

ON EXACT SAMPLING OF NONNEGATIVE INFINITELY DIVISIBLE RANDOM VARIABLES

ZHIYI CHI,* *University of Connecticut*

Abstract

Nonnegative infinitely divisible (i.d.) random variables form an important class of random variables. However, when this type of random variable is specified via Lévy densities that have infinite integrals on $(0, \infty)$, except for some special cases, exact sampling is unknown. We present a method that can sample a rather wide range of such i.d. random variables. A basic result is that, for any nonnegative i.d. random variable X with its Lévy density explicitly specified, if its distribution *conditional* on $X \leq r$ can be sampled exactly, where $r > 0$ is any fixed number, then X can be sampled exactly using rejection sampling, without knowing the explicit expression of the density of X . We show that variations of the result can be used to sample various nonnegative i.d. random variables.

Keywords: Infinitely divisible; rejection sampling; exact sampling; subordinator; Lévy process

2010 Mathematics Subject Classification: Primary 60E07

Secondary 60G50

1. Introduction

Nonnegative infinitely divisible (i.d.) random variables are important in many regards; however, their sampling in general is an involved issue. By the Lévy–Khintchine formula, a nonnegative i.d. random variable can be represented as a series of jump times of a Poisson process [4], [33]. On the one hand, the series representation is the basis of many exact or approximate sampling methods [3], [8], [11], [32]. On the other hand, if an i.d. random variable has an infinite Lévy measure on $(0, \infty)$ then, with probability 1, its series representation has infinitely many positive random terms, which cannot be summed in closed form but have to be added one by one. This rules out exact sampling of the random variable in a finite number of steps.

Conceptually, it is important and interesting to see which nonnegative i.d. random variables can be sampled exactly. There are actually many such random variables, the most familiar being gamma, Pareto, Fisher, F , Gumbel, Weibull, log-normal, and half-Cauchy variables [33]–[35]. It is known that positive random variables which have log-convex densities or are mixtures of gamma random variables with the same shape parameter in $(0, 2)$ are i.d. [6], [33], [35]. Since these random variables have analytically tractable density functions, they can be sampled exactly. A special class of i.d. random variables, known as Vervaat perpetuities, can also be sampled exactly by carefully exploiting their structures [15]–[17], [22]. It is also well known that nonnegative stable random variables and their exponentially tilted versions can be sampled exactly [8], [10], [13], [20], [21], [32]. These random variables not only have explicit density functions, but explicit Lévy measures as well. Generally speaking, in view of the

Received 10 May 2011; revision received 19 March 2012.

* Postal address: Department of Statistics, University of Connecticut, 215 Glenbrook Road, U-4120, Storrs, CT 06269, USA. Email address: zhiyi.chi@uconn.edu

Lévy–Khinchine formula, it is natural to specify an i.d. random variable via its Lévy measure; however, other than the aforementioned cases and perhaps a few others, for variables specified in this way, exact sampling is lacking.

In this paper we show that sampling is achievable for a rather wide range of nonnegative i.d. random variables specified via Lévy measures. Henceforth, by sampling we mean exact sampling, and we shall speak interchangeably of the sampling *for* a Lévy density, the sampling *from* the i.d. distribution with the Lévy density, and the sampling *of* an i.d. random variable with the Lévy density. To start with, suppose that we wish to sample a nonnegative i.d. random variable X with an upper truncated Lévy density $\lambda(x)\mathbf{1}\{x \leq r\}$, where $r > 0$ is fixed and λ itself is a Lévy density on $(0, \infty)$. Many functional relationships between the distribution of X and that associated with λ are known, which can be used to evaluate the (probability) density of X [11]. However, it is practically (and conceptually) more satisfactory if sampling can be done without the numerical evaluation of density functions [10]. With this in mind, our goal is to sample X without knowing its density explicitly.

Our approach starts with the following simple observation: if $Y \geq 0$ is an i.d. random variable with Lévy density λ then X *conditional* on $X \leq r$ and Y *conditional* on $Y \leq r$ are identically distributed. To be sure, represent Y as $X + Z$, where Z consists of jump times greater than r . Because Z is compound Poisson and independent of X , $P\{Z = 0\} > 0$ and $P\{X \leq x \mid X \leq r\} = P\{X \leq x, Z = 0 \mid X \leq r, Z = 0\}$. Since, for any $x \leq r$, the events $\{X \leq x, Z = 0\}$ and $\{Y \leq x\}$ are the same, we obtain $P\{X \leq x \mid X \leq r\} = P\{Y \leq x \mid Y \leq r\}$. Consequently, if we can choose λ appropriately so that we know how to sample Y , then we can sample X conditional on $X \leq r$ by sampling Y conditional on $Y \leq r$, using, say, the rejection method (cf. [13], [18], [19], and [27]).

The above observation is crucial; however, it cannot singularly lead to the sampling of X . There are two issues. First, $P\{X \leq r\}$ is often unknown, and, hence, the density of X on $(0, r]$ is known only up to a multiplicative factor. The second issue appears to be more serious. That is, there is little direct knowledge about the density of X on (r, ∞) . Since X consists of only jump times less than r , we have little use of the density of Y on (r, ∞) , as it involves large jump times nonexistent in X . Instead, we need to rely on certain relations between the density of X on $(0, r]$ and its density on (r, ∞) that are applicable for sampling.

To address the issues, in Section 3 we establish an integral series expansion of the density of X on (r, ∞) in terms of its density on $(0, r]$ and Lévy density. While the expansion contains infinitely many integrals, it allows exact sampling. In Section 4, a general procedure that applies rejection sampling to the integral series expansion is presented. Together, the two sections deliver the following method to sample a nonnegative i.d. random variable. First, decompose its Lévy density into $\varphi + \chi$ such that, on some $(0, r] \neq \emptyset$, φ is identical to a Lévy density λ which we know how to sample for, and $\chi \geq 0$ is integrable. Indeed, for φ , we need only $\varphi(t) = [1 + O(t)]\lambda(t)$ as $t \rightarrow 0$. Once such a decomposition is found, we can sample for φ using the procedure in Section 4. Meanwhile, since the i.d. distribution with Lévy density χ is compound Poisson, its sampling is more or less standard, although efficient algorithms can be found for special cases [13], [18]. The sum of the values independently sampled from the two i.d. distributions then follows the i.d. distribution we wish to sample.

We note that rejection sampling has long been used to sample i.d. random variables (cf. [12], [13], and [15]). However, unlike in the procedures of the above-cited works, we do not rely on explicit approximations to distributions. More precisely, if sampling for a Lévy density is available then it can be incorporated into our procedure to sample for a whole class of Lévy

densities without the need to derive new formulae for the densities. Therefore, our procedure is complementary to available sampling procedures, such as those in [10], [13], and [18].

In Sections 5 and 6 we consider two general designs for our procedure and illustrate them with examples. The first example deals with the Lévy density $\varphi(t) = ct^{-\alpha-1}\mathbf{1}\{0 < t \leq r\}$ with $\alpha \in (0, 1)$ and $r \in (0, \infty)$. While the sampling for the stable Lévy density $ct^{-\alpha-1}\mathbf{1}\{t > 0\}$ is well known [10], [13], [21], [32], to the best of the author’s knowledge, the sampling for φ is unknown. We will utilize the sampling of the stable distribution to sample for φ . As a more concrete example, we next consider the Lévy density $\varphi(t) = \psi(t)(e^t - 1)^{-1-\alpha}\mathbf{1}\{t > 0\}$ with $\alpha \in (0, 1)$ and $\psi(t) = 1 + O(t)$ as $t \rightarrow 0$. When $\psi(t) = e^{\beta t}$ with $\beta < \alpha + 1$, φ gives rise to a Lamperti-stable distribution [9], [25]. In the third example, we apply the procedure to gamma distributions. Since efficient sampling of gamma distributions has long been known (cf. [13] and [18]), the point here is that the procedure can be exploited to get some interesting theoretical results. The above examples are used to illustrate the first design. For the second design, we give two examples. In the first example, we consider the sampling of Vervaat perpetuities. Recently, efficient sampling procedures for Vervaat perpetuities have been discovered [16], [17], [22]. These procedures employ sophisticated coupling techniques for Markov chains and, depending on the parameter value of the distribution being sampled, their expected numbers of iterations range from 1 to ∞ . We show that, for any parameter value of the distribution, by using the second design, the expected number of iterations can be arbitrarily close to 1. In the second example, we consider the Lévy density $ce^{-t}(1 - t^a)\mathbf{1}\{0 < t \leq r\}/[t \ln(1/t)]$. We show that its sampling can be done by incorporating the sampling of Vervaat perpetuities and the trick of subordination. In all but the theoretical example, we provide some analysis on the complexity of the procedure.

Finally, for the integral series expansion in Section 3, local boundedness of the probability density is a required condition. Section 7 gives a simple criterion to check the condition.

2. Preliminaries

2.1. Notation

Henceforth, by Lévy densities we mean those of nonnegative i.d. random variables, which coincide with measurable functions $\varphi \geq 0$ with support in $[0, \infty)$ and $\int(1 \wedge t)\varphi(t) dt < \infty$ (cf. [4] and [33]). By i.d. random variable with Lévy density φ , we mean specifically a random variable X with Laplace transform $E(e^{-\theta X}) = \int(e^{-\theta t} - 1)\varphi(t) dt$, $\theta > 0$. We denote by $ID(\varphi)$ the distribution with the Laplace transform. We call X strictly positive if $P\{X > 0\} = 1$. Under this setup, $X \sim ID(\varphi)$ is strictly positive if and only if $\int \varphi = \infty$ [11], [30], and in this case, X has a density with support being the entire $[0, \infty)$ [33, Theorems 24.10 and 27.10]. On the other hand, if $\int \varphi < \infty$ then $X \sim ID(\varphi)$ is compound Poisson, with $P\{X = 0\} > 0$.

For brevity and when there is no possibility of confusion, a probability density will be referred to as a density. Denote by $Unif(0, 1)$ the uniform distribution on $(0, 1)$, and, for $r > 0$, $a > 0$, and $b > 0$, denote by $Gamma(a, r)$ the distribution with density $x^{a-1}e^{-x/r}\mathbf{1}\{x > 0\}/[r^a\Gamma(a)]$ and by $Beta(a, b)$ the distribution with density $x^{a-1}(1 - x)^{b-1}\mathbf{1}\{0 < x < 1\}/B(a, b)$, where $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a + b)$. Also, denote $Gamma(1, r)$ by $Exp(r)$. For $p \in (0, 1)$, denote by $Binomial(n, p)$ the law of the sum of n independent and identically distributed (i.i.d.) ξ_i with $P\{\xi_i = 0\} = 1 - P\{\xi_i = 1\} = 1 - p$.

2.2. Decomposition and exponential tilting

Sampling for a Lévy density can often be built upon the sampling for other Lévy densities. Suppose that we know how to sample for a Lévy density φ . If a Lévy density φ_1 can be

decomposed into $\varphi + \chi$, with $\chi \geq 0$ being integrable, then we can sample for φ_1 by generating $X + \xi$, where $X \sim \text{ID}(\varphi)$ and $\xi \sim \text{ID}(\chi)$ are independent. A standard method to sample for χ is as follows. Sample a Poisson process with intensity χ , which has a finite number of points with probability 1. The sum of the coordinates of the points then follows $\text{ID}(\chi)$ (cf. [13] and [18]).

On the other hand, if $\varphi_1 \leq \varphi$ then even if $\chi = \varphi - \varphi_1$ is integrable, in general, it is unclear how to sample for φ_1 based on φ . However, suppose that, for some $b > 0$, $\varphi_1(t) = e^{-bt}\varphi(t)$, i.e. φ_1 is an exponentially tilted version of φ . It is a standard result that exponential tilting of a Lévy density induces the same exponential tilting of the corresponding i.d. distribution (up to a normalizing constant; cf. [1], [4], [8], and [20]). We state the result for later use.

Lemma 2.1. *Let φ and φ_1 be two Lévy densities such that, for some $b > 0$, $\varphi_1(t) = e^{-bt}\varphi(t)$. Then, for $X_1 \sim \text{ID}(\varphi_1)$ and $X \sim \text{ID}(\varphi)$, $\mathbb{P}\{X_1 \in dx\} = e^{-bx} \mathbb{P}\{X \in dx\} / \mathbb{E}(e^{-bX})$.*

The sampling of exponentially tilted distributions is well known (cf. [18]). More generally, suppose that, for some $b \geq 0$, $e^{-bt}\varphi(t) \leq \varphi_1(t) \leq \varphi(t)$. Let $\chi(t) = \varphi_1(t) - e^{-bt}\varphi(t)$. We have $\chi \geq 0$ and, by $\chi(t) \leq (1 - e^{-bt})\varphi(t) \leq [(bt) \wedge 1]\varphi(t)$, $\chi \in L^1(0, \infty)$. Then $\text{ID}(\varphi_1)$ can be sampled as follows.

1. Keep sampling (U, X) until $U \leq e^{-bX}$, where $U \sim \text{Unif}(0, 1)$ and $X \sim \text{ID}(\varphi)$ are independent.
2. Sample $\xi \sim \text{ID}(\chi)$. Return $X + \xi$.

3. An integral series expansion of the density

In this section, let φ be a Lévy density with $\int \varphi = \infty$. From Section 2, we know that $\text{ID}(\varphi)$ has a density that has the entire $[0, \infty)$ as support. We shall derive an integral series expansion that expresses the density of $\text{ID}(\varphi)$ on (r, ∞) for a given $r > 0$ in terms of its density on $(0, r]$ and Lévy density. Denote by g the density of $\text{ID}(\varphi)$, and let $\varrho(t) = t\varphi(t)\mathbf{1}\{t > 0\}$.

3.1. Main result

Recall that a function is said to be locally bounded on $A \subset \mathbb{R}$ if it is uniformly bounded on any compact subset of A .

Theorem 3.1. *Fix $r > 0$. Suppose that, for some $M > r$, g is locally bounded on $(0, M)$. Define*

$$h(v, x) = \frac{\mathbf{1}\{x > 0\}\varrho(x - v)}{x}$$

and, for each $k \geq 1$,

$$h_k(v, x) = \int \mathbf{1}\{v_1 > r\}h(v, v_1)h(v_1, v_2) \cdots h(v_k, x) dv_1 \cdots dv_k.$$

Then, for $x \in (r, M)$,

$$g(x) = \int_0^r g(v) dv \left[h(v, x) + \sum_{k=1}^{\infty} h_k(v, x) \right]. \tag{3.1}$$

Remark. From $h(v, x) = 0$ for $x < v$, it follows that $h_k(v, x) = 0$ for $k \geq 1$ and $x \leq r$.

To prove Theorem 3.1, we need the following known result (see [33, Theorem 51.1] and [34, Corollary 4.2.2]).

Lemma 3.1. For any $x > 0$, $g(x) = \int_0^x g(v)h(v, x) dv$.

Proof of Theorem 3.1. Equipped with Lemma 3.1, we first show that, given $x \in (r, M)$, for each $n \geq 1$,

$$g(x) = \int_0^r g(v) dv \left[h(v, x) + \sum_{k=1}^n h_k(v, x) \right] + R_n(x), \tag{3.2}$$

with

$$R_n(x) = \int \mathbf{1}\{v > r\} g(v)h(v, v_1)h(v_1, v_2) \cdots h(v_n, x) dv dv_1 \cdots dv_n.$$

By Lemma 3.1 and noting that $h(v, x) = 0$ for $v > x$,

$$g(x) = \int_0^\infty g(v)h(v, x) dv = \int_0^r g(v)h(v, x) dv + \int_r^\infty g(v)h(v, x) dv. \tag{3.3}$$

For $g(v)$ in the last integral, as $v > r$, expansion (3.3) can be iterated once to yield

$$g(x) = \int_0^r g(v)h(v, x) dv + \int_r^\infty \left[\int_0^r g(w)h(w, v) dw + \int_r^\infty g(w)h(w, v) dw \right] h(v, x) dv.$$

Making some changes in the variables and regrouping the integrals yields

$$\begin{aligned} g(x) &= \int_0^r g(v) dv \left[h(v, x) + \int_r^\infty h(v, v_1)h(v_1, x) dv_1 \right] \\ &\quad + \int_r^\infty \int_r^\infty g(v)h(v, v_1)h(v_1, x) dv dv_1 \\ &= \int_0^r g(v) dv [h(v, x) + h_1(v, x)] + R_1(x), \end{aligned}$$

which shows (3.2) for $n = 1$. In general, for $g(v)$ in the integral expression of $R_n(x)$, as $v \geq r$, we can iterate expansion (3.3) to obtain

$$\begin{aligned} R_n(x) &= \int \mathbf{1}\{v > r\} \left[\int_0^r g(w)h(w, v) dw + \int_r^\infty g(w)h(w, v) dw \right] \\ &\quad \times h(v, v_1)h(v_1, v_2) \cdots h(v_n, x) dv dv_1 \cdots dv_n \\ &= \int_0^r g(w) dw \left[\int \mathbf{1}\{v > r\} h(w, v)h(v, v_1)h(v_1, v_2) \cdots h(v_n, x) dv dv_1 \cdots dv_n \right] \\ &\quad + \int g(w)\mathbf{1}\{w > r\}\mathbf{1}\{v > r\}h(w, v)h(v, v_1) \cdots h(v_n, x) dw dv dv_1 \cdots dv_n. \end{aligned}$$

The right-hand side is exactly $\int_0^r g(w)h_{n+1}(w, x) dw + R_{n+1}(x)$. Then, by induction, (3.2) holds.

