



# Large deviations of lp-blocks of regularly varying time series and applications to cluster inference

Gloria Buritic Á, Thomas Mikosch, Olivier Wintenberger

## ► To cite this version:

Gloria Buritic Á, Thomas Mikosch, Olivier Wintenberger. Large deviations of lp-blocks of regularly varying time series and applications to cluster inference. 2022. hal-03269176v4

**HAL Id: hal-03269176**

**<https://hal.science/hal-03269176v4>**

Preprint submitted on 20 Jul 2022

**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.

July 20, 2022

# LARGE DEVIATIONS OF $\ell^p$ -BLOCKS OF REGULARLY VARYING TIME SERIES AND APPLICATIONS TO CLUSTER INFERENCE

GLORIA BURITICÁ, THOMAS MIKOSCH, AND OLIVIER WINTENBERGER

**ABSTRACT.** In the regularly varying time series setting, a cluster of exceedances is a short period for which the supremum norm exceeds a high threshold. We propose to study a generalization of this notion considering short periods, or blocks, with  $\ell^p$ -norm above a high threshold. Our main result derives new large deviation principles of extremal  $\ell^p$ -blocks, which guide us to define and characterize spectral cluster processes in  $\ell^p$ . We then study cluster inference in  $\ell^p$  to motivate our results. We design consistent disjoint blocks estimators to infer features of cluster processes. Our estimators promote the use of large empirical quantiles from the  $\ell^p$ -norm of blocks as threshold levels which eases implementation and also facilitates comparison for different  $p > 0$ . Our approach highlights the advantages of cluster inference based on extremal  $\ell^\alpha$ -blocks, where  $\alpha > 0$  is the index of regular variation of the series. We focus on inferring important indices in extreme value theory, e.g., the *extremal index*.

## 1. INTRODUCTION

For various applications of extreme value statistics with stationary time series, it is natural to wonder how a recorded high level can affect the future behavior of the sequence or how time dependencies perturb inferential methodologies. For example, for high quantile marginal estimation it is well known that the inference procedures tailored for independent observations are disturbed by temporal dependencies and must be corrected to produce accurate estimates; cf. Leadbetter [24], Embrechts *et al.* [18]. We aim to model extremal time dependencies in the setting of  $\mathbb{R}^d$ -valued stationary regularly varying time series  $(\mathbf{X}_t)_{t \in \mathbb{Z}}$  with generic element  $\mathbf{X}_0$ ; see Section 2.1 for a definition, cf. Basrak and Segers [4]. In this framework, an exceedance of a high threshold by the norm  $|\mathbf{X}_t|$  at time  $t$  might trigger consecutive exceedances in some small time interval around  $t$ . These short periods with

---

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

*Key words and phrases.* Regularly varying time series, large deviation principles, cluster processes, extremal index.

Thomas Mikosch's research is partially supported by Danmarks Frie Forskningsfond Grant No 9040-00086B. Olivier Wintenberger would like to thank Riccardo Passeggeri for useful discussions on the topic. Gloria Buriticá and Olivier Wintenberger would like to acknowledge the support of the French Agence Nationale de la Recherche (ANR) under reference ANR20-CE40-0025-01 (T-REX project).

at least one exceedance were introduced implicitly in the seminal paper by Davis and Hsing [11]. We refer to them as *clusters (of exceedances)*. They were further reviewed in Basrak and Segers [4] and Basrak *et al.* [3].

The main motivation for studying clusters (of exceedances) can be traced back to Theorem 2.5 in Davis and Hsing [11]. For weakly dependent regularly varying time series, the point process with atoms at  $a_n^{-1}\mathbf{X}_{[0,n]} = a_n^{-1}(\mathbf{X}_0, \dots, \mathbf{X}_n)$  admits a limit distribution that can be characterized in terms of three features: the index of regular variation, the distribution of cluster (of exceedances), and the *extremal index* of  $(|\mathbf{X}_t|)$ , denoted by  $\theta_{|\mathbf{X}|}$ , where  $a_n$  are moderate threshold levels satisfying  $n\mathbb{P}(|\mathbf{X}_0| > a_n) \rightarrow 1$  as  $n \rightarrow \infty$ . In this setting, clusters are modeled as rare events of  $\mathbf{X}_{[0,n]}$  when its supremum norm exceeds the high level  $x_n$  such that  $\mathbb{P}(\|\mathbf{X}_{[0,n]}\|_\infty > x_n) \sim \theta_{|\mathbf{X}|} n \mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0$ . From the last relation we also see that  $\theta_{|\mathbf{X}|}$  arises when comparing the extremal behavior of blocks of maxima in  $(\mathbf{X}_t)$  with the corresponding behavior of the blocks in an iid sequence  $(\mathbf{X}'_t)$  with the same marginal distribution. In particular  $\theta_{|\mathbf{X}|}$  describes how the blocks of maxima reach high levels compared to the iid setting.

In view of the previous discussion a cluster (of exceedances) is tied together with the extremal index by the supremum norm. Our main theoretical result extends the aforementioned ideas from the  $\ell^\infty$ -norm to  $\ell^p$ -norms for  $p < \infty$ . In Theorem 2.1 we investigate the behavior of  $\mathbf{X}_{[0,n]}$  when its  $\ell^p$ -norm exceeds high levels  $(x_n)$  satisfying  $\mathbb{P}(\|\mathbf{X}_{[0,n]}\|_p > x_n) = \mathbb{P}(\sum_{t=1}^n |\mathbf{X}_t|^p > x_n^p) \rightarrow 0$  as  $n \rightarrow \infty$ . We call this a large deviation result since it describes the probability that the partial sums  $\|\mathbf{X}_{[0,n]}\|_p^p$  exceed the extreme threshold  $x_n^p$ . This leads us to a new definition of a *cluster process* in the space  $\ell^p = \ell^p(\mathbb{R}^d)$  and, in the limiting case  $p = \infty$ , one recovers the classical clusters (of exceedances). Similarly, large deviation principles for sums were considered by Nagaev [31], Cline and Hsing [9] in the independent heavy-tailed case, and by Mikosch and Wintenberger [27, 28, 29], Mikosch and Rodionov [30] in the dependent heavy-tailed case. We extend large deviation principles to  $\ell^p$ -norms  $\|\mathbf{X}_{[0,n]}\|_p$ , and extremal  $\ell^p$ -blocks, i.e., blocks  $\mathbf{X}_{[0,n]}$  with large  $\ell^p$ -norm.

We apply our findings to cluster inference. For this purpose we divide the sample  $\mathbf{X}_1, \dots, \mathbf{X}_n$  into disjoint blocks  $(\mathbf{B}_t)_{1 \leq t \leq \lfloor n/b_n \rfloor}$ ,  $\mathbf{B}_t := \mathbf{X}_{(t-1)b_n+1:b_n}$ , for a sequence of block lengths  $(b_n)$  such that  $b_n \rightarrow \infty$  and  $b_n/n \rightarrow 0$ . Then we select blocks whose  $\ell^p$ -norms exceed a high threshold  $x_{b_n}$ . Our goal is to infer features of the cluster processes from these extremal  $\ell^p$ -blocks. In Theorem 4.2 we design consistent disjoint blocks methods with thresholds chosen as order statistics of  $\ell^p$ -norms. Hereby we must choose a number  $k_n = k_n(p)$  of blocks with large  $\ell^p$ -norm such that  $n/(b_n k_n) \rightarrow \infty$ ,  $k_n \rightarrow \infty$ . The sequence  $(k_n)$  appeals to the classical bias-variance trade-off in extreme value statistics; see for example Resnick [33]. When choosing a small number of blocks  $k_n$  for inference the variance of the estimates increases while a large number  $k_n$  leads to strong bias. This calls for a rigorous

definition of extremal  $\ell^p$ -blocks with the goal of revealing how  $p$  plays a key role for tuning the sequence  $k_n = k_n(p)$ . Moreover, we can derive the same quantity by using extremal  $\ell^p$ -blocks for different values of  $p$  if we apply a change-of-norm technique. Our large deviations result allows us then to compare the different estimators through the tuning parameter  $k_n(p)$ . The key argument of our analysis is the relationship we stress between the sequence  $k_n = k_n(p)$  and the large deviations of  $\ell^p$ -norms.

One advantage of using empirical  $\ell^p$ -norm thresholds is that they adapt to the block lengths  $(b_n)$  and take into account the value of  $p$ . In the existing literature for  $p = \infty$  no detailed advice is given as to how  $(b_n)$  and  $(x_{b_n})$  must be chosen; see for example Drees, Rootzén [16], Drees, Neblung [15], Cissokho, Kulik [8], Drees *et al.* [14], who assume growth conditions on the sequence  $(x_n)$  such as  $n\mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0$  as  $n \rightarrow \infty$ . It is common practice to replace  $x_{b_n}$  by an upper order statistic of  $(|\mathbf{X}_t|)_{1 \leq t \leq n}$ ; see for example the blocks estimator of the *extremal index* proposed by Hsing [20]. In our setting, the order statistics of the  $\ell^p$ -norms adapt naturally to different values of  $p$ . Asymptotic normality of our estimators could be derived by combining arguments from Theorem 4.3 in Cissokho and Kulik [8] and the large deviation arguments developed below; this topic is the subject of ongoing work and will not be presented here.

The case when  $p$  and the index  $\alpha$  of regular variation of  $(\mathbf{X}_t)$  coincide is rather specific. The relation  $\mathbb{P}(\|\mathbf{X}_{[0,n]}\|_\alpha > x_n) \sim n\mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0$  indicates that serial dependence does not affect large deviations of the  $\ell^\alpha$ -norm. From this relation we see that the  $\ell^\alpha$ -norm of the series reaches high levels at the same rate as in the iid case. Consequently, when  $p$  coincides with the index  $\alpha$ , the temporal dependencies of the sequence do not perturb the number  $k_n = k_n(\alpha)$  of extremal  $\ell^\alpha$ -blocks we can consider for inference. In practice, this fact might ease tuning the parameter  $k_n$ . Hereby we focus on inferring classical indices of serial dependence based on extremal  $\ell^\alpha$ -blocks. We apply our inference procedure to estimate the extremal index using extremal  $\ell^\alpha$ -blocks. We also consider inference of *cluster indices* as defined by Mikosch and Wintenberger [28] on partial sum functionals by considering extremal  $\ell^\alpha$ -blocks. Our simulation study supports the fact that  $\ell^\alpha$ -cluster inference is robust regarding the number  $k_n$  of extremal  $\ell^\alpha$ -blocks we can choose.

The previous indices are based on functionals that are shift-invariant with respect to the backward shift in sequence spaces; see Kulik and Soulier [23] for details. We extend cluster inference to functionals acting on  $\ell^p$  by studying  $\alpha$ th-power sum functionals acting on  $\ell^p$ . The key argument for this extension is the random shift analysis of Janssen [22] expressed in terms of the  $\alpha$ th moment of the cluster process. A similar idea has been investigated in Drees *et al.* [17] and Davis *et al.* [10] for inference of the tail process. Here we focus on cluster inference.

**1.1. Outline of the paper.** In Section 2, after introducing preliminaries on regular variation, we present the main large deviation principle (Theorem 2.1). In Section 3 we study  $\ell^p$ -valued cluster processes which were introduced in Theorem 2.1. We apply this theorem in Section 4 where we deal with inference for shift-invariant functionals acting on these cluster processes (see Theorem 4.2), choosing thresholds as empirical quantiles of the  $\ell^p$ -norms of blocks. We continue with an in-depth analysis of the assumptions of Theorem 2.1; see Section 5. In Section 6 we consider inference for non-shift-invariant functionals. We also illustrate our approach of  $\ell^p$ -based cluster inference for  $p = \alpha$  and compare it with the case  $p = \infty$ ; see Section 7. We defer all proofs to Section 8.

**1.2. Notation.** For integers  $i$  and  $a < b$  we write  $i + [a, b] = \{i + a, \dots, i + b\}$ . It is convenient to embed the vectors  $\mathbf{x}_{[a,b]} \in \mathbb{R}^{d(b-a+1)}$  in  $(\mathbb{R}^d)^\mathbb{Z}$  by assigning zeros to indices  $i \notin [a, b]$ , and we then also write  $\mathbf{x}_{[a,b]} \in (\mathbb{R}^d)^\mathbb{Z}$ . We write  $\mathbf{x} := (\mathbf{x}_t) = (\mathbf{x}_t)_{t \in \mathbb{Z}}$ , and define truncation at level  $\varepsilon > 0$  from above and below by  $\underline{\mathbf{x}}_\varepsilon = (\underline{\mathbf{x}}_{t_\varepsilon})_{t \in \mathbb{Z}}$ ,  $\overline{\mathbf{x}}^\varepsilon = (\overline{\mathbf{x}}_{t^\varepsilon})_{t \in \mathbb{Z}}$ , where  $\underline{\mathbf{x}}_{t_\varepsilon} = \mathbf{x}_t \mathbb{1}(|\mathbf{x}_t| > \varepsilon)$ ,  $\overline{\mathbf{x}}_{t^\varepsilon} = \mathbf{x}_t \mathbb{1}(|\mathbf{x}_t| \leq \varepsilon)$ .

We focus on the sequence space  $\ell^p$ ,  $p \in (0, \infty]$  equipped with the metric

$$d_p(\mathbf{x}, \mathbf{y}) := \begin{cases} \|\mathbf{x} - \mathbf{y}\|_p = \left( \sum_{t \in \mathbb{Z}} |\mathbf{x}_t - \mathbf{y}_t|^p \right)^{1/p}, & p \in (1, \infty), \\ \|\mathbf{x} - \mathbf{y}\|_p^p, & p \in (0, 1), \end{cases} \quad \mathbf{x}, \mathbf{y} \in \ell^p,$$

and the supremum distance in the case  $p = \infty$ . We know that  $d_p$  makes  $\ell^p$  a separable Banach space for  $p \in (1, \infty]$ , and a separable complete metric space for  $p \in (0, 1)$ . Recall the *backshift operator* acting on  $\mathbf{x} \in (\mathbb{R}^d)^\mathbb{Z}$ :  $B^k \mathbf{x} = (\mathbf{x}_{t-k})_{t \in \mathbb{Z}}$ ,  $k \in \mathbb{Z}$ . Then we define the shift-invariant space  $\tilde{\ell}^p = \ell^p / \sim$  as the quotient space with respect to the equivalence relation  $\sim$  in  $\ell^p$ :  $\mathbf{x} \sim \mathbf{y}$  if there exists  $k \in \mathbb{Z}$  such that  $B^k \mathbf{x} = \mathbf{y}$ . An element of  $\tilde{\ell}^p$  is denoted by  $[\mathbf{x}] = \{B^k \mathbf{x} : k \in \mathbb{Z}\}$ . For ease of notation, we often write  $\mathbf{x}$  instead of  $[\mathbf{x}]$ , and we notice that any element in  $\ell^p$  can be embedded in  $\tilde{\ell}^p$  by using the equivalence relation. We define for  $[\mathbf{x}], [\mathbf{y}] \in \tilde{\ell}^p$ ,

$$\tilde{d}_p([\mathbf{x}], [\mathbf{y}]) := \inf_{k \in \mathbb{Z}} \{d_p(B^k \mathbf{a}, \mathbf{b}) : \mathbf{a} \in [\mathbf{x}], \mathbf{b} \in [\mathbf{y}]\}.$$

For  $p \in (0, \infty]$ ,  $\tilde{d}_p$  is a metric on  $\tilde{\ell}^p$  and turns it into a complete metric space; see Basrak *et al.* [3].

## 2. PRELIMINARIES AND MAIN RESULT

**2.1. About regular variation of time series.** We consider an  $\mathbb{R}^d$ -valued stationary process  $(\mathbf{X}_t)$ . Following Davis and Hsing [11], we call it regularly varying if the finite-dimensional distributions of the process are regularly varying. This notion involves the vague convergence of certain tail measures; see Resnick [33]. Avoiding the concept of vague convergence and infinite

limit measures, Basrak and Segers [4] showed that regular variation of  $(\mathbf{X}_t)$  is equivalent to the weak convergence relations: for every  $h \geq 0$ ,

$$\mathbb{P}(x^{-1}(\mathbf{X}_t)_{[-h,h]} \in \cdot \mid |\mathbf{X}_0| > x) \xrightarrow{w} \mathbb{P}(Y(\boldsymbol{\Theta}_t)_{[-h,h]} \in \cdot), \quad x \rightarrow \infty, \quad (2.1)$$

where  $Y$  is Pareto( $\alpha$ )-distributed, i.e., it has tail  $\mathbb{P}(Y > y) = y^{-\alpha}$ ,  $y > 1$ , independent of the vector  $(\boldsymbol{\Theta}_t)_{[-h,h]}$  in  $(\mathbb{R}^d)^{2h+1}$  and  $|\boldsymbol{\Theta}_0| = 1$ . According to Kolmogorov's consistency theorem, one can extend the latter finite-dimensional vectors to a sequence  $\boldsymbol{\Theta} = (\boldsymbol{\Theta}_t)_{t \in \mathbb{Z}}$  in  $(\mathbb{R}^d)^{\mathbb{Z}}$  called the *spectral tail process* of  $(\mathbf{X}_t)$ .

Following Planinić and Soulier [32], the spectral tail process  $(\boldsymbol{\Theta}_t)$  satisfies the *time-change formula*: for every measurable function  $f : (\ell^p, d_p) \rightarrow \mathbb{R}$  such that  $f(\lambda \mathbf{x}) = f(\mathbf{x})$  for all  $\lambda > 0$ , we have for all  $t, s \in \mathbb{Z}$ ,

$$(2.2) \quad \mathbb{E}[f(B^s(\boldsymbol{\Theta}_t))\mathbb{1}(\boldsymbol{\Theta}_{-s} \neq \mathbf{0})] = \mathbb{E}[|\boldsymbol{\Theta}_s|^\alpha f((\boldsymbol{\Theta}_t))].$$

The regular variation property of  $(\mathbf{X}_t)$ , denoted by  $\mathbf{RV}_\alpha$ , is determined by the (tail)-index  $\alpha > 0$  and the spectral tail process.

Furthermore, Segers *et al.* [34] characterized regular variation of random elements with values in star-shaped metric spaces. Their results are based on weak convergence in the spirit of (2.1). Our focus will be on a special star-shaped space: the sequence space  $(\ell^p, d_p)$ . Using the *p-modulus function*  $\|\cdot\|_p$ , the  $\ell^p$ -valued stationary process  $(\mathbf{X}_t)$  has the property  $\mathbf{RV}_\alpha$  if and only if relation (2.1) holds with  $|\mathbf{X}_0|$  replaced by  $\|\mathbf{X}_{[0,h]}\|_p$ . Equivalently, (Proposition 3.1 in Segers *et al.* [34]), for every  $h \geq 0$ ,

$$(2.3) \quad \mathbb{P}(x^{-1}\mathbf{X}_{[0,h]} \in \cdot \mid \|\mathbf{X}_{[0,h]}\|_p > x) \xrightarrow{w} \mathbb{P}(Y\mathbf{Q}^{(p)}(h) \in \cdot), \quad x \rightarrow \infty,$$

the Pareto( $\alpha$ ) variable  $Y$  is independent of  $\mathbf{Q}^{(p)}(h) \in \mathbb{R}^{d(h+1)}$ , such that  $\|\mathbf{Q}^{(p)}(h)\|_p = 1$  a.s. We call  $\mathbf{Q}^{(p)}(h)$  the *spectral component* of  $\mathbf{X}_{[0,h]}$  in  $\ell^p$ .

**2.2. Main result.** We start by giving our main result on large deviations of the sequence  $\mathbf{X}_{[0,n]}$ , that we embed in the space  $(\tilde{\ell}^p, \tilde{d}_p)$ . The proof is postponed to Section 8.1.

