# Remarks on Lévy Process Simulation


**Søren Asmussen**

**Abstract** Algorithms for simulation of a Lévy process $X(t)$ are discussed, with particular emphasis on two algorithms approximating jumps that are in some sense small. One is classical, defining small jumps as those of absolute value $<\varepsilon$. The other one appears to be new and relies on an completely monotone structure of the Lévy density $n(x)$. One then truncates the representing measure of $n(x)$ to $[0, A]$, meaning that jumps of mean $<1/A$ are left out. In both algorithms, the large jump part is simulated as compound Poisson and the small jumps are approximated. The standard choice of such an approximation is normal with the same mean and variance, but we also consider gamma approximations in two variants, and show that in some cases these perform substantially better. Other algorithms are briefly surveyed and we sketch a new one for simulation of a tempered stable (CGMY) process with infinite variation.

**Keywords** Acceptance-rejection · Complete monotonicity · Conditional Monte Carlo · Lévy measure · Tempered stable process

## 1 Introduction

A Lévy process $X(t)$ has the structure $X(t) = at + \sigma W(t) + J(t)$ where $W(t)$ is standard Brownian motion (BM) and $J(t)$ an independent pure jump process (see further below). This class of processes has been used in numerous application areas, of which we in particular mention finance [14, 35] and queueing [15].

Calculations for a Lévy process are, however, in general more difficult than for BM, and an abundance of expressions that are explicit for BM are not so even in the most popular parametric Lévy models. Simulation of $X(t)$ is therefore one of the main computational tools. For example in finance, it is most often the simplest vehicle for evaluating option prices of the form

S. Asmussen (✉)
Department of Mathematics, Aarhus University, Ny Munkegade, 8000 Aarhus C, Denmark
e-mail: bastin@iro.umontreal.ca



Z. Botev et al. (eds.), *Advances in Modeling and Simulation*,
https://doi.org/10.1007/978-3-031-10193-9_2

$$\mathbb{E}\Psi\big(X(0:T)\big) \tag{1}$$

where $T$ is the maturity time, $X(0:T)$ stands for the whole path $\big\{X(t)\big\}_{0\le t\le T}$, and $\Psi$ is a suitable path functional.

The simulation of the $at+\sigma W(t)$ component is straightforward, so we assume that $X(t)$ is a pure jump process. The main characteristic of such a process is its Lévy measure $\nu$, which with a few exceptions we throughout assume absolutely continuous with density $n(x)\ge 0$. Conditions needed on $\nu$ are $\int_{|x|>\varepsilon}\nu(\mathrm{d}x)<\infty$ and $\int_{|x|\le\varepsilon}x^2\nu(\mathrm{d}x)<\infty$ for some (and then all) $\varepsilon>0$. The process is said to have finite activity if $\lambda=\int_{\mathbb{R}}\nu(\mathrm{d}x)<\infty$ and is then a compound Poisson process with Poisson rate $\lambda$ and density $n(x)/\lambda$ of the jumps. Sample paths of $X(t)$ are of finite variation if and only if $\int_{|x|\le\varepsilon}x\nu(\mathrm{d}x)<\infty$. The picture is roughly that jumps in $[x,x+\mathrm{d}x)$ occur at Poisson rate $n(x)\,\mathrm{d}x$ and independently for different values of $x$. In the infinite variation case, $X(t)$ is, however, only completely specified by $n(x)$ up to a drift term (see further Sect. 2). The process is called spectrally positive or negative if $n(x)\equiv 0$ for $x<0$, resp. $x>0$; otherwise, we refer to it as two-sided. In finance, the most popular classes of jump processes are the NIG (Normal Inverse Gaussian), tempered stable (TS or CGMY), VG (Variance Gamma) and Meixner ones, and we survey these in Sect. 3.

Exact simulation of the whole path $X(0:T)$ is obviously impossible due to the presence of infinitely many jumps of the process. One could hope that one can perform exact simulation of $X(T)$ for any given $T$ and thereby a discrete skeleton $X(h), X(2h),\dots$ for any $h$. As surveyed briefly in Sect. 8, this is simple for VG, with a little added effort also possible for NIG and CGMY with finite variation, and presumably possible but quite tedious for Meixner. In general, this is however not feasible and we focus on two approximative alternatives. They both consist in simulating the finite number of jumps which are in some sense "big" as a compound Poisson process, and replacing the infinity of the remaining "small" ones with an easily simulated approximation. The path $X(0:T)$ can then by obtaining by assigning i.i.d. uniform $[0,T]$ location to the jumps and possibly filling in some information provided by the particular form of the approximation. The first of these approaches is classical and widely applied, and simply defines the big jumps as those of absolute value $>\varepsilon$; we refer to this as the $\varepsilon$-algorithm. These jumps are those coming from the part of $\nu$ concentrated on $\{|x|>\varepsilon\}$. By definition, this is a finite measure and so the corresponding contribution to $X$ can be simulated as a compound Poisson process. The second approach, which does not appear to have been considered in the simulation literature, relies on a completely monotone (CM) structure

$$n(x) \;=\; \int_0^\infty \mathrm{e}^{-xt}\,V(\mathrm{d}t) \;=\; \int_0^\infty \mathrm{e}^{-xt}v(t)\,\mathrm{d}t \tag{2}$$

of the Lévy density where $V$ is a Radon measure with density $v$. This holds in many main examples and represents the jumps as an infinite mixture of exponential($t$) jumps with the rate $t$ having weight $v(t)/t$ (see further Sect. 5). The compound

Poisson part is then obtained by restricting $V$ to $(0, A)$ for some $A < \infty$, meaning that exponential jumps with mean $< 1/A$ are left out. We refer to the method as the CM-algorithm. In both approaches, the computational effort as measured by the Poisson mean goes to infinity as $\varepsilon \to 0$, resp. $A \to \infty$. As for the approximation of the small jump part, the standard choice in the $\varepsilon$-algorithm is a normal with the same mean and variance and is substantiated in [6] by a limit result as $\varepsilon \to 0$ (further relevant references pertaining to this are [16, 37]). However, we shall also consider gamma alternatives in 2–3 variants and illustrate by examples that these perform at least as well, in some cases even convincingly better. Doing so, our point of view is largely empirical: for the practitioner, comparison of approaches as $\varepsilon \to 0$ matters less than performance for $\varepsilon$ so moderate that the computational effort is within reach. As $\varepsilon \to 0$, the small jumps contribute less, and hence limit results become less relevant. Similar remarks apply to the CM-algorithm. We also point out that in some types of applications, the approximation of the small jumps need not necessarily be simulated, but instead it may be used via conditional Monte Carlo for providing smooth density estimates and variance reduction.

## 2 Lévy Processes

For the general theory of Lévy processes, see e.g. [33] and [10]. A jump process is constructed from a Poisson random measure $L(\mathrm{d}t, \mathrm{d}x)$ on $(0, \infty) \times \mathbb{R}/\{0\}$ with intensity measure $\mathrm{d}t \otimes \nu(\mathrm{d}x)$. In the finite variation case $\int |x|\, \nu(\mathrm{d}x) < \infty$, one has

$$X(t) \;=\; \int_{s \le t,\, x \in \mathbb{R}} x\, L(\mathrm{d}s, \mathrm{d}x)\,, \quad \kappa(\theta) \;=\; \int_{-\infty}^{\infty} \left(\mathrm{e}^{\theta x} - 1\right) \nu(\mathrm{d}x) \tag{3}$$

at least for $\Re(\theta) = 0$ and in our examples in a strip containing the imaginary axis. Here $\kappa(\theta) = \log \mathbb{E}\mathrm{e}^{\theta X(1)}$ is the so-called Lévy exponent or cumulant function. In the infinite variation case, there are too many small jumps for these integrals to converge. Instead, so-called compensation is needed and consists in appropriate centerings and limits. Traditionally, jumps of absolute size $< 1$ are centered, which leads to

$$X(t) \;=\; at + \lim_{\varepsilon \to 0} \left\{ \int_{s \le t,\, \varepsilon < |x| < \infty} x\, L(\mathrm{d}s, \mathrm{d}x) - t \int_{\varepsilon < |x| \le 1} x\, \nu(\mathrm{d}x) \right\}, \tag{4a}$$

$$\kappa(\theta) \;=\; a + \int_{-\infty}^{\infty} \left(\mathrm{e}^{\theta x} - 1 - \theta x \mathbb{I}(|x| \le 1)\right) \nu(\mathrm{d}x) \tag{4b}$$

for some $a$. Obviously, taking 1 as truncation point is arbitrary, and other choices lead to different values of $a$. If the mean $\mathbb{E}X(1) = \kappa'(0)$ is finite, it may be more convenient to center all jumps, and one then has

$$X(t) \;=\; t\kappa'(0) + \lim_{\varepsilon\to 0}\left\{\int_{s\le t,\,|x|>\varepsilon} x\,L(\mathrm{d}s,\mathrm{d}x) - t\int_{|x|>\varepsilon} x\,\nu(\mathrm{d}x)\right\}, \tag{5a}$$

$$\kappa(\theta) \;=\; \kappa'(0) + \int_{-\infty}^{\infty}\left(\mathrm{e}^{\theta x} - 1 - \theta x\right)\nu(\mathrm{d}x)\,. \tag{5b}$$

The cumulants $\kappa_k$ of $X(1)$ are given as the $k$th derivatives $\kappa^{(k)}(0)$ of $\kappa(\theta)$ at $\theta = 0$. In particular, $\kappa_1 = \mathbb{E}X(1)$, $\kappa_2 = \mathbb{V}arX(1)$, and the skewness and (excess) kurtosis are $\kappa_3/\kappa_2^{3/2}$, resp. $\kappa_4/\kappa_2^2$. For $k \ge 2$, one alternatively has

$$\kappa_k \;=\; \int_{-\infty}^{\infty} x^k\,\nu(\mathrm{d}x), \tag{6}$$

and this expression is also valid for $k = 1$ in the finite variation case.

