

Smile with the Gaussian term structure model

Abdelkoddousse Ahdida, Aurélien Alfonsi*, Ernesto Palidda*

November 5, 2015

Abstract

We propose an affine extension of the Linear Gaussian term structure Model (LGM) such that the instantaneous covariation of the factors is given by an affine process on semidefinite positive matrices. First, we set up the model and present some important properties concerning the Laplace transform of the factors and the ergodicity of the model. Then, we present two main numerical tools to implement the model in practice. First, we obtain an expansion of caplets and swaptions prices around the LGM. Such a fast and accurate approximation is useful for assessing the model behavior on the implied volatility smile. Second, we provide a second order scheme for the weak error, which enables to calculate exotic options by a Monte-Carlo algorithm. These two pricing methods are compared with the standard one based on Fourier inversion.

Keywords : *Affine Term Structure Model, Linear Gaussian Model, Wishart processes, Price expansion, Discretization scheme, Caplets, Swaptions*

Motivation and overview of the paper

Affine Term Structure Models (ATSM) are an important class of models for interest rates that include the classical and pioneering models of Vasicek [34] and Cox-Ingersoll-Ross [11]. These models have been settled and popularized by the papers of Duffie and Kan [19], Dai and Singleton [15] and Duffie, Filipović and Schachermayer [18]. We refer to Filipović [22] for a textbook on these term structure models. The Linear Gaussian Model (LGM) is a simple but important subclass of ATSM that assumes that the underlying factors follow a Gaussian process. It has been considered by El Karoui and Lacoste [21] and El Karoui et al. [20], and has now become a market standard for pricing fixed income derivatives, thanks to its simplicity. However, this model has a main drawback to be calibrated to market data: it produces implied volatility smiles that are flat.

The goal of this paper is to present a quite natural extension of the LGM that keeps the affine structure and generates an implied volatility smile. To do so, we consider an affine diffusion of Wishart type on the set of semidefinite positive matrices and replace, roughly speaking, the constant volatility matrix by (a linear function of) this process. The dependence between the factors and their volatility is made through a specific covariation that keeps the affine structure and that has been proposed by Da Fonseca et al. [14] in an equity framework. Thanks to this, the proposed model which is a stochastic variance-covariance affine term structure model (see Definition 6), is able to produce an implied volatility smile. It has many parameters and may seem at first sight difficult to handle. For this reason, we present it as a perturbation of the LGM. Thus, the calibration of the model to market data can be made in two steps: first, one can calibrate the LGM and then calibrate the new parameters to the implied volatility smile. The calibration of this model is discussed on some cases in Palidda [27]. In the present paper, we do not tackle the practical calibration issue: our goal is just to set up the model and give the main numerical methods for a practical use of this model. Namely, we define in Section 2 the model and present some important properties such as the value of the Laplace transform under the initial and forward measures or the ergodicity property. Then, we give two tools that are important to implement the model in practice. First, we present in Section 3 a price expansion for caplets and swaptions around the LGM when the volatility of the volatility of the factor Y is small. These explicit formulas are useful to calculate quickly the impact of the parameters on the volatility cube and thus to calibrate the model. Second, we propose in Section 4 a discretization scheme for the model that is of second order for the weak error. Having an accurate scheme is important in practice since it allows to calculate exotic

*Université Paris-Est, CERMICS, Projet MATHRISK ENPC-INRIA-UMLV, 6-8 avenue Blaise Pascal, 77455 Marne La Vallée, France.
E-mails : ahdida.abdel@gmail.com, alfonsi@cermics.enpc.fr, ernesto.palidda@gmail.com.
Most of this work has been done when Ernesto was working for the Groupe de Recherche Opérationnelle of Crédit Agricole. We thank Christophe Michel, Victor Reutenauer, Anas Benabid and Nicole El Karoui for interesting and fruitful discussions around this paper. This research also benefited from the support of the "Chaire Risques Financiers", Fondation du Risque.

options by a Monte-Carlo algorithm. Besides, this scheme can be easily adapted to other models relying on the same affine structure such as the one of Da Fonseca et al. [14]. Last, Section 5 compares the expansion and the Monte-Carlo method with the classical Fourier technique popularized by Carr and Madan [9] and indicates the relevance of each method.

1 The Linear Gaussian Model (LGM) in a nutshell

The model that we present is meant to extend the classical LGM, and we need thus to recall briefly the LGM. We work under a risk-neutral measure \mathbb{P} , and consider a p -dimensional standard Brownian motion Z . Let Y be the solution of the following Ornstein-Uhlenbeck SDE

$$Y_t = y + \int_0^t \kappa(\theta - Y_s) ds + \int_0^t \sqrt{V} dZ_s, \quad (1)$$

where $\kappa \in \mathcal{M}_p(\mathbb{R})$ is a matrix of order p , V is a semidefinite positive matrix of order p and $\theta \in \mathbb{R}^p$. The LGM assumes that the spot rate is an affine function of the vector Y :

$$r_t = \varphi + \sum_{i=1}^p Y_t^i, \quad (2)$$

and the coordinates Y^i are usually called the factors of the model. It is not restrictive to assume that the weight of each factor in (2) is the same for all factors and equal to one: if we had $r_t = \varphi + \sum_{i=1}^p m_i Y_t^i$, we could check easily that $(m_1 Y^1, \dots, m_p Y^p)^\top$ is also an Ornstein-Uhlenbeck process. Affine Term Structure models generally assume that the parameters (here κ , θ and V) are fixed and are valid over a long time period, while the factors (here the vector Y) evolve and reflect the current state of the market. Therefore, one often assumes that the process Y is stationary to reflect some market equilibrium. Also, the factors are usually associated to different time scales: a factor with a small (resp. large) mean-reversion will influence the long-term (resp. short-term) behaviour of the interest rate. This leads to assume that

$$\kappa = \text{diag}(\kappa_1, \dots, \kappa_p) \text{ with } 0 < \kappa_1 < \dots < \kappa_p,$$

and we work under this assumption in the sequel. It can be easily checked (see for example Andersen and Piterbarg [3]) that any linear Gaussian model such that κ has distinct positive eigenvalues can be rewritten, up to a linear transformation of the factors, within the present parametrization.

Let $(\mathcal{F}_t)_{t \geq 0}$ denote the natural filtration of Z . For $0 \leq t \leq T$, the price $P_{t,T} = \mathbb{E} \left[\exp \left(- \int_t^T r_s ds \right) \middle| \mathcal{F}_t \right]$ at time t of the zero-coupon bond with maturity T is an exponential affine function of Y :

$$P_{t,T} = \exp(E(T-t) + B(T-t)^\top Y_t), \quad (3)$$

where $B(\tau) = -(\kappa^\top)^{-1}(I_p - e^{-\kappa^\top \tau})\mathbf{1}_p$ and $E(\tau) = -\varphi\tau + \int_0^\tau B(s)^\top \kappa \theta + \frac{B(s)^\top V B(s)}{2} ds$ for $\tau \geq 0$. Here, $\mathbf{1}_p$ stands for the vector in \mathbb{R}^p that has all its entries equal to one. The function $B(\tau)$ maps the factors variations ΔY on the yield curve variations and is often called the support function. The factors Y^i associated with the larger parameters κ_i impact on the short term behaviour of the yield curve while the one associated with the smaller parameters κ_i will drive more the long term behaviour.

We now briefly introduce some of the basic notions on the interest rates vanilla option market. The most liquid traded interest rates options are swaptions and caplets. They are respectively expressed with respect to the forward Libor rate and the forward swap rate, which are defined as follows for $0 \leq t \leq T$, $\delta > 0$ and $m \in \mathbb{N}^*$:

$$\begin{aligned} L_t(T, \delta) &= \frac{1}{\delta} \left(\frac{P_{t,T}}{P_{t,T+\delta}} - 1 \right) \\ S_t(T, m) &= \frac{P_{t,T} - P_{t,T+m\delta}}{\delta \sum_{i=1}^m P_{t,T+i\delta}}. \end{aligned}$$

The prices of caplets and swaptions are respectively given by

$$C_t(T, \delta, K) = \mathbb{E} \left[e^{-\int_t^{T+\delta} r_s ds} (L_T(T, \delta) - K)^+ \middle| \mathcal{F}_t \right]$$

$$\text{Swaption}_t(T, m, \delta, K) = \mathbb{E} \left[e^{-\int_t^T r_s ds} \sum_{i=1}^m \delta P_{t, T+i\delta} (S_T(T, m) - K)^+ \middle| \mathcal{F}_t \right].$$

Caplets are usually available for short tenors δ (up to 1 year) and swaptions are quoted for tenors $m\delta$ from 2 to 30 years. The market practice is to apply a standard change of numeraire technique (see Geman et al. [24]) and rewrite the above expressions as

$$C_t(T, \delta, K) = P_{t, T+\delta} \mathbb{E}^{T+\delta} \left[(L_T(T, \delta) - K)^+ \middle| \mathcal{F}_t \right] \quad (4)$$

$$\text{Swaption}_t(T, m, \delta, K) = \left(\sum_{i=1}^m \delta P_{t, T+i\delta} \right) \mathbb{E}^A \left[(S_T(T, m) - K)^+ \middle| \mathcal{F}_t \right], \quad (5)$$

where $\mathbb{E}^{T+\delta}$ (resp. \mathbb{E}^A) denotes the expectation taken with respect to the measure $T + \delta$ -forward neutral (resp. annuity) measure associated with the numeraire $P_{t, T+\delta}$ (resp. $\sum_{i=1}^m \delta P_{t, T+i\delta}$). The market prices are then quoted and analyzed in terms of either the log-normal or normal implied volatility obtained by inverting respectively the pricing formulas (4) and (5) w.r.t. the Black-Scholes and Bachelier formulas. Within the LGM model, the log-normal implied volatility of the caplet is given by

$$\int_t^T [B(T-u) - B(T+\delta-u)]^\top V [B(T-u) - B(T+\delta-u)] du,$$

which is a particular case of formula (28) below. This implied volatility does not depend on the strike. It shows that the mean-reversion parameter κ plays a role in shaping the form of the caplets volatility cube, according to the different time scales. The role of the diagonal coefficients of the matrix V is determined by the support functions $m_{ii}(\tau, \delta) = (\frac{1-e^{-\kappa_i \delta}}{\kappa_i})^2 \frac{1-e^{-2\kappa_i \tau}}{2\kappa_i \tau}$. The effect of off-diagonal elements of the variance-covariance matrix V is determined by the support functions $m_{ij}(\tau, \delta) = \frac{1-e^{-\kappa_i \delta}}{\kappa_i} \frac{1-e^{-\kappa_j \delta}}{\kappa_j} \frac{1-e^{-(\kappa_i + \kappa_j) \tau}}{(\kappa_i + \kappa_j) \tau}$. These functions are plotted in Figure 1.

Also, by using a standard approximation, we can obtain the normal implied volatility of the swaptions:

$$\int_t^T [B^S(u)]^\top V B^S(u) du,$$

with $B^S(u) = \omega_0^0 B(T-u) - \omega_0^m B(T+m\delta-u) - S_0(T, m, \delta) \sum_{k=1}^m \omega_0^k B(T+k\delta-u)$, $\omega_0^k = \frac{P_{0, T+k\delta}}{\sum_{i=1}^m P_{0, T+i\delta}}$. This is a particular case of formula (43) below. This implied volatility has a rather similar structure as the one of the caplets, but it is not time homogeneous. Both implied volatilities for caplets and swaptions do not depend on

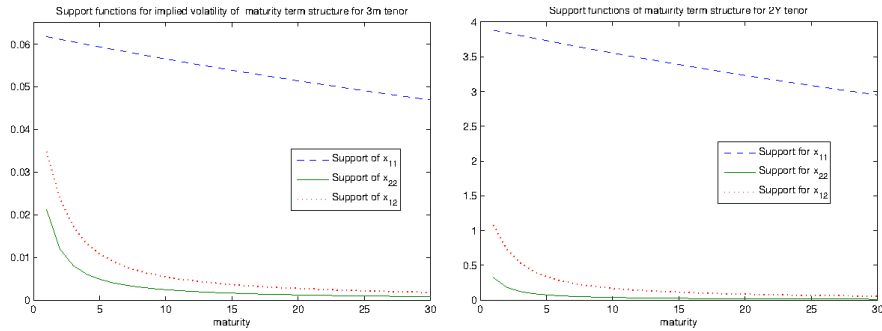


Figure 1: Support functions for the volatility term structure in a two factors model with $\kappa = \text{diag}(0.01, 1)$ for a 3 months (left) and 2 years (right) maturity.

the strike and give thus a flat smile, which is a well-known fact. This is unfortunate if one aims to reproduce the volatility cube observed on market data (i.e. the implied volatility with respect to the maturity T , the tenor δ or $m\delta$ and the strike K). The extension of the LGM that we introduce in Section 2 is meant to correct this drawback.

2 An affine extension of the LGM with stochastic covariance

This section is devoted to the definition of the model that we study in this paper. This model is a stochastic variance-covariance perturbation of the LGM. We chose a quite general specification that keeps the model affine and gives a stochastic instantaneous covariance for the factors, which will generate a smile for the Caplets and Swaptions. We first present the dynamics of the factor and then present some properties of the model that rely on its affine structure.

2.1 State variables dynamics

We consider W a d -by- d square matrix made of independent standard Brownian motions, and Z an independent Brownian motion of dimension p . We will denote by $(\mathcal{F}_t)_{t \geq 0}$ the filtration generated by (W, Z) . We consider the following SDE for the state variables (or factors)

$$Y_t = y + \int_0^t \kappa(\theta - Y_s) ds + \int_0^t c \sqrt{X_s} [\bar{\rho} dZ_s + dW_s \rho] \quad (6)$$

$$X_t = x + \int_0^t (\Omega + (d-1)\epsilon^2 I_d^n + bX_s + X_s b^\top) ds + \epsilon \int_0^t \sqrt{X_s} dW_s I_d^n + I_d^n dW_s^\top \sqrt{X_s}. \quad (7)$$

The matrix I_d^n is defined for $0 \leq n \leq d$ by $(I_d^n)_{i,j} = \mathbb{1}_{i=j \leq n}$, and the parameters are taken as follows

$$\begin{aligned} x, \Omega \in \mathcal{S}_d^+(\mathbb{R}), b \in \mathcal{M}_d(\mathbb{R}), \epsilon \in \mathbb{R}_+, y, \theta \in \mathbb{R}^p, \kappa = \text{diag}(\kappa_1, \dots, \kappa_p) \text{ with } \kappa_1, \dots, \kappa_p > 0, \\ c \in \mathcal{M}_{p \times d}(\mathbb{R}), \rho \in \mathbb{R}^d \text{ such that } |\rho|^2 := \sum_{i=1}^d \rho_i^2 \leq 1 \quad \text{and} \quad \bar{\rho} = \sqrt{1 - |\rho|^2}, \end{aligned} \quad (8)$$

where $\mathcal{S}_d^+(\mathbb{R})$, $\mathcal{M}_d(\mathbb{R})$, and $\mathcal{M}_{p \times d}(\mathbb{R})$ denote respectively the set of semidefinite positive matrices of order d , the set of square matrices of order d , and the set of matrices with p rows and d columns. The process X is an affine diffusion on $\mathcal{S}_d^+(\mathbb{R})$, and the instantaneous covariance at time t of the factors Y is given by $cX_t c^\top$. When $\epsilon = 0$ and $\Omega = -bx - xb^\top$, we have $X_t = x$ and get back the Gaussian model with $V = cxc^\top$. The dependence structure between Y and X through the driving Brownian motions is the same as the one proposed by Da Fonseca, Grasselli and Tebaldi [14]. As explained in [14], this is the most general way to get a non trivial instantaneous correlation between Y and X while keeping the affine structure. In particular, the instantaneous quadratic covariations are linear with respect to (Y, X) and we have for $1 \leq i, j, k, l \leq d$ and $1 \leq m, m' \leq p$:

$$\langle d(Y_t)_m, d(Y_t)_{m'} \rangle = (cX_t c^\top)_{m,m'} dt, \quad (9)$$

$$\langle d(X_t)_{i,j}, d(X_t)_{k,l} \rangle = \epsilon^2 [(X_t)_{i,k} \mathbb{1}_{j=l \leq n} + (X_t)_{i,l} \mathbb{1}_{j=k \leq n} + (X_t)_{j,k} \mathbb{1}_{i=l \leq n} + (X_t)_{j,l} \mathbb{1}_{i=k \leq n}] dt, \quad (10)$$

$$\langle d(Y_t)_m, d(X_t)_{i,j} \rangle = \epsilon [(cX_t)_{m,i} (I_d^n \rho)_j + (cX_t)_{m,j} (I_d^n \rho)_i] dt. \quad (11)$$

We notice that only the n first components of ρ matter, and we can assume without loss of generality that $\rho_{n+1} = \dots = \rho_d = 0$.

