Electronic Journal of Probability

Electron. J. Probab. **0** (2023), article no. 0, 1–43.
ISSN: 1083-6489 https://doi.org/10.1214/YY-TN

Sample-Path Large Deviations for Lévy Processes and Random Walks with Lognormal Increments[*][†]

Zhe Su[‡] Chang-Han Rhee[§]

Abstract

The large deviations theory for heavy-tailed processes has seen significant advances in the recent past. In particular, [27] and [2] established large deviation asymptotics at the sample-path level for Lévy processes and random walks with regularly varying and (heavy-tailed) Weibull-type increments, respectively. Out of the three most prominent classes of heavy-tailed distributions—i.e., regularly varying, Weibull, and lognormal—this leaves the lognormal case open. This article establishes the *extended large deviation principle* (extended LDP) at the sample-path level for one-dimensional Lévy processes and random walks with lognormal-type increments. Building on these results, we also establish the extended LDPs for multi-dimensional processes with independent coordinates. We demonstrate the sharpness of these results by constructing counterexamples w.r.t. the M_1' topology, thereby proving that our results cannot be strengthened to a standard LDP under J_1 topology and M_1' topology.

Keywords: Sample Path Large Deviation; Lévy Process; Random Walk; Lognormal Tails.
MSC2020 subject classifications: 60F10.
Submitted to EJP on Dec 14, 2024, final version accepted on TBD.

1 Introduction

This paper develops the sample-path large deviations for Lévy processes and random walks whose increment distributions have lognormal-type tails. Let $\{X(t), t \geq 0\}$ be a centered Lévy process. We assume its Lévy measure ν is light-tailed on the negative half-line, and heavy-tailed on the positive half-line. In particular, we focus on the lognormal case, i.e., $\nu[x,\infty) = \exp(-r(\log x))$ where $r(x)$ is a regularly varying function with index $\gamma > 1$ as $x \to \infty$ under a very mild regularity condition. Consider a scaled process $\bar{X}_n(t) = X(nt)/n$ for $t \in [0,1]$. Similarly, consider a scaled and centered random walk $\bar{W}_n(t) = \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor}(Z_i - \mathbb{E}[Z_1])$ where $\mathbb{P}(Z_1 > x) = \exp(-r(\log x))$. The standard lognormal

[*]Supported by the Institute of Mathematical Statistics (IMS) and the Bernoulli Society.
[†]Current maintainer of class file is VTeX, Lithuania. Please send all queries to latex-support@vtex.lt.
[‡]Northwestern University, United States of America. E-mail: zhesu2022@u.northwestern.edu
[§]Northwestern University, United States of America.
E-mail: chang-han.rhee@northwestern.edu https://chrhee.github.io

distribution corresponds to the case $\gamma = 2$. We establish the sample-path large deviations of $\bar{X}_n$ and $\bar{W}_n$ and extends these results to multidimensional cases. Moreover, we demonstrates the sharpness of such results by showing that the classical large deviation principle doesn't hold in our contexts.

When the increment distributions are light-tailed, the large deviations of $\bar{X}_n$ and $\bar{W}_n$ have been thoroughly studied in probability theory. The classical theory of large deviations [30, 31, 11, 9, 29, 13, 12, 15] provide powerful tools for analyzing rare events. In particular, the sample-path level large-deviations bounds allow one to systematically characterize how a stochastic system deviates from their nominal behaviors for a wide variety of rare events [21].

In the heavy-tailed context, the seminal papers [23, 24] initiated the analysis of the tail aysmptotics of $\bar{W}_n(1)$, followed by vigorous research activities in the extreme value theory literature; see, for example, [6, 10, 14]. In particular, [10] assumes a very general class of heavy-tailed distributions in X and describes in detail how fast x needs to grow with n for the asymptotic relation

$$\mathbb{P}\left(X(n) > x\right) = n\mathbb{P}\left(X(1) > x\right)(1 + o(1)) \tag{1.1}$$

to hold, as $n \to \infty$. If (1.1) is valid, it is referred to as the *principle of a single big jump*. This insight was later generalized to the sample-path level in [18]. On the other hand, other related works, including [5, 16, 33], investigated the asymptotics of $\mathbb{P}\left(f(X) \in A\right)$ for many functionals f and sets A using ad-hoc approaches, revealing that rare events can also be governed by multiple jumps, rather than just one big jump.

More recently, [27] and [2] established sample-path large deviations for any number of large jumps, rather than a single large jump: they established asymptotic estimates of $\mathbb{P}\left(\bar{X}_n \in A\right)$, with A being sufficiently general sets of cádlág functions , thus facilitating a systematic way of studying rare events defined in terms of continuous functions of $\bar{X}_n$. They also clarified how the large deviations in the heavy tailed settings are connected to the standard large-deviations approach. More specifically, [2] examined large deviations for Lévy measures with Weibull tails, offering the logarithmic asymptotics

$$-\inf_{\xi \in A^\circ} I(x) \le \liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{X}_n \in A\right)}{\log n} \le \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{X}_n \in A\right)}{\log n} \le -\lim_{\epsilon\downarrow 0} \inf_{\xi \in A^\epsilon} I(x)$$

with the rate function

$$I(\xi) = \begin{cases} \sum_{t:\xi(t)\neq\xi(t-)} (\xi(t) - \xi(t-))^\alpha & \text{if } \xi \text{ is a non-decreasing pure jump path} \\ \infty & \text{otherwise} \end{cases}$$

For regularly varying Lévy measures, [27] established an exact asymptotics

$$C_{\mathcal{J}(A)}(A^\circ) \le \liminf_{n\to\infty} \frac{\mathbb{P}\left(\bar{X}_n \in A\right)}{n^{-\mathcal{J}(A)(\alpha-1)}} \le \limsup_{n\to\infty} \frac{\mathbb{P}\left(\bar{X}_n \in A\right)}{n^{-\mathcal{J}(A)(\alpha-1)}} \le C_{\mathcal{J}(A)}(\bar{A}),$$

where α is the index of Lévy measure's regularly varying tail, $\mathcal{J}(A)$ is the smallest number of jumps for a step function to be contained in A, and $C_{\mathcal{J}(A)}(\cdot)$ is a measure on the space of cádlág functions with $\mathcal{J}(A)$ or less jumps. Both asymptotic bounds imply that the rare events are driven by big jumps, hence characterizes the *catastrophe principle*.

Among the most important classes of tail distributions for modeling heavy-tailed phenomena—regularly varying, heavy-tailed Weibull, and lognormal distributions—the characterization of the catastrophe principle in the lognormal case remains open. This paper addresses this gap and establishes the sample-path large deviations for Lévy processes and random walks with lognormal increments.

Specifically, in Section 3.1, we establish the extended large deviation principle (extended LDP) for $\bar{X}_n$ under the Skorokhod J_1 topology:

$$-\inf_{\xi\in A^\circ} I(x) \le \liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{X}_n \in A\right)}{r(\log n)} \le \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{X}_n \in A\right)}{r(\log n)} \le -\lim_{\epsilon\downarrow 0}\inf_{\xi\in A^\epsilon} I(x) \quad (1.2)$$

where

$$I(\xi) = \begin{cases} \xi\text{'s number of jumps} & \text{if } \xi \in \mathbb{D}_{<\infty} \\ \infty & \text{otherwise.} \end{cases} \quad (1.3)$$

Recall that we assume $r(\cdot)$ is regularly with index $\gamma > 1$. $\mathbb{D}_{<\infty}$ is the space of non-decreasing step functions with finite number of jumps vanishing at the origin and continuous at 1. We accomplish this by first establishing the extended LDP for $\hat{J}_n^{\leqslant k}$(with respect to n), where $\hat{J}^{\leqslant k}$ represents, roughly speaking, the process constructed by taking the k largest jumps of the Lévy process. Then, we argue that the asymptotic behavior of $\hat{J}_n^{\leqslant k}$ governs that of $\bar{X}_n$ for sufficiently large k's. This allows us to obtain the extended LDP for $\bar{X}_n$ from that of $\hat{J}_n^{\leqslant k}$'s by leveraging the approximation lemma in [2].

Extended LDP was first introduced in [7]. Despite its strong-sounding name, extended LDP is actually a weaker statement than the standard LDP, as the upper bound in (1.2) involves the ϵ-fattening A^ϵ of the set A, hence, increasing the value of the upper bound.

We show that the standard LDP cannot be satisfied for the lognormal tails even under the coarser Skorokhod M_1' topology. Section 3.4 constructs a closed set w.r.t. the M_1' topology for which the standard LDP upper bound is violated. This contrasts to the conclusion in [2], where the extended LDP of $\bar{X}_n$ under the J_1 topology is strenghened to the standard LDP under the M_1' topology.

We also derive an extended LDP for random walks in section 3.2, assuming a lognormal-type tail in its increment distribution. In contrast to the Lévy process setting, the rate function here takes finite values for step functions that are discontinuous at time 1. This difference is due to the fact that rescaled random walk has a jump at the right boundary with a probability bounded from below. See Theorem 3.7 for the precise statement.

Many applications require modeling multiple sources of uncertainties. In such cases, large deviations for multidimensional processes provide the means to model such systems. For example, [1, 3] analyze the queue length asymptotics for the multiple server queues and stochastic fluid networks with heavy-tailed Weibull service times based on the large deviations results in [2] for multi-dimensional Lévy processes and random walks. Section 3.3 obtains extended LDP for multi-dimensional processes with independent components $\bar{\boldsymbol{X}}_n = (\bar{X}_n^{(1)}, \cdots, \bar{X}_n^{(d)})$. Here $\bar{X}_n^{(i)}$'s are centered and scaled 1-dimensional Lévy processes or random walks independent of each other. As the rate function in (1.3) is not good—i.e., does not have compact level set—the standard results such as Theorem 4.14 of [17] do not apply directly in our context, and the derivation of the extended LDP for $\bar{\boldsymbol{X}}_n$ from those of $\bar{X}_n^{(1)}$ requires careful justification. We take advantage of the discrete nature of $\bar{X}_n$ and $\bar{W}_n$'s rate functions to establish the extended LDP for a d-fold product of probability measures, where each coordinate satisfies an extended LDP with a rate function that only takes at most countable values.

Section 2 provides preliminaries, Section 3 presents the main results of the paper, and Section 4 presents most of the proofs for the results in Sections 2 and 3.

2 Preliminaries

This section provides preliminary results useful for later sections. All proofs in this subsection are deferred to Section 4.1. We begin by introducing recurring notations. Let $(\mathcal{X}, d)$ denote a metric space $\mathcal{X}$ equipped with a metric d. For a set A let A^c denote the complement of A. Let $d(x,A) \triangleq \inf_{y\in A} d(x,y)$, $B_r(x) \triangleq \{y \in \mathcal{X} : d(x,y) < r\}$, $A^\epsilon \triangleq \{x \in \mathcal{X} : d(x,A) \le \epsilon\}$, and $A^{-\epsilon} \triangleq \Big(\big(A^c\big)^\epsilon\Big)^c$ denote the distance between x and A, the open ball with radius r centered at x, the closed ϵ-fattening of A, and the open ϵ-shrinking of A, respectively.

We denote the the space of real-valued càdlàg functions from $[0,1]$ to $\mathbb{R}$—i.e., the Skorokhod space—by $\mathbb{D}[0,1]$ or simply $\mathbb{D}$ when the domain $[0,1]$ is clear. Let

$$d_{\|\cdot\|}(\xi,\zeta) = \|\xi-\zeta\| = \sup_{t\in[0,1]} |\xi(t)-\zeta(t)|$$

denote the uniform metric and recall the J_1 metric

$$d_{J_1}(\xi,\zeta) = \inf_{\lambda\in\Lambda} \|\lambda - e\| \vee \|\xi\circ\lambda - \zeta\|,$$

where Λ is the collection of all non-decreasing homeomorphisms on $[0,1]$, and e is the identity map on $[0,1]$. The M_1' metric on $\mathbb{D}[0,1]$ is defined as follows. Let $\Gamma(\xi)$ denote the extended completed graph of $\xi \in \mathbb{D}$. That is,

$$\Gamma(\xi) = \{(y,t) \in \mathbb{R}\times[0,1] : y \in [\xi(t-)\wedge\xi(t), \xi(t-)\vee\xi(t)]\} \tag{2.1}$$

with the convention that $\xi(0-) \triangleq 0$. Note that if $\xi(0) \neq 0$, then one can consider ξ as a path with a jump at time 0. Roughly speaking, $\Gamma(\xi)$ is the union of the graph of ξ and the vertical lines that concatenate the connected pieces. Define an order '$\prec$' on $\Gamma(\xi)$ as follows: for two points on $\Gamma(\xi)$, we say that $(y_1,t_1) \prec (y_2,t_2)$ if either $t_1 < t_2$, or $t_1 = t_2$ and $|\xi(t_1-) - y_1| < |\xi(t_2-) - y_2|$. A continuous, nondecreasing (w.r.t. $\prec$), and surjective function $(u,r) : [0,1] \to \Gamma(\xi)$ is called a parametrization of $\Gamma(\xi)$. Let $\Pi(\xi)$ denote the collection of all parametrizations of $\Gamma(\xi)$. The M_1' metric is

$$d_{M_1'}(\xi,\zeta) = \inf_{\substack{(u_1,r_1)\in\Pi(\xi)\\(u_2,r_2)\in\Pi(\zeta)}} \|u_1-u_2\| \vee \|r_1-r_2\|. \tag{2.2}$$

We will often consider multiple paths in $\mathbb{D}$ and work with their extended completed graphs as subsets of $\mathbb{R}\times[0,1]$. We use the ℓ_∞ distance in this space, i.e., $d\big((x,t),(y,s)\big) = |x-y| \vee |t-s|$ for $(x,t),(y,s) \in \mathbb{R}\times[0,1]$. See [2] for the full details of M_1' metric on $\mathbb{D}[0,1]$ and [25] or [32] for M_1' metric on $\mathbb{D}[0,\infty)$. It is obvious from their definitions that $d_{\|\cdot\|} \ge d_{J_1}$. Note also that the M_1' distance is upper bounded by the M_1 distance, and hence, Theorem 12.3.2 in [32] implies that $d_{J_1} \ge d_{M_1'}$. The next two simple observations are useful for bounding the M_1' distances from below. Their proof are provided in Section 4.1.

Lemma 2.1. *Suppose that $\xi,\zeta \in \mathbb{D}$. Then for any $(u,r) \in \Gamma(\xi)$,*

$$d_{M_1'}(\xi,\zeta) \ge d\big((u,r),\Gamma(\zeta)\big).$$

Lemma 2.2. *Suppose that $\xi,\zeta \in \mathbb{D}$. If there exists $s,t \in [0,1]$ and $\delta > 0$ such that $|\zeta(t) - \zeta(s-)| \ge 2\delta$, and ξ is constant on $[s-\delta, t+\delta]\cap[0,1]$, then*

$$d_{M_1'}(\xi,\zeta) \ge \delta.$$

Throughout this paper, we consider the following subsets of $\mathbb{D}$:

$$\begin{aligned}
&\mathbb{D}^{\uparrow} \triangleq \{\xi \in \mathbb{D} : \xi \text{ is a non-decreasing, pure jump function}\};\\
&\mathbb{D}_{=j} \triangleq \{\xi \in \mathbb{D}^{\uparrow} : \xi(0) = 0,\ \xi(1) = \xi(1-),\ \xi \text{ has } j \text{ jumps}\};\\
&\mathbb{D}_{\leqslant j} \triangleq \cup_{i=0}^{j} \mathbb{D}_{=i}, \quad \mathbb{D}_{<\infty} \triangleq \cup_{j=0}^{\infty} \mathbb{D}_{\leqslant j};\\
&\tilde{\mathbb{D}}_{=j} \triangleq \{\xi \in \mathbb{D}^{\uparrow} : \xi(0) = 0,\ \xi \text{ has } j \text{ jumps}\};\\
&\tilde{\mathbb{D}}_{\leqslant j} \triangleq \cup_{i=0}^{j} \tilde{\mathbb{D}}_{=i}, \quad \tilde{\mathbb{D}}_{<\infty} \triangleq \cup_{j=0}^{\infty} \tilde{\mathbb{D}}_{\leqslant j};\\
&\hat{\mathbb{D}}_{=j} \triangleq \{\xi \in \mathbb{D}^{\uparrow} : \xi(0) \geq 0,\ \xi \text{ has } j \text{ jumps}\};\\
&\hat{\mathbb{D}}_{\leqslant j} \triangleq \cup_{i=0}^{j} \hat{\mathbb{D}}_{=i}, \quad \hat{\mathbb{D}}_{<\infty} \triangleq \cup_{j=0}^{\infty} \hat{\mathbb{D}}_{\leqslant j}.
\end{aligned}$$

It's clear that $\mathbb{D}_{=j} \subset \tilde{\mathbb{D}}_{=j} \subset \hat{\mathbb{D}}_{=j}$ for any $j \in \mathbb{N}$, and hence, $\mathbb{D}_{<\infty} \subset \tilde{\mathbb{D}}_{<\infty} \subset \hat{\mathbb{D}}_{<\infty}$. It can be verified that $\mathbb{D}_{\leqslant j}$, $\tilde{\mathbb{D}}_{\leqslant j}$ and $\hat{\mathbb{D}}_{\leqslant j}$ are closed subsets of $(\mathbb{D}, d_{J_1})$. In Lemma 4.3, we show that $\hat{\mathbb{D}}_{\leqslant j}$ is closed in $(\mathbb{D}, d_{M_1'})$ as well.

We conclude this section with a brief review of the formal definition of *extended large deviation principle* (extended LDP) introduced in [7]. Extended LDP is often useful in heavy-tailed contexts where the standard LDP fails to hold.

Definition 2.3. *A sequence of measures $\{\mu_n\}_{n\geq 1}$ satisfies the extened LDP with speed a_n on $(\mathcal{X}, d)$ if there is a rate function $I : \mathcal{X} \to \mathbb{R}_+$ and a sequence $\{a_n\}_{n\geq 1}$ with $a_n \to \infty$ such that, for any open set G and any closed set F, the following inequalities hold:*

$$-\inf_{x\in G} I(x) \leq \liminf_{n\to\infty} \frac{\log \mu_n(G)}{a_n} \qquad \textit{and} \qquad \limsup_{n\to\infty} \frac{\log \mu_n(F)}{a_n} \leq -\lim_{\epsilon\downarrow 0} \inf_{x\in F^\epsilon} I(x).$$

For a sequence of $\mathcal{X}$-valued random variables $\{X_n\}_{n\geq 1}$, we say that X_n satisfies the extended LDP on $(\mathcal{X}, d)$ if $\{\mathbb{P}(X_n \in \cdot)\}_{n\geq 1}$ satisfies the extended LDP on $(\mathcal{X}, d)$. The rate function I in Definition 2.3 is uniquely determined; see [7].

We present a couple of tools that are useful for establishing extended LDPs:

Lemma 2.4. *Let $\{X_n\}_{n\geq 1}$ be a sequence of càdlàg stochastic processes. Let $\mathbb{E} \subset \mathbb{D}$ be a closed set such that $\mathbb{P}(X_n \in \mathbb{E}) = 1$ for all $n \in \mathbb{N}$. If $\{X_n\}_{n\geq 1}$ satisfies the extended LDP on $(\mathbb{E}, d)$ with the rate function I and speed a_n. Then $\{X_n\}_{n\geq 1}$ satisfies the extended LDP on $(\mathbb{D}, d)$ with the same speed sequence and the rate function given by*

$$I'(x) = \begin{cases} I(x) & x \in \mathcal{E} \\ \infty & x \notin \mathcal{E} \end{cases}$$

Lemma 2.4 is proved in Section 4.1.

Proposition 2.5 (Proposition 2.1 of [3])**.** *Let $\{X_n\}_{n\geq 1}$ and $\{Y_n^k\}_{n\geq 1}$ for each $k \in \mathbb{N}$ be sequence of random objects in a metric space $(\mathcal{X}, d)$. Let I and I_k for $k \in \mathbb{N}$ be non-negative lower-semicontinuous functions. Suppose that the following conditions hold:*

(1) For each $k \in \mathbb{N}$, the sequence $\{Y_n^k\}_{n\geq 1}$ satisfy the extended LDP with the rate function I_k and the speed $\{a_n\}_{n\geq 1}$.

(2) For each closed set F,
$$\lim_{k\to\infty} \inf_{x\in F} I_k(x) \geq \inf_{x\in F} I(x). \tag{2.3}$$

(3) For each $\delta > 0$ and each open set G, there exists $\epsilon > 0$ and $K \geq 0$ such that $k \geq K$ implies
$$\inf_{x\in G^{-\epsilon}} I_k(x) \leq \inf_{x\in G} I(x) + \delta. \tag{2.4}$$

(4) For every $\epsilon > 0$, it holds that

$$\lim_{k\to\infty} \limsup_{n\to\infty} \frac{1}{a_n} \log \mathbb{P}\left(d(X_n, Y_n^k) > \epsilon\right) = -\infty. \tag{2.5}$$

Then $\{X_n\}_{n\geq 1}$ satisfy the extended LDP with the rate function I and speed a_n.

When the sequences $\{Y_n^k\}_{n\geq 1}$ and the rate functions I_k's do not change with k, the above proposition directly implies the following corollary.

Corollary 2.6 (Corollary 2.1 of [3])**.** *Assume both the sequence $\{X_n\}_{n\geq 1}$ and $\{Y_n\}_{n\geq 1}$ take values in a metric space $(\mathcal{X}, d)$ and the following relation holds:*

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(d(X_n, Y_n) > \epsilon\right)}{a_n} = -\infty. \tag{2.6}$$

If $\{X_n\}_{n\geq 1}$ satisfies the extended LDP with the rate function I and speed $\{a_n\}_{n\geq 1}$, then $\{Y_n\}_{n\geq 1}$ satisfies the extended LDP with the same rate function and speed.

3 Main Results

This section presents the main results of the paper. Section 3.1 establishes the extended LDP at the sample-path level for one-dimensional Lévy processes, while Section 3.2 establishes it for one-dimensional random walks. Section 3.3 establishes the extended LDP for multi-dimensional processes with independent coordinates. Section 3.4 constructs a counterexample that shows our extended LDPs cannot be strengthened to a standard LDP. Section 3.5 illustrates the applicability of our results with a boundary crossing problem.

3.1 Extended LDP for Lévy processes

According to the Lévy-Ito decomposition (see, for example, Chapter 2 of [28]), the Lévy process with triplets (a, b, ν) has the following distributional representation: for $t \in [0, \infty)$

$$X(t) = bt + aB(t) + \int_{|x|<1} x\left(\hat{N}([0,t]\times dx) - t\nu(dx)\right) + \int_{|x|\geq 1} x\hat{N}([0,t]\times dx), \tag{3.1}$$

where a and b are scalars, B is the standard Brownian motion, and ν is a Lévy measure, i.e., a σ-finite measure supported on $\mathbb{R}\backslash\{0\}$ that satisfies $\int \min(1, |x|^2)\nu(dx) < \infty$. $\hat{N}$ is a Poisson random measure on $[0,\infty)\times(0,\infty)$ with mean measure $\mathbf{LEB}\times\nu$, where LEB is the Lebesque measure on $\mathbb{R}$. Throughout the rest of this paper, we make the following tail assumption on the Lévy measure ν.