## 3 Main Examples

In the absolutely continuous case, define the Lévy density $n(x) = \mathrm{d}\nu(x)/\mathrm{d}x$ as the density of the Lévy measure w.r.t. Lebesgue measure.

The NIG process [9] has parameters $\alpha, \delta > 0$, $\beta \in (-\alpha, \alpha)$ and $\mu \in \mathbb{R}$. The Lévy density is

$$n(x) \;=\; \frac{\alpha\delta}{\pi|x|}K_1\big(\alpha|x|\big)\mathrm{e}^{\beta x}, \quad x \in \mathbb{R}, \tag{7}$$

where as usual $K_1(z)$ denotes the modified Bessel function of the third kind with index 1. The cumulant function and the density of $X(1)$ are, respectively,

$$\kappa(s) \;=\; \mu s + \delta\left(\sqrt{\alpha^2-\beta^2} - \sqrt{\alpha^2-(\beta+s)^2}\right),\ \alpha-\beta < \Re(s) < \alpha+\beta\,,$$

$$\frac{\alpha\delta}{\pi}\exp\big\{\delta\sqrt{\alpha^2-\beta^2} + \beta(x-\mu)\big\}\frac{K_1\big(\alpha\sqrt{\delta^2+(x-\mu)^2}\big)}{\sqrt{\delta^2+(x-\mu)^2}}\,.$$

The Meixner (MX) process [18, 28, 35] has parameters $a, d > 0$, $b \in (-\pi, \pi)$ and $m \in \mathbb{R}$. The Lévy density is

$$n(x) \;=\; d\frac{\exp\{bx/a\}}{x\sinh(\pi|x|/a)} \;=\; 2d\frac{\exp\big\{bx/a - \pi|x|/a\big\}}{|x|\big(1-\exp\{-2\pi|x|/a\}\big)}\,. \tag{8}$$

The cumulant function and the density of $X(1)$ are, respectively,

$$\kappa(s) \;=\; 2d\log\bigl(\cos(b/2)\bigr) - 2d\log\bigl(\cos(as+b)/2)\bigr) + ms\,, \quad \tfrac{\pi+b}{a} < \Re(s) < \tfrac{\pi-b}{a}\,,$$
$$\frac{(2\cos(b/2))^{2d}}{3a\pi\,\Gamma(2d)}\mathrm{e}^{b(x-m)/a}\bigl|\Gamma(d+\mathrm{i}(x-m)/a)\bigr|^2\,. \tag{9}$$

For the tempered stable (TS) process [3, 12, 24]

$$n(x) \;=\; \delta_{\pm}\mathrm{e}^{-\beta_{\pm}|x|}/|x|^{\alpha_{\pm}+1} \tag{10}$$

where $\delta_+, \beta_+$ are for $x > 0$ and $\delta_-, \beta_-$ for $x < 0$. When $\delta_+ = \delta_-, \alpha_+ = \alpha_-$, the TS process goes under the acronym CGMY process in particular in finance, where the traditional notation is $\delta_+ = \delta_- = C, \alpha_+ = \alpha_- = Y, G$ instead of $\beta_-$ and $M$ instead of $\beta_+$. Cf. the author names in [12]! In terms of the positive jumps, $\alpha_+ < 0$ corresponds to a compound process, $\alpha_+ = 0$ to a gamma process where $X(1)$ is gamma distributed with shape parameter $\delta_+$ and rate parameter $\beta_+$, $0 < \alpha_+ < 1$ to infinite activity but finite variation, and $1 \le \alpha_+ < 2$ to infinite variation. The cumulant function is

$$\kappa(s) \;=\; \delta_-\Gamma(-\alpha_-)\bigl((\beta_- + s)^{\alpha_-} - \beta_-^{\alpha_-}\bigr) + \delta_+\Gamma(-\alpha_+)\bigl((\beta_+ - s)^{\alpha_+} - \beta_+^{\alpha_+}\bigr)\,, \tag{11}$$

$-\beta_- < \Re(s) < \beta_+$. Here and at other places in the theory, exceptions apply when $\alpha_+$ or $\alpha_-$ or both equals 0 or 1. The case $\alpha_+ = \alpha_- = 0$ is the VG process (the difference between two gamma processes).

Starting from [12, 13], the density of $X(1)$ in the TS process has traditionally been computed by Fourier inversion via (11). However, it is pointed out in [3] that the density can be expressed as

$$f(x) \;=\; \exp\{-\beta x - \delta\Gamma(-\alpha)\beta^{\alpha}\}\, f_0(x) \tag{12}$$

where $f_0$ is the density of a strictly $\alpha$-stable distribution $S_\alpha(\sigma, 1, 0)$ distribution with $\sigma = \bigl(-\delta\Gamma(-\alpha)\cos(\pi\alpha/2)\bigr)^{1/\alpha}$. See also [27, 30]. Given the availability of software for stable distributions, (12) provides an easy approach to numerical computations.

In all these examples, one has

$$n(x) \;\sim\; \frac{\delta}{x^{1+\alpha^*}} \quad \text{as } x \downarrow 0 \tag{13}$$

for some $\delta$ and some $\alpha^* \in [0, 2)$ (subject to this, $\alpha^*$ is sometimes referred to as the Blumenthal-Getoor index). In fact, for TS this holds since $\mathrm{e}^{-\beta x} \to 1$, whereas one has $\alpha^* = 1$ for NIG and MX, as follows from known asymptotics of $K_1$, resp. $1 - \exp\{-2\pi x a\} \sim 2\pi x/a$.

## 4 The $\varepsilon$–Algorithm

Typically, the positive and negative jumps are simulated separately, so we consider only the spectrally positive case in the following.

When truncating the jumps to $[\varepsilon, \infty)$, the exactly simulated compound Poisson part of $X(1)$ is $X_{\varepsilon,\infty}(1) = \sum_1^N Y_n(\varepsilon)$ where $N$ is Poisson $\lambda(\varepsilon)$ and $Y_1(\varepsilon), Y_2(\varepsilon), \ldots$ are i.i.d. with density $g(x;\varepsilon)$ with

$$\lambda(\varepsilon) = \int_\varepsilon^\infty n(x)\,\mathrm{d}x\,, \quad g(x;\varepsilon) = \frac{n(x)}{\lambda(\varepsilon)}\,, \ \varepsilon < x < \infty.$$

Some approximation $\widehat{X}_{0,\varepsilon}(1)$ of jumps of value $< \varepsilon$ is then used, and one returns the r.v. $\widehat{X}_{0,\varepsilon}(1) + X_{\varepsilon,\infty}(1)$. For these approximations, one typically needs the cumulants of $X_{0,\varepsilon}(1)$ which according to (6) are $\kappa_{k;0,\varepsilon} = \int_0^\varepsilon x^k\,\nu(\mathrm{d}x)$ if either $k \ge 2$ or $k \ge 1$ and the process has finite variation; in the infinite variation case, $\kappa_{1;0,\varepsilon} = 0$ subject to (5a). In practice, $\int_0^\varepsilon x^k\,\nu(\mathrm{d}x)$ is seldom explicit, but needs to be evaluated by numerical integration. Alternatively, one may note that subject to (13), one has

$$\kappa_{k;0,\varepsilon} \;=\; \int_0^\varepsilon x^k\,\nu(\mathrm{d}x) \;\sim\; \delta\frac{\varepsilon^{k-\alpha^*}}{k-\alpha^*} \quad \text{if } \alpha^* < 1,\ k \ge 1 \text{ or } 1 \le \alpha^* < 2,\ k \ge 2. \tag{14}$$

The most naive choice is $\widehat{X}_{0,\varepsilon}(1) \equiv 0$. However, it was suggested in [11] and [32] to take $X_{0,\varepsilon}(t)$ as a BM with fitted mean and variance when $\varepsilon < 1$. Supporting limit theorems were given in [6], establishing the validity of this procedure when $X$ is not too close to the finite activity case $\int \nu(\mathrm{d}x) < \infty$ and $\nu$ satisfies some weak smoothness conditions (a simple proof under the stronger condition (13) follows by paralleling the proof of Proposition 3 below). We shall here suggest gamma alternatives in two variants.

