Heavy-tailed random walks, buffered queues and hidden large deviations

HARALD BERNHARD[*] and BIKRAMJIT DAS[**]

Singapore University of Technology and Design, Pillar of Engineering Systems and Design, 8 Somapah Road, Singapore 487372. E-mail: [*]bernhard.harald@gmail.com; [**]bikram@sutd.edu.sg

It is well-known that large deviations of random walks driven by independent and identically distributed heavy-tailed random variables are governed by the so-called principle of one large jump. We note that further subtleties hold for such random walks in the large deviations scale which we call hidden large deviations. Our results are illustrated using two examples. First, we apply this idea in the context of queueing processes with heavy-tailed service times and study approximations of probabilities of severe congestion times for (buffered) queues. We exhibit our techniques by using limit measures from different large deviation regimes to provide a unified estimate of rare event probabilities in a simulated queue. Furthermore, we use our result to provide probability estimates of rare events governed by more than one jump in case the innovations of a random walk have infinite mean.

Keywords: buffered queues; heavy-tails; large deviations; regular variation

1. Introduction

Stochastic processes with heavy-tailed components as building blocks are of interest in many areas of application, including, but not restricted to hydrology (Anderson and Meerschaert [2]), finance, insurance and risk management (Smith [37], Embrechts, Klüppelberg and Mikosch [18], Ibragimov, Jaffee and Walden [23]), tele-traffic data (Crovella, Bestavros and Taqqu [11]), queueing theory (Boxma and Cohen [10], Zwart [40]), social networks and random graphs (Bollobás *et al.* [9], Durrett [17]). The notion of heavy-tails in applied probability is often studied under the paradigm of regular variation. In this paper, we concentrate on investigating subtle properties of heavy-tailed random walks which helps us in understanding intricate underlying structures of queueing models with heavy-tailed service times under certain regularity conditions. In particular, we establish how queueing congestion (which we define in terms of long intense periods) may occur in a simple GI/G/1 queue, not only because of one large jump, but also in terms of further jumps occurring in the process. Moreover, our result also aids in computing probabilities of certain rare events in random walks with innovations having infinite mean.

It is well known that if $\{Z_i\}_{i \geq 1}$ are independent and identically distributed (i.i.d.) zero mean regularly varying random variables, then a large deviation in the partial sum $S_n = \sum_{i=1}^n Z_i$ is most likely to be the result of one of the random variables Z_i attaining a very large value, see Embrechts, Klüppelberg and Mikosch [18], Section 8.6, for further details. Early results on this notion popularly known as the *principle of one large jump* were obtained in Nagaev [31–33]. More formally, the notion of one large jump in this case can be written as

$$\mathbb{P}(|S_n| > x) = n\mathbb{P}(|Z_1| > x)(1 + o(1)), \quad x > b_n \tag{1}$$

for some choice of $b_n \uparrow \infty$ as $n \to \infty$. Similar large deviation results have been obtained in Denisov, Dieker and Shneer [16] under the more general assumption of the random variables being sub-exponential. In further generality, the notion of regular variation of càdlàg processes has been characterized in Hult and Lindskog [20]. It was aptly noted in Hult *et al.* [22] that large deviations for such processes with heavy-tailed margins are very closely related to the notion of regular variation.

A precise large deviation result for partial sum processes on the space $\mathbb{D} := \mathbb{D}([0, 1], \mathbb{R})$ of càdlàg functions was provided in Hult *et al.* [22], Theorem 2.1; in fact this result was obtained for d-dimensional processes. Let S^n denote the càdlàg embedding of $\{S_k\}_{k=1}^n$ into $\mathbb{D}$ with $S_0 = 0$, that is, $S^n = (S_{\lfloor nt \rfloor})_{t \in [0,1]}$, where $\lfloor x \rfloor$ denotes the largest integer no larger than x. In particular for $d = 1$, the authors establish that for suitably chosen sequences $\gamma_n > 0$ and $\lambda_n \uparrow \infty$ one may observe that

$$\gamma_n \mathbb{P}\big(S^n / \lambda_n \in \cdot\big) \xrightarrow{w^\#} \mu(\cdot), \quad n \to \infty, \tag{2}$$

for a non-null measure μ, where $\xrightarrow{w^\#}$ denotes convergence in the space of boundedly finite measures on $\overline{\mathbb{D}}_0$; see Daley and Vere-Jones [12], Appendix 2.6, and Hult *et al.* [22] for further details on the space and $w^\#$-convergence. In particular, the result shows that an appropriate choice of scaling is $\gamma_n = [n\mathbb{P}(Z_1 > \lambda_n)]^{-1}$ and the limit measure μ concentrates all its mass on step functions with exactly one jump discontinuity, which essentially retrieves the *one large jump principle*.

Hence, this indicates of a possibility, albeit rarer than the above case, that a large deviation of S_n may occur because two or more of the random variables $\{Z_i\}_{i=1}^n$ were large. The probabilities of such events, although negligible under the scaling $\gamma_n = [n\mathbb{P}(Z_1 > \lambda_n)]^{-1}$, are not exactly zero. In this paper, we aim to recover the rates at which such deviations happen and examine their structure. Furthermore, our goal is to use such results in the context of queueing to understand the behavior of what we call long intense periods in a large-deviation-type event. Analyses of hidden behavior of regularly varying sequences on $\mathbb{R}^d$ (Resnick [34], Das, Mitra and Resnick [13]) and more recently on $\mathbb{R}^\infty$ and Lévy processes on $\mathbb{D}$ (Lindskog, Resnick and Roy [28]) have been conducted under the name *hidden regular variation*.

In this paper, our first contribution is to extend the large deviation result in (2) to *hidden large deviations* in the spirit of hidden regular variation. We establish that the most probable way a large deviation event occurs, which is not the result of only one random variable being large, is actually when two random variables are large; resulting in a non-null limit measure as in (2) concentrating on processes having two jump discontinuities. For our analyses, we use the framework proposed in Lindskog, Resnick and Roy [28] and the notion of convergence used here is known as $\mathbb{M}_\mathbb{O}$-convergence which is closely related to the $w^\#$-convergence of boundedly finite measures and developed in Hult and Lindskog [21], Das, Mitra and Resnick [13], Lindskog, Resnick and Roy [28]. We recall in brief the required background on regular variation and $\mathbb{M}_\mathbb{O}$-convergence in Section 2. The results on hidden large deviations of random walks are dealt with in Section 3, where the key result is obtained in Theorem 3.5. In this theorem, we show that for each $j \geq 1$ and a suitably chosen sequence $\lambda_n \uparrow \infty$, there exists a non-zero limit measure

$\mu \in \mathbb{M}(\mathbb{D}\backslash\{\mathbb{D}_{\leq(j-1)}\})$ such that

$$\left[n\mathbb{P}\big(|Z_1| > \lambda_n\big)\right]^{-j}\mathbb{P}\big(S^n/\lambda_n \in \cdot\big) \to \mu(\cdot), \quad n \to \infty, \tag{3}$$

where $\mathbb{D}_{\leq j-1}$ denotes the set of all step functions in $\mathbb{D}$ with at most $(j-1)$ jumps and $\to$ denotes convergence in $\mathbb{M}(\mathbb{D}\backslash\{\mathbb{D}_{\leq(j-1)}\})$. The results in the literature most closely related to the results in this paper are Theorem 2.1 of Hult *et al.* [22], Theorem 5.1 of Lindskog, Resnick and Roy [28] and, most recently, Theorem 3.2 of Rhee, Blanchet and Zwart [36]. As mentioned above, the results in Hult *et al.* [22] include a version of (3) for the case of $j = 1$. Theorem 5.1 in Lindskog, Resnick and Roy [28] provides a result similar to (3) with S^n replaced by a Lévy process X on $\mathbb{D}$ with regularly varying Lévy measure ν which is assumed to concentrate on $(0, \infty)$. We paraphrase this result as

$$[\nu\big([\lambda_n, \infty)\big)]^{-j}\mathbb{P}(X/\lambda_n \in \cdot) \to \mu(\cdot), \quad n \to \infty. \tag{4}$$

Note that the time scale of X remains the same throughout. In their work, Rhee, Blanchet and Zwart [36] also study Lévy processes X with regularly varying Lévy measure ν, with both time and space rescaled by n, that is, $X^n = \{X(nt)/n, t \in [0, 1]\}$, and show that

$$[n\nu\big([n, \infty)\big)]^{-j}\mathbb{P}\big(X^n \in \cdot\big) \to \mu(\cdot), \quad n \to \infty. \tag{5}$$

Note that in (3) and (5) the rate term has an additional n (compared to (4)) to compensate for the larger time scales considered. In the latter paper, the authors further establish two-sided limit results for Lévy processes where the Lévy measure is regularly varying with potentially differing coefficients at positive and negative infinity (Rhee, Blanchet and Zwart [36], Theorem 3.5). Particularly, Rhee, Blanchet and Zwart [36] establishes (3) for $\lambda_n = n$ in Theorem 4.1. In comparison, our result allows for greater flexibility in the scaling function, that is, we assume $\lambda_n \in \mathcal{RV}_{-\rho}$ with $\rho > \frac{1}{2\wedge\alpha}$ while they consider the case $\lambda_n = n$ throughout. Their results are more general with respect to the processes considered, which are Lévy processes and random walks with possibly different coefficients of regular variation for positive (upward) and negative (downward) jumps.

We show an application of our results in understanding queueing congestion in a simple GI/G/1 queueing model. Queues with heavy-tailed service times have been of interest to researchers for a few years (Boxma and Cohen [10], Jelenković [26], Zwart [40]). Our interest lies in figuring out how often would we observe long busy or intense periods in a queue and how does it happen. In Zwart [40], the author shows that for a GI/G/1 queue with heavy-tailed service times, the most likely way a long busy period occurs is when one big service requirement arrives at the beginning of the busy period and the queue drifts back to zero linearly thereafter. Consequently, a large deviation of a queueing process also looks exactly the same. Jelenković [26] studies the steady state loss in buffered queues and shows that for large buffers K the steady state loss can be approximated by the expected loss due to one arrival A filling the buffer completely starting zero, that is the expected loss is approximately $\mathbb{E}[A - K]$; see Zwart [39] for similar results concerning fluid queues.

Equipped with Theorem 3.5 on *hidden large deviations* for random walks, in Section 4 we study queueing processes with heavy-tailed service times and finite capacity, which is a natural model to assume in many contexts. We define a *long intense period* as the fraction of time a

queue with buffer capacity $K > 0$ spends continuously above a level $\theta K, \theta \in (0, 1)$ for one sojourn and study the length of the longest such period for a given observation horizon. A closely related notion of *long strange segments*, defined in Mansfield, Rachev and Samorodnitsky [29] has also consequently been investigated in Hult *et al.* [22], Section 4, which examines the length of time the average process value spends in an unexpected regime. Considering hidden large deviations in such a setting provides further insight since we observe that the first large deviation approximation gives only a crude estimate of the distribution of the length of intense periods for large buffers. In Theorem 4.3, we derive an approximation to the distribution of the length of long intense periods in queues with large buffer sizes and conduct a simulation study in Section 4.5 to show the effectiveness of the result. In our simulation study of the buffered queue, we provide an estimate of a rare event probability using both regular and hidden large deviations which we believe has not been used in conjunction in the literature yet.

Moreover, in Section 5, we illustrate the usefulness of our result in computing probabilities of large deviation events where the step sizes in a random walk do not have finite mean. In such a case, computation of probabilities of events defined by one large jump has been discussed in Hult *et al.* [22]; we additionally exemplify the case governed by two large jumps. The theoretical results are verified using simulation. Finally, future directions for research are indicated with conclusions summarized in Section 6.

2. Notations and background

In this section, we provide a summary of frequently used notations and concepts along with a review of material necessary for the results in the following sections. We mostly adhere to the notations and definitions introduced in Lindskog, Resnick and Roy [28].

2.1. Basic notations

A few notations and concepts are summarized here. Detailed discussions are in the references provided. Unless otherwise specified, capital letters like X, Z, S with various subscripts and superscripts are reserved for real-valued (and sometimes vector-valued) random variables, whereas bold-symbolled capital letters like $\boldsymbol{X}$, $\boldsymbol{Z}$, $\boldsymbol{S}$ (again with various subscripts and superscripts) denote vector- or function-valued random elements. Small letters in bold, like $\boldsymbol{z}$, are vectors in a suitable Euclidean space where $\boldsymbol{z} = (z_1, \ldots, z_n)$ if $\boldsymbol{z} \in \mathbb{R}^n$. The following table lists notations which are often used in the paper.

$\mathcal{RV}_\beta$ Class of regularly varying functions with index $\beta \in \mathbb{R}$: that is, $f \in \mathcal{RV}_\beta$ if $f : \mathbb{R}_+ \mapsto \mathbb{R}_+$ satisfying $\lim_{t \to \infty} f(tx)/f(t) = x^\beta$, for $x > 0$. We abuse notation and write $X \in \mathcal{RV}_{-\alpha}$ for regularly varying random variables as in Definition 2.3, that is, if $1 - F \in \mathcal{RV}_\beta$ where $X \sim F$.

$\mathbb{M}(\mathbb{S}\backslash\mathbb{C})$ $\mathbb{M}(\mathbb{S}, \mathbb{C}) = \mathbb{M}(\mathbb{S}\backslash\mathbb{C})$ is the set of Borel measures on $\mathbb{S}\backslash\mathbb{C}$ that are finite on sets bounded away from $\mathbb{C}$.

$\mu_n \to \mu$ Convergence in $\mathbb{M}(\mathbb{S}\backslash\mathbb{C})$; see Definition 2.1.

$U_j^\uparrow$ $\{\boldsymbol{u} \in [0, 1]^j : 0 \le u_1 < \cdots < u_j \le 1\}$.

$a \wedge b; a \vee b$ $\quad$ $\min\{a, b\}$; $\max\{a, b\}$, respectively.

$\mathbb{D} = \mathbb{D}([0, 1], \mathbb{R})$ $\quad$ Space of all real-valued càdlàg functions on [0, 1] equipped with the Skorohod J_1-metric. Càdlàg functions are functions which are right continuous and have a left limit at every point of the domain.

