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 variables are 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 $X \le r$ 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, if an i.d. random variable has an infinite Lévy measure on $(0, \infty)$, then with probability one, 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 what nonnegative i.d. random variables can be sampled exactly. As a matter of fact, there are many such random variables, the most familiar ones including 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 have not only explicit density functions, but also explicit Lévy measures. Generally speaking, in view of the

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

Lévy-Khintchine formula, it is natural to specify an i.d. random variables 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 we wish to sample a nonnegative i.d. random variable X with an upper truncated Lévy density $\lambda(x)\mathbf{1}\{x \le r\}$, 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 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 $Y \ge 0$ is an i.d. random variable with Lévy density λ, then X *conditional* on $X \le r$ and Y *conditional* on $Y \le r$ 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, $\mathrm{P}\{Z = 0\} > 0$ and $\mathrm{P}\{X \le x \,|\, X \le r\} = \mathrm{P}\{X \le x, Z = 0 \,|\, X \le r, Z = 0\}$. Since for any $x \le r$, the events $\{X \le x,\, Z = 0\}$ and $\{Y \le x\}$ are the same, we get $\mathrm{P}\{X \le x \,|\, X \le r\} = \mathrm{P}\{Y \le x \,|\, Y \le r\}$. Consequently, if we can choose λ appropriately so that we know how to sample Y, then we can sample X conditional on $X \le r$ by simpling Y conditional on $Y \le r$, using, say, the rejection method (cf. [13, 18, 19, 27]).

The above observation is crucial, however, it alone cannot lead to the sampling of X. There are two issues. First, $\mathrm{P}\{X \le r\}$ 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 only consists of 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, ∞) which 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 $(0, r] \ne \emptyset$, φ is identical to a Lévy density λ which we know how to sample for, and $\chi \ge 0$ is integrable. Indeed, for φ, we only need $\varphi(t) = [1 + O(t)]\lambda(t)$ as $t \to 0$. 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, 15]). Unlike in the procedures of the above-cited works, we do not rely on explicit approximations to distributions; however, we do require we know how to sample for some i.d. distribution using the distributions at hand. That is, 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 formulas for the densities. Therefore, our procedure is complementary to available sampling procedures, such as those in [10, 13, 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 $\varphi(t) = ct^{-\alpha-1}\mathbf{1}\{0 < t \le r\}$ 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 my best 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 Lévy density $\varphi(t) = \psi(t)(\mathrm{e}^t-1)^{-1-\alpha}\mathbf{1}\{t > 0\}$ with $\alpha \in (0,1)$ and $\psi(t) = 1+O(t)$ as $t \to 0$. When $\psi(t) = \mathrm{e}^{\beta t}$ 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, 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 $c\mathrm{e}^{-t}(1-t^a)\mathbf{1}\{0 < t \le r\}/[t\ln(1/t)]$. 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 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 $\varphi \ge 0$ with support in $[0,\infty)$ and $\int(1\wedge t)\varphi(t)\,\mathrm{d}t < \infty$ (cf. [4, 33]). By i.d. random variable with Lévy density φ, we mean specifically a random variable X with Laplace transform $\mathrm{E}(\mathrm{e}^{-\theta X}) = \int(\mathrm{e}^{-\theta t}-1)\varphi(t)\,\mathrm{d}t$, $\theta > 0$. We denote by $\mathrm{ID}(\varphi)$ the distribution with the Laplace transform. We call X strictly positive if $\mathrm{P}\{X > 0\} = 1$. Under the setup, $X \sim \mathrm{ID}(\varphi)$ 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 & 27.10]. On the other hand, if $\int\varphi < \infty$, then $X \sim \mathrm{ID}(\varphi)$ is compound Poisson, with $\mathrm{P}\{X = 0\} > 0$.

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

2.2. Decomposition and exponential tilting

Sampling for a Lévy density can often be built upon the sampling for other Lévy densities. Suppose we know how to sample for a Lévy density φ. If a Lévy density φ_1 can be decomposed into $\varphi+\chi$, with $\chi \ge 0$ being integrable, then we can sample for φ_1 by generating $X+\xi$, where

$X \sim \mathrm{ID}(\varphi)$ and $\xi \sim \mathrm{ID}(\chi)$ 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 $\mathrm{ID}(\chi)$ (cf. [13, 18]).

On the other hand, if $\varphi_1 \le \varphi$, 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 for some $b > 0$, $\varphi_1(t) = \mathrm{e}^{-bt}\varphi(t)$, 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, 20]). We state the result as follows for later use.

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

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

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

3. An integral series expansion of density

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

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 for some $M > r$, g is locally bounded on $(0, M)$. Define*

$$h(v, x) = \mathbf{1}\{x > 0\}\, \varrho(x - v)/x \tag{3.1}$$

and for each $k \ge 1$,

$$h_k(v, x) = \int \mathbf{1}\{v_1 > r\}\, h(v, v_1)h(v_1, v_2) \cdots h(v_k, x)\, \mathrm{d}v_1 \cdots \mathrm{d}v_k. \tag{3.2}$$

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

$$g(x) = \int_0^r g(v)\, \mathrm{d}v \left[h(v, x) + \sum_{k=1}^{\infty} h_k(v, x) \right]. \tag{3.3}$$

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

Proof. We start with the following known result ([34], Corollary 4.2.2; [33], Theorem 51.1).

Lemma 3.1. *For any $x > 0$, $g(x) = \int_0^x g(v)h(v, x)\, \mathrm{d}v$.*

Equipped with the lemma, we first show that given $x \in (r, M)$, for each $n \geq 1$,

$$g(x) = \int_0^r g(v)\,\mathrm{d}v \left[h(v,x) + \sum_{k=1}^{n} h_k(v,x) \right] + R_n(x), \tag{3.4}$$

$$\text{with} \quad R_n(x) = \int \mathbf{1}\{v > r\}\, g(v) h(v,v_1) h(v_1,v_2) \cdots h(v_n,x)\,\mathrm{d}v\,\mathrm{d}v_1 \cdots \mathrm{d}v_n.$$

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

$$g(x) = \int_0^\infty g(v)h(v,x)\,\mathrm{d}v = \int_0^r g(v)h(v,x)\,\mathrm{d}v + \int_r^\infty g(v)h(v,x)\,\mathrm{d}v. \tag{3.5}$$

For $g(v)$ in the last integral, as $v > r$, the expansion (3.5) can be iterated once to yield

$$g(x) = \int_0^r g(v)h(v,x)\,\mathrm{d}v + \int_r^\infty \left[\int_0^r g(w)h(w,v)\,\mathrm{d}w + \int_r^\infty g(w)h(w,v)\,\mathrm{d}w \right] h(v,x)\,\mathrm{d}v.$$

Make some changes in symbol and regroup the integrals to get

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

which shows (3.4) for $n = 1$. In general, for $g(v)$ in the integral expression of $R_n(x)$, as $v \geq r$, we can iterate the expansion (3.5) to get

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

The right hand side is exactly $\int_0^r g(w)h_{n+1}(w,x)\,\mathrm{d}w + R_{n+1}(x)$. Then by induction, (3.4) holds.

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

$$R_n(x) \leq \sup_{v\in[r,x]} g(v) \times I_n(x) \tag{3.6}$$

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

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

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

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

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, ... be i.i.d. $\sim \psi$. Then by the above inequality and Markov inequality, for any $t > 0$,

$$\begin{aligned} I_n(x) &\le (\mu/r)^{n+1}\mathrm{P}\{\xi_1 + \cdots + \xi_{n+1} < x - r\} \\ &\le (\mu/r)^{n+1}\mathrm{E}[\mathrm{e}^{t(x-r-\xi_1-\cdots-\xi_{n+1})}] = \mathrm{e}^{t(x-r)}[\mu\mathrm{E}(\mathrm{e}^{-t\xi})/r]^{n+1}. \end{aligned}$$

Since $\mathrm{E}\mathrm{e}^{-t\xi} \to 0$ as $t \to \infty$, we can find $t > 0$ such that $\mu\mathrm{E}(\mathrm{e}^{-t\xi})/r < 1$. Fixing such t and letting $n \to \infty$, it is seen $I_n(x) \to 0$ as $n \to \infty$. By assumption, $\sup_{v\in[r,x]} g(v) < \infty$, so by (3.6), $R_n(x) \to 0$. This together with (3.4) and monotone convergence yields (3.3).

