Strongly Efficient Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities

Xingyu Wang

Industrial Engineering and Management Sciences, Northwestern University, xingyuwang2017@u.northwestern.edu

Chang-Han Rhee

Industrial Engineering and Management Sciences, Northwestern University, chang-han.rhee@northwestern.edu

Abstract. In this paper, we address rare-event simulation for heavy-tailed Lévy processes with infinite activities. The presence of infinite activities poses a critical challenge, making it impractical to simulate or store the precise sample path of the Lévy process. We present a rare-event simulation algorithm that incorporates an importance sampling strategy based on heavy-tailed large deviations, the stick-breaking approximation for the extrema of Lévy processes, the Asmussen-Rosiński approximation, and the randomized debiasing technique. By establishing a novel characterization for the Lipschitz continuity of the law of Lévy processes, we show that the proposed algorithm is unbiased and strongly efficient under mild conditions, and hence applicable to a broad class of Lévy processes. In numerical experiments, our algorithm demonstrates significant improvements in efficiency compared to the crude Monte-Carlo approach.

Key words: rare-event simulation, strong efficiency, regularly varying distributions, Lévy process, infinite activities, large deviations, unbiased estimation

1. Introduction In this paper, we propose a strongly efficient rare-event simulation algorithm for heavy-tailed Lévy processes with infinite activities. Specifically, the goal is to estimate probabilities of the form $\mathbf{P}(X \in A)$, where $X = \{X(t) : t \in [0,1]\}$ is a Lévy process in $\mathbb{R}$, A is a subset of the càdlàg space that

2

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

doesn't include the typical path of X so that $\mathbf{P}(X \in A)$ is close to 0, and the event $\{X \in A\}$ is "unsimulatable" due to the infinite number of activities within any finite time interval. The defining features of the problem are as follows.

- The increments of the Lévy process $X(t)$ are heavy-tailed. Throughout this paper, we characterize the heavy-tailed phenomenon through the notion of regular variation and assume that the tail cdf $\mathbf{P}(\pm X(t) > x)$ decays roughly at a power-law rate of $1/x^\alpha$; see Definition 1 for details. The notion of heavy tails provides the mathematical formulation for the extreme uncertainty that manifests in a wide range of real-world dynamics and systems, including the spread of COVID-19 (see, e.g., Cohen et al. (2022)), traffic in computer and communication networks (see, e.g., Li (2018)), financial assets (see, e.g., Embrechts et al. (2013), Borak et al. (2011)), and the training of deep neural networks (see, e.g., Gurbuzbalaban et al. (2021), Hodgkinson and Mahoney (2021)).

- A is a general subset of $\mathbb{D}$ (i.e., the space of the real-valued càdlàg functions over $[0,1]$) that involves the supremum of the path. For concreteness in our presentation, the majority of the paper focuses on

$$A = \left\{ \xi \in \mathbb{D} : \sup_{t \in [0,1]} \xi(t) \geq a; \ \sup_{t \in (0,1]} \xi(t) - \xi(t-) < b \right\}. \tag{1}$$

Intuitively speaking, this is closely related to ruin probabilities under reinsurance mechanisms, as $\{X \in A\}$ requires the supremum of the process $X(t)$ over $[0,1]$ to exceed some threshold a even though all upward jumps in $X(t)$ are bounded by b. Nevertheless, we stress that the algorithmic framework proposed in this paper is flexible enough to address more general form of events $\{X \in A\}$ that are of practical interest. For instance, we demonstrate in Section B of the Supplementary Materials that the framework can also address rare event simulation in the context of barrier option pricing.

- $X(t)$ possesses infinite activities; see Section 2.3 for the precise definition. Consequently, it is computationally infeasible to simulate or store the entire sample path of such processes. In other words, we focus on a computationally challenging case where $\mathbf{I}\{X \in A\}$ cannot be exactly simulated or evaluated. Addressing such "unsimulatable" cases is crucial due to the increasing popularity of Lévy models with

infinite activities in risk management and mathematical finance (see, e.g., Carr et al. (2002, 2003),

Rosiński (2007), Bianchi et al. (2011), Sabino (2022)), as they offer more accurate and flexible descrip-

tions for the price and volatility of financial assets compared to the classical jump-diffusion models

(see, e.g., Mijatović and Tankov (2016)).

In summary, our goal is to tackle a practically significant yet computationally challenging task, where the

nature of the rare events renders crude Monte Carlo methods highly inefficient, if not entirely infeasible, due

to the infinite activities in $X(t)$. To address these challenges, we integrate several mathematical machinery:

a design of importance sampling based on sample-path large deviations for heavy-tailed Lévy processes in

Rhee et al. (2019), the stick-breaking approximation in González Cázares et al. (2022) for Lévy processes

with infinite activities, and the randomized multilevel Monte Carlo debiasing technique in Rhee and Glynn

(2015). By combining these tools, we propose a rare event simulation algorithm for heavy-tailed Lévy

processes with infinite activities that attains strong efficiency (see Definition 2 for details).

As mentioned above, the first challenge is rooted in the nature of rare events as the crude Monte Carlo

methods can be prohibitively expensive when estimating a small $p = \mathbf{P}(X \in A)$. Instead, variance reduction

techniques are often employed for efficient rare event simulation. When the underlying uncertainties are

light-tailed, the exponential tilting strategy guided by large deviation theories has been successfully applied

in a variety of contexts; see, e.g., Bucklew et al. (1990), Boxma et al. (2019), Torrisi (2004), Dupuis

et al. (2007b). However, the exponential tilting approach falls short in providing a principled and provably

efficient design of the importance sampling estimators (see, for example, Bassamboo et al. (2007)) due

to fundamentally different mechanisms through which the rare events occur. Instead, different importance

sampling strategies (e.g., Blanchet and Glynn (2008), Dupuis et al. (2007a), Blanchet et al. (2008), Blanchet

and Liu (2008), Murthy et al. (2014), Blanchet et al. (2013)) and other variance reduction techniques such as

conditional Monte Carlo (e.g., Asmussen and Kroese (2006), Hult et al. (2016)) and Markov Chain Monte

Carlo (e.g., Gudmundsson and Hult (2014)) have been proposed to address problems associated with specific

types of processes or events.

Recent developments in heavy-tailed large deviations, such as those by Rhee et al. (2019) and Wang and

Rhee (2023), offer critical insights into the design of efficient and universal importance sampling schemes

4

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

for heavy-tailed systems. Central to this development is the discrete hierarchy of heavy-tailed rare events,

known as the catastrophe principle. The principle dictates that rare events in heavy-tailed systems arise due to

catastrophic failures of a small number of system components, and the number of such components governs

the asymptotic rate at which the associated rare events occur. This creates a discrete hierarchy among heavy-

tailed rare events. By combining the defensive importance sampling design with such hierarchy, strongly

efficient importance sampling algorithms have been proposed for a variety of rare events associated with

random walks and compound Poisson processes in Chen et al. (2019). See also Wang and Rhee (2024) for a

tutorial on this topic. In this paper, we adopt and extend this framework to encompass Lévy processes with

infinite activities. The specifics of the importance sampling distribution are detailed in Section 3.1.

Another challenge arises from the simulation of Lévy processes with infinite activities. While the design

of importance sampling algorithm in Chen et al. (2019) has been successfully applied to a wide range

of stochastic systems that are exactly simulatable (including random walks, compound Poisson processes,

iterates of stochastic gradient descent, and several classes of queueing systems), it cannot be implemented for

Lévy processes with infinite activities. More specifically, the simulation of the random vector $(X(t), M(t))$,

where $M(t) = \sup_{s \le t} X(t)$, poses a significant challenge in the case with infinite activities. As of now, exact

simulation of the extrema of Lévy processes (excluding the compound Poisson case) is only available for

specific cases (see, for instance, González Cázares et al. (2019), Cázares et al. (2023), Cázares et al. (2020)),

let alone the exact simulation of the joint law of $(X(t), M(t))$. We therefore approach the challenge by

considering the following questions: (i) Does there exist a provably efficient approximation algorithm for

$(X(t), M(t))$, and (ii) Are we able to remove the approximation bias while still attaining strong efficiency

in our rare-event simulation algorithm?

Regarding the first question, several classes of algorithms have been proposed for the approximate simu-

lation of the extrema of Lévy processes. This includes the random walk approximations based on Euler-type

discretization of the process (see e.g., Asmussen et al. (1995), Dia and Lamberton (2011), Giles and Xia

(2017)), the Wiener-Hopf approximation methods (see e.g. Kuznetsov et al. (2011), Ferreiro-Castilla et al.

(2014)) based on the fluctuation theory of Lévy processes, the jump-adapted Gaussian approximations

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

5

(see e.g. Dereich (2011), Dereich and Heidenreich (2011)), and the characteristic function approach in Boyarchenko and Levendorskii (2023a,b) based on efficient evaluation of joint cdf. Nevertheless, the approximation errors in the aforementioned methods are either unavailable or exhibit a polynomial rate of decay. Thankfully, the recently developed stick-breaking approximation (SBA) algorithm in González Cázares et al. (2022) provides a novel approach to the simulation of the joint law of $X(t)$ and $M(t)$. The theoretical foundation of SBA is the following description for the concave majorants of Lévy processes with infinite activities in Pitman and Bravo (2012):

$$\big(X(t),\ M(t)\big) \stackrel{d}{=} \Big(\sum_{j\geq 1} \xi_j,\ \sum_{j\geq 1} \max\{\xi_j, 0\} \Big).$$

Here, $(l_j)_{j\geq 1}$ is a sequence of iteratively generated non-negative RVs satisfying $\sum_{j\geq 1} l_j = t$ and $\mathbf{E}l_j = t/2^j\ \forall j \geq 1$; conditioned on the values of $(l_j)_{j\geq 1}$, ξ_j's are independently generated such that $\xi_j \stackrel{d}{=} X(l_j)$. While it is computationally infeasible to generate the entirety of the infinite sequences $(l_j)_{j\geq 1}$ and $(\xi_j)_{j\geq 1}$, by terminating the procedure at the m-th step we yield approximations of the form

$$\big(\hat{X}_m(t),\ \hat{M}_m(t)\big) \triangleq \Big(\sum_{j=1}^{m} \xi_j,\ \sum_{j=1}^{m} \max\{\xi_j, 0\} \Big). \tag{2}$$

We provide a review in Section 2.3. In particular, due to $\mathbf{E}\big[\sum_{j>m} l_j \big] = t/2^m$, with each extra step in (2) we expect to reduce the approximation error by half, thus leading to the geometric convergence rate of errors. See González Cázares et al. (2022) for analyses of the approximation errors for different types of functionals.

Additionally, while SBA can be considered sufficiently accurate for a wide range of tasks, eliminating the approximation errors is crucial in the context of rare-event simulation. Otherwise, any effort to efficiently estimate a small probability might be fruitless and could be overwhelmed by potentially large errors in the algorithm. In order to remove the approximation errors of SBA in (2), we employ the construction of unbiased estimators proposed in Rhee and Glynn (2015). This can be interpreted as a randomized version of the multilevel Monte Carlo scheme Heinrich (2001), Giles (2008) when a sequence of biased yet increasingly more accurate approximations is available. It allows us to construct an unbiased estimation

6

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

algorithm that terminates within a finite number of steps. By combining SBA, the randomized debiasing technique, and the design of importance sampling distributions based on heavy-tailed large deviations, we propose Algorithm 2 for rare-event simulation of Lévy processes with infinite activities. In case that the exact sampling of $X(t)$, and hence the increments ξ_j's in (2), is not available, we further incorporate the Asmussen-Rosiński approximation (ARA) in Asmussen and Rosiński (2001). This approximation replaces the small-jump martingale in the Lévy process $X(t)$ with a Brownian motion of the same variance, thus leading to Algorithm 3. We note that the combination of SBA and the randomized debiasing technique has been explored in González Cázares et al. (2022), and an ARA-incorporated version of SBA has been proposed in González Cázares and Mijatović (2022). However, the goal of proposing strongly efficient rare-event simulation algorithm adds another layer of difficulty and sets our work apart from the existing literature. In particular, the notion of strong efficiency demands that the proposed estimator remains efficient under the importance sampling algorithm w.r.t. not just a given task, but throughout a sequence of increasingly more challenging rare-event simulation tasks as $\mathbf{P}(X \in A)$ tends to 0. This introduces a new dimension into the theoretical analysis that is not presented in González Cázares et al. (2022), González Cázares and Mijatović (2022) and necessitates the development of new technical tools to characterize the performance of the algorithm when all these components (importance sampling, SBA, debiasing technique, and ARA) are in effect.

An important technical question in our analysis concerns the continuity of the law of the running supremum $M(t)$. To provide high-level descriptions, let us consider estimators for $\mathbf{P}(X \in A) = \mathbf{E}\big[\mathbf{I}\{X \in A\}\big]$ that admit the form $f(\hat{X})$ where $\hat{X}$ is some approximation to the Lévy process X and $f(\xi) = \mathbf{I}\{\xi \in A\}$. SBA and the debiasing technique allow us to construct $\hat{X}$ such that the deviation $\hat{X} - X$ has a small variance. Nevertheless, the estimation can be fallible if X concentrates on the boundary cases, i.e., X falls into a neighborhood of ∂A fairly often. Specializing to the case in (1), this requires obtaining sufficiently tight bounds for probabilities of form $\mathbf{P}(M(t) \in [x, x + \delta])$. Nevertheless, the continuity of the law of the supremum $M(t)$ remains an active area of study, with many essential questions left open. Recent developments regarding the law of $M(t)$ are mostly qualitative or focus on the cumulative distribution function (cdf); see, e.g., Chaumont (2013),

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

7

Coutin et al. (2018), Kwaśnicki et al. (2013), Michna (2012), Michna et al. (2014), Michna (2013). In short, addressing this aspect of the challenge requires us to establish novel and useful quantitative characterizations of the law of supremum $M(t)$.

For our purpose of efficient rare event simulation, particularly under the importance sampling scheme detailed in Section 3.1, the following condition proves to be sufficient:

$$\mathbf{P}\Big(X^{<z}(t) \in [x, x+\delta]\Big) \leq \frac{C}{t^{\lambda} \wedge 1}\delta \qquad \forall z \geq z_0, \; t > 0, \; x \in \mathbb{R}, \; \delta \in [0, 1]. \tag{3}$$

Here, $X^{<z}(t)$ is a modulated version of the process $X(t)$ where all the upward jumps with sizes larger than z are removed; see Section 3 for the rigorous definition. First, we establish in Theorem 1 (resp. Theorem 2) that Algorithm 2 (resp. Algorithm 3) does attain strong efficiency under condition (3). More importantly, we demonstrate in Section 4 that condition (3) is mild for Lévy processes with infinitive activities, as it only requires the intensity of jumps to approach ∞ (hence attaining infinite activities in X) at a rate that is not too slow. In particular, in Theorems 3 and 4 we provide two sets of sufficient conditions for (3) that are easy to verify. We note that the representation of concave majorants for Lévy processes developed in Pitman and Bravo (2012) proves to be a valuable tool for studying the law of $X(t)$ and $M(t)$. As will be elaborated in the proofs in Section 6, the key technical tool that allows us to connect condition 3 with the law of the supremum $M(t)$ is, again, the representation in (2). See also Cázares et al. (2023) for its application in studying the joint density of $X(t)$ and $M(t)$ of stable processes.

Some algorithmic contributions of this paper were presented in a preliminary form at a conference in Wang and Rhee (2021) without rigorous proofs. The current paper presents several significant extensions: (*i*) In addition to Algorithm 2, we also propose an ARA-incorporated version of the importance sampling algorithm (see Algorithm 3) to address the case where $X(t)$ cannot be exactly simulated; (*ii*) Rigorous proofs of strong efficiency are provided in Section 6 in this paper; (*iii*) We establish two sets of sufficient conditions for (3) in Section 4, leveraging the properties of regularly varying or semi-stable processes.

The rest of the paper is structured as follows. Section 2 reviews the theoretical foundations of our algorithms, including the heavy-tailed large deviation theories (Section 2.2), the stick-breaking approximations

8

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

(Section 2.3), and the debiasing technique (Section 2.4). Section 3 presents the importance sampling algorithms and establishes their strong efficiency. Section 4 investigates the continuity of the law of $X(t)$ and provides sufficient conditions for (3), a critical condition to ensure the strong efficiency of our importance sampling scheme. Numerical experiments are reported in Section 5. Section 6 collects the proofs of strong efficiency. In the Supplementary Materials, Section A collects the proofs of technical results in Section 4, and Section B extends the algorithmic framework to the context of barrier option pricing.

2. Preliminaries

In this section, we introduce some notations and results that will be frequently used when developing the strongly efficient rare-event simulation algorithm.

2.1. Notations

Let $\mathbb{N} = \{0, 1, 2, \ldots\}$ be the set of non-negative integers. For any positive integer k, let $[k] = \{1, 2, \ldots, k\}$. For any $x, y \in \mathbb{R}$, let $x \wedge y \triangleq \min\{x, y\}$ and $x \vee y \triangleq \max\{x, y\}$. For any $x \in \mathbb{R}$, we define $(x)^+ \triangleq x \vee 0$ as the positive part of x, and

$$\lfloor x \rfloor \triangleq \max\{n \in \mathbb{Z}: n \leq x\}, \qquad \lceil x \rceil \triangleq \min\{n \in \mathbb{Z}: n \geq x\}$$

as the floor and ceiling function. Given a measure space $(\mathcal{X}, \mathcal{F}, \mu)$ and any set $A \in \mathcal{F}$, we use $\mu|_A(\cdot) \triangleq \mu(A \cap \cdot)$ to denote restriction of the measure μ on A. For any random variable X and any Borel measureable set A, let $\mathscr{L}(X)$ be the law of X, and $\mathscr{L}(X|A)$ be the law of X conditioned on event A. Let $(\mathbb{D}_{[0,1],\mathbb{R}}, \boldsymbol{d})$ be the metric space of $\mathbb{D} = \mathbb{D}_{[0,1],\mathbb{R}}$ (i.e., the space of all real-valued càdlàg functions with domain $[0, 1]$) equipped with Skorokhod J_1 metric $\boldsymbol{d}$. Here, the metric $\boldsymbol{d}$ is defined by

$$\boldsymbol{d}(x, y) \triangleq \inf_{\lambda \in \Lambda} \sup_{t \in [0,1]} |\lambda(t) - t| \vee |x(\lambda(t)) - y(t)| \tag{4}$$

with Λ being the set of all increasing homeomorphisms from $[0, 1]$ to itself.

Henceforth in this paper, the heavy-tailedness of any random element will be captured by the notion of regular variation.

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

9

DEFINITION 1.   For any measurable function $\phi : (0, \infty) \to (0, \infty)$, we say that ϕ is **regularly varying**
as $x \to \infty$ with index β (denoted as $\phi(x) \in \mathcal{RV}_\beta(x)$ as $x \to \infty$) if $\lim_{x\to\infty} \phi(tx)/\phi(x) = t^\beta$ for all $t > 0$. We
also say that a measurable function $\phi(\eta)$ is regularly varying as $\eta \downarrow 0$ with index β if $\lim_{\eta\downarrow 0} \phi(t\eta)/\phi(\eta) = t^\beta$
for any $t > 0$. We denote this as $\phi(\eta) \in \mathcal{RV}_\beta(\eta)$ as $\eta \downarrow 0$.

For properties of regularly varying functions, see, for example, Chapter 2 of Resnick (2007).

Next, we discuss the Lévy-Ito decomposition of one-dimensional Lévy processes, i.e., $X(t) \in \mathbb{R}$. The
law of a one-dimensional Lévy process $\{X(t) : t \geq 0\}$ is completely characterized by its generating triplet
(c, σ, ν) where $c \in \mathbb{R}$ represents the constant drift, $\sigma \geq 0$ is the magnitude of the Brownian motion term, and
the Lévy measure ν characterizes the intensity of the jumps. More precisely,

$$X(t) \stackrel{d}{=} ct + \sigma B(t) + \int_{|x|\leq 1} x\left[N([0,t] \times dx) - t\nu(dx)\right] + \int_{|x|>1} xN([0,t] \times dx) \tag{5}$$

where $\mathrm{Leb}(\cdot)$ is the Lebesgue measure on $\mathbb{R}$, B is a standard Brownian motion, the measure ν satisfies $\int (|x|^2 \wedge 1)\nu(dx) < \infty$, and N is a Poisson random measure over $(0, \infty) \times \mathbb{R}$ with intensity measure $\mathrm{Leb}((0, \infty)) \times \nu$
and is independent of B. For standard references on this topic, see Chapter 4 of Sato et al. (1999).

Given two sequences of non-negative real numbers $(x_n)_{n\geq 1}$ and $(y_n)_{n\geq 1}$, we say that $x_n = \boldsymbol{O}(y_n)$ (as
$n \to \infty$) if there exists some $C \in [0, \infty)$ such that $x_n \leq Cy_n \ \forall n \geq 1$. Besides, we say that $x_n = \boldsymbol{o}(y_n)$ if
$\lim_{n\to\infty} x_n/y_n = 0$. The goal of this paper is described in the following definition of strong efficiency.

DEFINITION 2.   Let $(L_n)_{n\geq 1}$ be a sequence of random variables supported on a probability space
$(\Omega, \mathcal{F}, \mathbf{P})$ and $(A_n)_{n\geq 1}$ be a sequence of events (i.e., $A_n \in \mathcal{F} \ \forall n$). We say that $(L_n)_{n\geq 1}$ are unbiased and
strongly efficient estimators of $(A_n)_{n\geq 1}$ if

$$\mathbf{E}L_n = \mathbf{P}(A_n) \ \forall n \geq 1; \qquad \mathbf{E}L_n^2 = \boldsymbol{O}\left(\mathbf{P}^2(A_n)\right) \ \text{ as } n \to \infty.$$

We stress again that strongly efficient estimators $(L_n)_{n\geq 1}$ achieve uniformly bounded relative errors (i.e.,
the ratio between standard error and mean) for all $n \geq 1$.

10

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

2.2. Sample-Path Large Deviations for Regularly Varying Lévy Processes

The key ingredient of our importance sampling algorithm is the recent development of the sample-path large deviations for Lévy processes with regularly varying increments; see Rhee et al. (2019). To familiarize the readers with this mathematical machinery, we start by reviewing the results in the one-sided cases, and then move onto the more general two-sided results.

Let $X(t) \in \mathbb{R}$ be a centered Lévy process (i.e., $\mathbf{E}X(t) = 0 \ \forall t > 0$) with generating triplet (c, σ, ν) such that the Lévy measure ν is supported on $(0, \infty)$. In other words, all the discontinuities in X will be positive, hence one-sided. Moreover, we are interested in the heavy-tailed setting where the function $H_+(x) = \nu[x, \infty)$ is regularly varying as $x \to \infty$ with index $-\alpha$ where $\alpha > 1$. Define a scaled version of the process as $\bar{X}_n(t) \triangleq \frac{1}{n}X(nt)$, and let $\bar{X}_n \triangleq \{\bar{X}_n(t) : t \in [0, 1]\}$. Note that $\bar{X}_n$ is a random element taking values in $\mathbb{D}$.

For all $l \geq 1$, let $\mathbb{D}_l$ be the subset of $\mathbb{D}$ containing all the non-decreasing step functions that has l jumps (i.e., discontinuities) and vanishes at the origin. Let $\mathbb{D}_0 = \{\mathbf{0}\}$ be the set that only contains the zero function $\mathbf{0}(t) \equiv 0$. Let $\mathbb{D}_{<l} = \cup_{j=0,1,\cdots,l-1}\mathbb{D}_l$. For any $\beta > 0$, let ν_β be the measure supported on $(0, \infty)$ with $\nu_\beta(x, \infty) = x^{-\beta}$. For any positive integer l, let ν_β^l be the $l-$fold product measure of ν_β restricted on $\{\mathbf{y} = (y_1, \ldots, y_l) \in (0, \infty)^l : y_1 \geq y_2 \geq \cdots \geq y_l\}$. Define the measure (for $l \geq 1$)

$$\mathbf{C}_l(\cdot) \triangleq \mathbf{E}\left[\nu_\beta^l\left\{\mathbf{y} \in (0, \infty)^l : \sum_{j=1}^{l} y_j \mathbf{I}_{[U_j, 1]} \in \cdot\right\}\right]$$

where all U_j's are iid copies of Unif$(0, 1)$. In case that $l = 0$, we set $\mathbf{C}_\beta^0$ as the Dirac measure on $\mathbf{0}$. The following result provides sharp asymptotics for rare events associated with $\bar{X}_n$. Henceforth in this paper, all measurable sets are understood to be Borel measurable.

Result 1 (Theorem 3.1 of Rhee et al. (2019)) *Let $A \subset \mathbb{D}$ be measurable. Suppose that $\mathcal{J}(A) \triangleq \min\{j \in \mathbb{N} : \mathbb{D}_j \cap A \neq \emptyset\} < \infty$ and A is bounded away from $\mathbb{D}_{<\mathcal{J}(A)}$ in the sense that $\mathbf{d}(A, \mathbb{D}_{<\mathcal{J}(A)}) > 0$. Then*

$$\mathbf{C}_{\mathcal{J}(A)}(A^\circ) \leq \liminf_{n \to \infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n, \infty))^{\mathcal{J}(A)}} \leq \limsup_{n \to \infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n, \infty))^{\mathcal{J}(A)}} \leq \mathbf{C}_{\mathcal{J}(A)}(A^-) < \infty$$

where A°, A^- are the interior and closure of A respectively.

Intuitively speaking, Result 1 embodies a general principle that, in heavy-tailed systems, rare events arise due to several "large jumps". Here, $\mathcal{J}(A)$ denotes the minimum number of jumps required in $\bar{X}_n$ for event $\{\bar{X}_n \in A\}$ to occur. As shown above in Result 1, $\mathcal{J}(A)$ dictates the polynomial rate of decay of the probabilities of the rare events $\mathbf{P}(\bar{X}_n \in A)$. Furthermore, results such as Corollary 4.1 in Rhee et al. (2019) characterize the conditional limits of $\bar{X}_n$: conditioning on the occurrence of rare events $\{\bar{X}_n \in A\}$, the conditional law $\mathscr{L}(\bar{X}_n|\{\bar{X}_n \in A\})$ converges in distribution to that of a step function over $[0,1]$ with exactly $\mathcal{J}(A)$ jumps (of random sizes and arrival times) as $n \to \infty$. Therefore, $\mathcal{J}(A)$ also dictates the most likely scenarios of the rare events. This insight proves to be critical when we develop the importance sampling distributions for the rare events simulation algorithm in Section 3.

Results for the two-sided cases admit a similar yet slightly more involved form, where the Lévy process $X(t)$ exhibits both positive and negative jumps. Specifically, let $X(t)$ be a centered Lévy process such that for $H_+(x) = \nu[x, \infty)$ and $H_-(x) = \nu(-\infty, -x]$, we have $H_+(x) \in \mathcal{RV}_{-\alpha}(x)$ and $H_-(x) \in \mathcal{RV}_{-\alpha'}(x)$ as $x \to \infty$ for some $\alpha, \alpha' > 1$. Let $\mathbb{D}_{j,k}$ be the set containing all step functions in $\mathbb{D}$ vanishing at the origin that has exactly j upward jumps and k downward jumps. As a convention, let $\mathbb{D}_{0,0} = \{\mathbf{0}\}$. Given $\alpha, \alpha' > 1$, let $\mathbb{D}_{<j,k} \triangleq \bigcup_{(l,m) \in \mathbb{I}_{<j,k}} \mathbb{D}_{l,m}$ where $\mathbb{I}_{<j,k} \triangleq \{(l,m) \in \mathbb{N}^2 \setminus \{(j,k)\} : l(\alpha-1) + m(\alpha'-1) \leq j(\alpha-1) + k(\alpha'-1)\}$. Let $\mathbf{C}_{0,0}$ be the Dirac measure on $\mathbf{0}$. For any $(j,k) \in \mathbb{N}^2 \setminus \{(0,0)\}$ let

$$\mathbf{C}_{j,k}(\cdot) \triangleq \mathbf{E}\left[\nu_\alpha^j \times \nu_{\alpha'}^k \left\{ (\boldsymbol{x}, \boldsymbol{y}) \in (0,\infty)^j \times (0,\infty)^k : \sum_{l=1}^{j} x_l \mathbf{I}_{[U_l, 1]} - \sum_{m=1}^{k} y_m \mathbf{I}_{[V_m, 1]} \in \cdot \right\} \right] \tag{6}$$

where all U_l's and V_m's are iid copies of Unif$(0,1)$ RVs. Now, we are ready to state the two-sided result.

Result 2 (Theorem 3.4 of Rhee et al. (2019)) *Let $A \subset \mathbb{D}$ be measurable. Suppose that*

$$\left(\mathcal{J}(A), \mathcal{K}(A) \right) \in \underset{(j,k) \in \mathbb{N}^2, \ \mathbb{D}_{j,k} \cap A \neq \emptyset}{argmin} j(\alpha-1) + k(\alpha'-1) \tag{7}$$

and A is bounded away from $\mathbb{D}_{<\mathcal{J}(A),\mathcal{K}(A)}$. Then the argument minimum in (7) is unique, and

$$
\begin{aligned}
\mathbf{C}_{\mathcal{J}(A),\mathcal{K}(A)}(A^\circ) &\leq \liminf_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n,\infty))^{\mathcal{J}(A)}(n\nu(-\infty,-n])^{\mathcal{K}(A)}} \\
&\leq \limsup_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n,\infty))^{\mathcal{J}(A)}(n\nu(-\infty,-n])^{\mathcal{K}(A)}} \leq \mathbf{C}_{\mathcal{J}(A),\mathcal{K}(A)}(A^-) < \infty
\end{aligned}
$$

where A°, A^- are the interior and closure of A respectively.

2.3. Concave Majorants and Stick-Breaking Approximations of Lévy Processes with Infinite Activities

Next, we review the distribution of the concave majorant of a Lévy process with infinite activities characterized in Pitman and Bravo (2012), which paves the way to the stick-breaking approximation algorithm proposed in González Cázares et al. (2022). Let $X(t)$ be a Lévy process with generating triplet (c,σ,ν). We say that X has **infinite activities** if $\sigma > 0$ or $\nu(\mathbb{R}) = \infty$. Let $M(t) \triangleq \sup_{s\leq t} X(s)$ be the running supremum of $X(t)$. The results in Pitman and Bravo (2012) establishes a Poisson–Dirichlet distribution that underlies the joint law of $X(t)$ and $M(t)$. Specifically, we fix some $T > 0$ and let V_i's be iid copies of Unif(0, 1) RVs. Recursively, let

$$
l_1 = TV_1, \qquad l_j = V_j \cdot (T - l_1 - l_2 - \ldots - l_{j-1}) \quad \forall j \geq 2. \tag{8}
$$

Conditioning on the values of $(l_j)_{j\geq 1}$, let ξ_j be a random copy of $X(l_j)$, with all ξ_j being independently generated.

