

Lévy Models Robustness and Sensitivity

Fred Espen Benth *, Giulia Di Nunno *, and Asma Khedher *

Oslo, July 9th 2009

Abstract

We study the robustness of the sensitivity with respect to parameters in expectation functionals with respect to various approximations of a Lévy process. As sensitivity parameter, we focus on the *delta* of an European option as the derivative of the option price with respect to the current value of the underlying asset. We prove that the delta is stable with respect to natural approximations of a Lévy process, including approximating the small jumps by a Brownian motion. Our methods are based on the density method, and we propose a new *conditional density method* appropriate for our purposes. Several examples are given, including numerical examples demonstrating our results in practical situations.

1 Introduction

The dynamics of asset prices seems to be well modeled by Lévy processes and most of current research in mathematical finance is focused around this class (see e.g. Cont and Tankov [6]). When modeling the asset dynamics, it can be difficult to determine the most appropriate process among the wide variety in the Lévy family. For example, some authors promote the generalized hyperbolic class (see for instance Eberlein [8] and Barndorff-Nielsen [2]), while others prefer the CGMY model (see Carr *et al.* [4]). Furthermore, it is a philosophical question whether asset prices are driven by pure-jump processes, or if there is a diffusion in the non-Gaussian dynamics (see e.g. Eberlein and Keller [9] for a discussion). From a statistical point of view it may be very hard to determine whether a model should have a diffusion term or not.

In this paper we are dealing with the robustness of expectation functionals of models chosen within the Lévy family. In particular, we study the sensitivity with respect to parameters in the functionals with respect to various approximations of a Lévy process. As sensitivity parameter, we choose to focus on the *delta* of an European option as the derivative of the option price with respect to the current value of the underlying asset.

There are several methods of computation of the delta, but in this presentation we consider the so-called *density method* which is proper of Lévy models admitting a probability density absolutely continuous with respect to the Lebesgue measure. This method was introduced in Broadie and Glasserman [3] for Brownian models, but it is

*Centre of Mathematics for Applications and Department of Mathematics. University of Oslo. P.O. Box 1053 Blindern, N-0316 Oslo, Norway. Emails: fredb@math.uio.no, g.d.nunno@cma.uio.no, asma.khedher@cma.uio.no

not restricted to this case. Following the density method, the computation of the delta leads to a formula of the type:

$$\Delta := \frac{\partial}{\partial s_o} \mathbb{E} [g(S^{s_o}(T))] = \mathbb{E} [g(S^{s_o}(T)) \times \Pi] , \quad (1.1)$$

where $(S^{s_o}(t))_{t \geq 0}$, $(S(0) = s_o > 0)$, is the discounted price of the asset and $g(S^{s_o}(T))$ is the payoff of a European option with maturity $T > 0$. The random variable Π is called *weight*. The underlying idea is to move the differentiation with respect to the parameter s_o to the density function, thus the weight Π is given by the log-derivative of the density. Since most asset price models are defined as the exponential of a Lévy process (see for instance Eberlein [8]), we choose to work in this paper with expectation functionals of the form

$$F(x) = \mathbb{E} [f(x + L(T))] .$$

We easily recover the above by letting $x = \ln s_o$ and $f(y) = g(e^y)$. The application of the density method will yield a weight Π which is the log-derivative of the density of $L(T)$.

As a variation to this method we introduce the *conditional density method* which allows some flexibility in the computation when dealing with Lévy models not of Brownian nature. The conditional density method relies on the observation that we may use conditioning in order to separate out differentiable density in the expectation function. More precisely, if we have a random variable which may be represented as a sum of two independent random variable, where one possesses a differentiable density, we may use conditional expectation and the “classical” density approach to move the differentiation to this density. We recall from the Lévy-Kintchine representation of Lévy processes that any Lévy process can be represented as a pure-jump process and an independent drifted Brownian component. The application of the conditional density method provides *different* weights than the density method. The fact that the weights are not unique is well-known, as this appears also by application to other methods of computations, e.g. the so-called Malliavin methods. We stress that the delta is in any case the same, only the computation method is different. It is well-known that the density method provides an expression for the delta which has minimal variance. This is the meaning of *optimality* for weights. The weights derived by the conditional density method are not optimal.

From the point of view of robustness to model choice, our point of departure is the paper of Asmussen and Rosinski [1], where it is proven that the small jumps of a Lévy process $(L(t))_{t \geq 0}$ can be approximated by a Brownian motion scaled with the standard deviation of the small jumps, that is,

$$L(t) \approx \sigma(\varepsilon)B(t) + N^\varepsilon(t),$$

where N^ε is a Lévy process with jumps bigger than ε and $(B_t)_{t \geq 0}$ is an independent Brownian motion. Note that if the Lévy process has a continuous martingale part, N^ε includes it. The function $\sigma(\varepsilon)$ is the standard deviation of the jumps smaller than ε of the Lévy process, which can be computed as the integral of z^2 with respect to the Lévy measure in a ball of radius ε . Obviously, $\sigma(\varepsilon)$ tends to zero with ε . In fact, this

approximative Lévy process converges in distribution to the original one. Based on this approximation, which is popular when simulating the paths of different Lévy processes like the normal inverse Gaussian (see Rydberg [13]), we investigate the relationship of the deltas derived from the two respective models. From the results of Asmussen and Rosinski [1], we know that the respective option prices converge when ε goes to zero. The question is if the same holds true for the deltas, and in this paper we show that this is indeed the case.

In itself it is maybe not a priori surprising that the deltas are robust with respect to Lévy models which are approximately equal, but it turns out that for pure-jump Lévy processes one obtains weights for the approximating model which explode when ε tends to zero. Hence, the random variable inside the expectation diverges. However, due to an independence property in the limit which is not found in the classical setting of the density method, the delta converges anyhow. However, the variance of the expression explodes, which in turn implies that the weights are highly inefficient from a Monte Carlo point of view. The same problem does not occur for Lévy processes having a continuous martingale part. Hence, we conclude that even though the delta is robust towards these approximations, the resulting expressions for the deltas may become inefficient for practical simulation, at least in the pure-jump case. We study numerical examples discussing this problem. Also, we provide convergence rates for the approximative deltas.

Our presentation is organized as follows. After a short introduction on Lévy processes, we discuss the computational methods for the delta based on the existence of a probability density function: we revise the density method and we introduce the conditional density method. Then we discuss the problems related to model robustness and we present our results in connection to the analysis of sensitivity. Several examples are provided, including different classes of Lévy process and relevant functions f . A numerical study investigates our findings in a practical setting based on Monte Carlo simulations. Comments on our results and future research are given as conclusion.

2 Density based methods for the computation of derivatives

In this Section we introduce and analyze the so-called density method for the calculation of derivatives with respect to parameters in expectation functionals. The density method is a classical approach based on the representation of the derivative of the expectation functionals as an expectation involving the logarithmic derivative of the density function. We recall basic results in this area, and propose a *conditional density* approach which may be useful in certain contexts.

2.1 Some mathematical preliminaries

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space equipped with a filtration $\{\mathcal{F}_t\}_{t \in [0, T]}$ satisfying the *usual conditions*. We introduce the generic notation $L(t)$ for a Lévy process on the probability space, and denote by $B(t)$ a Brownian motion, with $t \in [0, T]$

and $L(0) = B(0) = 0$ by convention. In the sequel of the paper, we fix $T = 1$ for simplicity in notation. We work with the RCLL¹ version of the Lévy process, and let $\Delta L(t) = L(t) - L(t-)$. Denote the Lévy measure of $L(t)$ by $Q(dz)$. This is a σ -finite Borel measure on $\mathbb{R}_0 \triangleq \mathbb{R} - \{0\}$.

We recall the Lévy-Itô decomposition for a Lévy process (See e.g. Sato [14]).

Theorem 2.1 Lévy-Itô decomposition.

Let $L(t)$ be a Lévy process and Q its Lévy measure. Then we have:

- Q verifies

$$\int_{\mathbb{R}_0} \min(1, z^2) Q(dz) < \infty.$$

- The jump measure of $L(t)$, denoted by $N(dt, dz)$, is a Poisson random measure on $[0, \infty[\times \mathbb{R}$ with intensity measure $Q(dz) dt$.
- There exists a Brownian motion $W(t)$ such that

$$L(t) = at + bW(t) + Z(t) + \lim_{\varepsilon \downarrow 0} \tilde{Z}_\varepsilon(t), \quad (2.1)$$

where

$$Z(t) \triangleq \sum_{s \in [0, t]} \Delta L(s) \mathbf{1}_{\{|\Delta L(s)| \geq 1\}} = \int_{|z| \geq 1} z N(ds, dz)$$

and

$$\tilde{Z}_\varepsilon(t) \triangleq \sum_{s \in [0, t]} \Delta L(s) \mathbf{1}_{\{\varepsilon \leq |\Delta L(s)| < 1\}} - t \int_{\varepsilon \leq |z| < 1} z Q(dz) = \int_{\varepsilon \leq |z| < 1} z \tilde{N}(ds, dz),$$

where $\tilde{N}(dt, dz) := N(dt, dz) - Q(dz)dt$ is the compensated Poisson random measure of $L(t)$ and $a, b \in \mathbb{R}$ are two constants. The limit in (2.1) is meant with convergence almost sure and uniform in $t \in [0, 1]$, The components W, Z and \tilde{Z}_ε are independent.

