rewrite the dynamics of a given interpolated forward rate as a

weighted average of the dynamics of forward rates "bracketing" it. Let $\sigma_{S,T}(t)$ be the volatility of interpolated forward rates, their dynamics is thus given by

$$\begin{aligned}\tau_{S,T}\sigma_{S,T}(t)R(t,S,T)dW(t) &= h(t)\tau_j\sigma_j(t)g_j(t)R_j(t)dW(t) + f(t)\tau_{j+1}\sigma_{j+1}(t)g_{j+1}(t)R_{j+1}(t)dW(t) \\ &= h(t)dR_j(t) + f(t)dR_{j+1}(t) \end{aligned} \tag{3.18}$$

where

$$h(t) = \frac{\alpha_S(1+\tau_{j+1}R_{j+1}(t))}{1+\alpha_T\tau_{j+1}R_{j+1}(t)} \tag{3.19}$$

and

$$\begin{aligned} f(t) &= \frac{1+\alpha_S\tau_jR_j(t)}{1+\alpha_T\tau_{j+1}R_{j+1}(t)} - \frac{\alpha_T(1+\alpha_S\tau_jR_j(t))(1+\tau_{j+1}R_{j+1}(t))}{(1+\alpha_T\tau_{j+1}R_{j+1}(t))^2} \\ &= \frac{(1+\alpha_T\tau_{j+1}R_{j+1}(t))\,(1+\alpha_S\tau_jR_j(t)) - \alpha_T(1+\alpha_S\tau_jR_j(t))(1+\tau_{j+1}R_{j+1}(t))}{(1+\alpha_T\tau_{j+1}R_{j+1}(t))^2} \\ &= \frac{(1+\alpha_S\tau_jR_j(t))\,(1+\alpha_T\tau_{j+1}R_{j+1}(t) - \alpha_T - \alpha_T\tau_{j+1}R_{j+1}(t))}{(1+\alpha_T\tau_{j+1}R_{j+1}(t))^2} \\ &= \frac{(1-\alpha_T)(1+\alpha_S\tau_jR_j(t))}{(1+\alpha_T\tau_{j+1}R_{j+1}(t))^2} \end{aligned} \tag{3.20}$$

Following the spirit of Jäckel and Rebonato (2003), to approximate $\sigma_{S,T}$, we fix the forward rates at their time-0 values (which are available from the initial bond curve). Denote the estimate of $\sigma_{S,T}$ by $\widetilde{\sigma}_{S,T}(t)$. Rearranging equation (3.18) and assuming that interpolated rates have the same tenor as the standard forward rates (i.e. $\tau_{S,T} = \tau_j = \tau_{j+1}$), we have

$$\widetilde{\sigma}_{S,T}(t) = \frac{h(0)\sigma_j(t)g_j(t)R_j(0) + f(0)\sigma_{j+1}(t)g_{j+1}(t)R_{j+1}(0)}{R(0,S,T)} \tag{3.21}$$

with

$$h(0) = \frac{\alpha_S(1+\tau_{j+1}R_{j+1}(0))}{1+\alpha_T\tau_{j+1}R_{j+1}(0)} \quad \text{and} \quad f(0) = \frac{(1-\alpha_T)(1+\alpha_S\tau_jR_j(0))}{(1+\alpha_T\tau_{j+1}R_{j+1}(0))^2} \tag{3.22}$$

We now see that the implied volatility of the interpolated forward rate $R(t,S,T)$ is a function of the deterministic functions $g_j(t)$ and $g_{j+1}(t)$ and also $\sigma_j(t)$ and $\sigma_{j+1}(t)$. Note what happens when we let S and T be dates from the tenor structure, say T_j and T_{j+1} respectively, for some $j \in [1,...,M-1]$. We have $h(0) = 0$ and $f(0) = 1$ since $\alpha_S = \alpha_T = 0$ and therefore,

$$\widetilde{\sigma}_{S,T}(t) = \widetilde{\sigma}_{T_j,T_{j+1}}(t) = \frac{\sigma_{j+1}(t)g_{j+1}(t)R_{j+1}(0)}{R(0,T_j,T_{j+1})} = \sigma_{j+1}(t)g_{j+1}(t)$$

which indicates this approximation is consistent when the interpolation is applied to standard forward rates.

For non-standard dates, Figure 3.2 shows a plot of the implied volatility of $R(T^*,T,T+\gamma)$ (continuous lines) against $\widetilde{\lambda}_{T,T+\delta}(T^*)$ (dashed lines) as a function of T , where

$$\widetilde{\lambda}_{S,T}(t) = \sqrt{\frac{1}{t}\int_0^t \widetilde{\sigma}^2_{S,T}(s)ds}$$

is the average Black implied volatility constructed with our approximation and we set $T^* = T$ for the forward-looking caplets and $T^* = T + \gamma$ for the backward-looking caplets (if the approximation is correct, $\widetilde{\lambda}_{T,T+\delta}(T^*)$ will be close to the implied volatility of the caplet on the corresponding forward rate). We clearly see that for both types of rates in the FMM, the approximation is excellent. This is also the case for exponentially decreasing volatility which is shown in Figures 3.2c and 3.2d (forward and backward-looking caplets respectively). For the LMM, the approximation can only be used on forward-looking rates. Note that the LMM is a special case of the FMM when $g_j(t) = \mathcal{I}_{\{t \leq T_{j-1}\}}$ (and adapted numeraire). If we adapt $g_j(t)$ and $g_{j+1}(t)$ accordingly in (3.18), we see no reason against using this approximation in the LMM. Figures 3.2a and 3.2c indeed shows that this approximation is very accurate in both constant and exponentially decreasing instantaneous volatility of forward rates when applied to the LMM.

The approximation (3.21) naturally needs to be adjusted when we include a displacement in the model. To derive a "displacement-aware" approximation, we must follow the same steps as before and apply the product rule for stochastic processes to (3.17) instead of (3.11). These derivations are similar to what was done in this section but slightly more cumbersome, so we leave them in Appendix A. We nonetheless state the resulting formula in the following proposition.

Proposition 3.2.1. *The interpolated forward rate $\widetilde{R}(t,S,T)$ has a stochastic differential equation given by*

$$\begin{aligned}\tau_{S,T}\widetilde{\sigma}_{S,T}(t)\widetilde{R}(t,S,T)dW(t) = h(t)\tau_j\sigma_j(t)g_j(t)\widetilde{R}_j(t)dW(t) \\ + f(t)\tau_{j+1}\sigma_{j+1}(t)g_{j+1}(t)\widetilde{R}_{j+1}(t)dW(t)\end{aligned} \tag{3.23}$$

where

$$h(t) = \frac{\widehat{\alpha}_S\left(1+\tau_{j+1}R_{j+1}(t)\right)}{1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}} \tag{3.24}$$

and

$$f(t) = \frac{\left(1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}\right)\left(1-\widehat{\alpha}_T(1-\tau_{j+1}\delta)-\delta\tau_{t,T_{j+1}}\right)}{\left(1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}\right)^2} \tag{3.25}$$

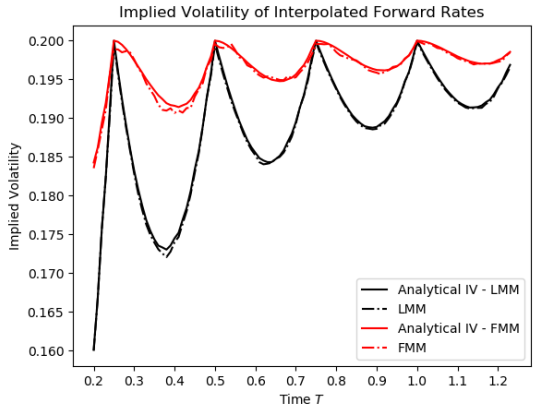
Furthermore, $\widetilde{\sigma}_{S,T}(t)$ can be approximated by

$$\widetilde{\sigma}_{S,T}(t) = \frac{h(0)\sigma_j(t)g_j(t)\widetilde{R}_j(0)+f(0)\sigma_{j+1}(t)g_{j+1}(t)\widetilde{R}_{j+1}(0)}{\widetilde{R}(0,S,T)} \tag{3.26}$$

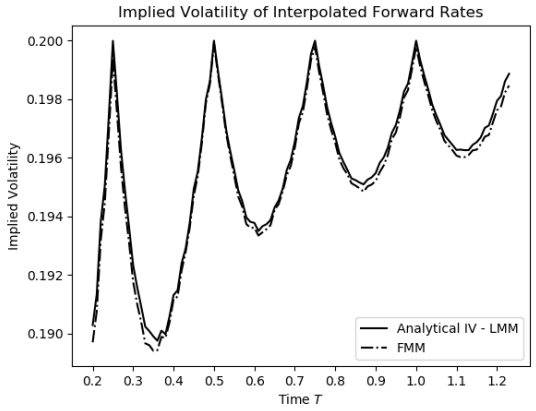
Note that if $\delta = 0$, i.e. there is no displacement in the model, $h(t)$ and $f(t)$ simplify to equations (3.19) and (3.20) respectively.

Proof. The proof is given in Appendix A. □

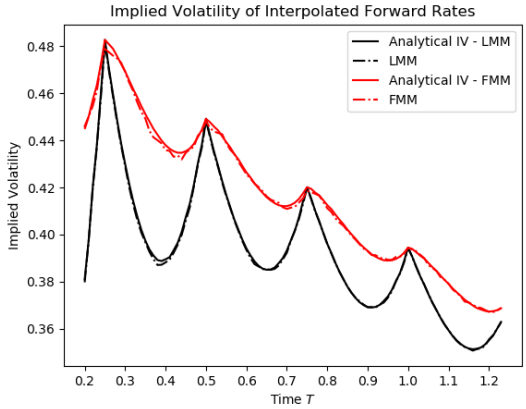
To corroborate the claim that (3.26) is indeed as accurate with displacement as it is without, we approximate the implied volatility of caplets on the same interpolated rates as in previous



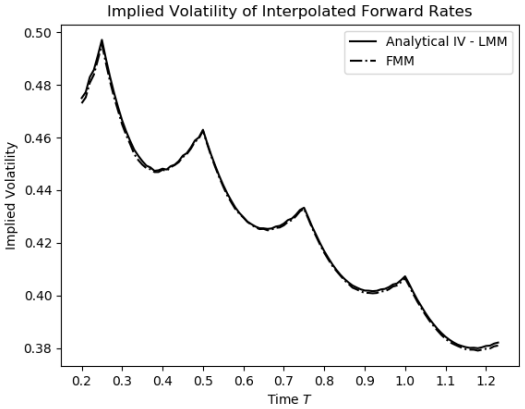
(a) Forward-looking rates with $\sigma_j(t) = 20\% \forall t$



(b) Backward-looking rates with $\sigma_j(t) = 20\%\ \forall t$



(c) Forward-looking rates with volatility $\sigma_j(t)$ given by (2.50)



(d) Backward-looking rates with volatility $\sigma_j(t)$ given by (2.50)

Figure 3.2: Implied volatility of forward rates and their approximations given by (3.18). Rates were simulated in the LMM and the FMM using 4000 simulations.

examples. Since they are nearly identical to the non-displaced case, we have left the results in the Appendix in Figure C.3. We consider again the same situations, forward and backward-looking, LMM and FMM, constant or exponentially decreasing volatility but each case drives us to the same conclusion: our approximation of the implied volatility of interpolated rates is very accurate.

3.3 Smooth-volatility interpolation

As mentioned earlier, the different figures which demonstrated our approximation is in fact showing

$$\widetilde{\lambda}_{S,T}(t) = \sqrt{\frac{1}{t}\int_0^t \widetilde{\sigma}^2_{S,T}(s)ds} \tag{3.27}$$

against the average Black Implied volatility of the corresponding caplet

$$\lambda_{S,T}(t) = \sqrt{\frac{1}{t}\int_0^t \sigma^2_{S,T}(s)ds} \tag{3.28}$$

where $\sigma_{S,T}$, the true interpolated rate instantaneous volatility, is in fact unknown. Jäckel (2006) describes a numerical solution which uses a root-search algorithm to obtain the implied volatility from caplet premiums. In the FMM, we have the freedom to specify the function $g_j(t)$ for each of the forward rates. As we also noted, previous research suggests that a setting which keeps simulating forward rates in the accrual period in combination with (3.3) solves the mispricing issue (Werpachowski (2010)). This would in fact correspond to setting $g_j(t) = \mathcal{I}_{\{t \leq T_j\}}$ in the FMM. There are two ways to confirm this or to find any other $g_j(t)$ (if any exists) that would solve this issue. Namely, repeating the experiments above with a different specification for $g_j(t)$ or implement a program that will numerical solve for $g_j(t)$ in (3.27) such that we set $\widetilde{\lambda}_{S,T}(t)$ equal to the desired $\lambda_{S,T}(t)$. This is one of the purposes of the analytical approximation. In the forward-looking case, we have been successful in this task, and both approaches leads to the conclusion that $g_j(t) = \mathcal{I}_{\{t \leq T_j\}}$ is optimal to remove the mispricing in caplets. For backward-looking caplets, trial-and-error and numerical solvers both failed the attempt to find an appropriate function $g_j(t)$. In fact, the numerical results suggested that such a function simply does not exist. Let us first focus on the forward-looking case.

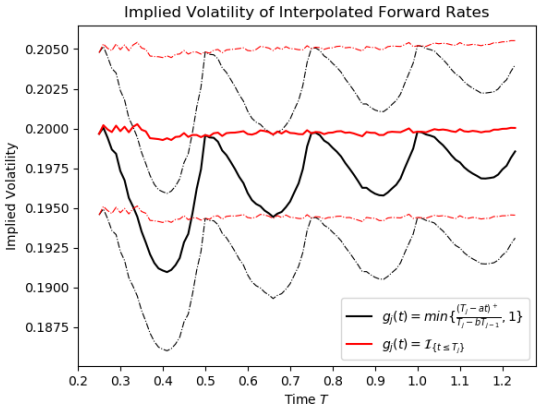
The mispricing brought by the NRB interpolation method, and which was known to be present in the LMM with other methods (see for instance Schlögl (2002)) is less serious in the FMM but equally present in forward-looking and backward-looking instruments. A previous thesis written by Bogt (2018) has researched a method to remedy this issue in the displaced LMM. Originally derived by Werpachowski (2010), that method explains that the interpolation involves a rate that has stopped evolving at some point and therefore has a zero volatility, which may explain the lower premium associated with caplets priced on such rates. His method therefore suggests to keep that forward rate alive for the sole purpose of the interpolation. The interpolated rates successfully ceased to exhibit a lower volatility and Bogt (2018) has extended his results to the displaced-LMM.

The method of Werpachowski (2010) is, in essence, an extension of the life of forward rates to include the accrual period. That is precisely what the FMM specifies. Furthermore, this method keeps simulating the rate without decaying the diffusion coefficients while the FMM lets the practioner freely select the desired function $g_j(t)$. Therefore, we test the interpolation method (3.14) by setting

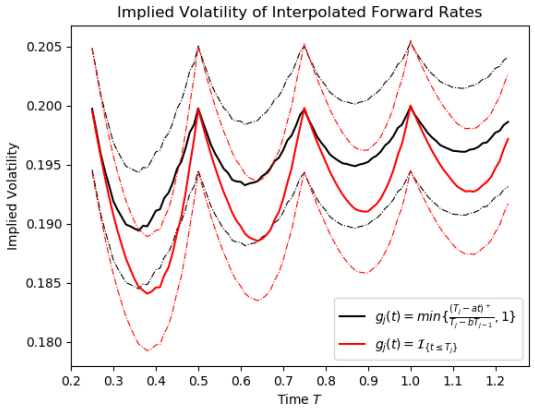
$$g_j(t) = \mathcal{I}_{\{t \leq T_j\}} \tag{3.29}$$

in (2.49). We now focus on the displaced model and the non-displaced case will follow from setting the displacement parameter $\delta = 0$.

