Strongly Efficient Rare-Event Simulation for Multiple-Jump Events in Regularly Varying Lévy Processes with Infinite Activities

Xingyu Wang, Chang-Han Rhee

September 25, 2023

Abstract

In this paper, we address rare-event simulation for heavy-tailed Lévy processes with infinite activities. Specifically, the presence of infinite activities poses a significant computational challenge, making it impractical to simulate or store the entire sample path. Building upon the importance sampling scheme in [14], we present a rare-event simulation algorithm that incorporates the sample path large deviations for heavy-tailed Lévy processes, the stick-breaking approximation for the extrema of Lévy processes, the Asmussen-Rosiński approximation for small-jump Lévy processes, and the randomized debiasing Monte Carlo scheme. Through a novel characterization for the Lipschitz continuity of the law of Lévy processes, we establish that the proposed algorithm is unbiased and strongly efficient under mild conditions, and hence applicable to a broad class of Lévy processes. In the numerical experiments, our algorithm demonstrates significant improvements in efficiency when compared to crude Monte-Carlo method.

1 Introduction

In this paper, we propose an importance sampling scheme designed to efficiently estimate the probability of rare events in heavy-tailed Lévy processes with infinite activities. In particular, the heavy-tailedness in the increments of the Lévy processes is captured by the notion of regular variation (see Definition 1). The prevalence of heavy-tailed phenomena extends across a diverse range of stochastic dynamics and systems, manifesting in crucial areas such as the spread of COVID-19 (see, e.g., [15]), traffic in computer and communication networks (see, e.g., [40, 26]), financial assets and risk processes (see, e.g., [24, 10]), and the training of deep neural networks (see, e.g., [33]). The estimation of the probability of rare events in heavy-tailed Lévy processes also holds considerable significance in many applications, including risk management [1], mathematical finances [50], and queueing systems [37].

More formally speaking, the objective of this paper is to estimate the probability of form $\mathbf{P}(A_n)$ where $A_n = \{\bar{X}_n \in A\}$, and $\bar{X}_n = \{\bar{X}_n(t) = \frac{1}{n}X(nt) :\ t \in [0,1]\}$ is a scaled version of some regularly varying Lévy process X. Here, A is subset of $\mathbb{D}$, the space of the real-valued RCLL functions over $[0,1]$. Two key features characterize the problem setup. First, the occurrence of the rare events A_n typically necessitates multiple large jumps in $\bar{X}_n$. To facilitate exposition, we focus on the scenario with

$$A = \Big\{\xi \in \mathbb{D} : \sup_{t\in[0,1]} \xi(t) \geq a;\ \sup_{t\in(0,1]} \xi(t) - \xi(t-) < b\Big\}. \tag{1.1}$$

Intuitively speaking, this characterization indicates that the supremum of the path ξ over $[0,1]$ exceeds the threshold a, even though no upward jump in ξ is larger than b. One can easily see that, for a step function (initialized at the origin) to fall into set A, the minimum number of jumps required is dictated by the ratio between a and b. Nevertheless, it is worth noticing that the framework developed in this paper can be effortlessly adapted to other multiple-jump rare events. Second, the regularly varying Lévy process X also exhibits infinite activities; see Section 2.4 for the rigorous definition. In simple terms, the existence of either Brownian motion components or infinitely many jumps in X makes it computationally infeasible to simulate or store the entire sample path of X within any given interval.

We discuss the three main challenges that arise in the current setup. Firstly, the nature of the rare events makes the crude Monte-Carlo simulation method extremely inefficient. For instance, when estimating $p_n = \mathbf{P}(\bar{X}_n \in A)$, the crude Monte-Carlo estimator $\mathbb{1}\{\bar{X}_n \in A\}$ attains a standard error of order $\sqrt{p_n}$. Consequently, the number of samples required to estimate the target probability to a given relative accuracy is of order $\sqrt{1/p_n}$. As $n \to \infty$, the crude Monte-Carlo method becomes prohibitively expensive due to $p_n \to 0$. To resolve this issue, we aim to design estimators L_n for probabilities $\mathbf{P}(\bar{X}_n \in A)$ that attains strong efficiency (see Definition 2). Such strongly efficient estimators maintain a uniformly bounded (w.r.t. n) relative error, thus maintaining the same level of efficiency (regarding the required number of simulation trials) regardless of the rarity of the target events.

To achieve this goal, we employ importance sampling, a frequently used variance reduction technique in Monte-Carlo simulation; see [2] for a standard treatment on this topic. The essence of importance sampling lies in the use of an alternative sampling distribution $\mathbf{Q}$ (instead of the nominal distribution $\mathbf{P}$), under which the rare event occurs more frequently. Then we correct the estimation by incorporating the likelihood ratio $d\mathbf{P}/d\mathbf{Q}$. Nevertheless, the performance of importance sampling technique is particularly sensitive to the choice of the importance sampling distribution $\mathbf{Q}$. Seemingly plausible yet theoretically unjustified choices of $\mathbf{Q}$ often fail to reduce the variance or even result in infinite variance during estimation; see, e.g., [29, 28]. Therefore, principled approaches are required to achieve theoretical guarantee in the designing of importance sampling estimators.

In light-tailed settings, importance sampling schemes guided by large deviation theories have proven to be highly effective. Typically, the law of the light-tailed system, when conditioning on the rare events, would (asymptotically) coincide with a modulated version of the dynamics where the law of all the increments is exponentially tilted. An importance sampling distribution is then crafted by applying the exponential change of measure to the increments (see, for instance, [11, 51]). Furthermore, in the context of queueing networks, [19] unveiled and capitalized the connections between importance sampling, large deviations, and differential games. This leads to an adaptive and state-dependent importance sampling algorithm, which is asymptotically optimal for rare-event simulation in queueing networks. Interested readers can find theoretical foundations and additional algorithmic developments for this approach in references such as [20, 21, 22].

In comparison, the conventional exponential tilting approach falls short in providing a principled and provably efficient design of the importance sampling estimators in heavy-tailed settings, as noted in references such as [5]. Variance reduction techniques, such as conditional Monte Carlo (e.g., [3, 36]) and Markov Chain Monte Carlo (e.g., [32]), have proven to be valuable when addressing certain types of heavy-tailed rare event simulation tasks. Moreover, different approaches have been explored to apply importance sampling in heavy-tailed systems. For example, [18] proposes a state-dependent importance sampling algorithm for the boundary-crossing probability of a random walk with regularly varying increments by progressively modifying the probability of observing a large jump at each step; in [6], the estimation of the first-passage probability of heavy-tailed random walks is carried out using a state-dependent importance sampling algorithm based on Doob's h-transform and Lyapunov inequalities (see also [7, 8, 9] for applications of the same technique in multidimensional and queueing contexts). Nevertheless, it is worth noticing that the previously mentioned works are often tailored for specific types of processes or specific rare events, and their generalizations (such as the identification of Lyapunov functions) can be highly non-trivial beyond the simple settings.

Fortunately, the recent developments of heavy-tailed large deviations, as exemplified by [47] and [53], have laid the foundation of an efficient and universal importance sampling estimators for heavy-tailed systems. At the core of theory is the catastrophe principle that characterizes the rarity and identifies the most likely causes for rare events in heavy-tailed systems. In heavy-tailed dynamics, rare event typically arise due to the catastrophic failures of a few components in the system. The exact number of such catastrophic failures dictates the asymptotic rate of decay and the most likely causes of the rare events, thus establishing a discrete hierarchy for rare events in heavy-tailed systems. More specifically, the results in [47] show that the asymptotics of form $\mathbf{P}(\bar{X}_n \in E) = \boldsymbol{O}\big((n\nu[n,\infty))^{l^*(E)}\big)$ (as $n \to \infty$) hold for all sets $E \subset \mathbb{D}$ satisfying a mild topological condition. Here, ν is the Lévy measure of the Lévy process X and $l^*(E)$ is the minimum number of jumps required for a step function to fall into set E. Consequently, for a given event $B \subseteq \mathbb{D}$ such that $l^*(B) = l^*(A)$ (recall that the target rare events $A_n = \{\bar{X}_n \in A\}$ are defined based on some $A \subseteq \mathbb{D}$), we can expect $\mathbf{P}(\bar{X}_n \in A)$ and $\mathbf{P}(\bar{X}_n \in B)$ to decay at a similar rate as $n \to \infty$, suggesting that the event $\{\bar{X}_n \in B\}$ can provide a reliable approximation to the target event $\{\bar{X}_n \in A\}$ when designing the importance sampling distribution. Building upon these ideas and employing defensive importance sampling with a carefully chosen set B, [14] has demonstrated a readily implementable and strongly efficient importance sampling scheme applicable to wide class of rare events in regularly varying random walks and compound Poisson processes. In this work, we adopt and extend this framework to encompass Lévy processes with infinite activities. The specifics of the set B (and hence the importance sampling distribution) are detailed in Section 3.1. Naturally, a crucial aspect in the implementation of the algorithm is the sampling from $\mathbf{P}(\ \cdot\ |\bar{X}_n \in B)$. This is addressed by Algorithm 1 in Section 3.4.

The second challenge lies in the simulation of the Lévy process X with infinite activities, which sets our work apart from existing ones such as [14] and introduces new technical obstacles in the design of the algorithm. As mentioned earlier, the presence of infinite activities makes it computationally infeasible to simulate or store the entire sample path of X. Our algorithm proposes a potential solution by first identifying all upward jumps with sizes exceeding a threshold $\gamma \in (0, b)$ and then simulating the supremum of the process X between the arrival times of these large jumps. Note that this still requires the simulation the suprema of Lévy processes with infinite activities. However, the law of the suprema of Lévy processes is generally unknown (see [43]). Besides, methods for the exact (or even ϵ-exact) simulation of these suprema or the value of the process at the first passage of a given threshold are only available for a few special cases; see, e.g., [31], [17], and [12]. Another direction that has been explored in existing literature utilizes the Winener–Hopf factorization in the fluctuation theory of Lévy processes; see [38, 25]. The drawback, however, is that this approach requires the capability of simulating $X(\boldsymbol{e}_\lambda)$, where $\boldsymbol{e}_\lambda$ is an exponential random variable with rate λ that is independent of the Lévy process X (i.e., simulating the Lévy process at a random exponential time). Unfortunately, an algorithm for the simulation $X(\boldsymbol{e}_\lambda)$ is also not available for most Lévy processes.

To overcome this challenge, we first construct a series of progressively more accurate approximations to the suprema of Lévy processes, and then remove the bias in the approximations. Specifically, we employ stick-breaking approximations (SBA) algorithms in [30] for the extrema of Lévy processes with infinite activities. This algorithm is built upon the theory of the convex minorants and concave majorants of Lévy processes in [45], which will be reviewed in Section 2.4. In simple terms, given a Lévy process $X(t)$ and its running suprema $M(t) = \sup_{s\in[0,t]} X(s)$, the joint law of $(X(t), M(t))$ admits the representation

$$\big(X(t), M(t)\big) \stackrel{d}{=} \Big(\sum_{j\geq 1} \xi_j,\ \sum_{j\geq 1} \max\{\xi_j, 0\}\Big). \tag{1.2}$$

Here, $(l_j)_{j\geq 1}$ is a sequence of non-negative RVs satisfying $\sum_{j\geq 1} l_j = 1$ and $\mathbf{E}l_j = t/2^j\ \forall j \geq 1$, and $\xi_j \stackrel{d}{=} X(l_j)$ are independently generated when conditioning on the values of $(l_j)_{j\geq 1}$. While it is computationally infeasible to generate the entire sequence of $(l_j)_{j\geq 1}$ or $(\xi_j)_{j\geq 1}$, by simulating l_j, ξ_j up to step m we obtain $\big(\sum_{j=1}^m \xi_j, \sum_{j=1}^m \max\{\xi_j, 0\}\big)$ as an approximation to $(X(1), M(1))$. In particular, the geometric rate of decay in $\mathbf{E}l_j$ ensures that, in expectation, the error in our m-step approximation decays exponentially fast as $m \to \infty$. Such an approximation naturally lend itself to the unbiased estimation technique in [48], which allows us to construct an unbiased estimator from a sequence of approximations that becomes progressively more accurate. Further details on the unbiased estimation technique are available in Section 2.3. This unbiased estimation technique can also be interpreted as a randomized version of the multilevel Monte Carlo scheme [34, 27]. For more information on the connection between unbiased estimators and multilevel Monte Carlo, refer to [52]. Combining the unbiased estimator with SBA, we are able to design an estimation algorithm that is unbiased for $\mathbf{P}(A_n)$ and terminates after generating finitely many l_j, ξ_j.

In practice, implementing the unbiased estimator using the stick-break approximation based on representation (1.2) requires the ability to sample $X(t)$ for arbitrary $t > 0$, which is generally not feasible for Lévy processes with infinite activities. Therefore, we appeal to the classical Asmussen-Rosiński approximation (ARA) in [4]. This approximation involves replacing the small jump martingale in the Lévy process X with a Brownian motion term of the same variance. Thus, we approximate X with a mixture of a Brownian motion term and a compound Poisson process (with drift). To maintain the unbiasedness of our estimator and remove the errors introduced in ARA, we also incorporate ARA into the unbiased estimation technique. The construction of these approximators will be detailed in Section 3.3. It is worth noticing that the combination of unbiased estimation and stick-breaking approximations has been mentioned in section 2.5 of [30]. However, our work is the first to carefully orchestrate SBA, ARA, and the unbiased estimation technique and we demonstrate the effectiveness of this combination in the contexts of rare-event simulation. Our efforts culminates in Theorem 3.4, which establishes the unbiasedness and strong efficiency of our proposed Algorithm 2.

Finally, the third major challenge pertains to the continuity of the law of the running supremum $M(t) = \sup_{s\in[0,t]} X(s)$. For the set A defined in (1.1), the indicator function $\mathbb{1}_A(\xi)$ is discontinuous at any $\xi \in \mathbb{D}$ with $\sup_{t\in[0,1]} \xi(t) = a$. As a result, when analyzing the estimation error of $\mathbf{P}(\bar{X}_n \in A)$ one must evaluate the probability that the supremum of $X(t)$, denoted by $M(t) = \sup_{s\in[0,t]} X(s)$, concentrates around na. Furthermore, to achieve strong efficiency we need to establish an upper bound on the relative error that is uniform w.r.t. all n. From the perspective of Hölder continuity, this roughly translates into such conditions: there exist some $C \in (0,\infty)$ and $\theta \in (0,1]$ such that $\mathbf{P}(M(n) \in [na-\delta, na+\delta]) < C\delta^\theta$ $\forall \delta \in (0,1]$ holds uniformly for all n large enough. 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 focusing on bounding the cumulative distribution function (cdf); see, e.g., [13, 16, 39, 41, 44, 42]. Most of these work do not provide quantitative characterization of the continuity, nor do they establish a version of Hölder continuity that is uniform with respect to time t or location x in $\mathbf{P}(M(t) \in [x-\delta, x+\delta])$. In short, tackling this aspect of the challenge requires establishing novel and useful quantitative characterizations of the law of supremum $M(t)$.

Again, the stick-breaking representation in (1.2) proves to be a valuable tool when analyzing the continuity of $M(t)$. Our strategy involves passing the continuity of $X(t)$ onto $M(t)$ through the convolution structure of $M(t) \stackrel{d}{=} \sum_{j\geq 1} \max\{X_j(l_j), 0\}$, where X_j are iid copies of the process X. In particular, it turns out that a sufficient condition for our rare-event simulation setup is

$$\mathbf{P}\Big(X^{<z}(t) \in [x, x+\delta]\Big) \leq \frac{C}{t^\lambda \wedge 1}\delta^\theta \qquad \forall z \geq z_0,\ t > 0,\ x \in \mathbb{R},\ \delta \in [0,1]. \tag{1.3}$$

Here, $X^{<z}$ is a modulated version of the process X where all the upward jumps with sizes larger than z are removed. Rigorous definitions can be found in Section 3. At first glance, the condition (1.3) may seem restrictive, as it claims a version of Hölder continuity that holds uniformly in t, x, and the truncation threshold z. Fortunately, in Section 4 we establish several sets of sufficient conditions for (1.3) that are easy to verify. The general idea is to examine the characteristic functions of the Lévy measure ν and prove its close resemblance to some α-stable or semi-stable processes that required continuity properties. In particular, we show that (1.3) is a mild condition for Lévy processes with infinitive activities, as it only requires the mass of ν to approach ∞ (hence attaining infinite activities in X) at a pace not too slow. This observation highlights the wide applicability of our approach to Lévy processes with infinite activities.

The paper is structured as follows. Section 2 provides a review of the theoretical foundations of our algorithms, including the heavy-tailed large deviation theories (Section 2.2), the debiasing technique

(Section 2.3), and the stick-breaking approximations (Section 2.4). Section 3 presents the importance sampling algorithm and establishes its strong efficiency. Section 4 investigates the continuity of the law of $X(t)$ and provides sufficient conditions for (1.3), a critical assumption behind our strongly efficient importance sampling scheme. Numerical experiments are reported in Section 5. The proofs of our technical results are collected in Section 6.

2 Preliminaries

In this section we first introduce a series of notations that will be frequently used in the remainder of the paper, and then review several important results as the building blocks for our strongly efficient rare-event simulation algorithm.

2.1 Notations

For any positive integer k, let $[k] = \{1, 2, \cdots, k\}$. 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 to get rounded-down or rounded-up integer values from x. Given a measure space $(\mathcal{X}, \mathcal{F}, \mu)$ and any $A \in \mathcal{F}$, we use $\mu|_A$ to denote restriction of the measure μ on A, which is defined as

$$\mu|_A(\cdot) \triangleq \mu(A \cap \cdot).$$

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 $\mathbb{D} = \mathbb{D}_{[0,1],\mathbb{R}}$, the space of all real-valued RCLL functions with domain $[0, 1]$, equipped with Skorokhod J_1 metric $\boldsymbol{d}$. Here the metric is defined by

$$\boldsymbol{d}(x, y) \triangleq \inf_{\lambda \in \Lambda} \sup_{t \in [0,1]} |\lambda(t) - t| \vee |x(\lambda(t)) - y(t)| \tag{2.1}$$

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

We introduce a few definitions. First, we start with the concept of regular variation, which is commonly used to capture the heavy-tailed phenomena.

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 $\phi \in \mathcal{RV}_\beta$) 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 $\phi(\eta) \in \mathcal{RV}_\beta(\eta)$.*

For details on the definition and properties of regularly varying functions, see, for example, chapter 2 of [46].

Next, we discuss the Lévy-Ito decomposition of one-dimensional Lévy processes. The law of any one-dimensional Lévy process $\{X(t) : t \geq 0\}$ can be 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, we have

$$X(t) \stackrel{d}{=} ct + \sigma B(t) + \int_{|x|\leq 1} x[N([0, t] \times dx) - t\nu(dx)] + \int_{|x|>1} xN([0, t] \times dx) \tag{2.2}$$

where 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 independent of B with intensity measure $\text{Leb}([0, \infty)) \times \nu$. For standard references on this topic, see chapter 4 of [49].

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

Definition 2. *Given a probability space $(\Omega, \mathcal{F}, \mathbf{P})$, let A_n be a sequence of events and L_n be a sequence of random variables. We say that $(L_n)_{n\geq 1}$ are **unbiased and strongly efficient** estimators of $(A_n)_{n\geq 1}$ if*

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

We stress that strongly efficient estimators $(L_n)_{n\geq 1}$ can achieve uniformly bounded relative errors for all $n \geq 1$. In other words, the number of Monte-Carlo simulation runs required to estimate the target probability $\mathbf{P}(A_n)$ to a given relative accuracy is uniformly bounded w.r.t. all n.

2.2 Sample-Path Large Deviations of Regularly Varying Lévy Processes

The main ingredient of the importance sampling distribution in our algorithm is the recent development of the sample-path large deviations of Lévy processes with regularly varying increments in [47]. In Section 2.2 we first familiarize the readers with the one-sided version of the results and then review the more complicated two-sided version. First, consider a Lévy process X that is centered (i.e., $\mathbf{E}X(t) = 0$) and has generating triplet (c, σ, ν) such that the Lévy measure ν is supported on $(0,\infty)$. In other words, all the jumps (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 $\alpha > 1$. We define a scaled version of X 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 $\mathbb{D}$.

To describe the sample-path large deviation of $\bar{X}_n$ in this one-sided setting, we introduce a few notations. 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 and vanishes at the origin. Let $\mathbb{D}_0 = \{\mathbf{0}\}$ where $\mathbf{0}(t) \equiv 0$ is the zero function. Let $\mathbb{D}_{<l} = \cup_{j=0,1,\cdots,l-1}\mathbb{D}_l$. 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 on $\{y \in (0,\infty)^l : \ y_1 \geq y_2 \geq \cdots \geq y_l\}$, and define the measure

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

where $(U_j)_{j\geq 1}$ is an iid sequence of Unif$(0,1)$; while for $l = 0$, let $\mathbf{C}_\beta^0$ be the Dirac measure on $\mathbf{0}$. The following large deviation results will be useful when designing rare-event simulation algorithms for $\bar{X}_n$. Throughout the rest of this paper, all measurable sets are understood to be Borel measurable.

Result 1 (Theorem 3.1 of [47])**.** *Let $A \subset \mathbb{D}$ be measurable. Let $l^* \triangleq \min\{l \in \mathbb{N} : \ \mathbb{D}_l \cap A \neq \emptyset\}$. If A is bounded away from $\mathbb{D}_{<l^*}$ in the sense that $\boldsymbol{d}(A, \mathbb{D}_{<l^*}) > 0$, then*

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

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

A similar large deviation result, albeit admitting a slightly more sophisticated form, can be developed for the two-sided cases where the Lévy process $X(t)$ has both positive and negative jumps. Now let $X(t)$ be a centered Lévy process such that for $H_+(x) = \nu[x,\infty)$ and $H_-(x) = \nu(-\infty,-x]$, we have $H_+ \in \mathcal{RV}_{-\alpha}$ and $H_-(x) \in \mathcal{RV}_{-\alpha'}$ 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 $\mathbb{I}_{<j,k} \triangleq \big\{(l,m) \in \mathbb{N}^2\backslash\{(j,k)\} : \ l(\alpha-1)+m(\alpha'-1) \leq j(\alpha-1)+k(\alpha'-1)\big\}$. Next, let $\mathbf{C}_{0,0}$ be the Dirac measure on $\mathbf{0}$, and for any $(j,k) \in \mathbb{N}^2\backslash\{(0,0)\}$ let

$$\mathbf{C}_{j,k}(\cdot) \triangleq \mathbf{E}\Bigg[\nu_\alpha^j \times \nu_{\alpha'}^k\Big\{(x,y) \in (0,\infty)^j \times (0,\infty)^k : \ \sum_{l=1}^{j} x_l \mathbb{1}_{[U_l,1]} - \sum_{m=1}^{k} y_m \mathbb{1}_{[V_m,1]} \in \cdot\Big\}\Bigg] \tag{2.3}$$

where U_l, V_m are two independent sequences of iid Unif$(0,1)$ RVs. Now we are ready to state the two-sided result.

Result 2 (Theorem 3.4 of [47])**.** *Let $A \subset \mathbb{D}$ be measurable. Let*

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

If A is bounded away from $\mathbb{D}_{<\mathcal{J}(A),\mathcal{K}(A)}$, then

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

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

2.3 Unbiased Estimators

As will be demonstrated soon, when designing the strongly efficient rare-event simulation algorithm we will first identify a series of fine approximations to the Lévy process X with infinite activities, and then remove the errors in the approximations so that the proposed estimator is indeed unbiased. To achieve unbiasedness in our algorithm, we apply the the debiasing techniques in [48]. In particular, due to τ being finite (almost surely) in Result 3 below, it is possible to simulate the estimator Z as it only requires the algorithm to generate $Y_0, Y_1, \cdots, Y_\tau$ instead of the infinite sequence $(Y_n)_{n\geq 0}$.

Result 3 (Theorem 1 in [48])**.** *Given a random variable Y, a sequence of random variables $(Y_n)_{n\geq 0}$ such that $\lim_{m\to\infty}\mathbf{E}Y_m = \mathbf{E}Y$, and a positive integer-valued random variable τ with unbounded support such that τ is independent of $(Y_m)_{m\geq 0}$ and Y, if*

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

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

2.4 Stick-Breaking Approximations for Lévy Processes

Next, we review the distribution of the concave majorant of a Lévy process with infinite activities [45], which paves the way to the stick-breaking approximation algorithm proposed in [30] for the joint law of the endpoint and maximum value of a Lévy process. For any Lévy process $X(t)$ 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 [45] provide an intriguing representation of the joint law of $X(t)$ and $M(t)$. Specifically, fix some $T > 0$. Let V_i be a sequence of iid Unif$(0,1)$ RVs, and recursively define

$$l_1 = TV_1, \qquad l_i = V_i(T - l_1 - \cdots - l_{i-1}) \ \forall i \geq 2. \tag{2.4}$$

Conditioning on the value of $(l_i)_{i\geq 1}$, let ξ_i be a random copy of $X(l_i)$ and all ξ_i be independently generated.

Result 4 (Theorem 1 in [45])**.** *If Lévy process X has infinite activities, then (with $(x)^+ = \max\{x, 0\}$)*

$$\big(X(T), M(T)\big) \stackrel{d}{=} \big(\sum_{i\geq 1} \xi_i, \sum_{i\geq 1} (\xi_i)^+\big). \tag{2.5}$$

By generating finitely many ξ_i instead of the entire sequence $(\xi_i)_{i\geq 1}$, we obtain the stick-breaking approximation to $X(T), M(T)$; see [30] for details. In particular, the geometric convergence rate of the approximation comes from the fact that $T - \mathbf{E}[l_1 + \cdots + l_n] = T/2^n$, that is, the part that has not been simulated up until step n is decaying geometrically fast in expectation. This stick-breaking approximation will be at the core of our algorithm when approximating the running supremum of a Lévy process with infinite activities. In particular, we utilize a coupling between multiple Lévy processes that is based on the stick-breaking representation above. For clarity of our description, here we focus on a case where we have 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. One can first generate l_i as described in (2.4). Conditioning on the values of $(l_i)_{i\geq 1}$, we can then independently generate ξ_i and $\widetilde{\xi}_i$ where ξ_i is a random copy of $X(l_i)$ and $\widetilde{\xi}_i$ is a random copy of $\widetilde{X}(l_i)$. Let $\widetilde{M}(t) \triangleq \sup_{s\leq t} \widetilde{X}(s)$. It then follows from Result 4 that we have identified a coupling between $X(T), M(T), \widetilde{X}(T), \widetilde{M}(T)$ where

$$\big(X(T), M(T), \widetilde{X}(T), \widetilde{M}(T)\big) \stackrel{d}{=} \big(\sum_{i\geq 1} \xi_i, \sum_{i\geq 1} (\xi_i)^+, \sum_{i\geq 1} \widetilde{\xi}_i, \sum_{i\geq 1} (\widetilde{\xi}_i)^+\big). \tag{2.6}$$

Remark 1. *Even though for the purpose of this paper we only need the information of $X(T)$, $M(T)$, $\widetilde{X}(T)$, and $\widetilde{M}(T)$, it is worth noticing that the method described above in fact allows us to construct a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ that supports the entire sample paths $X, \widetilde{X}$ whose endpoint values $X(T), \widetilde{X}(T)$ and suprema $M(T), \widetilde{M}(T)$ admit the joint law in (2.6). In particular, once we obtain l_i based on (2.4), 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 omit the details here and refer the readers to theorem 4 in [45]. The takeaway is that, whenever we apply the coupling described above, one can safely assume the existence of underlying Lévy processes X and $\widetilde{X}$ supported on the same probability space such that the law in (2.6) holds.*

3 Algorithm

In this section, we describe the structure of the rare events we are interested in and propose a strongly efficient simulation algorithm for such rare events. Throughout the rest of this paper, let $X(t)$ be a Lévy process with generating triplet (c_X, σ, ν) satisfying the following assumption on the heavy-tailedness in $X(t)$.