3.2. Formulation toward rejection sampling

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

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

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

$$p(x) = a^{-1}g(x)\mathbf{1}\{x \le r\}, \tag{3.7}$$

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

$$g(x) = a \int p(v)\,\mathrm{d}v \left[\delta(x-v) + h(v,x)\mathbf{1}\{x > r\} + \sum_{k=1}^{\infty} h_k(v,x)\right]. \tag{3.8}$$

The question is how to use (3.8) to sample from g. In the expansion, δ and h should be easy to handle; however, h_1, h_2, ... are defined by integrals. In general, for sampling, it is desirable not to evaluate integrals. With this in mind, we shall first consider densities of the form

$$g(x) = a \int p(v)\,\mathrm{d}v \left[\sum_i \int q_i(v,x,w_i)\,\nu_i(\mathrm{d}w_i)\right] \quad \text{for some constant } a > 0, \tag{3.9}$$

where the sum on the right hand side has at most countably many terms, $q_i \ge 0$ are known functions, and ν_i are σ-finite measures on some measurable spaces that may be different from each other. Besides including (3.8) as a special case, the expansion in (3.9) provides more flexibility for sampling. For example, if the integral that defines h_k in (3.8) is complicated, then we may consider using (3.9) to reformulate h_k as the sum of several integrals over disjoint regions that are easier to handle. This perspective will be incorporated in 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 π is specified as $\pi \propto q$, where $q \ge 0$ 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, such that ϕ can be sampled and $q(x) \le C\phi(x)$ for all x. Then the sampling for π proceeds as follows.

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

We now turn to the density in (3.9). Suppose we can find constants $C_i \ge 0$ and instrumental functions $\phi_i(\cdot,\cdot,\cdot) \ge 0$, such that for all v with $p(v) > 0$ and i,

$$C := \sum_i C_i < \infty, \quad q_i(v,\cdot,\cdot) \le C_i\phi_i(v,\cdot,\cdot), \quad \text{and}$$
$$\phi_i(v,\cdot,\cdot) \text{ is a probability density with respect to } \ell \times \nu_i,$$

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

Algorithm 4.1. (*Rejection sampling for densities of the form (3.9).*)

1. Sample $Z \sim p$ and κ independently, with $\mathrm{P}\{\kappa = i\} = C_i/C$.
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 $U \sim \mathrm{Unif}(0,1)$. If $C_\kappa U \phi_\kappa(Z,X,W) \le q_\kappa(Z,X,W)$, then output X and stop, else 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 (3.9).*

Proof. First, by integrating (3.9) over x and Fubini Theorem,

$$\begin{aligned} 1 = \int g(x)\,\mathrm{d}x &= a \int p(v)\,\mathrm{d}v \left[\sum_i \int q_i(v,x,w_i)\,\mathrm{d}x\,\nu_i(\mathrm{d}w_i)\right] \\ &\le a \int p(v)\,\mathrm{d}v \left[\sum_i C_i \int \phi_i(v,x,w_i)\,\mathrm{d}x\,\nu_i(\mathrm{d}w_i)\right] = aC. \end{aligned}$$

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,

$$\begin{aligned} &\mathrm{P}\{X \in A \text{ is output}\} \\ &= \int \sum_i \mathrm{P}\{Z \in \mathrm{d}v,\ \kappa = i\}\,\mathrm{P}\left\{X \in A,\ U \le \frac{q_i(v,X,W)}{C_i\phi_i(v,X,W)} \,\middle|\, Z = v,\ \kappa = i\right\} \\ &= \int p(v)\,\mathrm{d}v \sum_i \frac{C_i}{C}\left[\int \phi_i(v,x,w_i)\mathbf{1}\{x \in A\}\,\mathrm{d}x\,\nu_i(\mathrm{d}w_i)\mathrm{P}\left\{U \le \frac{q_i(v,x,w_i)}{C_i\phi_i(v,x,w_i)}\right\}\right]. \end{aligned}$$

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

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

As a result,

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

where the last equality is due to (3.9). 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, $\mathrm{P}\{X \in A \,|\, X \text{ is output}\} = \int_A g$, and on the other, with probability 1, the procedure eventually terminates. Because the iterations are independent, $\mathrm{P}\{X \in A \text{ at termination}\} = \int_A g$. Since this is true for all measurable A, we get $X \sim g$.

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 $\mathrm{ID}(\varphi)$ is locally bounded on $(0,\infty)$. Our goal is to sample for φ. We have in (3.8) 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., $p(x) = g(x)\mathbf{1}\{x \le r\}/a$, 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 (3.9), the domains of the increments may be partitioned into subregions, so that each can be treated conveniently in certain way.

For $k \ge 1$, let ℓ^k denote the Lebesgue measure on $\mathbb{R}^k$. Suppose for $k \ge 1$, we can find constants $C_{k1}, \ldots, C_{kn_k} \ge 0$ and measurable instrumental functions $\phi_{kj}(\cdot,\cdot)$ on $\mathbb{R}\times\mathbb{R}^k$, $j = 1,\ldots,n_k$, 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 $k \ge 1$,

$$\begin{gathered}\phi_{kj}(v,\cdot) \text{ are probability densities with respect to } \ell^k, \text{ such that}\\ \{s : \phi_{kj}(v,s) > 0\},\ j = 1,\ldots,n_k, \text{ are disjoint, and}\end{gathered} \tag{4.2}$$

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

Due 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

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

A rejection sampling procedure for g under the 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 $\mathrm{P}\{\kappa = k\} = C_k/C$, $k \geq 0$.
2. If $\kappa = 0$, then output $X = Z$ and stop, else sample η such that $\mathrm{P}\{\eta = j\} = C_{\kappa j}/C_\kappa$, $j = 1, \ldots, n_\kappa$.
3. Given Z, κ, and η, sample $S = (S_1, \ldots, S_\kappa)$ from the density $\phi_{\kappa\eta}(Z, \cdot)$.
4. Sample $U \sim \mathrm{Unif}(0,1)$. If $C_{\kappa\eta} U \phi_{\kappa\eta}(Z, S) \leq q_\kappa(Z, S)$, then output $X = Z + S_1 + \cdots + S_\kappa$ and stop, else 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 (3.8).*

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

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

where

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

We have

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

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

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

By (4.2) and (4.3), for each s, if $q_k(v, s) > 0$, then there is exactly one $j \in \{1, \ldots, n_k\}$, 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

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

which leads to

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

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

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

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

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

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

$$\begin{aligned} &\frac{1}{C} \int p(v) \, \mathrm{d}v \int \mathbf{1}\{x \in A\} \mathbf{1}\{v_1 > r\} h(v, v_1) \cdots h(v_{k-1}, x) \, \mathrm{d}v_1 \cdots \mathrm{d}v_{k-1} \, \mathrm{d}x \\ &= \frac{1}{C} \int \mathbf{1}\{x \in A\} \left[\int p(v) h_{k-1}(v, x) \, \mathrm{d}v \right] \mathrm{d}x. \end{aligned}$$

Combining the above results and using (3.8), in each iteration, $\mathrm{P}\{X \in A \text{ is output}\} = (aC)^{-1} \int_A g$. The proof is then complete following the argument for Proposition 4.1.

4.3. Conditional density *vs.* 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 its Lévy density being known. Given $r > 0$, if we can sample X conditional on $X \le r$, 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 $X \le r$ by rejection sampling.*

Basically, the result says that the sampling of an i.d. random variable boils down to the sampling of its conditional density on some $(0, r]$. The question is, 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 $r > 0$, such that $\varphi_1 = \varphi_2$ on $(0, r]$. Then, provided we can sample for one of them, we can sample for the other one using Algorithm 4.2.*

Proof. By assumption, φ_1 and φ_2 share an upper truncated version $\psi(t) = \varphi_1(t)\mathbf{1}\{t \le r\}$. Let $X_i \sim \mathrm{ID}(\varphi_i)$, $i = 1, 2$, and $Y \sim \mathrm{ID}(\psi)$. We have seen from Introduction that for $i = 1, 2$, the density of X_i conditional on $X_i \le r$ is identical to that of Y conditional on $Y \le r$. 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 $X_1 \le r$, and hence sample X_2 conditional on $X_2 \le r$. Then from Proposition 4.2, we can sample X_2 by applying Algorithm 4.2 to φ_2.

