

An invariance principle for sums and record times of regularly varying stationary sequences

Bojan Basrak¹  · Hrvoje Planinić¹ · Philippe Soulier²

Received: 2 September 2016 / Revised: 4 December 2017 / Published online: 1 February 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract We prove a sequence of limiting results about weakly dependent stationary and regularly varying stochastic processes in discrete time. After deducing the limiting distribution for individual clusters of extremes, we present a new type of point process convergence theorem. It is designed to preserve the entire information about the temporal ordering of observations which is typically lost in the limit after time scaling. By going beyond the existing asymptotic theory, we are able to prove a new functional limit theorem. Its assumptions are satisfied by a wide class of applied time series models, for which standard limiting theory in the space D of càdlàg functions does not apply. To describe the limit of partial sums in this more general setting, we use the space E of so-called decorated càdlàg functions. We also study the running maximum of partial sums for which a corresponding functional theorem can be still expressed in the familiar setting of space D . We further apply our method to analyze record times in a sequence of dependent stationary observations, even when their marginal distribution is not necessarily regularly varying. Under certain restrictions on dependence among the observations, we show that the record times after scaling converge to a relatively simple compound scale invariant Poisson process.

✉ Bojan Basrak
bbasrak@math.hr

Hrvoje Planinić
planinic@math.hr

Philippe Soulier
philippe.soulier@parisnanterre.fr

¹ Department of Mathematics, University of Zagreb, Bijenička 30, Zagreb, Croatia

² Université Paris Ouest Nanterre, 92000 Nanterre, France

Keywords Point process · Regular variation · Invariance principle · Functional limit theorem · Record times

Mathematics Subject Classification 60F17 · 60G52 · 60G55 · 60G70

1 Introduction

Donsker-type functional limit theorems represent one of the key developments in probability theory. They express invariance principles for rescaled random walks of the form

$$S_{[nt]} = X_1 + \cdots + X_{[nt]}, \quad t \in [0, 1]. \quad (1.1)$$

Many extension of the original invariance principle exist, most notably allowing dependence between the steps X_i , or showing, like Skorohod did, that non-Gaussian limits are possible if the steps X_i have infinite variance. For a survey of invariance principles in the case of dependent variables in the domain of attraction of the Gaussian law, we refer to [24], see also [10] for a thorough survey of mixing conditions. In the case of a non-Gaussian limit, the limit of the processes $(S_{[nt]})_{t \in [0, 1]}$ is not a continuous process in general. Hence, the limiting theorems of this type are placed in the space of càdlàg functions denoted by $D([0, 1])$ under one of the Skorohod topologies. The topology denoted by J_1 is the most widely used (often implicitly) and suitable for i.i.d. steps, but over the years many theorems involving dependent steps have been shown using other Skorohod topologies. Even in the case of a simple m -dependent linear process from a regularly varying distribution, it is known that the limiting theorem cannot be shown in the standard J_1 topology, see Avram and Taqqu [2]. Moreover, there are examples of such processes for which none of the Skorohod topologies work, see Sect. 4.

However, as we found out, for all those processes and many other stochastic models relevant in applications, random walks do converge, but their limit exists in an entirely different space. To describe the elements of such a space we use the concept of decorated càdlàg functions and denote the corresponding space by $E([0, 1])$, following Whitt [34]. See Sect. 4. Presentation of this new type of limit theorem is the main goal of our article. For the statement of our main result see Theorem 4.5 in Sect. 4. As a related goal we also study the running maximum of the random walk $S_{[nt]}$ for which, due to monotonicity, the limiting theorem can still be expressed in the familiar space $D([0, 1])$.

Our main analytical tool is the limit theory for point processes in a certain nonlocally compact space which is designed to preserve the order of the observations as we rescale time to interval $[0, 1]$ as in (1.1). Observe that due to this scaling, successive observations collapse in the limit to the same time instance. As the first result in this context, we prove in Sect. 2 a limit theorem related to large deviations results of the type shown recently by [28] (cf. also [19]) and offer an alternative probabilistic interpretation of these results. Using our setup, we can group successive observations in the sequence $\{X_i, i = 1, \dots, n\}$, in nonoverlapping clusters of increasing size to define a point process which completely preserves the information about the order among the observations. This allows us to show in a rather straightforward manner that

so constructed empirical point processes converge in distribution towards a Poisson point process on an appropriate state space. The corresponding theorem could be arguably considered as the key result of the paper. It motivates all the theorems in the later sections and extends point process limit theorems in [5, 14], see Sect. 3.

Additionally, our method allows for the analysis of records and record times in the sequence of dependent stationary observations X_i . By a classical result of Rényi, the number of records among first n iid observations from a continuous distribution grows logarithmically with n . Moreover it is known (see e.g. [30]) that record times rescaled by n tend to the so-called scale invariant Poisson process, which plays a fundamental role in several areas of probability, see [1]. For a stationary sequence with an arbitrary continuous marginal distribution, we show that the record times converge to a relatively simple compound Poisson process under certain restrictions on dependence. This form of the limit reflects the fact that for dependent sequences records tend to come in clusters, as one typically observes in many natural situations. This is the content of Sect. 5. Finally, proofs of certain technical auxiliary results are postponed to Sect. 6. In the rest of the introduction, we formally present the main ingredients of our model.

We now introduce our main assumptions and notation. Let $\|\cdot\|$ denote an arbitrary norm on \mathbb{R}^d and let \mathbb{S}^{d-1} be the corresponding unit sphere. Recall that a d -dimensional random vector X is regularly varying with index $\alpha > 0$ if there exists a random vector $\Theta \in \mathbb{S}^{d-1}$ such that

$$\frac{1}{\mathbb{P}(\|X\| > x)} \mathbb{P}(\|X\| > ux, X/\|X\| \in \cdot) \Rightarrow u^{-\alpha} \mathbb{P}(\Theta \in \cdot), \tag{1.2}$$

for every $u > 0$ as $x \rightarrow \infty$, where \Rightarrow denotes the weak convergence of measures, here on \mathbb{S}^{d-1} . An \mathbb{R}^d -valued time series $\{X_t, t \in \mathbb{Z}\}$ is regularly varying if all the finite-dimensional vectors $(X_k, \dots, X_l), k \leq l \in \mathbb{Z}$ are regularly varying, see [14] for instance. We will consider a stationary regularly varying process $\{X_t, t \in \mathbb{Z}\}$. The regular variation of the marginal distribution implies that there exists a sequence $\{a_n\}$ which for all $x > 0$ satisfies

$$n\mathbb{P}(\|X_0\| > a_n x) \rightarrow x^{-\alpha}. \tag{1.3}$$

If $d = 1$, it is known that

$$n\mathbb{P}(X_0/a_n \in \cdot) \xrightarrow{v} \mu, \tag{1.4}$$

where \xrightarrow{v} denotes vague convergence on $\mathbb{R} \setminus \{0\}$ with the measure μ on $\mathbb{R} \setminus \{0\}$ given by

$$\mu(dy) = p\alpha y^{-\alpha-1} \mathbb{1}_{(0,\infty)}(y)dy + (1-p)\alpha(-y)^{-\alpha-1} \mathbb{1}_{(-\infty,0)}(y)dy \tag{1.5}$$

for some $p \in [0, 1]$.

According to [6], the regular variation of the stationary sequence $\{X_t, t \in \mathbb{Z}\}$ is equivalent to the existence of an \mathbb{R}^d -valued time series $\{Y_t, t \in \mathbb{Z}\}$ called the tail process which satisfies $\mathbb{P}(\|Y_0\| > y) = y^{-\alpha}$ for $y \geq 1$ and, as $x \rightarrow \infty$,

$$\left(\{x^{-1} X_t, t \in \mathbb{Z}\} \mid \|X_0\| > x \right) \xrightarrow{\text{fi.di.}} \{Y_t, t \in \mathbb{Z}\}, \tag{1.6}$$

where $\xrightarrow{\text{fi.di.}}$ denotes convergence of finite-dimensional distributions. Moreover, the so-called spectral tail process $\{\Theta_t, t \in \mathbb{Z}\}$, defined by $\Theta_t = Y_t/\|Y_0\|, t \in \mathbb{Z}$, turns out to be independent of $\|Y_0\|$ and satisfies

$$\left(\{|X_0|^{-1} X_t, t \in \mathbb{Z}\} \mid \|X_0\| > x \right) \xrightarrow{\text{fi.di.}} \{\Theta_t, t \in \mathbb{Z}\}, \tag{1.7}$$

as $x \rightarrow \infty$. If $d = 1$, it follows that p from (1.5) satisfies $p = \mathbb{P}(\Theta_0 = 1) = 1 - \mathbb{P}(\Theta_0 = -1)$.

We will often assume in addition that the following condition, referred to as the anticlustering or finite mean cluster length condition, holds.

Assumption 1.1 There exists a sequence of integers $(r_n)_{n \in \mathbb{N}}$ such that $\lim_{n \rightarrow \infty} r_n = \lim_{n \rightarrow \infty} n/r_n = \infty$ and for every $u > 0$,

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\max_{m \leq |i| \leq r_n} \|X_i\| > a_n u \mid \|X_0\| > a_n u \right) = 0. \tag{1.8}$$

There are many time series satisfying the conditions above including several non-linear models like stochastic volatility or GARCH (see [26, Section 4.4]).

In the sequel, an important role will be played by the quantity θ defined by

$$\theta = \mathbb{P} \left(\sup_{t \geq 1} \|Y_t\| \leq 1 \right). \tag{1.9}$$

It was shown in [6, Proposition 4.2] that Assumption 1.1 implies that $\theta > 0$.

2 Asymptotics of clusters

Let l_0 be the space of double-sided \mathbb{R}^d -valued sequences converging to zero at both ends, i.e. $l_0 = \{x = \{x_i, i \in \mathbb{Z}\} : \lim_{|i| \rightarrow \infty} \|x_i\| = 0\}$. On l_0 consider the uniform norm

$$\|x\|_\infty = \sup_{i \in \mathbb{Z}} \|x_i\|,$$

which makes l_0 into a separable Banach space. Indeed, l_0 is the closure of all double-sided rational sequences with finitely many non zero terms in the Banach space of all bounded double-sided real sequences. Define the shift operator B on l_0 by $(Bx)_i = x_{i+1}$ and introduce an equivalence relation \sim on l_0 by letting $x \sim y$ if $y = B^k x$ for some $k \in \mathbb{Z}$. In the sequel, we consider the quotient space

$$\tilde{l}_0 = l_0 / \sim,$$

and define a function $\tilde{d} : \tilde{l}_0 \times \tilde{l}_0 \rightarrow [0, \infty)$ by

$$\tilde{d}(\tilde{x}, \tilde{y}) = \inf\{\|x' - y'\|_\infty : x' \in \tilde{x}, y' \in \tilde{y}\} = \inf\{\|B^k x - B^l y\|_\infty : k, l \in \mathbb{Z}\}.$$

for all $\tilde{x}, \tilde{y} \in \tilde{l}_0$, and all $x \in \tilde{x}, y \in \tilde{y}$. The proof of the following result can be found in Sect. 6.

Lemma 2.1 *The function \tilde{d} is a metric which makes \tilde{l}_0 a separable and complete metric space.*

One can naturally embed the set $\cup_{d \geq 1} (\mathbb{R}^d)^n \cup l_0$ into \tilde{l}_0 by mapping $x \in l_0$ to its equivalence class and an arbitrary finite sequence $x = (x_1, \dots, x_n) \in (\mathbb{R}^d)^n$ to the equivalence class of the sequence

$$(\dots, 0, 0, x, 0, 0, \dots),$$

which adds zeros in front and after it.

Let $\{Z_t, t \in \mathbb{Z}\}$ be a sequence distributed as the tail process $\{Y_t, t \in \mathbb{Z}\}$ conditionally on the event $\{\sup_{i \leq -1} \|Y_i\| \leq 1\}$ which, under Assumption 1.1, has a strictly positive probability (cf. [6, Proposition 4.2]). More precisely,

$$\mathcal{L}(\{Z_t, t \in \mathbb{Z}\}) = \mathcal{L}\left(\{Y_t, t \in \mathbb{Z}\} \mid \sup_{i \leq -1} \|Y_i\| \leq 1\right). \tag{2.1}$$

Since (1.8) implies that $\mathbb{P}(\lim_{|t| \rightarrow \infty} \|Y_t\| = 0) = 1$, see [6, Proposition 4.2], the sequence $\{Z_t\}$ in (2.1) can be viewed as a random element in l_0 and \tilde{l}_0 in a natural way. In particular, the random variable

$$L_Z = \sup_{j \in \mathbb{Z}} \|Z_j\|,$$

is a.s. finite and not smaller than 1 since $\mathbb{P}(\|Y_0\| > 1) = 1$. Due to regular variation and (1.8) one can show (see [7]) that for $v \geq 1$

$$\mathbb{P}(L_Z > v) = v^{-\alpha}.$$

One can also define a new sequence $\{Q_t, t \in \mathbb{Z}\}$ in \tilde{l}_0 as the equivalence class of

$$Q_t = Z_t / L_Z, \quad t \in \mathbb{Z}. \tag{2.2}$$

Consider now a block of observations (X_1, \dots, X_{r_n}) and define $M_{r_n} = \max_{1 \leq i \leq r_n} \|X_i\|$. It turns out that conditionally on the event that $M_{r_n} > a_n u$, the law of such a block has a limiting distribution and that L_Z and $\{Q_t\}$ are independent.

Theorem 2.2 *Under Assumption 1.1, for every $u > 0$,*

$$\mathcal{L} \left(\frac{X_1, \dots, X_{r_n}}{a_n u} \mid M_{r_n} > a_n u \right) \xrightarrow{d} \mathcal{L} (\{Z_t, t \in \mathbb{Z}\}),$$

as $n \rightarrow \infty$ in \tilde{l}_0 . Moreover, $\{Q_t, t \in \mathbb{Z}\}$ and L_Z in (2.2) are independent random elements with values in \tilde{l}_0 and $[0, \infty)$ respectively.

Proof Step 1. We write $\mathbf{X}_n(i, j) = (X_i, \dots, X_j)/a_n u$, $\mathbf{Y}(i, j) = (Y_i, \dots, Y_j)$ and $M_{k,l} = \max_{k \leq i \leq l} \|X_i\|$, $M_{k,l}^Y = \max_{k \leq i \leq l} \|Y_i\|$. By the Portmanteau theorem [8, Theorem 2.1], it suffices to prove that

$$\lim_{n \rightarrow \infty} \mathbb{E}[g(\mathbf{X}_n(1, r_n)) \mid M_{1,r_n} > a_n u] = \mathbb{E}[g(\{Y_t, t \in \mathbb{Z}\}) \mid M_{-\infty,-1}^Y \leq 1], \tag{2.3}$$

for every nonnegative, bounded and uniformly continuous function g on (\tilde{l}_0, \tilde{d}) .

Define the truncation $\tilde{\mathbf{x}}_\zeta$ at level ζ of $\tilde{\mathbf{x}} \in \tilde{l}_0$ by putting all the coordinates of $\tilde{\mathbf{x}}$ which are no greater in norm than ζ to zero, that is $\tilde{\mathbf{x}}_\zeta$ is the equivalence class of $(x_i \mathbb{1}_{\|x_i\| > \zeta})_{i \in \mathbb{Z}}$, where \mathbf{x} is a representative of $\tilde{\mathbf{x}}$. Note that by definition, $\tilde{d}(\tilde{\mathbf{x}}, \tilde{\mathbf{x}}_\zeta) \leq \zeta$.

For a function g on \tilde{l}_0 , define g_ζ by $g_\zeta(\tilde{\mathbf{x}}) = g(\tilde{\mathbf{x}}_\zeta)$. If g is uniformly continuous, then for each $\eta > 0$, there exists ζ such that $|g(\tilde{\mathbf{x}}) - g(\tilde{\mathbf{y}})| \leq \eta$ if $\tilde{d}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \leq \zeta$, that is, $\|g - g_\zeta\|_\infty \leq \eta$. Thus it is sufficient to prove (2.3) for g_ζ for $\zeta \in (0, 1)$.

One can now follow the steps of the proof of [6, Theorem 4.3]. Decompose the event $\{M_{1,r_n} > a_n u\}$ according to the smallest $j \in \{1, \dots, r_n\}$ such that $\|X_j\| > a_n u$. We have

$$\begin{aligned} &\mathbb{E}[g_\zeta(\mathbf{X}_n(1, r_n)); M_{1,r_n} > a_n u] \\ &= \sum_{j=1}^{r_n} \mathbb{E} [g_\zeta(\mathbf{X}_n(1, r_n)); M_{1,j-1} \leq a_n u < \|X_j\|]. \end{aligned} \tag{2.4}$$

Fix a positive integer m and let n be large enough so that $r_n \geq 2m + 1$. By the definition of g_ζ , for all $j \in \{m + 1, \dots, r_n - m\}$ we have that

$$M_{1,j-m-1} \vee M_{j+m+1,r_n} \leq a_n u \zeta \Rightarrow g_\zeta(\mathbf{X}_n(1, r_n)) = g_\zeta(\mathbf{X}_n(j - m, j + m)). \tag{2.5}$$

The proof is now exactly along the same lines as the proof of [6, Theorem 4.3] and we omit some details. Using stationarity, the decomposition (2.4), the relation (2.5) and the boundedness of g , we have,

$$\begin{aligned} &\mathbb{E} [g_\zeta(\mathbf{X}_n(1, r_n)) \mathbb{1}\{M_{1,r_n} > a_n u\}] \\ &\quad - r_n \mathbb{E} [g_\zeta(\mathbf{X}_n(-m, m)) \mathbb{1}\{M_{-m,-1} \leq a_n u\} \mathbb{1}\{\|X_0\| > a_n u\}] \\ &\leq 2m \|g\|_\infty \mathbb{P}(\|X_0\| > a_n u) + r_n \|g\|_\infty \mathbb{P}(M_{-r_n,-m-1} \\ &\quad \vee M_{m+1,r_n} > a_n u; \|X_0\| > a_n u). \end{aligned} \tag{2.6}$$

Next define $\theta_n = \mathbb{P}(M_{1,r_n} > a_n u) / \{r_n \mathbb{P}(\|X_0\| > a_n u)\}$. Under Assumption 1.1,

$$\lim_{n \rightarrow \infty} \theta_n = \mathbb{P}(\sup_{i \geq 1} \|Y_i\| \leq 1) = \theta, \tag{2.7}$$

where θ was defined in (1.9). See [6, Proposition 4.2]. Therefore, by Assumption 1.1, (2.6) and (2.7) we conclude that

$$\begin{aligned} & \lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \left| \mathbb{E}[g_\zeta(\mathbf{X}_n(1, r_n)) \mid M_{1,r_n} > a_n u] \right. \\ & \left. - \frac{1}{\theta_n} \mathbb{E}[g_\zeta(\mathbf{X}_n(-m, m)); M_{-m,-1} \leq a_n u \mid \|X_0\| > a_n u] \right| = 0. \end{aligned} \tag{2.8}$$

We now argue that, for every $m \geq 1$,

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E}[g_\zeta(\mathbf{X}_n(-m, m)); M_{-m,-1} \leq a_n u \mid \|X_0\| > a_n u] \\ & = \mathbb{E}[g_\zeta(\mathbf{Y}(-m, m)); M_{-m,-1}^Y \leq 1]. \end{aligned} \tag{2.9}$$

First observe that g_ζ , as a function on $(\mathbb{R}^d)^{2m+1}$, is continuous except maybe on the set $D_\zeta^{2m+1} = \{(x_1, \dots, x_{2m+1}) \in (\mathbb{R}^d)^{2m+1}; \|x_i\| = \zeta \text{ for some } i \in \{1, \dots, 2m+1\}\}$ and we have that $\mathbb{P}(\mathbf{Y}(-m, m) \in D_\zeta^{2m+1}) = 0$ since $\mathbf{Y} = \|Y_0\| \Theta$, $\|Y_0\|$ and Θ are independent and the distribution of $\|Y_0\|$ is Pareto therefore atomless. Observe similarly that the distribution of $M_{k,l}^Y = \|Y_0\| \max_{k \leq j \leq l} \|\Theta_j\|$ does not have atoms except maybe at zero. Therefore, since g_ζ is bounded, (2.9) follows by the definition of the tail process and the continuous mapping theorem.

