Extremes (2012) 15:231–256
DOI 10.1007/s10687-011-0135-9

Extremal dependence measure and extremogram: the regularly varying case

Martin Larsson · Sidney I. Resnick

Received: 30 March 2010 / Revised: 16 May 2011 /
Accepted: 24 May 2011 / Published online: 9 June 2011

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 or vector context which summarize extremal dependence in a function comparable to a 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.

Keywords Regular variation · Heavy tails · Asymptotic independence · Extremal dependence · EDM · Extremogram

AMS 2000 Subject Classifications 60G70 · 60F05 · 60G10

1 Introduction

How does one sensibly quantify serial tail dependence between large values in a heavy tailed stationary stochastic process? For Gaussian processes, the traditional correlation function is a useful, simple numerical summary of dependence but outside

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

M. Larsson (✉) · S. I. Resnick
School of Operations Research and Information Engineering, Cornell University,
Ithaca, NY 14853, USA
e-mail: mol23@cornell.edu

S. I. Resnick
email: sir1@cornell.edu

the Gaussian context, one should have modest expectations of usefulness for comparably simple numerical summaries of tail dependence. Two proposals for numerical summaries are the *extremal dependence measure (EDM)* (Resnick 2004) and the *extremogram* (Davis and Mikosch 2009). The goal of this paper is to compare the two concepts.

The extremogram is related to the *extreme dependence functions* and *extremal coefficient function* introduced by Fasen et al. (2010); see also Schlather and Tawn (2003). These concepts are generalizations in a stochastic processes context of the *coefficient of tail dependence* introduced by Ledford and Tawn (1996, 2003). The EDM was proposed as a statistical tool to investigate asymptotic independence, for which it was used in Hernandez-Campos et al. (2005) in a study involving Internet data. In D'Auria and Resnick (2008) it was used in a stochastic processes context to study long-range extremal dependence in data network models whose infinite variance precluded use of the correlation function.

A related way of describing extremal dependence in regularly varying stationary time series is via the tail and spectral processes; see Basrak and Segers (2009). An early, significant contribution to the understanding of extremal dependence is the work by Leadbetter (1983), who studied the *extremal index* with the goal of describing the clustering of extremes. In the stable process literature, the concept of *co-dimension* or *covariation* can be defined to provide numerical summaries of dependence (Samorodnitsky and Taqqu 1994). More discussion and references is in Heffernan (2000).

In this paper we investigate the decay of dependence between two observations in a regularly varying strictly stationary time series, as the time lag between the observations grows large. We review the extremogram and the EDM, and by means of some simple examples illustrate how they may exhibit very different asymptotic behavior, thereby highlighting 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 Samorodnitsky (2006) for a survey.

In contrast to the extremogram, the statistical properties of the EDM have not been as well studied. As a first step toward addressing this issue, we prove a central limit theorem for an estimator of the EDM in the simplified setting of iid regularly varying bivariate pairs. An interesting observation here is that the population version of the EDM (as opposed to some pre-asymptotic quantity) can be used to center the estimator, even in the absence of second order regular variation assumptions.

This paper is structured as follows. In Section 2 we briefly review some basic notions associated with multivariate regular variation. To facilitate the subsequent analysis, we introduce a concept of *equivalence* of dependence measures; for now, think of a dependence measure as a real-valued function acting on the law of a regularly varying vector. In Section 3 we review the EDM and the extremogram and derive some useful properties. Section 4 compares the rate of decay of the EDM and the extremogram in some examples. Concluding remarks, summaries and highlights of some contrasts as well as implications for the study of long-range tail dependence are discussed in Section 5. In the last section given as an Appendix we provide the central limit theorem mentioned previously.

2 Regular variation and measures of tail dependence

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 $\mathbb{E} = [\mathbf{0}, \infty]\backslash\{\mathbf{0}\}$ such that

$$tP\left(\frac{\boldsymbol{Z}}{b(t)} \in \cdot\right) \xrightarrow{v} \nu \qquad (t \to \infty). \tag{1}$$

Here $\xrightarrow{v}$ denotes vague convergence in $M_+(\mathbb{E})$, the space of Radon measures on $\mathbb{E}$. Restricting to the nonnegative orthant simplifies the exposition, but is not essential for our results. We also assume one-dimensional marginal distributions of ν are non-degenerate which means

$$\nu([0, \infty]^{i-1} \times (x, \infty] \times [0, \infty]^{d-i} = c_i x^{-\alpha}, \quad c_i > 0,\ x > 0,\ i = 1, \dots, d.$$