ν_α^j $\quad$ Product measure on $(\mathbb{R}\backslash\{0\})^j : \underbrace{\nu_\alpha \times \cdots \times \nu_\alpha}_{j \text{ times}}$ with ν_α as defined in (7).

d_{J_1} $\quad$ Skorohod J_1-metric on $\mathbb{D}$. If Λ denotes the class of strictly increasing continuous functions $\lambda : [0, 1] \to [0, 1]$ with $\lambda(0) = 0$, $\lambda(1) = 1$, then for $f, g \in \mathbb{D}$, we define

$$d_{J_1}(f, g) := \inf_{\lambda \in \Lambda} \max\left\{ \sup_{t \in [0,1]} \left| f(t) - g \circ \lambda(t) \right|, \sup_{t \in [0,1]} \left| \lambda(t) - t \right| \right\}$$

$$= \inf_{\lambda \in \Lambda} \| f - g \circ \lambda \| \vee \| \lambda - e \| \tag{6}$$

where $e(t) = t$, $\forall t \in [0, 1]$. See Billingsley [6] for further details on the J_1-topology. Note that similar definitions could be worked out if we take $\mathbb{D}_M = \mathbb{D}([0, M], \mathbb{R})$ in place of $\mathbb{D}$.

$\mathbb{D}_{=j}$ $\quad$ Space of all real-valued step functions on [0,1] with exactly j jumps, $j \geq 1$. Assume $\mathbb{D}_{=0}$ is the space containing only the constant function at 0. Moreover, $\mathbb{D}_{=j} \subset \mathbb{D}$.

$\mathbb{D}_{\leq j}$ $\quad$ Space of all step functions on [0,1] with j or less jumps, $\mathbb{D}_{\leq j} = \bigcup_{k=0}^{j} \mathbb{D}_{=k}$.

$\mathbb{C}_{=k}(\lambda)$ $\quad$ $\{z \in \mathbb{R}^n : |\{i : |z_i| > \lambda\}| = k\}$ for $\lambda > 0$, $1 \leq k \leq n$. It denotes the subset of $\mathbb{R}^n$ where exactly k co-ordinates are above $\lambda > 0$ in absolute value. Moreover, $\mathbb{R}^n = \bigcup_{k=0}^{n} \mathbb{C}_{=k}(\lambda)$. See Section 3 for further details.

2.2. Convergence in $\mathbb{M}_\mathbb{O}$

We state our results as convergence in $\mathbb{M}_\mathbb{O}$, a mode of convergence closely related to standard weak convergence of probability measures. The idea is as follows: to allow for hidden large deviations we need to exclude the set of all possible occurrences of (regular) large deviation events from the space under consideration in order to keep the limit measure non-degenerate in that region. This is similar to how large deviations avoid the law of large numbers for zero mean random variates: we need to exclude 0 from the non-negative real line to obtain a limit measure for $\mathbb{P}(|Z_1 + \cdots + Z_n| \geq nz)$, $z > 0$. Convergence in $\mathbb{M}_\mathbb{O}$ follows the same principle but we allow for the removal of an arbitrary closed set. As a consequence, we can define convergence in càdlàg spaces where we exclude certain types of step functions which form a closed set in $\mathbb{D}$.

In particular, let $\mathbb{S}$ be a complete separable metric space, $\mathcal{B}(\mathbb{S})$ the collection of Borel sets on $\mathbb{S}$ and $\mathbb{C} \subset \mathbb{S}$ a closed subset of $\mathbb{S}$. Then we denote $\mathbb{M}(\mathbb{S}, \mathbb{C}) = \mathbb{M}(\mathbb{S}\backslash\mathbb{C})$ the set of Borel measures on $\mathbb{S}\backslash\mathbb{C}$ which are finite on sets bounded away from $\mathbb{C}$, that is, the collection of sets $A \in \mathcal{B}(\mathbb{S})$ such that $\inf\{d(x, y) : x \in \mathbb{C}, y \in A\} > 0$ where d denotes the metric on $\mathbb{S}$. Finally, we call a sequence $\{\mu_n\}_{n \geq 0}$ convergent if the assigned values converge for a suitable class of test functions or sets. Denoting $\mathbb{O} = \mathbb{S}\backslash\mathbb{C}$ the support set we use the notation $\mathbb{M}_\mathbb{O} := \mathbb{M}(\mathbb{S}, \mathbb{C})$ as eponym for the mode of convergence and use the following definition of $\mathbb{M}_\mathbb{O}$-convergence.

Definition 2.1. A sequence of measures $\{\mu_n\}_{n\geq 0} \subset \mathbb{M}_{\mathbb{O}}$ converges to $\mu_0 \in \mathbb{M}_{\mathbb{O}}$ if for all closed sets F and open sets G in $\mathcal{B}(\mathbb{S})$ which are bounded away from $\mathbb{C}$ we have

$$\limsup_{n\to\infty} \mu_n(F) \leq \mu_0(F), \quad \text{and,}$$

$$\liminf_{n\to\infty} \mu_n(G) \geq \mu_0(G).$$

We write $\mu_n \to \mu_0$ in $\mathbb{M}_{\mathbb{O}}$ as $n \to \infty$, or simply $\mu_n \to \mu_0$.

The definition above states $\mathbb{M}_{\mathbb{O}}$-convergence in terms of open and closed subsets of $\mathbb{O}$. Theorem 2.1 of Lindskog, Resnick and Roy [28] provides several alternative characterizations of convergence in $\mathbb{M}_{\mathbb{O}}$. The corresponding version of a continuous mapping theorem in $\mathbb{M}_{\mathbb{O}}$ follows as Theorem 2.3 in the same publication. We state the continuous mapping theorem again for the sake of convenience. Let $\mathbb{S}'$ be a second complete separable metric space with a closed set $\mathbb{C}' \subset \mathbb{S}'$ and with the same properties as $\mathbb{S}$; we add primes ($'$s) to denote the corresponding elements of the second space.

Theorem 2.2 (Lindskog, Resnick and Roy [28], Theorem 2.3). *Let $h : \mathbb{S} \setminus \mathbb{C} \to \mathbb{S}' \setminus \mathbb{C}'$ be a measurable map such that for all sets $A' \in \mathcal{B}(\mathbb{S}') \cap h(\mathbb{S}\setminus\mathbb{C})$ bounded away from $\mathbb{C}'$, we also have $h^{-1}(A')$ is bounded away from $\mathbb{C}$. Also assume, $\mu(D_h) = 0$, where D_h is the set of discontinuity points of h. Then $\mu_n \to \mu$ in $\mathbb{M}(\mathbb{S}, \mathbb{C})$ implies $\mu_n \circ h^{-1} \to \mu \circ h^{-1}$ in $\mathbb{M}(\mathbb{S}', \mathbb{C}')$.*

Remark 1 (Relationship between $\mathbb{M}_{\mathbb{O}}$-convergence and $w^{\#}$ convergence). In Hult *et al.* [22], the large deviation result for random walks is stated in terms of $w^{\#}$ convergence, with boundedly finite measures. Specifically they consider the space $\overline{\mathbb{D}}_0 = (0, \infty] \times \mathbb{S}_{\mathbb{D}}$, where $\mathbb{S}_{\mathbb{D}}$ is the unit sphere in $\mathbb{D}$. The metric on the radial part is defined as $d_{(0,\infty]}(x, y) := |1/x - 1/y|$, thus making any set not bounded away from the zero function (in the usual J_1 metric) unbounded in the modified space. Unfortunately it is not immediately clear how to extend this theory to allow for the removal of more than just the zero function, whereas convergence in $\mathbb{M}_{\mathbb{O}}$ on the contrary is specifically designed for this purpose and hence we use this as our notion of convergence.

2.3. Regular variation and heavy-tailed large deviations

A measurable function $f : (0, \infty) \to (0, \infty)$ is regularly varying at infinity with index $\alpha \in \mathbb{R}$ if $\lim_{t\to\infty} f(tx)/f(t) = x^{\alpha}$ for all $x > 0$. A sequence of positive numbers $\{a_n\}_{n\geq 1}$ is regularly varying with index $\alpha \in \mathbb{R}$ if $\lim_{n\to\infty} a_{[cn]}/a_n = c^{\alpha}$ for all $c > 0$. Regular variation of unbounded random variables thus usually is defined in terms of regular variation of the tail of the corresponding cumulative distribution functions at infinity; see Bingham, Goldie and Teugels [7], de Haan and Ferreira [14], or Resnick [35] for related properties and examples. We work with an equivalent definition stated in terms of $\mathbb{M}_{\mathbb{O}}$-convergence Lindskog, Resnick and Roy [28], Section 3.2.

Definition 2.3. A random variable X is *regularly varying* at infinity if there exists a regularly varying sequence γ_n and a non-zero measure $\mu \in \mathbb{M}(\mathbb{R}\backslash\{0\})$ such that

$$\gamma_n \mathbb{P}(X/n \in \cdot) \to \mu(\cdot), \quad n \to \infty,$$

in $\mathbb{M}(\mathbb{R}\backslash\{0\})$.

Since the measure μ satisfies the scaling property $\mu(sA) = s^{-\alpha}\mu(A)$, $s > 0$, $A \in \mathcal{B}(\mathbb{R}\backslash\{0\})$ for some $\alpha \geq 0$ we write $X \in \mathcal{RV}_{-\alpha}$. In addition, Definition 2.3 implies that $\mathbb{P}(|X| > n) = L(n)n^{-\alpha}$ for some slowly varying function $L(n)$. Throughout this paper, we assume $\alpha > 0$. Moreover, for $X \in \mathcal{RV}_{-\alpha}$, we also assume that the following condition is satisfied:

$$\lim_{n\to\infty} \frac{\mathbb{P}(X > n)}{\mathbb{P}(|X| > n)} = p, \qquad \lim_{n\to\infty} \frac{\mathbb{P}(X < -n)}{\mathbb{P}(|X| > n)} = 1 - p := q,$$

for some $0 \leq p \leq 1$. This is called the *tail balance condition*. For univariate random variables $X \in \mathcal{RV}_{-\alpha}$ we stick to this choice unless otherwise specified. We also denote by ν_α, the following measure on $\mathbb{R}\backslash\{0\}$ for $x > 0$, $y > 0$,

$$\nu_\alpha\big((-\infty, -y) \cup (x, \infty)\big) = qy^{-\alpha} + px^{-\alpha}. \tag{7}$$

The limit measure $\mu(\cdot)$ in Definition 2.3 is equal to $\nu_\alpha(\cdot)$ if we choose the sequence γ_n to be $[\mathbb{P}(|X| > n)]^{-1}$. To clearly distinguish the concepts of regular variation and large deviations for heavy-tailed random variables, we provide the following definition.

Definition 2.4. A sequence of random variables $\{X_n\}_{n\geq 1} \subseteq \mathbb{R}$, with $X_n \to 0$ in probability, satisfies a *heavy-tailed large deviation type limit* on $\mathbb{R}\backslash\{0\}$ if there exists a positive sequence $\gamma_n \uparrow \infty$ and a non-zero measure $\mu \in \mathbb{M}(\mathbb{R}\backslash\{0\})$ such that as $n \to \infty$,

$$\gamma_n \mathbb{P}(X_n \in \cdot) \to \mu(\cdot),$$

in $\mathbb{M}(\mathbb{R}\backslash\{0\})$.

Remark 2. The similarity between the definitions of regular variation and heavy-tailed large deviation type limit (LDL) is quite evident here. One salient difference is that regular variation is defined for a single random element, whereas an LDL, for a sequence of random elements. The special case of $X_n = X/n$ shows that regular variation is a specific form of heavy tailed LDL according to our definition. The definition of an LDL implies that $\mathbb{P}(X_n \in \cdot) \to 0$ as $n \to \infty$ for all Borel sets in $\mathbb{R}\backslash\{0\}$.

We provide a more general definition of LDLs for random elements on a complete separable metric space $\mathbb{S}$, which helps in defining hidden large deviations eventually. Additionally, we no longer restrict to removing the zero element, but an arbitrary closed set $\mathbb{C} \subset \mathbb{S}$.

Definition 2.5. A sequence of random elements $\{X_n\}_{n\geq 1} \subset \mathbb{S}$ satisfies a *heavy-tailed large deviation type limit* on $\mathbb{S}\backslash\mathbb{C}$ for a closed set $\mathbb{C} \subset \mathbb{S}$ if there exists a positive sequence $\gamma_n \uparrow \infty$, and a non-zero measure $\mu \in \mathbb{M}(\mathbb{S}\backslash\mathbb{C})$ such that as $n \to \infty$,

$$\gamma_n \mathbb{P}(X_n \in \cdot) \to \mu(\cdot),$$

in $\mathbb{M}(\mathbb{S}\backslash\mathbb{C})$. We write $X_n \in \mathrm{LD}(\gamma_n, \mu, \mathbb{S}\backslash\mathbb{C})$.

The definition of heavy-tailed large deviations principle as given in Hult *et al.* [22], Definition 1.3, is equivalent to Definition 2.5 for stochastic processes with sample paths in $\mathbb{D}$. To avoid any confusion with regards to the established use of the term large deviations principle (see, e.g., Dembo and Zeitouni [15]), we use a variation of the term, namely, *large deviation type limit* (LDL). It has been observed, especially in the case of heavy-tailed random walks, that the limit measure μ obtained on $\mathbb{D}\backslash\{0\}$, concentrates only on step functions with one jump; see Hult *et al.* [22], Theorem 2.1. Hence, we enquire whether a different structure is observable if we do not allow one jump functions to be in the support of the limit measure μ above. Essentially, we are asking how often do we see events which are not governed by just exactly one jump in the space $\mathbb{D}$. The same question may be asked iteratively by removing the support set of a new found limit measure and examining the hidden structure of rarer and rarer events. Hence, one may be able to find a sequence $\{\mathbb{C}^{(j)}\}_{j\geq 1}$ of closed subsets of $\mathbb{S}$ with $\mathbb{C}^{(j+1)} \supset \mathbb{C}^{(j)}$ and positive sequences $\gamma_n^{(j)} \uparrow \infty$ with $\gamma_n^{(j+1)}/\gamma_n^{(j)} \to \infty$ as $n \to \infty$, and non-zero measures $\mu^{(j)} \in \mathbb{M}(\mathbb{S}\backslash\mathbb{C}^{(j)})$, for $j \geq 1$ such that

$$X_n \in \mathrm{LD}\big(\gamma_n^{(j)}, \mu^{(j)}, \mathbb{S}\backslash\mathbb{C}^{(j)}\big), \quad j \geq 1.$$

The limit measure $\mu^{(j)}$ necessarily concentrates on $\mathbb{C}^{(j+1)}\backslash\mathbb{C}^{(j)}$. Thus, the kth level LDL uncovers the structure of rare events which were *hidden* (i.e., negligible) under the scaling of the preceding jth level LDLs of the sequence with $j < k$.