Assumption 3.1. *ν is supported on $\mathbb{R}$.*

- *(lognormal-type right tail) For $x > 0$,*

$$\nu[x,\infty) = \exp(-r(\log x)) \tag{3.2}$$

where $r(\cdot)$ is a regularly varying function with index $\gamma > 1$. Additionally, there exists some $\gamma' \in (0,\gamma)$ such that

$$\liminf_{x\to\infty} \frac{r(x+c) - r(x)}{\exp\{-x^{\gamma'}\}} \geq 1 \tag{3.3}$$

for each $c > 0$.

- *(light left tail) For some $t > 0$,*

$$\int_{-\infty}^{0} e^{-tx}\nu(dx) < \infty. \tag{3.4}$$

That is, ν's left tail is light.

Remark 3.2. Under Assumption 3.1, the right tail of ν is heavier than those of any Weibull tails and lighter than any power law tails. Assumption 3.1 includes the case $\nu[x,\infty) = x^{\beta}\exp(-(\log x)^{\gamma})$ for some $\beta > 0$ and $\gamma > 1$. A special case, $\gamma = 2$, corresponds to the standard lognormal tail.

Remark 3.3. Note that the extra assumption (3.3) is very mild since it rapidly becomes less restrictive as x grows. However, Theorem 3.4 does not hold without any constraints, since the potential irregularity of the slowly varying function in $r(\cdot)$ can be amplified by the exponential function in (3.2), leading to pathological asymptotic behaviors.

Consider a scaled process

$$\bar{X}_n(t) = \frac{1}{n}\left(X(nt) - bnt - nt\int_{|x|\geq 1} x\nu(dx)\right). \tag{3.5}$$

The main result of this section is the following extended LDP for $\bar{X}_n$. Its proof is provided in Section 4.2.

Theorem 3.4. *The sequence $\{\bar{X}_n\}_{n\geq 1}$ of scaled processes defined in (3.5) satisfies the extended LDP in $(\mathbb{D}, d_{J_1})$ with the rate function*

$$I^{J_1}(\xi) = \begin{cases} \sum_{t\in[0,1]} \mathbb{1}\{\xi(t) \neq \xi(t-)\} & \textit{if } \xi \in \mathbb{D}_{<\infty} \\ \infty & \textit{otherwise} \end{cases} \tag{3.6}$$

and speed $r(\log n)$.

As is often the case with heavy-tailed large deviation results, Theorem 3.4 cannot be strengthened to the standard LDP w.r.t. the J_1 topology. However, one may naturally wonder if the standard LDP holds w.r.t. weaker topologies. For example, [2] establishes the standard LDP w.r.t. the M_1' topology, while showing that the standard LDP w.r.t. the J_1 topology is impossible in the heavy-tailed Weibull case. It turns out that the counterexample in Section 3.4 proves that even w.r.t. the M_1' topology, the standard LDP cannot be satisfied. Here we also mention that the rate function I^{J_1} fails to be lower-semicontinuous under the M_1' topology. To see this, consider $\xi = \mathbb{1}([0,1])$ and $\xi_n = \mathbb{1}([1/n,1])$ for $n \in \mathbb{N}$. It's straightforward to verify that $\xi_n \to \xi$ under the M_1' topology, but $\lim_{n\to\infty} I^{J_1}(\xi_n) < I^{J_1}(\xi)$.

Although the standard LDP cannot be satisfied, the extended LDP can be established w.r.t. the M_1' topology with a slightly different rate function. The next theorem establishes the extended LDP for $\bar{X}_n$ w.r.t. the M_1' topology. It should be clear from the proof of Theorem 3.5 that the M_1' extended LDP in the theorem doesn't provide any new useful bounds other than those that are already implied by the J_1 counterpart in Theorem 3.4. However, we state the next theorem here and prove it in Section 4.2 for the purpose of completeness.

Theorem 3.5. *The sequence $\{\bar{X}_n\}_{n\geq 1}$ defined in (3.5) satisfy the extended LDP in $(\mathbb{D}, d_{M_1'})$ with the rate function given by*

$$I^{M_1'}(\xi) = \begin{cases} \sum_{t\in[0,1]} \mathbb{1}\{\xi(t) \neq \xi(t-)\} & \xi \in \hat{\mathbb{D}}_{<\infty} \\ \infty & \textit{otherwise} \end{cases} \tag{3.7}$$

and speed $r(\log n)$.

3.2 Extended LDP for random walks

Let Z_i's be i.i.d. random variables, and consider the random walk $W(t) = \sum_{i=1}^{\lfloor t \rfloor} Z_i$ embedded in $\mathbb{D}$. Throughout the rest of the paper, we make the following assumption on the tail distributions of Z_i's:

Assumption 3.6. *Z_i is supported on $\mathbb{R}$.*

- *(lognormal-type right tail) For $x > 0$,*

$$\mathbb{P}(Z_i \geq x) = \exp(-r(\log x))$$

  where $r(\cdot)$ is a regularly varying function with index $\gamma > 1$. Additionally, there exists some $\gamma' \in (0, \gamma)$ such that

$$\liminf_{x\to\infty} \frac{r(x+c) - r(x)}{\exp\{-x^{\gamma'}\}} \geq 1.$$

  for each $c > 0$.

- *(light left tail) For some $t > 0$,*

$$\int_{-\infty}^{0} e^{-tx} \mathbb{P}(Z_i \in dx) < \infty.$$

We establish the extended LDP for the scaled processes:

$$\bar{W}_n(t) = \frac{1}{n} \sum_{i=1}^{\lfloor nt \rfloor} (Z_i - \mathbb{E}[Z_1]). \tag{3.8}$$

Unlike the Lévy process $\bar{X}_n$ in section 3.1, the random walks $\bar{W}_n$ always exhibit a jump at $t = 1$. This makes the rate function finite on $\tilde{\mathbb{D}}_{<\infty}$, not just $\mathbb{D}_{<\infty}$. To recall the definitions of $\mathbb{D}_{<\infty}$ and $\tilde{\mathbb{D}}_{<\infty}$, see section 2. The following is the main result of this subsection.

Theorem 3.7. *The sequence of random walk $\{\bar{W}_n\}_{n\geq 1}$ satisfies the extended large deviation principle on $(\mathbb{D}, J_1)$ with the rate function*

$$\tilde{I}(\xi) = \begin{cases} \sum_{t\in[0,1]} \mathbb{1}\{\xi(t) \neq \xi(t-)\} & \text{if } \xi \in \tilde{\mathbb{D}}_{<\infty} \\ \infty & \text{otherwise} \end{cases} \tag{3.9}$$

and speed $r(\log n)$.

The proof of Theorem 3.7 is provided in Section 4.3.

3.3 Extended LDP for multi-dimensional processes

This section establishes the extended LDP for multi-dimensional Lévy processes and random walks, each of whose coordinates are independent, and their increment distributions are lognormal-type. We accomplish this by building on the one-dimensional results in Section 3.1 and 3.2. Note that sice we are working with extended LDPs—as opposed to standard LDPs—and the rate functions that controls our extended LDPs are not good, standard results such as Theorem 4.14 of [17] cannot be directly applied in our context. However, the discrete nature of our rate function enables us to adapt the key proof ideas in Theorem 4.14 of [17] to establish the same result in our context in Proposition 3.8 despite the lack of goodness of our rate function. Theorem 3.8 is formulated in general metric spaces. Specifically, we consider the following sequence of random variables:

$$\boldsymbol{X}_n = (X_n^{(1)}, \cdots, X_n^{(k)})$$

where for each $i = 1, \ldots, k$,

(i) $X_n^{(i)}$ is a random object taking value in the metric space $(\mathcal{X}^{(i)}, d^{(i)})$.

(ii) $\{X_n^{(i)}\}_{n\geq 1}$ satisfies the extended LDP with the rate function $I^{(i)}$ and speed a_n

(iii) $I^{(i)}$ takes at most countable distinct values in $\mathbb{R}_+$ with no limit point.

Proposition 3.8. *Suppose that $\{\boldsymbol{X}_n\}_{n\geq 1}$ satisfies conditions (i), (ii), and (iii) above. Then $\{\boldsymbol{X}_n\}_{n\geq 1}$ satisfies the extended LDP on the product space $\left(\prod_{i=1}^k \mathcal{X}^{(i)}, \max_{i=1}^k d^{(i)}\right)$ with the rate function $\bar{I} = \sum_{i=1}^k I^{(i)}$ and speed a_n.*

The proof of Proposition 3.8 is provided in Section 4.4.

Moving back to Lévy processes, for $i = 1, \cdots, d$, let $X^{(i)}$ be an independent one-dimensional Lévy process with generating triplet (a_i, b_i, ν_i). Each ν_i satisfies Assumption 3.1 with $r(\cdot)$ replaced by $r_i(\cdot)$, and there exists a regularly varying $r_0(\cdot)$ such that $\lim_{x\to\infty} \frac{r_i(x)}{r_0(x)} = \lambda_i$. Consider the scaled and centered processes

$$\bar{X}_n^{(i)}(t) = \frac{1}{n}\big[X^{(i)}(nt) - b_i nt - nt\int_{|x|\geq 1} x\nu_i(dx)\big]$$

for $t \in [0,1]$ and $i = 1, \ldots, d$. Theorem 3.4 ensures that each $\{\bar{X}_n^{(i)}\}_{n\geq 1}$ on $(\mathbb{D}, d_{J_1})$ satisfies the extended LDP. By transforming extended LDP speed sequence of $\{\bar{X}_n^{(i)}\}_{n\geq 1}$'s to $\{r_0(\log n)\}_{n\geq 1}$, it is straightforward to see that the corresponding rate function is of the form $\lambda_i \cdot I^{J_1}$, which only takes values in the set $\{k\lambda_i : k \in \mathbb{N}\}$. Hence those d Lévy processes meet requirements (i), (ii) and (iii), and this directly leads to the following implication of Proposition 3.8:

Theorem 3.9. *On $\left(\prod_{i=1}^d \mathbb{D}, \sum_{i=1}^d d_{J_1}\right)$, the multidimensional Lévy process with independent coordinates $\bar{\boldsymbol{X}}_n = (\bar{X}_n^{(1)}, \cdots \bar{X}_n^{(d)})$ satisfies the extended LDP with the rate function $I^d : \prod_{i=1}^d \mathbb{D} \to \mathbb{R}$ and speed $r_0(\log n)$, where*

$$I^d(\xi_1, \cdots, \xi_d) \triangleq \begin{cases} \sum_{i=1}^d \lambda_i \cdot \sum_{t\in[0,1]} \mathbb{1}\{\xi_i(t) \neq \xi_i(t-)\} & \text{if } \boldsymbol{\xi} \in \prod_{i=1}^d \mathbb{D}_{<\infty} \\ \infty & \text{otherwise} \end{cases}. \tag{3.10}$$

Likewise, the extended LDP for the multidimensional random walks follows directly from Proposition 3.8.

Theorem 3.10. *On $\left(\prod_{i=1}^d \mathbb{D}, \sum_{i=1}^d d_{J_1}\right)$, the sequence $\{\bar{\boldsymbol{S}}_n = (\bar{S}_n^{(1)}, \cdots \bar{S}_n^{(d)})\}_{n\geq 1}$ with independent coordinates satisfies the extended LDP with speed $\log^\gamma n$ and the rate function $\tilde{I}^d : \prod_{i=1}^d \mathbb{D} \to \mathbb{R}_+$:*

$$\tilde{I}^d(\boldsymbol{\xi}) \triangleq \begin{cases} \sum_{i=1}^d \lambda_i \cdot \sum_{t\in[0,1]} \mathbb{1}\{\xi_i(t) \neq \xi_i(t-)\} & \text{if } \boldsymbol{\xi} \in \prod_{i=1}^d \tilde{\mathbb{D}}_{<\infty} \\ \infty & \text{otherwise} \end{cases} \tag{3.11}$$

3.4 Nonexistence of standard LDPs

In subsection 3.1, we established the extend LDP for Lévy processes under the J_1 topology and M_1' topology. A natural question is whether these results can be strengthened to establish the standard LDPs. Theorem 3.11 below constructs a counterexample and confirms that the extended LDP in Theorem 3.4 can not be strengthened to a standard LDP.

Theorem 3.11. *Recall $\bar{X}_n$ given in (3.5). Assume the Lévy measure associated with $\bar{X}_n$ satisfies Assumption 3.1 and is supported on $\mathbb{R}_+$. Then $\{\bar{X}_n\}_{n\geq 1}$ does not satisfy the LDP neither in the M_1' topology nor the J_1 topology. Since the M_1' metric is bounded by the J_1 metric, this also implies that the $\bar{X}_n$ cannot satisfy an LDP in the J_1 topology.*

Before proving Theorem 3.11, we set some notations and state two lemmata—Lemma 3.12 and 3.13—whose proofs are deferred to Section 4.5. Consider the following mapping:

$$\begin{aligned} \pi : \hat{\mathbb{D}}_{<\infty} &\longrightarrow \hat{\mathbb{D}}_{\leqslant 2} \\ \xi &\longmapsto \pi(\xi) = \begin{cases} \xi & \xi \in \hat{\mathbb{D}}_{\leqslant 2} \\ a_1 \mathbb{1}_{[t_1,1]} + a_2 \mathbb{1}_{[t_2,1]} & \text{otherwise} \end{cases}, \end{aligned} \tag{3.12}$$

where a_1 and a_2 are the first and second largest jump sizes of ξ, and t_1 and t_2 are the earliest times for those jumps. Consider the following sets:

$$A_n \triangleq \left\{ \xi = \sum_{i=1}^{2} z_i \mathbb{1}_{[v_i,1]} : z_1 \in [\log n, \infty), z_2 \in \left[\frac{1}{n^{1/3}}, \infty\right), z_1 \geq z_2, v_1 \in \left(\frac{1}{4}, \frac{1}{2}\right], v_2 \in \left(\frac{3}{4}, 1\right] \right\} \tag{3.13}$$

$$B_n \triangleq \pi^{-1}(A_n) \tag{3.14}$$

$$C_n \triangleq \left\{ \xi \in \mathbb{D} : d_{\|\cdot\|}(\xi, -\mu_1\nu_1 e) \leq \frac{1}{3}\frac{1}{n^{1/3}} \right\} \tag{3.15}$$

$$F_n \triangleq \left\{ \eta \in \mathbb{D} : \eta = \xi_1 + \xi_2, \xi_1 \in B_n, \xi_2 \in C_n \right\} \tag{3.16}$$

For some N such that $\log N - \frac{1}{2}\nu_1\mu_1 - \frac{1}{3}N^{-\frac{1}{3}} > 1$, we further define

$$F \triangleq \cup_{n=N}^{\infty} F_n \tag{3.17}$$

The following two lemmata are key to Theorem 3.11, and those lemmas' proof are left to Section 4.5.

Lemma 3.12. *On $(\mathbb{D}, d_{M_1'})$, the set F defined in (3.17) satisfies $\bar{F} \cap \hat{\mathbb{D}}_{\leqslant 1} = \emptyset$. Therefore,*

$$-\inf_{\xi \in \bar{F}} I^{M_1'}(\xi) \leq -2. \tag{3.18}$$

Lemma 3.13. *The set F defined in (3.17) satisfies*

$$-2 < \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{X}_n \in \bar{F}\right)}{r(\log n)}. \tag{3.19}$$

Proof of Theorem 3.11. Suppose that the LDP holds w.r.t. the M_1' topology. Then the upper bound of the LDP along with Lemma 3.12 and Lemma 3.13 imply the following contradictory statement:

$$-2 < \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{X}_n \in \bar{F}\right)}{r(\log n)} \leq -\inf_{\xi \in \bar{F}} I^{M_1'}(\xi) \leq -2.$$

Therefore, the LDP cannot hold w.r.t. the M_1' topology. Since the M_1' metric is bounded by the J_1 metric, this also implies that the $\bar{X}_n$ cannot satisfy an LDP in the J_1 topology. □

3.5 Example: boundary crossing with regulated jumps

This section illustrates an application of Theorem 3.4 to analyze the large deviation probabilities associated with level crossing events. Specifically, we consider the probability

$$\mathbb{P}\left(\sup_{t\in[0,1]} \bar{X}_n \geq b, \sup_{t\in[0,1]} |\bar{X}_n(t) - \bar{X}_n(t-)| \leq c \right). \tag{3.20}$$

This probability is closely related to the insolvency risk of reinsured insurance lines in actuarial science. To facilitate the use of extended LDP, we introduce the mapping

$$\begin{aligned}\phi : \mathbb{D} \longrightarrow & \mathbb{R}^2 \\ \xi \longmapsto & \Big(\sup_{t\in[0,1]} \xi(t), \ \sup_{t\in[0,1]} \big|\xi(t)-\xi(t-)\big| \Big),\end{aligned} \tag{3.21}$$

which captures the maximum value of a function ξ and its largest jump. The probability in (3.20) can be rephrased as follows:

$$\mathbb{P}\left(\phi(\bar{X}_n) \in [b,\infty)\times[0,c]\right). \tag{3.22}$$

The Lipschitz continuity of ϕ (as discussed in Section 4 of [2]) enables us to apply the contraction principle (Lemma B.3 in [3]) to the extended LDP of $\{\bar{X}_n\}_{n\geq 1}$ established in Theorem 3.4. Consequently, we deduce that $\{\phi(\bar{X}_n)\}_{n\geq 1}$ satisfies the extended LDP with the rate function given by

$$I_\phi(x,y) = \inf\left\{ I^{J_1}(\xi) : \sup_{t\in[0,1]} \xi(t) = x, \ \sup_{t\in[0,1]} |\xi(t)-\xi(t-)| = y \right\} = \left\lceil \frac{x}{y} \right\rceil.$$

That is,

$$\begin{aligned} -\inf_{(x,y)\in([b,\infty)\times[0,c])^\circ} \left\lceil \frac{x}{y} \right\rceil &\leq \liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\phi(\bar{X}_n)\in[b,\infty)\times[0,c]\right)}{r(\log n)} \\ &\leq \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\phi(\bar{X}_n)\in[b,\infty)\times[0,c]\right)}{r(\log n)} \leq -\lim_{\epsilon\downarrow 0} \inf_{(x,y)\in([b,\infty)\times[0,c])^\epsilon} \left\lceil \frac{x}{y} \right\rceil. \end{aligned}$$

If b/c is not an integer, this yields a tight asymptotic limit

$$\lim_{n\to\infty} \frac{\log \mathbb{P}\left(\phi(\bar{X}_n)\in[b,\infty)\times[0,c]\right)}{r(\log n)} = -\left\lceil \frac{b}{c} \right\rceil,$$

whereas if b/c is an integer, we get a slightly loser asymptotic limit

$$\begin{aligned} -\frac{b}{c} - 1 &\leq \liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\phi(\bar{X}_n)\in[b,\infty)\times[0,c]\right)}{r(\log n)} \\ &\leq \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\phi(\bar{X}_n)\in[b,\infty)\times[0,c]\right)}{r(\log n)} \leq -\frac{b}{c}. \end{aligned}$$

4 Proofs

4.1 Proofs for Section 2

Proof of Lemma 2.1. Fix a $(u,r)\in\Gamma(\xi)$. Let $(v_1,s_1)\in\Pi(\xi)$ and $(v_2,s_2)\in\Pi(\zeta)$ be given arbitrarily. There exists a t_0 such that $v_1(t_0)=u$ and $s_1(t_0)=r$.

$$\begin{aligned} d\big((u,r),\,\Gamma(\zeta)\big) &\leq d\Big(\big(v_1(t_0),s_1(t_0)\big),\,\big(v_2(t_0),s_2(t_0)\big)\Big) \\ &= \big|v_1(t_0)-v_2(t_0)\big| \vee \big|s_1(t_0)-s_2(t_0)\big| \leq \|v_1-v_2\| \vee \|s_1-s_2\|. \end{aligned}$$

Since the choice of (v_1,s_1) and (v_2,s_2) was arbitrary, we arrive at the conclusion of the lemma by taking the infimum over $(v_1,s_1)\in\Pi(\xi)$ and $(v_2,s_2)\in\Pi(\zeta)$. ☐

Proof of Lemma 2.2. Since $(\zeta(t),t)$ and $(\zeta(s-),s)$ belong to $\Gamma(\zeta)$, it is enough, in view of Lemma 2.1, to show that

$$d\big((\zeta(t),t),\Gamma(\xi)\big) \vee d\big((\zeta(s-),s),\Gamma(\xi)\big) \geq \delta. \tag{4.1}$$

To see that this is the case, note first that either

$$|\zeta(t)-\xi(t)|\geq\delta \quad \text{or} \quad |\zeta(s-)-\xi(s)|\geq\delta.$$

Suppose that $|\zeta(t)-\xi(t)|\geq\delta$. Note also that for any $(u,r)\in\Gamma(\xi)$, either $u=\xi(t)$ or $|t-r|\geq\delta$. Then,

$$d\big((\zeta(t),t),(u,r)\big)=|\zeta(t)-u|\vee|t-r|\geq\delta.$$

Taking infimum over $(u,v)\in\Gamma(\xi)$, we get $d\big((\zeta(t),t),\Gamma(\xi)\big)>\delta$. Similarly,

$$d\big((\zeta(s-),s),\Gamma(\xi)\big)>\delta$$

if $|\zeta(s-)-\xi(s)|\geq\delta$. Therefore, we have (4.1).

□

Proof of Lemma 2.4. Since $\mathbb{E}$ is a closed set, it is clear that newly defined function I' is lower-semicontinuous on $\mathbb{D}$. For any open set G in $\mathbb{D}$, $G\cap\mathbb{E}$ is open in $\mathbb{E}$ and I take ∞ value on $G\cap\mathbb{E}^c$. Therefore,

$$\begin{aligned}\liminf_{n\to\infty}\frac{\log\mathbb{P}\left(X_n\in G\right)}{a_n}&=\liminf_{n\to\infty}\frac{\log\mathbb{P}\left(X_n\in G\cap\mathbb{E}\right)}{a_n}\\&\geq-\inf_{x\in G\cap\mathbb{E}}I(x)=-\inf_{x\in G}I'(x)\end{aligned}$$

Also, for any closed set F in $\mathbb{D}$, $(F\cap\mathbb{E})^\epsilon\subset F^\epsilon$, hence

$$\begin{aligned}\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(X_n\in F\right)}{a_n}&=\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(X_n\in F\cap\mathbb{E}\right)}{a_n}\\&\leq-\lim_{\epsilon\downarrow0}\inf_{x\in(F\cap\mathbb{E})^\epsilon}I(x)\leq-\lim_{\epsilon\downarrow0}\inf_{x\in F^\epsilon}I'(x)\end{aligned}$$

Since the upper and lower bounds for the extended LDP are all satisfied with the lower-semicontinuous function I', the lemma is proved. □

4.2 Proofs for Section 3.1

This section proves Theorem 3.4 and Theorem 3.5. Before describing the structures of the proofs, we introduce some notations. Rearranging the terms in (3.1), we can decompose $\bar{X}_n$ as follows:

$$\bar{X}_n=\bar{Y}_n+\bar{R}_n \tag{4.2}$$

where

$$\bar{Y}_n(t)\triangleq\frac{1}{n}\int_{[1,\infty)}(x-\mu_1)\hat{N}([0,nt]\times dx), \tag{4.3}$$

$$\begin{aligned}\bar{R}_n(t)\triangleq\frac{aB(nt)}{n}&+\frac{1}{n}\int_{|x|<1}x\big(\hat{N}([0,nt]\times dx)-nt\nu(dx)\big)+\frac{\mu_1\hat{N}\big([0,nt]\times[1,\infty)\big)}{n}-t\nu_1\mu_1\\&+\frac{1}{n}\int_{(-\infty,1]}x\hat{N}([0,nt]\times dx)-t\int_{(-\infty,1]}x\nu(dx)\end{aligned} \tag{4.4}$$

for $t\in[0,1]$, where $\nu_1=\nu[1,\infty)$ and $\mu_1=\int_{[1,\infty)}x\frac{\nu(dx)}{\nu_1}$.