The 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 $X \sim \mathrm{ID}(\lambda)$ conditional on $X \le r$. The sampled value then follows the conditional distribution of $\mathrm{ID}(\varphi)$ 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

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

Notice that under the much stronger condition $\mathrm{e}^{-b_1 t}\varphi_1(t) \le \varphi_2(t) \le \mathrm{e}^{b_2 t}\varphi_1(t)$ for some b_1, $b_2 \ge 0$, the methods in Section 2 can be used to sample for one of 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 $t \to 0+$, 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 $\mathrm{e}^{-bt} \leq 1 - a_1 t$ for $0 < t \leq r$. If $a_1 = 0$, we can just let $b = 0$ and $r = r_0$. Define $\psi(t) = \mathrm{e}^{-bt}\varphi_1(t)\mathbf{1}\{t \leq r\}$. Then it is straightforward to check $\psi(t) \leq \varphi_2(t)\mathbf{1}\{t \leq r\}$. Let $\chi(t) = \varphi_2(t) - \psi(t)$. Consider the following procedure.

- Sample $X \sim \mathrm{ID}(\psi)$ and $\xi \sim \mathrm{ID}(\chi)$ 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,

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

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

Finally, notice 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 the center 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 $\mathrm{ID}(\varphi)$ 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, $\mathrm{P}\{X \text{ is output}\} = (aC)^{-1}$, and by (3.7), $a = \int_0^r g$.

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 $\kappa \geq 1$, η, S and U. Let $D(\omega)$ be the amount of time

to generate and process ω. For example, the computation of $\phi_{\kappa\eta}(Z,S)$ and $q_\kappa(Z,S)$ in step 4 is part of the processing of ω. Note that, after each iteration, whether or not the algorithm stops is determined by ω.

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

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 $\omega_1, \omega_2, \ldots$, $\mathrm{E}T = \mathrm{E}[D(\omega)] \times \mathrm{E}N$ (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 $\mathrm{ID}(\varphi)$ 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 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*

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

Then in Algorithm 4.2, one 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 design, step 4 of Algorithm 4.2 needs to compare $U \sim \mathrm{Unif}(0,1)$ 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 algorithm with probability ζ. As shown next, alternative methods to stop the algorithm can be found.

5.2. An example on 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 $\mathrm{ID}(\varphi)$ is smooth and bounded on $\mathbb{R}$ (cf. [33], p. 190). Let $X \sim \mathrm{ID}(\varphi)$. By 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 $c = 1$ in the following.

Let $\lambda(t) = t^{-\alpha-1}\mathbf{1}\{t > 0\}$. The sampling of $\mathrm{ID}(\lambda)$ 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 $Z \sim \mathrm{ID}(\lambda)$ conditional on $Z \leq r$ and κ such that $\mathrm{P}\{\kappa = k\} = b_k/\sum_{i\geq 0} b_i$, $k \geq 0$.
2. If $\kappa = 0$, then output $X = Z$ and stop.
3. Sample $T_1 \sim \mathrm{Beta}(\kappa\alpha, 1-\alpha)$ conditional on $T_1 < Z/r$. Set $S_1 = Z(1/T_1 - 1)$. If $\kappa > 1$, then, sequentially, for $i = 2, \ldots, \kappa$, sample $T_i \sim \mathrm{Beta}((\kappa - i + 1)\alpha, 1-\alpha)$ and set $S_i = (Z + S_1 + \cdots + S_{i-1})(1/T_i - 1)$.

4. Sample $U \sim \mathrm{Unif}(0,1)$ and $T \sim \mathrm{Beta}(\kappa\alpha, 1-\alpha)$. If $S_i \le r$ for all $i \le \kappa$ and $U \le [(r-rT)/(r-ZT)]^\alpha$, then output $X = Z + S_1 + \cdots + S_\kappa$ and stop, else go back to step 1.

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

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

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

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

To apply Proposition 5.1, let

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

Then $q_k(v,s) \le \bar{q}_k(v,s)$. To get $Q_k(v) = \int \bar{q}_k(v,s)\,\mathrm{d}s$, we use

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

which can be verified by change of variable $s = z(u^{-1} - 1)$ and properties of Beta functions. By integrating over s_k, s_{k-1}, …, s_1 one by one, we can get, for $k \ge 1$ and $j = 1, \ldots, k-1$,

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

and

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

where the last equality is by change of variable $s_1 = r - v + u$. It follows that

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

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

$$\sum_{k=1}^{\infty} b_k = \sum_{k=1}^{\infty} \frac{\alpha\theta^k}{(k-1)!} \int_0^\infty x^{k\alpha-1} \mathrm{e}^{-x}\, \mathrm{d}x = \alpha \int_0^\infty \theta x^{\alpha-1} \mathrm{e}^{-x} \sum_{k=0}^{\infty} \frac{\theta^k x^{\alpha k}}{k!}\, \mathrm{d}x$$
$$= \alpha \int_0^\infty \theta x^{\alpha-1} \mathrm{e}^{-x} \mathrm{e}^{\theta x^\alpha}\, \mathrm{d}x = \int_0^\infty \mathrm{e}^{z-(z/\theta)^{1/\alpha}}\, \mathrm{d}z = \int_0^\infty \mathrm{e}^{z-r[\alpha z/\Gamma(1-\alpha)]^{1/\alpha}}\, \mathrm{d}z < \infty. \quad (5.3)$$

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

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

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 $S = (S_1, \ldots, S_\kappa) \sim \phi_\kappa(Z, \cdot)$. 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 $T_1 \sim \mathrm{Beta}(\kappa\alpha, 1-\alpha)$ conditional on $T_1 < Z/r$. Thus, S_1 can be sampled as $Z(1/T_1 - 1)$. For $i > 1$, conditional on $S_1, \ldots, S_{i-1}$, it can be seen likewise that S_i has density $C\mathbf{1}\{x > 0\}\, x^{-\alpha}(Z + S_1 + \cdots + S_{i-1} + x)^{-1-(\kappa-i)\alpha}$ and thus can be sampled as $(Z + S_1 + \cdots + S_{i-1})(1/T_i - 1)$, with $T_i \sim \mathrm{Beta}((\kappa - i + 1)\alpha, 1 - \alpha)$. Step 3 is then established.

Finally, as remarked after Proposition 5.1, given Z, κ, and $S = (S_1, \ldots, S_\kappa)$, 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 change of variable $u = r(t^{-1} - 1)$ in (5.2). Then we can get $Q_\kappa(Z)/Q_\kappa(r) = \mathrm{E}\xi$, where $\xi = [(r - rT)/(r - ZT)]^\alpha$ with $T \sim \mathrm{Beta}(\kappa\alpha, 1-\alpha)$. Note $0 \leq \xi \leq 1$. Since $\mathrm{E}\xi = \mathrm{P}\{U \leq \xi\}$ for any ξ independent of U with $\mathrm{P}\{0 \leq \xi \leq 1\} = 1$, it is seen that step 4 is correct.

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

$$\mathrm{E}N = \exp(r^{-\alpha}/\alpha)\mathrm{P}\{Z \leq r\} \left\{1 + \int_0^\infty \mathrm{e}^{z-r[\alpha z/\Gamma(1-\alpha)]^{1/\alpha}}\, \mathrm{d}z\right\}.$$