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 $\boldsymbol{Z}$ (see, for instance, Resnick 2007, Theorem 6.1). A sequence $(\boldsymbol{Z}_n)_{n\in\mathbb{N}}$ of d-dimensional random vectors is called regularly varying if for every m and $n_1, \dots, n_m \in \mathbb{N}$, the dm-dimensional vector $\mathbf{vec}(\boldsymbol{Z}_{n_1}, \dots, \boldsymbol{Z}_{n_m})$ obtained by stacking the $\boldsymbol{Z}_{n_i}$ is regularly varying.

Now fix a norm $\|\cdot\|$ on $\mathbb{R}_+^d$ 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 (1) is, by Theorem 6.1 in Resnick (2007), the existence of a probability measure S on $\aleph_+ = \{\boldsymbol{x} \in \mathbb{E} : \|\boldsymbol{x}\| = 1\}$ and a constant $c > 0$ such that on $M_+((0, \infty] \times \aleph_+)$,

$$tP\left(\left(\frac{R}{b(t)}, \boldsymbol{\Theta}\right) \in \cdot\right) \xrightarrow{v} c\nu_\alpha \times S \qquad (t \to \infty), \tag{2}$$

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 by multiplicative constants. The constant c in Eq. 2 thus depends on which particular normalization one chooses. Since $c\nu_\alpha \times S = \nu \circ T^{-1}$, 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 $b(t)$ in Eq. 1 is the same for all components of $\boldsymbol{Z}$. This can be relaxed (Resnick 2007, Theorem 6.5), but is nonetheless appropriate in our setting.

Measures of extremal dependence proposed in the literature are often of the form $\nu(f) = \int_{\mathbb{E}} f(x)\nu(dx)$ for some fixed bounded, compactly supported $f : \mathbb{E} \mapsto \mathbb{R}_+$ which is continuous outside a ν-nullset; recall that the relatively compact sets in $\mathbb{E}$ are those bounded away from the origin. Usually $f = \mathbf{1}_C$ for some relatively compact measurable $C \subset \mathbb{E}$ with $\nu(\partial C) = 0$. This is the case for the extremogram.

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

$$\begin{aligned}\nu(f) &= \int_{\mathbb{E}} f(s)\nu(dx)\\ &= \int_{T(\mathbb{E})} f\circ T^{-1}(r,\boldsymbol{a})c\nu_\alpha(dr)S(d\boldsymbol{a}) = \int_{\aleph_+} k(\boldsymbol{a};\alpha)S(d\boldsymbol{a}),\end{aligned} \tag{3}$$

where

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

For extremal dependence measures with form (3), the following concept is a useful equivalence condition.

Definition 1 Fix a norm $\|\cdot\|$ on $\mathbb{R}_+^d$ for $d \geq 2$. Suppose ρ_i, $i = 1, 2$, are two $\mathbb{R}_+$-valued maps defined on the set of regularly varying random vectors $\boldsymbol{Z} = (Z_1, \ldots, Z_d)$, given by

$$\rho_i(\boldsymbol{Z}) = \int_{\aleph_+} k_i(\boldsymbol{a};\alpha)S(d\boldsymbol{a}) \tag{4}$$

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

$$m\rho_1(\boldsymbol{Z}) \leq \rho_2(\boldsymbol{Z}) \leq M\rho_1(\boldsymbol{Z}) \tag{5}$$

for every d-dimensional random vector $\boldsymbol{Z}$ that is regularly varying with index α.

Remarks This notion of equivalence is indeed an equivalence relation. Notice that the ρ_i in fact only depend on the law of $\boldsymbol{Z}$. To see why equivalence is relevant when considering the limiting decay of dependence measures, notice that the definition implies that

$$m \leq \liminf_{n\to\infty} \frac{\rho_2(\boldsymbol{Z}_n)}{\rho_1(\boldsymbol{Z}_n)} \leq \limsup_{n\to\infty} \frac{\rho_2(\boldsymbol{Z}_n)}{\rho_1(\boldsymbol{Z}_n)} \leq M \tag{6}$$

for any sequence $(\boldsymbol{Z}_n)$ such that each $\boldsymbol{Z}_n$ is a d-dimensional regularly varying random vector with index α and $\rho_1(\boldsymbol{Z}_n) > 0$. We also write Eq. 6 as $\rho_1(\boldsymbol{Z}_n) \asymp \rho_2(\boldsymbol{Z}_n)$ as $n \to \infty$.

