



**HAL**  
open science

# On the asymptotics of extremal $l_p$ -blocks cluster inference

Gloria Buriticá, Olivier Wintenberger

► **To cite this version:**

Gloria Buriticá, Olivier Wintenberger. On the asymptotics of extremal  $l_p$ -blocks cluster inference. 2024. hal-03912267v2

**HAL Id: hal-03912267**

**<https://hal.science/hal-03912267v2>**

Preprint submitted on 5 Jan 2024 (v2), last revised 10 Jan 2024 (v3)

**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.

January 5, 2024

## ON THE ASYMPTOTICS OF EXTREMAL $\ell^p$ -BLOCKS CLUSTER INFERENCE

GLORIA BURITICÁ AND OLIVIER WINTENBERGER

ABSTRACT. Extremes occur in stationary regularly varying time series as short periods with several large observations, known as extremal blocks. We study cluster statistics summarizing the behavior of functions acting on these extremal blocks. Examples of cluster statistics are the extremal index, cluster size probabilities, and other cluster indices. The purpose of our work is twofold. First, we state the asymptotic normality of block estimators for cluster inference based on consecutive observations with large  $\ell^p$ -norms, for  $p > 0$ . The case  $p = \alpha$ , where  $\alpha > 0$  is the tail index of the time series, has specific nice properties thus we analyze the asymptotic of blocks estimators when approximating  $\alpha$  using the Hill estimator. Second, we verify the conditions we require on classical models such as linear models and solutions of stochastic recurrence equations. Regarding linear models, we prove that the asymptotic variance of classical index cluster-based estimators is null as first conjectured in [28]. We illustrate our findings on simulations.

### 1. INTRODUCTION

We study stationary heavy-tailed time series  $(\mathbf{X}_t)$  in  $(\mathbb{R}^d, |\cdot|)$ , with regularly varying distributions, and tail index  $\alpha > 0$ ; cf. [5], a formal definition is conferred to Section 2.2. In this framework, extremal observations cluster: an extreme value triggers a short period with numerous large observations. This behavior is known to perturb classical inference procedures tailored for independent observations like high quantile inference; see [22]. This clustering effect can be summarized with the *extremal index*, initially introduced in [38] and [39]. We can interpret it as the inverse of the mean number of consecutive exceedances above a high threshold in a short period of time. In this article, we aim to infer statistics of the clustering effect by letting functionals act on consecutive observations with extremal behavior. For example, we can recover the extremal index from this setting and also other important indices of the extremes of the series.

For extremal cluster inference, we consider a sample  $\mathbf{X}_{[1,n]}$  together with a sequence  $(b_n)$ , and we define the sample of disjoint blocks  $(\mathcal{B}_j)_{j=1,\dots,m_n}$  as blocks of consecutive observations:

$$(1.1) \quad \mathcal{B}_j := (\mathbf{X}_{(j-1)b_n+1}, \dots, \mathbf{X}_{jb_n}) = \mathbf{X}_{(j-1)b_n+[1,b_n]},$$

---

2020 *Mathematics Subject Classification*. Primary 60G70 ; Secondary 60F10 62G32 60F05 60G57.

such that  $b_n \rightarrow \infty$ ,  $m_n = n/b_n \rightarrow \infty$ , as  $n \rightarrow \infty$ . Following the  $p$ -clusters theory developed in [12], the extremal behavior of the series is modeled by the  $p$ -clusters which we can recover from the conditional behavior of a block  $\mathcal{B}_j$  given that its  $\ell^p$ -norm is large:

$$\mathbb{P}(\mathcal{B}_1/x_{b_n} \in A \mid \|\mathcal{B}_1\|_p > x_{b_n}) \xrightarrow{w} \mathbb{P}(Y\mathbf{Q}^{(p)} \in A), \quad n \rightarrow \infty,$$

such that  $Y$  is independent of  $\mathbf{Q}^{(p)} \in \ell^p$ ,  $\mathbb{P}(Y > y) = y^{-\alpha}$ , for  $y > 1$ , and  $\|\mathbf{Q}^{(p)}\|_p = 1$  a.s., for  $p \in (0, \infty]$ . The weak convergence holds for a family of shift-invariant continuity sets  $A \subset \ell^p$ , and  $(x_n)$  is a suitable sequence satisfying  $\mathbb{P}(\|\mathcal{B}_1\|_p > x_{b_n}) \rightarrow 0$ , as  $n \rightarrow \infty$ .

In this article we study inference of  $p$ -cluster statistics of the form

$$(1.2) \quad f^{\mathbf{Q}}(p) = \mathbb{E}[f(Y\mathbf{Q}^{(p)})],$$

for suitable  $\ell^p$ -continuity functions  $f : \ell^p \rightarrow \mathbb{R}$  which are invariant to the shift operator of sequences. To infer the cluster statistic (1.2), we use the disjoint blocks estimators proposed in [12], and defined as

$$(1.3) \quad \widehat{f}^{\mathbf{Q}}(p) := \frac{1}{k} \sum_{t=1}^{m_n} f(\mathcal{B}_t/\|\mathcal{B}\|_{p,(k+1)}) \mathbb{1}(\|\mathcal{B}_t\|_p > \|\mathcal{B}\|_{p,(k+1)}),$$

where  $\|\mathcal{B}\|_{p,(1)} \geq \|\mathcal{B}\|_{p,(2)} \geq \dots \geq \|\mathcal{B}\|_{p,(m_n)}$ , denotes the sequence of order statistics of the  $\ell^p$ -norms of blocks defined in (1.1). Examples of cluster statistics are the extremal index, and other important cluster indices.

The case  $p = \alpha$  is particularly relevant for two reasons. First, the  $\ell^\alpha$ -norm's tail  $\mathbb{P}(\|\mathcal{B}_1\|_\alpha > x_{b_n})$  is equivalent to  $b_n \mathbb{P}(|\mathbf{X}_1| > x_{b_n})$ , which is the tail of any  $\ell^p$ -norm of blocks in the iid case. Thus the choice  $p = \alpha$  is ideal for tuning the parameters in the blocks estimator as it is less susceptible to local serial dependencies; see [12] for a discussion on that topic. Second, under mixing and anti-clustering conditions, choosing  $(a_n)$  satisfying  $n \mathbb{P}(|\mathbf{X}_1| > a_n) \rightarrow 1$  as  $n \rightarrow \infty$ , [13] prove that

$$(1.4) \quad \sum_{t=1}^n \varepsilon_{a_n^{-1} \mathbf{x}_t} \xrightarrow{d} \sum_{i=1}^{\infty} \sum_{j=-\infty}^{\infty} \varepsilon_{\Gamma_i^{-1/\alpha} \mathbf{Q}_{ij}}, \quad n \rightarrow \infty,$$

where  $(\Gamma_i)$  are the points of an homogeneous Poisson process,  $(\mathbf{Q}_{ij})_j$  are independent copies of the  $\alpha$ -cluster process  $\mathbf{Q} := \mathbf{Q}^{(\alpha)}$  independent of  $(\Gamma_i)$ . Thus the extremal cluster dependencies of the series are fully modeled using the spectral cluster process  $\mathbf{Q}$ , and from it one can recover the distribution of  $\mathbf{Q}^{(p)}$ , for every  $p \in (0, \infty]$  using the change-of-norms equation given by [12] and recalled in (2.12). This means in general we can estimate  $p$ -cluster statistics using  $\ell^\alpha$ -blocks estimators.

The main goal of this article is to establish the asymptotic normality of the blocks estimators from Equation (1.3), tailored for cluster inference. We state moment, mixing and bias assumptions yielding the existence of a sequence  $(k_n)$ , satisfying  $k = k_n \rightarrow \infty$ ,  $m_n/k_n \rightarrow \infty$  such that

$$(1.5) \quad \sqrt{k}(\widehat{f}^{\mathbf{Q}}(p) - f^{\mathbf{Q}}(p)) \xrightarrow{d} \mathcal{N}(0, \text{Var}(f(Y\mathbf{Q}^{(p)}))), \quad n \rightarrow \infty,$$

and the limit is a centered Gaussian distribution. Our inference methodology can be viewed as a Peak Over Threshold over order statistics of blocks. Moreover, for  $p = \alpha$ , fixing  $k$  and letting first  $n \rightarrow \infty$  in (1.3), we see from (1.4)

$$\widehat{f^{\mathbf{Q}}}(\alpha) \approx \frac{1}{k} \sum_{i=1}^k f((\Gamma_i/\Gamma_{k+1})^{-1/\alpha}(\mathbf{Q}_{it})_t).$$

Then, the simple expression of the asymptotic variance follows as  $(\Gamma_i/\Gamma_{k+1}) \stackrel{d}{=} (U_{k,i})$  where  $U_{k,1} < \dots < U_{k,k}$  are the ordered statistics of iid uniformly distributed  $U_j$ ,  $1 \leq j \leq k$ , and  $U^{-1/\alpha} \stackrel{d}{=} Y$ . This heuristic is extended to any  $p$ -cluster  $\mathbf{Q}^{(p)}$ ,  $p \in (0, \infty]$ , via the change-of-norms in (2.12).

In general for  $p$ -cluster inference, the function  $f(p)$  can involve the tail-index  $\alpha$  in its expression, meaning  $f(p) = f_\alpha(p)$ . Furthermore, we already mentioned that the choice of  $p = \alpha$  has some advantages, thus to implement these procedures we need to replace  $\alpha$  with an estimate  $\hat{\alpha}$ . We then show the asymptotic normality of  $p$ -cluster estimators of  $f_{\hat{\alpha}}(p)$  when we let  $1/\hat{\alpha}$  equal the classical Hill estimator, and we extend the analysis to cover  $\hat{\alpha}$ -cluster inference. Furthermore, we conduct simulations to illustrate that  $\ell^{\hat{\alpha}}$ -block estimators are competitive both in terms of bias and variance for finite sample sizes.

Our asymptotic results highlight how introducing  $\ell^p$ -norm block order statistics in (1.3), instead of order statistics of the sample  $(|\mathbf{X}_t|)$  as in [20, 14], can lead to a better asymptotic variance for cluster inference. We give examples of variance reduction in the case of linear models with short-range dependence, for inference of classical indices. In our examples, the asymptotic variance of linear models  $\text{Var}(f(Y\mathbf{Q}^{(p)}))$  is null because of the deterministic properties of the spectral  $p$ -cluster process of linear models. For linear models, the advantage of replacing thresholds with block maxima records was previously investigated in [28]. Existing works [18, 20, 14, 37] following [28] focus on cluster of exceedances inference such that  $p = \infty$ . Our asymptotic result comforts and extends the heuristics presented in [28] for  $p = \infty$  and linear models to the case  $p < \infty$  and general models. To prove the asymptotic normality of block estimators, we rely on the theory of empirical processes [53], but adapted to block estimators. For this purpose, we build on the previous work of [18], and the modern overview in [37]. To handle the asymptotics of extremal  $\ell^p$ -blocks, we build on the large deviation principles studied in [12], and appeal to the  $p$ -cluster processes theory therein.