The results for two different specifications of $g_j(t)$ are shown in Figure 3.3 (we now ignore interpolation in the LMM). For the forward-looking case, Figure 3.3a shows that the mispricing has disappeared, and the implied volatility function seems to be a smooth function of T (subject to numerical inaccuracy), which is in line with the results of Werpachowski (2010) (See Appendix C for results in a displaced model). A numerical solver as described also resulted in $g_j(t) = 1$ for $t \leq T_j$ (we omit implementation details but it is based on a root-search). We plot the approximation (3.26) for two specifications of $\sigma_j(t)$ in Figures 3.4 and show that the approximation is equally strong for displaced and non-displaced cases. Without the different colors of the line plot, it is actually almost impossible to distinguish



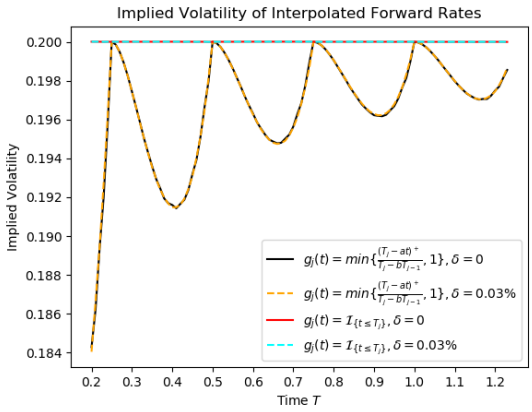
(a) Forward-looking rates with $\delta = 0$



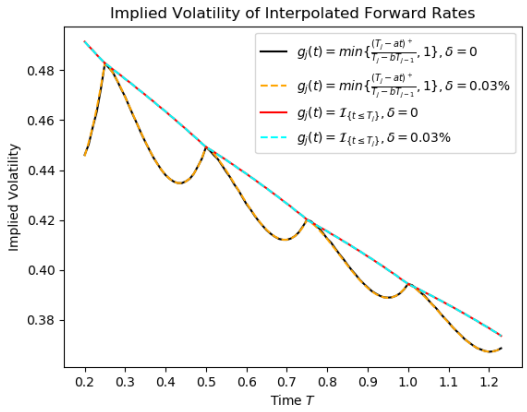
(b) Backward-looking rates with $\delta = 0$

Figure 3.3: Implied volatility of forward rates for different diffusion scaling coefficient $g_j(t)$, and the 95% Monte Carlo confidence interval. Rates were simulated using 10000 simulations.

between the two. The approximation also suggests that even when $\sigma_j(t)$ is an exponentially decreasing function, $g_j(t) = \mathcal{I}_{\{t \leq T_j\}}$ is sufficient to produce a smooth implied volatility function of forward-looking interpolated caplets.



(a) Forward-looking rates with constant $\sigma_j(t)$ for all t



(b) Forward-looking rates with exponentially decreasing $\sigma_j(t)$

Figure 3.4: Approximation of the implied volatility of interpolated forward rates for different specification of $g_j(t)$.

This straightforward solution is not sufficient in the case of backward-looking forward rates however. Figure 3.3b shows in fact that the opposite occurs, that is setting $g_j(t)$ to the indicator function in fact exacerbates the mispricing brought by the interpolation method. After experimenting with different shapes for $g_j(t)$ in the accrual period, we found that changing $g_j(t)$ in the context of the interpolation method (3.3) does not lead to a smooth implied volatility function on backward-looking caplets. Actually, changing $g_j(t)$ barely resulted in a different implied volatility function. Additionally, the numerical solver mentioned seemed to suggest that such a solution may not exist as it kept iterating without finding a root or constantly returned a root of $g_j(t) = 0$ for $t \in [T_{j-1}, T_j]$. However, we have found a method that substantially reduces the mispricing in such products although this method may carry a

contradicting nature in the way its derivation differs from the one we derived in the previous section. We present this method next.

3.3.1 Smooth implied volatility of interpolated backward-looking rates

By writing $R(t,S,T)$ in (3.11), we have allowed the interpolated rate to have a stochastic nature (or to be a process), which means we could decide to set this rate to be both a forward-looking rate or a backward-looking rate by setting a different value for t, and that is aligned with the hybrid characteristic of the FMM. However, we could also specify the interpolated backward-looking rate differently, namely by ignoring its stochasticity and setting $t = T$ in (3.2). Denote the interpolated forward rate under this interpolation method by $R^*(T,S,T)$, we have

$$R^*(T,S,T) = \frac{1}{\tau_{S,T}}\left[\frac{P(T,S)}{P(T,T)} - 1\right] = \frac{1}{\tau_{S,T}}\left[\frac{P(T,S)/N(T)}{P(T,T)/N(T)} - 1\right] = \frac{1}{\tau_{S,T}}\left[\frac{NRB(T,S)}{NRB(T,T)} - 1\right] \tag{3.30}$$

Now note that, from the definition of extended bond prices,

$$\frac{P(T,S)}{B(T)} = \frac{P(S,S)e^{\int_S^T r(u)du}}{B(S)e^{\int_S^T r(u)du}} = \frac{1}{B(S)} \qquad \text{and} \qquad \frac{P(T,T)}{B(T)} = \frac{1}{B(T)}$$

We can interpolate both quantities separately. First

$$\begin{aligned}
\frac{1}{B(S)} = \frac{P(S,t)}{B(S)} &= \alpha_S \frac{P(S,T_{\eta(S)-1})}{B(S)} + (1-\alpha_S)\frac{P(S,T_{\eta(S)})}{B(S)} \\
&= \alpha_S \frac{1}{B(T_{\eta(S)-1})} + (1-\alpha_S)\frac{P(S,T_{\eta(S)})}{B(T_{\eta(S)-1})P(S,T_{\eta(S)-1})} \\
&= \alpha_S \frac{1}{B(T_{\eta(S)-1})} + (1-\alpha_S)\frac{1}{B(T_{\eta(S)-1})}\frac{1}{1+\tau_{\eta(S)}R_{\eta(S)}(S)} \\
&= \frac{1}{B(T_{\eta(S)-1})}\left(\frac{1+\alpha_S\tau_{\eta(S)}R_{\eta(S)}(S)}{1+\tau_{\eta(S)}R_{\eta(S)}(S)}\right)
\end{aligned} \tag{3.31}$$

Note that we have applied the same interpolation scheme, namely a linear combination of the NRBs following no-arbitrage arguments, with the only difference is that we perform this interpolation at time T for the extended bond-price $P(T,S)$ scaled by the numeraire $B(T)$. We can obtain a similar expression for $\frac{1}{B(T)}$, and from (3.2), we take the ratio of the two to obtain

$$\begin{aligned}
1+\tau_{S,T}R^*(T,S,T) &= \frac{\frac{1}{B(T_{\eta(S)-1})}\left(\frac{1+\alpha_S\tau_{\eta(S)}R_{\eta(S)}(S)}{1+\tau_{\eta(S)}R_{\eta(S)}(S)}\right)}{\frac{1}{B(T_{\eta(T)-1})}\left(\frac{1+\alpha_T\tau_{\eta(T)}R_{\eta(T)}(T)}{1+\tau_{\eta(T)}R_{\eta(T)}(T)}\right)} \\
&= \frac{B(T_{\eta(T)-1})}{B(T_{\eta(S)-1})}\frac{1+\tau_{\eta(T)}R_{\eta(T)}(T)}{1+\tau_{\eta(S)}R_{\eta(S)}(S)}\frac{1+\alpha_S\tau_{\eta(S)}R_{\eta(S)}(S)}{1+\alpha_T\tau_{\eta(T)}R_{\eta(T)}(T)} \\
&= \left[\prod_{j=\eta(S)}^{\eta(T)-1} 1+\tau_j R_j(T_j)\right]\frac{1+\tau_{\eta(T)}R_{\eta(T)}(T)}{1+\tau_{\eta(S)}R_{\eta(S)}(S)}\frac{1+\alpha_S\tau_{\eta(S)}R_{\eta(S)}(S)}{1+\alpha_T\tau_{\eta(T)}R_{\eta(T)}(T)}
\end{aligned} \tag{3.32}$$

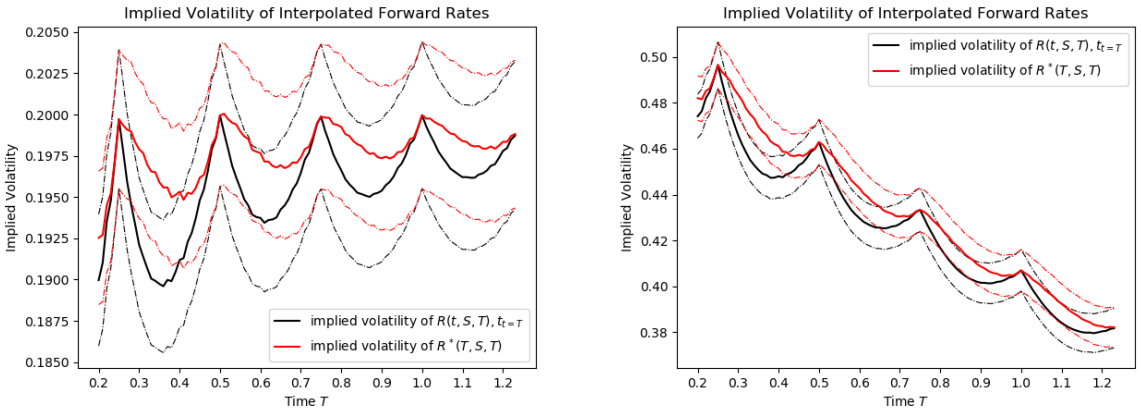
To highlight the contradiction arising from the two interpolation methods (both coming from an NRB interpolation scheme), we identify the differences between (3.11) and (3.32). They are

nearly identical, the only differences come from the time at which each term is observed. First, expand two terms out of the product operator in (3.11) and set $t = T$ to obtain

$$1 + \tau_{S,T} R(t,S,T)\Big|_{t=T} = \left[\prod_{j=\eta(S)+1}^{\eta(T)} 1 + \tau_j R_j(T)\right] \frac{1 + \alpha_S \tau_{\eta(S)} R_{\eta(S)}(T)}{1 + \alpha_T \tau_{\eta(T)} R_{\eta(T)}(T)}$$
$$= \left[\prod_{j=\eta(S)}^{\eta(T)-1} 1 + \tau_j R_j(T)\right] \frac{1 + \tau_{\eta(T)} R_{\eta(T)}(T)}{1 + \tau_{\eta(S)} R_{\eta(S)}(T)} \frac{1 + \alpha_S \tau_{\eta(S)} R_{\eta(S)}(T)}{1 + \alpha_T \tau_{\eta(T)} R_{\eta(T)}(T)}$$

where the last equality follows from the fact that backward-looking rates are constant after maturity and therefore $R_j(t) = R_j(T_j)$ for $t \geq T_j$. We can clearly see the difference arising from the time at which the two terms involving the rate $R_{\eta(S)}$ are observed. In the first interpolation method which allows stochasticity of interpolated rates, this rate is constant since $T > T_{\eta(S)}$ (T is larger than the maturity of $R_{\eta(S)}$), while in the second method this rate is observed at a point where it is still evolving and not fixed (although it is still backward-looking because $S > T_{\eta(S)-1}$). Although almost identical, these interpolation methods lead to very different results when altering $g_j(t)$ in (2.51). Furthermore, both methods are constructed following no arbitrage arguments and the fact that they differ presents an intriguing mathematical question. Note that the approximation derived in the previous can not be applied to (3.32) given the way it was constructed ($R^*(T,S,T)$ is a fixed quantity and has a zero instantaneous volatility).

We now compare the two interpolation methods in Figure 3.5 focusing only backward-looking forward rates. In both cases of the constant volatility (Figure 3.5a) and the exponentially decreasing volatility (Figure 3.5b), interpolation method (3.32) is an improvement over (3.11) and the mispricing is substantially reduced. For example, for caplets expiring soon in a constant volatility setting, the maximum pricing error (the absolute difference between the desired $\lambda_{S,T}$, equation (3.28), and the observed caplet implied volatility) shrank from 1.02% to 0.497%. In the exponential case, the curve of the implied volatility of interpolated rates looks considerably smoother, although there is still room for improvements, perhaps by varying the shape of $g_j(t)$.

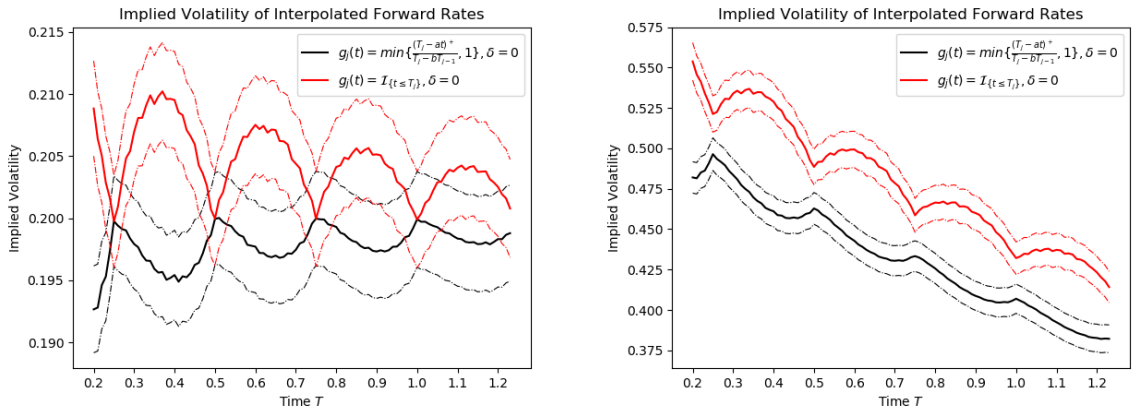


(a) Constant volatility setting and $g_j(t)$ as defined in (2.51)

(b) Exponentially decreasing volatility and $g_j(t)$ as defined in (2.51)

Figure 3.5: Implied volatility of forward rates with different interpolation methods (3.11) and (3.32), and the 95% Monte Carlo confidence interval. Rates were simulated in the FMM using 20000 simulations.

We investigate $g_j(t) = \mathcal{I}_{\{t \leq T_j\}}$ in Figure 3.6, where we now only consider interpolation (3.32). In forward-looking cases, this form for $g_j(t)$ removed the pricing error. In backward-looking cases however, this actually seems to result in overpriced caplets (the implied volatility of interpolated forward rate caplets is higher than we would expect) and a larger pricing error. This is the case for both constant and exponentially decreasing volatilities. However, it is an interesting result because at first, this suggests that there might exists a form for $g_j(t)$ such that the mispricing is completely removed. In other words, we hope to find such a function such that the implied volatility curve under this setting would lie in exactly between the underpriced and overpriced implied volatility curves, in Figures 3.6a and 3.6b. That is the motivation for introducing the constants a and b in (2.51), as we will experiment with different parameter to minimize as best as we can the pricing error observed here, assuming the existence of a solution based on our observations.