By Proposition 4.6, $\mathrm{E}T = \mathrm{E}D \times \mathrm{E}N$, where T is the amount of time required to sample one $X \sim \mathrm{ID}(\varphi)$, and D that required by a single iteration. Each iteration has to sample 1) one $Z \sim \mathrm{ID}(\lambda)$ conditional on $Z \leq r$, 2) one κ from a distribution that depends on r, and 3) provided $\kappa \geq 1$, one $T_1 \sim \mathrm{Beta}(\kappa\alpha, 1-\alpha)$ conditional $T_1 \leq Z/r$, and for each $i = 1, \ldots, \kappa$, a value from $\mathrm{Beta}(i\alpha, 1-\alpha)$. These samplings account for most of D. Denote by $D_1(r)$ the amount of time required for the conditional sampling of Z, $D_2(r)$ the amount of time required for the sampling of κ, and for $k \geq 1$, $z \in (0, 1]$, $D_3(k, z)$ the amount of time required to sample $\zeta \sim \mathrm{Beta}(k\alpha, 1-\alpha)$ conditional on $\zeta \leq z$. In Appendix 1, we show that it is possible to bound $\mathrm{E}D_1(r)$, $\mathrm{E}D_2(r)$, and $\mathrm{E}D_3(k, z)$ uniformly for $r > 0$, $k \geq 1$ and $z > 0$. Consequently, $\mathrm{E}D$ 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 $r \to \infty$, $\mathrm{E}N \to 1$ and $\mathrm{E}T \to \mathrm{E}D_1(\infty)$, the expected amount of time to sample one $Z \sim \mathrm{ID}(\lambda)$. On the other hand, as $r \to 0$, it can be shown $\mathrm{E}N \asymp \exp(r^{-\alpha}/\alpha)$, where for two functions f and g, $f \asymp g$ stands for $f = O(g)$ and $g = O(f)$; see Appendix 1. Therefore, if we directly apply the algorithm to $\varphi(t) = t^{-\alpha-1}\mathbf{1}\{t \le r\}$, then $\mathrm{E}N$ is extremely large for small r. However, by remark 2) following the algorithm, we can instead sample $O(r^{-\alpha})$ i.i.d. random variables with Lévy density $t^{-\alpha}\mathbf{1}\{t \le r'\}$ for some $r' \ge 1$ and then take their weighted sum. In this way, both $\mathrm{E}N$ and $\mathrm{E}T$ are lowered to the order of $O(r^{-\alpha})$.

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

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

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

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

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

If $\psi(t) = \mathrm{e}^{ct}$, where $c < \alpha+1$, then $\mathrm{ID}(\varphi)$ 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 is $c < \alpha + 1$, such that $\mathrm{e}^{ct} \le \psi(t)$ 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, $\inf_{t>0}[\mathrm{e}^{-ct}\psi(t)] = 0$, for example, e^{-t^2}, $\mathrm{e}^t[1 - \sin(t)]$, $(1-t)^2$. 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 $t \to 0+$, we indeed can find r and β to meet the requirement of the algorithm. From Section 2, the random variable Z sampled by step 1 has Lévy density $\mathrm{e}^{(\beta-\alpha-1)t}t^{-\alpha-1}\mathbf{1}\{t \le r\}$. From the choice of r and β, if $t > r$, then $\chi(t) = \varphi(t) \ge 0$, and if $0 < t \le r$, then, by $1 - \mathrm{e}^{-t} < t$ for all $t > 0$,

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

showing 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 separately. By Proposition 4.6, the expected amount of time required by step 1 is $\mathrm{E}T \times \mathrm{E}N$, where T is the amount of time to sample one observation from $\mathrm{ID}(\varphi_1)$ and N the total number of iterations required by the step. The analysis on $\mathrm{E}T$ is identical to Section 5.2. On the other hand, since in each iteration, the probability of acceptance is $\mathrm{P}\{U \le \mathrm{e}^{(\beta-\alpha-1)Z}\} = \mathrm{E}[\mathrm{e}^{(\beta-\alpha-1)Z}] = \exp\{\int[\mathrm{e}^{(\beta-\alpha-1)t} - 1]\varphi_1(t)\,\mathrm{d}t\}$, we have $\mathrm{E}N = \exp\{\int_0^r[1 - \mathrm{e}^{(\beta-\alpha-1)t}]t^{-\alpha-1}\,\mathrm{d}t\} \le \exp\{(\alpha+1-\beta)r^{1-\alpha}/(1-\alpha)\}$.

In step 2, ξ can be sampled as the sum of 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 by step 2 is of the same order as $\mathrm{E}K = \int \chi = \int [\psi(t) - \mathrm{e}^{(\beta-\alpha-1)t} t^{-\alpha-1} \mathbf{1}\{t \le r\}]\,\mathrm{d}t$.

5.4. A theoretical application

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

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

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

Algorithm 5.3. Define, for $k \ge 1$, $0 < v \le r$, $s = (s_1, \ldots, s_k)$, $s_i > 0$,

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

Set $\phi_k(v,s) = q_k(v,s)/Q_k(v)$, where $Q_k(v) = \int q_k(v,s)\,\mathrm{d}s$. Set $b_0 = 1$ and for $k \ge 1$, $b_k = Q_k(r)$. (We will show $Q_k(r) > Q_k(v)$ for all $v \in (0,r)$).

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

Algorithm 5.3 cannot be actually used to sample from Gamma$(c,1)$, since its step 1 relies on the sampling of 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 get $Q_k(v)$. Let $s_1' = s_1 + v - r$. Then

$$q_k(v,s) = c^k \mathbf{1}\{s_1' > 0, \text{ all } s_i > 0\}\, \mathrm{e}^{-(r-v)-(s_1'+s_2+\cdots+s_k)} \prod_{i=1}^{k} \frac{1}{r + s_1' + s_2 + \cdots + s_i}.$$

Integrating over s_1', s_2, …, s_k, it follows that for $k \ge 1$,

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

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

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

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

$$B = \frac{\mathrm{P}\{Z > r\}}{\mathrm{E}[\mathrm{e}^{-(r-Z)}\mathbf{1}\{Z \le r\}]} = c\Gamma(c) r^{-c} \mathrm{e}^{r} \mathrm{P}\{Z > r\} = \int_1^\infty \mathrm{e}^{r(1-t^{1/c})}\,\mathrm{d}t.$$

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 $\mathrm{ID}(\varphi)$ 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, $k \ge 1$. 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 $\mathrm{ID}(\varphi)$ on $(0, r]$. By upper truncating φ if necessary, assume $\int \varrho < \infty$, where $\varrho(t) = t\varphi(t)$. Moreover, by decreasing r if necessary, assume $\int_v^r \varphi > 0$ for any $v < r$. Indeed, $F(s) = \int_s^r \varphi$ is differentiable almost everywhere and by $\int_0^r \varphi = \infty$, 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 we can find a suitable Lévy density λ with support in $[0,\infty)$, such that

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

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

$$\psi(t) = t\lambda(t)/M. \tag{6.2}$$

Then ψ is a probability density with support in $[0,\infty)$. With a little abuse of notation, for $a > 0$, denote by $\psi(\cdot \,|\, V > a)$ the density of a generic variable $V \sim \psi$ conditional on $V > a$. By the assumptions on φ and λ, for any $a < r$, $\mathrm{P}\{V > a\} > 0$ 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 $V \sim \psi$, $\mathrm{P}\{V < \theta_\alpha\} = \alpha$. Denote $m_k = \lfloor k/2 \rfloor$ for $k \ge 1$.

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

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

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

$$n_1 = 1, \quad C_1 = C_{11} = M/r, \quad \phi_{11}(v, s) = \psi(s \,|\, V > r - v), \tag{6.3}$$

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

$$n_k = 2, \quad C_{k1} = (M/r)^k [4\alpha(1-\alpha)]^{m_k}, \quad C_{k2} = \frac{(M/r)^k (r/\theta_\alpha)^{m_k}}{m_k!}, \tag{6.4}$$

$$\phi_{k1}(v, s) = D_{k1}^{-1} \psi(s_1 \,|\, V > r - v) \prod_{i=2}^{k} \psi(s_i) \mathbf{1}\{N(s) \ge m_k\}, \tag{6.5}$$

$$\phi_{k2}(v, s) = D_{k2}^{-1} \psi(s_1 \,|\, V > r - v) \prod_{i=2}^{k} \psi(s_i) \mathbf{1}\{N(s) < m_k\}, \tag{6.6}$$

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

Remark 6.1. (i) In step 1 of Algorithm 4.2 we need to sample κ. From (6.3) and (6.4), $\mathrm{P}\{\kappa = 1 \,|\, m_\kappa = 0\} = 1$, and for $k \geq 1$, $\mathrm{P}\{\kappa = 2k \,|\, m_\kappa = k\} = 1 - \mathrm{P}\{\kappa = 2k+1 \,|\, m_\kappa = k\} = 1/(1 + M/r)$. 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, $h(s_2, \ldots, s_k) = D_{k1}^{-1} \prod_{i=2}^k \psi(s_i) \mathbf{1}\{N(s) \geq m_k\}$ 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 $T = (S_2, \ldots, S_k) \sim h$ independently. To sample T, first sample $\eta_k \sim \mathrm{Binomial}(k-1, \alpha)$ conditional on $\eta_k \geq m_k$; then, given η_k, sample $\xi_1, \ldots, \xi_{k-1}$ 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 $i > \eta$; finally set T as a random permutation of $\xi_1, \ldots, \xi_{k-1}$. We can sample from $\phi_{k2}(v, \cdot)$ similarly, except that we should sample $\eta_k \sim \mathrm{Binomial}(k-1, \alpha)$ conditional on $\eta_k < m_k$ instead.