By the expression for $h(v, x)$,

$$R_n(x) \leq \sup_{v \in [r, x]} g(v) \times I_n(x), \tag{3.4}$$

where, writing $v_0 = v$ and $v_{n+1} = x$,

$$I_n(x) = \int \mathbf{1}\{r < v_0 < v_1 < \dots < v_n < x\} \prod_{i=0}^n \frac{\varrho(v_{i+1} - v_i)}{v_{i+1}} dv_i.$$

Let $s_i = v_i - v_{i-1}$ for $i = 1, \dots, n + 1$. Then $v_i = x - s_{i+1} - \dots - s_{n+1}$, $i = 0, \dots, n$, and $\mathbf{1}\{r < v_0 < v_1 < \dots < v_n < x\} = \mathbf{1}\{\text{all } s_i > 0, s_1 + \dots + s_{n+1} < x - r\}$. As a result,

$$I_n(x) \leq r^{-n-1} \int \mathbf{1}\{s_1 + \dots + s_{n+1} < x - r\} \prod_{i=1}^{n+1} \varrho(s_i) ds_i.$$

Letting $\mu = \int_0^{x-r} \varrho$, $\psi(s) = \varrho(s)\mathbf{1}\{0 < s < x - r\}/\mu$ is a probability density. Let ξ, ξ_1, ξ_2, \dots be i.i.d. $\sim \psi$. Then, by the above inequality and Markov's inequality, for any $t > 0$,

$$\begin{aligned} I_n(x) &\leq \left(\frac{\mu}{r}\right)^{n+1} \text{P}\{\xi_1 + \dots + \xi_{n+1} < x - r\} \\ &\leq \left(\frac{\mu}{r}\right)^{n+1} \text{E}[\exp\{t(x - r - \xi_1 - \dots - \xi_{n+1})\}] \\ &= e^{t(x-r)} \left(\frac{\mu \text{E}(e^{-t\xi})}{r}\right)^{n+1}. \end{aligned}$$

Since $\text{E} e^{-t\xi} \rightarrow 0$ as $t \rightarrow \infty$, we can find $t > 0$ such that $\mu \text{E}(e^{-t\xi})/r < 1$. Fixing such t and letting $n \rightarrow \infty$, we obtain $I_n(x) \rightarrow 0$. By assumption, $\sup_{v \in [r,x]} g(v) < \infty$, so, by (3.4), $R_n(x) \rightarrow 0$. This, together with (3.2) and monotone convergence, yields (3.1).

3.2. Formulation toward rejection sampling

Let g be locally bounded on $(0, \infty)$. By Theorem 3.1, for all $x > 0$,

$$\begin{aligned} g(x) &= \mathbf{1}\{x \leq r\}g(x) + \mathbf{1}\{x > r\} \int_0^r g(v) dv \left[h(v, x) + \sum_{k=1}^{\infty} h_k(v, x) \right] \\ &= \int_0^r g(v) dv \left[\delta(x - v) + h(v, x)\mathbf{1}\{x > r\} + \sum_{k=1}^{\infty} h_k(v, x) \right], \end{aligned}$$

where δ is the delta function and the second equality is due to the fact that $h_k(v, x) = 0$ whenever $x \leq r$. Suppose that we know how to sample X conditional on $X \leq r$, i.e. from the density

$$p(x) = a^{-1}g(x)\mathbf{1}\{x \leq r\}, \tag{3.5}$$

where $a = \int_0^r g$. The value of a is often unavailable. Since the support of g is $[0, \infty)$, the conditional density is well defined. Using the above formula, we can rewrite g as

$$g(x) = a \int p(v) dv \left[\delta(x - v) + h(v, x)\mathbf{1}\{x > r\} + \sum_{k=1}^{\infty} h_k(v, x) \right]. \tag{3.6}$$

The question is how to use (3.6) to sample from g . In the expansion, δ and h should be easy to handle; however, h_1, h_2, \dots are defined by integrals. In general, for sampling, it is desirable

to avoid evaluating integrals. With this in mind, we shall first consider densities of the form

$$g(x) = a \int p(v) dv \left[\sum_i \int q_i(v, x, w_i) v_i(dw_i) \right] \text{ for some constant } a > 0, \tag{3.7}$$

where the sum on the right-hand side has at most countably many terms, the $q_i \geq 0$ are known functions, and the v_i are σ -finite measures on some measurable spaces that may be different from each other. Besides including (3.6) as a special case, the expansion in (3.7) provides more flexibility for sampling. For example, if the integral that defines h_k in (3.6) is complicated then we may consider using (3.7) to reformulate h_k as the sum of several integrals over disjoint regions that are easier to handle. This perspective will be incorporated into the following discussion.

4. Rejection sampling

4.1. Procedures for general densities

Rejection sampling is an exact sampling method [13], [18], [19], [27]. Let π be a density with respect to a σ -finite measure ν . Suppose that π is specified as $\pi \propto q$, where $q \geq 0$ is a known function. In standard rejection sampling, an ‘instrumental’ density function ϕ with respect to ν and a constant $C > 0$ have to be identified, so that ϕ can be sampled and $q(x) \leq C\phi(x)$ for all x . Then the sampling for π proceeds as follows.

- Keep sampling $X \sim \phi$ and $U \sim \text{Unif}(0, 1)$ independently until $CU\phi(X) \leq q(X)$. Then output X .

We now turn to the density in (3.7). Suppose that we can find constants $C_i \geq 0$ and instrumental functions $\phi_i(\cdot, \cdot, \cdot) \geq 0$ such that, for all v with $p(v) > 0$ and i ,

$$C := \sum_i C_i < \infty, \quad q_i(v, \cdot, \cdot) \leq C_i \phi_i(v, \cdot, \cdot),$$

and

$$\phi_i(v, \cdot, \cdot) \text{ is a probability density with respect to } \ell \times \nu_i,$$

where ℓ is the Lebesgue measure on \mathbb{R} . Under this setup, consider the procedure in the following algorithm.

Algorithm 4.1. (*Rejection sampling for densities of the form (3.7).*)

1. Sample $Z \sim p$ and κ independently, with $P\{\kappa = i\} = C_i/C$.
2. Given Z and κ , sample (X, W) from the density $\phi_\kappa(Z, \cdot, \cdot)$ with respect to $\ell \times \nu_\kappa$.
3. Sample $U \sim \text{Unif}(0, 1)$. If $C_\kappa U \phi_\kappa(Z, X, W) \leq q_\kappa(Z, X, W)$ then output X and stop; otherwise, go back to step 1 and start a new, independent iteration.

Proposition 4.1. *Algorithm 4.1 eventually terminates with probability 1, and its random output follows g in (3.7).*

Proof. First, by integrating (3.7) over x and Fubini's theorem,

$$\begin{aligned}
 1 &= \int g(x) \, dx \\
 &= a \int p(v) \, dv \left[\sum_i \int q_i(v, x, w_i) \, dx v_i(dw_i) \right] \\
 &\leq a \int p(v) \, dv \left[\sum_i C_i \int \phi_i(v, x, w_i) \, dx v_i(dw_i) \right] \\
 &= aC.
 \end{aligned}$$

Therefore, $C > 0$ and the sampling of κ in step 1 is well defined.

Given a measurable set $A \subset \mathbb{R}$, for each iteration,

$$\begin{aligned}
 &P\{X \in A \text{ is output}\} \\
 &= \int \sum_i P\{Z \in dv, \kappa = i\} P\left\{X \in A, U \leq \frac{q_i(v, X, W)}{C_i \phi_i(v, X, W)} \mid Z = v, \kappa = i\right\} \\
 &= \int p(v) \, dv \sum_i \frac{C_i}{C} \left[\int \phi_i(v, x, w_i) \mathbf{1}\{x \in A\} \, dx v_i(dw_i) P\left\{U \leq \frac{q_i(v, x, w_i)}{C_i \phi_i(v, x, w_i)}\right\} \right].
 \end{aligned}$$

Because $q_i(v, x, w_i) \leq C_i \phi_i(v, x, w_i)$, whether or not $C_i \phi_i(v, x, w_i)$ is positive,

$$C_i \phi_i(v, x, w_i) P\left\{U \leq \frac{q_i(v, x, w_i)}{C_i \phi_i(v, x, w_i)}\right\} = q_i(v, x, w_i).$$

As a result,

$$\begin{aligned}
 P\{X \in A \text{ is output}\} &= \frac{1}{C} \int p(v) \, dv \left[\sum_i \int q_i(v, x, w_i) \mathbf{1}\{x \in A\} \, dx v_i(dw_i) \right] \\
 &= \frac{1}{C} \int_A dx \left\{ \int p(v) \, dv \left[\sum_i \int q_i(v, x, w_i) v_i(dw_i) \right] \right\} \\
 &= \frac{1}{aC} \int_A g,
 \end{aligned}$$

where the last equality is due to (3.7). In particular, letting $A = \mathbb{R}$, for each iteration, the probability that X is output (and, hence, the procedure stops) is $1/(aC) > 0$. This shows that, on the one hand, in each iteration, $P\{X \in A \mid X \text{ is output}\} = \int_A g$ and, on the other hand, with probability 1, the procedure eventually terminates. Because the iterations are independent, $P\{X \in A \text{ at termination}\} = \int_A g$. Since this is true for all measurable A , we obtain $X \sim g$.

4.2. A procedure for positive i.d. random variables

Let φ be a Lévy density with $\int \varphi = \infty$ such that the density g of $ID(\varphi)$ is locally bounded on $(0, \infty)$. Our goal is to sample for φ . We have in (3.6) an integral series expansion of g in terms of $\varrho(t) = t\varphi(t)$ and its conditional version on some $(0, r] \neq \emptyset$, i.e. $p(x) = g(x)\mathbf{1}\{x \leq r\}/a$, with $a = \int_0^r g$. To apply Proposition 4.1 to the integral series expansion, our working hypothesis in this section is: for a given $r > 0$, p can be sampled exactly.

Observe that in the integral series expansion of g , given v , both $h(v, x)$ and $h_k(v, x)$ are defined in terms of the increments $x - v, v_i - v_{i-1}$. This suggests that we may sample X by sampling the increments. As noted after (3.7), the domains of the increments may be partitioned into subregions so that each can be treated conveniently in a certain way.

For $k \geq 1$, let ℓ^k denote the Lebesgue measure on \mathbb{R}^k . Suppose that, for $k \geq 1$, we can find constants $C_{k1}, \dots, C_{kn_k} \geq 0$ and measurable instrumental functions $\phi_{kj}(\cdot, \cdot)$ on $\mathbb{R} \times \mathbb{R}^k$, $j = 1, \dots, n_k$, such that the following conditions are satisfied. First,

$$\sum_{k=1}^{\infty} \sum_{j=1}^{n_k} C_{kj} < \infty. \tag{4.1}$$

Second, for each v with $p(v) > 0$ and $k \geq 1$,

$$\phi_{kj}(v, \cdot) \text{ are probability densities with respect to } \ell^k \text{ such that } \{s : \phi_{kj}(v, s) > 0\}, \tag{4.2}$$

$$j = 1, \dots, n_k, \text{ are disjoint}$$

and

$$q_k(v, s) := \mathbf{1}\{s_1 + v > r\} \prod_{i=1}^k \frac{\varrho(s_i)}{v + s_1 + \dots + s_i} \leq \sum_{j=1}^{n_k} C_{kj} \phi_{kj}(v, s). \tag{4.3}$$

Owing to condition (4.3), ϕ_{kj} should be designed according to $\varrho(s) = s\varphi(s)$. Naturally, it is desirable to have ϕ_{kj} that are easy to handle. Let

$$C_0 = 1; \quad C_k = \sum_{j=1}^{n_k} C_{kj}, \quad k \geq 1; \quad \text{and} \quad C = \sum_{k=0}^{\infty} C_k.$$

A rejection sampling procedure for g under this setup is given in the following algorithm.

Algorithm 4.2. (*Rejection sampling for a positive i.d. random variable.*)

1. Sample $Z \sim p$ and κ such that $P\{\kappa = k\} = C_k/C, k \geq 0$.
2. If $\kappa = 0$ then output $X = Z$ and stop; otherwise, sample η such that $P\{\eta = j\} = C_{k\eta}/C_\kappa, j = 1, \dots, n_\kappa$.
3. Given Z, κ , and η , sample $S = (S_1, \dots, S_\kappa)$ from the density $\phi_{\kappa\eta}(Z, \cdot)$.
4. Sample $U \sim \text{Unif}(0, 1)$. If $C_{\kappa\eta}U\phi_{\kappa\eta}(Z, S) \leq q_\kappa(Z, S)$ then output $X = Z + S_1 + \dots + S_\kappa$ and stop; otherwise, go back to step 1 and start a new, independent iteration.