The article is organized as follows. Preliminaries on mixing coefficients, regular variation, and the  $p$ -cluster theory of stationary time series are compiled in Section 2. In Section 3 we present our main result in Theorem 3.1, stating the asymptotic normality of the block estimators introduced in Equation (1.3). We work under mixing, moment, and bias conditions on the series that we also present in Section 3. Section 4 studies examples of extremal cluster inference such as estimation of the extremal index, the cluster size

probabilities, and the cluster index for sums. We conclude by verifying our conditions on classical models such as linear processes and stochastic recurrence equations in Section 5. In the case of linear models with short-range dependence, Theorem 5.6 states that the  $\ell^p$ -block estimators of all the aforementioned quantities have null-asymptotic variance. Thereby, they are super-efficient for cluster inference of important indices as conjectured by [28] for  $p = \infty$ . We illustrate the finite-sample performances of our estimators in Section 6. All proofs are deferred to Appendices A, B, C.

**1.1. Notation.** We consider stationary time series  $(\mathbf{X}_t)$  taking values in  $\mathbb{R}^d$ , that we endow with a norm  $|\cdot|$ . Let  $p > 0$ , and  $(\mathbf{x}_t) \in (\mathbb{R}^d)^{\mathbb{Z}}$ . Define the  $p$ -modulus function  $\|\cdot\|_p : (\mathbb{R}^d)^{\mathbb{Z}} \rightarrow [0, \infty]$  as

$$\|(\mathbf{x}_t)\|_p^p := \sum_{t \in \mathbb{Z}} |\mathbf{x}_t|^p,$$

and define the sequential space  $\ell^p$  as  $\ell^p := \{(\mathbf{x}_t) \in (\mathbb{R}^d)^{\mathbb{Z}} : \|(\mathbf{x}_t)\|_p^p < \infty\}$ , with the convention that, for  $p = \infty$ , the space  $\ell^\infty$  refers to sequences with finite supremum norm. For any  $p \in (0, \infty]$ , the  $p$ -modulus functions induce a distance  $d_p$  in  $\ell^p$ , and for  $p \in [1, \infty)$ , it defines a norm. Abusing notation, we call them all  $\ell^p$ -norms. Let  $\tilde{\ell}^p = \ell^p / \sim$  be the shift-invariant quotient space where:  $(\mathbf{x}_t) \sim (\mathbf{y}_t)$  if and only if there exists  $k \in \mathbb{Z}$  such that  $\mathbf{x}_{t-k} = \mathbf{y}_t$ ,  $t \in \mathbb{Z}$ . We also consider the metric space  $(\tilde{\ell}^p, \tilde{d}_p)$  such that for  $[\mathbf{x}], [\mathbf{y}] \in \tilde{\ell}^p$ ,

$$\tilde{d}_p([\mathbf{x}], [\mathbf{y}]) := \inf_{k \in \mathbb{Z}} \{d_p(\mathbf{x}_{t-k}, \mathbf{y}_t), (\mathbf{x}_t) \in [\mathbf{x}], (\mathbf{y}_t) \in [\mathbf{y}]\},$$

and without loss of generality, we write an element  $[\mathbf{x}]$  in  $\tilde{\ell}^p$  also as  $(\mathbf{x}_t)$ . Further details on the shift-invariant spaces are deferred to [12, 4].

The operator norm for  $d \times d$  matrices,  $\mathbf{A} \in \mathbb{R}^{d \times d}$ , is defined as  $|\mathbf{A}|_{op} := \sup_{|\mathbf{x}|=1} |\mathbf{A}\mathbf{x}|$ . The truncation operations of  $(\mathbf{x}_t)$  at the level  $\epsilon$ , for  $\epsilon > 0$ , are defined by

$$(\bar{\mathbf{x}}_t^\epsilon) := (\mathbf{x}_t \mathbb{1}_{|\mathbf{x}_t| \leq \epsilon}), \quad (\underline{\mathbf{x}}_{t_\epsilon}) := (\mathbf{x}_t \mathbb{1}_{|\mathbf{x}_t| > \epsilon}).$$

The notation  $a \wedge b$  denotes the minimum between two constants  $a, b \in \mathbb{Z}$ , and  $a \vee b$  denotes its maximum. We write  $\log^+(x) := \log(x) \vee 0$ , for  $x \in (0, \infty)$ . We sometimes write  $\mathbf{x}$  for the sequence  $\mathbf{x} := (\mathbf{x}_t) \in (\mathbb{R}^d)^{\mathbb{Z}}$ . Furthermore, for  $a, b \in \mathbb{R}$ , and  $a \leq b$ , we write as  $\mathbf{x}_{[a,b]}$  the vector  $(\mathbf{x}_t)_{t=a, \dots, b}$  taking values in  $(\mathbb{R}^d)^{b-a+1}$ . We write  $\mathbf{x}_{[a,b]} \in \tilde{\ell}^p$ , which means we take the natural embedding of  $\mathbf{x}_{[a,b]}$  in  $\tilde{\ell}^p$  defined by assigning zeros to undefined coefficients. It will be convenient to write  $\mathcal{G}_+(\tilde{\ell}^p)$  for the continuous non-negative functions on  $(\tilde{\ell}^p, \tilde{d}_p)$  which vanish in a neighborhood of the origin.

## 2. PRELIMINARIES

**2.1. Mixing coefficients.** Let  $(\mathbf{X}_t)$  be an  $\mathbb{R}^d$ -valued strictly stationary time series defined over a probability space  $((\mathbb{R}^d)^{\mathbb{Z}}, \mathcal{A}, \mathbb{P})$ . The properties

of stationary sequences are usually studied through mixing coefficients. Denote the past and future  $\sigma$ -algebras by

$$\mathcal{F}_{t \leq 0} := \sigma((\mathbf{X}_t)_{t \leq 0}), \quad \mathcal{F}_{t \geq h} := \sigma((\mathbf{X}_t)_{t \geq h}), \quad h \geq 1,$$

respectively. We recall the definition of the mixing coefficients  $(\beta_h)$  below

$$\beta_h := d_{TV}(\mathbb{P}_{\mathcal{F}_{t \leq 0} \otimes \mathcal{F}_{t \geq h}}, \mathbb{P}_{\mathcal{F}_{t \leq 0}} \otimes \mathbb{P}_{\mathcal{F}_{t \geq h}}),$$

where  $d_{TV}(\cdot, \cdot)$  is the total variation distance between two probability measures:  $((\mathbb{R}^d)^{\mathbb{Z}}, \mathcal{A}, \mathbb{P}_1)$ ,  $((\mathbb{R}^d)^{\mathbb{Z}}, \mathcal{A}, \mathbb{P}_2)$ , and  $\mathbb{P}_1 \otimes \mathbb{P}_2(A \times B) := \mathbb{P}_1(A)\mathbb{P}_2(B)$ , for  $A, B \in \mathcal{A}$ . For a summary on mixing conditions see [8, 17, 46].

**Remark 2.1.** A detailed interpretation of the  $\beta$ -mixing coefficients  $(\beta_h)$  in terms of the total variation distance can be found in Chapter 1.2 in [17]. These mixing coefficients are well adapted while working with Markov processes. Indeed, a strictly stationary Harris recurrent Markov chain  $(\mathbf{X}_t)$ , satisfies  $\beta_t \rightarrow 0$  exponentially fast as  $t \rightarrow \infty$ ; see Theorem 3.5 in [8].

**2.2. Regular variation.** We consider stationary time series  $(\mathbf{X}_t)$  taking values in  $(\mathbb{R}^d, |\cdot|)$  and regularly varying with tail index  $\alpha > 0$ : all its finite-dimensional vectors are multivariate regularly varying of the same index. In this case we write  $(\mathbf{X}_t)$  satisfies  $\mathbf{RV}_\alpha$ . Borrowing the ideas in [5],  $(\mathbf{X}_t)$  satisfies  $\mathbf{RV}_\alpha$  if and only if, for all  $h \geq 0$ , there exists a vector  $(\Theta_t)_{|t| \leq h}$ , taking values in  $(\mathbb{R}^d)^{2h+1}$  such that

$$(2.6) \quad \mathbb{P}(x^{-1}(\mathbf{X}_t)_{|t| \leq h} \in \cdot \mid |\mathbf{X}_0| > x) \xrightarrow{d} \mathbb{P}(Y(\Theta_t)_{|t| \leq h} \in \cdot), \quad x \rightarrow \infty,$$

where  $Y$  is independent of  $(\Theta_t)_{|t| \leq h}$  and  $\mathbb{P}(Y > y) = y^{-\alpha}, y > 1$ . We call the sequence  $(\Theta_t)$ , taking values in  $(\mathbb{R}^d)^{\mathbb{Z}}$ , the spectral tail process.

The time series  $(\Theta_t)$  does not inherit the stationarity property of the series. Instead, the time-change formula of [5] holds: for any  $s, t \in \mathbb{Z}, s \leq 0 \leq t$  and for any measurable bounded function  $f : (\mathbb{R}^d)^{t-s+1} \rightarrow \mathbb{R}$ ,

$$(2.7) \quad \mathbb{E}[f(\Theta_{s-i}, \dots, \Theta_{t-i}) \mathbf{1}(|\Theta_{-i}| \neq 0)] = \mathbb{E}[|\Theta_i|^\alpha f(\Theta_s/|\Theta_i|, \dots, \Theta_t/|\Theta_i|)].$$

**2.3.  $\ell^p$ -cluster processes.** Let  $(\mathbf{X}_t)$  be a stationary time series satisfying  $\mathbf{RV}_\alpha$ . For  $p > 0$ , we say the series admits a  $p$ -cluster process  $\mathbf{Q}^{(p)} \in \tilde{\ell}^p$  if there exists a sequence  $(x_n)$ , satisfying

$$(2.8) \quad \mathbb{P}(\|\mathbf{X}_{[1,n]}\|_p > x_n) \sim n c(p) \mathbb{P}(\|\mathbf{X}_1\| > x_n), \quad n \rightarrow \infty,$$

with  $c(p) \in (0, \infty)$ ,  $n \mathbb{P}(\|\mathbf{X}_1\| > x_n) \rightarrow 0$ , and

$$(2.9) \quad \mathbb{P}(\mathbf{X}_{[1,n]}/x_n \in \cdot \mid \|\mathbf{X}_{[1,n]}\|_p > x_n) \xrightarrow{w} \mathbb{P}(Y \mathbf{Q}^{(p)} \in \cdot), \quad n \rightarrow \infty,$$

where  $Y$  is independent of  $\mathbf{Q}^{(p)} \in \tilde{\ell}^p$ ,  $\mathbb{P}(Y > y) = y^{-\alpha}$ , for  $y > 1$ ,  $\|\mathbf{Q}^{(p)}\|_p = 1$  a.s., and the limit in (2.9) holds in  $(\tilde{\ell}^p, \tilde{d}_p)$ . We study below the anti-clustering and vanishing-small values conditions denoted  $\mathbf{AC}$ ,  $\mathbf{CS}_p$ , respectively, which guarantee the existence of  $\ell^p$ -clusters. We rephrase next the Theorem 2.1. of [12].

