LÉVY INTEREST RATE MODELS WITH A LONG MEMORY

Donatien Hainaut

REPRINT | 2022 / 01

UCLouvain

LIDAM
ISBA | Louvain Institute of Data Analysis and Modeling in economics and statistics

ISBA
Voie du Roman Pays 20 - L1.04.01
B-1348 Louvain-la-Neuve
Email : lidam-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/isba/publication.html

Article

Lévy interest rate models with a long memory

Donatien Hainaut [1]*

[1] UCLouvain, LIDAM; donatien.hainaut(at)uclouvain.be

† Current address: 20 voie du Roman Pays, 1348, Louvain-La-Neuve, Belgium

Version December 17, 2021 submitted to Risks

Abstract: This article proposes an interest rate model ruled by mean reverting Lévy processes with a sub-exponential memory of their sample path. This feature is achieved by considering an Ornstein-Uhlenbeck process in which the exponential decaying kernel is replaced by a Mittag-Leffler function. Based on a representation in term of an infinite dimensional Markov processes, we present the main characteristics of bonds and short-term rates in this setting. Their dynamics under risk neutral and forward measures are studied. Finally, bond options are valued with a discretization scheme and a discrete Fourier's transform.

Keywords: Interest rate, Lévy process, Mittag-Leffler function, Mean reverting process

1. Introduction

From the 1980s to nowadays, many interest rate models were proposed in the literature. Their common aim is to explain changes in bond or swap quotes and to replicate risks within the interest rates market. Three dominating frameworks coexist: short-term rate, forward rate and the Libor market models. In this last approach proposed by Brace et al. (1997) interest rates are driven by geometric processes. In the forward rate model pioneered by Heath et al. (1992), the term structure of rates is specified through instantaneous forward rates. Mercurio and Moraleda (2000) and Falini (2010) propose a forward model with a humped structure of volatilities. Li et al. 2020) develop a forward rate model unifying most of existing Gaussian models for interest rates. Short-term rate models such as promoted by Hull and White (1990), specify a mean reverting dynamic for the instantaneous risk-free rate. The framework developed in this article belongs to this third category.

The family of short-term rate models gathers multiple frameworks and we refer the reader e.g. to Boero and Torricelli (1996) or to Schmidt (2011) for a review. On the other hand, various processes were proposed to explain the interest rate risk. The literature is too vast for being exhaustive. We restrict our attention to contributions related to this work and do not attempt a general overview, referring instead to the Brigo and Mercurio (2006) for detailed accounts on the topic. For instance, Eberlein and Raible (1999) were among the firsts to propose a short-term rate model driven by Lévy processes. In their setting, the short-term rate is ruled by Ornstein-Uhlenbeck processes reverting in an exponential manner to a mean level. Within this framework, Eberlein and Kluge (2005) derive analytical formulae for the prices of caps and floors using bilateral Laplace transforms. In a similar setting, Hainaut and Macgilchrist (2010) propose a pentanomial tree for pricing derivatives. Hainaut (2013) studies the properties of a Gaussian short-term rate with a Markov Switching Multifractal volatility. Moreno and Platania (2015) propose a square-root model replicating economic cyles in the dynamic of interest rates. Hainaut (2016) and Hainaut and Njike Leunga (2020) develop unicurve and multicurve models in which the short-term rate is exposed to self-excitating jumps. Fontana et al. (2020) study a modelling framework exploiting the the self-exciting behavior of continuous-state branching processes with immigration (CBI).

A wide majority of short-term rate models are driven by Markov mean reverting processes. In this setting, the interest rate depends on its path through the integral of past occurrences weighted by an exponential decaying function, called the memory kernel. In this framework, the influence of previous variations of interest rates decreases then exponentially over time. In this article, we consider an alternative that consists to replace this memory kernel by a Mittag-Leffler decaying function. This function is sub-exponential and may be seen as a generalization of the exponential. In this setting, the interest rate remembers its previous occurrences for a longer period than in the exponential framework. The motivations for studying such a model are multiple. Firstly, his features replicate the persistence empirically observed in short term rates, as illustrated in Section 3 or in nominal yields and inflation, as shown by Golinski and Zaffaroni (2016). Secondly, the proposed model has a sufficient level of analytical tractability for most of financial applications such as option pricing. Finally, it also offers a novel alternative to existing approaches based on the fractional Brownian motion for introducing long memory, as for instance studied in Cheridito et al. (2003) or Maejima and Yamamoto (2003).

A consequence of this change of memory kernel is that the interest rate process is not Markov anymore. Bond prices depend then on the sample path of past rates. Nevertheless, we take advantage of the representation of the memory kernel as a Laplace Stieltjes integral to rewrite the short-term rate as an infinite dimensional Markov process. The last part of the article studies the properties of the model under a forward measure and proposes a pricing method for bond options.

2. A Lévy model with a Mittag Leffler kernel

We consider that the short-term rate is driven by $d \in \mathbb{N}$ independent Lévy processes, denoted by $(L_t^{(j)})_{t\geq 0}$ for $j = 1, ..., d$. These processes are defined on a probability space Ω, endowed with the natural filtration $(\mathcal{F}_t)_{t\geq 0}$ and a risk neutral measure, $\mathbb{Q}$. Each process $(L_t^{(j)})_{t\geq 0}$ has independent and stationary increments. It is fully described by a triplet $\left(\mu_j, \sigma_j^2, \nu_j\right)$ where $\mu_j \in \mathbb{R}$, $\sigma_j \in \mathbb{R}^+$ and $\nu_j(.)$ is a Lévy measure. The moment generating function (mgf) is denoted by $\phi_t^{(j)}(\omega)$ for $j = 1, ..., d$ and is equal to

$$\begin{aligned} \phi_t^{(j)}(\omega) &:= \mathbb{E}\left(\exp\left(\omega(L_t^{(j)} - L_0^{(j)})\right) \mid \mathcal{F}_0\right) \\ &= \exp\left(t\left(\mu_j\omega + \frac{1}{2}\omega^2\sigma_j^2 + \int_{\mathbb{R}}\left(e^{\omega z} - 1 - \omega z 1_{|z|<\epsilon}\right)\nu_j(dz)\right)\right) \\ &= \exp\left(t\psi_j(\omega)\right) , \end{aligned} \tag{1}$$

where $\epsilon > 0$. Notice that if the Lévy measure, $\nu_j(.)$, has no singularity at $z = 0$, we can set ϵ to zero. The function $\psi_j(\omega)$ is the characteristic exponent of this Lévy process that is always defined for pure imaginary numbers. Without loss of generality, we assume that $L_0^{(j)} = 0$. According to the Lévy Itô decomposition, each $L_t^{(j)}$ is the sum of three components : a deterministic drift $\mu_j\, t$, a Brownian motion with variance σ_j^2 and a jump process, $J_j(t\,, z)$, of intensity $\nu_j(dz)$ well defined on $[-\infty, 0) \cup (0, +\infty]$. This Lévy measure is such that the probability of observing k jumps between $[\tau_1, \tau_2]$ of a size included in a set $B \subset \mathbb{R}_{\setminus\{0\}}$ is given by

$$P\left(J_j([\tau_1, \tau_2] \times B) = k\right) = e^{-\int_{\tau_1}^{\tau_2}\int_B \nu_j(dz)dt}\frac{\left(\int_{\tau_1}^{\tau_2}\int_B \nu_j(dz)dt\right)^k}{k!}, \tag{2}$$

for $j = 1, ..., d$. If $\left(W_t^{(j)}\right)_{t\geq 0}$ for $j = 1, ..., d$, are independent Brownian motions on $(\Omega, \mathcal{F}, \mathbb{Q})$, $L_t^{(j)}$ may be split as the sum of a drift, a Brownian motion and a jump process

$$dL_t^{(j)} = \mu_j dt + \sigma_j dW_t^{(j)} + \int_{\mathbb{R}} z\,\tilde{J}_j(dt\,,\,dz), \tag{3}$$

where

$$\tilde{J}_j(dt\,,\,dz) = J_j(dt\,,\,dz) - \mathbf{1}_{|z|<\epsilon}\nu_j(dz)dt\,.$$

Without loss of generality, we assume that $\mathbb{E}\left(L_t^{(j)}\right) = 0$. This constraint implies that the drift compensates jumps larger than ϵ in absolute value:

$$\mu_j \quad = \quad -\int_{|z|\geq\epsilon} z\,\nu_j(dz) \quad , j = 1, ..., d.$$

From Cont and Tankov (2003, Lemma 15.1, pp 482), for any integrable function $f : \mathbb{R} \to \mathbb{C}$, the following relation holds

$$\mathbb{E}\left(e^{\omega\int_s^t f(u)dL_u^{(j)}}|\mathcal{F}_s\right) = \exp\left(\int_s^t \psi_j\left(\omega f(u)\right)\,du\right)\,,\; j = 1, ..., d. \tag{4}$$

This can be proved by approaching $f(.)$ with a stepwise function and by using the property of independence of increments. This property will be useful in later developments. We postulate that the risk-free rate, $(r_t)_{t\geq 0}$, is the sum of a deterministic function $\varphi(t) : \mathbb{R} \to \mathbb{R}$, and of d processes $\left(X_t^{(d)}\right)_{t\geq 0}$:

$$r_t = \varphi(t) + \sum_{j=1}^{d} X_t^{(j)}\,, \tag{5}$$

where the $X_t^{(j)}$ are such that:

$$X_t^{(j)} = g_j(t)X_0^{(j)} + \int_0^t g_j(t-u)\,dL_u^{(j)}\,. \tag{6}$$

The function $g_j(.) : \mathbb{R}^+ \to \mathbb{R}$ is continuously decreasing with an initial value $g_j(0) = 1$. As $g_j(.)$ determines the influence of past realizations of $L_t^{(j)}$ on the current value of the process, we call it "memory kernel" or "kernel function" in the remainder of this article. When the function $g_j(t)$ decreases exponentially, the processes $\left(X_t^{(j)}\right)_{j=1,...,d}$ are Markov and mean-reverting as recalled in Appendix A. Such a model was e.g. developed in Hainaut and Macgilchrist (2010) for pricing interest rate derivatives and will serve as benchmark in numerical illustrations. However $\left(X_t^{(j)}\right)_{t\geq 0}$ and $(r_t)_{t\geq 0}$ are in general not Markov for any other non-exponential kernels. This feature makes difficult the evaluation of bond prices at a given time $t > 0$ as their value depends on the whole sample path of interest rates up to t. Nevertheless, for a few kernel functions that admit a representation as a Laplace–Stieltjes integral, we will show that the interest rate may be represented as an infinite dimensional Markov process. In particular we consider two types of kernel functions, both based on the Mittag-Leffler function, denoted by $E_\alpha(.)$ where $\alpha \in [0, 1]$:

$$g_j(t) = E_{\alpha_j}(-\beta_j t) \quad \text{or} \quad g_j(t) = E_{\alpha_j}(-\beta_j t^{\alpha_j})\,. \tag{7}$$

To understand the motivation for working with such functions, we need to review the main properties of the Mittag-Leffler function. The Mittag Leffler function of order $\alpha > 0$ plays a fundamental role in the fractional calculus and can be considered as an extension of the exponential function. It is defined as an infinite sum

$$E_\alpha(t) = \sum_{n=0}^{\infty} \frac{t^n}{\Gamma(n\alpha+1)}\,.$$

In this article we assume that $\alpha \in [0,1]$. For this range of values, $E_\alpha(-t)$ is an intermediary between the power and the exponential decreasing functions:

$$E_0\,(-t) = \frac{1}{1+t}\,\forall |t| < 1 \text{ and } E_1(-t) = e^{-t}\,.$$

We refer the reader to the book of Gorenflo et al. (2020) for a detailed presentation of this function. For a complete presentation of fractional calculus, we recommend the book of Baleanu et al. (2012) and the book of Kochubei and Luchko (2019). Let us recall that a function $f : (0,\infty) \to \infty$ is called completely monotonic if it possesses derivatives $f^{(n)}(t)$ of any order $n = 0, 1, 2...$ and the derivatives are alternating in sign, i.e.

$$(-1)^n\, f^{(n)}(t) \geq 0\, \forall t \in ((0,\infty)\,.$$

The above property is equivalent to the existence of a representation of the function f in the form of a Laplace–Stieltjes integral with non-decreasing density and non-negative measure $d\gamma(.)$ such that

$$f(t) = \int_0^\infty e^{-ut} d\gamma(u)\,.$$

The Mittag-Leffler function of negative argument $E_\alpha(-\beta t)$ is completely monotonic for all $0 \leq \alpha \leq 1$ and $\beta \in \mathbb{R}^+$. A proof of this result is available in Gorenflo et al. (2020) on page 47. The authors show that

$$E_\alpha(-\beta t) \quad = \quad \int_0^\infty e^{-tu} d\gamma_{\alpha,\beta}(u)\,, \tag{8}$$

where the derivative of $\gamma_{\alpha,\beta}(u)$ is equal to

$$\frac{d\gamma_{\alpha,\beta}(u)}{du} \quad = \quad \frac{1}{\pi\alpha} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k!} \sin(\pi\alpha k)\, \Gamma(\alpha k+1) \frac{u^{k-1}}{\beta^k}\,. \tag{9}$$

Given that $E_\alpha(0) = 1$, we have that $\int_0^\infty d\gamma_{\alpha,\beta}(u) = 1$ and $\gamma_{\alpha,\beta}(.)$ is then a measure of probability on $\mathbb{R}^+$. In the remainder of this article, $E_\alpha(-\beta t)$ is called the decreasing Mittag-Leffler (ML) kernel. The DML behaves at short-term as

$$E_\alpha(-\beta t) \quad = \quad 1 - \frac{\beta t}{\Gamma(1+\alpha)} + ... \sim \exp\left(-\frac{\beta t}{\Gamma(1+\alpha)}\right)\,,$$

when $t \to 0$. From Haubold et al. (2011), we known that when $t \to \infty$, the DML converges to

$$E_\alpha(-\beta t) \sim \frac{(\beta\, t)^{-1}}{\Gamma\,(1-\alpha)}\,.$$

The function $E_\alpha(-\beta t)$ interpolates for intermediate times t between the decreasing exponential and the inverse power law. The exponential models fast decay for small time t, whereas the asymptotic inverse power law entails a slow decrease at long term. This point is illustrated in the left plot of Figure 1 that compares the ML kernel to the exponential and inverse power functions.