Theorem 4.1. *Algorithm 4.2 terminates eventually with probability 1, and its random output follows g in (3.6).*

Proof. Given a measurable set $A \subset \mathbb{R}$, in each iteration,

$$P\{X \in A \text{ is output}\} = P\{Z \in A, \kappa = 0\} + \int \sum_{k=1}^{\infty} \sum_{j=1}^{n_k} P\{Z \in dv, \kappa = k, \eta = j\} Q_{kj}(v),$$

where

$$Q_{kj}(v) = P\left\{v + S_1 + \dots + S_k \in A, U \leq \frac{q_k(v, S)}{C_{kj}\phi_{kj}(v, S)}\right\}, \quad S \sim \phi_{kj}(v, \cdot).$$

We have

$$P\{Z \in A, \kappa = 0\} = \frac{1}{C} \int \mathbf{1}\{v \in A\} p(v) \, dv$$

and, for each v with $p(v) > 0$, $k \geq 1$, and $j = 1, \dots, n_k$,

$$P\{Z \in dv, \kappa = k, \eta = j\} = \frac{C_{kj}}{C} p(v) \, dv.$$

By (4.2) and (4.3), for each s , if $q_k(v, s) > 0$ then there is exactly one $j \in \{1, \dots, n_k\}$ such that $\phi_{kj}(v, s) > 0$, and, conversely, if $\phi_{kj}(v, s) > 0$ for some j then $q_k(v, s) \leq C_{kj}\phi_{kj}(v, s)$ and $\phi_{kl}(v, s) = 0$ for all $l \neq j$. Thus, if $C_{kj} > 0$ then

$$\begin{aligned} Q_{kj}(v) &= \int \phi_{kj}(v, s) \mathbf{1}\{v + s_1 + \dots + s_k \in A\} P\left\{U \leq \frac{q_k(v, s)}{C_{kj}\phi_{kj}(v, s)}\right\} \, ds \\ &= \frac{1}{C_{kj}} \int q_k(v, s) \mathbf{1}\{\phi_{kj}(v, s) > 0\} \mathbf{1}\{v + s_1 + \dots + s_k \in A\} \, ds, \end{aligned}$$

which leads to

$$\begin{aligned} P\{Z \in dv, \kappa = k, \eta = j\} Q_{kj}(v) \\ = \frac{p(v) \, dv}{C} \int q_k(v, s) \mathbf{1}\{\phi_{kj}(v, s) > 0\} \mathbf{1}\{v + s_1 + \dots + s_k \in A\} \, ds. \end{aligned}$$

If $C_{kj} = 0$ then $P\{Z \in dv, \kappa = k, \eta = j\} = 0$ and $q_k(v, s) = 0$ for any s with $\phi_{kj}(v, s) > 0$. Therefore, the above equality still holds. As a result,

$$\begin{aligned} \int \sum_{j=1}^{n_k} P\{Z \in dv, \kappa = k, \eta = j\} Q_{kj}(v) \\ = \frac{1}{C} \int p(v) \, dv \int q_k(v, s) \sum_{j=1}^{n_k} \mathbf{1}\{\phi_{kj}(v, s) > 0\} \mathbf{1}\{v + s_1 + \dots + s_k \in A\} \, ds \\ = \frac{1}{C} \int p(v) \, dv \int q_k(v, s) \mathbf{1}\{v + s_1 + \dots + s_k \in A\} \, ds, \end{aligned} \tag{4.4}$$

where the last line is again due to conditions (4.2) and (4.3). Let $x = v + s_1 + \dots + s_k$, and, for $1 \leq i < k$, let $v_i = v + s_1 + \dots + s_i$. If $k = 1$ then $q_1(v, s) = \mathbf{1}\{x > r\}h(v, x)$ and, so, by a change of variable and Fubini's theorem, the iterated integral (4.4) is equal to

$$\frac{1}{C} \int \mathbf{1}\{x \in A\} \left[\int p(v) \mathbf{1}\{x > r\} h(v, x) \, dv \right] dx.$$

If $k > 1$ then it is easy to check that $q_k(v, s) = \mathbf{1}\{v_1 > r\}h(v, v_1) \dots h(v_{k-2}, v_{k-1})h(v_{k-1}, x)$ and, hence, by a change of variable and Fubini's theorem, the iterated integral (4.4) is equal to

$$\begin{aligned} \frac{1}{C} \int p(v) \, dv \int \mathbf{1}\{x \in A\} \mathbf{1}\{v_1 > r\} h(v, v_1) \dots h(v_{k-1}, x) \, dv_1 \dots \, dv_{k-1} \, dx \\ = \frac{1}{C} \int \mathbf{1}\{x \in A\} \left[\int p(v) h_{k-1}(v, x) \, dv \right] dx. \end{aligned}$$

Combining the above results and using (3.6), in each iteration, $P\{X \in A \text{ is output}\} = (aC)^{-1} \int_A g$. The proof is then completed by following the same argument as in the proof of Proposition 4.1.

4.3. Conditional density versus upper truncated Lévy density

We consider several implications of the results in Section 4.2. The statement below follows directly from Theorem 4.1.

Proposition 4.2. *Let X be a strictly positive i.d. random variable with known Lévy density. Given $r > 0$, if we can sample X conditional on $X \leq r$ then we can sample X using Algorithm 4.2. Conversely, as is well known, if we can sample X then, for any $r > 0$, we can sample X conditional on $X \leq r$ using rejection sampling.*

Basically, Proposition 4.2 states that the sampling of an i.d. random variable boils down to the sampling of its conditional density on some $(0, r]$. The questions are how to identify such r and how to sample from the corresponding conditional density.

Proposition 4.3. *Let φ_1 and φ_2 be Lévy densities with $\int \varphi_i = \infty$. Suppose that we can find an $r > 0$ such that $\varphi_1 = \varphi_2$ on $(0, r]$. Then, provided that we can sample for one of them, we can sample for the other using Algorithm 4.2.*

Proof. By assumption, φ_1 and φ_2 share an upper truncated version $\psi(t) = \varphi_1(t)\mathbf{1}\{t \leq r\}$. Let $X_i \sim \text{ID}(\varphi_i)$, $i = 1, 2$, and $Y \sim \text{ID}(\psi)$. As noted in the introduction, for $i = 1, 2$, the density of X_i conditional on $X_i \leq r$ is identical to that of Y conditional on $Y \leq r$. Thus, the conditional densities of X_i are equal to each other. If we can sample, say, X_1 , then we can sample X_1 conditional on $X_1 \leq r$, and, hence, sample X_2 conditional on $X_2 \leq r$. Then from Proposition 4.2 we can sample X_2 by applying Algorithm 4.2 to φ_2 .

The above proof provides the following answer to the question raised before Proposition 4.3. Given φ , find $r > 0$ and another Lévy density λ which we know how to sample for, such that $\varphi = \lambda$ on $(0, r]$. Sample $X \sim \text{ID}(\lambda)$ conditional on $X \leq r$. The sampled value then follows the conditional distribution of $\text{ID}(\varphi)$ on $(0, r]$.

As a further development along this line, suppose that, instead of sharing a common upper truncated version, φ_1 and φ_2 satisfy

$$\varphi_2(t) = [1 + O(t)]\varphi_1(t) \quad \text{as } t \rightarrow 0+. \tag{4.5}$$

Note that, under the much stronger condition $e^{-b_1 t}\varphi_1(t) \leq \varphi_2(t) \leq e^{b_2 t}\varphi_1(t)$ for some $b_1, b_2 \geq 0$, the methods in Section 2 can be used to sample for one φ_i based on the other.

Proposition 4.4. *For Lévy densities φ_1 and φ_2 that satisfy (4.5) with $\int \varphi_i = \infty$, if we can sample for one of them then we can also sample for the other using Algorithm 4.2, with possibly an extra step of sampling from a compound Poisson distribution.*

Proof. Since (4.5) is equivalent to $\varphi_1(t) = [1 + O(t)]\varphi_2(t)$ as $t \rightarrow 0+$, by symmetry, assume without loss of generality that we can sample for φ_1 . We can find constants $a_1 \geq 0$, $a_2 \geq 0$, and $r_0 > 0$ such that $(1 - a_1 t)\varphi_1(t) \leq \varphi_2(t) \leq (1 + a_2 t)\varphi_1(t)$ for $0 < t \leq r_0$. Fix $b \geq a_1$ and $0 < r \leq r_0$ such that $e^{-bt} \leq 1 - a_1 t$ for $0 < t \leq r$. If $a_1 = 0$, we can just let $b = 0$ and $r = r_0$. Define $\psi(t) = e^{-bt}\varphi_1(t)\mathbf{1}\{t \leq r\}$. Then it is straightforward to check that $\psi(t) \leq \varphi_2(t)\mathbf{1}\{t \leq r\}$. Let $\chi(t) = \varphi_2(t) - \psi(t)$. Consider the following procedure.

- Sample $X \sim \text{ID}(\psi)$ and $\xi \sim \text{ID}(\chi)$ independently. Return $X + \xi$.

From Proposition 4.3 we can sample for $\varphi_1(t)\mathbf{1}\{t \leq r\}$. Then by exponential tilting we can sample for $\psi(t)$; see Section 2.2. On the other hand,

$$\begin{aligned} 0 &\leq \chi(t) \\ &= \varphi_2(t)\mathbf{1}\{t \leq r\} - \psi(t) + \varphi_2(t)\mathbf{1}\{t > r\} \\ &\leq (1 + a_2t)\varphi_1(t)\mathbf{1}\{t \leq r\} - \psi(t) + \varphi_2(t)\mathbf{1}\{t > r\} \\ &= (1 + a_2t - e^{-bt})\varphi_1(t)\mathbf{1}\{t \leq r\} + \varphi_2(t)\mathbf{1}\{t > r\} \\ &\leq (a_2 + b)t\varphi_1(t)\mathbf{1}\{t \leq r\} + \varphi_2(t)\mathbf{1}\{t > r\}. \end{aligned}$$

Therefore, $\chi \in L^1(0, \infty)$, giving rise to a compound Poisson distribution.

Finally, note that if $\varphi_1(t)$ can be decomposed into

$$\varphi_1(t) = [1 + O(t)]\varphi_2(t) + \chi(t),$$

where φ_2 is a Lévy density with $\int \varphi_2 = \infty$ that we know how to sample for and $\chi \geq 0$ is integrable, then by Proposition 4.4 we can also sample for φ_1 .

In subsequent sections, all examples are based on Propositions 4.3 and 4.4. In most of the examples, the sampling for an upper-truncated Lévy density takes centre stage. This is natural. A generic approach to the sampling for a Lévy density φ is by sampling for $\varphi(t)\mathbf{1}\{t \leq r\}$ and $\varphi(t)\mathbf{1}\{t > r\}$ independently. After all, it is $\varphi(t)\mathbf{1}\{t \leq r\}$ that determines the conditional density of $ID(\varphi)$ on $(0, r]$, whose sampling is the crucial starting point of Algorithm 4.2. On the other hand, in principle, as long as we know how to sample from the conditional density, Algorithm 4.2 can be used whether or not the Lévy density is upper truncated. This observation is useful sometimes; see Section 5.4 for an example.

4.4. Some corollaries on complexity

The complexity of an algorithm can be measured in many ways (cf. [28]). For Algorithm 4.2, a useful measure is the number of iterations it needs to generate one output.

Proposition 4.5. *The number of iterations of Algorithm 4.2 follows the geometric distribution with mean value $C \int_0^r g$.*

Proof. From the last line of the proof of Theorem 4.1, at each iteration, $P\{X \text{ is output}\} = (aC)^{-1}$ and, by (3.5), $a = \int_0^r g$.

The amount of time required to generate one output is an important measure of complexity (cf. [13]). Denote by ω the vector of random values sampled in a single iteration of Algorithm 4.2, including Z, κ , and, provided that $\kappa \geq 1, \eta, S$, and U . Let $D(\omega)$ be the time taken to generate and process ω . For example, the computations of $\phi_{\kappa\eta}(Z, S)$ and $q_\kappa(Z, S)$ in step 4 are part of the processing of ω . Note that whether or not the procedure stops after the iteration is completely determined by ω .

Proposition 4.6. *Let T denote the total time taken for Algorithm 4.2 to generate a random output. Then $E T = E N \times E[D(\omega)] = (C \int_0^r g) E[D(\omega)]$.*

Proof. For $k \geq 1$, let ω_k denote the vector of random values sampled in iteration k . Then ω_k are i.i.d. $\sim \omega$. Let N be the number of iterations. Then $T = \sum_{k=1}^N D(\omega_k)$. Since N is a stopping time with respect to $\omega_1, \omega_2, \dots$, $E T = E[D(\omega)] \times E N$ (cf. [7, p. 101]). Then the result follows from Proposition 4.5.

5. A general design

5.1. Description of the design

Let φ be a Lévy density such that $\int \varphi = \infty$ and the density of $ID(\varphi)$ is locally bounded on $(0, \infty)$. In order to use Algorithm 4.2 to sample for φ , we need to find constants C_{kj} and instrumental functions ϕ_{kj} satisfying conditions (4.1)–(4.3) that are easy to compute and use. The following design is immediate from Theorem 4.1.

Proposition 5.1. *Let q_k be the functions in (4.3). Suppose that there are constants $b_k \geq 0$ with $\sum_{k=1}^\infty b_k < \infty$ and functions $\bar{q}_k(\cdot, \cdot)$ such that $q_k(v, \cdot) \leq \bar{q}_k(v, \cdot)$ for each v with $p(v) > 0$ and*

$$b_k \geq \sup_{p(v)>0} Q_k(v) \quad \text{with} \quad Q_k(v) = \int \bar{q}_k(v, s) \, ds.$$

Then in Algorithm 4.2, we can set $n_k = 1$, $C_{k1} = b_k$, and $\phi_{k1}(v, \cdot) = \bar{q}_k(v, \cdot)/Q_k(v)$ for $k \geq 1$.

Remark. Under the above setup, step 4 of Algorithm 4.2 involves a comparison between $U \sim \text{Unif}(0, 1)$ and $\zeta := [Q_\kappa(Z)/b_\kappa]q_\kappa(Z, S)/\bar{q}_\kappa(Z, S)$, which is hard if $Q_k(\cdot)$ is difficult to compute. However, given ζ , the whole purpose of the comparison is to stop the iteration with probability ζ . As we show next, alternative methods to stop the algorithm can be found.

5.2. An example on the upper-truncated stable Lévy density

Let $\alpha \in (0, 1)$. We consider how to apply Proposition 5.1 to sample for the Lévy density

$$\varphi(t) = ct^{-\alpha-1}\mathbf{1}\{0 < t \leq r\},$$

where $c, r \in (0, \infty)$. It can be shown that the density of $ID(\varphi)$ is smooth and bounded on \mathbb{R} (cf. [33, p. 190]). Let $X \sim ID(\varphi)$. By the Laplace transform, $X \sim c^{1/\alpha}X'$, where X' has Lévy density $t^{-1-\alpha}\mathbf{1}\{0 < t \leq rc^{-1/\alpha}\}$. Thus, without loss of generality, assume that $c = 1$ in the following.

Let $\lambda(t) = t^{-\alpha-1}\mathbf{1}\{t > 0\}$. The sampling of $ID(\lambda)$ is well known [10], [13], [21], [32]. This combined with Proposition 5.1 leads to the following algorithm to sample for φ .

Algorithm 5.1. Set $b_0 = 1$ and, for $k \geq 1$, $b_k = \alpha\theta^k\Gamma(k\alpha)/\Gamma(k)$ with $\theta = \Gamma(1 - \alpha)/(r^\alpha\alpha)$.

1. Sample $Z \sim ID(\lambda)$ conditional on $Z \leq r$ and κ such that $P\{\kappa = k\} = b_k/\sum_{i \geq 0} b_i$, $k \geq 0$.
2. If $\kappa = 0$ then output $X = Z$ and stop.
3. Sample $T_1 \sim \text{Beta}(k\alpha, 1 - \alpha)$ conditional on $T_1 < Z/r$. Set $S_1 = Z(1/T_1 - 1)$. If $\kappa > 1$ then, sequentially, for $i = 2, \dots, \kappa$, sample $T_i \sim \text{Beta}((\kappa - i + 1)\alpha, 1 - \alpha)$ and set $S_i = (Z + S_1 + \dots + S_{i-1})(1/T_i - 1)$.
4. Sample $U \sim \text{Unif}(0, 1)$ and $T \sim \text{Beta}(k\alpha, 1 - \alpha)$. If $S_i \leq r$ for all $i \leq \kappa$ and $U \leq [(r - rT)/(r - ZT)]^\alpha$, then output $X = Z + S_1 + \dots + S_\kappa$ and stop; otherwise, go back to step 1.

Remarks 5.1. (i) In step 3 of Algorithm 5.1, if an $S_i > r$ is generated at any time point, we can immediately terminate the current iteration and start a new one. Although this modification improves efficiency, for clarity, we do not implement it in what follows.

(ii) Using the Laplace transform, if $r \in (0, 1)$ then $X \sim r(Y_1 + \dots + Y_m)$, where $m = \lceil r^{-\alpha} \rceil \geq 1$ and Y_1, \dots, Y_m are i.i.d. with Lévy density $t^{-\alpha-1} \mathbf{1}\{0 < t \leq r'\}$, with $r' = rm^{1/\alpha} \geq 1$. While this suggests that it may suffice to consider only $r \geq 1$, for now we shall consider any $r > 0$.