Since $\hat{N}$ has a finite mean measure on $[1,\infty)\times[0,n]$, the process $\bar{Y}_n$ is a compound Poisson process. Before being centerd by μ_1 and scaled by $1/n$, the jump sizes of $\bar{Y}_n$ are independent and identically distributed (i.i.d.) random variables. Their distribution is the restriction of ν on $[1,\infty)$, normalized by ν_1. Let $Z_1,Z_2,\cdots$ be i.i.d. random variables following this distribution and let $N(t) \triangleq \hat{N}([0,t] \times [1,\infty))$. Then, $\bar{Y}_n$ has the following distributional representation:

$$\bar{Y}_n(t) \stackrel{\mathcal{D}}{=} \frac{1}{n} \sum_{i=1}^{N(nt)} (Z_i - \mu_1). \tag{4.5}$$

We further decompose (4.5) into two parts: the k largest jumps and the rest. Let $P_n(\cdot)$ be the random permutation of the indices $\{1, 2, \cdots, N(n)\}$ of Z_i's such that $P_n(i)$ is the rank of Z_i in the decreasing order among $Z_1, Z_2, \cdots, Z_{N(n)}$. Then

$$\bar{Y}_n(t) \stackrel{\mathcal{D}}{=} \underbrace{\frac{1}{n} \sum_{i=1}^{N(nt)} Z_i \mathbb{1}\{P_n(i) \le k\}}_{\bar{J}_n^{\leqslant k}(t)} + \underbrace{\frac{1}{n} \sum_{i=1}^{N(nt)} \big(Z_i \mathbb{1}\{P_n(i) > k\} - \mu_1\big)}_{\bar{H}_n^{\leqslant k}(t)} \tag{4.6}$$

Note that $\bar{J}_n^{\leqslant k}$ is the process that consists of the k largest jumps in $\bar{Y}_n$, whereas $\bar{H}_n^{\leqslant k}$ is the centered process that consists of the remaining jumps. Let E_i's be i.i.d. exponential random variables with mean 1, and $\Gamma_i \triangleq E_1 + \cdots + E_i$. Let U_i's be i.i.d. uniform random variables on $[0,1]$, independent from the Y_i's. Let

$$Q_n(\cdot) \triangleq n\nu[\cdot, \infty) \qquad \text{and} \qquad Q_n^{\leftarrow}(y) \triangleq \inf\{s > 0 : n\nu[s,\infty) < y\}. \tag{4.7}$$

According to the argument presented in [19] (p.305) and [26] (p.163), a coupling with respect to Y_i, U_i and $\hat{N}$ can be constructed by

$$\hat{N}([0,n] \times (0,\infty)) = \sum_{i=1}^{\infty} \epsilon_{(n \cdot U_i,\ Q_n^{\leftarrow}(\Gamma_i))} \tag{4.8}$$

Here $\epsilon_{(x,y)}$ denotes the Dirac measure at (x,y). The sequence $\{Q_n^{\leftarrow}(\Gamma_i), i \ge 1\}$ record the second coordinate of $\hat{N}$'s point masses in a decreasing order. The coupling in (4.8) facilitates a further decomposition of $\bar{J}_n^{\leqslant k}$ into $\hat{J}_n^{\leqslant k}$ + $\check{J}_n^{\leqslant k}$,where

$$\hat{J}_n^{\leqslant k}(t) \triangleq \frac{1}{n} \sum_{i=1}^{k} Q_n^{\leftarrow}(\Gamma_i) \mathbb{1}_{[U_i,1]}(t), \tag{4.9}$$

$$\check{J}_n^{\leqslant k}(t) \triangleq -\frac{1}{n} \mathbb{1}\{\tilde{N}_n < k\} \sum_{i=\tilde{N}_n+1}^{k} Q_n^{\leftarrow}(\Gamma_i) \mathbb{1}_{[U_i,1]}(t). \tag{4.10}$$

In (4.10), $\tilde{N}_n = \sum_{i=1}^{\infty} \mathbb{I}_{[0,n]\times[1,\infty)}\big((n \cdot U_i, Q_n^{\leftarrow}(\Gamma_i))\big)$ counts the number of point masses of $\hat{N}$ on the set $[0,n] \times [1,\infty)$. Roughly speaking, $\hat{J}_n^{\leqslant k}$ consists of the k largest jumps in $\bar{X}_n$, but if some of those jumps are of size smaller than 1, their offsets are recorded in $\check{J}_n^{\leqslant k}$.

We arrive at the distributional representation

$$\bar{X}_n \stackrel{\mathcal{D}}{=} \hat{J}_n^{\leqslant k} + \check{J}_n^{\leqslant k} + \bar{H}_n^{\leqslant k} + \bar{R}_n, \tag{4.11}$$

where $\hat{J}_n^{\leqslant k}$, $\check{J}_n^{\leqslant k}$, $\bar{H}_n^{\leqslant k}$, and $\bar{R}_n$ are defined in (4.9), (4.10), (4.6) and (4.4), respectively. This representation provides a clear road map for proving the main results of this subsection: Proposition 4.5 proves the extended LDP satisfied by the jump sizes of $\hat{J}_n^{\leqslant k}$. Based on Proposition 4.5, the extended LDP of $\hat{J}_n^{\leqslant k}$ is established in Proposition 4.1. It turns out that, for large k, the asymptotics of $\bar{X}_n$ is predominantly determined by $\hat{J}_n^{\leqslant k}$. This along with Proposition 2.5, we obtain the extended LDP of $\bar{X}_n$ in Theorem 3.4. To facilitate the verification the conditions of Proposition 2.5, we establish a bound for $\bar{H}_n^{\leqslant k}$ in Lemma 4.2. That is, the proof of Theorem 3.4 hinges on Proposition 4.1 and Lemma 4.2 below. Their proofs are provided in Section 4.2.1 and Section 4.2.2, respectively.

Proposition 4.1. *$\{\hat{J}_n^{\leqslant k}\}_{n\geq 1}$ satisfies the extended LDP on $(\mathbb{D}, d_{J_1})$ with the rate function*

$$\hat{I}_k(\xi) \triangleq \begin{cases} \sum_{t\in(0,1]} \mathbb{1}\{\xi(t) \neq \xi(t-)\} & \text{if } \xi \in \mathbb{D}_{\leqslant k} \\ \infty & \text{otherwise} \end{cases} \tag{4.12}$$

and speed $r(\log n)$.

Lemma 4.2. *Recall $\bar{H}_n^{\leqslant k}$ defined in (4.6). We conclude*

$$\lim_{k\to\infty} \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\|\bar{H}_n^{\leqslant k}\| > \epsilon\right)}{r(\log n)} = -\infty$$

With these in hand, we are ready to prove Theorem 3.4.

Proof of Theorem 3.4. We apply Proposition 2.5, with $\bar{X}_n$ and $\hat{J}_n^{\leqslant k}$ being X_n and Y_n^k, respectively, in the Proposition. To apply Proposition 2.5, we need to verify the following four conditions:

(1) For each $k \in \mathbb{N}$, the sequence $\{\hat{J}_n^{\leqslant k}\}_{n\geq 1}$ satisfies an extended LDP with rate $\hat{I}_k$ and speed $\{r(\log n)\}_{n\geq 1}$.

(2) For any closed set F,
$$\lim_{k\to\infty} \inf_{\xi\in F} \hat{I}_k(\xi) \geq \inf_{\xi\in F} I^{J_1}(\xi). \tag{4.13}$$

(3) For $\forall\delta > 0$ and any open set G, there exists $\epsilon > 0$ and $K \geq 0$ such that when $k \geq K$
$$\inf_{\xi\in G^{-\epsilon}} \hat{I}_k(\xi) \leq \inf_{\xi\in G} I^{J_1}(\xi) + \delta. \tag{4.14}$$

(4) For $\forall\epsilon > 0$,
$$\lim_{k\to\infty} \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(d_{J_1}(\bar{X}_n, \hat{J}_n^{\leqslant k}) > \epsilon\right)}{\log^\gamma n} = -\infty. \tag{4.15}$$

Condition (1) holds due to Proposition 4.1. Condition (2) is direct as $\hat{I}_k \geq I^{J_1}$ for any $k \in \mathbb{N}$. For the condition (3), we verify a stronger property that implies (4.14): for $\forall\delta > 0$ and any open set G, there exists $\epsilon > 0$ and $K \geq 0$ such that when $k \geq K$,

$$\inf_{\xi\in G^{-\epsilon}} \hat{I}_k(\xi) = \inf_{\xi\in G} I^{J_1}(\xi). \tag{4.16}$$

To verify (4.16), note first that the equality is obvious if $\inf_{\xi\in G} I^{J_1}(\xi) = \infty$ because $\hat{I}_k \geq I^{J_1}$. On the other hand, if $\inf_{\xi\in G} I^{J_1}(\xi) = m < \infty$, the following inequalities hold for any $\epsilon > 0$

$$\inf_{\xi\in G^{-\epsilon}} \hat{I}_k(\xi) \geq \inf_{\xi\in G} \hat{I}_k(\xi) \geq \inf_{\xi\in G} I^{J_1}(\xi) = m$$

Additionally, $\inf_{\xi\in G} I^{J_1}(\xi) = m$ suggests the existence of some $\eta \in G \cap \mathbb{D}_{=m}$. Given that G is open, there exists $\epsilon > 0$ such that $B_\epsilon(\eta) \subset G$, implying $\eta \in G^{-\epsilon}$. Therefore, for $k > m$, we have

$$\inf_{\xi\in G^{-\epsilon}} \hat{I}_k(x) \leq \hat{I}_k(\eta) = m$$

Therefore, $\inf_{\xi\in G^{-\epsilon}} \hat{I}_k(\xi) = \inf_{\xi\in G} I^{J_1}(\xi)$ for $k > m$.

We now move onto the condition (4). From the representation of $\bar{X}_n$ in (4.11), we have

$$\begin{aligned}\mathbb{P}\left(d_{J_1}(\tilde{X}_n, \hat{J}_n^{\leqslant k}) > \epsilon\right) &\leq \mathbb{P}\left(\|\tilde{X}_n - \hat{J}_n^{\leqslant k}\| > \epsilon\right) \\ &\leq \mathbb{P}\left(\|\bar{H}_n^{\leqslant k}\| > \frac{\epsilon}{3}\right) + \mathbb{P}\left(\|\check{J}_n^{\leqslant k}\| > \frac{\epsilon}{3}\right) + \mathbb{P}\left(\|\bar{R}_n\| > \frac{\epsilon}{3}\right).\end{aligned}$$

Therefore,

$$\begin{aligned}
&\lim_{k\to\infty}\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(d_{J_1}(\bar{Y}_n,\hat{J}_n^{\leqslant k})>\epsilon\right)}{r(\log n)}\\
&\le \max\Bigg\{\underbrace{\lim_{k\to\infty}\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\|\bar{H}_n^{\leqslant k}\|>\frac{\epsilon}{3}\right)}{r(\log n)}}_{\text{(I)}},\underbrace{\lim_{k\to\infty}\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\|\check{J}_n^{\leqslant k}\|>\frac{\epsilon}{3}\right)}{r(\log n)}}_{\text{(II)}},\\
&\qquad\underbrace{\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\|\bar{R}_n\|>\frac{\epsilon}{3}\right)}{r(\log n)}}_{\text{(III)}}\Bigg\}.
\end{aligned}\tag{4.17}$$

Among the three terms in (4.17), (I) $=-\infty$ by Lemma 4.2. For (II), by the definition of $\check{J}_n^{\leqslant k}$ in (4.10), we have

$$\text{(II)}=\mathbb{P}\left(\sup_{t\in[0,1]}\left|\frac{1}{n}\mathbb{1}\{\tilde{N}_n<k\}\sum_{i=\tilde{N}_n+1}^{k}Q_n^{\leftarrow}(\Gamma_i)\mathbb{1}_{[U_i,1]}(t)\right|>\frac{\epsilon}{3}\right)\le\mathbb{P}\left(\tilde{N}_n<k\right).$$

Since $\tilde{N}_n$ can be treated as n independent summation of Poisson random variables with mean $\nu[1,\infty)$, Cramér's theorem implies that $\mathbb{P}\left(\tilde{N}_n<k\right)\sim e^{-n\cdot C}$ for some constant $C>0$. This implies (II) $=-\infty$. For (III), $\bar{R}_n$ is a Lévy process such that $\bar{R}_n(1)$ has a finite moment generating function. Theorem 2.5 of [22] confirms that the numerator of (III) increases at a linear rate, and hence, (III) $=-\infty$. Hence condition (4) is satisfied and the proof is finished. □

Now we move on to the proof of Theorem 3.5. Before providing the proof, we establish the following lemma.

Lemma 4.3. *For any $j\in\mathbb{N}$, $\hat{\mathbb{D}}_{\leqslant j}$ is closed w.r.t the M_1' topology.*

Proof of Lemma 4.3. Note that $\mathbb{D}\backslash\hat{\mathbb{D}}_{\leqslant j}=A\cup B\cup C\cup D$ where

$$\begin{aligned}
A &\triangleq \{\eta\in\mathbb{D}:\eta(0)<0\}\\
B &\triangleq \{\eta\in\mathbb{D}:\eta(0)\ge0,\ \eta\text{ is not a non-decreasing function}\}\\
C &\triangleq \{\eta\in\mathbb{D}:\eta(0)\ge0,\ \eta\text{ is non-decreasing, but not a pure jump function}\}\\
D &\triangleq \{\eta\in\mathbb{D}:\eta(0)\ge0,\ \eta\text{ is a non-decreasing pure jump function with}\\
&\qquad\qquad\text{more than } j \text{ jumps}\}.
\end{aligned}$$

We argue that any given $\eta\notin\hat{\mathbb{D}}_{\leqslant j}$ is bounded away from $\hat{\mathbb{D}}_{\leqslant j}$ by considering η's in A, B, C, and D separately.

Suppose that $\eta\in A$. For any $\xi\in\hat{\mathbb{D}}_{\leqslant j}$, $(u,v)\in\Gamma(\xi)$ implies that $u\ge0$, and hence, $d\big((\eta(0),0),\Gamma(\xi)\big)\ge|\eta(0)|$. This along with Lemma 2.1 implies that $d_{M_1'}(\eta,\xi)>|\eta(0)|>0$. Since ξ was arbitrarily chosen in $\hat{\mathbb{D}}_{\leqslant j}$, we conclude that $d_{M_1'}(\eta,\hat{\mathbb{D}}_{\leqslant j})>|\eta(0)|>0$.

Suppose that $\eta\in B$. Then, there exists $\delta>0$ and $t_1,t_2\in[0,1]$ such that $t_2-t_1>4\delta$ and $\eta(t_1)-\eta(t_2)>4\delta$. We claim that for any $\xi\in\hat{\mathbb{D}}_{\leqslant j}$, $d_{M_1'}(\eta,\xi)>\delta$, and hence, $d_{M_1'}(\eta,\hat{\mathbb{D}}_{\leqslant j})>\delta$. To see why, suppose not. That is, suppose that there exists $\xi\in\hat{\mathbb{D}}_{\leqslant j}$ such that $d_{M_1'}(\eta,\xi)\le\delta$. Then, due to Lemma 2.1, there are $(u_1,s_1),(u_2,s_2)\in\Gamma(\xi)$ such that $d\big((\eta(t_1),t_1),(u_1,s_1)\big)\le2\delta$ where $u_1\in[\xi(s_1-),\,\xi(s_1)]$, and $d\big((\eta(t_2),t_2),(u_2,s_2)\big)\le2\delta$ where $u_2\in[\xi(s_2-),\,\xi(s_2)]$. Note that these imply

$$s_1\le t_1-2\delta<t_2+2\delta\le s_2$$

and

$$\xi(s_1) \geq u_1 \geq \eta(t_1) - 2\delta \geq \eta(t_2) + 2\delta \geq u_2 \geq \xi(s_2-),$$

which is contradictory to ξ being non-decreasing. This proves the claim.

Suppose that $\eta \in C$. Then there exists an interval within $[0,1]$ on which η is continuous and strictly increasing. By subdividing the increment over this interval into the ones with small enough increments, one can find a sufficiently small $\delta > 0$ and more than j non-overlapping subintervals of the form $[s-\delta, t+\delta]$ such that $t - s > 2\delta$ and $\eta(t) - \eta(s) > 2\delta$. For any $\xi \in \hat{\mathbb{D}}_{\leqslant j}$, it needs to be constant on at least one of these intervals. From Lemma 2.2, we see that $d_{M_1'}(\eta, \xi) \geq \delta$. Again, since this is for an arbitrary $\xi \in \hat{\mathbb{D}}_{\leqslant j}$, we conclude that $d_{M_1'}(\eta, \hat{\mathbb{D}}_{\leqslant j}) \geq \delta$.

Suppose that $\eta \in D$. Then $\eta = \sum_{i=1}^{k} z_i \mathbb{1}_{[t_1,1]}$ some $k > j$ and $0 \leq t_1 < t_2 < \cdots < t_k \leq 1$. Set $T_i \triangleq [t_i - \delta, t_i + \delta] \cap [0,1]$ and pick $\delta > 0$ small enough so that T_i's are disjoint and $2\delta \leq \min_{i=1,\cdots,k} \big(\eta(t_i) - \eta(t_i-)\big)$. Then, any path $\xi \in \hat{\mathbb{D}}_{\leqslant j}$ is constant on at least one of T_i's, and η jumps at t_i with jump size no smaller than 2δ. From Lemma 2.2, we have that $d_{M_1'}(\xi, \eta) \geq \delta/2$. Since ξ is arbitrary in path in $\hat{\mathbb{D}}_{\leqslant j}$, we conclude that $d_{M_1'}\big(\eta, \hat{\mathbb{D}}_{\leqslant j}\big) > \delta$. $\square$

Proof of Theorem 3.5. Note that $I^{M_1'}$ is obviously non-negative, and the sublevel sets $\Psi_{I^{M_1'}}(\alpha) = \hat{\mathbb{D}}_{\leqslant \lfloor \alpha \rfloor}$ are closed for any $\alpha \geq 0$ due to Lemma 4.3. Therefore, $I^{M_1'}$ is a rate function.

Consider a closed set F and an open set G under the M_1' topology. They are also closed and open under the J_1 topology, respectively. Due to Theorem 3.4, we have:

$$-\inf_{x \in G} I^{J_1}(x) \leq \liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{X}_n \in G\right)}{r(\log n)} \text{ and } \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{X}_n \in F\right)}{r(\log n)} \leq -\inf_{x \in F^{\epsilon,J_1}} I^{J_1}(x) \tag{4.18}$$

In the above inequalities, we use an extra superscript J_1 in F^{ϵ,J_1} to clarify that it is the ϵ-fattening w.r.t. J_1 topology. Likewise, we will denote the ϵ-fattening w.r.t. M_1' topology with $F^{\epsilon,M_1'}$.

Note that the M_1' metric is bounded by the J_1 metric, and hence, $F^{\epsilon,J_1} \subset F^{\epsilon,M_1'}$. Moreover, $I^{M_1'} \leq I^{J_1}$ immediately from their definitions. Hence,

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{X}_n \in F\right)}{r(\log n)} \leq -\inf_{x \in F^{\epsilon,J_1}} I^{J_1}(x) \leq -\inf_{x \in F^{\epsilon,M_1'}} I^{M_1'}(x),$$

which proves the upper bound of the extended LDP on $(\mathbb{D}, M_1')$.

We claim that for any open set G with respect to the M_1' topology, the following holds:

$$-\inf_{x\in G} I^{M_1'}(x) = -\inf_{x \in G} I^{J_1}(x). \tag{4.19}$$

The inequality $-\inf_{x\in G} I^{M_1'}(x) \geq -\inf_{x\in G} I^{J_1}(x)$ is trivial as $I^{M_1'} \leq I^{J_1}$. To show $-\inf_{x\in G} I^{M_1'}(x) \leq -\inf_{x\in G} I^{J_1}(x)$, note first that the inequality holds trivially if

$$-\inf_{x\in G} I^{M_1'}(x) = -\infty.$$

On the other hand, if $-\inf_{x\in G} I^{M_1'}(x) = -m > -\infty$ for some integer m, G contains some $\xi \in \hat{\mathbb{D}}_{=m}$. There are two possibilities:

- If no jumps of ξ occur at 0 or 1, $\xi \in \mathbb{D}_{=m}$ and

$$-\inf_{x\in G} I^{J_1}(x) \geq -I^{J_1}(\xi) = -m = -\inf_{x\in G} I^{M_1'}(x)$$

- If jumps of ξ occur at 0 or 1 or both, we can construct a new path $\xi' \in \mathbb{D}_{=m}$ by preserving the jump sizes of ξ while perturbing its jump times from 0 to δ and from 1 to $1-\delta$. Since $d_{M_1'}(\xi, \xi') \leq \delta$ and G is an open, we can make $\xi' \in G$ by choosing δ sufficiently small. This implies
$$-\inf_{x \in G} I^{J_1}(x) \geq -I^{J_1}(\xi') = -m = -\inf_{x \in G} I^{M_1'}(x)$$

Therefore, we arrive at (4.19). This along with the extended LDP w.r.t. the J_1 topology,

$$-\inf_{x \in G} I^{M_1'}(x) = -\inf_{x \in G} I^{J_1}(x) \leq \liminf_{n \to \infty} \frac{\log \mathbb{P}\left(\bar{X}_n \in F\right)}{r(\log n)}.$$

This proves the lower bound for the extended LDP on $(\mathbb{D}, M_1')$ and finishes the proof. $\square$

4.2.1 Proof of Proposition 4.1

Proposition 4.1 hinges on the extended LDP for $\left(Q_n^{\leftarrow}(\Gamma_1)/n, \cdots, Q_n^{\leftarrow}(\Gamma_k)/n\right)$, which is established in Proposition 4.5. The following lemma is useful in the proof of Proposition 4.5.

