Extremal clustering under moderate long range dependence and moderately heavy tails

Zaoli Chen
Gennady Samorodnitsky

Department of Mathematics
Cornell University
e-mail: zc288@cornell.edu

School of Operations Research and Information Engineering
Cornell University
e-mail: gs18@cornell.edu

Abstract: We study clustering of the extremes in a stationary sequence with subexponential tails in the maximum domain of attraction of the Gumbel We obtain functional limit theorems in the space of random sup-measures and in the space $D(0,\infty)$. The limits have the Gumbel distribution if the memory is only moderately long. However, as our results demonstrate rather strikingly, the "heuristic of a single big jump" could fail even in a moderately long range dependence setting. As the tails become lighter, the extremal behavior of a stationary process may depend on multiple large values of the driving noise.

MSC 2010 subject classifications: Primary 60G70, 60F17; secondary 60G57.
Keywords and phrases: Extreme value theory, long range dependence, random sup-measure, stable regenerative set, subexponential tails, extremal clustering, Gumbel domain of attraction.

1. Introduction

This paper is about a very unusual clustering of extreme values that can occur in certain types of stationary stochastic processes with long range dependence. It is useful to recall the basic definitions of the classical extreme value theory. A distribution H on $\mathbb{R}$ is in a maximum domain of attraction if there is a positive sequence (a_n) and a real sequence (b_n) such that the law of $(M_n^{(0)} - b_n)/a_n$ converges weakly as $n \to \infty$ to a nondegenerate distribution G. Here $M_n^{(0)} = \max(Y_1, \ldots, Y_n)$ is the largest value among n i.i.d. random variables $Y_1, Y_2, \ldots$ with the common distribution H. The distribution G is then, automatically, of the form $G(x) = G_\gamma(Ax + B)$, $x \in \mathbb{R}$ for some $A > 0, B \in \mathbb{R}$, and some $\gamma \in \mathbb{R}$. The "standard" distributions G_γ are the Fréchet $G_\gamma(x) = \exp\{-x^{-1/\gamma}\}$, $x \geq 0$

*This research was partially supported by the ARO grant W911NF-18 -10318 at Cornell University.

if $\gamma > 0$, the Gumbel $G_0(x) = \exp\{-e^{-x}\}$, $x \in \mathbb{R}$, and the Weibull $G_\gamma(x) = \exp\{-(-x)^{-1/\gamma}\}$, $x \le 0$ if $\gamma < 0$. See e.g. de Haan and Ferreira (2006).

The extreme values of an i.i.d. sequence, obviously, do not cluster. If, on the other hand, $X_1, X_2, \ldots$ is a stationary sequence with a common marginal distribution H, its extreme values may exhibit a clustering phenomenon. This is a well studied topic in the extreme value theory, where a numerical measure of clustering, *the extremal index*, goes back to Leadbetter (1983). Let $M_n = \max(X_1, \ldots, X_n)$, $n \ge 1$. As above, we denote by $M_n^{(0)}$ the largest of the first n *i.i.d.* observations $Y_1, Y_2, \ldots$ with the same marginal distribution H. Then the stationary sequence $X_1, X_2, \ldots$ has extremal index θ if for some nondegenerate distribution G we have both

$$(M_n^{(0)} - b_n)/a_n \Rightarrow G \ \text{ and } \ (M_n - b_n)/a_n \Rightarrow G^\theta \tag{1.1}$$

as $n \to \infty$; see e.g. de Haan and Ferreira (2006). An extremal index, if it exists, is in the range $0 < \theta \le 1$. Note that the first statement in (1.1) can be rewritten in the form

$$(M_{[n\theta]}^{(0)} - b_n)/a_n \Rightarrow G^\theta \,,$$

which, in conjunction with the second statement in (1.1), says that the largest among the first n observations from the stationary sequence is "similar" to the largest of the first $[n\theta]$ observations from the corresponding i.i.d. sequence. In fact, in most cases a stationary sequence satisfying (1.1) also satisfies

$$(M_n - b_{[n\theta]})/a_{[n\theta]} \Rightarrow G \,, \tag{1.2}$$

which emphasizes the similarity between M_n and $M_{[n\theta]}^{(0)}$ even more. It is also a part of the folklore in extreme value theory that the extremal index can be interpreted as the reciprocal of "the expected extremal cluster size", though we will introduce neither the exact definition of this object nor the conditions under which this interpretation is valid; see e.g. Embrechts et al. (1997). If we denote such expected extremal cluster size by κ, then an alternative expression of (1.2) is

$$(M_n - b_{[n/\kappa]})/a_{[n/\kappa]} \Rightarrow G \,. \tag{1.3}$$

A special situation occurs when the stationary sequence $X_1, X_2, \ldots$ exhibits long range dependence with respect to its extremes (see Samorodnitsky (2016)). In this case (1.1) may hold with $\theta = 0$ which, of course, only says that the centering and scaling of M_n should be changed to obtain a nondegenerate limit. Intuitively, the extremes cluster so much that the extremal clusters become unbounded. and one expects that (1.3) should be replaced by

$$(M_n - b_{m_n})/a_{m_n} \Rightarrow G \,, \tag{1.4}$$

where $m_n = [n/\kappa_n]$, and now κ_n is "the expected extremal cluster size" among the first n observations. This would allow $\kappa_n \to \infty$ and, therefore, $m_n = o(n)$ as $n \to \infty$. Hence, a change in the order of magnitude of scaling and/or centering for the partial maximum M_n.

In fact, it has turned out that (1.4) holds for certain stationary infinitely processes with regularly varying tails. In this case the marginal distributions are in the Fréchet maximum domain of attraction ($\gamma > 0$), which involves no centering ($b_n \equiv 0$); see Samorodnitsky (2004) and Lacaux and Samorodnitsky (2016). In this setup the sequence (m_n) in (1.4) was not obtained via the relation $m_n = [n/\kappa_n]$ but, rather, turned out to be a direct ingredient in the memory in the system. It is important to mention that in these cases an important distinction has appeared between "moderate" long range dependence and "extreme" long range dependence. In the former case the weak limit in (1.4) is, up to shifting and scaling, the standard Fréchet distribution, while in the latter case the limit is not one of the classical extreme value distributions; see Samorodnitsky and Wang (2019). In the former case an extreme of the process due to a single large value of the underlying noise, consistent with the "heuristic of a single big jump" for extreme events and large deviations of heavy tailed systems (e.g. Rhee et al. (2019)). On the other hand, in the latter case this heuristic fails.

Our goal in this paper is to understand how the extremes of a long memory stationary process cluster when the marginal tails are still heavy (so that the "heuristic of a single big jump" is still often the first guidance one has), but less heavy than the regularly varying tails considered earlier. The natural class of such marginal distributions is the class of *subexponential distributions*. Recall that a distribution H is subexponential if

$$\lim_{x\to\infty} \frac{\overline{H*H}(x)}{\overline{H}(x)} = 2\,, \tag{1.5}$$

where $\overline{H}(x) = 1 - H(x)$ (Chistyakov (1964)). Distributions with a regularly varying right tail are, of course, subexponential, but we are interested in the subexponential distributions whose tails are lighter than any regularly varying tail. Specifically, we are interested in the subexponential distributions in the maximum domain of attraction of the Gumbel distribution G_0. We will give the exact assumptions on the marginal tails in the sequel.

A surprising conclusion of our results that, while (1.4) does continue to hold in the class of long memory stationary processes with certain subexponential distributions in the maximum domain of attraction of the Gumbel distribution G_0 as the marginal distributions, (1.4) breaks down for such distributions once their tails become light enough. This may happen even when the memory is only "moderate long memory" (to be defined precisely in the sequel.) The moderate long memory case is the only one we consider in this paper. It turns out that when the tails become light enough, the centering in the left hand side of , (1.4) acquires another term, of a smaller order than b_{m_n} (but of a larger order than a_{m_n}.) This term arises because the "single big jump" heuristic breaks down once again. That is, that heuristic may break down not only when the memory is too long, but also when the tails are too light (while remaining subexponential, hence heavy!)

The paper is organized as follows. In Section 2 we review some essential facts on subexponential distributions in the Gumbel maximum domain of attraction, random closed sets, null-recurrent Markov chains underlying the infinitely divisible dynamics in our model and random sup-measures, and introduce the limiting random sup-measure later appearing in the main result. In Section 3, we introduce the stationary infinitely divisible processes we are considering and list the assumptions we are imposing. The main results, the extremal limit theorems in the space of random sup-measures and in the space of càdlàg functions are stated and proved in Section 4. This sections also contains two natural examples. The appendices A and B contain several arguments and verifications needed in the earlier parts of the paper.

We will use the following standard notation throughout the paper. Let $\{a_n\}_{n\in\mathbb{N}}$ and $\{b_n\}_{n\in\mathbb{N}}$ be two positive sequences. We describe the asymptotic relation between them by writing:

(a) $a_n \sim b_n$ if $\lim_{n\to\infty} a_n/b_n = 1$,
(b) $a_n \gg b_n$ if $\lim_{n\to\infty} a_n/b_n = \infty$, or equivalently $b_n \ll a_n$,
(c) $a_n \lesssim b_n$ if there exists $C > 0$ such that $a_n \leqslant Cb_n$ for large enough n, and analogously with $a_n \gtrsim b_n$,
(d) $a_n \asymp b_n$ if both $a_n \lesssim b_n$ and $a_n \gtrsim b_n$.

If $\{A_n\}_{n\in\mathbb{N}}$ and $\{B_n\}_{n\in\mathbb{N}}$ are two sequences of positive random variables, we will write

(a) $A_n = o_P(B_n)$ if $A_n/B_n \to 0$ in probability,
(b) $A_n \lesssim_P B_n$ if (A_n/B_n) is tight, and analogously with $A_n \gtrsim_P B_n$.

2. Preliminaries

This section is of the background nature. It collects a number of mostly well-known notions and results needed in this paper.

2.1. Subexponential distributions in the Gumbel maximum domain of attraction

Most of the material quoted in this section is in Resnick (1987); see also Goldie and Resnick (1988). Subexonentiality requires the distribution to have a support which is unbounded on the right, so we only consider such distributions. A distribution H is in the maximum domain of attraction of the Gumbel distribution if and only if $G = (1/(1-H))^{\leftarrow}$ is Π-varying, and for a non-decreasing function J the generalized inverse of J is defined as

$$J^{\leftarrow}(y) = \inf\{s : J(s) \geqslant y\}\,. \tag{2.1}$$

Furthermore, a non-negative, non-decreasing function V is said to be Π-varying if there exist functions $a(t) > 0, b(t) \in \mathbb{R}$ such that for $x > 0$

$$\lim_{t\to\infty} \frac{V(tx) - b(t)}{a(t)} = \log x. \tag{2.2}$$

Alternatively, H is in the maximum domain of attraction of the Gumbel distribution if and only if there exist $x_0 \in \mathbb{R}$ and $c(x) \to c > 0$ as $x \to \infty$ such that for $x_0 < x < \infty$

$$\overline{H}(x) = c(x) \exp\left\{ -\int_{x_0}^{x} \frac{1}{h(u)} du \right\} \tag{2.3}$$

where h (the so-called auxiliary function) is an absolutely continuous positive function on (x_0, ∞) with density h' satisfying $\lim_{u\to\infty} h'(u) = 0$. The function h must satisfy $h(x) = o(x)$ as $x \to \infty$, and subexponentiality of H requires also $\lim_{u\to\infty} h(u) = \infty$. For a distribution H satisfying (2.3), the centering and scaling required for the convergence $(M_n^{(0)} - b_n)/a_n \to G_0$ can be chosen as

$$b_n = \left(\frac{1}{1-H}\right)^{\leftarrow}(n), \;\; a_n = h(b_n)\,.$$

We will often use the following fact: if one replaces the function $c(\cdot)$ in (2.3) by an asymptotically equivalent function, and denotes the new normalizing sequences by $(\tilde{a}_n)$ and $(\tilde{b}_n)$, then

$$\lim_{n\to\infty} \frac{b_n - \tilde{b}_n}{a_n} = 0, \;\; \lim_{n\to\infty} \frac{\tilde{a}_n}{a_n} = 1\,. \tag{2.4}$$

2.2. Random closed sets

We use the notation $\mathcal{G}, \mathcal{F}$ and $\mathcal{K}$ for the families of open, closed and compact sets of $[0, 1]$ or $[0, \infty)$ (depending on the context), respectively. Details on most of the material in this section can be found in Molchanov (2017).

The Fell topology $\mathcal{B}(\mathcal{F})$ on $\mathcal{F}$ is the topology generated by the subbasis

$$\begin{aligned} \mathcal{F}_G &= \{F \in \mathcal{F} : F \cap G \neq \emptyset\}\,, \quad G \in \mathcal{G}, \\ \mathcal{F}^K &= \{F \in \mathcal{F} : F \cap K = \emptyset\}\,, \quad K \in \mathcal{K}. \end{aligned}$$

The Fell topology is metrizable and compact. A random closed set is a measurable mapping from a probability space to $\mathcal{B}(\mathcal{F})$.

For $\beta \in (0, 1)$, let $\{L_\beta(t)\}_{t\geqslant 0}$ be the standard β-stable subordinator, i.e. the increasing Lévy process with the Laplace transform $\mathbb{E}e^{-\theta L_\beta(t)} = e^{-t\theta^\beta}, \theta \geqslant 0$. We define the β-stable regenerative set Z to be the closure of the range of a β-subordinator, i.e.,

$$Z = \overline{\{L_\beta(t) : t \geqslant 0\}}. \tag{2.5}$$

It is a random closed subset of $[0, \infty)$. Much of the discussion in this paper revolves around a sequence of i.i.d. random closed subsets of $[0, 1]$ defined as follows.