Finally, since $\mathbf{Y}(-m, m) \rightarrow \mathbf{Y}$ a.s. in \tilde{l}_0 and since \mathbf{Y} has only finitely many coordinates greater than ζ , $g_\zeta(\mathbf{Y}(-m, m)) = g_\zeta(\mathbf{Y})$ for large enough m , almost surely. Thus, applying (2.8) and (2.9), we obtain by bounded convergence

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E}[g_\zeta(\mathbf{X}_n(1, r_n)) \mid M_{1,r_n} > a_n u] \\ & = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \frac{1}{\theta_n} \mathbb{E}[g_\zeta(\mathbf{X}_n(-m, m)); M_{-m,-1} \leq a_n u \mid \|X_0\| > a_n u] \\ & = \frac{1}{\theta} \lim_{m \rightarrow \infty} \mathbb{E}[g_\zeta(\mathbf{Y}(-m, m)); M_{-m,-1}^Y \leq 1] \\ & = \frac{1}{\theta} \mathbb{E}[g_\zeta(\mathbf{Y}); M_{-\infty,-1}^Y \leq 1]. \end{aligned}$$

Applying this to $g \equiv 1$ we obtain $\theta = \mathbb{P}(M_{-\infty,-1}^Y \leq 1)$. Hence (2.3) holds for g_ζ as we wanted to show.

Step 2. Observing that the mapping $\tilde{x} \mapsto (\tilde{x}, \|\mathbf{x}\|_\infty)$ is continuous on \tilde{l}_0 , we obtain for every $u > 0$,

$$\mathcal{L} \left(\frac{X_1, \dots, X_{r_n}}{a_n u}, \frac{M_{r_n}}{a_n u} \mid M_{r_n} > a_n u \right) \xrightarrow{d} \mathcal{L}(\{Z_t\}, L_Z). \tag{2.10}$$

Similarly, the mapping defined on $\tilde{l}_0 \times (0, \infty)$ by $(\tilde{x}, b) \mapsto \tilde{x}/b$ is again continuous. Hence, (2.10) implies

$$\mathcal{L} \left(\frac{X_1, \dots, X_{r_n}}{M_{r_n}}, \frac{M_{r_n}}{a_n u} \mid M_{r_n} > a_n u \right) \xrightarrow{d} \mathcal{L}(\{Q_t\}, L_Z) \tag{2.11}$$

by the continuous mapping theorem. To show the independence between L_Z and $\{Q_t\}$, it suffices to show

$$\mathbb{E} [g(\{Q_t\}) \mathbb{1}_{\{L_Z > v\}}] = \mathbb{E} [g(\{Q_t\})] P(L_Z > v), \tag{2.12}$$

for an arbitrary uniformly continuous function g on \tilde{l}_0 and $v \geq 1$.

By (2.11), the left-hand side of (2.12) is the limit of

$$\mathbb{E} \left[g \left(\frac{X_1, \dots, X_{r_n}}{M_{r_n}} \right) \mathbb{1}_{\{(a_n)^{-1} M_{r_n} > v\}} \mid M_{r_n} > a_n \right],$$

which further equals

$$\mathbb{E} \left[g \left(\frac{X_1, \dots, X_{r_n}}{M_{r_n}} \right) \mid M_{r_n} > a_n v \right] \frac{\mathbb{P}(M_{r_n} > a_n v)}{\mathbb{P}(M_{r_n} > a_n)}.$$

By (2.11), the first term in the product above tends to $\mathbb{E} [g(\{Q_t\})]$ as $n \rightarrow \infty$. On the other hand, by (2.7) and regular variation of $\|X_0\|$, the second term tends to $v^{-\alpha} = \mathbb{P}(L_Z > v)$. □

3 The point process of clusters

In this section we prove our main result on the point process asymptotics for the sequence $\{X_t, t \in \mathbb{Z}\}$. Prior to that, we discuss the topology of $w^\#$ -convergence.

3.1 Preliminaries on $w^\#$ -convergence

To study convergence in distribution of point processes on the non locally-compact space \tilde{l}_0 we use $w^\#$ -convergence and refer to [11, Section A2.6.] and [12, Section 11.1.] for details. Let \mathbb{X} be a complete and separable metric space and let $\mathcal{M}(\mathbb{X})$ denote the space of boundedly finite nonnegative Borel measures μ on \mathbb{X} , i.e. such that $\mu(B) < \infty$ for all bounded Borel sets B . The subset of $\mathcal{M}(\mathbb{X})$ of all point measures (that is measures μ such that $\mu(B)$ is a nonnegative integer for all bounded Borel sets B) is denoted by $\mathcal{M}_p(\mathbb{X})$. A sequence $\{\mu_n\}$ in $\mathcal{M}(\mathbb{X})$ is said to converge to μ in the $w^\#$ -topology, noted by $\mu_n \rightarrow_{w^\#} \mu$, if

$$\mu_n(f) = \int f d\mu_n \rightarrow \int f d\mu = \mu(f),$$

for every bounded and continuous function $f : \mathbb{X} \rightarrow \mathbb{R}$ with bounded support. Equivalently ([11, Proposition A2.6.II.]), $\mu_n \rightarrow_{w^\#} \mu$ refers to

$$\mu_n(B) \rightarrow \mu(B)$$

for every bounded Borel set B with $\mu(\partial B) = 0$. We note that when \mathbb{X} is locally compact, an equivalent metric can be chosen in which a set is relatively compact if and only if it is bounded, and $w^\#$ -convergence coincides with vague convergence. We refer to [20] or [30] for details on vague convergence. The notion of $w^\#$ -convergence is metrizable in such a way that $\mathcal{M}(\mathbb{X})$ is Polish ([11, Theorem A2.6.III.(i)]). Denote by $\mathcal{B}(\mathcal{M}(\mathbb{X}))$ the corresponding Borel sigma-field.

It is known, see [12, Theorem 11.1.VII], that a sequence $\{N_n\}$ of random elements in $(\mathcal{M}(\mathbb{X}), \mathcal{B}(\mathcal{M}(\mathbb{X})))$, converges in distribution to N , denoted by $N_n \xrightarrow{d} N$, if and only if

$$(N_n(A_1), \dots, N_n(A_k)) \xrightarrow{d} (N(A_1), \dots, N(A_k)) \text{ in } \mathbb{R}^k,$$

for all $k \in \mathbb{N}$ and all bounded Borel sets A_1, \dots, A_k in \mathbb{X} such that $N(\partial A_i) = 0$ a.s. for all $i = 1, \dots, k$.

Remark 3.1 As shown in [12, Proposition 11.1.VIII], this is equivalent to the pointwise convergence of the Laplace functionals, that is, $\lim_{n \rightarrow \infty} \mathbb{E}[e^{-N_n(f)}] = \mathbb{E}[e^{-N(f)}]$ for all bounded and continuous function f on \mathbb{X} with bounded support. It turns out that it is sufficient (and more convenient in our context) to verify the convergence of Laplace functionals for a smaller convergence determining family.

See the comments before Assumption 3.5.

3.2 Point process convergence

Consider now the space $\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$ with the subspace topology. Following [21], we metrize the space $\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$ with the complete metric

$$\tilde{d}'(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \left(\tilde{d}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) \wedge 1 \right) \vee |1/\|\tilde{\mathbf{x}}\|_\infty - 1/\|\tilde{\mathbf{y}}\|_\infty|$$

which is topologically equivalent to \tilde{d} , i.e. it generates the same (separable) topology on $\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$. However, a subset A of $\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$ is bounded for \tilde{d}' if and only if there exists an $\epsilon > 0$ such that $\tilde{\mathbf{x}} \in A$ implies $\|\tilde{\mathbf{x}}\|_\infty > \epsilon$. Therefore, for measures $\mu_n, \mu \in \mathcal{M}(\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$, $\mu_n \rightarrow_{w^\#} \mu$ if $\mu_n(f) \rightarrow \mu(f)$ for every bounded and continuous function f on $\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$ such that for some $\epsilon > 0$, $\|\tilde{\mathbf{x}}\|_\infty \leq \epsilon$ implies $f(\tilde{\mathbf{x}}) = 0$.

Remark 3.2 We note that under the metric \tilde{d}' , $w^\#$ -convergence coincides with the theory of M_0 -convergence introduced in [18], further developed in [22] and with the corresponding point process convergence recently studied by [35].

Take now a sequence $\{r_n\}$ as in Assumption 1.1, set $k_n = \lfloor n/r_n \rfloor$ and define

$$X_{n,i} = (X_{(i-1)r_n+1}, \dots, X_{ir_n})/a_n$$

for $i = 1, \dots, k_n$. As the main result of this section we show, under certain conditions, the point process of clusters N''_n defined by

$$N''_n = \sum_{i=1}^{k_n} \delta_{(i/k_n, X_{n,i})}$$

restricted to $[0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$ (i.e. we ignore indices i with $X_{n,i} = \mathbf{0}$), converges in distribution in $\mathcal{M}_p([0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$ to a suitable Poisson point process.

We first prove a technical lemma which is also of independent interest, see Remark 3.4. Denote by $\mathbb{S} = \{\tilde{\mathbf{x}} \in \tilde{l}_0 : \|\tilde{\mathbf{x}}\|_\infty = 1\}$ the unit sphere in \tilde{l}_0 and define the polar decomposition $\psi : \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\} \mapsto (0, +\infty) \times \mathbb{S}$ with $\psi(\tilde{\mathbf{x}}) = (\|\tilde{\mathbf{x}}\|_\infty, \tilde{\mathbf{x}}/\|\tilde{\mathbf{x}}\|_\infty)$.

Lemma 3.3 *Under Assumption 1.1, the sequence $\nu_n = k_n \mathbb{P}(X_{n,1} \in \cdot)$ in $\mathcal{M}(\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$ converges to $\nu = \theta (d(-y^{-\alpha}) \times \mathbb{P}_{\mathcal{Q}}) \circ \psi$ in $w^\#$ -topology and $\mathbb{P}_{\mathcal{Q}}$ is the distribution of $\{Q_j\}$ defined in (2.2).*

Proof Let f be a bounded and continuous function on $\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$ and $\epsilon > 0$ such that $f(\tilde{\mathbf{x}}) = 0$ if $\|\tilde{\mathbf{x}}\|_\infty \leq \epsilon$. Then $\mathbb{E}[f(X_{n,1})] = \mathbb{E}[f(X_{n,1})\mathbb{1}_{\{M_{1,r_n} > \epsilon a_n\}}]$, so by (1.3), (2.7) and Theorem 2.2 we get

$$\begin{aligned} \lim_{n \rightarrow \infty} \nu_n(f) &= \lim_{n \rightarrow \infty} k_n \mathbb{E}[f(X_{n,1})] \\ &= \lim_{n \rightarrow \infty} n \mathbb{P}(\|X_0\| > \epsilon a_n) \frac{\mathbb{P}(M_{1,r_n} > \epsilon a_n)}{r_n \mathbb{P}(\|X_0\| > \epsilon a_n)} \mathbb{E}[f(X_{n,1}) \mid M_{1,r_n} > \epsilon a_n] \\ &= \epsilon^{-\alpha} \theta \mathbb{E}[f(\epsilon \mathbf{Z})]. \end{aligned}$$

Applying Theorem 2.2, the last expression is equal to

$$\epsilon^{-\alpha} \theta \int_1^\infty \mathbb{E}[f(\epsilon y \mathcal{Q})] \alpha y^{-\alpha-1} dy = \theta \int_\epsilon^\infty \mathbb{E}[f(y \mathcal{Q})] \alpha y^{-\alpha-1} dy.$$

Finally, since $\|Q\|_\infty = 1$ a.s. and $f(\tilde{\mathbf{x}}) = 0$ if $\|\tilde{\mathbf{x}}\|_\infty \leq \epsilon$ we have that

$$\lim_{n \rightarrow \infty} \nu_n(f) = \theta \int_0^\infty \mathbb{E}[f(y \mathcal{Q})] \alpha y^{-\alpha-1} dy = \nu(f),$$

by definition of ν . □

Remark 3.4 The previous lemma is closely related to the large deviations result obtained in Mikosch and Wintenberger [28, Theorem 3.1]. For a class of functions f

called cluster functionals, which can be directly linked to the functions we used in the proof of Lemma 3.3, they showed that

$$\begin{aligned} &\lim_{n \rightarrow \infty} \frac{\mathbb{E}[f(a_n^{-1}X_1, \dots, a_n^{-1}X_{r_n})]}{r_n \mathbb{P}(\|X_0\| > a_n)} \\ &= \int_0^\infty \mathbb{E}[f(y\{\Theta_t, t \geq 0\}) - f(y\{\Theta_t, t \geq 1\})] \alpha y^{-\alpha-1} dy. \end{aligned} \tag{3.1}$$

However, for an arbitrary bounded measurable function $f : \tilde{l}_0 \rightarrow \mathbb{R}$ which is a.e. continuous with respect to ν and such that for some $\epsilon > 0$, $\|\tilde{x}\|_\infty \leq \epsilon$ implies $f(\tilde{x}) = 0$, Lemma 3.3 together with a continuous mapping argument and the fact that $k_n^{-1} \sim r_n \mathbb{P}(\|X_0\| > a_n)$ yields

$$\lim_{n \rightarrow \infty} \frac{\mathbb{E}[f(a_n^{-1}X_1, \dots, a_n^{-1}X_{r_n})]}{r_n \mathbb{P}(\|X_0\| > a_n)} = \lim_{n \rightarrow \infty} \nu_n(f) = \theta \int_0^\infty \mathbb{E}[f(y\mathcal{Q})] \alpha y^{-\alpha-1} dy,$$

which gives an alternative and arguably more interpretable expression for the limit in (3.1).

To show convergence of N''_n in $\mathcal{M}_p([0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$, we will need to assume that, intuitively speaking, one can break the dependent series $\{X_n, n \geq 1\}$ into asymptotically independent blocks.

Recall the notion of truncation of an element $\tilde{x} \in \tilde{l}_0$ at level $\epsilon > 0$ denoted by \tilde{x}_ϵ , see paragraph following (2.3). We denote by \mathcal{F}_+ the class of nonnegative functions f on $[0, 1] \times \tilde{l}_0$ which satisfy $f(t, \tilde{\mathbf{0}}) = 0$ and $f(t, \tilde{x}) = f(t, \tilde{x}_\epsilon)$ for some $\epsilon > 0$, and are continuous except maybe on the set $\{\tilde{x} \in \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\} : \|x_j\| = \epsilon \text{ for some } j \in \mathbb{Z} \text{ where } (x_i)_{i \in \mathbb{Z}} \in \tilde{x}\}$. Thus, $f \in \mathcal{F}_+$ depend only on coordinates greater in norm than ϵ . Using similar arguments as in the proof of Theorem 2.2, one can show that \mathcal{F}_+ is convergence determining (in the sense of Remark 3.1).

Assumption 3.5 There exists a sequence of integers $\{r_n, n \in \mathbb{N}\}$ such that $\lim_{n \rightarrow \infty} r_n = \lim_{n \rightarrow \infty} n/r_n = \infty$ and

$$\lim_{n \rightarrow \infty} \left(\mathbb{E}[e^{-N''_n(f)}] - \prod_{i=1}^{k_n} \mathbb{E} \left[\exp \left\{ -f \left(\frac{i}{k_n}, X_{n,i} \right) \right\} \right] \right) = 0,$$

for all $f \in \mathcal{F}_+$.

This assumption is somewhat stronger than the related conditions in [14] or [5], cf. Condition 2.2 in the latter paper, since we consider functions of the whole cluster. Still, as we show in Lemma 6.2, β -mixing implies Assumption 3.5. Since sufficient conditions for β -mixing are well studied and hold for many standard time series models (linear processes are notable exception, note), one can avoid cumbersome task of checking the assumption above directly. Linear processes are considered separately in Sect. 3.3.

It turns out that the choice of \tilde{l}_0 as the state space for clusters, together with the results above, does not only preserve the order of the observations within the cluster, but also makes the statement and the proof of the following theorem remarkably tidy and short.

