Proceedings of the 2020 Winter Simulation Conference
K.-H. Bae, B. Feng, S. Kim, S. Lazarova-Molnar, Z. Zheng, T. Roeder, and R. Thiesing, eds.

RARE-EVENT SIMULATION FOR MULTIPLE JUMP EVENTS IN HEAVY-TAILED LÉVY PROCESSES WITH INFINITE ACTIVITIES

Xingyu Wang

Department of Industrial Engineering
and Management Sciences
Northwestern University
2145 Sheridan Road
Evanston, IL 60208, USA

Chang-Han Rhee

Department of Industrial Engineering
and Management Sciences
Northwestern University
2145 Sheridan Road
Evanston, IL 60208, USA

ABSTRACT

In this paper we address the problem of rare-event simulation for heavy-tailed Lévy processes with infinite activities. We propose a strongly efficient importance sampling algorithm that builds upon the sample path large deviations for heavy-tailed Lévy processes, stick-breaking approximation of extrema of Lévy processes, and the randomized debiasing Monte Carlo scheme. The proposed importance sampling algorithm can be applied to a broad class of Lévy processes and exhibits significant improvements in efficiency when compared to crude Monte-Carlo method in our numerical experiments.

1 INTRODUCTION

In this paper, we propose a strongly efficient rare-event simulation algorithm for general Lévy processes with heavy-tailed jump distributions characterized by regular variation. Specifically, our goal is to estimate probabilities of the form $\mathbb{P}(A_n)$ for large n, where $A_n = \{\bar{X}_n \in A\}$, A is a subset of the Skorokhod path space, and $\bar{X}_n$ is a scaled Lévy process X with heavy-tailed jump distributions characterized by regular variation. Such problems arise in many different applications such as ruin and risk theory (Asmussen and Albrecher 2010), option pricing (Tankov 2003), and queuing networks (Debicki and Mandjes 2015).

Two major challenges arise when designing an efficient rare-event simulation algorithm for general Lévy processes with heavy-tailed jumps. First, the nature of the rare events renders the crude Monte-Carlo method extremely inefficient when n is large: assume the goal is to estimate $\mathbb{P}(A_n)$ with a given level of confidence on its relative error, then the number of samples required would approach ∞ as $n \to \infty$ and $\mathbb{P}(A_n)$ tends to 0. In the light-tailed case, one typical solution is to perform an exponential change of measure and analyze a properly tilted process that induces a much higher probability of occurrence for the desired event. The theory of large deviations can be used to determine the right amount of exponential tilting. For instance, when viewing risk processes from the perspective of large deviations, the asymptotic distribution of sample paths leading to ruination coincides with the distribution of the exponentially biased risk process parametrized by the solution of the Lundberg equation (Asmussen and Albrecher 2010); and under the guidance of large deviation principles, importance sampling algorithms have been proposed to asymptotically optimally simulate rare events in a dynamic fashion (Dupuis and Wang 2004), or simulate rare events in queuing networks with established bound on required computational efforts (Boxma et al. 2019). Similarly, for Lévy processes with heavy-tailed jumps, one would expect that the design of an efficient rare-event simulation algorithm entails the knowledge of large deviation results for the associated processes, since the large deviation principles not only characterize the decaying rate of $\mathbb{P}(A_n)$, but also describe the most likely scenario for the event to occur. Indeed, as revealed in (Rhee et al. 2019), by solving an optimization problem concerning the minimal number of jumps l^* required for a step function to trigger

978-1-7281-9499-8/20/$31.00 ©2020 IEEE

the target event, we see that, asymptotically, the sample paths leading to occurrence of rare events A_n are those with l^* large jumps. By exploiting this result to design a proper importance sampling distribution, (Chen et al. 2019) proposes a strongly efficient rare-event simulation algorithm for compound Poisson processes and random walks with regularly varying jumps. The current paper extends this framework, and presents an importance sampling algorithm for rare-event simulation of general Lévy processes with regularly varying jump distributions, beyond the compound Poisson processes.

The second difficulty lies in exact simulation of the sample path for general Lévy processes. Unlike compound Poisson processes or random walks, the sample path of a general Lévy process may not be exactly simulatable due to its infinite activities from the presence of either a Brownian motion or the infinitely many jumps within finite time intervals. Since many events A_n that arise in applications can be characterized in terms of the extreme behavior of process X within given time interval, one possible remedy is to simulate the extrema of the Lévy process instead of the entire sample path. However, an explicit expression of the distribution of the extrema, or an exact simulation algorithm for extrema, is not available for Lévy processes, except for a few specific cases such as spectrally one-sided processes (Michna et al. 2015)(Chaumont and Małecki 2018) or stable processes (Cázares et al. 2019).

In the current work, we address these issues by combining the stick-breaking approximation (SBA) idea for extrema of general Lévy processes proposed in (Cázares et al. 2018) and the debiasing technique from (Rhee and Glynn 2015) with the mixture importance sampling from (Chen et al. 2019). The foundation of SBA is the detailed study of (Pitman and Bravo 2012) on the concave majorants of Lévy processes, the distribution of which admits a Poisson-Dirichlet type of iterative structure, thus ensuring a geometrical convergence rate in SBA when estimating expectation of functionals on extrema of Lévy processes. By studying the distributional properties of Lévy processes, we show that our algorithm is strongly efficient for a broad class of Lévy processes.

The rest of the paper is organized as follows. We provide preliminaries of the work in Section 2, and detail the algorithm in Section 3. In Section 4 we establish a set of conditions under which the proposed algorithm is strongly efficient, and discuss the proper choice of parameters in the algorithm. In Section 5 we demonstrate the efficiency of the proposed importance sampling strategy with numerical experiments.

2 PRELIMINARIES

The importance sampling algorithm proposed in this paper builds upon the large deviations results for Lévy processes with regularly varying increments. The related notions and results are introduced below. We use $(\mathbb{D}, d)$ to denote the Skorokhod metric space of real-valued càdlàg functions with domain $[0,1]$. For any positive integer l, define

$$\mathbb{D}_l \triangleq \{\xi \in \mathbb{D} :\ \xi \text{ is a non-decreasing step function with } l \text{ jumps}, \xi(0) = 0\}.$$

For $l = 0$, let $\mathbb{D}_0 = \{\mathbf{0}\}$ where $\mathbf{0}(t) = 0\ \forall t \in [0,1]$. Furthermore, for each $l \in \mathbb{N}^+$, we define $\mathbb{D}_{<l} \triangleq \bigcup_{j=0}^{l-1} \mathbb{D}_j$. Any Lévy process $\{X(t) : t \geq 0\}$ is characterized by its generating triplet (c, σ^2, ν), where $c \in \mathbb{R}$ is the drift parameter, $\sigma \geq 0$ is the magitude of the Brownian motion term in $X(t)$, and ν is the Lévy measure of the process such that $\int (|x|^2 \wedge 1)\nu(dx) < \infty$. See chapter 4 of (Sato et al. 1999) for details.

