MARKOVIAN HEATH-JARROW-MORTON: VALUATION OF INTEREST RATE DERIVATIVES IN A TWO-FACTOR CHEYETTE GAUSSIAN MODEL

Guillermo Corredor Vegas

Trabajo de investigación 005/019

Máster en Banca y Finanzas Cuantitativas

Directores: Manuel Moreno Fuentes
Gregorio Vargas Martínez

Universidad Complutense de Madrid

Universidad del País Vasco

Universidad de Valencia

Universidad de Castilla-La Mancha

www.finanzascuantitativas.com

Markovian Heath-Jarrow-Morton: Valuation of interest rate derivatives in a Two-Factor Cheyette Gaussian model

Guillermo Corredor Vegas

Master's Thesis

UNIVERSIDAD
COMPLUTENSE
MADRID

Advisors: Manuel Moreno Fuentes,
Gregorio Vargas Martínez

Universidad Complutense de Madrid
July, 2019

Contents

1 Introduction

In the Heath-Jarrow-Morton framework [Heath et al., 1992], the model under the equivalent martingale measure is defined by specifying the volatility of the instantaneous forward rate and the initial forward curve. Some choices of the volatility structure may induce path-dependency on both the short and the instantaneous forward rates. This fact makes certain valuation methods like the use of *partial differential equation* through the *Feynman-Kac* theorem not available, and makes others, such as *Monte Carlo simulation* or *lattice* methods, more computationally intensive.

Several authors like [Carverhill, 1994] and [Ritchken et al., 1995], introduced restrictions on the volatility functions that led to a Markovian structure of the model. In this Thesis, we follow an approach originally proposed by [Cheyette, 1994], and then employed by [Beyna, 2013] in a deterministic volatility setting, where the instantaneous forward and short rates are normally distributed. In this *gaussian* Heath-Jarrow-Morton environment, analytical formulas are obtainable for some instruments and computations generally become easier.

We will start the Thesis with two introductory sections where we review the main interest rate instruments and the tools required to perform its valuation, and we derive the no-arbitrage condition in the Heath-Jarrow-Morton framework.

Then, after detailing the general Cheyette-Beyna specification, we will propose and analyze a Two-Factor Cheyette model with two state variables, in a similar fashion to other *gaussian* models such as the HJM G2++ of [Acar, 2009]. The particular choice of volatility selected could be extended by adding summands without increasing the number of state variables, however a somewhat simple form will be used in order to keep the model manageable. An analytical expression for the value of a caplet in our model is derived afterwards.

In the next section, a calibration procedure will be implemented to adjust model prices to EUR cap market data, obtaining the corresponding parameters by minimizing the squared sum of errors between them. A *simulated annealing* routine will be used to that effect.

The final section addresses the implementation of several numerical methods to value a selection of interest rate derivatives. The Euler scheme will be used to simulate paths of the state variables, and as a function of them, trajectories for the underlying of a barrier caplet. Making use of the *vanilla* caplet analytical formula previously obtained, a control variate estimator is implemented to increase the precision of the valuation procedure.

A European swaption will be valued by computing the expectation of a function of a bivariate normal variable, representing the payoff of the swaption in our model. The resulting two-dimensional integral will be evaluated using the composite Simpson's rule.

Regarding American-style options, several hybrid dynamic programming-simulation methods are available in literature, such as the *stochastic mesh* explained in [Broadie and Glasserman, 2004], or the *least squares Monte Carlo* approach introduced by [Longstaff and Schwartz, 2001]. We choose to implement the *random tree* of [Broadie and Glasserman, 1997] to value a Bermudan swaption. Although its computational complexity scales exponentially with the number of exercise dates, its implementation only requires the ability to simulate paths of the state variables and it greatly benefits from their Markovian property.

2 Fundamentals

2.1 Interest rates and basic instruments

The elementary fixed-income instruments and rates are defined in this Section, following mainly the notes on [Brigo and Mercurio, 2006]. This review will also set the notation for the rest of the Thesis.

Zero coupon bond The value at time t of an asset that pays one unit of currency at time T (with $t < T$) is denoted $P(t,T)$, noting the dependence on both t and T. By trivial no-arbitrage arguments, its value at maturity is $P(T,T) = 1$.

Spot rate (Continuously compounded) / Yield Given $P(t,T)$, the constant continuously compounded interest rate that satisfies the relationship $P(t,T)\exp\left[rate(T-t)\right] = 1$ defines the yield:

$$y(t,T) := -\frac{\log P(t,T)}{T-t} \tag{1}$$

Day-count fraction A function $\delta(t,T)$ that measures the time between two given dates taking into account the day-counting conventions of the contract. For simplicity we will set $\delta(t,T) := T - t$ where T and t are both real numbers. When we refer to a constant time step we will simply write $\delta(t,T) := \delta$.

Spot rate (Simply compounded) Given $P(t,T)$, the simply compounded interest rate that satisfies $P(t,T)[1 + rate(T-t)] = 1$ determines:

$$L(t,T) := \frac{1}{(T-t)}\left(\frac{1}{P(t,T)} - 1\right) \tag{2}$$

FRAs It is defined as a contract closed at time t that specifies a payoff at a later time T_2, of $Nom(T_2 - T_1)\left[K - L(T_1,T_2)\right]$, with $t < T_1 < T_2$. Its time t value[1] is

$$\mathbf{V_{FRA}}(t) = Nom\left[P(t,T_2)K(T_2 - T_1) - P(t,T_1) + P(t,T_2)\right] \tag{3}$$

[1] [Brigo and Mercurio, 2006] uses replication arguments to obtain the value of the FRA, without making use of the fundamental pricing equation which is introduced in Section 1.2. We prefer the latter approach so here we present the valuation results without proof.

The strike K that makes a FRA fair (i.e. its value equal to 0) at time t defines the forward simple rate:

Forward rate (Simple rate)

$$F(t,T_1,T_2) := \frac{1}{T_2 - T_1}\left(\frac{P(t,T_1)}{P(t,T_2)} - 1\right) \tag{4}$$

Considering the limit when T_2 approaches T_1, the instantaneous forward rate is characterized as:

Instantaneous Forward rate

$$f(t,T) := \lim_{T_2 \to T_1^+} F(t,T_1,T_2) = -\frac{\partial}{\partial T}\left[logP(t,T)\right] \tag{5}$$

Remark From equation (5), integrating from t to T with respect to the variable T, we can also note the relationship:

$$P(t,T) = \exp\left(-\int_t^T f(t,u)du\right) \tag{6}$$

When T approaches t in the spot rate (1) or (2), that is, $\lim_{T \to t^+} L(t,T)$ or $\lim_{T \to t^+} y(t,T)$, we obtain the instantaneous spot rate. A more convenient definition is the following:

Instantaneous Spot rate

$$r(t) := f(t,t) \tag{7}$$

Money market account Represents the value at time t of an investment of one unit of currency at time 0 in a savings account which accrues interest at the instantaneous spot rate. It is defined as:

$$B(t) := \exp\left(\int_0^t r(s)ds\right) \tag{8}$$

Swap (receiver forward start) The payments will take place at dates $T_2, T_3, ...T_{n+1}$ and the rates will reset at previous dates $T_1, T_2, ...T_n$, with $0 \leq ... < t < ... < T_1 < T_2 < ... < T_{n+1}$.

The payoff at every payment date T_i is:

$$\underbrace{Nom(T_i - T_{i-1})K}_{\text{receive fixed leg}} - \underbrace{Nom(T_i - T_{i-1})L(T_{i-1}, T_i)}_{\text{pay floating leg}} \tag{9}$$

Its value at time t, can be evaluated considering it as a portfolio of FRAs, resulting in:

$$\mathbf{V_{swap}}(t) = \sum_{i=2}^{n+1} NomP(t, T_i)(T_i - T_{i-1})K + NomP(t, T_{n+1}) - NomP(t, T_1) \tag{10}$$

2.2 No-arbitrage pricing in continuous time

The basic tools that will be required for valuation purposes will be outlined. We will start with several definitions and results from [Musiela and Rutkowski, 1997] that will set the stage for the statement of the most important result in the Section, the Fundamental Pricing Equation.

$$\mathbb{Q} \sim \mathbb{P} \text{ is a martingale measure } \iff \frac{S(t)}{B(t)} \text{is a } \mathbb{Q} \text{ - local martingale} \quad (11)$$

with $S(t)$ being a tradable asset.

$$\mathbb{Q} \sim \mathbb{P} \text{ is a spot martingale measure } \iff \frac{V_\phi(t)}{B(t)} \text{ is a } \mathbb{Q} \text{ - local martingale} \quad (12)$$

for every self financing trading strategy ϕ. [2]

The interesting result is that, under no dividends paid by the underlying, (11) $\iff$ (12), that is

$$\frac{S(t)}{B(t)} \text{ is a } \mathbb{Q} \text{ - local martingale } \iff \frac{V_\phi(t)}{B(t)} \text{ is a } \mathbb{Q} \text{ - local martingale} \quad (13)$$

so by finding an equivalent measure under which the tradable asset divided by a numeraire is a (local) martingale, we have automatically found a measure that also makes a (local) martingale the discounted value process of a self financing trading strategy.

Fundamental pricing equation Consider a self-financing portfolio, with t value $V(t)$, that replicates an $\mathcal{F}_T$-measurable claim C. Then, the time t price of the claim $\pi_C(t)$ has to verify:

$$\pi_C(t) = V(t) \quad \forall t \quad (14)$$

or else there exists an arbitrage opportunity.[3]

[2]We will assume that it satisfies any of the sufficient conditions in [Protter, 2003] to be a martingale. [Musiela and Rutkowski, 1997] address this issue by working with *admissible* strategies, defined as those whose discounted values $\frac{V(t)}{B(t)}$ follow martingales under $\mathbb{Q}$.

[3]This is clearly a situation that every valuation model has to forbid in order to be economically sound. An arbitrage opportunity is an strategy such that:

- $V(0) = 0$
- $V(T) \geq 0$
- $\mathbb{E}[V(T)] > 0$

Under the martingale measure, $\frac{V(t)}{B(t)}$ is a martingale (see (13)) so it holds that:

$$\frac{V(t)}{B(t)} = \mathbb{E}_{\mathbb{Q}}\left[\frac{V(T)}{B(T)}|\mathcal{F}_t\right] \tag{15}$$

As the existence of an equivalent martingale measure $\mathbb{Q}$ rules out arbitrage opportunities,[4] we can combine (14) and (15), and the fact that for a replicating strategy $V(T) = C$, to state:

$$\pi_C(t) = V(t) = B(t)\mathbb{E}_{\mathbb{Q}}\left[\frac{C}{B(T)}|\mathcal{F}_t\right] \tag{16}$$

The martingale measure plays therefore a crucial role, both forbidding arbitrage and allowing us to price a claim just by finding and expectation, rather than trying to figure out the composition of the portfolio strategy (the hedge portfolio, also of interest, should be found by other means).

Change of measure A practical statement of Girsanov's theorem applied to Brownian Motion, based on [Baxter and Rennie, 1996] is:

$W(t)$ is a $\mathbb{P}$-Brownian motion, and $\gamma(t)$ is and adapted process satisfying the Novikov[5] condition

$$\Longrightarrow$$

there exists a probability measure $\mathbb{Q} \sim \mathbb{P}$, such that $\widetilde{W}(t) := W(t) + \int_0^t \gamma(s)ds$ is a $\mathbb{Q}$-Brownian motion (in differential notation, we write $d\widetilde{W}(t) = dW(t) + \gamma(t)dt$).

So under certain conditions, we can *manipulate* by a change of measure the drift of a process, obtaining when possible the desired driftless dynamics required for the martingale measure in (11) .

[4] This important result is sometimes referred to as the First Fundamental Theorem of Asset Pricing. The proof in a intuitive discrete-time setting can be found in [Pliska, 1997], its continuous-time analogue, however, poses some challenges, as discussed in [Musiela and Rutkowski, 1997] or [Sondermann, 2006]. We don't delve into any detail and assume that the Theorem holds for all relevant situations.

[5] Novikov condition, i.e. $\mathbb{E}_{\mathbb{P}}\left[\exp\left(\frac{1}{2}\int_0^T \gamma^2(s)ds\right)\right] < \infty$, is a a sufficient condition for $\mathcal{E}(\gamma)(t)$ to be a martingale and to define a probability measure via $\mathcal{E}(\gamma) = \frac{d\mathbb{Q}}{d\mathbb{P}}$.

2.3 [Heath et al., 1992] framework: No-arbitrage condition

We briefly review the no-arbitrage condition in the Heath-Jarrow-Morton setup of [Heath et al., 1992]. The exposition is based on the notes of [Cairns, 2004] and [Baxter and Rennie, 1996]. We concentrate on the situation where $f(t,T)$ is driven by one factor for clarity's sake, although we will be dealing with a two-factor model in Section 5.

In this framework, the instantaneous forward rate is modeled under $\mathbb{P}$ as:

$$f(t,T) = f(0,T) + \int_0^t \alpha(u,T)du + \int_0^t \sigma(u,T)dW(u) \tag{17}$$

where, with fixed T, $\alpha(t,T)$ and $\sigma(t,T)$ are adapted processes in time t. The initial forward curve $f(0,T)$ is assumed to be known. Therefore, for a fixed T, $f(t,T)$ is a process satisfying[6]:

$$d_t f(t,T) = \alpha(t,T)dt + \sigma(t,T)dW(t) \tag{18}$$

Money market account From equations (8), (17) and (7) we have:

$$\begin{aligned} B(t) &= \exp\left(\int_0^t r(s)ds\right) \\ &= \exp\left(\int_0^t \left[f(0,s) + \int_0^s \alpha(u,s)du + \int_0^s \sigma(u,s)dW(u)\right] ds\right) \end{aligned}$$

And changing the order of integration on both double integrals: [7]

$$B(t) = \exp\left(\int_0^t f(0,s)ds + \int_0^t \int_u^t \alpha(u,s)dsdu + \int_0^t \int_u^t \sigma(u,s)dsdW(u)\right)$$

Zero-coupon bond By equation (6) and again changing order of integration:

$$\begin{aligned} P(t,T) &= \exp\left(-\int_t^T f(t,s)ds\right) \\ &= \exp\left(-\int_t^T f(0,s)ds - \int_0^t \int_t^T \alpha(u,s)dsdu - \int_0^t \int_t^T \sigma(u,s)dsdW(u)\right) \end{aligned}$$

[6]Subindex t included to remark the fact that $f(t,T)$ is a process over time t.

[7]Some technical conditions are required for changing the order of integration on both Riemann and Itô integrals, they are included in the assumptions made by [Heath et al., 1992] . The statement of the Fubini-style result for stochastic integrals can be found in [Filipovic, 2009].

We are interested in the discounted asset dynamics $Z(t)$ in order to define a measure $\mathbb{Q}$ under which $Z(t)$ is a martingale.

Discounted zero-coupon bond The discounted bond can be computed as:

$$\begin{aligned} Z(t,T) &:= \frac{P(t,T)}{B(t)} \\ &= \exp\left(-\int_0^T f(0,s)ds - \int_0^t \int_u^T \alpha(u,s)dsdu - \int_0^t \int_u^T \sigma(u,s)dsdW(u)\right) \end{aligned}$$

It is convenient to set[8]:

$$\Sigma(u,T) := -\int_u^T \sigma(u,s)ds \tag{19}$$

So we write:

$$Z(t,T) = \exp\left(-\int_0^T f(0,s)ds - \int_0^t \int_u^T \alpha(u,s)dsdu + \int_0^t \Sigma(u,T)dW(u)\right)$$

Its dynamics can be computed setting $Z(t,T) := \exp(Y(t,T))$ and applying Itô's lemma:

$$\begin{aligned} d_t Z(t,T) &= Z(t,T)dY(t,T) + \frac{1}{2}Z(t,T)d\langle Y,Y\rangle_t \\ &= Z(t,T)\left[\left(\int_t^T \alpha(t,s)ds\right)dt + \Sigma(t,T)dW(t)\right] + \frac{1}{2}Z(t,T)\Sigma^2(t,T)dt \\ &= Z(t,T)\left[\left[\left(\int_t^T \alpha(t,s)ds\right) + \frac{1}{2}\Sigma^2(t,T)\right]dt + \Sigma(t,T)dW(t)\right] \\ &= Z(t,T)\Sigma(t,T)\left[dW(t) + \left(\frac{1}{2}\Sigma(t,T) - \frac{1}{\Sigma(t,T)}\int_t^T \alpha(t,s)ds\right)dt\right] \end{aligned}$$

Considering $\gamma(t) := \frac{1}{2}\Sigma(t,T) - \frac{1}{\Sigma(t,T)}\int_t^T \alpha(t,s)ds$, we can apply Girsanov's theorem to define an equivalent probability measure $\mathbb{Q}$ under which:

$$\widetilde{W}(t) := W(t) + \int_0^t \gamma(u)du = W(t) + \int_0^t \frac{1}{2}\Sigma(u,T) - \frac{1}{\Sigma(u,T)}\left(\int_u^T \alpha(u,s)ds\right)du$$

is a $\mathbb{Q}$-Brownian motion.

[8] We will see later that this expression can be interpreted as the bond price volatility.

The discounted bond dynamics under $\mathbb{Q}$ become:

$$d_t Z(t,T) = Z(t,T)\Sigma(t,T)d\widetilde{W}(t) \tag{20}$$

so $Z(t,T)$ is, up to a technical condition, a $\mathbb{Q}$-martingale.

Rearranging the expression for $\gamma(t)$:

$$-\gamma(t)\Sigma(t,T) + \frac{1}{2}\Sigma(t,T)^2 = \int_t^T \alpha(t,s)ds \tag{21}$$

Differentiating both sides with respect to T (applying Leibniz's integral rule) and rearranging we obtain:

$$\alpha(t,T) = \gamma(t)\sigma(t,T) - \Sigma(t,T)\sigma(t,T) = \sigma(t,T)[\gamma(t) - \Sigma(t,T)] \tag{22}$$

With this relationship, that implies absence of arbitrage by the existence of $\mathbb{Q}$, we can now go back to the instantaneous forward rate $\mathbb{P}$-dynamics (18) and apply (22):

$$\begin{aligned} d_t f(t,T) &= \alpha(t,T)dt + \sigma(t,T)dW(t) \\ &= \sigma(t,T)[\gamma(t) - \Sigma(t,T)]dt + \sigma(t,T)dW(t) \\ &= \sigma(t,T)[\underbrace{dW(t) + \gamma(t)dt}_{d\widetilde{W}(t)} - \Sigma(t,T)]dt] \\ &= -\Sigma(t,T)\sigma(t,T)dt + \sigma(t,T)d\widetilde{W}(t) \end{aligned} \tag{23}$$

obtaining this way the dynamics of $f(t,T)$ under $\mathbb{Q}$.

