



HAL
open science

The tail empirical process of regularly varying functions of geometrically ergodic Markov chains

Rafal Kulik, Philippe Soulier, Olivier Wintenberger

► **To cite this version:**

Rafal Kulik, Philippe Soulier, Olivier Wintenberger. The tail empirical process of regularly varying functions of geometrically ergodic Markov chains. 2015. hal-01228825v1

HAL Id: hal-01228825

<https://hal.science/hal-01228825v1>

Preprint submitted on 16 Nov 2015 (v1), last revised 21 Sep 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

The tail empirical process of regularly varying functions of geometrically ergodic Markov chains

Rafał Kulik* Philippe Soulier† Olivier Wintenberger‡

Abstract

We consider a stationary regularly varying time series which can be expressed as a function of a geometrically ergodic Markov chain. We obtain practical conditions for the weak convergence of weighted versions of the multivariate tail empirical process. These conditions include the so-called geometric drift or Foster-Lyapunov condition and can be easily checked for most usual time series models with a Markovian structure. We illustrate these conditions on several models and statistical applications.

1 Introduction

Let $\{X_j, j \in \mathbb{Z}\}$ be a stationary, regularly varying univariate time series with marginal distribution function F and tail index α . This means that for each integer $h \geq 0$, there exists a non zero Radon measure $\nu_{0,h}$ on $\bar{\mathbb{R}}^{h+1} \setminus \{\mathbf{0}\}$ such that $\nu_{0,h}(\bar{\mathbb{R}}^{h+1} \setminus \mathbb{R}^{h+1}) = 0$ and

$$\lim_{t \rightarrow \infty} \frac{\mathbb{P}((X_0, \dots, X_h) \in tA)}{\mathbb{P}(X_0 > t)} = \nu_{0,h}(A),$$

for all relatively compact sets $A \in \bar{\mathbb{R}}^{h+1} \setminus \{\mathbf{0}\}$ satisfying $\nu_{0,h}(\partial A) = 0$. The measure $\nu_{0,h}$, called the exponent measure of (X_0, \dots, X_h) , is homogeneous with index $-\alpha$, i.e. $\nu_{0,h}(tA) = t^{-\alpha} \nu_{0,h}(A)$. This definition implies that $\nu_{0,h}((1, \infty) \times \mathbb{R}^h) = 1$. The purpose of this paper is to investigate statistical tools appropriate for the estimation of extremal quantities which can be derived from these exponent measures. The most important tool is the tail empirical process which we define now.

Let $\mathbf{X}_{i,j}$, $i \leq j$, denote the vector (X_i, \dots, X_j) . Let $\{u_n\}$ be an increasing sequence such that

$$\lim_{n \rightarrow \infty} \bar{F}(u_n) = \lim_{n \rightarrow \infty} \frac{1}{n \bar{F}(u_n)} = 0.$$

*University of Ottawa

†Université de Paris-Ouest

‡University of Copenhagen and Sorbonne Universités, UPMC University Paris 06

We define the (upper quadrant) tail empirical distribution (TED) function \tilde{M}_n by

$$\tilde{M}_n(\mathbf{v}) = \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^n 1_{\{\mathbf{X}_{j,j+h} \in u_n[-\infty, \mathbf{v}]^c\}} , \quad \mathbf{v} \in (\mathbf{0}, \infty) .$$

Let $M_n(\mathbf{v}) = \mathbb{E}[\tilde{M}_n(\mathbf{v})]$ and

$$\mathbb{M}_n(\mathbf{v}) = \sqrt{n\bar{F}(u_n)} \left\{ \tilde{M}_n(\mathbf{v}) - M_n(\mathbf{v}) \right\} , \quad \mathbf{v} \in (\mathbf{0}, \infty) .$$

In statistical applications, it is often useful to consider a weighted version of the tail empirical process (TEP). For a measurable function ψ defined on \mathbb{R}^{h+1} , define, for $\mathbf{v} \in (\mathbf{0}, \infty)$,

$$\tilde{M}_n^\psi(\mathbf{v}) = \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^n \psi \left(\frac{\mathbf{X}_{j,j+h}}{u_n} \right) 1_{(-\infty, \mathbf{v}]^c}(\mathbf{X}_{j,j+h}/u_n) , \quad (1.1)$$

$M_n^\psi(\mathbf{v}) = \mathbb{E}[\tilde{M}_n^\psi(\mathbf{v})]$ and

$$\mathbb{M}_n^\psi(\mathbf{v}) = \sqrt{n\bar{F}(u_n)} \left\{ \tilde{M}_n^\psi(\mathbf{v}) - M_n^\psi(\mathbf{v}) \right\} , \quad (1.2)$$

The investigation of the asymptotic behaviour of \mathbb{M}_n has a long and well known story and no longer necessitates any justification. See [Roo09] for references in the i.i.d. and weakly dependent univariate case. Naturally, when dealing with weakly dependent time series, some form of mixing condition is needed. The most convenient is absolute regularity or β -mixing, which allows to easily apply the blocking method. Many but not all time series models can be β -mixing under not too stringent conditions such as innovations with an absolutely continuous distribution. Notable exceptions come from integer valued time series or more generally, time series with a discrete valued input. It also excludes all long memory times series which require ad-hoc methods. See e.g. [KS11]. It is not the purpose of this paper to discuss these models.

In standard statistical problems, the β -mixing condition with a certain rate and moment conditions suffice to derive asymptotic distributions. But in extreme value theory for dependent data, further conditions are necessary. The most important one is the so-called anticlustering condition, introduced by [Smi92] as a sufficient condition for the extremal index of a time series to be positive. In the univariate case, it reads

$$\lim_{m \rightarrow \infty} \limsup_{n \rightarrow \infty} \sum_{j=m}^{r_n} \frac{\mathbb{P}(X_0 > u_n, X_j > u_n)}{\mathbb{P}(X_0 > u_n)} = 0 , \quad (1.3)$$

where r_n is an increasing sequence such that $r_n \mathbb{P}(X > u_n) \rightarrow 0$. Unfortunately, (1.3) is not implied by any temporal weak dependence condition, and is notably difficult to check. It has been checked in the literature by ad-hoc methods for several models. We refer to

[Roo09] for examples and further references. See also Section 3. Needless to say, only the simplest models have been investigated, and more complex time series such as threshold models remain to be studied in an extreme value context.

In the case of Markov chains, or functions of Markov chains, it was first (implicitly) proved by [RRSS06] that the so-called Foster-Lyapunov or geometric drift condition implies the anticlustering condition (1.3). It was later used by [MW13] to obtain large deviations and weak convergence to stable laws for heavy tailed functions of Markov chains. Let us mention, though we will not use this property here that the drift condition can also be used to check the asymptotic negligibility of small jumps in the case $1 \leq \alpha < 2$. It is well-known from the theory of Markov chains that this drift condition and irreducibility together imply β -mixing with geometric decay of the β -mixing coefficients. This in turn allows to applying the blocking technique, without any significant restriction on the number of order statistics involved in the definition of the extreme value statistics.

The main purpose of this paper is to show that the geometric drift condition can be used to prove weighted versions of the anticlustering condition, and ultimately to prove functional central limit theorems for the weighted, multivariate versions of the tail empirical process introduced above. We only consider finite dimensional tail empirical processes, that is, we do not study the full theory of cluster functionals developed in a general context by [DR10]. Such a theory is beyond the scope of the present paper. Still, our main result provides a tool for the investigation of extreme value statistics of all the time series which can be expressed as functions of irreducible, geometrically ergodic Markov chains.

The paper is organized as follows. In Section 2 we state our assumptions, including the geometric drift condition, and main result on weak convergence of the weighted tail empirical process of a function of a geometrically ergodic Markov chain. In Section 3, we illustrate the efficiency of our assumptions by studying two models, and also provide a counterexample which shows that geometric ergodicity, if not necessary, cannot be easily dispensed with. In Section 4, we illustrate our main theorem with some standard and less-standard statistical applications. The proof of the main result is in Section 5. The most important (and original) ingredient is, as already mentioned, to prove that the geometric drift condition implies a weighted anticlustering condition. This anticlustering condition (and the other assumptions) allow to apply the very general results of [DR10].

2 Weak convergence of the tail empirical process

Our context is a slight extension of the one in [MW13]. We now assume that $\{X_j, j \in \mathbb{N}\}$ is a function of a stationary Markov chain $\{\mathbb{Y}_j, j \in \mathbb{N}\}$, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with values in a measurable space (E, \mathcal{E}) . That is, there exists a measurable real valued function g such that $X_j = g(\mathbb{Y}_j)$.

Assumption 1. • *The Markov chain $\{\mathbb{Y}_j, j \in \mathbb{Z}\}$ is strictly stationary under \mathbb{P} .*

- *The sequence $\{X_j, j \in \mathbb{Z}\}$ defined by $X_j = g(\mathbb{Y}_j)$ is regularly varying with tail index $\alpha > 0$.*

- There exist a measurable function $V : E \rightarrow [1, \infty)$, $\gamma \in (0, 1)$, $x_0 \geq 1$ and $b > 0$ such that for all $y \in E$,

$$\mathbb{E}[V(\mathbb{Y}_1) \mid \mathbb{Y}_0 = y] \leq \gamma V(y) + b \mathbf{1}_{\{V(y) \leq x_0\}}. \quad (2.1)$$

- There exist an integer $m \geq 1$ and for all $x \geq x_0$, there exists a probability measure ν on (E, \mathcal{E}) and $\epsilon > 0$ such that, for all $y \in \{V \leq x\}$ and all measurable sets $B \in \mathcal{E}$,

$$\mathbb{P}(\mathbb{Y}_m \in B \mid \mathbb{Y}_0 = y) \geq \epsilon \nu(B). \quad (2.2)$$

- There exist $q \in (0, \alpha/2)$ and a constant c such that

$$|g|^q \leq cV. \quad (2.3)$$

- For every compact set $[a, b] \subset (0, \infty)$,

$$\limsup_{n \rightarrow \infty} \sup_{a \leq s \leq b} \frac{1}{u_n^q \bar{F}(u_n)} \mathbb{E} [V(\mathbb{Y}_0) \mathbf{1}_{\{s u_n < g(\mathbb{Y}_0)\}}] < \infty. \quad (2.4)$$

We will comment on these conditions in Section 2.1. We define formally the limiting covariances whose existence will be guaranteed by the assumptions of the theorem. In order to avoid trivialities, we assume that $\nu_{0,h}(\mathbb{R}_+^h) > 0$ for all $h \geq 0$.

$$\begin{aligned} c_j(\mathbf{v}, \mathbf{w}) &= \int_{\mathbb{R}^{h+j+1}} \mathbf{1}_{(-\infty, \mathbf{v}]^c}(\mathbf{x}_{0,h}) \mathbf{1}_{(-\infty, \mathbf{w}]^c}(\mathbf{x}_{j,j+h}) \nu_{0,j+h}(d\mathbf{x}), \\ c_j^\psi(\mathbf{v}, \mathbf{w}) &= \int_{\mathbb{R}^{j+h+1}} \psi(\mathbf{x}_{0,h}) \psi(\mathbf{x}_{j,j+h}) \mathbf{1}_{(-\infty, \mathbf{v}]^c}(\mathbf{x}_{0,h}) \mathbf{1}_{(-\infty, \mathbf{w}]^c}(\mathbf{x}_{j,j+h}) \nu_{0,j+h}(d\mathbf{x}), \\ C(\mathbf{v}, \mathbf{w}) &= c_0(\mathbf{v}, \mathbf{w}) + \sum_{j=1}^{\infty} \{c_j(\mathbf{v}, \mathbf{w}) + c_j(\mathbf{w}, \mathbf{v})\}, \end{aligned} \quad (2.5)$$

$$C^\psi(\mathbf{v}, \mathbf{w}) = c_0^\psi(\mathbf{v}, \mathbf{w}) + \sum_{j=1}^{\infty} \{c_j^\psi(\mathbf{v}, \mathbf{w}) + c_j^\psi(\mathbf{w}, \mathbf{v})\}. \quad (2.6)$$

Set $\mathbf{1} = (1, \dots, 1)$.

Theorem 2.1. *Let Assumption 1 hold and assume moreover that there exists $\eta > 0$ such that*

$$\lim_{n \rightarrow \infty} \log^{1+\eta}(n) \left\{ \bar{F}(u_n) + \frac{1}{\sqrt{nF(u_n)}} \right\} = 0. \quad (2.7)$$

Let $s_0 > 0$ be fixed.

- The process \mathbb{M}_n^ψ converges weakly in $\ell^\infty([s_0 \mathbf{1}, \infty))$ to a centered Gaussian process \mathbb{M}^ψ with covariance function C^ψ defined in (2.5).
- If $\psi : \mathbb{R}^{h+1} \rightarrow \mathbb{R}$ is such that

$$|\psi(\mathbf{x})| \leq \aleph((|x_0| \vee 1)^{q_0} + \dots + (|x_h| \vee 1)^{q_h}), \quad (2.8)$$

with $q_i + q_{i'} \leq q < \alpha/2$ for all $i, i' = 0, \dots, h$, then \mathbb{M}_n^ψ converges weakly to a centered Gaussian process \mathbb{M}^ψ with covariance function C^ψ defined in (2.6).