Result 3 (Theorem 1 in Pitman and Bravo (2012)) *Suppose that the Lévy process X has infinite activities. Then (with $(x)^+ = \max\{x, 0\}$)*

$$
(X(T), M(T)) \overset{d}{=} \Big(\sum_{j\geq 1} \xi_j, \sum_{j\geq 1} (\xi_j)^+ \Big). \tag{9}
$$

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

13

Based on the distribution characterized in (9), the stick-breaking approximation algorithm was proposed in González Cázares et al. (2022) where finitely many ξ_i's are generated in order to approximate $X(T)$ and $M(T)$. This approximation technique is a key component of our rare event simulation algorithm. In particular, we utilize a coupling between different Lévy processes based on the representation (9) above. For clarity of our description, we focus on two Lévy processes X and $\widetilde{X}$ with generating triplets (c, σ, ν) and $(\widetilde{c}, \widetilde{\sigma}, \widetilde{\nu})$, respectively. Suppose that both X and $\widetilde{X}$ have infinite activities. We first generate l_i's as described in (8). Conditioning on the values of $(l_i)_{i \geq 1}$, we then independently generate ξ_i and $\widetilde{\xi}_i$, which are random copies of $X(l_i)$ and $\widetilde{X}(l_i)$, respectively. Let $\widetilde{M}(t) \triangleq \sup_{s \leq t} \widetilde{X}(s)$. Applying Result 3, we identify a coupling between $X(T), M(T), \widetilde{X}(T), \widetilde{M}(T)$ such that

$$\big(X(T), M(T), \widetilde{X}(T), \widetilde{M}(T)\big) \overset{d}{=} \big(\sum_{i \geq 1} \xi_i, \sum_{i \geq 1} (\xi_i)^+, \sum_{i \geq 1} \widetilde{\xi}_i, \sum_{i \geq 1} (\widetilde{\xi}_i)^+\big). \tag{10}$$

REMARK 1. It is worth noticing that the method described above in fact implies the existence of a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ that supports the entire sample paths $\{X(t): t \in [0, T]\}$ and $\{\widetilde{X}(t): t \in [0, T]\}$, whose endpoint values $X(T), \widetilde{X}(T)$ and suprema $M(T), \widetilde{M}(T)$ admit the joint law in (10). In particular, once we obtain l_i based on (8), one can generate Ξ_i that are iid copies of the entire paths of X. That is, we generate a piece of sample path Ξ_i on the stick l_i, and the quantities ξ_i introduced earlier can be obtained by setting $\xi_i = \Xi_i(l_i)$. To recover the sample path of X based on the pieces Ξ_i, it suffices to apply Vervatt transform onto each Ξ_i and then reorder the pieces based on their slopes. We refer the readers to theorem 4 in Pitman and Bravo (2012). In summary, the method described above leads to a coupling between the sample paths of the underlying Lévy processes X and $\widetilde{X}$ such that (10) holds.

2.4. Randomized Debiasing Technique   To achieve unbiasedness in our algorithm and remove the errors in the stick-breaking approximations, we apply the randomized multi-level Monte-Carlo technique studied in Rhee and Glynn (2015). In particular, due to τ being finite (almost surely) in Result 4 below, the simulation of Z relies only on $Y_0, Y_1, \cdots, Y_\tau$ instead of the infinite sequence $(Y_n)_{n \geq 0}$.

14

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Result 4 (Theorem 1 in Rhee and Glynn (2015)) *Let random variables Y and $(Y_m)_{m\geq 0}$ be such that* $\lim_{m\to\infty} \mathbf{E} Y_m = \mathbf{E} Y$. *Let τ be a positive integer-valued random variable with unbounded support, independent of $(Y_m)_{m\geq 0}$ and Y. Suppose that*

$$\sum_{m\geq 1} \mathbf{E}|Y_{m-1}-Y|^2 \big/ \mathbf{P}(\tau \geq m) < \infty, \tag{11}$$

then $Z \triangleq \sum_{m=0}^{\tau}(Y_m - Y_{m-1})\big/\mathbf{P}(\tau \geq m)$ (with the convention $Y_{-1}=0$) satisfies

$$\mathbf{E} Z = \mathbf{E} Y, \qquad \mathbf{E} Z^2 = \sum_{m\geq 0} \bar{v}_m \big/ \mathbf{P}(\tau \geq m)$$

where $\bar{v}_m = \mathbf{E}|Y_{m-1}-Y|^2 - \mathbf{E}|Y_m - Y|^2$.

3. Algorithm

Throughout the rest of this paper, let $X(t)$ be a Lévy process with generating triplet (c_X, σ, ν) satisfying the following heavy-tailed assumption.

ASSUMPTION 1. $\mathbf{E} X(1) = 0$. $X(t)$ *is of infinite activity. The Blumenthal-Getoor index $\beta \triangleq \inf\{p > 0 : \int_{(-1,1)} |x|^p \nu(dx) < \infty\}$ satisfies $\beta < 2$. Besides, one of the two claims below holds for the Lévy measure ν.*

- *(One-sided cases) ν is supported on $(0,\infty)$, and function $H_+(x) = \nu[x,\infty)$ is regularly varying as $x \to \infty$ with index $-\alpha$ where $\alpha > 1$;*

- *(Two-sided cases) There exist $\alpha, \alpha' > 1$ such that $H_+(x) = \nu[x,\infty)$ is regularly varying as $x \to \infty$ with index $-\alpha$ and $H_-(x) = \nu(-\infty,-x]$ is regularly varying as $x \to \infty$ with index $-\alpha'$.*

The other assumption on $X(t)$ revolves around the continuity of the law of $X^{<z}$, which is the Lévy process with generating triplet $(c_X, \sigma, \nu|_{(-\infty,z)})$. That is, $X^{<z}$ is a modulated version of X where all the upward jumps with size larger than z are removed.

ASSUMPTION 2. *There exist $z_0, C, \lambda > 0$ such that*

$$\mathbf{P}\big(X^{<z}(t) \in [x, x+\delta]\big) \leq \frac{C\delta}{t^\lambda \wedge 1} \qquad \forall z \geq z_0,\ t > 0,\ x \in \mathbb{R},\ \delta > 0.$$

Assumption 2 can be interpreted as a uniform version of Lipschitz continuity in the law of $X^{<z}(t)$. In Section 4, we show that Assumption 2 is a mild condition for Lévy process with infinite activities and is easy to verify.

Next, we describe a class of target events $(A_n)_{n \geq 1}$ for which we propose a strongly efficient rare event simulation algorithm. Let $\bar{X}_n(t) = \frac{1}{n}X(nt)$ and $\bar{X}_n = \{\bar{X}_n(t) : t \in [0, 1]\}$ be the scaled version of the process. Define events

$$A \triangleq \{\xi \in \mathbb{D} : \sup_{t \in [0,1]} \xi(t) \geq a; \sup_{t \in (0,1]} \xi(t) - \xi(t-) < b\}, \qquad A_n \triangleq \{\bar{X}_n \in A\}. \tag{12}$$

In words, $\xi \in A$ means that the path ξ crossed barrier a even though no upward jumps in ξ is larger than b. For technical reasons, we also impose the following mild condition on the values of the constants $a, b > 0$.

ASSUMPTION 3. $a, b > 0$ and $a/b \notin \mathbb{Z}$.

In this section, we present a strongly efficient rare-event simulation algorithm for $(A_n)_{n \geq 1}$. Specifically, Section 3.1 presents the design of the importance sampling distribution $\mathbf{Q}_n$, Section 3.2 discusses how we apply the randomized Monte-Carlo debiasing technique in Result 4 in our algorithm, Section 3.3 discusses how we combine the debiasing technique with SBA in Result 3, and Section 3.4 explains how to sample from the importance sampling distribution $\mathbf{Q}_n$. Combining all these components in Section 3.5, we propose Algorithm 2 for rare-event simulation of $\mathbf{P}(A_n)$ and establish its strong efficiency in Theorem 1. Section 3.6 addresses the case where the exact simulation of $X^{<z}(t)$ is not available.

3.1. Importance Sampling Distributions $\mathbf{Q}_n$ At the core of our algorithm is a principled design of importance sampling strategies based on heavy-tailed large deviations. This can be seen as an extension of the framework proposed in Chen et al. (2019). First, note that

$$l^* \triangleq \lceil a/b \rceil \tag{13}$$

16

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

indicates the number of jumps required to cross the barrier a starting from the origin if no jump is allowed to be larger than b. Based on the sample-path large deviations reviewed in Section 2.2, we expect the events $A_n = \{\bar{X}_n \in A\}$ to be almost always caused by exactly l^* large upward jumps in $\bar{X}_n$. These insights reveal critical information regarding the conditional law $\mathbf{P}(\,\cdot\,|\bar{X}_n \in A)$. More importantly, they lead to a natural yet effective choice of importance sampling distributions to focus on the l^*-large-jump paths and provides sufficient approximations to $\mathbf{P}(\,\cdot\,|\bar{X}_n \in A)$. Specifically, for any $\gamma \in (0, b)$, define events $B_n^\gamma \triangleq \{\bar{X}_n \in B^\gamma\}$ with

$$B^\gamma \triangleq \left\{\xi \in \mathbb{D} : \#\{t \in [0,1] : \xi(t) - \xi(t-) \geq \gamma\} \geq l^*\right\}, \tag{14}$$

where, for any $\xi \in \mathbb{D}$, we define $\xi(t-) = \lim_{s \uparrow t} \xi(s)$ as the left-limit of ξ at time t. Intuitively speaking, the parameter $\gamma \in (0, b)$ acts as a threshold of "large jumps": any path $\xi \in B^\gamma$ has at least l^* upward jumps that are considered large relative to the threshold level γ. To prevent the likelihood ratio from blowing up to infinity, we then consider an importance sampling distribution with defensive mixtures (see Hesterberg (1995)) and define (for some $w \in (0,1)$)

$$\mathbf{Q}_n(\cdot) \triangleq w\mathbf{P}(\cdot) + (1-w)\mathbf{P}(\,\cdot\,|B_n^\gamma). \tag{15}$$

Sampling from $\mathbf{P}(\,\cdot\,|B_n^\gamma)$, and hence $\mathbf{Q}_n(\cdot)$, is straightforward and will be addressed in Section 3.4.

With the design of the importance sampling distribution $\mathbf{Q}_n$ in hand, one would naturally consider an estimator for $\mathbf{P}(A_n)$ of form $\mathbf{I}_{A_n} \cdot \frac{d\mathbf{P}}{d\mathbf{Q}_n}$. This is due to

$$\mathbf{E}^{\mathbf{Q}_n}\left[\mathbf{I}_{A_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n}\right] = \mathbf{E}[\mathbf{I}_{A_n}] = \mathbf{P}(A_n).$$

Here, we use $\mathbf{E}^{\mathbf{Q}_n}$ to denote the expectation operator under law $\mathbf{Q}_n$ and $\mathbf{E}$ for the expectation under $\mathbf{P}$. Nevertheless, the exact evaluation or simulation of $\mathbf{I}_{A_n} = \mathbf{I}\{\bar{X}_n \in A\}$ is generally not computationally feasible due to the infinite activities of the process X, making it computationally infeasible to simulate or store the

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

17

entire sample path with finite computational resources. This marks a significant difference from the tasks in

Chen et al. (2019), which focus on random walks or compound Poisson processes with constant drifts that

can be simulated exactly. To overcome this challenge, we instead consider estimators L_n in the form of

$$L_n = Z_n \frac{d\mathbf{P}}{d\mathbf{Q}_n} = \frac{Z_n}{w + \frac{1-w}{\mathbf{P}(B_n^\gamma)} \mathbf{I}_{B_n^\gamma}} \tag{16}$$

where Z_n can be simulated within finite computational resources and allows L_n to recover the right expec-

tation under the importance sampling distribution $\mathbf{Q}_n$, i.e., $\mathbf{E}^{\mathbf{Q}_n}[L_n] = \mathbf{P}(A_n)$. In Section 3.2, we elaborate

on the design of the estimators Z_n.

3.2. Estimators Z_n Intuitively speaking, the goal is to construct Z_n's that can be plugged into (16)

as unbiased estimators of $\mathbf{I}_{A_n}$. To this end, we consider the following decomposition of the Lévy process X.

For any $\xi \in \mathbb{D}$ and $t \geq 0$, let $\Delta\xi(t) = \xi(t) - \xi(t-)$ be the size of the discontinuity in ξ at time t. Recall that

$\gamma \in (0, b)$ is the threshold of large jumps in the definition of B^γ in (14). Let

$$J_n(t) \triangleq \sum_{s \in [0,t]} \Delta X(s) \mathbf{I}\big(\Delta X(s) \geq n\gamma\big), \tag{17}$$

$$\Xi_n(t) \triangleq X(t) - J_n(t) = X(t) - \sum_{s \in [0,t]} \Delta X(s) \mathbf{I}\big(\Delta X(s) \geq n\gamma\big).$$

We highlight several important facts regarding the decomposition $X(t) = J_n(t) + \Xi_n(t)$.

- By the definition of $\mathbf{Q}_n$, the law of Ξ_n remains unchanged under both $\mathbf{Q}_n$ and $\mathbf{P}$, which is identical to

  the law of $X^{<n\gamma}$, namely, a Lévy process with generating triplet $(c_X, \sigma, \nu|_{(-\infty, n\gamma)})$.

- Under $\mathbf{P}$, the process J_n admits the law of a Lévy process with generating triplet $(0, 0, \nu|_{[n\gamma, \infty)})$, which

  is a compound Poisson process.

- Under $\mathbf{Q}_n$, the path $\{J_n(t) : t \in [0, n]\}$ follows the same law as a Lévy process with generating triplet

  $(0, 0, \nu|_{[n\gamma, \infty)})$, conditioned on having at least l^* jumps over $[0, n]$.

- Under both $\mathbf{P}$ and $\mathbf{Q}_n$, the two processes J_n and Ξ_n are independent.

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

18

Let $\bar{J}_n(t) = \frac{1}{n}J_n(nt)$, $\bar{J}_n = \{\bar{J}_n(t) : t \in [0,1]\}$, $\bar{\Xi}_n(t) = \frac{1}{n}\Xi_n(nt)$, and $\bar{\Xi}_n = \{\bar{\Xi}_n(t) : t \in [0,1]\}$. We now discuss how the decomposition

$$\bar{X}_n = \bar{J}_n + \bar{\Xi}_n$$

can help us construct unbiased estimators of $\mathbf{I}_{A_n}$. First, recall that $\gamma \in (0, b)$. As a result, in the definition of events $A_n = \{\bar{X}_n \in A\}$ in (12), the condition $\sup_{t \in (0,1]} \xi(t) - \xi(t-) < b$ only concerns the large jump process $\bar{J}_n$ since any upward jump in $\bar{\Xi}_n$ is bounded by $\gamma < b$. Therefore, with

$$E \triangleq \{\xi \in \mathbb{D} : \sup_{t \in (0,1]} \xi(t) - \xi(t-) < b\}, \qquad E_n \triangleq \{\bar{J}_n \in E\} \tag{18}$$

and

$$M(t) \triangleq \sup_{s \leq t} X(s), \qquad Y_n^* \triangleq \mathbf{I}(M(n) \geq na),$$

we have

$$\mathbf{I}_{A_n} = Y_n^* \mathbf{I}_{E_n}.$$

As discussed above, the exact evaluation of Y_n^* is generally not computationally possible. Instead, suppose that we have access to a sequence of random variables $(\hat{Y}_n^m)_{m \geq 0}$ that only take values in $\{0, 1\}$ and provide progressively more accurate approximations to Y_n^* as $m \to \infty$. Then in light of the debiasing technique in Result 4, one can consider (under the convention that $\hat{Y}_n^{-1} \equiv 0$)

$$Z_n = \sum_{m=0}^{\tau} \frac{\hat{Y}_n^m - \hat{Y}_n^{m-1}}{\mathbf{P}(\tau \geq m)} \mathbf{I}_{E_n} \tag{19}$$

where τ is $\text{Geom}(\rho)$ for some $\rho \in (0, 1)$ and is independent of everything else. That is, $\mathbf{P}(\tau \geq m) = \rho^{m-1}$ for all $m \geq 1$. Indeed, this construction of Z_n is justified by the following proposition. We defer the proof to Section 6.1.

PROPOSITION 1.   *Let $C_0 > 0$, $\rho_0 \in (0,1)$, $\mu > 2l^*(\alpha - 1)$, and $\bar{m} \in \mathbb{N}$. Suppose that*

$$\mathbf{P}\left(Y_n^* \neq \hat{Y}_n^m \,\Big|\, \mathcal{D}(\bar{J}_n) = k\right) \leq C_0 \rho_0^m \cdot (k+1) \qquad \forall k \geq 0, n \geq 1, m \geq \bar{m} \tag{20}$$

where $\mathcal{D}(\xi)$ counts the number of discontinuities in ξ for any $\xi \in \mathbb{D}$. Besides, suppose that for all $\Delta \in (0,1)$,

$$\mathbf{P}\left(Y_n^* \neq \hat{Y}_n^m, \ \bar{X}_n \notin A^\Delta \,\Big|\, \mathcal{D}(\bar{J}_n) = k\right) \leq \frac{C_0 \rho_0^m}{\Delta^2 n^\mu} \qquad \forall n \geq 1, m \geq 0, k = 0, 1, \cdots, l^* - 1, \tag{21}$$

where $A^\Delta = \left\{\xi \in \mathbb{D} : \sup_{t \in [0,1]} \xi(t) \geq a - \Delta\right\}$. Then given $\rho \in (\rho_0, 1)$, there exists some $\bar{\gamma} = \bar{\gamma}(\rho) \in (0, b)$ such

*that for all $\gamma \in (0, \bar{\gamma})$, the estimators $(L_n)_{n \geq 1}$ specified in (16) and (19) are **unbiased and strongly efficient***

for $\mathbf{P}(A_n) = \mathbf{P}(\bar{X}_n \in A)$ under the importance sampling distribution $\mathbf{Q}_n$ in (15).

3.3. Construction of $\hat{Y}_n^m$ In light of Proposition 1, our next goal is to design $\hat{Y}_n^m$'s that provide

sufficient approximations to $Y_n^* = \mathbf{I}(M(n) \geq na)$ and satisfy the conditions (20) and (21).

Recall the decomposition of $X(t) = \Xi_n(t) + J_n(t)$ in (17). Under both $\mathbf{Q}_n$ and $\mathbf{P}$, the processes Ξ_n and J_n are

independent, and Ξ_n admits the law of $X^{<n\gamma}$, i.e., a Lévy process with generating triplet $(c_X, \sigma, \nu|_{(-\infty, n\gamma)})$.

This section discusses how, after sampling J_n from $\mathbf{Q}_n$, we approximate the supremum of Ξ_n. Specifically,

on event $\{\mathcal{D}(\bar{J}_n) = k\}$, i.e., the process J_n makes k jumps over $[0, n]$, J_n admits the form of ζ_k with

$$\zeta_k(t) = \sum_{i=1}^{k} z_i \mathbf{I}_{[u_i, n]}(t) \qquad \forall t \in [0, n] \tag{22}$$

for some $z_i \in [n\gamma, \infty)$ and $u_1 < u_2 < \cdots < u_k$. This allows us to partition $[0, n]$ into $k + 1$ disjoint intervals

$[0, u_1), [u_1, u_2), \ldots, [u_{k-1}, u_k), [u_k, 1]$. We adopt the convention $u_0 \equiv 0, u_{k+1} \equiv 1$ and set

$$I_i = [u_{i-1}, u_i) \quad \forall i \in [k], \qquad I_{k+1} = [u_k, 1]. \tag{23}$$

20

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

For $\zeta_k = \sum_{i=1}^k z_i \mathbf{I}_{[u_i, n]}$, define

$$M_n^{(i),*}(\zeta_k) \triangleq \sup_{t \in I_i} \Xi_n(t) - \Xi_n(u_{i-1}) \tag{24}$$

as the supremum of the fluctuations of $\Xi_n(t)$ over I_i. Define random function

$$Y_n^*(\zeta_k) \triangleq \max_{i \in [k+1]} \mathbf{I}\Big(\Xi_n(u_{i-1}) + \zeta_k(u_{i-1}) + M_n^{(i),*}(\zeta_k) \geq na\Big), \tag{25}$$

and note that $Y_n^*(J_n) = \mathbf{I}(\sup_{t \in [0,n]} X(t) \geq na)$.

In theory, the representation (25) provides an algorithm for the simulation of $\mathbf{I}(\sup_{t \in [0,n]} X(t) \geq na)$. Nevertheless, the exact simulation of the supremum $M_n^{(i),*}(\zeta_k)$ is generally not available. Instead, we apply SBA introduced in Section 2.3 to approximate $M_n^{(i),*}(\zeta_k)$, thus providing the construction of $\hat{Y}_n^m$. Specifically, define

$$l_1^{(i)} = V_1^{(i)} \cdot (u_i - u_{i-1}); \tag{26}$$

$$l_j^{(i)} = V_j^{(i)} \cdot (u_i - u_{i-1} - l_1^{(i)} - l_2^{(i)} - \cdots - l_{j-1}^{(i)}) \qquad \forall j \geq 2 \tag{27}$$

where each $V_j^{(i)}$ is an iid copy of Unif$(0,1)$. That is, for each $i \in [k+1]$, the sequence $(l_j^{(i)})_{j \geq 1}$ is defined under the recursion in (8), with $T = u_i - u_{i-1}$ set as the length of I_i. Then, conditioning on the values of $l_j^{(i)}$'s, we sample

$$\xi_j^{(i)} \sim \mathbf{P}\Big(\Xi_n(l_j^{(i)}) \in \ \cdot\ \Big), \tag{28}$$

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

21

i.e., $\xi_j^{(i)}$ is an independent copy of $\Xi_n(l_j^{(i)})$, with all $\xi_j^{(i)}$ being independently generated. Result 3 then implies $\big(\Xi_n(u_i) - \Xi_n(u_{i-1}),\, M_n^{(i),*}(\zeta_k)\big) \overset{d}{=} \big(\sum_{j\geq 1} \xi_j^{(i)},\, \sum_{j\geq 1} (\xi_j^{(i)})^+\big)$ for each $i \in [k+1]$. Furthermore, by summing up only finitely many $\xi_j^{(i)}$'s, we define

$$\hat{M}_n^{(i),m}(\zeta_k) = \sum_{j=1}^{m+\lceil \log_2(n^d) \rceil} (\xi_j^{(i)})^+ \tag{29}$$

as an approximation to $M_n^{(i),*}(\zeta_k)$ defined in (24). Here, $d > 0$ is another parameter of the algorithm. For technical reasons, we add an extra $\lceil \log_2(n^d) \rceil$ term in the summation in (29), which helps ensure that the algorithm achieves strong efficiency as $n \to \infty$ while only introducing a minor increase in the computational complexity.

Now, we are ready to present the design of the approximators $\hat{Y}_n^m$. For $\zeta_k = \sum_{i=1}^k z_i \mathbf{I}_{[u_i,n]}$, define the random function

$$\hat{Y}_n^m(\zeta_k) = \max_{i \in [k+1]} \mathbf{I}\left(\sum_{q=1}^{i-1} \sum_{j\geq 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q + \hat{M}_n^{(i)}(\zeta_k) \geq na \right). \tag{30}$$

Here, note that $\sum_{q=1}^{i-1} z_q = \zeta_k(u_{i-1})$. As a high-level description, the algorithm proceeds as follows. After sampling J_n from the importance sampling distribution $\mathbf{Q}_n$ defined in (15), we plug $\hat{Y}_n^m(J_n)$ into Z_n defined in (19), which in turn allows us to simulate $L_n = Z_n \frac{d\mathbf{P}}{d\mathbf{Q}_n}$ as the importance sampling estimator under $\mathbf{Q}_n$.

REMARK 2.   At first glance, one may get the impression that the simulation of $\hat{Y}_n^m$ involves the summation of infinitely many elements in $\sum_{q=1}^{i-1} \sum_{j\geq 0} \xi_j^{(q),m}$. Fortunately, the truncation index τ in Z_n (see (19)) is almost surely finite. Therefore, when running the algorithm in practice, after τ is decided, there is no need to simulate any $\hat{Y}_n^m$ beyond $m \leq \tau$. Given the construction of $\hat{M}_n^{(i),m}(\zeta_k)$ in (29), the simulation of $\hat{Y}_n^m(\zeta_k)$ only requires (for each $i \in [k+1]$) $\xi_1^{(i)}, \xi_2^{(i)}, \ldots, \xi_{\tau+\lceil \log_2(n^d) \rceil}^{(i)}$, as well as the sum $\sum_{j\geq \tau+\lceil \log_2(n^d) \rceil+1} \xi_j^{(i)}$. Furthermore, conditioning on the value of $u_i - u_{i-1} - \sum_{j=1}^{\tau+\lceil \log_2(n^d) \rceil} \xi_j^{(i)} = t$, the sum $\sum_{j\geq \tau+\lceil \log_2(n^d) \rceil+1} \xi_j^{(i)}$ admits the law of $\Xi_n(t)$ (see Result 3). This allows us to simulate $\sum_{j\geq \tau+\lceil \log_2(n^d) \rceil+1} \xi_j^{(i)}$ in one shot.

22

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Note that to implement the importance sampling algorithm and ensure strong efficiency, the following tasks still remain to be addressed.

(*i*) As mentioned above, the evaluation of $\hat{Y}_n^m(J_n)$ requires the ability to first sample J_n from the importance sampling distribution $\mathbf{Q}_n$ defined in (15). We address this in Algorithm 1 proposed in Section 3.4. In summary, the simulation algorithm of estimators L_n is detailed in Algorithm 2.

(*ii*) The strong efficiency of the proposed algorithm needs to be justified by verifying the conditions in Proposition 1. This will be done in Section 3.5 by establishing Theorem 1.

(*iii*) Simulating $\xi_j^{(i)}$'s requires the exact simulation of $X^{<n\gamma}(t)$, which may not be computationally feasible in certain cases. To address this challenge, Section 3.6 proposes Algorithm 3, which builds upon Algorithm 2 and incorporates another layer of approximation via ARA.

3.4. Sampling from $\mathbf{P}(\,\cdot\,|B_n^\gamma)$ In this section, we revisit the task of sampling from $\mathbf{P}(\,\cdot\,|B_n^\gamma)$, which is at the core of the implementation of the importance sampling distribution $\mathbf{Q}_n$ in (15).

Recall that under $\mathbf{P}$, the process J_n is a compound Poisson process with generating triplet $(0, 0, \nu|_{[n\gamma,\infty)})$. More precisely, let $\widetilde{N}_n(\cdot)$ be a Poisson process with rate $\nu[n\gamma,\infty)$, and we use $(S_i)_{i\geq 1}$ to denote the arrival times of jumps in $\widetilde{N}_n(\cdot)$. Let $(W_i)_{i\geq 1}$ be a sequence of iid random variables from

$$\nu_n^{\text{normalized}}(\cdot) = \frac{\nu_n(\cdot)}{\nu[n\gamma,\infty)}, \qquad \nu_n(\cdot) = \nu\big(\,\cdot\,\cap[n\gamma,\infty)\big)$$

and let W_i's be independent of $\widetilde{N}_n(\cdot)$. Under $\mathbf{P}$, we have

$$J_n(t) \stackrel{d}{=} \sum_{i=1}^{\widetilde{N}_n(t)} W_i = \sum_{i\geq 1} W_i \mathbf{I}_{[S_i,\infty)}(t) \qquad \forall t \geq 0.$$

Furthermore, for each $k \geq 0$, conditioning on $\{\widetilde{N}_n(n) = k\}$, the law of $S_1, \ldots, S_k$ is equivalent to that of the order statistics of k iid samples from $\text{Unif}(0, n)$, and W_i's are still independent of S_i's with the law unaltered. Therefore, the sampling of J_n from $\mathbf{P}(\,\cdot\,|B_n^\gamma)$ can proceed as follows. We first generate k from the distribution of $\text{Poisson}(n\nu[n\gamma,\infty))$, conditioning on $k \geq l^*$. Then, independently, we generate $S_1, \cdots, S_k$ as the order statistics of k iid samples from $\text{Unif}(0, n)$, and $W_1, \cdots, W_k$ as iid samples of law $\nu_n^{\text{normalized}}(\cdot)$. It is worth

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

23

mentioning that the sampling of W_i can be addressed with the help of the inverse measure. Specifically,

define $Q_n^\leftarrow(y) \triangleq \inf\{s > 0 : \nu_n[s, \infty) < y\}$ as the inverse of ν_n, and observe that

$$y \leq \nu_n[s, \infty) \qquad \Longleftrightarrow \qquad Q_n^\leftarrow(y) \geq s.$$

More importantly, for $U \sim \text{Unif}(0, \nu_n[n\gamma, \infty))$, the law of $Q_n^\leftarrow(U)$ is $\nu_n^{\text{normalized}}(\cdot)$. This leads to the steps

detailed in Algorithm 1.

Algorithm 1 Simulation of J_n from $\mathbf{P}(\ \cdot\ |B_n^\gamma)$

Require: $n \in \mathbb{N}, l^* \in \mathbb{N}, \gamma > 0$, the Lévy measure ν.

  1: Sample k from a Poisson distribution with rate $n\nu[n\gamma, \infty)$ conditioning on $k \geq l^*$

  2: Simulate $\Gamma_1, \cdots, \Gamma_k \overset{\text{iid}}{\sim} Unif(0, \nu_n[n\gamma, \infty))$

  3: Simulate $U_1, \cdots, U_k \overset{\text{iid}}{\sim} Unif(0, n)$

  4: **Return** $J_n = \sum_{i=1}^{k} Q_n^\leftarrow(\Gamma_i)\mathbf{I}_{[U_i, n]}$

3.5. Strong Efficiency and Computational Complexity With all the discussions above, we

propose Algorithm 2 for rare-event simulation of $\mathbf{P}(A_n)$. Specifically, here is a list of the parameters of the

algorithm.

- $\gamma \in (0, b)$: the threshold in B^γ defined in (14),

- $w \in (0, 1)$: the weight of the defensive mixture in $\mathbf{Q}_n$; see (15),

- $\rho \in (0, 1)$: the geometric rate of decay for $\mathbf{P}(\tau \geq m)$ in (19),

- $d > 0$: determining the $\log_2(n^d)$ term in (29).

The choice of $w \in (0, 1)$ won't affect the strong efficiency of the algorithm. Meanwhile, under proper

parametrization, Algorithm 2 meets conditions (20) and (21) stated in Proposition 1 and attains strong

efficiency. This is verified in Theorem 1.

THEOREM 1. *Let $d > \max\{2, 2l^*(\alpha - 1)\}$ and $w \in (0, 1)$. There exists $\rho_0 \in (0, 1)$ such that the following*

*claim holds: Given $\rho \in (\rho_0, 1)$, there exists $\bar{\gamma} \in (0, b)$ such that Algorithm 2 is **unbiased and strongly efficient***

under any $\gamma \in (0, \bar{\gamma})$.

24

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

We defer the proof to Section 6.2. In fact, in Section 3.6 we propose Algorithm 3, which can be seen as an extended version of Algorithm 2 with another layer of approximation. The strong efficiency of Algorithm 2 follows directly from that of Algorithm 3 (i.e., by setting $\kappa = 0$ in the proof of Theorem 2). The choices of $\bar{\gamma}$ and $\bar{\rho}$ that ensure strong efficiency are also specified at the end of Section 3.6.

REMARK 3. To conclude, we add a remark regarding the computational complexity of Algorithm 2 under the goal of attaining a given level of relative error at a specified confidence level. First, consider the case where the complexity of simulation of $X^{<z}(t)$ scales linearly with t (uniformly for all $z \in [z_0, \infty]$ for some constant z_0). This is a standard since the number of jumps we expect to simulate over $[0, t]$ grows linearly with t. Then, the complexity of the SBA steps at step 13 of Algorithm 2 also scales linearly with n, as the stick lengths of $l_j^{(i)}$'s, in expectation, grow linearly with n because we deal with the time horizon $[0, n]$ given the scale factor n. Next, since the same law for the truncation index τ (see step 8 of Algorithm 2) is applied for all n, the only other factor that is varying with n is $t_n = \lceil \log_2(n^d) \rceil$ in the loop at step 10. The strong efficiency of the algorithm then implies a computational complexity of order $O(n \cdot \log_2 n)$. If we instead assume that the cost of simulating $X^{<z}(t)$ is also uniformly bounded for all t, then the overall complexity of Algorithm 2 is further reduced to $O(\log_2 n)$.