Recall that the gamma distribution with shape parameter $r$ and rate parameter $b$ has density $b^r x^{r-1}\mathrm{e}^{-bx}/\Gamma(r)$ and cumulant function $\log\bigl(b/(b-z)\bigr)^r = -r\log(1-z/b)$ with $k$th derivative $r(k-1)!b^{-k}(1-z/b)^{-k}$. Thus the $k$th cumulant is $\kappa_k = r(k-1)!/b^k$; in particular the skewness is $(2r/b^3)/(r/b^2)^{3/2} = 2r^{-1/2}$. Given a distribution or a set of data with cumulants $\kappa_k^\#$, the most obvious possibility is to fit the mean and variance which leads to

$$b = \frac{\kappa_1^\#}{\kappa_2^\#}, \quad r \;=\; b\kappa_1^\# \;=\; \frac{\kappa_1^{\#2}}{\kappa_2^\#}\,. \tag{$\Gamma_1$}$$

One could also consider a three-parameter gamma family by allowing a shift $m$, and fitting the mean, variance and skewness then gives

$$r \;=\; \frac{4\kappa_2^{\#3}}{\kappa_3^{\#2}}\,, \quad b = \sqrt{\frac{r}{\kappa_2^\#}} \;=\; \frac{2\kappa_2^\#}{\kappa_3^\#}\,, \quad m \;=\; \kappa_1^\# - \frac{r}{b}. \tag{$\Gamma_2$}$$

Note that for a Lévy process, $(\Gamma_1)$ does not make sense in the infinite variation case since then $\kappa_{1;0,\varepsilon} = 0$ subject to (5). For a subordinator (a spectrally positive process with a non-negative linear drift), $(\Gamma_2)$ may be controversial because it may destroy the property of the process being non-decreasing. The normal approximation has the same problem, but not $(\Gamma_1)$. Both of $(\Gamma_1)$, $(\Gamma_2)$ asymptotically agree with the normal approximation as $\varepsilon \downarrow 0$. This follows since (14) implies that $b \to \infty$ in both cases, which implies a gamma distribution to be asymptotically normal.

Efficiently generating r.v.'s from the density $g(x;\varepsilon) = n(x)/\lambda(\varepsilon)$, $x > \varepsilon$ may not always be trivial. However, a general set-up covering many examples is

$$(n_1 n_2) \quad n(x) = \frac{n_1(x)}{n_2(x)} \text{ for } x > 0 \text{ with } n_1(x) \text{ strictly decreasing, } n_2'(x) > 0,$$
$$n_1(x) \text{ integrable on } (x_0, \infty) \text{ and } 1/n_2(x) \text{ on } (\varepsilon, x_0) \text{ for all } 0 < \varepsilon < x_0 < \infty.$$

In the TS situation, $n_1(x) = d\mathrm{e}^{-\beta x}$, $n_2(x) = x^{1+\alpha}$; for the positive jumps of MX, one may take $n_1(x) = 2d\exp\{bx/a - \pi x/a\}$, $n_2(x) = x\big(1 - \exp\{-2\pi x/a\}\big)$; etc.

Even for the TS case, the c.d.f. of $g(x;\varepsilon)$ is not explicitly available. Thus inversion is not feasible and acceptance-rejection (A-R) seems the reasonable approach. What suggests itself is to either use the exponential$(\beta)$ distribution on $(\varepsilon, \infty)$ as proposal and reject w.p. proportional to $1/x^{1+\alpha}$, or to use the Pareto$(\alpha)$ distribution on $(\varepsilon, \infty)$ as proposal and reject w.p. proportional to $\mathrm{e}^{-\beta x}$. However, the first procedure would lead to a high rejection rate for small or moderate $x$, and the second for large or moderate $x$. So, a reasonable compromise is to choose some threshold $x_0$ and use the Pareto proposal on $(\varepsilon, x_0)$ and the exponential on $(x_0, \infty)$. An equivalent formulation is to decompose $X_{\varepsilon,\infty}$ into two compound Poisson terms, one having jumps in $(\varepsilon, x_0]$ and the other having jumps in $(x_0, \infty)$. Note that the proposal on $(\varepsilon, x_0)$ (a truncated Pareto) is easily simulated by inversion as $\big(1/\varepsilon^\alpha - \alpha\mu_2(x_0)U\big)^{-1/\alpha}$ with $U$ uniform(0, 1), cf. [4, p. 39].

In order to analyze this A-R procedure in the general set-up of $(n_1 n_2)$, define for a fixed $\varepsilon > 0$

$$\lambda_1(x_0) = \int_\varepsilon^{x_0} n(x)\,\mathrm{d}x\,, \quad \mu_1(x_0) = \int_\varepsilon^{x_0} \frac{1}{n_2(x)}\,\mathrm{d}x\,, \quad C_1(x_0) = \frac{n_1(\varepsilon)\mu_1(x_0)}{\lambda_1(x_0)}\,,$$
$$\lambda_2(x_0) = \int_{x_0}^{\infty} n(x)\,\mathrm{d}x\,, \quad \mu_2(x_0) = \int_{x_0}^{\infty} n_1(x)\,\mathrm{d}x\,, \quad C_2(x_0) = \frac{\mu_2(x_0)}{\lambda_2(x_0)n_2(x_0)}\,.$$

The target distributions are then

$$f_1(x) = \frac{n_1(x)}{\lambda_1(x_0)n_2(x)},\ \varepsilon < x < x_0, \quad \text{and} \quad f_2(x) = \frac{n_1(x)}{\lambda_2(x_0)n_2(x)},\ x_0 < x < \infty,$$

and the proposals are

$$g_1(x) = \frac{1}{\mu_1(x_0)n_2(x)},\ \varepsilon < x < x_0, \quad \text{and} \quad g_2(x) = \frac{n_1(x)}{\mu_2(x_0)},\ x_0 < x < \infty.$$

Then $f_1(x) \le C_1(x_0)g_1(x)$ and $f_2(x) \le C_2(x_0)g_2(x)$, and we may use A-R with acceptance probabilities

$$\frac{f_1(x)}{C_1(x_0)g_1(x)} = \frac{n_1(x)\mu_1(x_0)}{\lambda_1(x_0)C_1(x_0)}, \quad \frac{f_2(x)}{C_2(x_0)g_2(x)} = \frac{\mu_2(x_0)}{\lambda_2(x_0)C_2(x_0)n_2(x)}$$

for r.v. generation from $f_1$, resp. $f_2$. This gives expected numbers $C_1(x_0)$, $C_2(x_0)$ of samplings from $g_1(x)$, resp. $g_2(x)$, and as measure $E(x_0)$ of the computational effort, we shall use the total number of these samplings, i.e.

$$E(x_0) \;=\; \lambda_1(x_0)C_1(x_0) + \lambda_2(x_0)C_2(x_0) \;=\; n_1(\varepsilon)\mu_1(x_0) + \frac{\mu_2(x_0)}{n_2(x_0)}\,.$$

Of course, if the costs to generate from $g_1(x)$, resp. $g_2(x)$ are very different, $E(x_0)$ needs to be reflected to reflect this disparity.

**Proposition 1** *Consider the function $E(x_0)$, $\varepsilon \le x_0 \le \infty$, If $n_2'(x_0)/n_2(x_0) \to 0$ as $x_0 \to \infty$, then $E(x_0)$ attains its minimum for some $\varepsilon < x_0^* < \infty$ satisfying $\psi(x_0^*) = 0$ where $\psi(x_0) = n_2(x_0)\big(n_1(\varepsilon) - n_1(x_0)\big) - \mu_2(x_0)n_2'(x_0)$. In particular, for the TS case $x_0^*$ is the unique solution in $(\varepsilon, \infty)$ of*

$$x_0^*(\mathrm{e}^{\beta(x_0^*-\varepsilon)} - 1) \;=\; \frac{1+\alpha}{\beta}. \tag{15}$$

***Proof*** We have $\frac{\mathrm{d}}{\mathrm{d}x_0}E(x_0) = \psi(x_0)/n_2(x_0)^2$. Here $\psi(\varepsilon) = -\mu_2(\varepsilon)n_2'(\varepsilon) < 0$. As $x_0 \to \infty$, we have $\liminf\big(n_1(\varepsilon) - n_1(x_0)\big) > 0$ and $\mu_2(x_0) \to 0$, and so $n_2'(x_0)/n_2(x_0) \to 0$ implies $\psi(x_0) > 0$ for all large $x_0$. This gives the first part of the result. For the second on the TS case, we get

$$\psi(x_0) \;=\; x_0^{1+\alpha}(d\mathrm{e}^{-\beta\varepsilon} - d\mathrm{e}^{-\beta x_0}) - d(\mathrm{e}^{-\beta x_0}/\beta)\cdot(1+\alpha)x_0^{\alpha}\,.$$

Multiplying by $\mathrm{e}^{-\beta x_0}$ and rearranging shows that $\psi(x_0^*) = 0$ is the same as (15). For uniqueness of the solution, note that the l.h.s. of (15) is strictly increasing in $x_0^*$ with limits 0 at $x_0^* = \varepsilon$ and $\infty$ at $x_0^* = \infty$. □

## 5 Using Complete Monotonicity Structure

Again, we consider only the spectrally positive case and assume the Lévy measure $n(x)$ to be completely monotone in the sense of (2). We refer to the measure $V(\mathrm{d}t)$ as the reference measure and to $v(t)$ as the reference density. See, e.g., [34] for background on complete monotonicity and a huge list of examples. Motivation and financial examples are in [12, 19, 21].

**Example 1** We check here that complete monotonicity holds in our main examples. We use the rule that if $m(x)$ is completely monotone with reference density $v(t)$, $t > 0$, then $\mathrm{e}^{-\beta x} m(x)$ is completely monotone with reference density $v(t-\beta)$ for $t > \beta$ and $= 0$ for $0 < t < \beta$.