2.1 Comments

(C1) Under Assumption 1, it is well known that the chain $\{\mathbb{Y}_j\}$ is irreducible and geometrically ergodic. This implies that the chain $\{\mathbb{Y}_j\}$ and the sequence $\{X_j\}$ are β -mixing and there exists $c > 1$ such that $\beta_n = O(e^{-cn})$; see [Bra05, Theorem 3.7]. This is a very strong requirement, however, it is satisfied by many usual time series models, under standard conditions. See Section 3. Moreover, the geometric drift condition has the following consequences.

- Let the stationary distribution of the chain $\{\mathbb{Y}_j\}$ be denoted by π . Then the drift condition implies that $\pi(V) < \infty$.
- It was proved in [MW13] that the geometric drift condition implies the anti-clustering condition (1.3). Under this anti-clustering condition, it is well known in particular that the extremal index of the sequence $\{X_j\}$ is positive. See [BS09, Proposition 4.2]. The geometric drift condition is not necessary for this anti-clustering condition to hold, but when the chain is not geometrically ergodic, nearly any asymptotic behavior of the tail empirical process is possible. See Section 3.3.

(C2) Condition (2.4) is an ad-hoc condition which has to be checked for each example. It is implied by the stronger condition

$$\limsup_{n \rightarrow \infty} \sup_{a \leq s \leq b} \frac{1}{u_n^q \bar{F}(u_n)} \mathbb{E} [V(\mathbb{Y}_0) 1_{\{su_n^q < V(\mathbb{Y}_0)\}}] < \infty. \quad (2.9)$$

It holds for instance if \mathbb{Y}_0 takes values in \mathbb{R}^d , is regularly varying with index α and $V(\mathbb{Y}_0) = 1 + \|\mathbb{Y}_0\|^q$, $g(\mathbb{Y}_0) = \mathbb{Y}_0^{(1)}$ (the first component of \mathbb{Y}_0).

In order to prove Theorem 2.1, we will use a weighted form of the classical anticlustering condition mentioned above. Precisely, for sequences $\{u_n\}$ and $\{r_n\}$, we will say that Condition $\mathcal{S}(u_n, r_n, \psi)$ holds if for every pair $\mathbf{v}, \mathbf{w} \in (\mathbf{0}, \infty)$,

$$\lim_{L \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{1}{\bar{F}(u_n)} \sum_{L < |j| \leq r_n} \mathbb{E} [|\psi(\mathbf{X}_{0,h}/u_n)| |\psi(\mathbf{X}_{j,j+h}/u_n)| 1_{[\infty, \mathbf{v}]^c}(\mathbf{X}_{0,h}/u_n) 1_{[-\infty, \mathbf{w}]^c}(\mathbf{X}_{j,j+h}/u_n)] = 0. \quad (\mathcal{S}(u_n, r_n, \psi))$$

In Lemma 5.3, we will prove that Assumption 1 and condition (2.8) on the function ψ imply this weighted anti-clustering condition. The proof of this result is rather straightforward but lengthy and we postpone it to Section 5. Here, to illustrate it, we prove that $\mathcal{S}(u_n, r_n, \psi)$ implies that the series in (2.6) is summable.

Lemma 2.2. *Assume that the sequence $\{X_j\}$ is regularly varying. Assume moreover that (2.8) and $\mathcal{S}(u_n, r_n, \psi)$ hold, then, for all $\mathbf{v}, \mathbf{w} \in (\mathbf{0}, \infty)$, the series $\sum_{j=1}^{\infty} |c_j^\psi(\mathbf{v}, \mathbf{w})|$ is summable.*

Proof. Fix positive integers $R > L \geq 1$ and set

$$\psi_{n,j}(\mathbf{v}) = \psi(\mathbf{X}_{j,j+h}/u_n) 1_{[\infty, \mathbf{v}]^c}(\mathbf{X}_{j,j+h}/u_n).$$

Then, by regular variation, (2.8) and since $2q < \alpha$,

$$\sum_{j=L}^R |c_j^\psi(\mathbf{v}, \mathbf{w})| = \lim_{n \rightarrow \infty} \sum_{j=L}^R \frac{\mathbb{E}[\psi_{n,0}(\mathbf{v})\psi_{n,j}(\mathbf{w})]}{\bar{F}(u_n)}.$$

Fix $\epsilon > 0$. Applying $\mathcal{S}(u_n, r_n, \psi)$, we can choose L such that, for every fixed $R \geq L$

$$\lim_{n \rightarrow \infty} \sum_{j=L}^R \frac{\mathbb{E}[\psi_{n,0}(\mathbf{v})\psi_{n,j}(\mathbf{w})]}{\bar{F}(u_n)} \leq \epsilon.$$

This yields that for every $\epsilon > 0$, for large enough L and all $R \geq L$, $\sum_{j=L}^R |c_j^\psi(\mathbf{v}, \mathbf{w})| \leq \epsilon$ and this precisely means that the series $\sum_{j=1}^\infty |c_j^\psi(\mathbf{v}, \mathbf{w})|$ is summable. \square

3 Two models and a counterexample

The convergence of the tail empirical process has been considered in the literature under mixing assumptions and additional conditions which have been checked for a few specific models such as solutions of stochastic recurrence equations (including GARCH processes) and linear processes. See e.g. [Dre00], [Dre03] and [DM09]. Our main result provides a simple condition for functions of geometrically ergodic Markov chains which include many usual time series models. In the following two subsections, we will prove that the assumptions of Theorem 2.1 hold (and are easily checked) for two models which have not been considered (or not fully investigated) in the earlier literature. In Section 3.3, we will illustrate on a counterexample the fact that the geomtric drift condition, though not necessary, cannot be innocuously dispensed with.

3.1 AR(p) with regularly varying innovations

Convergence of the tail empirical processes of exceedances for infinite order moving averages has been obtained in the case of finite variance innovation; for infinite variance innovations it was proved only in the case of an AR(1) process in [Dre03]. We next show that Assumption 1 holds for general causal invertible AR(p) models.

Corollary 3.1. *Assume that $\{X_j, j \in \mathbb{Z}\}$ is an AR(p) model*

$$X_j = \phi_1 X_{j-1} + \cdots + \phi_p X_{j-p} + \varepsilon_j, \quad j \geq 1,$$

that satisfies the following conditions:

- *the innovations $\{\varepsilon_j, j \in \mathbb{Z}\}$ are i.i.d. and regularly varying with index α ;*

- the innovations have a density f_ε not vanishing in a neighbourhood of zero;
- the spectral radius of the matrix

$$\Sigma = \begin{pmatrix} \phi_1 & \phi_2 & \cdots & \phi_p \\ 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 0 & \cdots & 0 \end{pmatrix},$$

is smaller than 1.

- if $\alpha \leq 2$, then $\sum_{i=1}^p |\phi_i|^q < 1$ for $q = \min\{1, \alpha\}$.

Then Assumption 1 holds.

Proof. The AR(p) process can be embedded into an \mathbb{R}^p -valued vector-autoregressive Markov chain

$$\mathbb{Y}_j = \Sigma \mathbb{Y}_{j-1} + Z_j \tag{3.1}$$

with

$$\mathbb{Y}_j = (X_j, \dots, X_{j-p+1})^T, \quad Z_j = (\varepsilon_j, 0, \dots, 0)^T.$$

Since the spectral radius of Σ is smaller than 1, the stationary solution to (3.1) exists and is given by $\mathbb{Y}_j = \sum_{k=0}^{\infty} \Sigma^k \mathbb{Y}_{j-k}$. Since the innovation $\{\varepsilon_j\}$ is regularly varying, the chain $\{\mathbb{Y}_j\}$ is also regularly varying with index α . The AR(p) process is recovered by taking $X_j = g(\mathbb{Y}_j)$ with $g(\mathbf{y}) = g(y_1, \dots, y_p) = y_1$ and $\{X_j, j \in \mathbb{Z}\}$ is also regularly varying. Hence, the first two items of Assumption 1 are fulfilled. Due to the assumption on the innovations density, the chain is an irreducible and aperiodic T -chain (see e.g. [FT85]). Thus, by [MT09, Theorem 6.0.1], all compact sets are small sets.

We now check the drift condition (2.1). Let λ be the spectral radius of the matrix Σ . Fix ϵ such that $\gamma = \lambda + \epsilon < 1$. Then there exists a norm $\|\cdot\|_\Sigma$ on \mathbb{R}^p such that the matrix norm of Σ with respect to this norm is at most γ , that is

$$\sup_{\substack{\mathbf{x} \in \mathbb{R}^p \\ \|\mathbf{x}\|_\Sigma = 1}} \|\Sigma \mathbf{x}\|_\Sigma \leq \gamma.$$

See [DMS14, Proposition 4.24 and Example 6.35] for more details. Choose such a norm and for $q < \alpha$, set $V_q(\mathbf{y}) = 1 + \|\mathbf{y}\|_\Sigma^q$ and $v_q = \|(1, 0, \dots, 0)\|_\Sigma^q$. If $q \leq 1$, then we will use the inequality $(x + y)^q \leq x^q + y^q$. If $q > 1$, then for every $\eta \in (0, 1)$, there exists a constant C_q such that for all $x, y \geq 0$,

$$(x + y)^q \leq (1 + \eta)x^q + C_q y^q. \tag{3.2}$$

(Take for instance $C_q = \{1 - (1 + \eta)^{-1/(1-q)}\}^{1-q}$.) Set $c_q = C_q = 1$ if $q \leq 1$ and $c_q = (1 + \eta)$ and C_q as above if $q > 1$. This yields

$$\begin{aligned} \mathbb{E}[V_q(\mathbb{Y}_1) \mid \mathbb{Y}_0 = \mathbf{y}] &\leq 1 + c_q \|\Sigma \mathbf{y}\|_\Sigma^q + C_q v_q \mathbb{E}[|\varepsilon_0|^q] \\ &\leq 1 + c_q \gamma^q \|\mathbf{y}\|_\Sigma^q + C_q v_q \mathbb{E}[|\varepsilon_0|^q] = \lambda_q V_q(x) + b_q \end{aligned}$$

where $\lambda_q = c_q \gamma^q$ is smaller than 1 if $q \leq 1$ or can be made smaller than 1 by choosing an appropriately small η if $q > 1$ and $b_q = 1 - \lambda_q + C_q v_q \mathbb{E}[|\varepsilon_1|^q]$. Since all compact sets are small, this yields (2.1). Furthermore, $|g(\mathbf{y})|^q \leq c V_q(\mathbf{y})$ and by regular variation,

$$\lim_{n \rightarrow \infty} \frac{1}{u_n \bar{F}(u_n)} \mathbb{E} [V_q(\mathbb{Y}_0) 1_{\{su_n < g(\mathbb{Y}_0)\}}] = \int_{x_p > s} \int_{\mathbb{R}^{p-1}} \{1 + \|\mathbf{x}\|_\Sigma^q\} \nu_{1,p}(d\mathbf{x}) .$$

Hence, Assumption 1 holds for all $q < \alpha$. Conditions (2.3) and (2.4) hold, see Comment (C2). \square

3.2 Threshold ARCH

Corollary 3.2. *Let $\xi \in \mathbb{R}$. Assume that $\{X_j\}$ is T -ARCH model*

$$X_j = (b_{10} + b_{11} X_{j-1}^2)^{1/2} Z_j 1_{\{X_{j-1} < \xi\}} + (b_{20} + b_{21} X_{j-1}^2)^{1/2} Z_j 1_{\{X_{j-1} \geq \xi\}} , \quad (3.3)$$

that satisfies the following conditions:

- $b_{ij} > 0$;
- the innovations $\{Z_j, j \in \mathbb{Z}\}$ are i.i.d. such that $\mathbb{E}[|Z_j^\beta|] < \infty$ for all $\beta > 0$;
- the innovations have a density f_Z not vanishing in a neighbourhood of zero and bounded;
- the Lyapunov exponent

$$\gamma = p \log b_{11}^{1/2} + (1 - p) \log b_{21}^{1/2} + \mathbb{E}[\log(|Z_1|)] ,$$

where $p = \mathbb{P}(Z_1 < 0)$, is strictly negative;

- $(b_{11} \vee b_{21})^{q/2} \mathbb{E}[|Z_0|^q] < 1$.

Then Assumption 1 holds.