In comparison, the crude Monte Carlo method requires a number of samples that is inversely proportional to the target probability $\mathbf{P}(A_n) \approx O(1/n^{l^*(\alpha-1)})$ (see Lemma 1) with $\alpha > 1$ being the heavy-tailed index in Assumption 1 and $l^* \geq 1$ defined in (13). Hypothetically, assuming that the evaluation of $\mathbb{I}_{A_n}$ (which at least requires the simulation of $X(t)$ and $M(t)$) is computationally feasible at a cost that scales linearly with n, we end up with a computational complexity of $O(n \cdot n^{l^*(\alpha-1)})$ (compared to the $O(n \cdot \log_2 n)$ cost of our algorithm). Similarly, if we assume that the cost of generating $\mathbb{I}_{A_n}$ is uniformly bounded for all n, then the complexity of the crude Monte-Carlo method is $O(n^{l^*(\alpha-1)})$ (compared to the $O(\log_2 n)$ cost of our algorithm). In summary, not only does the proposed importance sampling algorithm address Lévy processes with infinite activities that are not simulatable for crude Monte Carlo methods, but it also enjoys a significant improvement in terms of computational complexity, with the advantage becoming even more evident for multiple-jump events with large l^*.

Algorithm 2 Strongly Efficient Estimation of $\mathbf{P}(A_n)$

Require: $w \in (0,1)$, $\gamma > 0$, $d > 0$, $\rho \in (0,1)$ as the parameters of the algorithm; $a, b > 0$ as the characterization of the set A; (c_X, σ, ν) as the generating triplet of $X(t)$.

1:   Set $t_n = \lceil \log_2(n^d) \rceil$

2:   **if** $\mathrm{Unif}(0,1) < w$ **then**   ▷ Sample J_n from $\mathbf{Q}_n$
3:      Sample $J_n = \sum_{i=1}^k z_i \mathbf{I}_{[u_i, n]}$ from $\mathbf{P}$
4:   **else**
5:      Sample $J_n = \sum_{i=1}^k z_i \mathbf{I}_{[u_i, n]}$ from $\mathbf{P}(\,\cdot\,|B_n^\gamma)$ using Algorithm 1
6:   **end if**
7:   Set $u_0 = 0, u_{k+1} = n$.

8:   Sample $\tau \sim \mathrm{Geom}(\rho)$ ▷ Decide Truncation Index τ

9:   **for** $i = 1, 2, \ldots, k+1$ **do**   ▷ Stick-Breaking Procedure
10:      **for** $j = 1, 2, \ldots, t_n + \tau$ **do**
11:          Sample $V_j^{(i)} \sim \mathrm{Unif}(0,1)$
12:          Set $l_j^{(i)} = V_j^{(i)}(u_i - u_{i+1} - l_1^{(i)} - l_2^{(i)} - \ldots - l_{j-1}^{(i)})$
13:          Sample $\xi_j^{(i)} \sim \mathbf{P}(X^{<n\gamma}(l_j^{(i)}) \in \,\cdot\,)$
14:      **end for**
15:      Set $l_{t_n+\tau+1}^{(i)} = u_i - u_{i-1} - l_1^{(i)} - l_2^{(i)} - \ldots - l_{t_n+\tau}^{(i)}$
16:      Sample $\xi_{t_n+\tau+1}^{(i)} \sim \mathbf{P}(X^{<n\gamma}(l_{t_n+\tau+1}^{(i)}) \in \,\cdot\,)$
17:   **end for**

18:   **for** $m = 1, \cdots, \tau$ **do**   ▷ Evaluate $\hat{Y}_n^m$
19:      **for** $i = 1, 2, \ldots, k+1$ **do**
20:          Set $\hat{M}_n^{(i),m} = \sum_{q=1}^{i-1} \sum_{j=1}^{t_n+\tau+1} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q + \sum_{j=1}^{t_n+m} (\xi_j^{(i)})^+$
21:      **end for**
22:      Set $\hat{Y}_n^m = \mathbf{I}\{\max_{i=1,\ldots,k+1} \hat{M}_n^{(i),m} \geq na\}$
23:   **end for**
24:   Set $Z_n = \hat{Y}_n^1 + \sum_{m=2}^\tau (\hat{Y}_n^m - \hat{Y}_n^{m-1})/\rho^{m-1}$ ▷ Return the Estimator L_n
25:   **if** $\max_{i=1,\cdots,k} z_i > b$ **then**
26:      Return $L_n = 0$.
27:   **else**
28:      Set $\lambda_n = n\nu[n\gamma, \infty)$, $p_n = 1 - \sum_{l=0}^{l^*-1} e^{-\lambda_n} \frac{\lambda_n^l}{l!}$, $I_n = \mathbf{I}\{J_n \in B_n^\gamma\}$
29:      Return $L_n = Z_n/(w + \frac{1-w}{p_n} I_n)$
30:   **end if**

3.6. Construction of $\hat{Y}_n^m$ with ARA

As stressed earlier, implementing Algorithm 3.5 requires the ability to sample from $\mathbf{P}(X^{<n\gamma}(t) \in \,\cdot\,)$. The goal of this section is to address the challenge when the exact simulation of $X^{<n\gamma}(t)$ is not available. The plan is to incorporate the Asmussen-Rosiński approximation (ARA) in Asmussen and Rosiński (2001) into the design of the approximation $\hat{Y}_n^m$ proposed in Section 3.3.

To be specific, let

$$\kappa_{n,m} \triangleq \frac{\kappa^m}{n^r} \qquad \forall n \geq 1, \ m \geq 0 \tag{31}$$

where $\kappa \in (0,1)$ and $r > 0$ are two additional parameters of our algorithm. As a convention, we set $\kappa_{n,-1} \equiv 1$. Without loss of generality, we consider n large enough such that $n\gamma > 1 = \kappa_{n,-1}$. For the Lévy process

$\Xi_n = X^{<n\gamma}$ with the generating triplet $(c_X, \sigma, \nu|_{(-\infty, n\gamma)})$, consider the following decomposition (with $B(t)$ being a standard Brownian motion)

$$\Xi_n(t) = c_X t + \sigma B(t) + \underbrace{\sum_{s \le t} \Delta X(s) \mathbf{I}\Big(\Delta X(s) \in (-\infty, -1] \cup [1, n\gamma)\Big)}_{\triangleq J_{n,-1}(t)} \tag{32}$$

$$+ \sum_{m \ge 0} \underbrace{\left[\sum_{s \le t} \Delta X(s) \mathbf{I}\Big(|\Delta X(s)| \in [\kappa_{n,m}, \kappa_{n,m-1})\Big) - t \cdot \nu\Big((-\kappa_{n,m-1}, -\kappa_{n,m}] \cup [\kappa_{n,m}, \kappa_{n,m-1})\Big) \right]}_{\triangleq J_{n,m}(t)}.$$

Here, for any $m \ge 0$, $J_{n,m}$ is a martingale with $var[J_{n,m}(1)] = \bar{\sigma}^2(\kappa_{n,m-1}) - \bar{\sigma}^2(\kappa_{n,m})$ where

$$\bar{\sigma}^2(c) \triangleq \int_{(-c,c)} x^2 \nu(dx) \qquad \forall c \in (0,1]. \tag{33}$$

Generally speaking, the difficulty of implementing Algorithm 3.5 lies in the exact simulation of the martingale $\sum_{m \ge 0} J_{n,m}$. In particular, whenever we have $\nu((-\infty, 0) \cup (0, \infty)) = \infty$ for the Lévy measure ν, the expected number of jumps in $\sum_{m \ge 0} J_{n,m}$ (and hence $X^{<n\gamma}$ and X) will be infinite over any time interval with positive length. By applying ARA, our goal is to approximate the jump martingale $J_{n,m}$'s using Brownian motions, which yields a process that is amenable to exact simulation. To do so, let $(W^m)_{m \ge 1}$ be a sequence of iid copies of standard Brownian motions, which are also independent of $B(t)$. For each $m \ge 0$, define

$$\breve{\Xi}_n^m(t) \triangleq c_X t + \sigma B(t) + \sum_{q=-1}^{m} J_{n,q}(t) + \sum_{q \ge m+1} \sqrt{\bar{\sigma}^2(\kappa_{n,q-1}) - \bar{\sigma}^2(\kappa_{n,q})} \cdot W^q(t). \tag{34}$$

Here, the process $\breve{\Xi}_n^m$ can be interpreted as an approximation to Ξ_n, where the jump martingale (with jump sizes under $\kappa_{n,m}$) is substituted by a standard Brownian motion with the same variance. Note that for any $t > 0$, the random variable $\breve{\Xi}_n^m(t)$ is exactly simulatable, as it is a convolution of a compound Poisson process with constant drift and a Gaussian random variable.

Based on the approximations $\breve{\Xi}_n^m$ in (34), we apply SBA and reconstruct $\hat{M}_n^{(i),m}$ (originally defined in (29))

and $\hat{Y}_n^m$ (originally defined in (30)) as follows. Let $\zeta_k(t) = \sum_{i=1}^k z_i \mathbf{I}_{[u_i,n]}(t)$ be a piece-wise step function

with k jumps over $(0,n]$, i.e., admitting the form in (22). Recall that the jump times in ζ_k leads to a partition

of $[0,n]$ of $(I_i)_{i\in[k+1]}$ defined in (23). For any I_i, let the sequence $l_j^{(i)}$'s be defined as in (26)–(27). Next,

conditioning on $(l_j^{(i)})_{j\geq 1}$, one can sample $\xi_j^{(i),m}, \xi_j^{(i)}$ as

$$\left(\xi_j^{(i)}, \xi_j^{(i),0}, \xi_j^{(i),1}, \xi_j^{(i),2}, \ldots\right) \stackrel{d}{=} \left(\Xi_n(l_j^{(i)}), \breve{\Xi}_n^0(l_j^{(i)}), \breve{\Xi}_n^1(l_j^{(i)}), \breve{\Xi}_n^2(l_j^{(i)}), \ldots\right). \tag{35}$$

The coupling in (10) then implies

$$\left(\Xi_n(u_i) - \Xi_n(u_{i-1}), \sup_{t\in I_i}\Xi_n(t) - \Xi_n(u_{i-1}), \breve{\Xi}_n^0(u_i) - \breve{\Xi}_n^0(u_{i-1}), \sup_{t\in I_i}\breve{\Xi}_n^0(t) - \breve{\Xi}_n^0(u_{i-1}), \tag{36}\right.$$

$$\left.\breve{\Xi}_n^1(u_i) - \breve{\Xi}_n^1(u_{i-1}), \sup_{t\in I_i}\breve{\Xi}_n^1(t) - \breve{\Xi}_n^1(u_{i-1}), \ldots\right)$$

$$\stackrel{d}{=} \left(\sum_{j\geq 1}\xi_j^{(i)}, \sum_{j\geq 1}(\xi_j^{(i)})^+, \sum_{j\geq 1}\xi_j^{(i),0}, \sum_{j\geq 1}(\xi_j^{(i),0})^+, \sum_{j\geq 1}\xi_j^{(i),1}, \sum_{j\geq 1}(\xi_j^{(i),1})^+, \ldots\right).$$

Now, we define

$$\hat{M}_n^{(i),m}(\zeta_k) = \sum_{j=1}^{m+\lceil\log_2(n^d)\rceil}(\xi_j^{(i),m})^+ \tag{37}$$

as an approximation to $M_n^{(i),*}(\zeta_k) = \sup_{t\in I_i}\Xi_n(t) - \Xi_n(u_{i-1}) = \sum_{j\geq 1}(\xi_j^{(i)})^+$ using both ARA and SBA.

Compared to the original design in (29), the main difference in (37) is that we substitute $\xi_j^{(i)}$ with $\xi_j^{(i),m}$, and

the latter is exactly simulatable as, conditioning on the values of $l_j^{(i)}$'s, it admits the law of $\breve{\Xi}_n^m$. Similarly, let

$$\hat{Y}_n^m(\zeta_k) = \max_{i\in[k+1]}\mathbf{I}\left(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q),m} + \sum_{q=1}^{i-1}z_q + \hat{M}_n^{(i),m}(\zeta_k) \geq na\right); \tag{38}$$

28

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Again, the main difference between (38) and (30) is that we incorporate ARA and substitute $\xi_j^{(i)}$'s with $\xi_j^{(i),m}$'s.

Plugging the design of $\hat{Y}_n^m(\zeta_k)$ in (38) into the estimator Z_n in (19), we propose Algorithm 3 for rare-event simulation of $\mathbf{P}(A_n)$ when exact simulation of $X^{<n\gamma}$ is not available. Below is a summary of the parameters of the algorithm.

- $\gamma \in (0, b)$: the threshold in B^γ defined in (14),

- $w \in (0, 1)$: the weight of the defensive mixture in $\mathbf{Q}_n$; see (15),

- $\rho \in (0, 1)$: the geometric rate of decay for $\mathbf{P}(\tau \geq m)$ in (19),

- $\kappa \in [0, 1)$, $r > 0$: determining the truncation threshold $\kappa_{n,m}$ in (31),

- $d > 0$: determining the $\log_2(n^d)$ term in (37).

Theorem 2 justifies that, under proper parametrization, Algorithm 3 is unbiased and strongly efficient.

THEOREM 2. *Let $\mu > 2l^*(\alpha - 1)$ and $\beta_+ \in (\beta, 2)$ where $\alpha > 1$ is the heavy-tail index and $\beta \in (0, 2)$ is the Blumenthal-Getoor index in Assumption 1. Let $w \in (0, 1)$ and*

$$\kappa^{2-\beta_+} < \frac{1}{2}, \qquad r(2 - \beta_+) > \max\{2, \mu - 1\}, \qquad d > \max\{2, 2\mu - 1\}. \tag{39}$$

*There exists $\rho_0 \in (0, 1)$ such that the following claim holds: Given $\rho \in (\rho_0, 1)$, there exists $\bar{\gamma} \in (0, b)$ such that Algorithm 3 is **unbiased and strongly efficient** under any $\gamma \in (0, \bar{\gamma})$.*

In Section 6.2 we provide the proof, the key arguments of which are the verification of conditions (20) and (21) in Proposition 1. Here, we specify the choices of the parameters. First, pick $\alpha_3 \in (0, \frac{1}{\lambda})$, $\alpha_4 \in (0, \frac{1}{2\lambda})$ where $\lambda > 0$ is the constant in Assumption 2. Next, pick $\alpha_2 \in (0, \frac{\alpha_3}{2} \wedge 1)$ and $\alpha_1 \in (0, \frac{\alpha_2}{\lambda})$. Also, fix $\delta \in (1/\sqrt{2}, 1)$. This allows us to pick $\rho_0 \in (0, 1)$ such that

$$\rho_0 > \max\left\{\delta^{\alpha_1}, \ \frac{\kappa^{2-\beta_+}}{\delta^2}, \ \frac{1}{\sqrt{2}\delta}, \ \delta^{\alpha_2 - \lambda\alpha_1}, \ \delta^{1 - \lambda\alpha_3}, \delta^{-\alpha_2 + \frac{\alpha_3}{2}}\right\}.$$

After picking $\rho \in (\rho_0, 1)$, one can find some $q > 1$ such that $\rho_0^{1/q} < \rho$. Let $p > 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. Let $\Delta > 0$ be small enough such that $a - \Delta > (l^* - 1)b$. Then, we pick $\bar{\gamma} \in (0, b)$ small enough such that

$$\frac{a - \Delta - (l^* - 1)b}{\bar{\gamma}} + l^* - 1 > 2l^*p.$$

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

29

Again, the details of the parameter choices can be found at the beginning of Section 6. It is also worth mentioning that, by setting $\kappa = 0$, Algorithm 3 would reduce to Algorithm 2, as $\xi_j^{(i),m}$'s in (36) would reduce to $\xi_j^{(i)}$'s in (28); in other words, the ARA mechanism is effective only if the truncation threshold $\kappa_{n,m} = \kappa^m/n^r > 0$. As a result, Theorem 1 follows directly from Theorem 2 by setting $\kappa = 0$.

REMARK 4. While Algorithm 3 terminates within finite steps almost surely, its computational complexity may not be finite in expectation. This is partially due to the implementation of ARA as we approximate the jump martingale $J_{n,m}(t)$ in (32) using a independent Brownian motion term in (34). In theory, a potential remedy is to identify a better coupling between the jump martingales and Brownian motions; see, for instance, Theorem 9 of Mariucci and Reiß (2018). This would allow us to pick a larger κ for the truncation threshold $\kappa_{n,m}$ in ARA, under which the simulation algorithm generates significantly fewer jumps when sampling $\xi_j^{(i),m}$'s. However, to the best of our knowledge, there is no practical implementation of the coupling in Mariucci and Reiß (2018). We note that similar issues arise in works such as González Cázares and Mijatović (2022), where the coupling in Mariucci and Reiß (2018) imply a much tighter error bound in theory but cannot be implemented in practice.

4. Lipschitz Continuity of the Distribution of $X^{<z}(t)$

This section investigates the sufficient conditions for Assumption 2. That is, there exist z_0, C, $\lambda > 0$ such that

$$\mathbf{P}\big(X^{<z}(t) \in [x, x+\delta]\big) \le \frac{C\delta}{t^\lambda \wedge 1} \qquad \forall z \ge z_0,\ t > 0,\ x \in \mathbb{R},\ \delta > 0. \tag{40}$$

Here, recall that $X^{>z}$ is the Lévy process with generating triplet $(c_X, \sigma, \nu|_{(-\infty,z)})$. In other words, this is a modulated version of X where any the upward jump larger than z is removed.

30

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Algorithm 3 Strongly Efficient Estimation of $\mathbf{P}(A_n)$ with ARA

Require: $w \in (0,1)$, $\gamma > 0$, $r > 0$, $d > 0$, $\kappa \in [0,1)$, $\rho \in (0,1)$ as the parameters in the algorithm; $a, b > 0$ as the characterization of the set A; (c_X, σ, ν) as the generating triplet of $X(t)$; $\bar{\sigma}(\cdot)$ is defined in (33).

1: Set $t_n = \lceil \log_2(n^d) \rceil$ and $\kappa_{n,m} = \kappa^m / n^r$

2: **if** $\text{Unif}(0,1) < w$ **then**   ▷ Sample J_n from $\mathbf{Q}_n$
3:     Sample $J_n = \sum_{i=1}^k z_i \mathbf{I}_{[u_i, n]}$ from $\mathbf{P}$
4: **else**
5:     Sample $J_n = \sum_{i=1}^k z_i \mathbf{I}_{[u_i, n]}$ from $\mathbf{P}(\cdot \,|B_n^\gamma)$ using Algorithm 1
6: **end if**
7: Set $u_0 = 0, u_{k+1} = n$.

8: Sample $\tau \sim Geom(\rho)$ ▷ Decide Truncation Index τ

9: **for** $i = 1, 2, \ldots, k+1$ **do**   ▷ Stick-Breaking Procedure
10:     **for** $j = 1, 2, \ldots, t_n + \tau$ **do**
11:         Sample $V_j^{(i)} \sim \text{Unif}(0,1)$
12:         Set $l_j^{(i)} = V_j^{(i)} (u_i - u_{i+1} - l_1^{(i)} - l_2^{(i)} - \ldots - l_{j-1}^{(i)})$
13:     **end for**
14:     Set $l_{t_n+\tau+1}^{(i)} = u_i - u_{i-1} - l_1^{(i)} - l_2^{(i)} - \ldots - l_{t_n+\tau}^{(i)}$
15: **end for**

16: **for** $i = 1, \cdots, k+1$ **do**   ▷ Sample $\xi_j^{(i),m}$
17:     **for** $j = 1, 2, \cdots, t_n + \tau + 1$ **do**
18:         Sample $x_j^{(i)} \sim N(0, \sigma^2 \cdot l_j^{(i)})$
19:         Sample $y_j^{(i),-1} \sim \mathbf{P}(J_{n,-1}(l_j^{(i)}) \in \cdot)$
20:         **for** $m = 0, 1, \ldots, \tau$ **do**
21:             Sample $y_j^{(i),m} \sim \mathbf{P}(J_{n,m}(l_j^{(i)}) \in \cdot)$
22:             Sample $w_j^{(i),m} \sim N(0, (\bar{\sigma}^2(\kappa_{n,m-1}) - \bar{\sigma}^2(\kappa_{n,m})) \cdot l_j^{(i)})$
23:         **end for**
24:         Sample $w_j^{(i),\tau+1} \sim N(0, \bar{\sigma}^2(\kappa_{n,\tau}) \cdot l_j^{(i)})$
25:         **for** $m = 0, \ldots, \tau$ **do**
26:             Set $\xi_j^{(i),m} = c_X \cdot l_j^{(i)} + x_j^{(i)} + \sum_{q=-1}^m y_j^{(i),q} + \sum_{q=m+1}^{\tau+1} w_j^{(i),q}$
27:         **end for**
28:     **end for**
29: **end for**

30: **for** $m = 1, \cdots, \tau$ **do**   ▷ Evaluate $\hat{Y}_n^m$
31:     **for** $i = 1, 2, \ldots, k+1$ **do**
32:         Set $\hat{M}_n^{(i),m} = \sum_{q=1}^{i-1} \sum_{j=1}^{t_n+\tau+1} \xi_j^{(q),m} + \sum_{q=1}^{i-1} z_q + \sum_{j=1}^{t_n+m} (\xi_j^{(i),m})^+$
33:     **end for**
34:     Set $\hat{Y}_n^m = \mathbf{I}\{ \max_{i=1,\ldots,k+1} \hat{M}_n^{(i),m} \geq na \}$
35: **end for**

36: Set $Z_n = \hat{Y}_n^1 + \sum_{m=2}^\tau (\hat{Y}_n^m - \hat{Y}_n^{m-1}) / \rho^{m-1}$ ▷ Return the Estimator L_n
37: **if** $\max_{i=1,\cdots,k} z_i > b$ **then**
38:     Return $L_n = 0$.
39: **else**
40:     Set $\lambda_n = n\nu[n\gamma, \infty)$, $p_n = 1 - \sum_{l=0}^{l^*-1} e^{-\lambda_n} \frac{\lambda_n^l}{l!}$, $I_n = \mathbf{I}\{J_n \in B_n^\gamma\}$
41:     Return $L_n = Z_n / (w + \frac{1-w}{p_n} I_n)$
42: **end if**

To demonstrate our approach for establishing condition (40), we start by considering a simple case where

the Lévy process $X(t)$ has generating tripet (c_X, σ, ν) with $\sigma > 0$. This leads to the decomposition

$$X^{<z}(t) \overset{d}{=} \sigma B(t) + Y^{<z}(t) \qquad \forall t, z > 0$$

where B is a standard Brownian motion, $Y^{<z}$ is a Lévy process with generating triplet $(c_X, 0, \nu|_{(-\infty, z)})$, and the two processes are independent. Now, for any $x \in \mathbb{R}$, $t > 0$ and $\delta \in (0, 1)$,

$$
\begin{aligned}
\mathbf{P}(X^{<z}(t) \in [x, x+\delta]) &= \int_{\mathbb{R}} \mathbf{P}(\sigma B(t) \in [x-y, x-y+\delta]) \cdot \mathbf{P}(Y^{<z}(t) \in dy) \\
&= \int_{\mathbb{R}} \mathbf{P}\left(\frac{B(t)}{\sqrt{t}} \in \left[\frac{x-y}{\sigma\sqrt{t}}, \frac{x-y+\delta}{\sigma\sqrt{t}}\right]\right) \cdot \mathbf{P}(Y^{<z}(t) \in dy) \\
&\leq \frac{1}{\sigma\sqrt{2\pi}} \cdot \frac{\delta}{\sqrt{t}}.
\end{aligned}
\tag{41}
$$

The last inequality follows from the fact that a standard Normal distribution admits a density function bounded by $1/\sqrt{2\pi}$. Therefore, we verified Assumption 2 under $\lambda = 1/2, C = \frac{1}{\sigma\sqrt{2\pi}}$, and any $z_0 > 0$. The simple idea behind (41) is that continuity conditions such as (40) can be passed from one distribution to another through convolutional structures. To generalize this approach to the scenarios where $\sigma = 0$ in the generating triplet of the Lévy process X, we introduce the following definition.

DEFINITION 3.   Let μ_1 and μ_2 be Borel measures on $\mathbb{R}$. For any Borel set $A \subset \mathbb{R}$, we say that μ_1 **majorizes** μ_2 **when restricted on** A (denoted as $(\mu_1 - \mu_2)|_A \geq 0$) if $\mu(B \cap A) = \mu_1(B \cap A) - \mu_2(B \cap A) \geq 0$ for any Borel set $B \subset \mathbb{R}$. In other words $\mu|_A = (\mu_1 - \mu_2)|_A$ is a **positive** measure.

Now, let us consider the case where the generating triplet of X is $(c_X, 0, \nu)$. For the Lévy measure ν, if we can find some $z_0 > 0$, some Borel set $A \subseteq (-\infty, z_0)$ and some (positive) Borel measure μ such that $(\nu - \mu)|_A \geq 0$, then through a straightforward superposition of Poisson random measures, we obtain the decomposition (let $\mu_A = \mu|_A$)

$$
X^{<z}(t) \overset{d}{=} Y(t) + \widetilde{X}^{<z, -A}(t) \qquad \forall z \geq z_0
\tag{42}
$$

where $Y(t)$ is a Lévy process with generating triplet $(0, 0, \mu_A)$, $\widetilde{X}^{<z, -A}(t)$ is a Lévy process with generating triplet $(c_X, 0, \nu - \mu_A)$, and the two processes are independent. Furthermore, if Assumption 2 (conditions of form (40)) holds for the process $Y(t)$ with generating triplet $(0, 0, \mu_A)$, then by repeating the arguments in (41) we can show that Assumption 2 holds in $X^{<z}(t)$ for any $z \geq z_0$.

32

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Recall our running assumption that the Lévy process $X(t)$ is of infinite activities (see Assumption 1). In case that $\sigma = 0$, we must have $\nu((-\infty, 0) \cup (0, \infty)) = \infty$ for X to have infinite activity. Therefore, the key step is to identify the majorized measure μ such that

- $(\nu - \mu)|_A \geq 0$ holds for ν with infinite mass and some set A,

- condition (40) holds for the Lévy process $Y(t)$ in (42) with generating triplet $(0, 0, \mu|_A)$.

In the first main result of this section, we show that measures of form $\mu[x, \infty)$ that roughly increase at a power-law rate $1/x^\alpha$ (as $x \downarrow 0$) provide ideal choices for such majorized measures. In particular, the corresponding Lévy process $Y(t)$ in (42) is intimately related to α-stable processes that naturally satisfy continuity properties of form (40). The nature of the proof is less probabilistic and is independent of the main results of the paper (i.e., strong efficiency), so we collect it in Section A of the Supplementary Materials.

PROPOSITION 2. *Let $\alpha \in (0, 2), z_0 > 0$, and $\epsilon \in (0, (2-\alpha)/2)$. Suppose that $\mu[x, \infty)$ is regularly varying as $x \downarrow 0$ with index $-(\alpha + 2\epsilon)$. Then the Lévy process $Y(t)$ with generating triplet $(0, 0, \mu|_{(0,z_0)})$ has a continuous density function $f_{Y(t)}$ for each $t > 0$. Furthermore, there exists a constant $C < \infty$ such that*

$$\left\| f_{Y(t)} \right\|_\infty \leq \frac{C}{t^{1/\alpha} \wedge 1} \qquad \forall t > 0.$$

where $\|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$.

Equipped with Proposition 2, we obtain the following set of sufficient conditions for Assumption 2.

THEOREM 3. *Let (c_X, σ, ν) be the generating triplet of Lévy process X.*

(i) If $\sigma > 0$, then Assumption 2 holds for $\lambda = 1/2$ and any $z_0 > 0$.

(ii) If there exist Borel measure μ, some $z_0 > 0$, and some $\alpha' \in (0, 2)$ such that $(\nu - \mu)|_{(0,z_0)} \geq 0$ (resp., $(\nu - \mu)|_{(-z_0,0)} \geq 0$) and $\mu[x, \infty)$ (resp., $\mu(-\infty, x]$) is regularly varying with index $-\alpha'$ as $x \downarrow 0$, then Assumption 2 holds with $\lambda = 1/\alpha$ for any $\alpha \in (0, \alpha')$.

Proof Part (*i*) follows immediately from the calculations in (41). To prove part (*ii*), we fix some $\alpha \in (0, \alpha')$, and without loss of generality assume that $(\nu - \mu)|_{(0,z_0)} \geq 0$ and $\mu[x, \infty)$ is regularly varying with index α' as $x \downarrow 0$. This allows us to fix some $\epsilon = (\alpha' - \alpha)/2 \in (0, (2-\alpha)/2)$.

For any $z \geq z_0$, let $Y(t)$ and $\widetilde{X}^{<z,-A}(t)$ be defined as in (42) with $A = (0, z_0)$. First of all, applying Proposition 2, we can find $C > 0$ such that $\left\| f_{Y(t)} \right\|_\infty \leq \frac{C}{t^{1/\alpha \wedge 1}}$ $\forall t > 0$. Next, due to the independence between Y and $\widetilde{X}^{<z,-A}(t)$, it holds for all $x \in \mathbb{R}, \delta \geq 0$, and $t > 0$ that

$$\mathbf{P}(X^{<z}(t) \in [x, x+\delta]) = \int_{\mathbb{R}} \mathbf{P}(Y(t) \in [x-y, x-y+\delta]) \cdot \mathbf{P}(\widetilde{X}^{<z,-A}(t) \in dy) \leq \frac{C}{t^{1/\alpha} \wedge 1} \cdot \delta.$$

This concludes the proof. □

REMARK 5. It is worth noting that the conditions stated in Theorem 3 are mild for Lévy process $X(t)$ with infinite activities. In particular, for X to exhibit infinite activity, we must have either $\sigma > 0$ or $v(\mathbb{R}) = \infty$. Theorem 3 (i) deals with the case where $\sigma > 0$. On the other hand, when $\sigma = 0$ we must have either $\lim_{\epsilon \downarrow 0} v[\epsilon, \infty) = \infty$ or $\lim_{\epsilon \downarrow 0} v(-\infty, -\epsilon] = \infty$. To satisfy the conditions in part (ii) of Theorem 3, the only other requirement is that $v[\epsilon, \infty)$ (or $v(-\infty, -\epsilon]$) approaches infinity at a rate that is at least comparable to some power-law functions.

The next set of sufficient conditions for Assumption 2 revolves around another type of self-similarity structure in the Lévy measure v.