The heavy-tailed behavior of the positive jumps will be characterized by regular variation: recall that a Borel measurable function $\varphi : (0,\infty) \mapsto (0,\infty)$ is said to be regularly varying with index $\rho \in \mathbb{R}$ at $+\infty$ (denoted as $\varphi \in \mathrm{RV}_\rho$) if for any $t > 0$, $\lim_{x\to+\infty} \frac{\varphi(tx)}{\varphi(x)} = t^\rho$. For simplicity of the exposition, we focus on heavy-tailed behavior of positive jumps: In terms of the function $f(x) = \nu[x,\infty)$, we assume that $f \in RV_{-\alpha_+}$ with $\alpha_+ > 1$.

For any positive integer n, define the centered and scaled version of X as $\bar{X}_n(t) \triangleq \frac{1}{n} X_n(t) - ct - \mu_1 t$ where $\mu_1 = \int_{|x|\geq 1} x\nu(dx)$ and we assume $\mu_1 < \infty$. For any $\beta > 0$, let ν_β be the measure concentrated on $(0,\infty)$ with $\nu_\beta(x,\infty) = x^{-\beta}$. For any positive integer l, use ν_β^l to denote the $l-$fold product measure of ν_β restricted onto

$\{y \in (0,\infty)^l:\ y_1 \geq y_2 \geq \cdots \geq y_l\}$, and define the measure $C^l_\beta(\cdot) \triangleq \mathbb{E}\left[\nu^l_\beta\{y \in (0,\infty)^l:\ \sum_{j=1}^l y_j \mathbb{1}_{[U_j,1]} \in \cdot\}\right]$ where $(U_j)_{j\geq 1}$ is an i.i.d. sequence of Unif$(0,1)$; while for $l=0$, let C^0_β be the Dirac measure on $\mathbf{0}$. The following results describes the sample path large deviations for the corresponding scaled process $\bar{X}_n$.

Result 1 (Theorem 3.1/3.4 of (Rhee et al. 2019)) For any set A that is Borel measurable in $\mathbb{D}$ and is bounded away from $\mathbb{D}_{<l^*}$ where $l^* \triangleq \min\{l \in \mathbb{N}:\ \mathbb{D}_l \cap A \neq \emptyset\}$, we have

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

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

The rare events we concern in this paper are characterized by the extrema of the Lévy process X. For any $t>0$, we define the running supremum and infimum processes of X as $\bar{M}(t) = \sup_{s\in[0,t]} X(s)$. Results in (Pitman and Bravo 2012) provide useful tools for studying $\bar{M}$ using the iterative structure of the concave majorant of Lévy processes. Specifically, given any $t>0$ and a Lévy process X that is not a compound Poisson process with drift, the distribution of the pair $(\bar{M}(t), X(t))$ admits the following expression

$$(\bar{M}_t, X_t) \stackrel{d}{=} (\sum_{j\geq 1}(\xi_j)^+, \sum_{j\geq 1}\xi_j) \tag{1}$$

where the symbol $\stackrel{d}{=}$ denotes equivalence in distribution and $(\cdot)^+ \triangleq \max\{\cdot, 0\}$, ξ_j's are independent random variables such that $\xi_j \stackrel{d}{=} X(l_j)$ where $(l_j)_{j\geq 1}$ is the stick-breaking sequence defined by $(U_j)_{j\geq 1}$, a sequence of i.i.d. Unif$(0,1)$, as follows:

$$l_1 = tU_1,\quad l_j = U_j(t - l_1 - l_2 - \cdots l_{j-1})\ \forall j \geq 2.$$

A similar expression applies to the running infimum. See Theorem 1 in (Pitman and Bravo 2012) for details, and (Cázares et al. 2018) for the stick-breaking approximation (SBA) algorithm for efficient approximation of extrema of Lévy processes.

To achieve unbiasedness for the proposed estimators, we apply the debiasing techniques used in (Rhee and Glynn 2015):

Result 2 (Theorem 1 in (Rhee and Glynn 2015)) Given a random variable Y and a sequence of random variables $(Y_n)_{n\geq 0}$ such that $\lim_{n\to\infty}\mathbb{E}Y_n = \mathbb{E}Y$, and a positive integer-valued random variable N with unbounded support such that N is independent of $(Y_n)_{n\geq 0}$ and Y, if $\sum_{n\geq 1}\mathbb{E}|Y_{n-1}-Y|^2 \Big/ \mathbb{P}(N\geq n) < \infty$, then for

$$Z = \sum_{n=1}^N (Y_n - Y_{n-1}) \Big/ \mathbb{P}(N \geq n),$$

(with the convention $Y_{-1}=0$) Z is an element of L^2, and

$$\mathbb{E}Z = \mathbb{E}Y, \qquad \mathbb{E}Z^2 = \sum_{n\geq 0} \bar{v}_n \Big/ \mathbb{P}(N\geq n),$$

where $\bar{v}_n = \mathbb{E}|Y_{n-1}-Y|^2 - \mathbb{E}|Y_n - Y|^2$.

The goal of the work is to propose an importance sampling algorithm for rare-event simulation of heavy-tailed Lévy processes that achieves strong efficiency. Specifically, for a sequence of events $(A_n)_{n\geq 1}$ such that $\mathbb{P}(A_n)\to 0$ as $n\to\infty$, we say that a sequence of estimators $(L_n)_{n\geq 1}$ is unbiased and *strongly efficient* if we have $\mathbb{E}L_n = \mathbb{P}(A_n)$ for any $n\geq 1$, and $\mathbb{E}L_n^2 = \mathscr{O}(\mathbb{P}^2(A_n))$. Here, for two sequences of non-negative real numbers $(x_n)_{n\geq 1}$ and $(y_n)_{n\geq 1}$, we say $x_n = \mathscr{O}(y_n)$ if $\limsup_{n\to\infty}\frac{x_n}{y_n} < \infty$. Besides, we write $x_n = o(y_n)$ for the two positive real sequences if $\lim_{n\to\infty}\frac{x_n}{y_n} = 0$.

3 THE ALGORITHM

In this section, we describe the structure of the rare events we are interested in, and propose an importance-sampling algorithm for efficient estimation of their probability. For clarity, this section focuses on one running example which will be introduce shortly, and describes the algorithm tailored for the specific example. Nevertheless, it is worth mentioning that the principle underlying the algorithm proposed below enjoys greater flexibility and can be extended to more general cases.

3.1 The Rare Events $(A_n)_{n\geq 1}$ and the Process $X(t)$

Define the set

$$A = \{\xi \in \mathbb{D} : \sup_{t\in[0,1]} \xi(t) \geq a; \sup_{t\in(0,1]} \xi(t) - \xi(t-) < b\}. \tag{2}$$

In other words, $\xi \in A$ if the supremum of ξ has reached a, but no jump in ξ is larger than b. Furthermore, we make the following assumption about set A:

Assumption 1 $a,b > 0,\ a/b \notin \mathbb{Z}$.