Proof. Under the stated conditions, the Markov chain $\{X_j\}$ is an irreducible and aperiodic T -chain; see [Cli07]. Since the Lyapunov exponent is negative, [Cli07, Theorem 2.2] implies that the stationary distribution exists and the chain is geometrically ergodic. The stationary distribution is regularly varying and the index of regular variation of X_1 is obtained by solving

$$b_{11}^{\alpha/2} \mathbb{E}[|Z_1|^\alpha 1_{Z_1 < 0}] + b_{21}^{\alpha/2} \mathbb{E}[|Z_1|^\alpha 1_{Z_1 \geq 0}] = 1 ;$$

see again [Chi07]. To check the drift condition, let $q < \alpha$. Set $V(x) = 1 + |x|^q$. Using (3.2) we have

$$\begin{aligned} \mathbb{E}[V(X_0) \mid X_0 = x] &= 1 + \{(b_{10} + b_{11}x^2)^{q/2}1_{\{x < \xi\}} + (b_{20} + b_{21}x^2)^{q/2}1_{\{x \geq \xi\}}\} \mathbb{E}[|Z_0|^q] \\ &\leq 1 + (1 + \eta)\{b_{11}^{q/2}1_{\{x < \xi\}} + b_{21}^{q/2}1_{\{x \geq \xi\}}\} \mathbb{E}[|Z_0|^q] |x|^q + B_q, \end{aligned}$$

with $B_q = C_q\{b_{10}^{q/2}1_{\{x < \xi\}} + b_{20}^{q/2}1_{\{x \geq \xi\}}\} \mathbb{E}[|Z_0|^q]$. Finally, since we have here $V(x) = 1 + |x|^q$, Condition (2.4) holds by regular variation. \square

3.3 A Counterexample

The geometric drift condition is not a necessary condition for the conclusions of Theorem 2.1 to hold, but when it does not hold, it is easy to build counterexamples of non geometrically ergodic Markov chains which exhibit a highly non standard behaviour of their tail empirical process.

Let $\{Z_j, j \in \mathbb{Z}\}$ be a sequence of i.i.d. positive integer valued random variables with regularly varying right tail with index $\beta > 1$. Define the Markov chain $\{X_j, j \geq 0\}$ by the following recursion:

$$X_j = \begin{cases} X_{j-1} - 1 & \text{if } X_{j-1} > 1, \\ Z_j & \text{if } X_{j-1} = 1. \end{cases}$$

Since $\beta > 1$, the chain admits a stationary distribution π on \mathbb{N} given by

$$\pi(n) = \frac{\mathbb{P}(Z_0 \geq n)}{\mathbb{E}[Z_0]}, \quad n \geq 1.$$

To avoid confusion, we will denote the distributions functions of Z_0 and X_0 (when the initial distribution is π) by F_Z and F_X , respectively. The tail \bar{F}_X of the stationary distribution is then regularly varying with index $\alpha = \beta - 1$, since it is given by

$$\bar{F}_X(x) = \frac{\mathbb{E}[(Z_0 - [x])_+]}{\mathbb{E}[Z_0]} \sim \frac{x\bar{F}_Z(x)}{\beta\mathbb{E}[Z_0]}. \quad (3.4)$$

Assuming for simplicity that $\mathbb{P}(Z_0 = n) > 0$ for all $n \geq 1$, this chain is irreducible and aperiodic and the state 1 is a recurrent atom. The distribution of the return time τ_1 to the atom 1, when the chains starts from 1 is the distribution of Z_0 . Hence the chain is not geometrically ergodic since under the assumption on Z_0 , $\mathbb{E}_1[\kappa^{\tau_1}] = \mathbb{E}[\kappa^{Z_0}] = \infty$ for all $\kappa > 1$. Moreover, the extremal index of the chain is 0, by an application of [Roo88, Theorem 3.2 and Eq. 4.2].

Let $\{u_n\}$ be a scaling sequence and define the usual univariate tail empirical distribution function by

$$\tilde{T}_n(s) = \frac{1}{n\bar{F}_X(u_n)} \sum_{j=1}^n 1_{\{X_j > u_n s\}}, \quad (3.5)$$

and $T_n(s) = \mathbb{E}[\tilde{T}_n(s)] = \bar{F}_X(u_n s) / \bar{F}_X(u_n)$. Let $\{a_n\}$ be a scaling sequence such that $\lim_{n \rightarrow \infty} n\mathbb{P}(Z_0 > a_n) = 1$.

Proposition 3.3. • If $\lim_{n \rightarrow \infty} n\bar{F}_Z(u_n) = 0$, then $\lim_{n \rightarrow \infty} \mathbb{P}(\tilde{T}_n(s) \neq 0) = 0$.

- If $\beta \in (1, 2)$ and $\lim_{n \rightarrow \infty} n\bar{F}_Z(u_n) = \infty$, then there exists a β -stable random variable Λ such that for every $s > 0$, $a_n^{-1}n\bar{F}_X(u_n)\{\tilde{T}_n(s) - T_n(s)\} \xrightarrow{d} \Lambda$.
- If $\beta > 2$, $\lim_{n \rightarrow \infty} n\bar{F}_Z(u_n) = \infty$ and $s_0 > 0$ then the process $s \rightarrow \sqrt{n\bar{F}_Z(u_n)}\{\tilde{T}_n(s) - T_n(s)\}$ converges weakly in $\ell^\infty([s_0, \infty))$ to a centered Gaussian process \mathbb{G} with covariance function

$$C(s, t) = \frac{(\beta + 1)t^{1-\beta}}{\beta(\beta - 1)} - \frac{st^{-\beta}}{\beta}, \quad s < t.$$

Remark 3.4. • In the standard situation (for example, under the geometric drift condition), a non degenerate limit is expected if $n\bar{F}_X(u_n) \rightarrow \infty$. Since $\bar{F}_X(u_n) \sim u_n\bar{F}_Z(u_n)$, it may happen simultaneously that $n\bar{F}_X(u_n) \rightarrow \infty$ and $n\bar{F}_Z(u_n) \rightarrow 0$. The appropriate threshold is determined by the distribution of Z_0 and not by the stationary distribution of the chain.

- In the case $1 < \beta < 2$, $a_n^{-1}n\bar{F}_X(u_n) = \bar{F}_X(u_n) / \bar{F}_X(a_n) \rightarrow \infty$, thus the tail empirical distribution is consistent, but since the limiting distribution of the TEP does not depend on s , it might be useless for inference.

4 Statistical applications

In statistical applications the presence of the sequence $\{u_n\}$ is not desirable. Let k be an intermediate sequence and let the sequence $\{u_n\}$ be defined by $u_n = F^{\leftarrow}(1 - k/n)$ where F^{\leftarrow} is the left continuous generalized inverse of F . If F is continuous, then $k = n\bar{F}(u_n)$. Given a sample X_1, \dots, X_n , let $X_{n:1}, \dots, X_{n:n}$ be the increasing order statistics of the sample. In statistical applications, the sequence u_n is replaced by $X_{n:n-k}$, the $k+1$ largest observation in the sample.

In this section, we will give the asymptotic covariances in terms of the tail process $\{Y_j, j \in \mathbb{Z}\}$ or spectral tail process $\{\Theta_j, j \in \mathbb{Z}\}$, introduced by [BS09]. The regular variation of the time series $\{X_j, j \in \mathbb{Z}\}$ is equivalent to the existence of these processes defined as follows: for $j \leq k \in \mathbb{Z}$,

$$\mathbb{P}((Y_j, \dots, Y_k) \in \cdot) = \lim_{x \rightarrow \infty} \mathbb{P}(x^{-1}(X_{-j}, \dots, X_k) \in \cdot \mid |X_0| > x),$$

$$\Theta_j = Y_j / |Y_0|.$$

Then $|Y_0|$ is a standard Pareto random variable with tail index α , independent of the sequence $\{\Theta_j, j \in \mathbb{Z}\}$. The tail process provides in some cases convenient expressions of the asymptotic limits of the estimates, though in any case, these variances must be estimated.

Bias issues. When studying estimators, a bias term appears. In extreme value statistics, dealing with the bias means being able to make a suitable choice of the number k of order statistics that are used. Such a choice is always theoretically possible (see conditions (4.4), (4.14) and (4.8) below). The practical issue of a data-driven choice of k is not addressed here.

4.1 Convergence of order statistics

Consider the univariate TED and the univariate TEP

$$\tilde{T}_n = \tilde{M}_n(s, \infty, \dots, \infty) = \frac{1}{n\bar{F}(u_n)} \sum_{j=1}^n 1_{\{X_j > u_n s\}}, \quad (4.1)$$

$$T_n(s) = \mathbb{E}[\tilde{T}_n(s)] = \frac{\bar{F}(u_n s)}{\bar{F}(u_n)}, \quad (4.2)$$

$$\mathbb{G}_n(s) = \sqrt{n\bar{F}(u_n)} \{\tilde{T}_n(s) - T_n(s)\} = \mathbb{M}_n(s, \infty, \dots, \infty). \quad (4.3)$$

By Theorem 2.1, \mathbb{G}_n converges weakly to the Gaussian process \mathbb{G} defined by $\mathbb{G}(s) = \mathbb{M}(s, \infty, \dots, \infty)$. Note that $\bar{F}(u_n s)/\bar{F}(u_n)$ converges to $s^{-\alpha}$, uniformly on every interval $[s_0, \infty]$. Define

$$B_n(s_0) = \sup_{s \geq s_0} \left| \frac{\bar{F}(u_n s)}{\bar{F}(u_n)} - s^{-\alpha} \right|$$

and assume that

$$\lim_{n \rightarrow \infty} \sqrt{k} B_n(s_0) = 0. \quad (4.4)$$

Corollary 4.1. *Under the assumptions of Theorem 2.1 and if additionally (4.4) holds then*

$$\sqrt{k} \left\{ \frac{X_{n:n-k}}{u_n} - 1 \right\} \xrightarrow{d} \alpha^{-1} \mathbb{G}(1). \quad (4.5)$$

Moreover, the convergence holds jointly with that of \mathbb{M}_n^ψ to \mathbb{M}^ψ for any function ψ satisfying the assumption of Theorem 2.1.

The proof is a standard application of Theorem 2.1 and Vervaat's Lemma and is omitted (see e.g. [Roo09]; [KS11].) The autocovariance function C of the process \mathbb{G} is given by (2.5):

$$C(v, w) = c_0(v, w) + \sum_{j=1}^{\infty} \{c_j(v, w) + c_j(w, v)\},$$

where

$$c_j(v, w) = \lim_{n \rightarrow \infty} \frac{\mathbb{P}(X_0 > u_n v, X_j > u_n w)}{\bar{F}(u_n)} = \int_{\mathbb{R}^{j+1}} 1_{(v, \infty)}(x_0) 1_{(w, \infty)}(x_j) \boldsymbol{\nu}_{0,j}(\mathrm{d}\mathbf{x}).$$

In the language of [DM09], the sequence of coefficients $\{c_j\}$ is the extremogram of $\{X_j\}$ related to the sets $(v, \infty), (w, \infty)$. Using the tail process $\{Y_j\}$ or the spectral tail process $\{\Theta_j\}$, the coefficient c_j can be represented as

$$c_j(v, w) = v^{-\alpha} \mathbb{P}(Y_j > w/v \mid Y_0 > 1) = v^{-\alpha} \mathbb{E} [(\Theta_j v/w)_+^\alpha \wedge 1 \mid \Theta_0 = 1] .$$

This yields

$$\begin{aligned} \text{var}(\mathbb{G}(1)) &= C(1, 1) = 1 + 2 \sum_{j=1}^{\infty} \mathbb{P}(Y_j > 1 \mid Y_0 > 1) \\ &= 1 + 2 \sum_{j=1}^{\infty} \mathbb{E}[(\Theta_j)_+^\alpha \wedge 1 \mid \Theta_0 = 1] = \sum_{j \in \mathbb{Z}} \mathbb{E}[(\Theta_j)_+^\alpha \wedge 1 \mid \Theta_0 = 1] . \end{aligned}$$

In particular, if the sequence $\{X_j\}$ is extremally independent, which means that all the exponent measures $\nu_{0,j}$ are concentrated on the axes, then $\Theta_j = 0$ for $j \neq 0$ and the limiting distribution in (4.5) is normal with mean zero and variance α^{-2} .

Counterexample, continued. We now investigate the order statistics for the counterexample of Section 3.3. Consider the case $\beta > 2$ and $\lim_{n \rightarrow \infty} n \bar{F}_Z(u_n) = \infty$. An application of Vervaat's Lemma (see the argument in [Roo09] or [KS11]) yields

$$\sqrt{n \bar{F}_Z(u_n)} \left(T_n \circ \tilde{T}_n^{\leftarrow}(1) - 1 \right) \xrightarrow{d} -(\tilde{G} \circ T^{\leftarrow})(1) ,$$

where \tilde{T}_n, T_n are defined in (4.1), and $T(s) = s^{-\alpha}$, with $\alpha = \beta - 1$, the tail index of the stationary distribution. A Taylor expansion yields

$$T_n \circ \tilde{T}_n^{\leftarrow}(1) - 1 \approx T_n'(T_n^{\leftarrow}(1)) \left\{ \tilde{T}_n^{\leftarrow}(1) - T_n^{\leftarrow}(1) \right\} .$$

Set again $u_n = F_X^{\leftarrow}(1 - k/n)$. Then $\tilde{T}_n^{\leftarrow}(1) = X_{n:n-k}/u_n$, as well as $T_n^{\leftarrow}(1) = 1$. Under suitable regularity conditions which ensure that T_n' converges uniformly in the neighbourhood of 1, we have

$$\sqrt{n \bar{F}_Z(u_n)} \left\{ \frac{X_{n:n-k}}{u_n} - 1 \right\} \xrightarrow{d} \alpha^{-1} \tilde{\mathbb{G}}(1)$$

or equivalently,

$$\sqrt{k/u_n} \left\{ \frac{X_{n:n-k}}{u_n} - 1 \right\} \xrightarrow{d} \alpha^{-1} \tilde{\mathbb{G}}(1) .$$

The limiting distribution is normal with variance $2\alpha^{-4}(\alpha - 1)^{-1}$. If $n \bar{F}_Z(u_n) \rightarrow 1$, then $k \sim u_n \rightarrow \infty$ but in that case, for $j = o(u_n)$, $u_n^{-1} X_{n:n-j}$ converges weakly to one single Fréchet distribution with tail index β .