Let $\{V_j\}_{j\geqslant 1}$ be a family of i.i.d. random variables on $[0, 1]$ with

$$\mathbb{P}(V_1 \leqslant x) = x^{1-\beta}, \quad x \in [0, 1]\,, \tag{2.6}$$

independent of an i.i.d. sequence $\{Z_j\}_{j\geqslant 1}$ of β-stable regenerative sets. We define

$$\overline{R_j} = (V_j + Z_j) \cap [0,1]. \tag{2.7}$$

That is, each $\overline{R_j}$ is the restriction of a shifted stable regenerative set to $[0,1]$. It is, clearly, non-empty.

2.3. Null-recurrent Markov chains

We describe now some ergodic theoretic notions associated with certain null recurrent Markov chains. Our main reference here is Aaronson (1997). Let $\{Y_t\}_{t\in\mathbb{Z}}$ be an irreducible, aperiodic, and null recurrent Markov chain on $\mathbb{Z}$, and denote by $(\pi_i)_{i\in\mathbb{Z}}$ its unique invariant measure satisfying $\pi_0 = 1$. If $(E,\mathcal{E})$ is the "path space" $(\mathbb{Z}^{\mathbb{Z}}, \mathcal{B}(\mathbb{Z}^{\mathbb{Z}}))$, we can define an infinite σ-finite measure on $(E,\mathcal{E})$ by

$$\mu(\cdot) := \sum_{i\in\mathbb{Z}} \pi_i P_i(\cdot)\,, \tag{2.8}$$

where P_i is the probability law of $\{Y_t\}_{t\in\mathbb{Z}}$ on $(E,\mathcal{E})$ given $Y_0 = i$. The left shift operator on E by θ defined by

$$\theta : (\ldots, y_0, y_1, y_2, \ldots) \mapsto (\ldots, y_1, y_2, y_3, \ldots) \tag{2.9}$$

is a measure preserving, conservative and ergodic operator on $(E,\mathcal{E},\mu)$. See Harris and Robbins (1953). For $n \in \mathbb{Z}$ let

$$A_n := \{y \in E : y_n = 0\}. \tag{2.10}$$

The *wandering rate sequence* $\{w_n\}_{n\in\mathbb{N}}$ is then defined as

$$w_n := \mu\left(\bigcup_{k=0}^{n} A_k\right), \quad n \in \mathbb{N}. \tag{2.11}$$

Define the first visit time to state 0 by

$$\varphi(y) := \inf\{t \geqslant 1 : y_t = 0\}, y \in E. \tag{2.12}$$

The Markov chains we consider satisfy the following assumption.

Assumption 2.1. *There exists* $\beta \in (0,1)$ *and a slowly varying function* L *such that*

$$\overline{F}(n) := P_0[\varphi > n] = n^{-\beta} L(n) \in RV_{-\beta}\,. \tag{2.13}$$

Furthermore,

$$\sup_{n\geqslant 0} \frac{n\mathbb{P}_0(\varphi = n)}{\overline{F}(n)} < \infty\,. \tag{2.14}$$

Remark 2.1. Under the Assumption 2.1, as $n \to \infty$,

$$\begin{aligned} w_n &\sim \sum_{k=1}^{n} \mu(\varphi = k) \\ &= \sum_{k=1}^{n} \overline{F}(k-1) \sim \frac{n^{1-\beta}L(n)}{1-\beta} \in \mathrm{RV}_{1-\beta}\,. \end{aligned} \tag{2.15}$$

See Resnick et al. (2000) Lemma 3.3.

The times a sequence $y \in E$ visits state 0 under certain conditional versions of the measure μ are of crucial importance for us. Specifically, for each $n \in \mathbb{N}$, let

$$\mu_n(B) := \frac{\mu\left(B \cap \bigcup_{k=0}^{n} A_k\right)}{w_n},\ B \in \mathcal{E}. \tag{2.16}$$

Let $\{Y^{(j,n)}\}_{j\in\mathbb{N}}$ be a family of i.i.d. random elements in E with law μ_n. For each j we set

$$I_{j,n} := \left\{0 \leqslant t \leqslant n : Y_t^{(j,n)} = 0\right\}\,. \tag{2.17}$$

We further define

$$\widehat{I}_{1,n} := I_{1,n}, \tag{2.18}$$

$$\widehat{I}_{j,n} := I_{j,n} \cap \bigcap_{i=1}^{j-1} I_{i,n}^{c}, \quad j \geqslant 2. \tag{2.19}$$

The facts mentioned below are in Samorodnitsky and Wang (2019). First, by Theorem 5.4 *ibid.*,

$$\frac{1}{n} I_{j,n} \Rightarrow \overline{R_j},\ j = 1, 2, \dots \tag{2.20}$$

weakly in the space of random closed subsets of $[0,1]$, where $\overline{R_j}$ is defined in (2.7). In particular,

$$\lim_{n\to\infty} \mathbb{P}\left(I_{j,n} \cap nG \neq \emptyset\right) = \mathbb{P}\left(\overline{R_j} \cap G \neq \emptyset\right) > 0 \text{ for any } G \in \mathcal{G}([0,1]). \tag{2.21}$$

If, in addition, $0 < \beta < 1/2$, then for any two distinct $j_1, j_2 \in \mathbb{N}$

$$\lim_{n\to\infty} \mathbb{P}\left(I_{j_1,n} \cap I_{j_2,n} \neq \emptyset\right) = 0. \tag{2.22}$$

Therefore, for any $m \in \mathbb{N}$,

$$\lim_{n\to\infty} \mathbb{P}\left(I_{j,n} = \widehat{I}_{j,n},\, j = 1, \dots, m.\right) = 1. \tag{2.23}$$

In the sequel we will need estimates of how quickly the intersection probability in (2.22) and certain related probabilities converge to zero. For an open interval $T \subset [0,1]$ we define

$$p_{n,T} := \mathbb{P}\left(I_{1,n} \cap I_{2,n} \cap nT \neq \emptyset\right), \tag{2.24}$$

$$\overline{p}_{n,T} := \mathbb{P}\left(I_{1,n} \cap I_{2,n} \cap nT \neq \emptyset \middle| Y^{(1,n)}\right)\,, \tag{2.25}$$

with the latter probability being random. Clearly, $p_{n,T} = \mathbb{E}\overline{p}_{n,T}$. The following theorem may be of independent interest. It is proved in Appendix A.

Theorem 2.1. *Under Assumption 2.1 with* $0 < \beta < 1/2$, *for any open interval* T,

(i)

$$p_{n,T} \asymp \frac{n^\beta}{w_n L(n)} \,. \tag{2.26}$$

(ii) For any $C > 0$, *there exists* $c > 0$ *such that for every* $n \geq 1$

$$\mathbb{P}\left(\overline{p}_{n,T} \geqslant \frac{cn^\beta \log n}{w_n L(n)}\right) \leqslant n^{-C} \,. \tag{2.27}$$

(iii) For any $\gamma > (1-2\beta)^{-1}$ *and* $\epsilon > 0$, *there exists* $c > 0$ *such that*

$$\liminf_{n\to\infty} \mathbb{P}\left(\overline{p}_{n,T} \geqslant \frac{cn^\beta}{w_n L(n)} \cdot \frac{L\left((\log n)^\gamma\right)}{(\log n)^{\gamma\beta}} \,\Big|\, I_{1,n} \cap nT \neq \emptyset\right) \geqslant 1-\epsilon. \tag{2.28}$$

It follows immediately from part (iii) of the theorem that

$$\overline{p}_{n,[0,1]} \gtrsim_P \frac{n^\beta}{w_n L(n)} \cdot \frac{L\left((\log n)^\gamma\right)}{(\log n)^{\gamma\beta}} \,. \tag{2.29}$$

for any $\gamma > (1-2\beta)^{-1}$.

2.4. Random Sup-Measures

We continue to use the notation of Subsection 2.2. Our main reference is O'Brien et al. (1990); note that our sup-measures take values in $\overline{\mathbb{R}} = [-\infty, \infty]$.

A sup-measure is a mapping $m : \mathcal{G} \to \overline{\mathbb{R}}$ such that $m(\emptyset) = -\infty$ and $m(\cup_\alpha G_\alpha) = \vee_\alpha m(G_\alpha)$ for an arbitrary collection (G_α) of open sets. The sup-derivative $d^\vee m$ of m is defined by

$$d^\vee m(t) = \bigwedge_{t\in G} m(G),$$

it is an upper semicontinuous $\overline{\mathbb{R}}$-valued function of t. Given an $\overline{\mathbb{R}}$-valued function f, the sup-integral of f defined by

$$i^\vee f(G) = \bigvee_{t\in G} f(t), \quad G \in \mathcal{G};$$

it is automatically a sup-measure. It is always true that $m = i^\vee d^\vee m$, and we can extend the domain of a sup-measure to all Borel sets via

$$m(B) = \bigvee_{t\in B} d^\vee m(t), \quad B \;\text{ Borel.} \tag{2.30}$$

On the collection SM of all sup-measures one defines the sup-vague topology, in which a sequence of sup-measures $\{m_n\}_{n\geqslant 1}$ converges to a sup-measure m if and only if

$$\limsup_{n\to\infty} m_n(K) \leqslant m(K), \quad \text{for each } K \in \mathcal{K},$$
$$\liminf_{n\to\infty} m_n(G) \geqslant m(G), \quad \text{for each } G \in \mathcal{G}.$$

The space SM equipped with sup-vague topology is compact and metrizable.

A random sup-measure M is a measurable mapping from a probability space to SM. For a random sup-measure M, let $\mathscr{I}(M)$ be the collection of continuity intervals of M, defined by

$$\mathscr{I}(M) = \{I \text{ an open interval}: \ M(I) = M(\text{clos } I) \text{ a.s.}\}.$$

If $\{M_n\}_{n\geqslant 1}$ and M are random sup-measures, then $M_n \Rightarrow M$ if and only if

$$(M_n(I_1), \dots, M_n(I_m)) \Rightarrow (M(I_1), \dots, M(I_m)) \tag{2.31}$$

for arbitrary disjoint intervals $I_1, \dots, I_m \in \mathscr{I}(M)$.

We now define a family of random sup-measures that will arise naturally in the sequel. Let $\beta \in (0, 1/2)$ and consider a Poisson point process on $\mathbb{R} \times \mathbb{R}_+ \times \mathcal{F}(\mathbb{R}_+)$ with mean measure

$$e^{-u} du (1-\beta) v^{-\beta} dv \, dP_\beta \, ,$$

where P_β is the law of the β-stable regenerative set in (2.5). Let $(U_j, V_j^*, Z_j)_{j\in\mathbb{N}}$ be a measurable enumeration of points of this Poisson point process, and denote

$$R_j = V_j^* + Z_j, \quad j \in \mathbb{N}. \tag{2.32}$$

Since $\beta \in (0, 1/2)$, we have

$$\mathbb{P}(R_1 \cap R_2 = \emptyset) = 1\, ; \tag{2.33}$$

see Lemma 3.1 in Samorodnitsky and Wang (2019). It follows immediately that, on an event of probability 1, the function $\eta : \mathbb{R}_+ \to \overline{\mathbb{R}}$ defined by

$$\eta(t) = \bigvee_{j=1}^{\infty} U_j 1_{\{t\in R_j\}}$$

is an upper semicontinuous function. Hence, it is the sup-derivative of the random sup-measure

$$\mathcal{M}(B) = \bigvee_{j=1}^{\infty} U_j 1_{\{B\cap R_j \neq \emptyset\}} \, . \tag{2.34}$$

This measure is stationary, i.e.

$$\mathcal{M}(r + \cdot) \overset{d}{=} \mathcal{M}(\cdot) \tag{2.35}$$

for any $r \geqslant 0$; see Proposition 4.3 in Lacaux and Samorodnitsky (2016). Moreover, we claim that M is self-affine, i.e.

$$\mathcal{M}(a\cdot) \stackrel{d}{=} \mathcal{M}(\cdot) + (1-\beta)\log a \tag{2.36}$$

for any $a > 0$. Indeed, note that

$$\begin{aligned}
\mathcal{M}(aB) &= \bigvee_{j=1}^{\infty} U_j 1_{\{B\cap(a^{-1}V_j^*+a^{-1}Z_j)\neq\emptyset\}} \\
&\stackrel{d}{=} \bigvee_{j=1}^{\infty} U_j 1_{\{B\cap(a^{-1}V_j^*+Z_j)\neq\emptyset\}} \\
&\stackrel{d}{=} \bigvee_{j=1}^{\infty} (U_j + (1-\beta)\log a) 1_{\{B\cap(V_j^*+Z_j)\neq\emptyset\}} = \mathcal{M}(B) + (1-\beta)\log a\,;
\end{aligned}$$

see e.g. Proposition 4.1(b) in Samorodnitsky (2016) for the first distributional equality, while the second one holds because both the points $(U_j, a^{-1}V_j^*, Z_j)_{j\in\mathbb{N}}$ and the points $(U_j + (1-\beta)\log a, V_j^*, Z_j)_{j\in\mathbb{N}}$ form a Poisson point process with mean measure

$$a^{1-\beta}e^{-u}du(1-\beta)v^{-\beta}dv\,dP_\beta\,.$$

Suppose that $\{X_t\}_{t\in\mathbb{Z}}$ is a stationary process. It induces naturally a sequence of random sup-measures by setting for $n \in \mathbb{N}$

$$M_n(B) = \max_{t\in nB} X_t, \quad B \in \mathcal{B}(\mathbb{R}_+)\,. \tag{2.37}$$

If $\mathcal{M}$ is the random sup measure in (2.34), then the weak convergence

$$\frac{M_n(\cdot) - b_n}{a_n} \Rightarrow \mathcal{M}(\cdot)$$

in the space of random sup-measures on $[0,1]$ for some (a_n, b_n) guarantees also this weak convergence in the space of random sup-measures on $\mathbb{R}_+$. Furthermore, every open interval is a continuity interval for $\mathcal{M}$, since stable regenerative sets do not hit fixed points. By (2.31)

$$\left(\frac{M_n(I_1) - b_n}{a_n}, \ldots, \frac{M_n(I_m) - b_n}{a_n}\right) \Rightarrow (\mathcal{M}(I_1), \ldots, \mathcal{M}(I_m)) \tag{2.38}$$

for arbitrary disjoint open intervals $I_1, \ldots, I_m$ in $[0,1]$ is necessary and sufficient for weak convergence to $\mathcal{M}$.

