UNIVERSITÀ DEGLI STUDI DI PERUGIA
Dipartimento di Matematica e Informatica

CORSO DI LAUREA MAGISTRALE IN MATEMATICA

Tesi di Laurea

# Comparing three different valuation methods

*Laureando*
**Alessandro Bondi**

*Relatori*
**Ch.mo Prof. Thorsten Rheinländer**
**Ch.ma Prof.ssa Alessandra Cretarola**

**Anno Accademico 2018/2019**

# Introduction

Empirical studies suggest that financial markets are incomplete and hold unhedgeable and undiversifiable risks. Once the completeness assumption is removed, there typically exist several martingale measures consistent with the no-arbitrage principle that can be chosen as pricing measures (see, e.g., [19, 29]). In other words, the price of a derivative depends on the criteria followed to get it.

In this thesis, three methods for the simulation of European call option prices are introduced: pricing under the Esscher measure, calibration to real–traded securities with an entropic penalty term and nonparametric estimation of risk-neutral densities. Furthermore, the first two methods are implemented using the programming language $R$ and the simulated derivative prices are compared to real data retrieved from NASDAQ option chains. Thus, their relative performances are quantified, considering that the Esscher method could be the only feasible approach in new markets with few liquid securities. The work is structured as follows.

- The first chapter is devoted to the discussion of the theory that stands behind the main concepts of these methods. More precisely, classical topics regarding Lévy processes, such as the *Lévy–Kintchine Representation* and the *Lévy–Itô Decomposition theorem*, are treated and a connection between the generating triplet of a Lévy process and the more general concept of characteristics of semimartingales is presented.

- The second chapter analyses the *Esscher method* and is divided in two parts. The first presents the theory necessary to construct the Esscher measure for both exponential and linear processes. In the second, the dynamics of the stock prices are defined and sufficient conditions for the existence of the Esscher measure are established. Moreover, in the geometric case European call options prices are simulated using Esscher's as pricing measure and the results are compared to real data.

- The third chapter is intended to discuss the calibration method. The first part of it sets up the theoretical background, which consists in relative entropy of distributions and Lévy processes on the Skorokhod Space, while the second is devoted to the practical analysis.

- In Appendix A the original codes used to run simulations can be found.

- In Appendix B, the *nonparametric estimation method* proposed by M. Grith, W.K. Härdle, M. Schienle in [14] is presented and the idea of its application is briefly described.

# Contents

# Chapter 1

# Lévy Processes and Characteristics Of Semimartingales

In this chapter we introduce the theoretical results which are essential for the entire work. In particular we focus on Lévy processes, that will next be used to model stock prices, displaying the relation between their generating triplets and the more general concept of characteristics of a semimartingale.

## 1.1 Lévy Processes and Infinitely Divisible Distribution

**Definition 1.1.** A $\mathbb{R}^d$–valued stochastic process $X = \{X_t\}_{t\geq 0}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ is a *Lévy process* if the next conditions are satisfied:

i. for any choice of $n \in \mathbb{N}$ and $0 \leq t_0 < ... < t_n$, the random variables $X_{t_0}, X_{t_1} - X_{t_0}, ..., X_{t_n} - X_{t_{n-1}}$ are independent. This is the independent increments property;

ii. $X_0 = 0$ a.s.;

iii. the distribution of $X_{t+s} - X_s$ does not depend on $s$ for any $s, t \geq 0$. This is the temporal homogeneity, or stationary increments, property;

iv. $X$ is continuous in probability, that is, for every $t > 0$ and $\epsilon > 0$ it results $\lim_{s\to t} P(|X_s - X_t| > \epsilon) = 0$;

v. $X$ is right–continuous with left limits, i.e., it is a càdlàg process.

Dropping the condition (v.) we have a *Lévy process in law*; without the stationarity of increments the process is called *additive process*, and if it does not satisfy even the càdlàg property it will be an *additive process in law*.

**Definition 1.2.** The *characteristic function* $\widehat{\mu}$ of a *probability measure* (or *distribution*) $\mu$ on $\mathbb{R}^d$ is defined by

$$\widehat{\mu}(z) := \int_{\mathbb{R}^d} e^{i<z,x>} \mu(dx), \quad z \in \mathbb{R}^d.$$

The *characteristic function* of a random variable $X$ on $\mathbb{R}^d$ is defined by

$$\widehat{P_X}(z) := \int_{\mathbb{R}^d} e^{i<z,x>} P_X(dx), \quad z \in \mathbb{R}^d.$$

Given a distribution $\mu$ on $\mathbb{R}^d$, we have that $\widehat{\mu}(0) = 1, |\widehat{\mu}| \leq 1$ and $\widehat{\mu}$ is continuous in its domain. It is also well known that if two probability measures $\mu_1, \mu_2$ have the same characteristic function, i.e., $\widehat{\mu_1} = \widehat{\mu_2}$, then $\mu_1 = \mu_2$.

**Definition 1.3.** The *convolution* $\mu$ of two distributions $\mu_1$ and $\mu_2$ on $\mathbb{R}^d$ is denoted by $\mu = \mu_1 * \mu_2$ and it is the distribution on $\mathbb{R}^d$ defined by

$$\mu(B) := \int_{\mathbb{R}^d \times \mathbb{R}^d} 1_B(x+y)\,(\mu_1 \otimes \mu_2)(dx, dy), \quad B \in \mathcal{B}\left(\mathbb{R}^d\right),$$

where $\mu_1 \otimes \mu_2$ denotes the product measure on $\mathbb{R}^d \times \mathbb{R}^d$.

The convolution is a commutative and associative operation. It is possible to show that if $\mu = \mu_1 * \mu_2$, where $\mu_1$ and $\mu_2$ are two distributions on $\mathbb{R}^d$, then $\widehat{\mu} = \widehat{\mu_1}\,\widehat{\mu_2}$. We denote by $\mu^n$ the $n-$fold convolution of a distribution $\mu$ with itself, that is,

$$\mu^n = \underbrace{\mu * ... * \mu}_{n}.$$

**Definition 1.4.** A probability measure $\mu$ on $\mathbb{R}^d$ is *infinitely divisible* if for any $n \in \mathbb{N}$ there exists a distribution $\mu_n$ on $\mathbb{R}^d$ such that $\mu = \mu_n^n$.

A random variable $X$ is *infinitely divisible* if $P_X$ is so.

Equivalently, a probability measure $\mu$ on $\mathbb{R}^d$ is *infinitely divisible* if for any $n \in \mathbb{N}$ there exists a distribution $\mu_n$ on $\mathbb{R}^d$ such that $\widehat{\mu} = \widehat{\mu_n}^n$.
From the definition, it immediately follows that the convolution of two infinitely divisible distributions is infinitely divisible and that $\widehat{\mu}(z) \neq 0$ for any $z \in \mathbb{R}^d$, provided $\mu$ infinitely divisible. In the next example we recall the main features of the *Normal Inverse Gaussian* ($NIG$) distribution: it is important because a Lévy process $X = \{X_t\}_t$ with $X_1 \sim NIG$ provides a good fit to the data we are going to analyze.

**Example 1.1.** A probability measure $\mu$ on $\mathbb{R}$ is said to be $NIG$ with parameters $(\alpha, \beta, \mu, \delta)$, where $0 \leq |\beta| < \alpha$, $\mu \in \mathbb{R}$ and $\delta > 0$, if it has the following density:

$$f_{NIG}(x; \alpha, \beta, \mu, \delta) = \frac{\alpha}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x-\mu)\right) \frac{K_1\left(\alpha\delta\sqrt{1+\left(\frac{x-\mu}{\delta}\right)^2}\right)}{\sqrt{1+\left(\frac{x-\mu}{\delta}\right)^2}}, \tag{1.1}$$

with $K_1$ that denotes the modified Bessel function of the third kind with index 1 and $x \in \mathbb{R}$. For a definition of $K_1$ we refer to [1].

A $NIG$ distribution is infinitely divisible, since it is a particular *Generalized Hyperbolic* distribution (see, e.g., [5] and [13]). We refer to Remark 1.1 below for a thorough construction of such distribution. Given $X \sim NIG(\alpha, \beta, \mu, \delta)$, we have

$$E[X] = \mu + \frac{\delta\beta}{\sqrt{\alpha^2-\beta^2}}, \qquad \mathrm{Var}[X] = \frac{\delta\alpha^2}{(\alpha^2-\beta^2)^{3/2}},$$
$$\mathrm{Skew}[X] = 3\frac{\beta}{\alpha\left(\delta\sqrt{\alpha^2-\beta^2}\right)^{1/2}}.$$

Sometimes it is convenient to represent the four parameters of a $NIG$ distribution in the so–called *shape triangle*, introducing the couple $(\xi,\chi)$ defined by

$$\xi := \left(1+\delta\sqrt{\alpha^2-\beta^2}\right)^{-1/2}, \qquad \chi := \frac{\xi\beta}{\alpha}.$$

It follows that $0 \le |\chi| < \xi < 1$, so these new quantities actually lie in a triangle. Notice that $\chi$ is a measure of the skewness of the distribution: indeed, it has the same sign as $\beta$, which in turn determines the sign of Skew $[X]$.

In the case that $X$ is standard, its density simplifies to

$$f_X(x) = f_X(x;\xi,\chi), \quad x \in \mathbb{R},$$

i.e., the numbers of parameters reduces to 2. ///

**Remark 1.1.** A probability measure $\mu$ on $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ is said a *Generalized Inverse Gaussian (GIG)* distribution with parameters $\nu \in \mathbb{R}$, $\delta > 0$, $\gamma > 0$ if its density with respect to the Lebesgue measure is as follows:

$$f_{GIG}(x;\nu,\delta,\gamma) = \left(\frac{\gamma}{\delta}\right)^{\nu}\frac{1}{2K_\nu(\gamma\delta)}x^{\nu-1}\exp\left[-\frac{1}{2}\left(\delta^2x^{-1}+\gamma^2x\right)\right], \quad x>0. \tag{1.2}$$

In order to show that $f_{GIG}$ is a density function, we need the next representation ([30], formula 8, page 182) for $K_\nu$, the modified Bessel function of the third kind with index $\nu$:

$$K_\nu(x) = \frac{1}{2}\int_0^\infty y^{\nu-1}\exp\left[-\frac{x}{2}\left(y+\frac{1}{y}\right)\right]dy, \quad x>0. \tag{1.3}$$

In Figure 1.1 below there are the plots of the modified Bessel functions of the first three, nonnegative, integer orders. We have obtained them from a sample of randomly chosen points, in turn gotten with the software *R*. The computation of the next integral allows us to find the normalization constant in (1.2):

$$\int_0^\infty x^{\nu-1}e^{-\frac{1}{2}\left(\delta^2x^{-1}+\gamma^2x\right)}dx = \int_0^\infty x^{\nu-1}e^{-\frac{1}{2}\gamma\delta\left(\frac{\delta}{\gamma}x^{-1}+\frac{\gamma}{\delta}x\right)}dx$$
$$= \left(\frac{\delta}{\gamma}\right)^{\nu}\int_0^\infty y^{\nu-1}e^{-\frac{1}{2}\gamma\delta\left(y+\frac{1}{y}\right)}dy = 2\left(\frac{\delta}{\gamma}\right)^{\nu}K_\nu(\gamma\delta),$$

where in the second equality we made the substitution $y = \frac{\gamma}{\delta}x$. Besides, $K_\nu > 0$ in $\mathbb{R}^+$ from (1.3). Therefor $f_{GIG}$ is actually a density function on $\mathbb{R}^+$.

Let us fix other two parameters $\alpha$, $\beta \in \mathbb{R}$, with $0 \le |\beta| < \alpha$. For any $\mu \in \mathbb{R}$, $y \in \mathbb{R}^+$ we denote by $f_N(\cdot;\mu,\beta,y)$ the density of the probability measure on

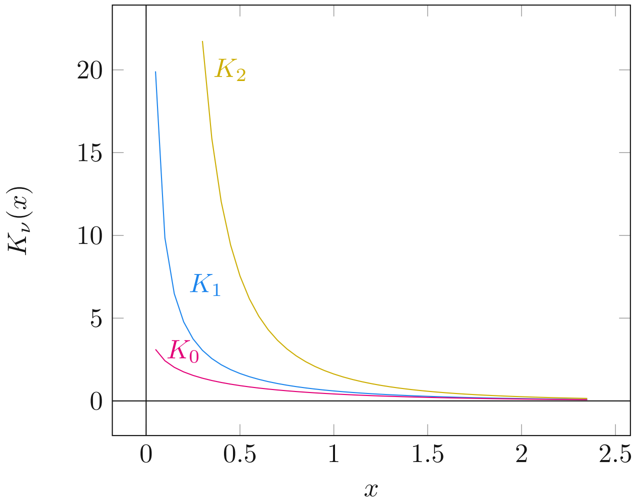



Figure 1.1: Modified Bessel functions of the first three, nonnegative, integer orders.

$(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ generated by a random variable $X \sim N(\mu + \beta y, y)$. It is possible to introduce a new distribution on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ considering the following function:

$$
\begin{aligned}
f_{GH}(x;\nu,\alpha,\beta,\mu,\delta) &:= \int_0^\infty f_N(x;\mu,\beta,y)\, f_{GIG}\left(y;\nu,\delta,\sqrt{\alpha^2-\beta^2}\right) dy \\
&= \frac{1}{\sqrt{2\pi}}\left(\frac{\sqrt{\alpha^2-\beta^2}}{\delta}\right)^{\nu} \frac{1}{2K_\nu\left(\delta\sqrt{\alpha^2-\beta^2}\right)} \\
&\qquad \int_0^\infty \frac{1}{\sqrt{y}} e^{-\frac{1}{2y}(x-\mu-\beta y)^2} y^{\nu-1} e^{-\frac{1}{2}\left(\delta^2 y^{-1} + \left(\alpha^2-\beta^2\right)y\right)} dy \\
&= \frac{1}{\sqrt{2\pi}}\left(\frac{\sqrt{\alpha^2-\beta^2}}{\delta}\right)^{\nu} \frac{1}{2K_\nu\left(\delta\sqrt{\alpha^2-\beta^2}\right)} e^{\beta(x-\mu)} \\
&\qquad \int_0^\infty y^{\nu-1-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left[\frac{1}{y}\left(\delta^2+(x-\mu)^2\right)+\alpha^2 y\right]\right\} dy \\
&= \frac{1}{\sqrt{2\pi}}\left(\frac{\sqrt{\alpha^2-\beta^2}}{\delta}\right)^{\nu} \frac{1}{2K_\nu\left(\delta\sqrt{\alpha^2-\beta^2}\right)} e^{\beta(x-\mu)} \\
&\qquad \left(\frac{\sqrt{\delta^2+(x-\mu)^2}}{\alpha}\right)^{\nu-\frac{1}{2}} \int_0^\infty z^{\left(\nu-\frac{1}{2}\right)-1} \exp\left[-\frac{1}{2}\left(z^{-1}+z\right)\alpha\sqrt{\delta^2+(x-\mu)^2}\right] dz \\
&= \frac{\left(\sqrt{\alpha^2-\beta^2}\right)^{\nu}}{\sqrt{2\pi}\delta^{\nu}\alpha^{\nu-\frac{1}{2}} K_\nu\left(\delta\sqrt{\alpha^2-\beta^2}\right)} e^{\beta(x-\mu)} \frac{K_{\nu-\frac{1}{2}}\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right)}{\left(\sqrt{\delta^2+(x-\mu)^2}\right)^{\frac{1}{2}-\nu}}, \quad x \in \mathbb{R},
\end{aligned}
\tag{1.4}
$$

where in the last but one equality we made the substitution $z = \frac{\alpha}{\sqrt{\delta^2+(x-\mu)^2}} y$ and in the last we used (1.3). A straightforward application of Tonelli's theorem shows that $f_{GH}(\cdot;\nu,\alpha,\beta,\mu,\delta)$ is a density function: the corresponding probability measure on $(\mathbb{R},\mathcal{B}(\mathbb{R}))$ is called *Generalized Hyperbolic* distribution.

We restore the $NIG$ density (1.1) taking $\nu = -\frac{1}{2}$ in (1.4). In fact, for every $\nu \in \mathbb{R}$, it results $K_{-\nu} = K_\nu$ and the representation ([1], formula 9.6.23, page 376)

$$K_\nu(x) = \frac{\sqrt{\pi}\left(\frac{1}{2}x\right)^\nu}{\Gamma\left(\nu+\frac{1}{2}\right)} \int_1^\infty e^{-xt}\left(t^2-1\right)^{\nu-\frac{1}{2}} dt, \quad x>0,$$

which holds for every $\nu > -\frac{1}{2}$, enables us to conclude

$$K_{-\frac{1}{2}}(x) = K_{\frac{1}{2}}(x) = -\sqrt{\frac{\pi x}{2}}\frac{1}{x}\left[e^{-xt}\right]_1^\infty = \sqrt{\frac{\pi}{2}}\frac{e^{-x}}{\sqrt{x}}, \quad x>0.$$

**Notation.** We denote with $\widehat{\mu}^t$ the $t$–th power of the characteristic function $\widehat{\mu}$, that is, $\widehat{\mu}^t(z) = (\widehat{\mu}(z))^t$ for any $z \in \mathbb{R}^d$.

The following result is due to K.–I. Sato (Lemma 7.9 in [27]).

**Proposition.** *If $\mu$ is infinitely divisible, for any $t \in [0,\infty)$ the function $\widehat{\mu}^t$ is the characteristic function of a probability measure on $\mathbb{R}^d$ which is infinitely divisible.*

The distribution whose existence is ensured by this proposition is obviously unique and it will be denoted by $\mu^t$.

There is a strict relation between Lévy processes and infinitely divisible distributions, which is shown in the next example.

**Example 1.2.** Given a Lévy process $X = \{X_t\}_t$, for every $t \geq 0$ the random variable $X_t$ is infinitely divisible. Indeed, the case $t=0$ is trivial, because by (ii.) in Definition 1.1 we have $P_{X_0} = \delta_0$, which is infinitely divisible as

$$\delta_0 = \delta_0^n \quad \text{for any } n \in \mathbb{N}.$$

Fix $t>0$ and $n \in \mathbb{N}$. Let $t_k := \frac{kt}{n}$ for $k = 0,...,n$ and denote by

$$\mu_n := P_{\left(X_{t_k}-X_{t_{k-1}}\right)},$$

which does not depend on the choice of $k \in \{1,...,n\}$ by temporal homogeneity. We can now read

$$X_t = X_0 + (X_{t_1}-X_0) + ... + \left(X_{t_n}-X_{t_{n-1}}\right),$$

so by independence of the increments we can conclude $\widehat{P_{X_t}} = \widehat{\mu_n}^n$, that is, $P_{X_t} = \mu_n^n$. ///

**Remark 1.2.** The previous rational still holds for Lévy processes in law, since we do not need the càdlàg condition. It is also possible to show (see Theorem 9.1 in [27]) that $X_t$ is infinitely divisible for any $t \geq 0$ when $X = \{X_t\}_t$ is just an additive process in law, so the temporal homogeneity is not necessary.

We end this section by presenting a classical theorem which strengthens the bond we have just introduced.

**Theorem 1.1.** *For every infinitely divisible distribution $\mu$ on $\mathbb{R}^d$, there exists a Lévy process $X = \{X_t\}_t$ such that $P_{X_1} = \mu$.*

*This process is unique up to identity in law and*

$$P_{X_t} = \mu^t, \quad t \geq 0.$$

We refer to Theorem 7.10 and Corollary 11.6 in [27] for a proof.

## 1.2 Lévy–Kintchine Representation

The following, renowned theorem provides a representation of the characteristic function of any infinitely divisible distribution. We denote by $D$ the unit ball in $\mathbb{R}^d$, that is, $D := \left\{x \in \mathbb{R}^d : |x| \leq 1\right\}$.

**Theorem 1.2.** a. *If $\mu$ is an infinitely divisible distribution on $\mathbb{R}^d$, then*

$$\widehat{\mu}(z) = \exp\left[-\frac{1}{2} < z, Az > +i < \gamma, z > + \int_{\mathbb{R}^d} \left(e^{i<z,x>} - 1 - i < z, x > 1_D(x)\right) \nu(dx)\right], \quad z \in \mathbb{R}^d, \tag{1.5}$$

*where $A$ is a symmetric, positive semidefinite $d \times d$ matrix, $\nu$ is a measure on $\mathbb{R}^d$ satisfying*

$$i.\ \nu(\{0\}) = 0 \qquad ii.\ \int_{\mathbb{R}^d} \left(|x|^2 \wedge 1\right) \nu(dx) < \infty$$

*and $\gamma \in \mathbb{R}^d$.*

b. *The representation (1.5) of $\widehat{\mu}$ by $(A, \nu, \gamma)$ is unique.*

**Definition 1.5.** The triplet $(A, \nu, \gamma)$ in Theorem 1.2 is called the *generating triplet* of $\mu$. In particular, $\nu$ is the *Lévy measure* of $\mu$.

We call *generating triplet* of a Lévy process $X = \{X_t\}_t$ the generating triplet of $P_{X_1}$. Considering an additive process $Y = \{Y_t\}_t$, we call *system of generating triplets* of $Y$ the set $\{(A_t, \nu_t, \gamma_t)\}_{t \geq 0}$, where $(A_t, \nu_t, \gamma_t)$ is the generating triplet of $P_{X_t}$ for every $t \geq 0$. The generating triplet of an infinitely divisible distribution on $\mathbb{R}$ will be denoted by $\left(\sigma^2, \nu, \gamma\right)$.

**Example 1.3.** Let $\mu$ be a $NIG$ distribution with parameters $(\alpha, \beta, \mu, \delta)$. Then its generating triplet is given by:

$$\begin{cases} \sigma^2 = 0 \\ \nu(dx) = \dfrac{\delta\alpha}{\pi|x|} e^{\beta x} K_1(\alpha|x|)\, dx \\ \gamma = \mu + \dfrac{2\delta\alpha}{\pi} \displaystyle\int_0^1 \sinh(\beta x) K_1(\alpha x)\, dx \end{cases}.$$

For an explicit calculation, we refer to the paper [4].

Since $\sigma^2 = 0$, a Lévy process $X = \{X_t\}_t$ corresponding to $\mu$, i.e., such that $P_{X_1} = \mu$, will be a *pure jump* process. ///

We are going to prove part (b.) of Theorem 1.2, as it is useful to show the correspondence between the generating triplet of a Lévy process and the more general concept of characteristics of semimartingale. We start by the following

**Lemma 1.1.** *A measure $\nu$ on $\mathbb{R}^d$ satisfying*

$$i.\ \nu(\{0\}) = 0 \qquad ii.\ \int_{\mathbb{R}^d} \left(|x|^2 \wedge 1\right) \nu(dx) < \infty$$

*is $\sigma$–finite.*

*Proof.* For any $n \in \mathbb{N}$, let $B_{1/n} := \left\{x \in \mathbb{R}^d : |x| \leq \frac{1}{n}\right\}$. Thanks to (ii.) it results $\nu\left({B_{1/n}}^c\right) < \infty$. Indeed,

$$\int_{\frac{1}{n}<|x|<1} \nu(dx) = n^2 \int_{\frac{1}{n}<|x|<1} \frac{1}{n^2} \nu(dx) \leq n^2 \int_{\frac{1}{n}<|x|<1} |x|^2 \nu(dx) < \infty,$$

therefore $\nu\left({B_{1/n}}^c\right) = \left\{\int_{\frac{1}{n}<|x|<1} + \int_{|x|\geq 1}\right\} \nu(dx) < \infty$. Hence, using also (i.), we can read $\mathbb{R}^d$ as a countable union of sets of finite measure:

$$\mathbb{R}^d = \{0\} \bigcup \left(\bigcup_{n\in\mathbb{N}} {B_{1/n}}^c\right).$$

This completes the proof. ■

In particular, the Lévy measure of any infinitely divisible distribution is $\sigma$–finite.

*Proof of Theorem 1.2, (b.).* It is known that for any $n \in \mathbb{N}$ there exists some $\theta_n \in \mathbb{C}$, with $|\theta_n| \leq 1$, such that

$$\exp(iu) = \sum_{k=0}^{n-1} \frac{(iu)^k}{k!} + \theta_n \frac{|u|^n}{n!}, \qquad u \in \mathbb{R}. \tag{1.6}$$

Assume that (1.5) holds. We note that for any $z \in \mathbb{R}^d$ it holds

$$\left|e^{i<z,x>} - 1 - i < z, x > 1_D(x)\right| \leq \frac{1}{2} |z|^2 |x|^2 1_D(x) + 2\, 1_{\{|x|>1\}}(x), \quad x \in \mathbb{R}^d,$$

by (1.6) and $\left|e^{i<z,\cdot>} - 1\right| \leq 2$. Since $\int_{\mathbb{R}^d} \left(|x|^2 \wedge 1\right) \nu(dx) < \infty$, by Lebesgue's convergence theorem the characteristic exponent

$$\psi(z) := -\frac{1}{2} < z, Az > +i < \gamma, z > + \int_{\mathbb{R}^d} \left(e^{i<z,x>} - 1 - i < z, x > 1_D(x)\right) \nu(dx), \quad z \in \mathbb{R}^d \tag{1.7}$$

is continuous in $z$. This implies that $\psi$ is uniquely determined by $\widehat{\mu}$. Now we choose $z \in \mathbb{R}^d$ and consider

$$\psi(sz) = -\frac{1}{2} s^2 < z, Az > +is < \gamma, z > + \int_{\mathbb{R}^d} \left(e^{i<sz,x>} - 1 - i < sz, x > 1_D(x)\right) \nu(dx), \quad s \in \mathbb{R};$$

applying again the dominated convergence theorem we infer

$$\frac{1}{s^2}\psi(sz) \to -\frac{1}{2} < z, Az > \quad \text{as} \quad s \to \infty.$$

Hence $A$ is uniquely determined by $\widehat{\mu}$.

Let $\eta(z) := \psi(z) + \frac{1}{2} < z, Az >, z \in \mathbb{R}^d$, and $C := [-1,1]^d$, a $d$–dimensional cube of side 2. For any $z \in \mathbb{R}^d$ we have

$$2^d \int_{\mathbb{R}^d} e^{i<z,x>} \left(1 - \prod_{j=1}^{d} \frac{\sin x_j}{x_j}\right) \nu(dx) = \int_C (\eta(z) - \eta(z+w))\, dw. \tag{1.8}$$

Indeed, if we fix $z \in \mathbb{R}^d$, then from the definition of $\eta$ we get

$$\begin{aligned}\eta(z) - \eta(z+w) &= -i < \gamma, w > \\ &\quad + \int_{\mathbb{R}^d} \left(e^{i<z,x>} - e^{i<z+w,x>} + i < w, x > 1_D(x)\right) \nu(dx)\end{aligned}$$

for any $w \in \mathbb{R}^d$. We further note that, by (1.6), it results

$$\begin{aligned}&\left|e^{i<z,x>} - e^{i<z+w,x>} + i < w, x >\right| = \left|e^{i<z,x>}\left(1 - e^{i<w,x>}\right) + i < w, x >\right| \\ &\le \left|e^{i<z,x>}\left(1 - e^{i<w,x>} + i < w, x >\right)\right| + \left|i < w, x > \left(1 - e^{i<z,x>}\right)\right| \\ &\le \frac{1}{2}|w|^2|x|^2 + |w||z||x|^2, \quad w \in \mathbb{R}^d, |x| \le 1.\end{aligned}$$

It then follows that

$$\begin{aligned}\left|e^{i<z,x>} - e^{i<z+w,x>} + i < w, x > 1_D(x)\right| &\le \left(\frac{1}{2}|w|^2|x|^2 + |w||z||x|^2\right) 1_D(x) \\ &\quad + 2\, 1_{\{|x|>1\}}(x), \quad w \in \mathbb{R}^d,\ x \in \mathbb{R}.\end{aligned}$$

Thus, by Tonelli's theorem the function

$$e^{i<z,x>} - e^{i<z+w,x>} + i < w, x > 1_D(x) \in L^1\left(\mathbb{R}^d \times C\right),$$

where in $\mathbb{R}^d \times C$ we consider the product measure $\nu(dx) \otimes dw$. As $\int_C < \gamma, w > dw = 0$, Fubini's theorem applies and we get

$$\begin{aligned}\int_C (\eta(z) - \eta(z+w))\, dw &= \int_{\mathbb{R}^d} \left[\int_C \left(e^{i<z,x>} - e^{i<z+w,x>}\right.\right. \\ &\qquad \left.\left. + i < w, x > 1_D(x)\right) dw\right] \nu(dx) \\ &= \int_{\mathbb{R}^d} e^{i<z,x>} \left(\int_C \left(1 - e^{i<w,x>}\right) dw\right) \nu(dx) \\ &= 2^d \int_{\mathbb{R}^d} e^{i<z,x>} \left(1 - \prod_{j=1}^{d} \frac{\sin x_j}{x_j}\right) \nu(dx),\end{aligned}$$

where in the last equality we have used that $|C| = 2^d$ and

$$e^{i(w_1 x_1 + \ldots + w_d x_d)} = (\cos w_1 x_1 + i \sin w_1 x_1) \ldots (\cos w_d x_d + i \sin w_d x_d),$$

hence

$$\int_C \left(1-e^{i<w,x>}\right) dw = 2^d - \prod_{j=i}^{d} \int_{[-1,1]} \left(\cos w_j x_j + i \sin w_j x_j\right) dw_j$$
$$= 2^d - \prod_{j=1}^{d} 2\frac{\sin x_j}{x_j} = 2^d \left(1 - \prod_{j=1}^{d} \frac{\sin x_j}{x_j}\right).$$

It has been possible to use Fubini's and Tonelli's theorems thanks to the $\sigma$–finiteness of $\nu$, which is assured by Lemma 1.1. At this point we define

$$\rho\left(dx\right) := 2^d \left(1 - \prod_{j=1}^{d} \frac{\sin x_j}{x_j}\right) \nu\left(dx\right);$$

since $\prod_{j=1}^{d} \frac{\sin x_j}{x_j} = 1 - \frac{|x|^2}{6} + o\left(|x|^3\right)$ for $x \to 0$, then $\rho$ is a finite measure on $\mathbb{R}^d$ with Fourier transform given by

$$\widehat{\rho}\left(z\right) = \int_{\mathbb{R}^d} e^{i<z,x>} \rho\left(dx\right) = \int_C \left(\eta\left(z\right) - \eta\left(z+w\right)\right) dw, \quad z \in \mathbb{R}^d.$$

As in the case of the characteristic functions for distributions, a finite Borel measure is solely indicated by its Fourier transform. Therefore $\widehat{\rho}$ (and $\rho$, as well) is uniquely determined by $\eta$, and so by $\widehat{\mu}$. Recalling that $\nu\left(\{0\}\right) = 0$ and the density of $\rho$ is strictly positive in $\mathbb{R}^d \setminus \{0\}$, we can affirm that $\nu$ depends only on $\widehat{\mu}$.