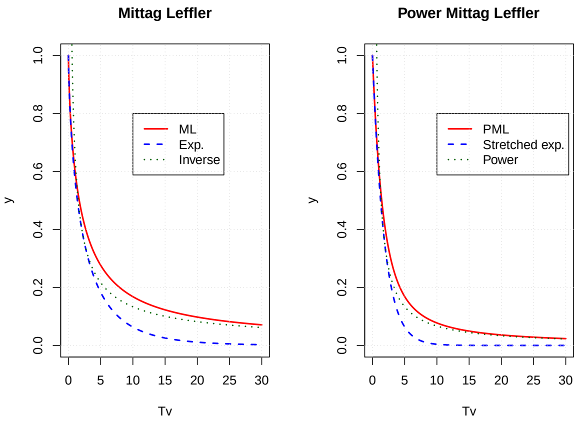

Figure 1. Left and right plots: Comparison of the Mittag Leffler (ML), $E_\alpha(-\beta t)$, and power Mittag Leffler (PML), $E_\alpha(-\beta t^\alpha)$, kernels with exponential and power decreasing functions ($\alpha = 0.7$ and $\beta = 0.5$).

The Mittag-Leffler function is also related to the fractional calculus. To explain this link, we recall that the Caputo's fractional derivative of order $\alpha \in (0,1)$ for a function $h(t) : \mathbb{R}^+ \to \mathbb{R}$, $\mathcal{C}^1$ with respect to t is defined by

$$\frac{\partial^\alpha}{\partial t^\alpha} h(t) \quad = \quad \frac{1}{\Gamma(1-\alpha)} \int_0^t (t-s)^{-\alpha} \frac{\partial}{\partial s} h(s) ds \tag{10}$$

When $\alpha = 1$, this derivative corresponds to the first order derivative. The solution of the fractional differential equation

$$\frac{\partial^\alpha}{\partial t^\alpha} y(t) = -\beta y(t) \qquad 0 < \alpha < 1$$

with the initial condition $y(0) = b_0$ is $y(t) = E_\alpha(-\beta t^\alpha)$. We abusively call this function the power decreasing Mittag-Leffler kernel (PML) . The Laplace transform of $E_\alpha(-\beta t^\alpha)$ is given by $\frac{s^{\alpha-1}}{s^\alpha+\beta}$ where $Re(s) > |\beta|^{\frac{1}{\alpha}}$. By inverting this transform, we can prove as in Gorenflo and Mainardi (1997) that the PML is the Laplace transform

$$E_\alpha(-\beta t^\alpha) = \int_0^\infty e^{-ut} \, d\gamma^p_{\alpha,\beta}(u) \, , \tag{11}$$

where the derivative of $\gamma^p_{\alpha,\beta}(u)$ is given by

$$\frac{d\gamma^p_{\alpha,\beta}(u)}{du} = \frac{1}{\pi} \frac{\beta \, u^{\alpha-1} \sin(\alpha\pi)}{u^{2\alpha} + 2\beta u^\alpha \cos(\alpha\pi) + \beta^2} \, . \tag{12}$$

Since $E_\alpha(0) = 1$, we have $\int_0^\infty d\gamma^p_{\alpha,\beta}(u) = 1$ and $\gamma^p_{\alpha,\beta}(u)$ is a measure of probability on $\mathbb{R}^+$. As underlined by Mainardi (2020), the PML behaves at short-term as

$$E_\alpha(-\beta t^\alpha) \quad = \quad 1 - \frac{\beta t^\alpha}{\Gamma(1+\alpha)} + ... \sim \exp\left(-\frac{\beta t^\alpha}{\Gamma(1+\alpha)}\right) \, .$$

when $t \to 0$. From Erdélyi et al. (1955), we known that when $t \to \infty$, the PML converges to

$$E_\alpha(-\beta t^\alpha) \sim \frac{\left(\beta^{1/\alpha} t\right)^{-\alpha}}{\Gamma\left(1-\alpha\right)} \, .$$

As a consequence the function $E_\alpha(-\beta t^\alpha)$ interpolates for intermediate times t between the stretched exponential and the negative power law. The stretched exponential models the very fast decay at short-term whereas the asymptotic power law is due to the very slow decay for large time t. The right plot of Figure 1 illustrates this convergence.

Figure 2 presents simulated sample paths of the short-term rate with a one dimensional ($d = 1$) model ruled by a Brownian motion. These paths are computed for various α with the same random occurrences in order to make comparison feasible. For the ML kernel, trajectories are nearly similar. An analysis of figures reveals that the sample path is smoother for $\alpha = 0.50$ than for $\alpha = 0.90$. This trend becomes visible if we choose a highest value for β. For PML sample paths, the difference is clearly visible. Decreasing α reduces the volatility of rates and smoothes the sample path.

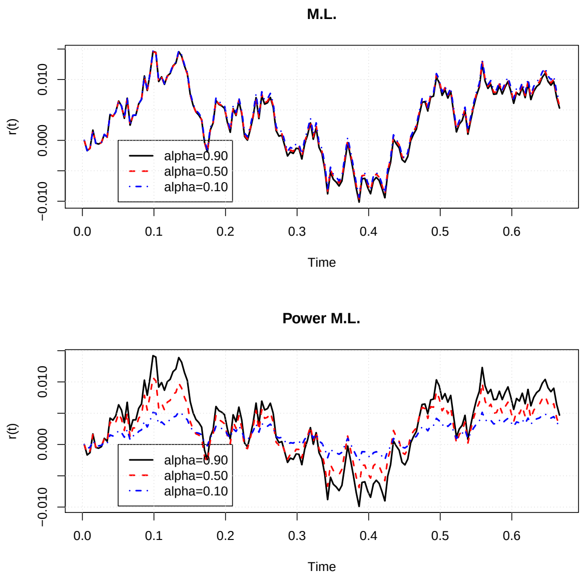

Figure 2. Upper and lower plots: Comparison of sample paths of $(r_t)_{t\geq 0}$ with Mittag Leffler (ML) and power Mittag Leffler (PML) kernels. $d = 1$, $\beta = 3$, $X_0^{(1)} = 0$, $\varphi(t) = 0$ and L_t is a Brownian motion with $\sigma_1 = 0.05$.

To conclude this section, we present the first two moments of the short-term rate and its autocovariance function. The ML and PML kernels defined in Equations (7) are continuous decreasing and integrable functions on any bounded interval of $\mathbb{R}$. From Equation (4), the mgf of $X_t^{(j)}$ for $j = 1, ..., d$ is then equal to:

$$\mathbb{E}\left(e^{\omega X_t^{(j)}} | \mathcal{F}_0\right) = \exp\left(\omega\, g_j(t)\, X_0^{(j)} + \int_0^t \psi_j\left(\omega\, g_j(t-u)\right) du\right) .$$

Deriving the mgf allows us to find the first moments of $X_t^{(j)}$ conditionally to the initial information. On the other hand, from the representation

$$r_t = \varphi(t) + \sum_{j=1}^{d} g_j(t) X_0^{(j)} + \sum_{j=1}^{d} \int_0^s g_j(t-u)\, dL_u^{(j)} + \sum_{j=1}^{d} \int_s^t g_j(t-u)\, dL_u^{(j)} , \tag{13}$$

of r_t, we infer that the conditional expectation of the short-term rate is given by

$$\mathbb{E}\left(r_t\, | \mathcal{F}_s\right) = \varphi(t) + \sum_{j=1}^{d} g_j(t) X_0^{(j)} + \sum_{j=1}^{d} \int_0^s g_j(t-u)\, dL_u^{(j)} , \tag{14}$$

Whereas the conditional variance of r_t is equal to the sum of variances of Lévy processes times the integral of the squared kernels

$$\mathbb{V}\left(r_t|\mathcal{F}_s\right)=\sum_{k=1}^{d}\left(\sigma_j^2+\int_{\mathbb{R}}z^2\nu_j(dz)\right)\int_s^t g_j(t-u)^2du\,. \tag{15}$$

Unfortunately the integral of $g_j(.)^2$ does not admit a closed form expression for the ML and PML kernels. Nevertheless, they can be numerically estimated. A direct calculation allows us to infer that the autocovariance between r_t and r_u for $u\leq t$ is given by

$$\begin{aligned}\mathbb{C}\left(r_t r_u|\mathcal{F}_s\right) &= \sum_{k=1}^{d}\mathbb{E}\left(\int_s^t g_j(t-v)\,dL_v^{(j)}\int_s^u g_j(u-v)\,dL_v^{(j)}\right) \\ &= \sum_{k=1}^{d}\left(\sigma_j^2+\int_{\mathbb{R}}z^2\nu_j(dz)\right)\int_s^u g_j(t-v)g_j(u-v)\,dv\,.\end{aligned} \tag{16}$$

Contrary to the exponential case, this covariance function between r_t and $r_{t-\Delta}$ does not admit any closed form expression when $t\to\infty$.

3. Empirical motivation

This short section provides some empirical arguments motivating the developments done in this article, in a particular the choice of a ML or PML memory kernel. We fit univariate Gaussian models ($d=1$) with exponential, ML and PML kernels to the Eonia time series from the 13th of March 2016 to the 1st of August 2019. The data set counts n_{obs} =866 daily observations. We select this time window mainly because the Eonia was relatively stable during this period as illustrated in the right plot of Figure 3. This allows us to assume that the trend function, $\varphi(t)$, is constant. Considering a larger dataset would raise the question of modelling a non-linear trend in view of the left plot of Figure 3.

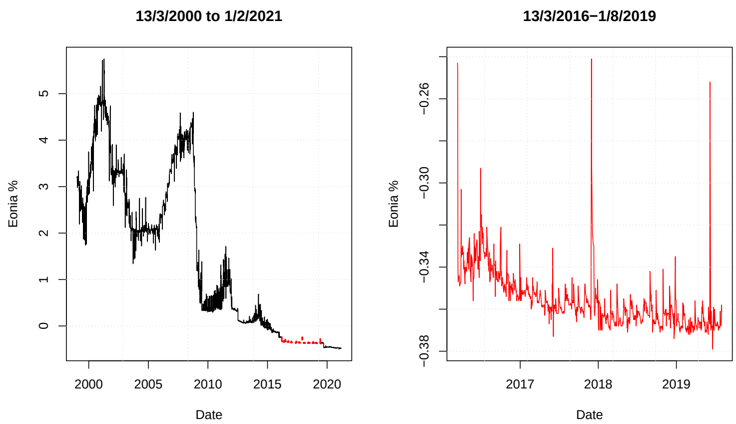

Figure 3. Left plot: Eonia time series from the 13/3/2000 to 1/2/2021. Right plot: dataset used to estimate parameters.

The dates of observations are denoted by $t_0,t_1,...,t_{n_{obs}-1}$ whereas the step of time between two successive observations is Δ_t. We first consider a one dimensional exponential kernel model as detailed in Appendix A. Under the assumption that $\varphi(t_k)=\varphi(t_k+\Delta_t)$ and $X_0=0$, we find that

$$\Delta L_{k+1} \quad = \quad L_{t_{k+1}}-L_{t_k}\approx r_{t_{k+1}}-r_{t_k}-\sum_{j=0}^{k}\left(e^{-\kappa(t_{k+1}-t_j)}-e^{-\kappa(t_k-t_j)}\right)\Delta L_j\,,$$

for $k=1,...,n_{obs}-1$. Under the assumption that L_t is a Brownian motion without drift and of variance σ^2, parameters are estimated by log-likelihood maximization. We next consider models with ML and PML kernel functions. In this cases, the variations of the underlying Lévy process are approached by

$$\Delta L_{k+1} \quad = \quad L_{t_{k+1}} - L_{t_k} \approx r_{t_{k+1}} - r_{t_k} - \sum_{j=0}^{k} \left(g(t_{k+1}-t_j) - g(t_k - t_j) \right) \, \Delta L_j \,,$$

for $k=1,...,n_{obs}-1$ where $g(.)$ is either the ML or the PML function. Under the assumption of normality, parameters are estimated by log-likelihood maximization whereas the goodness of fit is measured with the Akaike Information Criterion (AIC). The results of this procedure are reported in Table 1. In terms of log-likelihood, the exponential and ML models achieve the same goodness of fit but the exponential model is more parsimonious and then preferred. They share the same estimated volatilities but the parameters of reversion κ and β slightly differ. The best fit, based on the AIC, is obtained with a power Mittag Leffler kernel. The parameter α that drives the decay of the memory is small. This means that the process has a longer-term memory than the exponential model. On the other hand, the mean reversion speed is significantly higher than the ones of other models. We recall to the reader that estimated parameters reflect the dynamic of r_t under the real measure $\mathbb{P}$ and not the risk neutral one.