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

$$U \leq A_{kj} \mathrm{P}\{V > r - v\} \prod_{i=1}^k \frac{r\varphi(S_i)/\lambda(S_i)}{v + S_1 + \cdots + S_i}, \tag{6.7}$$

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 (6.7) as they are equal to 1 (with probability 1).

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

$$\begin{aligned} \mathrm{E}N &\leq 1 + (M/r) + \sum_{k=2}^\infty \left[(M/r)^k [4\alpha(1-\alpha)]^{m_k} + \frac{(M/r)^k (r/\theta_\alpha)^{m_k}}{m_k!} \right] \\ &= (1 + M/r) \left[\frac{4M^2 \alpha(1-\alpha)/r^2}{1 - 4M^2\alpha(1-\alpha)/r^2} + \mathrm{e}^{M^2/(r\theta_\alpha)} \right]. \end{aligned} \tag{6.8}$$

Proof. We first check that ϕ_{kj} satisfy 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 $N(s)$ only depends on $s_2, \ldots, s_k$,

$$\int \psi_{k1}(v, s) \, \mathrm{d}s = D_{k1}^{-1} \int \prod_{i=2}^k \psi(s_i) \mathbf{1}\{N(s) \geq m_k\} \, \mathrm{d}s_2 \cdots \mathrm{d}s_k = D_{k1}^{-1} \mathrm{P}\{\zeta \geq m_k\},$$

where $\zeta = \sum_{i=2}^k \mathbf{1}\{S_i < \theta_\alpha\}$, with S_i i.i.d. $\sim \psi$. Since $\zeta \sim \mathrm{Binomial}(k-1, \alpha)$, 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 next check that condition (4.3) is satisfied. By definition, $\varrho(t) \leq t\lambda(t) = M\psi(t)$. For $k = 1$ and v with $p(v) > 0$, by $v \leq r$,

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

Therefore, one can choose n_1, C_{11} and ϕ_{11} as in (6.3). For $k>1$, by (4.3),

$$q_k(v,s)=\mathbf{1}\{s_1+v>r\}\prod_{i=1}^{k}\frac{\varrho(s_i)}{v+s_1+\cdots+s_i}$$
$$\le\frac{\mathbf{1}\{s_1+v>r\}\,\varrho(s_1)}{r}\prod_{i=2}^{k}\frac{\varrho(s_i)}{r+s_2+\cdots+s_i}\le\frac{M^k\psi(s_1\,|\,V>r-v)}{r}\bar{\psi}(s),\qquad(6.9)$$

where

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

First, if $N(s)\ge m_k$, then by $\bar{\psi}(s)\le r^{-k+1}\prod_{i=2}^{k}\psi(s_i)\mathbf{1}\{N(s)\ge m_k\}$, (6.5) and (6.9),

$$q_k(v,s)\le(M/r)^k\psi(s_1\,|\,V>r-v)\prod_{i=2}^{k}\psi(s_i)\mathbf{1}\{N(s)\ge m_k\}=D_{k1}(M/r)^k\phi_{k1}(v,s).$$

Since $\eta_k\sim\text{Binomial}(k-1,\alpha)$, by Markov inequality, for $t\ge0$, $D_{k1}=\mathrm{P}\{\eta_k\ge m_k\}\le\mathrm{E}[\mathrm{e}^{t(\eta_k-m_k)}]=(1-\alpha+\alpha\mathrm{e}^t)^{k-1}\mathrm{e}^{-tm_k}\le(1-\alpha+\alpha\mathrm{e}^t)^{2m_k}\mathrm{e}^{-tm_k}$. Letting $t=\ln(1/\alpha-1)$, which is nonnegative since $\alpha\le1/2$, $D_{k1}\le[4\alpha(1-\alpha)]^{m_k}$. Thus, with C_{k1} as in (6.4), $q_k(v,s)\mathbf{1}\{N(s)\ge m_k\}\le C_{k1}\phi_{k1}(v,s)$.

Second, suppose $N(s)<m_k$. For each i, $r+s_2+\cdots+s_i\ge r+n_i\theta_\alpha$, where n_i is the total number of $2\le j\le i$ with $s_j\ge\theta_\alpha$. If s_{i_j}, $j=1,\ldots,k-1-N(s)$, are the members among s_2, $\ldots$, s_k that are no less than θ_α, then $n_{i_j}=j$. As a result,

$$\bar{\psi}(s)=\prod_{s_i<\theta_\alpha}\frac{1}{r+s_2+\cdots+s_i}\prod_{s_i\ge\theta_\alpha}\frac{1}{r+s_2+\cdots+s_i}\prod_{i=2}^{k}\psi(s_i)$$
$$\le r^{-N(s)}\prod_{j=1}^{k-1-N(s)}\frac{1}{r+j\theta_\alpha}\prod_{i=2}^{k}\psi(s_i)=r^{-(k-1)}\prod_{j=1}^{k-1-N(s)}\frac{1}{(1+j\theta_\alpha/r)}\prod_{i=2}^{k}\psi(s_i)$$
$$\overset{(a)}{\le} r^{-(k-1)}\prod_{j=1}^{m_k}\frac{1}{(1+j\theta_\alpha/r)}\prod_{i=2}^{k}\psi(s_i)\le\frac{r^{-(k-1)}(r/\theta_\alpha)^{m_k}}{m_k!}\prod_{i=2}^{k}\psi(s_i),$$

where (a) is due to $k-1-N(s)\ge m_k$. Therefore, by (6.6) and (6.9),

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

Since $D_{k2}\le1$, (4.3) is satisfied.

From the selection of α, $\sum_k C_{k1}<\infty$. It is straightforward to verify $\sum_k C_{k2}<\infty$. 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 $ct^{-1}\mathbf{1}\{0<t\le1\}$ (cf. [36]). From Laplace transform, rZ has Lévy density $ct^{-1}\mathbf{1}\{0<t\le r\}$. Efficient sampling methods for Vervaat perpetuities

are available [16, 17, 22]. For the procedure in [17], it is shown that as $c \to \infty$, the expected number of iterations is $\exp(c \ln c + O(c))$ as $c \to \infty$, while as $c \to 0$, the expected number tends to 1. In [16], for the special case $c = 1$, a procedure is discovered that has the expected number of iterations 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 $ct^{-1}\mathbf{1}\{0 < t \le 1\}$, we can instead sample for $ct^{-1}\mathbf{1}\{0 < t \le r\}$ 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 $ct^{-1}\mathbf{1}\{0 < t \le r\} = \varphi(t) + \chi(x)$, where

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

We then apply Proposition 6.1 to φ. It perhaps is not surprising that when r is large, the expected number of iterations needed is close to 1. Intuitively, this is because $\mathrm{ID}(\varphi)$ is so close to $\mathrm{Gamma}(c, 1)$, that with large probability, a single step to sample from $\mathrm{Gamma}(c, 1)$ 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 $L \ge 1$ is a parameter. Set

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

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

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

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

$$A_{k1} = \mathrm{P}\{\tau \ge m_k\}, \quad A_{k2} = \mathrm{P}\{\tau < m_k\} m_k! (\theta/r)^{m_k}.$$

1. Sample $Z \sim \mathrm{Gamma}(c, 1)$ conditional on $Z \le r$ and κ such that $\mathrm{P}\{\kappa = k\} = C_k/C$.
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 $\mathrm{P}\{\eta = j\} = C_{\kappa j}/C_\kappa$.
3. Sample $S_1 \sim \mathrm{Exp}(1)$ conditional on $S_1 \in (r - Z, r)$. If $\kappa > 1$, then do the following steps.
   (a) If $\eta = 1$, then sample $\tau \sim \mathrm{Binomial}(\kappa - 1, 1/2)$ conditional on $\tau \ge m_\kappa$, otherwise sample $\tau \sim \mathrm{Binomial}(\kappa - 1, 1/2)$ conditional on $\tau < m_\kappa$.
   (b) Sample $S_2, \ldots, S_\kappa$ independently, such that for $i \le \tau$, $S_i \sim \mathrm{Exp}(1)$ conditional on $S_i < \theta$ and for $i > \tau$, $S_i \sim \mathrm{Exp}(1)$ conditional on $\theta \le S_i < r$. Then randomly permute $S_2, \ldots, S_\kappa$.