Since we can obtain $\gamma$ from the expression of $\psi$ as a function of $A$ and $\nu$ by (1.7), the proof is complete. ■

**Remark 1.3.** It is not necessary to take $1_D\left(\cdot\right)$ to have integrability in (1.5). The next examples show two possible alternatives for such function.

**Example 1.4.** Let $c : \mathbb{R}^d \to \mathbb{R}$ be a bounded function such that

$$c\left(x\right) = 1 + o\left(|x|, |x| \to 0\right), \qquad c\left(x\right) = O\left(\tfrac{1}{|x|}, |x| \to \infty\right).$$

In this way for every $z \in \mathbb{R}^d$ we restore

$$\left|e^{i<z,x>} - 1 - i < z, x > c\left(x\right)\right| \leq |z|\, o\left(|x|^2, |x| \to 0\right) + \frac{1}{2}|z|^2 |x|^2, \quad x \in \mathbb{R}^d$$

and

$$\left|e^{i<z,x>} - 1 - i < z, x > c\left(x\right)\right| \leq 2 + |< z, x >|\, O\left(\frac{1}{|x|}, |x| \to \infty\right), \quad x \in \mathbb{R}^d.$$

In this setting, recalling that $c$ is bounded we can define

$$\gamma_c := \gamma + \int_{\mathbb{R}^d} x\left(c\left(x\right) - 1_D\left(x\right)\right) \nu\left(dx\right) \quad \text{componentwise},$$

so that the Lévy–Kintchine formula (1.5) becomes

$$\widehat{\mu}\left(z\right) = \exp\left[-\frac{1}{2} < z, Az > + i < \gamma_c, z > + \int_{\mathbb{R}^d} \left(e^{i<z,x>} - 1 - i < z, x > c\left(x\right)\right) \nu\left(dx\right)\right], \quad z \in \mathbb{R}^d. \tag{1.9}$$

Note that only $\gamma_c$ depends on the choice of $c$: neither $A$ nor $\nu$ does. ///

**Example 1.5.** Let $h : \mathbb{R}^d \to \mathbb{R}^d$ be bounded, with $h(x) = x$ in a neighborhood of 0. Later on, we will call such applications *truncation functions* (see Definition 1.12). It is immediate to show, for any $z \in \mathbb{R}^d$, the following:

$$\left| e^{i<z,x>} - 1 - i < z, h(x) > \right| \leq \frac{1}{2} |z|^2 |x|^2$$

in a neighborhood of 0 and

$$\left| e^{i<z,x>} - 1 - i < z, h(x) > \right| \leq 2 + C |z|, \quad x \in \mathbb{R}^d,$$

for some positive constant $C$ such that $h \leq C$ in $\mathbb{R}^d$. If we put

$$\gamma_h := \gamma + \int_{\mathbb{R}^d} (h(x) - x 1_D(x)) \, \nu(dx) \quad \text{componentwise},$$

then the Lévy–Kintchine formula (1.5) becomes

$$\widehat{\mu}(z) = \exp \Bigg[ -\frac{1}{2} < z, Az > + i < \gamma_h, z > \\ + \int_{\mathbb{R}^d} \left( e^{i<z,x>} - 1 - i < z, h(x) > \right) \nu(dx) \Bigg], \quad z \in \mathbb{R}^d. \quad (1.10)$$

Once again, only $\gamma_h$ depends on the choice of the truncation function. ///

We now want to extend, in some sense, the characteristic function (and exponent) of an infinitely divisible distribution to a subset of $\mathbb{C}^d$.

**Definition 1.6.** Let $g$ be a measurable, nonnegative function on $A \in \mathcal{B}(\mathbb{R}^d)$. The $g$–moment of a measure $\nu$ on $A$ is $\int_A g(x) \, \nu(dx)$.

The $g$–moment of a random variable $X$ on $\mathbb{R}^d$ is $E[g(X)] \left( = \int_{\mathbb{R}^d} g(x) P_X(dx) \right)$.

**Definition 1.7.** A function $g$ on $\mathbb{R}^d$ is said *submultiplicative* if it is nonnegative and there exists a constant $a > 0$ such that

$$g(x + y) \leq a \, g(x) \, g(y), \quad x, y \in \mathbb{R}^d.$$

We largely use these submultiplcative functions:

$$|x| \vee 1, \qquad \exp\left( |x|^\beta \right) \text{ for any } 0 < \beta \leq 1, \qquad \exp(< c, x >) \text{ for some } c \in \mathbb{R}^d.$$

Obviously, the product of two submultiplicative functions is submultiplicative, as well.

A property $\mathfrak{P}$ related to a distribution on $\mathbb{R}^d$ is said a *time independent distributional property* in the class of Lévy processes if, given a Lévy process $X = \{X_t\}_t$, the following equivalence holds:

$$P_{X_t} \text{ has } \mathfrak{P} \text{ for some } t > 0 \iff P_{X_t} \text{ has } \mathfrak{P} \text{ for every } t > 0.$$

The next result is the Theorem 25.3 in [27].

**Theorem 1.3.** *Let $g$ be a measurable, submultiplicative and locally bounded function on $\mathbb{R}^d$. Then the finiteness of the $g$–moment is a* time independent distributional property *in the class of Lévy processes. Moreover, if $X = \{X_t\}_t$ is a Lévy process on $\mathbb{R}^d$ with generating triplet $(A, \nu, \gamma)$, then $X_t$ has finite $g$–moment for every $t > 0$ if and only if $\nu\big|_{|x|>1}$ has finite $g$–moment.*

Thus, $E[g(X_t)] < \infty$ for every $t > 0$ if and only if $\int_{|x|>1} g(x)\,\nu(dx) < \infty$.

Let us take into account a Lévy process $X = \{X_t\}_t$ on $\mathbb{R}^d$ with generating triplet $(A, \nu, \gamma)$ and define for $c \in \mathbb{R}^d$ the submultiplicative function $g_c(x) := \exp(<c, x>)$, $x \in \mathbb{R}^d$, which is continuous, hence locally bounded and measurable. Given $u, w \in \mathbb{C}^d$, we consider

$$< u, w > := \sum_{j=1}^{d} u_j w_j,$$

so it is not the hermitian inner product. We introduce the following set:

$$C := \left\{ c \in \mathbb{R}^d : \int_{|x|>1} e^{<c,x>}\,\nu(dx) < \infty \right\}.$$

Of course $0 \in C$, moreover $C$ is convex. Indeed, letting $c_1, c_2 \in C$ and $t \in (0,1)$, we compute

$$\begin{aligned}\int_{|x|>1} e^{<c_1 t+(1-t)c_2,x>}\,\nu(dx) &= \int_{|x|>1} e^{t<c_1,x>} e^{(1-t)<c_2,x>}\,\nu(dx)\\ &\leq \left(\int_{|x|>1} e^{<c_1,x>}\,\nu(dx)\right)^t \left(\int_{|x|>1} e^{<c_2,x>}\,\nu(dx)\right)^{1-t} < \infty,\end{aligned}$$

where in the last passage we have used Holder's inequality. Theorem 1.3 states that $c \in C \Leftrightarrow E\left[e^{<c,X_t>}\right] < \infty$ for some $t > 0$ (equivalently, for any $t > 0$). At this point we set the function $\Psi : \widetilde{D} \to \mathbb{C}$ as follows:

$$\begin{aligned}\Psi(w) := \frac{1}{2} <w, Aw> + <\gamma, w>\\ + \int_{\mathbb{R}^d} \left(e^{<w,x>} - 1 - <w, x> 1_D(x)\right) \nu(dx), \quad w \in \widetilde{D}, \quad (1.11)\end{aligned}$$

where $\widetilde{D} := \left\{ w \in \mathfrak{C}^d : \mathfrak{Re}(w) \in C \right\}$. We immediately note that $iz \in \widetilde{D}$ for any $z \in \mathbb{R}^d$. In order to show that the integral in (1.11) is well posed, we fix $w \in \widetilde{D}$ and start off by considering $x \in D$; by (1.6) we get

$$\begin{aligned}&\left|e^{<w,x>} - 1 - < w, x > 1_D(x)\right| = \left|e^{<\mathfrak{Re}(w),x>} e^{i<\mathfrak{Im}(w),x>} - 1 - < w, x >\right|\\ &= \Bigg| \left(1 + < \mathfrak{Re}(w), x > + \frac{1}{2} < \mathfrak{Re}(w), x >^2 + O\left(|x|^2, |x| \to 0\right)\right)\\ &\qquad \left(1 + i < \mathfrak{Im}(w), x > + \frac{1}{2}\theta < \mathfrak{Im}(w), x >^2\right) - 1 - < w, x > \Bigg|,\end{aligned}$$

where $\theta \in \mathbb{C}$ is provided. So $|e^{<w,x>} - 1 - < w, x > 1_D(x)| = O\left(|x|^2, |x| \to 0\right)$. As far as the case $|x| > 1$ is concerned, recalling that $\mathfrak{Re}(w) \in C$ we immediately have

$$\int_{|x|>1} \left|e^{<w,x>}\right| \nu(dx) = \int_{|x|>1} e^{<\mathfrak{Re}(w),x>}\,\nu(dx) < \infty.$$

Therefore the function $\Psi$ is actually well defined in $\widetilde{D}$ and we also have

$$E\left[\left|e^{<w,X_t>}\right|\right] = E\left[e^{<\mathfrak{Re}(w),X_t>}\right] < \infty \quad \text{for any} \quad w \in \widetilde{D},\ t \geq 0.$$

Theorem 25.17 in [27] affirms that

$$E\left[e^{<w,X_t>}\right] = e^{t\Psi(w)} \quad \text{for any } \; w \in \widetilde{D},\, t \geq 0. \tag{1.12}$$

Henceforth, we call $\Psi$ the *cumulant function* of the Lévy process $X$. In particular, from the definition (1.11) of $\Psi$ we notice that

$$\Psi\left(iz\right) = \psi\left(z\right), \quad z \in \mathbb{R}^d.$$

In this sense $\Psi$ extends $\psi$ to a subset of $\mathbb{C}^d$.

**Remark 1.4.** With a similar reasoning to that of Remark 1.3, we can substitute the function $1_D\left(\cdot\right)$ in (1.11). For example, if we take a truncation function $h$ (see Definition 1.12), then $\Psi = \Psi_h$ in $\widetilde{D}$, where

$$\Psi_h\left(w\right) := \frac{1}{2} < w, Aw > + < \gamma_h, w > \\ + \int_{\mathbb{R}^d} \left(e^{<w,x>} - 1 - < w, h\left(x\right) >\right) \nu\left(dx\right), \quad w \in \widetilde{D}$$

and $\gamma_h$ is defined as in Example 1.5.

The following example will be crucial in developing the *Esscher measure*.

**Example 1.6.** Let $X = \{X_t\}_t$ be a $\mathbb{R}$–valued Lévy process defined on a probability space $(\Omega, \mathcal{F}, P)$, $\Psi$ be its cumulant function and $\theta \in \mathbb{R}$ be such that $E\left[\exp\left(\theta X_t\right)\right] < \infty$ for some $t > 0$ (equivalently, for any $t > 0$). Then we consider the process $M = \{M_t\}_t$, where

$$M_t := e^{\theta X_t - t\Psi(\theta)}, \quad t \geq 0.$$

We note that $M$ is integrable, by assumption and the fact that $t\Psi\left(\theta\right)$ is constant in $\Omega$ for every $t \geq 0$. Let us now construct the minimal augmented filtration $\mathbb{F} = \left(\mathcal{F}_t\right)_{t \geq 0}$ of $X$, i.e., $\mathcal{F}_t = \sigma\left(\mathcal{N} \bigcup \mathcal{F}_t^0\right)$ for any $t \geq 0$, where $\left(\mathcal{F}_t^0\right)_t$ is the natural filtration of the process and $\mathcal{N}$ is the collection of $\mathcal{F}$–negligible sets; obviously, $M$ is $\mathbb{F}$–adapted. If we fix $t > s \geq 0$, then using (1.12) and the properties of the increments we readily obtain

$$E\left[M_t | \mathcal{F}_s\right] = E\left[e^{\theta X_t - t\Psi(\theta)} \middle| \mathcal{F}_s\right] \overset{a.s.}{=} e^{\theta X_s - s\Psi(\theta)} E\left[e^{\theta(X_t - X_s)} \middle| \mathcal{F}_s\right] e^{-(t-s)\Psi(\theta)} \\ \overset{a.s.}{=} M_s\, E\left[e^{\theta X_{t-s}}\right] e^{-(t-s)\Psi(\theta)} = M_s.$$

Thus, $M$ is a martingale with mean 1.

The extension of this property to the $d$–dimensional case is straightforward. Indeed, if we consider $\theta \in C$ ($\Rightarrow E\left[e^{<\theta,X_t>}\right] < \infty$ for every $t > 0$), then the process $M = \{M_t\}_t$ defined by

$$M_t := e^{<\theta,X_t> - t\Psi(\theta)}, \quad t \geq 0$$

is a martingale with expectation 1:

$$E\left[M_t | \mathcal{F}_s\right] = E\left[e^{<\theta,X_t> - t\Psi(\theta)} \middle| \mathcal{F}_s\right] \\ \overset{a.s.}{=} e^{<\theta,X_s> - s\Psi(\theta)} E\left[e^{<\theta,X_t - X_s>} \middle| \mathcal{F}_s\right] e^{-(t-s)\Psi(\theta)} \\ \overset{a.s.}{=} M_s\, E\left[e^{<\theta,X_{t-s}>}\right] e^{-(t-s)\Psi(\theta)} = M_s.$$

**Definition 1.8.** Fix $T > 0$ and let $\theta \in C$. The probability measure $P^\theta$ on $\mathcal{F}_T$, with $P^\theta \sim P$ on $\mathcal{F}_T$, defined by $\frac{dP^\theta}{dP} := M_T$ is called the *Esscher transform* of $P$ with respect to $\theta$.

The density process is given by $\dfrac{dP^\theta\big|_{\mathcal{F}_t}}{dP\big|_{\mathcal{F}_t}} = M_t$, $t \in [0, T]$. ///

## 1.3 Lévy–Itô Decomposition

In this section we introduce and discuss one of the most important theorem in the theory of Lévy processes: the *Lévy–Itô decomposition of sample paths*, in the formulation suggested by K.–I. Sato in [27]. We start off by presenting the next

**Definition 1.9.** Let $(\Theta, \mathcal{B}, \rho)$ be a $\sigma$–finite measure space. A family of $\overline{\mathbb{N}}$–valued random variables $\{N(B), B \in \mathcal{B}\}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ is said a *Poisson random measure with intensity* $\rho$ if the following properties hold:

i. for every $B \in \mathcal{B}, N(B)$ is a Poisson random variable with mean $\rho(B)$;

ii. if $B_1, ..., B_n \in \mathcal{B}$ are disjoint, then $N(B_1), ..., N(B_n)$ are independent;

iii. for any $\omega \in \Omega$, $N(\cdot, \omega)$ is a measure on $\Theta$.

Moreover, we denote by $D_{a,b} := \left\{x \in \mathbb{R}^d : a < |x| \leq b\right\}$ and $H := (0, \infty) \times D_{0,\infty} = (0, \infty) \times \left(\mathbb{R}^d \setminus \{0\}\right)$, which we endow with the product $\sigma$–algebra $\mathcal{H} := \mathcal{B}(0, \infty) \otimes \mathcal{B}(D_{0,\infty})$.

**Theorem 1.4** ([27], Theorem 19.2)**.** *Let $X = \{X_t\}_t$ be an additive process on $\mathbb{R}^d$ defined on a probability space $(\Omega, \mathcal{F}, P)$ with system of generating triplets $\{(A_t, \nu_t, \gamma_t)\}_t$. Define the measure $\widetilde{\nu}$ on $(H, \mathcal{H})$ by*

$$\widetilde{\nu}\left((0, t] \times \widetilde{B}\right) := \nu_t\left(\widetilde{B}\right), \quad \text{for any } \widetilde{B} \in \mathcal{B}(D_{0,\infty}), t > 0.$$

*For every $B \in \mathcal{H}$ we set*

$$J(B, \omega) := \#\left\{s \in \mathbb{R}^+ : (s, X_s(\omega) - X_{s-}(\omega)) \in B\right\}, \quad \omega \in \Omega.$$

*Then:*

*a. $\{J(B), B \in \mathcal{H}\}$ is a Poisson random measure on $H$ with intensity $\widetilde{\nu}$;*

*b. there exists $\Omega_1 \in \mathcal{F}$, with $P(\Omega_1) = 1$, such that for any $\omega \in \Omega_1$*

$$X_t^1(\omega) := \lim_{\epsilon \to 0^+} \int_{(0,t] \times D_{\epsilon,1}} (x\, J(d(s, x), \omega) - x\widetilde{\nu}(d(s, x))) + \int_{(0,t] \times D_{1,\infty}} x\, J(d(s, x), \omega)$$

*is defined for all $t \in [0, \infty)$ and the convergence is uniform in $t$ on any bounded interval. The process $X^1 = \left\{X_t^1\right\}_t$ is additive on $\mathbb{R}^d$ with system of generating triplets $\{(0, \nu_t, 0)\}_t$;*

c. *define* $X^2 := X - X^1$ *in* $\Omega_1$. *There exists* $\Omega_2 \in \mathcal{F}$, *with* $P(\Omega_2) = 1$, *such that* $X_t^2(\omega)$ *is continuous in* $t$ *for any* $\omega \in \Omega_2$. *The process* $X^2 = \left\{X_t^2\right\}_t$ *is additive on* $\mathbb{R}^d$ *with system of generating triplets* $\{(A_t, 0, \gamma_t)\}_t$;

d. *the processes* $X^1$ *and* $X^2$ *are independent.*

We call $\lim_{\epsilon\to 0^+}\int_{(0,t]\times D_{\epsilon,1}}\left(x\,J\left(d\left(s,x\right),\omega\right)-x\widetilde{\nu}\left(d\left(s,x\right)\right)\right)$ the *compensated sum of jumps* of the process $X$.

We now discuss the previous theorem as far as a Lévy process $X = \{X_t\}_t$ with generating triplet $(A, \nu, \gamma)$ is concerned. By Theorem 1.1 its system of generating triplets is $\{(tA, t\nu, t\gamma)\}_t$ and the measure $\widetilde{\nu}$ simplifies to $\widetilde{\nu} = dt|_{(0,\infty)} \otimes \nu|_{\mathbb{R}^d\setminus\{0\}}$. This implies that $X^i = \left\{X_t^i\right\}_t$, for $i = 1, 2$, are Lévy processes, as well. In fact, considering a generic additive process $Y = \{Y_t\}_t$ with system of generating triplets $\{(A_t, \nu_t, \gamma_t)\}_t = \{(tA, t\nu, t\gamma)\}_t$, for any choice $0 \le s < t$ we have that $Y_t - Y_s$ is independent from $Y_s$ , hence

$$\widehat{P_{Y_t}} = \widehat{P_{Y_t - Y_s + Y_s}} = \widehat{P_{Y_t - Y_s}}\,\widehat{P_{Y_s}}.$$

Since $\widehat{P_{Y_s}}(z) \neq 0$ for every $z \in \mathbb{R}^d$, by (1.5) we can easily get

$$\begin{aligned}\widehat{P_{Y_t-Y_s}}\left(z\right)=\frac{\widehat{P_{Y_t}}\left(z\right)}{\widehat{P_{Y_s}}\left(z\right)} &= \exp\Bigg[-\frac{1}{2}\left(t-s\right)<z,Az>+i\left(t-s\right)<\gamma,z> \\ &\quad +\left(t-s\right)\int_{\mathbb{R}^d}\left(e^{i<z,x>}-1-i<x,z>1_D\left(x\right)\right)\nu\left(dx\right)\Bigg] \\ &= \widehat{P_{Y_{t-s}}}\left(z\right), \quad z\in\mathbb{R}^d.\end{aligned}$$

Thus, $X$ can be seen as the sum of two independent Lévy processes, at least in $\Omega_2$. In particular, $X_t^2$ has the same distribution as the homologous random variable of the sum between a constant drift and a Wiener process, that is,

$$X_t^2 \overset{d}{=} \gamma t + \sqrt{A}W_t, \quad t \ge 0,$$

where $W = \{W_t\}_t$ is a Brownian motion and $\sqrt{A}$ is the unique, positive semidefinite square root of the matrix $A$. Indeed, it is certainly true that $W_t \overset{d}{=} \sqrt{t}Z$ for every $t \ge 0$, where $Z \sim \mathcal{N}(0, Id)$. Hence defining $Y_t := \gamma t + \sqrt{A}W_t$ for $t \ge 0$ we get

$$Y_t \overset{d}{=} \gamma t + \sqrt{t}\sqrt{A}Z \sim \mathcal{N}(t\gamma, tA), \quad t > 0$$

so its characteristic function is well known and given by the following:

$$E\left[e^{i<z,Y_t>}\right] = \exp\left(it<\gamma,z> -\frac{1}{2}t<z,Az>\right), \quad t>0,\ z\in\mathbb{R}^d.$$

This implies $Y_t \overset{d}{=} X_t^2$ for every $t \ge 0$, as requested. Actually we can state something more: since $Y = \{Y_t\}_t$ is a Lévy process with generating triplet $(A, 0, \gamma)$, the same as $X^2$, according to Theorem 1.1 the processes $X^2$ and $Y$ are identical in law.

It is slightly more complicated to analyze the process $X^1$, specifically the term $N_t := \int_{(0,t]\times D_{1,\infty}} x\,J\left(d\left(s,x\right),\omega\right),\ t \ge 0$. Our final goal is to prove that $N = \{N_t\}_t$ is a compound Poisson process. We need the next result.

**Proposition 1.1** ([27], Proposition 19.5). *Let $(\Theta, \mathcal{B}, \rho)$ be a measure space with $\rho(\theta) < \infty$ and $\{N(B), B \in \mathcal{B}\}$ be a Poisson random measure with intensity $\rho$. Let $\phi : \Theta \to \mathbb{R}^d$ be a measurable function, and define*

$$Y := \int_\Theta \phi(\theta)\; N(d\theta) \quad \textit{componentwise.}$$

*Then*

$$E\left[e^{i<z,Y>}\right] = \exp\left[\int_\Theta \left(e^{i<z,\phi(\theta)>} - 1\right)\, \rho(d\theta)\right], \quad z \in \mathbb{R}^d. \tag{1.13}$$

In the same notation as the proposition, we define on $\Theta$ the probability measure

$$P(d\theta) := \frac{\rho(d\theta)}{\rho(\Theta)},$$

obviously assuming that $\rho(\Theta) > 0$. Denoting by $P\phi^{-1}(dx)$ the pushforward distribution on $\mathbb{R}^d$ (i.e., $P\phi^{-1}(B) := P\left(\phi^{-1}(B)\right)$ for any $B \in \mathcal{B}\left(\mathbb{R}^d\right)$), we can rewrite (1.13) as

$$\begin{aligned} E\left[e^{i<z,Y>}\right] &= \exp\left[\rho(\Theta)\left(\int_\Theta e^{i<z,\phi(\theta)>}\, P(d\theta) - 1\right)\right] \\ &= \exp\left[\rho(\Theta)\left(\int_{\mathbb{R}^d} e^{i<z,x>}\, P\phi^{-1}(dx) - 1\right)\right], \quad z \in \mathbb{R}^d. \end{aligned}$$

So if we suppose that $P\phi^{-1}(\{0\}) = 0$, we see that $Y$ has a compound Poisson distribution with constant $\rho(\Theta)$ and distribution $P\phi^{-1}(dx)$.

Let us consider again the process $N$ and fix $t > 0$. Taking the space $H_{t,1} := (0,t] \times D_{1,\infty}$ with measure $\left.\widetilde{\nu}\right|_{H_{t,1}}$, it results

$$\left.\widetilde{\nu}\right|_{H_{t,1}}(H_{t,1}) = t\nu(D_{1,\infty}) < \infty.$$

So, by the argument after the proposition above, if $\nu(D_{1,\infty}) > 0$ we arrive at

$$E\left[e^{i<z,N_t>}\right] = \exp\left[t\nu(D_{1,\infty})\left(\int_{\mathbb{R}^d} e^{i<z,x>}\left(\frac{\widetilde{\nu}}{t\nu(D_{1,\infty})}\right)\phi^{-1}(dx) - 1\right)\right], \quad z \in \mathbb{R}^d$$

with $\left(\frac{\widetilde{\nu}}{t\nu(D_{1,\infty})}\right)\phi^{-1}(dx)$ which can be expressed as

$$\left(\frac{\widetilde{\nu}}{t\nu(D_{1,\infty})}\right)\phi^{-1}(B) = \begin{cases} 0, & \text{if } B \subset D \\ \frac{\nu(B \cap D_{1,\infty})}{\nu(D_{1,\infty})}, & \text{otherwise} \end{cases}, \quad B \in \mathcal{B}\left(\mathbb{R}^d\right),$$

that is,

$$\left(\frac{\widetilde{\nu}}{t\nu(D_{1,\infty})}\right)\phi^{-1}(B) = \frac{\nu(B \cap D_{1,\infty})}{\nu(D_{1,\infty})}, \quad B \in \mathcal{B}\left(\mathbb{R}^d\right). \tag{1.14}$$

Thus, we can say that $N$ is a compound Poisson process with constant $\nu(D_{1,\infty})$ and distribution $\left(\frac{\widetilde{\nu}}{t\nu(D_{1,\infty})}\right)\phi^{-1}(dx)$.

Finally, we give a classical interpretation of the Lévy measure. For every $t > 0$, $B \in \mathcal{B}\left(\mathbb{R}^d \setminus \{0\}\right)$, it results that $J((0,t] \times B)$ is a Poisson random variable with mean $t\nu(B)$ and counts the number of jumps before time $t$ which lies in the set $B$. Taking $t = 1$, it follows $\nu(B) = E[J((0,1] \times B)]$, so the Lévy measure of a set $B$ can be thought of as the expected number of jumps of the Lévy process in $B$ in the unitary time interval.

We conclude this section with a remark which paves the way for the generalizations of the next section.

**Remark 1.5.** So far, we have developed the theory of Lévy processes in a simple probability space $(\Omega, \mathcal{F}, P)$, without a filtration. However, we can endow such $\Omega$ with a "nice" filtration in a natural way. Let us consider $X = \{X_t\}_t$ a Lévy process on $\Omega$; we denote by $\mathbb{F} = (\mathcal{F}_t)_t$ the augmented filtration generated by $X$, that is, $\mathcal{F}_t = \sigma\left(\mathcal{F}_t^0 \bigcup \mathcal{N}\right)$ for every $t \geq 0$, where $\left(\mathcal{F}_t^0\right)_t$ is the natural filtration of the process and $\mathcal{N}$ the collection of $\mathcal{F}$–negligible sets. According to [24], Theorem 31, $\mathbb{F}$ is right–continuous, so it satisfies the usual hypothesis, as well. Thinking of the Lévy process $X$ on the stochastic basis $(\Omega, \mathcal{F}, P; \mathbb{F})$, it obviously results to be adapted and the condition (i.) in Definition 1.1 can be substituted by the following one:

i'. $X_t - X_s$ is independent from $\mathcal{F}_s$ for every $0 \leq s < t$.

## 1.4 Characteristics Of Semimartingales

The aim of this section is to generalize the concept of generating triplet of a Lévy process to semimartingales. Here we mainly follow [28], Chapter II.

We start off by fixing a stochastic basis $(\Omega, \mathcal{F}, P; \mathbb{F})$, with $\mathbb{F}$ which satisfies the usual hypothesis (in this case it is not a big assumption, since we are going to work with the augmented filtration of a Lévy process, which fulfills this request according to Remark 1.5). Given two stopping times $S, T$, we denote by $[\![S, T]\!]$ the stochastic interval, i.e., the random set

$$[\![S, T]\!] \coloneqq \left\{(t, \omega) \in \mathbb{R}_0^+ \times \Omega : S(\omega) \leq t \leq T(\omega)\right\}.$$

Similarly, we can define the other three types of stochastic interval. In order to keep notation simple, we denote by $[\![T]\!] \coloneqq [\![T, T]\!]$. Recall that the optional $\sigma$–algebra $\mathcal{O}$ is the $\sigma$–algebra on $\mathbb{R}_0^+ \times \Omega$ generated by all the càdlàg, adapted processes, while the predictable $\sigma$–algebra $\mathcal{P}$ is the one generated by all càg, adapted processes.

**Definition 1.10.** A random set $A$ is called *thin* if it is of the form

$$A = \bigcup_n [\![T_n]\!],$$

where $(T_n)_n$ is a sequence of stopping times.

It is important to observe that the sections $\left\{t \in \mathbb{R}_0^+ : (t, \omega) \in A\right\}$, for $\omega \in \Omega$, are at most countable when $A$ is a thin set. We have the following, preliminary result.

**Proposition 1.2.** *If $X = \{X_t\}_t$ is a càdlàg, adapted process, then the random set $\{\Delta X \neq 0\}$ is thin.*

*Proof.* Let $n \in \mathbb{N}$, put $T_0^n := 0$ and define iteratively

$$T_{p+1}^n := \inf\left\{t > T_p^n : \left|X_t - X_{T_n^p}\right| > \frac{1}{2^n}\right\}, \quad p \in \mathbb{N}.$$

In order to prove that these objects are $\mathbb{F}^+$–stopping times (and then $\mathbb{F}$–stopping times to, as $\mathbb{F}$ is assumed to be right–continuous), we present an inductive argument. Of course, $T_0^n$ is a $\mathbb{F}^+$–stopping time; then we suppose that also $T_p^n$ is so for $p \in \mathbb{N}$ as inductive hypothesis. For every $t \in \mathbb{R}_0^+$ we denote by $Y_t$ the random variable

$$Y_t(\omega) := \left(X_t - X_{T_n^p(\omega)}\right)(\omega)\, 1_{[T_n^p(\omega),\infty)}(t), \quad \omega \in \Omega,$$

so $Y_t = \left(X_t - X_{T_n^p \wedge t}\right) 1_{\{T_n^p \leq t\}}$. Since $X$ is progressive, by adaptedness and right–continuity, and $\{T_n^p \leq t\} \in \mathcal{F}_{t+} = \mathcal{F}_t$, $t \geq 0$, by assumption, $Y = \{Y_t\}_t$ is an adapted process which inherits from $X$ the càdlàg property. Now we can express

$$T_{p+1}^n = \inf\left\{t > 0 : |Y_t| > \frac{1}{2^n}\right\},$$

hence by the right–continuity of $Y$ and the fact that $\left\{x \in \mathbb{R}^d : |x| > \frac{1}{2^n}\right\}$ is open we have

$$\{T_{p+1}^n < t\} = \bigcup_{s \in \mathbb{Q}, s<t} \left\{|Y_s| > \frac{1}{2^n}\right\} \in \mathcal{F}_t, \quad t > 0.$$

Therefore $T_{p+1}^n$ is a $\mathbb{F}^+$–stopping time, so a $\mathbb{F}$–stopping time, too.