	Exponential		ML	PML
σ	1.4763e-3	σ	1.4767e-3	1.3927e-3
κ	0.6926	α	0.9954	0.1924
		β	0.7218	1.2171
Log-likelihood	6 809.75	Log-likelihood	6 809.93	6 861.15
AIC	-13 615.50		-13 613.86	-13 716.30

Table 1. Parameter estimates by log-likelihood maximization of one dimensional models with exponential, ML and PML kernels.

4. Alternative formulation

If the memory is not exponential decaying, the interest rate process $(r_t)_{t\geq 0}$ is not Markov. This feature makes difficult the evaluation of bond prices or interest rate derivatives because their values depend on the whole history of the rate process. Nevertheless, if the memory kernel admits a representation as a Laplace–Stieltjes integral, we can convert the short-term rate into an infinite dimensional Markov process. Approximating this process allows us to infer the dynamic of bond prices. As developments are similar in both cases, we denote by $\gamma_j(.)$ the measures $\gamma^{(j)}_{\alpha_j,\beta_j}(.)$ or $\gamma^{p(j)}_{\alpha_j,\beta_j}(.)$ of $E_{\alpha_j}(-\beta_j t)$ and $E_{\alpha_j}(-\beta_j t^{\alpha_j})$. Let us recall that the function $g_j(t) = \int_0^\infty e^{-\xi t} d\gamma_j(\xi)$. Using the Fubini's theorem, the processes $X_t^{(j)}$ are then rewritten as

$$\begin{aligned} X_t^{(j)} &= X_0^{(j)} \int_0^\infty e^{-t\,\xi} d\gamma_j(\xi) + \int_0^t \int_0^\infty e^{-\xi(t-u)} d\gamma_j(\xi)\, dL_u^{(j)} \\ &= \int_0^\infty \underbrace{\left(e^{-t\,\xi} Y_0^{(j,\xi)} + \int_0^t e^{-\xi(t-u)}\, dL_u^{(j)} \right)}_{Y_t^{(j,\xi)}} d\gamma_j(\xi) \end{aligned}$$

for $j=1,...,d$. The $Y_t^{(j,\xi)}$ are Ornstein-Uhlenbeck processes for $j=1,...,d$ and for all $\xi \in \mathbb{R}^+$. Their initial values are $X_0^{(j)} = Y_0^{(j,\xi)}$ and they obey to the following stochastic differential equation (SDE):

$$dY_t^{(j,\xi)} = -\xi Y_t^{(j,\xi)} dt + dL_t^{(j)} \,. \tag{17}$$

As this SDE admits the solution

$$Y_t^{(j,\xi)} = e^{-\xi(t-s)} Y_s^{(j,\xi)} + \int_s^t e^{-\xi(t-u)} dL_u^{(j)} , \tag{18}$$

the short-term rate is then a sum of integrals of $Y_t^{(j,\xi)}$:

$$r_t = \varphi(t) + \sum_{j=1}^{d} \int_0^\infty Y_t^{(j,\xi)} d\gamma_j(\xi) . \tag{19}$$

Given that $\gamma_{\alpha_j,\beta_j}^{(j)}(.)$ and $\gamma_{\alpha_j,\beta_j}^{p(j)}(.)$ are probability measures, the differential of r_t is on the other hand equal to

$$dr_t = d\varphi(t) - \sum_{j=1}^{d} \int_0^\infty \xi Y_t^{(j,\xi)} d\gamma_j(\xi) + \sum_{j=1}^{d} \underbrace{\int_0^\infty d\gamma_j(\xi)}_{=1} dL_t^{(j)} .$$

Combining Equations (18) and (19) allow us to rewrite the short-term rate as the following sum

$$\begin{aligned} r_t &= \varphi(t) + \sum_{j=1}^{d} \int_0^\infty \left(e^{-\xi(t-s)} Y_s^{(j,\xi)} + \int_s^t e^{-\xi(t-u)} dL_u^{(j)} \right) d\gamma_j(\xi) \\ &= \varphi(t) + \sum_{j=1}^{d} \int_0^\infty e^{-\xi(t-s)} Y_s^{(j,\xi)} d\gamma_j(\xi) + \sum_{j=1}^{d} \int_s^t g_j(t-u) dL_u^{(j)} \end{aligned} \tag{20}$$

which does not depend anymore upon information prior to s. This relation emphasizes that r_t and $\left(Y_t^{(j,\xi)}\right)_{\xi\in\mathbb{R}^+, j\in\{1,\dots,d\}}$ form well an infinite dimensional Markov process. If we remember that $\mathbb{E}(L_t^{(j)}) = 0$, we infer an equivalent representation of Equation (14) for the conditional expectation of r_t in terms of $Y_t^{(j,\xi)}$:

$$\mathbb{E}\left(r_t | \mathcal{F}_s\right) = \varphi(t) + \sum_{j=1}^{d} \int_0^\infty e^{-\xi(t-s)} Y_s^{(j,\xi)} d\gamma_j(\xi) .$$

We will explain in the next section how to approximate the infinity of processes by discretizing the measures $\gamma_j(\xi)$. But before we evaluate zero coupon bonds and instantaneous forward rates in this setting.

5. Bond prices and forward rates

A zero-coupon bond of maturity t delivers an unit cash-flow at expiry. Its price at time s prior to t is denoted by $P(s,t)$ and in absence of arbitrage, is the expected discount factor under the risk neutral measure, $P(s,t) = \mathbb{E}\left(e^{-\int_s^t r_u\, du} \mid \mathcal{F}_s\right)$. The next proposition presents a semi-closed form expression for this price.

Proposition 1. *Let us respectively define the functions $B^{(\xi)}(s,t)$ and $h_j(s,t)$ as follows*

$$B^{(\xi)}(s,t) = \frac{1}{\xi}\left(1 - e^{-\xi(t-s)}\right) , \tag{21}$$

$$h_j(s,t) = \int_0^\infty B^{(\xi)}(s,t)\, d\gamma_j(\xi) , \tag{22}$$

for $s \leq t$ and $j = 1, ..., d$. The zero coupon bond price is equal to

$$P(s,t) = \exp\Bigg(-\int_s^t \varphi(u)du + \sum_{j=1}^{d} \int_s^t \psi_j(-h_j(u,t))du \qquad (23)$$
$$-\sum_{j=1}^{d}\left(X_0^{(j)} \int_s^t g_j(u)du + \int_0^s \int_0^\infty e^{-\xi(s-u)} B^{\xi}(s,t)\, d\gamma_j(\xi)\, dL_u^{(j)} \right)\Bigg) .$$

Proof. From Equation (13), we infer after a change of the integration order that

$$\int_s^t r_u\, du = \int_s^t \varphi(u)du + \sum_{j=1}^{d} X_0^{(j)} \int_s^t g_j(u)du +$$
$$\sum_{j=1}^{d} \int_0^s \int_s^t g_j(u-v)\, du\, dL_v^{(j)} + \sum_{j=1}^{d} \int_s^t \int_v^t g_j(u-v)\, du\, dL_v^{(j)} .$$

A direct calculation allows us to rewrite the integrals of $g_j(.)$ in this last expression as

$$\int_s^t g_j(u-v)\, du \; = \int_0^\infty e^{-\xi(s-v)} B^{\xi}(s,t)\, d\gamma_j(\xi) ,$$
$$\int_v^t g_j(u-v)\, du = \int_0^\infty B^{\xi}(v,t)\, d\gamma_j(\xi) = h_j(v,t) .$$

The zero-coupon bond price becomes then:

$$\mathbb{E}\left(e^{-\int_s^t r_u\, du} \mid \mathcal{F}_s \right) = \mathbb{E}\left(\exp\left(-\sum_{j=1}^{d} \int_s^t h_j(v,t)\, dL_v^{(j)} \right) \mid \mathcal{F}_s \right) \times$$
$$\exp\left(-\int_s^t \varphi(u)du - \sum_{j=1}^{d}\left(X_0^{(j)} \int_s^t g_j(u)du + \int_0^s \int_0^\infty e^{-(s-v)} B^{\xi}(s,t)\, d\gamma_j(\xi)\, dL_v^{(j)} \right)\right) .$$

Finally, the expectation in this last expression is calculated with Equation (4).

□

Remark that $\lim_{\xi\to\infty} B^{(\xi)}(v,t) = 0$ and that $\lim_{\xi\to 0} B^{(\xi)}(v,t) = t - v$. Furthermore the function $B^{(\xi)}(v,t)$ admits a single maximum $\xi^* < \infty$ such that that $B^{(\xi^*)}(v,t) \geq B^{(\xi)}(v,t)$ for al $\xi \in \mathbb{R}^+$. Therefore the integral $h_j(v,t) = \int_0^\infty B^{(\xi)}(v,t)\, d\gamma_j(\xi)$ is bounded by $B^{(\xi^*)}(v,t)$ and is well defined. Equation (23) emphasizes the dependence of the bond price to the history of Levy processes. A more convenient formulation in terms of $\left(Y_s^{(j,\xi)}\right)_{\xi\geq 0}$ for $j = 1, ..., d$, is presented in the next proposition. In this alternative valuation formula, the bond price is exclusively calculated with information available at the time of valuation.

Proposition 2. *The value at time $s > 0$ of a zero coupon bond price expiring at $t > s$ is also equal to*

$$P(s,t) \;=\; \exp\left(-\int_s^t \varphi(u)du - \sum_{j=1}^{d} \int_0^\infty Y_s^{(j,\xi)} B^{(\xi)}(s,t)\, d\gamma_j(\xi) \right) \qquad (24)$$
$$\times \prod_{j=1}^{d} \exp\left(\int_s^t \psi_j\left(-h_j(u,t)\right) du \right) ,$$

where $B^{(\xi)}(s,t)$ and $h_j(v,t)$ for $j = 1, ..., d$ are defined by Equations (21) and (22).

Proof. Starting from Equation (23), we rewrite the last terms as follows

$$
\begin{aligned}
& X_0^{(j)} \int_s^t g_j(u)du + \int_0^s \int_0^\infty e^{-\xi(s-u)} B^\xi(s,t)\, d\gamma_j(\xi)\, dL_u^{(j)} \\
&= \int_0^\infty B^\xi(s,t) \left(X_0^{(j)} e^{-s\xi} + \int_0^s e^{-\xi(s-v)}\, dL_v^{(j)} \right) d\gamma_j(\xi) \\
&= \int_0^\infty B^\xi(s,t) Y_s^{(j,\xi)}\, d\gamma_j(\xi)\,.
\end{aligned}
$$

We conclude that the bond price is well given by Equation (24).

□

The integrals in the bond price formula do not admit closed form expressions. Nevertheless, the integrals $\int_s^t \psi_j\left(-h_j(u,t)\right) du$ can be numerically computed without any particular difficulty whereas $\int_0^\infty Y_s^{(j,\xi)} B^{(\xi)}(s,t)\, d\gamma_j(\xi)$ is approached by a discretization scheme developped in the following section.

Let us recall that the initial values of processes $\left(Y_t^{(j,\xi)}\right)_{t\geq 0, j\in\{1,\ldots,d\}}$ are null, that is $Y_0^{(j,\xi)} = 0$, for all $\xi \in \mathbb{R}^+$ and $j = 1, ..., d$. Therefore, Equation (24) provides us a way to estimate the function $\varphi(.)$ such that the model perfectly matches the initial term structure of bond prices. More precisely, the integral of $\varphi(.)$ is such that

$$
\int_0^t \varphi(u)du \;=\; -\ln\left(P(0,t)\right) + \sum_{j=1}^d \int_0^t \psi_j\left(-h_j(v,t)\right) dv\,.
$$

Deriving this last expression leads to the following function:

$$
\varphi(t) \;=\; -\frac{\partial}{\partial t} \ln\left(P(0,t)\right) - \sum_{j=1}^d \psi_j\left(-h_j(0,t)\right)\,. \tag{25}
$$

The dynamic of interest rates can also be reformulated in terms of instantaneous forward rates. Let us recall that the instantaneous forward rate, denoted by $f(s,t)$, is such that $P(s,t) = \exp\left(-\int_s^t f(s,u)\, du\right)$ and is therefore equal to:

$$
f(s,t) = -\frac{\partial}{\partial t} \ln P(s,t) \qquad 0 \leq s \leq t.
$$

The next proposition states that the forward rate is the sum of the expected interest rate under the risk neutral measure and of an adjustment that directly depends upon the memory kernel.