From (7), we easily get

$$e^{\kappa t} Y_t = y + \int_0^t e^{\kappa s} \kappa \theta ds + \int_0^t e^{\kappa s} c \sqrt{X_s} [\bar{\rho} dZ_s + dW_s \rho].$$

Therefore, the process Y is uniquely determined once the processes Z, W and X are given. We know by Cuchiero et al. [12] that the SDE on X has a unique weak solution when $x \in \mathcal{S}_d^+(\mathbb{R})$ and $\Omega \in \mathcal{S}_d^+(\mathbb{R})$, and a unique strong solution if we assume besides that x is invertible and $\Omega - 2\epsilon^2 I_d^n \in \mathcal{S}_d^+(\mathbb{R})$. This leads to the following result.

Proposition 1 — *If $x \in \mathcal{S}_d^+(\mathbb{R})$, $\Omega \in \mathcal{S}_d^+(\mathbb{R})$ there exists a unique weak solution of the SDE (7). If we assume moreover that $\Omega - 2\epsilon^2 I_d^n \in \mathcal{S}_d^+(\mathbb{R})$ and $x \in \mathcal{S}_d^+(\mathbb{R})$ is positive definite, there is a unique strong solution to the SDE (7).*

The affine structure of the process (X, Y) allows us to give formulas for the Laplace transform of the marginal laws by means of Matrix Riccati Differential Equations (MRDE). Similar calculations have been made in equity modelling by Da Fonseca et al. [14] or Benabid et al. [5]. The following proposition states the precise result, which is useful for the pricing of Zero-Coupon bonds.

Proposition 2 — Let $\Lambda, \bar{\Lambda} \in \mathbb{R}^p$, $\Gamma, \bar{\Gamma} \in \mathcal{S}_d(\mathbb{R})$. For $t \geq 0$, we define $\lambda(t) \in \mathbb{R}^p$ by

$$\lambda_i(t) = \Lambda_i e^{-\kappa_i t} + \frac{\bar{\Lambda}_i}{\kappa_i} (1 - e^{-\kappa_i t}). \quad (12)$$

Let us assume that there exists $\Upsilon \in \mathcal{S}_d(\mathbb{R})$ such that

$$\Upsilon - \Gamma \in \mathcal{S}_d^+(\mathbb{R}), \quad (13)$$

$$\forall t \geq 0, - \left[2\epsilon^2 \Upsilon I_d^n \Upsilon + \Upsilon (b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c) + (b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c)^\top \Upsilon + \frac{1}{2} c^\top \lambda \lambda^\top c + \bar{\Gamma} \right] \in \mathcal{S}_d^+(\mathbb{R}) \quad (14)$$

Then, the following system of differential equations

$$\begin{cases} \dot{g} = 2\epsilon^2 g I_d^n g + g(b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c) + (b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c)^\top g + \frac{1}{2} c^\top \lambda \lambda^\top c + \bar{\Gamma}, & g(0) = \Gamma, \\ \dot{\eta} = \lambda^\top \kappa \theta + \text{Tr}(g(\Omega + \epsilon^2(d-1)I_d^n)), & \eta(0) = 0, \end{cases} \quad (15)$$

has a unique solution, which is defined on \mathbb{R}_+ . It satisfies $\Upsilon - g(t) \in \mathcal{S}_d^+(\mathbb{R})$ for any $t \geq 0$. Besides, we have for any $0 \leq t \leq T$:

$$\mathbb{E} \left[\exp \left(\text{Tr}(\Gamma X_T) + \Lambda^\top Y_T + \int_t^T \text{Tr}(\bar{\Gamma} X_s) + \bar{\Lambda}^\top Y_s ds \right) \middle| \mathcal{F}_t \right] = \exp(\eta(T-t) + \text{Tr}(g(T-t)X_t) + \lambda(T-t)^\top Y_t). \quad (16)$$

Proof: The proof is quite standard for affine diffusion. First, we notice that if (16) holds, we necessarily have that $M_t = \exp \left(\int_0^t \text{Tr}(\bar{\Gamma} X_s) + \bar{\Lambda}^\top Y_s ds \right) \exp(\eta(T-t) + \text{Tr}(g(T-t)X_t) + \lambda(T-t)^\top Y_t)$ is a martingale. We apply Itô's formula and use (9), (10) and (11). The martingale property yields to

$$\begin{aligned} & \bar{\Gamma} X_t + \bar{\Lambda}^\top Y_t - \dot{\eta}(T-t) - \text{Tr}(\dot{g}(T-t)X_t) - \dot{\lambda}(T-t)^\top Y_t + \text{Tr}(g(T-t)[\Omega + (d-1)\epsilon^2 I_d^n + bX_s + X_s b^\top]) \\ & + \lambda(T-t)^\top \kappa(\theta - Y_t) + 2\epsilon^2 \text{Tr}(Xg(T-t)I_d^n g(T-t)) + \frac{1}{2} \text{Tr}(Xc^\top \lambda(T-t)\lambda^\top(T-t)c) \\ & + \frac{\epsilon}{2} \text{Tr}(X[c^\top \lambda(T-t)\rho^\top I_d^n g(T-t) + g(T-t)I_d^n \rho \lambda^\top(T-t)c]) = 0. \end{aligned}$$

By identifying the constant term and the linear terms with respect to Y_t and X_t , we get (15) and $\dot{\lambda} = -\kappa\lambda + \bar{\Lambda}$, $\lambda(0) = \Lambda$, which leads to (12) since κ is diagonal with positive entries. By applying Proposition 1.1 of Dieci and Eirola¹ [17] to $\Upsilon - g$, the solution of (15) exists and is well defined for $t \geq 0$. Besides, $\Upsilon - g$ stays in $\mathcal{S}_d^+(\mathbb{R})$ by using (13) and (14).

Then, it remains to check that we have indeed (16), and it is sufficient to check it for $t = 0$. To do so, we apply Itô's formula to M and get

$$dM_s = M_s \left[\text{Tr}(g(T-s)[\sqrt{X_s} dW_s I_d^n + I_d^n dW_s^\top \sqrt{X_s}]) + \lambda(T-s)^\top c \sqrt{X_s} [\bar{\rho} dZ_s + dW_s \rho] \right].$$

Thus, M is a positive local martingale and thus a supermartingale, which gives $M_0 \geq \mathbb{E}[M_T]$. To prove that $M_0 = \mathbb{E}[M_T]$, we use the argument presented by Rydberg [30]. We define $N_t = M_t/M_0$ in order to work with probability measures. We define for $\mathbf{K} > 0$, $\tau_{\mathbf{K}} = \inf\{t \geq 0, \text{Tr}(X_t) \geq \mathbf{K}\}$, $\pi_{\mathbf{K}}(x) = \mathbb{1}_{\text{Tr}(x) \leq \mathbf{K}} x + \mathbb{1}_{\text{Tr}(x) \geq \mathbf{K}} \frac{\mathbf{K}}{\text{Tr}(x)} x$ for $x \in \mathcal{S}_d^+(\mathbb{R})$ and consider $N_t^{(\mathbf{K})}$ the solution of

$$\begin{aligned} dN_s^{(\mathbf{K})} &= N_s^{(\mathbf{K})} \left\{ \text{Tr}(g(T-s)[\sqrt{\pi_{\mathbf{K}}(X_s)} dW_s I_d^n + I_d^n dW_s^\top \sqrt{\pi_{\mathbf{K}}(X_s)}]) \right. \\ & \quad \left. + \lambda(T-s)^\top c \sqrt{\pi_{\mathbf{K}}(X_s)} [\bar{\rho} dZ_s + dW_s \rho] \right\}, \\ N_0^{(\mathbf{K})} &= 1. \end{aligned}$$

Clearly, $\mathbb{E}[N_T^{(\mathbf{K})}] = 1$, and under $\frac{d\mathbb{P}^{(\mathbf{K})}}{d\mathbb{P}} = N_T^{(\mathbf{K})}$,

$$dW_t^{(\mathbf{K})} = dW_t - 2\sqrt{\pi_{\mathbf{K}}(X_t)} g(T-t) I_d^n - \sqrt{\pi_{\mathbf{K}}(X_t)} c^\top \lambda(T-t) \rho^\top$$

¹We thank Martino Grasselli for pointing us this reference.

is a matrix Brownian motion under $\mathbb{P}^{(\mathbf{K})}$.

We now write $\mathbb{E}[N_T] = \mathbb{E}[N_T \mathbb{1}_{\tau_{\mathbf{K}} > T}] + \mathbb{E}[N_T \mathbb{1}_{\tau_{\mathbf{K}} \leq T}]$. By the dominated convergence theorem, we have $\mathbb{E}[N_T \mathbb{1}_{\tau_{\mathbf{K}} \leq T}] \xrightarrow{\mathbf{K} \rightarrow +\infty} 0$. Besides, $\mathbb{E}[N_T \mathbb{1}_{\tau_{\mathbf{K}} > T}] = \mathbb{E}[N_T^{(\mathbf{K})} \mathbb{1}_{\tau_{\mathbf{K}} > T}] = \mathbb{P}^{(\mathbf{K})}(\tau_{\mathbf{K}} > T)$, and we have to prove that this probability goes to 1. To do so, we focus on the following SDE

$$d\tilde{X}_t = (\Omega + (d-1)\epsilon^2 I_d^n + (b + 2\epsilon I_d^n g(T-t) + \epsilon I_d^n \rho \lambda^\top (T-t)c)\tilde{X}_t + \tilde{X}_t(b^\top + 2\epsilon g(T-t)I_d^n + \epsilon c^\top \lambda(T-t)\rho^\top I_d^n))dt \\ + \epsilon \left(\sqrt{\tilde{X}_t} dW_t I_d^n + I_d^n dW_t^\top \sqrt{\tilde{X}_t} \right)$$

starting from $\tilde{X}_0 = X_0$. We check that X solves before $\tau_{\mathbf{K}}$ and under $\mathbb{P}^{(\mathbf{K})}$ the same SDE as \tilde{X} under \mathbb{P} . This yields to $\mathbb{P}^{(\mathbf{K})}(\tau_{\mathbf{K}} > T) = \mathbb{P}(\inf\{t \geq 0, \text{Tr}(\tilde{X}_t) \geq \mathbf{K}\} > T)$. Since the SDE satisfied by \tilde{X} is the one of an affine diffusion on $\mathcal{S}_d^+(\mathbb{R})$, it is well defined for any $t \geq 0$. In particular $\max_{t \in [0, T]} \text{Tr}(\tilde{X}_t) < \infty$ a.s., which gives $\mathbb{P}(\inf\{t \geq 0, \text{Tr}(\tilde{X}_t) \geq \mathbf{K}\} > T) \xrightarrow{\mathbf{K} \rightarrow +\infty} 1$. \square

Remark 3 — The conditions (13) and (14) are satisfied for $\Upsilon = 0$ if, and only if $-\Gamma \in \mathcal{S}_d^+(\mathbb{R})$ and

$$\forall t \geq 0, -\bar{\Gamma} - \frac{1}{2}c^\top \lambda(t)\lambda^\top(t)c \in \mathcal{S}_d^+(\mathbb{R}). \quad (17)$$

Since $|\lambda_i(t)| \leq \max(|\Lambda_i|, |\bar{\Lambda}_i/\kappa_i|)$, we obtain $\lambda(t)^\top \lambda(t) \leq \sum_{i=1}^p \max(\Lambda_i^2, (\bar{\Lambda}_i/\kappa_i)^2)$. We therefore have $\sum_{i=1}^p \max(\Lambda_i^2, (\bar{\Lambda}_i/\kappa_i)^2)I_d - \lambda(t)^\top \lambda(t) \in \mathcal{S}_d^+(\mathbb{R})$ and then $\sum_{i=1}^p \max(\Lambda_i^2, (\bar{\Lambda}_i/\kappa_i)^2)c^\top c - c^\top \lambda(t)\lambda^\top(t)c \in \mathcal{S}_d^+(\mathbb{R})$. Therefore, a sufficient condition for (17) is

$$-\bar{\Gamma} - \frac{1}{2} \sum_{i=1}^p \max(\Lambda_i^2, (\bar{\Lambda}_i/\kappa_i)^2)c^\top c \in \mathcal{S}_d^+(\mathbb{R}).$$

With the Laplace transform (16), we have a mathematical tool to check if the process (X, Y) is stationary. This is important for our modeling perspective: unless for some transitory period, one may expect that the factors are stable around some equilibrium. The next proposition give a simple sufficient condition that ensures stationarity. It is proved in Appendix B.

Proposition 4 — If $-(b + b^\top) \in \mathcal{S}_d^+(\mathbb{R})$ is positive definite, the process (X, Y) is stationary.

Remark 5 — We chose to keep the dynamics of the process X in the space of positive semidefinite matrices as general as possible. Choosing a Wishart specification for X (which corresponds to $\Omega = \epsilon^2 \alpha I_d^n, \alpha > 0$) does not lead to a significant simplification of the model. While Wishart processes admit an explicit Laplace transform, this is not the case for the process (X, Y) defined by (7). The drift term Ω allows to account for a mean reversion behavior of the process X , we will typically consider a negative mean reversion matrix b , in which case we can set $\Omega = -bx^\infty - x^\infty b^\top$, so that the matrix process X mean reverts to x^∞ .

2.2 Model definition

Definition 6 — We assume that $(X_t, Y_t)_{t \geq 0}$ follows (6) and (7) under a risk-neutral measure. Then, we define the short interest rate by

$$r_t = \varphi + \sum_{i=1}^p Y_t^i + \text{Tr}(\gamma X_t), \quad (18)$$

with $\varphi \in \mathbb{R}$ and $\gamma \in \mathcal{S}_d(\mathbb{R})$.

From Proposition 2, we easily get the following result on the Zero-Coupon bonds.

Corollary 7 Bond reconstruction formula. Let $0 \leq t \leq T$ and $P_{t,T} = \mathbb{E}[\exp(-\int_t^T r_s ds) | \mathcal{F}_t]$ denote the price at time t of a zero-coupon bond with maturity T . Let us assume that

$$\gamma - \frac{1}{2} \left(\sum_{i=1}^p \frac{1}{\kappa_i^2} \right) c^\top c \in \mathcal{S}_d^+(\mathbb{R}). \quad (19)$$

Then, by using Remark 3, $P_{t,T}$ is given by

$$P_{t,T} = \exp(A(T-t) + \text{Tr}(D(T-t)X_t) + B(T-t)^\top Y_t), \quad (20)$$

with $A(t) = \eta(t) - \varphi t$, $D(t) = g(t)$ and $B(t) = \lambda(t)$, where (η, g, λ) is the solution of (15) with (12), $\Lambda = 0$, $\Gamma = 0$, $\bar{\Gamma} = -\gamma$ and $\bar{\Lambda} = -\mathbf{1}_p$ (i.e. $\bar{\Lambda}_i = -1$ for $1 \leq i \leq p$). In particular, we have $-D(T-t) \in \mathcal{S}_d^+(\mathbb{R})$.