Assumption 1. $\mathbf{E}X(1) = 0$. *Regarding the Lévy measure ν, 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.*

- *(One-sided case) 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 case) 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 $-\alpha$.*

We impose another key assumption that revolves around the continuity of the law of X. Specifically, for any $z > 0$ let $X^{<z}$ be the Lévy process with generating triplet $(c_X, \sigma, \nu|_{(-\infty,z)})$. That is, one can consider $X^{<z}$ as a modulated version of X where all the upward jumps with size larger than z are removed.

Assumption 2. *There exist $z_0, C, \lambda > 0$ and $\theta \in (0, 1]$ such that*

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

This assumption can be interpreted as a strengthened version of Hölder continuity of $X^{<z}(t)$. An in-depth discussion on sufficient conditions for Assumption 2 will be given in Section 4, where we show that Assumption 2 is a mild condition for Lévy process with infinite activities and is easy to verify.

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. Next, define events

$$A \triangleq \{\xi \in \mathbb{D} : \sup_{t\in[0,1]} \xi(t) \geq a;\ \sup_{t\in(0,1]} \xi(t) - \xi(t-) < b\}, \qquad A_n \triangleq \{\bar{X}_n \in A\}. \tag{3.1}$$

For technical reasons, we impose a mild condition on the values of the constants $a, b > 0$.

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

Note that if a RCLL path ξ belongs to the set A, that means the path ξ crossed barrier a even though no upward jumps in ξ is larger than b. Under the scaling of $\bar{X}_n(t) = \frac{1}{n}X(nt)$, it is worth noticing that $\bar{X}_n$ typically resembles the zero function $\mathbf{0}$ for large n. Therefore, it is rather unlikely to observe the barrier crossing phenomenon characterized in events $A_n = \{\bar{X}_n \in A\}$ for large n.

Due to the involved nature of our strongly efficient rare-event simulation algorithm, we will take one step at a time and scrutinize one component of the algorithm in each of the following subsections. The analysis culminates at Theorem 3.4, where we show that the proposed importance sampling algorithm (i.e., Algorithm 2) is unbiased and strongly efficient.

3.1 Importance Sampling Distributions $\mathbf{Q}_n$

At the core of our strongly efficient algorithm is the importance sampling strategy designed under the guidance of the large deviation theories. The framework we describe here can be viewed as an extension of the one in [14]. First, note that

$$l^* \triangleq \lceil a/b \rceil \tag{3.2}$$

characterizes the number of jumps required to cross the barrier a starting from the origin if no jump is allowed to be larger than b. For any $\gamma \in (0, b)$, define sets $B_n^\gamma \triangleq \{\bar{X}_n \in B^\gamma\}$ where

$$B^\gamma \triangleq \{\xi \in \mathbb{D} : \#\{t \in [0, 1] : \xi(t) - \xi(t-) \geq \gamma\} \geq l^*\}. \tag{3.3}$$

Intuitively speaking, the parameter $\gamma \in (0, b)$ in the algorithm is the threshold for *large* jumps, and any path ξ belonging to set B^γ has at least l^* upward jumps that are large. We then apply importance sampling with a defensive mixture (see [35]) and propose (for some $w \in (0, 1)$)

$$\mathbf{Q}_n(\cdot) \triangleq w\mathbf{P}(\cdot) + (1 - w)\mathbf{P}(\ \cdot\ | B_n^\gamma). \tag{3.4}$$

The sampling from $\mathbf{P}(\ \cdot\ |B_n^\gamma)$, and hence $\mathbf{Q}_n(\cdot)$, will be addressed in Section 3.4. Now that the design of the importance sampling distribution $\mathbf{Q}_n$ is clear, a natural choice of an estimator for $\mathbf{P}(A_n)$ is of form $\mathbb{1}_{A_n} \cdot \frac{d\mathbf{P}}{d\mathbf{Q}_n}$ since

$$\mathbf{E}^{\mathbf{Q}_n}\left[\mathbb{1}_{A_n}\frac{d\mathbf{P}}{d\mathbf{Q}_n}\right] = \mathbf{E}[\mathbb{1}_{A_n}] = \mathbf{P}(A_n).$$

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 $\mathbf{P}$. Indeed, this is the importance sampling estimator proposed in [14]. However, in their paper the object of study is compound Poisson process with constant drift, whose entire sample path over a fixed time horizon $[0,T]$ can be simulated exactly. In other words, the exact evaluation of $\mathbb{1}_{A_n}$ is computationally possible. We instead deal with Lévy processes X with infinite activities. The entirety of the sample path of X cannot be simulated exactly with finite computational resources, and the evaluation of $\mathbb{1}_{A_n}$ (i.e., verification of the barrier crossing with jumps bounded by b) is in general not computationally feasible. To sidestep this issue, one possibility is to consider estimators

$$L_n \triangleq Z_n \frac{d\mathbf{P}}{d\mathbf{Q}_n} = \frac{Z_n}{w + \frac{1-w}{\mathbf{P}(B_n^\gamma)}\mathbb{1}_{B_n^\gamma}} \tag{3.5}$$

such that Z_n can be simulated within finite computational resources and recovers the expectation of indicator $\mathbb{1}_{A_n}$ under the importance sampling distribution $\mathbf{Q}_n$. In the remainder of this section we elaborate the design of the estimators Z_n.

3.2 Estimators Z_n

Given $n \geq 1$, consider the following decomposition of the Lévy process X. Recall that $\gamma \in (0,b)$ is the parameter in (3.3) functioning as the threshold for *large* jumps. Let

$$\begin{aligned} J_n(t) &\triangleq \sum_{s\leq t}\Delta X(t)\mathbb{1}\big(\Delta X(t) \geq n\gamma\big), \\ \Xi_n(t) &\triangleq X(t) - J_n(t) = X(t) - \sum_{s\leq t}\Delta X(t)\mathbb{1}\big(\Delta X(t) \geq n\gamma\big). \end{aligned} \tag{3.6}$$

We highlight a few useful facts regarding this decomposition.

- $\mathbf{Q}_n$ only alters the law of J_n, so the law of Ξ_n stays the same under either $\mathbf{Q}_n$ or $\mathbf{P}$ and it is equivalent to that of $X^{<n\gamma}$, i.e. a Lévy process with generating triplet $(c_X, \sigma, \nu|_{(-\infty,n\gamma)})$;
- Under $\mathbf{P}$, the process J_n is a Lévy process with generating triplet $(0,0,\nu|_{[n\gamma,\infty)})$ (more precisely, it is a compound Poisson process);
- Under $\mathbf{Q}_n$, the process $\{J_n(t) :\ t \in [0,n]\}$ has the same law as that of Lévy process with generating triplet $(0,0,\nu|_{[n\gamma,\infty)})$ conditioning on the fact that this process has at least l^* jumps on $[0,n]$;
- Under either $\mathbf{P}$ or $\mathbf{Q}_n$, the two processes J_n and Ξ_n are independent.

Meanwhile, let $\bar{J}_n(t) = \frac{1}{n}J_n(nt), \bar{J}_n = \{\bar{J}_n(t) :\ t \in [0,1]\}$ and $\bar{\Xi}_n(t) = \frac{1}{n}\Xi_n(nt), \bar{\Xi}_n = \{\bar{\Xi}_n(t) :\ t \in [0,1]\}$. Due to $\gamma \in (0,b)$, in the definition of events $A_n = \{\bar{X}_n \in A\}$ in (3.1) 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, we define

$$E \triangleq \{\xi \in \mathbb{D} :\ \sup_{t\in(0,1]}\xi(t) - \xi(t-) < b\}, \qquad E_n \triangleq \{\bar{J}_n \in E\} \tag{3.7}$$

and use indicator $\mathbb{1}_{E_n}$ in our estimator to detect the condition $\sup_{t\in(0,1]}\xi(t)-\xi(t-)<b$ in A_n. Next, let

$$M(t)\triangleq\sup_{s\leq t}X(s),\qquad Y_n^*\triangleq\mathbb{1}\big(M(n)\geq na\big).$$

As discussed above, exact evaluation of Y_n^* is in general not computationally possible. Instead, suppose that we have a sequence of random variables $(\hat{Y}_n^m)_{m\geq 0}$ that only take values in $\{0,1\}$ and gradually approximate Y_n^* as $m\to\infty$. In light of the debiasing technique in Result 3, it is natural to consider the design of Z_n as (under the convention that $\hat{Y}_n^{-1}\equiv 0$)

$$Z_n=\sum_{m=0}^{\tau}\frac{\hat{Y}_n^m-\hat{Y}_n^{m-1}}{\mathbf{P}(\tau\geq m)}\mathbb{1}_{E_n}\tag{3.8}$$

where τ is $\mathrm{Geom}(\rho)$ for some $\rho\in(0,1)$ and is independent of everything else. This construction of Z_n is justified by the following proposition. The proof will be provided in Section 6.1.

Proposition 3.1. *Suppose there exist $C_0>0$, $\rho_0\in(0,1)$, $\mu>2l^*(\alpha-1)$, and $\bar{m}\geq 0$ such that*

$$\mathbf{P}\Big(Y_n^*\neq\hat{Y}_n^m\ \Big|\ \mathcal{D}(\bar{J}_n)=k\Big)\leq C_0\rho_0^m\cdot(k+1)\qquad\forall k\geq 0,n\geq 1,m\geq\bar{m}\tag{3.9}$$

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

$$\mathbf{P}\Big(Y_n^*\neq\hat{Y}_n^m,\ \bar{X}_n\notin A^\Delta\ \Big|\ \mathcal{D}(\bar{J}_n)=k\Big)\leq\frac{C_0\rho_0^m}{\Delta^2 n^\mu}\qquad\forall n\geq 1,m\geq 0,k=0,1,\cdots,l^*-1\tag{3.10}$$

*where $A^\Delta=\big\{\xi\in\mathbb{D}:\sup_{t\in[0,1]}\xi(t)\geq a-\Delta\big\}$. Then for all $\gamma\in(0,b)$ small enough and all $\rho\in(\rho_0,1)$, the estimators L_n specified in (3.5)(3.8) 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 (3.4).*

3.3 Construction of the Approximations $\hat{Y}_n^m$

In light of Proposition 3.1, our next goal is to design approximations $\hat{Y}_n^m$ to meet conditions (3.9) and (3.10). To this end, we discuss how to generate the sequence $\hat{Y}_n^m$ that approximates $Y_n^*\triangleq\mathbb{1}\big(M(n)\geq na\big)$. Recall that $a,b>0$ are the parameters defining the multiple jump barrier crossing rare events A_n (see (3.1)) and $\gamma\in(0,b)$ in the parameter of the algorithm that determines the large jump threshold in B_n^γ (see (3.3)) in our construction of the importance sampling distribution $\mathbf{Q}_n$. Now consider the decomposition of $X(t)=\Xi_n(t)+J_n(t)$ in (3.6). Under both $\mathbf{Q}_n$ and $\mathbf{P}$, we have that Ξ_n and J_n are independent and the law of Ξ_n is equivalent to that of $X^{<n\gamma}$, i.e. a Lévy process with generating triplet $(c_X,\sigma,\nu|_{(-\infty,n\gamma)})$. Therefore, one can first simulate $J_n(t)\triangleq\sum_{s\leq t}\Delta X(t)\mathbb{1}\big(\Delta X(t)\geq n\gamma\big)$ under $\mathbf{Q}_n$ and then sample Ξ_n (or at least approximate its supremum). Suppose that the process J_n has made k jumps over $[0,n]$ (i.e., $\mathcal{D}(\bar{J}_n)=k$) and admits the form

$$J_n(t)=\sum_{i=1}^{k}z_i\mathbb{1}_{[u_i,1]}(t)\qquad\forall t\in[0,n]\tag{3.11}$$

for some $z_i\in[n\gamma,\infty)$ and $u_i\in[0,n]$. Without loss of generality we assume the sequence $u_1<u_2<\cdots<u_k$ is ordered. This allows us to partition the timeline $[0,n]$ into $k+1$ disjoint intervals $[0,u_1),[u_1,u_2),\cdots,[u_{k-1},u_k),[u_k,1]$. We adopt the convention $u_0=0,u_{k+1}=1$ and set $I_i=[u_{i-1},u_i)$ for $i\in[k]$ and $I_{k+1}=[u_k,1]$.

Now let $\zeta(t)=\sum_{i=1}^{k}z_i\mathbb{1}_{[u_i,1]}(t)$. As a result, we can define

$$M_n^{(i)}(\zeta)\triangleq\sup_{t\in I_i}\Xi_n(t)-\Xi_n(u_{i-1}).\tag{3.12}$$

Then for the random function

$$Y_n^*(\zeta) \triangleq \max_{i \in [k+1]} \mathbb{1}\Big(\Xi_n(u_{i-1}) + \zeta(u_{i-1}) + M_n^{(i)} \geq na\Big), \tag{3.13}$$

we have $Y_n^*(J_n) = \mathbb{1}\big(\sup_{t\in[0,n]} X(t) \geq na\big)$. Naturally, to approximate such $M_n^{(i)}(\zeta)$ one would consider applying the stick-breaking approximations (SBA) introduced in Section 2.4. However, approximating the supremum of $X^{<n\gamma}$ on $[0,u]$ using SBA requires the capability to simulate $X^{<n\gamma}(t)$ for any $t > 0$. In general, the law of a Lévy process with infinite activities does not admit an explicit form, and it is unclear how to perform the exact simulation/sampling of Lévy process with infinite activities. To overcome this issue, we incorporate the the Asmussen-Rosiński approximation (ARA) proposed in [4]. The idea is to pick some small threshold level $\kappa \in (0,1)$ and substitute the jump martingale constituted by all jumps bounded by κ with a Brownian motion of the same variance.

Here we state our naming conventions of all the definitions below. We use n to denote the scale function, and m for the approximation level in ARA and SBA. The index i tells us which interval $[u_{i-1}, u_i]$ is concerned. For instance, Ξ_n has the law of the Lévy process $X^{<n\gamma}$, and the law of $\breve{\Xi}_n^m$ approximates that of $X^{<n\gamma}$ by substituting part of the small jump martingale with a Brownian motion component of the same variance; $(l_j^{(i)})_{j\geq 1}, (\xi_j^{(i)})_{j\geq 1}$ form a representation of type (2.5) on $[u_{i-1}, u_i]$ for Ξ_n while $(\xi_j^{(i),m})_{j\geq 1}$ constitutes the stick-breaking representation of $\breve{\Xi}_n^m$ on $[u_{i-1}, u_i]$. Specifically, let

$$\kappa_{n,m} = \frac{\kappa^m}{n^r} \qquad \forall n \geq 1, m \geq 0 \tag{3.14}$$

where $\kappa \in (0,1)$ and $r > 0$ are two other parameters in our algorithm. As a convention we set $\kappa_{n,-1} = 1$. Without loss of generality, we focus on 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)

$$\begin{aligned}\Xi_n(t) = c_X t + \sigma B(t) + &\underbrace{\sum_{s\leq t} \Delta X(s) \mathbb{1}\Big(\Delta X(s) \in (-\infty, -1] \cup [1, n\gamma)\Big)}_{\triangleq J_{n,-1}(t)} \\ &+ \sum_{m\geq 0} \Bigg[\underbrace{\sum_{s\leq t} \Delta X(s) \mathbb{1}\Big(|\Delta X(s)| \in [\kappa_{n,m}, \kappa_{n,m-1})\Big) - t \cdot \nu\Big((-\kappa_{n,m-1}, -\kappa_{n,m}] \cup [\kappa_{n,m}, \kappa_{n,m-1})\Big)}_{\triangleq J_{n,m}(t)}\Bigg].\end{aligned} \tag{3.15}$$

In particular, for any $m \geq 0$ we can see that $J_{n,m}$ is a martingale, and the variance of $J_{n,m}(1)$ is $\bar{\sigma}^2(\kappa_{n,m-1}) - \bar{\sigma}^2(\kappa_{n,m})$ where

$$\bar{\sigma}^2(c) \triangleq \int_{(-c,c)} x^2 \nu(dx) \qquad \forall c \in (0,1]. \tag{3.16}$$

To apply ARA, let $W^{(m)}$ be a sequence of iid copies of standard Brownian that is also independent of $B(t)$. Define

$$\breve{\Xi}_n^m(t) = c_X t + \sigma B(t) + J_{n,-1}(t) + \sum_{j=0}^{m} J_{n,j}(t) + \sum_{j\geq m+1} \sqrt{\bar{\sigma}^2(\kappa_{n,j-1}) - \bar{\sigma}^2(\kappa_{n,j})} \cdot W^{(j)}(t) \tag{3.17}$$

Here the process $\breve{\Xi}_n^m$ can be interpreted as an approximation to Ξ_n where the jump martingale of jumps bounded by $\kappa_{n,m}$ is substituted by a standard Brownian motion with the same variance. Note that (i) for all $m \geq 1$, the random variable $\breve{\Xi}^m_n(t)$ is exactly simulatable as it is a convolution of compound Poisson process (plus constant drift) with a Gaussian random variable; (ii) as $m \to \infty$, the law of $\breve{\Xi}^m_n$ approaches that of the Ξ_n.

Utilizing $\breve{\Xi}^m_n$ constructed under ARA, we are able to apply the SBA technique as follows. Given any step function ζ, we define random functions $\hat{Y}^m_n(\zeta)$ below. For the estimators Z_n in (3.8) we simply plug in $\hat{Y}^m_n(\bar{J}_n)$. Specifically, consider a step function $\zeta(t) = \sum_{i=1}^k z_i \mathbb{1}_{[u_i,n]}(t)$ with $z_i \in [n\gamma, \infty)$ and $u_i \in [0,t]$ with $u_1 < u_2 < \cdots < u_n$, Recall the definition of $M^{(i)}_n(\zeta)$ in (3.12) with $I_i = [u_{i-1}, u_i)$ for $i \in [k]$ and $I_{k+1} = [u_k, 1]$. We will abuse the notations a bit and omit the index n in the subscripts or superscripts of $l^{(i)}_j, \xi^{(i),m}_j, \xi^{(i)}_j$ defined above when there is no ambiguity. Now define

$$l^{(i)}_1 = V^{(i)}_1(u_{i+1} - u_i); \tag{3.18}$$

$$l^{(i)}_j = V^{(i)}_j(u_{i+1} - u_i - l^{(i)}_1 - l^{(i)}_2 - \cdots - l^{(i)}_{j-1}) \qquad \forall j \geq 2 \tag{3.19}$$

where $V^{(i)}_j$ is an iid sequence of Unif$(0,1)$. Next, conditioning on $(l^{(i)}_j)_{j\geq 1}$, one can sample $\xi^{(i),m}_j, \xi^{(i)}_j$ as

$$(\xi^{(i)}_j, \xi^{(i),1}_j, \xi^{(i),2}_j, \xi^{(i),3}_j, \cdots) \stackrel{d}{=} \Big(\Xi_n(l^{(i)}_j), \breve{\Xi}^1_n(l^{(i)}_j), \breve{\Xi}^2_n(l^{(i)}_j), \breve{\Xi}^3_n(l^{(i)}_j), \cdots\Big). \tag{3.20}$$

The coupling in (2.6) then implies

$$\begin{aligned}&\big(\Xi_n(u_i) - \Xi_n(u_{i-1}), M^{(i)}_n(\zeta), \breve{\Xi}^1_n(u_i) - \breve{\Xi}^1_n(u_{i-1}), \breve{M}^{(i),1}_n(\zeta), \breve{\Xi}^2_n(u_i) - \breve{\Xi}^2_n(u_{i-1}), \breve{M}^{(i),2}_n(\zeta), \cdots\big)\\ &\stackrel{d}{=} \Big(\sum_{j\geq 1}\xi^{(i)}_j, \sum_{j\geq 1}(\xi^{(i)}_j)^+, \sum_{j\geq 1}\xi^{(i),1}_j, \sum_{j\geq 1}(\xi^{(i),1}_j)^+, \sum_{j\geq 1}\xi^{(i),2}_j, \sum_{j\geq 1}(\xi^{(i),2}_j)^+, \cdots\Big).\end{aligned} \tag{3.21}$$

Lastly, by terminating the summation after $m + \lceil \log_2(n^d) \rceil$ steps, we get

$$\hat{M}^{(i),m}_n(\zeta) = \sum_{j=1}^{m+\lceil \log_2(n^d)\rceil} (\xi^{(i),m}_j)^+ \tag{3.22}$$

that can be simulated exactly and it approximates $\sum_{j\geq 1}(\xi^{(i),m}_j)^+ \stackrel{d}{=} \sup_{t\in I_i} \breve{\Xi}^m_n(t) - \breve{\Xi}^m_n(u_{i-1})$ as $m \to \infty$. Intuitively speaking, the extra $\lceil \log_2(n^d)\rceil$ term is introduced to guarantee the accuracy of SBA as $n \to \infty$ without significantly increasing the computational cost. Here $d > 0$ is another parameter in the algorithm.

In summary, we define the random function

$$\hat{Y}^m_n(\zeta) = \max_{i\in[k+1]} \mathbb{1}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q),m}_j + \sum_{q=1}^{i-1} z_q + \hat{M}^{(i),m}_n(\zeta) \geq na\Big); \tag{3.23}$$

here $\sum_{q=1}^{i-1} z_q = \zeta(u_{i-1})$ due to $\zeta(t) = \sum_{i=1}^k z_i \mathbb{1}_{[u_i,n]}(t)$, and $\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j \stackrel{d}{=} \breve{\Xi}^m_n(u_{i-1})$ due to the coupling in (3.21). For Z_n in (3.8), we plug in $\hat{Y}^m_n(J_n)$.

At first glance, one may get the impression that the simulation of $\hat{Y}^m_n$ is still computationally challenging due to the existence of infinite sequences. For example, in (3.23) we have $\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q),m}_j$, and the definition of $\breve{\Xi}^m_n$ in (3.17) involves infinitely many iid Brownian motions. Fortunately, the a.s. finite truncation index τ in Z_n guarantees that, when simulating Z_n, once τ is decided there is no need to simulate $\hat{Y}^m_n$ beyond $m \leq \tau$. As a result, in (3.17) one can always combine $\sum_{j\geq \tau+1}\sqrt{\bar{\sigma}^2(\kappa_{n,j-1}) - \bar{\sigma}^2(\kappa_{n,j})} \cdot W^{(j)}(t)$ into a single Brownian motion term. Similarly, to simulate $\hat{Y}^m_n$ for all $m \leq \tau$, when generating $\hat{M}^{(i),m}_n = \sum_{j=1}^m(\xi^{(i),m}_j)^+$ we only need $\xi^{(i),m}_j$ for all $j \leq \tau$. Therefore, for the infinite summation $\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j$ in (3.23), it is safe to combine $\sum_{j>\tau}\xi^{(q)}_j$ and generate the sum in one shot instead of individually sampling all the pieces $\xi_j^{(q)}$. In short, the estimator Z_n (and hence L_n) can be generated using finite computational resources; the steps are detailed in Algorithm 2.

As shown in the next two results, with appropriate parameterization, the algorithm proposed above fulfills conditions (3.9) and (3.10), thus achieving the strong efficiency as guaranteed by Proposition 3.1. We defer the proofs to Section 6.2.

Proposition 3.2. *Let $\beta_+ \in (\beta, 2)$ where $\beta < 2$ is the Blumenthal-Getoor index (see Assumption 1). Given any $\kappa \in (0,1)$ close enough to 0 and any $r, d > 0$ large enough such that*

$$\kappa^{2-\beta_+} < \frac{1}{2}, \qquad d \geq 2, \qquad 2(r - \beta_+) \geq 2,$$

there exists $C_0 > 0$ such that for all $\rho_0 \in (0,1)$ close enough to 1 and all $\gamma \in (0,b)$,

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

Proposition 3.3. *Let $\mu > 0$. Let $\beta_+ \in (\beta, 2)$ where $\beta < 2$ is the Blumenthal-Getoor index (see Assumption 1). Given any $\kappa \in (0,1)$ and any $r, d > 0$ large enough such that*

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

there exists $C_0 < \infty$ such that for all $\rho_0 \in (0,1)$ close enough to 1 and all $\gamma \in (0,b)$,

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

where $A^\Delta = \big\{\xi \in \mathbb{D} : \sup_{t \in [0,1]} \xi(t) \geq a - \Delta\big\}$.

3.4 Sampling from $\mathbf{P}(\ \cdot\ | B_n^\gamma)$

Now, we revisit the sampling of J_n from $\mathbf{P}(\ \cdot\ | B_n^\gamma)$, which is at the core of the implementation of the importance sampling distribution (with defensive mixture) $\mathbf{Q}_n$ in (3.4). First, 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 let $(S_i)_{i \geq 1}$ be the arrival times of jumps in $\widetilde{N}_n(\cdot)$. Let $(W_i)_{i \geq 1}$ be a sequence of iid random variables from distribution $\nu_n^{\text{normalized}}$, which is defined as

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

Under law $\mathbf{P}$, we have (for all $t \geq 0$)

$$J_n(t) = \sum_{i=1}^{\widetilde{N}_n(t)} W_i = \sum_{i \geq 1} W_i \mathbb{1}_{[S_i, \infty)}(t).$$

Furthermore, for each $k \geq 0$, conditioning on $\{\widetilde{N}_n(n) = k\}$, we know that the law of $S_1, \cdots, S_k$ is equivalent to that of the order statistics of k iid samples from Unif$(0, n)$, and W_i's are still independent of S_i's with the law unaltered. Therefore, to sample J_n from $\mathbf{P}(\ \cdot\ | B_n^\gamma)$, it suffices to first sample some k as Poisson$(n \cdot \nu_n[n\gamma, \infty))$, conditioning on $k \geq l^*$, and then independently generate $S_1, \cdots, S_k$ and $W_1, \cdots, W_k$ under the law of $\mathbf{P}(\cdot | \{\widetilde{N}_n(n) = k\})$. It is worth mentioning that the sampling of W_i (i.e., under the law of $\nu_n^{\text{normalized}}$ can be addressed with the help of the inverse of the measure. 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 $U \sim \text{Unif}(0, \nu_n[n\gamma, \infty))$, we have $Q_n^{\leftarrow}(U) \sim \nu_n^{\text{normalized}}$. See Algorithm 1 for the detailed steps.

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.
Sample k from a Poisson distribution with rate $n \cdot \nu_n[n\gamma, \infty)$ conditioning on $k \geq l^*$
Simulate $\Gamma_1, \cdots, \Gamma_k \overset{\text{iid}}{\sim} Unif(0, \nu_n[n\gamma, \infty))$
Simulate $U_1, \cdots, U_k \overset{\text{iid}}{\sim} Unif(0, n)$
Return $J_n = \sum_{i=1}^{k} Q_n^{\leftarrow}(\Gamma_i)\mathbb{1}_{[U_i, n]}$

3.5 Strong Efficiency and Computational Complexity

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

Require: $w \in (0,1), \gamma > 0, r > 0, d > 0, \kappa \in (0,1), \rho \in (0,1)$ for parameters in the algorithm; $a, b > 0$ as characterization of set A; (c_X, σ, ν) is the generating triplet of X; $\bar{\sigma}(\cdot)$ is defined in (3.16).

Set $t_n = \lceil \log_2(n^d) \rceil$ and $\kappa_{n,m} = \kappa^m / n^r$ for any $n \geq 1, m \geq 0$.

