Institut de Statistique, Biostatistique et Actuariat

Pricing of derivatives with memory.

Membres du jury:

Prof. Donatien Hainaut, *Promoteur*
Prof. Jérôme Barbarin

Mémoire présenté en vue de
l'obtention du mastère
en sciences actuarielles
(orientation sciences actuarielles)
par:

Maxime Dom

Louvain-La-Neuve
Août 2021

Preface.

This dissertation has been prepared in partial fulfillment of the requirements for the master degree in actuarial sciences delivered by the Université Catholique de Louvain.

Acknowledgment.

I would like to express my gratitude to all the people who supported me throughout the writing of this master thesis and especially to Mr Donatien Hainaut and Mr Charles Guy Njike Leunga for their valuable advice and technical contributions. Finally, I would like to thank Mr Jérôme Barbarin for the time he will spend on the review of this master thesis.

Contents

Chapter 1

Introduction

This thesis presents an overview of the pricing of (long and short) memory processes describing the commodity prices over time. These processes go beyond the rigid framework usually encountered in the literature since they require a relaxation of the strong Markov hypothesis. Although it makes the study more complex to approach, this generalisation brings it closer to the situations observed on the market in practice. This thesis takes its place in the continuation of the article developed by Benth and al. (2020) and is unique in the sense that this latter does not explore a practical case, which will be a substantial part of this framework.

After introducing the theoretical basis, i.e. the stochastic processes, the Lévy processes framework and the various properties that they are based on, we will pass to the main part of the thesis. We will use the generalized Langevin equation used by Benth (2020) to introduce models to describe the price dynamics of the underlying asset. To do so, Takahashi (1996) proposes different memory kernels with specific properties for each of them: a kernel with a short (instantaneous) memory which corresponds to a particular case of a well-known OU process, a negative Exponential kernel (i.e. rapid attenuation of the past history) and a negative Power kernel (i.e. strong residual of the past history) which, as we will see, shares some properties with fractional calculus. This area, which has been widely used in recent years to describe long time-dependent models, will not be addressed in this thesis in order to focus on processes that introduce a separate memory term into the process requiring the use of the Laplace-Fourier transform framework.

In a second step, we introduce the new change of measure that will be necessary in the practical part and in particular for the pricing of derivatives. Indeed, since we deal with memory processes that describe commodity prices, the market is no longer complete and that the risk-neutral probability measure is therefore no longer unique, as we shall see. To solve this problem, Benth (2020) introduces a risk measure that deviates from the risk-neutral one $\mathbb{Q}$, in the sense that it is necessary to introduce a new risk premium that could be associate to the convenience yield. Benth (2020) proposes to use memory terms to describe this. We also introduce the last theoretical points needed for the practical part such as implied volatility and the different numerical algorithms for pricing derivatives (i.e. Carr & Madan (1999)).

Finally, in the third and last part, we introduce the practical part which includes the different points discussed earlier. First, a discussion of commodity markets to confirm the presence of memory in prices. Then, a calibration of the different models on the four markets studied (Gold, Corn, Oil, Copper) to compare these different models and these different markets between them. Finally, we will discuss the pricing of derivatives, both vanilla and exotic options, on the different models used and analyse the sensitivity of the price of these options to the different models and parameters.

Chapter 2

Lévy Processes

2.1 Preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with usual conditions, with Ω, a sample space, $\mathcal{F}$, a sigma-algebra on Ω, $\mathbb{P}$, a probability measure on $(\Omega, \mathcal{F})$. An information flow on this probability space is an increasing family of sigma-algebra $(\mathcal{F}_t)_{0 \leq t \leq T}$ such as :

$$\forall\ 0 \leq s \leq t : \mathcal{F}_0 \subseteq \mathcal{F}_s \subseteq \mathcal{F}_t \subseteq \mathcal{F}$$

A function $f : [0, T] \to \mathbb{R}^d$ is càdlàg if it is right-continuous with existing left limits :

$$\forall\ t \in [0, T] : f(t^-) = \lim_{\substack{s \to t \\ s < t}} f(s) \quad \text{and} \quad f(t^+) = \lim_{\substack{s \to t \\ s > t}} f(s)$$

both exist and also that $f(t) = f(t^+)$. If the function includes a discontinuity at time t, this is called a « jump » and is noted as follows: $\Delta f(t) = f(t) - f(t^-)$.

2.2 Characteristic Function

The characteristic function (or Fourier-Stieltjes Transform) $\Phi : \mathbb{R} \to \mathbb{C}$ of a random variable X is the Fourier transform of its cumulative distribution function, $F_X(x) = \mathbb{P}(X \leq x)$, such that

$$\Phi_{X_t}(u) = \mathbb{E}[\exp(iuX)] = \int_{-\infty}^{\infty} \exp(iux) \mathrm{d}F_X(x), \tag{2.2.1}$$

where $i \in \mathbb{C}$, the imaginary unit. If the density function $f_X(x)$ is available, we can rewrite as

$$\Phi_{X_t}(u) = \mathbb{E}[\exp(iuX)] = \int_{-\infty}^{\infty} \exp(iux) f_X(x) \mathrm{d}x. \tag{2.2.2}$$

This characteristic function always exists, is uniformly continuous on the entire space and characterized uniquely the distribution function of X.

2.3 Definition

A càdlàg, right-continuity and left limits, stochastic process $(X_t)_{t \geq 0}$ with $X_0 = 0$, is said to be a Lévy process if it satisfied the following set of properties (Tankov, 2003) :

- has **independent increments** with respect to an increasing sequence of times, $t_0, t_1, ..., t_n$, such that $X_{t_0}, X_{t_1} - X_{t_0}, ..., X_{t_n} - X_{t_{n-1}}$ are independent,
- has **stationary increments** such that the law of $X_t - X_s$ is the same as the law of $X_{t-s} - X_0$, $\forall\ 0 \leq s \leq t$,
- has **stochastic continuity** such that $\forall\ \epsilon > 0, \lim_{s \to 0} \mathbb{P}(|X_{t+s} - X_t| \geq \epsilon) = 0$.

The latter does not imply that the path of the function is continuous but rather that the probability of observing a jump at a particular time is zero and allows to exclude the jump process with nonrandom times (e.g. calendar effects).

Let then introduce $(L_t)_{0\leq t\leq T}$ be a Lévy process with the characteristic triplet $(\mu,\, \sigma^2,\, \nu)$, where $\nu(A)$ is the Lévy measure for the set $A \subset \mathbb{R}_0$ and $\nu(A) = \lambda \times f(A)$ with λ, the intensity jump and $f(.)$, the jump size distribution. Assume that the Lévy measure of the L-process admits that $\int_{|z|\leq\epsilon} z^2\nu(\mathrm{d}z) < \infty$, i.e. L is a square-integrable Lévy Process, where $\epsilon \in \mathbb{R}^+$ (Applebaum, 2004). From this triplet composition, we can admit that L is a process which assume the following form

$$L_t = \mu t + \sigma W_t + \int_0^t \int_{\mathbb{R}} z\mathbb{1}_{\{|z|\leq\epsilon\}}\tilde{N}(\mathrm{d}s \times \mathrm{d}z) + \int_0^t \int_{\mathbb{R}} z\mathbb{1}_{\{|z|>\epsilon\}}N(\mathrm{d}s \times \mathrm{d}z), \qquad 0 \leq t \leq T, \tag{2.3.1}$$

where W_t is a standard Brownian motion and $\tilde{N}$ is a compensated poisson random measure such that $\tilde{N}(\mathrm{d}t \times \mathrm{d}z) = N(\mathrm{d}t \times \mathrm{d}z) - \nu(\mathrm{d}z)\mathrm{d}t$, where the latter term is $\mathbb{E}[N(\mathrm{d}t \times \mathrm{d}z)] = \nu(\mathrm{d}z)\mathrm{d}t$. Using a discrete vision of the jumps, N_t corresponds to the number of jumps with mean $\lambda \times t$, J_k the intensity of the jumps and $\mathbb{E}[J_k] = \kappa$. Note that J_k is an *iid* random variable and that all the sources of randomness are independent. It can be noticed that according to the Lévy Itô decomposition, each L_t can be computed as the sum of three components : a deterministic drift μt, a diffusion of variance σ^2 and a jump process $N(t, z)$ with an intensity of $\nu(\mathrm{d}z)$ defined on $[-\infty, 0) \cup (0, +\infty]$. This latter term is such that the probability of observing k jumps between the interval time $[t_1, t_2]$ of a size included in a set $B \in \mathbb{R}_0$ is

$$\mathrm{P}\Big[N([t_1, t_2] \times B) = k\Big] = \exp\Big(- \int_{t_1}^{t_2} \int_B \nu(\mathrm{d}z)\mathrm{d}t\Big) \times \frac{\Big(\int_{t_1}^{t_2}\int_B \nu(\mathrm{d}z)\mathrm{d}t\Big)^k}{k!}. \tag{2.3.2}$$

Concerning the pathwise properties of the Lévy process, this latter could be of finite/infinite variation or of finite/infinite activity. Let us make the difference between the two.

A lévy process is of **finite variation** if and only if its Lévy triplet satisfied (μ, σ, ν) :

$$\sigma = 0 \qquad \text{and} \qquad \int_{\mathbb{R}} (|x| \wedge 1)\nu(\mathrm{d}x) < +\infty, \tag{2.3.3}$$

As a consequence, a process of finite variation has no Brownian components since if $\sigma \neq 0$ or $\int_{\mathbb{R}}(|x| \wedge 1)\nu(\mathrm{d}x) = +\infty$, then the Lévy process is of infinite variation.

Some Lévy processes have **infinite activity**, i.e. $\nu(\mathbb{R}) = \infty$, in the sense that they jump at every moment, no matter how close we « zoom » in, thus becoming discontinuous involving that any time interval, even very small, contains an infinite number of jumps almost surely. In this case, this jump term must be compensated by its average otherwise the sum of these jumps (i.e. the integral of these jumps) will converge to infinity. Often, the truncation of the jumps at 1 is made (as we will see in the Lévy-Khintchine formula) but this choice is mainly conventional and one can use the threshold $\epsilon > 0$ to be more generic.

Defining $\varrho = \mu + \int_{|z|\geq\epsilon} z\nu(\mathrm{d}z)$, one can see that ϱ correspond to $\mathbb{E}[L_1]$ since the second and third terms after equality are martingales and have a zero mean in equation (2.3.1). It follows easily from this that L_t is a martingale if $\varrho = \mathbb{E}[L_1] = 0$, a submartingale if $\varrho > 0$ and a supermartingale if $\varrho < 0$. If the condition : $\int_{|z|\geq\epsilon} |z|^n\nu(\mathrm{d}z) < \infty$ with $n = \{1, 2\}$, is fulfilled, then the first two moments of the Lévy process L_t are :

$$\mathbb{E}[L_t] = \varrho t = \Big(\mu + \int_{|z|\geq\epsilon} z\nu(\mathrm{d}z)\Big)t \quad \text{and} \quad \mathbb{V}[L_t] = \Big(\sigma^2 + \int_{\mathbb{R}} z^2\nu(\mathrm{d}z)\Big)t. \tag{2.3.4}$$

2.4 Lévy-Khintchine Formula

A probability function F is said to be infinitely divisible if for every integer $n \geq 2$ there exists n independent and identically distributed random variables $Y_1, Y_2, ..., Y_n$ such that, the sum of these has a distribution F. Then, the characteristic function of an infinitely divisible distribution is called the infinitely divisible characteristic function. The characteristic function of the distribution F is also the n^{th} power of a characteristic function : $\Phi(u) = (\Phi_n(u))^n$.

Consider L_t a Lévy process, then $\forall t > 0$, its distribution is infinitely divisible with respect to the definition introduced previously[1]. This constrains the number of possible distributions : Gaussian, Poisson, Gamma, etc. Thanks to this, the characteristic function of L_t is generally described by the Lévy-Khintchine formula :

$$\Phi_{L_t}(u) := \mathbb{E}\Big[e^{iuL_t}\Big] = e^{t\left(\mu iu - \frac{\sigma^2 u^2}{2} + \int_{\mathbb{R}\setminus\{0\}}(e^{iux} - 1 - iux\mathbb{1}_{\{|x|<1\}})\nu(\mathrm{d}x)\right)}, \tag{2.4.1}$$

computed with the Lévy triplet (μ, σ^2, ν) where $\mu \in \mathbb{R}$, $\sigma^2 \geq 0$ and ν is the Lévy measure of F_L defined on $\mathbb{R}_0$ admitting

$$\int_{-\infty}^{\infty} \inf\{1, x^2\}\nu(\mathrm{d}x) < \infty. \tag{2.4.2}$$

The expression in (2.4.1) allows us to define the characteristic exponent of X as a continuous function $\psi(u) : \mathbb{R}^d \to \mathbb{R}$, such as :

$$\mathbb{E}[e^{iuX_t}] := e^{t\psi(u)}, \qquad u \in \mathbb{R}^d. \tag{2.4.3}$$

which corresponds, in (2.4.1), for a standard Lévy process as :

$$\psi(u) = iu\mu - \frac{\sigma^2 u^2}{2} + \int_{\mathbb{R}\setminus\{0\}} \big(e^{iux} - 1 - iux\mathbb{1}_{\{|x|<1\}}\big)\nu(\mathrm{d}x). \tag{2.4.4}$$

2.5 Quadratic Variation

Defining the partition $\Pi^n = \{t_0^n = 0 < t_1^n < ... < t_{n+1}^n = T\}$ on the interval $[0, T]$, such that $|\Pi^n| = \sup_k |t_k^n - t_{k-1}^n| \to 0$ when $n \to \infty$, the quadratic variation of the stochastic process X is given by :

$$[X, X]_t = \lim_{n\to\infty} \sum_{t_i \in \Pi^n}^{0 \leq t_i < t} (X_{t_{i+1}} - X_{t_i})^2. \tag{2.5.1}$$

By construction, we can see that $[X, X]_t$ is an increasing process which is null when X is continuous and has paths of finite variation. In particular, if X is a Lévy process with the corresponding characteristic triplet (μ, σ^2, ν), its quadratic variation process is such that :

$$[X, X]_t = \sigma^2 t + \sum_{\substack{s \in [0,t] \\ \Delta X_s \neq 0}} |\Delta X_s|^2 = \sigma^2 t + \int_{[0,t]\times\mathbb{R}} z^2 N(\mathrm{d}s \times \mathrm{d}z). \tag{2.5.2}$$

2.6 Law of Jumps

With regard to the distribution of the size of the jumps[2], Y, this is considered to follow an asymmetric double exponential law as proposed by Kou (2002) and as the density of Y takes the form:

$$f_Y(y) = p\eta_1 e^{-\eta_1 y}\mathbb{1}_{\{y\geq 0\}} + (1-p)\eta_2 e^{\eta_2 y}\mathbb{1}_{\{y<0\}}, \tag{2.6.1}$$

[1]Note that the opposite is also true.
[2]In the log-price.

where $\eta_1 > 1$ to ensure that the expectation of the asset price is finite and $\eta_2 > 0$. This choice has several advantages, including an asymmetric distribution of increments and heavier tails than in Merton's model, which is more in line with the reality of financial markets. Moreover, even if this distribution do not lead to an analytical form of the probability density of the Lévy process L_t, moments have one and also the moments of the jump. Thus, we note that

$$\mathbb{E}[Y] = \frac{p}{\eta_1} - \frac{(1-p)}{\eta_2} \qquad \mathbb{E}[Y^2] = \frac{2p}{\eta_1^2} + \frac{2(1-p)}{\eta_2^2}$$
$$\mathbb{E}[Y^3] = \frac{6p}{\eta_1^3} - \frac{6(1-p)}{\eta_2^3} \qquad \mathbb{E}[Y^4] = \frac{24p}{\eta_1^4} + \frac{24(1-p)}{\eta_2^4}$$

$$\mathbb{V}[Y] = p(1-p)\Big(\frac{1}{\eta_1} + \frac{1}{\eta_2}\Big)^2 + \Big(\frac{p}{\eta_1^2} + \frac{(1-p)}{\eta_2^2}\Big)$$

Kurtosis and skewness can be calculated using the different moments but we do not refer to them here due to their size and lack of interpretability. When the Lévy measure has no singularity at $z = 0$, i.e. when the number of jumps is a countable integer over a given time interval, which is the case for the double-exponential distribution, then we can set $\epsilon = 0$ such that we get

$$\begin{aligned} L_t &= \mu t + \sigma W_t + \int_0^t \int_{\mathbb{R}} z N(\mathrm{d}s \times \mathrm{d}z) \\ &= \mu t + \sigma W_t + \sum_{k=1}^{N_t} Y_k \end{aligned} \tag{2.6.2}$$

or in its differential form

$$\mathrm{d}L_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t + Y \mathrm{d}N_t. \tag{2.6.3}$$

Thanks to the Lévy-Khintchine Formula, we can compute the characteristic function of the process L_t :

$$\begin{aligned} \Phi_{L_t}(u) &= \mathbb{E}\Big[\exp\{iu(\mu t + \sigma W_t + \sum_{k=1}^{N_t} Y_k)\}\Big] \\ &= \exp\Big[iu\mu t - \frac{\sigma^2 u^2 t}{2}\Big] \exp\Big[\lambda t(\mathbb{E}[e^{iuY} - 1])\Big] \\ &= \exp\Big[iu\mu t - \frac{\sigma^2 u^2 t}{2} + \lambda t\Big\{\frac{p\eta_1}{\eta_1 - iu} + \frac{(1-p)\eta_2}{\eta_2 + iu} - 1\Big\}\Big] \end{aligned} \tag{2.6.4}$$

and according to the definition of the characteristic exponent in (2.4.3), we get :

$$\psi(u) = iu\mu - \frac{\sigma^2 u^2}{2} + \lambda\Big\{\frac{p\eta_1}{\eta_1 - iu} + \frac{(1-p)\eta_2}{\eta_2 + iu} - 1\Big\} \tag{2.6.5}$$

Kou (2002) lists various benefits of using this kind of jump distribution. One of the main reasons for the popularity of this distribution comes from behavioral finance. Indeed, it has been proven that the market tends to overreact and underreact to both good and bad news (e.g. Fama (1998)). Indeed, the double-exponential distribution offers both heavy tails (attributed to overreactions) and high peaks (attributed to underreactions). Thus, the fact that the model introduces this strong sensitivity to external news may lead to a leptokurtic distribution of returns. Some even interpret jumps in the price of the underlying asset only as a reaction of the market to external news.

Another important feature of this distribution is that it allows to reproduce different empirical elements observed in the market, on the one hand, the leptokurtic and asymmetric properties of returns as mentioned above and, on the other hand, the reproduction of the smile or the implied volatility skew observed in option prices.

On the contrary, jumps following a normal distribution in the returns as in the Merton model introduce less heavy tails, no asymmetry in the distribution, only a smile and not a skew in the implied volatility. However this kind of distribution allows nevertheless to obtain an approximation of the density of the complete Lévy process which can be very useful in the pricing of derivatives (e.g. the price of a call option can be seen as a weighted sum of several Black-Scholes options) which is not the case for the DED (double exponential distribution).

Chapter 3

Generalized Langevin Equation

Paul Langevin (1908) introduced the Langevin equation, well-known in physics and especially in molecular physics, which allows among other things to study the evolution of a system over time, on the one hand, from a macroscopic point of view with « slow relaxation » variables and on the other hand, from a microscopic point of view with « fast relaxation » variables. This second part corresponds to the stochastic nature of the equation. It takes the following form :

$$m\frac{\mathrm{d}x(t)}{\mathrm{d}t} = g(x(t), t) + \eta(t) \tag{3.0.1}$$

where $x(t)$ is a stochastic process and $\eta(t)$ a noise term. We can see in equation (3.0.1) that the changes in $x(t)$ depend on its own evolution but also on a random term. An extension of this equation is proposed by adding a path-dependent memory term :

$$m\frac{dx(t)}{\mathrm{d}t} = g(x(t), t) + \eta(t) + \int_{t_0}^{t} \zeta(x(u), u)M(t-u)\mathrm{d}u \tag{3.0.2}$$

where M is a weight function and ζ a memory function. This latter is a function of $x(t)$ and Yamada & Kunieda (1971) proposed to replace the function by $x(t)$ itself. This idea will be used in the following part of the report.

Therefore, let introduce $(S_t)_{0\leq t\leq T}$ the spot price in commodity market and X_t the log-return, such as

$$X_t := \log(S_t), \tag{3.0.3}$$

where $(X_t)_{0\leq t\leq T}$ is a generalized Langevin equation of the form

$$\mathrm{d}X_t = \beta(\theta - X_t)\mathrm{d}t + \Big(\int_0^t M(t-u)X_u\mathrm{d}u\Big)\mathrm{d}t + \chi(t-)\mathrm{d}L_t, \tag{3.0.4}$$

where the first term corresponds to the mean reverting part with β the speed of the reversion and θ the mean reversion level. M is a measurable function on $\mathbb{R}_+$ and χ is a deterministic càdlàg strictly positive function satisfying $\int_0^T \chi^2(s)\mathrm{d}s < \infty$. L_t is a Lévy process with the characteristic triplet (μ, σ^2, ν) and of the form as in (2.6.2). A comparison can be made between this expression and the equation (3.0.1) where $\eta(t) = \mathrm{d}L_t/\mathrm{d}t$.

In order to solve this SDE in (3.0.4), let introduce the Fourier-Laplace transform and its inverse denote by $\mathcal{L}[.]$ and $\mathcal{L}^{-1}[.]$, respectively. The Laplace Transform of a function $f(t)$ for all the real numbers $t \geq 0$, is defined by:

$$\mathcal{L}[f(t)] = \int_0^\infty f(t)e^{-st}\mathrm{d}t = F(s), \tag{3.0.5}$$

where s is a complex number such that, $s = \sigma + iw$, where σ and w are real numbers, see Widder (1945) for more details. So, starting from this expression of $F(s)$ we can find the function $f(t)$ thanks to the inverse of the Laplace transform which is defined as :

$$f(t) = \mathcal{L}^{-1}[F(s)] = \frac{1}{2\pi i} \lim_{T\to\infty} \int_{\gamma-iT}^{\gamma+iT} F(s)e^{st}\mathrm{d}s, \tag{3.0.6}$$

where γ is a real number so that the path of the integration boundary is in the convergence region of $F(s)$. Note that a function is uniquely determined by its Laplace Transform. Takahashi (1996) gave the following solution for the Laplace-Fourier transform of X_t :

$$\mathcal{L}[X_t] = \left[X_0 + \frac{\theta \times \beta}{s} + \chi(t-)\mathcal{L}\Big[\frac{\mathrm{d}L_t}{\mathrm{d}t}\Big]\right] H(s), \tag{3.0.7}$$

where

$$H(s) = \frac{1}{s - \mathcal{L}[M(s)] + \beta} \tag{3.0.8}$$

which allows to get an expression for X_t :

$$X_t = X_0\mathcal{L}^{-1}[H(s)] + \beta\theta \int_0^t \mathcal{L}^{-1}[H(s)]_{t-u}\mathrm{d}u + \int_0^t \chi(u-)\mathcal{L}^{-1}[H(s)]_{t-u}\frac{\mathrm{d}L_u}{\mathrm{d}u}\mathrm{d}u, \tag{3.0.9}$$

where the index $\mathcal{L}^{-1}[.]_{t-u}$ means that the Laplace transform are made with respect to $t-u$. At this point, the integrability of M is not yet necessary but the existence of $\mathcal{L}^{-1}[H(s)]$ is required (see Gripenberg and al, 1990) and the following assumption, for any $T > 0$,

$$\int_0^T \int_0^t \int_0^u M^2(t-u) \left(\mathcal{L}^{-1}\left[H(s)\right]_{v-u}\right)^2 \chi^2(v-)\ \mathrm{d}v\ \mathrm{d}u\ \mathrm{d}t < \infty. \tag{3.0.10}$$

At this stage, different orientations are possible depending on the choice of function M. In the following section, we will look at three different classes of functions for M and their implications for process memory.

We see that the SDE concerned in (3.0.4) belongs to the class of Volterra equations driven by a Lévy process. The collection of càdlàg functions on a given domain is known as Skorokhod space. Let us denote by $\mathbb{D}_{[0,T]}$ the Skorokhod space of càdlàg functions mapping $[0, T]$ to $\mathbb{R}$. Then, the solution of this SDE could be rewrite such that

$$\xi_t = \xi_0 + \int_0^t a(s, \xi)\mathrm{d}s + \int_0^t b(s-, \xi)\mathrm{d}L_s, \tag{3.0.11}$$

where ξ_0 is an $\mathcal{F}_0$-measurable random variable satisfying $P(|\xi_0| < \infty) = 1$ and $a(.,.)$, $b(.,.)$ are $\mathcal{B}([0,T]) \times \mathcal{B}(\mathbb{D}_{[0,T]})$-measurable functions. Protter (1985) shows that under some integrability conditions on the kernel M and its derivative thanks to the Hölder's inequality, the expression (3.0.4) has an unique solution.

3.1 Delta Function Kernel

The first kernel for the memory function is also the simplest: a function to introduce the direct influence of the past, i.e. the Dirac Delta function at zero of the form $M(t) = a\delta_0(t)$ such as

$$\begin{aligned}
\mathrm{d}X_t &= \beta(\theta - X_t)\mathrm{d}t + \Big(\int_0^t a\delta_0(t-u)X_u\mathrm{d}u\Big)\mathrm{d}t + \chi(t-)\mathrm{d}L_t \\
&= \beta(\theta - X_t)\mathrm{d}t + aX_t\mathrm{d}t + \chi(t-)\mathrm{d}L_t \\
&= (\beta - a)\Big(\frac{\beta\theta}{\beta - a} - X_t\Big)\mathrm{d}t + \chi(t-)\mathrm{d}L_t
\end{aligned} \tag{3.1.1}$$

A first very interesting observation is that by using this kernel which, let us point out, is nothing other than the instantaneous influence of the process, we fall back on a Markovian process in the sense that it satisfies the Markov property described in annex 9.2. Let introduce a process Y_t of the following form

$$Y_t = e^{(\beta-a)t}\Big(\frac{\beta\theta}{\beta-a} - X_t\Big), \tag{3.1.2}$$

or, in an equivalent form

$$\begin{aligned} \mathrm{d}Y_t &= (\beta-a)e^{(\beta-a)t}\Big(\frac{\beta\theta}{\beta-a} - X_t\Big)\mathrm{d}t - e^{(\beta-a)t}\mathrm{d}X_t \\ &= -e^{(\beta-a)t}\chi(t-)\mathrm{d}L_t \end{aligned} \tag{3.1.3}$$

By solving this last expression, Y_t could be rewritten such as

$$Y_t = Y_s - \int_s^t e^{(\beta-a)u}\chi(u-)\mathrm{d}L_u, \tag{3.1.4}$$

and by combining with the equation (3.1.2), we get :

$$\begin{aligned} X_t &= e^{-(\beta-a)(t-s)}X_s + \int_s^t \Big(\beta\theta e^{-(\beta-a)(t-u)}\Big)\mathrm{d}u + \int_s^t \chi(u-)e^{-(\beta-a)(t-u)}\mathrm{d}L_u \\ &= e^{-(\beta-a)(t-s)}X_s + \Big(\frac{\beta\theta}{\beta-a}\Big)\big(1-e^{-(\beta-a)(t-s)}\big) + \int_s^t \chi(u-)e^{-(\beta-a)(t-u)}\mathrm{d}L_u \end{aligned} \tag{3.1.5}$$

which is nothing but the Ornstein-Uhlenbeck process with a reversion level of $\frac{\beta\theta}{\beta-a}$ except that it is a Lévy process that is used, involving that this level would be modified as we will see by computing the long run mean. An alternative demonstration of the resolution of this SDE is given in Section 9.5. Considering that $\chi(t-) = 1$ the mean is :

$$\mathbb{E}[X_t] = e^{-(\beta-a)t}X_0 + \beta\theta\Big[\frac{1-e^{-(\beta-a)t}}{\beta-a}\Big] + \int_0^t e^{-(\beta-a)(t-u)}(\mu + \mathbb{E}[Y]\lambda)\mathrm{d}u, \tag{3.1.6}$$

by resolving the integral, rearranging and considering the long run level, such as $t \to \infty$,

$$\mathbb{E}[X] = \lim_{t\to\infty}\mathbb{E}[X_t] = \frac{\beta\theta + \mu + \mathbb{E}[Y]\lambda}{\beta-a} \tag{3.1.7}$$

About the Variance, using the quadratic bracket in (2.5.2), we get :

$$\begin{aligned} \mathbb{V}[X_t] &= \mathbb{E}\Big[\Big(\int_0^t e^{-(\beta-a)(t-u)}\mathrm{d}L_u\Big)^2\Big] \\ &= (\sigma^2 + \mathbb{E}[Y^2]\lambda)\Big[\frac{1-e^{-2(\beta-a)t}}{2(\beta-a)}\Big] \end{aligned} \tag{3.1.8}$$

and taking the limit for the long run variance, $t \to \infty$,

$$\mathbb{V}[X] = \lim_{t\to\infty}\mathbb{V}[X_t] = \frac{\sigma^2 + \mathbb{E}[Y^2]\lambda}{2(\beta-a)} \tag{3.1.9}$$

We can simulate some paths in order to have an overview of the process with previously chosen parameters. On the figure 3.1.1, we see that there is a strong mean reversion in this case, at least with the chosen parameters. The influence of past history is weak and the memory fades quickly. It can be seen that the significant lags in the ACF are related to reversion to the mean and then become insignificant once the reversion level is reached. Note that we are looking at the price of the underlying S_t and not the log-price X_t in our simulations. The method for the process simulation will be described in detail in the chapter on pricing and more specifically on the Monte Carlo algorithm (section 4.4).

