LONG-RANGE TAIL DEPENDENCE: EDM VS. EXTREMOGRAM

MARTIN LARSSON AND SIDNEY I. RESNICK

Abstract. The dependence of large values in a stochastic process is an important topic in risk, insurance and finance. The idea of risk contagion is based on the idea of large value dependence. The Gaussian copula notoriously fails to capture this phenomenon. Two notions in a process context which attempt to summarize this extremal dependence in a function comparable to a sample correlation function are the *extremal dependence measure* (EDM) and the *extremogram*. We review these ideas and compare the two tools and end with a central limit theorem for a natural estimator of the EDM which allows drawing confidence bands comparable to those provided by Bartlett's formula in a classical context of sample correlation functions.

1. Introduction

The problem of describing serial dependence in the tails of a stochastic process has been studied for some time, one of the earlier significant contributions being the work by Leadbetter [11], who studied the *extremal index* with the goal of describing the clustering of extremes. More recently, various attempts have been made to find extremal analogues of the autocorrelation and autocovariance functions from classical time series analysis. In particular we will review the *extremogram* suggested by Davis and Mikosch [5]. Related quantities are the *extreme dependence functions* and *extremal coefficient function* introduced by Fasen et al. [7], see also [18]. Both of these approaches may be viewed as generalizations in a stochastic processes context of the *coefficient of tail dependence* introduced by Ledford and Tawn [12, 13].

A slightly different, although related, approach is taken in [14], where the *extremal dependence measure*, or EDM, is introduced. The EDM was proposed as a statistical tool to investigate asymptotic independence, for which it was used in [8] in a study involving Internet data. In [3] it was used in a stochastic processes context to study long-range extremal dependence in certain data network models.

In this paper we are interested in understanding how the strength of dependence between two observations in a regularly varying strictly stationary time series decays as the time lag between the observations grows large. We will review the notions of serial extremal dependence mentioned above, namely the extremogram and the EDM, and by means of some simple examples illustrate how they may give rise to very different conclusions about long-range extremal dependence. The aim is to highlight the difficulty of devising one single universally valid notion of long-range tail dependence. This problem is prevalent in the literature on long-range dependence, where several different definitions have been suggested, all having their advantages and drawbacks. See [17] for a survey.

This paper is structured as follows. In Section 2 we briefly review some basic notions associated with multivariate regular variation. Influenced by the view that the extremal dependence structure

---

2000 *Mathematics Subject Classification.* 60G70 (Primary), 60F05, 60G10 (Secondary).

Key words and phrases. Regular variation, heavy tails, asymptotic independence, extremal dependence, EDM, extremogram.

S. Resnick was partially supported by ARO Contract W911NF-07-1-0078 at Cornell University.

is determined by the spectral measure, we introduce the concept of *weak asymptotic equivalence* of dependence measures; for now, think of a dependence measure as a real-valued function acting on a regularly varying vector. In Section 3 we review the EDM and the extremogram and derive some useful properties. Section 4 deals with *long-range tail dependence*, which we define based on the summability (or lack thereof) of the EDM and extremogram. Our aim is to use simple, tractable examples to illustrate some of the issues that any attempt at defining long-range tail dependence will have to deal with. Section 5 summarizes the preceding findings. Finally, Section 6 contains a central limit theorem for an estimator of the EDM in the case of iid, regularly varying bivariate pairs. This section deals with a statistical issue of practical importance.

2. Regular variation and measures of tail dependence

We say that a d-dimensional random vector $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ is regularly varying if there exist a function $b(t)$ tending to infinity and a non-null Radon measure ν on the punctured space $\mathbb{E} = [\boldsymbol{0}, \boldsymbol{\infty}]\backslash\{\boldsymbol{0}\}$ such that

$$tP\left(\frac{\boldsymbol{Z}}{b(t)} \in \cdot\right) \xrightarrow{\text{v}} \nu \qquad (t \to \infty). \tag{2.1}$$

Here $\xrightarrow{\text{v}}$ denotes vague convergence in $M_+(\mathbb{E})$, the space of Radon measures on $\mathbb{E}$. It is well-known (see, for instance, [15, Theorem 6.1]) that the limit measure ν has the scaling property $\nu(t\cdot) = t^{-\alpha}\nu(\cdot)$ for some $\alpha > 0$ called the index of regular variation of $\boldsymbol{Z}$.

Now fix a norm $\|\cdot\|$ on $\mathbb{E}$ and let $T : \boldsymbol{Z} \mapsto (\|\boldsymbol{Z}\|, \boldsymbol{Z}/\|\boldsymbol{Z}\|) = (R, \boldsymbol{\Theta})$ be the polar coordinate transformation. An equivalent formulation of the regular variation property (2.1) is [15, Theorem 6.1] the existence of a probability measure S on $\aleph_+ = \{\boldsymbol{x} \in \mathbb{E} : \|\boldsymbol{x}\| = 1\}$ and a constant $c > 0$ such that

$$tP\left(\left(\frac{R}{b(t)}, \boldsymbol{\Theta}\right) \in \cdot\right) \xrightarrow{\text{v}} c\nu_\alpha \times S \qquad (t \to \infty), \tag{2.2}$$

where ν_α is the measure on $(0, \infty]$ with density $\nu_\alpha(dx) = \alpha x^{-\alpha-1}dx$. The measure S is called the spectral or angular measure. Different normalizations $b(t)$ are possible, yielding different limit measures ν; however, all possible normalizations are asymptotically equivalent, and the limit measures only differ up to multiplicative constants. The constant c in (2.2) thus depends on which particular normalization one chooses. Using the well-known fact that $c\nu_\alpha \times S = \nu \circ T^{-1}$, it follows that $\nu \circ T^{-1}((1, \infty] \times \cdot) = c\nu_\alpha(1, \infty]S(\cdot) = cS(\cdot)$. Thus c is determined by the requirement that S be a probability measure: $c = cS(\aleph_+) = \nu\{\boldsymbol{x} : \|\boldsymbol{x}\| > 1\}$.

Note that the normalization $b(t)$ is the same for all components of $\boldsymbol{Z}$. This can be relaxed (see [15, Theorem 6.5]), but is appropriate when dealing with strictly stationary time series, where all one-dimensional marginal distributions are the same.

The extremal behavior of a regularly varying vector $\boldsymbol{Z}$ is determined by the limit measure ν. It is therefore reasonable that any quantity trying to capture tail properties should be a function of ν. In fact, the various measures proposed in the literature typically are of the form $\nu(f) = \int_{\mathbb{E}} f(x)\nu(dx)$ for some bounded, compactly supported $f : \mathbb{E} \mapsto \mathbb{R}$ which is continuous outside a ν-nullset (note that the relatively compact sets in $\mathbb{E}$ are those bounded away from the origin.) Usually $f = \boldsymbol{1}_C$ for some relatively compact measurable $C \subset \mathbb{E}$ with $\nu(\partial C) = 0$. One example is the extremogram which we will discuss momentarily.

By a change of variables, we may express $\nu(f)$ in terms of S and α as

$$\nu(f) = \int_{\mathbb{E}} f(s)\nu(dx) = \int_{T(\mathbb{E})} f \circ T^{-1}(r, \boldsymbol{a})c\nu_\alpha(dr)S(d\boldsymbol{a}) = \int_{\aleph_+} k(\boldsymbol{a}; \alpha)S(d\boldsymbol{a}),$$

where

$$k(\boldsymbol{a};\alpha) = c\int_0^\infty f \circ T^{-1}(r,\boldsymbol{a})\nu_\alpha(dr).$$

Note that we suppose the norm $\|\cdot\|$ has been fixed. We will later consider what happens when we switch between different norms.

It is a common view that the extremal dependence structure in a regularly varying vector is carried by the spectral measure S. Basing our approach on this philosophy, we take the representation $\int_{\aleph_+} k(\boldsymbol{a};\alpha)S(d\boldsymbol{a})$ as a natural starting point for considering measures of extremal dependence.

The focus of this paper is on serial dependence in stationary time series, especially over large time intervals. For this purpose the following concept is useful for comparing different dependence measures.

Definition 1. *Fix* $d \geq 2$. *Suppose* ρ_i, $i = 1, 2$, *are two real-valued functions defined on the set of regularly varying vectors* $\boldsymbol{Z} = (Z_1, \dots, Z_d)$ *of the form*

$$\rho_i(\boldsymbol{Z}) = \int_{\aleph_+} k_i(\boldsymbol{a};\alpha)S(d\boldsymbol{a}) \tag{2.3}$$

for some $k_i(\cdot,\alpha) : \aleph_+ \mapsto \mathbb{R}$, *where* $\alpha > 0$ *is the index of regular variation and* S *is the spectral measure of* $\boldsymbol{Z}$. *We call* ρ_1 *and* ρ_2 weakly asymptotically equivalent *if for each fixed* $\alpha > 0$ *there exist constants* $0 < m \leq M < \infty$ *such that*

$$m \leq \liminf_{n\to\infty} \frac{\rho_2(\boldsymbol{Z}_n)}{\rho_1(\boldsymbol{Z}_n)} \leq \limsup_{n\to\infty} \frac{\rho_2(\boldsymbol{Z}_n)}{\rho_1(\boldsymbol{Z}_n)} \leq M, \tag{2.4}$$

for any sequence $(\boldsymbol{Z}_n)$, *where each* $\boldsymbol{Z}_n$ *is a* d*-dimensional regularly varying random vector with index* α. *We also write (2.4) as* $\rho_1(\boldsymbol{Z}_n) \asymp \rho_2(\boldsymbol{Z}_n)$ *as* $n \to \infty$.

Remark 1. One readily checks that weak asymptotic equivalence is an equivalence relation.

Remark 2. One could imagine a stronger form of asymptotic equivalence. For given $\alpha > 0$, one could require the existence of a constant $M > 0$ such that $\rho(\boldsymbol{Z}_n) \sim M\rho_2(\boldsymbol{Z}_n)$ as $n \to \infty$. Although we will not need this stronger notion in this paper, we chose to include the qualifier *weak* in our definition of asymptotic equivalence.

If $(X_n)_{n\in\mathbb{N}}$ is a regularly varying strictly stationary sequence, weak asymptotic equivalence implies that $h \mapsto \rho_1(X_n, X_{n+h})$ is summable precisely when $h \mapsto \rho_2(X_n, X_{n+h})$ is. A major use of this concept will thus be to compare different dependence measures with respect to their conclusions about long-range tail dependence.