Sample $U \sim Unif(0,1)$ ▷ Sample J_n from $\mathbb{Q}$
if $U < w$ **then**
  Sample $J_n = \sum_{i=1}^{k} z_i \mathbb{1}_{[u_i, n]}$ as a compound Poisson process over $[0, n]$ with jump intensity measure $\nu|_{[n\gamma, \infty)}$
else
  Sample $J_n = \sum_{i=1}^{k} z_i \mathbb{1}_{[u_i, n]}$ using Algorithm 1
Let $u_0 = 0, u_{k+1} = n$.

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

for $i = 1, \cdots, k+1$ **do** ▷ Stick-breaking procedure
  Sample $V_1^{(i)} \sim Unif(0,1)$, and let $l_1^{(i)} = V_1^{(i)}(u_i - u_{i-1})$
  for $j = 2, 3, \cdots, t_n + \tau$ **do**
    Sample $V_j^{(i)} \sim Unif(0,1)$, and let $l_j^{(i)} = V_j^{(i)}(u_i - u_{i-1} - l_1^{(i)} - l_2^{(i)} - \cdots - l_{j-1}^{(i)})$
  Set $l_{t_n+\tau+1}^{(i)} = u_i - u_{i-1} - l_1^{(i)} - l_2^{(i)} - \cdots - l_{t_n+\tau}^{(i)}$

for $i = 1, \cdots, k+1$ **do** ▷ Decide $\xi_m^{(i),j}$—the increment on stick $l_j^{(i)}$ under ARA at level m
  for $j = 1, 2, \cdots, t_n + \tau + 1$ **do**
    Sample $B^{(i),j} \sim N(0, \sigma^2 \cdot l_j^{(i)})$
    Sample $J_{-1}^{(i),j} \sim F(l_j^{(i)}, \nu|_{(-\infty, n\gamma) \setminus (-\kappa_{n,0}, \kappa_{n,0})})$

    for $m = 0, 1, \cdots, \tau - 1$ **do**
      Sample $J_m^{(i),j} \sim F(l_j^{(i)}, \nu|_{(-\kappa_{n,m}, \kappa_{n,m}) \setminus (-\kappa_{n,m+1}, \kappa_{n,m+1})})$
      Sample $W_m^{(i),j} \sim N(0, (\bar{\sigma}^2(\kappa_{n,l}) - \bar{\sigma}^2(\kappa_{n,l+1})) \cdot l_j^{(i)})$
    Sample $W_\tau^{(i),j} \sim N(0, \bar{\sigma}^2(\kappa_{n,\tau}) \cdot l_j^{(i)})$

    for $m = 0, \cdots, \tau$ **do**
      Let $\xi_m^{(i),j} = c_X \cdot l_j^{(i)} + B^{(i),j} + \sum_{q=-1}^{m-1} J_q^{(i),j} + \sum_{q=m}^{\tau} W_q^{(i),j}$

for $m = 0, 1, \cdots, \tau$ **do** ▷ Evaluate $\hat{Y}_n^m$
  for $i = 1, 2, \cdots, k+1$ **do**
    Let $\hat{M}_m^{(i)} = \sum_{l=0}^{i-1} \sum_{j=1}^{t_n+\tau+1} \xi_m^{(i),j} + \sum_{j=1}^{t_n+\tau} (\xi_m^{(i),j})^+$
  Let $\hat{Y}_n^m = \mathbb{1}\{ \max_{i1, \cdots, k+1} [\hat{M}_m^{(i)} + J_n(u_{i-1})] \geq na \}$

Let $Z_n = \hat{Y}_n^0 + \sum_{m=1}^{\tau} (\hat{Y}_n^m - \hat{Y}_n^{m-1}) / \rho^{m-1}$ ▷ Return the Estimator L_n
if $\max_{i=1,\cdots,k} z_i > b$ **then**
  Return $L_n = 0$.
else
  Let $\lambda_n = n\nu[n\gamma, \infty),\ p_n = 1 - \sum_{l=0}^{l^*-1} e^{-\lambda_n} \frac{\lambda_n^l}{l!},\ I_n = \mathbb{1}\{J_n \in B_n^\gamma\}$
  Return $L_n = Z_n / (w + \frac{1-w}{p_n} I_n)$

All the discussions above lead to the importance sampling algorithm in Algorithm 2. Here for any $t > 0$ and any Borel measure μ with $\mu(\mathbb{R}) < \infty$, we use $F(t, \mu)$ to denote the law of the compound

Poisson process at time t with jump intensity measure μ; that is, the arrival rate of jumps is $\mu(\mathbb{R})$ and the law of jump sizes is $\mu(\cdot)/\mu(\mathbb{R})$. Below is a summary of all the parameters in the algorithm.

- $\gamma \in (0, b)$: the threshold in B^γ defined in (3.3)
- $w \in (0, 1)$: the weight of the defensive mixture in $\mathbf{Q}_n$; see (3.4)
- $\rho \in (0, 1)$: the geometric rate of decay for $\mathbf{P}(\tau \geq m)$ in (3.8)
- $\kappa \in (0, 1)$, $r > 0$: determining the truncation threshold $\kappa_{n,m}$; see (3.14)
- $d > 0$: determining the $\log_2(n^d)$ term in (3.22)

As shown in Theorem 3.4 below, Algorithm 2 is unbiased and strongly efficient when properly parametrized.

Theorem 3.4. *There exist $\bar{\kappa} \in (0, 1)$, $\bar{r} > 0$ and $\bar{d} > 0$ such that the following claim is valid: given any $\kappa \in (0, \bar{\kappa}]$, $r \geq \bar{r}$, $d \geq \bar{d}$ and any $w \in (0, 1)$, there is $\bar{\rho} \in (0, 1)$ such that Algorithm 2 is unbiased and strongly efficient under any $\gamma \in (0, b)$ small enough and any $\rho \in (\bar{\rho}, 1)$.*

Proof. Fix some $\mu > 2l^*(\alpha - 1)$. Fix some $\beta_+ \in (\beta, 2)$ where $\beta < 2$ is the Blumenthal-Getoor index (see Assumption 1). Set $\bar{\kappa} \in (0, 1)$ small enough and $\bar{r}, \bar{d}$ large enough such that

$$\bar{\kappa}^{2-\beta_+} < \frac{1}{2}, \qquad 2(\bar{r} - \beta_+) > 2 \vee (\mu - 1), \qquad \bar{d} \geq 2, \qquad \bar{d} > (2\mu - 1). \tag{3.24}$$

Now pick $\kappa \in (0, \bar{\kappa}]$, $r \geq \bar{r}$, $d \geq \bar{d}$ and $w \in (0, 1)$. Thanks to Propositions 3.2 and 3.3, we can find some $C_0 > 0$, $\bar{m} \geq 0$ and $\rho_0 \in (0, 1)$ such that conditions (3.9) and (3.10) hold for $\hat{Y}^m_n = \hat{Y}^m_n(J_n)$ (for its definition, see (3.23)) under the parameters specified above. It then follows immediately from Proposition 3.1 that the estimator L_n (and hence Algorithm 2) is unbiased and strongly efficient under any $\gamma \in (0, b)$ small enough and any $\rho \in (\rho_0, 1)$. To conclude the proof, one only needs to set $\bar{\rho} = \rho_0$. □

We stress that the exact range of the parameters for strong efficiency to hold in Theorem 3.4 is readily available. First, note that the choice of $w \in (0, 1)$ won't affect the unbiasedness and strong efficiency of the algorithm. Next, $\bar{\kappa}, \bar{r}$ and $\bar{d}$ can be determined using (3.24). After picking $\kappa \in (0, \bar{\kappa}]$, $r \geq \bar{r}$, $d > \bar{d}$ and any $w \in (0, 1)$, it is shown in the proof of Propositions 3.2 and 3.3 (specifically, in (6.24)–(6.36) and (6.43)) that the value of $\bar{\rho}$ can be determined as follows. First pick $\alpha_3 \in (0, \frac{\theta}{\lambda})$, $\alpha_4 \in (0, \frac{\theta}{2\lambda})$ where $\lambda > 0$ and $\theta \in (0, 1]$ are the constants in Assumption 2. Next, pick $\alpha_2 \in (0, \frac{\alpha_3}{2} \wedge 1)$ and $\alpha_1 \in (0, \frac{\theta\alpha_2}{\lambda})$. Also, fix $\delta \in (1/\sqrt{2}, 1)$. Since we require α_2 to be strictly less than 1, it is easy to see the existence of some integer $\bar{m}$ such that $\delta^{m\alpha_2} - \delta^m \geq \frac{\delta^{m\alpha_2}}{2}$ and $\delta^{m\alpha_2} < a$ $\forall m \geq \bar{m}$. Here $a > 0$ is the parameter in set A; see Assumption 3. This allows us to pick $\rho_1 \in (0, 1)$ such that

$$\rho_1 > \max\Big\{\delta^{\alpha_1},\ \frac{\kappa^{2-\beta_+}}{\delta^2},\ \frac{1}{\sqrt{2}\delta},\ \delta^{\theta\alpha_2 - \lambda\alpha_1},\ \delta^{\theta - \lambda\alpha_3}, \delta^{-\alpha_2 + \frac{\alpha_3}{2}}\Big\}.$$

Now we can fix some

$$\bar{\rho} \in \Big(\max\Big\{\frac{1}{\sqrt{2}},\ \kappa^{2-\beta_+},\ \rho_1\Big\}, 1\Big).$$

and pick a larger $\bar{m}$ if necessary to make sure that $m^2\rho_1^m \leq \bar{\rho}^m$ $\forall m \geq \bar{m}$. Lastly, the choice of γ is detailed in the proof of Proposition 3.1. Specifically, after picking $\rho \in (\bar{\rho}, 1)$, one can find some $q > 1$ such that $\bar{\rho}^{1/q} < \rho$. Let $p > 1$ be such that $\frac{1}{p} + \frac{1}{q} = 1$. For the algorithm to be unbiased and strongly efficient, one can pick any $\gamma > 0$ small enough such that

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

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

This section investigates the sufficient conditions for Assumption 2. In particular, we focus on the case where $\theta = 1$ so that Assumption 2 can be viewed as a strengthened and uniform version of Lipschitz continuity of the law of $X^{<z}(t)$.

To demonstrate the key technique in our approach, we start with a simple case where the Lévy process $X(t)$ has generating tripet (c_X, σ, ν) with $\sigma > 0$. This allows us to decompose the process into

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

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

$$\begin{aligned}
\mathbf{P}(X^{<z}(t) \in [x, x+\delta]) &= \int_{\mathbb{R}} \mathbf{P}(\sigma B(t) \in [x-y, x-y+\delta]) \cdot \mathbf{P}(Y^{<\gamma}(t) \in dy) \\
&= \int_{\mathbb{R}} \mathbf{P}\bigg(\frac{B(t)}{\sqrt{t}} \in \Big[\frac{x-y}{\sigma\sqrt{t}}, \frac{x-y+\delta}{\sigma\sqrt{t}}\Big]\bigg) \cdot \mathbf{P}(Y^{<\gamma}(t) \in dy) \\
&\leq \frac{1}{\sigma\sqrt{2\pi}} \cdot \frac{\delta}{\sqrt{t}}
\end{aligned} \tag{4.1}$$

since the density of a standard Normal distribution is bounded by $1/\sqrt{2\pi}$. This immediately verifies Assumption 2 under $\theta = 1,\ \lambda = 1/2, C = \frac{1}{\sigma\sqrt{2\pi}}$, and any $z_0 > 0$.

Now we consider the case where $\sigma = 0$. For any two (Borel) measures μ_1, μ_2 on $\mathbb{R}$, their difference $\mu = \mu_1 - \mu_2$ can be considered as a signed measure. 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. When $A = \mathbb{R}$ we simply write $\mu_1 - \mu_2 \geq 0$. Returning to the Lévy measure ν of the process $X(t)$, If we can find some $z_0 > 0$ and some (positive) Borel measure μ such that $(\nu - \mu)|_{(-z_0,z_0)} \geq 0$, then by splitting the underlying Poisson random measure for the jumps in the process $X(t)$, for any $z \geq z_0$ we have the decomposition

$$X^{<z}(t) \stackrel{d}{=} Y(t) + \widetilde{X}^{z_0,z}(t)$$

where $\mu_{z_0} = \mu|_{(-z_0,z_0)}$, Y is a Lévy process with generating triplet $(0, 0, \mu_{z_0})$, $\widetilde{X}^{z_0,z}$ is a Lévy process with generating triplet $(c_X, 0, \nu - \mu_{z_0})$, and the two processes are independent. Furthermore, if we can show that Assumption 2 holds for the process $Y(t)$ with generating triplet $(0, 0, \mu_{z_0})$, then analogous to the arguments in (4.1) we can pass the continuity conditions in Assumption 2 from $Y(t)$ to $X^{<z}(t)$ through the convolution of $X^{<z}(t) \stackrel{d}{=} Y(t) + \widetilde{X}^{z_0,z}(t)$.

The success of this technique hinges on the identification of the majorized measure μ, which naturally depends on the property of the majorizing measure ν. Recall the definition of regularly varying function in Definition 1. The following result provides sufficient conditions for Assumption 2 when $\nu[x, \infty)$ or $\nu(-\infty, x]$ is regularly varying as $x \downarrow 0$.

Proposition 4.1. *Let $\alpha \in (0, 2), z_0 > 0$, and $\epsilon \in (0, (2-\alpha)/2)$. Let ν be a Borel measure supported on $(0, \infty)$ such that $\nu[x, \infty)$ is regularly varying as $x \to 0$ with index $\alpha + 2\epsilon$. There exists a constant $C < \infty$ such that for the Levy process $\{Y(t) : t \geq 0\}$ with generating triplet $(0, 0, \nu|_{(-z_0,z_0)})$,*

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

where the L_∞ norm for a density f is defined as $\|f\|_\infty = \sup_{x \in \mathbb{R}} |f(x)|$, and $f_{Y(t)}$ is the density function of the distribution of $Y(t)$.

We give the detailed proof in Section 6.3. The main idea is that when ν is regularly varying with some index $\alpha' = \alpha + \epsilon$ at the origin, it dominates the measure $\nu_\alpha[x, \infty) = 1/x^\alpha$ eventually as $x \downarrow 0$. This allows us to factor out a polynomial component from ν around the origin and associate $Y(t)$ with an α-stable Lévy process. By comparing the characteristic functions of the two processes, we are able transfer the continuity of the α-stable Lévy process to the process $Y(t)$.

Equipped with Proposition 4.1, we establish the following set of sufficient conditions for Assumption 2.

Theorem 4.2. *Let (c_X, σ, ν) be the generating triplet of Lévy process X.*

(i) If $\sigma > 0$, then Assumption 2 holds for $\theta = 1$, $\lambda = 1/2$, 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 α' as $x \downarrow 0$, then Assumption 2 holds with $\theta = 1$ and $\lambda = 1/\alpha$ for any $\alpha \in (0, \alpha')$.

Proof. Part (i) follows immediately from the calculations in (4.1). 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 $\epsilon = (\alpha' - \alpha)/2 \in \big(0, (2 - \alpha)/2\big)$. For any $z \geq z_0$, consider the decomposition

$$X^{<z}(t) \stackrel{d}{=} Y(t) + \widetilde{X}^{z_0,z}(t)$$

where $\mu_{z_0} = \mu|_{(-z_0,z_0)}$, Y is a Lévy process with generating triplet $(0, 0, \mu_{z_0})$, $\widetilde{X}^{z_0,z}$ is a Lévy process with generating triplet $(c_X, \sigma, \nu - \mu_{z_0})$, and the two processes are independent. First of all, applying Proposition 4.1, we can find $C > 0$ such that $\|f_{Y(t)}\|_\infty \leq \frac{C}{t^{1/\alpha} \wedge 1}$ $\forall t > 0$. Next, due to the independence between Y and $\widetilde{X}^{z_0,z}$, it then holds for any $x \in \mathbb{R}, \delta \geq 0$, and $t > 0$ that

$$\begin{aligned}
\mathbf{P}(X^{<z}(t) \in [x, x + \delta]) &= \int_{\mathbb{R}} \mathbf{P}(Y(t) \in [x - y, x - y + \delta]) \cdot \mathbf{P}(\widetilde{X}^{z_0,z}(t) \in dy) \\
&\leq \frac{C}{t^{1/\alpha} \wedge 1} \cdot \delta \qquad \text{due to } \|f_{Y(t)}\|_\infty \leq \frac{C}{t^{1/\alpha} \wedge 1}.
\end{aligned}$$

This concludes the proof. ☐

Remark 2. *To understand why the conditions stated in Theorem 4.2 are considered mild, let us recall our underlying assumption that X exhibits infinite activity, implying either $\sigma > 0$ or $\nu(\mathbb{R}) = \infty$. Theorem 4.2 (i) deals with the scenario when $\sigma > 0$. On the other hand, when $\sigma = 0$ we must have either $\lim_{\epsilon \downarrow 0} \nu[\epsilon, \infty) = \infty$ or $\lim_{\epsilon \downarrow 0} \nu(-\infty, -\epsilon] = \infty$. To satisfy the conditions in part (ii) of Theorem 4.2, it only requires $\nu[\epsilon, \infty)$ (or $\nu(-\infty, -\epsilon]$) to approach infinity at a rate that matches or exceeds some polynomial function.*

Aside from the regularly varying structure exploited in Theorem 4.2 (ii), we discuss another type of self-similarity structure in the Lévy measure ν that can be utilized to verify Assumption 2. Given $\alpha \in (0, 2)$ and $b > 1$, we say that the process X is **α-semi-stable with span b** if its Lévy measure satisfies

$$\nu = b^{-\alpha} T_b \nu \tag{4.2}$$

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, note that X is **α-stable** if

$$\nu(dx) = \begin{cases} c_1 \frac{dx}{x^{1+\alpha}} & \forall x > 0 \\ c_2 \frac{dx}{x^{1+\alpha}} & \forall x < 0 \end{cases}$$

where $c_1, c_2 \geq 0, c_1 + c_2 > 0$. See Theorem 14.3 in [49] for details.

To see how α-semi-stability of ν is different from the concept of regular variations, consider the following examples. Given a Borel measure ν, suppose that $f(x) = \nu\big((-\infty, -x] \cup [x, \infty)\big)$ is regularly varying at 0 with index $\alpha > 0$. Even if ν satisfies the scaling-invariant property in (4.2) 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 $f(tx)/f(x)$ does not have a limit as x approaches 0. This confirms that $\nu[x, \infty)$ is not regularly varying as $x \downarrow 0$.

The usefulness of the concept of semi-similarity is demonstrated in Proposition 4.3 below. The proof is detailed in Section 6.3, and the key idea is to argue the similarity between the density of the non-trivial α-semi-stable process $Y(t)$ and the truncated $Y^{(-z_0, z_0)}(t)$ by comparing their characteristic functions.

Proposition 4.3. *Let $\alpha \in (0, 2)$ and ν be the Lévy measure of a non-trivial α-semi-stable process $Y(t)$ of span $b > 1$. Let $N \in \mathbb{Z}$. There exists some $C \in (0, \infty)$ such that, under $z_0 = b^N$,*

$$\Big\| f_{Y^{(-z_0, z_0)}(t)} \Big\|_\infty \leq \frac{C}{t^{1/\alpha} \wedge 1} \qquad \forall t > 0$$

where $\{Y^{(-z_0, z_0)}(t) : \ t > 0\}$ is the Lévy process with generating triplet $(0, 0, \nu|_{(-z_0, z_0)})$ and $f_{Y^{(-z_0, z_0)}(t)}$ is the density of distribution of $Y^{(-z_0, z_0)}(t)$.

Lastly, by applying Proposition 4.3, we yield another set of sufficient conditions for Assumption 2.

Theorem 4.4. *Let (c_X, σ, ν) be the generating triplet of Lévy process X. If there exist some Borel measure μ and some $z_0 > 0, \alpha \in (0, 2)$ such that $(\nu - \mu)|_{(-z_0, z_0)} \geq 0$ and μ is the Lévy measure of some α-semi-stable process, then Assumption 2 holds with $\theta = 1$ and $\lambda = 1/\alpha$.*

Proof. Let $b > 1$ be the span of the α-semi-stable process. Fix some $N \in \mathbb{Z}$ such that $b^N \leq z_0$. For any $z \geq z_0$, consider the decomposition

$$X^{<z}(t) \stackrel{d}{=} Y_*(t) + \widetilde{X}^z(t)$$

where $\mu_* = \mu|_{(-b^N, b^N)}$, Y_* is a Lévy process with generating triplet $(0, 0, \mu_*)$, $\widetilde{X}^z$ is a Lévy process with generating triplet $(c_X, \sigma, \nu - \mu_*)$, and the two processes are independent. First of all, applying

Proposition 4.3, we can find $C > 0$ such that $\|f_{Y_*(t)}\|_\infty \leq \frac{C}{t^{1/\alpha}\wedge 1}$ $\forall t > 0$. Next, due to the independence between Y_* and $\widetilde{X}^z$, it then holds for any $x \in \mathbb{R}, \delta \geq 0$, and $t > 0$ that

$$\begin{aligned}\mathbf{P}(X^{<z}(t) \in [x, x+\delta]) &= \int_{\mathbb{R}} \mathbf{P}(Y_*(t) \in [x-y, x-y+\delta]) \cdot \mathbf{P}(\widetilde{X}^z(t) \in dy) \\ &\leq \frac{C}{t^{1/\alpha}\wedge 1} \cdot \delta \qquad \text{due to } \|f_{Y_*(t)}\|_\infty \leq \frac{C}{t^{1/\alpha}\wedge 1}.\end{aligned}$$

This concludes the proof. □

5 Numerical Experiments

In this section, we apply the importance sampling strategy outlined in Algorithm 2 to conducty numerical experiments, showcasing (i) the performance of the importance sampling estimator under varying scaling factors and tail distributions, and (ii) the efficiency of the algorithm when compared to crude Monte Carlo methods.

We consider a Lévy process given by $X(t) = B(t) + \sum_{i=1}^{N(t)} W_i$, 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 iid samples from Pareto distribution with tail index $\alpha > 1$, i.e.,

$$\mathbf{P}(W_1 > x) = \frac{1}{\max\{x, 1\}^\alpha}.$$

For each $n \geq 1$, we define the scaled process $X_n(t) = \frac{X(nt)}{n}$. The objective is to estimate the probability of the events $A_n = \{X_n \in A\}$, where

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

with $a = 2$ and $b = 1.15$. To evaluate the performance of the importance sampling estimator under different scaling factors and tail distributions, we conduct experiments using $\alpha \in \{1.45, 1.6, 1.75\}$, and $n \in \{1000, 2000, \cdots, 10000\}$. The efficiency of an estimator is quantified by the *relative error*, namely the ratio of the standard deviation estimated from all samples to the estimated mean.

For the parameters in Algorithm 2, we set $\gamma = 0.2$, $w = 0.05$, $\rho = 0.95$, and $\bar{d} = 2$. Note that $l^* = \lceil a/b \rceil = 2$ in this case. We generate 500,000 independent samples for each combination of $\alpha \in \{1.45, 1.6, 1.75\}$ and $n \in \{1000, 2000, \cdots, 10000\}$. We also compare the efficiency of the importance sampling estimator against the crude Monte Carlo methods. For the number of simulation trials 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 estimated value of $\mathbf{P}(A_n)$ using Algorithm 2.

The results are summarized in Table 5.1 and Figure 5.1. In Table 5.1, it is evident that for a fixed α, the relative error of the importance sampling estimator remains constant regardless of the magnitude of n, demonstrating the strong efficiency established in Theorem 3.4. Figure 5.1 presents a comparison of relative errors between the two methods, highlighting that our importance sampling strategy significantly outperforms crude Monte Carlo methods by several orders of magnitude. In summary, when the importance sampling algorithm is appropriately parameterized, the efficiency of our importance sampling estimator becomes increasingly more evident against the vanilla Monte Carlo approach as the scaling factor n (and thus, the rarity of event A_n) grows larger.

6 Proofs

6.1 Proof of Proposition 3.1

We first prepare two technical lemmas based on the sample-path large deviations for heavy-tailed Lévy processes discussed in Section 2.2.

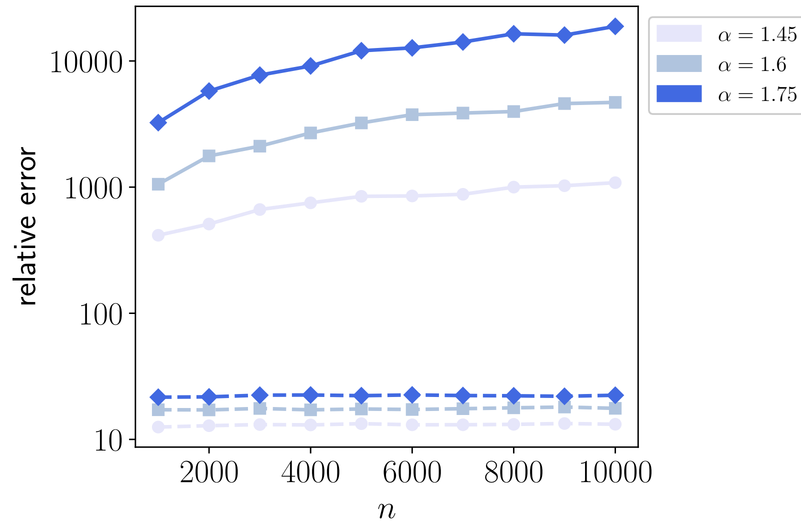

Figure 5.1: Relative errors of the proposed importance-sampling estimator (dashed lines) and crude Monte-Carlo methods (solid lines).

Table 5.1: Rare-event simulation results using Algorithm 2. First row: estimated probability of $\mathbf{P}(A_n)$; 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

Lemma 6.1. *There exist $c_A, C_A \in (0,\infty)$ such that*

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

Proof. In this proof we focus on the two-sided case in Assumption 1, but it is worth noticing that analysis to 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).

We claim that

(i) $\mathcal{J}(A) = l^*, \mathcal{K}(A) = 0$ is the only feasible solution to $\big(\mathcal{J}(A), \mathcal{K}(A)\big) \in \underset{(j,k)\in\mathbb{N}^2,\ \mathbb{D}_{j,k}\cap A\neq\emptyset}{\operatorname{argmin}} j(\alpha-1)+k(\alpha'-1)$; here $\mathbb{D}_{j,k}$ is the set containing all step functions in $\mathbb{D}$ vanishing at the origin that has exactly j upward jumps and k downward jumps;

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

(iii) the set A is bounded away from $\mathbb{D}_{<l^*,0} = \bigcup_{j\leq l^*-1} \mathbb{D}_{j,0}$; recall that $\mathbb{D}_{<j,k} \triangleq \bigcup_{(l,m)\in\mathbb{I}_{<j,k}} \mathbb{D}_{l,m}$ where $\mathbb{I}_{<j,k} \triangleq \big\{(l,m) \in \mathbb{N}^2\backslash\{(j,k)\} :\ l(\alpha-1)+m(\alpha'-1) \leq j(\alpha-1)+k(\alpha'-1)\big\}$.

Then one can apply Result 2 to conclude the proof. In particular, since the set A is bounded away from $\mathbb{D}_{<l^*,0}$ and $\big(\mathcal{J}(A),\mathcal{K}(A)\big)=(l^*,0)$, we yield

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

Now it only remains to prove claims (i), (ii), and (iii).

We start from 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

$$\xi(t)=\sum_{i=1}^j u_i\mathbb{1}_{[t_i,1]}(t)-\sum_{i=1}^k v_i\mathbb{1}_{[s_i,1]}(t). \tag{6.1}$$

We first show that $\mathbb{D}_{l^*,0}\cap A\neq\emptyset$. Recall the definition of $A=\{\xi\in\mathbb{D}:\sup_{t\in[0,1]}\xi(t)\geq a;\sup_{t\in(0,1]}\xi(t)-\xi(t-)<b\}$ in (3.1). Also, recall that $l^*=\lceil a/b\rceil$ and $a/b\notin\mathbb{Z}$; see Assumption 3. Therefore, $(l^*-1)b<a<l^*b$. This allows us to pick some $\epsilon>0$ small enough such that $l^*(b-\epsilon)>a$. By setting $j=l^*,k=0$, and $u_i=b-\epsilon$ for all $i\in[l^*]$ in (6.1), we can see that $\sup_{t\in[0,1]}\xi(t)=\sum_{i=1}^{l^*}u_i=l^*(b-\epsilon)>a$. This establishes $\mathbb{D}_{l^*,0}\cap A\neq\emptyset$.