**Theorem 2.1.** *Consider an  $\mathbb{R}^d$ -valued stationary time series  $(\mathbf{X}_t)$  satisfying  $\mathbf{RV}_\alpha$  for some  $\alpha > 0$ . For a given  $p > 0$ , assume that there exists a sequence  $(x_n)$  such that  $n\mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0$  as  $n \rightarrow \infty$ . Furthermore, assume that for every  $\delta > 0$ ,*

$$\mathbf{AC} : \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_{[k,n]}\|_\infty > \delta x_n \mid |\mathbf{X}_0| > \delta x_n) = 0,$$

$$\mathbf{CS}_p : \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{\mathbb{P}(\overline{\|x_n^{-1}\mathbf{X}_{[0,n]}\|_p}^\epsilon > \delta)}{n\mathbb{P}(|\mathbf{X}_0| > x_n)} = 0,$$

$n/x_n^p \rightarrow 0$  if  $p < \alpha$ , and there exists  $\kappa > 0$  such that  $n/x_n^{\alpha-\kappa} \rightarrow 0$  if  $p = \alpha$ . Then, there exists  $c(p) > 0$  such that

$$(2.4) \quad \lim_{n \rightarrow +\infty} \frac{\mathbb{P}(\|\mathbf{X}_{[0,n]}\|_p > x_n)}{n\mathbb{P}(|\mathbf{X}_0| > x_n)} = c(p).$$

Moreover,  $c(p) < \infty$  if  $p \geq \alpha$ , in particular,  $c(\infty) \leq c(p) \leq c(\alpha) = 1$ . If  $c(p) < \infty$  there exists  $\mathbf{Q}^{(p)} \in \tilde{\ell}^p$  such that  $\|\mathbf{Q}^{(p)}\|_p = 1$  a.s. and

$$(2.5) \quad \mathbb{P}(x_n^{-1}\mathbf{X}_{[0,n]} \in \cdot \mid \|\mathbf{X}_{[0,n]}\|_p > x_n) \xrightarrow{w} \mathbb{P}(Y\mathbf{Q}^{(p)} \in \cdot), \quad n \rightarrow \infty,$$

in the space  $(\tilde{\ell}^p, \tilde{d}_p)$  where  $Y$  is Pareto( $\alpha$ ) distributed, independent of  $\mathbf{Q}^{(p)}$ .

First, notice that under  $\mathbf{RV}_\alpha$ ,  $\mathbf{AC}$  and  $\mathbf{CS}_\alpha$  we obtain  $c(\alpha) = 1$ . This motivates the study of extremal  $\ell^\alpha$ -blocks since they reach high levels at a constant rate regardless of the temporal dependencies traced through  $c(p)$  in (2.4). Second, notice that for  $p > \alpha$  the result of Theorem 2.1 holds under  $\mathbf{RV}_\alpha$  and  $\mathbf{AC}$  solely. Indeed, condition  $\mathbf{CS}_p$  holds for  $p > \alpha$  by a Karamata-type argument; see Remark 5.1. We state Theorem 2.1 under the one-sided anti-clustering condition  $\mathbf{AC}$ . We use this condition together with a telescoping sum argument to compensate for the classical two-sided condition (5.1) used in Kulik and Soulier [23]. Conditions similar to  $\mathbf{CS}_p$  are standard when dealing with sum functionals acting on  $(\mathbf{X}_t)$ ; see e.g., Mikosch and Wintenberger [27]. We refer to Section 5 for a thorough discussion on the conditions  $\mathbf{AC}$ ,  $\mathbf{CS}_p$ , and the growth conditions imposed on  $(x_n)$ .

We refer to a relation of the type (2.4) as *large deviation probabilities* motivated by the following observation. Write  $S_k^{(p)} = \sum_{t=0}^k |\mathbf{X}_t|^p$ , for  $k \geq 1$ . Then  $|\mathbf{X}|^p$  is regularly varying with index  $\alpha/p$ . Relation (2.4) implies that

$$\begin{aligned} \mathbb{P}(\|\mathbf{X}_{[0,n]}\|_p > x_n) &= \mathbb{P}(S_n^{(p)} > x_n^p) \\ &\sim c(p) n \mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0, \quad n \rightarrow \infty. \end{aligned}$$

Thus the left-hand probability describes the rare event that the sum process  $S_n^{(p)}$  exceeds the extreme threshold  $x_n^p$ .

Relation (2.5) extends the large deviation result for  $\|\mathbf{X}_{[0,n]}\|_p$  in (2.4) to one for the process  $\mathbf{X}_{[0,n]}$  in the sequence space  $(\tilde{\ell}^p, \tilde{d}_p)$ . Motivated by inference of the spectral cluster process  $\mathbf{Q}^{(p)}$ , we establish (2.5) employing weak convergence in the spirit of the polar decomposition from (2.3).

**Remark 2.2.** Recall Hult and Lindskog [21] introduced regular variation for random elements assuming values in a general complete separable metric space by extending the vague convergence approach (see Resnick [33]) to  $M_0$ -convergence; see also Lindskog *et al.* [26]. Relation (2.5) provides a family of Borel sets in  $(\tilde{\ell}^p, \tilde{d}_p)$  for which the weak limit of the self-normalized blocks



$\mathbf{X}_{[0,n]}/\|\mathbf{X}_{[0,n]}\|_p$  exists. This result implies that the sequence of measures

$$\begin{aligned} \mu_n(\cdot) &:= \mathbb{P}(x_n^{-1}\mathbf{X}_{[0,n]} \in \cdot) / \mathbb{P}(\|\mathbf{X}_{[0,n]}\|_p > x_n) \\ \rightarrow \quad \mu(\cdot) &:= \int_0^\infty \mathbb{P}(y \mathbf{Q}^{(p)} \in \cdot) d(-y^{-\alpha}), \quad n \rightarrow \infty, \end{aligned}$$

in the  $M_0$ -sense in  $(\tilde{\ell}^p, \tilde{d}_p)$ . By the portmanteau theorem for measures (Theorem 2.4. in Hult and Lindskog [21])

$$\mu_n(A) = \mathbb{P}(x_n^{-1}\mathbf{X}_{[0,n]} \in A) / \mathbb{P}(\|\mathbf{X}_{[0,n]}\|_p > x_n) \rightarrow \mu(A),$$

for all Borel sets  $A$  in  $(\tilde{\ell}^p, \tilde{d}_p)$  satisfying  $\mu(\partial A) = 0$  and  $\mathbf{0} \notin \bar{A}$ . This approach is discussed in Kulik and Soulier [23].

### 3. SPECTRAL CLUSTER PROCESS REPRESENTATION

**3.1. The spectral cluster process in  $\ell^p$ .** From (2.1) recall the spectral tail process  $\Theta$  of a stationary sequence  $(\mathbf{X}_t)$  satisfying  $\mathbf{RV}_\alpha$ . We start by showing a representation of the spectral cluster process  $\mathbf{Q}^{(p)}$  from (2.5) in terms of  $\Theta$ . We deduce that the spectral cluster process is also well defined in  $(\ell^p, d_p)$ . The proof is deferred to Section 8.3.

**Proposition 3.1.** *For  $p, \alpha > 0$  assume  $\|\Theta\|_p + \|\Theta\|_\alpha < \infty$  a.s. Under the assumptions of Theorem 2.1 the constant  $c(p)$  defined in (2.4) admits the representation*

$$(3.1) \quad c(p) = \mathbb{E}[\|\Theta/\|\Theta\|_\alpha\|_p^\alpha].$$

*In addition, if  $c(p) < \infty$  then the distribution of the spectral cluster process  $\mathbf{Q}^{(p)}$  is given by*

$$(3.2) \quad \mathbb{P}(\mathbf{Q}^{(p)} \in \cdot) = (c(p))^{-1} \mathbb{E}[\|\Theta/\|\Theta\|_\alpha\|_p^\alpha \mathbf{1}(\Theta/\|\Theta\|_p \in \cdot)],$$

*in the space  $(\ell^p, d_p)$ .*

This result provides a new representation of the distribution of  $\mathbf{Q}^{(p)}$  for fixed  $p$ . In what follows, under the assumptions of Theorem 2.1, the spectral cluster processes are assumed to be defined in the space  $(\ell^p, d_p)$  via (3.2). Proposition 3.1 also relates distinct spectral cluster processes to each other by the change-of-norms transform in (3.2). In the next section we deal with the case  $p = \alpha$ .

**3.2. The spectral cluster process in  $\ell^\alpha$ .** In view of (3.2) the process  $\Theta/\|\Theta\|_\alpha$  is the candidate for the  $\ell^\alpha$ -spectral cluster process  $\mathbf{Q}^{(\alpha)}$  introduced in (2.5), and it plays a key role for characterizing  $\mathbf{Q}^{(p)}$  in general. The following result shows that  $\Theta/\|\Theta\|_\alpha$  is well defined in  $(\ell^\alpha, d_\alpha)$  under **AC**.

**Proposition 3.2.** *Let  $(\mathbf{X}_t)$  be a stationary sequence satisfying  $\mathbf{RV}_\alpha$  with spectral tail process  $(\Theta_t)$ . Then the following statements are equivalent:*

- i)  $\|\Theta\|_\alpha < \infty$  a.s. and  $\Theta/\|\Theta\|_\alpha$  is well defined in  $\ell^\alpha$ .
- ii)  $|\Theta_t| \rightarrow 0$  a.s. as  $t \rightarrow \infty$ .



- iii) The time of the largest record  $T^* := \inf\{s : s \in \mathbb{Z} \text{ such that } |\Theta_s| = \sup_{t \in \mathbb{Z}} |\Theta_t|\}$  is finite a.s.

Moreover, these statements hold under **AC**.

A proof of Proposition 3.2 is given in Lemma 3.6 of Buriticá *et al.* [7], appealing to results by Janssen [22].

From (2.3) recall the sequence of spectral components  $(\mathbf{Q}^{(\alpha)}(h))_{h \geq 0}$  of the vectors  $(\mathbf{X}_{[0,h]})_{h \geq 0}$ , satisfying the property  $\|\mathbf{Q}^{(\alpha)}(h)\|_\alpha = 1$  a.s. Our next result relates this sequence of spectral components to  $\Theta/\|\Theta\|_\alpha$ .

**Proposition 3.3.** *Let  $(\mathbf{X}_t)$  be a stationary time series satisfying  $\mathbf{RV}_\alpha$  and  $\lim_{t \rightarrow \infty} |\Theta_t| = 0$  a.s. Then  $\mathbf{Q}^{(\alpha)}(h) \xrightarrow{d} \mathbf{Q}^{(\alpha)}(\infty)$  as  $h \rightarrow \infty$  in  $(\tilde{\ell}^\alpha, \tilde{d}_\alpha)$  with  $\mathbf{Q}^{(\alpha)}(\infty) \stackrel{d}{=} \Theta/\|\Theta\|_\alpha$ .*

This result gives raise to the interpretation of  $\Theta/\|\Theta\|_\alpha$  as the spectral component of  $(\mathbf{X}_t)$  in  $\ell^\alpha$ . The proof is given in Section 8.3. We deduce the following almost sure relation in terms of the spectral cluster process in  $\ell^\alpha$ .

**Proposition 3.4.** *Let  $(\mathbf{X}_t)$  be a stationary sequence satisfying  $\mathbf{RV}_\alpha$  with spectral tail process  $(\Theta_t)$ . Under the assumptions **AC** and **CS** $_\alpha$ , we deduce the a.s. representations  $\mathbf{Q}^{(\alpha)} = \Theta/\|\Theta\|_\alpha$  and  $\Theta = \mathbf{Q}^{(\alpha)}/|\mathbf{Q}_0^{(\alpha)}|$  in  $(\tilde{\ell}^\alpha, \tilde{d}_\alpha)$ .*

Proposition 3.4 follows directly from Propositions 3.1 and 3.2.

#### 4. CONSISTENT CLUSTER INFERENCE BASED ON SPECTRAL CLUSTER PROCESSES

Let  $\mathbf{X}_1, \dots, \mathbf{X}_n$  be a sample from a stationary sequence  $(\mathbf{X}_t)$  satisfying  $\mathbf{RV}_\alpha$  for some  $\alpha > 0$ , and choose  $p > 0$ . We split the sample into disjoint blocks  $\mathbf{B}_t := \mathbf{X}_{(t-1)b+1:b}$ ,  $t = 1, \dots, m_n$ , where  $b = b_n \rightarrow \infty$  and  $m = m_n := \lfloor n/b_n \rfloor \rightarrow \infty$ . Throughout we assume that the sequence  $(x_n)$  satisfies the conditions of Theorem 2.1 for  $p > 0$ . We denote  $k = k_n := \lfloor m_n \mathbb{P}(\|\mathbf{B}\|_p > x_{b_n}) \rfloor \rightarrow \infty$ . Then, in particular  $\mathbb{P}(\|\mathbf{B}_1\|_p > x_b) \rightarrow 0$ ,  $m_n \rightarrow \infty$  and  $k_n \rightarrow \infty$ .

**4.1. Cluster functionals and mixing.** The real-valued function  $g$  on  $\tilde{\ell}^p$  is a *cluster functional* if it vanishes in some neighborhood of the origin and  $\mathbb{P}(Y\mathbf{Q}^{(p)} \in D(g)) = 0$  where  $D(g)$  denotes the set of discontinuity points of  $g$ . In what follows, it will be convenient to write  $\mathcal{G}_+(\tilde{\ell}^p)$  for the class of non-negative functions on  $\tilde{\ell}^p$  which vanish in some neighborhood of the origin.

For asymptotic theory we will need the following *mixing condition*.

**Condition  $\mathbf{MX}_p$ .** There exists an integer sequence  $b_n \rightarrow \infty$  such that  $m_n \rightarrow \infty$ ,  $k_n \rightarrow \infty$ , and for every Lipschitz-continuous  $f \in \mathcal{G}_+(\tilde{\ell}^p)$ , the sequence  $(x_n)$  satisfies

$$\mathbb{E}[e^{-\frac{1}{k} \sum_{i=1}^m f(x_b^{-1} \mathbf{B}_i)}] = \left( \mathbb{E}[e^{-\frac{1}{k} \sum_{i=1}^{\lfloor m/k \rfloor} f(x_b^{-1} \mathbf{B}_i)}] \right)^k + o(1), \quad n \rightarrow \infty. \quad (4.1)$$

with  $m_n := \lfloor n/b_n \rfloor$  and  $k_n := \lfloor m_n \mathbb{P}(\|\mathbf{B}\|_p > x_{b_n}) \rfloor$ .

If  $\mathbf{MX}_p$  is required in the sequel we will refer to the sequences  $(b_n)$ ,  $(m_n)$  and  $(k_n)$  chosen in this condition.

**Remark 4.1.** Condition  $\mathbf{MX}_p$  is similar to the mixing conditions  $\mathcal{A}$ ,  $\mathcal{A}'$  in Davis and Hsing [11], Basrak *et al.* [2], respectively. These are defined in terms of sequences  $(f(\mathbf{X}_t))$  while our functionals  $f$  act on blocks.  $\mathbf{MX}_p$  holds under mild conditions, for example, under strong mixing with quite general rate; cf. Lemma 6.2. in Basrak *et al.* [3].

**4.2. Consistent cluster inference.** The following result is the basis for an empirical procedure for spectral cluster inference built on disjoint blocks. The proof is given in Section 8.4.1.

**Theorem 4.2.** *Assume the conditions of Theorem 2.1 hold for  $p > 0$  with  $c(p) < \infty$  together with  $\mathbf{MX}_p$ . Then  $\|\mathbf{B}\|_{p,(k+1)}/x_b \xrightarrow{\mathbb{P}} 1$  and for every  $g \in \mathcal{G}_+(\tilde{\ell}^p)$ ,*

$$(4.2) \quad \frac{1}{k} \sum_{t=1}^m g(\|\mathbf{B}\|_{p,(k+1)}^{-1} \mathbf{B}_t) \xrightarrow{\mathbb{P}} \int_0^\infty \mathbb{E}[g(y\mathbf{Q}^{(p)})] d(-y^{-\alpha}), \quad n \rightarrow \infty.$$

such that  $\|\mathbf{B}\|_{p,(1)} \geq \|\mathbf{B}\|_{p,(2)} \geq \dots \geq \|\mathbf{B}\|_{p,(m)}$ .

By virtue of Proposition 3.1 we can derive the same spectral cluster statistic by letting the functionals  $g_p : \tilde{\ell}^p \rightarrow \mathbb{R}$  act on  $\mathbf{Q}^{(p)}$  for different pairs  $(p, g_p)$ . This opens the road to different ways to estimate the same constant, for example,  $c(q)$  for  $q > 0$ . To compare inference procedures tuned with different  $p$ , we observe that Theorem 4.2 promotes the use of order statistics of the sample of  $\ell^p$ -norms. The sequence  $(k_n)$  in (4.2) corresponds to the number of extreme blocks used for inference. The large deviation principles of Theorem 2.1 allow us then to compare the sequences  $k_n = k_n(p)$ . For inference through  $\mathbf{Q}^{(p)}$  the relation

$$(4.3) \quad k_n = \lfloor m_n \mathbb{P}(\|\mathbf{B}\|_p > x_{b_n}) \rfloor \sim c(p) n \mathbb{P}(|\mathbf{X}_0| > x_{b_n}),$$

justifies taking  $k_n$  larger as  $p$  decreases, for  $p \in (\alpha, \infty]$ , since  $c(\cdot)$  is a non-increasing function of  $p$ , and  $(x_n)$  is a sequence satisfying **AC** and  $n\mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0$ . For  $p \in (0, \alpha]$ , the sequence  $(x_n)$  must satisfy the additional condition **CS<sub>p</sub>**, which restricts the range of possible values for  $k_n$ , but allows us to consider continuous functionals on  $(\tilde{\ell}^p, \tilde{d}^p)$ . One advantage of choosing  $p = \alpha$  is that  $c(\alpha) = 1$ , thus the choice of  $k_n = k_n(\alpha)$  does not rely on the serial dependencies summarized in  $c(p)$ .

**4.3. Applications.** In this section we apply Theorem 4.2 for inference on some indices related to the extremes in a time-dependent sample and focus on cluster inference using  $\mathbf{Q}^{(\alpha)}$ . We illustrate our estimators for a regularly varying linear process in Section 7.

4.3.1. *The extremal index.* The extremal index of a regularly varying stationary time series has interpretation as a measure of clustering of serial exceedances, and was originally introduced in Leadbetter [24] and Leadbetter *et al.* [25]. If  $(\mathbf{X}'_t)$  is iid with the same marginal distribution as  $(\mathbf{X}_t)$  then the extremal index  $\theta_{|\mathbf{X}|}$  relates the expected number of serial exceedances of  $(|\mathbf{X}_t|)$  with the serial exceedances of  $(|\mathbf{X}'_t|)$ . Assuming **AC** and additional mixing assumptions (see e.g. Theorem 2.3. in [7]), the extremal index  $\theta_{|\mathbf{X}|}$  of  $(|\mathbf{X}_t|)$  exists and equals  $c(\infty)$ .

We aim at applying Theorem 4.2 with  $p = \alpha$ . In this setting, the change-of-norm formula in (3.1) leads to the identities

$$\theta_{|\mathbf{X}|} = c(\infty) = \mathbb{E} \left[ \frac{\|\Theta\|_\infty^\alpha}{\|\Theta\|_\alpha^\alpha} \right] = \mathbb{E}[\|\mathbf{Q}^{(\alpha)}\|_\infty^\alpha].$$

Then, letting  $p = \alpha$  and  $g(\mathbf{x}) = (\|\mathbf{x}\|_\infty^\alpha / \|\mathbf{x}\|_\alpha^\alpha) \mathbb{1}(\|\mathbf{x}\|_\alpha > 1)$  on the right-hand side of (4.2), we obtain

$$\begin{aligned} \int_0^\infty \mathbb{E}[g(y\mathbf{Q}^{(\alpha)})] d(-y^{-\alpha}) &= \int_0^\infty \mathbb{E} \left[ \frac{\|\mathbf{Q}^{(\alpha)}\|_\infty^\alpha}{\|\mathbf{Q}^{(\alpha)}\|_\alpha^\alpha} \mathbb{1}(\|\mathbf{Q}^{(\alpha)}\|_\alpha > y^{-\alpha}) \right] d(-y^{-\alpha}) \\ &= \mathbb{E}[\|\mathbf{Q}^{(\alpha)}\|_\infty^\alpha] = c(\infty). \end{aligned}$$

Next we introduce a new consistent disjoint blocks estimator of the extremal index defined from exceedances of  $\ell^\alpha$ -norm blocks.

**Corollary 4.3.** *Assume the conditions of Theorem 4.2 for  $p = \alpha$ . Then*

$$(4.4) \quad \frac{1}{k} \sum_{t=1}^m \frac{\|\mathbf{B}_t\|_\infty^\alpha}{\|\mathbf{B}_t\|_\alpha^\alpha} \mathbb{1}(\|\mathbf{B}_t\|_\alpha > \|\mathbf{B}\|_{\alpha, (k+1)}) \xrightarrow{\mathbb{P}} c(\infty), \quad n \rightarrow \infty.$$

An advantage of inferring the extremal index using extremal  $\ell^\alpha$ -blocks is that the tuning parameter  $k_n$  of the estimator does not rely on the clustering effect of the series since  $c(\alpha) = 1$  in Equation (4.3).

