Comparing Two Different Option Pricing Methods

Alessandro Bondi [1], Dragana Radojičić [2],* and Thorsten Rheinländer [2]

[1] Classe di Scienze, Scuola Normale Superiore di Pisa, 56126 Pisa, Italy; alessandro.bondi@sns.it
[2] TU Wien, Institute of Statistics and Mathematical Methods in Economics, 22180 Vienna, Austria; rheinlan@fam.tuwien.ac.at
* Correspondence: dragana@fam.tuwien.ac.at or gagaradojicic@gmail.com

Abstract: Motivated by new financial markets where there is no canonical choice of a risk-neutral measure, we compared two different methods for pricing options: calibration with an entropic penalty term and valuation by the Esscher measure. The main aim of this paper is to contrast the outcomes of those two methods with real-traded call option prices in a liquid market like NASDAQ stock exchange, using data referring to the period 2019–2020. Although the Esscher measure method slightly underperforms the calibration method in terms of absolute values of the percentage difference between real and model prices, it could be the only feasible choice if there are not many liquidly traded derivatives in the market.

1. Introduction

Empirical studies show that real financial markets are incomplete, which means that not all derivative securities like options are replicable by means of some initial capital plus the value process of some self-finance trading strategy, and therefore hold unhedgeable and undiversifiable risks. The second fundamental theorem of asset pricing implies that in incomplete markets there typically exist several martingale measures consistent with the no-arbitrage principle that can be chosen as pricing measures (see, e.g., Jeanblanc et al. 2009; Shreve 2004). In fact, the range of option prices composes the total range between the inf and the sup of expected values with respect to martingale measures, and is therefore too large for practical purposes, see (Eberlein and Hammerstein 2004, Theorem 11.55).

Therefore, it makes sense to select a particular martingale measure for pricing purposes. In this study, we introduce, discuss, and compare two methods for the simulation of European call option prices: one based on the Esscher martingale measure and the other on the calibration to real-traded securities with an entropic penalty term. We implemented them in the programming language *R* and contrast the simulated derivative prices with real data retrieved from NASDAQ option chains. The objective of the paper was to provide a good fit between real and simulated contingent claim prices, consequently suggesting a choice for risk-neutral pricing measures in incomplete markets. In order to fulfill this aim, we constructed a model rich enough to describe stock prices but feasible for practice and connected to option prices. In this context, it is natural to exogenously specify different, suitable underlying asset price dynamics on a case-by-case basis.

The Esscher transform, invented by the Swedish actuary F. Esscher and introduced in Esscher (1932), is a time-honored tool in actuarial science for approximating the distribution of the aggregate claims of a portfolio. It consists of an exponential tilting procedure. More recently, it has been used as a premium calculation principle, see Van Heerwaarden et al. (1989), in option pricing, see Gerber and Shiu (1994), and also in pricing defaultable assets, see Lee and Rheinländer (2012). In our financial

setting, the agent has the option to invest in a financial market via admissible strategies. Hereby we exclude strategies leading to arbitrage opportunities. The corresponding theory has been developed by Kallsen and Shiryaev, see Kallsen and Shiryaev (2002). In Benth and Schmeck (2014), the authors used an Esscher transform incorporated in a two-factor model to derive futures prices in electricity markets. Furthermore, pricing methodology for weather derivatives in a specific two-factor model is established in Hell et al. (2012).

The idea of calibration is to consider a set of liquid option prices, and to find a market martingale measure such that the expected values of the option payoffs are as close as possible to the observed prices. This approximation problem is typically ill-posed, and one has to include a regularization term, we opted for an entropic penalty term in this regard. A general exposition about the calibration method can be found in Cont and Tankov, see Cont and Tankov (2004a).

Sections 2 and 3 are divided into a theoretical and a numerical part. The second section is devoted to the geometric Esscher measure and Laplace cumulant processes. In this part, we modeled the log–prices of the stock with a *COGARCH* process (introduced in Klüppelberg et al. (2004)) and present a sufficient condition for the existence of such a measure. Assuming the driving Lévy process to have a *NIG* (normal inverse Gaussian) distribution, which is infinitely divisible and thus corresponds to a Lévy process, we explain the original procedure—relying on the canonical representation theorem for semimartingales—followed by simulating the paths of the stock pricesunder the martingale measure. We pinpoint that, as Lemma 1 shows, the choice of the *NIG* distribution in the *COGARCH* model is crucial, since it allows for generating the risk-neutral trajectories of the underlying asset price.

In the third section, the calibration method is explained; using the fundamental concepts of relative entropy of distributions and Lévy processes on the Skorohod space (presented in Appendix E), we chose a martingale measure under which the dynamics of the stock prices follow an exponential Lévy process. In this part we mainly follow Cont and Tankov (2004a), even if the mathematical structure was refined and generalized, and the numerical approximation machinery was tailored to our goals.

In order to make the readers familiar with the theory and the notation that stands behind the main concepts of these two methods, in the appendix we also present a construction of the *NIG* distribution and some essential results on Lévy processes, highlighting the connection with the characteristics of semimartingales.

The main point of the present paper is as follows: if there are many liquid options around, like plain vanilla calls and puts on liquid stocks, we would expect the calibration method to perform best, therefore, we chose it as the benchmark method. However, in new financial markets, like insurance derivatives, cryptocurrencies, energy, or electricity markets, there are often only a few derivatives or any at all available, or the underlying (like electricity) is difficult to trade in since it is not storable. In these cases the calibration method is not implementable, and there is not much choice other than to use the Esscher method which has been done, e.g., in the book by Benth, Benth, and Koekebakker about electricity and related markets, see Benth et al. (2008). Besides that, there are few, if at all, studies about practical implementation of hedging strategies in incomplete markets, therefore our study fills a gap in this direction. Our main result thus is that while the Esscher martingale measure based pricing method in a liquid market does underperform the calibration method, as is to be expected, it does so only by less than 5%. So it might be a feasible choice of method in new financial markets.

2. Esscher Measure Method

In this section, the dynamics of the stock prices are introduced and sufficient conditions for the existence of the Esscher measure, which was used to simulate European call options prices, are established. We refer to Appendix D for a brief summary of the theory, which is necessary to construct the geometric Esscher measure. In the following, we use notation that can be found, e.g., in the books (Rheinländer and Sexton 2011; Shiryaev and Jacod 2003), as well as references therein.

2.1. The Model

On a probability space $(\Omega, \mathcal{F}, P)$ we introduce a driving, $\mathbb{R}$-valued Lévy process $L = \{L_t\}_t$ with generating triplet $(\sigma^2, \nu, \gamma_h)$ with respect to a truncation function h (see Appendices B and C). We endow this space with $\mathbb{F}$, the augmented filtration of L. Define the dynamics of the spot prices by the process $S = \{S_t\}_t$, where $S := S_0 \exp(G)$, $S_0 \in \mathbb{R}^+$, with the log-prices $G = \{G_t\}_t$ which follow a *COGARCH* process, introduced in Klüppelberg et al. (2004):

$$dG_t = \sigma_{t-}dL_t, \quad G_0 = 0. \tag{1}$$

The volatility process $\sigma^2 = \{\sigma_t^2\}_t$ is given by

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

with $\sigma_0^2 \in \mathbb{R}^+$ and an adapted, right-continuous with left limits (RCLL) process $X = \{X_t\}_t$ defined by

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

where $k > 0$, $\eta > 0$, $\Phi \geq 0$. We recall that a process is said to be RCLL if it has, almost surely, right-continuous trajectories with left limits. Although σ^2 is a left-continuous process, as shown by (2), we use σ_- with the aim to highlight that we are working with a process which is adapted and left-continuous, therefore predictable.

For an adapted process $r = \{r_t\}_t$ modeling the interest rate, the discounted spot price process is given by $\widetilde{S} = \left\{\widetilde{S}_t\right\}_t$, where $\widetilde{S}_t := \exp\left(-\int_0^t r_s ds\right) S_t$, $t \geq 0$. Thus, we consider the process $G' = \{G'_t\}_t$, where $G'_t := -\int_0^t r_s\, ds + G_t$ for every $t \geq 0$, whose dynamics follow $dG'_t = -r_t dt + dG_t$. At this point, we can express $\widetilde{S} = S_0 \exp(G')$ and we want to find a process $\theta \in L(G')$ such that $\widetilde{S}$ is a P^θ–local martingale: P^θ denotes the geometric Esscher measure, or Esscher martingale transform, for the exponential process $\widetilde{S}$ (see Appendix D).

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

$$\mu^G(\omega; (0,t] \times A) = \sum_{s \leq t} \mathbb{1}_{\{x \neq 0\} \cap A}\,(\Delta G_s(\omega)) = \sum_{s \leq t} \mathbb{1}_{\{x \neq 0\}}(\Delta L_s\,(\omega))\,\mathbb{1}_{(\sigma_{s-}(\omega))^{-1}A}(\Delta L_s(\omega))$$

$$= \mu^L\left(\omega; (0,t] \times (\sigma_-\,(\omega))^{-1} A\right), \quad A \in \mathcal{B}\,(\mathbb{R}).$$

Furthermore $\mu^{G'} = \mu^G$, hence $F^{G'}_{(t,\omega)}\,(dy) = \nu\,(\sigma_{t-}\,(\omega)\,\cdot)^{-1}(dy)$, meaning that

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

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

$$\left(\theta_t + \frac{1}{2}\right)\sigma_{t-}^2 \sigma^2 - r_t + \sigma_{t-}\gamma_h + \int_{\mathbb{R}} \left(e^{(\theta_t+1)\sigma_{t-}y} - e^{\theta_t \sigma_{t-}y} - \sigma_{t-}h\,(y)\right)\nu\,(dy) = 0 \tag{3}$$

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

Note that G' is a 1-dimensional process, thus the geometric Esscher measure is unique.

Proof. Let

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

and $G'(h)_t := G'_t - \widetilde{G'}(h)_t$, $t \geq 0$. By Remark 1, we get

$$dG'(h)_t = dG'_t - d\widetilde{G'}(h)_t = -r_t dt + dG_t - (x - h(x)) \, \mu^G(dt, dx)$$
$$= -r_t dt + \sigma_{t-} dL_t - (\sigma_{t-}y - h(\sigma_{t-}y)) \, \mu^L(dt, dy).$$

Taking into account Lemma A1 in Appendix D, which provides

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

and considering that $(\sigma_- \cdot L)^c = \sigma_- \cdot L^c$ and $\langle \sigma_- \cdot L^c, \sigma_- \cdot L^c \rangle = \int \sigma_-^2 \sigma^2 dt$, we obtain the characteristics of G' under P with respect to h:

$$\begin{cases} b_t^{G'} = -r_t + \sigma_{t-}\gamma_h + \int_{\mathbb{R}} (h(\sigma_{t-}y) - \sigma_{t-}h(y)) \, \nu(dy) \\ c_t^{G'} = \sigma_{t-}^2 \sigma^2 \\ F_t^{G'}(dy) = \nu(\sigma_{t-} \cdot)^{-1}(dy) \end{cases}, \quad t \geq 0. \tag{4}$$

Thus, if θ satisfies $\widetilde{\kappa}(\theta + 1) - \widetilde{\kappa}(\theta) = 0$, thanks to Theorem A2 we can conclude that $\widetilde{S}$ is a P^θ-local martingale. In fact, using (A11) we have

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

where the last equality holds by (3). $\square$

The next lemma provides us with the candidate solutions to Equation (3) when L follows a *NIG* distribution (see Appendix A).