4. Sample $U \sim \mathrm{Unif}(0,1)$. If

$$U \le A_{\kappa\eta} \frac{\mathrm{e}^Z - 1}{\mathrm{e}^r - 1} \prod_{i=1}^{\kappa} \frac{r}{Z + S_1 + \cdots + S_i},$$

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

5. Sample $\xi \sim \mathrm{ID}(\chi)$. 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 $\mathrm{Gamma}(c,1)$ has Lévy density $ct^{-1}\mathrm{e}^{-t}\mathbf{1}\{t>0\}$. To implement the design in Proposition 6.1, let $\lambda = \varphi$ in (6.1). Then $t\lambda(t) = c\mathrm{e}^{-t}\mathbf{1}\{0 < t \le r\}$, giving $M = \int t\lambda(t)\,\mathrm{d}t = c(1-\mathrm{e}^{-r})$ as in the setup of the algorithm and $\psi(t) = \mathrm{e}^{-t}\mathbf{1}\{0 < t \le r\}/(1-\mathrm{e}^{-r})$. Note ψ is the density of $W \sim \mathrm{Exp}(1)$ conditional on $W \le r$. By letting $\alpha = 1/2$, θ in the setup of the algorithm is the median θ_α of ψ. Moreover, since $2(M/r)\sqrt{\alpha(1-\alpha)} = M/r < c/r \le 1/L \le 1$, $(M,\alpha) = (M, 1/2)$ satisfies the very first assumption of Proposition 6.1. The values of C_k and C_{kj} are set according to (6.3) and (6.4). Then steps 1 and 2 follow from the general procedure in Algorithm 4.2. By (6.3), for any $v \in (0,r)$ and s,

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

and by (6.5) and (6.6), for $k \ge 2$,

$$\phi_{k1}(v,s) = D_{k1}^{-1}\phi_{11}(v,s_1)\prod_{i=2}^{k}\frac{\mathrm{e}^{-s_i}\mathbf{1}\{0<s_i\le r\}}{1-\mathrm{e}^{-r}}\mathbf{1}\{N(s)\ge m_k\},$$

$$\phi_{k2}(v,s) = D_{k2}^{-1}\phi_{11}(v,s_1)\prod_{i=2}^{k}\frac{\mathrm{e}^{-s_i}\mathbf{1}\{0<s_i\le r\}}{1-\mathrm{e}^{-r}}\mathbf{1}\{N(s)< m_k\},$$

where $D_{k1} = 1 - D_{k2} = \mathrm{P}\{\eta_k \ge m_k\}$ with $\eta_k \sim \mathrm{Binomial}(k-1, 1/2)$. From the remark after Proposition 6.1, 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 (6.7).

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 $\alpha = 1/2$, by (6.8),

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

From $M/r < c/r \le 1/L$ and $M^2/(r\theta_\alpha) < c^2/(r\theta_\alpha) < 1/(L\theta_\alpha)$, it is seen $\mathrm{E}N \to 1$ as $L \to \infty$.

By Proposition 4.6, $\mathrm{E}T = \mathrm{E}D \times \mathrm{E}N$, where D is the amount of time required by an iteration. Each iteration has to sample 1) one $Z \sim \mathrm{Gamma}(c,1)$ conditional on $Z \le r$, 2) one κ, and 3) provided $\kappa \ge 1$, κ observations from various conditional distributions of $\mathrm{Exp}(1)$. Denote by D_i, $i = 1,2,3$, respectively, the amounts of time required by the samplings. Then $D \approx D_1 + D_2 + D_3$. First, if we use rejection sampling for 1), the expected number of iterations is $m_c = 1/\mathrm{P}\{Z \le r\}$, with $Z \sim \mathrm{Gamma}(c,1)$. Since $r \ge \max(c^2, 1)$ 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 $\mathrm{E}D_1 = O(m_c) = O(1)$. Second, the sampling of κ is standard, with $\mathrm{E}D_2$ uniformly bounded; see remark 1) following Proposition 6.1. Finally, since $\zeta \sim \mathrm{Exp}(1)$ conditional on $0 \le a < \zeta < b \le \infty$ can be sampled as $-\ln[(1-U)\mathrm{e}^{-a} + U\mathrm{e}^{-b}]$ with $U \sim \mathrm{Unif}(0,1)$, it is seen $\mathrm{E}(D_3 \,|\, \kappa)$ is proportional to κ, and thus $\mathrm{E}D_3$ is of the same order as $C_1/C + \sum_{k\ge 2} k(C_{k1} + C_{k2})/C$, which is $o(1)$ as $r \to 0$.

Now consider the sampling for χ. For $t \in [0,1]$, $\chi(t) \le c$, for $t \in [1,r]$, $\chi(t) \le c/t$, 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 $\mathrm{ID}(\chi)$. The number of points follows a Poisson distribution with mean

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

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

6.3. Another application

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

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

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

$$\varphi_b(t) - \varphi_a(t) = c\mathrm{e}^{-t}(t^{a-1} - t^{b-1})\mathbf{1}\{0 < t \le r\}/\ln(1/t)$$

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

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

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

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

We will see in a moment 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 $M = c_r(1-\mathrm{e}^{-r})$. Then set θ, C_k, C_{kj}, C, A_{kj} exactly as in the algorithm in Section 6.2.

1. Keep sampling (W, Z) until $Z \le r$, where, conditional on W, $Z \sim \mathrm{Gamma}(W, 1)$. Sample κ such that $\mathrm{P}\{\kappa = k\} = C_k/C$.
2. If $\kappa = 0$, then output $X = Z$ and stop. If $\kappa = 1$, set $\eta = 1$. Otherwise sample $\eta \in \{1, 2\}$, such that $\mathrm{P}\{\eta = j\} = C_{\kappa j}/C_\kappa$.
3. Sample $(S_1, \ldots, S_\kappa)$ exactly as step 3 of the algorithm in Section 6.2.

4. Sample $U \sim \mathrm{Unif}(0,1)$. If

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

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

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

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

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

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

Since $Z \sim G(W)$, the sampling of Z in step 1 follows. Since $r \le \mathrm{e}^{-c}$, by letting $c_r = c/\ln(1/r)$, we get $\varphi_a(t) \le c_r t^{-1}\mathrm{e}^{-t}\mathbf{1}\{0 < t \le r\}$. Then, to apply Proposition 6.1, set $\lambda(t) = c_r t^{-1}\mathrm{e}^{-t}\mathbf{1}\{0 < t \le r\}$ in (6.1) to get $M = c_r(1-\mathrm{e}^{-r})$ and $\psi(t) = \mathrm{e}^{-t}\mathbf{1}\{0 < t \le r\}/(1-\mathrm{e}^{-r})$. Since $M/r = c_r(1-\mathrm{e}^{-r})/r < 1$, we can set $\alpha = 1/2$ and get the same θ_α and all other constants as in the algorithm in Section 6.2, and execute its steps 2 and 3 without any change. Finally, step 4 follows from (6.7).

To justify the algorithm for W, let $\psi_0(t) = ct^{-1}\mathbf{1}\{0 < t \le a\}$. 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 by step 1, the value of aV it generates has Lévy density $\mathrm{e}^{-3t/2}\psi_0(t)$. On the other hand, recall $\Gamma(t)$ is convex on $(0,\infty)$, $\Gamma(t) \le \Gamma(1) = 1$ for $t \in [1,2]$, and $\Gamma'(1) = -\gamma$, where $\gamma < 3/5$ is Euler's constant (cf. [2]). Then, as $a \le 1$, $0 \le \psi_0(t) - \psi(t) = ct^{-1}[1-\Gamma(1+t)]\mathbf{1}\{0 < t \le a\} \le \gamma t\psi_0(t)$, giving $\psi_0(t) \ge \psi(t) \ge (1-\gamma t)\psi_0(t) \ge \mathrm{e}^{-3t/2}\psi_0(t)$. From this step 2 follows.