As we can see, the instantaneous forward rate $\mathbb{Q}$-dynamics are fully determined by the specification of $\sigma(t,T)$.

Bond dynamics under $\mathbb{Q}$ are easily computed using Itô's product rule realizing that $P(t,T) = Z(t,T)B(t)$:

$$\begin{aligned} d_t P(t,T) &= d_t[Z(t,T)B(t)] \\ &= Z(t,T)\Sigma(t,T)d\widetilde{W}(t)B(t) + r(t)B(t)dt\ Z(t,T) \\ &= P(t,T)[r(t)dt + \Sigma(t,T)d\widetilde{W}(t)] \end{aligned} \tag{24}$$

Equation (24) is the reason why $\Sigma(t,T)$ is known as the bond price volatility (under $\mathbb{Q}$).

2.4 Forward measure

The concept of change of numeraire will be thoroughly used in later Sections, so we proceed to describe the basics.

Numeraire Based on the exposition by [Privault, 2018] a suitable numeraire $N(t)$ needs to verify:

1. $N(t) > 0 \quad \forall t$
2. N is not a dividend paying asset
3. $\frac{N(t)}{B(t)}$ is a $\mathbb{Q}$-martingale

with $\mathbb{Q}$ representing the measure associated with the money market account numeraire $B(t)$.

Forward measure Given a numeraire $N(t)$, the forward measure $\widehat{\mathbb{P}}$ is defined through the Radon-Nikodym derivative as:

$$\frac{d\widehat{\mathbb{P}}}{d\mathbb{Q}} := \frac{\frac{N(T)}{N(0)}}{\frac{B(T)}{B(0)}} = \frac{N(T)}{N(0)} \frac{1}{B(T)} \tag{25}$$

Applicability of the forward measure [Brigo and Mercurio, 2006] enunciate a wider version of the result, but for our needs it is enough to state:

$$\frac{S(t)}{B(t)} \text{ is a martingale under } \mathbb{Q} \Rightarrow \frac{S(t)}{N(t)} \text{ is a martingale under } \widehat{\mathbb{P}} \text{ as defined in (25)}$$

Forward measure with zero-coupon bond numeraire In our analysis, the zero-coupon bonds (for different maturities) will be our assets, and a zero-coupon bond as well, but with a fixed maturity, our numeraire [9].

The fact that $P(T,T)$ is equal to 1, in contrast to $B(T)$, which needs information up to T to be known, makes convenient the use of the zero-coupon bond as a numeraire.

[9] The zero-coupon bond is indeed a suitable numerarie: it is reasonable to assume that its value, even in the worst scenario, is always greater than zero. Also it doesn't have intermediate payments and as seen in (20), it is a $\mathbb{Q}$-martingale when divided by the money market account.

We proceed to set $N(t) := P(t,T)$, and consider another bond $P(t,U)$ as the asset, with $U > T$. The bond price discounted by our new numeraire, should be a martingale under $\widehat{\mathbb{P}}^T$, so we will look for an appropriate change of measure,[10] setting:

$$Y(t,U) := \frac{P(t,U)}{P(t,T)}$$

Recalling the dynamics of the bond price under $\mathbb{Q}$ (see (24)), and using Itô's product rule, the dynamics of $Y(t,U)$ are:

$$d_tY(t,U) = Y(t,U)(\Sigma(t,U) - \Sigma(t,T))d\widetilde{W}(t) + Y(t,U)\Sigma^2(t,T)dt - Y(t,U)\Sigma(t,U)\Sigma(t,T)dt$$

Rearranging terms to clearly see the shape of the $\gamma_T(t)$[11] required to obtain driftless dynamics:

$$\begin{aligned}
d_tY(t,U) &= Y(t,U)(\Sigma(t,U) - \Sigma(t,T))\Bigg[d\widetilde{W}(t) \\
&+ \frac{\Sigma^2(t,T)}{\Sigma(t,U) - \Sigma(t,T)}dt - \frac{\Sigma(t,U)\Sigma(t,T)}{\Sigma(t,U) - \Sigma(t,T)}dt\Bigg] \\
&= Y(t,U)(\Sigma(t,U) - \Sigma(t,T))\left[d\widetilde{W}(t) - \frac{\Sigma(t,T)[\Sigma(t,T) - \Sigma(t,U)]}{\Sigma(t,T) - \Sigma(t,U)}dt\right] \\
&= Y(t,U)(\Sigma(t,U) - \Sigma(t,T))\left[d\widetilde{W}(t) - \Sigma(t,T)dt\right]
\end{aligned}$$

Setting $\gamma_T(t) := -\int_0^t \Sigma(s,T)ds$ and assuming that the required technical conditions hold, we can use Girsanov's theorem again to define an equivalent probability measure $\widehat{\mathbb{P}}^T$ such that:

$$d\widehat{W}^T(t) = d\widetilde{W}(t) - \Sigma(t,T)dt \tag{26}$$

is a $\widehat{\mathbb{P}}^T$-Brownian motion.

Finally, applying (26) we get the sought driftless dynamics:

$$d_tY(t,U) = Y(t,U)(\Sigma(t,U) - \Sigma(t,T))d\widehat{W}^T(t)$$

[10] The steps followed here are found in [Cairns, 2004], which seem the most intuitive for justifying the form of the forward measure with bond numeraire; first computing the discounted dynamics, and then defining a change of measure that will make it a (local) martingale under the new measure.

[11] We include a T subindex on $\gamma(t)$ to remark the dependence of the particular maturity of the bond used as numeraire.

Conditional expectation under $\widehat{\mathbb{P}}^T$ and the Fundamental Pricing Equation [Privault, 2018] shows that:

$$\frac{d\widehat{\mathbb{P}}_{\mathcal{F}_t}}{d\mathbb{Q}_{\mathcal{F}_t}} = \frac{B(t)}{B(T)}\frac{N(T)}{N(t)}$$

Then, for any $\mathcal{F}_T$-measurable payoff C:

$$\begin{aligned} \mathbb{E}_{\mathbb{Q}}\left[\frac{B(t)}{B(T)}C|\mathcal{F}_t\right] = \mathbb{E}_{\mathbb{Q}}\left[\frac{B(t)}{B(T)}\frac{N(t)}{N(t)}\frac{N(T)}{N(T)}C|\mathcal{F}_t\right] &= \mathbb{E}_{\mathbb{Q}}\left[\frac{d\widehat{\mathbb{P}}_{\mathcal{F}_t}}{d\mathbb{Q}_{\mathcal{F}_t}}\frac{N(t)}{N(T)}C|\mathcal{F}_t\right] \\ &= \mathbb{E}_{\widehat{\mathbb{P}}}\left[\frac{N(t)}{N(T)}C|\mathcal{F}_t\right] \end{aligned} \tag{27}$$

Considering the zero-coupon bond numeraire, the result in (27) becomes:

$$\mathbb{E}_{\mathbb{Q}}\left[\frac{B(t)}{B(T)}C|\mathcal{F}_t\right] = \mathbb{E}_{\widehat{\mathbb{P}}}\left[\frac{P(t,T)}{P(T,T)}C|\mathcal{F}_t\right] = \mathbb{E}_{\widehat{\mathbb{P}}}\left[P(t,T)C|\mathcal{F}_t\right]$$

So we have arrived at the relation:

$$B(t)\mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B(T)}C|\mathcal{F}_t\right] = P(t,T)\mathbb{E}_{\widehat{\mathbb{P}}}\left[C|\mathcal{F}_t\right] \tag{28}$$

where the left-hand side corresponds to the Fundamental Pricing Equation (16).

3 Valuation of interest rate derivatives: General results

The valuation of several interest rate derivatives will be outlined here, obtaining expressions that will be used in Section 7. We set the nominal of every instrument to 1 ($Nom := 1$) to make the exposition simpler.

3.1 *Vanilla* caplet

For a *vanilla* caplet with strike k, settle date T_1 and maturity T_2, the payoff is defined as:

$$\delta(F(T_1,T_1,T_2)-k)^+ = \delta(L(T_1,T_2)-k)^+$$

being the payoff $\mathcal{F}_{T_1}$-measurable

By the fundamental pricing equation (16):

$$\begin{aligned}
\mathbf{V_{caplet}}(t) &= B(t)\mathbb{E}_{\mathbb{Q}}\left[\frac{1}{B(T_2)}\delta(L(T_1,T_2)-k)^+|\mathcal{F}_t\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^{T_2} r(s)ds\right)\delta(L(T_1,T_2)-k)^+|\mathcal{F}_t\right]
\end{aligned}$$

Making use of iterated conditioning on $\mathcal{F}_{T_1}$ and noticing the fact that $P(T_1,T_2) = \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_{T_1}^{T_2} r(s)ds\right)|\mathcal{F}_{T_1}\right]$, we get:

$$\begin{aligned}
\mathbf{V_{caplet}}(t) &= \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^{T_1} r(s)ds\right)P(T_1,T_2)\delta[L(T_1,T_2)-k]^+|\mathcal{F}_t\right] \\
&= \mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^{T_1} r(s)ds\right)P(T_1,T_2)\delta\left[\frac{1}{\delta}\left(\frac{1}{P(T_1,T_2)}-1\right)-k\right]^+|\mathcal{F}_t\right]
\end{aligned}$$

Multiplying by $\dfrac{1+k\delta}{1+k\delta}$ and defining $K' := \dfrac{1}{1+k\delta}$:

$$\mathbf{V_{caplet}}(t) = \underbrace{1+k\delta}_{\frac{1}{K'}}\mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^{T_1} r(s)ds\right)\left[\underbrace{\frac{1}{1+k\delta}}_{K'}-P(T_1,T_2)\right]^+|\mathcal{F}_t\right]$$

We can see that this last expression is just $\frac{1}{K'}$ times the t price of a put option with strike K' and maturity T_1 written on the bond $P(T_1,T_2)$.

The computation of the expectation can be simplified by means of a change of measure as noted in Section 2.4. Using (28), we get the final expression for the value of a *vanilla* caplet:

$$\begin{aligned}\mathbf{V}_{\mathbf{caplet}}(t) &= \frac{1}{K'}\mathbb{E}_{\mathbb{Q}}\left[\exp\left(-\int_t^{T_1} r(s)ds\right)\left[K' - P(T_1,T_2)\right]^+ |\mathcal{F}_t\right] \\ &= \frac{1}{K'}P(t,T_1)\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\left[K' - P(T_1,T_2)\right]^+ |\mathcal{F}_t\right] \qquad (29)\end{aligned}$$

3.2 Barrier caplet

As an example of a path-dependent derivative, we consider a barrier down-and-out caplet with strike k, settle date T_1 and maturity $T_1 + \delta$ over the δ-period simple spot rate with barrier level B. The payoff of such instrument is:

$$\delta\Bigg(L(T_1, T_1+\delta) - k\Bigg)^+ \mathbb{1}\{\tau_B > T_1\}$$

with

$$\tau_B := \inf\{t_i \mid L(t_i, t_i+\delta) < B\} \qquad (30)$$

so τ_B represents the first moment the time-t_i simple forward rate falls below the barrier level. The expression $\{\tau_B > T_1\}$ states that the crossing event happens after the expiry of the option (so, for valuation purposes, the underlying does not cross the barrier and the indicator function takes value 1).

Being $\mathbb{1}\{\tau_B > T_1\}$ an $\mathcal{F}_{T_1}$ measurable random variable, we can use an analogous procedure to the previous one to value the barrier caplet, obtaining:

$$\mathbf{V}_{\mathbf{barriercaplet}}(t) = \frac{1}{K'}P(t,T_1)\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\left[K' - P(T_1,T_2)\right]^+ \mathbb{1}\{\tau_B > T_1\}|\mathcal{F}_t\right] \qquad (31)$$

3.3 European swaption

Considering the swap defined in Section 2.1, its value at T_1 (noting (10)) is:

$$\mathbf{V}_{\mathbf{swap}}(T_1) = \sum_{i=2}^{n+1} P(T_1,T_i)(T_i - T_{i-1})K + P(T_1,T_{n+1}) - P(T_1,T_1)$$

The option (European swaption) to enter the swap at T_1 with strike K has a T_1 payoff of:

$$\left(K\sum_{i=2}^{n+1} P(T_1,T_i)(T_i-T_{i-1})+P(T_1,T_{n+1})-1\right)^+ \tag{32}$$

Again, applying the fundamental pricing equation (16) over the payoff (32), we get:

$$\mathbf{V}_{\mathbf{ESwaption}}(t) \;=\; \mathbb{E}_{\mathbb{Q}}\left[\frac{B(t)}{B(T)}\left(K\sum_{i=2}^{n+1} P(T_1,T_i)(T_i-T_{i-1})+P(T_1,T_{n+1})-1\right)^+ |\mathcal{F}_t\right]$$

And finally, under $\widehat{\mathbb{P}}^{T_1}$:

$$\mathbf{V}_{\mathbf{ESwaption}}(t)=P(t,T)\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\left(K\sum_{i=2}^{n+1} P(T_1,T_i)(T_i-T_{i-1})+P(T_1,T_{n+1})-1\right)^+ |\mathcal{F}_t\right] \tag{33}$$

3.4 Bermudan swaption

American-style derivatives introduce a layer of complexity over the previous instruments, and cannot be handled directly with the fundamental pricing equation. Here we just formulate the problem here in Section 7 we will explore numerical procedures to approximate the value of the option.

Following [Glasserman, 2003] and considering the payoff in (32) the value of the Bermudan option at $t=0$ maturing at T_1 is:

$$\mathbf{V}_{\mathbf{BSwaption}}(0)=\sup_\tau\left[\mathbb{E}_{\mathbb{Q}}\left[\exp\left(\int_0^\tau r(s)ds\right)\left(K\sum_{i=2}^{n+1} P(\tau,T_i)(T_i-T_{i-1})+P(\tau,T_{n+1})-1\right)^+\right]\right]$$

where the $\sup_\tau$ represents the computation of the supremum over the set of exercise strategies τ taking values in $[0,T_1]$.

Again, by a change of measure and applying (28):

$$\mathbf{V}_{\mathbf{BSwaption}}(0)=\sup_\tau\left[P(0,\tau)\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\left(K\sum_{i=2}^{n+1} P(\tau,T_i)(T_i-T_{i-1})+P(\tau,T_{n+1})-1\right)^+\right]\right]$$

4 [Cheyette, 1994] approach: Justification and model specification

An adapted process $X(t)$ is Markov if

$$\mathbb{E}[h(X(u))|\mathcal{F}_t] = \mathbb{E}[h(X(u))|X_t], \ \forall\ 0 \leq t \leq u$$

A more practical statement about the Markov property is found in [Protter, 2003], where it is stated that (under certain technical conditions), a process with differential:

$$dX(t) = h(X(t), t)dt + g(X(t), t)dW(t) \tag{34}$$

with h and g being functions of $X(t)$ and t, is a Markov process.

In the Heath-Jarrow-Morton setting, the instantaneous short rate is:

$$r(t) = f(t, t) = f(0, t) + \int_0^t \alpha(u, t)du + \int_0^t \sigma(u, t)dW(u)$$

Focusing on the second integral:

$$A(t) := \int_0^t \sigma(u, t)dW(u)$$

This process has differential $dA(t) = \sigma(t, t)dW(t) + \left(\int_0^t \frac{\partial\sigma(u,t)}{\partial t}dW(u)\right)dt$ so, in general, cannot be written in the form of (34) and therefore is not Markovian. This has attracted some research and several authors have specified conditions where the Markovianity is preserved.[12].

In this Thesis, we follow the work of [Cheyette, 1994], however, we will restrict ourselves to the specification used by [Beyna, 2013] where the functions for the instantaneous forward rate volatility are deterministic in order to work on a *gaussian* Heath-Jarrow-Morton environment, therefore excluding stochastic volatility or dependence on the instantaneous forward rate.[13]

[12]For example in [Ritchken et al., 1995], they are able to express a one factor Heath-Jarrow-Morton structure as a two-state variables Markov process by imposing:

$$\sigma_f(t, T) = \sigma_r \exp\left(-\int_t^T \kappa(u)du\right)$$

.

[13]The results in [Cheyette, 1994] allow for the use of non-deterministic volatility functions $\sigma(t, T, \omega)$ while still preserving Markovianity.

4.1 Cheyette-Beyna model specification

The specification for each volatility function is as follows:[14]

$$\sigma_k(t,T) := \sum_{i=1}^{N_k} \frac{\alpha_{ik}(T)}{\alpha_{ik}(t)} \beta_{ik}(t), k = 1, 2...M \tag{35}$$

The total number of summands $\sum_{k=1}^{M} N_k$ of each function through all functions determines the number of state variables in the model[15].

An interesting feature is that the function $\beta(t)$ can be set to a polynomial of any order without increasing the complexity of the model in terms of factors or state variables, because it does not add additional summands in the volatility functions in (35).

The model relies on the definition of the state variables $X_{ij}(t)$. Other processes have to be defined as well:

Auxiliary processes

$$A_{ik}(t) := \int_0^t \alpha_{ik}(s)ds \tag{36}$$

[14]This specification nests some popular models. For example, Ho-Lee model in Heath-Jarrow-Morton specification reads:

$$\sigma(t,T) := \sigma$$

Which can be achieved in Cheyette form by setting:

$$\begin{aligned} \alpha_{11}(t) &:= 1 \\ \beta_{11}(t) &:= \sigma \\ M &:= 1 \\ N_1 &:= 1 \end{aligned}$$

So we obtain:

$$\sigma_1(t,T) = \sum_{i=1}^{1} \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \beta_{11}(t) = \frac{1}{1}\sigma = \sigma$$

[15]Not to confuse with the number of factors M, related to the number of independent Brownian motions in the specification of the instantaneous forward rate dynamics.