**Remark 4.4.** We can compare this estimator of  $c(\infty)$  with one based on the clusters (of exceedances). Motivated by the blocks estimator of the extremal index in Hsing [20], we let  $g(\mathbf{x}) := \sum_{j \in \mathbb{Z}} \mathbb{1}(|\mathbf{x}_t| > 1)$  act on large  $\ell^\infty$ -blocks. Choosing  $p = \infty$  and using this  $g$  on the right-hand side of (4.2), we can find an integer sequence  $k = k_n(\infty) \rightarrow \infty$  such that

$$(4.5) \quad \left( \frac{1}{k} \sum_{t=1}^n \mathbb{1}(|\mathbf{X}_t| > \|\mathbf{B}\|_{\infty, (k+1)}) \right)^{-1} \xrightarrow{\mathbb{P}} c(\infty), \quad n \rightarrow \infty.$$

Arguing as for (4.3),  $k_n \sim m_n \mathbb{P}(\|\mathbf{B}\|_\infty > x_b) \sim c(\infty) n \mathbb{P}(|\mathbf{X}_0| > x_b)$ . Thus, the number of extreme blocks used in (4.5) shrinks when  $c(\infty) < 1$ , compared to its implementation in an iid setting. In practice, this can make the choice of  $k_n$  sensitive to the temporal ties.

**4.3.2. A cluster index for sums.** In this section we assume that  $\alpha \in (0, 2)$  and  $\mathbb{E}[\mathbf{X}] = \mathbf{0}$  for  $\alpha \in (1, 2)$ . We study the partial sums  $\mathbf{S}_n := \sum_{t=1}^n \mathbf{X}_t$ ,  $n \geq 1$ , and introduce a normalizing sequence  $(a_n)$  such that  $n \mathbb{P}(|\mathbf{X}_0| > a_n) \rightarrow 1$ . Starting with Davis and Hsing [11],  $\alpha$ -stable central limit theory for  $(\mathbf{S}_n/a_n)$  was proved under suitable anti-clustering and mixing conditions.

In this setting, the quantity  $c(1)$  appears naturally and was coined cluster index in Mikosch and Wintenberger [28]. For  $d = 1$  it can be interpreted as an equivalent of the extremal index for partial sums rather than maxima. Indeed, consider a real-valued regularly varying stationary sequence  $(X_t)$  with index of regular variation  $\alpha \in (0, 2)$  satisfying  $\mathbb{P}(X \leq -x) = o(\mathbb{P}(X > x))$  or  $X \stackrel{d}{=} -X$ . Consider an iid sequence  $(X'_t)$  with  $X \stackrel{d}{=} X'$  and partial sums  $(S'_n)$ . Then  $a_n^{-1} S_n \xrightarrow{d} \xi_\alpha$  and  $a_n^{-1} S'_n \xrightarrow{d} \xi'_\alpha$ , both  $\xi_\alpha$  and  $\xi'_\alpha$  are  $\alpha$ -stable and

$$\mathbb{E}[e^{iu\xi_\alpha}] = (\mathbb{E}[e^{iu\xi'_\alpha}])^{c(1)}.$$

Under the assumptions of Proposition 3.1 and for  $p = \alpha$  we have  $c(1) = \mathbb{E}[\|\mathbf{Q}^{(\alpha)}\|_1^\alpha]$ . For  $\alpha \in (0, 1]$ , take  $p = \alpha$  and  $g(\mathbf{x}) = (\|\mathbf{x}\|_1^\alpha / \|\mathbf{x}\|_\alpha^\alpha) \mathbb{1}(\|\mathbf{x}\|_\alpha > 1)$  on the right-hand side of (4.2). Then an application of Theorem 4.2 with  $p = \alpha$  and  $g$  as mentioned yields a consistent estimator of  $c(1)$ .

**Corollary 4.5.** *We assume the conditions of Theorem 4.2 for  $p = \alpha$  and  $\alpha \in (0, 1]$ . Then we have for  $k = k_n \rightarrow \infty$ ,*

$$(4.6) \quad \frac{1}{k} \sum_{t=1}^m \frac{\|\mathbf{B}_t\|_1^\alpha}{\|\mathbf{B}_t\|_\alpha^\alpha} \mathbb{1}(\|\mathbf{B}_t\|_\alpha > \|\mathbf{B}\|_{\alpha, (k+1)}) \xrightarrow{\mathbb{P}} c(1), \quad n \rightarrow \infty.$$

The estimator on the left-hand side of (4.6) has the advantage that  $k_n \sim n \mathbb{P}(|\mathbf{X}_0| > x_{b_n})$ . Relation (4.6) holds by virtue of (4.3) regardless of the temporal dependence in the series.

**Remark 4.6.** For  $\alpha \in (1, 2)$  the function  $g$  applied in (4.6) to extremal  $\ell^\alpha$ -blocks is no longer bounded. If  $c(1) < \infty$  we can apply Theorem 4.2 with  $p = 1$  and  $g(\mathbf{x}) = (\|\mathbf{x}\|_\alpha^\alpha / \|\mathbf{x}\|_1^\alpha) \mathbb{1}(\|\mathbf{x}\|_1 > 1)$  to obtain a consistent estimator of  $c(1)$ . Indeed, the right-hand side of (4.2) turns into

$$\begin{aligned} \int_0^\infty \mathbb{E}[g(y\mathbf{Q}^{(1)})] d(-y^{-\alpha}) &= \int_0^\infty \mathbb{E}\left[\frac{\|\mathbf{Q}^{(1)}\|_\alpha^\alpha}{\|\mathbf{Q}^{(1)}\|_1^\alpha} \mathbb{1}(\|\mathbf{Q}^{(1)}\|_1^\alpha > y^{-\alpha})\right] d(-y^{-\alpha}) \\ &= \mathbb{E}[\|\mathbf{Q}^{(1)}\|_\alpha^\alpha] = (c(1))^{-1}, \end{aligned}$$

where the last identity follows from Proposition 3.1. Then Theorem 4.2 for  $p = 1$  and  $g$  as above yields a consistent estimator of  $c(1)$ . Note that  $c(1) \in [1, \infty)$  for  $\alpha \in (1, 2)$ . Hence the number  $k_n$  of extremal  $\ell^1$ -blocks for this estimator does not decrease in comparison with the iid case. This feature can also make this estimator robust for cluster inference.

**Remark 4.7.** Arguing as in Cissokho and Kulik [8], Kulik and Soulier [23], and assuming  $\mathbf{CS}_1$ , we can extend Theorem 4.2 for  $p = \infty$  to hold

for  $\ell^1$ -functionals. Then we can find  $k = k_n(\infty) \rightarrow \infty$  such that, with  $g(\mathbf{x}) := \mathbb{1}(\|\mathbf{x}\|_1 > 1)$  and  $p = \infty$  in (4.2),

$$(4.7) \quad \frac{\sum_{t=1}^m \mathbb{1}(\|\mathbf{B}_t\|_1 > \|\mathbf{B}\|_{\infty, (k+1)})}{\sum_{t=1}^n \mathbb{1}(\|\mathbf{X}_t\|_1 > \|\mathbf{B}\|_{\infty, (k+1)})} \xrightarrow{\mathbb{P}} c(1), \quad n \rightarrow \infty.$$

Here, following (4.3), we have  $k_n \sim c(\infty) n \mathbb{P}(|\mathbf{X}_0| > x_b)$ . This alternative estimator of  $c(1)$  based on extremal  $\ell^\infty$ -blocks is consistent for  $\alpha \in (0, 2)$ . Then, as in the extremal index example, the tuning parameter  $k_n$  in (4.7) is linked to the constant  $c(\infty) \in (0, 1]$  and must be chosen carefully in agreement with the clustering effect of the series.

Theorem 4.2 provides estimators of the parameters of the  $\alpha$ -stable limit  $\xi_\alpha$  of  $(\mathbf{S}_n/a_n)$ . Indeed, following the theory in Bartkiewicz *et al.* [1], we characterize the  $\alpha$ -stable limit in terms of  $\mathbf{Q}^{(1)}$ ; the proof is given in Section 8.4.2.

**Proposition 4.8.** *Consider a stationary regularly varying sequence  $(\mathbf{X}_t)$  with index  $\alpha \in (0, 1) \cup (1, 2)$ . We assume the mixing condition*

$$\mathbb{E}[e^{i\mathbf{u}^\top \mathbf{S}_n/a_n}] = (\mathbb{E}[e^{i\mathbf{u}^\top \mathbf{S}_{b_n}/a_n}])^{m_n} + o(1), \quad n \rightarrow \infty, \quad \mathbf{u} \in \mathbb{R}^d,$$

and the anti-clustering condition, for every  $\delta > 0$ ,

$$(4.8) \quad \lim_{l \rightarrow \infty} \limsup_{n \rightarrow \infty} n \sum_{t=l}^{b_n} \mathbb{E}[(|\mathbf{X}_t/a_n| \wedge \delta)(|\mathbf{X}_0/a_n| \wedge \delta)] = 0.$$

Then  $\mathbf{S}_n/a_n \xrightarrow{d} \xi_\alpha$  for an  $\alpha$ -stable random vector  $\xi_\alpha$  with characteristic function  $\mathbb{E}[\exp(i\mathbf{u}^\top \xi_\alpha)] = \exp(-c_\alpha \sigma_\alpha(\mathbf{u})(1 - i\beta(\mathbf{u})\tan(\alpha\pi/2)))$ ,  $\mathbf{u} \in \mathbb{R}^d$ , where  $c_\alpha := (\Gamma(2-\alpha)/|1-\alpha|)(1 \wedge \alpha) \cos(\alpha\pi/2)$ , and the scale and skewness parameters have representation

$$\begin{aligned} \sigma_\alpha(\mathbf{u}) &:= c(1) \mathbb{E}[|\mathbf{u}^\top \sum_{t \in \mathbb{Z}} \mathbf{Q}_t^{(1)}|^\alpha], \\ \beta(\mathbf{u}) &:= (\mathbb{E}[(\mathbf{u}^\top \sum_{t \in \mathbb{Z}} \mathbf{Q}_t^{(1)})_+^\alpha] - (\mathbf{u}^\top \sum_{t \in \mathbb{Z}} \mathbf{Q}_t^{(1)})_-^\alpha) / \mathbb{E}[|\mathbf{u}^\top \sum_{t \in \mathbb{Z}} \mathbf{Q}_t^{(1)}|^\alpha]. \end{aligned}$$

As for  $c(1)$ , an application of Theorem 4.2 with  $p = 1$  for  $\alpha \in (1, 2)$  and  $p = \alpha$  for  $\alpha \in (0, 1)$  yields natural estimators of the parameters  $(\sigma_\alpha(\mathbf{u}), \beta(\mathbf{u}))$  in the central limit theorem of Proposition 4.8.

## 5. A DISCUSSION OF THE ASSUMPTIONS OF THE LARGE DEVIATION PRINCIPLE IN THEOREM 2.1

Consider a stationary sequence  $(\mathbf{X}_t)$  satisfying  $\mathbf{RV}_\alpha$  and let  $(x_n)$  be a threshold sequences such that  $n \mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0$ . In the conditions **AC** and **CS<sub>p</sub>** below we refer to the same sequence  $(x_n)$ . In this section we will discuss the conditions of Theorem 2.1.

**5.1. Anti-clustering condition AC.** For every  $\delta > 0$ ,

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}(\|\mathbf{X}_{[k,n]}\|_\infty > \delta x_n \mid |\mathbf{X}_0| > \delta x_n) = 0.$$

Condition **AC** ensures that a large value at present time does not persist indefinitely in the extreme future of the time series. This anti-clustering is weaker than the more common two-sided one:

$$(5.1) \quad \lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \mathbb{P}\left(\max_{k \leq |t| \leq n} |\mathbf{X}_t| > \delta x_n \mid |\mathbf{X}_0| > \delta x_n\right) = 0.$$

A simple sufficient condition, which breaks block-wise extremal dependence into pair-wise, is given by

$$\lim_{k \rightarrow \infty} \limsup_{n \rightarrow \infty} \sum_{t=k}^n \mathbb{P}(|\mathbf{X}_t| > \delta x_n \mid |\mathbf{X}_0| > \delta x_n).$$

For  $m$ -dependent  $(\mathbf{X}_t)$  the latter condition turns into  $n \mathbb{P}(|\mathbf{X}_0| > \delta x_n) \rightarrow 0$  which is always satisfied.

If  $p \leq \alpha$  an extra assumption is required for controlling the accumulation of moderate extremes within a block.

**5.2. Vanishing-small-values condition  $\mathbf{CS}_p$ .** For  $p \in (0, \alpha]$  we assume that for a sequence  $(x_n)$  satisfying  $n \mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0$  and for every  $\delta > 0$ , we have

$$(5.2) \quad \lim_{\epsilon \rightarrow 0} \limsup_{n \rightarrow \infty} \frac{\mathbb{P}(\|\overline{x_n^{-1} \mathbf{X}_{[1,n]}}^\epsilon\|_p^p - \mathbb{E}[\|\overline{x_n^{-1} \mathbf{X}_{[1,n]}}^\epsilon\|_p^p] > \delta)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} = 0.$$

We refer to (5.2) as condition  $\mathbf{CS}_p$  in what follows. If  $\alpha < p < \infty$  then by Karamata's theorem (see Bingham *et al.* [5]) and since  $n \mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0$ ,

$$\mathbb{E}[\|\overline{x_n^{-1} \mathbf{X}_{[1,n]}}^\epsilon\|_p^p] = n \mathbb{E}[\|\overline{x_n^{-1} \mathbf{X}}^\epsilon\|^p] = o(1), \quad n \rightarrow \infty.$$

Also, if  $p < \alpha$ , then  $\mathbb{E}[|\mathbf{X}|^p] < \infty$ . If we also have  $n/x_n^p \rightarrow 0$  then

$$\mathbb{E}[\|\overline{x_n^{-1} \mathbf{X}_{[1,n]}}^\epsilon\|_p^p] \leq n x_n^{-p} \mathbb{E}[|\mathbf{X}|^p] \rightarrow 0, \quad n \rightarrow \infty.$$

If  $p = \alpha$ ,  $\mathbb{E}[|\mathbf{X}|^\alpha] < \infty$  and  $n/x_n^\alpha \rightarrow 0$  then the latter relation remains valid. If  $\mathbb{E}[|\mathbf{X}|^\alpha] = \infty$  then  $\mathbb{E}[\|\overline{x_n^{-1} \mathbf{X}}^\epsilon\|^\alpha] = x_n^{-\alpha} \ell(x_n)$  for some slowly varying function  $\ell$  depending on  $\epsilon$ , hence for every small  $\kappa > 0$  and large  $n$ ,  $\ell(x_n) \leq x_n^\kappa$ . Then the condition  $n x_n^{-\alpha+\kappa} \rightarrow 0$  also implies that  $\mathbb{E}[\|\overline{x_n^{-1} \mathbf{X}_{[1,n]}}^\epsilon\|^\alpha] = o(1)$ . Thus we retrieve  $\mathbf{CS}_p$  as used in Theorem 2.1. In sum, under the aforementioned additional growth conditions on  $(x_n)$  centering in (5.2) can be avoided. This is similar to condition  $\mathbf{CS}_p$  in Theorem 2.1.

We mentioned that conditions of a similar type as  $\mathbf{CS}_p$  are standard when dealing with sum functionals acting on  $(\mathbf{X}_t)$  (see for example Davis and Hsing [11], Bartkiewicz *et al.* [1], Mikosch and Wintenberger [27, 28, 29]), and are also discussed in Kulik and Soulier [23].

**Remark 5.1.** Assume  $\alpha < p < \infty$ . Then applications of Markov's inequality of order 1 and Karamata's theorem yield for  $\delta > 0$ , as  $n \rightarrow \infty$ ,

$$\begin{aligned} \frac{\mathbb{P}(\|\overline{x_n^{-1} \mathbf{X}_{[1,n]}}^\epsilon\|_p^p > \delta)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} &= \frac{\mathbb{P}(\sum_{t=1}^n |\overline{x_n^{-1} \mathbf{X}_t}^\epsilon|^p > \delta)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} \\ &\leq \frac{\mathbb{E}[\|\overline{x_n^{-1} \mathbf{X}_0}^\epsilon\|^p]}{\delta \mathbb{P}(|\mathbf{X}_0| > \epsilon x_n)} \frac{\mathbb{P}(|\mathbf{X}_0| > \epsilon x_n)}{\mathbb{P}(|\mathbf{X}_0| > x_n)} \rightarrow c \epsilon^{p-\alpha}. \end{aligned}$$

The right-hand side converges to zero as  $\epsilon \rightarrow 0$ . Here and in what follows,  $c$  denotes any positive constant whose value is not of interest. We conclude that (5.2) is automatic for  $p > \alpha$ .

**Remark 5.2.** Condition  $\mathbf{CS}_p$  is challenging to check for  $p \leq \alpha$ . For  $p/\alpha \in (1/2, 1]$ , by Čebyshev's inequality,

$$\begin{aligned} &\mathbb{P}(\|\overline{x_n^{-1} \mathbf{X}_{[1,n]}}^\epsilon\|_p^p - \mathbb{E}[\|\overline{x_n^{-1} \mathbf{X}_{[1,n]}}^\epsilon\|_p^p] > \delta) / [n \mathbb{P}(|\mathbf{X}_0| > x_n)] \\ &\leq \delta^{-2} \text{var}(\|\overline{x_n^{-1} \mathbf{X}_{[1,n]}}^\epsilon\|_p^p) / [n \mathbb{P}(|\mathbf{X}_0| > x_n)] \\ &\leq \delta^{-2} \frac{\mathbb{E}[\|\overline{x_n^{-1} \mathbf{X}_0}^\epsilon\|^{2p}]}{\mathbb{P}(|\mathbf{X}_0| > x_n)} \left[ 1 + 2 \sum_{h=1}^{n-1} |\text{corr}(\|\overline{x_n^{-1} \mathbf{X}_0}^\epsilon\|^p, \|\overline{x_n^{-1} \mathbf{X}_h}^\epsilon\|^p)| \right]. \end{aligned}$$

Now assume that  $(\mathbf{X}_t)$  is  $\rho$ -mixing with summable rate function  $(\rho_h)$ ; cf. Bradley [6]. Then the right-hand side is bounded by

$$\delta^{-2} \frac{\mathbb{E}[\|\overline{x_n^{-1} \mathbf{X}}^\epsilon\|^{2p}]}{\mathbb{P}(|\mathbf{X}_0| > x_n)} \left[ 1 + 2 \sum_{h=1}^{\infty} \rho_h \right] \sim \delta^{-2} \epsilon^{2p-\alpha} \left[ 1 + 2 \sum_{h=1}^{\infty} \rho_h \right], \quad \epsilon \rightarrow 0,$$

where we applied Karamata's theorem in the last step, and  $\mathbf{CS}_p$  follows. For Markov chains weaker assumptions such as the drift condition (DC) in Mikosch and Wintenberger [28, 29] can be used for checking  $\mathbf{CS}_p$ .

**Remark 5.3.** Condition  $\mathbf{CS}_p$  not only restricts the serial dependence of the time series  $(\mathbf{X}_t)$  but also the level of thresholds  $(x_n)$ . Indeed, for  $p/\alpha < 1/2$  and  $(\mathbf{X}_t)$  iid, since  $(\|n^{-1/2} \mathbf{X}'_{[1,n]}\|_p^p - \mathbb{E}[\|n^{-1/2} \mathbf{X}'_{[1,n]}\|_p^p])$  converges in distribution to a Gaussian limit by virtue of the central limit theorem,  $\mathbf{CS}_p$  implies necessarily that  $x_n/\sqrt{n} \rightarrow \infty$  as  $n \mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0$ .

**5.3. Threshold condition.** In Theorem 2.1 we assume growth conditions on  $(x_n)$ :  $n/x_n^p \rightarrow 0$  if  $p < \alpha$  and  $n/x_n^{\alpha-\kappa} \rightarrow 0$  for some  $\kappa > 0$  if  $p = \alpha$ .

For inference purposes it is tempting to decrease the threshold level  $x_n$  such that more exceedances are included in the estimators. Indeed, the assumptions on  $(x_n)$  can be relaxed, justified by results such as Nagaev's large deviation principle in [31], by adding a centering term as we will show in Lemma 5.4. However, in this section we aim at pointing at the difficulties that might arise while doing so in practice.