3. Hidden large deviations and random walks

Equipped with the terminology and tools in Section 2, we proceed to understand the structure of heavy-tailed random walks in this section. We look at heavy-tailed random walks as elements of $\mathbb{D}$. The key result for hidden large deviations for heavy-tailed random walks is in Theorem 3.5.

3.1. Bounds on sums of random variables

For random variables $Z_1, \ldots, Z_n$, denote their sum by $S_n = Z_1 + \cdots + Z_n$. Here S_n denotes the n-th step of a random walk. One of the key tools to bound movements in the random walk caused by "small" realizations will be Bernstein's inequality, see Bennett [5].

Lemma 3.1 (Bernstein's inequality). *Let* $Z_1, \ldots, Z_n$ *be i.i.d. bounded random variables with* $\mathbb{E}Z_1 = 0$, $\mathrm{Var}[Z_1] = \sigma^2$ *and* $|Z_1| \leq M$. *Then for any* $t > 0$,

$$\mathbb{P}\big(|S_n| \geq t\big) \leq 2\exp\left\{-\frac{t^2}{2n\sigma^2 + \frac{2}{3}Mt}\right\}.$$

We use this exponential bound on the absolute value of the sum to bound the probability of a large deviation in the sum of regularly varying random variables happening due to many variables attaining a small but non-negligible value. This bound, as we see hence, turns out to be exponential rather than polynomial in the deviation level λ_n.

Lemma 3.2. *Let* $\{Z_i\}_{i \geq 1}$ *be a sequence of i.i.d. random variables with* $Z_1 \in \mathcal{RV}_{-\alpha}$, $\alpha > 0$. *In case* $\mathbb{E}|Z_1| < \infty$, *we assume* $\mathbb{E}Z_1 = 0$. *Denote* $S_n = \sum_{k=1}^n Z_k$ *and let* $\lambda_n \in \mathcal{RV}_\rho$ *be a regularly varying sequence such that in case*

$$(1) \quad \mathrm{Var}[Z_1] < \infty, \quad \text{we have } \rho > \frac{1}{2}, \quad \text{and,}$$

$$(2) \quad \mathrm{Var}[Z_1] = \infty, \quad \text{we have } \alpha\rho > 1.$$

Then for any $\delta > 0$ *and* $\varepsilon_0 > 0$ *small enough, there exists a constant* $c > 0$ *such that for large enough* n,

$$\mathbb{P}\big(|S_n| > \delta\lambda_n, |Z_i| \leq \lambda_n^{1-\varepsilon_0}, \forall i \leq n\big) < 2\exp\big(-c\lambda_n^{\varepsilon_0}\big).$$

Remark 3. When Z_1 has finite variance, the condition, $\rho > 1/2$ guarantees that $\lambda_n \uparrow \infty$ fast enough such that we avoid the central limit regime. When Z_1 has infinite variance, then $\alpha\rho > 1$ ensures that the probability of at least one of the variables exceeding a large threshold on the scale of λ_n still tends to zero.

Proof of Lemma 3.2. We obtain the result by applying Lemma 3.1 to an appropriately truncated version of Z_i's (where the truncation still depends on n). Note that for $\mathrm{Var}[Z_1] = \infty$ we assume $\alpha\rho > 1$ and for $\mathrm{Var}[Z_1] < \infty$, we assume $\rho > 1/2$ and we know $\alpha \geq 2$, hence $\alpha\rho > 1$. First, we need to show the following auxiliary result, which follows from Karamata's theorem (Bingham, Goldie and Teugels [7]) and the assumption that $\alpha\rho > 1$. We claim that for $0 < \varepsilon_0 < 1$, with ε_0 small enough,

$$\left|\frac{n}{\lambda_n}\mathbb{E}[Z_1\mathbb{1}_{[|Z_1| \leq \lambda_n^{1-\varepsilon_0}]}]\right| \to 0, \quad n \to \infty. \tag{8}$$

We show (8) separately for the cases when $\mathbb{E}|Z_1| < \infty$ and $\mathbb{E}|Z| = \infty$.

1. First, assume that $\mathbb{E}|Z_1| < \infty$ and thus by assumption $\mathbb{E}[Z_1] = 0$. Then

$$\left| \frac{n}{\lambda_n} \mathbb{E}[Z_1 \mathbb{1}_{[|Z_1| \leq \lambda_n^{1-\varepsilon_0}]}] \right| = \left| \frac{n}{\lambda_n} (-1) \mathbb{E}[Z_1 \mathbb{1}_{[|Z_1| > \lambda_n^{1-\varepsilon_0}]}] \right|$$

$$\leq \frac{n}{\lambda_n} \mathbb{E}[|Z_1| \mathbb{1}_{[|Z_1| > \lambda_n^{1-\varepsilon_0}]}] = \frac{n}{\lambda_n} \int_{\lambda_n^{1-\varepsilon_0}}^{\infty} \mathbb{P}(|Z_1| > x) \, dx.$$

The final expression above is bounded above by a regularly varying function with index $1 - \rho + (-\alpha + 1)\rho(1 - \varepsilon_0)$ which for small enough ε_0 is negative since $\alpha\rho > 1$ by assumption.

2. In the other case $\mathbb{E}|Z_1| = \infty$. Hence $\alpha \leq 1$, so that the condition $\alpha\rho > 1$ implies $\rho > 1$, and we use the following bound:

$$\left| \frac{n}{\lambda_n} \mathbb{E}[Z_1 \mathbb{1}_{[|Z_1| \leq \lambda_n^{1-\varepsilon_0}]}] \right| \leq \frac{n}{\lambda_n} \mathbb{E}[|Z_1| \mathbb{1}_{[|Z_1| \leq \lambda_n^{1-\varepsilon_0}]}] \leq \frac{n}{\lambda_n} \int_0^{\lambda_n} \mathbb{P}(|Z_1| > x) \, dx.$$

Since $\alpha\rho > 1$ and $\rho > 1$, by assumption, for small enough ε_0, the final term above vanishes as $n \to \infty$.

Hence, (8) holds. Next, observe that given $\delta > 0$, for small enough ε_0 and large enough n,

$$\mathbb{P}(A_n) := \mathbb{P}(|S_n| > \delta\lambda_n, |Z_i| \leq \lambda_n^{1-\varepsilon_0} \forall 1 \leq i \leq n)$$

$$\leq \mathbb{P}\left(\left| \sum_{i=1}^n Z_i \mathbb{1}_{[|Z_i| \leq \lambda_n^{1-\varepsilon_0}]} \right| > \delta\lambda_n \right)$$

$$\leq \mathbb{P}\left(\left| \sum_{i=1}^n \left(Z_i \mathbb{1}_{[|Z_i| \leq \lambda_n^{1-\varepsilon_0}]} - \mathbb{E}[Z_i \mathbb{1}_{[|Z_i| \leq \lambda_n^{1-\varepsilon_0}]}] \right) \right| > \frac{\delta}{2}\lambda_n \right),$$

where the last inequality is obtained using (8). Using Lemma 3.1 to bound the sum of n zero mean random variables bounded in absolute value by $M = 2\lambda_n^{1-\varepsilon_0}$, we obtain that for large enough n,

$$\mathbb{P}(A_n) \leq 2\exp\left(-\frac{(\frac{\delta}{2}\lambda_n)^2}{2n \operatorname{Var}[Z_1 \mathbb{1}_{[|Z_1| \leq \lambda_n^{1-\varepsilon_0}]}] + \frac{4}{3}\lambda_n^{1-\varepsilon_0}\frac{\delta}{2}\lambda_n} \right)$$

$$\leq 2\exp\left(-\lambda_n^{\varepsilon_0} \frac{c_1}{c_2 + \beta(n)} \right),$$

where c_1, c_2 are positive constants and

$$\beta(n) = \frac{2n \operatorname{Var}[Z_1 \mathbb{1}_{[|Z_1| \leq \lambda_n^{1-\varepsilon_0}]}]}{\lambda_n^{2-\varepsilon_0}}.$$

Next, we show that $\beta(n) \to 0$ as $n \to \infty$ which implies that for large enough n, there exists $\zeta > 0$ such that

$$\mathbb{P}(A_n) \leq 2\exp\left(-\lambda_n^{\varepsilon_0} \frac{c_1}{c_2 + \zeta} \right) = 2\exp\left(-c\lambda_n^{\varepsilon_0} \right),$$

where $c = \frac{c_1}{c_2+\zeta}$, and thus the lemma is proven. We show $\beta(n) \to 0$ in the three different cases as follows.

1. If $\alpha \in (0, 2)$ (implying infinite variance and $\alpha\rho > 1$), using Karamata's theorem (Bingham, Goldie and Teugels [7], Proposition 1.5.8) for small enough ε_0, large enough n and constant $C > 0$ we have

$$\beta(n) \le \frac{2n}{\lambda_n^{2-\varepsilon_0}} \mathbb{E}\big[Z_1^2 \mathbb{1}_{[|Z_1| \le \lambda_n^{1-\varepsilon_0}]}\big]$$

$$\sim \frac{2n}{\lambda_n^{2-\varepsilon_0}} \times C\lambda_n^{2(1-\varepsilon_0)} \mathbb{P}\big(|Z_1| > \lambda_n^{1-\varepsilon_0}\big)$$

$$\sim 2Cn\lambda_n^{-\varepsilon_0} \mathbb{P}\big(|Z_1| > \lambda_n^{1-\varepsilon_0}\big) \to 0 \quad (n \to \infty).$$

2. If $\mathrm{Var}[Z_1] < \infty$, then $\beta(n) \le C\frac{n}{\lambda_n^{2-\varepsilon_0}}$ for some $C > 0$ and hence vanishes as $n \to \infty$ for small enough $\varepsilon_0 > 0$.
3. If $\alpha = 2$ and $\mathrm{Var}[Z_1] = \infty$, then the variance is a slowly varying function. Again, for ε_0 small enough $n\lambda_n^{-2+\varepsilon_0} \to 0$ at a polynomial rate and hence $\beta(n) \to 0$. $\square$

Remark 4. By considering sets of the form

$$A_{k,n} = \big\{|S_k| > \delta\lambda_n, |Z_i| \le \lambda_n^{1-\varepsilon_0} \,\forall 1 \le i \le n\big\}, \quad 1 \le k \le n,$$

the proof of Lemma 3.2 can easily be adapted to show that under the conditions of Lemma 3.2 we have

$$\mathbb{P}\big(|S_k| > \delta\lambda_n, |Z_i| \le \lambda_n^{1-\varepsilon_0}, \forall 1 \le i \le n\big) \le 2\exp\big(-c\lambda_n^{\varepsilon_0}\big), \quad 1 \le k \le n.$$

3.2. Random walks embedded in $\mathbb{D} := \mathbb{D}([0, 1], \mathbb{R})$

Now we embed the random walk $S_n = Z_1 + \cdots + Z_n$ in $\mathbb{D} = \mathbb{D}([0, 1], \mathbb{R})$ and discuss its large deviations. Let $\mathbf{Z}^{(n)} = (Z_1, \ldots, Z_n)$ where Z_i's are i.i.d. realizations having the same distribution as Z_1. For $t \in [0, 1]$ and $k \in \{1, \ldots, n\}$ define the functions $X_k^{(n)}(t) = Z_k\mathbb{1}_{[\frac{k}{n} \le t]}$ on $\mathbb{D}$. Now define

$$\mathbf{X}^{(n)}(t) := \sum_{k=1}^{n} X_k^{(n)}(t) = \sum_{k=1}^{\lfloor nt \rfloor} Z_k = S_{\lfloor nt \rfloor}$$

to be the embedding of the random walk induced by $Z_1, \ldots, Z_n$ into the space $\mathbb{D}$. Moreover, let $\mathbb{C}_{=0}(\lambda) := \{z \in \mathbb{R}^n : |z_i| \le \lambda, \forall i\}$. The following corollary is a consequence of Lemma 3.2.

Corollary 3.3. *Under the conditions of Lemma 3.2, for any $\delta > 0$ and $\varepsilon_0 > 0$ small enough, there exist a constant $c > 0$ such that for large enough n,*

$$\mathbb{P}\bigg(\sup_{t\in[0,1]} \big|\mathbf{X}^{(n)}\big| > \delta\lambda_n, \mathbf{Z}^{(n)} \in \mathbb{C}_{=0}\big(\lambda_n^{1-\varepsilon_0}\big)\bigg) \le 2\exp\big(-c\lambda_n^{\varepsilon_0/2}\big)$$

Proof. Observe that from Lemma 3.2, for any $\delta > 0$ and $\varepsilon_0 > 0$ small enough and for large enough n,

$$\mathbb{P}\Big(\sup_{t\in[0,1]}\big|X^{(n)}\big| > \delta\lambda_n, \mathbf{Z}^{(n)} \in \mathbb{C}_{=0}\big(\lambda_n^{1-\varepsilon_0}\big)\Big) = \mathbb{P}\Big(\sup_{t\in[0,1]}|S_{\lfloor nt\rfloor}| > \delta\lambda_n, |Z_i| \leq \lambda_n^{1-\varepsilon_0}, \forall i \leq n\Big)$$

$$\leq \mathbb{P}\Big(\sup_{1\leq k\leq n}|S_k| > \delta\lambda_n, |Z_i| \leq \lambda_n^{1-\varepsilon_0}, \forall i \leq n\Big)$$

$$\leq 2n\exp\big(-c\lambda_n^{\varepsilon_0}\big)$$

$$\leq 2\exp\big(-c\lambda_n^{\varepsilon_0/2}\big) \quad \text{(for large enough } n\text{)},$$

for some constant $c > 0$.   □

Now we define functions which relate vectors in $\mathbb{R}^n$ to càdlàg step functions in $\mathbb{D}$. For integers $j \in \mathbb{N}$, denote $U_j^\uparrow := \{\mathbf{u} \in [0,1]^j : 0 \leq u_1 < \cdots < u_j \leq 1\}$ and define functions

$$h_j : \big(\mathbb{R}\backslash\{0\}\big)^j \times U_j^\uparrow \to \mathbb{D},$$

$$h_j\big((\mathbf{z}, \mathbf{u})\big)(t) := \sum_{i=1}^{j} z_i\mathbb{1}_{[u_i \leq t]}.$$

The maps h_j allow us to define the collection of functions with exactly j jumps as subsets of $\mathbb{D}$. Hence, we define the following classes of càdlàg functions.

$$\mathbb{D}_{=0} := \big\{x \in \mathbb{D} : x(t) = 0, 0 \leq t \leq 1\big\} = \big\{\text{the zero function in [0,1]}\big\},$$

$$\mathbb{D}_{=j} := \big\{x \in \mathbb{D} : x(t) = h_j(\mathbf{z}, \mathbf{u})(t), 0 \leq t \leq 1, (\mathbf{z}, \mathbf{u}) \in \big(\mathbb{R}\backslash\{0\}\big)^j \times U_j^\uparrow\big)\big\},$$

$$\mathbb{D}_{\leq j} := \bigcup_{i=0}^{j}\mathbb{D}_{=i} = \{\text{càdlàg step functions with } j \text{ or less jumps}\}.$$

Lemma 3.4. *The map* $h_j : (\mathbb{R}\backslash\{0\})^j \times U_j^\uparrow \mapsto \mathbb{D}$ *is continuous for* $j \in \mathbb{N}$.