If $(X_n)_{n\in\mathbb{N}}$ is a regularly varying strictly stationary sequence, 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. This is useful in connection with long-range dependence, since the lack of summability of quantities measuring serial dependence is often used as a definition of long-range dependence.

There is an easy way of verifying equivalence, that will be used several times in the sequel.

Proposition 1 *Suppose* ρ_i, $i = 1, 2$, *are as in Eq. 4. Then* ρ_1 *and* ρ_2 *are equivalent if and only if for each* $\alpha > 0$ *there are constants* $0 < m \leq M < \infty$ *such that*

$$mk_1(\boldsymbol{a}; \alpha) \leq k_2(\boldsymbol{a}; \alpha) \leq Mk_1(\boldsymbol{a}; \alpha) \quad \textit{for all } \boldsymbol{a} \in \aleph_+. \tag{7}$$

Proof The necessity is clear; simply integrate Eq. 7 with respect to $S(d\boldsymbol{a})$ to get Eq. 5. For sufficiency, suppose Eq. 7 does not hold. Then for some $\alpha > 0$ there are $\boldsymbol{a}_n \in \aleph_+$, $n \geq 0$, such that either

$$\lim_{n\to\infty} k_2(\boldsymbol{a}_n; \alpha) = 0 \qquad \text{and} \qquad \limsup_{n\to\infty} k_1(\boldsymbol{a}_n; \alpha) > 0,$$

or

$$\lim_{n\to\infty} k_2(\boldsymbol{a}_n; \alpha) = \infty \qquad \text{and} \qquad \liminf_{n\to\infty} k_1(\boldsymbol{a}_n; \alpha) < \infty,$$

or one of the above with k_1 and k_2 interchanged. 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 Eq. 5 is violated. □

Suppose now that in addition to the norm $\|\cdot\|$ we consider a different norm $\|\cdot\|'$. The next result shows how to express a dependence measure $\rho(\boldsymbol{Z}) = \int_{\aleph_+} k(\boldsymbol{a}; \alpha) S(d\boldsymbol{a})$ in terms of the spectral measure S' associated with the norm $\|\cdot\|'$. Let $\aleph'_+ = \{\boldsymbol{x} \in \mathbb{E} : \|\boldsymbol{x}\|' = 1\}$ and define

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

Then g 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 g is a bijection with inverse

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

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

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

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

$$\int_{\aleph_{+}} f(\boldsymbol{a}) S(d\boldsymbol{a}) = \frac{1}{c} \int_{\mathbb{E}} f(\boldsymbol{x}/\|\boldsymbol{x}\|) \mathbf{1}_{(1,\infty}(\|\boldsymbol{x}\|) \nu(d\boldsymbol{x}).$$

The identity $\int_{\mathbb{E}} h(\boldsymbol{x}) \nu(d\boldsymbol{x}) = c' \int_0^{\infty} \int_{\aleph'_{+}} h(r\boldsymbol{a}') \alpha r^{-\alpha-1} dr S'(d\boldsymbol{a}')$ applied to the function $h(\boldsymbol{x}) = c^{-1} f(\boldsymbol{x}/\|\boldsymbol{x}\|) \mathbf{1}_{(1,\infty}(\|\boldsymbol{x}\|)$ then yields

$$\begin{aligned}
\int_{\aleph_{+}} f(\boldsymbol{a}) S(d\boldsymbol{a}) &= \frac{c'}{c} \int_0^{\infty} \int_{\aleph'_{+}} f\left(\frac{\boldsymbol{a}'}{\|\boldsymbol{a}'\|}\right) \mathbf{1}_{(1,\infty)}(r\|\boldsymbol{a}'\|) \alpha r^{-\alpha-1} dr S'(d\boldsymbol{a}') \\
&= \int_{\aleph'_{+}} \frac{c'}{c} f \circ g^{-1}(\boldsymbol{a}') \int_0^{\infty} \mathbf{1}_{(1,\infty)}(r\|\boldsymbol{a}'\|) \alpha r^{-\alpha-1} dr S'(d\boldsymbol{a}') \\
&= \int_{\aleph'_{+}} \frac{c'}{c} f \circ g^{-1}(\boldsymbol{a}') \|\boldsymbol{a}'\|^{\alpha} S'(d\boldsymbol{a}').
\end{aligned}$$

This suffices. □

Remarks Equation 8.38 in Beirlant et al. (2004) gives a similar conversion formula between spectral measures based on different norms.