There is an easy way of verifying weak asymptotic equivalence:

Proposition 1. *Suppose* ρ_i, $i = 1, 2$, *are as in* (2.3). *Then* ρ_1 *and* ρ_2 *are weakly asymptotically equivalent if and only if for each* $\alpha > 0$ *there are constants* $0 < m \leq M < \infty$ *such that*

$$mk_1(\boldsymbol{a};\alpha) \leq k_2(\boldsymbol{a};\alpha) \leq Mk_1(\boldsymbol{a};\alpha) \quad \textit{for all } \boldsymbol{a} \in \aleph_+. \tag{2.5}$$

Proof. If (2.5) holds, then for any spectral measure S,

$$m\int_{\aleph_+} k_1(\boldsymbol{a};\alpha)S(d\boldsymbol{a}) \leq \int_{\aleph_+} k_2(\boldsymbol{a};\alpha)S(d\boldsymbol{a}) \leq M\int_{\aleph_+} k_1(\boldsymbol{a};\alpha)S(d\boldsymbol{a}),$$

which implies (2.4), so that ρ_1 and ρ_2 are weakly asymptotically equivalent. Conversely, suppose (2.5) does not hold. Then for some $\alpha > 0$ there are $\boldsymbol{a}_n \in \aleph_+$, $n \geq 0$, such that

$$\lim_{n\to\infty} k_1(\boldsymbol{a}_n;\alpha)/k_2(\boldsymbol{a}_n;\alpha) = 0 \text{ or } +\infty.$$

Let $\boldsymbol{Z}_n$ be regularly varying with index α and spectral measure $S_n = \epsilon_{\boldsymbol{a}_n}$, a point mass at $\boldsymbol{a}_n$. Then $\rho_i(\boldsymbol{Z}_n) = k_i(\boldsymbol{a}_n)$, $i = 1, 2$, and thus (2.4) is violated. $\square$

Suppose now that in addition to the norm $\|\cdot\|$ we pick a different norm $\|\cdot\|'$. If we have a dependence measure $\rho(\boldsymbol{Z}) = \int_{\aleph_+} k(\boldsymbol{a}; \alpha) S(d\boldsymbol{a})$, where S is the spectral measure of $\boldsymbol{Z}$ with respect to $\|\cdot\|$, how can we express $\rho(\boldsymbol{Z})$ in terms of S', the spectral measure with respect to $\|\cdot\|'$? For the next result, we let $\aleph'_+ = \{\boldsymbol{x} \in \mathbb{E} : \|\boldsymbol{x}\|' = 1\}$, and define

$$h : \aleph_+ \mapsto \aleph'_+, \ \boldsymbol{a} \mapsto \frac{\boldsymbol{a}}{\|\boldsymbol{a}\|'}.$$

Note that h is the restriction to $\aleph_+$ of the second component of $T' : \boldsymbol{Z} \mapsto (\|\boldsymbol{Z}\|', \boldsymbol{Z}/\|\boldsymbol{Z}\|')$, the polar transformation corresponding to $\|\cdot\|'$. Moreover, one easily checks that h is a bijection with inverse

$$h^{-1} : \aleph'_+ \mapsto \aleph_+, \ \boldsymbol{a}' \mapsto \frac{\boldsymbol{a}'}{\|\boldsymbol{a}'\|}.$$

Proposition 2. *With the above notation,*

$$\int_{\aleph_+} k(\boldsymbol{a}; \alpha) S(d\boldsymbol{a}) = \rho(\boldsymbol{Z}) = \int_{\aleph'_+} k'(\boldsymbol{a}'; \alpha) S'(d\boldsymbol{a}'),$$

where

$$k'(\boldsymbol{a}'; \alpha) = \frac{c'}{c} \|\boldsymbol{a}'\|^{\alpha} k(h^{-1}(\boldsymbol{a}'); \alpha).$$

Here $c' = \nu\{\boldsymbol{x} : \|\boldsymbol{x}\|' > 1\}$, *so that* $\nu \circ (T')^{-1} = c'\nu_\alpha \times S'$.

Proof. For any $f : \aleph_+ \mapsto \mathbb{R}_+$, we have

$$\int_{\aleph_+} f(\boldsymbol{a}) S(d\boldsymbol{a}) = \int_{\aleph_+} f(\boldsymbol{a}) c^{-1} \nu\{\boldsymbol{x} : \|\boldsymbol{x}\| > 1, \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} \in d\boldsymbol{a}\}$$

and noting $\nu = \nu \circ (T')^{-1} \circ T'$ we have

$$= \int_{\aleph_+} f(\boldsymbol{a}) c^{-1} \nu \circ (T')^{-1} \{(r, \boldsymbol{a}') \in \mathbb{R}_+ \times \aleph'_+ : \|r\boldsymbol{a}'\| > 1, \frac{\boldsymbol{a}'}{\|\boldsymbol{a}'\|} \in d\boldsymbol{a}\}.$$

Replace $\nu \circ (T')^{-1}$ with $c'\nu_\alpha \times S'$ and transform the integration to integration on $\aleph'_+$ and we get

$$= \int_{\aleph'_+} f\big(\frac{\boldsymbol{\theta}'}{\|\boldsymbol{\theta}'\|}\big) \frac{c'}{c} \nu_\alpha \times S' \Big\{(r, \boldsymbol{a}') \in \mathbb{R}_+ \times \aleph'_+ : r > \frac{1}{\|\boldsymbol{a}'\|}, \frac{\boldsymbol{a}'}{\|\boldsymbol{a}'\|} \in \frac{d\boldsymbol{\theta}'}{\|\boldsymbol{\theta}'\|}\Big\}$$

$$= \int_{\aleph'_+} f \circ h^{-1}(\boldsymbol{\theta}') \frac{c'}{c} \|\boldsymbol{\theta}'\|^{\alpha} S'(d\boldsymbol{\theta}').$$

This suffices. $\square$

3. REVIEW OF DEPENDENCE MEASURES

3.1. **Extremal dependence measure (EDM).** Given a regularly varying bivariate random vector (Z_1, Z_2) with index α and spectral measure S, we define the *extremal dependence measure* (EDM) as

$$\text{EDM}(Z_1, Z_2) = \int_{\aleph_+} a_1 a_2 S(d\boldsymbol{a}). \tag{3.1}$$

It has the feature that it is zero if and only if S concentrates on

$$\{(1,0)/\|(1,0)\|,(0,1)/\|(0,1)\|\},$$

that is, the limit measure ν concentrates on the coordinate axes in $\mathbb{E}$. This means that (Z_1,Z_2) possesses asymptotic independence (see [15, Section 6.5.1]). Moreover, it is maximal if and only if S concentrates on $\{\boldsymbol{a}: a_1=a_2\}$, i.e. in the case of asymptotic full dependence. Note also that the EDM only depends on the spectral measure, which is consistent with our philosophy that the spectral measure determines the extremal dependence structure.

The EDM depends on the choice of norm. However, changing to a different norm does not affect the asymptotic behavior.

Proposition 3. *Let $\|\cdot\|$ and $\|\cdot\|'$ be two norms on $\mathbb{E}$. Then*

$$EDM(Z_1,Z_2)=\int_{\aleph_+} a_1a_2S(d\boldsymbol{a}) \qquad and \qquad EDM'(Z_1,Z_2)=\int_{\aleph'_+} a'_1a'_2S'(d\boldsymbol{a}')$$

are weakly asymptotically equivalent.

Proof. By Proposition 2, $\mathrm{EDM}'(Z_1,Z_2)=\int_{\aleph_+}\frac{c}{c'}a_1a_2(\|\boldsymbol{a}\|')^{\alpha-2}S(d\boldsymbol{a})$. Since all norms on finite-dimensional space are equivalent, there are constants $0<K_1\le K_2<\infty$ such that $K_1\le\|\boldsymbol{a}\|'\le K_2$ for all $\boldsymbol{a}\in\aleph_+$, where we used that $\|\boldsymbol{a}\|=1$. It follows that

$$ma_1a_2\le\frac{c}{c'}a_1a_2(\|\boldsymbol{a}\|')^{\alpha-2}\le Ma_1a_2 \qquad \text{for all } \boldsymbol{a}\in\aleph_+,$$

where $m=\frac{c}{c'}K_1^{\alpha-2}$ and $M=\frac{c}{c'}K_2^{\alpha-2}$. An application of Proposition 1 now gives the result. □

When $(X_n)_{n\in\mathbb{N}}$ is a regularly varying time series, the EDM can be used to study serial dependence in the tails. The idea is to look at how $\mathrm{EDM}(h):=\mathrm{EDM}(X_n,X_{n+h})$ behaves as a function of the lag. In [8], the EDM is used to study extremal dependence in Internet data. It is also applied in a stochastic process context in [3] to certain data network models. In the latter reference the authors conclude that for the models under consideration, $\mathrm{EDM}(h)$ decays as power function, and they connect this to the concept of long-range dependence. In this paper we define long-range tail dependence in the sense of the EDM to be present if $\mathrm{EDM}(h)$ is not summable.

When introduced in [14], the EDM was originally defined as

$$\mathrm{EDM}(Z_1,Z_2)=1-(4/\pi)^2\int_0^{\pi/2}\left(\theta-\frac{\pi}{4}\right)^2\bar{S}(d\theta)=(4/\pi)^2\int_0^{\pi/2}\theta(\pi/2-\theta)\bar{S}(d\theta), \tag{3.2}$$

where $\bar{S}$ is the spectral measure of (Z_1,Z_2) in polar coordinates. This expression depends on the particular parameterization, which is inconvenient in some situations. We will therefore use the definition (3.1), and remark that the two expressions are weakly asymptotically equivalent. Indeed, by Proposition 3, the EDM under any norm is weakly asymptotically equivalent to the EDM under the Euclidean norm. With respect to this norm, (3.2) becomes

$$\mathrm{EDM}(Z_1,Z_2)=(4/\pi)^2\int_{\aleph_+}\arctan\frac{a_2}{a_1}\left(\frac{\pi}{2}-\arctan\frac{a_2}{a_1}\right)S(d\boldsymbol{a}).$$

It is easy to check that $k_1(\boldsymbol{a};\alpha)=a_1a_2$ and $k_2(\boldsymbol{a};\alpha)=(4/\pi)^2\arctan\frac{a_2}{a_1}\left(\frac{\pi}{2}-\arctan\frac{a_2}{a_1}\right)$ satisfy the conditions of Proposition 1, so that the dependence measures in (3.1) and (3.2) are weakly asymptotically equivalent. This fact, together with the behavior in case of asymptotic independence and asymptotic full dependence, guarantee that the results given in [14] and [3] remain true for the EDM as defined in (3.1).

It is shown in [5, p. 3] that the extremogram has an interesting interpretation as the limit of a sequence of covariance functions. It turns out that the EDM has a similar interpretation.

Proposition 4. *Let* (Z_1, Z_2) *be regularly varying as in* (2.1)*. Then*

$$\text{EDM}(Z_1, Z_2) = \lim_{x\to\infty} \mathrm{E}\left(\frac{Z_1}{R}\frac{Z_2}{R} \mid R > x\right), \tag{3.3}$$

where $R = \|(Z_1, Z_2)\|$.