***Quadratic variation* processes** [16]

$$V_{ij,k}(t) := \int_0^t \frac{\alpha_{ik}(t)\alpha_{jk}(t)}{\alpha_{ik}(s)\alpha_{jk}(s)}\beta_{ik}(s)\beta_{jk}(s)ds = V_{ji,k}(t) \tag{37}$$

State variables Using the previous definitions, we set:

$$X_{ik}(t) := \int_0^t \frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)d\widetilde{W}_k(s) + \int_0^t \frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\left(\sum_{j=1}^{N_k}\frac{A_{jk}(t)-A_{jk}(s)}{\alpha_{jk}(s)}\beta_{jk}(s)\right)ds \tag{38}$$

Dynamics of the state variables The dynamics under $\mathbb{Q}$ are shown in [Beyna, 2013] to be:[17]

$$dX_{ik}(t) = \left(X_{ik}(t)\frac{\partial}{\partial t}(\log\alpha_{ik}(t)) + \sum_{j=1}^{N_k}V_{ij,k}(t)\right)dt + \beta_{ik}(t)d\widetilde{W}_k(t) \tag{39}$$

These state variables are indeed Markov processes (i.e., they verify (34)) so the model can be expressed as a Markovian system.[18]

We can now determine the form of the fundamental interest rate modeling elements within the model.

Instantaneous forward rate Inserting (35) in (17), and considering the newly defined (36), (37) and (38), the instantaneous forward rate under $\mathbb{Q}$ can be expressed as:

$$f(t,T) = f(0,T) + \sum_{k=1}^{M}\left(\sum_{j=1}^{N_k}\frac{\alpha_{jk}(T)}{\alpha_{jk}(t)}(X_{jk}(t) + \sum_{i=1}^{N_k}\frac{A_{ik}(T)-A_{ik}(t)}{\alpha_{ik(t)}}V_{ij,k}(t))\right) \tag{40}$$

A detailed derivation for our particular model (to be specified in Section 5), based on [Beyna, 2013], is presented in Appendix A.1.

[16] This terminology is used because the quadratic variation processes of $X_{11}(t)$ and $X_{12}(t)$ coincide with $V_{11,1}(t)$ and $V_{11,2}(t)$ respectively.

[17] We include a commented reproduction of the computations in Appendix B due to the importance of this result and the presence of some errata in Beyna's material.

[18] For instance, the short rate is now Markovian as it can be expressed as the sum of the state variables:

$$r(t) = f(0,t) + \sum_{k=1}^{M}\sum_{j=1}^{N_k}X_{jk}(t)$$

Zero-coupon bonds The formula for bonds can be explicitly computed (as shown in [Cheyette, 1994] and [Beyna, 2013]) using (6) and (40). After some calculations we get:[19]

$$\begin{aligned}
P(t,T) &= \exp\left(-\int_t^T f(t,u)du\right) \\
= \quad & \exp\left(-\int_t^T f(0,u) + \sum_{k=1}^{M}\left(\sum_{j=1}^{N_k}\frac{\alpha_{jk}(u)}{\alpha_{jk}(t)}(X_{jk}(t) + \sum_{i=1}^{N_k}\frac{A_{ik}(u)-A_{ik}(t)}{\alpha_{ik(t)}}V_{ij,k}(t)\right)du\right) \\
= \quad & \frac{P(0,T)}{P(0,t)}\exp\Bigg(-\sum_{k=1}^{M}\sum_{j=1}^{N_k}\frac{A_{jk}(T)-A_{jk}(t)}{\alpha_{jk(t)}}X_{jk}(t) \\
- \quad & \sum_{k=1}^{M}\sum_{i=1}^{N_k}\sum_{j=1}^{N_k}\frac{[A_{jk}(T)-A_{jk}(t)][A_{ik}(T)-A_{ik}(t)]}{2\alpha_{ik}(t)\alpha_{jk}(t)}V_{ij,k}(t)\Bigg) \qquad (41)
\end{aligned}$$

Equation (41) provides the link between state variables and bond prices, for this reason, it will be extensively used in Section 7.

[19] The details justifying second equality, following [Beyna, 2013], can be found for our model in Appendix A.2.

5 Two-factor Cheyette model

The particular model within the Cheyette-Beyna environment that will be analyzed in the rest of the Thesis is defined here. We try to find an specification simple enough so the computations in the last Section are straightforward but that still maintains a reasonable degree of flexibility. In his work, [Beyna, 2013] proposes a Three-factor version based on exponential volatility functions that turns out to be too cumbersome in calculations. After some trials, a simplified version of the HJM G2++ of [Acar, 2009] is chosen, where the exponential function of the first factor is replaced with a constant.

5.1 Specification

The model includes a second factor ($M := 2$) independent of the first. Then, we may specify the following in Cheyette-Beyna form, with $N_1 := 1$, $N_2 := 1$:

$$\alpha_{11}(t) := 1 \tag{42a}$$
$$\beta_{11}(t) := a \tag{42b}$$
$$\alpha_{12}(t) := \exp(-\theta t) \tag{42c}$$
$$\beta_{12}(t) := c \tag{42d}$$

The model involves two factors and also two state variables. The volatility functions turn into:

$$\sigma_1(t,T) = \sum_{i=1}^{1} \frac{\alpha_{i1}(T)}{\alpha_{i1}(t)} \beta_{i1}(t) = \frac{1}{1} a = a \tag{43a}$$
$$\sigma_2(t,T) = \sum_{i=1}^{1} \frac{\alpha_{i2}(T)}{\alpha_{i2}(t)} \beta_{i2}(t) = \frac{\exp(-\theta T)}{\exp(-\theta t)} c = \exp[-\theta(T-t)] c \tag{43b}$$

The parameters should verify the constraint imposed by the required positivity of the volatility functions, i.e. $\sigma_k(t,T) > 0 \;\; k = 1,2$. In this case, our model contains three parameters: $a, c \in \mathbb{R}^+$ and $\theta \in \mathbb{R}$.

Now we have to evaluate (36), (37) and (38), with our particular choice of volatility functions defined in (42a) through (42d).

First obtaining the auxiliary processes:

$$\begin{aligned} A_{11}(t) &= \int_0^t 1ds = t \\ A_{12}(t) &= \int_0^t \exp(-\theta s)ds = -\frac{1}{\theta}\left[\exp(-\theta t) - 1\right] \end{aligned}$$

Then, evaluating $V_{ij,k}(t)$:

$$\begin{aligned} V_{11,1}(t) &= \int_0^t a^2 ds = a^2 t \\ V_{11,2}(t) &= \int_0^t \frac{\exp(-\theta t)\exp(-\theta t)}{\exp(-\theta s)\exp(-\theta s)} c^2 ds = \frac{c^2}{2\theta}\left[1 - \exp(-2\theta t)\right] \end{aligned}$$

And also computing the state variables under $\mathbb{Q}$:

$$\begin{aligned} X_{11}(t) &= \int_0^t a d\widetilde{W}_1(s) + \int_0^t a\left[(t-s)a\right] ds = \int_0^t a d\widetilde{W}_1(s) + \frac{1}{2}a^2 t^2 \qquad (44) \\ X_{12}(t) &= \int_0^t \frac{\exp(-\theta t)}{\exp(-\theta s)} c\, d\widetilde{W}_2(s) \\ &+ \int_0^t \frac{\exp(-\theta t)}{\exp(-\theta s)} c \left(\frac{-\frac{1}{\theta}\left[\exp(-\theta t) - 1\right] + \frac{1}{\theta}\left[\exp(-\theta s) - 1\right]}{\exp(-\theta s)} c \right) ds \\ &= c\exp(-\theta t) \int_0^t \exp(\theta s) d\widetilde{W}_2(s) \\ &+ c^2 \exp(-\theta t) \frac{1}{\theta} \int_0^t \exp(2\theta s)\left[\exp(-\theta s) - \exp(-\theta t)\right] ds \\ &= c\exp(-\theta t) \int_0^t \exp(\theta s) d\widetilde{W}_2(s) + c^2 \frac{1}{\theta^2}\left[\frac{1}{2}\left[\exp(-2\theta t) - 1\right] + 1 - \exp(-\theta t)\right] \\ &= c\exp(-\theta t) \int_0^t \exp(\theta s) d\widetilde{W}_2(s) + c^2 \left(\frac{1}{2\theta^2} - \frac{\exp(-2t\theta)\left[2\exp(t\theta) - 1\right]}{2\theta^2} \right) \end{aligned}$$

Having in mind the instruments studied in Section 3, we are most interested in the dynamics of the state variables under the T_1-forward measure. To that regard, we first compute the bond price volatility (see (19)) for each factor in order to define the measure $\widehat{\mathbb{P}}^{T_1}$.

Bond price volatilities

$$\Sigma_1(t, T_1) = -\int_t^{T_1} \sigma_1(t, u) du = -\int_t^{T_1} a\ du = -a(T_1 - t) \qquad (45)$$

$$\Sigma_2(t,T_1) = -\int_t^{T_1} \sigma_2(t,u)du = -\int_t^{T_1} \exp\left[-\theta(u-t)\right]c\,du = \frac{c}{\theta}[\exp\left(-\theta(T_1-t)\right)-1] \quad (46)$$

We apply Multidimensional Girsanov's theorem to define an equivalent probability measure $\widehat{\mathbb{P}}^{T_1}$:

$$\begin{aligned} d\widehat{W_1}(t) &= d\widetilde{W}_1(t) - \Sigma_1(t,T_1)dt = d\widetilde{W}_1(t) + a(T_1-t)dt \\ d\widehat{W_2}(t) &= d\widetilde{W}_2(t) - \Sigma_2(t,T_1)dt = d\widetilde{W}_2(t) - \frac{c}{\theta}[\exp\left(-\theta(T_1-t)\right)-1] \quad (47) \end{aligned}$$

where $\widehat{W_1}(t), \widehat{W_2}(t)$ are independent Brownian motions under $\widehat{\mathbb{P}}^{T_1}$.

Applying this change of measure we can now compute the state variables under $\widehat{\mathbb{P}}^{T_1}$. It will also be useful for later calculations to evaluate both state variables at $t = T_1$ and determine its distribution.

State variables under $\widehat{\mathbb{P}}^{T_1}$

$$\begin{aligned} X_{11}(t) &= \int_0^t a[dW_1^{\widehat{\mathbb{P}}^{T_1}}(s) - a(T_1-s)ds] + \frac{1}{2}a^2t^2 \\ &= \int_0^t a\ dW_1^{\widehat{\mathbb{P}}^{T_1}}(s) - \int_0^t a^2(T_1-s)ds + \frac{1}{2}a^2t^2 \\ &= \int_0^t a\ dW_1^{\widehat{\mathbb{P}}^{T_1}}(s) - a^2T_1t + a^2t^2 \end{aligned}$$

At $t = T_1$,

$$X_{11}(T_1) = \int_0^{T_1} a\ dW_1^{\widehat{\mathbb{P}}^{T_1}}(s)$$

$$\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[X_{11}(T_1)\right] = 0$$

Applying Itô's isometry:

$$\mathbb{V}_{\widehat{\mathbb{P}}^{T_1}}\left[X_{11}(T_1)\right] = a^2T_1$$

Due to the normal distribution of stochastic integrals of deterministic integrands with respect to Brownian motion:

$$X_{11}(T_1) \overset{\widehat{\mathbb{P}}^{T_1}}{\sim} \mathcal{N}(0, a^2T_1)$$

$$\begin{aligned}
X_{12}(t) &= c\exp(-\theta t)\int_0^t \exp(\theta s) dW_2^{\widehat{\mathbb{P}}^T}(s) \\
&+ \frac{c^2}{\theta^2}\exp(-\theta t)\bigg(\frac{1}{2}\exp(-\theta T_1)\Big(\exp(2\theta t)-1\Big)-\exp(\theta t)+1\bigg) \\
&+ c^2\left(\frac{1}{2\theta^2}-\frac{\exp(-2t\theta)[2\exp(t\theta)-1]}{2\theta^2}\right)
\end{aligned}$$

At $t = T_1$,

$$X_{12}(T_1) = c\exp(-\theta T_1)\int_0^{T_1} \exp(\theta s) dW_2^{\widehat{\mathbb{P}}^T}(s)$$

$$\begin{aligned}
\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}[X_{12}(T_1)] &= 0 \\
\mathbb{V}_{\widehat{\mathbb{P}}^{T_1}}[X_{12}(T_1)] &= \frac{c^2}{2\theta}\Big[1-\exp(-2\theta T_1)\Big]
\end{aligned}$$

$$X_{12}(T_1) \overset{\widehat{\mathbb{P}}^{T_1}}{\sim} \mathcal{N}\bigg(0, \frac{c^2}{2\theta}\Big[1-\exp(-2\theta T_1)\Big]\bigg)$$

We compute the covariance between the state variables, applying the polarization identity:

$$\begin{aligned}
\mathbb{C}ov_{\widehat{\mathbb{P}}^{T_1}}[X_{11}(T_1), X_{12}(T_1)] &= \mathbb{C}ov_{\widehat{\mathbb{P}}^{T_1}}\left[\int_0^{T_1} a\ dW_1^{\widehat{\mathbb{P}}^{T_1}}(s)\ ,\ \int_0^{T_1} c\exp(-\theta T_1)\exp(\theta s) dW_2^{\widehat{\mathbb{P}}^T}(s)\right] \\
&= \mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\int_0^{T_1} a\ dW_1^{\widehat{\mathbb{P}}^{T_1}}(s)\cdot\int_0^{T_1} c\exp(-\theta T_1)\exp(\theta s) dW_2^{\widehat{\mathbb{P}}^T}(s)\right] \\
&= \mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\int_0^{T_1} a\ c\exp(-\theta T_1)\exp(\theta s)\ d\langle W_1^{\widehat{\mathbb{P}}^{T_1}}, W_2^{\widehat{\mathbb{P}}^T}\rangle(s)\right] \\
&= 0
\end{aligned}$$

Because the driving Brownian motions are independent and the integrand is deterministic, they are jointly normal. As the covariance between them is zero, we conclude that the two state variables are independent.

We can also compute their $\widehat{\mathbb{P}}^{T_1}$-dynamics, combining (39) with (47):

Dynamics of the state variables under $\widehat{\mathbb{P}}^{T_1}$

$$\begin{aligned} dX_{ik}(t) &= \left(X_{ik}(t)\frac{\partial}{\partial t}(\log \alpha_{ik}(t)) + \sum_{j=1}^{N_k} V_{ij,k}(t) \right) dt \;+\; \beta_{ik}(t)\left(d\widehat{W}_k^T(t) + \Sigma(t,T)dt \right) \\ &= \left(X_{ik}(t)\frac{\partial}{\partial t}(\log \alpha_{ik}(t)) + \sum_{j=1}^{N_k} V_{ij,k}(t) + \beta_{ik}(t)\Sigma(t,T) \right) dt \\ &+ \beta_{ik}(t) d\widehat{W}_k^T(t) \end{aligned}$$

In our model, they become:

$$\begin{aligned} dX_{11}(t) &= a^2\,(2t - T_1)\,dt \;+\; a\; dW_1^{\widehat{\mathbb{P}}^{T_1}}(t) \\ dX_{12}(t) &= \left(-\theta X_{12}(t) + \frac{c^2}{2\theta}\Big[1 - \exp\left(-2\theta t\right)\Big] + \frac{c^2}{\theta}\Big[\exp\left(-\theta(T_1 - t)\right) - 1\Big] \right) dt \\ &+ c dW_2^{\widehat{\mathbb{P}}^{T_1}}(t) \end{aligned} \tag{48}$$

The zero-coupon bond price will be the reference quantity used to value the instruments in Section 7.

Zero-coupon bond price Using (41), we can readily compute the expression for bond prices in our model:

$$\begin{aligned} P(t,T) &= \frac{P(0,T)}{P(0,t)} \exp\Bigg[-\left(\frac{A_{11}(T) - A_{11}(t)}{\alpha_{11}(t)} X_{11}(t) + \frac{A_{12}(T) - A_{12}(t)}{\alpha_{12}(t)} X_{12}(t) \right) \\ &- \Bigg(\frac{[A_{11}(T) - A_{11}(t)][A_{11}(T) - A_{11}(t)]}{2\alpha_{11}(t)\alpha_{11}(t)} V_{11,1}(t) \\ &+ \frac{[A_{12}(T) - A_{12}(t)][A_{12}(T) - A_{12}(t)]}{2\alpha_{12}(t)\alpha_{12}(t)} V_{11,2}(t) \Bigg) \Bigg] \\ &= \frac{P(0,T)}{P(0,t)} \exp\Bigg[-(T-t)X_{11}(t) - \frac{1}{\theta}\Big(1 - \exp[-\theta(T-t)] \Big) X_{12}(t) \\ &- \frac{1}{2}(T-t)^2 a^2 t \;-\; \frac{1}{2\theta^2}\Big(1 - \exp[-\theta(T-t)] \Big)^2 \frac{c^2}{2\theta}\Big[1 - \exp\left(-2\theta t\right) \Big] \Bigg] \end{aligned} \tag{49}$$

Note that given:

- the zero coupon initial curve, $T \mapsto P(0,T)$
- values for the parameters a, c, θ

the zero-coupon bond t price for any maturity T, is a function $P(t, T, X_{11}(t), X_{12}(t))$. The exponent happens to be the sum of independent normal random variables, therefore we conclude that the bond price is lognormally distributed.

5.2 Caplet analytical formula

Being in a *gaussian* Heath-Jarrow-Morton setting, we can obtain a closed-form expression for caplets [20] based on the lognormal distribution of bond prices.

We start by stating the following useful result for the valuation of call options, found for example in [Cairns, 2004]:

Expectation of maximum of lognormal random variable minus a positive constant. If X has a lognormal distribution under a probability measure $\mathbb{P}$, that is, $\log(X) \overset{\mathbb{P}}{\sim} \mathcal{N}\left(e, d^2\right)$, then for any constant $K > 0$:

$$\mathbb{E}_{\mathbb{P}}\left[(X-K)^+\right] = \mathbb{E}_{\mathbb{P}}[X]\Phi\left(\frac{e+d^2-\log K}{d}\right) - K\Phi\left(\frac{e-\log K}{d}\right) \tag{50}$$

where $\Phi(x)$ denotes the standard normal cumulative density function evaluated at x, and $e := \mathbb{E}_{\mathbb{P}}[\log X]$, $d^2 := \mathbb{V}_{\mathbb{P}}[\log X]$.

For our purposes, $X := P(T_1, T_2)$ and $e = \mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}[\log P(T_1, T_2)]$, $d^2 = \mathbb{V}_{\widehat{\mathbb{P}}^{T_1}}[\log P(T_1, T_2)]$. We first compute the variance, using the previously computed moments of the state variables.