Proof. The proof which is similar to Lemma 5.3 in Lindskog, Resnick and Roy [28] is skipped here.   □

Remark 5. The function h_j maps points in $(\mathbb{R}\backslash\{0\})^j \times U_j^\uparrow$ to functions in $\mathbb{D}_{=j} \subset \mathbb{D}$, which are càdlàg functions in $[0, 1]$ with exactly j jumps. Hence for any $F \subset \mathbb{D}$ bounded away from $\mathbb{D}_{\leq(j-1)}$, we have $h_j^{-1}(F) = h_j^{-1}(F \cap \mathbb{D}_{=j})$. Hence, $h_j \circ h_j^{-1}(F) = F \cap \mathbb{D}_{=j}$.

The following result is the main theorem of this paper and extends the large deviation result of Hult *et al.* [22] in the setting of Lindskog, Resnick and Roy [28] in order to obtain, what we call *hidden large deviations*. The special case of $j = 1$ in Theorem 3.5 corresponds to Hult *et al.*

[22], Theorem 2.1. The Lebesgue measure (in $\mathbb{R}^j$) is denoted Leb_j and ν_α^j is the j-fold product measure of ν_α (again in $\mathbb{R}^j$) with ν_α as defined in (7).

Theorem 3.5. *Let $j \geq 1$. Under the conditions of Lemma 3.2, for $\lambda_n \to \infty$ as $n \to \infty$,*

$$\gamma_n^{(j)} \mathbb{P}\big(X^{(n)}/\lambda_n \in \cdot\big) \to \big(\nu_\alpha^j \times \mathrm{Leb}_j\big) \circ h_j^{-1}(\cdot) \tag{9}$$

in $\mathbb{M}(\mathbb{D}\backslash\mathbb{D}_{\leq(j-1)})$ as $n \to \infty$, where $\gamma_n^{(j)} = [n\mathbb{P}(|Z_1| > \lambda_n)]^{-j}$.

Remark 6. First note that under the conditions of Lemma 3.2 we have $X^{(n)}/\lambda_n \to 0$ ($0 \in \mathbb{D}$) in probability.

Remark 7. The theorem states that the random element $X_n = X^{(n)}/\lambda_n$ satisfies a sequence of LDLs on $\{\mathbb{D}\backslash\mathbb{D}_{\leq j-1}\}_{j\geq 1}$, which means that the large deviations of the random walk concentrate on step functions with an increasing number of steps at increasingly faster rates. In particular, for polynomially bounded rates γ_n, large deviations of partial sum processes of i.i.d. regularly varying random variables will always concentrate on step functions, among all functions in $\mathbb{D}$. Naturally, one may ask whether any LDL on $\mathbb{D}\backslash\bigcup_{j=0}^\infty \mathbb{D}_{=j}$ can be found and at what rate would events in such a space occur. Evidently, it seems that the rate must be faster than all polynomials. However, such considerations are beyond the framework and scope of the current paper and we leave these for later investigations.

Proof of Theorem 3.5. We show convergence in $\mathbb{M}_\mathbb{O}$ according to Definition 2.1 for (9), starting with the upper bound for closed sets. The idea is to dissect the space $\mathbb{R}^n$, which contains the first n elements of the random walk, into a union of n disjoint sets that define which dimensions are allowed to be "big". For any $k = 0, 1, \ldots, n$, and $\lambda > 0$, define

$$\mathbb{C}_{=k}(\lambda) = \big\{z \in \mathbb{R}^n : \big|\{i : |z_i| > \lambda\}\big| = k\big\}.$$

Hence, $\mathbb{C}_{=k} \subset \mathbb{R}^n$ are all points in $\mathbb{R}^n$ which have exactly k-co-ordinates with absolute value greater than λ. Clearly

$$\mathbb{R}^n = \bigcup_{k=0}^n \mathbb{C}_{=k}(\lambda).$$

Upper bound. Let $F \subset \mathbb{D}$ be a closed set, bounded away from $\mathbb{D}_{\leq(j-1)}$. Then for small $\varepsilon_0 > 0$,

$$\mathbb{P}\big(X^{(n)}/\lambda_n \in F\big) = \mathbb{P}\Bigg(X^{(n)}/\lambda_n \in F, Z^{(n)} \in \bigcup_{k=0}^n \mathbb{C}_{=k}\big(\lambda_n^{1-\varepsilon_0}\big)\Bigg)$$

$$= \sum_{i=0}^n \mathbb{P}\big(X^{(n)}/\lambda_n \in F, Z^{(n)} \in \mathbb{C}_{=k}\big(\lambda_n^{1-\varepsilon_0}\big)\big) =: \sum_{i=0}^n \mathbb{P}(B_i).$$

We show that when multiplied by $\gamma_n^{(j)}$, all the probabilities above are negligible except $\mathbb{P}(B_j)$. Now, since F was chosen to be bounded away from $\mathbb{D}_{\leq(j-1)}$, there exists a $\delta_0 > 0$, such that

all elements of F have a minimum distance δ_0 to step functions with at most $j - 1$ jumps. In particular, F is bounded away from the zero element in $\mathbb{D}$.

1. Bounding $\mathbb{P}(B_0)$. Using Corollary 3.3, we have constants $c_0 > 0$ and $\varepsilon_0 > 0$ such that,

$$\mathbb{P}(B_0) \leq \mathbb{P}\big(\sup\big|X^{(n)}(t)/\lambda_n\big| > \delta_0/2, |Z_i| \leq \lambda_n^{1-\varepsilon_0}\big)$$
$$\leq 2\exp\big(-c_0\lambda_n^{\varepsilon_0}\big).$$

Hence, this term is exponentially bounded and goes to 0 when multiplied by $\gamma_n^{(j)}$.

2. Bounding $\mathbb{P}(B_i)$ *for* $1 \leq i \leq j - 1$. For $i \in \boldsymbol{I} := \{1, 2, \ldots, n\}$, denote by

$$\boldsymbol{K}(i) = \big\{\boldsymbol{k} = \{k_1, \ldots, k_i\} : 1 \leq k_1 < \cdots < k_i \leq n\big\},$$

all possible subsets of size i of the index set $\boldsymbol{I}$. We show that $\mathbb{P}(B_i)$ for $1 \leq i \leq j - 1$ are also exponentially bounded. With the same δ_0 as previously chosen, we have

$$\mathbb{P}(B_i) = \mathbb{P}\big(X^{(n)}/\lambda_n \in F, \boldsymbol{Z}^{(n)} \in \mathbb{C}_{=i}\big(\lambda_n^{1-\varepsilon_0}\big)\big)$$
$$= \sum_{\boldsymbol{k} \in \boldsymbol{K}(i)} \mathbb{P}\big(X^{(n)}/\lambda_n \in F, |Z_l| > \lambda_n^{1-\varepsilon_0}, \forall l \in \boldsymbol{k}, |Z_l| \leq \lambda_n^{1-\varepsilon_0}, \forall l \in \boldsymbol{I}\backslash\boldsymbol{k}\big)$$
$$\leq \sum_{\boldsymbol{k} \in \boldsymbol{K}(i)} \mathbb{P}\Bigg(\sup_{t \in [0,1]}\bigg|X^{(n)}(t) - \sum_{m=1}^{i} X_{k_m}^{(n)}(t)\bigg| > \lambda_n\frac{\delta_0}{2},$$
$$|Z_l| > \lambda_n^{1-\varepsilon_0}, \forall l \in \boldsymbol{k}, |Z_l| \leq \lambda_n^{1-\varepsilon_0}, \forall l \in \boldsymbol{I}\backslash\boldsymbol{k}\Bigg)$$
$$\leq \sum_{\boldsymbol{k} \in \boldsymbol{K}(i)} \mathbb{P}\Bigg(\sup_{t \in [0,1]}\bigg|\sum_{l \in \boldsymbol{I}\backslash\boldsymbol{k}} X_l^{(n)}(t)\bigg| > \lambda_n\frac{\delta_0}{2}, |Z_l| \leq \lambda_n^{1-\varepsilon_0}, \forall l \in \boldsymbol{I}\backslash\boldsymbol{k}\Bigg)$$
$$= \sum_{\boldsymbol{k} \in \boldsymbol{K}(i)} \mathbb{P}\Bigg(\sup_{m \in \boldsymbol{I}}\bigg|\sum_{1 \leq l \leq m, l \in \boldsymbol{I}\backslash\boldsymbol{k}} Z_l\bigg| > \lambda_n\frac{\delta_0}{2}, |Z_l| \leq \lambda_n^{1-\varepsilon_0}, \forall l \in \boldsymbol{I}\backslash\boldsymbol{k}\Bigg).$$

Now note that the probability on the right-hand side of the last inequality above is invariant under changing subsets $\boldsymbol{k} \in \boldsymbol{K}(i)$. Since the size of the set $\boldsymbol{K}(i)$ is $|\boldsymbol{K}(i)| = \binom{n}{i}$, we have

$$\mathbb{P}(B_i) \leq \binom{n}{i}\mathbb{P}\Bigg(\sup_{1 \leq j \leq n-i}|S_j| > \lambda_n\frac{\delta_0}{2}, |Z_l| \leq \lambda_n^{1-\varepsilon_0}, 1 \leq l \leq n - i\Bigg),$$
$$\leq 2\binom{n}{i}(n-i)\exp\big(-c_i\lambda_n^{\varepsilon_0}\big),$$

for some $c_i > 0$ according to Corollary 3.3. Since our choice of $\gamma_n^{(j)} = [n\mathbb{P}(|Z_1| > \lambda_n)]^{-j}$, clearly $\gamma_n^{(j)} \sum_{i=1}^{j-1} \mathbb{P}(B_i) \to 0$, as $n \to \infty$.

3. Bounding $\mathbb{P}(B_i)$ for $j + 1 \leq i \leq n$. We bound the quantity $\gamma_n^{(j)} \sum_{i=j+1}^{n} \mathbb{P}(B_i)$ together. We argue that when multiplied with $\gamma_n^{(j)}$, the probability of events with more than j large jumps is also negligible. Observe that

$$\gamma_n^{(j)} \sum_{i=j+1}^{n} \mathbb{P}(B_i) \leq \gamma_n^{(j)} \mathbb{P}\big(\exists k_1, \ldots, k_{j+1} \in I : |Z_{k_i}| > \lambda_n^{1-\varepsilon_0}, i = 1, \ldots, j+1\big)$$

$$\leq \gamma_n^{(j)} \binom{n}{j+1} \mathbb{P}\big(|Z_1| > \lambda_n^{1-\varepsilon_0}\big)^{j+1}$$

$$\leq cn \frac{\mathbb{P}(|Z_1| > \lambda_n^{1-\varepsilon_0})^{j+1}}{\mathbb{P}(|Z_1| > \lambda_n)^j} =: f_n,$$

for some $c > 0$. Now $f_n \in \mathcal{RV}_{r_0}$ with parameter

$$r_0 := 1 - (j+1)\rho\alpha + \varepsilon_0\rho(j+1)\alpha + j\rho\alpha = (1 - \alpha\rho) + (j+1)\varepsilon_0\rho\alpha,$$

see de Haan and Ferreira [14], Appendix, for details on operations on regular variation. Since by choice $\alpha\rho > 1$, for small enough ε_0, we have $r_0 < 0$. Hence,

$$\gamma_n^{(j)} \sum_{i=j+1}^{n} \mathbb{P}(B_i) \leq f_n \to 0$$

as $n \to \infty$.

4. Bounding $\mathbb{P}(B_j)$. Finally, we are left with the term $\gamma_n^{(j)} \mathbb{P}(B_j)$ which is the non-negligible contributing term for large n. For $\delta > 0$, let

$$F^\delta := \big\{x \in \mathbb{D} : d_{J_1}(x, F) \leq \delta\big\}$$

with δ small enough such that F^δ is still bounded away from $\mathbb{D}_{\leq(j-1)}$.

$$\mathbb{P}(B_j) = \mathbb{P}\big(X^{(n)}/\lambda_n \in F, Z^{(n)} \in \mathbb{C}_{=j}\big(\lambda_n^{1-\varepsilon_0}\big)\big)$$

$$= \sum_{k \in K(j)} \mathbb{P}\big(X^{(n)}/\lambda_n \in F, |Z_l| > \lambda_n^{1-\varepsilon_0}, \forall l \in k, |Z_k| \leq \lambda_n^{1-\varepsilon_0}, \forall l \in I \backslash k\big)$$

$$\leq \sum_{k \in K(j)} \mathbb{P}\Bigg(\sup_{t \in [0,1]} \bigg|X^{(n)}(t) - \sum_{m=1}^{j} X_{k_m}^{(n)}(t)\bigg| \leq \lambda_n\delta, X^{(n)}/\lambda_n \in F,$$

$$|Z_l| > \lambda_n^{1-\varepsilon_0}, \forall l \in k, |Z_l| \leq \lambda_n^{1-\varepsilon_0}, \forall l \in I \backslash k\Bigg)$$

$$+ \sum_{k \in K(j)} \mathbb{P}\Bigg(\sup_{t \in [0,1]} \bigg|X^{(n)}(t) - \sum_{m=1}^{j} X_{k_m}^{(n)}(t)\bigg| > \lambda_n\delta, X^{(n)}/\lambda_n \in F,$$

$$|Z_l| > \lambda_n^{1-\varepsilon_0}, \forall l \in \mathbf{k}, |Z_l| \le \lambda_n^{1-\varepsilon_0}, \forall l \in \mathbf{I} \backslash \mathbf{k} \Bigg)$$

$$\le \sum_{\mathbf{k} \in K(j)} \mathbb{P}\left(\sum_{m=1}^{j} X_{k_m}^{(n)} / \lambda_n \in F^{\delta} \right)$$

$$+ \binom{n}{j} \mathbb{P}\left(\sup_{t \in [0,1]} \left| \sum_{l \in \mathbf{I} \backslash \mathbf{k}} X_l^{(n)} \right| > \lambda_n \delta, |Z_l| \le \lambda_n^{1-\varepsilon_0}, \forall l \in \mathbf{I} \backslash \mathbf{k} \right)$$

$$= P_{j,1}^{(n)} + P_{j,2}^{(n)}. \tag{10}$$

Now, using Lemma 3.2, and arguments similar to the one for bounding $\mathbb{P}(B_i)$ for $1 \le i \le j-1$, we can check that the quantity $P_{j,2}^{(n)}$ is negligible at rate $\gamma_n^{(j)}$ and hence $\gamma_n^{(j)} P_{j,2}^{(n)} \to 0$ as $n \to \infty$.