The restriction of the sup measure $\mathcal{M}$ to subsets of $[0,1]$ has representation somewhat more transparent than (2.34). Let $\{V_j\}_{j\geqslant 1}$ be a family of i.i.d. random variables on $[0,1]$ with the law (2.6). Let $\{Z_j\}_{j\geqslant 1}$ be a family of i.i.d. β-stable regenerative sets in (2.5). Finally, let $\{\Gamma_j\}_{j\geqslant 1}$ be the sequence of arrival times of a unit rate Poisson processes on $(0,\infty)$. We assume that all three sequences are independent. Then the points $(-\log\Gamma_j, V_j, Z_j)_{j\in\mathbb{N}}$ form a Poisson point process on $\mathbb{R}\times[0,1]\times\mathcal{F}(\mathbb{R}_+)$ whose mean measure is the mean measure of restriction of the Poisson point process $(U_j, V_j^*, Z_j)_{j\in\mathbb{N}}$ to $\mathbb{R}\times[0,1]\times\mathcal{F}(\mathbb{R}_+)$. Therefore, if we define i.i.d. random nonempty compact sets by (2.7), then the following representation in law holds:

$$\mathcal{M}(B) = \bigvee_{t\in B} -\log\Gamma_j 1_{\{B\cap\overline{R_j}\neq\emptyset\}},\ B\in\mathcal{B}([0,1])\,. \tag{2.39}$$

3. A family of stationary subexponential infinitely divisible processes

We now define a family of stationary infinitely divisible processes for whom we will establish extremal limit theorems. Our processes will be of the form

$$X_n = \int_E f\circ\theta^n(x)\, M(dx),\ n\in\mathbb{Z}\,, \tag{3.1}$$

where θ is the left shift operator on $E=\mathbb{Z}^{\mathbb{Z}}$ in (2.9) and M is an infinitely divisible random measure on $(E,\mathcal{E})$ with a constant local characteristic triple (σ^2,ν,b) and control measure μ in (2.8), associated with an invariant measure of an irreducible, aperiodic, and null recurrent Markov chain on $\mathbb{Z}$; see Chapter 3 in Samorodnitsky (2016) for details in infinitely divisible random measures and integrals with respect to such measures. The function f must satisfy certain integrability conditions; if it does, the process $\{X_n\}_{n\in\mathbb{Z}}$ is automatically stationary, because the left shift θ preserves the control measure μ. In the sequel we will assume, for simplicity, that f is the indicator function

$$f(x) = \mathbf{1}(x_0=0)\ \text{ for }\ x=(\ldots,x_0,x_1,x_2,\ldots)\,, \tag{3.2}$$

but the results of this paper will undoubtedly hold for a more general class of functions f. The indicator function f in (3.2) automatically satisfies the integrability conditions and, in this case, each X_n is an infinitely divisible random variable with a characteristic triple (σ^2,ν,b); see Section 7 in Sato (2013). The key assumption we will impose in the sequel is that the distribution $(\nu(1,\infty))^{-1}[\nu]_{(1,\infty)}$ is subexponential, from which it immediately follows that

$$\mathbb{P}(X_n>x)\sim\nu(x,\infty)=:\bar\nu(x)\ \text{ as } x\to\infty \tag{3.3}$$

and, in particular, X_n has a subexponential distribution; see Embrechts et al. (1979). We will, in fact, impose a number of additional assumptions on the Lévy measure ν. These assumptions will guarantee that the tail of X_n is light enough to be in the maximum domain of attraction of the Gumbel distribution. On the other hand, they will also guarantee that this tail is not "too light".

Assumption 3.1. *The distribution* $(\nu(1,\infty))^{-1}[\nu]_{(1,\infty)}$ *is both subexponential and in the maximum domain of attraction of the Gumbel distribution. Furthermore, there is a distribution* $H_\#$ *satisfying* $\bar\nu(x) \sim a\overline{H_\#}$ *for* $a > 0$, *and which satisfies* (2.3) *with* $c \equiv 1$, *i.e.*

$$\overline{H_\#}(x) = \exp\left(-\int_{x_0}^{x} \frac{1}{h(u)} du\right) \text{ for } x > x_0\,, \tag{3.4}$$

and the auxiliary function h *with* $h' > 0$ *on* (x_0, ∞), *and such that*

$$\lim_{b\downarrow 1} \limsup_{x\to\infty} \frac{h(bx)}{h(x)} = 1\,. \tag{3.5}$$

Denoting

$$G(x) := \left(\frac{1}{1-H_\#}\right)^{\leftarrow}(x),\ x \ge 0\,, \tag{3.6}$$

we assume that the function G *is of the form*

$$G(x) = \exp\left\{\int_e^x \frac{\zeta(u)}{u\log u} du\right\},\ x > x_1, \quad \text{for some } x_1 > e, \tag{3.7}$$

where ζ *satisfies the following assumptions.*

(B1) ζ *is roughly increasing, i.e.,*

$$\zeta(x) \asymp \sup_{[1,x]} \zeta(u).$$

(B2) There exists some $\delta > 0$ *such that*

$$(\log\log u)^\delta \ll \zeta(u) \lesssim \frac{\log u}{\log\log u}.$$

(B3) For the $\delta > 0$ *in (B2) and for all small* $\rho > 0$,

$$\zeta\left(x^{1-\rho\big/(\log\log x)^{\delta\wedge 1}}\right) \gtrsim \zeta(x)\,.$$

(B4) For any $c > 0$,

$$\liminf_{x\to\infty} \int_{x^{1-c/\zeta(x)}}^{x} \frac{\zeta(u)}{u\log u} du > 0.$$

We check in Appendix B below that the following two important classes of Lévy measures with subexponential tails satisfy Assumption 3.1.

Example 3.1 (lognormal-type tails)**.**

$$\overline{\nu}(x) \sim cx^\beta (\log x)^\xi \exp\left(-\lambda(\log x)^\gamma\right) \quad \text{as } x\to\infty$$

for some $\gamma > 1$, $\lambda, c > 0$ *and* $\beta, \xi \in \mathbb{R}$.

Example 3.2 (super-lognormal-type tails)**.**

$$\overline{\nu}(x) \sim cx^\beta (\log x)^\xi \exp\left(\lambda(\log x)^\gamma\right) \exp\left(-\rho \exp\left(\mu(\log x)^\alpha\right)\right) \quad \text{as } x \to \infty$$

for some $\alpha \in (0,1)$*,* $c, \mu, \rho > 0$ *and* $\beta, \xi, \lambda, \gamma \in \mathbb{R}$*.*

Remark 3.1. *The semi-exponential-type tails such as* $\overline{\nu}(x) \sim \exp(-x^\alpha)$*,* $0 < \alpha < 1$*, unfortunately, do not satisfy the assumptions and, hence, are excluded from our analysis.*

The following proposition, proved in Appendix B, lists certain properties of Lévy measures satisfying Assumption 3.1. We will find these properties useful in the sequel. Let $\delta > 0$ as in Assumption 3.1 ($B2$).

Proposition 3.1. *(i)* $G(x) \gg \exp\left\{(\log\log u)^{1+\delta}/(1+\delta)\right\}$.
(ii) For any $\alpha_1 > \alpha_2 > 0$*, for any* $b < \log \alpha_1 - \log \alpha_2$*,*

$$\frac{G(x^{\alpha_1})}{G(x^{\alpha_2})} \gg \exp\left\{b(\log\log x)^\delta\right\} . \tag{3.8}$$

(iii) For any $H_i \in RV_{\alpha_i}$*,* $i = 1, 2$*,* $\alpha_1 > \alpha_2 > 0$*, for any* $b < \log \alpha_1 - \log \alpha_2$*,*

$$\frac{h \circ G(H_1(x))}{h \circ G(H_2(x))} \gg \exp\left\{b(\log\log x)^\delta\right\} . \tag{3.9}$$

(iv) For any $\alpha \neq 0$*,*

$$|G(x(\log x)^\alpha) - G(x)| \asymp (\log\log x) h \circ G(x) . \tag{3.10}$$

(v) For all sufficiently small $\rho > 0$*,*

$$\min_{1 \leqslant j \leqslant \rho \log x/\zeta(x)} \frac{G(x) - G\left(x2^{-j}\right)}{j} \gtrsim jh \circ G(x) . \tag{3.11}$$

4. Extremal limit theorems

Let $(X_t)_{t\in\mathbb{Z}}$ be a stationary infinitely divisible process (3.1), associated with an irreducible, aperiodic, and null recurrent Markov chain on $\mathbb{Z}$. Recall that we assume that the function f is the indicator function (3.2). Our main result in this section is a limit theorem for the sequence of random sup-measures defined by the process via (2.37). The Lévy measure ν of the infinitely divisible random measure M in (3.1) is assumed to satisfy Assumption 3.1. We denote

$$V(x) = (1/\overline{\nu})^{\leftarrow}(x), \; x > 0 . \tag{4.1}$$

The Markov chain underlying the process $(X_t)_{t\in\mathbb{Z}}$ is assumed to satisfy Assumption 2.1. We define for $n = 1, 2, \ldots$

$$b_n = V(w_n) + V\big(1/\overline{F}(n)\big), \quad a_n = h \circ V(w_n) , \tag{4.2}$$

with w_n is the wandering rate in (2.11), F is the first return time law in (2.13), and h the auxiliary function in (2.3).

Theorem 4.1. *Assume that Assumption 3.1 holds, and that Assumption 2.1 is satisfied with* $0 < \beta < 1/2$. *If* $(a_n), (b_n)$ *are given by* (4.2), *then*

$$\frac{M_n(\cdot) - b_n}{a_n} \Rightarrow \mathcal{M}(\cdot) \tag{4.3}$$

weakly in the space of sup-measures on $[0, 1]$, *where* (M_n) *are the random sup-measures in* (2.37), *and the limiting random sup-measure* $\mathcal{M}$ *is given by* (2.34).

There is a natural counterpart of Theorem 4.1 that establishes an extremal limit theorem in a function space. Recall that a standard Gumbel extremal process is a nondecreasing process $\big(X_{\mathrm{G}}(t),\, t > 0\big)$ satisfying

$$\mathbb{P}\big(X_{\mathrm{G}}(t_i) \le x_i,\, i = 1, \ldots, k\big) = \exp\left\{ -\sum_{i=1}^{k} (t_i - t_{i-1}) e^{-x_i} \right\}$$

for $0 < t_1 < \ldots < t_k$ and $x_1 \le \ldots \le x_k$. The process is continuous in probability and has a version in $D(0, \infty) = \cap_{\varepsilon > 0} D[\varepsilon, \infty)$; see Resnick and Rubinovitch (1973). It is immediate from the definition of the random sup-measure $\mathcal{M}$ in (2.34) that

$$\big(\mathcal{M}([0, t]),\, t > 0\big) \stackrel{d}{=} \big(X_{\mathrm{G}}(t^{1-\beta})\,; t > 0\big)\,, \tag{4.4}$$

see also Lacaux and Samorodnitsky (2016). Note that the finite-dimensional convergence part in the following theorem already follows from Theorem 4.1.

Theorem 4.2. *Under the assumptions of Theorem 4.1,*

$$\left(\frac{\max_{s \le nt} X_s - b_n}{a_n},\, t > 0 \right) \Rightarrow \big(X_{\mathrm{G}}(t^{1-\beta}),\, t > 0\big) \tag{4.5}$$

weakly in the Skorohod J_1 *topology on* $D(0, \infty)$.

Remark 4.1. Let us return to the discussion in Introduction of this paper and compare the statement of Theorems 4.1 and 4.2 to the "expected behavior" of the extreme values presented in (1.4). The results of Lacaux and Samorodnitsky (2016) and Samorodnitsky and Wang (2019) in the case of regularly varying tails suggest that $m_n = w_n$, and the centering and the normalization in (1.4) do not appear to be consistent with Theorema 4.1 and 4.2 due to the presence of an extra term $V\big(1/\overline{F}(n)\big)$ in the centering sequence. It turns out, however, that as long as the tails of the process $(X_t)_{t \in \mathbb{Z}}$ are "not too light" we have

$$\lim_{n \to \infty} \frac{V\big(1/\overline{F}(n)\big)}{a_n} = 0\,, \tag{4.6}$$

and so (1.4) does predict the correct centering and the normalization. Once the tails of the process become lighter, however, (4.6) may fail, and a different centering becomes necessary. We can see this phenomenon on Examples 3.1 and 3.2. In fact, for the lognormal-type tails of Example 3.1 the relation (4.6) holds, while for the super-lognormal-type tails of Example 3.2, 3.1 holds if $0 < \alpha < 1/2$ and fails if $1/2 < \alpha < 1$. These claims are verified in Appendix B.