(a) Constant volatility setting and $g_j(t)$ as defined in (3.29)

(b) Exponentially decreasing volatility and $g_j(t)$ as defined in (3.29)

Figure 3.6: Implied volatility of forward rates with interpolation method (3.32) and different forms for $g_j(t)$, and their 95% Monte Carlo confidence interval. Rates were simulated using 20000 simulations.

It will be easier to find an approximate solution, or equivalently to find optimal parameters a and b, in a flat volatility environment, since we know the true value of the interpolated caplet implied volatility that would translate into an optimal interpolation. This value is the constant instantaneous volatility of the forward rates in the model. Our observations this far suggest that any interpolation behavior that occurs in the constant volatility, propagates to the exponentially decreasing volatilities. Once we have found the optimal a and b in (2.51), we can safely test them in the exponentially decreasing volatility setting and draw conclusions accordingly. We have implemented a small algorithm based on a root search to minimize the sum over all caplets of the distance between its implied volatility and the constant $\sigma_j(t)$ input in our model (constant over t and j). This task was computationally intensive as it performed a set of Monte Carlo simulations at each iteration of the solver to compute the caplet premiums and obtain the implied volatilities. Furthermore, we are minimizing a multivariate function which exacerbates the computational effort of the algorithm. Any solution we find may not be unique nor optimal. We have restricted our search to a family of functions given by (2.51), but one could explore different families, such as functions which are increasing or quadratic in t. However, the range of possibilities would make it difficult to implement a systematic approach to finding a solution.

Our solver works as follows. We consider q caplets with expiries $s_i = s_{i-1} + 0.01, i = 2, \ldots, q$ and $s_1 = 0.25$ and maturities $s_i + \delta, i = 1, \ldots, Q$, and we solve

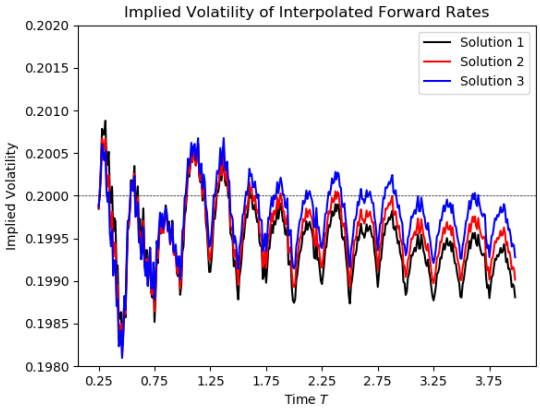
$$\min_{a,b} v(a, b, s_s, s_t) = \min_{a,b} \sum_{i=s}^{t} \left| \lambda_{s_i, s_i+\delta}(t) - \sigma^* \right| \tag{3.33}$$

where $\lambda_{s_i,s_i+\delta}(t)$ is the average implied volatility of the caplet with expiry s_i and maturity $s_i + \delta$ and σ^* is fixed to 20% . We repeated this experiment for caplets with expiries set on a different interval $[s_s, s_t]$, although. they all seem to roughly agree on the values of a and b, i.e. the optimal coefficients seem to be located around 0.8, with a being slightly larger than b. Three different solutions, along with their respective intervals on which they were solved, are shown in Table 3.1 along with the value of the objective function from (3.33). Furthermore, for each solution, we also compute the value of the objective function $v(a, b, 0.25, 4)$, to see if optimality in a sub-interval translates into optimality in a larger interval on which all pairs of coefficients are compared. Each interval has equal length although the intervals of the first and second solutions are overlapping. Solution 3 has the highest objective value on its sub-interval (worst score, since we aim to minimize 3.33), but achieves the best score on the whole interval $[0.25, 4]$.

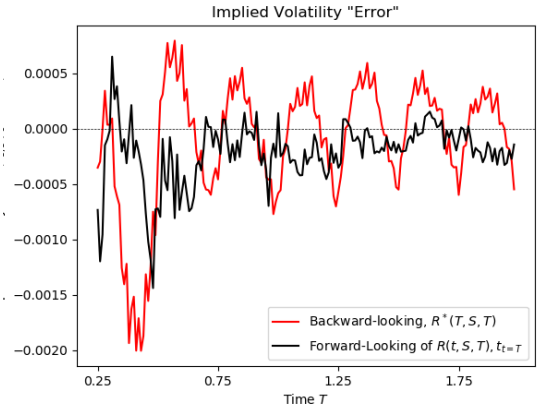
Solution	(a, b)	$[s_s, s_t]$	$v(a, b, s_s, s_t)$	$v(a, b, 0.25, 4)$
1	$(0.8406, 0.7859)$	$[0.25, 1.5]$	0.0476	0.2237
2	$(0.8557, 0.8079)$	$[1, 2.5]$	0.0475	0.1820
3	$(0.8663, 0.8269)$	$[2.5, 4]$	0.0479	0.1338^*

Table 3.1: Numerical solutions to the minimization problem (3.33)

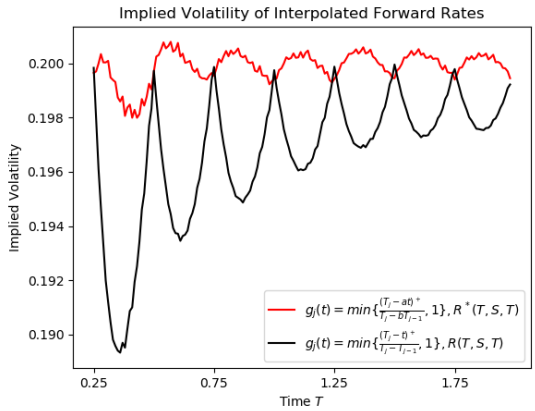
Next, we show these different solutions at work by plotting the implied volatility curves as we have done previously. Namely, we price caplets on the forward rate $R(t + \delta, t, t + \delta)$ for all $t \in [0.25, 4]$ using interpolation method (3.32) and compute the implied volatility for each of these caplets. Looking at Figure 3.7a, for caplets with expiries more distant from now, Solution 3 (given by the blue line) seems to fluctuate closer to the 20% line and its overall performance outweighs that of Solutions 1 and 2. These observations also suggest that the parameters a and b carry a calibration aspect and their choices may vary based on the number of forward rates modeled and their chosen tenors. Note also that, although the overall curve seem quite tortuous in Figure 3.7a, it is simply a consequence of the scale of the axis used to compare across the different pairs a and b (it would flatten if we would zoom out). To show that this interpolation method combined with Solution 3 is a satisfying remedy to the low volatility problem, let us compare the error produced by this method in the backward-looking case, to the error produced by (3.11) in the forward-looking case (where we establish that (3.11) was optimal when combined with $g_t = \mathcal{I}_{\{t \leq T_j\}}$). We plot the error, as defined by the difference between the caplet implied volatility and the target volatility of 20%, in Figure 3.7b. Clearly, the second solution (given by the black line) is "smoother" and reduces the mispricing more than the first solution (red line) but the difference is negligible. We are of course comparing across forward and backward-looking cases to simply provide a grasp on the magnitude of the pricing error of Solution 3 in the backward-looking case, and this example shows indeed that this error is small. Another way to support interpolation (3.32) with Solution 3 is to simply compare it against interpolation



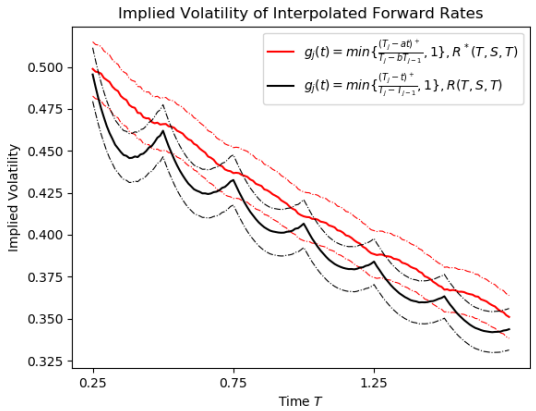
(a) Implied volatility for different solutions (a, b) given in Table (3.1)



(b) Error comparison of backward versus forward-looking ((3.32) versus (3.11))



(c) Comparison of interpolation methods (3.11) and (3.32) with constant volatility



(d) Comparison of interpolation methods (3.11) and (3.32) with exponential volatility

Figure 3.7: Implied volatility of forward rates with interpolation methods (3.11) and (3.32) and different coefficients a and b for $g_j(t)$ given by (2.51). Rates were simulated in the FMM using 8000 simulations.

method (3.11) which we saw performed poorly in backward-looking caplets. In Figure 3.7c, in a constant volatility setting, the mispricing is substantially smaller than before. As we expected, this error reduction seems to propagate nicely to exponentially decreasing volatility setting, shown in Figure 3.7d, where the alternative interpolation method (given by the red line) appear to be considerably smoother than interpolation method (3.11) (black line).

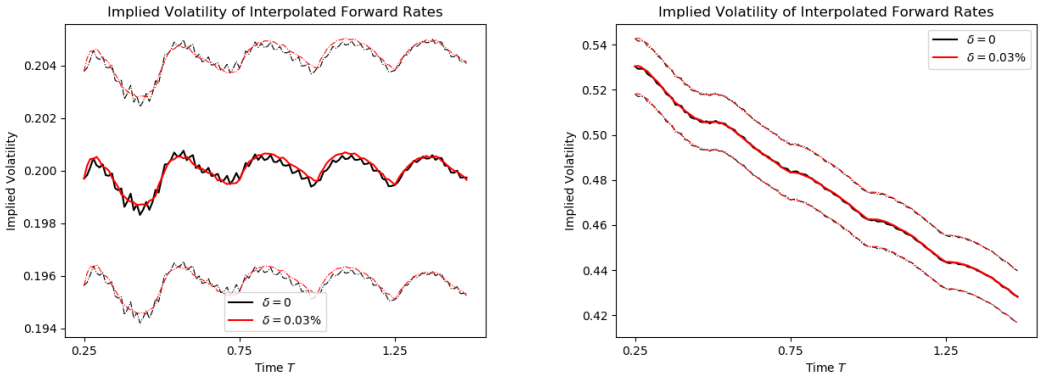
Finally, we can extend this method in the presence of a displacement in the same way we extended (3.3) to (3.14). Namely, in the presence of a displacement parameter δ, $\frac{1}{B(S)}$ becomes

$$\begin{aligned}
\frac{1}{B(S)} = \frac{P(S,t)}{B(S)} &= \widehat{\alpha}_S \frac{P(S,T_{\eta(S)-1})}{B(S)} + (1-\widehat{\alpha}_S)\frac{P(S,T_{\eta(S)})}{B(S)} + \delta\left(\widehat{\alpha}_S\tau_{\eta(S)} - \tau_{S,T_{\eta(S)}}\right)\frac{P(S,T_{\eta(S)})}{B(S)} \\
&= \widehat{\alpha}_S \frac{1}{B(T_{\eta(S)-1})} + \left(1-\widehat{\alpha}_S + \delta\left(\widehat{\alpha}_S\tau_{\eta(S)} - \tau_{S,T_{\eta(S)}}\right)\right)\frac{1}{B(T_{\eta(S)-1})}\frac{1}{1+\tau_{\eta(S)}R_{\eta(S)}(S)} \\
&= \frac{1}{B(T_{\eta(S)-1})}\left(\frac{1+\widehat{\alpha}_S\tau_{\eta(S)}R_{\eta(S)}(S) + \delta\left(\widehat{\alpha}_S\tau_{\eta(S)} - \tau_{S,T_{\eta(S)}}\right)}{1+\tau_{\eta(S)}R_{\eta(S)}(S)}\right) \\
&= \frac{1}{B(T_{\eta(S)-1})}\left(\frac{1+\widehat{\alpha}_S\tau_{\eta(S)}\widetilde{R}_{\eta(S)}(S) - \delta\tau_{S,T_{\eta(S)}}}{1+\tau_{\eta(S)}R_{\eta(S)}(S)}\right)
\end{aligned}$$

and similarly for $\frac{1}{B(T)}$ which leads to the following interpolated displaced forward rate

$$
\begin{aligned}
1+\tau_{S,T}\widetilde{R}^{*}(T,S,T) &= \frac{1/B(S)}{1/B(T)} = \frac{\frac{1}{B(T_{\eta(S)-1})}\left(\frac{1+\widehat{\alpha}_S\tau_{\eta(S)}\widetilde{R}_{\eta(S)}(S)-\delta\tau_{S,T_{\eta(S)}}}{1+\tau_{\eta(S)}R_{\eta(S)}(S)}\right)}{\frac{1}{B(T_{\eta(T)-1})}\left(\frac{1+\alpha_T\tau_{\eta(T)}\widetilde{R}_{\eta(T)}(T)-\delta\tau_{T,T_{\eta(T)}}}{1+\tau_{\eta(T)}R_{\eta(T)}(T)}\right)} \\
&= \frac{B(T_{\eta(T)-1})}{B(T_{\eta(S)-1})}\frac{1+\tau_{\eta(T)}R_{\eta(T)}(T)}{1+\tau_{\eta(S)}R_{\eta(S)}(S)}\frac{1+\alpha_S\tau_{\eta(S)}\widetilde{R}_{\eta(S)}(S)-\delta\tau_{S,T_{\eta(S)}}}{1+\alpha_T\tau_{\eta(T)}\widetilde{R}_{\eta(T)}(T)-\delta\tau_{T,T_{\eta(T)}}} \\
&= \left[\prod_{j=\eta(S)}^{\eta(T)-1}1+\tau_jR_j(T_j)\right]\frac{1+\tau_{\eta(T)}R_{\eta(T)}(T)}{1+\tau_{\eta(S)}R_{\eta(S)}(S)}\frac{1+\alpha_S\tau_{\eta(S)}\widetilde{R}_{\eta(S)}(S)-\delta\tau_{S,T_{\eta(S)}}}{1+\alpha_T\tau_{\eta(T)}\widetilde{R}_{\eta(T)}(T)-\delta\tau_{T,T_{\eta(T)}}}
\end{aligned}
\tag{3.34}
$$

which bears an obvious resemblance to the displaced interpolation method (3.17) with the differences being the same as when we compared (3.11) to (3.32). We show the implied volatility function of caplets priced under a displaced model in Figure 3.8. For the constant volatility setting in Figure 3.8a, it appears that the displaced case results in a smoother implied volatility function than the non-displaced case (the black lines displays "some saw-tooth" pattern, while the red line does not). We note that in our implementation, the non-displaced cases have been implemented with the displaced formula and setting $\delta = 0$.



(a) Interpolation methods (3.34) with constant volatility

(b) Interpolation methods (3.34) with exponential volatility

Figure 3.8: Implied volatility of forward rates with interpolation (3.34) with a and b given by Solution 3 in Table 3.1. Rates were simulated in the FMM using 16000 simulations.

Summary

This last step concludes the transition from a discrete time to a continuous time setting for the FMM. Forward-looking forward rates for any expiry-maturity pairs can be derived using the methods of this chapter, and as we showed with the implied volatility function of interpolated forward rates, there is no mispricing induced by the NRB interpolation when it is associated with a particular shape for the function $g_j(t)$. For backward-looking forward-rates, the interpolation is trickier and we have develop a method which, although appears to be model-dependent, substantially reduces the mispricing of non-linear derivatives referring those rates. In the next chapter, we will examine impacts of these interpolation methods on a practical CVA application.