Consider $l^* \triangleq \min\{l \in \mathbb{N} :\ \mathbb{D}_l \cap A \neq \emptyset\}$. In this case, we have $l^* = \lceil a/b \rceil$, and $l^* \geq 1$. Recall that for a step function ξ to belong to set A, ξ needs to have at least l^* jumps. Moreover, it is easy to see that, under Assumption 1, the set A is bounded away from $\mathbb{D}_{<l^*}$, and $C_\beta^{l^*}(A^\circ) > 0$ for any $\beta > 0$.

We study a Lévy process $\{X_t : t \geq 0\}$ with generating triplet (c_X, σ^2, ν). Since the case of compound Poisson processes were already treated in (Chen et al. 2019), we assume in this paper that X is not a compound Poisson process with linear drift, which implies that either $\sigma > 0$ or $\nu(-1,1) = \infty$. Furthermore, we reiterate several assumptions: (1) $\int_{|x|>1} |x|\nu(dx) < \infty$ so $X_t \in L_1$ for any $t \geq 0$; (2) as for the heavy-tail behavior of the positive jumps, the function $f(x) = \nu[x,\infty)$ is regularly varying at ∞ with index $-\alpha_+ < -1$; (3) the drift coefficient c_X is chosen specifically so that the process is already centered: $\mathbb{E}X_t = 0,\ \forall t > 0$. In this case, the scaled and centered version of X is $\bar{X}_n = \{X(nt)/n :\ t \in [0,1]\}$ for any $n \in \mathbb{Z}^+$.

Let $A_n \triangleq \{\bar{X}_n \in A\}$. The goal is to propose an algorithm for estimating $\mathbb{P}(A_n)$. To achieve unbiasedness and strong efficiency of the algorithm, we need the following assumption regarding distributions of $X(t)$. For a measure space $(\mathscr{X}, \mathscr{F}, \mu)$ and any $A \in \mathscr{F}$, denote the restriction of the measure μ on A as $\mu|_A(\cdot) \triangleq \mu(A \cap \cdot)$.

Assumption 2 For any $z_0 > 0$, there exist $C > 0, \alpha > 0, \theta \in (0,1]$ such that for any $t > 0, z \geq z_0, x \in \mathbb{R}, \delta \in [0,1]$, we have

$$\mathbb{P}(X^{<z}(t) \in [x, x+\delta]) \leq \frac{C}{t^\alpha \wedge 1}\delta^\theta;$$

where the process $X^{<z}$ is the Lévy process with the generating triplet $(c_X, \sigma^2, \nu|_{(-\infty,z)})$.

A process that has the same distribution as $X^{<z}$ can be obtained by removing all jumps larger than z from X. Similarly, we define $X^{\geqslant z}$ as the compound Poisson process with the generating triplet $(0,0,\nu_{[z,\infty)})$, and $X^{\geqslant z}$ is understood as the compound Poisson process generated merely by all jumps larger than z in X. Note that in Section 4, we show that Assumption 2 is a moderate condition.

3.2 Importance Sampling Strategy and Construction of the Unbiased Estimator

Our algorithm builds on the construction of the importance sampling distribution in (Chen et al. 2019). Consider the rare event simulation problem for some fixed scaling level $n \in \mathbb{Z}^+$. For any $\gamma > 0$, define sets $B_n^\gamma \triangleq \{\bar{X}_n \in B^\gamma\}$ where

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

namely, $\xi \in B^\gamma$ if and only if ξ has at least l^* jumps with size larger than γ. Now fix any $w \in (0,1)$, and define the following importance sampling distribution

$$\mathbb{Q}(\cdot) = w\mathbb{P}(\cdot) + (1-w)\mathbb{P}(\cdot|B_n^\gamma). \tag{3}$$

In the meantime, consider the following decomposition of X at point $n\gamma$: $X = \widetilde{X} + J_n$ where the two independent processes $\widetilde{X} = X^{<n\gamma}, J_n = X^{\geqslant n\gamma}$ can be though of as the *small-jump* and *large-jump* processes of X, with generating triplets $(c_X, \sigma, \nu|_{(-\infty,n\gamma)})$ and $(0,0,\nu|_{[n\gamma,\infty)})$ respectively.

Now let us observe two facts: first, $\mathbb{Q}$ is absolutely continuous w.r.t. $\mathbb{P}$ and vise versa; second, using the decomposition above, we see that $X^{<n\gamma}$ admits the same marginal distribution under $\mathbb{P}$ and $\mathbb{Q}$, as $\mathbb{Q}$ only alters the distribution of the *large-jump* process J_n. Therefore, to generate a sample path of X under $\mathbb{Q}$, we can sample the large jump process J_n from $\mathbb{Q}$, and then generate $\widetilde{X}$ under $\mathbb{P}$. To be more precise, we define set $E = \{\xi \in \mathbb{D} : \sup_{t\in[0,1]} \xi(t) - \xi(t-) < b\}$, and propose the following importance sampling estimator

$$L_n = Z_n(J_n)\mathbb{1}_E(J_n/n)\frac{d\mathbb{P}}{d\mathbb{Q}} = \frac{Z_n(J_n)\mathbb{1}_E(J_n/n)}{w + \frac{1-w}{\mathbb{P}(B_n^\gamma)}\mathbb{1}_{B_n^\gamma}(J_n)} \tag{4}$$

where J_n is sampled from $\mathbb{Q}$ and Z_n is a stochastic function such that for any step function ζ on $[0,n]$,

$$\mathbb{E}Z_n(\zeta) = \mathbb{P}\Big(\sup_{t\in[0,n]} \widetilde{X}(t) + \zeta(t) \geq na\Big).$$

Now it remains to describe: (a) the construction of Z_n; (b) the procedure of sampling J_n from $\mathbb{Q}$ (in particular, sampling J_n from $\mathbb{P}(\cdot|B_n^\gamma)$). For the first task, we combine the SBA algorithm with the debiasing technique as follows. To begin with, the nature of a jump process indicates the existence of some $k \in \{0,1,2,\cdots\}$ and sequences of real numbers $(z_i)_{i=1}^k, (u_i)_{i=1}^k$ with $u_i \in [0,n]$ and $(u_i)_{i=1}^k$ being distinct, such that $\zeta_k = \sum_{i=1}^k z_i \mathbb{1}_{[u_i,n]}$. From now on we use the subscript k to indicate the number of jumps in ζ. Given the representation $\zeta_k = \sum_{i=1}^k z_i \mathbb{1}_{[u_i,n]}$, the interval $(0,n]$ can be partitioned into $\{(u_i,u_{i+1}]\}_{i=0}^k$ with the convention that $u_0 = 0, u_{k+1} = n$. For each $i = 0,1,\cdots,k$, we conduct the following stick-breaking procedure on $(u_i,u_{i+1}]$:

$$l_1^{(i)} = U_1^{(i)}(u_{i+1} - u_i); \tag{5}$$

$$l_j^{(i)} = U_j^{(i)}(u_{i+1} - u_i - l_1^{(i)} - l_2^{(i)} - \cdots - l_{j-1}^{(i)}) \quad \forall j = 2,3,\cdots \tag{6}$$

where $(U_j^{(i)})_{j\geq1}$ is an i.i.d. sequence of Unif$(0,1)$. Next, for any given $0 \leq i \leq k, j \geq 1$, independently sample $\xi_j^{(i)} \sim F_{\widetilde{X}}(\cdot, l_j^{(i)})$ where we use $F_Y(\cdot,t)$ to denote the law of Y_t for any Lévy process Y. Let us define (for any $i = 0,1,\cdots,k$) $\widetilde{M}^{(i)} \triangleq \sum_{l=0}^{i-1}\sum_{j\geq1}\xi_j^{(l)} + \sum_{j\geq1}(\xi_j^{(i)})^+$ with the convention that $\sum_{i=0}^{-1}\cdot = 0$ and $(\cdot)^+ = \max\{0,\cdot\}$. Due to the coupling in (1), one can see that

$$\Big(\sup_{t\in(u_0,u_1]} \widetilde{X}_t, \sup_{t\in(u_1,u_2]} \widetilde{X}_t, \cdots, \sup_{t\in(u_k,u_{k+1}]} \widetilde{X}_t\Big) \stackrel{d}{=} \Big(\widetilde{M}^{(0)}, \widetilde{M}^{(1)}, \cdots, \widetilde{M}^{(k)}\Big).$$

Recall our current task: unbiased estimation for expectation of the indicator random variable

$$Y_n^*(\zeta_k) = \mathbb{1}\Big\{\max_{i=0,1,\cdots,k} \widetilde{M}^{(i)} + \zeta_k(u_i) \geq na\Big\}$$

given ζ_k. To apply the debiasing technique, the next step is to define a sequence of random variables $(Y_{n,m}(\zeta_k))_{m\geq 1}$ and approximate $Y_n^*(\zeta_k)$, where the subscript m indicates the approximation level of SBA employed by $Y_{n,m}$. Specifically, for any $m \geq 0$, define

$$\widetilde{M}_m^{(i)} = \sum_{l=0}^{i-1} \sum_{j\geq 0} \xi_j^{(l)} + \sum_{j=1}^{\lceil \log_2(n^2) \rceil + m} (\xi_j^{(i)})^+$$

where $\lceil x \rceil$ denotes the smallest integer that is larger than or equal to x.

Several remarks about the term $\widetilde{M}_m^{(i)}$: (a) As an approximation to $\widetilde{M}^{(i)} \stackrel{d}{=} \sup_{t\in(u_i,u_{i+1}]} \widetilde{X}_t$, $\widetilde{M}_m^{(i)}$ differs from $\widetilde{M}^{(i)}$ as it only inspects the increments of $\widetilde{X}$ on finitely many sticks; (b) Term $\lceil \log_2(n^2) \rceil$ dictates that: as far as SBA is concerned, the algorithm always performs at least $\lceil \log_2(n^2) \rceil$ SBA steps at the scaling level n; this choice serves to ensure the strong efficiency of the algorithm, and would not increase the expected computational time significantly.

Now, by defining $Y_{n,m}(\zeta_k) = \mathbb{1}\Big\{ \max_{i=0,1,\cdots,k} \widetilde{M}_m^{(i)} + \zeta_k(u_i) \geq na \Big\}$ for any $m \geq 0$ and let $Y_{n,-1}(\cdot) = 0$, we construct the desired unbiased estimator as follows:

$$Z_n(\zeta_k) = \sum_{m=0}^{\tau} \Big(Y_{n,m}(\zeta_k) - Y_{n,m-1}(\zeta_k) \Big) \Big/ \mathbb{P}(\tau \geq m), \tag{7}$$

where the randomized truncation index τ, independent of everything else, is chosen to be geometrically distributed with law $\mathbb{P}(\tau > m) = \rho^m$ for some $\rho \in (0,1)$ in our algorithm. Due to τ being finite almost surely, the number of $\xi_j^{(i)}$ we need to generate for evaluation of $Z_n(\zeta_k)$ is finite and depends on the value of τ. The said parametrization will be justified in Section 4 as we see that $(L_n)_{n\geq 1}$ is strongly efficient.

3.3 Sampling from $\mathbb{P}(\cdot|B_n^\gamma)$

Below we revisit the problem of sampling the *large-jump* process J_n from the conditional distribution $\mathbb{P}(\cdot|B_n^\gamma)$, and propose Algorithm 2. The rationale of the algorithm can be made clear once we observe the following facts, and the argument therein is a direct application of point transform and augmentation for Poisson random measures; for details, see Chapter 5 of (Resnick 2007).

First, to simulate J_n (under the original law $\mathbb{P}$), it suffices to simulate a Poisson random measure N_n on $[0,n] \times \mathbb{R}^+$ with intensity measure $\mathbf{Leb}[0,n] \times \nu_n$ where $\nu_n(\cdot) = \nu\big(\cdot \cap [n\gamma,\infty)\big)$. The Poisson random measure N_n admits the expression

$$N_n(\cdot) = \sum_{i=1}^{\widetilde{N}_n} \mathbb{1}\{(S_i,W_i) \in \ \cdot\ \}$$

where $\widetilde{N}_n \sim \text{Poisson}(n \cdot \nu[n\gamma,\infty))$ is the number of simulated points in N_n, $(S_i)_{i\geq 1}$ is an iid sequence of $\text{Unif}(0,n)$, and $(W_i)_{i\geq 1}$ is an iid sequence from the distribution $\nu_n(\cdot)/\nu_n[n\gamma,\infty)$; here we interpret S_i as the arrival time of the $i-$th large jump, W_i as its height, and $\widetilde{N}_n$ as the number of jumps in J_n on $[0.n]$. Next, consider the simulation of a Poisson random measure with intensity measure ν_n using the inversion function:

$$Q_n^{\leftarrow}(y) \triangleq \inf\{s > 0 : \nu_n[s,\infty) < y\}.$$

Algorithm 1 Efficient Estimation of $\mathbb{P}(A_n)$

Require: $w \in (0,1), \gamma > 0, \rho \in (0,1)$

```
if Unif(0,1) < w then                                          ▷ Sample J_n from Q
    Sample J_n = Σ_{i=1}^{k} z_i 1_[u_i,n] from P
else
    Sample J_n = Σ_{i=1}^{k} z_i 1_[u_i,n] from P(·|B_n^γ) using Algorithm 2
end if
Let u_0 = 0, u_{k+1} = n.
Sample τ ~ Geom(ρ)                                             ▷ Decide Truncation Index τ
for i = 0,1,···,k do                       ▷ Generate Stick Lengths, and Decide Increments
    Sample U_1^(i) ~ Unif(0,1). Let l_1^(i) = U_1^(i)(u_{i+1} − u_i)
    Sample ξ_{i,1} ~ F_X̃(·, l_1^(i))
    for j = 2,3,···,⌈log_2(n^2)⌉ + τ do
        Sample U_j^(i) ~ Unif(0,1). Let l_j^(i) = U_j^(i)(u_{i+1} − u_i − l_1^(i) − l_2^(i) − ··· − l_{j−1}^(i))
        Sample ξ_{i,j} ~ F_X̃(·, l_j^(i))
    end for
    Let l_{⌈log(n^2)⌉+τ+1}^(i) = u_{i+1} − u_i − l_1^(i) − l_2^(i) − ··· − l_{⌈log(n^2)⌉+τ}^(i)
    Sample ξ_{i,⌈log_2(n^2)⌉+τ+1} ~ F_X̃(·, l_{⌈log_2(n^2)⌉+τ+1}^(i))
end for
for m = 0,1,···,τ do                                           ▷ Evaluate Y_{n,m}
    for i = 0,1,2,···,k do
        Let M̃_m^(i) = Σ_{l=0}^{i−1} Σ_{j=1}^{⌈log_2(n^2)⌉+τ+1} ξ_{l,j}^m + Σ_{j=1}^{⌈log_2(n^2)⌉+τ} (ξ_{i,j}^m)^+
    end for
    Let Y_{n,m} = 1{ max_{i=0,1,···,k} M̃_m^(i) + J_n(u_i) ≥ na }
end for
Let Z_n = Y_{n,0} + Σ_{m=1}^{τ} (Y_{n,m} − Y_{n,m−1})/ρ^{m−1}        ▷ Return the Estimator L_n
if max_{i=1,···,k} z_i > b then
    Return L_n = 0.
else
    Let λ_n = nν[nγ,∞), p_n = 1 − Σ_{l=0}^{l*−1} e^{−λ_n} λ_n^l / l!, I_n = 1{J_n ∈ B_n^γ}
    Return L_n = Z_n/(w + (1−w)/p_n · I_n)
end if
```

Algorithm 2 Simulation of J_n under $\mathbb{P}(\cdot|B_n^\gamma)$

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

```
Sample k ~ Poisson(n · ν[nγ,∞)) conditioned on {k ≥ l*}
Sample Γ_1,···,Γ_k ~(i.i.d.) Unif[0, ν[nγ,∞)]
Sample U_1,···,U_k ~(i.i.d.) Unif[0,n]
Return J_n = Σ_{i=1}^{k} Q_n^←(Γ_i) 1_[U_i,n]
```

This inverse function has the property that $y \le \nu_n[s,\infty) \Leftrightarrow Q_n^{\leftarrow}(y) \ge s$. Therefore, for iid Exponential (with rate 1) random variables $\{E_i\}_{i\in\mathbb{Z}^+}$ and the corresponding running sum $\Gamma_i = \sum_{j=1}^{i} E_j$, it is known that

$$\sum_{i:\Gamma_i \le \nu_n[n\gamma,\infty)} \delta_{Q_n^{\leftarrow}(\Gamma_i)}$$

is the desired Poisson random measure, where δ_x denotes Dirac measure at x. Now, by augmenting $\{Q_n^{\leftarrow}(\Gamma_i)\}_{i\geq 1}$ with uniformly distributed random marks on $[0,n]$, we have that

$$N_n \stackrel{d}{=} \sum_{\Gamma_i \leq \nu_n[n\gamma,\infty)} \delta_{\left(U_i, Q_n^{\leftarrow}(\Gamma_i)\right)};\quad J_n \stackrel{d}{=} \sum_{\Gamma_i \leq \nu_n[n\gamma,\infty)} Q_n^{\leftarrow}(\Gamma_i)\mathbb{1}_{[U_i,n]}$$

where $(U_i)_{i\geq 1}$ is a sequence of iid Unif$(0,n)$ random variables that are independent of $\{\Gamma_i\}_{i\geq 1}$. Lastly, the condition $\bar{X}_n \in B^\gamma$ is equivalent to $\sup\{i : \Gamma_i \leq \nu_n[n\gamma,\infty)\} \geq l^*$. For any $k \geq l^*$, by further conditioning on the event $\big\{\sup\{i : \Gamma_i \leq \nu_n[n\gamma,\infty)\} = k\big\}$, the distribution of $(\Gamma_1,\cdots,\Gamma_k)$ is the same as that of the order statistics of k iid random variables from Unif$[0,\nu_n[n\gamma,\infty)]$. With the difficulty of sampling from $\mathbb{P}(\cdot|B_n^\gamma)$ resolved, we yield an importance sampling strategy that is readily implementable, and we detail the steps in Algorithm 1,

4 ANALYSIS OF THE ALGORITHM

This section is devoted to theoretical aspects of the proposed algorithm. We present Theorem 1 and 2 with proof sketches, and focus our discussion in this paper on their implications, the choice of parameters in the implementation, and the applicability of the proposed algorithm. The full details can be found in (Wang and Rhee 2020).

4.1 Strong Efficiency of $(L_n)_{n\geq 1}$

We state the main result regarding the efficiency of the importance sampling algorithm.

Theorem 1 Suppose that Assumption 1 and Assumption 2 are in force, and γ (which characterizes the set B^γ) and ρ (which determines the distribution of $\tau \sim$ Geom(ρ)) are as follows.

- Choose $\gamma \in (0, \frac{a-(l^*-1)b}{l^*})$ such that $\frac{a-(l^*-1)b}{\gamma}$ is not an integer.
- Let α,θ be the values stated in Assumption 2, and choose

$$\delta \in (1/\sqrt{2},1),\ \alpha_3 \in (0,\frac{\theta}{\alpha}),\ \alpha_4 \in (0,\frac{\theta}{2\alpha}),\ \alpha_2 \in (0,(\alpha_3/2)\wedge 1),\ \alpha_1 \in (0,\frac{\theta}{\alpha\alpha_2}).$$

Pick ρ such that

$$1 > \rho > \sqrt{\max\{\delta^\alpha, \frac{1}{\delta\sqrt{2}}, \delta^{\theta\alpha_2-\alpha\alpha_1}, \delta^{\theta-\alpha\alpha_3}, \delta^{-\alpha_2+\frac{\alpha_3}{2}}\}}.$$

Then, $(L_n)_{n\geq 1}$ is unbiased and strongly efficient for $(A_n)_{n\geq 1}$; namely;

$$\mathbb{E}^{\mathbb{Q}}[L_n] = \mathbb{P}(A_n),\quad \mathbb{E}^{\mathbb{Q}}[L_n^2] = \mathscr{O}(\mathbb{P}^2(A_n)).$$

4.2 Distributional Property of Small-Jump Processes $X^{<z}$

Below we provide a sufficient condition for Assumption 2, and show that a broad class of Lévy processes therefore can be addressed by the proposed algorithm. In particular, the conditions below verifies Assumption 2 with $\theta = 1$, which is equivalent to showing Lipschitz continuity of the law of $X^{<z}$.