Intuitively, starting from the equation (3.1.1), we understand that the sign of the parameter a has an influence on the way the memory is captured, in the sense that in the case where a is positive, a sudden

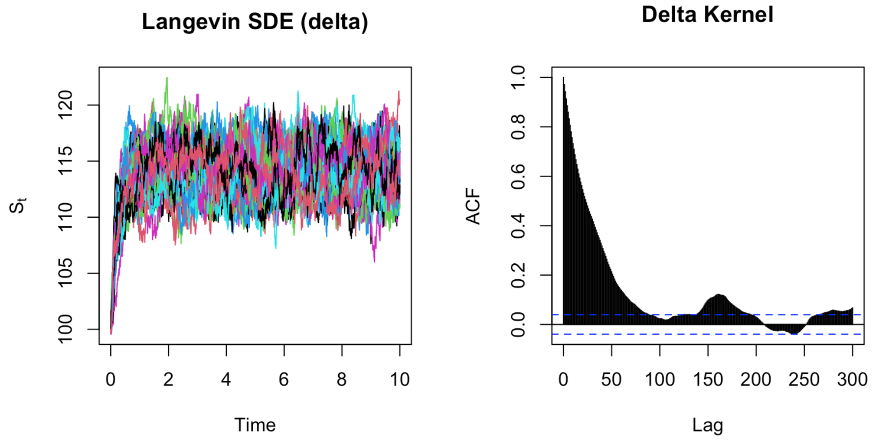

Figure 3.1.1: Simulation of 50 paths of the SDE according to the Delta kernel. The parameters used are $t \in [0, 10], a = -0.5, S_0 = 100, \theta = 0.15, \beta = 3, \lambda = 10, \eta_1 = 100, \eta_2 = 100, p = 0.5, \sigma = 0.1, \mu = 0, N = 2500$.

increase in the price of the underlying will tend to increase the value of the memory term as if the latter interpreted it as an increase that will persist over time. On the contrary, in the case where a is negative, this can be interpreted as: when the price of the underlying rises, it will only be temporary and the underlying will tend to fall back to its reversionary level afterwards. Indeed, we notice that the case where a is negative is more in line with the mean reversion term in terms of intuition. This observation is not observable due to the weakness of the memory in this case, but we will observe it for the next kernels.

About the characteristic function of the process, using the definition in the section 2.2, we get :

$$\begin{aligned}\Phi_{X_t}(u) &= \mathbb{E}[e^{iuX_t}|\mathcal{F}_s] \\ &= e^{iu\left(e^{-(\beta-a)(t-s)}X_s)+\beta\theta\left[\frac{1-e^{-(\beta-a)(t-s)}}{\beta-a}\right]\right)} \times \mathbb{E}[e^{iu\int_s^t e^{-(\beta-a)(t-v)}\mathrm{d}L_v}|\mathcal{F}_s]\end{aligned} \tag{3.1.10}$$

To compute the expectation term we need to use the result given by Eberlein and Raible (1999). Using the characteristic exponent of the Lévy process in (2.6.5) and defining

$$g(u) = e^{-(\beta-a)(t-u)}, \tag{3.1.11}$$

we could use the following properties, demonstrates in the appendix 9.8,

$$\begin{aligned}\mathbb{E}[e^{u\int_s^t g(v)\mathrm{d}L_v}|\mathcal{F}_s] &= e^{\int_s^t \psi(ug(v))dv} \\ &= \exp\left(\int_s^t \left(iug(v)\mu - \frac{\sigma^2u^2g(v)^2}{2} + \lambda\left\{\frac{p\eta_1}{\eta_1 - iug(v)} + \frac{(1-p)\eta_2}{\eta_2 + iug(v)} - 1\right\}\right)dv\right) \\ &= \exp\left(iu\mu\frac{1-e^{-(\beta-a)(t-s)}}{\beta-a} - \frac{\sigma^2u^2(1-e^{-2(\beta-a)(t-s)})}{4(\beta-a)} + \lambda\int_s^t\left\{\frac{p\eta_1}{\eta_1 - iue^{-(\beta-a)(t-v)}}\right.\right. \\ &\qquad\left.\left. + \frac{(1-p)\eta_2}{\eta_2 + iue^{-(\beta-a)(t-v)}} - 1\right\}dv\right) \\ &= \exp\left(iu\mu\frac{1-e^{-(\beta-a)(t-s)}}{\beta-a} - \frac{\sigma^2u^2(1-e^{-2(\beta-a)(t-s)})}{4(\beta-a)} + \lambda\left\{\frac{p}{\beta-a}\log\left(\frac{\eta_1 - iue^{-(\beta-a)(t-s)}}{\eta_1 - iu}\right)\right.\right. \\ &\qquad\left.\left. + \frac{(1-p)}{\beta-a}\log\left(\frac{\eta_2 + iue^{-(\beta-a)(t-s)}}{\eta_2 + iu}\right)\right\}\right)\end{aligned} \tag{3.1.12}$$

where the integral of the jump term is obtained such that

$$\begin{aligned}\int_s^t \frac{p\eta_1}{\eta_1 - iue^{-(\beta-a)(t-v)}}\mathrm{d}v &= p\Big[v - \frac{1}{\beta-a}\log\Big(1-\Big(\frac{iu}{\eta_1}e^{-(\beta-a)t}\Big)e^{(\beta-a)v}\Big)\Big]_{v=s}^{v=t} \\ &= p\Big[(t-s) + \log\Big(\frac{\eta_1 - iue^{-(\beta-a)(t-s)}}{\eta_1 - iu}\Big)\Big].\end{aligned} \tag{3.1.13}$$

Thanks to these relations, it is now possible to compute the final characteristic function of the process, considering that we are at inception, such as $s = 0$, we get :

$$\begin{aligned}\Phi_{X_t}(u) = \exp\Bigg(& iu\Big(e^{-(\beta-a)t}X_0 + \big[\beta\theta+\mu\big]\frac{1-e^{-(\beta-a)t}}{\beta-a}\Big) - \frac{\sigma^2u^2(1-e^{-2(\beta-a)t})}{4(\beta-a)} \\ & + \lambda\Big\{\frac{p}{\beta-a}\log\Big(\frac{\eta_1 - iue^{-(\beta-a)t}}{\eta_1 - iu}\Big) + \frac{(1-p)}{\beta-a}\log\Big(\frac{\eta_2 + iue^{-(\beta-a)t}}{\eta_2+iu}\Big)\Big\}\Bigg)\end{aligned} \tag{3.1.14}$$

Thanks to this characteristic function, we are able to easily extract a lot of information about the process as it allows for an unique definition of the process concerned. In fact, the k-th moments of the process X_t could be obtain by differentiating k times the characteristic function with respect to u, such as,

$$\mathbb{E}[X_t^k] := i^{-k}\,\frac{\partial^k}{\partial u^k}\Phi_{X_t}(u)\bigg|_{u=0} \tag{3.1.15}$$

Concerning the Delta kernel, by defining the function

$$\varphi(t) = \frac{\sigma^2}{2(\beta-a)}(1-e^{-2(\beta-a)t}) + \frac{\lambda}{\beta-a}\Big\{p\big(\frac{1-e^{-2(\beta-a)t}}{\eta_1^2}\big) + (1-p)\big(\frac{1-e^{-2(\beta-a)t}}{\eta_2^2}\big)\Big\}, \tag{3.1.16}$$

which is nothing else than the variance of the process X_t, as in (3.1.8), and the long run level

$$w = \frac{\beta\theta + \mu + \lambda\big(\frac{p}{\eta_1} - \frac{1-p}{\eta_2}\big)}{\beta-a}, \tag{3.1.17}$$

the first 4 moments are such that,

$$\begin{aligned}\mathbb{E}[X_t^1] &= e^{-(\beta-a)t}X_0 + w\big(1-e^{-(\beta-a)t}\big), \\ \mathbb{E}[X_t^2] &= \Bigg[e^{-(\beta-a)t}X_0 + w\big(1-e^{-(\beta-a)t}\big)\Bigg]^2 + \varphi(t), \\ \mathbb{E}[X_t^3] &= \Bigg[e^{-(\beta-a)t}X_0 + w\big(1-e^{-(\beta-a)t}\big)\Bigg]^3 + 3\Bigg[e^{-(\beta-a)t}X_0 + w\big(1-e^{-(\beta-a)t}\big)\Bigg]\varphi(t), \\ \mathbb{E}[X_t^4] &= \Bigg[e^{-(\beta-a)t}X_0 + w\big(1-e^{-(\beta-a)t}\big)\Bigg]^4 + 6\Bigg[e^{-(\beta-a)t}X_0 + w\big(1-e^{-(\beta-a)t}\big)\Bigg]^2\varphi(t) + 3\varphi(t)^2 \\ &\quad + \frac{6\lambda}{\beta-a}\Big\{p\big(\frac{1-e^{-4(\beta-a)t}}{\eta_1^4}\big) + (1-p)\big(\frac{1-e^{-4(\beta-a)}}{\eta_2^4}\big)\Big\}.\end{aligned} \tag{3.1.18}$$

These results allowing us to compute the mean, the variance, the Skewness and the Kurtosis :

$$\begin{aligned}\mathbb{E}[X_t] &= e^{-(\beta-a)t}X_0 + w\big(1-e^{-(\beta-a)t}\big), \\ \mathbb{V}[X_t] &= \Bigg(\frac{\sigma^2}{2} + \lambda\Big(\frac{p}{\eta_1^2} + \frac{1-p}{\eta_2^2}\Big)\Bigg) \times \Big(\frac{1-e^{-2(\beta-a)t}}{\beta-a}\Big), \\ \mathbb{S}[X_t] &= 0, \\ \mathbb{K}[X_t] &= 9 + \frac{6\lambda\Big\{\frac{p}{\eta_1^4} + \frac{1-p}{\eta_2^4}\Big\}\Big(\frac{1-e^{-4(\beta-a)t}}{\beta-a}\Big)}{\mathbb{V}^2[X_t]}.\end{aligned} \tag{3.1.19}$$

We are know interested in developing the autocovariance and the autocorrelation of the process. For ease of calibration in the practical part, we assume here that the mean of the diffusion term is zero. Since $\mathbb{E}[L_t] = 0$ if $\mu = -\lambda\mathbb{E}[Y]$, or equivalently, $\mu = 0$ and $\mathbb{E}[Y] = 0$, considering this case, by setting :

$$g(t-u) = e^{-(\beta-a)(t-u)} \tag{3.1.20}$$

for the Delta kernel, and using the variance of L_t in (2.3.4), if $s < t$ the autocovariance becomes :

$$\begin{aligned}
\mathbb{C}(X_t X_s) &= \mathbb{E}\Big[\int_0^t g(t-u)\mathrm{d}L_u \int_0^s g(s-u)\mathrm{d}L_u\Big] \\
&= \mathbb{E}\Big[\int_0^s g(t-u)\mathrm{d}L_u \int_0^s g(s-u)\mathrm{d}L_u\Big] + \mathbb{E}\Big[\int_0^t g(t-u)\mathrm{d}L_u\Big]\mathbb{E}\Big[\int_0^s g(s-u)\mathrm{d}L_u\Big] \\
&= \Big(\sigma^2 + \int_{\mathbb{R}} z^2\nu(\mathrm{d}z)\Big)\int_0^s g(t-u)g(s-u)\mathrm{d}u \\
&= \Big(\sigma^2 + \lambda\mathbb{E}[Y^2]\Big) \times \frac{e^{-(\beta-a)(t-s)} - e^{-(\beta-a)(t+s)}}{2(\beta-a)}
\end{aligned} \tag{3.1.21}$$

Using these results, it is now straightforward to find the autocorrelation of the process, which is described as

$$\rho_k = \frac{\mathbb{C}(X_t X_{t-k})}{\sqrt{\mathbb{V}[X_t])}\sqrt{\mathbb{V}[X_{t-k}])}}, \qquad k = 0, ..., N-1, \tag{3.1.22}$$

where k corresponds to the lag, i.e. each trading day in our case. Substituting the variance and covariance expressions derived above we find

$$\rho_k = \frac{e^{-(\beta-a)k} - e^{-(\beta-a)(2t-k)}}{\sqrt{1 - e^{-2(\beta-a)t} - e^{-2(\beta-a)(t-k)} + e^{-2(\beta-a)(2t-k)}}}. \tag{3.1.23}$$

We notice that when $k = 0$ the autocorrelation is 1 and decreasing when k increases. These observations correspond to what we expect since the autocovariance becomes equal to the variance at the inception.

3.2 Negative Exponential Kernel

This kernel is generally used when we are facing a persistent influence over a relatively short period of time. If $M(t) = ae^{-bt}$ with $b > 0$ and $b^2 - 4a > 0$, we get :

$$\mathrm{d}X_t = \beta(\theta - X_t)\mathrm{d}t + \Big(\int_0^t ae^{-b(t-u)}X_u\mathrm{d}u\Big)\mathrm{d}t + \chi(t-)\mathrm{d}L_t, \tag{3.2.1}$$

and one can see that the Markov property obtained so far is no longer valid in the sense that the integral in the memory term requires the use of the whole past history of the process and we can no longer be satisfied with the information in t. By studying the Laplace-Transform of the kernel and the inverse Laplace-Transform of the function $H(s)$ as presented in (3.0.8), it follows :

$$\begin{cases} \mathcal{L}[M(s)] = \dfrac{a}{s+b} \\ \mathcal{L}^{-1}[H(s)] = \mathcal{L}^{-1}\Big[\dfrac{1}{s - \mathcal{L}[M(s)] + \beta}\Big] \end{cases} \tag{3.2.2}$$

By replacing the first equation in the second one, developing and using a Laplace-Fourier table[1] it follows:

$$\begin{aligned}\mathcal{L}^{-1}[H(s)] &= \mathcal{L}^{-1}\Big[\frac{s+b}{s^2+(b+\beta)s+b\beta-a}\Big] \\ &= \frac{(s_1+b)e^{s_1t}-(s_2+b)e^{s_2t}}{s_1-s_2}\end{aligned} \tag{3.2.3}$$

where s_1, s_2 are the roots of the quadratic equation on the denominator. If $b^2 > 4a$ then the roots are two negative real numbers (and $s_1 \geq s_2$), and if $b^2 \leq 4a$ then they are complex numbers whose real parts are negative. In the case of complex roots, the model behavior is a periodic oscillation around that of the equal root $\exp(-t/2)$, so our interest is mainly in the case where the solution is real. The figure 3.2.1 shows the evolution of the Exponential kernel on the process, using the expression (3.2.3) and the SDE in (3.0.4) with $M(t) = ae^{-bt}$. This graph can be interpreted as a process that starts at 1 and will be influenced only by its memory term. The parameter a being negative, the process will converge to 0. We begin to understand the interest of combining a mean-reverting term and a memory process (two effects that contradict each other a priori), in order to avoid observing a convergence towards 0 (when $a < 0$) or a divergence of the price ($a > 0$), but rather to determine more or less precisely a level of convergence in the long term. This will become clearer when full simulations of the process are observed later.

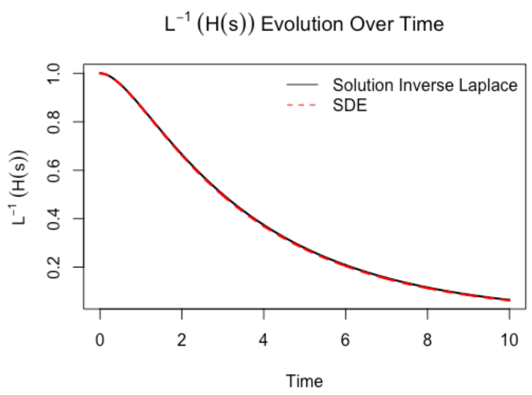

Figure 3.2.1: Evolution over time of the Exponential kernel.

The figure 3.2.1 use the following parameters : $a = -0.5, b = 2, t \in [0, 10]$. It can be seen that the two lines generally overlap and decrease quite slowly over time.

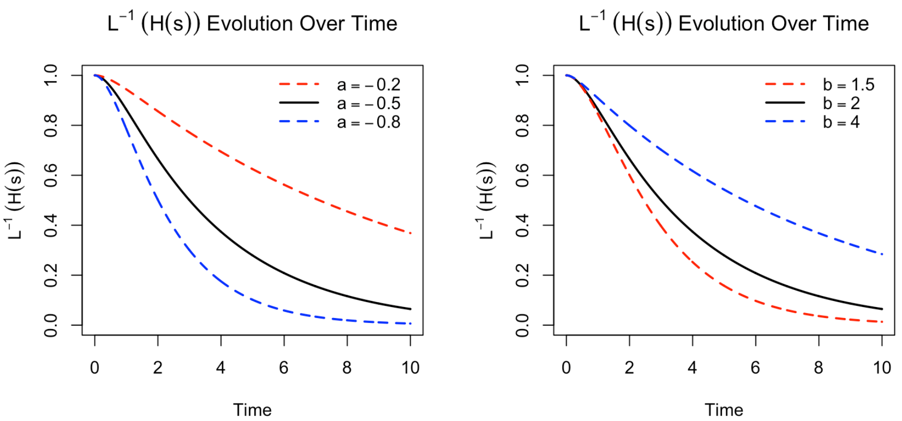

Figure 3.2.2: Sensitive analysis of inverse Laplace transform of the Exponential kernel for the parameters a and b.

[1]
$$\mathcal{L}^{-1}\Big[\frac{(s+d)}{(s+a)(s+b)}\Big] = \frac{(d-a)\exp(-at)-(d-b)\exp(-bt)}{(b-a)}$$

Figure 3.2.2 shows the sensitivity analysis of the inverse of the Laplace transform according to the parameters a and b. Let us emphasise that we observe the influence of the kernel on the process directly. It can be seen that the effects are quite similar: for a more and more negative a, the memory increases more and more and the latter will consider that the drop in price is persistent and will continue to decrease over time. The analysis is the opposite for a decrease of b since we observe a minus in front of the parameter, however, we notice that for shorter maturities, the effect of the parameter b is less pronounced than for a. This can be explained by the fact that b is multiplied by time and it is only when the latter becomes important that the parameter becomes significant.

Using the equation (3.2.3), the expression of the process X_t becomes :

$$
\begin{aligned}
X_t &= \frac{1}{s_1 - s_2} \times \Bigg(X_0\big[(s_1+b)e^{s_1 t} - (s_2+b)e^{s_2 t}\big] + \beta\theta \int_0^t (s_1+b)e^{s_1(t-u)} - (s_2+b)e^{s_2(t-u)} \mathrm{d}u \\
&\qquad + \int_0^t \chi(u-)(s_1+b)e^{s_1(t-u)} - (s_2+b)e^{s_2(t-u)} \mathrm{d}L_u \Bigg) \\
&= \frac{1}{s_1 - s_2} \times \Bigg(X_0\big[(s_1+b)e^{s_1 t} - (s_2+b)e^{s_2 t}\big] + \beta\theta \Big\{ \frac{(s_1+b)(e^{s_1 t}-1)}{s_1} - \frac{(s_2+b)(e^{s_2 t}-1)}{s_2} \Big\} \\
&\qquad + \int_0^t \chi(u-)(s_1+b)e^{s_1(t-u)} - (s_2+b)e^{s_2(t-u)} \mathrm{d}L_u \Bigg)
\end{aligned}
\tag{3.2.4}
$$

Same as the previous kernel, we can simulate some paths in order to have an overview of the process with previously chosen parameters. Once again, see the section 4.4 for more details.

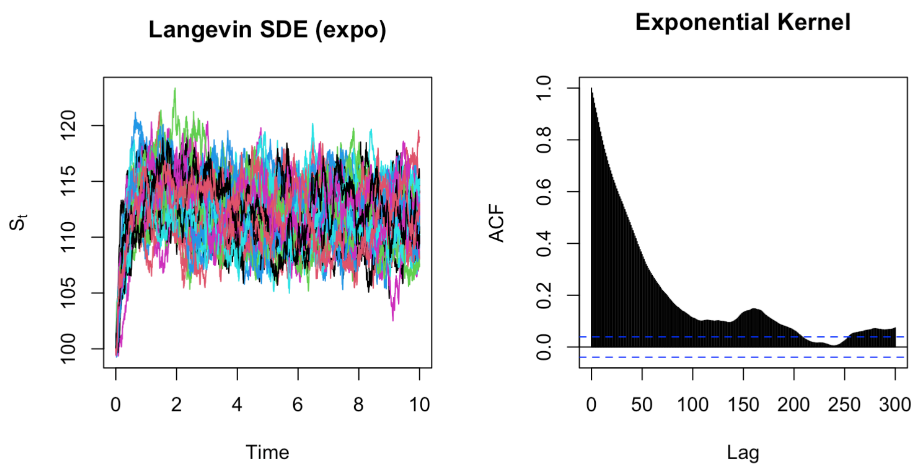

Figure 3.2.3: Simulation of 50 paths of the SDE according to the Exponential kernel. The parameters used are $t \in [0, 10], a = -0.5, S_0 = 100, \theta = 0.15, \beta = 3, \lambda = 10, \eta_1 = 100, \eta_2 = 100, p = 0.5, \sigma = 0.1, \mu = 0, b = 1, N = 2500$.

Figure 3.2.3 shows different trajectories of the underlying dynamics. Contrary to what was observed for the Delta kernel, we observe here that the process tends to go down after reaching the reversion level. This can be interpreted by the fact that initially the mean reversion term takes effect and then the memory term becomes more and more important and takes into account the fact that the process started from a lower level and will therefore tend to go back down. This interpretation could be the opposite depending on the sign of the parameter a. Indeed, if the latter was positive, then the process would have detected the fact that the price of the underlying asset has just risen sharply and that this increase is likely to persist. This can be seen in Figure 3.2.4 where all the parameters are unchanged except for the parameter a which goes from -0.5 to 0.5.

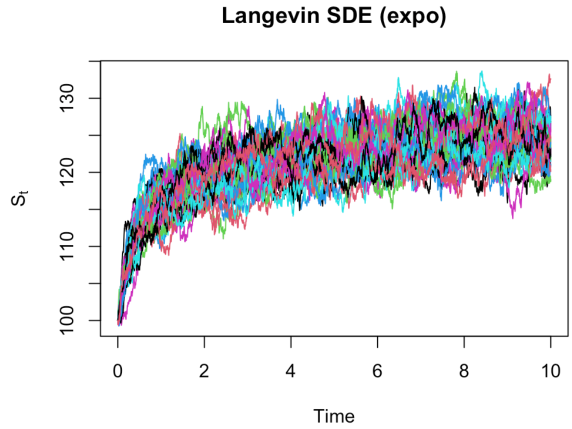

Figure 3.2.4: Simulation of 50 paths of the SDE according to the Exponential kernel with $a = 0.5$.

The mean of the process X_t are defined such as :

$$\mathbb{E}[X_t] = \frac{1}{s_1 - s_2} \times \left[X_0 \left[(s_1+b)e^{s_1 t} - (s_2+b)e^{s_2 t} \right] + \left(\beta\theta + \mu + \lambda\left(\frac{p}{\eta_1} - \frac{1-p}{\eta_2}\right) \right) \left\{ \frac{(s_1+b)(e^{s_1 t}-1)}{s_1} - \frac{(s_2+b)(e^{s_2 t}-1)}{s_2} \right\} \right], \tag{3.2.5}$$

with the final level, when $t \to \infty$,

$$\mathbb{E}[X] = \lim_{t\to\infty} \mathbb{E}[X_t] = \frac{b\left(\beta\theta + \mu + \lambda\left(\frac{p}{\eta_1} - \frac{1-p}{\eta_2}\right)\right)}{s_1 - s_2} \times \left(\frac{1}{s_2} - \frac{1}{s_1}\right). \tag{3.2.6}$$

Concerning the variance :

$$\begin{aligned} \mathbb{V}[X_t] &= \mathbb{E}\left[\left(\frac{1}{s_1 - s_2} \int_0^t \chi(u-)(s_1+b)e^{s_1(t-u)} - (s_2+b)e^{s_2(t-u)} \mathrm{d}L_u \right)^2 \right] \\ &= \mathbb{E}\left[\frac{1}{(s_1-s_2)^2} \int_0^t \left(\chi(u-)(s_1+b)^2 e^{2s_1(t-u)} - 2(s_1+b)(s_2+b)e^{(s_1+s_2)(t-u)} + (s_2+b)^2 e^{2s_2(t-u)} \right) d[L,L]_u \right], \end{aligned} \tag{3.2.7}$$

given that the quadratic variation of a Lévy process is defined as (2.5.2) and considering that $\chi(t-) = 1$

$$\begin{aligned} \mathbb{V}[X_t] &= \frac{1}{(s_1-s_2)^2} \times \mathbb{E}\left[\int_0^t \left((s_1+b)^2 e^{2s_1(t-u)} - 2(s_1+b)(s_2+b)e^{(s_1+s_2)(t-u)} + (s_2+b)^2 e^{2s_2(t-u)} \right) \times \left(\sigma^2 \mathrm{d}u + (Y_u)^2 \mathrm{d}N_u \right) \right] \\ &= \frac{(\sigma^2 + \mathbb{E}[Y^2]\lambda)}{(s_1-s_2)^2} \left((s_1+b)^2 \left[\frac{e^{2s_1 t} - 1}{2s_1} \right] - 2(s_1+b)(s_2+b)\left[\frac{e^{(s_1+s_2)t} - 1}{s_1+s_2} \right] + (s_2+b)^2 \left[\frac{e^{2s_2 t} - 1}{2s_2} \right] \right). \end{aligned} \tag{3.2.8}$$

with the final level, when $t \to \infty$,

$$\mathbb{V}[X] = \lim_{t\to\infty} \mathbb{V}[X_t] = \frac{(\sigma^2 + \mathbb{E}[Y^2]\lambda)}{(s_1-s_2)^2} \left(-\frac{(s_1+b)^2}{2s_1} + \frac{2(s_1+b)(s_2+b)}{s_1+s_2} - \frac{(s_2+b)^2}{2s_2} \right) \tag{3.2.9}$$

Once again, about the characteristic function of the process, using the same procedure as for the Delta kernel in (3.1.14) and defining

$$g_1(u) = \frac{(s_1+b)e^{s_1(t-u)}}{s_1 - s_2} \text{ and } g_2(u) = \frac{(s_2+b)e^{s_2(t-u)}}{s_1 - s_2} \tag{3.2.10}$$

we could compute the integral part in (3.2.4),

$$
\begin{aligned}
&\mathbb{E}[e^{u\int_s^t g_1(v)\mathrm{d}L_v - u\int_s^t g_2(v)\mathrm{d}L_v}|\mathcal{F}_s] = e^{\int_s^t \psi(ug_1(v))dv - \int_s^t \psi(ug_2(v))dv} \\
&\quad = \exp\Bigg(\frac{iu\mu}{s_1-s_2}\Big(\frac{(s_1+b)(e^{s_1(t-s)}-1)}{s_1} - \frac{(s_2+b)(e^{s_2(t-s)}-1)}{s_2}\Big) \\
&\qquad\qquad - \frac{u^2\sigma^2}{4(s_1-s_2)^2}\Big(\frac{(s_1+b)^2(e^{2s_1(t-s)}-1)}{s_1} - \frac{(s_2+b)^2(e^{2s_2(t-s)}-1)}{s_2}\Big) \\
&\qquad\qquad + \lambda\Bigg\{\frac{p}{s_1}\times\log\Big(\frac{\eta_1(s_1-s_2)-iu(s_1+b)}{\eta_1(s_1-s_2)-iu(s_1+b)e^{s_1(t-s)}}\Big) + \frac{(1-p)}{s_1}\times\log\Big(\frac{\eta_2(s_1-s_2)+iu(s_1+b)}{\eta_2(s_1-s_2)+iu(s_1+b)e^{s_1(t-s)}}\Big) \\
&\qquad\qquad - \frac{p}{s_2}\times\log\Big(\frac{\eta_1(s_1-s_2)-iu(s_2+b)}{\eta_1(s_1-s_2)-iu(s_2+b)e^{s_2(t-s)}}\Big) - \frac{(1-p)}{s_2}\times\log\Big(\frac{\eta_2(s_1-s_2)+iu(s_2+b)}{\eta_2(s_1-s_2)+iu(s_2+b)e^{s_2(t-s)}}\Big)\Bigg\}\Bigg)
\end{aligned}
\tag{3.2.11}
$$

Thanks to this equation, we compute the final characteristic function of the process, considering that we are at inception, such as $s = 0$, we get :

$$
\begin{aligned}
\Phi_{X_t}(u) = \exp\Bigg(&\frac{iu}{s_1-s_2}\times\Bigg[X_0[(s_1+b)e^{s_1t}-(s_2+b)e^{s_2t}] + (\beta\theta+\mu)\times\Big\{\frac{(s_1+b)(e^{s_1t}-1)}{s_1} - \frac{(s_2+b)(e^{s_2t}-1)}{s_2}\Big\}\Bigg] \\
&- \frac{u^2\sigma^2}{4(s_1-s_2)^2}\Big(\frac{(s_1+b)^2(e^{2s_1t}-1)}{s_1} - \frac{(s_2+b)^2(e^{2s_2t}-1)}{s_2}\Big) \\
&+ \lambda\Bigg\{\frac{p}{s_1}\times\log\Big(\frac{\eta_1(s_1-s_2)-iu(s_1+b)}{\eta_1(s_1-s_2)-iu(s_1+b)e^{s_1t}}\Big) + \frac{(1-p)}{s_1}\times\log\Big(\frac{\eta_2(s_1-s_2)+iu(s_1+b)}{\eta_2(s_1-s_2)+iu(s_1+b)e^{s_1t}}\Big) \\
&- \frac{p}{s_2}\times\log\Big(\frac{\eta_1(s_1-s_2)-iu(s_2+b)}{\eta_1(s_1-s_2)-iu(s_2+b)e^{s_2t}}\Big) - \frac{(1-p)}{s_2}\times\log\Big(\frac{\eta_2(s_1-s_2)+iu(s_2+b)}{\eta_2(s_1-s_2)+iu(s_2+b)e^{s_2t}}\Big)\Bigg\}\Bigg)
\end{aligned}
\tag{3.2.12}
$$