Theorem 3.6 *Let $\{X_t, t \in \mathbb{Z}\}$ be a stationary \mathbb{R}^d -valued regularly varying sequence with tail index $\alpha > 0$, satisfying Assumption 1.1 and 3.5 for the same sequence $\{r_n\}$. Assumption Then $N''_n \xrightarrow{d} N''$ in $\mathcal{M}_p([0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$ where N'' is a Poisson point process with intensity measure $Leb \times \nu$ which can be expressed as*

$$N'' = \sum_{i=1}^{\infty} \delta_{(T_i, P_i, \mathbf{Q}_i)}, \tag{3.2}$$

where

- (i) $\sum_{i=1}^{\infty} \delta_{(T_i, P_i)}$ is a Poisson point process on $[0, 1] \times (0, \infty]$ with intensity measure $Leb \times d(-\theta y^{-\alpha})$;
- (ii) $\{\mathbf{Q}_i, i \geq 1\}$ is a sequence of i.i.d. elements in \mathbb{S} , independent of $\sum_{i=1}^{\infty} \delta_{(T_i, P_i)}$ and with common distribution equal to the distribution of \mathbf{Q} in (2.2).

Proof Let, for every $n \in \mathbb{N}$, $\{\mathbf{X}_{n,i}^*, i = 1, \dots, k_n\}$ be independent copies of $\mathbf{X}_{n,1}$ and define

$$N_n^* = \sum_{i=1}^{k_n} \delta_{(i/k_n, \mathbf{X}_{n,i}^*)}. \tag{3.3}$$

Since by the previous lemma, $k_n \mathbb{P}(\mathbf{X}_{n,1} \in \cdot) \rightarrow_{w^\#} \nu$ in $\mathcal{M}(\tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$, the convergence of N_n^* to N'' in $\mathcal{M}_p([0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$ follows by an straightforward adaptation of [30, Proposition 3.21] (cf. [15, Lemma 2.2.(1)], see also [16, Theorem 2.4] or [33, Proposition 2.13]). Assumption 3.5 now yields that N''_n converge in distribution to the same limit since the convergence determining family \mathcal{F}_+ consists of functions which are a.e. continuous with respect to the measure $Leb \times \nu$. Finally, the representation of N'' follows easily by standard Poisson point process transformation argument (see [30, Section 3.3.2.]). □

Under the conditions of Theorem 3.6, as already noticed in [6, Remark 4.7], $\theta = \mathbb{P}(\sup_{i \geq 1} \|Y_i\| \leq 1)$ is also the extremal index of the time series $\{\|X_j\|\}$, i.e. $\lim_{n \rightarrow \infty} \mathbb{P}(a_n^{-1} \max_{1 \leq i \leq n} \|X_i\| \leq x) = e^{-\theta x^{-\alpha}}$ for all $x > 0$.

Remark 3.7 Note that the restriction to the time interval $[0, 1]$ is arbitrary and it could be substituted by an arbitrary closed interval $[0, T]$.

3.3 Linear processes

As we observed above, β -mixing offers a way of establishing Assumption 3.5 for a wide class of time series models. However, for linear processes, the truncation method offers an alternative and simpler way to obtain the point process convergence stated in the previous theorem.

Let $\{\xi_t, t \in \mathbb{Z}\}$ be a sequence of i.i.d. random variables with regularly varying distribution with index $\alpha > 0$. Consider the linear process $\{X_t, t \in \mathbb{Z}\}$ defined by

$$X_t = \sum_{j \in \mathbb{Z}} c_j \xi_{t-j},$$

where $\{c_j, j \in \mathbb{Z}\}$ is a sequence of real numbers such that $|c_0| > 0$ and

$$\sum_{j \in \mathbb{Z}} |c_j|^\delta < \infty \text{ with } \begin{cases} \delta < \alpha & \text{if } \alpha \in (0, 1], \\ \delta < \alpha & \text{if } \alpha \in (1, 2] \text{ and } \mathbb{E}[\xi_0] = 0, \\ \delta = 2 & \text{if } \alpha > 2 \text{ and } \mathbb{E}[\xi_0] = 0, \\ \delta = 1 & \text{if } \alpha > 1 \text{ and } \mathbb{E}[\xi_0] \neq 0. \end{cases} \tag{3.4}$$

These conditions imply that $\sum_{j \in \mathbb{Z}} |c_j|^\alpha < \infty$. Furthermore, it has been proved in [25, Lemma A3] that the sequence $\{X_t, t \in \mathbb{Z}\}$ is well defined, stationary and regularly varying with tail index α and

$$\lim_{u \rightarrow \infty} \frac{\mathbb{P}(|X_0| > u)}{\mathbb{P}(|\xi_0| > u)} = \sum_{j \in \mathbb{Z}} |c_j|^\alpha. \tag{3.5}$$

[25, Lemma A.4] proved that for $\alpha \in (0, 2]$, it is possible to take $\delta = \alpha$ in (3.4) at the cost of some restrictions on the distribution of ξ_0 which are satisfied for Pareto and stable distributions.

The spectral tail process $\{\Theta_t\}$ of the linear process was computed in [23]. It can be described as follows: let Θ^ξ be an $\{-1, 1\}$ -valued random variable with distribution equal to the spectral measure of ξ_0 . Then

$$\mathcal{L}(\{\Theta_t, t \in \mathbb{Z}\}) = \mathcal{L}\left(\left\{\frac{c_{t+K}}{|c_K|} \Theta^\xi, t \in \mathbb{Z}\right\}\right), \tag{3.6}$$

where K is an integer valued random variable, independent of Θ^ξ , such that

$$\mathbb{P}(K = n) = \frac{|c_n|^\alpha}{\sum_{j \in \mathbb{Z}} |c_j|^\alpha}, \quad n \in \mathbb{Z}. \tag{3.7}$$

It is also proved in [23] that the coefficient θ from (1.9) is given by

$$\theta = \frac{\max_{j \in \mathbb{Z}} |c_j|^\alpha}{\sum_{j \in \mathbb{Z}} |c_j|^\alpha}.$$

Since $\lim_{|j| \rightarrow \infty} c_j = 0$, random element $\mathbf{Q} = \{Q_j, j \in \mathbb{Z}\}$ in the space \mathbb{S} introduced in (2.2) is well defined and given by

$$Q = \left\{ \frac{\Theta^\xi c_j}{\max_{i \in \mathbb{Z}} |c_i|}, j \in \mathbb{Z} \right\}. \tag{3.8}$$

The following proposition can be viewed as an extension of [13, Theorem 2.4] and also as a version of Theorem 3.6 adapted to linear processes.

Proposition 3.8 *Let $\{r_n\}$ be a nonnegative non decreasing integer valued sequence such that $\lim_{n \rightarrow \infty} r_n = \lim_{n \rightarrow \infty} n/r_n = \infty$ and let $\{a_n\}$ be a non decreasing sequence such that $n\mathbb{P}(|X_0| > a_n) \rightarrow 1$. Then*

$$N''_n = \sum_{i=1}^{k_n} \delta_{(i/k_n, (X_{(i-1)r_{n+1}}, \dots, X_{ir_n})/a_n)} \xrightarrow{d} \sum_{i=1}^{\infty} \delta_{(T_i, P_i, Q_i)} \tag{3.9}$$

in $\mathcal{M}_p([0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$ where $\sum_{i=1}^{\infty} \delta_{(T_i, P_i)}$ is a Poisson point process on $[0, 1] \times (0, \infty)$ with intensity measure $Leb \times d(-\theta y^{-\alpha})$, independent of the i.i.d. sequence $\{Q_i, i \geq 1\}$ with values in \mathbb{S} and common distribution equal to the distribution of Q in (3.8).

Proof The proof of (3.11) is based on a truncation argument which compares X_t with

$$X_t^{(m)} = \sum_{j=-m}^m c_j \xi_{t-j}, \quad t \in \mathbb{Z}.$$

Let $\{b_n\}$ and $\{a_{m,n}\}$ be non decreasing sequences such that $n\mathbb{P}(|Z_0| > b_n) \rightarrow 1$ and $n\mathbb{P}(|X_0^{(m)}| > a_{m,n}) \rightarrow 1$. The limit (3.5) implies that $a_{m,n} \sim \left(\sum_{j=-m}^m |c_j|^\alpha\right)^{1/\alpha} b_n$. Let $N''_{m,n}$ be the point process of exceedences of the truncated sequence defined by

$$N''_{m,n} = \sum_{i=1}^{k_n} \delta_{(i/k_n, (X_{(i-1)r_{n+1}}^{(m)}, \dots, X_{ir_n}^{(m)})/b_n)}.$$

The process $\{X_t^{(m)}\}$ is $(2m)$ -dependent (hence β -mixing with $\beta_j = 0$ for $j > 2m$) and therefore satisfies the conditions of Theorem 3.6. Thus $N''_{m,n} \xrightarrow{d} N''_{(m)}$ with

$$N''_{(m)} = \sum_{i=1}^{\infty} \delta_{(T_i, P_i, \Theta_i^\xi \{c_j^{(m)}\})},$$

with $\sum_{i=1}^{\infty} \delta_{(T_i, P_i)}$ a Poisson point process on $[0, 1] \times (0, \infty)$ with mean measure $Leb \times d(-y^{-\alpha})$, $c_j^{(m)} = c_j$ if $|j| \leq m$ and $c_j = 0$ otherwise, and $\Theta_i^\xi, i \geq 1$ are i.i.d. copies of Θ^ξ . Since $\{c_j^{(m)}\}$ converges to $\{c_j\}$ in \tilde{l}_0 , it follows that

$$N''_{(m)} \rightarrow N''_{\infty} = \sum_{i=1}^{\infty} \delta_{(T_i, P_i \Theta_i^{\xi}(c_j))},$$

almost surely in $\mathcal{M}_p([0, 1] \times \tilde{l}_0 \setminus \{\mathbf{0}\})$.

Define now

$$N''_{\infty, n} = \sum_{i=1}^{k_n} \delta_{(i/k_n, (X_{(i-1)r_{n+1}}, \dots, X_{ir_n})/b_n)}.$$

Then, for every bounded Lipschitz continuous function f defined on $[0, 1] \times \tilde{l}_0 \setminus \{\mathbf{0}\}$ with bounded support,

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(|N''_{n,m}(f) - N''_{\infty,n}(f)| > \eta) = 0.$$

As in the proof of [13, Theorem 2.4], this follows from

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\max_{1 \leq i \leq n} |X_i^{(m)} - X_i| > b_n \gamma) = 0. \tag{3.10}$$

for all $\gamma > 0$ which is implied by (3.5); see [13, Lemma 2.3].

This proves that

$$N''_{\infty, n} \xrightarrow{d} N''_{\infty}. \tag{3.11}$$

and since N''_n and $N''_{\infty, n}$ differ only by a deterministic scaling of the points, this proves our result. □

4 Convergence of the partial sum process

In order to study the convergence of the partial sum process in cases where it fails to hold in the usual space D , we first introduce an enlarged space E . In the rest of the paper we restrict to the case of \mathbb{R} -valued time series.

4.1 The space of decorated càdlàg functions

To establish convergence of the partial sum process of a dependent sequence $\{X_n\}$ we will consider the function space $E \equiv E([0, 1], \mathbb{R})$ introduced in Whitt [34, Sections 15.4 and 15.5]. For the benefit of the reader, in what follows, we briefly introduce this space closely following the exposition in the previously mentioned reference.

The elements of E have the form

$$(x, J, \{I(t) : t \in J\})$$

where

- $x \in D([0, 1], \mathbb{R})$;
- J is a countable subset of $[0, 1]$ with $Disc(x) \subseteq J$, where $Disc(x)$ is the set of discontinuities of the càdlàg function x ;
- for each $t \in J$, $I(t)$ is a closed bounded interval (called the decoration) in \mathbb{R} such that $x(t), x(t-) \in I(t)$ for all $t \in J$.

Moreover, we assume that for each $\epsilon > 0$, there are at most finitely many times t for which the length of the interval $I(t)$ is greater than ϵ . This ensures that the graphs of elements in E , defined below, are compact subsets of \mathbb{R}^2 which allows one to impose a metric on E by using the Hausdorff metric on the space of graphs of elements in E .

Note that every triple $(x, J, \{I(t) : t \in J\})$ can be equivalently represented by a set-valued function

$$x'(t) := \begin{cases} I(t) & \text{if } t \in J, \\ \{x(t)\} & \text{if } t \notin J, \end{cases}$$

or by the graph of x' defined by

$$\Gamma_{x'} := \{(t, z) \in [0, 1] \times \mathbb{R} : z \in x'(t)\}.$$

In the sequel, we will usually denote the elements of E by x' .

Let m denote the Hausdorff metric on the space of compact subsets of \mathbb{R}^d (regardless of dimension) i.e. for compact subsets A, B ,

$$m(A, B) = \sup_{x \in A} \|x - B\|_\infty \vee \sup_{y \in B} \|y - A\|_\infty,$$

where $\|x - B\|_\infty = \inf_{y \in B} \|x - y\|_\infty$. We then define a metric on E , denoted by m_E , by

$$m_E(x', y') = m(\Gamma_{x'}, \Gamma_{y'}). \tag{4.1}$$

We call the topology induced by m_E on E the M_2 topology. This topology is separable, but the metric space (E, m_E) is not complete. Also, we define the uniform metric on E by

$$m^*(x', y') = \sup_{0 \leq t \leq 1} m(x'(t), y'(t)), \tag{4.2}$$

Obviously, m^* is a stronger metric than m_E , i.e. for any $x', y' \in E$,

$$m_E(x', y') \leq m^*(x', y'). \tag{4.3}$$

We will often use the following elementary fact: for $a \leq b$ and $c \leq d$ it holds that

$$m([a, b], [c, d]) \leq |c - a| \vee |d - b|. \tag{4.4}$$

By a slight abuse of notation, we identify every $x \in D$ with an element in E represented by

$$(x, Disc(x), \{[x(t-), x(t)] : t \in Disc(x)\}),$$

where for any two real numbers a, b by $[a, b]$ we denote the closed interval $[\min\{a, b\}, \max\{a, b\}]$. Consequently, we identify the space D with the subset D' of E given by

$$D' = \{x' \in E : J = \text{Disc}(x) \text{ and for all } t \in J, I(t) = [x(t-), x(t)]\}. \tag{4.5}$$

For an element $x' \in D'$ we have

$$\Gamma_{x'} = \Gamma_x,$$

where Γ_x is the completed graph of x . Since the M_2 topology on D corresponds to the Hausdorff metric on the space of the completed graphs Γ_x , the map $x \rightarrow (x, \text{Disc}(x), \{[x(t-), x(t)] : t \in \text{Disc}(x)\})$ is a homeomorphism from D endowed with the M_2 topology onto D' endowed with the M_2 topology. This yields the following lemma.

Lemma 4.1 *The space D endowed with the M_2 topology is homeomorphic to the subset D' in E with the M_2 topology.*

Remark 4.2 Because two elements in E can have intervals at the same time point, addition in E is in general not well behaved. However, problems disappear if one of the summands is a continuous function. In such a case, the sum is naturally defined as follows: consider an element $x' = (x, J, \{I(t) : t \in J\})$ in E and a continuous function b on $[0, 1]$, we define the element $x' + b$ in E by

$$x' + b = (x + b, J, \{I(t) + b(t) : t \in J\}).$$

We now state a useful characterization of convergence in (E, m_E) in terms of the local-maximum function defined for any $x' \in E$ by

$$M_{t_1, t_2}(x') := \sup\{z : z \in x'(t), t_1 \leq t \leq t_2\}, \tag{4.6}$$

for $0 \leq t_1 < t_2 \leq 1$.

Theorem 4.3 (Theorem 15.5.1 Whitt [34]) *For elements $x'_n, x' \in E$ the following are equivalent:*

- (i) $x'_n \rightarrow x'$ in (E, m_E) , i.e. $m_E(x'_n, x') \rightarrow 0$.
- (ii) For all $t_1 < t_2$ in a countable dense subset of $[0, 1]$, including 0 and 1,

$$M_{t_1, t_2}(x'_n) \rightarrow M_{t_1, t_2}(x') \text{ in } \mathbb{R}$$

and

$$M_{t_1, t_2}(-x'_n) \rightarrow M_{t_1, t_2}(-x') \text{ in } \mathbb{R}.$$

4.2 Invariance principle in the space E

Consider the partial sum process in $D([0, 1])$ defined by

$$S_n(t) = \sum_{i=1}^{\lfloor nt \rfloor} \frac{X_i}{a_n}, \quad t \in [0, 1],$$

and define also

$$V_n(t) = \begin{cases} S_n(t) & \text{if } 0 < \alpha < 1, \\ S_n(t) - \lfloor nt \rfloor \mathbb{E} \left(\frac{X_1}{a_n} \mathbb{1}_{\{|X_1|/a_n \leq 1\}} \right) & \text{if } 1 \leq \alpha < 2. \end{cases}$$

As usual, when $1 \leq \alpha < 2$, an additional condition is needed to deal with the small jumps.

Assumption 4.4 For all $\delta > 0$,

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P} \left(\max_{1 \leq k \leq n} \left| \sum_{i=1}^k \{X_i \mathbb{1}_{\{|X_i| \leq a_n \epsilon\}} - \mathbb{E}[X_i \mathbb{1}_{\{|X_i| \leq a_n \epsilon\}}]\} \right| > a_n \delta \right) = 0. \quad (4.7)$$

It is known from [14] that the finite dimensional marginal distributions of V_n converge to those of an α -stable Lévy process. This result is strengthened in [5] to convergence in the M_1 topology if $Q_j Q_{j'} \geq 0$ for all $j \neq j' \in \mathbb{Z}$, i.e. if all extremes within one cluster have the same sign. In the next theorem, we remove the latter restriction and establish the convergence of the process V_n in the space E .

For that purpose, we assume only regular variation of the sequence $\{X_t, t \in \mathbb{Z}\}$ and the conclusion of Theorem 3.6, i.e.

$$N''_n \xrightarrow{d} N'' = \sum_{i=1}^{\infty} \delta_{(T_i, P_i, Q_i)} \quad (4.8)$$

in $\mathcal{M}_p([0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$, where $\sum_{i=1}^{\infty} \delta_{(T_i, P_i)}$ is a Poisson point process on $[0, 1] \times (0, \infty]$ with intensity measure $Leb \times d(-\theta y^{-\alpha})$ with $\theta > 0$ and $Q_i = \{Q_{i,j}, j \in \mathbb{Z}\}$, $i \geq 1$ are i.i.d. sequences in \tilde{l}_0 , independent of $\sum_{i=1}^{\infty} \delta_{(T_i, P_i)}$ and such that $\mathbb{P}(\sup_{j \in \mathbb{Z}} |Q_{1,j}| = 1) = 1$. Denote by $Q = \{Q_j, j \in \mathbb{Z}\}$ a random sequence the with the distribution equal to the distribution of Q_1 . We also describe the limit of V_n in terms of the point process N'' .