4.2 Hill estimator

The classical Hill estimator of $\gamma = 1/\alpha$ is defined as

$$\hat{\gamma} = \frac{1}{k} \sum_{j=1}^n \log_+ \left(\frac{X_{n:n-j+1}}{X_{n:n-k}} \right).$$

Under the conditions ensuring that $X_{n:n-k}/u_n \rightarrow_P 1$, the order statistics appearing in the definition of the estimator are all positive with probability tending to 1, thus the estimator is well defined.

Corollary 4.2. *Under the assumptions of Theorem 2.1 and if moreover (4.4) holds, $\sqrt{k} \{\hat{\gamma} - \gamma\}$ converges weakly to a centered Gaussian distribution with variance*

$$\alpha^{-2} \left\{ 1 + 2 \sum_{j=1}^{\infty} \mathbb{P}(Y_j > 1 \mid Y_0 > 1) \right\} = \alpha^{-2} \left\{ 1 + 2 \sum_{j=1}^{\infty} \mathbb{E}[(\Theta_j)_+^\alpha \wedge 1 \mid \Theta_0 = 1] \right\}. \quad (4.6)$$

This result provides the asymptotic normality of the Hill estimator for all irreducible Markov chains that satisfy the geometric drift condition. The proof is again a standard application of Theorem 2.1 and is omitted. The expression for the limiting variance is justified in Section 5.5. In the case of an extremally independent time series where $\Theta_j = 0$ for $j \neq 0$, the variance is simply α^{-2} as in the case of an i.i.d. sequence.

4.3 Estimation of the extremal index

Set $A_h = \{x_0 \vee \dots \vee x_h > 1\}$ and consider

$$\nu_{0,h}(A_h) = \lim_{x \rightarrow \infty} \frac{\mathbb{P}(X_0 \vee \dots \vee X_h > x)}{\bar{F}(x)}.$$

By conditioning on the last index $j \leq h$ such that $X_j > x$, we can express this limit in terms of the spectral tail process:

$$\nu_{0,h}(A_h) = 1 + \sum_{j=1}^{h-1} \mathbb{P}(\vee_{i=1}^j Y_i \leq 1 \mid Y_0 > 1).$$

It has been proved in [BS09] that the anticlustering condition (1.3) implies that the right-tail extremal index θ_+ of the sequence $\{X_i\}$ is positive, and moreover, $\theta_+ = \mathbb{P}(\max_{j \geq 1} Y_j \leq 1 \mid Y_0 > 1)$. Since the drift condition (2.1) implies the anticlustering condition, we obtain

$$\theta_+ = \lim_{h \rightarrow \infty} \frac{\nu_{0,h}(A_h)}{h} \in [0, \infty).$$

Set $\theta_+(h) = h^{-1} \nu_{0,h}(A_h)$. Then a natural estimator of $\theta_+(h)$ can be defined by

$$\hat{\theta}_n(h) = \frac{1}{hk} \sum_{j=1}^{n-h} 1_{\{X_j \vee \dots \vee X_{j+h} > X_{n:n-k}\}} = h^{-1} \tilde{M}_n(u_n^{-1} X_{n:n-k} \mathbf{1}).$$

This yields, applying the homogeneity of the measure $\nu_{0,h}$,

$$\sqrt{k}\{\hat{\theta}_n(h) - \theta_+(h)\} = h^{-1}\mathbb{M}_{n,h}(u_n^{-1}X_{n:n-k}\mathbf{1}) \quad (4.7a)$$

$$+ \frac{n-h}{nh}\sqrt{k}\{M_n(u_n^{-1}X_{n:n-k}\mathbf{1}) - \nu_{0,h}(X_{n:n-k}A_h/u_n)\} \quad (4.7b)$$

$$+ \frac{n-h}{nh}\sqrt{k}\{(X_{n:n-k}/u_n)^{-\alpha} - 1\}\nu_{0,h}(A_h) + \frac{h}{n}\theta_+(h). \quad (4.7c)$$

In order to prove the convergence, we need an additional condition to deal with the bias term in (4.7b). For $s > 0$, define

$$B_n(h, s_0) = \sup_{s \geq s_0} \left| \frac{\mathbb{P}(X_0 \vee \dots \vee X_h > u_n s)}{\bar{F}(u_n)} - s^{-\alpha} \nu_{0,h}(A_h) \right|.$$

Since the function $s \rightarrow \mathbb{P}(X_0 \vee \dots \vee X_h > u_n s) / \bar{F}(u_n)$ is monotone and its limit is continuous, the convergence is also uniform on $[s_0, \infty]$ for every $s_0 > 0$. Therefore, we assume that the intermediate sequence k is chosen in such a way that

$$\lim_{n \rightarrow \infty} \sqrt{k} B_n(h, s_0) = 0, \quad (4.8)$$

for a fixed $s_0 \in (0, 1)$.

Corollary 4.3. *Under the assumptions of Theorem 2.1 and if moreover (4.4) and (4.8) hold, $\sqrt{k}(\hat{\theta}_n(h) - \theta_+(h))$ converges weakly to $h^{-1}\mathbb{M}_h(\mathbf{1}) - \theta_+(h)\mathbb{G}(1)$.*

Proof. Since the present assumptions subsume those of Corollary 4.1, $X_{n:n-k}/u_n \xrightarrow{p} 1$ and $\sqrt{k}(X_{n:n-k}/u_n)^{-\alpha} - 1 \rightarrow -\mathbb{G}(1)$. Moreover, Theorem 2.1 implies that $\mathbb{M}_n(u_n^{-1}X_{n:n-k}\mathbf{1}) \xrightarrow{d} \mathbb{M}(\mathbf{1})$, and the convergence holds jointly. Condition (4.8) implies that the bias term in (4.7b) is asymptotically vanishing. The result follows. \square

Note that we are not estimating the extremal index, but only the quantity $\theta_+(h)$ which converges to it as $h \rightarrow \infty$. Therefore, it is not devoid of interest; and for practical purposes, h is necessarily finite. Moreover, we can improve on this approximation. Since $\{\mathbb{P}(Y_1 \vee \dots \vee Y_h \leq 1 \mid Y_0 > 1), h \geq 1\}$ is a decreasing sequence with limit θ_+ , for each $h \geq 1$, $\mathbb{P}(Y_1 \vee \dots \vee Y_h \leq 1 \mid Y_0 > 1)$ is closer to its limit θ_+ (as $h \rightarrow \infty$) than its Cesaro mean. Therefore, it is probably a better idea to estimate this quantity rather than $\theta_+(h)$. This is indeed the case for certain models, as noted by [Hsi93, Example C]. Since by definition

$$\tilde{\theta}_+(h) = h\theta_+(h) - (h-1)\theta_+(h-1) = \lim_{x \rightarrow \infty} \frac{\mathbb{P}(X_0 \leq x, \dots, X_{h-1} \leq x, X_h > x)}{\bar{F}(x)},$$

an estimator of $\tilde{\theta}_+(h)$ is defined by

$$\tilde{\theta}_n(h) = \frac{1}{k} \sum_{i=1}^n 1_{\{X_i \vee \dots \vee X_{i+h-1} \leq X_{n:n-k}, X_{i+h} > X_{n:n-k}\}}.$$

This estimator was considered by [Hsi93] under ad-hoc summability assumption for the covariances of the indicators in the definition of the estimator which is implied by the anticlustering condition, hence by the drift condition. It can be similarly proved that under suitable bias conditions on k , $\sqrt{k}(\tilde{\theta}_n(h) - \theta_+(h))$ converges weakly to a centered Gaussian distribution which can be expressed as $\mathbb{M}_h(1, \dots, 1, \infty) - \{h\theta_+(h) - (h-1)\theta_+(h-1)\}\mathbb{G}(1)$.

4.4 Estimation of the cluster index

In a very similar way, with maxima replaced by sums, we can obtain the limiting distribution for an estimator of the cluster index. For $A_h = \{x_0 + \dots + x_h > 1\}$ we consider

$$b_+(h) = \frac{1}{h} \lim_{n \rightarrow \infty} \frac{\mathbb{P}(X_0 + \dots + X_h > x)}{\bar{F}(x)} = \frac{1}{h} \nu_{0,h}(A_h) .$$

It is shown in [MW14] that the drift condition (2.1) implies that

$$b_+ = \lim_{h \rightarrow \infty} b_+(h) \in [0, \infty) .$$

Define

$$\hat{b}_+(h) = \frac{1}{kh} \sum_{j=1}^{n-h} 1_{\{X_j + \dots + X_{j+h} > X_{n:n-k}\}} .$$

In order to prove the convergence, we need an additional condition. For $s > 0$, define

$$D_n(h, s_0) = \sup_{s \geq s_0} \left| \frac{\mathbb{P}(X_0 + \dots + X_h > u_n s)}{\bar{F}(u_n)} - s^{-\alpha} \nu_{0,h}(A_h) \right| .$$

We assume that the intermediate sequence k is chosen in such a way that

$$\lim_{n \rightarrow \infty} \sqrt{k} D_n(h, s_0) = 0 , \tag{4.9}$$

for a fixed $s_0 \in (0, 1)$.

Corollary 4.4. *Under the assumptions of Theorem 2.1 and if moreover (4.4) and (4.9) hold, $\sqrt{k}(\hat{b}_+(h) - b_+(h))$ converges weakly to a zero mean Gaussian random variable.*

The proof is similar to the proof of Corollary 4.3 and is omitted.

4.5 Conditional tail expectation

If the tail index of the time series $\{X_j\}$ is $\alpha > 1$, then the following limit exists:

$$\lim_{u \rightarrow \infty} \frac{1}{u} \mathbb{E}[X_h \mid X_0 > ux] = \int_{x_0=1}^{\infty} \int_{\mathbb{R}^h} x_h \nu_{0,h}(d\mathbf{x}) = \text{CTE}(h) . \tag{4.10}$$

The quantity above is the conditional tail expectation and when $h = 0$ it is being used (under the name expected shortfall) as a risk measure, a coherent alternative to the popular value-at-risk. In the case of bivariate i.i.d. vectors with the same distribution as (X, Y) , statistical procedures for estimating $\mathbb{E}[Y \mid X > Q_X(1 - p)]$ when $p \rightarrow 0$, where $Q_X(p) = F_X^{\leftarrow}(1 - p)$, were developed in [CEdHZ15]. The limit (4.10) yields the approximation

$$\mathbb{E}[X_h \mid X_0 > F^{\leftarrow}(1 - p)] \sim F^{\leftarrow}(1 - p)\text{CTE}(h) .$$

If $p \approx 1/n$, this becomes an extrapolation problem, related to the estimation of extreme quantiles. if $\hat{\alpha}_n$ is an estimator of the right tail index α of the marginal distribution, then an estimator of the quantile of order p is given by

$$X_{n:n-k} \left(\frac{k}{np} \right)^{1/\hat{\alpha}_n} .$$

The rationale for this estimator is the approximation $F^{\leftarrow}(1 - p) \sim F^{\leftarrow}(1 - k/n)(k/np)^{1/\alpha}$ and the convergence $X_{n:n-k}/F^{\leftarrow}(1 - k/n) \xrightarrow{p} 1$ established above. A simple estimator of $\text{CTE}(h)$ is given by

$$\hat{C}_n(h) = \frac{1}{kX_{n:n-k}} \sum_{j=1}^n X_{j+h} 1_{\{X_j > X_{n:n-k}\}}$$

This yields an estimator of $E_h(p) = \mathbb{E}[X_h \mid X_0 > F^{\leftarrow}(1 - p)]$:

$$\hat{E}_n = X_{n:n-k} \left(\frac{k}{np} \right)^{1/\hat{\alpha}_n} \hat{C}_n(h) = \left(\frac{k}{np} \right)^{1/\hat{\alpha}_n} \frac{1}{k} \sum_{j=1}^n X_{j+h} 1_{\{X_j > X_{n:n-k}\}} .$$

We will focus here on the asymptotic distribution of $\hat{C}_n(h) - \text{CTE}(h)$ suitably normalized. The application to the limiting distribution of $\sqrt{k} \left\{ \hat{E}_n / E_h(p) - 1 \right\}$ when $p = p_n \approx n^{-1}$ is straightforward, given additional ad-hoc bias assumptions. Define

$$\hat{T}_{n,h}(s) = \frac{1}{ku_n} \sum_{j=1}^n X_{j+h} 1_{\{X_j > u_n s\}} , \quad (4.11)$$

$$T_{n,h}(s) = \mathbb{E}[\hat{T}_{n,h}(s)] = \frac{1}{u_n \bar{F}(u_n)} \mathbb{E} \left[X_h 1_{\{X_0 > u_n s\}} \right] . \quad (4.12)$$