By using the same way of proceeding as for the kernel Delta (by deriving the characteristic function), we obtain easier the different moments of the process rather than computing them directly from the process. Concerning the Exponential kernel, by defining the function

$$
\varphi(t) = \frac{\Big(\sigma^2 + 2\lambda\big(\frac{p}{\eta_1^2}+\frac{1-p}{\eta_2^2}\big)\Big)}{(s_1-s_2)^2}\Bigg((s_1+b)^2\Big[\frac{e^{2s_1t}-1}{2s_1}\Big] - 2(s_1+b)(s_2+b)\Big[\frac{e^{(s_1+s_2)t}-1}{s_1+s_2}\Big] + (s_2+b)^2\Big[\frac{e^{2s_2t}-1}{2s_2}\Big]\Bigg)
\tag{3.2.13}
$$

which is the variance of the process X_t, as in (3.2.8), and the long run level

$$
w = \beta\theta + \mu + \lambda\Big(\frac{p}{\eta_1} - \frac{1-p}{\eta_2}\Big),
\tag{3.2.14}
$$

the first 4 moments are such that,

$$
\begin{aligned}
\mathbb{E}[X_t^1] &= \frac{1}{s_1-s_2}\times\Bigg[X_0\big[(s_1+b)e^{s_1t}-(s_2+b)e^{s_2t}\big] + w\Big\{\frac{(s_2+b)(1-e^{s_2t})}{s_2} - \frac{(s_1+b)(1-e^{s_1t})}{s_1}\Big\}\Bigg], \\
\mathbb{E}[X_t^2] &= \frac{1}{(s_1-s_2)^2}\times\Bigg[X_0\big[(s_1+b)e^{s_1t}-(s_2+b)e^{s_2t}\big] \\
&\qquad + w\Big\{\frac{(s_2+b)(1-e^{s_2t})}{s_2} - \frac{(s_1+b)(1-e^{s_1t})}{s_1}\Big\}\Bigg]^2 + \varphi(t), \qquad (3.2.15) \\
\mathbb{E}[X_t^3] &= \mathbb{E}^3[X_t^1] + 3\mathbb{E}[X_t^1]\varphi(t), \\
\mathbb{E}[X_t^4] &= \mathbb{E}^4[X_t^1] + 6\mathbb{E}^2[X_t^1]\varphi(t) + 3\varphi(t)^2 + \frac{24\lambda\big(\frac{p}{\eta_1^4}+\frac{1-p}{\eta_2^4}\big)}{(s_1-s_2)^2}\Bigg((s_1+b)^2\Big[\frac{e^{4s_1t}-1}{4s_1}\Big] \\
&\qquad - 2(s_1+b)(s_2+b)\Big[\frac{e^{2(s_1+s_2)t}-1}{2(s_1+s_2)}\Big] + (s_2+b)^2\Big[\frac{e^{4s_2t}-1}{4s_2}\Big]\Bigg).
\end{aligned}
$$

These results allowing us to compute the mean, the variance, the Skewness and the Kurtosis :

$$
\begin{aligned}
&\mathbb{E}[X_t] = \frac{1}{s_1 - s_2} \times \Bigg[X_0\big[(s_1+b)e^{s_1 t} - (s_2+b)e^{s_2 t}\big] + w\Big\{\frac{(s_2+b)(1-e^{s_2 t})}{s_2} - \frac{(s_1+b)(1-e^{s_1 t})}{s_1}\Big\}\Bigg],\\
&\mathbb{V}[X_t] = \frac{\Big(\sigma^2 + 2\lambda\big(\frac{p}{\eta_1^2} + \frac{1-p}{\eta_2^2}\big)\Big)}{(s_1-s_2)^2}\Bigg((s_1+b)^2\Big[\frac{e^{2s_1 t}-1}{2s_1}\Big] - 2(s_1+b)(s_2+b)\Big[\frac{e^{(s_1+s_2)t}-1}{s_1+s_2}\Big] + (s_2+b)^2\Big[\frac{e^{2s_2 t}-1}{2s_2}\Big]\Bigg),\\
&\mathbb{S}[X_t] = 0,\\
&\mathbb{K}[X_t] = 9 + \frac{\frac{24\lambda\left(\frac{p}{\eta_1^4}+\frac{1-p}{\eta_2^4}\right)}{(s_1-s_2)^2}\Bigg((s_1+b)^2\Big[\frac{e^{4s_1t}-1}{4s_1}\Big] - 2(s_1+b)(s_2+b)\Big[\frac{e^{2(s_1+s_2)t}-1}{2(s_1+s_2)}\Big] + (s_2+b)^2\Big[\frac{e^{4s_2t}-1}{4s_2}\Big]\Bigg)}{\mathbb{V}^2[X_t]}.
\end{aligned}
\tag{3.2.16}
$$

Concerning the autocovariance and the autocorrelation, we will assume here also that the diffusion term is zero. Since $\mathbb{E}[L_t] = 0$ if $\mu = -\lambda\mathbb{E}[Y]$, or equivalently, $\mu = 0$ and $\mathbb{E}[Y] = 0$, considering this case, by setting :

$$
g(t-u) = \frac{(s_1+b)e^{s_1(t-u)} - (s_2+b)e^{s_2(t-u)}}{s_1 - s_2} \tag{3.2.17}
$$

and using the variance of L_t in (2.3.4), if $s < t$ the autocovariance becomes, in a similar way to the Delta kernel :

$$
\begin{aligned}
\mathbb{C}(X_t X_s) &= \mathbb{E}\Big[\int_0^t g(t-u)\mathrm{d}L_u \int_0^s g(s-u)\mathrm{d}L_u\Big]\\
&= \mathbb{E}\Big[\int_0^s g(t-u)\mathrm{d}L_u \int_0^s g(s-u)\mathrm{d}L_u\Big] + \mathbb{E}\Big[\int_0^t g(t-u)\mathrm{d}L_u\Big]\mathbb{E}\Big[\int_0^s g(s-u)\mathrm{d}L_u\Big]\\
&= \Big(\sigma^2 + \int_{\mathbb{R}} z^2\nu(\mathrm{d}z)\Big)\int_0^s g(t-u)g(s-u)\mathrm{d}u\\
&= \frac{\Big(\sigma^2 + \lambda\mathbb{E}[Y^2]\Big)}{(s_1-s_2)^2} \times \Big[\frac{(s_1+b)^2}{2s_1}(e^{s_1(t+s)} - e^{s_1(t-s)}) - \frac{(s_1+b)(s_2+b)}{(s_1+s_2)}(e^{s_1t+s_2s} - e^{s_1(t-s)})\\
&\qquad - \frac{(s_1+b)(s_2+b)}{(s_1+s_2)}(e^{s_2t+s_1s} - e^{s_2(t-s)}) + \frac{(s_2+b)^2}{2s_2}(e^{s_2(t+s)} - e^{s_2(t-s)})\Big]
\end{aligned}
\tag{3.2.18}
$$

which leads to the following autocorrelation:

$$
\rho_k = \frac{\frac{(s_1+b)^2}{2s_1}(e^{s_1(2t-k)} - e^{s_1k}) - \frac{(s_1+b)(s_2+b)}{(s_1+s_2)}\Big((e^{s_1t+s_2(t-k)} - e^{s_1k}) + (e^{s_2t+s_1(t-k)} - e^{s_2k})\Big) + \frac{(s_2+b)^2}{2s_2}(e^{s_2(2t-k)} - e^{s_2k})}{\sqrt{\mathbb{V}[X_t]} \times \sqrt{\mathbb{V}[X_{t-k}]}}.
\tag{3.2.19}
$$

3.3 Negative Power Kernel

On the contrary to the Exponential negative kernel, a negative power function is used when the influence of the past persists over a longer period of time. If $M(t) = at^{-\alpha}$ with the condition[2] $0 < \alpha < 1$:

$$
\mathrm{d}X_t = \beta(\theta - X_t)\mathrm{d}t + \Big(\int_0^t a(t-u)^{-\alpha}X_u\mathrm{d}u\Big)\mathrm{d}t + \chi(t-)\mathrm{d}L_t, \tag{3.3.1}
$$

[2] If $\alpha < 0$, the function is inconsistent with the law of causality and if $\alpha > 1$, there is no Fourier-Laplace Transform available.

and which not anymore a Markovian process as with the negative Exponential kernel. By using the same procedure as for the previous section, it follows :

$$\begin{cases} \mathcal{L}[M(s)] = \dfrac{a\Gamma(1-\alpha)}{s^{1-\alpha}} \\ \mathcal{L}^{-1}[H(s)] = \mathcal{L}^{-1}\Big[\dfrac{1}{s - \mathcal{L}[M(s)] + \beta}\Big] \end{cases} \tag{3.3.2}$$

Once again by replacing the first equation in the second one and developing it follows:

$$\mathcal{L}^{-1}[H(s)] = \mathcal{L}^{-1}\Big[\frac{s^{1-\alpha}}{s^{2-\alpha} + \beta s^{1-\alpha} - a\Gamma(1-\alpha)}\Big]. \tag{3.3.3}$$

This expression does not allow to get an explicit formula of the inverse Laplace transform and must therefore be computed numerically. However, there are different solutions when we simplify the expression (3.3.1) and do not consider any mean-reverting term, such that:

$$\mathrm{d}X_t = \theta \mathrm{d}t + \Big(\int_0^t a(t-u)^{-\alpha} X_u \mathrm{d}u\Big)\mathrm{d}t + \chi(t-)\mathrm{d}L_t. \tag{3.3.4}$$

Before proceeding further this simplification, we can take the time to look at some simulations of equation (3.3.1) in Figure 3.3.1.

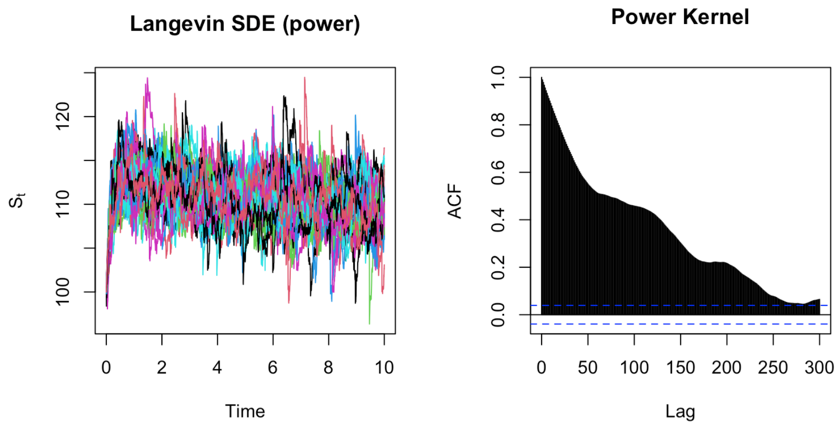

Figure 3.3.1: Simulation of 50 paths of the SDE according to the Exponential kernel. The parameters used are $t \in [0, 10], a = -0.5, S_0 = 100, \theta = 0.15, \beta = 3, \lambda = 10, \eta_1 = 100, \eta_2 = 100, p = 0.5, \sigma = 0.1, \mu = 0, \alpha = 0.4, N = 2500$.

First of all, we observe that, as announced, the dependence in time is more important in this case, with autocorrelations still significant, 300 days later (i.e. more than one year because 252 trading days per year). Then, we observe that the return to the mean is still strong and that the interpretation of the parameter a is the same as for the two other kernels: if $a < 0$ then a rise is seen as temporary in a way that is equivalent to the process retaining the fact that it started lower previously. Whereas if $a > 0$, a rise is seen as persistent over time and a real trend is observed. Figure 3.3.2 shows the opposite case, where the parameter is positive.

Case 1 : Benth & al (2020)

In the case where $0 < \alpha < 0.5$, Benth & al (2020) proposes a function that approximates this Laplace inverse function taking the following form :

$$\mathcal{L}^{-1}[H(s)] = \sum_{i=0}^{\infty} w_i(\alpha)(at^{2-\alpha})^i, \tag{3.3.5}$$

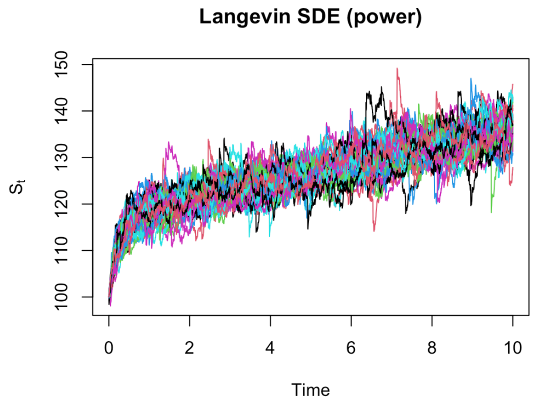

Figure 3.3.2: Simulation of 50 paths of the SDE according to the Exponential kernel with $a = 0.5$.

where $w_i(\alpha)$ represent the weights and have the following recursive relations :

$$w_i(\alpha) = w_{i-1}(\alpha)\frac{\Gamma\Big(1+(i-1)(2-\alpha)\Big)}{\Gamma\Big(1+i(2-\alpha)\Big)}\Gamma(1-\alpha), \qquad n = 1, 2, 3, ... \tag{3.3.6}$$

with $w_0(\alpha) = 1$ and where $\Gamma(.)$ is the Gamma-function. We can thus notice in equation (3.3.5) that the weights are decreasing and converge towards 0, contrary to the second part of the equation which will diverge towards infinity, more or less quickly according to the value of α and of a : the greater will be α, the more slowly the term will diverge and the greater will be a, the more quickly the term will diverge. The authors also suggest the use of Stirling's formula, which allows an approximation of the gamma function for values that become too large[3] : $\Gamma(1+x) \sim k\sqrt{x}(x/e)x$, which, in our case, becomes

$$\frac{\Gamma(1+(n-1)(2-\alpha))}{\Gamma(1+n(2-\alpha))} \sim \sqrt{\frac{n-1}{n}}\Big((1-\frac{1}{n})^n\frac{e}{(n-1)(2-\alpha)}\Big)^{2-\alpha}, \tag{3.3.7}$$

which tends to 0 when $n \to \infty$. The next graph shows the evolution of the Power kernel on the process in three different forms: the Laplace inverse measured numerically, the infinite sum suggested by Benth and the expression of the SDE presented in the equation (3.3.1) with $M(t) = at^{-\alpha}$. The figure 3.3.3 use the following parameters : $\alpha = 0.4, t \in [0, 10], a = -0.5$ and 100 parameters in the sum, which is sufficient. It can be seen that the numerical approach and the infinite sum merge, whereas the SDE deviates slightly from the line, which is consistent with intuition since it is itself a discretization of the process.

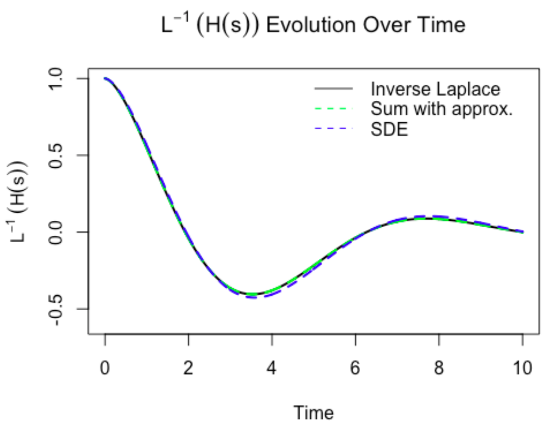

Figure 3.3.3: Evolution over time of the Power kernel.

[3] k is a positive constant.

Case 2 : Hainaut (2021)

A second approach to the negative Power kernel is introduced by Hainaut (2021) who proposes the following Power kernel :

$$M(t) = \frac{at^{\alpha-2}}{\Gamma(\alpha-1)}, \tag{3.3.8}$$

for any $1 < \alpha < 2$, since there is no closed formula for when $\alpha < 1$ and since we are looking for a decreasing path, we do not consider the case $\alpha > 2$. Now, by introducing the Mittag-Leffler function, which is well known when tackling Laplace transforms, we are able to proceed in a similar way to previous kernels:

$$\begin{cases} \mathcal{L}[M(s)] = \dfrac{a\Gamma(1-\alpha)}{s^{\alpha-1}\Gamma(1-\alpha)} = \dfrac{a}{s^{\alpha-1}} \\ \mathcal{L}^{-1}[H(s)] = \mathcal{L}^{-1}\Big[\dfrac{s^{\alpha-1}}{s^{\alpha}-a}\Big] = \mathcal{E}_{\alpha}[-at^{\alpha}] \end{cases} \tag{3.3.9}$$

where $\mathcal{E}(.)$ is the Mittag-Feller function, such as :

$$\mathcal{E}_{\alpha}[-at^{\alpha}] = \sum_{n=0}^{\infty} \frac{(-at^{\alpha})^n}{\Gamma(n\alpha+1)}. \tag{3.3.10}$$

This function, which is important in the theory of the fractional calculus, is a generalisation of the Exponential function when $\alpha = 1$, but also :

$$\mathcal{E}_0[-a] = \frac{1}{1+a} \qquad\qquad \mathcal{E}_1[-at] = \exp(-at). \tag{3.3.11}$$

Furthermore, Mainardi (2020) presents two asymptotic approximations, one for the short term and one for the long term:

$$\mathcal{E}_{\alpha}[-at^{\alpha}] = 1 - \frac{at^{\alpha}}{\Gamma(1+\alpha)} + ... \sim \exp\Big(-\frac{at^{\alpha}}{\Gamma(1+\alpha)}\Big), \quad t \to 0^+ \qquad\qquad \mathcal{E}_{\alpha}[-at^{\alpha}] \sim \frac{t^{-\alpha}}{a\Gamma(1-\alpha)}, \quad t \to \infty. \tag{3.3.12}$$

We are now able to compute the following solution,

$$X_t = X_0\mathcal{E}_{\alpha}[-at^{\alpha}] + \theta\int_0^t \mathcal{E}_{\alpha}[-a(t-u)^{\alpha}]\mathrm{d}u + \int_0^t \mathcal{E}_{\alpha}[-a(t-u)^{\alpha}]\mathrm{d}L_u. \tag{3.3.13}$$

Figure 3.3.4 presents a sensitive analysis of the inverse of the Laplace transform for the Power kernel introduced by Hainaut. It can be seen that the parameter a mainly affects the speed at which the memory fades/remains: the larger the parameter, the more the memory influences the further process. The different trajectories are subject mainly to a horizontal compression depending on time but practically no modification of the curve. Moreover, we can observe « bumps » in the trajectories which are justified by the fact that after the trajectory has fallen over time, the process « keeps the information » of the past fall and takes more time to rise again. On the contrary, the parameter α mainly affects the persistent of the memory influence. As can be seen in the graph on the left, the different curves undergo more changes on the y axis. In such a way that the « bumps » are now more and more (resp. less and less) frequent. This can be justified by the mathematical properties of the Mittag-Leffler function. In order to understand the usefulness and flexibility of the function, some of the possible manifestations of the ML function are given in figure 3.3.5.

Thus, the function ML interpolates for the intermediate time between the stretched Exponential and the negative power law. Since we know that the stretched Exponential models the very fast decay for a small time t, while the asymptotic power law is due to the very slow decay for a large time t, we understand that this ML function is a good combination between these two kernels.

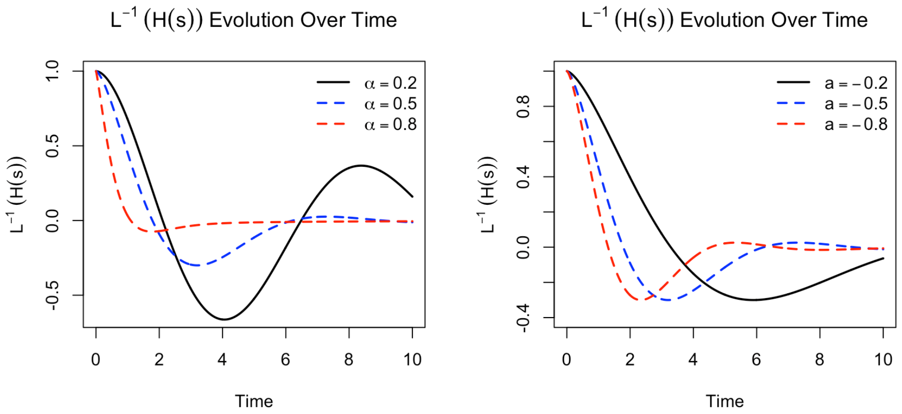

Figure 3.3.4: Sensitive analysis of inverse Laplace transform of the Power kernel for the parameters a and α according to the Hainaut's case.

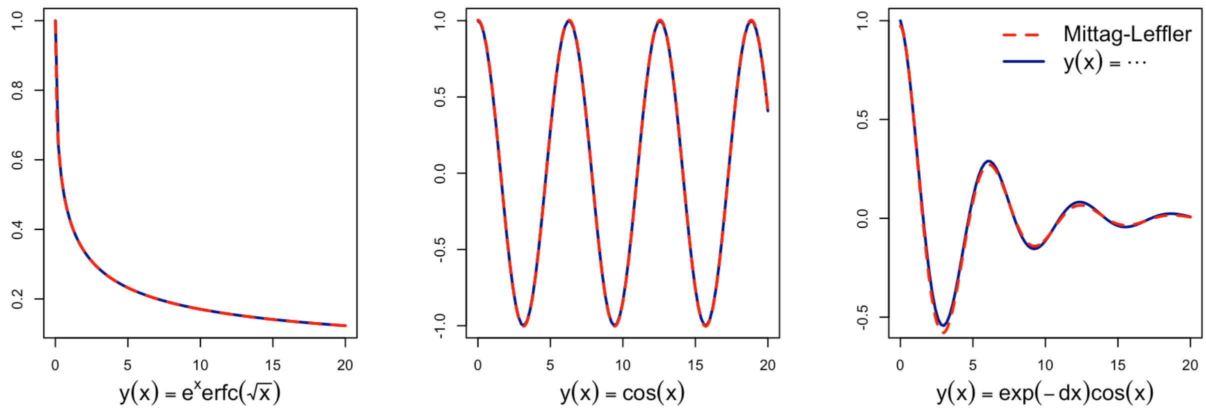

Figure 3.3.5: Various usual functions and the fit of the corresponding Mittag-Leffler function.

Then if we consider the moment generating function of the expression (3.3.13), we get :

$$\Phi_{X_t}(u) = \exp\left(iuX_0\mathcal{E}_\alpha[-at^\alpha] + iu\theta\int_0^t \mathcal{E}_\alpha[-a(t-u)^\alpha]\mathrm{d}u + \int_0^t \psi\Big(iu\mathcal{E}_\alpha[-a(t-u)^\alpha]\Big)\mathrm{d}u\right), \tag{3.3.14}$$

where ψ is the characteristic exponent. As explained earlier, by deriving this expression, we obtain the different moments of the process. Then, the conditional mean is computed such as :

$$\mathbb{E}[X_t|\mathcal{F}_s] = X_s\mathcal{E}_\alpha[-a(t-s)^\alpha] + \Big(\theta+\mu+\lambda(\frac{p}{\eta_1}-\frac{1-p}{\eta_2})\Big)\int_s^t \mathcal{E}_\alpha[-a(t-u)^\alpha]\mathrm{d}u \tag{3.3.15}$$

whereas the conditional variance of X_t is the variance of the Lévy process times the integral of the squared kernels :

$$\mathbb{V}[X_t|\mathcal{F}_s] = (\sigma^2+\lambda\mathbb{E}[Y^2])\int_s^t \Big(\mathcal{E}_\alpha[-a(t-u)^\alpha]\Big)^2\mathrm{d}u. \tag{3.3.16}$$

Finally, by considering the same approach than for the previous kernel, the autocovariance function are given by

$$\begin{aligned}\mathbb{C}(X_tX_s) &= \mathbb{E}\Big[\int_0^t g(t-u)\mathrm{d}L_u \int_0^s g(s-u)\mathrm{d}L_u\Big] \\ &= \Big(\sigma^2+\lambda\mathbb{E}[Y^2]\Big)\int_0^s \mathcal{E}_\alpha[-a(t-u)^\alpha]\mathcal{E}_\alpha[-a(s-u)^\alpha]\mathrm{d}u.\end{aligned} \tag{3.3.17}$$

Note that the autocovaricance function and the variance function do not have closed formulas and must be estimated numerically. The autocorrelation is such that, at inception

$$\rho_k = \frac{\int_0^{t-k} \mathcal{E}_\alpha[-a(t-u)^\alpha]\mathcal{E}_\alpha[-a(t-k-u)^\alpha]\mathrm{d}u}{\sqrt{\int_0^t \Big(\mathcal{E}_\alpha[-a(t-u)^\alpha]\Big)^2 \mathrm{d}u} \times \sqrt{\int_0^{t-k} \Big(\mathcal{E}_\alpha[-a(t-k-u)^\alpha]\Big)^2 \mathrm{d}u}}. \tag{3.3.18}$$

Case 3 : Takahashi (1996)

The third and last proposed Power kernel is the one introduced by Takahashi (1996) which takes the general form as presented earlier:

$$M(t) = at^{-\alpha}, (0 < \alpha < 1), \tag{3.3.19}$$

where, once again, the parameter α is bounded in order to propose, on the one hand, a Laplace-Fourier transform that exists ($\alpha < 1$), and on the other hand, so that the influence function is consistent with the causality law ($\alpha > 0$). In order to offer a closed[4] and tractable form, Takahashi suggested to use $\alpha = 0.5$, leading to :

$$\begin{cases} \mathcal{L}[M(s)] = \dfrac{a\Gamma(0.5)}{s^{0.5}} \\ \mathcal{L}^{-1}[H(s)] = \mathcal{L}^{-1}\Big[\dfrac{s^{0.5}}{s^{1.5} + a\Gamma(0.5)}\Big] \end{cases} \tag{3.3.20}$$

where we can rewrite the second relation such as

$$\begin{aligned} H(s) &= \frac{s^{0.5}}{s^{1.5} + a\sqrt{\pi}} \\ &= \frac{s_2}{(s_1 - s_3)(s_1 - s_2)}\Big(\frac{1}{\sqrt{s} - s_1} - \frac{1}{\sqrt{s} - s_3}\Big) - \frac{s_3}{(s_1 - s_3)(s_2 - s_3)}\Big(\frac{1}{\sqrt{s} - s_2} - \frac{1}{\sqrt{s} - s_3}\Big) \end{aligned} \tag{3.3.21}$$

where s_1, s_2, s_3 are the roots of the function $s^{1.5} + a\sqrt{\pi} = 0$ with respect to $\sqrt{s}$. Then, by using the linear properties of the inverse Laplace-Fourier transform such as:

$$\mathcal{L}^{-1}[aG(t) + bF(t)] = a\mathcal{L}^{-1}[G(t)] + b\mathcal{L}^{-1}[F(t)], \tag{3.3.22}$$

and the following result, with erfc(.), the complementary error function:

$$g(t, a) := \mathcal{L}^{-1}\Big[\frac{1}{\sqrt{s} - a}\Big] = \frac{1}{\sqrt{\pi t}} + ae^{a^2 t}\mathrm{erfc}(-a\sqrt{t}), \tag{3.3.23}$$

we are able to compute the solution, by defining:

$$f(t) := \frac{s_2\Big(g(t, s_1) - g(t, s_3)\Big)}{(s_1 - s_3)(s_1 - s_2)} - \frac{s_3\Big(g(t, s_2) - g(t, s_3)\Big)}{(s_1 - s_3)(s_2 - s_3)}, \tag{3.3.24}$$

we get:

$$X_t = X_0 f(t) + \theta \int_0^t f(t-u)\mathrm{d}u + \int_0^t f(t-u)\mathrm{d}L_u. \tag{3.3.25}$$

We notice that this case seems to propose a more exotic, less interpretable and less tractable form. Thus, we will mainly focus on the Delta and Exponential kernels in the rest of this work.

[4]By using a Laplace-Fourier transform table.

3.4 Pathwise Comparison

After introducing each of the different memory functions M in details, we will briefly compare the influence it can have on the price dynamics of the underlying asset. Figure 3.4.1 shows the same sample path according to each of the different models as well as the ACF of each of them.

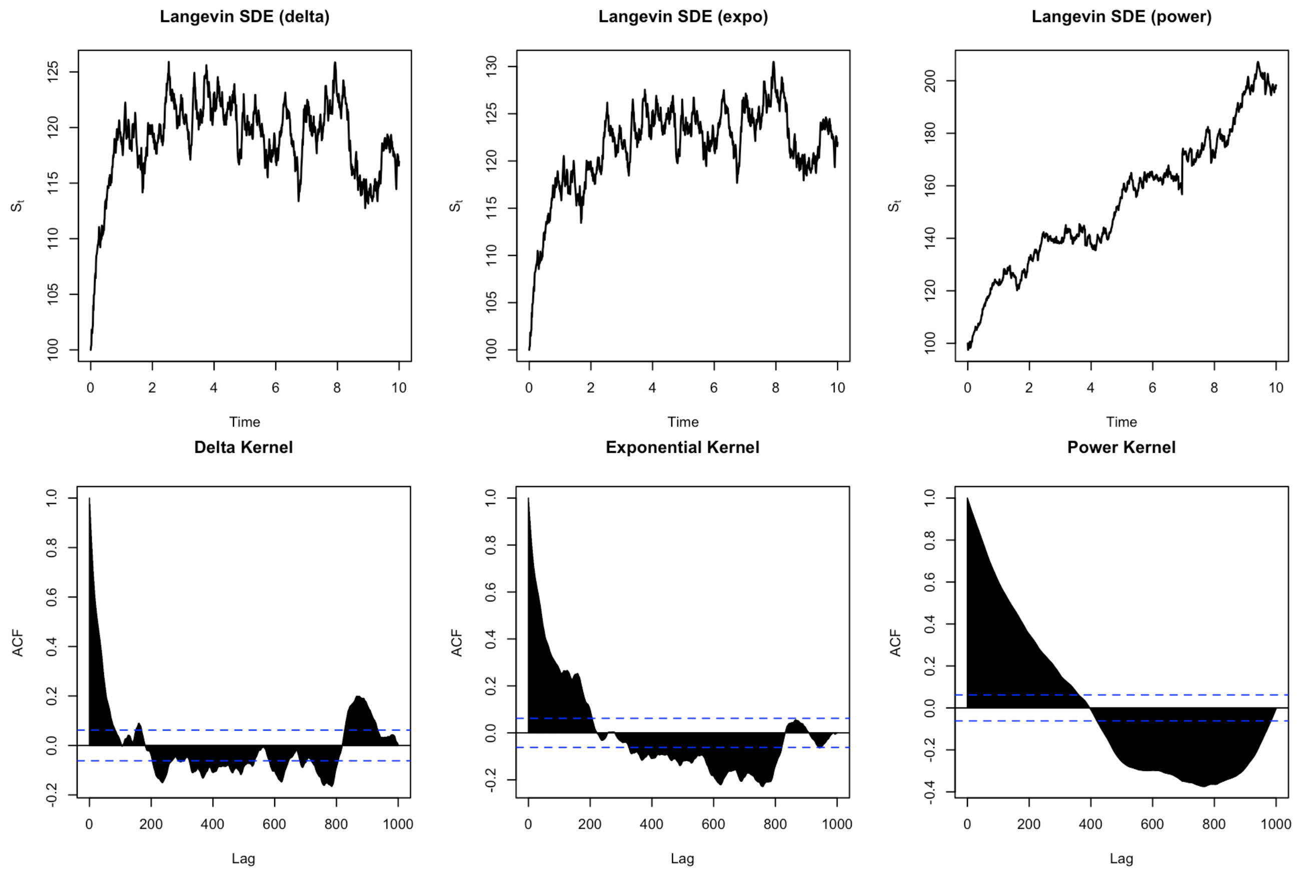

Figure 3.4.1: Simulation of one path of the SDE S_t for each of the different model introduced (Delta, Exponential, Power kernel) and their respective ACF. The parameters used are $\alpha = 0.7, t \in [0, 10], a = 0.5, S_0 = 100, \theta = 0.15, \beta = 3, \lambda = 10, \eta_1 = 100, \eta_2 = 100, p = 0.5, \sigma = 0.1, \mu = 0, b = 1, N = 1000$.