6.3.2. *Complexity.* First, consider the algorithm to sample for φ_a. Let N and T be respectively the number of iterations and amount of time required by the algorithm to generate an output. By (6.8),

$$1 \le \mathrm{E}N \le (1 + M/r)\left[\frac{M^2/r^2}{1 - M^2/r^2} + \mathrm{e}^{M^2/(r\theta)}\right].$$

As $r \to 0$, since $M/r = c_r(1-\mathrm{e}^{-r})/r = c(1-\mathrm{e}^{-r})/[r\ln(1/r)] \sim c/\ln(1/r)$ and $r/\theta = r/\ln[2/(1+\mathrm{e}^{-r})] \sim 2$, we get $\mathrm{E}N \to 1$ as $r \to 0$.

By Proposition 4.6, $\mathrm{E}T = \mathrm{E}N \times \mathrm{E}D$, where D is the amount of time required by an iteration in the algorithm. Let D_1, D_2 and D_3 be respectively the amounts of time to sample 1) one (W, Z) conditional on $Z \le r$, 2) one κ, and 3) provided $\kappa \ge 1$, $S_1, \ldots, S_\kappa$. Then $D \approx D_1 + D_2 + D_3$. Let Δ be the amount of time to sample a pair (W, Z). Then $\mathrm{E}D_1 = \mathrm{E}\Delta/\mathrm{P}\{Z \le r\}$. A bound for $\mathrm{E}\Delta$ will be given below. In Appendix 2, it is shown that given $c > 0$, as $r \to 0$, $\mathrm{P}\{Z \le r\} = (1+o(1))\mathrm{e}^{-cA}[\ln(1/r)]^{-c}$, where $A = A(a)$ is a constant. Then $\mathrm{E}D_1 \sim \mathrm{e}^{cA}[\ln(1/r)]^c\mathrm{E}\Delta$. The sampling of κ is standard, with $\mathrm{E}D_2$ uniformly bounded; see

remark 1) following Proposition 6.1. Finally, as in Section 6.2, $\mathrm{E}D_3$ is of the same order as $C_{11}/C + \sum_{k\ge 2} k(C_{k1}+C_{k2})/C$. It follows that $\mathrm{E}D_3 \to 0$ as $r \to 0$.

Next, consider the algorithm to sample W. Let T_W be the time it needs 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

$$\frac{\mathrm{E}T_V(c)}{\mathrm{E}(\mathrm{e}^{-3aV/2})} = \exp\left\{c\int_0^1 (1-\mathrm{e}^{-3at/2})t^{-1}\,\mathrm{d}t\right\}\mathrm{E}T_V(c) \le \mathrm{e}^{3ac/2}\mathrm{E}T_V(c)$$

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

$$\int \chi = c\int_0^a [\Gamma(t) - t^{-1}\mathrm{e}^{-3t/2}]\,\mathrm{d}t \le c\int_0^a t^{-1}(1-\mathrm{e}^{-3t/2})\,\mathrm{d}t \le 3ac/2.$$

As a result, $\mathrm{E}T_W = \mathrm{E}T_V(c)/\mathrm{E}(\mathrm{e}^{-3aV/2}) + O(\int\chi) \le \mathrm{e}^{3ac/2}\mathrm{E}T_V(c) + O(ac)$.

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

7. Conditions for local boundedness

Let φ be a Lévy density with $\int\varphi = \infty$. Recall that in order to get the integral series expansion in Theorem 3.1, the density g of $\mathrm{ID}(\varphi)$ has to be locally bounded. In some cases, the local boundedness can be directly checked using the explicit expression of g. Following the argument in Section 1, if $X \sim \mathrm{ID}(\varphi)$ with $\varphi(t) = \lambda(t)\mathbf{1}\{t\le r\}$, 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 density.)
Let $0<M<\infty$. Suppose $\varrho(t) = t\varphi(t)$ is locally bounded on $(0,M)$ and 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 $r<M$.

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

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

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 $x \in [r, c]$,

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

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 $v \in [0, \varepsilon]$, $x - v \ge r - \varepsilon$, and so the above inequality yields

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

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

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

Let $h \to 0$. Since $g_h \to g$ a.e. 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, $g_h(x) \le (2/r) \sup_{v \ge \varepsilon} \varrho(v) + \sup_{v \in [r/3, r]} g(v)$ for $x \ge r$. 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 $\delta \in (-a, a)$,

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

By dominated convergence,

$$\varlimsup_{\delta \to 0} |(x + \delta) g(x + \delta) - x g(x)| \le 2 \sup_{|u - x| \le a} g(u) \int_0^{2a} \varrho \le 2 \sup_{u \in [x/2, (x+M)/2]} g(u) \int_0^{2a} \varrho.$$

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

Acknowledgment 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. Probab.* **2**, 4, 951–972.

[2] Abramowitz, M. and Stegun, I. A. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*. National Bureau of Standards Applied Mathematics Series, Vol. **55**.

[3] Asmussen, S. and Rosiński, J. (2001). Approximations of small jumps of Lévy processes with a view towards simulation. *J. Appl. Probab.* **38**, 2, 482–493.

[4] Bertoin, J. (1996). *Lévy processes*. Cambridge Tracts in Mathematics, Vol. **121**. Cambridge University Press, Cambridge.

[5] Bertoin, J. and Yor, M. (2002). The entrance laws of self-similar Markov processes and exponential functionals of Lévy processes. *Potential Analysis* **17**, 4, 389–400.

[6] Bondesson, L. (1981). Classes of infinitely divisible distributions and densities. *Z. Wahrsch. Verw. Gebiete* **57**, 1, 39–71.

[7] Breiman, L. (1968). *Probability*. Addison-Wesley Publishing Company, Reading, Massachusetts.

[8] Brix, A. (1999). Generalized gamma measures and shot-noise Cox processes. *Adv. Appl. Probab. 31*, 929–953.

[9] Caballero, M. E., Pardo, J. C., and Pérez, J. L. (2010). On Lamperti stable processes. *Probability & Mathematical Statistics* **30**, 1, 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**, 354, 340–344.

[11] Covo, S. (2009). One-dimensional distributions of subordinators with upper truncated Lévy measure, and applications. *Adv. Appl. Probab.* **41**, 2, 367–392.

[12] Devroye, L. (1986a). An automatic method for generating random variates with a given characteristic function. *SIAM J. Appl. Math.* **46**, 4, 698–719.

[13] Devroye, L. (1986b). *Nonuniform random variate generation*. Springer-Verlag, New York.

[14] Devroye, L. (1987). A simple generator for discrete log-concave distributions. *Computing* **39**, 1, 87–91.

[15] Devroye, L. (2001). Simulating perpetuities. *Methodol. Comput. Appl. Probab.* **3**, 1, 97–115.

[16] Devroye, L. and Fawzi, O. (2010). Simulating the Dickman distribution. *Statist. Probab. Lett.* **80**, 3-4, 242–247.

[17] Fill, J. A. and Huber, M. L. (2010). Perfect simulation of Vervaat perpetuities. *Electron. J. Probab.* **15**, 4, 96–109.

[18] Glasserman, P. (2004). *Monte Carlo methods in financial engineering*. Applications of Mathematics (New York), Vol. **53**. Springer-Verlag, New York. Stochastic Modelling and Applied Probability.

[19] Hörmann, W., Leydold, J., and Derflinger, G. (2004). *Automatic nonuniform random variate generation*. Statistics and Computing. Springer-Verlag, Berlin.

[20] Hougaard, P. (1986). Survival models for heterogeneous populations derived from stable distributions. *Biometrika* **73**, 2, 387–396.

[21] Kanter, M. (1975). Stable densities under change of scale and total variation inequalities. *Ann. Probab.* **3**, 4, 697–707.

[22] Kendall, W. S. and Thönnes, E. (2004). Random walk CFTP. Tech. rep., University of Warwick, Department of Statistics.

[23] Kolassa, J. E. (1997). *Series approximation methods in statistics*, Second ed. Lecture Notes in Statistics, Vol. **88**. Springer-Verlag, New York.

[24] Kuznetsov, A., Kyprianou, A. E., Pardo, J., and van Schaik, K. (2011). A Wiener-Hopf Monte Carlo simulation technique for Lévy processes. *Ann. Appl. Probab.* **21**, 6, 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. Probab.* **20**, 2, 522–564.

[26] Lamperti, J. (1972). Semi-stable Markov processes. I. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete 22*, 205–225.

[27] Liu, J. S. (2001). *Monte Carlo strategies in scientific computing*. Springer Series in Statistics. Springer-Verlag, New York.

