EMPIRICAL CLT FOR CLUSTER FUNCTIONALS UNDER WEAK DEPENDENCE*

PAUL DOUKHAN[1,2] AND JOSÉ-GREGORIO GÓMEZ[1]

Abstract. We prove empirical central limit theorems (CLT) for extreme values cluster functionals empirical processes in the sense of the tough paper Drees and Rootzén (2010). Contrary to those authors we dont restrict to $\beta-$mixing samples. For this we use coupling properties enlightened for Dedecker & Prieur's $\tau-$dependence coefficients. We also explicit the asymptotic behavior of specific clusters. As an example we develop the number of excesses; it gives a complete example of a cluster functional for a non-mixing "reasonable model" (an AR(1)-process) for which results such as ours are definitely needed. In particular the expression of the limit Gaussian process is developed. Also we include in this paper some results of Drees (2011) for the extremal index and some simulations for this index to demonstrate the accuracy of this technique.

Keywords and phrases: Extremes, clustering of extremes, cluster functional of extremes, extremal index, uniform central limit theorem, τ-weak dependence, tail empirical process.

1. Introduction

We use the scheme in [Segers, 2003] and in [Drees & Rootzén, 2010] for defining clusters functionals. Let $d \geqslant 1$, the set $E \subset \mathbb{R}^d$ is measurable ($E \in \mathcal{B}(\mathbb{R}^d)$) and $0 \in E$. Let us consider E-valued normalized random variables $(X_{n,i})_{1\leqslant i\leqslant n, n\in\mathbb{N}}$, defined on some probability space $(\Omega, \mathcal{A}, \mathbb{P})$, which are row-wise stationary, that is $(X_{n,i})_{1\leqslant i\leqslant n}$ is stationary for each $n \in \mathbb{N}$. Those normalized random variables $X_{n,i}$, are built from another random process $(X_j)_{j\in\mathbb{N}}$, in such a way that the normalization maps all non-extreme values to zero. Additionally, it should satisfy that the sequence of conditional

Date: June 21, 2015.

*THIS RESEARCH HAS BEEN CONDUCTED AS PART OF THE PROJECT LABEX MME-DII (ANR11-LBX-0023-01)

[1]Department of Mathematics, University Cergy-Pontoise. 95000 Cergy - Pontoise, France.

[2]Institut Universitaire de France (IUF).

E-mails: doukhan@u-cergy.fr, jose.gomez@u-cergy.fr

AMS 2000 Subject Classifications: Primary 60G70; secondary 60F05, 60F17, 62G32.

distributions of $X_{n,i}$ given that $X_{n,i} \neq 0$ (*i.e.* $\overline{P}_n(x) = \mathbb{P}\{X_{n,i} > x | X_{n,i} \neq 0\}$) converge weakly to some non-degenerate limit.

For real-valued random variables a typical normalization is as follows. Let $(X_i)_{i\in\mathbb{N}}$ be a stationary process in $\mathbb{R}$ with marginal cumulative distribution function F, and let $(u_n)_{n\in\mathbb{N}}$ be an increasing sequence of thresholds such that $u_n \uparrow x_F$ with

$$x_F = \sup\{x \in \mathbb{R} : F(x) < 1\}, \qquad v_n := \mathbb{P}\{X_1 > u_n\} \underset{n\to\infty}{\longrightarrow} 0.$$

It is clear that, for each $x > 0$, the tail distribution function for X_i:

$$\overline{P}_n(x) = \mathbb{P}\{X_i - u_n > x | X_i > u_n\}$$

is degenerated as $n \to \infty$ (asymptotically this is a Dirac distribution). The following normalization is often used:

$$X_{n,i} = \left(\frac{X_i - u_n}{a_n}\right)_+ = \max\left\{\frac{X_i - u_n}{a_n}, 0\right\}, \qquad \text{for } 1 \leqslant i \leqslant n, \tag{1.1}$$

here $(a_n)_{n\in\mathbb{N}}$ is a sequence of positive norming constants, depending on u_n. Moreover for this case the sequence of cumulative distributions $\overline{P}_n(x) = \mathbb{P}\{X_{n,i} > x | X_i > u_n\}$ converges to a Pareto distribution for $x > 0$, see Section 4 of [Segers, 2003].

If the process $(X_i)_{i\in\mathbb{N}}$ is $\mathbb{R}^d$−valued ($d \geq 1$) then two applications write with:

$$X_{n,i} = \left(\frac{\|X_i\| - u_n}{a_n}\right)_+, \text{ for } 1 \leqslant i \leqslant n, \tag{1.2}$$

$$X_{n,i} = \left(\left(\frac{X_{1i} - u_n}{a_n}\right)_+, \left(\frac{X_{2i} - u_n}{a_n}\right)_+, \ldots, \left(\frac{X_{di} - u_n}{a_n}\right)_+\right), \tag{1.3}$$

for some norm $\|\cdot\|$ on the Euclidean space $\mathbb{R}^d$ $X_i = (X_{1i}, X_{2i}, \ldots, X_{di})$ and $(u_n)_{n\in\mathbb{N}}$, $(a_n)_{n\in\mathbb{N}}$ are defined as in eqn. (1.1). This could be interpreted as follows: we suppose that X_i is a vector of a cycle of d records per unit time i (for example, i is the i-th day and $d = 24$, hours per day). In particular if $\|X_i\| := \|X_i\|_1 = \sum_{j=1}^d |X_{ij}|$ then the expression in eqn. (1.2) is non-zero if the sum of the records exceeds the threshold u_n and for (1.3) $X_{n,i}$ this is the vector of excesses over a threshold u_n for each coordinate. Section 3 in [Drees & Rootzén, 2010] gives another example of a normalization of d consecutive excesses of real random variables $(X_i)_{i\in\mathbb{N}}$ *i.e.*

$$X_{n,i} = \left(\left(\frac{X_i - u_n}{a_n}\right)_+, \left(\frac{X_{i+1} - u_n}{a_n}\right)_+, \ldots, \left(\frac{X_{i+d-1} - u_n}{a_n}\right)_+\right). \tag{1.4}$$

Note that this example is interesting because it takes into account d consecutive extreme values, which contains useful information on the extremal dependence structure. That is is a general feature for many real applications. Namely:

(1) if d consecutive days of rain are observed in a given city, the volume of precipitated water may be larger than the volume of water that can be drained (through sewers, soil, rivers, etc.),
(2) If d very large claims are reported to an insurance company in a very small time interval (with respect to typical cases) this can be a risk with respect to the response capacity of the insurance company, and
(3) If d consecutive days of low temperatures a observed in a given city, then the power consumption (due to heating, etc.) endangers the response capacity of the company in charge of the energy distribution.

On the order hand, a typical example of a empirical process of extreme values cluster functionals is the tail empirical process:

$$T_n(x) = \frac{1}{\sqrt{nv_n}} \sum_{i=1}^{n} \big(\mathbb{1}\{X_{n,i} > x\} - \mathbb{P}(X_{n,1} > x) \big), \qquad x \geqslant 0, \tag{1.5}$$

with $X_{n,i}$ defined as in (1.1). This process has beed considerd by Drees and Rootzén under certain conditions (in particular for α and β-mixing conditions) they prove its uniform convergence to a Gaussian process T under additional conditions.
For example they prove the convergence of this tail empirical process for the cases of k-dependent sequences or stable AR(1)-processes [Rootzén, 1995], and ARCH(1)-processes [Drees, 2002, Drees, 2003] and applications for solutions of stochastic difference equations [Drees, 2000, Drees, 2002, Drees, 2003]. They use the extreme cluster functionals setting in [Segers, 2003] to generalize such empirical processes under β-mixing in [Drees & Rootzén, 2010].

However, note that the AR(1)-process, solution of the recursion:

$$X_k = \frac{1}{b}\big(X_{k-1} + \xi_k\big), \quad k \in \mathbb{Z}, \tag{1.6}$$

where $b \geqslant 2$ is an integer and $(\xi_k)_{k\in\mathbb{N}}$ are independent and uniformly distributed random variables on the set $\{0, 1, \ldots, b-1\}$ is not even $\alpha-$mixing, as this is shown in [Andrews, 1984] for $b = 2$ and in [Ango Nze & Doukhan, 2004] for $b > 2$. *The results in* [Drees & Rootzén, 2010] *can thus not be used here!* This is thus useful to improve on the CLT for empirical processes of extreme cluster functionals proposed by [Drees & Rootzén, 2010] for more general classes of weakly dependent processes. Here we make use of the coupling results of [Dedecker & Prieur, 2004a, Dedecker & Prieur, 2005] under τ-dependence assumptions.

This paper is organized as follows. In Section 2, we recall basic definitions and notations for cluster functionals and for the empirical process of cluster functionals. In Section 3 we define the τ-weak dependence coefficients and then Section 4 is concerned with the assumption to derive our main functional CLT for the cluster empirical process; namely asymptotic tightness, asymptotic equicontinuity and the assumptions to obtain a fidi CLT for empirical processes of cluster functionals. Those results are shown in Section 5. In Section 6 we develop a example similar to (1.5) for the multidimensional case for the case of AR(1)-inputs. Also a simulation study for the extremal index to demonstrate the accuracy of this technique. The proofs are reported in a last section.

2. Empirical Process of Cluster Functionals

As it was mentioned in the introduction, we consider a triangular array of row-wise stationary random variables $(X_{n,i})_{1\leqslant i\leqslant n, n\in\mathbb{N}}$ with normalized marginals taking its values in a measurable subset E of $(\mathbb{R}^d, \mathcal{B}(\mathbb{R}^d))$ and defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$.

Before defining the "empirical process of cluster functionals", we first define a cluster functional. The following definition is due to [Yun, 2000] and [Segers, 2003] for the univariate case, and to [Drees & Rootzén, 2010] for the more general multi-dimensional case:

Definition 2.1 (Cluster Functional, [Drees & Rootzén, 2010])**.**

- *We define the set of finite-length sequences with values in E:*
$$E_\cup := \bigcup_{r\in\mathbb{N}} E^r$$
equipped with the σ-field $\mathcal{E}_\cup$ induced by Borel-σ-fields on E^r, for $r\in\mathbb{N}$.
- *For $r\in\mathbb{N}$, let $Y=(X_1, X_2, \ldots, X_r)\in E^r\subset(\mathbb{R}^d)^r$. The* ***core*** *$Y^c\in E_\cup$ of Y is defined by*
$$Y^c = \begin{cases} (X_l)_{r_1\leqslant l\leqslant r_2}, & \text{if } Y\neq 0_r \text{ (the null element in } \mathbb{R}^r\text{)} \\ 0, & \text{otherwise} \end{cases}$$
here $r_1 := \min\{i\in\{1,\ldots,r\} : X_i\neq 0\}$ (first extreme value of the block Y) and $r_2 := \max\{i\in\{1,\ldots,r\} : X_i\neq 0\}$ (last extreme value of the block Y).
- *A measurable map $f : (E_\cup, \mathcal{E}_\cup) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is called a* ***cluster functional*** *if*
$$f(Y) = f(Y^c), \qquad \text{for all } Y\in E_\cup, \quad \text{and} \quad f(0_r)=0 \ (\forall r\geqslant 1).$$

Remark 1. Note that the cluster map $Y\longmapsto Y^c$ (defined in [Segers, 2003]) is not only a functional dependent on the extreme values set of the block, but it also depends on the null values set between the first and last extreme value of the block, *i.e.* the **cluster** is really the smallest sub-block Y^c containing all extreme values of the block and including also all the non-extreme values in-between, e.g. $f(0,1,1,0,0,1,0,0) = f(1,1,0,0,1)$.

Example 2.1. Consider the following notation. If $u = (u_1, \dots, u_d)$, $v = (v_1, \dots, v_d)$ are vectors in $\mathbb{R}^d$ say that $u \leq v$ if and only if $u_i \leq v_i$ for all $i = 1, \dots, d$. So, if $E = \mathbb{R}^d$ and $x \in T \subset E$, we have the following list of cluster functionals:

(2.1.1) Sum of excesses over x:

$$f_x(x_1, \dots, x_r) = \sum_{i=1}^{r} (x_i - x)\mathbb{1}\{x_i > x\},$$

(2.1.2) Number of excesses over x:

$$f_x(x_1, \dots, x_r) = \sum_{i=1}^{r} \mathbb{1}\{x_i > x\}.$$

Usually if $\phi : (E, \mathcal{E}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ denotes a function such that $\phi(0) = 0$ then

$$f_\phi(x_1, \dots, x_r) = \sum_{i=1}^{r} \phi(x_i) \tag{2.7}$$

is a cluster functional.