In the NIG case, this rule together with the standard formula $K_1(x) = x \int_1^\infty \mathrm{e}^{-xt}(t^2-1)^{1/2}\mathrm{d}t$ and elementary substitutions gives the expression

$$v(t) \;=\; \frac{\delta}{\pi}\sqrt{(t+\beta)^2-\alpha^2}\,, \quad t > \alpha - \beta,$$

for the reference density for the positive part of the Lévy measure. For MX, let $\chi(t) = \lceil t \rceil$ be the step function equal to $n+1$ for $t \in (n, n+1]$. Then

$$\begin{aligned}\frac{1}{1-\mathrm{e}^{-x}} \;&=\; 1 + \mathrm{e}^{-x} + \mathrm{e}^{-2x} + \cdots \;=\; 1 - \mathrm{e}^{-x} + 2(\mathrm{e}^{-x} - \mathrm{e}^{-2x}) + 3(\mathrm{e}^{-2x} - \mathrm{e}^{-3x}) + \cdots \\ &=\; \sum_{n=0}^{\infty} (n+1) x \int_n^{n+1} \mathrm{e}^{-xt}\,\mathrm{d}t \;=\; x \int_0^\infty \mathrm{e}^{-xt}\chi(t)\,\mathrm{d}t\end{aligned}$$

which gives

$$v(x) \;=\; 2d\,\chi\big(a(t-\pi/a+b/a)/(2\pi)\big)\,, \quad t > \pi/a - b/a.$$

Finally for the TS case, it is shown in [12] that $v(t) = \delta(t-\beta)^\alpha/\Gamma(1+\alpha)$, $t > \beta$, which in turn is an easy consequence of $\int_0^\infty \mathrm{e}^{-xt} t^\alpha\,\mathrm{d}t = \Gamma(1+\alpha)/x^{1+\alpha}$. ♢

In all three examples, the reference density $v(t)$ grows at rate $t^{\alpha^*}$ as $t \to \infty$, with $\alpha^*$ as in (13). This is in fact no coincidence since Feller's Tauberian theorem [17, p. 445] implies that $V(t) = \int_0^t v(s)\,\mathrm{d}s \sim \delta t^{1+\alpha^*}/\Gamma(2+\alpha^*)$. Hence by formal differentiation,

$$v(t) \;\sim\; \delta(1+\alpha^*)t^{\alpha^*}/\Gamma(2+\alpha^*) \;=\; \delta t^{\alpha^*}/\Gamma(1+\alpha^*). \tag{16}$$

We stress that this is formal: the known rigorous result in this direction requires (beyond existence of $v$) that $v$ is monotone, cf. [36]. However, we shall take (16) as an assumption for the further developments to follow.

In the following, we use that (2), (6) and Fubini's theorem give the representation

$$\int_0^\infty x^k n(x)\,\mathrm{d}x \;=\; \int_0^\infty \left(\int_0^\infty x^k \mathrm{e}^{-tx}\,\mathrm{d}x\right) v(t)\,\mathrm{d}t \;=\; \int_0^\infty \frac{k!}{t^{k+1}} v(t)\,\mathrm{d}t \tag{17}$$

of the cumulants for $k = 0, 1, \ldots$ As in Sect. 4, we decompose the Lévy density $n$ into two components, here taken as

$$n_{0,A}(x) \;=\; \int_0^A \mathrm{e}^{-xt} v(t)\,\mathrm{d}t\,, \quad n_{A,\infty}(x) \;=\; \int_A^\infty \mathrm{e}^{-xt} v(t)\,\mathrm{d}t\,.$$

The corresponding decomposition of $X$ is written as $X = X_{0,A} + X_{A,\infty}$. The key to our algorithm using complete monotonicity is the following:

**Proposition 2** *Assume the measure $V$ in (2) is finite and let*

$$\mu = V(\infty) = \int_0^\infty \frac{v(t)}{t}\,\mathrm{d}t\,, \quad \lambda = \int_0^\infty n(x)\,\mathrm{d}x\,.$$

*Then $\mu = \lambda$. Let further $T$ be standard exponential, $Y$ a independent r.v. with density $\frac{v(t)}{t\mu}$ and $Z$ one with density $n(x)/\lambda$. Then $T/Y = Z$ in distribution.*

***Proof*** Taking $k = 0$ in (17) gives $\lambda = \mu$. We then get

$$\begin{aligned}\mathbb{P}\big(T/Y \in \mathrm{d}x\big) &= \int_0^\infty \mathbb{P}\big(T/Y \in \mathrm{d}x \,\big|\, Y = t\big)\frac{v(t)}{t\mu}\,\mathrm{d}t \\ &= \int_0^\infty t\mathrm{e}^{-tx}\,\frac{v(t)}{t\mu}\,\mathrm{d}t \;=\; \frac{n(x)}{\mu} \;=\; \mathbb{P}(Z \in \mathrm{d}x)\,.\end{aligned}$$

□

This suggests that in the finite variation case, we can generate a r.v. $X$ approximately distributed as $X(1)$ as follows (more details on the individual steps are given below):

(1) Choose $A < \infty$, let $\lambda = \int_0^A v(t)/t\,\mathrm{d}t$ and generate $N$ as Poisson($\lambda$).
(2) Generate $X_1 = \sum_{n=1}^N T_n/Y_n(A)$ where the $T_n$ are standard exponential and the $Y_n(A)$ have density $v(t)/(\lambda t)$, $0 < t < A$.
(3) Generate $X_2$ as some approximation to $X_{A,\infty}(1)$.
(4) Return $X = X_1 + X_2$.
In the infinite variation case subject to (5), replace $X_1$ in (2) by

$$\sum_{n=1}^N \frac{T_n}{Y_n(A)} - \int_0^\infty x n_{0,A}(x)\,\mathrm{d}x \;=\; \sum_{n=1}^N \frac{T_n}{Y_n(A)} - \int_0^A \frac{v(t)}{t^2}\,\mathrm{d}t$$

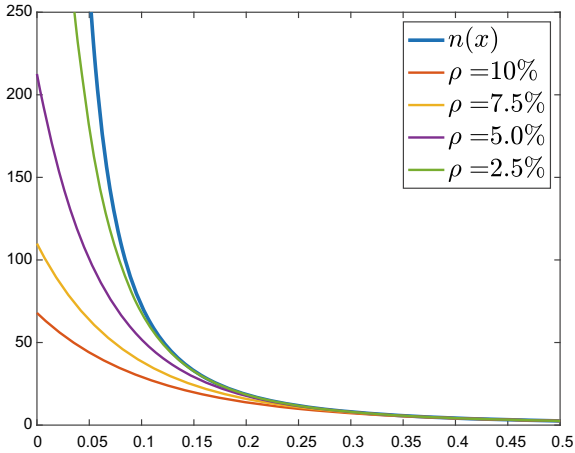
and $X$ in (4) by $\kappa'(0) + X_1 + X_2$. In both cases, $X \to X(1)$ as $A \to \infty$.

That $\lambda$ in (1) is finite follows by the Radon property of $V(\mathrm{d}x)$. The shape of the part $n_{0,A}$ of $n$ corresponding to the simulated large jumps is illustrated in Fig. 1, The process in the example is TS with $\alpha = 0.8$, variance $\kappa_2 = 1$, kurtosis $K = 2$ and there are 4 values of $A$ determined by the $\rho$ defined as the proportion $\mathbb{V}ar\big(X_{A,\infty}(1)\big)/\kappa_2$ of the total variance provided by the small jumps (see further Sect. 6).

As for the approximation in (3), the most obvious choice is a normal distribution with the correct mean and variance, and this is in fact justified by the following result (recall that $W$ denotes BM):

**Proposition 3** *Define $X^*_{A,\infty}(t) = \big(X_{A,\infty}(t) - t\mathbb{E}X_{A,\infty}(1)\big)\big/\sqrt{\mathbb{V}ar\,X_{A,\infty}(1)}$. Then $X^*_{A,\infty} \xrightarrow{\mathcal{D}} W$ in the Skorokhod space $D[0,\infty)$ as $A \to \infty$.*

**Fig. 1** $n(x)$ and $n_{0,A}(x)$




***Proof*** Let $\kappa_k^*$ be the $k$th cumulant of $X_{A,\infty}^*(1)$. Then $\kappa_k^*$ is of order $A^{\alpha-k}A^{(2-\alpha)k/2} = A^{\alpha((1-k/2)}$ for $k > 2$ since by (17)

$$\begin{aligned}\frac{\Gamma(1+\alpha)}{\delta}\int_0^\infty x^k n_{A,\infty}(x)\,\mathrm{d}x &= k!\int_A^\infty \frac{(t-\beta)^\alpha}{t^{k+1}}\,\mathrm{d}t \\ &\sim k!\int_A^\infty t^{\alpha-k-1}\,\mathrm{d}t = \frac{k!A^{\alpha-k}}{k-\alpha}.\end{aligned}$$

Hence $\kappa_k^* \to 0$ for $k > 2$ and obviously, $\kappa_1^* = 0$, $\kappa_2^* = 1$. Thus all cumulants and hence all moments of $X_{A,\infty}^*(1)$ converge to those of the standard normal r.v. $W(1)$. This implies $X_{A,\infty}^*(1) \xrightarrow{\mathcal{D}} W(1)$ (e.g. [22, Exercise 11 p. 101]), from which the asserted convergence in function space follows from Chap. 15 in [22]. □