The convergence (4.8) and Fatou’s lemma imply that $\theta \sum_{j \in \mathbb{Z}} \mathbb{E}[|Q_j|^\alpha] \leq 1$, as originally noted by [14, Theorem 2.6]. This implies that for $\alpha \in (0, 1]$,

$$\mathbb{E} \left[\left(\sum_{j \in \mathbb{Z}} |Q_j| \right)^\alpha \right] < \infty. \quad (4.9)$$

For $\alpha > 1$, this will have to be assumed. Furthermore, the case $\alpha = 1$, as usual, requires additional care. We will assume that

$$\mathbb{E} \left[\sum_{j \in \mathbb{Z}} |Q_j| \log \left(|Q_j|^{-1} \sum_{i \in \mathbb{Z}} |Q_i| \right) \right] < \infty, \tag{4.10}$$

where we use the convention $|Q_j| \log \left(|Q_j|^{-1} \sum_{i \in \mathbb{Z}} |Q_i| \right) = 0$ if $|Q_j| = 0$. Fortunately, it turns out that conditions (4.9) and (4.10) are satisfied in most examples. See Remark 4.8 below.

Theorem 4.5 *Let $\{X_t, t \in \mathbb{Z}\}$ be a stationary \mathbb{R} -valued regularly varying sequence with tail index $\alpha \in (0, 2)$ and assume that the convergence in (4.8) holds. If $\alpha \geq 1$ let Assumption 4.4 hold. For $\alpha > 1$, assume that (4.9) holds, and for $\alpha = 1$, assume that (4.10) holds. Then*

$$V_n \xrightarrow{d} V' = (V, \{T_i\}_{i \in \mathbb{N}}, \{I(T_i)\}_{i \in \mathbb{N}}),$$

with respect to M_2 topology on $E([0, 1], \mathbb{R})$, where

(i) V is an α -stable Lévy process on $[0, 1]$ given by

$$V(\cdot) = \sum_{T_i \leq \cdot} \sum_{j \in \mathbb{Z}} P_i Q_{i,j}, \quad 0 < \alpha < 1, \tag{4.11a}$$

$$V(\cdot) = \lim_{\epsilon \rightarrow 0} \left(\sum_{T_i \leq \cdot} \sum_{j \in \mathbb{Z}} P_i Q_{i,j} \mathbb{1}_{\{|P_i Q_{i,j}| > \epsilon\}} - (\cdot) \int_{\epsilon < |x| \leq 1} x \mu(dx) \right), \quad 1 \leq \alpha < 2, \tag{4.11b}$$

where the series in (4.11a) is almost surely absolutely summable and the holds uniformly on $[0, 1]$ almost surely (along some subsequence) with μ given in (1.5).

(ii) For all $i \in \mathbb{N}$,

$$I(T_i) = V(T_i-) + P_i \left[\inf_{k \in \mathbb{Z}} \sum_{j \leq k} Q_{i,j}, \sup_{k \in \mathbb{Z}} \sum_{j \leq k} Q_{i,j} \right].$$

Before proving the theorem, we make several remarks. We first note that for $\alpha < 1$, convergence of the point process is the only assumption of theorem. Further, an extension of Theorem 4.5 to multivariate regularly varying sequences would be possible at the cost of various technical issues (similar to those in [34, Section 12.3] for the extension of M_1 and M_2 topologies to vector valued processes) and one would moreover need to alter the definition of the space E substantially and introduce a new and weaker notion of M_2 topology.

Remark 4.6 If (4.9) holds, the sums $W_i = \sum_{j \in \mathbb{Z}} |Q_{i,j}|$ are almost surely well-defined and $\{W_i, i \geq 1\}$ is a sequence of i.i.d. random variables with $\mathbb{E}[W_i^\alpha] < \infty$. Furthermore, by independence of $\sum_{i=1}^\infty \delta_{P_i}$ and $\{W_i\}$, it follows that $\sum_{i=1}^\infty \delta_{P_i W_i}$ is a Poisson point process on $(0, \infty]$ with intensity measure $\theta \mathbb{E}[W_1^\alpha] \alpha y^{-\alpha-1} dy$. In particular, for every $\delta > 0$ there a.s. exist at most finitely many indices i such that $P_i W_i > \delta$. Also, this implies that

$$\sup_{i \in \mathbb{N}} \sum_{j \in \mathbb{Z}} P_i |Q_{i,j}| \mathbb{1}_{\{|P_i |Q_{i,j}| \leq \epsilon\}} \rightarrow 0 \tag{4.12}$$

almost surely as $\epsilon \rightarrow 0$. These facts will be used several times in the proof.

Remark 4.7 The Lévy process V from Theorem 4.5 is the weak limit in the sense of finite dimensional distributions of the partial sum process V_n , characterized by

$$\begin{aligned} & \log \mathbb{E}[e^{izV(1)}] \\ &= \begin{cases} iaz + \Gamma(1 - \alpha) \cos(\pi\alpha/2) \sigma^\alpha |z|^\alpha \{1 - i\phi \operatorname{sgn}(z) \tan(\pi\alpha/2)\} & \alpha \neq 1, \\ iaz - \frac{\pi}{2} \sigma |z| \{1 - i\frac{2}{\pi} \phi \operatorname{sgn}(z) \log(|z|)\} & \alpha = 1, \end{cases} \end{aligned} \tag{4.13}$$

with, denoting $x^{(\alpha)} = x|x|^{\alpha-1} = x_+^\alpha - x_-^\alpha$,

$$\sigma^\alpha = \theta \mathbb{E} \left[\left| \sum_{j \in \mathbb{Z}} Q_j \right|^\alpha \right], \quad \phi = \frac{\mathbb{E}[(\sum_{j \in \mathbb{Z}} Q_j)^{(\alpha)}]}{\mathbb{E}[|\sum_{j \in \mathbb{Z}} Q_j|^\alpha]}$$

and

- (i) $a = 0$ if $\alpha < 1$;
- (ii) $a = (\alpha - 1)^{-1} \alpha \theta \mathbb{E} \left[\sum_{j \in \mathbb{Z}} Q_j^{(\alpha)} \right]$ if $\alpha > 1$;
- (iii) if $\alpha = 1$, then

$$\begin{aligned} a = \theta & \left(c_0 \mathbb{E} \left[\sum_{j \in \mathbb{Z}} Q_j \right] - \mathbb{E} \left[\sum_{j \in \mathbb{Z}} Q_j \log \left(\left| \sum_{j \in \mathbb{Z}} Q_j \right| \right) \right] \right. \\ & \left. - \mathbb{E} \left[\sum_{j \in \mathbb{Z}} Q_j \log \left(|Q_j|^{-1} \right) \right] \right), \end{aligned}$$

with $c_0 = \int_0^\infty (\sin y - y \mathbb{1}_{(0,1]}(y)) y^{-2} dy$.

These parameters were computed in [14, Theorem 3.2] but with a complicated expression for the location parameter a in the case $\alpha = 1$ (see [14, Remark 3.3]). The explicit expression given here, which holds under the assumption (4.10), is new; it is obtained by direct calculation of the characteristic function of $V(1)$ following the steps of the proof of Lemma 6.4. As often done in the literature, if the sequence is assumed to be symmetric then assumption (4.10) is not needed and the location parameter is 0.

Remark 4.8 During the revision process for this paper, [29] showed that the quantity θ from (1.9) is positive whenever $\mathbb{P}(\lim_{|t| \rightarrow \infty} |Y_t| = 0) = 1$, and $\theta = \mathbb{P}(\sup_{i \leq -1} |Y_i| \leq 1)$. In particular, the sequence \mathbf{Q} from (2.2) is well defined in this case and moreover, by [29, Lemma 3.11], the condition (4.9) turns out to be equivalent to

$$\mathbb{E} \left[\left(\sum_{j=0}^{\infty} |\Theta_j| \right)^{\alpha-1} \right] < \infty \tag{4.14}$$

which is automatic if $\alpha \in (0, 1]$. Furthermore, if $\alpha = 1$, [29, Lemma 3.14] shows that the condition (4.10) is then equivalent to

$$\mathbb{E} \left[\log \left(\sum_{j=0}^{\infty} |\Theta_j| \right) \right] < \infty. \tag{4.15}$$

These conditions are easier to check than conditions (4.9) and (4.10) since it is easier to determine the distribution of the spectral tail process than the distribution of the process \mathbf{Q} from (2.2). In fact, it suffices to determine only the distribution of the forward spectral tail process $\{\Theta_j, j \geq 0\}$ which is often easier than determining the distribution of the whole spectral tail process. For example, it follows from the proof of [27, Theorem 3.2] that for functions of Markov chains satisfying a suitable drift condition, (4.14) and (4.15) hold for all $\alpha > 0$. Also, notice that for the linear process $\{X_t\}$ from Sect. 3.3 these conditions are satisfied if $\sum_{j \in \mathbb{Z}} |c_j| < \infty$.

Moreover, [29, Corollary 3.12 and Lemma 3.14] imply that the scale, skewness and location parameters from Remark 4.7 can also be expressed in terms of the forward spectral tail process as follows:

$$\begin{aligned} \sigma^\alpha &= \mathbb{E} \left[\left| \sum_{j=0}^{\infty} \Theta_j \right|^\alpha - \left| \sum_{j=1}^{\infty} \Theta_j \right|^\alpha \right], \\ \phi &= \sigma^{-\alpha} \mathbb{E} \left[\left(\sum_{j=0}^{\infty} \Theta_j \right)^{(\alpha)} - \left(\sum_{j=1}^{\infty} \Theta_j \right)^{(\alpha)} \right], \end{aligned}$$

$a = 0$ if $\alpha < 1$, $a = (\alpha - 1)^{-1} \alpha \mathbb{E}[\Theta_0]$ if $\alpha > 1$ (see (6.8)) and

$$a = c_0 \mathbb{E}[\Theta_0] - \mathbb{E} \left[\sum_{j=0}^{\infty} \Theta_j \log \left(\left| \sum_{j=0}^{\infty} \Theta_j \right| \right) - \sum_{j=1}^{\infty} \Theta_j \log \left(\left| \sum_{j=1}^{\infty} \Theta_j \right| \right) \right],$$

if $\alpha = 1$. It can be shown that these expressions coincide for $\alpha \neq 1$ with those in the literature, see e.g. [28, Theorem 4.3]. As already noted, the expression of the location parameter for $\alpha = 1$ under the assumption (4.10) (or (4.15)) is new.

Example 4.9 Consider again the linear process $\{X_t\}$ from Sect. 3.3. For infinite order moving average processes, [13] proved convergence of the finite dimensional distributions of the partial sum process; [2] proved the functional convergence in the M_1 topology (see [34, Section 12.3]) when $c_j \geq 0$ for all j ; using the S topology (which is weaker than the M_1 topology and makes the supremum functional not continuous), [3] proved the corresponding result under more general conditions on the sequence $\{c_j\}$ in the case $\alpha \leq 1$.

Our Theorem 4.5 directly applies to the case of a finite order moving average process. To consider the case of an infinite order moving average process, we assume for simplicity that $\alpha < 1$. Applying Theorem 4.5 to the point process convergence in (3.9), one obtains the convergence of the partial sum process $V_n \xrightarrow{d} V' = (V, \{T_i\}_{i \in \mathbb{N}}, \{I(T_i)\}_{i \in \mathbb{N}})$ in $(E, [0, 1])$ where

$$V(\cdot) = \frac{\sum_{j \in \mathbb{Z}} c_j}{\max_{j \in \mathbb{Z}} |c_j|} \sum_{T_i \leq \cdot} P_i \Theta_i^\xi,$$

and

$$I(T_i) = V(T_i-) + \frac{P_i \Theta_i^\xi}{\max_{j \in \mathbb{Z}} |c_j|} \left[\inf_{k \in \mathbb{Z}} \sum_{j \leq k} c_j, \sup_{k \in \mathbb{Z}} \sum_{j \leq k} c_j \right].$$

For an illustration consider the process

$$X_t = \xi_t + c\xi_{t-1}.$$

In the case $c \geq 0$, the convergence of partial sum process in M_1 topology follows from [2]. On the other hand, for negative c 's convergence fails in any of Skorohod's topology, but partial sums do have a limit in the sense described by our theorem as can be also guessed from Fig. 1.

Remark 4.10 We do not exclude the case $\sum_{j \in \mathbb{Z}} Q_j = 0$ with probability one, as happens for instance in Example 4.9 with $c = -1$. In such a case, the càdlàg component V is simply the null process.

Example 4.11 Consider a stationary GARCH(1, 1) process

$$X_t = \sigma_t Z_t, \quad \sigma_t^2 = \alpha_0 + \alpha_1 X_{t-1}^2 + \beta_1 \sigma_{t-1}^2, \quad t \in \mathbb{Z},$$

where $\alpha_0, \alpha_1, \beta_1 > 0$, and $\{Z_t, t \in \mathbb{Z}\}$ is a sequence of i.i.d. random variables with mean zero and variance one. Under mild conditions the process $\{X_t\}$ is regularly varying and satisfies Assumptions 1.1 and 3.5. These hold for instance in the case of standard normal innovations Z_t and sufficiently small parameters α_1, β_1 , see [26, Section 4.4]. Consider for simplicity such a stationary GARCH(1, 1) process with tail index $\alpha \in (0, 1)$. Since all the conditions of Theorem 4.5 are met, its partial sum process has a limit in the space E (cf. Fig. 2).

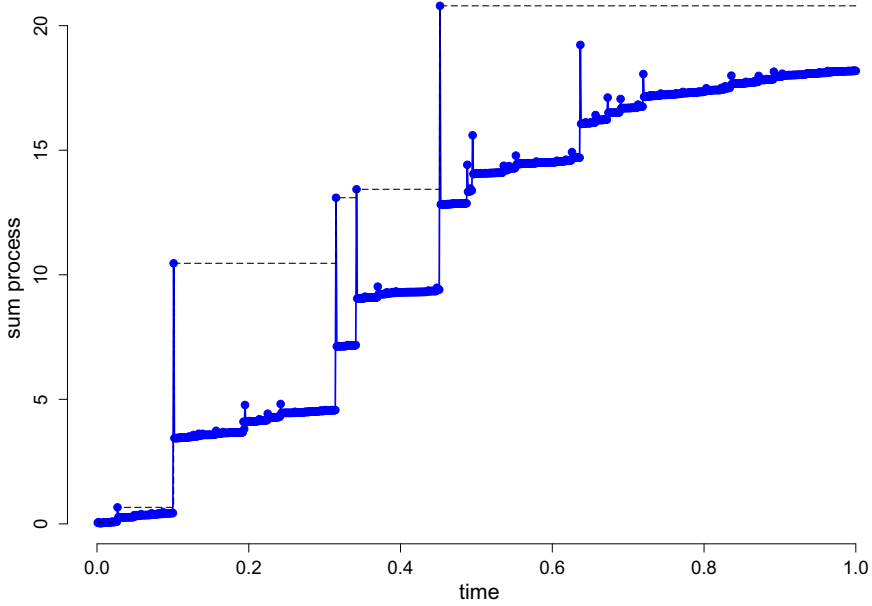


Fig. 1 A simulated sample path of the process S_n in the case of linear sequence $X_t = \xi_t - 0.7\xi_{t-1}$ with index of regular variation $\alpha = 0.7$ in blue. Observe that due to downward “corrections” after each large jump, in the limit the paths of the process S_n cannot converge to a càdlàg function

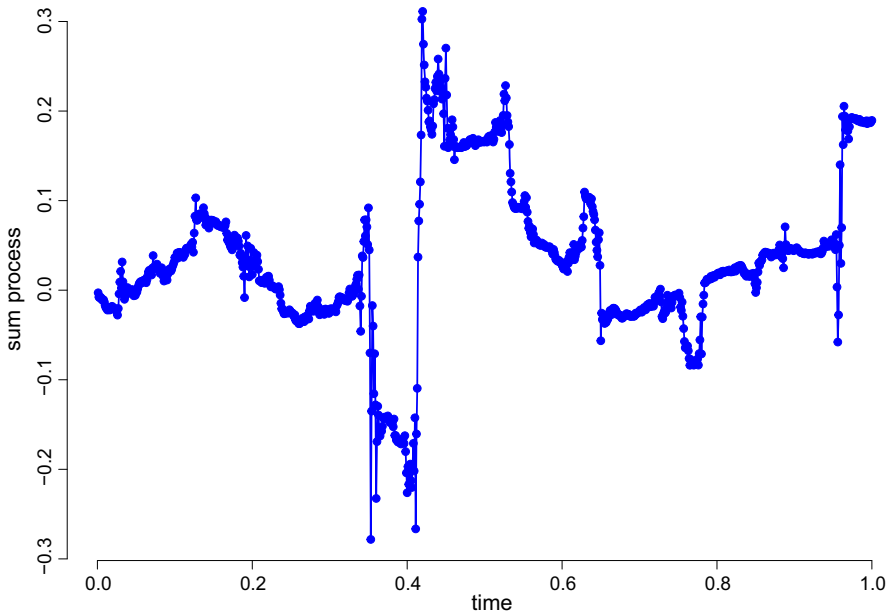


Fig. 2 A simulated sample path of the process S_n in the case of GARCH(1, 1) process with parameters $\alpha_0 = 0.01$, $\alpha_1 = 1.45$ and $\beta_1 = 0.1$, and tail index α between 0.5 and 1

Proof of Theorem 4.5. The proof is split into the cases $\alpha < 1$ which is simpler, and the case $\alpha \in [1, 2)$ where centering and truncation introduce additional technical difficulties.

(a) Assume first that $\alpha \in (0, 1)$. We divide the proof into several steps.