(2.1.3) Maximal excess over x:

$$f_x(x_1, \dots, x_r) = \max\{(x_1 - x)_+, \dots, (x_r - x)_+\},$$

where $(y - x)_+ := \|y - x\|\mathbb{1}\{y > x\}$;

(2.1.4) Number of up-crossings at x:

$$f_x(x_1, \dots, x_r) = \mathbb{1}\{x_1 > x\} + \mathbb{1}\{x_1 < x, x_2 > x\} + \dots + \mathbb{1}\{x_{r-1} < x, x_r > 0\}.$$

Example 2.2. For $E = \mathbb{R}^+$ and $x > 0$,

(2.2.5) Balanced periods[3] over x,

$$f_x(x_1, \dots, x_r) = \mathbb{1}\left\{\sum_{i=1}^{r} (x_i - x)\mathbb{1}\{x_i > 0\} = 0\right\}.$$

Example 2.3. We consider $E = \mathbb{R}^+$. If for each $p, q \in \{1, 2, \dots, r\}$, we denote $H_{p,q} = \{x_p, x_{p+1}, \dots, x_q\} \subseteq \{x_1, \dots, x_r\}$ such that $y > 0$, $\forall y \in H_{p,q}$. Then we can define the following functional:

(2.3.6) Maximum sum (greater than the level $u > 0$) of consecutive excesses over x,

$$f_{x,u}(x_1, \dots, x_r) = \max_{1 \leqslant p < q \leqslant r} \sum_{y \in H_{p,q}} (y - x)_+ \mathbb{1}\left\{\sum_{y \in H_{p,q}} (y - x)_+ > u\right\},$$

[3] In particular, many thanks are to Didier Dacunha-Castelle for managing electric consumptions questions, this functional is of an important use for electricity production problems.

Example 2.4. Let (E, d) be a metric space. If $\gamma_i : E \longrightarrow \mathbb{R}^d$ such that $\gamma_i(0) = 0$ for $i \in \{1, \dots, r\}$, we define functionals $g_x : E_\cup \longrightarrow \mathbb{R}$ for $x \in T$ as the previous cases:

$$g_x(x_1, \dots, x_r) = f_x(\gamma_1(x_1), \dots, \gamma_r(x_r)),$$

where f_x is any functional of the above list (2.1.1) - (2.3.6).

For applications's sake we expect that cluster functionals $f(\cdot)$ to satisfy:

(1) their domain is a vector of arbitrary length, with at least one non-zero component;
(2) it depends only on the extreme values and on the occurrence of non-extreme values in between two extreme values of the vector.

Remark 2. Note that the three first previous examples ((2.1.1)-(2.1.3)) satisfy the property of "invariance under permutations" *i.e.*:

$$f(x_1, \dots, x_r) = f(x_{\sigma(1)}, \dots, x_{\sigma(r)}),$$

for any permutation $\sigma : \{1, \dots, r\} \longrightarrow \{1, \dots, r\}$.
Therefore, these functionals (neither any cluster functional) will not give any information about the positions of the extreme values cluster.

Definition 2.2 (Empirical Process of Cluster Functionals [Drees & Rootzén, 2010])**.**

(1) *We define $Y_{n,j}$ as the j-th block of r_n consecutive values of the n-th row of $(X_{n,i})$. Thus there are $m_n := [n/r_n] = \max\{j \in \mathbb{N} : j \leqslant n/r_n\}$ blocks*

$$Y_{n,j} := (X_{n,i})_{(j-1)r_n+1 \leqslant i \leqslant jr_n} \quad \text{for} \quad 1 \leqslant j \leqslant m_n \tag{2.8}$$

of length r_n. Since $(X_{n,i})_{1 \leqslant i \leqslant n}$ is stationary for each n, then we can denote by Y_n to the "generic block" such that

$$Y_n \overset{\mathcal{D}}{=} Y_{n,1}. \tag{2.9}$$

Moreover the block-length r_n tends to infinity in such way that $r_n \ll n$.

(2) *We denote by $\mathcal{F}$ a class of "cluster functionals".*
Then the ***"empirical process Z_n of cluster functionals"*** *is the process $(Z_n(f))_{f \in \mathcal{F}}$ defined by*

$$Z_n(f) := \frac{1}{\sqrt{n v_n}} \sum_{j=1}^{m_n} (f(Y_{n,j}) - \mathbb{E} f(Y_{n,j})), \tag{2.10}$$

where $v_n := \mathbb{P}\{X_{n,1} \neq 0\} \longrightarrow 0$.

Drees and Rootzén [Drees & Rootzén, 2010] have proved CLTs for this process. In particular, they have proved a CLT for the Fidis of $(Z_n(f))_{f \in \mathcal{F}}$ by using the Bernstein blocks technique together with a β-mixing coupling condition to boil down convergence to convergence of sums over i.i.d. blocks through [Eberlein, 1984]'s technique involving the metric of total variation. [Drees & Rootzén, 2010] prove an uniform CLT by using [Van Der Vaart & Wellner, 1996]'s tightness criteria and asymptotic equicontinuity conditions together with a fidi CLT.

We aim at extending their CLT's for the empirical process $(Z_n(f))_{f\in\mathcal{F}}$, since the family of mixing processes is still very restrictive. One particular and really simple example of a non-mixing process is the AR(1)-process (1.6). We derive the same results as in [Drees & Rootzén, 2010] and some applications as in [Drees, 2011] under much weaker dependence conditions including *eg.* this example.
The $\tau-$weak dependence introduced by [Dedecker & Prieur, 2004a] holds for the Example in eqn. (1.6), as well as more generally for Bernoulli shifts processes and Markov chains.

3. τ-weak dependence [Dedecker & Prieur, 2004a]

Definition 3.1. *Let $(\Omega,\mathcal{A},\mathbb{P})$ be a probability space, and $\mathcal{M}$ a σ-algebra of $\mathcal{A}$. Let (E,δ) be a Polish space endowed with its metric. For any E-valued random variable $X\in\mathbb{L}^p$ (i.e. X satisfies $\|X\|_p := (\mathbb{E}|X|^p)^{1/p}<\infty$) Dedecker and Prieur defined the coefficient τ_p as:*

$$\tau_p(\mathcal{M},X) := \|\sup\{\mathbb{E}[h(X)|\mathcal{M}]-\mathbb{E}[h(X)] : h\in\Lambda(E,\delta)\}\|_p \tag{3.11}$$

where $\Lambda(E,\delta)$ denotes the class of all Lipschitz functions $h:E\longrightarrow\mathbb{R}$ such that

$$Lip(h) := \sup_{x\neq y}|h(x)-h(y)|/\delta(x,y)\leqslant 1.$$

Let $\mathcal{X}=(X_{n,i})_{1\leqslant i\leqslant n,n\in\mathbb{N}}$ be a triangular array of $\mathbb{L}^p$-integrable E-valued random variables, and $(\mathcal{M}_i)_{i\in\mathbb{Z}}$ be a sequence of σ-algebras of $\mathcal{A}$.
Then, for any $n\in\mathbb{N}$ Dedecker and Prieur defined the coefficient:

$$\tau_{p,n}(k) := \sup_{l\geqslant 1} l^{-1}\sup\{\tau_p(\mathcal{M}_i,(X_{n,j_1},\dots,X_{n,j_l})) : i+k\leqslant j_1<\cdots<j_l\leqslant n\}, \tag{3.12}$$

where we consider the distance $\delta_l(x,y)=\sum_{i=1}^l\delta(x_i,y_i)$ on E^l. Moreover, we say that $\mathcal{X}$ is τ_p-weakly dependent if

$$\lim_{k\longrightarrow\infty}\limsup_{n\longrightarrow\infty}\tau_{p,n}(k)=0. \tag{3.13}$$

Example 3.1 (Causal Bernoulli shifts)**.** Let $(\xi_i)_{i\in\mathbb{Z}}$ be a sequence of i.i.d.r.v's. (independent and identically distributed random variables) with values in a measurable space D. Assume that there exists a function $H: D^{\mathbb{N}}\longrightarrow\mathbb{R}$, such that $H(\xi_0,\xi_{-1},\dots)$ is defined almost surely. Then the stationary sequence $(X_i)_{i\geqslant 0}$ defined by $X_i=H(\xi_i,\xi_{i-1},\dots)$ is called a causal Bernoulli shifts.

Let $(\xi_i')_{i\in\mathbb{Z}}$ be an independent copy of the i.i.d. sequence $(\xi_i)_{i\in\mathbb{Z}}$. Consider a non-increasing sequence $(\Delta_{p,n}(i))_{i\geqslant 0}$, such that

$$\Delta_{p,n}(i) \geqslant \|X_{n,i} - X'_{n,i}\|_p, \tag{3.14}$$

for some $p \in [1,\infty]$, where $X_{n,i}$ and $X'_{n,i}$ denote the extreme normalizations of $X_i = H(\xi_i, \xi_{i-1}, \ldots)$ and $X'_i = H(\xi_i, \ldots, \xi_1, \xi'_0, \xi'_{-1}, \ldots)$, respectively.
If $\mathcal{M}_i = \sigma(X_j : j \leqslant i)$, then the coefficient $\tau_{1,n}$ of $(X_{n,i})_{1\leqslant i\leqslant n}$ is bounded above by $\Delta_{1,n}$.

In particular, if the triangular array $(X_{n,i})_{1\leqslant i\leqslant n, n\in\mathbb{N}}$, is defined as in (1.4) and if we consider a decreasing sequence $(\delta_{p,i})_{i\geqslant 0}$ such that

$$\|H(\xi_i, \xi_{i-1}, \ldots) - H(\xi_i, \ldots, \xi_1, \xi'_0, \xi'_{-1}, \ldots)\|_p \leqslant \delta_{p,i} \tag{3.15}$$

for $p \in [1,\infty]$. Then, $\tau_{1,n}(k)$ is bounded by

$$\tilde{\Delta}_n(k) := \frac{d}{a_n}\left(\delta_{p,k} v_n^{1/q} + 2\delta_{\infty,k}\mathbb{P}\{0 < X_0 - u_n \leqslant \delta_{\infty,k}\}\right), \tag{3.16}$$

where $p, q \in [1,\infty]$ such that $p^{-1} + q^{-1} = 1$.

Remark 3. If there exists positive real constants C, λ, α such that

$$\sup_{x\in\mathbb{R}} \mathbb{P}(x \leqslant X_0 \leqslant x + \lambda) \leqslant C\lambda^\alpha, \tag{3.17}$$

then

$$\tilde{\Delta}_n(k) \leqslant \frac{d}{a_n}\left(\delta_{p,k} v_n^{1/q} + 2C\delta_{\infty,k}^{1+\alpha}\right) \leqslant \frac{d}{a_n}\delta_{\infty,k}\left(v_n + 2C\delta_{\infty,k}^{\alpha}\right). \tag{3.18}$$

Application 1 (Causal linear processes)**.** *Let $D = \mathbb{R}$ and*

$$X_i = \sum_{j=0}^{\infty} b_j \xi_{i-j}.$$

Here we set $\delta'_{p,i} = 2\|\xi_0\|_p \sum_{j\geqslant i}|b_j| \geqslant \delta_{p,i}$ in case $\|\xi_0\|_p < \infty$.
The model (1.6) writes with $b_j = b^{-j-1}$ for some integer $b \geqslant 2$ and ξ_0 uniformly distributed on $\{0, \ldots, b-1\}$; in this case X_0 is uniformly distributed over $[0,1]$ and $\delta_{\infty,i} \leqslant b^{-i}$.

Example 3.2 (Markov models)**.** Let $G : (\mathbb{R}^l, \mathcal{B}(\mathbb{R}^l)) \times (D, \mathcal{D}) \longrightarrow (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ be a measurable function and let $(X_i)_{i\geqslant 1-l}$ be a sequence of random variables with values in $\mathbb{R}$ such that

$$X_i = G(X_{i-1}, X_{i-2}, \ldots, X_{i-l}; \xi_i), \qquad \forall i \geqslant 1, \tag{3.19}$$

for some sequence $(\xi_i)_{i\in\mathbb{N}}$ of i.i.d.r.v's. with values in a measurable space D and independent of $(X_0,\dots,X_{1-l})$. Then the random variables $Y_i=(X_i,X_{i-1},\dots,X_{i-l+1})$ defines a Markov chain such that $Y_i=F(Y_{i-1};\xi_i)$ with

$$F(x_l,\dots,x_1;\xi):=(G(x_l,\dots,x_1;\xi),x_l,x_{l-1},\dots,x_2). \tag{3.20}$$

Assume that $Y_0=(X_0,\dots,X_{1-l})$ is a stationary solution of (3.19). Let $Y_0'=(X_0',\dots,X_{1-l}')$ be independent of $(Y_0,(\xi_i)_{i\in\mathbb{N}})$ and distributed as Y_0. Then setting

$$X_i'=G(X_{i-1}',\dots,X_{i-l}';\xi_i), \tag{3.21}$$

X_i' is distributed as X_i and it is independent of $\mathcal{M}_0=\sigma(X_0,\dots,X_{1-l})$, for all $i\in\mathbb{N}$. As for the latter example, let $(\Delta_{p,n}(i))_{i\geqslant 0}$ be a non increasing sequence such that (3.14) holds, where $X_{n,i}$ and $X_{n,i}'$ are the normalizations of X_i and X_i' defined in (3.19) and (3.21), respectively. Hence one can apply the result of Lemma 3 in [Dedecker & Prieur, 2004a], and we obtain that $\tau_{1,n}(k)\leqslant\Delta_{1,n}(k)$.