First of all, if $\sigma > 0$, then for a fixed $\gamma_0 > 0$ and any $\gamma \geq \gamma_0$, we have the decomposition $X^{<\gamma}(t) = \sigma B(t) + Y^{<\gamma}(t)$ where B is a standard Brownian motion, $Y^{<\gamma}$ is a Levy process with generating triplet $(0,0,\nu|_{(-\gamma,\gamma)})$, and the two processes are independent. Now for any $x\in\mathbb{R}$ and $\delta\in(0,1)$, we have

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

Therefore, Assumption 2 holds with $\theta = 1, \alpha = 1/2$. From now on, we focus on the case where $\sigma = 0$. In addition, we also assume that $\nu(\mathbb{R}) = \infty$, because otherwise X is a compound Poisson process and this case has already been addressed by (Chen et al. 2019). We say that any measurable function $h : (0,\infty) \mapsto (0,\infty)$ is regularly varying at 0 with index ρ if, for $\varphi(x) = h(1/x)$, we have $\varphi \in \mathrm{RV}_{-\rho}$.

Theorem 2 For a fixed $\gamma_0 > 0$ and a Lévy process $\{X(t) : t \geq 0\}$ with generating triplet $(0,0,\nu)$, suppose that we have some Borel measure μ such that

- $(\nu - \mu)|_{(-\gamma_0,\gamma_0)}$ is a positive measure;
- the function $f : (0,\infty) \mapsto (0,\infty)$ defined as $f(x) = \mu\big((-\infty,-x] \cup [x,\infty)\big)$ is regularly varying at 0 with index $-(\alpha+\varepsilon)$ where $\alpha \in (0,2), \varepsilon \in (0, 2-\alpha)$.

Then there exists some $C < \infty$ such that for any $\gamma \geq \gamma_0$

$$\big\| f_{X^{<\gamma}(t)} \big\|_\infty \leq \frac{C}{t^{1/\alpha} \wedge 1} \quad \forall t > 0$$

where $\{X^{<\gamma}(t) :\ t > 0\}$ is the Lévy process with generating triplet $\big(0,0,\nu|_{(-\gamma,\gamma)}\big)$ and $f_{X^{<\gamma}(t)}$ is the density of distribution of $X^{<\gamma}(t)$.

An immediate consequence is as follows. Define a function $g(x) = \nu\big((\infty,-x) \cup (x,\infty)\big)$. If g is regularly varying at 0 with index $\beta > 0$, then Assumption 2 holds, and the proposed algorithm is strongly efficient. Intuitively, since we are excluding the simpler cases where $\sigma > 0$ or $\nu(\mathbb{R}) < \infty$, we must have $\lim_{x \downarrow 0} g(x) = \infty$. As long as $g(\cdot)$ approaches ∞ at a faster rate than some $1/x^\beta$ with $\beta > 0$, Assumption 2 is valid.

4.3 Sketch of Proof for Theorem 1

By performing a change of measure and plugging in the exact value of $d\mathbb{Q}/d\mathbb{P}$ (see (3)):

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}}[L_n^2] &= \int Z_n^2(J_n)\mathbb{1}_E(J_n/n)\frac{d\mathbb{P}}{d\mathbb{Q}}\frac{d\mathbb{P}}{d\mathbb{Q}}d\mathbb{Q} = \int Z_n^2(J_n)\mathbb{1}_E(J_n/n)\frac{d\mathbb{P}}{d\mathbb{Q}}d\mathbb{P} \\
&= \int Z_n^2(J_n)\mathbb{1}_{E\cap B_n^\gamma}(J_n/n)\frac{d\mathbb{P}}{d\mathbb{Q}}d\mathbb{P} + \int Z_n^2(J_n)\mathbb{1}_{E\cap (B_n^\gamma)^c}(J_n/n)\frac{d\mathbb{P}}{d\mathbb{Q}}d\mathbb{P} \leq \frac{\mathbb{P}(B_n^\gamma)}{1-w}\mathbb{E}[Z_{n,1}^2] + \frac{1}{w}\mathbb{E}[Z_{n,2}^2],
\end{aligned}$$

where $Z_{n,1} = Z_n(J_n)\mathbb{1}_{E\cap B^\gamma}(J_n/n), Z_{n,2} = Z_n(J_n)\mathbb{1}_{E\cap (B^\gamma)^c}(J_n/n)$. Using Result 1, we have $\mathbb{P}(B_n^\gamma) = \mathscr{O}(\mathbb{P}(A_n))$ as both B^γ and A are bounded away from $\mathbb{D}_{<l^*}$. Then strong efficiency follows immediately once we have

$$\mathbb{E}Z_{n,1}^2 = \mathscr{O}(\mathbb{P}(A_n)); \tag{8}$$

$$\mathbb{E}Z_{n,2}^2 = \mathscr{O}(\mathbb{P}^2(A_n)). \tag{9}$$

Fix some notations: we use ζ denote a step function, and save the index k to indicate the number of large jumps. For instance, the event $\{J_n = \zeta_k\}$ is equivalent to the event that J_n has k jumps. Note that on this set, J_n admits the representation $J_n \stackrel{d}{=} \zeta_k = \sum_{i=1}^k z_i \mathbb{1}_{[u_i,n]}$ where $z_1,\cdots,z_k$ are i.i.d. samples from the distribution $\nu(\cdot \cap [n\gamma,\infty))/\nu[n\gamma,\infty)$, and $u_1 \leq u_2 \leq \cdots \leq u_k$ are order statistics of k i.i.d. Unif$(0,n)$. Now note that

$$\mathbb{E}Z_{n,1}^2 \leq \sum_{k \geq l^*} \mathbb{E}[Z_n^2(J_n) \mid J_n = \zeta_k]\mathbb{P}(J_n \text{ has k jumps}) = \sum_{k \geq l^*} \mathbb{E}[Z_n^2(\zeta_k)]e^{-\lambda_n}\lambda_n^k/k! \tag{10}$$

with $\lambda_n = n\nu[n\gamma,\infty)$. Therefore, to show (8), it suffices to show the existence of a constant C such that $\mathbb{E}Z_n^2(\zeta_k) \leq Ck$ for any $k = 1,2,\cdots$. To see this, by plugging this bound into R.H.S. of (10) we will get

$$\mathbb{E}Z_{n,1}^2 \leq C\sum_{k \geq l^*} ke^{-\lambda_n}\lambda_n^k/k! \leq C\lambda_n^{l^*}\sum_{k \geq l^*} e^{-\lambda_n}\frac{\lambda_n^{k-l^*}}{(k-l^*)!} = C\cdot(n\nu[n\gamma,\infty))^{l^*} \tag{11}$$

and (8) follows immediately from large deviation principles (Result 1) and the fact that ν is regularly varying. To bound $\mathbb{E}Z_n^2(\zeta_k)$, recall that Z_n is an unbiased estimator, so from Result 2 we have $\mathbb{E}Z_n^2(\zeta_k) \leq \sum_{m\geq 0}\mathbb{P}\Big(Y_{n,m}(\zeta_k) \neq Y_n^*(\zeta_k)\Big)\Big/\mathbb{P}(\tau \geq m)$, where $Y_{n,m}$ and Y_n^* are indicator functions defined in Section 3.2. Then it remains to bound $\mathbb{P}\Big(Y_{n,m}(\zeta_k) \neq Y_n^*(\zeta_k)\Big)$, the probability that the supreme of a (non-compound-Poisson) Lévy process crossed a certain barrier while its m step SBA estimation did not.