We will prove the two theorems in the remainder of this section, beginning with Theorem 4.1. We start with a preliminary analysis that will split the proof into several steps. First of all, by (2.38), we need to prove that for arbitrarily disjoint open intervals $I_1, \ldots, I_m$ in $[0,1]$,

$$\left(\frac{M_n(I_i) - b_n}{a_n}\right)_{i=1,\ldots,m} \Rightarrow \left(\mathcal{M}(I_i)\right)_{i=1,\ldots,m} \tag{4.7}$$

weakly in $\mathbb{R}^m$. Note, further, that for any $0 < \varepsilon < a$ the function V in (4.1) satisfies

$$G\big(x(a-\varepsilon)\big) \le V(x) \le G\big(x(a+\varepsilon)\big) \tag{4.8}$$

for all x large enough. Next, we decompose the stationary process $(X_t)_{t\in\mathbb{Z}}$ as follows. Let $M^{(1)}$ and $M^{(2)}$ be two independent infinitely divisible random measures on $(E,\mathcal{E})$, both with with the same control measure μ as the measure M in (3.1). With (σ^2, ν, b) being the local characteristic triple of M, we set the local characteristic triple of $M^{(1)}$ to be $(0, [\nu]_{(x_0,\infty)}, 0)$, and the local characteristic triple of $M^{(2)}$ to be $(\sigma^2, [\nu]_{(-\infty,x_0]}, b)$, with x_0 as in (3.4). If we define for each $t \in \mathbb{Z}$

$$X_t^{(1)} = \int_E f\circ\theta^t(x)\, M^{(1)}(dx), \quad X_t^{(2)} = \int_E f\circ\theta^t(x)\, M^{(2)}(dx)\,, \tag{4.9}$$

then $\{X_t^{(i)}\}_{t\in\mathbb{Z}}, i = 1,2$ are two independent stationary infinitely divisible processes such that $\{X_t\}_{t\in\mathbb{Z}} \overset{d}{=} \{X_t^{(1)} + X_t^{(2)}\}_{t\in\mathbb{Z}}$. For $i = 1,2$ we let $M_n^{(i)}(\cdot)$ be the random sup-measure defined for $\{X_t^{(i)}\}_{t\in\mathbb{Z}}$ as in (2.37). The following proposition shows that $M_n^{(2)}$ is asymptotically negligible with our scaling.

Proposition 4.1. *$M_n^{(2)}([0,1])/a_n \to 0$ as $n \to \infty$.*

Proof. Since the Lévy measure of $X_0^{(2)}$ is bounded on the right, $\mathbb{P}(X_0^{(2)} > r) = o(e^{-cr})$ for any $c > 0$ see e.g. Theorem 26.1 in Sato (2013). Using the fact that $\zeta(x) \to \infty$ we use (4.8) and part (i) of Proposition 3.1 to see that for any $p > 0$, for all large n,

$$\begin{aligned} a_n \ge h\circ G(aw_n/2) &= \frac{G(aw_n/2)\zeta(aw_n/2)}{\log(aw_n/2)} \\ &\gg \frac{G(aw_n/2)}{\log(aw_n/2)} \gg (\log w_n)^p\,, \end{aligned}$$

therefore, taking $p > 1$ we have for any $\epsilon > 0$

$$\mathbb{P}\left(M_n^{(2)}([0,1]) > \epsilon a_n\right) \leq n\mathbb{P}\left(X_0^{(2)} > \epsilon(\log w_n)^p\right) = o\left(n \cdot \exp\left\{-\epsilon(\log w_n)^p\right\}\right) \to 0$$

by (2.15). □

Proposition 4.1 implies that, in order to show (4.7), we need to prove that

$$\left(\frac{M_n^{(1)}(I_i) - b_n}{a_n}\right)_{i=1,\ldots,m} \Rightarrow \left(\mathcal{M}(I_i)\right)_{i=1,\ldots,m}\ , \tag{4.10}$$

which we now carry out. Consider a probability space (which we will denote, with some abuse of notation, by $(\Omega, \mathcal{F}, \mathbb{P})$) supporting i.i.d. random elements $\{Y^{(j,n)}\}_{j\in\mathbb{N}}$ distributed with the law μ_n in (2.16) for each $n = 1, 2, \ldots$, as well as i.i.d. random closed subsets of $[0,1]$, $\{\overline{R_j}\}_{j\in\mathbb{N}}$ distributed as in (2.7) such that, with $I_{j,n}$ defined by (2.17) we have

$$\frac{1}{n} I_{j,n} \to \overline{R_j} \quad \text{a.s. as} \quad n \to \infty \quad \text{for each} \quad j \in \mathbb{N}; \tag{4.11}$$

this is possible by (2.20) and the Skorohod embedding. The same probability space also supports a sequence $\{\Gamma_j\}_{j\in\mathbb{N}}$ of the arrival times of a unit rate Poisson process on $\mathbb{R}_+$, independent of $\{I_{j,n}, \overline{R_j} : j, n \in \mathbb{N}\}$. The following series representation of the process $\{X_t^{(1)}\}_{t\in\mathbb{Z}}$ is the key for our argument. It follows from Theorem 3.4.1 in Samorodnitsky (2016). For each $n \in \mathbb{N}$,

$$\left(X_t^{(1)}\right)_{0\leqslant t\leqslant n} \stackrel{d}{=} \left(\sum_{j=1}^{\infty} \tilde{V}\left(w_n/\Gamma_j\right) \mathbf{1}_{\{t\in I_{j,n}\}}\right)_{0\leqslant t\leqslant n}\ , \tag{4.12}$$

where

$$\tilde{V}(y) = \begin{cases} V(y) & \text{for } y > 1/\bar{\nu}(x_0) \\ 0 & \text{otherwise} \end{cases}\ . \tag{4.13}$$

When proving (4.10) we will simply assume that the process $\{X_t^{(1)}\}_{t\in\mathbb{Z}}$ is given by the right hand side of (4.12). Furthermore, for notational simplicity we will drop the "tilde" over V in the sequel, while keeping in mind that it vanishes for small values of the argument, as in (4.13). We now state several propositions that will prove (4.10).

For $k \in \mathbb{N}$ we define, in the notation of (2.18) and (2.19),

$$M_{n,(k)}(B) = \max_{t\in nB\cap\widehat{I}_{k,n}} \sum_{j=1}^{\infty} V\left(w_n/\Gamma_j\right) \mathbf{1}_{\{t\in I_{j,n}\}},$$

$$\mathcal{M}_{(k)}(B) = \begin{cases} -\log\Gamma_k & \text{if} \quad \overline{R_k} \cap B \neq \emptyset \\ -\infty & \text{otherwise} \end{cases}\ .$$

Proposition 4.2. *For each $k \in \mathbb{N}$ and each open interval I in $[0,1]$,*

$$\frac{M_{n,(k)}(I) - b_n}{a_n} \xrightarrow{P} \mathcal{M}_{(k)}(I). \tag{4.14}$$

We define, further, for $K \in \mathbb{N}$,

$$M_{n,K}(B) = \bigvee_{k=1}^{K} M_{n,(k)}(B),$$
$$\mathcal{M}_K(B) = \bigvee_{k=1}^{K} \mathcal{M}_{(k)}(B).$$

It follows from Proposition 4.2 that for each K and each open interval I in $[0,1]$,

$$\frac{M_{n,K}(I) - b_n}{a_n} \xrightarrow{P} \mathcal{M}_K(I)\,. \tag{4.15}$$

Since it is also clear that for any open interval I in $[0,1]$, as $K \to \infty$,

$$\mathcal{M}_K(I) \longrightarrow \mathcal{M}(I) \quad \text{a.s.}$$

if the limting sup-measure $\mathcal{M}$ is defined on the same probability space $(\Omega, \mathcal{F}, \mathbb{P})$ by (2.39), then the only remaining step to establish (4.10) is the following claim.

Proposition 4.3. *For any open interval I in $[0,1]$ and $\epsilon > 0$,*

$$\lim_{K\to\infty} \lim_{n\to\infty} \mathbb{P}\left(|M_{n,K}(I) - M_n(I)| \geqslant \epsilon\right) = 0\,. \tag{4.16}$$

We now prove Propositions 4.2 and 4.3.

Proof of Proposition 4.2. We will only consider the case $I = (a,b)$ for some $0 \leqslant a < b \leqslant 1$, the other cases being similar. Note that, by (2.22) it is enough to prove the proposition for $k = 1$. Since the normalized tail $\bar{\nu}$ is in the Gumbel maximum domain of attraction, the function V is Π-varying, so by (2.2) and (2.4) we have

$$\frac{V(w_n/\Gamma_1) - V(w_n)}{h \circ V(w_n)} \longrightarrow -\log \Gamma_1 \quad \text{a.s. as } n \to \infty. \tag{4.17}$$

Furthermore, by (4.11)

$$\mathbf{1}_{\{I_{1,n} \cap nI \neq \emptyset\}} \longrightarrow \mathbf{1}_{\{\overline{R_1} \cap I \neq \emptyset\}} \quad \text{a.s. as } n \to \infty.$$

Since $h \circ V(w_n) = o(V(w_n))$, we have

$$\frac{V\left(w_n/\Gamma_1\right) \mathbf{1}_{\{I_{1,n} \cap nI \neq \emptyset\}} - V(w_n)}{h \circ V(w_n)} \longrightarrow \mathcal{M}_{(1)}(I) \quad \text{a.s.}.$$

If we denote

$$S_{n,(1)}(I) = M_{n,(1)}(I) - V\big(w_n/\Gamma_1\big)\mathbf{1}_{\{I_{1,n}\cap nI\neq\emptyset\}}\,,$$

then the claim of the proposition will follow from the following two statements:

$$\limsup_{n\to\infty}\frac{S_{n,(1)}(I) - V\big(1/\overline{F}(n)\big)}{h\circ V(w_n)} \leqslant 0 \quad \text{in probability} \tag{4.18}$$

and

$$\mathbb{P}\left(\liminf_{n\to\infty}\frac{S_{n,(1)}(I) - V\big(1/\overline{F}(n)\big)}{h\circ V(w_n)} \geqslant 0 \,\Big|\, \overline{R_1}\cap I\neq\emptyset\right) = 1\,, \tag{4.19}$$

which we proceed to prove. We start with (4.18). Note that

$$0 \leqslant S_{n,(1)}(I) \leqslant \max_{t\in I_{1,n}}\sum_{j=2}^{\infty} V\big(w_n/\Gamma_j\big)\mathbf{1}_{\{t\in I_{j,n}\}} =: S_{n,(1)}.$$

Let $c_1, c_2 > 0$ be positive constants to be determined later and write $A_{c_1,n} = c_1 \log n/\overline{F}(n)$. Then

$$\begin{aligned}
&\mathbb{P}\left(S_{n,(1)} \geqslant V\left(A_{c_1,n}\right) + c_2 h\circ V\left(A_{c_1,n}\right)\right)\\
\leqslant & A_{c_1,n}\cdot\mathbb{P}\left(\sum_{j=2}^{\infty} V\big(w_n/\Gamma_j\big)\mathbf{1}_{\{0\in I_{j,n}\}} \geqslant V\left(A_{c_1,n}\right) + c_2 h\circ V\left(A_{c_1,n}\right)\right)\\
&+\mathbb{P}\left(\#I_{1,n} > A_{c_1,n}\right)\\
\leqslant & A_{c_1,n}\cdot\mathbb{P}\left(X_{0,n} \geqslant V\left(A_{c_1,n}\right) + c_2 h\circ V\left(A_{c_1,n}\right)\right) + \mathbb{P}\left(\#I_{1,n} > A_{c_1,n}\right)\\
=: & A_{c_1,n}\cdot B_1 + B_2.
\end{aligned}$$

By (A.9) $B_2 \to 0$ if c_1 is large enough. Further, $A_{c_1,n}\cdot B_1 \to e^{-c_2}$ by Proposition 0.9 in Resnick (1987). Therefore, fix any $\epsilon\in(0,1)$, we can choose $c_1, c_2 > 0$ such that

$$\mathbb{P}\left(S_{n,(1)} \leqslant V\left(A_{c_1,n}\right) + c_2 h\circ V\left(A_{c_1,n}\right)\right) \geqslant 1-\epsilon.$$

The claim (4.18) follows by (2.4) and parts (ii), (iii), (iv) of Proposition 3.1,

$$\begin{aligned}
V(A_{c_1,n}) - V\big(1/\overline{F}(n)\big) &= G(A_{c_1,n}) - G\big(1/\overline{F}(n)\big)\\
&\quad + o(h\circ G(A_{c_1,n})) + o\left(h\circ G\big(1/\overline{F}(n)\big)\right)\\
&\lesssim (\log\log n) h\circ G\big(1/\overline{F}(n)\big)\\
&\quad + o(h\circ G(A_{c_1,n})) + o\left(h\circ G\big(1/\overline{F}(n)\big)\right)\\
&= o(h\circ V(w_n))\,.
\end{aligned}$$

We now prove (4.19). Let $\Omega_1 = \{I_{1,n}\cap nI\neq\emptyset\}$. For a fixed $\omega_1\in\Omega_1$ we view $\big\{\mathbf{1}_{\{I_{1,n}(\omega_1)\cap I_{j,n}\cap nI\neq\emptyset\}} : j=2,3,\dots\big\}$ as a Bernoulli sequence with the success probability $\overline{p}_{n,T}(\omega_1)$ in (2.25). By Theorem 2.1 (ii) and (iii), for every $0 < \epsilon < 1$ and $\gamma > (1-2\beta)^{-1}$ we can choose new $c_1, c_2 > 0$ such that the event

$$D_1 := \left\{ \frac{c_1 L((\log n)^\gamma)}{w_n \overline{F}(n)(\log n)^{\gamma\beta}} \leqslant \overline{p}_{n,T} \leqslant \frac{c_2 \log n}{w_n \overline{F}(n)} \right\}$$

satisfies $\mathbb{P}\left(D_1 \,|\, \Omega_1\right) \geqslant 1-\epsilon$ for all n large enough.