Gamma distributions fitted by $(\Gamma_1)$ or $(\Gamma_2)$ are appealing alternatives to the normal approximation and perform again significantly better in the numerical examples to be given in Sect. 6. A gamma form of $n_{A,\infty}(x)$ comes up directly: one can use (16) and standard asymptotics of the upper incomplete gamma function to infer that

$$\begin{aligned}n_{A,\infty}(x) &\sim \int_A^\infty \mathrm{e}^{-tx}\delta t^\alpha/\Gamma(1+\alpha)\,\mathrm{d}t = \frac{\delta}{x^{1+\alpha}\Gamma(1+\alpha)}\int_{Ax}^\infty \mathrm{e}^{-y}y^\alpha\,\mathrm{d}y \\ &\sim \frac{\delta}{x^{1+\alpha}\Gamma(1+\alpha)}(Ax)^\alpha\mathrm{e}^{-Ax} \sim \frac{\delta}{\Gamma(1+\alpha)}\frac{A^\alpha}{x}\mathrm{e}^{-Ax}\end{aligned}$$

for any given fixed $x$. However, the first $\sim$ is not valid if $Ax$ is small or moderate, and in fact the gamma distribution with shape parameter $\delta A^\alpha/\Gamma(1+\alpha)$ and rate parameter $A$ substantially underestimates the order of $X_{A,\infty}(1)$. For example, its mean is 1.2 for $\alpha = 0.8$, $\kappa_2 = 2$, $K = 2$ and $\rho = 0.75$, whereas the correct value is $\mathbb{E}X_{A,\infty}(1) = 5.5$.

## 6 Numerical Examples

As illustration of the $\varepsilon$- and CM-algorithms, we considered spectrally positive TS processes with varying parameters. Such a process can be parametrized with the variance $\kappa_2$, the kurtosis $K$ and $\alpha$, and one then has

$$\beta = \sqrt{\frac{(2-\alpha)(3-\alpha)}{\kappa_2^2 K}}\,, \quad \delta = \frac{\kappa_2}{\Gamma(2-\alpha)}\beta^{2-\alpha}\,,$$

cf. [3]. We considered three values 0.2, 0.8, 1.4 of $\alpha$ and three 1/2, 2, 8 of $K$, and normalized by taking $\kappa_2 = 1$. We further considered the normal as well as the two gamma approximations $(\Gamma_1)$, $(\Gamma_2)$ of the small jumps, and as measure of performance, we took the $L^2$–distance

$$d \;=\; \int_0^\infty \big(f(x) - \widehat{f}(x)\big)^2 \mathrm{d}x \tag{18}$$

between the true density $f(x)$ of $X(1)$ and an estimate $\widehat{f}(x)$ provided by simulation. Here $f(x)$ was evaluated by (12), using the MATLAB routines for stable distributions. For $\widehat{f}(x)$, we simulated $M = 10^6$ replicates $Z_1, \ldots, Z_M$ of $X_{\varepsilon,\infty}(1)$ and used the conditional Monte Carlo estimator

$$\widehat{f}(x) \;=\; \frac{1}{M}\sum_{m=1}^{M} \xi(x - Z_m) \tag{19}$$

where $\xi(\cdot)$ is the density in the approximation in question for the density of $X_{(0,\varepsilon)}(1)$. Cf. e.g. [4, p. 146] and [2] (see also [26] for more sophisticated applications of the technique), but note also that conditional Monte Carlo can not universally replace generation of a r.v. distributed according to $\xi(\cdot)$; e.g., this is needed when simulating a discrete skeleton. Numerically, (18) was computed by a discrete approximation with step length 0.01 in the interval $\mathbb{E}X(1) \pm 3$ (recall that $X(1)$ was normalized to standard deviation 1).

The truncation parameters $\varepsilon$, resp. $A$, for the two algorithms were chosen such that the variance of the approximated small jumps equaled various fractions $\rho$ of the total variance $\kappa_2 = 1$ of all jumps. For the $\varepsilon$-algorithm, this means that for a given $\rho$

$$\rho \;=\; \int_0^\varepsilon x^2 \frac{\delta \mathrm{e}^{-\beta x}}{x^{1+\alpha}}\,\mathrm{d}x \;=\; \frac{\delta}{\beta^{2-\alpha}}\int_0^{\varepsilon\beta} y^{2-\alpha-1}\mathrm{e}^{-y}\,\mathrm{d}y = \frac{\delta}{\beta^{2-\alpha}}\Gamma(\varepsilon\beta, 2-\alpha)\Gamma(2-\alpha)$$

where $\Gamma(\cdot; 2-\alpha)$ is the lower incomplete Gamma function with parameter $2-\alpha$. Thus

$$\varepsilon \;=\; \frac{1}{\beta}\Gamma^{-1}\left(\frac{\rho\beta^{2-\alpha}}{\delta\Gamma(2-\alpha}; \; 2-\alpha\right)\,.$$

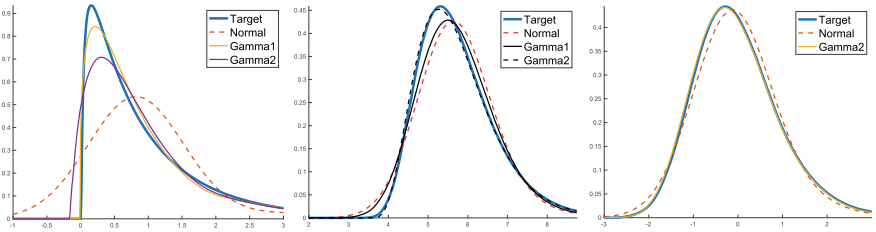



**Fig. 2** Left: $\varepsilon$-alg., $\alpha = 0.2$, $K = 2$, $\rho = 0.50$, $d_N = 1.3\,e^{-1}$, $d_{\Gamma 1} = 6.7\,e^{-3}$, $d_{\Gamma 2} = 3.5\,e^{-2}$; middle: CM-alg,. $\alpha = 0.8$, $K = 8$, $\rho = 0.75$, $d_N = 7.6\,e^{-3}$, $d_{\Gamma 1} = 2.6\,e^{-3}$, $d_{\Gamma 2} = 5.3\,e^{-5}$ right: $\varepsilon$-alg., $\alpha = 1.4$, $K = 2$, $\rho = 0.75$, $d_N = 2.0\,e^{-2}$, $d_{\Gamma 2} = 1.1\,e^{-4}$

For the CM-algorithm, we have instead by (17) that

$$\rho = \int_0^\infty x^2\,\mathrm{d}x \int_A^\infty \mathrm{e}^{-tx}\frac{\delta(t-\beta)^\alpha}{\Gamma(1+\alpha)}\,\mathrm{d}t = \frac{\delta}{\Gamma(1+\alpha)}\int_A^\infty \frac{(t-\beta)^\alpha}{t^3}\,\mathrm{d}t$$
$$= \frac{\delta}{\Gamma(1+\alpha)}\int_B^\infty \frac{y^\alpha}{(\beta+y)^3}\,\mathrm{d}y$$

where $B = A - \beta$, and this equation was solved numerically.

Here $S$ is the skewness of $X(1)$ and $\lambda$ is the Poisson mean in the compound Poisson sum of the simulated "large" jumps, that is,

$$\lambda = \int_\varepsilon^\infty n(x)\,\mathrm{d}x = \int_\varepsilon^\infty \frac{\delta \mathrm{e}^{-\beta x}}{x^{1+\alpha}}\,\mathrm{d}x\,, \quad \lambda = \int_0^A \frac{v(t)}{t}\,\mathrm{d}t = \int_0^B \frac{\delta t^\alpha}{(t+\beta)\Gamma(1+\alpha)}\,\mathrm{d}t$$

in the two cases. The $L_2$ distances in (18) are denoted by $d_N$ for the normal approximation and by $d_{\Gamma_1}$, $d_{\Gamma_2}$ for the two gamma ones. Graphs of $f(x)$ and the $\widehat{f}(x)$ are in Fig. 2 for some selected the parameter combinations in Table 2.

Our interpretation of Fig. 2 is that an $L_2$-distance of $e^{-4}$ or less corresponds to an almost perfect fit, whereas one of order $e^{-3}$ is sufficient for most practical purposes, one of order $e^{-2}$ or more inadequate. With this in mind, we were quite surprised to see how well both algorithms perform already for so large values of $\rho$ as 75% and 50%, or equivalently for so small values of $\lambda$ as those reported in the Tables 1 and 2. One further notes that both algorithms improve as $K$ gets smaller or $\alpha$ larger, which is in agreement with limit theorems given in [3] stating roughly that the distribution of $X(1)$ gets closer to normal in the two cases.