Concerning the price, we notice first of all that each of them observe a rather fast growth in the first months and that the Delta kernel and Exponential kernel models tend to stabilise. However, the Exponential model seems to approach higher values than the Delta model, which can be justified mathematically by the fact that the memory term takes more weight in the process than for the Delta model. These two models, after having reached their long-term level, seem to have a rather erratic trajectory. It can be seen that it is not an easy task to define a precise mean reversion level since the memory term also influences this level. It will therefore be more preferable to focus on the long term level rather than on the θ parameter itself.

Intuitively, one can interpret this more pronounced growth in the case of the Exponential model such that after a fairly steep growth, the model tries to keep the price fairly high and maintain growth. Indeed, the price memory assimilates this increase as something long-lasting and which seems to persist ($a > 0.5$). However, as the years go by and growth becomes less marked, the price tends to stabilise. This can also be seen in the ACFs. Indeed, we observe that the autocorrelation of the Exponential model is stronger and persists more over time compared to the Delta model.

On the contrary, the Power kernel model seems to keep growing over time, a real trend is observed along the years. This can be explained by the fact that the autocorrelation is so strong over time, the growth of the process between its initial level and its level of return to the mean is assimilated to a persistent growth over time and does not seem to stop, the process seems to diverge with this particular parametrisation.

In conclusion, we can see that each model seems to respect its respective features: the Delta model with an instantaneous memory, the Exponential model with a longer term memory but which tends to fade quite quickly and finally the power model with a strong memory which fades slowly.

Chapter 4

Options Pricing

Having a process characterizing the evolution of the spot[1] price S_T of some underlying asset, we would like to valuate option price : an European call of maturity T. However, the Black & Scholes framework is no more available in this case because we are not facing a log-normal distribution, forcing us to use a different way to compute the option valuation. In fact, the process L_t that we present in the previous section follows an unknown law since it sums a normal distribution with another distribution (the jump part). Carr & Madan (1999) proposed to solve this problem with a numerical method to obtain option prices and to do so, they use the well-known FFT and therefore the characteristic function of the log-price $X_T = \log(S_T)$. This pricing method benefits from the fact that it only requires the characteristic function of the model and can therefore be easily transposed to any other model. Their method assumes that this characteristic function is analytically known, which is often the case for a Lévy process ensuring us that it should work in our context since the log-price is driven by a Lévy process.

4.1 Carr & Madan Formula

The purpose of this section is to get analytical expressions of option prices using the Fourier transform in order to use the FFT algorithm on these expressions in the next section to provide a final option price. To this aim, consider k the log-strike of K, $C(K,T)$ the European call option of maturity T and of strike K and $q_T(.)$ the risk-neutral density of the log-price X_T. The characteristic function of the log-price at T and in point u is defined by

$$\Phi_{S_T}(u) = \mathbb{E}\Big[\exp\big(iu\log(S_T)\big)\Big] = \int_{-\infty}^{\infty} e^{ius} q_T(s)\mathrm{d}s, \tag{4.1.1}$$

and the initial Call value $C(k,T)$ is related to the risk-neutral density such as

$$C(k,T) = \int_{k}^{\infty} e^{-rT}(e^s - e^k) q_T(s)\mathrm{d}s. \tag{4.1.2}$$

However, this approach is problematic since allowing the option price to depend on k makes the equation not square-integrable as it tends to a positive constant S_0 if k tends to $-\infty$, implying that we can not use Fourier theory. To solve this, Carr & Madan (1999) proposes to consider the Fourier transform of a modified log-strike function :

$$c(k,T) = \exp(\alpha k) C(k,T) \tag{4.1.3}$$

where fixed $\alpha > 0$. Then, by defining $\psi_T(v)$ as the Fourier transform of $c(k,T)$. We are able to compute the call prices by using the inverse transform

$$\begin{aligned} C(K,T) &= \frac{\exp(-\alpha\log(K))}{2\pi} \int_{-\infty}^{\infty} e^{-iv\log(K)} \psi_T(v) dv \\ &= \frac{\exp(-\alpha\log(K))}{\pi} \int_{0}^{\infty} e^{-iv\log(K)} \psi_T(v) dv, \end{aligned} \tag{4.1.4}$$

[1] A proxy of the spot price in our case as we will see later.

where and takes the following form

$$\psi_T(\upsilon) = \frac{e^{rT}\Phi_{S_T}(\upsilon - (\alpha+1)i)}{\alpha^2 + \alpha - \upsilon^2 + i(2\alpha+1)\upsilon}. \tag{4.1.5}$$

where $\Phi_S(.)$ comes from (4.1.1). By substituting (4.1.5) in (4.1.4) and by resolving the integral, we get the price of the call option. As mentioned above, the FFT will help us to do this and in particular to solve the integral.

4.2 Fast Fourier Transforms

The Fast Fourier Transform is an algorithm that computes the discrete Fourier transform precisely and faster since it allows the complexity of the algorithm to be reduced from $\mathcal{O}(N^2)$ to $\mathcal{O}(N\log(N))$, which is a significant improvement. This algorithm (Schoutens, 2009) is based on a vector α_u to a vector transform β_u where $u = 1, ..., N$, such as :

$$\beta_u = \sum_{j=1}^{N} \exp\Big(-\frac{i2\pi(j-1)(u-1)}{N}\Big)\alpha_j \tag{4.2.1}$$

where N is typically a power of 2 and i the imaginary unit. This relationship will allow us to solve the equation (4.1.4) on N points-grid $(0, \eta, 2\eta, 3\eta, ..., (N-1)\eta)$:

$$C(k,T) \approx \frac{\exp(-\alpha k)}{\pi}\sum_{u=1}^{N}\exp\big(-i\upsilon_j k\big)\psi(\upsilon_j)\eta, \qquad \upsilon_j = \eta(j-1) \tag{4.2.2}$$

We calculate the prices for N log-strikes ranging from $-d$ to d : $k_u = -d + \frac{2d(n-1)}{N}$, and we require that $d\eta = \pi$ it follows :

$$\begin{aligned} C(k_u,T) &\approx \frac{\exp(-\alpha k_u)}{\pi}\sum_{i=1}^{N}\exp\Big(-i\upsilon_j\big(-d+\frac{2d(u-1)}{N}\big)\Big)\psi(\upsilon_j)\eta \\ &= \frac{\exp(-\alpha k_u)}{\pi}\sum_{i=1}^{N}\exp\Big(-\frac{i\upsilon_j 2d(u-1)}{N}\Big)\exp(i\upsilon_j d)\psi(\upsilon_j)\eta \\ &= \frac{\exp(-\alpha k_u)}{\pi}\sum_{i=1}^{N}\exp\Big(-\frac{i2\pi(j-1)(u-1)}{N}\Big)\exp(i\upsilon_j d)\psi(\upsilon_j)\eta \end{aligned} \tag{4.2.3}$$

This last equation is equivalent to (4.2.1). Carr & Madan (1999) use the following values to get their results : $\alpha = 1.5$, $\eta = 0.25$ and $N = 4096$. We will look at these parameters to make sure that they are also relevant in our calculations.

It can be seen that the most important advantage of this way of proceeding is in the fact that one can calculate the prices for a whole range of strikes in one run. The price in between the computed strikes are obtained via interpolation. Note that an additional modification, the Simpson's rule, can be introduced in order to increase the accuracy, as follows:

$$C(k_u,T) = \frac{\exp(-\alpha k_u)}{\pi}\sum_{i=1}^{N}\exp\Big(-\frac{i2\pi(j-1)(u-1)}{N}\Big)\exp(i\upsilon_j d)\psi(\upsilon_j)\frac{\eta}{3}[3+(-1)^j-\delta_{j-1}], \tag{4.2.4}$$

where δ_{j-1} is the Kronecker Delta function and takes 1 if $j=1$ and 0 otherwise.

4.3 Density Function

An alternative way to compute the option pricing through the estimation of the neutral risk density by inversion of the characteristic function (Hainaut, 2017). In order to do so, we start from the classical definition of the European call option :

$$C_t(k,T) = e^{-rT}\int_k^{\infty}(e^s - e^k)f_{X_t}(s,T)\mathrm{d}s, \tag{4.3.1}$$

where k is the log-strike and $f_{X_t}(.,T)$ is the risk-neutral density at maturity T of the log-price X_t observed in t. In order to compute this latter, given the definition of the characteristic function in section 2.2, we get

$$\Phi_{X_t}(iu) = 2\pi\mathcal{F}^{-1}(f_{X_t}) \iff \mathcal{F}^{-1}(f_{X_t}) = \frac{\Phi_{X_t}(iu)}{2\pi} \tag{4.3.2}$$

such that,

$$f_{X_t}(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\Phi_{X_t}(iu)e^{-iux}\mathrm{d}u. \tag{4.3.3}$$

Considering N the number of steps in the DFT, $\Delta_x = \frac{2x_{max}}{N-1}$ the step of discretization and defining $\Delta_z = \frac{2\pi}{N\Delta_x}$ such that,

$$z_j = \Big(-\frac{N}{2}\Delta_z + (j-1)\Delta_z\Big), \tag{4.3.4}$$

and the density f_{X_t}, at points :

$$x_k = \Big(-\frac{N}{2}\Delta_x + (k-1)\Delta_x\Big), \tag{4.3.5}$$

are approximated by

$$f_{X_t}(x_k) \approx \frac{\Delta_z}{2\pi}\exp\Big[-i\frac{N\pi}{2}\Big]\ \exp\Big[i\pi(k-1)\Big]\sum_{j=1}^{N}\delta_j\Phi_{X_t}(iz_j)\exp\Big[i\pi(j-1)\Big]\exp\Big[-i(k-1)(j-1)\frac{2\pi}{N}\Big] \tag{4.3.6}$$

where

$$\delta_j = \frac{1}{2}\mathbb{1}_{\{j=1\}} + \mathbb{1}_{\{j\neq1\}}. \tag{4.3.7}$$

It is now possible to compute the expression in (4.3.1). The density in between the computed densities are obtained via interpolation.

4.4 Monte Carlo

A third and final pricing algorithm that we will propose is Monte Carlo. To do this, we start by defining the method to be used to simulate the price of a process. Since we know that the size of the jumps Y in the log-price follows a double-exponential distribution with a density according to equation (2.6.1), its distribution function is such that

$$F_Y(y) = \begin{cases}(1-p)e^{\eta_2 y} & \text{if } y < 0,\\ (1-p) + p(1-e^{-\eta_1 y}) & \text{if } y \geq 0.\end{cases} \tag{4.4.1}$$

Thus, an easy way to simulate a random variable with a known distribution function $F_Y(y)$, is to inverse it and evaluate it into a uniform law random variable U over the interval $[0,1]$, such that $U \sim Uniform(0,1)$

(Denuit and Charpentier, 2004). Hence, based on a realization u of a random variable U, this boils down to finding the y-solution of the equation

$$F_Y(y) = u, \tag{4.4.2}$$

corresponding to

$$Y = \begin{cases} \dfrac{\log\left(\frac{u}{1-p}\right)}{\eta_2} & \text{if } u < (1-p), \\ \dfrac{\log\left(\frac{p}{1-u}\right)}{\eta_1} & \text{if } u \geq (1-p). \end{cases} \tag{4.4.3}$$

Using this property, we can simulate the entire Lévy process. Starting from the usual expression of the Lévy process in (2.6.2) and reordering the terms to obtain the Lévy-Ito decomposition :

$$L_t = \mu t + \sigma W_t + \sum_{s \leq t} \Delta L_s \mathbb{1}_{\{|\Delta L_s| \geq 1\}} + \lim_{\epsilon \to 0} N_t^\epsilon \tag{4.4.4}$$

where N_t^ϵ represents the restriction of the process itself for jumps size between ϵ and 1, and the compensation term such as :

$$N_t^\epsilon = \sum_{s \leq t} \Delta L_s \mathbb{1}_{\{\epsilon \leq |\Delta L_s| < 1\}} - \int_{\epsilon \leq |z| < 1} z\nu(\mathrm{d}z) \tag{4.4.5}$$

An approximation is possible such that :

$$L_t \approx L_t^\epsilon = \mu t + \sigma W_t + \sum_{s \leq t} \Delta L_s \mathbb{1}_{\{|\Delta L_s| \geq 1\}} + N_t^\epsilon \tag{4.4.6}$$

where the residuals $R_t^\epsilon = L_t - L_t^\epsilon = -N_t^\epsilon + \lim_{\delta \to 0}$ are a jump process with the Lévy triplet $(0, 0, \mathbb{1}_{\{|z|<\epsilon\}}\nu(\mathrm{d}z))$ and with a zero mean, allowing us to re-write :

$$\begin{aligned} L_t^\epsilon &= \mu t + \sigma W_t + \sum_{s \leq t} \Delta L_s \mathbb{1}_{\{|\Delta L_s| \geq 1\}} + \sum_{s \leq t} \Delta L_s \mathbb{1}_{\{\epsilon \leq |\Delta L_s| < 1\}} - t \int_{\epsilon \leq |z| < 1} z\nu(\mathrm{d}z) \\ &= \mu t - t \int_{|z|<1} z\nu(\mathrm{d}z) + \sigma W_t + \sum_{s \leq t} \Delta L_s \mathbb{1}_{\{|\Delta L_s| \geq \epsilon\}} + t \int_{|z|<\epsilon} z\nu(\mathrm{d}z) \\ &= bt + \sigma W_t + \sum_{s \leq t} \Delta L_s \mathbb{1}_{\{|\Delta L_s| \geq \epsilon\}} + \mathbb{E}\Big[\sum_{s \leq t} \Delta L_s \mathbb{1}_{\{|\Delta L_s| < \epsilon\}} \Big] \end{aligned} \tag{4.4.7}$$

where $b = \mu - \int_{|z|<1} z\nu(\mathrm{d}z)$. Since jumps following a double exponential distribution do not have infinite activity, the term $\int_{|z|<1} z\nu(\mathrm{d}z)$ is not necessary and this last expression is easy to simulate, by using the Euler discretization scheme, for a maturity T, we have :

$$L_{t_i} = \mu \Delta t + \sigma \eta \sqrt{\Delta t} + Y \Delta N_{t_i} \tag{4.4.8}$$

for $i = 0, 1, ..., n$, with $t_0 = 0$, $t_n = T$ where n is the number of time steps, $\Delta t = T/n$, Y correspond to the expression in (4.4.3), $\Delta N_{t_i} \sim Poi(\lambda \Delta t)$ and η is a random standard normal, $\mathcal{N}(0, 1)$. We are therefore able to simulate the price of a commodity under the probability measure $\mathbb{Q}$ that will be defined in the next section, by using the Euler discretization scheme:

$$\mathrm{d}S_{t_i} = S_{t_i} \left\{ 1 + \left(r + \beta(\theta - X_{t_i}) + \sum_{u \leq i} M(t_i - t_u) X_{t_u} \Delta t - \lambda \Big(\frac{p\eta_1}{\eta_1 - 1} + \frac{(1-p)\eta_2}{\eta_2 + 1} - 1 \Big) \right) \Delta t + \sigma \eta \sqrt{\Delta t} + (e^Y - 1) \Delta N_{t_i} \right\}, \tag{4.4.9}$$

for $i = 0, 1, ..., n$, with $t_0 = 0$, $t_n = T$ where n is the number of time steps, $\Delta t = T/n$ and $M(.)$ is the kernel function. We are now able to calculate the price of a European call (and put) option by simulating this process m times and averaging the payoff of each of these simulations:

$$C_t(K,T) = \mathbb{E}^{\mathbb{Q}}\left[e^{-r(T-t)}(S_T - K)_+|\mathcal{F}_t\right] = e^{-r(T-t)}\ \frac{1}{m}\sum_{k=1}^{m}(S_T^k - K)_+, \tag{4.4.10}$$

where S_T^k is the k-th simulation of the commodity price S at maturity T computed under the measure $\mathbb{Q}$.

Chapter 5

Change of Measure

In order to price a derivative in an arbitrage-free setting, we need to deal with the risk-neutral probability measure. Pascucci (2011) defined the first and second Fundamental Theorem of asset pricing, respectively as :

- *« A discrete market is arbitrage-free if and only if there exists at least one equivalent martingale measure. »*
- *« An arbitrage free market (S,B) is complete if and only if there exists a unique equivalent martingale measure with numeraire B. »*

Thus, in a complete market, with a single probability measure, we have unique arbitrage-free price dynamics. Equivalently, any asset in this market is replicable. Although this feature is highly desirable, in practice many financial markets are not, and in particular the commodity market, is not. For most commodities the spot market does not exist, or the futures/forwards markets are much more traded and therefore much more liquid, or they are markets that are only available to physical economic agents. Since we are facing very incomplete markets in the sense that the underlying spot cannot be traded, Benth (2020) suggests a different measure from the risk-free one, in order to deal with the specific features of the commodity market. Indeed, in practice, the operators of this kind of market often only speculate and never really hold the commodity. Thus, they agree to pay a price that differs from the one discounted at the risk-free rate since they do not have to deal physically with the commodity. The incompleteness is justified by these markets frictions and also by the non-Gaussian feature in the price dynamics. Then there is no unique risk-neutral probability measure in this context.

Contrary to standard financial products (stock market indices, share prices, etc.), commodities do not require the payment of a dividend, but costs related to the storage of the commodities, insurance, warehousing, etc. The person who holds the commodity therefore faces additional costs that are often considered as an additional interest rate to be paid. It should be noted, that holding the commodities can however present a major advantage for the negotiation of a contract since the seller has more flexibility by having the commodities at his disposal. This advantage explains why we speak of « convenience yield » when we mention commodities.

Therefore, this convenience yield makes it possible to justify the difference between the capitalized spot price and the price actually paid to buy the commodity in the future in the sense that the underlying assets will now have an average rate of return corresponding to the sum of the risk-free rate, but also an additional amount that can be associated with a risk premium. We introduce the additional premium that the holder of the commodity has to pay by $(\rho_t)_{0\geq t\geq T}$ to assume the risk, grouping storage costs, illiquidity and/or convenience yield.

However, although in theory we should associate this convenience yield with market storage costs, this assumption is rarely valid in practice. Therefore, Benth suggests that we investigate the possibility that this amount could be associate with memory terms evolving over time and taking into account the past evolution of the asset price.

In order to do so and to price options, we need to know the dynamics of S under a pricing measure $\mathbb{Q}$ which deviates from the risk neutral one in such a way that the asset price has a return corresponding to the risk-free rate and the convenience yield. Then, the dynamics of the discounted price, which is supposed to be a martingale when the risk-neutral measure is unique, is no longer martingale in this case, which is directly interpreted as a risk premium. Thus, let us begin by applying Ito's theorem, which can be found in the appendix to equation (9.3.3) on the equation of the process (3.0.4), in order to find the SDE that characterizes the price of the commodity under the measure $\mathbb{P}$. Without loss of generality in our case, we consider $\chi(t-) = 1$. We define $f(t, X) = e^X$ and it follows

$$\frac{\mathrm{d}S_t}{S_t} = \left(\beta\theta - \beta X_t + \int_0^t M(t-u)X_u \mathrm{d}u + \mu + \frac{\sigma^2}{2} + \lambda\mathbb{E}[e^Y - 1 - Y]\right)\mathrm{d}t + \sigma \mathrm{d}W_t + (e^Y - 1)d\tilde{N}_t \tag{5.0.1}$$

and by defining the numeraire of this pricing measure such that

$$B_t := \exp\left\{\int_0^t r_s \mathrm{d}s\right\} \tag{5.0.2}$$

we get the discounted price $\tilde{S}_t = \exp(-\int_0^t r_s \mathrm{d}s)S_t$ with $t \in [0, T]$. Considering the Girsanov theorem in Øksendal and Sulem (2009), such that $\mathrm{d}\mathbb{Q} = Z(T)\mathrm{d}\mathbb{P}$ with density process

$$\begin{aligned} Z(t) := \exp\Big\{ -\int_0^t \varphi(s, S)\mathrm{d}W_s - \frac{1}{2}\int_0^t \varphi(s, S)^2 \mathrm{d}s + &\int_0^t \int_{\mathbb{R}} \log(1 - \vartheta(s-, z))\tilde{N}(\mathrm{d}s, \mathrm{d}z) \\ &+ \int_0^t \int_{\mathbb{R}} \big[\log(1 - \vartheta(s, z)) + \vartheta(s, z)\big]\nu(\mathrm{d}z)\mathrm{d}s\Big\} \end{aligned} \tag{5.0.3}$$

where

$$\begin{aligned} \varphi(t, S) &= \frac{1}{\sigma}\Big(\beta\theta - \beta X_t + \int_0^t M(t-u)X_u \mathrm{d}u + \mu + \frac{\sigma^2}{2} - r_t - \rho_t\Big) \\ \vartheta(t, z) &= \frac{e^z - 1 - z}{e^z - 1} = 1 - \frac{z}{e^z - 1} \end{aligned} \tag{5.0.4}$$

where

$$\rho_s = \beta(\theta - X_s) + \int_0^s M(s-u)X_u \mathrm{d}u. \tag{5.0.5}$$

When the measure $\mathbb{Q}$ exists, we have

$$\begin{aligned} \mathrm{d}W_t^{\mathbb{Q}} &= \varphi(t, S)\mathrm{d}t + \mathrm{d}W_t \\ \tilde{N}^{\mathbb{Q}}(\mathrm{d}t, \mathrm{d}z) &= \vartheta(t, z)\nu(\mathrm{d}z)\mathrm{d}t + \tilde{N}(\mathrm{d}t, \mathrm{d}z) \end{aligned} \tag{5.0.6}$$

which are a Brownian motion and compensated jump measure under $\mathbb{Q}$, respectively, where the compensator is $\tilde{\nu}_t(dz)\mathrm{d}t = (1 - \vartheta(t, z))\nu(\mathrm{d}z)\mathrm{d}t$. The dynamics of the discounted commodity price under the pricing measure $\mathbb{Q}$ becomes

$$d\tilde{S}_t = \tilde{S}_t\left\{\rho_t \mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} + \int_{\mathbb{R}} (e^z - 1)\tilde{N}^{\mathbb{Q}}(\mathrm{d}t, \mathrm{d}z)\right\}, \tag{5.0.7}$$

and the commodity price

$$\mathrm{d}S_t = S_t\left\{\left(r_t + \beta(\theta - X_t) + \int_0^t M(t-u)X_u \mathrm{d}u\right)\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} + \int_{\mathbb{R}} (e^z - 1)\tilde{N}^{\mathbb{Q}}(\mathrm{d}t, \mathrm{d}z)\right\}, \tag{5.0.8}$$

which has a return equal to the sum of the risk-free rate and of the convenience yield as expected.

Before pursuing, it is important to mention that in the following sections we will benchmark the models by the Vasicek and Kou models. These models are chosen because they will allow us to see the influence of jumps on the calibration and pricing of options through Vasicek's model since the latter has a mean reversion term (like the Delta and Exponential kernel models) and thus an instantaneous influence of its past dynamics (like the Delta kernel model) but does not have any jump term. On the other hand, Kou's model seems to be relevant since it does include jumps but does not have a mean reversion term and thus no memory (even instantaneous) in its dynamics. With these two models, we will be able to modify one characteristic of our models at a time in order to see the effect of this characteristic on the price of a call for example. However, it is first necessary to express them under our pricing measure that we have just defined in order to show that these two models are compatible with this risk measure in order to « compare comparable values ». First, the Vasicek model, which does not include any jumps, under the $\mathbb{P}$ measure is defined as

$$\mathrm{d}X_t = \beta(\theta - X_t)\mathrm{d}t + \sigma \mathrm{d}W_t \tag{5.0.9}$$

and becomes under the pricing measure $\mathbb{Q}$ with $a = 0$ for each of the two kernels $\big(M(t) = a\delta_0(t)$ and $M(t) = ae^{-bt}\big)$ such that

$$\mathrm{d}X_t = r_t \mathrm{d}t + \beta(\theta - X_t)\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} \tag{5.0.10}$$

Then, the Kou model is developed under the measure $\mathbb{P}$ such that

$$\mathrm{d}X_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t + Y \mathrm{d}N_t \tag{5.0.11}$$

and becomes under the pricing measure $\mathbb{Q}$ with $a = \beta = 0$ for each of the two kernels $(M(t) = a\delta_0(t)$ and $M(t) = ae^{-bt})$ such that

$$\mathrm{d}X_t = \Big\{r - \frac{\sigma^2}{2} - \lambda\Big(\frac{p\eta_1}{\eta_1 - 1} + \frac{(1-p)\eta_2}{\eta_2 + 1} - 1\Big)\Big\}\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} + Y \mathrm{d}N_t^{\mathbb{Q}}. \tag{5.0.12}$$

Fortunately, by using the Ito lemma, we get the dynamic of the asset price under the pricing measure $\mathbb{Q}$

$$\mathrm{d}S_t = \Big\{r - \lambda\Big(\frac{p\eta_1}{\eta_1 - 1} + \frac{(1-p)\eta_2}{\eta_2 + 1} - 1\Big)\Big\}\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} + (e^Y - 1)\mathrm{d}N_t^{\mathbb{Q}}. \tag{5.0.13}$$

We refer to Annex 9.7 and 9.6 for more information on the measurement changes for these two models.

Chapter 6

Implied Volatility

We now introduce the notion of implied volatility using the well-known Black and Scholes model. For each of the call option prices available on the market at a fixed maturity T and for a certain strike K, we can calculate the unique parameter $\sigma(T, K)$ that allows the theoretical price (of the B&S model) to match with the empirical price. This volatility is called the implied volatility and is defined as the following relation is fulfilled :

$$C^{B\&S}(S_0, K, r, T-t, \hat{\sigma}_t^{imp}) = C_t^{market} \tag{6.0.1}$$

where C_t^{market} is the market price of the call option at time t and $C^{B\&S}$ is the European option call defining in the Black & Scholes framework under the risk-neutral measure $\mathbb{Q}$, given by

$$C^{B\&S}(S_t, K, r, T-t, \sigma) = S_t N(d_1) - Ke^{-r(T-t)} N(d_2) \tag{6.0.2}$$

where $\mathcal{N}(.)$ is the standard Normal CDF, $d_1 = \frac{1}{\sigma\sqrt{T-t}}\big[\log(\frac{S_0}{K}) + (r + \frac{\sigma^2}{2})(T-t)\big]$ and $d_2 = d_1 - \sigma\sqrt{T-t}$. However, it should be mentioned that there is no closed formula for inverting the B&S call price formula and that only numerical methods can be used such as the Newton-Raphson algorithm. Then, this volatility can be calculated for different strikes and different maturities to obtain an implied volatility « surface ».

It is also possible to obtain implied volatility through the Carr-Madan formula. Indeed, starting from the equation 4.1.5, and adapting $\psi_T(v)$ with the characteristics of the B&S characteristic function, which gives

$$\psi_{\text{BS},T}(v) = \frac{\exp\Big(-rT + iv(\log S_0 + rT)\Big)\ \exp\Big(-\frac{1}{2}(v-(\alpha+1)i)(v-\alpha i)\sigma_{\text{BS}}^2 T\Big)}{\alpha^2 + \alpha - v^2 + i(2\alpha+1)v} \tag{6.0.3}$$

with σ_{BS} is the B&S volatility. Then, since we know that we must have $C^{B\&S}(S_0, K, r, T-t, \hat{\sigma}_t^{imp}) = C_t^{market}$, we can deduce the following relationship :

$$\int_0^\infty \exp\big(-ivk\big)\Big(\psi_T(v) - \psi_{\text{BS},T}(v)\Big)\mathrm{d}v = 0. \tag{6.0.4}$$

Solving this equation with respect to σ_{BS} using the FFT algorithm allows us to numerically derive the implied volatility surface for the different models introduced previously.

Under the assumptions of the Black & Scholes model, this volatility surface should be flat, but in practice we observe that this condition is often violated since we observe rather a « smile », i.e. a convex function on the strike axis, or a « skew », i.e. a decreasing pattern on this axis as we can see on the figure 6.0.1. Note that the smile thus implies a minimum on the implied volatility curve which often occurs when K is close to the forward price at $t = 0$, when we are at-the-money. Some financial assets show a « frown » in their implied volatility, i.e. an inverse smile but this is not common and does not concern us. This shows that this model does not correspond to the reality of the market. One can imagine that when we fix the maturity, we go from a 3-dimensional to a 2-dimensional plan such as : $K \mapsto \hat{\sigma}^{imp}(T, K)$, as shown in Figure 6.0.1.