For $\omega_1 \in \Omega_1$ we denote $j_1 = j_1(\omega_1) = \inf\{j \geqslant 2 : I_{j,n} \cap I_{1,n}(\omega_1) \cap nI \neq \emptyset\}$ and note that $S_{n,(1)}(I) \geqslant V(w_n/\Gamma_{j_1})$. Therefore, for any $c_3 > 0$ we have

$$\begin{aligned}
&\mathbb{P}\left(S_{n,(1)}(I) \geqslant V\left(\frac{c_1}{c_3}\cdot\frac{L((\log n)^\gamma)}{\overline{F}(n)(\log n)^{\gamma\beta}}\right) \,\Big|\, \overline{R_1}\cap I \neq \emptyset\right) \\
\geqslant&\mathbb{P}\left(D_1 \cap \left\{V(w_n/\Gamma_{j_1}) \geqslant V\left(\frac{c_1}{c_3}\cdot\frac{L((\log n)^\gamma)}{\overline{F}(n)(\log n)^{\gamma\beta}}\right)\right\} \,\Big|\, \overline{R_1}\cap I \neq \emptyset\right) \\
\geqslant&\mathbb{P}\left(D_1 \cap \left\{\Gamma_{j_1} \leqslant c_3(\overline{p}_{n,T})^{-1}\right\} \,\Big|\, \overline{R_1}\cap I \neq \emptyset\right) \\
\geqslant&\mathbb{P}\left(D_1 \cap \left\{j_1 \leqslant (c_3/2)(\overline{p}_{n,T})^{-1}\right\} \,\Big|\, \overline{R_1}\cap I \neq \emptyset\right) - \epsilon \\
\geqslant&\mathbb{P}\big(D_1 \,|\, \overline{R_1}\cap I \neq \emptyset\big) - 2\epsilon \geqslant \mathbb{P}\big(D_1 \,|\, \Omega_1\big) - 3\epsilon \geqslant 1-4\epsilon\,,
\end{aligned}$$

for large n, where the 3rd inequality follows from the law of large numbers, the 4th inequality follows from the Markov inequality if c_3 large enough, and the penultimate inequality follows from (4.11). Since we can take ϵ as small as we wish, it is enough to show that

$$\left| V\left(\frac{c_1}{c_3}\cdot\frac{L((\log n)^\gamma)}{\overline{F}(n)(\log n)^{\gamma\beta}}\right) - V\left(1/\overline{F}(n)\right)\right| = o\left(h\circ V(w_n)\right)\,. \tag{4.20}$$

To this end, choose any $\alpha > \gamma\beta$ and note that for large n, by parts (iii) and (iv) of Proposition 3.1, the expression in the left-hand side does not exceed

$$\begin{aligned}
&V\left(1/\overline{F}(n)\right) - V\left((\log n)^{-\alpha}/\overline{F}(n)\right) \\
\lesssim&(\log\log n)\, h\circ V\left(1/\overline{F}(n)\right) \ll h\circ V(w_n)\,,
\end{aligned}$$

as required. □

Proof of Proposition 4.3. We start by fixing a small constant ρ and setting

$$i_n = \left\lfloor \frac{\rho\log w_n}{\zeta(w_n)}\right\rfloor. \tag{4.21}$$

The first step is to establish the following claim, that shows that for large k, $M_{n,(k)}(I)$ is not likely to become the overall maximum $M_n(I)$.

$$\lim_{i_0\to\infty,\,K\to\infty}\ \limsup_{n\to\infty}\mathbb{P}\left(\max_{2^{i_0}\leqslant k<2^{i_n}} M_{n,(k)}(I) > M_{n,K}(I)\right) = 0\,. \tag{4.22}$$

To this end, observe that, by Proposition 4.2, for any $\epsilon \in (0,1)$ we can choose $C_1 > 0$ large enough so that for all K large enough,

$$\lim_{n\to\infty} \mathbb{P}\left(M_{n,K}(I) \geqslant b_n - C_1 a_n\right) \geqslant 1-\epsilon\,.$$

Next, for $c>0$ let $A_{c,n} = c\log n/\overline{F}(n)$. For $i_0 \leqslant j \leqslant i_n - 1$, let H_j be the event

$$\bigcap_{k=2^j}^{2^{j+1}-1} \left\{M_{n,(k)}(I) \leqslant V\big(2w_n/k\big) + V\left(A_{c,n}\right) + 2jh\circ V(A_{c,n})\right\}.$$

We claim that, given $0<\epsilon<1$, we can find $c>0$ such that

$$\lim_{i_0\to\infty} \liminf_{n\to\infty} \mathbb{P}\left(\bigcap_{j=i_0}^{i_n-1} H_j\right) \geqslant 1-\epsilon. \tag{4.23}$$

Assuming, for a moment, that this is true, the claim (4.22) will follow once we check that for all j and n large enough,

$$V\left(w_n\right) - V\big(w_n/2^{j-1}\big) - C_1 a_n > V(A_{c,n}) + 2jh\circ V(A_{c,n}) - V(1/\overline{F}(n)). \tag{4.24}$$

Indeed, by (2.4) and part (v) of Proposition 3.1,

$$\begin{aligned}
&V(w_n) - V\big(w_n/2^{j-1}\big) - C_1 a_n\\
=&G(w_n) - G\big(w_n/2^{j-1}\big) - (C_1 + o(1))h\circ G(w_n)\\
\gtrsim&\, jh\circ G(w_n)\,,
\end{aligned}$$

while by part (iv) of Proposition 3.1,

$$\begin{aligned}
&V(A_{c,n}) + 2jh\circ V(A_{c,n}) - V(1/\overline{F}(n))\\
=&G(A_{c,n}) + 2jh\circ G(A_{c,n}) - G(1/\overline{F}(n)) + o(jh\circ G(A_{c,n}))\\
\lesssim&\,(j+\log\log n)h\circ V(A_{c,n})\,.
\end{aligned}$$

By part (iii) of Proposition 3.1 this gives (4.24), and, hence, (4.22), so we now prove (4.23). Switching to the complements, we will show that

$$\lim_{i_0\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\bigcup_{j=i_0}^{i_n-1} H_j^c\right) \leqslant \epsilon\,. \tag{4.25}$$

Recall that

$$M_{n,(k)} \leqslant V\big(w_n/\Gamma_k\big) + S_{n,(k)}, \quad S_{n,(k)} := \max_{t\in I_{k,n}} \sum_{j=k+1}^{\infty} V\big(w_n/\Gamma_j\big)1_{\{t\in I_{j,n}\}}.$$

Therefore, for each $i_0 \leqslant j < i_n$, $H_j^c \subseteq \bigcup_{k=2^j}^{2^{j+1}-1}\left(U_k \cup D_k \cup L_k\right)$ with

$$\begin{aligned}
U_k &= \left\{\Gamma_k \leqslant k/2\right\},\\
D_k &= \left\{\#I_{k,n} > A_{c,n}\right\},\\
L_k &= \left\{S_{n,(k)} \geqslant V(A_{c,n}) + 2jh\circ V(A_{c,n}),\, \#I_{k,n} \leqslant A_{c,n}\right\}.
\end{aligned}$$

Trivially,

$$\sum_{k=1}^{\infty} \mathbb{P}(U_k) < \infty\,, \tag{4.26}$$

and, if c is large enough, then by (A.9) we also have

$$\sum_{k=1}^{\infty} \mathbb{P}(D_k) < \infty\,. \tag{4.27}$$

Next, as $\mathbb{P}(X_0^{(1)} > x) \sim \overline{\nu}(x)$ by subexponentiality and $(A_{c,n})^{-1} \asymp \overline{\nu}(V(A_{c,n}))$, we have for $2^j \leqslant k < 2^{j+1}$,

$$\begin{aligned}
\mathbb{P}(L_k) \leqslant & A_{c,n} \cdot \mathbb{P}\left(X_0^{(1)} \geqslant V(A_{c,n}) + 2jh\circ V(A_{c,n})\right) \\
\lesssim & \frac{\overline{\nu}\left(V(A_{c,n}) + 2jh\circ V(A_{c,n})\right)}{\overline{\nu}\left(V(A_{c,n})\right)} \\
\lesssim & \exp\left\{-\int_0^{2j} \frac{h\circ V(A_{c,n})}{h\big[V(A_{c,n}) + uh\circ V(A_{c,n})\big]} du\right\} \\
\lesssim & \exp\left\{-\frac{2jh\circ V(A_{c,n})}{h\big[V(A_{c,n}) + 2(i_n-1)h\circ V(A_{c,n})\big]}\right\}\,.
\end{aligned}$$

Note that by Assumption 3.1 ($B1$), for some constant C, for large n,

$$\begin{aligned}
2(i_n-1)h\circ V(A_{c,n}) \sim & 2i_n h\circ G(A_{c,n}) \\
& \sim \frac{2\rho \log w_n}{\zeta(w_n)} \cdot \frac{\zeta(A_{c,n})}{\log A_{c,n}} V(A_{c,n}) \leqslant C\rho V(A_{c,n})\,,
\end{aligned}$$

so we can choose ρ small enough so that

$$\mathbb{P}(L_k) \leqslant \exp\left\{-2j\frac{h\circ V(A_{c,n})}{h\big[(1+C\rho)V(A_{c,n})\big]}\right\} \leqslant e^{-j}$$

because h is assumed to satisfy (3.5). It follows that

$$\sum_{k=1}^{\infty} \mathbb{P}(L_k) < \infty$$

which, together with (4.26) and (4.27), proves (4.25), so we have established (4.22). Now the claim of Proposition 4.3 will follow from the following statement that we prove next.

We claim that, with i_n given, once again, by (4.21),

$$\lim_{K\to\infty} \limsup_{n\to\infty} \mathbb{P}\left(\max_{k\geqslant 2^{i_n}} M_{n,(k)} < M_{n,K}(I)\right) = 1\,.$$

Since $b_n \sim G(w_n)$ and $a_n = o(b_n)$, by (4.15) it is enough to show that for some $\eta \in (0,1)$,

$$\lim_{n\to\infty} \mathbb{P}\left(\max_{k\geqslant 2^{i_n}} M_{n,(k)} \leqslant \eta G(w_n) \right) = 1\,. \tag{4.28}$$

To this end, choose $0 < r < (1-2\beta)/2$ and write

$$\begin{aligned}\max_{k\geqslant 2^{i_n}} M_{n,(k)} &\leqslant \max_{0\leqslant t\leqslant n} \sum_{j=2^{i_n}}^{\lfloor n^r \rfloor} V\big(w_n/\Gamma_j\big)\mathbf{1}_{\{t\in I_{j,n}\}} \\ &\quad + \max_{0\leqslant t\leqslant n} \sum_{j=\lfloor n^r\rfloor+1}^{\infty} V\big(w_n/\Gamma_j\big)\mathbf{1}_{\{t\in I_{j,n}\}} =: T_{1,n} + T_{2,n}\,.\end{aligned}$$

By the choice of r,

$$\mathbb{P}\left(\max_{0\leqslant t\leqslant n} \sum_{j=2^{i_n}}^{\lfloor n^r \rfloor} \mathbf{1}_{\{t\in I_{j,n}\}} \geqslant 2 \right) \leqslant n\mathbb{P}\left(\sum_{j=1}^{\lfloor n^r \rfloor} \mathbf{1}_{\{0\in I_{j,n}\}} \geqslant 2 \right) \lesssim \frac{n^{2r+1}}{w_n^2} \to 0.$$

Therefore, with probability increasing to 1,

$$T_{1,n} \leqslant V\big(w_n/\Gamma_{2^{i_n}}\big) \lesssim G\big(w_n/2^{i_n-1}\big)\,.$$

By Assumption 3.1 ($B4$),

$$\limsup_{n\to\infty} \frac{G(w_n/2^{i_n-1})}{G(w_n)} < 1\,,$$

so (4.28) will be established once we prove that for any $\epsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}\left(T_{2.n} > \epsilon G(w_n)\right) = 0\,.$$

The latter statement will follow from the following claim:

$$\lim_{n\to\infty} n\cdot \mathbb{P}\left(\sum_{\Gamma_j > n^r} V\big(w_n/\Gamma_j\big)\mathbf{1}_{\{0\in I_{j,n}\}} > \epsilon G(w_n) \right) = 0\,.$$

Since for some $s > 0$ $V(x) \leq G(sx)$ for all x, we will prove instead that

$$\lim_{n\to\infty} n\cdot \mathbb{P}\left(\sum_{\Gamma_j > n^r} G\big(sw_n/\Gamma_j\big)\mathbf{1}_{\{0\in I_{j,n}\}} > \epsilon G(w_n) \right) = 0\,. \tag{4.29}$$