**Proposition 2.2.** *Let  $(\mathbf{X}_t)$  be a stationary time series satisfying  $\mathbf{RV}_\alpha$ . Let  $(x_n)$  be a sequence such that  $n\mathbb{P}(|\mathbf{X}_1| > x_n) \rightarrow 0$ , as  $n \rightarrow \infty$ , and  $p > 0$ . For all  $\delta > 0$ , assume*

$$\mathbf{AC}: \lim_{s \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_{[s,n]}\|_\infty > \epsilon x_n \mid |\mathbf{X}_1| > \epsilon x_n) = 0, \quad \epsilon > 0,$$

$$\mathbf{CS}_p: \lim_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} \frac{\mathbb{P}(\|\overline{\mathbf{X}}_{[1,n]}/x_n^\epsilon\|_p^p > \delta)}{n\mathbb{P}(|\mathbf{X}_1| > x_n)} = 0$$

Then, if  $p \geq \alpha$ , Equation (2.8) holds with  $c(\infty) \leq c(p) \leq c(\alpha) = 1$ , and  $(\mathbf{X}_t)$  admits a  $p$ -cluster process  $\mathbf{Q}^{(p)}$  in the sense of (2.9). If  $p < \alpha$ , existence of the  $p$ -cluster process holds if  $\mathbb{E}[\|\mathbf{Q}^{(\alpha)}\|_p^\alpha] < \infty$ . In this case, Equation (2.8) holds with  $c(p) = \mathbb{E}[\|\mathbf{Q}^{(\alpha)}\|_p^\alpha]$ .

We see from Proposition 2.2 assuming  $\mathbf{AC}$  and  $\mathbf{CS}_\alpha$  implies the time series  $(\mathbf{X}_t)$  admits an  $\alpha$ -cluster  $\mathbf{Q}^{(\alpha)}$ , where  $\alpha > 0$ , denotes the tail index. In this case, appealing to Proposition 3.1. in [12], we have

$$(2.10) \quad \mathbf{Q} := \mathbf{Q}^{(\alpha)} \stackrel{d}{=} \boldsymbol{\Theta} / \|\boldsymbol{\Theta}\|_\alpha, \quad \in \tilde{\ell}^\alpha,$$

where  $(\boldsymbol{\Theta}_t)$  is the spectral tail process from Equation (2.6). Moreover, if  $\mathbf{CS}_p$ ,  $\mathbf{CS}_{p'}$ , and  $\mathbb{E}[\|\mathbf{Q}\|_p^\alpha] + \mathbb{E}[\|\mathbf{Q}\|_{p'}^\alpha] < \infty$  also hold, then the  $p, p'$ -clusters exist and are related by the change-of-norms formula below

$$(2.11) \quad \begin{aligned} & \mathbb{P}(\mathbf{Q}^{(p)} \in \cdot) \\ &= c(p)^{-1} \mathbb{E}[\|\mathbf{Q}\|_p^\alpha \mathbf{1}(\mathbf{Q}/\|\mathbf{Q}\|_p \in \cdot)] \end{aligned}$$

$$(2.12) \quad = \frac{c(p')}{c(p)} \mathbb{E}[\|\mathbf{Q}^{(p')}\|_p^\alpha \mathbf{1}(\mathbf{Q}^{(p')}/\|\mathbf{Q}^{(p')}\|_p \in \cdot)].$$

Since  $\|\mathbf{Q}^{(p)}\|_p = 1$  a.s. for any  $p > 0$ , then  $c(\alpha) = 1$ , and  $\mathbb{E}[\|\mathbf{Q}^{(p')}\|_p^\alpha] = c(p)/c(p')$ , where  $c(p), c(p')$ , are as in Equation (2.8). In the following we denote by  $\mathbf{Q}$  the  $\alpha$ -cluster as in (2.10).

**Remark 2.3.** *Using the monotonicity of norms, we see  $\mathbf{CS}_p$  implies  $\mathbf{CS}_{p'}$ , for  $p' > p > 0$ . If  $p > \alpha$ , condition  $\mathbf{CS}_p$  is always satisfied for sequences  $(x_n)$  such that  $n\mathbb{P}(|\mathbf{X}_1| > x_n) \rightarrow 0$ , as  $n \rightarrow \infty$ . If  $\alpha/2 < p \leq \alpha$ ,  $\mathbf{CS}_p$  still holds for short-range dependence models and sequences  $(x_n)$  such that there exists  $\kappa > 0$ , satisfying  $n/x_n^{p \wedge (\alpha - \kappa)} \rightarrow 0$ , as  $n \rightarrow \infty$  (see remarks 5.1. and 5.2. in [12]). We verify this condition for classical models in Section 5.*

**2.4. Tail-index estimation.** To estimate the tail-index  $\alpha$  of the series, we use the Hill estimator:

$$(2.13) \quad \frac{1}{\hat{\alpha}^n} := \frac{1}{\hat{\alpha}^n(k')} := \frac{1}{k'} \sum_{t=1}^n \log(|\mathbf{X}_t|/|\mathbf{X}|_{(k'+1)}),$$

where  $|\mathbf{X}|_{(1)} \geq |\mathbf{X}|_{(2)} \geq \dots \geq |\mathbf{X}|_{(n)}$ , and  $k' = k'(n)$  is a tuning sequence for (2.13) satisfying  $k' \rightarrow \infty$ ,  $n/k' \rightarrow \infty$ , as  $n \rightarrow \infty$ . To study the asymptotic properties of the Hill estimator, we write it as a cluster statistic and consider the functional  $h: \tilde{\ell}^p \rightarrow \mathbb{R}_{\geq 0}$  given by

$$(2.14) \quad h(\mathbf{x}_t) = \sum_{t \in \mathbb{Z}} \log(|\mathbf{x}_t|) \mathbb{1}(|\mathbf{x}_t| > 1).$$

It is easy to see  $h(\mathbf{x}_t) = h(\mathbf{x}_t) \mathbb{1}(\|\mathbf{x}_t\|_p > 1)$  for every  $p > 0$ , and

$$h^{\mathbf{Q}}(p) := \int_0^\infty \mathbb{E}[\sum_{t \in \mathbb{Z}} \log(y|\mathbf{Q}_t|) \mathbb{1}(y|\mathbf{Q}_t| > 1)] d(-y^{-\alpha}) = \frac{1}{\alpha}.$$

With this interpretation  $p > 0$  is arbitrary and can be chosen as  $p = \alpha$  for convenience. We also introduce the counts of exceedances functional  $e : \tilde{\ell}^p \rightarrow (0, \infty)$  given by

$$(2.15) \quad e(\mathbf{x}_t) := \sum_{j \in \mathbb{Z}} \mathbb{1}(|\mathbf{x}_t| > 1),$$

which also satisfies  $e(\mathbf{x}_t) = e(\mathbf{x}_t) \mathbb{1}(\|\mathbf{x}_t\|_p > 1)$  and

$$e^{\mathbf{Q}}(\alpha) = \int_0^\infty \mathbb{E}[\sum_{t \in \mathbb{Z}} \mathbb{1}(y|\mathbf{Q}_t| > 1)] d(-y^{-\alpha}) = 1.$$

### 3. ASYMPTOTIC NORMALITY

**3.1. Blocks estimation.** Let  $(\mathbf{X}_t)$  be an  $\mathbb{R}^d$ -valued stationary time series satisfying  $\mathbf{RV}_\alpha$ . We fix  $p \in (0, \infty]$ , and assume the conditions of Proposition 2.2 hold for  $p$ , thus the series admits a  $p$ -cluster process  $\mathbf{Q}^{(p)} \in \tilde{\ell}^p$ , and (2.9) holds for a sequence of high levels  $(x_n)$  satisfying  $\mathbb{P}(\|\mathbf{X}_{[1,n]}\|_p > x_n) \rightarrow 0$ , as  $n \rightarrow \infty$ . Given a functional  $f(p) : \tilde{\ell}^p \rightarrow \mathbb{R}$ , recall the block estimator in (1.3) is tuned with the block lengths  $(b_n)$ , and the number  $(k_n)$  of extremal blocks. The total number of disjoint blocks in a sample is denoted  $(m_n)$  with  $m_n = \lfloor n/b_n \rfloor$ . We assume the relation between  $(k_n)$ ,  $(x_n)$  and  $(b_n)$ :

$$(3.16) \quad \begin{aligned} k_n &:= k_n(p) = \lfloor m_n \mathbb{P}(\|\mathcal{B}_1\|_p > x_{b_n}) \rfloor \\ &\sim n c(p) \mathbb{P}(\|\mathbf{X}_1\|_p > x_{b_n}), \quad n \rightarrow \infty, \end{aligned}$$

holds, where  $c(p) \in (0, \infty)$  are as in (2.8), and  $c(\alpha) = 1$ . The relation (3.16) is key to derive the asymptotics of the blocks estimator.

In what follows, if the sequences  $(x_n)$ ,  $(b_n)$ ,  $(k_n)$ ,  $(m_n)$ , appear then they coincide with the ones mentioned here.

It will be useful to consider the functional  $1(p) : \tilde{\ell}^p \rightarrow \mathbb{R}$  defined by

$$(3.17) \quad 1(p)(\mathbf{x}_t) := \mathbb{1}(\|\mathbf{x}_t\|_p > 1),$$

which satisfies  $1^{\mathbf{Q}}(p) = \mathbb{P}(Y \|\mathbf{Q}^{(p)}\|_p > 1) = 1$ . In numerous examples, the function  $f(p)$  might depend on  $\alpha$  on its expression, meaning  $f(p) = f_\alpha(p)$ . Hence, it will also be useful to consider the family of functions  $f_q(p) : \tilde{\ell}^p \rightarrow \mathbb{R}$ , indexed by  $q$ , and we take  $q$  in a neighbourhood of  $\alpha$ .

**3.2. Main result.** Our main result is presented in Theorem 3.1 stating the asymptotic normality of the  $\ell^p$ -blocks estimator in (1.3) under the Lindeberg, bias and mixing assumptions below. We cover the implementation of



$\ell^p$ -blocks estimators for functionals  $f_\alpha(p)$  where we replace  $\alpha$  with an estimate  $\hat{\alpha}$ . We extend the result to  $\hat{\alpha}$ -cluster inference and choose extremal blocks as those with the largest  $\ell^{\hat{\alpha}}$ -norm where  $1/\hat{\alpha}$  is the Hill estimator. We consider below additional moment and smoothness assumptions for studying the asymptotics of these inference procedures.