Proposition 3. *The instantaneous forward rate at time s and of maturity t is given by:*

$$
\begin{aligned}
f(s,t) \;=\; & \mathbb{E}\left(r_t|\mathcal{F}_s\right) + \sum_{j=1}^d \int_s^t g_j(t-v)\left(\mu_j - h_j(v,t)\,\sigma_j^2\right) dv \\
& + \sum_{j=1}^d \int_s^t g_j(t-v) \int_{\mathbb{R}} \left(z\, e^{-h_j(v,t)\, z} - z 1_{|z|<\epsilon}\right) \nu_j(dz)\, dv\,,
\end{aligned} \tag{26}
$$

where $\mathbb{E}\left(r_t|\mathcal{F}_s\right) = \varphi(t) + \sum_{j=1}^{d}\int_0^\infty e^{-\xi(t-s)}Y_s^{(j,\xi)}d\gamma_j(\xi)$ *. Its dynamic is independent from processes* $Y_s^{(j,\xi)}$ *and is equal to*

$$\begin{aligned} df(s,t) &= \sum_{j=1}^{d} g_j(t-s)\,dL_s^{(j)} - \sum_{j=1}^{d} g_j(t-s)\left(\mu_j - h_j(v,t)\,\sigma_j^2\right)ds \\ &\quad - \sum_{j=1}^{d} g_j(t-s)\left(\int_{\mathbb{R}}\left(z\,e^{-h_j(v,t)\,z} - z1_{|z|<\epsilon}\right)\nu_j(dz)\right)ds\,. \end{aligned}$$

Proof. From Equation (24) and $\frac{\partial B^{(\xi)}(s,t)}{\partial t} = e^{-\xi(t-s)}$, we have that

$$\begin{aligned} &-\frac{\partial}{\partial t}\ln P(s,t) \\ &= \mathbb{E}\left(r_t|\mathcal{F}_s\right) - \sum_{j=1}^{d}\int_s^t \frac{\partial}{\partial t}\psi_j\left(-\int_0^\infty B^{(\xi)}(v,t)\,d\gamma_j(\xi)\right)dv\,. \end{aligned}$$

If we remember the expression (24) of the characteristic exponent, we immediately infer that

$$\begin{aligned} &\frac{\partial}{\partial t}\psi_j\left(-\int_0^\infty B^{(\xi)}(v,t)\,d\gamma_j(\xi)\right) = \\ &-\mu_j\int_0^\infty e^{-\xi(t-v)}\,d\gamma_j(\xi) + \left(\int_0^\infty B^{(\xi)}(v,t)\,d\gamma_j(\xi)\right)\int_0^\infty e^{-\xi(t-v)}\,d\gamma_j(\xi)\,\sigma_j^2 \\ &+\int_{\mathbb{R}} z\left(-\int_0^\infty e^{-\xi(t-v)}\,d\gamma_j(\xi)e^{-\int_0^\infty B^{(\xi)}(v,t)\,d\gamma_j(\xi)\,z} + \int_0^\infty e^{-\xi(t-v)}\,d\gamma_j(\xi)\,1_{|z|<\epsilon}\right)\nu_j(dz). \end{aligned}$$

As $\int_0^\infty e^{-\xi(t-v)}\,d\gamma_j(\xi) = g_j(t-v)$, we obtain Equation (26). The differential of $f(s,t)$ is obtained by applying the Itô's lemma for Lévy processes. □

This last proposition emphasizes that our short-term rate model can be reformulated as a forward rate model in which the forward rate dynamic is independent from processes $\left(Y_s^{(j,\xi)}\right)_{\xi\geq 0}$ for $j = 1,...,d$.

6. Discretization scheme

Instead of considering an infinity of processes $\left(Y_t^{(j,\xi)}\right)_{t\geq 0, j\in\{1,...,d\}}$, we approach the model with a finite number of equivalent processes. This presents several advantages. Firstly, this makes possible implementing our model. Secondly, we can rely on the Itô's calculus in most of developments. This allows us to deduce the dynamic of bond prices both in the approximated and original models. The key step consists to approximate the $\gamma_j(.)$'s by discrete measures with a finite numbers of atoms. For this purpose we consider a partition $\mathcal{E}^{(n)} := \{0 < \xi_0^{(n)} < \xi_1^{(n)} < ... < \xi_n^{(n)} < \infty\}$. On each interval $(\xi_i^{(n)}, \xi_{i+1}^{(n)})$, we define the barycenter of $\gamma_j(.)$:

$$b_{k+1}^{(j)} = \frac{\int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} z\,d\gamma_j(z)}{\int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} d\gamma_j(z)} \quad j = 1,...,d \tag{27}$$

and the mass of corresponding atoms is defined as the measure of intervals of the partition:

$$m_{k+1}^{(j)} = \int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} d\gamma_j(z)\,, \tag{28}$$

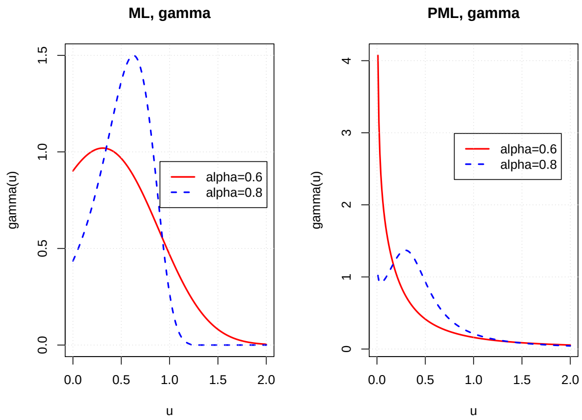

Figure 4. Left and right plots: measures $\gamma(.)$ of the Mittag Leffler (ML) and power Mittag Leffler (PML) kernels. $\beta = 0.5$.

for $k = 0, ..., n-1$. In practice, we choose $\xi_0^{(n)} = 0$ and set $\xi_n^{(n)}$ to a percentile of the density $\gamma_j(z)$ (e.g. 95%). The discrete measure for a partition of size n is defined as follows

$$\gamma_j^{(n)}(z) = \sum_{k=1}^{n} m_k^{(j)} \delta_{b_k^{(j)}}(z)\,, \tag{29}$$

where $\delta_{b_k^{(j)}}(z)$ is the Dirac measure located at point $b_k^{(j)}$. We consider that the following assumption holds for the partition $\mathcal{E}^{(n)}$:

- $\xi_0^{(n)} \to 0$ and $\xi_n^{(n)} \to \infty$ when $n \to \infty$.
- $\max |\xi_{i+1}^{(n)} - \xi_i^{(n)}| \to 0$ when $n \to \infty$.
- $\mathcal{E}^{(n)} \subset \mathcal{E}^{(n+1)}$.

In this case, for any function $f(.)$ integrable with respect to $\gamma_j(.)$, we have that $\lim_{n\to\infty} \int_0^\infty f(z)\, d\gamma_j^{(n)}(z) = \int_0^\infty f(z)\, d\gamma_j(z)$. Figure 4 shows the differential of measures $d\gamma(.)$ of the ML and PML kernels. For the ML kernels (left plot), decreasing the α clearly makes fatter the right tail of $d\gamma(.)$. The right plot also reveals that the construction of the partition requires a particular care for the PML kernel as we have to integrate numerically the measure $\gamma_{\alpha,\beta}^{(p)}(z)$ that is not defined for $z = 0$. The two next propositions respectively present the expressions of integrals needed for calculating the discrete equivalent distributions of $\gamma_j(.)$ in the ML and PML cases.

Proposition 4. *For the ML kernel, the discrete probability mass of atoms is given by*

$$\begin{aligned} m_{k+1}^{(j)} &= \int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} d\gamma_j(z) \\ &= \frac{1}{\pi\alpha_j} \sum_{l=1}^{\infty} \frac{(-1)^{l-1}}{l\, \beta_j^l\, l!} \sin(\pi\alpha_j l)\, \Gamma(\alpha_j l + 1) \left[\left(\xi_{k+1}^{(n)}\right)^l - \left(\xi_k^{(n)}\right)^l \right], \end{aligned} \tag{30}$$

whereas the numerator in the expression of barycenters (27) is

$$\int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} z d\gamma_j(z) = \frac{1}{\pi\alpha_j} \sum_{l=1}^{\infty} \frac{(-1)^{l-1}}{\beta^l\, (l+1)!} \sin(\pi\alpha_j l)\, \Gamma(\alpha_j l + 1) \left[\left(\xi_{k+1}^{(n)}\right)^{l+1} - \left(\xi_k^{(n)}\right)^{l+1} \right]. \tag{31}$$

This result is a direct consequence of the definition of the ML function. To the best of our knowledge, these expressions does not admit any other representations. In the PML case, barycenters have a closed-form expression.

Proposition 5. *For the PML kernel, the discrete probability mass of atoms is given by*

$$\begin{aligned} m_{k+1}^{(j)} = \int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} d\gamma_j(z) &= \frac{1}{\pi\alpha_j}\arctan\left(\frac{\left(\xi_{k+1}^{(n)}\right)^{\alpha_j} + \beta_j\cos(\pi\alpha_j)}{\beta_j\sin(\pi\alpha_j)}\right) \\ &\quad -\frac{1}{\pi\alpha_j}\arctan\left(\frac{\left(\xi_{k}^{(n)}\right)^{\alpha_j} + \beta_j\cos(\pi\alpha_j)}{\beta_j\sin(\pi\alpha_j)}\right). \end{aligned} \tag{32}$$

The numerator in the expression of barycenters (27) does not admit a closed form expression but may be approached by:

$$\begin{aligned} \int_{\xi_k^{(n)}}^{\xi_{k+1}^{(n)}} z d\gamma_j(z) \approx \left(\frac{\xi_{k+1}^{(n)} + \xi_k^{(n)}}{2\pi\alpha_j}\right)\Bigg[&\arctan\left(\frac{\left(\xi_{k+1}^{(n)}\right)^{\alpha} + \beta_j\cos(\pi\alpha_j)}{\beta_j\sin(\pi\alpha_j)}\right) \\ &-\arctan\left(\frac{\left(\xi_{k}^{(n)}\right)^{\alpha} + \beta_j\cos(\pi\alpha_j)}{\beta_j\sin(\pi\alpha_j)}\right)\Bigg]. \end{aligned} \tag{33}$$

Proof. The expression of the probability mass $m_{k+1}^{(j)}$ comes from the relation

$$\frac{\beta\ \sin(\alpha\pi)}{\pi}\int\frac{z^{\alpha-1}}{z^{2\alpha}+2\beta z^{\alpha}\cos(\alpha\pi)+\beta^2}dz = \frac{1}{\pi\alpha}\arctan\left(\frac{z^{\alpha}+\beta\cos(\pi\alpha)}{\beta\sin(\pi\alpha)}\right)$$

The second result is obtained with an integration by parts in which the integral is approached with the trapezoidal method.

□

To lighten further developments, we adopt the following notations $Y_s^{(j,k)} := Y_s^{(j,b_k^{(n)})}$ for $k = 1, ..., n$ and $j = 1, ..., d$. Let us first recall that $Y_0^{(j,k)} = 0$ and therefore $Y_s^{(j,k)}$ is an integral from zero to s with respect to the j^{th} Lévy process

$$Y_s^{(j,k)} = \int_0^s e^{-b_k^{(j)}(s-u)} dL_u^{(j)}. \tag{34}$$

Replacing the continuous measures $\left(\gamma_j\right)_{j=1...d}$ by their discrete equivalents leads to the approximation $r_s^{(n)}$ of the short-term rate r_s:

$$r_s^{(n)} = \varphi^{(n)}(s) + \sum_{j=1}^{d}\sum_{k=1}^{n} m_k^{(j)} Y_s^{(j,k)}. \tag{35}$$

We adopt the following notations: $B^{(j,k)}(s,t) = \frac{1}{b_k^{(j)}}\left(1 - e^{-b_k^{(j)}(t-s)}\right)$ for $k = 1, ..., n$ and $j = 1, ..., d$. The bond price in the discretized model with partitions of size n is denoted by $P^{(n)}(s,t)$ and is equal to

$$\begin{aligned} P^{(n)}(s,t) &= \exp\left(-\int_s^t \varphi^{(n)}(u)du - \sum_{j=1}^{d}\sum_{k=1}^{n} Y_s^{(j,k)} B^{(j,k)}(s,t)\, m_k^{(j)}\right) \\ &\times \prod_{j=1}^{d} \exp\left(\int_s^t \psi_j\left(-\sum_{k=1}^{n} B^{(j,k)}(v,t)\, m_k^{(j)}\right) dv\right), \end{aligned} \tag{36}$$

where $\varphi^{(n)}(t)$ is the discretized version of $\varphi(t)$:

$$\varphi^{(n)}(t) = -\frac{\partial}{\partial t}\ln\left(P(0,t)\right) - \sum_{j=1}^{d}\psi_j\left(-\sum_{k=1}^{n} B^{(j,k)}(0,t)\, m_k^{(j)}\right). \tag{37}$$

As the number of processes in the discretized model is finite, we can apply the Itô's lemma in order to establish the dynamic of $P^{(n)}(s,t)$.