This lets $\text{EDM}(Z_1, Z_2)$ be understood as the limit of conditional normalized cross-moments of Z_1 and Z_2. It says that the EDM is large if, conditionally on Z_1 and/or Z_2 being large, both are large simultaneously. The EDM is small if whenever Z_1 is large, Z_2 is small, and vice versa.

Proof. Let S be the spectral measure of (Z_1, Z_2), and continue to let T denote the polar coordinate transformation, so that $T(Z_1, Z_2) = (R, \mathbf{\Theta})$, where $R = \|(Z_1, Z_2)\|$ and $\mathbf{\Theta} = (Z_1/R, Z_2/R)$. Changing variables to $x = b(t)$, the regular variation condition (2.2) becomes

$$b^{\leftarrow}(x)\mathrm{P}\left(\left(\frac{R}{x}, \mathbf{\Theta}\right) \in \cdot\right) \xrightarrow{\mathrm{v}} c\nu_\alpha \times S \qquad (x \to \infty). \tag{3.4}$$

One valid choice of normalization is $b^{\leftarrow}(x) = 1/\mathrm{P}(R > x)$, and with this choice we obtain

$$S_x(\cdot) := b^{\leftarrow}(x)\mathrm{P}(x^{-1}R > 1, \mathbf{\Theta} \in \cdot) = \mathrm{P}(\mathbf{\Theta} \in \cdot \mid R > x) \xrightarrow{\mathrm{v}} S(\cdot),$$

since $(1, \infty] \times \Lambda$ is relatively compact in $\mathbb{E}$ for any measurable $\Lambda \subset \aleph_+$. Note that $c = 1$ for this particular $b(\cdot)$, since $S_x(\aleph_+) = 1$ for all x. In the compact space $\aleph_+$ all continuous functions are compactly supported, so we may apply $f : \aleph_+ \to \mathbb{R}$, $(a_1, a_2) \mapsto a_1a_2$ to get

$$S_x(f) \to S(f) = \int_{\aleph_+} a_1a_2S(d\boldsymbol{a}) \qquad (x \to \infty). \tag{3.5}$$

On the other hand,

$$S_x(f) = \int_{\aleph_+} a_1a_2\mathrm{P}(\mathbf{\Theta} \in d\boldsymbol{a} \mid R > x) = \mathrm{E}(\Theta_1\Theta_2 \mid R > x) = \mathrm{E}\left(\frac{Z_1}{R}\frac{Z_2}{R} \mid R > x\right),$$

which gives the result. □

Formula (3.3) suggests a natural estimator for the EDM. If $\boldsymbol{Z}_i = (Z_{1i}, Z_{2i})$, $i = 1, \ldots, n$, is an iid sample distributed as (Z_1, Z_2), we could try to estimate $\text{EDM}(Z_1, Z_2)$ by

$$\widehat{\text{EDM}}(Z_1, Z_2) = \frac{1}{N_n}\sum_{i=1}^{n} \frac{Z_{1i}}{R_i}\frac{Z_{2i}}{R_i}\mathbf{1}_{[R_i>x]}, \tag{3.6}$$

where $R_i = \|\boldsymbol{Z}_i\|$, and

$$N_n = \sum_{i=1}^{n} \mathbf{1}_{[\|\boldsymbol{Z}_i\|>x]}$$

is the number of exceedances over the (suitably chosen) threshold x. Note that this is precisely the estimator that results from integrating $f(a_1, a_2) = a_1a_2$ with respect to the empirical spectral measure,

$$\hat{S}_n = \frac{\sum_{i=1}^n \epsilon_{(R_i/b(t),\mathbf{\Theta}_i)}((1, \infty] \times \cdot)}{\sum_{i=1}^n \epsilon_{R_i/b(t)}(1, \infty]} = \frac{\sum_{i=1}^n \mathbf{1}_{[R_i>b(t)]}\epsilon_{\mathbf{\Theta}_i}}{\sum_{i=1}^n \mathbf{1}_{[R_i>b(t)]}},$$

where $(R_i, \boldsymbol{\Theta}_i) = T(\boldsymbol{Z}_i)$ and $b(t) = x$ is the normalization in (3.4). See [15, Chapter 9.2] for more on how to estimate the spectral measure. In Section 6 we consider the question of asymptotic normality of $\widehat{\text{EDM}}$, extending the discussion in [14].

We conclude by noting that the EDM can be expressed directly in terms of the limit measure ν of (Z_1, Z_2). A direct calculation using that $\nu \circ T^{-1} = c\nu_\alpha \times S$, where $c = \nu\{\|\boldsymbol{x}\| > 1\}$, shows that

$$\text{EDM}(Z_1, Z_2) = \frac{1}{\nu\{\|\boldsymbol{x}\| > 1\}} \int_{\{\|\boldsymbol{x}\|>1\}} \frac{x_1 x_2}{\|\boldsymbol{x}\|^2} \nu(d\boldsymbol{x}).$$

Compare this to (3.3) in Proposition 4.

3.2. **Extremogram.** The second dependence measure we consider is the *extremogram*, introduced in [5]. We first review this concept for a general state space $\mathbb{R}^d$, although we later specialize the setup to $\mathbb{R}_+$. Given a $\mathbb{R}^d$-valued strictly stationary regularly varying time series $(\boldsymbol{X}_n)_{n\in\mathbb{N}}$, let $\nu_{0,\dots,h}$ denote the limit measure of $(\boldsymbol{X}_0, \dots, \boldsymbol{X}_h)$ under the normalization $b(\cdot)$ given by $\mathrm{P}(\|\boldsymbol{X}_0\| > b(t)) \sim t^{-1}$ as $t \to \infty$. For $A, B \subset \mathbb{R}^d$ such that $C = A \times \mathbb{R}^{d(h-1)} \times B$ is bounded away from the origin and ∂C is a continuity set of $\nu_{0,\dots,h}$, the extremogram $\gamma_{AB}(h)$ is defined as

$$\gamma_{AB}(h) = \nu_{0,\dots,h}(C) = \nu_{0,\dots,h}(A \times \mathbb{R}^{d(h-1)} \times B) = \nu_h(A \times B),$$

where ν_h is the limit measure of $(\boldsymbol{X}_0, \boldsymbol{X}_h)$. Note that the extremogram, unlike the EDM, does not require choosing a norm on $\mathbb{E}$. Instead, one must choose which particular sets A and B to use.

Further on we will express the extremogram in terms of the spectral measure S_h, and for this the relation $\nu_h \circ T^{-1} = c_h \nu_\alpha \times S_h$ will be needed. Note that the constant c_h may potentially depend on h. As in the proof of Proposition 4 one can check, using the same transformation as in (3.4), that c_h is given by

$$c_h = \lim_{x\to\infty} \frac{\mathrm{P}(R > t)}{\mathrm{P}(\|\boldsymbol{X}_0\| > t)},$$

where $R = \|(\boldsymbol{X}_0, \boldsymbol{X}_h)\|$ is the $2d$-dimensional norm. As it turns out, the fact that c_h depends on h does not affect our asymptotic analysis. Indeed, by equivalence of norms, there are $m, M > 0$ such that

$$m\|\boldsymbol{X}_0\| \vee \|\boldsymbol{X}_h\| \le R \le M\|\boldsymbol{X}_0\| \vee \|\boldsymbol{X}_h\|.$$

(It is straightforward to check that $\|\cdot\| \vee \|\cdot\|$ is a norm on $\mathbb{R}^d \times \mathbb{R}^d$.) Therefore,

$$\mathrm{P}(R > x) \ge \mathrm{P}(m\|\boldsymbol{X}_0\| \vee \|\boldsymbol{X}_h\| > x) \ge \mathrm{P}(\|\boldsymbol{X}_0\| > x/m),$$

and

$$\begin{aligned}\mathrm{P}(R > x) \le& \mathrm{P}(M\|\boldsymbol{X}_0\| \vee \|\boldsymbol{X}_h\| > x) \le \mathrm{P}\left(\{\|\boldsymbol{X}_0\| > x/M\} \cup \{\|\boldsymbol{X}_h\| > x/M\}\right) \\ \le& 2\mathrm{P}(\|\boldsymbol{X}_0\| > x/M).\end{aligned}$$

Regular variation of $\mathrm{P}(\|\boldsymbol{X}_0\| > x)$ then gives

$$m^\alpha \le c_h \le 2M^\alpha,$$

regardless of h. Any statement about the asymptotic rate of decay or summability of $\gamma_{AB}(h)$ is thus independent of the precise behavior of c_h. In our analysis, therefore, we will for convenience leave the normalization $b(\cdot)$ unspecified, allowing it to be such that either $\mathrm{P}(\|\boldsymbol{X}_0\| > b(t)) \sim t^{-1}$ or $\mathrm{P}(R > b(t)) \sim t^{-1}$.

As it depends on the sets A and B, the extremogram is really a family of dependence measures. In [5], the extremogram is computed in the case $d = 1$ with $A = (u, \infty)$ and $B = (w, \infty)$ for a number of commonly encountered processes (e.g. the stochastic volatility model, GARCH, ARMA, etc.) For the purpose of comparison with the EDM, we will focus on this case, with $u, w > 0$, and moreover assume that each $\boldsymbol{X}_n = X_n$ is $\mathbb{R}_+$-valued.

In [5] the authors also propose a notion of long-range extremal dependence, namely the lack of summability of $\gamma_{AB}(h)$. We use the same notion in our comparison with the EDM.

We remark that the *extremal coefficient function* described in [7] coincides with the extremogram, up to a multiplicative constant, when $d = 1$ and $A = B = (1, \infty)$. In the same paper, the authors also propose two *extreme dependence functions* to study dependence between several time lags $X_{h_1}, \dots, X_{h_d}$ of the process simultaneously. We will not discuss them here, but mention only that they have essentially the same structure as the extremogram: both equal the limit measure of $(X_{h_1}, \dots, X_{h_d})$ evaluated in intersections and unions of rectangles.

The extremogram fits into the framework of Section 2 and may thus be expressed in terms of α and S. Here and in what follows, $\alpha > 0$ denotes the index of regular variation of the strictly stationary time series (X_n), and S_h is the spectral measure of (X_n, X_{n+h}).

