Projet d'approfondissement
Master Statistique et Finance
Barlcays

Multi-Factor Short Rate Models

Team
Hamza Bernoussi - Ensimag - MSF
Cédric Nogue Nogha - Ecole Polytechnique - MSF
Xinglong Tian - Ensta - MSF
Zhongxiu Zhang - Ensta - MSF

May 15, 2019

Contents

1 Introduction

The increasing complexity of hybrid products has significantly increased in recent years. We assist to the rising in equity-rates derivatives which provides coupons linked to the observation of equity and of many rates. An example of this kind of derivative is the dual/triple range on CMS spread, CMS and Indice equity. These complex structures need to catch the exact dynamics of rates and correlation between theses rates. Models such as short rate (Hull-White, HJM, ...) with one factor are not adapted to the pricing of these types of products. On a hedging/risks perspective it is important for a desk equity to price the vanilla-rate structure in a consistent way with the rates desk.
Our work consists in the analysis of different rates models to choose which of these will be the more adapted to price structured hybrid equity-rates that require the observation of many rates parameters (Libor, CMS, CMS spread).
We aimed to realize this project in three core steps: the analysis of short rates multi-factor models in which we will be developing the calibration methods of these models on the market instruments. Then we will build a process to calibrate the hybrid equity-rates model using Monte-Carlo and particles methods. Finally, we will do a comparative study of hybrid models for different Equity-rates products (hybrid swaps, swaptions...).
We have significantly worked on the general framework of multifactor rates modelling where we explain the fundamental ideas behind multifactor rates model and all the theoretical requirements (Parts 1 to 6) . The core elements of this approach are explained in the HJM model. To see the limitations of a one-factor model ant therefore the motivation of multifactor rates models, we used the Hull White one-factor model to develop a Monte-Carlo pricing of zero-coupons and the G1++ to price fictitious European call options on zero coupons, then we calibrate those prices with a set of given prices. That being done, we enlighten the advantages of the G2++ (two factors) to catch short term and long term structure, then we price swaptions using Monte-Carlo algorithm(Parts 7 and 8).
The main steps of our work are presented in details in the following parts,the language used for any numerical result is C++, a full project of the numerical implementation is provided with this report.

2 Reminders of the fundamentals on the interest rate modeling

2.1 Bonds and Forward Rates

$P(t,T)$ denote the time t **price of a zero-coupon bond** delivering for certain \$ 1 at maturity T.

$P(t,T,T+\tau)$ denote the time t **forward price for the zero-coupon bond** spanning $[T,T+\tau]$.

$$P(t,T,T+\tau) = P(t,T+\tau)/P(t,T)$$

$y(t,T,T+\tau)$ denote the **continuously compounded forward yield**, defined by:

$$e^{-y(t,T,T+\tau)\tau} = P(t,T,T+\tau)$$

$L(t,T,T+\tau)$ denote the **simple forward rate**, defined by:

$$1+\tau L(t,T,T+\tau) = 1/P(t,T,T+\tau)$$

$L(T,T,T+\tau)$ denote the **LIBOR (London Interbank Offered Rate) rates** quoted in the interbank market.

$f(t,T) := lim_{\tau\to 0}L(t,T,T+\tau)$ denote the time t **instantaneous forward rate** to time T, defined by:

$$f(t,T) := lim_{\tau\to 0}L(t,T,T+\tau)$$

$$P(t,T,T+\tau) = e^{-\int_T^{T+\tau} f(t,u)du}$$

$$f(t,T) = -\frac{\partial InP(t,T)}{\partial T}$$

$r(t)$ denote the **short rate** or sometimes the **spot rate**, defined by:

$$r(t) := f(t,t)$$

2.2 Fixed Income Probability Measures

We assume that the market is complete, and we use $V(t)$ to denote the time t price of a derivative security making an $\mathcal{F}_T$ measurable payment of $V(T)$.

2.2.1 Risk Neutral Measure

The **risk-neutral measure** Q is defined by the numeraire $\beta(t)$.

$\beta(t)$ denote the **continuously compounded money market account**, defined by:

$$\beta(t) := e^{\int_0^t r(u)du}$$

In the absence of arbitrage, the numeraire-deflated process $V(t)/\beta(t)$ must be a martingale, implying that:

$$\frac{V(t)}{\beta(t)} = E_t^Q\left(\frac{V(T)}{\beta(t)}\right)$$

$$V(t) = E_t^Q(V(T)e^{-\int_t^T r(u)du})$$

If we apply to the special case of $V(T) = 1$, we obtain a fundamental bond pricing formula:

$$P(t,T) = E_t^Q(e^{-\int_t^T r(u)du})$$

2.2.2 T-Forward measure

The T **Forward measure** Q^T is defined by the numeraire T-maturity zero-coupon bond:

$$\frac{V(t)}{P(t,T)} = E_t^T\left(\frac{V(T)}{P(T,T)}\right)$$

$$V(t) = P(t,T)E_t^T\left(V(T)\right), P(T,T) = 1$$

Change measures between Risk Neutral Measure and T Forward Measure:

$$\frac{dQ^T}{dQ} = \frac{P(T,T)}{P(0,T)}\frac{\beta(0)}{\beta(T)}$$

$$E_t^Q\left(\frac{dQ^T}{dQ}\right) = \frac{P(t,T)/P(0,T)}{\beta(t)}$$

$$\frac{dQ^T}{dQ}|_{\mathcal{F}_t} = \frac{\beta(t)/\beta(T)}{P(t,T)}$$

3 Swaption Pricing

A European swaption gives the holder a right, but not an obligation to enter a swap at a future date at a given fixed rate.

A payer swaption is an option to pay the fixed leg on a fixed-floating swap; A receiver swaption is an option to receive the fixed leg.

Assuming the underlying swap starts on the expiry date T_0 of the swaption, the payoff for a payer swaption at time T_0 equals:

$$V_{swaption}(T_0) = (V_{swap}(T_0))^+ = \left(\sum_{n=0}^{N-1} \tau_n P(T_0, T_{n+1})(L_n(T_0) - k)\right)^+$$

We could rewrite the swaption payout to:

$$V_{swaption}(T_0) = A(T_0)(S(T_0) - c)^+$$

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

$$S(t) = \frac{P(t, T_0) - P(t, T_N)}{A(t)}$$

Let Q^A be the measure induced by using A(t) as the numeraire, such that:

$$V_{swaption}(0) = A(0)E^A\left((S(T_0) - c)^+\right)$$

We know that S(t) is a martingale in Q^A, and due to our Markov setting, a deterministic function of $x(t)$, i.e. S(t)=S(t,x(t)). It follows from Ito's lemma that:

$$dS(t) = q(t, x(t))^T \sigma_x(t)^T dW(t)$$

where $W(t)$ is a Movement Brownian in the measure Q^A.

$q(t, x)$ is a d-dimensional column vector with elements:

$$q_j(t, x) = \frac{\partial S(t)}{\partial x_j} = \frac{\partial}{\partial x_j} \frac{P(t, T_0, x) - P(t, T_N, x)}{\sum_{i=0}^{N-1} \tau_i P(t, T_{i+1}, x)}$$

From the reconstitution formula of $P(t, T, x)$, we can evaluate the partial derivatives explicitly, yielding:

$$q_j(t, x) = -\frac{P(t, T_0, x)G_j(t, T_0) - P(t, T_N, x)G_j(t, T_N)}{A(t, x)} - \frac{S(t, x)\sum_{i=0}^{N-1} P(t, T_{i+1}, x)G_j(t, T_{i+1})}{A(t, x)}$$

The functions q_j can be experimentally verified to be close to a constant, so as a good approximation, we can write

$$q_j(t, x(t)) \approx q_j(t, \overline{x(t)})$$

where $\overline{x(t)}$ is some deterministic proxy for the random vector x(t). A reasonable approach is to set $\overline{x(t)} = 0$.