Lemma 1. *Let $L = \{L_t\}_t$ be a NIG-distributed Lévy process with parameters $(\alpha, \beta, \mu, \delta)$ and define the truncation function $h(x) := x\mathbb{1}_D(x)$, $x \in \mathbb{R}$. If a process $\theta = \{\theta_t\}_t$ such that*

$$(\sigma_{t-}\theta_t)(\omega), \; \sigma_{t-}(\theta_t + 1)(\omega) \in [-\alpha - \beta, \alpha - \beta], \quad t \in \mathbb{R}_0^+, \; \omega \in \Omega \tag{5}$$

fulfills Equation (3), then for every $t \geq 0$ we have

$$\theta_t^{1,2} = \frac{1}{2\left(R_t^2 + \delta^2\sigma_{t-}^2\right)\sigma_{t-}} \left(-\delta^2\sigma_{t-}^3 - 2\beta\delta^2\sigma_{t-}^2 - R_t^2(\sigma_{t-} + 2\beta) \right.$$

$$\left. \pm \sqrt{4\alpha^2\delta^2 R_t^2\sigma_{t-}^2 + 4R_t^4\alpha^2 - R_t^2\delta^2\sigma_{t-}^4 - \frac{R_t^6}{\delta^2} - 2\sigma_{t-}^2 R_t^4} \right), \tag{6}$$

with $R_t := -r_t + \mu\sigma_{t-}$.

Proof. Since $L_1 \sim NIG\,(\alpha, \beta, \mu, \delta)$, we apply (A6) and (A9) (see Appendices A and B) to obtain

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

where $\gamma_1 := \dfrac{2\delta\alpha}{\pi} \displaystyle\int_0^1 \sinh\,(\beta x)\, K_1\,(\alpha x)\, dx$ and $\nu\,(dx) = \dfrac{\delta\alpha}{\pi\,|x|} e^{\beta x} K_1\,(\alpha\,|x|)\, dx$. Using (A5), we get

$$\delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + z)^2}\right) = \gamma_1 z + \int_{\mathbb{R}} (e^{zx} - 1 - zx 1_D\,(x))\, \nu\,(dx) \tag{7}$$

for $z \in [-\alpha - \beta, \alpha - \beta]$. Proceeding as in the proof of Theorem 1, we have

$$0 = \widetilde{\kappa}\,(\theta + 1)_t - \widetilde{\kappa}\,(\theta)_t$$
$$= R_t + \sigma_{t-}\gamma_1 + \int_{\mathbb{R}} \left[e^{(\theta_t + 1)\sigma_{t-}x} - 1 - \sigma_{t-}\,(\theta_t + 1)\, x 1_D\,(x)\right]\nu\,(dx)$$
$$- \int_{\mathbb{R}} \left(e^{\theta_t \sigma_{t-}x} - 1 - \sigma_{t-}\theta_t x 1_D\,(x)\right)\nu\,(dx), \quad t \geq 0. \tag{8}$$

Since $\sigma_-\gamma_1 = \sigma_-\left[(\theta + 1) - \theta\right]\gamma_1$, we expand on the chain of equalities in (8):

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

where the last equality is obtained by combining (5) and (7). At this point it remains to find the possible solutions for every $t \geq 0$ to the equation

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

which can be written as $\sqrt{\alpha^2 - (\beta + \sigma_{t-}\,(\theta_t + 1))^2} = \dfrac{R_t}{\delta} + \sqrt{\alpha^2 - (\beta + \sigma_{t-}\theta_t)^2}$. Squaring both sides we get

$$2\frac{R_t}{\delta}\sqrt{\alpha^2 - (\beta + \sigma_{t-}\theta_t)^2} = -\left(\sigma_{t-}^2 + 2\beta\sigma_{t-} + \frac{R_t^2}{\delta^2}\right) - 2\sigma_{t-}^2\theta_t, \tag{9}$$

and repeating once again the same operation of squaring we end up with an equation of second degree in θ:

$$4\sigma_{t-}^2\left(\sigma_{t-}^2 + \frac{R_t^2}{\delta^2}\right)\theta_t^2 + 4\sigma_{t-}\left(\sigma_{t-}^3 + 2\beta\sigma_{t-}^2 + \frac{R_t^2}{\delta^2}\sigma_{t-} + 2\beta\frac{R_t^2}{\delta^2}\right)\theta_t + c_t = 0, \tag{10}$$

where $c_t := \left(\sigma_{t-}^2 + 2\beta\sigma_{t-} + \frac{R_t^2}{\delta^2}\right)^2 - 4\frac{R_t^2}{\delta^2}\,(\alpha^2 - \beta^2)$. Computing algebraically the solutions of (10) we get the expressions in (6). $\square$

We have empirically tested both possible branches of solutions in (6), and we have chosen

$$\theta_t = \frac{1}{2\left(R_t^2 + \delta^2\sigma_{t-}^2\right)\sigma_{t-}}\left(-\delta^2\sigma_{t-}^3 - 2\beta\delta^2\sigma_{t-}^2 - R_t^2\,(\sigma_{t-} + 2\beta)\right.$$
$$\left. - \sqrt{4\alpha^2\delta^2 R_t^2\sigma_{t-}^2 + 4R_t^4\alpha^2 - R_t^2\delta^2\sigma_{t-}^4 - \frac{R_t^6}{\delta^2} - 2\sigma_{t-}^2 R_t^4}\right) \tag{11}$$

to run the simulation because it prevents the risk-neutral dynamics from "exploding", i.e., from skyrocketing to unrealistically high prices in a short time.

Remark 2. *Unfortunately, it does not seem possible for us to prove, from an analytical point of view, that either candidate solution in* (6) *solves* (3), *i.e., that it is a true solution. What the simulations suggest is that only the process in* (11) *is indeed a solution of* (3). *This insight is confirmed by the numerical experiments to a greater extent. In fact, it turns out that both processes $\theta^{1,2}$ satisfy Condition* (5). *However, θ^1 makes the right hand side in* (9) *negative, and this implies that it does not solve* (3). *On the contrary, the process θ^2—the one we pick in* (11)—*makes such term positive, hence it solves* (3), *indeed.*

2.1.1. Simulation of the P^θ-Dynamics

The goal of this section is to carefully describe a procedure to generate the risk-neutral dynamics of the underlying asset price. This allows us to compute the simulated European call option prices with the Monte Carlo method.

Let $r \in \mathbb{R}^+$ represent the constant annual interest rate. We assume that the driving Lévy process L is NIG-distributed and that the hypothesis of Theorem 1 hold (with respect to the truncation function $\overline{h}(x) = x1_D(x)$, $x \in \mathbb{R}$). Using the R-package *yuima* (see Iacus 2011) we estimate the parameters of the model from the time series of the underlying asset log-prices. Next, we simulated a sufficiently large number of paths of the variance process σ^2 with the found parameters: by (11), from each of them we can get a trajectory of θ. The issue reduces to generating the paths of G, but note that, after the change of measure, L is not a Lévy process anymore because $F_t^\theta(dx) = e^{\theta_t x}\nu(dx)$, $t \geq 0$. Consequently, G is not a $COGARCH$ process under the martingale measure. Then, in order to simulate its trajectories we use the canonical representation for semimartingales (Shiryaev and Jacod (2003), Chapter II, Theorem 2.34): for a d-dimensional semimartingale $X = \{X_t\}_t$ with characteristics (B, C, ν^X) relative to the truncation function h, it holds:

$$X = X_0 + X^c + B + h(x) \star (\mu^X - \nu^X) + (x - h(x)) \star \mu^X. \tag{12}$$

The symbol $\star$ denotes the integration with respect to a random measure: we refer to (Shiryaev and Jacod (2003), Chapter II, Section 1) for an extensive study of this topic. For the reader's convenience, we provide the basic definition.

Definition 1. *Let μ be a random measure on $\mathbb{R}_0^+ \times \mathbb{R}$ and $W : \Omega \times \mathbb{R}_0^+ \times \mathbb{R} \to \mathbb{R}$ be measurable function with respect to the product σ-algebra $\mathcal{O} \otimes \mathcal{B}(\mathbb{R})$, where $\mathcal{O}$ denotes the optional σ-algebra on $\Omega \times \mathbb{R}_0^+$. The integral process $W \star \mu$ is defined by*

$$W \star \mu_t(\omega) := \begin{cases} \int_{[0,t] \times \mathbb{R}} W(\omega; s, x) \, \mu(\omega; ds, dx), & \text{if } \int_{[0,t] \times \mathbb{R}} |W(\omega; s, x)| \, \mu(\omega; ds, dx) < \infty \\ \infty, & \text{otherwise} \end{cases}.$$

For an arbitrarily chosen $\epsilon \leq 1$, fix the truncation function $h(x) := x1_{\{|z| < \epsilon\}}(x)$ for $x \in \mathbb{R}$, and let $(0, \nu, \gamma_h)$ be the corresponding generating triplet of L. From the proof of Theorem 1, specifically from (4), we readily get the characteristics $\left(b_t, 0, \nu\left(\sigma_{t-} \cdot\right)^{-1}(dx)\right)$ of G under P, and invoking (A12) and Remark A2 in Appendix D we find them under P^θ:

$$\begin{cases} b_t^\theta = b_t + \int_\mathbb{R} h(x)\left(e^{\theta_t x} - 1\right) F_t(dx) \\ c_t^\theta = 0 \\ F_t^\theta(dx) = e^{\theta_t x}\nu(\sigma_{t-} \cdot)^{-1}(dx) \end{cases}, \quad t \geq 0. \tag{13}$$

We now represent G by (12), noting that $G_0 = 0$ and $G^c = 0$ since $c^\theta = 0$. As regards the term $B = \{B_t\}_t$, we expand on the computations in (13), getting

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

$$- \int_{(-\epsilon,\epsilon)} \sigma_{t-}y \frac{\delta\alpha}{\pi|y|} e^{\beta y} K_1(\alpha|y|)\,dy, \quad t \geq 0,$$

and obtain the variables of the process B by $B_t = \int_{(0,t)} b_s^\theta\,ds$, $t \geq 0$.

At this point we need to approximate a trajectory of the process $h(x) \star (\mu^G - \nu^G)$. In order to do so, we neglect the contribution of the jumps of G with absolute value smaller than ϵ, focusing just on the term

$$h(x) \star (\mu^G - \nu^G) \approx -h(x) \star \nu^G = -\int_{(0,t)} ds \int_{\mathbb{R}} x\,\mathbb{1}_{\{|z|<\epsilon\}}(x)\,F_s^\theta(dx)$$

$$= -\int_{(0,t)} ds \int_{|x|<\epsilon/\sigma_{s-}} \sigma_{s-}y \frac{\delta\alpha}{\pi|y|} e^{(\theta_s\sigma_{s-}+\beta)y} K_1(\alpha|y|)\,dy, \quad t \geq 0,$$

which in turn cancels out with one of the addends defining B. Heuristically speaking, in this approach we are assuming that the small jumps of the underlying asset price do not contribute to the determination of the option value in a significant way. However, we may refer to Asmussen and Rosiński (2001) for a more sophisticated procedure taking into account the variation of such small jumps in the case of Lévy processes.

Finally we concentrate on the term $(x - h(x)) \star \mu^G$: we simulate its paths similarly to the Lévy–Itô decomposition theorem (see Sato 1999, Theorem 19.2). According to this result, if $\widetilde{L}$ is a Lévy process with generating triplet (A, ν, γ) and $D_{a,\infty} := \{x \in \mathbb{R}^d : a < |x| < \infty\}$ for $a > 0$, then $\left\{\left((x - x\mathbb{1}_D(x)) \star \mu^{\widetilde{L}}\right)_t\right\}_t$ is a compound Poisson process with constant $\nu(D_{1,\infty})$ and distribution

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

Thus, we take a nonhomogeneous Poisson process with time-varying intensity

$$\lambda_t := F_t^\theta(\{|x| > \epsilon\}) = \int_{\mathbb{R}} e^{\theta_t\sigma_{t-}y}\,\mathbb{1}_{D_{\epsilon,\infty}}(\sigma_{t-}y)\frac{\delta\alpha}{\pi|y|} e^{\beta y} K_1(\alpha|y|)\,dy$$

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

and impose the time-varying jumps sizes to follow

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

The jump times of this process have been simulated by the thinning algorithm described in Lewis and Shedler (1979).