To this end, denote for $n \in \mathbb{N}$,

$$\widetilde{x}_n = \epsilon G(w_n), \quad x_n = G(sw_n/n^r), \quad m_0 = \lfloor \widetilde{x}_n / x_n \rfloor,$$

and define $H_n : (x_0,\infty)^n \to \mathbb{R}$ by

$$H_n(z_1,\dots,z_n) = \int_{x_0}^{z_1} \frac{du}{h(u)} + \dots + \int_{x_0}^{z_n} \frac{du}{h(u)} := q(z_1) + \dots + q(z_n)\,.$$

Let N_n be a Poisson random variable with mean $s(1-n^r s^{-1} w_n^{-1})$ (positive for large n). If $\{\xi_i\}_{i=1}^{\infty}$ is a family of i.i.d. random variables independent of N_n with distribution equal to normalized $H_\#$ restricted to the interval (x_0,x_n). Then

$$\sum_{\Gamma_j>n^r} G\big(sw_n/\Gamma_j\big)\mathbf{1}_{\{0\in I_{j,n}\}} \overset{d}{=} \sum_{i=1}^{N_n} \xi_i\,, \tag{4.30}$$

so that

$$\begin{aligned}&\mathbb{P}\left(\sum_{\Gamma_j>n^r} G\big(sw_n/\Gamma_j\big)\mathbf{1}_{\{0\in I_{j,n}\}} > \widetilde{x}_n\right)\\ &= \sum_{d=1}^{\infty} \mathbb{P}\left(N_n = m_0+d\right)\mathbb{P}\left(\sum_{i=1}^{m_0+d} \xi_i > \widetilde{x}_n\right) =: \sum_{d=1}^{\infty} B_d\cdot Q_d\,. \end{aligned}\tag{4.31}$$

Clearly,

$$B_d \leqslant s^{m_0+d}/(m_0+d)!\,. \tag{4.32}$$

On the other hand, for some constant $c>0$,

$$\begin{aligned}Q_d =& \int_{(x_0,x_n)^{m_0+d}} \mathbf{1}_{\{\sum_{i=1}^{m_0+d} z_i > \widetilde{x}_n\}} \prod_{i=1}^{m_0+d} P_{\xi_i}(dz_i)\\ \leqslant& c^{m_0+d}\int_{(x_0,x_n)^{m_0+d}} \mathbf{1}_{\{\sum_{i=1}^{m_0+d} z_i > \widetilde{x}_n\}} \prod_{i=1}^{m_0+d} H_\#(dz_i)\\ =& c^{m_0+d}\int_{(x_0,x_n)^{m_0+d}} \mathbf{1}_{\{\sum_{i=1}^{m_0+d} z_i > \widetilde{x}_n\}} \prod_{i=1}^{m_0+d} \exp\{-q(z_i)\}q'(z_i)\,dz_i\\ \leqslant& \left(cq(x_n)\right)^{m_0+d}\exp\left(-\inf\left\{\sum_{i=1}^{m_0+d} q(z_i) : \sum_{i=1}^{m_0+d} z_i > \widetilde{x}_n, x_0<z_i<x_n\right\}\right)\,.\end{aligned}$$

To evaluate the infimum inside the above exponential, note that that the function H_n is increasing and concave in all of its variables. Hence its infimum is achieved at a boundary point which will have, say, k_d coordinates equal to x_n, m_0+d-k_d-1 coordinates equal to x_0, and a final coordinate that makes the sum of all coordinates equal to $\widetilde{x}_n$, for the smallest possible value of k_d that makes it possible. That means that

$$\inf\left\{\sum_{i=1}^{m_0+d} q(z_i) : \sum_{i=1}^{m_0+d} z_i > \widetilde{x}_n, x_0<z_i<x_n\right\} \geq k_d q(x_n)\,. \tag{4.33}$$

Clearly,

$$k_d = \left[\left[\frac{\widetilde{x}_n - x_n - (m_0 + d - 1)x_0}{x_n - x_0}\right]\right]_+ \geqslant \frac{\widetilde{x}_n - x_n}{2(x_n - x_0)} \tag{4.34}$$

if

$$d \leqslant \frac{(\widetilde{x}_n - x_n)}{2x_0} - m_0 + 1 \,. \tag{4.35}$$

Notice that by (4.32), the part of the sum in (4.31) corresponding to d outside of the above range does not exceed, for large n,

$$\sum_{d \geqslant \widetilde{x}_n / x_0} s^{m_0+d}/(m_0+d)! = o(1/n)$$

by part (i) of Proposition 3.1. On the other hand, for large n, for d in the range (4.35), $k_d \geq m_0/3$ by (4.34). Therefore, the part of the sum in (4.31) corresponding to d in the range (4.35) can be bounded by

$$\sum_{d=1}^{\infty} \frac{(csq(x_n))^{m_0+d}}{(m_0+d)!} \exp\{-m_0 q(x_n)/3\} = \exp\big\{ - q(x_n)(m_0/3 - cs)\big\} \,.$$

Since $m_0 \gg \log\log n \to \infty$ by part (ii) of Proposition 3.1, and

$$q(x_n) = -\log \overline{H_{\#}}(G(sw_n/n^r)) \sim (1-\beta-r)\log n \,,$$

the part of the sum in (4.31) corresponding to d in the range (4.35) is also $o(1/n)$, proving (4.29) and, hence, completing the proof of Proposition 4.3. □

Proof of Theorem 4.2. We need to prove that for any fixed $0 < T_1 < T_2 < \infty$,

$$\left(\frac{\max_{s \leq nt} X_s - b_n}{a_n}, \, T_1 \leqslant t \leqslant T_2 \right) \Rightarrow \big(X_{\mathrm{G}}(t^{1-\beta}), \, T_1 \leqslant t \leqslant T_2 \big)$$

weakly in the Skorohod J_1 topology on $D[T_1, T_2]$, and without loss of generality we assume that $T_2 \leq 1$. According to (4.4) and Proposition 4.1, is the same as proving

$$\left(\frac{M_n^{(1)}([0,t]) - b_n}{a_n}, \, T_1 \leqslant t \leqslant T_2 \right) \Rightarrow \big(\mathcal{M}([0,t]), \, T_1 \leqslant t \leqslant T_2 \big) \tag{4.36}$$

in the same space. We construct all the random objects in (4.5) on the same probability space as in the proof of Theorem 4.1 and prove a.s convergence in $D[T_1, T_2]$. In the course of the proof of the latter theorem we have shown that for every $\varepsilon > 0$ there is $K \geq 1$ such that

$$\limsup_{n\to\infty} \mathbb{P}\Big[\big(M_n^{(1)}([0,t]), \, T_1 \leqslant t \leqslant T_2 \big) \neq \big(M_{n,K}([0,t]), \, T_1 \leqslant t \leqslant T_2 \big) \Big] \leq \varepsilon \,.$$

Since, clearly,

$$\lim_{K\to\infty} \mathbb{P}\Big[\big(\mathcal{M}_K([0,t]), \, T_1 \leqslant t \leqslant T_2 \big) = \big(\mathcal{M}([0,t]), \, T_1 \leqslant t \leqslant T_2 \big) \Big] = 1 \,.$$

(4.36) will follow once we prove that for every $K = 1, 2, \dots$

$$\left(\frac{M_{n,K}([0,t]) - b_n}{a_n}, \, T_1 \leqslant t \leqslant T_2 \right) \to \big(\mathcal{M}_K([0,t]), \, T_1 \leqslant t \leqslant T_2 \big)$$

a.s. in $D[T_1, T_2]$ as $n \to \infty$. The stochastic process in the right hand side may take the value $-\infty$; the probability of this converges to zero as $K \to \infty$. For nondecreasing functions the value of $-\infty$ introduces no difficulties in the J_1 topology if one interpretes $(-\infty) - (-\infty)$ as zero. The assumption $0 < \beta < 1/2$ implies that the stable regenerative sets $(\overline{R}_j)$ are a.s. disjoint, so the latter statement will follows from

$$\left(\frac{M_{n,(k)}([0,t]) - b_n}{a_n}, \, T_1 \leqslant t \leqslant T_2 \right) \to \big(\mathcal{M}_{(k)}([0,t]), \, T_1 \leqslant t \leqslant T_2 \big)$$

a.s. in $D[T_1, T_2]$ for every $k \geqslant 1$ and, as before, it is enough to consider the case $k = 1$. As in the proof of Proposition 4.2, we only need to check that

$$\left(\frac{V\,(w_n/\Gamma_1)\,\mathbf{1}_{\{I_{1,n} \cap [0,nt] \neq \emptyset\}} - V(w_n)}{h \circ V(w_n)}, \, T_1 \leqslant t \leqslant T_2 \right) \to \big(\mathcal{M}_{(1)}([0,t]), \, T_1 \leqslant t \leqslant T_2 \big) \tag{4.37}$$

a.s.. However, it follows from (4.11) that, a.s.,

$$\inf\{I_{1,n}/n\} \to \inf\{\overline{R}_1\} \,.$$

Together with (4.17) this establishes (4.37), as required. □

Appendix A: Random Walks with Regularly Varying Tails

Among the major goals of this appendix is to prove Theorem 2.1. We start with recalling certain results on the ranges of the random walks from Barlow and Taylor (1992) and Samorodnitsky and Wang (2019). We consider a random walk $\{S_n\}_{n \geqslant 0}$ with N_0-valued steps $\{\xi_n\}_{n \geqslant 1}$ whose distribution F satisfies Assumption 2.1. Recall the standard notions

(a) the *range* $A = \{S_n : n = 0, 1, 2, \dots\}$,
(b) the *sojourn time* in F up to time k, $T_F(k) = \#\{0 \leqslant n \leqslant k : S_n \in F\}$, for $F \subset \mathbb{N}_0$, $k \in \mathbb{N} \cup \{\infty\}$.

The following properties are well-known; see e.g. Appendix A in Samorodnitsky and Wang (2019). As $n \to \infty$,

$$\mathbb{E}_0 \, T_{[0,n]}(\infty) \sim \frac{n^\beta}{\Gamma(1+\beta)\Gamma(1-\beta)L(n)} \,, \tag{A.1}$$

$$\mathbb{P}_0 \left(A \cap \{n\} \neq \emptyset \right) \sim \frac{n^{\beta-1}\mathbb{P}(\xi_1 > 0)}{\Gamma(\beta)\Gamma(1-\beta)L(n)} \,. \tag{A.2}$$

For $F \subset \mathbb{Z}$ we denote by $D(F) := \{x - y : x, y \in F\}$ its *difference set*. For every $\delta \in (0,1)$, there exists $c_0 = c_0(\delta) > 0$ such that for every F and every $k \in \mathbb{N} \cup \{\infty\}$ with $0 < \mathbb{E}_0(T_{D(F)}(k)) < \infty$, we have

$$\mathbb{P}_x\left(T_F(k) \geqslant c\,\mathbb{E}_0(T_{D(F)}(k))\right) \leqslant e^{-c\delta} \quad \text{for each} \quad x \in F \text{ and } c > c_0\,, \tag{A.3}$$

see e.g. Lemma 3.1 in Pruitt and Taylor (1969). Choosing, in particular, $F = F_n = \{0, 1, \ldots, n\}$, we have by (A.1),

$$\mathbb{E}_0 T_{D(F_n)}(\infty) = \mathbb{E}_0 T_{F_n}(\infty) \lesssim \frac{n^\beta}{L(n)}.$$

Therefor, choosing in (A.3) $\delta = 1/2$, we see that for any $C > 0$ we can choose $c > 0$ so that for all $n \in \mathbb{N}$,

$$\mathbb{P}_0\left(\#(A \cap [0,n]) \geqslant \frac{cn^\beta \log n}{L(n)}\right) \leqslant \mathbb{P}_0\left(T_{F_n}(\infty) \geqslant \frac{cn^\beta \log n}{L(n)}\right) \leqslant n^{-C}. \tag{A.4}$$

Furthermore, for a sufficiently large c, for each n, we have simultaneously for all $k = 0, 1, \ldots, n$ that

$$\mathbb{P}_0\left(\#(A \cap [0, 2^k)) \geqslant \frac{cn2^{\beta k}}{L(2^k)}\right) \leqslant e^{-n}\,. \tag{A.5}$$

Lemma A.1. *Assume that Assumption 2.1 holds and $S_0 = 0$. Then*

$$\limsup_{n_0 \to \infty} \sup_{n > n_0} \max_{0 \leqslant k \leqslant n-1} \max_{m \in \mathbb{Z}} \frac{\#(A \cap [m, m+2^k) \cap [2^{n_0}, 2^n))}{n2^{\beta k}/L(2^k)} < \infty \quad a.s. \tag{A.6}$$

Proof. Let c be such that (A.5) holds. Then

$$\begin{aligned}
&\mathbb{P}\left(\sup_{n > n_0} \max_{0 \leqslant k \leqslant n-1} \max_{m \in \mathbb{Z}} \frac{\#(A \cap [m, m+2^k) \cap [2^{n_0}, 2^n))}{n2^{\beta k}/L(2^k)} \geq c\right) \\
&\leq \sum_{n > n_0} n2^n \cdot \max_{\substack{0 \leqslant k \leqslant n-1 \\ [m, m+2^k) \subset [2^{n_0}, 2^n)}} \mathbb{P}\left(\#(A \cap [m, m+2^k) \cap [2^{n_0}, 2^n)) \geq \frac{cn2^{\beta k}}{L(2^k)}\right) \\
&\leq \sum_{n > n_0} n2^n \max_{0 \leqslant k \leqslant n-1} \mathbb{P}\left(\#(A \cap [0, 2^k)) \geqslant \frac{cn2^{\beta k}}{L(2^k)}\right) \leq \sum_{n > n_0} n(2/e)^n\,,
\end{aligned}$$

where the 1st inequality follows by the union bound since

$$n2^n \geqslant \#\{[m, m+2^k) : [m, m+2^k) \subset [2^{n_0}, 2^n), k = 0, \ldots, n-1.\},$$

the 2nd inequality follows from the strong Markov property, and the last one follows from (A.5). Since the last expression is summable in n_0, (A.6) follows by the first Borel-Cantelli Lemma.