With this approximation, we get the following swaption pricing formula:

$$V_{swaption}(0) = A(0)\left[(S(0) - c)\Phi(d) + \sqrt{v}\phi(d)\right]$$

where

$$d = \frac{S(0) - c}{\sqrt{v}}, v = \int_0^{T_0} \|q(t, \overline{x(t)})^T \sigma_x(t)^T\|^2 dt$$

4 HJM Model

4.1 Bond price dynamics

Let's place ourselves in a finit horizon [0, T]. Let $P(t, T)$ be the price bond and $\beta(t)$ be the continuously rolled money market account defined by

$$d\beta(t) = r(t)\beta(t)dt$$

Let's define $P_\beta(t, T)$ by

$$P_\beta(t, T) = \frac{P(t, T)}{\beta(t)}$$

Assuming that there is absence of arbitrage, $P_\beta(t, T)$ is a martingale under the risk-neutral measure Q. Using the martingale representation theorem, we have:

$$dP_\beta(t, T) = -P_\beta(t, T)\sigma_P(t, T)^T dW(t)$$

with $\sigma_P(t, T)$ a d-dimension vector and W(t) a d-dimensional Q Brownian motion.
Using the above, we have

$$\frac{dP(t, T)}{P(t, T)} = r(t)dt - \sigma_P(t, T)^T dW(t)$$

proof: Applying Ito's Lemma for $P(t, T) = P_\beta(t, T)\beta(t)$, we have

$$\begin{aligned} dP(t, T) &= r(t)\beta(t)P_\beta(t, T)dt + \beta(t)dP_\beta(t, T) \\ &= r(t)P(t, T)dt - \sigma_P(t, T)^T P(t, T)dW(t) \end{aligned}$$

Therefore,

$$\frac{dP(t, T)}{P(t, T)} = r(t)dt - \sigma_P(t, T)^T dW(t)$$

4.2 Forward rate dynamics

Using Ito's formula for $ln(P(t,T))$, we have

$$\begin{aligned} dln(P(t,T)) &= \frac{dP(t,T)}{P(t,T)} - \frac{d < P(t,T) >}{P(t,T)^2} \\ &= r(t)dt - \sigma_P(t,T)^T dW(t) - \sigma_P(t,T)\sigma_P(t,T)^T dt \\ &= O(dt) - \sigma_P(t,T)^T dW(t) \end{aligned}$$

And knowing that $f(t,T) = -\frac{\partial ln(P(t,T))}{\partial T}$

$$df(t,T) = \mu_f(t,T)dt + \sigma_f(t,T)^T dW(t)$$

with $\sigma_f(t,T) = \frac{\partial \sigma_P(t,T)}{\partial T}$.
Let's determine now $\mu_f(t,T)$ We know that $f(t,T)$ is a martingale under the forward measure Q^F. Therefore, we have

$$df(t,T) = \sigma_f(t,T)^T dW^F(t)$$

and using the fact that

$$dW^F(t) = dW(t) + \sigma_P(t,T)dt$$

we conclude that, under the risk-neutral measure

$$df(t,T) = \sigma_f(t,T)^T \sigma_P(t,T)dt + \sigma_f(t,T)^T dW(t)$$

Therefore,

$$df(t,T) = \sigma_f(t,T)^T \left(\int_t^T \sigma_f(t,u)du\right)dt + \sigma_f(t,T)^T dW(t)$$

4.3 Classical development

In this part, instead of using a separability condition, we are supposing that $r(t)$ is an affine function of a set of states variables verifying a linear system. We can write $r(t)$ as

$$r(t) = b(t) + c(t)^T q(t)$$

where $b(t) \in \mathbf{R}$, $c(t) \in \mathbf{R}^d$ being deterministic, and $q(t)$ verifying under the risk neutral measure:

$$dq(t) = k(t)(m(t) - q(t))dt + \sigma(t)dW(t)$$

where $m(t) \in \mathbf{R}^d$, $k(t), \sigma(t) \in \mathbf{R}^{d\times d}$ being deterministic.
By defining $J_k(t)$ a $d \times d$ matrix satisfying:

$$\frac{dJ_k(t)}{dt} = -k(t)J_k(t)$$

we apply Ito's lemma to $u(t) = J_k(t)^{-1}$ q(t)

$$du(t) = J_k(t)^{-1}k(t)m(t)dt + J_k(t)^{-1}\sigma(t)dW(t)$$

Therefore,

$$q(t) = J_k(t)\left(q(0) + \int_0^t J_k(s)^{-1}k(s)m(s)ds + \int_0^t J_k(t)^{-1}\sigma(s)dW(s)\right)$$

if now we set $z(t)$ as

$$z(t) = J_k(t)^{-1}q(t) - \int_0^t J_k(s)^{-1}k(s)m(s)ds$$

By applying Ito:

$$dz(t) = J_k(t)^{-1}\sigma(t)dW(t)dz(t) = \sigma_z(t)^T dW(t)$$

$$dz(t) = \sigma_z(t)^T dW(t)$$

with $\sigma_z(t) = \sigma(t)^T \left(J_k(t)^{-1}\right)^T$ Therefore,

$$r(t) = b_z(t) + c_z(t)^T z(t)$$

with $b_z(t) = b(t) + c(t)^T J_k(t) \int_0^t J_k(s)^{-1}k(s)m(s)ds$ and $c_z(t) = J_k(t)^T c(t)$

5 Development from separability condition

5.1 Development for the Gaussian Model

From the HJM model we can get

$$dP(t,T)/P(t,T) = r(t)dt - \sigma_P(t,T)^\top dW(t)$$

$P(t,T)$ is the time t price of a zero-coupon bond, $\sigma_P(t,T)$ is a bounded d-dimensional function of time, and $W(t)$ a d-dimensional Brownian motion in the risk-neutral measure Q. We can also get that, for the forward rate $f(t,T)$

$$\begin{aligned} df(t,T) &= \sigma_f(t,T)^\top \sigma_P(t,T)dt + \sigma_f(t,T)^\top dW(t) \\ &= \sigma_f(t,T)^\top \int_t^T \sigma_f(t,u)dudt + \sigma_f(t,T)^\top dW(t) \end{aligned}$$

and for the short rate $r(t)$

$$r(t) = f(t,t) = f(0,t) + \int_0^t \sigma_f(u,t)^\top \int_u^t \sigma_f(u,s)dsdu + \int_0^t \sigma_f(u,t)^\top dw(u)$$

The process for $f(t,T)$ and $r(t)$ is generally not Markovian, as the integral from 0 to t concerns the state before t. So we introduce the separability condition to transform them, and we separate the two parameters t and T in $\sigma_f(t,T)$.

Proposition 5.1 *Assume that $\sigma_f(t,T)$ is separable, in the sense that it can be written as*

$$\sigma_f(t,T) = g(t)h(T)$$

where g is a $d \times d$ deterministic matrix-valued function, and H is a d-dimensional deterministic vector. Then

$$f(t,T) = f(0,T) + \Omega(t,T) + h(T)^\top z(t)$$

where $\Omega(t,T)$ is a deterministic scalar

$$\Omega(t,T) = h(T)^\top \int_0^t g(s)^\top g(s) \int_s^T h(u)duds$$

$z(t)$ is a d-dimensional random vector

$$dz(t) = g(t)^\top dW(t), z(0) = 0$$

In particular, we have

$$r(t) = f(0,t) + \Omega(t,t) + h(t)^\top z(t)$$

*Proof. Inserting (5.1) into (**??**) and integrating over time, we can directly get the results.*

5.2 Mean-reverting state variable

Proposition 5.1 demonstrates that if $\sigma_f(t,T)$ is separable, then the forward curve can be reconstructed from d Gaussian martingale variables $z_i(t)$, $i = 1, 2, ..., d$ with joint SDE (5.1). However, the choice of d state variables is not unique, and may in fact have disadvantages in a numerical implementation since often the components of $g(t)$ grow exponentially with times. As a result, it's common to shift variables to explicitly have a mean-reverting drift. Let's demonstrate one particular construction, set

$$H(t) = diag(h(t)) = \begin{pmatrix} h_1(t) & 0 & ... & 0 \\ 0 & h_2(t) & ... & ... \\ ... & ... & ... & 0 \\ 0 & ... & 0 & h_d(t) \end{pmatrix}$$

Assuming that for all t we have $h_i \neq 0, i = 1, ..., d$ then $H(t)$ is invertible, and we define a diagonal $d \times d$ matrix $\varkappa(t)$ by

$$\varkappa(t) = -\frac{dH(t)}{dt} H(t)^{-1}$$

We also set

$$x(t) = H(t) \int_0^t g(s)^\top g(s) \int_s^t h(u)duds + H(t)z(t)$$

$$y(t) = H(t)(\int_0^t g(s)^\top g(s)ds)H(t)$$

Here $x(t)$ is a d-dimensional random vector, $y(t)$ is a deterministic $d \times d$ symmetric matrix. It can be verified that

$$dy(t)/dt = H(t)g(t)^\top g(t)H(t) - \varkappa(t)y(t) - y(t)\varkappa(t)$$

Proposition 5.2 *Let the forward rate volatility be separable, as in Proposition 12.1. Let $\varkappa(t)$, $x(t)$ and $y(t)$ be defined as in (5.2)-(5.2), and assume that $H(t) = diag(h(t))$ is invertible. Also define $\mathbf{1}= (1, 1, ..., 1)^\top \in \mathbb{R}^d$. Then*

$$dx(t) = (y(t)\mathbf{1} - \varkappa(t)x(t))dt + \sigma_x^\top dW(t), \quad \sigma_x(t) = g(t)H(t)$$

and, with $M(t, T) := H(T)H(t)^{-1}\mathbf{1}$

$$f(t, T) = f(0, T) + M(t, T)^\top (x(t) + y(t) \int_t^T M(t, u)du)$$

In particular, we have

$$r(t) = f(t, t) = f(0, t) + \mathbf{1}^\top x(t) = f(0, t) + \Sigma_{i=1}^d x_i(t)$$