Lemma 4.4. *Suppose that* $x = (x_1, \ldots, x_k) \in \mathbb{R}_+^{k\downarrow}$ *and and* $\iota = \check{I}_k(x)$. *Let*

$$E(x;\delta) \triangleq \{y \in \mathbb{R}_+^k : y_i \in E_j(x, y; \delta)\}$$

where

$$E_j(x, y; \delta) = \begin{cases} \left(Q_n(n(x_1 + \delta), \, Q_n(n(x_1 - \delta)\right) & \text{for } j = 1; \\ \left(Q_n(n(x_j + \delta) - \sum_{i=1}^{j-1} y_i, \, Q_n(n(x_j - \delta) - \sum_{i=1}^{j-1} y_i\right) & \text{for } j = 2, \ldots, \iota; \\ \left(Q_n(n\delta) - Q_n(n(x_\iota + \delta)), \, \infty\right) & \text{for } j = \iota + 1; \\ \left[0, \infty\right) & \text{for } j > \iota + 1. \end{cases}$$

Then $y \in E(x;\delta)$ *implies that*

$$\left(\frac{Q_n^{\leftarrow}(y_1)}{n}, \cdots, \frac{Q_n^{\leftarrow}(y_1 + \cdots + y_k)}{n}\right) \in \prod_{i=1}^{\iota} [x_i - \delta, x_i + \delta] \times [0, \delta]^{(k-\iota)} \triangleq D(x;\delta).$$

Proof of Lemma 4.4. Note that by the construction of $E_j(x, y; \delta)$, it is straightforward to check by induction that if $y_i \in E_i(x, y; \delta)$ for $i = 1, \ldots, j$, then

$$\sum_{i=1}^{j} y_j \in \left(Q_n(n(x_j + \delta)), Q_n(n(x_j - \delta))\right) \tag{4.20}$$

for $j \leq \iota$, and

$$\sum_{i=1}^{j} y_j \in [Q_n(n\delta), \infty) \tag{4.21}$$

for $j > \iota$. The conclusion of the lemma follows from the fact that $c \in \left(Q_n(b), Q_n(a)\right]$ if and only if $Q_n^{\leftarrow}(c) \in [a, b)$ for any c and $b > a \geq 0$. $\square$

Consider $\mathbb{R}_+^{k\downarrow} \triangleq \{(x_1, \cdots, x_k) \in \mathbb{R}^k : x_1 \geq x_2 \geq \cdots \geq x_k \geq 0\}$. For $n \geq 1$, define the follwoing measure on $\mathbb{R}_+^{k\downarrow}$:

$$\check{\mu}_n^k(\cdot) \triangleq \mathbb{P}\left(\left(\frac{Q_n^{\leftarrow}(\Gamma_1)}{n}, \cdots, \frac{Q_n^{\leftarrow}(\Gamma_k)}{n}\right) \in \cdot\right). \tag{4.22}$$

In the above definition, $\Gamma_i = E_1 + \cdots + E_i$, and E_i's are iid $\mathrm{Exp}(1)$. Therefore, Lemma 4.4 implies

$$\int_{D(x;\delta)} \check{\mu}_n(dx) > \int_{E(x;\delta)} e^{-\sum_{i=1}^k y_i} dy_1 \cdots dy_k. \tag{4.23}$$

Proposition 4.5. *$\{\check{\mu}_n^k\}_{n\geq 1}$ satisfies the extended LDP on $\mathbb{R}_+^{k\downarrow}$ with the rate function*

$$\check{I}_k(\boldsymbol{x}) \triangleq \sum_{i=1}^{k} \mathbb{1}\{x_i \neq 0\} \text{ for } \boldsymbol{x} = (x_1, \cdots, x_k) \in \mathbb{R}_+^{k\downarrow} \tag{4.24}$$

and the speed $r(\log n)$.

Proof of Proposition 4.5. We begin with showing $\check{I}_k(\cdot)$ is lower semi-continuous. Consider an $\alpha < k$ (since the case $\alpha \geq k$ is trivial) and the corresponding sublevel set $\Psi_{\check{I}_k}(\alpha) \triangleq \{x \in \mathbb{R}_+^k : \check{I}_k(x) \leq \alpha\}$. Note that $\Psi_{\check{I}_k}(\alpha) = \{x \in \mathbb{R}_+^k :\ x_{\lfloor\alpha\rfloor+1} = 0\}$. This is a closed set, because for any y in $\Psi_{\check{I}_k}(\alpha)^c$, we have $y_{\lfloor\alpha\rfloor+1} > 0$, and hence, $B(y;\delta)$ with $\delta = y_{\lfloor\alpha\rfloor+1}/2$ is a neighborhood of y within $\Psi_{\check{I}_k}(\alpha)^c$. This proves the lower semi-continuity of $\check{I}_k(\cdot)$.

We then verify the lower bound of the extended LDP is satisfied by $\{\check{\mu}_n^k\}_{n\geq 1}$, i.e. for any open set $G \subset \mathbb{R}_+^{k\downarrow}$,

$$-\inf_{x\in G} \check{I}_k(x) \leq \liminf_{n\to\infty} \frac{\log \check{\mu}_n(G)}{r(\log n)}. \tag{4.25}$$

For any given open set G, fix an arbitrary $x \in G$ and let $\iota = \check{I}_k(x)$. We first observe that for any given $x \in G$, we can find $x' \in G$ such that $\check{I}_k(x') = \iota = \check{I}_k(x)$ and $x_1' > x_2' > \cdots > x_\iota' > 0$. Therefore, for the purpose of establishing (4.25), we can assume w.l.o.g. that

$$x_1 > \ldots > x_\iota > x_{\iota+1} = 0. \tag{4.26}$$

Note that we can find a continuous function $f : \mathbb{R}_+^{k\downarrow} \to [0,1]$ such that $f(x) = 1$ and $f(y) = 0$ for $y \in G^c$ since $\mathbb{R}_+^{k\downarrow}$ is a completely regular topological space. Further, define $f_m(\cdot) = m(f(\cdot) - 1)$. Since f_m is $-m$ on G^c and at most 0 elsewhere,

$$\int_{\mathbb{R}_+^{k\downarrow}} e^{r(\log n) f_m(s)} \check{\mu}_n(ds) \leq e^{-mr(\log n)} \check{\mu}_n(G^c) + \check{\mu}_n(G),$$

and hence,

$$\begin{aligned} \liminf_{n\to\infty} \frac{\log \int_{\mathbb{R}_+^{k\downarrow}} e^{r(\log n) f_m(s)} \check{\mu}_n(ds)}{r(\log n)} &\leq \liminf_{n\to\infty} \frac{\log\left(e^{-mr(\log n)} + \check{\mu}_n(G)\right)}{r(\log n)} \\ &= \max\left\{ -m,\ \liminf_{n\to\infty} \frac{\log \check{\mu}_n(G)}{r(\log n)} \right\}. \end{aligned} \tag{4.27}$$

On the other hand, we claim that the following holds:

$$-\check{I}_k(x) \leq \liminf_{n\to\infty} \frac{\log \int_{\mathbb{R}_+^{k\downarrow}} e^{r(\log n) f_m(s)} \check{\mu}_n(ds)}{r(\log n)}. \tag{4.28}$$

Combining (4.27) and (4.28), then taking the limit $m \to \infty$ and the infimum over $x \in G$, we arrive at the conclusion of the proposition.

Now we are left with the proof of the claim (4.28). Recall the neighborhood $D(x;\delta) = \prod_{i=1}^{\iota}[x_i - \delta, x_i + \delta] \times [0,\delta]^{(k-\iota)}$ of x defined in Lemma 4.4. Pick an arbitrary $\epsilon > 0$. Since G is open and f_m is continuous, assumption (4.26) allows us to choose a small enough $\delta = \delta(x,\epsilon)$ so that

(i) $[x_i - \delta, x_i + \delta] \cap [x_j - \delta, x_j + \delta] = \emptyset$ for any i and j such that $i < j \leq \iota$;

(ii) $f_m(y) > -\epsilon$ for $y \in D(x;\delta)$.

From (ii) and (4.23)

$$\liminf_{n\to\infty}\frac{\log\int_{\mathbb{R}_+^{k\downarrow}}e^{r(\log n)f_m(x)}\check{\mu}_n(dx)}{r(\log n)}\geq\liminf_{n\to\infty}\frac{\log\left(e^{r(\log n)(-\epsilon)}\int_{D(x;\delta)}\check{\mu}_n(dx)\right)}{r(\log n)}$$
$$\geq-\epsilon+\liminf_{n\to\infty}\frac{\log\int_{E(x;\delta)}e^{-\sum_{i=1}^k y_i}dy_1\cdots dy_k}{r(\log n)}\qquad(4.29)$$

From the forms of $E(x;\delta)$ and $E_i(x,y;\delta)$'s in Lemma 4.4, we see that the integral in (4.29) can be decomposed as follows:

$$\begin{aligned}&\log\int_{E(x;\delta)}e^{-\sum_{i=1}^k y_i}dy_1\cdots dy_k\\&=\log\int_{E'(x;\delta)}e^{-\sum_{i=1}^\iota y_i}dy_1\cdots dy_\iota+\log\int_{Q_n(n\delta)-Q_n(n(x_\iota+\delta))}^\infty e^{-x}dx+\sum_{i=\iota+2}^k\cdot\log\int_0^\infty e^{-x}dx\\&=\underbrace{\log\int_{E'(x;\delta)}e^{-\sum_{i=1}^\iota y_i}dy_1\cdots dy_\iota}_{\text{(I)}}-\big(Q_n(n\delta)-Q_n(n(x_\iota+\delta))\big)\end{aligned}\qquad(4.30)$$

where $E'(x;\delta)=\{y\in\mathbb{R}_+^\iota: y_i\in E_j(x,y;\delta)\text{ for }j=1,\dots,\iota\}$. To bound (I), note first that the integrand is bounded from below due to (4.20):

$$e^{-\sum_{i=1}^\iota y_i}\geq e^{-Q_n(n(x_\iota-\delta))}\quad\text{on }E'(x;\delta)$$

and the length of the domain of integral in i^{th}coordinate is $Q_n(n(x_j-\delta))-Q_n(n(x_j+\delta))$ for each $i\leq\iota$, and hence,

$$\int_{E'(x;\delta)}dy_\iota\cdots dy_1=\prod_{i=1}^\iota\big(Q_n(n(x_j-\delta))-Q_n(n(x_j+\delta))\big).$$

Therefore,

$$\begin{aligned}(I)=\int_{E'(x;\delta)}e^{-\sum_{i=1}^\iota y_i}dy_\iota\cdots dy_1&\geq e^{-Q_n(n(x_\iota-\delta))}\int_{E'(x;\delta)}dy_\iota\cdots dy_1\\&=e^{-Q_n(n(x_\iota-\delta))}\cdot\prod_{i=1}^\iota\big(Q_n(n(x_i-\delta))-Q_n(n(x_i+\delta))\big).\end{aligned}\qquad(4.31)$$

From (4.29), (4.30), and (4.31),

$$\begin{aligned}&\liminf_{n\to\infty}\frac{\log\int_{\mathbb{R}_+^{k\downarrow}}e^{r(\log n)f_m(x)}\check{\mu}_n(dx)}{r(\log n)}\\&\geq-\epsilon+\lim_{n\to\infty}\frac{-\big(Q_n(n\delta)-Q_n(n(x_\iota+\delta))\big)}{r(\log n)}\\&+\lim_{n\to\infty}\frac{-Q_n(n(x_\iota-\delta))}{r(\log n)}+\sum_{i=1}^\iota\lim_{n\to\infty}\frac{\log\big(Q_n(n(x_i-\delta))-Q_n(n(x_i+\delta))\big)}{r(\log n)}\\&=-\epsilon-\iota,\end{aligned}$$

The last equality above can be derived by plugging in the form $Q_n(x)=n\nu[x,\infty)=n\exp\{-r(\log x)\}$ and leverage and the limit results (4.89) and (4.88). Taking $\epsilon\to0$, we arrive at (4.28), and this concludes the proof of (4.25).

Finally, we turn to show the upper bound of the extended LDP, i.e for any closed set $F \subset \mathbb{R}_+^{k\downarrow}$,

$$\limsup_{n\to\infty} \frac{\log \check{\mu}_n(F)}{r(\log n)} \leq -\lim_{\epsilon \downarrow 0} \inf_{x\in F^\epsilon} \check{I}_k(x). \tag{4.32}$$

For any given closed set F, let $\iota \triangleq \inf_{x\in F^\epsilon} \check{I}_k(x)$. Then, it is straightforward to see that there exists $r > 0$ such that $x_\iota > r$ for all $x \in F$. Therefore,

$$\begin{aligned}
\limsup_{n\to\infty} \frac{\log \check{\mu}_n(F)}{r(\log n)} &= \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\left(\frac{Q_n^{\leftarrow}(\Gamma_1)}{n}, \frac{Q_n^{\leftarrow}(\Gamma_2)}{n}, \cdots \frac{Q_n^{\leftarrow}(\Gamma_k)}{n}\right) \in F\right)}{r(\log n)} \\
&\leq \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\{Q_n^{\leftarrow}(\Gamma_\iota) \geq nr\}\right)}{r(\log n)} = \lim_{n\to\infty} \frac{\log \mathbb{P}\left(\{\Gamma_\iota \leq Q_n(nr)\}\right)}{r(\log n)} \\
&= \lim_{n\to\infty} \frac{\log \mathbb{P}\left(\{\Gamma_\iota \leq n \exp\{-r(\log nr)\}\}\right)}{r(\log n)} = -\iota,
\end{aligned}$$

where the last equality is from (4.92). This finishes the proof of the Proposition. □

We are now ready to prove Proposition 4.1.

Proof of Proposition 4.1. We first prove the extended LDP on $(\mathbb{D}_{\leqslant k}, d_{J_1})$, then lift it to the larger space $(\mathbb{D}, d_{J_1})$ via Lemma 2.4. Note that the conditions in Lemma 2.4 are satisfied as $\mathbb{P}\big(\hat{J}_n^{\leqslant k} \in \mathbb{D}_{\leqslant k}\big) = 1$, and $\mathbb{D}_{\leqslant k}$ is closed.

To see that $\hat{I}_k$ is a legitimate rate function, note that $\Psi_{\hat{I}_k}(c) = \mathbb{D}_{\leqslant \lfloor c \rfloor \wedge k}$ for any $c \in \mathbb{R}_+$, which is closed, and hence, $\hat{I}_k$ is lower-semicontinuous. To verify the lower bound, we show that

$$\liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\hat{J}_n^{\leqslant k} \in G\right)}{r(\log n)} \geq -\hat{I}_k(\xi) \tag{4.33}$$

for any open set $G \subset \mathbb{D}_{\leqslant k}$ and any $\xi \in G$. Fix a $\xi \in G$ and set $j \triangleq \hat{I}_k(\xi) \leq k$. Then $\xi = \sum_{i=1}^{j} x_i \mathbb{1}_{[u_i,1]}$, where $x_1 \geq \cdots \geq x_j > 0$ and $u_i \in (0,1)$ for $i = 1,\dots,j$. We construct a neighborhood of ξ in $\mathbb{D}_{\leqslant k}$ by perturbing ξ's jump sizes and times. Specifically, for some $w^* \in (0,1)$ that does not align with ξ's jump time, consider

$$C_\delta := \{\sum_{i=1}^{k} y_i \mathbb{1}_{[w_i,1]} \in \mathbb{D}_{\leqslant k} : y_i \in Y_i \text{ and } w_i \in W_i \text{ for } i = 1,\cdots,k\}$$

where Y_i's and W_i's are defined as:

$$Y_i = \begin{cases} (x_i - \delta, x_i + \delta) & i = 1,\cdots j \\ [0,\delta) & i = j+1,\cdots,k \end{cases}, \quad W_i = \begin{cases} (u_i - \delta, u_i + \delta) & i = 1,\cdots j \\ (w^* - \delta, w^* + \delta) & i = j+1,\cdots,k \end{cases}$$

It is evident that $d_{J_1}(\xi,\eta) < k\delta$ for any $\eta \in C_\delta$. Indeed, we could find a suitable time homeomorphism ρ such that $\eta \circ \rho$'s first j jumps and ξ's jumps have aligned locations. Then the uniform metric between ξ and $\eta \circ \rho$ is bounded by k accumulation of size δ. We can choose δ small enough so that (1) $C_\delta \subset G$; (2)for $p \leq j$, the p-th largest jump of η happens on $(u_p - \delta, u_p + \delta)$ with size in range $(x_p - \delta, x_p + \delta)$. Such a choice of δ allows for the following derivation::

$$\begin{aligned}
&\mathbb{P}\left(\hat{J}_n^{\leqslant k} \in G\right) \overset{\text{due to (1)}}{\geq} \mathbb{P}\left(\hat{J}_n^{\leqslant k} \in C_\delta\right) = \mathbb{P}\left(\sum_{i=1}^{k} \frac{Q_n^{\leftarrow}(\Gamma_i)}{n} \mathbb{1}_{[U_i,1]} \in C_\delta\right) \\
&\overset{\text{due to (2)}}{=} \mathbb{P}\left(\Big(\frac{Q_n^{\leftarrow}(\Gamma_1)}{n}, \cdots, \frac{Q_n^{\leftarrow}(\Gamma_k)}{n}\Big) \in \big(\prod_{i=1}^{k} Y_i\big) \cap \mathbb{R}_+^{k\downarrow}, (U_1, \cdots, U_k) \in \big(\prod_{i=1}^{k} W_i\big)\right) \\
&= \mathbb{P}\left(\Big(\frac{Q_n^{\leftarrow}(\Gamma_1)}{n}, \cdots, \frac{Q_n^{\leftarrow}(\Gamma_k)}{n}\Big) \in \big(\prod_{i=1}^{k} Y_i\big) \cap \mathbb{R}_+^{k\downarrow}\right) \cdot \text{Const} \\
&= \check{\mu}_n^k\big(\big(\prod_{i=1}^{k} Y_i\big) \cap \mathbb{R}_+^{k\downarrow}\big) \cdot \text{Const.}
\end{aligned}$$

Note that $\prod_{i=1}^{k} Y_i$ and $\prod_{i=1}^{k} W_i$ represent the product space of Y_i's and W_i's, respectively. Given that $\left(\Pi_{i=1}^{k} A_i\right) \cap \mathbb{R}_+^{k\downarrow}$ is an open set, the extended LDP of $\{\check{\mu}_n^k\}_{n\geq 1}$ as unraveled by Theorem 4.5 implies

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\hat{J}_n^{\leqslant k} \in G\right)}{r(\log n)} \geq \liminf_{n\to\infty} \frac{\log \check{\mu}_n^k\big(\big(\prod_{i=1}^{k} Y_i\big) \cap \mathbb{R}_+^{k\downarrow}\big)}{r(\log n)} + \liminf_{n\to\infty} \frac{\log \text{Const}}{r(\log n)} \\
&\geq - \inf_{x \in \left(\Pi_{i=1}^{k} Y_i\right) \cap \mathbb{R}_+^{k\downarrow}} \check{I}_k\big(x\big) + 0 \\
&= -j = -\hat{I}_k(\xi).
\end{aligned}$$

This establishes inequality (4.33), and, consequently, the lower bound of extended LDP is proved.

Then, we verify the upper bound of extended LDP: for a closed set $F \subset \mathbb{D}_{\leqslant k}$:

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\hat{J}_n^{\leqslant k} \in F\right)}{r(\log n)} \leq -\lim_{\epsilon\downarrow 0} \inf_{\xi \in F^\epsilon} \hat{I}_k(\xi) \tag{4.34}$$

Note that F^ϵ in the above inequality is the fattening set of F restricted to $\mathbb{D}_{\leqslant k}$. If F contains the zero function, then $\lim_{\epsilon\downarrow 0} \inf_{\xi\in F^\epsilon} \hat{I}_k(\xi) = 0$, and the upper bound (4.34) holds trivially. For the rest of the proof, we assume F does not contain the zero function. Define the index i_F as follows:

$$i_F = \max\{j \in \{0, 1, 2, \cdots k-1\} : d_{J_1}(F, \mathbb{D}_{\leqslant j}) > 0\} + 1 \tag{4.35}$$

The definition of i_F has two implications:(i) $F^\epsilon \cap \mathbb{D}_{\leqslant i_F - 1} = \emptyset$ for sufficiently small ϵ; (ii) $F^\epsilon \cap \mathbb{D}_{\leqslant i_F} \neq \emptyset$ for any $\epsilon > 0$. Therefore,

$$\lim_{\epsilon\downarrow 0} \inf_{\xi\in F^\epsilon} \hat{I}_k(\xi) = i_F \tag{4.36}$$

Given that $d_{J_1}(F, \mathbb{D}_{\leqslant i_F - 1}) > 0$, there exist an $r > 0$ such that any path ξ in F has its i_F-th largest jump greater than r. Therefore, F is a subset of C', with C' defined as:

$$\begin{aligned}
C' := \{\sum_{i=1}^{k} y_i \mathbb{1}_{[w_i,1]} :& y_1 \geq y_2 \geq \cdots \geq y_k \geq 0, y_{i_F} > r, \\
& w_i \in (0,1) \text{ for } i = 1, \cdots, k\}
\end{aligned}$$

The set inclusion $F \subset C'$ implies

$$\mathbb{P}\left(\hat{J}_n^{\leqslant k} \in F\right) \leq \mathbb{P}\left(\hat{J}_n^{\leqslant k} \in C'\right) = \mathbb{P}\left(\sum_{i=1}^{k} \frac{Q_n^{\leftarrow}(\Gamma_i)}{n} \mathbb{1}_{[U_i,1]} \in C'\right)$$
$$= \mathbb{P}\left(\frac{Q_n^{\leftarrow}(\Gamma_{i_F})}{n} > r\right) = \mathbb{P}\left(\Gamma_{i_F} \leq Q_n(nr)\right).$$

Taking the limit of both sides and employing limit result (4.93), we deduce

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\hat{J}_n^{\leqslant k} \in F\right)}{r(\log n)} \leq \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\Gamma_{i_F} \leq Q_n(nr)\right)}{r(\log n)} = -i_F \tag{4.37}$$

In view of the inequalities (4.37) and (4.36), the upper bound (4.34) is established, thus concluding the proof.

□

4.2.2 Proof of Lemma 4.2

Lemma 4.6 and 4.7 below are useful in the proof of Lemma 4.2.

Lemma 4.6. *Consider O_i's a sequence of i.i.d. random variables taking non-negative value. Assume $\mathbb{P}(O_1 > x) = \exp(-r(\log x))$, where $r(\cdot)$ being a regularly varying function with index $\gamma > 1$. Define $O_i^{\leqslant n\delta} = O_i \mathbb{1}\{O_i \leqslant n\delta\}$. Then for any given $\epsilon, \delta > 0$ and $M \in \mathbb{N}$, the following inequality holds*

$$\limsup_{n\to\infty} \frac{\log \max_{j=1,\cdots,M\cdot n} \mathbb{P}\left(\sum_{i=1}^{j}\left(O_i^{\leqslant n\delta} - \mathbb{E}\left[O_1\right]\right) > n\epsilon\right)}{r(\log n)} < -\frac{\epsilon}{3\delta} \tag{4.38}$$

Proof of Lemma 4.6. Since the truncated random variables $O_i^{\leqslant n\delta}$ have well-defined moment-generating functions, we can apply the Cramér-Chernoff method to derive an upper bound for the probability term in (4.38). Specifically, for $s > 0$, we have

$$\begin{aligned}\mathbb{P}\left(\sum_{i=1}^{j}\left(O_i^{\leqslant n\delta} - \mathbb{E}\left[O_1\right]\right) > n\epsilon\right) &= \mathbb{P}\left(\exp\left\{s\sum_{i=1}^{j} O_i^{\leqslant n\delta}\right\} > \exp\left\{ns\epsilon + js\mathbb{E}\left[O_1\right]\right\}\right)\\ &\leq \exp\left\{-ns\epsilon - js\mathbb{E}\left[O_1\right]\right\}\mathbb{E}\left[\exp\left\{s\sum_{i=1}^{j} O_i^{\leqslant n\delta}\right\}\right]\\ &= \exp\left\{-ns\epsilon - js\mathbb{E}\left[O_1\right] + \sum_{i=1}^{j}\log\mathbb{E}\left[\exp(sO_i^{\leqslant n\delta})\right]\right\}.\end{aligned} \tag{4.39}$$

Recall $r(\cdot)$ (in Assumption 3.6) is a regularly varying function, we can express it as $r(n) = \lambda r'(n)$ with some $\lambda > 0$ and $r'(n)$ also regularly varying with index γ. Select $0 < \lambda' < \frac{\lambda}{2}$ and set $s = \frac{\lambda' r'(\log n\delta)}{n\delta}$. For sufficiently large n, this choice unsures that $s > 1/n\delta$.