**L:** Fix  $\epsilon > 0$ , and let  $f = f(p) \in \mathcal{G}_+(\tilde{\ell}^p)$ , such that  $u \mapsto f(\mathbf{x}_t)/u$  is a non-increasing, and there exists  $\delta > 0$  such that, for all  $u > 0$ , the following Lindeberg-type condition holds

$$(3.18) \quad \lim_{n \rightarrow \infty} \frac{\mathbb{E}[(f(\mathbf{X}_{[1,n]}/(u x_n)))^{2+\delta} \mathbf{1}(\|\mathbf{X}_{[1,n]}/x_n\|_p > u)]}{\mathbb{P}(\|\mathbf{X}_{[1,n]}\|_\alpha > x_n)} < \infty,$$

and  $\sup_{q \in [\alpha-\epsilon, \alpha+\epsilon]} \mathbb{E}[(f(Y\mathbf{Q}^{(q)}))^2] < \infty$ .

**B**( $k_n$ ): Fix  $\epsilon > 0$ , consider  $f$  satisfying **L**, and assume the bias condition:

$$(3.19) \quad \lim_{n \rightarrow \infty} \sqrt{k_n} \sup_{u \in [1-\epsilon, 1+\epsilon]} \left| \frac{\mathbb{E}[f(\mathcal{B}_1/(u x_{b_n})) \mathbf{1}(\|\mathcal{B}_1/x_{b_n}\|_p > u)]}{\mathbb{P}(\|\mathcal{B}_1\|_p > x_{b_n})} - u^{-\alpha} f^{\mathbf{Q}}(p) \right| = 0,$$

holds where  $f^{\mathbf{Q}}(p)$  is as in (1.2).

**B** $_\alpha$ ( $k_n$ ): Fix  $\epsilon > 0$ , consider  $f$  satisfying **L**, and assume the bias condition:

$$(3.20) \quad \lim_{n \rightarrow \infty} \sqrt{k_n} \sup_{\substack{u \in [1-\epsilon, 1+\epsilon], \\ q \in [\alpha-\epsilon, \alpha+\epsilon]}} \left| \frac{\mathbb{E}[f(\mathcal{B}_1/(u x_{b_n})) \mathbf{1}(\|\mathcal{B}_1/x_{b_n}\|_q > u)]}{\mathbb{P}(\|\mathcal{B}_1\|_q > x_{b_n})} - u^{-\alpha} f^{\mathbf{Q}}(q) \right| = 0.$$

**MX** $_\beta$ : Consider  $f$  satisfying **L**, and let  $\delta$  be such that (3.18) hold. Assume that the mixing coefficients  $(\beta_t)$  satisfy for some sequence  $(\ell_n)$  satisfying  $\ell_n \rightarrow \infty$ , and  $\ell_n/b_n \rightarrow 0$ ,  $m_n \beta_{\ell_n}/k_n \rightarrow 0$ ,  $\ell_n/b_n \rightarrow 0$ , as  $n \rightarrow \infty$ ,

$$(3.21) \quad \lim_{n \rightarrow \infty} \sum_{t=2}^{m_n} (m_n \beta_{t b_n} / k_n)^{\frac{\delta}{2+\delta}} = 0.$$

If  $f$  is bounded, assume  $\sum_{t=2}^{m_n} m_n \beta_{t b_n} / k_n \rightarrow 0$  instead of (3.21).

**M:** Fix  $\epsilon > 0$ , and assume that the moment condition below holds

$$\mathbb{E}[\|\mathbf{Q}\|_{\alpha-\epsilon}^{\alpha+\epsilon}] < \infty,$$

which implies  $q \mapsto c(q)$  is a continuous function at  $\alpha$ .

**S:** Consider  $f$  satisfying **L**. If  $f$  has  $\alpha$  on its expression, i.e.,  $f = f_\alpha$ , then the function obtained by substituting  $\alpha$  by  $q$ , denoted  $f_q$ , admits the Taylor development:

$$(3.22) \quad f_q(\mathbf{x}) = f_\alpha(\mathbf{x}) + (q - \alpha) \frac{\partial f_q}{\partial q} \Big|_{q=\alpha}(\mathbf{x}) + (q - \alpha)^2 \frac{\partial^2 f_q}{\partial q^2} \Big|_{q=\xi}(\mathbf{x}),$$

where  $\xi = \xi(\mathbf{x})$  is a real value in  $[\alpha - \epsilon, \alpha + \epsilon]$ . Moreover, assume  $\partial f_q / \partial q|_{q=q'}$  and  $\partial^2 f_q / \partial q^2|_{q=q'}$  are  $\tilde{\ell}^\alpha$ -continuous, and there exists  $\delta > 0$  such that

$$\frac{\mathbb{E}\left[\sup_{q' \in [\alpha - \epsilon, \alpha + \epsilon]} \left(\frac{\partial f_q}{\partial q}\Big|_{q=q'} + \frac{\partial^2 f_q}{\partial q^2}\Big|_{q=q'}\right)^{1+\delta} (\mathbf{X}_{[1,n]}/x_n) \mathbf{1}(\|\mathbf{X}_{[1,n]}\|_{q'} > x_n)\right]}{\mathbb{P}(\|\mathbf{X}_{[1,n]}\|_\alpha > x_n)} < \infty.$$

We state in Theorem 3.1 below our main result on the asymptotic normality of the blocks estimator following the notation from equations (1.2) and (1.3). We defer its proof to Section A.

**Theorem 3.1.** *Let  $(\mathbf{X}_t)$  be a stationary time series satisfying  $\mathbf{RV}_\alpha$ . Assume the conditions of Proposition 2.2 hold and the series admits a  $p$ -cluster process  $\mathbf{Q}^{(p)} \in \tilde{\ell}^p$ . Consider  $f = f(p) : \tilde{\ell}^p \rightarrow \mathbb{R}$ , and assume  $f(p)$  and  $1(p)$  satisfy  $\mathbf{B}(k_n)$ , and  $h, e$ , in (2.14), (2.15) satisfy  $\mathbf{B}(k'_n)$ . Assume  $\mathbf{MX}_\beta$  holds. If  $f(p) = f_\alpha(p)$ , assume  $\mathbf{M}$  and  $\mathbf{S}$  hold. Then,*

$$(3.23) \quad \sqrt{k_n} \left( \widehat{f_\alpha^{\mathbf{Q}}}(p) - f_\alpha^{\mathbf{Q}}(p) \right) \xrightarrow{d} \mathcal{N}\left(0, \text{Var}(f_\alpha(Y\mathbf{Q}^{(p)})) + \kappa^2 (\partial f_q / \partial q|_\alpha^{\mathbf{Q}})^2 \alpha^2 \sigma_\alpha^2\right), \quad n \rightarrow \infty,$$

with  $\sigma_\alpha^2 := \mathbb{E}[(\alpha h(Y\mathbf{Q}) - e(Y\mathbf{Q}))^2]$ , and  $k_n/k'_n \rightarrow \kappa$ , with  $\kappa \geq 0$ ,  $Y$  independent of  $\mathbf{Q}^{(p)}$ , and  $\mathbb{P}(Y > y) = y^{-\alpha}$ , for  $y > 1$ . Moreover, for  $p = \alpha$ , assume  $f(\alpha)$  and  $1(\alpha)$  satisfy  $\mathbf{B}_\alpha(k_n)$ , and assume  $\mathbf{M}$  holds. Then, (3.23) continues to hold replacing  $p$  by  $\hat{\alpha}$  in the estimator.

The choice of  $k_n, k'_n$  are subject to the bias conditions  $\mathbf{B}(k_n)$  and  $\mathbf{B}(k'_n)$ . Actually, it is common practice to choose  $k'_n$  larger than  $k_n$ , and the numerical results from Section 6 support this practice. When we use fewer blocks  $k_n$  for  $p$ -cluster inference, compared to the number of records  $k'_n$  we use to tune the Hill estimator, precisely if  $k_n/k'_n \rightarrow 0$ , as  $n \rightarrow \infty$ , the variance term simplifies to

$$\sqrt{k} \left( \widehat{f_\alpha^{\mathbf{Q}}}(p) - f_\alpha^{\mathbf{Q}}(p) \right) \xrightarrow{d} \mathcal{N}\left(0, \text{Var}(f_\alpha(Y\mathbf{Q}^{(p)}))\right), \quad n \rightarrow \infty.$$

This expression also holds when the functional  $f$  doesn't include  $\alpha$  on its expression.

**Remark 3.2.** *To plug-in  $\hat{\alpha}$  in the place of  $\alpha$  for  $p$ -cluster inference we require the bias condition  $\mathbf{B}(k'_n)$  for  $h, e$ . To do so, the Hill estimator is seen as a block estimator as in (1.3) that evaluates the block functional  $h$  on  $(\mathcal{B}_t/x'_{b_n})$ ,  $t = 1, \dots, m_n$ , replacing the high threshold level  $(x'_{b_n})$  by  $k'$ -order statistic from the sample  $(|\mathbf{X}_t|)$  where*

$$k'_n(\alpha) \sim n \mathbb{P}(|\mathbf{X}_1| > x'_{b_n}), \quad n \rightarrow \infty.$$

Then the bias condition  $\mathbf{B}(k'_n)$  can be rewritten as

$$\lim_{n \rightarrow \infty} \sqrt{k'_n} \sup_{u \in [1-\epsilon, 1+\epsilon]} \left| \frac{\mathbb{E}[\log(|\mathbf{X}_1|/(u x'_{b_n})) \mathbf{1}(|\mathbf{X}_1| > (u x'_{b_n}))]}{\mathbb{P}(|\mathbf{X}_1| > x'_{b_n})} - u^{-\alpha} \frac{1}{\alpha} \right| = 0,$$

which is no longer a condition on blocks, but on  $|\mathbf{X}_1|$ . This type of condition was considered in [19, 15]. Notice that the dependence of the threshold  $x'_{b_n}$  with  $b_n$  is an artifact of our notation for  $\mathbf{B}$  which is common for  $(k_n)$  extremal blocks or  $(k'_n)$  extremal components.

**Remark 3.3.** The mixing condition  $\mathbf{MX}_\beta$  is comparable to the classical mixing conditions for central limit theory for block estimators, notably, (B1) and (B2) in [18]. In our notation we can write  $k_n = nv_n$ ,  $b_n = r_n/k_n$ , and in both cases we can take  $\ell_n = o(b_n)$ . Then, their assumption reads as  $n\beta\ell_n/(k_nb_n) \rightarrow 0$ , as  $n \rightarrow \infty$ , which is the first part of  $\mathbf{MX}_\beta$ . Actually, in the models we consider in Section 5, this is the leading term of our mixing assumption. Instead, our gain comes from the way we deal with the entropy of the cluster functionals, which is novel here. In [18], the authors require additionally (D2) or (D2'). In this case, for example, to deal with the extremal index (see e.g. [20]), the authors also require  $b_n\sqrt{k_n} \rightarrow 0$ , as  $n \rightarrow \infty$ , which is no longer an assumption in our theory. Instead, we show in Section 5 that in the classical model, we can take  $k_n$  as in (5.36), provided that the bias assumptions are met.