Proof. Note that the Leibniz integration gives

$$\frac{d}{dx}(\int_{a(x)}^{b(x)} f(x, t)dt) = f(x, b(x))\frac{d}{dx}b(x) - f(x, a(x))\frac{d}{dx}a(x) + \int_{a(x)}^{b(x)} \frac{\partial}{\partial x}f(x, t)dt$$

We write $V(t) = \int_0^t g(s)^\top g(s) \int_s^t h(u)duds$, then $x(t) = H(t)V(t) + H(t)z(t)$. We have $dx(t) = (H'V + HV' + H'z)dt + Hdz$.
The derivative for $V(t)$ is not so obvious. We write $V(t) = \int_0^t g(s)^\top g(s)\varphi(t, s)ds = f(t, \varphi(t, s))$, where $\varphi(t, s) = \int_s^t h(u)du$. Then

$$\frac{df(t, \varphi(t, s))}{dt} = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial \varphi}\frac{\partial \varphi}{\partial t}$$

the first term equals $g(t)g(t)^\top \varphi(t, t)$ according to the Leibniz integration. As $\varphi(t, t) = 0$ only the second term left, which equals $\int_0^t g(s)g(s)h(t)ds$. So we have

$$\begin{aligned}
dx(t) = &[\frac{dH(t)}{dt} \int_0^t g(s)^\top g(s) \int_s^t h(u)duds]dt \\
&+ [H(t) \int_0^t g(s)g(s)h(t)ds]dt + \frac{dH(t)}{dt}z(t)dt + H(t)dz(t) \\
= &\frac{dH(t)}{dt}H(t)^{-1}x(t)dt + [H(t)(\int_0^t g(s)g(s)ds)H(t)\mathbf{1}]dt \\
&+ H(t)g(t)^\top dW(t) \\
= &(y(t)\mathbf{1} - \varkappa(t)x(t))dt + \sigma_x(t)^\top dW(t).
\end{aligned}$$

The expression of $f(t, T)$ can be easily deducted by inserting the terms into the formula.

The meaning of this transformation is to transform (5.1) and (5.1) into a simplified form. As we can see $\Omega(t,T) = h(T)^\top V(t)$ is in fact the sum of the product of each pair of components of $h(T)$ and $V(t)$, we can introduce the diagonal matrix $H(t)$, as $H(t)V(t)$ is also a vector of the above products.

Also, we can see that $x(t)$ is somewhat a mean-reverting process, as $y(t)$ can be regarded as the mean level and $\varkappa(t)$ as the reverting rate.

We also need to deduct the expression of the zero-coupon. We can get the following corollary:

Corollary 5.3 *In the setting of Proposition 5.2, define*

$$G(t,T) = \int_t^T M(t,u)du$$

Then

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

Proof. We know that

$$P(t,T) = exp(-\int_t^T f(t,u)du)$$

By inserting (5.2) into (5.2) and doing some simple transformation we can get (5.2).

Let us examine some of the matrices involved in the multi-dimensional Gaussian model. As $\varkappa(t)$ is diagonal, we must have

$$\varkappa(t) = diag((\varkappa_1(t), \varkappa_2(t), ..., \varkappa_d(t))^\top)$$

Then we get

$$h(t) = (e^{-\int_0^t \varkappa_1(s)ds}, e^{-\int_0^t \varkappa_2(s)ds}, ..., e^{-\int_0^t \varkappa_d(s)ds})^\top.$$

And also

$$\begin{aligned} M(t,T) &= H(T)H(t)^{-1}\mathbf{1} \\ &= (e^{-\int_t^T \varkappa_1(s)ds}, e^{-\int_t^T \varkappa_2(s)ds}, ..., e^{-\int_t^T \varkappa_d(s)ds}) \end{aligned}$$

$$\begin{aligned} y(t) &= \int_0^t H(t)H(s)^{-1}\sigma_x(s)^\top \sigma_x(s) H(s)^{-1} H(t) ds \\ &= \int_0^t diag(M(s,t))\sigma_x(s)^\top \sigma_x(s) diag(M(s,t)) ds \end{aligned}$$

Here $\sigma_x(t)$ is the volatility of $x(t)$ and we have $\sigma_x(t) = H(t)g(t)^\top$. We know that $x(t)$ can be constructed from $y(t)$, and then $f(t,T)$ and $P(t,T)$. So that the specification of the d deterministic mean reversions $\varkappa_1(t), \varkappa_2(t), ..., \varkappa_d(t)$

and the $d \times d$ volatility matrix $\sigma_x(t)$ determines our $d-$dimensional Gaussian model. For the zero-coupon, we know that $G(t,T)$ takes the form

$$G(t,T) = (\int_t^T e^{-\int_t^u \varkappa_1(s)ds} du, ..., \int_t^T e^{-\int_t^u \varkappa_d(s)ds} du)^\top$$

which can be rewritten as

$$\int_t^T e^{-\int_t^u \varkappa_i(s)ds} du = (\int_0^T e^{-\int_0^u \varkappa_i(s)ds} du - \int_0^t e^{-\int_0^u \varkappa_i(s)ds} du) e^{\int_0^t \varkappa_i(s)ds}$$
$$:= (\Lambda_i(T) - \Lambda_i(t)) e^{\int_0^t \varkappa_i(s)ds}$$

In implementation we would typically pre-cache the 2d scalar functions

$$\Lambda_1(t), \Lambda_2(t), ..., \Lambda_d(t), exp(\int_0^t \varkappa_1(s)ds), ..., exp(\int_0^t \varkappa_d(s)ds)$$

on a suitable time grid, allowing subsequent discount bond pricing to be done quickly and conveniently for arbitrary t and T. Remark that the risk-neutral process for the zero-coupon $P(t,T)$ is log-normal

$$dP(t,T)/P(t,T) = r(t)dt - G(t,T)^\top \sigma_x(t)^\top dW(t)$$

which comes from (5.2).

By using a mean reverting matrix, we have written the short rate $r(t)$ as the forward value $f(0,t)$ plus a straight sum of d Gaussian mean-reverting variables $x_i(t)$, with each variable having a drift depending only on itself, as in (5.2). In fact we are not limited in this kind of expression, we can use any mean reversion matrix - diagonal or not - that we would like to get a more complicated form of $r(t)$. First we illustrate some extra notation. Consider the generic homogeneous ODE system

$$\frac{dp(t)}{dt} = -Q(t)p(t)$$

where Q(t) is a deterministic $d \times d$ matrix and p a $d-$dimensional (column) vector. The solution of (34) can always be represented as

$$p(T) = J_Q(T)p(0)$$

where $J_Q(T)$ is a $d \times d$ deterministic matrix satisfying

$$\frac{dJ_Q(t)}{dt} = -Q(t)J_Q(t)$$

The matrix $J_Q(t)$ is commutable by classical ODE methods and satisfies the boundary condition $J_Q(0) = I$, where I is the identity matrix. When Q is independent of time, we have $J_Q(t) = exp(-Qt)$. And notice that

$$\frac{d(J_Q(t)^{-1})}{dt} = (J_Q(t)^{-1})Q(t)$$

Lemma 5.4 *In the setting of Proposition 12.1, let's introduce some $d \times d$ mean reversion matrix $k(t)$ and assume that $J_k(t)$ as in (36) exists and is invertible for all t. Then*

$$r(t) = f(0,t) + \Omega(t,t) + h(t)^\top J_k(t)^{-1} x(t)$$

where
$$dx(t) = -k(t)x(t)dt + \sigma_x(t)^\top dW(t), \sigma_x(t) = g(t)J_x(t)^\top$$

Proof. Set $x(t) = J_k(t)z(t)$ such that $z(t) = J_k(t)^{-1}x(t)$. Apply the Ito formula for $x(t)$, we get

$$dx(t) = z(t)dJ_k(t) + J_k(t)dz(t) \qquad = -k(t)x(t)dt + J_k(t)g(t)^\top dW(t)$$

as we have $dJ_k(t)/dt = -k(t)J_k(t)$.

The lemma shows that we can incorporate essentially any mean reversion matrix k into the basic martingale setup in Proposition 5.1. Once we choose the form of k, J_k is determined, so if we have g, σ_x is also determined, then the whole process for the short rate is constructed. For numerical applications, the best choice of mean reversion is typically one that leaves both $k(t)$ and $\sigma_x(t)$ close to constant.