We introduce the following notation for the variation of the Lévy process $L(t)$ close to the origin. For $0 < \varepsilon \leq 1$,

$$\sigma^2(\varepsilon) \triangleq \int_{|z| < \varepsilon} z^2 Q(dz). \quad (2.2)$$

Since any Lévy measure $Q(dz)$ integrates z^2 in an open interval around zero, we have that $\sigma^2(\varepsilon)$ is finite for any $\varepsilon > 0$. The $\sigma^2(\varepsilon)$ represents the variance of the jumps smaller than ε of $L(t)$ in the case it is symmetric and has mean zero. We will frequently make use of $\sigma^2(\varepsilon)$ for our studies. But first, we recall a result of Orey [12] which relates the asymptotic behavior of the Lévy measure at zero (that is, the asymptotic behavior of $\sigma^2(\varepsilon)$ as ε tends to zero) to the smoothness of the probability density of $L(t)$.

Theorem 2.2 Let $L(t)$ be a Lévy process, then it follows:

¹Right-continuous with left limits, also called *càdlàg*.

- If $b > 0$ or $Q(\mathbb{R}_0) = \infty$, then $L(t)$ has a continuous probability density $p_t(\cdot)$ on \mathbb{R} .
- If there exists $\gamma \in]0, 2[$ such that $Q(dz)$ satisfies

$$\liminf_{\varepsilon \rightarrow 0} \frac{\sigma^2(\varepsilon)}{\varepsilon^\gamma} > 0, \quad (2.3)$$

then the probability density p_t of $L(t)$ is infinitely continuously differentiable and for all $n \geq 1$,

$$\lim_{|x| \rightarrow \infty} \frac{\partial^n p_t}{\partial x^n}(x) = 0.$$

We observe that both the α -stable and the normal inverse Gaussian (NIG) Lévy processes satisfy condition (2.3) ensuring the existence of a smooth density. Indeed, the Lévy measure of an α -stable process with $\alpha \in]0, 2[$ is (see for instance Sato [14])

$$Q(dz) = c_1 |z|^{-1-\alpha} \mathbf{1}_{\{z < 0\}} dz + c_2 z^{-1-\alpha} \mathbf{1}_{\{z > 0\}} dz,$$

with $c_1, c_2 \geq 0$ and $c_1 + c_2 > 0$. Therefore,

$$\sigma^2(\varepsilon) = \frac{c_1 + c_2}{2 - \alpha} \varepsilon^{2-\alpha}.$$

Hence, choose $\gamma = 2 - \alpha$ to verify condition (2.3). The NIG Lévy process has Lévy measure (see Barndorff-Nielsen [2]),

$$Q(dz) = \frac{\alpha \delta}{\pi |z|} K_1(\alpha |z|) e^{\beta z} dz,$$

where α, β, δ are parameters satisfying $0 \leq \beta \leq \alpha$ and $\delta > 0$, and $K_1(z)$ is the modified Bessel function of the third kind with index 1. Using properties of the Bessel functions (see Assmussen and Rosinski [1], one finds

$$\sigma^2(\varepsilon) = \frac{2\delta}{\pi} \varepsilon.$$

Hence, letting $\gamma = 1$ we readily verify condition (2.3) also for the NIG Lévy process.

Thm. 2.2 is useful in our analysis since it ensures that the density function of a Lévy process is differentiable, which is the basic requirement for the applicability of the so-called density method which we study next.

2.2 The density method

In this paper we are concerned with the derivative of the expectation of functionals of the form

$$F(x) \triangleq \mathbb{E}[f(x + Y)], \quad (2.4)$$

for a random variable Y and a measurable function f such that $f(x + Y) \in L^1(\mathbb{P})$ for each $x \in \mathbb{R}$ (or in some subset of \mathbb{R}). Here, we denote by $L^1(\mathbb{P})$ the space of all random

variables which are integrable with respect to \mathbb{P} . In most of our forthcoming analysis, Y will be a Lévy process $L(t)$ or some approximation of such including Brownian motion. We call a random variable Π a *weight* if $f(x + Y)\Pi \in L^1(\mathbb{P})$ for $x \in \mathbb{R}$ and

$$F'(x) \triangleq \frac{dF(x)}{dx} = \mathbb{E}[f(x + Y)\Pi]. \quad (2.5)$$

A straightforward derivation inside the expectation operator would lead to $F'(x) = \mathbb{E}[f'(x + Y)]$, so a sensitivity weight can be viewed as the result after a kind of “integration-by-parts” operation. The advantage with an expression of the form (2.5) is that we can consider the derivative of expectation functionals where the function f is not differentiable. Examples where this is relevant include the calculation of delta-hedge ratios in option pricing for “payoff-functions” f being non-differentiable (digital options, say). Other examples are the sensitivity of risk measures with respect to a parameter, where the risk measure may be a non-differentiable function of the risk (Value-at-Risk, say, which is a quantile measure).

There exist by now at least two methods to derive sensitivity weights for functionals like $F(x)$. The classical approach is the density method, which transfers the dependency of x to the density function of Y , and then differentiate. An alternative method is the Malliavin approach, applying the tools from Malliavin calculus to perform an integration-by-parts utilizing the Malliavin derivative rather than classical differentiation. We refer to Fournié *et al.* [11] for more information on this approach.

Let us discuss the density method (see Broadie and Glasserman [3] for applications to finance). Suppose Y has a density p_Y with respect to the Lebesgue measure dt . Then, from classical probability theory, we have that

$$F(x) = \int_{\mathbb{R}} f(x + y)p_Y(y) dy = \int_{\mathbb{R}} f(y)p_Y(y - x) dy. \quad (2.6)$$

Hence, the expectation functional $F(x)$ can be expressed as a convolution between f and p_Y . Recalling Thm. 8.10 in Folland [10], as long as $f \in L^1(\mathbb{R})$ and $p_Y \in C_0^n(\mathbb{R})$, F is n times continuously differentiable and its derivatives can be expressed as

$$F^{(k)}(x) = \int_{\mathbb{R}} f(y)(-1)^k \frac{d^k}{dy^k} p_Y(y - x) dy,$$

for $k \leq n$ and $F^{(k)}$ denoting the k 'th derivative of F . Here we have denoted the space of Lebesgue integrable functions on \mathbb{R} by $L^1(\mathbb{R})$ and the space of differentiable (up to order n) functions on \mathbb{R} vanishing at infinity by $C_0^n(\mathbb{R})$.

Restricting our attention to $n = 1$, and assuming that $p_Y(y) > 0$ for $y \in \mathbb{R}$, we find that

$$F'(x) = \int_{\mathbb{R}} f(x + y) \left(-\frac{d}{dy} \ln p_Y(y)\right) p_Y(y) dy = \mathbb{E}[f(x + Y)(-\partial \ln p_Y(Y))].$$

Thus, the density method yields a weight $\Pi = -\partial \ln p_Y(Y)$, the logarithmic derivative of the density. As we see from the above, under very mild assumptions on the density

of Y and the function f , we can find a weight Π for calculating the derivative of F without having to differentiate f .

Assuming that $f \in L^1(\mathbb{R})$ is rather strict in many applications. We can relax the conditions on f considerably as follows. Suppose that p_Y is differentiable and strictly positive, and $f(\cdot)p'_Y(\cdot - x)$ is bounded uniformly in x by an integrable function on \mathbb{R} . Then, according to Thm. 2.27 in Folland [10], we have

$$\begin{aligned} F'(x) &= \frac{d}{dx} \int_{\mathbb{R}} f(y)p_Y(y-x) dy \\ &= \int_{\mathbb{R}} f(y)(-1)p'_Y(y-x) dy \\ &= \int_{\mathbb{R}} f(x+y)\left(-\frac{d}{dy} \ln p_Y(y)\right)p_Y(y) dy \\ &= \mathbb{E}[f(x+Y)(-\partial \ln p_Y(Y))] . \end{aligned}$$

We obtain the same weight $\Pi = -\partial \ln p_Y(Y)$ as above, naturally. However, we can include functions f which can grow at infinity as long as the density (and its derivative) dampens this growth sufficiently. This ensures that we can apply the density method in financial contexts like calculating the delta of a call option.

2.3 The conditional density method

Another case which is important in our analysis will be the situation where we have two independent strictly positive random variables Y and Z with densities p_Y and p_Z , respectively. In some situations that we will encounter in the sequel, only one of the two densities may be known, or one of the two may be simpler to be used for computational purposes. We propose a *conditional density method* for such cases.

Obviously, if the density of $Y + Z$ is known, we are in the situation described in the previous subsection. Under the hypotheses stated there, we may apply the standard density method in order to find the derivative of the functional

$$F(x) = \mathbb{E}[f(x + Y + Z)] .$$

In this case we find

$$F'(x) = F'_{Y+Z}(x) = \mathbb{E}[f(x + Y + Z)(-\partial \ln p_{Y+Z}(Y + Z))] .$$

We use the notation $F'_{Y+Z}(x)$ to emphasize that we apply the density method to the sum $Y + Z$.