Analogously to the latter example if $(\delta_{p,i})_{i\geqslant 0}$ is a decreasing sequence such that $\delta_{p,i}\geqslant\|X_i-X_i'\|_p$ and if we consider the normalization (1.4) then $\tau_{1,n}(k)\leqslant\tilde{\Delta}_n(k)$ with $\tilde{\Delta}_n$ defined in (3.16).

In particular if G is such that

$$\|G(x;\xi_1)-G(y;\xi_1)\|_p\leqslant\sum_{i=1}^{l}a_i|x_i-y_i|,\quad\text{with }\sum_{i=1}^{l}a_i<1, \tag{3.22}$$

then $\delta_{p,i}\leqslant Ca^i$ for some $a\in[0,1)$ and $C>0$. (see [Dedecker et al., 2007], page 34).

Application 2 (Contracting Markov chain)**.** *Let $X_i=G(X_{i-1},\xi_i)$ be a Markov chain such that $G:(\mathbb{R},\mathcal{B}(\mathbb{R}))\times(D,\mathcal{D})\longrightarrow(\mathbb{R},\mathcal{B}(\mathbb{R}))$ is a measurable function and*

$$A=\|G(0;\xi_1)\|_p<\infty\ \text{and}\ \|G(x;\xi_1)-G(y;\xi_1)\|_p\leqslant a|x-y|, \tag{3.23}$$

for some $a\in(0,1)$ and $p\in[1,\infty]$. Then, $(X_i)_{i\in\mathbb{N}}$ has a stationary solution with p-th order finite moment as this is proved on page 35 of [Dedecker et al., 2007]*. Moreover under this condition: $\delta_{p,i}=\|X_0'-X_0\|_p\cdot a^i$.*

Remark 4. Note that in particular if $G(u;\xi)=A(u)+B(u)\xi$ for suitable Lipschitz functions $A(u)$ and $B(u)$ with $u\in\mathbb{R}$ then the corresponding iterative model (ARCH-type process) $X_i=G(X_{i-1};\xi_i)$ satisfies (3.23) with $a=\mathrm{Lip}(A)+\|\xi_1\|_p\mathrm{Lip}(B)<1$.

Remark 5. Note that the stationary iterative models $X_i=G(X_{i-1},\xi_i)$ are causal Bernoulli shifts if the condition (3.23) holds; this is proved in Proposition 3.2 in [Dedecker et al., 2007].

Application 3 (Nonlinear AR(l)-models)**.** *Let $l\geq 1$ and $(X_i)_i$ be the stationary solution of some equation*

$$X_i=R(X_{i-1},\dots,X_{i-l})+\xi_i$$

for some measurable function $R : \mathbb{R}^l \to \mathbb{R}$. The process $(X_i)_i$ is then called a stationary real nonlinear autoregressive model of order l. If $\|\xi_1\|_p < \infty$ and

$$|R(u_1,\dots,u_l) - R(v_1,\dots,v_l)| \leqslant \sum_{i=1}^{l} a_i |u_i - v_i|, \text{ for } a_1,\dots,a_l \geqslant 0 \text{ with } \sum_{i=1}^{l} a_i < 1,$$

and for all $(u_1,\dots,u_l),(v_1,\dots,v_l) \in \mathbb{R}^l$ then the function $G : \mathbb{R}^{l+1} \to \mathbb{R}$ by $G(u;\xi) = R(u) + \xi$ satisfies Condition (3.22) and therefore the sequence $(\delta_{p,i})_i$ admits an exponential decay rate.

4. Assumptions

This section addresses the assumptions useful to derive a FCLT for the Empirical cluster process.

4.1. **Fidi convergence conditions.** The technique used to prove the convergence of the empirical process (2.10) is the Bernstein blocks technique. To do this, we need to extract of each block $Y_{n,j}$ of length r_n a sub-block of length l_n such that $l_n = o(r_n)$, and combine this with suitable hypothesis on $\tau-$dependence.
Thus precisely $(X_{n,i})_{1 \leqslant i \leqslant n, n \in \mathbb{N}}$ is row-wise stationary such that

(B.1) $l_n \ll r_n \ll v_n^{-1} \ll n$ with $l_n \longrightarrow \infty$ and $nv_n \longrightarrow \infty$ as $n \to \infty$.
(B.2) $\tau_{1,n}(l_n) = o(r_n^{-1})$ and $\mathbb{E}(\|X_{n,1}\| \, | X_{n,1} \neq 0) < \infty$.

It is necessary consider the following notations used all over this paper.

Notation 1. Let $Y = (X_1, X_2, \dots, X_r)$. We will use the notation $Y^{(l:k)}$ as follows

$$Y^{(l:k)} = \begin{cases} 0, & \text{if } r < l, \\ (X_l,\dots,X_k), & \text{if } 1 \leqslant l \leqslant k \leqslant r, \\ Y, & \text{if } k > r. \end{cases}$$

and $Y^{(k)} := Y^{(1:k)}$. Moreover, if $f \in \mathcal{F}$ is a cluster functional, then we denote

$$\Delta_n(f) := f(Y_n) - f(Y_n^{(r_n - l_n)}), \tag{4.24}$$

where r_n is the length of the block Y_n such that $l_n \ll r_n$.

To prove CLT's for the fidis of the cluster functionals empirical process $(Z_n(f))_{f \in \mathcal{F}}$ as (2.10) with $r_n \ll v_n^{-1} \ll n$, we should take in account the following essential convergence assumptions:

(C.1) For all $f \in \mathcal{F}$,

$$\begin{aligned} \mathbb{E}|\Delta_n(f) - \mathbb{E}\Delta_n(f)|^2 \mathbb{1}\left\{|\Delta_n(f) - \mathbb{E}\Delta_n(f)| \leqslant \sqrt{nv_n}\right\} &= o(r_n v_n) \\ \mathbb{P}\left\{|\Delta_n(f) - \mathbb{E}\Delta_n(f)| > \sqrt{nv_n}\right\} &= o(r_n / v_n), \end{aligned}$$

(C.2) $\mathbb{E}\left[(f(Y_n) - \mathbb{E}f(Y_n))^2 \mathbb{1}\left\{|f(Y_n) - \mathbb{E}f(Y_n)| > \epsilon\sqrt{nv_n}\right\}\right] = o(r_n v_n)$, for all $\epsilon > 0$, and for all $f \in \mathcal{F}$.

(C.3) $(r_n v_n)^{-1}\mathrm{Cov}\left(f(Y_n), g(Y_n)\right) \longrightarrow c(f,g)$, for all $f, g \in \mathcal{F}$.

Remark 6. Note that Assumptions (C.1) and (C.2) are difficult to check in general, for that reason, consider the following (more restrictive but easier to verify) alternatives conditios:

(A.1) $\mathrm{Var}(\Delta_n(f)) = o(r_n v_n)$ for all $f \in \mathcal{F}$.

(A.2) $\mathbb{E}(f(Y_n))^{2+\delta} = O(r_n v_n)$ for some $\delta > 0$ and for all $f \in \mathcal{F}$.

Lemma 1. *The conditions (A.1) and (A.2) implies the conditions (C.1) and (C.2), respectively.*

4.2. **Tools for uniform convergence.** To prove uniform convergence, we use either asymptotic tightness of Z_n in the space $\ell^\infty(\mathcal{F})$, or asymptotic equicontinuity conditions, by some results in § 2.11 of [Van Der Vaart & Wellner, 1996]. Those results need independence therefore a argument of coupling for the blocks $(Y_{n,j})_{1\leqslant j\leqslant m_n, n\in\mathbb{N}}$ is used. [Dedecker & Prieur, 2004a] and Chapter 5 from [Dedecker et al., 2007] yield a suitable coupling argument under τ-weak dependence.

4.2.1. *Asymptotic tightness.*

Definition 4.1. *The sequence $(Z_n)_{n\in\mathbb{N}}$ is asymptotically tight if for every $\epsilon > 0$ there exists a compact set $K \in \ell^\infty(\mathcal{F})$ such that*

$$\limsup_{n\longrightarrow\infty} \mathbb{P}^*(Z_n \notin K^\delta) < \epsilon, \qquad \text{for every } \delta > 0,$$

where $K^\delta = \{f \in \ell^\infty(\mathcal{F}) : d_{\mathcal{F}}(f,K) < \delta\}$ is the "$\delta-$enlargement" around K and $\mathbb{P}^*$ denotes the outer probability.

Definition 4.2. *The bracketing number $N_{[\cdot]}(\epsilon, \mathcal{F}, L_2^n)$ is defined as the smallest number N_ϵ such that for each $n \in \mathbb{N}$ there exits a partition $(\mathcal{F}_{n,k}^\epsilon)_{1\leqslant k\leqslant N_\epsilon}$ of $\mathcal{F}$ such that*

$$\mathbb{E}^* \sup_{f,g\in\mathcal{F}_{n,k}^\epsilon} (f(Y_n) - g(Y_n))^2 \leqslant \epsilon^2 r_n v_n, \qquad \text{for} \quad 1 \leqslant k \leqslant N_\epsilon,$$

where $\mathbb{E}^*$ denotes the outer expectation.

In order to use Theorem 2.11.9 in [Van Der Vaart & Wellner, 1996] we need:

(T.1) The set $\mathcal{F}$ of cluster functionals is such that for each $f \in \mathcal{F}$ the expression $\mathbb{E}f^2(Y_n)$ is finite for all $n \in \mathbb{N}$ and such that the envelope function satisfies:

$$F(x) := \sup_{f\in\mathcal{F}} |f(x)| < \infty, \qquad \forall x \in E_\cup.$$

(T.2) $\mathbb{E}^*\left(F(Y_n)\mathbb{1}\{F(Y_n) > \epsilon\sqrt{nv_n}\}\right) = o(r_n\sqrt{v_n/n})$ for all $\epsilon > 0$.

Note that for a sequence of monotonically increasing positive functions $(h_n(\delta))_{n\geqslant 1}$ the convergence of $h_n(\delta_n)$ to zero $\forall \delta_n \downarrow 0$ is equivalent to

$$\lim_{\delta\downarrow 0} \limsup_{n\longrightarrow\infty} h_n(\delta) = 0,$$

thus the Assumptions 2 and 3 of Theorem 2.11.9 from [Van Der Vaart & Wellner, 1996] are reformulated as follows:

(T.3) There exists a semi-metric ρ on $\mathcal{F}$ such that $\mathcal{F}$ is totally bounded with respect to (w.r.t.) ρ and

$$\lim_{\delta\downarrow 0} \limsup_{n\longrightarrow\infty} \sup_{f,g\in\rho(f,g)<\delta} \frac{1}{r_n v_n} \mathbb{E}\left(f(Y_n) - g(Y_n)\right)^2 = 0.$$

(T.4)

$$\lim_{\delta\downarrow 0} \limsup_{n\longrightarrow\infty} \int_0^\delta \sqrt{\log N_{[\cdot]}(\epsilon, \mathcal{F}, L_2^n)}d\epsilon = 0.$$

4.2.2. *Asymptotic equicontinuity.*

Definition 4.3. *The sequence $(Z_n)_{n\in\mathbb{N}}$ is asymptotically equicontinuous with respect to a semi-metric ρ if for any $\epsilon > 0$ and $\eta > 0$ there exists some $\delta > 0$ such that:*

$$\limsup_{n\longrightarrow\infty} \mathbb{P}^* \left(\sup_{f,g\in\mathcal{F}:\rho(f,g)<\delta} |Z_n(f) - Z_n(g)| > \epsilon \right) < \eta.$$

We use Theorem 2.11.1 in [Van Der Vaart & Wellner, 1996] to prove asymptotic equicontinuity. For this, we define the semi-metric ρ_n on $\mathcal{F}$ as follows. Let $(Y^*_{n,j})_{1\leqslant j\leqslant m_n}$ be the independent blocks coupled to the original blocks $(Y_{n,j})_{1\leqslant j\leqslant m_n}$.
We define ρ_n as:

$$\rho_n(f,g) := \sqrt{\frac{1}{nv_n} \sum_{j=1}^{m_n} (f(Y^*_{n,j}) - g(Y^*_{n,j}))^2}. \tag{4.25}$$

So we denote by $N(\epsilon, \mathcal{F}, \rho)$ the "covering number", the minimum number of balls (with respect to the semi-metric ρ) with radius $\epsilon > 0$ necessary to cover $\mathcal{F}$. We need the following assumptions:

(T.4') For $k = 1, 2$ the map

$$(Y^*_{n,1}, \ldots, Y^*_{n,[m_n/2]}) \longmapsto \sup_{f,g\in\mathcal{F}:\rho(f,g)<\delta} \sum_{j=1}^{[m_n/2]} e_j \left(f(Y^*_{n,j}) - g(Y^*_{n,j}) \right)^k$$

is measurable for each $\delta > 0$ each vector $(e_1, \ldots, e_{[m_n/2]}) \in \{-1, 0, 1\}^{[m_n/2]}$ and each $n \in \mathbb{N}$.

