A large deviations approach to limit theory for heavy-tailed time series

T Mikosch, O Wintenberger

A LARGE DEVIATIONS APPROACH TO LIMIT THEORY FOR HEAVY-TAILED TIME SERIES

T. MIKOSCH AND O. WINTENBERGER

ABSTRACT. In this paper we propagate a large deviations approach for proving limit theory for (generally) multivariate time series with heavy tails. We make this notion precise by introducing regularly varying time series. We provide general large deviation results for functionals acting on a sample path and vanishing in some neighborhood of the origin. We study a variety of such functionals, including large deviations of random walks, their suprema, the ruin functional, and further derive weak limit theory for maxima, point processes, cluster functionals and the tail empirical process. One of the main results of this paper concerns bounds for the ruin probability in various heavy-tailed models including GARCH, stochastic volatility models and solutions to stochastic recurrence equations.

1. PRELIMINARIES AND BASIC MOTIVATION

In the last decades, a lot of efforts has been put into the understanding of limit theory for dependent sequences, including Markov chains (Meyn and Tweedie [42]), weakly dependent sequences (Dedecker et al. [21]), long-range dependent sequences (Doukhan et al. [23], Samorodnitsky [54]), empirical processes (Dehling et al. [22]) and more general structures (Eberlein and Taqqu [25]), to name a few references. A smaller part of the theory was devoted to limit theory under extremal dependence for point processes, maxima, partial sums, tail empirical processes. Resnick [49, 50] started a systematic study of the relations between the convergence of point processes, sums and maxima; see also Resnick [51] for a recent account. He advocated the use of *multivariate regular variation* as a flexible tool to describe heavy-tail phenomena combined with advanced continuous mapping techniques. For example, maxima and sums are understood as functionals acting on an underlying point process; if the point process converges these functionals converge as well and their limits are described in terms of the points of the limiting point process.

Davis and Hsing [13] recognized the power of this approach for limit theory of point processes, maxima, sums, and large deviations for dependent *regularly varying* processes, i.e., stationary sequences whose finite-dimensional distributions are regularly varying with the same index. Before [13], limit theory for particular regularly varying stationary sequences was studied for the sample mean, maxima, sample autocovariance and autocorrelation functions of linear and bilinear processes with iid regularly varying noise and extreme value theory was considered for regularly varying ARCH processes and solutions to stochastic recurrence equation; see Rootzén [53], Davis and

Resnick [16, 17, 18, 19], de Haan et al. [30]. Davis and Hsing [13] introduced regular variation of a random sequence as a general flexible tool for proving limit theory for heavy-tail phenomena. The theory in [13] is a benchmark for the results of the present paper. We quote the main result of [13] on convergence of point processes for reasons of comparison (see Theorem 2.3 below); we will also give a short alternative proof in this paper. The main result of [13] has been used to derive the central limit theorem with infinite variance stable limit via the mapping theorem. When studying other functionals of the sample path than sums and maximas, this approach is limited by the continuity condition in the mapping theorem. For example, the asymptotics for the supremum of the partial sums follows only under additional restrictions from the point process approach; see Basrak et al. [8].

We introduce a new approach to bypass this restriction. We essentially follow an argument of Jakubowski [32, 33] and Jakubowski and Kobus [34], using a telescoping sum approach. Under suitable anti-clustering conditions, this argument can be applied to Laplace functionals, characteristic functions of sums, distribution functions of maxima, etc. This approach turned out to be fruitful in our previous work; see Bartkiewicz et al. [5], Mikosch and Wintenberger [44, 45]. A careful study of related work such as Davis and Hsing [13], Basrak and Segers [9], Segers [55], Balan and Louhichi [3, 4] and Yun [57], shows that the telescoping approach has been used in these papers as well. The aim of this paper is to understand the common structural properties of these results and their close relationship with large deviation theory.

The framework of this paper is the one of regularly varying stationary processes that we introduce now. We commence with a random vector X with values in $\mathbb{R}^d$ for some $d \geqslant 1$. We say that this vector (and its distribution) are *regularly varying with index* $\alpha > 0$ if the following relation holds as $x \to \infty$:

$$(1.1) \qquad \frac{\mathbb{P}(|X| > ux, X/|X| \in \cdot)}{\mathbb{P}(|X| > x)} \xrightarrow{w} u^{-\alpha} \mathbb{P}(\Theta \in \cdot), \quad u > 0\,.$$

Here $\xrightarrow{w}$ denotes weak convergence of finite measures and Θ is a vector with values in the unit sphere $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$ of $\mathbb{R}^d$. Its distribution is the *spectral measure* of regular variation and depends on the choice of the norm. However, the definition of regular variation does not depend on any concrete norm; for convenience we always refer to the Euclidean norm. An equivalent way to define regular variation of X is to require that there exists a non-null Radon measure μ_X on the Borel σ-field of $\overline{\mathbb{R}}_0^d = \overline{\mathbb{R}}^d \setminus \{\mathbf{0}\}$ such that

$$(1.2) \qquad n\,\mathbb{P}(a_n^{-1} X \in \cdot) \xrightarrow{v} \mu_X\,,$$

where the sequence (a_n) can be chosen such that $n\,\mathbb{P}(|X| > a_n) \sim 1$ and $\xrightarrow{v}$ refers to vague convergence. The limit measure μ_X necessarily has the property $\mu_X(u\cdot) = u^{-\alpha}\mu_X(\cdot)\,, u > 0$, which explains the relation with the index α. We refer to Bingham et al. [10] for an encyclopedic treatment of one-dimensional regular variation and Resnick [50, 51] for the multivariate case.

Next consider a strictly stationary sequence $(X_t)_{t \in \mathbb{Z}}$ of $\mathbb{R}^d$-valued random vectors with a generic element X. It is *regularly varying with index* $\alpha > 0$ if every lagged vector $(X_1, ..., X_k)$, $k \geqslant 1$, is regularly varying in the sense of (1.1); see Davis and Hsing [13]. An equivalent description of a regularly varying sequence (X_t) is achieved by exploiting (1.2): for every $k \geqslant 1$, there exists a non-null Radon measure μ_k on the Borel σ-field of $\overline{\mathbb{R}}_0^{dk}$ such that

$$(1.3) \qquad n\,\mathbb{P}(a_n^{-1}(X_1, \ldots, X_k) \in \cdot) \xrightarrow{v} \mu_k\,,$$

where (a_n) is chosen such that $n\,\mathbb{P}(|X_0| > a_n) \sim 1$.

A convenient characterization of a regularly varying sequence (X_t) was given in Theorem 2.1 of Basrak and Segers [9]: there exists a sequence of $\mathbb{R}^d$-valued random vectors $(Y_t)_{t \in \mathbb{Z}}$ such that

$\mathbb{P}(|Y_0| > y) = y^{-\alpha}$ for $y > 1$ and for $k \geqslant 0$,

$$\mathbb{P}(x^{-1}(X_{-k}, \ldots, X_k) \in \cdot \mid |X_0| > x) \overset{w}{\to} \mathbb{P}((Y_{-k}, \ldots, Y_k) \in \cdot), \quad x \to \infty.$$

The process (Y_t) is the *tail process* of (X_t). Writing $\Theta_t = Y_t/|Y_0|$ for $t \in \mathbb{Z}$, one also has for $k \geqslant 0$,

$$(1.4) \qquad \mathbb{P}(|X_0|^{-1}(X_{-k}, \ldots, X_k) \in \cdot \mid |X_0| > x) \overset{w}{\to} \mathbb{P}((\Theta_{-k}, \ldots, \Theta_k) \in \cdot), \quad x \to \infty.$$

We identify $|Y_0| (Y_t/|Y_0|)_{|t| \leqslant k} = |Y_0| (\Theta_t)_{|t| \leqslant k}$, $k \geqslant 0$. Then $|Y_0|$ is independent of $(\Theta_t)_{|t| \leqslant k}$ for every $k \geqslant 0$. We refer to $(\Theta_t)_{t \in \mathbb{Z}}$ as the *spectral tail process* of (X_t). In what follows, we will make heavy use of the tail and spectral tail processes: most of our results will be expressed in terms of them. We will refer to either condition (1.3) and the equivalent tail and spectral tail conditions as $(\mathbf{RV}_\alpha)$.

The condition $(\mathbf{RV}_\alpha)$ is equivalent to the fact that for any $\varepsilon > 0$, any continuous bounded function $f(x_0, x_1, \ldots)$ on $(\mathbb{R}^d)^{\mathbb{N}}$ which vanishes for $|x_0| \leqslant \varepsilon$ the following relation holds:

$$\frac{\mathbb{E}[f(x^{-1}(X_0, \ldots, X_k, 0, 0, \ldots))]}{\mathbb{P}(|X_0| > x)} \to \int_0^\infty \mathbb{E}[f(y\Theta_0, \ldots, y\Theta_k, 0, 0, \ldots)] \, d(-y^{-\alpha}), \qquad k \geqslant 0.$$

When considering the extremal properties of a sample path $(X_0, \ldots, X_n)$ for large n, it is not natural to assume that X_0 is large. To overcome this restriction a telescoping argument helps. It shows that for any $\varepsilon > 0$, any continuous bounded function f on $(\mathbb{R}^d)^{\mathbb{N}}$ which vanishes if $|x_t| \leqslant \varepsilon$ for all $t \geqslant 0$, the following relation holds:

$$\frac{\mathbb{E}[f(x^{-1}(X_0, \ldots, X_n, 0, 0, \ldots))]}{\mathbb{P}(|X_0| > x)} \to \sum_{j=0}^n \int_0^\infty \mathbb{E}[f(\underbrace{0, \cdots, 0}_{j}, y\Theta_0, \ldots, y\Theta_{n-j}, 0, 0, \ldots)$$

$$(1.5) \qquad\qquad - f(\underbrace{0, \cdots, 0}_{j+1}, y\Theta_1, \ldots, y\Theta_{n-j}, 0, 0, \ldots)] \, d(-y^{-\alpha}), \qquad n \geqslant 0.$$

The right-hand side is no longer well-behaved when $n \to \infty$. To study the asymptotic extremal properties of the sample path a Cèsaro argument is required; under suitable assumptions the right-hand side will converge after renormalization with $n + 1$. In addition, the mean ergodic theorem can be used if we assume conditions such as

$$f(\underbrace{0, \cdots, 0}_{j}, y\Theta_0, \ldots, y\Theta_{n-j}, 0, 0, \ldots) = f(y\Theta_0, \ldots, y\Theta_{n-j}, 0, 0, \ldots).$$

Such a relation is satisfied under our condition $(\mathbf{C}_\varepsilon)$; see Section 3. If the spectral tail process is also ergodic we obtain

$$\lim_{n \to \infty} \lim_{x \to \infty} \frac{\mathbb{E}[f(x^{-1}(X_0, \ldots, X_n, 0, 0, \ldots))]}{(n+1)\mathbb{P}(|X_0| > x)} = \int_0^\infty \mathbb{E}[f(y(\Theta_t)_{t \geqslant 0}) - f((y\Theta_t)_{t \geqslant 1})] \, d(-y^{-\alpha}).$$

(1.6)

Our large deviation approach can be seen as an equicontinuity argument applied to the intermediate result (1.5). Under suitable assumptions it is possible that (1.6) holds uniformly on a region Λ_n of x-values when $n \to \infty$. We will use an anti-clustering condition to enforce the equicontinuity and ergodicity of the spectral tail process. Thanks to this new approach we can characterize the large deviations for various functionals of the sample path. For example, we obtain the large deviations of the supremum of the partial sums and we derive the asymptotic behavior of the ruin probability.

This new approach describes the limiting behavior of extremes of regularly varying sequences in term of their spectral tail processes. We refer to calculations of the spectral tail processes for concrete examples such that certain Markov, stochastic volatility, GARCH(1, 1) processes, solutions to stochastic recurrence equations, max-stable processes, and other examples in Basrak et al. [9, 8], Mikosch and Wintenberger [45], Davis et al. [15]; see also Examples 4.12–4.14 below.

The paper is organized as follows. In Section 2 we provide some probabilistic tools used in the article and formulate the main result of Davis and Hsing [13]. In Section 3 we formulate our main result about the large deviations for functionals acting on the sample paths of a regularly varying sequence; see Theorem 3.1. In Section 4.2 we show two other main results of this paper. In Theorem 4.5 we give a uniform large deviation bound for the suprema of a random walk constructed from a regularly varying sequence. A modification of the proof of Theorem 4.5 is then used to give bounds for the tails of the ruin functional; see Theorem 4.9. We apply the latter result to solutions to stochastic recurrence equations, GARCH$(1,1)$ and stochastic volatility processes. In Section 4.3 we show how the large deviation approach helps to prove results for cluster functionals and in Section 4.4 we apply large deviations to get results for the tail empirical process of a regularly varying sequence. Finally, the proofs of the results are provided in Section 5.

2. Preliminaries

2.1. **Anti-clustering conditions.** Davis and Hsing [13] introduced a condition that avoids "long-range dependence" of high level exceedances of the process (X_t):

Anti-clustering condition (**AC**): Let $m = m_n \to \infty$ be an integer sequence such that $m_n = o(n)$ as $n \to \infty$ and (a_n) the normalizing sequence from (**RV**$_\alpha$). They assume that

$$(2.7) \qquad \lim_{k \to \infty} \limsup_{n \to \infty} \mathbb{P}\big(\widetilde{M}_{k,m_n} > \delta a_n \mid |X_0| > \delta a_n\big) = 0, \quad \delta > 0,$$

where

$$(2.8) \qquad M_{s,t} = \max_{s \leqslant i \leqslant t} |X_i|, \quad s \leqslant t, \qquad \text{and} \qquad \widetilde{M}_{s,t} = \max_{s \leqslant |i| \leqslant t} |X_i|, \quad s \leqslant t.$$

This condition assures that extremal clusters of (X_t) get separated from each other when time goes by, i.e., the influence of an extremal shock at some time does not last forever. Conditions of this type are common in the extreme value literature, e.g. the popular condition $D'(a_n)$; see Leadbetter et al. [39], cf. Section 4.4 in Embrechts et al. [26]. It is often easy to verify (**AC**) by checking

$$\lim_{k \to \infty} \limsup_{n \to \infty} \sum_{k \leqslant |t| \leqslant m_n} \mathbb{P}(|X_t| > \delta a_n \mid |X_0| > \delta a_n) = 0, \quad \delta > 0.$$

The following result can be found in Segers [55] and Basrak and Segers [9]; see also O'Brien [47].

Proposition 2.1. *Let (X_t) be a non-negative strictly stationary sequence which is regularly varying with index $\alpha > 0$ and satisfies* (**AC**). *Then the limit*

$$\lim_{n \to \infty} \mathbb{P}(M_{1,m_n} \leqslant \delta a_n \mid X_0 > \delta a_n) = \gamma = P\big(\sup_{t \geqslant 1} Y_t \leqslant 1\big) = P\big(\sup_{t \geqslant 1} Y_{-t} \leqslant 1\big),$$

exists for every $\delta > 0$, it is positive and γ is the extremal index of (X_t).

The extremal index of a real-valued stationary sequence is often interpreted as reciprocal of the expected cluster size of high level exceedances. This intuition can be made precise; see for example the monographs Leadbetter et al. [39] and Embrechts et al. [26], Section 8.1.