Denote by N the number of iterations we run and by S^i, for $i = 1, \ldots, N$, the corresponding, simulated trajectories of the spot price under the pricing measure P^θ. We obtain the value of an European call option with strike K and maturity T following the Monte Carlo method, i.e., computing the sample mean of the vector of components $e^{-rT}\left(S^i(T) - K\right)^+$, $i = 1, \ldots, N$.

2.1.2. Empirical Results

We have empirically tested the Esscher method on the prices of call options on Apple Inc. stock (ticker symbol: AAPL) with fixed strike at $320. In Figure 1a, the simulated option prices are shown as function of their maturities. The average of the absolute values of percentage difference is 3.1646%.
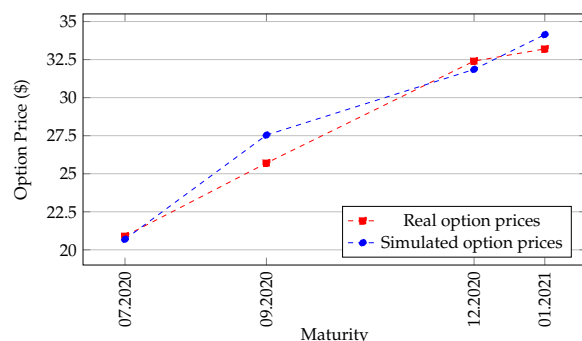
Furthermore, we applied this method to the prices of call options on Microsoft Corporation stock (ticker symbol: MSFT) with strike $190: we plot the outcomes in Figure 1b. In this case, the mean of absolute values of percentage difference settles down at 5.0518%.

Finally we report the results of a test on call options with strike $910 with Tesla, Inc. (ticker symbol: TSLA) as underlying stock. In Figure 1c we can notice a big discrepancy between the real price and the predicted value for the shortest-term maturity taken into account, the latter being 37.2755% smaller than the former. However, if we neglect this first term we recover a result in line with the previous experiments, namely an average for the absolute values of percentage difference of 4.0555%.
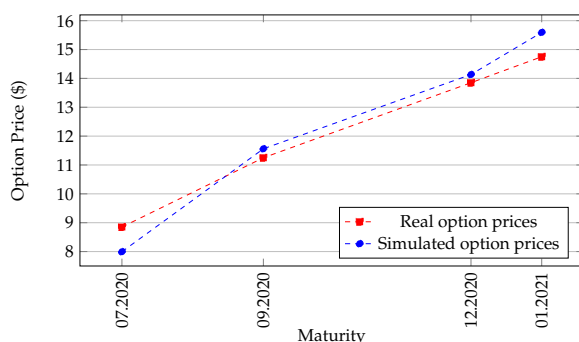
In all the three cases, we set the number of Monte Carlo iterations $N = 10^4$ and we used the time series of the log$-$prices associated to the trading year before 19 February 2020, to estimate the parameters of the *COGARCH* model (see Section 2.1.1). Therefore, the simulations were run prior to Apple's stock split (on 28 August 2020) and Tesla's stock split (on 31 August 2020).

Remark 3. *The real data was obtained from* https://old.nasdaq.com, *where American options are traded. Although our approximation generates the prices in an European model, it can be accepted as meaningful since:*

- *the analyzed derivatives have short-term maturities, so we are allowed to ignore the dividend yield (however, TSLA does not pay dividends);*
- *the time values of the options were always positive, implying that it is convenient to sell the option rather than exercising the call right.*

Maturity	Real data ($)	Model Price ($)
July 17, 2020	20.90	20.69
Sept. 18, 2020	25.70	27.54
Dec. 18, 2020	32.40	31.86
Jan. 15, 2021	33.20	34.14

(a) AAPL call options, Strike: $320.

Maturity	Real data ($)	Model Price ($)
July 17, 2020	8.85	7.99
Sept. 18, 2020	11.25	11.56
Dec. 18, 2020	13.85	14.13
Jan. 15, 2021	14.75	15.59

(b) MSFT call options, Strike: $190.

Figure 1. *Cont.*

Maturity	Real data ($)	Model Price ($)
July 17, 2020	158.17	99.21
Sept. 18, 2020	171.00	157.24
Oct. 16, 2020	180.00	178.13
Jan. 15, 2021	199.66	205.80
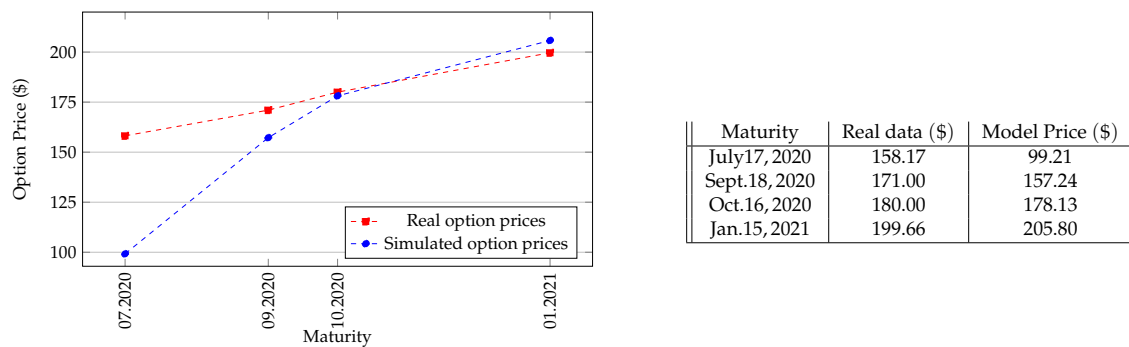
(c) TSLA call options, Strike: $910.

Figure 1. Numerical experiments for the Esscher method.

3. Calibration with Entropic Penalty Term Method

In the Esscher method, the density process is completely defined by the spot prices: this is an intuitive drawback in liquid markets, because it does not use all the available information (e.g., the time series of derivative prices). This is a motivation to consider the Calibration Method , which consists in modeling the spot prices dynamics directly under a martingale measure "chosen" by the market.

3.1. The Model

We start the practical analysis by taking into account a driving, $\mathbb{R}$-valued Lévy process $L = \{L_t\}_t$ with generating triplet (σ^2, ν, γ) on a probability space $(\Omega, \mathcal{F}, Q)$ endowed with the augmented filtration of L, which satisfies the usual conditions and is denoted by $\mathbb{G}$. We describe the stock prices $S = \{S_t\}_t$ with an *exponential Lévy model*: $S_t := S_0 \exp(rt + L_t)$, $t \geq 0$, with $S_0 \in \mathbb{R}^+$ and $r \in \mathbb{R}^+$ representing the constant annual interest rate. The discounting process $R = \{R_t\}_t$ reduces to the deterministic function $R_t := \exp(-rt)$, $t \geq 0$. Since we need Q to be a martingale measure for S we require:

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

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

With these assumptions, Remark A1 in Appendix B proves that $\{\exp(L_t)\}_t$ is a martingale with $E^Q[\exp(L_t)] = 1$ for every $t > 0$. Considering the discounted spot prices process $\widetilde{S} = \{\widetilde{S}_t\}_t$, defined by $\widetilde{S}_t := R_t S_t = S_0 \exp(L_t)$, $t \geq 0$, we can state that it is a Q-martingale with constant expectation equal to S_0.

In our discussion the driving Lévy process L will be a pure jump process, so its generating triplet simplifies as $\sigma^2 = 0$. In particular, due to assumption (ii) we get the next relation:

$$\gamma = -\int_{\mathbb{R}} (e^x - 1 - x\mathbb{1}_D(x)) \nu(dx). \tag{14}$$

Assume that there are N European call options with fixed maturity T available in the market, and for each $j = 1, \ldots, N$ let K_j and C^j be the strike price and the observed price of the j-th derivative, respectively. It is convenient to consider the quantities $k_j = \log(K_j)$, $j = 1, \ldots, N$. 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: $C_\nu^j := \exp(-rT) E^Q \left[(S_T - \exp(k_j))^+ \right]$, $j = 1, \ldots, N$. Denoting by Q_{L_T} the pushforward distribution on $\mathbb{R}$ generated by L_T we have

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

Thus, knowing the Lévy measure ν, the whole generating triplet could be recovered and therefore the option prices computed.

The reasoning which leads the calibration is choosing ν such that C_ν^j is "close" to C^j for any $j = 1, ..., N$. Hence we pick the Lévy measure of the model by a least-squares procedure:

$$\bar{\nu} = \arg\min_\nu \left\{ \sum_{j=1}^N \left| C_\nu^j - C^j \right|^2 \right\}. \tag{15}$$

Since the problem expressed in (15) is ill posed (see Cont and Tankov 2004a, Chp. 13), we introduce a regularization term. Let $L^0 = \{L_t^0\}_t$ be a, pure jump, driving Lévy process with generating triplet $(0, \nu_0, \gamma_0)$ statistically estimated from the time series of the underlying asset price. Furthermore, let Q^{L^0} and Q^L be the distributions on the Skorohod space $(\mathfrak{D}, \mathcal{F}_\mathfrak{D}; \mathbb{F})$ generated by L^0 and L, respectively, which are supposed to be equivalent on $\mathcal{F}_t$ for every $t > 0$ (see Appendix E). We use the relative entropy as a measure of the distance from Q^{L_0}, so that by Remark A3 the issue is reduced to solving the optimization problem:

$$\bar{\nu} = \arg\min_{\nu \in \mathbf{Q}} \left\{ \sum_{j=1}^N \left| C_\nu^j - C^j \right|^2 + \alpha \int_\mathbb{R} \left(\frac{d\nu}{d\nu^0} \log \frac{d\nu}{d\nu^0} + 1 - \frac{d\nu}{d\nu^0} \right) d\nu^0 \right\}, \tag{16}$$

for $\mathbf{Q} := \left\{ \nu : Q^L \big|_{\mathcal{F}_t} \sim Q^{L^0} \big|_{\mathcal{F}_t}, \, t > 0 \right\}$. The term α is a regularization parameter : the higher it is, the more we trust the initial distribution and the less importance we give to calibration. The existence of a solution to (16) has been studied, among others, in the paper Cont and Tankov (2004b). We are going to relax the assumptions in (16), considering the minimization in the set of the Lévy measures equivalent to ν^0.

3.1.1. Numerical Approximation

The goal of this section is to describe in each and every detail a discretization procedure that allows us to tackle the optimization problem in (16). To this aim, it is necessary to express the objective functional in (16) as a function of the masses of the discretized Lévy measures. The steps of our argument are the following:

I.   estimate the parameters of the prior Lévy process L^0 from the time series of log–returns;
II. introduce a discretization grid for the prior Lévy measure ν^0 and the driving Lévy measure ν (in what follows, the points of such a grid are denoted by $y_1 < \cdots < y_{N_d}$). Their discretized versions are denoted by ν_d^0 and ν_d, respectively;
III. compute the discrete version of the entropy term in (16) as a function of the masses of ν_d;
IV. use an approach based on the Fourier inversion theorem to get an approximation of the modified time values of the options at the log–strikes under scrutiny. This allows us to obtain (23), a discretized version of the objective functional in (16);
V.   calculate explicitly the derivatives of the discretized objective functional to speed up the simulations;
VI. choose the regularization parameter α.