(T.5)

$$\lim_{\delta\downarrow 0}\limsup_{n\longrightarrow\infty}\mathbb{P}^*\left(\int_0^\delta \sqrt{\log N(\epsilon,\mathcal{F},\rho_n)}d\epsilon>\xi\right)=0,\qquad \forall\xi>0.$$

5. Results

Theorem 1. *Suppose that (B.1), (B.2), (C.1), (C.2) and (C.3) holds. Then the fidis of the cluster functionals empirical process $(Z_n(f))_{f\in\mathcal{F}}$ converge to the fidis of a centered Gaussian process $(Z(f))_{f\in\mathcal{F}}$ with covariance function c defined in assumption (C.3).*

The following lemma is a version of the condition (2) in [Segers, 2003], useful to get an alternative expression of $\mathbb{E}[f(Y_n)|Y_n\neq 0]$ (Proposition 1 below). Moreover, with this lemma we derive immediately weak convergence of $\mathbb{P}_{f(Y_n)|Y_n\neq 0}$ to a probability measure $\mathbb{P}_{f,W}$ in Proposition 2, applying the tools of Segers again and part of the proof of Lemma 2.5 in [Drees & Rootzén, 2010]. Finally, using this propositions we give an alternative expression to the function c defined in the assumption (C.3). Such expression is (5.31) - Corollary 1.

Lemma 2. *If there exists $p,q,r\geqslant 1$ with $p^{-1}+q^{-1}=1$, such that*

$$\lim_{l\to\infty}\limsup_{n\to\infty}\frac{\tau_{p,n}(l)}{r_n v_n^{p^{-1}+r^{-1}}}=0 \tag{5.26}$$

and $r_n v_n^{1/q}\longrightarrow 0$ as $n\to\infty$, then

$$\lim_{l\to\infty}\limsup_{n\to\infty}\mathbb{P}\{Y_{n,1}^{(l+1,r_n)}\neq 0|X_{n,1}\neq 0\}=0. \tag{5.27}$$

Remark 7. Note that condition (5.26) is the condition which replaces the condition (B3) of [Drees & Rootzén, 2010].

Proposition 1. *Under assumption (B.1), suppose that there exists $p,q,r\geqslant 1$ (with $p^{-1}+q^{-1}=1$) such that (5.26) hold. Then*

$$\mathbb{E}\left[f(Y_n)|Y_n\neq 0\right]=\frac{1}{\theta_n}\mathbb{E}\left[f(Y_{n,1})-f(Y_{n,1}^{(2:r_n)})|X_{n,1}\neq 0\right]+o(1), \tag{5.28}$$

where $o(1)$ converges to 0 as $n\to\infty$ uniformly for all bounded cluster functionals $f\in\mathcal{F}$, and

$$\theta_n:=\frac{\mathbb{P}\{Y_n\neq 0\}}{r_n v_n}=\mathbb{P}\{Y_{n,1}^{(2:r_n)}=0|X_{n,1}\neq 0\}(1+o(1)). \tag{5.29}$$

Consider the following alternative condition:

(A.3) There is a sequence $W = (W_i)_{i\geqslant 1}$ of E-valued random variables such that, for all $k \in \mathbb{N}$, the joint conditional distribution

$$P_{(X_{n,i},\mathbb{1}\{X_{n,i}=0\})_{1\leqslant i\leqslant k}|X_{n,1}\neq 0}$$

converges weakly to $P_{(W_i,\mathbb{1}\{W_i=0\})}$, and for all $f \in \mathcal{F}$ are a.s. continuous with respect to the distribution of $W^{(k)} = (W_1, \dots, W_k)$ and $W^{(2:k)} = (W_2, \dots, W_k)$ for all k, that is,

$$\mathbb{P}\{W^{(2:k)} \in D_{f,k-1}, W_i = 0,\ \forall i > k\} = \mathbb{P}\{W^{(k)} \in D_{f,k}, W_i = 0,\ \forall i > k\} = 0$$

where we denote by $D_{f,k}$ the set of discontinuities of $f|_{E^k}$.

Remark 8. The existence of such sequence W is guaranteed in particular from Theorem 2 in [Segers, 2003] with $E = \mathbb{R}$ and the normalization (1.1). There, Segers has shown that if

$$\mathbb{P}_{((X_{n,i})_{1\leqslant i\leqslant k}|X_1>u_n)} \underset{n\to\infty}{\longrightarrow} -\log G_k,$$

where G_k is some $k-$dimensional extreme value distribution for all $k \in \mathbb{N}$, then there exists such sequence "*tail chain*" $W = (W_i)_{i\in\mathbb{N}}$ such that

$$P_{((X_{n,i},\mathbb{1}\{X_{n,i}=0\})_{1\leqslant i\leqslant k}|X_1>u_n)} \underset{n\to\infty}{\overset{w}{\longrightarrow}} P_{(W_i,\mathbb{1}\{W_i=0\})_{1\leqslant i\leqslant k}}, \tag{5.30}$$

for all $k \in \mathbb{N}$.

Proposition 2. *Suppose that (B.1), (A.3) are satisfied and that there exist* $p, q, r \geqslant 1$ *(with* $p^{-1} + q^{-1} = 1$*) such that (5.26) hold.*
Then

$$\begin{array}{ccl} m_W & := & \sup\{i \geqslant 1 : W_i \neq 0\} < \infty, \\ \theta_n & \underset{n\to\infty}{\longrightarrow} & \theta := \mathbb{P}\{W_i = 0, \forall i \geqslant 2\} = \mathbb{P}\{m_W = 1\} > 0, \\ P_{f(Y_n)|Y_n\neq 0} & \underset{n\to\infty}{\overset{w}{\longrightarrow}} & \dfrac{1}{\theta}\left(\mathbb{P}\{f(W) \in \cdot\} - \mathbb{P}\{f(W^{(2:\infty)}) \in \cdot, m_W \geqslant 2\}\right). \end{array}$$

Corollary 1. *Suppose that*

$$\mathcal{F} = \{f|(f(Y_n)^2)_{n\in\mathbb{N}} \text{ is uniformly integrable under } P(\cdot)/r_n v_n\}.$$

Assume that (B.2) and Proposition 2's hypothesis hold. If moreover the assumptions (C.1) and (C.3) are satisfied then the fidis of the cluster functionals empirical process $(Z_n(f))_{f\in\mathcal{F}}$ *converge to the fidis of a centered Gaussian process* $(Z(f))_{f\in\mathcal{F}}$ *with covariance function c defined by*

$$c(f, g) = \mathbb{E}\left[(fg)(W) - (fg)(W^{(2:\infty)})\right]. \tag{5.31}$$

There are many cases in which $\|f\|_\infty = \sup_{x\in E_\cup} |f(x)| < \infty$, for all $f \in \mathcal{F}$. Under this condition, it is clear that the conditions (C.1) and C.2) are satisfied. Therefore, it is important to note the following corollary.

Corollary 2. *Suppose that (B.2) and Proposition 2's hypothesis are satisfied. Then, if* $\|f\|_\infty = \sup_{x\in E_\cup} |f(x)| < \infty$ *for all* $f \in \mathcal{F}$*, the fidis of the cluster functionals empirical process* $(Z_n(f))_{f\in\mathcal{F}}$ *converges to the fidis of a centered Gaussian process* $(Z(f))_{f\in\mathcal{F}}$ *with covariance function* c *defined by (5.31).*

Theorem 2. *Suppose that (B.1) and (B.2) hold and that (T.1)-(T.4) are satisfied. Then the empirical process* $(Z_n)_{n\in\mathbb{N}}$ *is asymptotically tight in* $\ell^\infty(\mathcal{F})$*. Moreover, if the assumptions (C.1)-(C.3) hold, then* Z_n *converges to a centered Gaussian process* Z *with covariance function* c *in (C.3).*

Theorem 3. *Suppose that (B.1) and (B.2) hold and that (T.1), (T.2), (T.3), (T.4') and (T.5) are satisfied. Then the empirical process* $(Z_n)_{n\in\mathbb{N}}$ *is asymptotically equicontinuous. Moreover if the assumptions (C.1)-(C.3) hold, then* Z_n *converges to a centered Gaussian process* Z *with covariance function* c *in (C.3).*

Application 4 (Blocks estimator of the extremal index)**.** Let $(X_i)_{i\in\mathbb{N}}$ be a real stationary time series with distribution function F. Now consider the index defined in (5.29) with the extreme normalization (1.1) and $u_n := F^{\leftarrow}(1-v_n t)$, for all $t \in [0,1]$, *i.e.*

$$\theta_{n,t} := \frac{\mathbb{P}\{Y_n \neq 0\}}{r_n v_n t} = \frac{\mathbb{P}\{\max_{1\leqslant i\leqslant r_n} X_i > u_n\}}{r_n v_n t}, \quad \text{with } t \in [0,1]. \tag{5.32}$$

From Proposition 2, if (B.1) is satisfied and if there exist $p, q, r \geqslant 1$ (with $p^{-1}+q^{-1} = 1$)) such that (5.26) hold, then there is a number (*extremal index*) $\theta \in (0,1]$ such that

$$\theta_{n,t} \underset{n\to\infty}{\longrightarrow} \theta \quad \text{uniformly for } t \in [0,1]. \tag{5.33}$$

Given the convergence (5.33), Drees has suggested in his paper [Drees, 2011] to estimate θ replacing the unknown probability $\mathbb{P}\{\max_{1\leqslant i\leqslant r_n} X_i > u_n\}$ and the unknown expectation $r_n v_n t = \mathbb{E}\left[\sum_{i=1}^{r_n} \mathbb{1}\{X_i > u_n\}\right]$ by a empirical expression for $\theta_{n,t}$:

$$\widehat{\theta}_{n,t} := \frac{\sum_{j=1}^{m_n} \mathbb{1}\{\max_{(j-1)r_n < i \leqslant jr_n} X_i > u_n\}}{\sum_{j=1}^{m_n}\sum_{i=(j-1)r_n+1}^{jr_n} \mathbb{1}\{X_i > u_n\}}, \tag{5.34}$$

where $m_n = [n/r_n]$ such that $1 \ll r_n \ll v_n^{-1} \ll n$ but $nv_n \longrightarrow \infty$. Thus, such estimator (5.34) (called *blocks estimator of the extremal index*) can be expressed in terms of two empirical processes of cluster functionals $(Z_n(f_t), Z_n(g_t))_{0\leqslant t\leqslant 1}$ defined in (2.10). For this, suppose without loss of generality that the random variables $(X_i)_{1\leqslant i\leqslant n}$ are uniformly distributed on $[0,1]$ (otherwise, just consider the transformation $U_i = F(X_i)$, $1 \leqslant i \leqslant n$, where F is the distribution function of X_1, see [Drees, 2011]). Then, with the normalization (1.1) such that $a_n = v_n = 1 - u_n$ and the blocks $(Y_{n,j})_{1\leqslant j\leqslant m_n}$ defined in (2.8), we have that

$$\widehat{\theta}_{n,t} = \frac{m_n^{-1}\sum_{j=1}^{m_n} f_t(Y_{n,j})}{m_n^{-1}\sum_{j=1}^{m_n} g_t(Y_{n,j})} = \frac{\mathbb{E}f_t(Y_{n,1}) + (nv_n)^{1/2}m_n^{-1}Z_n(f_t)}{\mathbb{E}g_t(Y_{n,1}) + (nv_n)^{1/2}m_n^{-1}Z_n(g_t)}, \tag{5.35}$$

where

$$f_t(x_1,\ldots,x_r) := \mathbb{1}\{\max_{1\leqslant i\leqslant r} x_i > 1-t\} \tag{5.36}$$

$$g_t(x_1,\ldots,x_r) := \sum_{i=1}^{r} \mathbb{1}\{x_i > 1-t\}. \tag{5.37}$$

For this particular case, we consider the following assumptions:

(C.3.1) $(r_nv_n)^{-1}\mathrm{Cov}(g_s(Y_n), g_t(Y_n)) \longrightarrow c_g(s,t)$, for all $0\leqslant s,t\leqslant 1$.

(C.3.2) $(r_nv_n)^{-1}\mathrm{Cov}(f_s(Y_n), g_t(Y_n)) \longrightarrow c_{fg}(s,t)$, for all $0\leqslant s,t\leqslant 1$.