#### 4. CLUSTER STATISTICS

In view of Theorem 3.1, we derive asymptotic normality of the classical cluster index estimators in extreme value theory.

**4.1. The extremal index.** Let  $(\mathbf{X}_t)$  be a stationary time series in  $(\mathbb{R}^d, |\cdot|)$  satisfying  $\mathbf{RV}_\alpha$ . The extremal index  $\theta_{|\mathbf{X}|}$  of the series  $(|\mathbf{X}_t|)$  is a measure of serial clustering introduced in [38] and [39]. We recall the extremal index estimator proposed in [13], based on extremal  $\ell^\alpha$ -blocks.

**Corollary 4.1.** Consider  $f_\alpha : \tilde{\ell}^\alpha \rightarrow \mathbb{R}$  to be the function  $\mathbf{x} \mapsto \|\mathbf{x}\|_\infty^\alpha / \|\mathbf{x}\|_\alpha^\alpha$ . Assume the conditions of Theorem 3.1 hold for  $p = \alpha$ , and  $k/k' \rightarrow 0$ , as  $n \rightarrow \infty$ . Let  $\theta_{|\mathbf{X}|} = \mathbb{E}[\|\mathbf{Q}\|_\infty^\alpha]$ , hence we deduce an estimator

$$(4.24) \quad \widehat{\theta}_{|\mathbf{X}|} := \frac{1}{k} \sum_{t=1}^m \frac{\|\mathcal{B}_t\|_\infty^{\widehat{\alpha}}}{\|\mathcal{B}_t\|_{\widehat{\alpha}}^{\widehat{\alpha}}} \mathbb{1}(\|\mathcal{B}_t\|_{\widehat{\alpha}} > \|\mathcal{B}\|_{\widehat{\alpha},(k+1)}),$$

such that

$$\sqrt{k}(\widehat{\theta}_{|\mathbf{X}|} - \theta_{|\mathbf{X}|}) \xrightarrow{d} \mathcal{N}(0, \text{Var}(\|\mathbf{Q}\|_\infty^\alpha)), \quad n \rightarrow \infty.$$

*Proof.* The proof of Corollary 4.1 follows directly as  $f_\alpha \in \mathcal{G}_+(\tilde{\ell}^\alpha)$  is a bounded continuous function satisfying  $\mathbf{L}$ .  $\square$

For comparison, we also review the blocks estimator based on extremal  $\ell^\infty$ -blocks proposed in [27]:

$$(4.25) \quad \widehat{\theta}_{|\mathbf{X}|}^B := \frac{1}{k} \sum_{t=1}^{m_n} \mathbb{1}(\|\mathcal{B}_t\|_\infty > |\mathbf{X}|_{(k+1)}).$$

Direct computations from Example 10.4.2 in [37] yield

$$\sqrt{k}(\widehat{\theta}_{|\mathbf{X}|}^B - \theta_{|\mathbf{X}|}) \xrightarrow{d} \mathcal{N}(0, \sigma_\theta^2), \quad n \rightarrow \infty,$$

where  $\sigma_\theta^2 \in [0, \infty)$ , and

$$\begin{aligned} \sigma_\theta^2 &:= \theta_{|\mathbf{X}|}^2 \sum_{j \in \mathbb{Z}} \mathbb{E}[|\Theta_j|^\alpha \wedge 1] - \theta_{|\mathbf{X}|} \\ (4.26) \quad &:= \theta_{|\mathbf{X}|}^2 \sum_{j \in \mathbb{Z}} \sum_{t \in \mathbb{Z}} \mathbb{E}[|\mathbf{Q}_{j+t}|^\alpha \wedge |\mathbf{Q}_t|^\alpha] - \theta_{|\mathbf{X}|}. \end{aligned}$$

The last equality follows appealing to the time-change formula in (2.7) and Equation (2.10). As a result, we can compare the asymptotic variances of  $\widehat{\theta}_{|\mathbf{X}|}$  and  $\widehat{\theta}_{|\mathbf{X}|}^B$  in the cases where  $\mathbf{Q}$  is known. This is the topic of Section 5.

**Remark 4.2.** *An alternative  $\alpha$ -cluster estimator of the extremal index corresponds to the block functional  $f'(\mathbf{x}) = \mathbb{1}(\|\mathbf{x}\|_\infty > 1)$ . A similar asymptotic normality result applies but with an asymptotic variance  $\text{Var}(f'(Y\mathbf{Q}))$  larger than  $\text{Var}(f_\alpha(Y\mathbf{Q}))$ . It motivates the use of  $\widehat{\theta}_{|\mathbf{X}|}$  although it requires the estimation of  $\alpha$ . The latter using Hill's estimator is harmless choosing  $k'$  sufficiently large with respect to  $k$ , which is always the case in practice.*

**Remark 4.3.** *To check condition **S** for the functional  $f_\alpha(\mathbf{x}) = \|\mathbf{x}\|_\infty^\alpha / \|\mathbf{x}\|_\alpha^\alpha$ , note the Taylor expansions:*

$$\begin{aligned} &\|\mathbf{x}\|_q^q - \|\mathbf{x}\|_\alpha^\alpha \\ &= (q - \alpha) \sum_{t \in \mathbb{Z}} |\mathbf{x}_t|^\alpha \log(|\mathbf{x}_t|) + (q - \alpha)^2 \sum_{t \in \mathbb{Z}} |\mathbf{x}_t|^{q'} \log^2(|\mathbf{x}_t|), \\ &\|\mathbf{x}\|_\infty^q - \|\mathbf{x}\|_\infty^\alpha \\ &= (q - \alpha) \|\mathbf{x}\|_\infty^\alpha \log(\|\mathbf{x}\|_\infty) + (q - \alpha)^2 \|\mathbf{x}\|_\infty^{q''} \log^2(\|\mathbf{x}\|_\infty), \end{aligned}$$

hold for some  $q', q'' \in [\alpha \wedge q, \alpha \vee q]$ . Hence,  $f_q$  satisfies

$$\begin{aligned} &f_q(\mathbf{x}) - f_\alpha(\mathbf{x}) \\ &= (q - \alpha) \underbrace{\frac{\|\mathbf{x}\|_\infty^\alpha}{\|\mathbf{x}\|_\alpha^\alpha} \sum_{t \in \mathbb{Z}} \frac{|\mathbf{x}_t|^\alpha}{\|\mathbf{x}\|_\alpha^\alpha} \log(\|\mathbf{x}\|_\infty / |\mathbf{x}_t|)}_{:= \partial f_q / \partial q|_{q=\alpha}} \\ &\quad + (q - \alpha)^2 \frac{1}{2} \underbrace{\frac{\|\mathbf{x}\|_\infty^{q'}}{\|\mathbf{x}\|_{q'}^{q'}} \sum_{t \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} \frac{|\mathbf{x}_t|^{q'} |\mathbf{x}_j|^{q'}}{\|\mathbf{x}\|_{q'}^{q'} \|\mathbf{x}\|_{q'}^{q'}} \log(\|\mathbf{x}\|_\infty / |\mathbf{x}_j|) \log(|\mathbf{x}_j| / |\mathbf{x}_t|)}_{:= \partial^2 f_q / \partial q^2|_{q=q'}}. \end{aligned}$$

As mentioned, this expansion is helpful to verify condition **S** on the models from Section 5.

**4.2. The cluster index for sums.** Let  $(\mathbf{X}_t)$  be a stationary time series with values in  $(\mathbb{R}^d, |\cdot|)$  satisfying  $\mathbf{RV}_\alpha$ . We recall [43] coined the constant

$c(1)$  in (2.8) as the cluster index for sums. We review a cluster-based estimator of it, introduced in [12], based on extremal  $\ell^\alpha$ -blocks.

**Corollary 4.4.** *Consider  $f_\alpha : \tilde{\ell}^\alpha \rightarrow \mathbb{R}$  to be the function  $\mathbf{x} \mapsto \|\mathbf{x}\|_1^\alpha / \|\mathbf{x}\|_\alpha^\alpha$ . Assume the conditions of Theorem 3.1 hold for  $p = \alpha \wedge 1$ , and  $\alpha < 2$ , and  $k/k' \rightarrow 0$ , as  $n \rightarrow \infty$ . Let  $c(1) = \mathbb{E}[\|\mathbf{Q}\|_1^\alpha] < \infty$ , hence one deduces an estimator*

$$(4.27) \quad \widehat{c}(1) := \frac{1}{k} \sum_{t=1}^m \frac{\|\mathcal{B}_t\|_{\widehat{\alpha}}}{\|\mathcal{B}_t\|_{\widehat{\alpha}}} \mathbf{1}(\|\mathcal{B}_t\|_{\widehat{\alpha}} > \|\mathcal{B}\|_{\widehat{\alpha},(k+1)}),$$

such that

$$\sqrt{k}(\widehat{c}(1) - c(1)) \xrightarrow{d} \mathcal{N}(0, \text{Var}(\|\mathbf{Q}\|_1^\alpha)), \quad n \rightarrow \infty,$$

and  $c(1)$  is as in (2.8) with  $p = 1$ .

*Proof.* The proof of Corollary 4.4 follows directly from Theorem 3.1 as  $f_\alpha \in \mathcal{G}_+(\ell^\alpha)$  is a bounded continuous function satisfying **L**.  $\square$

Another sums index cluster-based estimator we can consider is the one proposed in [37] based on extremal  $\ell^\infty$ -blocks:

$$(4.28) \quad \widehat{c}^B(1) = \frac{1}{kb_n} \sum_{t=1}^{m_n} \mathbf{1}(\|\mathcal{B}_t\|_1 > |\mathbf{X}|_{(k+1)}).$$

Then, relying on Example 10.4.2 in [37],

$$\sqrt{k}(\widehat{c}^B - c(1)) \xrightarrow{d} \mathcal{N}(0, \sigma_{c(1)}^2), \quad n \rightarrow \infty,$$

for a constant  $\sigma_{c(1)}^2 \in [0, \infty)$  defined by

$$(4.29) \quad \sigma_{c(1)}^2 = c(1)^2 \sum_{j \in \mathbb{Z}} \sum_{t \in \mathbb{Z}} \mathbb{E}[|\mathbf{Q}_{j+t}|^\alpha \wedge |\mathbf{Q}_t|^\alpha] - c(1).$$

Similarly as in Example 4.1, whenever  $\mathbf{Q}$  is known, we can directly compare the asymptotic variances relative to the estimators  $\widehat{c}(1)$  and  $\widehat{c}^B(1)$ . Section 5 covers this topic for classical models where the cluster process is known.

Moreover, we can use the computations in Remark 4.3 to verify condition **S** holds. In this case it suffices to replace  $\|\mathbf{x}\|_\infty$  by  $\|\mathbf{x}\|_1$  in all its appearances.