5.2.1. *Justification.* In step 1 of Algorithm 5.1, it is clear that Z conditional on $Z \leq r$ is identically distributed as X conditional on $X \leq r$. To justify the sampling of κ , we need some calculations. For $v \in (0, r]$ and $s = (s_1, \dots, s_k)$, since $q(s_i) = s_i \varphi(s_i) = \mathbf{1}\{0 < s_i \leq r\} s_i^{-\alpha}$, by (4.3),

$$q_k(v, s) = \mathbf{1}\{s_1 + v > r\} \prod_{i=1}^k \frac{\mathbf{1}\{0 < s_i \leq r\} s_i^{-\alpha}}{v + s_1 + \dots + s_i}.$$

To apply Proposition 5.1, let

$$\bar{q}_k(v, s) = \mathbf{1}\{s_1 + v > r\} \prod_{i=1}^k \frac{\mathbf{1}\{s_i > 0\} s_i^{-\alpha}}{v + s_1 + \dots + s_i}.$$

Then $q_k(v, s) \leq \bar{q}_k(v, s)$. To obtain $Q_k(v) = \int \bar{q}_k(v, s) ds$, we use

$$\int_0^\infty \frac{s^{-\alpha} ds}{(z + s)^{1+\beta}} = z^{-\alpha-\beta} \frac{\Gamma(1-\alpha)\Gamma(\alpha+\beta)}{\Gamma(\beta+1)}, \quad z > 0, \alpha \in (0, 1), \beta \geq 0,$$

which can be verified by making the change of variable $s = z(u^{-1} - 1)$ and using the properties of beta functions. Integrating over s_k, s_{k-1}, \dots, s_1 , we obtain, for $k \geq 1$ and $j = 1, \dots, k - 1$,

$$\begin{aligned} \int \bar{q}_k(v, s) ds_k \dots ds_{k-j+1} &= \prod_{i=1}^j \frac{\Gamma(1-\alpha)\Gamma(i\alpha)}{\Gamma(1+(i-1)\alpha)} \mathbf{1}\{s_1 + v > r\} \left(\prod_{i=1}^{k-j} \frac{\mathbf{1}\{s_i > 0\} s_i^{-\alpha}}{v + s_1 + \dots + s_i} \right) \\ &\quad \times \frac{1}{(v + s_1 + \dots + s_{k-j})^{j\alpha}} \end{aligned} \tag{5.1}$$

and

$$\begin{aligned} Q_k(v) &= \prod_{i=1}^{k-1} \frac{\Gamma(1-\alpha)\Gamma(i\alpha)}{\Gamma(1+(i-1)\alpha)} \int \mathbf{1}\{s_1 + v > r\} \frac{s_1^{-\alpha} ds_1}{(v + s_1)^{1+(k-1)\alpha}} \\ &= \prod_{i=1}^{k-1} \frac{\Gamma(1-\alpha)\Gamma(i\alpha)}{\Gamma(1+(i-1)\alpha)} \int_0^\infty \frac{(r - v + u)^{-\alpha} du}{(r + u)^{1+(k-1)\alpha}}, \end{aligned} \tag{5.2}$$

making the change of variable $s_1 = r - v + u$. It follows that

$$Q_k(v) \leq Q_k(r) = r^{-\alpha k} \prod_{i=1}^k \frac{\Gamma(1-\alpha)\Gamma(i\alpha)}{\Gamma(1+(i-1)\alpha)} = \frac{r^{-\alpha k} [\Gamma(1-\alpha)]^k \Gamma(k\alpha)}{\alpha^{k-1} \Gamma(k)} = \frac{\alpha \theta^k \Gamma(k\alpha)}{\Gamma(k)}.$$

Clearly, $b_k = Q_k(r)$. As $0 < \alpha < 1$,

$$\begin{aligned} \sum_{k=1}^\infty b_k &= \sum_{k=1}^\infty \frac{\alpha \theta^k}{(k-1)!} \int_0^\infty x^{k\alpha-1} e^{-x} dx \\ &= \alpha \int_0^\infty \theta x^{\alpha-1} e^{-x} \sum_{k=0}^\infty \frac{\theta^k x^{\alpha k}}{k!} dx \\ &= \alpha \int_0^\infty \theta x^{\alpha-1} e^{-x} e^{\theta x^\alpha} dx \end{aligned}$$

$$\begin{aligned}
 &= \int_0^\infty e^{z-(z/\theta)^{1/\alpha}} dz \\
 &= \int_0^\infty \exp\left\{z - r \left[\frac{\alpha z}{\Gamma(1-\alpha)}\right]^{1/\alpha}\right\} dz \\
 &< \infty.
 \end{aligned}
 \tag{5.3}$$

Then by Proposition 5.1 we obtain the sampling of κ in step 1.

Step 2 directly follows from the general procedure in Algorithm 4.2. To justify step 3, according to Algorithm 4.2, given $Z \in (0, r]$ and $\kappa \geq 1$, we need to sample from the density

$$\phi_\kappa(Z, s) = \frac{\bar{q}_\kappa(Z, s)}{Q_\kappa(Z)} = C \mathbf{1}\{s_1 + Z > r\} \prod_{i=1}^\kappa \frac{\mathbf{1}\{s_i > 0\} s_i^{-\alpha}}{Z + s_1 + \dots + s_i},$$

where C is the normalizing constant. The exact value of C is not important and may change from line to line in the following. Let $S = (S_1, \dots, S_\kappa) \sim \phi_\kappa(Z, \cdot)$. From (5.1), S_1 has density $C \mathbf{1}\{x > r - Z\} x^{-\alpha} (Z + x)^{-1-(\kappa-1)\alpha}$ at $x \in \mathbb{R}$. By calculation, $Z/(Z + S_1)$ has density $C \mathbf{1}\{0 < x < Z/r\} (1-x)^{-\alpha} x^{\kappa\alpha-1}$, the same as that of $T_1 \sim \text{Beta}(\kappa\alpha, 1-\alpha)$ conditional on $T_1 < Z/r$. Thus, S_1 can be sampled as $Z(1/T_1 - 1)$. For $i > 1$, conditional on S_1, \dots, S_{i-1} , it can likewise be seen that S_i has density $C \mathbf{1}\{x > 0\} x^{-\alpha} (Z + S_1 + \dots + S_{i-1} + x)^{-1-(\kappa-i)\alpha}$ and, thus, can be sampled as $(Z + S_1 + \dots + S_{i-1})(1/T_i - 1)$, with $T_i \sim \text{Beta}((\kappa - i + 1)\alpha, 1 - \alpha)$. Step 3 is then established.

Finally, as remarked after Proposition 5.1, given Z, κ , and $S = (S_1, \dots, S_\kappa)$, all we need to do in step 4 is to stop the iteration with probability $[Q_\kappa(Z)/b_\kappa] q_\kappa(Z, S)/\bar{q}_\kappa(Z, S)$, which is equal to $\mathbf{1}\{\text{all } S_i \leq r\} Q_\kappa(Z)/Q_\kappa(r)$. However, $Q_\kappa(Z)$ is not easy to compute. To get around the problem, make the change of variable $u = r(t^{-1} - 1)$ in (5.2). Then we obtain $Q_\kappa(Z)/Q_\kappa(r) = E \xi$, where $\xi = [(r - rT)/(r - ZT)]^\alpha$ with $T \sim \text{Beta}(\kappa\alpha, 1 - \alpha)$. Note that $0 \leq \xi \leq 1$. Since $E \xi = P\{U \leq \xi\}$ for any ξ independent of U with $P\{0 \leq \xi \leq 1\} = 1$, step 4 is justified.

5.2.2. Complexity. Let N be the number of iterations required by Algorithm 5.1 to sample one $X \sim \text{ID}(\varphi)$. By Proposition 4.5, $EN = P\{X \leq r\} \sum_{k \geq 0} b_k$. Let $\chi(t) = t^{-\alpha-1} \mathbf{1}\{t > r\}$ and $\xi \sim \text{ID}(\chi)$ be independent of X . Then $Z \sim X + \xi$ and $P\{X \leq r\} = P\{Z \leq r\}/P\{\xi = 0\}$. The expression of $P\{Z \leq r\}$ is known [10], while $P\{\xi = 0\} = \exp\{-\int_r^\infty \chi\} = \exp\{-r^{-\alpha}/\alpha\}$. Together with (5.3) this gives

$$EN = \exp\left\{\frac{r^{-\alpha}}{\alpha}\right\} P\{Z \leq r\} \left(1 + \int_0^\infty \exp\left\{z - r \left(\frac{\alpha z}{\Gamma(1-\alpha)}\right)^{1/\alpha}\right\} dz\right).$$

By Proposition 4.6, $ET = ED \times EN$, where T is the amount of time required to sample one $X \sim \text{ID}(\varphi)$ and D is the amount of time required to complete a single iteration. Each iteration has to sample (i) one $Z \sim \text{ID}(\lambda)$ conditional on $Z \leq r$, (ii) one κ from a distribution that depends on r , and (iii) provided that $\kappa \geq 1$, one $T_1 \sim \text{Beta}(\kappa\alpha, 1 - \alpha)$ conditional on $T_1 \leq Z/r$ and, for each $i = 1, \dots, \kappa$, a value from $\text{Beta}(i\alpha, 1 - \alpha)$. These samplings account for most of D . Denote by $D_1(r)$ the amount of time required for the conditional sampling of Z , by $D_2(r)$ the amount of time required for the sampling of κ , and, for $k \geq 1$ and $z \in (0, 1]$, by $D_3(k, z)$ the amount of time required to sample $\zeta \sim \text{Beta}(k\alpha, 1 - \alpha)$ conditional on $\zeta \leq z$. In Appendix A we show that it is possible to bound $ED_1(r)$, $ED_2(r)$, and $ED_3(k, z)$ uniformly for $r > 0, k \geq 1$, and $z > 0$. Consequently, ED is of the same order as $\sum_{k \geq 0} (1 + k)b_k / \sum_{k \geq 0} b_k$.

It is easy to see that, as $r \rightarrow \infty$, $E N \rightarrow 1$ and $E T \rightarrow E D_1(\infty)$, the expected amount of time to sample one $Z \sim \text{ID}(\lambda)$. On the other hand, as $r \rightarrow 0$, it can be shown that $E N \asymp \exp\{r^{-\alpha}/\alpha\}$, where, for two functions f and g , $f \asymp g$ stands for $f = O(g)$ and $g = O(f)$; see Appendix A. Therefore, if we directly apply Algorithm 5.1 to $\varphi(t) = t^{-\alpha-1}\mathbf{1}\{t \leq r\}$ then $E N$ is extremely large for small r . However, by Remarks 5.1(ii) we can instead sample $O(r^{-\alpha})$ i.i.d. random variables with Lévy density $t^{-\alpha}\mathbf{1}\{t \leq r'\}$ for some $r' \geq 1$ and then take their weighted sum. This way, both $E N$ and $E T$ are lowered to $O(r^{-\alpha})$.

5.3. An application to a class of i.d. distributions

As an application of the result in Section 5.2, consider Lévy densities of the form

$$\varphi(t) = \psi(t)(e^t - 1)^{-\alpha-1}\mathbf{1}\{t > 0\}, \quad \alpha \in (0, 1),$$

where $\psi(t) \geq 0$ is a measurable function on $(0, \infty)$ such that $\int_c^\infty \psi(t)e^{-(\alpha+1)t} dt < \infty$ for any $c > 0$ and $\psi(t) = 1 + O(t)$ as $t \rightarrow 0+$. An algorithm to sample for φ is as follows.

Algorithm 5.2. Set $r \in (0, \infty]$ and $\beta < \alpha + 1$ such that $\psi(t) \geq e^{\beta t}$ for $t \in (0, r]$. Note that r can be ∞ .

1. Keep sampling (U, Z) until $U \leq e^{(\beta-\alpha-1)Z}$, where $U \sim \text{Unif}(0, 1)$ and $Z \sim \text{ID}(\varphi_1)$ with $\varphi_1(t) = t^{-\alpha-1}\mathbf{1}\{t \leq r\}$ are independent.
2. Sample $\xi \sim \text{ID}(\chi)$, with $\chi(t) = \varphi(t) - e^{(\beta-\alpha-1)t}\varphi_1(t) \geq 0$ being integrable. Return $X = Z + \xi$.

If $\psi(t) = e^{ct}$, where $c < \alpha + 1$, then $\text{ID}(\varphi)$ belongs to Lamperti-stable distributions [9], which arise from positive self-similar Markov processes and related processes (cf. [5], [9], [24]–[26]). In this case, we can simply set $r = \infty$ and $\beta = c$. Although the sampling of Lamperti-stable distributions with $\alpha \in (0, 1)$ is quite simple, somewhat surprisingly, it seems that it has not been explicitly stated in the literature.

More generally, if there exists a $c < \alpha + 1$ such that $e^{ct} \leq \psi(t)$ for all $t > 0$ then we can set $r = \infty$ and $\beta = c$. However, it is easy to find simple functions ψ such that, for any c , $\inf_{t>0}[e^{-ct}\psi(t)] = 0$; for example, e^{-t^2} , $e^t[1 - \sin(t)]$, and $(1 - t)^2$. For these functions, we need to select $r < \infty$ and $\beta < \alpha + 1$ accordingly.

5.3.1. *Justification.* Since $\psi(t) = 1 + O(t)$ as $t \rightarrow 0+$, we can indeed find r and β to meet the requirement of Algorithm 5.2. From Section 2, the random variable Z sampled in step 1 has Lévy density $e^{(\beta-\alpha-1)t}t^{-\alpha-1}\mathbf{1}\{t \leq r\}$. From the choices of r and β , if $t > r$ then $\chi(t) = \varphi(t) \geq 0$, and if $0 < t \leq r$ then, as $1 - e^{-t} < t$ for all $t > 0$,

$$\begin{aligned} \chi(t) &= e^{-(\alpha+1)t}[\psi(t)(1 - e^{-t})^{-\alpha-1} - e^{\beta t}t^{-\alpha-1}] \\ &\geq e^{(\beta-\alpha-1)t}[(1 - e^{-t})^{-\alpha-1} - t^{-\alpha-1}] \\ &> 0, \end{aligned}$$

showing that χ is a Lévy density. Following the proof of Proposition 4.4, it can be shown that χ is integrable. Thus, the correctness of step 2 is established.

5.3.2. *Complexity.* We consider steps 1 and 2 of Algorithm 5.2 separately. By Proposition 4.6, the expected amount of time required to complete step 1 is $E T \times E N$, where T is the amount of time taken to sample one observation from $\text{ID}(\varphi_1)$ and N is the total number of iterations required by the step. The analysis on $E T$ is identical to that given in Section 5.2. Although, since

in each iteration the probability of acceptance is $P\{U \leq e^{(\beta-\alpha-1)Z}\} = E(e^{(\beta-\alpha-1)Z}) = \exp\{\int (e^{(\beta-\alpha-1)t} - 1)\varphi_1(t) dt\}$, we have $EN = \exp\{\int_0^r (1 - e^{(\beta-\alpha-1)t})t^{-\alpha-1} dt\} \leq \exp\{(\alpha + 1 - \beta)r^{1-\alpha}/(1 - \alpha)\}$.

In step 2, ξ can be sampled as the sum of the coordinates of the points in a Poisson process with intensity χ . Let K be the number of points in the process. Then the expected amount of time required to complete step 2 is of the same order as $E K = \int \chi = \int (\psi(t) - e^{(\beta-\alpha-1)t}t^{-\alpha-1} \times \mathbf{1}\{t \leq r\}) dt$.

5.4. A theoretical application

As a theoretical application of Proposition 5.1, we evaluate, for $c > 0$ and $r > 0$,

$$B := \sum_{k=1}^{\infty} c^k E\left(\prod_{i=1}^k \frac{1}{r + S_1 + \dots + S_i}\right), \quad S_1, S_2, \dots \text{ i.i.d. } \sim \text{Exp}(1).$$

Recall that the Gamma($c, 1$) distribution has density $g(x) = x^{c-1}e^{-x}\mathbf{1}\{x > 0\}/\Gamma(c)$ and Lévy density $\varphi(t) = ct^{-1}e^{-t}\mathbf{1}\{t > 0\}$ (cf. [30]). Using the design in Proposition 5.1 with $\varrho(t) = t\varphi(t) = ce^{-t}\mathbf{1}\{t > 0\}$, Algorithm 4.2 can be reformulated as follows.

Algorithm 5.3. Define, for $k \geq 1, 0 < v \leq r$, and $s = (s_1, \dots, s_k), s_i > 0$,

$$q_k(v, s) = \mathbf{1}\{s_1 + v > r\} \prod_{i=1}^k \frac{\varrho(s_i)}{v + s_1 + \dots + s_i} = \mathbf{1}\{s_1 + v > r\} \prod_{i=1}^k \frac{ce^{-s_i} \mathbf{1}\{s_i > 0\}}{v + s_1 + \dots + s_i}.$$

Set $\phi_k(v, s) = q_k(v, s)/Q_k(v)$, where $Q_k(v) = \int q_k(v, s) ds$. Set $b_0 = 1$ and, for $k \geq 1, b_k = Q_k(r)$. (We will show that $Q_k(r) > Q_k(v)$ for all $v \in (0, r)$.)