Step 1 For every $\epsilon > 0$, consider the functions s^ϵ, u^ϵ and v^ϵ defined on \tilde{l}_0 by

$$s^\epsilon(\tilde{\mathbf{x}}) = \sum_j x_j \mathbb{1}_{\{|x_j|>\epsilon\}}, \quad u^\epsilon(\tilde{\mathbf{x}}) = \inf_k \sum_{j \leq k} x_j \mathbb{1}_{\{|x_j|>\epsilon\}},$$

$$v^\epsilon(\tilde{\mathbf{x}}) = \sup_k \sum_{j \leq k} x_j \mathbb{1}_{\{|x_j|>\epsilon\}},$$

and define the mapping $T^\epsilon : \mathcal{M}_p([0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}) \rightarrow E$ by setting, for $\gamma = \sum_{i=1}^\infty \delta_{t_i, \tilde{\mathbf{x}}^i}$,

$$T^\epsilon \gamma = \left(\left(\sum_{t_i \leq t} s^\epsilon(\tilde{\mathbf{x}}^i) \right)_{t \in [0,1]}, \{t_i : \|\tilde{\mathbf{x}}^i\|_\infty > \epsilon\}, \{I(t_i) : \|\tilde{\mathbf{x}}^i\|_\infty > \epsilon\} \right),$$

where

$$I(t_i) = \sum_{t_j < t_i} s^\epsilon(\tilde{\mathbf{x}}^j) + \left[\sum_{t_k = t_i} u^\epsilon(\tilde{\mathbf{x}}^k), \sum_{t_k = t_i} v^\epsilon(\tilde{\mathbf{x}}^k) \right].$$

Since m belongs $\mathcal{M}_p([0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$, there is only a finite number of points $(t_i, \tilde{\mathbf{x}}^i)$ such that $\|\tilde{\mathbf{x}}^i\|_\infty > \epsilon$ and furthermore, every $\tilde{\mathbf{x}}^i$ has at most finitely many coordinates greater than ϵ . Therefore, the mapping T^ϵ is well-defined, that is, $T^\epsilon \gamma$ is a proper element of E .

Next, we define the subsets of $\mathcal{M}_p([0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$ by

$$\Lambda_1 = \left\{ \sum_{i=1}^\infty \delta_{t_i, \tilde{\mathbf{x}}^i} : |x_j^i| \neq \epsilon, i \geq 1, j \in \mathbb{Z} \right\},$$

$$\Lambda_2 = \left\{ \sum_{i=1}^\infty \delta_{t_i, \tilde{\mathbf{x}}^i} : 0 < t_i < 1 \text{ and } t_i \neq t_j \text{ for every } i > j \geq 1 \right\}.$$

We claim that T^ϵ is continuous on the set $\Lambda_1 \cap \Lambda_2$. Assume that $\gamma_n \xrightarrow{w^\#} \gamma = \sum_{i=1}^\infty \delta_{t_i, \tilde{\mathbf{x}}^i} \in \Lambda_1 \cap \Lambda_2$. By an adaptation of [30, Proposition 3.13], this convergence implies that the finitely many points of γ_n in every set B bounded for \tilde{d}' and such that $\gamma(\partial B) = 0$ converge pointwise to the finitely many points of γ in B . In particular, this holds for $B = \{(t, \tilde{\mathbf{x}}) : \|\tilde{\mathbf{x}}\|_\infty > \epsilon\}$ and it follows that for all $t_1 < t_2$ in $[0, 1]$ such that $\gamma(\{t_1, t_2\} \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}) = 0$,

$$M_{t_1, t_2}(T^\epsilon(\gamma_n)) \rightarrow M_{t_1, t_2}(T^\epsilon(\gamma)) \text{ in } \mathbb{R}$$

and

$$M_{t_1, t_2}(-T^\epsilon(\gamma_n)) \rightarrow M_{t_1, t_2}(-T^\epsilon(\gamma)) \text{ in } \mathbb{R},$$

whith the local-maximum function M_{t_1, t_2} defined as in (4.6). Since the set of all such times is dense in $[0, 1]$ and includes 0 and 1, an application of Theorem 4.3 gives that

$$T^\epsilon(\gamma_n) \rightarrow T^\epsilon(\gamma)$$

in E endowed with the M_2 topology.

Recall the point process $N'' = \sum_{i=1}^\infty \delta_{(T_i, P_i, Q_i)}$ from (4.8). Since the mean measure of $\sum_{i=1}^\infty \delta_{(T_i, P_i)}$ does not have atoms, it is clear that $N'' \in \Lambda_1 \cap \Lambda_2$ a.s. Therefore, by the convergence $N''_n \xrightarrow{d} N''$ and the continuous mapping argument

$$\tilde{S}'_{n, \epsilon} \xrightarrow{d} S'_\epsilon,$$

where $\tilde{S}'_{n, \epsilon} = T^\epsilon(N''_n)$ and $S'_\epsilon = T^\epsilon(N'')$.

Step 2 Recall that $W_i = \sum_{j \in \mathbb{Z}} |Q_{i, j}|$ and $\sum_{i=1}^\infty \delta_{P_i W_i}$ is a Poisson point process on $(0, \infty]$ with intensity measure $\theta \mathbb{E}[W_i^\alpha] \alpha y^{-\alpha-1} dy$ (see Remark 4.6). Since $\alpha < 1$, one can sum up the points $\{P_i W_i\}$, i.e.

$$\sum_{i=1}^\infty P_i W_i = \sum_{i=1}^\infty \sum_{j \in \mathbb{Z}} P_i |Q_{i, j}| < \infty \text{ a.s.} \tag{4.16}$$

Therefore, defining $s(\vec{x}) = \sum_j x_j$, we obtain that the process

$$V(t) = \sum_{T_i \leq t} s(P_i Q_i), \quad t \in [0, 1],$$

is almost surely a well-defined element in D and moreover, it is an α -stable Lévy process. Further, we define an element V' in $E([0, 1], \mathbb{R})$ by

$$V' = (V, \{T_i\}_{i \in \mathbb{N}}, \{I(T_i)\}_{i \in \mathbb{N}}), \tag{4.17}$$

where

$$\begin{aligned} I(T_i) &= V(T_i-) + [u(P_i Q_i), v(P_i Q_i)], \\ u(\vec{x}) &= \inf_k \sum_{j \leq k} x_j, \quad v(\vec{x}) = \sup_k \sum_{j \leq k} x_j. \end{aligned}$$

Since for every $\delta > 0$ there are at most finitely many points $P_i W_i$ such that $P_i W_i > \delta$ and $diam(I(T_i)) = v(P_i Q_i) - u(P_i Q_i) \leq P_i W_i$, V' is indeed a proper element of E a.s.

We now show that, as $\epsilon \rightarrow 0$, the limits S'_ϵ from the previous step converge to V' in (E, m) almost surely. Recall the uniform metric m^* on E defined in (4.2). By (4.16) and dominated convergence theorem

$$m^*(S'_\epsilon, V') = \sup_{0 \leq t \leq 1} m(S'_\epsilon(t), V'(t)) \leq \sum_{i=1}^\infty \sum_{j \in \mathbb{Z}} P_i |Q_{ij}| \mathbb{1}_{\{|P_i Q_{ij}| \leq \epsilon\}} \rightarrow 0, \tag{4.18}$$

almost surely as $\epsilon \rightarrow 0$. Indeed, let S_ϵ be the càdlàg part of S'_ϵ , i.e.

$$S_\epsilon(t) = \sum_{T_i \leq t} s^\epsilon(P_i Q_i) = \sum_{T_i \leq t} \sum_{j \in \mathbb{Z}} P_i Q_{i,j} \mathbb{1}_{\{|P_i Q_{i,j}| > \epsilon\}}, \quad t \in [0, 1].$$

If $t \notin \{T_i\}$ then

$$m(S'_\epsilon(t), V'(t)) = |S_\epsilon(t) - V(t)| \leq \sum_{T_i \leq t} \sum_{j \in \mathbb{Z}} P_i |Q_{i,j}| \mathbb{1}_{\{|P_i Q_{i,j}| \leq \epsilon\}}.$$

Further, when $t = T_k$ for some $k \in \mathbb{Z}$, by using (4.4) we obtain

$$\begin{aligned} m(S'_\epsilon(t), V'(t)) &\leq |(S_\epsilon(T_k-) + v^\epsilon(P_k Q_k)) - (V(T_k-) + v(P_k Q_k))| \\ &\quad \vee |(S_\epsilon(T_k-) + u^\epsilon(P_k Q_k)) - (V(T_k-) + u(P_k Q_k))|. \end{aligned} \tag{4.19}$$

The first term on the right-hand side of the equation above is bounded by

$$\begin{aligned} &\left| \sum_{T_i < T_k} \sum_{j \in \mathbb{Z}} P_i Q_{i,j} \mathbb{1}_{\{|P_i Q_{ij}| \leq \epsilon\}} \right| + \left| \sup_{l \in \mathbb{Z}} \sum_{j \leq l} P_k Q_{k,j} \mathbb{1}_{\{|P_i Q_{i,j}| > \epsilon\}} - \sup_{l \in \mathbb{Z}} \sum_{j \leq l} P_k Q_{k,j} \right| \\ &\leq \left| \sum_{T_i < T_k} \sum_{j \in \mathbb{Z}} P_i Q_{i,j} \mathbb{1}_{\{|P_i Q_{ij}| \leq \epsilon\}} \right| + \sup_{l \in \mathbb{Z}} \left| \sum_{j \leq l} P_k Q_{k,j} \mathbb{1}_{\{|P_i Q_{k,j}| > \epsilon\}} - \sum_{j \leq l} P_k Q_{k,j} \right| \\ &\leq \sum_{T_i \leq T_k} \sum_{j \in \mathbb{Z}} P_i |Q_{i,j}| \mathbb{1}_{\{|P_i Q_{i,j}| \leq \epsilon\}}. \end{aligned}$$

Since, by similar arguments, one can obtain the same bound for the second term on the right-hand side of (4.19), (4.18) holds.

It now follows from (4.3) that

$$S'_\epsilon \rightarrow V'$$

almost surely in (E, m) .

Step 3 Recall that

$$X_{n,i} = (X_{(i-1)r_n+1}, \dots, X_{ir_n})/a_n$$

for $i = 1, \dots, k_n$ and let \tilde{S}'_n be an element in E defined by

$$\left(\left(\sum_{i/k_n \leq t} s(X_{n,i}) \right)_{t \in [0,1]}, \{i/k_n\}_{i=1}^{k_n}, \{I(i/k_n)\}_{i=1}^{k_n} \right),$$

where

$$I(i/k_n) = \sum_{j < i} s(X_{n,j}) + [u(X_{n,i}), v(X_{n,i})].$$

By [8, Theorem 4.2] and the previous two steps, to show that

$$\tilde{S}'_n \xrightarrow{d} V'$$

in (E, m_E) , it suffices to prove that, for all $\delta > 0$,

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(m_E(\tilde{S}'_{n,\epsilon}, \tilde{S}'_n) > \delta) = 0. \tag{4.20}$$

Note first that, by the same arguments as in the previous step, we have

$$m^*(\tilde{S}'_{n,\epsilon}, \tilde{S}'_n) \leq \sum_{j=1}^{k_n r_n} \frac{|X_j|}{a_n} \mathbb{1}_{\{|X_j| \leq a_n \epsilon\}}.$$

By (4.3), Markov’s inequality and Karamata’s theorem ([9, Proposition 1.5.8])

$$\begin{aligned} \limsup_{n \rightarrow \infty} \mathbb{P}(m_E(\tilde{S}'_{n,\epsilon}, \tilde{S}'_n) > \delta) &\leq \limsup_{n \rightarrow \infty} \frac{k_n r_n}{\delta a_n} \mathbb{E} [|X_1| \mathbb{1}_{\{|X_1| \leq a_n \epsilon\}}] \\ &= \lim_{n \rightarrow \infty} \frac{n}{\delta a_n} \cdot \frac{\alpha a_n \epsilon \mathbb{P}(|X_1| > a_n \epsilon)}{1 - \alpha} \\ &= \frac{\alpha}{(1 - \alpha)\delta} \epsilon^{1-\alpha}. \end{aligned}$$

This proves (4.20) since $\alpha < 1$ and hence

$$\tilde{S}'_n \xrightarrow{d} V'$$

in (E, m_E) .

Step 4 Finally, to show that the original partial sum process S_n (and therefore V_n since $\alpha \in (0, 1)$) also converges in distribution to V' in (E, m_E) , by Slutsky argument it suffices to prove that

$$m_E(S_n, \tilde{S}'_n) \xrightarrow{P} 0. \tag{4.21}$$

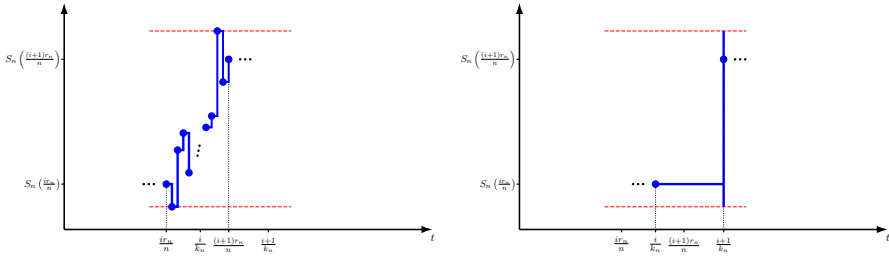


Fig. 3 Restrictions of graphs Γ_{S_n} and $\Gamma_{\tilde{S}'_n}$ on time intervals $(\frac{ir_n}{n}, \frac{(i+1)r_n}{n}]$ and $(\frac{i}{k_n}, \frac{i+1}{k_n}]$, respectively

Recall that $k_n = \lfloor n/r_n \rfloor$ so $\frac{ir_n}{n} \leq \frac{i}{k_n}$ for all $i = 0, 1, \dots, k_n$ and moreover

$$\frac{i}{k_n} - \frac{ir_n}{n} = \frac{i}{k_n} \left(1 - \frac{k_n r_n}{n} \right) \leq 1 - \frac{\lfloor n/r_n \rfloor}{n/r_n} = 1 - \left(1 - \frac{\{n/r_n\}}{n/r_n} \right) \leq \frac{r_n}{n}. \tag{4.22}$$

Let $d_{n,i}$ for $i = 0, \dots, k_n - 1$ be the Hausdorff distance between restrictions of graphs Γ_{S_n} and $\Gamma_{\tilde{S}'_n}$ on time intervals $(\frac{ir_n}{n}, \frac{(i+1)r_n}{n}]$ and $(\frac{i}{k_n}, \frac{i+1}{k_n}]$, respectively (see Fig. 3).

First note that, by (4.22), the time distance between any two points on these graphs is at most $2r_n/n$. Further, by construction, S_n and \tilde{S}'_n have the same range of values on these time intervals. More precisely,

$$\begin{aligned} \bigcup_{t \in (\frac{ir_n}{n}, \frac{(i+1)r_n}{n}]} \{z \in \mathbb{R} : (t, z) \in \Gamma_{S_n}\} &= \bigcup_{t \in (\frac{i}{k_n}, \frac{i+1}{k_n}]} \{z \in \mathbb{R} : (t, z) \in \Gamma_{\tilde{S}'_n}\} \\ &= \tilde{S}'_n((i + 1)/k_n). \end{aligned}$$

Therefore, the distance between the graphs comes only from the time component, i.e.

$$d_{n,i} \leq \frac{2r_n}{n} \rightarrow 0, \text{ as } n \rightarrow \infty,$$

for all $i = 0, 1, \dots, k_n - 1$.

Moreover, if we let d_{n,k_n} be the Hausdorff distance between the restriction of the graph Γ_{S_n} on $(\frac{k_n r_n}{n}, 1]$ and the interval $(1, \tilde{S}'_n(1))$, it holds that

$$d_{n,k_n} \leq \frac{r_n}{n} \vee \sum_{j=k_n r_n + 1}^n \frac{|X_j|}{a_n} \xrightarrow{P} 0,$$

as $n \rightarrow \infty$. Hence, (4.21) holds since

$$m_E(S_n, \tilde{S}'_n) \leq \bigvee_{i=0}^{k_n} d_{n,i},$$

and this finishes the proof in the case $\alpha \in (0, 1)$.

(b) Assume now that $\alpha \in [1, 2)$. As shown in Step 1. in the proof of (a), it holds that

$$\tilde{S}'_{n,\epsilon} \xrightarrow{d} S'_\epsilon \tag{4.23}$$

in E , where $\tilde{S}'_{n,\epsilon} = T^\epsilon(N''_n)$ and $S'_\epsilon = T^\epsilon(N'')$.

For every $\epsilon > 0$ define a càdlàg process $S_{n,\epsilon}$ by setting, for $t \in [0, 1]$,

$$S_{n,\epsilon}(t) = \sum_{i=1}^{\lfloor nt \rfloor} \frac{X_i}{a_n} \mathbb{1}_{\{|X_i|/a_n > \epsilon\}}.$$

Using the same arguments as in Step 4. in the proof of (a), it holds that, as $n \rightarrow \infty$,

$$m_E(S_{n,\epsilon}, \tilde{S}'_{n,\epsilon}) \xrightarrow{P} 0. \tag{4.24}$$

Therefore, by Slutsky argument it follows from (4.24) and (4.23) that

$$S_{n,\epsilon} \xrightarrow{d} S'_\epsilon \tag{4.25}$$

in (E, m_E) .

Since $\alpha \in [1, 2)$ we need to introduce centering, so we define the càdlàg process $V_{n,\epsilon}$ by setting, for $t \in [0, 1]$,

$$V_{n,\epsilon}(t) = S_{n,\epsilon}(t) - \lfloor nt \rfloor \mathbb{E} \left(\frac{X_1}{a_n} \mathbb{1}_{\{\epsilon < |X_1|/a_n \leq 1\}} \right).$$

From (1.4) we have, for any $t \in [0, 1]$, as $n \rightarrow \infty$,

$$\lfloor nt \rfloor \mathbb{E} \left(\frac{X_1}{a_n} \mathbb{1}_{\{\epsilon < |X_1|/a_n \leq 1\}} \right) \rightarrow t \int_{\{x: \epsilon < |x| \leq 1\}} x \mu(dx). \tag{4.26}$$

Since the limit function above is continuous and the convergence is uniform on $[0, 1]$, by Lemma 6.3 and (4.25) it follows that

$$V_{n,\epsilon} \xrightarrow{d} V'_\epsilon \tag{4.27}$$

in E , where V'_ϵ is given by (see Remark 4.2)

$$V'_\epsilon(t) = S'_\epsilon(t) - t \int_{\{x: \epsilon < |x| \leq 1\}} x \mu(dx).$$

Let V_ϵ be the càdlàg part of V'_ϵ , i.e.,

$$V_\epsilon(t) = \sum_{T_i \leq t} s^\epsilon(P_i Q_i) - t \int_{\{x: \epsilon < |x| \leq 1\}} x \mu(dx). \tag{4.28}$$

By Lemma 6.4, there exist an α -stable Lévy process V such that V_ϵ converges uniformly almost surely (along some subsequence) to V .