Lemma 1. *With $A = (u, \infty)$ and $B = (w, \infty)$ and under the normalization $b^{\leftarrow}(x) = 1/\mathrm{P}(\|(X_0, X_h)\| > x)$,*

$$\gamma_{AB}(h) = \int_{\aleph_+} \left(\frac{a_0}{u}\right)^{\alpha} \wedge \left(\frac{a_h}{w}\right)^{\alpha} S_h(d\boldsymbol{a}). \tag{3.7}$$

Proof. Note that $A \times B = \{(x_0, x_h) : \frac{x_0}{u} \wedge \frac{x_h}{w} > 1\}$. With T being the polar transformation $T(\boldsymbol{x}) = (\|\boldsymbol{x}\|, \boldsymbol{x}/\|\boldsymbol{x}\|)$,

$$T(A \times B) = \left\{(r, (a_0, a_h)) : \frac{ra_0}{u} \wedge \frac{ra_h}{w} > 1\right\} = \left\{(r, (a_0, a_h)) : r > \left(\frac{a_0}{u} \wedge \frac{a_h}{w}\right)^{-1}\right\}.$$

Under the given normalization, $\nu_h \circ T^{-1} = \nu_\alpha \times S_h$, and so

$$\begin{aligned}\nu_h(A \times B) &= (\nu_h \circ T^{-1})(T(A \times B)) = \nu_\alpha \times S_h(T(A \times B)) \\ &= \int_{\aleph_+} \int_{\left\{r > \left(\frac{a_0}{u} \wedge \frac{a_h}{w}\right)^{-1}\right\}} \nu_\alpha(dr) S_h(d\boldsymbol{a}) = \int_{\aleph_+} \left(\frac{a_0}{u} \wedge \frac{a_h}{w}\right)^{\alpha} S_h(d\boldsymbol{a}) \\ &= \int_{\aleph_+} \left(\frac{a_0}{u}\right)^{\alpha} \wedge \left(\frac{a_h}{w}\right)^{\alpha} S_h(d\boldsymbol{a}).\end{aligned}$$

□

In particular, this immediately implies that for $u, w > 0$, the extremogram is zero if and only if (X_n, X_{n+h}) possesses asymptotic independence, so in this respect it is similar to the EDM. Moreover, the integrand in (3.7) is maximal precisely when $u^{-1} a_0 = w^{-1} a_h$. Thus $\gamma_{AB}(h)$ is maximal if and only if S_h concentrates on $\{\boldsymbol{a} : a_0/a_h = u/w\}$. If $u = w$ this is $\{\boldsymbol{a} : a_0 = a_h\}$, so that $\gamma_{AB}(h)$ is maximal if and only if asymptotic full dependence is present. Another property that will let us simplify things later is the following.

Proposition 5. *Let $A_i = (u_i, \infty)$, $B_i = (w_i, \infty)$, $i = 1, 2$, and suppose $u_i, w_i > 0$. Then $\gamma_{A_1 B_1}$ and $\gamma_{A_2 B_2}$ are weakly asymptotically equivalent.*

Proof. For $\boldsymbol{a}$ near $(0, 1)/\|(0, 1)\|$, the ratio between the integrands in the representation (3.7) of $\gamma_{A_1 B_1}$ and $\gamma_{A_2 B_2}$ is equal to $(u_2/u_1)^\alpha > 0$. Similarly, if $\boldsymbol{a}$ is near $(1, 0)/\|(1, 0)\|$, the ratio is $(w_2/w_1)^\alpha > 0$. Both integrands are nonzero on the interior of $\aleph_+$, and since the ratio is continuous, finite and bounded away from zero, this gives weak asymptotic equivalence via Proposition 1. □

The next result tells us when the EDM and the extremogram are weakly asymptotically equivalent.

Proposition 6. *The EDM and the extremogram (with $A = (u, \infty)$, $B = (w, \infty)$, $u, w > 0$) are weakly asymptotically equivalent if and only if $\alpha = 1$, i.e. in the* standard case.

Proof. The proof is similar to that of Proposition 5. As $\boldsymbol{a} \to (0,1)/\|(0,1)\|$, the integrand in (3.7) is eventually equal to $u^{-\alpha} a_0^{\alpha}$. Moreover, the integrand in (3.1) is $a_0 a_h \sim a_0 \|(0,1)\|$. Hence if $\alpha \neq 1$, their ratio tends either to 0 or $+\infty$ as $\boldsymbol{a} \to (0,1)/\|(0,1)\|$, and thus one of the required constants in Proposition 1 fails to be positive respectively finite. Hence weak asymptotic equivalence does not hold. Conversely, if $\alpha = 1$, the ratio tends to $u\|(0,1)\|$ as $\boldsymbol{a} \to (0,1)/\|(0,1)\|$, and a similar argument shows that the limit is $w\|(1,0)\|$ when $\boldsymbol{a} \to (1,0)/\|(1,0)\|$. Furthermore, both the numerator and the denominator are nonzero on the interior of $\aleph_+$. Since the ratio is continuous we conclude that it is finite and bounded away from zero on $\aleph_+$. By Proposition 1, we have weak asymptotic equvalence. $\square$

Remark. If, as an example, $(X_n)_{n\in\mathbb{N}}$ is a regularly varying ARMA model, α does not only determine the tails of the marginals X_n, but may also affect the spectral measure of (X_n, X_{n+h}). This may potentially influence the rate of decay of $\mathrm{EDM}(X_n, X_{n+h})$. In concrete models, therefore, α may be a parameter that affects the spectral measures S_h. Thus if $\gamma_{AB}(h)$ is available as a function of α, one cannot simply set $\alpha = 1$, extract the rate of decay of $\gamma_{AB}(h)$, and use Proposition 6 to conclude that the EDM will decay at that rate for all values of α. This is elaborated on in Section 4.

The extremogram has the interesting property that its asymptotic decay is invariant when the observations of the time series are transformed to the standard case $\alpha = 1$.

Proposition 7. *Let (X_n) be a strictly stationary regularly varying time series with extremogram $\gamma_{AB}(h)$, $A = (u, \infty)$, $B = (w, \infty)$, $u, w > 0$. Define $X_n^* = b^{\leftarrow}(X_n)$, and let $\gamma^*_{AB}(h)$ be the extremogram of the (standard regularly varying, strictly stationary) time series (X_n^*). Then γ_{AB} and γ^*_{AB} are weakly asymptotically equivalent.*

Proof. By Proposition 5, we may assume without loss of generality that $u = w = 1$. As usual, let ν_h be the limit measure of (X_n, X_{n+h}), $\alpha > 0$ the index of regular variation, and S_h the spectral measure. Then we have

$$\begin{aligned}\nu_h^*(\mathbf{1}, \boldsymbol{\infty}] =& \lim_{t\to\infty} \mathrm{P}\left(\frac{X_0^*}{t} > 1, \frac{X_h^*}{t} > 1\right) = \lim_{t\to\infty} \mathrm{P}(X_0 > b(t), X_h > b(t)) \\ =& \nu_h(\mathbf{1}, \boldsymbol{\infty}].\end{aligned}$$

So γ_{AB} and γ^*_{AB} are trivially weakly asymptotically equivalent. $\square$

The following result is a consequence of the previous two propositions.

Proposition 8. *Let (X_n) be a strictly stationary regularly varying time series and define $X_n^* = b^{\leftarrow}(X_n)$. Let $\gamma_{AB}(h)$, $A = (u, \infty)$, $B = (w, \infty)$, $u, w > 0$, be the extremogram of (X_n) and let EDM*(h) be the EDM of (X_n^*). Then γ_{AB} and EDM* are weakly asymptotically equivalent.*

Proof. Let $\gamma^*_{AB}(h)$ be the extremogram of (X_n^*). By Proposition 6, γ^*_{AB} and EDM^* are weakly asymptotically equivalent. Furthermore, Proposition 7 implies that γ_{AB} and γ^*_{AB} are weakly asymptotically equivalent. The result now follows since weak asymptotic equivalence is an equivalence relation. $\square$

4. Long-range tail dependence; some examples

Long-range dependence can be vaguely understood as the persistence over large time intervals of the influence of specific outcomes on the evolution of the process. Formulated differently, a long-range dependent process has long memory in the sense that realized outcomes are "remembered" for a long time. See Samorodnitsky [17] for a survey of this topic, and Beran [1] for issues related to statistical inference.

Several different notions of long-range dependence have been suggested, the most common ones being based on second-order quantities such as autocovariance and autocorrelation functions. Usually one considers long-range dependence to be present if the autocorrelation function is not summable, and absent otherwise. As we have seen, in the heavy-tail context we now have objects akin to autocorrelation functions available, such as the extremogram and the EDM. It therefore appears natural to base a notion of long-range tail dependence on summability of these quantites. This has indeed been suggested in [5] for the extremogram. In [3] the EDM was used to measure the decay of tail dependence in certain data network models.

The purpose of this section is to look at some simple examples to get an understanding of the difficulties associated with finding a universal definition of long-range tail dependence. It turns out that the extremogram and the EDM, which appear similar at first sight, may yield very different conclusions about long-range dependence.

4.1. **Max-moving averages.**

Max-moving average processes, or, more generally, Max-ARMA processes, were studied, among others, by Hsing [9], Davis and Resnick [6], and Zhang and Smith [19]. We use the following formulation. Let Z_n, $n \in \mathbb{Z}$, be iid Fréchet with parameter $\alpha > 0$, i.e. $\mathrm{P}(Z_n \le x) = \Phi_\alpha(x) = \exp(-x^{-\alpha})$. Let $(\psi_i)_{i=0,1,\dots}$ be a sequence of nonnegative numbers, and define

$$X_n = \bigvee_{i=0}^{\infty} \psi_i Z_{n-i}, \quad n \in \mathbb{N}. \tag{4.1}$$

We require that $\sigma := \sum_{i=0}^{\infty} \psi_i^\alpha < \infty$. This ensures that the process exists and is strictly stationary (cf. [9]). We first collect some useful facts.