1. Sample $Z \sim \text{Gamma}(c, 1)$ conditional on $Z \leq r$ and κ from $\{0, 1, 2, \dots\}$ such that $P\{\kappa = k\} = b_k / \sum_{i=0}^{\infty} b_i$.
2. If $\kappa = 0$ then output $X = Z$ and stop; otherwise, continue.
3. Given Z and $\kappa \geq 1$, sample $S = (S_1, \dots, S_\kappa)$ from the density $\phi_k(Z, \cdot)$.
4. Sample $U \sim \text{Unif}(0, 1)$. If $b_k U \phi_k(Z, S) \leq q_k(Z, S)$ then output $X = Z + S_1 + \dots + S_\kappa$ and stop; otherwise, go back to step 1.

Algorithm 5.3 cannot be used to actually sample from Gamma($c, 1$), since its step 1 relies on the sampling from Gamma($c, 1$) itself. Nevertheless, by Proposition 5.1, its random output X follows Gamma($c, 1$). We use this fact to compute B .

First, we need to obtain $Q_k(v)$. Let $s'_1 = s_1 + v - r$. Then

$$q_k(v, s) = c^k \mathbf{1}\{s'_1 > 0, \text{ all } s_i > 0\} e^{-(r-v)-(s'_1+s_2+\dots+s_k)} \prod_{i=1}^k \frac{1}{r + s'_1 + s_2 + \dots + s_i}.$$

Integrating over s'_1, s_2, \dots, s_k , it follows that, for $k \geq 1$,

$$Q_k(v) = e^{-(r-v)} Q_k(r), \quad \text{with} \quad Q_k(r) = c^k E\left(\prod_{i=1}^k \frac{1}{r + S_1 + \dots + S_i}\right). \quad (5.4)$$

In each iteration, if $\kappa = 0$ then the algorithm stops and outputs $X = Z$. This is the case only where the algorithm outputs a value in $(0, r]$. If $\kappa \geq 1$ then, in order for the algorithm to stop, there has to be $b_\kappa U \phi_\kappa(Z, S) \leq q_\kappa(Z, S)$ in step 4. Since the event is equivalent to $U \leq Q_\kappa(Z)/b_\kappa = Q_\kappa(Z)/Q_\kappa(r)$, by (5.4), it has probability $e^{-(r-Z)}$ conditional on Z and κ . Observe that $B = \sum_{k=1}^\infty Q_k(r)$. Consequently,

$$P\{X \leq r\} = \frac{P\{\kappa = 0\}}{P\{\kappa = 0\} + \sum_{k=1}^\infty P\{\kappa = k\} E(e^{-(r-Z)} \mid Z \leq r)} = \frac{1}{1 + B E(e^{-(r-Z)} \mid Z \leq r)}.$$

Since $X \sim Z$, $P\{X \leq r\} = P\{Z \leq r\}$. Then, after some calculation,

$$B = \frac{P\{Z > r\}}{E(e^{-(r-Z)} \mathbf{1}\{Z \leq r\})} = c\Gamma(c)r^{-c}e^r P\{Z > r\} = \int_1^\infty e^{r(1-t^{1/c})} dt.$$

6. Another general design

6.1. Description of the design

Let φ be a Lévy density such that $\int \varphi = \infty$ and the density of $ID(\varphi)$ is locally bounded on $(0, \infty)$. In this section we consider a design for Algorithm 4.2 that employs two ϕ_{kj} to handle each q_k , $k \geq 1$. In contrast, the design in Section 5 uses one ϕ_{kj} for each q_k . Let $r > 0$ be fixed such that we know how to sample from the conditional density of $ID(\varphi)$ on $(0, r]$. By upper truncating φ if necessary, assume that $\int \varrho < \infty$, where $\varrho(t) = t\varphi(t)$. Moreover, by decreasing r if necessary, assume that $\int_v^r \varphi > 0$ for any $v < r$. Indeed, $F(s) = \int_s^r \varphi$ is differentiable almost everywhere and, as $\int_0^r \varphi = \infty$, the set of $s \in (0, r]$ with $F'(s) = -\varphi(s) < 0$ is nonempty [31]. We can reset r to any such s if necessary. Suppose that we can find a suitable Lévy density λ with support in $[0, \infty)$ such that

$$\varphi(t) \leq \lambda(t), \quad M := \int t\lambda(t) dt < \infty. \tag{6.1}$$

While $\lambda = \varphi$ clearly satisfies (6.1), to facilitate sampling, we sometimes need $\lambda \neq \varphi$. Let

$$\psi(t) = \frac{t\lambda(t)}{M}.$$

Then ψ is a probability density with support in $[0, \infty)$. With a little abuse of notation, for $a > 0$, denote by $\psi(\cdot \mid V > a)$ the density of a generic variable $V \sim \psi$ conditional on $V > a$. By the assumptions on φ and λ , for any $a < r$, $P\{V > a\} > 0$ and, hence, the conditional density is well defined. For $\alpha \in (0, 1)$, denote by θ_α the α th quantile of the distribution with density ψ . Then, for $V \sim \psi$, $P\{V < \theta_\alpha\} = \alpha$. Define $m_k = \lfloor k/2 \rfloor$ for $k \geq 1$.

Proposition 6.1. Fix $\alpha \in (0, \frac{1}{2}]$ with $2(M/r)\sqrt{\alpha(1-\alpha)} < 1$. Define

$$N(s) = \sum_{i=2}^k \mathbf{1}\{s_i < \theta_\alpha\}, \quad s = (s_1, \dots, s_k), \quad k \geq 1.$$

Then Algorithm 4.2 can sample for φ if we set, given $v \in (0, r)$, for $k = 1$ and $s \in \mathbb{R}$,

$$n_1 = 1, \quad C_1 = C_{11} = \frac{M}{r}, \quad \phi_{11}(v, s) = \psi(s \mid V > r - v), \tag{6.2}$$

and, for $k > 1$ and $s = (s_1, \dots, s_k)$,

$$n_k = 2, \quad C_{k1} = \left(\frac{M}{r}\right)^k [4\alpha(1 - \alpha)]^{m_k}, \quad C_{k2} = \frac{(M/r)^k (r/\theta_\alpha)^{m_k}}{m_k!}, \quad (6.3)$$

$$\phi_{k1}(v, s) = D_{k1}^{-1} \psi(s_1 | V > r - v) \prod_{i=2}^k \psi(s_i) \mathbf{1}\{N(s) \geq m_k\}, \quad (6.4)$$

$$\phi_{k2}(v, s) = D_{k2}^{-1} \psi(s_1 | V > r - v) \prod_{i=2}^k \psi(s_i) \mathbf{1}\{N(s) < m_k\}, \quad (6.5)$$

where $D_{k1} = P\{\eta_k \geq m_k\}$ and $D_{k2} = P\{\eta_k < m_k\}$, with $\eta_k \sim \text{Binomial}(k - 1, \alpha)$.

Remarks 6.1. (i) In step 1 of Algorithm 4.2 we need to sample κ . From (6.2) and (6.3), $P\{\kappa = 1 | m_\kappa = 0\} = 1$ and, for $k \geq 1$, $P\{\kappa = 2k | m_\kappa = k\} = 1 - P\{\kappa = 2k + 1 | m_\kappa = k\} = 1/(1 + M/r)$. Therefore, once m_κ is sampled, κ can be sampled conditional on m_κ . On the other hand, m_κ follows a mixture of the degenerate distribution at 0, a geometric distribution, and a Poisson distribution, with the latter two conditional on positive integers. It is known that, using rejection sampling, the distributions can be sampled with the expected number of iterations uniformly bounded for all the parameters involved [13], [14], [19].

(ii) In step 3 of Algorithm 4.2 we need to sample from ϕ_{kj} . As seen below, $h(s_2, \dots, s_k) = D_{k1}^{-1} \prod_{i=2}^k \psi(s_i) \mathbf{1}\{N(s) \geq m_k\}$ is a density. Given $v \in (0, r]$, to sample from $\phi_{k1}(v, \cdot)$, we can sample $S_1 \sim \psi(\cdot | V > r - v)$ and $T = (S_2, \dots, S_k) \sim h$ independently. To sample T , first sample $\eta_k \sim \text{Binomial}(k - 1, \alpha)$ conditional on $\eta_k \geq m_k$; then, given η_k , sample ξ_1, \dots, ξ_{k-1} independently such that $\xi_i \sim \psi(\cdot | V < \theta_\alpha)$ for $i \leq \eta_k$ and $\xi_i \sim \psi(\cdot | V \geq \theta_\alpha)$ for $i > \eta_k$; finally, set T as a random permutation of ξ_1, \dots, ξ_{k-1} . We can sample from $\phi_{k2}(v, \cdot)$ similarly, except that we should sample $\eta_k \sim \text{Binomial}(k - 1, \alpha)$ conditional on $\eta_k < m_k$ instead.

(iii) In step 4 in Algorithm 4.2 we need to check whether, given $U \sim \text{Unif}(0, 1)$ and $S \sim \phi_{kj}(v, \cdot)$, $C_{kj} U \phi_{kj}(v, S) \leq q_k(v, S)$. It can be seen that the inequality is equivalent to

$$U \leq A_{kj} P\{V > r - v\} \prod_{i=1}^k \frac{r\varphi(S_i)/\lambda(S_i)}{v + S_1 + \dots + S_i}, \quad (6.6)$$

where $A_{11} = 1$, $A_{k1} = D_{k1}[4\alpha(1 - \alpha)]^{-m_k}$, $A_{k2} = D_{k2}m_k!(\theta_\alpha/r)^{m_k}$ for $k \geq 2$, and $V \sim \psi$. Note that, since S is a sampled value, factors such as $\mathbf{1}\{S_1 > r - v\}$ and $\mathbf{1}\{N(s) \geq m_k\}$ are unnecessary in (6.6) as they are equal to 1 (with probability 1).

(iv) Let N denote the number of iterations required by Algorithm 4.2 under the above design. Then, by Proposition 4.5,

$$\begin{aligned} E N &\leq 1 + \frac{M}{r} + \sum_{k=2}^{\infty} \left(\left(\frac{M}{r}\right)^k [4\alpha(1 - \alpha)]^{m_k} + \frac{(M/r)^k (r/\theta_\alpha)^{m_k}}{m_k!} \right) \\ &= \left(1 + \frac{M}{r}\right) \left(\frac{4M^2\alpha(1 - \alpha)/r^2}{1 - 4M^2\alpha(1 - \alpha)/r^2} + e^{M^2/(r\theta_\alpha)} \right). \end{aligned} \quad (6.7)$$

Proof of Proposition 4.1. We first check that ϕ_{kj} satisfies condition (4.2). The proof for ϕ_{11} is trivial. Let $k \geq 2$. Since $\psi(\cdot | V > r - v)$ is a probability density and $N(s)$ depends only

on s_2, \dots, s_k ,

$$\int \psi_{k1}(v, s) \, ds = D_{k1}^{-1} \int \prod_{i=2}^k \psi(s_i) \mathbf{1}\{N(s) \geq m_k\} \, ds_2 \cdots ds_k = D_{k1}^{-1} \mathbf{P}\{\zeta \geq m_k\},$$

where $\zeta = \sum_{i=2}^k \mathbf{1}\{S_i < \theta_\alpha\}$, with S_i i.i.d. $\sim \psi$. Since $\zeta \sim \text{Binomial}(k - 1, \alpha)$, the above integral is 1 and $\psi_{k1}(v, \cdot)$ is a probability density. Similarly, $\psi_{k2}(v, \cdot)$ is a probability density. Since $\psi_{k1}(v, s)$ and $\psi_{k2}(v, s)$ clearly cannot both be positive, condition (4.2) is satisfied.

We now check that condition (4.3) is satisfied. By definition, $\varrho(t) \leq t\lambda(t) = M\psi(t)$. For $k = 1$ and v with $p(v) > 0$, as $v \leq r$,

$$\begin{aligned} q_1(v, s) &= \mathbf{1}\{s + v > r\} \frac{\varrho(s)}{v + s} \\ &\leq \mathbf{1}\{s + v > r\} \frac{\varrho(s)}{r} \\ &\leq \frac{M \mathbf{1}\{s > r - v\} \psi(s)}{r} \\ &\leq \frac{M}{r} \psi(s \mid V > r - v). \end{aligned}$$

Therefore, we can choose n_1, C_{11} , and ϕ_{11} as in (6.2). For $k > 1$, by (4.3),

$$\begin{aligned} q_k(v, s) &= \mathbf{1}\{s_1 + v > r\} \prod_{i=1}^k \frac{\varrho(s_i)}{v + s_1 + \cdots + s_i} \\ &\leq \frac{\mathbf{1}\{s_1 + v > r\} \varrho(s_1)}{r} \prod_{i=2}^k \frac{\varrho(s_i)}{r + s_2 + \cdots + s_i} \\ &\leq \frac{M^k \psi(s_1 \mid V > r - v)}{r} \bar{\psi}(s), \end{aligned} \tag{6.8}$$

where

$$\bar{\psi}(s) = \prod_{i=2}^k \frac{\psi(s_i)}{r + s_2 + \cdots + s_i}.$$

First, if $N(s) \geq m_k$ then, as $\bar{\psi}(s) \leq r^{-k+1} \prod_{i=2}^k \psi(s_i) \mathbf{1}\{N(s) \geq m_k\}$, and by (6.4) and (6.8),

$$q_k(v, s) \leq \left(\frac{M}{r}\right)^k \psi(s_1 \mid V > r - v) \prod_{i=2}^k \psi(s_i) \mathbf{1}\{N(s) \geq m_k\} = D_{k1} \left(\frac{M}{r}\right)^k \phi_{k1}(v, s).$$

Since $\eta_k \sim \text{Binomial}(k - 1, \alpha)$, by the Markov inequality, for $t \geq 0$,

$$D_{k1} = \mathbf{P}\{\eta_k \geq m_k\} \leq \mathbf{E}(e^{t(\eta_k - m_k)}) = (1 - \alpha + \alpha e^t)^{k-1} e^{-tm_k} \leq (1 - \alpha + \alpha e^t)^{2m_k} e^{-tm_k}.$$

Letting $t = \ln(1/\alpha - 1)$, which is nonnegative since $\alpha \leq \frac{1}{2}$, $D_{k1} \leq [4\alpha(1 - \alpha)]^{m_k}$. Thus, with C_{k1} as in (6.3), $q_k(v, s) \mathbf{1}\{N(s) \geq m_k\} \leq C_{k1} \phi_{k1}(v, s)$.

Second, suppose that $N(s) < m_k$. For each $i, r + s_2 + \cdots + s_i \geq r + n_i \theta_\alpha$, where n_i is the total number of $2 \leq j \leq i$ with $s_j \geq \theta_\alpha$. If the $s_{i_j}, j = 1, \dots, k - 1 - N(s)$, are the members

among s_2, \dots, s_k that are no less than θ_α , then $n_{i_j} = j$. As a result,

$$\begin{aligned} \bar{\psi}(s) &= \prod_{s_i < \theta_\alpha} \frac{1}{r + s_2 + \dots + s_i} \prod_{s_i \geq \theta_\alpha} \frac{1}{r + s_2 + \dots + s_i} \prod_{i=2}^k \psi(s_i) \\ &\leq r^{-N(s)} \prod_{j=1}^{k-1-N(s)} \frac{1}{r + j\theta_\alpha} \prod_{i=2}^k \psi(s_i) \\ &= r^{-(k-1)} \prod_{j=1}^{k-1-N(s)} \frac{1}{1 + j\theta_\alpha/r} \prod_{i=2}^k \psi(s_i) \\ &\leq r^{-(k-1)} \prod_{j=1}^{m_k} \frac{1}{1 + j\theta_\alpha/r} \prod_{i=2}^k \psi(s_i) \\ &\leq \frac{r^{-(k-1)}(r/\theta_\alpha)^{m_k}}{m_k!} \prod_{i=2}^k \psi(s_i), \end{aligned}$$

where the second inequality follows since $k - 1 - N(s) \geq m_k$. Therefore, by (6.5) and (6.8),

$$q_k(v, s) \leq \frac{M^k \psi(s_1 \mid V > r - v)}{r} \frac{r^{-(k-1)}(r/\theta_\alpha)^{m_k}}{m_k!} \prod_{i=2}^k \psi(s_i) = C_{k2} D_{k2} \phi_{k2}(v, s).$$

Since $D_{k2} \leq 1$, (4.3) is satisfied.