5.3 The two-factor Gaussian model

Given the general Gaussian model, here we specify a two-factor Gaussian model. In the setting of Proposition 5.1, assume $\varkappa(t) = diag((\varkappa_1(t), \varkappa_2(t))^\top)$, then we have

$$h(t) = \begin{pmatrix} e^{-\int_0^t \varkappa_1(u)du} \\ e^{-\int_0^t \varkappa_1(u)du} \end{pmatrix}, g(t) = \begin{pmatrix} \sigma_{11}(t)e^{\int_0^t \varkappa_1(u)du} & \sigma_{12}(t)e^{\int_0^t \varkappa_2(u)du} \\ \sigma_{21}(t)e^{\int_0^t \varkappa_1(u)du} & \sigma_{22}(t)e^{\int_0^t \varkappa_2(u)du} \end{pmatrix}$$

The form of $g(t)$ is chosen in the way that the volatility of $x(t)$ could be written as a simple matrix. We set $\sigma_{12}(t) = 0$ because if it is not the case, we could change variables via Cholesky decomposition to make it so. And we have

$$r(t) = f(0, t) + x_1(t) + x_2(t)$$

where $x(t) = (x_1(t), x_2(t))^\top$ satisfies

$$x(0) = 0, dx(t) = (y(t)\mathbf{1} - \varkappa(t)x(t))dt + \sigma_x(t)^\top dW(t), \sigma_x(t) = \begin{pmatrix} \sigma_{11}(t) & 0 \\ \sigma_{21}(t) & \sigma_{21}(t) \end{pmatrix}$$

Notice that the instantaneous correlation between $x_1(t)$ and $x_2(t)$ is

$$\rho_x(t) = \frac{\sigma_{22}(t)\sigma_{21}(t)}{\sqrt{\sigma_{11}(t)^2 + \sigma_{21}(t)^2}\sqrt{\sigma_{22}(t)^2}}$$

so for convenience we can rewrite our model to change the variance matrix to a diagonal
$$dx(t) = (y(t)\mathbf{1} - \varkappa(t)x(t))dt + \sigma_x^* dW^*(t)$$

where $\langle dW_1^*(t), dW_1^*(t\rangle = \rho_x(t)dt$ and $\sigma_x^*(t)$ is diagonal with non-negative elements,

$$\sigma_x^*(t) = \begin{pmatrix} \sqrt{\sigma_{11}(t)^2 + \sigma_{21}(t)^2} & 0 \\ 0 & |\ \sigma_{22}(t)\ | \end{pmatrix} := diag((\sigma_1(t), \sigma_2(t))^\top)$$

And we can construct the yield curve

$$
\begin{aligned}
f(t,T) &= f(0,T) + M(t,T)^\top (x(t) + y(t)G(t,T)),\\
P(t,T) &= \frac{P(0,T)}{P(0,T)} exp(-G(t,T)^\top x(t) - \frac{1}{2}G(t,T)^\top y(t)G(t,T))
\end{aligned}
$$

where

$$
G(t,T) = \int_t^T M(t,u)du, M(t,T) = (e^{-\int_t^T \varkappa_1(u)du}, e^{-\int_t^T \varkappa_2(u)du})
$$

Note that at the beginning we have $\varkappa_1(t), \varkappa_2(t)$ and $\sigma_x(t)$ i.e. $\sigma_{11}(t), \sigma_{21}(t), \sigma_{22}(t)$ to complete this model. After rewrite our model we need specify 5 functions of time: $\rho_x(t)$, $\varkappa_1(t), \varkappa_2(t), \sigma_1(t)$ and $\sigma_2(t)$.

5.4 Variance and correlation structure

One important motivation for the introduction of a multi-factor interest rate model is the ability to control correlations among various points on the forward curve. Let $\rho(t, T_1, T_2)$ denote the time t instantaneous correlation between the forward rates $f(t, T_1)$ and $f(t, T_2)$. From the representation in Proposition 5.1, we get the following result for the correlation

$$
\rho(t, T_1, T_2) = \frac{h(T_1)^\top g(t)^\top g(t) h(T_2)}{\sqrt{h(T_1)^\top g(t)^\top g(t) h(T_1)}\sqrt{h(T_2)^\top g(t)^\top g(t) h(T_2)}}
$$

In a particular model, we generally would strongly prefer for the correlation structure to be time-stationary, in the sense that ρ does not depend outright on t, but only on time to maturity $T_1 - t$ and $T_2 - t$, i.e.

$$
\rho(t, T_1, T_2) = \rho(T_1 - t, T_2 - t)
$$

To find out the constraints imposed from this restriction, we know that the forward rate process satisfies

$$
df(t,T) = O(dt) + \begin{pmatrix} \sigma_1(t)e^{-\int_t^T \varkappa_1(u)du} \\ \sigma_2(t)e^{-\int_t^T \varkappa_2(u)du} \end{pmatrix}^\top dW^*(t)
$$

From this representation, we get the following lemma
Lemma 5.5 *For the model (5.3), let*

$$
\begin{aligned}
b(t, T_1, T_2) = 1 &+ \rho_x \frac{\sigma_2(t)}{\sigma_1(t)}(e^{\int_t^{T_1}(\varkappa_2(u) - \varkappa_1(u))du} + e^{\int_t^{T_2}(\varkappa_2(u) - \varkappa_1(u))du})\\
&+ (\frac{\sigma_2(t)}{\sigma_1(t)})^2 e^{\int_t^{T_1}(\varkappa_2(u) - \varkappa_1(u))du - \int_t T_2(\varkappa_2(u) - \varkappa_1(u))du}
\end{aligned}
$$

Then

$$
\begin{aligned}
Var_t(df(t,T)) &= \sigma_1(t)^2 e^{-2\int_t^T \varkappa_1(u)du} b(t,T,T),\\
\rho(t, T_1, T_2) &= Corr_t(df(t,T_1), df(t,T_2)) = \frac{b(t,T_1,T_2)}{\sqrt{b(t,T_1,T_1)b(t,T_2,T_2)}}
\end{aligned}
$$

In particular, $\rho(t, T_1, T_2)$ is time-stationary if $\rho_x(t)$, $\varkappa_2(t) - \varkappa_1(t)$ and $\sigma_2(t)/\sigma_1(t)$ are all constant.
This lemma can be easily driven from the expression of $b(t, T_1, T_2)$. Then we have only two functions of time (say $\sigma_1(t)$ and $\varkappa_1(t)$) and three constants to be specified freely, rather than five functions of time as we mentioned in the precedent sector.

5.5 Volatility hump

Besides allowing us to properly model the correlation between various points of the forward curve, the use of two-factor Gaussian model is also able to produce time-stationary and non-monotonic term structure of forward rate volatilities.

Here to show the ability of two-factor model to provide this volatility hump, assume that $\varkappa_1$ and $\varkappa_2$ are fixed non-negative constant values, with at least one being positive, and σ_1 and σ_2 fixed constant positive values. From Lemma 12.5 we get

$$\begin{aligned} Var_t(df(t,T)) &= \sigma_1^2 e^{-2\varkappa_1(T-t)} + \sigma_2^2 e^{-2\varkappa_2(T-t)} + 2\rho_x \sigma_1 \sigma_2 e^{-(\varkappa_1+\varkappa_2)(T-t)} \\ &:= g(T-t) = g(\tau) \end{aligned}$$

where

$$\frac{1}{2}\frac{dg(\tau)}{d\tau} = -\varkappa_1 \sigma_1^2 e^{-2\varkappa_1 \tau} - \varkappa_2 \sigma_2^2 e^{-2\varkappa_2 \tau} - \rho_x \sigma_1 \sigma_2 (\varkappa_1 + \varkappa_2) e^{-(\varkappa_1+\varkappa_2)\tau}$$

Obviously for positive values of ρ_x, the forward rate variance term structure will thus always be downward-sloping. However, if we set ρ_x sufficiently negative, there may be intermediate values for which the variance will increase in τ, and we realize the volatility hump.

6 Multi-Factor statistical gaussian model

We can rewrite the equation of the volatility hump as:

$$df(t, t+\tau) = O(dt) + l(\tau)dz(t)$$

Where:

$$l(\tau) = \frac{\varkappa_r}{\varkappa_r - \varkappa_\varepsilon}(e^{\varkappa_\varepsilon \tau} - e^{\varkappa_r \tau})$$

$$dz(t) = \sigma_\varepsilon dW_\varepsilon(t)$$

We can interpret $z(t)$ as the single factor that affects the movements of the forward rate curve $f(t, t+\tau)_{\tau \geq 0}$ and the function $l(\tau)$ as the response function or a loading whose value at time τ determines the impact of the factor shock on a rate of tenor τ only. This justifies why we use a statistical approach to estimate the movements of the yield curve.
There follow some tools we will be using:

Stone-Weierstrass Theorem Any continuous function on $[0,T]$ can be uniformly approximated by functions of the set $\{\tau \longrightarrow e^{-\varkappa\tau}\}_{\varkappa\in R}$. That being said, for any function $l(\tau)$, there exists scalars $\{v_i\}_i$ and exponents $\{\varkappa_i\}_i$ such that:

$$l(\tau) = \sum_{i=1}^{n} v_i e^{-\varkappa_i\tau}$$

Thus considering the previous, we can see that the loading can be represented as an n-state single brownian motion Gaussian model as formalized in the next proposition:

$$df(t,t+\tau) = O(dt) + l(\tau)dz(t),$$
$$l(\tau) = \sum_{i=1}^{n} v_i e^{-\varkappa_i\tau}, dz(t) = \sigma_1(t)dW_1(t)$$