Let us make general comments on the model. It is close but slightly different from the model proposed in the PhD Thesis of Bensusan [6]. Nonetheless, our presentation as a perturbation of the LGM model enables us to have a better understanding of the model parameters. Thus, the vector process Y can be interpreted as in the LGM model, meaning it is assumed to be the main driver of the yield curve. The individual factors are viewed as principal components movements of the yield curve. We chose to specify the model such that the matrix process X admits a similar interpretation. Typically we will consider $\Omega = -bx^\infty - x^\infty b^\top$ with b symmetric negative to have a mean-reversion toward a given covariance matrix x^∞ . The parameter ϵ measures the level of the perturbation around the LGM. The matrix process X plays the role of a stochastic variance-covariance matrix of the main movements of the yield curve. It is possible to define the diffusion parameter c such that the diagonal factors of the matrix X play the role of the instantaneous stochastic variance of the yield curve movement and the off-diagonal terms play the role of the instantaneous covariance between two yield curve movements. The vector ρ is a correlation parameter between the processes Y and X . In a first approximation², interest rates options are options on linear combinations of the factors Y , and instantaneous variance of these linear combinations are linear combinations of the factors X . Therefore, the correlation parameter ρ will drive the skew of interest rates options. We now make more precise comments on the model.

- In order to keep the same factors as in the LGM, one would like to take $\gamma = 0$. However, this choice is possible only if the perturbation around the LGM is small enough provided that $-(b + b^\top)$ is positive definite, see Remark 8. Besides, even if $P_{t,T}$ may be well defined for $T - t$ small enough, it would be then given by the same formula, and therefore the yield curve dynamics depends anyway on the factor X .
- In order to have a clear interpretation of the volatility factor X on the factor Y , a possible choice is to consider $d = q \times p$ with $q \in \mathbb{N}^*$ and $c_{i,j} = \mathbb{1}_{(i-1) \times p < j \leq i \times p}$. Thus, from (9), the principal matrix $(X_{k,l})_{(i-1) \times p < k, l \leq i \times p}$ rules the instantaneous quadratic variation of the factor Y_i while the submatrix $(X_{k,l})_{(i-1) \times p < k \leq i \times p, (j-1) \times p < l \leq j \times p}$ rules the instantaneous covariation between the factors Y^i and Y^j .
- The model does not prevent from having a negative short rate or from having $\mathbb{E}[|P_{t,T}|^k] = \infty$ for any $k > 0$, unless we consider the degenerated case ($p = 0$) where the yield curve is driven by the volatility factors X and the factors Y are null. This particular model has been studied by Gnoatto in [25].
- Affine Term Structure models generally consider constant parameters that are fixed over a large period and reflect the market behaviour, while the current value of factors are fitted to market data. This is why we consider constant parameters here. However, in order to fit exactly Zero-Coupon Bond prices, it is possible to take a time-dependent function φ while keeping the tractability of the model.

Remark 8 — The condition (19) is sufficient to get that $P_{t,T}$ is well-defined. However, this condition does not depend on ϵ while we know that for $\epsilon = 0$, $P_{t,T}$ is well-defined since X is deterministic and Y is a Gaussian process. We can get a complementary sufficient condition when $-(b + b^\top)$ is positive definite, which is a reasonable assumption since it leads to a stationary process by Proposition 4. In this case, there exists $\mu > 0$ such that $-(b + b^\top) - \mu I_d \in \mathcal{S}_d^+(\mathbb{R})$. By using Proposition 2 with $\Upsilon = \frac{\mu}{4\epsilon^2} I_d$, we get that (14) is satisfied if we have

$$\forall t \geq 0, \frac{\mu^2}{8\epsilon^2} I_d - \frac{\mu}{8\epsilon} (I_d^n \rho \lambda^\top c + I_d^n \rho \lambda^\top c) - \frac{1}{2} c^\top \lambda \lambda^\top c + \gamma \in \mathcal{S}_d^+(\mathbb{R}).$$

²Note that this is not completely true, even in the simple LGM model. One important characteristic of short rate/factorial interest rates model is that the yield curve depends not only on the (stochastic) state variables of the model, but also on the volatility of the state variables. Therefore the volatility factors X appear in the payoff of interest rates options.

Since for $t \geq 0$, $\lambda(t)$ takes values in a compact subset of \mathbb{R}^p , there is $\epsilon_0 > 0$ such that this condition is satisfied for any $\epsilon \in (0, \epsilon_0)$.

Remark 9 — Let $a \in \mathcal{M}_d(\mathbb{R})$, and consider the model $r_t = \varphi + \sum_{i=1}^p Y_t^i + \text{Tr}(\tilde{\gamma} \tilde{X}_t)$ with

$$\begin{aligned} Y_t &= y + \int_0^t \kappa(\theta - Y_s) ds + \int_0^t \tilde{c} \sqrt{\tilde{X}_s} \left[\sqrt{1 - |\tilde{\rho}|^2} dZ_s + dW_s \tilde{\rho} \right] \\ \tilde{X}_t &= \tilde{x} + \int_0^t \left(\tilde{\Omega} + (d-1)\epsilon^2 a^\top a + \tilde{b} \tilde{X}_s + \tilde{X}_s \tilde{b}^\top \right) ds + \epsilon \int_0^t \sqrt{\tilde{X}_s} dW_s a + a^\top dW_s^\top \sqrt{\tilde{X}_s}, \end{aligned}$$

and $\tilde{\gamma} \in \mathcal{S}_d(\mathbb{R})$, $\tilde{x}, \tilde{\Omega} \in \mathcal{S}_d^+(\mathbb{R})$, $\tilde{c}, \tilde{b} \in \mathcal{M}_d(\mathbb{R})$, $\tilde{\rho} \in \mathbb{R}^d$ such that $|\tilde{\rho}| \leq 1$. This model may seem a priori more general, but this is not the case. In fact, let n be the rank of a and $u \in \mathcal{M}_d(\mathbb{R})$ be an invertible matrix such that $a^\top a = (u^{-1})^\top I_d^n (u^{-1})$. Then, $X_t = u^\top \tilde{X}_t u$ solves

$$dX_t = [\Omega + (d-1)\epsilon^2 I_d^n + bX_t + X_t b^\top] dt + u^\top \sqrt{\tilde{X}_t} dW_t a u + u^\top a^\top dW_t^\top \sqrt{\tilde{X}_t} u,$$

with $b = u^\top \tilde{b} (u^{-1})^\top$, $\Omega = u^\top \tilde{\Omega} u \in \mathcal{S}_d^+(\mathbb{R})$ and starting from $x = u^\top \tilde{x} u \in \mathcal{S}_d^+(\mathbb{R})$. After some calculations, we obtain $\langle d(Y_t)_m, d(Y_t)_{m'} \rangle = (\tilde{c} \tilde{X}_t \tilde{c}^\top)_{m,m'} dt = (c X_t c^\top)_{m,m'} dt$ with $c = \tilde{c} (u^{-1})^\top$; $\langle d(X_t)_{i,j}, d(X_t)_{k,l} \rangle = \epsilon^2 [(X_t)_{i,k} \mathbb{1}_{j=l \leq n} + (X_t)_{i,l} \mathbb{1}_{j=k \leq n} + (X_t)_{j,k} \mathbb{1}_{i=l \leq n} + (X_t)_{j,l} \mathbb{1}_{i=k \leq n}] dt$ and

$$\begin{aligned} \langle d(Y_t)_m, d(X_t)_{i,j} \rangle &= \epsilon \left[(u^\top \tilde{X}_t \tilde{c}^\top)_{m,i} (u^\top a^\top \tilde{\rho})_j + (u^\top \tilde{X}_t \tilde{c}^\top)_{m,j} (u^\top a^\top \tilde{\rho})_i \right] dt \\ &= \epsilon \left[(X_t c^\top)_{m,i} (u^\top a^\top \tilde{\rho})_j + (X_t c^\top)_{m,j} (u^\top a^\top \tilde{\rho})_i \right] dt. \end{aligned}$$

Since the law of (X, Y) is characterized by its infinitesimal generator, we can assume without loss of generality that $\tilde{\rho} \in \ker(u^\top a^\top)^\perp = \text{Im}(a u)$. Therefore, there is $\rho' \in \mathbb{R}^d$ such that $\tilde{\rho} = a u \rho'$, and we set $\rho_i = \rho'_i$ for $i \leq n$ and $\rho_i = 0$ for $n < i \leq d$. We have $|\rho|^2 = (\rho')^\top I_d^n \rho' = |\tilde{\rho}|^2 \leq 1$, and therefore (X, Y) follows the same law as the solution of (6) and (7), and we have $r_t = \varphi + \sum_{i=1}^p Y_t^i + \text{Tr}(\gamma X_t)$ with $\gamma = u^{-1} \tilde{\gamma} (u^{-1})^\top$.

2.3 Change of measure and Laplace transform

In the fixed income market, the pricing of vanilla products is often (if not always) made under a suitably chosen equivalent martingale measure different from the risk-neutral measure. It is thus important to characterize the distribution of the underlying state variables under these measures. The forward-neutral measures are probably the most important example of such pricing measures. In this paragraph, we will see that the dynamics of the factors remains affine and keeps the same structure under the forward measures.

2.3.1 Dynamics under the forward-neutral measures

We assume that the condition (19) holds. Let Q^U denote the U -forward neutral probability, which is defined on \mathcal{F}_U by

$$\frac{dQ^U}{d\mathbb{P}} = \frac{e^{-\int_0^U r_s ds}}{P_{0,U}}.$$

This is the measure associated with the numeraire $P_{t,U}$. It comes from the martingale property of discounted asset prices that for $t \in (0, U)$,

$$\begin{aligned} \frac{d \left(e^{-\int_0^t r_s ds} P_{t,U} \right)}{e^{-\int_0^t r_s ds} P_{t,U}} &= 2\epsilon \text{Tr}(D(U-t) \sqrt{X_t} dW_t I_d^n) + B(U-t)^\top c \sqrt{X_t} dW_t \rho + \bar{\rho} B(U-t)^\top c \sqrt{X_t} dZ_t \\ &= \text{Tr}([2\epsilon I_d^n D(U-t) \sqrt{X_t} + \rho B(U-t)^\top c \sqrt{X_t}] dW_t) + \bar{\rho} B(U-t)^\top c \sqrt{X_t} dZ_t. \end{aligned}$$

From Girsanov's theorem, the processes

$$\begin{aligned} dW_t^U &= dW_t - \sqrt{X_t} (2\epsilon D(U-t) I_d^n + c^\top B(U-t) \rho^\top) dt \\ dZ_t^U &= dZ_t - \bar{\rho} \sqrt{X_t} c^\top B(U-t) dt \end{aligned}$$

are respectively matrix and vector valued Brownian motions under Q^U and are independent. This yields to the following dynamics for Y and X under Q^U :

$$dX_t = (\Omega + (d-1)\epsilon^2 I_d^n + b^U(t)X_t + X_t(b^U(t))^\top)dt + \epsilon \left(\sqrt{X_t} dW_t^U I_d^n + I_d^n (dW_t^U)^\top \sqrt{X_t} \right) \quad (21)$$

$$dY_t = \kappa(\theta - Y_t)dt + cX_t c^\top B(U-t)dt + 2\epsilon cX_t D(U-t)I_d^n \rho dt + c\sqrt{X_t}(dW_t^U \rho + \bar{\rho} dZ_t^U), \quad (22)$$

with $b^U(t) = b + 2\epsilon^2 I_d^n D(U-t) + \epsilon I_d^n \rho B(U-t)^\top c$.

2.3.2 Laplace transforms

We are now interested in calculating the law of (X_T, Y_T) under the U -forward measure for $T \leq U$. More precisely, we calculate $\mathbb{E}^{Q^U} \left[\exp(\text{Tr}(\Gamma X_T) + \Lambda^\top Y_T) | \mathcal{F}_t \right]$ for $t \in [0, T]$ by using again Proposition 2. We assume that condition (19) holds and have

$$\begin{aligned} & \mathbb{E}^{Q^U} \left[\exp(\text{Tr}(\Gamma X_T) + \Lambda^\top Y_T) | \mathcal{F}_t \right] \\ &= \frac{1}{P_{t,U}} \mathbb{E} \left[\exp(\text{Tr}(\Gamma X_T) + \Lambda^\top Y_T - (U-t)\varphi - \int_t^U \mathbf{1}_p^\top Y_s ds - \int_t^U \text{Tr}(\gamma X_s) ds) | \mathcal{F}_t \right] \\ &= \frac{\mathbb{E} \left[\exp(\text{Tr}(\Gamma + D(U-T)X_T) + (\Lambda + B(U-T))^\top Y_T + A(U-T) - (T-t)\varphi - \int_t^T \mathbf{1}_p^\top Y_s ds - \int_t^T \text{Tr}(\gamma X_s) ds) | \mathcal{F}_t \right]}{\exp(A(U-t) + \text{Tr}(D(U-t)X_t) + B(U-t)^\top Y_t)}. \end{aligned}$$

We consider $\Gamma \in \mathcal{S}_d(\mathbb{R})$ and $\Lambda \in \mathbb{R}^p$ such that

$$-\Gamma \in \mathcal{S}_d^+(\mathbb{R}) \text{ and } |\Lambda_i| \leq e^{-\kappa_i(U-T)}/\kappa_i, \text{ for } 1 \leq i \leq p,$$

in order to have $|\Lambda_i + B_i(U-T)| \leq 1/\kappa_i$ and $-(\Gamma + D(U-T)) \in \mathcal{S}_d^+(\mathbb{R})$. By Proposition 2, condition (19) and Remark 3, we get that the expectation is finite and that

$$\mathbb{E}^{Q^U} \left[\exp(\text{Tr}(\Gamma X_T) + \Lambda^\top Y_T) | \mathcal{F}_t \right] = \exp(A^U(t, T) + \text{Tr}(D^U(t, T)X_t) + B^U(t, T)^\top Y_t), \quad (23)$$

with $F^U(t, T) = \tilde{F}(T-t) + F(U-T) - F(U-t)$ for $F \in \{A, D, B\}$, where $(\tilde{B}, \tilde{D}, \tilde{A})$ is the solution of (15) with $\tilde{B}(0) = \Lambda + B(U-T)$, $\tilde{D}(0) = \Gamma + D(U-T)$, $\tilde{A}(0) = 0$, $\tilde{\Lambda} = \mathbf{1}_p$ and $\tilde{\Gamma} = -\gamma$.

Corollary 10 *Let (19) hold. For $\Gamma \in \mathcal{S}_d(\mathbb{R})$ and $\Lambda \in \mathbb{R}^p$ such that $-\Gamma \in \mathcal{S}_d^+(\mathbb{R})$ and $|\Lambda_i| \leq e^{-\kappa_i(U-T)}/\kappa_i$ for $1 \leq i \leq p$, $\mathbb{E}^{Q^U} \left[\exp(\text{Tr}(\Gamma X_T) + \Lambda^\top Y_T) | \mathcal{F}_t \right] < \infty$ a.s. for any $t \in [0, T]$ and is given by (23).*

Let us mention that in practice, the formula above for $A^U(t, T)$, $D^U(t, T)$ and $B^U(t, T)$ requires to solve two different ODEs. It may be more convenient to use the following one that can be easily deduced from dynamics of (X, Y) under the U -forward measure:

$$\begin{cases} \frac{\partial B^U}{\partial t}(t, T) = \kappa^\top B^U(t, T), \quad B^U(T, T) = \Lambda, \\ -\frac{\partial D^U}{\partial t}(t, T) = 2\epsilon^2 D^U I_d^n D^U + D^U (b^U(t) + \epsilon I_d^n \rho (B^U)^\top c) + (b^U(t) + \epsilon I_d^n \rho (B^U)^\top c)^\top D^U + \frac{1}{2} c^\top B^U (B^U)^\top c \\ \quad + c^\top B^U B(U-t)^\top c + \epsilon D(U-t) I_d^n \rho (B^U)^\top c + \epsilon c^\top B^U \rho^\top I_d^n D(U-t), \quad D^U(T, T) = \Gamma, \\ -\frac{\partial A^U}{\partial t}(t, T) = B^U(t, T)^\top \kappa \theta + \text{Tr}(D^U(t, T)(\Omega + \epsilon^2(d-1)I_d^n)), \quad A^U(T, T) = 0. \end{cases} \quad (24)$$

3 Expansion of the volatility smile around the LGM

The goal of this section is to provide the asymptotic behaviour of the Caplet and Swaption prices when the volatility parameter ϵ is close to zero. The practical interest of these formulas is to give quickly a proxy for these prices. Thus, they give a tool to calibrate the model parameters to the smile. Let us mention here that expansions of Gram-Charlier type can be also be applied to price caplets and swaptions thanks to the affine structure of the

model, see for example Collin-Dufresne and Goldstein [10] and Tanaka et al. [33]. Some numerical examples are presented in [27] for the pricing of caplets in our model. Here, we only present the expansion with respect to ϵ since it is in accordance with our presentation of the model as a perturbation of the LGM.