Since $X$ is adapted and has the càdlàg property, the process $\Delta X$ is optional. It follows that

$$A_p^n := \left\{\Delta X_{T_p^n} \neq 0, T_p^n < \infty\right\} \in \mathcal{F}_{T_p^n}, \quad (n,p) \in \mathbb{N} \times \mathbb{N},$$

so we can introduce other $\mathbb{F}$–stopping times

$$T_p^{n\prime} := \begin{cases} T_p^n, & \text{in } A_p^n \\ \infty, & \text{otherwise} \end{cases}, \quad (n,p) \in \mathbb{N} \times \mathbb{N}.$$

We complete the proof showing the random sets equality

$$\{\Delta X \neq 0\} = \bigcup_{n,p} [\![T_p^{n\prime}]\!].$$

The inclusion "$\supset$" is the easiest one. Indeed, let $(t,\omega) \in \bigcup_{n,p} [\![T_p^{n\prime}]\!]$. Then there exists a couple $(n,p) \in \mathbb{N} \times \mathbb{N}$ such that $(t,\omega) \in [\![T_p^{n\prime}]\!]$, so $t = T_p^{n\prime}(\omega) < \infty$ implies $\omega \in A_p^n$ and $t = T_p^n(\omega)$, whence $\Delta X_t(\omega) \neq 0$.

In order to prove also the other inclusion "$\subset$", let $(t,\omega) \in \{\Delta X \neq 0\}$. Since $X$ is càdlàg, then $T_p^n \nearrow \infty$ as $p \to \infty$ pointwise, for any $n \in \mathbb{N}$; moreover $T_1^n > 0$. As $\Delta X_t(\omega) \neq 0$, then there exists $n \in \mathbb{N}$ such that

$$|X_t(\omega) - X_{t-}(\omega)| > \frac{1}{2^n}.$$

This implies that $T_1^{(n+1)}(\omega) \leq t$. Indeed, assuming by contradiction that $T_1^{(n+1)}(\omega) > t$, then

$$|X_t(\omega) - X_{t-}(\omega)| \leq |X_t(\omega) - X_0(\omega)| + |X_{t-}(\omega) - X_0(\omega)| \leq \frac{1}{2^n},$$

which is evidently an absurdity. At this point, if $T_1^{(n+1)}(\omega) = t$, then $(t, \omega) \in [\![T_1^{(n+1)'}]\!]$ and we are done; if instead $T_1^{(n+1)}(\omega) < t$, using that $T_p^{(n+1)} \nearrow \infty$ as $p \to \infty$ with an analogous argument we can state again the existence of a $p \in \mathbb{N}$ such that $T_p^{(n+1)}(\omega) = t$, hence $(t, \omega) \in [\![T_p^{(n+1)'}]\!]$. ■

**Definition 1.11.** Let $X = \{X_t\}_t$ be an adapted, càdlàg, $\mathbb{R}^d$–valued process. The measure $\mu^X$ on $\mathbb{R}_0^+ \times \mathbb{R}^d$ defined by

$$\mu^X(\omega; dt, dx) = \sum_s 1_{\{x \neq 0\}}(\Delta X_s(\omega))\, \delta_{(s, \Delta X_s(\omega))}(dt, dx), \quad \omega \in \Omega, \tag{1.15}$$

where $\delta_{(\overline{a})}$ denotes the Dirac measure at a point $\overline{a}$, is called *measure associated to its jumps*.

Note that this definition makes sense thanks to the previous proposition: the set $\{\Delta X \neq 0\}$ is thin, so its sections are at most countable and we know how to interpret the sum in (1.15). The measure owes its name to the following fact: if we fix $t \in \mathbb{R}^+$ and $B \in \mathcal{B}(\mathbb{R}^d)$, then

$$\mu^X(\omega; (0, t] \times B) = \sum_{s \leq t} 1_{\{x \neq 0\} \cap B}(\Delta X_s(\omega)), \quad \omega \in \Omega,$$

so it counts the number of jumps of $X$ which lie in $B$ before time $t$.

For a thorough, and yet rather linear, introduction to general random measure we refer to [28], in particular Chapter II, §1a, 1b and we try to stick to the notation therein. It is possible to prove that $\mu^X$ is an integer–valued random measure on $\mathbb{R}_0^+ \times \mathbb{R}^d$ ([28], Proposition II, 1.16).

We now consider a process $X = \{X_t\}_t$ which is a $d$–dimensional, $\mathbb{F}$–adapted semimartingale, highlighting that in this work every semimartingale is càdlàg.

**Definition 1.12.** A bounded function $h : \mathbb{R}^d \to \mathbb{R}^d$ such that $h(x) = x$ in a neighborhood of 0 is called *truncation function*.

We denote by $C_t^d$ the set of all truncation functions.

**Notation.** The set of the $\mathbb{R}$–valued, càdlàg, adapted, nondecreasing (resp. finite variation) processes starting at 0 is denoted by $\mathcal{V}^+$ (resp. $\mathcal{V}$). For a process $A \in \mathcal{V}^+$, it is well–defined $A_\infty := \lim_{t \to \infty} A_t$: the set of $A \in \mathcal{V}^+$ such that $A_\infty$ is integrable is denoted by $\mathcal{A}^+$. The space of the càdlàg, local martingales starting at 0 is indicated by $\mathcal{L}$. The symbol $\cdot$ denotes the integration, in a sense which is clear from the contest time by time.

Let $h \in C_t^d$, then $\Delta X - h(\Delta X) \neq 0$ only if $|\Delta X| > b$ for some $b > 0$. Define

$$\begin{cases} \widetilde{X}(h)_t := \sum_{s \leq t} [\Delta X_s - h(\Delta X_s)] \\ X(h)_t := X_t - \widetilde{X}(h)_t \end{cases}, \quad t \geq 0. \tag{1.16}$$

We notice $\widetilde{X}(h) = \left\{\widetilde{X}(h)_t\right\}_t$ is a $d$–dimensional process in $\mathcal{V}^d$ (its components are in $\mathcal{V}$), while $X(h) = \{X(h)_t\}_t$ is a semimartingale with $\Delta X(h) = h(\Delta X)$, which is bounded. Therefore $X(h)$ is a special semimartingale and its canonical decomposition will be

$$X(h) = X_0 + M(h) + B(h), \tag{1.17}$$

where $B(h)$ is a predictable process in $\mathcal{V}^d$ and $M(h) \in \mathcal{L}^d$ (its components are in $\mathcal{L}$).

**Definition 1.13.** Fix $h \in C_t^d$. The *characteristics* of $X$ with respect to $h$ are the triplet $\left(B, C, \nu^X\right)$, where:

i. $B = \left(B^i\right)_{i \leq d}$ is a predictable process in $\mathcal{V}^d$, namely $B = B(h)$ in (1.17);

ii. $C = \left(C^{ij}\right)_{i,j \leq d}$ is a continuous process in $\mathcal{V}^{d \times d}$, namely $C^{ij} = < X^{i,c}, X^{j,c} >$, where $X^{i,c}$ is the continuous martingale part of $X^i$;

iii. $\nu^X$ is a predictable random measure on $\mathbb{R}_0^+ \times \mathbb{R}^d$, namely the compensator of $\mu^X$.

It is clear that $C$ and $\nu^X$ do not depend on the choice of $h$ while $B$ does. Furthermore, from the definition it follows that the characteristics are unique up to a $P$–null set. This allows for a "good version" of them, according to the next result.

**Proposition 1.3** ([28], Proposition II, 2.9)**.** *There exists a version of the characteristics* $\left(B, C, \nu^X\right)$ *of* $X$ *of the form*

$$\begin{cases} B^i = b^i \cdot A, & i = 1, ..., d \\ C^{ij} = c^{ij} \cdot A, & i, j = 1, ..., d \;, \\ \nu^X(\omega; dt, dx) = dA_t(\omega) F_{(t,\omega)}(dx) & P\text{–}a.s. \end{cases}$$

*where* $\nu^X\left(\{t\} \times \mathbb{R}^d\right) \leq 1$ *identically and*

*a.* $A$ *is a predictable process in* $\mathcal{A}_{loc}^+$*;*

*b.* $b = \left(b^i\right)_{i \leq d}$ *is a d–dimensional predictable process;*

*c.* $c = \left(c^{ij}\right)$ *is a predictable process with values in the set of all symmetric, positive semidefinite,* $d \times d$ *matrices;*

*d.* $F_{(t,\omega)}(dx)$ *is a transition kernel from* $\left(\mathbb{R}_0^+ \times \Omega, \mathcal{P}\right)$ *into* $\left(\mathbb{R}^d, \mathcal{B}\left(\mathbb{R}^d\right)\right)$ *satisfying, among others, the following:*

$$i.\ F_{(t,\omega)}(\{0\}) = 0 \qquad ii.\ \int_{\mathbb{R}^d} \left(|x|^2 \wedge 1\right) F_{(t,\omega)}(dx) \leq 1$$

*for every* $t \in \mathbb{R}_0^+$, $\omega \in \Omega$.

We are now ready to show the relation between characteristics of semimartingale and generating triplet of a Lévy process.

**Definition 1.14.** A process $X = \{X_t\}_t$ is a *PII* process if it is càdlàg, adapted, starts at 0 and has independent increments, that is, $X_t - X_s$ is independent from the $\sigma$–algebra $\mathcal{F}_s$ for every $0 \leq s < t$. Further, if it also has stationary increments, then it is a *PIIS* process.

Of course, a Lévy process is a *PIIS* process (see Remark 1.5). Given $X = \{X_t\}_t$ a *PII* process, for every $z \in \mathbb{R}^d$ we introduce the function

$$g(z)_t := E\left[e^{i<z, X_t>}\right], \quad t \geq 0.$$

It is possible to prove ([28], Theorem II, 4.14) that $X$ is a semimartingale if and only if the functions $t \mapsto g(z)_t$ have finite variation over finite intervals for any $z \in \mathbb{R}^d$.

**Lemma 1.2.** *Every $PIIS$ process is a semimartingale.*

*Proof.* Let $X = \{X_t\}_t$ be a $PIIS$ process and fix $z \in \mathbb{R}^d$. Then

$$\begin{aligned} g(z)_{t+s} &= E\left[e^{i<z,X_{t+s}>}\right] = E\left[e^{i<z,X_{t+s}-X_s>}\right] E\left[e^{i<z,X_s>}\right] \\ &= E\left[e^{i<z,X_t>}\right] E\left[e^{i<z,X_s>}\right] = g(z)_s \, g(z)_t\,, \quad s,t \in \mathbb{R}_0^+, \end{aligned}$$

where in the second equality we have used the independence of increments and in the third their temporal homogeneity. Since $g(z)_0 = 1$ and $t \mapsto g(z)_t$ is right–continuous, then there exists $\lambda_z \in \mathbb{C}$ such that $g(z)_\cdot = \exp(\lambda_z \cdot)$, which has locally finite variation. This completes the proof. ■

In particular, every Lévy process is a semimartingale, so it makes sense to study its characteristics and their relation with the generating triplet. The next theorem provides us with the link we are searching for.

**Theorem 1.5** ([28], Corollary II, 4.19)**.** *A $d$–dimensional process $X = \{X_t\}_t$ is a $PIIS$ if and only if it is a semimartingale admitting a version $\left(B, C, \nu^X\right)$ of its characteristics of the form*

$$\begin{cases} B_t(\omega) = bt \\ C_t(\omega) = ct \\ \nu^X(\omega; dt, dx) = dt\, F(dx) \end{cases}, \quad t \geq 0,\, \omega \in \Omega,$$

*where $b \in \mathbb{R}^d$, $c$ is a symmetric, positive semidefinite $d \times d$ matrix, and $F(dx)$ is a measure on $\mathbb{R}^d$ such that*

$$i.\ F(\{0\}) = 0; \qquad ii.\ \int_{\mathbb{R}^d} \left(|x|^2 \wedge 1\right) F(dx) < \infty\ .$$

*Moreover, for any $t \in \mathbb{R}_0^+$ and $z \in \mathbb{R}^d$, the Lévy–Kintchine formula holds:*

$$E\left[e^{i<z,X_t>}\right] = \exp\left[t\left(-\frac{1}{2} < z, cz > +i < b, z > + \int_{\mathbb{R}^d} \left(e^{i<z,x>} - 1 - i < z, h(x) >\right) F(dx)\right)\right]. \quad (1.18)$$

We note that a $PIIS$ process has a version of its characteristic which is deterministic, with $\nu^X(dt, dx) = dt \otimes F(dx)$.

Finally, we take into account a Lévy process $L = \{L_t\}_t$ with generating triplet $(A, \nu, \gamma_h)$ with respect to a truncation function $h$ (see Example 1.5). Considering $t = 1$ in (1.18), by uniqueness of the Lévy–Kintchine representation we get:

$$\text{a. } b = \gamma_h \qquad \text{b. } c = A \qquad \text{c. } F(dx) = \nu(dx)\ .$$

Hence, the characteristics of $L$ relative to $h$ are given by

$$(\gamma_h t, At, dt \otimes \nu(dx))\,.$$

# Chapter 2

# Esscher Measure

In this chapter we first present the theory necessary to construct the *Esscher measure* for both exponential and linear processes. Later, we specify the dynamics of the stock prices and discuss sufficient conditions for the existence of the Esscher measure, which will be considered as a locally equivalent martingale measure. Moreover, in the geometric case we use it to simulate the prices of European call options, comparing the results to real data.

## 2.1 Laplace Cumulant Processes

Fix a stochastic basis $(\Omega, \mathcal{F}, P; \mathbb{F})$, with $\mathbb{F}$ which satisfies the usual hypothesis throughout this chapter.

**Definition 2.1.** A random set $A \subset \mathbb{R}_0^+ \times \Omega$ is called *evanescent* if

$$P\left(\{\omega \in \Omega : \exists t \in \mathbb{R}_0^+ \text{ such that } (t, \omega) \in A\}\right) = 0.$$

Two stochastic processes $X = \{X_t\}_t$ and $Y = \{Y_t\}_t$ are equal *up to evanescence*, or equivalently, *up to indistinguishability*, if the random set $\{X \neq Y\}$ is evanescent, i.e.,

$$P\left(\{\omega \in \Omega : \exists t \in \mathbb{R}_0^+ \text{ such that } X_t(\omega) \neq Y_t(\omega)\}\right) = 0.$$

**Definition 2.2.** The real–valued semimartingale $X$ is said to be *exponentially special* if $\exp(X - X_0)$ is a special semimartingale.

**Notation.** The symbol $\mathcal{M}$ denotes the space of càdlàg martingales.

If $X$ is a special semimartingale with canonical decomposition $X = X_0 + A + M$, then the process $A$ can be called *additive compensator*, or *drift*, of $X$ and it is the unique (up to evanescence) predictable process in $\mathcal{V}$ such that $X - X_0 - A \in \mathcal{L}$. By analogy, we set:

**Definition 2.3.** Let $X$ be a real–valued semimartingale. A predictable process $V \in \mathcal{V}$ is called an *exponential compensator* of $X$ if $\exp(X - X_0 - V) \in \mathcal{M}_{loc}$.

We have the following, simple result.

**Proposition 2.1.** *A real–valued semimartingale $X$ has an exponential compensator if and only if it is exponentially special.*

*The exponential compensator is unique up to evanescence.*

In order to prove this proposition, we need some information on the multiplicative decomposition of a semimartingale.

**Lemma 2.1** ([28], Theorem II, 8.21)**.** *Let $X$ be a semimartingale with $X_0 = 1$ such that $X$ and $X_-$ take their values in $(0, \infty)$. Then $X = LD$, where $L$ is a positive local martingale, $D$ is a positive predictable process with locally finite variation and $L_0 = D_0 = 1$, if and only if $X$ is a special semimartingale.*

*In this case, the multiplicative decomposition is unique up to evanescence.*

*Proof of Proposition 2.1.* We first assume that $X$ has an exponential compensator $V \in \mathcal{V}$. We can define

$$D := \exp(V) \quad \text{and} \quad L := \exp(X - X_0 - V) \in \mathcal{M}_{loc}.$$

As the function $\exp(\cdot)$ is Lip–continuous in any compact interval and $V \in \mathcal{V}$, the process $D$ has finite variation and $V_0 = 0$ implies $D_0 = 1$. From the continuity of $\exp(x)$ and the fact that $V$ is predictable it is obvious that $D$ is predictable, as well. Therefore

$$\exp(X - X_0) = LD$$

is a special semimartingale, by Lemma 2.1.

If instead we take $X$ to be an exponentially special semimartingale, then $\exp(X - X_0)$ is a special semimartingale and Lemma 2.1 states that

$$\exp(X - X_0) = LD,$$

where $D$ is a positive predictable process of locally finite variation and $L \in \mathcal{M}_{loc}$, with $D_0 = L_0 = 1$. Defining

$$V := \log(D) \in \mathcal{V},$$

we readily get that $V$ is an exponential compensator of $X$.

In each case, the multiplicative decomposition of a special semimartingale is unique up to evanescence, and so is the exponential compensator. ■

We now turn our attention to the Laplace cumulant process. Let us consider a $d$–dimensional semimartingale $X = \left(X^1, ..., X^d\right)$. We denote by $\mathcal{E}$ and $\mathcal{L}og$ the stochastic exponential and the stochastic logarithm, respectively.

**Definition 2.4.** Let $\theta \in L(X)$ be such that $\theta \cdot X$ is exponentially special. The *Laplace cumulant* $\widetilde{K}^X(\theta)$ of $X$ at $\theta$ is defined as the additive compensator of the real–valued, special semimartingale

$$\mathcal{L}og(\exp(\theta \cdot X)).$$

The *modified Laplace cumulant* $K^X(\theta)$ of $X$ at $\theta$ is the process

$$K^X(\theta) := \log\left(\mathcal{E}\left(\widetilde{K}^X(\theta)\right)\right). \tag{2.1}$$

We note that this definition is well posed.
Indeed, the process $\mathcal{L}og\,(\exp(\theta \cdot X))$ is a special semimartingale, since it satisfies

$$\mathcal{L}og\,(\exp(\theta \cdot X)) = \frac{1}{(\exp(\theta \cdot X))_-} \cdot \exp(\theta \cdot X)\,,$$

hence the additive compensator is predictable considering that $\exp(\theta \cdot X)$ is special and $\frac{1}{(\exp(\theta \cdot X))_-}$ is predictable. This shows that $\widetilde{K}^X(\theta)$ is well defined. As far as $K^X(\theta)$ is concerned, Theorem III, 7.4 in [28] shows that

$$\Delta \widetilde{K}^X(\theta)_t = \int_{\mathbb{R}^d} \left(e^{<\theta_t, x>} - 1\right) \nu^X(\{t\} \times dx) > -1, \quad t \geq 0,$$

since we are working with a version of the characteristics of $X$ such that

$$\nu^X\left(\{t\} \times \mathbb{R}^d\right) \leq 1$$

identically. Thus, the stochastic exponential in (2.1) is strictly positive.

Theorem III, 7.14 in [28] shows that if $\theta \cdot X$ is exponentially special, then $K^X(\theta)$ is its exponential compensator. This implies that the process $Z^\theta = \{Z_t^\theta\}_t$, defined by

$$Z^\theta := \exp\left(\theta \cdot X - K^X(\theta)\right) \in \mathcal{M}_{loc},$$

with $Z_0^\theta = 1$. We try to find conditions under which it is possible to define a probability measure $P^\theta$ in $\mathcal{F}$ which is locally absolutely continuous with respect to $P$ (in symbols, $P^\theta \overset{loc}{\ll} P$) such that $Z^\theta$ is the density process of $P^\theta$ relative to $P$. The following theorem provides us with one possible solution.

**Theorem.** *If $X = \{X_t\}_t$ is an uniformly integrable (UI) martingale, then $X_t \to X_\infty$ a.s. and in $L^1$. Moreover*

$$E\left[X_\infty | \mathcal{F}_t\right] = X_t \text{ a.s. for any } t \in \mathbb{R}_0^+.$$

Hence if we assume that $Z^\theta$ is a $UI$ martingale, then we can define

$$P^\theta(d\omega) := Z_\infty^\theta P(d\omega).$$

It is trivial to show that $Z^\theta$ is the density of $P^\theta$ relative to $P$. Indeed, for any $t \in \mathbb{R}_0^+$ it results:

$$P^\theta(F) = E^P\left[Z_\infty^\theta 1_F\right] = E^P\left[E^P\left[Z_\infty^\theta | \mathcal{F}_t\right] 1_F\right] = E^P\left[Z_t^\theta 1_F\right], \quad F \in \mathcal{F}_t.$$

**Remark 2.1.** Noting that $Z_t^\theta > 0$ for every $t \in \mathbb{R}_0^+$, we can state that $P^\theta$ and $P$ are actually locally equivalent.

## 2.2 Geometric Esscher Measure

We now want to impose some hypothesis such that for any $i = 1, ..., d$ the process

$$S^i := S_0^i \exp\left(X^i\right),$$

with $S_0^i \in \mathbb{R}^+$, is a $P^\theta$–local martingale. In order to do so, we need the following, widely used result.

**Lemma 2.2.** *Let $P, P'$ be two probability measures on $(\Omega, \mathcal{F}; \mathbb{F})$ such that $P' \overset{loc}{\ll} P$ and denote by $Z = \{Z_t\}_t$ the Radon–Nikodym density process. Let $X = \{X_t\}_t$ be a $\mathbb{R}$–valued, càdlàg, adapted process. Then:*

*a. $X$ is a $P'$–martingale if and only if $XZ$ is a $P$–martingale;*

*b. $X$ is a $P'$–local martingale if and only if $XZ$ is a $P$–local martingale.*

*Proof.* a. For any $t \in \mathbb{R}_0^+$, obviously $X_t \in L^1(P') \Leftrightarrow X_t Z_t \in L^1(P)$. Let us suppose that the process $X$ is $P'$–integrable (i.e., $X_t \in L^1(P')$ for any $t \in \mathbb{R}_0^+$). We fix $t \in \mathbb{R}_0^+$ and note that

$$E^{P'}[X_t 1_A] = E^P[X_t Z_t 1_A], \quad A \in \mathcal{F}_t.$$

It follows that, for any $0 \le s < t$, we have $E^{P'}[X_t - X_s | \mathcal{F}_s] \overset{a.s.}{=} 0$ if and only if

$$E^{P'}[(X_t - X_s) 1_A] = E^P[X_t Z_t 1_A] - E^P[X_s Z_s 1_A] = 0, \quad A \in \mathcal{F}_s.$$

This is equivalent to

$$E^P[X_t Z_t - X_s Z_s | \mathcal{F}_s] \overset{a.s.}{=} 0.$$

b. Without loss of generality we may assume that $X_0 = 0$. Fix $\tau$ a $\mathbb{F}$–stopping time and $t > 0$. Since $Z$ is a càdlàg $P$–martingale, Doob's optional sampling theorem applies and we get $E^P[Z_t | \mathcal{F}_\tau] = Z_{t\wedge\tau}$ almost surely. Therefore

$$E^P[|X_t^\tau Z_t|] = E^P[|X_{\tau\wedge t} Z_t|] = E^P[|(XZ)_t^\tau|].$$

Indeed, the process $X$ is progressive (as càdlàg and adapted), hence for any $\mathbb{F}$–stopping time $\sigma$ the variable $X_\sigma$ is $\mathcal{F}_\sigma$–measurable. In our case, this implies that $X_{t\wedge\tau}$ is $\mathcal{F}_{t\wedge\tau}$–measurable, and in particular $\mathcal{F}_\tau$–measurable. Thus,

$$E^P[|X_{\tau\wedge t}| Z_t] = E^P\left[E^P[|X_{\tau\wedge t}| Z_t | \mathcal{F}_\tau]\right] = E^P[|X_{\tau\wedge t}| Z_{\tau\wedge t}].$$

So we conclude that $(XZ)^\tau$ is $P$–integrable $\Leftrightarrow$ $X^\tau Z$ is $P$–integrable $\Leftrightarrow$ $X^\tau$ is $P'$–integrable.
Let us assume that $X^\tau Z$ is $P$–integrable and read

$$X^\tau Z = [X^\tau Z - (XZ)^\tau] + (XZ)^\tau.$$

We call $M := X^\tau Z - (XZ)^\tau$ and show that $M = \{M_t\}_t$ is a $P$–martingale. For any $0 \le s < t$ it results

$$\begin{aligned}
E^P[M_t | \mathcal{F}_s] &= E^P[X_t^\tau Z_t - (XZ)_t^\tau | \mathcal{F}_s] = E^P[(Z_t - Z_{\tau\wedge t}) X_{\tau\wedge t} | \mathcal{F}_s] \\
&\overset{a.s.}{=} E^P\left[E^P\left[(Z_t - Z_{\tau\wedge t}) X_{\tau\wedge t} | \mathcal{F}_{(\tau\vee s)\wedge t}\right] | \mathcal{F}_s\right] \\
&\overset{a.s.}{=} E^P\left[E^P\left[Z_t - Z_{\tau\wedge t} | \mathcal{F}_{(\tau\vee s)\wedge t}\right] X_{\tau\wedge t} | \mathcal{F}_s\right] \\
&\overset{a.s.}{=} E^P\left[\left(Z_{(\tau\vee s)\wedge t} - Z_{\tau\wedge t}\right) X_{\tau\wedge t} | \mathcal{F}_s\right] \\
&= E^P\left[(Z_s - Z_\tau) X_\tau 1_{\{\tau\le s\}} | \mathcal{F}_s\right] = E^P\left[(Z_s - Z_{\tau\wedge s}) X_{\tau\wedge s} 1_{\{\tau\le s\}} | \mathcal{F}_s\right] \\
&\overset{a.s.}{=} (Z_s - Z_{\tau\wedge s}) X_{\tau\wedge s} 1_{\{\tau\le s\}} = M_s.
\end{aligned}$$

It follows that $(XZ)^{\tau}$ is a $P$–martingale $\Leftrightarrow$ $X^{\tau}Z$ is a $P$–martingale. By part (a.), this amounts to saying that $X^{\tau}$ is $P'$–martingale, and the proof is complete.

■

At this point the next theorem is immediate.

**Theorem 2.1.** *Let $\theta \in L(X)$ be such that $\theta \cdot X$ is exponentially special and $Z^{\theta}$ is a UI martingale. Set*

$$\theta_j^{(i)} := \begin{cases} \theta_j, & j \neq i \\ \theta_i + 1, & j = i \end{cases}.$$

*Then the processes $S^i = S_0^i \exp\left(X^i\right)$ are $P^{\theta}$–local martingales if and only if $\theta^{(i)} \cdot X$ is exponentially special and $K^X\left(\theta^{(i)}\right) = K^X(\theta)$ up to evanescence for any $i = 1, ..., d$.*

In this case, we call $P^{\theta}$ *geometric Esscher measure*, or *Esscher martingale transform* for exponential processes.

*Proof.* Fix $i = 1, ..., d$. By Lemma 2.2, $\exp\left(X^i\right)$ is a $P^{\theta}$–local martingale if and only if

$$\exp\left(X^i\right) Z^{\theta} = \exp\left(X^i\right) \exp\left(\theta \cdot X - K^X(\theta)\right) = \exp\left(\theta^{(i)} \cdot X - K^X(\theta)\right)$$

is a $P$–local martingale. Thanks to the uniqueness of the exponential compensator, this amounts to requiring not only that $\theta^{(i)}$ is exponentially special, but also that $K^X\left(\theta^{(i)}\right) = K^X(\theta)$ up to evanescence. ■

Theorem 4.2 in [20] states that if $d = 1$, then the geometric Esscher measure is unique, provided its existence.

**Example 2.1.** Let $X = \{X_t\}_t$ be a $\mathbb{R}^d$–valued Lévy process with generating triplet $(A, \nu, \gamma_h)$ with respect to the truncation function $h$. Suppose that there exists $\theta \in \mathbb{R}^d$ such that

$$E\left[\exp\left(<\theta, X_t>\right)\right] < \infty \quad \text{for some } t > 0.$$

By Remark 1.4 and Example 1.6 we know that $M = \{M_t\}_t$ is a martingale with expectation 1, where

$$M_t := \exp\left(<\theta, X_t> - t\Psi_h(\theta)\right), \quad t \geq 0.$$

It follows that $\{<\theta, X_t>\}_t$ is exponentially special with $K^X(\theta)_t = t\Psi_h(\theta), t \geq 0$, up to evanescence, and in particular the modified Laplace cumulant is a deterministic process. In order to define the Esscher transform $P^{\theta}$ in $\mathcal{F}$ we consider $M$ to be uniformly integrable, so $P^{\theta}(d\omega) = M_{\infty} P(d\omega)$. Thus, if also

$$E\left[\exp\left(<\theta^{(i)}, X_t>\right)\right] < \infty, \quad i = 1, ..., d,\ t > 0,$$

then Theorem 2.1 applies and states that $S^i = S_0^i \exp\left(X^i\right)$ are $P^\theta$–local martingales if and only if $K^X(\theta^{(i)}) = K^X(\theta)$ up to indistinguishability, $i = 1, ..., d$, equality that in this setting reduces to

$$\gamma_h^i + \frac{1}{2}A^{ii} + A^{i\cdot}\theta + \int_{\mathbb{R}} \left[\exp\left(< \theta^{(i)}, x >\right) - \exp\left(< \theta, x >\right) - h^i(x)\right] \nu(dx)$$
$$= \gamma_h^i + \frac{1}{2}A^{ii} + A^{i\cdot}\theta + \int_{\mathbb{R}} \left[e^{<\theta,x>}\left(e^{x^i} - 1\right) - h^i(x)\right] \nu(dx) = 0, \quad i = 1, ..., d,$$

recalling that $A$ is a symmetric matrix. In this case, the geometric Esscher measure is the Esscher transform. ///

## 2.3 Linear Esscher Measure

It is common to model the price process $S$ of a security in a financial market with a stochastic exponential, instead of a classical exponential. However, these two concepts are strictly related by the next result.