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

$$\text{EDM}(Z_1, Z_2) = \int_{\aleph_{+}} a_1 a_2 S(d\boldsymbol{a}). \tag{8}$$

Note $\text{EDM}(Z_1, Z_2) = 0$ if and only if S concentrates on

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

that is, iff the limit measure ν concentrates on the coordinate axes in $\mathbb{E}$. This happens precisely when (Z_1, Z_2) possesses asymptotic independence (Resnick 2007, Section 6.5.1). Moreover, if the norm is symmetric, EDM 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.

The EDM depends on the choice of norm. However, as the following result shows, under a change of norm the EDM still belongs to the same equivalence class in the sense of Definition 1.

Proposition 3 *Let $\|\cdot\|$ and $\|\cdot\|'$ be two norms on $\mathbb{R}^2_+$. Then*

$$EDM(Z_1, Z_2) = \int_{\aleph_+} a_1 a_2 S(d\boldsymbol{a}) \qquad and \qquad EDM'(Z_1, Z_2) = \int_{\aleph'_+} a'_1 a'_2 S'(d\boldsymbol{a}')$$

are equivalent.

Proof By Proposition 2, $\mathrm{EDM}'(Z_1, Z_2) = \int_{\aleph_+} \frac{c}{c'} a_1 a_2 (\|\boldsymbol{a}\|')^{\alpha-2} S(d\boldsymbol{a})$. Since all norms on finite-dimensional space are equivalent, there are constants $0 < K_1 \leq K_2 < \infty$ such that $K_1 \leq \|\boldsymbol{a}\|' \leq K_2$ for all $\boldsymbol{a} \in \aleph_+$, where we used that $\|\boldsymbol{a}\| = 1$. It follows that

$$m a_1 a_2 \leq \frac{c}{c'} a_1 a_2 (\|\boldsymbol{a}\|')^{\alpha-2} \leq M a_1 a_2 \qquad \text{for all } \boldsymbol{a} \in \aleph_+,$$

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 stationary time series, the EDM is used as a tool to study serial tail dependence by examining how $\mathrm{EDM}(h) := \mathrm{EDM}(X_n, X_{n+h})$ behaves as a function of lag h. In Hernandez-Campos et al. (2005), an empirical estimator of EDM is used to study extremal dependence in Internet data and D'Auria and Resnick (2008) use $\mathrm{EDM}(h)$ to measure dependence decay for infinite variance data network models and conclude presence of a kind of long range dependence.

When introduced in Resnick (2004), the EDM was originally defined as

$$\begin{aligned} \mathrm{EDM}(Z_1, Z_2) &= 1 - (4/\pi)^2 \int_0^{\pi/2} \left(\theta - \frac{\pi}{4}\right)^2 \bar{S}(d\theta) \\ &= (4/\pi)^2 \int_0^{\pi/2} \theta(\pi/2 - \theta) \bar{S}(d\theta), \end{aligned} \tag{9}$$

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 and we prefer definition (8). The two expressions are equivalent in the sense of Definition 1. Indeed, by Proposition 3, the EDM under any norm is equivalent to the EDM under the Euclidean norm under which Eq. 9 becomes

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

It is easy to check that

$$k_1(\boldsymbol{a}; \alpha) = a_1 a_2 \quad \text{and} \quad k_2(\boldsymbol{a}; \alpha) = (4/\pi)^2 \arctan\left(\frac{a_2}{a_1}\right)\left(\frac{\pi}{2} - \arctan\left(\frac{a_2}{a_1}\right)\right)$$

satisfy the conditions of Proposition 1, so that the dependence measures in Eqs. 8 and 9 are equivalent. This fact, together with the behavior in case of asymptotic independence and asymptotic full dependence, guarantee that the results given in Resnick (2004) and D'Auria and Resnick (2008) remain true for the EDM as defined in Eq. 8.