To evaluate $\mathbb{E}\left[\exp(sO_i^{\leqslant n\delta})\right]$, we consider the following decomposition:

$$\mathbb{E}\left[e^{sO_i^{\leqslant n\delta}}\right] = \int_{[0,1/s]} e^{sy} d\mathbb{P}\left(O_1 \leq y\right) + \int_{(1/s,n\delta]} e^{sy} d\mathbb{P}\left(O_1 \leq y\right) + \int_{(n\delta,\infty)} 1 d\mathbb{P}\left(O_1 \leq y\right). \tag{4.40}$$

Using the inequality $e^x \le 1 + x + x^2$ that holds on $x \in [0,1]$, the first integral in (4.40) can be bounded as:

$$\int_{[0,1/s]} e^{sy} d\mathbb{P}\left(O_1 \le y\right) \le \int_{[0,1/s]} (1+sy+s^2y^2) d\mathbb{P}\left(O_1 \le y\right) \le 1+s\mathbb{E}\left[O_1\right]+s^2\mathbb{E}\left[O_1^2\right]. \quad (4.41)$$

Note that the above bound is valid since $\mathbb{E}\left[O_1\right]$ and $\mathbb{E}\left[O_1^2\right]$ are well-defined due to O_1 satisfies Assumption 3.6.

For the second integral in (4.40), we have for sufficiently large n,

$$\begin{aligned}
&\int_{(1/s,n\delta]} e^{sy} d\mathbb{P}\left(O_1 \le y\right) = e^{sy}\mathbb{P}\left(O_1 \le y\right)\Big|_{1/s}^{n\delta} - s\int_{1/s}^{n\delta} e^{sy}\mathbb{P}\left(O_1 \le y\right) dy \\
&= e^{sn\delta}\mathbb{P}\left(O_1 \le n\delta\right) - e\mathbb{P}\left(O_1 \le 1/s\right) - s\int_{1/s}^{n\delta} e^{sy} dy + s\int_{[1/s,n\delta]} e^{sy}\mathbb{P}\left(O_1 > y\right) dy \\
&\le e\big(1 - \mathbb{P}\left(O_1 \le 1/s\right)\big) + s\int_{[1/s,n\delta]} e^{sy}\mathbb{P}\left(O_1 > y\right) dy \\
&\le e\mathbb{P}\left(O_1 > 1/s\right) + sn\delta e^{2sn\delta}\mathbb{P}\left(O_1 > n\delta\right) \qquad (4.42)
\end{aligned}$$

where for the last inequality we used

$$\mathbb{P}\left(O_1 > \frac{1}{s}\right) < e^{sn\delta}\mathbb{P}\left(O_1 > n\delta\right). \quad (4.43)$$

To see the above relation holds for sufficiently large n, it is helpful to realize that

$$\lim_{n\to\infty} \frac{\log e^{sn\delta}\mathbb{P}\left(Z > n\delta\right)}{\log \mathbb{P}\left(Z > 1/s\right)} = \lim_{n\to\infty} \frac{-(\lambda-\lambda')r'(\log n\delta)}{-\lambda r'(\log n\delta - \log\left(\lambda' r'(\log n\delta)\right))} = \frac{\lambda-\lambda'}{\lambda} < 1.$$

Note that the limit computation above uses the property (4.90).

The third integral in (4.40) is $\mathbb{P}\left(O_1 > n\delta\right)$. Combining this with the bounds from (4.41) and (4.42), and leveraging the relation $\mathbb{P}\left(O_1 > n\delta\right) \le \mathbb{P}\left(O_1 > 1/s\right) \le s^2\mathbb{E}\left[O_1^2\right]$ yield that, for sufficiently large n,

$$\mathbb{E}\left[e^{sO_i^{\leqslant n\delta}}\right] \le 1 + sn\delta \cdot e^{2sn\delta}\mathbb{P}\left(O_1 > n\delta\right) + s\mathbb{E}\left[O_1\right] + s^2(e+2)\mathbb{E}\left[O_1^2\right].$$

Inserting the above bound into (4.39), and utilizing the inequality $\log(1+x) \le x$ for $x > 0$, we obtain

$$\begin{aligned}
&\mathbb{P}\left(\sum_{i=1}^{j}\left(O_i^{\leqslant n\delta} - \mathbb{E}\left[O_1\right]\right) > n\epsilon\right) \\
&\le \exp\{-ns\epsilon - js\mathbb{E}\left[O_1\right] + j\log\left(1 + sn\delta\cdot e^{2sn\delta}\mathbb{P}\left(O_1 > n\delta\right) + s\mathbb{E}\left[O_1\right] + s^2(e+2)\mathbb{E}\left[O_1^2\right]\right)\} \\
&\le \exp\{-ns\epsilon - js\mathbb{E}\left[O_1\right] + j\big(sn\delta\cdot e^{2sn\delta}\mathbb{P}\left(O_1 > n\delta\right) + s\mathbb{E}\left[O_1\right] + s^2(e+2)\mathbb{E}\left[O_1^2\right]\big)\} \\
&\le \exp\{-ns\epsilon + j\big(sn\delta\cdot e^{2sn\delta}\mathbb{P}\left(O_1 > n\delta\right) + s^2(e+2)\mathbb{E}\left[O_1^2\right]\big)\}.
\end{aligned}$$

Maximizing over $j = 1, \cdots, M\cdot n$ and taking logarithms gives

$$\begin{aligned}
&\log \max_{j=1,\cdots,M\cdot n} \mathbb{P}\left(\sum_{i=1}^{j}\left(O_i^{\leqslant n\delta} - \mathbb{E}\left[O_1\right]\right) > n\epsilon\right) \\
&\le -ns\epsilon + M\cdot n\big(sn\delta\cdot e^{2sn\delta}\mathbb{P}\left(O_1 > n\delta\right) + s^2(e+2)\mathbb{E}\left[O_1^2\right]\big) \\
&= -\lambda' r'(\log n\delta)\cdot\frac{\epsilon}{\delta} + Mn\lambda' r'(\log n\delta)e^{(2\lambda'-\lambda)r'(\log n\delta)} + \frac{M\lambda'^2(r'(\log n\delta))^2}{n\delta^2}(e+2)\mathbb{E}\left[O_1^2\right].
\end{aligned}$$

The second and third term in the last line above approach to zero as $n \to \infty$. Therefore,

$$\limsup_{n\to\infty} \frac{\log \max_{j=1,\cdots,M\cdot n} \mathbb{P}\left(\sum_{i=1}^{j}\left(O_i^{\leqslant n\delta} - \mathbb{E}\left[Z\right]\right) > n\epsilon\right)}{r(\log n)} \leq \lim_{n\to\infty} \frac{-\lambda' r'(\log n\delta)\cdot\frac{\epsilon}{\delta}}{\lambda r'(\log n)} = -\frac{\epsilon}{\delta}\cdot\frac{\lambda'}{\lambda}.$$

Choosing λ' close to $\lambda/2$ allows us to conclude the lemma. □

Lemma 4.7. *Consider O_i's and $O_i^{\leqslant n\delta}$'s being the same as in the statement of Lemma 4.6. Then for any given $\epsilon, \delta > 0$ and $M \in \mathbb{N}$, the following inequality holds*

$$\limsup_{n\to\infty} \frac{\log \max_{j=1,\cdots,M\cdot n} \mathbb{P}\left(\sum_{i=1}^{j}\left(\mathbb{E}\left[O_1\right] - O_i^{\leqslant n\delta}\right) > n\epsilon\right)}{r(\log n)} = -\infty \tag{4.44}$$

Proof of Lemma 4.7. Since O_1 has the first moment, for large n, $M\mathbb{E}\left[O_1 \mathbb{1}\{O_1 > n\delta\}\right] < \epsilon/2$. For such n and $j = 1, \cdots, M\cdot n$, we have

$$\begin{aligned} &\mathbb{P}\left(\sum_{i=1}^{j}\left(\mathbb{E}\left[O_1\right] - O_i^{\leqslant n\delta}\right) > n\epsilon\right) \\ &= \mathbb{P}\left(\sum_{i=1}^{j}\left(\mathbb{E}\left[O_1^{\leqslant n\delta}\right] - O_i^{\leqslant n\delta}\right) > n\epsilon - j\cdot\mathbb{E}\left[O_1\mathbb{1}\{Q_1 > n\delta\}\right]\right) \\ &\leq \mathbb{P}\left(\sum_{i=1}^{j}\left(\mathbb{E}\left[O_1^{\leqslant n\delta}\right] - O_i^{\leqslant n\delta}\right) > \frac{n\epsilon}{2}\right) \end{aligned} \tag{4.45}$$

It's direct to check that $\mathbb{E}\left[O_i^{\leqslant n\delta}\right] - O_i^{\leqslant n\delta} \leq \mathbb{E}\left[O_1\right]$ and for large n, Var $\left[\mathbb{E}\left[O_i^{\leqslant n\delta}\right] - O_i^{\leqslant n\delta}\right]$ is no greater than Var $\left[O_1\right] + 2\mathbb{E}\left[O_1\right]$. Therefore, we can apply the Bernstein's inequality (see Lemma 4.16 in appendix Section 4.6.2), which yields, for large n,

$$\begin{aligned} \mathbb{P}\left(\sum_{i=1}^{j}\left(\mathbb{E}\left[O_i^{\leqslant n\delta}\right] - O_i^{\leqslant n\delta}\right) > \frac{n\epsilon}{2}\right) &\leq \exp\left\{-\frac{n^2\epsilon^2/4}{2(\sum_i^j \text{Var}\left[\mathbb{E}\left[O_i^{\leqslant n\delta}\right] - O_i^{\leqslant n\delta}\right] + \frac{n\epsilon\mathbb{E}[O_1]}{6}}\right\} \\ &\leq \exp\left\{-\frac{n^2\epsilon^2/4}{2(j(\text{Var}\left[O_1\right] + 2\mathbb{E}\left[O_1\right]) + \frac{n\epsilon\mathbb{E}[O_1]}{6}}\right\}. \end{aligned}$$

Combining the above bound with (4.45) and maximize over $j = 1, \cdots, M\cdot n$, we have, for large n,

$$\log \max_{j=1,\cdots,M\cdot n} \mathbb{P}\left(\sum_{i=1}^{j}\left(\mathbb{E}\left[O_1\right] - O_i^{\leqslant n\delta}\right) > n\epsilon\right) \leq -\frac{n^2\epsilon^2/4}{2(M\cdot n(\text{Var}\left[Z\right] + 2\mathbb{E}\left[O_1\right]) + \frac{n\epsilon\mathbb{E}[O_1]}{6}} \tag{4.46}$$

We can conclude the lemma by realizing the upper bound in (4.46) decrease at least by linear rate while the denominator in (4.44) increase with rate $(\log n)^\gamma$. □

With Lemma 4.6 and 4.7 in hand, we are ready to prove Lemma 4.2.

Proof of Lemma 4.2. For any fixed $k \in \mathbb{N}$, chose $\delta > 0$ such that $k\delta < \epsilon/2$. We begin by analyzing $\|\{\bar{H}_n^{\leqslant k}\| > \epsilon\}$ conditioning on the event $\{N(n) \geq k, Z_{P_n^{-1}(k)} \leq n\delta\}$ and its complement. This leads to the following probability upper bound:

$$
\begin{aligned}
&\mathbb{P}\left(\|\bar{H}_n^{\leqslant k}\| > \epsilon\right) = \mathbb{P}\left(\sup_{t\in[0,1]}\Big|\sum_{i=1}^{N(nt)}\left(Z_i\mathbb{1}(P_n(i) > k) - \mu_1\right)\Big| > n\epsilon\right)\\
&\le \mathbb{P}\left(\max_{j=1,\cdots,N(n)}\Big|\sum_{i=1}^{j}\left(Z_i\mathbb{1}(P_n(i) > k) - \mu_1\right)\Big| > n\epsilon, Z_{P_n^{-1}(k)} \le n\delta, N(n) \ge k\right)\\
&+ \mathbb{P}\left(\{Z_{R_n^{-1}(k)} \le n\delta, N(n) \ge k\}^c\right)\\
&\le \mathbb{P}\left(\max_{j=1,\cdots,N(n)}\sum_{i=1}^{j}\left(Z_i\mathbb{1}(P_n(i) > k) - \mu_1\right) > n\epsilon, Z_{P_n^{-1}(k)} \le n\delta, N(n) \ge k\right) &(4.47)\\
&+ \mathbb{P}\left(\max_{j=1,\cdots,N(n)}\sum_{i=1}^{j}\left(\mu_1 - Z_i\mathbb{1}(P_n(i) > k)\right) > n\epsilon, Z_{P_n^{-1}(k)} \le n\delta, N(n) \ge k\right) &(4.48)\\
&+ \mathbb{P}\left(Z_{P_n^{-1}(k)} > n\delta\right) + \mathbb{P}\left(N(n) < k\right). &(4.49)
\end{aligned}
$$

We proceed with investigating the terms (4.47) and (4.48) separately. Given that $\{P_n(i) > k\}$ is a subset of $\{Z_i \le n\delta\}$ under the condition $Z_{P_n^{-1}(k)} \le n\delta$ and $N(n) \ge k$, the sum $\sum_{i=1}^{j}\left(Z_i\mathbb{1}(Z_i \le n\delta) - \mu_1\right)$ includes more positive Z_i's than $\sum_{i=1}^{j}\left(Z_i\mathbb{1}(P_n(i) > k) - \mu_1\right)$. Therefore, we can upper bound the probability (4.47) as

$$
\begin{aligned}
&(4.47) \le \mathbb{P}\left(\max_{j=1,\cdots,N(n)}\sum_{i=1}^{j}\left(Z_i\mathbb{1}(Z_i \le n\delta) - \mu_1\right) > n\epsilon, Z_{P_n^{-1}(k)} \le n\delta, N(n) \ge k\right)\\
&\le \mathbb{P}\left(\max_{j=1,\cdots,N(n)}\sum_{i=1}^{j}\left(Z_i\mathbb{1}(Z_i \le n\delta) - \mu_1\right) > n\epsilon\right)\\
&\le \mathbb{P}\left(j = 1,\cdots,n(\lfloor\nu_1\rfloor + 1)\max\sum_{i=1}^{j}\left(Z_i\mathbb{1}(Z_i \le n\delta) - \mu_1\right) > n\epsilon\right)\\
&\qquad + \mathbb{P}\left(N(n) > n(\lfloor\nu_1\rfloor + 1)\right). &(4.50)
\end{aligned}
$$

To evaluate (4.48), note that $\mathbb{1}(Z_j \le n\delta) - \mathbb{1}(P_n(j) > k) = 1$ if and only if Z_j is among the k-largest value of Z_i's and $Z_j \le n\delta$, hence $\sum_{i=1}^{j} Z_i\big(\mathbb{1}(Z_i \le n\delta) - \mathbb{1}(P_n(i) > k)\big) \le kn\delta$ for any j. This leads to the following bound for (4.48):

$$
\begin{aligned}
&(4.48) = \mathbb{P}\Big(\max_{j=1,\cdots,N(n)}\sum_{i=1}^{j}\left(\mu_1 - Z_i\mathbb{1}(Z_i \le n\delta)\right) + \sum_{i=1}^{j} Z_i\big(\mathbb{1}(Z_i \le n\delta) - \mathbb{1}(P_n(i) > k)\big) > n\epsilon,\\
&Z_{P_n^{-1}(k)} \le n\delta, N(n) \ge k\Big)\\
&\le \mathbb{P}\Big(\max_{j=1,\cdots,N(n)}\sum_{i=1}^{j}\left(\mu_1 - Z_i\mathbb{1}(Z_i \le n\delta)\right) > n(\epsilon - k\delta), Z_{P_n^{-1}(k)} \le n\delta, N(n) \ge k\Big)\\
&\le \mathbb{P}\Big(\max_{j=1,\cdots,N(n)}\sum_{i=1}^{j}\left(\mu_1 - Z_i\mathbb{1}(Z_i \le n\delta)\right) > \frac{n\epsilon}{2}\Big)\\
&\le \mathbb{P}\left(\max_{j=1,\cdots,n(\lfloor\nu_1\rfloor+1)}\sum_{i=1}^{j}\left(\mu_1 - Z_i\mathbb{1}(Z_i \le n\delta)\right) > \frac{n\epsilon}{2}\right) + \mathbb{P}\left(N(n) > n(\lfloor\nu_1\rfloor + 1)\right) &(4.51)
\end{aligned}
$$

Having established the upper bounds (4.50) and (4.51) for (4.47) and (4.48), and considering the additional terms in (4.49), we can conclude that

$$\mathbb{P}\left(\|\bar{H}_n^{\leqslant k}\| > \epsilon\right) \leq 2\mathbb{P}\left(\max_{j=1,\cdots,n(\lfloor\nu_1\rfloor+1)}\left|\sum_{i=1}^{j}\left(Z_i\mathbb{1}(Z_i \leq n\delta) - \mu_1\right)\right| > \frac{n\epsilon}{2}\right)$$
$$+ \mathbb{P}\left(Z_{P_n^{-1}(k)} > n\delta\right) + \mathbb{P}\left(N(n) < k\right) + 2\mathbb{P}\left(N(n) > n(\lfloor\nu_1\rfloor + 1)\right). \quad (4.52)$$

For the third and fourth terms in (4.52), since $N(n)$ is distributed as the sum of n i.i.d. Poisson(ν_1) random variables and Cramér's theorem confirms those two terms decay as $e^{-n\cdot\mathrm{const}}$. Consequently,

$$\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(N(n) < k\right)}{r(\log n)} = -\infty \text{ and } \limsup_{n\to\infty}\frac{\log 2\mathbb{P}\left(N(n) > n(\lfloor\nu_1\rfloor + 1)\right)}{r(\log n)} = -\infty$$

For the second term in (4.52), the k-th largest Z_i's is distributed as $Q_n^{-1}(\Gamma_k)$. By leveraging the limit result (4.93) in appendix, this term satisfies

$$\begin{aligned}\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(Z_{P_n^{-1}(k)} > n\delta\right)}{r(\log n)} &= \limsup_{n\to\infty}\frac{\log\mathbb{P}\left(Q_n^{\leftarrow}(\Gamma_k) > n\delta\right)}{r(\log n)}\\ &\leq \limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\Gamma_k \leq Q_n(n\delta)\right)}{r(\log n)} = -k\end{aligned}$$

For the first term in (4.52), we employ Etemadi's inequality (see Lemma 4.15 in appendix Section 4.6.2) to externalize the maximum from the probability. Then, invoking Lemma 4.6 and Lemma 4.7 from the appendix, we analyze the term as follows:

$$\begin{aligned}&\limsup_{n\to\infty}\frac{\log 2\mathbb{P}\left(\max_{j=1,\cdots,n(\lfloor\nu_1\rfloor+1)}\left|\sum_{i=1}^{j}\left(Z_i\mathbb{1}(Z_i \leq n\delta) - \mu_1\right)\right| > \frac{n\epsilon}{2}\right)}{r(\log n)}\\
&\leq \limsup_{n\to\infty}\frac{\log 6\max_{j=1,\cdots,n(\lfloor\nu_1\rfloor+1)}\mathbb{P}\left(\left|\sum_{i=1}^{j}\left(Z_i\mathbb{1}(Z_i \leq n\delta) - \mu_1\right)\right| \geq \frac{n\epsilon}{6}\right)}{r(\log n)}\\
&\leq 0 + \max\left\{\limsup_{n\to\infty}\frac{\log\max_{j=1,\cdots,n(\lfloor\nu_1\rfloor+1)}\mathbb{P}\left(\sum_{i=1}^{j}\left(Z_i\mathbb{1}(Z_i \leq n\delta) - \mu_1\right) > \frac{n\epsilon}{7}\right)}{r(\log n)},\right.\\
&\left.\limsup_{n\to\infty}\frac{\log\max_{j=1,\cdots,n(\lfloor\nu_1\rfloor+1)}\mathbb{P}\left(\sum_{i=1}^{j}\left(\mu_1 - Z_i\mathbb{1}(Z_i \leq n\delta)\right) > \frac{n\epsilon}{7}\right)}{r(\log n)}\right\} \leq -\frac{\epsilon}{21\delta}.\end{aligned}$$

Returning to (4.52), for each fixed k, we can choose δ small enough such that $-\epsilon/21\delta < -k$. Therefore, we obtain

$$\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\|\bar{H}_n^{\leqslant k}\| > \epsilon\right)}{r(\log n)} \leq \max\{-\frac{\epsilon}{21\delta}, -k, -\infty, -\infty\} = -k$$

The conclusion of the lemma follows by considering $k \to \infty$. □

4.3 Proofs for Section 3.2

This section gives the proof of Theorem 3.7. Recall the definition of $\bar{W}_n$ in (3.8). Depending on the sign of the increments, we can decompose $\bar{W}_n = \bar{W}_n^+ + \bar{W}_n^-$, where for $t \in [0,1]$,

$$\bar{W}_n^+(t) = \frac{1}{n}\sum_{i=1}^{\lfloor nt\rfloor}\left(Z_i\mathbb{1}\{Z_i \geq 0\} - \mathbb{E}\left[Z_1\mathbb{1}\{Z_1 \geq 0\}\right]\right) \quad (4.53)$$

and

$$\bar{W}_n^-(t) = \frac{1}{n}\sum_{i=1}^{\lfloor nt \rfloor}(Z_i \mathbb{1}\{Z_i < 0\} - \mathbb{E}\left[Z_1 \mathbb{1}\{Z_1 < 0\}\right]). \tag{4.54}$$

Note that $\bar{W}_n^+$ and $\bar{W}_n^-$ are random walks with only positive and negative jumps, respectively. Since we assume Z_i's satisfy Assumption 3.6, the increments of $\bar{W}_n^-$ follow a distribution supported on $(-\infty, 0)$ with a light tail towards $-\infty$.