Now to conclude the proof of claim (i), it suffices to show that $j\geq l^*$ is a necessary condition for $\mathbb{D}_{j,k}\cap A\neq\emptyset$. Indeed, knowing that $j\geq l^*$ is a necessary condition, one can see that

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

Also, recall that we have just shown $\mathbb{D}_{l^*,0}\cap A\neq\emptyset$. Because of $\alpha,\alpha'>1$ (see Assumption 1), we can then show that $\underset{(j,k)\in\mathbb{N}^2,\ \mathbb{D}_{j,k}\cap A\neq\emptyset}{\operatorname{argmin}} j(\alpha-1)+k(\alpha'-1)=\{(l^*,0)\}$ and conclude the proof of claim (i). In other words, to prove claim (i) one only needs to show that $j\geq l^*$ is a necessary condition for $\mathbb{D}_{j,k}\cap A\neq\emptyset$. To this end, we proceed with a proof by contradiction. Suppose there is some $j\leq l^*-1,k\geq 0$ and some $\xi\in\mathbb{D}_{j,k}\cap A$. Then by definition of set A, in the representation (6.1) of ξ we must have $u_i<b$ for all $i\in[k]$. As a result,

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

This leads to the contradiction that $\xi\notin A$ and allows us to conclude the proof of claim (i).

To prove claim (ii), recall that we can pick some $\epsilon>0$ small enough such that $l^*(b-\epsilon)>a$. Therefore, given any $u_i\in(b-\epsilon,b)$ and $0<t_1<t_2<\cdots<t_{l^*}<1$, for the step function $\xi(t)=\sum_{i=l^*}^j u_i\mathbb{1}_{[t_i,1]}(t)$ we must have $\sup_{t\in[0,1]}\xi(t)\geq\xi(1)>a$, thus implying $\xi\in A$. This observation leads to (for the definition of $\mathbf{C}_{j,k}$, see (2.3))

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

This establishes claim (ii).

Moving onto claim (iii), recall the bound $a>(l^*-1)b$. It then holds for any $\epsilon>0$ small enough that $a-\epsilon>(l^*-1)(b+\epsilon)$. Fix such $\epsilon>0$. It suffices to show that

$$\boldsymbol{d}(\xi,\xi')\geq\epsilon\qquad\forall j=0,1,\cdots,l^*-1,\ \xi\in\mathbb{D}_{j,0},\ \xi'\in A. \tag{6.2}$$

Here $\boldsymbol{d}$ is the Skorokhod J_1 metric; see (2.1) for the definition. Now fix some $j=0,1,\cdots,l^*-1,\ \xi\in\mathbb{D}_{j,0}$, and $\xi'\in A$. Also, fix some λ that is an increasing homeomorphisms from $[0,1]$ to itself. Let

$\widetilde{\xi}(t) = \xi\big(\lambda(t)\big)$. It suffices to show that $\sup_{t\in[0,1]}|\xi'(t) - \widetilde{\xi}(t)| \geq \epsilon$. First, using the representation in (6.3), we know there is some $(u_i)_{i\leq j} \in (0,\infty)^j$ and some $0 < \widetilde{t}_1 < \cdots < \widetilde{t}_j < 1$ such that

$$\widetilde{\xi}(t) = \sum_{i=1}^{j} u_i \mathbb{1}_{[\widetilde{u}_i,1]}(t).$$

Next, we consider two different cases. If $u_i > b+\epsilon$ for some $i \in [j]$, then by definition of A (in particular, $\xi'(t) - \xi'(t-) < b$ for all $t \in [0,1]$ and $\xi' \in A$), we have $\sup_{t\in[0,1]}|\xi'(t) - \widetilde{\xi}(t)| \geq \epsilon$. On the other hand, if $u_i \in (0, b+\epsilon]$ for all $i \in [j]$, then

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

Due to $\sup_{t\in[0,1]}\xi'(t) \geq a$ for $\xi' \in A$, we yield $\sup_{t\in[0,1]}|\xi'(t) - \widetilde{\xi}(t)| \geq \epsilon$ again. This establishes (6.2), and hence claim (iii). $\square$

Lemma 6.2. *Let $p > 1$. Let $\Delta > 0$ be such that $a - \Delta > (l^*-1)b$ and $\big[a - \Delta - (l^*-1)b\big]/\gamma \notin \mathbb{Z}$. Let*

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

If $(J_\gamma + l^* - 1)/p > 2l^*$, then

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

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

Proof. Similar to the proof of Lemma 6.1, we focus on the two-sided case in Assumption 1, but it is worth noticing that analysis to 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).

First, recall that $\gamma \in (0,b)$ is the parameter in the definition of set B^γ in (3.3). Also, recall that $J_n(t) = \sum_{s\leq t}\Delta X(t)\mathbb{1}\big(\Delta X(t) \geq n\gamma\big)$ and $\bar{J}_n = \{\frac{1}{n}J_n(nt) :\ t \in [0,1]\}, \bar{X}_n = \{\frac{1}{n}X(nt) :\ t \in [0,1]\}$. Therefore, $\mathbf{P}\big(\bar{X}_n \in A^\Delta,\ \mathcal{D}(\bar{J}_n) \leq l^*-1\big) = \mathbf{P}\big(\bar{X}_n \in E(\Delta)\big)$ where

$$E(\Delta) = \Big\{\xi \in \mathbb{D} :\ \sup_{t\in[0,1]}\xi(t) \geq a - \Delta;\ \sup_{t\in(0,1]}\xi(t) - \xi(t-) < b,$$
$$\#\big\{t \in [0,1] :\ \xi(t) - \xi(t-) \geq \gamma\big\} \leq l^*-1\Big\}.$$

Furthermore, we claim that

(i) $\mathcal{J}\big(E(\Delta)\big) = l^*-1+J_\gamma, \mathcal{K}\big(E(\Delta)\big) = 0$ is the only feasible solution to

$$\Big(\mathcal{J}\big(E(\Delta)\big), \mathcal{K}\big(E(\Delta)\big)\Big) \in \underset{(j,k)\in\mathbb{N}^2,\ \mathbb{D}_{j,k}\cap E(\Delta)\neq\emptyset}{\operatorname{argmin}} j(\alpha-1) + k(\alpha'-1)$$

where $\mathbb{D}_{j,k}$ is the set containing all step functions in $\mathbb{D}$ vanishing at the origin that has exactly j upward jumps and k downward jumps;

(ii) The set $E(\Delta)$ is bounded away from $\mathbb{D}_{<l^*-1+J_\gamma,0} = \bigcup_{j\leq l^*+J_\gamma-2}\mathbb{D}_{j,0}$; recall that $\mathbb{D}_{<j,k} \triangleq \bigcup_{(l,m)\in\mathbb{I}_{<j,k}}\mathbb{D}_{l,m}$ where $\mathbb{I}_{<j,k} \triangleq \big\{(l,m) \in \mathbb{N}^2\backslash\{(j,k)\} :\ l(\alpha-1)+m(\alpha'-1) \leq j(\alpha-1)+k(\alpha'-1)\big\}$.

These two claims allows us to apply Result 2 and immediately get

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

as $n \to \infty$. To conclude, recall that we have imposed the condition that $(J_\gamma + l^* - 1)/p > 2l^*$, which implies $l^* - 1 + J_\gamma > 2l^*p$ and hence $\big(n\nu[n,\infty)\big)^{l^*-1+J_\gamma} = \boldsymbol{o}\Big(\big(n\nu[n,\infty)\big)^{2l^*p}\Big)$ due to $n\nu[n,\infty) \in \mathcal{RV}_{-(\alpha-1)}(n)$ as $n \to \infty$ with $\alpha > 1$.

Now it only remains to prove claims (i) and (ii). In order to prove claim (i), we first show that $(l^* - 1 + J_\gamma, 0)$ is a feasible solution to $\{(j,k) \in \mathbb{N}^2 :\ \mathbb{D}_{j,k} \cap E(\Delta) \neq \emptyset\}$. 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

$$\xi(t) = \sum_{i=1}^{j} u_i \mathbb{1}_{[t_i,1]}(t) - \sum_{i=1}^{k} v_i \mathbb{1}_{[s_i,1]}(t). \tag{6.3}$$

Since $\big[a - \Delta - (l^*-1)b\big]/\gamma \notin \mathbb{Z}$, we must have $J_\gamma \cdot \gamma > a - \Delta - (l^*-1)b$. It then holds for all $\epsilon > 0$ small enough that $a - \Delta < J_\gamma(\gamma - \epsilon) + (l^*-1)(b-\epsilon)$. By setting $j = l^* - 1 + J_\gamma$, $k = 0$, $u_i = b - \epsilon$ for all $i \in [l^*-1]$, and $u_i = \gamma - \epsilon$ for all $i = l^*, l^*+1, \cdots, l^* - 1 + J_\gamma$ in (6.3), we get $\xi \in \mathbb{D}_{l^*-1+J_\gamma,0} \cap E(\Delta)$. This proves that $(l^* - 1 + J_\gamma, 0)$ is a feasible solution to $\{(j,k) \in \mathbb{N}^2 :\ \mathbb{D}_{j,k} \cap E(\Delta) \neq \emptyset\}$.

Furthermore, if we can show that $j \geq l^* - 1 + J_\gamma$ is the necessary condition for $\mathbb{D}_{j,k} \cap E(\Delta) \neq \emptyset$, then we get

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

Because of $\alpha, \alpha' > 1$ (see Assumption 1), we can then conclude that $\underset{(j,k)\in\mathbb{N}^2,\ \mathbb{D}_{j,k}\cap E(\Delta)\neq\emptyset}{\operatorname{argmin}} j(\alpha-1) + k(\alpha'-1) = \{(l^* - 1 + J_\gamma, 0)\}$. In other words, regarding claim (i) the only remaining task is to show that $j \geq l^* - 1 + J_\gamma$ is the necessary condition for $\mathbb{D}_{j,k} \cap E(\Delta) \neq \emptyset$. We proceed with a proof by contradiction. Suppose there is some $\xi \in \mathbb{D}_{j,k} \cap E(\Delta)$ with $j \leq l^* + J_\gamma - 2$. Then by definition of $E(\Delta)$, we know that all elements in $\{u_i :\ i \in [l^* + J_\gamma - 2]\}$ are upper bounded by b, and among them there are at most $l^* - 1$ of there are larger than γ. As a result

$$\sup_{t\in[0,1]} \xi(t) \leq \sum_{i=1}^{l^*+J_\gamma-2} u_i \leq (l^*-1)b + (J_\gamma - 1)\gamma.$$

However, since $\big[a - \Delta - (l^*-1)b\big]/\gamma \notin \mathbb{Z}$, we have $(J_\gamma - 1)\gamma < a - \Delta - (l^*-1)b < J_\gamma\gamma$, which implies $\sup_{t\in[0,1]}\xi(t) \leq (l^*-1)b + (J_\gamma-1)\gamma < a - \Delta$ and hence $\xi \notin E(\Delta)$. With this contraction obtained, we conclude the proof of claim (i).

Moving onto claim (ii), recall the bound $(J_\gamma - 1)\gamma < a - \Delta - (l^*-1)b$ we have just applied. It then holds for all $\epsilon > 0$ small enough such that

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

Fix such $\epsilon > 0$. By establishing

$$\boldsymbol{d}(\xi,\xi') \geq \epsilon \qquad \forall j = 0,1,\cdots,\ l^* + J_\gamma - 2,\ \xi \in \mathbb{D}_{j,0},\ \xi' \in E(\Delta) \tag{6.4}$$

(here $\boldsymbol{d}$ is the Skorokhod J_1 metric; see (2.1) for the definition), we get $\boldsymbol{d}\Big(\bigcup_{j\leq l^*+J_\gamma-2}\mathbb{D}_{j,0}, E(\Delta)\Big) \geq \epsilon > 0$ and conclude the proof of claim (ii). To prove (6.4), we fix some $j = 0,1,\cdots,\ l^*+J_\gamma-2,\ \xi \in \mathbb{D}_{j,0}$, and $\xi' \in E(\Delta)$. First, it suffices to show that for any λ that is an increasing homeomorphisms from $[0,1]$ to itself, we have $\sup_{t\in[0,1]}\big|\xi'(t) - \xi\big(\lambda(t)\big)\big| > \epsilon$. Specifically, we can fix some λ and set

$\widetilde{\xi}(t) = \xi\big(\lambda(t)\big)$. Using the representation in (6.3), we know there is some $(u_i)_{i\leq j} \in (0,\infty)^j$ and some $0 < \widetilde{t}_1 < \cdots < \widetilde{t}_j < 1$ such that

$$\widetilde{\xi}(t) = \sum_{i=1}^{j} u_i \mathbb{1}_{[\widetilde{u}_i,1]}(t).$$

We proceed studying the three cases that exhaust all the possibilities for the jumps u_i in $\widetilde{\xi}$. First, suppose there is some $u_i \geq b+\epsilon$. Then from the definition of $E(\Delta)$, we know that $\xi'(t) - \xi'(t-) \leq b$ for all $t \in [0,1]$, and hence $\sup_{t\in[0,1]} |\xi'(t) - \widetilde{\xi}(t)| \geq \epsilon$. Next, suppose that $u_i \in (0, b+\epsilon)$ for all $i \in [j]$, but there are at least l^* elements in $\{u_i : \ i \in [j]\}$ that are larger than or equal to $\gamma+\epsilon$. Then from the definition of $E(\Delta)$ again, we know $\sup_{t\in[0,1]} |\xi'(t) - \widetilde{\xi}(t)| \geq \epsilon$. Lastly, suppose that $u_i \in (0, b+\epsilon)$ for all $i \in [j]$, and there are at most l^*-1 elements in $\{u_i : \ i \in [j]\}$ that are larger than or equal to $\gamma+\epsilon$. Then observe that

$$\begin{aligned}
\sup_{t\in[0,1]} \widetilde{\xi}(t) &\leq \sum_{i=1}^{j} u_i \leq (l^*-1)(b+\epsilon) + \big[j - (l^*-1)\big](\gamma+\epsilon) \\
&\leq (l^*-1)(b+\epsilon) + (J_\gamma - 1)(\gamma+\epsilon) \qquad \text{due to } j \leq l^* - 2 + J_\gamma \\
&< a - \Delta - \epsilon \qquad \text{due to our choice of } \epsilon > 0.
\end{aligned}$$

Again, from the definition of $E(\Delta)$, we have $\sup_{t\in[0,1]} |\xi(t) - \widetilde{\xi}(t)| \geq \epsilon$. In summary, we have shown that $\sup_{t\in[0,1]} \big|\xi'(t) - \xi\big(\lambda(t)\big)\big| \geq \epsilon$ for any $j = 0, 1, \cdots, \ l^* + J_\gamma - 2, \ \xi \in \mathbb{D}_{j,0}, \ \xi' \in E(\Delta)$ and any λ that is an increasing homeomorphisms from $[0,1]$ to itself. This establishes (6.4) and completes the proof of claim (ii). $\square$

Now we are ready to prove Proposition 3.1.

Proof of Proposition 3.1. We start by proving the unbiasedness of the importance sampling estimator L_n. Note that

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

Recall that $\tau \sim \text{Geom}(\rho)$ under $\mathbf{Q}_n$ (that is, the importance sampling distribution $\mathbf{Q}_n$ does not alter the law of τ) and τ is independent of everything else. Furthermore, we claim that (for any $n \geq 1$ and $\rho \in (\rho_0, 1)$)

$$\sum_{m\geq 1} \mathbf{E}^{\mathbf{Q}_n} \left| \hat{Y}_n^{m-1} \mathbb{1}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} - Y_n^* \mathbb{1}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} \right|^2 \Big/ \mathbf{P}(\tau \geq m) < \infty. \tag{6.5}$$

Then due to $\mathbf{P}(\tau \geq m) = \rho^{m-1}$, we have $\mathbf{E}^{\mathbf{Q}_n} \big| \hat{Y}_n^{m-1} \mathbb{1}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} - Y_n^* \mathbb{1}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} \big|^2 \to 0$ as $m \to \infty$. This $\mathcal{L}_2$ convergence result then implies the $\mathcal{L}_1$ convergence of $\mathbf{E}^{\mathbf{Q}_n} \big| \hat{Y}_n^{m-1} \mathbb{1}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} - Y_n^* \mathbb{1}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} \big| \to 0$ as $m \to \infty$. Applying Result 3, we conclude the proof of the unbiasedness of L_n under $\mathbf{Q}_n$ (for any $n \geq 1$). Returning to claim (6.5), observe that

$$\begin{aligned}
&\mathbf{E}^{\mathbf{Q}_n} \left| \hat{Y}_n^{m-1} \mathbb{1}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} - Y_n^* \mathbb{1}_{E_n} \frac{d\mathbf{P}}{d\mathbf{Q}_n} \right|^2 \\
&= \mathbf{E}^{\mathbf{Q}_n} \left[|\hat{Y}_n^{m-1} - Y_n^*|^2 \cdot \mathbb{1}_{E_n} \cdot \left(\frac{d\mathbf{P}}{d\mathbf{Q}_n}\right)^2 \right] \leq \mathbf{E}^{\mathbf{Q}_n} \left[|\hat{Y}_n^{m-1} - Y_n^*|^2 \cdot \left(\frac{d\mathbf{P}}{d\mathbf{Q}_n}\right)^2 \right] \\
&= \mathbf{E}\left[|\hat{Y}_n^{m-1} - Y_n^*|^2 \cdot \frac{d\mathbf{P}}{d\mathbf{Q}_n} \right]
\end{aligned}$$

$$\leq \frac{1}{w}\mathbf{E}|\hat{Y}_n^{m-1} - Y_n^*|^2 \qquad \text{due to } \frac{d\mathbf{P}}{d\mathbf{Q}_n} \leq \frac{1}{w}, \text{ see (3.5).}$$

Therefore, given any $n \geq 1$, it suffices to show that $\mathbf{E}|\hat{Y}_n^m - Y_n^*|^2 = \boldsymbol{o}(\rho^m)$ as $m \to \infty$. In particular, since $\hat{Y}_n^m$ and Y_n^* only take values in $\{0,1\}$, we have $\mathbf{E}|\hat{Y}_n^m - Y_n^*|^2 = \mathbf{P}(\hat{Y}_n^m \neq Y_n^*)$. Now observe that (for any $m \geq \bar{m}$)

$$\begin{aligned}
\mathbf{P}(\hat{Y}_n^m \neq Y_n^*) &= \sum_{k\geq 0} \mathbf{P}\Big(Y_n^* \neq \hat{Y}_n^m \ \Big|\ \mathcal{D}(\bar{J}_n) = k\Big)\mathbf{P}\Big(\mathcal{D}(\bar{J}_n) = k\Big) \\
&\leq \sum_{k\geq 0} C_0\rho_0^m \cdot (k+1) \cdot \mathbf{P}\Big(\mathcal{D}(\bar{J}_n) = k\Big) \qquad \text{due to the upper bound in (3.9)} \\
&= C_0\rho_0^m \sum_{k\geq 0} \cdot (k+1) \cdot \exp\big(-n\nu[n\gamma,\infty)\big)\frac{\big(n\nu[n\gamma,\infty)\big)^k}{k!} \\
&= C_0\rho_0^m \cdot \mathbf{E}\Big[1 + \text{Poisson}\big(n\nu[n\gamma,\infty)\big)\Big] \\
&= C_0\rho_0^m \cdot \big(1 + n\nu[n\gamma,\infty)\big).
\end{aligned}$$

In particular, due to $\nu(x) \in \mathcal{RV}_{-\alpha}(x)$ as $x \to \infty$ with $\alpha > 1$, we have $n\nu[n\gamma,\infty) \in \mathcal{RV}_{-(\alpha-1)}(n)$ as $n \to \infty$, thus implying $\lim_{n\to\infty} n\nu[n\gamma,\infty) = 0$. As a result, by setting $C_0^* = C_0 \cdot \sup_{n\geq 1}\big(1 + n\nu[n\gamma,\infty)\big) < \infty$, we get $\mathbf{P}(\hat{Y}_n^m \neq Y_n^*) \leq C_0^*\rho_0^m$ for any $n \geq 1$ and $m \geq \bar{m}$. In light of the choice $\rho \in (\rho_0, 1)$, we get $\mathbf{P}(\hat{Y}_n^m \neq Y_n^*) = \boldsymbol{o}(\rho^m)$ as $m \to \infty$. This concludes the proof of claim (6.5), and hence the unbiasedness of L_n.

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

$$\mathbf{E}^{\mathbf{Q}_n}[L_n^2] = \int Z_n^2 \frac{d\mathbf{P}}{d\mathbf{Q}_n}\frac{d\mathbf{P}}{d\mathbf{Q}_n}d\mathbf{Q}_n = \int Z_n^2 \frac{d\mathbf{P}}{d\mathbf{Q}_n}d\mathbf{P} = \int Z_n^2 \mathbb{1}_{B_n^\gamma}\frac{d\mathbf{P}}{d\mathbf{Q}_n}d\mathbf{P} + \int Z_n^2 \mathbb{1}_{(B_n^\gamma)^c}\frac{d\mathbf{P}}{d\mathbf{Q}_n}d\mathbf{P}.$$

Meanwhile, from (3.5) one can see that $\frac{d\mathbf{P}}{d\mathbf{Q}_n} \leq \frac{1}{w}$ on event $(B_n^\gamma)^c$ and $\frac{d\mathbf{P}}{d\mathbf{Q}_n} \leq \frac{\mathbf{P}(B_n^\gamma)}{1-w}$ on event B_n^γ, which leads to

$$\mathbf{E}^{\mathbf{Q}_n}[L_n^2] \leq \frac{\mathbf{P}(B_n^\gamma)}{1-w}\mathbf{E}[Z_n^2\mathbb{1}_{B_n^\gamma}] + \frac{1}{w}\mathbf{E}[Z_n^2\mathbb{1}_{(B_n^\gamma)^c}]. \tag{6.6}$$

Now we claim the existence of some $c_A, C_A, C_B \in (0,\infty)$ such that

$$c_A \leq \liminf_{n\to\infty} \frac{\mathbf{P}(A_n)}{\big(n\nu[n,\infty)\big)^{l^*}} \leq \limsup_{n\to\infty} \frac{\mathbf{P}(A_n)}{\big(n\nu[n,\infty)\big)^{l^*}} \leq C_A, \tag{6.7}$$

$$\mathbf{P}(B_n^\gamma) \leq C_B \cdot \big(n\nu[n,\infty)\big)^{l^*} \qquad \forall n \geq 1. \tag{6.8}$$

Let $Z_{n,1} \triangleq Z_n\mathbb{1}_{B_n^\gamma}$ and $Z_{n,2} \triangleq Z_n\mathbb{1}_{(B_n^\gamma)^c}$. Then it suffices to prove that for any $\gamma \in (0,b)$ close enough to 0 and any $\rho \in (\rho_0, 1)$, the following claims hold (as $n \to \infty$)

$$\mathbf{E}[Z_{n,1}^2] = \boldsymbol{O}\big((n\nu[n,\infty))^{l^*}\big), \tag{6.9}$$

$$\mathbf{E}[Z_{n,2}^2] = \boldsymbol{o}\big((n\nu[n,\infty))^{2l^*}\big). \tag{6.10}$$

Indeed, using (6.8) and (6.9) we get $\mathbf{P}(B_n^\gamma)\mathbf{E}[Z_n^2\mathbb{1}_{B_n^\gamma}] = \boldsymbol{O}\big((n\nu[n,\infty))^{l^*}\big)\cdot\boldsymbol{O}\big((n\nu[n,\infty))^{l^*}\big) = \boldsymbol{O}\big((n\nu[n,\infty))^{2l^*}\big) = \boldsymbol{O}\big(\mathbf{P}^2(A_n)\big)$. The last equality follows from (6.7). Similarly, from (6.7) and (6.10) we get $\mathbf{E}[Z_n^2\mathbb{1}_{(B_n^\gamma)^c}] = \boldsymbol{o}\big((n\nu[n,\infty))^{2l^*}\big) = \boldsymbol{o}\big(\mathbf{P}^2(A_n)\big)$. Therefore, in (6.6) we have $\mathbf{E}^{\mathbf{Q}_n}[L_n^2] = \boldsymbol{O}\big(\mathbf{P}^2(A_n)\big)$, thus establishing the strong efficiency. It now remains to prove claims (6.7)(6.8)(6.9)(6.10).

Proof of Claim (6.7):

It follows directly from Lemma 6.1.

Proof of Claim (6.8):

We will make use of the following preliminary result. For any $c > 0, k \in \mathbb{N}$,

$$\mathbf{P}\Big(\text{Poisson}(c) \geq k\Big) = \sum_{j \geq k} \exp(-c)\frac{c^j}{j!} = c^k \sum_{j \geq k} \exp(-c)\frac{c^{j-k}}{j!} \leq c^k \sum_{j \geq k} \exp(-c)\frac{c^{j-k}}{(j-k)!} = c^k. \quad (6.11)$$

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

$$\begin{aligned}
\mathbf{P}(B_n^\gamma) &\triangleq \mathbf{P}\Big(\#\big\{t \in [0,n] : \ X(t) - X(t-) \geq n\gamma\big\} \geq l^*\Big) \qquad \text{due to } \bar{X}_n(t) = X(nt)/n \\
&= \sum_{k \geq l^*} \exp\big(-n\nu[n\gamma,\infty)\big)\frac{\big(n\nu[n\gamma,\infty)\big)^k}{k!} \leq \big(n\nu[n\gamma,\infty)\big)^{l^*} \qquad \text{due to (6.11)}.
\end{aligned}$$

Lastly, since the function $H_+(x) = \nu[x,\infty)$ is regularly varying as $x \to \infty$ with index $-\alpha$ (see Assumption 1), we have $\lim_{n\to\infty} \frac{(n\nu[n\gamma,\infty))^{l^*}}{(n\nu[n,\infty))^{l^*}} = 1/\gamma^{\alpha l^*} \in (0,\infty)$, and hence $\mathbf{P}(B_n^\gamma) = \boldsymbol{O}\big((n\nu[n,\infty))^{l^*}\big)$.