To motivate the results of this section we start by considering an iid sequence  $(\mathbf{X}_t)$  satisfying  $\mathbf{RV}_\alpha$  for some  $\alpha > 0$ . Then, for  $p > \alpha$ , (2.4) holds



with limit  $c(p) = 1$  and  $S_n^{(p)} = \sum_{t=1}^n |\mathbf{X}_t|^p$  has infinite expectation. If  $p < \alpha$  the process  $(S_n^{(p)})$  has finite expectation and by the law of large numbers, for  $n/x_n^p \rightarrow 0$ ,

$$(5.3) \quad \begin{aligned} & \mathbb{P}(\|\mathbf{X}_{[0,n]}\|_p > x_n (n x_n^{-p} \mathbb{E}[|\mathbf{X}|^p] + 1)^{1/p}) \\ &= \mathbb{P}(S_n^{(p)} - \mathbb{E}[S_n^{(p)}] > x_n^p (1 + o(1))) \rightarrow 0. \end{aligned}$$

Following Nagaev [31], a large deviation result for the centered process holds:

$$\mathbb{P}(S_n^{(p)} - \mathbb{E}[S_n^{(p)}] > x_n^p) \sim n \mathbb{P}(|\mathbf{X}_0| > x_n), \quad n \rightarrow \infty,$$

provided  $n/x_n^{\alpha-\kappa} \rightarrow 0$  for  $p/\alpha \in (1/2, 1)$  and some  $\kappa > 0$ , and  $\sqrt{n \log n}/x_n^p \rightarrow 0$  for  $p/\alpha < 1/2$ . These conditions are satisfied for extreme thresholds satisfying  $n/x_n^p \rightarrow 0$ . In this case the centering term  $\mathbb{E}[S_n^{(p)}]$  in (5.3) is always negligible which allows us to derive (2.4). Next, we extend the previous ideas to regularly varying time series.

**Lemma 5.4.** *Consider an  $\mathbb{R}^d$ -valued stationary process  $(\mathbf{X}_t)$  satisfying the conditions  $\mathbf{RV}_\alpha$ ,  $\mathbf{AC}$ ,  $\mathbf{CS}_p$  and  $c(p) < \infty$  for some  $p > 0$ . If  $p < \alpha$  then*

$$(5.4) \quad \lim_{n \rightarrow \infty} \frac{\mathbb{P}(\|\mathbf{X}_{[0,n]}\|_p > x_n (n x_n^{-p} \mathbb{E}[|\mathbf{X}|^p] + 1)^{1/p})}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} = c(p).$$

If  $p = \alpha$  then

$$(5.5) \quad \lim_{n \rightarrow \infty} \frac{\mathbb{P}(\|\mathbf{X}_{[0,n]}\|_\alpha > x_n (n \mathbb{E}[|\overline{\mathbf{X}}/x_n|^\alpha] + 1)^{1/\alpha})}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} = c(\alpha) = 1.$$

Moreover, if also  $\mathbb{E}[|\mathbf{X}|^\alpha] < \infty$  then equation (5.4) holds for  $p = \alpha$ .

The proof is given in Section 8.2. Now the restrictions on the level of the thresholds  $(x_n)$  are the ones implicitly implied by condition  $\mathbf{CS}_p$  in (5.2); see Remark 5.3.

We define an auxiliary sequence of levels:

$$z_n := z_n(p) = \begin{cases} x_n (n x_n^{-p} \mathbb{E}[|\mathbf{X}|^p] + 1)^{1/p} & \text{if } p < \alpha, \\ x_n (n \mathbb{E}[|\overline{\mathbf{X}}/x_n|^\alpha] + 1)^{1/\alpha} & \text{if } p = \alpha, \\ x_n & \text{if } p > \alpha. \end{cases}$$

For thresholds satisfying the growth conditions  $n/x_n^p \rightarrow 0$  we have  $z_n \sim x_n$ , while for moderate thresholds satisfying  $\mathbf{CS}_p$  and  $n/x_n^p \rightarrow \infty$  this is no longer the case.

For the purposes of inference Lemma 5.4 is not as satisfactory as (2.4) in Theorem 2.1. Indeed, the level  $z_n$  in the selection of the exceedances is not the original threshold  $x_n$ . For any moderate threshold  $x_n$  with  $z_n/x_n \rightarrow \infty$  the use of  $x_n$  instead of  $z_n$  might yield to a different limit. As a toy example,

consider the problem of inferring the constant  $c(q)/c(p)$  for  $p < \alpha$ ,  $q > p$ . Then an application of Lemma 5.4 ensures that

$$\begin{aligned} \mathbb{P}(\|\mathbf{X}_{[0,n]}\|_q > z_n(q) \mid \|\mathbf{X}_{[0,n]}\|_p > z_n(p)) \\ \rightarrow \mathbb{P}(\|Y\mathbf{Q}^{(p)}\|_q > 1) = \mathbb{E}[\|\mathbf{Q}^{(p)}\|_q^\alpha] = c(q)/c(p), \quad n \rightarrow \infty. \end{aligned}$$

However, choosing the same moderate threshold  $z_n = z_n(q)$ , we would have

$$\begin{aligned} \mathbb{P}(\|\mathbf{X}_{[0,n]}\|_q > z_n \mid \|\mathbf{X}_{[0,n]}\|_p > z_n) &\sim \frac{\mathbb{P}(\|\mathbf{X}_{[0,n]}\|_q > z_n)}{\mathbb{P}(n^{1/p}\mathbb{E}[\|\mathbf{X}\|^p]^{1/p} > z_n)} \\ &\rightarrow \begin{cases} 1 & \text{if } q < \alpha, \\ 0 & \text{if } q > \alpha, \end{cases} \quad n \rightarrow \infty. \end{aligned}$$

By this argument, the growth conditions on  $(x_n)$  are justified to simplify inference procedures. Otherwise, the choice of the threshold sequence becomes delicate.

## 6. INFERENCE BEYOND SHIFT-INVARIANT FUNCTIONALS

So far we only considered inference for shift-invariant functionals acting on  $(\tilde{\ell}^p, \tilde{d}_p)$  such as maxima and sums. Following the shift-projection ideas in Janssen [22], jointly with continuous mapping arguments, we extend inference to functionals on  $(\ell^p, d_p)$ .

**6.1. Inference for cluster functionals in  $(\ell^p, d_p)$ .** Let  $g : (\ell^p, d_p) \rightarrow \mathbb{R}$  be a bounded measurable function. We define the functional  $\psi_g : (\ell^p, \tilde{d}_p) \rightarrow \mathbb{R}$  by

$$(6.1) \quad [\mathbf{z}] \mapsto \psi_g([\mathbf{z}]) := \sum_{j \in \mathbb{Z}} |\mathbf{z}_{-j}^*|^\alpha g((B^j \mathbf{z}_t^*)_{t \in \mathbb{Z}}),$$

where  $\mathbf{z}_t^* := \mathbf{z}_{t-T^*(\mathbf{z})}$ , for  $t \in \mathbb{Z}$ , such that  $T^*(\mathbf{z}) := \inf\{s \in \mathbb{Z} : |\mathbf{z}_s| = \|\mathbf{z}\|_\infty\}$  and  $B : \ell^p \rightarrow \ell^p$  is the backward-shift map.

We link the distribution of the spectral cluster process  $\mathbf{Q}^{(\alpha)}$  from Equation (3.2) and the distribution of the class  $[\mathbf{Q}^{(\alpha)}]$  through the mappings (6.1) in the next proposition whose proof is given in Section 8.5.1.

**Proposition 6.1.** *The following relation holds for every real-valued bounded measurable function  $g$  on  $\ell^\alpha$*

$$\mathbb{E}[g(\mathbf{Q}^{(\alpha)})] = \mathbb{E}[\psi_g([\mathbf{Q}^{(\alpha)}])],$$

where  $\psi_g$  is as in (6.1). This relation remains valid if  $\alpha$  is replaced by  $p$ , whenever the spectral cluster process in  $\ell^p$  is well defined.

For  $p \leq \alpha$  the mappings in (6.1) are continuous functionals on  $(\tilde{\ell}^p, \tilde{d}_p)$  and we can extend Theorem 4.2 to continuous functionals on  $(\ell^p, d_p)$  evaluated at the spectral cluster process  $\mathbf{Q}^{(p)}$  taking values in  $(\ell^p, d_p)$ .

**Theorem 6.2.** *Assume the conditions of Theorem 4.2 for  $p \leq \alpha$ . Then for any continuous bounded function  $g : \ell^p \cap \{\mathbf{x} : \|\mathbf{x}\|_p = 1, |\mathbf{x}_0| > 0\} \rightarrow \mathbb{R}$ ,*

$$(6.2) \quad \hat{g}^{(p)} := \frac{1}{k} \sum_{t=1}^m \underbrace{\sum_{j=1}^b W_{j,t}(p) g\left(\frac{B^{j-1} \mathbf{B}_t}{\|\mathbf{B}_t\|_p}\right)}_{=: \psi_g(\mathbf{B}_t / \|\mathbf{B}_t\|_p)} \mathbb{1}(\|\mathbf{B}_t\|_p > \|\mathbf{B}\|_{p,(k+1)})$$

$$\xrightarrow{\mathbb{P}} \mathbb{E}[g(\mathbf{Q}^{(p)})], \quad n \rightarrow \infty,$$

where  $W_{j,t}(p) = |\mathbf{X}_{(t-1)b+j}|^\alpha / \|\mathbf{B}_t\|_p^\alpha$  for all  $j = 1, \dots, b$ .

The proof is given in Section 8.5.2.

**6.2. Applications.** Examples of non-shift-invariant functionals on  $(\ell^p, d_p)$  are measures of serial dependence, probabilities of large deviations such as the supremum of a random walk and ruin probabilities, and functionals of the spectral tail process  $\Theta$ . We study these examples in the remainder of this section.

**6.2.1. Measures of serial dependence.** Define  $g_h(\mathbf{x}_t) = |\mathbf{x}_h|^\alpha \frac{\mathbf{x}_0^\top}{|\mathbf{x}_0|} \frac{\mathbf{x}_h}{|\mathbf{x}_h|}$ . Then the following result is straightforward from Theorem 6.2.

**Corollary 6.3.** *Assume the conditions of Theorem 6.2 for  $p = \alpha$ . Then*

$$\hat{g}_h^{(\alpha)} := \frac{1}{k} \sum_{t=1}^m \underbrace{\sum_{j=1}^{b-h} W_{j,t} W_{j+h,t} \frac{\mathbf{X}_{j,t}^\top}{|\mathbf{X}_{j,t}|} \frac{\mathbf{X}_{j+h,t}}{|\mathbf{X}_{j+h,t}|}}_{=: \psi_{g_h}(\mathbf{B}_t / \|\mathbf{B}_t\|_p)} \mathbb{1}(\|\mathbf{B}_t\|_\alpha > \|\mathbf{B}\|_{\alpha,(k+1)}).$$

$$\xrightarrow{\mathbb{P}} \mathbb{E}[g_h(\mathbf{Q}^{(a)})], \quad n \rightarrow +\infty,$$

where the weights  $W_{j,t} = W_{j,t}(\alpha)$  are defined in Theorem 6.2, satisfying  $\sum_{j=1}^b W_{j,t} = 1$ , and  $\mathbf{X}_{j,t} := \mathbf{X}_{(t-1)b+j}$  for  $j = 1, \dots, b$ .

The function  $g_h$  gives a summary of the magnitude and direction of the time series  $h$  lags after recording a high-level exceedance of the norm, and satisfies the relation  $\sum_{h \in \mathbb{Z}} \mathbb{E}[g_h(\mathbf{Q}^{(\alpha)})] = 1$ .

**Example 6.4.** Let  $(X_t)$  be a linear process satisfying the assumptions in Example 7, then

$$\mathbb{E}[g_h(\mathbf{Q}^{(\alpha)})] = \frac{\sum_{t \in \mathbb{Z}} |\varphi_t|^\alpha |\varphi_{t+h}|^\alpha \text{sign}(\varphi_t) \text{sign}(\varphi_{t+h})}{(\|\varphi\|_\alpha^\alpha)^2}, \quad h \in \mathbb{Z}.$$

This function is proportional to the autocovariance function of a finite variance linear process with coefficients  $(|\varphi_t|^\alpha \text{sign}(\varphi_t))$ . In particular, for  $\alpha = 1$  it is proportional to the autocovariance function of a finite variance linear process with coefficients  $(\varphi_t)$ .

**6.2.2. Large deviations for the supremum of a random walk.** We start by reviewing Theorem 4.5 in Mikosch and Wintenberger [29]; the proof is given in Section 8.5.3.

**Proposition 6.5.** *Consider a univariate stationary sequence  $(X_t)$  satisfying  $\mathbf{RV}_\alpha$  for some  $\alpha \geq 1$ ,  $\mathbf{AC}$ ,  $\mathbf{CS}_1$ , and  $c(1) < \infty$ . Then for all  $p \geq 1$ ,*

$$(6.3) \quad \left| \frac{\mathbb{P}(\sup_{1 \leq t \leq n} S_t > x_n)}{n\mathbb{P}(|X_1| > x_n)} - c(p) \mathbb{E} \left[ \lim_{s \rightarrow \infty} \left( \sup_{t \geq -s} \sum_{i=-s}^t Q_i^{(p)} \right)_+^\alpha \right] \right| \rightarrow 0, \quad n \rightarrow \infty.$$

If  $\alpha \geq 1$ , then  $\|Q^{(1)}\|_\alpha^\alpha \leq \|Q^{(1)}\|_1^\alpha = 1$  and a consistent estimator of  $c(1) = 1/\mathbb{E}[\|Q^{(1)}\|_\alpha^\alpha]$  was suggested in Section 4.3.2. A consistent estimator of the term in (6.3) is given next.

**Corollary 6.6.** *Assume the conditions of Theorem 6.2 for  $p = 1$ . Then*

$$\left| \frac{\sum_{t=1}^m \left( \sup_{1 \leq j \leq b} \frac{X_{t,j}}{\|\mathbf{B}_t\|_1} \right)_+^\alpha \mathbb{1}(\|\mathbf{B}_t\|_1 > \|\mathbf{B}\|_{1,(k+1)})}{\sum_{t=1}^m \frac{\|\mathbf{B}_t\|_\alpha^\alpha}{\|\mathbf{B}_t\|_1^\alpha} \mathbb{1}(\|\mathbf{B}_t\|_1 > \|\mathbf{B}\|_{1,(k+1)})} - c(1) \mathbb{E} \left[ \lim_{s \rightarrow \infty} \left( \sup_{t \geq -s} \sum_{i=-s}^t Q_i^{(1)} \right)_+^\alpha \right] \right| \xrightarrow{\mathbb{P}} 0, \quad n \rightarrow \infty,$$

where  $X_{t,j} := X_{(t-1)b+j}$ , for  $1 \leq j \leq b$ ,  $1 \leq t \leq m$ .

Following the same ideas and using Theorem 4.9 in [29], one can also derive a consistent estimator for the constant in the related ruin problem.

**6.2.3. Application: a cluster-based method for inference on  $(\Theta_t)$ .** Exploiting the relation  $(\mathbf{Q}_t^{(\alpha)})/|\mathbf{Q}_0^{(\alpha)}| \stackrel{d}{=} (\Theta_t)$  discussed in Section 3.2, we propose cluster-based estimation methods for the spectral tail process.

Cluster-based approaches with the goal to improve inference on  $\Theta_1$  for Markov chains were considered in Drees *et al.* [17]; see also Davis *et al.* [10] and Drees *et al.* [14] for related cluster-based procedures on  $(\Theta_t)_{|t| \leq h}$  for fixed  $h \geq 0$ . Our approach can be seen as an extension for inference on the  $\ell^\alpha$ -valued sequence  $(\Theta_t)$ .

Consider the continuous re-normalization function  $\zeta(\mathbf{x}) = \mathbf{x}/|\mathbf{x}_0|$  on  $\{\mathbf{x} \in \ell^\alpha : |\mathbf{x}_0| > 0\}$ . We derive the following result from Theorem 6.2; the proof is given in Section 8.5.4.

**Proposition 6.7.** *Assume the conditions of Theorem 6.2 for  $p = \alpha$ . Let  $\rho : (\ell^\alpha, d_\alpha) \rightarrow \mathbb{R}$  be a homogeneous continuous function and  $\rho_\zeta(\mathbf{x}) := (\rho^\alpha \wedge 1) \circ \zeta(\mathbf{x})$ . Then for  $k = k_n \rightarrow \infty$ ,*

$$\hat{\rho}_\zeta^{(\alpha)} := \frac{1}{k} \sum_{t=1}^m \psi_{\rho_\zeta}(\mathbf{B}_t) \mathbb{1}(\|\mathbf{B}_t\|_p > \|\mathbf{B}\|_{p,(k+1)}) \xrightarrow{\mathbb{P}} \mathbb{P}(\rho(Y \Theta) > 1), \quad n \rightarrow \infty,$$

where  $\psi_{\rho_\zeta}(\mathbf{B}_t)$  is defined in (6.2) and the Pareto( $\alpha$ ) random variable  $Y$  is independent of  $\Theta$ .

Classical examples of such functionals are  $\rho(\mathbf{x}) = \max_{i \geq 0, j \geq i} (x_i - x_j)_+$ , functionals related to large deviations such as  $\rho(\mathbf{x}) = \sup_{t \geq 0} (\sum_{i=0}^t x_i)_+$ , or measures of serial dependence such as  $\rho(\mathbf{x}) = |\mathbf{x}_h|$ .

## 7. CLUSTER INFERENCE IMPLEMENTATION FOR REGULARLY VARYING LINEAR PROCESS

In this section we illustrate the index estimators of Corollaries 4.3 and 4.5 for a regularly varying linear process

$$X_t := \sum_{j \in \mathbb{Z}} \varphi_j Z_{t-j}, \quad t \in \mathbb{Z},$$

where  $(Z_t)$  is an iid real-valued regularly varying sequence with (tail)-index  $\alpha > 0$ , and  $(\varphi_j)$  are real coefficients such that  $\sum_{j \in \mathbb{Z}} |\varphi_j|^{1 \wedge (\alpha - \varepsilon)} < \infty$  for some  $\varepsilon > 0$ .

In this setting,  $(X_t)$  is regularly varying with the same (tail)-index  $\alpha > 0$ , and the distributions of  $Z_t$  and  $X_t$  are tail-equivalent; see Davis and Resnick [12]. The spectral cluster process of  $(X_t)$  is given by  $Q_t^{(\alpha)} = (\varphi_{t+J} / \|(\varphi_t)\|_\alpha) \Theta_0^Z$ ,  $t \in \mathbb{Z}$ , where  $\lim_{x \rightarrow \infty} \mathbb{P}(\pm Z_0 > x) / \mathbb{P}(|Z_0| > x) = \mathbb{P}(\Theta_0^Z = \pm 1)$ ,  $\Theta_0^Z$  is independent of a random shift  $J$  with distribution  $\mathbb{P}(J = j) = |\varphi_j|^\alpha / \|(\varphi_t)\|_\alpha^\alpha$ ; see Kulik and Soulier [23], (15.3.9). Then

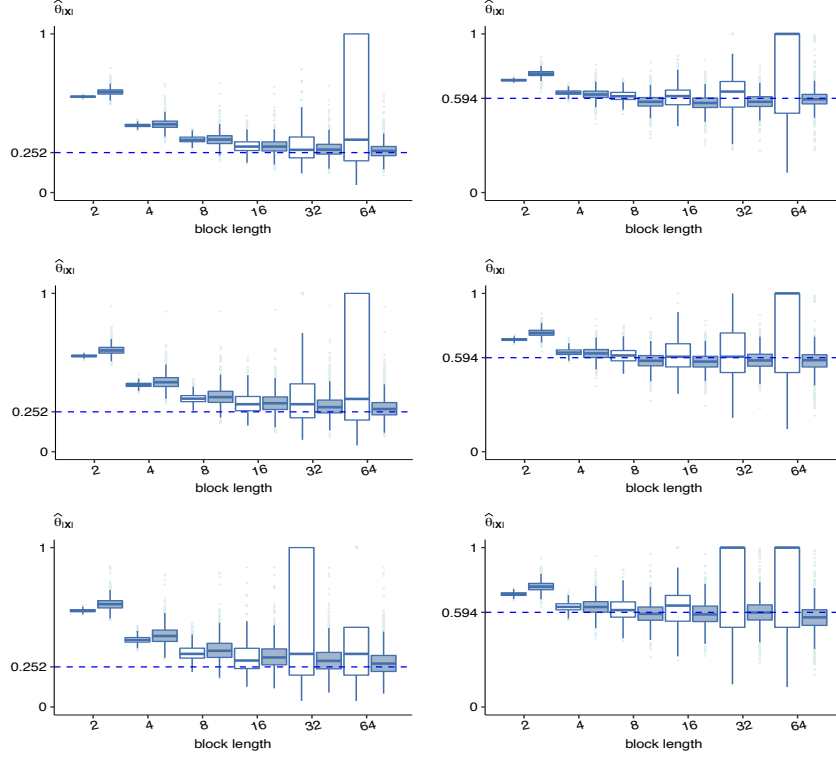
$$c(\infty) = \max_{t \in \mathbb{Z}} |\varphi_t|^\alpha / \|\varphi\|_\alpha^\alpha, \quad c(1) = (\sum_{t \in \mathbb{Z}} |\varphi_t|^\alpha) / \|\varphi\|_\alpha^\alpha.$$

For the causal AR(1) model given by  $X_t = \varphi X_{t-1} + Z_t$ ,  $t \in \mathbb{Z}$ ,  $|\varphi| < 1$ , one retrieves  $\theta_{|X|} = c(\infty) = 1 - |\varphi|^\alpha$  and  $c(1) = (1 - |\varphi|^\alpha) / (1 - |\varphi|)^\alpha$ .