Proposition 9. *Let (X_n) be the max-moving average process in (4.1). Then*

(i) $\mathrm{P}(X_n \le x) = \Phi_{\alpha,\sigma}(x) = \exp\{-\sigma x^{-\alpha}\}$ *(Fréchet marginals)*

(ii) The limit measure ν_h of (X_n, X_{n+h}) is

$$\nu_h([0,(x,y)]^c) = \sigma \sum_{i=0}^{\infty} \left(\frac{\psi_i}{x} \vee \frac{\psi_{i+h}}{y}\right)^\alpha + \sigma \sum_{i=0}^{h-1} \left(\frac{\psi_i}{y}\right)^\alpha .$$

(iii) Let $\|\cdot\| = \|\cdot\|_\infty$ be the supremum norm on $\mathbb{R}_+^2$. With respect to $\|\cdot\|$, the spectral measure corresponding to ν_h is

$$S_h = \frac{1}{c}\left(\sum_{i=0}^{\infty} (\psi_i \vee \psi_{i+h})^\alpha \epsilon_{(\psi_i,\psi_{i+h})/(\psi_{i+h}\vee\psi_i)} + \sum_{i=0}^{h-1} \psi_i^\alpha \epsilon_{(0,1)}\right),$$

where

$$c = \sum_{i=0}^{\infty} (\psi_i \vee \psi_{i+h})^\alpha + \sum_{i=0}^{h-1} \psi_i^\alpha .$$

Proof. These computations are similar to others appearing in the literature; see for example [6]. We first compute the two-dimensional marginals $F_h(x,y) = \mathrm{P}(X_n \le x, X_{n+h} \le y)$ (by stationarity, F_h does not depend on n).

$$\begin{aligned} F_h(x,y) &= \mathrm{P}(X_n \le x,\ X_{n+h} \le y) = \mathrm{P}(\psi_i Z_{n-i} \le x,\ \psi_i Z_{n+h-i} \le y;\ i = 0,1,\dots) \\ &= \mathrm{P}\left\{Z_{n-i} \le \frac{x}{\psi_i} \wedge \frac{y}{\psi_{i+h}};\ i = 0,1,\dots\right\} \cap \left\{Z_{n+h-i} \le \frac{y}{\psi_i};\ i = 0,\dots,h-1\right\} \end{aligned}$$

$$
\begin{aligned}
&= \prod_{i=0}^{\infty} \Phi_{\alpha,\sigma}\left(\frac{x}{\psi_i} \wedge \frac{y}{\psi_{i+h}}\right) \times \prod_{i=0}^{h-1} \Phi_{\alpha,\sigma}\left(\frac{y}{\psi_i}\right) \\
&= \exp\left\{-\sigma \sum_{i=0}^{\infty}\left(\frac{\psi_i}{x} \vee \frac{\psi_{i+h}}{y}\right)^{\alpha} - \sigma \sum_{i=0}^{h-1}\left(\frac{\psi_i}{y}\right)^{\alpha}\right\}.
\end{aligned}
$$

Clearly $F_h(t^{-1/\alpha}x, t^{-1/\alpha}y) = F_h^t(x,y)$, so F_h is max-stable. Thus F_h is its own extreme value limit, and the limit measure ν_h is given by $\nu_h([0,(x,y)]^c) = -\log F_h(x,y)$, from which (ii) follows. Next,

$$
\mathrm{P}(X_n \le x) = F_h(x,\infty) = \exp\left\{-x^{-\alpha}\sum_{i=0}^{\infty}\psi_i^{\alpha}\right\} = \Phi_{\alpha,\sigma}(x),
$$

which gives (i). Finally, it is well-known (see e.g. [15, Proposition 6.4]) that

$$
\nu_h([0,(x,y)]^c) = \hat{c}\int_{\aleph_+}\left(\frac{x}{a_0}\right)^{-\alpha} \vee \left(\frac{y}{a_h}\right)^{-\alpha} S_h(d\boldsymbol{a}) = \hat{c}\int_{\aleph_+}\left(\frac{a_0}{x} \vee \frac{a_h}{y}\right)^{\alpha} S_h(d\boldsymbol{a})
$$

for some constant $\hat{c}$. With S_h as in (iii) and $\hat{c} = \sigma c$, the right side is

$$
\sigma\sum_{i=0}^{\infty}(\psi_i \vee \psi_{i+h})^{\alpha}\left(\frac{\psi_i/(\psi_i \vee \psi_{i+h})}{x} \vee \frac{\psi_{i+h}/(\psi_i \vee \psi_{i+h})}{y}\right)^{\alpha} + \sigma\sum_{i=0}^{h-1}\left(\frac{\psi_i}{y}\right)^{\alpha},
$$

which indeed equals $\nu_h([0,(x,y)]^c)$. Since ν_h is determined by its values on sets of the form $[0,(x,y)]^c$, and since the spectral measure is unique, we obtain (iii). Notice also that c is the correct normalization: $cS_h(\aleph_+) = \sum_{i=0}^{\infty}(\psi_i \vee \psi_{i+h})^{\alpha} + \sum_{i=0}^{h-1}\psi_i^{\alpha} = c$. ☐

Remark. The limit measure in Proposition 9 uses the normalization $b(t) = t^{1/\alpha}$. It can be seen from the proof that this yields $\nu \circ T^{-1} = \sigma c \nu_\alpha \times S$.

What happens to the limit measure ν_h in Proposition 9 (ii) as h becomes large? Observe that $\lim_{h\to\infty}\sum_{i=0}^{h-1}\psi_i^{\alpha} = \sum_{i=0}^{\infty}\psi_i^{\alpha} = \sigma$. Moreover, $\lim_{h\to\infty}\psi_{i+h} = 0$, so

$$
\frac{\psi_i}{x} \vee \frac{\psi_{i+h}}{y} \to \frac{\psi_i}{x} \qquad (h \to \infty).
$$

Since also

$$
\sum_{i=0}^{\infty}\left(\frac{\psi_i}{x} \vee \frac{\psi_{i+h}}{y}\right)^{\alpha} \le (x^{-\alpha} + y^{-\alpha})\sum_{i=0}^{\infty}\psi_i^{\alpha} < \infty,
$$

dominated convergence yields

$$
\sum_{i=0}^{\infty}\left(\frac{\psi_i}{x} \vee \frac{\psi_{i+h}}{y}\right)^{\alpha} \to \sum_{i=0}^{\infty}\left(\frac{\psi_i}{x}\right)^{\alpha} = \sigma x^{-\alpha} \qquad (h \to \infty).
$$

Hence $\nu_h([0,(x,y)]^c) \to \sigma x^{-\alpha} + \sigma y^{-\alpha}$ as $h \to \infty$. This means that ν_h converges vaguely to a limit which concentrates on the axes, so that X_n and X_{n+h} are approximately asymptotically independent for large h. We are interested in how fast the strength of dependence decays.

Proposition 10. *For the max-moving average process (X_n) in (4.1) we have the following.*

(i) The extremogram at lag h with $A = (1,\infty)$, $B = (1,\infty)$ is

$$
\gamma_{AB}(h) = 2\sigma - \sum_{i=0}^{\infty}(\psi_i \vee \psi_{i+h})^{\alpha} - \sum_{i=0}^{h-1}(\psi_i)^{\alpha}.
$$

(ii) The EDM at lag h with respect to the norm $\|\cdot\|_\infty$ is

$$\mathit{EDM}(h)=\frac{\sum_{i=0}^{\infty}\psi_i^{\alpha-1}\psi_{i+h}}{\sum_{i=0}^{\infty}(\psi_i\vee\psi_{i+h})^{\alpha}+\sum_{i=0}^{h-1}\psi_i^{\alpha}}.$$

Proof. Proposition 9 implies $\nu_h((x,\infty)\times[0,\infty))=\sigma x^{-\alpha}$ and $\nu_h([0,\infty)\times(y,\infty))=\sigma y^{-\alpha}$. Thus

$$\begin{aligned}\gamma_{AB}(h) &= \nu_h(A\times B)=\nu_h((1,\infty)\times[0,\infty))+\nu_h([0,\infty)\times(1,\infty))-\nu_h([\mathbf{0},\mathbf{1}]^c)\\ &= \sigma+\sigma-\sum_{i=0}^{\infty}(\psi_i\vee\psi_{i+h})^{\alpha}-\sum_{i=0}^{h-1}(\psi_i)^{\alpha}.\end{aligned}$$

The expression for the EDM follows from (3.1) and Proposition 9 (*iii*). □

4.1.1. *Analysis of the extremogram.* Consider first the extremogram. We confine attention to $A=B=(1,\infty)$ without loss of generality (cf. Proposition 5.) If we also assume that (ψ_i) is a monotonically decreasing sequence, then

$$\gamma_{AB}(h)=2\sigma-\sum_{i=0}^{\infty}\psi_i^{\alpha}-\sum_{i=0}^{h-1}\psi_i^{\alpha}=\sum_{i=h}^{\infty}\psi_i^{\alpha}.$$

Let us consider the particular case where $\psi_i=(1+i)^{-\beta}$, with $\beta>1/\alpha$ to ensure summability of ψ_i^{α}. In this case

$$\gamma_{AB}(h)=\sum_{i=h}^{\infty}\psi_i^{\alpha}\sim\int_h^{\infty}x^{-\alpha\beta}dx=\frac{1}{\alpha\beta-1}h^{-\alpha\beta+1}\quad(h\to\infty).$$

We conclude that

$$\gamma_{AB}(h)\sim\frac{1}{\alpha\beta-1}h^{-\alpha\beta+1}\quad(h\to\infty).$$

In particular this means that $\gamma_{AB}(h)$ is summable if and only if $\alpha\beta>2$. Thus, if one takes summability of $\gamma_{AB}(h)$ as lack of long-range dependence, then (X_n) exhibits long-range dependence if and only if $\alpha\beta\le 2$.

Consider now the transformed series (X_n^*), where $X_n^*=X_n^{\alpha}$. Then (X_n^*) is a max-moving average process with innovations $Z_n^*=Z_n^{\alpha}\sim\Phi_{1,\sigma}(x)$ and coefficients $\psi_i^*=\psi_i^{\alpha}$. The extremogram $\gamma_{AB}^*(h)$ of (X_n^*) is thus

$$\gamma_{AB}^*(h)=\sum_{i=h}^{\infty}(\psi_i^*)^1=\sum_{i=h}^{\infty}\psi_i^{\alpha},$$

which is equal to $\gamma_{AB}(h)$. Hence the rate of decay of the extremogram is invariant under this type of transformation. In view of Proposition 7, this is no surprise, since $b(t)=t^{1/\alpha}$ is a possible choice of normalization in the case of Fréchet innovations.

4.1.2. *Analysis of the EDM.* The EDM requires a slightly more sophisticated analysis. We continue to look at the example where $\psi_i=(1+i)^{-\beta}$, $\beta>1/\alpha$. Since the coefficient sequence is decreasing, we get

$$\mathrm{EDM}(h)=\frac{\sum_{i=0}^{\infty}\psi_i^{\alpha-1}\psi_{i+h}}{\sum_{i=0}^{\infty}\psi_i^{\alpha}+\sum_{i=0}^{h-1}\psi_i^{\alpha}}=\frac{\sum_{i=0}^{\infty}\psi_i^{\alpha-1}\psi_{i+h}}{\sigma+\sum_{i=0}^{h-1}\psi_i^{\alpha}},$$

and because $\sum_{i=0}^{h-1}\psi_i^\alpha \to \sigma$ as $h \to \infty$, $\mathrm{EDM}(h) \sim \frac{1}{2\sigma}\sum_{i=0}^{\infty}\psi_i^{\alpha-1}\psi_{i+h}$ as $h \to \infty$. With our particular choice of ψ_i, we have, as $n \to \infty$,

$$\mathrm{EDM}(h) \sim \frac{1}{2\sigma}\sum_{i=0}^{\infty}[(1+i)^{-\beta}]^{\alpha-1}(1+i+h)^{-\beta} = \frac{1}{2\sigma}\sum_{i=1}^{\infty}[i^{\alpha-1}(i+h)]^{-\beta}. \tag{4.2}$$

Define $I(h) = \int_1^\infty [x^{\alpha-1}(x+h)]^{-\beta}dx$ and observe that

$$I(h) \ \le \ \sum_{i=1}^{\infty}[i^{\alpha-1}(i+h)]^{-\beta} \ \le \ (1+h)^{-\beta} + I(h). \tag{4.3}$$

We are interested in the asymptotic behavior of $I(h)$. Changing variables to $t = h/x$ yields

$$I(x) = h^{-\alpha\beta+1}\int_0^h t^{\alpha\beta-\beta-2}(1+t^{-1})^{-\beta}dt.$$

Divide into three different cases according to the relation between α and β.