In the remaining term $P_{j,1}^{(n)}$, we use the inverse of the map h_j to measure the probability. For any set $F^* \subset \mathbb{D}$, define

$$\left(J(F^*), T(F^*) \right) := h_j^{-1}\left(F^* \cap \mathbb{D}_{=j} \right) \subset \left(\mathbb{R} \backslash \{0\} \right)^j \times U_j^{\uparrow}$$

to be the pre-image of $F^* \cap \mathbb{D}_{=j}$ under the map h_j broken into the *jump* part and the *time* part. Clearly, $\sum_{m=1}^{j} X_{k_m}^{(n)} / \lambda_n \in F^{\delta}$ is equivalent to $\sum_{m=1}^{j} X_{k_m}^{(n)} / \lambda_n \in F^{\delta} \cap \mathbb{D}_{=j}$, as F^{δ} is bounded away from $\mathbb{D}_{\le j-1}$. Thus,

$$P_{j,1}^{(n)} = \sum_{\mathbf{k} \in K(j)} \mathbb{P}\left(\sum_{m=1}^{j} X_{k_m}^{(n)} / \lambda_n \in F^{\delta} \right)$$

$$= \sum_{\mathbf{k} \in K(j)} \mathbb{P}\left(\sum_{m=1}^{j} X_{k_m}^{(n)} / \lambda_n \in F^{\delta} \cap \mathbb{D}_{=j} \right)$$

$$= \sum_{\mathbf{k} \in K(j)} \mathbb{P}\left(\sum_{m=1}^{j} Z_{k_m} \mathbb{1}_{[k_m/n]}(t) / \lambda_n \in F^{\delta} \cap \mathbb{D}_{=j} \right)$$

$$= \sum_{1 \le k_1, \dots, k_j \le n} \mathbb{P}\left(\left(\frac{Z_{k_m}}{\lambda_n} \right)_{1 \le m \le j} \in J(F^{\delta}) \right) \mathbb{1}\left(\left(\frac{k_1}{n}, \dots, \frac{k_j}{n} \right) \in T(F^{\delta}) \right)$$

$$= \mathbb{P}\left((Z_1, \dots, Z_j) / \lambda_n \in J(F^{\delta}) \right) \sum_{1 \le k_1 < \dots < k_j \le n} \mathbb{1}\left(\left(\frac{k_1}{n}, \dots, \frac{k_j}{n} \right) \in T(F^{\delta}) \right).$$

Note that as $n \to \infty$,

$$\mathbb{P}\left(|Z_1| > \lambda_n \right)^{-j} \mathbb{P}\left((Z_1, \dots, Z_j) / \lambda_n \in J(F^{\delta}) \right) \to \nu_{\alpha}^j\left(J(F^{\delta}) \right), \tag{11}$$

which, in case $J(F^\delta)$ is a rectangle, follows from the assumption that the Z_i are i.i.d. regularly varying random variables. The general case then follows from the fact that ν_α has no discontinuity points and agrees with the limit measure on all rectangles (see, e.g., Resnick [35], Lemma 6.1 for a similar argument). Similarly, for $T(F^\delta)$, we obtain for $n \to \infty$,

$$n^{-j} \sum_{1 \le k_1 < \cdots < k_j \le n} \mathbb{1}\left(\left(\frac{k_1}{n}, \ldots, \frac{k_j}{n}\right) \in T(F^\delta)\right) \to \mathrm{Leb}_j\left(T(F^\delta)\right). \tag{12}$$

Hence using (11) and (12), we have as $n \to \infty$,

$$\gamma_n^{(j)} P_{j,1}^{(n)} \to \left(\nu_\alpha^j \times \mathrm{Leb}_j\right)\left(J(F^\delta), T(F^\delta)\right) = \left(\nu_\alpha^j \times \mathrm{Leb}_j\right) \circ h_j^{-1}(F^\delta).$$

Therefore,

$$\limsup_{n\to\infty} \gamma_n^{(j)} \mathbb{P}(B_j) \le \left(\nu_\alpha^j \times \mathrm{Leb}_j\right) \circ h_j^{-1}(F^\delta),$$

for $\delta > 0$. Summing up all the bounds we obtained, we have

$$\limsup_{n\to\infty} \gamma_n^{(j)} \mathbb{P}\left(X^{(n)}/\lambda_n \in F\right) \le \left(\nu_\alpha^j \times \mathrm{Leb}_j\right)\left(h_j^{-1}(F^\delta)\right).$$

Since $h_j^{-1}(F) = \bigcap_{\delta>0} h_j^{-1}(F^\delta)$, letting $\delta \to 0$ gives us the required upper bound

$$\limsup_{n\to\infty} \gamma_n^{(j)} \mathbb{P}\left(X^{(n)}/\lambda_n \in F\right) \le \left(\nu_\alpha^j \times \mathrm{Leb}_j\right) \circ h_j^{-1}(F).$$

Lower bound. Let G be open and bounded away from $\mathbb{D}_{\le(j-1)}$. Now define, $G^{-\delta} \subset G$,

$$G^{-\delta} = \left\{f \in G : d_{J_1}(f, g) < \delta \text{ implies } g \in G\right\}.$$

Choose δ small enough such that $G^{-\delta}$ is non-empty. It is still open and bounded away from $\mathbb{D}_{\le(j-1)}$. Searching for a lower bound, we shrink the set G to its bare minimum,

$$\mathbb{P}\left(X^{(n)}/\lambda_n \in G\right) \ge \sum_{1 \le k_1 < \cdots < k_j \le n} \mathbb{P}\left(\sum_{i=1}^j X_{k_i}^{(n)}/\lambda_n \in G^{-\delta}, \sup\left|X^{(n)} - \sum_{i=1}^j X_{k_i}^{(n)}\right| < \lambda_n \delta\right)$$

$$= \sum_{1 \le k_1 < \cdots < k_j \le n} \mathbb{P}\left(\sum_{i=1}^j X_{k_i}^{(n)}/\lambda_n \in G^{-\delta}\right) \mathbb{P}\left(\sup\left|X^{(n)} - \sum_{i=1}^j X_{k_i}^{(n)}\right| < \lambda_n \delta\right).$$

The second factor converges to one since $S_n/\lambda_n \to 0$ in probability as $n \to \infty$. For the first factor, we are able to proceed in the same fashion as we found the convergence rate of $P_{j,1}^{(n)}$ (defined in (10)) for the upper bound, to obtain the lower bound. $\quad\square$

Remark 8. Note that instead of our functions being in $\mathbb{D} := \mathbb{D}([0, 1], \mathbb{R})$, we can easily extend Theorem 3.5 to càdlàg functions in $\mathbb{D}_M = \mathbb{D}([0, M], \mathbb{R})$ for some number $M > 0$, with minor

modifications to the proof. Hence, all the results obtained in this section hold if we amend the definitions of the spaces $\mathbb{D}$, $\mathbb{D}_{=j}$, $\mathbb{D}_{\leq j}$ accordingly. Without loss of generality, we refer to these results as if they hold for $\mathbb{D}_M$ and its appropriate subsets from now on.

3.3. Random walks with a constant drift

The conclusion in Theorem 3.5 assumes that the random variables are centered. If the random variables driving the random walk have a finite but non-zero mean, a similar result holds. By setting $\lambda_n = n$ we are able to preserve the drift in the limit. Theorem 3.5 can be modified easily to incorporate such a drift term. For completeness and subsequent use, we state the result without proof as Corollary 3.6 below. A formal proof of this version of Theorem 3.5 with $\lambda_n = n$ and non-trivial drift can be found in Theorem 4.1 of Rhee, Blanchet and Zwart [36].

Corollary 3.6. *Let $\{Z_i\}_{i=1}^{\infty}$ be a sequence of i.i.d. random variables with $Z_1 \in \mathcal{RV}_{-\alpha}$, $\alpha > 1$. Denote $m = \mathbb{E}[Z_1]$ and define*

$$h_j^m : \left(\mathbb{R}\backslash\{0\}\right)^j \times U_j^{\uparrow} \to \mathbb{D},$$

$$h_j^m\big((z, \boldsymbol{u})\big)(t) := \sum_{i=1}^{j} z_i \mathbb{1}_{[u_i \leq t]} + mt,$$

and correspondingly $\mathbb{D}_{=j}^m := h_j^m(\mathbb{R}^j\backslash\{0\} \times U_j^{\uparrow})$. Then, as $n \to \infty$,

$$\gamma_n^{(j)}\mathbb{P}\big(X^{(n)}/n \in \cdot\big) \to \big(\nu_\alpha^j \times \mathrm{Leb}_j\big) \circ \big(h_j^m\big)^{-1}(\cdot),$$

in $\mathbb{M}(\mathbb{D}\backslash\mathbb{D}_{\leq(j-1)}^m)$.

Remark 9. The space $\mathbb{D}_{=j}^m$ is defined as the space of step functions with exactly j jump discontinuities and a constant drift term "mt". In particular, $\mathbb{D}_{=0}^m = \{x(t) = mt, t \in [0, M]\}$. Theorem 3.5 allowed for scalings λ_n that grow fast enough such that $X^{(n)}/\lambda_n$ stays close to zero for large n. Note that in Corollary 3.6 we restrict to $\alpha > 1$ and specify $\lambda_n = n$ to preserve the drift term. Necessarily, we observe for sets A bounded away from $\mathbb{D}_{=0}^m$ that $\mathbb{P}(X^{(n)}/n \in A) \to 0$ as $n \to \infty$. Hence, we examine (a sequence of) large deviation type limits on $\mathbb{D}\backslash\mathbb{D}_{\leq j-1}^m$. This result becomes particularly applicable in the queueing context we discuss next.

4. Application to finite buffer queues

In this section, we apply the results of Theorem 3.5 and Corollary 3.6 to the modified Lindley recursion; see (13) below. This formula is usually interpreted as describing the evolution of the queue length in a queue with finite buffer. First we derive a result on large deviations for what we call long intense periods, defined as the maximum time a queue-size process spends continuously above a certain threshold. These large deviations concern limits where both the threshold level

and the buffer size approach infinity while the arrival process is sped up appropriately. Secondly, we present a simulation study which combines two of the derived large deviation type limits to provide a simple analytical approximation and explanation for the empirical distribution of extremal lengths of long intense periods.

4.1. Queueing processes

We study recursions of the form

$$Q_n^K = \min\{\max\{Q_{n-1}^K + A_n - C_n, 0\}, K\}, \tag{13}$$

for $n \geq 1$ with $Q_0 \geq 0$ where $\{A_n\}_{n\geq 1}$ and $\{C_n\}_{n\geq 1}$ are two sequences of i.i.d. non-negative random variables. This is the modified version of Lindley's recursion (Lindley [27]) to accommodate queues with finite buffers of size K. The recursion in (13) can be interpreted in many different ways. For example, in the context of network traffic, A_n may be interpreted as the number of packets arriving in the time interval $C_n - C_{n-1}$, whereas Q_{n-1} describes the amount of work previously in the buffer of a single server processing work at a fixed rate. Any number of packets arriving at a full buffer are immediately discarded. For example, Jelenković [26] studies (13) under the assumption that $\int_0^x \mathbb{P}(A_1 > z)\,dz/\mathbb{E}A_1$ follows a subexponential distribution to conclude that the stationary loss rate is essentially due to one large observation when the buffer size approaches infinity. Denote Q^K the stationary queue length. Then

$$\mathbb{E}\big[(Q^K + A_1 - C_1 - K) \vee 0\big] = \mathbb{E}\big[(A_1 - K) \vee 0\big]\big(1 + o(1)\big), \quad K \to \infty.$$

Sample path large deviation principles for queueing processes with both infinite and finite buffers are studied in Ganesh, O'Connell and Wischik [19] mostly under the assumption that the moment generating function exists. We work with regularly varying random variables $A_1 \in \mathcal{RV}_{-\alpha}, \alpha > 0$ throughout which do not satisfy this assumption.

To study the queueing recursion (13), we follow the continuous mapping approach. First, we define a suitable embedding of the sequences $\{A_n\}$ and $\{C_n\}$ in the space $\mathbb{D}_M$ and then employ a continuous map to obtain a process which agrees with the queueing recursion at specified discrete time stamps. See, for example, Whitt [38], Asmussen [3] or Andersen *et al.* [1] for more on this approach applied to queueing processes. This map is usually called a reflection map or the Skorohod map. We briefly recall the required results. For a process $x \in \mathbb{D}_M$ with $x(0) = 0$, we call $\{v(t), l(t), u(t)\}$ the solution to the Skorohod problem if $v(t) \in [0, K]$ and

$$v(t) = x(t) + l(t) - u(t), \qquad \int_0^\infty v(t)\,dl(t) = 0, \qquad \int_0^\infty \big(K - v(t)\big)\,du(t) = 0,$$

and both l, u are non-negative non-decreasing functions. Denote $\psi_0^K : \mathbb{D}_M \to \mathbb{D}_M$ the reflection map on the interval $[0, K]$ as

$$\psi_0^K : x \mapsto v, \tag{14}$$

where υ denotes the resulting regulated process of the Skorohod problem. This map is (Lipschitz-) continuous on $\mathbb{D}_M$ equipped with the J_1 metric, see Andersen *et al.* [1], Lemma 4.6. To facilitate the discussion let $C_n = c$, $n \geq 1$ for some $c > 0$ and denote

$$A(t) := \sum_{i=1}^{\infty} (A_i - c)\mathbb{1}_{[t \geq i]}, \quad t \geq 0$$

the embedding of the random walk induced by $A_n - c$ in $\mathbb{D}_M$. Then

$$Q^K := \psi_0^K(A) \tag{15}$$

is an embedding of Q_n^K into $\mathbb{D}_M$ satisfying $Q^K(t) = Q_t^K$ for $t \in \mathbb{N}_0$. Consequently we call Q^K a queueing process with buffer K.