2.2. **Mixing conditions.** Davis and Hsing [13] assumed a mixing condition in terms of Laplace functionals of the point processes

$$(2.9) \qquad N_{nj} = \sum_{t=1}^{j} \varepsilon_{a_n^{-1} X_t}, \quad j = 1, \dots, n, \quad N_{nn} = N_n, \quad n \geqslant 1,$$

with state space $\overline{\mathbb{R}}_0^d = \overline{\mathbb{R}}^d \setminus \{0\}$, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$. This condition reads as follows:

Condition $\mathcal{A}(a_n)$: For the same sequence (m_n) as in **(AC)** and with $k_n = [n/m] \to \infty$,

$$\mathbb{E}\mathrm{e}^{-\int f dN_n} - \left(\mathbb{E}\mathrm{e}^{-\int f dN_{n,m_n}}\right)^{k_n} \to 0, \quad f \in \mathbb{C}_K^+,$$

where $\mathbb{C}_K^+$ is the set of non-negative continuous functions with compact support.

Boundedness of a subset of $\overline{\mathbb{R}}_0^d$ means that it is bounded away from zero, in particular, compact sets in $\overline{\mathbb{R}}_0^d$ are bounded away from zero. Davis and Hsing [13] assumed a slightly more general version of $\mathcal{A}(a_n)$: their functions f are any non-negative step functions on $\overline{\mathbb{R}}_0^d$ with bounded support. For both classes of functions, the convergence of the Laplace functionals $\mathbb{E}\mathrm{e}^{-\int f dQ_n} \to \mathbb{E}\mathrm{e}^{-\int f dQ}$ for point processes $(Q_n), Q$ is equivalent to $Q_n \xrightarrow{d} Q$; see Kallenberg [35]. The restriction to $f \in \mathbb{C}_K^+$ is common in the literature, e.g. Resnick [49, 50, 51]. Various papers which build on Davis and Hsing [13] also assume $\mathcal{A}(a_n)$ for $f \in \mathbb{C}_K^+$; see Basrak and Segers [9], Balan and Louhichi [3]. Condition $\mathcal{A}(a_n)$ for suitable sequences (m_n) follows from both strong mixing and weak dependence in the sense of Dedecker and Doukhan [20].

Remark 2.2. Let $B_\delta^c = \{x \in \overline{\mathbb{R}}_0^d : |x| > \delta\}$, $\delta > 0$. Under regular variation of (X_t), $P(N_{nm}(B_\delta^c) > 0) \leqslant m_n \mathbb{P}(|X_1| > \delta a_n) \to 0$ and therefore an iid sequence $(\widetilde{N}_{nm}^{(i)})$ of copies of N_{nm} is a *null-array* in the sense of Kallenberg [35]. Then, according to Theorem 6.1 in [35], the sequence of point processes $\widetilde{N}_n = \sum_{i=1}^{k_n} \widetilde{N}_{nm}^{(i)}$ is relatively compact and the subsequential limits are infinitely divisible, possibly null. By virtue of $\mathcal{A}(a_n)$, $N_n \xrightarrow{d} N$ for some infinitely divisible point process N if and only if $\widetilde{N}_n \xrightarrow{d} N$.

2.3. **Weak convergence of point processes.** Now we are in the position to formulate one of the main results in Davis and Hsing [13]. The result was proved in the case $d = 1$ but immediately translates to the case $d > 1$; see Davis and Mikosch [14].

Theorem 2.3. *Assume that the strictly stationary $\mathbb{R}^d$-valued sequence (X_t) satisfies*
1. *the regular variation condition* $(\mathbf{RV}_\alpha)$ *for some $\alpha > 0$,*
2. *the anti-clustering condition* $(\mathbf{AC})$,
3. *the mixing condition* $\mathcal{A}(a_n)$.

Then $N_n = \sum_{t=1}^n \varepsilon_{a_n^{-1} X_t} \xrightarrow{d} N$ in the space of point measures on $\overline{\mathbb{R}}_0^d$ equipped with the vague topology and the infinitely divisible limiting point process N has representation

$$N = \sum_{i=1}^\infty \sum_{j=1}^\infty \varepsilon_{\Gamma_i^{-1/\alpha} Q_{ij}},$$

(Γ_i) is an increasing enumeration of a homogeneous Poisson process on $(0, \infty)$ with intensity γ, $(Q_{ij})_{j \geqslant 1}$, $i = 1, 2, \ldots$, are iid sequences of points Q_{ij} such that $\max_{j \geqslant 1} |Q_{ij}| = 1$ a.s. and γ is the extremal index of the sequence $(|X_t|)$. The distribution of $\sum_{j=1}^\infty \varepsilon_{Q_{1j}}$ is given in Theorem 2.7 of Davis and Hsing [13] in terms of the limit measures μ_k, $k \geqslant 1$; see (1.3).

Remark 2.4. Basrak and Segers [9] gave an alternative proof of this result. They also showed that N has Laplace functional in terms of the spectral tail chain (Θ_t) given by

$$\mathbb{E}\mathrm{e}^{-\int f dN} = \exp\left\{-\int_0^\infty \mathbb{E}\left(\mathrm{e}^{-\sum_{i=1}^\infty f(y\Theta_i)} - \mathrm{e}^{-\sum_{i=0}^\infty f(y\Theta_i)}\right) d(-y^{-\alpha})\right\}, \quad f \in \mathbb{C}_K^+.$$

Moreover, they showed that the extremal index γ is positive and has representation

$$(2.10) \qquad \gamma = \mathbb{E}\left(\sup_{t \geqslant 0} |\Theta_t|^\alpha - \sup_{t \geqslant 1} |\Theta_t|^\alpha\right).$$

3. A LARGE DEVIATIONS APPROACH TO LIMIT THEORY FOR HEAVY-TAILED TIME SERIES

We provide a general large deviation result for functionals acting on a regularly varying sequence and vanishing in some neighborhood of the origin. The latter property means that only large values of the sequence make a contribution to the limiting quantities.

We consider a complex-valued function f on $(\mathbb{R}^d)^{\mathbb{N}}$ an denote its restrictions to $(\mathbb{R}^d)^l$ by f_l, $f_0 = 0$. We will say that f (or equivalently (f_l)) satisfies $(\mathbf{C}_\varepsilon)$:

$(\mathbf{C}_\varepsilon)$ For $l \geqslant 1$, the a.e. continuous function f_l is bounded uniformly and

$$f_l(x_1, \ldots, x_{j_1-1}, x_{j_1}, \ldots, x_{j_2}, x_{j_2+1}, \ldots, x_l) = f_{j_2-j_1+1}(x_{j_1}, \ldots, x_{j_2}), \qquad 1 \leqslant j_1 \leqslant j_2 \leqslant l,$$

provided $|x_i| \leqslant \varepsilon$, $i = 1, \ldots, j_1 - 1, j_2 + 1, \ldots, l$.

We suppress the dependence on ε in the notation and we often write f instead of f_l; it will be clear from the number of arguments which f_l we are dealing with.

For a stationary sequence (X_t) we follow an approach which was advocated by Jakubowski and Kobus [34] and Jakubowski [32, 33] in the context of α-stable limit theory for sums of infinite variance random variables and was exploited in Bartkiewicz et al. [5], Balan and Louhichi [3], Mikosch and Wintenberger [44, 45] for proving limit theory for point processes, sums of regularly varying sequences with infinite variance stable limit laws, large deviation probabilities and other results. The main idea of this approach is to use suitable telescoping sums involving the differences $\mathbb{E}\big[f(x^{-1}X_0, \ldots, x^{-1}X_k) - f(x^{-1}X_1, \ldots, x^{-1}X_k)\big]$, $k \geqslant 1$.

Theorem 3.1. *Consider a strictly stationary $\mathbb{R}^d$-valued sequence (X_t) satisfying*

    1. *$(\mathbf{RV}_\alpha)$ for some $\alpha > 0$,*
    2. *the uniform anti-clustering condition*

$$(3.11) \qquad \lim_{k \to \infty} \limsup_{n \to \infty} \sup_{x \in \Lambda_n} \mathbb{P}(M_{k,n} > x\delta \mid |X_0| > x\delta) = 0, \quad \delta > 0,$$

for some sequence of Borel subsets $\Lambda_n \subset (0, \infty)$ such that $x_n = \inf \Lambda_n \to \infty$ and $n\mathbb{P}(|X_0| > x_n) \to 0$.

Let f be a complex-valued function satisfying $(\mathbf{C}_\varepsilon)$ for some $\varepsilon > 0$. Then the following relation holds:

$$\sup_{x \in \Lambda_n} \left| \frac{\mathbb{E}[f(x^{-1}X_1, \ldots, x^{-1}X_n)]}{n\mathbb{P}(|X_0| > x)} - \int_0^\infty \mathbb{E}[f(y(\Theta_t)_{t \geqslant 0}) - f(y(\Theta_t)_{t \geqslant 1})]d(-y^{-\alpha}) \right| \to 0.$$

(3.12)

The proof is given in Section 5.1.

Remark 3.2. If we consider one-point sets Λ_n the anti-clustering condition (3.11) can be compared with the anti-clustering condition $(\mathbf{AC})$ of Davis and Hsing [13]; see (2.7). Indeed, if we choose $x_n = \delta a_n$ for $\delta > 0$ and replace $M_{k,n}$ by $\widetilde{M}_{k,m_n}$ for some sequence $m_n \to \infty$, $m_n = o(n)$, then (3.11) is related to $(\mathbf{AC})$. However, there is one major difference: (3.11) does not involve maxima over sets of negative integers. If the stronger condition $(\mathbf{AC})$ holds, Basrak and Segers [9] showed that the extremal index $\gamma_{|X|}$ of the sequence $(|X_t|)$ is positive. It is also the case under the less restrictive condition (3.11) because $Y_t \overset{P}{\to} 0$ when $t \to \infty$ a.s. from the proof of Proposition 4.2 of [9] and $\gamma_{|X|} = \mathbb{P}(|Y_t| \leqslant 1, t > 0)$.

Recall the notion of a k-dependent stationary sequence (X_t) for some integer $k \geqslant 0$, i.e., the σ-fields $\sigma(X_t, t \leqslant 0)$ and $\sigma(X_t, t \geqslant k+1)$ are independent. In this case, Theorem 3.1 simplifies.

Theorem 3.3. *Consider a stationary $\mathbb{R}^d$-valued k-dependent sequence (X_t) satisfying $(\mathbf{RV}_\alpha)$ for some $\alpha > 0$ and a complex-valued function f satisfying $(\mathbf{C}_\varepsilon)$ for some $\varepsilon > 0$. Let (x_n) be a real-valued sequence such that $n\,\mathbb{P}(|X_0| > x_n) \to 0$. Then, as $n \to \infty$,*

$$\sup_{x \geqslant x_n} \left| \frac{\mathbb{E}[f_n(x^{-1}X_1, \ldots, x^{-1}X_n)]}{n\mathbb{P}(|X_0| > x)} - \lambda_k \int_0^\infty \mathbb{E}[f_{k+1}(y\widetilde{\Theta}_0, \ldots, y\widetilde{\Theta}_k)]d(-y^{-\alpha}) \right| \to 0\,.$$

(3.13)

with $\lambda_k = \mathbb{P}(\Theta_{-j} = 0, j = 1, \ldots, k) > 0$ and

(3.14) $$\mathbb{P}\big((\widetilde{\Theta}_j)_{j=0,\ldots,k} \in \cdot\big) = \mathbb{P}\big((\Theta_j)_{j=0,\ldots,k} \in \cdot \mid \Theta_{-l} = 0, l = 1, \ldots, k\big)\,.$$

The proof is given in Section 5.2.

Remark 3.4. The limiting expression in (3.12) does in general not coincide with the "naive" limit by letting $k \to \infty$ in (3.13), i.e.,

$$\mathbb{P}(\Theta_{-j} = 0, j \geqslant 1) \int_0^\infty \mathbb{E}[f(y(\tilde{\Theta}_t)_{t \geqslant 0})]d(-y^{-\alpha})$$

because $\mathbb{P}(\Theta_{-j} = 0, j \geqslant 1) = 0$ is possible. However, if $\mathbb{P}(|\Theta_{-l}| \leqslant \delta, l \geqslant 1) > 0$ for some $\delta > 0$ then we can define $(\widetilde{\Theta}_t^\delta)$ through the relation

$$\mathbb{P}\big((\widetilde{\Theta}_t^\delta)_{t \geqslant 0} \in \cdot\big) = \mathbb{P}\big((\Theta_t)_{t \geqslant 0} \in \cdot \mid |\Theta_{-l}| \leqslant \delta, l \geqslant 1\big)\,, \qquad \varepsilon \geqslant \delta > 0\,.$$

In view of condition $(\mathbf{C}_\varepsilon)$ and the proof of Theorem 3.3, we also have

$$\int_0^\infty \mathbb{E}[f(y(\Theta_t)_{t \geqslant 0}) - f(y(\Theta_t)_{t \geqslant 1})]d(-y^{-\alpha})$$

$$= \int_\varepsilon^\infty \mathbb{E}[f(y(\Theta_t)_{t \geqslant 0}) - f(y(\Theta_t)_{t \geqslant 1})]d(-y^{-\alpha})$$

$$= \int_\varepsilon^\infty \mathbb{E}f(y(\Theta_t)_{t \geqslant 0})d(-y^{-\alpha}) - \int_\varepsilon^\infty \mathbb{E}f(y(\Theta_t)_{t \geqslant 1})]d(-y^{-\alpha})$$

$$= \mathbb{P}(|\Theta_{-l}| \leqslant \delta, l \geqslant 1) \int_\varepsilon^\infty \mathbb{E}f(y(\tilde{\Theta}_t^\delta)_{t \geqslant 0})d(-y^{-\alpha}),$$

and both quantities on the right-hand side in the third line are finite and involve only a finite number of Θ_t a.s. because $\Theta_t \xrightarrow{P} 0$.

4. Applications

In this section we will provide various applications of Theorems 3.1 and 3.3 to limit theorems of large deviation-type and weak convergence results of various kinds.

4.1. **Limit theory for point processes and partial sums.**

4.1.1. *Weak convergence of point processes.* We re-prove the point process result of Davis and Hsing [13] on point process convergence given as Theorem 2.3 above, formulated in the language of Basrak and Segers [9]; see Remark 2.4. A careful analysis of [9] shows that their proofs use ideas which are close to those in the proof of Theorem 3.1; see also Balan and Louhichi [3] who apply Jakubowski's ideas to point process convergence of more general triangular arrays. Recall the definition of the point processes N_n from (2.9).

Theorem 4.1. *Assume that the strictly stationary $\mathbb{R}^d$-valued sequence (X_t) satisfies*

1. $(\mathbf{RV}_\alpha)$ *for some $\alpha > 0$,*
2. *the mixing condition $\mathcal{A}(a_n)$,*

3. *the anti-clustering condition* (3.11) *for* $\Lambda_n = \{a_n\}$.