[20] Computing unconditional expectation is enough for the purpose of valuation of claims at $t = 0$ (which is what we will be doing in Section 7), and it is easier than the more general moments conditional on $\mathcal{F}_t$.

$$\mathbb{V}_{\widehat{\mathbb{P}}^{T_1}}[\log P(T_1,T_2)] = \mathbb{V}_{\widehat{\mathbb{P}}^{T_1}}\left[-\left(\frac{A_{11}(T_2)-A_{11}(T_1)}{\alpha_{11}(T_1)}X_{11}(T_1)+\frac{A_{12}(T_2)-A_{12}(T_1)}{\alpha_{12}(T_1)}X_{12}(T_1)\right)\right]$$
$$= \left(\frac{A_{11}(T_2)-A_{11}(T_1)}{\alpha_{11}(T_1)}\right)^2\mathbb{V}_{\widehat{\mathbb{P}}^{T_1}}[X_{11}(T_1)]+\left(\frac{A_{12}(T_2)-A_{12}(T_1)}{\alpha_{12}(T_1)}\right)^2\mathbb{V}_{\widehat{\mathbb{P}}^{T_1}}[X_{12}(T_1)]$$
$$+2\left(\frac{A_{11}(T_2)-A_{11}(T_1)}{\alpha_{11}(T_1)}\right)\left(\frac{A_{12}(T_2)-A_{12}(T_1)}{\alpha_{12}(T_1)}\right)\mathbb{C}ov_{\widehat{\mathbb{P}}^{T_1}}[X_{11}(T_1),X_{12}(T_1)]$$
$$= (T_2-T_1)^2a^2T_1+\left(\frac{1-\exp\left(-\theta(T_2-T_1)\right)}{\theta}\right)^2\frac{c^2}{2\theta}\Big[1-\exp\left(-2\theta T_1\right)\Big] \quad (51)$$

A result in [Cairns, 2004] linking expectation and variance in this setting allows us to quickly compute the expectation as:

$$\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}[\log P(T_1,T_2)|] = \log\frac{P(0,T_2)}{P(0,T_1)}+\frac{1}{2}\mathbb{V}_{\widehat{\mathbb{P}}^{T_1}}[\log P(T_1,T_2)|] \quad (52)$$

The pull-call parity for European bond options states that:

$$\mathbf{Call_{bond}}(t,T_1,T_2,K)-P(t,T_2)+KP(t,T_1) = \mathbf{Put_{bond}}(t,T_1,T_2,K) \quad (53)$$

Knowing the relationship (53), we can restate the value of the caplet previously determined in (29), in terms of the price of a call bond option so we can apply (50):

$$\begin{aligned}\mathbf{V_{caplet}}(0) &= \frac{1}{K'}P(0,T_1)\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\left[K'-P(T_1,T_2)\right]^+\right] \qquad (54)\\ &= \frac{1}{K'}\left(P(0,T_1)\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\left[P(T_1,T_2)-K'\right]^+\right]-P(0,T_2)+K'P(0,T_1)\right)\end{aligned}$$

Noting that:

$$\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[P(T_1,T_2)\right] = \mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\frac{P(T_1,T_2)}{P(T_1,T_1)}\right] = \frac{P(0,T_2)}{P(0,T_1)}$$

The expectation in (54) then becomes:

$$\begin{aligned}
\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\left[P(T_1,T_2)-K'\right]^+\right] &= \mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[P(T_1,T_2)\right]\Phi\left(\frac{e+d^2-\log K'}{d}\right)-K'\Phi\left(\frac{e-\log K'}{d}\right)\\
&= \frac{P(0,T_2)}{P(0,T_1)}\Phi\left(\frac{e+d^2-\log K'}{d}\right)-K'\Phi\left(\frac{e-\log K'}{d}\right) \qquad (55)
\end{aligned}$$

Putting all the pieces together, we arrive at:

$$\begin{aligned}
\mathbf{V_{caplet}}(0) &= \frac{1}{K'}\Bigg(P(0,T_1)\Bigg[\frac{P(0,T_2)}{P(0,T_1)}\Phi\left(\frac{e+d^2-\log K'}{d}\right)\\
&- K'\Phi\left(\frac{e-\log K'}{d}\right)\Bigg]-P(0,T_2)+K'P(0,T_1)\Bigg)\\
&= \frac{1}{K'}\Bigg(P(0,T_2)\Phi\left(\frac{e+d^2-\log K'}{d}\right)\\
&- K'P(0,T_1)\Phi\left(\frac{e-\log K'}{d}\right)-P(0,T_2)+K'P(0,T_1)\Bigg) \qquad (56)
\end{aligned}$$

where $e=\mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}[\log P(T_1,T_2)]$ and $d^2=\mathbb{V}_{\widehat{\mathbb{P}}^{T_1}}[\log P(T_1,T_2)]$ have been computed in (51) and (52).

The analytical formula obtained is convenient for the calibration procedure that will be explained in Section 6 and for the implementation of a control variate estimator in Section 7.

6 Calibration. *Simulated annealing*

The process of adjusting model parameters to market reality is known as calibration. There exist several approaches to it and usually only a fraction of the market is considered. In order to calibrate the proposed model to market transactions, we choose a somewhat simplified approach where only OIS rates and a set of EUR caps with a particular strike will be taken into account.[21] We consider a set of market caplet prices[22] with strike $k = 0.005$, settle dates $\{0.25, 0.5, 0.75, ..., 28.5, 29, 29.5\}$ and maturities $\{0.5, 0.75, 1, ..., 29, 29.5, 30\}$. We have a total of 63 prices that can be arranged in a vector $capletsMKT_{63\times1}$.

The process is meant to find parameters that minimize the sum of squared errors, SSE, between market and model prices[23] under said parameters.

The minimization procedure will be achieved using a *simulated annealing* routine. It is a derivative-free optimization method whose adequacy for models of this class is hinted at in [Beyna, 2013]. Given an initial point, the general structure involves searching randomly for neighbouring points which offer a lower value of the objective function, while also introducing a probability (determined by the difference in the objective function and a parameter called temperature) of going to a *worse* point, which gives the procedure the ability of not getting stuck at a local minimum.

The main elements of the procedure are:

- Objective function f (SSE in our case). The change in the value of f, when evaluated in another point, is noted as Λ.
- Acceptance function, which will define the probability of accepting a *worse* point, that is, a point which increases the value of the objective function when compared to the previous point ($\Lambda > 0$).

[21] We take a simplified route calibrating the models to caps of a particular strike, instead of a more complex approach using several strikes of caps and including swaptions.

[22] The procedure to obtain such prices and the particularities of the EUR cap market are treated in Appendix C.

[23] Some authors [Beyna, 2013] recommend calibration to implied volatilities, as they are not as affected by maturity and strike effects, but for the sake of brevity we have not taken that path. It is also worth noting that ATM volatility is usually the preferred choice due to its liquidity but we did not recover it in the *stripping* process.

A popular choice is the negative exponential function $\exp\left(\frac{-\Lambda}{Temp}\right)$,[24] which assigns greater probability when the temperature is high and/or the change in the objective function is small.

- Initial and final temperature, T_0 and T_{min}
- Temperature reduction parameter, λ
- *Neighbour* selection, in our case through a $\mathcal{N}(0,1)$ extraction, so values around 0 are more likely than extreme ones, without completely excluding the possibility of large numbers being obtained. A scale factor allows to adjust the variance of the normal random extractions so it can better fit the magnitude of the parameters.

The implementation [25] used here is inspired on [Press et al., 2007]. The annealing schedule will be reset M times, and the search of new points in each temperature is repeated N times, storing the best point found so far. Also, bounds for the values of the parameters are incorporated to make sure they verify $a, c \in \mathbb{R}^+$.

Defining a function $caplets2Fac(a, c, \theta)$, that, taking the model parameters as input, returns a vector$_{63\times 1}$ with the Cheyette-Two-Factor model prices for caplets with the same set of settle dates, maturities and strike as the ones in the market (using (56)), we can set the following procedure, detailed in pseudocode in Algorithm 1.

After some experimentation, taking $Temp_0 = 0.01$, $p_0 = (\underbrace{0.35}_{a}\ \underbrace{0.25}_{c}\ \underbrace{0.05}_{\theta})$, $\lambda = 0.95$, M=5, N=50, $Temp_{min} = 0.0001$, scale $= 0.0001$ and parameter bounds $(0,\infty)$,$(0,\infty)$,$(-\infty,\infty)$, yields the results:

$$a_{opt} = 0.506898, \quad c_{opt} = 0.083819, \quad \theta_{opt} = 0.104966$$

$$SSE(a_{opt}, c_{opt}, \theta_{opt}) = 0.01495$$

The goodness of fit of the results can be seen in Figure 1, comparing market and model prices[26]:

[24] It is indeed a negative exponential function because $Temp > 0$, and $\Lambda > 0$ when the acceptance function is relevant. This also implies that its resulting value is never greater than 1, so it can be rightfully interpreted as a probability.

[25] The random numbers required here and in the Valuation section have been obtained using *Matlab 2019a* implementation of the *Mersenne Twister* algorithm for $\mathcal{U}(0,1)$ variates and the *Ziggurat* algorithm for obtaining $\mathcal{N}(0,1)$ distributed numbers.

[26] An alternate calibration output was found where shorter maturities were nicely adjusted at the expense of a high error in longer maturities. Although it might have been an alternative in view of the instruments of Section 7, it was discarded due to the high resulting output of the SSE function.

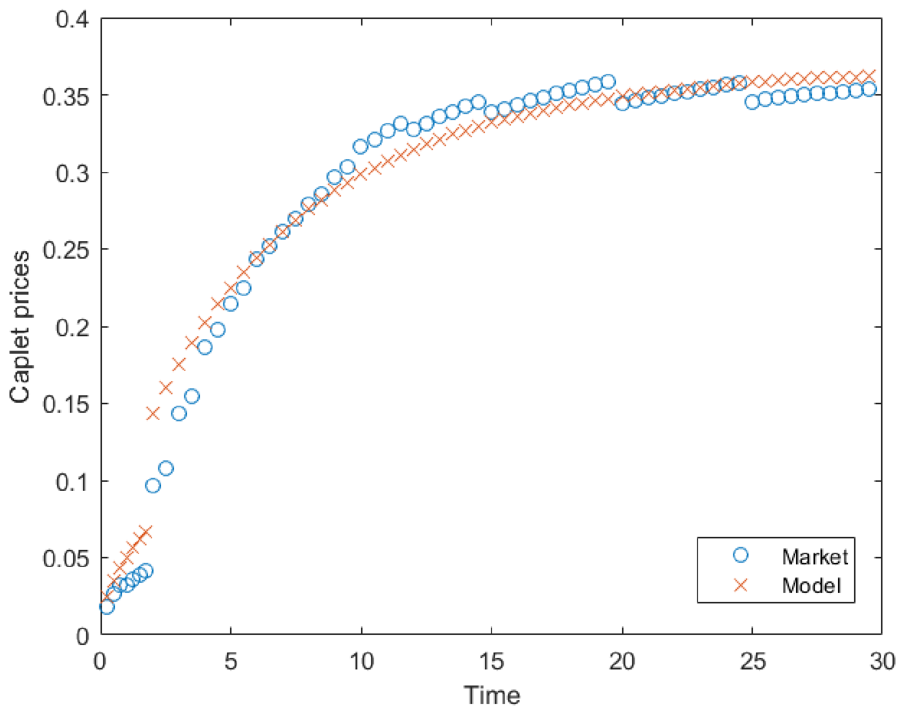

Figure 1: Differences between caplet market prices and caplet prices in the Two-Factor Cheyette model with parameters $a = 0.506898$, $c = 0.083819$, $\theta = 0.104966$.

The values found for the parameters allow us to define the volatility functions:

$$\begin{aligned} \sigma_1(t,T) &:= 0.506898 \\ \sigma_2(t,T) &:= \exp\Big(-0.104966(T-t)\Big)0.083819 \end{aligned}$$

Model specification also requires $f(0,T)$, which is obtained through the splines that describe the yield curve based on market data and relations (1) and (5)[27].

In the next section we assume that model parameters $\{a, c, \theta\}$ as well as $f(0,T)$ are known, so the volatility functions $\sigma_1(t,T)$, $\sigma_2(t,T)$ are explicitly defined and bond prices at $t = 0$ are known for every possible maturity.

[27]See Appendix C for details, where the function $T \to P(0,T)$ is computed using cubic splines and relation (1). Considering (5), it is equivalent to defining the initial instantaneous forward curve $f(0,T)$.

Algorithm 1 Minimization of sum of squared model errors respect to caplet market prices using Simulated Annealing

Require: $capletsMKT,\ caplets2Fac(a, c, \theta),\ Temp_0,\ p_0,\ \alpha,\ M,\ N,\ Temp_{min},\ scale$

Ensure: $a_{opt}, c_{opt}, \theta_{opt}, f(p_{opt})$

$f(p) = (caplets_{MKT} - caplets_{2Fac}(p))' \cdot (caplets_{MKT} - caplets_{2Fac}(p))$ ▷ SSE function
$p_{best} = p_0$
for $m = 1$ to M **do** ▷ M resets of the annealing schedule
  $Temp = Temp_0$
  $p_{old} = p_{best}$
  while $Temp > Temp_{min}$ **do**
    for $n = 1$ to N **do** ▷ N explorations at each temperature
      $element$ = Select at random an element of the vector
      $p_{old}[element] = p_{old}[element] + scale * random_{\mathcal{N}(0,1)}$
      if $p_{old}[1] < 0$ **then** ▷ Ensure $a, c \in \mathbb{R}^+$
        $p_{old}[1] = |random_{\mathcal{N}(0,1)}|$
      else if $p_{old}[2] < 0$ **then**
        $p_{old}[2] = |random_{\mathcal{N}(0,1)}|$
      end if
      $p_{new} = p_{old}$
      $\Lambda = f(p_{new}) - f(p_{old})$
      if $\Lambda < 0$ **then** ▷ Accept better point
        $p_{old} = p_{new}$
      else ▷ Accept worse point with probability $prob$
        $prob = \exp\left(\dfrac{-\Delta}{Temp}\right)$
        if $random_{\mathcal{U}(0,1)} < prob$ **then**
          $p_{old} = p_{new}$
        end if
      end if
      if $f(p_{new}) < f(p_{best})$ **then** ▷ Store best point found
        $p_{best} = p_{new}$
      end if
    end for n
    $Temp = \lambda * Temp$ ▷ Decrease temperature
  end while
end for m
$p_{opt} = p_{best};\ a_{opt} = p_{opt}[1];\ c_{opt} = p_{opt}[2];\ \theta_{opt} = p_{opt}[3];\ f(p_{opt})$
return $a_{opt}, c_{opt}, \theta_{opt}, f(p_{opt})$

7 Valuation of interest rate derivatives: Numerical methods

In this Section, we choose specific examples of the derivatives analyzed in Section 3, as a means to explain and test several numerical methods. A suitable method will be used for each derivative and its $t = 0$ value will be calculated.

The chosen instruments are:

- Barrier down-and-out caplet over the 3-month simple spot rate (i.e. $L(t, t + \delta)$ with $\delta := 0.25$), settled at $T_1 := 1$, maturing at $T_2 = 1.25$, with strike $k = 0.005$ and barrier set at $B = -0.15$.
- European swaption, maturing at $T_1 := 1$ for entering a receiver swap at $K = 0.005$ with reset dates $\{1, 1.25, 1.5, 1.75\}$ and payment dates $\{1.25, 1.5, 1.75, 2\}$.
- Bermudan swaption, for entering a receiver swap at $K = 0.005$ with the same tenor structure as the previous one and possible exercise dates $\{0.25, 0.5, 0.75, 1\}$.

7.1 Monte Carlo path simulation with control variate. Barrier caplet

In order to simulate paths of the underlying, we implement a first order discretization (Euler scheme). Because of the functional form of $\beta_{ik}(t)$,the diffusion term never depends on $X(t)$, so the more precise Milstein scheme collapses into Euler scheme in the Cheyette-Beyna specification we are working with. The main advantage of discretization schemes is their generality, although they introduce discretization error, so in the case of the *vanilla* caplet, we will test the results against known caplet analytical values to check their adequacy.

We will partition the interval $[0, T_1]$ in $N_{int} = 1000$ subintervals. Each one will have the same length, namely $\frac{T_1}{N_{int}}$, and $N_{int} + 1$ time points will be defined $0 = t_0 < t_1 < ... < t_{N_{int}-1} < t_{N_{int}} = T_1$.

The scheme for the two state variables, given the dynamics (48), is then:

$$\begin{aligned}
X_{11}(t_0) &= X_{11}(0) = 0 \\
X_{11}(t_{i+1}) &= X_{11}(t_i) + a^2\left(2t_i - T_1\right)\Delta t \;+\; a\sqrt{\Delta t}Z^1_{i+1} \\
X_{12}(t_0) &= X_{12}(0) = 0 \\
X_{12}(t_{i+1}) &= X_{12}(t_i) + \Bigg(-\theta X_{12}(t_i) + \frac{c^2}{2\theta}\Big[1 - \exp\left(-2\theta t_i\right)\Big] \\
&+ \frac{c^2}{\theta}\Big[\exp\left(-\theta(T_1 - t_i)\right) - 1\Big]\Bigg)\Delta t \;+\; c\sqrt{\Delta t}Z^2_{i+1}
\end{aligned} \tag{57}$$

for $i = 0, 1, ..., N_{int} - 1$, where $Z^1_i \overset{iid}{\sim} \mathcal{N}(0,1)$, $Z^2_i \overset{iid}{\sim} \mathcal{N}(0,1)$ are also independent with respect to each other.

***Vanilla* caplet** As a reference, we compute the value of a caplet with the same characteristics as the barrier caplet except for its barrier feature. Recalling (29), we write:

$$\mathbf{V_{caplet}}(0) = \mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\frac{1}{K'}P(0,T_1)\left[K' - P(T_1,T_2)\right]^+\right] \tag{58}$$

Applying the discretization scheme in (57) we are able to obtain realizations of the state variables at T_1. The payoff is a function of the bond price $P(T_1, T_2)$, which is itself a function of X_{11} and X_{12}, and therefore it can be computed for each realization using (49). Then we can estimate the expectation in (58) by calculating their sample mean,[28] obtaining the results shown in Table 1.