Proof of Claim (6.9):

By definition of Z_n in (3.8),

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

We make one more observation. For any $k = 0, 1, \cdots, l^*-1$, on event $\{\mathcal{D}(\bar{J}_n) = k\} = \big\{\#\{t \in [0,1] : \bar{X}_n(t) - \bar{X}_n(t-) \geq \gamma\} = k\big\}$, due to our choice of $\gamma \in (0,b)$ we have $\#\{t \in [0,1] : \ \bar{X}_n(t) - \bar{X}_n(t-) > b\} < l^*$. As a result, we must have $\mathbb{1}_{B_n^\gamma} = 0$, and hence $Z_{n,1} = 0$. By applying Result 3, we yield

$$\begin{aligned}
&\mathbf{E}Z_{n,1}^2 \\
&\leq \sum_{m \geq 1} \frac{\mathbf{E}\Big[\big|Y_n^* \mathbb{1}_{E_n \cap B_n^\gamma} - \hat{Y}_n^{m-1}\mathbb{1}_{E_n \cap B_n^\gamma}\big|^2\Big]}{\mathbf{P}(\tau \geq m)} \\
&= \sum_{m \geq 1} \frac{\mathbf{E}\Big[\big|Y_n^* - \hat{Y}_n^{m-1}\big|^2 \mathbb{1}_{E_n \cap B_n^\gamma}\Big]}{\mathbf{P}(\tau \geq m)} \leq \sum_{m \geq 1} \frac{\mathbf{E}\Big[\big|Y_n^* - \hat{Y}_n^{m-1}\big|^2 \mathbb{1}_{B_n^\gamma}\Big]}{\mathbf{P}(\tau \geq m)} \\
&= \sum_{m \geq 1} \frac{\mathbf{E}\Big[\mathbb{1}\big(Y_n^* \neq \hat{Y}_n^{m-1}\big)\cdot \mathbb{1}_{B_n^\gamma}\Big]}{\mathbf{P}(\tau \geq m)} \qquad \text{because } \hat{Y}_n^m \text{ and } Y_n^* \text{ only take values in } \{0,1\} \\
&= \sum_{m \geq 1}\sum_{k \geq 0} \frac{\mathbf{E}\Big[\mathbb{1}\big(Y_n^* \neq \hat{Y}_n^{m-1}\big)\cdot \mathbb{1}_{B_n^\gamma} \ \Big|\ \{\mathcal{D}(\bar{J}_n) = k\}\Big]}{\mathbf{P}(\tau \geq m)} \cdot \mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big) \\
&= \sum_{m \geq 1}\sum_{k \geq l^*} \frac{\mathbf{E}\Big[\mathbb{1}\big(Y_n^* \neq \hat{Y}_n^{m-1}\big)\cdot \mathbb{1}_{B_n^\gamma} \ \Big|\ \{\mathcal{D}(\bar{J}_n) = k\}\Big]}{\mathbf{P}(\tau \geq m)} \cdot \mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big) \ \text{ due to } \mathbb{1}_{B_n^\gamma} = 0 \text{ on } \{\mathcal{D}(\bar{J}_n) < l^*\} \\
&\leq \sum_{k \geq l^*} \mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big) \cdot \sum_{m \geq 1} \frac{\mathbf{P}\Big(Y_n^* \neq \hat{Y}_n^{m-1} \ \Big|\ \{\mathcal{D}(\bar{J}_n) = k\}\Big)}{\mathbf{P}(\tau \geq m)} \\
&\leq \sum_{k \geq l^*} \mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big) \cdot \left[\sum_{m=1}^{\bar{m}} \frac{1}{\rho^{m-1}} + \sum_{m \geq \bar{m}+1} \frac{C_0\rho_0^{m-1}\cdot(k+1)}{\rho^{m-1}}\right] \qquad \text{due to (3.9) and } \tau \sim \text{Geom}(\rho)
\end{aligned}$$

$$\leq \sum_{k\geq l^*}\mathbf{P}\big(\mathcal{D}(\bar{J}_n)=k\big)\cdot(k+1)\cdot\underbrace{\Big[\sum_{m=1}^{\bar{m}}\frac{1}{\rho^{m-1}}+\sum_{m\geq \bar{m}+1}\frac{C_0\rho_0^{m-1}}{\rho^{m-1}}\Big]}_{\triangleq \widetilde{C}_{\rho,1}}.$$

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

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

Lastly, due to the regular varying nature of $H_+(x)=\nu[x,\infty)$, we can conclude that $\mathbf{E}Z_{n,1}^2=\boldsymbol{O}\big((n\nu[n,\infty))^{l^*}\big)$ as $n\to\infty$.

Proof of Claim (6.10):

We start by fixing some $\rho\in(\rho_0,1)$ and specifying a few constants. Due to $\rho_0<\rho<1$, one can find some $q>1$ such that $\rho_0^{1/q}<\rho$. Let $p>1$ be such that $\frac{1}{p}+\frac{1}{q}=1$. Due to Assumption 3, we have $a>(l^*-1)b$. By picking any $\gamma\in(0,b)$ small enough, we have $(\hat{J}_\gamma+l^*-1)/p>2l^*$ where

$$\hat{J}_\gamma=\frac{a-\Delta-(l^*-1)b}{\gamma}.$$

Then one can pick some $\Delta>0$ small enough such that $\big[a-\Delta-(l^*-1)b\big]/\gamma\notin\mathbb{Z}$, $a-\Delta>(l^*-1)b$, and $(J_\gamma+l^*-1)/p>2l^*$ where

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

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

$$Z_{n,2}\triangleq Z_n\mathbb{1}_{(B_n^\gamma)^c}=\underbrace{Z_n\mathbb{1}_{A_n^\Delta\cap(B_n^\gamma)^c}}_{\triangleq Z_{n,3}}+\underbrace{Z_n\mathbb{1}_{(A_n^\Delta)^c\cap(B_n^\gamma)^c}}_{\triangleq Z_{n,4}}.$$

For term $Z_{n,3} = \sum_{m=0}^{\log_2(n^2)+\tau} \frac{\hat{Y}_n^m \mathbb{1}_{A_n^\Delta \cap E_n \cap (B_n^\gamma)^c} - \hat{Y}_n^{m-1}\mathbb{1}_{A_n^\Delta \cap E_n \cap (B_n^\gamma)^c}}{\mathbf{P}(\tau \geq m)}$, applying Result 3 we get

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

Applying Lemma 6.2, we get $\mathbf{P}\big(A_n^\Delta \cap E_n \cap (B_n^\gamma)^c\big) = \boldsymbol{O}\Big(\big(n\nu[n,\infty)\big)^{2pl^*}\Big)$ as $n \to \infty$. As a result, $\Big(\mathbf{P}\big(A_n^\Delta \cap E_n \cap (B_n^\gamma)^c\big)\Big)^{1/p} = \boldsymbol{o}\Big(\big(n\nu[n,\infty)\big)^{2l^*}\Big)$ as $n \to \infty$. On the other hand, for any $n \geq 1$ and $m \geq \bar{m}$,

$$\begin{aligned}
&\mathbf{P}\big(Y_n^* \neq \hat{Y}_n^{m-1}\big) \\
&= \sum_{k\geq 0} \mathbf{P}\big(Y_n^* \neq \hat{Y}_n^{m-1} \bigm| \mathcal{D}(\bar{J}_n) = k\big)\mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big)\big) \\
&\leq \sum_{k\geq 0} C_0\rho_0^{m-1}\cdot (k+1)\cdot \exp\big(-n\nu[n\gamma,\infty)\big)\frac{\big(n\nu[n\gamma,\infty)\big)^k}{k!} \qquad \text{using (3.9)} \\
&= C_0\rho_0^{m-1}\cdot \Bigg[\sum_{k\geq 1}\exp\big(-n\nu[n\gamma,\infty)\big)\frac{\big(n\nu[n\gamma,\infty)\big)^k}{(k-1)!} + \sum_{k\geq 0}\exp\big(-n\nu[n\gamma,\infty)\big)\frac{\big(n\nu[n\gamma,\infty)\big)^k}{k!}\Bigg] \\
&= C_0\rho_0^{m-1}\cdot \Bigg[n\nu[n\gamma,\infty)\cdot\sum_{k\geq 1}\exp\big(-n\nu[n\gamma,\infty)\big)\frac{\big(n\nu[n\gamma,\infty)\big)^{k-1}}{(k-1)!} + \sum_{k\geq 0}\exp\big(-n\nu[n\gamma,\infty)\big)\frac{\big(n\nu[n\gamma,\infty)\big)^k}{k!}\Bigg] \\
&= C_0\rho_0^{m-1}\cdot\big(n\nu[n\gamma,\infty)+1\big).
\end{aligned}$$

Since $n\nu[n\gamma,\infty) \in \mathcal{RV}_{-(\alpha-1)}$ as $n \to \infty$ and $\alpha > 1$, we have $\lim_{n\to\infty} n\nu[n\gamma,\infty) = 0$. This allows us to pick some $C_\gamma < \infty$ such that $\sup_{n\geq 1} n\nu[n\gamma,\infty) + 1 \leq C_\gamma$. In short, we have obtained that $\Big(\mathbf{P}\big(Y_n^* \neq \hat{Y}_n^{m-1}\big)\Big)^{1/q} \leq (C_0C_\gamma)^{1/q}\cdot(\rho_0^{1/q})^{m-1}$ for all $n \geq 1, m \geq \bar{m}$. Now we get

$$\begin{aligned}
\mathbf{E}Z_{n,3}^2 &\leq \boldsymbol{o}\Big(\big(n\nu[n,\infty)\big)^{2l^*}\Big)\cdot\Bigg[\sum_{m=1}^{\bar{m}}\frac{1}{\mathbf{P}(\tau\geq m)} + \sum_{m\geq\bar{m}+1}\frac{(C_0C_\gamma)^{1/q}\cdot(\rho_0^{1/q})^{m-1}}{\mathbf{P}(\tau\geq m)}\Bigg] \\
&= \boldsymbol{o}\Big(\big(n\nu[n,\infty)\big)^{2l^*}\Big)\cdot\underbrace{\Bigg[\sum_{m=1}^{\bar{m}}\frac{1}{\rho^{m-1}} + \sum_{m\geq\bar{m}+1}\frac{(C_0C_\gamma)^{1/q}\cdot(\rho_0^{1/q})^{m-1}}{\rho^{m-1}}\Bigg]}_{\triangleq \widetilde{C}_{\rho,2}} \qquad \text{due to } \tau \sim \mathrm{Geom}(\rho) \\
&\qquad\qquad \text{note that } \widetilde{C}_{\rho,2} < \infty \text{ due to } \rho \in (\rho_0, 1) \\
&= \widetilde{C}_{\rho,2}\cdot\boldsymbol{o}\Big(\big(n\nu[n,\infty)\big)^{2l^*}\Big) = \boldsymbol{o}\Big(\big(n\nu[n,\infty)\big)^{2l^*}\Big). \qquad (6.12)
\end{aligned}$$

The last line follows from our choice of $\rho_0^{1/q} < \rho$. Moving on, we can apply Result 3 again and bound the second-order moment of term $Z_{n,4} = \sum_{m=0}^{\log_2(n^2)+\tau} \frac{\hat{Y}_n^m \mathbb{1}_{(A_n^\Delta)^c \cap E_n \cap (B_n^\gamma)^c} - \hat{Y}_n^{m-1} \mathbb{1}_{(A_n^\Delta)^c \cap E_n \cap (B_n^\gamma)^c}}{\mathbf{P}(\tau \geq m)}$ with

$$
\begin{aligned}
&\mathbf{E}Z_{n,4}^2 \\
&\leq \sum_{m\geq 1} \frac{\mathbf{E}\Big[\big|Y_n^* \mathbb{1}_{(A_n^\Delta)^c \cap E_n \cap (B_n^\gamma)^c} - \hat{Y}_n^{m-1} \mathbb{1}_{(A_n^\Delta)^c \cap E_n \cap (B_n^\gamma)^c}\big|^2\Big]}{\mathbf{P}(\tau \geq m)} = \sum_{m\geq 1} \frac{\mathbf{E}\Big[\big|Y_n^* - \hat{Y}_n^{m-1}\big|^2 \mathbb{1}_{(A_n^\Delta)^c \cap E_n \cap (B_n^\gamma)^c}\Big]}{\mathbf{P}(\tau \geq m)} \\
&= \sum_{m\geq 1} \frac{\mathbf{E}\Big[\mathbb{1}\big(Y_n^* \neq \hat{Y}_n^{m-1}\big) \cdot \mathbb{1}_{(A_n^\Delta)^c \cap E_n \cap (B_n^\gamma)^c}\Big]}{\mathbf{P}(\tau \geq m)} \qquad \text{because } \hat{Y}_n^m \text{ and } Y_n^* \text{ only take values in } \{0,1\} \\
&\leq \sum_{m\geq 1} \frac{\mathbf{P}\Big(\big\{Y_n^* \neq \hat{Y}_n^{m-1},\ \bar{X}_n \notin A^\Delta\big\} \cap (B_n^\gamma)^c\Big)}{\mathbf{P}(\tau \geq m)} \qquad \text{due to } A_n^\Delta = \{\bar{X}_n \in A^\Delta\} \\
&= \sum_{m\geq 1} \frac{\mathbf{P}\Big(\big\{Y_n^* \neq \hat{Y}_n^{m-1},\ \bar{X}_n \notin A^\Delta\big\} \cap \{\mathcal{D}(\bar{J}_n) < l^*\}\Big)}{\mathbf{P}(\tau \geq m)} \\
&\qquad\qquad\qquad\qquad \text{due to } B_n^\gamma = \{\bar{X}_n \in B^\gamma\} = \{\bar{J}_n \in B^\gamma\} = \{\mathcal{D}(\bar{J}_n) \geq l^*\} \\
&= \sum_{m\geq 1} \sum_{k=0}^{l^*-1} \frac{\mathbf{P}\big(Y_n^* \neq \hat{Y}_n^{m-1},\ \bar{X}_n \notin A^\Delta \mid \{\mathcal{D}(\bar{J}_n) = k\}\big)}{\mathbf{P}(\tau \geq m)} \cdot \mathbf{P}\big(\mathcal{D}(\bar{J}_n) = k\big) \\
&\leq \sum_{m\geq 1} \sum_{k=0}^{l^*-1} \frac{C_0 \rho_0^{m-1}}{\Delta^2 n^\mu \cdot \rho^{m-1}} \qquad \text{due to (3.10)} \\
&= l^* \sum_{m\geq 1} \frac{C_0 \rho_0^{m-1}}{\Delta^2 n^\mu \cdot \rho^{m-1}} = \frac{C_0 l^*}{\Delta \cdot (1 - \frac{\rho_0}{\rho})} \cdot \frac{1}{n^\mu}.
\end{aligned}
$$

Lastly, recall the condition that $\mu > 2l^*(\alpha - 1)$. Since $n\nu[n,\infty) \in \mathcal{RV}_{-(\alpha-1)}$ as $n \to \infty$, we have $\big(n\nu[n,\infty)\big)^{2l^*} \in \mathcal{RV}_{-2l^*(\alpha-1)}$. As a result, $1/n^\mu = \boldsymbol{o}\Big(\big(n\nu[n,\infty)\big)^{2l^*}\Big)$ and hence

$$
\mathbf{E}Z_{n,4}^2 = \boldsymbol{o}\Big(\big(n\nu[n,\infty)\big)^{2l^*}\Big) \qquad \text{as } n \to \infty. \tag{6.13}
$$

Combining (6.12)(6.13) with the preliminary bound $(x+y)^2 \leq 2x^2 + 2y^2$, we yield $\mathbf{E}Z_{n,2}^2 \leq 2\mathbf{E}Z_{n,3}^2 + 2\mathbf{E}Z_{n,4}^2 = \boldsymbol{o}\Big(\big(n\nu[n,\infty)\big)^{2l^*}\Big)$ and conclude the proof of (6.10). $\square$

6.2 Proofs of Propositions 3.2 and 3.3

The proof of Propositions 3.2 and 3.3 will be based on the technical tools developed below. First, we collect two useful results.

Result 5 (Lemma 9 in [30])**.** *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 for the Blumenthal-Getoor index $\beta \triangleq \inf\{p > 0 :\ I_0^p(\nu) < \infty\}$ we have $\beta < 2$. Then for any $\beta_+ \in (\beta, 2)$,*

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

The next result can be obtained by setting $T = t$ in Lemma 10 in [30].

Result 6 (Lemma 10 in [30]). *Let X be a Lévy process with generating triplet (c, σ, ν). Suppose that $I_+^1(\nu) \triangleq \int_{[1,\infty)} x\nu(dx) < \infty$ and for the Blumenthal-Getoor index $\beta \triangleq \inf\{p > 0 : I_0^p(\nu) < \infty\}$ we have $\beta < 2$. For any $t > 0, \beta_+ \in (1 \vee \beta, 2)$,*

$$\mathbf{E}\big[M(t)\big] \leq \Big(|\sigma|\sqrt{\frac{2}{\pi}} + 2\sqrt{I_0^{\beta_+}(\nu)}\Big)\sqrt{t} + \Big(c^+ + I_+^1(\nu) + 2I_0^{\beta_+}(\nu)\Big)t$$

where $M(t) \triangleq \sup_{s \leq t} X(s)$ and $I_0^p(\nu) \triangleq \int_{(-1,1)} |x|^p \nu(dx)$.

Next, we study the expectations regarding the supremum of Ξ_n (see (3.6) for the definition) and the difference between Ξ_n and $\breve{\Xi}_n^m$ (see (3.17)).

Lemma 6.3. *There exists a constant $C_X < \infty$ only depending on the law of Lévy process X such that*

$$\mathbf{E}\Big[\sup_{s\in[0,t]} \Xi_n(t)\Big] \leq C_X(\sqrt{t} + t) \qquad \forall t > 0,\ n \geq 1.$$

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

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

where $x^+ = x \vee 0$ for any $x \in \mathbb{R}$ and $I_+^1(\nu) \triangleq \int_{[1,\infty)} x\nu(dx), I_0^p(\nu) \triangleq \int_{(-1,1)} |x|^p \nu(dx)$. Meanwhile, recall that for any $n \geq 1, \gamma > 0$, the process 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 Result 6 that $\mathbf{E}\sup_{s\in[0,t]} \Xi_n(t) \leq \Big(|\sigma|\sqrt{\frac{2}{\pi}} + 2\sqrt{I_0^{\beta_+}(\nu_n)}\Big)\sqrt{t} + \Big(c_X^+ + I_+^1(\nu_n) + 2I_0^{\beta_+}(\nu_n)\Big)t$ holds for any $t > 0, n \geq 1$. In particular, since $I_0^{\beta_+}(\nu_n) = \int_{(-1,1)} |x|^p \nu_n(dx) = \int_{(-1,1)\cap(-\infty,n\gamma)} |x|^p \nu(dx) \leq I_0^{\beta_+}(\nu)$ and $I_+^1(\nu_n) = \int_{[1,\infty)} x\nu_n(dx) = \int_{[1,\infty)\cap(-\infty,n\gamma)} x\nu(dx) = I_+^1(\nu)$, we obtain

$$\mathbf{E}\Big[\sup_{s\in[0,t]} \Xi_n(t)\Big] \leq C_X(\sqrt{t} + t) \qquad \forall t > 0,\ n \geq 1$$

and conclude the proof. □

Lemma 6.4. *Let $\beta_+ \in (\beta, 2)$ where $\beta < 2$ is the Blumenthal-Getoor index (see Assumption 1). There exists some $C_{\beta_+} \in (0, \infty)$ that only depends on β_+ the law of Lévy process X such that*

$$\mathbf{P}\Big(\sup_{t\in[0,n]} \Big|\Xi_n(t) - \breve{\Xi}_n^m(t)\Big| > c\Big) \leq \frac{C_{\beta_+}\kappa^{m(2-\beta_+)}}{c^2 \cdot n^{2(r-\beta_+)-1}} \qquad \forall c > 0, n \geq 1$$

where r is the parameter in the truncation threshold $\kappa_{n,m} = \kappa^m/n^r$ (see (3.14)).

Proof. From the definitions of Ξ_n and $\breve{\Xi}_n^m$ in (3.15)(3.17),

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

where (for any $c \in (0, 1]$) $X^{(-c,c)}$ 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 $var X^{(-\kappa_{n,m},\kappa_{n,m})}(1) = \bar{\sigma}^2(\kappa_{n,m})$; here $\bar{\sigma}^2(c) \triangleq \int_{(-c,c)} x^2 \nu(dx)$ is defined in (3.16). Therefore,

$$\mathbf{P}\Big(\sup_{t\in[0,n]} \Big|\Xi_n(t) - \breve{\Xi}_n^m(t)\Big| > c\Big)$$

$$
\begin{aligned}
&\leq \frac{1}{c^2}\mathbf{E}\Big|\Xi_n(n) - \breve{\Xi}^m_n(n)\Big|^2 = \frac{1}{c^2}\mathbf{E}\Big|X^{(-\kappa_{n,m},\kappa_{n,m})}(n) - \bar{\sigma}(\kappa_{n,m})B(n)\Big|^2 \qquad \text{using Doob's inequality}\\
&= \frac{2n}{c^2}\bar{\sigma}^2(\kappa_{n,m}) \qquad \text{because } X^{(-\kappa_{n,m},\kappa_{n,m})} \text{ and } B \text{ are independent}\\
&\leq \frac{2n}{c^2}\cdot \kappa_{n,m}^{2-\beta_+} I_0^{\beta_+}(\nu) \qquad \text{using Result 5}\\
&= \frac{2I_0^{\beta_+}(\nu)}{c^2}\cdot \frac{n\kappa^{m(2-\beta_+)}}{n^{r(2-\beta_+)}} = \frac{2I_0^{\beta_+}(\nu)}{c^2}\cdot \frac{\kappa^{m(2-\beta_+)}}{n^{r(2-\beta_+)-1}} \qquad \text{due to } \kappa_{n,m} = \kappa^m/n^r.
\end{aligned}
$$

To conclude the proof, note that the constant $C_{\beta_+} = 2I_0^{\beta_+}(\nu) = 2\int_{(-1,1)}\int |x|^{\beta_+}\nu(dx)$ only depends on β_+. □

To present the next few lemmas we introduce a slightly more general version of the stick-breaking procedure described in (3.18)(3.19)(3.20). For any $l > 0$, let

$$
l_1(l) = V_1 \cdot l, \tag{6.14}
$$

$$
l_j(l) = V_j \cdot (l - l_1 - l_2 - \cdots - l_{j-1}) \ \forall j \geq 2 \tag{6.15}
$$

where V_j is an iid sequence of Unif$(0,1)$ RVs. For any $n \geq 1$, let $\Xi_n, \breve{\Xi}^m_n$ be L evy processes with joint law specified in (3.15)(3.17) and let them be independent of the sequence V_j. Conditioning on the values of $l_j(l)$, define $\xi^{[n]}_j(l), \xi^{[n];m}_j(l)$ using

$$
\big(\xi^{[n]}_j(l), \xi^{[n],1}_j(l), \xi^{[n],2}_j(l), \xi^{[n],3}_j(l), \cdots\big) = \Big(\Xi_n\big(l_j(l)\big), \breve{\Xi}^1_n\big(l_j(l)\big), \breve{\Xi}^2_n\big(l_j(l)\big), \breve{\Xi}^3_n\big(l_j(l)\big), \cdots\Big) \ \forall j \geq 1. \tag{6.16}
$$

Lemma 6.5. *Let $n \in \mathbb{Z}_+$ and $l \in [0,n]$. Let $l_j(l)$ and $\xi^{[n]}_j(l), \xi^{[n],m}_j(l)$ be defined as in (6.14)-(6.16). Let $\beta_+ \in (\beta, 2)$ where $\beta < 2$ is the Blumenthal-Getoor index (see Assumption 1). There exists some $C_{\beta_+} \in (0,\infty)$ that only depends on β_+ the law of Lévy process X such that*

$$
\mathbf{P}\Big(\Big|\sum_{j=1}^{m+\lceil \log_2(n^d)\rceil}\big(\xi^{[n]}_j(l)\big)^+ - \sum_{j=1}^{m+\lceil \log_2(n^d)\rceil}\big(\xi^{[n],m}_j(l)\big)^+\Big| > c\Big) \leq \frac{C_{\beta_+}\kappa^{m(2-\beta_+)}}{c^2\cdot n^{2(r-\beta_+)-1}} \qquad \forall c,d > 0, m \geq 0, n \geq 1
$$

where r is the parameter in the truncation threshold $\kappa_{n,m} = \kappa^m/n^r$ (see (3.14)).

Proof. Let $k(n) = \lceil \log_2(n^d)\rceil$. First, due to $|x^+ - y^+| \leq |x - y|$,

$$
\begin{aligned}
&\mathbf{P}\Big(\Big|\sum_{j=1}^{m+k(n)}\big(\xi^{[n]}_j(l)\big)^+ - \sum_{j=1}^{m+k(n)}\big(\xi^{[n],m}_j(l)\big)^+\Big| > c\Big)\\
&\leq \mathbf{P}\Big(\sum_{j=1}^{m+k(n)}\Big|\big(\xi^{[n]}_j(l)\big)^+ - \big(\xi^{[n],m}_j(l)\big)^+\Big| > c\Big) \leq \mathbf{P}\Big(\sum_{j=1}^{m+k(n)}\Big|\underbrace{\xi^{[n]}_j(l) - \xi^{[n],m}_j(l)}_{\triangleq q_j}\Big| > c\Big).
\end{aligned} \tag{6.17}
$$

Let $\chi = 2^{1/4}$. Due to $\frac{1}{\chi} + \frac{1}{\chi^2} + \cdots + \frac{1}{\chi^{m+k(n)}} \leq \frac{1}{\chi - 1}$, we have

$$
1 \geq (\chi - 1)\Big(\frac{1}{\chi} + \frac{1}{\chi^2} + \cdots + \frac{1}{\chi^{m+k(n)}}\Big).
$$

As a result,

$$
\mathbf{P}\Big(\sum_{j=1}^{k(n)+m}|q_j| > c\Big) \leq \mathbf{P}\Big(\sum_{j=1}^{k(n)+m}|q_j| > c(\chi-1)\big(\sum_{j=1}^{k(n)+m}\frac{1}{\chi^j}\big)\Big)
$$

$$\leq \sum_{j=1}^{k(n)+m} \mathbf{P}\Big(|q_j| > c\frac{\chi-1}{\chi^j}\Big) \tag{6.18}$$

To bound (6.18) we fix $j \in \big[\lceil \log_2(n^2) \rceil + m\big]$ and study $\mathbf{P}(|q_j| > c\frac{\chi-1}{\chi^j})$. Conditioning on $l_j(l) = t$ (for any $t \in [0,l]$), we get

$$\begin{aligned}
&\mathbf{P}\Big(|q_j| > c\frac{\chi-1}{\chi^j}\ \Big|\ l_j(l) = t\Big) \\
&= \mathbf{P}\Big(\Big|\Xi_n(t) - \breve{\Xi}_n^m(t)\Big| > c\frac{\chi-1}{\chi^j}\Big) \qquad \text{due to (6.16)} \\
&\leq \frac{\chi^{2j}}{c^2(\chi-1)^2}\mathbf{E}\Big|\Xi_n(t) - \breve{\Xi}_n^m(t)\Big|^2 = \frac{\chi^{2j}}{c^2(\chi-1)^2}\mathbf{E}\Big|X^{(-\kappa_{n,m},\kappa_{n,m})}(t) - \bar{\sigma}(\kappa_{n,m})B(t)\Big|^2
\end{aligned}$$

where (for any $y \in (0,1]$) $X^{(-y,y)}$ is the Lévy process with generating triplet $(0,0,\nu|_{(-y,y)})$, $\kappa_{n,m} = \kappa^m/n^r$, 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 $var X^{(-\kappa_{n,m},\kappa_{n,m})}(1) = \bar{\sigma}^2(\kappa_{n,m})$; here $\bar{\sigma}^2(c) \triangleq \int_{(-c,c)} x^2\nu(dx)$ is defined in (3.16). This leads to $\mathbf{E}\big|X^{(-\kappa_{n,m},\kappa_{n,m})}(t) - \bar{\sigma}(\kappa_{n,m})B(t)\big|^2 = 2t \cdot \bar{\sigma}^2(\kappa_{m,n})$. We now have, unconditionally,

$$\begin{aligned}
\mathbf{P}\Big(|q_j| > c\frac{\chi-1}{\chi^j}\Big) &= \frac{\chi^{2j}}{c^2(\chi-1)^2} \cdot 2\bar{\sigma}^2(\kappa_{m,n}) \cdot \mathbf{E}[l_j(l)] \\
&= \frac{\sqrt{2^j}}{c^2(2^{1/4}-1)^2} \cdot 2\bar{\sigma}^2(\kappa_{m,n}) \cdot \mathbf{E}[l_j(l)] \qquad \text{due to } \chi = 2^{1/4} \\
&= \frac{\sqrt{2^j}}{c^2(2^{1/4}-1)^2} \cdot 2\bar{\sigma}^2(\kappa_{m,n}) \cdot \frac{l}{2^j} \qquad \text{by definition of } l_j(l) \text{ in (6.14)(6.15)} \\
&\leq \frac{2n}{c^2(2^{1/4}-1)^2\sqrt{2^j}} \cdot \bar{\sigma}^2(\kappa_{m,n}) \qquad \text{due to } l \leq n.
\end{aligned}$$

Therefore, in (6.17) we obtain

$$\begin{aligned}
&\mathbf{P}\Big(\Big|\sum_{j=1}^{m+t(n)} (\xi_j^{[n]}(l))^+ - \sum_{j=1}^{m+t(n)} (\xi_j^{[n],m}(l))^+\Big| > c\Big) \\
&\leq \bar{\sigma}^2(\kappa_{m,n}) \cdot \sum_{j=1}^{m+t(n)} \frac{2n}{c^2(2^{1/4}-1)^2\sqrt{2^j}} \leq \bar{\sigma}^2(\kappa_{m,n}) \cdot \sum_{j\geq 0} \frac{2n}{c^2(2^{1/4}-1)^2\sqrt{2^j}} \\
&= n\bar{\sigma}^2(\kappa_{m,n}) \cdot \frac{2\sqrt{2}}{c^2(2^{1/4}-1)^2(\sqrt{2}-1)}.
\end{aligned}$$

Lastly, using Result 5, we have

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

where $I_0^{\beta_+}(\nu) = \int_{(-1,1)} \int |x|^{\beta_+}\nu(dx)$. To conclude the proof, note that $I_0^{\beta_+}(\nu)$ only depends on β_+. $\square$

Lemma 6.6. *Let $n \in \mathbb{Z}_+$ and $l \in [0,n]$. Let $C_X < \infty$ be the constant characterized in Lemma 6.3 that only depends on the law of Lévy process X. The inequality*

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

holds for all $c, d > 0$, $n \geq 1$, and $m \geq 0$.