Then $N_n \overset{d}{\to} N$, *where* (N_n) *is defined in* (2.9), *and* N *has Laplace functional*

$$\mathbb{E}\mathrm{e}^{-\int_{\mathbb{R}_0^d} g\,dN} = \exp\left(-\int_0^\infty \mathbb{E}(\mathrm{e}^{-\sum_{j=1}^\infty g(y\Theta_j)} - \mathrm{e}^{-\sum_{j=0}^\infty g(y\Theta_j)})\,d(-y^{-\alpha})\right), \quad g \in \mathbb{C}_K^+.$$

Proof. We prove the convergence of the logarithms of the Laplace functionals $\log \mathbb{E}\mathrm{e}^{-\int g\,dN_n} \to \log \mathbb{E}\mathrm{e}^{-\int g\,dN}$ for $g \in \mathbb{C}_K^+$. We assume that $g(x) = 0$ for $|x| \leqslant \varepsilon$ for some $\varepsilon > 0$. In view of the mixing condition we have for some $m = m_n \to \infty$ and $k_n = [n/m] \to \infty$,

$$\log \mathbb{E}\mathrm{e}^{-\int g\,dN} = \log \mathbb{E}\mathrm{e}^{-\sum_{t=1}^n g(a_n^{-1}X_t)} \sim k_n \log \mathbb{E}\mathrm{e}^{-\sum_{t=1}^{m_n} g(a_n^{-1}X_t)}.$$

By a Taylor expansion, since $\mathbb{E}\sum_{t=1}^{m_n} g(a_n^{-1}X_t) \leqslant Cm_n\mathbb{P}(|X| > \varepsilon a_n) \to 0$,

$$-k_n \log \mathbb{E}\exp(-\sum_{t=1}^{m_n} g(a_n^{-1}X_t)) \sim k_n\mathbb{E}(f_{m_n}(a_n^{-1}X_1, \ldots, a_n^{-1}X_{m_n})),$$

where

$$f_l(x_1, \ldots, x_l) = 1 - \exp\left(-\sum_{t=1}^l g(x_t)\right), \quad 1 \leqslant l.$$

Notice that (f_l) satisfies $(\mathbf{C}_\varepsilon)$. Now an application of Theorem 3.1 with $\Lambda = \{a_n\}$ and n replaced by m_n yields

$$k_n\,\mathbb{E}(f(a_n^{-1}X_1, \ldots, a_n^{-1}X_{m_n})) \to \int_0^\infty \left[\mathbb{E}[f(y(\Theta_i)_{i\geqslant 0}) - f(y(\Theta_i)_{i\geqslant 1})]\right]d(-y^{-\alpha}).$$

The limit is the desired logarithm of the Laplace functional N. Combining the arguments, we proved the corollary. $\qquad\square$

Remark 4.2. For a k-dependent regularly varying sequence (X_t), the mixing and anti-clustering conditions of Theorem 4.1 are trivialy satisfied. Moreover, we conclude from Theorem 3.3 that N has Laplace functional

$$E\mathrm{e}^{-g\,dN} = \exp\left(-\lambda_k \int_0^\infty \mathbb{E}(1 - \mathrm{e}^{-\sum_{j=0}^k g(y\widetilde{\Theta}_j)})\,d(-y^{-\alpha})\right), \quad g \in \mathbb{C}_K^+.$$

Calculation shows that the infinitely divisible limiting point process N has representation

$$N = \sum_{i=1}^\infty \sum_{j=0}^k \varepsilon_{\Gamma_i^{-1/\alpha}\widetilde{\Theta}_{ij}},$$

where (Γ_i) is an increasing enumeration of a homogeneous Poisson process on $(0, \infty)$ with intensity λ_k, independent of an iid sequence $(\widetilde{\Theta}_{ij})_{0\leqslant j\leqslant k}$, $i = 1, 2, \ldots$, with generic element $(\widetilde{\Theta}_j)_{0\leqslant j\leqslant k}$.

4.1.2. *The α-stable central limit theorem.* In this section we consider the truncated random variables

$$\overline{X}_t = X_t\mathbb{1}_{|X_t|\leqslant\varepsilon a_n}, \quad \underline{X}_t = X_t - \overline{X}_t, \quad t \in \mathbb{Z},$$

and the corresponding partial sums $\overline{S}_n$ and $\underline{S}_n$, where we suppress the dependence on $\varepsilon > 0$ and n in the notation.

Theorem 4.3. *Consider a stationary* $\mathbb{R}^d$*-valued sequence* (X_t) *satisfying*

1. $(\mathbf{RV}_\alpha)$ *for some* $\alpha \in (0, 2)\backslash\{1\}$,
2. *the mixing condition*

$$(4.1) \qquad\qquad \mathbb{E}\mathrm{e}^{is'\underline{S}_n/a_n} - \left(\mathbb{E}\mathrm{e}^{is'\underline{S}_m/a_n}\right)^{k_n} \to 0, \quad s \in \mathbb{R}^d, \quad n \to \infty,$$

3. *the anti-clustering condition* (3.11) *for* $\Lambda_n = \{a_n\}$,
4. *for* $\alpha \in (1,2)$, *in addition,* $\mathbb{E}X = 0$, *the vanishing-small-values condition*

$$(4.2) \qquad \lim_{\varepsilon \downarrow 0} \limsup_{n \to \infty} \mathbb{P}(a_n^{-1}|\overline{S}_n - \mathbb{E}\overline{S}_n| > \delta) = 0, \quad \delta > 0,$$

and $\sum_{j=1}^{\infty} \mathbb{E}|\Theta_j| < \infty$.

Then $a_n^{-1}S_n \xrightarrow{d} \xi_\alpha$, *where the limit is an α-stable random variable with log-characteristic function*

$$(4.3) \qquad \int_0^\infty \mathbb{E}\Big(e^{iys'\sum_{j=0}^\infty \Theta_j} - e^{iys'\sum_{j=1}^\infty \Theta_j} - iys'A \Big) d(-y^{-\alpha}), \qquad s \in \mathbb{R}^d,$$

where $A = 0$ *for* $\alpha \in (0,1)$ *and* $A = \Theta_0$ *for* $\alpha \in (1,2)$.

The proof is given in Section 5.3.

Remark 4.4. Condition (4.2) for $\alpha \in (1,2)$ is standard in central limit theory; see the discussions in Davis and Hsing [13], Bartkiewicz et al. [5], Basrak et al. [8]. Sufficient conditions are k-dependence of (X_t) and conditional independence. For concrete models such as stochastic volatility models, GARCH and certain Markov chains, see the references above and [44, 45]. For similar characterizations of the α-stable limiting laws as in (4.3), see Mirek [46]. It coincides with the limit law given in [5] as shown by the computations of Louhichi and Rio [40].

In the k-dependent case, Jakubowski and Kobus [34] and Kobus [36] got related α-stable limit theory under the assumption that $(X_0, \dots, X_k)$ is regularly varying with index α. In view of Proposition 5.1, the latter condition is equivalent to condition $(\mathbf{RV}_\alpha)$. Extensions of the α-stable central limit theorem to the stationary case were considered in Jakubowski [32, 33].

4.2. **Large deviations for suprema of a random walk and ruin bounds.**

4.2.1. *Large deviations for the supremum of a random walk.* In this section we derive a result for the suprema of a univariate random walk (S_n). We write for any $x, \varepsilon > 0$,

$$\overline{X}_t = X_t \mathbb{1}_{|X_t| \leqslant \varepsilon x}, \quad \underline{X}_t = X_t - \overline{X}_t, \quad t \in \mathbb{Z},$$

and

$$S_0 = 0, \quad \overline{S}_t = \sum_{i=1}^t \overline{X}_i, \quad \underline{S}_t = S_t - \overline{S}_t, \quad t \in \mathbb{Z}.$$

Here we suppress the dependence of these quantities on x, ε in the notation.

Theorem 4.5. *Consider a stationary $\mathbb{R}$-valued sequence (X_t) satisfying the following conditions*
1. $(\mathbf{RV}_\alpha)$ *for some $\alpha > 0$,*
2. *the anti-clustering condition* (3.11).

If $\alpha > 1$ we also assume
3. *the vanishing-small-values condition*

$$(4.4) \qquad \lim_{\varepsilon \downarrow 0} \limsup_{n \to \infty} \sup_{x \in \Lambda_n} \frac{\mathbb{P}(x^{-1}\sup_{t \leqslant n}|\overline{S}_t| > \delta)}{n\,\mathbb{P}(|X| > x)} = 0, \quad \delta > 0,$$

4. $\mathbb{E}\Big(\sum_{i=1}^\infty |\Theta_i| \Big)^{\alpha-1} < \infty$.

Then

$$(4.5) \quad \sup_{x \in \Lambda_n} \Big| \frac{\mathbb{P}\Big(\sup_{i \leqslant n} S_i > x \Big)}{n\,\mathbb{P}(|X| > x)} - \mathbb{E}\Big[\Big(\Theta_0 + \sup_{t \geqslant 1} \sum_{i=1}^t \Theta_i \Big)_+^\alpha - \Big(\sup_{t \geqslant 1} \sum_{i=1}^t \Theta_i \Big)_+^\alpha \Big] \Big| \to 0, \quad n \to \infty.$$

The proof is given in Section 5.4.

Remark 4.6. We observe that the limit in (4.5) can be 0, for example for the large deviations of the telescoping sum S_n with $X_t = Y_t - Y_{t-1}$ with $Y_t \geqslant 0$ iid regularly varying. Then $\Theta_0 = -\Theta_1 = 1$, $\Theta_t = 0$ for $t \geqslant 2$ and the limiting constant is zero.

Condition (4.4) can often be verified by using maximal inequalities for sums, for example in the case of regenerative Markov chains or conditionally independent random variables; see for example [45]. For a k-dependent sequence one can verify this condition as well (see Lemma 5.2), resulting in the following corollary.

Corollary 4.7. *Assume that (X_t) is a k-dependent univariate strictly stationary sequence which is also regularly varying with index $\alpha > 0$. In addition, we assume the following conditions:*

1. $\mathbb{E}X = 0$ *if* $\mathbb{E}|X| < \infty$.
2. *If* $\alpha = 1$ *and* $\mathbb{E}|X| = \infty$, *we have*

$$(4.6) \qquad \limsup_{n \to \infty} \sup_{x \geqslant x_n} n x^{-1} |\mathbb{E}\overline{X}| = 0 \, .$$

Let (a_n) be any sequence such that $n\,\mathbb{P}(|X| > a_n) \to 1$ as $n \to \infty$. Then

$$\sup_{x \geqslant x_n} \Big| \frac{\mathbb{P}\Big(\sup_{t \leqslant n} S_t > x \Big)}{n\mathbb{P}(|X| > x)} - \lambda_k \mathbb{E}\Big(\sup_{t \leqslant k} \sum_{i=0}^{t} \widetilde{\Theta}_i \Big)_+^{\alpha} \Big| \to 0,$$

where $x_n \to \infty$ is any sequence such that $x_n/a_n \to \infty$ if $\alpha < 2$, $x_n/n^{0.5+\delta} \to \infty$ for some $\delta > 0$ if $\alpha = 2$ and $\mathbb{E}X^2 = \infty$, and $x_n \geqslant C\sqrt{n \log n}$ for sufficiently large $C > 0$ if $\mathbb{E}X^2 < \infty$.

The proof is given in Section 5.5.

Remark 4.8. The proof of Corollary 4.7 immediately extends to certain subadditive functionals acting on the random walk (S_n) which are more general than suprema. Indeed, let g_l be a sequence of real-valued functions on $\mathbb{R}^l$, $l \geqslant 1$. Assume that, for any $l \geqslant 1$,

- g_l is *continuous and positively homogeneous*, i.e., $g_l(c\mathbf{x}) = c g_l(\mathbf{x})$ for any $\mathbf{x} \in \mathbb{R}^l$ and $c > 0$,
- *subadditive*, i.e., $g_l(\mathbf{x} + \mathbf{y}) \leqslant g_l(\mathbf{x}) + g_l(\mathbf{y})$, $\quad \mathbf{x}, \mathbf{y} \in \mathbb{R}^l$,
- a *domination property* holds: for $s_t = x_1 + \cdots + x_t$, $\mathbf{s}_l = (s_1, \ldots, s_l)$, there exists a constant $C > 0$ not depending on l such that $|g_l(\mathbf{s}_l)| \leqslant C \sup_{1 \leqslant i \leqslant l} |s_i|$.

Finally, write $\underline{s}_t = \sum_{i=1}^{t} x_i \mathbb{1}_{|x_i| > \varepsilon}$ and assume that $(\mathbb{1}_{g_l(\underline{s}_1, \ldots, \underline{s}_l) > 1})$ satisfies $(\mathbf{C}_\varepsilon)$. Then, under the conditions and with the notation of Corollary 4.7, the following result holds:

$$\sup_{x \geqslant x_n} \Big| \frac{\mathbb{P}\Big(g_n\big((S_t)_{t=1,\ldots,n} \big) > x \Big)}{n\mathbb{P}(|X| > x)} - \lambda_k \mathbb{E}\Big[\Big(g_{k+1}\big((\sum_{i=0}^{t} \widetilde{\Theta}_i)_{t=0,\ldots,k} \big) \Big)_+^{\alpha} \Big] \Big| \to 0 \, .$$

For example, with $g_l(s_1, \ldots, s_l) = \sup_{t \leqslant l} |s_t|$, $l \geqslant 1$, we obtain

$$\sup_{x \geqslant x_n} \Big| \frac{\mathbb{P}\Big(\sup_{t \leqslant n} |S_t| > x \Big)}{n\mathbb{P}(|X| > x)} - \lambda_k \mathbb{E} \sup_{t \leqslant k} \Big| \sum_{i=0}^{t} \widetilde{\Theta}_i \Big|^{\alpha} \Big| \to 0 \, .$$

With $g_l(s_1, \ldots, s_l) = s_l$ we get a large deviation result for sums:

$$\sup_{x \geqslant x_n} \Big| \frac{\mathbb{P}\Big(S_n > x \Big)}{n\mathbb{P}(|X| > x)} - \lambda_k \mathbb{E}\Big(\sum_{i=0}^{k} \widetilde{\Theta}_i \Big)_+^{\alpha} \Big| \to 0 \, .$$

Other functionals of this kind are given by $g_l(s_1, \ldots, s_l) = \max_{i=1,\ldots,l}(s_i - s_l)_+$ and $g_l(s_1, \ldots, s_l) = \max_{i=1,\ldots,l} s_i - \min_{i=1,\ldots,l} s_i$, $g_l(s_1, \ldots, s_l) = \max_{i,j=1,\ldots,l}(s_i - s_j)$.

For regularly varying moving averages with index $\alpha < 2$, Basrak and Krizmanič [7] studied the M_2-functional convergence of the partial sums when $\sum_{i=0}^{k} \widetilde{\Theta}_i$ coincides with $\sup_{0 \leqslant t \leqslant k} \sum_{i=0}^{t} \widetilde{\Theta}_i$. They derived the limiting law of the supremum of the partial sums; it is the supremum of the α-stable Lévy process $(\xi_t)_{t \in [0,1]}$ where ξ_1 has the same distribution as the limiting law of the partial sums. In view of the results above, a similar phenomenon can be observed under the condition

$$(4.7) \qquad \mathbb{E}\Big(\sum_{i=0}^{k} \widetilde{\Theta}_i \Big)_+^{\alpha} = \mathbb{E}\Big(\sup_{0 \leqslant t \leqslant k} \sum_{i=0}^{t} \widetilde{\Theta}_i \Big)_+^{\alpha}$$

for the large deviations of sums and their suprema:

$$\sup_{x \geqslant x_n} \Big| \frac{\mathbb{P}(S_n > x)}{n\mathbb{P}(|X| > x)} - \frac{\mathbb{P}(\sup_{1 \leqslant t \leqslant n} S_t > x)}{n\mathbb{P}(|X| > x)} \Big| \to 0 \,.$$

However, in the cases where (4.7) does not hold, the functional central limit theorem cannot hold for any topology for which the supremum is a continuous function.

4.2.2. *Ruin probabilities.* In this section we assume that (X_t) is a univariate strictly stationary sequence which is also regularly varying with index $\alpha > 1$. The latter condition ensures that $\mathbb{E}|X| < \infty$. We will also assume that $\mathbb{E}X = 0$. In what follows, we will study the asymptotic behavior of the tail probability $\mathbb{P}(\sup_{t \geqslant 0}(S_t - \rho t) > x)$ for $\rho > 0$ as $x \to \infty$. We will refer to this probability as *ruin probability* since similar expressions appear in the context of non-life insurance mathematics; see Asmussen and Albrecher [1] and Embrechts et al. [26], Chapter 1. We use the notation of Section 4.2.1.