Where $W_1(t)$ is a one-dimensional Brownian motion and $\sigma_1(t)$ is a one-dimensional function of time. Then this model model admits a Markovian representation in n state variables.

$$r(t) = f(0,t) + \sum_{i=1}^{n} x_i(t)$$

Where with $x(t) = (x_1(t),...,x_n(t))^T$ and $\varkappa = diag((x_1,...,x_n)^T)$, we have

$$dx(t) = (y(t)1 - \varkappa x(t))dt + \sigma_x(t)^T dW(t),$$
$$\sigma_x(t) = \sigma_1(t)\begin{pmatrix} v_1 & v_2 & ... & v_n \\ 0 & 0 & ... & 0 \\ . & . & ... & 0 \\ . & . & ... & 0 \\ 0 & 0 & ... & 0 \\ 0 & 0 & ... & 0 \end{pmatrix}, W(t) = (W_1(t),0,0,...,0)^T,$$

and

$$y(t) = H(t)(\int_0^t \sigma_1(s)^2 H(s)^{-1} U H(s)^{-1} ds)H(t),$$
$$H(t) = diag((e^{-\varkappa_1 t},...,e^{-\varkappa_n t})^T),$$
$$U = v_k v_j{}_{k,j=1}^{n}$$

Proof We have:
$df(t,T) = O(dt) + \sum_{i=1}^{n} e^{-\varkappa_i T}(e^{\varkappa_i t}\sigma_1(t)v_i dW_i(t))$ Hence the model can be written in a separable form with
$h(t) = ((e^{-\varkappa_1 t},...,e^{-\varkappa_n t})^T), g(t) = \sigma_x(t)H(t)^{-1}$
$H(t)$ and $\sigma_x(t)$ being given in the statement of the proposition and obtained using the result of separability case 12.1.2 to y(t) and the fact that:

$$\sigma_x(t)^T\sigma_x(t) = \sigma_1(t)^2 U$$

The model allows for an arbitrary loading function $l(\tau)$ but employs only one factor to describe the dynamics of an interest rate curve. Assume we are given m loading functions, each describing the (linear) response of the forward rate curve to a given factor. Approximating each loading by a linear combination of exponential, we arrive at a model of the form:

$$
\begin{aligned}
df(t, t+\tau) &= O(dt) + \sum_{j=1}^{m} l_j(\tau) dz_j(t), \\
l_j(\tau) &= \sum_{j=1}^{n_j} v_{j,i} e^{-\varkappa_{j,i}\tau}, \\
dz_j(t) &= \sigma_j(t) dW_j(t),
\end{aligned}
$$

This process is thus Markovian with a total number of states $n = \sum_{i=1}^{m} n_j$
As for the one-dimension case, this model admits a Markovian representation

$$r(t) = f(0,t) + 1^T x(t),$$

Where $x(t) = (x_1(t), ..., x_n(t))^T$ satisfies

$$dx(t) = (y(t) - \varkappa x(t))dt + \sigma_x(t) dW(t),$$

with

$$
\begin{aligned}
\varkappa &= diag((\varkappa_{1,1}, ..., \varkappa_{1,n_1}, \varkappa_{2,1}, ..., \varkappa_{2,n_2}, ..., \varkappa_{m,n_m})^T), \\
\sigma(t) &= diag((\sigma_1(t), ..., \sigma_1(t), ..., \sigma_2(t), ..., \sigma_m(t))^T), \\
h(t) &= (e^- \varkappa_{1,1} t, ..., e^- \varkappa_{1,n_1} t, e^- \varkappa_{2,1} t, ..., e^- \varkappa_{m,n_m} t)^T
\end{aligned}
$$

and $\sigma_x(t) = v\sigma(t)$
where

$$
v = \begin{pmatrix}
v_{1,1} & ... & v_{1,n_1} & 0 & ... & 0 & ... & 0 & ... & 0 \\
0 & ... & 0 & v_{2,1} ... & v_{2,n_2} & ... & ... & ... & ... & ... \\
... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\
... & ... & ... & ... & ... & ... & ... & ... & ... & ... \\
... & ... & ... & ... & ... & ... & ... & 0 & ... & 0 \\
0 & ... & 0 & ... & 0 & ... & . & v_{m,1} & ... & v_{m,n_m}
\end{pmatrix}
$$

$W(t)$ being a m-dimensional vector of independent Brownian motions
$W(t) = (W_1(t), ..., W_m(t))^T$
and
$y(t) = H(t)(\int_0^t H(s)^-1\sigma(s)v^T\sigma(s)H(s)^-1 ds)H(t)$, $H(t) = diag(h(t))$
The representation (97) allows us to link the interest rate model parametrization to stastitical properties of the movements of the yield curve. In fact, if we fix N_τ, the number of tenors of interest, and specify a set of tenors $\{\tau_1, ..., \tau_{N_\tau}\}$, then we can observe from history how the vector of rates $f(t) = (f(t, t+\tau_1), ..., f(t, t+\tau_{N_\tau}))^T$ changes over time. With the application of principal components analysis, we can identify $m \leq N_\tau$ factors $\zeta(t) = (\zeta_1(t), ..., \zeta_m(t))^T$ and m loadings $\lambda_j = (\lambda_j(\tau_1), ..., \lambda_j(\tau_{N_\tau}))$ that we use to represent:

$$\Delta f(t) \approx \sum_{j=1}^{m} \lambda_j \Delta\zeta_j(t);$$

7 Some practical cases

Let's now consider some practical cases where we can easily price products on rates and compute calibration processes.

7.1 Hull White 1 factor model

The Hull and White model assumes that, under the risk neutral measure, the short rate follows the dynamic (with a constant volatility parameter):

$$dr(t) = (\theta(t) - ar(t))\,dt + \sigma dW(t)$$

If we take $\gamma(s,t) = -\sigma(s)exp(-a(t-s))$, we can check that the HJM Model is equivalent to the Hull and White Model.

The equations representing the short rate and the forward rate in the Hull and White Model can be deduced from the equations presented above in the more generic case of the HJM model. We get the following equations:

$$r_t = f(0,t) + \frac{\sigma^2}{2a^2}\left[1 - exp(-at)\right]^2 + \sigma exp(-at)\int_0^t exp(as)dW_s$$

$$\begin{aligned} f(t,T) =& f(0,T) + \frac{\sigma^2}{a^2}exp(at)\left[exp(-at) - 1\right] - \frac{\sigma^2}{2a^2}exp(-2at)\left[exp(2at) - 1\right] \\ &+ \sigma exp(-aT)\int_0^t exp(as)dW_s \end{aligned}$$

Beta represents the money market account, that is the amount of money that one would get by placing 1 unit of currency at the risk neutral rate. We calculate the Beta by the following equation:

$$\beta(t) = exp(\int_0^t r_s ds)$$

The value of a Zero coupon bond at t can be calculated using the formula:

$$B(t,T) = exp(-\int_t^T f(t,s)ds)$$

For the integral calculation, we will discretise the integral and to compute the values of a finite number to approximate it.

A swap curve is given as a first input to calibrate this model. By interpolation and bootstrapping, it is transformed into a Zero Coupon curve where a Zero Coupon value is known for every maturity by time steps of 1 month. This allows us to compute the forward rate $f(0,t)$, as the derivative of this Zero Coupon curve.

7.1.1 Numerical implementation

We have implemented in C++ through Monte-Carlo methods, the pricing of a call on Zero Coupon. Basically, we use the dynamics of the rates as described in the Hull-White one factor model to calculate the coupons and derive the price of call on ZC coupon. We compare this price to the price of closed formula we can derive by the analytic calculus of the price of the call. As standard input parameters we used a number of Monte-Carlo paths of 500, a strike of K = 0.7, the option maturity being 2 years and the zero coupon maturity being 5 years. The results can be summarized in the following graphs:

We can see that though for obvious reasons of calculus time, we chose a number

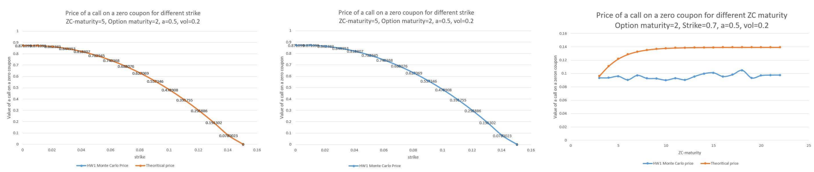

Figure 1: Comparison of MC and the analytic closed formula prices of a ZC call with strikes and spreads

of MC paths of 500, the prices with the MC method and those with the analytic closed formulas matched quite well. And we can retrieve general behaviour of common markets prices curves of ZC calls such as the non-increasing growth in the strike K.