We aim to illustrate the estimators of  $\theta_{|X|}$  and  $c(1)$  built on extremal  $\ell^\alpha$ -blocks for the causal AR(1) model with student( $\alpha$ ) noise. Guided by (4.3), we take  $k = k_n = \lfloor n/b_n^2 \rfloor$  as

$$k_n = \lfloor m_n \mathbb{P}(\|\mathbf{B}\|_\alpha > x_{b_n}) \rfloor \sim n \mathbb{P}(|\mathbf{X}_0| > x_{b_n}) = o(n/b_n^{1+\kappa}),$$

for  $\kappa > 0$  sufficiently small using the Potter bound. For estimation of  $\alpha$ , we follow the bias-correction procedure in de Haan *et al.* [19]. This estimator is plugged into (4.4), (4.6), resulting in the estimators  $\hat{\theta}_{|X|}$ ,  $\hat{c}(1)$ , as a function of block lengths. Figures 7.1 and 7.2 present boxplots (in blue) of these estimators as a function of  $b_n$  and for different sample sizes  $n$ . For comparison, we also show boxplots (in white) of the estimators in (4.5) and (4.7) based on extremal  $\ell^\infty$ -blocks. Inference based on  $\ell^\alpha$ -block, coupled with a Hill-type estimate of  $\alpha$ , seems to be robust compared to the  $\ell^\infty$ -blocks approach. In all examples the block length  $b = 32$  gives nice results for the  $\ell^\alpha$ -approach in terms of bias and dispersion. Instead, the  $\ell^\infty$ -estimator appears to be highly sensitive to the block length choice. Also, notice that the bias for large block lengths decreases as  $n$  increases. Indeed, if we fix  $n$ , the relation  $\lfloor n/b^2 \rfloor \rightarrow 0$  as  $b \rightarrow \infty$  restricts the block length for small sample sizes. We also refer to Buriticá *et al.* [7] for further simulation experiences showing that the estimator of the extremal index in (4.4) compares favorably with various classical estimators as regards bias.



**Figure 7.1.** Boxplot of estimates  $\hat{\theta}_{|X|}$  as a function of  $b_n$  from (4.4) for inference through  $\mathbf{Q}^{(\alpha)}$  (in blue) and from (4.5) through  $\mathbf{Q}^{(\infty)}$  (in white). 1 000 simulated samples  $(X_t)_{t=1,\dots,n}$  from a causal AR(1) model with student( $\alpha$ ) noise with  $\alpha = 1.3$  and  $\varphi = 0.8$  (left column),  $\varphi = 0.5$  (right column) were considered. Rows correspond to results for  $n = 8\,000, 4\,000, 2\,000$  from top to bottom.

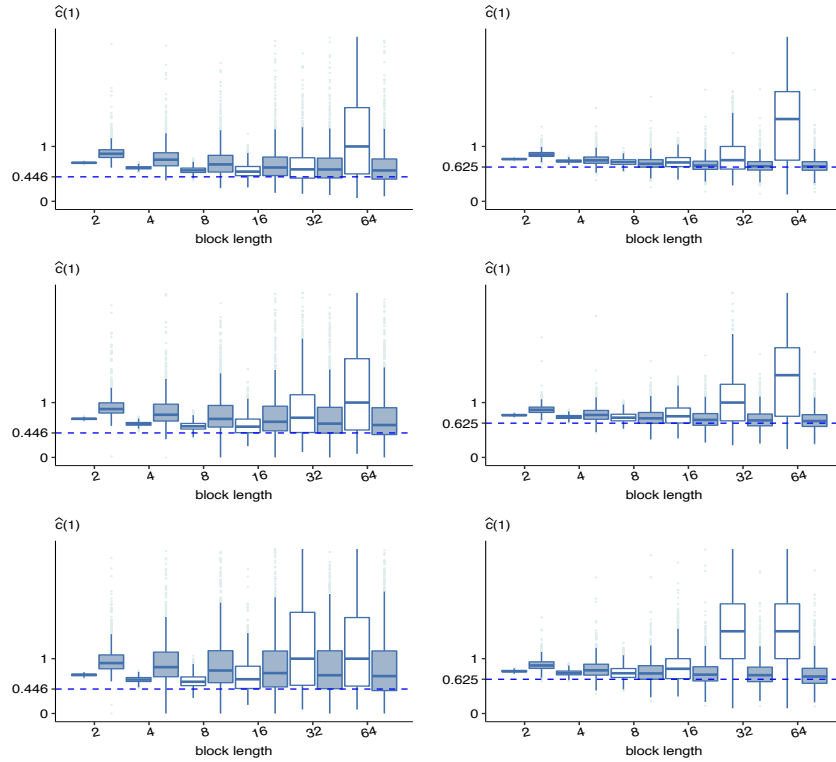
## 8. PROOFS

**8.1. Proof of Theorem 2.1.** Recall the properties of the sequence  $(x_n)$  from Section 5, in particular  $n \mathbb{P}(|\mathbf{X}_0| > x_n) \rightarrow 0$ . The main result in Theorem 2.1 follows by applications of Lemma 8.1 and Proposition 8.2 below; their proofs are given at the end of this section.

**Lemma 8.1.** *Consider an  $\mathbb{R}^d$ -valued stationary time series  $(\mathbf{X}_t)$  satisfying the conditions  $\mathbf{RV}_\alpha$ ,  $\mathbf{AC}$ ,  $\mathbf{CS}_p$ . If  $p < \alpha$ , assume also  $n/x_n^p \rightarrow 0$  and, if  $p = \alpha$ ,  $n/x_n^{\alpha-\kappa} \rightarrow 0$  for some  $\kappa > 0$ . Then the following relation holds*

$$(8.1) \quad \lim_{n \rightarrow \infty} \frac{\mathbb{P}(\|\mathbf{X}_{[0,n]}\|_p > x_n)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} = c(p),$$

where  $c(p)$  is given in (3.1).



**Figure 7.2.** Boxplot of estimates  $\hat{c}(1)$  as a function of  $b_n$  from (4.6) for inference through  $\mathbf{Q}^{(\alpha)}$  (in blue) and from (4.7) for inference through  $\mathbf{Q}^{(\infty)}$  (in white). We simulate 1000 samples  $(X_t)_{t=1,\dots,n}$  from an AR(1) model with student( $\alpha$ ) for  $\alpha = 0.7$  and  $\varphi = 0.8$  (left column),  $\varphi = 0.5$  (right column) were considered. Rows correspond to  $n = 8000, 4000, 2000$  from top to bottom.

We recall from Remark 5.1 that (5.2) in  $\mathbf{CS}_p$  is always satisfied for  $p > \alpha$ . Moreover, for  $p \leq \alpha$ , under the growth conditions on  $(x_n)$  in Theorem 2.1, centering with the expectation in (5.2) is not necessary.

**Proposition 8.2.** *Assume the conditions of Lemma 8.1. Then,*

$$\mathbb{P}(x_n^{-1} \mathbf{X}_{[0,n]} \in \cdot \mid \|\mathbf{X}_{[0,n]}\|_p > x_n) \xrightarrow{w} \mathbb{P}(Y \mathbf{Q}^{(p)} \in \cdot), \quad n \rightarrow \infty, \quad (8.2)$$

in the space  $(\tilde{\ell}^p, \tilde{d}_p)$  where the Pareto( $\alpha$ ) random variable  $Y$  and  $\mathbf{Q}^{(p)}$  are independent.

**Proof of Lemma 8.1.** Choose some  $\epsilon > 0$ ,  $\delta \in (0, 1)$ . Since  $\|a_n^{-1} \mathbf{X}_{[0,n]}\|_p^p$  is a sum of non-negative random variables we have the following bounds via



truncation

$$\begin{aligned}
 & \mathbb{P}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}\|_p^p > 1) \leq \mathbb{P}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}\|_p^p > 1) \\
 (8.3) \quad & \leq \mathbb{P}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}\|_p^p > (1 - \delta^p)) + \mathbb{P}(\|\overline{\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}}\|_p^p > \delta^p).
 \end{aligned}$$

By  $\mathbf{CS}_p$  and in view of Remark 5.1 we have

$$(8.4) \quad \lim_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} \mathbb{P}(\|\overline{\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}}\|_p^p > \delta^p) / (n\mathbb{P}(|\mathbf{X}_0| > x_n)) = 0.$$

Now, for any choice of  $u > 0$ , it remains to determine the limits of the terms  $\mathbb{P}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}\|_p^p > u) / (n\mathbb{P}(|\mathbf{X}_0| > x_n))$ . We start with a telescoping sum representation

$$\begin{aligned}
 & \mathbb{P}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}\|_p^p > u) - \mathbb{P}(\|\underline{x_n^{-1}\mathbf{X}_0}\|_p^p > u) \\
 &= \sum_{i=1}^n (\mathbb{P}(\|\underline{x_n^{-1}\mathbf{X}_{[0,i]_\epsilon}}\|_p^p > u) - \mathbb{P}(\|\underline{x_n^{-1}\mathbf{X}_{[0,i-1]_\epsilon}}\|_p^p > u)) \\
 &= \sum_{i=1}^n \mathbb{E}[(\mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[0,i]_\epsilon}}\|_p^p > u) - \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[1,i]_\epsilon}}\|_p^p > u)) \mathbb{1}(|\mathbf{X}_0| > \epsilon x_n)],
 \end{aligned}$$

where we used stationarity in the last step and the fact that the difference of the indicator functions vanishes on  $\{|\mathbf{X}_0| \leq \epsilon x_n\}$ . We also observe that the second term on the left-hand side is of the order  $o(n\mathbb{P}(|\mathbf{X}_0| > x_n))$ . For any fixed  $k$  write  $A_k = \{\max_{k \leq t \leq n} |\mathbf{X}_t| > \epsilon x_n\}$ . Regular variation of  $(\mathbf{X}_t)$  ensures that, as  $n \rightarrow \infty$ ,

$$\begin{aligned}
 & \frac{\mathbb{P}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}\|_p^p > u)}{n\mathbb{P}(|\mathbf{X}_0| > x_n)} \\
 & \sim \epsilon^{-\alpha} \frac{1}{n} \sum_{i=1}^n \mathbb{E}[\mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[0,i]_\epsilon}}\|_p^p > u) - \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[1,i]_\epsilon}}\|_p^p > u) \mid |\mathbf{X}_0| > \epsilon x_n] \\
 & \sim \epsilon^{-\alpha} \mathbb{E}[(\mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[0,k-1]_\epsilon}}\|_p^p > u) - \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[1,k-1]_\epsilon}}\|_p^p > u)) \\
 & \quad \times \mathbb{1}(A_k^c) \mid |\mathbf{X}_0| > \epsilon x_n] + \epsilon^{-\alpha} O(\mathbb{P}(A_k \mid |\mathbf{X}_0| > \epsilon x_n)),
 \end{aligned}$$

where the second term vanishes, first letting  $n \rightarrow \infty$  and then  $k \rightarrow \infty$ , by virtue of  $\mathbf{AC}$ . Now the regular variation property of  $(\mathbf{X}_t)$  implies that

$$\begin{aligned}
 & \lim_{n \rightarrow \infty} \frac{\mathbb{P}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}\|_p^p > u)}{n\mathbb{P}(|\mathbf{X}_0| > x_n)} \\
 &= \lim_{k \rightarrow \infty} \epsilon^{-\alpha} (\mathbb{P}(\sum_{t=0}^{k-1} |\epsilon Y \boldsymbol{\Theta}_t|^p \mathbb{1}(|Y \boldsymbol{\Theta}_t| > 1) > u) \\
 & \quad - \mathbb{P}(\sum_{t=1}^{k-1} |\epsilon Y \boldsymbol{\Theta}_t|^p \mathbb{1}(|Y \boldsymbol{\Theta}_t| > 1) > u)),
 \end{aligned}$$

by a change of variable this term equals

$$\begin{aligned}
 &= \lim_{k \rightarrow \infty} \mathbb{E}[\int_{\epsilon}^{\infty} (\mathbb{1}(\sum_{t=0}^{k-1} |y \boldsymbol{\Theta}_t|^p \mathbb{1}(|y \boldsymbol{\Theta}_t| > \epsilon) > u) \\
 & \quad - \mathbb{1}(\sum_{t=1}^{k-1} |y \boldsymbol{\Theta}_t|^p \mathbb{1}(|y \boldsymbol{\Theta}_t| > \epsilon) > u)) d(-y^{-\alpha})] \\
 &= \lim_{k \rightarrow \infty} \mathbb{E}[\int_0^{\infty} (\mathbb{1}(\|\underline{y \boldsymbol{\Theta}_{[0,k-1]_\epsilon}}\|_p^p > u) - \mathbb{1}(\|\underline{y \boldsymbol{\Theta}_{[1,k-1]_\epsilon}}\|_p^p > u)) d(-y^{-\alpha})].
 \end{aligned}$$

In the last step we used the fact that the integrand vanishes for  $0 \leq y \leq \epsilon$ . The integrand is non-negative and bounded by  $\mathbb{1}(y > \epsilon)$  which is integrable. Thus we may take the limit as  $k \rightarrow \infty$  inside the integral to derive the quantity

$$\mathbb{E} \left[ \int_0^\infty \left( \mathbb{1}(\|y \underline{\Theta}_{[0,\infty]_\epsilon}\|_p^p > u) - \mathbb{1}(\|y \underline{\Theta}_{[1,\infty]_\epsilon}\|_p^p > u) \right) d(-y^{-\alpha}) \right].$$

By monotone convergence as  $\epsilon \downarrow 0$  we get the limit

$$(8.5) \quad u^{-\alpha/p} \mathbb{E} [\|\Theta_{[0,\infty]}\|_p^\alpha - \|\Theta_{[1,\infty]}\|_p^\alpha] = u^{-\alpha/p} c(p).$$

An application of this formula and a telescoping sum argument yield

$$\begin{aligned} (8.6) \quad & \mathbb{E} \left[ \|\Theta_{t \geq 0}\|_p^\alpha - \|\Theta_{t \geq 1}\|_p^\alpha \right] \\ &= \mathbb{E} \left[ \|\Theta\|_\alpha^\alpha (\|\Theta_{t \geq 0}\|_\alpha^\alpha - \|\Theta_{t \geq 1}\|_\alpha^\alpha) \right] \\ &= \sum_{s \in \mathbb{Z}} \mathbb{E} \left[ |\Theta_s|^\alpha (\|\Theta_{t \geq 0}\|_\alpha^\alpha - \|\Theta_{t \geq 1}\|_\alpha^\alpha) \right] \\ &= \sum_{s \in \mathbb{Z}} \mathbb{E} \left[ (\|\Theta_{t \geq -s}\|_\alpha^\alpha - \|\Theta_{t \geq -s+1}\|_\alpha^\alpha) \right] \\ &= \mathbb{E} [\|\Theta\|_p^\alpha / \|\Theta\|_\alpha^\alpha] = c(p). \end{aligned}$$

Now an appeal to (8.3) with  $u = 1$  and  $u = 1 - \delta^p$  yields

$$\begin{aligned} c(p) &\leq \liminf_{n \rightarrow \infty} \frac{\mathbb{P}(\|x_n^{-1} \mathbf{X}_{[0,n]}\|_p^p > 1)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} \\ &\leq \limsup_{n \rightarrow \infty} \frac{\mathbb{P}(\|x_n^{-1} \mathbf{X}_{[0,n]}\|_p^p > 1)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} \leq (1 - \delta^p)^{-\alpha/p} c(p). \end{aligned}$$

The limit relation (8.1) follows as  $\delta \downarrow 0$ .

**Proof of Proposition 8.2.** Consider any bounded Lipschitz-continuous function  $f : (\tilde{\ell}^p, \tilde{d}_p) \rightarrow \mathbb{R}$ . The statement is proved if we can show that

$$\begin{aligned} & \lim_{n \rightarrow \infty} \mathbb{E}[f(x_n^{-1} \mathbf{X}_{[0,n]}) \mid \|\mathbf{X}_{[0,n]}\|_p > x_n] \\ &= c(p)^{-1} \mathbb{E} [\|\Theta / \|\Theta\|_\alpha\|_p^\alpha f(Y \Theta / \|\Theta\|_p)] . \end{aligned}$$

In view of Lemma 8.1 it suffices to show

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{\mathbb{E}[f(x_n^{-1} \mathbf{X}_{[0,n]}) \mathbb{1}(\|\mathbf{X}_{[0,n]}\|_p > x_n)]}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} \\ &= \mathbb{E} [\|\Theta / \|\Theta\|_\alpha\|_p^\alpha f(Y \Theta / \|\Theta\|_p)] . \end{aligned}$$

In these limit relations we may replace  $f(x_n^{-1}\mathbf{X}_{[0,n]})$  by  $f(\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon)$  since by (8.4) for any  $\delta > 0$ , some  $K_f > 0$ ,

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} \frac{\mathbb{P}(|f(x_n^{-1}\mathbf{X}_{[0,n]}) - f(\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon)| > \delta, \|\mathbf{X}_{[0,n]}\|_p > x_n)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} \\ & \leq \lim_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} \frac{\mathbb{P}(K_f d_p(\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon, \mathbf{0}) > \delta)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} \\ & \leq \lim_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} \frac{\mathbb{P}(K_f \|\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon\|_p^p > \delta(p))}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} = 0, \end{aligned}$$

where  $\delta(p) = \delta^p$  for  $p \geq 1$  and  $= \delta$  for  $p \in (0, 1)$ . We also have for  $\delta \in (0, 1)$ ,

$$\begin{aligned} G_{n,\epsilon} &= \frac{\mathbb{E}[f(\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon) | \mathbb{1}(\|\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon\|_p > 1) - \mathbb{1}(\|\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon\|_p > 1)]}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} \\ &\leq c \frac{\mathbb{P}(\|\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon\|_p > 1 \geq \|\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon\|_p)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} \\ &\leq c \frac{\mathbb{P}(\|\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon\|_p > \delta)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} + c \frac{\mathbb{P}(1 \geq \|\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon\|_p > 1 - \delta)}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} \\ &= G_{n,\epsilon,\delta}^{(1)} + G_{n,\epsilon,\delta}^{(2)}. \end{aligned}$$

Applying (8.4) to  $G_{n,\epsilon,\delta}^{(1)}$  and using the calculations in the proof of Lemma 8.1 leading to (8.5) for  $G_{n,\epsilon,\delta}^{(2)}$ , we conclude that

$$\begin{aligned} \lim_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} G_{n,\epsilon} &\leq \lim_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} G_{n,\epsilon,\delta}^{(1)} + \lim_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} G_{n,\epsilon,\delta}^{(2)} \\ &= 0 + c((1 - \delta)^{-\alpha/p} - 1) \downarrow 0, \quad \delta \downarrow 0. \end{aligned}$$

Thus it suffices to show

$$\begin{aligned} & \lim_{\epsilon \downarrow 0} \lim_{n \rightarrow \infty} \frac{\mathbb{E}[f(\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon) \mathbb{1}(\|\overline{x_n^{-1}\mathbf{X}_{[0,n]}}_\epsilon\|_p > 1)]}{n \mathbb{P}(|\mathbf{X}_0| > x_n)} \\ (8.7) \quad &= \mathbb{E}[\|\boldsymbol{\Theta}/\|\boldsymbol{\Theta}\|_\alpha\|_p^\alpha f(Y \boldsymbol{\Theta}/\|\boldsymbol{\Theta}\|_p)] . \end{aligned}$$

This is the goal of the remaining proof.