Fix a maturity T and define the grid $x_j \in \mathbb{R}^+$, $j = 1, \ldots, N$, with $x_1 < x_2 < \ldots < x_N$: these points represent the log–strike prices of the options available on the market. Moreover, from the time series of the log–returns we obtain the parameters of the historical, *NIG*-distributed Lévy process $L^0 = \{L_t^0\}_t$ by a maximum likelihood procedure using the *R*-package *ghyp*. Note that, in this case, one can select any pure jump, infinitely divisible distribution: we opt for the *NIG* by analogy with the Esscher's method, and our choice is satisfactory, as the final goodness of fit between real data and model prices shows. First we introduce a discretization grid consisting in the points y_h, $h = 1, \ldots, N_d$,

with $y_1 < y_2 < \ldots < y_{N_d}$, which constitutes a partition of the interval $[-M, M]$ for some $M > 0$, and then 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 $\delta_{(\bar{a})}$ denotes the Dirac measure at a point $\bar{a}$ and

$$\nu_1^0 = \int_{(-\infty, y_1]} d\nu^0; \quad \nu_h^0 = \int_{(y_{h-1}, y_h]} d\nu^0, h = 2, \ldots, N_d - 1; \quad \nu_{N_d}^0 = \int_{(y_{N_d-1}, \infty)} d\nu^0.$$

Now we take into account another measure which has the same mass points as the previous one: $\nu_d (dy) = \sum_{h=1}^{N_d} \nu_h \delta_{(y_h)} (dy)$, with $\nu_h \in \mathbb{R}^+$ for every $h = 1, \ldots, N_d$. We note that the calibrated measures ν_d and ν_d^0 are equivalent, but ν_d and the prior ν^0 are not. We understand the measure ν_d as the discretized Lévy measure of the driving, pure jump Lévy processes $L = \{L_t\}_t$. We now want to compute the Radon–Nikodym derivative $\frac{d\nu_d}{d\nu_d^0}$ with the aim to explicitly express the entropy term of (16) in the discrete version

$$\int_{\mathbb{R}} \left(\frac{d\nu_d}{d\nu_d^0} \log \frac{d\nu_d}{d\nu_d^0} + 1 - \frac{d\nu_d}{d\nu_d^0} \right) d\nu_d^0. \tag{17}$$

Set $y_0 = -\infty$ and define the function

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

For every $A \in \mathcal{B}(\mathbb{R})$ it results

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

$$= \sum_{h=1}^{N_d} \frac{\nu_h}{\nu_h^0} \nu_h^0 \mathbb{1}_A (y_h) = \nu_d (A),$$

as $\nu_d^0 ((y_{h-1}, y_h] \cap A) = \nu_h^0 \mathbb{1}_A (y_h), h = 1, \ldots, N_d$. Therefore, $\frac{d\nu_d}{d\nu_d^0} = f$. Moving back to the integral (17), we get:

$$\int_{\mathbb{R}} \left(\frac{d\nu_d}{d\nu_d^0} \log \frac{d\nu_d}{d\nu_d^0} + 1 - \frac{d\nu_d}{d\nu_d^0} \right) d\nu_d^0 = \int_{(-\infty, y_{N_d}]} \left(\frac{d\nu_d}{d\nu_d^0} \log \frac{d\nu_d}{d\nu_d^0} + 1 - \frac{d\nu_d}{d\nu_d^0} \right) d\nu_d^0$$

$$= \sum_{h=1}^{N_d} \left[\frac{\nu_h}{\nu_h^0} \log \frac{\nu_h}{\nu_h^0} + 1 - \frac{\nu_h}{\nu_h^0} \right] \nu_d^0 ((y_{h-1}, y_h]) = \sum_{h=1}^{N_d} \left[\nu_h \left(\log \nu_h - \log \nu_h^0 \right) + \nu_h^0 - \nu_h \right].$$

Following the Carr and Madan approach (see Carr and Madan 1999, Section 3.2) for further details), we are able to to express the quantity $\sum_{j=1}^N \left| C_\nu^j - C^j \right|^2$ as a function of $\nu_1, \ldots, \nu_{N_d}$. The option price $C_\nu (k)$ is not an integrable function, as a swift application of Lebesgue's convergence theorem shows that it converges to S_0 as $k \to -\infty$. Therefore, we focus on the modified time value z_T, defined by $z_T (k) := C_\nu (k) - \left(S_0 - e^{k-rT} \right)^+$, $k \in \mathbb{R}$. We suppose that z_T and its inverse Fourier transform ζ_T are integrable, so by inversion (see, e.g., Rudin 1987, Theorem 9.11) we have

$$z_T (k) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-iku} \zeta_T (u) \, du \quad a.e. \tag{18}$$

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

$$\left(S_0 - e^{k-rT}\right)^+ = \left(S_0 - e^{k-rT}\right) 1_{\{z \le \log S_0 + rT\}}(k)$$

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

In a similar way we get

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

This provides the following expression for $z_T(k)$:

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

The assumption (ii) ensures that $\int_{\mathbb{R}} e^y \, Q_{L_T}(dy) = 1$, so we finally obtain

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

The insight here is to use the Fourier transform and its inverse to estimate the values $z_T(x_j)$ for every $j = 1, \dots, N$. Hence we fix a point $u \in \mathbb{R}$ and introduce the inverse Fourier transform

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

Under suitable conditions which allow us to switch the order of integration we have

$$\zeta_T(u) = -\frac{e^{-rT}}{\sqrt{2\pi}} \int_{(-\infty,0]} \left[\int_{(y+\log S_0 + rT, \log S_0 + rT)} e^{iuk} \left(S_0 e^{rT+y} - e^k\right) dk \right] Q_{L_T}(dy)$$

$$+ \frac{e^{-rT}}{\sqrt{2\pi}} \int_{(0,\infty)} \left[\int_{(\log S_0 + rT, y+\log S_0 + rT)} e^{iuk} \left(S_0 e^{rT+y} - e^k\right) dk \right] Q_{L_T}(dy). \tag{19}$$

We turn our attention on the computation of the first term of the sum (19). An explicit calculation of the inner integral (respect to the Lebesgue measure) of such addend, indicated by I_1, gives

$$I_1(y) = S_0 \frac{e^{rT + y + iu(\log S_0 + rT)}}{iu(iu+1)} \left[iu \left(1 - e^{-y} + 1 - e^{iuy}\right) \right]$$

$$= S_0 \frac{e^{rT + iu(\log S_0 + rT)}}{iu+1}(e^y - 1) + S_0 \frac{e^{rT + iu(\log S_0 + rT)}}{iu(iu+1)} e^y - S_0 \frac{e^{iu \log S_0} e^{(iu+1)rT}}{iu(iu+1)} e^{y+iuy}, \quad y \in (-\infty, 0].$$

The same strategy allows us to compute the inner integral I_2 of the second addend in (19), as well. Precisely, for every $y \in (0, \infty)$ we get

$$I_2(y) = -S_0 \frac{e^{rT + iu(\log S_0 + rT)}}{iu+1}(e^y - 1) - S_0 \frac{e^{rT + iu(\log S_0 + rT)}}{iu(iu+1)} e^y + S_0 \frac{e^{iu \log S_0} e^{(iu+1)rT}}{iu(iu+1)} e^{y+iuy}.$$

Therefore we are able to represent the inverse Fourier transform as

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

Both (A9) in Appendix B and assumption (i) make us conclude that

$$\zeta_T (u) = \frac{S_0}{\sqrt{2\pi}} \frac{e^{iu(\log S_0 + rT)}}{iu\,(iu+1)} \left[e^{T\Psi(iu+1)} - 1 \right], \quad u \in \mathbb{R}. \tag{20}$$

By the definition of Ψ in (A8) (see Appendix B), substituting the discrete version ν_d for the Lévy measure ν associated to L, as a consequence of (14) we get, for every $u \in \mathbf{R}$,

$$\Psi(iu+1) \simeq \int_{\mathbb{R}} \left(e^{(iu+1)y} - iue^y - e^y + iu \right) \nu_d(dy) = \sum_{h=1}^{N_d} e^{y_h} \left(e^{iuy_h} - 1 \right) \nu_h + iu \sum_{h=1}^{N_d} (1 - e^{y_h})\nu_h. \tag{21}$$

Plugging (21) into (20) we obtain, for $u \in \mathbb{R}$, the final approximation

$$\zeta_T(u) \simeq \frac{S_0}{\sqrt{2\pi}} \frac{e^{iu(\log S_0 + rT)}}{iu\,(iu+1)} \left[\exp\left(T \sum_{h=1}^{N_d} e^{y_h} \left(e^{iuy_h} - 1 \right)\nu_h + iuT \sum_{h=1}^{N_d} (1 - e^{y_h})\nu_h \right) - 1 \right]. \tag{22}$$

With the aim to approximate z_T at the points x_j, for $j = 1, \dots, N$, we construct a new, uniform grid with mesh d containing the log-strike one. Specifically, fixed $\widetilde{N} \in \mathbb{N}$, we set

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

for $\Delta := \frac{A}{\widetilde{N}}$, with $A := \frac{2\pi}{d}$, which is the size of the discretization interval. We carry out our construction so that for every $j = 1, \dots, N$ and $h = 1, \dots, N_d$ there exists a $n_{hj} \in \left\{ -\widetilde{N} + 1, \dots, \widetilde{N} - 1 \right\}$ such that $x_j - y_h = \widetilde{x}_{n_{hj}}$. We eventually introduce the points of the discretization grid as

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

We then compute

$$z_T(\widetilde{x}_n) \simeq \frac{1}{\sqrt{2\pi}} \int_{(-A/2, A/2)} e^{-iu\widetilde{x}_n} \zeta_T(u)\, du \simeq \frac{1}{\sqrt{2\pi}} \frac{A}{\widetilde{N}} \sum_{k=0}^{\widetilde{N}-1} e^{-iu_k \widetilde{x}_n} \widetilde{w}_k \zeta_T(u_k)$$

$$= \frac{1}{\sqrt{2\pi}} \frac{A}{\widetilde{N}} e^{i\frac{A}{2}\widetilde{x}_n} \sum_{k=0}^{\widetilde{N}-1} \exp\left(-i\frac{2\pi n}{\widetilde{N}} k \right) \widetilde{w}_k \zeta_T(u_k), \quad n = 0, \dots, \widetilde{N} - 1,$$

where $\widetilde{w}_k$ are chosen by the trapezoidal rule, i.e.,

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

Therefore, knowing $\zeta_T(u_k)$ for $k = 0, \dots, \widetilde{N} - 1$, we can use a fast Fourier transform (FFT) to estimate the values $z_T(\widetilde{x}_n)$, for $n = 0, \dots, \widetilde{N} - 1$, and, due to the symmetry of the grid, an inverse discrete Fourier transform to get $z_T(\widetilde{x}_n)$ for $n = -\widetilde{N} + 1, \dots, -1$. Thus, we reduce the optimization problem (16) to the minimization in $(\mathbb{R}^+)^d$ of the objective functional:

$$\mathcal{F}(\nu_1, \dots, \nu_{N_d}) = \sum_{j=1}^{N} \left| z_t(x_j) + \left(S_0 - e^{x_j - rT} \right)^+ - C^j \right|^2 + \alpha \sum_{h=1}^{N_d} \left[\nu_h \left(\log \nu_h - \log \nu_h^0 \right) + \nu_h^0 - \nu_h \right]. \tag{23}$$

The optimization is run under the *L-BFGS-B* algorithm, so we compute the derivatives of $\mathcal{F}$ to speed it up. We focus on deriving the first addend in (23), since calculating the entropic term is straightforward. Denote by $g\left(v_1,\ldots,v_{N_d}\right)$ the argument of the exponential in (22) and define

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

Under suitable integrability condition, for every $u \in \mathbb{R}$ and $h = 1, \ldots, N_d$ we have

$$\frac{\partial \zeta_T(u)}{\partial v_h}\left(v_1,\ldots,v_{N_d}\right) = C_u T e^{g\left(v_1,\ldots,v_{N_d}\right)}\left[e^{y_h}\left(e^{iuy_h}-1\right)+iu\left(1-e^{y_h}\right)\right]$$

$$= \frac{S_0}{\sqrt{2\pi}}\frac{T}{iu+1}e^{iu(\log S_0+rT)}\left(1-e^{y_h}\right)e^{g\left(v_1,\ldots,v_{N_d}\right)} + Te^{y_h}\zeta_T(u)\left(e^{iuy_h}-1\right)+Te^{y_h}C_u\left(e^{iuy_h}-1\right).$$