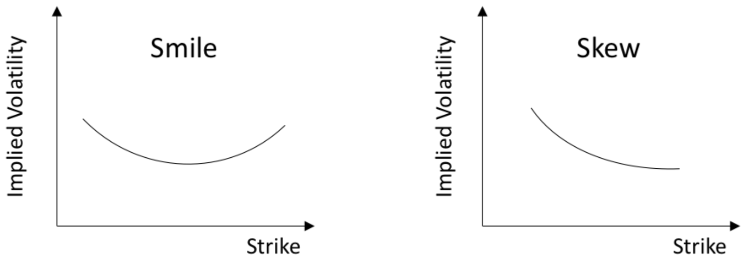

Figure 6.0.1: Smile and skew of the implied volatility surface for a certain maturity.

Since Black Monday in 1987, OTM puts have become much more attractive to buyers due to the possibility of making a huge profit. In addition, these options have become essential as portfolio insurance in the event of a market decline. This has led to a significant increase in demand on a permanent basis resulting in higher prices and thus higher implied volatility. It is this effect that explains why the « smile » of implied volatility has now been replaced by a « skew ». Thus, implied volatility is significantly higher for OTM options with a lower strike price than the underlying, while implied volatility is significantly lower for OTM options with a higher strike price than the underlying.

The implied volatility surface is studied over a maturity interval ranging from a few weeks to one year and for a moneyness (Strike/Price) around 1, i.e. where the option is most liquid. Tankov (2003) mentions some features regarding the study of implied volatility in exponential-Lévy models and these will also be relevant for the study of our processes as we will see later :

- For quite low maturities, when we are out of the money (moneyness is < 1), we observe a skewness pattern and implied volatilities are lower than those in the money. This effect is enhanced for models with jumps in the price.
- Processes with negatively skewed jump distributions present a skew in their implied volatility surface. However, if the variance of this jump distribution is substantial, it tends to generate a smile in the implied volatilities.
- In general, these effects tend to disappear as the maturity increases and we observe a flattening of the surface. This is mainly explained by the effect of the central limit theorem which implies that in the case of a Lévy process with finite variance, as the maturity increases, the distribution of $(X_T - \mathbb{E}[X_T])/\sqrt{T}$ tends towards a normal distribution. This implies that when the maturity grows tremendously, $T \to \infty$, the implied volatility becomes flat and the underlying normality assumption of the Black & Scholes model is satisfied.

Moreover, as we shall see, many commodities are associated with mean-reverting dynamics, a characteristic that tends to introduce, on the one hand, convexity in the implied volatility surface as maturity increases and, on the other hand, volatility inversely proportional to moneyness. It will be interesting to observe whether or not all these elements are respected in the practical application.

Chapter 7

Practical Applications

In this chapter we shall now turn to the practical study of different commodity markets. First, we will confirm the influence of past evolution on the price of each of them. Then, we will present the procedure to calibrate the Delta kernel model and the Exponential kernel model on these different markets. Once done, we will analyse the results and compare the models retained. We will finally move on to the pricing of vanilla derivatives in a first step and exotic derivatives in a second step.

7.1 Market Data

As explained in the chapter on measurement change, for most commodities the spot market does not exist, or the futures/forwards markets are much more traded and therefore much more liquid, or they are markets that are only available to physical economic agents. Based on this observation, it is not possible to consider the spot price directly in this practical study, but rather we will use the futures market for each of the commodities. Indeed, futures (generally for the first month) are often referred to as a proxy for the spot price and moreover, the futures price is directly modelled/influenced by the dynamics of the spot price whether it is available or not since we know that the prices of these contracts converge to the spot price when approaching maturity as follows

$$\lim_{t \to T} F(t, T) = S_T. \tag{7.1.1}$$

This assumption would imply that we are looking at real-world projections, whereas if we consider this price as a future price contract, we would be dealing with projections under the risk-neutral measure (or more precisely under the pricing measure in our case). We prefer the conservative assumption and we will therefore use the futures price as a proxy for the spot price of the different commodities, however, we will misemploy the language by confusing the two names in the following.

We will calibrate the different models introduced previously on the data available on the commodities market. We consider the following four markets for our study: Gold, Copper, Corn and Oil. The data used are the historical prices on the Yahoo Finance platform of the different commodities over the last 5 years (1245 data): from November 2015 to November 2020. We will calibrate our models to determine whether the various characteristics we have introduced are consistent with market dynamics. We will also study more « standard » models to use them as a benchmark and examine whether or not our models add value to the analysis of the empirical data. The method used will be the same for each model. The objective is to compare the results obtained by the different models in order to highlight the one that fits the best with the market. We expect to obtain parameters to be able to price in the next chapter different financial products according to the different price dynamics introduced so far. Figure 7.1.1 shows the prices of the different futures for each commodity.

It should be noted that the Oil market experienced abnormal fluctuations in April 2020 due to market friction caused by the coronavirus crisis. However, our models do not, by definition, allow for negative prices. Thus, for the sake of the analysis, we will consider that this was an exceptional case and we

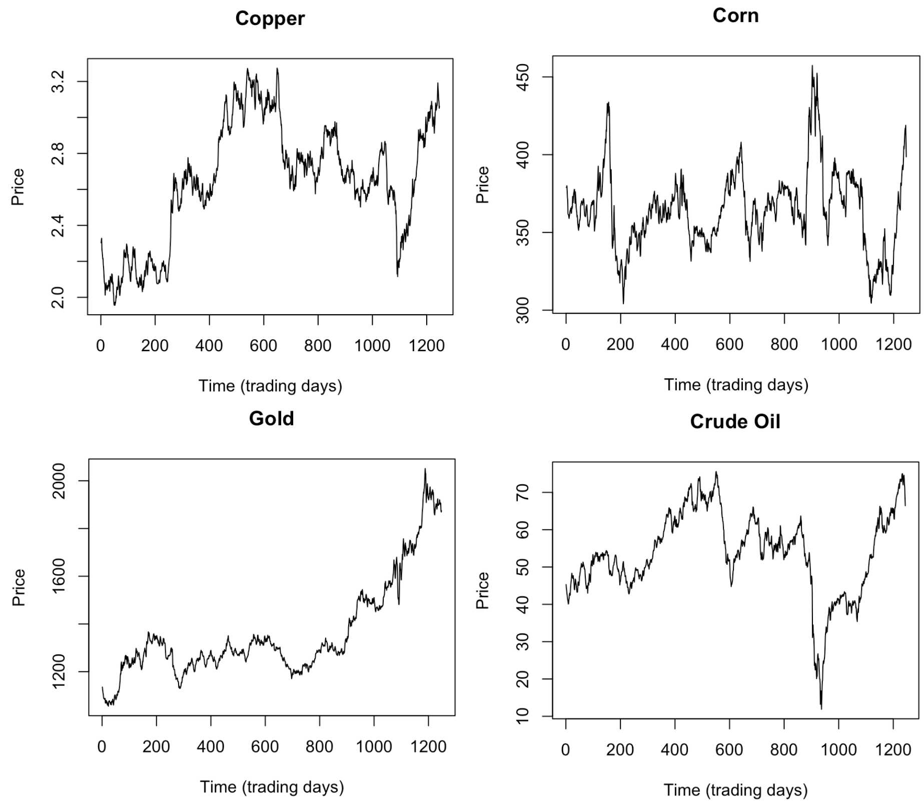

Figure 7.1.1: Prices of the different commodities markets studied. They are all reported in USD except for Corn, which is reported in US cents.

will not include these days in the analysis (two trading days) so that this irregularity does not bias the results.

7.2 Past History in The Market

There are different definitions widely used to describe the long-term dependence of a process, but the two most common are the following. A stochastic process is said long-range dependent (LRD), which describes systems with memory and persistence to past events,

- if the autocorrelation[1] of the process are not summable, such that,

$$\sum_{h=0}^{\infty} |\rho(h)| = \infty. \tag{7.2.1}$$

  In other words, a random process is said to have a long-memory if it has an autocorrelation[1] function that is not integrable.

- if the autocovariance function of the process meets the condition :

$$\gamma(h) = L(h)h^{2d-1}, \quad h = 0, 1, ..., \tag{7.2.2}$$

  where $d \in (0, 0.5)$, refers to the slow decay[2] of the ACF function and where $L(.)$ is a slowly varying function at infinity, i.e. if it is positive on $[b, \infty)$ with $b \geq 0$ and for any $a > 0$:

$$\lim_{u \to \infty} \frac{L(au)}{L(u)} = 1. \tag{7.2.3}$$

[1] Some references rather mention the autocovariance.
[2] A slower decay than an exponential, i.e. a decay like a power-law with an exponent smaller than 1.

There is an long-standing debate about whether commodities are stationary or not. For example, Samuelson (2016) insists that there is stationarity for economic reasons while Barndorff-Nielsen (2013) shows that although there is stationarity in spot energy prices, futures and forwards prices are not constant in the long run which should be a consequence of a stationary spot. Other empirical studies on various indices/stocks have shown that the autocorrelation function remains positive and slowly decreases while remaining significantly positive even after several weeks. Volterra models (including memory) driven by a Lévy process as introduced earlier are popular for energy, gas and Oil markets (see Benth & al. (2014), Barndorff-Nielsen & al. (2013), Paschke and Prokopczuk (2010)).

Thus, in order to account for this characteristic, Schwartz and Smith (2000) include a non-stationary term in the evolution of the commodity price. Ortiz-Latorre (2017) proposes a second possibility: to consider a change of measure from the spot measure $\mathbb{P}$ to a non-stationary measure $\mathbb{Q}$, which is the same, which is in line with what Benth & al. (2020) recently proposed in their paper and is also the idea that is supported in the writing of this thesis as explained in the chapter on measurement change. We will use these considerations to build models to account for this non-stationarity in the long run.

Hurst (1951) introduces the classical rescaled range test (R/S test) which seeks to compare the minimum and maximum values of running sums of deviations from the sample mean, renormalized by the sample standard deviation. In fact, we observe that the deviations are larger for long-memory processes than for the short ones. However, Lo (1991) shows that this test is too weak and assumes too often that a time series is a long memory process. Then, he introduces a modified version of the test which is stronger. Consider a sample time series X_1, X_2, ..., X_n with $\bar{X}_n$, $\hat{\sigma}_x^2$, $\hat{\gamma}_x$, the sample mean, the sample variance and the autocovariance, respectively. The modified rescaled range statistic $Q_n(q)$ is defined by

$$Q_n(q) = \frac{1}{\hat{\sigma}_n(q)} \left[\max_{1\leq k\leq n} \sum_{j=1}^{k} (X_j - \bar{X}_n) - \min_{1\leq k\leq n} \sum_{j=1}^{k} (X_j - \bar{X}_n) \right], \tag{7.2.4}$$

where

$$\sigma_n^2(q) = \hat{\sigma}_x^2 + 2\sum_{j=1}^{q} w_j(q)\hat{\gamma}_j, \qquad w_j(q) = 1 - \frac{j}{q+1}, \tag{7.2.5}$$

with $q < n$. It is noteworthy that only the denominator changes between the classical and modified versions of the test. In the classical version, the denominator is the sample variance $\hat{\sigma}_x^2$. The result is sensitive to the value of q that is chosen, Lo suggests the value $q = [k_n]$, with

$$k_n = \left(\frac{3n}{2}\right)^{\frac{1}{3}} \left(\frac{2\hat{\rho}}{1-\hat{\rho}^2}\right)^{\frac{2}{3}}, \tag{7.2.6}$$

where $[x]$ is the greatest integer less than or equal to the value x and $\hat{\rho}$ is the sample first-order autocorrelation coefficient of the data. Lo informs us that the null hypothesis can be rejected at the 5% threshold when the test statistic $Q_n(q) > 1.747$. In fact, Montanari & al. (1999) have shown that this test is too severe and cannot reject the null hypothesis of short range dependence with moderate value of the Hurst exponent ($H \approx 0.6$). This leads us to an interest in fractional calculus. However, this is not the subject of this thesis and there is already a very detailed literature on this topic. We will therefore stick to the basics. Long memory is indeed discussed in terms of the Hurst exponent: a positively correlated process with long memory is defined by a Hurst exponent in the interval between 0.5 and 1. It is known that for a Gaussian process which, by definition, has increments that do not display long memory, the standard deviation increases asymptotically as $t^{\frac{1}{2}}$. On the contrary, diffusion processes with long memory increments have a standard deviation that increases asymptotically as $t^H L_t$, with $1/2 < H < 1$, and L_t a slow-varying function. Then when the exponent is $0 < H < 1/2$, the process has a very short memory (it is very excited and will generate very rough trajectories), when $H = 0.5$, it is equivalent to a classical random walk (Brownian motion) and when $1/2 < H < 1$, the process is long-range dependent and it will be much smoother. In short, this exponent is a measure of the extent of long-range dependence in a time series.

The greater the strength of the reversion to the mean, the more H tends towards 0 (anti-persistent). On the contrary, when the time series is persistent, it is, by definition, trend reinforcing, which implies that the direction of the next value is generally in the same direction as the current value. As the strength of the trend increases, H tends towards 1. Most financial time series are persistent ($H > 0.5$) which implies a strong trend. As explained, commodities are often associated with mean-reverting trajectories ($H < 1/2$) whereas, as we will see, different commodities seem to have a strong trend and long-range dependency leading to higher H ($> 1/2$). As explained earlier, it can be seen that mean reversion and long range dependency seem to be two contradictory features and this may therefore seem to be a weakness of the modelling but may also be a strength in taking into account antinomian movements since it can allow to combine the effect. Moreover, as we shall see, different models and different types of memory are studied, which makes it possible to include one or the other characteristic and not to force the model to require both. If we look at Figure 7.1.1, we can see that only the price of Gold seems to show a real trend in its evolution over time. The other markets seem to be more stationary. It will therefore be interesting to see whether this observation is justified and verified in the study of future models.

However, determining the Hurst exponent is not an easy task and there is a wide range of tests and methods for estimating it, including the periodogram method, the R/S method and the fit of the autocorrelation function. It is interesting to study each of these methods in order to obtain a critical opinion on the exact value of the exponent. Again, we will not go into detail about each method and will only relate the R/S test to the Hurst exponent. Indeed, we know that the latter is defined as the asymptotic behaviour of the rescaled range :

$$\mathbb{E}\Big[\frac{R(n)}{S(n)}\Big] = Cn^H, n \to \infty \tag{7.2.7}$$

where $Q_n = R(n)/S(n)$, C is a constant and n is the time span of the observation. Then the exponent estimation can be done by regressing linearly $\log\big(\mathbb{E}[R(n)/S(n)]\big)$ on $\log(n)$ since

$$\log\Big(\mathbb{E}\Big[\frac{R(n)}{S(n)}\Big]\Big) = H\log(n) + \log(C). \tag{7.2.8}$$

All these methods allow us to learn more about the different markets we are going to consider: Gold, Copper, Oil and Corn. We will try to find out more about the type of memory each market has. To do this, Table 7.2.1 shows the results of the various tests for each market.

Commodity Market	R/S test	Rescaled R/S test	Empirical Hurst	H: autocorrelation
Gold	2.583	2.111	0.787	0.873
Copper	1.689	1.338	0.598	0.623
Corn	1.466	1.049	0.452	0.587
Oil	1.995	1.812	0.746	0.837

Table 7.2.1: Review of the memory in the price of different markets.

First of all we notice that the classic R/S test is less strict than the rescaled version. However, we notice that the conclusions are almost the same: we reject the long memory property in both cases for the Corn and Copper market. The other two columns allow us to judge that there is a weak memory for the Copper market. We observe that the Corn market even seems to be anti-persistent ($H < 1/2$). Conversely, the Gold and Oil markets appear to have clear long memory characteristics in their structure. At first glance, it seems that the Oil and Gold markets will be more likely to be driven by long-range dependency, while the Corn and Copper markets will be more likely to be mean-reverting. The third column is obtained by performing a linear regression such as in (7.2.8). The last column is obtained by fitting the d parameter seen as a negative power of the empirical autocorrelation according to the number of lags: $\rho(k) \sim k^{2d-1}$.

7.3 General Method

First, let us make a point of vocabulary: the difference between model estimation and calibration[3]. Model estimation is the process of choosing the best model type and structure, given certain data. This step may include calibration. On the contrary, calibration is the process of finding the coefficients that allow a model (whose type and structure are already determined) to reflect most accurately a particular known data set. For financial domains, estimation is often performed on historical data (real measure of probability) while calibration is performed on pricing data (risk-free measure of probability). The latter is therefore based on less uncertain data (financial derivatives, implied volatility matrix, etc.) rather than on historical data (price of the underlying over time). Indeed, future price dynamics are difficult to predict with real past data: who knows how much the price of Oil will be worth in 3 years? However, commodity derivatives data is not available online or at least not for free, so we have to fall back on historical data. Furthermore note, that the derivative pricing chapter of this work uses a different probability measure than the classical risk-neutral measure, so it will be necessary to keep in mind that the parameters obtained characterize the real dynamics of the underlyings and not the parameters related to the pricing under another probability measure. It will therefore be necessary to introduce a risk premium (and a convenience yield) later on when considering pricing as explained in the chapter on the change of measure.

The parameters will be estimated by log-likelihood maximisation. However, this first requires some modifications and treatments on the process in order to obtain less time consuming results. Then, the approach used to calibrate is as follows:

1. Firstly, it is needed to compute the log-returns of the process such that $X_j = \log(P_j/P_0)$ where P_j is the historic price of the commodity on the j^{th} day. We now consider that X_j follows the dynamics of the models introduced earlier.

2. Secondly, we use the respective autocorrelation expressions for each model that we developed earlier to capture the non-Markovian feature of the process. We use here the assumption introduced in the previous part, which assumes that the mean of the diffusion term is zero in order to simplify the calculations. Note that this assumption is only slightly restrictive and does not change significantly the results as we will see later. These relations are used to obtain the parameters linked to this autocorrelation: β and a in the case of the Delta kernel and a, b and β in the case of the negative Exponential kernel. They are obtained by squared error minimisation on the autocorrelation of the empirical financial data obtained using the *acf*[4] function in R. The loss function used for this purpose is the weighted Root Mean Square Error the function that will compute the difference between the autocorrelation in the market and the autocorrelation computed with the model, the aim being to minimise this distance. It takes the following expression :

$$RMSE = \sqrt{\sum_{j=1}^{N} \frac{1}{N}(\rho_j^{market} - \hat{\rho}_j^{model})^2}. \tag{7.3.1}$$

The RMSE has the advantage of the same unit as what we measure and can therefore be more easily interpreted. Note that we have tried to use other loss functions such as the Mean Absolute Error or the Mean Absolute Percentage Error, but the results are essentially the same. We can define the optimal set $\Theta_1^* = (a^*, \beta^*)$ and $\Theta_1^* = (a^*, b^*, \beta^*)$, respectively, which contains the different values of the parameters allowing to reproduce the features of the current[5] market in the best possible way. Thus, we try to find :

$$\Theta_1^* = \arg\min_{\Theta_1} \mathcal{L}(\rho_j^{market}, \hat{\rho}_j^{model}(\Theta_1)) \tag{7.3.2}$$

where $\mathcal{L}$ is the RMSE in (7.3.1). Note that since the function used in R is sensitive to the starting points chosen for each of the parameters, thus it is necessary to have some knowledge of the range

[3]It should be noted, however, that in this work we make the mistake of confusing calibration with estimation.

[4]The function has commands to adjust the way of computing and in particular to remove the de-meaning terms, which is our case, given the assumption done previously.

[5]Between November 2015 and November 2020.

of parameters and/or use a grid running through different values in order to find the initial values to minimise the loss function.

3. Thirdly, we can now subtract the mean-reverting term so that we define the process Y_j for the Delta kernel as
$$Y_j = X_j - \Big(\frac{\beta^*\theta}{\beta^* - a^*}\Big)\big(1 - e^{-(\beta^*-a^*)t_j}\big), \tag{7.3.3}$$
and for the Exponential kernel as,
$$Y_j = X_j - \frac{\beta^*\theta}{s_1^* - s_2^*}\Big\{\frac{(s_1^*+b^*)(e^{s_1^* t_j}-1)}{s_1^*} - \frac{(s_2^*+b^*)(e^{s_2^* t_j}-1)}{s_2^*}\Big\}. \tag{7.3.4}$$
However, these relationships do not yet allow the parameter to be estimated precisely, as we shall see in the following point.

4. Fourthly, by defining t the number of years at the time the price $j-1$ is observed and Δ_t as one trading day and according to equation (3.1.5), we can define the process Z_j for the Delta kernel as
$$\begin{aligned} Z_j &= Y_j - Y_{j-1}e^{-(\beta^*-a^*)\Delta_t} \\ &= \int_0^{t+\Delta t} e^{-(\beta^*-a^*)(t+\Delta_t-u)}\mathrm{d}L_u - \int_0^t e^{-(\beta^*-a^*)(t-u)}\mathrm{d}L_u \\ &\approx \int_0^{\Delta_t} e^{-(\beta^*-a^*)(\Delta_t-u)}\mathrm{d}L_u \end{aligned} \tag{7.3.5}$$
and for the Exponential kernel as,
$$\begin{aligned} Z_j &= Y_j - Y_{j-1}\left(\frac{1}{s_1^*-s_2^*}\times\Big[(s_1^*+b^*)e^{s_1^*\Delta_t} - (s_2^*+b^*)e^{s_2^*\Delta_t}\Big]\right) \\ &= \frac{1}{s_1^*-s_2^*}\times\left\{\int_0^{t+\Delta_t}\Big((s_1^*+b^*)e^{s_1^*(t+\Delta_t-u)} - (s_2^*+b^*)e^{s_2^*(t+\Delta_t-u)}\Big)\mathrm{d}L_u\right. \\ &\qquad\left. - \int_0^t \Big((s_1^*+b^*)e^{s_1^*(t-u)} - (s_2^*+b^*)e^{s_2^*(t-u)}\Big)\mathrm{d}L_u\right\} \\ &\approx \frac{1}{s_1^*-s_2^*}\times\int_0^{\Delta_t}\Big((s_1^*+b^*)e^{s_1^*(\Delta_t-u)} - (s_2^*+b^*)e^{s_2^*(\Delta_t-u)}\Big)\mathrm{d}L_u. \end{aligned} \tag{7.3.6}$$
We see that this relationship is obtained by a simplification and the introduction of an assumption: the stationarity of the increments over time. This corresponds in reality to assuming that each day corresponds to the first observed price X_0. This is an important constraint in the context of this framework: the study of memory processes with time dependence. It is for this reason that the interest of autocorrelation is essential in order to capture the long term effect of the process.

We can now estimate the parameter θ. Once again, we use the minimisation of the squared errors in order to compute the new optimal set $\Theta_2^* = (\theta^*)$ by using these equations. In this particular case, for simplifications of the computation of the autocorrelation and the computation time, we will assume that the mean of the Lévy term is null, we will estimate the minimization of this RMSE of the process Z_j around zero. One can notice that it was not possible to carry out this centering around 0 before knowing this process Z_j.

5. Fifth and last, it is now possible to numerically invert the characteristic function of the integral term of the two last expressions in order to reconstruct the density of this process Z_j. These characteristic functions are given in equation (3.1.12) and (3.2.11) respectively and the algorithm

allowing us to invert these functions in section 4.3. In order to estimate these optimal parameters, we maximise the log-likelihood on this distribution with

$$\Theta_3^* = \arg\max_{\Theta_3} \ \log \hat{f}(\Theta_3; \mathbf{Z}), \tag{7.3.7}$$

where $\mathbf{Z} = \{Z_1, Z_2, ..., Z_n\}$ is the observations set obtained at the previous point and $\hat{f}(\Theta_3; \mathbf{Z})$ is the density function evaluated at the points $\mathbf{Z}$ for the parameters set $\Theta_3 = (\lambda, p, \eta_1, \eta_2, \sigma)$. The final optimal set of parameters is such that $\Theta^* = \left\{\Theta_1^*, \Theta_2^*, \Theta_3^*\right\}$.

7.4 Calibration of Delta Kernel

Following the steps presented in the previous section, in order to obtain an optimal set of parameters, we first calculate the observed autocorrelation of the respective markets and try to fit the corresponding autocorrelation from the model with the Delta kernel. The results are presented in Figure 7.4.1 and Table 7.4.1. The RMSE in the table is the one obtained from the parameter values also in the table.

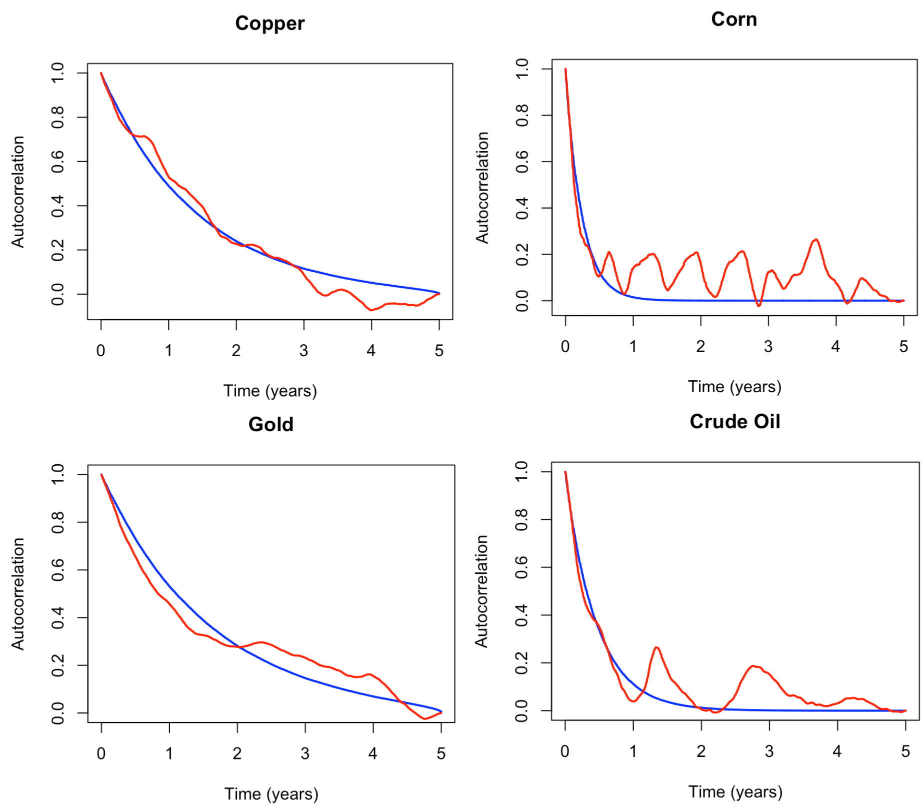

Figure 7.4.1: Empirical (red) and calibrated (blue) autocorrelation for each commodity market.

A first observation is that the plot of the autocorrelation for the Gold and Copper markets seem to be quite close to the empirical autocorrelation, which encourages us to study this model for these two markets. On the opposite, the Crude Oil and Corn markets do not seem to fit this model, which only allows for Exponential autocorrelation. However, we want to point out that, for example for the Gold market, under the conditions of this model, the parameters β and a satisfy only one equation: $\beta - a \approx 0.63$, and since we have 2 unknowns for a single equation: we have an infinite number of possible solutions. Thus, we notice here that it is possible to replace these two parameters by one and we fall back on a configuration introduced by the Ornstein–Uhlenbeck model. This is consistent with the intuition that one can have since these two models satisfy an instantaneous influence of the past evolution (Markov property). We note that all the parameter β is positive and larger than the parameter a, which makes it possible to obtain a negative Exponential and thus a convergence of the mean-reversion term. Moreover, we observe that the parameter a is negative for the Corn market but this has no real impact on the interpretability

Commodity Market	Model	RMSE	a	β
Gold	Delta Kernel	0.0647	0.9361	1.5639
	Vasicek	0.0647	\	0.6278
Copper	Delta Kernel	0.0587	1.1447	1.8553
	Vasicek	0.0587	\	0.7105
Corn	Delta Kernel	0.1169	-0.9994	3.1994
	Vasicek	0.1169	\	4.1987
Oil	Delta Kernel	0.0791	0.1467	2.3533
	Vasicek	0.0791	\	2.2065

Table 7.4.1: Parameters obtained from the minimisation of the RMSE for the model with the Delta kernel.

since, for the Delta kernel, only the result of the difference of the parameters matters.

However, we know that the long-run level is different from the one in the Ornstein–Uhlenbeck model, so one may wonder if the value of β and a would have an influence on the quality of the estimation of the θ parameter? In fact, once again the minimisation of the RMSE will simply aim to have the value of the long-run level equal to a certain level, forcing the θ parameter to adapt, according to the values of β and a, in order to fill this equation. This confirms the intuition leading to the fact that the instantaneous memory term is, in the case of a model with a Delta kernel, redundant. Thus, it should be noted that the value of the parameter θ is more difficult to interpret as it is influenced by the value of β and a. It is better to interpret the long run level. The results for the study of the parameter θ are presented in table 7.4.2.

Commodity Market	Model	θ	Long-Run Level
Gold	Delta Kernel	0.1265	0.3151
	Vasicek	0.3152	0.3152
Copper	Delta Kernel	0.0905	0.2363
	Vasicek	0.2363	0.2363
Corn	Delta Kernel	-0.0493	-0.0376
	Vasicek	-0.0376	-0.0376
Oil	Delta Kernel	0.1644	0.1754
	Vasicek	0.1754	0.1754

Table 7.4.2: Parameters obtained from the minimisation of the RMSE for the model with the Delta kernel.

The higher long-term levels for the Gold and Copper markets can be easily interpreted by the fact that they grow throughout the period under consideration, unlike Oil and Corn which have more erratic paths. We can point out here what seems to be a weakness of the model here: the parameter θ assumes a mean reversion which is constant. However, the introduction of a memory parameter (in the case of long memory models, i.e. Exponential kernel) allows us to overcome this problem by the fact that the model will take into account the previous evolution in its future evolution: if the model grows during months, it will take it into account in its future evolution before stabilizing towards its reference level (return to the average). Even if it is difficult to estimate with precision the value of this return to the mean, the introduction of the term memory can solve partially this problem by introducing a change in the level of reversion by making it dynamic over time.

The characteristic function of Z_j can now be inverted and its distribution evaluated in order to find the optimal parameters. As mentioned before, we had to set the condition implying that the mean of the process Z_j is zero in order to facilitate the calculation of the autocorrelation in particular. However, this condition is not very restrictive as we will see later. However, for the Kou model, this constraint is not necessary since the calibration method does not require passing through the autocorrelation. This is therefore a major advantage of this model. We can observe the values of the optimal parameters in table 7.4.3. We will look at the empirical and theoretical densities later in this chapter.