**4.3. The cluster sizes.** In general, a classical approach to model serial exceedances is using point processes as in [39] and [28]. For the levels  $(a_n)$ , satisfying  $n\mathbb{P}(|\mathbf{X}_1| > a_n) \rightarrow 1$ , as  $n \rightarrow \infty$ , and for every fixed  $x > 0$  consider the point process of exceedances with state space  $(0, 1]$ :

$$\eta_{n,x}(\cdot) := \overline{N}_n(\{\mathbf{y} : |\mathbf{y}| > x\} \times \cdot) = \sum_{i=1}^n \varepsilon_{i/n}(\cdot) \mathbf{1}(|\mathbf{X}_i| > x a_n).$$

Under mixing and anti-clustering conditions, for fixed  $x > 0$ , we can express the limiting point process in [28] such as

$$\begin{aligned} \eta_{m,x}(\cdot) &\xrightarrow{d} \eta_x(\cdot) := \overline{N}(\{\mathbf{y} : |\mathbf{y}| > x\} \times \cdot) \\ &= \sum_{i=1}^{\infty} \sum_{j \in \mathbb{Z}} \mathbb{1}(\Gamma_i^{-1/\alpha} |\mathbf{Q}_{ji}| > x) \varepsilon_{U_i}(\cdot), \end{aligned}$$

where the points  $(U_i)$  are iid uniformly distributed on  $(0, 1)$ ,  $(\Gamma_i)$  are the points of a standard homogeneous Poisson process, and  $(\mathbf{Q}_{\cdot i})$  are iid copies of the cluster process  $\mathbf{Q}$ . Using the independence among these three processes, one can easily rewrite the limit as

$$(4.30) \quad \eta_x((0, t]) := \sum_{i=1}^{N_x(t)} \xi_i, \quad 0 < t \leq 1,$$

where

- $N_x$  is a homogeneous Poisson process on  $(0, 1]$  with intensity  $x^{-\alpha}$ ,
- for an iid sequence  $(Y_i)$  of Pareto( $\alpha$ )-distributed random variables which is also independent of  $(\mathbf{Q}_i)$ ,

$$\xi_i := \sum_{j \in \mathbb{Z}} \mathbb{1}(Y_i |\mathbf{Q}_{ij}| > 1),$$

- $N_x, (\xi_i)$  are independent.

Relying on the point process of exceedances representation in (4.30), the random variables  $(\xi_i)$  can be interpreted as counts of serial exceedances from one cluster. Furthermore, we deduce the relation  $\mathbb{P}(\xi_1 > 0) = \mathbb{E}[\|\mathbf{Q}\|_{\infty}^{\alpha}] = \theta_{|\mathbf{x}|}$ , and also get an expression for the cluster size probabilities

$$(4.31) \quad \mathbb{P}(\xi_1 = j) = \mathbb{E}[|\mathbf{Q}|_{(j)}^{\alpha} - |\mathbf{Q}|_{(j+1)}^{\alpha}] = \pi_j, \quad j \geq 1.$$

The statistic  $\pi_j$  can be understood as the probability of recording a cluster of length  $j$ . The blocks estimators provide natural estimators of these quantities

$$(4.32) \quad \hat{\pi}_j := \frac{1}{k} \sum_{t=1}^m \frac{|\mathcal{B}_t|_{(j)}^{\hat{\alpha}} - |\mathcal{B}_t|_{(j+1)}^{\hat{\alpha}}}{\|\mathcal{B}_t\|_{\hat{\alpha}}^{\hat{\alpha}}} \mathbb{1}(\|\mathcal{B}_t\|_{\hat{\alpha}} > \|\mathcal{B}\|_{\hat{\alpha},(k+1)}),$$

$|\mathcal{B}_t|_{(1)} \geq |\mathcal{B}_t|_{(2)} \geq \dots \geq |\mathcal{B}_t|_{(m)}$  are the order statistics of  $\mathcal{B}_t$ , the  $t$ -th block.

**Corollary 4.5.** *Consider the function  $\pi_j : \tilde{\ell}^{\alpha} \rightarrow \mathbb{R}$  defined by  $\pi_j(\mathbf{x}) := (|\mathbf{x}|_{(j)}^{\alpha} - |\mathbf{x}|_{(j+1)}^{\alpha}) / \|\mathbf{x}\|_{\alpha}^{\alpha}$ , where  $|\mathbf{x}|_{(1)} \geq |\mathbf{x}|_{(2)} \geq \dots$ . Assume the conditions of Theorem 3.1 hold for  $p = \alpha$  and  $k/k' \rightarrow 0$ , as  $n \rightarrow \infty$ . Then, for all  $j \geq 1$  we have*

$$(4.33) \quad \sqrt{k}(\hat{\pi}_j - \pi_j) \xrightarrow{d} \mathcal{N}(0, \text{Var}(\pi_j^{\mathbf{Q}}(\mathbf{Q}))), \quad n \rightarrow \infty.$$

Corollary 4.5 provides a novel procedure for estimating cluster size probabilities based on extremal  $\ell^{\alpha}$ -blocks. As in the previous examples, the asymptotic variance can be computed as long as  $\mathbf{Q}$  is known. This allows

for comparison with the other cluster-based inference procedures provided in [27, 23, 49]. One advantage of our methodology is that we can straightforwardly infer the asymptotic variances of cluster sizes since we express them as cluster statistics in (4.33). Moreover, inference through extremal  $\ell^\alpha$ -blocks has already proven to be useful in [12] for fine-tuning the hyperparameters of the estimators, see also the discussion in Section 6.

As before, we can use the computations in Remark 4.3 to verify Condition **S**, it suffices to replace  $\|\mathbf{x}\|_\infty$  by  $(\|\mathbf{x}\|_{(j)} - \|\mathbf{x}\|_{(j+1)})$  in the equations therein.

## 5. MODELS

**5.1. Linear  $m_0$ -dependent sequences.** We consider  $(\mathbf{X}_t)$  to be a  $m_0$ -dependent time series with values in  $(\mathbb{R}^d, |\cdot|)$  satisfying  $\mathbf{RV}_\alpha$ .

**Example 5.1.** *The time series  $(\mathbf{X}_t)$  is a linear moving average of order  $m_0 \geq 1$  if it satisfies*

$$(5.34) \quad \mathbf{X}_t := \mathbf{Z}_t + \varphi_1 \mathbf{Z}_{t-1} + \cdots + \varphi_{m_0} \mathbf{Z}_{t-m_0}, \quad t \in \mathbb{Z},$$

with  $\mathbb{R}^d$ -variate iid innovations  $(\mathbf{Z}_t)$  satisfying  $\mathbf{RV}_\alpha$ , and  $(\varphi_j) \in \mathbb{R}^{m_0}$ .

Alternatively, the max-moving average of order  $m_0 \geq 1$  satisfies

$$(5.35) \quad X_t := \max\{Z_t, \varphi_1 Z_{t-1}, \dots, \varphi_{m_0} Z_{t-m_0}\}, \quad t \in \mathbb{Z},$$

with  $\mathbb{R}_+$ -variate iid innovations  $(Z_t)$  satisfying  $\mathbf{RV}_\alpha$ , and  $(\varphi_j) \in \mathbb{R}_+^{m_0}$ . Then both moving averages satisfy  $\mathbf{RV}_\alpha$  with  $|\mathbf{Q}|$  admitting the same deterministic expression  $(|\varphi_t|/||(\varphi_j)||_\alpha)$  in  $\tilde{\ell}^\alpha$ , see for instance Proposition 3.1. in [12] and Chapter 5 of [37].

Let  $p > \alpha/2$ . For all  $\kappa > 0$ , a sequence satisfying  $x_n = O(b_n^{\kappa+1/(p \wedge \alpha)})$  verifies **AC** and **CS<sub>p</sub>**. This is a consequence of Remark 2.3. Choosing  $(x_n)$  in this way implies there exist  $\kappa' > 0$ , and  $(k_n)$  satisfying

$$(5.36) \quad k_n = O(n b_n^{-\kappa' - \alpha/(p \wedge \alpha)}),$$

such that Equation (3.16) holds from an application of Potter's bound. Since  $\kappa$  can be chosen arbitrarily small,  $\kappa'$  can also be arbitrarily close to zero.

Keeping this in mind, we can state the Proposition below. The proof is postponed to Section C.

**Proposition 5.2.** *Consider  $(\mathbf{X}_t)$  to be an  $m_0$ -dependent time series with values in  $(\mathbb{R}^d, |\cdot|)$ . Consider  $p > \alpha/2$ , and sequence  $(k_n)$  and  $(k'_n)$  satisfying (5.36), such that  $k = k_n \rightarrow \infty$ ,  $m_n/k_n \rightarrow \infty$ , and  $k' = k'_n \rightarrow \infty$ ,  $k/k' \rightarrow 0$ . Consider  $f_\alpha(p) : \tilde{\ell}^p \rightarrow \mathbb{R}$ , and assume **L**, **S** hold. Then,*

$$\sqrt{k}(f_{\hat{\alpha}}^{\mathbf{Q}}(p) - f_\alpha^{\mathbf{Q}}(p)) \xrightarrow{d} \mathcal{N}(0, \text{Var}(f_\alpha(Y^{\mathbf{Q}}(p))))), \quad n \rightarrow \infty,$$

under the bias conditions  $\mathbf{B}_\alpha(k)$ ,  $\mathbf{B}(k')$ , and the result extends to  $\hat{\alpha}$ -cluster inference. In particular the  $\hat{\alpha}$ -cluster based estimators from Section 4 in (4.24) (4.27), and (4.32) are asymptotically normally distributed, and in the

case of the moving averages of Example 5.1

$$\sqrt{k}(\widehat{f_{\hat{\alpha}}^{\mathbf{Q}}}(\hat{\alpha}) - f_{\alpha}^{\mathbf{Q}}(\alpha)) \xrightarrow{\mathbb{P}} 0, \quad n \rightarrow \infty.$$

**5.2. Linear processes.** In this section we consider stationary linear processes  $(\mathbf{X}_t)$  with values in  $(\mathbb{R}^d, |\cdot|)$  satisfying  $\mathbf{RV}_{\alpha}$ .

**Example 5.3.** Consider  $(\mathbf{X}_t)$  to be an  $\mathbb{R}^d$ -variate sequence satisfying

$$(5.37) \quad \mathbf{X}_t = \sum_{j \in \mathbb{Z}} \varphi_j \mathbf{Z}_{t-j}, \quad t \in \mathbb{Z},$$

for a sequence of iid innovations  $(\mathbf{Z}_t)$  satisfying  $\mathbf{RV}_{\alpha}$ , and a sequence  $(\varphi_j)$  in  $\mathbb{R}^{\mathbb{Z}}$ . Moreover, assume there exists  $\kappa > 0$  such that  $\|(\varphi_j)\|_{(\alpha-\kappa) \wedge 2} < \infty$ .