Proposition 6. *The zero coupon bond price $P^{(n)}(s,t)$ is a geometric Lévy processes solution of the SDE:*

$$\begin{aligned} \frac{dP^{(n)}(s,t)}{P^{(n)}(s,t)} &= r_s^{(n)}ds - \sum_{j=1}^{d}\sigma_j\sum_{k=1}^{n} B^{(j,k)}(s,t)\, m_k^{(j)} dW_s^{(j)} \\ &+ \sum_{j=1}^{d}\int_{\mathbb{R}}\left(e^{-z\left(\sum_{k=1}^{n} m_k^{(j)} B^{(j,k)}(s,t)\right)} - 1\right)\left(J_j(ds\,,dz) - \nu_j(dz)ds\right). \end{aligned} \tag{38}$$

with the terminal condition $P^{(n)}(t,t) = 1$.

Proof. If we remember that $dY_s^{(j,k)} = dL_s^{(j)} - b_k^{(j)} Y_s^{(j,k)} ds$, the Itô's lemma gives us the following differential for $P^{(n)}(s,t)$:

$$\begin{aligned} dP^{(n)}(s,t) &= \frac{\partial P^{(n)}(s,t)}{\partial s}ds - \sum_{j=1}^{d}\sum_{k=1}^{n}\frac{\partial P^{(n)}(s,t)}{\partial Y^{(j,k)}} b_k^{(j)} Y_s^{(j,k)} ds \\ &+ \sum_{j=1}^{d}\sum_{k=1}^{n}\frac{\partial P^{(n)}(s,t)}{\partial Y^{(j,k)}}\mu_j ds + \sum_{j=1}^{d}\sum_{k=1}^{n}\frac{\partial P^{(n)}(s,t)}{\partial Y^{(j,k)}}\int_{\mathbb{R}} z\,\tilde{J}_j(dt\,,dz) \\ &+ \sum_{j=1}^{d}\sum_{k=1}^{n}\frac{\partial P^{(n)}(s,t)}{\partial Y^{(j,k)}}\sigma_j dW_s^{(j)} + \frac{1}{2}\sum_{j=1}^{d}\sum_{k=1}^{n}\sum_{l=1}^{n}\frac{\partial^2 P^{(n)}(s,t)}{\partial Y^{(j,k)}\partial Y^{(j,l)}}\sigma_j^2 ds \\ &+ \sum_{j=1}^{d}\int_{\mathbb{R}} P^{(n)}(s,t)\left(e^{-z\left(\sum_{k=1}^{n} m_k^{(j)} B^{(j,k)}(s,t)\right)} - 1\right) - z\sum_{k=1}^{n}\frac{\partial P^{(n)}(s,t)}{\partial Y^{(j,k)}} J_j(ds\,,dz). \end{aligned} \tag{39}$$

The first and second order partial derivatives of $P^{(n)}(s,t)$ are equal to:

$$\frac{\partial P^{(n)}(s,t)}{\partial Y^{(j,k)}} = -P^{(n)}(s,t)\, B^{(j,k)}(s,t)\, m_k^{(j)}, \tag{40}$$

$$\frac{\partial^2 P^{(n)}(s,t)}{\partial Y^{(j,k)}\partial Y^{(i,l)}} = P^{(n)}(s,t)\, B^{(j,k)}(s,t) B^{(i,l)}(s,t)\, m_k^{(j)} m_l^{(i)}. \tag{41}$$

Whereas its partial derivative with respect to time is given by

$$\frac{\partial P^{(n)}(s,t)}{\partial s} = P^{(n)}(s,t)\left(\varphi^{(n)}(s) + \sum_{j=1}^{d}\sum_{k=1}^{n} Y_s^{(j,k)} e^{-b_k^{(j)}(t-s)}\, m_k^{(j)} - \sum_{j=1}^{d}\psi_j\left(-\sum_{k=1}^{n} B^{(j,k)}(s,t)\, m_k^{(j)}\right)\right), \tag{42}$$

where the characteristic exponent is developed as the following sum

$$\psi_j\left(-\sum_{k=1}^{n} B^{(j,k)}(s,t)\, m_k^{(j)}\right) = -\mu_j \sum_{k=1}^{n} B^{(j,k)}(s,t)\, m_k^{(j)} + \frac{1}{2}\left(\sum_{k=1}^{n} B^{(j,k)}(s,t)\, m_k^{(j)}\right)^2 \sigma_j^2 \tag{43}$$
$$\int_{\mathbb{R}}\left(e^{-\left(\sum_{k=1}^{n} B^{(j,k)}(s,t)\, m_k^{(j)}\right)z} - 1 + \left(\sum_{k=1}^{n} B^{(j,k)}(s,t)\, m_k^{(j)}\right) z 1_{|z|<\epsilon}\right)\nu_j(dz)\,.$$

Combining Equations (39), (40), (41), (42) and (43) allows us to rewrite the dynamic of bond prices as follows:

$$\begin{aligned} dP^{(n)}(s,t) &= P^{(n)}(s,t)\left(\varphi^{(n)}(s) + \sum_{j=1}^{d}\sum_{k=1}^{n} Y_s^{(j,k)} e^{-b_k^{(j)}(t-s)}\, m_k^{(j)}\right) ds \\ &\quad + P^{(n)}(s,t)\sum_{j=1}^{d}\sum_{k=1}^{n} Y_s^{(j,k)} B^{(j,k)}(s,t)\, m_k^{(j)} b_k^{(j)} ds \\ &\quad - P^{(n)}(s,t)\sum_{j=1}^{d}\sigma_j \sum_{k=1}^{n} B^{(k)}(s,t)\, m_k^{(n)} dW_s^{(j)} \\ &\quad + P^{(n)}(s,t)\sum_{j=1}^{d}\int_{\mathbb{R}}\left(e^{-z\left(\sum_{k=1}^{n} m_k^{(j)} B^{(j,k)}(s,t)\right)} - 1\right)\left(J_j(ds\,,\,dz) - \nu_j(dz)ds\right)\,. \end{aligned}$$

We infer the result if we remember Equation (35) and notice that

$$\sum_{k=1}^{n} Y_s^{(j,k)} e^{-b_k^{(j)}(t-s)}\, m_k^{(j)} + \sum_{k=1}^{n} Y_s^{(j,k)} B^{(k)}(s,t)\, m_k^{(j)} b_k^{(j)} = \sum_{k=1}^{n} Y_s^{(j,k)}\, m_k^{(j)}\,.$$

□

As by construction $\lim_{n\to\infty}\gamma_j^{(n)}(z) = \gamma_j(z)$, we immediately infer the dynamic of the bond price in the non-discretized model by considering the limit of Equation (38).

Corollary 1. *If the short-term rate is driven by Equation (19), the bond price is solution of the following SDE:*

$$\frac{dP(s,t)}{P(s,t)} = r_s ds - \sum_{j=1}^{d}\sigma_j h_j(s,t) dW_s^{(j)} + \sum_{j=1}^{d}\int_{\mathbb{R}}\left(e^{-zh_j(s,t)} - 1\right)\left(J_j(ds\,,\,dz) - \nu_j(dz)ds\right)\,. \tag{44}$$

where $h_j(s,t) = \int_0^\infty B^{(\xi)}(s,t)\, d\gamma_j(\xi)$ and with the terminal condition $P(t,t) = 1$.

In order to understand the role played by the memory kernel in the evolution of the term structure of interest rates, we analyze the expectation and variance of bond yields. The bond yield of a zero-coupon bond in the discretized model is defined as

$$y^{(n)}(s,t) = -\frac{\ln P^{(n)}(s,t)}{t-s}\,.$$

From Equation (36), we infer that this yield is proportional to a weighted sum of mean reverting processes $Y_s^{(j,k)}$. Since $\mathbb{E}\left(Y_s^{(j,k)}|\mathcal{F}_0\right)=0$ by construction, the expected yield conditionally to the initial filtration is equal to

$$\mathbb{E}\left(y^{(n)}(s,t)\,|\,\mathcal{F}_0\right) = \frac{1}{t-s}\left(\int_s^t \varphi^{(n)}(u)du - \sum_{j=1}^{d}\int_s^t \psi_j\left(-\sum_{k=1}^{n} B^{(j,k)}(v,t)\,m_k^{(j)}\right)dv\right)\,. \tag{45}$$

The variance of the yield is the sum of d variances of a linear combination of processes $Y_s^{(j,k)}$. If we remember that the covariance $Y_s^{(j,k)}$ and $Y_s^{(l,k)}$ is equal to

$$\begin{aligned}\mathbb{C}\left(Y_s^{(j,k)}Y_s^{(j,l)}|\mathcal{F}_0\right) &= \left(\sigma_j^2+\int_{\mathbb{R}} z^2\nu_j(dz)\right)\int_0^s e^{-\left(b_k^{(j)}+b_l^{(j)}\right)(s-u)}du \\ &= \frac{\sigma_j^2+\int_{\mathbb{R}} z^2\nu_j(dz)}{b_k^{(j)}+b_l^{(j)}}\left(1-e^{-\left(b_k^{(j)}+b_l^{(j)}\right)s}\right)\,,\end{aligned} \tag{46}$$

the variance of the yield, conditionally to $\mathcal{F}_0$, is the following triple sum:

$$\begin{aligned}\mathbb{V}\left(y^{(n)}(s,t)\,|\,\mathcal{F}_0\right) &= \sum_{j=1}^{d}\mathbb{V}\left(\sum_{k=1}^{n}\frac{B^{(j,k)}(s,t)\,m_k^{(j)}}{t-s}Y_s^{(j,k)}|\mathcal{F}_0\right) \\ &= \sum_{j=1}^{d}\left[\sum_{k=1}^{n}\sum_{l=1}^{n}\frac{B^{(j,k)}(s,t)\,m_k^{(j)}}{t-s}\frac{B^{(j,l)}(s,t)\,m_l^{(j)}}{t-s}\mathbb{C}\left(Y_s^{(j,k)}Y_s^{(j,l)}|\mathcal{F}_0\right)\right]\,.\end{aligned} \tag{47}$$

To illustrate these results, we consider an univariate jump-diffusion model ($d=1$) in which a single Lévy process is ruled by the following SDE

$$dL_t^{(1)} \;=\; -\lambda_1\eta_1 dt+\sigma_1 dW_t^{(1)}+\eta_1\,dN_t^{(1)}\,. \tag{48}$$

$N_t^{(1)}$ is a Poisson process with constant intensity λ_1. The characteristic exponent is in this case equal to

$$\psi_1(\omega) \;=\; -\lambda_1\eta_1\omega+\frac{1}{2}\omega^2\sigma_1^2+\lambda_1\left(e^{\omega\eta_1}-1\right), \tag{49}$$

whereas the Lévy measure is $\nu_1(z)=\lambda_1\delta_{\eta_1}(z)$ where $\delta_{\eta_1}(z)$ is the Dirac measure located at $z=\eta_1$. The instantaneous variance of $dL_t^{(1)}$ is equal to

$$\sigma_1^2+\int_{\mathbb{R}} z^2\nu_1(dz) \;=\; \sigma_1^2+\lambda_1\eta_1^2\,.$$

Of course, we can consider other type of Lévy processes like the variance gamma and the Normal inverse gaussian but this does not fundamentally modify the conclusions drawn in this section. We first fit the curve $\varphi(t)$ to the term structure of zero-coupon bond yields, bootstrapped from the ICE swap rates on the 26/2/21 at noon. Since this function is proportional to instantaneous forward rates, we have to interpolate bond yields. For this purpose, we use a Nelson-Siegel (NS) model. Appendix B recalls this model, reports the swap rates curve on the 26/2/21 and parameter estimates. There exist more advanced interpolation methods, as e.g. splines or kriging techniques as detailed in Cousin et al. (2016) but the NS model is sufficiently accurate for our purpose that is to understand the dynamic of yields over time.

We calculate the term structure of expected yields in 5, 10 and 15 years with n =40 atoms. Increasing n does not significantly modify the results. A sensitivity analysis with respect to the number of atoms is proposed in Section 7. For the parameters, we set $\beta = 1.5$, $\lambda_1 = 0.5$, $\sigma_1 = 0.01$ and $\eta_1 = -0.0002$, which are possible realistic values in view of results from Section 3. We remind to the reader that α is the parameter tuning the memory of the model. If $\alpha = 1$, the memory kernel is exponential and the model forgets in an exponential manner past fluctuations of interest rates. Whereas for lower values of α, the model forgets these variations according to a power decaying function. Figures 5 and 6 show expected yields computed with the ML and PML models for $\alpha = 0.5$ and $\alpha = 0.9$. We use a discretization scheme with 40 atoms that cover the ML and PML density up to their 90% percentiles. For both models, the range between short and long-term yields is narrowing with the time horizon. In the ML model, the lower is α, the longer is the memory and the quicker is the convergence to a bumped (but nearly flat) yield curve. In the PML model, expected yield curves with a low α dominates those with a high α and have a more pronounced curvature.