Direct calculation from (18) leads, for every $k \in \mathbb{R}$ and $h = 1, \ldots, N_d$, to

$$\frac{\partial z_T(k)}{\partial v_h}\left(v_1,\ldots,v_{N_d}\right) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} e^{-iuk}\frac{\partial \zeta_T(u)}{\partial v_h}\left(v_1,\ldots,v_{N_d}\right)du$$

$$= \frac{S_0}{2\pi}T\left(1-e^{y_h}\right)\int_{\mathbb{R}}\frac{e^{-iuk}}{iu+1}e^{iu(\log S_0+rT)}e^{g\left(v_1,\ldots,v_{N_d}\right)}du$$

$$+ Te^{y_h}\left[z_T\left(k-y_h\right)+\left(S_0-e^{k-rT-y_h}\right)^+ - z_T(k)-\left(S_0-e^{k-rT}\right)^+\right]. \tag{24}$$

Taking $k = \widetilde{x}_n$, for $n = 0, \ldots, \widetilde{N}-1$, we can approximate the first addend in (24) with an FFT:

$$\int_{\mathbb{R}}\frac{e^{-iu\widetilde{x}_n}}{iu+1}e^{iu(rT+\log S_0)}e^{g\left(v_1,\ldots,v_{N_d}\right)} \simeq \frac{A}{\widetilde{N}}\sum_{k=0}^{\widetilde{N}-1}e^{-iu_k\widetilde{x}_n}\widetilde{w}_k f(u_k) = \frac{A}{\widetilde{N}}e^{i\frac{A}{2}\widetilde{x}_n}\sum_{k=0}^{\widetilde{N}-1}\exp\left(-i\frac{2\pi n}{\widetilde{N}}k\right)\widetilde{w}_k f(u_k),$$

where

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

Summarizing, we can numerically calculate the derivatives of $\mathcal{F}$:

$$\frac{\partial \mathcal{F}}{\partial v_h}\left(v_1,\ldots,v_{N_d}\right) = 2\sum_{j=1}^{N}\left[z_T\left(x_j\right)+\left(S_0-e^{x_j-rT}\right)^+ - C^j\right]\frac{\partial z_T\left(x_j\right)}{\partial v_h}\left(v_1,\ldots,v_{N_d}\right)+\alpha\left(\log v_h - \log v_h^0\right)$$

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

The regularization parameter α should be eventually determined; to this aim, we follow a procedure loosely inspired by the Morozov discrepancy principle (see the classical Morozov (1966)). We start off by minimizing the quadratic pricing error (15), ignoring the entropy term and using the discretized Lévy measure of L^0 as starting point. The value ϵ_0 of the functional at the found minimum $\overline{v}$ is interpreted as the distance between the market and the selected model class. After that we consider the bid–ask spread of the options on the market and denote by ϵ its Euclidean norm. The absolute value of the difference between these two quantities is then allowed to be slightly greater, although it must maintain the same order of magnitude. This leads to the introduction of another term $\widetilde{\epsilon} := c\,|\epsilon_0 - \epsilon|$, where c is a positive constant to be picked. At this point the functional in (16) is strictly increasing in α, so we choose the regularization parameter as follows:

$$\alpha := \widetilde{\epsilon}\left(\int_{\mathbb{R}}\left(\frac{d\overline{v}}{dv^0}\log\frac{d\overline{v}}{dv^0}+1-\frac{d\overline{v}}{dv^0}\right)dv^0\right)^{-1}.$$

For our applications, taking $c \approx 2$ has proven to be a satisfactory choice in terms of goodness of fit.

3.1.2. Empirical Results

The final step is to empirically apply this method on prices of real–traded options. In particular, we have tested it with derivatives on the same stocks as those considered for the Esscher's method. In order to estimate the parameters of the historical Lévy process, we used the same time series of log–prices as those of the Esscher's method (see Section 2.1.2) All the call options under scrutiny expire in January 2021: 11 months from the time of simulation.

In the case of call options on Apple Inc. stock (AAPL) we obtained an average of the absolute values of percentage difference of 3.5794%: Figure 2a displays the results of such implementation. As regards call options on Microsoft Corporation stock (MSFT), the absolute values of percentage difference is 1.6425% and the results are shown in Figure 2b. Finally we calibrated our method to the prices of call options on Tesla, Inc. stock (TSLA): Figure 2c shows the outcomes. Here the mean of absolute values of percentage difference settles down at 0.9148%.

This time we retrieved the real prices from https://finance.yahoo.com, but the same considerations as Remark 3 apply.
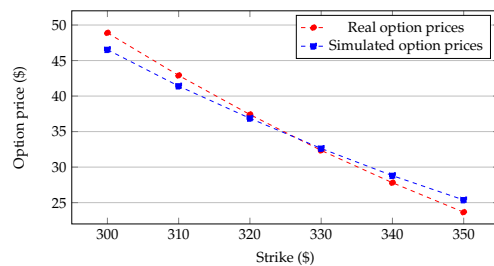
Strike ($)	Real data ($)	Model Price ($)
300	48.90	46.53
310	42.90	41.41
320	37.40	36.87
330	32.35	32.62
340	27.80	28.80
350	23.65	25.38

(**a**) Apple Inc. stock (AAPL) call option, Maturity: Jan. 2021.
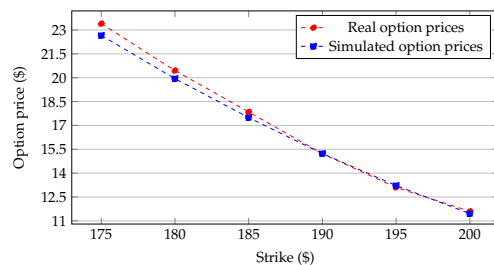
Strike ($)	Real data ($)	Model Price ($)
175	23.40	22.66
180	20.45	19.94
185	17.85	17.50
190	15.22	15.23
195	13.10	13.22
200	11.60	11.46

(**b**) Microsoft Corporation stock (MSFT) call option, Maturity: Jan. 2021.
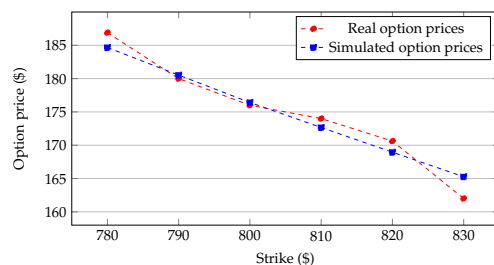
Strike ($)	Real data ($)	Model Price ($)
780	186.90	184.67
790	180.00	180.53
800	176.00	177.42
810	174.00	172.66
820	170.62	168.95
830	162.10	165.27

(**c**) Tesla, Inc. stock (TSLA) call option, Maturity: Jan. 2021.

Figure 2. Numerical experiments for the calibration method.

4. Conclusions

The main purpose of this research is to quantify the performances of two option pricing methods in a liquid market like NASDAQ. The first approach is based on the Esscher measure, the second one on the calibration to real-traded call option prices with an entropic penalty term.

Our experimental analysis shows that, from a computational point of view, the Esscher method with 10^4 Monte Carlo iterations is far faster than the calibration method: the table in Figure 3 reports the exact execution times that we have obtained with an Asus ZenBook UX31A. This fact is especially noticeable when we need to simulate the prices of options with different maturities and same strike. On the other hand, the calibration method allows generating the entire option chain for a fixed maturity with great precision. Moreover, it is more stable than the previous one, as it provides results closer to real data for a larger range of stock prices.

Comparing averages of absolute values for percentage difference between real and simulated prices, the calibration procedure offers better outcomes than Esscher's, with the only exception of AAPL call options, where the latter outperforms the former by approximately 42 basis points. 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 an appealing and efficient solution to price call options in illiquid—or low liquid—financial markets.

	Esscher's method (T_E)	Calibration method (T_C)	$\frac{T_C}{T_E}$
AAPL	1873 s	3477 s	186%
MSFT	1917 s	3782 s	197%
TSLA	1809 s	2926 s	162%

Figure 3. Computational times (in seconds) of both methods (with an Asus ZenBook UX31A).

5. Future Research

Considering the future work on the subject, one of the directions worth investigating is the Linear Esscher measure method. The importance of focusing on this approach is given by its intrinsic link to the minimal entropy Hellinger martingale measure, as indicated by Rheinländer and Sexton in (Rheinländer and Sexton 2011, Chp. 9), and, in much more detail, by Choulli and Stricker in Choulli and Stricker (2006). Moreover, it would be interesting to extend our study to other asset classes like metal futures, and to compare it with the results obtained in Chen (2011).

Author Contributions: Conceptualization, A.B. and T.R.; methodology, T.R.; software, A.B.; validation, A.B., D.R. and T.R.; formal analysis, A.B., D.R. and T.R.; investigation, A.B., D.R. and T.R.; resources, A.B., D.R. and T.R.; data curation, A.B., D.R.; writing—original draft preparation, A.B., D.R.; writing—review and editing, A.B., D.R. and T.R.; visualization, A.B., D.R.; supervision, T.R.; project administration, T.R.; funding acquisition, T.R. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Normal Inverse Gaussian Distribution

In our work, we extensively use the *NIG* distribution, because it provides a good fit for the analyzed data. A probability measure μ 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:

$$f_{NIG}(x; \alpha, \beta, \mu, \delta) = \frac{\alpha}{\pi} \exp\left(\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)\right) \frac{K_1\left(\alpha\delta\sqrt{1 + \left(\frac{x-\mu}{\delta}\right)^2}\right)}{\sqrt{1 + \left(\frac{x-\mu}{\delta}\right)^2}}, \tag{A1}$$

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 Abramowitz and Stegun (1970). The next procedure is the classical construction of such a distribution.

A probability measure μ on $(\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ is said to be 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

$$f_{GIG}(x; \nu, \delta, \gamma) = \left(\frac{\gamma}{\delta}\right)^\nu \frac{1}{2K_\nu(\gamma\delta)} x^{\nu-1} \exp\left[-\frac{1}{2}\left(\delta^2 x^{-1} + \gamma^2 x\right)\right], \quad x > 0. \tag{A2}$$

In order to show that f_{GIG} is a density function, we need the following representation for K_ν (Watson 1966, Formula (8), p. 182), the modified Bessel function of the third kind with index ν:

$$K_\nu(x) = \frac{1}{2} \int_0^\infty y^{\nu-1} \exp\left[-\frac{x}{2}\left(y + \frac{1}{y}\right)\right] dy, \quad x > 0. \tag{A3}$$

The computation of the next integral allows us to find the normalization constant in (A2):

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

where in the second equality we made the substitution $y = \frac{\gamma}{\delta} x$. Since $K_\nu > 0$ in $\mathbb{R}^+$ from (A3), we can conclude that f_{GIG} is actually a density function on $\mathbb{R}^+$.
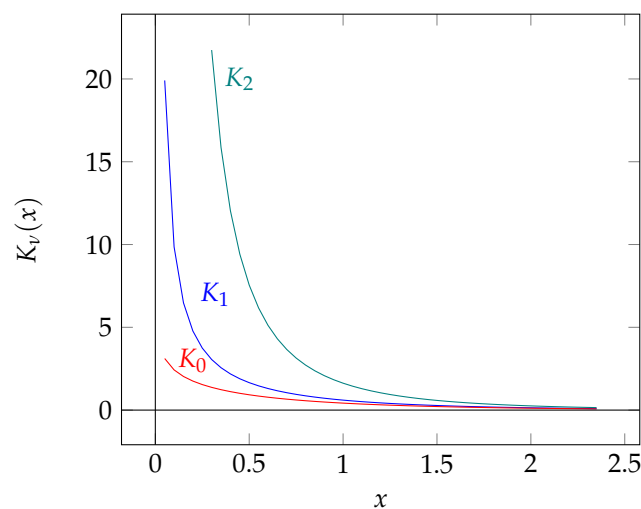

Figure A1. K_0, K_1, K_2 are the modified Bessel functions of the first three, nonnegative, integer orders, respectively.