Remark 10. We could also work with other embeddings, for example,

$$B(t) := \sum_{n=1}^{\infty} A_n \mathbb{1}_{[t \geq n]} - ct, \quad t \geq 0,$$

and define $Q_B^K := \psi_0^K(B)$ to allow for more nuanced interpretations of the queueing process Q. But since this work focuses on scaled versions of the queueing process with both time t and space $Q^K(t)$ scaled appropriately, the exact form of the interpolation is mostly irrelevant for the limit.

4.2. Long intense periods

We adopt the position that a queueing process with the queue size close to the buffer K corresponds to an undesirable state. In such a state, the service quality (of which $Q^K(t)$ is a proxy) is perceived as suboptimal. In the following, we introduce and study the longest period an observed queueing process spends above a certain threshold θK during the observation horizon $[0, M]$. We call such intervals long intense periods and investigate their length.

Definition 4.1 (Long intense period). For a càdlàg function $x \in \mathbb{D}_M$ and a fixed level $\eta \in \mathbb{R}_+$ we define

$$L^{\eta} : \mathbb{D}_M \to \mathbb{R}_+,$$
$$x \mapsto \sup_{0 \leq s < t \leq M} \{t - s : x(u) > \eta \; \forall u \in (s, t)\}. \tag{16}$$

For a queueing process Q with buffer K we call $L^{\theta K}(Q^K)$ the length of the intense period at level $\theta \in (0, 1)$.

How useful is it to calculate large deviations for long intense periods in queues? Can we gain more insight into waiting times in queues with this information? To illustrate the applicability of

such results, we use a simulation study which investigates the distribution of long intense periods for large threshold levels $\theta \in (0, 1)$.

Example 4.2. The object of our simulation study is a queueing process $Q^K(t)$ with $N = 50{,}000$ arrival variables following a power law distribution with tail index $\alpha = 1.44$ and expectation $m = 0.5$. The queue has a finite buffer $K = 20{,}000$ and any additional service requirements will be lost. We assume the server works at a fixed rate $c = 1$ with A_i describing the amount of service requirements arriving in one unit of time. We study long intense periods above the level $\theta = 0.85$, that is $Q^K(t) > 17{,}000$ is considered intense. The queueing process is observed on $[0, M]$ with $M = N$. In this example, we treat service time distributions that still have finite means but infinite variance. The particular α value corresponds to the tail parameter of file sizes in Internet traffic reported in Jelenković and Momčilović [25]. To be precise, we consider the arrival distribution

$$\mathbb{P}(A_1 > z) = \left(\frac{z}{(\alpha - 1)m} + 1 \right)^{-\alpha}, \quad z > 0.$$

Clearly, $\mathbb{P}(A_1 > z) \in \mathcal{RV}_{-\alpha}$.

Figure 1 contains a histogram of the realized lengths of the long intense periods in queueing processes with the parameters stated above. It is based on 22,000 observations which exhibit a strictly positive long intense period, which means, Figure 1 shows a histogram of $L^{\theta K}(Q)$ conditioned on $L^{\theta K}(Q) > 0$. We would like to understand the shape of the histogram that we observe here; why is there a peak in the middle and a decay afterwards? We revisit the histogram
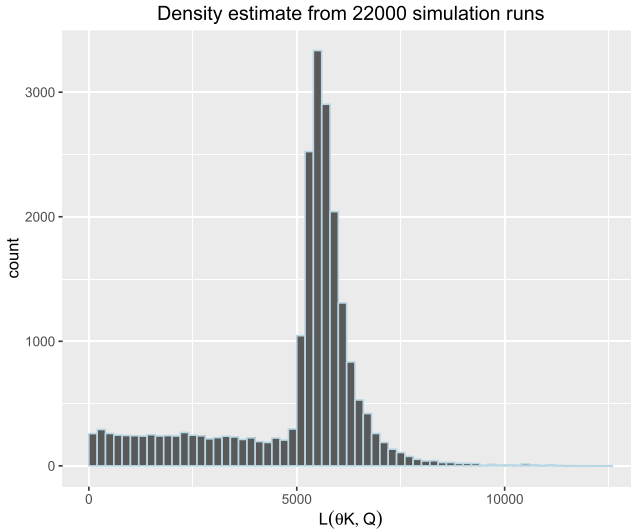

Figure 1. Histogram for $L^{\theta K}(Q)|L^{\theta K}(Q) > 0$ generated by 22,000 realizations of a queueing process with 50,000 arrivals at integer time points and maximum capacity $K = 20{,}000$. The critical level was set to $\theta = 0.85$. The red vertical line marks the location of the theoretical point mass at $L = \frac{1-\theta}{c-m} K$. The service time distribution follows an exact power law with $-\alpha = -1.44$ and mean $1/2$.

at the end of this section, accompanied by an explanation for its shape, based on (hidden) large deviations.

4.3. Large deviations for long intense periods

We work out the corresponding sequence of LDLs for long intense periods of queueing processes.

Theorem 4.3. *Let $\{A_i\}_{i\geq 1}$ be a sequence of i.i.d. non-negative regularly varying random variables with $A_1 \in \mathcal{RV}_{-\alpha}$, $\alpha > 1$. Assume $c > m := \mathbb{E}[A_1]$ and define the queueing process $Q^{K,(n)}(t) := Q^{nK}(nt)/n$, $t \in [0, M]$, $n \geq 1$ with Q^K defined as in (15). Denote $\kappa := \frac{1-\theta}{c-m}K$. The intense periods $L_n := L^{\theta K}(Q^{K,(n)})$ of the queueing process $Q^{K,(n)}$ observed on $[0, M]$ satisfy a sequence of LDLs on $[0, M]\backslash[0, (j-1)\kappa]$ with the limit measure $\mu_L^{(j)}$ concentrating it's mass on $((j-1)\kappa, j\kappa]$. Specifically,*

$$L_n \in \mathrm{LD}\big(\gamma_n^{(j)}, \mu_L^{(j)}, [0, M]\backslash[0, (j-1)\kappa]\big), \quad 1 \leq j \leq \left\lfloor \frac{M}{\kappa} \right\rfloor,$$

where the limit measure is given by

$$\mu_L^{(j)} = \big(\nu_\alpha^j \times \mathrm{Leb}_j\big) \circ \big(h_j^{m-c}\big)^{-1} \circ \big(\psi_0^K\big)^{-1} \circ \big(L^{\theta K}\big)^{-1}.$$

Remark 11. The assumption $c > \mathbb{E}[A_1]$ ensures that the process drifts in the negative direction on an average, such that the process being close to its buffer is actually a rare event. At the first level for $j = 1$, the theorem states that the long intense periods of a queueing process with buffer K and negative drift may be approximated by summing over all one-jump functions that contain a jump of size at least θK. Since the measure concentrates on one-jump functions, the maximum attainable long intense period is attained by a single jump that exceeds the buffer limit K, with the process drifting in the negative direction at a rate $m - c$ afterwards. Thus, no matter the size of the jump, the process will leave the intense region at most κ time units after the jump.

Remark 12. Measuring the longest connected interval of time spent above a certain threshold is not a continuous operation for càdlàg processes. For example, consider for $M > 2$ the function $x \in \mathbb{D}_M$ such that

$$x(t) := \begin{cases} 1 - t & \text{if } t \in [0, 1), \\ 2 - t & \text{if } t \in [1, M]. \end{cases}$$

Adding a small constant via $\phi_c(x)(t) := x(t) + c$ we obtain for $c < 0$: $L^0(\phi_c(x)) = (1 - |c|) \wedge 0$ but $L^0(x) = L^0(\phi_0(x)) = 2$, while at the same time $\phi_c(x) \to x$ as $c \to 0$. Consequently, L is not continuous. Nevertheless $L^{\theta K}$ is continuous almost everywhere with respect to the limit measure $\mu = \nu_\alpha \times \mathrm{Leb}_1 \circ (h_1^{m-c})^{-1} \circ (\psi_0^K)^{-1}$ on $\mathbb{D}_M$ as the only way to obtain a discontinuity is through the jump at the end of the long intense interval. But the jump position is uniformly distributed hence the measure of that set is zero. Additionally, for our purposes, there is no need to consider functions outside the support of μ.

We need the following lemma to prove Theorem 4.3.

Lemma 4.4. *Let $j \in \mathbb{N}$. Denote*

$$E_j := \left\{ x \in \mathbb{D}_M : \big[x(t) = 0\big] \text{ OR } \big[x(t+s) = x(t) - s(c-m), |s| \text{ small enough}\big] \right.$$
$$\left. \text{for all but } j \text{ points } t. \text{ Additionally } x(t) \geq x(t-) \; \forall t \in [0, M] \right\},$$

$$D_j := \left\{ x \in E_j : \exists t \in \{\text{discontinuity points of } x\} \right.$$
$$\left. \text{such that } x(t-) = \theta K \text{ OR } x(t) = \theta K \text{ OR } t \in \{0, M\} \right\}.$$

Then $L^{\theta K}$ is continuous on $E_j \backslash D_j$.

Remark 13. Note that E_j contains all càdlàg functions which contain exactly j positive jumps and decrease at a rate $c - m$ otherwise, regulated to take values in $[0, K]$. The set D_j further restricts to those functions whose jumps are bounded away from the critical level θK.

Proof of Lemma 4.4. To show the claim, we need to introduce additional notations. We define an intense period as a period during which the function $x \in E_j$ stays continuously above the critical level θK and enumerate all such periods. Subsequently, we show that the length of each such period cannot change much in case x is not perturbed too much. Define

$$_s L(l, x, v), {}_t L(l, x, v) : \mathbb{R}_+ \times \mathbb{D}_M \times [0, M] \to \mathbb{R}_+$$
$$_s L(l, x, v) := \inf\big\{ u \in (v, M] : x(u) > l \big\},$$
$$_t L(l, x, v) := \inf\big\{ u \in ({}_s L(l, x, v), M] : x(u) < l \text{ OR } u = M \big\}.$$

We assume, as is usually the case, that $\inf \varnothing = \infty$. Next, we recursively record the start and end times of what we call intense periods, starting at zero.

$$s_1 := {}_s L(l, x, 0), t_1 := {}_t L(l, x, 0),$$
$$s_i := {}_s L(l, x, t_{i-1}), t_i := {}_t L(l, x, t_{i-1}), \quad i \geq 2,$$
$$n_x := \max\{i : s_i < \infty\}.$$

In case the tuple s_i, t_i are finite, we call $t_i - s_i$ the length of the ith intense period of x. Note that for $x \in E_j$ there are exactly n_x intense periods, with $0 \leq n_x \leq j$. The length of the longest of these corresponds to what we defined above in (16) as the length of the long intense period of x.

Let $x \in E_j \backslash D_j$. Then the set of time points at which x is above the critical level can be partitioned as

$$\big\{ u : x(u) > \theta K \big\} = \bigcup_{i=1}^{n_x} [s_i, t_i).$$

Since all jump discontinuities of x have values bounded away from the critical level, all of the intervals $[s_i, t_i)$ and $[t_i, s_{i+1})$ are of positive length. Moreover, denoting

$$\Delta_J^{\theta K}(x) := \min\{|x(u^-) - \theta K| \wedge |x(u) - \theta K| \wedge u \wedge M - u : u \text{ is a discontinuity point of } x\},$$

we find that for all $0 < \delta < \Delta_J^{\theta K}(x)$,

$$x(u) \in (\theta K - \delta, \theta K + \delta)$$
$$\Leftrightarrow \quad u \in \left(t_i - \frac{\delta}{c - m}, t_i + \frac{\delta}{c - m}\right) \cap [0, M] \quad \text{for some } 1 \leq i \leq n_x. \tag{17}$$

Next, we show that for all $\varepsilon > 0$, small enough such that $\varepsilon(c - m) < \Delta_J^{\theta K}(x)$, there exists a ζ such that whenever $d_{J_1}(x, y) < \zeta$ we have

$$u \in [s_i + \varepsilon, t_i - \varepsilon) \quad \Rightarrow \quad y(u) > \theta K,$$
$$u \in [t_i + \varepsilon, s_{i+1} - \varepsilon) \quad \Rightarrow \quad y(u) < \theta K. \tag{18}$$

This implies that any y close enough to x has similar intense periods as x, ignoring any negligible intense periods of y. Hence, $L^{\theta K}$ is continuous at x in $(\mathbb{D}_M, d_{J_1})$. We proceed by showing that the above claim holds for $\zeta = \frac{\varepsilon((c-m)\wedge 1)}{3}$. Then there exists a $\lambda \in \Lambda$ such that

$$\|x - y \circ \lambda\| < \frac{\varepsilon((c - m) \wedge 1)}{2}, \tag{19}$$

$$\|\lambda - e\| < \frac{\varepsilon((c - m) \wedge 1)}{2}. \tag{20}$$

Now (19) combined with (17) (where $\delta = \varepsilon(c - m)/2$) implies

$$u \in [s_i, t_i - \varepsilon/2] \quad \Rightarrow \quad (y \circ \lambda)(u) > \theta K + \frac{\varepsilon(c - m)}{2} - \frac{\varepsilon((c - m) \wedge 1)}{2} \geq \theta K,$$
$$u \in (t_i - \varepsilon/2, s_{i+1}) \quad \Rightarrow \quad (y \circ \lambda)(u) < \theta K. \tag{21}$$

Accounting for the time change introduced through λ, we infer from (20) that the last two implications in (21) hold when the two intervals get reduced by a further $\varepsilon/2$ on each side. In turn, this proves the statement in (18) and thus implying the continuity of $L^{\theta K}$ on $E_j \backslash D_j$ for all $j \geq 1$. $\square$

Proof of Theorem 4.3. We apply the continuous mapping argument in Theorem 2.2 twice. First, using the Skorohod map of (14), the large deviations result in Corollary 3.6 and continuous mapping yield

$$\gamma_n^{(j)} \mathbb{P}\big(Q^{nK}(nt)/n \in \cdot\big) \rightarrow (\nu_\alpha \times \text{Leb}_j) \circ \big(h_j^{m-c}\big)^{-1} \circ \big(\psi_0^K\big)^{-1}(\cdot), \quad n \rightarrow \infty \tag{22}$$

in $\mathbb{M}(\mathbb{D}_M \backslash \psi_0^K((\mathbb{D}_M)_{\leq j-1}^{m-c}))$. This is due to the definition in (14) satisfying $\psi_0^K(x/n) = \psi_0^{nK}(x)/n$. The Lipschitz continuity of ψ_0^K ensures that the "bounded away" condition of Theorem 2.2 is satisfied.