4 Application to Credit Value Adjustment

An important exercise of interest rate models and in particular the one described in Section 2.2.3 is the calculation of value adjustments in the price of interest rate derivatives. Financial products are always priced in a risk-neutral setting, that is using expectations under the $\mathbb{P}^{T_j}$-measure, given no-arbitrage arguments. When pricing an interest rate swap (IRS) for example (or the FRA of section 2.1.2), it is implicitly assumed that each counterparty will meet their obligations and not fail to deliver payments to the other counterparty. In reality, however, each counterparty is subject to default and might not be able to meet such future obligations. If a certain company enters a payer IRS with a bank (the company pays the fixed rate and receives the floating rate) and faces a certain probability of default before the maturity of the contract, the bank will likely require its client to pay a premium over the fixed rate to account for that risk of default, i.e. the credit risk. A higher chance of default will result in a higher fixed rate. Similarly, a company that enters a receiver IRS (it receives the fixed rate and pays the floating) will receive a lower fixed rate than would an otherwise risk-free counterparty. In practice, the fixed rate is not computed directly but instead solved through the computation of the so-called Credit Value Adjustment (CVA) which is added to the net present value (NPV) of the swap. The fixed rate of the swap is then such that the NPV is 0.

CVA is part of broader group of value adjustments attached to the price of financial derivatives and which are jointly referred to as XVA. For example, the Debt Value Adjustment, or DVA, is the CVA but seen from the perspective of the bank and is usually lower than CVA. In this section, we discuss CVA in its simplest form and analyze it in the context of the FMM and also compare CVA across different interpolation methods introduced in the previous section. For a thorough discussion of the main XVAs, we refer the reader to Brigo et al. (2013). The work from Bogt (2018) appears to be the first to suggest the comparison of CVA of swaps across interpolation methods. Our focus extends to comparing CVA between forward-looking and backward-looking swaps. As we will explain, we expect backward-looking swaps to yield a higher CVA than forward-looking swaps.

4.1 Theory of Counterparty Credit Risk

As is clear from the discussion above, chance of default of the counterparties play a major role in the computation of CVA. Reasonably enough, both counterparties in a swap will carry some credit risk. Even sovereigns, which are generally considered risk-free, can default on their obligations as we witnessed during the 2011 Eurozone crisis. A CVA set up where both counterparties are defaultable is called bilateral. Here, we focus on unilateral CVA, where one counterparty, typically the bank, is considered default free (the counterparties in a swap are usually a bank and a corporate, but can also very well be between two banks). This assumption simplifies the computation of CVA and provides the basis on which to construct a bilateral model. Furthermore, interest rate swaps often come with a Credit Support Annex (CSA) that obliges the counterparties to post collateral which mitigates their credit risk since it provides the collateral holder with an "insurance" against default of the other counterparty. Therefore,

posting collateral reduces CVA but does not completely remove it due to factors such as the type and quality of the collateral (cash, bonds, currency, ...) or the day lag between the computation of CVA and the exchange of collateral, etc. (see Darbyshire (2017) for more details). In this thesis, we assume that there is no CSA agreement in place. We briefly provide a few results on Unilateral CVA from the work of Brigo et al. (2013), which we use to compute CVA. We refer to their book for more details.

Consider the probability space $(\Omega, \mathcal{F}, \mathbb{P})$, introduced in Section 2.1.1. Recall that $\mathcal{F}$, the Borel field of measurable events, has a filtration $\mathcal{F}_t$ which represents default-free market observable information up to time t. We must expand the probability space with a $\sigma-$algebra and filtration which also includes default events. Let $\mathcal{D}_t$ be the filtration generated by the default events in the market, and let $\mathcal{G}_t := \mathcal{F}_t \wedge \mathcal{D}_t$ be the filtration of the $\sigma-$algebra $\mathcal{G}$ of measurable default-free and default events. Our probability space becomes $(\Omega, \mathcal{G}, \mathbb{P})$.

We denote the time of the default event of the counterparty by τ (we price from the perspective of the bank or investor and compute CVA for the counterparty) and assume its cumulative distribution to be given by

$$\mathbb{Q}\left(\tau \leq T\right) = 1 - \exp\left(-\int_0^T \beta_s ds\right) \tag{4.1}$$

Typically, the parameter β_s, called the intensity, is obtained from the Credit Default Swap (CDS) quotes of the counterparty. This may not always be possible or appropriate if, for example, there is no CDS traded on that counterparty or if they are illiquid. Alternatively, one could simply assume a constant intensity. We briefly explain how to obtain β_s for a given set of CDS quotes in Appendix B.

Let $\Pi\left(t, T\right)$ be the net cash flows of the product under consideration at time t, where T is the maturity or the last payment date, and which is traded with a default-free counterparty. CVA does not apply exclusively to swaps, but in the case of a payer IRS with notional N, we have

$$\Pi(t, \widetilde{T}_\alpha)^{\text{IRS}} = \sum_{i=\widetilde{\eta}(t)}^{\alpha} N\left(\widetilde{T}_i - \widetilde{T}_{i-1}\right) P(t, \widetilde{T}_i)\left(L(t, \widetilde{T}_i - \widetilde{T}_{i-1}) - K\right) \tag{4.2}$$

where $\widetilde{T}_0 < \cdots < \widetilde{T}_\alpha$ are the tenor dates of the IRS and $\widetilde{\eta}(t) := \{j : \widetilde{T}_{j-1} < j < \widetilde{T}_j\}$.

If we denote by $\bar{\Pi}\left(t, T\right)$ the same payoff but where the counterparty now has a positive probability of default, we have

$$\begin{aligned}\bar{\Pi}\left(t, T\right) =& \mathcal{I}_{\{\tau > T\}} \Pi\left(t, T\right) + \mathcal{I}_{\{t \leq \tau \leq T\}} \Bigg[\Pi\left(t, \tau\right) \\ &+ P(t, \tau)\left(R\left(\mathbb{E}\left[\Pi\left(\tau, T\right) \,\middle|\, \mathcal{G}_\tau\right]\right)^+ - \left(-\mathbb{E}\left[\Pi\left(\tau, T\right) \,\middle|\, \mathcal{G}_\tau\right]\right)^+\right)\Bigg]\end{aligned} \tag{4.3}$$

where R is the recovery of the counterparty, that is the fraction of the original cashflows that is to be recovered in case of default. It is usually deterministic and assumed to be around 40%. The quantity $1 - R$ is, intuitively enough, called the loss given default (LGD).

Without default before maturity, $\mathcal{I}_{\{\tau > T\}} = 1$ and the cashflows with a defaultable

counterparty is the same as with a non-defaultable counterparty $\bar{\Pi}(t,T) = \Pi(t,T)$. However, if there is an early default, the payments occur normally till $\tau \leq T$, which gives the term

$$\Pi(t,\tau)$$

and if the residual value of the product is positive, the investor receives what can be recovered

$$R\left(\mathbb{E}\left[\Pi(\tau,T) \;\middle|\; \mathcal{G}_\tau\right]\right)^+$$

If it is negative, the investor pays it entirely (since it is default-free)

$$-\left(-\mathbb{E}\left[\Pi(\tau,T) \;\middle|\; \mathcal{G}_\tau\right]\right)^+$$

From equation (4.3), Brigo et al. (2013) derive the following *unilateral* counterparty risk pricing formula

$$\begin{aligned}
\mathbb{E}\left[\bar{\Pi}(t,T) \;\middle|\; \mathcal{G}_t\right] &:= \mathbb{E}_t\left[\bar{\Pi}(t,T)\right] \\
&= \mathbb{E}_t\left[\Pi(t,T)\right] - \mathbb{E}_t\Big[\mathcal{I}_{\{t<\tau\leq T\}}(1-R)P(t,\tau)\left(\mathbb{E}_\tau\left[\Pi(\tau,T)\right]\right)^+\Big] \\
&= \mathbb{E}_t\left[\Pi(t,T)\right] - \textsc{Ucva}(t,T) \qquad (4.4) \\
\textit{with} \qquad \textsc{Ucva}(t,T) &:= \mathbb{E}_t\Big[\mathcal{I}_{\{t<\tau\leq T\}}(1-R)P(t,\tau)\left(\mathbb{E}_\tau\left[\Pi(\tau,T)\right]\right)^+\Big]
\end{aligned}$$

Since $P(t,\tau) = \mathbb{E}\left[e^{-\int_t^\tau r(u)du} \;\middle|\; \mathcal{G}_t\right]$, we can write

$$\begin{aligned}
\textsc{Ucva}(t,T) &= \mathbb{E}\left[\mathcal{I}_{\{t<\tau\leq T\}}(1-R)\,\mathbb{E}\left[e^{-\int_t^\tau r(u)du} \;\middle|\; \mathcal{G}_t\right]\left(\mathbb{E}_\tau\left[\Pi(\tau,T)\right]\right)^+ \;\middle|\; \mathcal{G}_t\right] \\
&= \mathbb{E}_t\left[\mathcal{I}_{\{t<\tau\leq T\}}(1-R)\frac{B(t)}{B(\tau)}\left(\mathbb{E}_\tau\left[\Pi(\tau,T)\right]\right)^+\right]
\end{aligned}$$

Furthermore, by defining grid points $t_0 < \cdots < t_\beta$ and assuming that the default event is postponed to the next grid point, Brigo et al. (2013) show that, at time $t = 0$, equation (4.4) can be approximated as

$$\begin{aligned}
\mathbb{E}\left[\bar{\Pi}(0,t_\beta)\right] &\approx \mathbb{E}\left[\Pi(0,t_\beta)\right] - (1-R)\sum_{j=1}^{\beta}\mathbb{E}\left[\mathcal{I}_{\{t_{j-1}<\tau\leq t_j\}}\frac{B(t)}{B(t_j)}\left(\mathbb{E}_{t_j}\left[\Pi(t_j,t_\beta)\right]\right)^+\right] \\
&= \mathbb{E}\left[\Pi(0,t_\beta)\right] - (1-R)\sum_{j=1}^{\beta}\mathbb{Q}\left(t_{j-1}<\tau\leq t_j\right)\mathbb{E}_{t_j}\left[\frac{B(0)}{B(t_j)}\left(\mathbb{E}_{t_j}\left[\Pi(t_j,t_\beta)\right]\right)^+\right] \qquad (4.5)
\end{aligned}$$

where we assumed independence between the default event τ and $\Pi(t,T)$ (the expectation operator factors and $\mathbb{E}\left[\mathcal{I}_{\{t<\tau\leq T\}}\right] = \mathbb{Q}\left(T_{j-1}<\tau\leq T_j\right)$). We see that the price of the derivative with a defaultable counterparty, is the risk-neutral price of the same derivative minus an average of options on the residual value of the asset weighted by default probabilities.

This additional term which has an optionality characteristic, is the Unilateral CVA, i.e.

$$\text{UCVA}(0, t_\beta) = (1-R)\sum_{j=1}^{\beta} \mathbb{Q}\left(t_{j-1} < \tau \leq t_j\right) \mathbb{E}_{t_j}\left[\frac{B(0)}{B(t_j)}\left(\mathbb{E}_{t_j}\left[\Pi\left(t_j, t_\beta\right)\right]\right)^+\right] \tag{4.6}$$

4.2 CVA results for vanilla payer IRS

We can apply the CVA formula (4.6) to an IRS and analyze how CVA differs across forward and backward-looking swaps and interpolation methods.

Consider the payer IRS given in equation (4.2) and let us choose the grid points such that they contain the tenors dates of the IRS $\widetilde{T}_0 < \cdots < \widetilde{T}_\alpha$. For this swap, the unilateral CVA is given by

$$\begin{aligned}
\text{UCVA}(0, \widetilde{T}_\alpha) &= (1-R)\sum_{j=1}^{\beta} \mathbb{Q}\left(t_{j-1} < \tau \leq t_j\right) \mathbb{E}_{t_j}\left[\frac{B(0)}{B(t_j)}\left(\mathbb{E}_{t_j}\left[\Pi(t_j, \widetilde{T}_\alpha)^{\text{IRS}}\right]\right)^+\right] \\
&= (1-R)\sum_{j=1}^{\beta} q_j \mathbb{E}_{t_j}\left[\frac{B(0)}{B(t_j)}\left(N\sum_{i=\widetilde{\eta}(t_j)}^{\alpha} P(t_j, \widetilde{T}_i)(\widetilde{T}_i - \widetilde{T}_{i-1})\left(R(t_j, \widetilde{T}_{i-1}, \widetilde{T}_i) - K\right)\right)^+\right] \\
&= (1-R)N\sum_{j=1}^{\beta} q_j \mathbb{E}_{t_j}\left[\left(\sum_{i=\widetilde{\eta}(t_j)}^{\alpha} \frac{P(t_j, \widetilde{T}_i)}{B(t_j)}(\widetilde{T}_i - \widetilde{T}_{i-1})\left(R(t_j, \widetilde{T}_{i-1}, \widetilde{T}_i) - K\right)\right)^+\right]
\end{aligned} \tag{4.7}$$

where $q_j := \mathbb{Q}\left(t_{j-1} < \tau \leq t_j\right) = \mathbb{Q}\left(\tau > t_j\right) - \mathbb{Q}\left(\tau > t_{j-1}\right)$. In the equation above, we recognise the NRB obtainable with one of the interpolation methods of the previous section, which also provide us with the interpolated forward rates. Therefore, the Unilateral CVA is easily computed by replacing $\mathbb{E}_{t_j}$ by a Monte Carlo estimate in equation (4.7). In Appendix B, we explain how to obtain default probabilities and in this exercise, we use the ones presented in Table B.1.

We consider three IRSs running 4 years each and which differ by their accrual period. We compute the CVA for each swap as a percentage of the notional. Results for forward-looking and backward-looking swaps are shown in Table 4.1. The modeled forward rates are quarterly rates (they have a tenor of 0.25 years). The first swap has quarterly payments, the second swap has semi-annual payments and the last one has five payments per year (the accrual is 0.2 years). This requires interpolation for the annual swap and the swap with the 0.2 years accrual. The reason for these choices is to see how interpolation performs in two distinctive cases, those are when the interpolated rate tenor is larger than the modeled tenor, and when it is smaller. In the forward-looking case, we simulate forward rates under three different set-ups, namely in the FMM with $g_j(t)$ given by (2.51), the FMM with $g(t)$ in (3.29) and the LMM. In each case, we use a displacement of 0.03% and 2000 Monte Carlo paths. The interpolated forward rates are constructed with the NRB interpolation method (3.14). In the backward-looking case, forward rates are simulated in the FMM. The same $g_j(t)$ will be used with interpolation method (3.14) but we also analyze CVA when interpolation method (3.32) is used (in combination with the

optimal coefficient a and b given by solution 3 in Table 3.1). In any case, the volatility process of each forward rate $\sigma_i(t), i = 1, \ldots, 16$, is constant and equal to 20% (Since the modeled tenor is 0.25 years, there are 16 tenor dates: $T_0 = 0, T_1 = 0.25, T_2 = 0.50, \ldots, T_{16} = 4$). We simulate on the grid $t_0, \ldots, t_\beta$ given by $0, 0.05, 0.1, \ldots, 4.95, 5$. The probabilities of default are calculated as described in the previous section, with the intensities from Table B.1. Finally, the recovery rate is again set to 40%.