Let us fix two other parameters $\alpha, \beta \in \mathbb{R}$ such that $0 \leq |\beta| < \alpha$. 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 $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ 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:

$$
f_{GH}(x; \nu, \alpha, \beta, \mu, \delta) := \int_0^\infty f_N(x; \mu, \beta, y) f_{GIG}\left(y; \nu, \delta, \sqrt{\alpha^2 - \beta^2}\right) dy
$$

$$
= \frac{1}{\sqrt{2\pi}} \left(\frac{\sqrt{\alpha^2 - \beta^2}}{\delta}\right)^\nu \frac{1}{2K_\nu\left(\delta\sqrt{\alpha^2 - \beta^2}\right)}
$$

$$
\int_0^\infty \frac{1}{\sqrt{y}} e^{-\frac{1}{2y}(x - \mu - \beta y)^2} y^{\nu - 1} e^{-\frac{1}{2}\left(\delta^2 y^{-1} + (\alpha^2 - \beta^2)y\right)} dy
$$

$$
= \frac{1}{\sqrt{2\pi}} \left(\frac{\sqrt{\alpha^2 - \beta^2}}{\delta}\right)^\nu \frac{1}{2K_\nu\left(\delta\sqrt{\alpha^2 - \beta^2}\right)} e^{\beta(x - \mu)}
$$

$$
\int_0^\infty y^{\nu - 1 - \frac{1}{2}} \exp\left\{-\frac{1}{2}\left[\frac{1}{y}\left(\delta^2 + (x - \mu)^2\right) + \alpha^2 y\right]\right\} dy
$$

$$
= \frac{1}{\sqrt{2\pi}} \left(\frac{\sqrt{\alpha^2 - \beta^2}}{\delta}\right)^\nu \frac{1}{2K_\nu\left(\delta\sqrt{\alpha^2 - \beta^2}\right)} e^{\beta(x - \mu)}
$$

$$
\left(\frac{\sqrt{\delta^2 + (x - \mu)^2}}{\alpha}\right)^{\nu - \frac{1}{2}} \int_0^\infty z^{\left(\nu - \frac{1}{2}\right) - 1} \exp\left[-\frac{1}{2}\left(z^{-1} + z\right) \alpha\sqrt{\delta^2 + (x - \mu)^2}\right] dz
$$

$$
= \frac{\left(\sqrt{\alpha^2 - \beta^2}\right)^\nu}{\sqrt{2\pi}\delta^\nu \alpha^{\nu - \frac{1}{2}} K_\nu\left(\delta\sqrt{\alpha^2 - \beta^2}\right)} e^{\beta(x - \mu)} \frac{K_{\nu - \frac{1}{2}}\left(\alpha\sqrt{\delta^2 + (x - \mu)^2}\right)}{\left(\sqrt{\delta^2 + (x - \mu)^2}\right)^{\frac{1}{2} - \nu}}, \quad x \in \mathbb{R}, \tag{A4}
$$

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 (A3).

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 recover the *NIG* density (A1) taking $\nu = -\frac{1}{2}$ in (A4). In fact, for every $\nu \in \mathbb{R}$, it results $K_{-\nu} = K_\nu$ and the representation (Abramowitz and Stegun (1970), Formula (9.6.23))

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

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

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

As a particular case of Generalized Hyperbolic distribution, the *NIG* probability measure is infinitely divisible (see, e.g., Barndorff-Nielsen and Halgreen 1977; Eberlein and Hammerstein 2004). The moment generating function for a random variable $X \sim NIG(\alpha, \beta, \mu, \delta)$ is given by

$$
E\left[e^{zX}\right] = \exp\left[\mu z + \delta\left(\sqrt{\alpha^2 - \beta^2} - \sqrt{\alpha^2 - (\beta + z)^2}\right)\right], \quad -\alpha - \beta \leq z \leq \alpha - \beta, \tag{A5}
$$

and its generating triplet (see Appendix B for the definition of this concept) is given by

$$\begin{cases} \sigma^2 = 0 \\ \nu\,(dx) = \dfrac{\delta\alpha}{\pi\,|x|}e^{\beta x}K_1\,(\alpha\,|x|)\,dx \\ \gamma = \mu + \dfrac{2\delta\alpha}{\pi}\displaystyle\int_0^1 \sinh\,(\beta x)\,K_1\,(\alpha x)\,dx \end{cases}. \tag{A6}$$

For an explicit calculation, we refer to the paper Barndorff-Nielsen (1998).

Appendix B. Cumulant Function of Lévy Processes

Fix a probability space $(\Omega, \mathcal{F}, P)$. It is well known that, if $X = \{X_t\}_t$ is a Lévy process, then X_t is infinitely divisible for every $t \in \mathbb{R}_0^+$. On the other hand, for every infinitely divisible probability measure μ on $\mathbb{R}^d$, there exists a Lévy process $X = \{X_t\}_t$, which is unique up to identity in law, such that $X_1 \sim \mu$.

The Lévy–Khintchine representation theorem states that, given an infinitely divisible distribution μ on $\mathbb{R}^d$, its characteristic function $\hat{\mu}$ can be written as:

$$\hat{\mu}\,(z) = \exp\left[-\frac{1}{2}\langle z, Az\rangle + i\langle \gamma, z\rangle + \int_{\mathbb{R}^d}\left(e^{i\langle z,x\rangle} - 1 - i\langle z, x\rangle 1_D\,(x) \right)\nu\,(dx) \right], \quad z \in \mathbb{R}^d, \tag{A7}$$

where $D := \left\{ x \in \mathbb{R}^d : |x| \leq 1 \right\}$, $\gamma \in \mathbb{R}^d$, ν is a measure on $\mathbb{R}^d$ satisfying

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

and A is a symmetric, positive semidefinite $d \times d$ matrix. The representation (A7) of $\hat{\mu}$ by (A, ν, γ) is unique and the triplet (A, ν, γ) is called generating triplet of the distribution μ. The generating triplet of a Lévy process $X = \{X_t\}_t$ is the generating triplet of X_1.

Note that it is not necessary to take 1_D to have integrability in (A7). In fact, let $h : \mathbb{R}^d \to \mathbb{R}^d$ be a bounded function, with $h\,(x) = x$ in a neighborhood of 0: we call it a truncation function. Since, for every $z \in \mathbb{R}^d$, in a neighborhood of 0 we have $\left| e^{i\langle z,x\rangle} - 1 - i\langle z, h\,(x)\rangle \right| \leq \frac{1}{2}|z|^2\,|x|^2$ and, further, $\left| e^{i\langle z,x\rangle} - 1 - i\langle z, h\,(x)\rangle \right| \leq 2 + C\,|z|$, $x \in \mathbb{R}^d$, for some positive constant C such that $h \leq C$ in $\mathbb{R}^d$, if we put $\gamma_h := \gamma + \int_{\mathbb{R}^d}(h\,(x) - x1_D\,(x))\nu\,(dx)$ component wise, then the Lévy–Khintchine formula with respect to h becomes:

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

We note that only γ_h depends on the choice of the truncation function.

It is possible to "extend" the argument of the exponential function in (A7), called *characteristic exponent* and denoted by ψ, to a subset of $\mathbb{C}^d$. Precisely, taking into account the set

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

we can define the function $\Psi : \widetilde{D} \to \mathbb{C}$ as follows:

$$\Psi\,(w) := \frac{1}{2}\langle w, Aw\rangle + \langle \gamma, w\rangle + \int_{\mathbb{R}^d}\left(e^{\langle w,x\rangle} - 1 - \langle w, x\rangle 1_D\,(x) \right)\nu\,(dx), \quad w \in \widetilde{D}, \tag{A8}$$

where $\widetilde{D} := \left\{ w \in \mathfrak{C}^d : \mathfrak{Re}\,(w) \in C \right\}$. We readily note that $iz \in \widetilde{D}$ for every $z \in \mathbb{R}^d$ and the following equality holds: $\Psi\,(iz) = \psi\,(z)$, $z \in \mathbb{R}^d$. In this sense Ψ extends ψ to a subset of $\mathbb{C}^d$. We pinpoint that,

in this paper, we set $\langle u, w \rangle := \sum_{j=1}^{d} u_j w_j$, $u, w \in \mathbb{C}^d$, so $\langle \cdot, \cdot \rangle$ is not the $\mathbb{C}^d$-Hermitian inner product. Theorem 25.17 in Sato (1999) affirms that, if $X = \{X_t\}_t$ is a Lévy process generated by (A, ν, γ), then

$$E\left[e^{\langle w, X_t \rangle}\right] = e^{t\Psi(w)} \quad \text{for any } w \in \widetilde{D}, \ t > 0 : \tag{A9}$$

we call Ψ the cumulant function of X. As we have previously done, we can introduce a truncation function h and express $\Psi = \Psi_h$ in $\widetilde{D}$, where

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

Remark A1. *Let $X = \{X_t\}_t$ be a $\mathbb{R}^d$-valued Lévy process defined on a probability space $(\Omega, \mathcal{F}, P)$, Ψ be its cumulant function and $\theta \in \mathbb{R}^d$ be such that $E\left[\exp\left(\langle \theta, X_t \rangle\right)\right] < \infty$ for some $t > 0$. By Theorem 25.3 in Sato (1999), this is equivalent to require that $\theta \in C$. We introduce the random variables*

$$M_t := \exp\left(\langle \theta, X_t \rangle - t\Psi(\theta)\right), \quad t \geq 0.$$

Note that $M = \{M_t\}_t$ is integrable, by assumption and the fact that $t\Psi(\theta)$ is constant in Ω for every $t \geq 0$. Let us construct the minimal augmented filtration $\mathbb{F} = (\mathcal{F}_t)_{t \geq 0}$ of X, i.e., $\mathcal{F}_t = \sigma\left(\mathcal{N} \cup \mathcal{F}_t^0\right)$ for any $t \geq 0$, where $\left(\mathcal{F}_t^0\right)_t$ is the natural filtration of the process and $\mathcal{N}$ the collection of $\mathcal{F}$-negligible sets; obviously, M is $\mathbb{F}$-adapted. According to (Protter 2005, Chapter I, Theorem 31) $\mathbb{F}$ is right-continuous, so it satisfies the usual conditions, as well. If we fix $t > s \geq 0$, then using (A9) and the properties of the Lévy-increments we see that M is a martingale with mean 1:

$$E\left[M_t | \mathcal{F}_s\right] = E\left[e^{\langle \theta, X_t \rangle - t\Psi(\theta)} \Big| \mathcal{F}_s\right] \stackrel{a.s.}{=} e^{\langle \theta, X_s \rangle - s\Psi(\theta)} E\left[e^{\langle \theta, X_t - X_s \rangle} \Big| \mathcal{F}_s\right] e^{-(t-s)\Psi(\theta)}$$

$$\stackrel{a.s.}{=} M_s E\left[e^{\langle \theta, X_{t-s} \rangle}\right] e^{-(t-s)\Psi(\theta)} = M_s.$$

Definition A1. *Fix $T > 0$ and let $\theta \in C$. The probability measure P^θ 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 θ. The density process is given by $\frac{dP^\theta|_{\mathcal{F}_t}}{dP|_{\mathcal{F}_t}} = M_t$, $t \in [0, T]$.*

Appendix C. 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 (Shiryaev and Jacod 2003, Chp. II).

We start off by fixing a stochastic basis $(\Omega, \mathcal{F}, P; \mathbb{F})$, with $\mathbb{F}$ which satisfies the usual conditions. Given two stopping times S, T, the stochastic interval is the random set

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

and $[\![T]\!] := [\![T, T]\!]$.

Definition A2. *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 $\{t \in \mathbb{R}_0^+ : (t, \omega) \in A\}$, for $\omega \in \Omega$, are at most countable when A is a thin set.

Theorem A1. *If $X = \{X_t\}_t$ is a RCLL, adapted process, then the random set $\{\Delta X \neq 0\}$ is thin.*

We refer to (He et al. 2018, Theorem 3.32) for a proof. Thanks to this result we can easily introduce the next concept.