(C.4) For some bounded function $h:(0,1]\longrightarrow\mathbb{R}$ such that $\lim_{t\to 0} h(t) = 0$

$$(r_nv_n)^{-1}\mathbb{E}\left(f_s(Y_{n,1}) - f_t(Y_{n,1})\right)^2 \leqslant h(t-s), \quad \forall 0\leqslant s<t\leqslant 1,$$

for all n sufficiently large.

The following are a slight variation of the first two results of [Drees, 2011], in the sense that we replace the β-mixing condition with τ-dependence condition. The rest of the results of such paper also are true if we change the assumptions (C1) and (C2) for our assumptions (B.1) and (B.2). However, those results will not develop here because it is not the aim of this work.

Proposition 3.

(3.1) *Suppose that $\tau_{1,n}(l_n) = o(r_n^{-1})$ where l_n, r_n are such that (B.1) is satisfied. Then $(Z_n(f_t))_{0\leqslant t\leqslant 1}$ converges weakly to $Z_f := (\sqrt{\theta}B_t)_{0\leqslant t\leqslant 1}$, where B denote a standard Brownian motion.*

(3.2) *If additionally (C.3.1) and (C.4) are satisfied and $r_n = o(\sqrt{nv_n})$, then the sequence of processes $(Z_n(g_t))_{0\leqslant t\leqslant 1}$ converges weakly to a centered Gaussian process $(Z_g(t))_{0\leqslant t\leqslant 1}$ with covariance function c_g.*

(3.3) *Under all the hypothesis of (3.1) and (3.2), if moreover (C.3.2) holds, then $(Z_n(f_t), Z_n(g_t))_{0\leqslant t\leqslant 1}$ converge weakly to $(Z_f(t), Z_g(t))_{0\leqslant t\leqslant 1}$ with*

$$\begin{aligned} Cov(Z_f(s), Z_f(t)) &= \theta(s\wedge t),\\ Cov(Z_g(s), Z_g(t)) &= c_g(s,t),\\ Cov(Z_f(s), Z_g(t)) &= c_{fg}(s,t), \quad 0\leqslant s,t\leqslant 1. \end{aligned} \tag{5.38}$$

Using the same argument in Remark 8, we can find explicit expressions for the covariance functions c_g and c_{fg} as functions of the "tail chains" of $(X_i)_{i\in\mathbb{N}}$. This is, if for every $k \in \mathbb{N}$ the distribution function of $(X_1, \dots, , X_k)$ belongs to the domain of attraction of an extreme-value distribution, then there exist a sequence $W = (W_i)_{i\in\mathbb{N}}$ such that (5.30) hold. In such case:

$$c_g(s,t) = s \wedge t + \sum_{i=1}^{\infty} (\mathbb{P}\{W_1 > 1-s, W_{i+1} > 1-t\} + \mathbb{P}\{W_1 > 1-t, W_{i+1} > 1-s\})$$

$$c_{fg}(s,t) = \begin{cases} \mathbb{P}\{W_1 > 1-t, \max_{j\geqslant 1} W_j > 1-s\} & \\ +\sum_{i=1}^{\infty} \mathbb{P}\{W_1 > 1-s, W_{i+1} > 1-t, \max_{j\geqslant 2} W_j \leqslant 1-s\}, & s < t, \\ t & s \geqslant t. \end{cases}$$

Corollary 3. *Under Proposition 3 - (3.3)'s assumptions,*

$$(\sqrt{nv_n}t(\widehat{\theta}_{n,t} - \theta_{n,t}))_{0<t\leqslant 1} \underset{n\to\infty}{\xrightarrow{w}} Z := Z_f - \theta Z_g, \tag{5.39}$$

where Z is a Gaussian process such that $\mathbb{E}Z(t) = 0$ and

$$Cov(Z(s), Z(t)) = \theta(s \wedge t - c_{fg}(s,t) - c_{fg}(t,s) + \theta^2 c_g(s,t)). \tag{5.40}$$

6. Examples and Simulations

6.1. **AR(1)-process with the functional "number of excesses over x".** We consider the AR(1)-process (1.6) where $b \geqslant 2$ is an integer, $(\xi_i)_{i\in\mathbb{N}}$ are i.i.d. and uniformly distributed on the set $U(b) := \{0, 1, \dots, b-1\}$.
It is clear that X_0 is uniformly distributed on $[0,1]$. Moreover we define the normalized random variables $(X_{n,i})_{1\leqslant i\leqslant n, n\in\mathbb{N}}$ as in eqn. (1.4) with $a_n = v_n = 1 - u_n$. We set $(x_1, \dots, x_d) \leqslant (y_1, \dots, y_d)$ if and only if $x_i \leqslant y_i$, for all $i = 1, \dots, d$ in case $x, y \in [0,1]^d$. Then

$$\begin{aligned} &\mathbb{P}\{X_{n,1} > x | X_{n,1} \neq 0\} \\ &= \frac{1}{b^d \bar{v}_n} \sum_{j_1,\dots,j_d \in U(b)} \left(\max_{i=1,\dots,d} \left\{ 1 - b^i + \sum_{s=1}^{i} b^{s-1} j_s + b^i v_n (1-x_i) \right\}_+ \wedge 1 \right) \\ &\underset{n\to\infty}{\longrightarrow} \max_{i=1,\dots,d} \{b^{i-d}(1-x_i)\}, \end{aligned} \tag{6.41}$$

where $\bar{v}_n := \mathbb{P}\{X_{n,1} \neq 0\} \sim \ v_n = \mathbb{P}\{X_1 > u_n\} \longrightarrow 0$.

Consider $\mathcal{F}$ the family of cluster functionals f as in Example (2.1.2) *i.e.*

$$\mathcal{F} = \left\{ f_x,\ x \in [0,1]^d \right\}, \quad \text{with} \quad f_x(x_1, \dots, x_r) = \sum_{i=1}^{r} \mathbb{1}\{x_i > x\} \tag{6.42}$$

For this case, we obtain the covariance function c of (C.3):

$$c(x,y) = \min\left(\max_{k=1,\dots,d}\{b^k(1-x_k)\}, \max_{k=1,\dots,d}\{b^k(1-y_k)\}\right) + \sum_{i=1}^{\infty} H_{b,i}(x,y) + \sum_{i=1}^{\infty} H_{b,i}(y,x), \tag{6.43}$$

where, for $i \geqslant d$

$$H_{b,i}(x,y) := \frac{1}{b^i}\min\left(\max_{k=1\dots,d}\{b^k(1-x_k)\}, \max_{k=1,\dots,d}\{b^{k+i}(1-y_k)\}\right) \tag{6.44}$$

and for $1 \leqslant i < d$,

$$H_{b,i}(x,y) := \frac{1}{b^i}\min\left(\max_{k=1\dots,i}\{b^k(1-x_k)\}, \max_{k=i+1,\dots,d}\{b^k \min(1-x_k, 1-y_k)\}, \max_{k=d-i,\dots,d}\{b^{k+i}(1-y_k)\}\right)$$

Conditions (C.1), (C.2), (T.1) - (T.4) hold for uniformly distributed random variables and for the same family $\mathcal{F}$, see page 2177 and 2178 in [Drees & Rootzén, 2010].

Thus, under assumption (B.1), setting r_n such that $b^{-2r_n} = o(v_n)$ (or $b^{-r_n} = O(v_n)$ (see eqn. (3.18) and Application 1), then the empirical process $(Z_n(x))_{x\in[0,1]^d}$ defined as:

$$Z_n(x) := \frac{1}{\sqrt{nv_n}}\sum_{i=1}^{r_n m_n}\left(\mathbb{1}\{X_{n,i} > x\} - \mathbb{P}\{X_{n,i} > x\}\right) \sim \frac{1}{\sqrt{nv_n}}\sum_{i=1}^{n}\left(\mathbb{1}\{X_{n,i} > x\} - \mathbb{P}\{X_{n,i} > x\}\right) \tag{6.45}$$

converges to a centered Gaussian process Z with covariance function (6.43).

6.2. **Simulation study.** The experiment is to estimate the extremal index θ through the blocks estimator of the extremal index (5.35).

Let us consider the AR(1)-process (1.6). Here, as X_0 is uniformly distributed on $[0,1]$ and $X_i = \frac{X_0}{b^i} + \sum_{s=1}^{i}\frac{\xi_s}{b^{i-s+1}}$ for all $i \geqslant 1$, we obtain a theoretic expression for the index (5.32) with $u_n = 1 - v_n t$ for $t \in (0,1]$:

$$\theta_{n,t} = \frac{1}{b^{r_n} r_n v_n t}\sum_{j_1,\dots,j_{r_n}\in U(b)} \min\left(\max_{i=1,\dots,r_n}\left\{1 - b^i(1-v_n t) + \sum_{s=1}^{i} b^{s-1} j_s\right\}_+, 1\right), \tag{6.46}$$

which converges to some $\theta = \theta(b) \in (0,1)$ if (B.1) is satisfied with $b^{-2l_n} = o(r_n v_n^{\beta})$, for some $\beta \in [1,3]$.

Now, we simulate $AR(1)-$processes (1.6) for $b = 2, 3$ and their blocks estimators (5.35) respective with the normalization (1.1) and $a_n = v_n = 1 - u_n$ to make the comparison of the results estimated with the theoretic model (6.46).

To simulate the blocks estimators, we made $N = 60$ sequences of length $n = 10000$ with $v_n = n^{-1/2}$ and $r_n = [\log(n)]$. In Figure 1 we showed a polygonal curve $(t, \theta_{n=10000,t})_{t=0.1,0.2,\ldots,1}$ (blue curve) of $(t, \theta_{n=10000,t})_{0 \leqslant t \leqslant 1}$ and a mean polygonal curve $(t, \widehat{\theta}_{n=10000,t})_{t=0.1,0.2,\ldots,1}$ (black curve) of $(t, \widehat{\theta}_{n=10000,t})_{t=0 \leqslant t1}$. Moreover, the confidence intervals I_t of $\widehat{\theta}_t$ with a confidence level $1 - \alpha = 0.95$ (red curves).

As expected, the estimated value through the blocks estimator is quite close to the index theoretical (6.46), with $n = 10^4$. The numerical results are shown in Tables 1 and 2, for the cases $b = 2$ and $b = 3$, respectively.

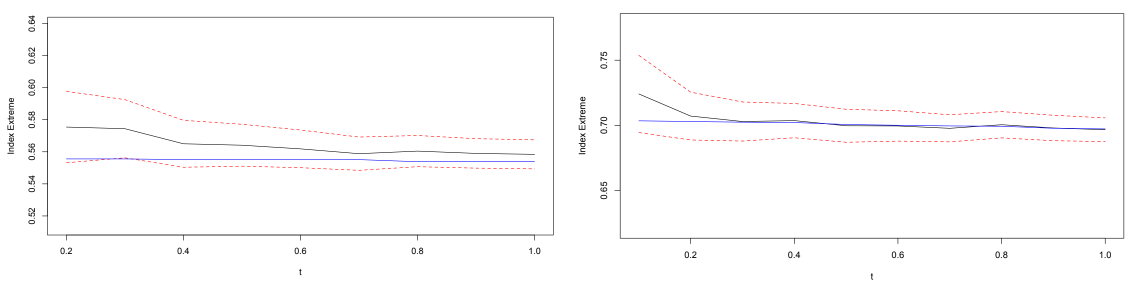

Figure 1. Extremal index for the AR(1)-input
Left: $\theta_{n=10000,t}$ is the blue curve, $\widehat{\theta}_{n=10000,t}$ is the black curve and the confidence intervals $I_t = (CI_i(t), CI_s(t))$ are the red curves, for the AR(1)-input (1.6) with $b = 2$. Right: the same situation but with $b = 3$

t	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$\theta_{n,t}$	0.555	0.555	0.555	0.555	0.555	0.555	0.554	0.554	0.554
$\widehat{\theta}_{n,t}$	0.575	0.574	0.565	0.564	0.562	0.559	0.560	0.559	0.559
CI_s	0.598	0.593	0.580	0.577	0.574	0.570	0.570	0.568	0.567
CI_i	0.553	0.556	0.550	0.551	0.550	0.548	0.551	0.550	0.549

Table 1. Comparison between the blocks estimation and the theoretical approximation (6.46), for the AR(1)-input with $b = 2$ and $n = 10^4$.