The arguments that we use in this section to obtain the expansion have been developed in the book of Fouque et al. [23]. They rely on an expansion of the infinitesimal generator with respect to ϵ . Recently, this technique was applied by Bergomi and Guyon [7] to provide approximation under a multi factor model for the forward variance. Here, we have to take into account some specific features of the fixed income and work under the appropriate probability measure to apply these arguments. Not surprisingly the zero order term in the expansion is exactly the volatility of the LGM with a time-dependent variance-covariance matrix. More interestingly the higher order terms allow to confirm the intuitions on the role of the parameters and factors that determine the shape and dynamics of the volatility.

Last, we have to mention that the calculations presented in this section are rather formal. In particular, we implicitly assume that the caplet and swaption prices are smooth enough and admit expansions with respect to ϵ . A rigorous proof of these expansions is beyond the scope of this paper.

3.1 Price and volatility expansion for Caplets

From (4), the only quantity of interest in order to understand the Caplets volatility cube is what we call the forward Caplet price

$$\text{FCaplet}(t, T, \delta) = \mathbb{E}^{T+\delta} \left[(L_T(T, \delta) - K)^+ \mid \mathcal{F}_t \right],$$

which can be rewritten as a call option on the forward zero coupon bond $\frac{P_{t,T}}{P_{t,T+\delta}}$

$$\text{FCaplet}(t, T, \delta) = \frac{1}{\delta} \mathbb{E}^{T+\delta} \left[\left(\frac{P_{T,T}}{P_{T,T+\delta}} - (1 + \delta K) \right)^+ \mid \mathcal{F}_t \right].$$

Since (X, Y) is a Markov process, $\text{FCaplet}(t, T, \delta)$ is a function of (X_t, Y_t) and therefore we can define the forward price function

$$P(t, x, y) = \mathbb{E}^{T+\delta} \left[\left(\frac{P_{T,T}}{P_{T,T+\delta}} - (1 + \delta K) \right)^+ \mid X_t = x, Y_t = y \right]. \quad (25)$$

The goal of Subsection 3.1 is to obtain the second order expansion (27) of P with respect to ϵ .

3.1.1 A convenient change of variable

We want to get an expansion of the caplet price with respect to ϵ . To do so, we need a priori to get an expansion to ϵ of the infinitesimal generator of the process (X, Y) under the probability $Q^{T+\delta}$. However, we can make before a change of variable that simplifies this approach. Thus, we define

$$H_t = \Delta A(t, T, \delta) + \text{Tr}(\Delta D(t, T, \delta) X_t) + \Delta B(t, T, \delta)^\top Y_t,$$

with

$$\begin{aligned} \Delta A(t, T, \delta) &= A(t, T) - A(t, T + \delta) \\ (\Delta B, \Delta D)(t, T, \delta) &= (B, D)(T - t) - (B, D)(T + \delta - t) \end{aligned}$$

Thus, we have $\frac{P_{t,T}}{P_{t,T+\delta}} = e^{H_t}$. It is well known that $\frac{P_{t,T}}{P_{t,T+\delta}}$ is a martingale under $Q^{T+\delta}$, see e.g. Proposition 2.5.1 in Brigo and Mercurio [8]. Thus, we get by Itô calculus from (21) and (22) that (X, H) solve the following SDE

$$\begin{aligned} dX_t &= (\Omega + \epsilon^2(d-1)I_d^n + b^{T+\delta}(t)X_t + X_t(b^{T+\delta}(t))^\top)dt + \epsilon\sqrt{X_t}dW_t^{T+\delta}I_d^n + \epsilon I_d^n(dW_t^{T+\delta})^\top\sqrt{X_t}, \\ dH_t &= -\frac{1}{2}(\Delta B^\top cX_t c^\top \Delta B + 4\epsilon^2 \text{Tr}(\Delta D I_d^n \Delta D X_t) + 2\epsilon(\Delta B^\top cX_t \Delta D I_d^n \rho))dt \\ &\quad + \Delta B^\top c\sqrt{X_t}(dW_t^{T+\delta}\rho + \bar{\rho}dZ_t^{T+\delta}) + 2\epsilon \text{Tr}(\Delta D \sqrt{X_t}dW_t^{T+\delta}I_d^n). \end{aligned} \quad (26)$$

Therefore, $P(t, x, y) = \mathbb{E}^{T+\delta} \left[(e^{H_T} - (1 + \delta K))^+ | X_t = x, Y_t = y \right]$ only depends on (x, y) through (x, h) where $h = \Delta B(t, T, \delta)^\top y + \text{Tr}(\Delta D(t, T, \delta)x) + \Delta A(t, T, \delta)$, and we still denote by a slight abuse of notations

$$P(t, x, h) = \mathbb{E}^{T+\delta} \left[(e^{H_T} - (1 + \delta K))^+ | X_t = x, H_t = h \right].$$

Let us emphasize that this change of variable is crucial in order to apply an expansion procedure similar to the one of Bergomi and Guyon [7]. It allows to reduce the dimensionality of the underlying state variable. The variable H is one-dimensional and it is the only variable that appears in the payoff of the caplet. Though this is obvious from the definition of the model, we insist on the fact that the implied volatility of caplets is a function of the factors X only. This appears clearly in the SDE (26), H_t can be viewed as continuous version of the forward Libor rate and its volatility depends on the factors X only.

3.1.2 Expansion of the price

From the SDE (26), (9), (10) and (11), we get the following PDE representation of P :

$$\begin{aligned} \partial_t P + \mathcal{L}(t)P &= 0 \\ P(T, x, h) &= (e^h - (1 + \delta K))^+ \end{aligned}$$

where $\mathcal{L}(t)$ is the infinitesimal generator of (26). We assume that P admits a second order expansion

$$P = P_0 + \epsilon P_1 + \epsilon^2 P_2 + o(\epsilon^2). \quad (27)$$

Our goal is to calculate in a quite explicit way the value of P_0 , P_1 and P_2 . We assume in our derivations that these functions P_0 , P_1 and P_2 are smooth enough. To determine the value of P_0 , P_1 and P_2 , we proceed as Bergomi and Guyon [7] and make an expansion of the generator $\mathcal{L}(t) = \mathcal{L}_0(t) + \epsilon \mathcal{L}_1(t) + \epsilon^2 \mathcal{L}_2(t) + \dots$ in order to obtain the PDEs satisfied by P_0 , P_1 and P_2 . Namely, we obtain

$$\begin{aligned} \partial_t P_0 + \mathcal{L}_0(t)P_0 &= 0, & P_0(T, x, h) &= (e^h - (1 + \delta K))^+, \\ \partial_t P_1 + \mathcal{L}_0(t)P_1 + \mathcal{L}_1(t)P_0 &= 0, & P_1(T, x, h) &= 0, \\ \partial_t P_2 + \mathcal{L}_0(t)P_2 + \mathcal{L}_2(t)P_0 + \mathcal{L}_1(t)P_1 &= 0, & P_2(T, x, h) &= 0. \end{aligned}$$

Thus, we can solve first the PDE for P_0 , then for P_1 and so on. Let $\text{BS}(h, v) = \mathbb{E} \left[(\exp(h - \frac{1}{2}v + \sqrt{v}G) - (1 + \delta K))^+ \right]$ with $G \sim N(0, 1)$ denote the Black-Scholes price with realized volatility v . We obtain easily that

$$P_0(t, x, h) = \text{BS}(h, v(t, T, \delta, x)),$$

with

$$v(t, T, \delta, x) = \int_t^T \Delta B(u, T, \delta)^\top c X_{u-t}^0(x) c^\top \Delta B(u, T, \delta) du, \quad (28)$$

$$X_s^0(x) = e^{bs} \left(x + \int_0^s e^{-bu} \Omega e^{-b^\top u} du \right) e^{b^\top s}. \quad (29)$$

The higher order terms are given by³

$$P_1(t, x, h) = (c_1(t, T, \delta, x)(\partial_h^3 - \partial_h^2) + c_2(t, T, \delta, x)(\partial_h^2 - \partial_h)) P_0(t, x, h) \quad (30)$$

and

$$\begin{aligned} P_2(t, x, h) = & \left[(d_1(t, T, \delta, x)(\partial_h^2 - \partial_h)^2 + d_2(t, T, \delta, x)(\partial_h^2 - \partial_h)\partial_h + d_3(t, T, \delta, x)(\partial_h^2 - \partial_h)) \right. \\ & + \left(e_1(t, T, \delta, x)(\partial_h^2 - \partial_h)^2 \partial_h^2 + e_2(t, T, \delta, x)(\partial_h^2 - \partial_h)^2 \partial_h + e_3(t, T, \delta, x)(\partial_h^2 - \partial_h)^2 \right. \\ & \left. \left. + e_4(t, T, \delta, x)(\partial_h^2 - \partial_h)\partial_h^2 + e_5(t, T, \delta, x)(\partial_h^2 - \partial_h)\partial_h + e_6(t, T, \delta, x)(\partial_h^2 - \partial_h) \right) \right] P_0(t, x, h). \end{aligned} \quad (31)$$

³The details of these simple but tedious calculations are available online <http://arxiv.org/abs/1412.7412> in the first draft of this paper for the caplets and swaptions.

The coefficients c_i , d_i and e_i are given in Appendix A.1. We recall that the derivatives $\partial_h^i P_0$ of P_0 with respect to h can be calculated explicitly, so that the expansion is very efficient from the point of view of the computational time, see Section 5.

Remark 11 — *It is easy to obtain then the expansion $v_{\text{Imp}} = v_0 + \epsilon v_1 + \epsilon^2 v_2 + o(\epsilon^2)$ of the implied volatility defined by $\delta\text{FCaplet}(t, T, \delta) = \text{BS}(h, v_{\text{Imp}})$. We obtain as expected $v_0 = v(t, T, \delta, x)$ and*

$$\frac{v_1}{2} = c_2(t, T, \delta, x) + c_1(t, T, \delta, x) \left(\frac{1}{2} - \frac{h - \log(1 + \delta K)}{v_0} \right). \quad (32)$$

Since neither c_1 nor c_2 depend on the strike, the skew is at the first order in ϵ proportional to c_1 , that is at its turn a linear function of ρ . We have in particular a flat smile at the first order when $\rho = 0$, as one may expect.

3.2 Price and volatility expansion for Swaptions

From (5), the only quantity of interest in order to understand the swaptions volatility cube is what we call the annuity-forward swaption price

$$\text{AFSwaption}(t, T, m, \delta) = \mathbb{E}^A \left[(S_t(T, m, \delta) - K)^+ | \mathcal{F}_t \right].$$

It is standard to view swaptions as a basket option of forward Libor rates with stochastic weights, we have

$$S_t(T, m, \delta) = \sum_{i=1}^m \omega_t^i L_t(T + (i-1)\delta, T + i\delta) \quad (33)$$

$$\omega_t^i = \frac{P_{t, T+i\delta}}{\sum_{i=1}^m P_{t, T+i\delta}}. \quad (34)$$

The difficulty here comes from the fact that forward Libor rates, and the stochastic weights are complicated functions of the state variables (X, Y) . The first implication is that the change of measure between \mathbb{P} and Q^A is also complicated and the dynamics of the state variables under this new measure is quite unpleasant to work with. The second implication is that we cannot directly operate a convenient change of variable as we did for caplets. In order to derive an expansion for swaptions we thus proceed stepwise. First, we use a standard approximation that freezes the weights at their initial value (see for example Brigo and Mercurio [8] p. 239, d'Aspremont [16] and Piterbarg [29]). This is justified by the fact the variation of the weights is less important than the variation of the forward Libor rates⁴. Second, we use a similar approximation for the swap rate. Thus, the approximated swap rate is an affine function of the underlying state variables, which enables us to take advantage of the affine structure of the model. Let us mention that this technique is similar to the quadratic approximation of the swap rate proposed by Piterbarg in [29]. Finally we perform our expansion on the affine approximation of the swap rates and obtain the second order expansion (41), which is the main result of Subsection 3.2.

3.2.1 Dynamics of the factors under the annuity measure

The annuity measure knowing the information up to date t , $Q^A | \mathcal{F}_t$ is defined by

$$\frac{dQ^A}{d\mathbb{P}} \Big|_{\mathcal{F}_t} = e^{-\int_t^T r_s ds} \frac{A_T(T, m, \delta)}{A_t(T, m, \delta)}.$$

It comes from the martingale property of discounted asset prices under the risk neutral measure that

$$\frac{d \left(e^{-\int_0^t r_s ds} A_t(T, m, \delta) \right)}{e^{-\int_0^t r_s ds} A_t(T, m, \delta)} = \sum_{i=1}^m \omega_t^i \left(B(T + i\delta - t)^\top c \sqrt{X_t} (dW_t \rho + \bar{\rho} dZ_t) + 2\epsilon \text{Tr} \left(D(T + i\delta - t) \sqrt{X_t} dW_t I_d^n \right) \right). \quad (35)$$

⁴To the best of our knowledge there have been very few attempts to quantify either theoretically or numerically this statement. In [16] d'Aspremont investigates the accuracy of the approximation for pricing swaptions in the log-normal BGM model, he shows that the approximation is less efficient for long maturities and long tenors.

From Girsanov's theorem, the change of measure is given by

$$\begin{aligned} dW_t^A &= dW_t - \sqrt{X_t} \left(2\epsilon \sum_{i=1}^m \omega_t^i D(T + i\delta - t) I_d^n + c^\top B(T + i\delta - t) \rho^\top \right) dt, \\ dZ_t^A &= dZ_t - \bar{\rho} \sqrt{X_t} c^\top \sum_{i=1}^m \omega_t^i B(T + i\delta - t) dt. \end{aligned}$$

This allows us to calculate from (6) and (7) the dynamics of the state variables under the annuity measure Q^A :

$$\begin{aligned} dY_t &= \left(\kappa(\theta - Y_t) + cX_t c^\top \sum_{k=1}^m \omega_t^k B(T + k\delta - t) + 2\epsilon cX_t \sum_{k=1}^m \omega_t^k D(T + k\delta - t) I_d^n \rho \right) dt \\ &\quad + c\sqrt{X_t}(\bar{\rho}dZ_t^A + dW_t^A \rho), \end{aligned} \quad (36)$$

$$dX_t = (\Omega + \epsilon^2(d-1)I_d^n + b^A(t)X_t + X_t(b^A(t))^\top)dt + \epsilon \left(\sqrt{X_t}dW_t^A I_d^n + I_d^n(dW_t^A)^\top \sqrt{X_t} \right), \quad (37)$$

where $b^A(t) = b + \epsilon I_d^n \rho \sum_{k=1}^m \omega_t^k B(T + k\delta - t)^\top c + 2\epsilon^2 I_d^n \sum_{k=1}^m \omega_t^k D(T + k\delta - t)$.