From the selection of α , $\sum_k C_{k1} < \infty$. It is straightforward to verify that $\sum_k C_{k2} < \infty$. Therefore, (4.1) is satisfied. Thus, Algorithm 4.2 can be used to sample for φ .

6.2. An application to Vervaat perpetuity

We begin with some background. A Vervaat perpetuity with parameter $c > 0$ is an i.d. random variable Z with Lévy density $ct^{-1}\mathbf{1}\{0 < t \leq 1\}$ (cf. [36]). From the Laplace transform, rZ has Lévy density $ct^{-1}\mathbf{1}\{0 < t \leq r\}$. Efficient sampling methods for Vervaat perpetuities are available [16], [17], [22]. For the algorithm in [17], it was shown that, as $c \rightarrow \infty$, the expected number of iterations is $\exp(c \ln c + O(c))$ as $c \rightarrow \infty$, while, as $c \rightarrow 0$, the expected number of iterations tends to 1. In [16], for the special case $c = 1$, an algorithm was given whose expected number of iterations was no greater than 2.32. These sampling procedures all use the coupling-from-the-past paradigm for Markov chains [29]. Since a Vervaat perpetuity Z with parameter $c > 1$ has the same distribution as the sum of $\lfloor c \rfloor$ independent Vervaat perpetuities with parameter 1 and, when c is noninteger, one independent Vervaat perpetuity with parameter $c - \lfloor c \rfloor$, the results imply that Z can actually be sampled with the expected total number of iterations no greater than $2.32\lfloor c \rfloor + c_0$, where c_0 is a constant.

We show that, by using the design in Proposition 6.1, the expected number of iterations required to sample a Vervaat perpetuity can be arbitrarily close to 1. As a trade-off, we have to sample from an increasingly complicated compound Poisson distribution. However, the treatment of the latter is standard.

The idea is quite simple. First, as noted above, to sample for $ct^{-1}\mathbf{1}\{0 < t \leq 1\}$, we can instead sample for $ct^{-1}\mathbf{1}\{0 < t \leq r\}$ and return ζ/r , where ζ is the sampled value. Here, the artificial parameter r is introduced to control the expected number of iterations. Second, we

decompose $ct^{-1}\mathbf{1}\{0 < t \leq r\} = \varphi(t) + \chi(x)$, where

$$\varphi(t) = ct^{-1}e^{-t}\mathbf{1}\{0 < t \leq r\}, \quad \chi(t) = ct^{-1}(1 - e^{-t})\mathbf{1}\{0 < t \leq r\}.$$

We then apply Proposition 6.1 to φ . It is perhaps not surprising that, when r is large, the expected number of iterations needed is close to 1. Intuitively, this is because $ID(\varphi)$ is so close to $\text{Gamma}(c, 1)$ that, with large probability, a single step to sample from $\text{Gamma}(c, 1)$ would be enough. The Lévy density χ on the other hand gives rise to a compound Poisson distribution. The algorithm we shall verify is the following.

Algorithm 6.1. Set $r = L \max(c^2, 1)$ and φ, χ accordingly, where $L \geq 1$ is a parameter. Set

$$M = c(1 - e^{-r}), \quad \theta = \ln 2 - \ln(1 + e^{-r}).$$

Set $C_0 = 1, C_1 = C_{11} = M/r$, and, for $k \geq 2$, set $m_k = \lfloor k/2 \rfloor$ and

$$C_{k1} = \left(\frac{M}{r}\right)^k, \quad C_{k2} = \frac{(M/r)^k (r/\theta)^{m_k}}{m_k!}, \quad C_k = C_{k1} + C_{k2}.$$

Set $C = \sum_{k=0}^{\infty} C_k$. Set $A_{11} = 1$ and, for $k \geq 2$, letting $\tau \sim \text{Binomial}(k - 1, \frac{1}{2})$, set

$$A_{k1} = P\{\tau \geq m_k\}, \quad A_{k2} = P\{\tau < m_k\}m_k! \left(\frac{\theta}{r}\right)^{m_k}.$$

1. Sample $Z \sim \text{Gamma}(c, 1)$ conditional on $Z \leq r$ and κ such that $P\{\kappa = k\} = C_k/C$.
2. If $\kappa = 0$ then set $X = Z$ and go to step 5. If $\kappa = 1$, set $\eta = 1$. Otherwise, sample $\eta \in \{1, 2\}$ such that $P\{\eta = j\} = C_{kj}/C_\kappa$.
3. Sample $S_1 \sim \text{Exp}(1)$ conditional on $S_1 \in (r - Z, r)$. If $\kappa > 1$ then carry out the following steps.
 - (a) If $\eta = 1$ then sample $\tau \sim \text{Binomial}(\kappa - 1, \frac{1}{2})$ conditional on $\tau \geq m_\kappa$; otherwise, sample $\tau \sim \text{Binomial}(\kappa - 1, \frac{1}{2})$ conditional on $\tau < m_\kappa$.
 - (b) Sample S_2, \dots, S_κ independently such that, for $i \leq \tau, S_i \sim \text{Exp}(1)$ conditional on $S_i < \theta$ and, for $i > \tau, S_i \sim \text{Exp}(1)$ conditional on $\theta \leq S_i < r$. Then randomly permute S_2, \dots, S_κ .
4. Sample $U \sim \text{Unif}(0, 1)$. If

$$U \leq A_{\kappa\eta} \frac{e^Z - 1}{e^r - 1} \prod_{i=1}^{\kappa} \frac{r}{Z + S_1 + \dots + S_i}$$

then set $X = Z + S_1 + \dots + S_\kappa$; otherwise, go back to step 1.

5. Sample $\xi \sim ID(\chi)$. Then return $(X + \xi)/r$.

6.2.1. *Justification.* It suffices to show that steps 1–4 of Algorithm 6.1 indeed sample for φ . The sampling of Z in step 1 follows from the fact that $\text{Gamma}(c, 1)$ has Lévy density $ct^{-1}e^{-t} \times \mathbf{1}\{t > 0\}$. To implement the design in Proposition 6.1, let $\lambda = \varphi$ in (6.1). Then $t\lambda(t) = ce^{-t}\mathbf{1}\{0 < t \leq r\}$, giving $M = \int t\lambda(t) dt = c(1 - e^{-r})$, as in the setup of the algorithm, and $\psi(t) = e^{-t}\mathbf{1}\{0 < t \leq r\}/(1 - e^{-r})$. Note that ψ is the density of $W \sim \text{Exp}(1)$ conditional

on $W \leq r$. By letting $\alpha = \frac{1}{2}$, θ in the setup of the algorithm is the median θ_α of ψ . Moreover, since $2(M/r)\sqrt{\alpha(1-\alpha)} = M/r < c/r \leq 1/L \leq 1$, $(M, \alpha) = (M, \frac{1}{2})$ satisfies the very first assumption of Proposition 6.1. The values of C_k and C_{kj} are set according to (6.2) and (6.3). Then steps 1 and 2 follow from the general procedure in Algorithm 4.2. By (6.2), for any $v \in (0, r)$ and s ,

$$\phi_{11}(v, s) = \frac{\mathbf{1}\{r - v < s < r\}e^{-s}}{e^{-(r-v)} - e^{-r}},$$

and, by (6.4) and (6.5), for $k \geq 2$,

$$\begin{aligned} \phi_{k1}(v, s) &= D_{k1}^{-1} \phi_{11}(v, s_1) \prod_{i=2}^k \frac{e^{-s_i} \mathbf{1}\{0 < s_i \leq r\}}{1 - e^{-r}} \mathbf{1}\{N(s) \geq m_k\}, \\ \phi_{k2}(v, s) &= D_{k2}^{-1} \phi_{11}(v, s_1) \prod_{i=2}^k \frac{e^{-s_i} \mathbf{1}\{0 < s_i \leq r\}}{1 - e^{-r}} \mathbf{1}\{N(s) < m_k\}, \end{aligned}$$

where $D_{k1} = 1 - D_{k2} = P\{\eta_k \geq m_k\}$ with $\eta_k \sim \text{Binomial}(k - 1, \frac{1}{2})$. From Remarks 6.1(ii), it can be seen that step 3 samples from $\phi_{11}(v, \cdot)$, $\phi_{k1}(v, \cdot)$, and $\phi_{k2}(v, \cdot)$. Finally, step 4 follows from (6.6).

6.2.2. *Complexity.* We consider the samplings of X and ξ separately. Let N and T be respectively the number of iterations and amount of time required to generate X . Since $\alpha = \frac{1}{2}$, by (6.7),

$$1 \leq \mathbb{E} N \leq C = \left(1 + \frac{M}{r}\right) \left(\frac{M^2/r^2}{1 - M^2/r^2} + e^{M^2/(r\theta_\alpha)}\right).$$

From the facts that $M/r < c/r \leq 1/L$ and $M^2/(r\theta_\alpha) < c^2/(r\theta_\alpha) < 1/(L\theta_\alpha)$, it is seen that $\mathbb{E} N \rightarrow 1$ as $L \rightarrow \infty$.

By Proposition 4.6, $\mathbb{E} T = \mathbb{E} D \times \mathbb{E} N$, where D is the amount of time required to complete an iteration. Each iteration has to sample (i) one $Z \sim \text{Gamma}(c, 1)$ conditional on $Z \leq r$, (ii) one κ , and (iii) provided that $\kappa \geq 1$, κ observations from various conditional distributions of $\text{Exp}(1)$. Denote by D_i , $i = 1, 2, 3$, the amounts of time required by the samplings. Then $D \approx D_1 + D_2 + D_3$. First, if we use rejection sampling for (i), the expected number of iterations is $m_c = 1/P\{Z \leq r\}$, with $Z \sim \text{Gamma}(c, 1)$. Since $r \geq \max(c^2, 1)$ and both the mean and variance of Z are equal to c , it is not hard to see that m_c is bounded for $c > 0$. On the other hand, the expected amount of time required to sample one Z can be uniformly bounded for c [13]. Then $\mathbb{E} D_1 = O(m_c) = O(1)$. Second, the sampling of κ is standard, with $\mathbb{E} D_2$ uniformly bounded; see Remarks 6.1(i). Finally, since $\zeta \sim \text{Exp}(1)$ conditional on $0 \leq a < \zeta < b \leq \infty$ can be sampled as $-\ln[(1-U)e^{-a} + Ue^{-b}]$ with $U \sim \text{Unif}(0, 1)$, then $\mathbb{E}(D_3 \mid \kappa)$ is proportional to κ , and, thus, $\mathbb{E} D_3$ is of the same order as $C_1/C + \sum_{k \geq 2} k(C_{k1} + C_{k2})/C$, which is $o(1)$ as $r \rightarrow 0$.

Now consider the sampling for χ . For $t \in [0, 1]$, $\chi(t) \leq c$; for $t \in [1, r]$, $\chi(t) \leq c/t$; and, for $t > r$, $\chi(t) = 0$. It is then easy to sample a Poisson process with intensity χ . The sum of the coordinates of the sampled points follows $\text{ID}(\chi)$. The number of points follows a Poisson distribution with mean

$$\int_0^r \chi(t) dt \leq \int_0^1 c dt + \int_1^r ct^{-1} dt = c(1 + \ln r) = c[1 + \ln L + \ln \max(c^2, 1)].$$

Therefore, even for quite large L and c , the amount of time required to sample for χ is manageable.

6.3. Another application

Fix $c > 0$. We now apply Proposition 6.1 to sample for the Lévy density

$$\varphi_a(t) = \frac{ce^{-t}(1-t^a)\mathbf{1}\{0 < t \leq r\}}{t \ln(1/t)}, \quad a > 0.$$

Based on the sampling for φ_a , we can sample for a variety of Lévy densities φ , such as $\varphi(t) = c(1-t^a)\mathbf{1}\{0 < t \leq r\}/[t \ln(1/t)]$, provided that $\varphi - \varphi_a \geq 0$ is integrable. Also, for $b \in (a, \infty]$,

$$\varphi_b(t) - \varphi_a(t) = \frac{ce^{-t}(t^{a-1} - t^{b-1})\mathbf{1}\{0 < t \leq r\}}{\ln(1/t)}$$

is the Lévy density of a compound Poisson distribution. Therefore, if we can sample for φ_a for $0 < a \leq 1$ then we can do so for all $0 < a \leq \infty$. Thus, we shall consider only $0 < a \leq 1$.

For simplicity, let $r \leq e^{-c}$. This condition is restrictive if we directly sample for φ_a with large c . However, note that $ID(\varphi_a)$ can be represented by $Y_1 + \dots + Y_m$, where the Y_i are i.i.d. and each have Lévy density $(c/m)e^{-t}(1-t^a)\mathbf{1}\{0 < t \leq r\}/[t \ln(1/t)]$. By sampling instead for the latter with large enough m , the condition on r becomes mild.

Let W be an i.d. random variable with Lévy density

$$\psi(t) = c\Gamma(t)\mathbf{1}\{0 < t \leq a\}.$$

We will see below that W can be sampled using the sampling of a Vervaat perpetuity. An algorithm to sample for φ_a is given next.

Algorithm 6.2. Set $c_r = c/\ln(1/r)$ and $M = c_r(1 - e^{-r})$. Then set θ, C_k, C_{kj}, C , and A_{kj} exactly as in the algorithm in Section 6.2.

1. Keep sampling (W, Z) until $Z \leq r$, where, conditional on $W, Z \sim \text{Gamma}(W, 1)$. Sample κ such that $P\{\kappa = k\} = C_k/C$.
2. If $\kappa = 0$ then output $X = Z$ and stop. If $\kappa = 1$, set $\eta = 1$. Otherwise, sample $\eta \in \{1, 2\}$ such that $P\{\eta = j\} = C_{kj}/C_\kappa$.
3. Sample (S_1, \dots, S_κ) exactly as in step 3 of Algorithm 6.1.
4. Sample $U \sim \text{Unif}(0, 1)$. If

$$U \leq A_{\kappa\eta} \frac{e^Z - 1}{e^r - 1} \prod_{i=1}^{\kappa} \frac{cr(1 - S_i^a)/\ln(1/S_i)}{v + S_1 + \dots + S_i}$$

then output $X = Z + S_1 + \dots + S_\kappa$ and stop. Otherwise, go back to step 1.

An algorithm to sample W is as follows. Recall that $0 < a \leq 1$.

1. Keep sampling (U, V) until $U \leq e^{-3aV/2}$, where U and V are independent, such that $U \sim \text{Unif}(0, 1)$ and $V > 0$ is a Vervaat perpetuity with parameter c .
2. Sample $\xi \sim ID(\chi)$, with $\chi(t) = \psi(t) - ct^{-1}e^{-3t/2}\mathbf{1}\{0 < t \leq a\}$. Return $aV + \xi$.

6.3.1. *Justification.* First, consider Algorithm 6.2. Let $G(s)$ be a gamma process independent of W with $G(s) \sim \text{Gamma}(s, 1)$, $s > 0$. Denote by g_s the density of $G(s)$. By the properties of subordination (cf. [33]), $G(W)$ is i.d. with Lévy density

$$\int g_s(t)c\Gamma(s)\mathbf{1}\{0 < s \leq a\} ds = c \int_0^a t^{s-1}e^{-t} ds = \frac{ce^{-t}(1-t^a)}{t \ln(1/t)}, \quad t > 0.$$

Since $Z \sim G(W)$, the sampling of Z in step 1 follows. Since $r \leq e^{-c}$, by letting $c_r = c/\ln(1/r)$, we obtain $\varphi_a(t) \leq c_r t^{-1}e^{-t}\mathbf{1}\{0 < t \leq r\}$. Then, to apply Proposition 6.1, set $\lambda(t) = c_r t^{-1}e^{-t}\mathbf{1}\{0 < t \leq r\}$ in (6.1) to obtain $M = c_r(1 - e^{-r})$ and $\psi(t) = e^{-t}\mathbf{1}\{0 < t \leq r\}/(1 - e^{-r})$. Since $M/r = c_r(1 - e^{-r})/r < 1$, we can set $\alpha = \frac{1}{2}$ and obtain the same θ_α and all other constants as in Algorithm 6.1, and execute its steps 2 and 3 without any change. Finally, step 4 follows from (6.6).