DEFINITION 4. Given $\alpha \in (0, 2)$ and $b > 1$, a Lévy process X is α-**semi-stable with span** b if its Lévy measure v satisfies

$$v = b^{-\alpha} T_b v \tag{43}$$

where the transformation T_r $(\forall r > 0)$ onto a Borel measure ρ on $\mathbb{R}$ is given by $(T_r \rho)(B) = \rho(r^{-1} B)$. As a special case of semi-stable processes, note that X is α-stable if

$$v(dx) = c_1 \cdot \frac{dx}{x^{1+\alpha}} \mathbb{I}\{x > 0\} + c_2 \cdot \frac{dx}{|x|^{1+\alpha}} \mathbb{I}\{x < 0\}$$

where $c_1, c_2 \geq 0$, $c_1 + c_2 > 0$. See Theorem 14.3 in Sato et al. (1999) for details. However, it is worth noting that the Lévy processes with regularly varying Lévy measures v studied in Proposition 2 are not strict subsets of the semi-stable processes introduced in Definition 4. For instance, given a Borel measure v, suppose that $f(x) = v\big((-\infty, -x] \cup [x, \infty)\big)$ is regularly varying at 0 with index $\alpha > 0$. Even if v satisfies

the scaling-invariant property in (43) for some $b > 1$, we can fix a sequence of points $\{x_n = \frac{1}{b^n}\}_{n \geq 1}$ and assign an extra mass of $\ln n$ onto ν at each point x_n. In doing so, we break the scaling-invariant property but still maintain the regular variation of ν. On the other hand, to show that semi-stable processes may not have regularly varying Lévy measure (when restricted on some neighborhood of the origin), let us consider a simple example. For some $b > 1$ and $\alpha \in (0, 2)$, define the following measure:

$$\nu(\{b^{-n}\}) = b^{n\alpha} \quad \forall n \geq 0; \qquad \nu\big(\mathbb{R}\backslash\{b^n : n \in \mathbb{N}\}\big) = 0.$$

Clearly, ν can be seen as the restriction of the Lévy measure (restricted on $(-1, 1)$) of some α-semi-stable process. Now define function $f(x) = \nu[x, \infty)$ on $(0, \infty)$. For any $t > 0$,

$$\frac{f(tx)}{f(x)} = \frac{\sum_{n=0}^{\lfloor \log_b (1/tx) \rfloor} b^{n\alpha}}{\sum_{n=0}^{\lfloor \log_b (1/x) \rfloor} b^{n\alpha}} = \frac{b^{\lfloor \log_b (1/tx) \rfloor + 1} - 1}{b^{\lfloor \log_b (1/x) \rfloor + 1} - 1}.$$

As $x \to 0$, we see that $f(tx)/f(x)$ will be very close to

$$b^{\alpha(\lfloor \log_b (1/tx) \rfloor - \lfloor \log_b (1/x) \rfloor)}.$$

As long as we didn't pick $t = b^k$ for some $k \in \mathbb{Z}$, asymptotically, the value of $f(tx)/f(x)$ will repeatedly cycle through the following three different values

$$\{b^{\alpha \lfloor \log_b (1/t) \rfloor}, b^{\alpha \lfloor \log_b (1/t) \rfloor + \alpha}, b^{\alpha \lfloor \log_b (1/t) \rfloor - \alpha}\},$$

thus implying that $f(tx)/f(x)$ does not converge as x approaches 0. This confirms that $\nu[x, \infty)$ is not regularly varying as $x \downarrow 0$.

In Proposition 3, we show that semi-stable processes, as well as their truncated counterparts, satisfy continuity conditions of form (40). We say that the process $Y(t)$ is non-trivial if it is not a deterministic linear function (i.e., $Y(t) \equiv ct$ for some $c \in \mathbb{R}$). The proof is again detailed in Section A.

PROPOSITION 3. *Let $\alpha \in (0, 2)$ and $N \in \mathbb{Z}$. Suppose that μ is the Lévy measure of a non-trivial α-semi-stable process $Y'(t)$ of span $b > 1$. Then under $z_0 = b^N$, the Lévy process $\{Y(t) : t > 0\}$ with generating triplet $(0, 0, \mu|_{(-z_0, z_0)})$ has a continuous density function $f_{Y(t)}$ for any $t > 0$. Furthermore, there exists some $C \in (0, \infty)$ such that*

$$\big\|f_{Y(t)}\big\|_\infty \leq \frac{C}{t^{1/\alpha} \wedge 1} \qquad \forall t > 0.$$

Lastly, by applying Proposition 3, we yield another set of sufficient conditions for Assumption 2.

THEOREM 4. *Let (c_X, σ, ν) be the generating triplet of Lévy process X. Suppose that there exist some Borel measure μ and some $z' > 0$, $\alpha \in (0, 2)$ such that $(\nu - \mu)|_{(-z', z')} \geq 0$, and μ is the Lévy measure of some α-semi-stable process. Then Assumption 2 holds for $\lambda = 1/\alpha$.*

Proof   Let $b > 1$ be the span of the α-semi-stable process. Fix some $N \in \mathbb{Z}$ such that $z_0 \triangleq b^N \leq z'$. For any $z \geq z_0$, let $Y(t)$ and $\widetilde{X}^{<z, -A}(t)$ be defined as in (42) with $A = (-z_0, z_0)$. First of all, applying Proposition 3, we can find $C > 0$ such that $\left\| f_{Y(t)} \right\|_\infty \leq \frac{C}{t^{1/\alpha \wedge 1}}$ $\forall t > 0$. Next, due to the independence between Y and $\widetilde{X}^{<z, -A}(t)$, it holds for all $x \in \mathbb{R}, \delta \geq 0$, and $t > 0$ that

$$\mathbf{P}(X^{<z}(t) \in [x, x + \delta]) = \int_{\mathbb{R}} \mathbf{P}(Y(t) \in [x - y, x - y + \delta]) \cdot \mathbf{P}(\widetilde{X}^{<z, -A}(t) \in dy) \leq \frac{C}{t^{1/\alpha} \wedge 1} \cdot \delta.$$

This concludes the proof. $\square$

5. Numerical Experiments   In this section, we apply the importance sampling strategy outlined in Algorithms 2 and 3 and demonstrate that (i) the performance of the importance sampling estimators under different scaling factors and tail distributions, and (ii) the strong efficiency of the proposed algorithms when compared to crude Monte Carlo methods. Specifically, consider a Lévy process $X(t) = B(t) + \sum_{i=1}^{N(t)} W_i$, where $B(t)$ is the standard Brownian motion, N is a Poisson process with arrival rate 0.5, and $\{W_i\}_{i \geq 1}$ is a sequence of iid random variables with law (for some $\alpha > 1$)

$$\mathbf{P}(W_1 > x) = \mathbf{P}(-W_1 > x) = \frac{0.5}{(1 + x)^\alpha}, \qquad \forall x > 0.$$

For each $n \geq 1$, we define the scaled process $\bar{X}_n(t) = \frac{X(nt)}{n}$. The goal is to estimate the probability of $A_n = \{X_n \in A\}$, where the set A is defined as in (12) with $a = 2$ and $b = 1.15$. Note that this is a case with $l^* = \lceil a/b \rceil = 2$.

To evaluate the performance of the importance sampling estimator under different scaling factors and tail distributions, we run experiments under $\alpha \in \{1.45, 1.6, 1.75\}$, and $n \in \{100, 200, \cdots, 1000\}$. The efficiency is evaluated by the relative error of the algorithm, namely the ratio between the standard deviation and the estimated mean. In Algorithm 2, we set $\gamma = 0.25$, $w = 0.05$, $\rho = 0.97$, and $d = 4$. In Algorithm 3, we further

FIGURE 1. Relative errors of the proposed importance sampling estimator.
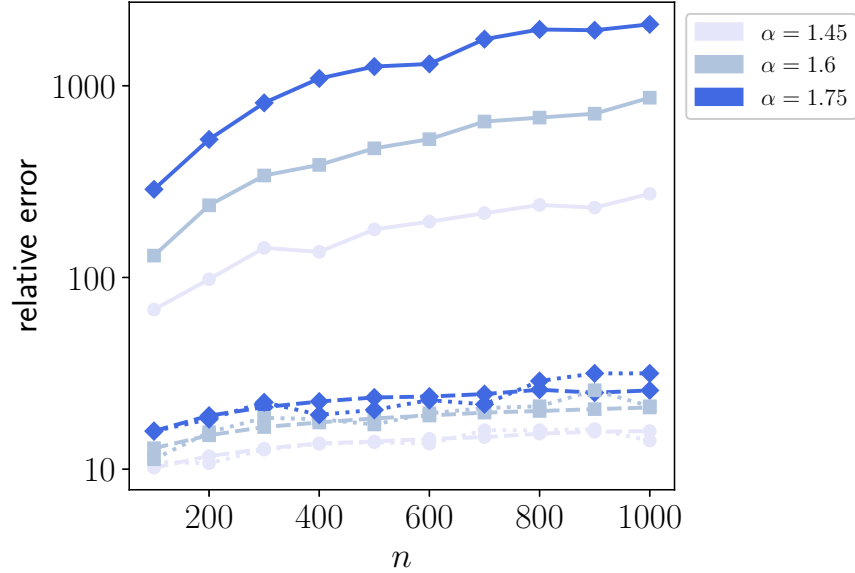
Note.   Results are plotted under log scale. **Dashed lines**: the importance sampling estimator in Algorithm 2, **Dotted lines**: the importance sampling estimator with ARA in Algorithm 3, **Solid lines**: the crude Monte-Carlo methods (solid lines).

set $\kappa = 0.5$ and $r = 1.5$. For both algorithms, we generate 10,000 independent samples for each combination of $\alpha \in \{1.45, 1.6, 1.75\}$ and $n \in \{1000, 2000, \cdots, 10000\}$. For the number of samples in crude Monte Carlo estimation, we ensure that at least $64/\hat{p}_{\alpha,n}$ samples are generated, where $\hat{p}_{\alpha,n}$ is the probability estimated by Algorithm 2.

The results are summarized in Table 1 and Figure 1. In Table 1, we see that for a fixed α, the relative error of the importance sampling estimators stabilizes around a constant level as n increases. This aligns with the strong efficiency established in Theorems 1 and 2. In comparison, the relative error of crude Monte Carlo estimators continues to increase as n tends to infinity. Figure 1 further highlights that our importance sampling estimators significantly outperform crude Monte Carlo methods by orders of magnitude. In summary, when Algorithms 2 and 3 are appropriately parameterized, their efficiency becomes increasingly evident when compared against the crude Monte Carlo approach as the scaling factor n grows larger and the target probability approaches 0.

6. Proofs of Section 3

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

37

TABLE 1. Relative errors of Algorithm 2 (first row), Algorithm 3 (second row), and crude Monte Carlo (third row).

n	200	400	600	800	1000
	11.70	13.65	14.40	15.33	15.82
$\alpha = 1.45$	10.74	13.57	13.57	15.98	14.11
	97.86	136.03	195.40	238.81	273.13
	15.03	17.53	19.06	20.12	20.98
$\alpha = 1.6$	15.59	18.23	19.59	21.30	21.30
	237.82	386.35	526.13	681.79	866.02
	19.03	22.54	23.94	25.97	25.77
$\alpha = 1.75$	18.23	19.22	22.92	28.85	31.61
	524.78	1091.29	1298.98	1965.22	2089.82

6.1. Proof of Proposition 1 We first prepare two technical lemmas using the sample-path large

deviations for heavy-tailed Lévy processes reviewed in Section 2.2.

LEMMA 1. *For the set $A \subset \mathbb{D}$ defined in (12) and the quantity l^* defined in (13),*

$$0 < \liminf_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n,\infty))^{l^*}} \leq \limsup_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n,\infty))^{l^*}} < \infty.$$

Proof In this proof, we focus on the two-sided case in Assumption 1. It is worth noticing that analysis for

the one-sided case is almost identical, with the only major difference being that we apply Result 1 (i.e., the

one-sided version of the large deviations of $\bar{X}_n$) instead of Result 2 (i.e., the two-sided version). Specifically,

we claim that

(*i*) $(l^*, 0) \in \underset{(j,k)\in\mathbb{N}^2,\ \mathbb{D}_{j,k}\cap A\neq\emptyset}{\operatorname{argmin}} j(\alpha - 1) + k(\alpha' - 1);$

(*ii*) $\mathbf{C}_{l^*,0}(A^\circ) > 0;$

(*iii*) the set A is bounded away from $\mathbb{D}_{<l^*,0}$.

Then by applying Result 2, we yield

$$0 < \mathbf{C}_{l^*,0}(A^\circ) \leq \liminf_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n,\infty))^{l^*}} \leq \limsup_{n\to\infty} \frac{\mathbf{P}(\bar{X}_n \in A)}{(n\nu[n,\infty))^{l^*}} \leq \mathbf{C}_{l^*,0}(A^-) < \infty$$

38

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

and conclude the proof. Now, it remains to prove claims (*i*), (*ii*), and (*iii*).

Proof of Claim (*i*). By definitions of $\mathbb{D}_{j,k}$, for any $\xi \in \mathbb{D}_{j,k}$ there exist $(u_i)_{i=1}^j \in (0, \infty)^j$, $(t_i)_{i=1}^j \in (0, 1]^j$ and $(v_i)_{i=1}^k \in (0, \infty)^k$, $(s_i)_{i=1}^k \in (0, 1]^k$ such that

$$\xi(t) = \sum_{i=1}^j u_i \mathbf{I}_{[t_i,1]}(t) - \sum_{i=1}^k v_i \mathbf{I}_{[s_i,1]}(t) \qquad \forall t \in [0,1]. \tag{44}$$

First, from Assumption 3, one can choose $\epsilon > 0$ small enough such that $l^*(b - \epsilon) > a$. Then for the case with $(j, k) = (l^*, 0)$ in (44), by picking $u_i = b - \epsilon$ for all $i \in [l^*]$, we have $\sup_{t \in [0,1]} \xi(t) = \sum_{i=1}^{l^*} u_i = l^*(b - \epsilon) > a$, and hence $\xi \in A$. This verifies $\mathbb{D}_{l^*,0} \cap A \neq \emptyset$.

Next, suppose we can show that $j \geq l^*$ is a necessary condition for $\mathbb{D}_{j,k} \cap A \neq \emptyset$. Then we get

$$\left\{ (j, k) \in \mathbb{N}^2 : \mathbb{D}_{j,k} \cap A \neq \emptyset \right\} \subseteq \left\{ (j, k) \in \mathbb{N}^2 : j \geq l^*, \ k \geq 0 \right\},$$

which immediately verifies claim (*i*) due to $\alpha, \alpha' > 1$; see Assumption 1. Now, to show that $j \geq l^*$ is a necessary condition for $\mathbb{D}_{j,k} \cap A \neq \emptyset$, note that from (44), it holds for any $\xi \in \mathbb{D}_{j,k} \cap A$ that $a < \sup_{t \in [0,1]} \xi(t) \leq \sum_{i=1}^j u_i < jb$. As a result, we must have $j > a/b$ and hence $j \geq l^* = \lceil a/b \rceil$ due to $a/b \notin \mathbb{Z}$; see Assumption 3. This concludes the proof of claim (*i*).

Proof of Claim (*ii*). Again, choose some $\epsilon > 0$ small enough such that $l^*(b - \epsilon) > a$. Given any $u_i \in (b - \epsilon, b)$ and $0 < t_1 < t_2 < \cdots < t_{l^*} < 1$, the step function $\xi(t) = \sum_{i=1}^{l^*} u_i \mathbf{I}_{[t_i,1]}(t)$ satisfies $\sup_{t \in [0,1]} \xi(t) \geq l^*(b - \epsilon) > a$, thus implying $\xi \in A$. Therefore, (for the definition of $\mathbf{C}_{j,k}$, see (6))

$$\mathbf{C}_{l^*,0}(A^\circ) \geq \nu_\alpha^{l^*}\left((b - \epsilon, b)^{l^*} \right) = \frac{1}{l^*!} \left[\frac{1}{(b - \epsilon)^\alpha} - \frac{1}{b^\alpha} \right]^{l^*} > 0.$$

Proof of Claim (*iii*). Assumption 3 implies that $a > (l^* - 1)b$, allowing us to choose $\epsilon > 0$ small enough that $a - \epsilon > (l^* - 1)(b + \epsilon)$. It suffices to show that

$$d(\xi, \xi') \geq \epsilon \qquad \forall \xi \in \mathbb{D}_{<l^*,0}, \ \xi' \in A. \tag{45}$$

Here, d is the Skorokhod J_1 metric; see (4) for the definition. To prove (45), we start with the following observation: due to claim (i), for any $(j, k) \in \mathbb{N}^2$ with $(j, k) \in \mathbb{I}_{<l^*, 0}$, we must have $j \leq l^* - 1$. Now, we proceed with a proof by contradiction. Suppose there is some $\xi \in \mathbb{D}_{j,k}$ with $j \leq l^* - 1$ and some $\xi' \in A$ such that $d(\xi, \xi') < \epsilon$. Due to $\xi' \in A$ (and hence no upward jump in ξ' is larger than b) and $d(\xi, \xi') < \epsilon$, under the representation (44) we must have $u_i < b + \epsilon$ $\forall i \in [j]$. This implies $\sup_{t \in [0,1]} \xi(t) \leq \sum_{i=1}^{j} u_i < j(b + \epsilon) \leq (l^* - 1)(b + \epsilon)$. Due to $d(\xi, \xi') < \epsilon$ again, we yield the contradiction that $\sup_{t \in [0,1]} \xi'(t) < (l^* - 1)(b + \epsilon) + \epsilon < a$ (and hence $\xi' \notin A$). This concludes the proof of claim (iii). □

LEMMA 2.   *Let $p > 1$. Let $\Delta > 0$ be such that $a - \Delta > (l^* - 1)b$ and $[a - \Delta - (l^* - 1)b]/\gamma \notin \mathbb{Z}$. Suppose that $(J_\gamma + l^* - 1)/p > 2l^*$ holds for*

$$J_\gamma \triangleq \lceil \frac{a - \Delta - (l^* - 1)b}{\gamma} \rceil. \tag{46}$$

Then

$$\mathbf{P}\big(\bar{X}_n \in A^\Delta \cap E, \ \mathcal{D}(\bar{J}_n) \leq l^* - 1\big) = o\Big(\big(n\nu[n, \infty)\big)^{2pl^*}\Big) \qquad as \ n \to \infty$$

where $A^\Delta = \{\xi \in \mathbb{D} : \sup_{t \in [0,1]} \xi(t) \geq a - \Delta\}$, $E \triangleq \{\xi \in \mathbb{D} : \sup_{t \in (0,1]} \xi(t) - \xi(t-) < b\}$ and the function $\mathcal{D}(\xi)$ counts the number of discontinuities in $\xi \in \mathbb{D}$.

Proof   Similar to the proof of Lemma 1, we focus on the two-sided case in Assumption 1. Still, it is worth noticing that the proof of the one-sided case is almost identical, with the only major difference being that we apply Result 1 instead of Result 2.

First, observe that $\mathbf{P}(\bar{X}_n \in A^\Delta \cap E, \ \mathcal{D}(\bar{J}_n) \leq l^* - 1) = \mathbf{P}(\bar{X}_n \in F)$ where

$$F \triangleq \Big\{ \xi \in \mathbb{D} : \ \sup_{t \in [0,1]} \xi(t) \geq a - \Delta; \ \sup_{t \in (0,1]} \xi(t) - \xi(t-) < b,$$

$$\#\{t \in [0, 1] : \ \xi(t) - \xi(t-) \geq \gamma\} \leq l^* - 1 \Big\}.$$

Furthermore, we claim that

(i) $(J_\gamma + l^* - 1, 0) \in \underset{(j,k)\in\mathbb{N}^2,\ \mathbb{D}_{j,k}\cap F\neq\emptyset}{\operatorname{argmin}} j(\alpha - 1) + k(\alpha' - 1);$

(ii) the set F is bounded away from $\mathbb{D}_{<J_\gamma + l^* - 1, 0}$.

Then we are able to apply Result 2 and obtain

$$\mathbf{P}(\bar{X}_n \in A^\Delta \cap E,\ \mathcal{D}(\bar{J}_n) \leq l^* - 1) = \mathbf{P}(\bar{X}_n \in F) = \boldsymbol{O}\big((n\nu[n,\infty))^{J_\gamma + l^* - 1}\big) \qquad \text{as } n \to \infty.$$

Lastly, by our assumption $(J_\gamma + l^* - 1)/p > 2l^*$, we get $(n\nu[n,\infty))^{l^* - 1 + J_\gamma} = \boldsymbol{o}\big((n\nu[n,\infty))^{2pl^*}\big)$ and conclude

the proof. Now, it remains to prove claims (i) and (ii).

Proof of Claim (i). By definition of $\mathbb{D}_{j,k}$, given any $\xi \in \mathbb{D}_{j,k}$ there exist $(u_i)_{i=1}^j \in (0,\infty)^j, (t_i)_{i=1}^j \in (0,1]^j$

and $(v_i)_{i=1}^k \in (0,\infty)^k, (s_i)_{i=1}^k \in (0,1]^k$ such that the representation (44) holds. By assumption $[a - \Delta - (l^* -$

$1)b]/\gamma \notin \mathbb{Z}$, for J_γ defined in (46) we have

$$(J_\gamma - 1)\gamma < a - \Delta - (l^* - 1)b < J_\gamma \cdot \gamma. \tag{47}$$

It then holds for all $\epsilon > 0$ small enough that $a - \Delta < J_\gamma(\gamma - \epsilon) + (l^* - 1)(b - \epsilon)$. As a result, for the

case with $(j,k) = (l^* - 1 + J_\gamma, 0)$ in (44), by picking $u_i = b - \epsilon$ for all $i \in [l^* - 1]$, and $u_i = \gamma - \epsilon$ for all

$i = l^*, l^* + 1, \cdots, l^* - 1 + J_\gamma$, we get $\sup_{t\in[0,1]} \xi(t) = J_\gamma(\gamma - \epsilon) + (l^* - 1)(b - \epsilon) > a - \Delta$. This proves that

$\xi \in \mathbb{D}_{l^* - 1 + J_\gamma, 0} \cap F$, and hence $\mathbb{D}_{l^* - 1 + J_\gamma, 0} \cap F \neq \emptyset$.

Next, suppose we can show that $j \geq l^* - 1 + J_\gamma$ is the necessary condition for $\mathbb{D}_{j,k} \cap F \neq \emptyset$. Then, we get

$$\{(j,k) \in \mathbb{N}^2 :\ \mathbb{D}_{j,k} \cap F \neq \emptyset\} \subseteq \{(j,k) \in \mathbb{N}^2 :\ j \geq l^* - 1 + J_\gamma,\ k \geq 0\},$$

which immediately verifies claim (i) due to $\alpha, \alpha' > 1$; see Assumption 1. Now, to show that $j \geq l^* - 1 + J_\gamma$ is a

necessary condition, note that, from (44), it holds for any $\xi \in \mathbb{D}_{j,k} \cap F$ that $a - \Delta < \sup_{t\in[0,1]} \xi(t) \leq \sum_{i=1}^j u_i$.

Furthermore, by the definition of the set F, we must have (here, w.l.o.g., we order u_i's by $u_1 \geq u_2 \geq \ldots \geq u_j$)

$u_i < b$ for all $i \in [l^* - 1]$ and $u_i < \gamma$ for all $i = l^*, l^* + 1, \cdots, j$. This implies $(l^* - 1)b + (j - l^* + 1)\gamma > a - \Delta$,

and hence $j > \frac{a - \Delta - (l^* - 1)b}{\gamma} + l^* - 1$, which is equivalent to $j \geq J_\gamma + l^* - 1$.

Proof of Claim (ii). From (47), we can fix some $\epsilon > 0$ small enough such that

$$a - \Delta - \epsilon > (l^* - 1)(b + \epsilon) + (J_\gamma - 1)(\gamma + \epsilon). \tag{48}$$

It suffices to show that

$$\boldsymbol{d}(\xi, \xi') \geq \epsilon \qquad \forall \xi \in \mathbb{D}_{<J_\gamma + l^* - 1, 0}, \ \xi' \in F. \tag{49}$$

Here, $\boldsymbol{d}$ is the Skorokhod J_1 metric; see (4) for the definition. To prove (49), we start with the following observation: using claim (i), for any $(j, k) \in \mathbb{N}^2$ with $(j, k) \in \mathbb{I}_{<J_\gamma + l^* - 1, 0}$, we must have $j \leq J_\gamma + l^* - 2$. Next, we proceed with a proof by contradiction. Suppose there is some $\xi \in \mathbb{D}_{j,k}$ with $j \leq J_\gamma + l^* - 2$ and some $\xi' \in F$ such that $\boldsymbol{d}(\xi, \xi') < \epsilon$. By the definition of the set F above, any upward jump in ξ' is bounded by b, and at most $l^* - 1$ of them is larger than γ. Then from $\boldsymbol{d}(\xi, \xi') < \epsilon$, we know that any upward jump in ξ is bounded by $b + \epsilon$, and at most $l^* - 1$ of them is larger than $\gamma + \epsilon$. Through (44), we now have

$$\sup_{t \in [0,1]} \xi(t) \leq \sum_{i=1}^{j} u_i \leq (l^* - 1)(b + \epsilon) + (J_\gamma - 1)(\gamma + \epsilon) < a - \Delta - \epsilon.$$

The last inequality follows from (48). Using $\boldsymbol{d}(\xi, \xi') < \epsilon$ again, we yield the contraction that $\sup_{t \in [0,1]} \xi'(t) < a - \Delta$ and hence $\xi' \notin F$. This concludes the proof of (49). $\square$

Now, we are ready to prove Proposition 1.

Proof of Proposition 1   We start by proving the **unbiasedness** of the importance sampling estimator

$$L_n = Z_n \frac{d\mathbf{P}}{d\mathbf{Q}_n} = \sum_{m=0}^{\tau} \frac{\hat{Y}_n^m \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} - \hat{Y}_n^{m-1} \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n}}{\mathbf{P}(\tau \geq m)}.$$

under $\mathbf{Q}_n$. Note that under both $\mathbf{P}$ and $\mathbf{Q}_n$, we have $\tau \sim \text{Geom}(\rho)$ (i.e., $\mathbf{P}(\tau \geq m) = \rho^{m-1}$) and that τ is independent of everything else. In light of Result 4, it suffices to verify $\lim_{m \to \infty} \mathbf{E}^{\mathbf{Q}_n}[Y_m] = \mathbf{E}^{\mathbf{Q}_n}[Y]$ and condition (11) (under $\mathbf{Q}_n$) with the choice of $Y_m = \hat{Y}_n^m \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n}$ and $Y = Y_n^* \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n}$. In particular, it suffices to show that (for any $n \geq 1$)

$$\sum_{m \geq 1} \mathbf{E}^{\mathbf{Q}_n}\left[\left|\hat{Y}_n^{m-1} \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} - Y_n^* \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n}\right|^2\right] \Big/ \mathbf{P}(\tau \geq m) < \infty. \tag{50}$$

42

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

To see why, note that (50) directly verifies condition (11). Furthermore, it implies $\lim_{m\to\infty} \mathbf{E}^{\mathbf{Q}_n} |\hat{Y}_n^{m-1} \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} - Y_n^* \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n}|^2 = 0$. The $\mathcal{L}_2$ convergence then implies the $\mathcal{L}_1$ convergence, i.e., $\lim_{m\to\infty} \mathbf{E}^{\mathbf{Q}_n} [\hat{Y}_n^{m-1} \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n}] = \mathbf{E}^{\mathbf{Q}_n} [Y_n^* \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n}]$.

To prove claim (50), observe that

$$
\begin{aligned}
\mathbf{E}^{\mathbf{Q}_n} \left[\left| \hat{Y}_n^{m-1} \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} - Y_n^* \mathbf{I}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} \right|^2 \right] &\leq \mathbf{E}^{\mathbf{Q}_n} \left[|\hat{Y}_n^{m-1} - Y_n^*|^2 \cdot \left(\frac{d\mathbf{P}}{d\mathbf{Q}_n} \right)^2 \right] \\
&= \mathbf{E} \left[|\hat{Y}_n^{m-1} - Y_n^*|^2 \cdot \frac{d\mathbf{P}}{d\mathbf{Q}_n} \right] \\
&\leq \frac{1}{w} \mathbf{E} |\hat{Y}_n^{m-1} - Y_n^*|^2 \qquad \text{due to } \frac{d\mathbf{P}}{d\mathbf{Q}_n} \leq \frac{1}{w}, \text{ see (16).}
\end{aligned}
$$

In particular, since $\hat{Y}_n^m$ and Y_n^* only take values in $\{0, 1\}$, we have $\mathbf{E}|\hat{Y}_n^m - Y_n^*|^2 = \mathbf{P}(\hat{Y}_n^m \neq Y_n^*)$, and

$$
\begin{aligned}
\mathbf{P}(\hat{Y}_n^m \neq Y_n^*) &= \sum_{k \geq 0} \mathbf{P}(Y_n^* \neq \hat{Y}_n^m \mid \mathcal{D}(\bar{J}_n) = k) \mathbf{P}(\mathcal{D}(\bar{J}_n) = k) \\
&\leq \sum_{k \geq 0} C_0 \rho_0^m \cdot (k+1) \cdot \mathbf{P}(\mathcal{D}(\bar{J}_n) = k) \qquad \text{for all } m \geq \bar{m} \text{ due to (20)} \\
&= C_0 \rho_0^m \cdot \mathbf{E} \left[1 + \text{Poisson}(n\nu[n\gamma, \infty)) \right] = C_0 \rho_0^m \cdot (1 + n\nu[n\gamma, \infty)).
\end{aligned}
\tag{51}
$$

The last line in the display above follows from the definition of $\bar{J}_n(t) = \frac{1}{n} J(nt)$ in (17). To conclude, note that $\nu(x) \in \mathcal{RV}_{-\alpha}(x)$ and hence $n\nu[n\gamma, \infty) \in \mathcal{RV}_{-(\alpha-1)}(n)$ with $\alpha > 1$, thus implying $\sup_{n \geq 1} n\nu[n\gamma, \infty) < \infty$; also, as prescribed in Proposition 1 we have $\rho \in (\rho_0, 1)$.

The rest of the proof is devoted to establishing the **strong efficiency** of L_n. Observe that

$$
\mathbf{E}^{\mathbf{Q}_n} [L_n^2] = \int Z_n^2 \frac{d\mathbf{P}}{d\mathbf{Q}_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} d\mathbf{Q}_n = \int Z_n^2 \frac{d\mathbf{P}}{d\mathbf{Q}_n} d\mathbf{P} = \int Z_n^2 \mathbf{I}_{B_n^\gamma} \frac{d\mathbf{P}}{d\mathbf{Q}_n} d\mathbf{P} + \int Z_n^2 \mathbf{I}_{(B_n^\gamma)^c} \frac{d\mathbf{P}}{d\mathbf{Q}_n} d\mathbf{P}.
$$

By definitions in (16), on event $(B_n^\gamma)^c$ we have $\frac{d\mathbf{P}}{d\mathbf{Q}_n} \leq \frac{1}{w}$, while on event B_n^γ we have $\frac{d\mathbf{P}}{d\mathbf{Q}_n} \leq \frac{\mathbf{P}(B_n^\gamma)}{1-w}$. As a result,

$$
\mathbf{E}^{\mathbf{Q}_n} [L_n^2] \leq \frac{\mathbf{P}(B_n^\gamma)}{1-w} \mathbf{E}[Z_n^2 \mathbf{I}_{B_n^\gamma}] + \frac{1}{w} \mathbf{E}[Z_n^2 \mathbf{I}_{(B_n^\gamma)^c}].
\tag{52}
$$

Meanwhile, Lemma 1 implies that

$$0 < \liminf_{n \to \infty} \frac{\mathbf{P}(A_n)}{(n\nu[n, \infty))^{l^*}} \le \limsup_{n \to \infty} \frac{\mathbf{P}(A_n)}{(n\nu[n, \infty))^{l^*}} < \infty. \tag{53}$$

Let $Z_{n,1} \triangleq Z_n \mathbf{I}_{B_n^\gamma}$ and $Z_{n,2} \triangleq Z_n \mathbf{I}_{(B_n^\gamma)^c}$. Furthermore, given $\rho \in (\rho_0, 1)$, we claim the existence of some $\bar{\gamma} = \bar{\gamma}(\rho) \in (0, b)$ such that for any $\gamma \in (0, \bar{\gamma})$,

$$\mathbf{P}(B_n^\gamma) = \boldsymbol{O}\big((n\nu[n, \infty))^{l^*}\big), \tag{54}$$

$$\mathbf{E}[Z_{n,1}^2] = \boldsymbol{O}\big((n\nu[n, \infty))^{l^*}\big), \tag{55}$$

$$\mathbf{E}[Z_{n,2}^2] = \boldsymbol{o}\big((n\nu[n, \infty))^{2l^*}\big), \tag{56}$$

as $n \to \infty$. Then, using (54) and (55) we get $\mathbf{P}(B_n^\gamma)\mathbf{E}[Z_n^2 \mathbf{I}_{B_n^\gamma}] = \boldsymbol{O}\big((n\nu[n, \infty))^{2l^*}\big) = \boldsymbol{O}\big(\mathbf{P}^2(A_n)\big)$. The last equality follows from (53). Similarly, from (53) and (56) we get $\mathbf{E}[Z_n^2 \mathbf{I}_{(B_n^\gamma)^c}] = \boldsymbol{o}\big((n\nu[n, \infty))^{2l^*}\big) = \boldsymbol{o}\big(\mathbf{P}^2(A_n)\big)$. Therefore, in (52) we have $\mathbf{E}^{\mathbf{Q}_n}[L_n^2] = \boldsymbol{O}\big(\mathbf{P}^2(A_n)\big)$, thus establishing the strong efficiency. Now, it remains to prove claims (54), (55), and (56).