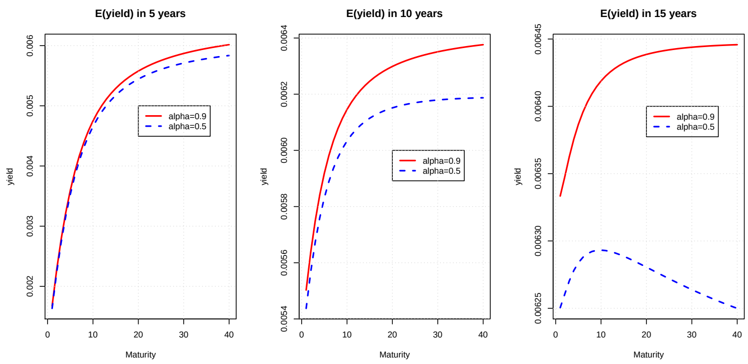

Figure 5. Expected yield curves in 5, 10 and 15 years, α =0.5 and 0.9, β = 1.5, λ_1=0.5 η_1= -0.002, $\sigma_1 = 0.01$. ML kernel with n =40 atoms.

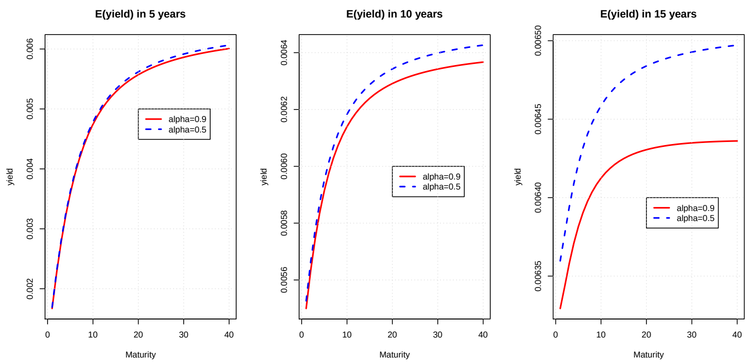

Figure 6. Expected yield curves in 5, 10 and 15 years, α =0.5 and 0.9, β = 1.5, λ_1=0.5 η_1= -0.002, $\sigma_1 = 0.01$. PML kernel with n =40 atoms.

Figures 7 and 8 report the expected standard deviations of future yields calculated with the ML and PML models. In both cases, the term structures of expected standard deviations do not significantly evolve with the time horizon and are decreasing functions of the maturity. We also notice that standard deviations are respectively inversely and directly proportional to α in the ML and PML models.

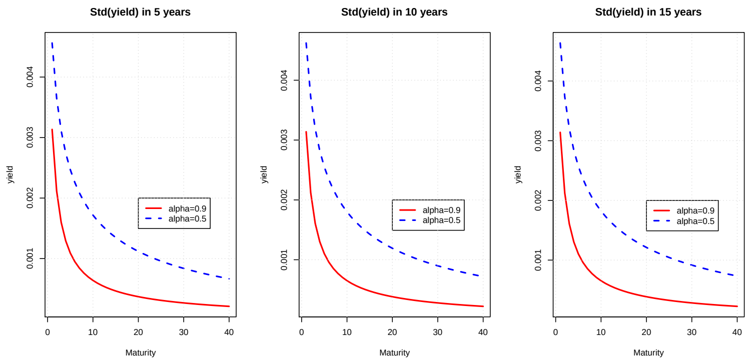

Figure 7. Expected standard deviations of yields in 5, 10 and 15 years, α =0.5 and 0.9, $\beta = 1.5$, λ_1=0.5 η_1= -0.002, $\sigma_1 = 0.01$. ML kernel with n =40 atoms.

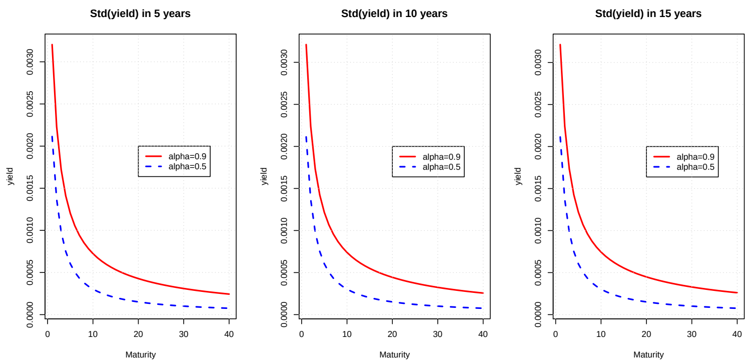

Figure 8. Expected standard deviations of yields in 5, 10 and 15 years, α =0.5 and 0.9, $\beta = 1.5$, λ_1=0.5 η_1= -0.002, $\sigma_1 = 0.01$. PML kernel with n =40 atoms.

7. Pricing of bond options

The purpose of this section is to illustrate how long-term memory models can be used for the pricing of derivatives. We investigate this for European options written on a zero coupon bond. Our choice is motivated by the fact that these are the building blocks of many structured products. The underlying bond has a maturity noted t_2 while the option expires at time $t_1 \leq t_2$ with a strike price K. Without loss of generality we focus on a call option but all developments remain valid for any other type of European options. The value at time $s \leq t_1 \leq t_2$ of the call option is denoted by $C(s)$ and is equal to the expected discounted payoff under the risk neutral measure $\mathbb{Q}$:

$$C(s) \quad = \quad \mathbb{E}\left(e^{-\int_s^{t_1} r_u \, du} \left(P(t_1,t_2) - K\right)_+ \mid \mathcal{F}_s\right) .$$

This price does not admit any closed form expression. Nevertheless we can apply the standard technique of changes of numeraire combined with a discrete Fourier's transform (DFT) to evaluate this call option. The first stage consists to define a forward measure, denoted by $\mathbb{F}$ that has $P(s,t_1)$ as numeraire, i.e. every assets discounted by $P(s,t_1)$ is a martingale. If $S_t^{(0)} = e^{\int_0^t r_u \, du}$ is the cash account (numeraire of $\mathbb{Q}$), the Radon-Nykodym derivative defining this change of measure is $\left.\frac{d\mathbb{F}}{d\mathbb{Q}}\right|_{t_1} = \frac{S_0^{(0)}}{S_{t_1}^{(0)} P(0,t_1)}$. Using standard arguments, the call price is equal to the product of a discount bond and of the expected payoff under this forward measure,

$$C(s) \quad = \quad P(s,t_1)\mathbb{E}^{\mathbb{F}}\left(\left(P(t_1,t_2) - K\right)_+ \mid \mathcal{F}_s\right) . \tag{50}$$

The next proposition details the dynamics of underlying Lévy processes under the forward measure.

Proposition 7. *Under the forward measure $\mathbb{F}$ with $P(s,t_1)$ as numeraire, the process $\left(W_s^{(j)\mathbb{F}}\right)_{s\geq 0}$ such that*

$$dW_s^{(j)\mathbb{F}} = dW_t^{(j)} + h_j(u,t_1)du\,. \tag{51}$$

for $j=1,...d$, is a Brownian motion. Furthermore, the random measure

$$\tilde{J}_j^{\mathbb{F}}(du,dz) \quad = \quad \tilde{J}_j(du,dz) - 1_{|z|<\epsilon}\left(e^{-h_j(u,t_1)z}-1\right)\nu_j(dz)du\,, \tag{52}$$

for $j=1,...d$, defines a jump process under $\tilde{\mathbb{P}}$ with the time-varying measure $\nu_j^{\mathbb{F}}(dt,dz)=e^{-h_j(u,t_1)z}\nu_j(dz)$.

Proof. From Applebaum (2004, chapter 5, sections 2 and 4), the forward change of measure

$$\begin{aligned}\left.\frac{d\mathbb{F}}{d\mathbb{Q}}\right|_{t_1} &= \frac{S_0^{(0)}}{S_{t_1}^{(0)}P(0,t_1)} \qquad (53)\\ &= \exp\left(\sum_{j=1}^{d}\int_0^{t_1} -h_j(u,t_1)\sigma_j dW_u^{(j)} - \frac{1}{2}h_j(u,t_1)^2\sigma_j^2 du\right.\\ &\quad -\sum_{j=1}^{d}\int_0^{t_1}\int_{\mathbb{R}}\left(e^{-h_j(u,t_1)z}-1+h_j(u,t_1)z1_{|z|<\epsilon}\right)\nu_j(dz)du\\ &\quad \left.+\sum_{j=1}^{d}\int_0^{t_1}\int_{|z|\geq\epsilon} -h_j(u,t_1)z\,\tilde{J}_j(du\,,\,dz)\right)\end{aligned}$$

is a martingale under the measure $\mathbb{Q}$ which defines an equivalent measure $\mathbb{F}$ under which $\left(W_s^{(j)\mathbb{F}}\right)_{s\geq 0}$ and $\tilde{J}_j^{\mathbb{F}}(du,dz)$ for $j=1,...d$ are respectively Brownian motions and random jump measures satisfying Equations (51) and (52).
□

Combining Corollary 1 and Proposition 7 allows us to infer that a zero-coupon bond of maturity $t\geq t_1$ is ruled under the forward measure by the SDE

$$\begin{aligned}\frac{dP(s,t)}{P(s,t)} \quad &= \quad \left(r_s + \sum_{j=1}^{d}\sigma_j h_j(s,t_1)^2 + \sum_{j=1}^{d}\int_{\mathbb{R}}\left(e^{-zh_j(s,t_1)}-1\right)^2\nu_j(dz)\right)ds \qquad (54)\\ &\quad + \sum_{j=1}^{d}\int_{\mathbb{R}}\left(e^{-zh_j(s,t_1)}-1\right)\left(J_j^{\mathbb{F}}(ds\,,\,dz) - e^{-h_j(s,t_1)z}\nu_j(dz)ds\right)\\ &\quad - \sum_{j=1}^{d}\sigma_j h_j(s,t_1)dW_s^{(j)\mathbb{F}}\,.\end{aligned}$$

This equation emphasizes that the average bond return under the forward measure is the risk-free rate plus a positive time varying premium that converges toward zero when $s\to t_1$. The next proposition reveals that the processes $Y_t^{(j,\xi)}$ do not revert anymore toward zero under $\mathbb{F}$.

Proposition 8. *Let us define the function*

$$\theta^{(j,\xi)}(u) = \begin{cases} -\left(h_j(u,t_1)\sigma_j + \int_{\mathbb{R}} z\left(1-e^{-h_j(u,t_1)z}1_{|z|<\epsilon}\right)\nu_j(dz)\right) & \xi>0 \\ 0 & \xi=0 \end{cases}. \tag{55}$$

Under the forward measure $\mathbb{F}$, $Y_t^{(j,\xi)}$ is a process reverting to $\theta^{(j,\xi)}(t)/\xi$ at a speed ξ, such that

$$Y_t^{(j,\xi)} = e^{-\xi(t-s)}Y_s^{(j)} + \int_s^t e^{-\xi(t-u)}\theta^{(j,\xi)}(u)du \qquad (56)$$
$$+\sigma_j \int_s^t e^{-\xi(t-u)}dW_u^{(j)\mathbb{F}} + \int_s^t e^{-\xi(t-u)}\int_{\mathbb{R}} z\tilde{J}_j^{\mathbb{F}}(du,dz) \quad ,$$

for all $j = 1, ..., d$ and $\xi \in \mathbb{R}^+$.

Proof. By definition of $Y_s^{(j,\xi)}$ and from Corollary 1, we immediately infer its dynamic under the forward measure.

$$dY_s^{(j,\xi)} = -\xi Y_s^{(j,\xi)}ds + \sigma_j dW_s^{(j)\mathbb{F}} + \int_{\mathbb{R}} z\tilde{J}_j^{\mathbb{F}}(ds,dz) \qquad (57)$$
$$+\left(\mu_j - h_j(s,t_1)\sigma_j + \int_{|z|<\epsilon} z\left(e^{-h_j(s,t_1)z}-1\right)\nu_j(dz)\right)ds\,.$$

Since we impose that $\mu_j = -\int_{\mathbb{R}}\left(z - z1_{|z|<\epsilon}\right)\nu_j(dz)$ for $j = 1, ..., d$ in order to ensure that $\mathbb{E}\left(L_t^{(j)}\right) = 0$, we deduce that $Y_s^{(j,\xi)}$ reverts to $\theta^{(j,\xi)}(t)/\xi$:

$$dY_s^{(j,\xi)} \;=\; \xi\left(\frac{1}{\xi}\theta^{(j,\xi)}(s) - Y_s^{(j,\xi)}\right)ds + \sigma_j dW_s^{(j)\mathbb{F}} + \int_{\mathbb{R}} z\tilde{J}_j^{\mathbb{F}}(ds,dz)\,,$$

and the solution of this SDE is well Equation (56). □

The equations (53) and (57) are useful to simulate sample paths of interest rates and bond prices under the forward measure e.g. for pricing exotic options. One solution to price a call option on a zero-coupon bond consists to numerically invert the Fourier's transform of the density of its bond log-return under the forward measure. This step is performed with a standard discrete Fourier's transform algorithm (DFT) that approximates the probability density function of the log-return under the forward measure $\mathbb{F}$. The expectation of the payoff under the forward measure is next computed with this distribution. The moment generating function of the bond log-return, presented in the next proposition, admits a closed form formula and is a key result to implement the DFT procedure.