Definition A3. *Let* $X = \{X_t\}_t$ *be an adapted, RCLL,* $\mathbb{R}^d$*-valued process. The measure* μ^X *on* $\mathbb{R}_0^+ \times \mathbb{R}^d$ *defined by*

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

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

We now consider a process $X = \{X_t\}_t$ which is a d-dimensional semimartingale, highlighting that we restrict our attention to RCLL and $\mathbb{F}$-adapted semimartingales. Let h be a truncation function; using the measure μ^X it is possible to derive the characteristics of X (see Shiryaev and Jacod 2003, Chp. II, Definition 2.6), which we denote by the triplet (B, C, ν^X). They are defined up to a P-null set and only B depends on the choice of h.

Every Lévy process—once we endow the probability space in which it is defined with its minimal augmented filtration—is a *PIIS* process (a RCLL, adapted process which starts at 0 and has independent and stationary increments), which in turn is a semimartingale. Thus, considering a Lévy process $L = \{L_t\}_t$ with generating triplet (A, ν, γ_h) relative to a truncation function h, we can apply (Shiryaev and Jacod 2003, Chp. II, Corollary 4.19) together with the uniqueness of the Lévy–Khintchine representation to express its characteristics as $(\gamma_h t, At, dt \otimes \nu(dx))$.

Appendix D. Laplace Cumulant and Geometric Esscher Measure

On a stochastic basis $(\Omega, \mathcal{F}, P; \mathbb{F})$, with the filtration $\mathbb{F}$ which satisfies the usual conditions of right-continuity and completeness, we define a d–dimensional semimartingale $X = (X^1, \dots, X^d)$, i.e., a process which can be decomposed as $X = M + A$, a sum of a local martingale M and a process A of locally finite variation which is called drift, or additive compensator, of the semimartingale, and with characteristics (B, C, ν^X) relative to a truncation function h (see Appendix B for the definition). In particular, a special semimartingale is a semimartingale where its finite variation part is predictable; such a special semimartingale has a unique semimartingale decomposition. In view of (Shiryaev and Jacod 2003, Chp. II, Proposition 2.9) we can assume, without loss of generality, that $\nu^X(\{t\} \times \mathbb{R}^d) \leq 1$ identically. An exposition on semimartingale characteristics is outside of the scope of this paper, for this we refer to Shiryaev and Jacod (2003).

Now we introduce the stochastic logarithm and the stochastic exponential, denoted by $\mathcal{L}$ and $\mathcal{E}$, respectively. The symbol $\cdot$ denotes stochastic integration, the exact meaning can vary according to the context. Let X be a semimartingale; the set of stochastic processes which are integrable with respect to X is denoted by $L(X)$. This notion is quite intricate, and it is not the purpose of this article to give an overview on the theory of stochastic integration. For this, we refer to (Protter 2005, Chp. IV, Section 2). The solution of the stochastic integral equation

$$Y = 1 + Y_- \cdot X$$

is denoted as the stochastic exponential $\mathcal{E}(X)$ of the semimartingale X. Here, the process $Y_- = \{Y_-(t)\}_t$ is defined by $Y_-(t) := Y_{t-} = \lim_{s \to t^-} Y_s$ for $t > 0$, with $Y_-(0) := Y_0$. Conversely, if Y is a semimartingale such that both Y and Y_- do not vanish, then the process

$$X = \frac{1}{Y_-} \cdot Y,$$

denoted by $\mathcal{L}(Y)$, is called the stochastic logarithm of Y, and is the unique semimartingale such that

$$Y = Y_0\, \mathcal{E}(X).$$

For a detailed exposition of these concepts we refer to (Shiryaev and Jacod 2003, Chp. II, Section 8)

Definition A4. *Let $\theta \in L(X)$ be an admissible strategy such that $\theta \cdot X$ is exponentially special, i.e., $\exp(X)$ is a special semimartingale. The Laplace cumulant $\widetilde{K}^X(\theta)$ of X at θ is the additive compensator of the real-valued, special semimartingale*

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

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

$$K^X(\theta) := \log\left(\mathcal{E}\left(\widetilde{K}^X(\theta)\right)\right). \tag{A10}$$

Note that the additive compensator of the process

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

is predictable because $\exp(\theta \cdot X)$ is special and $\frac{1}{(\exp(\theta \cdot X))_-}$ is predictable. Therefore, $\mathcal{L}(\exp(\theta \cdot X))$ is a special semimartingale and the definition of $\widetilde{K}^X(\theta)$ is well posed. As far as $K^X(\theta)$ is concerned, recalling (Shiryaev and Jacod 2003, Chp. III, Theorem 7.4) we have

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

Thus, the stochastic exponential in (A10) is strictly positive and the process $K^X(\theta)$ is well defined, as well.

If $\theta \cdot X$ is exponentially special, then $K^X(\theta)$ is its exponential compensator, see (Shiryaev and Jacod 2003, Chp. III, Theorem 7.14). This means that the process $Z^\theta = \{Z_t^\theta\}_t$, defined by $Z_t^\theta := \exp(\theta \cdot X_t - K^X(\theta)_t)$, $t \geq 0$, is a local martingale starting at $Z_0^\theta = 1$. We now recall the concept of uniform integrability.

Definition A5. *A non-empty set Φ of real-valued random variable defined on a probability space $(\Omega, \mathcal{F}, P)$ is uniformly integrable if*

$$\lim_{n \to \infty} \sup_{X \in \Phi} E\left[|X| \mathbf{1}_{\{|X| \geq n\}}\right] = 0.$$

Supposing that Z^θ is a uniformly integrable martingale, then we can set $P^\theta(d\omega) := Z_\infty^\theta P(d\omega)$, where $Z_t \to Z_\infty$ a.s. and in L^1 as $t \to \infty$. It is straightforward to show that Z^θ is the density of P^θ relative to P; besides, these two distributions are locally equivalent as $Z_t^\theta > 0$, $t \geq 0$.

The following result (Shiryaev and Jacod 2003, Chp. III, Theorem 7.18) establishes a necessary and sufficient condition such that the process $S^i := S_0^i \exp(X^i)$, $S_0^i \in \mathbb{R}^+$ is a P^θ-local martingale for every $i = 1, \ldots, d$.

Theorem A2. *Let $\theta \in L(X)$ be such that $\theta \cdot X$ is exponentially special and Z^θ is a uniformly integrable martingale. Define*

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

Then the processes $S^i = S_0^i \exp(X^i)$ are P^θ-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 every $i = 1, \ldots, d$.

We call P^θ geometric Esscher measure, or *Esscher martingale transform* for exponential processes. For $d = 1$, in case the geometric Esscher measure exists, (Kallsen and Shiryaev 2002, Theorem 4.2)

provides its uniqueness: this means that, if we find another process $\tilde{\theta} \in L(X)$ such that $S = S_0 \exp(X)$ is a $P^{\tilde{\theta}}$-local martingale, then $P^{\tilde{\theta}} = P^{\theta}$.

We conclude with some technical results. Let $\theta \in L(X)$ be such that $\theta \cdot X$ is exponentially special. By (Shiryaev and Jacod 2003, Chp. III, Theorem 7.4) and (Shiryaev and Jacod 2003, Chp. II, Proposition 2.9) we can write $\widetilde{K}^X(\theta) = \widetilde{\kappa}(\theta) \cdot A$, where

$$\widetilde{\kappa}(\theta)_t := \langle \theta_t, b_t \rangle + \frac{1}{2} \langle \theta_t, c_t \theta_t \rangle + \int_{\mathbb{R}^d} \left(e^{\langle \theta_t, x \rangle} - 1 - \langle \theta_t, h(x) \rangle \right) F_t(dx), \quad t \geq 0. \tag{A11}$$

The next lemma (Kallsen and Shiryaev 2002, Lemma 2.11) allows us to express the drift process as a function of the characteristics of the semimartingale.

Lemma A1. *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*

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

Finally, we use Girsanov's theorem to compute the characteristics $\left(B^{\theta}, C^{\theta}, \nu^{X^{\theta}} \right)$ of X under P^{θ}—provided its existence—relative to the same h:

$$\begin{cases} B^{\theta^i} = B^i + c^{i \cdot} \theta \cdot A + h^i(x) \left(\dfrac{e^{\langle \theta_t, x \rangle}}{1 + \hat{W}(\theta)_t} - 1 \right) \star \nu^X, \quad i = 1, \ldots, d \\[2mm] C^{\theta} = C \\[2mm] \nu^{X^{\theta}}(dt, dx) = \dfrac{e^{\langle \theta_t, x \rangle}}{1 + \hat{W}(\theta)_t} \nu^X(dt, dx) \end{cases} \tag{A12}$$

where $\hat{W}(\theta)_t := \int_{\mathbb{R}^d} \left(e^{\langle \theta_t, x \rangle} - 1 \right) \nu^X(\{t\} \times dx)$, $t \geq 0$. Note that $\hat{W}(\theta)_t = \Delta \widetilde{K}^X(\theta)_t > -1$, $t \geq 0$.

Remark A2. *For $\omega \in \Omega$, $t \geq 0$ and $G \in \mathcal{B}\left(\mathbb{R}^d \right)$, we have:*

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

$$= \int_{\{t\}} dA_s(\omega) \int_{\mathbb{R}^d} \mathbb{1}_G(x) F_{(t, \omega)}(dx) = F_{(t, \omega)}(G) \int_{\{t\}} dA_s(\omega).$$

Moreover, as $\int_{\{t\}} dA_s(\omega) = A_t(\omega) - A_{t-}(\omega)$, if the function $A_{\cdot}(\omega)$ is continuous in t, then $\nu^X(\omega; \{t\} \times dx)$ is the null measure on $\mathcal{B}\left(\mathbb{R}^d \right)$, which implies that $\hat{W}(\theta)_t(\omega) = 0$.

Appendix E. Lévy Processes on Skorohod Space and Relative Entropy of Distributions

Let $(\Omega, \mathcal{F}, P)$ be a probability space which carries a $\mathbb{R}^d$-valued, additive process $\{X_t\}_t$ with system of generating triplets $\{(A_t, \nu_t, \gamma_t)\}_t$. For every $t \geq 0$ we define $x_t : \mathfrak{D} \to \mathbb{R}^d$, where $\mathfrak{D}$ is the *Skorohod Space*, as $x_t(\xi) := \xi(t)$, $\xi \in \mathfrak{D}$, and we introduce the σ-algebra $\mathcal{F}_{\mathfrak{D}} := \sigma(\{x_t, t \geq 0\})$. Now we are in the position to define a natural filtration $\mathbb{F} := (\mathcal{F}_t)_{t \geq 0}$ on the measurable space $(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}})$, where $\mathcal{F}_t := \sigma(\{x_s, 0 \leq s \leq t\})$, $t \geq 0$. Since $\{X_t\}_t$ is an RCLL process, we set the map $\phi : \Omega \to \mathfrak{D}$ as $\phi(\omega) := X_{\cdot}(\omega)$, $\omega \in \Omega$, where $X_{\cdot}(\omega) : [0, \infty) \to \mathbb{R}^d$, with $X_{\cdot}(\omega)(t) := X_t(\omega)$ for any $t \geq 0$. The function ϕ is $\mathcal{F} / \mathcal{F}_{\mathfrak{D}}$ measurable, since

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

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

$$\phi^{-1}\left(x_t^{-1}(B)\right) = \left\{\omega \in \Omega : \phi(\omega) \in x_t^{-1}(B)\right\} = \left\{\omega \in \Omega : X.(\omega) \in x_t^{-1}(B)\right\}$$
$$= \left\{\omega \in \Omega : X_t(\omega) \in B\right\} = X_t^{-1}(B) \in \mathcal{F}.$$

This enables us to construct the pushforward measure $P^{\mathfrak{D}}$ on $(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}})$, that is,

$$P^{\mathfrak{D}}(A) := P\phi^{-1}(A) = P\left(\phi^{-1}(A)\right), \quad A \in \mathcal{F}_{\mathfrak{D}}. \tag{A13}$$

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

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

for some $t_1 < t_2 < \ldots < t_n$, $B_1, \ldots, B_n \in \mathcal{B}\left(\mathbb{R}^d\right)$ and $n \in \mathbb{N}$. By (A13), we have

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

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 $\{X_t\}_t$. Specifically, if $\{X_t\}_t$ were a Lévy process, then $\{x_t\}_t$ would inherit the temporal homogeneity, so it would be a Lévy process, as well.