On the other hand, if only one of the two densities p_Y or p_Z is known or better fitting computations, we can apply the *conditional density method* as follows. Since by conditioning we have

$$F(x) = \mathbb{E}[\mathbb{E}[f(x + Y + Z) | Y]] = \mathbb{E}[\mathbb{E}[f(x + Y + Z) | Z]] ,$$

we find (see Sato, Prop. 1.16)

$$F(x) = \int_{\mathbb{R}} \mathbb{E}[f(y + Z)] p_Y(y-x) dy = \int_{\mathbb{R}} \mathbb{E}[f(z + Y)] p_Z(z-x) dz .$$

This holds as long as $\mathbb{E}[f(\cdot + Z)]p_Y(\cdot - x)$ is integrable (or, symmetrically, $\mathbb{E}[f(\cdot + Y)]p_Z(\cdot - x)$ is integrable). Strictly speaking, the Proposition 1.16 in Sato is only valid under boundedness conditions, however, these can be relaxed by standard limiting arguments. The expressions of $F(x)$ can be used in two ways to derive the derivative $F'(x)$: First, we find

$$F'_Y(x) = \mathbb{E}[f(x + Y + Z)(-\partial \ln p_Y(Y))] ,$$

as long as p_Y is differentiable and $\mathbb{E}[f(\cdot - Z)]p'_Y(\cdot - x)$ is bounded by an integrable function uniformly in x , say. Symmetrically, we obtain

$$F'_Z(x) = \mathbb{E}[f(x + Y + Z)(-\partial \ln p_Z(Z))] ,$$

whenever $\mathbb{E}[f(\cdot - Y)]p'_Z(\cdot - x)$ is bounded by an integrable function uniformly in x . Obviously, $F'_Y(x) = F'_Z(x) = F'_{Y+Z}(x) = F'(x)$, however, the three different calculations lead to three different weights, being, respectively,

$$\begin{aligned} \Pi_{Y+Z} &\triangleq -\partial \ln p_{Y+Z}(Y + Z) \\ \Pi_Y &\triangleq -\partial \ln p_Y(Y) \\ \Pi_Z &\triangleq -\partial \ln p_Z(Z) . \end{aligned}$$

The two last are resulting from the conditional density method, while the first one is from the density method. These three weights are genuinely different.

3 Robustness of the delta to model choice

In this Section we will analyze the sensitivity of expectation functionals with respect to Lévy processes and their approximations. Our main focus will be on cases where a Lévy process $L(t)$ and its approximation $L_\varepsilon(t)$ are indistinguishable in practical contexts for small ε . Hence, in a concrete application, we may think of two models $L(t)$ and $L_\varepsilon(t)$ for the same random phenomenon which we cannot in practical terms separate. For instance, we may think of two speculators in a financial market who want to price an option. The first investor believes in a model given by $L(t)$, while the other chooses a model $L_\varepsilon(t)$, being slightly different than the former. The distributions of the two models will be very close, and thus also the derived option price. However, the main question we want to analyze in this paper is whether the same holds true for the sensitivities (or the Greeks in financial terminology). We refer to this question as a problem of robustness of sensitivities to model choice.

We analyze a particular class of approximations of $L(t)$, namely the one introduced by Asmussen and Rosinski [1] where the small jumps of $L(t)$ are substituted by an appropriately scaled Brownian motion. Before analyzing the sensitivity parameter delta for such approximations, we include for the convenience of the reader some details on small jump approximations and convergence.

3.1 Small jump approximations of Lévy processes

In applications of Lévy processes, it is often useful to approximate the small jumps by a Brownian motion. This approximation was advocated in Rydberg [13] as a way to

simulate the path of a Lévy process with NIG distributed increments, and later studied in detail by Asmussen and Rosinski [1].

Recall the Lévy-Itô decomposition of a Lévy process $L(t)$ as given in (2.1) and introduce now an approximating Lévy process

$$L_\varepsilon(t) \triangleq at + bW(t) + \sigma(\varepsilon)B(t) + Z(t) + \tilde{Z}_\varepsilon(t), \quad (3.1)$$

with $\sigma^2(\varepsilon)$ defined in (2.2) and $B(t)$ being a Brownian motion independent of $L(t)$ (which in fact also means independent of $W(t)$). From the definition of \tilde{Z}_ε , we see that we have substituted the small jumps (compensated by their expectation) in $L(t)$ by a Brownian motion scaled with $\sigma(\varepsilon)$. Hence, we approximate the small jumps by a Brownian motion with the same variance as the compensated small jumps. We have the following result:

Proposition 3.1 *Let the processes $L(t)$ and $L_\varepsilon(t)$ be defined as in equation (2.1) and (3.1), respectively. Then, for every t , we have:*

$$\lim_{\varepsilon \rightarrow 0} L_\varepsilon(t) = L(t) \quad \text{the convergence is in } L^1 \text{ and } \mathbb{P} - \text{a.s.}$$

Proof. Whenever $\varepsilon \rightarrow 0$, it follows that

$$\begin{aligned} \mathbb{E} [|L_\varepsilon(t) - L(t)|] &= \mathbb{E} \left[\left| \sigma(\varepsilon)B(t) - \int_0^t \int_{0 \leq |z| \leq \varepsilon} z \tilde{N}(ds, dz) \right| \right] \\ &\leq \sigma(\varepsilon) \mathbb{E} [|B(t)|] + \mathbb{E} \left[\left| \int_0^t \int_{0 \leq |z| \leq \varepsilon} z \tilde{N}(ds, dz) \right| \right] \\ &\leq \sigma(\varepsilon) \mathbb{E} [B^2(t)]^{1/2} + \mathbb{E} \left[\left(\int_0^t \int_{0 \leq |z| \leq \varepsilon} z \tilde{N}(ds, dz) \right)^2 \right]^{1/2} \\ &\leq 2\sigma(\varepsilon)\sqrt{t} \rightarrow 0, \end{aligned}$$

from the triangle and Cauchy-Schwarz inequalities. This proves the convergence in L^1 . As for the \mathbb{P} -a.s. convergence, this follows directly from the proof of the Lévy-Kintchine formula (See Thm. 19.2 in Sato [14]). \square

The study in Asmussen and Rosinski [1] gives a central limit type of result for the approximation of the small jumps. It says that the small jumps are, after scaling by $\sigma(\varepsilon)$, indeed close to be standard normally distributed. We note that the above result only says that, for every t , the two random variables $L(t)$ and $L_\varepsilon(t)$ are close in distribution, but nothing about the asymptotic distribution of the small jumps in the limit. Indeed, under an asymptotic condition on $\sigma(\varepsilon)$, the result in [1] is:

Theorem 3.2 *If*

$$\lim_{\varepsilon \rightarrow 0} \frac{\sigma(\varepsilon)}{\varepsilon} = \infty, \quad (3.2)$$

then

$$\lim_{\varepsilon \rightarrow 0} \sigma^{-1}(\varepsilon) \tilde{Z}_\varepsilon = B,$$

where B is a Brownian motion and the convergence is in distribution.

This results supports the choice of using a Brownian motion and the scale $\sigma(\varepsilon)$ for the small jumps of a Lévy process.

3.2 Robustness to model choice

Assume that $f : \mathbb{R} \mapsto \mathbb{R}$ is a measurable function and that for each x belonging to a compact set of \mathbb{R} , there exists a random variable $U \in L^1(\mathbb{P})$ such that $|f(x + L_\varepsilon(1))| \leq U$ for all ε . Without loss of generality, we can consider $x \in [x_1, x_2]$, for some $x_1, x_2 \in \mathbb{R}$. Since $f(x + L_\varepsilon(1))$ converges almost surely to $f(x + L(1))$, by dominated convergence it holds that

$$\lim_{\varepsilon \downarrow 0} \mathbb{E}[f(x + L_\varepsilon(1))] = \mathbb{E}[f(x + L(1))] = F(x). \quad (3.3)$$

Such expectation functionals arise in pricing of options, where f is the payoff function from the option and $x + L(1)$ is the state of the underlying asset at exercise time 1. If $f(x) = \mathbf{1}_{\{x < q\}}$, we may view the expectation as coming from a simple quantile risk measure on the random variable $x + L(1)$, where the x is the initial state of the system under consideration. For notational simplification, we introduce

$$F_\varepsilon(x) \triangleq \mathbb{E}[f(x + L_\varepsilon(1))] \quad (3.4)$$

and set $F_0(x) = F(x)$. We analyze $F'_\varepsilon(x)$ and its convergence to $F'(x)$.

To differentiate $F_\varepsilon(x)$, we have in fact a multiple of different approaches. Motivated from the Malliavin method of Davis and Johansson [7] for jump diffusions, it is natural to use the conditional density method with respect to the Brownian motion. However, this leads to three possibilities. Either we can differentiate with respect to the original Brownian motion $W(t)$, or with respect to $B(t)$, or finally with respect to a *new* Brownian motion defined as the sum of the two. This will lead to three new expectation operators, which we would like to converge to $F'(x)$ when $\varepsilon \downarrow 0$. As we will see, this is indeed the case, however, the three choices have different properties. Obviously, if the distribution of $L_\varepsilon(t)$ is available, one would prefer to use the density method directly on this. However, due to the truncation of the jumps at ε , this will not, in most practical applications, be available.