Theorem 4.9. *Assume that (X_t) is a univariate strictly stationary sequence which is also regularly varying with index $\alpha > 1$, has mean zero and satisfies the conditions of Theorem 4.5 with $\Lambda_n = [C_1 n, C_2 n]$ for any possible choice of positive constants $C_1 < C_2$. Then we have for any $\rho > 0$,*

$$\frac{\mathbb{P}(\sup_{t \geqslant 0}(S_t - \rho t) > x)}{x\, \mathbb{P}(|X| > x)} \sim \frac{\mathbb{E}\Big[\big(\sup_{t \geqslant 0} \sum_{i=0}^{t} \Theta_i \big)_+^{\alpha} - \big(\sup_{t \geqslant 1} \sum_{i=1}^{t} \Theta_i \big)_+^{\alpha} \Big]}{(\alpha - 1)\rho}, \quad x \to \infty \,.$$

(4.8)

The proof is given in Section 5.6.

Remark 4.10. Notice that the right-hand side of (4.8) is of the form

$$\frac{\mathbb{E}[(1 + \sum_{i=1}^{\infty} \Theta_i)^{\alpha} - (\sum_{i=1}^{\infty} \Theta_i)^{\alpha}]}{(\alpha - 1)\rho}$$

provided that the Θ_t, $t \geqslant 0$, are non-negative.

Corollary 4.11. *Assume that (X_t) is a univariate strictly stationary k-dependent sequence which is also regularly varying with index $\alpha > 1$ and has mean zero. Then we have for any $\rho > 0$,*

$$\frac{\mathbb{P}(\sup_{t \geqslant 0}(S_t - \rho t) > x)}{x\, \mathbb{P}(|X| > x)} \sim \frac{\lambda_k}{(\alpha - 1)\rho} \mathbb{E}\Big(\sup_{t \leqslant k} \sum_{i=0}^{t} \widetilde{\Theta}_i \Big)_+^{\alpha}, \quad x \to \infty \,.$$

The proof is given in Section 5.7.

Example 4.12. Consider the stochastic recurrence equation $X_t = A_t X_{t-1} + B_t$, $t \in \mathbb{Z}$, where (A_t, B_t), $t \in \mathbb{Z}$, constitute an $\mathbb{R}^2$-valued iid sequence. We assume that (X_t) constitutes a strictly stationary Markov chain.

The Goldie case: We assume the conditions of Goldie [28] are satisfied, ensuring that X_0 is regularly varying with index $\alpha > 0$. In particular, we have $A \geqslant 0$ a.s., $\mathbb{E}[A^\alpha] = 1$ for some positive α and we also need some further conditions on the distribution of the sequence (A_t, B_t) to ensure that one has a Nummelin regeneration scheme. Then (X_t) is regularly varying of order α, $\Theta_0 = 1$ and $\Theta_t = \Pi_t = A_1 \cdots A_t$, $t \geqslant 1$; see Basrak and Segers [9]. Proceeding as in the proof of Theorem 7.2 in [45] and using the drift condition[1] $(\mathbf{DC}_p)$ with $p < \alpha$, one can show the anti-clustering condition (3.11). In the proof of Theorem 4.6 in [44] we showed that the vanishing-small-values condition (4.10) without the supremum is satisfied under $(\mathbf{DC}_p)$ for $p < \alpha$. An inspection of the proof also shows that one may restrict oneself to the absolute values $|\overline{X}_t|$, implying the vanishing-small-values condition for the supremum as well. From Theorem 4.9 and Remark 4.10 we conclude that

$$\frac{\mathbb{P}(\sup_{t \geqslant 0}(S_t - (\rho + \mathbb{E}X)t) > x)}{x\, \mathbb{P}(|X| > x)} \sim \frac{\mathbb{E}\left[\left(1 + \sum_{i=1}^\infty \Pi_i\right)^\alpha - \left(\sum_{i=1}^\infty \Pi_i\right)^\alpha\right]}{(\alpha - 1)\rho}, \quad x \to \infty.$$

(4.9)

This result recovers Theorem 4.1 in Buraczewski et al. [12] in the case of a Nummelin regeneration scheme. The method of proof in [12] is completely different from ours.

We also mention that $Y_t \stackrel{d}{=} 1 + \sum_{i=1}^\infty \Pi_i$, where (Y_t) is the strictly stationary causal solution to the sequence $Y_t = A_t Y_{t-1} + 1$, $t \in \mathbb{Z}$, which is regularly varying with index α. Then the constant on the right-hand side of (4.9) can be written as

$$\frac{\mathbb{E}\left[Y_0^\alpha - (Y_0 - 1)^\alpha\right]}{(\alpha - 1)\rho}.$$

The Grey case: We assume now that the conditions of Grey [29] are satisfied. This means that A may assume real values, $\mathbb{E}|A|^\alpha < 1$ and B regularly varying with index α for some $\alpha > 0$. Then the unique strictly stationary solution to the stochastic recurrence equation exists and is regularly varying with index α and $\Theta_t/\Theta_0 = \Pi_t$ as above. Following Segers [56], we have in this case $\mathbb{P}(\Theta_{-j} = 0, j \geqslant 1) = \mathbb{P}(\Theta_{-1} = 0) = 1 - \mathbb{E}|\Theta_1|^\alpha > 0$ because $|\Theta_1| = |\Theta_0||A_1| = |A_1|$. Thus, one can turn to the simpler alternative expression of the right-hand side of (4.8)

$$(1 - \mathbb{E}|A|^\alpha)\frac{\mathbb{E}\left[\sup_{t \geqslant 0}\left(\sum_{i=0}^t \widetilde{\Theta}_i\right)_+^\alpha\right]}{(\alpha - 1)\rho},$$

where $\widetilde{\Theta}_t/\widetilde{\Theta}_0 = \Pi_t$, $\mathbb{P}(\widetilde{\Theta}_0 = 1) = \lim_{x \to \infty} \mathbb{P}(B > x)/\mathbb{P}(|B| > x) = p$ and $\mathbb{P}(\widetilde{\Theta}_0 = -1) = \lim_{x \to \infty} \mathbb{P}(B \leqslant -x)/\mathbb{P}(|B| > x) = q$. Then, we obtain

$$(1 - \mathbb{E}|A|^\alpha)\frac{\mathbb{E}\left[p \sup_{t \geqslant 0}\left(1 + \sum_{i=1}^t \Pi_i\right)_+^\alpha + q \sup_{t \geqslant 0}\left(1 + \sum_{i=1}^t \Pi_i\right)_-^\alpha\right]}{(\alpha - 1)\rho}.$$

We recover the result of Konstantinides and Mikosch [37] for $A \geqslant 0$ a.s., $p = 1$ and the one of Mikosch and Samorodnitsky [43] in the AR(1) case when $A = \phi$ for some $|\phi| < 1$. If $A \geqslant 0$ then the constant turns into

$$(1 - \mathbb{E}A^\alpha)\frac{p\, \mathbb{E}\left[\left(1 + \sum_{i=1}^t \Pi_i\right)^\alpha\right]}{(\alpha - 1)\rho}.$$

---

[1]We say that a function $X_t = g(\Phi_t)$ of a Markov chain (Φ_t) satisfies the $(\mathbf{DC}_p)$ condition if $\mathbb{E}[|g(\Phi_1)|^p \mid \Phi_0 = x] \leqslant \beta|g(x)|^p + b\mathbb{1}_A(x)$ where $0 < \beta < 1$ and A is a small set; see [42] for details.

Ruin bounds in the case of general linear processes $X_t = \sum_j \psi_j Z_{t-j}$ for iid regularly varying (Z_t) can be derived in a similar fashion using the computations of the spectral tail process given in Meinguet and Segers [41] recovering the results in [43].

Example 4.13. Consider the GARCH(1,1) model $X_t = \sigma_t Z_t$, $t \in \mathbb{Z}$; see Bollerslev [11]. Here (Z_t) is an iid mean zero and unit variance sequence of random variables and (σ_t^2) satisfies the stochastic recurrence equation $\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1) \sigma_{t-1}^2$, $t \in \mathbb{Z}$, where α_0, α_1 and β_1 are positive constants chosen such that (σ_t^2) is strictly stationary. Moreover, we assume that the above stochastic recurrence equation for (σ_t^2) with $B_t = \alpha_0$ and $A_t = \alpha_1 Z_{t-1}^2 + \beta_1$ satisfies the Goldie conditions of Example 4.12, ensuring that (σ_t^2) is regularly varying with index $\alpha/2$. Then an application of Breiman's multivariate results (see Basrak et al. [6]) implies that (X_t) is regularly varying with index α. As before, we write $\Pi_t = A_1 \cdots A_t$. Following [45], Section 5.4, we observe that as $x \to \infty$,

$$\frac{\mathbb{P}(|(X_0,\ldots,X_t) - \sigma_0(Z_0, \Pi_1^{0.5} Z_1, \ldots, \Pi_t^{0.5} Z_t)| > x)}{\mathbb{P}(\sigma > x)} = o(1)\,.$$

An application of the multivariate Breiman result yields

$$\frac{\mathbb{P}(x^{-1}\sigma_0(Z_0, Z_1 \Pi_1^{0.5}, \ldots, Z_t \Pi_t^{0.5}) \in \cdot)}{\mathbb{P}(|X| > x)} \overset{v}{\to} \frac{1}{\mathbb{E}|Z_0|^\alpha} \int_0^\infty \mathbb{P}(y(Z_0, Z_1 \Pi_i^{0.5}, \ldots, Z_t \Pi_t^{0.5}) \in \cdot)\, d(-y^{-\alpha}).$$

Then

$$\mathbb{P}(x^{-1}(X_0,\ldots,X_t) \in \cdot \mid |X_0| > x)$$
$$\overset{w}{\to} \frac{1}{\mathbb{E}|Z_0|^\alpha} \int_0^\infty \mathbb{P}(y(Z_0, Z_1 \Pi_i^{0.5}, \ldots, Z_t \Pi_t^{0.5}) \in \cdot, y|Z_0| > 1)\, d(-y^{-\alpha})$$
$$= \frac{1}{\mathbb{E}|Z_0|^\alpha} \mathbb{E}\Big[|Z_0|^\alpha \mathbb{1}_{|Y_0|(Z_0, Z_1 \Pi_i^{0.5}, \ldots, Z_t \Pi_t^{0.5}) \in |Z_0| \cdot}\Big]\,,$$

where $|Y_0|$ is Pareto distributed with index α and independent of (Z_t). By direct calculation, we obtain

$$\frac{\mathbb{E}\Big[\big(\sup_{t \geqslant 0} \sum_{i=0}^t \Theta_i\big)_+^\alpha - \big(\sup_{t \geqslant 1} \sum_{i=1}^t \Theta_i\big)_+^\alpha\Big]}{(\alpha-1)\rho}$$
$$= \frac{\mathbb{E}\Big[\sup_{t \geqslant 0}(Z_0 + \sum_{i=1}^t Z_i \Pi_i^{0.5})_+^\alpha - \sup_{t \geqslant 1}(\sum_{i=1}^t Z_i \Pi_i^{0.5})_+^\alpha\Big]}{\mathbb{E}|Z_0|^\alpha(\alpha-1)\rho}\,.$$

Thus we derived the scaling constant for the ruin probability in (4.8) in the case of a GARCH(1, 1) process. We also mention that the other conditions of Theorem 4.9 are satisfied. Indeed, the drift $(\mathbf{DC}_p)$ for $p < \alpha$ is satisfied, implying the anti-clustering and vanishing-small-values conditions as in the case of Example 4.12; see [45] for details.

Example 4.14. Consider a stochastic volatility model $X_t = \sigma_t Z_t$, $t \in \mathbb{Z}$, where (σ_t) is a strictly stationary sequence with lognormal marginals independent of an iid regularly varying sequence (Z_t). Then (X_t) is regularly varying with the same index and it is not difficult to see that $\Theta_t = 0$ for $t \neq 0$. Now an application of Theorem 4.9 yields the same ruin bound as in the iid case. This result supports the general theory of such models whose extremal behaviour mimics the one of an iid regularly varying sequence.

4.2.3. *Large deviations for multivariate sums on half-spaces.* The same techniques as in the previous section can be used to prove the following large deviation result for multivariate sums.

Theorem 4.15. *Consider a stationary $\mathbb{R}^d$-valued sequence (X_t) satisfying the following conditions*
  1. *($\mathbf{RV}_\alpha$) for some $\alpha > 0$,*

2. *the anti-clustering condition* (3.11).

If $\alpha > 1$ we also assume

3. *the vanishing-small-values condition*

$$(4.10) \qquad \lim_{\varepsilon \downarrow 0} \limsup_{n \to \infty} \sup_{x \in \Lambda_n} \frac{\mathbb{P}(x^{-1}|\overline{S}_n| > \delta)}{n\,\mathbb{P}(|X| > x)} = 0\,, \quad \delta > 0\,,$$

4. $\mathbb{E}\Big(\sum_{i=1}^{\infty} |\Theta_i| \Big)^{\alpha-1} < \infty.$

Then for every $\theta \in \mathbb{S}^{d-1}$,

$$\sup_{x \in \Lambda_n} \Big| \frac{\mathbb{P}\big(\theta' S_n > x\big)}{n\,\mathbb{P}(|X| > x)} - \mathbb{E}\Big[\Big(\theta' \sum_{i=0}^{\infty} \Theta_i \Big)_+^{\alpha} - \Big(\theta' \sum_{i=1}^{\infty} \Theta_i \Big)_+^{\alpha} \Big] \Big| \to 0\,, \quad n \to \infty\,.$$

In addition, if $\alpha \notin \mathbb{N}$ or X is symmetric, then there exists a unique Radon measure μ_α on $\overline{\mathbb{R}}_0^d$ such that $\mu_\alpha(t\cdot) = t^{-\alpha}\mu_\alpha(\cdot)$, $t > 0$, and for any sequence (x_n) such that $x_n \in \Lambda_n$, $n \geqslant 1$,

$$\frac{\mathbb{P}\ (x_n^{-1} S_n \in \cdot)}{n\,\mathbb{P}(|X| > x_n)} \xrightarrow{v} \mu_\alpha(\cdot)\,, \quad n \to \infty\,.$$

Moreover, μ_α is determined by its values on the subsets $\{y \in \mathbb{R}^d : \theta'y > 1\}$ for any $\theta \in \mathbb{S}^{d-1}$.

The proof of the last part follows by the same arguments as for Theorem 4.3 in Mikosch and Wintenberger [45].

4.3. **Large deviations for cluster functionals.** Following Yun [57] and Segers [55], we call a sequence of non-negative functions (c_l) on $\mathbb{R}^l$ a *cluster functional* if it satisfies $(\mathbf{C}_0)$. As for the restrictions f_l, we will often suppress the dependence of c_l on the index l; it will be clear from the context.

Simple examples of cluster functionals are

$$c_l(x_1, \ldots, x_l) \quad = \quad \sum_{i=1}^{l} \phi(x_i)\,, \quad l \geqslant 1\,,$$

with $\phi : \mathbb{R}^d \to \mathbb{R}^+$ satisfying $\phi(0) = 0$ and, for $d = 1$,

$$c_l(x_1, \ldots, x_l) \quad = \quad \sum_{i=1}^{l} (x_i - z)_+\,, \quad l \geqslant 1\,,$$

$$c_l(x_1, \ldots, x_l) \quad = \quad \max_{1 \leqslant i \leqslant l} (x_i - z)_+\,, \quad l \geqslant 1\,,$$

for some $z \geqslant 0$; see [24, 57, 55] for further examples.