**Proposition 2.2.** *Let $X$ and $\overline{X}$ be two real–valued semimartingales. The following holds:*

a. *if $\exp(X) = \mathcal{E}\left(\overline{X}\right)$, then $\overline{X} - \overline{X}_0 = \mathcal{L}og\left(\exp(X)\right)$ up to indistinguishability. Furthermore $X_0 = 0$ and $\Delta\overline{X} > -1$;*

b. *if $\overline{X} = \mathcal{L}og\left(\exp(X)\right)$, then $\exp(X - X_0) = \mathcal{E}\left(\overline{X}\right)$.*

*Proof.* Let $Y := \exp(X)$, so we can state $Y > 0$ and $Y_- > 0$.

a. If $Y = \mathcal{E}\left(\overline{X}\right) = \mathcal{E}\left(\overline{X} - \overline{X}_0\right)$, then $Y_0 = 1$, which in turn implies $X_0 = 0$. Moreover $\mathcal{E}\left(\overline{X}\right) > 0$, so by the expression of the stochastic exponential (see, for example, Proposition 2.41 in [25]) we get $\Delta\overline{X} > -1$. Since $\mathcal{L}og(Y)$ is the unique (up to evanescence) semimartingale starting at 0 such that
$$Y = Y_0\mathcal{E}\left(\mathcal{L}og(Y)\right) = \mathcal{E}\left(\mathcal{L}og(Y)\right),$$
we have $\overline{X} - \overline{X}_0 = \mathcal{L}og\left(\exp(X)\right)$ up to indistinguishability.

b. On the other hand, if $\overline{X} = \mathcal{L}og(Y)$, then it suffices to use the property of the stochastic logarithm we have just cited to obtain
$$\mathcal{E}\left(\overline{X}\right) = \mathcal{E}\left(\mathcal{L}og(Y)\right) = \frac{Y}{Y_0} = \exp(X - X_0).$$

The proof is now complete. ■

For every $i = 1, ..., d$ we set $S^i := S_0^i\mathcal{E}\left(X^i\right)$, where $S_0^i \in \mathbb{R}^+$ and $X = \left(X^1, ..., X^d\right)$ is a $d$–dimensional semimartingale with characteristics $(B, C, \nu^X)$ with respect to a truncation function $h$. We also assume that $\Delta X^i > -1$, so that $S^i > 0$ and can represent the price of a security in the market. Since $dS^i = S_-^i dX^i$, then $S^i$ is a local martingale if $X^i$ is so. Vice versa, it results that $X^i - X_0^i = \frac{1}{S_-^i} \cdot S^i \left(= \mathcal{L}og\left(S^i\right)\right)$ up to evanescence, hence $X^i$ is a local martingale when $S^i$ is so, as well. Similarly to the case of geometric Esscher measure, we search for a process $\theta$ such that $X^i$ (hence $S^i$, as well) is a $P^\theta$–local martingale for every $i = 1, ..., d$. In order to reach the main theorem of the section, we need several technical results taken from [20, 28].

Let $\theta \in L(X)$ be such that $\theta \cdot X$ is exponentially special. Theorem III, 7.4 in [28] states that $\widetilde{K}^X(\theta) = \widetilde{\kappa}(\theta) \cdot A$, where

$$\widetilde{\kappa}(\theta)_t := < \theta_t, b_t > + \frac{1}{2} < \theta_t, c_t\theta_t > + \int_{\mathbb{R}^d} \left(e^{<\theta_t,x>} - 1 - < \theta_t, h(x) >\right) F_t(dx), \quad t \geq 0. \tag{2.2}$$

The following lemma ([20], Lemma 2.11) is crucial, because it allows us to express the drift process as a function of the semimartingale characteristics.

**Lemma 2.3.** *Let $\theta \in L(X)$ be such that $\theta \cdot X$ is a special semimartingale. Then its drift process $D^X(\theta) = \delta(\theta) \cdot A$, where*

$$\delta(\theta)_t := < \theta_t, b_t > + \int_{\mathbb{R}^d} < \theta_t, x - h(x) > F_t(dx), \quad t \geq 0.$$

As a consequence of *Girsanov's theorem* for semimartingales (see, for example, Theorem III, 3.24 in [28]), we can explicitly express the characteristics $\left(B^\theta, C^\theta, \nu^{X^\theta}\right)$ –always with respect to the same $h$– of $X$ under $P^\theta$:

$$\begin{cases} B^{\theta^i} = B^i + c^{i\cdot}\theta \cdot A + h^i(x)\left(\dfrac{e^{<\theta_t,x>}}{1+\widehat{W}(\theta)_t} - 1\right) \star \nu^X, & i = 1, ..., d \\ C^\theta = C \\ \nu^{X^\theta}(dt, dx) = \dfrac{e^{<\theta_t,x>}}{1+\widehat{W}(\theta)_t}\nu^X(dt, dx) \end{cases}, \tag{2.3}$$

where

$$\widehat{W}(\theta)_t := \int_{\mathbb{R}^d} \left(e^{<\theta_t,x>} - 1\right) \nu^X(\{t\} \times dx), \quad t \geq 0.$$

Note that for every $t \in \mathbb{R}_0^+$ it results $\widehat{W}(\theta)_t = \Delta\widetilde{K}^X(\theta)_t > -1$.

**Remark 2.2.** Let us fix $\omega \in \Omega$, $t \geq 0$ and $G \in \mathcal{B}(\mathbb{R}^d)$. We have:

$$\begin{aligned} \nu^X(\omega; \{t\} \times G) &= \int_{\mathbb{R}_0^+} dA_s(\omega) \int_{\mathbb{R}^d} 1_{\{t\}\times G}(s, x) F_{(s,\omega)}(dx) \\ &= \int_{\{t\}} dA_s(\omega) \int_{\mathbb{R}^d} 1_G(x) F_{(t,\omega)}(dx) = F_{(t,\omega)}(G) \int_{\{t\}} dA_s(\omega). \end{aligned}$$

Since

$$\int_{\{t\}} dA_s(\omega) = A_t(\omega) - A_{t-}(\omega),$$

if the function $A_\cdot(\omega)$ is continuous in $t$ then $\nu^X(\omega; \{t\} \times dx)$ is the null measure on $\mathcal{B}(\mathbb{R}^d)$, implying that $\widehat{W}(\theta)_t(\omega) = 0$.

Finally, we introduce the derivatives of cumulant processes, also mentioning Proposition 2.25 in [20].

**Definition 2.5.** Let $\theta \in L(X)$ be such that $\theta \cdot X$ is an exponentially special semimartingale and

$$\left|x^i e^{<\theta_t, x>} - h^i(x)\right| \star \nu^X \in \mathcal{V}, \quad i = 1, ..., d.$$

i. The derivative of $\widetilde{K}^X$ in $\theta$ is the $\mathbb{R}^d$–valued process

$$D\widetilde{K}^X(\theta) = \left(D_1\widetilde{K}^X(\theta), ..., D_d\widetilde{K}^X(\theta)\right),$$

where $D_i\widetilde{K}^X(\theta) := D_i\widetilde{\kappa}(\theta) \cdot A$, with

$$D_i\widetilde{\kappa}(\theta)_t := b_t^i + c_t^{i\cdot}\theta_t + \int_{\mathbb{R}^d} \left(x^i e^{<\theta_t, x>} - h^i(x)\right) F_t(dx), \quad t \geq 0, \tag{2.4}$$

for every $i = 1, ..., d$.

ii. The derivative of $K^X$ in $\theta$ is the $\mathbb{R}^d$–valued process

$$DK^X(\theta) = \left(D_1K^X(\theta), ..., D_dK^X(\theta)\right),$$

where

$$D_iK^X(\theta) := \frac{1}{1 + \widehat{W}(\theta)} \cdot D_i\widetilde{K}^X(\theta), \quad i = 1, ..., d.$$

Let us just note that (2.4) can be gotten by formally differentiating in $\theta$ (2.2), using that $c$ is symmetric.

**Proposition 2.3.** *Under the hypothesis of Definition 2.5, the derivatives of* $\widetilde{K}^X$ *and* $K^X$ *in* $\theta$ *are predictable processes in* $\mathcal{V}^d$ *and it results* $D_iK^X(\theta) = D_i\kappa(\theta) \cdot A$*, where*

$$D_i\kappa(\theta)_t := b_t^i + c_t^{i\cdot}\theta_t + \int_{\mathbb{R}^d} \left(\frac{x^i e^{<\theta_t, x>}}{1 + \widehat{W}(\theta)_t} - h^i(x)\right) F_t(dx), \quad t \geq 0, \tag{2.5}$$

*for every* $i = 1, ..., d$.

At this point we are ready to present the analogue of Theorem 2.1 for stochastic exponential.

**Theorem 2.2.** *Let* $\theta \in L(X)$ *be such that* $\theta \cdot X$ *is exponentially special and* $Z^\theta$ *is a UI martingale. Then the processes* $S^i = S_0^i \mathcal{E}(X^i)$ *are* $P^\theta$*–local martingales if and only if* $\left|x^i e^{<\theta_t, x>} - h^i(x)\right| \star \nu^X \in \mathcal{V}$ *and* $D_iK^X(\theta) = 0$ *for any* $i = 1, ..., d$.

In this case, we call $P^\theta$ *linear Esscher measure*, or *Esscher martingale transform* for linear processes.

*Proof.* Assume that $S^i$, and so also $X^i$, is a $P^\theta$–local martingale for any $i = 1, ..., d$. By (2.3), we can get the characteristics $\left(B^\theta, C^\theta, \nu^{X^\theta}\right)$ of $X$ under $P^\theta$. In particular,

$$\begin{cases} b^\theta = b + c\theta + \int_{\mathbb{R}^d} h(x) \left(\frac{e^{<\theta, x>}}{1 + \widehat{W}(\theta)} - 1\right) F(dx), & i = 1, ..., d \\ c^\theta = c \\ F^\theta(dx) = \frac{e^{<\theta, x>}}{1 + \widehat{W}(\theta)} F(dx) \end{cases},$$

where the integral in the expression of $b^\theta$ must be read componentwise. Fix $i = 1, ..., d$; since $X^i$ is a $P^\theta$–local martingale, we choose the deterministic processes

$$e_j^{(i)} := \begin{cases} 0, & j \neq i \\ 1, & j = i \end{cases}$$

and express $X^i = e^{(i)} \cdot X$. By Lemma 2.3 we deduce

$$0 = b^{\theta^i} \cdot A + \left(x^i - h^i(x)\right) \star \nu^{X^\theta} = b^{\theta^i} \cdot A + \left(x^i - h^i(x)\right) \frac{e^{<\theta_t, x>}}{1 + \widehat{W}(\theta)_t} \star \nu^X$$

$$= B^i + c^{i \cdot}\theta \cdot A + \left(x^i \frac{e^{<\theta_t, x>}}{1 + \widehat{W}(\theta)_t} - h^i(x)\right) \star \nu^X, \quad i = 1, ..., d.$$

As $\left|x^i e^{<\theta_t, x>} - h^i(x)\right| \star \nu^X \in \mathcal{V}$ for every $i = 1, ..., d$ (see [20], Theorem 4.4 for the detailed proof of this technical result), a straightforward application of Proposition 2.3 gives $DK^X(\theta) = 0$.

Vice versa, assuming $DK^X(\theta) = 0$, the same proposition yields

$$0 = b^i \cdot A + c^{i \cdot}\theta \cdot A + \left(\frac{x^i e^{<\theta_t, x>}}{1 + \widehat{W}(\theta)_t} - h^i(x)\right) \star \nu^X$$

$$= b^{\theta^i} \cdot A + \left(\frac{x^i e^{<\theta_t, x>}}{1 + \widehat{W}(\theta)_t} - h^i(x) \frac{e^{<\theta_t, x>}}{1 + \widehat{W}(\theta)_t}\right) \star \nu^X$$

$$= b^{\theta^i} \cdot A + \left(x^i - h^i(x)\right) \star \nu^{X^\theta}.$$

Hence Lemma 2.3 shows that $X^i$ is a $P^\theta$–local martingale for every $i = 1, ..., d$, and so is $S^i$. ■

Theorem 4.5 in [20] states that the *linear Esscher measure* is unique, provided its existence.

**Example 2.2.** Let $X = \{X_t\}_t$ be a $\mathbb{R}^d$–valued Lévy process with generating triplet $(A, \nu, \gamma_h)$ with respect to the truncation function $h$. Suppose that there exists $\theta \in \mathbb{R}^d$ such that $E[\exp(<\theta, X_t>)] < \infty$ for some $t > 0$. Then $\{<\theta, X_t>\}_t$ is exponentially special with $K^X(\theta)_t = t\Psi_h(\theta)$, $t \geq 0$, up to evanescence, as argued in Example 2.1. We further consider $M = \{M_t\}_t$, with $M_t := \exp(<\theta, X_t> - t\Psi(\theta))$, $t \geq 0$, to be uniformly integrable, so it is possible to define the Esscher transform $P^\theta$ in $\mathcal{F}$ as $P^\theta(d\omega) = M_\infty P(d\omega)$.

Take the submultiplicative function

$$g(x) := (|x| \vee 1) \exp(<\theta, x>), \quad x \in \mathbb{R}^d,$$

and assume that $X_t$ has finite $g$–moment for some $t > 0$. By Theorem 1.3 this implies that $\int_{|x|>1} |x| \exp(<\theta, x>)\, \nu(dx) < \infty$. Hence we also have

$$\int_{\mathbb{R}^d} \left|x^i \exp(<\theta, x>) - h^i(x)\right| \nu(dx) < \infty, \quad i = 1, ..., d,$$

so for every $t \geq 0$ and $i = 1, ..., d$ it results

$$\left(\left|x^i \exp(<\theta, x>) - h^i(x)\right| \star \nu^X\right)_t = t \int_{\mathbb{R}^d} \left|x^i \exp(<\theta, x>) - h^i(x)\right| \nu(dx).$$

Thus, the process $\left|x^i e^{<\theta_t,x>} - h^i(x)\right| \star \nu^X \in \mathcal{V}$ for any $i = 1, ..., d$.

Now Theorem 2.2 applies and states that $S_0^i \mathcal{E}\left(X^i\right)$ is a $P^\theta$–local martingale if and only if $D_i K^X(\theta) = 0$ for every $i$. By Proposition 2.3, this equality holds if and only if

$$\gamma_h^i + A^{i\cdot}\theta + \int_{\mathbb{R}^d} \left(x^i e^{<\theta,x>} - h^i(x)\right) \nu(dx) = 0, \quad i = 1, ..., d, \tag{2.6}$$

because here $\widehat{W}(\theta) = 0$ (see Remark 2.2). Recalling that

$$\Psi_h(\theta) =< \gamma_h, \theta > + \frac{1}{2} < \theta, A\theta > + \int_{\mathbb{R}^d} \left(e^{<\theta,x>} - 1 - < \theta, h(x) >\right) \nu(dx)$$

and that $A$ is a symmetric matrix, we note that (2.6) can be obtained by formally differentiating $\Psi_h(\theta)$ in $\theta^i$ for any $i$. In this case, the linear Esscher measure is the Esscher transform.

Furthermore, by (2.3) the characteristics of $X$ relative to $P^\theta$ are given by

$$\begin{cases} B_t^{\theta^i} = \gamma_h^i t + A^{i\cdot}\theta t + \int_{\mathbb{R}^d} h^i(x)\left(e^{<\theta,x>} - 1\right)\nu(dx)\, t, \\ C_t^\theta = At \\ \nu^{X^\theta}(dt, dx) = e^{<\theta,x>}\nu^X(dt, dx) = dt\, e^{<\theta,x>}\nu(dx) \end{cases}, \quad t \geq 0,$$

so Theorem 1.5 states that it is a $PIIS$ process even under $P^\theta$ (cf. Proposition 9.6 in [25]). ///

## 2.4 Applications

### 2.4.1 Geometric Esscher Measure

In this section we apply the theory developed so far to a specific price process. Let $(\Omega, \mathcal{F}, P)$ be a probability space, $L = \{L_t\}_{t\geq 0}$ be a driving, $\mathbb{R}$–valued Lévy process with generating triplet $\left(\sigma^2, \nu, \gamma_h\right)$ with respect to a fixed truncation function $h$. Denote by $\mathbb{F}$ the augmented filtration of $L$, so it fulfills the usual hypothesis. The dynamics of the spot prices are given by the process $S = \{S_t\}_t$, defined by

$$S := S_0 \exp(G), \quad S_0 \in \mathbb{R}^+,$$

where the log–prices $G = \{G_t\}_t$ follow

$$dG_t = \left(\rho^G - \lambda^G G_{t-}\right) dt + \sigma_{t-} dL_t, \quad G_0 = 0, \tag{1}$$

with $\rho^G, \lambda^G \in \mathbb{R}$. The volatility process $\sigma^2 = \left\{\sigma_t^2\right\}_t$ is defined as in [22] by

$$\sigma_t^2 := \left(k \int_0^t e^{X_s} ds + \sigma_0^2\right) e^{-X_{t-}}, \quad t \geq 0, \tag{2}$$

where $\sigma_0^2 \in \mathbb{R}^+$ and $X = \{X_t\}_t$ is the càdlàg, adapted process given by

$$X_t := \eta \log \delta - \sum_{0<s\leq t} \log\left[1 + \Phi(\Delta L_s)^2\right], \quad t \geq 0,$$

with $k>0$, $\eta>0$, $\Phi\geq 0$. Proposition 3.2 in [22] shows

$$\sigma_t^2 = kt - \eta\int_0^t \sigma_s^2\, ds + \Phi \sum_{0<s<t} \sigma_s^2 \left(\Delta L_s\right)^2 + \sigma_0^2, \quad t\geq 0.$$

It is evident from (2) that $\sigma^2$ is a left–continuous and adapted process. In any case, we prefer writing $\sigma_-$ instead of just $\sigma$ to highlight the fact that we are working with a càg and adapted, hence predictable, process.

Given an adapted process $r=\{r_t\}_t$ representing the interest rate, the discount factor is $\{r_t-\delta_t\}_t$, with $\delta_t := \lambda^G G_{t-}$, $t\geq 0$. In this setting, the discounted spot prices $\widetilde{S}=\left\{\widetilde{S}_t\right\}_t$ become

$$\widetilde{S}_t := \exp\left(-\int_0^t (r_s-\delta_s)\, ds\right) S_t, \quad t\geq 0.$$

We can interpret $\delta=\{\delta_t\}_t$ as the convenience yield (i.e., the difference between storage expenses and consumption value per unit of time) in the case of commodities or the dividend yield for stocks (see [7]).
Thus, we define the process $G'=\{G'_t\}_t$, where $G'_t := -\int_0^t (r_s-\delta_s)\, ds + G_t$ for every $t\geq 0$, whose dynamics are $dG'_t = -(r_t-\delta_t)\, dt + dG_t$. We can now express $\widetilde{S}=S_0\exp(G')$. Therefore, we want to find a process $\theta\in L(G')$ such that $\widetilde{S}$ is a $P^\theta$–local martingale (provided the existence of $P^\theta$, obviously).

**Remark 2.3.** From the dynamics of $G$ in (1), it follows that $G$ jumps at the same times as $L$ does, and $\Delta G_t = \sigma_{t-}\Delta L_t$, $t\geq 0$. Fixed $t\in\mathbb{R}^+$ and $\omega\in\Omega$, the measure associated to its jumps is

$$\begin{aligned}\mu^G(\omega;(0,t]\times A) &= \sum_{s\leq t} 1_{\{x\neq 0\}\cap A}\left(\Delta G_s(\omega)\right)\\ &= \sum_{s\leq t} 1_{\{x\neq 0\}}\left(\Delta L_s(\omega)\right) 1_{(\sigma_{s-}(\omega))^{-1}A}\left(\Delta L_s(\omega)\right)\\ &= \mu^L\left(w;(0,t]\times\left(\sigma_-(\omega)\right)^{-1}A\right), \quad A\in\mathcal{B}(\mathbb{R}).\end{aligned}$$

Moreover it is clear that $\mu^{G'}=\mu^G$, therefore $F^{G'}_{(t,\omega)}(dy) = \nu\left(\sigma_{t-}(\omega)\cdot\right)^{-1}(dy)$, meaning that

$$F^{G'}_{(t,\omega)}(A) = \int_{\mathbb{R}} 1_A\left(\sigma_{t-}(\omega)\, y\right)\nu(dy), \quad A\in\mathcal{B}(\mathbb{R}).$$

**Theorem 2.3.** *Let $\theta\in L(G')$ be such that $\theta\cdot G'$ is exponentially special and $Z^\theta$ is a UI martingale. If also $(\theta+1)\cdot G'$ is exponentially special and $\theta$ satisfies*

$$\left(\theta_t+\frac{1}{2}\right)\sigma_{t-}^2\sigma^2 - r_t + \rho^G + \sigma_{t-}\gamma_h + \int_{\mathbb{R}}\left(e^{(\theta_t+1)\sigma_{t-}y} - e^{\theta_t\sigma_{t-}y} - \sigma_{t-}h(y)\right)\nu(dy) = 0 \tag{2.7}$$

*for every $t\geq 0$, then $\widetilde{S}$ is a $P^\theta$–local martingale.*

Note that $G'$ is 1–dimensional, therefore the *geometric Esscher measure* is unique (if it exists).

*Proof.* We first determine the characteristics of $G'$ under $P$ with respect to $h$ as function of the generating triplet of the Lévy process $L$. We define

$$\widetilde{G'}(h)_t := \sum_{s\leq t}\left[\Delta G'_s - h(\Delta G'_s)\right] = \left((x-h(x)) \star \mu^G\right)_t, \quad t\geq 0$$

and $G'(h)_t := G'_t - \widetilde{G'}(h)_t$, $t \geq 0$. Thanks to the previous remark, we can write

$$\begin{aligned} dG'(h)_t &= dG'_t - d\widetilde{G'}(h)_t = -(r_t-\delta_t)\,dt + dG_t - (x-h(x))\;\mu^G(dt,dx) \\ &= -\left(r_t-\rho^G\right)dt + \sigma_{t-}dL_t - (\sigma_{t-}y - h(\sigma_{t-}y))\;\mu^L(dt,dy). \end{aligned}$$

As $(\sigma_-\cdot L)^c = \sigma_-\cdot L^c$ and $<\sigma_-\cdot L^c, \sigma_-\cdot L^c>= \int \sigma_-^2\sigma^2 dt$, considering that Lemma 2.3 gives

$$\delta^L(\sigma_-)_t = \sigma_{t-}\gamma_h + \int_{\mathbb{R}}\sigma_{t-}(y-h(y))\;\nu(dy), \quad t\geq 0$$

we get

$$\begin{cases} b_t^{G'} = -r_t + \rho^G + \sigma_{t-}\gamma_h + \int_{\mathbb{R}}(h(\sigma_{t-}y) - \sigma_{t-}h(y))\;\nu(dy) \\ c_t^{G'} = \sigma_{t-}^2\sigma^2 \\ F_t^{G'}(dy) = \nu(\sigma_{t-}\cdot)^{-1}(dy) \end{cases}, \quad t\geq 0. \tag{2.8}$$

At this point, if $\theta$ satisfies $\widetilde{\kappa}(\theta+1) - \widetilde{\kappa}(\theta) = 0$, Theorem 2.1 applies and we can affirm that $\widetilde{S}$ is a $P^\theta$–local martingale. In fact, using the expression of $\widetilde{\kappa}(\theta)$ provided by (2.2) it results

$$\begin{aligned} &\widetilde{\kappa}(\theta+1)_t - \widetilde{\kappa}(\theta)_t \\ &= b_t^{G'} + \frac{1}{2}\sigma_{t-}^2\sigma^2 + \theta_t\sigma_{t-}^2\sigma^2 + \int_{\mathbb{R}}\left(e^{(\theta_t+1)x} - e^{\theta_t x} - h(x)\right)F_t^{G'}(dx) \\ &= \left(\theta_t + \frac{1}{2}\right)\sigma_{t-}^2\sigma^2 - r_t + \rho^G + \sigma_{t-}\gamma_h + \int_{\mathbb{R}}\left(e^{(\theta_t+1)\sigma_{t-}y} - e^{\theta_t\sigma_{t-}y} - \sigma_{t-}h(y)\right)\nu(dy) \\ &= 0, \quad t\geq 0, \end{aligned}$$

by (2.7). This completes the proof. ∎

If $L$ follows a $NIG$ distribution, then the solutions to Equation (2.7) can be explicitly expressed, as shown in the next result. We just need to recall that, if a random variable $X \sim NIG(\alpha,\beta,\mu,\delta)$, then its moment generating function is provided by

$$E\left[e^{zX}\right] = \exp\left[\mu z + \delta\left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+z)^2}\right)\right], \quad -\alpha-\beta\leq z\leq \alpha-\beta. \tag{2.9}$$

**Lemma 2.4.** *Let $L = \{L_t\}_t$ be a Lévy process following a $NIG$ distribution with parameters $(\alpha,\beta,\mu,\delta)$ and set the truncation function*

$$h(x) := x1_D(x), \quad x\in\mathbb{R}.$$

*If a process $\theta = \{\theta_t\}_t$ such that*

$$(\sigma_{t-}\theta_t)(\omega),\; \sigma_{t-}(\theta_t+1)(\omega) \in [-\alpha-\beta, \alpha-\beta], \quad t\in\mathbb{R}_0^+,\; \omega\in\Omega \tag{2.10}$$

*fulfills Equation* (2.7), *then for every* $t \geq 0$ *we have*

$$\theta_t^{1,2} = \frac{1}{2\left(R_t^2 + \delta^2\sigma_{t-}^2\right)\sigma_{t-}}\Bigg(-\delta^2\sigma_{t-}^3 - 2\beta\delta^2\sigma_{t-}^2 - R_t^2\left(\sigma_{t-} + 2\beta\right) \pm\sqrt{4\alpha^2\delta^2R_t^2\sigma_{t-}^2 + 4R_t^4\alpha^2 - R_t^2\delta^2\sigma_{t-}^4 - \frac{R_t^6}{\delta^2} - 2\sigma_{t-}^2R_t^4}\Bigg), \tag{2.11}$$

*with* $R_t := -r_t + \rho^G + \mu\sigma_{t-}$.

*Proof.* Since $L_1 \sim NIG\left(\alpha, \beta, \mu, \delta\right)$, (1.12) applies and along with Example 1.3 states that for every $z \in [-\alpha - \beta, \alpha - \beta]$ it results

$$E\left[e^{zL_1}\right] = \exp\left[\mu z + \gamma_1 z + \int_{\mathbb{R}} \left(e^{zx} - 1 - zx1_D\left(x\right)\right)\nu\left(dx\right)\right],$$

where

$$\gamma_1 := \frac{2\delta\alpha}{\pi}\int_0^1 \sinh\left(\beta x\right)K_1\left(\alpha x\right)dx$$

and

$$\nu\left(dx\right) = \frac{\delta\alpha}{\pi\left|x\right|}e^{\beta x}K_1\left(\alpha\left|x\right|\right)dx.$$

Now (2.9) yields, for the same $z \in [-\alpha - \beta, \alpha - \beta]$, the next equality:

$$\delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - \left(\beta + z\right)^2}\right) = \gamma_1 z + \int_{\mathbb{R}}\left(e^{zx} - 1 - zx1_D\left(x\right)\right)\nu\left(dx\right). \tag{2.12}$$

Following the foregoing proof we have

$$\begin{aligned} 0 &= \widetilde{\kappa}\left(\theta+1\right)_t - \widetilde{\kappa}\left(\theta\right)_t \\ &= R_t + \sigma_{t-}\gamma_1 + \int_{\mathbb{R}}\left[e^{(\theta_t+1)\sigma_{t-}x} - 1 - \left(e^{\theta_t\sigma_{t-}x} - 1\right) - \sigma_{t-}x1_D\left(x\right)\right]\nu\left(dx\right) \\ &= R_t + \sigma_{t-}\gamma_1 + \int_{\mathbb{R}}\left[e^{(\theta_t+1)\sigma_{t-}x} - 1 - \sigma_{t-}\left(\theta_t+1\right)x1_D\left(x\right)\right]\nu\left(dx\right) \\ &\qquad - \int_{\mathbb{R}}\left(e^{\theta_t\sigma_{t-}x} - 1 - \sigma_{t-}\theta_t x1_D\left(x\right)\right)\nu\left(dx\right), \quad t \geq 0. \end{aligned} \tag{2.13}$$

Considering that $\sigma_-\gamma_1 = \sigma_-\left[\left(\theta+1\right) - \theta\right]\gamma_1$ trivially holds, we can continue expanding on the chain of equalities in (2.13) arriving at:

$$\begin{aligned} 0 &= R_t + \left[\sigma_{t-}\left(\theta_t+1\right)\gamma_1 + \int_{\mathbb{R}}\left(e^{(\theta_t+1)\sigma_{t-}x} - 1 - \sigma_{t-}\left(\theta_t+1\right)x1_D\left(x\right)\right)\nu\left(dx\right)\right] \\ &\qquad - \left[\sigma_{t-}\theta_t\gamma_1 + \int_{\mathbb{R}}\left(e^{\theta_t\sigma_{t-}x} - 1 - \sigma_{t-}\theta_t x1_D\left(x\right)\right)\nu\left(dx\right)\right] \\ &= R_t + \delta\left(\sqrt{\alpha^2 - \left(\beta + \sigma_{t-}\theta_t\right)^2} - \sqrt{\alpha^2 - \left(\beta + \sigma_{t-}\left(\theta_t+1\right)\right)^2}\right), \quad t \geq 0, \end{aligned}$$

where in the last passage we used (2.12) together with the technical assumption (2.10). In order to complete the proof we therefore need to find, for any $t \geq 0$, the possible solutions of the following equation:

$$R_t + \delta\left(\sqrt{\alpha^2 - \left(\beta + \sigma_{t-}\theta_t\right)^2} - \sqrt{\alpha^2 - \left(\beta + \sigma_{t-}\left(\theta_t+1\right)\right)^2}\right) = 0. \tag{2.14}$$

We write (2.14) as

$$\sqrt{\alpha^2 - (\beta + \sigma_{t-}(\theta_t + 1))^2} = \frac{R_t}{\delta} + \sqrt{\alpha^2 - (\beta + \sigma_{t-}\theta_t)^2}$$

and then we square both sides, getting

$$2\frac{R_t}{\delta}\sqrt{\alpha^2 - (\beta + \sigma_{t-}\theta_t)^2} = -\left(\sigma_{t-}^2 + 2\beta\sigma_{t-} + \frac{R_t^2}{\delta^2}\right) - 2\sigma_{t-}^2\theta_t.$$

Repeating once again the same operation of squaring we end up with an equation of second degree in $\theta$, namely

$$4\sigma_{t-}^2\left(\sigma_{t-}^2 + \frac{R_t^2}{\delta^2}\right)\theta_t^2 + 4\sigma_{t-}\left(\sigma_{t-}^3 + 2\beta\sigma_{t-}^2 + \frac{R_t^2}{\delta^2}\sigma_{t-} + 2\beta\frac{R_t^2}{\delta^2}\right)\theta_t + c_t = 0, \tag{2.15}$$

where the coefficient $c_t := \left(\sigma_{t-}^2 + 2\beta\sigma_{t-} + \frac{R_t^2}{\delta^2}\right)^2 - 4\frac{R_t^2}{\delta^2}\left(\alpha^2 - \beta^2\right)$. The solutions to (2.15) can be algebraically computed (admittedly, the calculation is quite tricky) and are provided by (2.11). This completes the proof. ■

Lemma 2.4 does not add anything to the theory developed so far, but it is crucial in applications. In fact, having an explicit expression for $\theta$ enables us to simulate the spot prices under the martingale measure, given in this case by the geometric Esscher measure. Both branches of the solutions (2.11) have been tested empirically, but only

$$\theta_t = \frac{1}{2\left(R_t^2 + \delta^2\sigma_{t-}^2\right)\sigma_{t-}}\Bigg(-\delta^2\sigma_{t-}^3 - 2\beta\delta^2\sigma_{t-}^2 - R_t^2\left(\sigma_{t-} + 2\beta\right) - \sqrt{4\alpha^2\delta^2R_t^2\sigma_{t-}^2 + 4R_t^4\alpha^2 - R_t^2\delta^2\sigma_{t-}^4 - \frac{R_t^6}{\delta^2} - 2\sigma_{t-}^2R_t^4}\Bigg) \tag{2.16}$$

prevents the risk–neutral dynamics from exploding.