Proof. Conditioning on $l_{m+\lceil \log_2(n^d)\rceil}(l) = t$, it follows directly from Result 4 that

$$\mathscr{L}\Big(\sum_{j>m+\lceil \log_2(n^d)\rceil} \big(\xi_j^{[n]}(l)\big)^+ \ \Big|\ l_{m+\lceil \log_2(n^d)\rceil}(l) = t\Big) = \mathscr{L}\big(\sup_{s\in[0,t]} \Xi_n\big).$$

It then follows from Markov property and the Lemma 6.3 that

$$\mathbf{P}\Big(\sum_{j>m+\lceil \log_2(n^d)\rceil} \big(\xi_j^{[n]}(l)\big)^+ > c\ \Big|\ l_{m+\lceil \log_2(n^d)\rceil}(l) = t\Big) \leq \frac{C_X}{c}(\sqrt{t}+t) \qquad \forall c,t>0.$$

Therefore, unconditionally,

$$\begin{aligned}\mathbf{P}\Big(\sum_{j>m+\lceil \log_2(n^d)\rceil} \big(\xi_j^{[n]}(l)\big)^+ > c\Big) &\leq \frac{C_X}{c}\mathbf{E}\Big[\sqrt{l_{m+\lceil \log_2(n^d)\rceil}(l)} + l_{m+\lceil \log_2(n^d)\rceil}(l)\Big] \\ &\leq \frac{C_X}{c}\Big[\sqrt{\mathbf{E}l_{m+\lceil \log_2(n^d)\rceil}(l)} + \mathbf{E}l_{m+\lceil \log_2(n^d)\rceil}(l)\Big]\end{aligned}$$

because of Jensen's inequality, i.e., $\mathbf{E}\sqrt{W} \leq \sqrt{\mathbf{E}W}$ for any non-negative random variable W. Lastly, by definition of $l_j(l)$ in (6.14)(6.15), $\mathbf{E}l_{m+\lceil \log_2(n^d)\rceil}(l) \leq \frac{l}{2^m\cdot n^d} \leq \frac{n}{2^m\cdot n^d} = \frac{1}{n^{d-1}\cdot 2^m}$; here we applied $l \leq n$. This concludes the proof. □

Lemma 6.7. *Let $n \in \mathbb{Z}_+$ and $l \in [0,n]$. Let $C, z_0, \lambda > 0, \theta \in (0,1]$ be the constants in Assumption 2. Let $C_X < \infty$ be the constant characterized in Lemma 6.3 that only depends on the law of Lévy process X. Given $y_0, \delta, \alpha_3, \alpha_4 > 0$, the inequality*

$$\mathbf{P}\Big(\sum_{j=1}^{m+\lceil \log_2(n^d)\rceil} \big(\xi_j^{[n]}(l)\big)^+ \in [y, y+c]\Big) \leq C\frac{(m + (\lceil \log_2(n^d)\rceil) n^{\alpha_4\lambda}}{\delta^{\alpha_3\lambda}}c^\theta + 4C_X\big(m^2 + (\lceil \log_2(n^d)\rceil)^2\big)\frac{\delta^{\alpha_3/2}}{y_0\cdot n^{\alpha_4/2}}.$$

holds for all $y \geq y_0$ and $c,d > 0$.

Proof. To simplify notations, let $k(n) \triangleq \lceil \log_2(n^d)\rceil$ and write $l_j \triangleq l_j(l)$ when there is no ambiguity. For the sequence of random variables $(l_1,\cdots,l_{m+k(n)})$, let $\tilde{l}_1 \geq \tilde{l}_2 \geq \cdots \geq \tilde{l}_{m+k(n)}$ be the order statistics. For any $t_1 \geq t_2 \geq \cdots \geq t_{m+k(n)} > 0$, by conditioning on $\tilde{l}_j = t_j\ \forall j \in [m+k(n)]$, it follows from (6.16) that

$$\mathscr{L}\Big(\sum_{j=1}^{m+k(n)} \big(\xi_j^{[n]}(l)\big)^+ \ \Big|\ \tilde{l}_j = t_j\ \forall j \in [m+k(n)]\Big) = \mathscr{L}\Big(\sum_{j=1}^{m+k(n)} \big(\Xi'_{n,j}(t_j)\big)^+\Big)$$

where $\Xi'_{n,j}$ are iid copies of the Lévy processes $\Xi_n = X^{<n\gamma}$. Next, fix

$$\eta = \delta^{m\alpha_3}/n^{\alpha_4}.$$

Also, given the ordered sequence t_j, we define $J \triangleq \#\{j \in [m+k(n)]:\ \tilde{t}_j \geq \eta\}$ as the number elements in $t_1 \geq t_2 \geq \cdots \geq t_{m+k(n)}$ that are no less than η. Note that if $t_1 < \eta$ we have $J = 0$. By considering a decomposition of event based on the first j such that $\Xi'_{n,j}(t_j) > 0$ (and hence $\big(\Xi'_{n,j}(t_j)\big)^+ > 0$), we get

$$
\begin{aligned}
&\mathbf{P}\Big(\sum_{j=1}^{m+k(n)}\big(\xi^{[n]}_j(l)\big)^+\in[y,y+c]\ \Big|\ \tilde{l}_j=t_j\ \forall j\in[m+k(n)]\Big)\\
&=\sum_{j=1}^{J}\underbrace{\mathbf{P}\Big(\Xi'_{n,i}(t_i)\le 0\ \forall i\in[j-1];\ \Xi'_{n,j}(t_j)>0;\ \sum_{i=j}^{m+k(n)}\big(\Xi'_{n,i}(t_i)\big)^+\in[y,y+c]\Big)}_{\triangleq p_j}\\
&\quad+\underbrace{\mathbf{P}\Big(\Xi'_{n,i}(t_i)\le 0\ \forall i\in[J];\ \sum_{j=J+1}^{m+k(n)}\big(\Xi'_{n,j}(t_j)\big)^+\in[y,y+c]\Big)}_{\triangleq p_*}.
\end{aligned}
\tag{6.19}
$$

We first bound the term p_j. For any $j\in[J]$, observe that

$$
\begin{aligned}
p_j&\le \mathbf{P}\Big(\Xi'_{n,j}(t_j)>0;\ \sum_{i=j}^{m+k(n)}\big(\Xi'_{n,i}(t_i)\big)^+\in[y,y+c]\Big)\\
&=\int_{\mathbb{R}}\mathbf{P}\Big(\Xi'_{n,j}(t_j)\in[y-x,y-x+c]\cap(0,\infty)\Big)\mathbf{P}\Big(\sum_{i=j+1}^{m+k(n)}\big(\Xi'_{n,i}(t_i)\big)^+\in dx\Big)\\
&\le \frac{C}{t_j^{\lambda}\wedge 1}c^{\theta}\qquad\text{due to Assumption 2}\\
&\le \frac{Cn^{\alpha_4\lambda}}{\delta^{m\alpha_3\lambda}}c^{\theta}\qquad\text{due to } j\le J\implies t_j\ge\eta=\delta^{m\alpha_3}/n^{\alpha_4}.
\end{aligned}
\tag{6.20}
$$

Moving on, for term p_*, we have

$$
\begin{aligned}
p_*&\le \mathbf{P}\Big(\sum_{j=J+1}^{m+k(n)}\big(\Xi'_{n,j}(t_j)\big)^+\in[y,y+c]\Big)\le \mathbf{P}\Big(\sum_{j=J+1}^{m+k(n)}\big(\Xi'_{n,j}(t_j)\big)^+\ge y_0\Big)\qquad\text{due to } y\ge y_0>0\\
&\le \sum_{j=J+1}^{m+k(n)}\mathbf{P}\Big(\Xi'_{n,j}(t_j)\ge y_0/N\Big)\qquad\text{where } N\triangleq k(m)+m-J\\
&\le \sum_{j=J+1}^{m+k(n)}\frac{C_X(\sqrt{t_j}+t_j)\cdot N}{y_0}\qquad\text{using Markov's inequality and Lemma 6.3}\\
&\le \sum_{j=J+1}^{m+k(n)}\frac{C_X(\sqrt{\eta}+\eta)\cdot N}{y_0}\qquad\text{due to } j>J\implies t_j<\eta=\delta^{\alpha_3}/n^{\alpha_4}\\
&=N^2\cdot\frac{C_X(\sqrt{\eta}+\eta)}{y_0}\le\big(m+k(n)\big)^2\frac{C_X(\sqrt{\eta}+\eta)}{y_0}\qquad\text{due to } N\le m+k(n)\\
&\le 2C_X\big(m^2+(\lceil\log_2(n^d)\rceil)^2\big)\frac{\sqrt{\eta}+\eta}{y_0}\qquad\text{using } (u+v)^2\le 2(u^2+v^2)\\
&\le 4C_X\big(m^2+(\lceil\log_2(n^d)\rceil)^2\big)\frac{\sqrt{\eta}}{y_0}=4C_X\big(m^2+(\lceil\log_2(n^d)\rceil)^2\big)\frac{\delta^{m\alpha_3/2}}{y_0\cdot n^{\alpha_4/2}}.
\end{aligned}
\tag{6.21}
$$

In the last line of the display above, we applied $\eta=\delta^{m\alpha_3}/n^{\alpha_4}\in(0,1)$. Plugging (6.20)(6.21) into

(6.19), we yield

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

In particular, in the last line of the display above, all quantities involved do not depend on the value of sequence t_j, so unconditionally we have

$$\mathbf{P}\Big(\sum_{j=1}^{m+k(n)}\big(\xi_j^{[n]}(l)\big)^+ \in [y, y+c]\Big) \leq C\frac{(m + (\lceil\log_2(n^d)\rceil)n^{\alpha_4\lambda}}{\delta^{m\alpha_3\lambda}}c^\theta + 4C_X\big(m^2 + (\lceil\log_2(n^d)\rceil)^2\big)\frac{\delta^{m\alpha_3/2}}{y_0\cdot n^{\alpha_4/2}}.$$

This completes the proof. □

Now we are ready to provide the proof of Propositions 3.2 and 3.3.

Proof of Proposition 3.2. We start by specifying parameters $\kappa, r, d, C_0, \rho_0$ and $\bar{m}$. First, let $\beta \in [0, 2)$ be the Blumenthal-Getoor index of X; see Assumption 1. Fix some

$$\beta_+ \in (\beta, 2). \tag{6.22}$$

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

$$d \geq 2, \qquad 2(r - \beta_+) \geq 2. \tag{6.23}$$

Let $\lambda > 0, \theta \in (0, 1]$ be the constants in Assumption 2. First, choose

$$\alpha_3 \in (0, \frac{\theta}{\lambda}),\ \ \alpha_4 \in (0, \frac{\theta}{2\lambda}). \tag{6.24}$$

Next, fix

$$\alpha_2 \in (0, \frac{\alpha_3}{2} \wedge 1), \tag{6.25}$$

and based on the chosen value of α_2, fix

$$\alpha_1 \in (0, \frac{\theta\alpha_2}{\lambda}). \tag{6.26}$$

Fix

$$\delta \in (1/\sqrt{2}, 1). \tag{6.27}$$

Since we require α_2 to be strictly less than 1, it is easy to see the existence of some integer $\bar{m}$ such that

$$\delta^{m\alpha_2} - \delta^m \geq \frac{\delta^{m\alpha_2}}{2} \text{ and } \delta^{m\alpha_2} < a \qquad \forall m \geq \bar{m} \tag{6.28}$$

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

$$\kappa^{2-\beta_+} < \frac{1}{2} < \delta^2 \tag{6.29}$$

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

$$\delta^{\alpha_1} < \rho_1, \tag{6.30}$$

$$\frac{\kappa^{2-\beta_+}}{\delta^2} < \rho_1 \tag{6.31}$$

$$\frac{1}{\sqrt{2}\delta} < \rho_1 \tag{6.32}$$

$$\delta^{\theta\alpha_2 - \lambda\alpha_1} < \rho_1 \tag{6.33}$$

$$\delta^{\theta - \lambda\alpha_3} < \rho_1 \tag{6.34}$$

$$\delta^{-\alpha_2 + \frac{\alpha_3}{2}} < \rho_1. \tag{6.35}$$

Lastly, for any $\rho_0 \in (\rho_1, 1)$, by picking a larger $\bar{m}$ if necessary, we can ensure that

$$m^2 \rho_1^m \leq \rho_0^m \qquad \forall m \geq \bar{m}. \tag{6.36}$$

We kick off the analysis by characterizing the law of J_n under $\mathbf{P}$ conditioning on event $\{\mathcal{D}(\bar{J}_n) = k\}$, i.e., there are k jumps over $[0,n]$ in the process J_n. Note that under $\mathbf{P}$, the process J_n is a Lévy process with generating triplet $(0, 0, \nu|_{[n\gamma,\infty)})$. Therefore, $\mathcal{L}(J_n | \{\mathcal{D}(\bar{J}_n) = k\}) = \mathcal{L}(\zeta_k)$ where

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

$0 < u_1 < u_2 < \cdots < u_k$ are the order statistics of k iid copies of Unif$(0,n)$, and z_i are iid with law $\nu\big(\cdot \cap [n\gamma,\infty)\big)/\nu[n\gamma,\infty)$. Recall the random functions $Y_n^*(\cdot)$ and $\hat{Y}_n^m(\cdot)$ in (3.13)(3.23). We now have (for all $n \geq 1, k, m \geq 0$)

$$\begin{aligned}\mathbf{P}\Big(Y_n^*(J_n) \neq \hat{Y}_n^m(J_n) \ \Big|\ \mathcal{D}(\bar{J}_n) = k\Big) &= \mathbf{P}\Big(Y_n^*(\zeta_k) \neq \hat{Y}_n^m(\zeta_k)\Big) \\ &\leq \mathbf{P}(u_1 < n\delta^{m\alpha_1}) + \mathbf{P}\Big(Y_n^*(\zeta_k) \neq \hat{Y}_n^m(\zeta_k),\ u_1 \geq n\delta^{m\alpha_1}\Big).\end{aligned}$$

For term $\mathbf{P}(u_1 < n\delta^{m\alpha_1})$, suppose that $0 < u_1 < u_2 < \cdots < u_k$ are the order statistics of $(v_i)_{i=1}^k$ that are iid copies of Unif$(0,n)$, then

$$\begin{aligned}\mathbf{P}(u_1 < n\delta^{m\alpha_1}) &= \mathbf{P}(v_i < n\delta^{m\alpha_1} \text{ for some } i \in [k]) \leq k \cdot \mathbf{P}\big(\text{Unif}(0,n) < n\delta^{m\alpha_1}\big) \\ &= k \cdot \delta^{m\alpha_1} < k \cdot \rho_0^m \qquad \text{due to (6.30)}.\end{aligned}$$

Therefore, it suffices to find C_0 such that

$$\mathbf{P}\Big(Y_n^*(\zeta_k) \neq \hat{Y}_n^m(\zeta_k),\ u_1 \geq n\delta^{m\alpha_1}\Big) \leq C_0 \rho_0^m \cdot (k+1) \qquad \forall n \geq 1, m, k \geq 0.$$

For notation simplicity, let $t(n) = \lceil \log_2(n^d) \rceil$. Due to the coupling between $\xi_j^{(i)}, \xi_j^{(i),m}$ in (3.20)(3.21) and the definitions of $Y_n^*(\cdot)$ and $\hat{Y}_n^m(\cdot)$ in (3.13)(3.23), we have

$$Y_n^*(\zeta_k) = \max_{i \in [k+1]} \underbrace{\mathbb{1}\Big\{\sum_{q=1}^{i-1}\sum_{j \geq 0} \xi_j^{(q)} + \sum_{q=1}^{i-1} z_q + \sum_{j \geq 1} (\xi_j^{(i)})^+ \geq na\Big\}}_{\triangleq Y_n^{(i),*}(\zeta_k)},$$

$$\hat{Y}_n^m(\zeta_k) = \max_{i \in [k+1]} \underbrace{\mathbb{1}\Big\{\sum_{q=1}^{i-1}\sum_{j \geq 0} \xi_j^{(q),m} + \sum_{q=1}^{i-1} z_q + \sum_{j=1}^{m+t(n)} (\xi_j^{(i),m})^+ \geq na\Big\}}_{\triangleq \hat{Y}_n^{(i),m}(\zeta_k)}$$

where (under the convention $u_0 = 0, u_{k+1} = n$)

$$l_1^{(i)} = V_1^{(i)}(u_{i+1} - u_i), \qquad l_j^{(i)} = V_j^{(i)}(u_{i+1} - u_i - l_1^{(i)} - l_2^{(i)} - \cdots - l_{j-1}^{(i)}) \ \ \forall j \geq 2,$$
$$\big(\xi_j^{(i)}, \xi_j^{(i),1}, \xi_j^{(i),2}, \xi_j^{(i),3}, \cdots\big) \stackrel{d}{=} \Big(\Xi_n(l_j^{(i)}), \breve{\Xi}_n^1(l_j^{(i)}), \breve{\Xi}_n^2(l_j^{(i)}), \breve{\Xi}_n^3(l_j^{(i)}), \cdots\Big).$$

with $V_j^{(i)}$ all being iid copies of Unif$(0,1)$. Therefore, $\mathbf{P}\big(Y_n^*(\zeta_k) \neq \hat{Y}_n^m(\zeta_k),\ u_1 \geq n\delta^{m\alpha_1}\big) \leq \sum_{i=1}^{k+1} \mathbf{P}\big(Y_n^{(i),*}(\zeta_k) \neq \hat{Y}_n^{(i),m}(\zeta_k),\ u_1 \geq n\delta^{m\alpha_1}\big)$. Next, define events

$$E_{n,1}^{(i),m} \triangleq \Big\{\Big|\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} - \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q),m}\Big| \leq \frac{\delta^m}{3\sqrt{n}}\Big\},$$
$$E_{n,2}^{(i),m} \triangleq \Big\{\Big|\sum_{j=1}^{m+t(n)}(\xi_j^{(i)})^+ - \sum_{j=1}^{m+t(n)}(\xi_j^{(i),m})^+\Big| \leq \frac{\delta^m}{3\sqrt{n}}\Big\},$$
$$E_{n,3}^{(i),m} \triangleq \Big\{\sum_{j\geq m+t(n)+1}(\xi_j^{(i)})^+ \leq \frac{\delta^m}{3\sqrt{n}}\Big\}.$$

Now we observe a few important facts on $E_{n,1}^{(i),m} \cap E_{n,2}^{(i),m} \cap E_{n,3}^{(i),m} \cap \big\{Y_n^{(i),*}(\zeta_k) \neq \hat{Y}_n^{(i),m}(\zeta_k),\ u_1 \geq n\delta^{m\alpha_1}\big\}$. First, let

$$W_n^{(i),*} \triangleq \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} + \sum_{q=1}^{i-1} z_q + \sum_{j\geq 1}(\xi_j^{(i)})^+,$$
$$\widetilde{W}_n^{(i),m} \triangleq \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} + \sum_{q=1}^{i-1} z_q + \sum_{j=1}^{m+t(n)}(\xi_j^{(i)})^+,$$
$$\hat{W}_n^{(i),m} \triangleq \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q),m} + \sum_{q=1}^{i-1} z_q + \sum_{j=1}^{m+t(n)}(\xi_j^{(i),m})^+.$$

On $E_{n,1}^{(i),m} \cap E_{n,2}^{(i),m} \cap E_{n,3}^{(i),m}$, we have

$$|W_n^{(i),*} - \widetilde{W}_n^{(i),m}| \vee |W_n^{(i),*} - \hat{W}_n^{(i),m}| \vee |\widetilde{W}_n^{(i),m} - \hat{W}_n^{(i),m}| \leq \delta^m/\sqrt{n}.$$

However, due to $Y_n^{(i),*}(\zeta_k) \neq \hat{Y}_n^{(i),m}(\zeta_k)$, we must have $\widetilde{W}_n^{(i),m} \in [na - \frac{\delta^m}{\sqrt{n}}, na + \frac{\delta^m}{\sqrt{n}}]$: otherwise, when combined with $|W_n^{(i),*} - \widetilde{W}_n^{(i),m}| \vee |\widetilde{W}_n^{(i),m} - \hat{W}_n^{(i),m}| \leq \delta^m/\sqrt{n}$, the fact that $\widetilde{W}_n^{(i),m} < na - \frac{\delta^m}{\sqrt{n}}$ (resp. $\widetilde{W}_n^{(i),m} > na + \frac{\delta^m}{\sqrt{n}}$) immediately implies that $\hat{W}_n^{(i),m} < na, W_n^{(i),*} < na$ (resp. $\hat{W}_n^{(i),m} > na, W_n^{(i),*} > na$), and hence $Y_n^{(i),*}(\zeta_k) = \hat{Y}_n^{(i),m}(\zeta_k) = 0$ (resp. $Y_n^{(i),*}(\zeta_k) = \hat{Y}_n^{(i),m}(\zeta_k) = 1$). In short, we have shown that

$$E_{n,1}^{(i),m} \cap E_{n,2}^{(i),m} \cap E_{n,3}^{(i),m} \cap \big\{Y_n^{(i),*}(\zeta_k) \neq \hat{Y}_n^{(i),m}(\zeta_k),\ u_1 \geq n\delta^{m\alpha_1}\big\}$$
$$\subseteq \Big\{\widetilde{W}_n^{(i),m} \in [na - \frac{\delta^m}{\sqrt{n}}, na + \frac{\delta^m}{\sqrt{n}}],\ u_1 \geq n\delta^{m\alpha_1}\Big\}.$$

This decomposition of events leads to

$$\mathbf{P}\Big(Y_n^{(i),*}(\zeta_k) \neq \hat{Y}_n^{(i),m}(\zeta_k),\ u_1 \geq n\delta^{m\alpha_1}\Big)$$
$$\leq \mathbf{P}\Big(\big(E_{n,1}^{(i),m}\big)^c\Big) + \mathbf{P}\Big(\big(E_{n,2}^{(i),m}\big)^c\Big) + \mathbf{P}\Big(\big(E_{n,3}^{(i),m}\big)^c\Big) + \mathbf{P}\Big(\widetilde{W}_n^{(i),m} \in [na - \frac{\delta^m}{\sqrt{n}}, na + \frac{\delta^m}{\sqrt{n}}],\ u_1 \geq n\delta^{m\alpha_1}\Big). \tag{6.37}$$

Furthermore,

$$
\begin{aligned}
&\mathbf{P}\Big(\widetilde{W}^{(i),m}_n \in [na - \frac{\delta^m}{\sqrt{n}}, na + \frac{\delta^m}{\sqrt{n}}],\ u_1 \geq n\delta^{m\alpha_1}\Big) \\
&= \mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j + \sum_{q=1}^{i-1} z_q \in [na - \delta^{m\alpha_2}, na + \delta^{m\alpha_2}],\ \widetilde{W}^{(i),m}_n \in [na - \frac{\delta^m}{\sqrt{n}}, na + \frac{\delta^m}{\sqrt{n}}],\ u_1 \geq n\delta^{m\alpha_1}\Big) \\
&\qquad + \mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j + \sum_{q=1}^{i-1} z_q \notin [na - \delta^{m\alpha_2}, na + \delta^{m\alpha_2}],\ \widetilde{W}^{(i),m}_n \in [na - \frac{\delta^m}{\sqrt{n}}, na + \frac{\delta^m}{\sqrt{n}}],\ u_1 \geq n\delta^{m\alpha_1}\Big) \\
&\leq \mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j + \sum_{q=1}^{i-1} z_q \in [na - \delta^{m\alpha_2}, na + \delta^{m\alpha_2}],\ u_1 \geq n\delta^{m\alpha_1}\Big) \\
&\qquad + \mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j + \sum_{q=1}^{i-1} z_q \notin [na - \delta^{m\alpha_2}, na + \delta^{m\alpha_2}],\ \widetilde{W}^{(i),m}_n \in [na - \frac{\delta^m}{\sqrt{n}}, na + \frac{\delta^m}{\sqrt{n}}]\Big) \\
&\leq \mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j + \sum_{q=1}^{i-1} z_q \in [na - \delta^{m\alpha_2}, na + \delta^{m\alpha_2}],\ u_1 \geq n\delta^{m\alpha_1}\Big) \\
&\qquad + \int_{\mathbb{R}\setminus[na-\delta^{m\alpha_2}, na+\delta^{m\alpha_2}]} \mathbf{P}\Big(\sum_{j=1}^{m+t(n)} (\xi^{(i)}_j)^+ \in [na - x - \frac{\delta^m}{\sqrt{n}}, na - x + \frac{\delta^m}{\sqrt{n}}]\Big)\mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j + \sum_{q=1}^{i-1} z_q \in dx\Big) \\
&= \mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j + \sum_{q=1}^{i-1} z_q \in [na - \delta^{m\alpha_2}, na + \delta^{m\alpha_2}],\ u_1 \geq n\delta^{m\alpha_1}\Big) \\
&\qquad + \int_{(-\infty, na-\delta^{m\alpha_2}]} \mathbf{P}\Big(\sum_{j=1}^{m+t(n)} (\xi^{(i)}_j)^+ \in [na - x - \frac{\delta^m}{\sqrt{n}}, na - x + \frac{\delta^m}{\sqrt{n}}]\Big)\mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j + \sum_{q=1}^{i-1} z_q \in dx\Big).
\end{aligned}
$$

In the last equality of the display above, we applied the simple fact that $\mathbf{P}\big(\sum_{j\geq 1}(\xi^{(i)}_j)^+ \geq 0\big) = 1$. Furthermore, we claim that for all $i \in [k+1]$, $n \geq 1$, and $m \geq \bar{m}$,

$$
\mathbf{P}\Big(\big(E^{(i),m}_{n,1}\big)^c\Big) = \mathbf{P}\Big(\Big|\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j - \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q),m}_j\Big| > \frac{\delta^m}{3\sqrt{n}}\Big) \leq C_1\rho_0^m, \tag{6.38}
$$

$$
\mathbf{P}\Big(\big(E^{(i),m}_{n,2}\big)^c\Big) = \mathbf{P}\Big(\Big|\sum_{j=1}^{m+t(n)} (\xi^{(i)}_j)^+ - \sum_{j=1}^{m+t(n)} (\xi^{(i),m}_j)^+\Big| > \frac{\delta^m}{3\sqrt{n}}\Big) \leq C_2\rho_0^m, \tag{6.39}
$$

$$
\mathbf{P}\Big(\big(E^{(i),m}_{n,3}\big)^c\Big) = \mathbf{P}\Big(\sum_{j\geq m+t(n)+1} (\xi^{(i)}_j)^+ > \frac{\delta^m}{3\sqrt{n}}\Big) \leq C_3\rho_0^m, \tag{6.40}
$$

$$
\mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j + \sum_{q=1}^{i-1} z_q \in [na - \delta^{m\alpha_2}, na + \delta^{m\alpha_2}],\ u_1 \geq n\delta^{m\alpha_1}\Big) \leq C_4\rho_0^m, \tag{6.41}
$$

$$
\mathbf{P}\Big(\sum_{j=1}^{m+t(n)} (\xi^{(i)}_j)^+ \in [y, y + \frac{2\delta^m}{\sqrt{n}}]\Big) \leq C_5\rho_0^m \qquad \forall y \geq \delta^{m\alpha_2} - \frac{\delta^m}{\sqrt{n}} \tag{6.42}
$$

where the values of constants $C_1, \cdots, C_5$ do not depend on n, m, k, i. Then by setting $C_0 = \sum_{j=1}^5 C_j$, we have in (6.37) that $\mathbf{P}\big(Y^{(i),*}_n(\zeta_k) \neq \hat{Y}^{(i),m}_n(\zeta_k),\ u_1 \geq n\delta^{m\alpha_1}\big) \leq C_0\rho^m$ for all $n \geq 1, m \geq \bar{m}, i \in$

$[k+1]$. In conclusion, we yield (for all $n \geq 1,\ m \geq \bar{m}$)

$$\begin{aligned}&\mathbf{P}\Big(Y_n^*(\zeta_k) \neq \hat{Y}_n^m(\zeta_k),\ u_1 \geq n\delta^{m\alpha_1}\Big)\\ &\leq \sum_{i=1}^{k+1}\mathbf{P}\big(Y_n^{(i),*}(\zeta_k) \neq \hat{Y}_n^{(i),m}(\zeta_k),\ u_1 \geq n\delta^{m\alpha_1}\big) \leq C_0\rho^m \cdot (k+1)\end{aligned}$$

and conclude the proof. Now, it only remains to prove claims (6.38)(6.39)(6.40)(6.41)(6.42).