Let $\tilde{Q}(x)$ be the tail distribution function of $Z_1 \mathbb{1}\{Z_1 \geq 0\}$ and $\tilde{Q}^{\leftarrow}(x)= \inf\{s \geq 0 : \tilde{Q}(s) < x\}$ denotes the generalized inverse of $\tilde{Q}$. Consider i.i.d uniform $[0,1]$ random variables $V_1, V_2, \cdots, V_{n-1}$ and let $V_{(1)}, V_{(2)}, \cdots, V_{(n-1)}$ be their decreasing order statistics. Then $(\tilde{Q}^{\leftarrow}(V^{(1)}), \cdots \tilde{Q}^{\leftarrow}(V^{(n-1)}))$ is distributed as the order(decreasing) statistics of $n-1$ independent samples from the distribution of $Z_1 \mathbb{1}\{Z_1 \geq 0\}$. Consider another sets of i.i.d uniform $[0,1]$ random variables $U_1, \cdots, U_{n-1}$ that are independent from V_i's. Let $R_1, \cdots, R_{n-1}$ indicate the increasing rank of $U_1, \cdots, U_{n-1}$, i.e U_i is the R_i-th smallest element among $U_1, \cdots, U_{n-1}$. Then $(R_1, \cdots R_{n-1})$ has a uniform distribution on the permutation set of index $\{1, \cdots, n-1\}$. With $\tilde{Q}^{\leftarrow}(V^{(i)})$'s and R_i/n standing for the jump size and location respectively, $\bar{W}_n^+$ has the following distributional representation:

$$\bar{W}_n^+ \overset{\mathcal{D}}{=} \frac{1}{n}\sum_{i=1}^{n-1}\left(\tilde{Q}^{\leftarrow}(V_{(i)}) - \mathbb{E}\left[Z_1 \mathbb{1}\{Z_1 \geq 0\}\right]\right)\mathbb{1}_{[\frac{R_i}{n},1]} + \frac{1}{n}(Z_n - \mathbb{E}\left[Z_1 \mathbb{1}\{Z_1 \geq 0\}\right])\mathbb{1}_{\{1\}}. \tag{4.55}$$

We further define $\tilde{S}_n$ by replacing the jump location R_i/n in the above representation with U_i:

$$\tilde{S}_n = \frac{1}{n}\sum_{i=1}^{n-1}\left(\tilde{Q}^{\leftarrow}(V_{(i)}) - \mathbb{E}\left[Z_1 \mathbb{1}\{Z_1 \geq 0\}\right]\right)\mathbb{1}_{[U_i,1]} + \frac{1}{n}(Z_n - \mathbb{E}\left[Z_1 \mathbb{1}\{Z_1 \geq 0\}\right])\mathbb{1}_{\{1\}}. \tag{4.56}$$

It is direct to see $\tilde{S}_n = \tilde{J}_n^k + \tilde{H}_n^k$, where

$$\tilde{J}_n^k \triangleq \frac{1}{n}\sum_{i=1}^{k}\tilde{Q}^{\leftarrow}(V_{(i)})\mathbb{1}_{[U_i,1]} + \frac{1}{n}Z_n \mathbb{1}_{\{1\}} \tag{4.57}$$

and

$$\tilde{H}_n^k \triangleq \frac{1}{n}\sum_{i=k+1}^{n-1}\tilde{Q}^{\leftarrow}(V_{(i)})\mathbb{1}_{[U_i,1]} - \frac{1}{n}\mathbb{E}\left[Z\right]\sum_{i=1}^{n-1}\mathbb{1}_{[U_i,1]} - \frac{1}{n}\mathbb{E}\left[Z\right]\mathbb{1}_{\{1\}}. \tag{4.58}$$

We will conclude the extended LDP of $\{\tilde{S}_n\}_{n\geq 1}$ with the sequence of Lemma 4.8, 4.9 and 4.10 stated below. Those lemmas are parallel to Proposition 4.5, Proposition 4.1 and Theorem 3.4 in Section 3.1, sharing a comparable proof framework with subtle differences in computation. Then the proof of Theorem 3.7 proceed with showing the equivalence between $\tilde{S}_n$, $\bar{W}_n^+$ and $\bar{W}_n$ in the sense of Corollary 2.6. The proofs of Lemmas 4.8, 4.9, and 4.10 will follow afterward.

Lemma 4.8. *For any $k \in \mathbb{N}$, recall that $\mathbb{R}_+^{k\downarrow} = \{(x_1, \cdots, x_k) \in \mathbb{R}^k : x_1 \geq x_2 \geq \cdots \geq x_k \geq 0\}$. The sequence of measures $\{\tilde{\mu}_n\}_{n\geq 1}$ on $\mathbb{R}_+^{k\downarrow}$ defined by*

$$\tilde{\mu}_n(\cdot) \triangleq \mathbb{P}\left(\left(\frac{\tilde{Q}^{\leftarrow}(V_{(1)})}{n}, \cdots, \frac{\tilde{Q}^{\leftarrow}(V_{(k)})}{n}\right) \in \cdot\right) \tag{4.59}$$

satisfies the extended LDP with the rate $\tilde{I}^k : \mathbb{R}_+^{k\downarrow} \to \mathbb{R}$ given by

$$\tilde{I}^k(\boldsymbol{x}) \triangleq \sum_{i=1}^{k}\mathbb{1}\{x_i \neq 0\} \text{ with } = (x_1, \cdots, x_k) \in \mathbb{R}_+^{k\downarrow} \tag{4.60}$$

and speed $r(\log n)$.

Lemma 4.9. *The sequence $\{\tilde{J}_n^k\}_{n\geq 1}$ defined in (4.57) satisfies the extended LDP on $(\mathbb{D}, d_{J_1})$ with rate function $\tilde{I}_k : \mathbb{D} \to \mathbb{R}$ given by*

$$\tilde{I}_k(\xi) \triangleq \begin{cases} \sum_{t\in(0,1]} \mathbb{1}\{\xi(t) \neq \xi(t-)\} & \text{if } \xi \in \tilde{\mathbb{D}}_{\leqslant k} \\ \infty & \text{otherwise} \end{cases} \tag{4.61}$$

and speed $r(\log n)$.

Lemma 4.10. *The sequence of processes $\{\tilde{S}_n\}_{n\geq 1}$ satisfies the extended large deviation on $(\mathbb{D}, d_{J_1})$ with the rate function given by (3.9) and speed $r(\log n)$.*

Proof of Theorem 3.7. Using the distributional representation of $\bar{W}_n^+$ in (4.55) and definition of $\tilde{S}_n$ in (4.56), we have

$$\begin{aligned}\mathbb{P}\left(d_{J_1}(\bar{W}+_n, \tilde{S}_n) > \epsilon\right) &\leq \mathbb{P}\left(\sup_{1\leq i\leq n-1} \left|\frac{R_i}{n} - U_i\right| > \epsilon\right) = \mathbb{P}\left(\sup_{1\leq i\leq n-1} \left|\frac{i}{n} - U_{(i)}\right| > \epsilon\right) \\ &\leq \mathbb{P}\left(\sqrt{n} \sup_{x\in[0,1]} \left|\frac{1}{n}\sum_{i=1}^n \mathbb{1}(U_i \leq x) - x\right| > \sqrt{n}\epsilon\right) \leq 2e^{-2\epsilon^2 n},\end{aligned}$$

where the last inequality above is due to Corollary 1 of [20]. This further implies

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(d_{J_1}(\bar{W}_n, \tilde{S}_n)\right)}{r(\log n)} \leq -\lim_{n\to\infty} \frac{2\epsilon^2 n}{r(\log n)} = -\infty.$$

Therefore, $\{\tilde{S}_n\}_{n\geq 1}$ and $\{\bar{W}_n\}_{n\geq 1}$ satisfies the extended LDP with the same rate function and speed sequence by Corollary 2.6. Also, we have

$$\mathbb{P}\left(d_{J_1}(\bar{W}_n, \bar{W}_n^+) > \epsilon\right) \leq \mathbb{P}\left(\|\bar{W}_n^-\| > \epsilon\right)$$

Note that $\bar{W}_n^-$ is a random walk with increments of light tails. According to Mogulskii Theorem (see Theorem 5.1.2 in [9]), the term $\mathbb{P}\left(\|\bar{W}_n^-\| > \epsilon\right)$ decays geometrically with respect to n. Therefore, by Corollary 2.6 again, $\{\bar{W}_n^+\}_{n\geq 1}$ and $\{\bar{W}_n\}_{n\geq 1}$ satisfies the extended LDP with the same rate function and speed sequence. This complete the proof of the theorem. $\square$

Proof of Lemma 4.8. The rate function $\tilde{I}^k$ is the same as $\check{I}_k$ in (4.24), which was shown to be lower semi-continuous in Proposition 4.5. For the rest, we focus on verifying the lower and upper bounds of the extended LDP.

For the lower bound of the extended LDP, consider an open set $G \subset \mathbb{R}_+^{k\downarrow}$. Fix an arbitrary $x \in G$ and let $\tilde{I}^k(x) = \iota$. We can assume w.l.o.g that

$$x_1 > \cdots > x_\iota > x_{\iota+1} = 0. \tag{4.62}$$

Using a similar argument for (4.27) in Proposition 4.5, we can find continuous functions $f_m : \mathbb{R}_+^{k\downarrow} \to [-m, 0]$ for each $m \in \mathbb{N}$ such that $f_m(x) = 0$ and $f_m(y) = -m$ for $y \in G^c$, and the following inequality is satisfied:

$$\liminf_{n\to\infty} \frac{\log \int_{\mathbb{R}_+^{k\downarrow}} e^{r(\log n) f_m(x)} \tilde{\mu}_n(dx)}{r(\log n)} \leq \max\{\liminf_{n\to\infty} \frac{\log \tilde{\mu}_n(G)}{r(\log n)}, -m\} \tag{4.63}$$

If the following inequality holds:

$$-\tilde{I}^k(x) \leq \liminf_{n\to\infty} \frac{\log \int_{\mathbb{R}_+^{k\downarrow}} e^{r(\log n) f_m(x)} \tilde{\mu}_n(dx)}{r(\log n)}, \tag{4.64}$$

then the lower bound of the extended LDP is obtained by combining (4.63) and (4.64), then taking the limit $m \to \infty$ and the infimum over $x \in G$.

Therefore, to complete the proof of the lower bound of the extended LDP, it remains to show (4.64). Recall the neighborhood $D(x;\delta) = \prod_{i=1}^{\iota}[x_i - \delta, x_i + \delta] \times [0,\delta]^{(k-\iota)}$ of x defined in Lemma 4.4. Pick an arbitrary $\epsilon > 0$. Since G is open and f_m is continuous, assumption (4.62) allows us to choose a small enough $\delta = \delta(x,\epsilon)$ so that

(i) $[x_i - \delta, x_i + \delta] \cap [x_j - \delta, x_j + \delta] = \emptyset$ for any i and j such that $i < j \le \iota$;

(ii) $f_m(y) > -\epsilon$ for $y \in D(x;\delta)$.

From (ii), we have

$$\liminf_{n\to\infty} \frac{\log \int_{\mathbb{R}_+^{k\downarrow}} e^{r(\log n) f_m(x)} \tilde{\mu}_n(dx)}{r(\log n)} \ge \liminf_{n\to\infty} \frac{\log \int_{D_\delta} e^{\tilde{Q}(n)\epsilon} \tilde{\mu}_n(dx)}{r(\log n)} \ge -\epsilon + \liminf_{n\to\infty} \frac{\log \tilde{\mu}_n(D_\delta)}{r(\log n)} \tag{4.65}$$

According to D_δ's form, the term $\tilde{\mu}_n(D_\delta)$ satisfies:

$$\begin{aligned}
\tilde{\mu}_n(D_\delta) \ge \mathbb{P}\Big(\{&\frac{\tilde{Q}^{\leftarrow}(V_{(1)})}{n} \in [x_1 - \delta, x_1 + \delta], \cdots, \frac{\tilde{Q}^{\leftarrow}(V_{(\tilde{I}^k(\hat{\boldsymbol{x}}))})}{n} \in [x_\iota - \delta, x_\iota + \delta], \\
&\frac{\tilde{Q}^{\leftarrow}(V_{(j)})}{n} \in [0,\delta) \text{ for } j = \iota+1, \cdots, n-1\}\Big) \\
= \binom{n-1}{k} &\cdot \prod_{i=1}^{\iota} \big(\tilde{Q}(n(x_i - \delta) - \tilde{Q}(n(x_i + \delta))\big) \cdot \big(1 - \tilde{Q}(n\delta)\big)^{n-1-\iota},
\end{aligned}$$

and this implies

$$\begin{aligned}
\liminf_{n\to\infty} \frac{\log \tilde{\mu}_n(D_\delta)}{r(\log n)} \ge \lim_{n\to\infty} \frac{\log \binom{n-1}{k}}{r(\log n)} + \sum_{i=1}^{\iota} \lim_{n\to\infty} \frac{\log \big(\tilde{Q}(n(x_i - \delta) - \tilde{Q}(n(x_i + \delta))\big)}{r(\log n)} \\
+ \lim_{n\to\infty} \frac{(n-1-\iota) \cdot \log \big(1 - \tilde{Q}(n\delta)\big)}{r(\log n)} = -\iota.
\end{aligned}$$

The last equality above can be derived by inserting $\tilde{Q}(x) = \exp(-r(\log x))$ and leveraging $\log \binom{n-1}{k} \sim \log(n-1)^k \sim k \log n$, the limit result (4.91) and (4.88). Plugging the above inequality into (4.65) and taking $\epsilon \to 0$ yields (4.64).

For the upper bound of the extended LDP, consider a closed set $F \subset \mathbb{R}_+^{k\downarrow}$ and let $\iota = \lim_{\epsilon\downarrow 0} \inf_{x \in F^\epsilon} \check{I}_k(x)$. Then it is direct to see that there exists $r > 0$ such that $x_\iota > r$ for all $x \in F$. Therefore,

$$\begin{aligned}
\limsup_{n\to\infty} \frac{\log \check{\mu}_n(F)}{r(\log n)} &= \limsup_{n\to\infty} \frac{\log \mathbb{P}\Big(\Big(\frac{\tilde{Q}^{\leftarrow}(V_{(1)})}{n}, \frac{\tilde{Q}^{\leftarrow}(V_{(2)})}{n}, \cdots \frac{\tilde{Q}^{\leftarrow}(V_{(k)})}{n}\Big) \in F\Big)}{r(\log n)} \\
&\le \limsup_{n\to\infty} \frac{\log \mathbb{P}\Big(\{\tilde{Q}^{\leftarrow}(V_{(\iota)}) \ge nr\}\Big)}{r(\log n)} = \limsup_{n\to\infty} \frac{\log \mathbb{P}\Big(\{V_{(\iota)} \le \tilde{Q}(nr)\}\Big)}{r(\log n)} \\
&\le -\iota = -\lim_{\epsilon\downarrow 0} \inf_{x \in F^\epsilon} \tilde{I}^k(x).
\end{aligned}$$

where the last inequality is due to $\tilde{Q}(x) = \exp(-r(\log x))$ and (4.93). $\square$

Proof of Lemma 4.9. We only prove the extended LDP on $(\tilde{\mathbb{D}}_{\leqslant k}, d_{J_1})$, as that can be turned to an extended LDP on $(\mathbb{D}, d_{J_1})$ with the help of Lemma 2.4. The conditions to apply Lemma 2.4 are met, as $\mathbb{P}\left(\tilde{J}_n^k \in \tilde{\mathbb{D}}_{\leqslant k}\right) = 1$ and $\tilde{\mathbb{D}}_{\leqslant k}$ is closed.

The rate function $\tilde{I}_k$ is lower-semicontinuous, as its level set $\Psi_{\tilde{I}_k}(c) = \tilde{\mathbb{D}}_{\leqslant \min\{\lfloor c \rfloor, k\}}$ is closed. It remains to show the extened LDP's lower bound: for any G as a subset $\tilde{\mathbb{D}}_{\leqslant k}$ and open

$$\liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in G\right)}{r(\log n)} \geq -\inf_{\xi\in G} \tilde{I}_k(\xi) \tag{4.66}$$

and extended LDP's upper bound: for any F a subset $\tilde{\mathbb{D}}_{\leqslant k}$ and closed

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in F\right)}{r(\log n)} \leq -\lim_{\epsilon\downarrow 0} \inf_{\xi\in F^\epsilon} \tilde{I}_k(\xi) \tag{4.67}$$

(Here F^ϵ is the set $\{\eta \in \tilde{\mathbb{D}}_{\leqslant k} : d_{J_1}(\eta, F) \leq \epsilon\}$).

For the lower bound (4.66), we discuss the following two cases:

Case 1: If ξ has j jumps for some $j \geq k$ and all those jumps fall on $(0,1)$. In such case, $\xi = \sum_{i=1}^j x_i \mathbb{1}_{[u_i,1]}$ with x_i's being the jump sizes of non-increasing order and $u_i \in (0,1)$ being the corresponding jump times. We construct a neighbourhood of ξ of the following form

$$\begin{aligned} C_\delta \triangleq \Big\{\eta = \sum_{i=1}^k y_i \mathbb{1}_{[w_i,1]} + y_{k+1}\mathbb{1}_{\{1\}} \in \mathbb{D} : \boldsymbol{y} = (y_1, \cdots, y_{k+1}) \in \big((\prod_{i=1}^k Y_i) \cap \mathbb{R}_+^{k\downarrow}\big) \times Y_{k+1}, \\ \boldsymbol{w} = (w_1, \cdots, w_k) \in \prod_{i=1}^k W_i, \Big\} \end{aligned} \tag{4.68}$$

with the sets Y_i's and W_i's defined as below

$$Y_i = \begin{cases} (x_i - \delta, x_i + \delta) & i = 1, \cdots, j \\ [0, \delta) & i = j+1, \cdots, k+1 \end{cases}, \quad W_i = \begin{cases} (u_i - \delta, u_i + \delta) & i = 1, \cdots, j \\ (w^* - \delta, w^* + \delta) & i = j+1, \cdots, k \end{cases}$$

In the above, w^* is a fixed time that does not belong to $\{u_1, \cdots, u_j\}$. It is not hard to see for any $\eta \in C_\delta$ $d_{J_1}(\xi, \eta) < (k+1)\delta$ is satisfied, hence we can make $C_\delta \in G$ by choosing δ small.

Based on $C_\delta \in G$, we have

$$\begin{aligned} &\mathbb{P}\left(\tilde{J}_n^k \in G\right) \geq \mathbb{P}\left(\tilde{J}_n^k \in C_\delta\right) \\ =\ &\mathbb{P}\left(\frac{1}{n}\sum_{i=1}^k \tilde{Q}^{\leftarrow}(V_{(i)})\mathbb{1}_{[U_i,1]} + \frac{1}{n} Z_n \mathbb{1}_{\{1\}} \in C_\delta\right) \\ =\ &\mathbb{P}\left(\Big(\frac{\tilde{Q}^{\leftarrow}(V_{(1)})}{n}, \cdots, \frac{\tilde{Q}^{\leftarrow}(V_{(k)})}{n}\Big) \in (\prod_{i=1}^k Y_i) \cap \mathbb{R}_+^{k\downarrow}\right) \mathbb{P}\left(\frac{Z_n}{n} \in Y_{k+1}\right) \\ &\mathbb{P}\left((U_1, \cdots, U_k) \in B\right) \\ =\ &\tilde{\mu}_n((\prod_{i=1}^k Y_i) \cap \mathbb{R}_+^{k\downarrow}) \cdot \mathbb{P}\left(Z_n < n\delta\right) \cdot \text{Const.} \end{aligned}$$

Therefore,

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in G\right)}{r(\log n)} \\
&\geq \liminf_{n\to\infty} \frac{\log \tilde{\mu}_n((\prod_{i=1}^k Y_i) \cap \mathbb{R}_+^{k\downarrow}) + \log \mathbb{P}\left(Z_n < n\delta\right)}{r(\log n)} \\
&= \liminf_{n\to\infty} \frac{\log \tilde{\mu}_n\big((\prod_{i=1}^k Y_i) \cap \mathbb{R}_+^{k\downarrow}\big)}{r(\log n)} + \liminf_{n\to\infty} \frac{\log\big(1-\tilde{Q}(n\delta)\big)}{r(\log n)} &(4.69)\\
&\geq - \inf_{\boldsymbol{x}\in(\prod_{i=1}^k Y_i)\cap \mathbb{R}_+^{k\downarrow}} \tilde{I}^k\big(\boldsymbol{x}\big) + 0 \geq -\tilde{I}^k\big((x_1,\cdots,x_j,0,\cdots,0)\big) = -j = -\tilde{I}_k(\xi). &(4.70)
\end{aligned}$$

Note that to lower bound two limit terms in from (4.69), we use the conclusion from Lemma 4.8 such that $\{\tilde{\mu}_n\}_{n\geq 1}$ satisfying the extended LDP, and the limit results (4.90). The second inequality in to (4.70) holds due to $(x_1,\cdots,x_j,0,\cdots,0) \in (\prod_{i=1}^k Y_i) \cap \mathbb{R}_+^{k\downarrow}$.

Case 2: If ξ has j jumps for some $j \geq k$ and one of those jump occur at $t=1$. In such case, $\xi = \sum_{i=1}^{j-1} x_i \mathbb{1}_{[u_i,1]} + x_j \mathbb{1}_{\{1\}}$ with $x_1,\cdots,x_{j-1}$ being the jump sizes of non-increasing order. We construct the set C_δ as (4.68), but modify Y_i's and W_i's definitions as below

$$Y_i = \begin{cases} (x_i-\delta, x_i+\delta) & i=1,\cdots,j-1 \\ [0,\delta) & i=j,\cdots,k \\ (x_j-\delta,x_j+\delta) & i=k+1 \end{cases}, \quad B_i = \begin{cases} (u_i-\delta,u_i+\delta) & i=1,\cdots,j-1 \\ (\frac{1}{2}-\delta, \frac{1}{2}+\delta) & i=j,\cdots,k \end{cases}$$

Similar to the case 1 above, C_δ in this case is a neighborhood of ξ and we can choose δ small enough to make it as a subset of G. By the extended LDP of $\{\tilde{\mu}_n\}_{n\geq 1}$ and $(x_1,\cdots,x_{j-1},0,\cdots,0) \in (\prod_{i=1}^k Y_i) \cap \mathbb{R}_+^{k\downarrow}$, we can conclude

$$\begin{aligned}
&\liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in G\right)}{r(\log n)} \geq \liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in C_\delta\right)}{r(\log n)} \\
&= \liminf_{n\to\infty} \frac{\log \tilde{\mu}_n\big((\prod_{i=1}^k Y_i) \cap \mathbb{R}_+^{k\downarrow}\big) + \log \mathbb{P}\left(\frac{Z_n}{n} \in Y_{k+1}\right) + \log \mathbb{P}\left((U_1,\cdots,U_k) \in B\right)}{r(\log n)} \\
&\geq \liminf_{n\to\infty} \frac{\log \tilde{\mu}_n(A\cap \mathbb{R}_+^{k\downarrow}) + \log \mathbb{P}\left(n(x_j-\delta) < Z_n < n(x_j+\delta)\right)}{r(\log n)} \\
&\geq \liminf_{n\to\infty} \frac{\log \tilde{\mu}_n(A\cap\mathbb{R}_+^{k\downarrow})}{r(\log n)} + \liminf_{n\to\infty} \frac{\log\big(\tilde{Q}(n(x_j-\frac{\delta}{2})) - \tilde{Q}(n(x_j+\frac{\delta}{2}))\big)}{r(\log n)} &(4.71)\\
&\geq - \inf_{\boldsymbol{x}\in(\prod_{i=1}^k Y_i)\cap \mathbb{R}_+^{k\downarrow}} \tilde{I}^k\big(\boldsymbol{x}\big) - 1 \geq -j \geq -\tilde{I}^k\big((x_1,\cdots,x_{j-1},0,\cdots,0)\big) - 1 = -\tilde{I}_k(\xi).
\end{aligned}$$

Note that the second limit term in (4.71) equals -1 is due to the limit result (4.87). Combining the two above cases yield

$$\liminf_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in G\right)}{r(\log n)} \geq -\tilde{I}_k(\xi)$$

and the lower bound (4.66) is established by taking supremum over all $\xi \in G$.

Now we turn to prove the upper bound (4.67). If the closed set F contains the zero function, then $\lim_{\epsilon\downarrow 0}\inf_{\xi\in F^\epsilon} \tilde{I}_k(\xi) = 0$, and the upper bound (4.67) is trivial. Hence we consider F that does not contain the zero function and define

$$i^* = \max\{j \in \{0,1,2,\cdots k-1\},: d_{J_1}(F, \tilde{D}_{\leqslant j}) > 0\} + 1 \tag{4.72}$$

Based on this definition, F is bounded away from $\tilde{\mathbb{D}}_{\leqslant i^*-1}$, which implies

$$\lim_{\epsilon \downarrow 0} \inf_{\xi \in F^\epsilon} \tilde{I}_k(\xi) = i^* \tag{4.73}$$

Also, paths in F has at least i^* jumps, and due to $d_{J_1}(F, \tilde{D}_{\leqslant i^*-1}) > 0$, we can find some $r > 0$ such that paths in F should have their i^*-th largest jump size no smaller than r.