Next, as in Step 2. in the proof of (a) we can define a element V' in E by

$$V' = (V, \{T_i\}_{i \in \mathbb{N}}, \{I(T_i)\}_{i \in \mathbb{N}}), \tag{4.29}$$

where

$$I(T_i) = V(T_i-) + [u(P_i Q_i), v(P_i Q_i)], \tag{4.30}$$

Also, one can argue similarly to the proof of (4.18) to conclude that

$$m^*(V', V'_\epsilon) \leq \|V - V_\epsilon\|_\infty + \sup_{i \in \mathbb{N}} \sum_{j \in \mathbb{Z}} P_i |Q_{i,j}| \mathbb{1}_{\{|P_i Q_{i,j}| \leq \epsilon\}}.$$

Now it follows from (4.3), (4.12) and a.s. uniform convergence of V_ϵ to V that

$$V'_\epsilon \rightarrow V' \text{ a.s.} \tag{4.31}$$

in (E, m_E) .

Finally, by (4.27), (4.31) and [8, Theorem 4.2], to show that

$$V_n \xrightarrow{d} V', \tag{4.32}$$

in (E, m_E) , it suffices to prove that, for all $\delta > 0$,

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(m_E(V_{n,\epsilon}, V_n) > \delta) = 0. \tag{4.33}$$

But this follows from Assumption 4.4 and (4.3) since

$$m^*(V_{n,\epsilon}, V_n) \leq \max_{1 \leq k \leq n} \left| \sum_{i=1}^k \left(\frac{X_i}{a_n} \mathbb{1}_{\{|X_i|/a_n \leq \epsilon\}} - \mathbb{E} \left[\frac{X_i}{a_n} \mathbb{1}_{\{|X_i|/a_n \leq \epsilon\}} \right] \right) \right|.$$

Hence (4.32) holds and this finishes the proof. □

4.3 Supremum of the partial sum process

We next show that the supremum of the partial sum process converges in distribution in D endowed with the M_1 topology, where the limit is the “running supremum” of the limit process V' from Theorem 4.5.

Let V be the Lévy process defined in (4.11) and define the process V^+ on $[0, 1]$ by

$$V^+(t) = \begin{cases} V(t), & t \notin \{T_j\}_{j \in \mathbb{N}} \\ V(t-) + \sup_{k \in \mathbb{Z}} \sum_{j \leq k} P_i Q_{i,j}, & t = T_i \text{ for some } i \in \mathbb{N}. \end{cases}$$

Define V^- analogously using infimum instead of supremum. Note that V^+ and V^- need not be right-continuous at the jump times T_j . However, their partial supremum or infimum are càdlàg functions.

Theorem 4.12 *Under the same conditions as in Theorem 4.5, it holds that*

$$\left(\sup_{s \leq t} V_n(s) \right)_{t \in [0,1]} \xrightarrow{d} \left(\sup_{s \leq t} V^+(s) \right)_{t \in [0,1]},$$

and

$$\left(\inf_{s \leq t} V_n(s) \right)_{t \in [0,1]} \xrightarrow{d} \left(\inf_{s \leq t} V^-(s) \right)_{t \in [0,1]},$$

jointly in $D([0, 1], \mathbb{R})$ endowed with the M_1 topology.

Proof We prove the result only for the supremum of the partial sum process since the infimum case is completely analogous and joint convergence holds since we are applying the continuous mapping argument to the same process.

Define the mapping $\text{sup} : E([0, 1], \mathbb{R}) \rightarrow D([0, 1], \mathbb{R})$ by

$$\text{sup}(x')(t) = \sup\{z : z \in x'(s), 0 \leq s \leq t\}.$$

Note that $\text{sup}(x')$ is non-decreasing and since for every $\delta > 0$ there are at most finitely many times t for which the $\text{diam}(x'(t))$ is greater than δ , by [34, Theorem 15.4.1.] it follows easily that this mapping is well-defined, i.e. that $\text{sup}(x')$ is indeed an element in D . Also, by construction,

$$\text{sup}(V') = \left(\sup_{s \leq t} V^+(s) \right)_{t \in [0,1]}$$

and

$$\text{sup}(V_n) = \left(\sup_{s \leq t} V_n(s) \right)_{t \in [0,1]}.$$

Define the subset of E by

$$\Lambda = \{x' \in E : x'(0) = \{0\}\}$$

and assume that $x'_n \rightarrow x'$ in (E, m_E) , where $x'_n, x' \in \Lambda$. By Theorem 4.3 it follows that

$$\text{sup}(x'_n)(t) = M_{0,t}(x'_n) \rightarrow M_{0,t}(x') = \text{sup}(x')(t)$$

for all t in a dense subset of $(0,1]$, including 1. Also, the convergence trivially holds for $t = 0$ since $\sup(x'_n)(0) = \sup(x')(0) = 0$ for all $n \in \mathbb{N}$. Since $\sup(x')$ is non-decreasing for all $x' \in E$, we can apply [34, Corollary 12.5.1] and conclude that

$$\sup(x'_n) \rightarrow \sup(x')$$

in D endowed with M_1 topology. Since $V_n, V' \in \Lambda$ almost surely, by Theorem 4.5 and continuous mapping argument it follows that

$$\left(\sup_{s \leq t} V_n(s) \right)_{t \in [0,1]} \xrightarrow{d} \left(\sup_{s \leq t} V^+(s) \right)_{t \in [0,1]}$$

in D endowed with M_1 topology. □

Remark 4.13 Note that when $\sum_{j \in \mathbb{Z}} Q_j = 0$ a.s., the limit for the supremum of the partial sum process in Theorem 4.12 is simply a so called Fréchet extremal process. For an illustration of the general limiting behavior of running maxima in the case of a linear processes, consider again the moving average of order 1 from Example 4.9. Figure 1 shows a path (dashed line) of the running maxima of the MA(1) process $X_t = \xi_t - 0.7\xi_{t-1}$.

4.4 M_2 convergence of the partial sum process

We can now characterize the convergence of the partial sum process in the M_2 topology in $D([0, 1])$ by an appropriate condition on the tail process of the sequence $\{X_t, t \in \mathbb{Z}\}$.

Assumption 4.14 The sequence $\{Q_j, j \in \mathbb{Z}\}$ satisfies

$$-\left(\sum_{j \in \mathbb{Z}} Q_j \right)_- = \inf_{k \in \mathbb{Z}} \sum_{j \leq k} Q_j \leq \sup_{k \in \mathbb{Z}} \sum_{j \leq k} Q_j = \left(\sum_{j \in \mathbb{Z}} Q_j \right)_+ \quad \text{a.s.} \quad (4.34)$$

i.e. $-s(Q)_- = u(Q) \leq v(Q) = s(Q)_+$ a.s.

Note that this assumption ensures that $\sum_{j \in \mathbb{Z}} Q_j \neq 0$ and that the limit process V' from Theorem 4.5 has sample paths in the subset D' of E which was defined in (4.5). By Lemma 4.1, Theorem 4.5 and the continuous mapping theorem, the next result follows immediately.

Theorem 4.15 *If, in addition to conditions in Theorem 4.5, Assumption 4.14 holds, then*

$$V_n \xrightarrow{d} V$$

in $D([0, 1], \mathbb{R})$ endowed with the M_2 topology.

Since the supremum functional is continuous with respect to the M_2 topology, this result implies that the limit of the running supremum of the partial sum process is the running supremum of the limiting α -stable Lévy process as in the case of i.i.d. random variables.

Example 4.16 For the linear process $X_t = \sum_{j \in \mathbb{Z}} c_j \xi_{t-j}$ from Sect. 3.3 and Example 4.9, the corresponding sequence $\{Q_j\}$ was given in (3.8). It follows that the condition (4.34) can be expressed as

$$-\left(\sum_{j \in \mathbb{Z}} c_j\right)_- = \inf_{k \in \mathbb{Z}} \sum_{j \leq k} c_j \leq \sup_{k \in \mathbb{Z}} \sum_{j \leq k} c_j = \left(\sum_{j \in \mathbb{Z}} c_j\right)_+. \tag{4.35}$$

This is exactly [4, Condition 3.2]. Note that (4.35) implies that

$$\left| \sum_{j \in \mathbb{Z}} c_j \right| > 0.$$

5 Record times

In this section we study record times in a stationary sequence $\{X_t, t \in \mathbb{Z}\}$. Since record times remain unaltered after a strictly increasing transformation, the main result below holds for stationary sequences with a general marginal distribution as long as they can be monotonically transformed into a regularly varying sequence.

We start by introducing the notion of records for sequences in \tilde{l}_0 . For $y \geq 0$ and $\mathbf{x} = \{x_j\} \in \tilde{l}_0$ define

$$R^{\mathbf{x}}(y) = \sum_{j=-\infty}^{\infty} \mathbb{1}_{\{x_j > y \vee \sup_{i < j} x_i\}},$$

representing the number of records in the sequence \mathbf{x} larger than y , which is finite for $\mathbf{x} \in \tilde{l}_0$. For notational simplicity, we suppress notation $\tilde{\mathbf{x}}$ in this section.

Let $\gamma = \sum_{i=1}^{\infty} \delta_{t_i, \mathbf{x}^i} \in \mathcal{M}_p([0, \infty) \times \tilde{l}_0 \setminus \{\mathbf{0}\})$, where $\mathbf{x}^i = \{x_j^i\} \in \tilde{l}_0$. Define, for $t > 0$

$$M^\gamma(t) = \sup_{t_i \leq t} \|\mathbf{x}^i\|_\infty, \quad M^\gamma(t-) = \sup_{t_i < t} \|\mathbf{x}^i\|_\infty,$$

where we set $\sup \emptyset = 0$ for convenience. Next, let R_γ be the (counting) point process on $(0, \infty)$ defined by

$$R_\gamma = \sum_i \delta_{t_i} R^{\mathbf{x}^i}(M^\gamma(t_i-)),$$

hence for arbitrary $0 < a < b$

$$R_\gamma(a, b) = \sum_{a < t_i \leq b} \sum_{j=-\infty}^{\infty} \mathbb{1}_{\{x_j^i > M^\gamma(t_i-) \vee \sup_{k < j} x_k^i\}}.$$

Consider the following subset of $\mathcal{M}_p([0, \infty) \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$

$$A = \left\{ \gamma = \sum_{i=1}^{\infty} \delta_{t_i, x^i} : \text{such that } M^\gamma(t) > 0 \text{ for all } t > 0, \text{ while} \right. \\ \left. \text{all } t_i \text{'s are mutually different as well as all nonzero } x_j^i \text{'s} \right\}.$$

The space $\mathcal{M}_p((0, \infty))$ is endowed with the $w^\#$ topology which is equivalent to the usual vague topology since $(0, \infty)$ is locally compact and separable.

Lemma 5.1 *The mapping $\gamma \mapsto R_\gamma$ from $\mathcal{M}_p([0, \infty) \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$ to $\mathcal{M}_p((0, \infty))$ is continuous at every $\gamma \in A$.*

Proof Fix an arbitrary $\gamma \in A$, and assume a sequence $\{\gamma_n\}$ in $\mathcal{M}_p([0, \infty) \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$ satisfies $\gamma_n \rightarrow_{w^\#} \gamma$. We must prove that R_{γ_n} converges vaguely to R_γ in $\mathcal{M}_p((0, \infty))$. By the Portmanteau theorem, it is sufficient to show that that for all $0 < a < b \in \{t_i\}^c$

$$R_{\gamma_n}(a, b) \rightarrow R_\gamma(a, b),$$

as $n \rightarrow \infty$. For $0 < a < b \in \{t_i\}^c$ there are finitely many time instances, say $t_{i_1}, \dots, t_{i_k} \in (a, b)$ such that $\|x^{i_l}\|_\infty > M^\gamma(a) > 0$, hence

$$R_\gamma(a, b) = \sum_{t_i \in (a, b)} R^{x^i}(M^\gamma(t_i-)) = \sum_{l=1}^k R^{x^{i_l}}(M^\gamma(t_{i_l})). \tag{5.1}$$

For all $\gamma_n = \sum_{i=1}^{\infty} \delta_{t_i^n, x^{n,i}}$ with n large enough, there also exist exactly k (depending on a and b) time instances $t_{i_1}^n, \dots, t_{i_k}^n \in (a, b)$ such that $\|x^{n,i_l}\|_\infty > M^\gamma(a)$. Moreover, they satisfy $x^{n,i_l} \rightarrow x^{i_l}$ and $t_{i_l}^n \rightarrow t_{i_l}$ for $l = 1, \dots, k$ as $n \rightarrow \infty$. Hence for n large enough

$$R_{\gamma_n}(a, b) = \sum_{t_i^n \in (a, b)} R^{x^{n,i}}(M^{\gamma_n}(t_i^n-)) = \sum_{l=1}^k R^{x^{n,i_l}}(M^{\gamma_n}(t_{i_l}^n-)). \tag{5.2}$$

Assume $y_n \rightarrow y > 0$ and $x^n \rightarrow x = \{x_j\}$ where the non zero x_j are pairwise distinct and $x_j \neq y$ for all $j \in \mathbb{Z}$. Then, it is straightforward to check that

$$R^{x^n}(y_n) \rightarrow R^x(y).$$

Observe further that for the choice of $t_{i_l}, t_{i_l}^n$ we made above, it holds that $M^{\gamma_n}(t_{i_l}^n -) \rightarrow M^\gamma(t_{i_l} -)$ since t_i 's are all different. Together with (5.1) and (5.2) this yields

$$R_{\gamma_n}(a, b] \rightarrow R_\gamma(a, b]$$

as $n \rightarrow \infty$. □

Since we are only interested in records, for convenience we consider a nonnegative stationary regularly varying sequence $\{X_t, t \in \mathbb{Z}\}$.

Adopting the notation of Sect. 3.2, we denote

$$N_n'' = \sum_{i=1}^{\infty} \delta_{(i/k_n, X_{n,i})}. \tag{5.3}$$

We will also need the point process N_n defined by

$$N_n = \sum_{i=1}^{\infty} \delta_{(i/n, X_i/a_n)}.$$

The process N_n can be viewed as a point process on $[0, \infty) \times \mathbb{R} \setminus \{0\}$, but since \mathbb{R} can be embedded in \tilde{l}_0 (by identifying a real number $x \neq 0$ to a sequence with exactly one nonzero coordinate equal to x), in the sequel we treat it as a process on the space $[0, \infty) \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$.

As in the previous section, we will assume that

$$N_n'' \xrightarrow{d} N'' = \sum_{i=1}^{\infty} \delta_{(T_i, P_i Q_i)}, \tag{5.4}$$

as $n \rightarrow \infty$, where N'' has the same form as in (4.8), but on the space $\mathcal{M}_p([0, \infty) \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\})$. For β -mixing and linear processes, this convergence follows by direct extension of results in Sect. 3 from the state space $[0, 1] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$ to $[0, T] \times \tilde{l}_0 \setminus \{\tilde{\mathbf{0}}\}$ for arbitrary $T > 0$.

Theorem 5.2 *Let $\{X_t, t \in \mathbb{Z}\}$ be a stationary regularly varying sequence with tail index $\alpha > 0$. Assume that the convergence in (5.4) holds and moreover that*

$$\mathbb{P}(\text{all nonzero } Q_{1,j} \text{'s are mutually different}) = 1.$$

Then

$$R_{N_n} \xrightarrow{d} R_{N''},$$

in $\mathcal{M}_p((0, \infty))$. Moreover, the limiting process is a compound Poisson process with representation

$$R_{N''} = \sum_{i \in \mathbb{Z}} \delta_{\tau_i} \kappa_i,$$

where $\sum_{i \in \mathbb{Z}} \delta_{\tau_i}$ is a Poisson point process on $(0, \infty)$ with intensity measure $x^{-1} dx$ and $\{\kappa_i\}$ is a sequence of i.i.d. random variables independent of it with the same distribution as the integer valued random variable $R_{Q_1}(1/\zeta)$ where ζ is a Pareto random variable with tail index α , independent of Q_1 .

Proof Since they are constructed from the same sequence X_1, X_2, \dots , the record values of the point process N''_n in (5.3) correspond to the record values of the point process N_n . Using (5.4) and the additional assumption on the Q 's, by Lemma 5.1 it follows that

$$R_{N''_n} \xrightarrow{d} R_{N''}. \tag{5.5}$$

Note that record times i/n of the process N_n appear at slightly altered times $(\lfloor i/r_n \rfloor + 1)/k_n$ in the process N''_n . However, asymptotically the record times are very close. Indeed, take $f \in C^+_K(0, \infty)$, then f has a support on the set of the form $(a, b]$, $0 < a < b$, which can be enlarged slightly to a set $(a - \epsilon, b + \epsilon]$, where $0 < a - \epsilon$ for a sufficiently small $\epsilon > 0$. Clearly, f is continuous on that set, even uniformly. Note further that for any such ϵ there is an integer n_0 such that for $n \geq n_0$, $i/n \in (a, b]$ implies $(\lfloor i/r_n \rfloor + 1)/k_n \in (a - \epsilon, b + \epsilon]$, and vice versa. Moreover, by uniform continuity of the function f , n_0 can be chosen such that for $n \geq n_0$,

$$\left| f\left(\frac{i}{n}\right) - f\left(\frac{\lfloor i/r_n \rfloor + 1}{k_n}\right) \right| \leq \epsilon.$$

Consider now the difference between Laplace functionals of the point processes $R_{N''_n}$ and R_{N_n} for a function $f \in C^+_K(0, \infty)$ as above. Since ϵ above can be made arbitrarily small, it follows that

$$\left| \mathbb{E}[e^{-f(R_{N''_n})}] - \mathbb{E}[e^{-f(R_{N_n})}] \right| \rightarrow 0,$$

which together with (5.5) yields the convergence statement of the theorem.