To illustrate the idea, we henceforth focus on a simplified scenario. Fix a constant $c > 0$ and recall that small-jump process $\widetilde{X}$ is a (non-compound-Poisson) Lévy process with bounded jumps and satisfies Assumption 2. Define $M = \sup_{0\leq t\leq 1}\tilde{X}(t)$. Using the coupling in (1), we have $M \stackrel{d}{=} \sum_{i\geq 1}(\xi_i)^+$ where $(\xi_i \stackrel{d}{=} \tilde{X}(l_i))_{i\geq 1}$ are independent conditioned on the stick length sequence $(l_i)_{i\geq 1}$ with $\sum_i l_i = 1$. Furthermore, we use $Y^* = \mathbb{1}\{M \geq c\}$ to indicate whether the supreme of $\tilde{X}$ on [0,1] exceeds c, while $Y_m = \mathbb{1}\{M_m \triangleq \sum_{i=1}^m (\xi_i)^+ \geq c\}$ as its counterpart for the $m-$step SBA estimation. To bound $\mathbb{P}(Y^* \neq Y_m)$, fix some $\varepsilon \in (0,c)$ and notice that if $\{Y^* \neq Y_m\}$ occurs, then so does at least one of the following two events: (a) $|M - M_m| \geq \varepsilon$; (b) $M \in [c-\varepsilon, c+\varepsilon]$. For the former, the geometric convergence rate of SBA gives an effective bound (in particular, see Lemma 10 in (Cázares et al. 2018)). Now the proof hinges on bounding the latter, which boils down to analyzing probability of the form $\mathbb{P}(M_j \in [c, c+\delta])$ for $c, \delta > 0$.

To this end, recall that $M_j \stackrel{d}{=} (\xi_1)^+ * (\xi_2)^+ * \cdots * (\xi_j)^+$ where $*$ is the convolution operator, and observe the following facts. First, convolution operation preserves the smoothness of any distribution involved; for instance $X * Y$ is continuous if either X or Y is continuous, and $X * Y$ has a density bounded by a constant K if so does either X or Y. Besides, given $(l_i)_{i\geq 1}$, the law of $\xi_i \stackrel{d}{=} \tilde{X}(l_i)$ satisfies Assumption 2, and as long as $\xi_i > 0$, we will have $\xi_i = (\xi_i)^+$ so the same smoothness property can be passed to $(\xi_i)^+$, hence M_j. To apply these facts, fix some $h > 0$ as a threshold value and note that: if there exists $i = 1, \cdots, j$ such that $l_i > 0$ and $\xi_i > 0$, then by further conditioning on this event we can use Assumption 2 to provide a bound with $t \geq h$; otherwise, the supreme M is equal to sum of increments only on sticks shorter than h, the total length of which is less than jh. If h (and hence jh) is indeed a small value, then it is unlikely that Lévy process $\tilde{X}$ reached the barrier c within such a short period of time jh (again, see Lemma 10 in (Cázares et al. 2018)), let along crossing the barrier c and staying in $[c, c+\delta]$. By carefully choosing j and h, we can establish a useful upper bound for $\mathbb{P}(M \in [c, c+\delta])$, and eventually for $\mathbb{E}Z_n^2(\zeta_k)$.

The argument for (9) will be analogous, except that we need to notice the following fact: since $Y_n^* \geq Y_{n,m}$, for $Y_n^* \neq Y_{n,m}$ to occur we need to at least ensure $Y_n^* = 1$, which is equivalent to the condition $\{\bar{X}_n \in A\}$. Combining this with Hölder's inequality when using Result 2 and the bound $\mathbb{E}Z_n^2(\zeta_k) \leq Ck$ above, we will have $\mathbb{E}Z_{n,2}^2 \leq C_1\sqrt{\mathbb{P}(\bar{X}_n \in A \cap (B^\gamma)^c)}$ where $C_1 < \infty$ is some constant. Thus, by picking γ small enough and invoking large deviation principles (Result 1) again, we then have (9) and conclude the proof.

5 SIMULATION EXPERIMENTS

In this section, we apply the proposed importance sampling strategy in Algorithm 1 to the following setting and use numerical experiments to demonstrate: (1) the performance of the importance sampling estimator over a range of different scaling factor n for a number of tail distributions; (2) the efficiency of the algorithm when compared to crude Monte-Carlo methods.

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 $\lambda = 0.1$, and $\{W_i\}_{i\geq 1}$ is a sequence of i.i.d. samples from Pareto distribution with $\mathbb{P}(W_1 > x) = 1/\max\{x,1\}^\alpha$ where the tail index $\alpha > 1$. For each $n \geq 1$, define the scaled process $X_n(t) = \frac{X(nt)}{n}$, and we are interested in the probability of the event $A_n = \{X_n \in A\}$ where

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

with $a = 2, b = 1.15$. As stressed above, we aim to showcase the performance of the importance sampling estimator under different n and α. Specifically, in our experiments we use $\alpha = 1.45, 1.6, 1.75$, and

$n = 1000, 2000, \cdots, 10000$. To quantify the efficiency of an estimator, we report the *relative error*: the ratio between the standard deviation estimated by all samples of the estimator and the estimated mean.

In implementing the importance sampling estimator, we used $\gamma = 0.2, w = 0.05, \rho = 0.95$ (and note that $l^* = 2$ in this case). For each $\alpha \in \{1.45, 1.6, 1.75\}$ and $n \in \{1000, 2000, \cdots, 10000\}$, we generated 500,000 independent samples. To compare the efficiency of the proposed importance sampling estimator against the crude Monte-Carlo methods, we generated at least $64/\hat{p}_{\alpha,n}$ crude Monte-Carlo samples where $\hat{p}_{\alpha,n}$ is the estimated value of $\mathbb{P}(A_n)$ using Algorithm 1 as described above.

The results of the experiment 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 estimator stays at a constant level regardless of how large n is. This is as expected in view of the strong efficiency of the estimator established in Theorem 1. Therefore, if the goal is to achieve a certain level of standard error, the number of samples required for Algorithm 1 is bounded and does not increase with n.