Corollary 4.16. *Assume that the strictly stationary $\mathbb{R}$-valued sequence (X_t) satisfies $(\mathbf{RV}_\alpha)$, the uniform anti-clustering condition* (3.11) *and that*

$$f_l(x_1, \ldots, x_l) = \mathbb{1}_{c_l((x_1-1)_+, \ldots, (x_l-1)_+) > 1}$$

satisfies $(\mathbf{C}_1)$. Then

$$\sup_{x \in \Lambda_n} \Big| \frac{\mathbb{P}(c((x^{-1}X_i - 1)_+)_{1 \leqslant i \leqslant n}) > 1)}{n\mathbb{P}(|X| > x)}$$

$$(4.11) \qquad -[\mathbb{P}(c(((|Y_0|\Theta_t - 1)_+)_{t \geqslant 0}) > 1) - \mathbb{P}(c(((|Y_0|\Theta_t - 1)_+)_{t \geqslant 1}) > 1)\Big| \to 0\,, \quad n \to \infty\,.$$

Proof. We apply Theorem 3.3 to (f_l) satisfying $(\mathbf{C_1})$ and we obtain the uniform limit for $x \in \Lambda_n$ as $n \to \infty$:

$$\int_0^\infty \left[\mathbb{P}(c(((y\Theta_i - 1)_+)_{i\geqslant 0}) > 1) - \mathbb{P}(c(((y\Theta_i - 1)_+)_{i\geqslant 1}) > 1) \right] d(-y^{-\alpha}).$$

By assumption, $|\Theta_0| = 1$ and therefore we can restrict the area of integration to $[1, \infty)$, recovering the desired limit. $\qquad\square$

Remark 4.17. In the k-dependent case, the anti-clustering condition is trivially satisfied. In view of Theorem 3.3 relation (4.11) then turns into

$$\sup_{x \geqslant x_n} \left| \frac{\mathbb{P}\Big(c_n \big(([x^{-1}X_i - 1]_+)_{1 \leqslant i \leqslant n}\big) > 1 \Big)}{n \, \mathbb{P}(|X| > x)} - \lambda_k \, \mathbb{P}\Big(c_{k+1}\big(([\,|Y_0|\widetilde{\Theta}_i - 1]_+)_{0 \leqslant i \leqslant k}\big) > 1\Big) \right| \to 0 \,.$$

Remark 4.18. Corollary 4.16 is formulated for univariate sequences (X_t). However, one can generalize this result in various ways; see for example Drees and Rootzén [24]. For example, let (X_t) be an $\mathbb{R}^d$-valued sequence satisfying the conditions of Theorem 3.3 and $A \subset \overline{\mathbb{R}}_0^d$ be a Borel set whose distance to the origin is $\geqslant \varepsilon > 0$. Moreover, let $g : \mathbb{R}^d \to \mathbb{R}$ be a measurable function. If g is a.e. continuous and A is a continuity set with respect to Lebesgue measure then the same argument as for Corollary 4.16 now yields

$$\sup_{x \in \Lambda_n} \left| \frac{\mathbb{P}(c_n(g(x^{-1}X_i)\mathbb{1}_{x^{-1}X_i \in A})_{1 \leqslant i \leqslant n}) > 1)}{n \mathbb{P}(|X| > x)} \right.$$
$$\left. - \left[\mathbb{P}(c((g(|Y_0|\Theta_i)\mathbb{1}_{|Y_0|\widetilde{\Theta}_i \in A})_{i \geqslant 0}) > 1) - \mathbb{P}(c((g(|Y_0|\Theta_i)\mathbb{1}_{|Y_0|\widetilde{\Theta}_i \in A})_{i \geqslant 1}) > 1) \right] \right| \to 0 \,.$$

4.4. Beyond condition $(\mathbf{C_\varepsilon})$: Convergence of the tail empirical point process.
In this section we consider an example, where the condition $(\mathbf{C_\varepsilon})$ in Theorem 3.1 is not satisfied but Jakubowski's telescoping sum approach is also applicable, yielding a limit result. Related theory was developed in Balan and Louhichi [3] beyond the framework of regularly varying sequences, in the context of triangular arrays of strictly stationary sequences and infinitely divisible limit laws; see also Jakubowski and Kobus [34].

Recycling notation, we define the random measures

$$(4.12) \qquad N_{nj} = k_n^{-1} \sum_{t=1}^j \varepsilon_{a_m^{-1} X_t} \,, \quad j = 1, \ldots, n, \quad N_{nn} = N_n \,,$$

where $m = m_n \to \infty$ and $k_n = [n/m] \to \infty$. The *tail empirical point process* N_n plays an important role in extreme value statistics; see Resnick and Stărică [52], Resnick [51], Drees and Rootzén [24] and the references therein.

Theorem 4.19. *Consider a strictly stationary $\mathbb{R}^d$-valued sequence (X_t), satisfying the following conditions:*

1. *$(\mathbf{RV}_\alpha)$ for some $\alpha > 0$,*
2. *the mixing condition $\mathcal{A}(a_n)$ modified for the random measures (4.12),*
3. *the anti-clustering condition (3.11) with $\Lambda_n = \{a_n\}$.*

Then the relation $N_n \xrightarrow{P} \mu_1$ holds in the space of random measures on $\overline{\mathbb{R}}_0^d$ equipped with the vague topology, where $n \, \mathbb{P}(a_n^{-1} X \in \cdot) \xrightarrow{v} \mu_1$ as $n \to \infty$.

The proof is given in Section 5.8.

Remark 4.20. Closely related results can be found in Resnick and Stărică [52] in the 1-dimensional case. Their mixing and anti-clustering conditions are slightly different and they prove results for triangular arrays under a vague tightness condition (which is satisfied for $(\mathbf{RV}_\alpha)$), similar to Balan and Louhichi [3]. It follows from the results in Resnick and Stărică [52] that Theorem 4.19 implies the consistency of the Hill estimator of α in the case of positive random variables (X_t), i.e., if $m_n \to \infty$ and $m_n/n \to 0$ then

$$\Big(\sum_{t=1}^{m_n-1} \log \Big(X_{(n-i+1)}/X_{(n-m_n+1)} \Big) \Big)^{-1} \xrightarrow{P} \alpha\,, \quad n \to \infty\,,$$

where $X_{(1)} \leqslant \cdots \leqslant X_{(n)}$ a.s. is the ordered sample of $X_1, \ldots, X_n$.

5. Proofs

5.1. **Proof of Theorem 3.1.** For a given integer $n \geqslant 1$ and $\delta > 0$, we define

$$c_x(j_1, j_2) \;=\; \begin{cases} f(x^{-1}X_{j_1}, \ldots, x^{-1}X_{j_2}) & 1 \leqslant j_1 \leqslant j_2 \leqslant n \\ 0 & j_1 > j_2, \end{cases}$$

$$h_k(x^{-1}X_j, \ldots, x^{-1}X_n) \;=\; \begin{cases} \mathbb{1}_{|X_j| \leqslant x\delta} & k = 0 \\ \mathbb{1}_{M_{j+k,n} \leqslant x\delta} & k \geqslant 1\,. \end{cases}\;, \quad j \leqslant n\,.$$

By a telescoping sum argument, we decompose

$$c_x(1, n) = \sum_{j=1}^{n} [c_x(j, n) - c_x(j+1, n)]\,.$$

We denote $\Delta_x(n) = c_x(1, n) - \sum_{j=1}^{n} [c_x(j, j+k) - c_x(j+1, j+k)]$ for some $k \geqslant 1$. By stationarity, we have

$$\mathbb{E}\Delta_x(n) = \mathbb{E}f(x^{-1}X_1, \ldots, x^{-1}X_n) - n\left[\mathbb{E}f(x^{-1}X_0, \ldots, x^{-1}X_k) - \mathbb{E}f(x^{-1}X_1, \ldots, x^{-1}X_k)\right].$$

Moreover, as f satisfies $(\mathbf{C}_\varepsilon)$ and assuming without loss of generality that $|f| \leqslant 1$, we obtain for any $\delta \leqslant \varepsilon$ that

$$\begin{aligned} |\Delta_x(n)| &= \left| \sum_{j=1}^{n} [c_x(j,n) - c_x(j+1,n)] - [c_x(j,j+k) - c_x(j+1,j+k)] \right| \\ &\leqslant \left| \sum_{j=1}^{n} c_x(j,n)(1 - h_0(x^{-1}X_j, \ldots, x^{-1}X_n))(1 - h_k(x^{-1}X_j, \ldots, x^{-1}X_n)) \right| \\ &\leqslant \sum_{j=1}^{n} \mathbb{1}_{|X_j| > x\delta} \mathbb{1}_{M_{j+k,n} > x\delta}\,. \end{aligned}$$

Using stationarity, we obtain

$$\begin{aligned} \mathbb{E}|\Delta_x(n)| &\leqslant \sum_{j=1}^{n} \mathbb{P}(|X_j| > x\delta\,, M_{j+k,n} > x\delta) \\ &\leqslant n\, \mathbb{P}(|X_0| > x\delta\,, M_{k,n} > x\delta)\,. \end{aligned}$$

Using the uniform anti-clustering condition (3.11), we conclude that

$$\lim_{k \to \infty} \limsup_{n \to \infty} \sup_{x \in \Lambda_n} \frac{\mathbb{E}|\Delta_x(n)|}{n\, \mathbb{P}(|X_0| > x)} = 0\,, \quad \delta > 0\,.$$

It remains to prove the existence of the limit, uniformly for $x \in \Lambda_n$,

$$\lim_{k\to\infty} \lim_{n\to\infty} \frac{\mathbb{E}[f(x^{-1}X_0, \ldots, x^{-1}X_k) - f(x^{-1}X_1, \ldots, x^{-1}X_k)]}{\mathbb{P}(|X_0| > x)},$$

and to identify it. By regular variation and $(\mathbf{C}_\varepsilon)$, we have

$$\sup_{x\in\Lambda_n} \left| \frac{\mathbb{E}[f(x^{-1}X_0, \ldots, x^{-1}X_k) - f(x^{-1}X_1, \ldots, x^{-1}X_k)]}{\mathbb{P}(|X_0| > x)} \right.$$
$$\left. - \int_0^\infty \mathbb{E}[f(y\Theta_0, \ldots, y\Theta_k) - f(0, y\Theta_1, \ldots, y\Theta_k)]\, d(-y^{-\alpha}) \right|$$
$$= \sup_{x\in\Lambda_n} \left| \frac{\mathbb{E}\left[\left(f(x^{-1}X_0, \ldots, x^{-1}X_k) - f(x^{-1}X_1, \ldots, x^{-1}X_k)\right) \mathbb{1}_{|X_0|>x} \right]}{\mathbb{P}(|X_0| > x)} \right.$$
$$\left. - \int_0^\infty \mathbb{E}[f(y\Theta_0, \ldots, y\Theta_k) - f(0, y\Theta_1, \ldots, y\Theta_k)]\, d(-y^{-\alpha}) \right|$$
$$\to \quad 0, \quad n \to \infty.$$

Finally, for every $k \geqslant 1$, in view of $(\mathbf{C}_\varepsilon)$,

$$\int_0^\infty \mathbb{E}[f(y\Theta_0, \ldots, y\Theta_k) - f(0, y\Theta_1, \ldots, y\Theta_k)]\, d(-y^{-\alpha})$$

(5.13)
$$= \int_\varepsilon^\infty \mathbb{E}[f(y\Theta_0, \ldots, y\Theta_k) - f(0, y\Theta_1, \ldots, y\Theta_k)]\, d(-y^{-\alpha}).$$

The absolute value of the integrand is bounded by 2, hence integrable on $[\varepsilon, \infty)$. Moreover, under the anti-clustering condition (3.11) we have $\Theta_k \overset{P}{\to} 0$ as $k \to \infty$ as follows from the Remark 3.2 above. Therefore and by $(\mathbf{C}_\varepsilon)$ the limits $\lim_{k\to\infty} f(y\Theta_0, \ldots, y\Theta_k)$ exist and are finite for $y > 0$. Dominated convergence implies that one may let $k \to \infty$ in (5.13) and interchange the limit and the integral.

   Combining the partial limit results above, we conclude that the theorem is proved.

5.2. **Proof of Theorem 3.3.** We start by collecting some properties of the sequence (Θ_t) which are specific for a k-dependent regularly varying sequence. In particular, we will show that the conditional probability laws (3.14) are well defined.

Proposition 5.1. *Assume the stationary $\mathbb{R}^d$-valued k-dependent sequence (X_t) is such that $(X_0, \ldots, X_k)$ is regularly varying with index α then (X_t) satisfies $(\mathbf{RV}_\alpha)$. The following properties also hold:*

- $\mathbb{P}(\Theta_t = 0) = 1$ *for* $|t| \geqslant k + 1$,
- *For* $1 \leqslant |t| \leqslant k$, $\mathbb{P}(\Theta_t \neq 0, \Theta_{t+j} \neq 0) = 0$ *for* $|j| \geqslant k + 1$,
- $\mathbb{P}(\Theta_{-t} = 0, t = 1, \ldots, k) > 0$.

Proof of Proposition 5.1. Assume that $(X_0, \ldots, X_k)$ is regularly varying and that (X_t) is k-dependent. Then, for any $|t| > k$, we have by independence of X_t and X_0 that $\mathbb{P}(\Theta_t = 0) = 1$. As the limit law is degenerate, the convergence in the definition of Θ_t holds also in probability; $\mathbb{P}(|X_t|/|X_0| \leqslant \varepsilon \mid |X_0| > x) \to 1$ for $\varepsilon > 0$. By an application of a Slutsky argument, as $(X_0, \ldots, X_k)/|X_0|$ converges in distribution, conditionally on $|X_0| > x$, then it is also true that $(X_0, \ldots, X_t)/|X_0|$ given $|X_0| > x$ converges to $(\Theta_0, \ldots, \Theta_t) = (\Theta_0, \ldots, \Theta_k, 0, \ldots, 0)$.