Taking $\lambda$ as measure of computational effort is certainly not unambiguous. On top comes the effort in generating from the r.v.'s $Y_n(\varepsilon)$, $Y_n(A)$ with densities proportional to $n(x)$, $\varepsilon < x < \infty$, resp. $v(t)/t$, $0 < t < A$. However, this issue is largely implementation dependent. We have given one suggestion (based on (15)) for the $\varepsilon$-algorithm in Sect. 4 and give a similar A-R scheme for the CM-algorithm and TS case in the appendix. Both are certainly amenable to improvement. Comparison of the $\varepsilon$- and CM algorithms show that $\lambda$ is slightly higher for the CM algorithm. However, the values of $\lambda$ reported in the tables are quite small and thus $1 + \lambda$ could

**Table 1** $\rho = 75\%$

| $\alpha$ | $K$ | $S$ | $\varepsilon$-algorithm | | | | CM-algorithm | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\lambda$ | $d_N$ | $d_{\Gamma_1}$ | $d_{\Gamma_2}$ | $\lambda$ | $d(F, F_N)$ | $d(F, F_{\Gamma_1})$ | $d(F, F_{\Gamma_2})$ |
| 0.2 | 1/2 | 0.57 | 0.23 | $1.4\,e^{-3}$ | $8.2\,e^{-5}$ | $4.0\,e^{-5}$ | 1.6 | $2.0\,e^{-2}$ | $2.8\,e^{-5}$ | $4.0\,e^{-5}$ |
| | 2 | 1.13 | 0.06 | $1.4\,e^{-2}$ | $1.7\,e^{-4}$ | $2.2\,e^{-4}$ | 0.40 | $1.5\,e^{-2}$ | $2.8\,e^{-4}$ | $1.4\,e^{-4}$ |
| | 8 | 2.27 | 0.01 | $1.6\,e^{-1}$ | $1.3\,e^{-2}$ | $3.7\ e^{-2}$ | 0.10 | $1.6\,e^{-1}$ | $2.6\,e^{-2}$ | $7.1\,e^{-2}$ |
| 0.8 | 1/2 | 0.52 | 0.24 | $8.8\,e^{-4}$ | $2.3\,e^{-4}$ | $4.3\,e^{-5}$ | 1.3 | $1.2\,e^{-2}$ | $4.6\,e^{-4}$ | $5.3\,e^{-5}$ |
| | 2 | 1.04 | 0.06 | $7.6\,e^{-3}$ | $2.6\,e^{-3}$ | $5.3\,e^{-5}$ | 0.33 | $8.2\,e^{-3}$ | $3.5\,e^{-3}$ | $3.3\,e^{-4}$ |
| | 8 | 2.09 | 0.01 | $5.0\,e^{-2}$ | $2.4\,e^{-2}$ | $2.8\,e^{-3}$ | 0.08 | $5.0\,e^{-2}$ | $2.5\,e^{-2}$ | $2.0\,e^{-2}$ |
| 1.4 | 1/2 | 0.43 | 0.30 | $2.9\,e^{-4}$ | – | $2.2\,e^{-5}$ | 1.23 | $3.7\,e^{-4}$ | – | $4.8\,e^{-5}$ |
| | 2 | 0.77 | 0.07 | $2.0\,e^{-3}$ | – | $1.1\,e^{-4}$ | 0.31 | $2.4\,e^{-3}$ | – | $2.2\,e^{-4}$ |
| | 8 | 1.73 | 0.02 | $1.1\,e^{-2}$ | – | $1.0\,e^{-3}$ | 0.08 | $1.1\,e^{-2}$ | – | $3.6\,e^{-3}$ |

**Table 2** $\rho = 50\%$

| $\alpha$ | $K$ | $S$ | $\varepsilon$-algorithm | | | | CM-algorithm | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\lambda$ | $d_N$ | $d_{\Gamma_1}$ | $d_{\Gamma_2}$ | $\lambda$ | $d(F, F_N)$ | $d(F, F_{\Gamma_1})$ | $d(F, F_{\Gamma_2})$ |
| 0.2 | 1/2 | 0.57 | 0.96 | $2.9\,e^{-4}$ | $8.8\,e^{-5}$ | $3.9\,e^{-5}$ | 4.1 | $5.7\,e^{-4}$ | $2.8\,e^{-5}$ | $4.6\,e^{-5}$ |
| | 2 | 1.13 | 0.24 | $6.1\,e^{-3}$ | $4.4\,e^{-4}$ | $1.4\,e^{-4}$ | 1.0 | $7.5\,e^{-3}$ | $1.4\,e^{-4}$ | $1.0\,e^{-4}$ |
| | 8 | 2.27 | 0.060 | $1.3\,e^{-1}$ | $6.7\,e^{-3}$ | $3.5\,e^{-2}$ | 0.26 | $1.3\,e^{-1}$ | $1.9\,e^{-2}$ | $4.4\,e^{-2}$ |
| 0.8 | 1/2 | 0.52 | 1.20 | $1.2\,e^{-4}$ | $2.6\,e^{-5}$ | $4.3\,e^{-5}$ | 4.35 | $2.5\,e^{-4}$ | $8.0\,e^{-5}$ | $3.6\,e^{-5}$ |
| | 2 | 1.04 | 0.30 | $5.1\,e^{-4}$ | $5.1\,e^{-4}$ | $4.7\,e^{-5}$ | 1.09 | $2.9\,e^{-3}$ | $1.1\,e^{-3}$ | $1.2\,e^{-4}$ |
| | 8 | 2.09 | $7.5\,e^{-2}$ | $2.3\,e^{-2}$ | $1.1\,e^{-2}$ | $6.2\,e^{-4}$ | $2.7\,e^{-2}$ | $3.0\,e^{-2}$ | $1.4\,e^{-2}$ | $4.2\,e^{-3}$ |
| 1.4 | 1/2 | 0.43 | 2.42 | $2.0\,e^{-5}$ | – | $4.0\,e^{-5}$ | 7.66 | $2.8\,e^{-5}$ | – | $3.0\,e^{-5}$ |
| | 2 | 0.77 | 0.61 | $1.8\,e^{-4}$ | – | $6.1\,e^{-5}$ | 1.92 | $3.1\,e^{-4}$ | – | $5.8\,e^{-5}$ |
| | 8 | 1.73 | 0.15 | $2.3\,e^{-2}$ | – | $1.3\,e^{-4}$ | 0.48 | $2.9\,e^{-3}$ | – | $3.1\,e^{-4}$ |

be a more fair measure than $\lambda$, taking into account also the generation of the Poisson r.v.'s in addition to the $Y$. This makes the difference even smaller. As for precision, values of order $e-5$ should not be compared as they do not improve by increasing $\rho$, which could presumably be due to the discretization. Once this is said, the $\Gamma_2$ scheme gives most often better precision than the $\Gamma_1$ one, and both improve upon the normal, in some cases even significantly. The $\varepsilon$-algorithm gives slightly more precise estimates for the given $\rho$ than the CM one, but most often not that much. Altogether, which one to prefer may depend on case-dependent issues such as the facility to generate the $Y_n(\varepsilon)$ or $Y_n(A)$.

Concerning the chosen values $1/2, 2, 8$ of the kurtosis $K$, we remark that in financial log-return data $K$ is most often of order 1–3 for daily log-returns series, but higher values occur when calibrating parameters, cf. Table 1 in [3]. Sampling at higher frequencies than daily will also increase $K$, and hence one may expect that larger values of $\lambda$ than the ones in our tables will be needed for good precision.

## 7 Exact Simulation of $X(h)$ and other Methods

In our main examples, it is fairly straightforward to generate a r.v. distributed as $X(h)$ in a NIG process. Indeed, one description of the process is as subordinate to a BM $W$ with drift $\beta$ w.r.t. an inverse Gaussian subordinator $\chi(t)$. In more detail, if $W_1$ is another independent BM with drift $\gamma$ and $\chi(t) = \inf\{s > 0 : W_1(t) > \delta t\}$, then $W(\chi(t)) + \mu t$ is distributed as $X(1)$ in a NIG$(\delta, \alpha, \beta, \mu)$ process with $\alpha = \sqrt{\beta^2 + \gamma^2}$. Here a r.v. distributed as $\chi(t)$ need not be simulated via the relation to $W_1$ but can be directly generated. For $X(h)$, just replace $\delta$ by $\delta h$ and $\mu$ my $\mu h$. These facts are surveyed in, e.g., [4, p. 343] and implemented in, e.g., [25]. A similar but easier exact subordination construction applies to the VG process. Asymptotic subordination algorithms for TS and MX are in [27].

For the spectrally positive TS process with finite variation ($\alpha < 1$), it was noted in [3] that a r.v. distributed as $X(1)$ can be generated by an A-R scheme, using (12) with the $S_\alpha(\sigma, 1, 0)$ r.v. $Z$ as proposal and acceptance probability $\mathrm{e}^{-\beta z}$ when $Z = z$; for the standard algorithm to generate $Z$, see [4, p. 332]. Two-sided processes are of course generated by taking the difference between the positive and negative parts. The simplicity of this scheme should be compared to other approaches in the literature, e.g. [8, 23]. It was also remarked in [3] that the situation is more complicated when $\alpha \geq 1$, since then $X(1)$ is supported by the whole of $\mathbb{R}$ and $\mathrm{e}^{-\beta z}$ is unbounded there. We suggest here an exact scheme based on asymptotic properties of stable densities. The details are in the Appendix but are included more for the sake of completeness than because we think the scheme is more attractive than the simple and efficient $\varepsilon$- and CM-algorithms.

A general comment on the method of discrete skeletons is that it gives little information on the whole path $X(0:T)$ unless one uses a skeleton with a quite small $h$ and thereby a considerable computational effort.

We are not aware of methods for exact simulation of $X(h)$ in the MX process. One could potentially use the explicit form of the density, cf. (9), via A-R, but a difficulty is to find suitable bounds for the complex gamma function.