To get some intuition on the problem we are facing, let us consider a trivial case of a Brownian motion

$$L(t) = bW(t).$$

Of course, in this case we have by a straightforward application of the density method

$$F'(x) = \mathbb{E} \left[f(x + bW(1)) \frac{W(1)}{b} \right].$$

Now, introduce

$$L_\varepsilon(t) = \sqrt{b^2 - \varepsilon^2} W(t) + \varepsilon B(t)$$

where $B(t)$ is an independent Brownian motion. Note that $L_\varepsilon(t)$ in distribution is identical to $L(t)$, so it is in this sense not an approximation of $L(t)$ as above for the Lévy case. However, we are mimicking the approximation procedure above by representing

the “small jumps” of $W(t)$ by a new Brownian motion $B(t)$. If we first apply the conditional density method on $W(t)$, we find

$$\begin{aligned} F'_\varepsilon(x) &= \frac{d}{dx} \int_{\mathbb{R}} \mathbb{E}[f(u + \varepsilon B(1))] p_{\sqrt{b^2 - \varepsilon^2}W(1)}(u - x) dx \\ &= \mathbb{E} \left[f(x + L_\varepsilon(1)) \frac{W(1)}{\sqrt{b^2 - \varepsilon^2}} \right], \end{aligned}$$

where we recall that p_X denotes the probability density of the random variable X . We see easily that $F'_\varepsilon(x)$ converges nicely to $F'(x)$. Next, differentiating using the distribution of $\sqrt{b^2 - \varepsilon^2}W(1) + \varepsilon B(1)$ leads similarly to

$$F'_\varepsilon(x) = \mathbb{E} \left[f(x + L_\varepsilon(1)) \frac{\sqrt{b^2 - \varepsilon^2}W(1) + \varepsilon B(1)}{b^2} \right].$$

This will again converge nicely to $F'(x)$. Finally, apply the procedure with respect to $B(t)$ to find

$$F'_\varepsilon(x) = \mathbb{E} \left[f(x + L_\varepsilon(1)) \frac{B(1)}{\varepsilon} \right].$$

This results in a sensitivity weight $B(1)/\varepsilon$ which explodes when $\varepsilon \downarrow 0$, and it is not immediately clear by direct inspection of the functional if it is nicely behaving when taking the limit. However, since all the three approaches above lead to the same derivative $F'_\varepsilon(x)$, we are ensured that the limit also in this case is equal to $F'(x)$. However, from a practical perspective the weight will have a very high variance compared to the two first approaches, and therefore it is not useful in numerical simulations. This illustrates that the approach of Davis and Johansson [7] is not necessarily leading to sensitivity weights which are “good”. In particular we should notice that in the case of pure-jump Lévy processes we will not have any $W(t)$ -term, and we are somehow “forced” to use the density method with respect to $B(t)$, the approximating Brownian motion. This may lead to problems when understanding the limit since we will not have any comparison.

Let us go back to the general case, where we first suppose that $b > 0$ in (2.1) and apply the density method on the combination of $W(t)$ and $B(t)$. To distinguish between the different sensitivity weights, we introduce some notation. Let the derivative of $F_\varepsilon(x)$ with respect to x resulting from applying the density method on $W(t)$, which is the Brownian motion in the Lévy-Kintchine representation of $L(t)$, be denoted by $F'_{\varepsilon,W}(x)$. Further, we use the notation $F'_{\varepsilon,B}(x)$ and $F'_{\varepsilon,B,W}(x)$ for the derivative when we use the density method with respect to the small-jump approximating process B or $bW(t) + \sigma(\varepsilon)B(t)$, respectively. Note that even though we may have the density of $L(t)$, it may be very hard to find the density of $L_\varepsilon(t)$, and thus to apply the density method on the approximating process directly.

We denote by C_b^k the space of k -times continuously differentiable functions with all derivatives bounded, C_b^0 will be denoted by C_b , the space of bounded and continuous functions. It is simple to derive the following result:

Proposition 3.3 Suppose $f \in C_b$. For every $\varepsilon > 0$, we have that

$$\begin{aligned} F'_{\varepsilon,B}(x) &= \mathbb{E} \left[f(x + L_\varepsilon(1)) \frac{B(1)}{\sigma(\varepsilon)} \right] \\ F'_{\varepsilon,B,W}(x) &= \mathbb{E} \left[f(x + L_\varepsilon(1)) \frac{bW(1) + \sigma(\varepsilon)B(1)}{b^2 + \sigma^2(\varepsilon)} \right]. \end{aligned}$$

If $b > 0$, we have in addition that

$$F'_{\varepsilon,W}(x) = \mathbb{E} \left[f(x + L_\varepsilon(1)) \frac{W(1)}{b} \right].$$

Proof. Using the conditional density method applied to $B(t)$, we get

$$F'_\varepsilon(x) = \frac{\partial}{\partial x} \int_{\mathbb{R}} \mathbb{E}[f(u + a + bW(1) + Z(1) + \tilde{Z}_\varepsilon(1))] p_{\sigma(\varepsilon)B(1)}(u - x) dx.$$

Here we can dominate the density $P_{\sigma(\varepsilon)B(1)}(u - x)$ uniformly in x by an integrable function which is a sufficient condition to take the derivative inside the integral if f is bounded. Applying the conditional density method to $bW(t) + \sigma(\varepsilon)B(t)$ and $W(t)$, respectively, and using the same arguments above to take the derivative inside the integral, we get the result. \square

Note that $F'_{\varepsilon,B}(x) = F'_{\varepsilon,B,W}(x) = F'_{\varepsilon,W}(x)$ for all $\varepsilon > 0$. Moreover, we have the following robustness result when $b > 0$, that is, when the Lévy process $L(t)$ has a continuous martingale term.

Proposition 3.4 Suppose that the diffusion coefficient $b > 0$ and that $f \in C_b$. Then we have

$$\lim_{\varepsilon \downarrow 0} F'_{\varepsilon,W}(x) = \lim_{\varepsilon \downarrow 0} F'_{\varepsilon,B,W}(x) = \lim_{\varepsilon \downarrow 0} F'_{\varepsilon,B}(x) = \mathbb{E} \left[f(x + L(1)) \frac{W(1)}{b} \right] = F'(x).$$

Proof. This hinges on the fact that,

$$F'(x) = \mathbb{E} \left[f(x + L(1)) \frac{W(1)}{b} \right].$$

Now, by the assumption on $f(x + L_\varepsilon(1))W(1)$ and the dominated convergence theorem, we find that

$$\lim_{\varepsilon \downarrow 0} F'_{\varepsilon,W}(x) = F'(x).$$

Furthermore, since $F'_{\varepsilon,B}(x) = F'_{\varepsilon,B,W}(x) = F'_{\varepsilon,W}(x)$, we have that the limit of $F'_{\varepsilon,B,W}(x)$ and $F'_{\varepsilon,B}(x)$ also exist and are equal to $F'(x)$. This proves the result. \square

Remark that although we cannot bound $B(1)/\sigma(\varepsilon)$ by some integrable random variable, we still obtain the convergence. This depends on the fact that the derivative $F'_{\varepsilon,B}(x)$ is equal to $F'_{\varepsilon,W}(x)$ when $b > 0$. When $b = 0$, we can not use this argument anymore, however, we have the following simple result when f is smooth.

Proposition 3.5 Suppose $f \in C_b^1$ and that there exists a random variable $U \in L^1(\mathbb{P})$ such that $|f'(x + L_\varepsilon(1))| \leq U$ uniformly in x and ε . Then

$$\lim_{\varepsilon \rightarrow 0} F'_{\varepsilon, B}(x) = F'(x) = \mathbb{E} [f'(x + L(1))] .$$

Proof. First, observe that $|f'(x + L(1))| \leq U$ uniformly in x by choosing $\varepsilon = 0$ in the assumption. Hence, by Thm. 2.27 in Folland [10], $F(x)$ is differentiable, and we can move the differentiation inside the expectation operator to obtain

$$F'(x) = \mathbb{E} [f'(x + L(1))] .$$

This proves the second equality. Next, by the same argument, we have that

$$F'_\varepsilon(x) = \frac{d}{dx} \mathbb{E} [f(x + L_\varepsilon(1))] = \mathbb{E} [f'(x + L_\varepsilon(1))] .$$

From the conditional density method, we know that $F'_{\varepsilon, B}(x) = F'_\varepsilon(x)$. By dominated convergence, it holds that

$$\lim_{\varepsilon \rightarrow 0} F'_\varepsilon(x) = F'(x)$$

and the proof is complete. \square

Note that the result holds for all $b \geq 0$, and we could have used it to prove the limits for $b > 0$ as well in the smooth case of f .

In many applications, like for instance in finance, the assumption that f should be continuous and bounded is too restrictive. For example, a call option will lead to an unbounded function, whereas a digital option gives a discontinuous f . Hence, it is natural to look for extensions of the above results to classes of functions where the conditions on f are weakened. One natural approach is to look at classes of functions f which can be approximated by functions in C_b . Another path, which we shall take here, is to apply Fourier methods.