It is shown by Davis and Mikosch (2009, p. 980), that the extremogram, whose definition we will recall shortly, has an interesting interpretation as the limit of a sequence of covariance functions. Not surprisingly the EDM has a similar interpretation.

Proposition 4 *Let* (Z_1, Z_2) *be regularly varying as in Eq. 1. Then*

$$EDM(Z_1, Z_2) = \lim_{x \to \infty} \mathrm{E}\left(\frac{Z_1}{R}\frac{Z_2}{R} \mid R > x\right), \tag{10}$$

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

This lets $\mathrm{EDM}(Z_1, Z_2)$ 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 $T(Z_1, Z_2) = (R, \boldsymbol{\Theta})$, where $R = \|(Z_1, Z_2)\|$ and $\boldsymbol{\Theta} = (Z_1/R, Z_2/R)$. Changing variables to $x = b(t)$, the regular variation condition (2) becomes

$$b^{\leftarrow}(x)\mathrm{P}\left(\left(\frac{R}{x}, \boldsymbol{\Theta}\right) \in \cdot\right) \xrightarrow{v} c\nu_\alpha \times S \qquad (x \to \infty). \tag{11}$$

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

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

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 $f : \aleph_+ \to \mathbb{R}$, $(a_1, a_2) \mapsto a_1 a_2$ to get

$$S_x(f) \to S(f) = \int_{\aleph_+} a_1 a_2 S(d\boldsymbol{a}) \qquad (x \to \infty). \tag{12}$$

On the other hand,

$$S_x(f) = \int_{\aleph_+} a_1 a_2 \mathrm{P}(\boldsymbol{\Theta} \in d\boldsymbol{a} \mid R > x) = \mathrm{E}(\Theta_1 \Theta_2 \mid R > x) = \mathrm{E}\left(\frac{Z_1}{R}\frac{Z_2}{R} \mid R > x\right),$$

which gives the result. □

Formula (10) 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 set $R_i = \|\boldsymbol{Z}_i\|$, and $N_n = \sum_{i=1}^n \mathbf{1}_{[R_i > x]}$, the number of exceedances over a high threshold x and estimate $\mathrm{EDM}(Z_1, Z_2)$ by

$$\widehat{\mathrm{EDM}}(Z_1, Z_2) = \frac{1}{N_n} \sum_{i=1}^n \frac{Z_{1i}}{R_i} \frac{Z_{2i}}{R_i} \mathbf{1}_{[R_i > x]}. \tag{13}$$

This is precisely the estimator that results from integrating $f(a_1, a_2) = a_1 a_2$ with respect to the empirical spectral measure,

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

where $(R_i, \boldsymbol{\Theta}_i) = T(\boldsymbol{Z}_i)$ and $b(t) = x$ is the normalization in Eq. 11. Cf. Resnick (2007, Chapter 9.2). We discuss the asymptotic normality of $\widehat{\mathrm{EDM}}$ in the Appendix, extending the discussion in Resnick (2004).

Of course the EDM can be expressed directly in terms of the limit measure ν of (Z_1, Z_2) and a calculation using $\nu \circ T^{-1} = c\nu_\alpha \times S$, where $c = \nu\{\|\boldsymbol{x}\| > 1\}$, shows

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

Compare this to Eq. 10 in Proposition 4.

3.2 Extremogram

The second dependence measure we consider is the *extremogram*, introduced in Davis and Mikosch (2009). We first review this concept for a general state space

$\mathbb{R}^d$, although we later specialize the setup to $\mathbb{R}^1_+$. Given a $\mathbb{R}^d$-valued strictly stationary regularly varying time series $(\boldsymbol{X}_n)_{n\in\mathbb{N}}$, let $\nu_{0,\ldots,h}$ denote the limit measure of $(\boldsymbol{X}_0,\ldots,\boldsymbol{X}_h)$ 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 $\nu_{0,\ldots,h}$, the extremogram $\gamma_{AB}(h)$ is defined as

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

where ν_h is the limit measure of $(\boldsymbol{X}_0, \boldsymbol{X}_h)$. Modulo a multiplicative scalar coming from the normalization $b(\cdot)$, the extremogram, unlike the EDM, does not require choice of a norm. Instead, one must choose particular sets A and B. It should be pointed out that slightly different versions of the extremogram appear in the literature. The one given above is consistent with Davis and Mikosch (2009).