In the setting of Example 5.3, a stationary solution  $(\mathbf{X}_t)$  exists and satisfies  $\mathbf{RV}_{\alpha}$  (c.f. [16, 41, 30]). Proposition 5.4 below demonstrates conditions  $\mathbf{AC}$ ,  $\mathbf{CS}_p$  hold for  $p > \alpha/2$ , and a suitable sequence  $(x_n)$  such that  $n\mathbb{P}(|\mathbf{X}_1| > x_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Therefore, the time series  $(\mathbf{X}_t)$  admits an  $\alpha$ -cluster process  $\mathbf{Q}$ , which we can compute in terms of the filter  $(\varphi_j)$ , and the spectral measure of the random variable  $\mathbf{Z}_0$ , denoted by  $\boldsymbol{\Theta}_0^{\mathbf{Z}}$ , with  $|\boldsymbol{\Theta}_0^{\mathbf{Z}}| = 1$  a.s. We obtain the expression, cf. Chapter 5 of [37],

$$(5.38) \quad \mathbf{Q} \stackrel{d}{=} (\varphi_t / \|(\varphi_j)\|_{\alpha}) \boldsymbol{\Theta}_0^{\mathbf{Z}}, \quad \in \tilde{\ell}^{\alpha}.$$

Note again that the norm of the  $\alpha$ -cluster process, i.e.,  $|\mathbf{Q}|$ , is deterministic in  $\tilde{\ell}^{\alpha}$ . Assuming  $\|(\varphi_j)\|_p < \infty$ , we can compute the indices  $c(p)$  in (2.8) by

$$(5.39) \quad c(p) = \mathbb{E}[\|\mathbf{Q}\|_p^{\alpha}] = \|(\varphi_j)\|_p^{\alpha} / \|(\varphi_j)\|_{\alpha}^{\alpha} < \infty.$$

Classic examples of these heavy-tailed linear models are auto-regressive moving averages, i.e., ARMA processes, with iid regularly varying noise; cf. [10].

The proposition below guarantees that the assumptions of Proposition 2.2 hold. We defer its proof to Section C.1.

**Proposition 5.4.** Consider  $(\mathbf{X}_t)$  to be a linear process with values in  $(\mathbb{R}^d, |\cdot|)$ , as in Example 5.3. Consider  $p > \alpha/2$ , and a sequence  $(x_n)$  such that  $n/x_n^{p \wedge (\alpha-\kappa)} \rightarrow 0$ ,  $n \rightarrow \infty$ , for some  $\kappa > 0$ . Then it holds for all  $\delta > 0$

$$(5.40) \quad \lim_{s \rightarrow \infty} \limsup_{n \rightarrow \infty} \frac{\mathbb{P}(\|\mathbf{X}_{[1,n]}/x_n - \mathbf{X}_{[1,n]}^{(s)}/x_n\|_p^p > \delta)}{n\mathbb{P}(|\mathbf{X}_1| > x_n)} = 0,$$

where  $\mathbf{X}_t^{(s)} := \sum_{|j| \leq s} \varphi_j \mathbf{Z}_{t-j}$ . Thus  $\mathbf{AC}$  and  $\mathbf{CS}_p$  are satisfied.

We now review the mixing properties of a linear process. We recall below the statement in Theorem 2.1. in [45] (see Lemma 15.3.1. in [37]).

**Proposition 5.5.** Consider  $(\mathbf{X}_t)$  to be a causal linear process with values in  $(\mathbb{R}^d, |\cdot|)$ , as in Example 5.3 with  $\varphi_j = 0$ , for  $j < 0$ . Assume the distribution of  $\mathbf{Z}_0$  is absolutely continuous with respect to the Lebesgue measure in  $\mathbb{R}^d$ , and has a density  $g_{\mathbf{Z}}$  satisfying

$$\text{i) } \int |g(\mathbf{x} - \mathbf{y}) - g(\mathbf{x})| d\mathbf{x} = O(|\mathbf{y}|), \text{ for all } \mathbf{y} \in \mathbb{R}^d,$$



- ii)  $\varphi_t = O(t^{-\rho})$ , for  $t \geq 0$ , and  $\rho > 2 + 1/\alpha$ ,
- iii)  $\sum_{j=0}^{\infty} \varphi_j \mathbf{x}^j \neq 0$ , for all  $\mathbf{x} \in \mathbb{R}^d$  with  $|\mathbf{x}| < 1$ ,

Then, for all  $0 < \varepsilon < \alpha$ , the mixing coefficients  $(\beta_t)$  satisfy

$$(5.41) \quad \beta_t = O\left(t^{1 - \frac{(\rho-1)(\alpha-\varepsilon)}{1+\alpha-\varepsilon}}\right).$$

Combining Propositions 5.4 and 5.5, we state below the asymptotic normality of the  $p$ -cluster based estimators for linear processes in Theorem 5.6. We defer its proof to Section C.2.

**Theorem 5.6.** *Consider  $(\mathbf{X}_t)$  to be a causal linear process with values in  $(\mathbb{R}^d, |\cdot|)$ , as in Example 5.3. Let  $\rho > 0$ , and assume the conditions of Proposition 5.5 hold with  $\varphi_t = O(t^{-\rho})$ , for  $t > 0$ . Consider  $p > \alpha/2$ , and sequence  $(k_n)$  and  $(k'_n)$  satisfying (5.36), such that  $k = k_n \rightarrow \infty$ ,  $m_n/k_n \rightarrow \infty$ , and  $k' = k'_n \rightarrow \infty$ ,  $k/k' \rightarrow 0$ . Consider  $f_\alpha(p) : \tilde{\ell}^p \rightarrow \mathbb{R}$ , and assume **L**, **S** hold. Assume that for  $\delta > 0$  as in (3.18),*

$$(5.42) \quad \rho > 3 + \frac{2}{\alpha} + \frac{2}{\delta}(1 + \frac{1}{\alpha}).$$

If  $f_\alpha(p)$  is bounded, condition (5.42) can be replaced by  $\rho > 3 + 2/\alpha$ . Then,

$$\sqrt{k}(f_{\hat{\alpha}}^{\widehat{\mathbf{Q}}}(p) - f_{\alpha}^{\mathbf{Q}}(p)) \xrightarrow{d} \mathcal{N}(0, \text{Var}(f_\alpha(Y^{\mathbf{Q}}(p)))) , \quad n \rightarrow \infty,$$

under the bias conditions  $\mathbf{B}_\alpha(k)$ ,  $\mathbf{B}(k')$ , and the result extends to  $\hat{\alpha}$ -cluster inference. In particular the  $\hat{\alpha}$ -cluster based estimators from Section 4 in (4.24) (4.27), and (4.32), satisfy

$$\sqrt{k}(f_{\hat{\alpha}}^{\widehat{\mathbf{Q}}}(\hat{\alpha}) - f_{\alpha}^{\mathbf{Q}}(\alpha)) \xrightarrow{\mathbb{P}} 0, \quad n \rightarrow \infty.$$

Regarding cluster inference in the case of linear models, the  $\alpha$ -cluster approach has an optimal asymptotic variance for shift-invariant functionals since we use the  $\ell^\alpha$ -norm order statistics. For this reason, it compares favourably with state-of-the-art blocks estimator. For example, for the extremal index, the super-efficient estimator in (4.24) has a lower asymptotic variance than the blocks estimator in (4.25). Indeed the asymptotic variance  $\sigma_\theta^2$  of the latter, computed in (4.26), is not necessarily null. For example, for the autoregressive process of order one AR(1) one has  $\sigma_\theta^2 = 1 - \theta_{|\mathbf{X}|} > 0$ .

**5.3. Affine stochastic recurrence equation solution under Kesten's conditions.** In this section we focus on the causal solution to the affine stochastic recurrence equation SRE under Kesten's conditions. To guarantee the existence of a solution  $(\mathbf{X}_t)$ , with values in  $(\mathbb{R}^d, |\cdot|)$  as in (5.43) satisfying **RV** $_\alpha$ , we rely on Theorem 2.1. and Theorem 2.4 in [2]. For an overview, we refer to [11]. In what follows, we study time series  $(\mathbf{X}_t)$  as in the Example 5.7 below.

**Example 5.7.** *Consider  $(\mathbf{X}_t)$  to be a sequence with values in  $\mathbb{R}^d$  satisfying*

$$(5.43) \quad \mathbf{X}_t = \mathbf{A}_t \mathbf{X}_{t-1} + \mathbf{B}_t, \quad t \in \mathbb{Z},$$

where  $((\mathbf{A}_t, \mathbf{B}_t))$  is an iid sequence of non-negative random  $d \times d$  matrices with generic element  $\mathbf{A}$ , and non-negative random vectors with generic element  $\mathbf{B}$  taking values in  $\mathbb{R}^d$ . For the existence of a causal stationary solution, we assume

- i)  $\mathbb{E}[\log^+ |\mathbf{A}|_{op}] + \mathbb{E}[\log^+ |\mathbf{B}|] < \infty$ ,
- ii) under i), assume the Lyapunov exponent of  $(\mathbf{A}_t)$ , denoted  $\gamma$ , satisfies

$$\gamma := \lim_{n \rightarrow \infty} n^{-1} \log |\mathbf{A}_t \cdots \mathbf{A}_1|_{op} < 0, \quad \text{a.s.}$$

To guarantee the heavy-tailedness condition  $\mathbf{RV}_\alpha$ , we also assume

- iii)  $\mathbf{B} \neq \mathbf{0}$  a.s., and  $\mathbf{A}$  has no zero rows a.s.
- iv) there exists  $\kappa > 0$  such that  $\mathbb{E}[|\mathbf{A}|_{op}^\kappa] < 1$ ,
- v) the set  $\Gamma$  from Equation (5.44) generates a dense group on  $\mathbb{R}$ ,

$$(5.44) \quad \Gamma = \{ \log |\mathbf{a}_n \cdots \mathbf{a}_1|_{op} : n \geq 1, \mathbf{a}_n \cdots \mathbf{a}_1 > 0, \\ \mathbf{a}_n, \dots, \mathbf{a}_1 \text{ are in the support of } \mathbf{A}'\text{s distribution} \},$$

- vi) there exists  $\kappa_1 > 0$  such that  $\mathbb{E}[(\min_{i=1, \dots, d} \sum_{j=1}^d A_{ij})^{\kappa_1}] \geq d^{\kappa_1/2}$ , and  $\mathbb{E}[|\mathbf{A}|_{op}^{\kappa_1} \log^+ |\mathbf{A}|_{op}] < \infty$ .
- vii) under i) – vi), there exists a unique  $\alpha > 0$  such that

$$(5.45) \quad \lim_{n \rightarrow \infty} n^{-1} \log \mathbb{E}[|\mathbf{A}_n \cdots \mathbf{A}_1|_{op}^\alpha] = 0,$$

and  $\mathbb{E}[|\mathbf{B}|^\alpha] < \infty$ . If  $d > 1$  assume  $\alpha$  is not an even integer.