   The second property of the spectral tail process follows from the fact that

$$\mathbb{P}(|Y_t| > \varepsilon, |Y_{t-j}| > \varepsilon) = 0, \quad \varepsilon > 0, \quad |j| \geqslant k + 1.$$

Indeed, by independence of X_t and X_{t-j} for $|j| > k$,

$$
\mathbb{P}(|X_t| \wedge |X_{t-j}| > \varepsilon a_n \mid |X_0| > a_n) \leqslant \frac{\mathbb{P}(|X_t| \wedge |X_{t-j}| > \varepsilon a_n)}{\mathbb{P}(|X_0| > a_n)}
$$

$$
= \frac{[\mathbb{P}(|X_0| > \varepsilon a_n)]^2}{\mathbb{P}(|X_0| > a_n)} \sim \varepsilon^{-2\alpha} \mathbb{P}(|X| > a_n) \to 0 \,.
$$

The second property implies that

$$
\mathbb{P}(\max_{-k \leqslant t < 0} |\Theta_t| > 0, \Theta_k \neq 0) = 0 \,.
$$

Then

$$
\mathbb{P}(\Theta_k \neq 0) = \mathbb{P}(\max_{-k \leqslant t < 0} |\Theta_t| = 0, \Theta_k \neq 0) \leqslant \mathbb{P}(\max_{-k \leqslant t < 0} |\Theta_t| = 0) \,,
$$

and the third property follows if $\mathbb{P}(\Theta_k \neq 0) > 0$. Now assume that $\mathbb{P}(\Theta_k \neq 0) = 0$. By the time change formula in Basrak and Segers [9], $\mathbb{P}(\Theta_k \neq 0) = \mathbb{E}|\Theta_{-k}|^\alpha$ and $\mathbb{P}(\Theta_{-k} = 0) = 1$. Thus $\max_{-k \leqslant t < 0} |\Theta_t| = \max_{-k < t < 0} |\Theta_t|$ a.s. A recursive argument yields the third property in the general case. $\qquad \square$

Proof of Theorem 3.3. We apply Theorem 3.1 for $\Lambda_n = [x_n, \infty)$. By virtue of k-dependence the anti-clustering condition is trivially satisfied. In view of Proposition 5.1 it remains to show that

$$
\lambda_k \int_0^\infty \mathbb{E}(f_{k+1}(y\widetilde{\Theta}_0, \ldots, y\widetilde{\Theta}_k)) d(-y^{-\alpha})
$$

(5.14)
$$
= \int_0^\infty \mathbb{E}[f_{k+1}(y\Theta_0, \ldots, y\Theta_k) - f_{k+1}(0, y\Theta_1, \ldots, y\Theta_k)] d(-y^{-\alpha}) \,.
$$

We observe that

$$
f_{k+1}(y\Theta_0, \ldots, y\Theta_k) \mathbb{1}_{\Theta_{-j}=0, j=1,\ldots,k}
$$

(5.15)
$$
= f_{k+1}(y\Theta_0, \ldots, y\Theta_k) - \sum_{t=1}^k f_{k+1}(y\Theta_0, \ldots, y\Theta_k) \mathbb{1}_{\Theta_{-k+t} \neq 0, \Theta_{-k+t+1}=0, \ldots, \Theta_{-1}=0} \,.
$$

From k-dependence we conclude that $\Theta_t \neq 0$ implies $\Theta_{t+j} = 0$, $j \geqslant k+1$. Therefore

$$
f_{k+1}(y\Theta_0, \ldots, y\Theta_k) \mathbb{1}_{\Theta_{-k+t} \neq 0, \Theta_{-k+t+1}=0, \ldots, \Theta_{-1}=0}
$$
$$
= f_{k+1}(y\Theta_0, \ldots, y\Theta_t, 0, \ldots, 0) \mathbb{1}_{\Theta_{-k+t} \neq 0, \Theta_{-k+t+1}=0, \ldots, \Theta_{-1}=0} \,.
$$

By the time change formula in Theorem 3.1 (iii) in Basrak and Segers [9] and the property $(\mathbf{C_\varepsilon})$,

$$
\mathbb{E}[f_{k+1}(y\Theta_0, \ldots, y\Theta_t, 0, \ldots, 0) \mathbb{1}_{\Theta_{-k+t} \neq 0, \Theta_{-k+t+1}=0, \ldots, \Theta_{-1}=0}]
$$
$$
= \mathbb{E}[f_{k+1}(y\Theta_{k-t}|\Theta_{k-t}|^{-1}, \ldots, y\Theta_k|\Theta_{k-t}|^{-1}, 0, \ldots, 0)|\Theta_{k-t}|^\alpha \mathbb{1}_{\Theta_0 \neq 0, \Theta_1=0, \ldots, \Theta_{k-t-1}=0}]
$$
$$
= \mathbb{E}[f_{k+1}(y\Theta_{k-t}|\Theta_{k-t}|^{-1}, \ldots, y\Theta_k|\Theta_{k-t}|^{-1})|\Theta_{k-t}|^\alpha \mathbb{1}_{\Theta_1=0, \ldots, \Theta_{k-t-1}=0}] \,.
$$

Now, by Fubini's Theorem when $\Theta_{k-t} \neq 0$, first integrating with respect to y and then changing variables:

$$
\int_0^\infty \mathbb{E} f_{k+1}(y\Theta_{k-t}|\Theta_{k-t}|^{-1}, \ldots, y\Theta_k|\Theta_{k-t}|^{-1}) \mathbb{1}_{\Theta_1=0, \ldots, \Theta_{k-t-1}=0, \Theta_{k-t} \neq 0} \, d(-(y|\Theta_{k-t}|^{-1})^{-\alpha})
$$
$$
= \int_0^\infty \mathbb{E} f_{k+1}(y\Theta_{k-t}, \ldots, y\Theta_k) \mathbb{1}_{\Theta_1=0, \ldots, \Theta_{k-t-1}=0, \Theta_{k-t} \neq 0} \, d(-y^{-\alpha}) \,.
$$

Taking into account (5.15) and the previous identities, we obtain

$$\int_0^\infty \mathbb{E}(f_{k+1}(y\Theta_0,\ldots,y\Theta_k)\mathbb{1}_{\Theta_{-j}=0,j=1,\ldots,k})d(-y^{-\alpha})$$

$$= \int_0^\infty \mathbb{E}[f_{k+1}(y\Theta_0,\ldots,y\Theta_k)]d(-y^{-\alpha})$$

$$- \sum_{i=1}^k \int_0^\infty \mathbb{E}[f_{k+1-i}(y\Theta_i,\ldots,y\Theta_k)\mathbb{1}_{\Theta_j=0,j=1,\ldots,i-1,\Theta_i\neq0}]d(-y^{-\alpha})\,.$$

Since (f_ℓ) satisfies the consistency property $(\mathbf{C}_\varepsilon)$ the right-hand side turns into

$$\int_0^\infty \mathbb{E}[f_{k+1}(y\Theta_0,\ldots,y\Theta_k)]d(-y^{-\alpha}) - \int_0^\infty \mathbb{E}\Big[f_{k+1}(0,y\Theta_1,\ldots,y\Theta_k)\sum_{i=1}^k \mathbb{1}_{\Theta_j=0,j=1,\ldots,i-1,\Theta_i\neq0}\Big]d(-y^{-\alpha})$$

and, as $f(0,\ldots,0)=0$, the desired result follows. $\qquad\square$

5.3. **Proof of Theorem 4.3.** We start with the case $\alpha \in (0,1)$. In this case, the negligibility condition

$$\lim_{\varepsilon\downarrow0}\limsup_{n\to\infty}\mathbb{P}(|a_n^{-1}\overline{S}_n| > \delta) = 0\,,\quad \delta > 0\,,$$

is satisfied. Indeed, an application of Markov's inequality and Karamata's theorem yields

$$\mathbb{P}(|a_n^{-1}\overline{S}_n| > \delta) \leqslant \frac{\mathbb{E}|X|\mathbb{1}_{|X|\leqslant\varepsilon a_n}}{a_n\varepsilon\mathbb{P}(|X| > \varepsilon a_n)}\varepsilon n\,\mathbb{P}(|X| > \varepsilon a_n) \to c\,\varepsilon^{1-\alpha}\,,\quad n\to\infty\,,$$

and the right-hand side vanishes as $\varepsilon \downarrow 0$. Therefore we may focus on the limit behavior of the sequence $(a_n^{-1}\underline{S}_n)$. Fix a small value $\varepsilon \in (0,1)$. The mixing condition (4.1) implies that

$$\log\mathbb{E}e^{is'\underline{S}_n/a_n} \sim k_n\log\mathbb{E}e^{is'\underline{S}_m/a_n} \sim -k_n\Big(1 - \mathbb{E}e^{is'\underline{S}_m/a_n}\Big) \sim \frac{\mathbb{E}e^{is'\underline{S}_m/a_n} - 1}{m\,\mathbb{P}(|X| > a_n)}\,.$$

We define the functions

$$f_l(x_1,\ldots,x_l) = \exp\Big(iys'\sum_{t=1}^l \underline{x}_t\Big) - 1\,,\quad l \geqslant 0\,,$$

where $\underline{x}_t = x_t\mathbb{1}_{|x_t|>\varepsilon a_n}$. The sequence (f_l) satisfies $(\mathbf{C}_\varepsilon)$ and $m\mathbb{P}(|X| > a_n) \to 0$ as $n \to \infty$. An application of Theorem 4.3 yields

$$\frac{\mathbb{E}e^{is'\underline{S}_m/a_n} - 1}{m\,\mathbb{P}(|X| > a_n)} \to \int_0^\infty \mathbb{E}\Big[e^{iys'\sum_{j=0}^\infty \underline{\Theta}_j} - e^{iys'\sum_{j=1}^\infty \underline{\Theta}_j}\Big]d(-y^{-\alpha})\,,\quad s\in\mathbb{R}^d\,,\quad n\to\infty\,.$$

Here $\underline{\Theta}_i = \Theta_i\mathbb{1}_{y|\Theta|>\varepsilon}$. To complete the proof we have to justify that we can let $\varepsilon \downarrow 0$ in the limiting expression. We observe that the integrand vanishes on the event $\{|y\Theta_0| \leqslant \varepsilon\} = \{y \leqslant \varepsilon\}$. Therefore the right-hand side turns into

$$\int_\varepsilon^\infty \mathbb{E}\Big[\big(e^{iys'\Theta_0} - 1\big)e^{iys'\sum_{j=1}^\infty \underline{\Theta}_j}\Big]d(-y^{-\alpha})\,.$$

The integrand is uniformly bounded and therefore integrable at infinity. In order to apply dominated convergence as $\varepsilon \downarrow 0$, we observe that for some constant $c > 0$,

$$\int_\varepsilon^1 \mathbb{E}\Big|\big(e^{iys'\Theta_0} - 1\big)e^{iys'\sum_{j=1}^\infty \underline{\Theta}_j}\Big|d(-y^{-\alpha}) \leqslant \int_\varepsilon^1 \mathbb{E}|ys'\Theta_0|\,d(-y^{-\alpha}) \leqslant c\int_0^1 y^{-\alpha}\,dy\,.$$

Now an application of the dominated convergence theorem as $\varepsilon \downarrow 0$ yields the desired log-characteristic function of an α-stable law.

The proof in the case $\alpha \in (1,2)$ is similar. In view of the negligibility condition (4.2) we may focus on the limit behavior of $(a_n^{-1}(\underline{S}_n - \mathbb{E}\underline{S}_n))$. Since $\mathbb{E}\underline{S}_n/a_n$ converges, the mixing condition remains valid for the corresponding centered sums. Then

$$\log \mathbb{E}e^{\,is'(\underline{S}_n - \mathbb{E}\underline{S}_n)/a_n} \sim \frac{\mathbb{E}e^{\,is(\underline{S}_m - \mathbb{E}\underline{S}_m)/a_n} - 1}{m\,\mathbb{P}(|X > a_n)}\,.$$

We also have

$$\left|\left(\mathbb{E}e^{\,is'(\underline{S}_m - \mathbb{E}\underline{S}_m)/a_n} - 1\right) - \left(\mathbb{E}e^{\,is'\underline{S}_m/a_n} - 1 - is'\mathbb{E}\underline{S}_m/a_n\right)\right|$$

$$= \left|\left(\mathbb{E}e^{\,is'\underline{S}_m/a_n} - 1\right)\left(e^{-is'\mathbb{E}\underline{S}_m/a_n} - 1\right) + \left(e^{-is'\mathbb{E}\underline{S}_m/a_n} - 1 + is'\mathbb{E}\underline{S}_m/a_n\right)\right|$$

$$\leqslant c\left(\mathbb{E}|\underline{S}_m|/a_n\right)^2$$

$$\leqslant c\left(m\mathbb{E}|X/a_n|\mathbb{1}_{|X|>\varepsilon a_n}\right)^2 = O(m\,\mathbb{P}(|X| > a_n)) = o(1)\,, \quad n \to \infty\,.$$

and

$$\begin{aligned}
\frac{a_n^{-1}s'\mathbb{E}\underline{S}_m}{m\,\mathbb{P}(|X| > a_n)} &= \frac{s'\mathbb{E}\underline{X}_0}{a_n\mathbb{P}(|X| > a_n)} \\
&\sim \mathbb{E}(s'X_0/(\varepsilon a_n) \mid |X_0| > \varepsilon a_n)\varepsilon^{1-\alpha} \\
&\to s'\mathbb{E}Y_0\,\varepsilon^{1-\alpha} \\
&= \mathbb{E}|Y_0|\,s'\mathbb{E}\Theta_0\,\varepsilon^{1-\alpha} \\
(5.16) \qquad\qquad &= s'\mathbb{E}\Theta_0\frac{\alpha}{\alpha-1}\varepsilon^{1-\alpha}\,.
\end{aligned}$$

An application of Theorem 3.1 yields as $n \to \infty$,

$$\frac{\mathbb{E}\left[e^{\,is'\underline{S}_m/a_n} - 1 - is'\mathbb{E}\underline{S}_m/a_n\right]}{m\,\mathbb{P}(|X| > a_n)}$$

$$\to \int_0^\infty \mathbb{E}\left[\left(e^{\,iys'\sum_{j=0}^\infty \underline{\Theta}_j} - 1 - iys'\sum_{j=0}^\infty \underline{\Theta}_j\right) - \left(e^{\,iys'\sum_{j=1}^\infty \underline{\Theta}_j} - 1 - iys'\sum_{j=1}^\infty \underline{\Theta}_j\right)\right]d(-y^{-\alpha})$$

$$= \int_\varepsilon^\infty \mathbb{E}\left[\left(e^{\,iys'\Theta_0} - 1\right)\left(e^{\,iys'\sum_{j=1}^\infty \underline{\Theta}_j} - 1\right) + \left(e^{\,iys'\Theta_0} - 1 - iys'\Theta_0\right)\right]d(-y^{-\alpha})\,.$$

The integrand is bounded by cy for large values of y and therefore integrable at infinity. We also have

$$\int_\varepsilon^1 \mathbb{E}\left|\left(e^{\,iys'\Theta_0} - 1\right)\left(e^{\,iys'\sum_{j=1}^\infty \underline{\Theta}_j} - 1\right) + \left(e^{\,iys'\Theta_0} - 1 - iys'\Theta_0\right)\right|d(-y^{-\alpha})$$

$$\leqslant c\left(\sum_{j=1}^\infty \mathbb{E}|\Theta_j| + 1\right)\int_0^1 y^2\,d(-y^{-\alpha}) < \infty$$

Therefore an application of the dominated convergence theorem yields the desired log-characteristic function of a stable law in the case $\alpha \in (1,2)$.