[28] We are considering the sample mean of the n independent simulations as the estimator of the expectation. By the Central Limit Theorem, the error of the estimation is normally distributed $\hat{V}_n - V \sim \mathcal{N}\left(0, \frac{\sigma_V^2}{n}\right)$, so its standard deviation is $\frac{\sigma_V}{\sqrt{n}}$. Substituting σ_V for the sample standard deviation $s_V = \sqrt{\frac{1}{n-1}\sum_{i=1}^n (V_i - \hat{V}_n)^2}$, allows us to compute the **Error** shown in the Table.

	Monte Carlo Simulation			**Analytical**
	$n = 100$	$n = 1000$	$n = 10000$	
Value	0.0535	0.0498	0.0508	0.0504
s_V	0.0771	0.0699	0.0645	-
Error ($\frac{s_V}{\sqrt{n}}$)	0.0077	0.0022	0.0006	-

Table 1: Monte Carlo simulations ($n = 100, 1000, 10000$) and analytical formula comparison for a *vanilla* caplet with $k = 0.005$, $T_1 = 1$, $T_2 = 1.25$.

Barrier caplet Turning our attention to the barrier caplet, it is possible to restate (30) in a more convenient way in terms of bond prices employing (2):

$$\tau_B = \inf\left\{t_i \mid \frac{1}{B\delta + 1} < P(t_i, t_i + \delta)\right\} \tag{59}$$

And the indicator function as:

$$\mathbb{1}\{\tau_B > T_1\} = \mathbb{1}\left\{\frac{1}{B\delta + 1} > P(t, t + \delta)\ \forall\ 0 \leq t \leq T_1\right\}$$

We can compute the value of the indicator function for each path by simulating bond prices at each t with maturity at $t + \delta$ and then determine τ_B. This doesn't pose any problem as, given a realization of X_{11} and X_{12} for a time t, we can obtain t bond prices for any maturity using (49).

The value of the caplet barrier is then formulated, based on (31), as :

$$\mathbf{V_{barriercaplet}}(0) =$$
$$= \mathbb{E}_{\widehat{\mathbb{P}}^{T_1}}\left[\frac{1}{K'}P(0, T_1)\left[K' - P(T_1, T_2)\right]^+ \mathbb{1}\left\{\frac{1}{B\delta + 1} > P(t, t + \delta),\quad \forall\ 0 \leq t \leq T_1\right\}\right]$$

which can be interpreted as constant times an up-and-out put on a bond with strike K.

Considering time discretization of the interval $[0, T_1]$ and performing n independent simulations we can compute the following approximation:[29]

$$\frac{1}{n}\sum_{j=1}^{n}\left(\frac{1}{K'}P(0,T_1)[K'-P_j(T_1,T_2)]^+\mathbb{1}\left\{\frac{1}{B\delta+1} > P_j(t_k,t_k+\delta),\ 0 = t_0 < t_1 < ... < t_k < ...\ T_1\right\}\right) \tag{60}$$

[29]We are considering a continuously monitored barrier, but due to the discretization of the time interval $[0, T_1]$, some error is introduced as we are not checking the barrier at every instant.

Obtaining the results appearing in Table 2:

Monte Carlo simulation			
	n=100	n=1000	**n=5000**
Value	0.03111	0.02616	**0.02585**
s_V	0.06274	0.05997	**0.05996**
Error ($\frac{s_V}{\sqrt{n}}$)	0.00627	0.00190	**0.00085**

Table 2: Monte Carlo simulations ($n = 100, 1000, 5000$ for the barrier down-and-out caplet over $L(t, t+\delta)$, with $\delta = 0.25$; settled at $T_1 = 1$, maturing at $T_2 = 1.25$, with strike $k = 0.005$ and $B = -0.15$.

Control variate In order to reduce estimation error, we can increase the number of simulations (aiming on increasing $\sqrt{n}$ so $\frac{1}{\sqrt{n}}$ decreases), or try to lower standard deviation s_V. A method of reducing standard deviation when an analytical expression is available for a related[30] derivative (such as the *vanilla* caplet in our case), is the control variate estimator. The control variate estimator for our case is defined as:

$$PO_{control} = PO_{barrier} - \frac{Cov(PO_{vanilla}, PO_{barrier})}{\mathbb{V}[PO_{vanilla}]} \Big(PO_{vanilla} - \mathbb{E}[PO_{vanilla}] \Big) \tag{61}$$

Then, instead of computing the expectation of the payoff for the barrier option, we calculate the expectation of the newly defined *control payoff* to obtain an estimation of the value of the option.

Summarizing, the steps would be:

- Simulate n paths of both state variables under the T_1-Forward measure $\widehat{\mathbb{P}}^{T_1}$ in the interval $[0, T_1]$ via the Euler scheme in (57).
- Compute bond prices, using (49), at every t_i in $[0, T_1]$ with maturity $t_i + \delta$ for each path.
- Compute the indicator function (59) for each path (i.e. check if the barrier is hit for each path comparing each bond price at every t_i with $\frac{1}{B\delta+1}$).

[30] In this case relatedness refers to correlation between the payoffs of both instruments

- Compute the payoff for both vanilla (using the last bond price of each path, that is $P(T_1, T_1+\delta) = P(T_1, T_2)$) and barrier caplets (taking into account the last bond price and the value of the indicator function for each path).

- Define the control payoff as in (61), where $\mathbb{E}[PO_{vanilla}]$ can be analytically computed using (58), but $\mathbb{V}[PO_{vanilla}]$ and $Cov(PO_{vanilla}, PO_{barrier})$ have to be replaced with their sample counterparts.

- With n independent observations of the control payoff, calculate its sample mean and its sample standard deviation in order to obtain an estimation of the value of the barrier caplet and a measure of its error.

[Glasserman, 2003] points out that a $\rho(PO_{vanilla}, PO_{barrier})$ around 0.7 (the sample correlation coefficient between payoffs $\hat{\rho}(PO_{vanilla}, PO_{barrier})$ in our analysis is about 0.71) decreases the number of simulations required to obtain the same error to about half, with respect to the original case without control variate. Following this idea, we try $n = 2500$ and check that this is indeed the case, obtaining approximately the same error without control variate and $n = 5000$, than using control variate and setting $n = 2500$. These findings are summarized in Table 3.

Monte Carlo simulation (control variate)				
	n=100	n=1000	**n=2500**	n=5000
Value	0.02778	0.02510	**0.02546**	0.02521
s_V	0.04749	0.04309	**0.04245**	0.04232
Error $(\frac{s_V}{\sqrt{n}})$	0.00475	0.00136	**0.00084**	0.00060

Table 3: Monte Carlo simulation results for the Barrier down-and-out caplet over $L(t, t+\delta)$, with $\delta = 0.25$; settled at $T_1 = 1$, maturing at $T_2 = 1.25$, with strike $k = 0.005$ and $B = -0.15$ with control variate for $n = 100, 1000, 2500, 5000$.

7.2 Numerical integration. European swaption

For the European swaption, we decide to use numerical integration to compute its value due to the path-independent nature of this derivative. We recall the distribution of the state variables under $\widehat{\mathbb{P}}^{T_1}$ at T_1. As stated in Section 5, the two state variables are jointly normal, so they follow a bivariate normal distribution.

Considering:

$$
\begin{aligned}
X_{11}(T_1) &\overset{\widehat{\mathbb{P}}^{T_1}}{\sim} \mathcal{N}(0, a^2 T_1) \overset{d}{=} a\sqrt{T_1}\ \underbrace{\mathcal{N}(0,1)}_{x} \\
X_{12}(T_1) &\overset{\widehat{\mathbb{P}}^{T_1}}{\sim} \mathcal{N}\left(0, \frac{c^2}{2\theta}\left[1-\exp\left(-2\theta T_1\right)\right]\right) \overset{d}{=} \\
&\overset{d}{=} \frac{c}{\sqrt{2\theta}}\sqrt{1-\exp\left(-2\theta T_1\right)}\ \underbrace{\mathcal{N}(0,1)}_{y}
\end{aligned}
$$

we can work with the bivariate standard normal instead. The joint probability density function under $\widehat{\mathbb{P}}^{T_1}$ of X, Y is therefore:

$$
f^{\widehat{\mathbb{P}}^{T_1}}(x,y) = \frac{1}{2\pi}\exp\left[-\frac{1}{2}\left(x^2+y^2\right)\right]
$$

The payoff for the European swaption (see (32)) in our model can be expressed as a function of x and y, as

$$
\begin{aligned}
g(x,y) &= \left[K\left(\frac{P(0,T_2)}{P(0,T_1)}\exp\left[-(T_2-T_1)a\sqrt{T_1}x\right.\right.\right. \\
&- \frac{1}{\theta}\left(1-\exp[-\theta(T_2-T_1)]\right)\frac{c}{\sqrt{2\theta}}\sqrt{1-\exp\left(-2\theta T_1\right)}y \\
&- \left.\left.\left.\frac{1}{2}(T_2-T_1)^2 a^2 T_1 - \frac{1}{2\theta^2}\left(1-\exp\left[-\theta(T_2-T_1)\right]\right)^2\frac{c^2}{2\theta}\left[1-\exp\left(-2\theta T_1\right)\right]\right]\right]0.25
\end{aligned}
$$

$$
\begin{aligned}
+ & \frac{P(0,T_3)}{P(0,T_1)} \exp \Bigg[- (T_3 - T_1) a \sqrt{T_1} x \\
- & \frac{1}{\theta} \Big(1 - \exp[-\theta(T_3 - T_1)] \Big) \frac{c}{\sqrt{2\theta}} \sqrt{1 - \exp(-2\theta T_1)} y \\
- & \frac{1}{2}(T_3 - T_1)^2 a^2 T_1 - \frac{1}{2\theta^2} \Big(1 - \exp[-\theta(T_3 - T_1)] \Big)^2 \frac{c^2}{2\theta} \Big[1 - \exp(-2\theta T_1) \Big] \Bigg] 0.25 \\
+ & \frac{P(0,T_4)}{P(0,T_1)} \exp \Bigg[- (T_4 - T_1) a \sqrt{T_1} x \\
- & \frac{1}{\theta} \Big(1 - \exp[-\theta(T_4 - T_1)] \Big) \frac{c}{\sqrt{2\theta}} \sqrt{1 - \exp(-2\theta T_1)} y \\
- & \frac{1}{2}(T_4 - T_1)^2 a^2 T_1 - \frac{1}{2\theta^2} \Big(1 - \exp[-\theta(T_4 - T_1)] \Big)^2 \frac{c^2}{2\theta} \Big[1 - \exp(-2\theta T_1) \Big] \Bigg] 0.25 \\
+ & \frac{P(0,T_5)}{P(0,T_1)} \exp \Bigg[- (T_5 - T_1) a \sqrt{T_1} x \\
- & \frac{1}{\theta} \Big(1 - \exp[-\theta(T_5 - T_1)] \Big) \frac{c}{\sqrt{2\theta}} \sqrt{1 - \exp(-2\theta T_1)} y \\
- & \frac{1}{2}(T_5 - T_1)^2 a^2 T_1 - \frac{1}{2\theta^2} \Big(1 - \exp[-\theta(T_5 - T_1)] \Big)^2 \frac{c^2}{2\theta} \Big[1 - \exp(-2\theta T_1) \Big] \Bigg] 0.25 \Bigg) \\
+ & \frac{P(0,T_5)}{P(0,T_1)} \exp \Bigg[- (T_5 - T_1) a \sqrt{T_1} x \\
- & \frac{1}{\theta} \Big(1 - \exp[-\theta(T_5 - T_1)] \Big) \frac{c}{\sqrt{2\theta}} \sqrt{1 - \exp(-2\theta T_1)} y \\
- & \frac{1}{2}(T_5 - T_1)^2 a^2 T_1 - \frac{1}{2\theta^2} \Big(1 - \exp[-\theta(T_5 - T_1)] \Big)^2 \frac{c^2}{2\theta} \Big[1 - \exp(-2\theta T_1) \Big] \Bigg] - 1 \Bigg]^+
\end{aligned}
$$

The value of the European swaption (see (33)) defined at the beginning of Section 7 can then be then expressed as:

$$
\begin{aligned}
\mathbf{V}_{\mathbf{ESwaption}}(0) &= P(0,T) \mathbb{E}_{\widehat{\mathbb{P}}^{T_1}} \left[\left(K \sum_{i=2}^{5} P(T_1,T_i)(T_i - T_{i-1}) + P(T_1,T_5) - 1 \right)^+ \right] \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y) \; f^{\widehat{\mathbb{P}}^{T_1}}(x,y) \; dx \; dy
\end{aligned}
$$

Being a two dimensional integral, we decide to use a quadrature scheme to evaluate the integral. We choose a composite Simpson's scheme. Additionally, we set the truncations for the upper and lower limits of the integral, establishing the approximation:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)f(x,y)\ dx\ dy \approx \int_{L}^{U}\int_{L}^{U} g(x,y)f^{\widehat{\mathbb{P}}^{T_1}}(x,y)dxdy \qquad (62)$$

with $L=-5$ and $U=5$.[31].

Following the presentation in [Holton, 2003], we establish:

$$\int_{L}^{U}\int_{L}^{U} g(x,y)f^{\widehat{\mathbb{P}}^{T_1}}(x,y)dxdy \approx \sum_{i=0}^{N_x}\sum_{j=0}^{N_y} g(x_i,y_j)\ f^{\widehat{\mathbb{P}}^{T_1}}(x_i,y_i)\ w_x(i)\ w_y(j) \qquad (63)$$

where $w_x(i)\ w_y(j)$ denote the weights.

We set an equally spaced partition $p_x = x_0, x_1, ...x_{N_x}$ and $p_y = y_0, y_1, ...y_{N_y}$ and define $\Delta x := \frac{U-L}{N_x}$ and $\Delta y := \frac{U-L}{N_y}$.

Choosing the weights as $w_x = \left\{\frac{\Delta x}{3}, 4\frac{\Delta x}{3}, 2\frac{\Delta x}{3}, \dots, 2\frac{\Delta x}{3}, 4\frac{\Delta x}{3}, \frac{\Delta x}{3}\right\}$ and $w_y = \left\{\frac{\Delta y}{3}, 4\frac{\Delta y}{3}, 2\frac{\Delta y}{3}, \dots, 2\frac{\Delta y}{3}, 4\frac{\Delta y}{3}, \frac{\Delta y}{3}\right\}$ we implement Simpson's composite rule for the two dimensional case.

For a large $N = N_x = N_y$ (i.e. 5000, so we can consider this a reasonably precise approximation), the value of the integral computing (63) is 0.20618126. After some experimentation, we find that $n = 110$ and $L=-3, U=3$ yields the same result up to the fifth decimal place, and requires way less function evaluations, so is obviously faster to compute. Other values of N and their corresponding results are shown for comparison. All these results are summed up in Table 4.

[31] In a bivariate standard normal distribution, the probability enclosed by $x \in [-5,5]$, $y \in [-5,5]$ is 0.9999988, so it seems wasteful to expand the limits any further.

Numerical integration				
Truncation	$L=-5, U=5$			$L=-3, U=3$
N	$N=100$	$N=1000$	$N=5000$	$N=110$
Value	0.20100217	0.20618123	0.20618126	0.20618771

Table 4: Numerical integration results for the European swaption maturing at $T_1 := 1$ for entering a receiver swap at $K\prime = 0.005$ with reset dates {1, 1.25, 1.5, 1.75} and payment dates {1.25, 1.5, 1.75, 2}.

7.3 *Random Tree.* Bermudan swaption

The *random tree* methodology developed by [Broadie and Glasserman, 1997] will be implemented to value the Bermudan swaption described before. It involves simulating the discounted payoff of the option (computed through the previous simulation of the state variables) in a tree structure and then computing estimators that will allow us to establish boundaries for its value. It requires the simulation of paths conditional on the value they took on the previous node, so it clearly benefits from the Markovianity of the state variables.

This method is best suited for Bermudan-style options with a relatively low number of exercise dates. Its use for American options would be excessively costly due to the exponential scaling of the number of nodes with respect to the amount of exercise dates.

An upper bound for the value of the option will be constructed using an estimator that is biased high (High Estimator), and a lower bound with another which is biased low (Low Estimator). The consistency of the estimators is discussed in [Broadie and Glasserman, 1997]. The procedure to compute these will be described step-by-step, with its corresponding pseudocode detailed in Algorithm 2.

The tree has an structure where b branches come out of each node, with this process being repeated m times. A greater number of branches per node increases the precision of the estimation, whereas the extension of the tree m is set by the possible exercise dates. A relevant feature (different from *lattice* methods) is that the value of each is node is random, as it stems from simulations of the state variables.

We provide a graphical representation of a tree with $m = 2$, $b = 2$, to clarify the idea.

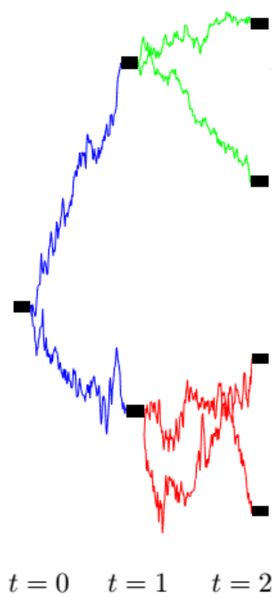

Figure 2: Example of one simulation of the tree structure for a given underlying process, with $m = 2$, $b = 2$ and time step of 1. After the initial node at $t = 0$, we have b nodes at $t = 1$ and b^2 nodes at $t = 2$.

Now we will explain the tree construction procedure for our model and the computation of the High and Low estimators for the value of the Bermudan swaption.

Trees for the state variables The tree structure for the state variables is built as follows: for each node before the terminal ones, simulate b independent replications of the trajectory of the state variable up to the next node using (57), conditional on the value of the starting node. Then repeat the process for the second state variable.