Choose any  $\epsilon > 0$ . Noticing that  $\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}} = \underline{x_n^{-1}\mathbf{X}_{[1,n]_\epsilon}}$  on  $\{|x_n^{-1}\mathbf{X}_0| \leq \epsilon\}$ , we have

$$\begin{aligned}
I &:= \mathbb{E}[f(\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}\|_p > 1)] \\
&= \mathbb{E}\left[\left(f(\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}\|_p > 1) \right. \right. \\
&\quad \left. \left. - f((0, \underline{x_n^{-1}\mathbf{X}_{[1,n]_\epsilon}})) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[1,n]_\epsilon}\|_p > 1)\right) \mathbb{1}(|\mathbf{X}_0| > \epsilon x_n)\right] \\
&\quad + \mathbb{E}[f((0, \underline{x_n^{-1}\mathbf{X}_{[1,n]_\epsilon}})) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[1,n]_\epsilon}\|_p > 1)] \\
&= \mathbb{E}\left[\left(f(\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}}) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n]_\epsilon}\|_p > 1) \right. \right. \\
&\quad \left. \left. - f((0, \underline{x_n^{-1}\mathbf{X}_{[1,n]_\epsilon}})) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[1,n]_\epsilon}\|_p > 1)\right) \mathbb{1}(|\mathbf{X}_0| > \epsilon x_n)\right] \\
&\quad + \mathbb{E}[f((0, \underline{x_n^{-1}\mathbf{X}_{[0,n-1]_\epsilon}})) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[0,n-1]_\epsilon}\|_p > 1)] .
\end{aligned}$$

where we used the stationarity in the last step. Using the same idea recursively, we obtain

$$\begin{aligned}
I &= \sum_{j=1}^n \mathbb{E}\left[\left(f((\mathbf{0}^{n-j}, \underline{x_n^{-1}\mathbf{X}_{[0,j]_\epsilon}})) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[0,j]_\epsilon}\|_p > 1) \right. \right. \\
&\quad \left. \left. - f((\mathbf{0}^{n-j+1}, \underline{x_n^{-1}\mathbf{X}_{[1,j]_\epsilon}})) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[1,j]_\epsilon}\|_p > 1)\right) \mathbb{1}(|\mathbf{X}_0| > \epsilon x_n)\right] \\
&\quad + \mathbb{E}[f((\mathbf{0}^n, \underline{x_n^{-1}\mathbf{X}_0})) \mathbb{1}(|\mathbf{X}_0| > \epsilon x_n)] ,
\end{aligned}$$

where  $\mathbf{0}^k := \{0\}^k$  for  $k \geq 1$ . By regular variation of  $\mathbf{X}_0$  the last right-hand term is  $o(n\mathbb{P}(|\mathbf{X}_0| > x_n))$ . Therefore by regular variation of  $(\mathbf{X}_t)$  we obtain as  $n \rightarrow \infty$ ,

$$\begin{aligned}
&I / (n\mathbb{P}(|\mathbf{X}_0| > x_n)) \\
&\sim \frac{\epsilon^{-\alpha}}{n} \sum_{j=1}^n \mathbb{E}\left[\left(f((\mathbf{0}^{n-j}, \underline{x_n^{-1}\mathbf{X}_{[0,j]_\epsilon}})) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[0,j]_\epsilon}\|_p > 1) \right. \right. \\
&\quad \left. \left. - f((\mathbf{0}^{n-j+1}, \underline{x_n^{-1}\mathbf{X}_{[1,j]_\epsilon}})) \mathbb{1}(\|\underline{x_n^{-1}\mathbf{X}_{[1,j]_\epsilon}\|_p > 1)\right) \mathbb{1}(|\mathbf{X}_0| > \epsilon x_n)\right] \\
&=: II .
\end{aligned}$$

Write  $A_k = \{\|\mathbf{X}_{[k,n]}\|_\infty > \epsilon x_n\}$  for fixed  $k \geq 1$ . By **AC**,  $\mathbb{P}(A_k \mid |\mathbf{X}_0| > \epsilon x_n)$  vanishes by first letting  $n \rightarrow \infty$  then  $k \rightarrow \infty$ . Since each of the summands in II is uniformly bounded in absolute value we may restrict the summation

to  $j \in \{k-1, \dots, n\}$  for any fixed  $k \geq 1$ . Therefore we have as  $n \rightarrow \infty$ ,

$$\begin{aligned}
II &= O(\mathbb{P}(A_k \mid |\mathbf{X}_0| > \epsilon x_n)) \\
&\sim \frac{\epsilon^{-\alpha}}{n} \sum_{j=k-1}^n \mathbb{E}[f(\underline{(\mathbf{0}^{n-j}, x_n^{-1} \mathbf{X}_{[0,j]_\epsilon})}) \mathbb{1}(\{\|x_n^{-1} \mathbf{X}_{[0,j]_\epsilon}\|_p > 1\}) \\
&\quad - f(\underline{(\mathbf{0}^{n-j+1}, x_n^{-1} \mathbf{X}_{[1,j]_\epsilon})}) \mathbb{1}(\{\|x_n^{-1} \mathbf{X}_{[1,j]_\epsilon}\|_p > 1\}) \mid |\mathbf{X}_0| > \epsilon x_n] \\
&= \frac{\epsilon^{-\alpha}}{n} \sum_{j=k-1}^n \mathbb{E}[f(\underline{(\mathbf{0}^{n-j}, x_n^{-1} \mathbf{X}_{[0,k-1]_\epsilon}, \mathbf{0}^{j-k})}) \mathbb{1}(\{\|x_n^{-1} \mathbf{X}_{[0,k-1]_\epsilon}\|_p > 1\}) \\
&\quad - f(\underline{(\mathbf{0}^{n-j+1}, x_n^{-1} \mathbf{X}_{[1,k-1]_\epsilon}, \mathbf{0}^{j-k})}) \mathbb{1}(\{\|x_n^{-1} \mathbf{X}_{[1,k-1]_\epsilon}\|_p > 1\}) \mid |\mathbf{X}_0| > \epsilon x_n].
\end{aligned}$$

Next we apply shift-invariance and regular variation in  $\tilde{\ell}^p$ :

$$\begin{aligned}
&\sim \epsilon^{-\alpha} \mathbb{E}[f(\underline{x_n^{-1} \mathbf{X}_{[0,k-1]_\epsilon})} \mathbb{1}(\{\|x_n^{-1} \mathbf{X}_{[0,k-1]_\epsilon}\|_p > 1\}) \\
&\quad - f(\underline{x_n^{-1} \mathbf{X}_{[1,k-1]_\epsilon})} \mathbb{1}(\{\|x_n^{-1} \mathbf{X}_{[1,k-1]_\epsilon}\|_p > 1\}) \mid |\mathbf{X}_0| > \epsilon x_n] \\
&\rightarrow \epsilon^{-\alpha} \mathbb{E}[f(\underline{\epsilon Y \boldsymbol{\Theta}_{[0,k-1]_\epsilon})} \mathbb{1}(\{\|\epsilon Y \boldsymbol{\Theta}_{[0,k-1]_\epsilon}\|_p > 1\}) \\
&\quad - f(\underline{\epsilon Y \boldsymbol{\Theta}_{[1,k-1]_\epsilon})} \mathbb{1}(\{\|\epsilon Y \boldsymbol{\Theta}_{[1,k-1]_\epsilon}\|_p > 1\})] \\
&= \mathbb{E}[f(\underline{\epsilon Y \boldsymbol{\Theta}_{[0,k-1]_1})} \mathbb{1}(\{\|\epsilon Y \boldsymbol{\Theta}_{[0,k-1]_1}\|_p > 1\}) \\
&\quad - f(\underline{\epsilon Y \boldsymbol{\Theta}_{[1,k-1]_1})} \mathbb{1}(\{\|\epsilon Y \boldsymbol{\Theta}_{[1,k-1]_1}\|_p > 1\})] =: J_{k,\epsilon}.
\end{aligned}$$

By Proposition 3.2 we have  $\|\boldsymbol{\Theta}\|_\alpha < \infty$  a.s.,  $|\boldsymbol{\Theta}_t| \xrightarrow{\text{a.s.}} 0$  as  $|t| \rightarrow \infty$ , hence  $T := \inf_{t \geq 0} \{t : Y|\boldsymbol{\Theta}_t| < 1\} < \infty$  a.s. Then by monotone convergence as  $k \rightarrow \infty$ ,

$$\begin{aligned}
J_{k,\epsilon} &= \epsilon^{-\alpha} \mathbb{E}[(f(\underline{\epsilon Y \boldsymbol{\Theta}_{[0,\infty]_1})} \mathbb{1}(\{\|\epsilon Y \boldsymbol{\Theta}_{[0,\infty]_1}\|_p > 1\}) \\
&\quad - f(\underline{\epsilon Y \boldsymbol{\Theta}_{[1,\infty]_1})} \mathbb{1}(\{\|\epsilon Y \boldsymbol{\Theta}_{[1,\infty]_1}\|_p > 1\})) \mathbb{1}(T < k)] \\
&\quad + O(\mathbb{P}(T \geq k)) \\
&\rightarrow \epsilon^{-\alpha} \mathbb{E}[f(\underline{\epsilon Y \boldsymbol{\Theta}_{[0,\infty]_1})} \mathbb{1}(\{\|\epsilon Y \boldsymbol{\Theta}_{[0,\infty]_1}\|_p > 1\}) \\
&\quad - f(\underline{\epsilon Y \boldsymbol{\Theta}_{[1,\infty]_1})} \mathbb{1}(\{\|\epsilon Y \boldsymbol{\Theta}_{[1,\infty]_1}\|_p > 1\})] \\
&= \int_0^\infty \mathbb{E}[f(\underline{y \boldsymbol{\Theta}_{[0,\infty]_\epsilon})} \mathbb{1}(\{\|y \boldsymbol{\Theta}_{[0,\infty]_\epsilon}\|_p > 1\}) \\
&\quad - f(\underline{y \boldsymbol{\Theta}_{[1,\infty]_\epsilon})} \mathbb{1}(\{\|y \boldsymbol{\Theta}_{[1,\infty]_\epsilon}\|_p > 1\})] d(-y^{-\alpha}) =: J_\epsilon.
\end{aligned}$$

In the last step we changed variables,  $u = \epsilon y$ , and observed that the integrand vanishes for  $y < \epsilon$ .

Finally, we want to let  $\epsilon \downarrow 0$ . We start by inter-changing expectation and integral in  $J_\epsilon$ , and change variables,  $u = y\|\boldsymbol{\Theta}_{[0,\infty]_\alpha}\|_\alpha$ , in the first term of the integrand and then proceed similarly for the second term with the

convention that it is zero on  $\{\|\Theta_{[1,\infty]}\|_\alpha = 0\}$ :

$$\begin{aligned}
J_\epsilon &= \mathbb{E} \left[ \int_0^\infty \left( \|\Theta_{[0,\infty]}\|_\alpha^\alpha f \left( \frac{y \Theta_{[0,\infty]}}{\|\Theta_{[0,\infty]}\|_\alpha} \right) \mathbb{1} \left( \left\| \frac{y \Theta_{[0,\infty]}}{\|\Theta_{[0,\infty]}\|_\alpha} \right\|_p > 1 \right) \right. \right. \\
&\quad \left. \left. - \|\Theta_{[1,\infty]}\|_\alpha^\alpha f \left( \frac{y \Theta_{[1,\infty]}}{\|\Theta_{[1,\infty]}\|_\alpha} \right) \mathbb{1} \left( \left\| \frac{y \Theta_{[1,\infty]}}{\|\Theta_{[1,\infty]}\|_\alpha} \right\|_p > 1 \right) \right) d(-y^{-\alpha}) \right] \\
&= \sum_{t=1}^\infty \int_0^\infty \mathbb{E} \left[ |\Theta_t|^\alpha \left( f \left( \frac{y \Theta_{[0,\infty]}}{\|\Theta_{[0,\infty]}\|_\alpha} \right) \mathbb{1} \left( \left\| \frac{y \Theta_{[0,\infty]}}{\|\Theta_{[0,\infty]}\|_\alpha} \right\|_p > 1 \right) \right. \right. \\
&\quad \left. \left. - f \left( \frac{y \Theta_{[1,\infty]}}{\|\Theta_{[1,\infty]}\|_\alpha} \right) \mathbb{1} \left( \left\| \frac{y \Theta_{[1,\infty]}}{\|\Theta_{[1,\infty]}\|_\alpha} \right\|_p > 1 \right) \right) \right] d(-y^{-\alpha}) \\
&\quad + \mathbb{E} \left[ \int_0^\infty f \left( \frac{y \Theta_{[0,\infty]}}{\|\Theta_{[0,\infty]}\|_\alpha} \right) \mathbb{1} \left( \left\| \frac{y \Theta_{[0,\infty]}}{\|\Theta_{[0,\infty]}\|_\alpha} \right\|_p > 1 \right) d(-y^{-\alpha}) \right].
\end{aligned}$$

Next we apply the time-change formula (2.2) to each summand.

$$\begin{aligned}
J_\epsilon &= \sum_{t=1}^\infty \int_0^\infty \mathbb{E} \left[ f \left( \frac{y \Theta_{[-t,\infty]}}{\|\Theta_{[-t,\infty]}\|_\alpha} \right) \mathbb{1} \left( \left\| \frac{y \Theta_{[-t,\infty]}}{\|\Theta_{[-t,\infty]}\|_\alpha} \right\|_p > 1 \right) \right. \\
&\quad \left. - f \left( \frac{y \Theta_{[1-t,\infty]}}{\|\Theta_{[1-t,\infty]}\|_\alpha} \right) \mathbb{1} \left( \left\| \frac{y \Theta_{[1-t,\infty]}}{\|\Theta_{[1-t,\infty]}\|_\alpha} \right\|_p > 1 \right) \right] d(-y^{-\alpha}) \\
&\quad + \mathbb{E} \left[ \int_0^\infty f \left( \frac{y \Theta_{[0,\infty]}}{\|\Theta_{[0,\infty]}\|_\alpha} \right) \mathbb{1} \left( \left\| \frac{y \Theta_{[0,\infty]}}{\|\Theta_{[0,\infty]}\|_\alpha} \right\|_p > 1 \right) d(-y^{-\alpha}) \right].
\end{aligned}$$

This is a telescoping sum in  $t$  with value

$$J_\epsilon = \mathbb{E} \left[ \int_0^\infty f \left( \frac{y \Theta}{\|\Theta\|_\alpha} \right) \mathbb{1} \left( \left\| \frac{y \Theta}{\|\Theta\|_\alpha} \right\|_p > 1 \right) d(-y^{-\alpha}) \right].$$

By monotone convergence we have

$$(8.8) \quad \lim_{\epsilon \downarrow 0} J_\epsilon = \mathbb{E} \left[ \int_0^\infty f \left( y \frac{\Theta}{\|\Theta\|_\alpha} \right) \mathbb{1} \left( \left\| y \frac{\Theta}{\|\Theta\|_\alpha} \right\|_p > 1 \right) d(-y^{-\alpha}) \right].$$

Combining the arguments above, we proved (8.7) as desired.  $\square$

**8.2. Proof of Lemma 5.4. The case  $p < \alpha$ .** Choose some  $\epsilon > 0$ ,  $\delta \in (0, 1)$ . We have the following bounds via truncation

$$\begin{aligned}
I_1 - I_2 &:= \mathbb{P}(\|x_n^{-1} \mathbf{X}_{[0,n]_\epsilon}\|_p^p - \mathbb{E}[\|x_n^{-1} \mathbf{X}_{[0,n]_\epsilon}\|_p^p] > 1 + \delta^p) \\
&\quad - \mathbb{P}(\|x_n^{-1} \mathbf{X}_{[0,n]_\epsilon}\|_p^p - \mathbb{E}[\|x_n^{-1} \mathbf{X}_{[0,n]_\epsilon}\|_p^p] \leq -\delta^p) \\
&\leq \mathbb{P}(\|x_n^{-1} \mathbf{X}_{[0,n]_\epsilon}\|_p^p - \mathbb{E}[\|x_n^{-1} \mathbf{X}_{[0,n]_\epsilon}\|_p^p] > 1) \\
&\leq \mathbb{P}(\|x_n^{-1} \mathbf{X}_{[0,n]_\epsilon}\|_p^p - \mathbb{E}[\|x_n^{-1} \mathbf{X}_{[0,n]_\epsilon}\|_p^p] > 1 - \delta^p) \\
&\quad + \mathbb{P}(\|x_n^{-1} \mathbf{X}_{[0,n]_\epsilon}\|_p^p - \mathbb{E}[\|x_n^{-1} \mathbf{X}_{[0,n]_\epsilon}\|_p^p] > \delta^p) =: I_3 + I_4.
\end{aligned}$$

Taking into account  $\mathbf{CS}_p$  for  $p < \alpha$ , we have

$$\lim_{\epsilon \downarrow 0} \limsup_{n \rightarrow \infty} (I_2 + I_4) / (n \mathbb{P}(|\mathbf{X}_0| > x_n)) = 0.$$

Moreover, we observe that by Karamata's theorem for  $p < \alpha$

$$\begin{aligned} \mathbb{E}[\|x_n^{-1} \mathbf{X}_{[0,n]}\|_p^p] &= n \mathbb{E}[|\mathbf{X}_0/x_n|^p \mathbf{1}(|\mathbf{X}_0| > \epsilon x_n)] \\ &= O(n \mathbb{P}(|\mathbf{X}_0| > \epsilon x_n)) = o(1). \end{aligned}$$

Thus centering in  $I_1$  and  $I_3$  is not needed, and one can follow the lines of the proof of Lemma 8.1 to conclude.

**The case  $p = \alpha$ .** It requires only slight changes; we omit details.  $\square$

### 8.3. Proofs of the results of Section 3.

8.3.1. *Proof of Proposition 3.1.* The representation (3.2) follows by identifying  $\lim_{\epsilon \downarrow 0} J_\epsilon$  as on the right-hand side of (8.8). In particular, taking  $f$  as the constant map  $(\mathbf{x}_t) \mapsto 1$  in (3.2) we obtain the representation of the constant  $c(p)$  in (3.1).

8.3.2. *Proof of Proposition 3.3.* Our goal is first to relate the sequence of spectral components  $(\mathbf{Q}^{(p)}(h))_{h \geq 0}$  to  $(\Theta_t)$ . We start with two auxiliary results whose proofs are given at the end of this section.

**Lemma 8.3.** *Let  $(\mathbf{X}_t)$  be a stationary time series satisfying  $\mathbf{RV}_\alpha$ . Then for  $h \geq 0$ ,*

$$\mathbb{P}(\mathbf{Q}^{(p)}(h) \in \cdot) = \frac{1}{c(p, h)} \sum_{k=0}^h \mathbb{E} \left[ \frac{\|\Theta_{-k+[0,h]}\|_p^\alpha}{\|\Theta_{-k+[0,h]}\|_\alpha^\alpha} \mathbf{1} \left( \frac{\Theta_{-k+[0,h]}}{\|\Theta_{-k+[0,h]}\|_p} \in \cdot \right) \right], \quad (8.9)$$

where  $c(p, h) := \sum_{k=0}^h \mathbb{E}[\|\Theta_{-k+[0,h]}\|_p^\alpha / \|\Theta_{-k+[0,h]}\|_\alpha^\alpha]$ . In particular,  $c(\alpha, h) = h + 1$  and

$$\mathbb{P}(\mathbf{Q}^{(\alpha)}(h) \in \cdot) = \mathbb{P}(\Theta_{-U^{(h)}+[0,h]} / \|\Theta_{-U^{(h)}+[0,h]}\|_\alpha \in \cdot), \quad (8.10)$$

where  $U^{(h)}$  is uniformly distributed on  $\{0, \dots, h\}$  and independent of  $\Theta$ .