Remark 9. *In June 2015 at the Workshop "Mathematical Foundations of Heavy Tailed Analysis" in Copenhagen, Philippe Soulier has detailed calculations of θ for*

t	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
$\theta_{n,t}$	0.703	0.703	0.702	0.702	0.701	0.700	0.700	0.699	0.698	0.697
$\widehat{\theta}_{n,t}$	0.724	0.707	0.703	0.704	0.700	0.700	0.698	0.700	0.698	0.697
CI_s	0.754	0.725	0.718	0.717	0.712	0.711	0.708	0.711	0.708	0.706
CI_i	0.694	0.689	0.688	0.690	0.687	0.688	0.687	0.690	0.688	0.688

TABLE 2. Comparison between the blocks estimation and the theoretical approximation (6.46), for the AR(1)-input with $b = 3$ and $n = 10^4$

AR(1)-processes: $X_i = b^{-1}X_{i-1} + Z_i$ *with* $b > 1$ *under the assumption that the innovations* $(Z_i)_{i\in\mathbb{N}}$ *are i.i.d. regularly varying with index* $\alpha > 1$, $RV(\alpha)$. *In this case,*

$$\theta = \theta(b, \alpha) = 1 - \frac{1}{b^\alpha}. \tag{6.47}$$

However, note that the AR(1)-process (1.6) has innovations $Z_i = \dfrac{\xi_i}{b}$ *regularly varying with index* $\alpha = 1$. *Thus we can not use the approximation (6.47) in this case. Naturally one can consider an approximation* $\alpha \longrightarrow 1^+$, *then* $\theta(2, 1) = 1/2$ *and* $\theta(3, 1) = 2/3$ *for the AR(1)-process (1.6) with* $b = 2$ *and* $b = 3$, *respectively. But those results are not comparable with our results, since for* n *sufficiently large* ($n = 10^6$), *the theoretical value (6.46) is* $\theta_{n,t} \approx \theta(b)$, *where* $\theta(2) = 0.5384$ *and* $\theta(3) = 0.6953$. *Moreover, note that* $\theta(2, 1)$ *and* $\theta(3, 1)$ *are outside the confidence region of the blocks estimator* $\widehat{\theta}_{n=10^4,t}$ *(see Tables 1 and 2 for* $b = 2$ *and* $b = 3$, *respectively).*

7. PROOFS

Proof of Example 3.1. Since the r.v's $(X_i)_{i\geqslant 1}$ and $(X'_i)_{i\geqslant 1}$ have the same distributions the $(X_{n,i})_{1\leqslant i\leqslant n}$ and $(X'_{n,i})_{1\leqslant i\leqslant n}$ also have the same distributions. Moreover the r.v's $(X'_{n,i})_{1\leqslant i\leqslant n}$ are independent on $\mathcal{M}_0$.
From Lemma 3 in [Dedecker & Prieur, 2004a] we get:

$$\begin{aligned} \tau_1(\mathcal{M}_0, (X_{n,j_1}, X_{n,j_2}, \ldots, X_{n,j_l})) &\leqslant \sum_{i=1}^{l} \left\| X_{n,j_i} - X'_{n,j_i} \right\|_1 \\ &\leqslant l\Delta_{1,n}(k), \end{aligned} \tag{7.48}$$

for $k < j_1 < j_2 < \cdots < j_l \leq n$.
The rest of the proof follows from the definition of $\tau_{1,n}$. $\square$

Proof of (3.16): Define

$$X'_{n,i} := \left(\left(\frac{X'_i - u_n}{a_n} \right)_+ , \left(\frac{X'_{i+1} - u_n}{a_n} \right)_+ , \ldots , \left(\frac{X'_{i+d-1} - u_n}{a_n} \right)_+ \right), \tag{7.49}$$

where $X'_i = H(\xi_i, \xi_{i-1}, \ldots, \xi_1, \xi'_0, \xi'_{-1}, \ldots)$. Recall that the r.v's $(X'_{n,i})_{1 \leqslant i \leqslant n, n \in \mathbb{N}}$ are distributed as $(X_{n,i})_{1 \leqslant i \leqslant n, n \in \mathbb{N}}$ and independent of $\mathcal{M}_0$: hence from Lemma 3 in [Dedecker & Prieur, 2004a]), for $k < j_1 < j_2 < \ldots < j_l \leqslant n$ we obtain:

$$\begin{aligned}
\tau_1(\mathcal{M}_0, (X_{n,j_1}, \ldots, X_{n,j_l})) &\leqslant \sum_{i=1}^{l} \left\| X_{n,j_i} - X'_{n,j_i} \right\|_1 \\
&= \frac{1}{a_n} \sum_{i=1}^{l} \left\| \sum_{j=1}^{d} \left((X_{j_i+j-1} - u_n)_+ - (X'_{j_i+j-1} - u_n)_+ \right) e_j \right\|_1 \\
&\leqslant \frac{1}{a_n} \sum_{i=1}^{l} \sum_{j=1}^{d} \|(X_{j_i+j-1} - u_n)_+ - (X'_{j_i+j-1} - u_n)_+\|_1 \\
&\leqslant \frac{1}{a_n} \sum_{i=1}^{l} \sum_{j=1}^{d} \big(\mathbb{E}|X_{j_i+j-1} - X'_{j_i+j-1}| \mathbb{1}\{X_{j_i+j-1} > u_n, X'_{j_i+j-1} > u_n\} \\
&\qquad + \mathbb{E}|X_{j_i+j-1} - u_n| \mathbb{1}\{X_{j_i+j-1} > u_n, X'_{j_i+j-1} \leqslant u_n\} \\
&\qquad\qquad + \mathbb{E}|X'_{j_i+j-1} - u_n| \mathbb{1}\{X_{j_i+j-1} \leqslant u_n, X'_{j_i+j-1} > u_n\} \big).
\end{aligned} \tag{7.50}$$

Note that:

$$\begin{aligned}
&\mathbb{E}|X_{j_i+j-1} - X'_{j_i+j-1}| \mathbb{1}\{X_{j_i+j-1} > u_n, X'_{j_i+j-1} > u_n\} \\
&\quad \leqslant \|X_{j_i+j-1} - X'_{j_i+j-1}\|_p \mathbb{P}^{1/q}\{X_{j_i+j-1} > u_n, X'_{j_i+j-1} > u_n\} \\
&\qquad \leqslant \|X_{j_i+j-1} - X'_{j_i+j-1}\|_p v_n^{1/q} \leqslant \delta_{p,k} v_n^{1/q},
\end{aligned} \tag{7.51}$$

for some $p, q \in [1, \infty]$ such that $p^{-1} + q^{-1} = 1$. On the other hand,

$$\begin{aligned}
&\mathbb{E}|X_{j_i+j-1} - u_n| \mathbb{1}\{X_{j_i+j-1} > u_n, X'_{j_i+j-1} \leqslant u_n\} \\
&\quad \leqslant \mathbb{E}|X_{j_i+j-1} - u_n| \mathbb{1}\{u_n < X_{j_i+j-1} \leqslant u_n + \delta_{\infty, j_i+j-1}\} \\
&\qquad \leqslant \delta_{\infty, j_i+j-1} \mathbb{P}\{0 < X_{j_i+j-1} - u_n \leqslant \delta_{\infty, j_i+j-1}\} \\
&\qquad\qquad \leqslant \delta_{\infty,k} \mathbb{P}\{0 < X_0 - u_n \leqslant \delta_{\infty,k}\}.
\end{aligned} \tag{7.52}$$

Similarly, we have that

$$\begin{aligned}
&\mathbb{E}|X'_{j_i+j-1} - u_n| \mathbb{1}\{X_{j_i+j-1} \leqslant u_n, X'_{j_i+j-1} > u_n\} \\
&\qquad\qquad \leqslant \delta_{\infty,k} \mathbb{P}\{0 < X_0 - u_n \leqslant \delta_{\infty,k}\}
\end{aligned} \tag{7.53}$$

Therefore, (7.50) is bounded by

$$\frac{dl}{a_n}\left(\delta_{p,k} v_n^{1/q} + 2\delta_{\infty,k}\mathbb{P}\{0 < X_0 - u_n \leqslant \delta_{\infty,k}\}\right). \tag{7.54}$$

Finally, applying the definition we have that $\tau_{1,n}(k) \leqslant \tilde{\Delta}_n(k)$. □

Lemma 3 (Coupling)**.** *Suppose that the random variables $(X_{n,i})_{1\leqslant i\leqslant n}$ are such that*

$$\mathbb{E}\left(\|X_{n,1}\| | X_{n,1} \neq 0\right) < \infty.$$

We consider together even and odd block sizes, by using $k = 0$ or 1 according to the parity. Assume that for each $j \in \{1, \ldots, [m_n/2]\}$ there is a random variable $\delta_{k,j}$ uniformly distributed on $[0,1]^{r_n}$ and independent of $\mathcal{M}^k_{n,j-1} = \sigma(Y_{n,2-k}, \ldots, Y_{n,2(j-1)-k})$ and $\sigma(Y_{n,2j-k})$. Then there exists a random block $\dot{Y}_{n,2j-k}$ measurable with respect to $\mathcal{M}^k_{n,j-1} \vee \sigma(Y_{n,2j-k}) \vee \sigma(\delta_{k,j})$, independent of $\mathcal{M}^k_{n,j-1}$, distributed as $Y_{n,2j-k}$, and such that

$$\left\| \mathbb{E}\left(d_{r_n}(Y_{n,2j-k}, \dot{Y}_{n,2j-k})\right) \Big| \mathcal{M}^k_{n,j-1} \right\|_p \leqslant r_n \tau_{p,n}(r_n), \tag{7.55}$$

where $\delta_r((x_1, \ldots, x_r), (y_1, \ldots, y_r)) = \sum_{i=1}^r \|x_i - y_i\|$.
In particular, if $\mathcal{M}^k_{n,j-1} = \sigma(\dot{Y}_{n,2-k}, \ldots, \dot{Y}_{n,2(j-1)-k})$ then the blocks $(\dot{Y}_{n,2j-k})_{1\leqslant j\leqslant[m_n/2]}$ are independent.

Proof: We set here $k = 0$ (for even block sizes) since the steps are similar if $k = 1$. If $j \in \{1, 2, \ldots, [m_n/2]\}$ denote $A_j = \{(2j-1)r_n + 1, \ldots, 2jr_n\}$. Assume that for each $i \in A_j$ there is a random variable δ_i uniformly distributed on $[0,1]$ independent of $\mathcal{M}^0_{n,j-1}$ with $\sigma(Y_{n,2j}) \supset \sigma(X_{n,i})$ (without loss of generality, we can assume that the variables $(\delta_i)_{i\in A_j}$ are independent). From [Dedecker et al., 2007]'s Lemma 5.3, there exists a random variable $\dot{X}_{n,i}$ measurable with respect to $\mathcal{M}^0_{n,j-1} \vee \sigma(X_{n,i}) \vee \sigma(\delta_i)$ independent of $\mathcal{M}^0_{n,j-1}$ distributed as $X_{n,i}$ and such that

$$\tau_p(\mathcal{M}^0_{n,j-1}, X_{n,i}) = \left\|\mathbb{E}(\|X_{n,i} - \dot{X}_{n,i}\| \,|\mathcal{M}^0_{n,j-1})\right\|_p. \tag{7.56}$$

Set $\dot{Y}_{n,2j} = (\dot{X}_{n,(2j-1)r_n+1}, \ldots, \dot{X}_{2jr_n})$ and $\delta_{0,j} = (\delta_{n,(2j-1)r_n+1}, \ldots, \delta_{2jr_n})$ then the random block $\dot{Y}_{n,2j}$ is measurable with respect to $\mathcal{M}^0_{n,j-1} \vee \sigma(Y_{n,2j}) \vee \sigma(\delta_{0,j})$ independent of $\mathcal{M}^0_{n,j-1}$ and distributed as $Y_{n,2j}$; moreover:

$$
\begin{aligned}
(7.57)\quad & \left\|\mathbb{E}(d_{r_n}\left(Y_{n,2j},\dot{Y}_{n,2j}\right)\,|\mathcal{M}^0_{n,j-1})\right\|_p \\
& \leqslant \sum_{i=1}^{r_n}\left\|\mathbb{E}(\|X_{n,(2j-1)r_n+i}-\dot{X}_{n,(2j-1)r_n+i}\|\,|\mathcal{M}^0_{n,j-1})\right\|_p \\
& = \sum_{i=1}^{r_n}\tau_p(\mathcal{M}^0_{nj-1},X_{n,(2j-1)r_n+i}) \\
& \leqslant \sum_{i=1}^{r_n}\tau_p(\mathcal{M}^0_{nj-1},X_{n,(2j-1)r_n+1}) \\
& \leq r_n\tau_{p,n}(r_n). \quad \square
\end{aligned}
$$

Lemma 4 (Coupling under $\tau-$weakly dependence: the sub-blocks)**.** *Suppose that $\mathbb{E}\left(\|X_{n,1}\|\,|X_{n,1}\neq 0\right)<\infty$ for $1\leqslant i\leqslant n$. Moreover assume that, for $j\in\{2,\dots,m_n\}$, there exists a random variable δ_j uniformly distributed on $[0,1]^{r_n-l_n}$ independent of the $\sigma-$algebras $\mathcal{M}_{n,j-1}=\sigma(Y^{(r_n-l_n)}_{n,1},\dots,Y^{(r_n-l_n)}_{n,j-1})$ and $\sigma(Y^{(r_n-l_n)}_{n,j})$. Then there exists a random block $\dot{Y}^{(r_n-l_n)}_{n,j}$, measurable with respect to $\mathcal{M}_{n,j-1}\vee\sigma(Y^{(r_n-l_n)}_{n,j})\vee\sigma(\delta_j)$, independent of $\mathcal{M}_{n,j-1}$ and distributed as $Y^{(r_n-l_n)}_{n,j}$ such that*

$$
(7.58)\qquad \left\|\mathbb{E}\left(d_{r_n-l_n}(Y^{(r_n-l_n)}_{n,j},\dot{Y}^{(r_n-l_n)}_{n,j})\right)\middle|\,\mathcal{M}_{n,j-1}\right\|_p\leqslant r_n\tau_{p,n}(l_n).
$$

If $\mathcal{M}_{n,j-1}=\sigma(\dot{Y}^{(r_n-l_n)}_{n,1},\dots,\dot{Y}^{(r_n-l_n)}_{n,j-1})$ then the blocks $(\dot{Y}^{(r_n-l_n)}_{n,j})_{1\leqslant j\leqslant m_n}$ are independent.