3.2.2 An affine approximation of the forward swap rate

The forward swap rate is a martingale under the annuity measure Q^A . Therefore, we can only focus on the martingale terms when applying Itô's formula to $\frac{P_{t,T} - P_{t,T+m\delta}}{\sum_{i=1}^m P_{t,T+i\delta}}$, and we get from (20) that

$$\begin{aligned} dS_t(T, m, \delta) &= \\ &= \left[\omega_t^0 B(T - t)^\top - \omega_t^m B(T + m\delta - t)^\top - S_t(T, m, \delta) \sum_{k=1}^m \omega_t^k B(T + k\delta - t)^\top \right] c\sqrt{X_t}(dW_t^A \rho + \bar{\rho}dZ_t^A) \\ &\quad + 2\epsilon \text{Tr} \left(\left[\omega_t^0 D(T - t) - \omega_t^m D(T + m\delta - t) - S_t(T, m, \delta) \sum_{k=1}^m \omega_t^k D(T + k\delta - t) \right] \sqrt{X_t} dW_t^A I_d^n \right) \end{aligned} \quad (38)$$

By a slight abuse of notations, we will now drop the (T, m, δ) dependence of the swap rate and simply denote by S_t its time t value. We now use the standard approximation that consists in freezing the weights ω_t^k and the value of the swap rate S_t in the right-hand side to their value at zero. We then have

$$dS_t = B^S(t)^\top c\sqrt{X_t}(dW_t^A \rho + \bar{\rho}dZ_t^A) + 2\epsilon \text{Tr} \left(D^S(t) \sqrt{X_t} dW_t^A I_d^n \right), \quad (39)$$

where

$$(B, D)^S(t) = \omega_0^0(B, D)(T - t) - \omega_0^m(B, D)(T + m\delta - t) - S_0(T, m, \delta) \sum_{k=1}^m \omega_0^k(B, D)(T + k\delta - t).$$

These coefficients are time-dependent and deterministic. We do the same approximation on X and get

$$dX_t = (\Omega + \epsilon^2(d-1)I_d^n + b_0^A(t)X_t + X_t(b_0^A(t))^\top)dt + \epsilon \left(\sqrt{X_t}dW_t^A I_d^n + I_d^n(dW_t^A)^\top \sqrt{X_t} \right), \quad (40)$$

where

$$b_0^A(t) = b + \epsilon I_d^n \rho \sum_{k=1}^m \omega_0^k B(T + k\delta - t)^\top c + 2\epsilon^2 I_d^n \sum_{k=1}^m \omega_0^k D(T + k\delta - t).$$

Thanks to this approximation, we remark that the process, that we still denote by (S_t, X_t) for simplicity, is now affine. This enables us to use again the same argument as for the Caplet prices to get an expansion of the price. The only difference lies in the fact the expansion is around the Gaussian model rather than around the log-normal model.

3.2.3 The swaption price expansion

Let $P^S(t, x, s) = \mathbb{E}^A \left[(S_t - K)^+ | S_t = s, X_t = x \right]$ denote the price of the Swaption at time $t \in [0, T]$. It solves the following pricing PDE

$$\partial_t P^S + \mathcal{L}(t) P^S = 0, \quad t \in (0, T), \quad P^S(T, x, s) = (s - K)^+,$$

where \mathcal{L} is the infinitesimal generator of the SDE (39) and (40). Again, we assume that P^S admits a second order expansion

$$P^S = P_0^S + \epsilon P_1^S + \epsilon^2 P_2^S + o(\epsilon^2) \quad (41)$$

and that the functions P_0^S , P_1^S and P_2^S are smooth enough. Let $\text{BH}(s, v) = \mathbb{E} \left[(s + \sqrt{v}G - K)^+ \right]$ with $G \sim N(0, 1)$ denote the European call price with strike K in the Bachelier model with realized volatility $v > 0$ and spot price $s \in \mathbb{R}$. We obtain

$$P_0^S(t, x, s) = \text{BH}(s, v^S(t, T, x)), \quad (42)$$

where

$$v^S(t, T, x) = \int_t^T B^S(r)^\top c X_{r-t}^0(x) c^\top B^S(r) dr, \quad (43)$$

and $X_s^0(x)$ is defined by (29). The higher order term are

$$P_1^S(t, x, s) = (c_1^S(t, T, x) \partial_s^3 + c_2^S(t, T, x) \partial_s^2) \text{BH}(s, v^S(t, T, x)), \quad (44)$$

$$\begin{aligned} P_2^S(t, x, s) = & \left[d_1^S(t, T, x) \partial_s^4 + d_2^S(t, T, x) \partial_s^3 + d_3^S(t, T, x) \partial_s^2 \right. \\ & + e_1^S(t, T, x) \partial_s^6 + e_2^S(t, T, x) \partial_s^5 + e_3^S(t, T, x) \partial_s^4 \\ & \left. + e_4^S(t, T, x) \partial_s^4 + e_5^S(t, T, x) \partial_s^3 + e_6^S(t, T, x) \partial_s^2 \right] \text{BH}(s, v^S(t, T, x)), \end{aligned} \quad (45)$$

where the coefficients c_i^S , d_i^S and e_i^S are given in Appendix A.2. Again, the derivatives of P^S with respect to s can be calculated explicitly, which makes this formula very efficient from a computational point of view.

3.3 Numerical results

We now assess on some examples the accuracy of the expansions we have developed. In practice we are interested in knowing up to what level of parameters and for what set of maturities and tenors the accuracy of the expansion is satisfactory. Let us recall that our expansion for caplets results from the combination of two expansions, the first on the support matrix function D up to the order 1 in ϵ is given by (52) and (53), the second on the infinitesimal generator of the Markov process (X, H) defined by (26). By construction the approximation of $D(\tau)$ will be more accurate for a small τ . As a consequence, for a given set of parameters, the full expansion will likely to be more accurate for short maturities, short tenors caplets. The expansion for swaptions results from a supplementary approximation step, which consists in freezing the weights ω^i in the diffusion of the Markov process (X, S) defined by (37) and (38). This approximation can be inaccurate for long maturities and long tenors swaptions. Therefore, we expect the full expansion to be more accurate for short maturities, short tenors swaptions.

We assess the quality of the price expansion for caplets and swaptions. We compare the expanded price with the price computed using Monte Carlo simulation and the discretization scheme 1 described in Section 4 on a regular time grid. The expanded prices and the Monte Carlo prices are compared in terms of the normal implied volatility of the forward Libor rate for caplets and of the forward swap rate for swaptions. The implied volatility is given in basis points (10^{-4}). In abscissa is indicated the difference between the strike and the at-the-money value, and the unit is one percent. A $6M \times 2Y$ caplet will denote a caplet with maturity $T = 2$ years and tenor $\delta = 0.5$ years, while a $5Y \times 2Y$ swaption will denote a swaption with maturity $T = 2$ years and tenor $m\delta = 5$ years.

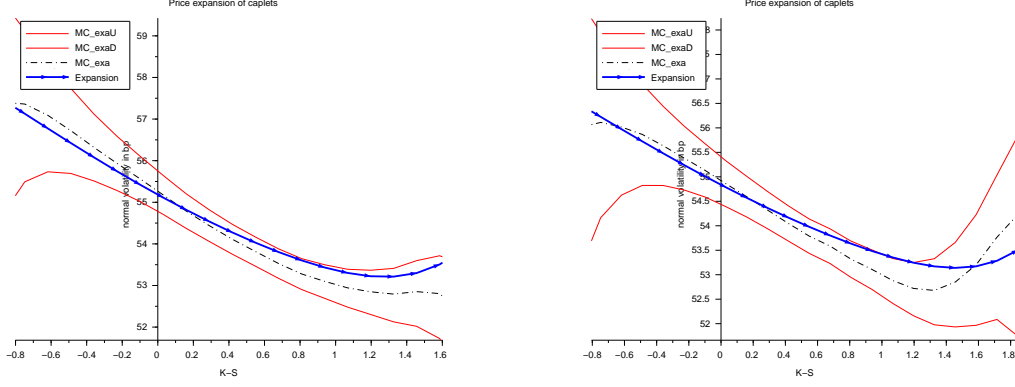


Figure 2: $\rho^\top = (-0.4, -0.2)$. Plot of the expanded smile of a $1Y \times 1Y$ caplet against the Monte Carlo smile obtained with 100000 paths and a discretization grid of 4 points for different values of the parameter ϵ , respectively from left to right $\epsilon = 0.002$ and $\epsilon = 0.0015$. The forward Libor rate value is $L(0, 1Y, 1Y) = 1.02\%$.

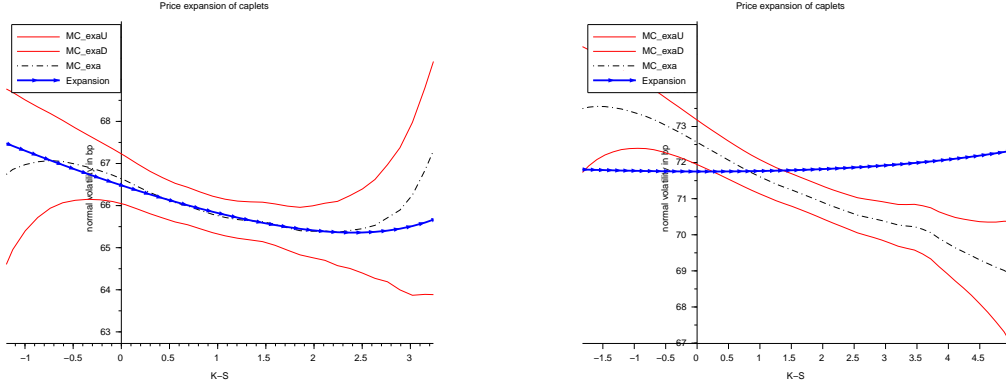


Figure 3: $\rho^\top = (-0.4, -0.2)$ and $\epsilon = 0.0015$. Left: plot of the expanded smile of a $6M \times 2Y$ caplet against the Monte Carlo smile. Right: plot of the expanded smile of a $6M \times 5Y$ caplet against the Monte Carlo smile. The Monte Carlo smile is obtained with 100000 paths and a discretization grid of 8 points. The forward Libor rates values are $L(0, 6M, 2Y) = 1.14\%$ and $L(0, 6M, 5Y) = 1.35\%$.

We have tested different sets of model parameters. The parameters values have been chosen in such a way that the yield curve and volatility levels generated by the model are in line with today's US and EUR interest rates market levels. Here, we only consider the following parameter set with $p = 2$ and $d = 2$:

$$\begin{aligned} \kappa &= \text{diag}(0.1, 1), \quad c = I_d, \quad b = -\text{diag}(0.41, 0.011), \quad \Omega = -(bx_\infty + x_\infty b^\top) + 0.4I_d, \quad \gamma = 0.001I_d \\ x &= 10^{-4} \begin{pmatrix} 2.25 & -1.2 \\ -1.2 & 1 \end{pmatrix}, \quad x_\infty = 10^{-4} \begin{pmatrix} 1 & -0.125 \\ -0.125 & 0.25 \end{pmatrix}. \end{aligned} \quad (46)$$

We note that $-(b + b^\top) = -2b$ is positive definite. We know from Remark 8 that the condition of non-explosion will be verified in general for these set of parameters when ϵ is small enough, and we have checked that the yield curve given by this parameter set is well defined up to 50 years.

In all the graphics the dotted line gives the Monte Carlo smile obtained with 100000 simulation paths, the solid line with small arrows is the expanded smile, the two continuous solid lines are the upper and lower bounds of the 95% confidence interval of the Monte Carlo price. Figures 2 and 3 show the accuracy of the expansion for the valuation of caplets. The approximation is accurate for expiries up to 2 years and less accurate with the same parameters for longer expiries. For maturities up to 2 years, the at-the-money volatility of the expanded smile is

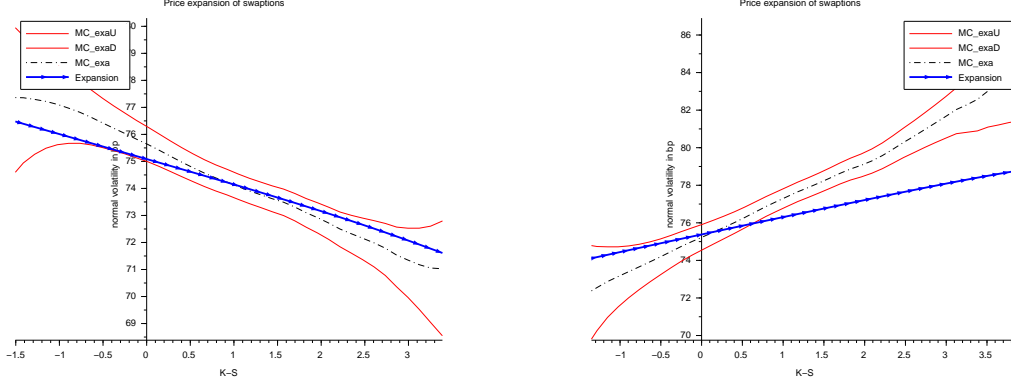


Figure 4: $\epsilon = 0.0015$. Plot of the expanded smile of a $5Y \times 2Y$ swaption with coupon payment frequency of 6 months against the Monte Carlo smile obtained with 100000 paths and a discretization grid of 8 points for different values of the parameter ρ , from left to right $\rho^\top = (-0.4, -0.2)$ and $\rho^\top = (0.4, 0.2)$. The forward swap rate value is $S(0, 5Y, 2Y) = 1.3\%$.

almost identical to the Monte Carlo smile and the whole expanded smile stays within the 95% confidence interval. Figure 4 shows the accuracy of the expansion for the valuation of swaptions. We observe that the expansion is more accurate for negative values of the correlation parameters ρ (a similar behaviour is observed for Caplets). This can be intuitively understood from the Riccati equation (15): a negative ρ pushes D to zero while a positive one pushes D away from zero, and the expansion that we use on D (see (52) and (53)) is then less accurate. Overall the expansion is accurate at-the-money and is much less accurate out-of-the-money. For example, the graphic on the right hand side of Figure 3 shows that the expanded smile of the 6 months maturity 5 years expiry smile is quite inaccurate and the expanded smile fails to fit the skew of the Monte Carlo smile. However, the difference in the at-the-money volatility between the expanded price and Monte Carlo is around 1 bp.

To sum up, the second order expansion is basically accurate for small perturbations and small maturities. Otherwise, one should be careful and rely on other methods such as the Monte-Carlo method or Fourier inversion method. Nonetheless, as discussed in Section 5, the calculation of the expansion is much faster than the other methods. It may thus be relevant to start a calibration routine and select a reasonable set of parameters.

4 Second order discretization schemes for Monte Carlo simulation

The goal of this section is to construct discretization schemes for the process (X, Y) defined by (6) and (7). It is crucial to have an efficient way to simulate the model in order to use it in practice. Ideally, the model should be calibrated to market data to vanilla options such as caplets and swaptions and then be used to calculate exotic option prices. The calculation of these prices is generally made with a Monte-Carlo algorithm which requires to simulate the process (X, Y) .