7.2 The G1++ model

The simple case of a rate depending on one factor is the G1++ model, as suggested by the general framework, the dynamics of the instantaneous short-rate under the risk-neutral measure is:

$$r(t) = x(t) + f(t)$$

Where the process $(x_t)_{t\geq 0}$ satisfies the stochastic differential equation

$$dx(t) = -ax(t)dt + \sigma(t)dW(t), x(0) = 0$$

with a being a positive constant and $\sigma(t)$ a deterministic function of time. Integrating this equation, we get

$$r(t) = x(s)e^{-a(t-s)} + \int_s^t \sigma(s)e^{-a(t-u)}dW(u) + f(t)$$

Therefore conditional to F_s the information available at time s, the market up to time s in normally distributed with mean and variance given by:

$$E[r(t)|F_s] = x(s)e^{-a(t-s)} + f(t)$$

$$VAR[r(t)|F_s] = \sigma^2 \int_s^t e^{-2a(t-u)} du = \frac{\sigma^2}{2a}(1 - e^{-2a(t-s)})$$

We denote by $P(t,T)$ the price at time t of a zero coupon bond maturing at time T with unit nominal value. The existence value of risk-neutral measure Q implies that

$$P(t,T) = E[e^{\int_t^T r(s)ds}|F_t]$$

Given the previous distribution of $r(t)$ we can derive that

$$P(0,T) = E[e^{-\int_0^T r(t)dt}] = E[e^{\int_0^T x(t)+f(t)dt}] = e^{-\int_0^T f(t)dt}E[e^{-\int_0^T x(t)dt}]$$

Using the Fubini's theorem for stochastic integrals we can deduce that

$$P(0,T) = e^{-\int_0^T f(t)dt} e^{\frac{1}{2}V(0,T)}$$

Where we have

$$V(0,T) = \frac{\sigma^2}{a^2}(T - 2\frac{1 - e^{-aT}}{a} + \frac{1 - e^{-2aT}}{2a})$$

For a a bond starting at t and with maturity T, we have a similar formula for the price

$$P(t,T) = e^{-\int_0^T f(s)ds} e^{\int_0^t f(s)ds} e^{-((1-e^{-a(T-t)})/a)x(t)} e^{1/2V(t,T)}$$

7.2.1 Pricing and calibration in the G1++

If we want to fit the model to the market price, we can use the market instantaneous forward rate $f^M(t)$ defined as $P^M(0,T) = e^{-\int_0^T f^M(t)dt}$ and we can rewrite the bond as

$$P(t,T) = \frac{P^M(0,T)}{P^M(0,t)} e^{\frac{1}{2}(V(t,T)-V(0,T)+V(0,t))} e^{-((1-e^{-a(T-t)})/a)x(t)}$$
$$= A(t,T)e^{-B(t,T)x(t)}$$

Where,

$$A(t,T) = \frac{P^M(0,T)}{P^M(0,t)} e^{\frac{1}{2}(V(t,T)-V(0,T)+V(0,t))}$$

,

$$B(t,T) = (1 - e^{-a(T-t)})/a$$

For simulation, we can thus generate $P(t,T)$ only by using the market discount curve and the process x(t). And from P(t,T) we can derive all the other rates (forward rates, swap rates etc.).
In order to derive an explicit formula of the price of an European Swaption, we need to change the measure. Let's denote by Q^T the T-forward measure defined

by using a zero coupon bond with maturity T as numeraire. By the Girsanov theorem the dynamics is now

$$dx(t) = -ax(t)dt - \sigma^2(t)B(t,T)dt + \sigma dW(t),$$

Now if we want to price a European swaption with maturity T, strike K and nominal value N which gives the holder the right to enter at time $T = t_0$ into a swap with payment dates $(t_1, ..., t_k), t_0 < t_1 < ... < t_k$ where he pays the fixed rate K and receive the Libor rate . The price of the swaption is thus at t_0:

$$ES(0,T,t_k,K,N) = NP(0,T)E^T[(1 - \sum_{i=1}^{k} c_i P(T,t_i))^+ | F_0]$$

The calculation in the case where $\sigma(t)$ is a constant gives:

$$ES(0,T,t_k,K,N) = NP(0,T)(\Phi(-\bar{y_1}) - \sum_{i=1}^{k} c_i A(T,t_i) e^{-B(T,t_i)M(0,T)} e^{B^2(T,t_i)(\sigma^2(0,T)/2)} \Phi(-\bar{y_2}(i)))$$

Where Φ is the standard normal cumulative function, $\bar{y_1}$ and $\bar{y_2}(i)$ are given by:

$$\bar{y_1} = \frac{\bar{y} - M(0,T)}{\sigma(0,T)},$$

$$\bar{y_2}(i) = \bar{y_1} + B(T,t_i)\sigma(0,T)$$

Where $\bar{y}$ is the unique solution of

$$\sum_{i=1}^{k} c_i A(T,t_i) e^{-B(T,t_i)\bar{y}} = 1$$

7.2.2 Numerical results for G1++ Model

- Pricing swaptions:
  We implemented both Monte-Carlo and closed formula pricing for swaptions.
  For the closed formula, we used Newton Algorithm to compute $\bar{y}$ as explained above. We are sure to obtain a global minimum since the equation to minimize is convex function of $\bar{y}$.
  The following graph shows the put 5Y/ 10Y swaption prices obtained with the closed formula for different strike rates.

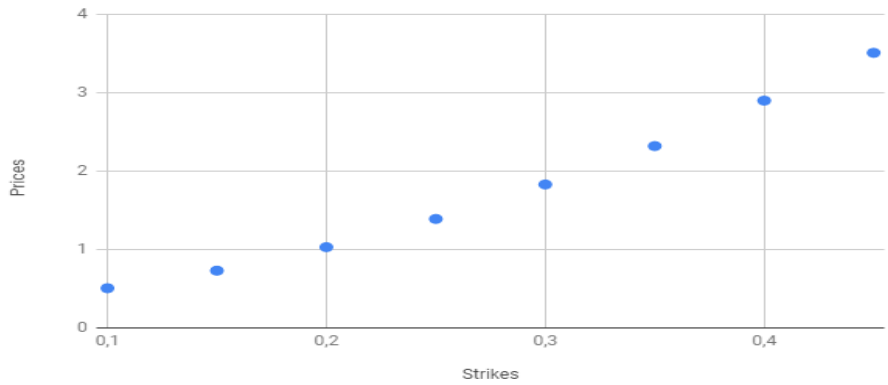

Figure 2: Theoretical prices of 5Y/10Y put swaption

- Model Calibration:
The objective of the calibration is to estimate the mean reversion coefficient a and the volatility σ considered as constant. The calibration is performed with minimizing the mean squared error between market prices and model prices:

$$MSE = \Sigma_i (PV_i^{mod} - PV_i^{mkt})^2$$

For numerical applications, minimization is performed using the gradient descent algorithm with a constant learning rate, partial derivatives being calculated with the finite difference method.
Let's point out the importance of initial guesses for the model parameters, since this method doesn't ensure finding a global minimum.
Actually, for numerical applications, we set model's parameters as follows:

$$a = 0.5$$
$$\sigma = 0.2$$

Then, we compute theoretical prices for 5 swaptions with same maturity and tenor but with different strikes, that we consider as our market data. afterwards, we make initial guesses of the parameters:

$$a_0 = 0.2$$
$$\sigma_0 = 0.07$$

After performing the gradient descent algorithm as described above, we obtain after approximately 2000 iterations an SDE of 10^{-4} order and optimal models parameters as follows:

$$a_{opt} = 0.48$$
$$\sigma_{opt} = 0.17$$

Let's point out that we tested our calibration using different strike and maturities for swaptions and different initials guesses to validate the efficiency of our calibration.

7.3 Pricing and calibration in the G2++

The motivation to use the a two factor rate model comes from the fact that in an affine term-structure model, $f(t, T_1)$ and $f(t, T_2)$ with $T_1 = t + 1$ and $T_2 = t + 100$ (short rate and long rate) are perfectly correlated, meaning their correlation is equal to 1, which is not realistic.
As introduced in the general framework, in the case of the G2++, we have

$$r(t) = x_1(t) + x_2(t) + \phi(t)$$

where ϕ is deterministic and x_1 , x_2 are assumed to satisfy the stochastic problems

$$dx_1(t) = -k_1 x_1(t)dt + \sigma_1 dW_1(t), x_1(0) = 0,$$
$$dx_2(t) = -k_2 x_2(t)dt + \sigma_2 dW_2(t), x_2(0) = 0$$

where k_1, k_2, σ_1, $\sigma_2 > 0$, and W_1, W_2 are brownian motions under the risk-neutral measure such that

$$dW_1(t)dW_2(t) = \rho dt,$$

with $\rho \in [-1; 1]$
The short rate is then given by

$$r(t) = x_1(s)e^{-k_1(t-s)} + x_2(s)e^{-k_2(t-s)} + \phi(t) + \sigma_1 \int_s^t e^{-k_1(t-u)} dW_1(u) + \sigma_2 \int_s^t e^{-k_2(t-u)} dW_2(u)$$