**Lemma 8.4.** *Assume  $|\Theta_t| \rightarrow 0$  as  $t \rightarrow \infty$  and let  $f : \tilde{\ell}^\alpha \cap \{\mathbf{x} : \|\mathbf{x}\|_p = 1\} \rightarrow (0, \infty)$  be any bounded Lipschitz-continuous function in  $(\tilde{\ell}^\alpha, \tilde{d}_\alpha)$ . Then, for every  $p \geq \alpha$ ,*

$$\frac{c(p, h)}{h+1} \mathbb{E}[f(\mathbf{Q}^{(p)}(h))] \rightarrow \mathbb{E}[\|\Theta / \|\Theta\|_\alpha\|_p^\alpha f(\Theta / \|\Theta\|_p)], \quad (8.11)$$

as  $h \rightarrow +\infty$

We conclude from (8.11) for  $f(\mathbf{x}) \equiv 1$  that  $\lim_{h \rightarrow \infty} c(p, h)/(h+1) = c(p)$ . If  $0 < c(p) < \infty$ ,

$$\lim_{h \rightarrow \infty} \mathbb{E}[f(\mathbf{Q}^{(p)}(h))] = c(p)^{-1} \mathbb{E}[\|\Theta / \|\Theta\|_\alpha\|_p^\alpha f(\Theta / \|\Theta\|_p)]. \quad (8.12)$$

Finally, the portmanteau theorem yields  $\mathbf{Q}^{(p)}(h) \xrightarrow{d} \mathbf{Q}^{(p)}(\infty)$  in  $(\tilde{\ell}^p \cap \{\mathbf{x} : \|\mathbf{x}\|_p = 1\}, \tilde{d}_p)$  where  $\mathbf{Q}^{(p)}(\infty)$  is well defined in view of the right-hand side of (8.12). This finishes the proof of the proposition.  $\square$



*Proof of Lemma 8.3.* If  $\|\mathbf{X}_{[0,h]}/x\|_p > 1$  then for sufficiently small  $\epsilon > 0$ ,  $\|\mathbf{X}_{[0,h]}/x\|_\infty > \epsilon$ . Therefore, on  $\{\|\mathbf{X}_{[0,h]}/x\|_p > 1\}$ ,

$$\sum_{i=0}^h |\mathbf{X}_i/x|^\alpha \mathbb{1}(|\mathbf{X}_i/x| > \epsilon) > 0.$$

Using stationarity, we obtain

$$\begin{aligned} & \mathbb{P}(\|\mathbf{X}_{[0,h]}/x\|_p > 1) \\ &= \sum_{i=0}^h \mathbb{E} \left[ \frac{|\mathbf{X}_i/x|^\alpha \mathbb{1}(|\mathbf{X}_i/x| > \epsilon)}{\sum_{t=0}^h |\mathbf{X}_t/x|^\alpha \mathbb{1}(|\mathbf{X}_t/x| > \epsilon)} \mathbb{1}(\|\mathbf{X}_{[0,h]}/x\|_p > 1) \right] \\ &= \sum_{i=0}^h \mathbb{E} \left[ \frac{|\mathbf{X}_0/x|^\alpha \mathbb{1}(|\mathbf{X}_0/x| > \epsilon)}{\sum_{t=-i}^{h-i} |\mathbf{X}_t/x|^\alpha \mathbb{1}(|\mathbf{X}_t/x| > \epsilon)} \mathbb{1}(\|\mathbf{X}_{[-i,h-i]}/x\|_p > 1) \right] \\ &= \mathbb{P}(|\mathbf{X}_0| > x\epsilon) \sum_{i=0}^h \mathbb{E} \left[ \frac{|\mathbf{X}_0/x|^\alpha \mathbb{1}(\|\mathbf{X}_{[-i,h-i]}/x\|_p > 1)}{\sum_{t=-i}^{h-i} |\mathbf{X}_t/x|^\alpha \mathbb{1}(|\mathbf{X}_t/x| > \epsilon)} \mid |\mathbf{X}_0| > x\epsilon \right]. \end{aligned}$$

Applying the definition (2.1) of regular variation and dominated convergence, we obtain as  $x \rightarrow \infty$ ,

$$\begin{aligned} \frac{\mathbb{P}(\|\mathbf{X}_{[0,h]}/x\|_p > 1)}{\mathbb{P}(|\mathbf{X}_0| > x)} &\rightarrow \epsilon^{-\alpha} \sum_{i=0}^h \mathbb{E} \left[ \frac{|\epsilon Y \boldsymbol{\Theta}_0|^\alpha \mathbb{1}(\|\epsilon Y \boldsymbol{\Theta}_{[-i,h-i]}\|_p > 1)}{\sum_{t=-i}^{h-i} |\epsilon Y \boldsymbol{\Theta}_t|^\alpha \mathbb{1}(|Y \boldsymbol{\Theta}_t| > 1)} \right] \\ &= \sum_{i=0}^h \int_\epsilon^\infty \mathbb{E} \left[ \frac{\mathbb{1}(y \|\boldsymbol{\Theta}_{[-i,h-i]}\|_p > 1)}{\sum_{t=-i}^{h-i} |\boldsymbol{\Theta}_t|^\alpha \mathbb{1}(y |\boldsymbol{\Theta}_t| > \epsilon)} \right] d(-y^{-\alpha}). \end{aligned}$$

The left-hand side does not depend on  $\epsilon$ . Therefore, letting  $\epsilon \downarrow 0$ , we arrive at

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{\mathbb{P}(\|\mathbf{X}_{[0,h]}/x\|_p > 1)}{\mathbb{P}(|\mathbf{X}_0| > x)} &= \sum_{i=0}^h \int_0^\infty \mathbb{E} \left[ \frac{\mathbb{1}(y \|\boldsymbol{\Theta}_{[-i,h-i]}\|_p > 1)}{\|\boldsymbol{\Theta}_{[-i,h-i]}\|_\alpha^\alpha} \right] d(-y^{-\alpha}) \\ (8.13) \quad &= \sum_{i=0}^h \mathbb{E} \left[ \frac{\|\boldsymbol{\Theta}_{-i+[0,h]}\|_p^\alpha}{\|\boldsymbol{\Theta}_{-i+[0,h]}\|_\alpha^\alpha} \right] = c(p, h). \end{aligned}$$

This constant is finite since  $\|\boldsymbol{\Theta}_{i+[0,h]}\|_p \leq (h+1)\|\boldsymbol{\Theta}_{i+[0,h]}\|_\infty$ .

Next we prove (8.9). For this reason, let  $A$  be a continuity set with respect to the limit law in (8.9). An appeal to (8.13) yields

$$\begin{aligned} & c(p, h) \mathbb{P}(x^{-1} \mathbf{X}_{[0,h]} \in A \mid \|\mathbf{X}_{[0,h]}\|_p > x) \\ & \sim \frac{\mathbb{P}(x^{-1} \mathbf{X}_{[0,h]} \in A, \|\mathbf{X}_{[0,h]}\|_p > x)}{\mathbb{P}(\|\mathbf{X}_{[0,h]}\|_p > x)} =: I(x). \end{aligned}$$

Proceeding as for the derivation of (8.13), we obtain

$$\begin{aligned} I(x) &\sim \int_0^\infty \sum_{i=0}^h \mathbb{E} \left[ \frac{\mathbb{1}(y \|\Theta_{[-i, h-i]}\|_p > 1)}{\|\Theta_{[-i, h-i]}\|_\alpha^\alpha} \mathbb{1}(y \Theta_{[-i, h-i]} \in A) \right] d(-y^{-\alpha}) \\ &= \int_1^\infty \sum_{i=0}^h \mathbb{E} \left[ \frac{\|\Theta_{[-i, h-i]}\|_p^\alpha}{\|\Theta_{[-i, h-i]}\|_\alpha^\alpha} \mathbb{1}\left(y \frac{\Theta_{[-i, h-i]}}{\|\Theta_{[-i, h-i]}\|_p} \in A\right) \right] d(-y^{-\alpha}). \end{aligned}$$

In the last step we changed the variable,  $u = y \|\Theta_{[-i, h-i]}\|_p > 0$  a.s., observing that  $\|\Theta_{[-i, h-i]}\|_p \geq \|\Theta_0\|_p = 1$ . This proves (8.9) and the lemma.  $\square$

*Proof of Lemma 8.4.* Assume  $f : \tilde{\ell}^\alpha \cap \{\mathbf{x} : \|\mathbf{x}\|_p = 1\} \rightarrow (0, \infty)$  is any bounded Lipschitz-continuous function in  $(\tilde{\ell}^\alpha, \tilde{d}_\alpha)$ . By Lemma 8.3 we have for all  $p \geq \alpha$ ,

$$\begin{aligned} &\frac{c(p, h)}{h+1} \mathbb{E}[f(\mathbf{Q}^{(p)}(h))] - c(p) \mathbb{E}[f(\mathbf{Q}^{(p)})] = \\ &= \frac{1}{h+1} \sum_{k=0}^h \mathbb{E} \left[ \left( \frac{\|\Theta_{-k+[0, h]}\|_p^\alpha}{\|\Theta_{-k+[0, h]}\|_\alpha^\alpha} - \frac{\|\Theta\|_p^\alpha}{\|\Theta\|_\alpha^\alpha} \right) f(\Theta_{-k+[0, h]}/\|\Theta_{-k+[0, h]}\|_p) \right] + \\ &\quad + \mathbb{E} \left[ \frac{\|\Theta\|_p^\alpha}{\|\Theta\|_\alpha^\alpha} \left( f(\Theta_{-k+[0, h]}/\|\Theta_{-k+[0, h]}\|_p) - f(\Theta/\|\Theta\|_p) \right) \right] \\ &=: I + II. \end{aligned}$$

We will prove that  $I$  and  $II$  vanish as  $h \rightarrow \infty$ . Since  $p \geq \alpha$  subadditivity yields for  $k \in [0, h]$ ,

$$\begin{aligned} &\left| \frac{\|\Theta_{-k+[0, h]}\|_p^\alpha}{\|\Theta_{-k+[0, h]}\|_\alpha^\alpha} - \frac{\|\Theta\|_p^\alpha}{\|\Theta\|_\alpha^\alpha} \right| = \frac{\left| \|\Theta_{-k+[0, h]}\|_\alpha^\alpha \|\Theta_{-k+[0, h]}\|_p^\alpha - \|\Theta_{-k+[0, h]}\|_\alpha^\alpha \|\Theta\|_p^\alpha \right|}{\|\Theta_{-k+[0, h]}\|_\alpha^\alpha \|\Theta\|_\alpha^\alpha} \\ &\leq \frac{\left| \|\Theta_{-k+[0, h]}\|_\alpha^p \|\Theta_{-k+[0, h]}\|_p^p - \|\Theta_{-k+[0, h]}\|_\alpha^p \|\Theta\|_p^p \right|^{\alpha/p}}{\|\Theta_{-k+[0, h]}\|_\alpha^\alpha \|\Theta\|_\alpha^\alpha}. \end{aligned}$$

Moreover,

$$\begin{aligned} &\left| \|\Theta_{-k+[0, h]}\|_\alpha^p \|\Theta_{-k+[0, h]}\|_p^p - \|\Theta_{-k+[0, h]}\|_\alpha^p \|\Theta\|_p^p \right| \\ &\leq \|\Theta_{-k+[0, h]}\|_\alpha^p \left( \sum_{t=-\infty}^{-k-1} |\Theta_t|^p + \sum_{t=-k+h+1}^{+\infty} |\Theta_t|^p \right) \\ &\quad + \left| \|\Theta\|_\alpha - \|\Theta_{-k+[0, h]}\|_\alpha \right|^p \sum_{t=-k}^{-k+h} |\Theta_t|^p. \end{aligned}$$

Thus,  $|I|$  is bounded from above by

$$\begin{aligned} & \frac{1}{h+1} \|f\|_\infty \sum_{k=0}^h \left( \mathbb{E} \left[ \frac{\|\Theta_{[-\infty, -k-1]}^\alpha\|_p^\alpha}{\|\Theta\|_\alpha^\alpha} \right] + \mathbb{E} \left[ \frac{\|\Theta_{[-k+h+1, \infty]}^\alpha\|_p^\alpha}{\|\Theta\|_\alpha^\alpha} \right] \right. \\ & \quad \left. + \mathbb{E} \left[ \frac{\|\Theta_{[-\infty, -k-1]}^\alpha\|_p^\alpha + \|\Theta_{[-k+h+1, \infty]}^\alpha\|_p^\alpha}{\|\Theta\|_\alpha^\alpha} \frac{\|\Theta_{-k+[0, h]}^\alpha\|_p^\alpha}{\|\Theta_{-k+[0, h]}^\alpha\|_\alpha^\alpha} \right] \right) \\ & \leq \frac{1}{h+1} \|f\|_\infty \sum_{k=0}^h \left( \mathbb{E} \left[ \frac{\|\Theta_{[-\infty, -(k+1)]}^\alpha\|_p^\alpha + \|\Theta_{[-\infty, -(k+1)]}^\alpha\|_\alpha^\alpha}{\|\Theta\|_\alpha^\alpha} \right] \right. \\ & \quad \left. + \mathbb{E} \left[ \frac{\|\Theta_{[k+1, \infty]}^\alpha\|_p^\alpha + \|\Theta_{[k+1, \infty]}^\alpha\|_\alpha^\alpha}{\|\Theta\|_\alpha^\alpha} \right] \right). \end{aligned}$$

Taking the limit as  $h \rightarrow \infty$ , the Cèsaro limit on the right-hand side converges to zero.

We use the Lipschitz-continuity of  $f$  to obtain an upper bound of  $|II|$ :

$$|II| \leq \frac{1}{h+1} c \sum_{k=0}^h \mathbb{E} \left[ \frac{\|\Theta_{-k+[0, h]}^\alpha\|_p^\alpha}{\|\Theta_{-k+[0, h]}^\alpha\|_\alpha^\alpha} \tilde{d}_\alpha \left( \frac{\Theta_{-k+[0, h]}}{\|\Theta_{-k+[0, h]}\|_p}, \frac{\Theta}{\|\Theta\|_p} \right) \right].$$

Similar arguments as for  $|I| \rightarrow 0$  show that  $|II| \rightarrow 0$ .  $\square$

#### 8.4. Proofs of the results of Section 4.

8.4.1. *Proof of Theorem 4.2.* We start with a version of Theorem 4.2 for deterministic thresholds  $(x_b)$ .

**Lemma 8.5.** *Assume the conditions of Theorem 4.2. Then for every  $g \in \mathcal{G}_+(\tilde{\ell}^p)$ ,*

$$(8.14) \quad \frac{1}{k} \sum_{t=1}^m g(x_b^{-1} \mathbf{B}_t) \xrightarrow{\mathbb{P}} \int_0^\infty \mathbb{E}[g(y \mathbf{Q}^{(p)})] d(-y^{-\alpha}), \quad n \rightarrow \infty,$$

holds for sequences  $k_n \rightarrow \infty$  and  $m_n := [n/b_n] \rightarrow \infty$  as in  $\mathbf{MX}_p$ .

*Proof.* If  $\mathbf{MX}_p$  holds for Lipschitz-continuous  $f \in \mathcal{G}_+(\tilde{\ell}^p)$ , then it holds for functions  $g \in \mathcal{G}_+(\tilde{\ell}^p)$  of the form  $g(\mathbf{x}_t) = \mathbb{1}(\mathbf{x}_t \in A)$  where  $A$  is a continuity-set of  $\tilde{\ell}^p$  and  $\mathbf{0} \notin \bar{A}$ . It suffices to prove that

$$(8.15) \quad \left( \mathbb{E} \left[ e^{-\frac{1}{k} \sum_{t=1}^{\lfloor m/k \rfloor} g(x_b^{-1} \mathbf{B}_t)} \right] \right)^k \rightarrow e^{-\mathbb{E} \left[ \int_0^\infty g(y \mathbf{Q}^{(p)}) d(-y^{-\alpha}) \right]}.$$

By stationarity,

$$(8.16) \quad \mathbb{E} \left[ 1 - e^{-\frac{1}{k} \sum_{t=1}^{\lfloor m/k \rfloor} g(x_b^{-1} \mathbf{B}_t)} \right] = O(k^{-2} m \mathbb{E}[g(x_b^{-1} \mathbf{B}_1)]).$$

Since  $g$  vanishes in some neighborhood of the origin there exists  $c_g > 0$  such that  $g(\mathbf{x}) = g(\mathbf{x}) \mathbb{1}(\|\mathbf{x}\|_p > c_g)$ . Therefore and by virtue of Proposition 8.2 the right-hand side of (8.16) vanishes as  $n \rightarrow \infty$ . Now a Taylor expansion argument shows that the left-hand side of (8.15) is of the asymptotic order  $\sim \exp\{-(m/k) \mathbb{E}[g(x_b^{-1} \mathbf{B}_t)]\}$ , and another application of Proposition 8.2 yields

(8.15). We conclude by the portmanteau theorem for  $M_0(\tilde{\ell}^p)$ -convergence in Hult and Lindskog [21], Theorem 2.4. that (8.14) holds.  $\square$

We continue with the proof of Theorem 4.2. Lemma 8.5 implies convergence of the empirical measures in  $M_0(\tilde{\ell}^p)$ :

$$P_n(\cdot) := \frac{1}{k} \sum_{t=1}^m \mathbb{1}(x_b^{-1} \mathbf{B}_t \in \cdot) \xrightarrow{\mathbb{P}} P(\cdot) := \int_0^\infty \mathbb{P}(y \mathbf{Q}^{(p)} \in \cdot) d(-y^{-\alpha}).$$

Using the argument in Resnick [33], p. 81, we may conclude  $\|\mathbf{B}\|_{p,(k+1)}/x_b \xrightarrow{\mathbb{P}} 1$ , and thus the joint convergence in  $(P_n, \|\mathbf{B}\|_{p,(k+1)}/x_b) \xrightarrow{\mathbb{P}} (P, 1)$  in  $M_0(\tilde{\ell}^p) \times \mathbb{R}_+$  follows. Now (4.2) follows by an application of the continuous mapping theorem to the scaling function  $s(P(\cdot), t) = P(t \cdot)$ . To prove continuity of  $s$  we use again the portmanteau theorem for  $M_0(\tilde{\ell}^p)$ -convergence in Hult and Lindskog [21], Theorem 2.4. Thus it suffices to check whether the limit  $P_n f(\cdot/t) \xrightarrow{\mathbb{P}} P f$  holds as  $(n, t) \rightarrow (\infty, 1)$  for Lipschitz-continuous  $f \in \mathcal{G}_+(\tilde{\ell}^p)$ . But we have with Lemma 8.5

$$\begin{aligned} |P_n f(\cdot/t) - P f| &\leq |P_n f(\cdot/t) - P_n f| + |P_n f - P f| \\ &= |P_n f(\cdot/t) - P_n f| + o_{\mathbb{P}}(1), \quad n \rightarrow \infty. \end{aligned}$$

Then, for all  $0 < t_0 \leq t < 2$ , for  $t_0 \leq 1$ , setting  $g(\mathbf{x}) = (\|\mathbf{x}\|_p \wedge \|f\|_\infty) \mathbb{1}(\{\mathbf{x} : \|\mathbf{x}\|_p > c_f/t_0\})$ , we have

$$\begin{aligned} |P_n f(\cdot/t) - P f| &\leq |t^{-1} - 1| P_n g + o_{\mathbb{P}}(1) \\ &\leq |t^{-1} - 1| (c + o_{\mathbb{P}}(1)) + o_{\mathbb{P}}(1), \end{aligned}$$

for some  $c > 0$ ,  $c_f > 0$  as above. Letting  $t \rightarrow 1$ , continuity of  $s$  follows.  $\square$

8.4.2. *Proof of Proposition 4.8.* The result follows by a direct application of Theorem 3.1 in Bartkiewicz *et al.* [1] on  $\mathbf{u}^\top \mathbf{S}_n$  for every  $\mathbf{u} \in \mathbb{R}^d$  such that  $|\mathbf{u}| = 1$  by checking their conditions **(AC)**, **(TB)**. Condition (4.8) implies that for all  $\delta > 0$ ,

$$\lim_{l \rightarrow \infty} \limsup_{n \rightarrow \infty} n \sum_{t=l}^{b_n} \mathbb{P}(|\mathbf{X}_t| > \delta a_n, |\mathbf{X}_0| > \delta a_n),$$

from which **(AC)** is immediate. This condition also implies **(TB)**. We show this in two steps. First, we identify the coefficients  $b(v)$  in **(TB)** in terms of the spectral tail process. Mikosch and Wintenberger [27] showed that

$$b_{\pm}(v) - b_{\pm}(v-1) = \mathbb{E} \left[ \left( \sum_{j=0}^v \mathbf{u}^\top \boldsymbol{\Theta}_j \right)_{\pm}^{\alpha} \right] - \mathbb{E} \left[ \left( \sum_{j=1}^v \mathbf{u}^\top \boldsymbol{\Theta}_j \right)_{\pm}^{\alpha} \right],$$

where we suppress in the notation the dependence of the left-hand side on  $\mathbf{u}$  in what follows. **(TB)** amounts to verifying that  $b_{\pm}(v) - b_{\pm}(v-1)$  converges as  $v \rightarrow \infty$ . For  $\alpha \in (0, 1)$  this follows by concavity since  $\|\boldsymbol{\Theta}\|_{\alpha} < \infty$  a.s. For