Discounted payoff tree We consider the payoff of the swaption in each node, given the bond prices determined by the simulated state variables at that point. That is, we compute, for each node, the payoff (32) particularized[32] for our instrument:

$$\begin{aligned}
\mathbf{Payoff_{a,c,\theta}} \quad & (t_i, X_{11}(t_i), X_{12}(t_i)), \delta) := \Bigg[K\Bigg(P_{a,c,\theta}(t_i, t_i+\delta, X_{11}(t_i), X_{12}(t_i))\delta \\
+ \quad & P_{a,c,\theta}(t_i, t_i+2\delta, X_{11}(t_i), X_{12}(t_i))\delta + P_{a,c,\theta}(t_i, t_i+3\delta, X_{11}(t_i), X_{12}(t_i))\delta \\
+ \quad & P_{a,c,\theta}(t_i, t_i+4\delta, X_{11}(t_i), X_{12}(t_i))\delta\Bigg) \\
+ \quad & P_{a,c,\theta}(t_i, t_i+4\delta, X_{11}(t_i), X_{12}(t_i)) - 1\Bigg]^+
\end{aligned}$$

Following [Glasserman, 2003], when simulating under the t_m forward measure ($\widehat{\mathbb{P}}^{T_1}$ in our case), we set up a discount factor:

$$D_0(i) := \frac{P(0,t_m)}{P(t_i,t_m)} = \frac{P(0,T_1)}{P(t_i,T_1)},$$

The discounted payoff tree is obtained simply by multiplying the corresponding time t_i discount factor $D_0(i)$ to each node at each time t_i, $i = \{1, ..., m\}$.

High estimator The value of the estimator at the terminal nodes is equal to the payoff of the swaption at the corresponding terminal nodes.

$$\hat{\Theta}_H^{j_1,j_2,...j_m}(m) = DPO_{swaption}^{j_1,j_2,...j_m}(m, X_{11}(m), X_{12}(m)),$$

Then, working backwards, we obtain the values for the estimator at nodes $i = m-1$,
$m-2, ..., 1$ by calculating:[33]

$$\hat{\Theta}_H^{j_1,j_2,...j_i}(i) = max\Bigg(DPO_{swaption}^{j_1,j_2,...j_i}(i, X_{11}(i), X_{12}(i)), \frac{1}{b}\sum_{j=1}^{b}\hat{\Theta}_H^{j_1,j_2,...j_{i+1}}(i+1, j)\Bigg)$$

[32] $P_{a,c,\theta}(t, T, X_{11}(t), X_{12}(t))$ refers to a function that computes (49) with the parameter values For a, c, θ obtained in Section 6.

[33] The notation used by Glasserman is not particularly clear. He positions a node describing the path followed by taking the j-th branch in each node, with $j = 1, ...b$, so a sequence of $j_1, j_2, ...j_i$ is required. In contrast, in our pseudocode implementation we denote by j the vertical position in the tree, so $j = 1, ..., b^i$ for each node i, which seems more intuitive.

Low estimator The estimator at the terminal nodes is again the payoff:

$$\hat{\Theta}_L^{j_1,j_2,\ldots j_m}(m) = DPO_{swaption}^{j_1,j_2,\ldots j_m}(m, X_{11}(m), X_{12}(m))$$

For each one of the previous nodes, a preliminary calculation (a variable named *aux* is used for this in our code) is required. We consider, for each $k = 1, ..., b$:

$$\frac{1}{b-1} \sum_{j=1, j\neq k}^{b} \hat{\Theta}_L^{j_1,j_2,\ldots j_{i+1}}(i+1, j) \leq DPO_{swaption}^{j_1,j_2,\ldots j_i}(i, X_{11}(i), X_{12}(i)) \quad (64)$$

If (64) holds, then set:

$$\xi^{j_1,j_2,\ldots j_m}(i,k) := DPO_{swaption}^{j_1,j_2,\ldots j_i}(i, X_{11}(i), X_{12}(i))$$

Otherwise:

$$\xi^{j_1,j_2,\ldots j_m}(i,k) := \hat{\Theta}_L^{j_1,j_2,\ldots j_{i+1}}(i+1, k)$$

Then, the Low Estimator at nodes $i = m-1, m-2, ..., 1$ becomes:

$$\hat{\Theta}_L^{j_1,j_2,\ldots j_i}(i) = \frac{1}{b} \sum_{k=1}^{b} \xi^{j_1,j_2,\ldots j_i}(i,k)$$

The steps explained so far comprise one simulation of the *random tree* estimators.

Algorithm 2 Simulation of one instance of the *random tree*

Require: $P_{a,c,\theta}(t, T, X_{11}(t), X_{12}(t)), P(0,T), b, m, N_{int}, T_1, \delta$

Ensure: $\hat{\Theta}_L$,$\hat{\Theta}_H$

1: $\Delta t = \dfrac{T_1}{N_{int}}$

2: $N_{branch} = \dfrac{N_{int}}{m}$

3: $X_{11}^{nodes} = \mathbf{0}_{b^m \times (m+1)}$

4: $X_{12}^{nodes} = \mathbf{0}_{b^m \times (m+1)}$

▷ Trees for the state variables

for $i = 2$ to $m + 1$ **do**
  for $j = 1$ to $b^{(i-1)}$ **do**
    counter $= \left\lceil \frac{j}{b} \right\rceil$
    $X_{11}^{simul}[1] = X_{11}^{nodes}[\text{counter}, i-1]$
    $X_{12}^{simul}[1] = X_{12}^{nodes}[\text{counter}, i-1]$
    for $k = 1$ to $N_{branch} - 1$ **do**
      $X_{11}^{simul}[k+1] = X_{11}^{simul}[k] + a^2\Bigg(2(N_{branch}(i-2)+k) - T_1\Bigg)\Delta t$
      $+\ a\sqrt{\Delta t}\ random_{\mathcal{N}(0,1)}$
      $X_{12}^{simul}[k+1] = X_{12}^{simul}[k] + \Bigg(-\theta X_{12}^{simul}[k] + \frac{c^2}{2\theta}\Bigg(1$
      $-\exp(-2\theta\Delta t(N_{branch} * (i-2)+k))\Bigg)$
      $+\frac{c^2}{\theta}\Bigg(\exp(-\theta(T_1 - \Delta t(N_{branch} * (i-2)+k))) - 1\Bigg)\Bigg)\Delta t$
      $+\ c\ \sqrt{\Delta t}\ random_{\mathcal{N}(0,1)}$
    end for k
    $X_{11}^{nodes}[j,i] = X_{11}^{simul}[N_{branch}]$
    $X_{12}^{nodes}[j,i] = X_{12}^{simul}[N_{branch}]$
  end for j
end for i
$PO_{swaption}^{nodes} = \mathbf{0}_{b^m \times m+1}$

▷ Discounted payoff tree

for $i = 1$ to $m + 1$ **do**
  for $j = 1$ to $b^{(i-1)}$ **do**
    counter $= \left\lceil \frac{j}{b} \right\rceil$
    $PO_{swaption}^{nodes}[j,i] = \mathbf{Payoff}_{\mathbf{a,c,\theta}}(\Delta t\ N_{branch}(i-1),$
    $, X_{11}^{nodes}[j,i], X_{12}^{nodes}[j,i], \delta)$
  end for j
end for i
$P_{T_1-mat} = \mathbf{0}_{b^m \times m+1}$; $P_{T_1-mat}[1,1] = P_{a,c,\theta}[0, T_1, 0, 0]$
for $i = 2$ to $m + 1$ **do**
  for $j = 1$ to $b^{(i-1)}$ **do**
    counter $= \left\lceil \frac{j}{b} \right\rceil$
    $P_{T_1-mat}[j,i] = P_{a,c,\theta}(\Delta t N_{branch}(i-1), T_1, X_{11}^{nodes}[j,i], X_{12}^{nodes}[j,i])$
  end for j
end for i

$DPO^{nodes}_{swaption} = \mathbf{0}_{b^m \times m+1}$; $DPO^{nodes}_{swaption}[1,1] = PO^{nodes}_{swaption}[1,1]$
for $i = 2$ to $m+1$ **do**
  for $j = 1$ to $b^{(i-1)}$ **do**
    counter $= \left\lceil \dfrac{j}{b} \right\rceil$
    disc $= \dfrac{P_{a,c,\theta}(0, T_1, 0, 0)}{P_{T_1-mat}[j,i]}$
    $DPO^{nodes}_{swaption}[j,i] = \text{disc}PO^{nodes}_{swaption}[j,i]$
  end for j
end for i

▷ High Estimator

$HE^{nodes} = \mathbf{0}_{b^m \times m+1}$; $HE^{nodes}[\,:\,, m+1] = DPO^{nodes}_{swaption}[\,:\,, m+1]$
for $i = (m+1) - 1$ to 1 **step** -1 **do**
  for $j = 1$ to $b^{(i-1)}$ **do**
    $HE^{nodes}[j,i] = max(DPO[j,i], \dfrac{1}{b}\Sigma(HE[jb-(b-1):jb, i+1]]))$
  end for j
end for i
$\Theta_H = HE^{nodes}[1,1]$

▷ Low Estimator

$LE^{nodes} = \mathbf{0}_{b^m \times m+1}$; $LE^{nodes}[\,:\,, m+1] = DPO^{nodes}_{swaption}[\,:\,, m+1]$
for $i = (m+1) - 1$ to 1 **step** -1 **do**
  for $j = 1$ to $b^{(i-1)}$ **do**
    for $k = 1$ to b **do**
      aux $= \mathbf{0}_{k \times 1}$
      if $DPO^{nodes}_{swaption}[j,i] > \dfrac{1}{b-1}\Big(\Sigma(HE[jb-(b-1):jb, i+1])$
      $-HE[j*b+k-b, i+1]\Big)$ **then**
        aux $[k]$ $= DPO^{nodes}_{swaption}[j,i]$
      else
        aux $[k]$ $= LE^{nodes}[j+k-1, i+1]$
      end if
      $LE^{nodes}[j,i] = \dfrac{1}{b}\Sigma(\text{aux})$
    end for k
  end for j
end for i
$\hat{\Theta}_L = LE^{nodes}[1,1]$
return $\hat{\Theta}_L$,$\hat{\Theta}_H$

Confidence interval for the value of the Bermudan swaption The confidence interval defined in [Broadie and Glasserman, 1997] sets the upper bound as: [34]

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\Theta}_{H,\{i\}} + \frac{\Phi(\frac{\eta}{2})s_{\hat{\Theta}_H}}{\sqrt{n}}$$

And the lower bound:

$$\frac{1}{n}\sum_{i=1}^{n}\hat{\Theta}_{L,\{i\}} - \frac{\Phi(\frac{\eta}{2})s_{\hat{\Theta}_L}}{\sqrt{n}}$$

Repeating the process n times to obtain n independent replications and computing the sample mean and sample standard deviation of $\{\hat{\Theta}_{L(1)}, \hat{\Theta}_{L(2)}, \ldots, \hat{\Theta}_{L(n)}\}$ and $\{\hat{\Theta}_{H(1)}, \hat{\Theta}_{H(2)}, \ldots, \hat{\Theta}_{H(n)}\}$, we can obtain a confidence interval for $\mathbf{V_{BSwaption}}(0)$.

In Table 5 we show the upper and lower bounds that define the confidence intervals for $\eta = 0.05$, with $n = 500$ simulations of the tree, and $b = 2, 3, 4, 5$.

	Parameter b			
	2	3	4	5
$\frac{1}{500}\sum_{i=1}^{500}\Theta_{L,\{i\}}$	0.2418	0.2673	0.2615	0.2625
$\frac{1}{500}\sum_{i=1}^{500}\Theta_{H,\{i\}}$	0.3252	0.3018	0.2818	0.2744
$s_{\hat{\Theta}_L}$	0.2726	0.2800	0.2791	0.2840
$s_{\hat{\Theta}_H}$	0.2679	0.1869	0.1422	0.1248
Lower bound	0.2316	0.2569	0.2511	0.2519
Upper bound	0.3352	0.3088	0.2871	0.2791

Table 5: *Random tree* simulations ($n = 500$) for the Bermudan swaption with different values of $b = 2, 3, 4, 5$.

[34]The notation $\Theta_{H,\{i\}}$ refers to the value of the High Estimator in simulation i.

8 Concluding remarks

Some of the limitations and possible extensions of the work done so far are outlined:

- Calibration could have been made to a broader set of instruments. During the process it was seen that many different combinations of parameters were able to approximately reproduce the market price of caplets for the particular strike chosen, indicating that there is still room within the model for considering further market data.
- A great deal of bond price volatility is present in our simulations, that could indicate that other values of the parameters, yet to be found, might be more adequate. To that regard, additional optimization methods, such as the Nelder-Mead algorithm, could have been tested.
- The multicurve approach to valuation of interest rate derivatives was not considered in order to simplify the exposition, and it is certainly an extension that should be included.
- Instantaneous correlation between the two Brownian motions driving the model could be introduced to allow for grater flexibility. More dramatic extensions such as stochastic volatility would require a complete rework of the exposition.
- A PDE formulation of the valuation of the instruments could have been introduced using Feynman-Kac (as in [Valero et al., 2011] or [Beyna, 2013]), making further use of the Markovianity of the state variables.
- As noted during the work, the $\beta(t)$ function in (35) can be freely chosen (although being easily differentiable and integrable is highly recommended) without adding any state variables, so a more complex function could be chosen in order the improve the capabilities of the model, at the expense of lengthier computations.

A Computations in the Two-Factor Cheyette model

A.1 Instantaneous forward rate

In our Two-Factor model, we have $M = 2$, $N_1 = 1$ and $N_2 = 1$, so inserting the Cheyette volatility (35) in the integrated two factor version of (23) we get:

$$
\begin{aligned}
f(t,T) \quad = \quad & f(0,T) \ + \int_0^t \left(\frac{\alpha_{11}(T)}{\alpha_{11}(s)} \beta_{11}(s) \left[\int_s^T \frac{\alpha_{11}(u)}{\alpha_{11}(s)} \beta_{11}(s) du \right] \right) ds \\
+ \quad & \int_0^t \left(\frac{\alpha_{12}(T)}{\alpha_{12}(s)} \beta_{12}(s) \left[\int_s^T \frac{\alpha_{12}(u)}{\alpha_{12}(s)} \beta_{12}(s) du \right] \right) ds \\
+ \quad & \int_0^t \left(\frac{\alpha_{11}(T)}{\alpha_{11}(s)} \beta_{11}(s) \right) dW_1(s) + \int_0^t \left(\frac{\alpha_{12}(T)}{\alpha_{12}(s)} \beta_{12}(s) \right) dW_2(s)
\end{aligned}
$$

Pulling terms out of the Riemann integrals:

$$
\begin{aligned}
= f(0,T) \quad + \quad & \int_0^t \left(\frac{\alpha_{11}(T)}{\alpha_{11}(s)} \beta_{11}(s) \left[\frac{\beta_{11}(s)}{\alpha_{11}(s)} \int_s^T \alpha_{11}(u) du + \right] \right) ds \\
+ \quad & \int_0^t \left(\frac{\alpha_{12}(T)}{\alpha_{12}(s)} \beta_{12}(s) \left[\frac{\beta_{12}(s)}{\alpha_{12}(s)} \int_s^T \alpha_{12}(u) du \right] \right) ds \\
+ \quad & \int_0^t \left(\frac{\alpha_{11}(T)}{\alpha_{11}(s)} \beta_{11}(s) \right) dW_1(s) + \int_0^t \left(\frac{\alpha_{12}(T)}{\alpha_{12}(s)} \beta_{12}(s) \right) dW_2(s)
\end{aligned}
$$

$$
\begin{aligned}
= f(0,T) \quad + \quad & \int_0^t \left(\frac{\alpha_{11}(T)\beta_{11}(s)\beta_{11}(s)}{\alpha_{11}(s)\alpha_{11}(s)} \int_s^T \alpha_{11}(u) du \right) ds \\
+ \quad & \int_0^t \left(\frac{\alpha_{12}(T)\beta_{12}(s)\beta_{12}(s)}{\alpha_{12}(s)\alpha_{12}(s)} \int_s^T \alpha_{12}(u) du \right) ds \\
+ \quad & \int_0^t \left(\frac{\alpha_{11}(T)}{\alpha_{11}(s)} \beta_{11}(s) \right) dW_1(s) + \int_0^t \left(\frac{\alpha_{12}(T)}{\alpha_{12}(s)} \beta_{12}(s) \right) dW_2(s)
\end{aligned}
$$

Multiplying the integrals by $\frac{\alpha_{1i}(t)}{\alpha_{1i}(t)}$, noting that $\int_s^T \alpha_{1i}(u)du = A_{1i}(T) - A_{1i}(s)$, and adding $\pm A_{1i}(t)$ inside the Riemann integrals:

$$\begin{aligned}
= f(0,T) \quad & + \quad \int_0^t \left(\frac{\alpha_{11}(T)\beta_{11}(s)\beta_{11}(s)}{\alpha_{11}(s)\alpha_{11}(s)} [A_{11}(T) - A_{11}(s)] \right. \\
& + \quad \left. \frac{\alpha_{12}(T)\beta_{12}(s)\beta_{12}(s)}{\alpha_{12}(s)\alpha_{12}(s)} [A_{12}(T) - A_{12}(s)] \right) ds \\
& + \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \int_0^t \frac{\alpha_{11}(t)}{\alpha_{11}(s)} \beta_{11}(s) dW_1(s) + \frac{\alpha_{12}(T)}{\alpha_{12}(t)} \int_0^t \frac{\alpha_{12}(t)}{\alpha_{12}(s)} \beta_{12}(s) dW_2(s)
\end{aligned}$$

$$\begin{aligned}
= f(0,T) \quad & + \quad \int_0^t \left(\frac{\alpha_{11}(t)}{\alpha_{11}(t)} \frac{\alpha_{11}(T)\beta_{11}(s)\beta_{11}(s)}{\alpha_{11}(s)\alpha_{11}(s)} [A_{11}(t) - A_{11}(s)] \right. \\
& + \quad \left. \frac{\alpha_{12}(t)}{\alpha_{12}(t)} \frac{\alpha_{12}(T)\beta_{12}(s)\beta_{12}(s)}{\alpha_{12}(s)\alpha_{12}(s)} [A_{12}(t) - A_{12}(s)] \right) ds \\
& + \quad \int_0^t \left(\frac{\alpha_{11}(t)}{\alpha_{11}(t)} \frac{\alpha_{11}(T)\beta_{11}(s)\beta_{11}(s)}{\alpha_{11}(t)\alpha_{11}(s)} [A_{11}(T) - A_{11}(t)] \right. \\
& + \quad \left. \frac{\alpha_{12}(t)}{\alpha_{12}(t)} \frac{\alpha_{12}(T)\beta_{12}(s)\beta_{12}(s)}{\alpha_{12}(s)\alpha_{12}(s)} [A_{12}(T) - A_{12}(t)] \right) ds \\
& + \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \int_0^t \frac{\alpha_{11}(t)}{\alpha_{11}(s)} \beta_{11}(s) dW_1(s) + \frac{\alpha_{12}(T)}{\alpha_{12}(t)} \int_0^t \frac{\alpha_{12}(t)}{\alpha_{12}(s)} \beta_{12}(s) dW_2(s)
\end{aligned}$$

$$\begin{aligned}
= f(0,T) \quad & + \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \int_0^t \left(\frac{\alpha_{11}(t)\beta_{11}(s)\beta_{11}(s)}{\alpha_{11}(s)\alpha_{11}(s)} [A_{11}(t) - A_{11}(s)] \right) ds \\
& + \quad \frac{\alpha_{12}(T)}{\alpha_{12}(t)} \int_0^t \left(\frac{\alpha_{12}(t)\beta_{12}(s)\beta_{12}(s)}{\alpha_{12}(s)\alpha_{12}(s)} [A_{12}(t) - A_{12}(s)] \right) ds \\
& + \quad \int_0^t \left(\frac{\alpha_{11}(t)}{\alpha_{11}(t)} \frac{\alpha_{11}(T)\beta_{11}(s)\beta_{11}(s)}{\alpha_{11}(s)\alpha_{11}(s)} [A_{11}(T) - A_{11}(t)] \right) ds \\
& + \quad \int_0^t \left(\frac{\alpha_{12}(t)}{\alpha_{12}(t)} \frac{\alpha_{12}(T)\beta_{12}(s)\beta_{12}(s)}{\alpha_{12}(s)\alpha_{12}(s)} [A_{12}(T) - A_{12}(t)] \right) ds \\
& + \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \int_0^t \frac{\alpha_{11}(t)}{\alpha_{11}(s)} \beta_{11}(s) dW_1(s) + \frac{\alpha_{12}(T)}{\alpha_{12}(t)} \int_0^t \frac{\alpha_{12}(t)}{\alpha_{12}(s)} \beta_{12}(s) dW_2(s)
\end{aligned}$$