Consider two Lévy processes $(\{x_t\}_t, P)$ and $(\{x_t\}_t, P')$ on the Skorohod space $(\mathfrak{D}, \mathcal{F}_{\mathfrak{D}})$ endowed with the filtration $\mathbb{F}$. The next theorem (Sato 1999, Theorems 33.1 & 33.2) provides us with conditions which ensure $P\big|_{\mathcal{F}_t} \sim P'\big|_{\mathcal{F}_t}$ for any $t > 0$.

Theorem A3. *Let $(\{x_t\}_t, P)$, $(\{x_t\}_t, P')$ be Lévy processes on $\mathbb{R}^d$ with generating triplets (A, ν, γ) and (A', ν', γ'), 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:* $A = A'$, $\nu \sim \nu'$.

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

$$\int_{\mathbb{R}^d}\left(e^{\phi(x)/2} - 1\right)^2 \nu(dx) < \infty, \qquad \gamma' - \gamma - \int_{|x| \leq 1} x\,(\nu' - \nu)(dx) \in \left\{Ay, \, y \in \mathbb{R}^d\right\}.$$

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 satisfies the following properties:

(i) *U is a P-Lévy process on $\mathbb{R}$ with generating triplet*

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

(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 $t > 0$.*

An explicit expression of the process U, which is unique up to identity in law, can be retrieved in (Sato 1999, Theorem 33.2).

Definition A6. *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*

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

Recall that given two distributions P, P' on $(\Omega, \mathcal{F})$ we have $H(P, P') \geq 0$, with equality if and only if $P = P'$.

We finally present a theorem which explicitly computes the relative entropy of two equivalent Lévy processes in the Skorohod space as a function of their generating triplets.

Theorem A4. *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, ν, γ) and (A', ν', γ'), respectively. Suppose that $P\big|_{\mathcal{F}_t} \sim P'\big|_{\mathcal{F}_t}$ for every $t > 0$ and choose $\eta \in \mathbb{R}^d$ such that*

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

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

$$H\left(P'\big|_{\mathcal{F}_T}, P\big|_{\mathcal{F}_T}\right) = \frac{T}{2}\langle \eta, A\eta \rangle + T \int_{\mathbb{R}^d} \left(\frac{d\nu'}{d\nu} \log \frac{d\nu'}{d\nu} + 1 - \frac{d\nu'}{d\nu}\right) d\nu. \tag{A14}$$

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

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

We introduce the moment generating function $M_{U_T}(z) := E^P\left[e^{zU_T}\right] = e^{T\Psi(z)}$, $z \in (0, 1)$, where Ψ is the cumulant function of the Lévy process U, i.e.,

$$\Psi(z) = \frac{1}{2}\sigma_U^2 z^2 + \gamma_U z + \int_{\mathbb{R}} (e^{zx} - 1 - zx\mathbf{1}_D(x)) \nu_U(dx), \quad z \in (0, 1), \tag{A15}$$

and the last equality is ensured by (A9) in Appendix B. Actually M_{U_T} is well defined for $z = 1$ too, with $M_{U_T}(1) = E^P[e^{U_T}] = e^{T\Psi(1)} = 1$, by A3 in Theorem A3. We can see that M_{U_T} is differentiable in $(0, 1)$ with $M_{U_T}'(\cdot) = E^P[U_T e^{\cdot U_T}]$. Indeed, for every $z \in (0, 1)$, $M_{U_T}(z) = \int_{\mathbb{R}} e^{zx} P_{U_T}(dx)$ and we can derive under integral sign since

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

At this point the dominated convergence theorem readily shows that

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

On the other hand, we introduce the function $f(z) := e^{T\Psi(z)}$, $z \in (0, 1)$.

Even in this case we can affirm that f is differentiable in its domain, with derivative provided by $f' = Te^{T\Psi}\Psi'$. In particular the following equality is true:

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

In fact, we can derive under the integral sign in (A15) as, for every $z \in (0, 1)$, it results that $|xe^{zx} - x| \leq 1 + e$, $x \in D$, with $xe^{zx} - x \leq Cx^2$ in a neighborhood of 0 for some constant $C > 0$

which is independent of z. Moreover, $|x| e^{zx} \le g(x)$ for $|x| > 1$, with $\int_{|x|>1} |x| e^{|x|} \nu_U(dx) < \infty$ by Theorem 25.3 in Sato (1999). Applying another time the Lebesgue's convergence theorem we arrive at

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

Therefore

$$E^P \left[U_T e^{U_T} \right] = T e^{T\Psi(1)} \left(\sigma_U^2 + \gamma_U + \int_{\mathbb{R}} (x e^x - x \mathbf{1}_D(x)) \, \nu_U(dx) \right)$$

$$= \frac{T}{2} \langle \eta, A\eta \rangle + T \int_{\mathbb{R}^d} \left(\frac{d\nu'}{d\nu} \log \frac{d\nu'}{d\nu} + 1 - \frac{d\nu'}{d\nu} \right) d\nu,$$

using the expression of $(\sigma_U^2, \nu_U, \gamma_U)$ in A3 of Theorem A3. Since $\dfrac{dP'\big|_{\mathcal{F}_t}}{dP\big|_{\mathcal{F}_t}} = e^{U_t}$ for every $t > 0$ (see A3 in Theorem A3) and

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

we obtain (A14) and complete the proof. $\square$

Remark A3. *In the case of two $\mathbb{R}$-valued Lévy processes $(\{x_t\}_t, P)$, $(\{x_t\}_t, P')$ with generating triplets (σ^2, ν, γ) and $\left(\sigma^{2'}, \nu', \gamma' \right)$, respectively, under the hypothesis of the previous theorem (A14) reduces to*

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

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 have

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

restoring (Cont and Tankov 2004a, Proposition 9.10).

References

Abramowitz, Milton, and Irene A. Stegun. 1970. *Handbook of Mathematical Functions: with Formulas, Graphs, and Mathematical Tables.* Washington, DC: US Government Printing Office, vol. 55.

Asmussen, Søren, and Jan Rosiński. 2001. Approximations of Small Jumps of Lévy Processes with a View towards Simulation. *Journal of Applied Probability* 38: 482–93. [CrossRef]

Barndorff-Nielsen, Ole E. 1998. Processes of Normal Inverse Gaussian Type. *Finance and Stochastics* 2: 41–68. [CrossRef]

Barndorff-Nielsen, Ole, and Christian Halgreen. 1977. Infinite Divisibility of the Hyperbolic and Generalized Inverse Gaussian Distributions. *Probability Theory and Related Fields* 38: 309–12. [CrossRef]

Benth, Fred Espen, and Maren Diane Schmeck. 2014. Pricing futures and options in electricity markets. *The Interrelationship Between Financial and Energy Markets.* Berlin/Heidelberg: Springer, pp. 233–60.

Benth, Fred Espen, Jurate Saltyte Benth, and Steen Koekebakker. 2008. *Stochastic Modelling of Electricity and Related Markets.* Singapore: World Scientific, vol. 11.

Carr, Peter, and Dilip Madan. 1999. Option Valuation using the Fast Fourier Transform. *Journal of Computational Finance* 2: 61–73. [CrossRef]

Chen, Dan. 2011. Three Essays on Pricing and Hedging in Incomplete Markets. Ph.D. thesis, The London School of Economics and Political Science (LSE), London.

Choulli, Tahir, and Christophe Stricker. 2006. More on minimal entropy–Hellinger martingale measure. *Mathematical Finance* 16: 1–19. [CrossRef]

Cont, Rama, and Peter Tankov. 2004a. *Financial Modelling with Jump Processes*. Boca Raton: Chapman and Hall/CRC.

Cont, Rama, and Peter Tankov. 2004b. Nonparametric calibration of jump-diffusion option pricing models. *Journal of Computational Finance, Incisive Media* 7: 1–49. [CrossRef]

Eberlein, Ernst, and Ernst August V. Hammerstein. 2004. Generalized hyperbolic and inverse Gaussian distributions: limiting cases and approximation of processes. In *Seminar on Stochastic Analysis, Random Fields and Applications IV*. Berlin: Springer, pp. 221–64.

Esscher, F. 1932. On the Probability Function in the Collective Theory of Risk. *Skandinavisk Aktuarietidskrift* 15: 175–95.

Gerber, Hans U., and Elias SW Shiu. 1994. Option Pricing by Esscher Transforms. *Transactions of the Society of Actuaries* 46: 99–191.

Hell, Philipp, Thilo Meyer-Brandis, and Thorsten Rheinländer. 2012. Consistent factor models for temperature markets. *International Journal of Theoretical and Applied Finance* 15: 1250027. [CrossRef]

He, Sheng-wu, Jia-gang Wang, and Jia-an Yan. 2018. *Semimartingale Theory and Stochastic Calculus*. New York: Routledge.

Iacus, Stefano Maria. 2011. *Option Pricing and Estimation of Financial Models with R*. New York: John Wiley and Sons.

Jeanblanc, Monique, Marc Yor, and Marc Chesney. 2009. *Mathematical Methods for Financial Markets*. Berlin: Springer.

Kallsen, Jan, and Albert N. Shiryaev. 2002. The Cumulant Process and Esscher's Change of Measure. *Finance and Stochastics* 6: 397–428. [CrossRef]

Klüppelberg, Claudia, Alexander Lindner, and Ross Maller. 2004. A Continuous Time GARCH Process Driven by a Lévy Process: Stationarity and Second Order Behaviour. *Journal of Applied Probability* 41: 601–622. [CrossRef]

Lee, Young, and Thorsten Rheinländer. 2012. Optimal martingale measures for defaultable assets. *Stochastic Processes and their Applications* 122: 2870–84. [CrossRef]

Lewis, PA W., and Gerald S. Shedler. 1979. Simulation of Nonhomogeneous Poisson Processes by Thinning. *Naval Research Logistics Quarterly* 26: 403–13. [CrossRef]

Morozov, Vladimir Alekseevich. 1966. On the solution of functional equations by the method of regularization. *Doklady Akademii Nauk* 167: 510–12.

Protter, P. E. 2005. *Stochastic Integration and Differential Equations*, 2nd ed. Stochastic Modelling and Applied Probability. Berlin: Springer, vol. 21.

Rheinländer, Thorsten, and Jenny Sexton. 2011. *Hedging Derivatives*. Singapore: World Scientific, vol. 15.

Rudin, Walter. 1987. *Real and Complex Analysis*, 3rd ed. Mathematics Series. New York: McGraw-Hill Higher Education.

Sato, Ken-Iti. 1999. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge: Cambridge University Press.

Shiryaev, Albert, and Jean J. Jacod. 2003. *Limit Theorems for Stochastic Processes*. A Series of Comprehensive Studies in Mathematics. Berlin: Germany, vol. 288.

Shreve, Steven E. 2004. *Stochastic Calculus for Finance II (Continuous–Time Models)*. Springer Finance Textbooks. Berlin: Springer, vol. 11.

Van Heerwaarden, Angela E., Rob Kaas, and Marc J. Goovaerts. 1989. Properties of the Esscher premium calculation principle. *Insurance Mathematics and Economics 8.4* 335: 261–67. [CrossRef]

Watson, George Neville. 1966. *A Treatise on the Theory of Bessel Functions*, 2nd ed. Cambridge: Cambridge University Press.