Proposition 9. *The moment generating function (mgf) of $\ln P(t_1,t_2)$ under the measure $\mathbb{F}$, conditionally to the filtration $\mathcal{F}_s$, is given by*

$$\mathbb{E}^{\mathbb{F}}\left(e^{\omega \ln P(t_1,t_2)}|\mathcal{F}_s\right) = \qquad (58)$$
$$\exp\left(-\sum_{j=1}^d \int_s^{t_1}\psi_j\left(-h_j(u,t_1)\right)du - \omega\int_{t_1}^{t_2}\varphi(u)du + \omega\sum_{j=1}^d\int_{t_1}^{t_2}\psi_j\left(-h_j(u,t_2)\right)du\right)$$
$$\times\exp\left(-\omega\sum_{j=1}^d\int_0^\infty Y_s^{(j,\xi)}e^{-\xi(t_1-s)}B^{(\xi)}(t_1,t_2)\,d\gamma_j(\xi)\right)$$
$$\exp\left(\sum_{j=1}^d\int_s^{t_1}\psi_j\left(-\int_0^\infty\frac{1}{\xi}\left(1-(1+\omega)\,e^{-\xi(t_1-u)}+\omega e^{-\xi(t_2-u)}\right)d\gamma_j(\xi)\right)du\right).$$

Proof. The mgf can be rewritten as an expectation under the risk neutral measure using the Bayes' rule

$$\mathbb{E}^{\mathbb{F}}\left(e^{\omega \ln P(t_1,t_2)}|\mathcal{F}_s\right) \quad = \quad \frac{\mathbb{E}\left(\left.\frac{d\mathbb{F}}{d\mathbb{Q}}\right|_{t_1} e^{\omega \ln P(t_1,t_2)}|\mathcal{F}_s\right)}{\mathbb{E}\left(\left.\frac{d\mathbb{F}}{d\mathbb{Q}}\right|_{t_1}|\mathcal{F}_s\right)}, \tag{59}$$

where $\left.\frac{d\mathbb{F}}{d\mathbb{Q}}\right|_{t_1}$ is detailed in Equation (53). From Equation (24) and due to the independence between the processes $Y_{t_1}^{(j,\xi)}$ and $Y_{t_1}^{(k,\xi)}$ for $j \neq k$, the expectation (59) becomes

$$\begin{aligned}
&\mathbb{E}^{\mathbb{F}}\left(e^{\omega \ln P(t_1,t_2)}|\mathcal{F}_s\right) = \\
&\exp\left(-\sum_{j=1}^{d}\int_s^{t_1}\psi_j\left(-h_j(u,t_1)\right)du-\omega\int_{t_1}^{t_2}\varphi(u)du+\omega\sum_{j=1}^{d}\int_{t_1}^{t_2}\psi_j\left(-h_j(u,t_2)\right)du\right)\\
&\times\prod_{j=1}^{d}\mathbb{E}\left(\exp\left(-\int_s^{t_1}h_j(u,t_1)\,dL_u^{(j)}-\omega\int_0^{\infty}Y_{t_1}^{(j,\xi)}B^{(\xi)}(t_1,t_2)\,d\gamma_j(\xi)\right)\mid\mathcal{F}_s\right).
\end{aligned}$$

From Equation (18) and after a change of integration order, the integral of processes $Y_{t_1}^{(j,\xi)}$ is :

$$\begin{aligned}
&\int_0^{\infty}Y_{t_1}^{(j,\xi)}B^{(\xi)}(t_1,t_2)\,d\gamma_j(\xi)\\
&=\int_0^{\infty}Y_s^{(j,\xi)}e^{-\xi(t_1-s)}B^{(\xi)}(t_1,t_2)\,d\gamma_j(\xi)+\int_0^{\infty}\int_s^{t_1}e^{-\xi(t_1-u)}dL_u^{(j)}B^{(\xi)}(t_1,t_2)\,d\gamma_j(\xi)\\
&=\int_0^{\infty}Y_s^{(j,\xi)}e^{-\xi(t_1-s)}B^{(\xi)}(t_1,t_2)\,d\gamma_j(\xi)+\int_s^{t_1}\int_0^{\infty}e^{-\xi(t_1-u)}B^{(\xi)}(t_1,t_2)\,d\gamma_j(\xi)\,dL_u^{(j)}.
\end{aligned}$$

This allows us to rewrite $\mathbb{E}^{\mathbb{F}}\left(e^{\omega \ln P(t_1,t_2)}|\mathcal{F}_s\right)$ as follows:

$$\begin{aligned}
&\mathbb{E}^{\mathbb{F}}\left(e^{\omega \ln P(t_1,t_2)}|\mathcal{F}_s\right) = \\
&\exp\left(-\sum_{j=1}^{d}\int_s^{t_1}\psi_j\left(-h_j(u,t_1)\right)du-\omega\int_{t_1}^{t_2}\varphi(u)du+\omega\sum_{j=1}^{d}\int_{t_1}^{t_2}\psi_j\left(-h_j(u,t_2)\right)du\right)\\
&\qquad\times\exp\left(-\omega\sum_{j=1}^{d}\int_0^{\infty}Y_s^{(j,\xi)}e^{-\xi(t_1-s)}B^{(\xi)}(t_1,t_2)\,d\gamma_j(\xi)\right)\\
&\times\prod_{j=1}^{d}\mathbb{E}\left(\exp\left(-\int_s^{t_1}\left(h_j(u,t_1)-\omega\int_0^{\infty}e^{-\xi(t_1-u)}B^{(\xi)}(t_1,t_2)\,d\gamma_j(\xi)\right)dL_u^{(j)}\right)\mid\mathcal{F}_s\right).
\end{aligned}$$

By definition of $h_j(u,t) := \int_0^{\infty}B^{(\xi)}(u,t)\,d\gamma_j(\xi)$, we have that

$$\begin{aligned}
&h_j(u,t_1)-\omega\int_0^{\infty}e^{-\xi(t_1-u)}B^{(\xi)}(t_1,t_2)\,d\gamma_j(\xi) = \\
&\int_0^{\infty}\left(B^{(\xi)}(u,t_1)-\omega e^{-\xi(t_1-u)}B^{(\xi)}(t_1,t_2)\right)d\gamma_j(\xi),
\end{aligned}$$

wherein the integrand is equal to

$$\begin{aligned}
&B^{(\xi)}(u,t_1)-\omega e^{-\xi(t_1-u)}B^{(\xi)}(t_1,t_2)\\
&\quad=\frac{1}{\xi}\left(1-(1+\omega)\,e^{-\xi(t_1-u)}+\omega e^{-\xi(t_2-u)}\right).
\end{aligned}$$

Combining these three last equations lead to the result. □

The expression (58) of the mgf involves integrals of $Y_s^{(j,\xi)}$ with respect to $\xi \in \mathbb{R}^+$. These integrals do not admit analytical formulas and are in practice approached with the discretization scheme introduced in Section 6. Let n be the number of atoms of the discrete approximation of γ_j for $j = 1, ..., d$. To lighten future developments, we adopt the following notations

$$\begin{aligned} h_j^{(n)}(u,t) &:= \sum_{k=1}^{n} \frac{m_k^{(j)}}{b_k^{(j)}} \left(1 - e^{-b_k^{(j)}(t-u)}\right), \\ g_j^{(n)}(\omega,u,t_1,t_2) &:= \sum_{k=1}^{n} \frac{m_k^{(j)}}{b_k^{(j)}} \left(1 - (1+\omega)\, e^{-b_k^{(j)}(t_1-u)} + \omega e^{-b_k^{(j)}(t_2-u)}\right), \end{aligned}$$

for $j = 1, ..., d$. The discretized version of the mgf of the bond log-return under the forward measure is as follows

$$\begin{aligned} &\mathbb{E}^{\mathbb{F}}\left(e^{\omega \ln P^{(n)}(t_1,t_2)} | \mathcal{F}_s\right) = \\ &\exp\left(\sum_{j=1}^{d} \int_s^{t_1} \psi_j\left(-g_j^{(n)}(\omega,u,t_1,t_2)\right) - \psi_j\left(-h_j^{(n)}(u,t_1)\right) du\right) \\ &\times \exp\left(\int_{t_1}^{t_2} \omega \sum_{j=1}^{d} \psi_j\left(-h_j^{(n)}(u,t_2)\right) - \omega\, \varphi^{(n)}(u)\, du\right) \\ &\times \exp\left(-\sum_{j=1}^{d}\sum_{k=1}^{n} m_k^{(j)} Y_s^{(j,k)} e^{-b_k^{(j)}(t_1-s)} B^{(j,k)}(t_1,t_2)\right), \end{aligned}$$

where $\varphi^{(n)}(t)$, the discretized version of $\varphi(t)$, is detailed in Equation (37). The integrals with respect to times are numerically computed e.g. with a Simpson's rule. Let us denote by $f^{(n)}(x)$, the probability density of $\ln P^{(n)}(t_1,t_2)$ conditionally to $\mathcal{F}_s$ and

$$\Upsilon^{(n)}(i\omega) = \mathbb{E}^{\mathbb{F}}\left(e^{i\,\omega \ln P^{(n)}(t_1,t_2)} \mid \mathcal{F}_s\right) = \int e^{i\omega x} f^{(n)}(x) dx,$$

its Fourier's transform. The probability density function can therefore be expressed as the real part of the inverse Fourier's transform:

$$\begin{aligned} f^{(n)}(x) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Upsilon^{(n)}(i\omega)\, e^{-i\,x\,\omega} d\omega \\ &= \frac{1}{\pi} Re\left(\int_0^{+\infty} \Upsilon^{(n)}(i\omega) e^{-i\,x\,\omega} d\omega\right), \end{aligned} \tag{60}$$

and is numerically computed with the algorithm detailed in the next proposition.

Proposition 10. *Let M be the number of steps used in the Discrete Fourier Transform (DFT) and $\Delta_x = \frac{2x_{max}}{M-1}$ be this step of discretization. Let us denote $\Delta_\omega = \frac{2\pi}{M\,\Delta_x}$ and*

$$\omega_m = (m-1)\Delta_\omega,$$

for $m = 1...M$. The values of $f^{(n)}(x)$ the pdf of $\ln P^{(n)}(t_1,t_2)|\mathcal{F}_s$ at points $x_k = -\frac{M}{2}\Delta_x + (k-1)\Delta_x$ are approached by the sum:

$$f^{(n)}(x_k) \approx \frac{2}{M\,\Delta_x} Re\left(\sum_{m=1}^{M} \varrho_m \Upsilon^{(n)}(i\,\omega_m)\,(-1)^{m-1} e^{-i\frac{2\pi}{M}(m-1)(k-1)}\right). \tag{61}$$

where $\varrho_m = \frac{1}{2}1_{\{m=1\}} + 1_{\{m\neq 1\}}$.

The proof of this result is based on the trapezoidal approximation of the integral in Equation (60).

Figure 9 shows the probability density functions of $P^{(n)}(5,10)$ under the forward measure of numeraire $P^{(n)}(0,5)$. We consider for this illustration a one factor Lévy model ($d=1$) in which the driving process is a jump-diffusion such as presented in Equation (48), with the same parameters as those of Section 6. The DFT parameters are $x_{max} = 0.2$ and $M = 2^{10}$. The plots reveals that variances under $\mathbb{F}$ of bond prices display the same sensitivity to the memory parameter α as under the risk neutral measure. The PML and ML bond variances are respectively directly and inversely proportional to α.

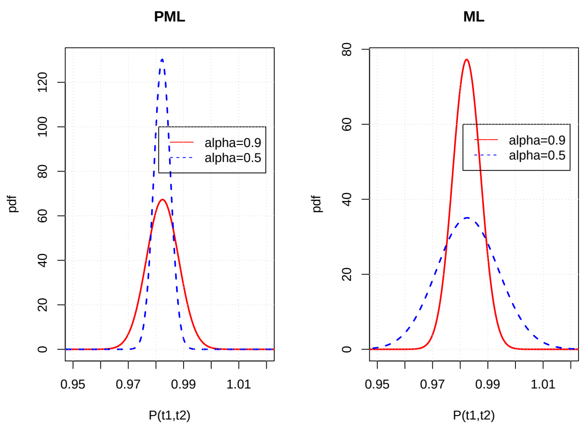

Figure 9. Probability density function of $P^{(n)}(5,10)$ under the forward measure of numeraire $P^{(n)}(0,5)$, α =0.5 and 0.9, $\beta = 1.5$, λ_1=0.5 η_1= -0.002, $\sigma_1 = 0.01$. ML and PML kernels with n =40 atoms.

After computation of the log-bond density, the call price is calculated by the following sum:

$$\begin{aligned} C(s) &= P(s,t_1)\mathbb{E}^{\mathbb{F}}\left(\left(P(t_1,t_2)-K\right)_+ \mid \mathcal{F}_s\right) \\ &\approx P(s,t_1)\sum_{k=1}^{M} \mathbf{1}_{\{x_k\geq \ln K\}}\left(e^{x_k}-K\right) f^{(n)}(x_k)\,\Delta_x\,, \end{aligned}$$

where $\mathbf{1}_{\{x_k\geq \ln K\}}$ is an indicator variable equal to one on $[\ln L, +\infty)$ and zero otherwise. Figure 10 shows call prices on $P^{(n)}(5,10)$ for various strike prices in the ML and PML kernels. In the PML model, increasing the memory parameter α drives up the option values whatever the strike price. For the ML kernel, we observe the opposite trend. This is relevant with our previous conclusions about the variance that is the main driving factor of option prices.