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 Eq. 11, that c_h is given by

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

where $R = \|(\boldsymbol{X}_0, \boldsymbol{X}_h)\|$ is the $2d$-dimensional norm of $(\boldsymbol{X}_0, \boldsymbol{X}_h)\in\mathbb{R}^{2d}$. 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 Davis and Mikosch (2009), 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 Davis and Mikosch (2009) 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 Fasen et al. (2010) 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 $X_{h_1}, \ldots, X_{h_d}$ 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 $(X_{h_1}, \ldots, X_{h_d})$ 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}). The following relationship between $\gamma_{AB}(h)$ and S_h is familiar from extreme value theory; see for instance Resnick (2007, p. 167) and the following is for completeness.

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

$$\gamma_{AB}(h) = \int_{\aleph_+} \left(\frac{a_0}{u}\right)^{\alpha} \wedge \left(\frac{a_h}{w}\right)^{\alpha} S_h(d\boldsymbol{a}). \tag{14}$$

Proof Note that $A \times B = \left\{(x_0, x_h) : \frac{x_0}{u} \wedge \frac{x_h}{w} > 1\right\}$. 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 Eq. 14 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, provided the norm is symmetric.

Since the extremogram is defined in Cartesian coordinates, its definition does not depend on a choice of norm so the analogue of Proposition 3 is clear. The next result indicates that equivalence is preserved by changing the sets A, B.

Proposition 5 *Let $A_i = (u_i, \infty)$, $B_i = (w_i, \infty)$, $i = 1, 2$, and suppose $u_i, w_i > 0$. Then $\gamma_{A_1B_1}$ and $\gamma_{A_2B_2}$ are equivalent.*

Proof For $\boldsymbol{a}$ near $(0, 1)/\|(0, 1)\|$, the ratio between the integrands in the representation (14) of $\gamma_{A_1B_1}$ and $\gamma_{A_2B_2}$ 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 equivalence via Proposition 1. □

Are the EDM and the extremogram always equivalent? No!

Proposition 6 *The EDM and the extremogram (with $A = (u, \infty)$, $B = (w, \infty)$, $u, w > 0$) are 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 Eq. 14 is eventually equal to $u^{-\alpha}a_0^\alpha$. Moreover, the integrand in Eq. 8 is $a_0a_h \sim a_0\|(0, 1)\|$. 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 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 equivalence. □

Remarks If, as an example, $(X_n)_{n\in\mathbb{N}}$ is a regularly varying ARMA model, α not only determines the tails of the marginals X_n, but is also connected to the behavior of the spectral measure associated with (X_n, X_{n+h}). The value of α may therefore influence the rate of decay of EDM(X_n, X_{n+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.

Transformation to the standard case is not an equivalence operation for the EDM; as we will see next, it is for the extremogram. Thus, for the extremogram, asymptotic decay as a function of lag is invariant when the observations of the time series are transformed to the standard case $\alpha = 1$ and the extremogram rate of decay is a function of the copulas.

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 identical.*

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

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

So γ_{AB} and γ_{AB}^* are identical, as claimed. □

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 equivalent.*

Proof Let $\gamma_{AB}^*(h)$ be the extremogram of (X_n^*). By Proposition 6, γ_{AB}^* and EDM* are equivalent. Furthermore, Proposition 7 implies that γ_{AB} and γ_{AB}^* are equivalent. The result follows. □

Finally, we remark that there is no simple expression for how the EDM behaves under standardization of the marginals.

4 Aymptotic decay as a function of lag; some examples

We now compare the asymptotic behavior of the EDM and the extremogram as a function of lag for some examples in order and learn that these dependence measures can decay at substantially different rates. This is important if one wants to define a notion of long-range tail dependence. Indeed, one common way of defining (non-extremal) long-range dependence for stationary sequences is through summability, or lack thereof, of the autocovariance function when it exists. A natural approach in the extremal case would therefore be to do the same using the EDM or the extremogram—this has in fact been suggested in Davis and Mikosch (2009). We also mention that in a recent paper, Kulik and Soulier (2011) introduce the idea of letting a finite threshold and time lag tend to infinity simultaneously. They show, among other things, how this allows one to establish asymptotic normality of Hill estimators, even in the presence of long-range dependence. There are other approaches to long-range dependence in the literature; see Samorodnitsky (2006) for a survey and Beran (1994) for issues related to statistical inference.