□

Lemma A.2. *Assume that Assumption 2.1 holds and $S_0 = 0$. For any $\eta, \gamma > 0$,*

$$\#\left\{k : S_k \leqslant \eta n,\, \xi_k \geqslant (\log n)^\gamma\right\} \gtrsim_P \frac{n^\beta L\left((\log n)^\gamma\right)}{(\log n)^{\gamma\beta} L(n)}\,. \tag{A.7}$$

Proof. Let $N_t = \max\{k : S_k \leqslant t\} + 1$, $t \geq 0$. Then for each $x > 0$, as $m \to \infty$,

$$\mathbb{P}\left(\overline{F}(m) N_m \geqslant x^{-\beta}\right) \to J_\beta(x),$$

where J_β is an $\mathbb{R}_+$-supported strictly β-stable distribution; see e.g. Feller (1966) XI .5 (5.6). Therefore, for any $\epsilon > 0$ we can choose $c > 0$ so small that with $m_n = \lceil c n^\beta / L(n) \rceil$ we have $\liminf_{n\to\infty} P(B_n) > 1 - \epsilon$ for the events $B_n = \{N_{\eta n} \geqslant m_n\}$, $n \geq 1$. Consider also the events

$$D_n = \left\{ \frac{1}{m_n} \sum_{k=1}^{m_n} 1_{\{\xi_k > (\log n)^\gamma\}} < \frac{\overline{F}\left((\log n)^\gamma\right)}{2} \right\},\ n = 1, 2, \ldots .$$

By Chebyshev's inequality, as $n \to \infty$,

$$\mathbb{P}(D_n) \lesssim \frac{m_n \cdot \mathrm{var}(1_{\{\xi_k > (\log n)^\gamma\}})}{\left(m_n \overline{F}\left((\log n)^\gamma\right)\right)^2} \lesssim \frac{L(n)}{n^\beta} \cdot \frac{(\log n)^{\gamma\beta}}{L\left((\log n)^\gamma\right)} \to 0\,.$$

Hence $\liminf_{n\to\infty} \mathbb{P}(B_n \cap D_n^c) \geqslant 1 - \epsilon$. However, on the event $B_n \cap D_n^c$

$$\#\left\{k : S_k \leqslant \eta n,\, \xi_k \geqslant (\log n)^\gamma\right\} \geqslant \frac{c}{2} \frac{n^\beta L\left((\log n)^\gamma\right)}{(\log n)^{\gamma\beta} L(n)}\,,$$

leading to the desired conclusion. □

Proof of Theorem 2.1 (i). We use the notation $T = (a, b)$ throughout the proof. Let $\{Y^{(1)}\}_{t\in\mathbb{Z}}$ and $\{Y^{(2)}\}_{t\in\mathbb{Z}}$ be i.i.d. Markov chains on $\mathbb{Z}$ starting at 0, satisfying Assumption 2.1. The simultaneous visit times of the two chains to 0,

$$\varphi_j^* = \inf\{n \geqslant \varphi_{j-1} + 1 : Y_n^{(1)} = Y_n^{(2)} = 0\},\ j = 1, 2, \ldots$$

with $\varphi_0^* = 0$ and $\varphi_j^* = \infty$ if $\varphi_{j-1}^* = \infty$, form a terminating (since $\beta < 1/2$) renewal process. We denote by $\overline{F^*}$ the tail distribution of $\varphi^* := \varphi_1^*$. By the last entrance decomposition,

$$\begin{aligned} p_{n,T} &= \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \mathbb{P}\left(Y_k^{(1)} = Y_k^{(2)} = 0,\ \text{no simultaneous returns after } k \text{ in } nT\right) \\ &= \sum_{k=\lceil na \rceil}^{\lfloor nb \rfloor} \frac{1}{w_n^2} \overline{F^*}(\lfloor nb \rfloor - k) \asymp \frac{n}{w_n^2} \end{aligned}$$

since $\overline{F^*}(\infty) > 0$. This proves (2.26).

In the remaining part of the proof we will use the following simple observation that allows us to use the basic facts about random walks with regularly varying tails describing earlier in this section. For a fixed $n \in \mathbb{N}$ we construct a random walk $\{S_k^{(n)}\}_{k\geqslant 0}$ by choosing the initial state distributed as $\min I_{1,n}$ and the steps with the distribution F in (2.13). Recall that

$$\mathbb{P}(\min I_{1,n} \le nx) = \frac{w_{[nx]}}{w_n} \quad \text{for } 0 < x < 1. \tag{A.8}$$

The range A_n of $\{S_k^{(n)}\}_{k\geqslant 0}$, obviously, satisfies

$$A_n \cap [0,m] \stackrel{d}{=} I_{1,n} \cap [0,m] \text{ for all } m \le n\,.$$

By conditioning on $S_0^{(n)}$ and using (A.4), we see that for any $C > 0$ we can choose $c > 0$ so that for all $n \in \mathbb{N}$,

$$\mathbb{P}\left(\#I_{1,n} \geqslant \frac{cn^\beta \log n}{L(n)}\right) \leqslant n^{-C}\,. \tag{A.9}$$

Proof of Theorem 2.1 (ii). For $c > 0$ denote $a_{c,n} := cn^\beta \log n/L(n)$ and consider the event $B_n = \{\#(I_{1,n} \cap nT) \leqslant a_{c,n}\}$. On B_n, the "union bound method" shows that for large n,

$$\overline{p}_{n,T} \leqslant \frac{\#(I_{1,n} \cap nT)}{w_n} \leqslant \frac{a_{c,n}}{w_n}.$$

Therefore it suffices to show that $\mathbb{P}(B_n^c) \leqslant n^{-C}$ if c is large enough. This, however, follows immediately from (A.9).

Proof of Theorem 2.1 (iii). By the measure preserving property of the shift θ it is enough to consider intervals of the form $T = (0,b)$. Let $v_1 < v_2 < v_3 < \cdots$ be the enumeration of the points of $I_{1,n}$ in the increasing order. We construct a subset of by $I_{1,n}$

$$I_{1,n,\gamma} := \{v_i \in I_{1,n} : v_{i+1} - v_i \geqslant (\log n)^\gamma\} \tag{A.10}$$

(not including the last point in $I_{1,n}$.) For an $\omega_1 \in \{I_{1,n} \cap nT \neq \emptyset\}$, a lower bound for $\overline{p}_{n,T}(\omega_1)$ is derived below, where for typographical convenience we use the notation $\mathbb{P}_2$ to denote the probability measure associated with $Y^{(2,n)}$.

$$\begin{aligned}
\overline{p}_{n,T}(\omega_1) &\geqslant \mathbb{P}_2\left(I_{1,n,\gamma}(\omega_1) \cap I_{2,n} \cap nT \neq \emptyset\right) \\
&\geqslant \sum_{u \in I_{1,n,\gamma}(\omega_1) \cap nT} \mathbb{P}_2\left(u = \max(I_{1,n,\gamma}(\omega_1) \cap I_{2,n})\right) \\
&= \frac{1}{w_n} \sum_{u \in I_{1,n,\gamma}(\omega_1) \cap nT} \mathbb{P}_2\left(I_{1,n,\gamma}(\omega_1) \cap I_{2,n} \cap (u,\infty) \cap nT = \emptyset \,\middle|\, u \in I_{2,n}\right)\,.
\end{aligned}$$

Now the claim of part (iii) of the theorem follows from the following two statements.

For any $\epsilon \in (0,1)$, there exists $c = c(\epsilon) > 0$ such that for all $\gamma > 0$

$$\liminf_{n\to\infty} \mathbb{P}\left(\#(I_{1,n,\gamma} \cap nT) \geqslant d_n \,\middle|\, I_{1,n} \cap nT \neq \emptyset\right) \geqslant 1-\epsilon\,, \tag{A.11}$$

where

$$d_n = \frac{cn^\beta L\left((\log n)^\gamma\right)}{(\log n)^{\gamma\beta} L(n)}\,.$$

Further we claim that, if $\gamma > (1-2\beta)^{-1}$, then for every $0 < \epsilon < 1$ there is an event B with $\mathbb{P}(B) > 1-\epsilon$ such that for every $w_1 \in B$,

$$\sup_{u\in I_{1,n,\gamma}(\omega_1)\cap nT} \mathbb{P}_2\left(I_{1,n,\gamma}(\omega_1) \cap I_{2,n} \cap (u,\infty) \cap nT \neq \emptyset \,\middle|\, u \in I_{2,n}\right) = o_P(1)\,. \tag{A.12}$$

These two statements are proved in the remainder of this section.

Fix $\epsilon \in (0,1)$. For any $\eta \in (0,1)$ we have by (2.20) and (2.6),

$$\begin{aligned}&\mathbb{P}\left(\frac{1}{n}I_{1,n} \cap \eta nT \neq \emptyset \,\middle|\, \frac{1}{n}I_{1,n} \cap nT \neq \emptyset\right)\\ \to&\mathbb{P}\left(\overline{R_1} \cap \eta T \neq \emptyset \,\middle|\, \overline{R_1} \cap T \neq \emptyset\right) = \eta^{1-\beta}.\end{aligned}$$

Therefore, if η is sufficiently close to 1,

$$\liminf_{n\to\infty} \mathbb{P}\left(\frac{1}{n}I_{1,n} \cap \eta nT \neq \emptyset \,\middle|\, \frac{1}{n}I_{1,n} \cap nT \neq \emptyset\right) \geqslant \sqrt{1-\epsilon}. \tag{A.13}$$

Note that

$$\begin{aligned}&\mathbb{P}\left(\#(I_{1,n,\gamma} \cap nT) \geqslant d_n \,\middle|\, \frac{1}{n}I_{1,n} \cap \eta nT \neq \emptyset\right)\\ =&\sum_{i\in\eta nT} P\big(S_0 = i\big|S_0 \in \eta nT\big)\, \mathbb{P}_i\left(\#\{k : \xi_k \geqslant (\log n)^\gamma, S_k \leqslant nb - i\} \geqslant d_n\right)\\ \geqslant&\mathbb{P}_0\left(\#\{k : \xi_k \geqslant (\log n)^\gamma, S_k \leqslant \lfloor n(1-\eta)b\rfloor\} \geqslant d_n\right)\,.\end{aligned}$$

We conclude by Lemma A.2 that a fixed $\eta \in (0,1)$ for which (A.13) holds, we can choose c such that

$$\liminf_{n\to\infty} \mathbb{P}\left(\#(I_{1,n,\gamma} \cap nT) \geqslant d_n \,\middle|\, I_{1,n} \cap \eta nT \neq \emptyset\right) \geqslant \sqrt{1-\epsilon}\,. \tag{A.14}$$

Clearly, (A.13) and (A.14) give us (A.11), so it remains to prove (A.12).

Let k_1 and k_2 be such that $2^{k_1} \leqslant (\log n)^\gamma < 2^{k_1+1}$ and $2^{k_2-1} \leqslant n < 2^{k_2}$. Let $u \in I_{1,n,\gamma}(\omega_1) \cap nT$ and denote by $\overline{q}_n(u|\omega_1)$ the probability in the left hand side of (A.12). We have

$$\begin{aligned}&\overline{q}_n(u|\omega_1)\\ \leqslant&\sum_{k=k_1}^{k_2} \mathbb{P}\left(s \in I_{2,n} \text{ for some } s \in [u+2^k, u+2^{k+1}) \cap I_{1,n}(\omega_1) \,\middle|\, u \in I_{2,n}\right)\\ \leqslant&\sum_{k=k_1}^{k_2} \#\left([u+2^k, u+2^{k+1}) \cap I_{1,n}(\omega_1)\right) \cdot \max_{i\in[u+2^k,u+2^{k+1})} \mathbb{P}\left(i \in I_{2,n} \,|\, u \in I_{2,n}\right).\end{aligned}$$

Fix any $\epsilon \in (0,1)$. By (A.8) and Lemma A.1, there is $C > 0$ and an event B with probability higher than $1-\epsilon$ such that for all n large enough, all $w_1 \in B$, all $u \in I_{1,n,\gamma}(\omega_1) \cap nT$ and all $k \geq k_1$,

$$\#\left([u+2^k, u+2^{k+1}) \cap I_{1,n}(\omega_1)\right) \leqslant C \log n \frac{2^{\beta k}}{L(2^k)} . \tag{A.15}$$

Further, by (A.2),

$$\sup_{u\geq 0} \max_{i\in[u+2^k,u+2^{k+1})} \mathbb{P}\left(i \in I_{2,n} \mid u \in I_{2,n}\right) \lesssim \frac{2^{-(1-\beta)k}}{L(2^k)} . \tag{A.16}$$

Combining (A.15), (A.16), and Potter's bounds, we see that for any $w_1 \in B$

$$\max_{u\in I_{1,n,\gamma}(\omega_1)\cap nT} \overline{q}_n(u|\omega_1) \lesssim \log n \sum_{k=k_1}^{k_2} \frac{2^{-(1-2\beta)k}}{(L(2^k))^2} \lesssim (\log n)^{1+\alpha\gamma}$$

for any $\alpha > 2\beta - 1$. By the choice of γ, we can select α in such a way that $1 + \alpha\gamma < 0$. This proves (A.12).