Consider now a point measure $\gamma = \sum_{i=1}^\infty \delta_{t_i, \mathbf{x}^i} \in \mathcal{M}_p([0, \infty) \times \tilde{l}_0 \setminus \{\mathbf{0}\})$, but such that all \mathbf{x}^i have only nonnegative components and all t_i 's are mutually different. We say that a point measure γ has a record at time t if $(t, \mathbf{x}) \in \{(t_i, \mathbf{x}^i) : i \geq 1\}$ and $M^\gamma(t-) = \sup_{t_i < t} \|\mathbf{x}^i\|_\infty < \|\mathbf{x}\|_\infty$. Taking the order into account, at time t we will see exactly $R^\gamma(M^\gamma(t-))$ records. Similarly, a point measure $\eta = \sum_{i=1}^\infty \delta_{t_i, x_i} \in \mathcal{M}_p([0, \infty) \times [0, \infty))$ has a record at time t with corresponding record value x , if $(t, x) \in \{(t_i, x_i) : i \geq 1\}$ and $\eta([0, t) \times [x, \infty)) = 0$.

To prove the representation of the limit, observe that $N'' = \sum_{i=1}^{\infty} \delta_{(T_i, P_i, Q_i)}$ has records at exactly the same time instances as the process $M_0 = \sum_{i=1}^{\infty} \delta_{(T_i, P_i)}$, since by the assumptions of the theorem and by definition of the sequences Q_i , all of their components are in $[0, 1]$ with one of them being exactly equal to 1. Because M_0 is a Poisson point process on $[0, \infty) \times (0, \infty]$ with intensity measure $dx \times d(-\theta y^{-\alpha})$, it has infinitely many points of in any set of the form $[a, b] \times [0, \epsilon]$ with $a < b$ and $\epsilon > 0$. Hence, one can a.s. write the record times of M_0 as a double sided sequence $\tau_n, n \in \mathbb{Z}$, such that $\tau_n < \tau_{n+1}$ for each n . Fix an arbitrary $s > 0$, and assume without loss of generality that τ_1 represents the first record time strictly greater than s , i.e. $\tau_1 = \inf\{\tau_i : \tau_i > s\}$. Denote the corresponding successive record values by U_n ; they clearly satisfy $U_n < U_{n+1}$ and $U_n = \sup_{T_i \leq \tau_n} P_i$ and $U_0 = \sup_{T_i \leq s} P_i$. According to [30, Proposition 4.9], $\sum_{n \in \mathbb{Z}} \delta_{\tau_n}$ is a Poisson point process with intensity $x^{-1} dx$ on $(0, \infty)$. Apply now [30, Proposition 4.7 (iv)] (note that U_0 corresponds to $Y(s)$ in the notation of that proposition) to prove that $\{U_n/U_{n-1}, n \geq 1\}$ is a sequence of i.i.d. random variables with a Pareto distribution with tail index α . Because the record times τ_n and record values U_n for $n \geq 1$ of the point process M_0 match the records of point process $N'' = \sum_{i=1}^{\infty} \delta_{(T_i, P_i, Q_i)}$ on the interval (s, ∞) , we just need to count how many of them appear at any give time τ_n which are larger than the previous record U_{n-1} . If, say, $\tau_n = T_i$, that number corresponds to the number of Q_{ij} 's which after multiplication by the corresponding $U_n = P_i$ represent a record larger than U_{n-1} . Hence, that random number has the same distribution as

$$\kappa = R^Q(U_1/U_0).$$

Recall that $s > 0$ was arbitrary. Now since the point process $\sum_{i=1}^{\infty} \delta_{(T_i, P_i)}$ and therefore sequence $\{U_n/U_{n-1}, n \geq 1\}$ is independent of the i.i.d. random elements $\{Q_i\}$ and since U_1/U_0 has a Pareto distribution with tail index α , the claim follows. \square

Example 5.3 For an illustration of the previous theorem, consider the moving average process of order 1

$$X_t = \xi_t + c\xi_{t-1},$$

for a sequence of i.i.d. nonnegative random variables $\{\xi_t, t \in \mathbb{Z}\}$ with regularly varying distribution and the tail index $\alpha > 0$. Assume further that $c > 1$. By (3.8), the sequence $\{Q_j\}$, as a random element in the space \tilde{l}_0 , is in this case equal to the deterministic sequence $\{\dots, 0, 1/c, 1, 0, \dots\}$. Intuitively speaking, in each cluster of extremely large values, there are exactly two successive extreme values with the second one c times larger than the first. Therefore, each such cluster can give rise to at most 2 records. By straightforward calculations, the random variables κ_i from Theorem 5.2 have the following distribution

$$\mathbb{P}(\kappa_i = 2) = \mathbb{P}(1/\zeta \leq 1/c) = \mathbb{P}(\zeta \geq c) = \frac{1}{c^\alpha} = 1 - \mathbb{P}(\kappa_i = 1).$$

6 Lemmas

6.1 Metric on the space \tilde{l}_0

Let (\mathbb{X}, d) be a metric space, we define the distance between $x \in \mathbb{X}$ and subset $B \subset \mathbb{X}$ by $d(x, B) = \inf\{d(x, y) : y \in B\}$. Let \sim be an equivalence relation on \mathbb{X} and let $\tilde{\mathbb{X}}$ be the induced quotient space. Define a function $\tilde{d} : \tilde{\mathbb{X}} \times \tilde{\mathbb{X}} \rightarrow [0, \infty)$ by:

$$\tilde{d}(\tilde{x}, \tilde{y}) = \inf\{d(x', y') : x' \in \tilde{x}, y' \in \tilde{y}\},$$

for all $\tilde{x}, \tilde{y} \in \tilde{\mathbb{X}}$.

Lemma 6.1 *Let (\mathbb{X}, d) be a complete separable metric space. Assume that for all $\tilde{x}, \tilde{y} \in \tilde{\mathbb{X}}$ and all $x, x' \in \tilde{x}$ we have*

$$d(x, \tilde{y}) = d(x', \tilde{y}). \tag{6.1}$$

Then \tilde{d} is a pseudo-metric which makes $\tilde{\mathbb{X}}$ a separable and complete pseudo-metric space.

Proof To prove that \tilde{d} is a pseudo-metric, the only nontrivial step is to show that \tilde{d} satisfies the triangle inequality, but that is implied by Condition (6.1). Separability is easy to check and it remains to prove that $(\tilde{\mathbb{X}}, \tilde{d})$ is complete.

Let $\{\tilde{x}_n\}$ be a Cauchy sequence in $(\tilde{\mathbb{X}}, \tilde{d})$. Then we can find a strictly increasing sequence of nonnegative integers $\{n_k\}$ such that

$$\tilde{d}(\tilde{x}_m, \tilde{x}_n) < \frac{1}{2^{k+1}},$$

for all integers $m, n \geq n_k$ and for every integer $k \geq 1$. We define a sequence of elements $\{y_n\}$ in \mathbb{X} inductively as follows:

- Let y_1 be an arbitrary element of \tilde{x}_{n_1} .
- For $k \geq 1$ let y_{k+1} be an element of $\tilde{x}_{n_{k+1}}$ such that $d(y_k, y_{k+1}) < \frac{1}{2^{k+1}}$. Such an y_{k+1} exists by Condition (6.1).

Then the sequence $\{y_n\}$ is a Cauchy sequence in (\mathbb{X}, d) . Indeed, for every $k \geq 1$ and for all $m, n \geq k$ we have that

$$d(y_m, y_n) \leq \sum_{l=m \wedge n}^{m \vee n - 1} d(y_l, y_{l+1}) < \sum_{l=k}^{\infty} \frac{1}{2^{l+1}} = \frac{1}{2^k}.$$

Since (\mathbb{X}, d) is complete, the sequence $\{y_n\}$ converges to some $x \in \mathbb{X}$. Let \tilde{x} be the equivalence class of x . It follows that the sequence $\{\tilde{x}_{n_k}\}$ converges to \tilde{x} because $\tilde{d}(\tilde{x}_{n_k}, \tilde{x}) \leq d(y_k, x)$ by definition of \tilde{d} . Finally, since $\{\tilde{x}_n\}$ is a Cauchy sequence, it follows easily that the whole sequence $\{\tilde{x}_n\}$ also converges to \tilde{x} , hence $(\tilde{\mathbb{X}}, \tilde{d})$ is complete. □

Proof of Lemma 2.1. Since we have $\|\theta^k \mathbf{x} - \theta^l \mathbf{y}\|_\infty = \|\theta^{k-l} \mathbf{x} - \mathbf{y}\|_\infty$ for all $\mathbf{x}, \mathbf{y} \in l_0$ and $k, l \in \mathbb{Z}$, it follows that

$$\tilde{d}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = \inf\{\|\theta^k \mathbf{x} - \mathbf{y}\|_\infty : k \in \mathbb{Z}\} = \inf\{\|\mathbf{x}' - \mathbf{y}\|_\infty : \mathbf{x}' \in \tilde{\mathbf{x}}\},$$

for all $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in \tilde{l}_0$, and all $\mathbf{x} \in \tilde{\mathbf{x}}, \mathbf{y} \in \tilde{\mathbf{y}}$. In view of Lemma 6.1 it only remains to show that \tilde{d} is a metric, rather than just a pseudo-metric.

Assume that $\tilde{d}(\tilde{\mathbf{x}}, \tilde{\mathbf{y}}) = 0$ for some $\tilde{\mathbf{x}}, \tilde{\mathbf{y}} \in \tilde{l}_0$. Then, for arbitrary $\mathbf{x} \in \tilde{\mathbf{x}}, \mathbf{y} \in \tilde{\mathbf{y}}$, there exists a sequence of integers $\{k_n\}$ such that $\|\theta^{k_n} \mathbf{x} - \mathbf{y}\|_\infty \rightarrow 0$, as $n \rightarrow \infty$. It suffices to show that the sequence $\{k_n\}$ is bounded. Indeed, by passing to a convergent subsequence it follows that there exists an integer k such that $\mathbf{y} = \theta^k \mathbf{x}$, hence $\tilde{\mathbf{x}} = \tilde{\mathbf{y}}$. Suppose now that the sequence $\{k_n\}$ is unbounded and that $\mathbf{y} \neq \mathbf{0}$ (the case when $\mathbf{y} = \mathbf{0}$ is trivial). Without loss of generality, we can assume that $k_n \rightarrow \infty$, as $n \rightarrow \infty$. Since $\mathbf{y} \neq \mathbf{0}$ and $\lim_{|i| \rightarrow \infty} \|y_i\| = 0$, there exists integers i_0 and $N > 0$ such that $\|y_{i_0}\| = \|\mathbf{y}\|_\infty > 0$ and $\|y_i\| < \|\mathbf{y}\|_\infty/4$ for $|i| \geq N$. Since $\|\theta^{k_n} \mathbf{x} - \mathbf{y}\|_\infty \rightarrow 0$ there exists an integer $n_0 > 0$ such that $\|\theta^{k_n} \mathbf{x} - \mathbf{y}\|_\infty < \|\mathbf{y}\|_\infty/4$ for $n \geq n_0$. By our assumption, we can find an integer $n \geq n_0$ such that $k_n - k_{n_0} + i_0 \geq N$, and it follows that

$$\frac{3}{4} \|\mathbf{y}\|_\infty < \|(\theta^{k_n} \mathbf{x})_{i_0}\| = \|x_{k_n+i_0}\| = \|(\theta^{k_{n_0}} \mathbf{x})_{k_n-k_{n_0}+i_0}\| < \frac{1}{2} \|\mathbf{y}\|_\infty,$$

which is a contradiction. Hence, the sequence $\{k_n\}$ is bounded. □

6.2 Assumption 3.5 is a consequence of β -mixing

The β -mixing coefficients of the sequence $\{X_j, j \in \mathbb{Z}\}$ are defined by

$$\beta_n = \frac{1}{2} \sup_{\mathcal{A}, \mathcal{B}} \sum_{i \in I} \sum_{j \in J} |\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)|,$$

where the supremum is taken over all finite partitions $\mathcal{A} = \{A_i, i \in I\}$ and $\mathcal{B} = \{B_j, j \in J\}$ such that the sets A_i are measurable with respect to $\sigma(X_k, k \leq 0)$ and the sets B_j are measurable with respect to $\sigma(X_k, k \geq n)$. See [32, Section 1.6].

Lemma 6.2 *Assume that the sequence $\{X_j\}$ is β -mixing with coefficients $\{\beta_j, j \in \mathbb{N}\}$. Assume that there exists a sequence r_n satisfying Assumption 1.1 and a sequence ℓ_n such that*

$$\lim_{n \rightarrow \infty} \frac{\ell_n}{r_n} = \lim_{n \rightarrow \infty} \frac{n}{r_n} \beta_{\ell_n} = 0. \tag{6.2}$$

Then Assumption 3.5 holds.

Proof Write $X_{i,j}$ for (X_i, \dots, X_j) . Set $k_n = \lfloor n/r_n \rfloor$ and let $\tilde{\mathbf{X}}_n$ be the vector of length $k_n(r_n - \ell_n)$ which concatenates all the subvectors $X_{(j-1)r_n+1, jr_n-\ell_n}, j = 1, \dots, k_n$.

Let $\tilde{\mathbf{X}}_n^*$ be the vector build with independent blocks $X_{(j-1)r_n+1, jr_n-\ell_n}^*$ which each have the same distribution has the corresponding original blocks $X_{(j-1)r_n+1, jr_n-\ell_n}$. Applying [17, Lemma 2] and (6.2), we obtain

$$d_{TV}(\mathcal{L}(\tilde{\mathbf{X}}_n), \mathcal{L}(\tilde{\mathbf{X}}_n^*)) \leq k_n \beta_{\ell_n} = o(1). \tag{6.3}$$

Set $\tilde{X}_{n,i} = a_n^{-1} X_{(j-1)r_n+1, jr_n-\ell_n}$ and $\tilde{X}_{n,i}^* = a_n^{-1} X_{(j-1)r_n+1, jr_n-\ell_n}^*$ and define the following point processes

$$\tilde{N}_n'' = \sum_{i=1}^{k_n} \delta_{(i/k_n, \tilde{X}_{n,i})}, \quad \tilde{N}_n^* = \sum_{i=1}^{k_n} \delta_{(i/k_n, X_{n,i}^*)}.$$

Let f be a nonnegative function defined on $[0, 1] \times \tilde{\ell}_0$. Since the exponential of a negative function is less than 1, by definition of the total variation distance, the bound (6.3) yields

$$\left| \mathbb{E} \left[e^{-\tilde{N}_n''(f)} \right] - \mathbb{E} \left[e^{-\tilde{N}_n^*(f)} \right] \right| \leq d_{TV}(\mathcal{L}(\tilde{\mathbf{X}}_n), \mathcal{L}(\tilde{\mathbf{X}}_n^*)) = o(1). \tag{6.4}$$

We must now check that the same limit holds with the full blocks instead of the truncated blocks. Under Assumption 1.1 (which holds for any sequence smaller than r_n hence for ℓ_n), we know by [6, Proposition 4.2] that for every $\epsilon > 0$ and every sequence $\{\ell_n\}$ such that $\ell_n \rightarrow \infty$ and $\ell_n \bar{F}(a_n) \rightarrow 0$,

$$\lim_{n \rightarrow \infty} \frac{\mathbb{P}(\max_{1 \leq i \leq \ell_n} \|X_i\| > \epsilon a_n)}{\ell_n \mathbb{P}(\|X_0\| > a_n)} = \theta \epsilon^{-\alpha}. \tag{6.5}$$

Then, applying (6.5) yields,

$$\begin{aligned} \mathbb{P} \left(\max_{1 \leq j \leq k_n} \max_{1 \leq i \leq \ell_n} \|X_{jr_n-i+1}\| > \epsilon a_n \right) &\leq k_n \mathbb{P} \left(\max_{1 \leq i \leq \ell_n} \|X_i\| > \epsilon a_n \right) \\ &= O(k_n \ell_n \mathbb{P}(\|X_0\| > a_n)) \\ &= O(\ell_n / r_n) = o(1). \end{aligned}$$

Assume now that f depends only on the components greater than some $\epsilon > 0$ in norm. Then $N_n''(f) = \tilde{N}_n''(f)$ unless at least one component at the end of one block is greater than ϵ in norm. This yields

$$\left| \mathbb{E}[e^{-N_n''(f)}] - \mathbb{E}[e^{-\tilde{N}_n''(f)}] \right| \leq \mathbb{P} \left(\max_{1 \leq j \leq k_n} \max_{1 \leq i \leq \ell_n} \|X_{jr_n-i+1}\| > \epsilon a_n \right) = o(1).$$

The same relation also holds for the independent blocks. Therefore, Assumption 3.5 holds. □

6.3 On continuity of addition in E

The next lemma gives sufficient conditions for continuity of addition in the space $E([0, 1], \mathbb{R})$.

Lemma 6.3 *Suppose that $\{x_n, n \in \mathbb{N}\}$ is a sequence in $D([0, 1], \mathbb{R})$ and $x' = (x, S, \{I(t) : t \in S\})$ an element in E such that $x_n \rightarrow x'$ in E . Suppose also that $\{b_n, n \in \mathbb{N}\}$ is a sequence in $D([0, 1], \mathbb{R})$ which converges uniformly to a continuous function b on $[0, 1]$. Then the sequence $\{x_n - b_n\}$ converges in (E, m_E) to an element $x' - b \in E$ defined by*

$$x' - b = (x - b, S, \{I(t) - b(t) : t \in S\}).$$

Proof Recall the definition of m_E given in (4.1). By Whitt [34, Theorem 15.5.1.] to show that $x_n - b_n \rightarrow x' - b$ in E , it suffices to prove that

$$\sup_{(t,z) \in \Gamma_{x_n - b_n}} \|(t, z) - \Gamma_{x' - b}\|_\infty \rightarrow 0. \tag{6.6}$$

Take an arbitrary $\epsilon > 0$. Note that b is uniformly continuous so by the conditions of the lemma there exists $0 < \delta \leq \epsilon$ and $n_0 \in \mathbb{N}$ such that

- (i) $|t - s| < \delta \Rightarrow |b(t) - b(s)| < \epsilon$,
- (ii) $m_E(x_n, x') < \delta$, for all $n \geq n_0$ and
- (iii) $|b_n(t) - b(t)| < \epsilon$, for all $t \in [0, 1]$.

Also, since b is continuous, it easily follows that $|b_n(t) - b_n(t-)| \leq 2\epsilon$ for all $n \geq n_0$ and $t \in [0, 1]$.