The standard errors of the parameters estimated by the maximum log-likelihood are calculated as the square root of the inverse of the Fisher information matrix, i.e. the negative of the Hessian log-likelihood matrix :

$$\text{Std. Error}(\Theta) = \sqrt{-\Big(\frac{\partial^2 \log L(\Theta)}{\partial \Theta^2}\Big)^{-1}}. \tag{7.4.1}$$

Market	Model	Log-likelihood	λ	p	η_1	η_2	σ
Gold	Delta Kernel	4376	30.789	0.756	142.281	45.921	0.101
	Kou	**4398**	47.026 (0.040)	0.563 (0.055)	123.409 (0.014)	143.213 (0.194)	0.094 (0.003)
	Vasicek	4322	/	/	/	/	0.112
Copper	**Delta Kernel**	**3920**	49.907 (0.079)	0.473 (0.093)	117.233 (0.134)	130.356 (0.081)	0.149 (0.000)
	Kou	3918	53.292	0.591	128.798	121.089	0.148
	Vasicek	3908	/	/	/	/	0.166
Corn	Delta Kernel	3900	225.508	0.440	98.152	124.698	0.098
	Kou	**3907**	159.194 (0.122)	0.489 (0.027)	113.543 (0.069)	118.509 (0.089)	0.106 (0.004)
	Vasicek	3838	/	/	/	/	0.177
Oil	**Delta Kernel**	**3233**	36.642 (0.217)	0.227 (0.045)	15.203 (0.074)	51.687 (0.140)	0.224 (0.005)
	Kou	3199	72.772	0.580	60.836	44.112	0.184
	Vasicek	3128	/	/	/	/	0.299

Table 7.4.3: Parameters and their standard error obtained by maximisation of log-likelihood for the model with the Delta kernel, the model Kou and the model Vasicek.

Let us start by drawing attention to the fact that, as far as the Delta model is concerned, the calibration of the density of the log-returns is done only through the variance and the kurtosis. Indeed, we have previously assumed that the log-returns have a zero mean (after having centred them at 0 through the mean reversion term). This constraint also imposes that the skewness of the density is zero. For the Kou model, we no longer impose this constraint since we no longer calibrate the autocorrelation.

Various observations are interesting. First of all, let us notice that the Vasicek model is never retained, implying that the introduction of jumps in the model seems relevant for each of the markets despite the constraint that imposes that the jumps be of zero mean for the Delta kernel. We note also that two markets retain the Delta kernel model and two markets retain the Kou model. This can be explained by the fact that the latter two are characterised by a non-zero mean and skewness as we will see later. Indeed, the constraint on the mean also has an impact on the skewness and constrains it towards 0.

We note that all the parameters are significant and relatively stable with the exception of the p parameter for the Copper market which has a very wide confidence interval $[0.291; 0.655]_{95\%}$ due to its rather high standard error and therefore cannot conclude whether there are more positive or negative jumps on average. This problem could probably be partially handled by increasing the sample size of the data, by increasing the accuracy of the gradient descent calculation in the optimization algorithm or by introducing regularization. However, in order to obtain a really accurate output, the computation time becomes important, so it is necessary to put the balance between accuracy, stability and efficiency.

We notice that the different models tend to capture a lot of jumps for the Corn and Oil markets ($\lambda >> 0$) with up to 225 jumps per year for the Corn market for the Delta model, i.e. almost one jump per day (252 trading days per year). As mentioned earlier, these two markets seem to show more erratic patterns. However, as the η_1 and η_2 parameters are quite high (> 100), these jumps are often small. The Delta kernel model used for the Oil market has 3 times more negative jumps than positive ones ($p = 0.227$), but when a positive jump appears it tends to be quite large ($\eta_1 = 15.203$). We notice that the speed of reversion ($\beta = 2.3533$) as well as the volatility parameter ($\sigma = 0.224$) are both quite high which suggests that there are quite important movements in the process: when the noise term kicks in, the price tends to move quite far away from its previous value, then the reversion towards the long term level is quite fast. Then, we observe kind of swings in the process.

Each of these models satisfies the Markov property and has no memory or instantaneous memory. In the next section we will see whether a non-Markov model can improve on the results obtained so far.

7.5 Calibration of Exponential Kernel

The calibration of the Exponential kernel model was harder to perform, mainly due to its more exotic form which caused numerical instabilities as discussed below.

Concerning the study of the autocorrelation in the model, the introduction of an additional parameter and the more flexible form of the difference of two exponentials allow us to outperform the fit of the Delta kernel for each of the different markets. However, for some markets (i.e. Copper), the optimal parameters allowing to minimize the RMSE included a parameter β and/or b close to 0 ($\approx 10^{-7}$), leading, in the following steps, to errors in the derivation of the FFT and the inversion of the characteristic function. It was therefore necessary to bound the parameter in order to make it large enough for the numerical approximations to be carried out without concern : it was necessary to introduce the condition such that $\beta, b > 0.001$. We can however interpret the parameter b, being very small, as trying to bring the exponential closer to 1 ($= e^{0\times t}$), which suggests that the negative exponential (or the difference between two exponentials for the Exponential kernel) is not the best function to fit the past influence of the process. One can interpret the fact that the model sets β to 0 such that the mean reversion term does not help to explain the dynamics of the underlying. However, by bounding the parameters further, we will explore this model to see if it can improve on the results obtained previously. Figure 7.5.1 shows the autocorrelation of the different markets obtained according to the Exponential or Delta kernel.

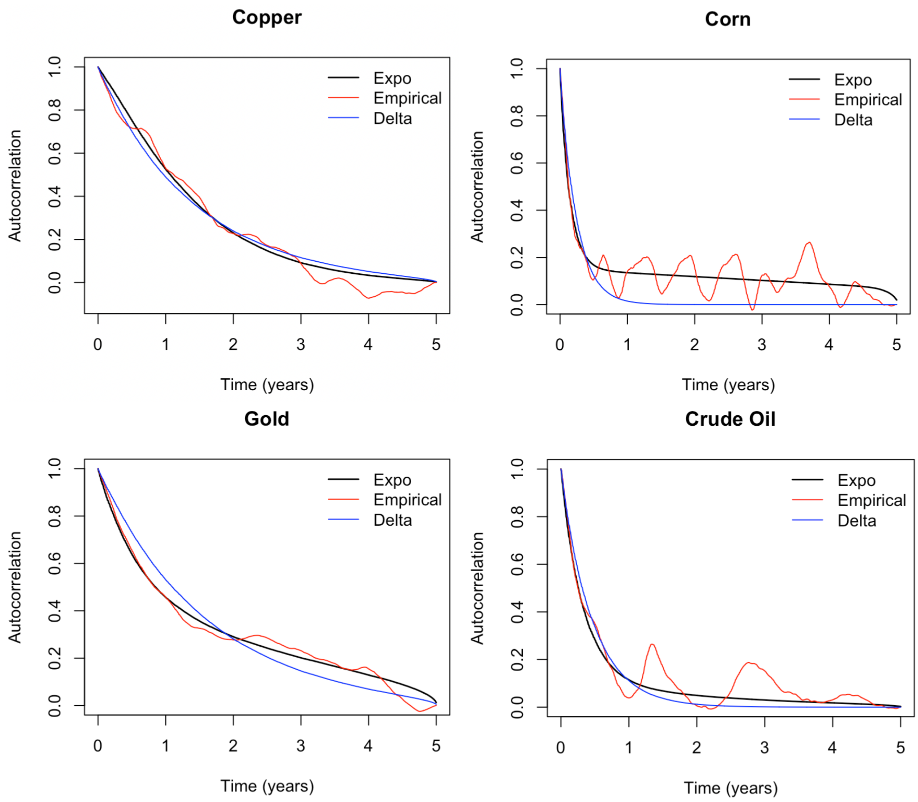

Figure 7.5.1: Autocorrelation of the commodities markets according to the Exponential kernel model, Delta.

Commodity Market	RMSE	β	b	a
Gold	0.0324	1.3343	0.6168	0.4245
Copper	0.0477	0.0010	2.3445	-1.3730
Corn	0.0615	6.8127	0.2466	1.7931
Oil	0.0570	2.9602	0.2801	0.4090

Table 7.5.1: Parameters obtained from the minimisation of the RMSE for the model with the Exponential kernel.

First of all, we observe that the introduction of the memory term in the Exponential kernel model makes it possible to improve the fit of the autocorrelation since the RMSE are smaller than previously for each of the markets studied. We observe that the RMSE obtained is quite far from the one obtained for the Corn market and this is easily visible on the graph since the Delta kernel, by its constrained form, assumed that autocorrelation is close to zero after 6 months whereas the Exponential kernel manages to capture autocorrelation better over the years. However, there is a sort of recurrent wave for the Corn and Oil market. Therefore, it could be interesting to study the kernel Power here since it was able to approximate the periodic function as we have seen before. In practice, this is a difficult task since it is necessary to calculate an RMSE (or MSE, etc.) by means of infinite sum integrals. The Gold and Copper markets appear to have more persistent autocorrelation over time.

Furthermore, it is more difficult to interpret directly the value of the parameter b and thus the persistence or not of the memory in the price of the underlying since the parameter β comes into play. It is easier to interpret directly the graphs presented in figure 7.5.1. However, we can still conclude that the Corn market seems to have a very high speed of reversion to the mean, unlike the Copper market, which seems to have no reversion to the mean in the price of the commodity.

Finally, it is noteworthy that the continuation of the calibration approach used until now is no longer efficient since the parameterisation obtained thanks to the autocorrelation does not make it possible to invert the characteristic function to obtain the density of the process. Indeed, due to its exotic form, the Exponential kernel requires a particular attention on these three parameters: it was necessary to impose the condition such as $\beta < b$. However, this condition seems too restrictive and does not allow us to obtain good results. Thus, it seems necessary to change the calibration approach. To do so, we introduce a simplified approach such that we start from the scaled log-prices and introduce the estimation of all the parameters in the likelihood maximisation on the density of log-returns. Indeed, the difference with the previous approach lies in the fact that we no longer take into account autocorrelation and that all the parameters will be estimated in a single step: the maximisation of the log-likelihood of the density of log-returns.

This approach, which seems less restrictive and more generic, suffers from two weaknesses: on the one hand, the parameterisation obtained does not in any way allow for any form of memory in the process, as was the case with autocorrelation, and on the other hand, it attempts to estimate 9 parameters at a time, which raises concerns about the precision of the calibration. Indeed, the gradient descent techniques that rely on these parameter estimates are becoming more and more sensitive to the starting values and more and more heterogeneous sets of parameters are proposed as solutions. However, on the other hand, this new approach has two considerable advantages: firstly, it is no longer necessary to set the constraint such that $\mathbb{E}[\mathrm{d}L_t] = 0$ which, in addition to introducing a non-zero mean, also allows the introduction of skewness which can be very interesting in certain markets as we will see. Secondly, this approach will make it possible to leave the parameters a, b, β more free and even to drop them to 0 which will be the case for certain markets.

Table 7.5.2 shows the parameters obtained by maximisation of log-likelihood according to the new method of calibration. It is no longer really possible to interpret the autocorrelation obtained with the new parametrisation since it is no longer a calibration criteria. Indeed, the assumption of stationarity of

Commodity Market	Gold	Oil	Corn	Copper
Log-likelihood	**4403**	3200	3855	3914
a	0.3748 (0.005)	-0.744	-2.747	0.377
b	2.335 (0.020)	1.825	3.595	7.792
θ	0.2472 (0.017)	0	0	0
β	1.0507 (0.011)	0	0	0
λ	29.681 (0.007)	110.364	135.482	29.658
p	0.7331 (0.008)	0.2364	0.442	0.4977
η_1	128.294 (0.021)	40.101	115.302	102.227
η_2	148.617 (0.190)	53.019	119.141	110.223
σ	0.1212 (0.001)	0.1421	0.1012	0.1622

Table 7.5.2: Parameters and their standard error obtained by maximisation of log-likelihood for the model with the Exponential kernel.

the log-returns is even more present under this calibration technique: the parameters a, b, β obtained make it possible to obtain the optimal density of the daily log-returns. We observe that only the Gold market has retained the mean-reverting term and it is also the only market that allows to improve its log-likelihood thanks to the kernel Exponential model. The parameterizations of the other models fail to explain the dynamics of the log-returns. The interpretations of the value of the different parameters are relatively close to those made in the previous section: the Corn market has many jumps and in general more negative ones, as well as the Oil market which has three times more negative jumps than positive ones, the opposite of the Gold market (three times more positive jumps).

We also notice that the parameter a of the Corn and Oil markets is negative which implies that the memory term has a negative effect on the future of the process: when the price of the underlying increases in t, then the memory term will pull the price down in $t+1$. In other words, the past prices of the commodity have a negative influence on the price of the commodity today. As the parameter b is very high for the Copper market, this can be interpreted as the fact that the memory of the process fades very quickly: only the prices of the last few days have an influence on today's price and on those that will come in the following days. In contrast, the Oil market seems to retain its price memory for a long time. In this context, since the β parameter is zero, it is easier to interpret the value of the b parameter directly. However, as explained above, we must interpret these parameters with caution since autocorrelation is not used in this new calibration.

Commodity Market	Empirical				Model			
	$\mathbb{E}[\Delta X_t]$	$\sqrt{\mathbb{V}[\Delta X_t]}$	$\mathbb{S}[\Delta X_t]$	$\mathbb{K}[\Delta X_t]$	$\mathbb{E}[\Delta X_t]$	$\sqrt{\mathbb{V}[\Delta X_t]}$	$\mathbb{S}[\Delta X_t]$	$\mathbb{K}[\Delta X_t]$
Gold (EXPO)	0.0003	0.0075	0.1809	7.778	0.0003	0.0075	0.2571	6.973
Copper (DELTA)	0.0001	0.0107	-0.0017	4.911	0	0.0105	0	4.707
Corn (KOU)	0	0.0112	-0.0476	5.363	0	0.0110	-0.0537	5.215
Oil (DELTA)	0.0002	0.0052	-0.0838	3.964	0	0.0048	0	3.871

Table 7.5.3: Moments of the daily log-returns of the markets and the fitted optimal models.

Table 7.5.3 presents the detailed study of the different empirical moments with those obtained for each of the optimal models (Delta kernel, Exponential kernel, Kou). First of all, let us recall that the calibration methods being different according to the models, some constraints are taken into account or not, i.e. the diffusion term has a zero mean for the Delta kernel model. Moreover, note that the mean of the log-returns of each model are close to 0 while they all have a mean reversion term since we assume that we are at the reversion level at inception, in $t = 0$, the mean reversion term no longer affecting the mean. Except for the Gold market, all markets have negative skewness which is often observed in practice. We notice that the kurtosis is highest for the Gold market, as can be seen in the figure 7.5.2: the density of Gold rises very high and shows heavy-tailed distribution.

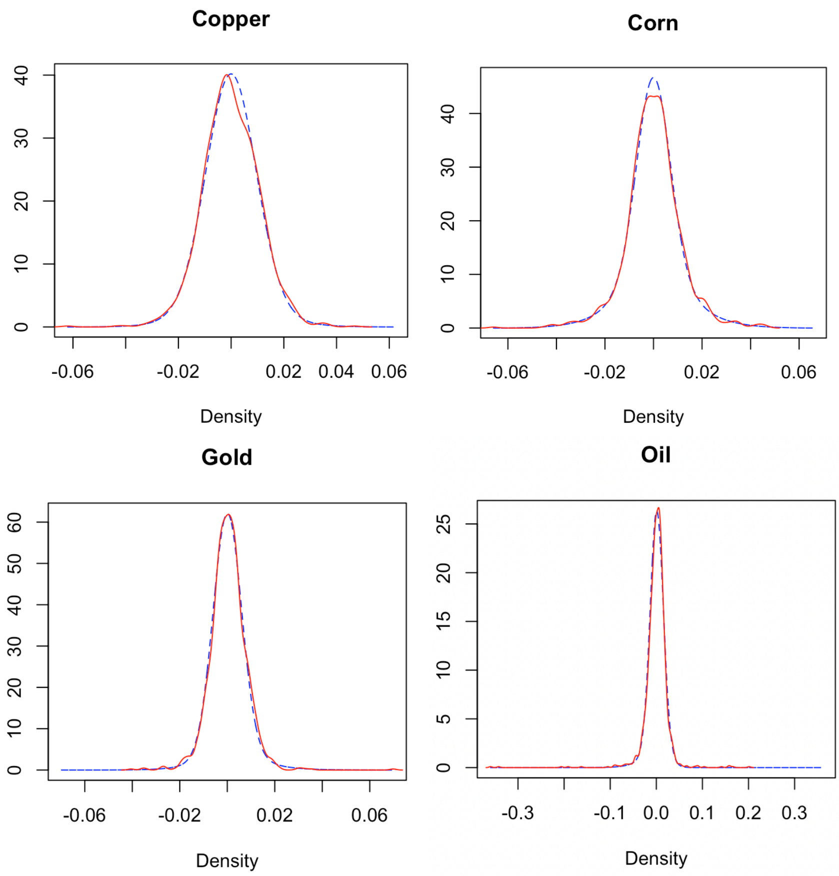

Figure 7.5.2: The empirical densities and those approximated by the optimal model for each market.

Figure 7.5.3 shows price simulations according to the models and parameters chosen for each of the commodities. It can be seen that the simulations follow the observed prices fairly closely. These simulations make it possible to become aware of the different characteristics of each commodity: greater variance of Corn and Copper, stationarity of Oil over time, constant evolution of Gold, etc. It is clear that the fall in the price of Oil in April 2020 is an isolated event and that it is practically impossible to predict this type of situation.

Let us emphasise that these simulations are carried out according to the real probability measure $\mathbb{P}$ in order to have a legitimate comparison with the observed prices. However, as explained above, these results should be considered with great caution due to the uncertainty that resides in the projections of the real world.

Finally, we also observe at figure 7.5.4 the projections of prices according to the last available price on 1 November 2020 over a one-year horizon as well as the 0.5% quantile of the lowest prices in the distribution. We notice that for the Copper and Gold markets, this « boundary » price obtained in 1 year with a probability of half a percent is none other than the price observed today leading to the observation that in 0.5% of cases, the price observed in 1 year will be smaller than or equal to the price observed today. This is consistent with the fact that both markets seem to have grown throughout the last 5 years.

7.6 Yield Curve

When we are looking at the pricing of commodity derivatives, the main risk is on the underlying, which will be the most sensitive and will have the most impact on the risk taken by the financial players.

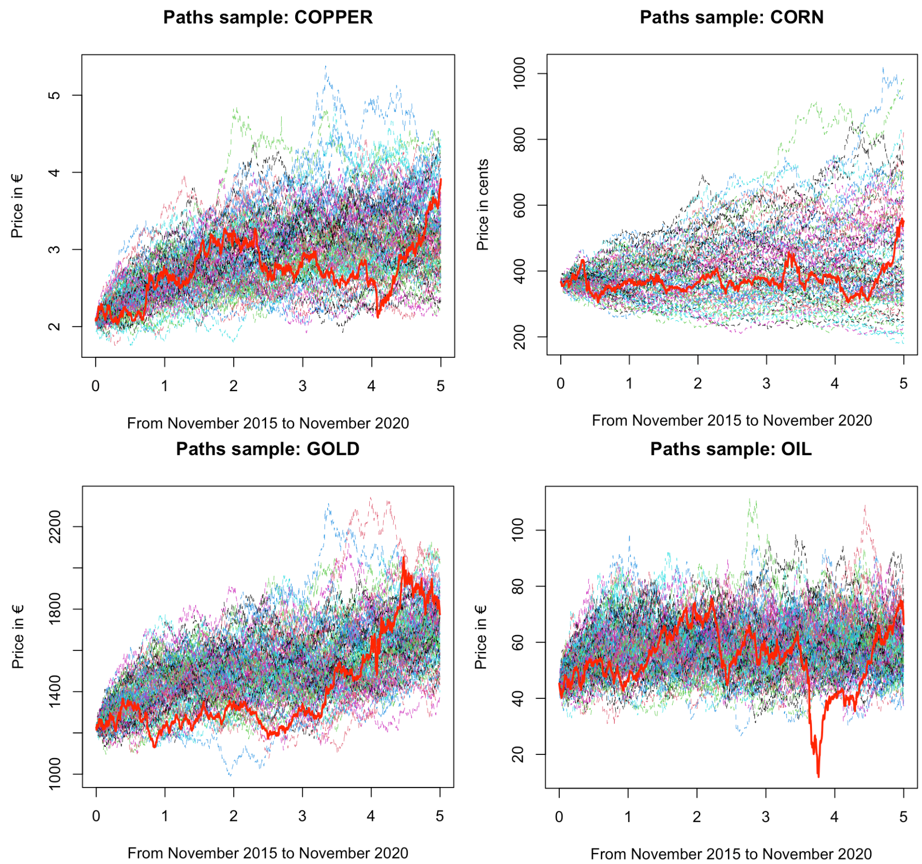

Figure 7.5.3: Simulation of 100 paths sample of the price of each commodity over the 5 years of observation and comparison with the price actually observed on the market.

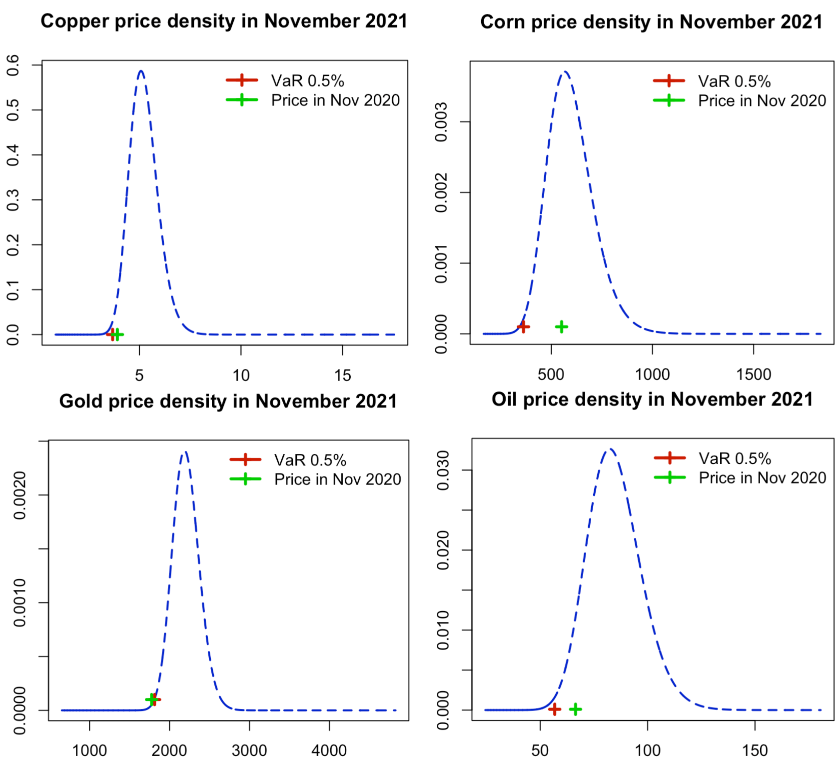

Figure 7.5.4: Price density in November 2021 (in blue) according to the models and parameters chosen previously, the actual price (in green) and the 0.5% quantile (in red).

Nowadays, the interest rate risk is quite low, so we will not consider a stochastic rate model.

In order to obtain a risk-free interest rate, we will use the LIBOR as a proxy, and more precisely LIBOR USD, it is the most common choice in the derivative industry. The LIBOR is the short-term borrowing rate of AA-rated institutions on the international interbank market for short-term loans. Although the 2009 crisis has shown that this rate is not truly risk-free, we will favor its use since it is not easy (or free) to access other rates used on the market (such as OIS).

Note that we cannot use this curve as it stands since it is expressed in simple interest rates as the maturity is less than 1 year. However, since we price our derivatives using compound interest rates, it is therefore necessary to make the following modification in order to obtain the compound interest rate curve:

$$r^c(T) = \frac{1}{T}\ln(1 + r^s T), \tag{7.6.1}$$

where r^c is the compound interest rate, r^s the simple interest rate and T the maturity (in years). We obtain the curve shown in Figure 7.6.1. It can be seen that the rates are very low. This can be explained by the very low rates that are currently on the market. We can therefore already see that the impact of the risk-free rate on the pricing of derivatives will be very small. Therefore, some investors may want to use investments with a maturity of more than one year to achieve higher returns. In order to do this, we will use the Nelson-Siegel model, calibrated in $t = 0$ on the LIBOR USD rates for larger maturities. Although this method lacks precision and does not correspond to market reality[6], it makes sense in a context where, once again, the data are hard to access. The model takes the following form:

$$r^{NS}(T) = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{pmatrix}^{\top} \begin{pmatrix} 1 \\ \frac{\tau(1-e^{\frac{-(T-t)}{\tau}})}{(T-t)} \\ \frac{\tau(1-e^{\frac{-(T-t)}{\tau}})}{(T-t)} - e^{\frac{-(T-t)}{\tau}} \end{pmatrix}, \tag{7.6.2}$$

The 3 parameters of interest represent the level, the slope and the curvature of the yield curve, respectively. They are obtained by least square minimization errors and in our case : $\beta_1 = 0.00396, \beta_2 = -0.00308, \beta_3 = -3e-05$ and $\tau = 0.35529$. We can see that when the maturity tends to infinity, the long run level[7] is 0.00352 ($= \beta_1$). This gives the curve shown in Figure 7.6.1.

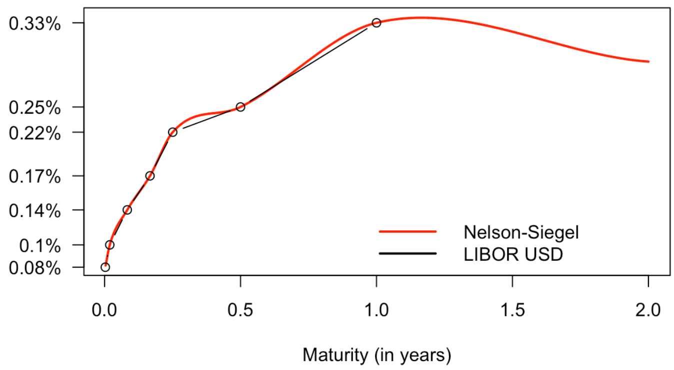

Figure 7.6.1: LIBOR USD term structure on the November 1, 2020 (black) and Nelson-Siegel LIBOR USD term structure at inception, $t = 0$ (red).

7.7 Implied Volatility

As mentioned before, the implied volatility surface is the representation adopted by the market for prices of call (and put) options. In order to analyse the impact of memory parameters and mean reversion

[6]It is not consistent that the rate at 2 years is lower than at 1 year.
[7]Note that this level is reached when T > 10 000 years.

on implied volatility, the prices of European options call are fitted thanks to the parameters obtained previously. We will make sure that the models retained are able to reproduce the features present in the financial market. In order to further understand the impact of the characteristics of each model and its specificities, a sensitive analysis on the parameters is performed by varying one parameter at a time for a fixed maturity. Let us first observe the different global surface for the optimal models defined above.

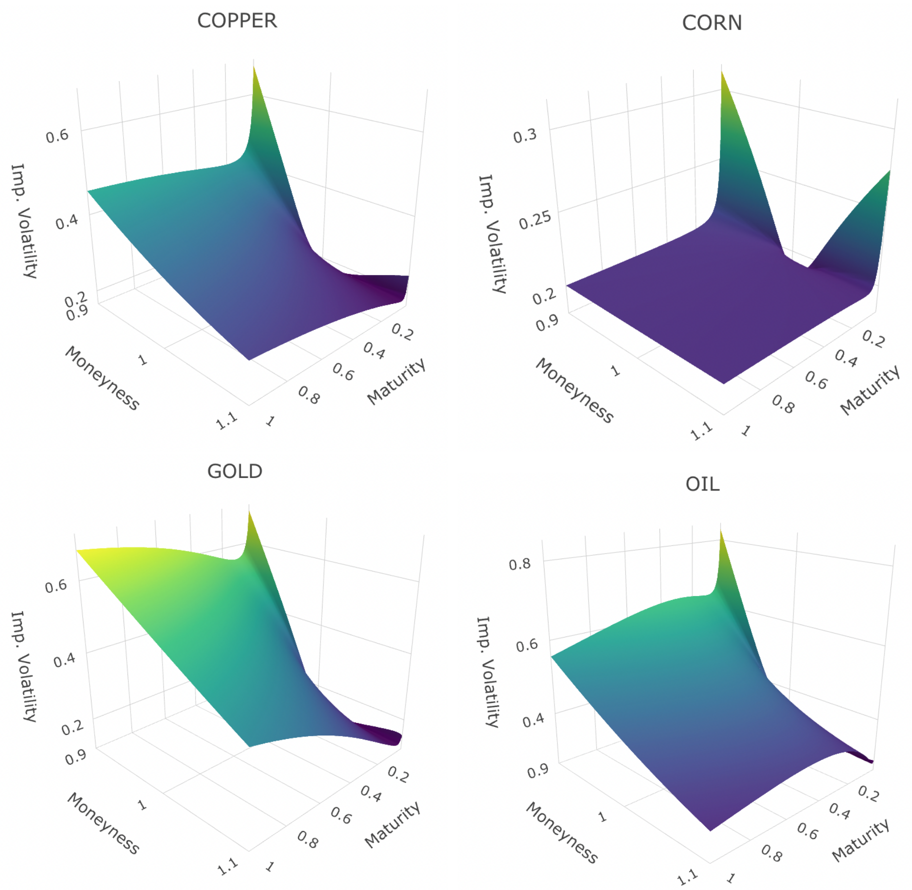

Figure 7.7.1: Implied volatility surface of Call Options with the retained models.

One can notice that the implied volatility is strongly sensitive to the evolution of the moneyness, mainly for short maturities, but it remains so even when the maturity increases. As announced in the section corresponding to implied volatility, we notice that when we are out of the money, volatility draws a skew. About the Copper surface, contrary to what is often observed for mean-reverting models, we observe a smile and not a skew with a local minimum ATM, a feature that is usually more attributed to the equity market. However, it should be noted that this smile tends to disappear very quickly as the maturity increases.