	Forward-Looking			***Backward-Looking***	
Interpolation method	(3.14)	(3.14)	(3.14)	(3.14)	(3.34)
Specification of $g_j(t)$	LMM	$\mathcal{I}_{\{t \leq T_j\}}$	$\frac{(T_j - t)^+}{T_j - T_{j-1}}$	$\mathcal{I}_{\{t \leq T_j\}}$	$(\frac{(T_j - at)^+}{T_j - bT_{j-1}}$
Tenor of Forward Rates	**Calculated CVA**				
0.25 years	0.0389	0.0389	0.0389	0.0392	0.0391
Monte Carlo Upper Bound	0.0408	0.0408	0.0408	0.0412	0.0411
Monte Carlo Lower Bound	0.0370	0.0370	0.0370	0.0373	0.0372
0.20 years	0.0362	0.0366	0.0365	0.0367	0.0365
Monte Carlo Upper Bound	0.0381	0.0384	0.0383	0.0385	0.0384
Monte Carlo Lower Bound	0.0344	0.0347	0.0346	0.0348	0.0347
0.50 years	0.0400	0.0401	0.0401	0.0414	0.0412
Monte Carlo Upper Bound	0.0421	0.0422	0.0422	0.0435	0.0433
Monte Carlo Lower Bound	0.0380	0.0381	0.0381	0.0394	0.0391

Table 4.1: Unilateral CVA of forward and backward-looking IRS with different interpolation schemes and accruals. The values are presented as a percentage of the notional. The Monte Carlo 95% Window is also given.

Firstly, when comparing forward against backward-looking swaps, the results in Table 4.1 show that, as we might expect, backward-looking products generally produce a higher CVA. For a given interpolation and $g_j(t)$, backward-looking swaps yield a slightly higher CVA. The magnitude of this increase becomes larger as the tenor of the forward rate increase, but this could also be due to the effect of interpolation. In fact, the same CVA computations on the swap with a tenor of 0.5 years, when the modeled forwards have a tenor of 0.5 years, yield a CVA of 0.0411% in the forward-looking case and 0.0423% in the backward-looking case, suggesting that all interpolation methods here leads to a noteworthy underestimation of CVA. Secondly, we cannot conclude from these results that the *choice* of interpolation method significantly affects CVA results, within both types of swaps, since they are very comparable.

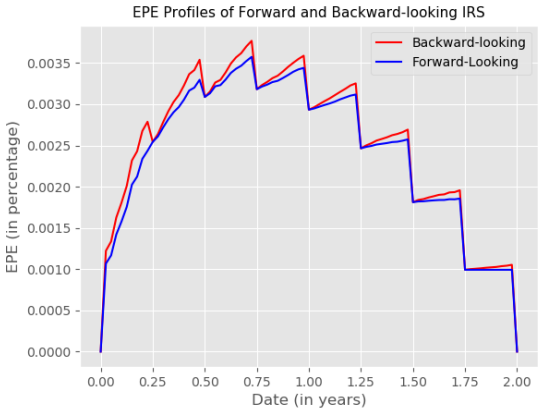
The reason why forward-looking swaps generally have a lower CVA than backward-looking swaps is due to the continued evolution of forward rates beyond the expiry. An important quantity involved in the CVA computation is the Expected Positive Exposure (EPE), which is a function of time and is given by

$$\mathrm{EPE}(t) = \mathbb{E}_t \left[\left(\sum_{i=\widetilde{\eta}(t)}^{\alpha} \textsc{Nrb}(t, \widetilde{T}_i)(\widetilde{T}_i - \widetilde{T}_{i-1}) \left(R(t, \widetilde{T}_{i-1}, \widetilde{T}_i) - K \right) \right)^+ \right] \tag{4.8}$$

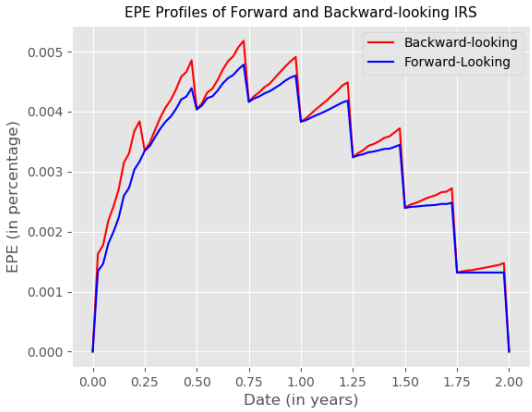
At a given time point, we consider all paths of forward rates and compute the discounted remaining value of the swap payments. We average only the paths where the residual value of

this swap is positive and this leads to the EPE. (The Expected Negative Exposure is defined similarly, taking negative residual values instead). At time 0, the EPE is 0 since the fixed rate of the swap is such that the swap has value 0. At $t \geq \widetilde{T}_\alpha$, the EPE is again 0 since there is no payment left in the swap.

We look at EPE profiles of forward and backward-looking swaps. For IRS with the same tenor structure as the model, the results are shown in Figure 4.8 (Note that interpolation is still used to obtain the NRBs, but not the forward rates). We simulated forward rates under the two volatility settings we have considered so far, namely the exponentially decreasing volatility and the constant volatility (set to 30%). As expected, the backward-looking swaps have a higher EPE than the forward-looking swaps. This is likely resulting from the fact $R(t, \widetilde{T}_{i-1}, \widetilde{T}_i)$ continues to evolve beyond $\widetilde{\eta}(t)$ in the backward-looking case, as we just mentioned. Since CVA is a weighted average of EPEs, higher EPEs translate into a higher CVA. Certainly this continued volatility will also generate paths where the residual value will be smaller. However, the max function and the optionality characteristic of the EPE will result in more cases where the payoff is higher. A plot of some paths, which are then averaged to compute the EPE, is shown in Appendix C, namely Figures C.4a and C.4b. It appears that forward-looking positive exposures generally lie below the backward-looking positive exposures. Note that the staircase shape of EPEs results from its definition. At every time point $t = \widetilde{T}_i$, $i = 1, \ldots, \alpha$, $\widetilde{\eta}(t) = i + 1$, which implies that a payment drops out of formula (4.8) and is no longer included in the CVA and hence the sudden drop in the EPE. Indeed, we are no longer exposed to this payment as it has been realised.



(a) Constant volatility setting and $g_j(t)$ as defined in (3.29)

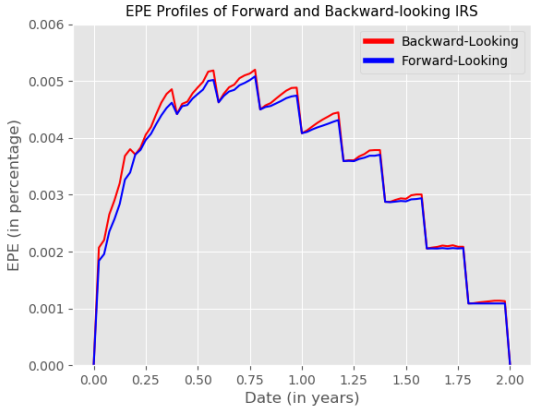


(b) Exponentially decreasing volatility and $g_j(t)$ as defined in (3.29)

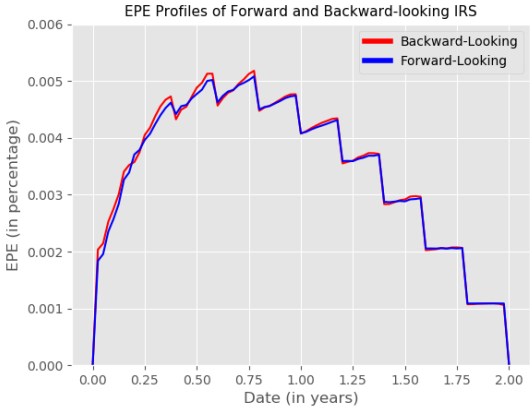
Figure 4.1: EPE and MC paths of Positive Exposures of forward and backward-looking IRS with tenor of 0.25 years. Rates were simulated in the FMM using 2000 simulations.

This behavior (higher CVA in backward-looking swaps) should also appear in IRS priced with interpolated forward rates. Therefore, we now plot the EPE profiles for the interpolated IRS from Table 4.1 (given the higher CVA of backward-looking interpolated IRS, there is no reason to expect this behavior to change since CVA is a weighted average of EPEs). We shorten the IRS to be running over 2 years. We focus our attention to forward rates simulated with a constant volatility of 30%. For forward-looking swaps, we construct interpolated forward

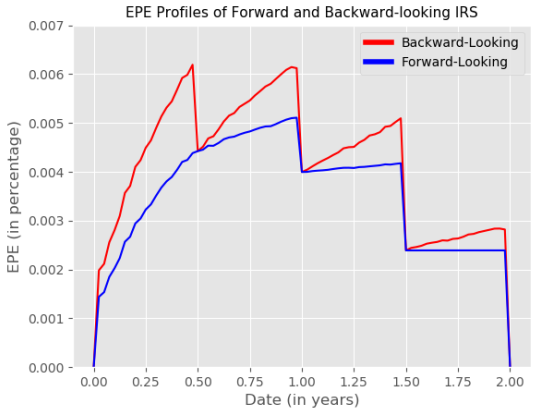
rates with interpolation method (3.14) and $g_j(t)$ is set to (3.29). Backward-looking interpolated swaps are constructed on the one hand with the same scheme, and on the other hand with interpolaton method (3.34) and $g_j(t)$ given by (2.51) and a, b as previously set. The resulting EPEs are shown in Figure 4.2. The second interpolation method of backward-looking rates is applied to EPE shown in Figures 4.2b and 4.2d. EPEs of backward-looking swaps, regardless of the interpolation method, are generally higher than forward-looking swaps. When the backward-looking rates are interpolated according the method described in Section 3.3.1 however, the EPE is almost identical to the EPE of the forward-looking swap. For backward-looking forward rates with a tenor shorter than that of the model, this may lead to an underestimated CVA as a result. For the larger tenor however, even though this interpolation method leads to a slightly lower EPE profile (Figures 4.2c and 4.2d), the backward-looking IRS has a larger EPE than the forward-looking IRS, as it should be.



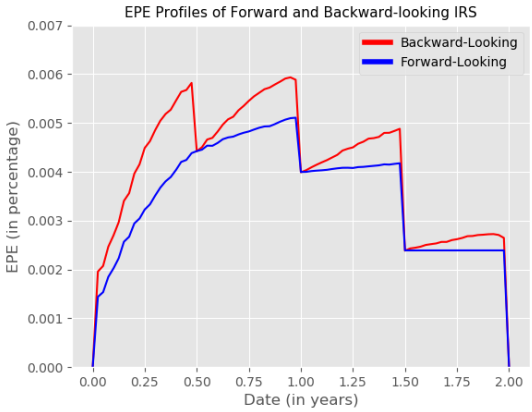
(a) IRS with tenor of 0.2 years. Backward rates are interpolated by (3.14) and $g_j(t)$ given by (3.29)



(b) IRS with tenor of 0.2 years. Backward rates are interpolated by (3.34) and $g_j(t)$ given by (3.29)



(c) IRS with tenor of 0.5 years. Backward rates are interpolated by (3.14) and $g_j(t)$ given by (3.29)



(d) IRS with tenor of 0.5 years. Backward rates are interpolated by (3.34) and $g_j(t)$ given by (3.29)

Figure 4.2: EPE of forward and backward-looking interpolated IRS with different tenors and interpolation of backward-looking rates. Rates were simulated in the FMM using 2000 simulations.

5 Conclusion

As banks and other financial institutions must adapt their internal systems to accommodate for the discontinuation of IBOR rates and the introduction of new overnight RFRs such as €STR or SOFR, it is of paramount importance that they carefully select the tools to accurately price the products contingent on these rates and manage the inherent risk appropriately. Many products will begin to reference these new rates, as is already the case for interest rate swaps and will propagate to caps, floors and swaptions. The FMM developed by Lyashenko and Mercurio (2019) and presented in this thesis is one such tool that can assist institutions in their transition. It conveniently extends the existing LMM to allow for the diffusion of such overnight RFRs. Given that LMM is often the model of choice in many operations such as the computation of XVA, the FMM can be regarded as natural replacement of the LMM to perform these task in products based on overnight RFRs. For example, our analysis on CVA show what initially resulted from intuition, namely that CVA of backward-looking products appear to be higher than forward-looking equivalents. This observation generally does not seem to be strongly impacted by interpolation methods, although the interpolation of backward-looking rates derived in section 3.3.1 may result in an underestimated CVA.

The FMM as an extension of the LMM was motivated by a mathematical observation that goes as follows. On a given expiry-maturity pair, we can define two distinct forward rates (convention tells us that for most financial product, the payment date is the maturity). The first forward rate is forward-looking in the sense that the payment of a specific product based on this forward rate is fixed at the expiry. The second is backward-looking because its value is derived from the daily compounding of overnight rates published between the expiry and the maturity, and one has to wait until maturity to know the applicable payment value. When setting those quantities to be the value of the fixed rate in two FRAs (forward and backward-looking respectively) such that it has value 0 at the time of issuance, rules of probability measures and expectations indicate that these two rates are identical before the expiry, as we showed. Afterwards, the forward-looking rate "dies" but the backward-looking continues to exist until maturity after which it becomes constant. Backward-looking rates are reconciled mathematically with the traditional discretely compounded forward rates thanks to the so-called extended bond price. The FMM is then simply a model which incorporates the LMM up to the expiry, but allows the forward rates to evolve further till maturity. Any pricing application based on this model can then select the relevant forward rate (forward or backward-looking) from the simulations by observing the rate at the relevant date.

As we noted, the expiry-maturity pairs of the forward rates modelled will likely differs from the dates of the products we need to price, simply because there might be too many products to include all expiry-maturity pairs in the model without significantly impacting computational performance. As we explained, a solution is to implement an interpolation scheme to transform a discrete-time model to a continuous time model, in the maturity dimension. However, interpolation of backward-looking rates has not yet been studied since there was never a need to do so and we decided to carry extra attention to the treatment of such rates, while also considering interpolation of forward-looking forward rates in the FMM.

This thesis showed that an interpolation method following from no-arbitrage arguments and commonly used in the LMM (we called it the NRB interpolation), applies equally to the FMM and hence, to backward-looking rates. We have also incorporated the presence of a displacement with important results previously derived by Bogt (2018) (the displacement-aware interpolation would not have been implementable without his results).

This interpolation method is however known to lead to interpolated forward rates which have a lower instantaneous volatility compared to standard forward rates. This translates into a mispricing on products that refer interpolated forward rates. An alternative interpolation method which solved this issue was derived in the LMM and the displaced LMM by Werpachowski (2010) and Bogt (2018) respectively. Their solution is somewhat already incorporated in the FMM because of the continued diffusion in the accrual period. Therefore, we achieved a smooth implied volatility of interpolated caplets in forward-looking products by modifying the decaying behaviour of forward rate volatilities such that it mimics the solution of Werpachowski (2010).

Unfortunately, for backward-looking products, this simple solution is not effective. Indeed, the mispricing was exacerbated and changing the behaviour of forward rates in the accrual period merely impacted the mispricing. This was confirmed by our derivation of a very accurate approximation of the instantaneous volatility of forward rates for rates constructed under the NRB interpolation, which suggested that a smooth implied volatility function was not possible under this interpolation. The definition of extended-bond prices led us to explore the NRB interpolation applied to backward-looking bonds rather than treating these bonds as stochastic processes and subsequently observing the forward rates constructed from those bonds at maturity. We showed that this derivation remains consistent with no-arbitrage arguments as well as extended bond prices. This approach along with an optimization routine to determine the volatility structure in the accrual period of forward rates, significantly reduced the mispricing in interpolated caplets. Finally, our results in a simplistic CVA application are in line with our expectations. Backward-looking swaps yield larger CVA than their forward-looking counterparts, whether the payment dates correspond to the tenor structure or whether they are not and interpolation is required.