Let now $f \in L^1(\mathbb{R})$. The Fourier transform of f is defined by

$$\widehat{f}(u) = \int_{\mathbb{R}} f(y) e^{iuy} dy . \quad (3.5)$$

Suppose in addition that $\widehat{f} \in L^1(\mathbb{R})$. Then the inverse Fourier transform is well-defined, and we have

$$f(y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iuy} \widehat{f}(u) du . \quad (3.6)$$

With these two definitions at hand, we can do the following calculation taken from Carr and Madan [5]. Assume for every x that $f(x + \cdot)$ is integrable with respect to the distribution $p_{L(1)}(dy)$ of $L(1)$. Then

$$\mathbb{E}[f(x + L(1))] = \int_{\mathbb{R}} f(x + y) p_{L(1)}(dy) .$$

Invoking the inverse Fourier transformed representation of f in (3.6), and applying Fubini-Tonelli to commute the integration, we find

$$\begin{aligned}\mathbb{E}[f(x + L(1))] &= \int_{\mathbb{R}} \left\{ \frac{1}{2\pi} \int_{\mathbb{R}} e^{-i(x+y)u} \widehat{f}(u) du \right\} p_{L(1)}(dy) \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} \left\{ \int_{\mathbb{R}} e^{-iuy} p_{L(1)}(dy) \right\} \widehat{f}(u) du.\end{aligned}$$

Thus, it follows that

$$\mathbb{E}[f(x + L(1))] = \frac{1}{2\pi} \int_{\mathbb{R}} e^{-iux} \varphi_{L(1)}(u) \widehat{f}(u) du, \quad (3.7)$$

where $\varphi_{L(1)}$ is the characteristic function of $L(1)$ defined from the Lévy-Kintchine formula as

$$\phi_{L(1)}(u) = \exp \left(iau - \frac{1}{2} b^2 u^2 + \int_{\mathbb{R}_0} e^{iuz} - 1 - iuz 1_{|z| < 1} Q(dz) \right). \quad (3.8)$$

We have the following Lemma for the delta.

Lemma 3.6 *Under the condition $u\widehat{f}(u) \in L^1(\mathbb{R})$ we have*

$$F'(x) = \frac{\partial}{\partial x} \mathbb{E}[f(x + L(1))] = \frac{1}{2\pi} \int_{\mathbb{R}} -iue^{-iux} \varphi_{L(1)}(u) \widehat{f}(u) du.$$

Proof. We differentiate the integrand in (3.7) and dominate it uniformly in x :

$$\left| \frac{\partial}{\partial x} e^{-iux} \varphi_{L(1)}(u) \widehat{f}(u) \right| = \left| -iu \frac{e^{-iux}}{2\pi} \phi_{L(1)}(u) \widehat{f}(u) \right| \leq |u\widehat{f}(u)|.$$

The result follows by appealing to Prop. 2.27 in Folland [10]. \square

Note that the condition $u\widehat{f}(u) \in L^1(\mathbb{R})$ is related to the derivative of f , since as long as f is differentiable we have $\widehat{f}'(u) = u\widehat{f}(u)$ whenever $f' \in L^1(\mathbb{R})$.

We finally reach the desired stability result for non-smooth f 's.

Proposition 3.7 *Suppose that*

$$u\widehat{f}(u) \in L^1(\mathbb{R}).$$

Then we have

$$\lim_{\varepsilon \rightarrow 0} \frac{\partial}{\partial x} \mathbb{E}[f(x + L_\varepsilon(1))] = \frac{\partial}{\partial x} \mathbb{E}[f(x + L(1))].$$

Proof. From Lemma 3.6 applied to $L_\varepsilon(1)$ we have

$$\frac{\partial}{\partial x} \mathbb{E}[f(x + L_\varepsilon(1))] = \frac{1}{2\pi} \int_{\mathbb{R}} -iue^{-iux} \varphi_{L_\varepsilon(1)}(u) \widehat{f}(u) du.$$

But,

$$\left| -iue^{-iux} \varphi_{L_\varepsilon(1)}(u) \widehat{f}(u) \right| \leq |u\widehat{f}(u)|$$

which, from the assumption, permits us to take the limit inside the integral and the result follows by Prop. 2.24 in Folland [10]. \square

Observe that in the Proposition above we handle $b \geq 0$, and there is no need to differentiate between the cases $b = 0$ and $b > 0$. There is no requirement of continuity of f in the above arguments. However, the integrability restriction excludes unbounded functions f , like for instance those coming from option pricing. However, we can easily deal with such by introducing a damped function f in the following manner. Define for $\alpha > 0$ the function

$$g_\alpha(y) = e^{-\alpha y} f(y). \quad (3.9)$$

Assuming that $g_\alpha \in L^1(\mathbb{R})$ and $\widehat{g}_\alpha \in L^1(\mathbb{R})$ for some $\alpha > 0$, we can apply the above results for g_α . To translate to f , observe that

$$f(y) = \frac{1}{2\pi} \int_{\mathbb{R}} e^{(\alpha - iu)y} \widehat{g}_\alpha(u) du,$$

and

$$\widehat{g}_\alpha(u) = \widehat{f}(u + i\alpha).$$

Hence, Prop. 3.7 holds for any f such that there exists $\alpha > 0$ for which we have the following assumptions

$$(\alpha - iu)\widehat{f}(u + i\alpha) \in L^1(\mathbb{R}) \quad \text{and} \quad e^{\alpha y} p_{L(1)}(dy) \in L^1(\mathbb{R}).$$

As illustration we consider two examples. First, let f be the payoff from a call option written on an asset with price defined as $S(t) = S(0) \exp(L(t))$ ($S(0) > 0$). Then, with $x = \ln S(0)$, we have

$$f(y) = \max(e^y - K, 0)$$

for $K > 0$ being the strike price. For $\alpha > 1$, we have that $g_\alpha \in L^1(\mathbb{R})$. Moreover,

$$\widehat{g}_\alpha(u) = \frac{K e^{(iu - \alpha) \ln K}}{(iu - \alpha)(iu - \alpha + 1)},$$

which is in $L^1(\mathbb{R})$. By a direct calculation, we find that

$$(\alpha - iu)\widehat{f}(u + i\alpha) = \frac{K^{1+iu-\alpha}}{1 + iu - \alpha},$$

which belongs to $L^1(\mathbb{R})$. Hence, Prop. 3.7 ensures that the approximation $L_\varepsilon(1)$ gives a delta which converges to the delta resulting from the model with $L(1)$.

We consider now a digital option written on an asset with price defined as $S(t) = S(0) \exp(L(t))$ ($S(0) > 0$). Then, with $x = \ln S(0)$, we have

$$f(y) = 1_{\{e^y > B\}}, \quad B \in \mathbb{R}_+.$$

For $\alpha > 0$, we have that $g_\alpha \in L^1(\mathbb{R})$. Moreover,

$$\widehat{g}_\alpha(u) = \frac{-B^{iu-\alpha}}{iu - \alpha},$$

which is in $L^1(\mathbb{R})$. By a direct calculation, we find that

$$(\alpha - iu)\widehat{f}(u + i\alpha) = B^{iu-\alpha},$$

which belongs to $L^1(\mathbb{R})$.

3.3 Robustness to smoothing of a Lévy process

In the above analysis we have focused entirely on the approximation of the small jumps in a Lévy process. However, we can apply our analysis also in a slightly different situation. Suppose that we are dealing with is a Lévy process for which the density of $L(t)$ (and its log-derivative) may be hard to compute, or may not even be existent analytically. In this case one may approximate the derivative of $F(x)$ by considering the following Lévy process:

$$\widehat{L}_\varepsilon(t) \triangleq L(t) + \widehat{\sigma}(\varepsilon)B(t), \quad (3.10)$$

where $B(t)$ is a Brownian motion independent of $L(t)$ and

$$\lim_{\varepsilon \downarrow 0} \widehat{\sigma}(\varepsilon) = 0.$$

We call $\widehat{L}_\varepsilon(t)$ a *smoothing* of $L(t)$, since we add an independent Brownian motion which has a smooth density, and thus $\widehat{L}_\varepsilon(t)$ will possess a smooth density as well.

Using the same proof as in Prop. 3.1, we have that $\widehat{L}_\varepsilon(t)$ converges in $L^1(\mathbb{P})$ to $L(t)$. Furthermore, since obviously $\widehat{\sigma}(\varepsilon)B(1)$ converges *a.s.* to zero, $\widehat{L}_\varepsilon(1)$ converges *a.s.* to $L(1)$. We have

$$\widehat{F}'_\varepsilon(x) = \mathbb{E} \left[f(x + \widehat{L}_\varepsilon(1)) \frac{B(1)}{\widehat{\sigma}(\varepsilon)} \right]. \quad (3.11)$$

Tracing through the arguments in the preceding subsection and assuming the right conditions on f , we find that

$$\lim_{\varepsilon \rightarrow 0} \widehat{F}'_\varepsilon(x) = F'(x). \quad (3.12)$$

This provides us with another stability result. The derivative of $F(x)$ is continuous with respect to perturbation of $L(t)$ by $\widehat{\sigma}(\varepsilon)B(t)$. As a curiosity, we can do the following: By independence of $L(1)$ and $B(1)$, we have

$$\begin{aligned} \mathbb{E} \left[f(x + \widehat{L}_\varepsilon(1)) \frac{B(1)}{\widehat{\sigma}(\varepsilon)} \right] &= \mathbb{E} \left[\left(f(x + \widehat{L}_\varepsilon(1)) - f(x + L(1)) \right) \frac{B(1)}{\widehat{\sigma}(\varepsilon)} \right] \\ &\quad + \mathbb{E} \left[f(x + L(1)) \frac{B(1)}{\widehat{\sigma}(\varepsilon)} \right] \\ &= \mathbb{E} \left[B^2(1) \frac{f(x + \widehat{\sigma}(\varepsilon)B(1) + L(1)) - f(x + L(1))}{\widehat{\sigma}(\varepsilon)B(1)} \right]. \end{aligned}$$

Notice, that the fraction on the right is close to a Malliavin derivative, since we in fact are looking at a derivative of f along B . Loosely speaking, when taking the limit we are looking at a derivative of $f(x + L(1))$ in the direction of $B(1)$, which resembles the idea of Malliavin differentiation. Furthermore, it is to be expected that this limit will be independent of $B(1)$, which has variance equal to 1. Informally, we have therefore given a link between the Malliavin derivative based on Brownian motion and the delta for Lévy processes. Hence, this motivates that the approach by Davis and Johansson [7] may be extended to more general Lévy processes than merely Brownian motion and

Poisson processes as is the case in their paper. The formalization of this procedure is left to a future study.