We note that the Corn market, which does not contain a mean reversion term, shows a smile in its surface for low maturities. The curvature of this smile and the one of the Copper surface are inversely proportional to time to maturity. Contrary to equities, note also that the Gold and Oil markets show a skew and not a smile without a local minimum for ATM options. This is justified by the elements presented in the section on implied volatility: processes with skewed jump distributions present a skew in their implied volatility surface. But also that unlike diffusion models, which propose a rather weak skew for low maturities, models with jumps in price dynamics (and exponential-Lévy models) introduce a very pronounced skew for low maturities. Instead, volatilities are convex and inversely proportional to the moneyness. Given that a similar trend is not observed for the equity volatility surface, this behavior may be associated to the mean reversion of returns.

Regarding the sensitive analysis on the influence of the parameters on the implied volatility, we observe in figure 7.7.2 the effect of the variation of $\theta, \eta_1, \eta_2, p, b, \lambda, \sigma, \beta$ on the implied volatility for the Exponential kernel model with a fixed maturity of 1 month. Note that the findings for the Delta model are substantially the same. A first observation may be that at this maturity, the volatility skew has decreased a little, there is a less steep slope. Then, it is obvious that the mean reversion parameters, i.e. θ and β respectively, influence implied volatility almost exclusively when we are with moneyness lower than 1 (OTM puts, ITM calls[8]). Their effect is similar as their influence on the process is quite similar. However, we notice that the effect of the speed of reversion (β) is null far from ITM calls while the level of reversion (θ) is significant regardless of the price of the underlying asset.

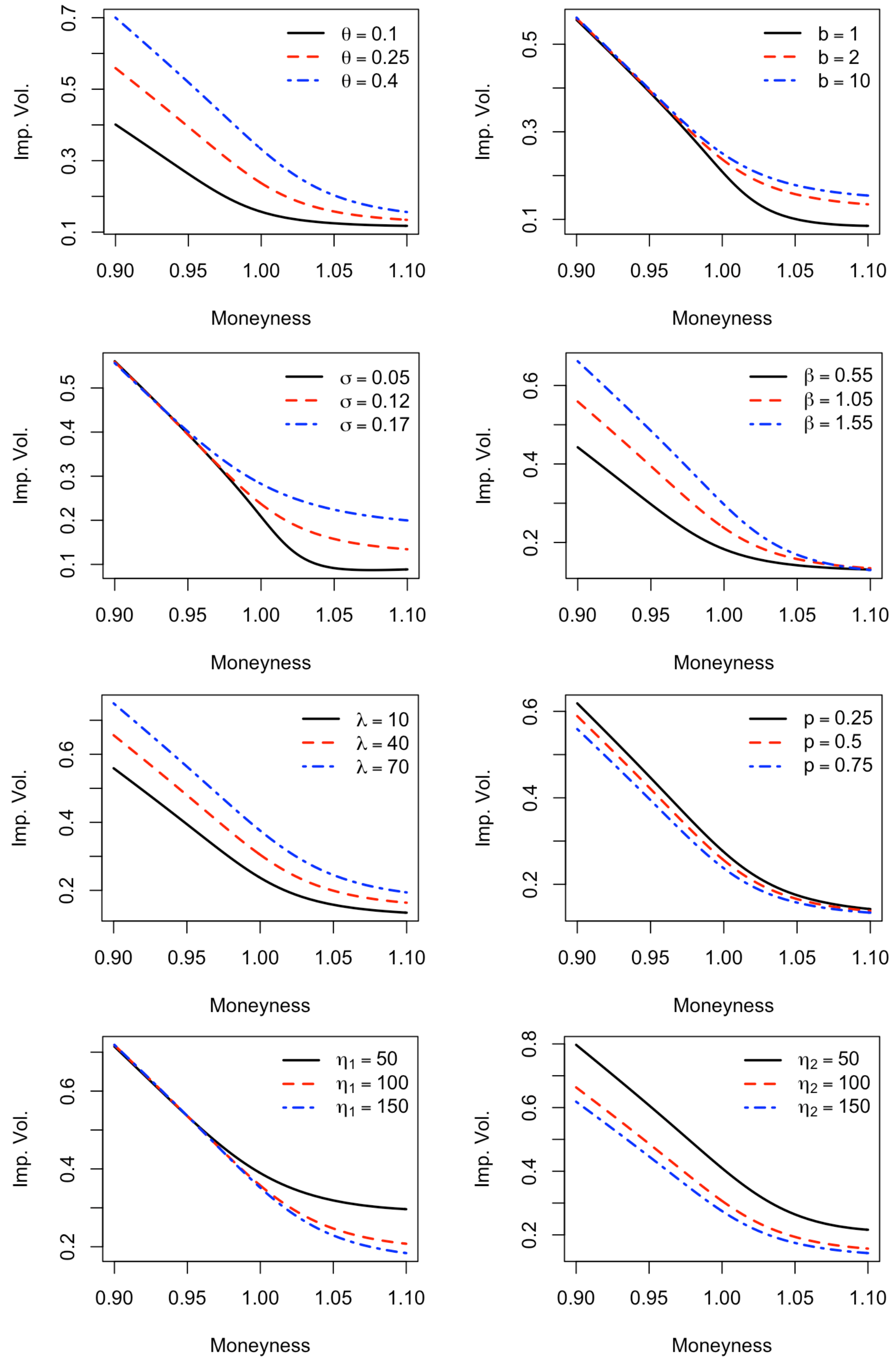

Figure 7.7.2: Implied volatility sensitivity of Call Options on Exponential kernel Gold with maturities of 1 month.

[8] Recall that we have defined moneyness as K/S.

Next, we notice that the effects of λ, η_2 and p all tend to influence the implied volatility by a translation and not by a real change in curvature. On the contrary, σ, b, and η_1 have an influence when one is close to the money, especially for σ, and in the money, especially for b and η_1. We see a modification of the curvature for these three parameters. Note also that when σ decreases and tends to 0, we tend to find a slight implied volatility smile while the curve becomes almost flat when σ becomes large. When the volatility tends to 0, the variance also tends to 0 (there will be some variance left because of the jumps) and we can see that we are approaching a straight line which takes the well-known form of a hockey stick. The effect of an increase in b is less and less important since b is involved through a negative exponential. It can therefore be seen that the effect of introducing memory into the price of the underlying tends to affect the skew of implied volatility and to increase the slope when the memory is longer term.

7.8 Pricing of Derivatives

We will start by calculating the price of derivatives (Option Call) according to the different pricing methods presented in section 4 in order to compare the difference in the price, and then compare them and observe the influence of different market characteristics on the final prices. To do this, we calculate the price of an option according to the 3 methods and calculate the absolute value of the difference between the prices. The results are shown in Figure 7.8.1 (right).

A first observation is that each of the three methods gives fairly similar results. However, we note that the Monte Carlo simulation method is still quite volatile due to its construction and that it would be necessary to use even more simulations in order to converge even more towards the values obtained by the Carr-Madan method and by the density reconstruction. However, the calculation time is already very high compared to the other two methods. Thus, for reasons of accuracy and computation time, we will prefer the methods using the FFT algorithm. In order to judge precisely the difference between these two methods, let us study figure 7.8.1 (left).

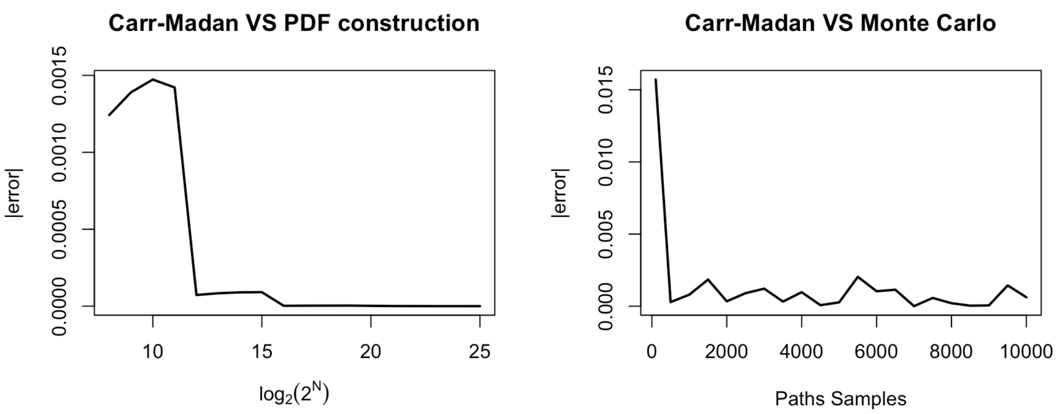

Figure 7.8.1: Spread between the price of a call option according to the Carr and Madan formula and according to the algorithm for the construction of the neutral density.

We notice that quite intuitively when N (which corresponds to the number of points in its density) grows the spread decreases until it seems to stabilise. In reality, it keeps decreasing by a factor of 10 every time the step power increases by one. The method of numerically integrating the reconstructed density seems more volatile and more sensitive to the number of steps chosen. Moreover, it requires a much longer calculation time than the Carr-Madan formula. Thus, we will prefer the Carr-Madan formula for the remainder of this section and in particular for the following sensitivity analysis.

Concerning the analysis of the price obtained for the different markets and models, it is essential to point out here that the parameters were calibrated on the log-returns of the prices, therefore some of them (i.e. θ) are dependent on the price used in $t = 0$, i.e. $X_0 = 0$. Thus, it is required to price the option with $S_0 = 1$ and to choose the strike such that it will allow to obtain the desired moneyness then multiply the price of the option obtained by the price actually observed at the inception. Let us take an example to illustrate this: if the price in $S_0 = 100$ and we want to set the strike at 90, then we must price with $S_0 = 1$, $K = 0.9$ and then multiply the price of the option obtained by 100. In this case, in order to facilitate the comparison between the option prices obtained between the markets, we will consider $S_0 = 100$ for each of the commodities. The call prices for each of the markets are shown in table 7.8.1.

Maturity	Gold (Expo)	Corn (Kou)	Oil (Delta)	Copper (Delta)
$T = 1$ month	11.461	9.185	12.223	10.483
$T = 6$ months	21.337	10.901	21.410	16.617

Table 7.8.1: Call option prices computed with the optimal parameters derived earlier for each of the markets and the optimal model for a maturity of one and six months, respectively.

Several observations emerge: first, we notice that models including memory terms tend to increase the option price, whether this memory is instantaneous or long. This is mainly due to the inclusion of a mean-reverting term which greatly influences the option price as we will see in the next section. Note however that the price of Gold and Oil options vary greatly over time, which can be explained by the fact that their market price has been increasing steadily over the last 5 years in the case of Gold and over the last 6 months in the case of Oil, with each model taking into account memory terms (long for Gold and short for Oil). We note that the price of Corn has been relatively stable and stationary for years and this is reflected in the price of derivatives. However, it remains a difficult task to interpret the price of options that way, so in order to better understand the impact of each parameter on the prices we observe, we will perform a sensitivity analysis.

Finally, consider the comparison between the different models for one and the same market in Table 7.8.2. We know that vanilla options are only focused on marginal distributions at maturity, which has the consequence of not taking into account the « path-dependent » effects that characterise memory processes. Thus, quite intuitively, the different prices obtained should be similar regardless of the model considered. However, since the models have not been calibrated according to the same assumptions (nullity of the mean of the noise term, etc.), some differences may exist. It should be noted that price differences can still be significant as maturity increases: the price according to Kou's model may diverge greatly from other prices since Kou's model does not offer a return to the mean and a large trend in price. This effect has a very large influence on the price as we will see in the next section. In addition, the various models were calibrated using observed prices and not options prices, thus not being able to account for pricing under the risk neutral probability measure (or pricing measure). Indeed, we notice that some prices are quite different, especially for the Exponential kernel models, since the latter has been « well » calibrated only for the Gold market, other markets having failed to obtain accurate and robust parameters. For the Gold market, all prices are relatively similar. Later in this chapter, we will look at the pricing of exotic derivatives in this market in order to highlight the path-dependent effect.

Model	Gold	Corn	Oil	Copper
Kou	10.945	9.185	12.223	10.284
Delta Kernel	11.178	8.983	12.077	10.483
Expo Kernel	11.461	9.918	11.310	9.401

Table 7.8.2: Call option prices computed with the optimal parameters derived earlier for each of the markets and each of the models for a maturity of one month.

7.8.1 Sensitive Analysis

In order to learn more about the influence that the different parameters can have on the option price, we will vary, for a particular market, each of the parameters one by one and fix the others at the optimal values defined earlier. We will carry out this study on the Copper market for the Delta kernel model and on the Gold market for the Exponential kernel model, i.e. for the parameter b. Figure 7.8.2 shows the sensitive analysis of the Delta kernel model for the parameters: $\beta, \theta, \eta_1, \eta_2, p, \lambda$ and σ and of the Exponential kernel model for the parameter b.

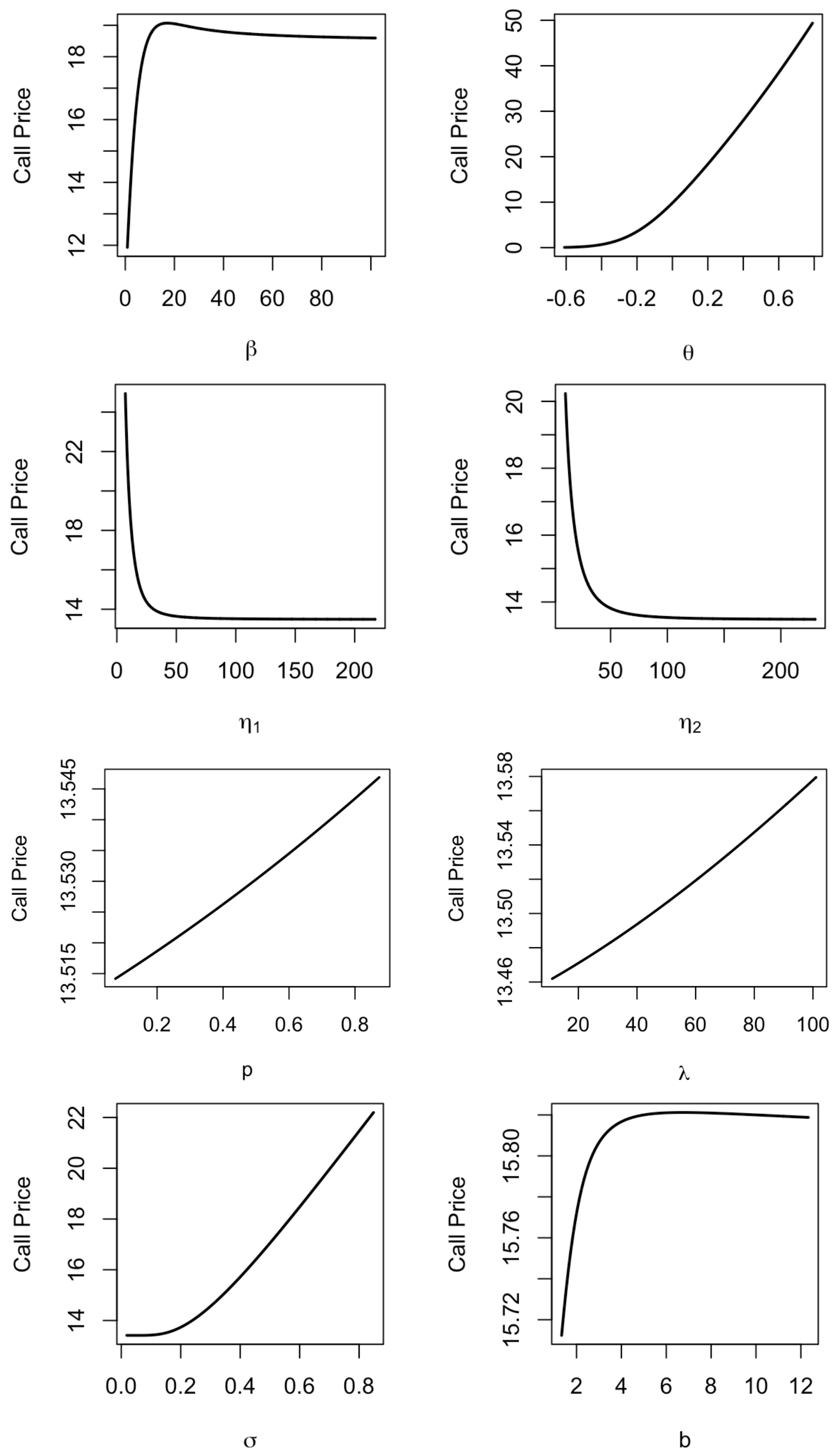

Figure 7.8.2: Sensitivity analysis of Call Options on Copper (and Gold for b) with maturities of 3 months.

First of all, the trajectory that the call price takes when β increases is totally in line with the intuition that one can have since this curve corresponds to that of a function of the type $1 - e^{-x}$. Indeed, when β tends to infinity, on the one hand, the Lévy term tends to 0, and on the other hand, the mean reversion term will tend towards its reversion level which explains why the option price is flat. Obviously, when β becomes zero, we find the Kou model.

On the contrary, concerning the θ parameter, we notice that its effect seems linear or even slightly exponential on the option price, which is coherent since its effect is linear on the value of the log-returns and thus becomes exponential on the value of the underlying price. Therefore, for longer maturities and larger θ values, this effect becomes more pronounced. We notice that the option price tends towards 0 when the mean reversion level becomes too low, which is consistent.

Concerning the different jump parameters, we notice that η_1 and η_2 seem to follow a negative exponential and the option price becomes flat when the jumps become small in amplitude. As regards the parameters p and λ, we note that their influence remains rather weak on the value of the option, contrary to the other parameters. When the probability of recording a positive jump (p) increases, the option price increases, which seems consistent: the more likely it is to have positive jumps in the price of the underlying asset, the more the price of a call increases. This intuition is also relevant for the λ parameter. However, it is notable that this effect can vary depending on the value of the other jump parameters: if the average jump were to become negative, increasing the jump frequency would decrease the value of the option and vice versa.

Regarding volatility, its effect is initially quite flat and then becomes convex. Its effect on the call price is similar to what can be observed for a classic Black & Scholes option: when volatility tends towards 0, the option price is less and less uncertain and it tends towards a simple discounting of the price between maturity and the current date. This observation is similar in our case except that the variance is not zero when σ is zero since there is still uncertainty from the jump term.

Finally, with regard to the memory parameter b, the sensitive analysis was obtained on the Gold market. It can be seen that the influence seems similar to that of the β parameter and that the option price varies little because of b. Since when the parameter b decreases, the memory is more persistent over time, it can be seen that a « long » memory tends to decrease the option price. However, this particular univariate analysis is dangerous: let us start again from the memory process introduced in equation (3.2.1). We observe that when b decreases, the total value of the process tends to increase since the integral of the memory term tends to increase. Therefore, if the parameter b increases, the value of the other parameters should adapt in order to find an equilibrium. It can therefore be seen that in the case where a becomes negative, the opposite observation holds.

7.9 Exotic Pricing

Although all models have been calibrated and give fairly close vanilla option prices (for a given market), exotic options can lead to drastically different prices. Indeed, contrary to vanilla options which are only interested in the marginal distribution at maturity, exotic options consider the whole process and can be path-dependent. Therefore, we will consider the pricing of a derivative product that allows to take into account the evolution of the process (in this case the price of the commodity) throughout the contract. Indeed, as the memory processes retained so far have trajectories influenced by the past evolution of the process, it may be interesting to observe whether these have an influence on the price.

One feature that could be interesting in the case of commodities is the introduction of coupons during the life of the contract since the possession of the latter suggests, on the one hand, possible costs of storage, merchandise, etc., and, on the other hand, does not allow for dividends, as is the case for financial shares. Thus, offering a product that allows for regular remuneration (notes) to cover these costs seems to be an important benefit. However, the price of this kind of product could quickly become very high and it may be interesting to introduce a condition allowing to decrease the price of the product, for example a barrier. Let us introduce a product that would satisfy these two features: the Barrier Reverse Convertible (BRC).

First, we will define the product in details in order to understand its weaknesses, its potential and its design. Note that this is an European contract and that there is therefore no surrender during the life of the contract. It consists of 2 components, on the one hand, a debt instrument (i.e. notes) generally higher than those offered on the market and on a monthly or quarterly basis. On the other hand, a derivative in the form of a put option which, at maturity (in $t = T$), gives the issuer the right to repay the principal or part of the principal. The different scenarios can be described as follows depending on the value of the stock S_T:

- In the case where the commodity price never reached or exceeded a lower barrier H but ended up above of its initial value S_0, then the investor gets back his initial investment in addition to the coupons paid during the contract and the amount is paid out in cash.
- However, if the price of the commodity crosses the lower barrier H at least once and still ends up below the starting price S_0 at maturity, then the investor only receives part of his investment. The amount he receives is in this case delivered in shares.

The payoff at maturity is therefore determined as follows:

$$\begin{aligned}\text{Payoff}^{BRC} &= N \times \mathbb{1}_{\{L>H\}} + N \times \mathbb{1}_{\{L\leq H\}} \times \mathbb{1}_{\{S_T>S_0\}} + \frac{N \times S_T}{S_0} \times \mathbb{1}_{\{L\leq H\}} \times \mathbb{1}_{\{S_0\geq S_T\}} \\ &= \frac{N}{S_0} \times \left[S_0 - \mathbb{1}_{\{L\leq H\}} \times \Big(S_0 - S_T \Big)_+ \right]\end{aligned} \tag{7.9.1}$$

where L is for the lowest value of the underlying during the life of the contract. The second term in the equation corresponds to a well-known product: Down-in-Barrier Put (DIBP) where the strike is equal to the initial price S_0.

Several points are important to mention regarding the advantages and disadvantages of the product. First of all, it is important to realise that it is a complex product which, in addition to the risks of « traditional » bonds (i.e. default risk, inflation risk, ...), also has risks related to the underlying. Indeed, it is possible to lose a substantial part of the principal (or even all of it in the worst case), as explained above. In addition, the product does not participate in any positive financial return on the price of the underlying in question. Then, the total return is limited to the stated coupon interest rate. Thus, the issuer of this product hopes that the value of the share will depreciate and the buyer hopes that it will remain stable or even increase in value.

However, intuitively, when one is willing to concede risk, one can get something back. The BRC product, and Reverse Convertible Notes in general, offer the opportunity to earn higher financial returns than most government bonds and other conventional products through higher coupons. This product is therefore more suitable for investors who are looking to invest part of their capital in riskier products with higher returns. It is also suitable for investors who bet that the price of the underlying will be relatively stable in the near future. However, it should be noted that in the case of a low-volatility share, the value of the coupons may be lower and the main attraction of the product may be lost.

Finally, it is important to choose the maturities proposed as well as the value and frequency of the coupons and the value of the barrier. Regarding the maturity, these products are generally of short term maturity (between 3 months and 1 year) with coupons received monthly or quarterly. We will look at these different possibilities, as well as the value of the coupons and the barrier to be chosen in order to obtain a contract that is both attractive for the issuer and for the investors.

We can write the Fair Price, thanks to (7.9.1) and by defining t_i as the times of payment of the notes with $t_n = T$, as follows :

$$\begin{aligned}FV_{t=0} &= C \sum_{i=1}^{n} \Big(e^{-r(t_i)\ t_i} \Big) + \frac{N}{S_0} e^{-r(T)T} \mathbb{E}^{\mathbb{Q}} \left[S_0 - \mathbb{1}_{\{L\leq H\}} \times \Big(S_0 - S_T \Big)_+ \right] \\ &= C \sum_{i=1}^{n} \Big(e^{-r(t_i)\ t_i} \Big) + N e^{-r(T)\ T} - \frac{N}{S_0} \times \text{DIBP}_{t=0}\end{aligned} \tag{7.9.2}$$

Several degrees of freedom are available and left to the issuer: the maturity T, the number n of coupons, the value C of the coupon and finally the value of the barrier H. We can already note that a variation of the barrier will affect the price of the DIBP option which will directly influence the value or the frequency of the coupons and inversely. We can also note that in order to add security, the issuer may also charge additional fees which should, however, remain reasonable in order to remain competitive. If we consider a fee of $x\%$, the product is sell for N and its fair value becomes $(1-x\%)N$. In this type of short-term product, the additional costs must be large enough to ensure a profitable product but not too large to remain competitive in the market and attractive to the customer. That is why we will consider fee of 1%.

In order to estimate the different features of the contract, let us start by defining some rules that will allow us to simplify our selection. To facilitate the analysis, let us start by assuming that $N = 1$,

- The present value must necessarily be smaller than or equal to $(1-x)N$ in order to receive a cash inflow of $x\%$ in fees for the issuer. Thus, with a fee of 1%, the maximum possible price is 0.99$ respectively.
- The barrier value and maturity should be set in such a way that they minimise the exercise of the DIBP option so that the product remains attractive.

Let us define ζ as the proportion of cases where the client will receive the full nominal amount at maturity as :

$$\zeta = 1 - \frac{\sum_{i=1}^{m} \mathbb{1}_{\{L \leq H\}} \times \mathbb{1}_{\{S_T \leq S_0\}}}{m} \tag{7.9.3}$$

Figure 7.9.1 shows the evolution of this proportion according to the chosen maturity and the barrier level. We observe that the increase of the maturity and the increase of the barrier tend to decrease the proportion ζ which is coherent since the price has more chance to cross the barrier when the maturity and the barrier value are higher.

The value of the coupons is therefore deduced from the equilibrium that exists in the relationship (7.9.2). Thus, it can be seen that, depending on his risk aversion, the customer is able to choose a higher or lower barrier and thus higher or lower coupons as well. Note that we will consider maturities of 6 months, 1 year, 1.5 years and 2 years to reflect the « short-term » nature of this type of product but also to see how the product reacts for maturities greater than one year. The coupon frequency will be every 3 months[9] and the barrier will range from 1450$ to 1700$, since $S_0 = 1775.3\$$. The value of the coupons will therefore be determined as follows (for $N = 1$):

$$C = \frac{(1-x) - e^{-r(T)T} + \frac{1}{S_0}\text{DIBP}_{t=0}}{\sum_{i=1}^{n}\left(e^{-r(t_i)\ t_i}\right)}. \tag{7.9.4}$$

We can observe the evolution of the value of the DIBP option and the value of the notes according to the barrier in Figure 7.9.1. Therefore, we observe that, as expected, when the maturity increases, the risk of crossing the barrier increases, increasing the proportion of obtaining a cash flow. Therefore, since the payoff corresponds to « minus » this option, when the maturity increases, the price of the BRC product decreases since the client has less chance of receiving the full amount of his nominal at maturity.

Concerning the value of the coupons, the barrier effect is straightforward: when the barrier increases, the value of DIBP and therefore of the coupons increases. However, with regard to the effect of maturity, two effects face each other. On the one hand, when the maturity increases, the value of DIBP increases and on the other hand, the number of coupons increases. It can therefore be seen that the marginal effect of

[9] As mentioned earlier, the payment frequency could vary but will only have an effect on the value of the coupons so that the fair value equilibrium condition is met.

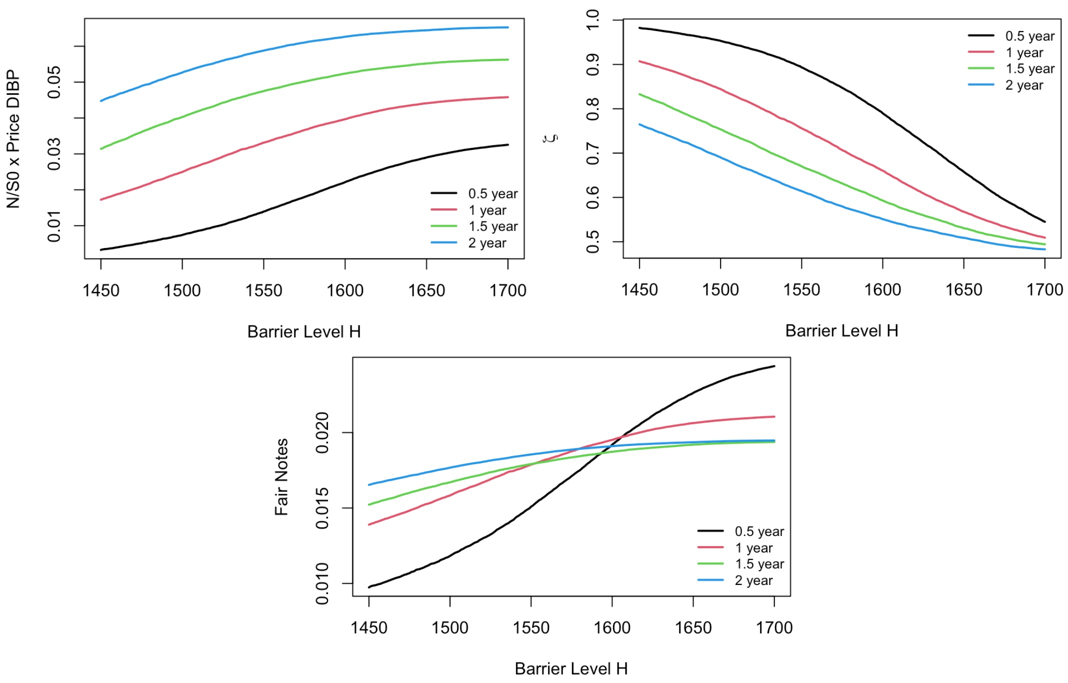

Figure 7.9.1: Price of $N/S_0\times$ DIBP with $N = 1$ (left), the proportion ζ (right) according to the barrier level H and the maturity T and the coupon rate (down).

the barrier tends to flatten out as the maturity increases.

We also notice that when the maturity increases further, the value of the coupons tends to converge (blue and green curves) but the proportion of clients who will receive their nominal at maturity tends to decrease. These two effects tend to cancel each other out as clients will receive more and more coupons but less and less of their nominal (on average).

Thus, the previous observations will impose that the fair value will be fixed and that it is the coupons that will vary according to the maturity and the barrier level. Therefore, several products are possible depending on the client's expectations and in particular his risk aversion. In order to understand what is possible with this BRC product, we propose 3 different products whose characteristics are shown in table 7.9.1:

- a « dynamic » product with a higher barrier and therefore higher or more frequent coupons but with a possible loss in the principal delivered at the end.
- a « secure » product with a lower barrier but lower coupons during the contract and/or less frequent and thus a greater chance of recovering the entire initial investment.
- a « balanced » product which corresponds to a kind of in-between: a medium barrier with medium value coupons.