1. $\alpha < 1 + 1/\beta$: Then $\alpha\beta - \beta - 2 < -1$, so
$$\int_0^h t^{\alpha\beta-\beta-2}(1+t^{-1})^{-\beta}dt \le \int_0^h t^{\alpha\beta-\beta-2}dt \to \text{const} < \infty \qquad (h \to \infty).$$
Therefore, since $\int_0^h t^{\alpha\beta-\beta-2}(1+t^{-1})^{-\beta}dt$ is increasing in h a finite limit exists, and thus $I(h) \sim \text{const}\cdot h^{-\alpha\beta+1}$. Moreover, $\alpha < 1+1/\beta$ implies $\beta > \alpha\beta - 1$, so that $(1+h)^{-\beta}/I(h) \to 0$. Then from (4.2) and (4.3) we obtain
$$\mathrm{EDM}(h) \sim \frac{1}{2\sigma}\text{const}\cdot h^{-\alpha\beta+1}.$$
2. $\alpha > 1 + 1/\beta$: Since $t \mapsto (1+t^{-1})^{-\beta}$ is slowly varying, Karamata's Theorem yields
$$\int_0^h t^{\alpha\beta-\beta-2}(1+t^{-1})^{-\beta}dt \sim \frac{h^{\alpha\beta-\beta-1}(1+t^{-1})^{-\beta}}{\alpha\beta-\beta-1},$$
so that $I(h) \sim \text{const}\cdot h^{-\beta}$. Since $(1+h)^{-\beta}/h^{-\beta} \to 1$, (4.3) only lets us conclude, together with (4.2), that
$$\mathrm{EDM}(h) \asymp h^{-\beta} \qquad (h \to \infty).$$
3. $\alpha = 1 + 1/\beta$: In this case $I(h) = h^{-\beta}\int_0^h t^{-1}(1+t^{-1})^{-\beta}dt =: h^{-\beta}L(h)$. By Karamata's Theorem L is slowly varying. Note that $L(h) \to \infty$ as $h \to \infty$, so $(1+h)^{-\beta}/I(h) \to 0$ and we obtain
$$\mathrm{EDM}(h) \sim \frac{1}{2\sigma}h^{-\beta}L(h) \qquad (h \to \infty).$$

The three cases case together say that the EDM decays at a rate equal to $\min(\beta, \alpha\beta - 1)$, which is positive due to the stationarity requirement $\alpha > 1/\beta$.

Thus, as measured by the EDM, extremal dependence decays at a rate that depends on the innovations when α is relatively small, but not when α is relatively large. That is, very heavy innovation tails tend to distort the amount of memory in the process, whereas in the presence of lighter-tailed innovations, the memory is governed solely by the decay of the coefficients ψ_i.

In the regime of Case 1 above, where α is relatively small, the EDM and the extremogram decay at the same rate, and thus agree in their conclusion about long-range tail dependence. Furthermore, if $\beta \le 1$, long-range dependence in the sense of the EDM will always be present. Indeed, in Case 2 and 3, $\beta \le 1$ implies that $\mathrm{EDM}(h)$ is not summable. In Case 1, we have $\alpha\beta < \beta + 1 \le 2$, and so $\mathrm{EDM}(h)$ is again not summable.

Now consider again the transformed sequence (X_n^*), where $X_n^* = X_n^\alpha$. The new sequence is max-moving average with $\Phi_{1,\sigma}$-innovations and coefficients $\psi_i^* = (1+i)^{-\alpha\beta}$. The previous argument then shows that the EDM decays with rate $\min(\alpha\beta, \alpha\beta - 1) = \alpha\beta - 1$. Note that this is the same as the rate of decay of the extremogram for the original sequence, which is precisely what we expect based on Proposition 8.

We emphasize that this example also provides an illustration to the remark after Proposition 6. Indeed, for many choices of α and β, the extremogram and the EDM have different rates of decay. However, if $\alpha = 1$ we are always in Case 1. Then the two measures decay at the same rate, as is required from Proposition 6.

4.2. **MA(∞) and AR(1) with heavy-tailed innovations.**

Let us consider a heavy-tailed MA process $(X_n)_{n\in\mathbb{N}}$ with coefficients ψ_i, $i \in \mathbb{N}$, i.e.,

$$X_n = \sum_{i=0}^{\infty} \psi_i Z_{n-i}, \qquad n \in \mathbb{N}, \tag{4.4}$$

where $\{Z_n, n \in \mathbb{Z}\}$ are iid and $\mathrm{P}(Z_n > x) \in \mathrm{RV}_{-\alpha}$ for some $\alpha > 0$. See Davis and Resnick [4] for more details on such processes. We consider the special case where the Z_n and ψ_i are all nonnegative. A strictly stationary version of (X_n) exists under mild conditions on the coefficients ψ_i (cf Hult and Samorodnitsky [10].) As remarked by Hsing [9, p. 56], the point process limit of the process (X_n) coincides with that of the max-moving average process with coefficients ψ_i and Fréchet innovations. The limit measures of the finite-dimensional marginals are thus the same, so Proposition 10 gives us the extremogram and EDM also for (nonnegative) regularly varying MA processes of the form (4.4).

Specialize to the case where (X_n) is an AR(1) process with parameter $\phi \in (0,1)$. Its MA representation (4.4) has $\psi_i = \phi^i$. We get

$$\mathrm{EDM}(h) = \frac{\sum_{i=0}^{\infty} \phi^{i\alpha + h}}{\sum_{i=0}^{\infty} \phi^{i\alpha} + \sum_{i=0}^{h-1} \phi^{i\alpha}} = \frac{\phi^h \frac{1}{1-\phi^\alpha}}{\frac{1}{1-\phi^\alpha} + \frac{1-\phi^{h\alpha}}{1-\phi^\alpha}} = \frac{\phi^h}{2 - \phi^{h\alpha}}.$$

With $A = B = (1, \infty)$ we recover results (up to a constant) from [5, Section 2.6] and [7, Theorem 3.3]:

$$\gamma_{AB}(h) = \sum_{i=h}^{\infty} \phi^{i\alpha} = \phi^{h\alpha} \sum_{i=0}^{\infty} \phi^{i\alpha} = \frac{\phi^{h\alpha}}{1 - \phi^\alpha}.$$

This example does not exhibit long-range tail dependence. Nonetheless, there are a few interesting observations to be made. Notice that for fixed h, both $\gamma_{AB}(h)$ and $\mathrm{EDM}(h)$ increase as α decreases. Heavier innovation tails thus yield stronger serial dependence. However, in the case of the EDM, the dependence on α is *local* in the sense that whatever the value of α, we have $\mathrm{EDM}(h) \sim \phi^h/2$ as $h \to \infty$.

The extremogram, on the other hand, behaves differently. Here the rate of decay depends crucially on α. Moreover, as $\alpha \to \infty$, meaning that the innovation tails become lighter, the extremogram tends to zero for all lags, approaching a situation with no extremal dependence at all. The same is true for the max-moving average example, where the model coefficients had a power type decay. There too the extremogram tends to zero for each lag as $\alpha \to \infty$. Extrapolating to the case with light-tailed innovations, this indicates that the extremogram, suitably defined for this non-regularly varying case, should be zero for all positive lags. In particular this would exclude long-range tail dependence in such models.

5. SUMMARY

In this section we give a concise summary of the previous findings, with the aim of highlighting the similarities and differences between the EDM and the extremogram.

Notions of long-range tail dependence differ. The max-moving average example shows that the EDM and the extremogram decay at different rates in general. In the model where the coefficients decay as $i^{-\beta}$ we obtained

$$\gamma_{AB}(h) \sim \text{const} \cdot h^{-\alpha\beta+1} \qquad \text{and} \qquad \text{EDM}(h) \asymp h^{-(\alpha\beta-1)\wedge\beta}.$$

In the case of geometrically decaying coefficients (the AR(1) model with parameter ϕ):

$$\gamma_{AB}(h) \sim \text{const} \cdot \phi^{h\alpha} \qquad \text{and} \qquad \text{EDM}(h) \sim \text{const} \cdot \phi^{h}.$$

One can summarize the difference as follows: the decay rate of the extremogram always depends crucially on α. The decay rate of the EDM depends of α only when tails are sufficiently heavy (α is sufficiently small) *and* the coefficients decay slowly, e.g. as a power function (so far we have of course only shown this for the particular case of max-moving averages.)

Behavior under standardization. By Proposition 7, the the extremogram has invariant asymptotics when the time series is standardized. This is not true for the EDM; a counterexample is the max-moving average model considered in Section 4. However, since the EDM after standardization agrees with the extremogram (Proposition 8), one way to reconcile the differences would be to always standardize before computing dependence measures. There have been objections to using such a procudure, for instance in [7, p. 16]. The argument is that standardization in general changes the model, and thus the dependence structure. A consequence of this view and Proposition 7 is that the extremogram fails to capture aspects of the (asymptotic) dependence structure that change in the standardization process. Note, however, that adopting this view has the consequence that one views the copula transformation as changing dependence structure.

Basic invariance. Although the EDM depends on the choice of norm, we have seen that the particular choice does not influence the asymptotics, see Proposition 3. Similarly, the asymptotics of the extremogram for the sets $A = (u, \infty)$, $B = (w, \infty)$ does not depend on the particular choice of u and w, as long as they are strictly positive, see Proposition 5.

Max-stable limits. Let $\boldsymbol{X} = (X_n)$ be an $\mathbb{R}_+$-valued regularly varying sequence, and let $\{\boldsymbol{X}(j) : j = 1, 2, \ldots\}$ be iid copies of the sequence $\boldsymbol{X}$. In [14, Theorem 2] it is shown that if EDM(h) is zero for all h beyond some lag h_0, then

$$\frac{\bigvee_{j=1}^{m} \boldsymbol{X}(j)}{b(m)} \Rightarrow \boldsymbol{M} \qquad (m \to \infty)$$

in $\mathbb{R}^\infty$ for some max-stable process $\boldsymbol{M} = (M_n)$ with the property that M_n and M_{n+h} are independent for $h \geq h_0$. We do not intend to discuss this result further, but only mention that the proof relies on the fact that EDM$(h) = 0$ if and only if X_n and X_{n+h} are asymptotically independent. The same holds for the extremogram with $A = (u, \infty)$, $B = (w, \infty)$, $u, w > 0$ (see the comments after Lemma 1), so the result in [14] is still true under the alternative assumption $\gamma_{AB}(h) = 0$ for all $h \geq h_0$.