5.4. **Proof of Theorem 4.5.** We start with the case $\alpha > 1$. We have for any $\delta > 0$,

$$\mathbb{P}\left(\sup_{t\leqslant n}\underline{S}_t > (1+\delta)x\right) - \mathbb{P}\left(\sup_{t\leqslant n}|\overline{S}_t| > \delta x\right)$$

$$(5.17) \qquad \leqslant \mathbb{P}\left(\sup_{t\leqslant n}S_t > x\right) \leqslant \mathbb{P}\left(\sup_{t\leqslant n}\underline{S}_t > (1-\delta)x\right) + \mathbb{P}\left(\sup_{t\leqslant n}|\overline{S}_t| > \delta x\right)\,.$$

In view of condition (4.4), the limiting behavior will be determined by the ratios

$$\frac{\mathbb{P}(\sup_{t\leqslant n}\underline{S}_t > (1\pm\delta)x)}{n\mathbb{P}(|X_0|>x)} = \frac{\mathbb{E}(f_n(x^{-1}X_1,\ldots,x^{-1}X_n))}{n\mathbb{P}(|X_0|>x)}$$

with the functions $f_l(x_1,\ldots,x_l) = \mathbb{1}_{\sup_{t\leqslant l}(\underline{x}_1+\cdots+\underline{x}_t)>(1\pm\delta)}$ for fixed small $\delta,\varepsilon>0$. The functions f_l satisfy the condition (C_ε) and an application of Theorem 3.1 yields for $\varepsilon<1$ as $n\to\infty$,

$$\sup_{x\in\Lambda_n}\Big|\frac{\mathbb{P}\big(\sup_{i\leqslant n}\underline{S}_i>(1\pm\delta)x\big)}{n\,\mathbb{P}(|X|>x)}$$
$$-(1\pm\delta)\int_0^\infty\Big[\mathbb{P}\Big(\Theta_0+\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i>y^{-1}\Big)-\mathbb{P}\Big(\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i>y^{-1}\Big)\Big]d(-y^{-\alpha})\Big|\to 0\,.$$

Finally, we can take limits as $\varepsilon,\delta\downarrow 0$, using a domination argument. The domination argument is justified because we have

$$\int_0^\infty\Big[\mathbb{P}\Big(\Theta_0+\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i>y^{-1}\Big)-\mathbb{P}\Big(\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i>y^{-1}\Big)\Big]d(-y^{-\alpha})$$

$$= \alpha\int_0^\infty\Big[\mathbb{P}\Big(\Theta_0+\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i>z\Big)-\mathbb{P}\Big(\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i>z\Big)\Big]z^{\alpha-1}\,dz$$

$$= \alpha\int_0^\infty\mathbb{E}\Big[\Big(\mathbb{1}_{\Theta_0=1,1+\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i>z}-\mathbb{1}_{\Theta_0=1,\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i>z}\Big)$$
$$+\Big(\mathbb{1}_{\Theta_0=-1,-1+\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i>z}-\mathbb{1}_{\Theta_0=-1,\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i>z}\Big)\Big]z^{\alpha-1}\,dz$$

$$= \mathbb{E}\Big[\mathbb{1}_{\Theta_0=1}\Big[\Big(1+\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i\Big)_+^\alpha-\Big(\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i\Big)_+^\alpha\Big]$$
$$-\mathbb{1}_{\Theta_0=-1}\Big[\Big(\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i\Big)_+^\alpha-\Big(-1+\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i\Big)_+^\alpha\Big]\Big]$$

$$= \mathbb{E}\Big[\Big(\Theta_0+\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i\Big)_+^\alpha-\Big(\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i\Big)_+^\alpha\Big]\,.$$

Also observe that

$$(5.18)\qquad \mathbb{E}\Big|\Big(\Theta_0+\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i\Big)_+^\alpha-\Big(\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i\Big)_+^\alpha\Big|$$
$$\leqslant c\Big[1+\mathbb{E}\Big(\sum_{i=1}^\infty|\underline{\Theta}_i|\Big)^{\alpha-1}\Big]\leqslant c\Big[1+\mathbb{E}\Big(\sum_{i=1}^\infty|\Theta_i|\Big)^{\alpha-1}\Big]\,.$$

Under the assumptions of the theorem, the right-hand side is finite. An application of Lebesgue dominate convergence yields that

$$\mathbb{E}\Big[\Big(\Theta_0+\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i\Big)_+^\alpha-\Big(\sup_{t\geqslant 1}\sum_{i=1}^t\underline{\Theta}_i\Big)_+^\alpha\Big]\to\mathbb{E}\Big[\Big(\Theta_0+\sup_{t\geqslant 1}\sum_{i=1}^t\Theta_i\Big)_+^\alpha-\Big(\sup_{t\geqslant 1}\sum_{i=1}^t\Theta_i\Big)_+^\alpha\Big],\quad \varepsilon\to 0\,.$$

This concludes the proof in the case $\alpha>1$.

In the case $\alpha \in (0,1]$ one can follow the lines of the proof but instead of (5.18) one can use concavity to obtain the bound

$$\mathbb{E}\Big|\Big(\Theta_0 + \sup_{t\geqslant 1}\sum_{i=1}^{t}\underline{\Theta}_i\Big)_+^{\alpha} - \Big(\sup_{t\geqslant 1}\sum_{i=1}^{t}\underline{\Theta}_i\Big)_+^{\alpha}\Big| \leqslant c\,.$$

An application of Lebesgue dominated convergence finishes the proof.

5.5. **Proof of Corollary 4.7.** We apply Theorem 4.5. For a k-dependent sequence, the anti-clustering condition (3.11) is trivially satisfied and we also have $\mathbb{E}|\Theta_i|^{\alpha} < \infty$ for all i. The vanishing-small-values condition is verified in Lemma 5.2 below. Now one can follow the lines of the proof of Theorem 4.5 but instead of Theorem 3.1 apply Theorem 3.3 to obtain

$$\sup_{x\geqslant x_n}\Big|\frac{\mathbb{P}(\sup_{t\leqslant n}\underline{S}_t > (1\pm\delta)x)}{n\mathbb{P}(|X_0| > x)}$$
$$-(1\pm\delta)^{-\alpha}\lambda_k\int_0^{\infty}\mathbb{P}\Big(\sup_{t\leqslant k}y\sum_{i=0}^{t}\Theta_i\mathbb{1}_{y|\Theta_i|>\varepsilon} > 1\Big)d(-y^{-\alpha})\Big| \to 0\,.$$

Finally, we can take limits as $\varepsilon, \delta \downarrow 0$, using a domination argument and the fact that $\mathbb{E}|\Theta_i|^{\alpha} < \infty$ for all i.

Lemma 5.2. *Assume the conditions of Corollary 4.7. Then*

$$(5.19) \qquad \lim_{\varepsilon\downarrow 0}\limsup_{n\to\infty}\sup_{x\geqslant x_n}\frac{\mathbb{P}\Big(\sup_{t\leqslant n}|\overline{S}_t| > x\Big)}{n\mathbb{P}(|X| > x)} = 0\,.$$

Proof of Lemma 5.2. We observe that

$$\sup_{t\leqslant n}|\overline{S}_t| \leqslant \sum_{j=1}^{k+1}\sup_{s(k+1)\leqslant n}\Big|\sum_{t=0}^{s}\overline{X}_{(k+1)t+j}\Big|\,.$$

Therefore for $\delta > 0$,

$$\frac{\mathbb{P}\Big(\sup_{t\leqslant n}|\overline{S}_t| > \delta x\Big)}{n\mathbb{P}(|X_0| > x)} \quad\leqslant\quad \frac{(k+1)\mathbb{P}\Big(\sup_{s(k+1)\leqslant n}|\sum_{t=0}^{s}\overline{X}_{(k+1)t}| > \delta x/(k+1)\Big)}{n\mathbb{P}(|X_0| > x)} = I(x)\,.$$

If $\alpha \in (0,1)$, we have by Markov's inequality,

$$\sup_{x\geqslant x_n}I(x) \quad\leqslant\quad \sup_{x\geqslant x_n}\frac{(k+1)\mathbb{P}\Big(\sum_{t=1}^{[n/(k+1)]}|\overline{X}_{kt}| > \delta x/(k+1)\Big)}{n\mathbb{P}(|X_0| > x)}$$
$$\leqslant\quad c\sup_{x\geqslant x_n}\frac{\mathbb{E}|\overline{X}_0|}{x\mathbb{P}(|X_0| > x)}$$
$$\leqslant\quad c\sup_{x\geqslant x_n}\frac{\mathbb{E}(|X_0|\mathbb{1}_{|X_0|\leqslant x\varepsilon})}{\mathbb{E}(|X_0|\mathbb{1}_{|X_0|\leqslant x})}\sup_{x\geqslant x_n}\frac{\mathbb{E}(|X_0|\mathbb{1}_{|X_0|\leqslant x})}{x\mathbb{P}(|X_0| > x)}\,.$$

The second supremum is bounded in view of Karamata's theorem and the first supremum is bounded by $c\varepsilon^{1-\alpha}$ uniformly for small ε is view of the uniform convergence theorem for regularly varying functions. Hence the right-hand side converges to zero by first letting $n \to \infty$ and then $\varepsilon \downarrow 0$.

If $\alpha \geqslant 1$, we have

$$I(x) \leqslant \frac{(k+1)\mathbb{P}\Big(\sup_{s(k+1)\leqslant n}|\sum_{t=0}^{s}(\overline{X}_{(k+1)t} - \mathbb{E}\overline{X})| > \delta(x - \delta^{-1}nx^{-1}|\mathbb{E}\overline{X}|)/(k+1)\Big)}{n\mathbb{P}(|X_0| > x)}\,.$$

If $\mathbb{E}|X| < \infty$ and $\mathbb{E}X = 0$ we have by Karamata's theorem and uniform convergence

$$nx^{-1}|\mathbb{E}\overline{X}| \leqslant nx^{-1}\mathbb{E}|X|\mathbb{1}_{|X|>\varepsilon x} \leqslant c(\varepsilon)n\mathbb{P}(|X| > x) = o(1)\,, \quad n \to \infty\,.$$

The last identity follows from the fact that $n\mathbb{P}(|X| > x_n) = o(1)$ as $n \to \infty$. In the case $\alpha = 1$ and $\mathbb{E}|X| = \infty$ we require the corresponding condition (4.6). Therefore we have

$$I(x) \leqslant \frac{(k+1)\mathbb{P}\Big(\sup_{s(k+1)\leqslant n}|\sum_{t=0}^{s}(\overline{X}_{(k+1)t} - \mathbb{E}\overline{X})| > 0.5\delta x/(k+1)\Big)}{n\mathbb{P}(|X_0| > x)}\,.$$

uniformly for $x \geqslant x_n$ and large n. To ease notation, we will assume without loss of generality that $E\overline{X} = 0$. Now an application of the Fuk-Nagaev inequality (see Petrov [48], p. 78) for $p > \alpha \vee 2$ yields,

$$I(x) \leqslant \frac{c\Big(nx^{-p}\mathbb{E}|\overline{X}|^p + \mathrm{e}^{-cx^2(n\mathbb{E}\overline{X}^2)^{-1}}\Big)}{n\mathbb{P}(|X| > x)} \leqslant c\frac{\mathbb{E}|\overline{X}|^p}{x^p\mathbb{P}(|X| > x)} + c\frac{\mathrm{e}^{-cx^2(n\mathbb{E}\overline{X}^2)^{-1}}}{n\mathbb{P}(|X| > x)}\,.$$

The first summand on the right-hand side converges to $\varepsilon^{p-\alpha}$ uniformly for small ε and uniformly for $x \geqslant x_n$. The second term is uniformly negligible for $x \geqslant x_n$ if $\alpha < 2$ in view of the relation $x^2(n\mathbb{E}\overline{X}^2)^{-1} \sim c(n\mathbb{P}(|X| > x))^{-1}$. If $\alpha = 2$ and $\mathbb{E}X^2 = \infty$, Karamata's theorem yields that $\mathbb{E}(\overline{X}/x)^2/\mathbb{P}(|X| > x)$ is a slowly varying function converging to infinity as $x \to \infty$, and then the growth condition $x_n/n^{0.5+\delta} \to \infty$ ensures that the second term is negligible. If $\mathbb{E}|X|^2 < \infty$, writing $x = \sqrt{n\log n}y$ for $y > 0$ sufficiently large, we obtain

$$\frac{\mathrm{e}^{-cx^2(n\mathbb{E}\overline{X}^2)^{-1}}}{n\mathbb{P}(|X| > x)} \leqslant n^{\alpha+\varepsilon-cy^2-1}y^{\alpha+\varepsilon} \leqslant n^{-\eta}$$

for some $\eta > 0$ and thus the right-hand side converges to 0 uniformly for $x \geqslant x_n$. $\square$

5.6. **Proof of Theorem 4.9.** We start by proving that

$$(5.20) \qquad \lim_{C\to\infty}\limsup_{x\to\infty}\frac{\mathbb{P}(\sup_{t>Cx}(S_t - \rho t) > x)}{x\,\mathbb{P}(|X| > x)} \;=\; 0\,.$$

Assume that $[Cx] \in D_k = (2^k, 2^{k+1}]$ for some $k \geqslant 1$. Then

$$\mathbb{P}\big(\sup_{t>Cx}(S_t - \rho t) > x\big) \;\leqslant\; \sum_{l=k}^{\infty}\mathbb{P}\big(\sup_{t\in D_l}S_t > x + \rho\,2^l\big)$$

$$\leqslant\; \sum_{l=k}^{\infty}\mathbb{P}\big(\sup_{t\leqslant 2^{l+1}}S_t > x + \rho\,2^l\big)\,.$$

We observe that $x + \rho\,2^l \in D_l$ and therefore we are in the range where we may apply the large deviation results for suprema in Theorem 4.5. Then the right-hand side above can be bounded uniformly by

$$c\sum_{l=k}^{\infty}2^l\,\mathbb{P}(|X| > x + \rho\,2^l) \;\leqslant\; c\int_{Cx}^{\infty}\mathbb{P}(|X| > \rho y + x)\,dy$$

$$\leqslant\; c\,(Cx)\mathbb{P}(|X| > Cx) \sim c\,C^{1-\alpha}\,x\,\mathbb{P}(|X| > x)\,.$$

In the last step we used Karamata's theorem. Relation (5.20) follows by observing that $\alpha > 1$.

In view of (5.20) it suffices to bound the probabilities $\mathbb{P}(\sup_{t\leqslant Cx}(S_t - \rho t) > x)$ as $x \to \infty$ for any large value of C. Moreover, in view of condition (4.4) we may replace the random walk (S_t) in the ruin probability by the truncated version $(\underline{S}_t)$, letting $\varepsilon \downarrow 0$ in the final bound. We then adapt the

proof of Theorem 3.1 to the case where the functional depends on the last index larger than $\varepsilon > 0$. More specifically, denoting

$$c_{x,j_2}(j_1, j_2) = 1_{\sup_{j_1 \leqslant t \leqslant j_2} (\sum_{i=j_1}^t \underline{X}_i - \rho t) > x}, \quad 1 \leqslant j_1 \leqslant j_2 \leqslant [Cx]$$

we follow the reasoning of Section 5.1. Observing that

$$\mathbb{E}\big([c_{x,[Cx]}(j, [Cx]) - c_{x,[Cx]}(j+1, [Cx])] - [c_{x,j+k}(j, j+k) - c_{x,j+k}(j+1, j+k)]\big)$$
$$\leqslant \mathbb{P}(|X_j| > x\varepsilon, M_{j+k,[Cx]} > x\varepsilon),$$

we use the stationarity and the anti-clustering condition (3.11) to estimate the bound by

$$(5.21) \qquad\qquad \mathbb{P}(M_{k,[Cx]} > \varepsilon x, |X_0| > \varepsilon x) \leqslant C_k \, \mathbb{P}(|X| > x)$$

for large x and a real sequence (C_k) (depending on C) such that $C_k \to 0$ as $k \to \infty$. Therefore it suffices to study the limiting behavior of the difference of the two terms emerging from the telescoping argument when $x \to \infty$ and then $k \to \infty$; we indicate how we treat the first one:

$$\sum_{j=1}^{[Cx]} \mathbb{E}(c_{x,j+k}(j, j+k)) = \sum_{j=1}^{[Cx]} \mathbb{P}\Big(\sup_{j \leqslant t \leqslant j+k} \Big(\sum_{i=j}^t \underline{X}_i - \rho t \Big) > x \Big) = \sum_{j=1}^{[Cx]} \mathbb{P}\Big(\sup_{0 \leqslant t \leqslant k} \Big(\sum_{i=0}^t \underline{X}_i - \rho(t+j) \Big) > x \Big).$$

We have

$$\mathbb{P}\big(\sup_{0 \leqslant t \leqslant k} (\underline{S}_t + \underline{X}_0) > x + \rho k \big) \leqslant \mathbb{P}\big(\sup_{0 \leqslant t \leqslant k} (\underline{S}_t + \underline{X}_0 - \rho(t+j)) > x \big) \leqslant \mathbb{P}\big(\sup_{0 \leqslant t \leqslant k} (\underline{S}_t + \underline{X}_0) > x + \rho j \big).$$

We obtain the following sandwich bound

$$x \int_{1/x}^{C} \mathbb{P}\big(\sup_{0 \leqslant t \leqslant k} (\underline{S}_t + \underline{X}_0) > (1 + u\rho)x + \rho k \big) du \leqslant \sum_{j=1}^{[Cx]} \mathbb{P}\big(\sup_{0 \leqslant t \leqslant k} (\underline{S}_t + \underline{X}_0 - \rho(t+j)) > x \big)$$
$$\leqslant x \int_0^{C+x^{-1}} \mathbb{P}\big(\sup_{0 \leqslant t \leqslant k} (\underline{S}_t + \underline{X}_0) > (1 + u\rho)x \big) du.$$

By first letting $x \to \infty$, using classical arguments from regular variation theory, in particular the uniform convergence theorem, we obtain

$$\frac{\sum_{j=1}^{[Cx]} \mathbb{P}(\sup_{0 \leqslant t \leqslant k}(\underline{S}_t + \underline{X}_0 - \rho(t+j)) > x)}{x\mathbb{P}(|X| > x)} du \sim \int_0^C \frac{\mathbb{E}(\sup_{t \leqslant k} \sum_{i=0}^t \Theta_i)_+^\alpha}{(1 + u\rho)^\alpha} du.$$

We then determine the limit of the difference of the two terms emerging from the telescoping argument. For any $k \geqslant 1$ we have

$$\frac{\mathbb{P}(\sup_{t \leqslant Cx}(S_t - \rho t) > x)}{x\mathbb{P}(|X| > x)} = \int_0^C \frac{\mathbb{E}[(\sup_{t \leqslant k} \sum_{i=0}^t \Theta_i)_+^\alpha - (\sup_{1 \leqslant t \leqslant k} \sum_{i=1}^t \Theta_i)_+^\alpha]}{(1 + u\rho)^\alpha} du + O(CC_k).$$

The last term vanishes when $k \to \infty$ for any $C > 0$. Moreover, one can use uniform convergence and let $\varepsilon \to 0$. Finally, letting $C \to \infty$ and observing that $\int_0^\infty (1 + \rho u)^{-\alpha} du = \rho^{-1}(\alpha - 1)^{-1}$, we obtain the desired ruin bound.