4.1 Max-moving averages

Max-moving average processes, or, more generally, Max-ARMA processes, were studied, among others, by Hsing (1986), Davis and Resnick (1989), and Zhang and Smith (2004). 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

$$X_n = \bigvee_{i=0}^{\infty} \psi_i Z_{n-i}, \quad n \in \mathbb{N}. \tag{15}$$

We require that $\sigma := \sum_{i=0}^{\infty} \psi_i^\alpha < \infty$. This ensures that the process exists and is strictly stationary (cf. Hsing 1986). We first collect some useful facts.

Proposition 9 *Let (X_n) be the max-moving average process in Eq. 15. 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 This result follows from computations similar to others appearing in the literature; see for example Davis and Resnick (1989). □

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 as a function of lag.

Proposition 10 *For the max-moving average process* (X_n) *in Eq. 15 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*

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

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 Eq. 8 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 $\alpha\beta \leq 2$.

Consider the transformed series (X_n^*), where $X_n^* = X_n^\alpha$. This is max-moving average with innovations $Z_n^* = Z_n^\alpha \sim \Phi_{1,\sigma}(x)$ and coefficients $\psi_i^* = \psi_i^\alpha$. By Proposition 7, its extremogram γ_{AB}^* coincides with γ_{AB}, as can also be seen from a direct calculation,

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

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 $n \to \infty$,

$$\mathrm{EDM}(h) \sim \frac{1}{2\sigma} \sum_{i=0}^{\infty} \left[(1+i)^{-\beta}\right]^{\alpha-1} (1+i+h)^{-\beta} = \frac{1}{2\sigma} \sum_{i=1}^{\infty} \left[i^{\alpha-1}(i+h)\right]^{-\beta}. \quad (16)$$

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

$$I(h) \leq \sum_{i=1}^{\infty} \left[i^{\alpha-1}(i+h)\right]^{-\beta} \leq (1+h)^{-\beta} + I(h). \tag{17}$$

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

$$\int_0^h t^{\alpha\beta-\beta-2}(1+t^{-1})^{-\beta} dt \leq \int_0^h t^{\alpha\beta-\beta-2} dt \to \text{const} < \infty \qquad (h \to \infty).$$

   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 Eqs. 16 and 17 we obtain

$$\text{EDM}(h) \sim \frac{1}{2\sigma} \text{const} \cdot h^{-\alpha\beta+1}.$$

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 $(1+h)^{-\beta}/h^{-\beta} \to 1$, Eq. 17 only lets us conclude, together with Eq. 16, that

$$\text{EDM}(h) \asymp h^{-\beta} \qquad (h \to \infty).$$

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

$$\text{EDM}(h) \sim \frac{1}{2\sigma} h^{-\beta} L(h) \qquad (h \to \infty).$$

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, 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.

Let (X_n^*) with $X_n^* = X_n^\alpha$ be the standardized sequence, which is max-moving average with $\Phi_{1,\sigma}$-innovations and coefficients $\psi_i^* = (1+i)^{-\alpha\beta}$. The EDM thus decays with rate $\min(\alpha\beta, \alpha\beta - 1) = \alpha\beta - 1$, like the extremogram for the original sequence, in accordance with Proposition 8.

This complements 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, and 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.,

$$X_n = \sum_{i=0}^{\infty} \psi_i Z_{n-i}, \qquad n \in \mathbb{N}, \tag{18}$$

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 (1985) 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 (c.f. Cline 1983; Hult and Samorodnitsky 2008). As remarked by Hsing (1986, 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 (18).

Specialize to the case where (X_n) is an AR(1) process with parameter $\phi \in (0, 1)$. Its MA representation (18) 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 Section 2.6 in Davis and Mikosch (2009) and Theorem 3.3 in Fasen et al. (2010):

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

This example does not exhibit long-range tail dependence in the sense that both the EDM and the extremogram are summable for all α. 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. 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, has a rate of decay that depends crucially on α.

5 Summary and concluding remarks

We now provide 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

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

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

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

For these models, 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 on α only when tails are sufficiently heavy (α is small) *and* the coefficients decay slowly, e.g. as a power function. (We have, of course, only shown this for the particular case of max-moving averages.)