Let us consider an example where the smoothing of $L(t)$ may be an attractive procedure. The so-called CGMY distribution was suggested in Carr *et al.* [4] to model asset price returns. It does not have any explicit density function, but is defined through its cumulant function²

$$\psi_{\text{CGMY}}(\theta) = C\Gamma(-Y) \{ (M - i\theta)^Y + (G - i\theta)^Y - G^Y \}, \quad (3.13)$$

with $\Gamma(x)$ being the Gamma-function and constants C, G, M and Y . We suppose that C, G and M are positive, and $Y \in [0, 2)$. The CGMY distribution is infinitely divisible, and we can define a CGMY-Lévy process $L(t)$ with Lévy measure

$$Q(dz) = C|z|^{-1-Y} \exp(-(G1(z < 0) + M1(z > 0))|z|) dz. \quad (3.14)$$

Since we do not have explicitly the density function of $L(t)$, the density method can not be used for deriving sensitivity estimates $F'(x)$. Instead we can apply the density method on the Brownian motion after smoothing $L(t)$. We first verify the condition of Thm. 2.2 for the CGMY-Lévy process, showing that a smooth density indeed exists.

We have

$$\sigma^2(\varepsilon) = C \int_{|z| < \varepsilon} |z|^{1-Y} \exp(-(G1(z < 0) + M1(z > 0))|z|) dz,$$

and thus it is sufficient to verify the condition in Thm. 2.2 for $z > 0$. Note that $e^{-Mz} \geq M\varepsilon$ for $0 \leq z \leq \varepsilon$, and it follows that

$$\int_0^\varepsilon z^{1-Y} e^{-Mz} dz \geq \frac{1}{2-Y} \varepsilon^{2-Y} e^{-M\varepsilon}.$$

Let $\gamma = 2 - Y$ in Thm. 2.2, and we see that

$$\liminf_{\varepsilon \downarrow 0} \frac{\sigma^2(\varepsilon)}{\varepsilon^\gamma} \geq \frac{1}{2-Y} > 0,$$

as long as $Y > 0$. Hence, there exists a smooth density for the CGMY-distribution when $Y \in (0, 2)$ and, if we would have this available, we could calculate $F'(x)$ via its logarithmic derivative. By smoothing we can approximate $F'(x)$ by

$$\widehat{F}'_\varepsilon(x) = \mathbb{E} \left[f(x + \widehat{L}_\varepsilon(1)) \frac{B(1)}{\sigma(\varepsilon)} \right].$$

The sensitivity weight will have a large variance for small ε , but it provides us with an expression that can be calculated using Monte Carlo simulations based on sampling of the CGMY-distribution and an independent normal distribution.

As an application, we consider an example from insurance. Let the loss of an insurance company be described by $L(t)$, and x being the premium charged by the

²The cumulant function is here the logarithm of the characteristic function

company to accept this risk. The question for the insurance company is to find a level x such that the net loss $x + L(1)$ is acceptable. A simple measure could be that the insurance company can only bear losses which are above a certain threshold, K say. Given a premium x , they want to calculate the probability of falling below the threshold K , which can be expressed by $P(x + L(1) < K)$. We find

$$P(x + L(1) < K) = \mathbb{E} [1_{\{x+L(1)<K\}}] ,$$

which therefore is an expectation functional on the form we have analyzed in this paper with $f(z) = 1_{\{z<K\}}$. Consider the derivative of this probability with respect to x , which we call the *marginal premium rate*:

$$F'(x) = \frac{d}{dx} \mathbb{E} [1_{\{x+L(1)<K\}}] .$$

The marginal premium rate tells us how sensitive the loss probability is with respect to the premium. Of course, if we know the density of $L(1)$, $p_{L(1)}$, and this is differentiable, the marginal premium rate is straightforwardly calculated to be

$$F'(x) = -p_{L(1)}(K - x) .$$

Thus, changing the premium by dx leads to a change in the loss probability of $-p_{L(1)}(K - x) dx$. However, if now the density of $L(1)$ is not known as is the case for the CGMY-distribution, we can not perform this simple calculation. By smoothing $L(t)$, we find the approximation

$$\widehat{F}'_{\varepsilon}(x) = \mathbb{E} \left[1_{\{x+\widehat{L}_{\varepsilon}(1)<K\}} \frac{B(1)}{\widehat{\sigma}(\varepsilon)} \right] .$$

Computations using conditional expectation lead to

$$\begin{aligned} \widehat{F}'_{\varepsilon}(x) &= \mathbb{E} \left[\mathbb{E} \left[1_{\{x+\widehat{\sigma}(\varepsilon)B(1)+L(1)<K\}} \frac{B(1)}{\widehat{\sigma}(\varepsilon)} \mid L(1) \right] \right] \\ &= -\mathbb{E} [p_{\widehat{\sigma}(\varepsilon)B(1)}(K - L(1) - x)] , \end{aligned}$$

with $p_{\widehat{\sigma}(\varepsilon)B(1)}$ being the density function of $\widehat{\sigma}(\varepsilon)B(1)$. Thus, also the approximation can be expressed as a density evaluated in $K - x$, however, in this case we need to take the expectation over $L(1)$. Furthermore, the density is singular when going to the limit.

Using the theory of distribution functions, we give a direct argument for the convergence of $\widehat{F}'_{\varepsilon}(x)$ to $F'(x)$. In fact by integration-by-parts, we have

$$\widehat{F}'_{\varepsilon,B}(x) = - (p_{\widehat{\sigma}(\varepsilon)B(1)}(K - x - \cdot), p_{L(1)})_2 = - (p_{\widehat{\sigma}(\varepsilon)B(1)}, p_{L(1)}(K - x - \cdot))_2$$

where $(\cdot, \cdot)_2$ is the inner product in $L^2(\mathbb{R})$, the space of square-integrable functions on \mathbb{R} . Since $p_{\widehat{\sigma}(\varepsilon)B(1)} \rightarrow \delta_0$ when $\widehat{\sigma}(\varepsilon) \rightarrow 0$, we find

$$\lim_{\varepsilon \rightarrow 0} \widehat{F}'_{\varepsilon,B}(x) = - (\delta_0, p_{L(1)}(K - x - \cdot))_2 = p_{L(1)}(K - x) = F'(x) .$$

This procedure may be carried through rigorously by using Schwartz distribution theory.

3.4 Some numerical issues

In the last part of this Section we turn our attention to some numerical issues concerning the use of the conditional density method for the above approximations.

For numerical purposes, it is of interest to know the rate of convergence of $F'_\varepsilon(x)$ to $F'(x)$. We have the following convergence speed for $F'_{\varepsilon,B,W}(x)$:

Proposition 3.8 *Suppose $b > 0$, $L(t)$ having finite variance and f being a Lipschitz continuous function. Then there exists a constant C depending on x , b , the Lipschitz constant of f , and the variance of $L(1)$ such that*

$$|F'_{\varepsilon,B,W}(x) - F'(x)| \leq C\sigma(\varepsilon).$$

Proof. From the triangle and Cauchy-Schwarz inequalities, we have

$$\begin{aligned} & |F'_{\varepsilon,B,W}(x) - F'(x)| \\ & \leq \mathbb{E} \left[|f(x + L_\varepsilon(1)) - f(x + L(1))| \frac{bW(1) + \sigma(\varepsilon)B(1)}{b^2 + \sigma^2(\varepsilon)} \right] \\ & \quad + \mathbb{E} \left[|f(x + L(1))| \left| \frac{bW(1) + \sigma(\varepsilon)B(1)}{b^2 + \sigma^2(\varepsilon)} - \frac{W(1)}{b} \right| \right] \\ & \leq \mathbb{E} [|f(x + L_\varepsilon(1)) - f(x + L(1))|^2]^{1/2} \frac{\mathbb{E} [(bW(1) + \sigma(\varepsilon)B(1))^2]^{1/2}}{b^2 + \sigma^2(\varepsilon)} \\ & \quad + \mathbb{E} [f^2(x + L(1))]^{1/2} \mathbb{E} \left[\left| \frac{bW(1) + \sigma(\varepsilon)B(1)}{b^2 + \sigma^2(\varepsilon)} - \frac{W(1)}{b} \right|^2 \right]^{1/2}. \end{aligned}$$

Letting K being the Lipschitz constant (which we assume being equal to the growth constant of f for convenience), we get

$$\begin{aligned} |F'_{\varepsilon,B,W}(x) - F'(x)| & \leq \frac{K}{\sqrt{b^2 + \sigma^2(\varepsilon)}} \mathbb{E} [|L_\varepsilon(1) - L(1)|^2]^{1/2} + K \mathbb{E} [(1 + x + L(1))^2]^{1/2} \\ & \quad \times \mathbb{E} \left[\left| \left(\frac{b}{b^2 + \sigma^2(\varepsilon)} - \frac{1}{b} \right) W(1) + \frac{\sigma(\varepsilon)}{b^2 + \sigma^2(\varepsilon)} B(1) \right|^2 \right]^{1/2}. \end{aligned}$$

Since $W(1)$ and $B(1)$ are independent, we find the last expectation to be (after taking the square-root)

$$\sigma(\varepsilon)/b\sqrt{(b^2 + \sigma^2(\varepsilon))}.$$

Moreover,

$$L_\varepsilon(1) - L(1) = \sigma(\varepsilon)B(T) + \tilde{Z}_\varepsilon(1) - \lim_{\varepsilon \downarrow 0} \tilde{Z}_\varepsilon(1).$$

Note that the difference between $\tilde{Z}_\varepsilon(1)$ and $\lim_{\varepsilon \downarrow 0} \tilde{Z}_\varepsilon(1)$ is the jumps between 0 and ε . Due to independence of the jumps and the Brownian motion B , we get

$$\mathbb{E} [|L_\varepsilon(1) - L(1)|^2] = 2\sigma^2(\varepsilon).$$

Hence, the result follows. \square

We note that with minor modifications of the above proof we can show that

$$|F'_{\varepsilon,B}(x) - F'(x)| \leq C\sigma(\varepsilon),$$

where C is a positive constant (not necessarily equal to the constant in the Proposition above). To show this result, we can simply let $b = 0$ in the proof and modify accordingly. Finally, it holds true for $\widehat{F}'_{\varepsilon}(x)$ as well by similar arguments.