Regular variation and $\alpha > 1$ imply

$$T_h(s) = \lim_{n \rightarrow \infty} T_{n,h}(s) = s^{-\alpha} \int_{x_0=1}^{\infty} \int_{\mathbb{R}^h} x_h \nu_{0,h}(\mathrm{d}\mathbf{x}_{0,h}) = s^{-\alpha} \text{CTE}(h) . \quad (4.13)$$

We shall assume that

$$\lim_{n \rightarrow \infty} \sqrt{k} \sup_{s \geq s_0} |T_{n,h}(s) - T_h(s)| = 0 . \quad (4.14)$$

Then

$$\sqrt{k} \left\{ \hat{C}_n(h) - \text{CTE}(h) \right\} = \frac{u_n}{X_{n:n-k}} \sqrt{k} \left\{ \hat{T}_{n,h}(X_{n:n-k}/u_n) - T_{n,h}(X_{n:n-k}/u_n) \right\} \quad (4.15a)$$

$$+ \frac{u_n}{X_{n:n-k}} \sqrt{k} \left\{ T_{n,h}(X_{n:n-k}/u_n) - (X_{n:n-k}/u_n)^{-\alpha} \text{CTE}(h) \right\} \quad (4.15b)$$

$$+ \sqrt{k} \left\{ (X_{n:n-k}/u_n)^{-\alpha-1} - 1 \right\} \text{CTE}(h) . \quad (4.15c)$$

If (4.4) holds then we can apply Corollary 4.1. In particular, $X_{n:n-k}/u_n \xrightarrow{P} 1$. As usual, the term in (4.15b) is a bias term which vanishes thanks to (4.14). If $\alpha > 2$, applying Theorem 2.1, the term in (4.15a) has a Gaussian limit which can be expressed as $\mathbb{M}_h^\psi(1)$ with $\psi(x) = x$. The last terms converges to $-\alpha^{-1}(\alpha + 1)\mathbb{G}(1)$.

Corollary 4.5. *Under the assumptions of Theorem 2.1 and if moreover (4.4) and (4.14) hold, $\sqrt{k} \left\{ \hat{C}_n - \text{CTE}(h) \right\}$ converges weakly to $\mathbb{M}_h^\psi(1) - \alpha^{-1}(\alpha + 1)\text{CTE}(h)\mathbb{G}(1)$.*

5 Proofs

5.1 Proof of Condition $\mathcal{S}(u_n, r_n, \psi)$

As mentioned in Section 2.1, the main ingredient of the proof is Condition $\mathcal{S}(u_n, r_n, \psi)$. In order to prove that it is implied by the drift condition (2.1), we recall some consequences of the geometric drift condition. Under condition (2.2), the chain $\{\mathbb{Y}_j\}$ can be embedded into an extended Markov chain $\{(\mathbb{Y}_j, B_j)\}$ such that the latter chain possesses an atom A , that is $\bar{P}(s, \cdot) = \bar{P}(t, \cdot)$ for every $s, t \in A$, where \bar{P} is the transition kernel of the extended chain. This existence is due to the Nummelin splitting technique (see [MT09, Chapter 5]). Denote by \mathbb{E}_A the expectation conditionally to $(\mathbb{Y}_0, B_0) \in A$ and let τ_A be the first return time to A of the chain $\{(\mathbb{Y}_j, B_j), j \geq 0\}$. Note that τ_A is a stopping time with respect to the extended chain, but not with respect to the chain $\{\mathbb{Y}_j\}$. We assume that the extended chain is defined on the original probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and that the extended chain is stationary under \mathbb{P} . Then, there exist $\kappa > 1$ and a constant \aleph such that for all $y \in E$,

$$\mathbb{E} \left[\sum_{j=1}^{\tau_A} \kappa^j |X_j|^q \mid \mathbb{Y}_0 = y \right] \leq c \mathbb{E} \left[\sum_{j=1}^{\tau_A} \kappa^j V(\mathbb{Y}_j) \mid \mathbb{Y}_0 = y \right] \leq \aleph V(y) . \quad (5.1)$$

By Jensen's inequality, this implies that for all $q_1 \leq q$, there exists $\kappa_1 \in (1, \kappa)$ such that

$$\mathbb{E} \left[\sum_{j=1}^{\tau_A} \kappa_1^j |X_j|^{q_1} \mid \mathbb{Y}_0 = y \right] \leq \aleph V^{q_1/q}(y) . \quad (5.2)$$

Moreover, Kac's formula [MT09, Theorem 15.0.1] gives an expression of the stationary distribution in terms of the return time to A . For every bounded measurable function f ,

it holds that

$$\mathbb{E}[f(\mathbb{Y}_0)] = \frac{1}{\mathbb{E}_A[\tau_A]} \mathbb{E}_A \left[\sum_{i=0}^{\tau_A-1} f(\mathbb{Y}_i) \right]. \quad (5.3)$$

Since $V \geq 1$, the inequality (5.1) integrated with respect to the stationary distribution implies that $\mathbb{E}[\kappa^{\tau_A}] < \infty$. For $q \geq 0$ and $0 < s < t \leq \infty$ define

$$Q_n(s) = \frac{1}{u_n^q \bar{F}(u_n)} \mathbb{E} [V(\mathbb{Y}_0) 1_{\{su_n < X_0\}}] ,$$

and $Q_n(s, t) = Q_n(s) - Q_n(t)$.

Lemma 5.1. *Let Assumption 1 holds. For every $s_0 > 0$, there exists a constant \aleph such that for $q_1 + q_2 \leq q$, $L > h$ and $t > s \geq s_0$,*

$$\frac{1}{\bar{F}(u_n)} \mathbb{E} \left[\sum_{j=L}^{\tau_A} 1_{\{su_n < X_0 \leq tu_n\}} |X_h/u_n|^{q_1} |X_j/u_n|^{q_2} \right] \leq \aleph \kappa^{-L} \{Q_n^{q_2/q}(s, t) + Q_n(s, t)\}. \quad (5.4)$$

Proof. Let the left hand side of (5.4) be denoted by $S_n(s, t)$. Splitting the expectation between $\{|X_h| \leq u_n\}$ and $\{|X_h| > u_n\}$ yields

$$\begin{aligned} S_n(s, t) &\leq \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[\sum_{j=L}^{\tau_A} 1_{\{su_n < X_0 \leq tu_n\}} |X_j/u_n|^{q_2} \right] \\ &\quad + \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[\sum_{j=L}^{\tau_A} 1_{\{su_n < X_0 \leq tu_n\}} 1_{\{u_n < |X_h|\}} |X_h/u_n|^{q_1} |X_j/u_n|^{q_2} \right] \\ &= S_{n,1}(s, t) + S_{n,2}(s, t). \end{aligned}$$

Applying the bound (5.2) and Jensen's inequality (since $q_2 \leq q$), we obtain

$$S_{n,1}(s, t) \leq \frac{1}{u_n^{q_2} \bar{F}(u_n)} \mathbb{E} [V^{q_2/q}(\mathbb{Y}_0) 1_{\{su_n < X_0 \leq tu_n\}}] \leq \aleph \kappa^{-L} Q_n^{q_2/q}(s, t).$$

By the Markov property and the bound (5.2) (and noting that $q_2 \leq q - q_1$), we obtain

$$\begin{aligned}
S_{n,2}(s, t) &= \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[\sum_{j=L}^{\tau_A} 1_{\{su_n < X_0 \leq tu_n\}} 1_{\{u_n < |X_h|\}} |X_h/u_n|^{q_1} |X_j/u_n|^{q_2} \right] \\
&= \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[1_{\{su_n < X_0 \leq tu_n\}} 1_{\{u_n < |X_h|\}} |X_h/u_n|^{q_1} \mathbb{E} \left[\sum_{j=L}^{\tau_A} |X_j/u_n|^{q_2} \mid \mathbb{Y}_h \right] \right] \\
&= \frac{\kappa_2^{-L}}{\bar{F}(u_n)} \mathbb{E} \left[1_{\{su_n < X_0 \leq tu_n\}} 1_{\{u_n < |X_h|\}} |X_h/u_n|^{q_1} \mathbb{E} \left[\sum_{j=L}^{\tau_A} \kappa_2^j |X_j/u_n|^{q_2} \mid \mathbb{Y}_h \right] \right] \\
&\leq \aleph \frac{\kappa_2^{-L}}{u_n^{q_2} \bar{F}(u_n)} \mathbb{E} \left[1_{\{su_n < X_0 \leq tu_n\}} 1_{\{u_n < |X_h|\}} |X_h/u_n|^{q_1} V^{q_2/q}(\mathbb{Y}_h) \right] \\
&\leq \aleph \frac{\kappa_2^{-L}}{u_n^{q_2} \bar{F}(u_n)} \mathbb{E} \left[1_{\{su_n < X_0 \leq tu_n\}} 1_{\{u_n < |X_h|\}} |X_h/u_n|^{q_1 - q_2} V^{q_2/q}(\mathbb{Y}_h) \right] \\
&\leq \aleph \frac{\kappa_2^{-L}}{u_n^q \bar{F}(u_n)} \mathbb{E} \left[1_{\{su_n < X_0 \leq tu_n\}} V(\mathbb{Y}_h) \right] .
\end{aligned}$$

Iterating the drift condition (2.1) (and since $V \geq 1$), we obtain that

$$\mathbb{E}[V(\mathbb{Y}_h) \mid \mathbb{Y}_0 = y] \leq \gamma^h V(y) + \frac{b}{1 - \gamma} \leq \aleph V(y) .$$

This yields

$$S_{n,2}(s, t) \leq \aleph \frac{\kappa_2^{-L}}{u_n^q \bar{F}(u_n)} \mathbb{E} \left[1_{\{su_n < X_0 \leq tu_n\}} V(\mathbb{Y}_0) \right] \leq \aleph \kappa^{-L} Q_n(s, t) .$$

□

Lemma 5.2. *If Assumption 1 holds, $r_n \bar{F}(u_n) = o(1)$, and $q_1 + q_2 \leq q < \alpha$, then*

$$\frac{1}{\bar{F}(u_n)} \mathbb{E} \left[1_{\{\tau_A > h\}} \sum_{j=\tau_A+1}^{r_n} 1_{\{su_n < X_0 \leq tu_n\}} 1_{\{su_n < X_j \leq tu_n\}} |X_h/u_n|^{q_1} |X_j/u_n|^{q_2} \right] = o(1) . \quad (5.5)$$

Proof. Let the left handside of (5.5) be denoted R_n . Then, by the strong Markov property,

$$\begin{aligned}
R_n &\leq \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[1_{\{su_n < X_0\}} |X_h/u_n|^{q_1} \mathbb{E} \left[\sum_{j=\tau_A+1}^{r_n} 1_{\{su_n < X_j\}} |X_j/u_n|^{q_2} \mid \mathcal{F}_{\tau_A} \right] \right] \\
&\leq \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[1_{\{su_n < X_0\}} |X_h/u_n|^{q_1} 1_{\{\tau_A > h\}} \mathbb{E} \left[\sum_{j=\tau_A+1}^{r_n + \tau_A} 1_{\{su_n < X_j\}} |X_j/u_n|^{q_2} \mid \mathcal{F}_{\tau_A} \right] \right] \\
&\leq \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[1_{\{su_n < X_0\}} |X_h/u_n|^{q_1} \right] \mathbb{E}_A \left[\sum_{j=1}^{r_n} 1_{\{su_n < X_j\}} |X_j/u_n|^{q_2} \right] .
\end{aligned}$$

By classical regenerative arguments, Kac's formula (5.3) and regular variation, we obtain as $n \rightarrow \infty$,

$$\begin{aligned} \mathbb{E}_A \left[\sum_{j=1}^{r_n} 1_{\{su_n < X_j\}} |X_j/u_n|^{q_2} \right] &\sim \frac{r_n}{\mathbb{E}_A[\tau_A]} \mathbb{E}_A \left[\sum_{j=1}^{\tau_A} 1_{\{su_n < X_j\}} |X_j/u_n|^{q_2} \right] \\ &\leq r_n \bar{F}(u_n) \frac{\mathbb{E}[1_{\{X_0 > u_n s_0\}} |X_0/u_n|^{q_2}]}{\bar{F}(u_n)} = O(r_n \bar{F}(u_n)) . \end{aligned}$$

This yields, applying again the drift condition, Condition (2.4) and Jensen's inequality,

$$\begin{aligned} R_n &\leq O(r_n) \mathbb{E} [1_{\{X_0 > u_n s_0\}} |X_h/u_n|^{q_1}] \\ &= O(r_n \bar{F}(u_n)) \frac{\mathbb{E} [V^{q_1/q}(\mathbb{Y}_0) | X_0 > u_n s_0]}{u_n^{q_1}} \\ &= O(r_n \bar{F}(u_n)) \left(\frac{\mathbb{E} [V(\mathbb{Y}_0) | X_0 > u_n s_0]}{u_n^q} \right)^{q_1/q} = o(1) . \end{aligned}$$