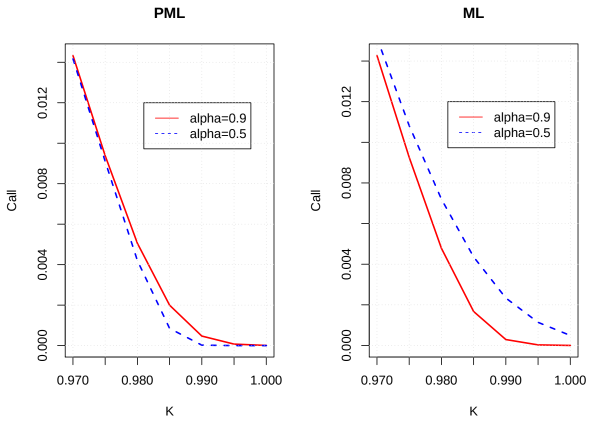

Figure 10. Call prices on $P^{(n)}(5,10)$, α =0.5 and 0.9, β = 1.5, λ_1=0.5 η_1= -0.002, σ_1 = 0.01. ML and PML kernels with n =40 atoms.

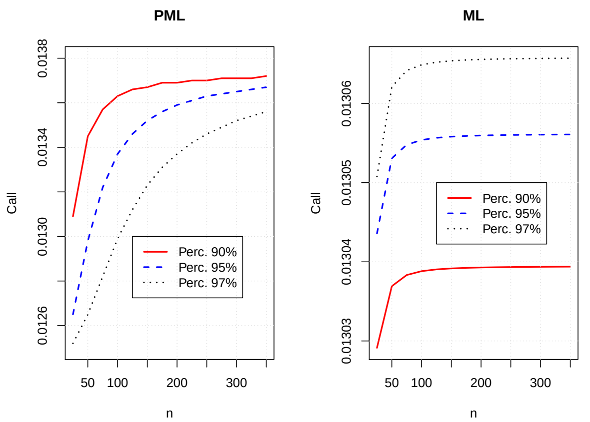

Figure 11. Call prices on $P^{(n)}(5,10)$ in function of n, α =0.5 and 0.9, β = 1.5, λ_1=0.5 η_1= -0.002, $\sigma_1 = 0.01$. Strike= 0.97.

Figure 11 allows us to confirm the convergence of the call option when the number, n, of atoms in the discretization scheme growths. We recall that the partition ranges from $\xi_0 = 0$ up to ξ_n that is a percentile of $\gamma_j(z)$. Here percentiles 90%, 95% and 97% are considered. For the ML kernel (right plot), the convergence is quick and the difference between prices computed with the 90% and 97% is negligle (around 2.6e-5). For the PML model (left plot), prices are higher (and then conservative) with a 90% percentile and converge with less atoms than with a partition covering 97% of $\gamma_j(.)$.

8. Conclusions

This article contributes to the literature on interest rate models by proposing an alternative to standard mean reverting processes. The later have an exponential decaying memory of the interest rate sample path. Motivated by observations of Section 3, our approach instead considers models in which the memory kernels are sub-exponential functions. In this setting, the short-term rate remembers its recent history for a longer period than in the exponential framework.

Even if the interest rate process is not Markov anymore, the model preserves a high level of analytical tractability by taking advantage of the Laplace-Stieltjes representation of the memory function. The short-term rate can indeed be reformulated as an infinite dimensional Markov process. In this setting, bond prices admit closed form expressions and a discretization scheme allows to infer their dynamics. Furthermore, Section 7 proposes a procedure based on a change of numeraire to evaluate European bond options. The numerical analysis underlines the differences between models with Mittag-Leffler (ML) or power Mittag-Leffler (PML) memory kernels. The main one being that standard deviations of bond yields are respectively directly and inversely proportional to the parameter α.

This article paves the way of further research. For instance, the model could be adapted to a multi-curves framework or could serve as basis for a deeper empirical analysis. Other possible extensions are the modulation of parameters by a Markov chain to capture economic regimes or the introduction of clustered jumps.

Appendix A

When the memory kernel is exponentially decreasing, i.e. $g_j(t-u) = e^{-\kappa_j(t-u)}$, the $\left(X_t^{(j)}\right)_{t\geq 0}$'s are Ornstein-Uhlenbeck (OU) processes defined by

$$X_t^{(j)} = e^{-\kappa_j(t-s)} X_s^{(j)} + \int_s^t e^{-\kappa_j(t-u)} dL_u^{(j)} \quad , j = 1, ..., d. \tag{62}$$

Their differential is such that

$$dX_t^{(j)} = -\kappa_j X_t^{(j)} dt + dL_t^{(j)} \quad , j = 1, ..., d \tag{63}$$

and emphasizes that $\left(X_t^{(j)}\right)_{t\geq 0}$ processes revert to zero at a speed κ_j. In this particular case, the influence on $X_t^{(j)}$ of the past sample paths of $L_t^{(j)}$ decays in an exponential manner. If the short term rate is the sum of a deterministic function $\varphi(t)$ and of $X_t^{(j)}$ for $j = 1, ..., d$, as formulated in Equation (5), the conditional variance of the short-term rate is a function of the time horizon:

$$\mathbb{V}\left(r_t | \mathcal{F}_s\right) = \sum_{j=1}^{d} \left(\sigma_j^2 + \int_{\mathbb{R}} z^2 \nu_j(dz)\right) \frac{\left(1 - e^{-2\kappa_j(t-s)}\right)}{2\kappa_j}, \tag{64}$$

whereas the autocovariance of r_t and r_u for $t \geq u \geq s$ is given by:

$$\mathbb{C}\left(r_t\, r_u | \mathcal{F}_s\right) = \sum_{j=1}^{d} \left(\sigma_j^2 + \int_{\mathbb{R}} z^2 \nu_j(dz)\right) \frac{e^{-\kappa_j(t-u)} - e^{-\kappa_j(t+u-2s)}}{2\kappa_j}.$$

This last equation reveals that the autocovariance between the current short rate and the future one decays exponentially. Notice that the covariance between r_t and $r_{t-\Delta}$ for any $\Delta > 0$ converges to a function of Δ when $t \to \infty$:

$$\lim_{t\to\infty} \mathbb{C}\left(r_t\, r_{t-\Delta} | \mathcal{F}_0\right) = \sum_{j=1}^{d} \left(\sigma_j^2 + \int_{\mathbb{R}} z^2 \nu_j(dz)\right) \frac{e^{-\kappa_j \Delta}}{2\kappa_j}. \tag{65}$$

Appendix B

The Nelson-Siegel model postulates the following representation of the curve of instantaneous forward rates:

$$f(0,t) = b_0 + (b_{10} + b_{11} t) \exp\left(-c_1 t\right), \tag{66}$$

the parameters b_0, b_{10} and b_{11} respectively determine the mean level, the slope and the curvature of the forward rates curve. The zero-coupon rate in the Nelson-Siegel model (66) is obtained by integrating these forward rates and by dividing the result by the maturity:

$$y_m(t) \quad = \quad b_0 + \frac{1}{t}\frac{b_{10}}{c_1}\left(1 - e^{-c_1 t}\right) + \frac{1}{t}\frac{b_{11}}{(c_1)^2}\left(1 - (c_1 t + 1)\, e^{-c_1 t}\right) . \tag{67}$$

Parameters are estimated by minimizing the quadratic spread between market and model zero-coupon yields. In the numerical illustration, this model is fitted to ICE Euro swap rates on the 26/2/21 at 12 o'clock, presented in Table 2. The corresponding Nelson-Siegel parameters are reported in Table 3.

Maturity	Swap rates (%)	Maturity	Swap rates (%)
1	-0.466	8	-0.027
2	-0.419	9	0.036
3	-0.363	10	0.148
4	-0.298	12	0.275
5	-0.229	15	0.376
6	-0.160	20	0.393
7	-0.092	25	0.373

Table 2. ICE Euro swap rates on the 26/2/21, 12:00. Source: https://fred.stlouisfed.org/

Parameters	Values
b_0	0.00504905
b_{10}	-0.00892662
b_{11}	-0.00350623
c_1	0.29428630

Table 3. Nelson-Siegel parameters for the ICE Euro swap rates on the 26/2/21.

Acknowledgment

The author is grateful to the "Fonds de la Recherche Scientifique - FNRS" for financial support under Grant number 33658713.

Applebaum D. Lévy processes and stochastic calculus. Cambridge studies in advanced mathematics.2004, Cambridge University Press.

Baleanu D., Diethelm K., Scalas E., Trujillo J.J. Fractional calculus: models and numerical methods. Series on Complexity, Nonlinearity and Chaos, Vol 3, World Scientific, 2012.

Boero G., Torricelli C. A comparative evaluation of alternative models of the term structure of interest rates. European Journal of Operational Research, 1996, 93 (1), p. 205-223.

Brace A., Gatarek M., Musiela M. The market model of interest rate dynamics. Mathematical Finance, 1997, 7 (2), p. 127-147.

Brigo D., Mercurio F. Interest Rate Models: Theory and Practice: With Smile, Inflation. 2006, Springer-Verlag Berlin.

Cheridito P., Kawaguchi H., Maejima M. Fractional Ornstein-Uhlenbeck processes, 2003, 8, p. 1-14.

Cont R., Tankov P. Financial modelling with jump processes. Chapman & Hall CRC, 2003, financial mathematics series.

Cousin A., Maatouk H., Rullière D. Kriging of financial term-structures. European Journal of Operational Research, 2016, 255 (2), p. 631-648.

Eberlein E. Raible S. Term structure models driven by general Lévy processes. Mathematical Finance, 1999, 9, p. 31-53.

Eberlein E. Kluge W. 2005. Exact pricing formulae for caps and swaptions in a Lévy term structure model. Journal of Computational Finance 9 (2), p. 99-125.

Erdélyi A., Magnus W., Oberhettinger F., Tricomi F. Higher Transcendental Functions; McGraw-Hill: New York, NY, USA, 1955, Volume 3.

Falini, J. Pricing caps with HJM models: The benefits of humped volatility. European Journal of Operational Research, 2010, 207 (3), p. 1358-1367

Fontana C., Gnoatto A., Szulda G. Multiple yield curve modelling with CBI processes. Mathematics and Financial Economics, 2020, https://link.springer.com/article/10.1007/s11579-020-00289-4

Golinski A., Zaffaroni P. Long memory affine term structure models. Journal of Econometrics, 2016, 191 (1), p.33-56.

Gorenflo R., Kilbas A.A., Mainardi F. and Rogosin S.V. Mittag-Leffler Functions, Related Topics and Applications. Second Edition, 2020, Springer-Verlag Berlin Heidelberg.

Gorenflo R., Mainardi F. Fractional Calculus. In: Carpinteri A., Mainardi F. (eds) Fractals and Fractional Calculus in Continuum Mechanics. International Centre for Mechanical Sciences (Courses and Lectures), 1997 vol 378. Springer.

Hainaut D., Macgilchrist R. An interest rate tree driven by a Lévy process. Journal of derivatives 2010, 18 (2), p. 33-45.

Hainaut D. A model for interest rates with clustering effects. Quantitative Finance, 2016, 16(8), p. 1203-1218

Hainaut D. A fractal version of the Hull-White interest rate model. Economic Modelling, 2013, 31, p 323-334.

Haubold H.J., Mathai A.M, Saxena R.K. Mittag-Leffler Functions and Their Applications. 2011Journal of Applied Mathematics, 2011, Article ID 298628.

Heath D., Jarrow R., Morton, A. Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. Econometrica, 1992, 60 (1), p. 77-105.

Hull J. and White A. Pricing interest rate derivative securities. Review of financial studies, 1990, 3 (4), p. 573-592.

Kochubei A., Luchko Y. Handbook of fractional calculus with applications. Volume 1: basic theory. 2019, De Gruyter, Berlin.

Li H., Ye X., Yu F. Unifying Gaussian dynamic term structure models from a Heath–Jarrow–Morton perspective. European Journal of Operational Research, 2020, 286 (3), p. 1153-1167

Maejima M., Yamamoto K. Long-Memory Stable Ornstein-Uhlenbeck Processes, Electronic Journal of Probability, 2003, 8, p. 1-18.

Mainardi F. Why the Mittag-Leffler Function Can Be Considered the Queen Function of the Fractional Calculus? Entropy, 2020, 22, 1359.

Mercurio F., Moraleda J. An analytically tractable interest rate model with humped volatility. European Journal of Operational Research, 2000, 120 (1), p.109-121.

Moreno M., Platania F. A cyclical square-root model for the term structure of interest rates. European Journal of Operational Research, 2015, 241, p.109-121.

Njike Leunga C., Hainaut D. Interbank credit risk modelling with self-exciting jump processes. International Journal of Theoretical and Applied Finance, 2020, 23 (6).

Schmidt W.M. Interest rate term structure modelling. European Journal of Operational Research, 2011, 214 (1), p.1-14.