### Simulation of the $P^\theta$–dynamics

Since we are going to consider options on stocks and not on commodities, the convenience yield $\delta$ is redundant, so we remove it by setting $\lambda^G = 0$. For the sake of simplicity we consider a constant annual interest rate $r \in \mathbb{R}^+$ and put $\rho^G = 0$: in this way, under the historical probability measure $P$, the log–prices $G$ follow a $COGARCH$ process, as introduced in [22], and (1) reduces to

$$dG_t = \sigma_{t-}dL_t, \quad G_0 = 0.$$

Such types of processes are generalized by the so–called $COGARCH(p,q)$, which have been introduced in [6]: Theorem 2.2 therein displays the connection with the model we are adopting. In our approach the driving Lévy process $L$ is assumed to have a $NIG$ distribution and we further suppose that the hypothesis of Theorem 2.3 are satisfied (with truncation function $\overline{h}(x) = x1_D(x)$, $x \in \mathbb{R}$), so that the geometric Esscher measure $P^\theta$ for $\widetilde{S}$ is uniquely determined. We estimate all the parameters of the model from the time series of the underlying asset log–prices using the $R$–package *yuima* (see, e.g., [16]), which computes them according to the procedure explained in [17]. At this point, we use the found coefficients to simulate a sufficiently large number of paths of the variance process $\sigma^2$: thanks to (2.16), from each of them we can recover a trajectory of $\theta$, as well. The problem therefore reduces to generating the paths of $G$ from those we have just obtained. After the change of measure, $L$ is not a Lévy process anymore: indeed, by (2.3) and Remark 2.2 we have

$$F_t^\theta\left(dx\right) = e^{\theta_t x}\nu\left(dx\right), \quad t \geq 0$$

and Theorem 1.5 ensures that $L$ is not even a $PIIS$ process under $P^\theta$. As a consequence of this fact, $G$ is not a $COGARCH$ process under the martingale measure and the simulation of its trajectories gets complicated. The idea we follow to overcome this setback is to use the canonical representation for semimartingales ([28], Theorem II, 2.34): given a $d$–dimensional semimartingale $X = \{X_t\}_t$ with characteristics $\left(B, C, \nu^X\right)$ relative to the truncation function $h$, then the next representation holds:

$$X = X_0 + X^c + B + h\left(x\right) \star \left(\mu^X - \nu^X\right) + \left(x - h\left(x\right)\right) \star \mu^X. \tag{2.17}$$

Fix the truncation function $h\left(x\right) := x\, 1_{\{|z|<\epsilon\}}\left(x\right)$, where $\epsilon \leq 1$ is arbitrarily chosen. Considering Examples 1.3 and 1.5, straightforward computations provide the generating triplet of $L$ under $P$ with respect to $h$:

$$\begin{cases} \sigma^2 = 0 \\ \nu\left(dx\right) = \dfrac{\delta\alpha}{\pi\left|x\right|} e^{\beta x} K_1\left(\alpha\left|x\right|\right) dx \\ \gamma_h = \gamma - \int_{\epsilon<|x|<1} x\, \nu\left(dx\right) \end{cases},$$

where $\gamma = \mu + \dfrac{2\delta\alpha}{\pi}\displaystyle\int_0^1 \sinh\left(\beta x\right) K_1\left(\alpha x\right) dx$. From the proof of Theorem 2.3, specifically from (2.8), we readily get the characteristics of $G$ under $P$:

$$\begin{cases} b_t = \sigma_{t-}\gamma_h + \int_{\mathbb{R}} \left(h\left(\sigma_{t-}x\right) - \sigma_{t-}h\left(x\right)\right)\, \nu(dx) \\ c_t = 0 \\ F_t(dx) = \nu(\sigma_{t-}\cdot)^{-1}\left(dx\right) \end{cases}, \quad t \geq 0.$$

We invoke once again (2.3) together with Remark 2.2 to find them under $P^\theta$:

$$\begin{cases} b_t^\theta = b_t + \int_{\mathbb{R}} h\left(x\right)\left(e^{\theta_t x} - 1\right)\, F_t(dx) \\ c_t^\theta = 0 \\ F_t^\theta(dx) = e^{\theta_t x}\nu(\sigma_{t-}\cdot)^{-1}\left(dx\right) \end{cases}, \quad t \geq 0. \tag{2.18}$$

We now focus on the representation of $G$ given by (2.17). By construction $G_0 = 0$, and obviously $G^c = 0$, as well, since $c_t^\theta = 0$, $t \geq 0$. As regards the term $B = \{B_t\}_t$, it can be easily obtained by

$$B_t = \int_{(0,t)} b_s^\theta\, ds, \quad t \geq 0,$$

hence we expand on the computations in (2.18):

$$b_t^\theta = b_t + \int_{\mathbb{R}} \sigma_{t-} y \, 1_{\{|z|<\epsilon\}} (\sigma_{t-} y) \left( e^{\theta_t \sigma_{t-} y} - 1 \right) \nu (dy) , \quad t \geq 0.$$

Substituting the expression of $b_t$ we conclude

$$b_t^\theta = \sigma_{t-} \gamma_h + \int_{|x|<\epsilon/\sigma_{t-}} \sigma_{t-} y \frac{\delta\alpha}{\pi |y|} e^{(\theta_t \sigma_{t-} + \beta) y} K_1 (\alpha |y|) \, dy - \int_{(-\epsilon,\epsilon)} \sigma_{t-} y \frac{\delta\alpha}{\pi |y|} e^{\beta y} K_1 (\alpha |y|) \, dy, \quad t \geq 0.$$

We next replace the term $h(x) \star (\mu^G - \nu^G)$ with another one representing the variation of the small jumps of $G$, as suggested in [3] in the case of Lévy processes. In particular, we introduce the process $\sigma^2(\epsilon) = \left\{ \sigma^2(\epsilon)_t \right\}_t$, defined by

$$\sigma^2(\epsilon)_t := \int_{|x|<\epsilon} x^2 \, F_t^\theta (dx) , \quad t \geq 0.$$

With the usual calculations we get

$$\begin{aligned} \sigma^2(\epsilon)_t &= \int_{\mathbb{R}} x^2 e^{\theta_t x} \, 1_{\{|z|<\epsilon\}} (x) \, F_t (dx) = \int_{\mathbb{R}} \sigma_{t-}^2 y^2 e^{\theta_t \sigma_{t-} y} \, 1_{\{|z|<\epsilon\}} (\sigma_{t-} y) \, \nu (dy) \\ &= \int_{|x|<\epsilon/\sigma_{t-}} \sigma_{t-}^2 y^2 \frac{\delta\alpha}{\pi |y|} e^{(\theta_t \sigma_{t-} + \beta) y} K_1 (\alpha |y|) \, dy, \quad t \geq 0. \end{aligned}$$

Taking a Brownian motion $W = \{W_t\}_t$, we approximate a trajectory of $h(x) \star (\mu^G - \nu^G)$ with one of $\sigma^2(\epsilon) \, W$.

Finally we turn our attention to the term $(x - h(x)) \star \mu^G$. In order to simulate its paths, we proceed in analogy to the Lévy–Itô decomposition theorem presented in Section 1.3 (Theorem 1.4). In that instance, we proved that, if $L$ is a Lévy process, then

$$\begin{aligned} \int_{(0,t]\times D_{(1,\infty)}} x \, J(ds, dx) &= \int_{[0,t]\times \mathbb{R}^d} (x - x 1_D (x)) \, \mu^L (ds, dx) \\ &= \left( (x - x \, 1_D (x)) \star \mu^L \right)_t , \quad t \geq 0 \end{aligned}$$

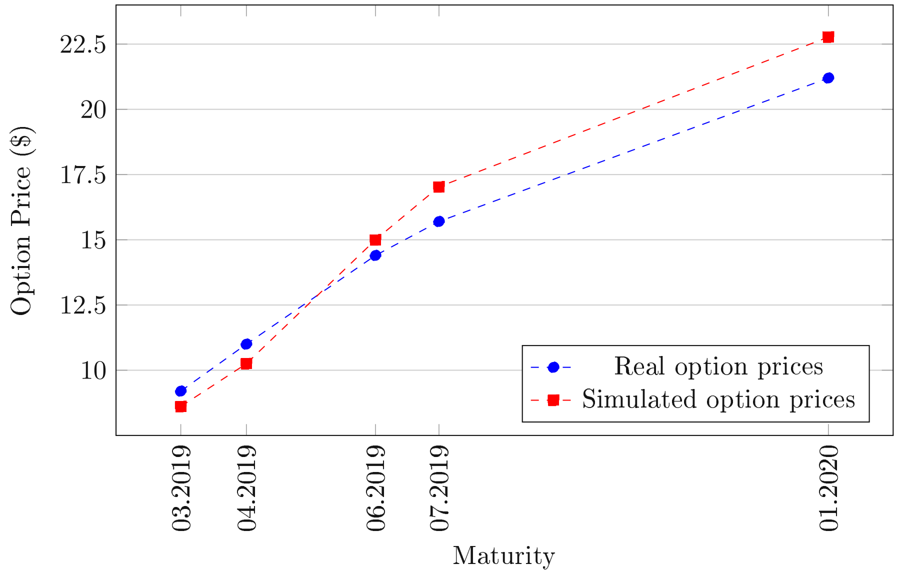
is a compound Poisson process with distribution $\left( \frac{\widetilde{\nu}}{t\nu(D_{1,\infty})} \right) \phi^{-1} (dx)$ (see (1.14)) and constant $\nu (D_{1,\infty})$. In this case, we consider a nonhomogeneous Poisson process with time–varying intensity

$$\begin{aligned} \lambda_t := F_t^\theta (\{|x| > \epsilon\}) &= \int_{\mathbb{R}} e^{\theta_t \sigma_{t-} y} \, 1_{D_{\epsilon,\infty}} (\sigma_{t-} y) \frac{\delta\alpha}{\pi |y|} e^{\beta y} K_1 (\alpha |y|) \, dy \\ &= \int_{|x|>\epsilon/\sigma_{t-}} \frac{\delta\alpha}{\pi |y|} e^{(\theta_t \sigma_{t-} + \beta) y} K_1 (\alpha |y|) \, dy, \quad t \geq 0. \end{aligned}$$

The jumps time of such process have been simulated with a *thinning algorithm*, as proposed in [23]. Instead, the time–varying jumps sizes are assumed to be

$$\begin{aligned} c_t &:= \int_{\mathbb{R}} x \, 1_{D_{\epsilon,\infty}} (x) \, F_t^\theta (dx) \\ &= \int_{|x|>\epsilon/\sigma_{t-}} \sigma_{t-} y \frac{\delta\alpha}{\pi |y|} e^{(\theta_t \sigma_{t-} + \beta) y} K_1 (\alpha |y|) \, dy, \quad t \geq 0. \end{aligned}$$

Figure 2.1: AAPL call option, Strike 165$.




Following these steps we can get as many trajectories of $G$ (and so of $S$) under $P^\theta$ as we want. Let $N$ be the number of iterations we run and $S^i$, for $i = 1, ..., N$, be the corresponding, simulated trajectories of the spot price. Recalling that $P^\theta$ is a martingale measure for $S$ which is locally equivalent to $P$, we use it as pricing measure. This means that we obtain the price of an European call option with strike $K$ and maturity $T$ by computing the sample mean of the vector of components

$$e^{-rT}\left(S^i\left(T\right) - K\right)^+, \quad i = 1, ..., N.$$

### Empirical Results

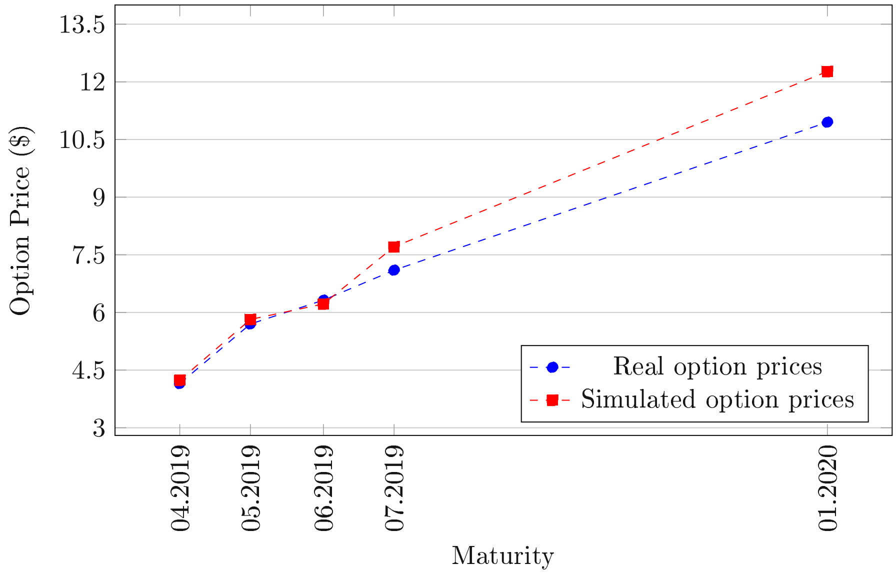
We have empirically tested the Esscher method with the prices of call options on *Apple Inc.* stock (ticker symbol: AAPL) with fixed strike at 165$. Figure 2.1 above shows the simulated option prices as a function of their maturities. We refer to Appendix A, Section A.1, for the used code. The average of the absolute values of percentage difference is 6.6525%.

We also tried to apply the geometric Esscher measure to the prices of call options on *Microsoft Corporation* stock (ticker symbol: MSFT) with strike 110$: Figure 2.2 below displays the outcomes. In this case, the mean of absolute values of percentage difference settles down at 5.2701%.

**Remark 2.4.** In both the simulations the real data was obtained from the website `https://www.nasdaq.com`, where American options are traded. Our argument generates the option prices assuming they are European, instead. In any case, this approximation can be accepted as meaningful for two reasons:

- the analyzed derivatives have short-term maturities, so we are allowed to ignore the dividend yield;

Figure 2.2: MSFT call option, Strike 110$.




- the time values of the options were always positive, implying that it is convenient to sell the option rather than exercising its call right.

### 2.4.2 Linear Esscher Measure

As we have already argued at the beginning of Section 2.3, sometimes it is convenient to use the stochastic exponential $\mathcal{E}$ instead of the standard exponential exp to model prices. In the previous subsection, we had the processes $S = \{S_t\}_t$, with $S_t = S_0 \exp(G_t)$, $t \geq 0$, which described the spot prices and $\widetilde{S} = \left\{\widetilde{S}_t\right\}_t$, with $\widetilde{S}_t = S_0 \exp(G'_t)$, $t \geq 0$, for the discounted spot prices. Considering that $G_0 = G'_0 = 0$, thanks to Proposition 2.2 we can express these processes by stochastic exponential as follows:

$$S_t = S_0 \mathcal{E}\left(\overline{G}\right)_t, \; \widetilde{S}_t = S_0 \mathcal{E}\left(\overline{G'}\right)_t, \quad t \geq 0,$$

where $\overline{G} \coloneqq \mathcal{L}og(\exp(G))$ and $\overline{G'} \coloneqq \mathcal{L}og(\exp(G'))$. The next theorem ([28], Theorem II, 8.10) shows how to derive the characteristics of $\overline{G'}$ from those of $G'$.

**Theorem 2.4.** *Let $X$ and $\overline{X}$ be two real–valued semimartingales such that $\overline{X} = \mathcal{L}\text{og}(\exp(X))$. Denote by $\left(B, C, \nu^X\right)$ the characteristics of $X$ with respect to a truncation function $h$. Then there exists a version $\left(\overline{B}, \overline{C}, \nu^{\overline{X}}\right)$ of the characteristics of $\overline{X}$, always relative to $h$, satisfying*

$$\begin{cases} \overline{B} = B + \frac{C}{2} + \left(h\left(e^x - 1\right) - h(x)\right) \star \nu^X \\ \overline{C} = C \\ 1_G(x) \star \nu^{\overline{X}} = 1_G\left(e^x - 1\right) \star \nu^X, \quad G \in \mathcal{B}(\mathbb{R}) \end{cases} . \tag{2.19}$$

In particular, recalling that

$$\begin{cases} B = b \cdot A \\ C = c \cdot A \\ \nu^X(\omega; dt, dx) = dA_t(\omega) F_{(t,\omega)}(dx) \quad \mathrm{P}\text{–}a.s. \end{cases},$$

from (2.19) we get

$$\begin{cases} \overline{b}_t = b_t + \frac{c_t}{2} + \int_{\mathbb{R}} (h(e^x - 1) - h(x)) \, F_t(dx) \\ \overline{c}_t = c_t \\ \overline{F}_t(dx) = F_t(e^{\cdot} - 1)^{-1}(dx) \end{cases}, \quad t \geq 0, \tag{2.20}$$

where the last expression can be explicitly written in this way:

$$\overline{F}_t(G) = \int_{\mathbb{R}} 1_G(e^x - 1) F_t(dx), \quad G \in \mathcal{B}(\mathbb{R}), t \geq 0.$$

With the new expression of the process $\widetilde{S}$, the idea is to apply Theorem 2.2 in order to find another sufficient condition to make $P^\theta$ a martingale measure for $S$. We obtain the following result:

**Theorem 2.5.** *Let $\theta \in L(\overline{G'})$ be such that $\theta \cdot \overline{G'}$ is exponentially special and $Z^\theta$ is a UI martingale. If the process $\left|x\, e^{\theta_t x} - h(x)\right| \star \nu^{\overline{G'}} \in \mathcal{V}$ and $\theta$ satisfies*

$$\left(\theta_t + \frac{1}{2}\right)\sigma_{t-}^2 \sigma^2 - r_t + \rho^G + \sigma_{t-}\gamma_h + \int_{\mathbb{R}} \left[\left(e^{\sigma_{t-}y} - 1\right) e^{\theta_t\left(e^{\sigma_{t-}y} - 1\right)} - \sigma_{t-} h(y)\right] \nu(dy) = 0 \tag{2.21}$$

*for every $t \geq 0$, then $\widetilde{S}$ is a $P^\theta$–local martingale.*

Note that in this case $P^\theta$ is the *Esscher martingale transform* for linear processes, so it is unique.

*Proof.* We first determine the characteristics of $\overline{G'}$ under $P$ with respect to the truncation function $h$. From the proof of Theorem 2.3 (see, in particular, (2.8)) we already know those of $G'$; since $\overline{G'} = \mathcal{L}og(\exp(G'))$, the previous theorem applies and (2.20) states

$$\begin{cases} \overline{b}_t = b_t^{G'} + \frac{1}{2} c_t^{G'} + \int_{\mathbb{R}} (h(e^x - 1) - h(x)) \, F_t^{G'}(dx) \\ \overline{c}_t = c_t^{G'} \\ \overline{F}_t(dx) = F_t^{G'}(e^{\cdot} - 1)^{-1}(dx) \end{cases}, \quad t \geq 0. \tag{2.22}$$

At this point, if $DK^{\overline{G'}}(\theta) = 0$, then Theorem 2.2 enables us to conclude that $P^\theta$ is a martingale measure for $S$, as desired. Since $DK^{\overline{G'}}(\theta) = \int D\kappa(\theta) \, dt$, we focus on the term

$$\begin{aligned} D\kappa(\theta)_t &= \overline{b}_t + \overline{c}_t \theta_t + \int_{\mathbb{R}} \left( \frac{x\, e^{\theta_t x}}{1 + \widehat{W}(\theta)_t} - h(x) \right) \overline{F}_t(dx) \\ &= \left(\theta_t + \frac{1}{2}\right)\sigma_{t-}^2 \sigma^2 - r_t + \rho^G + \sigma_{t-}\gamma_h + \int_{\mathbb{R}} (h(\sigma_{t-}y) - \sigma_{t-} h(y)) \, \nu(dy) \\ &+ \int_{\mathbb{R}} (h(e^x - 1) - h(x)) F_t^{G'}(dx) + \int_{\mathbb{R}} \left(x\, e^{\theta_t x} - h(x)\right) F_t^{G'}(e^{\cdot} - 1)^{-1}(dx), \quad t \geq 0 \end{aligned} \tag{2.23}$$

noting that $\widehat{W}(\theta) = 0$ (see Remark 2.2). Now for every $t \geq 0$ it results

$$\int_{\mathbb{R}} \left(h\left(e^{x}-1\right)-h\left(x\right)\right) F_t^{G'}\left(dx\right) = \int_{\mathbb{R}} \left(h\left(e^{\sigma_{t-}y}-1\right)-h\left(\sigma_{t-}y\right)\right) \nu\left(dy\right);$$

furthermore

$$\int_{\mathbb{R}} \left(x\, e^{\theta_t x} - h\left(x\right)\right) F_t^{G'}\left(e^{\cdot}-1\right)^{-1}(dx) = \int_{\mathbb{R}} \Big[\left(e^{z}-1\right) e^{\theta_t\left(e^{z}-1\right)} - h\left(e^{z}-1\right)\Big] F_t^{G'}(dz)$$
$$= \int_{\mathbb{R}} \Big[\left(e^{\sigma_{t-}y}-1\right) e^{\theta_t\left(e^{\sigma_{t-}y}-1\right)} - h\left(e^{\sigma_{t-}y}-1\right)\Big] \nu\left(dy\right).$$

Plugging these two terms into (2.23) we end up with

$$\left(\theta_t + \frac{1}{2}\right)\sigma_{t-}^2\sigma^2 - r_t + \rho^G + \sigma_{t-}\gamma_h + \int_{\mathbb{R}} \Big[\left(e^{\sigma_{t-}y}-1\right) e^{\theta_t\left(e^{\sigma_{t-}y}-1\right)} - \sigma_{t-}h\left(y\right)\Big]\nu(dy)$$
$$= D\kappa\left(\theta\right)_t = 0, \quad t \geq 0,$$

by (2.21). This completes the proof. ■

**Remark 2.5.** Even if the theory developed for the linear Esscher measure allowed us to get Theorem 2.5, which is the natural analogon of Theorem 2.3, unfortunately we could not simulate option prices modeling the stock dynamics with a stochastic exponential. In fact, we were not able to find an appropriate driving Lévy process for the explicit computation of the solutions to (2.21). This drawback prevented us from obtaining an analogous of Lemma 2.4, and consequently from simulating a trajectory for the process $\theta$, which was essential to start the procedure described in subsection 2.4.1.

# Chapter 3

# Calibration To Derivative Prices

Reasoning on the construction of the Esscher measure, one might note that the density process is defined solely by the spot prices. This fact can represent an intuitive setback especially in liquid markets, where the time series of derivative prices is available. In order to solve this problem, we can approach it by the so–called *Calibration Method*: in few words, it consists in modeling the dynamics of the spot prices directly under a martingale measure, which is customarily said to be "chosen" by the market.
In the next sections we introduce some theoretical topics aiming to explain in which sense the market determines ("chooses") the measure.

## 3.1 Lévy Processes On Skorokhod Space

Let us consider a $\mathbb{R}^d$–valued, additive process $\{X_t\}_t$ with system of generating triplets $\{(A_t, \nu_t, \gamma_t)\}_t$ on a probability space $(\Omega, \mathcal{F}, P)$. We introduce the *Skorokhod Space* $\mathfrak{D}$ as the space of the $\mathbb{R}^d$–valued, càdlàg functions defined on $\mathbb{R}_0^+$. Specifically, we have $\mathfrak{D} := D\left([0,\infty), \mathbb{R}^d\right)$. For any $t \geq 0$ denote by $x_t : \mathfrak{D} \to \mathbb{R}^d$ the following function:

$$x_t(\xi) := \xi(t), \quad \xi \in \mathfrak{D}.$$

We can now endow the space $\mathfrak{D}$ with the $\sigma$–algebra $\mathcal{F}_{\mathfrak{D}} := \sigma(\{x_t, t \geq 0\})$. It is also natural to consider a filtration $\mathbb{F}$ on the measurable space $(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}})$, namely

$$\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}, \quad \text{where } \mathcal{F}_t := \sigma(\{x_s, 0 \leq s \leq t\}), \quad t \geq 0. \tag{3.1}$$

Since $\{X_t\}_t$ is a càdlàg process, we can define the map $\phi : \Omega \to \mathfrak{D}$ given by

$$\phi(\omega) := X_.(\omega), \quad \omega \in \Omega,$$

where $X_.(\omega) : [0,\infty) \to \mathbb{R}^d$, with $X_.(\omega)(t) := X_t(\omega)$ for any $t \geq 0$. The function $\phi$ is $\mathcal{F}/\mathcal{F}_{\mathfrak{D}}$ measurable: to see this, it suffices to read

$$\mathcal{F}_{\mathfrak{D}} = \sigma(\{x_t, t \geq 0\}) = \sigma\left(\left\{x_t^{-1}(B), B \in \mathcal{B}\left(\mathbb{R}^d\right), t \geq 0\right\}\right).$$

Hence for every $B \in \mathcal{B}(\mathbb{R}^d)$ and $t \geq 0$ we have

$$\begin{aligned}\phi^{-1}\left(x_t^{-1}(B)\right) &= \left\{\omega \in \Omega : \phi(\omega) \in x_t^{-1}(B)\right\} = \left\{\omega \in \Omega : X_.(\omega) \in x_t^{-1}(B)\right\} \\ &= \left\{\omega \in \Omega : X_t(\omega) \in B\right\} = X_t^{-1}(B) \in \mathcal{F},\end{aligned}$$

and the desired measurability of $\phi$ follows immediately. This allows us to construct the pushforward measure $P^{\mathfrak{D}}$ on $(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}})$, that is,

$$P^{\mathfrak{D}}(A) := P\phi^{-1}(A) = P\left(\phi^{-1}(A)\right), \quad A \in \mathcal{F}_{\mathfrak{D}}. \tag{3.2}$$

We now focus on the stochastic process $\{x_t\}_t$ defined on the probability space $\left(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}}, P^{\mathfrak{D}}\right)$. Fix a cylinder set $C \in \mathcal{F}_{\mathfrak{D}}$, i.e.,

$$C = \{\xi \in \mathfrak{D} : \xi(t_1) \in B_1, ..., \xi(t_n) \in B_n\},$$

for some $t_1 < t_2 < ... < t_n$, $B_1, ..., B_n \in \mathcal{B}(\mathbb{R}^d)$ and $n \in \mathbb{N}$. By (3.2), we get

$$\begin{aligned}P^{\mathfrak{D}}(x_{t_1} \in B_1, ..., x_{t_n} \in B_n) = P^{\mathfrak{D}}(C) &= P\left(\phi^{-1}(C)\right) \\ &= P\left(X_{t_1} \in B_1, ..., X_{t_n} \in B_n\right).\end{aligned}$$

Thus, $\{x_t\}_t$ and $\{X_t\}_t$ are identical in law, whence $\{x_t\}_t$ is an additive process with the same system of generating triplets as $\{X_t\}_t$. Furthermore, if $\{X_t\}_t$ were a Lévy process, then the temporal homogeneity would be inherited by $\{x_t\}_t$, which would be a Lévy process, as well.

Let us now take into account two Lévy processes $(\{x_t\}_t, P)$ and $(\{x_t\}_t, P')$, both defined on the Skorokhod space $(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}})$ endowed with the filtration $\mathbb{F}$ in (3.1). We can think of $P$ and $P'$ as two probability measures on $(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}})$ determined by two Lévy processes defined on a pair of probability spaces (not necessarily the same space), as shown by the previous argument. We want to find out what conditions must be assumed in order to have

$$P\big|_{\mathcal{F}_t} \sim P'\big|_{\mathcal{F}_t} \quad \text{for every } t > 0.$$

This problem was first solved by Skorokhod, Kunita, Watanabe and Neumann, among others. We summarize their results in the next theorem.