Another approximate method is based on using a series expansion of the form $X(T) = \sum_1^\infty \{H(\Gamma_n, V_n) - c_n T\}$ where the $\Gamma_n$ are the order epochs of a standard Poisson process, and the $V_n$ independent i.i.d. (possibly multivariate) r.v.'s., see the surveys in [31] and [4] XII.4. In the implementation, ones truncates to $n \leq N$ terms. Since $H(\cdot, v)$ is typically decreasing for fixed $v$, this method is hardly intrinsically different from the $\varepsilon$-algorithm. Calculation of $H$ is not always straightforward. We are not aware of systematic studies of the error term $\sum_{N+1}^\infty \ldots$.

## 8 Maxima, Minima and Other Path Functionals

In Sects. 4–6 and 7, we have concentrated on simulation of $X(T)$ alone, say $T = h$ or $T = 1$ (there is no loss of generality in taking $T = 1$ since $X(T) = X_T(1)$ where $X_T$ is the process obtained by replacing the Lévy measure $\nu$ by $T\nu$). In the financial

context, this covers European options, where $\Psi$ in (1) is a function of $X(T)$ alone. E.g. $\Psi\big(X(0:T) = \mathrm{e}^{-rT}\big[Z(0)\mathrm{e}^{X(T)} - K\big]^+$ for a European call with strike $K$. For many other options, $\Psi$ is, however, more complicated. E.g. for an down-and-in barrier option

$$\Psi\big(X(0:T)\big) \;=\; \mathrm{e}^{-rT}\big[Z(0)\mathrm{e}^{X(T)} - K\big]^+ \mathbb{I}\big(Z(0)\mathrm{e}^{X(t)} \le L \text{ for some } t \le T\big)\,.$$

One therefore needs to know also the minimum $m_T = \inf_{t\le T} X(t)$ of $X(0:T)$, which typically is close to the value at some negative jump. Minima or maxima also come up in the context of queues modeled by Lévy input, where key processes $Y$ such as workload, queue length etc. are obtained by reflecting the input $X$ at 0. This means

$$Y(T) \;=\; \big(Y(0) + X(T)\big) \vee \max_{t\le T}\big(X(T) - X(t)\big)\,.$$

In particular, $Y(T) \stackrel{d}{=} M_T$ where $M_T = \sup_{t\le T} X(t)$ in the case $Y(0) = 0$ of an initially empty queue. If $X$ is simulated as a discrete skeleton with step size $h$, the path of $Y$ is approximated by $Y_h(0) = Y(0)$ and the Lindley recursion

$$Y_h\big((n+1)h\big) \;=\; \big[Y_h(nh) + X\big((n+1)h\big) - X(nh)\big]^+\,,$$

leading to

$$Y_h(Nh) \;=\; \big(Y(0) + X(Nh)\big) \vee \max_{n\le N}\big(X(Nh) - X(nh)\big) \;\stackrel{d}{=}\; \max_{n\le N} X(nh)$$

where the final $\stackrel{d}{=}$ requires $Y(0) = 0$. For these facts, see Sects. III.6–7, IX.2 of [1].

We mention several strategies to access a minimum or maximum, say $m(T)$, without recommending any particular one (in fact, such a choice may depend on the particular application context and a more extensive numerical study). One strategy is just to simulate a sufficiently fine skeleton exactly, when possible, and then take the minimum along the skeleton. This may be supplemented with continuity corrections as developed in [5, 20], that is, r.v.'s approximating

$$\min_{nh\le t\le (n+1)h} X(t) \;\big|\; X_{(n+1)h}, X_{nh}\,.$$

If exact simulation of a skeleton is not feasible, one may instead generate the skeleton approximately by one of the compound Poisson algorithms of Sects. 4, 5, allocate uniform $[0, T]$ locations to the Poisson jump times $\tau_n$ , and supplement the minimum along the $\tau_n$ by invoking bridge r.v.'s of the form

$$\min_{\tau_n\le t\le \tau_{n+1}} \big(\widehat{X}_{0:\varepsilon}(t) - \widehat{X}_{0:\varepsilon}(\tau_n)\big) \;\big|\; \widehat{X}_{0:\varepsilon}(\tau_n)\,.$$

The distribution and hence generation of such bridge minima is standard when $\widehat{X}_{0:\varepsilon}$ is generated by using the normal approximation. For our gamma approximations, they may be efficiently generated by invoking the relation between gamma, beta and Dirichlet distributions, as developed in [7] and implemented in [25].

Similar remarks apply to other and possibly more complicated path functionals. For example, for a Parisian option one needs to know the first time the asset price $e^{X(t)}$ makes an excursion of length $> D$ below some level $L$.

**Acknowledgements** I am grateful to the reviewers for useful comments and corrections, and to Alexey Kuznetsov for useful hints related to Example 1.

## Appendix

### *An A-R Scheme for Generation from $v(t)/t$ in the TS Case*

We need to generate a r.v. $Z$ with density proportional to $u(t)/t$ , $\beta < t < A$ where $u(t) = (t-\beta)^\alpha$. To this end, write $Z = \beta + Z_0$ where $Z_0$ has density proportional to $u(t+\beta)/(t+\beta) = t^\alpha/(t+\beta)$ , $0 < t < B$ where $B = A - \beta$. Here $Y = 1/Z_0$ has density proportional to $1/y^{1+\alpha}/(1+\beta y)$, $1 < 1/B < y < \infty$, and can therefore be generated by A-R with either a Pareto($\alpha$) proposal and acceptance probability proportional to $1/(1+\beta y)$ (high for small $y$) or a Pareto($1+\alpha$) proposal and acceptance probability proportional to $y/(1+\beta y)$ (high for large $y$). As in Sect. 4, we use a mixture, corresponding to breaking the compound Poisson part in 2) above into two. So, let

$$\lambda_1 = \int_{1/B}^{y_0} \frac{1}{y^{1+\alpha}(1+\beta y)}\,\mathrm{d}y\,, \quad \mu_1 = \int_{1/B}^{y_0} \frac{1}{y^{1+\alpha}}\,\mathrm{d}y = \frac{1}{\alpha}[B^\alpha - 1/y_0^\alpha]\,,$$
$$\lambda_2 = \int_{y_0}^{\infty} \frac{1}{y^{1+\alpha}(1+\beta y)}\,\mathrm{d}y\,, \quad \mu_2 = \int_{y_0}^{\infty} \frac{1}{y^{2+\alpha}}\,\mathrm{d}y = \frac{1}{(1+\alpha)y_0^{1+\alpha}}\,,$$
$$C_1 = \frac{\mu_1}{\lambda_1(1+\beta/B)}\,, \quad C_2 = \frac{\mu_2}{\beta\lambda_2}$$

The target densities are then

$$f_1(y) = \frac{1}{\lambda_1 y^{1+\alpha}(1+\beta y)},\ 1/B < y < y_0, \text{ and } f_2(y) = \frac{1}{\lambda_2 y^{1+\alpha}(1+\beta y)},\ y_0 < y < \infty,$$

and chosen with probabilities $\lambda_1/(\lambda_1+\lambda_2)$, resp. $\lambda_2/(\lambda_1+\lambda_2)$, and the proposals are

$$g_1(y) = \frac{1}{\mu_1 y^{1+\alpha}},\ 1/B < y < y_0, \text{ and } g_2(y) = \frac{1}{\mu_2 y^{2+\alpha}},\ y_0 < y < \infty,$$

Then $f_1(y) \le C_1(y_0)g_1(y)$ and $f_2(y) \le C_2(y_0)g_2(y)$, and we may use A-R with acceptance probabilities

$$\frac{f_1(y)}{C_1(y_0)g_1(y)} = \frac{1}{1+\beta y}, \quad \frac{f_2(y)}{C_2(y_0)g_2(y)} = \frac{y}{1+\beta y}$$

for r.v. generation from $f_1$, resp. $f_2$. This gives expected numbers $C_1(y_0)$, $C_2(y_0)$ of samplings from $g_1(y)$, resp. $g_2(y)$, and as measure $E(y_0)$ of the computational effort, we shall use the total number of these samplings, i.e.

$$E(y_0) \;=\; \lambda_1 C_1 + \lambda_2 C_2 \;=\; \frac{\mu_1}{1+\beta/B} + \frac{\mu_2}{\beta} \;=\; \frac{\beta\mu_1 + (1+\beta/B)\mu_2}{\beta(1+\beta/B)} \qquad (20)$$

**Proposition 4** *The function $E(y_0)$, $1/B < y_0 < \infty$, is minimized for $y_0 = y_0^* = 1/\beta + 1/B$.*

***Proof*** In (20), $\beta$ and $B$ as well as term $B^\alpha/\alpha$ in $\mu_1$ do not depend on $y_0$, so we are left with the minimization of $-\beta/\alpha/y_0^\alpha + (1+\beta/B)/(1+\alpha)/y_0^{1+\alpha}$. The derivative is $1/y_0^{1+\alpha}/(1+\beta/B) - 1/y_0^{2+\alpha}$ which changes sign from negative to positive at $y_0^*$. From this the result follows. □

## *An A-R Scheme for Spectrally Positive Infinite Variation TS Processes*

Let $f_0$ be the density of a strictly $\alpha$-stable distribution $S_\alpha(\sigma, 1, 0)$ distribution with $\sigma = \big(-\delta\Gamma(-\alpha)\cos(\pi\alpha/2)\big)^{1/\alpha}$. The goal is to generate a r.v. $X$ from the density $f(x) = \exp\{-\beta x - \psi\} f_0(x)$ in the case $\alpha > 1$ where $f$ and $f_0$ have support on the whole of $\mathbb{R}$; here $\psi = \delta\Gamma(-\alpha)\beta^\alpha$. We use that $f_0(x)$ has the asymptotics [29, p. 100]

$$f_0(x) \;\sim\; \frac{c_1}{|x|^\ell} \exp\big\{-c_2|x|^\eta\big\} \;\text{ as } x \to -\infty \qquad (21)$$

for suitable (explicit) constants) $c_1, c_2$ and $\ell = \alpha/(2\alpha-2)$, $\eta = \alpha/(\alpha-1)$.