Clearly, F can be represented by $F_1 \cup F_2$, where

$$F_1 = \{\xi \in F, \xi(1) - \xi(1-) \text{ does not belong to } \xi's\ i^* \text{ largest jumps sizes}\}$$
$$F_2 = \{\xi \in F, \xi(1) - \xi(1-) \text{ belongs to } \xi's\ i^* \text{ largest jumps sizes}\}.$$

Therefore,

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in F\right)}{r(\log n)} = \max\Bigg\{\underbrace{\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in F_1\right)}{r(\log n)}}_{\text{(I)}}, \underbrace{\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in F_2\right)}{r(\log n)}}_{\text{(II)}}\Bigg\} \tag{4.74}$$

For the term (I), we have

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in F_1\right)}{r(\log n)} = \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^k \tilde{Q}^{\leftarrow}(V_{(i)})\mathbb{1}_{[U_i,1]} + \frac{1}{n}Z\mathbb{1}_{\{1\}} \in F_1\right)}{r(\log n)} \\
&\leq \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{Q}^{\leftarrow}(V_{(i^*)}) \geq nr, Z \in \mathbb{R}_+, U_i \in (0,1) \text{ for } i = 1, \cdots, k\right)}{r(\log n)} \\
&= \lim_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{Q}^{\leftarrow}(V_{(i^*)}) \geq nr\right)}{r(\log n)} = \lim_{n\to\infty} \frac{\log \mathbb{P}\left(V_{(i^*)} \geq \tilde{Q}(nr)\right)}{r(\log n)} = -i^*.
\end{aligned} \tag{4.75}$$

For the term (II), we have

$$\begin{aligned}
&\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in F_2\right)}{r(\log n)} = \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\frac{1}{n}\sum_{i=1}^k \tilde{Q}^{\leftarrow}(V_{(i)})\mathbb{1}_{[U_i,1]} + \frac{1}{n}Z\mathbb{1}_{\{1\}} \in F_2\right)}{r(\log n)} \\
&\leq \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{Q}^{\leftarrow}(V_{(i^*-1)}) \geq nr, Z \geq nr, U_i \in (0,1) \text{ for } i = 1, \cdots, k\right)}{r(\log n)} \\
&= \lim_{n\to\infty} \frac{\log \mathbb{P}\left(V_{(i^*-1)} \geq \tilde{Q}(nr)\right)}{r(\log n)} + \lim_{n\to\infty} \frac{\log \tilde{Q}(nr)}{r(\log n)} = -(i^*-1) - 1 = -i^*.
\end{aligned} \tag{4.76}$$

Note that to yield the equation in (4.75) and (4.76), we plugin the form of $\tilde{Q}$ and use the limit result (4.93)) and (4.90).

Combining the calculation for (I) and (II), (4.74) becomes

$$\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\tilde{J}_n^k \in F_2\right)}{r(\log n)} \leq -i^*.$$

With (4.73), the upper bound (4.67)is established. This finishes the proof. □

Proof of Lemma 4.10. Since $\bar{S}_n$ and $\tilde{S}_n$ has the same distribution, we will infer the extended LDP of $\{\tilde{S}_n\}_{n\geq 1}$. For this purpose, we apply the approximation lemma 2.5. As the extended LDP of $\{\tilde{J}_n^k\}_{n\geq 1}$ is confirm by Lemma 4.9, it remains to verify Lemma 2.5's conditions (2.3), (2.4) and (2.5), with $\tilde{I}_k$, $\tilde{I}$, $\tilde{J}_n^k$, $\tilde{S}_n$, $r(\log n)$ being I_k, I, Y_n^k,X_n, a_n according to Lemma 2.5's notations.

Since $\tilde{I}_k(\cdot) \geq \tilde{I}(\cdot)$ on $\mathbb{D}$, the condition (2.3) is clearly met. In Theorem 3.4, we have shown the condition (2.4) is satisfied w.r.t $\hat{I}_k$ and I^{J_1}. Here, we use the same argument the condition (2.4) is satisfied w.r.t $\tilde{I}_k$ and $\tilde{I}$ holds. To see the condition (2.5) is satisfied, note that

$$\begin{aligned}
&\mathbb{P}\left(d_{J_1}(\tilde{S}_n, \tilde{J}_n^k) > \epsilon\right) \leq \mathbb{P}\left(\|\tilde{H}_n^k\| > \epsilon\right) \\
&= \mathbb{P}\left(\|\frac{1}{n}\sum_{i=k+1}^{n-1} \tilde{Q}^{\leftarrow}(V_{(i)})\mathbb{1}_{[U_i,1]} - \frac{1}{n}\mathbb{E}\left[Z\right]\sum_{i=1}^{n-1} I_{[U_i,1]} - \frac{1}{n}\mathbb{E}\left[Z\right]\mathbb{1}_{\{1\}}\| > \epsilon\right) \\
&\leq \mathbb{P}\left(\|\frac{1}{n}\sum_{i=k+1}^{n-1} \tilde{Q}^{\leftarrow}(V_{(i)})\mathbb{1}_{[U_i,1]} - \frac{1}{n}\mathbb{E}\left[Z\right]\sum_{i=1}^{n-1} I_{[U_i,1]}\| > \frac{\epsilon}{2}\right) \\
&+ \mathbb{P}\left(\|\frac{1}{n}\mathbb{E}\left[Z\right]\mathbb{1}_{\{1\}}\| > \frac{\epsilon}{2}\right).
\end{aligned}$$

It's clear the term $\mathbb{P}\left(\|\frac{1}{n}\mathbb{E}\left[Z\right]\mathbb{1}_{\{1\}}\|_\infty > \frac{\epsilon}{2}\right)$ becomes zero for large n. For such large n,

$$\begin{aligned}
&\mathbb{P}\left(d_{J_1}(\tilde{S}_n, \tilde{J}_n^k) > \epsilon\right) \\
&\leq \mathbb{P}\left(\|\frac{1}{n}\sum_{i=k+1}^{n-1} \tilde{Q}^{\leftarrow}(V_{(i)})\mathbb{1}_{[U_i,1]} - \frac{1}{n}\mathbb{E}\left[Z\right]\sum_{i=1}^{n-1} I_{[U_i,1]}\| > \frac{\epsilon}{2}, \tilde{Q}^{\leftarrow}(V_{(k)}) \geq n\delta\right) \\
&+ \mathbb{P}\left(\|\frac{1}{n}\sum_{i=k+1}^{n-1} \tilde{Q}^{\leftarrow}(V_{(i)})\mathbb{1}_{[U_i,1]} - \frac{1}{n}\mathbb{E}\left[Z\right]\sum_{i=1}^{n-1} I_{[U_i,1]}\| > \frac{\epsilon}{2}, \tilde{Q}^{\leftarrow}(V_{(k)}) \leq n\delta\right) \\
&\leq \mathbb{P}\left(\tilde{Q}^{\leftarrow}(V_{(k)}) \geq n\delta\right) \\
&+ \mathbb{P}\left(\|\frac{1}{n}\sum_{i=k+1}^{n-1} \tilde{Q}^{\leftarrow}(V_{(i)})\mathbb{1}_{[U_i,1]} - \frac{1}{n}\mathbb{E}\left[Z\right]\sum_{i=1}^{n-1} I_{[U_i,1]}\| > \frac{\epsilon}{2}, \tilde{Q}^{\leftarrow}(V_{(k)}) \leq n\delta\right). \qquad (4.77)
\end{aligned}$$

Let $Z_1', \cdots, Z_{n-1}'$ be $n-1$ independent copies of the generic random variable Z and $P_{n-1}'(\cdot)$ be a random permutation on $\{1, \cdots, n-1\}$ such that $Z'_{P'_{n-1}(i)}$ is the i-th largest item among $Z_1', \cdots, Z_{n-1}'$. Since $\tilde{Q}^{\leftarrow}(V_{(1)}), \cdots, \tilde{Q}^{\leftarrow}(V_{(n-1)})$ has the same distribution of the ascending order of $Z_1', \cdots, Z_{n-1}'$,

$$\begin{aligned}
(4.77) =&\mathbb{P}\left(\max_{j=1,\cdots,n-1}\Big|\sum_{i=1}^{j}\left(Z_i'\mathbb{1}\{P_{n-1}'(i) > k\} - \mathbb{E}\left[Z\right]\right)\Big| > \frac{n\epsilon}{2}, Z'_{P'_{n-1}(k)} \leq n\delta\right) \\
&\leq \mathbb{P}\left(\max_{j=1,\cdots,n-1}\sum_{i=1}^{j}\left(Z_i'\mathbb{1}\{P_{n-1}'(i) > k\} - \mathbb{E}\left[Z\right]\right) > \frac{n\epsilon}{2}, Z'_{P'_{n-1}(k)} \leq n\delta\right) \\
&+ \mathbb{P}\left(\max_{j=1,\cdots,n-1}\sum_{i=1}^{j}\left(\mathbb{E}\left[Z\right] - Z_i'\mathbb{1}\{P_{n-1}'(i) > k\}\right) > \frac{n\epsilon}{2}, Z'_{P'_{n-1}(k)} \leq n\delta\right) \\
&\leq \mathbb{P}\left(\max_{j=1,\cdots,n-1}\sum_{i=1}^{j}\left(Z_i'\mathbb{1}\{Z_i' \leq n\delta\} - \mathbb{E}\left[Z\right]\right) > \frac{n\epsilon}{2}\right) \\
&+ \mathbb{P}\left(\max_{j=1,\cdots,n-1}\sum_{i=1}^{j}\left(\mathbb{E}\left[Z\right] - Z_i'\mathbb{1}\{Z_i' \leq n\delta\}\right) + n\delta k > \frac{n\epsilon}{2}\right).
\end{aligned}$$

We can choose δ small to satisfy $\delta k < \epsilon/4$ in the last line above. This, along with the Etemadi's inequality (see Lemma 4.15 in Section 4.6.2), enable us to further upper bound

(4.77) by

$$
\begin{aligned}
(4.77) \leq & \mathbb{P}\left(\max_{j=1,\cdots,n-1} \Big| \sum_{i=1}^{j} \big(Z_i' \mathbb{1}\{Z_i' \leq n\delta\} - \mathbb{E}\,[Z] \big) \Big| > \frac{n\epsilon}{4} \right) \\
\leq & 3 \max_{j=1,\cdots,n} \mathbb{P}\left(\Big| \sum_{i=1}^{j} \big(Z_i' \mathbb{1}\{Z_i' \leq n\delta\} - \mathbb{E}\,[Z] \big) \Big| > \frac{n\epsilon}{12} \right) \\
\leq & 3 \max_{j=1,\cdots,n} \mathbb{P}\left(\sum_{i=1}^{j} \big(Z_i' \mathbb{1}\{Z_i' \leq n\delta\} - \mathbb{E}\,[Z] \big) > \frac{n\epsilon}{12} \right) \\
+ & 3 \max_{j=1,\cdots,n} \mathbb{P}\left(\sum_{i=1}^{j} \big(\mathbb{E}\,[Z] - Z_i' \mathbb{1}\{Z_i' \leq n\delta\} \big) > \frac{n\epsilon}{12} \right)
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
& \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(d_{J_1}(\tilde{S}_n, \tilde{J}_n^k) > \epsilon \right)}{r(\log n)} \\
\leq & \max\Bigg\{ \limsup_{n\to\infty} \frac{\log \mathbb{P}\left(V_{(k)} \leq \tilde{Q}(n\delta) \right)}{r(\log n)}, \\
& \limsup_{n\to\infty} \frac{\log 3 \max_{j=1,\cdots,n} \mathbb{P}\left(\sum_{i=1}^{j} \left(Z_i' \mathbb{1}\{Z_i' \leq n\delta\} - EZ \right) > \frac{n\epsilon}{12} \right)}{r(\log n)}, \\
& \limsup_{n\to\infty} \frac{\log 3 \max_{j=1,\cdots,n} \mathbb{P}\left(\sum_{i=1}^{j} \left(EZ - Z_i' \mathbb{1}\{Z_i' \leq n\delta\} \right) > \frac{n\epsilon}{12} \right)}{r(\log n)} \Bigg\} \qquad (4.78) \\
= & \max\{-k, -\frac{\epsilon}{36\delta}, -\infty\}. \qquad (4.79)
\end{aligned}
$$

To achieve (4.79) in the above, we have used the limit result (4.93) along with $\tilde{Q}(x) = \mathbb{P}\,(Z > x) = \exp\{-r(\log x)\}$, Lemma 4.6 and Lemma 4.7 respectively to bound the three limsup terms in the (4.78). We can choose arbitrary small δ to make the value in (4.79) to $-k$, and letting $k \to \infty$ confirms the condition (2.5). This finishes the proof. □

4.4 Proofs for Section 3.3

We start with a lemma that will be useful.

Lemma 4.11. *If the extended large deviation upper bounds hold for X_n and Y_n with rate functions I and J, respectively, and suppose that*

$$F \triangleq \bigcap_{m=1}^{M} (C_m \times \mathcal{Y}) \cup (\mathcal{X} \times D_m) \qquad (4.80)$$

where M is a finite integer, and $C_n \subseteq \mathcal{X}$ and $D_n \subseteq \mathcal{Y}$ are not necessarily closed sets. Suppose that X_n and Y_n satisfy the extended LDP with the rate functions I and J. Then the extended LDP upper bound w.r.t. F holds for (X_n, Y_n) with the rate function $K(x, y) = I(x) + J(y)$, i.e.,

$$\limsup_{n\to\infty} a_n^{-1} \mathbb{P}\big((X_n, Y_n) \in F\big) \leq -\lim_{\epsilon\to 0} \inf_{(x,y)\in F^\epsilon} K(x, y)$$

Proof of Lemma 4.11. Note that

$$F = \bigcup_{\ell\in\{0,1\}^M} \Big(\bigcap_{m:\ell_m=0} C_m \times \bigcap_{m:\ell_m=1} D_m \Big). \qquad (4.81)$$

Therefore,

$$
\begin{aligned}
&\limsup_{n\to\infty} a_n^{-1}\log \mathbb{P}\big((X_n,Y_n)\in F\big)\\
&=\limsup_{n\to\infty} a_n^{-1}\log \sum_{\ell\in\{0,1\}^M} \mathbb{P}\big(X_n\in \bigcap_{m:\ell_m=0} C_m\big)\cdot \mathbb{P}\big(Y_n\in \bigcap_{m:\ell_m=1} D_m\big)\\
&\le \sup_{\ell\in\{0,1\}^N}\left\{\limsup_{n\to\infty} a_n^{-1}\log \mathbb{P}\big(X_n\in \bigcap_{m:\ell_m=0} C_m\big)\cdot \mathbb{P}\big(Y_n\in \bigcap_{m:\ell_m=0} D_m\big)\right\}\\
&= \sup_{\ell\in\{0,1\}^N}\left\{-\lim_{\epsilon\to 0}\inf_{x\in\left(\bigcap_{m:\ell_m=0} C_m\right)^\epsilon} I(x)-\lim_{\epsilon\to 0}\inf_{x\in\left(\bigcap_{m:\ell_m=1} D_m\right)^\epsilon} J(x)\right\}\\
&=-\lim_{\epsilon\to 0}\inf_{\ell\in\{0,1\}^N}\inf_{(x,y)\in\left(\bigcap_{m:\ell_m=0} C_m\right)^\epsilon\times\left(\bigcap_{m:\ell_m=1} D_m\right)^\epsilon} I(x)+J(x)\\
&=-\lim_{\epsilon\to 0}\inf_{(x,y)\in F^\epsilon} I(x)+J(y)
\end{aligned}
$$

□

Now we are ready to prove Proposition 3.8.

Proof of Proposition 3.8. By induction, it is enough to show this for $d=2$. To ease the notation, we will denote $X_n^{(1)}$ with X_n, $X_n^{(2)}$ with Y_n, $\mathcal{X}^{(1)}$ with $\mathcal{X}$, $\mathcal{X}^{(2)}$ with $\mathcal{Y}$, $I^{(1)}$ with I, $I^{(2)}$ with J, and $\bar{I}$ with K.

For the lower semicontinuity of K, note that the sublevel set of K can be written as follows:

$$
\begin{aligned}
\Phi_K(\alpha) &\triangleq \{(x,y)\in\mathcal{X}\times\mathcal{Y}: K(x,y)\le\alpha\}\\
&=\bigcap_{n=1}^{\infty}\bigcup_{m=0}^{n+1}\underbrace{\left\{x\in\mathcal{X}: I(x)\le\frac{m}{n}\alpha\right\}}_{\Psi_I\left(\frac{m}{n}\alpha\right)}\times\underbrace{\left\{y\in\mathcal{Y}: J(y)\le\frac{n+1-m}{n}\alpha\right\}}_{\Psi_J\left(\frac{n+1-m}{n}\alpha\right)},
\end{aligned}
$$

which is closed since I and J are lower semicontinuous, and each of the unions involve only finite number of closed sets.

The proof of the lower bound of the extended LDP is identical to that of the lower bound in Theorem 4.14 in [17], but we still provide it here for the sake of completeness. Given an open set $G\subseteq\mathcal{X}\times\mathcal{Y}$ and its element $(x,y)\in G$, there exist open neighborhoods $G_\mathcal{X}\subseteq\mathcal{X}$ and $G_\mathcal{Y}\subseteq\mathcal{Y}$ of x and y such that $G_\mathcal{X}\times G_\mathcal{Y}\subseteq G$.

$$
\begin{aligned}
\liminf_{n\to\infty} a_n^{-1}\log\mathbb{P}\big((X_n,Y_n)\in G\big) &\ge \liminf_{n\to\infty} a_n^{-1}\log\mathbb{P}(X_n\in G_\mathcal{X})\cdot\mathbb{P}(Y_n\in G_\mathcal{Y})\\
&=\liminf_{n\to\infty} a_n^{-1}\log\mathbb{P}(X_n\in G_\mathcal{X})+\liminf_{n\to\infty} a_n^{-1}\log\mathbb{P}(Y_n\in G_\mathcal{Y})\\
&\ge-\inf_{x'\in G_\mathcal{X}} I(x')-\inf_{y'\in G_\mathcal{Y}} I(y')\\
&=-K(x,y)
\end{aligned}
$$

Taking infimum over $(x,y)\in G$, we arrive at the desired lower bound.

For the upper bound of the extended LDP, we introduce a few extra notations. Let $\mathcal{I}\triangleq\{I(x): x\in\mathcal{X}\}$, $\mathcal{J}\triangleq\{J(x): x\in\mathcal{Y}\}$, and $\mathcal{K}\triangleq\{K(x,y): x\in\mathcal{X}, y\in\mathcal{Y}\}$ denote the ranges of I, J, and K. Fix an $F\subseteq\mathcal{X}\times\mathcal{Y}$ and set

$$
k'\triangleq\lim_{\epsilon\to 0}\inf_{(x,y)\in F^\epsilon} K(x,y) \tag{4.82}
$$

If $k = 0$, then the extended LDP upper bound is trivially satisfied, and hence, we assume w.l.o.g. $k > 0$. From the assumption (iii), there exists the largest element $k \in \mathcal{K}$ such that $k < k'$. Note also that

$$\Psi_K(k) = \bigcup_{(i,j)\in B(k)} \Psi_I(i) \times \Psi_J(j)$$

where $B(k) = \{(i,j) \in \mathcal{I} \times \mathcal{J} : i + j \le k\}$. Note also that (4.82) implies that there exists $\delta > 0$ such that $d(F, \Psi_K(k)) > \delta$. This, in turn, implies that

$$F \subseteq \big(\mathcal{X} \times \mathcal{Y}\big) \setminus \big(\Psi_K(k)\big)^\delta = \bigcap_{(i,j)\in B(k)} \Big(\big(\mathcal{X} \setminus \Psi_I(i)^\delta\big) \times \mathcal{Y}\Big) \cup \Big(\mathcal{X} \times \big(\mathcal{Y} \setminus \Psi_J(j)^\delta\big)\Big).$$

From this and Lemma 4.11, we get

$$\begin{aligned} \liminf_{n\to\infty} a_n^{-1} \log \mathbb{P}\big((X_n, Y_n) \in F\big) &\le \liminf_{n\to\infty} a_n^{-1} \log \mathbb{P}\big((X_n, Y_n) \in \big(\mathcal{X} \times \mathcal{Y}\big) \setminus \big(\Psi_K(k)\big)^\delta\big) \\ &\le -\lim_{\epsilon\to 0} \inf_{(x,y)\in\big((\mathcal{X}\times\mathcal{Y})\setminus(\Psi_K(k))^\delta\big)^\epsilon} K(x,y) < -k. \end{aligned}$$

Since $-k \le -k' = -\lim_{\epsilon\to 0} \inf_{(x,y)\in F^\epsilon} K(x,y)$, this implies the extended LDP upper bound:

$$\liminf_{n\to\infty} a_n^{-1} \log \mathbb{P}\big((X_n, Y_n) \in F\big) \le -\lim_{\epsilon\to 0} \inf_{(x,y)\in F^\epsilon} K(x,y).$$

□

4.5 Proofs for Section 3.4

Proofs of Lemma 3.12. Suppose that $\eta \in \hat{\mathbb{D}}_{\leqslant 1}$. It is sufficient to show that $\eta \notin \bar{F}$. We start with the following observations:

(1) If $\xi \in F_n$, then $\xi(\frac{1}{2}) \ge \log n - \frac{1}{2}\nu_1\mu_1 - \frac{1}{3}n^{-\frac{1}{3}}$. In particular, $\xi(\frac{1}{2}) > 1$ for any $\xi \in F$.

(2) If $\xi \in F_n$, then ξ has a jump between $(\frac{3}{4}, 1]$ of size no smaller than $\frac{1}{3}n^{-\frac{1}{3}}$

Note that since (1) implies that F_n's are bounded away from the zero function. Therefore, we will focus on the case $\eta \not\equiv 0$ w.l.o.g. This allows us to write $\eta = z\mathbb{1}_{[v,1]}$ for $z > 0$ and $v \in [0,1]$. We conclude the proof by considering the two possible cases—$v \in (1/2, 1]$ and $v \in [0, 1/2]$—separately:

Suppose that $v \in (1/2, 1]$. For any $\xi \in F$, set $z \triangleq (\xi(1/2), 1/2) \in \Gamma(\xi)$. From (1), we can easily check $d\big(z, \Gamma(\eta)\big) \ge v - 1/2$. From this and Lemma 2.1, we have $d_{M_1'}(\eta, \xi) \ge v - 1/2$. Since this is true for any $\xi \in F$, we conclude that $\eta \notin \bar{F}$.

Suppose that $v \in [0, 1/2]$. We will proceed with proof by contradiction. Suppose that $\eta \in \bar{F}$ so that there exists a sequence $\xi_n \in F$ such that $d_{M_1'}(\xi_n, \eta) \to 0$ as $n \to \infty$. We claim that $\xi_n \in F_{m_n}$ for some m_n such that $m_n \to \infty$ as $n \to \infty$. To see this, recall (2) and the fact that η is constant between 1/2 and 1. Therefore, Lemma 2.2 implies that $d_{M_1'}(\xi_n, \eta)$ is bounded away from 0 if the claim doesn't hold. On the other hand,

$$d_{M_1'}(\xi_n, \eta) \ge d\big((\xi_n(1/2), 1/2), \Gamma(\eta)\big) \ge |\xi_n(1/2) - z| \ge \log m_n - \frac{1}{2}\nu_1\mu_1 - \frac{1}{3}m_n^{-\frac{1}{3}} - z \to \infty,$$

where the second inequality is from Lemma 2.1 and the third inequality is from the claim. This is contradictory to our earlier assumption, and hence, η cannot be in $\bar{F}$.