As it was shown in the G1++, we can deduce that the price of the zero-coupon bond in the G2++ is

$$P(t, T) = exp(-\int_t^T \phi(u)du - M(t, T) + \frac{1}{2}V^2(t, T))$$

where,

$$M(t, T) = x_1(t)B_1(t, T) + x_2(t)B_2(t, T)$$

and

$$V^2(t, T) = \frac{\sigma_1^2}{k_1^2}(T - t - B_1(t, T) - \frac{1}{2}B_1^2(t, T) + \frac{\sigma_2^2}{k_2^2}(T - t - B_2(t, T) - \frac{1}{2}B_2^2(t, T) + \frac{2\sigma_1\sigma_2\rho}{k_1 k_2}(T - t - B_1(t, T) - B_2(t, T)B$$

where,

$$B_1(t, T) = \frac{1 - e^{-k_1(T-t)}}{k_1}, B_2(t, T) = \frac{1 - e^{-k_2(T-t)}}{k_2}$$

and,

$$B_{12}(t,T) = \frac{1 - e^{-(k_1+k_2)(T-t)}}{k_1 + k_2}$$

Using the T-forward measure, we can derive the price of a European call option with strike K and maturity T and written on a zero-coupon bond with maturity S at time $t \in [0,T]$ which is given by

$$ZBC(t,T,S,K) = P(t,S)\Phi(h) - KP(t,T)\Phi(h - \hat{\sigma}),$$

where

$$\hat{\sigma} = \frac{\sigma_1^2}{2k_1}(1-e^{-2k_1(T-t)})B_1^2(T,S)+\frac{\sigma_2^2}{2k_2}(1-e^{-2k_2(T-t)})B_2^2(T,S)+2\sigma_1\sigma_2\rho B_1(T,S)B_2(T,S)B_{12}(t,T)$$

and

$$h = \frac{1}{\hat{\sigma}} ln(\frac{P(t,S)}{P(t,T)K}) + \frac{\hat{\sigma}}{2}$$

For European swaptions, the price at time 0 of a European swaption of maturity $T_\alpha = T$ on an IRS depending on the notional value N, the fixed rate K, and the set of times $\mathcal{T} = \{T_{\alpha+1}, ..., T_\beta\}$ is given by

$PS(0,T,\mathcal{T},N,K) = NP(0,T)\int_{-\infty}^{\infty} \frac{e^{\frac{-(x-\bar{\mu}_1)^2}{2\bar{\sigma}_1^2}}}{\bar{\sigma}_1\sqrt{2\pi}}[\Phi(-h_1(x)) - \sum_{i=\alpha+1}^{\beta} \lambda_i(x)e^{\kappa_i(x)}\Phi(-h_2(x))]dx$,
Where $\mu_1, h_1, \lambda_i, \sigma_1, h_2, \kappa$ are determistic functions of known parameters.

7.3.1 Numerical results for G2++ model

We implement both the theoretical pricing formula and the monte carlo procedure for European call on zero-coupon and for swaption. The model parameters are $k_1 = 0.2$, $k_2 = 0.2$, $\sigma_1 = 0.05$, $\sigma_2 = 0.05$, $\rho = -0.7$. We set the zero-coupon maturity as 10 years, the option maturity from 1 year to 8 years, the strike from 0.5 to 0.8. The prices we obtained by theoretical formula and monte carlo respectively are as in the following graphs Figure 2 and Figure 3

The theoretical prices and the monte carlo prices are coherent, and the variation of the prices with respect to strike and maturity is logical. We also implement the theoretical fomula and the monte carlo for the swaption price, the two prices are coherent, but as it takes long time to run the code, we don't present the graphs of the prices here.

The target of the calibration is to find the parameters to minimize the mean squared error between model prices and market prices, i.e. $\Sigma_i(PV_i^{mod} - PV_i^{mkt})^2$.

As both the procedures of calculation of theoretical swaption price and the monte carlo simulation are too slow, in G2++ we use the European call option on zero-coupon to do the calibration of the model. As we have five parameters

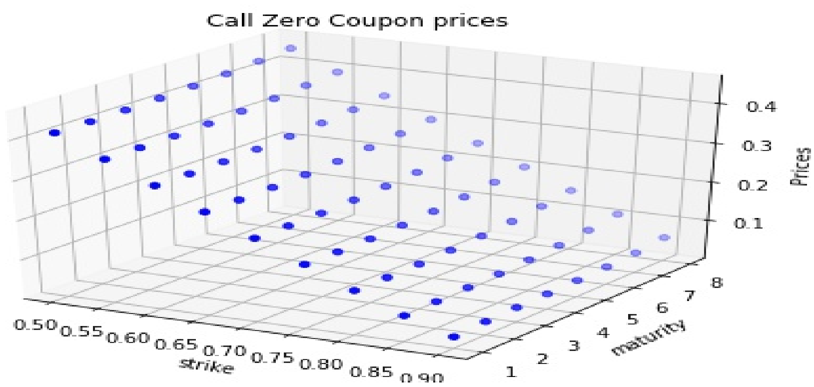

Figure 3: Theoretical prices of Call on zero-coupon

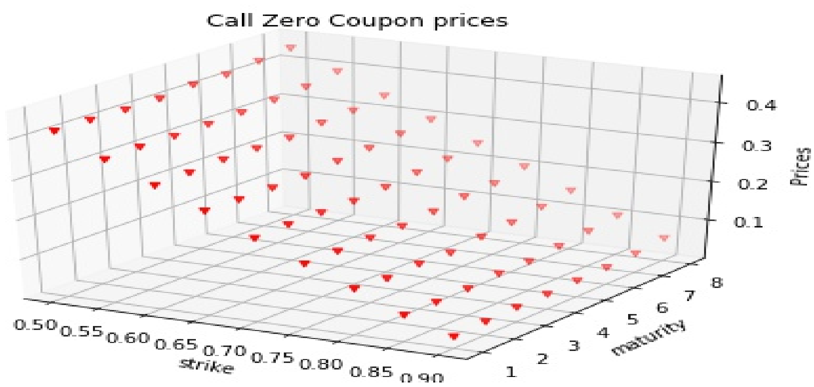

Figure 4: Monte Carlo prices of Call on zero-coupon

to calibrate in G2++, and different combination of parameters can give very close prices, here to verify the validation of our calibration, we first fix the parameters k_2, σ_1, σ_2, ρ. We set $k_1 = 0.1$, $k_2 = 0.2$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$, $\rho = 0.5$, for the calibration we choose two call on zero-coupon with the same weight: the first zero-coupon maturity is 4 years, option maturity is 2 years, and strike 0.95; the second zero-coupon maturity is 6 years, option maturity is 2 years, and strike 0.9. The prices given by our model with the above parameters are respectively 0.302888 and 0.575440. We input these prices as market prices into our calibration procedure, and we fix $k_2 = 0.2$, $\sigma_1 = 0.2$, $\sigma_2 = 0.3$, $\rho = 0.5$ and adjust k_1 to minimize the mean squared error. The procedure is down by the gradient descent method, the learing rate is set as 0.05. We start from $k_1 = 0.5$ and after 1000 steps we refind $k_1 = 0.0999502$, which is quite close to the correct parameter, and the mean squared error is $1.53752e-09$. The validation of our calibration is verified.

Then we modify the four parameters $k_1, k_2, \sigma_1, \sigma_2$ to do the calibration. As negative correlation in calibration could cause no valid square root in the theoretical pricing formula, we still fix $\rho = 0.5$. We start from $k_1 = 0.5, k_2 = 0.5, \sigma_1 = 0.5, \sigma_2 = 0.5$ and after 1000 steps we get the optimal parameters $k_1 = 0.116692, k_2 = 0.202299, \sigma_1 = 0.249934, \sigma_2 = 0.250382$. The mean squared error is $2.00201e-07$. Although our results are different from the original ones, we still constate that our calibration is valid, because as we mentioned before different combination of parameters could lead to close result, and the mean squared error is small enough.

Short rate models with only one single driving Brownian motion imply that the instantaneous correlation between forward rates at different maturities is one, which is contrary to reality. So one important motivation for the introduction of a multi-factor interest rate model is the ability to control correlations among various points of interest rate curve. To demonstrate that, we calculate and draw the correlation term structures generated by our two-factor Gaussian model.

We set $k_1 = 0.2, k_2 = 0.2, \sigma_1 = 0.5, \sigma_2 = 0.5$. We change the parameter ρ to see the change of correlation term structure of zero-coupon. The following figures show the correlation term structures of zero-coupon with different ρ. Let $\rho(t, T_1, T_2)$ denotes the time t instantaneous correlation between the zero-coupon $P(t, T_1)$ and $P(t, T_2)$.

In Figure 4 we fix $T_1 = 5y$ and $T_2 = 8y$ and we change the t. As t increases the correlation between the two zero coupons decreases. When $\rho = -0.9$, the correlation term structure is rather flat, because at that time the two-factor model degrades to a nearly one factor model.