Proof of Claim (54). We show that the claim holds for all $\gamma \in (0, b)$. For any $c > 0$ and $k \in \mathbb{N}$, note that

$$\mathbf{P}\big(\text{Poisson}(c) \ge k\big) = \sum_{j \ge k} \exp(-c)\frac{c^j}{j!} = c^k \sum_{j \ge k} \exp(-c)\frac{c^{j-k}}{j!} \le c^k \sum_{j \ge k} \exp(-c)\frac{c^{j-k}}{(j-k)!} = c^k. \tag{57}$$

Recall that $B_n^\gamma = \{\bar{X}_n \in B^\gamma\}$ and $B^\gamma \triangleq \{\xi \in \mathbb{D} : \#\{t \in [0,1] : \xi(t) - \xi(t-) \ge \gamma\} \ge l^*\}$. Therefore,

$$\mathbf{P}(B_n^\gamma) = \mathbf{P}\big(\#\{t \in [0,n] : X(t) - X(t-) \ge n\gamma\} \ge l^*\big) \qquad \text{due to } \bar{X}_n(t) = X(nt)/n$$

$$= \sum_{k \ge l^*} \exp\big(-n\nu[n\gamma, \infty)\big)\frac{\big(n\nu[n\gamma, \infty)\big)^k}{k!} \le \big(n\nu[n\gamma, \infty)\big)^{l^*} \qquad \text{due to (57)}.$$

44

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Lastly, the regularly varying nature of $\nu[x, \infty)$ (see Assumption 1) implies $\lim_{n\to\infty} \frac{(n\nu[n\gamma, \infty))^{l^*}}{(n\nu[n, \infty))^{l^*}} = 1/\gamma^{\alpha l^*} \in$

$(0, \infty)$, and hence $\mathbf{P}(B_n^\gamma) = \boldsymbol{O}\big((n\nu[n, \infty))^{l^*}\big)$.

Proof of Claim (55). Again, we prove the claim for all $\gamma \in (0, b)$. By the definition of Z_n in (19),

$$Z_{n,1} = Z_n \mathbf{I}_{B_n^\gamma} = \sum_{m=0}^{\tau} \frac{\hat{Y}_n^m \mathbf{I}_{E_n \cap B_n^\gamma} - \hat{Y}_n^{m-1} \mathbf{I}_{E_n \cap B_n^\gamma}}{\mathbf{P}(\tau \geq m)}.$$

Meanwhile, by the definition of B_n^γ, we have $\mathbf{I}_{B_n^\gamma} = 0$ on $\{\mathcal{D}(\bar{J}_n) < l^*\}$, where $\mathcal{D}(\xi)$ counts the number of

discontinuities for any $\xi \in \mathbb{D}$. By applying Result 4 under the choice of $Y_m = \hat{Y}_n^m \mathbf{I}_{E_n \cap B_n^\gamma}$ and $Y = Y_n^* \mathbf{I}_{E_n \cap B_n^\gamma}$,

we yield

$$
\begin{aligned}
\mathbf{E}Z_{n,1}^2 &\leq \sum_{m \geq 1} \frac{\mathbf{E}\left[\left|Y_n^* \mathbf{I}_{E_n \cap B_n^\gamma} - \hat{Y}_n^{m-1} \mathbf{I}_{E_n \cap B_n^\gamma}\right|^2\right]}{\mathbf{P}(\tau \geq m)} \\
&\leq \sum_{m \geq 1} \sum_{k \geq l^*} \frac{\mathbf{E}\left[\mathbf{I}\big(Y_n^* \neq \hat{Y}_n^{m-1}\big) \,\Big|\, \{\mathcal{D}(\bar{J}_n) = k\}\right]}{\mathbf{P}(\tau \geq m)} \cdot \mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big) \quad \text{due to } \mathbf{I}_{B_n^\gamma} = 0 \text{ on } \{\mathcal{D}(\bar{J}_n) < l^*\} \\
&\leq \sum_{k \geq l^*} \mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big) \cdot \sum_{m \geq 1} \frac{\mathbf{P}\big(Y_n^* \neq \hat{Y}_n^{m-1} \,\big|\, \{\mathcal{D}(\bar{J}_n) = k\}\big)}{\mathbf{P}(\tau \geq m)} \\
&\leq \sum_{k \geq l^*} \mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big) \cdot \left[\sum_{m=1}^{\bar{m}} \frac{1}{\rho^{m-1}} + \sum_{m \geq \bar{m}+1} \frac{C_0 \rho_0^{m-1} \cdot (k+1)}{\rho^{m-1}}\right] \qquad \text{by condition (20)} \\
&\leq \sum_{k \geq l^*} \mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big) \cdot (k+1) \cdot \underbrace{\left[\sum_{m=1}^{\bar{m}} \frac{1}{\rho^{m-1}} + \sum_{m \geq \bar{m}+1} \frac{C_0 \rho_0^{m-1}}{\rho^{m-1}}\right]}_{\triangleq \widetilde{C}_{\rho,1}}. \qquad (58)
\end{aligned}
$$

In particular, given $\rho \in (\rho_0, 1)$, we have $\widetilde{C}_{\rho,1} < \infty$, and hence

$$
\begin{aligned}
\mathbf{E} Z_{n,1}^2 &\leq \widetilde{C}_{\rho,1} \sum_{k \geq l^*} (k+1) \cdot \mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big) \\
&= \widetilde{C}_{\rho,1} \sum_{k \geq l^*} (k+1) \cdot \exp\big(-n\nu[n\gamma, \infty)\big) \frac{\big(n\nu[n\gamma, \infty)\big)^k}{k!} \\
&\leq 2\widetilde{C}_{\rho,1} \sum_{k \geq l^*} k \cdot \exp\big(-n\nu[n\gamma, \infty)\big) \frac{\big(n\nu[n\gamma, \infty)\big)^k}{k!} \qquad \text{due to } l^* \geq 1 \implies \frac{k+1}{k} \leq 2 \; \forall k \geq l^* \\
&\leq 2\widetilde{C}_{\rho,1} \cdot \big(n\nu[n\gamma, \infty)\big)^{l^*} \sum_{k \geq l^*} \exp\big(-n\nu[n\gamma, \infty)\big) \frac{\big(n\nu[n\gamma, \infty)\big)^{k-l^*}}{(k-l^*)!} \qquad \text{due to } l^* \geq 1 \\
&= 2\widetilde{C}_{\rho,1} \cdot \big(n\nu[n\gamma, \infty)\big)^{l^*}.
\end{aligned}
$$

Again, the regular varying nature of $\nu[x, \infty)$ allows us to conclude that $\mathbf{E} Z_{n,1}^2 = \boldsymbol{O}\big((n\nu[n, \infty))^{l^*}\big)$.

Proof of Claim (56). Fix some $\rho \in (\rho_0, 1)$ and some $q > 1$ such that $\rho_0^{1/q} < \rho$. Let $p > 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. By Assumption 3, we can pick some $\Delta_0 > 0$ small enough such that $a - \Delta_0 > (l_1^*)b$. This allows us to pick $\bar{\gamma} \in (0, b)$ small enough such that $(\hat{J} + l^* - 1)/p > 2l^*$ where

$$
\hat{J} \triangleq \frac{a - \Delta_0 - (l^* - 1)b}{\bar{\gamma}}. \tag{59}
$$

We prove the claim for all $\gamma \in (0, \bar{\gamma})$. Specifically, given any $\gamma \in (0, \bar{\gamma})$, one can pick $\Delta \in (0, \Delta_0)$ such that $[a - \Delta - (l^* - 1)b]/\gamma \notin \mathbb{Z}$. Due to our choice of γ and Δ, it follows from (59) that $(J_\gamma + l^* - 1)/p > 2l^*$ where

$$
J_\gamma \triangleq \lceil \frac{a - \Delta - (l^* - 1)b}{\gamma} \rceil.
$$

Let $A^\Delta = \{\xi \in \mathbb{D} : \sup_{t \in [0,1]} \xi(t) \geq a - \Delta\}$ and $A_n^\Delta = \{\bar{X}_n \in A^\Delta\}$. Also, note that

$$
Z_{n,2} = Z_n \mathbf{I}_{(B_n^\gamma)^c} = \underbrace{Z_n \mathbf{I}_{A_n^\Delta \cap (B_n^\gamma)^c}}_{\triangleq Z_{n,3}} + \underbrace{Z_n \mathbf{I}_{(A_n^\Delta)^c \cap (B_n^\gamma)^c}}_{\triangleq Z_{n,4}}.
$$

46

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Specifically, $Z_{n,3} = \sum_{m=0}^{\tau} \frac{\hat{Y}_n^m \mathbf{I}_{A_n^\Delta \cap E_n \cap (B_n^\gamma)^c} - \hat{Y}_n^{m-1} \mathbf{I}_{A_n^\Delta \cap E_n \cap (B_n^\gamma)^c}}{\mathbf{P}(\tau \geq m)}$. Analogous to the calculations in (58), by applying

Result 4 under the choice of $Y_m = \hat{Y}_n^m \mathbf{I}_{A_n^\Delta \cap E_n \cap (B_n^\gamma)^c}$ and $Y = Y_n^* \mathbf{I}_{A_n^\Delta \cap E_n \cap (B_n^\gamma)^c}$, we yield

$$
\begin{aligned}
\mathbf{E}Z_{n,3}^2 &\leq \sum_{m \geq 1} \frac{\mathbf{E}\left[\left|Y_n^* - \hat{Y}_n^{m-1}\right|^2 \mathbf{I}_{A_n^\Delta \cap E_n \cap (B_n^\gamma)^c}\right]}{\mathbf{P}(\tau \geq m)} \\
&= \sum_{m \geq 1} \frac{\mathbf{E}\left[\mathbf{I}(Y_n^* \neq \hat{Y}_n^{m-1}) \cdot \mathbf{I}_{A_n^\Delta \cap E_n \cap (B_n^\gamma)^c}\right]}{\mathbf{P}(\tau \geq m)} \qquad \text{because } \hat{Y}_n^m \text{ and } Y_n^* \text{ only take values in } \{0,1\} \\
&\leq \sum_{m \geq 1} \frac{\left(\mathbf{P}(Y_n^* \neq \hat{Y}_n^{m-1})\right)^{1/q} \cdot \left(\mathbf{P}(A_n^\Delta \cap E_n \cap (B_n^\gamma)^c)\right)^{1/p}}{\mathbf{P}(\tau \geq m)} \qquad \text{by Hölder's inequality.}
\end{aligned}
$$

Applying Lemma 2, we get $\left(\mathbf{P}(A_n^\Delta \cap E_n \cap (B_n^\gamma)^c)\right)^{1/p} = o\left((n\nu[n,\infty))^{2l^*}\right)$. On the other hand, it has been

shown in (51) that for any $n \geq 1$ and $m \geq \bar{m}$, we have $\mathbf{P}(Y_n^* \neq \hat{Y}_n^m) \leq C_0 C_\gamma \rho_0^m$ where $C_\gamma \triangleq \sup_{n \geq 1} n\nu[n\gamma,\infty) +$

$1 < \infty$. In summary,

$$
\mathbf{E}Z_{n,3}^2 \leq o\left((n\nu[n,\infty))^{2l^*}\right) \cdot \underbrace{\left[\sum_{m=1}^{\bar{m}} \frac{1}{\rho^{m-1}} + \sum_{m \geq \bar{m}+1} \frac{(C_0 C_\gamma)^{1/q} \cdot (\rho_0^{1/q})^{m-1}}{\rho^{m-1}}\right]}_{\triangleq \widetilde{C}_{\rho,2}}. \tag{60}
$$

Note that $\widetilde{C}_{\rho,2} < \infty$ due to our choice of $\rho_0^{1/q} < \rho$.

Similarly, to bound the second-order moment of $Z_{n,4} = \sum_{m=0}^{\tau} \frac{\hat{Y}_n^m \mathbf{I}_{(A_n^{\Delta})^c \cap E_n \cap (B_n^{\gamma})^c} - \hat{Y}_n^{m-1} \mathbf{I}_{(A_n^{\Delta})^c \cap E_n \cap (B_n^{\gamma})^c}}{\mathbf{P}(\tau \geq m)}$, we apply Result 4 again and get

$$
\begin{aligned}
\mathbf{E}Z_{n,4}^2 &\leq \sum_{m \geq 1} \frac{\mathbf{E}\left[\left|Y_n^* - \hat{Y}_n^{m-1}\right|^2 \mathbf{I}_{(A_n^{\Delta})^c \cap E_n \cap (B_n^{\gamma})^c}\right]}{\mathbf{P}(\tau \geq m)} \\
&= \sum_{m \geq 1} \frac{\mathbf{E}\left[\mathbf{I}(Y_n^* \neq \hat{Y}_n^{m-1}) \cdot \mathbf{I}_{(A_n^{\Delta})^c \cap E_n \cap (B_n^{\gamma})^c}\right]}{\mathbf{P}(\tau \geq m)} && \text{because } \hat{Y}_n^m \text{ and } Y_n^* \text{ only take values in } \{0,1\} \\
&\leq \sum_{m \geq 1} \frac{\mathbf{P}\left(\{Y_n^* \neq \hat{Y}_n^{m-1}, \bar{X}_n \notin A^{\Delta}\} \cap (B_n^{\gamma})^c\right)}{\mathbf{P}(\tau \geq m)} && \text{due to } A_n^{\Delta} = \{\bar{X}_n \in A^{\Delta}\} \\
&= \sum_{m \geq 1} \frac{\mathbf{P}\left(\{Y_n^* \neq \hat{Y}_n^{m-1}, \bar{X}_n \notin A^{\Delta}\} \cap \{\mathcal{D}(\bar{J}_n) < l^*\}\right)}{\mathbf{P}(\tau \geq m)} && \text{due to } B_n^{\gamma} = \{\mathcal{D}(\bar{J}_n) \geq l^*\} \\
&= \sum_{m \geq 1} \sum_{k=0}^{l^*-1} \frac{\mathbf{P}\left(Y_n^* \neq \hat{Y}_n^{m-1}, \bar{X}_n \notin A^{\Delta} \mid \{\mathcal{D}(\bar{J}_n) = k\}\right)}{\mathbf{P}(\tau \geq m)} \cdot \mathbf{P}\left(\mathcal{D}(\bar{J}_n) = k\right) \\
&\leq \sum_{m \geq 1} \sum_{k=0}^{l^*-1} \frac{C_0 \rho_0^{m-1}}{\Delta^2 n^{\mu} \cdot \rho^{m-1}} && \text{due to (21)} \\
&= l^* \sum_{m \geq 1} \frac{C_0 \rho_0^{m-1}}{\Delta^2 n^{\mu} \cdot \rho^{m-1}} = \frac{C_0 l^*}{\Delta \cdot (1 - \frac{\rho_0}{\rho})} \cdot \frac{1}{n^{\mu}} = \boldsymbol{o}\left((n\nu[n,\infty))^{2l^*}\right).
\end{aligned}
$$
(61)

The last equality follows from the condition $\mu > 2l^*(\alpha - 1)$ prescribed in Proposition 1 and the fact that $n\nu[n,\infty) \in \mathcal{RV}_{-(\alpha-1)}(n)$ as $n \to \infty$. Combining (60) and (61) with the preliminary bound $(x+y)^2 \leq 2x^2 + 2y^2$, we yield $\mathbf{E}Z_{n,2}^2 \leq 2\mathbf{E}Z_{n,3}^2 + 2\mathbf{E}Z_{n,4}^2 = \boldsymbol{o}\left((n\nu[n,\infty))^{2l^*}\right)$ and conclude the proof of (56). $\quad\square$

6.2. Proof of Theorems 1 and 2 We stress again that Theorem 1 follows directly from Theorem 2 with $\kappa = 0$ (i.e., by disabling ARA from Algorithm 3). We devote the remainder of this section to proving Theorem 2.

Throughout Section 6.2, we fix the following constants and parameters. First, let $\beta \in [0,2)$ be the Blumenthal-Getoor index of $X(t)$ and $\alpha > 1$ be the regularly varying index of $\nu[x,\infty)$; see Assumption 1. Fix some

$$
\beta_+ \in (\beta, 2), \qquad \mu > 2l^*(\alpha - 1). \tag{62}
$$

48

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

This allows us to pick d, r large enough such that

$$r(2 - \beta_+) > \max\{2, \mu - 1\}, \qquad d > \max\{2, 2\mu - 1\} \tag{63}$$

for d in (37) and r in (31). Let $\lambda > 0$ be the constant in Assumption 2. Choose

$$\alpha_3 \in (0, \frac{1}{\lambda}), \qquad \alpha_4 \in (0, \frac{1}{2\lambda}). \tag{64}$$

Next, fix

$$\alpha_2 \in (0, \frac{\alpha_3}{2} \wedge 1). \tag{65}$$

Based on the chosen value of α_2, fix

$$\alpha_1 \in (0, \frac{\alpha_2}{\lambda}). \tag{66}$$

Pick

$$\delta \in (1/\sqrt{2}, 1). \tag{67}$$

Since we require α_2 to be strictly less than 1, there is some integer $\bar{m}$ such that

$$\delta^{m\alpha_2} - \delta^m \geq \frac{\delta^{m\alpha_2}}{2} \text{ and } \delta^{m\alpha_2} < a \qquad \forall m \geq \bar{m} \tag{68}$$

where $a > 0$ is the parameter in set A; see Assumption 3. Based on the values of δ and β_+, it holds for all $\kappa \in [0, 1)$ small enough that

$$\kappa^{2-\beta_+} < \frac{1}{2} < \delta^2 \tag{69}$$

Then, based on all previous choices, it holds for all $\rho_1 \in (0,1)$ close enough to 1 such that

$$\delta^{\alpha_1} < \rho_1, \tag{70}$$

$$\frac{\kappa^{2-\beta_+}}{\delta^2} < \rho_1 \tag{71}$$

$$\frac{1}{\sqrt{2}\delta} < \rho_1 \tag{72}$$

$$\delta^{\alpha_2 - \lambda\alpha_1} < \rho_1 \tag{73}$$

$$\delta^{1-\lambda\alpha_3} < \rho_1 \tag{74}$$

$$\delta^{-\alpha_2 + \frac{\alpha_3}{2}} < \rho_1, \tag{75}$$

$$(1/\sqrt{2}) \vee \kappa^{2-\beta_+} < \rho_1. \tag{76}$$

Lastly, pick $\rho_0 \in (\rho_1, 1)$. By picking a larger $\bar{m}$ if necessary, we can ensure that

$$m^2 \rho_1^m \le \rho_0^m \qquad \forall m \ge \bar{m}. \tag{77}$$

Next, we make a few observations. Given some non-negative integer k, let

$$\zeta_k(t) = \sum_{i=1}^{k} z_i \mathbf{I}_{[u_i,n]}(t) \tag{78}$$

where $0 < u_1 < u_2 < \ldots < u_k < n$ are the order statistics of k iid samples of $\text{Unif}(0,n)$, and z_i's are iid samples from $\nu(\cdot \cap [n\gamma, \infty))/\nu[n\gamma, \infty)$. We adopt the convention that $u_0 \equiv 0$ and $u_{k+1} \equiv 1$. Note that when $k = 0$, we set $\zeta_0(t) \equiv 0$ as the zero function, and set $I_1 = [0,n]$, $u_0 = 0$, and $u_1 = n$.

For $Y_n^*(\cdot)$ defined in (25) and $\hat{Y}_n^m(\cdot)$ defined in (38), note that

$$Y_n^*(\zeta_k) = \max_{i \in [k+1]} \mathbf{I}\{W_n^{(i),*}(\zeta_k) \ge na\}, \qquad \hat{Y}_n^m(\zeta_k) = \max_{i \in [k+1]} \mathbf{I}\{\hat{W}_n^{(i),m}(\zeta_k) \ge na\} \tag{79}$$

50

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

where

$$W_n^{(i),*}(\zeta_k) \triangleq \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} + \sum_{q=1}^{i-1}z_q + \sum_{j\geq 1}(\xi_j^{(i)})^+, \tag{80}$$

$$\hat{W}_n^{(i),m}(\zeta_k) \triangleq \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q),m} + \sum_{q=1}^{i-1}z_q + \sum_{j=1}^{m+\lceil \log_2(n^d)\rceil}(\xi_j^{(i),m})^+. \tag{81}$$

See (26)–(28) and (35) for the definitions $\xi_j^{(i)}$'s and $\xi_j^{(i),m}$'s, respectively. Also, define

$$\widetilde{W}_n^{(i),m}(\zeta_k) \triangleq \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} + \sum_{q=1}^{i-1}z_q + \sum_{j=1}^{m+\lceil \log_2(n^d)\rceil}(\xi_j^{(i)})^+. \tag{82}$$

As intermediate steps for the proof of Theorem 2, we present the following two results. Proposition 4 states that, using $\widetilde{W}_n^{(i),m}(\zeta_k)$ as an anchor, we see that $W_n^{(i),*}(\zeta_k)$ and $\hat{W}_n^{(i),m}(\zeta_k)$ would stay close enough with high probability, especially for large m. Proposition 5 then shows that it is unlikely for the law of $\widetilde{W}_n^{(i),m}(\zeta_k)$ to concentrate around any $y \in \mathbb{R}$.

PROPOSITION 4. *There exists some constant $C_1 \in (0, \infty)$ such that the inequality*

$$\mathbf{P}\left(\left|W_n^{(i),*}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k)\right| \vee \left|\hat{W}_n^{(i),m}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k)\right| > x\right) \leq \frac{C_1 \kappa^{m(2-\beta_+)}}{x^2 \cdot n^{r(2-\beta_+)-1}} + \frac{C_1}{x}\sqrt{\frac{1}{n^{d-1}\cdot 2^m}}$$

holds for any $k \in \mathbb{N}$, $i \in [k+1]$, $n \geq 1$, $m \in \mathbb{N}$, and $x > 0$.

PROPOSITION 5. *There exists some constant $C_2 \in (0, \infty)$ such that the inequality*

$$\mathbf{P}\left(\widetilde{W}_n^{(i),m}(\zeta_k) \in \left[y - \frac{\delta^m}{\sqrt{n}}, y + \frac{\delta^m}{\sqrt{n}}\right] \text{ for some } i \in [k+1]\right) \leq (k+1)\cdot C_2\rho_0^m$$

holds for any $k \in \mathbb{N}$, $i \in [k+1]$, $n \geq 1$, $m \geq \bar{m}$, and $y > \delta^{m\alpha_2}$.

Equipped with Propositions 4 and 5, we are able to prove the main results of Section 3.6, i.e., Theorem 2.

Proof of Theorem 2   In light of Proposition 1, it suffices to verify conditions (20) and (21).

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

51

Verification of (20). Conditioning on $\{\mathcal{D}(\bar{J}_n) = k\}$, the conditional law of $J_n = \{J_n(t) : t \in [0, n]\}$ is

the same as the law of the process ζ_k specified in (78). This implies $\mathbf{P}\big(Y_n^*(J_n) \neq \hat{Y}_n^m(J_n) \mid \mathcal{D}(\bar{J}_n) = k\big) =$

$\mathbf{P}\big(Y_n^*(\zeta_k) \neq \hat{Y}_n^m(\zeta_k)\big)$. Next, on event

$$
\bigcap_{i \in [k+1]} \left(\left\{ \left| W_n^{(i),*}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k) \right| \vee \left| \hat{W}_n^{(i),m}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k) \right| \le \frac{\delta^m}{\sqrt{n}} \right\} \right.
$$
$$
\left. \cap \left\{ \widetilde{W}_n^{(i),m}(\zeta_k) \notin \left[na - \frac{\delta^m}{\sqrt{n}}, na + \frac{\delta^m}{\sqrt{n}} \right] \right\} \right),
$$

we must have (for any $i \in [k + 1]$) $W_n^{(i),*}(\zeta_k) \vee \hat{W}_n^{(i),m}(\zeta_k) < na$ or $W_n^{(i),*}(\zeta_k) \wedge \hat{W}_n^{(i),m}(\zeta_k) > na$. It then

follows from (79) that, on this event, we have $Y_n^*(\zeta_k) = \hat{Y}_n^m(\zeta_k)$. Therefore,

$$
\mathbf{P}\big(Y_n^*(\zeta_k) \neq \hat{Y}_n^m(\zeta_k)\big) \tag{83}
$$
$$
\le \sum_{i \in [k+1]} \mathbf{P}\left(\left| W_n^{(i),*}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k) \right| \vee \left| \hat{W}_n^{(i),m}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k) \right| > \frac{\delta^m}{\sqrt{n}} \right)
$$
$$
+ \mathbf{P}\left(\widetilde{W}_n^{(i),m}(\zeta_k) \in \left[na - \frac{\delta^m}{\sqrt{n}}, na + \frac{\delta^m}{\sqrt{n}} \right] \text{ for some } i \in [k+1] \right).
$$

Applying Proposition 4 (with $x = \delta^m/\sqrt{n}$), we get (for any $i \in [k + 1]$)

$$
\mathbf{P}\left(\left| W_n^{(i),*}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k) \right| \vee \left| \hat{W}_n^{(i),m}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k) \right| > \frac{\delta^m}{\sqrt{n}} \right)
$$
$$
\le C_1 \cdot \left[\frac{\kappa^{m(2-\beta_+)} \cdot n}{\delta^{2m} \cdot n^{r(2-\beta_+)-1}} + \frac{\sqrt{n}}{(\sqrt{2}\delta)^m \cdot \sqrt{n^{d-1}}} \right]
$$
$$
= C_1 \cdot \left[\left(\frac{\kappa^{2-\beta_+}}{\delta^2} \right)^m \cdot \frac{1}{n^{r(2-\beta_+)-2}} + \left(\frac{1}{\sqrt{2}\delta} \right)^m \cdot \sqrt{\frac{1}{n^{d-2}}} \right]
$$
$$
\le C_1 \cdot \left[\left(\frac{\kappa^{2-\beta_+}}{\delta^2} \right)^m + \left(\frac{1}{\sqrt{2}\delta} \right)^m \right] \qquad \text{due to the choices of } d \text{ and } r \text{ in (63)}
$$
$$
\le 2C_1 \rho_0^m \qquad \text{due to the choices in (71) and (72), and } \rho_0 \in (\rho_1, 1). \tag{84}
$$

On the other hand, due to (68), we have $na - \delta^m \alpha_2 \geq a - \delta^m \alpha_2 > 0$. for all $n \geq 1$ and $m \geq \bar{m}$. This allows us to apply Proposition 5 (with $y = na$) and yield (for any $i \in [k+1]$)

$$\mathbf{P}\left(\widetilde{W}_n^{(i),m}(\zeta_k) \in \left[na - \frac{\delta^m}{\sqrt{n}}, na + \frac{\delta^m}{\sqrt{n}} \right] \text{ for some } i \in [k+1] \right) \leq (k+1) \cdot C_2 \rho_0^m \qquad \forall m \geq \bar{m}. \qquad (85)$$

Plugging (84) and (85) into (83), we conclude the proof by setting $C_0 = 2C_1 + C_2$.

Verification of (21). Fix some $\Delta \in (0,1)$ and $k = 0, 1, \ldots, l^* - 1$. Again, conditioning on $\{\mathcal{D}(\bar{J}_n) = k\}$, the conditional law of $J_n = \{J_n(t) : t \in [0, n]\}$ is the same as the law of the process ζ_k specified in (78). This implies

$$\mathbf{P}\left(Y_n^*(J_n) \neq \hat{Y}_n^m(J_n), \ \bar{X}_n \notin A^\Delta \ \Big| \ \mathcal{D}(\bar{J}_n) = k \right)$$

$$= \mathbf{P}\left(Y_n^*(J_n) \neq \hat{Y}_n^m(J_n), \ \sup_{t \in [0,n]} X(t) < n(a - \Delta) \ \Big| \ \mathcal{D}(\bar{J}_n) = k \right) \qquad \text{by definition of set } A^\Delta$$

$$= \mathbf{P}\left(\max_{i \in [k+1]} \hat{W}_n^{(i),m}(\zeta_k) \geq na, \ \max_{i \in [k+1]} W_n^{(i),*}(\zeta_k) < n(a - \Delta) \right)$$

$$\leq \sum_{i \in [k+1]} \mathbf{P}\left(\left| \hat{W}_n^{(i),m}(\zeta_k) - W_n^{(i),*}(\zeta_k) \right| > n\Delta \right)$$

$$\leq \sum_{i \in [k+1]} \mathbf{P}\left(\left| W_n^{(i),*}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k) \right| \vee \left| \hat{W}_n^{(i),m}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k) \right| > \frac{n\Delta}{2} \right)$$

$$\leq (k+1) \cdot \left[\frac{4C_1}{\Delta^2 n^2} \cdot \frac{\kappa^{m(2-\beta_+)}}{n^{r(2-\beta_+)-1}} + \frac{2C_1}{\Delta} \cdot \frac{1}{n} \sqrt{\frac{1}{n^{d-1} \cdot 2^m}} \right] \qquad \text{by Proposition 4}$$

$$= (k+1) \cdot \left[\frac{4C_1}{\Delta^2} \cdot \frac{\kappa^{m(2-\beta_+)}}{n^{r(2-\beta_+)+1}} + \frac{2C_1}{\Delta} \cdot \frac{(1/\sqrt{2})^m}{n^{\frac{d+1}{2}}} \right]$$

$$\leq \frac{k+1}{n^\mu} \cdot \left[\frac{4C_1}{\Delta^2} \cdot \kappa^{m(2-\beta_+)} + \frac{2C_1}{\Delta} \cdot (1/\sqrt{2})^m \right] \qquad \text{by the choices of } r \text{ and } d \text{ in (63)}$$

$$\leq \frac{k+1}{n^\mu} \cdot \left[\frac{4C_1}{\Delta^2} \cdot \rho_0^m + \frac{2C_1}{\Delta} \cdot \rho_0^m \right] \qquad \text{due to the choice of } \rho_1 \text{ in (76) and } \rho_0 \in (\rho_1, 1).$$

Due to $\Delta \in (0, 1)$ (and hence $\frac{1}{\Delta} < \frac{1}{\Delta^2}$) and $k \leq l^* - 1$, we conclude the proof by setting $C_0 = 6l^* C_1$. $\quad \square$

The rest of this section is devoted to proving Propositions 4 and 5. First, we collect a useful result.