The disadvantage of this approach however, is that there is no way to determine the analytical instantaneous volatility of interpolated forward rates since we treat those forward rates as fixed and no longer stochastic. Additionally, it may seem that the numerical solution arising from this approach is somewhat bound to the chosen model and tenor, which would require several set of simulations when pricing different types of products (forward or backward-looking). This is inconvenient and undermines the purpose of interpolation. A more robust alternative is therefore needed but in the absence of such a solution, we proved that our interpolation method eliminates most of the mispricing. An alternative idea follows from mirroring the solution of Werpachowski (2010) to backward-looking forward rates. To be specific, Werpachowski extended the simulation of LIBOR rates to include the accrual period. Perhaps the backward-looking mispricing issue can be solved analogously, that is by extending the simulation of backward-looking rates beyond maturity (solely for the purpose of interpolation, naturally). We have not explored this option and further research is needed to prove or refute this claim, which merely follows from intuition.

References

Andersen, L. and Piterbarg, V. *Interest Rate Modeling. Volume I: Foundations and Vanilla Models.* Atlantic Financial Press, 2010.

Baxter, M. and Rennie, A. *Financial Calculus: An introduction to derivative pricing.* Cambridge University Press, 1996.

Beveridge, C. and Joshi, M. Interpolation schemes in the displaced-diffusion LIBOR market model. *SIAM Journal on Financial Mathematics*, 3:593–604, 2012.

Black, F. The pricing of commodity contracts. *Journal of Financial Economics*, 3:167–179, 1976.

Bogt, T. *Arbitrage-Free Interpolation in the LIBOR Market Model.* Master Thesis. University of Amsterdam, 2018.

Brace, A., Gątarek, D., and Rutkowski, M. The market model of interest rate dynamics. *Mathematical Finance*, 7:127–154, 1997.

Brigo, D. and Mercurio, F. *Interest Rate Models: Theory and Practice.* Springer Verlag, 2nd edition, 2006.

Brigo, D., Morini, M., and Pallavicini, A. *Counterparty: credit risk, collateral and funding.* John Willey and Sons, 2013.

Darbyshire, J. *Pricing and Trading Interest Rate Derivatives: A Practical Guide to Swaps.* Aitch and Dee Limited, 2017.

Ding, D. The martingale approach for credit-risky exchange option pricing. *Applied Mathematical Sciences*, 3:129–140, 2009.

European Central Bank. Report by the working group on euro risk-free rate: On the impact of the transition from EONIA to ESTR on cash and derivatives products. 2019.

Glasserman, P. and Zhao, X. Arbitrage-free discretization of lognormal forward libor and swap rate models. *Finance and Stochastics*, pages 35–68, 2000.

Heath, D., Jarrow, R., and Morton, A. Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica*, 60:77 – 105, 1992.

Henrard, M. A quant perspective on ibor fallback consultation results. *SSRN*, 2019.

Jamishidian, F. Libor and swap market models and measures. *Finance and Stochastics*, 1: 293–330, 1997.

Jäckel, P. *Monte Carlo Methods in Finance.* Wiley Finance, 2002.

Jäckel, P. By implication. *Willmot magazine*, pages 60–66, 2006.

Jäckel, P. and Rebonato, R. Linking caplet and swaption volatilities in a bgm/j framework: Approximate solutions and empirical evidence. *Journal of Computation Finance*, 6:41–59, 2003.

Koopman, S., Mallee, M., and van der Wel, M. Analyzing the term structure of interest rates using the dynamic nelson-siegel model with time-varying parameters. *Journal of Business and Economic Statistics*, 28:329–343, 2010.

Lyashenko, A. and Mercurio, F. Looking forward to backward-looking rates. *SSRN*, 2019.

Mercurio, F. A simple multi-curve model for pricing sofr futures and other derivatives. *SSRN*, 2018.

Rebonato, R. *Volatility and Correlations. The Perfect Hedger and the Fox.* John Wiley and Sons, second edition, 2004.

Schlögl, E. Arbitrage-free interpolation in models of markets observable interest rates. *Advances in Finance and Stochastics*, 2002.

Shreve, S. E. *Stochastic Calculus for Finance II. Continuous Time Models.* Springer Finance, 2008.

Skorohod, A. *Integration in Hilbert Space.* Springer-Verlag Berlin Heidelberg, 1974. doi: 10.1007/978-3-642-65632-3.

Tchuindjio, L. An extended heath-jarrow-morton risk neutral drift. *Applied Mathematics Letters*, 22:396–400, 2009.

Werpachowski, R. Arbitrage-free rate interpolation scheme for libor market model with smooth volatility term structure. *SSRN*, 2010.

Zeitsch, P. The economics of XVA trading. *Journal of Mathematical Finance*, 7:239–274, 2017.

A Proofs

This section contains the proof of some results derived in the main text. We have restated the results for convenience.

Theorem 2.2.2. *Under the $\widetilde{\mathbb{P}}$-measure, the bond price dynamics are given by*

$$dP(t,T) = P(t,T)\Big(r(t)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\Big)$$

The solution of this stochastic differential equation is

$$\begin{aligned} P(t,T) &= P(0,T)\exp\left(\int_0^t r(u)du - \int_0^t \frac{1}{2}\mathbf{\Sigma}^*(u,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(u,T)du - \int_0^t \mathbf{\Sigma}^*(u,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(u)\right) \\ &= P(0,T)B(t)\exp\left(-\int_0^t \frac{1}{2}\mathbf{\Sigma}^*(u,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(u,T)du - \int_0^t \mathbf{\Sigma}^*(u,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(u)\right) \end{aligned}$$

Proof. We apply Itô-Doeblin with $f(t,X_t) = \exp(X_t)$ and

$$X_t = \int_0^t r(u)du - \int_0^t \frac{1}{2}\mathbf{\Sigma}^*(u,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(u,T)du - \int_0^t \mathbf{\Sigma}^*(u,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(u)$$

The differential of X_t is given by

$$dX_t = r(t)dt - \frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)$$

Since $P(t,T) = P(0,T)\exp(X_t)$, Itô-Doeblin gives

$$\begin{aligned} dP(t,T) &= P(0,T)\exp(X_t)dX_t + \frac{1}{2}P(0,T)\exp(X_t)dX_t dX_t \\ &= P(t,T)\Bigg(r(t)dt - \frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t) \\ &\qquad + \frac{1}{2}\left(-\mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\right)\left(-\mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\right)^\top\Bigg) \\ &= P(t,T)\Bigg(r(t)dt - \frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t) \\ &\qquad + \frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\left(d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\right)^\top \mathbf{\Sigma}^*(t,T)\Bigg) \\ &= P(t,T)\Bigg(r(t)dt - \frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\, \mathbf{\Sigma}^*(t,T)dt \\ &\qquad + \frac{1}{2}\mathbf{\Sigma}^*(t,T)^\top \boldsymbol{\rho}\mathbf{\Sigma}^*(t,T)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\Bigg) \\ &= P(t,T)\Big(r(t)dt - \mathbf{\Sigma}^*(t,T)^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\Big) \end{aligned}$$

□

Proposition 2.2.3. *The dynamics of $\widetilde{R}_j(t)$ are given by*

$$d\widetilde{R}_j(t) = \left(R_j(t) + \frac{1}{\tau_j}\right)\left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right]^\top \left\{\boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j)dt + d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\right\}$$

Proof. Using Itô-Doeblin formula for Itô processes with

$$\begin{aligned} X(t) &= \int_0^t \frac{1}{2}\left[\boldsymbol{\Sigma}^*(u,T_j)^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(u,T_j) - \boldsymbol{\Sigma}^*(u,T_{j-1})^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(u,T_{j-1})\right] du \\ &\quad + \int_0^t \left[\boldsymbol{\Sigma}^*(u,T_j) - \boldsymbol{\Sigma}^*(u,T_{j-1})\right]^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(u) \\ f(x) &= f_x(x) = f_{xx}(x) = e^x \end{aligned}$$

and

$$\begin{aligned} dX_t &= \frac{1}{2}\left[\boldsymbol{\Sigma}^*(t,T_j)^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_{j-1})\right] dt + \left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right]^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t) \\ dX_t dX_t &= \left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right]^\top \boldsymbol{\rho}\left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right] dt \\ &= \left[\boldsymbol{\Sigma}^*(t,T_j)^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j) - 2\ \boldsymbol{\Sigma}^*(t,T_{j-1})^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j) + \boldsymbol{\Sigma}^*(t,T_{j-1})^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_{j-1})\right] dt \end{aligned}$$

the dynamics for $R_j(t)$ is

$$\begin{aligned} d\widetilde{R}_j(t) &= d\left(\widetilde{R}_j(t) + \frac{1}{\tau_j}\right) \\ &= \left(R_j(t) + \frac{1}{\tau_j}\right)\left\{\frac{1}{2}\left[\boldsymbol{\Sigma}^*(t,T_j)^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_{j-1})\right] dt\right. \\ &\quad + \left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right]^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t) \\ &\quad \left. + \frac{1}{2}\left[\boldsymbol{\Sigma}^*(t,T_j)^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j) - 2\ \boldsymbol{\Sigma}^*(t,T_{j-1})^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j) + \boldsymbol{\Sigma}^*(t,T_{j-1})^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_{j-1})\right] dt\right\} \\ &= \left(R_j(t) + \frac{1}{\tau_j}\right)\left\{\left[\boldsymbol{\Sigma}^*(t,T_j)^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})^\top \boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j)\right] dt\right. \\ &\quad \left. + \left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right]^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\right\} \\ &= \left(R_j(t) + \frac{1}{\tau_j}\right)\left\{\left[\boldsymbol{\Sigma}^*(t,T_j)^\top - \boldsymbol{\Sigma}^*(t,T_{j-1})^\top\right]\boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j)dt\right. \\ &\quad \left. + \left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right]^\top d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\right\} \\ &= \left(R_j(t) + \frac{1}{\tau_j}\right)\left[\boldsymbol{\Sigma}^*(t,T_j) - \boldsymbol{\Sigma}^*(t,T_{j-1})\right]^\top \left\{\boldsymbol{\rho}\ \boldsymbol{\Sigma}^*(t,T_j)dt + d\mathbf{W}^{\widetilde{\mathbb{P}}}(t)\right\} \end{aligned}$$ □

Theorem 2.3.1. *A caplet (floorlet) with expiry T_{j-1} and maturity T_j is equivalent to a swaplet where the buyer of this financial product can choose to exercise the swaplet only if it has a positive*

value. The payoff at maturity is given by

$$\tau_j\Big(R(T^*,T_{j-1},T_j)-K\Big)^+=\tau_j\Big(\widetilde{R}(T^*,T_{j-1},T_j)-(K+\delta)\Big)^+$$

where K is the strike price of the caplet and $T^*=T_{j-1}$ if the caplet is forward looking and $T^*=T_j$ if it is backward-looking.
The price of this contract at time t can be shown to be

$$\boldsymbol{Bl}^{\,caplet}(t)=\gamma P(t,T_j)\Big(\widetilde{R}(t,T_{j-1},T_j)\Phi(d_1)-(K+\delta)\Phi(d_2)\Big)$$

where δ is the displacement parameter and

$$d_1=\frac{\log\left(\frac{\widetilde{R}_j(t)}{K+\delta}\right)+\frac{1}{2}\int_t^{T^*}\|\mathbf{\Gamma}(s,T_j)\|^2\,ds}{\sqrt{\int_t^{T^*}\|\mathbf{\Gamma}(s,T_j)\|^2\,ds}},\qquad d_2=d_1-\sqrt{\int_t^{T^*}\|\mathbf{\Gamma}(s,T_j)\|^2\,ds}$$

Proof. Since $R_j(t)$ is martingale under the extended $\mathbb{P}^{T_j}$ measure, the risk-neutral price of the caplet with payoff given by (2.63) is

$$\begin{aligned}\mathbf{Bl}^{\,\text{caplet}}(t)&=\mathbb{E}\left[e^{-\int_t^{T_j}r(u)du}\tau_j\Big(\widetilde{R}(T^*,T_{j-1},T_j)-(K+\delta)\Big)^+\,\Big|\,\mathcal{F}_t\right]\\&=\mathbb{E}^{\mathbb{P}^{T_j}}\left[\frac{P(0,T_j)}{P(T_j,T_j)}\tau_j\Big(\widetilde{R}_j(T^*)-(K+\delta)\Big)^+\,\Big|\,\mathcal{F}_t\right]\\&=P(0,T_j)\,\mathbb{E}^{\mathbb{P}^{T_j}}\left[\tau_j\Big(\widetilde{R}_j(T^*)-(K+\delta)\Big)^+\right]\end{aligned}\tag{A.1}$$

where we have changed the measure from $\mathbb{Q}$ to $\mathbb{P}^{T_j}$ and the last line follows from the payoff's independence from $\mathcal{F}_t$. From (2.47), we see that $\widetilde{R}_j(T^*)$ can be written as $f(X)$ with $f(x)=R_j(t)e^x$ and where X is a normal distribution with mean $\mu=-\frac{1}{2}\int_t^{T^*}\|\mathbf{\Gamma}(u,T_j)\|^2\,du$ and variance

$$(\sigma^*)^2=\left(\int_t^{T^*}\mathbf{\Gamma}(t,T_j)^\top d\mathbf{W}^{T_j}(u)\right)^2=\int_t^{T^*}\mathbf{\Gamma}(t,T_j)^\top\mathbf{\Gamma}(t,T_j)dt=\int_t^{T^*}\|\mathbf{\Gamma}(t,T_j)\|^2\,du$$

where we have used Itô's isometry (Shreve (2008). Therefore, we have

$$\begin{aligned}\mathbb{E}^{\mathbb{P}^{T_j}}\left[\tau_j\Big(\widetilde{R}_j(T^*)-(K+\delta)\Big)^+\right]&=\mathbb{E}^{\mathbb{P}^{T_j}}\left[\tau_j\left(\widetilde{R}_j(T^*)-(K+\delta)\right)\mathcal{I}_{\{(\widetilde{R}_j(T^*)-(K+\delta))>0\}}\right]\\&=\int_{-\infty}^{\infty}\tau_j\left(\widetilde{R}_j(T^*)-(K+\delta)\right)\mathcal{I}_{\{(\widetilde{R}_j(T^*)-(K+\delta))>0\}}\phi(x)dx\end{aligned}$$

where $\phi(x)$ is the density function of a standard normal distribution. The integral above is non-zero when

$$\begin{aligned}\widetilde{R}_j(T^*)-(K+\delta)>0&\Leftrightarrow\widetilde{R}_j(t)e^{\mu+\sigma^*\varepsilon}-(K+\delta)>0\\&\Leftrightarrow\varepsilon>-\frac{\log\left(\frac{\widetilde{R}_j(t)}{(K+\delta)}\right)+\mu}{\sigma^*}:=-d_2\end{aligned}$$

where ε has density function $\phi(X)$. Therefore,

$$\begin{aligned}
\mathbb{E}^{\mathbb{P}^{T_j}}\left[\tau_j\Big(\widetilde{R}_j(T^*)-(K+\delta)\Big)^+\right] &= \int_{-\infty}^{\infty}\tau_j\left(\widetilde{R}_j(T^*)-(K+\delta)\right)\mathcal{I}_{\{(\widetilde{R}_j(T^*)-(K+\delta))>0\}}\phi(x)dx \\
&= \int_{-d_2}^{\infty}\tau_j\left(\widetilde{R}_j(T^*)-(K+\delta)\right)\phi(x)dx \\
&= \tau_j\int_{-d_2}^{\infty}\widetilde{R}_j(T^*)\phi(x)dx-\tau_j(K+\delta)\int_{-d_2}^{\infty}\phi(x)dx \\
&= \tau_j\widetilde{R}_j(t)\int_{-d_2}^{\infty}e^{\mu+\sigma^*x}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}dx-\tau_j(K+\delta)(1-\Phi(-d_2)) \\
&= \tau_j\widetilde{R}_j(t)e^{\mu+\frac{1}{2}\sigma^*}\int_{-d_2}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-\sigma^*)^2}dx-\tau_j(K+\delta)(\Phi(d_2)) \\
&= \tau_j\widetilde{R}_j(t)\left(1-\Phi(-d_2+\sigma^*)\right)-\tau_j(K+\delta)\Phi(d_2) \\
&= \tau_j\left(\widetilde{R}_j(t)\Phi(d_1)-(K+\delta)\Phi(d_2)\right)
\end{aligned}$$

where $\Phi(x)=\int_{\infty}^{x}\phi(t)dt$ and

$$\begin{aligned}
d_1 = d_2-\sigma^* &= \frac{\log\left(\frac{\widetilde{R}_j(t)}{(K+\delta)}\right)-\frac{1}{2}\int_0^{T^*}\|\mathbf{\Gamma}(u,T_j)\|^2\,du}{\sqrt{\int_0^{T^*}\|\mathbf{\Gamma}(u,T_j)\|^2\,du}}-\int_0^{T^*}\|\mathbf{\Gamma}(u,T_j)\|^2\,du \\
&= \frac{\log\left(\frac{\widetilde{R}_j(t)}{(K+\delta)}\right)+\frac{1}{2}\int_0^{T^*}\|\mathbf{\Gamma}(u,T_j)\|^2\,du}{\sqrt{\int_0^{T^*}\|\mathbf{\Gamma}(u,T_j)\|^2\,du}}
\end{aligned}$$