**Theorem 3.1.** *Let $(\{x_t\}_t, P)$, $(\{x_t\}_t, P')$ be Lévy processes on $\mathbb{R}^d$ with generating triplets $(A, \nu, \gamma)$ and $(A', \nu', \gamma')$, respectively. Then the following properties are equivalent:*

*a.* $P\big|_{\mathcal{F}_t} \sim P'\big|_{\mathcal{F}_t}$ *for every* $t > 0$;

*b. the generating triplets satisfy*

$$A = A', \qquad \nu \sim \nu'.$$

*Besides, considering the function $\phi : \mathbb{R}^d \to \mathbb{R}$ defined by $\phi := \log\left(\frac{d\nu'}{d\nu}\right)$, it results:*

$$\int_{\mathbb{R}^d} \left(e^{\phi(x)/2} - 1\right)^2 \nu(dx) < \infty \tag{3.3}$$

*and*

$$\gamma' - \gamma - \int_{|x|\leq 1} x(\nu' - \nu)(dx) \in \left\{Ay,\, y \in \mathbb{R}^d\right\}. \tag{3.4}$$

*In this case, chosen* $\eta \in \mathbb{R}^d$ *such that* $\gamma' - \gamma - \int_{|x|\leq 1} x\,(\nu' - \nu)\,(dx) = A\eta$, *there exists a process* $U = \{U_t\}_t$ *defined on* $\mathfrak{D}$ *which fulfills:*

*i.* $U$ *is a* $P$*–Lévy process on* $\mathbb{R}$ *with generating triplet* $\left(\sigma_U^2, \nu_U, \gamma_U\right)$ *given by*

$$\begin{cases} \sigma_U^2 = <\eta, A\eta> \\ \nu_U = \nu\phi^{-1}\big|_{\mathbb{R}\setminus\{0\}} \\ \gamma_U = -\frac{1}{2} < \eta, A\eta > - \int_{\mathbb{R}} \left(e^y - 1 - y\,1_D\,(y)\right)\nu\phi^{-1}\,(dy) \end{cases};$$

*ii.* $E^P\left[e^{U_t}\right] = E^{P'}\left[e^{-U_t}\right] = 1$ *for every* $t \geq 0$*;*

*iii.* $e^{U_t} = \dfrac{dP'\big|_{\mathcal{F}_t}}{dP\big|_{\mathcal{F}_t}}$ $P - a.s.$ *for every* $t > 0$.

We refer to [27], Theorems 33.1 and 33.2 for a proof as well as an explicit expression of the process $U$. We just note that, according to Theorem 1.1, by (i.) the process $U$ is unique up to identity in law and that the components of the integral in (3.4) are finite by (3.3):

$$\left|\int_{|x|\leq 1} x^i\,(\nu' - \nu)\,(dx)\right| = \left|\int_{|x|\leq 1} x^i \left(e^{\phi(x)} - 1\right)\nu\,(dx)\right|$$
$$\leq \int_{|x|\leq 1} |x| \left|e^{\phi(x)} - 1\right| \nu\,(dx) < \infty, \quad i = 1, ..., d.$$

## 3.2 Relative Entropy Of Probability Measures

We present the definition of relative entropy as proposed in [18].

**Definition 3.1.** Given two probability measures $P, P'$ on a measurable space $(\Omega, \mathcal{F})$, the relative entropy $H\,(P, P')$ of $P$ with respect to $P'$ is defined by

$$H\,(P, P') := \begin{cases} \int_\Omega \log\left(\frac{dP}{dP'}\,(\omega)\right) P\,(d\omega), & \text{if } P \ll P' \\ \infty, & \text{otherwise} \end{cases}.$$

First of all, we have to discuss whether or not this definition is well posed. To this aim, we have to prove that the function $\log\left(\frac{dP}{dP'}\right)$ is defined $P$–a.s. in $\Omega$ when $P \ll P'$.
This is obvious when $P \sim P'$, because in this case $\frac{dP}{dP'} > 0$ $P'$–a.s., and therefore $P$–a.s., as well. If instead $P \ll P'$, but $P \nsim P'$, then $\frac{dP}{dP'} = 0$ is a subset of $\Omega$ with positive $P'$–probability (so, for example, it would not be possible to define $\int_\Omega \log\left(\frac{dP}{dP'}\,(\omega)\right) P'\,(d\omega)$). We put

$$\Omega' := \left\{\omega \in \Omega : \frac{dP}{dP'}\,(\omega) > 0\right\} \in \mathcal{F};$$

it results $P'\,(\Omega') < 1$. Considering the measurable space $\left(\Omega', \mathcal{F}\big|_{\Omega'}\right)$, then $P\big|_{\Omega'} := P\big|_{\mathcal{F}|_{\Omega'}}$ and $P'\big|_{\Omega'} := P'\big|_{\mathcal{F}|_{\Omega'}}$ are equivalent measures (note that $P'\big|_{\Omega'}$ is certainly not a probability measure), because

$$\frac{dP\big|_{\Omega'}}{dP'\big|_{\Omega'}} = \frac{dP}{dP'}\bigg|_{\Omega'} > 0.$$

It follows that

$$P(\Omega') = \int_{\Omega'} \frac{dP}{dP'}\bigg|_{\Omega'} (\omega)\, P'(d\omega) = \int_{\Omega} \frac{dP}{dP'}(\omega)\, P'(d\omega) = P(\Omega) = 1, \quad \text{as desired.}$$

However, it is still not clear if we can write $\int_\Omega \log\left(\frac{dP}{dP'}(\omega)\right) P(d\omega)$, since the integrand has not constant sign so the integral could not make sense. For this reason, we consider

$$\left(\log\left(\frac{dP}{dP'}(\omega)\right)\right)^- = -\log\left(\frac{dP}{dP'}(\omega)\right) 1_{\{z\leq 1\}}\left(\frac{dP}{dP'}(\omega)\right), \quad \omega \in \Omega'.$$

Then, recalling that $\log(x)\, 1_{\{z\geq 1\}}(x) \leq x$ for any $x > 0$, we get

$$\begin{aligned}
&\int_{\Omega'} \left(\log\left(\frac{dP}{dP'}(\omega)\right)\right)^- P(d\omega) = \int_{\Omega'} \left(-\log\left(\frac{dP}{dP'}(\omega)\right)\right) 1_{\{z\leq 1\}}\left(\frac{dP}{dP'}(\omega)\right) P(d\omega)\\
&= \int_{\Omega'} \log\left(\frac{dP'|_{\Omega'}}{dP|_{\Omega'}}(\omega)\right) 1_{\{z\geq 1\}}\left(\frac{dP'|_{\Omega'}}{dP|_{\Omega'}}(\omega)\right) P(d\omega) \leq \int_{\Omega'} \frac{dP'|_{\Omega'}}{dP|_{\Omega'}}(\omega)\, P(d\omega)\\
&= P'(\Omega') \leq 1 < \infty,
\end{aligned}$$

where in the second equality we have used $\frac{dP'|_{\Omega'}}{dP|_{\Omega'}} = \left(\frac{dP}{dP'}\big|_{\Omega'}\right)^{-1}$. Therefore $H(P, P')$ is actually a quantity we can handle.

From now on we fix a measurable space $(\Omega, \mathcal{F})$. The following property of the relative entropy is as simple as important.

**Proposition 3.1.** *For any pair $P, P'$ of probability measures on $(\Omega, \mathcal{F})$ it results*

$$H(P, P') \geq 0.$$

*Equality holds if and only if $P = P'$.*

*Proof.* We just study the nontrivial case, i.e., we assume $P \ll P'$. Recalling that $\log x \leq x - 1$ for any $x > 0$, we have

$$\begin{aligned}
H(P, P') &= \int_{\Omega'} \log\left(\frac{dP}{dP'}(\omega)\right) P(d\omega) = -\int_{\Omega'} \log\left(\frac{dP'|_{\Omega'}}{dP|_{\Omega'}}(\omega)\right) P(d\omega)\\
&\geq \int_{\Omega'} \left(1 - \frac{dP'|_{\Omega'}}{dP|_{\Omega'}}(\omega)\right) P(d\omega) = P(\Omega') - P'(\Omega') = 1 - P'(\Omega') \geq 0.
\end{aligned}$$

If $P \nsim P'$, then $P'(\Omega') < 1$, as already discussed above, so $H(P, P') > 0$. If instead we assume $P \sim P'$ (hence $P'(\Omega') = 1$), bearing in mind that $\log x = x - 1 \Leftrightarrow x = 1$, we can state that the equality

$$-\int_{\Omega'} \log\left(\frac{dP'}{dP}(\omega)\right) P(d\omega) = \int_{\Omega'} \left(1 - \frac{dP'}{dP}(\omega)\right) P(d\omega)$$

holds if and only if $\frac{dP'}{dP} = 1$ $P$–a.s. in $\Omega'$, and therefore in $\Omega$, too. In turn, this is equivalent to say $P = P'$ in $\mathcal{F}$, completing the proof. ■

The fact that $H$ is nonnegative enables us to understand it as an expression of the similarity of two probability measures. For this reason, the relative entropy is also called distance (or divergence), namely the *Kullback–Leibler* distance.

The following inequality is well known:

$$\sum_{i=1}^{m} a_i \log \frac{a_i}{b_i} \geq a \log \frac{a}{b} \quad \text{for any } a_i, b_i > 0,\ m \in \mathbb{N}, \tag{3.5}$$

where $a := \sum_{i=1}^{m} a_i$ and analogously $b := \sum_{i=1}^{m} b_i$. More precisely, equality in (3.5) holds if and only if

$$\frac{a_1}{b_1} = \frac{a_2}{b_2} = ... = \frac{a_m}{b_m}.$$

This elementary tool suffices to prove a crucial property of relative entropy: the strict convexity.

**Proposition 3.2.** *Given three probability measures $P, Q$ and $P'$ on $(\Omega, \mathcal{F})$ satisfying $P \sim P', Q \sim P'$ and $P \neq Q$, then for every $\alpha \in (0,1)$ it results*

$$H\left(\alpha P + (1-\alpha) Q, P'\right) < \alpha H\left(P, P'\right) + (1-\alpha) H\left(Q, P'\right).$$

*Proof.* Let us fix a generic $\alpha \in (0,1)$ and define the new probability measure

$$\mu := \alpha P + (1-\alpha) Q;$$

we have $\mu \sim P'$, obviously. We define

$$f := \frac{dP}{dP'}, \quad g := \frac{dQ}{dP'}, \quad h := \alpha f + (1-\alpha) g,$$

so it is immediate to note that $\frac{d\mu}{dP'} = h$. We claim the existence of a subset $M$ of $\{\omega \in \Omega : f(\omega) > 0, g(\omega) > 0, f(\omega) \neq g(\omega)\}$ such that $P'(M) > 0$. If that were true, then in the set $\{f > 0, g > 0\}$ the following would be satisfied:

$$\alpha f \log \frac{\alpha f}{\alpha} + (1-\alpha) g \log \frac{(1-\alpha) g}{(1-\alpha)} \geq h \log h,$$

by (3.5). In particular, this inequality would be strict in a set with positive $P'$–probability, hence

$$\int_{\{f>0,g>0\}} \alpha f \log f \, dP' + \int_{\{f>0,g>0\}} (1-\alpha) g \log g \, dP' > \int_{\{f>0,g>0\}} h \log h \, dP'.$$

Since

$$\{h > 0\} = \{f > 0, g = 0\} \bigcup \{f > 0, g > 0\} \bigcup \{f = 0, g > 0\}$$

and furthermore $\{f > 0, g = 0\}, \{f = 0, g > 0\}$ are two $P'$–null sets given $P \sim P' \sim Q$, we could conclude

$$\begin{aligned} H\left(\mu, P'\right) =& \int_{\{h>0\}} h \log h \, dP' < \int_{\{f>0\}} \alpha f \log f \, dP' + \int_{\{g>0\}} (1-\alpha) g \log g \, dP' \\ &= \alpha H\left(P, P'\right) + (1-\alpha) H\left(Q, P'\right). \end{aligned}$$

In order to show our claim, assume by contradiction that $f = g$ $P'$–a.s. in $\{f > 0, g > 0\}$. Therefore, for every $A \in \mathcal{F}$ we have

$$\begin{aligned} P(A) &= E^{P'}[f\,1_A] = E^{P'}\left[f\,1_{A\cap\{f>0,g>0\}}\right] + E^{P'}\left[f\,1_{A\cap\{f>0,g=0\}}\right] \\ &= E^{P'}\left[g\,1_{A\cap\{f>0,g>0\}}\right], \end{aligned}$$

as $P'(\{f > 0, g = 0\}) = 0$. In the same way we get

$$Q(A) = E^{P'}\left[g\,1_{A\cap\{f>0,g>0\}}\right].$$

This implies that $P = Q$, which is an absurdity since we started off with two different probability measures. The proof is now complete. ■

We finally show a result which computes explicitly the relative entropy of two equivalent Lévy processes in the Skorokhod space as a function of their generating triplets.

**Theorem 3.2.** *Let $(\{x_t\}_t, P)$, $(\{x_t\}_t, P')$ be Lévy processes on $\mathbb{R}^d$ defined on $(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}})$ with generating triplets $(A, \nu, \gamma)$ and $(A', \nu', \gamma')$, respectively. Suppose that $P|_{\mathcal{F}_t} \sim P'|_{\mathcal{F}_t}$ for every $t > 0$ and choose $\eta \in \mathbb{R}^d$ such that*

$$\gamma' - \gamma - \int_{|x|\le 1} x\,(\nu' - \nu)(dx) = A\eta.$$

*Assume also that $E^P[g(U_t)] < \infty$ for some $t > 0$, where $g(x) := (|x| \vee 1)\,e^{|x|}$ for any $x \in \mathbb{R}$. Then for every $T > 0$ it results*

$$H\left(P'|_{\mathcal{F}_T}, P|_{\mathcal{F}_T}\right) = \frac{T}{2} < \eta, A\eta > +T\int_{\mathbb{R}^d}\left(\frac{d\nu'}{d\nu}\log\frac{d\nu'}{d\nu} + 1 - \frac{d\nu'}{d\nu}\right)d\nu. \tag{3.6}$$

*Proof.* Let us fix a finite time horizon $T > 0$. For any $z \in (0, 1)$, by assumption we have

$$E^P\left[e^{zU_T}\right] = \int_{\mathbb{R}} e^{zx}\,P_{U_T}(dx) \le \int_{\mathbb{R}} g(x)\,P_{U_T}(dx) < \infty.$$

We introduce the moment generating function

$$MGF_{U_T}(z) := E^P\left[e^{zU_T}\right] = e^{T\Psi(z)}, \quad z \in (0,1),$$

where $\Psi$ is the cumulant function of the Lévy process $U$, i.e.,

$$\Psi(z) = \frac{1}{2}\sigma_U^2 z^2 + \gamma_U z + \int_{\mathbb{R}}\left(e^{zx} - 1 - zx1_D(x)\right)\nu_U(dx), \quad z \in (0,1), \tag{3.7}$$

and the last equality is ensured by (1.12). Actually $MGF_{U_T}$ is well defined also in $z = 1$, with $MGF_{U_T}(1) = E^P\left[e^{U_T}\right] = e^{T\Psi(1)} = 1$, by (ii.) in Theorem 3.1.

We can see that $MGF_{U_T}$ is differentiable in $(0,1)$ with $MGF_{U_T}{}'(\cdot) = E^P\left[U_T e^{\cdot U_T}\right]$. Indeed, for every $z \in (0,1)$, $MGF_{U_T}(z) = \int_{\mathbb{R}} e^{zx}\,P_{U_T}(dx)$ and we can derive under integral sign since

$$|x|\,e^{zx} \le g(x) \in L^1(P_{U_T}), \quad z \in (0,1), x \in \mathbb{R}.$$

At this point the dominated convergence theorem readily shows that

$$\lim_{z\to 1^-} MGF_{U_T}{}'(z) = \int_{\mathbb{R}} xe^x\, P_{U_T}(dx) = E^P\left[U_T e^{U_T}\right].$$

On the other hand, we introduce the function

$$f(z) := e^{T\Psi(z)}, \quad z \in (0,1).$$

Even in this case we can affirm that $f$ is differentiable in its domain, with derivative provided by

$$f'(z) = Te^{T\Psi(z)}\Psi'(z), \quad z \in (0,1).$$

In particular the following equality is true:

$$\Psi'(z) = \sigma_U^2 z + \gamma_U + \int_{\mathbb{R}} \left(xe^{zx} - x\,1_D(x)\right)\nu_U(dx), \quad z \in (0,1).$$

In fact, we can derive under the integral sign in (3.7) as, for every $z \in (0,1)$, it results that $|xe^{zx} - x| \le 1 + e$, $x \in D$, with $xe^{zx} - x \le Cx^2$ in a neighborhood of 0 for some constant $C > 0$ not depending on $z$. Moreover, $|x|\, e^{zx} \le g(x)$ for $|x| > 1$, with $\int_{|x|>1} |x|\, e^{|x|}\,\nu_U(dx) < \infty$ by Theorem 1.3. Applying another time the Lebesgue's convergence theorem we arrive at

$$\lim_{z\to 1^-} MGF_{U_T}{}'(z) = Te^{T\Psi(1)}\left(\sigma_U^2 + \gamma_U + \int_{\mathbb{R}}\left(xe^x - x\,1_D(x)\right)\nu_U(dx)\right).$$

Therefore

$$\begin{aligned} E^P\left[U_T e^{U_T}\right] &= Te^{T\Psi(1)}\left(\sigma_U^2 + \gamma_U + \int_{\mathbb{R}}\left(xe^x - x\,1_D(x)\right)\nu_U(dx)\right) \\ &= \frac{T}{2} < \eta, A\eta > + T\int_{\mathbb{R}^d}\left(\frac{d\nu'}{d\nu}\log\frac{d\nu'}{d\nu} + 1 - \frac{d\nu'}{d\nu}\right)d\nu, \end{aligned}$$

using the expression of $\left(\sigma_U^2, \nu_U, \gamma_U\right)$ in (i.) of Theorem 3.1. Since $\frac{dP'|_{\mathcal{F}_t}}{dP|_{\mathcal{F}_t}} = e^{U_t}$ for every $t > 0$ (see (iii.) in Theorem 3.1) and

$$H\left(P'|_{\mathcal{F}_T}, P|_{\mathcal{F}_T}\right) = \int_{\mathfrak{D}} \frac{dP'|_{\mathcal{F}_T}}{dP|_{\mathcal{F}_T}} \log\frac{dP'|_{\mathcal{F}_T}}{dP|_{\mathcal{F}_T}}\, dP = E^P\left[U_T e^{U_T}\right],$$

we get (3.6) and we are done. ■

**Remark 3.1.** In the case of two $\mathbb{R}$–valued Lévy processes $(\{x_t\}_t, P)$, $(\{x_t\}_t, P')$ with generating triplets $\left(\sigma^2, \nu, \gamma\right)$ and $\left(\sigma^{2\prime}, \nu', \gamma'\right)$, respectively, under the hypothesis of the previous theorem (3.6) reduces to

$$H\left(P'|_{\mathcal{F}_T}, P|_{\mathcal{F}_T}\right) = \frac{T}{2}\sigma^2\eta^2 + T\int_{\mathbb{R}}\left(\frac{d\nu'}{d\nu}\log\frac{d\nu'}{d\nu} + 1 - \frac{d\nu'}{d\nu}\right)d\nu.$$

If we are dealing with pure jump processes (e.g., $NIG$ processes), the first term in the sum of the right–hand side is 0; if instead $\sigma^2 > 0$, then we can express

$$\frac{T}{2}\sigma^2\eta^2 = \frac{T}{2\sigma^2}\left(\gamma' - \gamma - \int_{|x|\le 1} x\left(\nu' - \nu\right)(dx)\right)^2,$$

restoring Proposition 9.10 in [11].

## 3.3 Application To European Call Options

Let $(\Omega, \mathcal{F}, Q)$ be a probability space and $L = \{L_t\}_t$ be a driving, $\mathbb{R}$–valued Lévy process with generating triplet $\left(\sigma^2, \nu, \gamma\right)$. We endow this space with $\mathbb{G}$, the augmented filtration of $L$, so it satisfies the usual hypothesis. Denote by $\mu^L$ the measure associated to the jumps of $L$ and $\nu^L$ its compensator, namely $\nu^L(dx) = dt \otimes \nu(dx)$. We now describe the dynamics of the stock prices $S = \{S_t\}_t$ with an *exponential Lévy–model*, meaning that

$$S_t := S_0 \exp(rt + L_t), \quad t \geq 0, \tag{3}$$

with $S_0 \in \mathbb{R}^+$ and $r \in \mathbb{R}^+$ representing the constant annual interest rate. In this way, the discounting process $R = \{R_t\}_t$ reduces to the deterministic function

$$R_t := \exp(-rt), \quad t \geq 0.$$

As already said at the beginning of this chapter, we want $Q$ to be a martingale measure for $S$, therefore we add two hypothesis:

i. there exists $t > 0$ such that $E^Q[\exp(L_t)] < \infty$;

ii. $\Psi(1) = \frac{1}{2}\sigma^2 + \gamma + \int_{\mathbb{R}} (e^x - 1 - x1_D(x))\, \nu(dx) = 0$.

With these assumptions, Example 1.6 ensures that $\{\exp(L_t)\}_t$ is a martingale with $E^Q[\exp(L_t)] = 1$ for every $t > 0$. Taking into account the discounted spot prices process $\widetilde{S} = \left\{\widetilde{S}_t\right\}_t$, defined by

$$\widetilde{S}_t := R_t S_t = S_0 \exp(L_t), \quad t \geq 0,$$

we can state that it is a $Q$ martingale with constant expectation equal to $S_0$. Thus, we have actually modeled the spot prices directly under a martingale measure.

We are going to consider the driving Lévy process $L$ to be a pure jump process, so its generating triplet simplifies considering that $\sigma^2 = 0$. In particular, due to assumption (ii.) we get the next relation:

$$\gamma = -\int_{\mathbb{R}} (e^x - 1 - x1_D(x))\, \nu(dx). \tag{3.8}$$

Let us suppose that in the market there are $N$ European call options with fixed maturity $T$, and denote by $K_j$ and $C^j$ the strike price and observed price of the $j$–th derivative for each $j = 1, ..., N$, respectively. In our model it is convenient to consider the quantities

$$k_j = \log(K_j), \quad j = 1, ..., N.$$

In this setting, $S$ is the price process of the underlying asset. Since $Q$ is a martingale measure for $S$, we can use it to price options as follows:

$$C^j_\nu := \exp(-rT)\, E^Q\left[(S_T - \exp(k_j))^+\right], \quad j = 1, ..., N.$$

Denoting by $Q_{L_T}$ the pushforward distribution on $\mathbb{R}$ generated by $L_T$ we have

$$C^j_\nu = \exp(-rT) \int_{\mathbb{R}} \left(S_0 e^{rT+y} - e^{k_j}\right)^+ Q_{L_T}(dy), \quad j = 1, ..., N.$$

Hence knowing the Lévy measure $\nu$, we can retrieve the whole generating triplet and compute the option prices.

The very reasonable and intuitive idea behind the calibration is to choose $\nu$ such that $C_\nu^j$ is "close" to $C^j$ for any $j = 1, ..., N$. For example, we can choose as Lévy measure of the model the following:

$$\overline{\nu} = \arg\min_{\nu} \left\{ \sum_{j=1}^{N} w_j \left| C_\nu^j - C^j \right|^2 \right\}, \tag{3.9}$$

where the $w_j$–s are factors which represent the relative weights of the contracts, so they reflect the confidence in single data point. In this case, they can be assessed from the bid–ask spread:

$$w_j := \frac{1}{\left| C_{BID}^j - C_{ASK}^j \right|^2}, \quad j = 1, ..., N.$$

Although a solution to (3.9) was found, it could not be unique and the functional $\sum_{j=1}^N w_j \left| C_\nu^j - C^j \right|^2$ could present flat regions, i.e., it is not sensitive to variations in model parameters. Hence the problem expressed in (3.9) is ill posed. We refer to [11], Chapter 13, for an empirical analysis of this situation.

For this reason we want to introduce a regularization term. Suppose that a *historical* (or *prior*), pure jump, driving Lévy process with generating triplet $(0, \nu_0, \gamma_0)$ has been statistically estimated from the time series of the underlying asset price. Let then $L^0 = \left\{ L_t^0 \right\}_t$ be this prior process: it makes sense to require $Q^{L^0}$ and $Q^L$, the distributions on the Skorokhod space $(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}}; \mathbb{F})$ generated by $L^0$ and $L$, respectively, to be equivalent on $\mathcal{F}_t$ for every $t > 0$. It appears clear that the regularization term should penalize those models whose generated probability measure on $(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}})$ is far –in some sense– from $Q^{L^0}$. This reasoning leads us to use the relative entropy as a measure of the diversity from $Q^{L^0}$. In conclusion, since by Remark 3.1 the equality

$$H\left( Q^L\Big|_{\mathcal{F}_t}, Q^{L^0}\Big|_{\mathcal{F}_t} \right) = t \int_{\mathbb{R}} \left( \frac{d\nu}{d\nu^0} \log \frac{d\nu}{d\nu^0} + 1 - \frac{d\nu}{d\nu^0} \right) d\nu^0$$

holds, the problem to be solved becomes:

$$\overline{\nu} = \arg\min_{\nu \in \mathbf{Q}} \left\{ \sum_{j=1}^{N} w_j \left| C_\nu^j - C^j \right|^2 + \alpha \int_{\mathbb{R}} \left( \frac{d\nu}{d\nu^0} \log \frac{d\nu}{d\nu^0} + 1 - \frac{d\nu}{d\nu^0} \right) d\nu^0 \right\}, \tag{3.10}$$

where

$$\mathbf{Q} := \left\{ \nu : Q^L\Big|_{\mathcal{F}_t} \sim Q^{L^0}\Big|_{\mathcal{F}_t}, t > 0 \right\}$$

and $\alpha$ is called *regularization parameter*: the higher $\alpha$ is, the more we trust the initial distribution and the less importance we give to calibration. The existence of a solution to (3.10) has been studied, among others, in [12] and [21] . For the sake of simplicity we are going to relax the assumptions in (3.10), considering the minimization in the set of the Lévy measures $\nu \sim \nu^0$.

### Numerical Approximation

In order to tackle the optimization problem (3.10), we start by fixing a maturity $T$ and the points $x_j \in \mathbb{R}^+$, for $j = 1, ..., N$, with $x_1 < x_2 < ... < x_N$: they represent the grid of the log–strike prices of the options available on the market. We then estimate the parameters of the historical, driving Lévy process $L^0 = \{L_t^0\}_t$, assumed to have a $NIG$ distribution, from the time series of the underlying asset with the *generalized method of moments* (see [15]). At this point, we introduce a discretization grid consisting in the points $y_h$, for $h = 1, ..., N_d$, with $-\infty = y_0 < y_1 < y_2 < ... < y_{N_d}$, which constitutes a partition of the interval $[-M, M]$ for some $M > 0$, and we approximate the Lévy measure of $L^0$ with the discrete version

$$\nu_d^0 (dy) = \sum_{h=1}^{N_d} \nu_h^0 \delta_{(y_h)} (dy) ,$$

where $\delta_{(\overline{a})}$ is, as usual, the Dirac measure at a point $\overline{a}$ and

$$\nu_h^0 = \int_{(y_{h-1}, y_h]} d\nu^0, \; h = 1, ..., N_d - 1; \qquad \nu_{N_d}^0 = \int_{(y_{N_d - 1}, \infty)} d\nu^0.$$

Now we take another measure with the same mass points as the previous one, namely

$$\nu_d (dy) = \sum_{h=1}^{N_d} \nu_h \delta_{(y_h)} (dy) ,$$

with $\nu_h \in \mathbb{R}^+$ for every $h = 1, ..., N_d$. We remark that the calibrated measure $\nu_d$ is equivalent to $\nu_d^0$, but not to the prior $\nu^0$. We interpret these measures $\nu_d$ as the discretized Lévy measures of driving, pure jump Lévy processes $L = \{L_t\}_t$. Aiming to explicitly compute the discrete version of the entropy term in (3.10)

$$\int_{\mathbb{R}} \left( \frac{d\nu_d}{d\nu_d^0} \log \frac{d\nu_d}{d\nu_d^0} + 1 - \frac{d\nu_d}{d\nu_d^0} \right) d\nu_d^0 \tag{3.11}$$

we need to find the Radon–Nikodym derivative $\frac{d\nu_d}{d\nu_d^0}$. So we fix a generic $A \in \mathcal{B}(\mathbb{R})$ and note that the next equalities trivially hold:

$$\nu_d (A) = \sum_{h=1}^{N_d} \nu_h \delta_{(y_h)} (A) = \sum_{h=1}^{N_d} \nu_h 1_A (y_h) = \sum_{h=1}^{N_d} \frac{\nu_h}{\nu_h^0} \nu_h^0 1_A (y_h) . \tag{3.12}$$

On the other hand, define the function

$$f(y) := \begin{cases} \frac{\nu_h}{\nu_h^0}, & \text{if } y_{h-1} < y \le y_h, \quad h = 1, ..., N_d \\ 0, & \text{otherwise} \end{cases} .$$

It results

$$\begin{aligned} \int_{\mathbb{R}} f 1_A \, d\nu_d^0 &= \sum_{h=1}^{N_d} \frac{\nu_h}{\nu_h^0} \int_{(y_{h-1}, y_h]} 1_A (y) \, \nu_d^0 (dy) = \sum_{h=1}^{N_d} \frac{\nu_h}{\nu_h^0} \nu_d^0 ((y_{h-1}, y_h] \cap A) \\ &= \sum_{h=1}^{N_d} \frac{\nu_h}{\nu_h^0} \nu_h^0 1_A (y_h) , \end{aligned} \tag{3.13}$$

using that $\nu_d^0\left((y_{h-1},y_h]\cap A\right)=\nu_h^0 1_A\left(y_h\right)$. Comparing the last terms in (3.12) and (3.13) we can state that $\frac{d\nu_d}{d\nu_d^0}=f$. Moving back to the integral (3.11) we get

$$\begin{aligned}
&\int_{\mathbb{R}}\left(\frac{d\nu_d}{d\nu_d^0}\log\frac{d\nu_d}{d\nu_d^0}+1-\frac{d\nu_d}{d\nu_d^0}\right)d\nu_d^0=\int_{\left(-\infty,y_{N_d}\right]}\left(\frac{d\nu_d}{d\nu_d^0}\log\frac{d\nu_d}{d\nu_d^0}+1-\frac{d\nu_d}{d\nu_d^0}\right)d\nu_d^0\\
&=\left\{\int_{(-\infty,y_1]}+\ldots+\int_{\left(y_{N_d-1},y_{N_d}\right]}\right\}\left(\frac{d\nu_d}{d\nu_d^0}\log\frac{d\nu_d}{d\nu_d^0}+1-\frac{d\nu_d}{d\nu_d^0}\right)d\nu_d^0\\
&=\sum_{h=1}^{N_d}\left[\frac{\nu_h}{\nu_h^0}\log\frac{\nu_h}{\nu_h^0}+1-\frac{\nu_h}{\nu_h^0}\right]\nu_d^0\left((y_{h-1},y_h]\right)\\
&=\sum_{h=1}^{N_d}\left[\nu_h\left(\log\nu_h-\log\nu_h^0\right)+\nu_h^0-\nu_h\right].
\end{aligned}$$