6. A CENTRAL LIMIT THEOREM FOR THE EDM

In [14], asymptotic normality is proven for an estimator of the EDM in the case of iid observations. More precisely, for $\boldsymbol{Z}_i = (Z_{1,i}, Z_{2,i})$ iid and regularly varying, EDM$(Z_{1,1}, Z_{2,1})$ can be estimated by a quantity which, after proper centering and scaling, converges weakly to a standard normal. One generalization of this result would be to allow for dependent $\boldsymbol{Z}_i$, in particular letting $\boldsymbol{Z}_i = (X_i, X_{i+h})$

for some lag h and a regularly varying sequence (X_i). Davis and Mikosch [5] resolve this case for the extremogram under α-mixing. We will go in a different direction. We keep the iid assumption, but remove the need of knowing the normalization $b(\cdot)$, which is replaced by an order statistic. We remark that the central limit results in [14] and [5] both assume that $b(\cdot)$ is known.

The setup for this section is as follows. Let $\boldsymbol{Z}_i = (Z_{1,i}, Z_{2,i})$, $i = 1, 2, \ldots$, be an iid sequence of $\mathbb{R}_+^2$-valued regularly varying vectors with $\alpha > 0$ being the index of regular variation. We will often write $\gamma = 1/\alpha$. Let S be the spectral measure of $\boldsymbol{Z}_i$ with respect to some fixed norm $\|\cdot\|$. Define

$$R_i = \|\boldsymbol{Z}_i\|, \qquad \boldsymbol{\Theta}_i = (\Theta_{1,i}, \Theta_{2,i}) = \frac{\boldsymbol{Z}_i}{R_i},$$

and let F be the distribution function of R_i. We want to estimate $\mathrm{EDM}(Z_{1,1}, Z_{2,1}) = \int_{\aleph_+} a_1 a_2 S(d\boldsymbol{a})$ using the estimator $\hat{\rho}_n$ given by

$$\hat{\rho}_n = \frac{1}{k} \sum_{i=1}^{n} h(\boldsymbol{\Theta}_i) \mathbf{1}_{[R_i \geq R_{(k)}]},$$

where $h(\boldsymbol{\Theta}_i) = \Theta_{1,i}\Theta_{2,i}$. As usual, $k = k(n)$ is such that $\lim_{n\to\infty} k/n = 0$, and $R_{(k)}$ denotes the k:th upper order statistic in the sample of size n (we suppress the dependence on the sample size in our notation.) Choosing the number k in practice can be notoriously difficult. Note that $\hat{\rho}_n$ is precisely the estimator given in (3.6) with $x = R_{(k)}$, since $\sum_{i=1}^n \mathbf{1}_{R_i \geq R_{(k)}} = k$. For convenience we also introduce $\tilde{\boldsymbol{\Theta}} \sim S$, so that $\mathrm{EDM}(Z_{1,1}, Z_{2,1}) = \mathrm{E}h(\tilde{\boldsymbol{\Theta}})$.

The aim of this section is to prove asymptotic normality of $\hat{\rho}_n$ under appropriate conditions. Specifically, we will prove

Theorem 1. *Suppose that*

$$\lim_{n\to\infty} \sqrt{k}\left(\frac{n}{k}\mathrm{E}(h(\boldsymbol{\Theta}_1)\mathbf{1}_{[R_1/b(\frac{n}{k}) \geq t^{-\gamma}]}) - \mathrm{E}h(\tilde{\boldsymbol{\Theta}})\frac{n}{k}\bar{F}(b(n/k)t^{-\gamma})\right) = 0 \tag{6.1}$$

holds locally uniformly for $t \in [0, \infty)$, *and assume that* $\sigma^2 = Var(h(\tilde{\boldsymbol{\Theta}})) > 0$. *Then*

$$\sqrt{k}\left(\hat{\rho}_n - \mathrm{E}h(\tilde{\boldsymbol{\Theta}})\right) \Rightarrow N(0, \sigma^2).$$

Remark. Note that the assumption $\sigma^2 > 0$ excludes the asymptotically independent case. If $\sigma = 0$, the statement is still true, but tells us only that the factor $\sqrt{k}$ is too small to yield a non-degenerate limit.

Remark. No second-order regular variation condition is made; instead we use assumption (6.1). Intuitively, this condition guarantees that the dependence between $\boldsymbol{\Theta}_i$ and R_i decays sufficiently fast with n, as R_i is conditioned to lie above $b(n/k)$. In the limit they are of course independent. This can be viewed as a rate condition on $k = k(n)$. Note that if R_1 and $\boldsymbol{\Theta}_1$ are already independent, the left side of (6.1) is zero for all n. The condition is thus automatically satisfied in this case.

A sufficient condition for (6.1) is that

$$\sqrt{k}\left[\frac{n}{k}\mathrm{P}\left(\left(\frac{R_1}{b(n/k)}, \boldsymbol{\Theta}_1\right) \in \cdot\right) - \frac{n}{k}\mathrm{P}\left(\frac{R_1}{b(n/k)} \in \cdot\right) \times S\right] \xrightarrow{\mathrm{v}} 0 \qquad (n \to \infty)$$

in $M_+((0,\infty] \times \aleph_+)$. This condition only involves the distribution of $(R_1, \boldsymbol{\Theta}_1)$, and not the particular function h.

Proof. Consider the process

$$W_n(t) = \frac{1}{\sigma\sqrt{k}} \sum_{i=1}^{n} \left(h(\boldsymbol{\Theta}_i) - \mathrm{E}h(\tilde{\boldsymbol{\Theta}})\right) \mathbf{1}_{[R_i/b(\frac{n}{k}) \geq t^{-\gamma}]}.$$

The main step is to prove that $W_n \Rightarrow W$ in $D[0,\infty)$, where W is Brownian motion. Suppose this is done. It is well-known that

$$\frac{R_{(k)}}{b(n/k)} \xrightarrow{\mathrm{P}} 1,$$

see for instance [15, p. 81]. Since the limit is a constant, W_n and $R_{(k)}/b(n/k)$ converge jointly, and we may thus apply the composition map $D[0,\infty)\times\mathbb{R}\mapsto\mathbb{R}$, $(f,c)\mapsto f(c)$, to deduce

$$W_n\left(\left(\frac{R_{(k)}}{b(n/k)}\right)^{-\alpha}\right) \Rightarrow W(1).$$

The result then follows upon noting that

$$\sigma W_n\left(\left(\frac{R_{(k)}}{b(n/k)}\right)^{-\alpha}\right) = \sqrt{k}\left(\hat\rho_n - \mathrm{E}h(\tilde{\boldsymbol{\Theta}})\right).$$

The proof that $W_n \Rightarrow W$ in $D[0,\infty)$ uses a classical technique based on finite-dimensional marginal convergence and tightness.

We start with the finite-dimensional distributions. For a given interval $(s,t]\subset[0,\infty)$, let $i(j,n)$ be the j:th index i for which $R_i/b(n/k)\in[t^{-\gamma},s^{-\gamma})$. Define

$$N_n = \sum_{i=1}^{n} \boldsymbol{\epsilon}_{R_i/b(\frac{n}{k})}[t^{-\gamma},s^{-\gamma}),$$

and write

$$\begin{aligned} W_n(t)-W_n(s) &= \frac{1}{\sigma\sqrt{k}}\sum_{j=1}^{N_n}\left(h(\boldsymbol{\Theta}_{i(j,n)})-\mathrm{E}h(\tilde{\boldsymbol{\Theta}})\right) \\ &= \frac{1}{\sigma\sqrt{k}}\sum_{j=1}^{N_n}\left(h(\boldsymbol{\Theta}_{i(j,n)})-\mathrm{E}h(\boldsymbol{\Theta}_{i(1,n)})\right) + \frac{1}{\sigma\sqrt{k}}N_n\left(\mathrm{E}h(\boldsymbol{\Theta}_{i(1,n)})-\mathrm{E}h(\tilde{\boldsymbol{\Theta}})\right) \\ &= C_n + D_n. \end{aligned}$$

First consider D_n. We have

$$\begin{aligned} \sigma D_n =& \frac{N_n}{k}\sqrt{k}\left(\mathrm{E}(h(\boldsymbol{\Theta}_1)\mid R_1/b(n/k)\in[t^{-\gamma},s^{-\gamma})) - \mathrm{E}h(\tilde{\boldsymbol{\Theta}})\right) \\ =& \frac{N_n/k}{\frac{n}{k}F(b(\frac{n}{k})[t^{-\gamma},s^{-\gamma}))}\sqrt{k}\left(\mathrm{E}(h(\boldsymbol{\Theta}_1)\mathbf{1}_{[t^{-\gamma}\le R_1/b(\frac{n}{k})<s^{-\gamma}]}) - \mathrm{E}h(\tilde{\boldsymbol{\Theta}})\frac{n}{k}F(b(n/k)[t^{-\gamma},s^{-\gamma}))\right) \\ =& \frac{N_n/k}{\frac{n}{k}F(b(\frac{n}{k})[t^{-\gamma},s^{-\gamma}))}(B_n(t)-B_n(s)). \end{aligned}$$

It follows from [15, Theorem 6.2 (9)] that $N_n/k \xrightarrow{\mathrm{P}} \nu_\alpha[t^{-\gamma},s^{-\gamma}) = t-s$, and regular variation of $1-F$ similarly implies that $\frac{n}{k}F(b(\frac{n}{k})[t^{-\gamma},s^{-\gamma}))\to t-s$. Assumption (6.1) says that $B_n\to 0$ locally uniformly, so $D_n \xrightarrow{\mathrm{P}} 0$.

To deal with C_n, let $\sigma_n^2 = \mathrm{Var}(h(\boldsymbol{\Theta}_{i(1,n)})) = \mathrm{Var}(h(\boldsymbol{\Theta}_1)\mid R_1/b(n/k)\in[t^{-\gamma},s^{-\gamma}))$ and define the process Y_n by

$$Y_n(r) = \frac{1}{\sigma_n\sqrt{k}}\sum_{j=1}^{kr}\left(h(\boldsymbol{\Theta}_{i(j,n)})-\mathrm{E}h(\boldsymbol{\Theta}_{i(j,n)})\right). \tag{6.2}$$

By the *Découpage de Lévy*, the sequence $\{\boldsymbol{\Theta}_{i(j,n)}\}$ is iid, so the functional central limit theorem for triangular arrays implies that $Y_n \Rightarrow Y$ in $D[0,\infty)$, where Y is Brownian motion. (See the proof of Theorem 3 in [14] for more details on this technique.) As before, $N_n/k \xrightarrow{\mathrm{P}} t-s$, a deterministic limit, so we have joint convergence and may apply composition to obtain

$$Y_n(N_n/k) = \frac{1}{\sigma_n\sqrt{k}}\sum_{j=1}^{N_n}\big(h(\boldsymbol{\Theta}_{i(j,n)}) - \mathrm{E}h(\boldsymbol{\Theta}_{i(j,n)})\big) \Rightarrow Y(t-s).$$

The left-hand side equals $\frac{\sigma}{\sigma_n}C_n$. Moreover, regular variation implies that $\sigma_n \to \sigma$, and since $\sigma > 0$, we obtain $C_n \Rightarrow Y(t-s) \sim N(0, t-s)$.