□

Lemma 5.3. *Let Assumption 1 and (2.8) hold and $r_n \bar{F}(u_n) = o(1)$. Then Condition $\mathcal{S}(u_n, r_n, \psi)$ holds.*

Proof. By assumption (2.8), and for every $\mathbf{v}, \mathbf{w} \in (\mathbf{0}, \infty)$, there exists $\epsilon > 0$ such that for $j \geq h$,

$$\begin{aligned} 0 &\leq 1_{[\infty, \mathbf{v}]^c}(\mathbf{X}_{0,h}/u_n) 1_{[\infty, \mathbf{w}]^c}(\mathbf{X}_{j,j+h}/u_n) |\psi(\mathbf{X}_{0,h}/u_n)| |\psi(\mathbf{X}_{j,j+h}/u_n)| \\ &\leq \aleph u_n^{-q_{i_2} - q_{i_4}} \sum_{i_1, i_2, i_3, i_4=0}^h 1_{\{\epsilon u_n < X_{i_1}\}} |X_{i_2}|^{q_{i_2}} 1_{\{\epsilon u_n < X_{j+i_3}\}} |X_{j+i_4}|^{q_{i_4}} . \end{aligned} \quad (5.6)$$

For all i, i' , we can write

$$1_{\{\epsilon u_n < X_i\}} |X_{i'}/u_n|^q \leq 1_{\{\epsilon u_n < X_i\}} |X_i/u_n|^q + 1_{\{u_n \epsilon < X_{i'}\}} |X_{i'}/u_n|^q .$$

Thus, we can restrict the sum in (5.6) to the set of indices (i_1, i_2, i_3, i_4) such that $i_1 = i_2$ and $i_3 = i_4$ and since h is fixed, there is no loss of generality in restricting further to the cases $i_1 = i_2 = i_3 = i_4 = 0$. For an integer $L > 0$, splitting the sum at τ_A and applying Lemmas 5.1 and 5.2, we obtain, with $q_1 + q_2 \leq q$,

$$\begin{aligned} &\frac{1}{\bar{F}(u_n)} \mathbb{E} \left[\sum_{j=L}^{r_n} 1_{\{su_n < X_0\}} 1_{\{su_n < X_j\}} |X_0/u_n|^{q_1} |X_j/u_n|^{q_2} \right] \\ &\leq \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[\sum_{j=L}^{\tau_A} 1_{\{su_n < X_0\}} 1_{\{su_n < X_j\}} |X_0/u_n|^{q_1} |X_j/u_n|^{q_2} \right] \\ &+ \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[\sum_{j=\tau_A+1}^{r_n} 1_{\{su_n < X_0\}} 1_{\{su_n < X_j\}} |X_0/u_n|^{q_1} |X_j/u_n|^{q_2} \right] \\ &\leq \aleph \kappa^{-L} \{Q_n^{q_2/q}(s) + Q_n(s)\} + o(1) . \end{aligned}$$

This proves that $\mathcal{S}(u_n, r_n, \psi)$ holds since Q_n is asymptotically locally uniformly bounded by (2.4). \square

5.2 Adapting Drees and Rootzen 2010

The proof of Theorem 2.1 consists in applying [DR10, Theorem 2.8]. We introduce new assumptions. Let $\{r_n\}$ be an intermediate sequence, that is $r_n \rightarrow \infty$ and $r_n/n \rightarrow 0$, and $\{u_n\}$ be a scaling sequence, that is $u_n \rightarrow \infty$. The sequence $\{r_n\}$ is the size of blocks in the blocking method.

Theorem 5.4. *Let $\{X_j, j \in \mathbb{Z}\}$ be a strictly stationary regularly varying sequence with a continuous marginal distribution function F , $\{u_n\}$ be a scaling sequence and $\{r_n\}$ be an intermediate sequence such that Condition $\mathcal{S}(u_n, r_n, \psi)$ holds. Assume that the sequence $\{X_j, j \in \mathbb{Z}\}$ is absolutely regular (i.e. beta-mixing) with coefficients $\{\beta_n, n \geq 1\}$ and there exists a sequence $\{\ell_n\}$ such that*

$$\ell_n \rightarrow \infty, \ell_n/r_n \rightarrow 0, \lim_{n \rightarrow \infty} n\beta_{\ell_n}/r_n = 0. \quad (5.7)$$

Assume that there exists $\delta, \eta > 0$ such that

$$\lim_{n \rightarrow \infty} n\bar{F}(u_n) = \infty, \quad \lim_{n \rightarrow \infty} r_n \bar{F}(u_n) = \frac{r_n}{\{n\bar{F}(u_n)\}^{\delta/2}} = 0, \quad (5.8)$$

$$\sup_{n \geq 1} \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[\left| \psi \left(\frac{\mathbf{X}_{0,h}}{u_n} \right) \right|^{2+\delta} 1_{[0,v]^c}(\mathbf{X}_{0,h}/u_n) \right] < \infty. \quad (5.9)$$

Then, for each $s_0 > 0$, the process \mathbb{M}_n^ψ converges weakly to the centered Gaussian process \mathbb{M}^ψ with covariance function C^ψ defined in (2.6).

Proof of Theorem 5.4. We will check the assumptions of [DR10, Theorem 2.8], that is, conditions (B1), (B2), (C1), (C2), (C3), (D1), (D2'), (D3), (D5) and (D6') therein.

- Conditions (B1) and (B2) hold by stationarity and Condition (5.7).
- Lemmas 5.5 and 5.6 imply (C2) and (C3) and hence the finite dimensional convergence. Since the functionals we consider are sums, condition (C1) is straightforward given the conditions on l_n .
- Condition (D1) (finiteness of the envelope function) holds and the bound (5.12) implies (D2').
- Lemma 5.7 implies (D3).
- As shown in [DR10, Example 3.8], conditions (D5) and (D6') hold for finite dimensional sets of parameters.

\square

We now state and prove the needed lemmas. For conciseness, set, for $\mathbf{v}, \in (\mathbf{0}, \infty)$,

$$\psi_{n,j}(\mathbf{v}) = \psi(\mathbf{X}_{j,j+h}/u_n)1_{(-\infty, \mathbf{v}]^c}(\mathbf{X}_{j,j+h}/u_n), \quad S_n(\mathbf{v}) = \sum_{j=1}^{r_n} \psi_{n,j}(\mathbf{v}).$$

Lemma 5.5. *If Conditions $\mathcal{S}(u_n, r_n, \psi)$ and (5.9) hold, then the series $\sum_{j=1}^{\infty} |c_j^\psi(\mathbf{v}, \mathbf{w})|$ is summable. If moreover Condition (5.8) holds, then, for all $\mathbf{v}, \mathbf{w} \in (\mathbf{0}, \infty)$,*

$$\lim_{n \rightarrow \infty} \frac{\text{cov}(S_n(\mathbf{v}), S_n(\mathbf{w}))}{r_n \bar{F}(u_n)} = C^\psi(\mathbf{v}, \mathbf{w}).$$

Proof. The first statement was already proved in Lemma 2.2. For a fixed integer L , we have

$$\frac{\text{cov}(S_n(\mathbf{v}), S_n(\mathbf{w}))}{r_n \bar{F}(u_n)} = \sum_{1 \leq j, j' \leq r_n} \frac{\mathbb{E}[\psi_{n,j}(\mathbf{v})\psi_{n,j'}(\mathbf{w})]}{r_n \bar{F}(u_n)} + \sum_{1 \leq j, j' \leq r_n} \frac{\mathbb{E}[\psi_{n,j}(\mathbf{v})]\mathbb{E}[\psi_{n,j'}(\mathbf{w})]}{r_n \bar{F}(u_n)}.$$

The terms with products of expectations are negligible since by regular variation the normalization which makes these terms convergent is $\bar{F}^2(u_n)$. Thus we only deal with the main terms. Fix an integer $L > h$. Then, by stationarity,

$$\begin{aligned} \sum_{1 \leq j, j' \leq r_n} \frac{\mathbb{E}[\psi_{n,j}(\mathbf{v})\psi_{n,j'}(\mathbf{w})]}{r_n \bar{F}(u_n)} &= \frac{\mathbb{E}[\psi_{n,0}(\mathbf{v})\psi_{n,0}(\mathbf{w})]}{\bar{F}(u_n)} \\ &+ \sum_{j=1}^L (1 - j/r_n) \frac{\mathbb{E}[\psi_{n,0}(\mathbf{v})\psi_{n,j}(\mathbf{w})] + \mathbb{E}[\psi_{n,j}(\mathbf{v})\psi_{n,0}(\mathbf{w})]}{\bar{F}(u_n)} \\ &+ \sum_{j=L+1}^{r_n} (1 - |j|/r_n) \frac{\mathbb{E}[\psi_{n,0}(\mathbf{v})\psi_{n,j}(\mathbf{w})] + \mathbb{E}[\psi_{n,0}(\mathbf{v})\psi_{n,j}(\mathbf{w})]}{\bar{F}(u_n)}. \end{aligned}$$

By Condition $\mathcal{S}(u_n, r_n, \psi)$, for every $\epsilon > 0$, we can choose L such that

$$\limsup_{n \rightarrow \infty} \sum_{j=L+1}^{r_n} \frac{\mathbb{E}[|\psi_{n,0}(\mathbf{v})\psi_{n,j}(\mathbf{w})|] + \mathbb{E}[|\psi_{n,0}(\mathbf{v})\psi_{n,j}(\mathbf{w})|]}{\bar{F}(u_n)} \leq \epsilon.$$

By regular variation and (5.9), we have

$$\begin{aligned} \lim_{n \rightarrow \infty} \left(\frac{\mathbb{E}[\psi_{n,0}(\mathbf{v})\psi_{n,0}(\mathbf{w})]}{\bar{F}(u_n)} + \sum_{j=1}^L (1 - j/r_n) \frac{\mathbb{E}[\psi_{n,j}(\mathbf{v})\psi_{n,j'}(\mathbf{w})]}{\bar{F}(u_n)} \right) \\ = c_0^\psi(\mathbf{v}, \mathbf{w}) + \sum_{j=1}^L \{c_j^\psi(\mathbf{v}, \mathbf{w}) + c_j^\psi(\mathbf{w}, \mathbf{v})\}. \end{aligned}$$

Since we have proved that the series $\sum_{j=1}^{\infty} |c_j^\psi(\mathbf{v}, \mathbf{w})|$ is convergent, choosing L large enough yields

$$\limsup_{n \rightarrow \infty} \left| \sum_{1 \leq j, j' \leq r_n} \frac{\mathbb{E}[\psi_{n,j}(\mathbf{v})\psi_{n,j'}(\mathbf{w})]}{r_n \bar{F}(u_n)} - C^\psi(\mathbf{v}, \mathbf{w}) \right| \leq \epsilon .$$

Since ϵ is arbitrary, this concludes the proof. \square

Lemma 5.6. *If Conditions $\mathcal{S}(u_n, r_n, \psi)$ (5.8) and (5.9) hold then for all $\epsilon > 0$ and all $\mathbf{v} \in (, \infty]^{h+1}$,*

$$\lim_{n \rightarrow 0} \frac{1}{r_n \bar{F}(u_n)} \mathbb{E} \left[S_n^2(\mathbf{v}) 1_{\{|S_n(\mathbf{v})| > \epsilon \sqrt{n \bar{F}(u_n)}\}} \right] = 0 .$$

Proof. By monotonicity, the second statement is equivalent to the first one with $\mathbf{v} =$ Write

$$\begin{aligned} Z_n(\mathbf{v}, \epsilon) &= \frac{1}{r_n \bar{F}(u_n)} \mathbb{E} \left[S_n^2(\mathbf{v}) 1_{\{|S_n(\mathbf{v})| > \epsilon \sqrt{n \bar{F}(u_n)}\}} \right] \\ &= \frac{1}{r_n \bar{F}(u_n)} \mathbb{E} \left[\sum_{j=1}^{r_n} \psi_{n,j}^2(\mathbf{v}) 1_{\{|S_n(\mathbf{v})| > \epsilon \sqrt{n \bar{F}(u_n)}\}} \right] \\ &\quad + \frac{2}{r_n \bar{F}(u_n)} \sum_{i=1}^{r_n} \sum_{j=i+1}^{r_n} \mathbb{E} \left[\psi_{n,i}(\mathbf{v}) \psi_{n,j}(\mathbf{v}) 1_{\{|S_n(\mathbf{v})| > \epsilon \sqrt{n \bar{F}(u_n)}\}} \right] \\ &= R_n + R_n^* . \end{aligned}$$

Let $\delta \in (0, 1)$ be as in (5.8). Using the elementary bound

$$1_{\{|S_n(\mathbf{v})| > \epsilon \sqrt{n \bar{F}(u_n)}\}} \leq \frac{1}{(\epsilon \sqrt{n \bar{F}(u_n)})^\delta} \sum_{i=1}^{r_n} \psi_{n,i}^\delta(\mathbf{v}) , \quad (5.11)$$

we obtain

$$R_n \leq \frac{1}{r_n \bar{F}(u_n) \{\epsilon \sqrt{n \bar{F}(u_n)}\}^\delta} \mathbb{E} \sum_{j=1}^{r_n} \sum_{i=1}^{r_n} [\psi_{n,j}^2(\mathbf{v}) \psi_{n,i}^\delta(\mathbf{v})] .$$