Proof of Claim (6.38)

From the coupling between $\xi_j^{(i)}, \xi_j^{(i),m}$ in (3.20)(3.21), we have

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

where the laws of processes $\Xi_n, \breve{\Xi}_n^m$ are stated in (3.15)(3.17). Applying Lemma 6.4, we yield (for all $n \geq 1, m, k \geq 0, i \in [k+1]$)

$$\begin{aligned}&\mathbf{P}\Big(\Big|\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} - \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q),m}\Big| > \frac{\delta^m}{3\sqrt{n}}\Big)\\ &= \mathbf{P}\Big(|\Xi_n(u_{i-1}) - \breve{\Xi}_n^m(u_{i-1})| > \frac{\delta^m}{3\sqrt{n}}\Big) \leq \mathbf{P}\Big(\sup_{t\in[0,n]}|\Xi_n(t) - \breve{\Xi}_n^m(t)| > \frac{\delta^m}{3\sqrt{n}}\Big)\\ &\leq \frac{9C_{\beta_+}n}{\delta^{2m}\cdot n^{2(r-\beta_+)-1}} \cdot \kappa^{m(2-\beta_+)} = 9C_{\beta_+} \cdot \frac{1}{n^{2(r-\beta_+)-2}} \cdot \left(\frac{\kappa^{2-\beta_+}}{\delta^2}\right)^m\\ &\leq 9C_{\beta_+} \cdot \left(\frac{\kappa^{2-\beta_+}}{\delta^2}\right)^m \qquad \text{due to (6.23)}\\ &\leq 9C_{\beta_+} \cdot \rho_0^m \qquad \text{due to (6.31) and } \rho_0 \in (\rho_1, 1)\end{aligned}$$

where $C_{\beta_+} < \infty$ is the constant characterized in Lemma 6.4 that only depends on β_+ and the law of the Lévy process X. To conclude the proof of claim (6.38), we pick $C_1 = 9C_{\beta_+}$.

Proof of Claim (6.39)

It follows directly from Lemma 6.5 that

$$\begin{aligned}&\mathbf{P}\Big(\Big|\sum_{j=1}^{m+t(n)}(\xi_j^{(i)})^+ - \sum_{j=1}^{m+t(n)}(\xi_j^{(i),m})^+\Big| > \frac{\delta^m}{3\sqrt{n}}\Big)\\ &\leq \frac{9C_{\beta_+}n}{\delta^{2m}\cdot n^{2(r-\beta_+)-1}} \cdot \kappa^{m(2-\beta_+)} = 9C_{\beta_+} \cdot \frac{1}{n^{2(r-\beta_+)-2}} \cdot \left(\frac{\kappa^{2-\beta_+}}{\delta^2}\right)^m\\ &\leq 9C_{\beta_+} \cdot \left(\frac{\kappa^{2-\beta_+}}{\delta^2}\right)^m \qquad \text{due to (6.23)}\\ &\leq 9C_{\beta_+} \cdot \rho_0^m \qquad \text{due to (6.31) and } \rho_0 \in (\rho_1, 1)\end{aligned}$$

where $C_{\beta_+} < \infty$ is the constant characterized in Lemma 6.5 that only depends on β_+ and the law of the Lévy process X. To conclude the proof of claim (6.39), we pick $C_2 = 9C_{\beta_+}$.

Proof of Claim (6.40)

Using Lemma 6.6, we get

$$\mathbf{P}\Big(\sum_{j\geq m+t(n)+1}(\xi_j^{(i)})^+ > \frac{\delta^m}{3\sqrt{n}}\Big) \leq \frac{3C_X\sqrt{n}}{\delta^m} \cdot \left[\sqrt{\frac{1}{n^{d-1}\cdot 2^m}} + \frac{1}{n^{d-1}\cdot 2^m}\right] \leq \frac{6C_X\sqrt{n}}{\delta^m} \cdot \sqrt{\frac{1}{n^{d-1}\cdot 2^m}}$$

$$
\begin{aligned}
&= 6C_X \cdot \sqrt{\frac{1}{n^{d-2}}} \cdot \Big(\frac{1}{\sqrt{2}\delta}\Big)^m \\
&\leq 6C_X \cdot \Big(\frac{1}{\sqrt{2}\delta}\Big)^m \qquad \text{due to (6.23)} \\
&\leq 6C_X \cdot \rho_0^m \qquad \text{due to (6.32) and } \rho_0 \in (\rho_1, 1)
\end{aligned}
$$

where $C_X < \infty$ is the constant characterized in Lemma 6.6 that only depends on the law of the Lévy process X. By setting $C_3 = 6C_X$, we conclude the proof of claim (6.40).

Proof of Claim (6.41)

Due to the independence between z_i and $\xi_j^{(i)}$,

$$
\begin{aligned}
&\mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} + \sum_{q=1}^{i-1} z_q \in [na - \delta^{m\alpha_2}, na + \delta^{m\alpha_2}],\ u_1 \geq n\delta^{m\alpha_1}\Big) \\
&= \int_{\mathbb{R}} \mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} \in [na - x - \delta^{m\alpha_2}, na - x + \delta^{m\alpha_2}],\ u_1 \geq n\delta^{m\alpha_1}\Big)\mathbf{P}(\sum_{q=1}^{i-1} z_q \in dx) \\
&\leq \int_{\mathbb{R}} \mathbf{P}\Big(\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} \in [na - x - \delta^{m\alpha_2}, na - x + \delta^{m\alpha_2}]\ \Big|\ u_1 \geq n\delta^{m\alpha_1}\Big)\mathbf{P}(\sum_{q=1}^{i-1} z_q \in dx) \\
&= \int_{\mathbb{R}} \mathbf{P}\Big(X^{<n\gamma}(u_{i-1}) \in [na - x - \delta^{m\alpha_2}, na - x + \delta^{m\alpha_2}]\ \Big|\ u_1 \geq n\delta^{m\alpha_1}\Big)\mathbf{P}(\sum_{q=1}^{i-1} z_q \in dx)
\end{aligned}
$$

where $(u_i)_{i=1}^k$ are independent of the Lévy process $X^{<n\gamma}$; see the law of $\xi_j^{(i)}$ in (3.20). Now let us consider two different cases depending on the value of $i \in [k+1]$. If $i \geq 2$, then due to the nature of the order statistics $0 = u_0 < u_1 < u_2 < \cdots < u_k < u_{k+1} = 1$, 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

$$
\begin{aligned}
&\mathbf{P}\Big(X^{<n\gamma}(u_{i-1}) \in [na - x - \delta^{m\alpha_2}, na - x + \delta^{m\alpha_2}]\ \Big|\ u_1 \geq n\delta^{m\alpha_1}\Big) \\
&\leq \frac{C}{n^\lambda \cdot \delta^{m\alpha_1\lambda}} \cdot (2\delta^{m\alpha_2})^\theta \leq 2^\theta C \cdot \Big(\frac{\delta^{\theta\alpha_2}}{\delta^{\lambda\alpha_1}}\Big)^m \\
&\leq 2^\theta C \cdot \rho_0^m \qquad \text{due to (6.33) and } \rho_0 \in (\rho_1, 1).
\end{aligned}
$$

In case that $i = 1$, we get

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

For any $m \geq \bar{m}$, due to (6.28) we must have $na - \delta^{m\alpha_2} \geq a - \delta^{m\alpha_2}0$ for all $n \geq 1, m \geq \bar{m}$, thus implying $\mathbb{1}\big(0 \in [na - \delta^{m\alpha_2}, na + \delta^{m\alpha_2}]\big) = 0$. To conclude, it suffices to pick $C_4 = 2^\theta C$ and note that $C < \infty, \theta \in (0, 1]$ are constants in Assumption 2 that only depend on the law of the Lévy process X.

Proof of Claim (6.42)

Applying Lemma 6.7 with $y_0 = \delta^{m\alpha_2} - \frac{\delta^m}{\sqrt{n}}$ and $c = \frac{2\delta^m}{\sqrt{n}}$, we get (for all $n \geq 1, m \geq \bar{m}, y \geq y_0$)

$$
\begin{aligned}
&\mathbf{P}\Big(\sum_{j=1}^{m+t(n)} (\xi_j^{(i)})^+ \in [y, y + \frac{2\delta^m}{\sqrt{n}}]\Big) \\
&\leq C\frac{(m + (\lceil \log_2(n^d) \rceil)n^{\alpha_4\lambda}}{\delta^{m\alpha_3\lambda}} \cdot \frac{2^\theta\delta^{m\theta}}{n^{\theta/2}} + 4C_X\big(m^2 + (\lceil \log_2(n^d) \rceil)^2\big)\frac{\delta^{m\alpha_3/2}}{(\delta^{m\alpha_2} - \frac{\delta^m}{\sqrt{n}}) \cdot n^{\alpha_4/2}}
\end{aligned}
$$

$$
\begin{aligned}
&\leq C\frac{(m+(\lceil\log_2(n^d)\rceil)n^{\alpha_4\lambda}}{\delta^{m\alpha_3\lambda}}\cdot\frac{2^\theta\delta^{m\theta}}{n^{\theta/2}}+8C_X\big(m^2+(\lceil\log_2(n^d)\rceil)^2\big)\frac{\delta^{m\alpha_3/2}}{\delta^{m\alpha_2}\cdot n^{\alpha_4/2}} \\
&\qquad\text{due to (6.28)} \\
&=\underbrace{2^\theta C\cdot\frac{m}{n^{\frac{\theta}{2}-\lambda\alpha_4}}\cdot\left(\frac{\delta^\theta}{\delta^{\lambda\alpha_3}}\right)^m}_{\triangleq p_{n,m,1}}+\underbrace{2^\theta C\cdot\frac{\lceil\log_2(n^d)\rceil}{n^{\frac{\theta}{2}-\lambda\alpha_4}}\cdot\left(\frac{\delta^\theta}{\delta^{\lambda\alpha_3}}\right)^m}_{\triangleq p_{n,m,2}} \\
&\qquad+\underbrace{8C_X\cdot\frac{m^2}{n^{\alpha_4/2}}\cdot\left(\frac{\delta^{\alpha_3/2}}{\delta^{\alpha_2}}\right)^m}_{\triangleq p_{n,m,3}}+\underbrace{8C_X\cdot\frac{\big(\lceil\log_2(n^d)\rceil\big)^2}{n^{\alpha_4/2}}\cdot\left(\frac{\delta^{\alpha_3/2}}{\delta^{\alpha_2}}\right)^m}_{\triangleq p_{n,m,4}}
\end{aligned}
$$

where $C_X < \infty$ is the constant characterized in Lemma 6.3 that only depends on the law of Lévy process X, and $C \in (0,\infty), \lambda > 0, \theta \in (0,1]$ be the constants in Assumption 2. For term $p_{n,m,1}$, note that for any $n \geq 1, m \geq \bar{m}$,

$$
\begin{aligned}
p_{n,m,1} &\leq 2^\theta C\cdot m\left(\frac{\delta^\theta}{\delta^{\lambda\alpha_3}}\right)^m \qquad \text{due to } \frac{\theta}{2} > \lambda\alpha_4\text{; see (6.24)} \\
&\leq 2^\theta C\cdot m\rho_1^m \qquad \text{due to (6.34)} \\
&\leq 2^\theta C\cdot\rho_0^m \qquad \text{due to (6.36)}.
\end{aligned}
$$

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

$$
p_{n,m,2} \leq 2^\theta CC_{d,1}\cdot\left(\frac{\delta^\theta}{\delta^{\lambda\alpha_3}}\right)^m \leq 2^\theta CC_{d,1}\cdot\rho_0^m \qquad \text{due to (6.34) and } \rho_0\in(\rho_1,1).
$$

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

$$
\begin{aligned}
p_{n,m,3} &\leq 8C_X\cdot m^2\left(\frac{\delta^{\alpha_3/2}}{\delta^{\alpha_2}}\right)^m \leq 8C_X\cdot m^2\rho_1^m \qquad \text{due to (6.35)} \\
&\leq 8C_X\cdot\rho_0^m \qquad \text{due to (6.36)}.
\end{aligned}
$$

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 $n\geq 1, m\geq 0$)

$$
p_{n,m,4} \leq 8C_XC_{d,2}\cdot\left(\frac{\delta^{\alpha_3/2}}{\delta^{\alpha_2}}\right)^m \leq 8C_XC_{d,2}\cdot\rho_0^m \qquad \text{due to (6.34) and } \rho_0\in(\rho_1,1).
$$

To conclude the proof, one can simply set $C_5 = \max\{2^\theta C,\ 2^\theta CC_{d,1},\ 8C_X,\ 8C_XC_{d,2}\}$. □

Proof of Proposition 3.3. Throughout this proof, we only consider $\rho_0 \in (0,1)$ large enough such that

$$
\rho_0 > \frac{1}{\sqrt{2}}, \qquad \rho_0 > \kappa^{2-\beta_+}. \tag{6.43}
$$

Under $\mathbf{P}$, the process J_n is a Lévy process with generating triplet $(0,0,\nu|_{[n\gamma,\infty)})$. Therefore, $\mathcal{L}(J_n|\{\mathcal{D}(\bar{J}_n)=k\}) = \mathcal{L}(\zeta_k)$ where

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

$0 < u_1 < u_2 < \cdots < u_k$ are the order statistics of k iid copies of $\text{Unif}(0,n)$, and z_i are iid with law $\nu\big(\,\cdot \cap [n\gamma,\infty)\big)/\nu[n\gamma,\infty)$.

For notation simplicity, let $t(n) = \lceil \log_2(n^d) \rceil$. Due to the coupling between $\xi^{(i)}_j, \xi^{(i),m}_j$ in (3.20)(3.21) and the definitions of $Y^*_n(\cdot)$ and $\hat{Y}^m_n(\cdot)$ in (3.13)(3.23), we have

$$Y^*_n(\zeta_k) = \max_{i\in[k+1]} \mathbb{1}\big\{W^{(i),*}_n \geq na\big\}, \qquad \hat{Y}^m_n(\zeta_k) = \max_{i\in[k+1]} \mathbb{1}\big\{\hat{W}^{(i),m}_n \geq na\big\}$$

where

$$W^{(i),*}_n \triangleq \sum_{q=1}^{i-1}\sum_{j\geq 0} \xi^{(q)}_j + \sum_{q=1}^{i-1} z_q + \sum_{j\geq 1}(\xi^{(i)}_j)^+,$$
$$\hat{W}^{(i),m}_n \triangleq \sum_{q=1}^{i-1}\sum_{j\geq 0} \xi^{(q),m}_j + \sum_{q=1}^{i-1} z_q + \sum_{j=1}^{m+t(n)}(\xi^{(i),m}_j)^+.$$

In particular, on event $\{\mathcal{D}(\bar{J}_n) = k\}$ (i.e., J_n admits the representation of ζ_k over $[0,n]$), we have $\sup_{t\in[0,n]} X(t) = \max_{i=1}^{k+1} W^{(i),*}_n$. As a result, for all $n \geq 1$, $m \geq 0$, and $k = 0,1,\cdots,l^*-1$,

$$\begin{aligned}
\mathbf{P}\Big(Y^*_n(J_n) \neq \hat{Y}^m_n(J_n),\ \bar{X}_n \notin A^\Delta \ \Big|\ \mathcal{D}(\bar{J}_n) = k\Big) &= \mathbf{P}\Big(Y^*_n(J_n) \neq \hat{Y}^m_n(J_n),\ \sup_{t\in[0,n]} X(t) < n(a-\Delta) \ \Big|\ \mathcal{D}(\bar{J}_n) = k\Big) \\
&= \mathbf{P}\Big(\max_{i\in[k+1]} \hat{W}^{(i),m}_n \geq na,\ \max_{i\in[k+1]} W^{(i),*}_n < n(a-\Delta)\Big) \\
&\leq \sum_{i\in[k+1]} \mathbf{P}\Big(\big|\hat{W}^{(i),m}_n - W^{(i),*}_n\big| > n\Delta\Big).
\end{aligned} \tag{6.44}$$

To proceed, define events

$$E^{(i),m}_{n,1} \triangleq \Big\{\Big|\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j - \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q),m}_j\Big| \leq n\Delta/3\Big\},$$
$$E^{(i),m}_{n,2} \triangleq \Big\{\Big|\sum_{j=1}^{m+t(n)}(\xi^{(i)}_j)^+ - \sum_{j=1}^{m+t(n)}(\xi^{(i),m}_j)^+\Big| \leq n\Delta/3\Big\},$$
$$E^{(i),m}_{n,3} \triangleq \Big\{\sum_{j\geq m+t(n)+1}(\xi^{(i)}_j)^+ \leq n\Delta/3\Big\}.$$

Note that on $E^{(i),m}_{n,1} \cap E^{(i),m}_{n,2} \cap E^{(i),m}_{n,3}$ we must have $\big|\hat{W}^{(i),m}_n - W^{(i),*}_n\big| \leq n\Delta$, which implies $\{|\hat{W}^{(i),m}_n - W^{(i),*}_n| > n\Delta\} \cap \big(E^{(i),m}_{n,1} \cap E^{(i),m}_{n,2} \cap E^{(i),m}_{n,3}\big) = \emptyset$ and hence

$$\{|\hat{W}^{(i),m}_n - W^{(i),*}_n| > n\Delta\} \subseteq \big(E^{(i),m}_{n,1} \cap E^{(i),m}_{n,2} \cap E^{(i),m}_{n,3}\big)^c = \big(E^{(i),m}_{n,1}\big)^c \cup \big(E^{(i),m}_{n,2}\big)^c \cup \big(E^{(i),m}_{n,3}\big)^c. \tag{6.45}$$

Furthermore, we claim that for all $n \geq 1, m \geq 0, k = 0,1,\cdots,l^*-1$ and $i \in [k+1]$,

$$\mathbf{P}\Big(\big(E^{(i),m}_{n,1}\big)^c\Big) = \mathbf{P}\Big(\Big|\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q)}_j - \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi^{(q),m}_j\Big| > n\Delta/3\Big) \leq \frac{C_1\rho_0^m}{n^\mu\Delta^2}, \tag{6.46}$$

$$\mathbf{P}\Big(\big(E^{(i),m}_{n,2}\big)^c\Big) = \mathbf{P}\Big(\Big|\sum_{j=1}^{m+t(n)}(\xi^{(i)}_j)^+ - \sum_{j=1}^{m+t(n)}(\xi^{(i),m}_j)^+\Big| > n\Delta/3\Big) \leq \frac{C_2\rho_0^m}{n^\mu\Delta^2}, \tag{6.47}$$

$$\mathbf{P}\Big(\big(E^{(i),m}_{n,3}\big)^c\Big) = \mathbf{P}\Big(\sum_{j\geq m+t(n)+1}(\xi^{(i)}_j)^+ > n\Delta/3\Big) \leq \frac{C_3\rho_0^m}{n^\mu\Delta^2} \tag{6.48}$$

where the constants C_1, C_2, C_3 do not depend on n, m, k, i. By combining these claims with the decomposition of events (6.45), we have in (6.44) that (with $C' = C_1 + C_2 + C_3$)

$$\mathbf{P}\Big(Y_n^*(J_n) \neq \hat{Y}_n^m(J_n),\ \bar{X}_n \notin A^\Delta\ \Big|\ \mathcal{D}(\bar{J}_n) = k\Big) \leq (k+1)\frac{C'\rho_0^m}{n^\mu \Delta^2} \leq l^* \cdot \frac{C'\rho_0^m}{n^\mu \Delta^2}$$

for all $n \geq 1, m \geq 0, k = 0, 1, \cdots, l^* - 1$. By setting $C_0 = l^* C'$ we conclude the proof. Now it only remains to prove (6.46)(6.47)(6.48).

Proof of Claim (6.46)

From the coupling between $\xi_j^{(i)}, \xi_j^{(i),m}$ in (3.20)(3.21), we have

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

where the laws of processes $\Xi_n, \breve{\Xi}_n^m$ are stated in (3.15)(3.17). Applying Lemma 6.4, we yield (for all $n \geq 1, m, k \geq 0, i \in [k+1]$)

$$\begin{aligned}
&\mathbf{P}\Big(\Big|\sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q)} - \sum_{q=1}^{i-1}\sum_{j\geq 0}\xi_j^{(q),m}\Big| > n\Delta/3\Big) \\
&= \mathbf{P}\Big(|\Xi_n(u_{i-1}) - \breve{\Xi}_n^m(u_{i-1})| > n\Delta/3\Big) \leq \mathbf{P}\Big(\sup_{t\in[0,n]}|\Xi_n(t) - \breve{\Xi}_n^m(t)| > n\Delta/3\Big) \\
&\leq \frac{9C_{\beta_+}}{\Delta^2 \cdot n^{2(r-\beta_+)+1}} \cdot \kappa^{m(2-\beta_+)} \\
&\leq \frac{9C_{\beta_+}}{\Delta^2} \cdot \frac{\rho_0^m}{n^\mu} \qquad \text{due to } 2(r-\beta_+)+1 > \mu \text{ and } \kappa^{2-\beta_+} < \rho_0
\end{aligned}$$

where $C_{\beta_+} < \infty$ is the constant characterized in Lemma 6.4 that only depends on β_+ and the law of the Lévy process X. To conclude the proof of claim (6.46), we pick $C_1 = 9C_{\beta_+}$.

Proof of Claim (6.47)

It follows directly from Lemma 6.5 that

$$\begin{aligned}
\mathbf{P}\Big(\Big|\sum_{j=1}^{m+t(n)}(\xi_j^{(i)})^+ - \sum_{j=1}^{m+t(n)}(\xi_j^{(i),m})^+\Big| > n\Delta/3\Big) &\leq \frac{9C_{\beta_+}\kappa^{m(2-\beta_+)}}{\Delta^2 \cdot n^2 \cdot n^{2(r-\beta_+)-1}} = \frac{9C_{\beta_+}\kappa^{m(2-\beta_+)}}{\Delta^2 \cdot n^{2(r-\beta_+)+1}} \\
&\leq \frac{9C_{\beta_+}}{\Delta^2} \cdot \frac{\rho_0^m}{n^\mu}.
\end{aligned}$$

In the last line of the display above, we again applied $2(r-\beta_+)+1 > \mu$ and $\kappa^{2-\beta_+} < \rho_0$; here $C_{\beta_+} < \infty$ is the constant characterized in Lemma 6.5 that only depends on β_+ and the law of the Lévy process X. To conclude the proof of claim (6.47), we pick $C_2 = 9C_{\beta_+}$.

Proof of Claim (6.48)

Using Lemma 6.6, we get

$$\begin{aligned}
\mathbf{P}\Big(\sum_{j\geq m+t(n)+1}(\xi_j^{(i)})^+ > n\Delta/3\Big) &\leq \frac{3C_X}{n\Delta} \cdot \Big[\sqrt{\frac{1}{n^{d-1}\cdot 2^m}} + \frac{1}{n^{d-1}\cdot 2^m}\Big] \leq \frac{6C_X}{n\Delta} \cdot \sqrt{\frac{1}{n^{d-1}\cdot 2^m}} \\
&= \frac{6C_X}{\Delta} \cdot \frac{(\sqrt{1/2})^m}{n^{\frac{d+1}{2}}} \\
&\leq \frac{6C_X}{\Delta} \cdot \frac{\rho_0^m}{n^\mu} \qquad \text{due to } \frac{d+1}{2} > \mu \text{ and } \frac{1}{\sqrt{2}} < \rho_0
\end{aligned}$$

$$< \frac{6C_X}{\Delta^2} \cdot \frac{\rho_0^m}{n^\mu} \qquad \text{due to } \Delta \in (0,1)$$

where $C_X < \infty$ is the constant characterized in Lemma 6.6 that only depends on the law of the Lévy process X. By setting $C_3 = 6C_X$, we conclude the proof of claim (6.48). □

6.3 Proofs of Propositions 4.1 and 4.3

The proof of Proposition 4.1 is based on the inversion formula of the characteristic functions (see, e.g., theorem 3.3.5. of [23]). Specifically, we compare the characteristic function of $Y(t)$ with an α-stable process and establish the similarities between their laws.