Consider now an arbitrary number of disjoint intervals $(s_m, t_m]$, $m = 1,\dots,M$. Similarly as before, we define $i_m(j,n)$ to be the j:th index i for which $R_i/b(n/k) \in [t_m^{-\gamma}, s_m^{-\gamma})$, and we set

$$N_n^m = \sum_{i=1}^n \boldsymbol{\epsilon}_{R_i/b(\frac{n}{k})}[t_m^{-\gamma}, s_m^{-\gamma}).$$

In the same manner as above, and with obvious notation, we decompose the M increments as

$$W_n(t_m) - W_n(s_m) = C_n^m + D_n^m.$$

We again obtain $D_n^m \xrightarrow{\mathrm{P}} 0$ for each m. Next, for each m, define processes Y_n^m as in (6.2), but with $i_m(j,n)$ instead of $i(j,n)$. The Découpage de Lévy implies that the M sequences $\{\boldsymbol{\Theta}_{i_m(j,n)} : j = 1,2,\dots\}$ are independent for fixed n, and hence the processes Y_n^m are also independent. The previously established convergence result, which was proven for one single sequence of processes $\{Y_n\}$, thus holds jointly:

$$(Y_n^1,\dots,Y_n^M) \Rightarrow (Y^1,\dots,Y^M),$$

where the limit is M-dimensional Brownian motion. Composition with $N_n^m/k \xrightarrow{\mathrm{P}} t_m - s_m$ for $m = 1,\dots,M$, lets us conclude that

$$(C_n^1,\dots,C_n^M) \Rightarrow N(0, \mathrm{diag}(t_1 - s_1,\dots,t_M - s_M)).$$

This proves finite-dimensional convergence.

Now, since the limit process W has continuous paths, Theorem 13.5 in Billingsley [2] yields the result as soon as we verify

$$\mathrm{E}(|W_n(t) - W_n(s)|^2|W_n(s) - W_n(r)|^2) \le (t-r)^2$$

for all $0 \le r \le s \le t$ and all n. However, a careful look at the arguments leading up to this result shows that it suffices to prove

$$\limsup_n \mathrm{E}(|W_n(t) - W_n(s)|^2|W_n(s) - W_n(r)|^2) \le (t-r)^2 \tag{6.3}$$

for all $0 \le r \le s \le t$. This fact was also used in [16]. So fix r, s, t and write

$$\begin{aligned}\alpha_i &= \Big(h(\boldsymbol{\Theta}_i) - \mathrm{E}(h(\tilde{\boldsymbol{\Theta}}))\Big)\mathbf{1}_{[R_i/b(\frac{n}{k})\in[t^{-\gamma}, s^{-\gamma})]}\\ \beta_i &= \Big(h(\boldsymbol{\Theta}_i) - \mathrm{E}(h(\tilde{\boldsymbol{\Theta}}))\Big)\mathbf{1}_{[R_i/b(\frac{n}{k})\in[s^{-\gamma}, r^{-\gamma})]}.\end{aligned}$$

Using that $\alpha_i\beta_i = 0$, it is straightforward to check that

$$\sigma^4\mathrm{E}(|W_n(t) - W_n(s)|^2|W_n(s) - W_n(r)|^2) = \frac{1}{k^2}\mathrm{E}\left(\sum_i \alpha_i\right)^2\left(\sum_i \beta_i\right)^2$$

$$\begin{aligned}&= \frac{n(n-1)}{k^2}\mathrm{E}(\alpha_1^2\beta_2^2) + \frac{n(n-1)(n-2)}{k^2}\mathrm{E}(\alpha_1^2\beta_2\beta_3)\\ &\quad + \frac{n(n-1)(n-2)}{k^2}\mathrm{E}(\alpha_1\alpha_2\beta_3^2) + \frac{n(n-1)(n-2)(n-3)}{k^2}\mathrm{E}(\alpha_1\alpha_2\beta_3\beta_4).\end{aligned}$$

All expectations factorize by independence, and using simple bounds for the coefficients we get

$$\begin{aligned}&\sigma^4\mathrm{E}(|W_n(t)-W_n(s)|^2|W_n(s)-W_n(r)|^2)\\ &\qquad\le \frac{n^2}{k^2}\mathrm{E}(\alpha_1^2)\mathrm{E}(\beta_1^2) + \frac{n^3}{k^2}\mathrm{E}(\alpha_1^2)(\mathrm{E}(\beta_1))^2\\ &\qquad\qquad + \frac{n^3}{k^2}(\mathrm{E}(\alpha_1))^2\mathrm{E}(\beta_1^2) + \frac{n^4}{k^2}(\mathrm{E}(\alpha_1))^2(\mathrm{E}(\beta_1))^2.\end{aligned}$$

Note that the function $g : (0,\infty] \times \aleph_+ \to \mathbb{R}$ given by

$$g(r,\boldsymbol{\theta}) = \Big(h(\boldsymbol{\theta}) - \mathrm{E}h(\tilde{\boldsymbol{\Theta}})\Big)^2 \mathbf{1}_{[r\in[t^{-\gamma},s^{-\gamma})]}$$

is $\nu_\alpha \times S$-a.e. continuous and compactly supported in $(0,\infty]\times\aleph_+$. (As before, $\nu_\alpha \in M_+((0,\infty])$ is given by $\nu_\alpha(x,\infty] = x^{-\alpha}$.) By regular variation, and since $\mathrm{E}(\alpha_1^2) = \mathrm{E}g(R_1/b(n/k),\boldsymbol{\Theta}_1)$, we get

$$\frac{n}{k}\mathrm{E}(\alpha_1^2) \to \mathrm{E}\left(h(\tilde{\boldsymbol{\Theta}}) - \mathrm{E}h(\tilde{\boldsymbol{\Theta}})\right)^2 \nu_\alpha[t^{-\gamma},s^{-\gamma}) = \sigma^2(t-s).$$

Similarly, $\frac{n}{k}\mathrm{E}(\beta_1^2) \to \sigma^2(s-r)$. Moreover,

$$\begin{aligned}\sqrt{k}\frac{n}{k}\mathrm{E}(\alpha_1) &= \sqrt{k}\left(\mathrm{E}(h(\boldsymbol{\Theta}_1)\mathbf{1}_{[t^{-\gamma}\le R_1/b(\frac{n}{k})<s^{-\gamma}]}) - \mathrm{E}h(\tilde{\boldsymbol{\Theta}})\frac{n}{k}F(b(n/k)[t^{-\gamma},s^{-\gamma}))\right)\\ &\to 0 \qquad (n\to\infty)\end{aligned}$$

by assumption (6.1). Thus $\frac{n^2}{k}(\mathrm{E}(\alpha_1))^2 = (\sqrt{k}\frac{n}{k}\mathrm{E}(\alpha_1))^2 \to 0$, and similarly we also get $\frac{n^2}{k}(\mathrm{E}(\beta_1))^2 \to 0$. Combining these results yields

$$\begin{aligned}\sigma^4\limsup_n \mathrm{E}(|W_n(t)-W_n(s)|^2|W_n(s)-W_n(r)|^2) &\le \sigma^2(t-s)\sigma^2(s-r)\\ &\le \sigma^4(t-r)^2.\end{aligned}$$

We conclude that (6.3) holds. □

References

[1] J. Beran. *Statistics for Long-Memory Processes.* Chapman & Hall, New York, 1994.

[2] P. Billingsley. *Convergence of Probability Measures.* Wiley Series in Probability and Statistics. Wiley, second edition, 1999.

[3] B. D'Auria and S. I. Resnick. The influence of dependence on data network models. *Adv. Appl. Prob.*, 40:60–94, 2008.

[4] R. Davis and S. I. Resnick. Limit theory for moving averages of random variables with regularly varying tail probabilities. *Ann. Probability*, 13:179–195, 1985.

[5] R. A. Davis and T. Mikosch. The extremogram: A correlogram for extreme events. *Preprint*, 2008.

[6] R. A. Davis and S. I. Resnick. Basic properties and prediction of max-arma processes. *Advances in Applied Probability*, 21(4):781–803, December 1989.

[7] V. Fasen, C. Klüppelberg, and M. Schlather. High-level depence in time series models. *Preprint*, 2009.

[8] F. Hernandez-Campos, K. Jeffay, C. Park, J. S. Marron, and S. I. Resnick. Extremal dependence: Internet traffic applications. *Stochastic Models*, 21(1):1–35, 2005.

[9] T. Hsing. Extreme value theory for suprema of random variables with regularly varying tail probabilities. *Stochastic Processes and their Applications*, 22:51–57, 1986.

[10] H. Hult and G. Samorodnitsky. Tail probabilities for infinite series of regularly varying random vectors. *Bernoulli*, 14:838–864, 2008.

[11] M. Leadbetter. Extremes and local dependence in stationary sequences. *Z. Wahrscheinligkeitstheorie verw. Gebiete*, 65(2):291–306, June 1983.

[12] A. Ledford and J. Tawn. Statistics for near independence in multivariate extreme values. *Biometrika*, 83(169-187), 1996.

[13] A. Ledford and J. Tawn. Diagnostics for dependence within time series extremes. *J. Roy. Statist. Soc. Ser. B*, 65(521-543), 2003.

[14] S. I. Resnick. The extremal dependence measure and asymptotic independence. *Stochastic Models*, 20(2):205–227, 2004.

[15] S. I. Resnick. *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer, 2007.

[16] S. I. Resnick and C. Stărică. Smoothing the hill estimator. *Advances in Applied Probability*, 29(1):271–293, March 1997.

[17] G. Samorodnitsky. Long range dependence. *Foundations and Trends in Stochastic Systems*, 1(3):163–257, 2006.

[18] M. Schlather and J. Tawn. A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika*, 90(1):139–156, 2003.

[19] Z. Zhang and R. L. Smith. The behavior of multivariate maxima of moving maxima processes. *Journal of Applied Probability*, 41(4):1113–1123, December 2004.

Martin Larsson, School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853

E-mail address: `mol23@cornell.edu`

Sidney I. Resnick, School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853

E-mail address: `sir1@cornell.edu`