Proof. The same argument previous proof. However note that the sub-blocks $(Y^{(r_n-l_n)}_{n,j})_{j=1,\dots,m_n}$ are separated by l_n variables. $\square$

Proof of Theorem 1. Let $(Y_{n,j})_{1\leqslant j\leqslant m_n}$ be the blocks built from $(X_{n,i})_{1\leqslant i\leqslant n}$. For $k\in\{0,1\}$, we consider the independent blocks $(\dot{Y}_{n,2j-k})_{1\leqslant j\leqslant[m_n/2]}$ coupled to the original blocks $(Y_{n,2j-k})_{1\leqslant j\leqslant[m_n/2]}$, from Lemma 3. Therefore, if we define $\Delta^*_{n,j}:=f(\dot{Y}_{n,j})-f(\dot{Y}^{(r_n-l_n)}_{n,j})$, for $j=1,\dots,m_n$, we have that $\Delta^*_{n,j}(f)\overset{\mathcal{D}}{=}\Delta_{n,j}(f)\overset{\mathcal{D}}{=}\Delta_n(f)$, for each j, where $\Delta_{n,j}(f):=f(Y_{n,j})-f(Y^{(r_n-l_n)}_{n,j})$ and $\Delta_n(f)$ is defined in (4.24). Now, if we consider the assumption (C.1), we can apply [Petrov, 1975]'s Theorem 1

(Section IX.1) to the i.i.d.r.v's $X_{n,j} := (nv_n)^{-1/2}\Delta^*_{n,j}(f)$, so

$$D\dot{Z}_n^{(k)}(f) := \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{[m_n/2]} \left(\Delta^*_{n,2j-k}(f) - \mathbb{E}\Delta^*_{n,2j-k}(f)\right) = o_P(1) \tag{7.59}$$

for $k = 0, 1$. In consequence,

$$DZ_n(f) := \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{m_n} (\Delta_{n,j}(f) - \mathbb{E}\Delta_{n,j}(f)) = o_P(1) \tag{7.60}$$

On the other hand, by Lemma 4, we have that

$$BZ_n(f) := \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{m_n} \left(f(Y_{n,j}^{(r_n-l_n)}) - \mathbb{E}f(Y_{n,j}^{(r_n-l_n)})\right) \tag{7.61}$$

converge weakly in fidis if, and only if

$$B\dot{Z}_n(f) := \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{m_n} \left(f(\dot{Y}_{n,j}^{(r_n-l_n)}) - \mathbb{E}f(\dot{Y}_{n,j}^{(r_n-l_n)})\right) \tag{7.62}$$

converge weakly in fidis (and in this case the limit distributions are the same). The latter holds because $B\dot{Z}_n(f) = \dot{Z}_n(f) - D\dot{Z}_n(f)$ and from the assumptions (C.2) and (C.3).

Finally, as $Z_n(f) = BZ_n(f) + DZ_n(f)$ $\forall f \in \mathcal{F}$, we get the result. $\square$

Proof of Theorem 2. We consider blocks $Y_{n,1}, Y_{n,2}, \ldots, Y_{n,m_n}$. Using Lemma 3, for $k \in \{0, 1\}$ we build independent blocks $(\dot{Y}_{n,2j-k})_{1\leqslant j\leqslant[m_n/2]}$ coupled to the original blocks $(Y_{n,2j-k})_{1\leqslant j\leqslant[m_n/2]}$; that is, $Y_{n,2j-k} \overset{\mathcal{D}}{=} \dot{Y}_{n,2j-k}$ and $\mathbb{E}\delta_{r_n}(Y_{n,2j-k}, \dot{Y}_{n,2j-k}) \to 0$, as $n \to \infty$ for each $j = 1, \ldots, [m_n/2]$ and $k = 0, 1$. Thus Z_n is asymptotically tight iff

$$\dot{Z}_n^{(k)}(f) := \frac{1}{\sqrt{nv_n}} \sum_{j=1}^{[m_n/2]} \left(f(\dot{Y}_{n,2j-k}) - \mathbb{E}f(\dot{Y}_{n,2j-k})\right) \tag{7.63}$$

is asymptotically tight for $k = 0, 1$. The latter is true due to Theorem 2.11.9 in [Van Der Vaart & Wellner, 1996] by setting $Z_{nj}(f) = f(Y_{n,j})$ and $[m_n/2]$ instead of m_n. For the remaining assertion we use Theorem 1. $\square$

Proof of Theorem 3. We follow the same lines as in the previous proof. From the triangle inequality Z_n is asymptotically equicontinuous if $\dot{Z}_n^k$ from eqn. (7.63) is asymptotically equicontinuous for each $k \in \{0, 1\}$; this holds true from Theorem

2.11.1 in [Van Der Vaart & Wellner, 1996].
The uniform convergence to a Gaussian process now follows from Theorem 1. ☐

Proof of Lemma 2. Use the following chain of inequalities:

$$
\begin{aligned}
\mathbb{P}\{Y_n^{(l+1:r_n)} \neq 0 | X_{n,1} \neq 0\} &= v_n^{-1}\mathbb{P}\{Y_n^{(l+1:r_n)} \neq 0, X_{n,1} \neq 0\} \\
&= v_n^{-1} \int_{\{X_{n,1}\neq 0\}} \mathbb{P}\{Y_n^{(l+1:r_n)} \neq 0 | \sigma(X_{n,1})\} d\mathbb{P} \\
&= v_n^{-1} \int_{\{X_{n,1}\neq 0\}} \mathbb{E}[\mathbb{1}\{Y_n^{(l+1:r_n)} \neq 0\} | \sigma(X_{n,1})] d\mathbb{P} \\
&= v_n^{-1} \lim_{a\longrightarrow 0} \int_{\{X_{n,1}\neq 0\}} \mathbb{E}\left[\frac{g_a(X_{n,l+1},\dots,X_{n,r_n})}{(r_n-l)a} | \sigma(X_{n,1})\right] d\mathbb{P} \\
&\leqslant \lim_{a\longrightarrow\infty} \frac{1}{(r_n-l)av_n} \|\mathbb{E}[g_a(X_{n,l+1},\dots,X_{n,r_n})|\sigma(X_{n,1})]\|_p \|\mathbb{1}\{X_{n,1}\neq 0\}\|_q \\
&\leqslant \lim_{a\longrightarrow\infty} \frac{1}{(r_n-l)v_n^{p^{-1}}a} \|\mathbb{E}[g_a(X_{n,l+1},\dots,X_{n,r_n})|\sigma(X_{n,1})] - \mathbb{E}[g_a(X_{n,l+1},\dots,X_{n,r_n})]\|_p \\
&\quad + \lim_{a\longrightarrow 0} \frac{1}{(r_n-l)v_n^{p^{-1}}a} \mathbb{E}|g_a(X_{n,l+1},\dots,X_{n,r_n})| \\
&\leqslant \lim_{a\longrightarrow 0} \frac{\tau_p(\sigma(X_{n,1}),(X_{n,l+1},\dots,X_{n,r_n}))}{(r_n-l)v_n^{p^{-1}}a} + \frac{\mathbb{P}\{Y_n^{(l+1:r_n)} \neq 0\}}{v_n^{p^{-1}}} \\
&\leqslant \frac{\tau_{p,n}(l)}{(r_n-l)v_n^{p^{-1}}a} + (r_n-l)v_n^{1/q},
\end{aligned}
$$

where $p^{-1}+q^{-1}=1$ and $\frac{g_a(x_{l+1},\dots,x_{r_n})}{(r_n-l)a}$ is a Lipschitz function that approximates to $\mathbb{1}\{(x_{l+1},\dots,x_{r_n})\neq 0\}$. Therefore setting $a = v_n^{r^{-1}}$ then the result will follow from assumption (5.26). ☐

Proof of Proposition 1. Use only Lemma 2 and the remaining steps are those in [Drees & Rootzén, 2010], Lemma 2.5. ☐

Proof of Proposition 2. Follows the lines of the proof of Proposition 1. ☐

Proof of Proposition 3. The steps are the same that in the proof of Theorem 2.1 in [Drees, 2011], but replacing the assumptions (C1) and (C2) of his paper by our assumptions (B.1) and (B.2). □

Proof of Corollary 3. We follow the same lines as in the previous proof replacing the assumptions (C1) and (C2) of [Drees, 2011] by our assumptions (B.1) and (B.2) in the proof of Corollary 2.3 of the previous paper. □

Proof of (6.41). Let $(X_i)_{i\geqslant 0}$ be the AR(1)-process (1.6). Note that for each $i \in \mathbb{N}$

$$X_i = \frac{X_0}{b^i} + \sum_{s=1}^{i} \frac{\xi_s}{b^{i-s+1}} \tag{7.64}$$

If n is sufficiently large such that $b^d v_n < 1$, then for $x \in [0,1]^d$:

$$\begin{aligned}
&\mathbb{P}\{X_{n,1} > x, X_{n,1} \neq 0\} = \mathbb{P}\{X_i > a_n x_i + u_n, \text{ for some } i = 1, \dots, d\} \\
&= \mathbb{P}\left\{X_0 > b^i(a_n x_i + u_n) - \sum_{s=1}^{i} b^{s-1}\xi_s, \text{ for some } i = 1, \dots, d\right\} \\
&= \mathbb{P}\left\{X_0 > \min_{i=1,\dots,d}\left\{b^i(a_n x_i + u_n) - \sum_{s=1}^{i} b^{s-1}\xi_s\right\}\right\} \\
&= \sum_{j_1,\dots,j_d \in U(b)} \mathbb{P}\left\{X_0 > \min_{i=1,\dots,d}\left\{b^i(a_n x_i + u_n) - \sum_{s=1}^{i} b^{s-1} j_s\right\}, (\xi_1, \dots, \xi_d) = (j_1, \dots, j_d)\right\} \\
&= \frac{1}{b^d} \sum_{j_1,\dots,j_d \in U(b)} \mathbb{P}\left\{X_0 > \min_{i=1,\dots,d}\left\{b^i(a_n x_i + u_n) - \sum_{s=1}^{i} b^{s-1} j_s\right\}\right\} \\
&= \frac{1}{b^d} \sum_{j_1,\dots,j_d \in U(b)} \left(\max_{i=1,\dots,d}\left\{1 - b^i + \sum_{s=1}^{i} b^{s-1} j_s + b^i v_n(1 - x_i)\right\}_+ \wedge 1\right) \\
&= \frac{1}{b^d} \max_{i=1,\dots,d}\left\{b^i v_n(1 - x_i)\right\},
\end{aligned} \tag{7.65}$$

since $\mu_b(j_1, \dots, j_d; i) := 1 - b^i + \sum_{s=1}^{i} b^{s-1} j_s \leqslant -1$ for all $(j_1, \dots, j_d) \in U^d(b) \setminus \{(b-1, \dots, b-1)\}$ and $\mu_b(b-1, b-1, \dots, b-1) = 0$.