Next, note that the limit measure in (22) concentrates all its mass on E_j. Additionally, note that $\mu_L^{(j)}(D_j) = 0$, $j \geq 1$ as $(h_j^{m-c})^{-1} \circ (\psi_0^K)^{-1}(D_j) \subset \mathbb{R}^{2j}$ is not of full dimension; the condition of having $x(t)$ or $x(t^-) = \theta K$ amounts to imposing restrictions linking the time and value of a jump through relations of the form

$$x(t_{i-1}) - (t_i - t_{i-1})(c - m) + J(t_i) = \theta K,$$

$$\text{OR } x(t_{i-1}) - (t_i - t_{i-1})(c - m) = \theta K,$$

$$\text{OR } J(t_i) = \theta K, x\left(t_i^-\right) = 0.$$

where t_i denotes the time of the ith jump and $J(t_i)$ denotes the corresponding size of the jump. Hence, Lemma 4.4 above combined with another application of Theorem 2.2 (continuous mapping) yields the result. $\qquad\square$

4.4. Calculating explicit limit measures

In the following, we compute the limit measures $\mu_L^{(1)}$ and $\mu_L^{(2)}$ explicitly. Assume $M > 2\kappa$ throughout. For $j = 1$, we obtain

$$\mu_L^{(1)}\big((l, \infty)\big) = \begin{cases} (M - l)\big(l(c - m) + \theta K\big)^{-\alpha} & \text{if } l \in (0, \kappa] \\ 0 & \text{otherwise.} \end{cases}$$

In other words, the measure $\mu_L^{(1)}$ is the sum of a point mass at $l = \kappa$ with value $K^{-\alpha}(M - \kappa)$ and an absolutely continuous part on $(0, \kappa)$. Considering this initial large deviations estimate on its own we would approximate $\mathbb{P}(L^{\theta K}(Q^K) > \kappa) \approx 0$. For any finite buffer non-limit scenario this may be too coarse. A more refined estimate based on hidden large deviations allows for more accuracy. Namely on $[0, M] \setminus [0, \kappa]$, we have

$$\gamma_n^{(2)} \mathbb{P}\big(L^{\theta n K}\big(Q^{nK}(nt)\big) \in \cdot\big) \to \mu_L^{(2)}(\cdot), \quad n \to \infty,$$

in $\mathbb{M}([0, M] \setminus [0, \kappa])$, which concentrates on $(\kappa, 2\kappa]$. This again can be explained by the rate $\gamma_n^{(2)}$ only allowing for at most two jumps in the random walk. Any intense period with length $L < \kappa$ is more likely to happen due to one jump; hence processes containing only one jump must be excluded in the hidden large deviation limit. Long intense periods with length $L > 2\kappa$ are not possible since the maximum length is achieved if the buffer is filled at some initial time $t_0 < M - 2\kappa$ starting the long intense period and an additional jump at time $t_0 + \kappa$ of size at least $(1 - \theta)K$ occurs. We compute the limit measure for the events $\{L > l\}$, $l \in (\kappa, 2\kappa]$:

$$\mu_{L, \theta, K}^{(2)}\big((l, \infty)\big) = \mu^{(2)}\big(\{\text{All two jump functions with } L > l\}\big)$$

$$= (M - l) \int_{\theta K}^{\infty} \nu_\alpha(dj_1) \int_{l - \kappa}^{(K \wedge j_1 - \theta K)/(c - m)} du_2 \int_{l(c - m) - (K \wedge j_1 - \theta K)}^{\infty} \nu_\alpha(dj_2)$$

$$
\begin{aligned}
&= (M - l) \int_{\theta K}^{\infty} \nu_\alpha(\mathrm{d}j_1) \mathbb{1}_{\left[\frac{K \wedge j_1 - \theta K}{c-m} > l - \kappa\right]} \frac{\frac{K \wedge j_1 - \theta K}{c-m} - (l - \kappa)}{(l(c-m) - (K \wedge j_1 - \theta K))^\alpha} \\
&= \frac{M - l}{c - m} \int_{\theta K + l(c-m) - (1-\theta)K}^{K} \alpha x^{-\alpha-1} \frac{x - \theta K - l(c-m) + (1-\theta K)}{(l(c-m) + \theta K - x)^\alpha} \, \mathrm{d}x \\
&\quad + \frac{M - l}{c - m} K^{-\alpha} \frac{2(1-\theta)K - l(c-m)}{(l(c-m) - (1-\theta)K)^\alpha}.
\end{aligned}
\tag{23}
$$

Remark 14. Further limit measures can be computed but the explicit derivation becomes more cumbersome as the level increases; we leave this to the user's discretion to compute according to their chosen precision level.

4.5. Simulation study – combining the first two LDLs

In this section, we provide some insights on the practical relevance of hidden large deviations. We find that in the setting of the simulation study described in Example 4.2 we are able to numerically validate the rate and limit measure of hidden large deviations. The previous section established large deviation limits for long intense periods for any interval $[(j-1)\kappa, j\kappa]$ with $j \leq \lfloor M/\kappa \rfloor$, each with its own rate. And indeed, for the theory of LDLs we may only treat these limit measures separately due to the different magnitudes of the rates $\gamma_n^{(j)}$. In practice however, for any finite observation period of a queue with finite buffer size, several of the limit measures might be relevant for a single statistic. The simulation study will examine the interplay of different rates in a single probability estimate. We proceed to construct the two estimates involving the first and second level LDLs separately.

One jump. According to traditional large deviation estimates for heavy tailed queueing processes with "large" buffers, the long intense period will be due to a single jump reaching above the threshold level θK and the queue drifting in direction $-(c-m)$ thereafter. To use Theorem 4.3, we need to choose a queue sequence number n. We thus obtain the following approximation.

$$
\begin{aligned}
\mathbb{P}\big(L^{\theta K}\big(\boldsymbol{Q}^K\big) > l\big) &= \mathbb{P}\big(L^{\theta K/n}\big(\boldsymbol{Q}^{K/n,(n)}\big) > l/n\big) \\
&\approx \frac{1}{\gamma_n^{(1)}} \mu_L^{(1)}\big((l/n, \infty)\big) \\
&= \begin{cases} n^\alpha \mathbb{P}(A_1 > n)(M - l)\big(l(c-m) + \theta K\big)^{-\alpha} & \text{if } l \in (0, \kappa], \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

The point mass at κ yields

$$
\mathbb{P}\big(L^{\theta K}\big(\boldsymbol{Q}^K\big) \in (\kappa - \varepsilon, \kappa + \varepsilon)\big) \approx \left(\frac{n}{K}\right)^\alpha \mathbb{P}(A_1 > n)(M - \kappa).
$$

One or two jumps. The *two jump measure* can be approximated in the same fashion as the approximation for one jump above using equation (23) instead. To get a single estimate for the
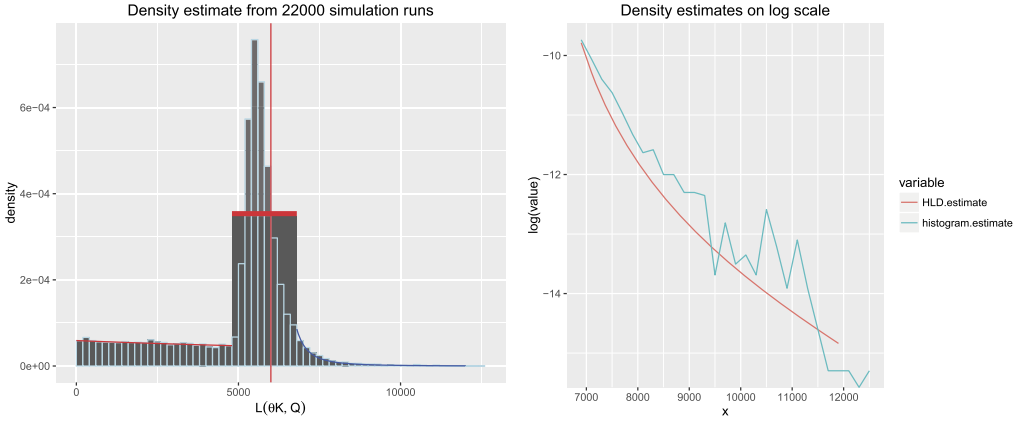
Figure 2. Left: Histogram for $L^{\theta K}(\boldsymbol{Q})|L^{\theta K}(\boldsymbol{Q}) > 0$ generated by 22,000 realizations of a queueing process with 50,000 arrivals at integer time points and maximum capacity $K = 20,000$. The critical level was set to $\theta = 0.85$. The red vertical line marks the location of the theoretical point mass at $L = \frac{1-\theta}{c-m}K$. The service time distribution follows an exact power law with $-\alpha = -1.44$ and mean $1/2$. Approximations of the densities with the large deviations estimate (in red) and the hidden large deviations estimate (in blue) are added on top of the histogram. Right: Same data as on the left restricted to $L > 7000$. Density estimate with HLD compared to empirical values of the histogram on logarithmic scale.

distribution of the long intense periods of the queueing process $\boldsymbol{Q}^K$, we propose to combine the two estimates into a single approximation.

$$\mathbb{P}\big(L^{\theta K}\big(\boldsymbol{Q}^K\big) > l\big) \approx \begin{cases} n\mathbb{P}(A_1 > n)\mu_L^{(1)}\big((l/n, \infty)\big) & \text{if } l \in (0, \kappa] \\ \big[n\mathbb{P}(A_1 > n)\big]^2 \mu_L^{(2)}\big((l/n, \infty)\big) & \text{if } l \in (\kappa, 2\kappa] \\ 0 & \text{otherwise,} \end{cases} \tag{24}$$

where the buffer size K and observation horizon M are scaled accordingly in the limit measures.

In Figure 2, we plot the same histogram as in Figure 1, and view it as an estimate of the density

$$\mathbb{P}\big(L^{\theta K}\big(\boldsymbol{Q}^K\big) \in \mathrm{d}l | L^{\theta K}\big(\boldsymbol{Q}^K\big) > 0\big), l > 0.$$

The limit measure $\mu_L^{(1)}$ puts zero mass on values beyond the vertical red line which marks the location of the point mass of $\mu_L^{(1)}$. Since n is finite, we expect some values immediately to the right of the point mass caused by only a finite number of random variables approximating the mean rate of decrease for the queue content. Nevertheless, concerning the values on the far right we believe an explanation via Hidden Large deviations (HLD) is best suited for the distribution of $L^{\theta K}(\boldsymbol{Q}^K)$. Hence, we add the estimate in (24) to the plot. To visualize the point mass, we fix two $\varepsilon_1, \varepsilon_2 > 0$ such that the point mass at κ gets distributed over the area $(\kappa - \varepsilon_1, \kappa + \varepsilon_2)$. Outside of this region we approximate the measure with the corresponding densities. Additionally, we provide a plot of the tail of the distribution on a log scale to better visualize the fit for the hidden large deviation estimate. The figures clearly show how our hidden large

deviation estimates closely approximate the histogram observed; it becomes more evident in the right-hand plot of Figure 2 which is plotted on a log-scale.

5. A random walk with very heavy-tailed innovations

In general, for processes driven by innovations $Z_i \in \mathcal{RV}_{-\alpha}$ with $\alpha > 1$ one may always be able to choose a scaling $\lambda_n = cn$ (as in Theorem 3.5) for some constant $c > 0$ for computing large deviation probabilities. The application in Section 4 was a special case which required setting $\lambda_n = n$, meaning $\rho = 1$, in Theorem 3.5, so that we could retain the drift structure of the queue while obtaining the limit probability of congestion events. Interestingly, our result actually allows us to compute probabilities where $\lambda_n \neq n$. This of course is useful if $\alpha \leq 1$ implying that S_n/n does not converge to 0 necessarily where $S_n = Z_1 + \cdots + Z_n$ is the partial sum process. We illustrate with a concrete example as follows. Suppose $\{Z_i\}_{i \geq 1}$ are i.i.d. random variables with

$$\mathbb{P}(Z_1 > x) = \mathbb{P}(Z_1 < -x) = \frac{1}{2}x^{-1/2}, \quad x \geq 1.$$

Clearly, $Z_i \in \mathcal{RV}_{-\alpha}$ where $\alpha = 1/2$. Moreover, $\mathbb{E}[|Z_i|] = \infty$. Hence, Theorem 3.5 is applicable, but requires a sequence $\lambda_n \in \mathcal{RV}_\rho$ with $\alpha\rho = \rho/2 > 1$ to obtain a proper limit distribution in (9). Let $\lambda_n = n^3$, $n \geq 1$, meaning $\rho = 3$. Applying Theorem 3.5 for a fixed $j \geq 1$, we get:

$$n^{j/2}\mathbb{P}\left(\frac{X^{(n)}(t)}{n^3} \in \cdot\right) \to \left(v_{1/2}^j \times \mathrm{Leb}_j\right) \circ h_j^{-1}(\cdot),$$

as $n \to \infty$ in the appropriate space. Let us take as examples two particular types of limit sets with a fixed parameter $\ell > 0$:

$$A_1 := \left\{x \in \mathbb{D} : \sup_{0 \leq t \leq 1} x(t) > \ell\right\};$$

$$A_2 := \left\{x \in \mathbb{D} : x \in A_1 \text{ and } \inf_{0 \leq t \leq 1} x(t) < -\ell\right\}.$$