Proof of Proposition 4.1. Applying Lévy-Khintchine formula, we yield the following formula for the characteristic function of $\varphi_t(z) = \mathbf{E}\exp(izY(t))$:

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

Note that for $\phi(z,x)$ and its complex conjugate $\overline{\phi(z,x)}$, we have

$$\begin{aligned}\phi(z,x) &= \cos(zx) - 1 + i\big(\sin(zx) - zx\mathbb{1}_{(0,1]}(x)\big),\\ \overline{\phi(z,x)} &= \cos(zx) - 1 - i\big(\sin(zx) - zx\mathbb{1}_{(0,1]}(x)\big).\end{aligned}$$

As a result,

$$|\varphi_t(z)| = \exp\Big(-t\int_{(0,z_0)}\big(1-\cos(zx)\big)\nu(dx)\Big) \qquad \forall z \in \mathbb{R},\ t > 0. \tag{6.49}$$

Furthermore, we claim the existence of some $\widetilde{M}, \widetilde{C} \in (0,\infty)$ such that

$$\int_{(0,z_0)}\big(1-\cos(zx)\big)\nu(dx) \geq \widetilde{C}|z|^\alpha \qquad \forall |z| \geq \widetilde{M}. \tag{6.50}$$

By plugging (6.50) in (6.49), we obtain 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 3.3.5. of [23]) that for all $t > 0$,

$$\begin{aligned}\|f_{Y_t}\|_\infty &\leq \frac{1}{2\pi}\int|\varphi_t(z)|dz\\ &\leq \frac{1}{2\pi}\Big(2\widetilde{M} + \int_{|z|\geq\tilde{M}}\exp\big(-t\widetilde{C}|z|^\alpha\big)dz\Big)\\ &\leq \frac{1}{2\pi}\Big(2\widetilde{M} + \int_{|z|>0}\exp\big(-t\widetilde{C}|z|^\alpha\big)dz\Big)\\ &= \frac{1}{2\pi}\Big(2\widetilde{M} + \frac{1}{t^{1/\alpha}}\int\exp(-\widetilde{C}|x|^\alpha)dx\Big) \qquad \text{by letting } x = zt^{1/\alpha}\\ &= \frac{\widetilde{M}}{\pi} + \frac{C_1}{t^{1/\alpha}} \qquad \text{where } C_1 = \frac{1}{2\pi}\int\exp(-\widetilde{C}|x|^\alpha)dx < \infty.\end{aligned}$$

To conclude the proof, one only needs to pick $C = \frac{\widetilde{M}}{\pi} + C_1$. Now, it only remains to prove (6.50).

Proof of Claim (6.50)

We start by fixing some constants. Let

$$C_0 = \int_0^\infty(1-\cos y)\frac{dy}{y^{1+\alpha}}. \tag{6.51}$$

On $y \in (0,1]$, note that $1 - \cos y \leq y^2/2$, and hence $\frac{|1-\cos y|}{y^{1+\alpha}} \leq \frac{1}{2y^{\alpha-1}}$. On $y \in (1,\infty)$, note that $1 - \cos y \in [0,1]$ so $\frac{|1-\cos y|}{y^{1+\alpha}} \leq 1/y^{\alpha+1}$. 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:

$$\frac{\theta^{2-\alpha}}{2(2-\alpha)} \leq \frac{C_0}{8}; \tag{6.52}$$

$$\frac{\delta}{\theta^\alpha} \leq \frac{C_0}{8}. \tag{6.53}$$

For any $M > 0$ and $z \neq 0$, observe that

$$\begin{aligned}
\frac{\int_{x \geq \frac{M}{|z|}} \big(1-\cos(zx)\big)\frac{dx}{x^{1+\alpha}}}{|z|^\alpha} &= \frac{\int_{x \geq \frac{M}{|z|}} \big(1-\cos(|z|x)\big)\frac{dx}{x^{1+\alpha}}}{|z|^\alpha} \\
&= \int_M^\infty \big(1-\cos y\big)\frac{dy}{y^{1+\alpha}} \qquad \text{by letting } y = |z|x.
\end{aligned}$$

Therefore, by fixing some $M > \theta$ large enough, we must have

$$\frac{1}{|z|^\alpha}\int_{x \geq M/|z|} \big(1-\cos(zx)\big)\frac{dx}{x^{1+\alpha}} \leq C_0/4 \qquad \forall z \neq 0. \tag{6.54}$$

Moving on, we fix some $c > 0$ and consider the difference between $\int_{(0,z_0)}\big(1-\cos(zx)\big)\nu(dx)$ and $\int_0^{M/z}\big(1-\cos(zx)\big)\frac{c\,dx}{x^{1+\alpha}}$. For any z such that $|z| > M/z_0$,

$$\begin{aligned}
&\frac{1}{|z|^\alpha}\bigg[\int_{(0,z_0)}\Big(1-\cos(zx)\Big)\nu(dx) - \int_0^{M/|z|}\Big(1-\cos(zx)\Big)c\frac{dx}{x^{1+\alpha}}\bigg] \\
&\geq \frac{1}{|z|^\alpha}\bigg[\int_{(0,M/|z|)}\Big(1-\cos(zx)\Big)\nu(dx) - \int_0^{M/|z|}\Big(1-\cos(zx)\Big)c\frac{dx}{x^{1+\alpha}}\bigg] \\
&\geq -\underbrace{\frac{1}{|z|^\alpha}\int_0^{\theta/|z|}\Big(1-\cos(zx)\Big)c\frac{dx}{x^{1+\alpha}}}_{\triangleq I_1(z)} \qquad \text{due to our choice of } M > \theta \\
&\quad + \underbrace{\frac{1}{|z|^\alpha}\bigg[\int_{[\theta/|z|,M/|z|)}\Big(1-\cos(zx)\Big)\nu(dx) - \int_{[\theta/|z|,M/|z|)}\Big(1-\cos(zx)\Big)c\frac{dx}{x^{1+\alpha}}\bigg]}_{\triangleq I_2(z)}.
\end{aligned} \tag{6.55}$$

First, for any $z \neq 0$,

$$\begin{aligned}
I_1(z) &\leq \frac{c}{|z|^\alpha}\int_0^{\theta/|z|}\frac{z^2x^2}{2}\frac{dx}{x^{1+\alpha}} \qquad \text{due to } 1-\cos w \leq \frac{w^2}{2}\ \forall w \in \mathbb{R} \\
&= \frac{c}{2}\int_0^\theta y^{1-\alpha}dy \qquad \text{by setting } y = |z|x \\
&= \frac{c}{2}\cdot\frac{\theta^{2-\alpha}}{2-\alpha} \leq c\cdot\frac{C_0}{8} \quad \text{due to (6.52)}.
\end{aligned} \tag{6.56}$$

Next, to analyze $I_2(z)$, let us focus on the function

$$h(z) = 1 - \cos z.$$

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 $M = x_0 > x_1 > \cdots > x_N = \theta$ such that

$$\begin{aligned} &\frac{x_{j-1}}{x_j} = t_0 \qquad \forall j = 1, 2, \cdots, N, \\ &|h(x) - h(y)| < \delta \qquad \forall j = 1, 2, \cdots, N, \ x, y \in [x_j, x_{j-1}]. \end{aligned} \tag{6.57}$$

In other words, we use a geometric sequence $\{x_0, x_1, \cdots, x_N\}$ to partition $[\theta, M]$ into $N+1$ intervals, on any of which the value of $h(z) = 1 - \cos z$ fluctuates within range δ in (6.53). Now fix some $\Delta > 0$ such that

$$(1 - \Delta) t_0^{\alpha+\epsilon} > 1. \tag{6.58}$$

Recall that $\nu[x, \infty)$ is regularly varying as $x \to 0$ with index $\alpha + 2\epsilon$. In other words, for $g(y) = \nu[1/y, \infty)$, we have $g \in RV_{\alpha+2\epsilon}$. By Potter's bounds (see proposition 2.6 in [46]), we know the existence of some $\bar{y}_1 > 0$ such that

$$\frac{g(ty)}{g(y)} \geq (1 - \Delta) t^{\alpha+\epsilon} \qquad \forall y \geq \bar{y}_1, \ t \geq 1. \tag{6.59}$$

On the other hand, define

$$\widetilde{g}(y) = c y^{\alpha}, \qquad \nu_c(dx) = c \mathbb{1}_{(0,\infty)}(x) \frac{dx}{x^{1+\alpha}}.$$

Note that $\widetilde{g}(y) = \nu_c(1/y, \infty)$. Due to $g \in RV_{\alpha+2\epsilon}$, we can find some $\bar{y}_2 > 0$ such that

$$g(y) \geq \frac{t_0^{\alpha} - 1}{(1 - \Delta) t_0^{\alpha+\epsilon} - 1} \cdot \widetilde{g}(y) \quad \forall y \geq \bar{y}_2. \tag{6.60}$$

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 $j = 1, 2, \cdots, N$, the mass of ν on $[x_j, x_{j-1})$ satisfies

$$\begin{aligned} \nu[x_j/|z|, x_{j-1}/|z|) &= g(|z|/x_j) - g(|z|/x_{j-1}) \qquad \text{due to } g(y) = \nu[1/y, \infty) \\ &= g(t_0|z|/x_{j-1}) - g(|z|/x_{j-1}) \qquad \text{due to } x_{j-1} = t_0 x_j, \text{ see (6.57)} \\ &\geq g(|z|/x_{j-1}) \cdot \Big((1 - \Delta) t_0^{\alpha+\epsilon} - 1 \Big) \qquad \text{due to } \frac{|z|}{x_j} \geq \bar{y}_1 \vee \bar{y}_2 \text{ and (6.59)} \\ &\geq \widetilde{g}(|z|/x_{j-1}) \cdot (t_0^{\alpha} - 1) \qquad \text{due to (6.60)}, \end{aligned}$$

whereas the mass of ν_c on $[x_j, x_{j-1})$ is

$$\nu_c[x_j/|z|, x_{j-1}/|z|) = \widetilde{g}(|z|/x_j) - \widetilde{g}(|z|/x_{j-1}) = \widetilde{g}(|z|/x_{j-1}) \cdot (t_0^{\alpha} - 1).$$

Therefore, given any $z \in \mathbb{R}$ such that $|z| \geq \widetilde{M}$, we have $\nu\big(E_j(z)\big) \geq \nu_c\big(E_j(z)\big)$ for all $j \in [N]$ where $E_j(z) = [x_j/|z|, x_{j-1}/|z|)$. This leads to

$$\begin{aligned} &I_2(z) \\ &= \frac{1}{|z|^{\alpha}} \sum_{j=1}^{N} \Big[\int_{E_j(z)} \Big(1 - \cos(zx)\Big) \nu(dx) - \int_{E_j(z)} \Big(1 - \cos(zx)\Big) c \frac{dx}{x^{1+\alpha}} \Big] \\ &\geq \frac{1}{|z|^{\alpha}} \sum_{j=1}^{N} \Big[\underline{m}_j \cdot \nu\big(E_j(z)\big) - \bar{m}_j \cdot \nu_c\big(E_j(z)\big) \Big] \end{aligned}$$

with $\bar{m}_j = \max\{h(z): \ z \in [x_j, x_{j-1}]\}$, $\underline{m}_j = \min\{h(z): \ z \in [x_j, x_{j-1}]\}$

$$
\begin{aligned}
&= \frac{1}{|z|^\alpha} \sum_{j=1}^N \Big[\underline{m}_j \cdot \nu\big(E_j(z)\big) - \underline{m}_j \cdot \nu_c\big(E_j(z)\big)\Big] + \frac{1}{|z|^\alpha} \sum_{j=1}^N \Big[\underline{m}_j \cdot \nu_c\big(E_j(z)\big) - \bar{m}_j \cdot \nu_c\big(E_j(z)\big)\Big] \\
&\geq 0 + \frac{1}{|z|^\alpha} \sum_{j=1}^N \Big[\underline{m}_j \cdot \nu_c\big(E_j(z)\big) - \bar{m}_j \cdot \nu_c\big(E_j(z)\big)\Big] \qquad \text{due to } \nu\big(E_j(z)\big) \geq \nu_c\big(E_j(z)\big) \\
&\geq -\frac{\delta}{|z|^\alpha} \sum_{j=1}^N \nu_c\big(E_j(z)\big) = -\frac{\delta}{|z|^\alpha} \nu_c[\theta/|z|, M/|z|) \qquad \text{due to (6.57)} \\
&= -\frac{\delta}{|z|^\alpha} \int_{\theta/|z|}^{M/|z|} c\frac{dx}{x^{1+\alpha}} \\
&\geq -\frac{\delta}{|z|^\alpha} \int_{\theta/|z|}^{\infty} c\frac{dx}{x^{1+\alpha}} = -c\delta/\theta^\alpha \\
&\geq -c \cdot \frac{C_0}{8} \qquad \text{due to (6.53).}
\end{aligned} \tag{6.61}
$$

Plugging (6.56) and (6.61) back into (6.55), we have shown that for all $|z| \geq \widetilde{M}$,

$$
\begin{aligned}
&\frac{1}{|z|^\alpha} \int_{(0,z_0)} \big(1 - \cos(zx)\big)\nu(dx) \\
&\geq -\frac{cC_0}{4} + \frac{1}{|z|^\alpha} \int_0^{M/|z|} \Big(1 - \cos(zx)\Big)c\frac{dx}{x^{1+\alpha}} \\
&= -\frac{cC_0}{4} - \frac{1}{|z|^\alpha} \int_{M/|z|}^{\infty} \Big(1 - \cos(zx)\Big)c\frac{dx}{x^{1+\alpha}} + \frac{1}{|z|^\alpha} \int_0^{\infty} \Big(1 - \cos(zx)\Big)c\frac{dx}{x^{1+\alpha}} \\
&\geq -\frac{cC_0}{2} + \frac{1}{|z|^\alpha} \int_0^{\infty} \Big(1 - \cos(zx)\Big)c\frac{dx}{x^{1+\alpha}} \qquad \text{using (6.54)} \\
&= -\frac{cC_0}{2} + \int_0^{\infty} \Big(1 - \cos y\Big)c\frac{dy}{y^{1+\alpha}} \qquad \text{by setting } y = |z|x \\
&= -\frac{cC_0}{2} + cC_0 = \frac{cC_0}{2} \qquad \text{by definition of } C_0 = \int_0^\infty (1 - \cos y)\frac{dy}{y^{1+\alpha}}.
\end{aligned}
$$

To conclude the proof of claim (6.50), we set $\widetilde{C} = cC_0/2$. $\square$

Again, the proof of Proposition 4.3 makes heavy use of the inversion formula. This time, we argue that the law of $Y^{(-z_0,z_0)}(t)$ is close enough to that of a semi-stable process.

Proof of Proposition 4.3. Recall that $z_0 = b^N$ for some $N \in \mathbb{Z}$. It suffices to prove the result for the Lévy process $\widetilde{Y} = Y^{(-z_0,z_0)}$. Let us denote the characteristic functions of $\widetilde{Y}(t)$ and $Y(t)$ as $\widetilde{\varphi}_t$ and φ_t respectively. The arguments about complex conjugates in (6.49) can be applied again to obtain

$$|\widetilde{\varphi}_t(z)| = \exp\Big(- t\int_{|x|<b^N} \big(1 - \cos(zx)\big)\nu(dx)\Big).$$

As for φ_t, using proposition 14.9 in [49], we get

$$|\varphi_t(z)| = \exp\Big(- t|z|^\alpha \eta(z)\Big) \tag{6.62}$$

where $\eta(z)$ is a non-negative function continuous on $\mathbb{R}\backslash\{0\}$ satisfying $\eta(bz) = \eta(z)$ and

$$\eta(z) = \frac{\int_{\mathbb{R}} \big(1 - \cos(zx)\big)\nu(dx)}{|z|^\alpha} \qquad \forall z \neq 0.$$

Furthermore, we claim that there is some $c > 0$ such that

$$\eta(z) \geq c \qquad \forall z \in [1, b]. \tag{6.63}$$

Then due to the self-similarity (i.e., $\eta(bz) = \eta(z)$), we have $\eta(z) \geq c$ for all $z \neq 0$. In the meantime, note that

$$\frac{1}{|z|^\alpha} \int_{|x| \geq b^N} \big(1 - \cos(zx)\big)\nu(dx) \leq \frac{\nu(|x| \geq b^N)}{|z|^\alpha}$$

and b, N are fixed. By picking $M > 0$ large enough, it holds for any $|z| \geq M$ that

$$\frac{1}{|z|^\alpha} \int_{|x| \geq b^N} \big(1 - \cos(zx)\big)\nu(dx) \leq \frac{c}{2}. \tag{6.64}$$

Therefore, for any $|z| \geq M$,

$$\begin{aligned}
\int_{|x| < b^N} \big(1 - \cos(zx)\big)\nu(dx) &= \int_{x \in \mathbb{R}} \big(1 - \cos(zx)\big)\nu(dx) - \int_{|x| \geq b^N} \big(1 - \cos(zx)\big)\nu(dx) \\
&= \eta(z) \cdot |z|^\alpha - \int_{|x| \geq b^N} \big(1 - \cos(zx)\big)\nu(dx) \\
&\geq c|z|^\alpha - \frac{c}{2}|z|^\alpha = \frac{c}{2}|z|^\alpha. \qquad \text{using (6.63) and (6.64)},
\end{aligned}$$

and hence $|\widetilde{\varphi}_t(z)| \leq \exp\big(-\frac{c}{2}t|z|^\alpha\big)$ for all $|z| \geq M$. Applying Inversion Formula (see theorem 3.3.5. of [23]), we get (for any $t > 0$)

$$\begin{aligned}
\Big\| f_{\widetilde{Y}(t)} \Big\|_\infty &\leq \frac{1}{2\pi} \int |\widetilde{\varphi}_t(z)| dz \\
&\leq \frac{1}{2\pi} \Big[2M + \int_{|z| \geq M} |\widetilde{\varphi}_t(z)| dz \Big] \\
&\leq \frac{M}{\pi} + \frac{1}{2\pi} \int \exp\Big(-\frac{c}{2} t |z|^\alpha\Big) dz \\
&= \frac{M}{\pi} + \frac{1}{2\pi} \cdot \frac{1}{t^{1/\alpha}} \int \exp\Big(-\frac{c}{2}|x|^\alpha\Big) dx \qquad \text{using } x = t^{1/\alpha} \cdot z \\
&\leq \frac{M}{\pi} + \frac{C_1}{t^{1/\alpha}} \qquad \text{where } C_1 = \frac{1}{2\pi} \int \exp\Big(-\frac{c}{2}|x|^\alpha\Big) dx.
\end{aligned}$$

To conclude the proof, we set $C = \frac{M}{\pi} + C_1$. Now it only remains to prove claim (6.63).

Proof of Claim (6.63)

We proceed with a proof by contradiction. Suppose there exists some $z \in [1, b]$ such that

$$\int_{\mathbb{R}} \big(1 - \cos(zx)\big)\nu(dx) = 0.$$

Now for any $\epsilon > 0$, define the following sets:

$$\begin{aligned}
S &= \{x \in \mathbb{R} : \ 1 - \cos(zx) > 0\} = \mathbb{R} \backslash \{\frac{2\pi}{z}k : \ k \in \mathbb{Z}\}; \\
S_\epsilon &= \{x \in \mathbb{R} : \ 1 - \cos(zx) \geq \epsilon\}.
\end{aligned}$$

Observe that

- For any $\epsilon > 0$, we have $\epsilon \cdot \nu(S_\epsilon) \le \int_{S_\epsilon} \big(1 - \cos(zx)\big)\nu(dx) \le \int_{\mathbb{R}} \big(1 - \cos(zx)\big)\nu(dx) = 0$, which implies $\nu(S_\epsilon) = 0$;

- As a result, $\lim_{\epsilon \to 0} \nu(S_\epsilon) = \nu(S)$.

Together with the fact that $\nu(\mathbb{R}) > 0$ (so that the process is non-trivial), we know the existence of some $m \in \mathbb{Z}, \delta > 0$ such that

$$\nu(\{\frac{2\pi}{z}m\}) = \delta > 0.$$

Besides, from $\nu(S) = 0$ we know that $\nu(-\frac{2\pi}{z}, \frac{2\pi}{z}) = 0$. However, due to (4.2) we know that $\nu = b^{-\alpha} T_b \nu$ 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

$$\nu(\{\frac{2\pi m}{z} b^{-k}\}) > 0 \quad \forall k = 1, 2, 3, \cdots$$

which contradicts $\nu(-\frac{2\pi}{z}, \frac{2\pi}{z}) = 0$ since, eventually for all $b > 0$ large enough, we have $\frac{2\pi m}{z} b^{-k} \in (-\frac{2\pi}{z}, \frac{2\pi}{z})$. This concludes the proof of $\eta(z) > 0$ for all $z \in [1, b]$. Lastly, since η is continuous on $z \in [1, b]$, we can find a strictly positive lower bound $c > 0$ such that $\eta(z) \ge c > 0$ for all $z \in [1, b]$. $\square$

References

[1] S. Asmussen and H. Albrecher. *Ruin Probabilities*. WORLD SCIENTIFIC, 2nd edition, 2010.

[2] S. Asmussen and P. W. Glynn. *Variance-Reduction Methods*, pages 126–157. Springer New York, New York, NY, 2007.

[3] S. Asmussen and D. P. Kroese. Improved algorithms for rare event simulation with heavy tails. *Advances in Applied Probability*, 38(2):545–558, 2006.

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

[5] A. Bassamboo, S. Juneja, and A. Zeevi. On the inefficiency of state-independent importance sampling in the presence of heavy tails. *Operations Research Letters*, 35(2):251–260, 2007.

[6] J. Blanchet and P. Glynn. Efficient rare-event simulation for the maximum of heavy-tailed random walks. *The Annals of Applied Probability*, 18(4):1351 – 1378, 2008.

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

[8] J. Blanchet and J. Liu. Efficient importance sampling in ruin problems for multidimensional regularly varying random walks. *Journal of Applied Probability*, 47(2):301–322, 2010.

[9] J. H. Blanchet and J. Liu. State-dependent importance sampling for regularly varying random walks. *Advances in Applied Probability*, 40(4):1104–1128, 2008.

[10] S. Borak, A. Misiorek, and R. Weron. *Models for heavy-tailed asset returns*, pages 21–55. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[11] O. Boxma, E. Cahen, D. Koops, and M. Mandjes. Linear networks: rare-event simulation and markov modulation. *Methodology and Computing in Applied Probability*, 2019.

[12] J. I. G. Cázares, A. Mijatović, and G. U. Bravo. ε-strong simulation of the convex minorants of stable processes and meanders. *Electronic Journal of Probability*, 25(none):1 – 33, 2020.

[13] L. Chaumont. On the law of the supremum of Lévy processes. *The Annals of Probability*, 41(3A):1191 – 1217, 2013.

[14] B. Chen, J. Blanchet, C.-H. Rhee, and B. Zwart. 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, 2019.

[15] J. E. Cohen, R. A. Davis, and G. Samorodnitsky. 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, 2022.

[16] L. Coutin, M. Pontier, and W. Ngom. Joint distribution of a lévy process and its running supremum. *Journal of Applied Probability*, 55(2):488–512, 2018.

[17] J. I. G. Cázares, F. Lin, and A. Mijatović. Fast exact simulation of the first passage of a tempered stable subordinator across a non-increasing function, 2023.

[18] P. Dupuis, K. Leder, and H. Wang. Importance sampling for sums of random variables with regularly varying tails. *ACM Trans. Model. Comput. Simul.*, 17(3):14–es, jul 2007.

[19] P. Dupuis, A. D. Sezer, and H. Wang. Dynamic importance sampling for queueing networks. *The Annals of Applied Probability*, 17(4):1306 – 1346, 2007.

[20] P. Dupuis and H. Wang. Importance sampling, large deviations, and differential games. *Stochastics and Stochastic Reports*, 76(6):481–508, 2004.

[21] P. Dupuis and H. Wang. On the convergence from discrete to continuous time in an optimal stopping problem. *The Annals of Applied Probability*, 15(2):1339 – 1366, 2005.

[22] P. Dupuis and H. Wang. Importance sampling for jackson networks. *Queueing Systems*, 62(1-2):113–157, 2009.

[23] R. Durrett. *Probability: Theory and Examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, 2019.

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

[25] A. Ferreiro-Castilla, A. Kyprianou, R. Scheichl, and G. Suryanarayana. Multilevel monte carlo simulation for lévy processes based on the wiener–hopf factorisation. *Stochastic Processes and their Applications*, 124(2):985–1010, 2014.

[26] S. Foss, D. Korshunov, S. Zachary, et al. *An introduction to heavy-tailed and subexponential distributions*, volume 6. Springer New York, NY, 2011.

[27] M. B. Giles. Multilevel monte carlo path simulation. *Operations Research*, 56(3):607–617, 2008.

[28] P. Glasserman and S.-G. Kou. Analysis of an importance sampling estimator for tandem queues. *ACM Trans. Model. Comput. Simul.*, 5(1):22–42, jan 1995.

[29] P. Glasserman and Y. Wang. Counterexamples in importance sampling for large deviations probabilities. *The Annals of Applied Probability*, 7(3):731 – 746, 1997.

[30] J. I. González Cázares, A. Mijatović, and G. Uribe Bravo. Geometrically convergent simulation of the extrema of lévy processes. *Mathematics of Operations Research*, 47(2):1141–1168, 2022.

[31] J. I. González Cázares, A. Mijatović, and G. U. Bravo. Exact simulation of the extrema of stable processes. *Advances in Applied Probability*, 51(4):967–993, 2019.

[32] T. Gudmundsson and H. Hult. Markov chain monte carlo for computing rare-event probabilities for a heavy-tailed random walk. *Journal of Applied Probability*, 51(2):359–376, 2014.

[33] M. Gurbuzbalaban, U. Simsekli, and L. Zhu. The heavy-tail phenomenon in sgd. In M. Meila and T. Zhang, editors, *Proceedings of the 38th International Conference on Machine Learning*, volume 139 of *Proceedings of Machine Learning Research*, pages 3964–3975. PMLR, 18–24 Jul 2021.

[34] S. Heinrich. Multilevel monte carlo methods. In S. Margenov, J. Waśniewski, and P. Yalamov, editors, *Large-Scale Scientific Computing*, pages 58–67, Berlin, Heidelberg, 2001. Springer Berlin Heidelberg.

[35] T. Hesterberg. Weighted average importance sampling and defensive mixture distributions. *Technometrics*, 37(2):185–194, 1995.

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

[37] M. M. Krzysztof Debicki. *Queues and Lévy Fluctuation Theory*. Springer Cham, 2015.

[38] A. Kuznetsov, A. E. Kyprianou, J. C. Pardo, and K. van Schaik. A Wiener–Hopf Monte Carlo simulation technique for Lévy processes. *The Annals of Applied Probability*, 21(6):2171 – 2190, 2011.

[39] M. Kwaśnicki, J. Małecki, and M. Ryznar. Suprema of Lévy processes. *The Annals of Probability*, 41(3B):2047 – 2065, 2013.

[40] Y. Li. Queuing theory with heavy tails and network traffic modeling. working paper or preprint, Oct. 2018.

[41] Z. Michna. Formula for the supremum distribution of a spectrally positive lévy process, 2012.

[42] Z. Michna. Explicit formula for the supremum distribution of a spectrally negative stable process. *Electronic Communications in Probability*, 18(none):1 – 6, 2013.

[43] Z. Michna and W. law Poland. Formula for the supremum distribution of a spectrally positive lévy process. *arXiv preprint arXiv:1104.1976*, 2011.

[44] Z. Michna, Z. Palmowski, and M. Pistorius. The distribution of the supremum for spectrally asymmetric lévy processes, 2014.

[45] J. Pitman and G. U. Bravo. The convex minorant of a Lévy process. *The Annals of Probability*, 40(4):1636 – 1674, 2012.

[46] S. I. Resnick. *Heavy-tail phenomena: probabilistic and statistical modeling*. Springer Science & Business Media, 2007.

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

[48] C.-H. Rhee and P. W. Glynn. Unbiased estimation with square root convergence for sde models. *Operations Research*, 63(5):1026–1043, 2015.

[49] K.-i. Sato, S. Ken-Iti, and A. Katok. *Lévy processes and infinitely divisible distributions*. Cambridge university press, 1999.

[50] P. Tankov. *Financial modelling with jump processes*. Chapman and Hall/CRC, 2003.

[51] G. Torrisi. Simulating the ruin probability of risk processes with delay in claim settlement. *Stochastic Processes and their Applications*, 112(2):225–244, 2004.

[52] M. Vihola. Unbiased estimators and multilevel monte carlo. *Operations Research*, 66(2):448–462, 2018.

[53] X. Wang and C.-H. Rhee. Large deviations and metastability analysis for heavy-tailed dynamical systems, 2023.