Take $n \geq n_0$ and a point $(t, z) \in \Gamma_{x_n - b_n}$, i.e.

$$z \in [(x_n(t-) - b_n(t-)) \wedge (x_n(t) - b_n(t)), (x_n(t-) - b_n(t-)) \vee (x_n(t) - b_n(t))].$$

Since $|b_n(t) - b_n(t-)| \leq 2\epsilon$ there exists $z' \in [x_n(t-) \wedge x_n(t), x_n(t-) \vee x_n(t)]$ (i.e. $(t, z') \in \Gamma_{x_n}$), such that

$$|(z' - b_n(t)) - z| \leq 2\epsilon.$$

Next, since $m_E(x_n, x') < \delta$, there exists a point $(s, y) \in \Gamma_{x'}$ such that

$$|s - t| \vee |y - z'| < \delta.$$

Note that $(s, y - b(s)) \in \Gamma_{x' - b}$ and by previous arguments

$$\begin{aligned} |(y - b(s)) - z| &= |(y - b(s)) - z + (z' - b_n(t)) - (z' - b_n(t)) + b(t) - b(t)| \\ &\leq |y - z'| + |b(t) - b(s)| + |b_n(t) - b(t)| + |(z' - b_n(t)) - z| \\ &\leq \delta + \epsilon + \epsilon + 2\epsilon \\ &\leq 5\epsilon. \end{aligned}$$

Also, $|s - t| < \delta \leq \epsilon$. Hence, for all $n \geq n_0$,

$$\sup_{(t,z) \in \Gamma_{x_n - b_n}} \|(t, z) - \Gamma_{x' - b}\|_\infty \leq 5\epsilon.$$

and since ϵ was arbitrary, (6.6) holds. □

6.4 A lemma for partial sum convergence in E

Lemma 6.4 *Let $\alpha \in [1, 2)$ and let the assumptions of Theorem 4.5 hold. Then there exists an α -stable Lévy process V on $[0, 1]$ such that, as $\epsilon \rightarrow 0$, the process V_ϵ defined in (4.28) converges uniformly a.s. (along some subsequence) to V .*

Proof Recall that

$$V_\epsilon(t) = \sum_{T_i \leq t} s^\epsilon(P_i; Q_i) - t \int_{\{x: \epsilon < |x| \leq 1\}} x\mu(dx)$$

where

$$\mu(dx) = p\alpha x^{-\alpha-1} \mathbb{1}_{(0, \infty)}(x)dx + (1 - p)\alpha(-x)^{-\alpha-1} \mathbb{1}_{(-\infty, 0)}(x)dx$$

for $p = \mathbb{P}(\Theta_0 = 1)$. We first show that the centering term can be expressed as an expectation of a functional of the limiting point process N'' . More precisely, we show that for all $\epsilon > 0$

$$\int_{\{x: \epsilon < |x| \leq 1\}} x\mu(dx) = \theta \int_0^\infty \mathbb{E} \left[y \sum_{j \in \mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon < y|Q_j| \leq 1\}} \right] \alpha y^{-\alpha-1} dy. \tag{6.7}$$

First, as shown in [14, Theorem 3.2, Equation (3.13)] it holds that

$$\theta \mathbb{E} \left[\sum_{j \in \mathbb{Z}} Q_j |Q_j|^{\alpha-1} \right] = 2p - 1, \tag{6.8}$$

so by Fubini’s theorem, if $\alpha > 1$

$$\begin{aligned} \theta \int_0^\infty \mathbb{E} \left[y \sum_{j \in \mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon < y|Q_j| \leq 1\}} \right] \alpha y^{-\alpha-1} dy &= \alpha \theta \mathbb{E} \left[\sum_{j \in \mathbb{Z}} Q_j \int_{\epsilon|Q_j|^{-1}}^{|Q_j|^{-1}} y^{-\alpha} dy \right] \\ &= \frac{\alpha}{\alpha - 1} (\epsilon^{-\alpha+1} - 1) \theta \mathbb{E} \left[\sum_{j \in \mathbb{Z}} Q_j |Q_j|^{\alpha-1} \right] \\ &= \frac{\alpha}{\alpha - 1} (\epsilon^{-\alpha+1} - 1) (2p - 1), \end{aligned}$$

and if $\alpha = 1$ the same term equals $\log(\epsilon^{-1})(2p - 1)$. Note that the use of Fubini's theorem is justified since the same calculation as above shows that the above integral converges absolutely since $\mathbb{E}[\sum_{j \in \mathbb{Z}} |Q_j|^\alpha] < \infty$. The equality in (6.7) now follows by the definition of the measure μ . Hence, for all $t \in [0, 1]$

$$V_\epsilon(t) = \sum_{T_i \leq t} s^\epsilon(P_i Q_i) - t\theta \int_0^\infty \mathbb{E} \left[y \sum_{j \in \mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon < y|Q_j| \leq 1\}} \right] \alpha y^{-\alpha-1} dy.$$

Recall from Remark 4.6 that we can define $W = \sum_{j \in \mathbb{Z}} |Q_j|$, $W_i = \sum_{j \in \mathbb{Z}} |Q_{i,j}|$ so that $\{W_i, i \geq 1\}$ is a sequence of i.i.d. random variables with the same distribution as W and $\mathbb{E}[W_i^\alpha] < \infty$ and that $\sum_{i=1}^\infty \delta_{P_i W_i}$ is a Poisson point process on $(0, \infty)$ with intensity measure $\theta \mathbb{E}[W_1^\alpha] \alpha y^{-\alpha-1} dy$. In particular, for every $\delta > 0$ there are almost surely at most finitely many points $P_i W_i$ such that $P_i W_i > \delta$. For $\delta, \epsilon > 0$, define

$$m_{\epsilon, \delta} = \theta \int_0^\infty \mathbb{E} \left[y \sum_{j \in \mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon < y|Q_j| \leq 1, \delta < yW\}} \right] \alpha y^{-\alpha-1} dy.$$

Note that $\lim_{\epsilon \rightarrow 0} m_{\epsilon, \delta} = m_{0, \delta}$ for all $\delta > 0$ by the dominated convergence theorem. Indeed, if $\alpha = 1$ we have that

$$\begin{aligned} & \theta \int_0^\infty \mathbb{E} \left[y \sum_{j \in \mathbb{Z}} |Q_j| \mathbb{1}_{\{y|Q_j| \leq 1, \delta < yW\}} \right] \alpha y^{-\alpha-1} dy \leq \theta \mathbb{E} \left[\sum_{j \in \mathbb{Z}} |Q_j| \int_{\frac{\delta \wedge 1}{W}}^{\frac{1}{|Q_j|}} y^{-1} dy \right] \\ & = \theta \mathbb{E} \left[\sum_{j \in \mathbb{Z}} |Q_j| \log(|Q_j|^{-1}) + W \log W + \log((\delta \wedge 1)^{-1} W) \right], \end{aligned}$$

which is finite by assumption (4.10), and if $\alpha > 1$ similar calculation using the assumption $\mathbb{E}[W^\alpha] < \infty$ justifies the use of the dominated convergence theorem.

Since for every $\delta > 0$ there a.s. exists at most finitely many points $P_i W_i$ such that $P_i W_i > \delta$, for every $\epsilon \geq 0$ we can define the process $V_{\epsilon, \delta}$ in $D[0, 1]$ by

$$\begin{aligned} V_{\epsilon, \delta}(t) &= \sum_{T_i \leq t} s^\epsilon(P_i Q_i) \mathbb{1}_{\{\delta < P_i W_i\}} - t m_{\epsilon, \delta} \\ &= \sum_{T_i \leq t} \sum_{j \in \mathbb{Z}} P_i Q_{i,j} \mathbb{1}_{\{\epsilon < P_i |Q_{i,j}|, \delta < P_i W_i\}} - t m_{\epsilon, \delta}. \end{aligned}$$

Furthermore, for every fixed $\delta > 0$, as $\epsilon \rightarrow 0$, $V_{\epsilon, \delta}$ converges uniformly almost surely to $V_{0, \delta}$.

Next, we prove that for any positive sequence $\{\delta_k\}$ with $\delta_k \searrow 0$ as $k \rightarrow \infty$, V_{0, δ_k} converges uniformly almost surely to a process V in $D([0, 1])$. Note first that by [14, Theorem 3.1] the finite dimensional distributions of $V_{0, \delta}$ converge to those of an α -stable Lévy process.

Since $\sum_{i \geq 1} \delta_{T_i, P_i, Q_i}$ is a Poisson point process on $[0, 1] \times (0, \infty) \times \tilde{I}_0$, the process $V_{0,\delta}$ has independent increments with respect to δ , that is for every $\delta < \delta'$, $V_{0,\delta} - V_{0,\delta'}$ is independent of $V_{0,\delta'}$. Moreover, since $V_{0,\delta} - V_{0,\delta'}$ is a Poisson integral, we have that

$$\begin{aligned} \text{var}(V_{0,\delta}(1) - V_{0,\delta'}(1)) &= \theta \int_0^\infty y^2 \mathbb{E} \left[\left(\sum_{j \in \mathbb{Z}} Q_j \right)^2 \mathbb{1}_{\{\delta < yW \leq \delta'\}} \right] \alpha y^{-\alpha-1} dy \\ &\leq \theta \mathbb{E} \left[W^2 \int_0^{\delta'/W} \alpha y^{-\alpha+1} dy \right] = \frac{\theta \alpha (\delta')^{2-\alpha}}{(2-\alpha)} \mathbb{E}[W^\alpha]. \end{aligned}$$

Therefore, $\lim_{\delta' \rightarrow 0} \text{var}(V_{0,\delta}(1) - V_{0,\delta'}(1)) = 0$ and now arguing exactly as in the proof of [31, Proposition 5.7, Property 2] shows that for any positive sequence $\{\delta_k\}$ with $\delta_k \searrow 0$, $\{V_{0,\delta_k}\}$ is almost surely a Cauchy sequence in $D([0, 1])$ with respect to the supremum metric $\|\cdot\|_\infty$. Since the space $D([0, 1])$ is complete under this metric, we obtain the existence of the process $V = \{V(t), t \in [0, 1]\}$ with paths in $D([0, 1])$ almost surely and such that $\lim_{k \rightarrow \infty} \|V_{0,\delta_k} - V\|_\infty = 0$ almost surely.

There only remains to prove that for all $u > 0$,

$$\lim_{\delta \rightarrow 0} \limsup_{\epsilon \rightarrow 0} \mathbb{P}(\|V_\epsilon - V_{\epsilon,\delta}\|_\infty > u) = 0. \tag{6.9}$$

Indeed, this would imply that $\|V_\epsilon - V\|_\infty \rightarrow 0$ in probability and hence that, along some subsequence, V_ϵ converges to V uniformly almost surely. Since for $\delta \leq 1$, $yW = \sum_{j \in \mathbb{Z}} y|Q_j| \leq \delta$ implies that $y|Q_j| \leq \delta \leq 1$ for all $j \in \mathbb{Z}$, we have that

$$\begin{aligned} V_\epsilon(t) - V_{\epsilon,\delta}(t) &= \sum_{T_i \leq t} \sum_{j \in \mathbb{Z}} P_i Q_{i,j} \mathbb{1}_{\{\epsilon < P_i |Q_{i,j}|, P_i W_i \leq \delta\}} \\ &\quad - t\theta \int_0^\infty \mathbb{E} \left[y \sum_{j \in \mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon < y|Q_j|, yW \leq \delta\}} \right] \alpha y^{-\alpha-1} dy. \end{aligned}$$

The process $V_\epsilon - V_{\epsilon,\delta}$ is a càdlàg martingale, thus applying Doob-Meyer’s inequality yields

$$\begin{aligned} \mathbb{P}(\|V_\epsilon - V_{\epsilon,\delta}\|_\infty > u) &\leq u^{-2} \text{var}(V_\epsilon(1) - V_{\epsilon,\delta}(1)) \\ &= u^{-2} \theta \int_0^\infty y^2 \mathbb{E} \left[\left(\sum_{j \in \mathbb{Z}} Q_j \mathbb{1}_{\{\epsilon < y|Q_j|, yW \leq \delta\}} \right)^2 \right] \alpha y^{-\alpha-1} dy \\ &\leq u^{-2} \theta \mathbb{E} \left[W^2 \int_0^{\delta/W} \alpha y^{-\alpha+1} dy \right] = \frac{\theta \alpha \delta^{2-\alpha}}{u^2 (2-\alpha)} \mathbb{E}[W^\alpha] \end{aligned}$$

and hence (6.9) holds. □

Acknowledgements Parts of this paper were written when Bojan Basrak visited the Laboratoire MODAL'X at Université Paris Nanterre. Bojan Basrak takes pleasures in thanking MODAL'X and for excellent hospitality and financial support, as well as Johan Segers for useful discussions over the years. The work of Bojan Basrak and Hrvoje Planinić has been supported in part by Croatian Science Foundation under the project 3526. The work of Philippe Soulier was partially supported by LABEX MME-DII.

References

1. Arratia, R.: On the central role of scale invariant Poisson processes on $(0, \infty)$. In: *Microsurveys in discrete probability* (Princeton, NJ, 1997), vol. 41 of DIMACS Ser. Discrete Math. Theoret. Comput. Sci., pp. 21–41. Am. Math. Soc., Providence, RI, (1998)
2. Avram, F., Taqqu, M.S.: Weak convergence of sums of moving averages in the α -stable domain of attraction. *Ann. Probab.* **20**(1), 483–503 (1992)
3. Balan, R., Jakubowski, A., Louhichi, S.: Functional convergence of linear processes with heavy-tailed innovations. *J. Theor. Probab.* **29**(2), 491–526 (2016)
4. Basrak, B., Krizmanić, D.: A limit theorem for moving averages in the α -stable domain of attraction. *Stoch. Process. Appl.* **124**(2), 1070–1083 (2014)
5. Basrak, B., Krizmanić, D., Segers, J.: A functional limit theorem for dependent sequences with infinite variance stable limits. *Ann. Probab.* **40**(5), 2008–2033 (2012)
6. Basrak, B., Segers, J.: Regularly varying multivariate time series. *Stoch. Process. Appl.* **119**(4), 1055–1080 (2009)
7. Basrak, B., Tafro, A.: A complete convergence theorem for stationary regularly varying multivariate time series. *Extremes* **19**(3), 549–560 (2016)
8. Billingsley, P.: *Convergence of Probability Measures*. Wiley, New York (1968)
9. Bingham, N.H., Goldie, C.M., Teugels, J.L.: *Regular variation*, Encyclopedia of Mathematics and its Applications, vol. 27. Cambridge University Press, Cambridge (1989)
10. Bradley, R.C.: *Introduction to Strong Mixing Conditions*, vol. 1. Kendrick Press, Heber City (2007)
11. Daley, D.J., Vere-Jones, D.: *An Introduction to the Theory of Point Processes*. Vol. I. Probability and its Applications: Elementary Theory and Methods, 2nd edn. Springer, New York (2003)
12. Daley, D.J., Vere-Jones, D.: *An Introduction to the Theory of Point Processes*. Vol. II. Probability and its Applications: General Theory and Structure, 2nd edn. Springer, New York (2008)
13. Davis, R., Resnick, S.: Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **13**(1), 179–195 (1985)
14. Davis, R.A., Hsing, T.: Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23**(2), 879–917 (1995)
15. Davis, R.A., Mikosch, T.: Extreme value theory for space-time processes with heavy-tailed distributions. *Stoch. Process. Appl.* **118**(4), 560–584 (2008)
16. de Haan, L., Lin, T.: On convergence toward an extreme value distribution in $C[0, 1]$. *Ann. Probab.* **29**(1), 467–483 (2001)
17. Eberlein, E.: Weak convergence of partial sums of absolutely regular sequences. *Stat. Probab. Lett.* **2**(5), 291–293 (1984)
18. Hult, H., Lindskog, F.: Regular variation for measures on metric spaces. *Publ. Inst. Math. (Beograd) (N.S.)* **80**(94), 121–140 (2006)
19. Hult, H., Samorodnitsky, G.: Large deviations for point processes based on stationary sequences with heavy tails. *J. Appl. Probab.* **47**(1), 1–40 (2010)
20. Kallenberg, O.: *Random Measures*. Akademie-Verlag, Berlin (1983)
21. Kallenberg, O.: *Random Measures, Theory and Applications*, vol. 77. Springer, Berlin (2017)
22. Lindskog, F., Resnick, S.I., Roy, J.: Regularly varying measures on metric spaces: hidden regular variation and hidden jumps. *Probab. Surv.* **11**, 270–314 (2014)
23. Meinguet, T., Segers, J.: Regularly varying time series in Banach spaces. [arXiv:1001.3262](https://arxiv.org/abs/1001.3262), (2010)
24. Merlevède, F., Peligrad, M., Utev, S.: Recent advances in invariance principles for stationary sequences. *Probab. Surv.* **3**, 1–36 (2006)
25. Mikosch, T., Samorodnitsky, G.: The supremum of a negative drift random walk with dependent heavy-tailed steps. *Ann. Appl. Probab.* **10**(3), 1025–1064 (2000)
26. Mikosch, T., Wintenberger, O.: Precise large deviations for dependent regularly varying sequences. *Probab. Theory Relat. Fields* **156**(3–4), 851–887 (2013)

27. Mikosch, T., Wintenberger, O.: The cluster index of regularly varying sequences with applications to limit theory for functions of multivariate Markov chains. *Probab. Theory Relat Fields* **159**(1–2), 157–196 (2014)
28. Mikosch, T., Wintenberger, O.: A large deviations approach to limit theorem for heavy-tailed time series. *Probab. Theory Relat Fields* **166**, 233–269 (2016)
29. Planinić, H., Soulier, P.: The tail process revisited. *ArXiv e-prints*, (June 2017)
30. Resnick, S.I.: *Extreme Values, Regular Variation and Point Processes*. Applied Probability, vol. 4. Springer, New York (1987)
31. Resnick, S.I.: *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York (2007)
32. Rio, E.: *Théorie asymptotique des processus aléatoires faiblement dépendants*, *Mathématiques & Applications (Berlin) [Mathematics and Applications]*, vol. 31. Springer, Berlin (2000)
33. Roueff, F., Soulier, P.: Convergence to stable laws in the space D . *J. Appl. Probab.* **52**(1), 1–17 (2015)
34. Whitt, W.: *Stochastic-Process Limits*. Springer, New York (2002)
35. Zhao, Y.: Point processes in a metric space. *arXiv preprint* [arXiv:1611.06995](https://arxiv.org/abs/1611.06995), (2016)