Referring back to Remark 5, we know that $h_i^{-1}(A_i) = h_i^{-1}(A_i \cap \mathbb{D}_{=i})$. For $i = 1$, the set $A_1 \cap \mathbb{D}_{=1}$ corresponds to the set of all one jump functions with positive jump size larger than ℓ. For $i = 2$, the set $A_2 \cap \mathbb{D}_{=2}$ consists of all two jump functions with one positive and one negative jump of size at least ℓ such that the sum of the two jumps is larger than ℓ in absolute value. It is of course possible to compute probabilities of events governed by three or more large jumps; but we restrict to at most two jumps for illustrative purposes. An application of Theorem 3.5 now yields the following approxima-

tions.

$$\text{For } j = 1: \quad \mathbb{P}\big(X^{(n)}(t) \in n^3 \times A_1\big) \approx n^{-1/2} \times \big(\nu_{1/2}(l, \infty) \times \text{Leb}_1[0, 1]\big)$$

$$= \frac{1}{2\sqrt{nl}}, \quad \text{and,}$$

$$\text{for } j = 2: \quad \mathbb{P}\big(X^{(n)}(t) \in n^3 \times A_2\big) \approx n^{-2/2} \times \text{Leb}_2\big(U_2^\uparrow\big)$$

$$\times 2 \iint_{x_1 > l, x_2 < -x_1 - 2l} d\nu_{1/2}(x_2)\, d\nu_{1/2}(x_1) \qquad (25)$$

$$= \frac{1}{2n} \times \frac{1}{2} \int_l^\infty (x_1 + 2l)^{-1/2} \frac{1}{2} x_1^{-3/2} \, dx_1$$

$$= \frac{\sqrt{3} - 1}{8nl}.$$

These values obtained can also be verified using simulation. We use Monte Carlo simulation with $N = 2 \times 10^5$ realizations of random walks with n steps each to get a probability estimate which is averaged over $R = 100$ repetitions to yield the probability point estimates presented in Table 1 for A_1, A_2 and different values of n and ℓ. In general, regular Monte Carlo estimates of such large deviation events in a heavy-tailed regime often introduce a certain bias; efficient computation of probabilities in such rare event scenarios is of interest and has been addressed in the literature; see Asmussen, Binswanger and Højgaard [4], Blanchet and Liu [8]. We observe that our simulation estimates are pretty close to the theoretical large deviation estimates and hence we do not pursue a further refined technique for simulation here. Also note that further large deviation events relating to three or more jumps can be computed using formulas similar to (25), but computing Monte Carlo estimates would require a larger value of N.

Table 1. Monte Carlo and large deviation estimates for scaled probabilities of sets A_1 and A_2

n	ℓ	$p_{\text{LD}}(A_1)$	$\hat{p}(A_1)$	$\hat{\sigma}(p(A_1))$	$p_{\text{LD}}(A_2)$	$\hat{p}(A_2)$	$\hat{\sigma}(p(A_2))$
200	0.5	0.0500	0.0488	0.00048	0.000915	0.001014	0.000065
200	1.0	0.0354	0.0347	0.00035	0.000458	0.000498	0.000046
200	2.0	0.0250	0.0247	0.00036	0.000229	0.000253	0.000038
500	0.5	0.0316	0.0311	0.00040	0.000366	0.000405	0.000054
500	1.0	0.0224	0.0221	0.00037	0.000183	0.000207	0.000033
500	2.0	0.0158	0.0157	0.00027	0.000092	0.000102	0.000022
1000	0.5	0.0224	0.0221	0.00032	0.000183	0.000207	0.000029
1000	1.0	0.0158	0.0157	0.00030	0.000092	0.000104	0.000020
1000	2.0	0.0112	0.0111	0.00023	0.000046	0.000050	0.000016

Here $\hat{p}(A_j)$ denotes the estimated probability for $\mathbb{P}(X^{(n)} \in n^3 A_j(l))$, $j = 1, 2$ from $R = 100$ repetitions with $\hat{\sigma}(p(A_j))$ denoting the estimated standard deviation based on $R = 100$ computations of $\hat{p}(A_j)$; and $p_{\text{LD}}(A_j)$ denotes the approximated value obtained in (25).

6. Conclusion and further remarks

We provide limit measures for successively rarer large deviations of random walks with regularly varying i.i.d. increments. By scaling time and space appropriately we are able to obtain limit measures for large deviations of queueing processes which preserve the drift term in the limit. We have exhibited that hidden large deviations at the second level, though happening at the squared rate of the first large deviation, are numerically observable. Our simulation example exhibits that hidden large deviation estimates in conjunction with regular heavy-tailed large deviation estimates perform quite well in approximating the behavior observed in the data histogram – even at the tail (on a log scale). We have also computed large deviation probabilities for random walk processes without a finite mean and shown the accuracy of our estimates via Monte Carlo simulation.

For future directions of study, one may explore large deviations on a space $\mathbb{D}\setminus\bigcup_{j=1}^{\infty}\mathbb{D}_{=j}$ which we have not ventured into, mostly since the exact structure of such deviations remains largely an open question. Similarly, we have not explored situations where the "i.i.d." assumption is relaxed. A jth level LDL happens at a rate which is the jth power of the rate of the first LDL. This clearly is a consequence of the independence among the random variables driving the random walk. Researchers have looked into large deviations for regular variation under dependence (Mikosch and Wintenberger [30]) and hidden regular variation with asymptotic independence (not under a large deviation scaling) (Janssen and Drees [24]), but such results are still under investigation for a large deviation scaling assuming hidden regular variation.

Acknowledgements

References

[1] Andersen, L.N., Asmussen, S., Glynn, P.W. and Pihlsgård, M. (2015). Lévy processes with two-sided reflection. In *Lévy Matters*. *V. Lecture Notes in Math.* **2149** 67–182. Cham: Springer. MR3467345 https://doi.org/10.1007/978-3-319-23138-9_2

[2] Anderson, P.L. and Meerschaert, M.M. (1998). Modeling river flows with heavy tails. *Water Resour. Res.* **34** 2271–2280.

[3] Asmussen, S. (2003). *Applied Probability and Queues*: *Stochastic Modelling and Applied Probability*, 2nd ed. *Applications of Mathematics* (*New York*) **51**. New York: Springer. MR1978607

[4] Asmussen, S., Binswanger, K. and Højgaard, B. (2000). Rare events simulation for heavy-tailed distributions. *Bernoulli* **6** 303–322. MR1748723 https://doi.org/10.2307/3318578

[5] Bennett, G. (1962). Probability inequalities for the sum of independent random variables. *J. Amer. Statist. Assoc.* **57** 33–45.

[6] Billingsley, P. (1999). *Convergence of Probability Measures*, 2nd ed. *Wiley Series in Probability and Statistics*: *Probability and Statistics*. New York: Wiley. A Wiley-Interscience Publication. MR1700749 https://doi.org/10.1002/9780470316962

[7] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1989). *Regular Variation*. *Encyclopedia of Mathematics and Its Applications* **27**. Cambridge: Cambridge Univ. Press. MR1015093

[8] Blanchet, J.H. and Liu, J. (2008). State-dependent importance sampling for regularly varying random walks. *Adv. in Appl. Probab.* **40** 1104–1128. MR2488534

[9] Bollobás, B., Borgs, C., Chayes, J. and Riordan, O. (2003). Directed scale-free graphs. In *Proceedings of the Fourteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (*Baltimore*, *MD*, 2003) 132–139. New York: ACM. MR1974912

[10] Boxma, O.J. and Cohen, J.W. (1999). The single server queue: Heavy tails and heavy traffic. In *Self-Similar Network Traffic and Performance Evaluation* (K. Park and W. Willinger, eds.) Wiley, New York.

[11] Crovella, M., Bestavros, A. and Taqqu, M.S. (1999). Heavy-tailed probability distributions in the world wide web. In *A Practical Guide to Heavy Tails*: *Statistical Techniques for Analysing Heavy Tailed Distributions* (M.S. Taqqu, R. Adler and R. Feldman, eds.). Boston: Birkhäuser.

[12] Daley, D.J. and Vere-Jones, D. (2003). *An Introduction to the Theory of Point Processes*. *Vol. I*: *Elementary Theory and Methods*, 2nd ed. *Probability and Its Applications* (*New York*). New York: Springer. MR1950431

[13] Das, B., Mitra, A. and Resnick, S. (2013). Living on the multidimensional edge: Seeking hidden risks using regular variation. *Adv. in Appl. Probab.* **45** 139–163. MR3077544

[14] de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory*: *An Introduction*. *Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2234156 https://doi.org/10.1007/0-387-34471-3

[15] Dembo, A. and Zeitouni, O. (2010). *Large Deviations Techniques and Applications*. *Stochastic Modelling and Applied Probability* **38**. Berlin: Springer. Corrected reprint of the second (1998) edition. MR2571413 https://doi.org/10.1007/978-3-642-03311-7

[16] Denisov, D., Dieker, A.B. and Shneer, V. (2008). Large deviations for random walks under subexponentiality: The big-jump domain. *Ann. Probab.* **36** 1946–1991. MR2440928 https://doi.org/10.1214/07-AOP382

[17] Durrett, R. (2010). *Random Graph Dynamics*. *Cambridge Series in Statistical and Probabilistic Mathematics* **20**. Cambridge: Cambridge Univ. Press. MR2656427

[18] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997). *Modelling Extremal Events*: *For Insurance and Finance*. *Applications of Mathematics* (*New York*) **33**. Berlin: Springer. MR1458613 https://doi.org/10.1007/978-3-642-33483-2

[19] Ganesh, A., O'Connell, N. and Wischik, D. (2004). *Big Queues*. *Lecture Notes in Math*. **1838**. Berlin: Springer. MR2045489 https://doi.org/10.1007/b95197

[20] Hult, H. and Lindskog, F. (2005). Extremal behavior of regularly varying stochastic processes. *Stochastic Process. Appl.* **115** 249–274. MR2111194 https://doi.org/10.1016/j.spa.2004.09.003

[21] Hult, H. and Lindskog, F. (2006). Regular variation for measures on metric spaces. *Publ. Inst. Math*. (*Beograd*) (*N.S.*) **80(94)** 121–140. MR2281910 https://doi.org/10.2298/PIM0694121H

[22] Hult, H., Lindskog, F., Mikosch, T. and Samorodnitsky, G. (2005). Functional large deviations for multivariate regularly varying random walks. *Ann. Appl. Probab.* **15** 2651–2680. MR2187307 https://doi.org/10.1214/105051605000000502

[23] Ibragimov, R., Jaffee, D. and Walden, J. (2011). Diversification disasters. *J. Financ. Econ.* **99** 333–348.

[24] Janssen, A. and Drees, H. (2016). A stochastic volatility model with flexible extremal dependence structure. *Bernoulli* **22** 1448–1490. MR3474822 https://doi.org/10.3150/15-BEJ699

[25] Jelenković, P. and Momčilović, P. (2003). Asymptotic loss probability in a finite buffer fluid queue with heterogeneous heavy-tailed on-off processes. *Ann. Appl. Probab.* **13** 576–603. MR1970278 https://doi.org/10.1214/aoap/1050689595

[26] Jelenković, P.R. (1999). Subexponential loss rates in a $GI/GI/1$ queue with applications. *Queueing Syst.* **33** 91–123. MR1748640 https://doi.org/10.1023/A:1019167927407

[27] Lindley, D.V. (1952). The theory of queues with a single server. *Proc. Camb. Philos. Soc.* **48** 277–289. MR0046597 https://doi.org/10.1017/s0305004100027638

[28] Lindskog, F., Resnick, S.I. and Roy, J. (2014). Regularly varying measures on metric spaces: Hidden regular variation and hidden jumps. *Probab. Surv.* **11** 270–314. MR3271332 https://doi.org/10.1214/14-PS231

[29] Mansfield, P., Rachev, S.T. and Samorodnitsky, G. (2001). Long strange segments of a stochastic process. *Ann. Appl. Probab.* **11** 878–921. MR1865027 https://doi.org/10.1214/aoap/1015345352

[30] Mikosch, T. and Wintenberger, O. (2013). Precise large deviations for dependent regularly varying sequences. *Probab. Theory Related Fields* **156** 851–887. MR3078288 https://doi.org/10.1007/s00440-012-0445-0

[31] Nagaev, A.V. (1969). Integral limit theorems with regard to large deviations when Cramér's condition is not satisfied. I. *Teor. Veroyatn. Primen.* **14** 51–63. MR0247651

[32] Nagaev, A.V. (1969). Integral limit theorems with regard to large deviations when Cramér's condition is not satisfied. II. *Teor. Veroyatn. Primen.* **14** 203–216. MR0247652

[33] Nagaev, A.V. (1969). Limit theorems that take into account large deviations when Cramér's condition is violated. *Izv. Akad. Nauk UzSSR Ser. Fiz.-Mat. Nauk* **13** 17–22. MR0282396

[34] Resnick, S. (2002). Hidden regular variation, second order regular variation and asymptotic independence. *Extremes* **5** 303–336. MR2002121 https://doi.org/10.1023/A:1025148622954

[35] Resnick, S.I. (2007). *Heavy-Tail Phenomena*: *Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering*. New York: Springer. MR2271424

[36] Rhee, C.-H., Blanchet, J. and Zwart, B. (2017). Sample path large deviations for heavy-tailed Lévy processes and random walks. ArXiv E-prints arXiv:1606.02795v2.

[37] Smith, R.L. (2003). Statistics of extremes, with applications in environment, insurance and finance. In *SemStat*: *Seminaire Europeen de Statistique*, *Exteme Values in Finance*, *Telecommunications*, *and the Environment* (B. Finkenstadt and H. Rootzén, eds.) 1–78. London: Chapman-Hall.

[38] Whitt, W. (2002). *Stochastic-Process Limits*: *An Introduction to Stochastic-Process Limits and Their Application to Queues. Springer Series in Operations Research*. New York: Springer. MR1876437

[39] Zwart, A.P. (2000). A fluid queue with a finite buffer and subexponential input. *Adv. in Appl. Probab.* **32** 221–243. MR1765162 https://doi.org/10.1239/aap/1013540031

[40] Zwart, A.P. (2001). Tail asymptotics for the busy period in the $GI/G/1$ queue. *Math. Oper. Res.* **26** 485–493. MR1849881 https://doi.org/10.1287/moor.26.3.485.10584