It is worth to recall that the standard Euler-Maruyama scheme is not well defined for square-root diffusions even in dimension one, see Alfonsi [2]. We have then to consider a different scheme. We use here the splitting technique that is already used by Ahdida and Alfonsi [1] for Wishart processes. We explain here briefly the main line of this method and refer to [2] for precise statements in a framework that embeds affine diffusions. Let us consider that we want to approximate an SDE ξ with infinitesimal generator \mathcal{L} on the regular time grid $t_i = iT/N$, for $i = 0, \dots, N$. A scheme is fully described by a probability law $\hat{p}_x(t, dz)$ that approximates the law of ξ_t given $\xi_0 = x$. We denote by $\hat{\xi}_t^x$ a random variable following this law. Then, the law of the corresponding discretization scheme $(\hat{\xi}_{t_i}, 0 \leq i \leq N)$ is as follows: $\hat{\xi}_0 = \xi_0$ and $\hat{p}_{\hat{\xi}_{t_i}}(T/N, dz)$ is the conditional law of $\hat{\xi}_{t_{i+1}}$ given $(\hat{\xi}_{t_j}, 0 \leq j \leq i)$. Then, one would like to know the error made when using the approximation scheme instead of the original process ξ . We have basically the following result, up to technical details that are given

in [2]. If $\hat{\xi}_t^x$ satisfies the following expansion

$$\mathbb{E}[f(\hat{\xi}_t^x)] = f(x) + t\mathcal{L}f(x) + \frac{t^2}{2}\mathcal{L}^2f(x) + O(t^3) \quad (47)$$

for any smooth function f , then

$$\exists C > 0, |\mathbb{E}[f(\hat{\xi}_{t_N})] - \mathbb{E}[f(\xi_{t_N})]| \leq C/N^2.$$

Thus, to get a weak error of order 2, we mainly have to construct a scheme $\hat{\xi}_t^x$ that satisfies (47). We can construct iteratively second order schemes by splitting the infinitesimal generator. In fact, let us assume that $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$ and that $\hat{\xi}_t^{i,x}$ is a second order scheme for \mathcal{L}_i . Let \mathbf{B} be an independent Bernoulli variable with parameter $1/2$. Then, the following schemes

$$\hat{\xi}_{t/2}^{1,\hat{\xi}_t^{2,\hat{\xi}_t^{1,x}}} \text{ and } \mathbf{B}\hat{\xi}_t^{2,\hat{\xi}_t^{1,x}} + (1 - \mathbf{B})\hat{\xi}_t^{1,\hat{\xi}_t^{2,x}} \quad (48)$$

satisfy (47) and are thus second order schemes for \mathcal{L} . Therefore, a strategy to construct a second order scheme is to split the infinitesimal generator into elementary pieces for which second order schemes or even exact schemes are known.

To use this splitting technique, we first have to calculate the infinitesimal generator of (X, Y) . It is defined for a \mathcal{C}^2 function $f : \mathcal{M}_d(\mathbb{R}) \times \mathbb{R}^p \rightarrow \mathbb{R}$ by $\mathcal{L}f(x, y) = \lim_{t \rightarrow 0+} \frac{\mathbb{E}[f(X_t, Y_t)] - f(x, y)}{t}$. From (9), (10) and (11), we easily get

$$\begin{aligned} \mathcal{L} = & \sum_{m=1}^p (\kappa(\theta - y))_m \partial_{y_m} + \sum_{1 \leq i, j \leq d} (\Omega + (d-1)\epsilon^2 I_d^n + bx + xb^\top)_{i,j} \partial_{x_{i,j}} \\ & + \frac{1}{2} \sum_{m,m'=1}^p (cxc^\top)_{m,m'} \partial_{y_m} \partial_{y_{m'}} + \frac{1}{2} \sum_{m=1}^p \sum_{1 \leq i, j \leq d} \epsilon [(cx)_{m,i} (I_d^n \rho)_j + (cx)_{m,j} (I_d^n \rho)_i] \partial_{x_{i,j}} \partial_{y_m} \\ & + \frac{1}{2} \sum_{1 \leq i, j, k, l \leq d} \epsilon^2 [x_{i,k} (I_d^n)_{j,l} + x_{i,l} (I_d^n)_{j,k} + x_{j,k} (I_d^n)_{i,l} + x_{j,l} (I_d^n)_{i,k}] \partial_{x_{i,j}} \partial_{x_{k,l}}. \end{aligned}$$

Here, ∂_{y_m} denotes the partial derivative with respect to the m -th coordinate in \mathbb{R}^p and $\partial_{x_{i,j}}$ the partial derivative with respect to the element at the i -th row and j -th column. When $\rho = 0$, this operator is simply the sum of the infinitesimal generators for X and the generator for Y when X is frozen. We know from [1] a second order scheme for X . When X is frozen, Y follows an Ornstein-Uhlenbeck process and the law of Y_t is a Gaussian vector that can be sampled exactly. By using the composition rule (48), we get a second order scheme for (X, Y) .

Thus, the difficulty here comes from the correlation between X and Y that has to be handled with care. We first make some simplifications. The first term $\sum_{m=1}^p (\kappa(\theta - y))_m \partial_{y_m}$ is the generator of the linear Ordinary Differential Equation $y'(t) = \kappa(\theta - y(t))$ that is solved exactly by $y(t) = e^{-\kappa t} y(0) + (I_p - e^{-\kappa t})\theta$. Therefore, it is sufficient to have a second order scheme for $\mathcal{L} - \sum_{m=1}^p (\kappa(\theta - y))_m \partial_{y_m}$, which is the generator of (6) and (7) when $\kappa = 0$. When $\kappa = 0$, we have $Y_t = y + c(\tilde{Y}_t - \tilde{Y}_0)$ with

$$\tilde{Y}_t = \tilde{Y}_0 + \int_0^t \sqrt{X_s} [\bar{\rho} dZ_s + dW_s \rho].$$

We can then focus on getting a second order scheme for (X, \tilde{Y}) , which amounts to work with $p = d$ and $c = I_d$. It is therefore sufficient to find a second order scheme for the SDE

$$\begin{aligned} Y_t &= y + \int_0^t \sqrt{X_s} [\bar{\rho} dZ_s + dW_s \rho], \\ X_t &= x + \int_0^t (\Omega + (d-1)\epsilon^2 I_d^n + bX_s + X_s b^\top) ds + \epsilon \int_0^t \sqrt{X_s} dW_s I_d^n + I_d^n dW_s^\top \sqrt{X_s}, \end{aligned}$$

with the infinitesimal generator

$$\begin{aligned} \mathcal{L} = & \sum_{1 \leq i, j \leq d} (\Omega + (d-1)\epsilon^2 I_d^n + bx + xb^\top)_{i,j} \partial_{x_{i,j}} + \frac{1}{2} \sum_{m=1}^d \sum_{1 \leq i, j \leq d} \epsilon [x_{m,i} (I_d^n \rho)_j + x_{m,j} (I_d^n \rho)_i] \partial_{x_{i,j}} \partial_{y_m} \\ & + \frac{1}{2} \sum_{m,m'=1}^d x_{m,m'} \partial_{y_m} \partial_{y_{m'}} + \frac{1}{2} \sum_{1 \leq i, j, k, l \leq d} \epsilon^2 [x_{i,k} (I_d^n)_{j,l} + x_{i,l} (I_d^n)_{j,k} + x_{j,k} (I_d^n)_{i,l} + x_{j,l} (I_d^n)_{i,k}] \partial_{x_{i,j}} \partial_{x_{k,l}}. \end{aligned} \quad (49)$$

4.1 A second order scheme

For $1 \leq q \leq d$, we define $e_d^q \in \mathcal{S}_d^+(\mathbb{R})$ by $(e_d^q)_{k,l} = \mathbb{1}_{k=l=q}$ and $g_d^q \in \mathbb{R}^d$ by $(g_d^q)_k = \mathbb{1}_{q=k}$ so that $I_d^n = \sum_{q=1}^n e_d^q$ and $I_d^n \rho = \sum_{q=1}^n \rho_q g_d^q$. We define

$$\begin{aligned} \mathcal{L}_q^c = & \epsilon^2(d-1)\partial_{x_{q,q}} + \frac{1}{2} \sum_{m=1}^d \sum_{1 \leq i,j \leq d} \epsilon \rho_q [x_{m,i}(g_d^q)_j + x_{m,j}(g_d^q)_i] \partial_{x_{i,j}} \partial_{y_m} + \frac{\rho_q^2}{2} \sum_{m,m'=1}^d x_{m,m'} \partial_{y_m} \partial_{y_{m'}} \\ & + \frac{1}{2} \sum_{1 \leq i,j,k,l \leq d} \epsilon^2 [x_{i,k}(e_d^q)_{j,l} + x_{i,l}(e_d^q)_{j,k} + x_{j,k}(e_d^q)_{i,l} + x_{j,l}(e_d^q)_{i,k}] \partial_{x_{i,j}} \partial_{x_{k,l}}. \end{aligned} \quad (50)$$

We consider the splitting $\mathcal{L} = \mathcal{L}' + \mathcal{L}'' + \sum_{q=1}^n \mathcal{L}_q^c$ of the operator (49), with

$$\begin{aligned} \mathcal{L}' &= \sum_{1 \leq i,j \leq d} (\Omega + bx + xb^\top)_{i,j} \partial_{x_{i,j}}, \\ \mathcal{L}'' &= \left(1 - \sum_{q=1}^n \rho_q^2\right) \frac{1}{2} \sum_{m,m'=1}^d x_{m,m'} \partial_{y_m} \partial_{y_{m'}}. \end{aligned}$$

The operator \mathcal{L}' is the one of the linear ODE $x'(t) = \Omega + (d-1)\epsilon^2 I_d^n + bx + xb^\top$ that can be solved exactly and stays in the set of semidefinite positive matrices, see Lemma 27 in [1]. The operator \mathcal{L}'' is the one of $Y_t'' = y'' + \sqrt{1 - \sum_{q=1}^n \rho_q^2} \sqrt{x} Z_t$, which can be sampled exactly since it is a Gaussian vector with mean y'' and covariance matrix $(1 - |\rho|^2)tx$. The operator \mathcal{L}_q^c is the infinitesimal generator of the following SDE

$$\begin{cases} Y_t &= y + \rho_q \int_0^t \sqrt{X_s} dW_s g_d^q, \\ X_t &= x + \int_0^t (d-1)\epsilon^2 e_d^q ds + \epsilon \int_0^t \sqrt{X_s} dW_s e_d^q + e_d^q dW_s^\top \sqrt{X_s}. \end{cases} \quad (51)$$

Thus, X follows an elementary Wishart process and stays in $\mathcal{S}_d^+(\mathbb{R})$. Using the notation of [1], X_t follows the law $WIS_d(x, d-1, 0, e_d^q, \epsilon^2 t)$. Theorems 9 and 16 in [1] gives respectively an exact and a second (or higher) discretization scheme for this process. We now explain how to calculate Y_t once that X_t has been sampled. From (51), we have for $1 \leq i \leq d$,

$$\begin{aligned} d(Y_t)_i &= \rho_q \sum_{j=1}^d (\sqrt{X_t})_{i,j} (dW_t)_{j,q}, \\ d(X_t)_{q,i} &= \epsilon \sum_{j=1}^d (\sqrt{X_t})_{i,j} (dW_t)_{j,q} + \mathbb{1}_{i=q} \left[(d-1)\epsilon^2 dt + \sum_{j=1}^d (\sqrt{X_t})_{q,j} (dW_t)_{j,q} \right]. \end{aligned}$$

This yields to

$$\begin{aligned} (Y_t)_i &= y_i + \frac{\rho_q}{\epsilon} ((X_t)_{q,i} - x_{q,i}), \text{ if } i \neq q, \\ (Y_t)_q &= y_i + \frac{\rho_q}{2\epsilon} [(X_t)_{q,q} - x_{q,q} - \epsilon^2(d-1)t]. \end{aligned}$$

Using these formula together with the exact (resp. second order) scheme for X_t , we get an exact (resp. second order) scheme for (51). By using the composition rules (48), we get a second order scheme for (49).

4.2 A faster second order scheme when $\Omega - \epsilon^2 I_d^n \in \mathcal{S}_d^+(\mathbb{R})$

As explained in [1], the sampling of each elementary Wishart process in \mathcal{L}_q requires a Cholesky decomposition that has a time complexity of $O(d^3)$. Since the second order scheme proposed above calls $n \leq d$ times this routine, the whole scheme requires at most $O(d^4)$ operations. However, by adapting an idea that has been already used

in [1] for Wishart processes, it is possible to get a faster scheme if we assume in addition that $\Omega - \epsilon^2 I_d^n \in \mathcal{S}_d^+(\mathbb{R})$. We now present this alternative scheme that only requires $O(d^3)$ operations.

We consider the splitting $\mathcal{L} = \tilde{\mathcal{L}}' + \tilde{\mathcal{L}}'' + \hat{\mathcal{L}}$ of the operator (49), with

$$\begin{aligned}\tilde{\mathcal{L}}' &= \sum_{1 \leq i, j \leq d} (\Omega - \epsilon^2 I_d^n + bx + xb^\top)_{i,j} \partial_{x_{i,j}} \\ \hat{\mathcal{L}} &= \sum_{1 \leq i \leq n} d\epsilon^2 \partial_{x_{i,i}} + \frac{1}{2} \sum_{m=1}^d \sum_{1 \leq i, j \leq d} \epsilon [x_{m,i} (I_d^n \rho)_j + x_{m,j} (I_d^n \rho)_i] \partial_{x_{i,j}} \partial_{y_m} + \frac{\sum_{q=1}^n \rho_q^2}{2} \sum_{m, m'=1}^d x_{m, m'} \partial_{y_m} \partial_{y_{m'}} \\ &\quad + \frac{1}{2} \sum_{1 \leq i, j, k, l \leq d} \epsilon^2 [x_{i,k} (I_d^n)_{j,l} + x_{i,l} (I_d^n)_{j,k} + x_{j,k} (I_d^n)_{i,l} + x_{j,l} (I_d^n)_{i,k}] \partial_{x_{i,j}} \partial_{x_{k,l}}.\end{aligned}$$

Again, $\tilde{\mathcal{L}}'$ is the operator of the linear ODE $x'(t) = \Omega - \epsilon^2 I_d^n + (d-1)\epsilon^2 I_d^n + bx + xb^\top$ that can be solved exactly and stays in the set of semidefinite positive matrices by Lemma 27 in [1] since $\Omega - \epsilon^2 I_d^n \in \mathcal{S}_d^+(\mathbb{R})$. We have already seen above that the generator \mathcal{L}'' can be sampled exactly, and we focus now on the sampling of $\hat{\mathcal{L}}$. It relies on the following result.

Lemma 12 — For $x \in \mathcal{S}_d^+(\mathbb{R})$ we consider $c \in \mathcal{M}_d(\mathbb{R})$ such that $c^\top c = x$. We define $U_t = c + \epsilon W_t I_d^n$, $X_t = U_t^\top U_t$ and $Y_t = y + \int_0^t U_s^\top dW_s I_d^n \rho$. Then, the process (X, Y) has the infinitesimal generator $\hat{\mathcal{L}}$.