Using Hölder's inequality with $p = (2 + \delta)/2$ and $q = (2 + \delta)/\delta$ yields

$$R_n \leq \frac{r_n}{(\epsilon \sqrt{n \bar{F}(u_n)})^\delta} \frac{1}{\bar{F}(u_n)} \mathbb{E} [\psi_{n,0}^{2+\delta}(\mathbf{v})] .$$

This bound, (5.8) and (5.9) prove that $\limsup_{n \rightarrow \infty} R_n = 0$. Splitting the sum in j in R_n^*

at $i + L$ where L is an arbitrary integer, and using (5.11) we obtain

$$\begin{aligned}
R_n^* &\leq \frac{2}{r_n \bar{F}(u_n)} \sum_{i=1}^{r_n} \sum_{j=i+1}^{i+L} \mathbb{E} \left[\psi_{n,i}(\mathbf{v}) \psi_{n,j}(\mathbf{v}) 1_{\{|S_n(\mathbf{v})| \geq \epsilon \sqrt{n \bar{F}(u_n)}\}} \right] + \frac{2}{\bar{F}(u_n)} \sum_{j=L}^{r_n} \mathbb{E} [\psi_{n,0}(\mathbf{v}) \psi_{n,j}(\mathbf{v})] \\
&\leq \frac{2}{r_n \bar{F}(u_n)} \frac{1}{(\epsilon \sqrt{n \bar{F}(u_n)})^\delta} \sum_{i=1}^{r_n} \sum_{j=i+1}^{i+L} \sum_{k=1}^{r_n} \mathbb{E} [\psi_{n,i}(\mathbf{v}) \psi_{n,j}(\mathbf{v}) \psi_{n,k}^\delta(\mathbf{v})] \\
&\quad + \frac{2}{\bar{F}(u_n)} \sum_{j=L}^{r_n} \mathbb{E} [\psi_{n,0}(\mathbf{v}) \psi_{n,j}(\mathbf{v})] = R_{n,1}^*(L) + R_{n,2}^*(L).
\end{aligned}$$

Fix $\eta > 0$. By Condition $\mathcal{S}(u_n, r_n, \psi)$ we can choose L such that $\limsup_{n \rightarrow \infty} R_{n,2}^*(L) \leq \eta$. Then, by Hölder's inequality

$$\limsup_{n \rightarrow \infty} R_{n,1}^*(L) \leq \limsup_{n \rightarrow \infty} L \frac{r_n}{(\epsilon \sqrt{n \bar{F}(u_n)})^\delta} \frac{1}{\bar{F}(u_n)} \mathbb{E} [\psi_{n,0}^{2+\delta}(\mathbf{v})] = 0.$$

Altogether, we have proved that $\limsup_{n \rightarrow \infty} Z_n(\mathbf{v}, \epsilon) \leq \eta$ for every $\eta > 0$. This concludes the proof. \square

Fix $s_0 > 0$ and set $S = [s_0 \mathbf{1}, \infty]$ (with $\mathbf{1} = (1, \dots, 1)$) and

$$S_n^* = \sup_{\mathbf{v} \in S} \sum_{i=1}^{r_n} \psi_{n,i}(\mathbf{v}) = \sum_{i=1}^{r_n} \psi_{n,i}(s_0 \mathbf{1}).$$

Thus Lemma 5.6 implies that

$$\lim_{n \rightarrow 0} \frac{1}{r_n \bar{F}(u_n)} \mathbb{E} \left[(S_n^*)^2 1_{\{S_n^* > \epsilon \sqrt{n \bar{F}(u_n)}\}} \right] = 0. \tag{5.12}$$

Define now

$$\rho(\mathbf{v}, \mathbf{w}) = \frac{1}{\bar{F}(u_n)} \int_{\mathbb{R}^{h+1}} \psi^2(\mathbf{x}) |1_{(-\infty, v]c}(\mathbf{x}) - 1_{(-\infty, w]c}(\mathbf{x})| \nu_{0,h}(\mathbf{d}\mathbf{x})$$

Then ρ is a metric on $[0, \infty) \setminus \{0\}$ and

$$\rho(\mathbf{v}, \mathbf{w}) = \lim_{n \rightarrow \infty} \frac{1}{\bar{F}(u_n)} \mathbb{E} [\psi^2(\mathbf{X}_{0,h}/u_n) |1_{(-\infty, v]c}(\mathbf{X}_{0,h}/u_n) - 1_{(-\infty, w]c}(\mathbf{X}_{0,h}/u_n)|],$$

for all $\mathbf{v}, \mathbf{w} \in [0, \infty) \setminus \{0\}$.

Lemma 5.7. *If Conditions $\mathcal{S}(u_n, r_n, \psi)$ and (5.9) hold then, for $s_0 > 0$,*

$$\lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \sup_{\substack{\mathbf{v}, \mathbf{w} \in S \\ \rho(\mathbf{v}, \mathbf{w}) \leq \epsilon}} \frac{1}{r_n \bar{F}(u_n)} \mathbb{E} [(S_n(\mathbf{v}) - S_n(\mathbf{w}))^2] = 0.$$

Proof. Set $\psi_{n,j}(\mathbf{v}, \mathbf{w}) = \psi_{n,j}(\mathbf{v}) - \psi_{n,j}(\mathbf{w})$ for $\mathbf{v}, \mathbf{w} \in (\mathbf{0}, \infty)$. For every integer $L >$, by stationarity we have

$$\begin{aligned} \frac{1}{r_n \bar{F}(u_n)} \mathbb{E}[\{S_n(\mathbf{v}) - S_n(\mathbf{w})\}^2] &\leq \frac{2}{\bar{F}(u_n)} \sum_{j=0}^L \mathbb{E}[|\psi_{n,0}(\mathbf{v}, \mathbf{w})| |\psi_{n,j}(\mathbf{v}, \mathbf{w})|] \\ &\quad + \frac{2}{\bar{F}(u_n)} \sum_{j=L+1}^{r_n} \mathbb{E}[|\psi_{n,0}(\mathbf{v}, \mathbf{w})| |\psi_{n,j}(\mathbf{v}, \mathbf{w})|] . \end{aligned}$$

By Condition $\mathcal{S}(u_n, r_n, \psi)$, for every $\epsilon > 0$, we can choose L such that

$$\limsup_{n \rightarrow \infty} \sup_{\mathbf{v}, \mathbf{w} \in S} \frac{2}{\bar{F}(u_n)} \sum_{j=L+1}^{r_n} \mathbb{E}[|\psi_{n,0}(\mathbf{v}, \mathbf{w})| |\psi_{n,j}(\mathbf{v}, \mathbf{w})|] \leq \epsilon .$$

By Hölder inequality and stationarity we have

$$\frac{2}{\bar{F}(u_n)} \sum_{j=0}^L \mathbb{E}[|\psi_{n,0}(\mathbf{v}, \mathbf{w})| |\psi_{n,j}(\mathbf{v}, \mathbf{w})|] \leq 2(L+1) \frac{1}{\bar{F}(u_n)} \mathbb{E}[\psi_{n,0}^2(\mathbf{v}, \mathbf{w})] . \quad (5.13)$$

This yields

$$\limsup_{n \rightarrow \infty} \frac{2}{\bar{F}(u_n)} \sum_{j=0}^L \mathbb{E}[|\psi_{n,0}(\mathbf{v}, \mathbf{w})| |\psi_{n,j}(\mathbf{v}, \mathbf{w})|] \leq 2(L+1) \rho(\mathbf{v}, \mathbf{w}) ,$$

uniformly on S , hence

$$\limsup_{n \rightarrow \infty} \sup_{\substack{\mathbf{v}, \mathbf{w} \in S \\ \rho(\mathbf{v}, \mathbf{w}) \leq \epsilon}} \frac{2}{\bar{F}(u_n)} \sum_{j=0}^L \mathbb{E}[|\psi_{n,0}(\mathbf{v}, \mathbf{w})| |\psi_{n,j}(\mathbf{v}, \mathbf{w})|] \leq 2(L+1)\epsilon . \quad (5.14)$$

Gathering (5.13) and (5.14) proves Lemma 5.7. \square

5.3 Proof of Theorem 2.1

The proof of Theorem 2.1 consists in showing that Assumption 1, (2.7) and (2.8) imply the conditions of Theorem 5.4.

- Under Assumption 1, the chain $\{\mathbb{Y}_j\}$ is irreducible and geometrically ergodic. This implies that the chain $\{\mathbb{Y}_j\}$ and the sequence $\{X_j\}$ are β -mixing and there exists $c > 1$ such that $\beta_n = O(e^{-cn})$; see [Bra05, Theorem 3.7]. Hence Condition (5.7) holds if we set $l_n = c^{-1} \log(n)$ and $r_n = \log^{1+\eta}(n)$ for an arbitrarily small $\eta > 0$.
- Condition (2.7) and the choice $r_n = \log^{1+\eta}(n)$ imply

$$\lim_{n \rightarrow \infty} n \bar{F}(u_n) = \infty , \quad \lim_{n \rightarrow \infty} r_n \bar{F}(u_n) = \frac{r_n}{\{n \bar{F}(u_n)\}^{\delta/2}} = 0 . \quad (5.15)$$

- Lemma 5.3 shows that Assumption 1, (2.8) and (5.15) imply Condition $\mathcal{S}(u_n, r_n, \psi)L$
- Condition (5.9) follows from regular variation and (2.8). Indeed, for $\mathbf{v} \in [0, \infty) \setminus \{\mathbf{0}\}$, we have

$$\begin{aligned} \limsup_{n \rightarrow \infty} \frac{1}{\bar{F}(u_n)} \mathbb{E} \left[\left| \psi \left(\frac{\mathbf{X}_{0,h}}{u_n} \right) \right|^{2+\delta} 1_{(-\infty, \mathbf{v}]^c}(\mathbf{X}_{0,h}) \right] \\ \leq \limsup_{n \rightarrow \infty} \sum_{0 \leq i, i' \leq h} \mathbb{E}[|X_i/u_n|^{q_i(2+\delta)} 1_{\{X_{i'} > u_n \epsilon\}}] \\ = \sum_{0 \leq i, i' \leq h} \int_{\mathbb{R}^{h+1}} |x|^{q_i(2+\delta)} 1_{\{x_{i'} > \epsilon\}} \boldsymbol{\nu}_{0,h}(dx) . \end{aligned}$$

All these integrals are finite since by assumption, $q_i(2 + \delta) < \alpha$ for δ small enough.

5.4 Proof of Proposition 3.3

Let N_n be the number of returns to the state 1 before time n , that is

$$N_n = \sum_{j=0}^n 1_{\{X_j=1\}} .$$

Set also $T_{-1} = -\infty$, $T_0 = X_0$ and $T_n = X_0 - 1 + Z_1 + \dots + Z_n$ for $n \geq 1$. Then, $\{N_n\}$ is the counting process associated to the delayed renewal process $\{T_n\}$. That is, for $n, k \geq 0$,

$$N_n = k \Leftrightarrow T_{k-1} \leq n < T_k ,$$

Since $\mathbb{E}[Z_0] < \infty$, setting $\lambda = 1/\mathbb{E}[Z_0]$, we have $N_n/n \rightarrow \lambda$ a.s. With this notation, we have, for every $s > 0$,

$$\sum_{j=0}^n 1_{\{X_j > u_n s\}} = (X_0 - [u_n s])_+ + \sum_{j=1}^{N_n} (Z_j - [u_n s])_+ + \zeta_n , \quad (5.16)$$

where $\zeta_n = (n - T_{N_n}) \wedge (Z_{N_n} - [u_n s])_+ - (Z_j - [u_n s])_+$ is a correcting term accounting for the possibly incomplete last portion of the path. Since $\zeta_n = O_P(1)$, it does not play any role in the asymptotics.

- Consider the case $\lim_{n \rightarrow \infty} n \bar{F}_Z(u_n) = 0$. Then, for an integer $m > \lambda$,

$$\begin{aligned} \mathbb{P} \left(\sum_{j=1}^{N_n} (Z_j - [u_n s])_+ \neq 0 \right) &\leq \mathbb{P}(N_n > mn) + \mathbb{P} \left(\sum_{j=1}^{mn} (Z_j - [u_n s])_+ \neq 0 \right) \\ &\leq \mathbb{P}(N_n > mn) + \mathbb{P}(\exists j \in \{1, \dots, mn\}, Z_j > [u_n s]) \\ &\leq \mathbb{P}(N_n > mn) + mn \bar{F}_Z([u_n s]) \rightarrow 0 . \end{aligned}$$

Furthermore,

$$\mathbb{P}((X_0 - [u_n s])_+ \neq 0) = \bar{F}_X([u_n s]) \rightarrow 0 .$$

This proves our first claim.