In practice, one uses Monte Carlo methods in order to calculate $F'_{\varepsilon}(x)$. We consider the case $F'_{\varepsilon,B}(x)$, and recall that the estimated value of this based on N Monte Carlo simulations is

$$F'_{\varepsilon,B}(x) \approx \sum_{n=1}^N f(x + l_{\varepsilon,n}) \frac{b_n}{\sigma(\varepsilon)}$$

where b_n and $l_{\varepsilon,n}$ are independent random draws of $B(1)$ and $L_{\varepsilon}(1)$, respectively. Note that in order to draw from $L_{\varepsilon}(1)$, we use the draw from $B(1)$. The Monte Carlo error (or rather the standard deviation of the error) is given by

$$\text{std}(f(x + L_{\varepsilon}(1))B(1)) / (\sqrt{N}\sigma(\varepsilon)).$$

Assume now for technical simplicity that f is bounded. Then, from dominated convergence and independence of L and B , we find

$$\begin{aligned} \lim_{\varepsilon \rightarrow 0} \text{Var} \left[f(x + L_{\varepsilon}(1))B(1) \right] &= \lim_{\varepsilon \rightarrow 0} \left\{ \mathbb{E} \left[f^2(x + L_{\varepsilon}(1))B^2(1) \right] - \mathbb{E} \left[f(x + L_{\varepsilon}(1))B(1) \right]^2 \right\} \\ &= \mathbb{E} \left[f^2(x + L(1)) \right]. \end{aligned}$$

Hence,

$$\lim_{\varepsilon \rightarrow 0} \text{Var} \left(f(x + L_{\varepsilon}(1)) \frac{B(1)}{\sigma(\varepsilon)} \right) = \infty.$$

From this we can conclude the following: If we decide to use the conditional density method on pure-jump Lévy processes after first doing an approximation, the expression to simulate will have a large variance for small ε . Indeed, when ε tends to zero the variance explodes. This means that for close approximations of $x + L(1)$ we will have an expression to simulate which has a very high variance, and therefore we need a very high number of samples to get a confident estimate of the delta. In conclusion, the method may become very inefficient and unstable, and variance-reducing techniques are called for in order to get reliable estimates.

4 Numerical examples

We consider some examples to illustrate the conditional density method and our findings on approximations.

Let us assume that $L(t)$ is an NIG-Lévy process, that is, a Lévy process with NIG-distributed increments. Supposing $L(1)$ being NIG distributed with parameters α , β , δ and μ , the density is (see Barndorff-Nielsen [2])

$$p_{\text{NIG}}(x; \alpha, \beta, \delta, \mu) = \frac{\alpha\delta}{\pi} e^{\delta\sqrt{\alpha^2 - \beta^2} + \beta(x - \mu)} \frac{K_1 \left(\alpha\sqrt{\delta^2 + (x - \mu)^2} \right)}{\sqrt{\delta^2 + (x - \mu)^2}}. \quad (4.1)$$

Here, K_λ is the modified Bessel function of the second order with parameter λ , which can be represented by the integral

$$K_\nu(z) = \frac{\sqrt{\pi}z^\nu}{2^\nu\Gamma(\nu + \frac{1}{2})} \int_1^\infty e^{-zt}(t^2 - 1)^{\nu-\frac{1}{2}} dt,$$

for $\nu > -\frac{1}{2}$ and $z > 0$. We apply the density method to find a sensitivity weight Π : A direct differentiation gives

$$-\partial \ln P_{\text{NIG}}(x) = -\beta + \frac{x - \mu}{\delta^2 + (x - \mu)^2} \left\{ 1 - \frac{\alpha\sqrt{\delta^2 + (x - \mu)^2} K_1'(\alpha\sqrt{\delta^2 + (x - \mu)^2})}{K_1(\alpha\sqrt{\delta^2 + (x - \mu)^2})} \right\}.$$

We can now use the recursive relation for the derivative of the Bessel function K_λ , saying (see e.g. Rydberg [13])

$$K_1'(x) = -\frac{1}{2}K_0(x) - \frac{1}{2}K_2(x).$$

Using the recursion $K_2(z) = K_0(z) + (2/z)K_1(z)$ we reach

$$K_1'(z) = -\frac{1}{z}K_1(z) - K_0(z).$$

Inserting this into the expression of $-\partial \ln p_{\text{NIG}}$ yields,

$$\begin{aligned} \Pi &= -\partial \ln p_{\text{NIG}}(L(1)) \\ &= -\beta + \frac{L(1) - \mu}{\delta^2 + (L(1) - \mu)^2} \left\{ 2 + \alpha\sqrt{\delta^2 + (L(1) - \mu)^2} \frac{K_0(\alpha\sqrt{\delta^2 + (L(1) - \mu)^2})}{K_1(\alpha\sqrt{\delta^2 + (L(1) - \mu)^2})} \right\}. \end{aligned} \tag{4.2}$$

Since this is a function of $L(1)$, it will be a variance optimal weight.

In applying the Monte Carlo simulation technique, it may be rather cumbersome to calculate the two modified Bessel functions K_1 and K_2 in order to calculate an outcome of the sensitivity weight Π . In fact, for each draw we must perform such a calculation, which makes the method very inefficient due to the heavy computational burden involved in calculating Bessel functions. An alternative will then be to use an approximation, like for instance considering the smoothed random variable $\widehat{L}_\varepsilon(1)$ defined in (3.10). Using the conditional density argument, we find that the delta can be calculated by the expectation operator

$$\widehat{F}'_\varepsilon(x) = \mathbb{E} \left[f(x + L(1) + \varepsilon B(1)) \frac{B(1)}{\varepsilon} \right].$$

Hence, rather than doing numerical calculation of Bessel functions, we simulate from a normal distribution. From the analysis in this paper, letting $\varepsilon \rightarrow 0$ brings us back to the derivative we are interested in. Hence, for small ε 's, $\widehat{F}'_\varepsilon(x)$ should be reasonably close to $F'(x)$. We have tested this numerically in the following examples.

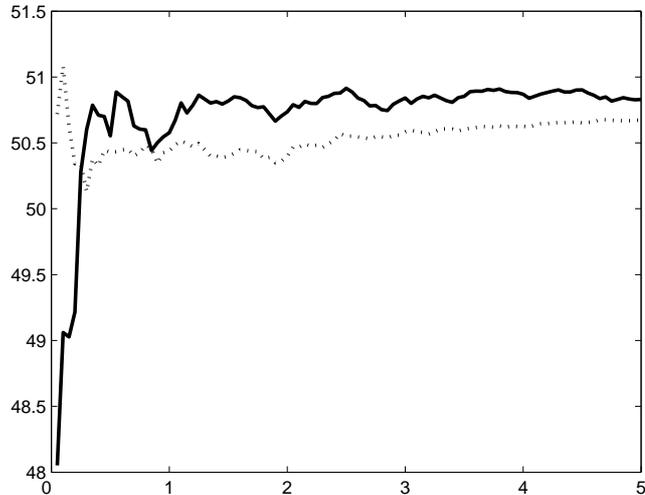


Figure 1: The estimated derivative based on the density method (broken line) versus an approximation using an added Brownian motion (solid line) as a function of the number of samples (in 10^5).

Let $\alpha = 50$, $\beta = \mu = 0$ and $\delta = 0.015$. These figures are not unreasonable estimates for the logreturns of a stock price on a daily scale, see Rydberg [13]. Further, we consider a function f being the payoff from a call option with strike $K = 100$, that is,

$$f(x) = \max(0, \exp(x) - 100).$$

We implemented to density method in Matlab by sampling a NIG-distribution using the technique in Rydberg [13] and calculating the Bessel functions $K_0(z)$ and $K_1(z)$ using the built-in Matlab function `besselk`. The approximation $\widehat{F}'_\varepsilon(x)$ was calculated by drawing samples from a standard normal distribution.