To justify the algorithm for W , let $\psi_0(t) = ct^{-1}\mathbf{1}\{0 < t \leq a\}$. From Section 6.2, if V is a Vervaat perpetuity with parameter c then aV has Lévy density ψ_0 . Due to the exponential tilting implemented in step 1, the value of aV it generates has Lévy density $e^{-3t/2}\psi_0(t)$. However, recall that $\Gamma(t)$ is convex on $(0, \infty)$, $\Gamma(t) \leq \Gamma(1) = 1$ for $t \in [1, 2]$, and $\Gamma'(1) = -\gamma$, where $\gamma < \frac{3}{5}$ is Euler’s constant (cf. [2]). Then, as $a \leq 1$, $0 \leq \psi_0(t) - \psi(t) = ct^{-1}[1 - \Gamma(1+t)]\mathbf{1}\{0 < t \leq a\} \leq \gamma t\psi_0(t)$, giving $\psi_0(t) \geq \psi(t) \geq (1 - \gamma t)\psi_0(t) \geq e^{-3t/2}\psi_0(t)$. From this, step 2 follows.

6.3.2. *Complexity.* First, consider Algorithm 6.2. Let N and T respectively be the number of iterations and the amount of time required by the algorithm to generate an output. By (6.7),

$$1 \leq EN \leq \left(1 + \frac{M}{r}\right) \left[\frac{M^2/r^2}{1 - M^2/r^2} + e^{M^2/(r\theta)}\right].$$

As $r \rightarrow 0$, since $M/r = c_r(1 - e^{-r})/r = c(1 - e^{-r})/[r \ln(1/r)] \sim c/\ln(1/r)$ and $r/\theta = r/\ln[2/(1 + e^{-r})] \sim 2$, we obtain $EN \rightarrow 1$ as $r \rightarrow 0$.

By Proposition 4.6, $ET = EN \times ED$, where D is the amount of time required by an iteration in the algorithm. Let D_1, D_2 , and D_3 respectively be the amounts of time to sample (i) one (W, Z) conditional on $Z \leq r$, (ii) one κ , and (iii) provided that $\kappa \geq 1, S_1, \dots, S_\kappa$. Then $D \approx D_1 + D_2 + D_3$. Let Δ be the amount of time taken to sample a pair (W, Z) . Then $ED_1 = E\Delta/P\{Z \leq r\}$. A bound for $E\Delta$ will be given below. In Appendix B, it is shown that, given $c > 0$, as $r \rightarrow 0$, $P\{Z \leq r\} = (1 + o(1))e^{-cA}[\ln(1/r)]^{-c}$, where $A = A(a)$ is a constant. Then $ED_1 \sim e^{cA}[\ln(1/r)]^c E\Delta$. The sampling of κ is standard, with ED_2 uniformly bounded; see Remarks 6.1(i). Finally, as in Section 6.2, ED_3 is of the same order as $C_{11}/C + \sum_{k \geq 2} k(C_{k1} + C_{k2})/C$. It follows that $ED_3 \rightarrow 0$ as $r \rightarrow 0$.

Next, consider the algorithm to sample W . Let T_W be the time taken to generate one output. If $T_V(c)$ denotes the amount of time required to sample a Vervaat perpetuity with parameter c then the expected amount of time required by step 1 of the algorithm is

$$\frac{ET_V(c)}{E(e^{-3aV/2})} = \exp\left\{c \int_0^1 (1 - e^{-3at/2})t^{-1} dt\right\} ET_V(c) \leq e^{3ac/2} ET_V(c),$$

while the expected amount of time required by step 2 is of the same order as

$$\int \chi = c \int_0^a [\Gamma(t) - t^{-1}e^{-3t/2}] dt \leq c \int_0^a t^{-1}(1 - e^{-3t/2}) dt \leq \frac{3ac}{2}.$$

As a result, $ET_W = ET_V(c)/E(e^{-3aV/2}) + O(\int \chi) \leq e^{3ac/2} ET_V(c) + O(ac)$.

We now can obtain a bound for $E \Delta$. In order to sample one pair of (W, Z) , we first sample W and then $Z \sim \text{Gamma}(W, 1)$. Since the expected amount of time to sample from $\Gamma(r, 1)$ can be uniformly bounded for $r > 0$ [13], this gives $E \Delta = E T_W + O(1)$.

7. Conditions for local boundedness

Let φ be a Lévy density with $\int \varphi = \infty$. Recall that in order to obtain the integral series expansion in Theorem 3.1, the density g of $ID(\varphi)$ has to be locally bounded. In some cases, the local boundedness can be directly checked using the explicit expression of g . Following the same argument as in Section 1, if $X \sim ID(\varphi)$ with $\varphi(t) = \lambda(t)\mathbf{1}\{t \leq r\}$, where λ is a Lévy density that gives rise to a locally bounded density, then the density of X is locally bounded in $(0, r]$. This is the case in all the examples in Sections 5 and 6. Thus, to apply Algorithm 4.2 to X , all we need to do is to make sure its density is locally bounded outside $(0, r]$. The following result provides a simple criterion for this.

Proposition 7.1. (Local boundedness of the density.) *Let $0 < M < \infty$. Suppose that $\varrho(t) = t\varphi(t)$ is locally bounded on $(0, M)$ and that g is locally bounded in $(0, r]$ for some $r > 0$.*

- (a) *g is locally bounded on $(0, M)$. Moreover, if ϱ is bounded on $[a, \infty)$ for any $a > 0$ then g is bounded on $[a, \infty)$ for any $a > 0$.*
- (b) *Under the extra assumption that ϱ is continuous on $(0, M)$, g is continuous on $(0, M)$.*

Proof. Without loss of generality, assume that $r < M$.

(a) Fix $0 < \varepsilon < r/2$ with $\int_0^\varepsilon \varrho < r/2$. Fix $c \in (r, M)$. By Lemma 3.1, for any $x \in [r, c]$,

$$\begin{aligned}
 g(x) &= \int_0^x \frac{\varrho(v)}{x} g(x-v) \, dv \leq \frac{1}{r} \int_0^x \varrho(v) g(x-v) \, dv \\
 &\leq \frac{1}{r} \int_0^\varepsilon \varrho(v) g(x-v) \, dv + \frac{1}{r} \sup_{v \in [\varepsilon, c]} \varrho(v) \int_\varepsilon^x g(x-v) \, dv \\
 &\leq \frac{1}{r} \int_0^\varepsilon \varrho(v) g(x-v) \, dv + \frac{1}{r} \sup_{v \in [\varepsilon, c]} \varrho(v).
 \end{aligned}
 \tag{7.1}$$

Given $h > 0$, let $g_h(x) = h^{-1} \int_{x-h}^x g$. Then g_h is continuous on \mathbb{R} and from (7.1), for $x \in [r, c]$,

$$g_h(x) \leq \frac{1}{r} \int_0^\varepsilon \varrho(v) g_h(x-v) \, dv + \frac{1}{r} \sup_{v \in [\varepsilon, c]} \varrho(v).$$

Let $S = \{x \in [r, c]: g_h(x) > g_h(v) \text{ for all } v \in [r-\varepsilon, x)\}$. If $S \neq \emptyset$ then, for any $x \in S$ and $v \in [0, \varepsilon]$, $x-v \geq r-\varepsilon$, and so the above inequality yields

$$g_h(x) \leq \frac{g_h(x)}{r} \int_0^\varepsilon \varrho + \frac{1}{r} \sup_{v \in [\varepsilon, c]} \varrho(v) \leq \frac{g_h(x)}{2} + \frac{1}{r} \sup_{v \in [\varepsilon, c]} \varrho(v).$$

It follows that, for any $x \in [r, c]$, $g_h(x) \leq (2/r) \sup_{v \in [\varepsilon, c]} \varrho(v)$. On the other hand, if $S = \emptyset$ then, for $h < r/6$ and $x \in [r, c]$, $g_h(x) \leq \sup_{v \in [r-\varepsilon, r]} g_h(v) \leq \sup_{v \in [r-\varepsilon-h, r]} g(v) \leq \sup_{v \in [r/3, r]} g(v)$. In any case, for all small $h > 0$ and $x \in [r, c]$,

$$g_h(x) \leq \frac{2}{r} \sup_{v \in [\varepsilon, c]} \varrho(v) + \sup_{v \in [r/3, r]} g(v).$$

Let $h \rightarrow 0$. Since $g_h \rightarrow g$ almost everywhere with respect to the Lebesgue measure, g is bounded on $[r, c] \setminus A$ for some set A that has zero Lebesgue measure. Then, by the integral relation in Lemma 3.1, g is uniformly bounded on $[r, c]$. Since $c \in (r, M)$ is arbitrary and, by assumption, g is locally bounded on $(0, r]$, g is locally bounded on $(0, M)$. Furthermore, from the above inequality, $g_h(x) \leq (2/r) \sup_{v \geq \varepsilon} \varrho(v) + \sup_{v \in [r/3, r]} g(v)$ for $x \geq r$. If ϱ is bounded on $[a, \infty)$ for any $a > 0$ then it is not hard to see that g is bounded on $[a, \infty)$ for any $a > 0$.

(b) Fix $x \in (0, M)$. Given $a \in (0, \min(x, M - x)/2)$, by Lemma 3.1, for any $\delta \in (-a, a)$,

$$\begin{aligned} |(x + \delta)g(x + \delta) - xg(x)| &= \left| \int_0^{x+\delta} \varrho(v)g(x + \delta - v) \, dv - \int_0^x \varrho(v)g(x - v) \, dv \right| \\ &\leq \int_0^{a+\delta} \varrho(v)g(x + \delta - v) \, dv + \int_0^a \varrho(v)g(x - v) \, dv \\ &\quad + \left| \int_{a+\delta}^{x+\delta} \varrho(v)g(x + \delta - v) \, dv - \int_a^x \varrho(v)g(x - v) \, dv \right| \\ &\leq 2 \sup_{|u-x| \leq a} g(u) \int_0^{2a} \varrho + \int_a^x |\varrho(v + \delta) - \varrho(v)|g(x - v) \, dv. \end{aligned}$$

By dominated convergence,

$$\overline{\lim}_{\delta \rightarrow 0} |(x + \delta)g(x + \delta) - xg(x)| \leq 2 \sup_{|u-x| \leq a} g(u) \int_0^{2a} \varrho \leq 2 \sup_{u \in [x/2, (x+M)/2]} g(u) \int_0^{2a} \varrho.$$

Since a is arbitrary, the limit is 0. As $x > 0$, this implies that g is continuous at x .

Appendix A

In Section 5.2, for fixed $\alpha > 0$, we made the following claims.

- (C1) $E D_1(r)$ is uniformly bounded for $r > 0$, with $D_1(r)$ the amount of time required to sample $Z \sim \text{ID}(\lambda)$ conditional on $Z \leq r$, with $\lambda(t) = t^{-\alpha-1} \mathbf{1}\{t > 0\}$.
- (C2) $E D_2(r)$ is uniformly bounded for $r > 0$, with $D_2(r)$ the amount of time required to sample κ from the probability mass function $p_k = b_k / \sum_{i \geq 0} b_i$, $k \geq 0$, where $b_0 = 1$, and, for $k \geq 1$, $b_k = \alpha \theta^k \Gamma(k\alpha) / \Gamma(k)$ with $\theta = \Gamma(1 - \alpha) / (r^\alpha \alpha)$.
- (C3) $E D_3(k, z)$ is uniformly bounded for $k \geq 1$ and $z \in (0, 1]$, where $D_3(k, z)$ denotes the amount of time required to sample $\zeta \sim \text{Beta}(k\alpha, 1 - \alpha)$ conditional on $\zeta \leq z$.
- (C4) $E N \asymp \exp\{r^{-\alpha} / \alpha\}$ as $r \rightarrow 0$, where N is the number of iterations of the algorithm in the section.

As the proof of (C1) depends on part of the proof of (C4), we first establish (C2)–(C4).

Proof of (C2). Define $p_k = 0$ for $k < 0$. All we need is p_k to be log-concave, i.e. $p_k^2 \geq p_{k-1} p_{k+1}$ for any k , as it allows rejection sampling of p_k with no more than five iterations on average [14], [19]. One issue here is the difficulty to evaluate the normalizing constant $C = C(r) = \sum_k b_k$. However, it is relatively easy to find $C_i = C_i(r)$, $i = 1, 2$, such that $C_1 < C < C_2$ and $0 < \inf_r (C_1/C) < \sup_r (C_2/C) < \infty$. Then the rejection sampling can be modified by using $b_k/C_2 < p_k < b_k/C_1$, while still achieving a uniformly bounded expected number of iterations. For example, the dominating function in Section 3 of [14] can be modified to $(b_m/C_1) \min(1, e^{1-(b_m/C_2)(|x|-1/2)})$, where $m = \arg \max_k b_k$.

Let $f(x) = \Gamma(\alpha x) / \Gamma(x)$, $x > 0$. Since $\Gamma(x) \sim 1/x$ as $x \rightarrow 0+$, f can be continuously extended to 0 with $f(0) = 1/\alpha$, yielding $b_k = \alpha \theta^k f(k)$ for all $k \geq 0$. Thus, to show that p_k is log-concave, it suffices to show that $(\ln f)''(x) < 0$ for $x > 0$. We have $(\ln f)''(x) = \alpha^2 \psi(\alpha x) - \psi(x)$, where ψ is a polygamma function [2]. From the integral representation of ψ ,

$$(\ln f)''(x) = \int_0^\infty \frac{\alpha^2 t e^{-\alpha x t}}{1 - e^{-t}} dt - \int_0^\infty \frac{t e^{-x t}}{1 - e^{-t}} dt = \int_0^\infty t e^{-x t} \left(\frac{1}{1 - e^{-t/\alpha}} - \frac{1}{1 - e^{-t}} \right) dt,$$

which is indeed negative since $\alpha < 1$.

Proof of (C3). The particular parametrization of the beta distributions is not important. Let $T_{a,b}(r)$ denote the amount of time taken to sample $\zeta \sim \text{Beta}(a, b)$ conditional on $\zeta \leq r$. Clearly, $D_3(k, r) = T_{k\alpha, 1-\alpha}(r)$. Given a_0 and $b \in (0, 1)$, consider the following procedure.

- Fix $c \in (0, a_0)$. If $r \geq 1 - c/a$ then keep sampling $\zeta \sim \text{Beta}(a, b)$ until $\zeta \leq r$; otherwise, keep sampling U, V i.i.d. $\sim \text{Unif}(0, 1)$ until $V \leq [(1 - r)/(1 - \zeta)]^{1-b}$, where $\zeta = rU^{1/a}$.

We show that $E T_{a,b}(r)$ is uniformly bounded for $a \geq a_0$ and $r \in (0, 1]$ by using this procedure. First, it is clear that $P\{\zeta \leq 1 - c/a\}$ is a positive continuous function in $a \geq a_0$. As $a \rightarrow \infty$, $B(a, b) = \Gamma(a)\Gamma(b) / \Gamma(a + b) \sim \Gamma(b)a^{-b}$, giving

$$\begin{aligned} P\left\{\zeta \leq 1 - \frac{c}{a}\right\} &\sim \frac{a^b}{\Gamma(b)} \int_0^{1-c/a} x^{a-1} (1-x)^{b-1} dx \\ &= \frac{1}{\Gamma(b)} \int_c^\theta \left(1 - \frac{t}{a}\right)^{a-1} t^{b-1} dt \\ &\rightarrow \frac{1}{\Gamma(b)} \int_c^\infty e^{-x} t^{b-1} dt \\ &> 0. \end{aligned}$$

Thus, letting $p_0 = \inf_{a \geq a_0} P\{\zeta \leq 1 - c/a\}$, we have $p_0 > 0$. If $r \geq 1 - c/a$ then in each iteration of the rejection sampling the probability of acceptance is $P\{\zeta \leq r\} \geq P\{\zeta \leq 1 - c/a\} \geq p_0$, which leads to $E T_{a,b}(r) \leq p_0^{-1} E T_{a,b}(1)$. It is known that, using suitable rejection sampling, $\sup_{a,b} E T_{a,b}(1) < \infty$ [13]. As a result, $E T_{a,b}(r)$ is bounded for $a \geq a_0$ and $r \geq 1 - c/a$. On the other hand, if $r \in (0, 1 - c/a)$ then, as the density of $rU^{1/a}$ is $\mathbf{1}\{0 < t \leq r\} a t^{a-1} / r^a$, the rejection sampling indeed generates $\zeta \sim \text{Beta}(a, b)$ conditional on $\zeta \leq r$. In each iteration, the probability of acceptance is

$$\begin{aligned} (1 - r)^{1-b} \int_0^1 (1 - r t^{1/a})^{b-1} dt &\geq (1 - r)^{1-b} \int_{e^{-1}}^1 (1 - r t^{1/a})^{b-1} dt \\ &\geq (1 - r)^{1-b} (1 - e^{-1}) (1 - r e^{-1/a})^{b-1} \\ &\geq (1 - e^{-1}) \left(\frac{c/a}{1 - (1 - c/a)e^{-1/a}} \right)^{1-b}. \end{aligned}$$

The right-hand side is a positive continuous function of $a \geq a_0$ and can be shown to converge to $(1 - e^{-1})[c/(c + 1)]^{1-b} > 0$ as $a \rightarrow \infty$. This implies that the probability of acceptance is uniformly bounded away from 0 for $a \geq a_0$ and $r \in (0, c/a)$. As a result, $E T_{a,b}(r)$ is uniformly upper bounded.