Product	Barrier Level	Recovery Frequency	Coupon Rate	Maturity	Price
Secure	1500$	95.32%	1.18%	0.5 year	0.99$
Balanced	1600$	78.98%	1.92%	0.5 year	0.99$
Dynamic	1700$	54.50%	2.44%	0.5 year	0.99$

Table 7.9.1: The three final products proposed and their features.

We note that each of the 3 products promotes the lowest maturity. In the case where one seeks to maximise the chances of recovering the initial investment (secure), one seeks to maximise ζ and it is the 6-month maturity that allows this (with a low barrier). When one seeks to maximise the return on investment (dynamic), one seeks to maximise the value of the coupons and it is also the 6-month maturity that allows this (with a high barrier). Finally, when we choose an intermediate solution (balanced), we take a medium barrier, where the coupon curves intersect, and we note that it seems relevant in this case to favour the lowest maturity as well, since it allows us to offer the same return on investment while maximising the probability of recovering our initial investment. Therefore, we understand that the nature of the BRC is indeed to be a short-term product.

We can now look at the influence of the model concerned on the pricing of this product. Since we have imposed that the price of the coupons are designed to match the fair value of the product, we know that whatever the model, the price of the product is fixed and will be the same regardless of the model considered. However, Figure 7.9.2 shows that the structure of the product, i.e. the coupon rate, can vary greatly depending on the model.

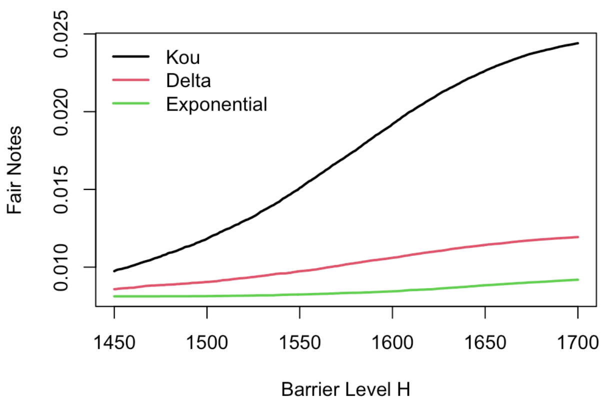

Figure 7.9.2: Comparison of coupon rates according to the model considered for the Gold market.

We can therefore see that, unlike the vanilla options, which obtained similar prices depending on the model, the exotic derivatives that are path-dependent have completely different prices, or at least with regard to the guarantees offered (coupons) concerning the BRC product contract. We notice that the model with the Exponential kernel is the one that offers the lowest coupons, followed by the model with the Delta kernel and finally the Kou model. This can perhaps be justified by the fact that the Exponential kernel model assumes the highest convenience yield according to the assumptions with which the latter was defined previously (i.e., the process memory), followed by the Delta kernel model and finally the Kou model, which does not. However, as mentioned before, since we have seen that the model with the Exponential kernel was the most consistent, it is the one we will retain. Nevertheless, let us observe that the difference seems to be substantial, which highlights the interest of studying the calibration of processes with memory in order to try to stick as well as possible to the dynamics of the underlying.

It is therefore important to ensure that the coupons obtained from the Exponential kernel model are sufficient to adjust the convenience yield (storage, transport, etc.). To increase the value of the coupons it may be possible to decrease the fees included in the product[10]. Decreasing the frequency and/or decreasing the number of coupons will not maximise the revenue obtained over the total period under consideration, rather it corresponds to features that can be chosen by the customer according to the costs he has to realise: matching the time of payment for the transport of the commodities with the time of payment of the note etc.

[10]The issuer must however remain vigilant to maintain its profitability.

Chapter 8

Conclusion

The main motivation for this thesis is based on the following observation. Since the commodity market is characterised by a non-Gaussianity in the price dynamics and often does not allow to trade its underlying spot directly, we are in highly incomplete market leading to a risk measure that is no longer unique. Thus, it is necessary to introduce a probability measure that deviates from the risk neutral one. As in Benth (2020), we have studied this « pricing measure » by introducing a risk premium through a convenience yield assimilated to memory terms.

Based on this point, this master thesis reviewed the main memory models through the generalized Langevin equation. We first introduced a model with instantaneous memory through a Delta kernel, for which the past history is based only on the last observation. Then, we developed a model with short-term persistent past influence form the past history, the Exponential kernel. Finally, we presented a model with strong residual of the past history that can be embedded in a long-term dependence of the underlying, the Power kernel. We know that this memory feature as well as the mean reversion in the process are two effects that can be assimilated to commodity markets. Although they seem contradictory, they can actually provide more exotic features that can be interesting in the markets in practice: reversion to the mean followed by an upward or downward trend depending on whether the market considers this decline to be persistent or not, etc., thus allowing in a way to consider a reversion to the mean that evolves over time.

The main contribution of this thesis is based on the study of a practical case on the pricing of derivatives using these memory processes. As we studied the different markets considered (Gold, Copper, Oil, Corn), it became noticeable that all of them showed quite distinct patterns. Indeed, the Copper and Oil markets were characterised by instantaneous memory models (Delta kernel), whereas the model adopted for the Corn market is the Kou model used as a benchmark, and finally the Gold model seems to retain a model with a price dynamic that takes into account its past evolution (Exponential kernel). In this practical part, we have observed that the prices of vanilla options as well as exotic options (through fair notes) vary considerably according to the models used.

Despite these findings, future tracks can be explored in order to possibly bring the different markets even closer to the reality of the markets:

- The implementation of a whole model with regard to convenience yield and interest rate. Indeed, in his article, Benth (2020) suggests the possibility of developing bivariate Ornstein-Uhlenbeck (OU) dynamics for these two processes.
- Study the calibration performed in this thesis using financial products (i.e. vanilla options, often paid data) in order to limit the uncertainty linked to real world projections.

Chapter 9

Appendix

9.1 Martingale, Local Martingale & Semi-martingale

Consider an usual probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with an information flow $\mathcal{F}_t$. First, a stochastic process $(X_t)_{t\geq 0}$ is a **martingale** with respect to an filtration $\mathcal{F}_t$ if it satisfies the following conditions:

- X_t is $\mathcal{F}_t$-adapted (X is non anticipating),
- $\mathbb{E}[|X_t|] < \infty \ \forall \ t \in [0, T]$,
- $\mathbb{E}[X_t|\mathcal{F}_s] = X_s$ for any $s < t$.

Thus, $(X_t)_{t\geq 0}$ is a martingale if and only if $\mathbb{E}[X_T] = \mathbb{E}[X_0]$ with T, all the stopping times. T is a stopping time if $\{w : T(w) \leq t\} = \{T \leq t\} \in \mathcal{F}_t \ \forall \ t > 0$.

Furthermore, a stochastic process $(X_t)_{t\geq 0}$ is a **local martingale** if it satisfies the properties of a martingale at least locally. If there exists a sequence of stopping times T_n that almost surely (a.s.) tends towards infinity such as, $T_n < T$ a.s. with $T = \lim_{n\to\infty} T_n$ a.s. and therefore $X_{t\wedge T_n}$ is a martingale for each n. Obviously, every martingale is a local martingale.

Finally, a stochastic process $(X_t)_{t\geq 0}$ is a **semi-martingale** if it assumes a Dood-Meyer decomposition such as:

$$X_t = X_0 + M_t + A_t \tag{9.1.1}$$

where $M_0 = A_0 = 0$, X_0 is $\mathcal{F}_0$-measurable, M_t is a $\mathcal{F}_t$-adapted local martingale and A_t is a càdlàg adapted process locally square integrable martingale. From the Lévy-Itô decomposition, it comes that any Lévy process is a semi-martingale with

$$M_t = \sigma W_t + \int_0^t \int_{|z|<1} z\tilde{N}(\mathrm{d}s \times \mathrm{d}z) \qquad \text{and} \qquad A_t = \mu t + \int_0^t \int_{|z|\geq 1} zN(\mathrm{d}s \times \mathrm{d}z). \tag{9.1.2}$$

9.2 Markov Property

The Markov property (or memoryless property) is the property that a stochastic process is memoryless, i.e. it is not possible to predict the future more accurately with any additional past information than the present information. This can be expressed as follows: in a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, let us define a stochastic process $(X_t)_{0\leq t\leq T}$ adapted to the information set $\mathcal{F}_t$, one can say that this process is a Markovian one if for each event $B \subset \Omega$ the next relation is fulfilled :

$$\mathrm{P}\big(X_t \in B|\mathcal{F}_s\big) = \mathrm{P}\big(X_t \in B|X_s\big). \tag{9.2.1}$$

9.3 Itô's-Lemma

Let X_t be a Jump-Diffusion process defining by the following expression :

$$X_t = X_0 + \int_0^t \mu_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + \sum_{i=1}^{N_t} Y_i, \tag{9.3.1}$$

where the first three terms represent the continuous part of the process, X_t^c, and where μ_t and σ_t are $\mathcal{F}_t$-adapted (nonanticipating) with $\mathbb{E}[\int_0^T |\mu_s|\mathrm{d}s] < \infty$ and $\mathbb{E}[\int_0^T \sigma_s^2 \mathrm{d}s] < \infty$.

Consider now a $C^{1,2}$ function $f(t,x)$ from $[0,T] \times \mathbb{R} \to \mathbb{R}$ with X_t a Jump-Diffusion as presented in (9.3.1), then :

$$\begin{aligned} f(t, X_t) = f(0, X_0) + \int_0^t \frac{\partial f(u, X_u)}{\partial u}\mathrm{d}u + \int_0^t \frac{\partial f(u, X_u)}{\partial x} dX_u^c \\ + \frac{1}{2}\int_0^t \frac{\partial^2 f(u, X_u)}{\partial x^2} d[X^c, X^c]_u + \sum_{i=0}^{N_t}\Big(f(t_i, X_{t_i} + Y_i) - f(t_i, X_{t_i})\Big). \end{aligned} \tag{9.3.2}$$

The equivalent in differential notation is such that :

$$\mathrm{d}f(t, X_t) = \frac{\partial f(t, X_t)}{\partial t}\mathrm{d}t + \frac{\partial f(t, X_t)}{\partial x}\mathrm{d}X_t^c + \frac{1}{2}\frac{\partial^2 f(t, X_t)}{\partial x^2} d[X^c, X^c]_t + \Big(f(t, X_t + Y) - f(t, X_t)\Big)\mathrm{d}N_t. \tag{9.3.3}$$

9.4 Put-Call Parity

An important property of European options is the put-call parity, which enhances the relationship between a call and a put for the same strike K and maturity T:

$$P(t, S_t, K) + S_t = C(t, S_t, K) + Ke^{-r(T-t)}. \tag{9.4.1}$$

Without this relationship, an opportunity for arbitration is created, but note that in practice some costs (e.g. transaction costs) mean this relationship will not exactly hold.

9.5 Alternative Demonstration : Delta Kernel

There is an alternative demonstration to the one presented in the paper with regard to the process having the Delta function as the memory function. Considering the SDE process with the Delta kernel such as $M(t) = a\delta_0(t)$, the Fourier-Laplace transforms are in this case such as :

$$\begin{cases} \mathcal{L}[M(s)] = a \\ \mathcal{L}^{-1}[H(s)] = \mathcal{L}^{-1}\Big[\dfrac{1}{s - \mathcal{L}[M(s)] + \beta}\Big] \end{cases} \tag{9.5.1}$$

Using the Laplace-Fourier table, i.e. $\mathcal{L}^{-1}\Big[\frac{1}{s+a}\Big] = e^{-at}$:

$$\begin{aligned} \mathcal{L}^{-1}[H(s)] &= \mathcal{L}^{-1}\Big[\frac{1}{s - a + \beta}\Big] \\ &= e^{-(\beta - a)t} \end{aligned} \tag{9.5.2}$$

The process X_t in (3.0.9) becomes :

$$\begin{aligned} X_t &= e^{-(\beta-a)t}X_0 + \beta\theta\int_0^t e^{-(\beta-a)(t-u)}\mathrm{d}u + \int_0^t \chi(u-)e^{-(\beta-a)(t-u)}\mathrm{d}Lu \\ &= e^{-(\beta-a)t}X_0 + \beta\theta\Big[\frac{1 - e^{-(\beta-a)t}}{\beta - a}\Big] + \int_0^t \chi(u-)e^{-(\beta-a)(t-u)}\mathrm{d}Lu \end{aligned} \tag{9.5.3}$$

which is the same expression as the equation (3.1.5) and allows us to conclude.

9.6 Demonstration : Kou

We assume that the dynamics of the log-returns are driven by a Kou model under the measure $\mathbb{P}$ such that :

$$\mathrm{d}X_t = \mu \mathrm{d}t + \sigma \mathrm{d}W_t + Y \mathrm{d}N_t \tag{9.6.1}$$

which, thanks to Ito's lemma with $f(t,x) = \exp(x)$, has equivalent price dynamics

$$\frac{\mathrm{d}S_t}{S_t} = \Big(\mu + \frac{\sigma^2}{2}\Big)\mathrm{d}t + \sigma \mathrm{d}W_t + (e^Y - 1)\mathrm{d}N_t \tag{9.6.2}$$

and becomes under the pricing measure $\mathbb{Q}$ with $a = \beta = 0$ for each of the two kernels ($M(t) = a\delta_0(t)$ and $M(t) = ae^{-bt}$) such that

$$\frac{\mathrm{d}S_t}{S_t} = \Big\{r - \lambda\Big(\frac{p\eta_1}{\eta_1 - 1} + \frac{(1-p)\eta_2}{\eta_2 + 1} - 1\Big)\Big\}\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} + (e^Y - 1)\mathrm{d}N_t^{\mathbb{Q}}. \tag{9.6.3}$$

Thus, thanks to Ito's lemma with $f(t,x) = \log(x)$, we can find the dynamics of log-returns under the pricing measure $\mathbb{Q}$:

$$\mathrm{d}X_t = \Big\{r - \frac{\sigma^2}{2} - \lambda\Big(\frac{p\eta_1}{\eta_1 - 1} + \frac{(1-p)\eta_2}{\eta_2 + 1} - 1\Big)\Big\}\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}} + Y \mathrm{d}N_t^{\mathbb{Q}}. \tag{9.6.4}$$

Finally, by direct integration from inception to t we find the expected result and this allows us to conclude :

$$X_t = X_0 + \Big\{r - \frac{\sigma^2}{2} - \lambda\Big(\frac{p\eta_1}{\eta_1 - 1} + \frac{(1-p)\eta_2}{\eta_2 + 1} - 1\Big)\Big\}t + \sigma W_t^{\mathbb{Q}} + \sum_{i=1}^{N_t} Y_i. \tag{9.6.5}$$

9.7 Demonstration : Ornstein-Uhlenbeck

We assume that the dynamics of the log-returns are driven by an Ornstein-Uhlenbeck model under the measure $\mathbb{P}$ such that :

$$\mathrm{d}X_t = \beta(\theta - X_t)\mathrm{d}t + \sigma \mathrm{d}W_t \tag{9.7.1}$$

which, thanks to Ito's lemma with $f(t,x) = \exp(x)$, has equivalent price dynamics

$$\frac{\mathrm{d}S_t}{S_t} = \Big(\beta(\theta - X_t) + \frac{\sigma^2}{2}\Big)\mathrm{d}t + \sigma \mathrm{d}W_t \tag{9.7.2}$$

and becomes under the pricing measure $\mathbb{Q}$ with $a = 0$ for each of the two kernels $\big(M(t) = a\delta_0(t)$ and $M(t) = ae^{-bt}\big)$ such that

$$\frac{\mathrm{d}S_t}{S_t} = \Big(r_t + \beta(\theta - X_t)\Big)\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}}. \tag{9.7.3}$$

Thus, thanks to Ito's lemma with $f(t,x) = \log(x)$, we can find the dynamics of log-returns under the pricing measure $\mathbb{Q}$:

$$\mathrm{d}X_t = \Big(r_t - \frac{\sigma^2}{2} + \beta(\theta - X_t)\Big)\mathrm{d}t + \sigma \mathrm{d}W_t^{\mathbb{Q}}. \tag{9.7.4}$$

Finally, by using the demonstration presented previously for the construction of the kernel Delta model and applying it to this particular model, where the reversion level is $\theta + r - \frac{\sigma^2}{2}$, we find the expected result and can conclude :

$$\begin{aligned} X_t &= e^{-\beta(t-s)} X_s + \Big(\beta\theta + r - \frac{\sigma^2}{2}\Big) \int_s^t e^{-\beta(t-u)} \mathrm{d}u + \int_s^t e^{-\beta(t-u)} \mathrm{d}W_u \\ &= e^{-\beta(t-s)} X_s + \Big(\theta + \frac{r - \frac{\sigma^2}{2}}{\beta}\Big) \times \Big(1 - e^{-\beta(t-s)}\Big) + \int_s^t e^{-\beta(t-u)} \mathrm{d}W_u. \end{aligned} \tag{9.7.5}$$

9.8 Demonstration : Characteristic Function Lévy

Eberlein and Raible (1999) proposes that for any integrable function $f : \mathbb{R} \to \mathbb{C}$, the following relation holds

$$\mathbb{E}\Big[\exp\Big(w \int_s^t f(u) \mathrm{d}L_u\Big) | \mathcal{F}_s\Big] = \exp\Big(\int_s^t \psi\big(wf(u)\big) \mathrm{d}u\Big) \tag{9.8.1}$$

where $\psi(.)$ is the characteristic exponent. This result could be derive from the fact that

$$\begin{aligned} \mathbb{E}\Big[\exp\Big(w \int_s^t f(u) \mathrm{d}L_u\Big) | \mathcal{F}_s\Big] &= \lim_{n\to\infty} \mathbb{E}\Big[\exp\Big(w \sum_{j=1}^n f(u_j)(L_{u_j} - L_{u_{j-1}})\Big) | \mathcal{F}_s\Big] \\ &= \lim_{n\to\infty} \prod_{j=1}^n \mathbb{E}\Big[\exp\Big(wf(u_j)(L_{u_j} - L_{u_{j-1}})\Big) | \mathcal{F}_s\Big] \end{aligned} \tag{9.8.2}$$

and by using the nested expectation,

$$\mathbb{E}[.|\mathcal{F}_s] = \mathbb{E}\big[\mathbb{E}[.|\mathcal{F}_{u_{j-1}}] | \mathcal{F}_s\big], \tag{9.8.3}$$

it leads to,

$$\mathbb{E}\Big[\exp\Big(wf(u_j)(L_{u_j} - L_{u_{j-1}})\Big) | \mathcal{F}_{u_{j-1}}\Big] = \exp\Big((u_j - u_{j-1})\psi(wf(u_j))\Big), \tag{9.8.4}$$

and finally,

$$\begin{aligned} \mathbb{E}\Big[\exp\Big(w \int_s^t f(u) \mathrm{d}L_u\Big) | \mathcal{F}_s\Big] &= \lim_{n\to\infty} \prod_{j=1}^n \mathbb{E}\Big[\exp\Big(\psi\big(wf(u_j)\big)\Delta j\Big) | \mathcal{F}_s\Big] \\ &= \mathbb{E}\Big[\exp\Big(\sum_{j=1}^n \psi\big(wf(u_j)\big)\Delta j\Big) | \mathcal{F}_s\Big] \\ &= \exp\Big(\int_s^t \psi\big(wf(u)\big) \mathrm{d}u\Big) \end{aligned} \tag{9.8.5}$$

which allows us to find the announced result and to conclude.

9.9 Study of the normality of log returns.

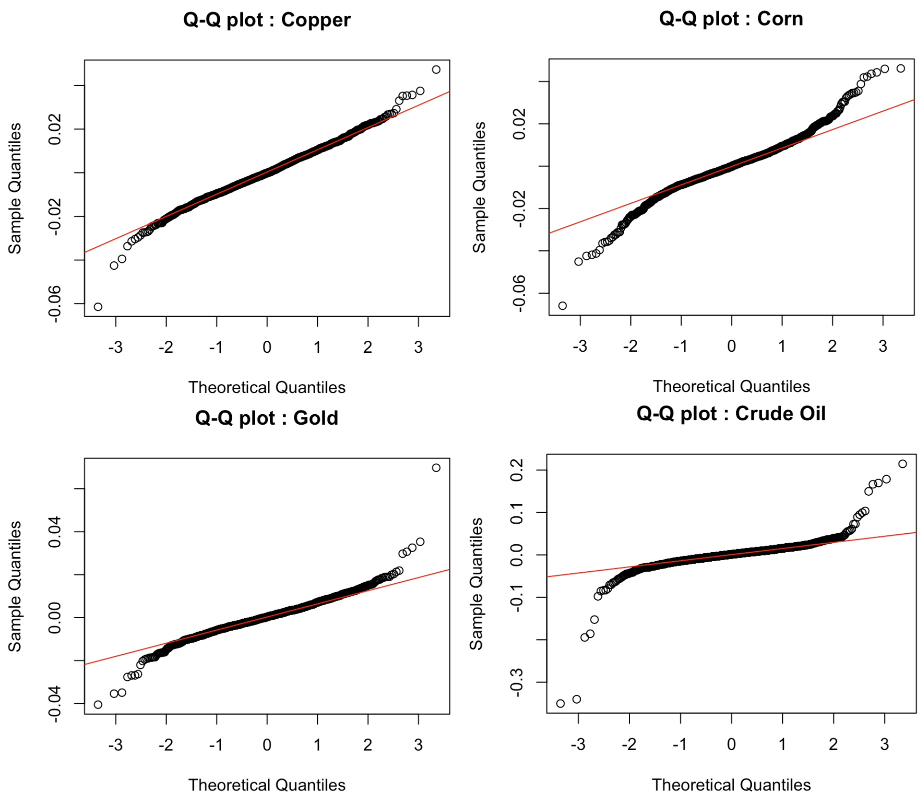

Figure 9.9.1: QQ-plot of the different markets allow to highlight the non-normality of the log-returns.

Bibliography

[1] David Applebaum. *Lévy processes and stochastic calculus.* Number 93 in Cambridge studies in advanced mathematics. Cambridge University Press, Cambridge, UK ; New York, 2004.

[2] Ole E. Barndorff-Nielsen, Fred Espen Benth, and Almut E. D. Veraart. Modelling energy spot prices by volatility modulated Lévy-driven Volterra processes. *Bernoulli*, 19(3), August 2013.

[3] Fred Espen Benth, Asma Khedher, and Michèle Vanmaele. Pricing of Commodity Derivatives on Processes with Memory. *Risks*, 8(1):8, January 2020.

[4] Fred Espen Benth, Claudia Klüppelberg, Gernot Müller, and Linda Vos. Futures pricing in electricity markets based on stable CARMA spot models. *Energy Economics*, 44:392–406, July 2014.

[5] Mats Brodén and Peter Tankov. Tracking errors from disc hedging in exponential Lévy models. *International Journal of Theoretical and Applied Finance*, 14(06):803–837, September 2011.

[6] Peter Carr and Dilip Madan. Option valuation using the fast Fourier transform. *The Journal of Computational Finance*, 2(4):61–73, 1999.

[7] Peter Carr and Dilip Madan. Saddlepoint methods for option pricing. *The Journal of Computational Finance*, 13(1):49–61, September 2009.

[8] Roy Cerqueti and Viviana Fanelli. Long memory and crude oil's price predictability. *Annals of Operations Research*, 299(1-2):895–906, April 2021.

[9] Ricardo Crisóstomo. Speed and biases of Fourier-based pricing choices: a numerical analysis. *International Journal of Computer Mathematics*, 95(8):1565–1582, August 2018.

[10] Michel Denuit and Arthur Charpentier. *Mathématiques de l'assurance non-vie. T. 1: Principes fondamentaux de théorie du risque.* Collection économie et statistiques avancées. Economica, Paris, 2004.

[11] Jean-Loup Dupret and Donatien Hainaut. Portfolio insurance under rough volatility and Volterra processes. 2021.

[12] Ernst Eberlein and Sebastian Raible. Term structure models driven by general levy processes. *Mathematical Finance*, 9(1):31–53, January 1999.

[13] Eugene F. Fama. Market efficiency, long-term returns, and behavioral finance. *SSRN Electronic Journal*, 1998.

[14] Gustaf Gripenberg, Londen Stig-Olof, and Olof Staffans. *Volterra Integral and Functional Equations.* Cambridge University Press, Cambridge, 1990.

[15] Donatien Hainaut. Clustered Lévy processes and their financial applications. *Journal of Computational and Applied Mathematics*, 319:117–140, August 2017.

[16] Donatien Hainaut. Continuous Mixed-Laplace Jump Diffusion Models for Stocks and Commodities. *Quantitative Finance and Economics*, 1(2):145–173, 2017.

[17] Donatien Hainaut. Calendar Spread Exchange Options Pricing with Gaussian Random Fields. *Risks*, 6(3):77, August 2018.

[18] Donatien Hainaut. Lévy interest rate models with a long memory. 2021.

[19] H. E. Hurst. Long-Term Storage Capacity of Reservoirs. *Transactions of the American Society of Civil Engineers*, 116(1):770–799, 1951.

[20] Eduardo Abi Jaber, Martin Larsson, and Sergio Pulido. Affine Volterra processes. October 2019.

[21] Ralf Korn, Elke Korn, and Gerald Kroisandt. *Monte Carlo methods and models in finance and insurance*. Chapman & Hall/CRC financial mathematics series. CRC Press/Taylor & Francis, Boca Raton, FL, 2010. OCLC: ocn463639246.

[22] S. G. Kou. A Jump-Diffusion Model for Option Pricing. *Management Science*, 48(8):1086–1101, August 2002.

[23] S. G. Kou and Hui Wang. Option Pricing Under a Double Exponential Jump Diffusion Model. *Management Science*, 50(9):1178–1192, September 2004.

[24] Andreas E. Kyprianou, Wim Schoutens, and Paul Wilmott, editors. *Exotic option pricing and advanced Lévy models*. John Wiley, Chichester, England ; Hoboken, NJ, 2005. OCLC: ocm59877855.

[25] Alan L. Lewis. A Simple Option Formula for General Jump-Diffusion and Other Exponential Lévy Processes. *SSRN Electronic Journal*, 2001.

[26] Fabrizio Lillo and J. Doyne Farmer. The Long Memory of the Efficient Market. *Studies in Nonlinear Dynamics & Econometrics*, 8(3), January 2004.

[27] Andrew W. Lo. Long-Term Memory in Stock Market Prices. *Econometrica*, 59(5):1279–1313, 1991.

[28] Jacques Lévy Véhel and Evelyne Lutton, editors. *Fractals in engineering: new trends in theory and applications*. Springer, London, 1st ed edition, 2005.

[29] Francesco Mainardi. Why the mittag-leffler function can be considered the queen function of the fractional calculus? *Entropy 22*, 9(1):1359, 2020.

[30] Thomas Mikosch, Jens-Peter Kreiß, Richard A. Davis, and Torben Gustav Andersen, editors. *Handbook of Financial Time Series*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.

[31] A. Montanari, M.S. Taqqu, and V. Teverovsky. Estimating long-range dependence in the presence of periodicity: An empirical study. *Mathematical and Computer Modelling*, 29(10-12):217–228, May 1999.

[32] Salvador Ortiz-Latorre. A New Pricing Measure in the Barndorff-Nielsen–Shephard Model for Commodity Markets. In Josep Díaz, Lefteris Kirousis, Luis Ortiz-Gracia, and Maria Serna, editors, *Extended Abstracts Summer 2015*, volume 6, pages 133–139. Springer International Publishing, Cham, 2017.

[33] Raphael Paschke and Marcel Prokopczuk. Commodity derivatives valuation with autoregressive and moving average components in the price dynamics. *Journal of Banking & Finance*, 34(11):2742–2752, November 2010.

[34] Andrea Pascucci. *PDE and Martingale Methods in Option Pricing*. Bocconi & Springer Series. Springer Milan, Milano, 2011.

[35] Philip Protter. Volterra Equations Driven by Semimartingales. *The Annals of Probability*, 13(2), May 1985.

[36] Piergiacomo Sabino and Nicola Cufaro Petroni. Fast Pricing of Energy Derivatives with Mean-Reverting Jump-diffusion Processes. *Applied Mathematical Finance*, pages 1–22, April 2021.

[37] Paul A. Samuelson. Proof That Properly Discounted Present Values of Assets Vibrate Randomly. *The Bell Journal of Economics and Management Science*, 4(2):369–374, 1973.

[38] Wim Schoutens. *Lévy processes in finance: pricing financial derivatives.* Wiley series in probability and statistics. J. Wiley, Chichester, West Sussex ; New York, 2003.

[39] Wim Schoutens and Jessica Cariboni. *Lévy processes in credit risk.* Wiley finance. John Wiley & Sons, Chichester, U.K, 2009. OCLC: ocn312636198.

[40] Wim Schoutens, Erwin Simons, and Jurgen Tistaert. A perfect calibration! now what? *Wilmott*, 2004(2):66–78, March 2004.

[41] Eduardo Schwartz and James E. Smith. Short-Term Variations and Long-Term Dynamics in Commodity Prices. *Management Science*, 46(7):893–911, July 2000.

[42] Stavros J. Sioutis. Calibration and Filtering of Exponential Lévy Option Pricing Models. May 2017. arXiv: 1705.04780.

[43] Masafumi Takahashi. Non-ideal Brownian motion, generalized Langevin Equation and its application to the security market. *Financial Engineering and the Japanese Markets*, 3(2):87–119, July 1996.

[44] Peter Tankov. *Financial Modelling with Jump Processes.* Chapman and Hall/CRC, 0 edition, December 2003.

[45] D. V. Widder. What is the Laplace Transform? *The American Mathematical Monthly*, 52(8):419–425, October 1945.

[46] N. Yamada and T. Kunieda. Laplace transformation and operational calculus (in japanese). *Corona Publishing*, 1971.

[47] Bernt Øksendal and Agnès Sulem. *Applied Stochastic Control of Jump Diffusions.* Springer Berlin Heidelberg, Berlin, Heidelberg, 2007.