Rearranging and pulling elements out of the integrals:

$$\begin{aligned}
= \ f(0,T) \quad &+ \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \int_0^t \left(\frac{\alpha_{11}(t)\beta_{11}(s)}{\alpha_{11}(s)} \frac{[A_{11}(t) - A_{11}(s)]}{\alpha_{11}(s)} \beta_{11}(s) \right) ds \\
&+ \quad \frac{\alpha_{12}(T)}{\alpha_{12}(t)} \int_0^t \left(\frac{\alpha_{12}(t)\beta_{12}(s)}{\alpha_{12}(s)} \frac{[A_{12}(t) - A_{12}(s)]}{\alpha_{12}(s)} \beta_{12}(s) \right) ds \\
&+ \quad \frac{\alpha_{11}(T)[A_{11}(T) - A_{11}(t)]}{\alpha_{11}(t)\alpha_{11}(t)} \int_0^t \left(\frac{\alpha_{11}(t)\alpha_{11}(t)}{\alpha_{11}(s)\alpha_{11}(s)} \beta_{11}(s)\beta_{11}(s) \right) ds \\
&+ \quad \frac{\alpha_{12}(T)[A_{12}(T) - A_{12}(t)]}{\alpha_{12}(t)\alpha_{12}(t)} \int_0^t \left(\frac{\alpha_{12}(t)\alpha_{12}(t)}{\alpha_{12}(s)\alpha_{12}(s)} \beta_{12}(s)\beta_{12}(s) \right) ds \\
&+ \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \int_0^t \frac{\alpha_{11}(t)}{\alpha_{11}(s)} \beta_{11}(s) dW_1(s) + \frac{\alpha_{12}(T)}{\alpha_{12}(t)} \int_0^t \frac{\alpha_{12}(t)}{\alpha_{12}(s)} \beta_{12}(s) dW_2(s)
\end{aligned}$$

$$\begin{aligned}
= \ f(0,T) \quad &+ \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \int_0^t \left(\frac{\alpha_{11}(t)\beta_{11}(s)}{\alpha_{11}(s)} \frac{[A_{11}(t) - A_{11}(s)]}{\alpha_{11}(s)} \beta_{11}(s)) ds \right. \\
&+ \quad \frac{\alpha_{12}(T)}{\alpha_{12}(t)} \int_0^t \left(\frac{\alpha_{12}(t)\beta_{12}(s)}{\alpha_{12}(s)} \frac{[A_{12}(t) - A_{12}(s)]}{\alpha_{12}(s)} \beta_{12}(s) \right) ds \\
&+ \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \frac{[A_{11}(T) - A_{11}(t)]}{\alpha_{11}(t)} \int_0^t \left(\frac{\alpha_{11}(t)\alpha_{11}(t)}{\alpha_{11}(s)\alpha_{11}(s)} \beta_{11}(s)\beta_{11}(s) \right) ds \\
&+ \quad \frac{\alpha_{12}(T)}{\alpha_{12}(t)} \frac{[A_{12}(T) - A_{12}(t)]}{\alpha_{12}(t)} \int_0^t \left(\frac{\alpha_{12}(t)\alpha_{12}(t)}{\alpha_{12}(s)\alpha_{12}(s)} \beta_{12}(s)\beta_{12}(s) \right) ds \\
&+ \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \int_0^t \frac{\alpha_{11}(t)}{\alpha_{11}(s)} \beta_{11}(s) dW_1(s) + \frac{\alpha_{12}(T)}{\alpha_{12}(t)} \int_0^t \frac{\alpha_{12}(t)}{\alpha_{12}(s)} \beta_{12}(s) dW_2(s)
\end{aligned}$$

Taking into account the definitions of $A_{ij}(t)$ and $V_{ij,k}(t)$ stated in (36) and (37) respectively, and also collecting terms:

$$\begin{aligned}
= \ f(0,T) \quad &+ \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)} \left[\int_0^t \left(\frac{\alpha_{11}(t)\beta_{11}(s)}{\alpha_{11}(s)} \frac{[A_{11}(t) - A_{11}(s)]}{\alpha_{11}(s)} \beta_{11}(s) \right) ds \right. \\
&+ \quad \left. \frac{[A_{11}(T) - A_{11}(t)]}{\alpha_{11}(t)} V_{11,1}(t) + \int_0^t \frac{\alpha_{11}(t)}{\alpha_{11}(s)} \beta_{11}(s) dW_1(s) \right] \\
&+ \quad \frac{\alpha_{12}(T)}{\alpha_{12}(t)} \left[\int_0^t \left(\frac{\alpha_{12}(t)\beta_{12}(s)}{\alpha_{12}(s)} \frac{[A_{12}(t) - A_{12}(s)]}{\alpha_{12}(s)} \beta_{12}(s) \right) ds \right. \\
&+ \quad \left. \frac{[A_{12}(T) - A_{12}(t)]}{\alpha_{12}(t)} V_{11,2}(t) + \int_0^t \frac{\alpha_{12}(t)}{\alpha_{12}(s)} \beta_{12}(s) dW_2(s) \right]
\end{aligned}$$

Applying the definition for the state variable, $X_{jk}(t)$, in (38):

$$\begin{aligned} = f(0,T) \quad &+ \quad \frac{\alpha_{11}(T)}{\alpha_{11}(t)}\left[X_{11}(t) + \frac{[A_{11}(T) - A_{11}(t)]}{\alpha_{11}(t)}V_{11,1}(t)\right] \\ &+ \quad \frac{\alpha_{12}(T)}{\alpha_{12}(t)}\left[X_{12}(t) + \frac{[A_{12}(T) - A_{12}(t)]}{\alpha_{12}(t)}V_{11,2}(t)\right] \end{aligned}$$

Collecting terms in summations we arrive to the following expression, with $k = 1, 2$ and $N_1 = 1, N_2 = 1$:

$$f(t,T) = f(0,T) \; + \sum_{k=1}^{2}\left[\frac{\alpha_{1k}(T)}{\alpha_{1k}(t)}\left[X_{1k}(t) + \frac{[A_{1k}(T) - A_{1k}(t)]}{\alpha_{1k}(t)}V_{11,k}(t)\right]\right] \tag{65}$$

which is easily related to the general case expression of (40).

A.2 Zero-coupon bond price

Using the expression for the zero-coupon bond in terms of the instantaneous forward rate (6), and (65) we have:

$$\begin{aligned}
P(t,T) &= \exp\Big(-\int_t^T f(t,u)du\Big)\\
&= \exp\Big(-\int_t^T f(0,u) + \frac{\alpha_{11}(u)}{\alpha_{11}(t)}\Big[X_{11}(t) + \frac{[A_{11}(u)-A_{11}(t)]}{\alpha_{11}(t)}V_{11,1}(t)\Big]\\
&+ \frac{\alpha_{12}(u)}{\alpha_{12}(t)}\Big[X_{12}(t) + \frac{[A_{12}(u)-A_{12}(t)]}{\alpha_{12}(t)}V_{11,2}(t)\Big]du\Big)\\
&= \frac{P(0,T)}{P(0,t)}\exp\Big(-\int_t^T\Big(\frac{\alpha_{11}(u)}{\alpha_{11}(t)}X_{11}(t) + \frac{\alpha_{12}(u)}{\alpha_{12}(t)}X_{12}(t)\Big)du\\
&- \int_t^T\Big(\frac{\alpha_{11}(u)}{\alpha_{11}(t)}\frac{[A_{11}(u)-A_{11}(t)]}{\alpha_{11}(t)}V_{11,1}(t) + \frac{\alpha_{12}(u)}{\alpha_{12}(t)}\frac{[A_{12}(u)-A_{12}(t)]}{\alpha_{12}(t)}V_{11,2}(t)\Big)du\Big)
\end{aligned}$$

Recalling that by the definition of $A_{ij}(t)$ we have $A_{1j}(u) - A_{1j}(t) = \int_t^u \alpha_{1j}(s)ds$, and extracting terms out of the integrals:

$$\begin{aligned}
&= \frac{P(0,T)}{P(0,t)}\exp\Big(-\int_t^T\Big(\frac{\alpha_{11}(u)}{\alpha_{11}(t)}X_{11}(t) + \frac{\alpha_{12}(u)}{\alpha_{12}(t)}X_{12}(t)\Big)du\\
&- \int_t^T\Big(\frac{\alpha_{11}(u)}{\alpha_{11}(t)}\Big(\int_t^u\frac{\alpha_{11}(s)}{\alpha_{11}(t)}ds\Big)V_{11,1}(t) + \frac{\alpha_{12}(u)}{\alpha_{12}(t)}\Big(\int_t^u\frac{\alpha_{12}(s)}{\alpha_{12}(t)}ds\Big)V_{11,2}(t)\Big)du\Big)\\
&= \frac{P(0,T)}{P(0,t)}\exp\Big(-\int_t^T\frac{\alpha_{11}(u)}{\alpha_{11}(t)}X_{11}(t)du - \int_t^T\frac{\alpha_{12}(u)}{\alpha_{12}(t)}X_{12}(t)du\\
&- \frac{1}{\alpha_{11}(t)\alpha_{11}(t)}V_{11,1}(t)\int_t^T\alpha_{11}(u)[A_{11}(u)-A_{11}(t)]du\\
&- \frac{1}{\alpha_{12}(t)\alpha_{12}(t)}V_{11,2}(t)\int_t^T\alpha_{12}(u)[A_{12}(u)-A_{12}(t)]du\Big)
\end{aligned}$$

Computing the integrals we arrive at:

$$
\begin{aligned}
= \quad & \frac{P(0,T)}{P(0,t)} \exp\Bigg(- \frac{X_{11}(t)}{\alpha_{11}(t)} \int_t^T \alpha_{11}(u)du - \frac{X_{12}(t)}{\alpha_{12}(t)} \int_t^T \alpha_{12}(u)du \\
- \quad & \frac{(A_{11}(T) - A_{11}(t))^2}{2\alpha_{11}(t)\alpha_{11}(t)} V_{11,1}(t) - \frac{(A_{12}(T) - A_{12}(t))^2}{2\alpha_{12}(t)\alpha_{12}(t)} V_{11,2}(t) \Bigg) \\
= \quad & \frac{P(0,T)}{P(0,t)} \exp\Bigg(- \frac{A_{11}(T) - A_{11}(t)}{\alpha_{11}(t)} X_{11}(t) - \frac{A_{12}(T) - A_{12}(t)}{\alpha_{12}(t)} X_{12}(t) \\
- \quad & \frac{(A_{11}(T) - A_{11}(t))^2}{2\alpha_{11}(t)\alpha_{11}(t)} V_{11,1}(t) - \frac{(A_{12}(T) - A_{12}(t))^2}{2\alpha_{12}(t)\alpha_{12}(t)} V_{11,2}(t) \Bigg)
\end{aligned}
$$

B Dynamics of the state variables

Computing the differential of (38), noting that the dependence on t of the integrands requires the use of Leibniz's integral rule, we get:

$$
\begin{aligned}
dX_{ik}(t) &= \frac{\alpha_{ik}(t)}{\alpha_{ik}(t)}\beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\right] d\widetilde{W}_k(t)\right) dt \\
&+ \left(\int_0^t \frac{\alpha_{ik}(t)}{\alpha_{ik}(t)}\beta_{ik}(t)\left(\sum_{j=1}^{N_k}\frac{A_{jk}(t)-A_{jk}(t)}{\alpha_{jk}(t)}\beta_{jk}(t)\right) ds\right) dt \\
&+ \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\left(\sum_{j=1}^{N_k}\frac{A_{jk}(t)-A_{jk}(s)}{\alpha_{jk}(s)}\beta_{jk}(s)\right)\right] ds\right) dt \\
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\right] d\widetilde{W}_k(t)\right) dt \\
&+ \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\left(\sum_{j=1}^{N_k}\frac{A_{jk}(t)-A_{jk}(s)}{\alpha_{jk}(s)}\beta_{jk}(s)\right)\right] ds\right) dt
\end{aligned}
$$

Using the definition of $A_{jk}(t)$:

$$
\begin{aligned}
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\right] d\widetilde{W}_k(t)\right) dt \\
&+ \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\left(\sum_{j=1}^{N_k}\frac{\int_s^t \alpha_{jk}(u)du}{\alpha_{jk}(s)}\beta_{jk}(s)\right)\right] ds\right) dt
\end{aligned}
$$

Extracting summation by the additive property of integrals and derivatives:

$$
\begin{aligned}
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\right] d\widetilde{W}_k(t)\right) dt \\
&+ \left(\sum_{j=1}^{N_k}\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\left(\frac{\int_s^t \alpha_{jk}(u)du}{\alpha_{jk}(s)}\beta_{jk}(s)\right)\right] ds\right) dt
\end{aligned}
$$

Pulling out terms of the last integral:

$$
\begin{aligned}
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\right] d\widetilde{W}_k(t)\right) dt \\
&+ \left(\sum_{j=1}^{N_k}\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\left(\int_s^t \alpha_{jk}(u)du\right)\right] ds\right) dt
\end{aligned}
$$

Extracting terms out of the derivative:

$$
\begin{aligned}
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\right] d\widetilde{W}_k(t)\right) dt \\
&+ \left(\sum_{j=1}^{N_k}\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\frac{\partial}{\partial t}\left[\alpha_{ik}(t)\left(\int_s^t \alpha_{jk}(u)du\right)\right] ds\right) dt
\end{aligned}
$$

Computing the derivative in the second summand using the product rule:

$$
\begin{aligned}
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\right] d\widetilde{W}_k(t)\right) dt \\
&+ \left(\sum_{j=1}^{N_k}\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\left[\frac{\partial\alpha_{ik}(t)}{\partial t}\int_s^t \alpha_{jk}(u)du + \alpha_{ik}(t)\alpha_{jk}(t)\right] ds\right) dt \\
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\right] d\widetilde{W}_k(t)\right) dt \\
&+ \left(\sum_{j=1}^{N_k}\left(\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\frac{\partial\alpha_{ik}(t)}{\partial t}\left(\int_s^t \alpha_{jk}(u)du\right) ds\right.\right. \\
&+ \left.\left.\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\alpha_{ik}(t)\alpha_{jk}(t)ds\right)\right)dt
\end{aligned}
$$

Rearranging derivatives and integrals:

$$
\begin{aligned}
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\right] d\widetilde{W}_k(t)\right) dt \\
&+ \left(\sum_{j=1}^{N_k}\left(\frac{\partial\alpha_{ik}(t)}{\partial t}\int_0^t\int_s^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\alpha_{jk}(u)duds + \alpha_{ik}(t)\alpha_{jk}(t)\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}ds\right)\right)dt \\
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial}{\partial t}\left[\frac{\alpha_{ik}(t)}{\alpha_{ik}(s)}\beta_{ik}(s)\right] d\widetilde{W}_k(t)\right) dt \\
&+ \left(\sum_{j=1}^{N_k}\left(\frac{\partial\alpha_{ik}(t)}{\partial t}\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\int_s^t \alpha_{jk}(u)duds + \alpha_{ik}(t)\alpha_{jk}(t)\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}ds\right)\right) dt
\end{aligned}
$$

Decomposing into two summands and substituting $V_{ij,k}(t)$:

$$\begin{aligned}
= \quad & \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\int_0^t \frac{\partial \alpha_{ik}(t)}{\partial t}\left[\frac{\beta_{ik}(s)}{\alpha_{ik}(s)}\right]d\widetilde{W}_k(t)\right)dt \\
+ \quad & \left(\sum_{j=1}^{N_k}\left(\frac{\partial \alpha_{ik}(t)}{\partial t}\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\int_s^t \alpha_{jk}(u)duds\right)\right. \\
+ \quad & \left.\sum_{j=1}^{N_k}\left(\underbrace{\alpha_{ik}(t)\alpha_{jk}(t)\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}ds}_{V_{ij,k}(t)}\right)\right)dt \\
= \quad & \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\frac{\partial \alpha_{ik}(t)}{\partial t}\int_0^t \frac{\beta_{ik}(s)}{\alpha_{ik}(s)}d\widetilde{W}_k(t)\right)dt \\
+ \quad & \left(\frac{\partial \alpha_{ik}(t)}{\partial t}\sum_{j=1}^{N_k}\left(\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\int_s^t \alpha_{jk}(u)duds\right) + \sum_{j=1}^{N_k}(V_{ij,k}(t))\right)dt \\
= \quad & \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\sum_{j=1}^{N_k}(V_{ij,k}(t))\right)dt + \left(\frac{\partial \alpha_{ik}(t)}{\partial t}\int_0^t \frac{\beta_{ik}(s)}{\alpha_{ik}(s)}d\widetilde{W}_k(t)\right)dt \\
+ \quad & \left(\frac{\partial \alpha_{ik}(t)}{\partial t}\sum_{j=1}^{N_k}\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\int_s^t \alpha_{jk}(u)duds\right)dt
\end{aligned}$$