5.7. **Proof of Corollary 4.11.** The proof follows the lines of the proof of Theorem 4.9. We mention that condition 3.11 is trivially satisfied and the vanishing-small-values condition holds in view of Lemma 5.2.

5.8. **Proof of Theorem 4.19.** As in the proof of Theorem 4.1, we have for any $g \in \mathbb{C}_K^+$ with $g(x) = 0$ for $|x| \leqslant \delta$ for some $\delta > 0$, in view of the mixing condition $\mathcal{A}(a_n)$,

$$-\log \mathbb{E}\mathrm{e}^{-\int g dN_n} \sim k_n\big(1 - \mathbb{E}\mathrm{e}^{-\int g dN_{nm}}\big), \quad n \to \infty.$$

Now we can proceed as in the beginning of the proof of Theorem 3.1. For $k \geqslant 2$,

$$\left| k_n \mathbb{E}\big(1 - \mathrm{e}^{-\int g dN_{nm}}\big) - n\big[\mathbb{E}\big(1 - \mathrm{e}^{-\int g dN_{n,k-1}}\big) - \mathbb{E}\big(1 - \mathrm{e}^{-\int g dN_{n,k}}\big)\big] \right|$$
$$\leqslant \quad \mathbb{P}(M_{k,m} > \delta a_m \mid |X_0| > a_m\delta).$$

The right-hand side is negligible by virtue of (3.11). Applying a Taylor expansion, we have for some random $\xi \in (0,1)$,

$$n\mathbb{E}\big(1 - \mathrm{e}^{-\int g dN_{nk}}\big) \quad = \quad n\Big[\mathbb{E}\int g dN_{nk} - 0.5\mathbb{E}\Big[\mathrm{e}^{-\xi \int g dN_{nk}}\Big(\int g\, dN_{nk}\Big)^2\Big]\Big] = I_1 - I_2.$$

Let $\mu_{0,l}$, $l \geqslant 1$, be the limit measures of regular variation for $a_n^{-1}(X_0, X_l)$, then we have

$$I_1 \quad = \quad m_n k\, \mathbb{E}g(a_m^{-1}X) \to k \int g\, d\mu_1,$$

$$I_2 \quad \leqslant \quad nk_n^{-2}\mathbb{E}\Big(\sum_{t=1}^{k} g(a_m^{-1}X_t)\Big)^2$$

$$= \quad k_n^{-1}\Big[k\, m_n\, \mathbb{E}g^2(a_m^{-1}X) + 2\sum_{h=1}^{k-1}(k-h)\, m_n\, \mathbb{E}[g(a_m^{-1}X_0)g(a_m^{-1}X_h)]\Big]$$

$$\sim \quad k_n^{-1}\Big[k \int g^2\, d\mu_1 + 2\sum_{h=1}^{k-1}(k-h)\int g(x)\, g(y)\, \mu_{0,h}(dx, dy)\Big] \to 0, \quad n \to \infty.$$

Finally, we have

$$\lim_{n\to\infty} k_n \mathbb{E}(1 - \mathrm{e}^{-\int g\, dN_{nm}}) \quad = \quad \lim_{k\to\infty}\lim_{n\to\infty} n\Big[\mathbb{E}\big(1 - \mathrm{e}^{-\int g dN_{n,k-1}}\big) - \mathbb{E}\big(1 - \mathrm{e}^{-\int f dN_{n,k}}\big)\Big]$$
$$= \quad \int g\, d\mu_1.$$

This proves the convergence of the Laplace functionals

$$\mathbb{E}\mathrm{e}^{-\int g\, dN_n} \to \mathrm{e}^{-\int g\, d\mu_1}, \quad g \in \mathbb{C}_K^+,$$

hence $N_n \xrightarrow{P} \mu_1$.

Remark 5.3. The proof shows that the condition $(\mathbf{RV}_\alpha)$ is not really needed in this case. It suffices that X_0 is regularly varying and $n\,\mathbb{P}(a_n^{-1}(X_0, X_h) \in \cdot) \xrightarrow{v} \mu_{0,h}$ for every $h \geqslant 1$, where $\mu_{0,h}$ is a Radon measure on $\overline{\mathbb{R}}_0^d$ which, possibly, is the null measure.

References

[1] Asmussen, S. and Albrecher, H. (2010) *Ruin Probabilities, 2nd edition* World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ.

[2] Avram, F. and Taqqu, M.S. (1992) Weak convergence of sums of moving averages in the α-stable domain of attraction. *Ann. Probab.*, **20**, 483–503.

[3] Balan, R.M. and Louhichi, S. (2009) Convergence of point processes with weakly dependent points. *J. Theor. Probab.* **22**, 955–982.

[4] BALAN, R. AND LOUHICHI, S. (2010) Explicit conditions for the convergence of point processes associated to statonary arrays. *Electronic Commun. in Probab.* **15**, 428–441.

[5] BARTKIEWICZ, K., JAKUBOWSKI, A., MIKOSCH, T. AND WINTENBERGER, O. (2011) Stable limits for sums of dependent infinite variance random variables. *Probab. Th. Rel. Fields* **150**, 337–372.

[6] BASRAK, B., DAVIS, R.A. AND MIKOSCH. T. (2002) A characterization of multivariate regular variation. *Ann. Appl. Probab.* **12**, 908–920.

[7] BASRAK, B. AND KRIZMANIĆ, D. (2014) A limit theorem for moving averages in the α-stable domain of attraction, *Stoch. Proc. Appl.*, **124**, 1070–1083,

[8] BASRAK, B., KRIZMANIĆ, D. AND SEGERS, J. (2012) A functional limit theorem for dependent sequences with infinite variance stable limits. *Ann. Probab.* **40**, 2008–2033.

[9] BASRAK, B. AND SEGERS, J. (2009) Regularly varying multivariate time series. *Stoch. Proc. Appl.* **119**, 1055–1080.

[10] BINGHAM, N.H., GOLDIE, C.M. AND TEUGELS, J.L. (1987) *Regular Variation.* Cambridge University Press, Cambridge.

[11] BOLLERSLEV, T. (1986) Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**, 307–327.

[12] BURACZEWSKI, D., DAMEK, E., MIKOSCH, T. AND ZIENKIEWICZ, J. (2013) Large deviations for solutions to stochastic recurrence equations under Kesten's condition. *Ann. Probab.* **41**, 2401–3050.

[13] DAVIS, R.A. AND HSING, T. (1995) Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Prob.* **23**, 879–917.

[14] DAVIS, R.A. AND MIKOSCH, T. (1998) Limit theory for the sample ACF of stationary process with heavy tails with applications to ARCH. *Ann. Statist.* **26**, 2049–2080.

[15] DAVIS, R.A., MIKOSCH, T. AND ZHAO, Y. (2013) Measures of serial extremal dependence and their estimation. *Stoch. Proc. Appl.* **123**, 2575–2602.

[16] DAVIS, R.A. AND RESNICK, S.I. (1985) Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13**, 179–195.

[17] DAVIS, R.A. AND RESNICK, S.I. (1985) More limit theory for the sample correlation function of moving averages. *Stoch. Proc. Appl.* **20** , 257–279.

[18] DAVIS, R.A. AND RESNICK, S.I. (1986) Limit theory for the sample covariance and correlation functions of moving averages. *Ann. Statist.* **14**, 533–558.

[19] DAVIS, R.A. AND RESNICK, S.I. (1996) Limit theory for bilinear processes with heavy-tailed noise. *Ann. Appl. Probab.* **6**, 1191–1210.

[20] DEDECKER,J. AND DOUKHAN, P. (2003) A new covariance inequality and applications. *Stoch. Proc. Appl.* **106**, 63–80.

[21] DEDECKER, J., DOUKHAN, P., LANG, G., LEÓN, R., LOUHICHI, S. AND PRIEUR, C. (2007) *Weak Dependence: With Examples and Applications.* Lecture Notes in Statistics **190**. Springer, New York.

[22] DEHLING, H., MIKOSCH, T. AND SØRENSEN, M. (EDS.) (2002) *Empirical Process Techniques for Dependent Data.* Birkhäuser, Boston.

[23] DOUKHAN, P., OPPENHEIM, G. AND TAQQU, M.S. (EDS.) (2003) *Theory and Applications of Long-Range Dependence.* Birkhäuser, Boston.

[24] DREES, H. AND ROOTZÉN, H. (2010) Limit theorems for empirical processes of cluster functionals. *Ann. Statist.* **38**, 2145–2186.

[25] EBERLEIN, E. AND TAQQU, M.S. (EDS.) (1986) *Dependence in Probability and Statistics. A Survey of Recent Results. Progress in Probability and Statistics* **11**. Birkhäuser, Boston.

[26] EMBRECHTS, P., KLÜPPELBERG, C. AND MIKOSCH, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[27] FELLER, W. (1971) *An Introduction to Probability Theory and Its Applications.* Vol. II. Second edition. Wiley, New York.

[28] GOLDIE, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1**, 126–166.

[29] GREY, D.R. (1994) Regular variation in the tail behaviour of solutions of random difference equations. *Ann. Appl. Probab.* **4**, 169–183.

[30] HAAN, L. DE, RESNICK, S.I., ROOTZÉN, H. AND VRIES, C.G. DE (1989) Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stoch. Proc. Appl.* **32**, 213–224.

[31] HSING, T. (1993) On some estimates based on sample behavior near high level excursions *Probab. Th. Rel. Fields* **95**, 331–356

[32] JAKUBOWSKI, A. (1993) Minimal conditions in *p*-stable limit theorems. *Stoch. Proc. Appl.* **44**, 291–327.

[33] JAKUBOWSKI, A. (1997) Minimal conditions in *p*-stable limit theorems - II. *Stoch. Proc. Appl.* **68**, 1–20.

[34] JAKUBOWSKI, A. AND KOBUS, M. (1989) α-stable limit theorems for sums of dependent random vectors. *J. Multivariate Anal.* **29**, 219–251.

[35] KALLENBERG, O. (1983) *Random Measures,* 3rd edition. Akademie–Verlag, Berlin.

[36] KOBUS, M. (1995) Generalized Poisson distributions as limits of sums for arrays of dependent random vectors. *J. Multivariate Anal.* **52**, 199–244.

[37] KONSTANTINIDES, D. AND MIKOSCH, T. (2005) Large deviations and ruin probabilities for solutions to stochastic recurrence equations with heavy-tailed innovations. *Ann. Probab.* **33**, 1992–2035.

[38] LEADBETTER, M.R. (1983) Extremes and local dependence of stationary sequences. *Z. Wahrscheinlichkeitsthoerie verw. Gebiete* **65**, 291–306.

[39] LEADBETTER, M.R., LINDGREN, G. AND ROOTZÉN, H. (1983) *Extremes and Related Properties of Random Sequences and Processes.* Springer, Berlin.

[40] LOUHICHI, S. AND RIO, E. (2011) Functional convergence to stable Lévy motions for iterated random Lipschitz mappings. *Electr. J. Probab.* **16**, 2452–2480.

[41] MEINGUET, T. AND SEGERS, J. (2010) Regularly varying time series in Banach spaces. *arXiv preprint* arXiv:1001.3262.

[42] MEYN, S.P. AND TWEEDIE, R.L. (1993) *Markov Chains and Stochastic Stability.* Springer, London.

[43] MIKOSCH, T. AND SAMORODNITSKY, G. (2000) The supremum of a negative drift random walk with dependent heavy-tailed steps. *Ann. Appl. Probab.* **10**, 1025–1064.

[44] MIKOSCH, T. AND WINTENBERGER, O. (2013) Precise large deviations for dependent regularly varying sequences. *Probab. Th. Rel. Fields,* **156**, 851–887.

[45] MIKOSCH, T. AND WINTENBERGER, O. (2013) The cluster index of regularly varying sequences with applications to limit theory for functions of multivariate Markov chains. *Probab. Theory Rel. Fields* (2014) **159**, 157–196.

[46] MIREK, M. (2011) Heavy tail phenomenon and convergence to stable laws for iterated Lipschitz maps. *Probab. Rel. Fields,* **151**, 705–734.

[47] O'BRIEN, G.L. (1987) Extreme values for stationary and Markov sequences. *Ann. Probab.* **15**, 281–291.

[48] PETROV, V.V. (1995) *Limit Theorems of Probability Theory.* Oxford University Press, Oxford (UK).

[49] RESNICK, S.I. (1986) Point processes, regular variation and weak convergence. *Adv. Appl. Probab.* **18**, 66–138.

[50] RESNICK, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[51] RESNICK, S.I. (2007) *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer, New York.

[52] RESNICK, S.I. AND STĂRICĂ, C. (1998) Tail index estmation for dependent data. *Ann. Appl. Probab.* **8**, 1156–1183.

[53] ROOTZÉN, H. (1976) Extremes of moving averages of stable processes. *Ann. Probab.* **6**, 847–869.

[54] SAMORODNITSKY, G. (2006) Long range dependence. *Found. Trends Stoch. Syst.* **1**, 163–257.

[55] SEGERS, J. (2005) Approximate distributions of clusters of extremes. *Statist. Probab. Letters* **74**, 330–336.

[56] SEGERS, J. (2007) Multivariate regular variation of heavy-tailed Markov chains. *arXiv preprint* math/0701411.

[57] YUN, S. (2000) The distributions of cluster functionals of extreme events in a dth order Markov chain. *J. Appl. Probab.* **37**, 29–44.

T. MIKOSCH, UNIVERSITY OF COPENHAGEN, UNIVERSITETSPARKEN 5, 2100 COPENHAGEN, DENMARK
E-mail address: `mikosch@math.ku.dk`

O. WINTENBERGER, SORBONNE UNIVERSITÉS, UPMC UNIV PARIS 06, LSTA, CASE 158 4 PLACE JUSSIEU, 75005 PARIS, FRANCE
E-mail address: `olivier.wintenberger@upmc.fr`