Proof of (C4). The proof of the asymptotic of

$$E N = \exp\left\{\frac{r^{-\alpha}}{\alpha}\right\} P\{Y \leq r\} \left(1 + \int_0^\infty e^{z-r[\alpha z/\Gamma(1-\alpha)]^{1/\alpha}} dz\right)$$

as $r \rightarrow 0$ is an exercise in the saddle point method [23]. Define $\varepsilon = r[\alpha/\Gamma(1-\alpha)]^{1/\alpha}$ and $M = (\alpha/\varepsilon)^{\alpha/(1-\alpha)}$. By the change of variable $z = M(1-t)$,

$$\begin{aligned} \int_0^\infty e^{z-r[\alpha z/\Gamma(1-\alpha)]^{1/\alpha}} dz &= \int_0^\infty e^{z-\varepsilon z^{1/\alpha}} dz \\ &= M \int_{-\infty}^1 e^{M(1-t)-\varepsilon M^{1/\alpha}(1-t)^{1/\alpha}} dt \\ &= M \int_{-\infty}^1 e^{M[(1-t)-\alpha(1-t)^{1/\alpha}]} dt. \end{aligned}$$

The function $f(t) = 1-t-\alpha(1-t)^{1/\alpha}$ is smooth on $(-\infty, 1)$ and maximized uniquely at $t = 0$ with $f(0) = 1-\alpha$ and $f''(0) = 1-1/\alpha < 0$. As $r \rightarrow 0$, $M \rightarrow \infty$. Then, by the saddle point method,

$$\begin{aligned} M \int_{-\infty}^1 e^{M[(1-t)-\alpha(1-t)^{1/\alpha}]} dt &= M[1 + o(1)] \sqrt{\frac{2\pi}{M|f''(0)|}} e^{Mf(0)} \\ &= [1 + o(1)] \sqrt{\frac{2\pi M}{1/\alpha - 1}} e^{M(1-\alpha)}. \end{aligned}$$

Next, since Y has Lévy density $\lambda(t) = t^{-\alpha-1}\mathbf{1}\{t > 0\}$, $Y \sim [\Gamma(1-\alpha)/\alpha]^{1/\alpha} S$, where S has Laplace transform $\exp\{-\theta^\alpha\}$. Then, by [10],

$$\begin{aligned} P\{Y \leq r\} &= P\{S \leq \varepsilon\} \\ &= \frac{1}{\pi} \int_0^\pi \exp\{-\varepsilon^{-\alpha/(1-\alpha)} h(t)\} dt \\ &= \frac{1}{\pi} \int_0^\pi \exp\{-M\alpha^{-\alpha/(1-\alpha)} h(t)\} dt, \end{aligned}$$

where h is a function defined as

$$h(0) = (1-\alpha)\alpha^{\alpha/(1-\alpha)}, \quad h(t) = \frac{\sin((1-\alpha)t)[\sin(\alpha t)]^{\alpha/(1-\alpha)}}{(\sin t)^{1/(1-\alpha)}} \mathbf{1}\{0 < t < \pi\}, \quad t \neq 0.$$

By concavity, $\sin(\alpha t) > a \sin(t)$ for $a \in (0, 1)$ and $t \in (0, \pi)$. Thus, h is minimized uniquely at $t = 0$. Using $\sin t = t - t^3/3! + \dots$, it is straightforward to check that $h \in C^\infty[0, \pi)$, $h'(0+) = 0$, and $h''(0+) > 0$. Noting that the integral for $P\{Y \leq r\}$ is over an interval to the right of 0, by the saddle point method,

$$P\{Y \leq r\} = [1 + o(1)] \sqrt{\frac{\alpha^{\alpha/(1-\alpha)}}{2\pi M h''(0+)}} e^{-M\alpha^{-\alpha/(1-\alpha)} h(0)}.$$

Since $\alpha^{-\alpha/(1-\alpha)} h(0) = 1-\alpha$, we then obtain $E N \asymp \exp(r^{-\alpha}/\alpha)$, as claimed.

Proof of (C1). Let S be as above. By scaling, it suffices to show that the expected time to sample S conditional on $S \leq r$ is uniformly bounded for $r > 0$. From [10], S can be embedded into a random vector (ξ, S) such that $\xi \sim \text{Unif}(0, 1)$ and, conditional on ξ , $P\{S \leq r \mid \xi\} = \exp\{-r^{-\alpha/(1-\alpha)}h(\xi)\}$ and $S \sim [h(\xi)/W]^{(1-\alpha)/\alpha}$, with $W \sim \text{Exp}(1)$. Therefore, to sample S conditional on $S \leq r$, we can first sample ξ conditional on $S \leq r$, and then, given ξ , sample $W \sim \text{Exp}(1)$ and set

$$S = r \left[\frac{h(\xi)}{h(\xi) + Wr^{\alpha/(1-\alpha)}} \right]^{(1-\alpha)/\alpha}.$$

To establish (C1), it suffices to show that K_r , the expected number of iterations required to sample ξ conditional on $S \leq r$, is bounded for $r > 0$. By Bayes formula, conditional on $S \leq r$, the density of ξ is in proportion to $q_r(t) := \exp\{-r^{-\alpha/(1-\alpha)}h(t)\}$. It is easy to see that, using rejection sampling, we can have $\sup_{r \geq r_0} K_r < \infty$ for any fixed $r_0 > 0$. Therefore, we need to show only that $\sup_{r \leq r_0} K_r < \infty$ for small $r_0 > 0$. From the proof of (C4), there exists $c > 1/h''(0+)$ such that $h(t) \geq h(0) + t^2/(2c)$ for all $t \in (0, \pi)$. Consider the following procedure: keep sampling $X \sim N(0, cr^{\alpha/(1-\alpha)})$ and $U \sim \text{Unif}(0, 1)$ until $U \leq \exp\{-r^{-\alpha/(1-\alpha)}[h(|X|) - h(0) - X^2/(2c)]\}$ and then return $\xi = |X|$. This procedure samples ξ conditional on $S \leq r$. The probability of acceptance of each iteration is

$$\begin{aligned} & \frac{1}{\sqrt{2\pi cr^{\alpha/(1-\alpha)}}} \int_{-\infty}^{\infty} \exp\{-r^{-\alpha/(1-\alpha)}[h(|t|) - h(0)]\} dt \\ &= \frac{1}{\sqrt{2\pi cr^{\alpha/(1-\alpha)}}} [1 + o(1)] \sqrt{\frac{2\pi r^{\alpha/(1-\alpha)}}{h''(0+)}} \\ &\rightarrow \frac{1}{\sqrt{ch''(0+)}} \\ &> 0 \quad \text{as } r \rightarrow 0, \end{aligned}$$

where the last three lines are due to the saddle point method. It is then easy to show that, using the above procedure, K_r is uniformly bounded for all small r .

Appendix B

In Section 6.3 we made the following claim. Let $\lambda(t) = ce^{-t}(1 - t^a)\mathbf{1}\{t > 0\}/[t \ln(1/t)]$ with $c > 0$ and $a > 0$, and let $Z \sim \text{ID}(\lambda)$. Then, for a constant $A = A(a)$, as $r \rightarrow 0$ while c is fixed,

$$P\{Z \leq r\} = [1 + o(1)]e^{-cA} \left[\ln\left(\frac{1}{r}\right) \right]^{-c}.$$

To prove the claim, for any $r < e^{-1}$, $Z \sim \eta_r + X_r + Y$, where η_r , X_r , and Y are independent i.d. random variables with Lévy densities $\mathbf{1}\{t \leq r\}\lambda(t)$, $\mathbf{1}\{r \leq t < e^{-1}\}\lambda(t)$, and $\mathbf{1}\{t > e^{-1}\}\lambda(t)$, respectively. We seen that $P\{Z \leq r\} = P\{\eta_r \leq r\} P\{X_r = 0\} P\{Y = 0\}$. First, by the Markov inequality, for any $s > 0$,

$$\begin{aligned} P\{\eta_r > r\} &\leq E(e^{s\eta_r/r-s}) \\ &= \exp\left\{ \int_0^r (e^{st/r} - 1) \frac{ce^{-t}(1 - t^a)}{t \ln(1/t)} dt - s \right\} \\ &\leq \exp\left\{ \frac{c}{\ln(1/r)} \int_0^r t^{-1}(e^{st/r} - 1) dt - s \right\}. \end{aligned}$$

By the convexity of the exponential function,

$$P\{\eta_r > r\} \leq \exp\left\{\frac{c}{\ln(1/r)} \int_0^r \frac{s}{r} e^{st/r} dt - s\right\} = \exp\left\{\frac{c(e^s - 1)}{\ln(1/r)} - s\right\}.$$

Letting $r \rightarrow 0$ followed by $s \rightarrow \infty$, we see that $P\{\eta_r > r\} \rightarrow 0$ and, hence, $P\{\eta_r \leq r\} \rightarrow 1$.

Next, by the property of the Poisson process,

$$P\{X_r = 0\} = \exp\left\{-\int_r^{e^{-1}} \lambda(t) dt\right\} = e^{cI(r)} \exp\left\{-\int_r^{e^{-1}} \frac{c dt}{t \ln(1/t)}\right\} = e^{cI(r)} \left[\ln\left(\frac{1}{r}\right)\right]^{-c},$$

where

$$I(r) = \int_r^{e^{-1}} \frac{1 - e^{-t}(1 - t^a)}{t \ln(1/t)} dt \rightarrow I(0) = \int_0^{e^{-1}} \frac{1 - e^{-t}(1 - t^a)}{t \ln(1/t)} dt < \infty.$$

Finally, $P\{Y = 0\} = e^{-cJ}$, where $J = \int_{e^{-1}}^{\infty} \lambda < \infty$. Combining the results, we obtain $P\{Z \leq r\} \sim e^{c(J-I(0))} [\ln(1/r)]^{-c}$, as claimed.

Acknowledgement

The author would like to thank an anonymous referee for many helpful and stimulating suggestions, and for pointing out several important references.

References

- [1] AALEN, O. O. (1992). Modelling heterogeneity in survival analysis by the compound Poisson distribution. *Ann. Appl. Prob.* **2**, 951–972.
- [2] ABRAMOWITZ, M. AND STEGUN, I. A. (1964). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables* (Nat. Bureau Stand. App. Math. Ser. **55**). Government Printing Office, Washington, DC.
- [3] ASMUSSEN, S. AND ROSIŃSKI, J. (2001). Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Prob.* **38**, 482–493.
- [4] BERTOIN, J. (1996). *Lévy Processes* (Camb. Tracts Math. **121**). Cambridge University Press.
- [5] BERTOIN, J. AND YOR, M. (2002). The entrance laws of self-similar Markov processes and exponential functionals of Lévy processes. *Potential Anal.* **17**, 389–400.
- [6] BONDESSON, L. (1981). Classes of infinitely divisible distributions and densities. *Z. Wahrscheinlichkeitsth.* **57**, 39–71.
- [7] BREIMAN, L. (1968). *Probability*. Addison-Wesley, Reading, MA.
- [8] BRIX, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. Appl. Prob.* **31**, 929–953.
- [9] CABALLERO, M. E., PARDO, J. C. AND PÉREZ, J. L. (2010). On Lamperti stable processes. *Prob. Math. Statist.* **30**, 1–28.
- [10] CHAMBERS, J. M., MALLOWS, C. L. AND STUCK, B. W. (1976). A method for simulating stable random variables. *J. Amer. Statist. Assoc.* **71**, 340–344.
- [11] COVO, S. (2009). One-dimensional distributions of subordinators with upper truncated Lévy measure, and applications. *Adv. Appl. Prob.* **41**, 367–392.
- [12] DEVROYE, L. (1986). An automatic method for generating random variates with a given characteristic function. *SIAM J. Appl. Math.* **46**, 698–719.
- [13] DEVROYE, L. (1986). *Nonuniform Random Variate Generation*. Springer, New York.
- [14] DEVROYE, L. (1987). A simple generator for discrete log-concave distributions. *Computing* **39**, 87–91.
- [15] DEVROYE, L. (2001). Simulating perpetuities. *Methodology Comput. Appl. Prob.* **3**, 97–115.
- [16] DEVROYE, L. AND FAWZI, O. (2010). Simulating the Dickman distribution. *Statist. Prob. Lett.* **80**, 242–247.
- [17] FILL, J. A. AND HUBER, M. L. (2010). Perfect simulation of Vervaat perpetuities. *Electron. J. Prob.* **15**, 96–109.
- [18] GLASSERMAN, P. (2004). *Monte Carlo Methods in Financial Engineering* (Appl. Math. **53**). Springer, New York.
- [19] HÖRMANN, W., LEYDOLD, J. AND DERFLINGER, G. (2004). *Automatic Nonuniform Random Variate Generation*. Springer, Berlin.

- [20] HOUGAARD, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73**, 387–396.
- [21] KANTER, M. (1975). Stable densities under change of scale and total variation inequalities. *Ann. Prob.* **3**, 697–707.
- [22] KENDALL, W. S. AND THÖNNES, E. (2004). Random walk CFTP. Tech. Rep., Department of Statistics, University of Warwick.
- [23] KOLASSA, J. E. (1997). *Series Approximation Methods in Statistics* (Lecture Notes Statist. **88**), 2nd edn. Springer, New York.
- [24] KUZNETSOV, A., KYPRIANOU, A. E., PARDO, J. AND VAN SCHAİK, K. (2011). A Wiener-Hopf Monte Carlo simulation technique for Lévy processes. *Ann. Appl. Prob.* **21**, 2171–2190.
- [25] KYPRIANOU, A. E., PARDO, J. C. AND RIVERO, V. (2010). Exact and asymptotic n -tuple laws at first and last passage. *Ann. Appl. Prob.* **20**, 522–564.
- [26] LAMPERTI, J. (1972). Semi-stable Markov processes. I. *Z. Wahrscheinlichkeitsth.* **22**, 205–225.
- [27] LIU, J. S. (2001). *Monte Carlo Strategies in Scientific Computing*. Springer, New York.
- [28] PAPADIMITRIOU, C. H. (1994). *Computational Complexity*. Addison-Wesley, Reading, MA.
- [29] PROPP, J. G. AND WILSON, D. B. (1996). Exact sampling with coupled Markov chains and applications to statistical mechanics. *Random Structures Algorithms* **9**, 223–252.
- [30] REYNOLDS, D. S. AND SAVAGE, I. R. (1971). Random wear models in reliability theory. *Adv. Appl. Prob.* **3**, 229–248.
- [31] RUDIN, W. (1987). *Real and Complex Analysis*, 3rd edn. McGraw-Hill, New York.
- [32] SAMORODNITSKY, G. AND TAQQU, M. S. (1994). *Stable Non-Gaussian Random Processes*. Chapman & Hall, New York.
- [33] SATO, K.-I. (1999). *Lévy Processes and Infinitely Divisible Distributions* (Camb. Stud. Adv. Math. **68**). Cambridge University Press.
- [34] STEUTEL, F. W. (1970). *Preservation of Infinite Divisibility under Mixing and Related Topics* (Math. Centre Tracts **33**). Mathematisch Centrum, Amsterdam.
- [35] STEUTEL, F. W. (1979). Infinite divisibility in theory and practice. *Scand. J. Statist.* **6**, 57–64.
- [36] VERVAAT, W. (1979). On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv. Appl. Prob.* **11**, 750–783.