[28] Papadimitriou, C. H. (1994). *Computational complexity*. Addison-Wesley Publishing Company, 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**, 1-2, 223–252.

[30] Reynolds, D. S. and Savage, I. R. (1971). Random wear models in reliability theory. *Adv. Appl. Probab. 3*, 229–248.

[31] Rudin, W. (1987). *Real and complex analysis*, Third ed. McGraw-Hill Book Co., New York.

[32] Samorodnitsky, G. and Taqqu, M. S. (1994). *Stable non-Gaussian random processes*. Stochastic Modeling. Chapman & Hall, New York. Stochastic models with infinite variance.

[33] Sato, K.-I. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge Studies in Advanced Mathematics, Vol. **68**. Cambridge University Press, Cambridge. Translated from the 1990 Japanese original, Revised by the author.

[34] Steutel, F. W. (1970). *Preserevation of infinite divisibility under mixing and related topics*. Mathematical Centre Tracts, Vol. **33**. Mathematisch Centrum, Amsterdam.

[35] Steutel, F. W. (1979). Infinite divisibility in theory and practice. *Scand. J. Stat.* **6**, 2, 57–64.

[36] Vervaat, W. (1979). On a stochastic difference equation and a representation of nonnegative infinitely divisible random variables. *Adv. Appl. Probab.* **11**, 4, 750–783.

Appendix 1

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

1) $\mathrm{E}D_1(r)$ is uniformly bounded for $r > 0$, with $D_1(r)$ the amount of time required to sample $Z \sim \mathrm{ID}(\lambda)$ conditional on $Z \leq r$, with $\lambda(t) = t^{-\alpha-1}\mathbf{1}\{t > 0\}$;
2) $\mathrm{E}D_2(r)$ is uniformly bounded for $r > 0$, with $D_2(r)$ the amount of time required to sample κ from 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 $\theta = \Gamma(1-\alpha)/(r^\alpha\alpha)$;

3) $\mathrm{E}D_3(k,z)$ 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 $\zeta \sim \mathrm{Beta}(k\alpha, 1-\alpha)$ conditional on $\zeta \leq z$;

4) $\mathrm{E}N \asymp \exp(r^{-\alpha}/\alpha)$ as $r \to 0$, where N is the number of iterations of the algorithm in the section.

We establish the claims in the order of 2), 3), 4), and 1), where the proof of 1) depends on part of the proof of 4).

Claim 2). Define $p_k = 0$ for $k < 0$. All we need is p_k being 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 5 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 $(b_m/C_1)\min(1, \mathrm{e}^{1-(b_m/C_2)(|x|-1/2)})$, where $m = \arg\max_k b_k$.

Let $f(x) = \Gamma(\alpha x)/\Gamma(x)$, $x > 0$. By $\Gamma(x) \sim 1/x$ as $x \to 0+$, 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 p_k is log-concave, it suffices to show $(\ln f)''(x) < 0$ 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,

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

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

Claim 3). The particular parametrization of the Beta distributions is not important. Let $T_{a,b}(r)$ denote the amount of time to sample $\zeta \sim \mathrm{Beta}(a,b)$ conditional on $\zeta \leq r$. Evidently, $D_3(k,r) = T_{k\alpha,1-\alpha}(r)$. 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 $\zeta \sim \mathrm{Beta}(a,b)$ until $\zeta \leq r$; else keep sampling U, V i.i.d. $\sim \mathrm{Unif}(0,1)$ until $V \leq [(1-r)/(1-\zeta)]^{1-b}$, where $\zeta = rU^{1/a}$.

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

$$\begin{aligned}\mathrm{P}\{\zeta \leq 1 - c/a\} &\sim \frac{a^b}{\Gamma(b)}\int_0^{1-c/a} x^{a-1}(1-x)^{b-1}\,\mathrm{d}x \\ &= \frac{1}{\Gamma(b)}\int_c^\theta (1-t/a)^{a-1}t^{b-1}\,\mathrm{d}t \to \frac{1}{\Gamma(b)}\int_c^\infty \mathrm{e}^{-x}t^{b-1}\,\mathrm{d}t > 0.\end{aligned}$$

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

rejection sampling indeed generates $\zeta \sim \mathrm{Beta}(a,b)$ conditional on $\zeta \le r$. In each iteration, the probability of acceptance is

$$\begin{aligned}(1-r)^{1-b}\int_0^1 (1-rt^{1/a})^{b-1}\,\mathrm{d}t &\ge (1-r)^{1-b}\int_{\mathrm{e}^{-1}}^1 (1-rt^{1/a})^{b-1}\,\mathrm{d}t \\ &\ge (1-r)^{1-b}(1-\mathrm{e}^{-1})(1-r\mathrm{e}^{-1/a})^{b-1} \ge (1-\mathrm{e}^{-1})\left[\frac{c/a}{1-(1-c/a)\mathrm{e}^{-1/a}}\right]^{1-b}.\end{aligned}$$

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

Claim 4). The proof of the asymptotic of

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

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

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

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 $r \to 0$, $M \to \infty$. Then by the saddle point method,

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

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],

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

where h is a function defined as

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

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

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

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

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

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

To establish claim 1), it suffices to show that K_r, the expected number of iterations required to sample ξ conditional on $S \le r$, is bounded for $r > 0$. By Bayes formula, conditional on $S \le r$, 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 $\sup_{r \ge r_0} K_r < \infty$ for any fixed $r_0 > 0$. Therefore, we only need to show $\sup_{r \le r_0} K_r < \infty$ for small $r_0 > 0$. From the proof of claim 4), there is $c > 1/h''(0+)$, such that $h(t) \ge h(0) + t^2/(2c)$ for all $t \in (0,\pi)$. Consider the following procedure: keep sampling $X \sim N(0, cr^{\alpha/(1-\alpha)})$ and $U \sim \mathrm{Unif}(0,1)$ until $U \le \exp\{-r^{-\alpha/(1-\alpha)}[h(|X|) - h(0) - X^2/(2c)]\}$ and then return $\xi = |X|$. It is seen the procedure samples ξ conditional on $S \le r$. The probability of acceptance of each iteration is

$$\begin{aligned}&\frac{1}{\sqrt{2\pi c r^{\alpha/(1-\alpha)}}}\int_{-\infty}^{\infty}\exp\{-r^{-\alpha/(1-\alpha)}[h(|t|) - h(0)]\}\,\mathrm{d}t\\ &\qquad = \frac{1}{\sqrt{2\pi c r^{\alpha/(1-\alpha)}}} \times [1 + o(1)]\sqrt{\frac{2\pi r^{\alpha/(1-\alpha)}}{h''(0+)}} \to \frac{1}{\sqrt{ch''(0+)}} > 0, \quad r \to 0,\end{aligned}$$

where the second line is due to the saddle point method. It is then easy to get that, by using the procedure, K_r is uniformly bounded for all small r.

Appendix 2

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

$$\mathrm{P}\{Z \le r\} = [1 + o(1)]\mathrm{e}^{-cA_a}[\ln(1/r)]^{-c}.$$

To prove the claim, for any $r < \mathrm{e}^{-1}$, $Z \sim \eta_r + X_r + Y$, where η_r, X_r and Y are independent i.d. random variables with Lévy densities $\mathbf{1}\{t \le r\}\lambda(t)$, $\mathbf{1}\{r \le t < \mathrm{e}^{-1}\}\lambda(t)$, and $\mathbf{1}\{t > \mathrm{e}^{-1}\}\lambda(t)$,

respectively. It is seen $\mathrm{P}\{Z \le r\} = \mathrm{P}\{\eta_r \le r\}\mathrm{P}\{X_r = 0\}\mathrm{P}\{Y = 0\}$. First, by Markov inequality, for any $s > 0$,

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

By the convexity of the exponential function,

$$\mathrm{P}\{\eta_r > r\} \le \exp\left\{\frac{c}{\ln(1/r)}\int_0^r (s/r)\mathrm{e}^{st/r}\,\mathrm{d}t - s\right\} = \exp\left\{\frac{c(\mathrm{e}^s - 1)}{\ln(1/r)} - s\right\}.$$

Letting $r \to 0$ followed by $s \to \infty$, it is seen $\mathrm{P}\{\eta_r > r\} \to 0$ and hence $\mathrm{P}\{\eta_r \le r\} \to 1$.

Next, by the property of Poisson process,

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

where

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

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