$1 < \alpha < 2$  this will follow by a convexity argument if  $\mathbb{E}[(\sum_{j \geq 0} |\Theta_j|)^{\alpha-1}] < \infty$ . By subadditivity and Jensen's inequality, it is enough to check

$$(8.17) \quad \sum_{j=0}^{\infty} (\mathbb{E}[|\Theta_j|^{\alpha-1} \mathbb{1}(|\Theta_j| > 1)] + \mathbb{E}[|\Theta_j| \wedge 1]) < +\infty.$$

We start by showing

$$(8.18) \quad \sum_{j=0}^{\infty} \mathbb{E}[|\Theta_j| \wedge 1] < \infty.$$

Condition (4.8) implies

$$\lim_{l \rightarrow \infty} \limsup_{n \rightarrow \infty} n \sum_{j=l}^{b_n} \mathbb{E}[ (|a_n^{-1} \mathbf{X}_j| \wedge 1) \mathbb{1}(|\mathbf{X}_0| > a_n) ] = 0,$$

which yields the following Cauchy criterion: for every  $\varepsilon > 0$  there exists  $K$  sufficiently large such that for  $l \geq K, h \geq 0$ ,

$$\limsup_{n \rightarrow \infty} n \sum_{j=l}^{l+h} \mathbb{E}[ (|a_n^{-1} \mathbf{X}_j| \wedge 1) \mathbb{1}(|\mathbf{X}_0| > a_n) ] = \sum_{j=l}^{l+h} \mathbb{E}[|Y \Theta_j| \wedge 1] \leq \varepsilon,$$

where we used regular variation of  $(\mathbf{X}_t)$  in the last step. Then, we conclude (8.18) holds. By stationarity we can show similarly

$$(8.19) \quad \sum_{j=0}^{\infty} \mathbb{E}[|\Theta_{-j}| \wedge 1] < +\infty.$$

Then, by the time-change formula in (2.2) we deduce

$$\begin{aligned} \infty &> \sum_{j=0}^{\infty} \mathbb{E}[|\Theta_{-j}| \wedge 1] = \sum_{j=0}^{\infty} \mathbb{E}[|\Theta_j|^{\alpha} (|\Theta_j|^{-1} \wedge 1)] \\ &= \sum_{j=0}^{\infty} \mathbb{E}[|\Theta_j|^{\alpha-1} \wedge |\Theta_j|^{\alpha}] > \sum_{j=0}^{\infty} \mathbb{E}[|\Theta_j|^{\alpha-1} \mathbb{1}(|\Theta_j| > 1)], \end{aligned}$$

and (8.17) holds. This finishes the proof of the fact that  $\mathbb{E}[(\sum_{j \geq 0} |\Theta_t|)^{\alpha-1}] < +\infty$ , in particular  $c(1) < \infty$ . Applying the mean value theorem and dominated convergence we arrive at the relation

$$b_{\pm}(v) - b_{\pm}(v-1) \rightarrow \mathbb{E} \left[ \left( \sum_{j=0}^{\infty} \mathbf{u}^{\top} \Theta_j \right)_{\pm}^{\alpha} - \left( \sum_{j=1}^{\infty} \mathbf{u}^{\top} \Theta_j \right)_{\pm}^{\alpha} \right], \quad v \rightarrow \infty.$$

Reasoning for the limit as for (8.6) and recalling that  $c(1) < \infty$ , we identify

$$\begin{aligned}
& \mathbb{E} \left[ \left( \sum_{j=0}^{\infty} \mathbf{u}^\top \boldsymbol{\Theta}_j \right)_\pm^\alpha - \left( \sum_{j=1}^{\infty} \mathbf{u}^\top \boldsymbol{\Theta}_j \right)_\pm^\alpha \right] \\
&= \mathbb{E} \left[ \left( \sum_{j=-\infty}^{\infty} \mathbf{u}^\top \boldsymbol{\Theta}_j \right)_\pm^\alpha / \|\boldsymbol{\Theta}\|_\alpha^\alpha \right] = \mathbb{E} \left[ \left( \sum_{j=-\infty}^{\infty} \mathbf{u}^\top \mathbf{Q}_j^{(\alpha)} \right)_\pm^\alpha \right] \\
&= c(1) \mathbb{E} \left[ \left( \sum_{j=-\infty}^{\infty} \mathbf{u}^\top \mathbf{Q}_j^{(1)} \right)_\pm^\alpha \right].
\end{aligned}$$

### 8.5. Proofs of the results of Section 6.

8.5.1. *Proof of Proposition 6.1.* Notice that  $\psi_g$  is bounded and measurable. For  $p = \alpha$  we have  $\mathbf{Q}^{(\alpha)} \stackrel{d}{=} \boldsymbol{\Theta} / \|\boldsymbol{\Theta}\|_\alpha$ . Then the result follows from Proposition 3.6 in Janssen [22]. For  $p > 0$ , assuming the spectral cluster process  $\mathbf{Q}^{(p)}$  is well defined we have  $\|\boldsymbol{\Theta}\|_p < \infty$  a.s. and  $c(p) < \infty$ . Then, we introduce the Radon-Nikodym derivative of  $\mathcal{L}(\mathbf{Q}^{(p)})$  with respect to  $\mathcal{L}(\boldsymbol{\Theta} / \|\boldsymbol{\Theta}\|_p)$  which by (3.2) is the function  $h : \ell^p \cap \{\mathbf{x} : \|\mathbf{x}\|_p = 1\} \rightarrow \mathbb{R}_{\geq 0}$  defined by  $h(\mathbf{y} / \|\mathbf{y}\|_p) := \|\mathbf{y}\|_\alpha / \|\mathbf{y}\|_p$ . Finally, the result follows by another application of Proposition 3.6 in Janssen [22].  $\square$

8.5.2. *Proof of Theorem 6.2.* The proof is given for  $p = \alpha$  only; the case  $p \leq \alpha$  extends in a natural way. Let  $g : \ell^\alpha \rightarrow \mathbb{R}$  be a continuous bounded function. We start by proving that  $\psi_g$  defined in (6.1) is a continuous bounded function on  $\tilde{\ell}^\alpha$ . Fix  $\epsilon > 0$  and  $[\mathbf{z}] \in \tilde{\ell}^\alpha \cap \{[\mathbf{y}] : \|\mathbf{y}\|_\alpha = 1\}$ . Then for all  $[\mathbf{x}] \in \tilde{\ell}^\alpha \cap \{[\mathbf{y}] : \|\mathbf{y}\|_\alpha = 1\}$ ,  $k \in \mathbb{Z}$  and  $N \in \mathbb{N}$ , we have

$$\begin{aligned}
& |\psi_g(\mathbf{x}) - \psi_g(\mathbf{z})| = \left| \sum_{j \in \mathbb{Z}} |\mathbf{x}_j^*|^\alpha g((\mathbf{x}_{j+t}^*)_t) - \sum_{j \in \mathbb{Z}} |\mathbf{z}_j^*|^\alpha g((\mathbf{z}_{j+t}^*)_t) \right| \\
&= \left| \sum_{j \in \mathbb{Z}} |\mathbf{x}_j^* - \mathbf{z}_{j-k}^*|^\alpha g((\mathbf{x}_{j+t}^*)_t) - \sum_{j \in \mathbb{Z}} |\mathbf{z}_j^*|^\alpha (g((\mathbf{z}_{j+t}^*)_t) - g((\mathbf{x}_{j+t+k}^*)_t)) \right| \\
&\leq \|g\|_\infty d_\alpha^\alpha(B^{-k} \mathbf{z}^*, \mathbf{x}^*) + 2\|g\|_\infty d_\alpha^\alpha(\mathbf{z}^*, \mathbf{z}_{[-N, N]}^*) \\
&\quad + \sum_{|j| < N} |\mathbf{z}_j^*|^\alpha |g((\mathbf{z}_{j+t}^*)_t) - g((\mathbf{x}_{j+t+k}^*)_t)|.
\end{aligned}$$

If  $[\mathbf{x}]$  satisfies  $\tilde{d}_\alpha^\alpha(\mathbf{z}, \mathbf{x}) < \epsilon(3\|g\|_\infty)^{-1}$  then there exists  $k_0 \in \mathbb{Z}$  such that

$$\tilde{d}_\alpha^\alpha(\mathbf{z}, \mathbf{x}) < d_\alpha^\alpha(B^{-k_0} \mathbf{z}^*, \mathbf{x}^*) < \epsilon(3\|g\|_\infty)^{-1}.$$

Furthermore, choose  $N_0 \geq 0$  such that  $d_\alpha^\alpha(\mathbf{z}^*, \mathbf{z}_{[-N_0, N_0]}^*) < \epsilon(2 \times 3\|g\|_\infty)^{-1}$  and consider the finite set  $C_{[\mathbf{z}]} \subset \ell^\alpha \cap \{[\mathbf{y}] : \|\mathbf{y}\|_\alpha = 1\}$ , defined by  $C_{[\mathbf{z}]} := \{(\mathbf{z}_{j+t}^*)_t \in \ell^\alpha : |j| < N_0, |\mathbf{z}_j^*| > 0\}$ . Notice that for every  $\tilde{\mathbf{z}} \in C_{[\mathbf{z}]}$  there exists  $\delta(\tilde{\mathbf{z}})$  such that if  $d_\alpha^\alpha(\tilde{\mathbf{z}}, \mathbf{x}) < \delta(\tilde{\mathbf{z}})$  implies  $|g(\tilde{\mathbf{z}}) - g(\mathbf{x})| < \epsilon/3$ . Finally, define  $\eta(\mathbf{z}) := \min\{\delta(\tilde{\mathbf{z}}) : \tilde{\mathbf{z}} \in C\} \wedge \epsilon(3\|g\|_\infty)^{-1}$ . Then, noticing that

$\sum_{|j| < N_0} |\mathbf{z}_j^*|^\alpha \leq \|\mathbf{z}\|_\alpha^\alpha = 1$ , we also obtain a bound for the last term. Hence, for every  $[\mathbf{x}] \in \tilde{\ell}^\alpha$  satisfying  $d_\alpha^\alpha(\mathbf{z}, \mathbf{x}) < \eta(\mathbf{z})$  we have  $|\psi_g(\mathbf{x}) - \psi_g(\mathbf{z})| < \epsilon$ .

This finishes the proof of the continuity of the function  $\psi_g$  on  $\tilde{\ell}^\alpha \cap \{\mathbf{y} : \|\mathbf{y}\|_\alpha = 1\}$ . We conclude with applications of Lemma 8.5 and Proposition 6.1.  $\square$

8.5.3. *Proof of Proposition 6.5.* Theorem 4.5 in Mikosch and Wintenberger [29] yields immediately

$$\left| \frac{\mathbb{P}(\sup_{1 \leq t \leq n} S_t > x_n)}{n\mathbb{P}(|X_1| > x_n)} - \mathbb{E} \left[ \left( \sup_{t \geq 0} \sum_{i=0}^t \Theta_i \right)_+^\alpha - \left( \sup_{t \geq 1} \sum_{i=1}^t \Theta_i \right)_+^\alpha \right] \right| \rightarrow 0, \quad n \rightarrow \infty, \quad (8.20)$$

and  $n\mathbb{P}(|X_1| > x_n) \rightarrow 0$ . We multiply the function inside the limiting expected value by the constant  $1 = \|\Theta\|_\alpha^\alpha / \|\Theta\|_\alpha^\alpha$ . Moreover, since  $c(1) < \infty$ , then  $\mathbb{E}[(\sum_{t=1}^\infty |\Theta_t|)^{\alpha-1}] < \infty$ ; see Lemma 3.11 in Planinić and Soulier [32]. Then, by Fubini's theorem,

$$\begin{aligned} & \mathbb{E} \left[ \left( \sup_{t \geq 0} \sum_{i=0}^t \Theta_i \right)_+^\alpha - \left( \sup_{t \geq 1} \sum_{i=1}^t \Theta_i \right)_+^\alpha \right] \\ &= \sum_{j \in \mathbb{Z}} \mathbb{E} \left[ |\Theta_j|^\alpha \left( \left( \sup_{t \geq 0} \sum_{i=0}^t \frac{\Theta_i}{\|\Theta\|_\alpha} \right)_+^\alpha - \left( \sup_{t \geq 1} \sum_{i=1}^t \frac{\Theta_i}{\|\Theta\|_\alpha} \right)_+^\alpha \right) \right]. \end{aligned}$$

At this point we apply the time-change formula for positive measurable functions of  $\Theta$  at every term of the sum in  $j \in \mathbb{Z}$ ; see Corollary 2.8. in Dombry *et al.* [13]. By the same argument as in the proof of Proposition 8.2 we obtain the representation of the expectation in (8.20) in terms of the univariate spectral cluster process  $Q^{(p)}$ .

Now we apply Theorem 6.2 to  $f(\mathbf{x}) := \lim_{k \rightarrow \infty} (\sup_{t \geq -k} \sum_{i=-k}^t x_i)_+^\alpha$  on  $\ell^1$ . It is uniformly continuous and bounded by one on the sphere of  $\ell^p$ , hence (6.2) holds for  $f$ . Similarly, the constant  $c(1)^{-1}$  can be estimated by employing the function  $g(\mathbf{x}) := \|\mathbf{x}\|_\alpha$  on  $\ell^1$  which is bounded by one on the unitary  $\ell^1$ -sphere for  $\alpha \geq 1$ .  $\square$

8.5.4. *Proof of Proposition 6.7.* The re-normalization function  $\zeta$  is continuous on the unit sphere of  $(\ell^\alpha, d_\alpha)$ , except for sequences with  $\mathbf{x}_0 = 0$ . Then

$$\begin{aligned} \mathbb{P}(\rho(Y \Theta) > 1) &= \mathbb{E}[\rho(\Theta_t)^\alpha \wedge 1] = \mathbb{E}[\rho(\mathbf{Q}_t^{(\alpha)} / |\mathbf{Q}_0^{(\alpha)}|)^\alpha \wedge 1] \\ &= \mathbb{E}[(\rho^\alpha \wedge 1) \circ \zeta(\mathbf{Q}^{(\alpha)})]. \end{aligned}$$

The proof is finished by an application of Theorem 6.2.  $\square$

## REFERENCES

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



- [2] BASRAK, B., KRIZMANIĆ, D. AND SEGERS, J. (2012) A functional limit theorem for dependent sequences with infinite variance stable limits. *Ann. Probab.* **40**, 2008–2033.
- [3] BASRAK, B., PLANINIĆ, H. AND SOULIER, P. (2018) An invariance principle for sums and record times of regularly varying stationary sequences. *Probab. Th. Rel. Fields* **172**, 869–914.
- [4] BASRAK, B. AND SEGERS, J. (2009) Regularly varying multivariate time series. *Stoch. Proc. Appl.* **119**, 1055–1080.
- [5] BINGHAM, N., GOLDIE, C.M. AND TEUGELS, J. (1987) *Regular Variation*. Cambridge University Press, Cambridge (UK).
- [6] BRADLEY, R.C. (2005) Basic properties of strong mixing conditions. A survey and some open questions. *Probability Surveys* **2**, 107–144.
- [7] BURITICÁ, G. MIKOSCH, T. MEYER, N. AND WINTENBERGER, O. (2021) Some variations on the extremal index. *Zap. Nauchn. Semin. POMI. Volume 501, Probability and Statistics*. **30**, 5277. To be translated in J.Math.Sci. (Springer).
- [8] CISSOKHO, Y. AND KULIK, R. (2021) Estimation of cluster functionals for regularly varying time series: sliding blocks estimators. *Electronic Journal of Statistics* **15**, 2777–2831.
- [9] CLINE, D.B.H. AND HSING, T. (1998) *Large deviation probabilities for sums of random variables with heavy or subexponential tails*, Technical Report, Texas A& M University.
- [10] DAVIS, R.A., DREES, H., SEGERS, J. AND WARCHOL, M. (2018) Inference on the tail process with applications to financial time series modeling. *J. Econometrics* **205**, 508–525.
- [11] DAVIS, R.A. AND HSING, T. (1995) Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Probab.* **23**, 879–917.
- [12] DAVIS, R.A., AND RESNICK, S. (1985) Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probab.* **3**, 179–195.
- [13] DOMBRY, C., HASHORVA, E. AND SOULIER, P. (2018) Tail measure and spectral tail process of regularly varying time series. *Ann. Appl. Probab.* **28**, 3884–3921.
- [14] DREES, H., JANSSEN, A. AND NEBLUNG, S. (2021) Cluster based inference for extremes of time series. *arXiv preprint arXiv:2103.08512*.
- [15] DREES, H. AND NEBLUNG, S. (2021) Asymptotics for sliding blocks estimators of rare events. *Bernoulli* **27**, 1239–1269.
- [16] DREES, H. AND ROOTZÉN, H. (2010) Limit theorems for empirical processes of cluster functionals. *Ann. Stat.* **38**, 2145–2186.
- [17] DREES, H., SEGERS, J., AND WARCHOL, M. (2015) Statistics for tail processes of Markov chains. *Extremes* **18**, 369–402.
- [18] EMBRECHTS, P., KLÄEPPPELBERG, C., AND MIKOSCH, T. (2013). *MODELLING EXTREMAL EVENTS: FOR INSURANCE AND FINANCE*. **33**, Springer Science & Business Media.
- [19] HAAN, L. DE, MERCADIER, C. AND ZHOU, C. (2016) Adapting extreme value statistics to financial time series: dealing with bias and serial dependence. *Finance and Stochastics* **20**, 321–354.
- [20] HSING, T. (1993) Extremal index estimation for a weakly dependent stationary sequence. *Ann. Stat.* **21**, 2043–2071.
- [21] HULT, H. AND LINDSKOG, F. (2006) Regular variation for measures on metric spaces. *Publ. de l’Institut Math. (Belgrade)* **80** (94), 121–140.
- [22] JANSSEN, A. (2019) Spectral tail processes and max-stable approximations of multivariate regularly varying time series. *Stoch. Proc. Appl.* **129**, 1993–2009.
- [23] KULIK, R. AND SOULIER, P. (2020) *Heavy-Tailed Time Series*. Springer, New York.
- [24] LEADBETTER, M.R. (1983) Extremes and local dependence in stationary sequences. *Probab. Th. Relat. Fields* **65**, 291–306.

- [25] LEADBETTER, M.R., LINDGREN, G., AND ROOTZÉN, H. (1983) *Extremes and related properties of random sequences and processes*. Springer, Berlin.
- [26] LINDSKOG, F., RESNICK, S.I., AND ROY, J. (2014) Regularly varying measures on metric spaces: hidden regular variation and hidden jumps. *Probab. Surveys* **11**, 270–314.
- [27] MIKOSCH, T. AND WINTENBERGER, O. (2013) Precise large deviations for dependent regularly varying sequences. *Probab. Th. Rel. Fields* **156**, 851–887.
- [28] MIKOSCH, T. AND WINTENBERGER, O. (2014) The cluster index of regularly varying sequences with applications to limit theory for functions of multivariate Markov chains. *Probab. Th. Rel. Fields* **159**, 157–196.
- [29] MIKOSCH, T. AND WINTENBERGER, O. (2016) A large deviations approach to limit theory for heavy-tailed time series. *Probab. Th. Rel. Fields* **166**, 233–269.
- [30] MIKOSCH, T. AND RODIONOV, I. (2021) Precise large deviations for dependent subexponential variables. *Bernoulli*. **27**, 1319–1347.
- [31] NAGAEV, S.V. (1979) Large deviations of sums of independent random variables. *Ann. Probab.* **7**, 745–789.
- [32] PLANINIĆ, H. AND SOULIER, P. (2018) The tail process revisited. *Extremes* **21**, 551–579.
- [33] RESNICK, S.I. (2007) *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer, New York.
- [34] SEGERS, J., ZHAO, Y. AND MEINGUET, T. (2017) Polar decomposition of regularly varying time series in star-shaped metric spaces. *Extremes* **20**, 539–566.

LPSM, SORBONNE UNIVERSITÉ, UPMC UNIVERSITÉ PARIS 06, F-75005, PARIS, FRANCE

*Email address:* gloria.buritica@upmc.fr

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF COPENHAGEN, UNIVERSITETSPARKEN 5, DK-2100 COPENHAGEN, DENMARK

*Email address:* mikosch@math.ku.dk

LPSM, SORBONNE UNIVERSITÉS, UPMC UNIVERSITÉ PARIS 06, F-75005, PARIS, FRANCE

*Email address:* olivier.wintenberger@upmc.fr