□

Proof of Lemma 3.13. Based on the Lévy-Ito decomposition (3.1), $\bar{X}_n$ has the following distribution representation:

$$\bar{X}_n \stackrel{\mathcal{D}}{=} \bar{J}_n + \bar{H}_n \tag{4.83}$$

where

$$
\bar{J}_n(t) \triangleq \frac{1}{n}\int_{[1,\infty)} x\hat{N}([0,nt]\times dx) \tag{4.84}
$$

$$
\bar{H}_n(t) \triangleq \frac{aB(nt)}{n} + \frac{1}{n}\int_{(0,1]} x\big(\hat{N}([0,nt]\times dx) - nt\nu(dx)\big) - t\int_{[1,\infty)} x\nu(dx) \tag{4.85}
$$

Since $\bar{J}_n$ and $\bar{H}_n$ are independent, we have

$$
\begin{aligned}
\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\bar{X}_n\in F\right)}{r(\log n)} &\geq \limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\bar{X}_n\in F_n\right)}{r(\log n)} \geq \limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\bar{J}_n\in B_n,\bar{H}_n\in C_n\right)}{r(\log n)} \\
&\geq \limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\bar{J}_n\in B_n\right)+\log\mathbb{P}\left(\bar{H}_n\in C_n\right)}{r(\log n)} \\
&= \underbrace{\limsup_{n\to\infty}\frac{\log\mathbb{P}\left(\bar{J}_n\in B_n\right)}{r(\log n)}}_{\text{(I)}} + \underbrace{\lim_{n\to\infty}\frac{\log\mathbb{P}\left(\bar{H}_n\in C_n\right)}{r(\log n)}}_{\text{(II)}}.
\end{aligned} \tag{4.86}
$$

Due to Lemma 4.12, we have (II)$=0$. To evaluate (I), recall $Q_n^{\leftarrow}(\cdot)$'s definition in (4.7) and $\hat{N}$'s representation in (4.8). $\bar{J}_n$ has the following representation.

$$
\bar{J}_n(t) \stackrel{d}{=} \sum_{i=1}^{\infty} Q_n^{\leftarrow}(\Gamma_i)\mathbb{1}_{[1,\infty)}(Q_n^{\leftarrow}(\Gamma_i))\mathbb{1}_{[U_i,1]}(t)
$$

Here $\Gamma_i = E_1+\cdots+E_i$ and E_i's are i.i.d. Exp(1) variables. From the definitions of A_n and B_n's in (3.13) and (3.14),

$$
\begin{aligned}
&\{E_1\in\big(Q_n(n\cdot 2\log n),Q_n(n\log n)\big],E_2\in\big[0,Q_n(n\cdot n^{-\frac{1}{3}})-Q_n(n\log n)\big),\\
&\quad U_1\in(\frac{1}{4},\frac{1}{2}],U_2\in(\frac{3}{4},1]\}\\
&\subset\{E_1\in\big(Q_n(n\cdot 2\log n),Q_n(n\log n)\big],E_1+E_2\in\big[0,Q_n(n\cdot n^{-\frac{1}{3}})\big],\\
&\quad U_1\in(\frac{1}{4},\frac{1}{2}],U_2\in(\frac{3}{4},1]\}\\
&\subset\{\frac{Q_n^{\leftarrow}(E_1)}{n}\in[\log n,2\log n)\text{ and }\frac{Q_n^{\leftarrow}(E_1+E_2)}{n}\in[n^{-\frac{1}{3}},\infty),U_1\in(\frac{1}{4},\frac{1}{2}],U_2\in(\frac{3}{4},1]\}\\
&\subset\{\pi(\bar{J}_n)\in A_n\}=\{\bar{J}_n\in\pi^{-1}(A_n)=B_n\}.
\end{aligned}
$$

where the second inclusion is due the fact that $c\in\big(Q_n(b),Q_n(a)\big]$ if and only if $Q_n^{\leftarrow}(c)\in$

$[a, b)$ for any c and $b > a \geq 0$. Therefore, the term (I) by can be lower-bounded as below:

$$
\begin{aligned}
&\limsup_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{J}_n \in B_n\right)}{r(\log n)} \\
&\geq \limsup_{n\to\infty} \\
&\quad \frac{\log \mathbb{P}\left(E_1 \in (Q_n(n\cdot 2\log n), Q_n(n\log n)]\right) + \log \mathbb{P}\Big(E_2 \in [0, Q_n(n\cdot n^{-\frac{1}{3}}) - Q_n(n\log n))\Big)}{r(\log n)} \\
&= \limsup_{n\to\infty} \frac{1}{r(\log n)}\Big[\ \log \int_{Q_n(n\cdot 2\log n)}^{Q_n(n\log n)} e^{-y_1} dy_1 + \log \int_0^{Q_n(n^{\frac{2}{3}})-Q_n(n\log n)} e^{-y_2} dy_2\Big] \\
&\geq -\underbrace{\lim_{n\to\infty} \frac{Q_n(n\log n)}{r(\log n)}}_{\text{(III)}} + \underbrace{\lim_{n\to\infty} \frac{\log\big(Q_n(n\log n) - Q_n(n\cdot 2\log n)\big)}{r(\log n)}}_{\text{(IV)}} \\
&\quad + \underbrace{\lim_{n\to\infty} \frac{\log\big(1 - e^{-Q_n(n^{\frac{2}{3}})+Q_n(n\log n)}\big)}{r(\log n)}}_{\text{(V)}}.
\end{aligned}
$$

Now we are left with identifying (III), (IV), and (V).

Recall $Q_n(x) = n\nu[x, \infty) = n\exp(-r(\log x))$. For (III), we have

$$
\text{(III)} = \lim_{n\to\infty} \frac{Q_n(n\log n)}{r(\log n)} = \lim_{n\to\infty} \frac{n\exp\big(-r(\log n + \log\log n)\big)}{r(\log n)} = 0
$$

Turning to (IV),

$$
\begin{aligned}
\text{(IV)} &= \lim_{n\to\infty} \frac{\log n + \log\Big(\exp\big(-r(\log n + \log\log n)\big) - \exp\big(-r(\log n + \log\log n + \log 2)\big)\Big)}{r(\log n)} \\
&= 0 - \lim_{n\to\infty} \frac{r(\log n + \log\log n)}{r(\log n)} \\
&\quad + \lim_{x\to\infty} \frac{\log\Big(1 - \exp\big(r(x) - r(x + \log 2)\big)\Big)}{r(x)} \cdot \lim_{n\to\infty} \frac{r(\log n + \log\log n)}{r(\log n)} \\
&= -1,
\end{aligned}
$$

where the last equality is due to (4.87) and (4.90). Lastly, for (V), we have

$$
\begin{aligned}
\text{(V)} &= \lim_{n\to\infty} \frac{\log n + \log\big(\exp(-r(\frac{2}{3}\log n)) - \exp(-r(\log n + \log\log n)\big)}{r(\log n)} \\
&= 0 - \lim_{n\to\infty} \frac{r(\frac{2}{3}\log n)}{r(\log n)} + \lim_{n\to\infty} \frac{\log\Big(1 - \exp\big(r(\frac{2}{3}\log n) - r(\log n + \log\log n)\big)\Big)}{r(\log n)} \\
&= -(\frac{2}{3})^{\gamma}.
\end{aligned}
$$

where the last equality is due to $r(\frac{2}{3}\log n) - r(\log n + \log\log n) \leq r(\frac{2}{3}\log n) - r(\log n) \sim \big((\frac{2}{3})^{\gamma} - 1\big) r(\log n) \to -\infty$ when $n \to \infty$. Combining the value of (III), (IV), and (V) directly yield the conclusion. □

Lemma 4.12. *For C_n in (3.15) and $\bar{H}_n$ in (4.85),*

$$
\lim_{n\to\infty} \frac{\log \mathbb{P}\left(\bar{H}_n \in C_n\right)}{r(\log n)} = 0
$$

Proof of Lemma 4.12. It is enough to shows that $\mathbb{P}\left(\bar{H}_n \notin C_n\right) \to 0$ as $n \to \infty$.

$$\begin{aligned}
\mathbb{P}\left(\bar{H}_n \notin C_n\right) &= \mathbb{P}\left(\sup_{t\in[0,1]} |\bar{H}_n(t) + t\nu_1\mu_1| > \frac{1}{3}\frac{1}{n^{\frac{1}{3}}}\right) \\
&= \mathbb{P}\left(\sup_{t\in[0,1]} \left|\frac{aB(nt)}{n} + \frac{1}{n}\int_{(0,1]} x\big(\hat{N}([0,nt]\times dx) - nt\nu(dx)\big)\right| > \frac{1}{3}\frac{1}{n^{\frac{1}{3}}}\right) \\
&= \mathbb{P}\left(\sup_{t\in[0,1]} \left(aB(nt) + \int_{(0,1]} x\big(\hat{N}([0,nt]\times dx) - nt\nu(dx)\big)\right)^2 > \frac{1}{9}n^{\frac{4}{3}}\right) \\
&\le \frac{\mathbb{E}\left[\left(aB(n) + \int_{(0,1]} x\big(\hat{N}([0,n]\times dx) - nt\nu(dx)\big)\right)^2\right]}{\frac{1}{9}n^{\frac{4}{3}}} \\
&= \frac{n\mathbb{E}\left[\left(aB(1) + \int_{(0,1]} x\big(\hat{N}([0,1]\times dx) - t\nu(dx)\big)\right)^2\right]}{\frac{1}{9}n^{\frac{4}{3}}} = \frac{\text{Const}}{n^{\frac{1}{3}}} \to 0,
\end{aligned}$$

where the inequality is due to the Doob's submartingale inequality. ☐

4.6 Supporting lemmata

4.6.1 Asymptotic limits associated with regularly varying $r(\cdot)$

Recall Γ_i is the summation of i independent Exp(1) random variables. Also, $V_{(i)}$ represents the i-th order statistic among n independent Uniform[0,1] random variables.

Lemma 4.13. *Assume $r(\cdot)$ is a nondecreasing regularly varying function with index $\gamma > 1$. In addition, assume there exists some $\gamma' \in (0,\gamma)$ such that*

$$\liminf_{x\to\infty} \frac{r(x+c) - r(x)}{\exp\{-x^{\gamma'}\}} \ge 1$$

for any $c > 0$. Then the following asymptotics hold:

$$\lim_{x\to\infty} \frac{\log\big(1 - \exp\{r(x) - r(x+c)\}\big)}{r(x)} = 0 \text{ for any } c > 0 \tag{4.87}$$

$$\lim_{n\to\infty} \frac{\log\Big(\exp\big(-r(\log nx_1)\big) - \exp\big(-r(\log nx_2)\big)\Big)}{r(\log n)} = -1 \text{ for } 0 < x_1 < x_2 \tag{4.88}$$

Proof of Lemma 4.13. For (4.87),

$$\begin{aligned}
0 &\ge \lim_{x\to\infty} \frac{\log\big(1 - \exp(r(x) - r(x+c))\big)}{r(x)} \\
&\ge \lim_{x\to\infty} \frac{\log\big(1 - \exp(-\exp(-x^{\gamma'}))\big)}{r(x)} = \lim_{x\to\infty} \frac{-x^{\gamma'}}{x^{\gamma}L(x)} = 0,
\end{aligned}$$

where the last inequality above has leveraged $e^y \sim y+1$ when y close to 0.

Turing to (4.88),

$$\begin{aligned}
&\lim_{n\to\infty} \frac{\log\big(\exp(-r(\log nx_1)) - \exp(-r(\log nx_2))\big)}{r(\log n)} \\
&= \lim_{n\to\infty} \frac{\log\Big(\exp\big(-r(\log nx_1)\big)\cdot\big(1-\exp\big\{r(\log nx_1) - r(\log nx_2)\big\}\big)\Big)}{r(\log n)} \\
&= \lim_{n\to\infty} \frac{\log\big(\exp\big(-r(\log nx_1)\big)\big)}{r(\log n)} \\
&+ \lim_{n\to\infty} \frac{\log\big(1-\exp\big\{r(\log n + \log x_1) - r(\log n + \log x_1 + \log x_2 - \log x_1)\big\}\big)}{r(\log n + \log x_1)} \\
&\quad \cdot \frac{r(\log n + \log x_1)}{r(\log n)} \\
&= -1
\end{aligned}$$

Note that the last equality holds due to (4.87). ☐

Lemma 4.14. *Assume $r(\cdot)$ is a regularly varying function with index $\gamma > 1$. The following asymptotics hold:*

$$\lim_{n\to\infty} n\exp\big(-r(\log nx)\big) = 0 \text{ for } x > 0 \tag{4.89}$$

$$\lim_{x\to\infty} \frac{r(x+o(x))}{r(x)} = 1 \tag{4.90}$$

$$\lim_{n\to\infty} \frac{n\log\Big(1-\exp\big(-r(\log nx)\big)\Big)}{r(\log n)} = 0 \text{ for } x > 0 \tag{4.91}$$

$$\lim_{n\to\infty} \frac{\log \mathbb{P}\big(\Gamma_i \le n\exp\big(-r(\log nx)\big)\big)}{r(\log n)} = -i \text{ for } x > 0 \tag{4.92}$$

$$\lim_{n\to\infty} \frac{\log \mathbb{P}\left(V_{(i)} \le \exp\big(-r(\log nx)\big)\right)}{r(\log n)} \le -i \text{ for } x > 0 \tag{4.93}$$

Proof of Lemma 4.14. The first three (4.89), (4.90) and (4.91) are trivial.

For (4.92), recall that Γ_i follows the Erlang-i distribution, and hence,

$$\lim_{n\to\infty} \frac{\log \mathbb{P}\left(\Gamma_i \le n\exp\big(-r(\log nx)\big)\right)}{r(\log n)} = \lim_{n\to\infty} \frac{\log\left(\int_0^{n\exp\left(-r(\log nx)\right)} s^{i-1}e^{-s}ds\right)}{r(\log n)}. \tag{4.94}$$

Note that for $y > 0$ small enough,

$$\frac{1}{2i}y^i \le \frac{1}{2}\int_0^y s^{i-1}ds \le \int_0^y s^{i-1}e^{-s}ds \le \int_0^y s^{i-1}ds \le \frac{1}{i}y^i$$

Combining this with (4.94), we arrive at:

$$\lim_{n\to\infty} \frac{\log\left(\int_0^{n\exp\left(-r(\log nx)\right)} s^{i-1}e^{-s}ds\right)}{r(\log n)} = \lim_{n\to\infty} \frac{i\Big(\log n - r(\log nx)\Big)}{r(\log n)} = -i$$

where last equality is due to (4.90). For the last result (4.93), recall $V_{(i)}$ has $Beta(i, n-i)$ distribution, hence

$$
\begin{aligned}
&\lim_{n\to\infty} \frac{\log \mathbb{P}\left(V_{(i+1)} \le \exp\left(-r(\log nx)\right)\right)}{r(\log n)} \\
&= \limsup_{n\to\infty} \frac{\log \int_0^{\exp\left(-r(\log nx)\right)} \frac{\Gamma(n-1)}{\Gamma(i)\Gamma(n-i)} y^{i-1}(1-y)^{n-i-1} dy}{r(\log n)} \\
&\le \limsup_{n\to\infty} \frac{\log \int_0^{\exp\left(-r(\log nx)\right)} \frac{\Gamma(n-1)}{\Gamma(i)\Gamma(n-i)} y^{i-1} dy}{r(\log n)} \\
&= \limsup_{n\to\infty} \frac{\log \left(\frac{\Gamma(n-1)}{\Gamma(i)\Gamma(n-i)i}\left(\exp\left(-r(\log nx)\right)\right)^i\right)}{r(\log n)} \\
&= \lim_{n\to\infty} \frac{-i \cdot r(\log nx)}{r(\log n)} = -i
\end{aligned}
$$

where the last line uese (4.90). This finishes the proof. □

4.6.2 Basic concentration inequalities

This section reviews two well-known concentration inequalities. For reference, see Theorem 22.5 in [4] and Section 2.7 in [8], respectively.

Lemma 4.15. *[Etemadi's inequality] Let $X_1, \cdots, X_n$ be independent real-valued random variables defined on some common probability space, and let $x \ge 0$. Let S_k denote the partial sum $S_k = X_1 + \cdots + X_k$. Then*

$$
\mathbb{P}\left(\max_{1\le k\le n} |S_k| \ge 3x\right) \le 3 \max_{1\le k\le n} \mathbb{P}\left(|S_k| \ge x\right)
$$

Lemma 4.16. *[Bernstein's inequality] Let $X_1, \cdots, X_n$ be independent real-valued random variables with finite variance such that $X_i \le b$ for some $b > 0$ almost surely for all $i \le n$. Let $S = \sum_{i=1}^n (X_i - \mathbb{E}[X_i])$ and $v = \sum_{i=1}^n \mathbb{E}\left[X_i^2\right]$, then for any $t > 0$*

$$
\mathbb{P}\left(S \ge t\right) \le \exp\left\{-\frac{t^2}{2(v + bt/3)}\right\}
$$

References

[1] Mihail Bazhba, Jose Blanchet, Chang-Han Rhee, and Bert Zwart. Queue length asymptotics for the multiple-server queue with heavy-tailed Weibull service times. *Queueing Syst.*, 93(3-4):195–226, 2019.

[2] Mihail Bazhba, Jose Blanchet, Chang-Han Rhee, and Bert Zwart. Sample path large deviations for Lévy processes and random walks with Weibull increments. *Ann. Appl. Probab.*, 30(6):2695–2739, 2020.

[3] Mihail Bazhba, Chang-Han Rhee, and Bert Zwart. Large deviations for stochastic fluid networks with Weibullian tails. *Queueing Syst.*, 102(1-2):25–52, 2022.

[4] Patrick Billingsley. *Probability and measure*. Wiley Series in Probability and Statistics. John Wiley & Sons, Inc., Hoboken, NJ, anniversary edition, 2012. With a foreword by Steve Lalley and a brief biography of Billingsley by Steve Koppes.

[5] Jose Blanchet and Yixi Shi. Measuring systemic risks in insurance-reinsurance networks. *Preprint*, 2012.

[6] A. A. Borovkov and K. A. Borovkov. *Asymptotic analysis of random walks*, volume 118 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 2008. Heavy-tailed distributions, Translated from the Russian by O. B. Borovkova.

[7] A. A. Borovkov and A. A. Mogulskii. On large deviation principles in metric spaces. *Sibirsk. Mat. Zh.*, 51(6):1251–1269, 2010.

[8] Stéphane Boucheron, Gábor Lugosi, and Pascal Massart. *Concentration inequalities*. Oxford University Press, Oxford, 2013. A nonasymptotic theory of independence, With a foreword by Michel Ledoux.

[9] Amir Dembo and Ofer Zeitouni. *Large deviations techniques and applications*, volume 38 of *Stochastic Modelling and Applied Probability*. Springer-Verlag, Berlin, 2010. Corrected reprint of the second (1998) edition.

[10] D. Denisov, A. B. Dieker, and V. Shneer. Large deviations for random walks under subexponentiality: the big-jump domain. *Ann. Probab.*, 36(5):1946–1991, 2008.

[11] Jean-Dominique Deuschel and Daniel W. Stroock. *Large deviations*, volume 137 of *Pure and Applied Mathematics*. Academic Press, Inc., Boston, MA, 1989.

[12] Paul Dupuis and Richard S. Ellis. *A weak convergence approach to the theory of large deviations*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, 1997. A Wiley-Interscience Publication.

[13] Richard S. Ellis. *Entropy, large deviations, and statistical mechanics*. Classics in Mathematics. Springer-Verlag, Berlin, 2006. Reprint of the 1985 original.

[14] Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. *Modelling extremal events*, volume 33 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1997. For insurance and finance.

[15] Jin Feng and Thomas G. Kurtz. *Large deviations for stochastic processes*, volume 131 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2006.

[16] Sergey Foss and Dmitry Korshunov. On large delays in multi-server queues with heavy tails. *Math. Oper. Res.*, 37(2):201–218, 2012.

[17] Ayalvadi Ganesh, Neil O'Connell, and Damon Wischik. *Big queues*, volume 1838 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2004.

[18] Henrik Hult, Filip Lindskog, Thomas Mikosch, and Gennady Samorodnitsky. Functional large deviations for multivariate regularly varying random walks. *Ann. Appl. Probab.*, 15(4):2651–2680, 2005.

[19] Filip Lindskog, Sidney I. Resnick, and Joyjit Roy. Regularly varying measures on metric spaces: hidden regular variation and hidden jumps. *Probab. Surv.*, 11:270–314, 2014.

[20] P. Massart. The tight constant in the Dvoretzky-Kiefer-Wolfowitz inequality. *Ann. Probab.*, 18(3):1269–1283, 1990.

[21] A. A. Mogulskii. Large deviations for the trajectories of multidimensional random walks. *Teor. Verojatnost. i Primenen.*, 21(2):309–323, 1976.

[22] A. A. Mogulskii. Large deviations for processes with independent increments. *Ann. Probab.*, 21(1):202–215, 1993.

[23] A. V. Nagaev. Limit theorems that take into account large deviations when Cramér's condition is violated. *Izv. Akad. Nauk UzSSR Ser. Fiz.-Mat. Nauk*, 13(6):17–22, 1969.

[24] A. V. Nagaev. A property of sums of independent random variables. *Teor. Verojatnost. i Primenen.*, 22(2):335–346, 1977.

[25] Anatolii A. Puhalskii and Ward Whitt. Functional large deviation principles for first-passage-time processes. *Ann. Appl. Probab.*, 7(2):362–381, 1997.

[26] Sidney I. Resnick. *Heavy-tail phenomena*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2007. Probabilistic and statistical modeling.

[27] Chang-Han Rhee, Jose Blanchet, and Bert Zwart. Sample path large deviations for Lévy processes and random walks with regularly varying increments. *Ann. Probab.*, 47(6):3551–3605, 2019.

[28] Ken-iti Sato. *Lévy processes and infinitely divisible distributions*, volume 68 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2013. Translated from the 1990 Japanese original, Revised edition of the 1999 English translation.

[29] Adam Shwartz and Alan Weiss. *Large deviations for performance analysis*. Stochastic Modeling Series. Chapman & Hall, London, 1995. Queues, communications, and computing, With an appendix by Robert J. Vanderbei.

[30] S. R. S. Varadhan. Asymptotic probabilities and differential equations. *Comm. Pure Appl. Math.*, 19:261–286, 1966.

[31] S. R. S. Varadhan. *Large deviations and applications*, volume 46 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1984.

[32] Ward Whitt. *Stochastic-process limits*. Springer Series in Operations Research. Springer-Verlag, New York, 2002. An introduction to stochastic-process limits and their application to queues.

[33] Bert Zwart, Sem Borst, and Michel Mandjes. Exact asymptotics for fluid queues fed by multiple heavy-tailed on-off flows. *Ann. Appl. Probab.*, 14(2):903–957, 2004.

Acknowledgments. This work was supported by the National Science Foundation under Grant Number CMMI-2146530. Both authors are grateful for insightful discussions with Bert Zwart.