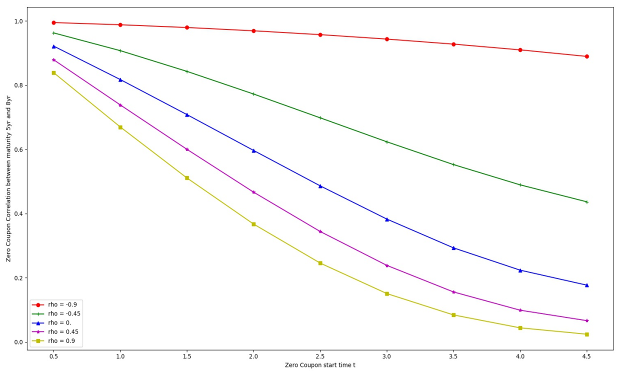

Figure 5: Correlation term structure when changing t

In Figure 5 we fix $t = 1y$ and $T_1 = 2y$ and we change the T_2. We can see as the zero-coupon maturity difference increases, the correlation between the two zero coupons decreases, and when the difference becomes small, the correlation is close to 1, which accords with the theoretical analysis. Again, when $\rho = -0.9$, the correlation term structure becomes rather flat.

In Figure 6 we fix $t = 2, T_1 = 3, T_2 = 5, k_1 = 0.05, \sigma_1 = 0.5, \sigma_2 = 0.5$ and change k_2 to see the correlation term structure with different ρ. We can see that different parameters of the model lead to different shapes of correlation term structure. So we can manipulate the parameters to get the required correlation

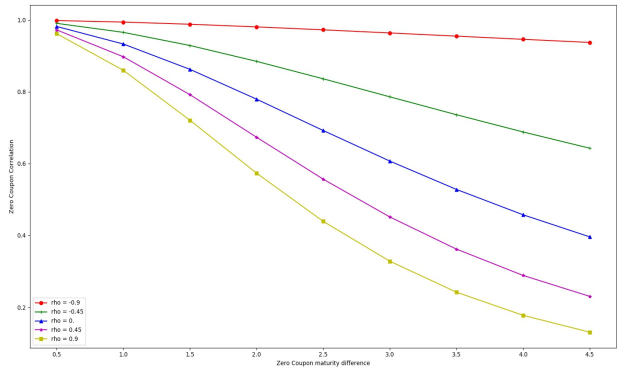

Figure 6: Correlation term structure when changing $T_2 - T_1$

term structure.

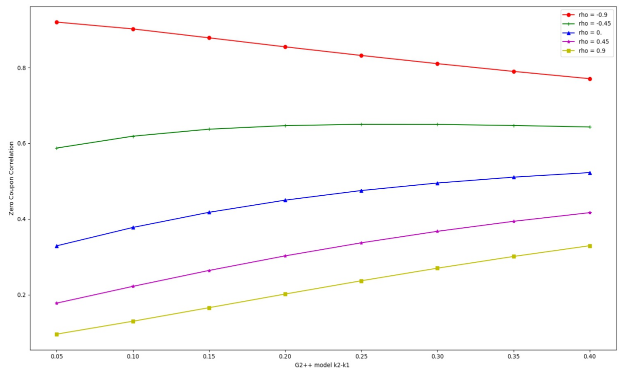

Figure 7: Correlation term structure when changing $k_2 - k_1$

In Figure 7 we fix $t = 2, T_1 = 3, T_2 = 5, k_1 = 0.2, k_2 = 0.2, \sigma_1 = 0.1$ and change $sigma_2$ to see the correlation term structure with different ρ. We find that when the difference between σ increases, the correlation between zero-coupons decreases.

From these analyzes, the G2 ++ model allows control correlations between various points of the interest rate curve. This would be more appropriate for pricing hybrid equity rate products requiring the observation of multiple interest rates.

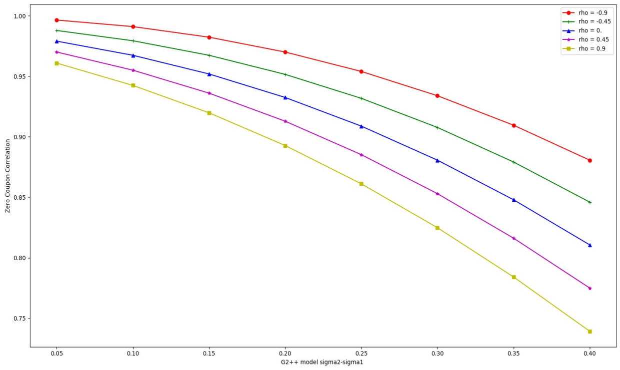

Figure 8: Correlation term structure when changing $sigma_2 - sigma_1$

8 Models enhanced with local volatility

Even if the two factor model developed above represents a more realistic modeling of zero coupon correlations, the model is not able to capture the whole market smile.
In this part, we consider an extension of the one and two factor gaussian short rate with local volatility to a better fit of the smile.

8.1 Cheyette Model

Its shown in 5, that a one factor Markovian Gaussian Model is obtained by imposing a separability condition on the volatility structure of the instantaneous forward rates. The Cheyette model is obtained by retaining the separability condition but with allowing the function g to be stochastic:

$$\sigma_f(t, T, w) = g(t, w)h(T)$$

Given the separability condition above, the Cheyette model is defined by the process $x(t)$ and $y(t)$ such as:

$$\begin{aligned} dx_t &= (y_t - ax_t)dt + \sigma_t dW_t \\ dy(t) &= (\sigma_t^2 - 2ay_t)dt \\ r_t &= f(0, t) + x_t \end{aligned}$$

σ_t being a stochastic process and a the mean reversion coefficient considered as constant.
The price of the zero coupon bond under Cheyette model is given by:

$$P_{t,T} = \frac{P_{0,T}}{P_{0,t}} e^{\left(\frac{e^{-a(T-t)}-1}{a}\right)x_t - \frac{1}{2}\left(\frac{e^{-a(T-t)}-1}{a}\right)^2 y_t}$$

8.2 The volatility structure

We consider a swap rate $s_t^{k,m}(t, x(t), y(t))$ between maturities T_k and T_m. Under the swap measure $Q^{k,m}$:

$$ds_t^{k,m} = \frac{\partial s_t^{k,m}}{\partial x} \sigma_t dW^{k,m}$$

By setting:

$$\sigma(t, s_t^{k,m}) = \frac{\partial s_t^{k,m}}{\partial x} \sigma_t$$

we get:

$$ds_t^{k,m} = \sigma(t, s_t^{k,m}) dW^{k,m}$$

8.3 Calibration

To calibrate the local volatility, we suggest a non parametric method based on the formula showed in [5]. The authors developed a Dupire like formula for swaption prices $C(t, K)$ (t being the maturity and K the strike rate) on a swap $s_t^{t,t+\theta}$

$$\sigma(t, K)^2 = \sigma_{loc}(t, K)^2 + 2\frac{E(t, K)}{\partial_K^2 C(t, K)}$$

with

$$\sigma_{loc}(t, K)^2 = 2\frac{\partial_t C^{mkt}(t, K) + K\partial_K C^{mkt}(t, K) + 2\int_k^\infty C^{mkt}(t, x)dx}{\partial_K^2 C^{mkt}(t, K)}$$

and,

$$E(t, K) = E^Q[e^{-\int_0^t r_s ds} \xi_t]$$

ξ_t being an explicit function of x_t and y_t:

$$\xi_t = \mathbb{1}_{s_t^{t+\theta} > K}(f_{t,t} - f_{t,t+\theta} P_{t,t+\theta} - s_t^{t+\theta}(1 - P_{t,t+\theta}))$$

Since the volatility structure depends on the t-marginals of x_t and y_t through $E(t, .)$, the particle method explained in [6] and [7] is then an appropriate method for calibration.

9 Conclusion

In this project, we have significantly worked on different short rate models: from the one factor Hull While model to the G2++ two factor model. We have studied the theoretical fundamentals of those models and have been able to calibrate them and price different interest rate derivatives with obtaining some interesting numerical results. We had also the possibility to explore, in a theoretical way, rate models enhanced with local volatility, unfortunately we didn't have the opportunity to perform some numerical applications to this class of models.
To conclude, we would like to thank Barclays Capital and Mr. Ralph Laviolette for giving us the opportunity to be a part of a such stimulating project.

10 Bibliography

- [1] “Interest Rate Modeling”, V. Piterbarg, L. Andersen 2010
- [2] “Equity Hybrid Derivatives”, H. Buehler, M. Overhaus, 2006
- [3] “Turbo Charging the Cheyette Model”, J. Andreasen, 2000
- [4] “Pricing and Hedging with Smiles”, B. Dupire, 1993
- [5]”Interest Rate Models Enhanced With Local Volatility”, P.H Labordere, Lingling Cao, 2016.
- [6] “The Smile Calibration Problem Solved”, P.H. Labordere, J. Guyon, 2011
- [7] “Being Particular About Calibration”, P.H. Labordere, J. Guyon, 2011