Now we need to express also the quantity $\sum_{j=1}^{N} w_j\left|C_\nu^j-C^j\right|^2$ as a function of $\nu_1,\ldots,\nu_{N_d}$. In order to do so, as suggested in [11], we follow the Carr and Madan approach (see the paper [8], Section 3.2, for further details). This means that we do not approximate directly the option prices $C_\nu(k)$, since a swift application of Lebesgue's convergence theorem shows that $C_\nu$ tends to $S_0$ as $k\to-\infty$, hence it is not integrable in $k$. Instead, we take into account the so–called *modified time value* $z_T$, which is defined by

$$z_T(k):=C_\nu(k)-\left(S_0-e^{k-rT}\right)^+,\quad k\in\mathbb{R}.$$

We are going to assume that $z_T$ and its inverse Fourier transform $\zeta_T$ are integrable, so that by inversion (see, e.g., Theorem 9.11 in [26]) we get

$$z_T(k)=\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}e^{-iku}\zeta_T(u)\,du\quad a.e. \tag{3.14}$$

Fix $k\in\mathbb{R}$; we first analyze the term

$$\begin{aligned}
\left(S_0-e^{k-rT}\right)^+&=\left(S_0-e^{k-rT}\right)1_{\{z\le\log S_0+rT\}}(k)\\
&=e^{-rT}\int_{\mathbb{R}}\left(e^{\log S_0+rT}-e^k\right)1_{\{z\le\log S_0+rT\}}(k)\,Q_{L_T}(dy);
\end{aligned}$$

considering that we similarly get the equality

$$C_\nu(k)=e^{-rT}\int_{\mathbb{R}}\left(e^{y+\log S_0+rT}-e^k\right)1_{\{z\le y+\log S_0+rT\}}(k)\,Q_{L_T}(dy)$$

we can express $z_T(k)$ as follows:

$$\begin{aligned}
z_T(k)=e^{-rT}\int_{\mathbb{R}}\Big[&\left(S_0e^{rT+y}-e^k\right)1_{\{z\ge k-\log S_0-rT\}}(y)\\
&-\left(S_0e^{rT}-e^k\right)1_{\{z\le\log S_0+rT\}}(k)\Big]\,Q_{L_T}(dy).
\end{aligned}$$

Again, recalling the assumption (ii.), according to which

$$\int_{\mathbb{R}}e^y\,Q_{L_T}(dy)=1,$$

we finally obtain

$$z_T(k) = e^{-rT} \int_{\mathbb{R}} \Big[ \left(S_0 e^{rT+y} - e^k\right) \left(1_{\{z \geq k - \log S_0 - rT\}}(y) - 1_{\{z \leq \log S_0 + rT\}}(k)\right) \Big] Q_{L_T}(dy), \quad k \in \mathbb{R}.$$

The idea is to get an estimation of the values $z_T(x_j)$, $j = 1, ..., N$, using the Fourier transform and its inverse. Therefore we fix a point $u \in \mathbb{R}$ and define the inverse Fourier transform

$$\zeta_T(u) := \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{iuk} z_T(k)\, dk.$$

Allowing ourselves to switch the order of integration we have

$$\begin{aligned}
\zeta_T(u) &= \frac{e^{-rT}}{\sqrt{2\pi}} \int_{\mathbb{R}} \left[ \int_{\mathbb{R}} e^{iuk} \left(S_0 e^{rT+y} - e^k\right) \left(1_{\{z \leq y + \log S_0 + rT\}}(k) - 1_{\{z \leq \log S_0 + rT\}}(k)\right) dk \right] Q_{L_T}(dy) \\
&= -\frac{e^{-rT}}{\sqrt{2\pi}} \int_{(-\infty,0]} \left[ \int_{(y+\log S_0 + rT, \log S_0 + rT)} e^{iuk} \left(S_0 e^{rT+y} - e^k\right) dk \right] Q_{L_T}(dy) \\
&\quad + \frac{e^{-rT}}{\sqrt{2\pi}} \int_{(0,\infty)} \left[ \int_{(\log S_0 + rT, y + \log S_0 + rT)} e^{iuk} \left(S_0 e^{rT+y} - e^k\right) dk \right] Q_{L_T}(dy). \qquad (3.15)
\end{aligned}$$

We focus on the computation of the first addend of (3.15). A quick explicit calculation of the inner integral (respect to the Lebesgue measure) of such term, which we indicate by $I_1$, gives

$$\begin{aligned}
I_1(y) &= S_0 \frac{e^{rT+y+iu(\log S_0 + rT)}}{iu(iu+1)} \left[iu\left(1 - e^{-y} + 1 - e^{iuy}\right)\right] \\
&= S_0 \frac{e^{rT+iu(\log S_0 + rT)}}{iu+1}\left(e^y - 1\right) + S_0 \frac{e^{rT+iu(\log S_0 + rT)}}{iu(iu+1)} e^y \\
&\quad - S_0 \frac{e^{iu \log S_0} e^{(iu+1)rT}}{iu(iu+1)} e^{y+iuy}, \quad y \in (-\infty, 0].
\end{aligned}$$

An identical procedure enables us to compute the inner integral $I_2$ of the second addend in (3.15), as well. In particular we arrive at

$$\begin{aligned}
I_2(y) &= -S_0 \frac{e^{rT+iu(\log S_0 + rT)}}{iu+1}\left(e^y - 1\right) - S_0 \frac{e^{rT+iu(\log S_0 + rT)}}{iu(iu+1)} e^y \\
&\quad + S_0 \frac{e^{iu \log S_0} e^{(iu+1)rT}}{iu(iu+1)} e^{y+iuy}, \quad y \in (0, \infty).
\end{aligned}$$

Hence it is possible to express the inverse Fourier transform as

$$\zeta_T(u) = \frac{S_0}{\sqrt{2\pi}} \frac{e^{iu(\log S_0 + rT)}}{iu+1} \left[ \int_{\mathbb{R}} (1 - e^y)\, Q_{L_T}(dy) - \frac{1}{iu} \int_{\mathbb{R}} e^y\, Q_{L_T}(dy) + \frac{1}{iu} \int_{\mathbb{R}} e^{y+iuy}\, Q_{L_T}(dy) \right].$$

Thanks to assumption (i.) and (1.12) we conclude that

$$\zeta_T(u) = \frac{S_0}{\sqrt{2\pi}} \frac{e^{iu(\log S_0 + rT)}}{iu(iu+1)} \left[ e^{T\Psi(iu+1)} - 1 \right], \quad u \in \mathbb{R}. \tag{3.16}$$

Recalling the definition of $\Psi$ in (1.11), we substitute the discrete version $\nu_d$ for the Lévy measure $\nu$ associated to $L$. As a consequence of (3.8) we obtain

$$\begin{aligned}\Psi(iu+1) &\simeq \int_{\mathbb{R}} \left( e^{(iu+1)y} - iue^y - e^y + iu \right) \nu_d(dy) \\ &= \int_{\mathbb{R}} e^y \left( e^{iuy} - 1 \right) \nu_d(dy) + iu \int_{\mathbb{R}} (1 - e^y) \nu_d(dy) \\ &= \sum_{h=1}^{N_d} e^{y_h} \left( e^{iuy_h} - 1 \right) \nu_h + iu \sum_{h=1}^{N_d} (1 - e^{y_h}) \nu_h, \quad u \in \mathbb{R}.\end{aligned}$$

Plugging this term into (3.16) we eventually end up with the next approximation:

$$\begin{aligned}\zeta_T(u) \simeq \frac{S_0}{\sqrt{2\pi}} \frac{e^{iu(\log S_0 + rT)}}{iu(iu+1)} \Bigg[ \exp \Bigg( T \sum_{h=1}^{N_d} e^{y_h} \left( e^{iuy_h} - 1 \right) \nu_h \\ + iuT \sum_{h=1}^{N_d} (1 - e^{y_h}) \nu_h \Bigg) - 1 \Bigg], \quad u \in \mathbb{R}. \end{aligned} \tag{3.17}$$

Let us estimate the modified time value $z_T$ at the points $x_j$, for $j = 1, ..., N$. In order to compute the integral in (3.14), we decide to construct another grid, this time uniform with mesh $d > 0$, which contains the points of the previous one of the log–strike prices. Specifically, chosen $\widetilde{N} \in \mathbb{N}$, we introduce

$$\widetilde{x}_n := \frac{2\pi n}{\widetilde{N}\Delta}, \quad n = -\widetilde{N}, ..., -1, 0, 1, ..., \widetilde{N},$$

where $A := \frac{2\pi}{d}$ is the size of the discretization interval while $\Delta := \frac{A}{\widetilde{N}}$. The construction is carried out so that for every $j = 1, ..., N$ and $h = 1, ..., N_d$ there exists a $n_{hj} \in \left\{ -\widetilde{N} + 1, ..., \widetilde{N} - 1 \right\}$ such that $x_j - y_h = \widetilde{x}_{n_{hj}}$. Finally we define the points of the discretization grid as

$$u_k := -\frac{A}{2} + k\Delta, \quad k = 0, ..., \widetilde{N}.$$

We then compute

$$\begin{aligned} z_T(\widetilde{x}_n) &\simeq \frac{1}{\sqrt{2\pi}} \int_{(-A/2, A/2)} e^{-iu\widetilde{x}_n} \zeta_T(u)\, du \simeq \frac{1}{\sqrt{2\pi}} \frac{A}{\widetilde{N}} \sum_{k=0}^{\widetilde{N}-1} e^{-iu_k \widetilde{x}_n} \widetilde{w}_k \zeta_T(u_k) \\ &= \frac{1}{\sqrt{2\pi}} \frac{A}{\widetilde{N}} e^{i\frac{A}{2}\widetilde{x}_n} \sum_{k=0}^{\widetilde{N}-1} \exp\left( -i\frac{2\pi n}{\widetilde{N}} k \right) \widetilde{w}_k \zeta_T(u_k), \quad n = 0, ..., \widetilde{N} - 1, \end{aligned}$$

where $\widetilde{w}_k$ are coefficients chosen according to the trapezoidal rule as

$$\widetilde{w}_k := \begin{cases} \frac{1}{2}, & \text{if } k = 0, \widetilde{N} - 1 \\ 1, & \text{if } k = 1, ..., \widetilde{N} - 2 \end{cases}.$$

Hence it is possible to get the values $z_T(\widetilde{x}_n)$, for $n = 0, ..., \widetilde{N}-1$, with a *fast Fourier transform* (FFT) once known $\zeta_T(u_k)$ for $k = 0, ..., \widetilde{N}-1$. Due to the symmetry of the grid, an analogous reasoning leads to retrieve $z_T(\widetilde{x}_n)$ for $n = -\widetilde{N}+1, ..., -1$, recalling to use an inverse discrete Fourier transform. In conclusion, Problem (3.10) translates into the minimization in $(\mathbb{R}^+)^d$ of the next objective functional:

$$\mathcal{F}(\nu_1, ..., \nu_{N_d}) = \sum_{j=1}^{N} w_j \left| z_t(x_j) + \left(S_0 - e^{x_j - rT}\right)^+ - C^j \right|^2 + \alpha \sum_{h=1}^{N_d} \left[\nu_h \left(\log \nu_h - \log \nu_h^0\right) + \nu_h^0 - \nu_h\right].$$

In order to speed up the optimization (we used the *L-BFGS-B* algorithm), we compute the derivatives of $\mathcal{F}$. The entropy term is easy to derive, hence we focus on the other one. To facilitate the notation, denote by

$$C_u := \frac{S_0}{\sqrt{2\pi}} \frac{e^{iu(\log S_0 + rT)}}{iu(iu+1)}, \quad u \in \mathbb{R}$$

and by $g(\nu_1, ..., \nu_{N_d})$ the argument of the exponential in (3.17). Under suitable integrability conditions, for every $u \in \mathbb{R}$ and $h = 1, ..., N_d$ we get

$$\begin{aligned} \frac{\partial \zeta_T(u)}{\partial \nu_h}(\nu_1, ..., \nu_{N_d}) &= C_u T e^{g(\nu_1, ..., \nu_{N_d})} \left[e^{y_h}\left(e^{iuy_h} - 1\right) + iu\left(1 - e^{y_h}\right)\right] \\ &= \frac{S_0}{\sqrt{2\pi}} \frac{T}{iu+1} e^{iu(\log S_0 + rT)} \left(1 - e^{y_h}\right) e^{g(\nu_1, ..., \nu_{N_d})} \\ &\quad + T e^{y_h} \zeta_T(u) \left(e^{iuy_h} - 1\right) + T e^{y_h} C_u \left(e^{iuy_h} - 1\right). \end{aligned}$$

From (3.14), by explicit calculation for every $k \in \mathbb{R}$ and $h = 1, ..., N_d$ we obtain

$$\begin{aligned} \frac{\partial z_T(k)}{\partial \nu_h}(\nu_1, ..., \nu_{N_d}) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iuk} \frac{\partial \zeta_T(u)}{\partial \nu_h}(\nu_1, ..., \nu_{N_d})\, du \\ &= \frac{S_0}{2\pi} T\left(1 - e^{y_h}\right) \int_{\mathbb{R}} \frac{e^{-iuk}}{iu+1} e^{iu(\log S_0 + rT)} e^{g(\nu_1, ..., \nu_{N_d})} du \\ &\quad + T e^{y_h}\Big[z_T(k - y_h) + \left(S_0 - e^{k - rT - y_h}\right)^+ - z_T(k) - \left(S_0 - e^{k - rT}\right)^+\Big]. \end{aligned} \tag{3.18}$$

As already done with $z_T$, we can approximate the first addend in (3.18) when $k = \widetilde{x}_n$, for $n = 0, ..., \widetilde{N}-1$, with a FFT:

$$\begin{aligned} \int_{\mathbb{R}} \frac{e^{-iu\widetilde{x}_n}}{iu+1} e^{iu(rT + \log S_0)} e^{g(\nu_1, ..., \nu_{N_d})} &\simeq \frac{A}{\widetilde{N}} \sum_{k=0}^{\widetilde{N}-1} e^{-iu_k \widetilde{x}_n} \widetilde{w}_k f(u_k) \\ &= \frac{A}{\widetilde{N}} e^{i\frac{A}{2}\widetilde{x}_n} \sum_{k=0}^{\widetilde{N}-1} \exp\left(-i\frac{2\pi n}{\widetilde{N}} k\right) \widetilde{w}_k f(u_k), \end{aligned}$$

where

$$f(u) := \frac{\sqrt{2\pi}}{S_0} iu\zeta_T(u) + \frac{e^{iu(\log S_0 + rT)}}{iu+1}, \quad u \in \mathbb{R}.$$

Putting together all the terms we have laboriously gotten we can numerically calculate the derivatives of the objective functional $\mathcal{F}$, which are provided by

$$\frac{\partial \mathcal{F}}{\partial \nu_h}(\nu_1, ..., \nu_{N_d}) = 2\sum_{j=1}^{N} w_j \Big[ z_T(x_j) + \big(S_0 - e^{x_j - rT}\big)^+ - C^j \Big] \frac{\partial z_T(x_j)}{\partial \nu_h}(\nu_1, ..., \nu_{N_d}) + \alpha \left(\log \nu_h - \log \nu_h^0\right)$$

for any $h = 1, ..., N_d$.

In our simulation we want to give the same importance to all option prices available in the market, meaning that we decide not to mirror the customers' preferences. This translates into assigning the same weight at each data point, explicitly

$$w_j = \frac{1}{N}, \quad j = 1, ..., N.$$

Finally the regularization parameter $\alpha$ must be picked. In order to do so, we interpret it as a proxy of the market error. Therefore, in a first moment we minimize the quadratic pricing error (3.9) without entropy term. Denoting by $\epsilon_0$ the value of the functional at the found minimum, such value can be thought to as a measure of the distance between the market and the selected model class. It is then satisfactory to take

$$\alpha := \epsilon_0.$$

Besides, we remark that, using the calibration procedure, a single simulation provides the whole option chain corresponding to the desired strike–prices at a fixed maturity $T$.

### Empirical Results

We have empirically tested this method with the prices of call options on *Alphabet Inc.* class C stock (ticker symbol: GOOG, no voting rights) expiring in Jan, 2020: 11 months from the time of simulation. Figure 3.1 shows the results of such implementation. We refer to Appendix A, Section A.2, for the used code. The average of the absolute values of percentage difference is 1.7360%.

We also tried to calibrate our method to the prices of call options on *Amazon.com, Inc.* stock (ticker symbol: AMZN): Figure 3.2 below displays the outcomes. In this case, the mean of absolute values of percentage difference settles down at 2.2827%.

**Remark 3.2.** We retrieved the real prices from the website `https://www.nasdaq.com`, where American options are traded. We have instead simulated the prices assuming that they were European. It is well known that an American option should be worth slightly more than its European analogon, since it gives the buyer more privilege. In our case, however, considering that *Alphabet Inc.* and *Amazon.com, Inc.* do not pay dividends and that the time values of the options were always positive, it makes sense to get an approximation of the prices of such derivatives under an European model. In fact, if it is convenient to sell an option rather than exercising the call before the maturity, why would someone be willing to pay more for a right they are never going to use?

Figure 3.1: GOOG call option, Jan 2020.

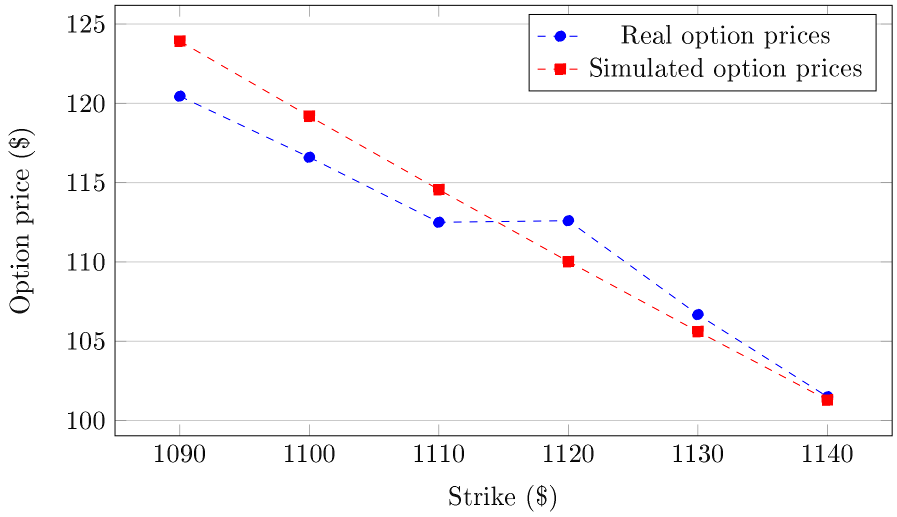
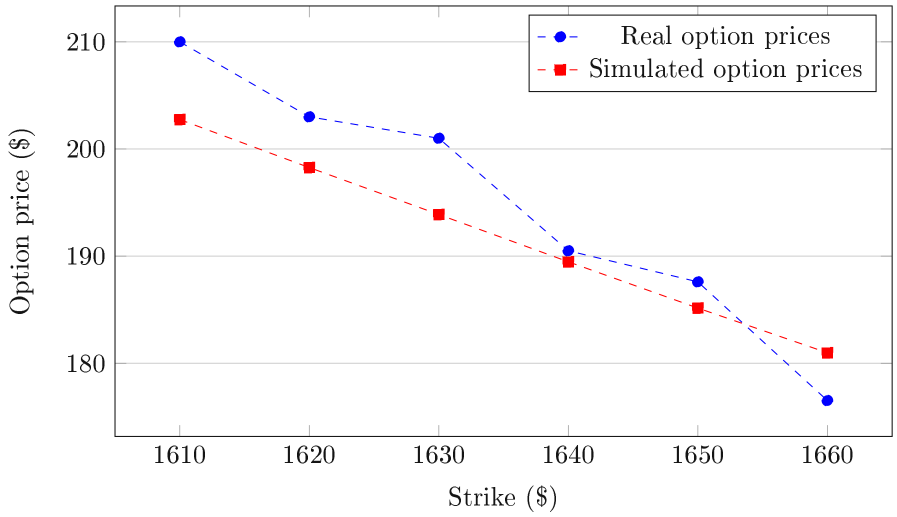



Figure 3.2: AMZN call option, Jan 2020.



# Conclusion

This research focuses on two option pricing methods: one based on the Esscher measure and one on the calibration to real–traded call option prices. Its main purpose is to quantify their performances in a liquid market like NASDAQ.

The empirical results for the Esscher method suggest that it is particularly reliable while dealing with options on stocks whose prices are in the order of hundreds of dollars. From a computational point of view, it is faster then the calibration method, especially when we need to simulate the prices of options with different maturities and same strike.

On the other hand, the calibration method allows to generate the entire option chain for a fixed maturity and it is more stable than the previous one, in the sense that it generates results close to real data for a larger range of stock prices.

Comparing the averages of absolute values of percentage difference between real and simulated prices, the calibration procedure offers better outcomes than Esscher's. Nevertheless, the Esscher method only needs the historical time series of the underlying asset price to be applied, so it is feasible also in markets with few derivatives. This study shows that the Esscher measure is a good solution to price call options in illiquid –or low liquid– financial markets.

# Bibliography


[1] Abramowitz, M., Stegun, I.A., *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables*, Dover Publications, 1972.

[2] Arfken, G.B., Weber, H.J., *Mathematical Methods for Physicists*, Elsevier Academic Press, 6th ed., 2005.

[3] Asmussen, S., Rosiński, J., *Approximations of Small Jumps of Lévy Processes with a View towards Simulation*, Journal of Applied Probability **38(2)**(2001), Applied Probability Trust, 482–493.

[4] Barndorff-Nielsen, O.E., *Processes of Normal Inverse Gaussian Type*, Finance and Stochastics **2** (1998), Springer–Verlag, 41–68.

[5] Barndorff-Nielsen, O.E., Halgreen, C., *Infinite Divisibility of the Hyperbolic and Generalized Inverse Gaussian Distributions*, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete **38(4)** (1977), Springer–Verlag, 309–312.

[6] Brockwell, P., Chadraa, E., Lindner, A., *Continuous-Time GARCH Processes*, The Annals of Applied Probability **16(2)** (2006), Institute of Mathematical Statistics, 790–826.

[7] Carmona, R., Ludkovski, M., *Spot Convenience Yield Models for Energy Markets*, Contemporary Mathematics **351** (2004), AMS, 65–80.

[8] Carr, P., Madan, D.B., *Option Valuation using the Fast Fourier Transform*, Journal of Computational Finance **2(4)** (1999), 61–73.

[9] Chen, D., *Three essays on pricing and hedging in incomplete markets*, PhD thesis, The London School of Economics and Political Science, 2011.

[10] Chen, D., Rheinländer, T., *Valuation of Metal Futures: a Study in Incomplete Markets*, London School of Economics, 2011.

[11] Cont, R., Tankov, P., *Financial Modelling with Jump Processes*, Financial Mathematics Series, Chapman and Hall/CRC, 2004.

[12] Cont, R., Tankov, P., *Nonparametric calibration of jump-diffusion option pricing models*, Journal of Computational Finance **7(3)** (2004), Incisive Media Ltd., 1–49.

[13] Eberlein, E., v. Hammerstein, E.A., *Generalized Hyperbolic and Inverse Gaussian Distributions: Limiting Cases and Approximation of Processes*, Progress in Probability **58** (2004), Birkhäuser–Verlag, 221–264.

[14] Grith, M., Härdle, W.K., Schienle, M., *Nonparametric Estimation of Risk–Neutral Densities*, SFB 649 Discussion Paper 2010–021, Humboldt-Universität zu Berlin, 2010.

[15] Hansen, L.P., *Large Sample Properties of Generalized Method Of Moments Estimators*, Econometrica **50(4)** (1982), The Econometric Society, 1029–1054.

[16] Iacus, S.M., *Option Pricing and Estimation of Financial Models with R*, John Wiley & Sons, 2011.

[17] Iacus, S.M., Mercuri, L., Rroji, E., *COGARCH(p, q): Simulation and Inference with the yuima Package*, Journal of Statistical Software **80(4)** (2017), Foundation for Open Access Statistics.

[18] Ihara, S., *Information Theory for Continuous Systems*, World Scientific, 1993.

[19] Jeanblanc, M., Yor, M., Chesney, M., *Mathematical Methods for Financial Markets*, Springer Finance Textbooks, Springer–Verlag, 2009.

[20] Kallsen, J., Shiryaev, A., *The Cumulant Process and Esscher's Change of Measure*, Finance and Stochastics **6** (2002), Springer–Verlag, 397–428.

[21] Keller–Ressel, M., *Affine Processes and Applications to Stochastic Volatility Modelling. Presentation Slides.*, Weierstrass Institute, Berlin, 2008.

[22] Klüppelberg, C., Lindner, A., Maller, R., *A Continuous Time GARCH Process Driven by a Lévy Process: Stationarity and Second Order Behaviour*, Journal of Applied Probability **41(3)** (2004), Applied Probability Trust, 601–622.

[23] Lewis, P.A.W., Shedler, G.S., *Simulation of Nonhomogeneous Poisson Processes by Thinning*, Naval research logistics quarterly **26(3)** (1979), 403–413.

[24] Protter, P.E., *Stochastic Integration and Differential Equations*, Stochastic Modelling and Applied Probability **21**, Springer–Verlag, 2nd ed., 2005.

[25] Rheinländer, T., Sexton, J., *Hedging Derivatives*, Advanced Series on Statistical Science & Applied Probability **15**, World Scientific, 2011.

[26] Rudin,W., *Real and Complex Analysis*, Mathematics Series, McGraw-Hill Higher Education , 3rd ed., 1987.

[27] Sato, K.-I., *Lévy Processes and Infinitely Divisible Distributions*, Cambridge studies in advanced mathematics **68**, Cambridge University Press, 1999.

[28] SHIRYAEV, A., JACOD, J., *Limit Theorems for Stochastic Processes*, A Series of Comprehensive Studies in Mathematics **288**, Springer–Verlag, 2003.

[29] SHREVE, S.E., *Stochastic Calculus for Finance II (Continuous–Time Models)*, Springer Finance Textbooks, Springer–Verlag, 2004.

[30] WATSON, G.N., *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, 2nd ed., 1966.

# Appendices

# Appendix A

# Simulation Codes

In this appendix we just present the codes we wrote to run the simulations of Chapters 2 and 3. We used the language *R*.