The  $\mathbb{R}^d$ -variate series  $(\mathbf{X}_t)$ , satisfying (5.43) and i) – vii), admits a causal stationary solution and satisfies  $\mathbf{RV}_\alpha$ , with  $\alpha > 0$  as in Equation (5.45).

The previous example is motivated by the seminal Kesten's paper [35]. We follow Theorem 2.1. in [2] to state conditions i) – ii) of Example 5.7. Under the conditions i) – ii), the unique solution  $(\mathbf{X}_t)$  of (5.43) has the a.s. causal representation

$$(5.46) \quad \mathbf{X}_t = \sum_{i \geq 0} \mathbf{A}_{t-i+1} \cdots \mathbf{A}_t \mathbf{B}_{t-i}, \quad t \in \mathbb{Z},$$

where the first summand is  $\mathbf{B}_t$  for  $i = 0$  by convention; for an overview see [11].

One of the main reasons why the solutions to SRE as in Example 5.7 have received strong interest, is because  $(\mathbf{X}_t)$  satisfies  $\mathbf{RV}_\alpha$  even when the innovations  $((\mathbf{A}_t, \mathbf{B}_t))$  are light-tailed. This feature was first noticed in [35] where the original Kesten's assumptions were introduced. In Kesten's framework, a causal stationary solution to the SRE exists as in (5.46), and the extremes of the series occur due to the sums of infinitely many terms of growing length products appearing in (5.46); see [6] for a review. Further, the community adopted the simplified Kesten's conditions stated by Goldie in [24] for univariate SRE. These conditions also aim to capture the heavy-tailed feature under lighter-tailed innovations. In Example 5.7, we borrow the conditions iii) – vii) established for the multivariate setting from Theorem 2.4 and Corollary 2.7. in [2]. Then, a solution  $(\mathbf{X}_t)$  as in Example 5.7 satisfies  $\mathbf{RV}_\alpha$ ,

for  $\alpha > 0$ , and the index of regular variation  $\alpha$  is the unique solution to the Equation (5.45). We are also interested in Example 5.7 because it models classic econometric time series such as the squared ARCH( $p$ ), and the volatility of GARCH( $p, q$ ) processes; see [11].

Concerning the extremes of  $(\mathbf{X}_t)$  in Example 5.7, the forward spectral tail process satisfies the relation

$$\Theta_t = \mathbf{A}_t \cdots \mathbf{A}_1 \Theta_0, \quad t \geq 0,$$

where  $(\mathbf{A}_t)$  is an iid sequence distributed as  $\mathbf{A}$ ; see [34]. The backward spectral tail process has a cumbersome representation that we omit here; c.f. [34]. We state in Proposition 5.8 sufficient conditions on  $(\mathbf{A}, \mathbf{B})$  yielding assumptions  $\mathbf{AC}$ ,  $\mathbf{CS}_p$  hold for  $p > \alpha/2$ , and a suitable sequence  $(x_n)$  such that  $n\mathbb{P}(\|\mathbf{X}_1\| > x_n) \rightarrow 0$  as  $n \rightarrow \infty$ . In this case the time series  $(\mathbf{X}_t)$  admits an  $\alpha$ -cluster process  $\mathbf{Q}$ . We recall the identity from Equation (8.6) of [12]:  $c(p) = \mathbb{E}[\|\mathbf{Q}\|_p^\alpha] = \mathbb{E}[\|(\Theta_t)_{t \geq 0}\|_p^\alpha - \|(\Theta_t)_{t \geq 1}\|_p^\alpha]$ , for  $c(p)$  as in (2.8). Then, letting  $p = \alpha/2$ , a straightforward computation yields

$$\begin{aligned} c(p) &\leq 2 \mathbb{E}[\|(\Theta_t)_{t \geq 0}\|_p^{\alpha-p}] = 2 \mathbb{E}[\sum_{t \geq 0} |\mathbf{A}_t \cdots \mathbf{A}_1 \Theta_0|^p] \\ &\leq 2s \sum_{t \geq 0} (\mathbb{E}[\|\mathbf{A}_s \cdots \mathbf{A}_1\|_{op}^p])^t, \end{aligned}$$

and  $\mathbb{E}[\|\mathbf{A}_s \cdots \mathbf{A}_1\|_{op}^p] < 1$ , for  $p < \alpha$  and  $s \geq 1$  fixed sufficiently large in the setting of Example 5.7. Hence, for  $p \in (\alpha/2, \alpha)$ ,  $c(p) < \infty$  in (2.8), and then the series admits a  $p$ -cluster process  $\mathbf{Q}^{(p)}$ .

We state now Proposition 5.8 which verifies conditions  $\mathbf{AC}$ ,  $\mathbf{CS}_p$  for the SRE equation. The proof is postponed to Section C.3.

**Proposition 5.8.** *Let  $(\mathbf{X}_t)$  be a stationary time series with values in  $(\mathbb{R}^d, |\cdot|)$ , as in Example 5.7. Let  $p > \alpha/2$ , and consider  $(x_n)$  such that there exists  $\kappa > 0$  satisfying  $n/x_n^{p \wedge (\alpha - \kappa)} \rightarrow 0$ , as  $n \rightarrow \infty$ . Then,  $(x_n)$  satisfies conditions  $\mathbf{AC}$  and  $\mathbf{CS}_p$ .*

In the setting of SRE equations, condition  $\mathbf{AC}$  has been shown in Theorem 4.17 in [42]. In [42], the authors already considered a condition similar to  $\mathbf{CS}_p$ . Parallel to their setting, we propose a proof of Proposition 5.8 which shows  $\mathbf{CS}_p$  holds over uniform regions  $\Lambda_n = (x_n, \infty)$  such that  $n/x_n^p \rightarrow 0$ , as  $n \rightarrow \infty$ , in the sense of (C.86). Thereby, our proof extends Theorem 4.17 in [42] to uniform regions  $\Lambda_n$  not having an upper bound.

Concerning the mixing properties of  $(\mathbf{X}_t)_{t \geq 0}$  as in Example 5.7, we use that it is a Markov chain and that  $\mathbf{X}_0$  has the stationary distribution. As mentioned in Remark 2.1, we can then use Markov chain's theory to compute its mixing coefficients; cf. [40]. We review Theorem 2.8. in [2], yielding an exponential decay of the mixing-coefficients  $(\beta_t)$  of the series. For a general treatment see Chapter 4.2 in [11].

**Proposition 5.9.** *Consider a time series  $(\mathbf{X}_t)$  with values in  $(\mathbb{R}^d, |\cdot|)$ , as in Example 5.7. Assume there exists a Borel measure  $\mu$  on  $(\mathbb{R}^d, |\cdot|)$ , such that the Markov chain  $(\mathbf{X}_t)_{t \geq 0}$  is  $\mu$ -irreducible, i.e., for all  $C \subset \mathbb{R}^d$  with*

$\mu(C) > 0$ ,

$$(5.47) \quad \sum_{t=0}^{\infty} \mathbb{P}(\mathbf{X}_t \in C \mid \mathbf{X}_0 = \mathbf{x}) > 0, \quad \mathbf{x} \in \mathbb{R}.$$

Then  $(\mathbf{X}_t)$  has mixing coefficients  $(\beta_t)$  satisfying  $\beta_t = O(\rho^t)$  for some  $\rho \in (0, 1)$ , and we say it is strongly mixing with geometric rate. Moreover,  $(\mathbf{X}_t)_{t \geq 0}$  is irreducible with respect to the Lebesgue measure if  $(\mathbf{A}, \mathbf{B})$  admits a density.

We can now state the asymptotic normality of cluster-based estimator for SRE solutions in Theorem 5.10 below. The proof is postponed to Section C.4.

**Theorem 5.10.** *Consider  $(\mathbf{X}_t)$  to be the causal solution to the SRE in (5.43) with values in  $(\mathbb{R}^d, |\cdot|)$ , as in Example 5.7. Assume the conditions of Proposition 5.9 hold. Consider  $p > \alpha/2$ , and sequence  $(k_n)$  and  $(k'_n)$  satisfying (5.36), such that  $k = k_n \rightarrow \infty$ ,  $m_n/k_n \rightarrow \infty$ , and  $k' = k'_n \rightarrow \infty$ ,  $k/k' \rightarrow 0$ . Consider  $f_\alpha(p) : \tilde{\ell}^p \rightarrow \mathbb{R}$ , and assume  $\mathbf{L}, \mathbf{S}$  hold. Then,*

$$\sqrt{k}(f_{\hat{\alpha}}^{\mathbf{Q}}(p) - f_{\alpha}^{\mathbf{Q}}(p)) \xrightarrow{d} \mathcal{N}(0, \text{Var}(f_{\alpha}(Y\mathbf{Q}^{(p)}))), \quad n \rightarrow \infty,$$

under the bias conditions  $\mathbf{B}_\alpha(k)$ ,  $\mathbf{B}(k')$ , and the result extends to  $\hat{\alpha}$ -cluster inference. In particular, the  $\hat{\alpha}$ -cluster based estimators from Section 4 in (4.24) (4.27), and (4.32), are asymptotically normally distributed.

**Remark 5.11.** In this example, the asymptotic variances of the  $\alpha$ -cluster based estimators from Section 4 in (4.24) (4.27), and (4.32) are non-null. The limiting variances in Theorem 5.10 are difficult to compare with the existing ones in the literature because of the complexity of the distribution of  $\mathbf{Q}^{(p)}$ . However, we provide simple  $\ell^\alpha$ -block estimators of the asymptotic variances in Section 6.

## 6. NUMERICAL EXPERIMENTS

This section aims to illustrate the finite-sample performance of the  $\hat{\alpha}$ -cluster estimators on time series  $(\mathbf{X}_t)$  with tail-index  $\alpha > 0$ . In all the models we consider in Section 5, we work under the assumption that the tuning parameters of the  $\alpha$ -cluster satisfy (5.36). We take  $\kappa' = 1$  in (5.36) which yields  $b = \sqrt{n/k}$ . In this case, the implementation of our estimators can be written solely as a function of  $k$  and  $k'$ . Recall  $k = k_n$  must satisfy  $k \rightarrow \infty$ ,  $m/k \rightarrow \infty$  with  $m = \lfloor n/b \rfloor$ ,  $n/k' \rightarrow \infty$ , and  $k/k' \rightarrow 0$ . Numerical comparisons of our  $\hat{\alpha}$ -cluster based approach with other existing estimators for the extremal index and the cluster index are at the advantage of our approach; see [13] and [12]. The code of all numerical experiments is available at: [https://github.com/GBuritica/cluster\\_functionals.git](https://github.com/GBuritica/cluster_functionals.git).

**6.1. Tuning the Hill estimator.** We recommend choosing the tuning sequence of the tail-index and of the cluster estimators as  $(k'_n)$ ,  $(k_n)$ , respectively, such that  $k/k' \rightarrow 0$ . Roughly speaking, the cluster statistics capture