Result 5 (Lemma 1 of González Cázares et al. (2022)) *Let ν be the Lévy measure of a Lévy process X. Let*
$I_0^p(\nu) \triangleq \int_{(-1,1)} |x|^p \nu(dx)$. *Suppose that $\beta < 2$ for the Blumenthal-Getoor index $\beta \triangleq \inf\{p > 0 : I_0^p(\nu) < \infty\}$.*
Then

$$\int_{(-\kappa,\kappa)} x^2 \nu(dx) \le \kappa^{2-\beta_+} I_0^{\beta_+}(\nu) \qquad \forall \kappa \in (0,1], \ \beta_+ \in (\beta, 2).$$

Next, we prepare two lemmas regarding the expectations of the supremum of Ξ_n (see (17) for the definition) and the difference between Ξ_n and $\breve{\Xi}_n^m$ (see (34)).

LEMMA 3. *There exists a constant $C_X < \infty$ (depending only on the law of Lévy process $X(t)$) such that*

$$\mathbf{E}\left[\sup_{s \in [0,t]} \Xi_n(t)\right] \le C_X(\sqrt{t} + t) \qquad \forall t > 0, \ n \ge 1.$$

Proof Recall that the generating triplet of X is (c_X, σ, ν) and for the Blumenthal-Getoor index $\beta \triangleq \inf\{p > 0 : \int_{(-1,1)} |x|^p \nu(dx) < \infty\}$ we have $\beta < 2$; see Assumption 1. Fix some $\beta_+ \in (1 \vee \beta, 2)$ in this proof. We prove the lemma for

$$C_X \triangleq \max\left\{|\sigma|\sqrt{\frac{2}{\pi}} + 2\sqrt{I_0^{\beta_+}(\nu)}, \ (c_X)^+ + I_+^1(\nu) + 2 I_0^{\beta_+}(\nu)\right\}$$

where $(x)^+ = x \vee 0$, $I_+^1(\nu) = \int_{[1,\infty)} x \nu(dx)$, and $I_0^p(\nu) = \int_{(-1,1)} |x|^p \nu(dx)$.

Recall that Ξ_n is a Lévy process with generating triplet $(c_X, \sigma, \nu|_{(-\infty, n\gamma)})$. Let $\nu_n \triangleq \nu|_{(-\infty, n\gamma)}$. It follows from Lemma 2 of González Cázares et al. (2022) (specifically, by setting $t = T$ in equation (26)) that, for all $t > 0$ and $n \ge 1$,

$$\mathbf{E}\sup_{s \in [0,t]} \Xi_n(t) \le \left(|\sigma|\sqrt{\frac{2}{\pi}} + 2\sqrt{I_0^{\beta_+}(\nu_n)}\right)\sqrt{t} + \left((c_X)^+ + I_+^1(\nu_n) + 2 I_0^{\beta_+}(\nu_n)\right)t. \tag{86}$$

Note that $I_0^{\beta_+}(\nu_n) = \int_{(-1,1)} |x|^p \nu_n(dx) = \int_{(-1,1) \cap (-\infty, n\gamma)} |x|^p \nu(dx) \le I_0^{\beta_+}(\nu)$ and $I_+^1(\nu_n) = \int_{[1,\infty)} x\nu_n(dx) = \int_{[1,\infty) \cap (-\infty, n\gamma)} x\nu(dx) \le I_+^1(\nu)$. Plugging these two bounds into (86), we conclude the proof. $\square$

LEMMA 4. *There exists some $C \in (0, \infty)$ (only depending on the choice of $\beta_+ \in (\beta, 2)$ in (62) and the law of Lévy process X) such that*

$$\mathbf{P}\left(\sup_{t \in [0,n]} \left|\Xi_n(t) - \breve{\Xi}_n^m(t)\right| > x\right) \le \frac{C\kappa^{m(2-\beta_+)}}{x^2 n^{r(2-\beta_+)-1}} \qquad \forall x > 0, \ n \ge 1, \ m \in \mathbb{N}$$

where r is the parameter in the truncation threshold $\kappa_{n,m} = \kappa^m / n^r$ (see (31)).

54

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Proof From the definitions of Ξ_n and $\breve{\Xi}_n^m$ in (32) and (34), respectively, we have

$$\Xi_n(t) - \breve{\Xi}_n^m(t) \stackrel{d}{=} X^{(-\kappa_{n,m}, \kappa_{n,m})}(t) - \bar{\sigma}(\kappa_{n,m}) B(t)$$

where $X^{(-c,c)}$ is the Lévy process with generating triplet $(0, 0, \nu|_{(-c,c)})$, $\kappa_{n,m} = \kappa^m/n^r$, and B is a standard Brownian motion independent of $X^{(-\kappa_{n,m}, \kappa_{n,m})}$. In particular, $X^{(-\kappa_{n,m}, \kappa_{n,m})}$ is a martingale with variance $var\left[X^{(-\kappa_{n,m}, \kappa_{n,m})}(1)\right] = \bar{\sigma}^2(\kappa_{n,m})$; see (33) for the definition of $\bar{\sigma}^2(\cdot)$. Therefore,

$$
\begin{aligned}
&\mathbf{P}\left(\sup_{t \in [0,n]} \left|\Xi_n(t) - \breve{\Xi}_n^m(t)\right| > x\right) \\
&\leq \frac{1}{x^2} \mathbf{E}\left|X^{(-\kappa_{n,m}, \kappa_{n,m})}(n) - \bar{\sigma}(\kappa_{n,m}) B(n)\right|^2 \qquad \text{using Doob's inequality} \\
&= \frac{2n}{x^2} \bar{\sigma}^2(\kappa_{n,m}) \qquad \text{due to the independence between } X^{(-\kappa_{n,m}, \kappa_{n,m})} \text{ and } B \\
&\leq \frac{2n}{x^2} \cdot \kappa_{n,m}^{2-\beta_+} I_0^{\beta_+}(\nu) \qquad \text{using Result 5} \\
&= \frac{2 I_0^{\beta_+}(\nu)}{x^2} \cdot \frac{n \kappa^{m(2-\beta_+)}}{n^{r(2-\beta_+)}} = \frac{2 I_0^{\beta_+}(\nu)}{x^2} \cdot \frac{\kappa^{m(2-\beta_+)}}{n^{r(2-\beta_+)-1}} \qquad \text{due to } \kappa_{n,m} = \kappa^m/n^r.
\end{aligned}
$$

To conclude the proof, we set $C = 2 I_0^{\beta_+}(\nu) = 2 \int_{(-1,1)} \int |x|^{\beta_+} \nu(dx)$. □

To facilitate the presentation of the next few lemmas, consider a slightly more general version of the stick-breaking procedure in (26)–(35) to account for arbitrary stick lengths. Specifically, for any $l > 0$, let

$$l_1(l) = V_1 \cdot l, \qquad l_j(l) = V_j \cdot \left(l - l_1(l) - l_2(l) - \cdots - l_{j-1}(l)\right) \quad \forall j \geq 2, \tag{87}$$

where V_j's are iid copies of $\text{Unif}(0,1)$. Independent of V_j's, for any n and m, let Ξ_n and $\breve{\Xi}_n^m$ be Lévy processes with joint law specified in (32) and (34), respectively. Conditioning on the values of $l_j(l)$, define $\xi_j^{[n]}(l), \xi_j^{[n],m}(l)$ using (for all $j \geq 1$)

$$\left(\xi_j^{[n]}(l), \xi_j^{[n],0}(l), \xi_j^{[n],1}(l), \xi_j^{[n],2}(l), \dots\right) = \left(\Xi_n(l_j(l)), \breve{\Xi}_n^0(l_j(l)), \breve{\Xi}_n^1(l_j(l)), \breve{\Xi}_n^2(l_j(l)), \dots\right). \tag{88}$$

LEMMA 5. *There exists some $C \in (0, \infty)$ (only depending on the choice of $\beta_+ \in (\beta, 2)$ in (62) and the law of Lévy process X) such that, for all $m \in \mathbb{N}$ and $n \geq 1$,*

$$\mathbf{P}\left(\left|\sum_{j=1}^{m+\lceil \log_2(n^d)\rceil} \left(\xi_j^{[n]}(l)\right)^+ - \sum_{j=1}^{m+\lceil \log_2(n^d)\rceil} \left(\xi_j^{[n],m}(l)\right)^+\right| > y\right) \leq \frac{C\kappa^{m(2-\beta_+)}}{y^2 n^{r(2-\beta_+)-1}} \qquad \forall y > 0, \ l \in [0, n]$$

where r is the parameter in the truncation threshold $\kappa_{n,m} = \kappa^m/n^r$ (see (31)) and $(x)^+ = x \vee 0$.

Proof   For notational simplicity, set $k(n) = \lceil \log_2(n^d)\rceil$. Due to $|(x)^+ - (y)^+| \leq |x - y|$,

$$\mathbf{P}\left(\left|\sum_{j=1}^{m+k(n)} \left(\xi_j^{[n]}(l)\right)^+ - \sum_{j=1}^{m+k(n)} \left(\xi_j^{[n],m}(l)\right)^+\right| > y\right)$$

$$\leq \mathbf{P}\left(\sum_{j=1}^{m+k(n)} \left|\left(\xi_j^{[n]}(l)\right)^+ - \left(\xi_j^{[n],m}(l)\right)^+\right| > y\right) \leq \mathbf{P}\left(\sum_{j=1}^{m+k(n)} \underbrace{\left|\xi_j^{[n]}(l) - \xi_j^{[n],m}(l)\right|}_{\triangleq q_j} > y\right). \tag{89}$$

Furthermore, we claim the existence of some constant $\tilde{C} \in (0, \infty)$ such that (for any $y, d > 0$, $l \in [0.n]$, and any $n \geq 1$, $m \in \mathbb{N}$)

$$\mathbf{P}\left(\sum_{j=1}^{m+k(n)} |q_j| > y\right) \leq \tilde{C} \cdot \frac{n\bar{\sigma}^2(\kappa_{n,m})}{y^2}. \tag{90}$$

Then using Result 5, we yield

$$n\bar{\sigma}^2(\kappa_{n,m}) \leq n \cdot \kappa_{n,m}^{2-\beta_+} I_0^{\beta_+}(\nu) = \frac{\kappa^{m(2-\beta_+)}}{n^{r(2-\beta_+)-1}} \cdot I_0^{\beta_+}(\nu)$$

where $I_0^{\beta_+}(\nu) = \int_{(-1,1)} \int |x|^{\beta_+} \nu(dx)$. Setting $C = \tilde{C} I_0^{\beta_+}(\nu)$, we conclude the proof.

Now, it only remains to prove claim (90). Let $\chi = 2^{1/4}$. Note that

$$1 = (\chi - 1)\sum_{j \geq 1} \frac{1}{\chi^j} \geq (\chi - 1)\left(\frac{1}{\chi} + \frac{1}{\chi^2} + \cdots + \frac{1}{\chi^{m+k(n)}}\right).$$

As a result,

$$\mathbf{P}\left(\sum_{j=1}^{k(n)+m} |q_j| > y\right) \leq \mathbf{P}\left(\sum_{j=1}^{k(n)+m} |q_j| > y(\chi - 1)\sum_{j=1}^{k(n)+m} \frac{1}{\chi^j}\right) \leq \sum_{j=1}^{k(n)+m} \mathbf{P}\left(|q_j| > y \cdot \frac{\chi - 1}{\chi^j}\right) \tag{91}$$

Next, we bound each $\mathbf{P}(|q_j| > y\frac{\chi-1}{\chi^j})$. Conditioning on $l_j(l) = t$ (for any $t \in [0, l]$), we get

$$
\begin{aligned}
\mathbf{P}\left(|q_j| > y\frac{\chi-1}{\chi^j} \,\middle|\, l_j(l) = t\right) &= \mathbf{P}\left(\left|\Xi_n(t) - \breve{\Xi}_n^m(t)\right| > y\frac{\chi-1}{\chi^j}\right) \qquad \text{due to (88)} \\
&\leq \frac{\chi^{2j}}{y^2(\chi-1)^2}\mathbf{E}\left|X^{(-\kappa_{n,m},\kappa_{n,m})}(t) - \bar{\sigma}(\kappa_{n,m})B(t)\right|^2 \\
&= \frac{\chi^{2j}}{y^2(\chi-1)^2} \cdot 2\bar{\sigma}^2(\kappa_{n,m})t \\
\implies \mathbf{P}\left(|q_j| > y\frac{\chi-1}{\chi^j}\right) &\leq \frac{\chi^{2j}}{y^2(\chi-1)^2} \cdot 2\bar{\sigma}^2(\kappa_{n,m}) \cdot \mathbf{E}[l_j(l)] \\
&= \frac{\sqrt{2}^j}{y^2(2^{1/4}-1)^2} \cdot 2\bar{\sigma}^2(\kappa_{m,n}) \cdot \mathbf{E}[l_j(l)] \qquad \text{due to } \chi = 2^{1/4} \\
&= \frac{\sqrt{2}^j}{y^2(2^{1/4}-1)^2} \cdot 2\bar{\sigma}^2(\kappa_{m,n}) \cdot \frac{l}{2^j} \qquad \text{by definition of } l_j(l) \text{ in (87)} \\
&\leq \frac{2}{(2^{1/4}-1)^2\sqrt{2}^j} \cdot \frac{n\bar{\sigma}^2(\kappa_{m,n})}{y^2} \qquad \text{due to } l \leq n.
\end{aligned}
$$

Therefore, in (91), we get

$$
\mathbf{P}\left(\sum_{j=1}^{k(n)+m}|q_j| > y\right) \leq \frac{n\bar{\sigma}^2(\kappa_{m,n})}{y^2}\sum_{j\geq 1}\frac{2}{(2^{1/4}-1)^2\sqrt{2}^j} = \frac{n\bar{\sigma}^2(\kappa_{m,n})}{y^2} \cdot \underbrace{\frac{2\sqrt{2}}{(2^{1/4}-1)^2(\sqrt{2}-1)}}_{\triangleq \bar{C}},
$$

thus establishing claim (90). □

LEMMA 6.   *Let $n \in \mathbb{Z}_+$ and $l \in [0, n]$. Let $C_X < \infty$ be the constant characterized in Lemma 3 that only depends on the law of Lévy process X. The inequality*

$$
\mathbf{P}\left(\sum_{j>m+\lceil\log_2(n^d)\rceil}\left(\xi_j^{[n]}(l)\right)^+ > x\right) \leq \frac{2C_X}{x}\sqrt{\frac{1}{n^{d-1}\cdot 2^m}}
$$

holds for all $x > 0$, $n \geq 1$, and $m \geq 0$, where $(x)^+ = x \vee 0$.

Proof   For this proof, we adopt the notation $\breve{l}_k(l) \triangleq l - l_1(l) - l_2(l) - \ldots - l_k(l)$ for the remaining stick length after the first k sticks. Conditioning on $\breve{l}_{m+\lceil\log_2(n^d)\rceil}(l) = t$,

$$
\begin{aligned}
\mathbf{P}\left(\sum_{j>m+\lceil\log_2(n^d)\rceil}\left(\xi_j^{[n]}(l)\right)^+ > x \,\middle|\, \breve{l}_{m+\lceil\log_2(n^d)\rceil}(l) = t\right) &= \mathbf{P}\left(\sup_{s\in[0,t]}\Xi_n(s) > x\right) \qquad \text{by Result 3} \\
&\leq \frac{C_X}{x}(\sqrt{t}+t) \qquad \text{using Lemma 3.}
\end{aligned}
$$

Therefore, unconditionally,

$$\mathbf{P}\Big(\sum_{j>m+\lceil\log_2(n^d)\rceil}\big(\xi_j^{[n]}(l)\big)^+ > x\Big) \le \frac{C_X}{x}\mathbf{E}\Big[\sqrt{\breve{l}_{m+\lceil\log_2(n^d)\rceil}(l)} + \mathbf{E}\breve{l}_{m+\lceil\log_2(n^d)\rceil}(l)\Big]$$

$$\le \frac{C_X}{x}\Big[\sqrt{\mathbf{E}\breve{l}_{m+\lceil\log_2(n^d)\rceil}(l)} + \mathbf{E}\breve{l}_{m+\lceil\log_2(n^d)\rceil}(l)\Big]$$

The last line follows from Jensen's inequality. Lastly, by definition of $l_j(l)$'s in (87), we have

$$\mathbf{E}\breve{l}_{m+\lceil\log_2(n^d)\rceil}(l) = \frac{l}{2^{m+\lceil\log_2(n^d)\rceil}} \le \frac{l}{2^m\cdot n^d} \le \frac{n}{2^m\cdot n^d} = \frac{1}{n^{d-1}\cdot 2^m} \qquad \text{due to } l \in [0,n].$$

This concludes the proof.   □

LEMMA 7.   *Let $n \in \mathbb{Z}_+$ and $l \in [0,n]$. Let C and λ be the constants in Assumption 2. Let $C_X < \infty$ be the*

constant characterized in Lemma 3 that only depends on the law of Lévy process X. The inequality

$$\mathbf{P}\Big(\sum_{j=1}^{m+\lceil\log_2(n^d)\rceil}\big(\xi_j^{[n]}(l)\big)^+ \in [y, y+c]\Big)$$

$$\le C\frac{(m + (\lceil\log_2(n^d)\rceil)n^{\alpha_4\lambda}}{\delta^{\alpha_3\lambda}}c + 4C_X\big(m^2 + (\lceil\log_2(n^d)\rceil)^2\big)\frac{\delta^{\alpha_3/2}}{y_0\cdot n^{\alpha_4/2}}.$$

holds for all $y \ge y_0 > 0$, $c > 0$, $n \ge 1$, and $m \in \mathbb{N}$.

Proof   To simplify notations, in this proof we set $k(n) = \lceil\log_2(n^d)\rceil$ and write $l_j = l_j(l)$ when there

is no ambiguity. For the sequence of random variables $(l_1, \cdots, l_{m+k(n)})$, let $\tilde{l}_1 \ge \tilde{l}_2 \ge \cdots \ge \tilde{l}_{m+k(n)}$ be its

order statistics. Given any ordered positive real sequence $t_1 \ge t_2 \ge \cdots \ge t_{m+k(n)} > 0$, by conditioning on

$\tilde{l}_j = t_j \ \forall j \in [m+k(n)]$, it follows from (88) that

$$\mathbf{P}\Big(\sum_{j=1}^{m+k(n)}\big(\xi_j^{[n]}(l)\big)^+ \in [y, y+c]\ \Big|\ \tilde{l}_j = t_j\ \forall j \in [m+k(n)]\Big) = \mathbf{P}\Big(\sum_{j=1}^{m+k(n)}\big(\Xi_n^{(j)}(t_j)\big)^+ \in [y, y+c]\Big) \qquad (92)$$

where $\Xi_n^{(j)}$'s are iid copies of the Lévy processes $\Xi_n = X^{<n\gamma}$. Next, fix $\eta = \delta^{m\alpha_3}/n^{\alpha_4}$. Given the sequence of

real numbers t_j's, we define $J \triangleq \#\{j \in [m+k(n)] : t_j > \eta\}$ as the number of elements in the sequence that

are larger than η. In case that $t_1 \leq \eta$, we set $J = 0$. With J defined, we consider a decomposition of events in

(92) based on the first $j \in [m+k(n)]$ such that $\Xi_n^{(j)}(t_j) > 0$ (and hence $\left(\Xi_n^{(j)}(t_j)\right)^+ > 0$), especially if such

t_j is larger than η or not. To be specific,

$$
\mathbf{P}\left(\sum_{j=1}^{m+k(n)} \left(\Xi_n^{(j)}(t_j)\right)^+ \in [y, y+c] \right)
$$

$$
= \underbrace{\sum_{j=1}^{J} \mathbf{P}\left(\Xi_n^{(i)}(t_i) \leq 0 \ \forall i \in [j-1]; \ \Xi_n^{(j)}(t_j) > 0; \ \sum_{i=j}^{m+k(n)} \left(\Xi_n^{(i)}(t_i)\right)^+ \in [y, y+c] \right)}_{\triangleq p_j}
$$

$$
+ \underbrace{\mathbf{P}\left(\Xi_n^{(i)}(t_i) \leq 0 \ \forall i \in [J]; \ \sum_{j=J+1}^{m+k(n)} \left(\Xi_n^{(j)}(t_j)\right)^+ \in [y, y+c] \right)}_{\triangleq p_*}. \tag{93}
$$

We first bound terms p_j's. For any $j \in [J]$, observe that

$$
p_j \leq \mathbf{P}\left(\Xi_n^{(j)}(t_j) > 0; \ \sum_{i=j}^{m+k(n)} \left(\Xi_n^{(i)}(t_i)\right)^+ \in [y, y+c] \right)
$$

$$
= \int_{\mathbb{R}} \mathbf{P}\left(\Xi_n^{(j)}(t_j) \in [y-x, y-x+c] \cap (0, \infty) \right) \mathbf{P}\left(\sum_{i=j+1}^{m+k(n)} \left(\Xi_n^{(i)}(t_i)\right)^+ \in dx \right)
$$

$$
\leq \frac{Cc}{t_j^{\lambda} \wedge 1} \qquad \text{by Assumption 2}
$$

$$
\leq \frac{Cn^{\alpha_4 \lambda}}{\delta^{m\alpha_3 \lambda}} \cdot c \qquad \text{due to } j \leq J, \text{ and hence } t_j > \eta = \delta^{m\alpha_3}/n^{\alpha_4}. \tag{94}
$$

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

59

On the other hand,

$$
\begin{aligned}
p_* &\leq \mathbf{P}\left(\sum_{j=J+1}^{m+k(n)} \left(\Xi_n^{(j)}(t_j) \right)^+ \in [y, y+c] \right) \leq \mathbf{P}\left(\sum_{j=J+1}^{m+k(n)} \left(\Xi_n^{(j)}(t_j) \right)^+ \geq y_0 \right) \qquad \text{due to } y \geq y_0 > 0 \\
&\leq \sum_{j=J+1}^{m+k(n)} \mathbf{P}\left(\Xi_n^{(j)}(t_j) \geq y_0/N \right) \qquad \text{with } N \triangleq m + k(n) - J \\
&\leq \sum_{j=J+1}^{m+k(n)} \frac{C_X(\sqrt{t_j} + t_j) \cdot N}{y_0} \qquad \text{by Lemma 3} \\
&\leq \sum_{j=J+1}^{m+k(n)} \frac{C_X(\sqrt{\eta} + \eta) \cdot N}{y_0} \qquad \text{due to } j > J, \text{ and hence } t_j \leq \eta = \delta^{m\alpha_3}/n^{\alpha_4} \\
&= N^2 \cdot \frac{C_X(\sqrt{\eta} + \eta)}{y_0} \leq \left(m + k(n) \right)^2 \cdot \frac{C_X(\sqrt{\eta} + \eta)}{y_0} \qquad \text{due to } N \leq m + k(n) \\
&\leq 2C_X\left(m^2 + (\lceil \log_2(n^d) \rceil)^2 \right) \frac{\sqrt{\eta} + \eta}{y_0} \qquad \text{using } (u+v)^2 \leq 2(u^2 + v^2) \\
&\leq 4C_X\left(m^2 + (\lceil \log_2(n^d) \rceil)^2 \right) \frac{\sqrt{\eta}}{y_0} = 4C_X\left(m^2 + (\lceil \log_2(n^d) \rceil)^2 \right) \frac{\delta^{m\alpha_3/2}}{y_0 \cdot n^{\alpha_4/2}}. \qquad (95)
\end{aligned}
$$

Plugging (94) and (95) into (93), we yield

$$
\begin{aligned}
&\mathbf{P}\left(\sum_{j=1}^{m+k(n)} \left(\xi_j^{[n]}(l) \right)^+ \in [y, y+c] \ \bigg| \ \tilde{l}_j = t_j \ \forall j \in [m+k(n)] \right) \\
&\leq J \cdot \frac{Cn^{\alpha_4\lambda}}{\delta^{m\alpha_3\lambda}} c + 4C_X\left(m^2 + (\lceil \log_2(n^d) \rceil)^2 \right) \frac{\delta^{m\alpha_3/2}}{y_0 \cdot n^{\alpha_4/2}} \\
&\leq C \frac{(m + \lceil \log_2(n^d) \rceil) n^{\alpha_4\lambda}}{\delta^{m\alpha_3\lambda}} c + 4C_X\left(m^2 + (\lceil \log_2(n^d) \rceil)^2 \right) \frac{m\delta^{\alpha_3/2}}{y_0 \cdot n^{\alpha_4/2}} \qquad \text{due to } J \leq m + \lceil \log_2(n^d) \rceil.
\end{aligned}
$$

To conclude the proof, just note that the inequality above holds when conditioning on any sequence of $t_1 \geq t_2 \geq \cdots \geq t_{m+k(n)} > 0$, so it would also hold unconditionally. $\quad \square$

Now, we are ready to prove Propositions 4 and 5.

Proof of Proposition 4 In this proof, we fix some $k \in \mathbb{N}$, $n \geq 1$, and $m \in \mathbb{N}$. Let the process ζ_k be defined as in (78). Recall the definitions of $W_n^{(i),*}(\zeta_k)$, $\hat{W}_n^{(i),m}(\zeta_k)$, and $\widetilde{W}_n^{(i),m}(\zeta_k)$ in (80)–(82). See also (26)–(35) for the definitions $\xi_j^{(i)}$'s and $\xi_j^{(i),m}$'s.

60

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

To simplify notations, define $t(n) = \lceil \log_2(n^d) \rceil$. Define events

$$E_1^{(i)}(x) \triangleq \left\{ \left| \sum_{q=1}^{i-1} \sum_{j \geq 0} \xi_j^{(q)} - \sum_{q=1}^{i-1} \sum_{j \geq 0} \xi_j^{(q),m} \right| \leq \frac{x}{2} \right\},$$

$$E_2^{(i)}(x) \triangleq \left\{ \left| \sum_{j=1}^{m+t(n)} (\xi_j^{(i)})^+ - \sum_{j=1}^{m+t(n)} (\xi_j^{(i),m})^+ \right| \leq \frac{x}{2} \right\},$$

$$E_3^{(i)}(x) \triangleq \left\{ \sum_{j \geq m+t(n)+1} (\xi_j^{(i)})^+ \leq x \right\}.$$

Note that on event $E_1^{(i)}(x) \cap E_2^{(i)}(x) \cap E_3^{(i)}(x)$, we must have $|W_n^{(i),*}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k)| \leq x$ and $|\hat{W}_n^{(i),m}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k)| \leq x$. As a result,

$$\mathbf{P}\left(\left| W_n^{(i),*}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k) \right| \vee \left| \hat{W}_n^{(i),m}(\zeta_k) - \widetilde{W}_n^{(i),m}(\zeta_k) \right| > x \right) \leq \sum_{q=1}^{3} \mathbf{P}\left(\left(E_q^{(i)}(x) \right)^c \right).$$

Furthermore, we claim the existence of constant $(\widetilde{C}_q)_{q=1,2,3}$, the values of which do not depend on x, k, n, and m, such that (for any $x > 0$ and $i \in [k+1]$)

$$\mathbf{P}\left(\left(E_1^{(i)}(x) \right)^c \right) = \mathbf{P}\left(\left| \sum_{q=1}^{i-1} \sum_{j \geq 0} \xi_j^{(q)} - \sum_{q=1}^{i-1} \sum_{j \geq 0} \xi_j^{(q),m} \right| > \frac{x}{2} \right) \leq \frac{\widetilde{C}_1 \kappa^{m(2-\beta_+)}}{x^2 n^{r(2-\beta_+)-1}}, \tag{96}$$

$$\mathbf{P}\left(\left(E_2^{(i)}(x) \right)^c \right) = \mathbf{P}\left(\left| \sum_{j=1}^{m+t(n)} (\xi_j^{(i)})^+ - \sum_{j=1}^{m+t(n)} (\xi_j^{(i),m})^+ \right| > \frac{x}{2} \right) \leq \frac{\widetilde{C}_2 \kappa^{m(2-\beta_+)}}{x^2 n^{r(2-\beta_+)-1}}, \tag{97}$$

$$\mathbf{P}\left(\left(E_3^{(i)}(x) \right)^c \right) = \mathbf{P}\left(\sum_{j \geq m+t(n)+1} (\xi_j^{(i)})^+ > x \right) \leq \frac{\widetilde{C}_3}{x} \sqrt{\frac{1}{n^{d-1} \cdot 2^m}}. \tag{98}$$

This allows us to conclude the proof by setting $C_1 = \sum_{q=1}^{3} \widetilde{C}_q$. Now, it remains to prove claims (96)–(98).

Proof of Claim (96). The claim is trivial if $i \leq 1$, so we only consider the case where $i \geq 2$. Due to the coupling between $\xi_j^{(i)}$ and $\xi_j^{(i),m}$ in (35)(36), we have

$$\left(\sum_{q=1}^{i-1} \sum_{j \geq 0} \xi_j^{(q)}, \sum_{q=1}^{i-1} \sum_{j \geq 0} \xi_j^{(q),m} \right) \stackrel{d}{=} \left(\Xi_n(u_{i-1}), \breve{\Xi}_n^m(u_{i-1}) \right)$$

where the laws of processes $\Xi_n, \breve{\Xi}_n^m$ are stated in (32) and (34), respectively. Applying Lemma 4, we yield

$$\mathbf{P}\left(\left| \sum_{q=1}^{i-1} \sum_{j \geq 0} \xi_j^{(q)} - \sum_{q=1}^{i-1} \sum_{j \geq 0} \xi_j^{(q),m} \right| > \frac{x}{2} \right) \leq \mathbf{P}\left(\sup_{t \in [0,n]} |\Xi_n(t) - \breve{\Xi}_n^m(t)| > \frac{x}{2} \right) \leq \frac{4C \kappa^{m(2-\beta_+)}}{x^2 n^{r(2-\beta_+)-1}},$$

where $C < \infty$ is the constant characterized in Lemma 4 that only depends on β_+ and the law of the Lévy

process X. To conclude the proof of claim (96), we pick $\widetilde{C}_1 = 4C$.

Proof of Claim (97). It follows directly from Lemma 5 that

$$\mathbf{P}\left(\left| \sum_{j=1}^{m+t(n)} (\xi_j^{(i)})^+ - \sum_{j=1}^{m+t(n)} (\xi_j^{(i),m})^+ \right| > \frac{x}{2} \right) \le \frac{4C\kappa^{m(2-\beta_+)}}{x^2 n^{r(2-\beta_+)-1}},$$

where $C < \infty$ is the constant characterized in Lemma 5 that only depends on β_+ and the law of the Lévy

process X. To conclude the proof of claim (97), we pick $\widetilde{C}_2 = 4C$.