Factoring out $\dfrac{\partial \alpha_{ik}(t)}{\partial t}$:

$$\begin{aligned}
= \quad & \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\sum_{j=1}^{N_k}(V_{ij,k}(t))\right)dt \\
+ \quad & \frac{\partial \alpha_{ik}(t)}{\partial t}\left[\int_0^t \frac{\beta_{ik}(s)}{\alpha_{ik}(s)}d\widetilde{W}_k(t) + \sum_{j=1}^{N_k}\int_0^t \frac{\beta_{ik}(s)\beta_{jk}(s)}{\alpha_{ik}(s)\alpha_{jk}(s)}\int_s^t \alpha_{jk}(u)duds\right]dt
\end{aligned}$$

Switching summation and intergal, computing the inner integral, and extracting i terms from the j-indexed summation we get:

$$\begin{aligned}
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\sum_{j=1}^{N_k}(V_{ij,k}(t))\right)dt \\
&+ \frac{\partial \alpha_{ik}(t)}{\partial t}\left[\int_0^t \frac{\beta_{ik}(s)}{\alpha_{ik}(s)}d\widetilde{W}_k(t) + \int_0^t \frac{\beta_{ik}(s)}{\alpha_{ik}(s)}\sum_{j=1}^{N_k}\frac{\beta_{jk}(s)}{\alpha_{jk}(s)}[A_{jk}(t) - A_{jk}(s)]ds\right]dt
\end{aligned}$$

Noting that $\dfrac{\partial \log \alpha_{ik}(t)}{\partial t} = \dfrac{\frac{\partial \alpha_{ik}(t)}{\partial t}}{\alpha_{ik}(t)}$, then $\dfrac{\partial \alpha_{ik}(t)}{\partial t} = \alpha_{ik}(t)\dfrac{\partial \log \alpha_{ik}(t)}{\partial t}$ and therefore:

$$\begin{aligned}
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\sum_{j=1}^{N_k}(V_{ij,k}(t))\right)dt \\
&+ \alpha_{ik}(t)\frac{\partial \log \alpha_{ik}(t)}{\partial t}\left[\int_0^t \frac{\beta_{ik}(s)}{\alpha_{ik}(s)}d\widetilde{W}_k(t) + \int_0^t \frac{\beta_{ik}(s)}{\alpha_{ik}(s)}\sum_{j=1}^{N_k}\frac{[A_{jk}(t) - A_{jk}(s)]}{\alpha_{jk}(s)}\beta_{jk}(s)ds\right]dt
\end{aligned}$$

Using the definition of $X_{ik}(t)$:

$$\begin{aligned}
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\sum_{j=1}^{N_k}(V_{ij,k}(t))\right)dt \\
&+ \frac{\partial \log \alpha_{ik}(t)}{\partial t}\left[\underbrace{\int_0^t \frac{\alpha_{ik}(t)\beta_{ik}(s)}{\alpha_{ik}(s)}d\widetilde{W}_k(t) + \int_0^t \frac{\alpha_{ik}(t)\beta_{ik}(s)}{\alpha_{ik}(s)}\sum_{j=1}^{N_k}\frac{[A_{jk}(t) - A_{jk}(s)]}{\alpha_{jk}(s)}\beta_{jk}(s)ds}_{X_{ik}(t)}\right]dt
\end{aligned}$$

Finally, grouping dt terms we obtain the expression in (39):

$$\begin{aligned}
dX_{ik}(t) &= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\sum_{j=1}^{N_k}(V_{ij,k}(t))\right)dt + \frac{\partial \log \alpha_{ik}(t)}{\partial t}\left[X_{ik}(t)\right]dt \\
&= \beta_{ik}(t)d\widetilde{W}_k(t) + \left(\sum_{j=1}^{N_k}(V_{ij,k}(t)) + \frac{\partial \log \alpha_{ik}(t)}{\partial t}X_{ik}(t)\right)dt
\end{aligned}$$

C Cubic spline yield-curve interpolation and caplet *stripping*

Dataset

The dataset used comprises *OIS zero rates* as seen at a particular date (3/21/2019) with maturities $T_{OIS} \in \{\frac{1}{12}, \frac{2}{12}, ... , 1, 2, ... , 10, 12, 15, 20, 25, 30\}$, and EUR caps with maturities $T_{caps} \in \{1, 1.5, 2, ... , 10, 12, 15, 20, 25, 30\}$. The settle dates for these are trimonthly until year 2, becoming semiannual after. Also, the very first caplet (which has no randomness attached to it) is not considered, so the relevant dates for caplets become $T_{settle,caplets} \in \{0.25, 0.5, 0.75, 1, 1.25, ... , 28, 28.5, 29, 29.5\}$ and $T_{maturity,caplets} \in \{0.5, 0.75, 1, 1.25, ... , 28, 28.5, 29, 29.5, 30\}$. Market quotes for caps are not prices but implied normal[35] flat volatilities, in order to retrieve cap prices we would need to plug them in Bachelier formula for caps. Even further, it would be preferable to work with caplets, the fundamental pieces, instead of caps, so we will try to deduce them in a process called *stripping*.

Cubic spline interpolation

In order to perform caplet *stripping*, we require discount factors not only at dates quoted in the market, but at every relevant caplet date: interpolation is therefore required. Cubic spline obtains a smooth function (a set of functions *pasted* together) that, by construction, includes all the data points supplied. The property of exactly recovering original data is the main reason for its use.

Considering a data set of $n + 1$ elements, a spline $S(x)$ is a piecewise function, whose n pieces are third order polynomials, so for $i = 1, ..., n$ we have:

$$S(x) = \begin{cases} S_1(x) = a_1(x - x_1)^3 + b_1(x - x_1)^2 + c_1(x - x_1) + d_1 & \text{If } x_0 \leq x \leq x_1, \\ \quad ... \\ S_i(x) = a_i(x - x_i)^3 + b_i(x - x_i)^2 + c_i(x - x_i) + d_i & \text{If } x_{i-1} \leq x \leq x_i, \\ \quad ... \\ S_n(x) = a_n(x - x_n)^3 + b_n(x - x_n)^2 + c_n(x - x_n) + d_n & \text{If } x_{n-1} \leq x \leq x_n \end{cases}$$

[35] The Bachelier model is one of the market standard quoting models since the appearance of negative rates in EUR, replacing Black model. It specifies normal dynamics for the instantaneous forward rate $d_t f(t, T) = \sigma dW(t)$, where the price at time 0 of a caplet is given by $(T_{i+1} - T_i)P(0, T_{i+1})(L(0, T_i, T_{i+1}) - K)\Phi(\frac{L(0,T_i,T_{i+1})-K}{\sigma\sqrt{T}}) + \sigma\sqrt{T_i}\phi(\frac{L(0,T_i,T_{i+1})-K}{\sigma\sqrt{T_i}})$.

We need to determine the 4 coefficients for each one of the n polynomials, and thus $4n$ unknowns. It has to satisfy the following properties:

- *Hit* all data points

$$S_i(x_i) = y_i, \ \forall i = 1, ..., n$$

$$S_i(x_{i-1}) = y_{i-1}, \ \forall i = 1, ..., n$$

- Be a *smooth* function altogether, so we require first and second derivatives to match at the endpoints of every piece:

$$S_i'(x_i) = S_{i+1}'(x_i), \ \forall i = 1, ..., n-1$$

$$S_i''(x_i) = S_{i+1}''(x_i), \ \forall i = 1, ..., n-1$$

In total being $4n-2$ restrictions, so we need to impose two additional ones in order to obtain a unique solution. An easy choice involves setting:

$$S_1''(x_0) = 0$$
$$S_n''(x_n) = 0$$

After some calculations, [36] the equations can be restated, in matrix form, as:

$$\begin{bmatrix} 2 & \lambda_1 & \dots & \dots & 0 \\ \mu_2 & 2 & \ddots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & 2 & \lambda_{n-2} \\ 0 & 0 & \dots & \mu_{n-1} & 2 \end{bmatrix} \begin{bmatrix} M_1 \\ \vdots \\ \vdots \\ \vdots \\ M_{n-1} \end{bmatrix} = \begin{bmatrix} d_1 \\ \vdots \\ \vdots \\ \vdots \\ d_{n-1} \end{bmatrix}$$

With:

$$\begin{aligned} M_i &= S''(x_i), \qquad i = 0, 1, ..., n \\ h_i &= x_i - x_{i-1}, \qquad i = 1, ..., n \\ \mu_i &= \frac{h_{i+1}}{h_i + h_{i+1}} \\ \lambda_i &= 1 - \mu_i \\ d_i &= 6\left(\frac{y_2 - y_1}{(x_2 - x_1)(x_2 - x_0)} - \frac{y_1 - y_0}{(x_1 - x_0)(x_2 - x_0)}\right) \end{aligned}$$

[36] See [Wikiversity, 2019] for details.

Each polynomial can be expressed as:

$$S_i(x) = M_{i-1}\frac{(x_i - x)^3}{6h_i} + M_i\frac{(x - x_{i-1})^3}{6h_i} + \left(y_{i-1} - \frac{M_{i-1}h_i^2}{6}\right)\frac{x_i - x}{h_i} + \left(y_i - \frac{M_i h_i^2}{6}\right)\frac{x - x_{i-1}}{h_i}$$

Using this last expression we obtain the polynomials required for the interpolation after solving the system of equations.

Following the notation introduced in Section 2, what has been computed is the yield function at $t = 0$:

$$y(0,T) = \begin{cases} S_1(T) & \text{If} \quad T \leq T_{\{1\}} \\ \quad ... & \\ S_n(T) & \text{If} \quad T_{\{n-1\}} \leq T \end{cases}$$

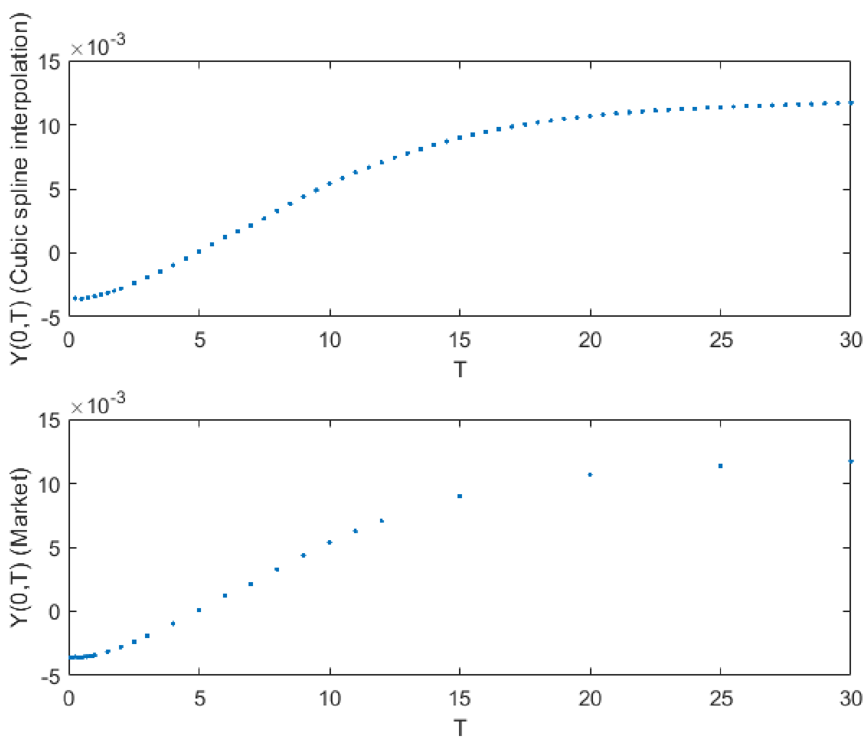

Figure 3: Market OIS yields for quoted dates and spline-interpolated yields for relevant caplet dates, i.e. $\{0.25, 0.5, ..., 29.5, 30\}$.

This allows to compute $t = 0$ bond prices for every maturity (by relationship (1)):

$$P(0,T) = \exp\Big(-T\ y(0,T)\Big)$$

Caplet *stripping*

A bootstrapping technique will be used to infer caplet volatilities, which involves finding iteratively the implied volatility that makes equal the sum of cap prices up to a certain date with the quoted volatility, and the sum of caplet prices using a different volatility for each set of caplets between cap maturities. The process allows for the recovery of caplet volatilities for every strike except ATM[37]. As noted in Section 6, we will focus on a single strike for calibration. The results for caplets with strike $k = 0.005$ are shown in Figure 4. Having recovered caplet implied volatilities, it is straightforward to obtain caplet prices by means of the Bachelier formula.

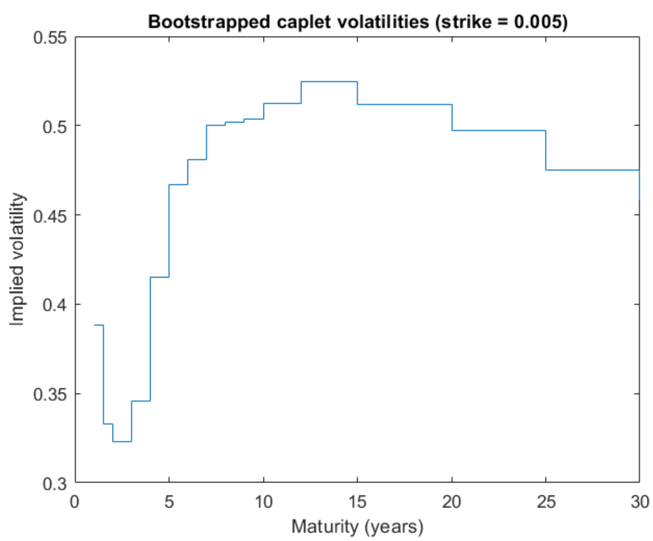

Figure 4: Stripped Bachelier volatilities for caplets with maturities $T_{maturity,caplets} \in \{0.5, 0.75, 1, 1.25, ..., 28, 28.5, 29, 29.5, 30\}$ and $k = 0.005$

[37]The ATM volatility won't be considered for simplicity, although is a liquid part of the market. The strike in this case is by definition equal to the swap rate being then maturity-varying, which adds a layer of complexity. It requires the calculation of prices for strikes not quoted in the market and therefore a strike-interpolating method is needed, as noted in [Atanasova, 2017].

References

[Acar, 2009] Acar, S. (2009). *A guide on the implementation of the Heath-Jarrow-Morton TwoFactor Gaussian Short Rate Model.* Berichte des Fraunhofer ITWM, Nr. 170.

[Atanasova, 2017] Atanasova, E. (2017). *Pricing and Hedging Fixed Income Derivatives under Negative Interest Rates: An SABR approach.* Master's Thesis, Universidad Complutense de Madrid.

[Baxter and Rennie, 1996] Baxter, M. and Rennie, A. (1996). *Financial calculus: An introduction to derivative pricing.* Cambridge University Press.

[Beyna, 2013] Beyna, I. (2013). *Interest Rate Derivatives: Valuation,Calibration and Sensitivity Analysis.* Springer.

[Brigo and Mercurio, 2006] Brigo, D. and Mercurio, F. (2006). *Interest Rate Models: Theory and Practice.* Springer Finance.

[Broadie and Glasserman, 1997] Broadie, M. and Glasserman, P. (1997). Pricing american-style securities using simulation. *Journal of Economics Dynamics and Control*, (21):1323 – 1352.

[Broadie and Glasserman, 2004] Broadie, M. and Glasserman, P. (2004). A stochastic mesh method for pricing high-dimensional american options. *The Journal of Computational Finance*, 7(4):35 – 72.

[Cairns, 2004] Cairns, A. (2004). *Interest Rate Models: An Introduction.* Princeton University Press.

[Carverhill, 1994] Carverhill, A. (1994). When is the short rate markovian? *Mathematical Finance*, 4(4):305 – 312.

[Cheyette, 1994] Cheyette, O. (1994). *Markov Representation of the Heath-Jarrow-Morton Model.* Presented at the UCLA Workshop on the Future of Fixed Income Financial Theory.

[Filipovic, 2009] Filipovic, D. (2009). *Term Structure Models: A Graduate Course.* Springer Finance.

[Glasserman, 2003] Glasserman, P. (2003). *Monte Carlo Methods in Financial Engineering.* Springer.

[Heath et al., 1992] Heath, D. et al. (1992). Bond pricing and the term structure of interest methodology for contingent claims valuation. *Econometrica*, 60(1):77 – 105.

[Holton, 2003] Holton, G. (2003). *Value at Risk: Theory and Practice*. Academic Press.

[Longstaff and Schwartz, 2001] Longstaff, F. and Schwartz, E. (2001). Valuing american options by simulation: A simple least-squares approach. *The Review of Financial Studies*, 14(1):113 – 147.

[Musiela and Rutkowski, 1997] Musiela, M. and Rutkowski, M. (1997). *Martingale Methods in Financial Modelling*. Springer.

[Pliska, 1997] Pliska, S. (1997). *Introduction to Mathematical Finance: Discrete time models*. John Wiley and Sons.

[Press et al., 2007] Press, H. et al. (2007). *Numerical Recipes: The Art of Scientific Computing (third edition)*. Cambridge University Press.

[Privault, 2018] Privault, N. (2018). *Notes on Stochastic Finance*. Nanyang Technological University.

[Protter, 2003] Protter, P. (2003). *Stochastic Integration and Differential Equations*. Springer.

[Ritchken et al., 1995] Ritchken, P. et al. (1995). Volatility structures of forward rates and the dynamics of the term structure. *Mathematical Finance*, 1(5):55 – 72.

[Sondermann, 2006] Sondermann, D. (2006). *Introduction to Stochastic Calculus for Finance: A new didactic approach*. Springer.

[Valero et al., 2011] Valero, E. et al. (2011). Fast resolution of a single factor heath-jarrow-morton model with stochastic volatility. *Journal of Computational and Applied Mathematics*, (236):1637 – 1655.

[Wikiversity, 2019] Wikiversity (2019). Cubic spline interpolation.