Appendix B: Calculations for Sections 3 and 4

We start by checking that the lognormal-type tails of Example 3.1 satisfy Assumption 3.1. The fact that $(\nu(1,\infty))^{-1}\nu(\cdot \cap (1,\infty))$ is a subexponential distribution follows from Theorem 4.1.17 in Samorodnitsky (2016). Next, let

$$\overline{H_{\#}}(x) = c_1 x^{\beta} (\log x)^{\xi} \exp\left(-\lambda(\log x)^{\gamma}\right)$$

for $x > x_0$ that is large enough so that this function is decreasing and c_1 is such that $\overline{H_{\#}}(x_0) = 1$. That is, (3.4) holds with

$$h(x) = \left(\frac{\lambda\gamma(\log x)^{\gamma-1}}{x} - \frac{\xi}{x \log x} - \frac{\beta}{x}\right)^{-1} .$$

Regular variation of h is clear, and so are the eventual positivity of h' and the fact that $\lim_{x\to\infty} h'(x) = 0$. In particular, $H_{\#}$ is in the maximum domain of attraction of the Gumbel distribution. Next, by the implicit function theorem, G is, for large values of the argument, of the form (3.7). The relation $\overline{H_{\#}} \circ G(x) = x^{-1}$ for $x > 1/\overline{H_{\#}}(x_0) := x_1$ means, in this case, that

$$c_1 G(x)^{\beta} (\log G(x))^{\xi} \exp\left(-\lambda(\log G(x))^{\gamma}\right) = x^{-1} ,$$

so $\log G(x) \sim (\log x/\lambda)^{1/\gamma}$ as $x \to \infty$. Denoting $g(x) = G'(x)$ we also have, for $x > x_1$,

$$\beta \frac{g(x)}{G(x)} + \xi \frac{g(x)}{G(x)\log G(x)} - \lambda\gamma \frac{(\log G(x))^{\gamma-1} g(x)}{G(x)} = -\frac{1}{x} ,$$

so that, as $x \to \infty$,

$$\zeta(x) = \frac{g(x)}{G(x)} x \log x \sim \gamma^{-1} \lambda^{-1/\gamma} (\log x)^{1/\gamma} . \tag{B.1}$$

The assumptions $(B1)$-$(B4)$ follow from (B.1).

Next we check that the super-lognormal-type tails of Example 3.2 satisfy Assumption 3.1. Once again, the fact that $(\nu(1,\infty))^{-1}\nu(\cdot \cap (1,\infty))$ is a subexponential distribution follows from Theorem 4.1.17 in Samorodnitsky (2016). Now we set

$$\overline{H_{\#}}(x) = c_1 x^{\beta} (\log x)^{\xi} \exp\left(\lambda (\log x)^{\gamma}\right) \exp\left(-\rho \exp\left(\mu (\log x)^{\alpha}\right)\right)$$

for or $x > x_0$ and appropriate x_0, c_1, and (3.4) holds with

$$h(x) = \left(\frac{\rho\alpha\mu(\log x)^{\alpha-1} \exp\left(\mu(\log x)^{\alpha}\right)}{x} - \frac{\lambda\gamma(\log x)^{\gamma-1}}{x} - \frac{\xi}{x \log x} - \frac{\beta}{x} \right)^{-1} .$$

All of the arguments we used in the previous example still work. In this case we have

$$\exp\left(\mu(\log G(x))^{\alpha}\right) \sim \log x/\rho \quad \text{as } x \to \infty,$$

so also $\log G(x) \sim (\log\log x/\mu)^{1/\alpha}$ as $x \to \infty$. Since

$$\begin{aligned} \beta \frac{g(x)}{G(x)} + \xi \frac{g(x)}{G(x)\log G(x)} &+ \lambda\gamma \frac{(\log G(x))^{\gamma-1} g(x)}{G(x)} \\ -\rho\mu\alpha \exp\left(\mu(\log x)^{\alpha}\right) &\frac{(\log G(x))^{\alpha-1} g(x)}{G(x)} = -\frac{1}{x} , \end{aligned}$$

we conclude that

$$\zeta(x) = \frac{g(x)}{G(x)} x \log x \sim \alpha^{-1} \mu^{-1/\alpha} (\log\log x)^{(1-\alpha)/\alpha} \quad \text{as } x \to \infty. \tag{B.2}$$

As before, the assumptions $(B1)$-$(B4)$ follow from (B.2).

Proof of Proposition 3.1. (i) and (ii) follow by direct integration. To show (iii), we note that the derivative g of G satisfies $h \circ G(x) = xg(x)$ for all large x. Therefore, for large x,

$$\begin{aligned} \frac{h \circ G(H_1(x))}{h \circ G(H_2(x))} &= \frac{G(H_1(x))}{G(H_2(x))} \cdot \frac{\zeta(H_1(x))}{\zeta(H_2(x))} \cdot \frac{\log(H_2(x))}{\log(H_1(x))} \\ &\asymp \frac{G(H_1(x))}{G(H_2(x))} \cdot \frac{\zeta(H_1(x))}{\zeta(H_2(x))} \gg \exp\left\{ b(\log\log x)^{\delta} \right\} \end{aligned}$$

by Assumption 3.1 $(B2)$, Potter's bounds and direct integration.

For part (iv), we only consider the case $\alpha > 0$. When $\alpha < 0$, a similar argument works. Write

$$G\left(x(\log x)^{\alpha}\right) - G(x) = \int_1^{(\log x)^{\alpha}} \frac{G(ux)\zeta(ux)}{u\log(ux)}du\,.$$

Dividing this identity by $h \circ G(x) = G(x)\zeta(x)/\log x$ gives us

$$\frac{G\left(x(\log x)^{\alpha}\right) - G(x)}{h \circ G(x)} = \int_1^{(\log x)^{\alpha}} \frac{G(ux)}{G(x)} \cdot \frac{\zeta(ux)}{\zeta(x)} \cdot \frac{\log x}{\log(ux)} \cdot \frac{du}{u}\,. \tag{B.3}$$

Denote $I = [1, (\log x)^{\alpha}]$. Clearly, $\log x \sim \log(ux)$ uniformly over $u \in I$. Further, by Assumption 3.1 $(B1), (B3)$, we see that $\zeta(x) \asymp \zeta(ux)$ uniformly over $u \in I$. Finally, for $u \in I$, by Assumption 3.1 $(B2)$,

$$\begin{aligned} 1 \le \frac{G(ux)}{G(x)} &= \exp\left\{\int_x^{ux} \frac{\zeta(v)}{v\log v}dv\right\} \\ &\le \exp\left\{C\int_x^{ux} \frac{1}{v\log\log v}dv\right\} \\ &\le \exp\left\{C\int_x^{x(\log x)^{\alpha}} \frac{1}{v\log\log v}dv\right\} \to e^{\alpha C}\,, \end{aligned}$$

where C is a suitable constant. The claim now follows from (B.3) since

$$\int_1^{(\log x)^{\alpha}} \frac{du}{u} = \alpha\log\log x\,.$$

The argument for (v) is similar to that for (iv), We start with

$$\frac{G(x) - G\left(x2^{-j}\right)}{h \circ G(x)} = \int_{2^{-j}}^1 \frac{G(xu)}{G(x)} \cdot \frac{\zeta(xu)}{\zeta(x)} \cdot \frac{\log x}{\log(xu)} \cdot \frac{du}{u}. \tag{B.4}$$

Denoting now $I = \left[2^{-\rho\log x/\zeta(x)}, 1\right]$. Due to $\zeta(\cdot) \to \infty$, it is clear that $\log x \sim \log(xu)$ uniformly over $u \in I$. Furthermore, $x2^{-\rho\log x/\zeta(x)} \to \infty$, so by Assumption 3.1 $(B1)$, $(B2)$ and $(B3)$,

$$\frac{\zeta(xu)}{\zeta(x)} \gtrsim \frac{\zeta\left(x2^{-\rho\log x/\zeta(x)}\right)}{\zeta(x)} \gtrsim \frac{\zeta\left(x2^{-\rho\log x/(\log\log x)^{\delta}}\right)}{\zeta(x)} \gtrsim 1\,,$$

uniformly over $u \in I$. Finally, for $u \in I$, by Assumption 3.1 $(B1)$, $(B2)$, for some constant C,

$$\begin{aligned} \frac{G(ux)}{G(x)} \gtrsim \frac{G\left(x2^{-\rho\log x/\zeta(x)}\right)}{G(x)} &= \exp\left(-\int_{x2^{-\rho\log x/\zeta(x)}}^x \frac{\zeta(u)}{u\log u}du\right) \\ &\geqslant \exp\left(-C\zeta(x)\int_{x2^{-\rho\log x/\zeta(x)}}^x \frac{du}{u\log u}\right) > 2^{\rho C - 1} \end{aligned}$$

for all x large, uniformly over $u \in I$. Therefore, by (B.4) and Assumption 3.1 $(B2)$,

$$\frac{G(x) - G\left(x2^{-j}\right)}{jh \circ G(x)} \gtrsim 1\,,$$

as required. □

We finish by checking the claims made in Remark 4.1, and we start with the lognormal-type tails of Example 3.1. To see that (4.6) holds, it is enough to check that for any $C > 0$,

$$\lim_{n\to\infty} \frac{G\big(1/\overline{F}(n)\big)}{h \circ G(cw_n)} = 0\,. \tag{B.5}$$

The ratio above is asymptotic to

$$\frac{G\big(1/\overline{F}(n)\big)}{G(cw_n)} \frac{\log w_n}{\zeta(Cw_n)} = \exp\left(\int_{w_n}^{1/\bar{F}(n)} \frac{\zeta(u)}{u \log u} du\right) \frac{\log w_n}{\zeta(Cw_n)}\,, \tag{B.6}$$

which converges to 0 as $n \to \infty$ by (B.1). Next, for the super-lognormal-type tails of Example 3.2 with $0 < \alpha < 1/2$ one checks that (4.6) holds in the same way as above, by using (B.2) instead of (B.1). Finally, to see that (4.6) fails when $1/2 < \alpha < 1$, one needs to prove that, in this case, for any $C > 0$ the limit in (B.5) is infinity instead of 0. To do so one uses, once again, (B.6). It is routine to see that the expression there converges to infinity by using (B.2).

References

J. Aaronson (1997): *An Introduction to Infinite Ergodic Theory*, volume 50 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence.

M. T. Barlow and S. Taylor (1992): Defining Fractal Subsets of $\mathbb{Z}^d$. *Proceedings of the London Mathematical Society* :125–152.

V. Chistyakov (1964): A theorem on sums of independent random variables and its applications to branching random processes. *Theory of Probability and its Applications* 9:640–648.

L. de Haan and A. Ferreira (2006): *Extreme Value Theory: An Introduction*. Springer, New York.

P. Embrechts, C. Goldie and N. Veraverbeke (1979): Subexponentiality and infinite divisibility. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 49:335–347.

P. Embrechts, C. Klüppelberg and T. Mikosch (1997): *Modelling Extremal Events for Insurance and Finance*. Springer-Verlag, Berlin.

W. Feller (1966): *An Introduction to Probability Theory and its Applications*, volume 2. Wiley, New York, 1st edition.

C. Goldie and S. Resnick (1988): Distributions that are both subexponential and in the domain of attraction of an extreme-value distribution. *Advances in Applied Probability* 20:706–718.

T. Harris and H. Robbins (1953): Ergodic theory of Markov chains admitting an infinite invariant measure. *Proceedings of the National Academy of Sciences* 39:860–864.

C. Lacaux and G. Samorodnitsky (2016): Time-changed extremal process as a random sup measure. *Bernoulli* 22:1979–2000.

M. Leadbetter (1983): Extremes and local dependence of stationary sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 65:291–306.

I. Molchanov (2017): *Theory of Random Sets*. Springer, London, 2nd edition.

G. O'Brien, P. Torfs and W. Vervaat (1990): Stationary self-similar extremal processes. *Probability Theory and Related Fields* 87:97–119.

W. Pruitt and S. Taylor (1969): Sample path properties of processes with stable components. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 12:267–289.

S. Resnick (1987): *Extreme Values, Regular Variation and Point Processes*. Springer-Verlag, New York.

S. Resnick and M. Rubinovitch (1973): The structure of extremal processes. *Advances in Applied Probability* 5:287–307.

S. Resnick, G. Samorodnitsky and F. Xue (2000): Growth rates of sample covariances of stationary symmetric α-stable processes associated with null recurrent Markov chains. *Stochastic Processes and Their Applications* 85:321–339.

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

G. Samorodnitsky (2004): Extreme value theory, ergodic theory, and the boundary between short memory and long memory for stationary stable processes. *The Annals of Probability* 32:1438–1468.

G. Samorodnitsky (2016): *Stochastic Processes and Long Range Dependence*. Springer, Cham, Switzerland.

G. Samorodnitsky and Y. Wang (2019): Extremal theory for long range dependent infinitely divisible processes. *Annals of Prbability* 47:2529–2562.

K. Sato (2013): *Lévy Processes and Infinitely Divisible Distributions*. Cambridge University Press, Cambridge, 2nd edition.