Proof : For $1 \leq i, j, m \leq d$, we have $d(X_t)_{i,j} = \epsilon \sum_{k=1}^d ((U_t)_{k,i} (dW_t)_{k,j} \mathbb{1}_{j \leq n} + (U_t)_{k,j} (dW_t)_{k,i} \mathbb{1}_{i \leq n}) + \mathbb{1}_{i=j \leq n} d\epsilon^2 dt$ and $d(Y_t)_m = \sum_{k,l=1}^d (U_t)_{k,m} (dW_t)_{k,l} (I_d^n \rho)_l$. This leads to

$$\begin{aligned}\langle d(Y_t)_m, d(Y_t)_{m'} \rangle &= \sum_{k,l=1}^d (U_t)_{k,m} (U_t)_{k,m'} (I_d^n \rho)_l^2 dt = \left(\sum_{l=1}^n \rho_l^2 \right) (X_t)_{m,m'} dt, \\ \langle d(Y_t)_m, d(X_t)_{i,j} \rangle &= \epsilon [(I_d^n \rho)_j (X_t)_{m,i} + (I_d^n \rho)_i (X_t)_{m,j}] dt, \\ \langle d(X_t)_{i,j}, d(X_t)_{k,l} \rangle &= \epsilon^2 [(X_t)_{i,k} (I_d^n)_{j,l} + (X_t)_{i,l} (I_d^n)_{j,k} + (X_t)_{j,k} (I_d^n)_{i,l} + (X_t)_{j,l} (I_d^n)_{i,k}] dt,\end{aligned}$$

which precisely gives the generator $\hat{\mathcal{L}}$. \square

Thanks to Lemma 12, it is sufficient to construct a second order scheme for (U, Y) . Since $\langle d(Y_t)_m, d(U_t)_{i,j} \rangle = \epsilon (U_t)_{i,m} (I_d^n \rho)_j dt$, the infinitesimal generator $\bar{\mathcal{L}}$ of (U, Y) is given by

$$\bar{\mathcal{L}} = \frac{\epsilon^2}{2} \sum_{i=1}^d \sum_{j=1}^n \partial_{x_{i,j}}^2 + \frac{\epsilon}{2} \sum_{i,m=1}^d \sum_{j=1}^n \rho_j x_{i,m} \partial_{x_{i,j}} \partial_{y_m} + \frac{\sum_{q=1}^n \rho_q^2}{2} \sum_{m, m'=1}^d (x^\top x)_{m, m'} \partial_{y_m} \partial_{y_{m'}}.$$

We use now the splitting $\bar{\mathcal{L}} = \sum_{q=1}^n \bar{\mathcal{L}}_q$ with

$$\bar{\mathcal{L}}_q = \frac{\epsilon^2}{2} \sum_{i=1}^d \partial_{x_{i,q}}^2 + \frac{\epsilon}{2} \sum_{i,m=1}^d \rho_q x_{i,m} \partial_{x_{i,q}} \partial_{y_m} + \frac{\rho_q^2}{2} \sum_{m, m'=1}^d (x^\top x)_{m, m'} \partial_{y_m} \partial_{y_{m'}}.$$

By straightforward calculus, we find that $\bar{\mathcal{L}}_q$ is the generator of the following SDE

$$dY_t = \rho_q U_t^\top dW_t g_d^q, \quad dU_t = \epsilon dW_t e_d^q.$$

We note that only the q^{th} row of U is modified. For $1 \leq i \leq d$ we have $d(U_t)_{i,q} = \epsilon (dW_t)_{i,q}$ and $d(Y_t)_m = \rho_q \sum_{j=1}^d (U_t)_{j,m} (dW_t)_{j,q}$. This yields to

$$\begin{aligned}(Y_t)_m &= (Y_0)_m + \rho_q \sum_{j=1}^d (U_0)_{j,m} (W_t)_{j,q} \text{ for } m \neq q, \\ (Y_t)_q &= (Y_0)_q + \rho_q \sum_{j=1}^d (U_0)_{j,q} (W_t)_{j,q} + \frac{\epsilon \rho_q}{2} \sum_{j=1}^d \{(W_t)_{j,q}^2 - t\}.\end{aligned}$$

By using these formulas, we can then sample exactly (U_t, Y_t) and then get a second order scheme for $\hat{\mathcal{L}}$. We note that the simulation cost of $\tilde{\mathcal{L}}_q$ requires $O(d)$ operations and then the one of $\tilde{\mathcal{L}}$ requires $O(d^2)$ operations. Since a matrix multiplication requires $O(d^3)$ operations, this second order scheme for $\tilde{\mathcal{L}}$ and then for \mathcal{L} requires $O(d^3)$ operations instead of $O(d^4)$ for the scheme described in Subsection 4.1.

Remark 13 — *As already mentioned, the dependence between the processes X and Y is the same as the one proposed by Da Fonseca, Grasselli and Tebaldi [14] for a model on asset returns. Therefore, we can use the same splittings as the one proposed in Subsections 4.1 and 4.2 to construct second order schemes for their model.*

4.3 Numerical results

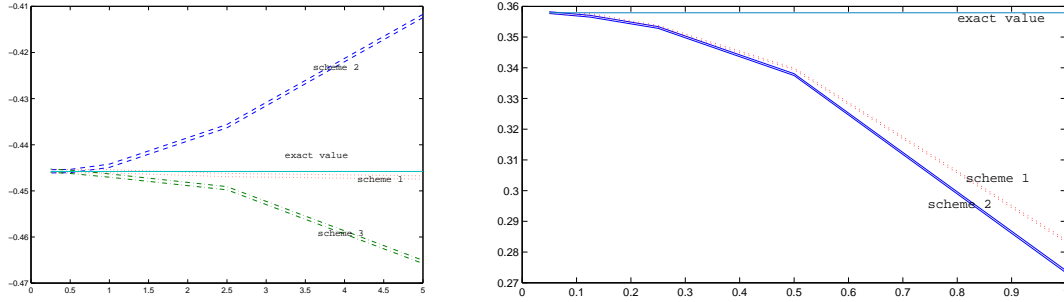


Figure 5: Weak error convergence. Parameters: $p = d = 3$, 10^7 Monte Carlo samples, $T = 5$. The real value of $\mathbb{E} [\exp (-i (\operatorname{Tr}(\Gamma X_T) + \Lambda^\top Y_T))]$, as a function of the time step T/N . Left: $\Gamma = 0.05I_d, \Lambda = 0.021_d$ and the diffusion parameters $x = 0.4I_d, y = 0.21_d, \Omega = 2.5I_d, n = d, \rho = 0, b = 0, \kappa = 0, c = I_d$. The value obtained by solving the ODE: -0.445787 . Right: $\Gamma = 0.2I_d + 0.04q, \Lambda = 0.21_d$ and the diffusion parameters $x = 0.4I_d + 0.2q, y = 0.21_d, \Omega = 0.5I_d, n = d, \rho = -0.31_d, b = -0.5I_d, \kappa = 0.1I_p, c = I_d$, where $q_{i,j} = \mathbb{1}_{i \neq j}$. The value obtained by solving the ODE: 0.357901 . For each scheme, the two curves represent the upper and lower bound of the 95% confidence interval.

We now turn to the empirical analysis of the convergence of the discretization schemes we have proposed. We will use the following notations.

- Scheme 1 is the second order scheme given in Subsection 4.1, where we use the exact sample of the Wishart part and the exact simulation the Gaussian variables.
- Scheme 2 is the second order scheme given in Subsection 4.1, where we use the second order scheme for the Wishart part and replacing the simulation of Gaussian variables by random variables that matches the five first moments, see Theorem 16 and equation (36) in [1].
- Scheme 3 is the second order scheme given in Subsection 4.2.

In order to assess that the potential second order schemes we have proposed for \mathcal{L} give indeed a weak error of order 2, we start by analyzing the weak error for quantities that we can compute analytically. Namely, we consider $\mathbb{E} [\exp (-i (\operatorname{Tr}(\Gamma X_T) + \Lambda^\top Y_T))]$, which can be calculated by solving a system of differential equations similar to (15). We then compare the values obtained by Monte Carlo simulation and the value obtained by solving the system of differential equation. As shown by Figures 5, we observe a weak error which is compatible with the rate of $O(1/N^2)$. When it is well defined, Scheme 3 has to be preferred since it is much faster than the others.

5 Comparison of the different numerical methods

The goal of this section is to compare the computational time needed to price vanilla instruments in the model by using the different numerical methods. We consider the case of a $6M \times 1Y$ caplet with strike 1% , which means

$T = 1$, $\delta = 1/2$ and its price is given by

$$\frac{1}{\delta} \mathbb{E}[e^{-\int_0^T r_s ds} (1 - (1 + K\delta)P_{T,T+\delta})^+] = \frac{P_{0,T}}{\delta} \mathbb{E}^T[(1 - (1 + K\delta)P_{T,T+\delta})^+].$$

We will compare the expansion and the Monte-Carlo method with respect to the Fourier inversion method presented by Carr and Madan [9] and Lee [26]. Their approach can be directly applied for Caplets by working with the forward Caplet price. Let us note that this method can be adapted for swaptions by making the same approximation as the one that we use for the expansion, see Schrager and Pelsser [31] and Singleton and Umantsev [32]. We consider here the four following numerical methods.

- The Monte-Carlo method that consists in using the second order scheme for (X, Y) with a time step of $1/8$ and 10000 paths in order to approximate $\frac{1}{\delta} \mathbb{E}[e^{-\int_0^T r_s ds} (1 - (1 + K\delta)P_{T,T+\delta})^+]$.
- The expansion up to order 2. The integrals that define the coefficients c_i , d_i and e_i are approximated by using a trapezoidal rule and a time step of $1/20$.
- The Fourier transform under \mathbb{P}^T . Starting from the expectation under the T -forward measure, we use the construction of Carr and Madan [9]. In equation (5) of [9], we use $\alpha = 1.25$, truncate the integral at 375 and use a Simpson's rule with a discretization step of $1/8$. Since we calculate here only one price, we do not use the FFT which would have generated further constraints between the discretization and strike grids.
- The Fourier transform under $\mathbb{P}^{T+\delta}$. This is the same method starting with formula (25), and we use the same parameters to approximate the integral and for α .

Pricing Method	Price (bp)	Cpu time (s)
MC price	51.75 ± 1.46 (95% CI)	43.3
Expansion	52.33	0.686
Fourier under P^T	53.84	31.6
Fourier under $P^{T+\delta}$	52.87	33.6

Table 1: Price of the 6M×1Y caplet with strike 1% using different methods with parameter set (46) and $\rho = (-0.4, -0.2)$. Computations are made on a personal laptop with 4Go RAM and a 2.13GHz CPU.

The striking fact is that the method based on the Fourier transform is not so efficient in this context, even though the Fourier inversion is in dimension one. The reason is that the evaluation of the Fourier transform requires to solve numerically matrix Riccati differential equations, for which we take a time step of $1/8$. Figure 6 indicates on our case that a minimum of 2000 evaluations is necessary to have a precision similar to the Monte-Carlo method. Thus, a basic application of the method of Carr and Madan is not very efficient: the bottleneck is to find a smarter way to calculate the characteristic function. In comparison, the Monte Carlo method is not much more time consuming and allows to calculate the price for all strikes and maturities at the same time. Last, we observe that the expansion method is much faster than the others, but is limited to short maturities as indicated in Subsection 3.3. It can therefore be a tool to calibrate quickly the model to some key features such as the at the money price and skew.

Conclusion

The contribution of our paper is twofold. First, the purpose of this paper is to define a Wishart driven affine term structure model for interest rates model, in which the parameters and state variables of the model admit a clear interpretation in terms of the yield curve dynamics, and to provide an efficient numerical framework to implement the model. Other affine term structure models involving Wishart processes have been proposed for example by Bensusan [6] or Gnoatto [25]. A pitfall of general affine term structure model is to offer an abundant parametrization with few intuitions for the practitioner. Here, we believe that presenting the model as a perturbation of the standard LGM model is a good way to get a grip on it, to have a better understanding of the parameters and to have a starting point for the calibration procedure. Let us mention here that getting a reliable and stable calibration procedure of the model is beyond the scope of this paper. In particular, the choice of the

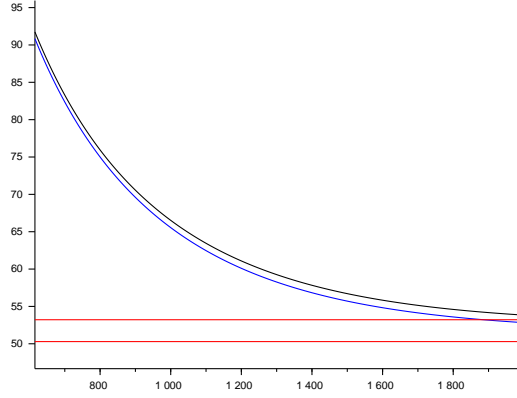


Figure 6: Convergence of the Fourier transform price of the $6M \times 1Y$ caplet with strike 1% and a time step of $1/8$, in function of the number of discretization steps n_{st} . The integration is thus made on $[0, n_{st}/8]$. The parallel lines indicate the 95% confidence interval obtained by MC.

dimensions p and d should be discussed on real data. Also, we have made the choice in this paper to present the model with constant (as opposite to time dependent) parameters: only the factors are meant to describe the state of the interest rate market. Thus, this version of the model has a priori a limited flexibility to calibrate to the swaption volatility cube compared to fully non-homogeneous term structure model with time dependent parameters such as the stochastic volatility forward Libor model of Piterbarg [28], and the stochastic volatility Cheyette model considered by Andreasen in [4]. A full discussion on the calibration of our model as well as the comparison to other models is left for further research.

The other contribution of the paper is to investigate different numerical methods for the model. We know that having efficient numerical methods is a prerequisite to use a model. Besides, our results can be interesting for other models based on Wishart dynamics. As the state variables dynamics is affine, their Fourier and Laplace transforms are tractable and can be obtained by solving Ordinary Differential Equations. Therefore Fourier transform pricing methods can be applied to price vanilla interest rates options in the model. However, the results of our numerical investigation suggest that standard Fourier based pricing methods suffer from numerical efficiency. This is due to the rather lengthy evaluation of the characteristic function together with a slow convergence rate of the Fourier transform discretization. A smarter way to evaluate the characteristic function and to solve the corresponding differential equation has to be investigated to make this method more attractive. As an alternative, we have developed a pricing method for vanilla interest rates options based on a perturbation of the infinitesimal generator of the state variables. This method provides a fast pricing tool for the products which would typically be used for model calibration. The method is particularly efficient for short expiries, but proves limitations for long dated options. Also, the expansion provides analytical expressions for the implied volatility of caplets and swaptions. This is important to confirm the intuitions on the role of the parameters and it can be used to initialize the calibration routine. Last, we propose a second order discretization scheme for the model, which is useful to run a Monte Carlo method. This scheme is easy to implement and very efficient in practice. Besides, it can be adapted easily to a wider range of financial models that use the same dependence structure between the vector Y and its instantaneous Wishart covariance matrix X , such as the Wishart affine stochastic correlation model developed by Da Fonseca et al. [13, 14]. Moreover, it is up to our knowledge the first second order discretization scheme that is able to handle this instantaneous covariance structure.