Therefore,

$$\mathbb{P}\{X_{n,1} > x | X_{n,1} \neq 0\} \underset{n\to\infty}{\longrightarrow} \max_{i=1,\dots,d}\{b^{i-d}(1 - x_i)\}. \tag{7.66}$$

□

Proof of (6.43). Let $x, y \in [0,1]^d$. Then as before for $i \geqslant 1$, if n is sufficiently large such that $b^{i+d} v_n < 1$, then we have:

$$
\begin{aligned}
(7.67)\quad & \mathbb{P}\{X_{n,1} > x, X_{n,i+1} > y\} \\
&= \mathbb{P}\left\{X_k > a_n x_k + u_n, X_{i+l} > a_n y_l + u_n,\ \text{for some } (k,j) \in \{1,\dots,d\}^2\right\} \\
&= \mathbb{P}\left\{X_0 > \min_{k=1,\dots,d}\left\{b^k(a_n x_k + u_n) - \sum_{s=1}^{k} b^{s-1}\xi_s\right\},\right. \\
&\qquad \left. X_i > \min_{l=1,\dots,d}\left\{b^l(a_n y_l + u_n) - \sum_{s=1}^{l} b^{s-1}\xi_{s+i}\right\}\right\} \\
&= \sum_{\substack{j_1,\dots,j_d \in U(b)\\ j_{i+1},\dots,j_{i+d} \in U(b)}} \mathbb{P}\left\{X_0 > \min_{k=1,\dots,d}\left\{b^k(a_n x_k + u_n) - \sum_{s=1}^{k} b^{s-1} j_s\right\},\right. \\
&\qquad X_i > \min_{l=1,\dots,d}\left\{b^l(a_n y_l + u_n) - \sum_{s=1}^{l} b^{s-1} j_{s+i}\right\}, \\
&\qquad \left.(\xi_1,\dots,\xi_d,\xi_{i+1},\dots,\xi_{i+d}) = (j_1,\dots,j_d,j_{i+1},\dots,j_{i+d})\right\} \\
&= \frac{1}{b^d} \sum_{\substack{j_1,\dots,j_d \in U(b)\\ j_{i+1},\dots,j_{i+d} \in U(b)}} \mathbb{P}\left\{X_0 > \min_{k=1,\dots,d}\left\{b^k(a_n x_k + u_n) - \sum_{s=1}^{k} b^{s-1} j_s\right\},\right. \\
&\qquad \left. X_i > \min_{l=1,\dots,d}\left\{b^l(a_n y_l + u_n) - \sum_{s=1}^{l} b^{s-1} j_{s+i}\right\}, (\xi_1,\dots,\xi_d) = (j_1,\dots,j_d)\right\} \\
&= \frac{1}{b^d} \sum_{\substack{j_1,\dots,j_d \in U(b)\\ j_{i+1},\dots,j_{i+d} \in U(b)}} \mathbb{P}\left\{X_0 > 1 - \max_{k=1,\dots,d}\left\{\mu_b(j_1,\dots,j_d;k) + b^k v_n(1-x_k)\right\}_+ \wedge 1,\right. \\
&\qquad \left. X_i > 1 - \max_{l=1,\dots,d}\left\{\mu_b(j_{i+1,\dots,j_{i+d}}) + b^l v_n(1-y_l)\right\}_+ \wedge 1, (\xi_1,\dots,\xi_d) = (j_1,\dots,j_d)\right\} \\
&= \frac{1}{b^d}\mathbb{P}\left\{X_0 > 1 - \max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right. \\
&\qquad \left. X_i > 1 - \max_{l=1,\dots,d}\left\{b^l v_n(1-y_l)\right\}, \xi_1 = \ldots = \xi_d = b-1\right\}
\end{aligned}
$$

since $\mu_b(j_1,\dots,j_d;i) := 1 - b^i + \sum_{s=1}^{i} b^{s-1} j_s \leqslant -1$ for all $(j_1,\dots,j_d) \in U^d(b) \setminus \{(b-1,\dots,b-1)\}$ and $\mu_b(b-1,b-1,\dots,b-1) = 0$.

Moreover, note that if $i > d$

$$\begin{aligned}
(7.68)\quad \mathbb{P}\{X_{n,1} > x, X_{n,i+1} > y\} &= \frac{1}{b^d}\mathbb{P}\left\{X_0 > 1 - \max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right.\\
&\left. X_i > 1 - \max_{l=1,\dots,d}\left\{b^l v_n(1-y_l)\right\}, \xi_1 = \ldots = \xi_d = b-1\right\}\\
&= \frac{1}{b^d}\mathbb{P}\left\{X_0 > 1 - \max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right.\\
&\left. X_0 > b^i - b^i \max_{l=1,\dots,d}\left\{b^l v_n(1-y_l)\right\} + 1 - b^d - \sum_{s=d+1}^{i} b^{s-1}\xi_s\right\}\\
&= \frac{1}{b^d}\sum_{j_{d+1},\dots,j_i\in U(b)}\mathbb{P}\left\{X_0 > 1 - \max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right.\\
&\left. X_0 > b^i - b^i \max_{l=1,\dots,d}\left\{b^l v_n(1-y_l)\right\} + 1 - b^d - \sum_{s=d+1}^{i} b^{s-1}j_s, (\xi_{d+1},\dots,\xi_i) = (j_{d+1},\dots,j_i)\right\}\\
&= \frac{1}{b^i}\sum_{j_{d+1},\dots,j_i\in U(b)}\mathbb{P}\left\{X_0 > 1 - \max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right.\\
&\left. X_0 > b^i - b^i \max_{l=1,\dots,d}\left\{b^l v_n(1-y_l)\right\} + 1 - b^d - \sum_{s=d+1}^{i} b^{s-1}j_s\right\}\\
&= \frac{1}{b^i}\sum_{j_{d+1},\dots,j_i\in U(b)}\min\left(\max_{k=1,\dots,d}\left\{b^k v_n(1-x_k)\right\},\right.\\
&\left.\max_{l=1,\dots,d}\left\{b^d + \sum_{s=d+1}^{i} b^{s-1}j_s + b^{l+i}v_n(1-y_l) - b^i\right\}_+, 1\right)\\
&= \frac{v_n}{b^i}\min\left(\max_{k=1,\dots,d}\left\{b^k(1-x_k)\right\}, \max_{l=1,\dots,d}\left\{b^{l+i}(1-y_l)\right\}\right) = v_n H_{b,i}(x,y)
\end{aligned}$$

Similarly for $1 \leqslant i < d$, we obtain that

$$\begin{aligned}(7.69)\quad &\mathbb{P}\{X_{n,1} > x, X_{n,i+1} > y\} \\ &= \frac{v_n}{b^i} \min\left(\max_{k=1\dots,i}\{b^k(1-x_k)\}, \max_{k=i+1,\dots,d}\{b^k \min(1-x_k, 1-y_k)\}, \max_{k=d-i,\dots,d}\{b^{k+i}(1-y_k)\}\right) \\ &= v_n H_{b,i}(x,y)\end{aligned}$$

From Lemma 5.2 - (iii) in [Drees & Rootzén, 2010], $\mathbb{E}|f(Y_n)| = o(\sqrt{nv_n})$. Thus, for n sufficiently large:

$$\begin{aligned}(7.70)\quad &\frac{\mathrm{Cov}\left(f_x(Y_n), f_y(Y_n)\right)}{r_n v_n} \sim \mathbb{P}\{X_{n,1} > x, X_{n,1} > y\} \\ &+ \sum_{i=1}^{r_n-1}\left(1 - \frac{i}{r_n}\right)\left(\mathbb{P}\{X_{n,1} > x, X_{n,i+1} > y\} + \mathbb{P}\{X_{n,1} > y, X_{n,i+1} > x\}\right) \\ &\underset{n\to\infty}{\longrightarrow} \min\left(\max_{k=1,\dots,d}\{b^k(1-x_k)\}, \max_{k=1,\dots,d}\{b^k(1-y_k)\}\right) + \sum_{i=1}^{\infty}\left(H_{b,i}(x,y) + H_{b,i}(y,x)\right).\end{aligned}$$

□

Proof of (6.46). The proof is similar to the proof of the expression (6.41). Indeed,

$$\begin{aligned}(7.71)\quad &\mathbb{P}\left\{\max_{1\leqslant i\leqslant r_n} X_i > 1 - v_n t\right\} = \mathbb{P}\left\{X_i > 1 - v_n t, \text{ for some } i = 1,\dots,r_n\right\} \\ &= \mathbb{P}\left\{X_0 > b^i(1-v_n t) - \sum_{s=1}^{i} b^{s-1}\xi_s, \text{ for some } i = 1,\dots,r_n\right\} \\ &= \mathbb{P}\left\{X_0 > \min_{1\leqslant i\leqslant r_n}\left\{b^i(1-v_n t) - \sum_{s=1}^{i} b^{s-1}\xi_s\right\}\right\} \\ &= \sum_{j_1,\dots,j_{r_n}\in U(b)} \mathbb{P}\left\{X_0 > \min_{1\leqslant i\leqslant r_n}\left\{b^i(1-v_n t) - \sum_{s=1}^{i} b^{s-1}\xi_s\right\}, (\xi_1,\dots,\xi_{r_n}) = (j_1,\dots,j_{r_n})\right\} \\ &= \frac{1}{b^{r_n}} \sum_{j_1,\dots,j_{r_n}\in U(b)} \mathbb{P}\left\{X_0 > \min_{1\leqslant i\leqslant r_n}\left\{b^i(1-v_n t) - \sum_{s=1}^{i} b^{s-1}j_s\right\}\right\} \\ &= \frac{1}{b^{r_n}} \sum_{j_1,\dots,j_{r_n}\in U(b)} \min\left(\max_{1\leqslant i\leqslant r_n}\left\{1 + \sum_{s=1}^{i} b^{s-1}j_s - b^i(1-v_n t)\right\}_+, 1\right).\end{aligned}$$

□

Aknowledgements. Warm thanks are due to the very constructive and friendly help of Olivier Wintenberger in the redaction and the finalization of this paper.
Very specials thanks are also due to an anonymous referee who pointed clearly the weaknesses of a previous version of this work and helped us to make it more adequate for publication.

References

[Andrews, 1984] Andrews, D. *Non strong mixing autoregressive processes.* J. Appl. Probab. 21, 930-934; (1984).

[Ango Nze & Doukhan, 2004] Ango Nze P., Doukhan P. *Weak dependence and applications to econometrics.* Econom Theory 20:9951045; (2004).

[Dedecker & Prieur, 2004a] Dedecker, J., Prieur, C. *Coupling for $\tau-$dependent sequences and applications.* Journal of Theor. Probab. 17-4, 861-885; (2004a).

[Dedecker & Prieur, 2004b] Dedecker, J., Prieur, C. *Couplage pour la distance minimale.* C. R. Acad. Sci. Paris Serie 1 338-10, 805-808; (2004b)

[Dedecker & Prieur, 2005] Dedecker, J., Prieur, C. *New dependence coefficients. Examples and applications to statistics.* Prob. Theor. and Rel. Fields 132, 203-236; (2005).

[Dedecker et al., 2007] Dedecker, J., Doukhan, P., Lang, G., León, J.R., Louhichi,S. & Prieur, C. *Weak dependence: With Examples and Applications* Lecture Notes in Statistics 190, Springer-Verlag; (2007).

[Doukhan, 1994] Doukhan, P. *Mixing: Properties and Examples.* Lect. Notes Statis. 85; (1994).

[Doukhan & Louhichi, 1999] Doukhan, P. & Louhichi, S. *A new weak dependence condition and applications to moment inequalities.*Stoch. Proc. Appl. 84. 313-342; (1999).

[Doukhan et al., 2009] Doukhan, P., Mayo, N. & Truquet L. *Weak dependence, models and some applications.* Metrika 69: 199-225; (2009).

[Doukhan, 2015] Doukhan, P. *Stochastic Models for Time Series.* Springer-Verlag; (2015).

[Drees, 2000] Drees, H. *Weighted Approximations of Tail Processes for β-Mixing Random Variables.* Ann. Appl. Probab. 10, 1274-1301; (2000).

[Drees, 2002] Drees, H. *Tail empirical processes under mixing conditions.* In: H.G. Dehling, T. Mikosch and M. Sorensen (eds.), Empirical Process Techniques for Dependent Data, 325-342. Birkhuser, Boston; (2002).

[Drees, 2003] Drees, H. *Extreme Quantile Estimation for Dependent Data with Applications to Finance.* Bernoulli 9, 617-657; (2003).

[Drees & Rootzén, 2010] Drees, H & Rootzén, H. *Limit Theorems for Empirical Processes of Cluster Functionals.* Ann. Stat. 4, 2145 - 2186; (2010).

[Drees, 2011] Drees, H. *Bias correction for blocks estimators of the extremal index.* Univ. Hamburg; (2011).

[Eberlein, 1984] Eberlein, E. *Weak convergence of partial sums of absolutely regular sequences.* Statist. Probab. Letters 2, 291 - 293; (1984).

[Petrov, 1975] Petrov, V.V. *Sums of Independent Random Variables.* Springer, Berlin; (1975).

[Segers, 2003] Segers, J. *Functionals of clusters of extremes.* Adv. Appl. Probab. 35, 1028 - 1045; (2003).

[Rootzén, 1995] *The tail empirical process for stationary sequences.* Technical report, Department of Mathematics, Chalmers University, Sweden; (1995).

[Rootzén, 2009] Rootzén, H. *Weak convergence of the tail empirical function for dependent sequences*. Stoch. Proc. Appl. 119, 468 - 490; (2009).

[Yun, 2000] Yun, S. *The distribution of cluster functionals of extreme events in a dth-order Markov chain*. J. Appl. Probab. 37, 29-44; (2000).

[Van Der Vaart & Wellner, 1996] Van Der Vaart, A.W. and Wellner, J. A. *Weak Convergence and Empirical Processes*. Springer, New York; (1996).