For initialization of the algorithm:

(1) Select $-A < 0$ and compute $p = \int_{-A}^{\infty} f(x)\,\mathrm{d}x$

(2) Select $c_3 < c_2$ and find $c_4 < \infty$ such that

$$h(x) \;=\; \frac{\mathrm{e}^{\beta|x|} f_0(x)}{|x|^{\eta-1}\exp\big\{-c_3|x|^\eta\big\}} \;\le\; c_4 \quad \text{for all } x < -A.$$

The algorithm is then as follows:

(3) Generate $I$ as Bernoulli($p$).

(4) If $I = 1$, generate $X \in (-A, \infty)$ having density $f(x)/p$, $-A < x < \infty$, by A-R with proposal $Z_0$ having a strictly $\alpha$-stable distribution $S_\alpha(\sigma, 1, 0)$ conditioned to $(-A, \infty)$ and acceptance probability $\mathrm{e}^{-b(z+A)}$ when $Z_0 = z$
(5) If $I = 0$, generate $X \in (-\infty, -A)$ with density $\widetilde{f}(x) = \mathrm{e}^{b|x|-\psi} f_0(x)/(1-p)$, $= \infty < x < -A$, by an A-R scheme defined as follows. As proposal, take a r.v. $Z_1$ distributed as $-Z_2$ given $Z_2 > A$ where $Z_2 > 0$ is Weibull with $\mathbb{P}(Z_2 > z) = \mathrm{e}^{-c_3 z^\eta}$. If $Z_1 = x$, accept w.p. $c_4 h(x)$.
(8) return $X$.
*Explanation:* Step (2) is possibly because (21), $c_2 > c_3$ and $\eta > 1$ imply $h(x) \to 0$ as $x \to -\infty$. In (5), the proposal density is $g(x) = \mathbb{P}\big(X_2 \in \mathrm{d}|x| \,\big|\, Z_2 > A\big) = c_3\eta|x|^{\eta-1}\mathrm{e}^{-c_3 z^\eta}/\mathrm{e}^{-c_3 A^\eta}$. Thus the ratio of the target density to the proposal density is

$$\frac{\widetilde{f}(x)}{g(x)} \;=\; c_5 h(x) \quad \text{where} \quad c_5 = \frac{\exp\{-\psi - c_3 A^\eta\}}{c_3\eta(1-p)}$$

Hence $\widetilde{f}(x)/g(x) \le c_0 h(x)$ where $c_0 = c_4 c_5$, and acceptance w.p. $\widetilde{f}(x)/c_0/g(x) = c_4 h(x)$ will produce the correct result. The conditioned sampling of proposals in (6) and (7) is straightforward by sampled by sampling a $S_\alpha(\sigma, 1, 0)$, resp. Weibull, r.v. until the conditioning requirement is met. Available software, say MATLAB or Nolan's STABLE package (see the Preface to [29]) accounts for computing $f_0(x)$ and generating the $S_\alpha(\sigma, 1, 0)$ r.v.'s. The Weibulls can be generated by inversion.

## References


1. Asmussen, S.: Applied Probability and Queues, 2nd edn. Springer, New York (2003)
2. Asmussen, S.: Conditional Monte Carlo for sums, with applications to insurance and finance. Ann Actuarial Sci **12**, 455–478 (2018)
3. Asmussen, S.: On the role of skewness and kurtosis in tempered stable (CGMY) Lévy models in finance. Finance and Stochastics **26**, 383–416 (2022)
4. Asmussen, S., Glynn, P.: Stochastic Simulation, Algorithms and Analysis. Springer, Berlin (2007)
5. Asmussen, S., Ivanovs, J.: Discretization error for a two-sided reflected Lévy process. Queueing Syst. **89**, 199–212 (2018)
6. Asmussen, S., Rosiński, J.: Approximations for small jumps of Lévy processes with a view towards simulation. J. Appl. Probab. **38**, 482–493 (2001)
7. Avramidis, A., L'Ecuyer, P., Tremblay, P.A.: Efficient simulation of Gamma and variance Gamma processes. In: Proceedings of the 2003 Winter Simulation Conference, pp. 319–326 (2003)
8. Ballotta, L., Kyriakou, I.: Monte Carlo simulation of the CGMY process and option pricing. J. Futur. Mark. **4**, 1095–1121 (2014)
9. Barndorff-Nielsen, O.: Processes of normal inverse Gaussian type. Financ. Stochast. **4**, 1095–1121 (1998)
10. Bertoin, J.: Lévy Processes. Cambridge University Press (1996)
11. Bondesson, L.: On simulation from infinitely divisible distributions. Adv. Appl. Probab. **14**, 855–869 (1982)
12. Carr, P., Geman, H., Madan, D., Yor, M.: The fine structure of asset returns: an empirical investigation. J. Bus. **75**, 305–332 (2002)

13. Carr, P., Madan, D.: Option valuation using the fast Fourier transform. J. Comput. Financ. **2**, 61–73 (1998)
14. Cont, R., Tankov, P.: Financial Modelling with Jump Processes. Chapman and Hall/CRC (2004)
15. Debicki, K., Mandjes, M.: Queues and Lévy Fluctuation Theory. Springer (2018)
16. Dia, E.: Error bounds for small jumps of Lévy processes. Adv. Appl. Probab. **45**, 86–105 (2013)
17. Feller, W.: An Introduction to Probability Theory and Its Applications, vol. 2, 2 edn. Wiley (1971)
18. Grigelionis, B.: Generalized z-distributions and related stochastic processes. Lietuvos Matematikos Rinkinys **41**, 239–251 (2001)
19. Hackmann, D., Kuznetsov, A.: Approximating Lévy processes with completely monotone jumps. Ann. Appl. Probab. **26**, 328–359 (2016)
20. Ivanovs, J.: Zooming in on a Lévy process at its supremum. Ann. Appl. Probab. **28**, 912–940 (2018)
21. Jeannin, M., Pistorius, M.: A transform approach to compute prices and Greeks of barrier options driven by a class of Lévy processes. Quant. Financ. **10**, 629–644 (2010)
22. Kallenberg, O.: Foundations of Modern Probability, 2nd ed. edn. Springer (2003)
23. Karlsson, P.: Finite element based Monte Carlo simulation of options on Lévy driven assets. Int. J. Financ. Eng. **5**, 1850013 (2018)
24. Küchler, U., Tappe, S.: Tempered stable distributions and processes. Stoch. Process. Their Appl. **123**, 4256–4293 (2013)
25. L'Ecuyer, P.: `https://github.com/umontreal-simul/ssj` (2015). Software package, Université de Montreal
26. L'Ecuyer, P., Puchhammer, F., Ben Abdellah, A.: Monte Carlo and quasi-Monte Carlo density estimation via conditioning. Inf. J. Comput. (2022). https://doi.org/10.1287/ijoc.2021.1135
27. Madan, D., Yor, M.: "CGMY and Meixner subordinators are absolutely continuous with respect to one sided stable subordinators". HAL Archives-Ouvert.fr hal-00016662v2 (2006)
28. Mozzola, M., Muliere, P.: Reviewing alternative characterizations of Meixner process. Probab. Surv. **8**, 127–154 (2011)
29. Nolan, J.: Univariate Stable Distributions Models for Heavy Tailed Data. Springer, Berlin (2020)
30. Poirot, J., Tankov, P.: Monte Carlo option pricing for tempered stable (CGMY) processes. Asia-Pac. Financ. Mark. **13**, 327–344 (2006)
31. Rosiński, J.: Series representations of Lévy processes from the perspective of point processes. In: Barndorff-Nielsen, O., Mikosch, T., Resnick, S. (eds.) Lévy Processes — Theory and Applications, pp. 401–415. Birkhäuser (2001)
32. Rydberg, T.: The normal inverse Gaussian Lévy process: simulation and approximations. Stoch. Model. **13**, 887–910 (1997)
33. Sato, K.: Lévy Processes and Infinitely Divisible Distributions. Cambridge University Press (1999)
34. Schilling, R., Song, R., Vondraček, Z.: Bernstein Functions. Theory and Applications, 2nd edn. de Gruyter (2012)
35. Schoutens, W.: Lévy Processes in Finance Pricing Financial Derivatives. Wiley (2003)
36. Seneta, E.: A Tauberian theorem of E. Landau and W. Feller. Annals of Probability **6**, 1057–1058 (1973)
37. Signahl, M.: On error rates to normal approximations and simulation schemes for Lévy processes. Stoch. Model. **19**, 287–298 (2003)