A Explicit formulas of the price expansion

A.1 Caplets price expansion

We first write the expansion up to order 1 for D , and we get from (15) that $D(t) = D_0(t) + \epsilon D_1(t) + O(\epsilon^2)$ with $\dot{D}_0 = D_0 b + b D_0 + \frac{1}{2} c^\top B B^\top c - \gamma$, $D_0(0) = 0$ and $\dot{D}_1 = D_1 b + b D_1 + \frac{1}{2} D_0 I_d^n \rho B^\top c + \frac{1}{2} c^\top B \rho^\top I_d^n D_0$,

$D_1(0) = 0$. We then obtain

$$D_0(t) = e^{b^\top t} \left(\int_0^t e^{-b^\top s} \left(\frac{1}{2} c^\top B(s) B(s)^\top c - \gamma \right) e^{-bs} ds \right) e^{bt} \quad (52)$$

$$D_1(t) = \frac{1}{2} e^{b^\top t} \left(\int_0^t e^{-b^\top s} (c^\top B(s) \rho^\top I_d^n D_0(s) + D_0(s) I_d^n \rho B(s)^\top c) e^{-bs} ds \right) e^{bt}. \quad (53)$$

We recall that $X_s^0(x)$ is defined by (29). The coefficients of formulas (30) and (31) are given by

$$c_1(t, T, \delta, x) = \int_t^T \Delta B^\top(s, T, \delta) c X_{s-t}^0(x) \partial_x v(s, T, \delta) I_d^n \rho ds,$$

$$c_2(t, T, \delta, x) = \int_t^T \Delta B^\top(s, T, \delta) c X_{s-t}^0(x) \Delta D_0(s, T, \delta) I_d^n \rho + B^\top(T + \delta - s) c X_{s-t}^0(x) \partial_x v(s, T, \delta) I_d^n \rho ds,$$

$$e_1(t, T, \delta, x) = \int_t^T c_1(s, T, \delta, x) \Delta B^\top(s, T, \delta) c X_{s-t}^0(x) \partial_x v(s, T, \delta) I_d^n \rho ds,$$

$$e_2(t, T, \delta, x) = \int_t^T c_1(s, T, \delta, x) [(\Delta B^\top(s, T, \delta) c X_{s-t}^0(x) \Delta D_0(s, T, \delta) I_d^n \rho) + B^\top(T + \delta - s) c X_{s-t}^0(x) \partial_x v(s, T, \delta) I_d^n \rho] \\ + c_2(s, T, \delta, x) \Delta B^\top(s, T, \delta) c X_{s-t}^0(x) \partial_x v(s, T, \delta) I_d^n \rho ds,$$

$$e_3(t, T, \delta, x) = \int_t^T c_2(s, T, \delta, x) [(\Delta B^\top(s, T, \delta) c X_{s-t}^0(x) \Delta D_0(s, T, \delta) I_d^n \rho) + B^\top(T + \delta - s) c X_{s-t}^0(x) \partial_x v(s, T, \delta) I_d^n \rho] ds,$$

$$e_4(t, T, \delta, x) = \int_t^T 2 \Delta B^\top(s, T, \delta) c X_{s-t}^0(x) \partial_x c_1(s, T, \delta) I_d^n \rho ds,$$

$$e_5(t, T, \delta, x) = \int_t^T 2 B^\top(T + \delta - s) c X_{s-t}^0(x) \partial_x c_1(s, T, \delta) I_d^n \rho + 2 \Delta B^\top(s, T, \delta) c X_{s-t}^0(x) \partial_x c_2(s, T, \delta) I_d^n \rho ds,$$

$$e_6(t, T, \delta, x) = \int_t^T 2 B^\top(T + \delta - s) c X_{s-t}^0(x) \partial_x c_2(s, T, \delta) I_d^n \rho ds.$$

and

$$d_1(t, T, \delta, x) = \int_t^T \frac{1}{2} \text{Tr} [I_d^n \partial_x v(s, T, \delta) X_{s-t}^0(x) \partial_x v(s, T, \delta)] ds$$

$$d_2(t, T, \delta, x) = \int_t^T 2 \text{Tr} [\Delta D_0(s, T, \delta) X_{s-t}^0(x) \partial_x v(s, T, \delta) I_d^n] ds$$

$$d_3(t, T, \delta, x) = \int_t^T (2 \text{Tr} (\Delta D_0(s, T, \delta) I_d^n \Delta D_0(s, T, \delta) X_{s-t}^0(x)) + (\Delta B^\top(s, T, \delta) c X_{s-t}^0(x) \Delta D_1(s, T, \delta) I_d^n \rho)) \\ + \frac{1}{2} \text{Tr} [(d-1) I_d^n + 4 X_{s-t}^0(x) D_0(T + \delta - s) I_d^n] \partial_x v(s, T, \delta) ds.$$

A.2 Swaption price expansion

We have $D^S = D_0^S + \epsilon D_1^S + o(\epsilon)$, with

$$D_i^S(t) = \omega_0^0 D_i(T-t) - \omega_0^m D_i(T+m\delta-t) - S_0(T, m, \delta) \sum_{k=1}^m \omega_0^k D_i(T+k\delta-t), \quad i = 0, 1,$$

where the functions D_0 and D_1 are given by (52) and (53). The coefficients of formulas (44) and (46) are given by

$$c_1^S(t, T, x) = \int_t^T B^S(u)^\top c X_{u-t}^0(x) \partial_x v^S(u, T) I_d^n \rho du,$$

$$c_2^S(t, T, x) = \int_t^T B^S(u)^\top c X_{u-t}^0(x) D_0^S(u) I_d^n \rho + B^A(u)^\top c X_{u-t}^0(x) \partial_x v^S(u, T) I_d^n \rho du,$$

$$\begin{aligned}
d_1^S(t, T, x) &= \int_t^T \frac{1}{2} \text{Tr}(I_d^n \partial_x v^S(u, T) X_{u-t}^0(x) \partial_x v^S(u, T)) du, \\
d_2^S(t, T, x) &= \int_t^T 2 \text{Tr}(D_0^S(u) X_{u-t}^0(x) \partial_x v^S(u, T) I_s^n) du, \\
d_3^S(t, T, x) &= \int_t^T [2 \text{Tr}(D_0^S(u) I_d^n D_0^S(u) X_{u-t}^0(x)) + (B^S(u))^\top c X_{u-t}^0(x) D_1^S(u) I_d^n \rho] \\
&\quad + \text{Tr} \left([2 X_{u-t}^0(x) D_0^A(u) I_d^n + \frac{1}{2} (d-1) I_d^n] \partial_x v^S(u, T) \right) du, \\
e_1^S(t, T, x) &= \int_t^T c_1^S(u, T, x) (B^S(u))^\top c X_{u-t}^0(x) \partial_x v^S(u, T) I_d^n \rho du, \\
e_2^S(t, T, x) &= \int_t^T c_1^S(u, T, x) [B^S(u)^\top c X_{u-t}^0(x) D_0^S(u) I_d^n \rho + B^A(u)^\top c X_{u-t}^0(x) \partial_x v^S(u, T) I_d^n \rho] \\
&\quad + c_2^S(u, T, x) B^S(u)^\top c X_{u-t}^0(x) \partial_x v^S(u, T) I_d^n \rho du, \\
e_3^S(t, T, x) &= \int_t^T c_2^S(u, T, x) [B^S(u)^\top c X_{u-t}^0(x) D_0^S(u) I_d^n \rho + B^A(u)^\top c X_{u-t}^0(x) \partial_x v^S(u, T) I_d^n \rho] \\
&\quad + 2 B^S(u)^\top c X_{u-t}^0(x) \partial_x c_1^S(u, T) I_d^n \rho du, \\
e_4^S(t, T, x) &= \int_t^T 2 B^A(u)^\top c X_{u-t}^0(x) \partial_x c_1^S(u, T) I_d^n \rho + 2 B^S(u)^\top c X_{u-t}^0(x) \partial_x c_2^S(u, T) I_d^n \rho du, \\
e_5^S(t, T, x) &= \int_t^T 2 B^A(u)^\top c X_{u-t}^0(x) \partial_x c_2^S(u, T) I_d^n \rho du.
\end{aligned}$$

B Proof of Proposition 4

We first recall the following useful result

$$\forall x, y \in \mathcal{S}_d^+(\mathbb{R}), \text{Tr}(xy) \geq 0, \quad (54)$$

which comes easily from $\text{Tr}(xy) = \text{Tr}(\sqrt{xy}\sqrt{x})$ and $\sqrt{xy}\sqrt{x} \in \mathcal{S}_d^+(\mathbb{R})$.

For $x, y \in \mathcal{S}_d(\mathbb{R})$, we use the notation $x \leq y$ if $y - x \in \mathcal{S}_d^+(\mathbb{R})$. By assumption, there is $\mu > 0$ such that $2\mu I_d \leq -(b + b^\top)$. We now apply Proposition 2 with $\bar{\Lambda} = 0$ and $\bar{\Gamma} = 0$. Since $\|\lambda(t)\| \leq \|\Lambda\|$, there is a constant $h > 0$ small enough such that for any $\Lambda \in \mathbb{R}^d$ satisfying $\|\Lambda\| < h$ we have

$$\forall t \geq 0, \mu I_d \leq -[b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c + (b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c)^\top] \text{ and } \frac{1}{2} c^\top \lambda \lambda^\top c \leq \frac{\mu^2}{8\epsilon^2} I_d$$

By choosing $\Upsilon = \frac{\mu}{4\epsilon^2} I_d$, we see that the condition (14) is satisfied since $\frac{\mu^2}{4\epsilon^2} I_d - \frac{\mu^2}{8\epsilon^2} I_d - \frac{1}{2} c^\top \lambda \lambda^\top c \in \mathcal{S}_d^+(\mathbb{R})$ for all $t \geq 0$. Thus, the conclusions of Proposition 2 hold for any $\Lambda \in \mathbb{R}^d$ and $\Gamma \in \mathcal{S}_d(\mathbb{R})$ such that $\|\Lambda\| < h$ and $\Gamma \leq \frac{\mu}{4\epsilon^2} I_d$, and we have $g(t) \leq \frac{\mu}{4\epsilon^2} I_d$ for any $t \geq 0$. We now want to prove that $\lambda(t) \xrightarrow{t \rightarrow +\infty} 0$, $g(t) \xrightarrow{t \rightarrow +\infty} 0$ and $\eta(t)$ converges when $t \rightarrow +\infty$. This will prove the convergence to the stationary law by Lévy's theorem.

From (15), we have

$$\frac{1}{2} \frac{d}{dt} \text{Tr}(g^2) = 2\epsilon^2 \text{Tr}(g I_d^n g^2) + \text{Tr}(g^2 [b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c + (b + \frac{\epsilon}{2} I_d^n \rho \lambda^\top c)^\top]) + \text{Tr}(g \frac{1}{2} c^\top \lambda \lambda^\top c).$$

By (54), we get

$$\frac{1}{2} \frac{d}{dt} \text{Tr}(g^2) \leq \frac{\mu}{2} \text{Tr}(g I_d^n g) - \mu \text{Tr}(g^2) + \frac{\mu}{4\epsilon^2} \text{Tr}(\frac{1}{2} c^\top \lambda \lambda^\top c).$$

Since $\text{Tr}(g I_d^n g) \leq \text{Tr}(g^2)$, we get by Gronwall's lemma

$$\frac{1}{2} \text{Tr}(g(t)^2) \leq \frac{1}{2} \text{Tr}(\Gamma^2) e^{-\mu t} + \frac{\mu}{4\epsilon^2} \int_0^t \text{Tr} \left(\left[\frac{1}{2} c^\top \lambda(s) \lambda^\top(s) c \right]^2 \right) e^{-\mu(t-s)} ds.$$

We now use that the entries of λ decay exponentially. Since $\int_0^t e^{-\mu' s} e^{-\mu(t-s)} ds \underset{t \rightarrow +\infty}{=} O(e^{-\frac{\min(\mu, \mu')}{2} t})$ for $\mu, \mu' > 0$, we get that there exists $C, \nu > 0$ such that $\frac{1}{2} \text{Tr}(g(t)^2) \leq C e^{-\nu t}$. This gives that $g(t) \underset{t \rightarrow +\infty}{\rightarrow} 0$ and that $\eta(t) = \int_0^t \lambda^\top(s) \kappa \theta + \text{Tr}(g(s)(\Omega + \epsilon^2(d-1)I_d^n)) ds$ converges. \square

References

- [1] A. Ahdida and A. Alfonsi. Exact and high-order discretization schemes for Wishart processes and their affine extensions. *Ann. Appl. Probab.*, 23(3):1025–1073, 2013.
- [2] A. Alfonsi. High order discretization schemes for the CIR process: application to affine term structure and Heston models. *Math. Comp.*, 79(269):209–237, 2010.
- [3] L. Andersen and V. Piterbarg. *Interest Rate Modeling*, volume Issue 2. Atlantic Financial Press, 2010.
- [4] J. Andreasen. Back to the future. *Risk Magazine*, 2005.
- [5] A. Benabid, H. Bensusan, and N. El Karoui. Wishart Stochastic Volatility: Asymptotic Smile and Numerical Framework. <https://hal.archives-ouvertes.fr/hal-00458014>, 2008.
- [6] H. Bensussan. Interest rate and longevity risks : dynamic modelling and applications to derivative products and life insurance. *PhD Thesis, Ecole Polytechnique*, 2010. <https://tel.archives-ouvertes.fr/pastel-00563792/>.
- [7] L. Bergomi and J. Guyon. Stochastic volatilities orderly smiles. *Risk Magazine*, 2012.
- [8] D. Brigo and F. Mercurio. *Interest rate models—theory and practice*. Springer Finance. Springer-Verlag, Berlin, second edition, 2006. With smile, inflation and credit.
- [9] P. Carr and D. Madan. Option valuation using fast fourier transform. *Journal of Computational Finance*, 1999.
- [10] P. Collin-Dufresne and R. Goldstein. Pricing swaptions within an affine framework. *The Journal of Derivatives*, 2002.
- [11] J. C. Cox, J. E. Ingersoll, Jr., and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53(2):385–407, 1985.
- [12] C. Cuchiero, D. Filipović, E. Mayerhofer, and J. Teichmann. Affine processes on positive semidefinite matrices. *Ann. Appl. Probab.*, 21(2):397–463, 2011.
- [13] J. Da Fonseca, M. Grasselli, and F. Ielpo. Estimating the Wishart affine stochastic correlation model using the empirical characteristic function. *Stud. Nonlinear Dyn. Econom.*, 18(3):253–289, 2014.
- [14] J. Da Fonseca, M. Grasselli, and C. Tebaldi. Option pricing when correlations are stochastic: an analytical framework. *Review of Derivatives Research*, 2008.
- [15] Q. Dai and K. J. Singleton. Specification analysis of affine term structure models. *The Journal of Finance*, 55(5):1943–1978, 2000.
- [16] A. D’Aspremont. Interest rate model calibration using semidefinite programming. *Applied Mathematical Finance*, 10(3):183–213, 2003.
- [17] L. Dieci and T. Eirola. Positive definiteness in the numerical solution of Riccati differential equations. *Numer. Math.*, 67(3):303–313, 1994.
- [18] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Ann. Appl. Probab.*, 13(3):984–1053, 2003.
- [19] D. Duffie and R. Kan. A yield-factor model of interest rates. *Mathematical Finance*, 6:379–406, 1996.
- [20] N. El Karoui and V. Lacoste. Multifactor models of the term structure of interest rates. 1992.

- [21] N. El Karoui, C. Lepage, R. Myneni, N. Roseau, and R. Wiswanathan. The valuation and hedging of contingent claims with markovian interest rates. 1991.
- [22] D. Filipović. *Term-structure models*. Springer Finance. Springer-Verlag, Berlin, 2009. A graduate course.
- [23] J.-P. Fouque, G. Papanicolaou, and K. R. Sircar. *Derivatives in financial markets with stochastic volatility*. Cambridge University Press, Cambridge, 2000.
- [24] H. Geman, N. El Karoui, and J.-C. Rochet. Changes of numéraire, changes of probability measure and option pricing. *J. Appl. Probab.*, 32(2):443–458, 1995.
- [25] A. Gnoatto. The Wishart short rate model. *Int. J. Theor. Appl. Finance*, 15(8):1250056, 24, 2012.
- [26] R. W. Lee. Option pricing by transform methods: Extensions, unification, and error control. *Journal of Computational Finance*, 2004.
- [27] E. Palidda. Managing interest rates derivatives with stochastic variance-covariance. *PhD Thesis, Ecole des Ponts ParisTech*, 2015. <https://hal.archives-ouvertes.fr/tel-01217655v1>.
- [28] V. Piterbarg. A Stochastic Volatility Forward Libor Model with a Term Structure of Volatility Smiles. *Social Science Research Network Working Paper Series*, Nov. 2003.
- [29] V. Piterbarg. Rates squared. *Risk Magazine*, 2009.
- [30] T. H. Rydberg. A note on the existence of unique equivalent martingale measures in a markovian setting. *Finance and Stochastics*, 1(3):251–257, 1997.
- [31] D. F. Schrager and A. A. J. Pelsser. Pricing swaptions and coupon bond options in affine term structure models. *Math. Finance*, 16(4):673–694, 2006.
- [32] K. J. Singleton and L. Umantsev. Pricing coupon-bond options and swaptions in affine term structure models. *Math. Finance*, 12(4):427–446, 2002.
- [33] K. Tanaka, T. Yamada, and T. Watanabe. Applications of Gram-Charlier expansion and bond moments for pricing of interest rates and credit risk. *Quant. Finance*, 10(6):645–662, 2010.
- [34] O. Vasicek. An equilibrium characterization of the term structure. *Journal of Financial Economics*, 5(2):177 – 188, 1977.