## A.1 Esscher Measure Code

The succeeding code refers to the simulation of AAPL option prices.

```
library(yuima)
require(yuima)
setwd("/home/alessandrobondi/Downloads/R␣warmup")
AAPL= read.csv("AAPLHistorical.csv", header = TRUE)
logdayprice <- AAPL$log.Prices
logdayprice=logdayprice[length(logdayprice):1]
dataI <- setData(as.matrix(logdayprice))
##initial data
NIGparam <- list(a1 = 0.038, b1 =0.0558, a0 = 2.1662e-06/0.0558, alpha =
    1.7182171, beta = -0.3124525, delta=1.6336979, mu = 0.3021202, y01 =
     0)
NIGmodel<- setCogarch(p = 1, q = 1,measure = list(df = "rNIG(z,alpha,beta
    ,delta,mu)"), measure.type = "code", XinExpr = TRUE)
##define the thinning algorithm
NHPP.sim <- function(lambda,t_max){
        s<-0
        t<-0
        interarrival<-numeric()
        lambda_star <- max(lambda[1:t_max+1])
        while(s<= t_max){
                u <- runif(1)
                s <- s - log(u)/lambda_star
                if (s<=t_max){
                        if(runif(1) < lambda[s+1]/lambda_star) {
                                t<-s
                                interarrival<- c(interarrival,t)
                        }
                }
        }
        if (t<=t_max) {
```

```
                return(interarrival)
        }
        else {
                lungo=length(interarrival)
                return(interarrival[1:lungo-1])
        }
}
NIGYuima <- setYuima(data = dataI, model = NIGmodel)
resNIG<- gmm(yuima = NIGYuima, start = NIGparam, Est.Incr = "IncrPar")
fin=data.frame(resNIG@coef)
r<-0.005/length(logdayprice)
epsilon=.01
##the vector "tempi" contains the maturities of the considered call
    options
tempi=c(23,40,88,106,231)
alpha<- fin[4,1]
beta <- fin[5,1]
delta<-fin[6,1]
mu <- fin[7,1]
estparam.NIG=list(a1=fin[2,1],b1=fin[1,1],a0=fin[3,1])
n_iter<-100
x<-list()
ord_def=list()
Term <- (length(logdayprice)-1)
num <- length(logdayprice)
set.seed(123)
gamma_NIG=mu+(2*alpha*delta/pi)*integrate(function(x) {sinh(beta*x)*
    besselK(alpha*x,1,expon.scaled = FALSE)},lower = 0,upper=1)$value
gamma_NIG_h=gamma_NIG-integrate(function(x) {x*delta*alpha/(pi*abs(x))*
    exp(beta*x)*besselK(alpha*abs(x),1,expon.scaled = FALSE)},lower =
    epsilon,upper=1)$value-integrate(function(x) {x*delta*alpha/(pi*abs(
    x))*exp(beta*x)*besselK(alpha*abs(x),1,expon.scaled = FALSE)},lower
    = -1,upper=-epsilon)$value
##start the iteration
for (j in 1:n_iter)
{
##simulate the COGARCH process with the estimated parameters
cog.NIG <- setCogarch(p = 1, q = 1, work = FALSE, measure = list(df="rNIG
    (z,0.924838508244,-0.03978538525997,0.961693831661,0.117699889207)")
    , measure.type = "code", Cogarch.var = "G", V.var = "V", Latent.var
    = "Y")
samp.NIG <- setSampling(Terminal = Term, n = num)
sim.NIG <- simulate(cog.NIG, true.parameter = estparam.NIG, sampling =
    samp.NIG,method = "euler")
sigma2 <- vector(mode="numeric", length=length(logdayprice))
##generate a path of the variance process
for (i in 1:length(logdayprice)) {
        sigma2[i]<-sim.NIG@data@original.data[i,2]
}
R=vector(mode="numeric", length=length(logdayprice))
for (t in 1:length(logdayprice)) {
        R[t]=-r+mu*sqrt(sigma2[t])
}
theta <- vector(mode="numeric", length=length(logdayprice))
```

```
##generate a path of the process theta
for (t in 1:length(logdayprice)) {
        theta[t]<-1/((2*(R[t]^2+sigma2[t]*delta^2))*sqrt(sigma2[t])*delta
             )*(-delta^3*sqrt(sigma2[t])^3-2*delta^3*beta*sigma2[t]-delta*
             sqrt(sigma2[t])*R[t]^2-2*delta*beta*R[t]^2-sqrt(-delta^4*
             sigma2[t]^2*R[t]^2-2*delta^2*sigma2[t]*R[t]^4+4*R[t]^4*delta
             ^2*alpha^2-R[t]^6+4*sigma2[t]*delta^4*R[t]^2*alpha^2))
}
##introduce the term for the variation of the small jumps
variationsmalljumps=vector(mode="numeric", length=length(logdayprice))
for (t in 1:length(logdayprice)) {
        variationsmalljumps[t]<- integrate(function(x) {sigma2[t]*x^2*
             delta*alpha/(pi*abs(x))*exp((beta+theta[t]*sqrt(sigma2[t]))*x
             )*besselK(alpha*abs(x),1,expon.scaled = FALSE)},lower = -
             epsilon/sqrt(sigma2[t]),upper=-0.00001)$value+integrate(
             function(x) {sigma2[t]*x^2*delta*alpha/(pi*abs(x))*exp((beta+
             theta[t]*sqrt(sigma2[t]))*x)*besselK(alpha*abs(x),1,expon.
             scaled = FALSE)},lower = 0.00001,upper=epsilon/sqrt(sigma2[t
             ]))$value
}
require(sde)
bm=BM(x=0,t0=0,T=(length(logdayprice)-1),N=length(logdayprice)-1)
addendo1=vector(mode="numeric", length=length(logdayprice))
for (t in 1:length(logdayprice)) {
        addendo1[t]= sqrt(variationsmalljumps[t])*(bm[t])
}
##simulate a trajectory of the "Poisson" component
##compute the intensity of the nonhomogeneous Poisson process
lambda=numeric()
for (t in 1:length(logdayprice)) {
        lambda[t]<-integrate(function(x) {delta*alpha/(pi*abs(x))*exp((
             beta+theta[t]*sqrt(sigma2[t]))*x)*besselK(alpha*abs(x),1,
             expon.scaled = FALSE)},lower = -Inf,upper=-epsilon*1/sqrt(
             sigma2[t]))$value+integrate(function(x) {delta*alpha/(pi*abs(
             x))*exp((beta+theta[t]*sqrt(sigma2[t]))*x)*besselK(alpha*abs(
             x),1,expon.scaled = FALSE)},lower = 1/sqrt(sigma2[t])*epsilon
             ,upper=Inf)$value
}
##apply the thinning algorithm to get the jumps time
run1=NHPP.sim(lambda,(length(logdayprice)-1))
##compute the time-varying jumps sizes
jumps=numeric()
for (t in 1:length(run1)) {
        jumps<-c(jumps,(integrate(function(x) {sqrt(sigma2[run1[t]+1])*x*
             delta*alpha/(pi*abs(x))*exp((beta+theta[run1[t]+1]*sqrt(
             sigma2[run1[t]+1]))*x)*besselK(alpha*abs(x),1,expon.scaled =
             FALSE)},lower = -Inf,upper=-epsilon/sqrt(sigma2[run1[t]+1]))$
             value+integrate(function(x) {sqrt(sigma2[run1[t]+1])*x*delta*
             alpha/(pi*abs(x))*exp((beta+theta[run1[t]+1]*sqrt(sigma2[run1
             [t]+1]))*x)*besselK(alpha*abs(x),1,expon.scaled = FALSE)},
             lower =epsilon/sqrt(sigma2[run1[t]+1]),upper=Inf)$value))
}
##put together, preserving the order, the jump times with the analyzed
    days
```

```
unite<-c(0:(length(logdayprice)-1),run1)
united=unite[!duplicated(unite)]
ord=sort(united,decreasing = FALSE)
c=c(which(ord==tempi[1]),which(ord==tempi[2]),which(ord==tempi[3]),which(
    ord==tempi[4]),which(ord==tempi[5]))
ord_def[[j]]=c
##Build B
b_theta=vector(mode="numeric", length=length(theta))
for (t in 1:length(theta)) {
       b_theta[t]=sqrt(sigma2[t])*gamma_NIG_h-integrate(function(y) {
           sqrt(sigma2[t])*y*delta*alpha/(pi*abs (y))*exp(beta*y)*
           besselK(alpha*abs(y),1,expon.scaled = FALSE)},lower=-epsilon,
           upper=-0.00001)$value-integrate(function(y) {sqrt(sigma2[t])*
           y*delta*alpha/(pi*abs(y))*exp(beta*y)*besselK(alpha*abs(y),1,
           expon.scaled = FALSE)},lower=0.00001,upper=epsilon)$value +
           integrate(function(y) {sqrt(sigma2[t])*y*exp(theta[t]*sqrt(
           sigma2[t])*y)*delta*alpha/(pi*abs(y))*exp(beta*y)*besselK(
           alpha*abs(y),1,expon.scaled = FALSE)},lower=0.00001,upper=
           epsilon/sqrt(sigma2[t]))$value+integrate(function(y) {sqrt(
           sigma2[t])*y*exp(theta[t]*sqrt(sigma2[t])*y)*delta*alpha/(pi*
           abs (y))*exp(beta*y)*besselK(alpha*abs(y),1,expon.scaled =
           FALSE)},lower=-epsilon/sqrt(sigma2[t]),upper=-0.00001)$value
}
##generate the risk-neutral path
B=cumsum(b_theta)
incr_brown=c(0,diff(addendo1,1))
g=vector(mode="numeric", length=length(ord))
g[1]=0
for (i in 2:length(ord)) {
       if (is.element(ord[i],0:(length(logdayprice)-1))) {
              g[i]=g[i-1]+incr_brown[ord[i]+1]
       }
       else {
              g[i]=g[i-1]+jumps[which(run1==ord[i])]
       }
}
for (i in 2:length(ord)) {
       g[i]=g[i]+B[ord[i]+1]
}
x[[j]]<-g
}
##simulate the call option prices
strike=165
predicted_prices=list()
for (i in 1:n_iter) {
       c=numeric()
       for (t in 1:5) {
              c=c(c,170.5*exp(x[[i]][ord_def[[i]][t]]))
       }
       predicted_prices[[i]]=c
}
TTM.call=list()
for (i in 1:n_iter) {
       c=numeric()
```

```
        for (t in 1:5) {
                c=c(c,exp(-r*tempi[t])*pmax(0,predicted_prices[[i]][t]-
                    strike))
        }
        TTM.call[[i]]=c
}
aux=list()
for (t in 1:5) {
        c=numeric()
        for (i in 1:n_iter) {
                c=c(c,TTM.call[[i]][t])
        }
        aux[[t]]=c
}
predicted_pricesOPTIONS=numeric()
for (t in 1:5) {
        predicted_pricesOPTIONS[t]=mean(aux[[t]])
}
predicted_pricesOPTIONS
C=c(9.20,11.00,14.40,15.70,21.20)
mean((abs(predicted_pricesOPTIONS-C))/C)
```

## A.2 Calibration Code

The following code refers to the simulation of GOOG option prices.

```
## construction of the objective functional to be minimized
d=0.001
A=2*pi/d
N_tilde=7451
N=6
Strikes=numeric(N)
for (j in 1:N) {
        Strikes[j]=1080+j*10
}
N_discretization_measure=8
r=0.008
T=11/12
C=c(120.45,116.6,112.5,112.6,106.68,101.5)
W=c(1,1,1,1,1,1)
W=1/N*W
S_0=1113.8
Delta<-A/N_tilde
## build the points of the bigger grid x
x=vector(mode="numeric",length = N_tilde+1)
for (n in 1:(N_tilde+1)) {
        x[n]=d*(n-1)
}
## build the discretization points of the Levy Measure
dis=vector(mode="numeric")
dis[1]=0.01
for (n in 2:(N_discretization_measure/2)) {
        dis[n]=x[min(which(x>0.03+dis[n-1]))]
```

```
}
dis=sort(c(-dis,dis))
dis
## define the discretization grid of points u_k
u=vector(mode="numeric",length = N_tilde+1)
for (k in 1:(N_tilde+1)) {
        u[k]=-A/2+(k-1)*Delta
}
##build the weights w according to the trapezoidal rule
w_tilde=vector(mode="numeric",length = N_tilde)
w_tilde[1]=1/2
w_tilde[N_tilde]=1/2
for (k in 2:(N_tilde-1)) {
        w_tilde[k]=1
}
## compute Zeta(u_k) for k=0,...,N_tilde.
## specify the smaller grid of log-strikes
x_restricted= vector(mode="numeric",length = N)
for (j in 1:N) {
        c=numeric(N_tilde+1)
        for (n in 1:(N_tilde+1)) {
                c[n]=abs(x[n]-log(Strikes)[j])
        }
        x_restricted[j]=x[which(c==min(c))]
}
exp(x_restricted)
i=sqrt(-1+0i)
## set the functions Zeta & zeta
Zeta=list()
for (k in 1:N_tilde) {
        Zeta[[k]] <- function(nu,k) {
                1/(sqrt(2*pi))*S_0*exp(i*u[k]*r*T+i*u[k]*log(S_0))/(i*u[k]
                    *(1+i*u[k]))*(exp(T*(i*u[k]*sum((1-exp(dis))*nu)+sum(
                    exp(dis)*(exp(i*u[k]*dis)-1)*nu)))-1)
        }
}
z_aux=list()
z_aux<-function(nu) {
        c=vector()
        for (k in 1:N_tilde) {
                c=c(c,w_tilde[k]*Zeta[[k]](nu,k))
        }
        c
}
g_pos<-function(nu) {fft(z_aux(nu))}
z=list()
for (n in 1:N_tilde) {
        z[[n]]<-function(nu,n){
                Re(1/(sqrt(2*pi))*A/N_tilde*exp(i*A/2*x[n])*g_pos(nu)[n])
        }
}
g_neg<-function(nu) {fft(z_aux(nu),inverse=TRUE)}
z_neg=list()
for (n in 1:N_tilde ) {
```

```
        z_neg[[n]]<-function(nu,n){
                Re(1/(sqrt(2*pi))*A/N_tilde*exp(-i*A/2*x[n])*g_neg(nu)[n])
        }
}
## evaluation of z in the grid of log(strikes)
z_restricted=list()
for (j in 1:N) {
        z_restricted[[j]]<-function(nu,j){
                z[[which(x==x_restricted[j])]](nu,which(x==x_restricted[j
                    ]))
        }
}
##Entropy term
alpha=1.39824945059989
Beta=-0.265735585367672
delta=1.46729435279966
mu=0.295853383567856
##determine the discretized values of the prior nu0[j]
nu0=numeric(length=N_discretization_measure)
nu0[1]=integrate(function(x){delta*alpha/(pi*abs(x))*exp(Beta*x)*besselK(
    alpha*abs(x),1,expon.scaled = FALSE)},lower = -Inf,upper=dis[1])$
    value
for (h in 2:(N_discretization_measure-1)) {
        if (h==5) {
                nu0[h]=integrate(function(x) {delta*alpha/(pi*abs(x))*exp(
                    Beta*x)*besselK(alpha*abs(x),1,expon.scaled = FALSE)},
                    lower = dis[h-1],upper=-.005)$value+integrate(function
                    (x) {delta*alpha/(pi*abs(x))*exp(Beta*x)*besselK(alpha
                    *abs(x),1,expon.scaled = FALSE)},lower = 0.005,upper=
                    dis[h])$value
        }
        else {
                nu0[h]=integrate(function(x) {delta*alpha/(pi*abs(x))*exp(
                    Beta*x)*besselK(alpha*abs(x),1,expon.scaled = FALSE)},
                    lower = dis[h-1],upper=dis[h])$value
        }
}
nu0[N_discretization_measure]=integrate(function(x){delta*alpha/(pi*abs(x
    ))*exp(Beta*x)*besselK(alpha*abs(x),1,expon.scaled = FALSE)},lower =
     dis[N_discretization_measure-1],upper=Inf)$value
nu0
##determine the entropy term (expressed by a discretized integral)
entropy_term<-function(nu) {
        c=vector()
        for (h in 1:N_discretization_measure){
                c=c(c,nu[h]*(log(nu[h]/nu0[h]))+nu0[h]-nu[h])
        }
        sum(c)
}
## this is the searched objective functional
addendo1_objective=list()
addendo1_objective<-function(nu) {
        c=vector()
        for (j in 1:N) {
```

```
                c=c(c,W[j]*(abs(z_restricted[[j]](nu,j)+pmax(0,S_0-exp(x_
                    restricted[j]-r*T))-C[j]))^2)
        }
        sum(c)
}
objective_functional<-function(nu) {addendo1_objective(nu)}

##computation of the gradient for gradient descent method
##entropy term
grad_entropy_term<-function(nu) {
        c=vector()
        for (h in 1:N_discretization_measure){
                c=c(c,log(nu[h]/nu0[h]))
        }
        c
}
## focus on the difficult gradient
addendo1_aux=list()
addendo1_aux<-function(nu,h) {
        c=vector()
        for (j in 1:N) {
                if (x_restricted[j]>=dis[h]) {
                        if (dis[h]>0) {
                        c=c(c,W[j]*(z_restricted[[j]](nu,j)+max(0,S_0-exp(x
                            _restricted[j]-r*T))-C[j])*(z[[which(x==x_
                            restricted[j])-which(x==dis[h])+1]](nu,which(x
                            ==x_restricted[j])-which(x==dis[h])+1)-z_
                            restricted[[j]](nu,j)))
                        }
                        else {
                        c=c(c,W[j]*(z_restricted[[j]](nu,j)+max(0,S_0-exp(x
                            _restricted[j]-r*T))-C[j])*(z[[which(x==x_
                            restricted[j])+which(x==abs(dis[h]))-1]](nu,
                            which(x==x_restricted[j])+which(x==abs(dis[h]))
                            -1)-z_restricted[[j]](nu,j)))
                        }
                }
                else {
                        c=c(c,W[j]*(z_restricted[[j]](nu,j)+max(0,S_0-exp(x
                            _restricted[j]-r*T))-C[j])*(z_neg[[which(x==dis
                            [h])-which(x==x_restricted[j])+1]](nu,which(x==
                            dis[h])-which(x==x_restricted[j])+1)-z_
                            restricted[[j]](nu,j)))
                }
        }
        sum(c)
}
addendo1=list()
for (h in 1:N_discretization_measure) {
        addendo1[[h]]<-function(nu,h) {
                2*T*exp(dis[h])*addendo1_aux(nu,h)
        }
}
addendo2_aux=list()
```

```
addendo2_aux<-function(nu,h) {
        c=vector()
                for (j in 1:N) {
                c=c(c,W[j]*(z_restricted[[j]](nu,j)+max(0,S_0-exp(x_
                    restricted[j]-r*T))-C[j])*(max(0,S_0-exp(x_restricted[
                    j]-r*T-dis[h]))-max(0,S_0-exp(x_restricted[j]-r*T))))
        }
        sum(c)
}
addendo2=list()
for (h in 1:N_discretization_measure) {
        addendo2[[h]]<-function(nu,h) {
                2*T*exp(dis[h])*addendo2_aux(nu,h)
        }
}
f_notes=list()
for (k in 1:N_tilde) {
        f_notes[[k]] <- function(nu,k) {
                sqrt(2*pi)/S_0*i*u[k]*Zeta[[k]](nu,k)+exp(i*u[k]*(log(S_0)
                    +r*T))/(i*u[k]+1)
        }
}
aux<-function(nu) {
        c=vector()
        for (k in 1:N_tilde) {
                c=c(c,w_tilde[k]*f_notes[[k]](nu,k))
        }
        c
}
g_again<-function(nu) {fft(aux(nu))}
addendo3_aux=list()
addendo3_aux<-function(nu) {
        c=vector()
        for (j in 1:N) {
                c=c(c,W[j]*(z_restricted[[j]](nu,j)+max(0,S_0-exp(x_
                    restricted[j]-r*T))-C[j])*exp(i*x_restricted[j]*A/2)*g
                    _again(nu)[which(x==x_restricted[j])])
        }
        sum(Re(c))
}
addendo3=list()
for (h in 1:N_discretization_measure) {
        addendo3[[h]]<-function(nu,h) {
                S_0*T/pi*A/N_tilde*(1-exp(dis[h]))*addendo3_aux(nu)
        }
}
## this is the gross gradient
grad<-function(nu) {
        c=vector()
                for (h in 1:N_discretization_measure) {
                        c=c(c,addendo1[[h]](nu,h)+addendo2[[h]](nu,h)+
                            addendo3[[h]](nu,h))
        }
        c
```

```
}
tizia_revenge=optim(nu0, objective_functional, grad, method = "L-BFGS-B",
    lower=1e-17)
tizia_revenge
alpha_regularization=tizia_revenge$value
objective_functional1<-function(nu) {
        addendo1_objective(nu)+alpha_regularization*entropy_term(nu)
}
grad1<-function(nu) {
        c=vector()
        for (h in 1:N_discretization_measure) {
                c=c(c,addendo1[[h]](nu,h)+addendo2[[h]](nu,h)+addendo3[[h
                    ]](nu,h)+alpha_regularization*grad_entropy_term(nu)[h
                    ])
        }
        c
}
tizia1_revenge=optim(nu0, objective_functional1, grad1, method = "L-BFGS-
    B",lower=1e-17)
tizia1_revenge\label{key}
predicted_optionprices=numeric(N)
for (j in 1:N) {
        predicted_optionprices[j]=z_restricted[[j]](tizia1_revenge$par,j)
            +pmax(0,S_0-exp(x_restricted[j]-r*T))
}
predicted_optionprices
mean((abs(predicted_optionprices-C))/C)
```

# Appendix B

# Nonparametric Estimation Of Risk–Neutral Densities

In this appendix we are going to present an option pricing method which is nonparametric. The remarkable advantage in this approach is that we can reduce the misspecification risk. In other terms, this model dispenses with the constraints given by the assumptions on the underlying asset price dynamics or the statistical family of distributions that the risk–neutral density is assumed to belong to. This will turn into the absence of dynamics for the price process $S$. Of course, nonparametric estimation techniques require larger sample sizes for the same accuracy as parametric estimation procedures, therefore the increasing availability of large data sets (e.g., intraday traded option prices) made them doable and feasible. In particular, we underscore the key passages of the paper [14]. We instead refer to [9, 10] for an adaptation of the same model to commodity futures markets.

Let us introduce a generic security whose price process is described by the $\mathbb{R}^+$–valued stochastic process $S = \{S_t\}_t$ defined on a reference probability space $(\Omega, \mathcal{F}, P)$. Besides, we assume that the market is dynamically complete and that it admits an equivalent martingale measure $Q$ for $S$. This means that the process $\widetilde{S} = \left\{\widetilde{S}_t\right\}_t$, defined by

$$\widetilde{S}_t \coloneqq e^{-\int_0^t r(s)\,ds} S_t, \quad t \geq 0,$$

with $r = r(t)\,,\, t \geq 0$, which is a deterministic function representing the risk–free interest rate, is a $Q$–local martingale. To ease the notation, we set $r_{u-t} \coloneqq \int_t^u r(s)\, ds$. Under these assumptions, if we consider an European call option on $S$ with maturity $\overline{T}$ and strike $K$, then its price at time $t \geq 0$ is determined by the *risk–neutral pricing formula*:

$$C_t\left(K, \overline{T}\right) = e^{-r_{\overline{T}-t}} \int_{(0,\infty)} (S_T - K)^+\, q\left(S_T | \overline{T}, r_{\overline{T}-t}, \delta_{\overline{T}-t}, S_t\right)\, dS_T. \tag{B.1}$$

In (B.1), the quantity $\delta_{\overline{T}-t}$ represents the dividend yield of the asset in the period $\left(t, \overline{T}\right)$ and $q\left(S_T | \overline{T}, r_{\overline{T}-t}, \delta_{\overline{T}-t}, S_t\right)$ is the conditional risk–neutral density. Implicitly we are assuming that these state variables contain all the needed information to estimate the option prices and $q$, while other market factors, such as the volatility, are considered to be negligible. Introducing the time to maturity $\tau := \overline{T} - t$, then (B.1) reduces to

$$C_t\left(K, \overline{T}\right) = e^{-r_\tau} \int_{(0,\infty)} \left(S_T - K\right)^+ q\left(S_T | \overline{T}, r_\tau, \delta_\tau, S_t\right)\, dS_T.$$

In order to keep the notation simple, we just write $q\left(S_T\right)$ instead of $q\left(S_T|\cdot\right)$. If we now assume that all the variables other than $K$ are fixed, then the price of an European call option expiring in $\tau$ units of time can be expressed under the historical probability density $p$ in the following way:

$$\begin{aligned} C_t\left(K\right) &= e^{-r_\tau} \int_{(0,\infty)} \left(S_T - K\right)^+ \frac{q\left(S_T\right)}{p\left(S_T\right)} p\left(S_T\right)\, dS_T \\ &= e^{-r_\tau} \int_{(0,\infty)} \left(S_T - K\right)^+ m\left(S_T\right) p\left(S_T\right)\, dS_T, \end{aligned} \tag{B.2}$$

where $m\left(\cdot\right)$ is defined by

$$q\left(S_T\right) = m\left(S_T\right) p\left(S_T\right), \quad S_T \in \mathbb{R}^+ \tag{B.3}$$

and is called *pricing kernel*, or *stochastic discount factor*. In financial mathematics this quantity is particularly important, since it summarizes the relationship between the physical measure $p$ and the risk–neutral density $(RND)$ $q$, hence information on the asset price.

## B.1 Estimation Of $RND$ Via Empirical Pricing Kernel

From an economic point of view, the pricing kernel $m$ describes the risk preferences of an agent in an exchange economy and in a lot of applications it is the object of interest. Our goal is to estimate $m$ directly, i.e., from observable option prices, as well as to evaluate $p$ from historical data. In this way, we would be enabled to asses an appraisal of $q$ using (B.3). As already happened for the $RND$, we cannot incorporate all the factors driving the form of the pricing kernel, so we consider its projection into the set of available payoff functions and denote it as $m^*$. In this way, $m^*$ is a function of only $S_T$. We have to require two technical assumptions:

i. the applications $m$ and $m^*$ need to be close in some sense. We therefore fix $\epsilon > 0$ sufficiently small and assume that

$$\|m - m^*\|^2 := \int_{\mathbb{R}^+} |m\left(x\right) - m^*\left(x\right)|\, dx < \epsilon;$$

ii. there exists a sequence $\left(\alpha_l\right)_l \subset \mathbb{R}$ such that the projection $m^*$ has the series expansion

$$m^*\left(S_T\right) = \sum_{l=1}^{\infty} \alpha_l g_l\left(S_T\right), \quad S_T \in \mathbb{R}^+, \tag{B.4}$$

where $\{g_l\}_l$ is a fixed collection of basis functions. For example, paper [14] adopts the Laguerre polynomials to conduct empirical analysis.

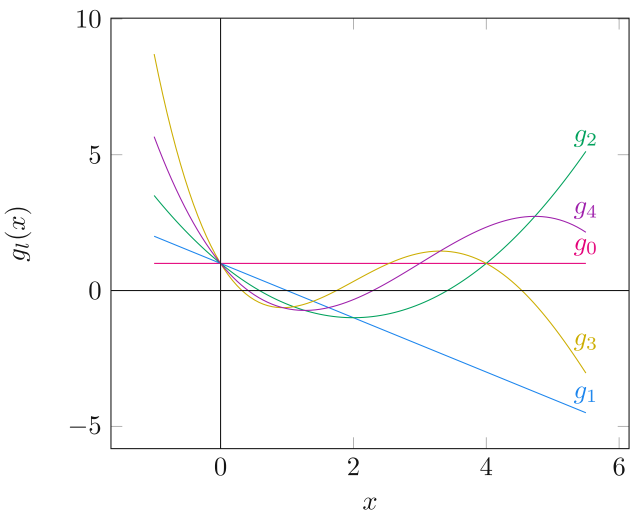



Figure B.1: First five terms of Laguerre polynomial sequence.

**Remark B.1.** The *Laguerre polynomials* are a polynomial sequence defined by recurrence. In particular, putting the first two terms

$$g_0(x) := 1, \qquad g_1(x) := 1 - x$$

for $x \in \mathbb{R}$, then

$$g_{l+1}(x) := \frac{(2l + 1 - x)\, g_l(x) - l g_{l-1}(x)}{l+1}, \qquad x \in \mathbb{R},\ l = 2, 3, \dots .$$

Hence, for example, we can easily get:

$$g_2(x) = \frac{1}{2}\left(x^2 - 4x + 2\right);$$
$$g_3(x) = \frac{1}{6}\left(-x^3 + 9x^2 - 18x + 6\right);$$
$$g_4(x) = \frac{1}{24}\left(x^4 - 16x^3 + 72x^2 - 96x + 24\right).$$

According to [1], formula 22.6.15, the functions $g_l$ are solutions to the Laguerre's differential equation

$$xy'' + (1 - x)y' + ly = 0, \qquad l \in \mathbb{N}$$

and they can be directly computed by (formula 13.60 in [2])

$$g_l(x) = \sum_{k=0}^{l} \binom{l}{k} \frac{(-1)^k}{k!} x^k, \qquad x \in \mathbb{R},$$

which shows that $g_l(0) = \binom{l}{0} = 1$ for any $l \in \mathbb{N}$. Considering the inner product $< \cdot, \cdot >: \mathbb{R}[x] \times \mathbb{R}[x] \to \mathbb{R}$ defined by

$$< f, g > := \int_0^\infty f(x)\, g(x)\, e^{-x}\, dx, \qquad f, g \in \mathbb{R}[x],$$

then we can use the *Rodrigues's formula* ([1], 22.11.6)

$$g_l(x) = \frac{e^x}{l!}\frac{d^l}{du^l}\left(u^l e^{-u}\right)\Big|_{u=x}, \quad x \in \mathbb{R}$$

to conclude that the Laguerre's polynomials constitute an orthonormal system.

In practice, very intuitively, we can only expand $m^*$ up to a finite number $L$. So, having fixed such $L$, if we were able to estimate the coefficients $\alpha_l$ (for $l = 1, ..., L$) directly from the market, for example using the least–squares procedure, then we would obtain the approximation

$$\widehat{m^*}(S_T) = \sum_{l=1}^{L} \widehat{\alpha}_l g_l(S_T), \quad S_T \in \mathbb{R}^+.$$

By the assumption (i.), such estimate would provide us with an approximation of $m$, as well. It is obvious that the choice of $L$, which is the point at which we truncate the series in (B.4), has deep aftermath in the quality of the estimation: the larger $L$, the better the fit, but the higher the computational cost and less robust the result. However, the pricing kernel is only indirectly determined by the price of an European call option through (B.2). Thus, the problem is now to find a feasible estimator of the vector $\alpha := (\alpha_1, ..., \alpha_L)^T$.

Let us suppose that there are $N$ options in the market at time $t$, with maturity $\overline{T_i}$, strike $K_i$ and price $C_t^i$, for $i = 1, ..., N$. Denote the vector of prices by

$$C_t := \left(C_t^1, ..., C_t^N\right)^T$$

and define, for every $l = 1, ..., L$, the quantities

$$\psi_{il} := e^{-r_\tau i} \int_{(0,\infty)} (S_{T_i} - K_i)^+ g_l(S_{T_i}) p(S_{T_i})\, dS_{T_i}, \quad i = 1, ..., N.$$

Now by (B.2) we should obtain

$$\begin{aligned} C_t^i &= e^{-r_\tau i} \int_{(0,\infty)} (S_{T_i} - K_i)^+ \left(\sum_{l=1}^{L} \widehat{\alpha}_l g_l(S_{T_i})\right) p(S_{T_i})\, dS_{T_i} + \epsilon_i \\ &= \sum_{l=1}^{L} \widehat{\alpha}_l \psi_{il} + \epsilon_i, \quad i = 1, ..., N, \end{aligned}$$

for some $\epsilon_i > 0$ small enough. Then, for $p$ known, the vector $\widehat{\alpha} := (\widehat{\alpha_1}, ..., \widehat{\alpha_L})^T$ could be obtained by a least–squares technique:

$$\widehat{\alpha} = \arg\min_{\alpha} \left\{ \sum_{i=1}^{N} \left| C_t^i - \sum_{l=1}^{L} \alpha_l \psi_{il} \right|^2 \right\}. \tag{B.5}$$

However even in this case $p$ is not known and can only be estimated. In order to do so, we replace it by a kernel density estimator $\widehat{p}$, arriving at

$$\hat{\psi}_{il} := e^{-r_\tau i} \int_{(0,\infty)} (S_{T_i} - K_i)^+ g_l(S_{T_i}) \widehat{p}(S_{T_i})\, dS_{T_i}, \quad i = 1, ..., N,$$

where $l = 1, ..., L$. Following this method we end up with an estimator $\widehat{\alpha}$ of $\alpha$ which is a feasible version of (B.5), namely

$$\widehat{\alpha} = \left(\widehat{\Psi}^T \widehat{\Psi}\right)^{-1} \widehat{\Psi}^T C_t,$$

where $\widehat{\Psi}$ is the $N \times L$ matrix given by $\widehat{\Psi} := \begin{bmatrix} \hat{\psi}_{11} & \hat{\psi}_{12} & \dots & \hat{\psi}_{1L} \\ \hat{\psi}_{21} & \hat{\psi}_{22} & \dots & \hat{\psi}_{2L} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\psi}_{N1} & \hat{\psi}_{N2} & \dots & \hat{\psi}_{NL} \end{bmatrix}$.

At this point, we retrace our steps getting an estimate of the pricing kernel provided by

$$\widehat{m}(s) = g(s)\,\widehat{\alpha}(s), \quad s \geq 0,$$

where $g := (g_1, ..., g_L)$. In conclusion, the risk–neutral density is estimated by (B.3) as

$$\widehat{q}(s) = \widehat{m}(s)\,\widehat{p}(s), \quad s \geq 0.$$