Therefore, inserting the above integral in (A.1) yield the desired result. □

Proposition 3.2.1. *The interpolated forward rate $\widetilde{R}(t,S,T)$ has a stochastic differential equation given by*

$$\begin{aligned}
\tau_{S,T}\widetilde{\sigma}_{S,T}(t)\widetilde{R}(t,S,T)dW(t) = {}& h(t)\tau_j\sigma_j(t)g_j(t)\widetilde{R}_j(t)dW(t) \\
&+ f(t)\tau_{j+1}\sigma_{j+1}(t)g_{j+1}(t)\widetilde{R}_{j+1}(t)dW(t)
\end{aligned}$$

where

$$h(t)=\frac{\widehat{\alpha}_S\left(1+\tau_{j+1}R_{j+1}(t)\right)}{1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}}$$

and

$$f(t)=\frac{\left(1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}\right)\left(1-\widehat{\alpha}_T(1-\tau_{j+1}\delta)-\delta\tau_{t,T_{j+1}}\right)}{\left(1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}\right)^2}$$

Furthermore, $\widetilde{\sigma}_{S,T}(t)$ can be approximated by

$$\widetilde{\sigma}_{S,T}(t)=\frac{h(0)\sigma_j(t)g_j(t)\widetilde{R}_j(0)+f(0)\sigma_{j+1}(t)g_{j+1}(t)\widetilde{R}_{j+1}(0)}{\widetilde{R}(0,S,T)}$$

Note that if $\delta=0$, i.e. there is no displacement in the model, $h(t)$ and $f(t)$ simplify to equations

(3.19) and (3.20) respectively.

Proof.

$$1+\tau_{S,T}\widetilde{R}(t,S,T)=\frac{1+\widehat{\alpha}_S\tau_{\eta(S)}\widetilde{R}_{\eta(S)}(t)-\delta\tau_{t,T_{\eta(S)}}}{1+\widehat{\alpha}_T\tau_{\eta(T)}\widetilde{R}_{\eta(T)}(t)-\delta\tau_{t,T_{\eta(T)}}}\prod_{j=\eta(S)+1}^{\eta(T)}1+\tau_jR_j(t)$$

As before, assume $T\in[T_{\eta(S)},T_{\eta(S)+1}]$, and for brevity let $\eta(S)=j$. Therefore, $\eta(T)=j+1$

$$1+\tau_{S,T}\widetilde{R}(t,S,T)=\frac{1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}}{1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}}1+\tau_{j+1}R_{j+1}(t)$$

Now, we apply the product rule to this equation

$$\begin{aligned}d(1+\tau_{S,T}\widetilde{R}(t,S,T))=&\left(\frac{1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}}{1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}}\right)\tau_{j+1}dR_{j+1}(t)\\&+(1+\tau_{j+1}R_{j+1}(t))d\left(\frac{1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}}{1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}}\right)\end{aligned}$$

Consider the diffusion of the term involving the ratio (and ignoring the terms involving dt), we have

$$\begin{aligned}&d\left(\frac{1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}}{1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}}\right)\\&=\frac{\widehat{\alpha}_S\tau_jd\widetilde{R}_j(t)}{1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}}-\frac{\widehat{\alpha}_T\tau_{j+1}d\widetilde{R}_{j+1}(t)\left(1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}\right)}{\left(1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}\right)^2}\\&=\frac{\widehat{\alpha}_S\tau_j\sigma_j(t)g_j(t)\widetilde{R}_j(t)dW(t)}{1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}}-\frac{\widehat{\alpha}_T\tau_{j+1}\sigma_{j+1}(t)g_{j+1}(t)\left(1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}\right)\widetilde{R}_{j+1}(t)dWt(t)}{\left(1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}\right)^2}\end{aligned}$$

Therefore,

$$\tau_{S,T}\widetilde{\sigma}_{S,T}(t)\widetilde{R}(t,S,T)dW(t)=h(t)\tau_j\sigma_j(t)g_j(t)R_j(t)dW(t)+f(t)\tau_{j+1}\sigma_{j+1}(t)g_{j+1}(t)R_{j+1}(t)dW(t)$$

where

$$h(t)=\frac{\widehat{\alpha}_S\tau_j\left(1+\tau_{j+1}R_{j+1}(t)\right)}{1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}}$$

and

$$
\begin{aligned}
f(t) &= \frac{1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}}{1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}} - \frac{\widehat{\alpha}_T\left(1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}\right)(1+\tau_{j+1}R_{j+1}(t))}{\left(1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}\right)^2} \\
&= \frac{\left(1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}\right)\left(1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}-\widehat{\alpha}_T-\widehat{\alpha}_T\tau_{j+1}R_{j+1}(t)\right)}{\left(1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}\right)^2} \\
&= \frac{\left(1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}\right)\left(1+\widehat{\alpha}_T\tau_{j+1}R_{j+1}(t)+\widehat{\alpha}_T\tau_{j+1}\delta-\delta\tau_{t,T_{j+1}}-\widehat{\alpha}_T-\widehat{\alpha}_T\tau_{j+1}R_{j+1}(t)\right)}{\left(1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}\right)^2} \\
&= \frac{\left(1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}\right)\left(1+\widehat{\alpha}_T\tau_{j+1}\delta-\delta\tau_{t,T_{j+1}}-\widehat{\alpha}_T\right)}{\left(1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}\right)^2} \\
&= \frac{\left(1+\widehat{\alpha}_S\tau_j\widetilde{R}_j(t)-\delta\tau_{t,T_j}\right)\left(1-\widehat{\alpha}_T(1-\tau_{j+1}\delta)-\delta\tau_{t,T_{j+1}}\right)}{\left(1+\widehat{\alpha}_T\tau_{j+1}\widetilde{R}_{j+1}(t)-\delta\tau_{t,T_{j+1}}\right)^2}
\end{aligned}
$$

□

B Bootstrapping the CDS Curve

In this section, we explain how to obtain probabilities of default for a party that has CDS actively traded in the market. This obviously will not be possible for some counterparties, and since it is beyond the scope of this thesis to explore such cases, we refer the reader to Brigo et al. (2013) who present many alternatives.

Let us introduce the mathematics of a CDS with minimal assumptions. A CDS is a contract where one party, the protection buyer, pays a premium to another party, the protection seller, in exchange for an insurance against the default of a predefined bond issuer. The contract consists of two legs. The premium leg consists of regular payments equal to the CDS rate s times the notional N. The CDS rate is quoted as an annual rate, but payments schedule can be defined differently. Assume that the payment dates are given by $\widehat{T}_0 < \cdots < \widehat{T}_n$. The value of the premium leg at time 0 is given by

$$\begin{aligned} V_{\text{PREM}}(0) &= N\sum_{i=1}^{n} s\,\mathbb{E}\left[P(0,\widehat{T}_i)\mathcal{I}_{\{\tau>\widehat{T}_i\}}(\widehat{T}_i-\widehat{T}_{i-1})\right] + N\sum_{i=0}^{n} s\,\mathbb{E}\left[P(0,\tau)\mathcal{I}_{\{\widehat{T}_{i-1}<\tau\le\widehat{T}_i\}}\left(\tau-\widehat{T}_i\right)\right] \\ &= Ns\sum_{i=1}^{n}\mathbb{Q}(\tau>\widehat{T}_i)(\widehat{T}_i-\widehat{T}_{i-1})P(0,T_i) + Ns\sum_{i=1}^{n}\int_{\widehat{T}_{i-1}}^{\widehat{T}_i} P(0,t)(t-\widehat{T}_i)d\mathbb{Q}(\tau<t) \\ &= Ns\left[\sum_{i=1}^{n}\mathbb{Q}(\tau>\widehat{T}_i)(\widehat{T}_i-\widehat{T}_{i-1})P(0,T_i) + \sum_{i=1}^{n}\int_{\widehat{T}_{i-1}}^{\widehat{T}_i}\beta_s P(0,t)(t-\widehat{T}_i)\mathbb{Q}(\tau>t)dt\right] \end{aligned} \tag{B.1}$$

where we have assumed independence between $r(t)$ and the default event τ. The second term in the first equation is due to the accrued premium payable if the default event occurs outside payment dates. The second equation simply follows from

$$d\mathbb{Q}(\tau<t) = d\left(1-\exp\left(-\int_0^t\beta_s ds\right)\right) = \beta_s\left(\exp\left(-\int_0^t\beta_s ds\right)\right)dt = \beta_s\mathbb{Q}(\tau>t)dt$$

The protection leg consists of a one-time payment to the protection buyer in case of default, where the payment is the fraction of the notional that is lost due to default (i.e. notional times loss given default). Its value at time 0 is given by

$$V_{\text{PROT}}(0) = (1-R)N\sum_{i=1}^{n}\mathbb{E}\left[P(0,\tau)\mathcal{I}_{\{\widehat{T}_{i-1}<\tau\le\widehat{T}_i\}}\right] = (1-R)N\sum_{i=1}^{n}\int_{\widehat{T}_{i-1}}^{\widehat{T}_i}\beta_s P(0,t)\mathbb{Q}(\tau>t)dt \tag{B.2}$$

The value of the spread is calculated such that the contract has value 0 when it is issued, i.e. such that

$$V_{\text{PREM}}(0) = V_{\text{PROT}}(0) \tag{B.3}$$

With numerical techniques, we can compute the values of β_s from equation (B.3), since we are given the spread s from market data. However, we assume a piecewise constant shape of the

intensities as a function of time, or put mathematically

$$h(t) = \int_0^t \beta_s ds := \sum_{i=1}^{\eta(t)-1} (T_i^* - T_{i-1}^*)\beta_i + \left(t - T_{\eta(t)-1}^*\right) \beta_{\eta(t)} \tag{B.4}$$

such that β_i is a constant on the interval (T_{i-1}^*, T_i^*) where the dates $T_i^*, i = 0, \ldots, n$ are the maturities of the CDS quotes. This feature is more apparent when plotting β_s as a function of s, done in Figure B.1. One could also assume, for example, piecewise linearity (Brigo et al. (2013)).

The technique we use with this assumption is called bootstrapping, because we have to solve for the default intensities in short-dated CDS before moving on the longer-dated CDS. To illustrate this, consider for example a CDS with a maturity of $T_1^* = 1$. For this contract, since we assume β_s to be constant on (T_0^*, T_1^*), we can easily solve equation (B.3) and the integrals that it involves and compute β_s. Next, consider a CDS maturity $T_2^* = 2$. Payments up to $T_1^* = 1$ are easily computed since we just solved for β_s in that interval, so we must solve for β_s in the interval (T_1^*, T_2^*). This is again easily implemented numerically since we know the relevant CDS spread from the market and β_s is assumed constant in that period. Repeating the same steps for every CDS maturity T_i^* available, we can derive the entire intensity curve.

We materialize the bootstrapping procedure described above in an example with hypothetical CDS quotes shown in Table B.1. The CDS quotes are given for 6 months, 1 to 5, 7, 10, 20 and 30 years and each CDS contract has quarterly payments. We assume the recovery rate to be 40%. The intensities stripped from this CDS curve are also presented. In this table, we show the resulting β_s, which are given for up to a certain maturity. So for example, the intensity applicable from 6 months to 1 year is 0.0161. We also compute various probability of survival with this set of intensities. The probability of the survival is simply one minus the probability of default. In other words, the probability of survival up to time t is the probability that default occurs beyond time t, $\mathbb{Q}(\tau > t)$. The date $T_0 = 0$ is set to be 13-feb-2020. This example can be considered medium to high risk, since the probability of survival for 20 and 30 years are quite low. Figure B.1 plots the intensities for the first 10 years to show the piecewise constant shape we have specified. We will not consider IRS with a maturity beyond than 30 years, but if we were, we could simply assume the intensity to be constant beyond that point.

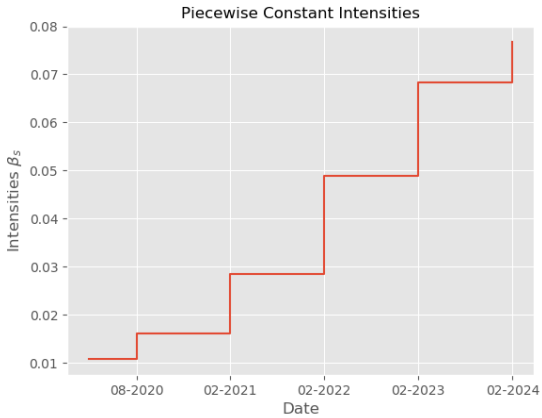
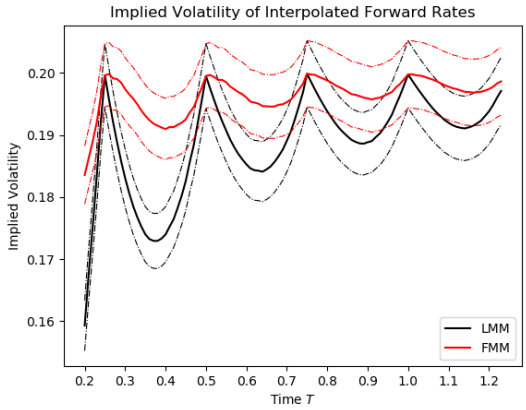


Figure B.1: Piecewise constant shape of intensities of default probabilities

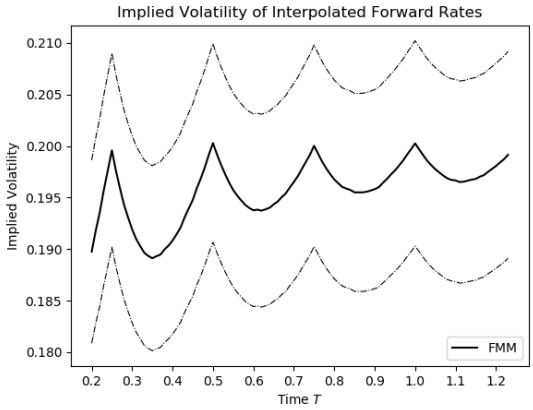
CDS Maturity (in years)	CDS quote (in basis points)	Intensity β_s	Date	Survival probability $\mathbb{Q}(\tau > t)$
0.5	64.388	0.0107	13-aug-2020	99.47%
1	80.481	0.0161	16-feb-2021	98.65%
2	116.451	0.0284	14-feb-2022	95.90%
3	149.814	0.0490	13-feb-2023	91.33%
4	181.751	0.0684	13-feb-2024	85.29%
5	211.014	0.0767	13-feb-2025	78.98%
7	251.895	0.0667	16-feb-2027	69.07%
10	273.457	0.0491	13-feb-2030	59.64%
20	289.547	0.0490	13-feb-2040	37.80%
30	292.952	0.0456	14-feb-2050	23.95%

Table B.1: CDS quotes and bootstrapped intensity parameters for a medium to high risk underlying company.