In Figure 1 we show the resulting derivatives for $x = \ln(S_0)$ with $S_0 = 100$ and $\varepsilon = 0.01$. Along the horizontal axis we have the number of samples (in 10^5) used in the estimation of the expectation operator, and the two expressions are calculated using common random numbers. The density method is depicted with a broken line, and we see that it has slightly less variance than the approximated derivative. But looking at the scale on the vertical axis, the approximation is pretty good, although it seems that it is slightly overestimating the true derivative. By reducing ε , we observe a convergence towards $F'(x)$, however, at the expense of a higher variance in the estimation of the expectation in $\widehat{F}'_\varepsilon(x)$. This is shown in Figure 2, where we plot the estimates as function of samples for three different values of ε , $\varepsilon = 0.01, 0.005$ and $\varepsilon = 0.001$. The smaller ε , the higher variance, which leads to a higher number of samples for ensuring accuracy of the estimate. We note that our numerical example covers the delta of an at-the-money call option on a stock, where the delta is calculated one time step (one day, say) prior to the exercise date of the option. We get the delta by dividing the derivatives $F'(x)$ and $\widehat{F}'_\varepsilon(x)$ by $S_0 = 100$, resulting from an application of the chain rule.

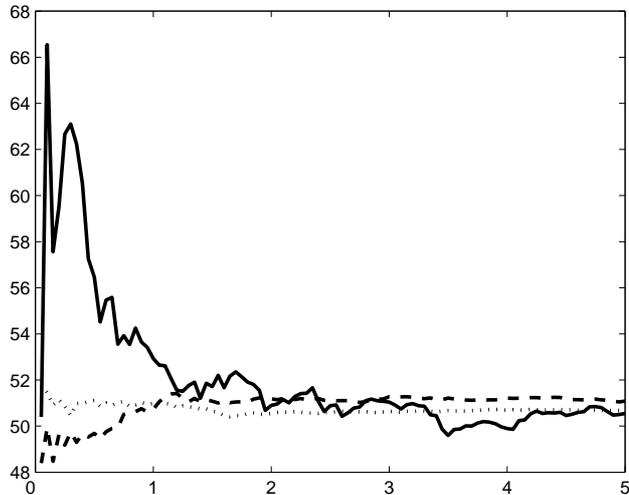


Figure 2: The estimated derivative based on an approximation using an added Brownian motion as a function of the number of samples (in 10^5). $\varepsilon = 0.01$ in dotted line, $\varepsilon = 0.005$ in broken line and $\varepsilon = 0.001$ in solid line.

To test our method on discontinuous functions f , we considered the above set-up for a digital option, that is, a payoff function $f(x) = 1(e^x > K)$ for some positive threshold K . The simulations showed that the derivative $\widehat{F}'_\varepsilon(x)$ had a significantly higher variance when estimated by Monte Carlo simulations. In fact, one needed to choose number of samples several scales above what was required for the call option in order to get reasonable estimates for the approximated derivative. Hence, from a numerical perspective, discontinuous functions f seem to behave badly under approximations by Brownian motions when one applies it to pure-jump Lévy processes. Variance reducing techniques like *quasi*-Monte Carlo simulations may be fruitful and speed up the convergence in such situations.

As a final note in this numerical subsection, let us briefly discuss the issues concerning approximating the small jumps of a Lévy process by a Brownian motion. Following the idea in this paper, the small jumps are approximated by $\sigma(\varepsilon)B(t)$ for a suitable scaling $\sigma(\varepsilon)$. The sensitivity weight is of the form

$$\Pi = \frac{B(1)}{\sigma(\varepsilon)},$$

if we have no continuous martingale part in the Lévy process and decide to use the density method with respect to the Brownian motion B . In order to simulate $F'_{\varepsilon,B}(x)$, we must sample from $B(1)$ and $L_\varepsilon(1)$. The latter is equivalent to sample from a compound Poisson process since we have only jumps of size bigger than ε . Indeed, we must sample from a compound Poisson process with jump size distribution given by

$$\mathbf{1}_{|z| \geq \varepsilon} Q(dz)/c$$

where the normalizing constant c is defined as

$$c = Q(|z| \geq \varepsilon).$$

The jump intensity will be c . This is in principle simple to simulate as long as one has a routine to sample for the truncated Lévy measure *and* knows the constant c . However, using for instance Markov Chain Monte Carlo methods, one can sample from the jump distribution without knowing the constant c .

5 Conclusion

In this study we have considered the problem of robustness of the sensitivity parameter delta to model choice. our models are selected within the Lévy family, but they differ according to how the presence of small jumps is taken into account.

First, following the study in Asmussen and Rosinski [1], we have considered models with small jumps, see $L(t)$ in (2.1) and their approximations given by models with of type $L_\varepsilon(t)$ (3.1), where a continuous martingale part with controlled standard deviation is replacing the small jumps. In this case both models have the same total variance. In this case $L_\varepsilon(t) \rightarrow L(t)$, for $\varepsilon \downarrow 0$. Secondly, we have considered a smoothing $\widehat{L}_\varepsilon(t)$ of the Lévy process $L(t)$. Also in this case we have $\widetilde{L}_\varepsilon(t) \rightarrow L(t)$, for $\varepsilon \downarrow 0$, but there is no control on the variances between the two models. The two situations can be usefully applied in different contexts.

In both cases we have addressed the question of the robustness of the parameter delta

$$\begin{aligned} F'(x) &= \frac{d}{dx} E[f(x + L(t))] \\ F'_\varepsilon(x) &= \frac{d}{dx} E[f(x + L_\varepsilon(t))] \\ \widehat{F}'_\varepsilon(x) &= \frac{d}{dx} E[f(x + \widehat{L}_\varepsilon(t))]. \end{aligned}$$

We have applied different methods of computation: the classical density method and the newly introduced conditional density method. The different computational techniques for the delta lead to different weights. However the values of the parameter is the same. Qualitatively, this last one is an application of computations similar to the ones in the density method, but applied after having performed some conditioning (this inspired by the Malliavin method á la Davis and Johansson [7]). In our analysis we have considered functions f with different degrees of regularity, always keeping in mind the needs coming from applications to finance and insurance. Our examples include, for example, also the digital option.

Indeed a robustness result is proved, i.e.

$$\begin{aligned} F'_\varepsilon(x) &\rightarrow F'(x), \quad \varepsilon \downarrow 0 \\ \widehat{F}'_\varepsilon(x) &\rightarrow F'(x), \quad \varepsilon \downarrow 0. \end{aligned}$$

If this is reassuring when coming to applications, we also remark that we experience some curious situations important from the numerical point of view. In fact, according

to the different methods applied some representations of the deltas turn out to be highly inefficient. This is evident when we consider models $L(t)$ with no original continuous martingale component (i.e. $b = 0$ in (2.1)) and we take the corresponding $L_\varepsilon(t)$ as approximating model. In this case the conditional density method shows an exploding variance of the random variable that must be simulated. This yielding to the need of a large number of samples to get some confident estimate of the delta.

Our future studies will direct in two directions. On one side we will consider the problem of robustness of the delta to model choice for path dependent options, i.e. for functional f . This will require techniques of Malliavin calculus. On the other side we will consider the problem of robustness of prices to model choice under change of measure. To explain, when dealing with no-arbitrage pricing an equivalent martingale measure (EMM) is required. This is a probability measure under which the discounted price processes are martingales. In incomplete market models, such as those driven by $L(t)$, the EMMs available are infinite. However, they do not correspond to the set of EMMs for $L_\varepsilon(t)$ or $\hat{L}_\varepsilon(t)$. Our goal is to understand the robustness of pricing when different models are considered.

References

- [1] Asmussen, S., and Rosinski, J. (2001). Approximations of small jump Lévy processes with a view towards simulation. *J. Appl. Prob.*, **38**, pp. 482–493.
- [2] Barndorff-Nielsen, O. E. (1998). Processes of normal inverse Gaussian type. *Finance Stoch.*, **2**(1), pp. 41–68.
- [3] Broadie, M., and Glasserman, P. (1996). Estimating security price derivatives using simulation. *Manag. Science*, **42**, pp. 169–285.
- [4] Carr, P., Geman, H., Madan, D., and Yor, M. (2002). The fine structure of asset returns: an empirical investigation. *J. Business*, **75**(2), pp. 305–332.
- [5] Carr, P., and Madan D. B. (1998). Option valuation using fast Fourier transform. *J. Comp. Finance*, **2**, pp. 61–73.
- [6] Cont, R., and Tankov, P. (2004). *Financial Modelling with Jump Processes*. Chapman Hall.
- [7] Davis, M. H. A., and Johansson, M. P.(2006). Malliavin Monte Carlo Greeks for jump diffusions. *Stoch. Processes. Appl.*, **116**(1), pp. 101–129.
- [8] Eberlein, E. (2001). Applications of generalized hyperbolic Lévy motion to finance. In *Lévy Processes – Theory and Applications*, Barndorff-Nielsen, O. E., Mikosch, T., and Resnick, S (eds.), Birkhäuser: Boston, 2001, pp. 319–336.
- [9] Eberlein, E., and Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli*, **1**, pp 281–299.

- [10] Folland, G. B. (1984). *Real Analysis – Modern Techniques and their Applications*. John Wiley & Sons.
- [11] Fournié, E., Lasry, J. M., Lébucheux, J., Lions, P. L., and Touzi, N. (1999). Applications of Malliavin calculus to Monte Carlo methods in finance, *Finance Stoch.*, **3**, pp 391–412.
- [12] Orey, S. (1968). On continuity properties of infinitely divisible distribution functions. *Annals Math. Statistics*, **39**, pp. 936–937.
- [13] Rydberg, T. H. (1997). The normal inverse Gaussian Lévy process: simulation and approximation, *Comm. Stat. Stoch. Models*, **13**, pp. 887–910.
- [14] Sato, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press.