Figure 1 plots the relative errors of the importance sampling estimators and the crude Monte Carlo estimators. This illustrates the benefit of the proposed importance sampling strategy. For crude Monte-Carlo scheme, the relative error grows polynomialy with n (to be more precise, roughly $\mathscr{O}(n^{l^*(\alpha-1)})$). In contrast, the relative error of the Algorithm 1 is nearly constant. Note that the expected cost to generate a single sample is $\mathscr{O}(n)$ (in terms of the expected number of jumps required to simulate) for both methods. The proposed importance sampling method always outperforms crude Monte-Carlo scheme for sufficiently rare events A_n (i.e., large n), and the difference in the performance grows as n grows.

6 CONCLUSIONS

We proposed a strongly efficient importance sampling algorithm for rare-event simulation of Lévy processes with heavy-tailed jump distributions and infinite activities, where the events are triggered by multiple jumps. The numerical experiments confirm the strong efficiency of the proposed algorithm, which outperforms crude Monte-Carlo method by orders of magnitude as the event of interest gets rarer. In our future works, we aim to extend the current framework to a more general class of rare events and stochastic processes.

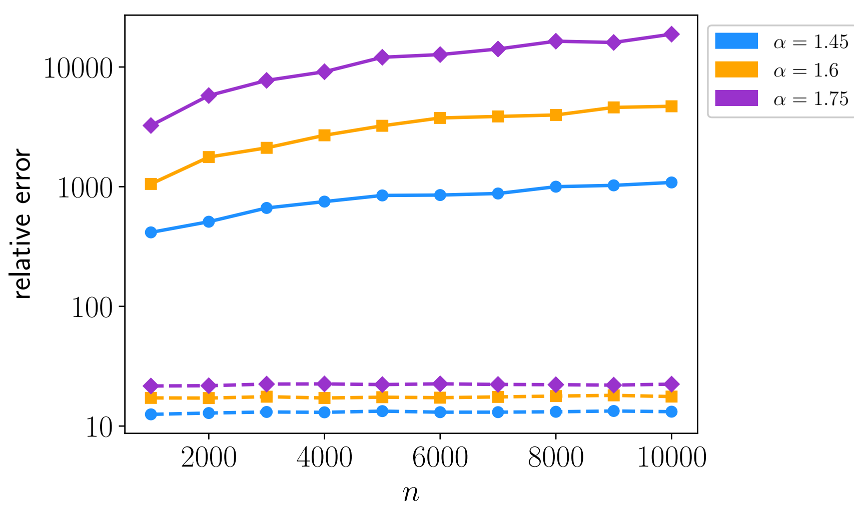

Figure 1: Comparison of relative errors between the proposed importance sampling estimator and crude Monte Carlo estimator. Solid lines: Crude Monte-Carlo estimator; Dashed lines: Importance-sampling estimator.

Table 1: Rare-event simulation results using Algorithm 1. First row: estimated probability of $\mathbb{P}(A_n)$; Second row: the relative error.

n	2000	4000	6000	8000	10000
$\alpha = 1.45$	3.53×10^{-6}	1.85×10^{-6}	1.28×10^{-6}	9.76×10^{-7}	7.96×10^{-7}
	12.84	13.02	13.06	13.16	13.19
$\alpha = 1.6$	3.34×10^{-7}	1.45×10^{-7}	8.84×10^{-8}	5.89×10^{-8}	4.60×10^{-8}
	17.13	17.16	17.26	17.80	17.63
$\alpha = 1.75$	3.46×10^{-8}	1.14×10^{-8}	6.21×10^{-9}	4.17×10^{-9}	2.92×10^{-9}
	21.74	22.50	22.53	22.16	22.40

REFERENCES

Asmussen, S., and H. Albrecher. 2010. *Ruin Probabilities*. Singapore: World Scientific.

Boxma, O., E. Cahen, D. Koops, and M. Mandjes. 2019. "Linear Stochastic Fluid Networks: Rare-Event Simulation and Markov Modulation". *Methodology and Computing in Applied Probability* 21(1):125–153.

Cázares, J. G., A. Mijatović, and G. U. Bravo. 2018. "Geometrically Convergent Simulation of the Extrema of Lévy Processes". *arXiv preprint arXiv:1810.11039*. https://arxiv.org/abs/1810.11039.

Cázares, J. I. G., A. Mijatović, and G. U. Bravo. 2019. "Exact Simulation of the Extrema of Stable Processes". *Advances in Applied Probability* 51(4):967–993.

Chaumont, L., and J. Małecki. 2018. "Short Proofs in Extrema of Spectrally One-Sided Lévy Processes". *Electronic Communications in Probability* 23.

Chen, B., J. Blanchet, C.-H. Rhee, and B. Zwart. 2019. "Efficient Rare-Event Simulation for Multiple Jump Events in Regularly Varying Random Walks and Compound Poisson Processes". *Mathematics of Operations Research*.

Debicki, K., and M. Mandjes. 2015. *Queues and Lévy Fluctuation Theory*. Springer.

Dupuis, P., and H. Wang. 2004. "Importance Sampling, Large Deviations, and Differential Games". *Stochastics: An International Journal of Probability and Stochastic Processes* 76(6):481–508.

Michna, Z., Z. Palmowski, and M. Pistorius. 2015. "The Distribution of the Supremum for Spectrally Asymmetric Lévy Processes". *Electronic Communications in Probability* 20.

Pitman, J., and G. U. Bravo. 2012. "The Convex Minorant of a Lévy Process". *The Annals of Probability* 40(4):1636–1674.

Resnick, S. I. 2007. *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Science and Business Media.

Rhee, C.-H., J. Blanchet, B. Zwart 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, C.-H., and P. W. Glynn. 2015. "Unbiased Estimation with Square Root Convergence for SDE Models". *Operations Research* 63(5):1026–1043.

Sato, K.-i., S. Ken-Iti, and A. Katok. 1999. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge university press.

Tankov, P. 2003. *Financial Modelling with Jump Processes*. Chapman and Hall/CRC.

Wang, X., and C.-H. Rhee. 2020. "Efficient Rare-Event Simulation for Multiple Jump Events in Regularly Varying Lévy Processes with Infinite Activities". *Working Paper. (arXiv:2007.08080)*. https://arxiv.org/abs/2007.08080.

AUTHOR BIOGRAPHIES

XINGYU WANG is a PhD candidate in the Department of Industrial Engineering and Management Sciences at Northwestern University. He received his B.S. degree in Psychology and Applied Mathematics from Peking University, China, and the M.S. degree in Analytics from Northwestern University. His research interests include stochastic simulation and applied probability. His e-mail address is xingyuwang2017@u.northwestern.edu.

CHANG-HAN RHEE is an Assistant Professor in Industrial Engineering and Management Sciences at Northwestern University. He received his Ph.D. in Computational and Mathematical Engineering from Stanford University. His research interests include stochastic simulation and applied probability. His e-mail address is chang-han.rhee@northwestern.edu, and his website is https://chrhee.github.io.