We proceed with the case $\lim_{n \rightarrow \infty} n\bar{F}_Z(u_n) = \infty$. Using (3.4) and (5.16) we have

$$\begin{aligned} \sum_{j=0}^n \{1_{\{X_j > u_n s\}} - \mathbb{P}(X_0 > u_n s)\} &= \sum_{j=1}^{N_n} \{(Z_j - [u_n s])_+ - \mathbb{E}[(Z_0 - [u_n s])_+]\} \\ &\quad + (X_0 - [u_n s])_+ + \{N_n - \lambda n\} \mathbb{E}[(Z_0 - [u_n s])_+] . \end{aligned} \quad (5.17)$$

- Consider the case $n\bar{F}_Z(u_n) \rightarrow \infty$ and $\beta \in (1, 2)$. Since $\lim_{n \rightarrow \infty} \mathbb{E}[(Z_0 - [u_n s])_+] = 0$, we obtain, for every $s > 0$,

$$\begin{aligned} a_n^{-1} \sum_{j=0}^n \{1_{\{X_j > u_n s\}} - \mathbb{P}(X_0 > u_n s)\} \\ = a_n^{-1} \sum_{j=1}^{N_n} \{Z_j - \mathbb{E}[Z_0]\} + a_n^{-1} \sum_{j=1}^{N_n} \{Z_j \wedge [u_n s] - \mathbb{E}[Z_0 \wedge [u_n s]]\} + o_P(1) . \end{aligned}$$

By regular variation of \bar{F}_Z , we obtain

$$\text{var} \left(\sum_{j=1}^n \{Z_j \wedge [u_n s] - \mathbb{E}[(Z_0 \wedge [u_n s])]\} \right) = O(u_n^2 n \bar{F}_Z(u_n)) .$$

The regular variation of \bar{F} and the conditions $n\bar{F}_Z(u_n) \rightarrow \infty$ and $n\bar{F}_Z(a_n) \rightarrow 1$ imply that $u_n/a_n \rightarrow 0$. Define $h(x) = x\sqrt{\bar{F}_Z(x)}$. The function h is regularly varying at infinity with index $1 - \beta/2 > 0$ and thus

$$\lim_{n \rightarrow \infty} \frac{u_n \sqrt{n\bar{F}_Z(u_n)}}{a_n} = \lim_{n \rightarrow \infty} \frac{u_n \sqrt{\bar{F}_Z(u_n)}}{a_n \sqrt{\bar{F}_Z(a_n)}} = \lim_{n \rightarrow \infty} \frac{h(u_n)}{h(a_n)} = 0 .$$

This yields

$$a_n^{-1} \sum_{j=0}^n \{1_{\{X_j > u_n s\}} - \mathbb{P}(X_0 > u_n s)\} = a_n^{-1} \sum_{j=1}^{N_n} \{Z_j - \mathbb{E}[Z_0]\} + o_P(1) ,$$

where the $o_P(1)$ term is locally uniform with respect to $s > 0$. Since the distribution of Z_0 is in the domain of attraction of the β -stable law and the sequence $\{Z_j\}$ is i.i.d., $\{a_n^{-1}(T_{[ns]} - \lambda^{-1}s), s > 0\} \Rightarrow \Lambda$, where Λ is a mean zero, totally skewed to the right β -stable Lévy process, and the convergence holds with respect to the J_1 topology on

compact sets of $(0, \infty)$. Since $\lim_{n \rightarrow \infty} N_n/n = \lambda$ a.s., and since a Lévy process is stochastically continuous, this yields, by [Whi02, Proposition 13.2.1],

$$a_n^{-1} \sum_{j=1}^{N_n} \{Z_j - \mathbb{E}[Z_0]\} \xrightarrow{d} \Lambda(\lambda)$$

This proves the second claim.

- Consider now the case $\beta > 2$. In that case, Vervaat's Lemma implies that $(N_n - \lambda n)/\sqrt{n}$ converges weakly to a gaussian distribution. Thus, (5.17) combined with $\mathbb{E}[(Z_0 - [u_n s])_+] = O(u_n \bar{F}_Z(u_n))$, yields

$$(N_n - \lambda n) \mathbb{E}[(Z_0 - [u_n s])_+] = O_P(u_n \bar{F}_Z(u_n) \sqrt{n}).$$

Next, we apply the Lindeberg central limit theorem for triangular arrays of independent random variables to prove that

$$\frac{1}{u_n \sqrt{n \bar{F}_Z(u_n)}} \sum_{j=1}^n \{(Z_j - [u_n s])_+ - \mathbb{E}[(Z_0 - [u_n s])_+]\} \xrightarrow{d} N\left(0, \frac{2s^{1-\alpha}}{\alpha(\alpha-1)}\right).$$

By regular variation of \bar{F}_Z , we have, for all $\delta \in [2, \beta)$,

$$\mathbb{E}[(Z_0 - u_n s)_+^\delta] \sim C_\delta u_n^\delta \bar{F}_Z(u_n) s^{\delta-\beta},$$

with $C_\delta = \delta \int_1^\infty (z-1)^{\delta-1} z^{-\beta} dz$. Set

$$Y_{n,j}(s) = \frac{1}{u_n \sqrt{n \bar{F}_Z(u_n)}} \{(Z_j - [u_n s])_+ - \mathbb{E}[(Z_0 - [u_n s])_+]\}.$$

The previous computations yield, for $\delta \in (2, \beta)$ and $s > 0$,

$$\begin{aligned} \lim_{n \rightarrow \infty} n \text{var}(Y_{n,1}(s)) &= \frac{2s^{2-\beta}}{(\beta-1)(\beta-2)}, \\ n \mathbb{E}[|Y_{n,1}|^\delta] &= O\left(\frac{n u_n^\delta \bar{F}_Z(u_n)}{u_n^\delta \{n \bar{F}_Z(u_n)\}^{\delta/2}}\right) = O(\{n \bar{F}_Z(u_n)\}^{1-\delta/2}) = o(1). \end{aligned}$$

We conclude that the Lindeberg central limit theorem holds. Convergence of the finite dimensional distribution is done along the same lines and tightness with respect to the J_1 topology on $(0, \infty)$ is proved by applying [Bil99, Theorem 13.5].

5.5 Variance of the Hill estimator

Let the processes \mathbb{W} and \mathbb{B} be defined on $[0, 1]$ by $\mathbb{W}(s) = \mathbb{G}(s^{-1/\alpha})$ and $\mathbb{B}(s) = \mathbb{W}(s) - s\mathbb{W}(1)$. Then, the limiting distribution of the Hill estimator is that of the random variable $\alpha^{-1}\mathbb{Z}$, with

$$\mathbb{Z} = \int_0^1 \mathbb{B}(s) \frac{ds}{s}.$$

In the case of extremal independence \mathbb{B} is the standard Brownian bridge and $\text{var}(\mathbb{Z}) = 1$. In the general case, let γ denote the autocovariance function of \mathbb{W} , i.e.

$$\gamma(s, t) = s \wedge t + \sum_{j=1}^{\infty} \mathbb{E} [(s\Theta_j^\alpha) \wedge t + (t\Theta_j^\alpha) \wedge s \mid \Theta_0 = 1] .$$

This yields $\gamma(1, 1) = 1 + 2 \sum_{j=1}^{\infty} \mathbb{E}[\Theta_j^\alpha \wedge 1 \mid \Theta_0 = 1]$ and

$$\begin{aligned} \text{var}(\mathbb{Z}) &= \int_0^1 \int_0^1 \frac{\gamma(s, t) - s\gamma(t, 1) - t\gamma(s, 1) + st\gamma(1, 1)}{st} dsdt \\ &= \gamma(1, 1) + \int_0^1 \int_0^1 \frac{\gamma(s, t)}{st} dsdt - 2 \int_0^1 \frac{\gamma(s, 1)}{s} ds . \end{aligned}$$

Note now that $\gamma(s, t) = \gamma(t, s)$ and $\gamma(s, t) = s\gamma(1, t/s)$, so that

$$\begin{aligned} \int_0^1 \int_0^1 \frac{\gamma(s, t)}{st} dsdt &= 2 \int_0^1 \int_0^t \frac{\gamma(s, t)}{st} dsdt = 2 \int_0^1 \int_0^t \frac{\gamma(s/t, 1)}{s} dsdt \\ &= 2 \int_0^1 \int_0^1 \frac{\gamma(u, 1)}{u} dudt = 2 \int_0^1 \frac{\gamma(u, 1)}{u} du . \end{aligned}$$

This shows that $\text{var}(\mathbb{Z}) = \gamma(1, 1) = 1 + 2 \sum_{j=1}^{\infty} \mathbb{E}[\Theta_j^\alpha \wedge 1 \mid \Theta_0 = 1]$, which proves (4.6).

Acknowledgements The research of Rafal Kulik was supported by the NSERC grant 210532-170699-2001. The research of Philippe Soulier and Olivier Wintenberger was partially supported by the ANR14-CE20-0006-01 project AMERISKA network. Philippe Soulier also acknowledges support from the LABEX MME-DII.

References

- [Bil99] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons Inc., New York, second edition, 1999.
- [Bra05] Richard C. Bradley. Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys*, 2:107–144, 2005.
- [BS09] Bojan Basrak and Johan Segers. Regularly varying multivariate time series. *Stochastic Process. Appl.*, 119(4):1055–1080, 2009.
- [CEdHZ15] Juan-Juan Cai, John J.H.J Einmahl, Laurens de Haan, and Chen Zhou. Estimation of the marginal expected shortfall: the mean when a related variable is extreme. *Journal of the Royal Statistical Society. Series B.*, 77(2):417–442, 2015.

- [Cli07] Daren B. H. Cline. Regular variation of order 1 nonlinear AR-ARCH models. *Stochastic Processes and their Applications*, 117(7):840–861, 2007.
- [DM09] Richard A. Davis and Thomas Mikosch. The extremogram: A correlogram for extreme events. *Bernoulli*, 38A:977–1009, 2009. Probability, statistics and seismology.
- [DMS14] Randal Douc, Eric Moulines, and David S. Stoffer. *Nonlinear time series*. Chapman & Hall/CRC Texts in Statistical Science Series. Chapman & Hall/CRC, Boca Raton, FL, 2014. Theory, methods, and applications with R examples.
- [DR10] Holger Drees and Holger Rootzén. Limit theorems for empirical processes of cluster functionals. *Ann. Statist.*, 38(4):2145–2186, 2010.
- [Dre00] Holger Drees. Weighted approximations of tail processes for β -mixing random variables. *The Annals of Applied Probability*, 10(4):1274–1301, 2000.
- [Dre03] Holger Drees. Extreme quantile estimation for dependent data, with applications to finance. *Bernoulli*, 9(4):617–657, 2003.
- [FT85] Paul D. Feigin and Richard L. Tweedie. Random coefficient autoregressive processes: a Markov chain analysis of stationarity and finiteness of moments. *J. Time Ser. Anal.*, 6(1):1–14, 1985.
- [Hsi93] Tailen Hsing. Extremal index estimation for a weakly dependent stationary sequence. *The Annals of Statistics*, 21(4):2043–2071, 1993.
- [KS11] Rafał Kulik and Philippe Soulier. The tail empirical process for long memory stochastic volatility sequences. *Stochastic Processes and their Applications*, 121(1):109 – 134, 2011.
- [MT09] Sean Meyn and Richard L. Tweedie. *Markov chains and stochastic stability*. Cambridge University Press, Cambridge, second edition, 2009.
- [MW13] Thomas Mikosch and Olivier Wintenberger. Precise large deviations for dependent regularly varying sequences. *Probability Theory and Related Fields*, 2013.
- [MW14] Thomas Mikosch and Olivier Wintenberger. The cluster index of regularly varying sequences with applications to limit theory for functions of multivariate Markov chains. *Probability Theory and Related Fields*, 159(1-2):157–196, 2014.
- [Roo88] Holger Rootzén. Maxima and exceedances of stationary Markov chains. *Adv. in Appl. Probab.*, 20(2):371–390, 1988.

- [Roo09] Holger Rootzén. Weak convergence of the tail empirical process for dependent sequences. *Stoch. Proc. Appl.*, 119(2):468–490, 2009.
- [RRSS06] Gareth O. Roberts, Jeffrey S. Rosenthal, Johann Segers, and Bruno Sousa. Extremal indices, geometric ergodicity of markov chains, and mcmc. *Extremes*, 9(3-4):213–229, 2006.
- [Smi92] Richard L. Smith. The extremal index for a Markov chain. *Journal of Applied Probability*, 29(1):37–45, 1992.
- [Whi02] Ward Whitt. *Stochastic-process limits*. Springer-Verlag, New York, 2002.

R. Kulik: rkulik@uottawa.ca
University of Ottawa,
Department of Mathematics and Statistics,
585 King Edward Av.,
K1J 8J2 Ottawa, ON, Canada

Ph. Soulier: philippe.soulier@u-paris10.fr
Université de Paris Ouest-Nanterre,
Département de mathématiques et informatique,
Laboratoire MODAL'X EA3454,
92000 Nanterre, France

O. Wintenberger: wintenberger@math.ku.dk
University of Copenhagen,
Department of Mathematics,
Universitetspark 5,
2100 Copenhagen, DENMARK