Proof of Claim (98). Using Lemma 6,

$$\mathbf{P}\left(\sum_{j \ge m+t(n)+1} (\xi_j^{(i)})^+ > x \right) \le \frac{2C_X}{x} \cdot \sqrt{\frac{1}{n^{d-1} \cdot 2^m}}$$

where $C_X < \infty$ is the constant characterized in Lemma 6 that only depends on the law of the Lévy process

X. By setting $\widetilde{C}_3 = 2C_X$, we conclude the proof of claim (98). □

Proof of Proposition 5 In this proof, we fix some $k \in \mathbb{N}$. Recall the representation $\zeta_k(t) =$

$\sum_{i=1}^k z_i \mathbf{I}_{[u_i,n]}(t)$ in (78) where $0 < u_1 < u_2 < \ldots < u_k < n$ are the order statistics of k iid samples of Unif$(0,n)$.

Recall the definition of $\widetilde{W}_n^{(i),m}(\zeta_k)$ in (82). See also (26)–(35) for the definitions $\xi_j^{(i)}$'s and $\xi_j^{(i),m}$'s.

We start with the following decomposition of events:

$$\mathbf{P}\left(\exists i \in [k+1] \ s.t. \ \widetilde{W}_n^{(i),m}(\zeta_k) \in \left[y - \frac{\delta^m}{\sqrt{n}}, y + \frac{\delta^m}{\sqrt{n}} \right] \right)$$

$$\le \mathbf{P}(u_1 < n\delta^{m\alpha_1}) + \mathbf{P}\left(\exists i \in [k+1] \ s.t. \ \widetilde{W}_n^{(i),m}(\zeta_k) \in \left[y - \frac{\delta^m}{\sqrt{n}}, y + \frac{\delta^m}{\sqrt{n}} \right], \ u_1 \ge n\delta^{m\alpha_1} \right).$$

First, $\mathbf{P}(u_1 < n\delta^{m\alpha_1}) \le k \cdot \mathbf{P}(\mathrm{Unif}(0, n) < n\delta^{m\alpha_1}) = k \cdot \delta^{m\alpha_1} < k \cdot \rho_0^m$. The last inequality follows from our choice of ρ_1 in (70) and $\rho_0 \in (\rho_1, 1)$. Furthermore, for each $i \in [k+1]$

$$
\mathbf{P}\left(\widetilde{W}_n^{(i),m}(\zeta_k) \in [y - \frac{\delta^m}{\sqrt{n}}, y + \frac{\delta^m}{\sqrt{n}}], \ u_1 \ge n\delta^{m\alpha_1} \right)
$$

$$
= \mathbf{P}\left(\sum_{q=1}^{i-1} \sum_{j \ge 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q \in [y - \delta^{m\alpha_2}, y + \delta^{m\alpha_2}], \ \widetilde{W}_n^{(i),m}(\zeta_k) \in [y - \frac{\delta^m}{\sqrt{n}}, y + \frac{\delta^m}{\sqrt{n}}], \ u_1 \ge n\delta^{m\alpha_1} \right)
$$

$$
+ \mathbf{P}\left(\sum_{q=1}^{i-1} \sum_{j \ge 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q \notin [y - \delta^{m\alpha_2}, y + \delta^{m\alpha_2}], \ \widetilde{W}_n^{(i),m}(\zeta_k) \in [y - \frac{\delta^m}{\sqrt{n}}, y + \frac{\delta^m}{\sqrt{n}}], \ u_1 \ge n\delta^{m\alpha_1} \right)
$$

$$
\le \mathbf{P}\left(\sum_{q=1}^{i-1} \sum_{j \ge 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q \in [y - \delta^{m\alpha_2}, y + \delta^{m\alpha_2}], \ u_1 \ge n\delta^{m\alpha_1} \right)
$$

$$
+ \mathbf{P}\left(\sum_{q=1}^{i-1} \sum_{j \ge 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q \notin [y - \delta^{m\alpha_2}, y + \delta^{m\alpha_2}], \ \widetilde{W}_n^{(i),m} \in [y - \frac{\delta^m}{\sqrt{n}}, y + \frac{\delta^m}{\sqrt{n}}] \right)
$$

$$
\le \mathbf{P}\left(\sum_{q=1}^{i-1} \sum_{j \ge 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q \in [y - \delta^{m\alpha_2}, y + \delta^{m\alpha_2}], \ u_1 \ge n\delta^{m\alpha_1} \right)
$$

$$
+ \int_{\mathbb{R} \setminus [y - \delta^{m\alpha_2}, y + \delta^{m\alpha_2}]} \mathbf{P}\left(\sum_{j=1}^{m+t(n)} (\xi_j^{(i)})^+ \in [y - x - \frac{\delta^m}{\sqrt{n}}, y - x + \frac{\delta^m}{\sqrt{n}}] \right) \mathbf{P}\left(\sum_{q=1}^{i-1} \sum_{j \ge 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q \in dx \right)
$$

$$
= \mathbf{P}\left(\sum_{q=1}^{i-1} \sum_{j \ge 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q \in [y - \delta^{m\alpha_2}, y + \delta^{m\alpha_2}], \ u_1 \ge n\delta^{m\alpha_1} \right)
$$

$$
+ \int_{(-\infty, y - \delta^{m\alpha_2}]} \mathbf{P}\left(\sum_{j=1}^{m+t(n)} (\xi_j^{(i)})^+ \in [y - x - \frac{\delta^m}{\sqrt{n}}, y - x + \frac{\delta^m}{\sqrt{n}}] \right) \mathbf{P}\left(\sum_{q=1}^{i-1} \sum_{j \ge 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q \in dx \right).
$$

The last equality follows from the simple fact that $\sum_{j \ge 1} (\xi_j^{(i)})^+ \ge 0$. Furthermore, we claim the existence of constants $\widetilde{C}_1$ and $\widetilde{C}_2$, the values of which do not vary with parameters n, m, k, y, i, such that for all $n \ge 1$ and $m \ge \bar{m}$,

$$
\mathbf{P}\left(\sum_{q=1}^{i-1} \sum_{j \ge 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q \in [y - \delta^{m\alpha_2}, y + \delta^{m\alpha_2}], \ u_1 \ge n\delta^{m\alpha_1} \right) \le \widetilde{C}_1 \rho_0^m \qquad \forall y > \delta^{m\alpha_2}, \tag{99}
$$

$$
\mathbf{P}\left(\sum_{j=1}^{m+t(n)} (\xi_j^{(i)})^+ \in [w, w + \frac{2\delta^m}{\sqrt{n}}] \right) \le \widetilde{C}_2 \rho_0^m \qquad \forall w \ge \delta^{m\alpha_2} - \frac{\delta^m}{\sqrt{n}}. \tag{100}
$$

Then, we conclude the proof by setting $C_2 = 1 + \widetilde{C}_1 + \widetilde{C}_2$. Now, we prove claims (99) and (100)

Proof of Claim (99). If $i \leq 1$, the claim is trivial due to $y > \delta^{m\alpha}$ and hence $0 \notin [y - \delta^{m\alpha_2}, y + \delta^{m\alpha_2}]$. Now,

we consider the case where $i \geq 2$. Due to the independence between z_i and $\xi_j^{(i)}$,

$$
\mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} + \sum_{q=1}^{i-1}z_q \in [y - \delta^{m\alpha_2}, y + \delta^{m\alpha_2}],\ u_1 \geq n\delta^{m\alpha_1}\Big)
$$

$$
= \int_{\mathbb{R}}\mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} \in [y - x - \delta^{m\alpha_2}, y - x + \delta^{m\alpha_2}],\ u_1 \geq n\delta^{m\alpha_1}\Big)\mathbf{P}\Big(\sum_{q=1}^{i-1}z_q \in dx\Big)
$$

$$
\leq \int_{\mathbb{R}}\mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} \in [y - x - \delta^{m\alpha_2}, y - x + \delta^{m\alpha_2}]\ \Big|\ u_1 \geq n\delta^{m\alpha_1}\Big)\mathbf{P}\Big(\sum_{q=1}^{i-1}z_q \in dx\Big)
$$

$$
= \int_{\mathbb{R}}\mathbf{P}\Big(X^{<n\gamma}(u_{i-1}) \in [y - x - \delta^{m\alpha_2}, y - x + \delta^{m\alpha_2}]\ \Big|\ u_1 \geq n\delta^{m\alpha_1}\Big)\mathbf{P}\Big(\sum_{q=1}^{i-1}z_q \in dx\Big)
$$

where $(u_i)_{i=1}^k$ are independent of the Lévy process $X^{<n\gamma}$. In particular, recall that $0 = u_0 < u_1 < u_2 < \ldots <$

$u_k < n$ are order statistics. Therefore, on event $\{u_1 \geq n\delta^{m\alpha_1}\}$ we must have $u_{i-1} \geq u_1 \geq n\delta^{m\alpha_1}$. It then follows

directly from Assumption 2 that

$$
\mathbf{P}\Big(X^{<n\gamma}(u_{i-1}) \in [y - x - \delta^{m\alpha_2}, y - x + \delta^{m\alpha_2}]\ \Big|\ u_1 \geq n\delta^{m\alpha_1}\Big)
$$

$$
\leq \frac{C}{(n^\lambda \delta^{m\alpha_1 \lambda}) \wedge 1} \cdot 2\delta^{m\alpha_2} \leq 2C \cdot \Big(\frac{\delta^{\alpha_2}}{\delta^{\lambda\alpha_1}}\Big)^m \leq 2C \cdot \rho_0^m \qquad \text{due to (73) and } \rho_0 \in (\rho_1, 1),
$$

where C and λ are the constants specified in Assumption 2. To conclude, it suffices to set $\widetilde{C}_1 = 2C$.

64

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Proof of Claim (100). Applying Lemma 7 with $y_0 = \delta^{m\alpha_2} - \frac{\delta^m}{\sqrt{n}}$ and $c = \frac{2\delta^m}{\sqrt{n}}$, we get (for all $n \geq 1, m \geq \bar{m}, y \geq y_0$)

$$\mathbf{P}\left(\sum_{j=1}^{m+t(n)} (\xi_j^{(i)})^+ \in [y, y + \frac{2\delta^m}{\sqrt{n}}] \right)$$

$$\leq C \frac{(m + (\lceil \log_2(n^d) \rceil)) n^{\alpha_4 \lambda}}{\delta^{m\alpha_3 \lambda}} \cdot \frac{2\delta^m}{\sqrt{n}} + 4C_X (m^2 + (\lceil \log_2(n^d) \rceil)^2) \frac{\delta^{m\alpha_3/2}}{(\delta^{m\alpha_2} - \frac{\delta^m}{\sqrt{n}}) \cdot n^{\alpha_4/2}}$$

$$\leq C \frac{(m + (\lceil \log_2(n^d) \rceil)) n^{\alpha_4 \lambda}}{\delta^{m\alpha_3 \lambda}} \cdot \frac{2\delta^m}{\sqrt{n}} + 8C_X (m^2 + (\lceil \log_2(n^d) \rceil)^2) \frac{\delta^{m\alpha_3/2}}{\delta^{m\alpha_2} \cdot n^{\alpha_4/2}} \qquad \text{due to (68)}$$

$$= \underbrace{2C \cdot \frac{m}{n^{\frac{1}{2} - \lambda\alpha_4}} \cdot \left(\frac{\delta}{\delta^{\lambda\alpha_3}} \right)^m}_{\triangleq p_{n,m,1}} + \underbrace{2C \cdot \frac{\lceil \log_2(n^d) \rceil}{n^{\frac{1}{2} - \lambda\alpha_4}} \cdot \left(\frac{\delta}{\delta^{\lambda\alpha_3}} \right)^m}_{\triangleq p_{n,m,2}}$$

$$+ \underbrace{8C_X \cdot \frac{m^2}{n^{\alpha_4/2}} \cdot \left(\frac{\delta^{\alpha_3/2}}{\delta^{\alpha_2}} \right)^m}_{\triangleq p_{n,m,3}} + \underbrace{8C_X \cdot \frac{(\lceil \log_2(n^d) \rceil)^2}{n^{\alpha_4/2}} \cdot \left(\frac{\delta^{\alpha_3/2}}{\delta^{\alpha_2}} \right)^m}_{\triangleq p_{n,m,4}}.$$

Here, $C_X < \infty$ is the constant in Lemma 3 that only depends on the law of Lévy process X, and $C \in (0, \infty), \lambda > 0$ are the constants in Assumption 2. First, for any $n \geq 1$ and $m \geq \bar{m}$,

$$p_{n,m,1} \leq 2C \cdot m \cdot \left(\frac{\delta}{\delta^{\lambda\alpha_3}} \right)^m \qquad \text{due to } \frac{1}{2} > \lambda\alpha_4; \text{ see (64)}$$

$$\leq 2C \cdot m\rho_1^m \qquad \text{due to (74)}$$

$$\leq 2C \cdot \rho_0^m \qquad \text{due to (77)}.$$

For term $p_{n,m,2}$, note that $\frac{\lceil \log_2(n^d) \rceil}{n^{\frac{1}{2} - \lambda\alpha_4}} \to 0$ as $n \to \infty$ due to $\frac{1}{2} > \lambda\alpha_4$. This allows us to fix some $C_{d,1} < \infty$ such that $\sup_{n=1,2,\cdots} \frac{\lceil \log_2(n^d) \rceil}{n^{\frac{1}{2} - \lambda\alpha_4}} \leq C_{d,1}$. As a result, for any $n \geq 1, m \geq 0$,

$$p_{n,m,2} \leq 2CC_{d,1} \cdot \left(\frac{\delta}{\delta^{\lambda\alpha_3}} \right)^m \leq 2CC_{d,1} \cdot \rho_0^m \qquad \text{due to (74) and } \rho_0 \in (\rho_1, 1).$$

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

65

Similarly, for all $n \geq 1$ and $m \geq \bar{m}$,

$$p_{n,m,3} \leq 8 C_X \cdot m^2 \cdot \left(\frac{\delta^{\alpha_3/2}}{\delta^{\alpha_2}} \right)^m \leq 8 C_X \cdot m^2 \rho_1^m \qquad \text{due to (75)}$$

$$\leq 8 C_X \cdot \rho_0^m \qquad \text{due to (77)}.$$

Besides, due to $\frac{(\lceil \log_2(n^d) \rceil)^2}{n^{\alpha_4/2}} \to 0$ as $n \to \infty$, we can find $C_{d,2} < \infty$ such that $\sup_{n=1,2,\cdots,} \frac{(\lceil \log_2(n^d) \rceil)^2}{n^{\alpha_4/2}} \leq C_{d,2}$.
This leads to (for all $n \geq 1, m \geq 0$)

$$p_{n,m,4} \leq 8 C_X C_{d,2} \cdot \left(\frac{\delta^{\alpha_3/2}}{\delta^{\alpha_2}} \right)^m \leq 8 C_X C_{d,2} \cdot \rho_0^m \qquad \text{due to (74) and } \rho_0 \in (\rho_1, 1).$$

To conclude the proof, we can simply set $\widetilde{C}_2 = 2C + 2C C_{d,1} + 8 C_X + 8 C_X C_{d,2}$. $\quad \square$

References

Asmussen S, Glynn P, Pitman J (1995) Discretization Error in Simulation of One-Dimensional Reflecting Brownian Motion. *The Annals of Applied Probability* 5(4):875 − 896, URL http://dx.doi.org/10.1214/aoap/1177004597.

Asmussen S, Kroese DP (2006) Improved algorithms for rare event simulation with heavy tails. *Advances in Applied Probability* 38(2):545–558, URL http://dx.doi.org/10.1239/aap/1151337084.

Asmussen S, Rosiński J (2001) Approximations of small jumps of lévy processes with a view towards simulation. *Journal of Applied Probability* 38(2):482–493.

Bassamboo A, Juneja S, Zeevi A (2007) On the inefficiency of state-independent importance sampling in the presence of heavy tails. *Operations Research Letters* 35(2):251–260, ISSN 0167-6377, URL http://dx.doi.org/https://doi.org/10.1016/j.orl.2006.02.002.

Bianchi ML, Rachev ST, Kim YS, Fabozzi FJ (2011) Tempered infinitely divisible distributions and processes. *Theory of Probability & Its Applications* 55(1):2–26.

Blanchet J, Glynn P (2008) Efficient rare-event simulation for the maximum of heavy-tailed random walks. *The Annals of Applied Probability* 18(4):1351 − 1378, URL http://dx.doi.org/10.1214/07-AAP485.

Blanchet J, Glynn P, Liu J (2008) Efficient rare event simulation for heavy-tailed multiserver queues. Technical report, Department of Statistics, Columbia University.

66

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Blanchet J, Hult H, Leder K (2013) Rare-event simulation for stochastic recurrence equations with heavy-tailed innovations. *ACM Trans. Model. Comput. Simul.* 23(4), ISSN 1049-3301, URL http://dx.doi.org/10.1145/2517451.

Blanchet JH, Liu J (2008) State-dependent importance sampling for regularly varying random walks. *Advances in Applied Probability* 40(4):1104–1128, URL http://dx.doi.org/10.1239/aap/1231340166.

Borak S, Misiorek A, Weron R (2011) *Models for heavy-tailed asset returns*, 21–55 (Berlin, Heidelberg: Springer Berlin Heidelberg), ISBN 978-3-642-18062-0, URL http://dx.doi.org/10.1007/978-3-642-18062-0_1.

Boxma O, Cahen E, Koops D, Mandjes M (2019) Linear networks: rare-event simulation and markov modulation. *Methodology and Computing in Applied Probability* URL http://dx.doi.org/https://doi.org/10.1007/s11009-018-9644-1.

Boyarchenko S, Levendorskii S (2023a) Efficient evaluation of joint pdf of a lévy process, its extremum, and hitting time of the extremum. URL https://arxiv.org/abs/2312.05222.

Boyarchenko S, Levendorskii S (2023b) Simulation of a lévy process, its extremum, and hitting time of the extremum via characteristic functions. URL https://arxiv.org/abs/2312.03929.

Bucklew JA, Ney P, Sadowsky JS (1990) Monte carlo simulation and large deviations theory for uniformly recurrent markov chains. *Journal of Applied Probability* 27(1):44–59, URL http://dx.doi.org/10.2307/3214594.

Carr P, Geman H, Madan D, Yor M (2002) The fine structure of asset returns: An empirical investigation. *The Journal of Business* 75(2):305–332, ISSN 00219398, 15375374, URL http://www.jstor.org/stable/10.1086/338705.

Carr P, Geman H, Madan DB, Yor M (2003) Stochastic volatility for lévy processes. *Mathematical Finance* 13(3):345–382, URL http://dx.doi.org/https://doi.org/10.1111/1467-9965.00020.

Cázares JIG, Kohatsu-Higa A, Mijatović A (2023) Joint density of the stable process and its supremum: Regularity and upper bounds. *Bernoulli* 29(4):3443 – 3469, URL http://dx.doi.org/10.3150/23-BEJ1590.

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

67

Cázares JIG, Mijatović A, Bravo GU (2020) ε-strong simulation of the convex minorants of stable processes and meanders. *Electronic Journal of Probability* 25(none):1 – 33, URL `http://dx.doi.org/10.1214/20-EJP503`.

Chaumont L (2013) On the law of the supremum of Lévy processes. *The Annals of Probability* 41(3A):1191 – 1217, URL `http://dx.doi.org/10.1214/11-AOP708`.

Chen B, Blanchet J, Rhee CH, Zwart B (2019) Efficient rare-event simulation for multiple jump events in regularly varying random walks and compound poisson processes. *Mathematics of Operations Research* 44(3):919–942, URL `http://dx.doi.org/10.1287/moor.2018.0950`.

Cohen JE, Davis RA, Samorodnitsky G (2022) Covid-19 cases and deaths in the united states follow taylor's law for heavy-tailed distributions with infinite variance. *Proceedings of the National Academy of Sciences* 119(38):e2209234119, URL `http://dx.doi.org/10.1073/pnas.2209234119`.

Coutin L, Pontier M, Ngom W (2018) Joint distribution of a lévy process and its running supremum. *Journal of Applied Probability* 55(2):488–512, URL `http://dx.doi.org/10.1017/jpr.2018.32`.

Cázares JIG, Lin F, Mijatović A (2023) Fast exact simulation of the first passage of a tempered stable subordinator across a non-increasing function.

Dereich S (2011) Multilevel Monte Carlo algorithms for Lévy-driven SDEs with Gaussian correction. *The Annals of Applied Probability* 21(1):283 – 311, URL `http://dx.doi.org/10.1214/10-AAP695`.

Dereich S, Heidenreich F (2011) A multilevel monte carlo algorithm for lévy-driven stochastic differential equations. *Stochastic Processes and their Applications* 121(7):1565–1587, ISSN 0304-4149, URL `http://dx.doi.org/https://doi.org/10.1016/j.spa.2011.03.015`.

Dia EHA, Lamberton D (2011) Connecting discrete and continuous lookback or hindsight options in exponential lévy models. *Advances in Applied Probability* 43(4):1136–1165, URL `http://dx.doi.org/10.1239/aap/1324045702`.

Dupuis P, Leder K, Wang H (2007a) Importance sampling for sums of random variables with regularly varying tails. *ACM Trans. Model. Comput. Simul.* 17(3):14–es, ISSN 1049-3301, URL `http://dx.doi.org/10.1145/1243991.1243995`.

68

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Dupuis P, Sezer AD, Wang H (2007b) Dynamic importance sampling for queueing networks. *The Annals of Applied Probability* 17(4):1306 – 1346, URL http://dx.doi.org/10.1214/105051607000000122.

Durrett R (2019) *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics (Cambridge University Press), ISBN 9781108473682, URL https://books.google.com/books?id=b22MDwAAQBAJ.

Embrechts P, Klüppelberg C, Mikosch T (2013) *Modelling extremal events: for insurance and finance*, volume 33 (Springer Science & Business Media).

Ferreiro-Castilla A, Kyprianou A, Scheichl R, Suryanarayana G (2014) Multilevel monte carlo simulation for lévy processes based on the wiener–hopf factorisation. *Stochastic Processes and their Applications* 124(2):985–1010, ISSN 0304-4149, URL http://dx.doi.org/https://doi.org/10.1016/j.spa.2013.09.015.

Giles MB (2008) Multilevel monte carlo path simulation. *Operations Research* 56(3):607–617, URL http://dx.doi.org/10.1287/opre.1070.0496.

Giles MB, Xia Y (2017) Multilevel monte carlo for exponential lévy models. *Finance and Stochastics* 21(4):995–1026.

González Cázares J, Mijatović A (2022) Simulation of the drawdown and its duration in lévy models via stick-breaking gaussian approximation. *Finance and Stochastics* 26(4):671–732.

González Cázares JI, Mijatović A, Uribe Bravo G (2022) Geometrically convergent simulation of the extrema of lévy processes. *Mathematics of Operations Research* 47(2):1141–1168, URL http://dx.doi.org/10.1287/moor.2021.1163.

González Cázares JI, Mijatović A, Bravo GU (2019) Exact simulation of the extrema of stable processes. *Advances in Applied Probability* 51(4):967–993, URL http://dx.doi.org/10.1017/apr.2019.39.

Gudmundsson T, Hult H (2014) Markov chain monte carlo for computing rare-event probabilities for a heavy-tailed random walk. *Journal of Applied Probability* 51(2):359–376, URL http://dx.doi.org/10.1239/jap/1402578630.

Gurbuzbalaban M, Simsekli U, Zhu L (2021) The heavy-tail phenomenon in sgd. Meila M, Zhang T, eds., *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of*

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

69

Machine Learning Research, 3964–3975 (PMLR), URL `https://proceedings.mlr.press/v139/gurbuzbalaban21a.html`.

Heinrich S (2001) Multilevel monte carlo methods. Margenov S, Waśniewski J, Yalamov P, eds., *Large-Scale Scientific Computing*, 58–67 (Berlin, Heidelberg: Springer Berlin Heidelberg), ISBN 978-3-540-45346-8.

Hesterberg T (1995) Weighted average importance sampling and defensive mixture distributions. *Technometrics* 37(2):185–194, ISSN 00401706, URL `http://www.jstor.org/stable/1269620`.

Hodgkinson L, Mahoney M (2021) Multiplicative noise and heavy tails in stochastic optimization. *International Conference on Machine Learning*, 4262–4274 (PMLR).

Hult H, Juneja S, Murthy K (2016) Exact and efficient simulation of tail probabilities of heavy-tailed infinite series .

Kuznetsov A, Kyprianou AE, Pardo JC, van Schaik K (2011) A Wiener–Hopf Monte Carlo simulation technique for Lévy processes. *The Annals of Applied Probability* 21(6):2171 – 2190, URL `http://dx.doi.org/10.1214/10-AAP746`.

Kwaśnicki M, Ma lecki J, Ryznar M (2013) Suprema of Lévy processes. *The Annals of Probability* 41(3B):2047 – 2065, URL `http://dx.doi.org/10.1214/11-AOP719`.

Li Y (2018) Queuing theory with heavy tails and network traffic modeling, URL `https://hal.science/hal-01891760`, working paper or preprint.

Mariucci E, Reiß M (2018) Wasserstein and total variation distance between marginals of Lévy processes. *Electronic Journal of Statistics* 12(2):2482 – 2514, URL `http://dx.doi.org/10.1214/18-EJS1456`.

Michna Z (2012) Formula for the supremum distribution of a spectrally positive lévy process.

Michna Z (2013) Explicit formula for the supremum distribution of a spectrally negative stable process. *Electronic Communications in Probability* 18(none):1 – 6, URL `http://dx.doi.org/10.1214/ECP.v18-2236`.

Michna Z, Palmowski Z, Pistorius M (2014) The distribution of the supremum for spectrally asymmetric lévy processes.

Mijatović A, Tankov P (2016) A new look at short-term implied volatility in asset price models with jumps. *Mathematical Finance* 26(1):149–183, URL `http://dx.doi.org/https://doi.org/10.1111/mafi.12055`.

70

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Murthy KRA, Juneja S, Blanchet J (2014) State-independent importance sampling for random walks with regularly varying increments. *Stochastic Systems* 4(2):321–374, URL http://dx.doi.org/10.1287/13-SSY114.

Pitman J, Bravo GU (2012) The convex minorant of a Lévy process. *The Annals of Probability* 40(4):1636 – 1674, URL http://dx.doi.org/10.1214/11-AOP658.

Resnick SI (2007) *Heavy-tail phenomena: probabilistic and statistical modeling* (Springer Science & Business Media).

Rhee CH, Blanchet J, Zwart B, et al. (2019) Sample path large deviations for lévy processes and random walks with regularly varying increments. *The Annals of Probability* 47(6):3551–3605.

Rhee CH, Glynn PW (2015) Unbiased estimation with square root convergence for sde models. *Operations Research* 63(5):1026–1043, URL http://dx.doi.org/10.1287/opre.2015.1404.

Rosiński J (2007) Tempering stable processes. *Stochastic Processes and their Applications* 117(6):677–707, ISSN 0304-4149, URL http://dx.doi.org/https://doi.org/10.1016/j.spa.2006.10.003.

Sabino P (2022) Pricing energy derivatives in markets driven by tempered stable and cgmy processes of ornstein–uhlenbeck type. *Risks* 10(8), ISSN 2227-9091, URL http://dx.doi.org/10.3390/risks10080148.

Sato Ki, Ken-Iti S, Katok A (1999) *Lévy processes and infinitely divisible distributions* (Cambridge university press).

Torrisi G (2004) Simulating the ruin probability of risk processes with delay in claim settlement. *Stochastic Processes and their Applications* 112(2):225–244, ISSN 0304-4149, URL http://dx.doi.org/https://doi.org/10.1016/j.spa.2004.02.007.

Wang X, Rhee CH (2021) Rare-event simulation for multiple jump events in heavy-tailed lévy processes with infinite activities. *Proceedings of the Winter Simulation Conference*, 409–420, WSC '20 (IEEE Press), ISBN 9781728194998.

Wang X, Rhee CH (2023) Large deviations and metastability analysis for heavy-tailed dynamical systems.

Wang X, Rhee CH (2024) Importance sampling strategy for heavy-tailed systems with catastrophe principle. *Proceedings of the Winter Simulation Conference*, 76–90, WSC '23 (IEEE Press), ISBN 9798350369663.

Online Companion (Appendix)

This online companion is organized as follows. In Section A, we provide the proofs for technical results in Section 4.

In Section B, we extend the algorithmic framework of the main paper to address rare-event simulation in the context

of barrier option pricing.

Appendix A: Proof of Propositions 2 and 3   The proof of Proposition 2 is based on the inversion formula of

the characteristic functions (see, e.g., Theorem 3.3.14 of Durrett (2019)). Specifically, we compare the characteristic

function of $Y(t)$ with an α-stable process to draw connections between their distributions.

Proof of Proposition 2   The Lévy-Khintchine formula (see e.g. Theorem 8.1 of Sato et al. (1999)) leads to the

following expression for the characteristic function of $\varphi_t(z) = \mathbf{E}\exp(izY(t))$:

$$\varphi_t(z) = \exp\left(t \int_{(0,z_0)} \underbrace{\left[\exp(izx) - 1 - izx\mathbf{I}_{(0,1]}(x) \right]}_{\triangleq \phi(z,x)} \mu(dx) \right) \qquad \forall z \in \mathbb{R},\ t > 0.$$

Note that

$$\phi(z,x) = \cos(zx) - 1 + i\big(\sin(zx) - zx\mathbf{I}_{(0,1]}(x) \big).$$

Then from $|e^{x+iy}| = e^x$ for all $x, y \in \mathbb{R}$,

$$|\varphi_t(z)| = \exp\left(-t \int_{(0,z_0)} \big(1 - \cos(zx)\big)\mu(dx) \right) \qquad \forall z \in \mathbb{R},\ t > 0. \tag{101}$$

Furthermore, we claim the existence of some $\widetilde{M}, \widetilde{C} \in (0, \infty)$ such that

$$\int_{(0,z_0)} \big(1 - \cos(zx)\big)\mu(dx) \geq \widetilde{C}|z|^\alpha \qquad \forall z \in \mathbb{R} \text{ with } |z| \geq \widetilde{M}. \tag{102}$$

Plugging (102) into (101), we yield that for all $|z| \geq \widetilde{M}$ and $t > 0$, $|\varphi_t(z)| \leq \exp(-t\widetilde{C}|z|^\alpha)$. It then follows directly

from the inversion formula (see Theorem 3.3.14 of Durrett (2019)) that, for all $t > 0$, $Y(t)$ admits a continuous density

2

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

function $f_{Y(t)}$ with a uniform bound

$$
\begin{aligned}
\left\| f_{Y(t)} \right\|_\infty &\le \frac{1}{2\pi} \int |\varphi_t(z)| dz \\
&\le \frac{1}{2\pi} \left(2\widetilde{M} + \int_{|z| \ge \widetilde{M}} \exp\left(-t\widetilde{C}|z|^\alpha \right) dz \right) \\
&\le \frac{1}{2\pi} \left(2\widetilde{M} + \frac{1}{t^{1/\alpha}} \int_{\mathbb{R}} \exp(-\widetilde{C}|x|^\alpha) dx \right) \qquad \text{by letting } x = zt^{1/\alpha} \\
&= \frac{\widetilde{M}}{\pi} + \frac{C_1}{t^{1/\alpha}} \qquad \text{where } C_1 = \frac{1}{2\pi} \int_{\mathbb{R}} \exp(-\widetilde{C}|x|^\alpha) dx < \infty.
\end{aligned}
$$

To conclude the proof, pick $C = \frac{\widetilde{M}}{\pi} + C_1$. Now, it only remains to prove claim (102).