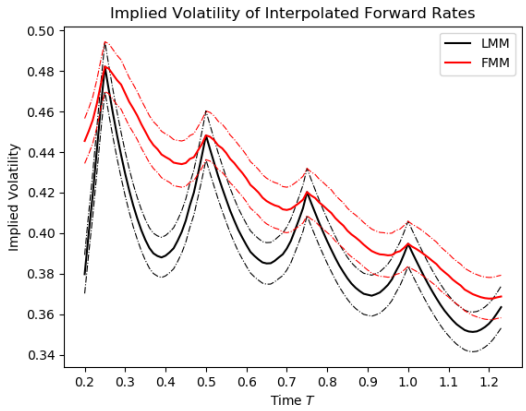
C Figures



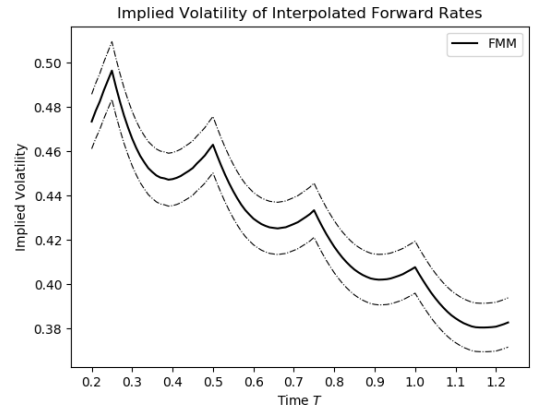
(a) Forward-looking rates with $\delta = 0.03\%$ and constant volatility



(b) Backward-looking rates with $\delta = 0.03\%$ and constant volatility

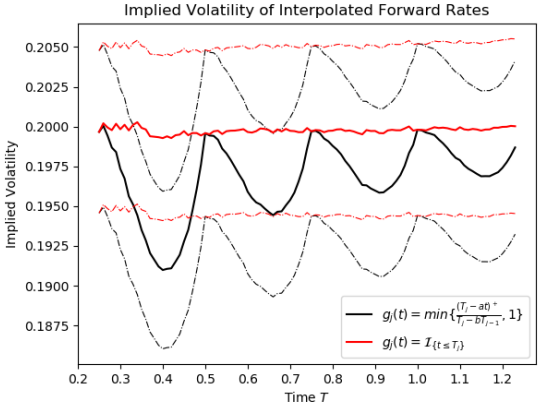


(c) Forward-looking rates with $\delta = 0.03\%$ and exponential volatility

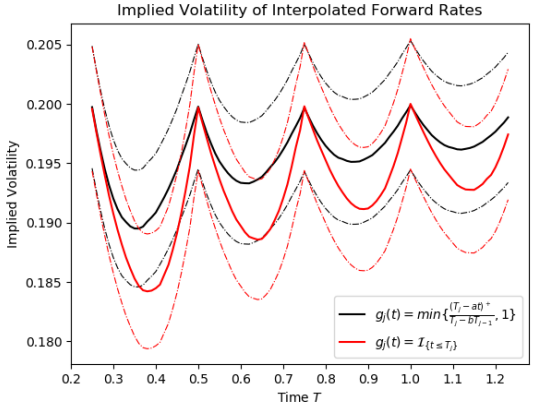


(d) Backward-looking rates with $\delta = 0.03\%$

Figure C.1: Implied volatility of caplets on interpolated forward rates. The dotted line shows the 95% Monte Carlo confidence intervals. Rates were simulated with 12000 simulations for constant volatility and 22000 for exponential volatility.

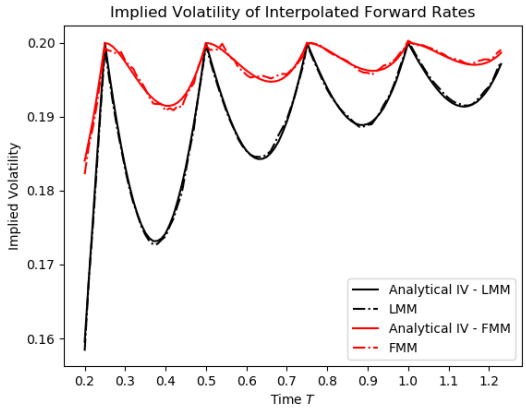


(a) Forward-looking rates with $\delta = 0.03\%$

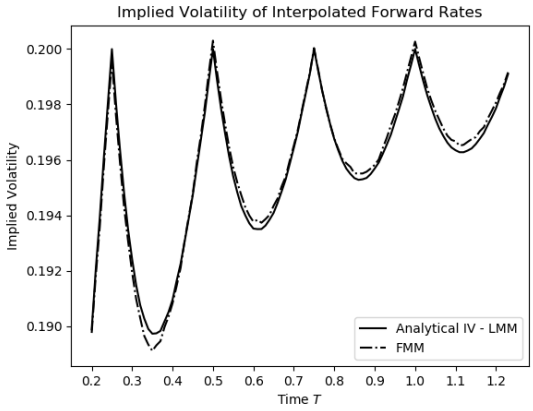


(b) Backward-looking rates with $\delta = 0.03\%$

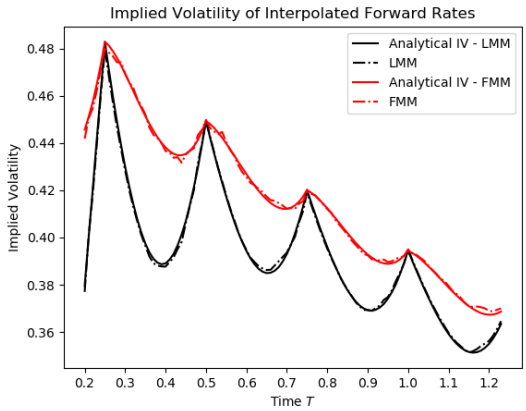
Figure C.2: Implied volatility of forward rates for different diffusion scaling coefficient $g_j(t)$, and the 95% Monte Carlo confidence interval. Rates were simulated using 10000 simulations.



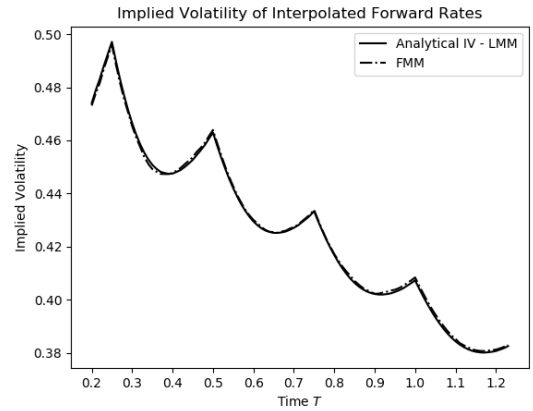
(a) Forward-looking rates with $\sigma_j(t) = 20\%\ \forall t$



(b) Backward-looking rates with $\sigma_j(t) = 20\%\ \forall t$

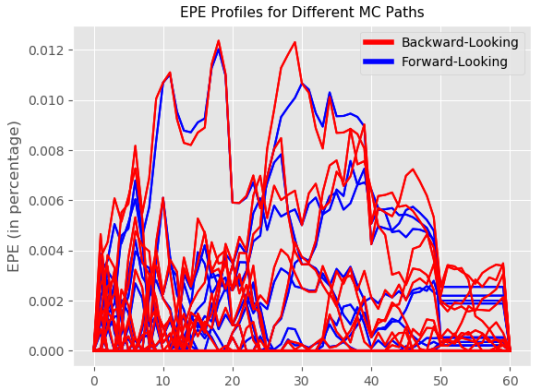


(c) Forward-looking rates with $\sigma_j(t)$ given by (2.50)

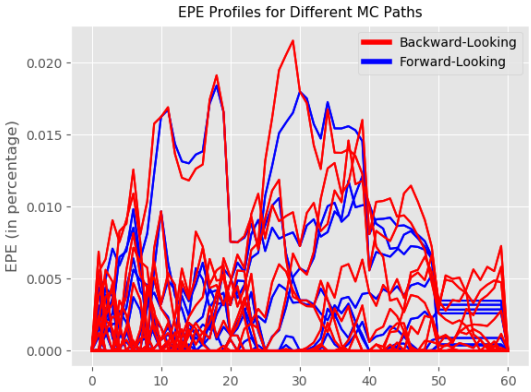


(d) Backward-looking rates with $\sigma_j(t)$ given by (2.50)

Figure C.3: Implied volatility of forward rates and their approximations given by (3.26). Rates were simulated in the LMM and the FMM using 3000 simulations, with a displacement of $\delta = 0.03\%$.



(a) Constant volatility setting and $g_j(t)$ as defined in (3.29)



(b) Exponentially decreasing volatility and $g_j(t)$ as defined in (3.29)

Figure C.4: EPE and MC paths of Positive Exposures of forward and backward-looking IRS with tenor of 0.25 years. Rates were simulated in the FMM using 2000 simulations.

D Numerical implementation

In this section we briefly consider numerical aspects of our model. When it comes to pricing instruments, we have to resort to numerical methods, because an analytical solution to equation (2.58) does not exist. Therefore, we need to approximate this solution and the standard practice is to simulate the forward rates in a Monte Carlo setting by discretizing (2.58) and generating sample paths for the forward rates. Instruments can then be priced by averaging payoffs across the different paths.

It is common in practical situations to diffuse the logarithm of forward rates $\log \widetilde{R}_j(t)$ instead of the rates themselves because it is known to increase numerical stability. Furthermore, the diffusion coefficient under this setting is deterministic (Brigo and Mercurio (2006)). Applying Itô-Doeblin to (2.58), we have

$$\begin{aligned} d\log \widetilde{R}_j(t) &= \frac{\partial \log \widetilde{R}_j(t)}{\partial \widetilde{R}_j(t)} d\widetilde{R}_j(t) + \frac{1}{2}\frac{\partial^2 \log \widetilde{R}_j(t)}{\partial (\widetilde{R}_j(t))^2} d\widetilde{R}_j(t) d\widetilde{R}_j(t) \\ &= \frac{1}{\widetilde{R}_j(t)} \left[g_j(t)\sigma_j(t)\widetilde{R}_j(t) \sum_{i=1}^{j} \rho_{ij} \frac{g_i(t)\sigma_i(t)\widetilde{R}_i(t)}{1+\tau_i R_i(t)} dt + g_j(t)\sigma_j(t)\widetilde{R}_j(t) d\widetilde{W}^{\mathbb{Q}}(t) \right] \\ &\quad + \frac{1}{2}\left(\frac{-1}{(\widetilde{R}_j(t))^2} \right) g_j^2(t)\sigma_j^2(t)\widetilde{R}_j^2(t) dt \\ &= g_j(t)\sigma_j(t) \left[\sum_{i=1}^{j} \rho_{ij} \frac{g_i(t)\sigma_i(t)\widetilde{R}_i(t)}{1+\tau_i R_i(t)} - \frac{1}{2} g_j(t)\sigma_j(t) \right] dt + g_j(t)\sigma_j(t) d\widetilde{W}^{\mathbb{Q}}(t) \end{aligned} \tag{D.1}$$

To discretize equation (D.1), many solution exists. For example we can simply apply a Euler discretization scheme. Let $\Delta t > 0$ be the time step, we can compute realizations of $\log R_j(t)$ as

$$\begin{aligned} \log \widetilde{R}_j(t+\Delta t) = \log \widetilde{R}_j(t) &+ g_j(t)\sigma_j(t) \left[\sum_{i=1}^{j} \rho_{ij} \frac{g_i(t)\sigma_i(t)\widetilde{R}_i(t)}{1+\tau_i R_i(t)} - \frac{1}{2} g_j(t)\sigma_j(t) \right] \Delta t \\ &+ g_j(t)\sigma_j(t)\sqrt{\Delta t}\, \varepsilon \end{aligned} \tag{D.2}$$

where ε is a standard normally distributed random variable $\varepsilon \sim \mathcal{N}(0,1)$, and consequently

$$R_j(t+\Delta) = e^{\log \widetilde{R}_j(t+\Delta)t} - \delta \tag{D.3}$$

where δ is the displacement parameter. It now becomes clear how the displacement is able to produce negative interest rates. The above discretization can be improved in several ways. The above discretization scheme may carry some inaccuracy in the resulting sample (discretization error). Glasserman and Zhao (2000) investigates alternative discretization schemes that focus on excluding arbitrage which may be introduced in schemes such as (D.2). Some authors argue that the inaccuracy of the discretization process (D.2) above is negligible. Alternatively, one can also substantially shrink the time step Δt toward 0 as to mimic a continuous setting as closely as possible, which naturally comes at the cost of a lower numerical speed.

There is a considerable number of other ways to improve numerical accuracy and reduce the number of paths required to obtain numerical convergence. For example, a technique recurrent

in a Monte Carlo context is the predictor-corrector method. The idea behind this method is that the discretization presented above incorrectly assumes the time-varying elements in (D.2) to be fixed at time t and constant over the whole interval $[t, t + \Delta t)$. The predictor-corrector method instead suggests to replace the increment with the average of these elements evaluated at t and Δt respectively. Let $\mu(t)$ denote the instantaneous drift of (D.2), i.e.

$$\mu(t) = g_j(t)\sigma_j(t) \left[\sum_{i=1}^{j} \rho_{ij} \frac{g_i(t)\sigma_i(t)\widetilde{R}_i(t)}{1 + \tau_i R_i(t)} - \frac{1}{2} g_j(t)\sigma_j(t) \right]$$

Then, the predictor-corrector method suggests to compute $\log R_j(t + \Delta t)$ by setting its drift and coefficient as a weighted average of the initial and terminal drift/coefficient of the interval $[t, t + \Delta t]$

$$\begin{aligned} \log R_j(t + \Delta t) = \log R_j(t) &+ \left(\frac{1}{2}\mu(t) + \frac{1}{2}\mu(t + \Delta t) \right) \Delta t \\ &+ \left(\frac{1}{2} g_j(t)\sigma_j(t) + \frac{1}{2} g_j(t + \Delta t)\sigma_j(t + \Delta t) \right) \sqrt{\Delta t}\varepsilon \end{aligned}$$

although it is not strictly necessary to put equal weights on each values. This method is one of many that are described by Jäckel (2002). We implemented two other of his methods which drastically reduced the number of paths required to achieve convergence. Since they are conceptually involved and beyond the scope of this thesis, we only briefly state them. The first method consists of improving the random number generation by using the so-called Sobol number generator, which improves the uniformity of random samples. The second one concerns the path construction. Since we know the covariance matrix of Brownian motion increments (which are independent and of variance equal to of their length of time), we can pre-compute the Brownian motions of each path with a spectral decomposition of the covariance matrix which is faster than an incremental implementation (computing them at each path).