Proof of Claim (102). We start by fixing some constants.

$$
C_0 = \int_0^\infty (1 - \cos y) \frac{dy}{y^{1+\alpha}}. \tag{103}
$$

For $y \in (0, 1]$, note that $1 - \cos y \le y^2/2$, and hence $\frac{|1-\cos y|}{y^{1+\alpha}} \le \frac{1}{2y^{\alpha-1}}$. For $y \in (1, \infty)$, note that $1 - \cos y \in [0, 1]$ and hence $\frac{|1-\cos y|}{y^{1+\alpha}} \le 1/y^{\alpha+1}$. Due to $\alpha \in (0, 2)$, we have $C_0 = \int_0^\infty (1 - \cos y) \frac{dy}{y^{1+\alpha}} \in (0, \infty)$. Next, choose positive real numbers θ, δ such that

$$
\frac{\theta^{2-\alpha}}{2(2-\alpha)} \le \frac{C_0}{4}, \tag{104}
$$

$$
\frac{\delta}{\alpha\theta^\alpha} \le \frac{C_0}{4}. \tag{105}
$$

For any $M > 0$ and $z \ne 0$, observe that (by setting $y = |z|x$ in the last step)

$$
\frac{\int_{x \ge \frac{M}{|z|}} \left(1 - \cos(zx)\right) \frac{dx}{x^{1+\alpha}}}{|z|^\alpha} = \frac{\int_{x \ge \frac{M}{|z|}} \left(1 - \cos(|z|x)\right) \frac{dx}{x^{1+\alpha}}}{|z|^\alpha} = \int_M^\infty \left(1 - \cos y\right) \frac{dy}{y^{1+\alpha}}.
$$

Therefore, by fixing some $M > \theta$ large enough, we have

$$
\frac{1}{|z|^\alpha} \int_{x \ge M/|z|} \left(1 - \cos(zx)\right) \frac{dx}{x^{1+\alpha}} \le \frac{C_0}{4} \qquad \forall z \ne 0. \tag{106}
$$

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

3

To proceed, we compare $\int_{(0,z_0)} \left(1 - \cos(zx)\right)\mu(dx)$ with $\int_0^{M/z} \left(1 - \cos(zx)\right)\frac{dx}{x^{1+\alpha}}$. Recall that z_0 is the constant prescribed in the statement of Proposition 2. For any $z \in \mathbb{R}$ such that $|z| > M/z_0$,

$$\frac{1}{|z|^\alpha}\left[\int_{(0,z_0)} \left(1 - \cos(zx)\right)\mu(dx) - \int_0^\infty \left(1 - \cos(zx)\right)\frac{dx}{x^{1+\alpha}}\right]$$

$$\geq \frac{1}{|z|^\alpha}\left[\int_{(\theta/|z|,M/|z|)} \left(1 - \cos(zx)\right)\mu(dx) - \int_0^\infty \left(1 - \cos(zx)\right)\frac{dx}{x^{1+\alpha}}\right]$$

due to our choice of $M > \theta$ and $|z| > M/z_0$

$$\geq -\underbrace{\frac{1}{|z|^\alpha}\int_0^{\theta/|z|} \left(1 - \cos(zx)\right)\frac{dx}{x^{1+\alpha}}}_{\triangleq I_1(z)} - \underbrace{\frac{1}{|z|^\alpha}\int_{M/|z|}^\infty \left(1 - \cos(zx)\right)\frac{dx}{x^{1+\alpha}}}_{\triangleq I_2(z)}$$

$$+ \underbrace{\frac{1}{|z|^\alpha}\left[\int_{[\theta/|z|,M/|z|)} \left(1 - \cos(zx)\right)\mu(dx) - \int_{[\theta/|z|,M/|z|)} \left(1 - \cos(zx)\right)\frac{dx}{x^{1+\alpha}}\right]}_{\triangleq I_3(z)}. \tag{107}$$

We bound the terms $I_1(z)$, $I_2(z)$, and $I_3(z)$ separately. First, for any $z \neq 0$,

$$I_1(z) \leq \frac{1}{|z|^\alpha}\int_0^{\theta/|z|} \frac{z^2x^2}{2}\frac{dx}{x^{1+\alpha}} \qquad \text{due to } 1 - \cos w \leq \frac{w^2}{2} \; \forall w \in \mathbb{R}$$

$$= \frac{1}{2}\int_0^\theta y^{1-\alpha}dy \qquad \text{by setting } y = |z|x$$

$$= \frac{1}{2}\cdot\frac{\theta^{2-\alpha}}{2-\alpha} \leq \frac{C_0}{4} \quad \text{due to (104).} \tag{108}$$

For $I_2(z)$, it follows immediately from (106) that

$$I_2(z) \leq \frac{C_0}{4} \qquad \forall z \neq 0. \tag{109}$$

Next, in order to bound $I_3(z)$, we consider the function $h(z) \triangleq 1 - \cos z$. Since $h(z)$ is uniformly continuous on $[\theta, M]$, we can find some $N \in \mathbb{N}$, $t_0 > 1$, and a sequence of real numbers $\theta = x_0 > x_1 > \cdots > x_N = M$ such that

$$\frac{x_{j-1}}{x_j} = t_0 \qquad \forall j = 1, 2, \cdots, N, \tag{110}$$

$$|h(x) - h(y)| < \delta \qquad \forall j = 1, 2, \cdots, N, \; x, y \in [x_j, x_{j-1}]. \tag{111}$$

4

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

In other words, we use a geometric sequence $\{x_0, x_1, \cdots, x_N\}$ to partition $[\theta, M]$ into N intervals. On any of these intervals, the fluctuations of $h(z) = 1 - \cos z$ is bounded by the constant δ fixed in (105). Now fix some $\Delta > 0$ such that (recall that $\epsilon > 0$ is prescribed in the statement of this proposition)

$$(1 - \Delta)t_0^{\alpha+\epsilon} > 1. \tag{112}$$

Since $\mu[x, \infty)$ is regularly varying as $x \to 0$ with index $-(\alpha + 2\epsilon)$, for $g(y) = \mu[1/y, \infty)$ we have $g \in \mathcal{RV}_{\alpha+2\epsilon}(y)$ as $y \to \infty$. By Potter's bound (see Proposition 2.6 in Resnick (2007)), there exists $\bar{y}_1 > 0$ such that

$$\frac{g(ty)}{g(y)} \geq (1 - \Delta)t^{\alpha+\epsilon} \qquad \forall y \geq \bar{y}_1, \ t \geq 1. \tag{113}$$

Meanwhile, define

$$\widetilde{g}(y) = y^\alpha, \qquad \nu_\alpha(dx) = \mathbf{I}_{(0,\infty)}(x)\frac{dx}{x^{1+\alpha}}$$

and note that $\widetilde{g}(y) = \nu_\alpha(1/y, \infty)$. Due to $g \in \mathcal{RV}_{\alpha+2\epsilon}$, we can find some $\bar{y}_2 > 0$ such that

$$g(y) \geq \frac{t_0^\alpha - 1}{(1 - \Delta)t_0^{\alpha+\epsilon} - 1} \cdot \widetilde{g}(y) \quad \forall y \geq \bar{y}_2. \tag{114}$$

Let $\widetilde{M} = \max\{M/z_0, M\bar{y}_1, M\bar{y}_2\}$. For any $|z| \geq \widetilde{M}$, we have $|z| \geq M/z_0$ and $\frac{|z|}{x_j} \geq \frac{|z|}{M} \geq \bar{y}_1 \vee \bar{y}_2$ for any $j = 0, 1, \cdots, N$. As a result, for $z \in \mathbb{R}$ with $|z| \geq \widetilde{M}$ and any $j = 1, 2, \cdots, N$,

$$\begin{aligned}
\mu[x_j/|z|, x_{j-1}/|z|) &= g(|z|/x_j) - g(|z|/x_{j-1}) &&\text{by definition of } g(y) = \mu[1/y, \infty)\\
&= g(t_0|z|/x_{j-1}) - g(|z|/x_{j-1}) &&\text{due to } x_{j-1} = t_0 x_j; \text{ see (111)}\\
&\geq g(|z|/x_{j-1}) \cdot \left((1 - \Delta)t_0^{\alpha+\epsilon} - 1\right) &&\text{due to } \frac{|z|}{x_j} \geq \bar{y}_1 \vee \bar{y}_2 \text{ and (113)}\\
&\geq \widetilde{g}(|z|/x_{j-1}) \cdot (t_0^\alpha - 1) &&\text{due to (114)}.
\end{aligned}$$

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

5

On the other hand, $\nu_\alpha[x_j/|z|, x_{j-1}/|z|) = \widetilde{g}(|z|/x_j) - \widetilde{g}(|z|/x_{j-1}) = \widetilde{g}(|z|/x_{j-1}) \cdot (t_0^\alpha - 1)$. Therefore, given any $z \in \mathbb{R}$

such that $|z| \geq \widetilde{M}$, we have $\mu(E_j(z)) \geq \nu_\alpha(E_j(z))$ for all $j \in [N]$ where $E_j(z) = [x_j/|z|, x_{j-1}/|z|)$. This leads to

$$
\begin{aligned}
&I_3(z) \\
&= \frac{1}{|z|^\alpha} \sum_{j=1}^N \left[\int_{E_j(z)} \left(1 - \cos(zx)\right) \mu(dx) - \int_{E_j(z)} \left(1 - \cos(zx)\right) \frac{dx}{x^{1+\alpha}} \right] \\
&\geq \frac{1}{|z|^\alpha} \sum_{j=1}^N \left[\underline{m}_j \cdot \mu(E_j(z)) - \bar{m}_j \cdot \nu_\alpha(E_j(z)) \right] \\
&\qquad \text{with } \bar{m}_j = \max\{h(z): \ z \in [x_j, x_{j-1}]\}, \ \underline{m}_j = \min\{h(z): \ z \in [x_j, x_{j-1}]\} \\
&= \frac{1}{|z|^\alpha} \sum_{j=1}^N \left[\underline{m}_j \cdot \mu(E_j(z)) - \underline{m}_j \cdot \nu_\alpha(E_j(z)) \right] + \frac{1}{|z|^\alpha} \sum_{j=1}^N \left[\underline{m}_j \cdot \nu_\alpha(E_j(z)) - \bar{m}_j \cdot \nu_\alpha(E_j(z)) \right] \\
&\geq 0 + \frac{1}{|z|^\alpha} \sum_{j=1}^N \left[\underline{m}_j \cdot \nu_\alpha(E_j(z)) - \bar{m}_j \cdot \nu_\alpha(E_j(z)) \right] \qquad \text{due to } \mu(E_j(z)) \geq \nu_\alpha(E_j(z)) \\
&\geq -\frac{\delta}{|z|^\alpha} \sum_{j=1}^N \nu_\alpha(E_j(z)) = -\frac{\delta}{|z|^\alpha} \nu_c[\theta/|z|, M/|z|) \qquad \text{due to (111)} \\
&= -\frac{\delta}{|z|^\alpha} \int_{\theta/|z|}^{M/|z|} \frac{dx}{x^{1+\alpha}} \\
&\geq -\frac{\delta}{|z|^\alpha} \int_{\theta/|z|}^{\infty} \frac{dx}{x^{1+\alpha}} = -\frac{\delta}{\alpha \theta^\alpha} \\
&\geq -\frac{C_0}{4} \qquad \text{due to (105).}
\end{aligned}
\tag{115}
$$

Plugging (108), (109), and (115) back into (107), we have shown that for all $|z| \geq \widetilde{M}$,

$$
\begin{aligned}
&\frac{1}{|z|^\alpha} \int_{(0, z_0)} \left(1 - \cos(zx)\right) \mu(dx) \\
&\geq -\frac{3C_0}{4} + \frac{1}{|z|^\alpha} \int_0^\infty \left(1 - \cos(zx)\right) \frac{dx}{x^{1+\alpha}} \\
&= -\frac{3C_0}{4} + \int_0^\infty \left(1 - \cos y\right) \frac{dy}{y^{1+\alpha}} \qquad \text{by setting } y = |z|x \\
&= -\frac{3C_0}{4} + C_0 = \frac{C_0}{4} \qquad \text{by definition of } C_0 = \int_0^\infty (1 - \cos y) \frac{dy}{y^{1+\alpha}}.
\end{aligned}
$$

To conclude the proof of claim (102), we set $\widetilde{C} = C_0/4$. $\quad\square$

Again, the proof of Proposition 3 makes use of the inversion formula.

Proof of Proposition 3 Let us denote the characteristic functions of $Y'(t)$ and $Y(t)$ by φ_t and $\widetilde{\varphi}_t$, respectively.

Repeating the arguments using complex conjugates in (101), we obtain

$$|\widetilde{\varphi}_t(z)| = \exp\left(-t\int_{|x|<b^N}\big(1-\cos(zx)\big)\mu(dx)\right).$$

As for φ_t, using proposition 14.9 in Sato et al. (1999), we get

$$|\varphi_t(z)| = \exp\left(-t|z|^\alpha\eta(z)\right) \tag{116}$$

where $\eta(z)$ is a non-negative function continuous on $\mathbb{R}\backslash\{0\}$ satisfying $\eta(bz) = \eta(z)$ and

$$\eta(z) = \frac{\int_{\mathbb{R}}\big(1-\cos(zx)\big)\mu(dx)}{|z|^\alpha} \qquad \forall z \neq 0.$$

This implies $\eta(z) = \eta(-z)$ for all $z \neq 0$. Furthermore, we claim the existence of some $c > 0$ such that

$$\eta(z) \geq c \qquad \forall z \in [1, b]. \tag{117}$$

Then due to the self-similarity of μ (i.e., $\eta(bz) = \eta(z)$), we have $\eta(z) \geq c$ for all $z \neq 0$. In the meantime, note that

$$\frac{1}{|z|^\alpha}\int_{|x|\geq b^N}\big(1-\cos(zx)\big)\mu(dx) \leq \frac{\mu\{x:|x|\geq b^N\}}{|z|^\alpha}.$$

By picking $M > 0$ large enough, it holds for any $|z| \geq M$ that

$$\frac{1}{|z|^\alpha}\int_{|x|\geq b^N}\big(1-\cos(zx)\big)\mu(dx) \leq \frac{c}{2}. \tag{118}$$

Therefore, for any $|z| \geq M$,

$$\int_{|x|<b^N}\big(1-\cos(zx)\big)\mu(dx) = \int_{x\in\mathbb{R}}\big(1-\cos(zx)\big)\mu(dx) - \int_{|x|\geq b^N}\big(1-\cos(zx)\big)\mu(dx)$$

$$= \eta(z)\cdot|z|^\alpha - \int_{|x|\geq b^N}\big(1-\cos(zx)\big)\mu(dx)$$

$$\geq c|z|^\alpha - \frac{c}{2}|z|^\alpha = \frac{c}{2}|z|^\alpha \qquad \text{using (117) and (118),}$$

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

7

and hence $|\widetilde{\varphi}_t(z)| \leq \exp\left(-\frac{c}{2}t|z|^\alpha\right)$ for all $|z| \geq M$. Applying inversion formula, we get (for any $t > 0$)

$$
\begin{aligned}
\|f_{Y(t)}\|_\infty &\leq \frac{1}{2\pi} \int |\widetilde{\varphi}_t(z)| dz \\
&\leq \frac{1}{2\pi}\left[2M + \int_{|z| \geq M} |\widetilde{\varphi}_t(z)| dz \right] \\
&\leq \frac{M}{\pi} + \frac{1}{2\pi} \int \exp\left(-\frac{c}{2}t|z|^\alpha\right) dz \\
&= \frac{M}{\pi} + \frac{1}{2\pi} \cdot \frac{1}{t^{1/\alpha}} \int \exp\left(-\frac{c}{2}|x|^\alpha\right) dx \qquad \text{using } x = t^{1/\alpha} \cdot z \\
&\leq \frac{M}{\pi} + \frac{C_1}{t^{1/\alpha}} \qquad \text{where } C_1 = \frac{1}{2\pi} \int \exp\left(-\frac{c}{2}|x|^\alpha\right) dx.
\end{aligned}
$$

To conclude the proof, we set $C = \frac{M}{\pi} + C_1$. Now it only remains to prove claim (117).

Proof of Claim (117)

We proceed with a proof by contradiction. If $\inf_{z \in [1,b]} \eta(z) = 0$, then by continuity of $\eta(z)$, there exists some $z \in [1, b]$ such that

$$
\int_{\mathbb{R}} \left(1 - \cos(zx)\right) \mu(dx) = 0.
$$

Now for any $\epsilon > 0$, define the following sets:

$$
S = \{x \in \mathbb{R}: \ 1 - \cos(zx) > 0\} = \mathbb{R} \backslash \{\frac{2\pi}{z} k: \ k \in \mathbb{Z}\};
$$

$$
S_\epsilon = \{x \in \mathbb{R}: \ 1 - \cos(zx) \geq \epsilon\}.
$$

Observe that

- For any $\epsilon > 0$, we have $\epsilon \cdot \mu(S_\epsilon) \leq \int_{S_\epsilon} \left(1 - \cos(zx)\right)\mu(dx) \leq \int_{\mathbb{R}} \left(1 - \cos(zx)\right)\mu(dx) = 0$, which implies $\mu(S_\epsilon) = 0$;

- Meanwhile, $\lim_{\epsilon \to 0} \mu(S_\epsilon) = \mu(S) = 0$.

Together with the fact that $\mu(\mathbb{R}) > 0$ (so that the process is non-trivial), there must be some $m \in \mathbb{Z}$, $\delta > 0$ such that

$$
\mu(\{\frac{2\pi}{z} m\}) = \delta > 0.
$$

8

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

Besides, from $\mu(S) = 0$ we know that $\mu(\{-\frac{2\pi}{z}, \frac{2\pi}{z}\} \setminus \{0\}) = 0$. However, by definition of semi-stable processes in

(43) we know that $\mu = b^{-\alpha} T_b \mu$ where the transformation T_r ($\forall r > 0$) onto a Borel measure ρ on $\mathbb{R}$ is defined as

$(T_r \rho)(B) = \rho(r^{-1}B)$. This implies

$$\mu(\{\frac{2\pi m}{z} b^{-k}\}) > 0 \;\; \forall k = 1, 2, 3, \cdots$$

which would contradict $\mu(\{-\frac{2\pi}{z}, \frac{2\pi}{z}\} \setminus \{0\}) = 0$ eventually for k large enough. This concludes the proof of $\eta(z) > 0$ for

all $z \in [1, b]$. $\quad\square$

Appendix B: Barrier Option Pricing

B.1. Problem Setting This section considers the estimation of probabilities $P(A_n)$ with $A_n = \{\bar{X}_n \in A\}$ and

$$A \triangleq \{\xi \in \mathbb{D} : \xi(1) \le -b, \; \sup_{t \le 1} \xi(t) + ct \ge a\},$$

which corresponds to rare-event simulation in the context of down-and-in option. Here, we assume that a, $b > 0$ and

$c < a$. We consider the two-sided case in Assumption 1. That is, $X(t)$ is a centered Lévy process with Lévy measures

ν, and there exists some α, $\alpha' > 1$ such that $\nu[x, \infty) \in \mathcal{RV}_{-\alpha}(x)$ and $\nu(-\infty, -x] \in \mathcal{RV}_{-\alpha'}(x)$ as $x \to \infty$. Also, we

impose an alternative version of Assumption 2 throughout. Let $X^{(-z,z)}(t)$ be the Lévy process with with generating

triplet $(c_X, \sigma, \nu|_{(-z,z)})$. That is, $X^{(-z,z)}(t)$ is a modulated version of X where all jumps with size larger than z are

removed.

ASSUMPTION 4. *There exist z_0, C, $\lambda > 0$ such that*

$$\mathbf{P}\big(X^{(-z,z)}(t) \in [x, x+\delta]\big) \le \frac{C\delta}{t^\lambda \wedge 1} \qquad \forall z \ge z_0, \; t > 0, \; x \in \mathbb{R}, \; \delta > 0.$$

B.2. Importance Sampling Algorithm Below, we present the design of the importance sampling algorithm. For

any $\xi \in \mathbb{D}$ and $t \in (0, 1]$, let $\Delta\xi(t) = \xi(t) - \xi(t-)$ be the discontinuity in ξ at time t, and we set $\Delta\xi(0) \equiv 0$. Let

$$B^\gamma = \Big\{\xi \in \mathbb{D} : \#\{t \in [0, 1] : \Delta\xi(t) \ge \gamma\} \ge 1, \; \#\{t \in [0, 1] : \Delta\xi(t) \le -\gamma\} \ge 1\Big\}$$

and let $B_n^\gamma = \{\bar{X}_n \in B^\gamma\}$. Intuitively speaking, on event B_n^γ there is at least one upward and one downward "large" jump

in $\bar{X}_n$, where $\gamma > 0$ is understood as the threshold for jump sizes to be considered "large".

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

9

Fix some $w \in (0, 1)$, and let

$$\mathbf{Q}_n(\cdot) = w\mathbf{P}(\cdot) + (1 - w)\mathbf{P}(\,\cdot\,|B_n^\gamma).$$

The algorithm samples

$$L_n = Z_n \frac{d\mathbf{P}}{d\mathbf{Q}_n} = \frac{Z_n}{w + \frac{1-w}{\mathbf{P}(B_n^\gamma)}\mathbf{I}_{B_n^\gamma}}$$

under $\mathbf{Q}_n$. Now, we discuss the design of Z_n to ensure the strong efficiency of L_n. Analogous to the decomposition in

(17), let

$$J_n(t) = \sum_{s \in [0,t]} \Delta X(s)\mathbf{I}\big(|\Delta X(s)| \geq n\gamma\big),$$

$$\Xi_n(t) = X(t) - J_n(t) = X(t) - \sum_{s \in [0,t]} \Delta X(s)\mathbf{I}\big(|\Delta X(s)| \geq n\gamma\big).$$

Let $\bar{J}_n(t) = \frac{1}{n}J_n(nt)$, $\bar{J}_n = \{\bar{J}_n(t) : t \in [0,1]\}$, $\bar{\Xi}_n(t) = \frac{1}{n}\Xi_n(nt)$, and $\bar{\Xi}_n = \{\bar{\Xi}_n(t) : t \in [0,1]\}$. Meanwhile, set

$$M_c(t) \triangleq \sup_{s \leq t} X(s) + cs, \qquad Y_{n;c}^* \triangleq \mathbf{I}\big(M_c(n) \geq na, \ X(n) \leq -nb\big),$$

We have $\mathbf{I}_{A_n} = Y_{n;c}^*$. Under the convention $\hat{Y}_n^{-1} \equiv 0$, consider estimators Z_n of form

$$Z_n = \sum_{m=0}^{\tau} \frac{\hat{Y}_{n;c}^m - \hat{Y}_{n;c}^{m-1}}{\mathbf{P}(\tau \geq m)} \tag{119}$$

where τ is Geom(ρ) for some $\rho \in (0, 1)$ and is independent of everything else. Analogous to Proposition 1, the

following result provides sufficient conditions on $\hat{Y}_{n;c}^m$ for L_n to attain strong efficiency.

PROPOSITION 6. *Let $C_0 > 0$, $\rho_0 \in (0, 1)$, $\mu > \alpha + \alpha' - 2$, and $\bar{m} \in \mathbb{N}$. Suppose that*

$$\mathbf{P}\Big(Y_{n;c}^* \neq \hat{Y}_{n;c}^m \ \Big| \ \mathcal{D}^+(\bar{J}_n) = k, \ \mathcal{D}^-(\bar{J}_n) = k'\Big) \leq C_0\rho_0^m \cdot (k + k' + 1) \qquad \forall k, \ k' \geq 0, \ n \geq 1, \ m \geq \bar{m} \tag{120}$$

where $\mathcal{D}^+(\xi)$ and $\mathcal{D}^-(\xi)$ count the number of discontinuities of positive and negative sizes in ξ, respectively. Besides,

suppose that for all $\Delta \in (0, 1)$,

$$\mathbf{P}\Big(Y_{n;c}^* \neq \hat{Y}_{n;c}^m, \ \bar{X}_n \notin A^\Delta \ \Big| \ \mathcal{D}^+(\bar{J}_n) = 0 \text{ or } \mathcal{D}^-(\bar{J}_n) = 0\Big) \leq \frac{C_0\rho_0^m}{\Delta^2 n^\mu} \qquad \forall n \geq 1, \ m \geq 0 \tag{121}$$

10

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

*where $A^\Delta = \left\{ \xi \in \mathbb{D} : \sup_{t \in [0,1]} \xi(t) + ct \geq a - \Delta, \, \xi(1) \leq -b \right\}$. Then given $\rho \in (\rho_0, 1)$, there exists some $\bar{\gamma} = \bar{\gamma}(\rho) \in (0, b)$ such that for all $\gamma \in (0, \bar{\gamma})$, the estimators $(L_n)_{n \geq 1}$ are **unbiased and strongly efficient** for $\mathbf{P}(A_n) = \mathbf{P}(\bar{X}_n \in A)$ under the importance sampling distribution $\mathbf{Q}_n$.*

The proof is almost identical to that of Proposition 1. In particular, the proof requires that

$$\mathbf{P}(A_n) = \boldsymbol{O}\left(n\nu[n, \infty) \cdot n\nu(-\infty, -n] \right)$$

and that, for any $\beta > 0$, it holds for all γ small enough that

$$\mathbf{P}(A_n^\Delta \setminus B_n^\gamma) = \boldsymbol{o}(n^\beta)$$

where $A_n^\Delta = \{\bar{X}_n \in A^\Delta\}$. These can be obtained directly using sample path large deviations for heavy-tailed Lévy processes in Result 2. The Proposition 6 is then established by repeating the arguments in Proposition 1 using Result 4 for randomized debiasing technique.

B.3. Construction of $\hat{Y}_{n;c}^m$ Next, we describe the construction of $\hat{Y}_{n;c}^m$ that can satisfy the conditions in Proposition 6. Specifically, we consider the case where ARA is involved. Let

$$\Xi_{n;c}(t) \triangleq \Xi_n(t) + ct.$$

Under both $\mathbf{P}$ and $\mathbf{Q}_n$, $\Xi_{n;c}(t)$ admits the law of a Lévy process with generating triplet $(c_X + c, \sigma, \nu|_{(-n\gamma, n\gamma)})$. This leads to the Lévy-Ito decomposition

$$\Xi_{n,c}(t) \overset{d}{=} (c_X + c)t + \sigma B(t) + \underbrace{\sum_{s \leq t} \Delta X(s) \mathbf{I}\left(\Delta X(s) \in (-n\gamma, -1] \cup [1, n\gamma) \right)}_{\triangleq J_{n,-1}(t)}$$

$$+ \sum_{m \geq 0} \left[\underbrace{\sum_{s \leq t} \Delta X(s) \mathbf{I}\left(|\Delta X(s)| \in [\kappa_{n,m}, \kappa_{n,m-1}) \right) - t \cdot \nu\left((-\kappa_{n,m-1}, -\kappa_{n,m}] \cup [\kappa_{n,m}, \kappa_{n,m-1}) \right)}_{\triangleq J_{n,m}(t)} \right]$$

with $\kappa_{n,m}$ defined in (31). Besides, let $\bar{\sigma}^2(\cdot)$ be defined as in (33). For each $n \geq 1$ and $m \geq 0$, consider the approximation

$$\breve{\Xi}_{n;c}^m(t) \triangleq (c_X + c)t + \sigma B(t) + \sum_{q=-1}^{m} J_{n,q}(t) + \sum_{q \geq m+1} \sqrt{\bar{\sigma}^2(\kappa_{n,q-1}) - \bar{\sigma}^2(\kappa_{n,q})} \cdot W^q(t)$$

where $(W^m)_{m \geq 1}$ is a sequence of iid copies of standard Brownian motions independent of everything else.

Wang and Rhee: *Rare-Event Simulation for Regularly Varying Lévy Processes with Infinite Activities*
Article submitted to *Mathematics of Operations Research*

11

Next, we discuss how to apply SBA and construct approximators $\hat{Y}_{n;c}^m$'s in (119). Let $\zeta_k(t) = \sum_{i=1}^{k} z_i \mathbf{I}_{[u_i,n]}(t)$ be a piece-wise step function with k jumps over $(0,n]$, where $0 < u_1 < u_2 < \ldots < u_k \leq n$, and $z_i \neq 0$ for each $i \in [k]$. Recall that the jump times in ζ_k leads to a partition of $[0,n]$ of $(I_i)_{i \in [k+1]}$ defined in (23). For any I_i, let the sequence $l_j^{(i)}$'s be defined as in (26)–(27). Conditioning on $(l_j^{(i)})_{j \geq 1}$, one can then sample $\xi_{j;c}^{(i),m}, \xi_{j;c}^{(i)}$ using

$$\left(\xi_{j;c}^{(i)}, \xi_{j;c}^{(i),0}, \xi_{j;c}^{(i),1}, \xi_{j;c}^{(i),2}, \ldots \right) \overset{d}{=} \left(\Xi_{n;c}(l_j^{(i)}), \breve{\Xi}_{n;c}^0(l_j^{(i)}), \breve{\Xi}_{n;c}^1(l_j^{(i)}), \breve{\Xi}_{n;c}^2(l_j^{(i)}), \ldots \right).$$

The coupling in (10) then implies

$$\left(\Xi_{n;c}(u_i) - \Xi_{n;c}(u_{i-1}), \sup_{t \in I_i} \Xi_{n;c}(t) - \Xi_{n;c}(u_{i-1}), \breve{\Xi}_{n;c}^0(u_i) - \breve{\Xi}_{n;c}^0(u_{i-1}), \sup_{t \in I_i} \breve{\Xi}_{n;c}^0(t) - \breve{\Xi}_{n;c}^0(u_{i-1}), \right.$$

$$\left. \breve{\Xi}_{n;c}^1(u_i) - \breve{\Xi}_{n;c}^1(u_{i-1}), \sup_{t \in I_i} \breve{\Xi}_{n;c}^1(t) - \breve{\Xi}_{n;c}^1(u_{i-1}), \ldots \right)$$

$$\overset{d}{=} \left(\sum_{j \geq 1} \xi_{j;c}^{(i)}, \sum_{j \geq 1} (\xi_{j;c}^{(i)})^+, \sum_{j \geq 1} \xi_{j;c}^{(i),0}, \sum_{j \geq 1} (\xi_{j;c}^{(i),0})^+, \sum_{j \geq 1} \xi_{j;c}^{(i),1}, \sum_{j \geq 1} (\xi_{j;c}^{(i),1})^+, \ldots \right).$$

Now, we define

$$\hat{M}_{n;c}^{(i),m}(\zeta_k) = \sum_{j=1}^{m+\lceil \log_2(n^d) \rceil} (\xi_{j;c}^{(i),m})^+$$

as an approximation to $M_{n;c}^{(i),*}(\zeta_k) = \sup_{t \in I_i} \Xi_{n;c}(t) - \Xi_{n;c}(u_{i-1}) = \sum_{j \geq 1} (\xi_{j;c}^{(i)})^+$. Now, set

$$\hat{Y}_{n;c}^m(\zeta_k) = \left[\max_{i \in [k+1]} \mathbf{I}\left(\sum_{q=1}^{i-1} \sum_{j \geq 0} \xi_{j;c}^{(q),m} + \sum_{q=1}^{i-1} z_q + \hat{M}_{n;c}^{(i),m}(\zeta_k) \geq na \right) \right] \cdot \mathbf{I}\left(\sum_{q=1}^{k+1} \sum_{j \geq 0} \xi_{j;c}^{(q),m} + \sum_{q=1}^{k} z_q - cn \leq -nb \right).$$

In (119), we plug in $\hat{Y}_{n;c}^m = \hat{Y}_{n;c}^m(J_n)$.

The proof of the strong efficiency is almost identical to that of Theorem 2. The only major difference is that in Lemma 7, we apply Assumption 4 instead of Assumption 2.