Behavior under standardization By Proposition 7, the large lag asymptotics of the extremogram are invariant under standardization of the time series. This is not true for the EDM; a counterexample is the max-moving average model. 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 are objections to using such a procedure, for instance in Fasen et al. (2010, p. 16). They argue that standardization in general changes the model, and thus the dependence structure. This view has the consequence (see Proposition 7) that the extremogram fails to capture aspects of the large lag asymptotic dependence since it is invariant under standardization. This view also has the consequence that one views the copula transformation as changing the dependence structure.

Of course one could attempt to cure the lack of invariance of the EDM under standardization by changing the definition, as the integrand in the current definition does not depend on α. This is an approach adopted by *covariation*; see Samorodnitsky and Taqqu (1994). The presumption when the EDM was first defined by Resnick (2004), was that standardization had been applied to the data since when examining for instance bivariate heavy tailed data, the two tail indices are almost never estimated to be equal and hence without some sort of standardization procedure, the spectral measure will not exist.

Basic invariance Although the EDM depends on the choice of norm, we have seen in Proposition 3 that the particular choice does not influence the large lag asymptotics. Similarly, large lag asymptotics of the extremogram for the sets $A = (u, \infty)$, $B = (w, \infty)$ depend neither on the particular choice of u and w, as long as they are strictly positive (Proposition 5) nor on choice of norm.

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 $\boldsymbol{X}$. Theorem 2 in Resnick (2004) shows 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 Resnick (2004) is still true under the alternative assumption $\gamma_{AB}(h) = 0$ for all $h \geq h_0$.

Appendix: A central limit theorem for the EDM

In Resnick (2004), 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 (2009) 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 Resnick (2004) and Davis and Mikosch (2009) 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 Eq. 13 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*

$$\lim_{n\to\infty} \sqrt{k} \left(\frac{n}{k} \mathrm{E}(h(\boldsymbol{\Theta}_1) \mathbf{1}_{[R_1/b(\frac{n}{k}) \geq t^{-\gamma}]}) - \mathrm{E}h(\tilde{\boldsymbol{\Theta}}) \frac{n}{k} \bar{F}(b(n/k)t^{-\gamma}) \right) = 0 \qquad (19)$$

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

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

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 (19). 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 Eq. 19 is zero for all n. The condition is thus automatically satisfied in this case.

A sufficient condition for Eq. 19 is that

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

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

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

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 Resnick (2007, 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

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

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

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

First consider D_n. We have

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

It follows from Theorem 6.2 (9) in Resnick (2007) 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 (19) says that $B_n \to 0$ locally uniformly, so $D_n \xrightarrow{\mathrm{P}} 0$.

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

$$Y_n(r) = \frac{1}{\sigma_n\sqrt{k}}\sum_{j=1}^{kr}\left(h\left(\mathbf{\Theta}_{i(j,n)}\right) - \mathrm{E}h\left(\mathbf{\Theta}_{i(j,n)}\right)\right). \tag{20}$$

By the *Découpage de Lévy*, the sequence $\{\mathbf{\Theta}_{i(j,n)}\}$ 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 Resnick (2004) 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

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

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 $R_i/b(n/k) \in \left[t_m^{-\gamma}, s_m^{-\gamma}\right)$, and we set

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

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 $D_n^m \xrightarrow{\text{P}} 0$ for each m. Next, for each m, define processes Y_n^m as in Eq. 20, but with $i_m(j,n)$ instead of $i(j,n)$. The Découpage de Lévy implies that the M sequences $\{\boldsymbol{\Theta}_{i_m(j,n)} : j = 1, 2, \ldots\}$ 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:

$$\big(Y_n^1, \ldots, Y_n^M\big) \Rightarrow \big(Y^1, \ldots, Y^M\big),$$

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

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

This proves finite-dimensional convergence.

Now, since the limit process W has continuous paths, Theorem 13.5 in Billingsley (1999) yields the result as soon as we verify

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

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

$$\limsup_n \mathrm{E}\big(|W_n(t) - W_n(s)|^2 |W_n(s) - W_n(r)|^2\big) \le (t-r)^2 \tag{21}$$

for all $0 \le r \le s \le t$. This fact was also used in Resnick and Stărică (1997). So fix r, s, t and write

$$\alpha_i = \big(h(\boldsymbol{\Theta}_i) - \mathrm{E}(h(\tilde{\boldsymbol{\Theta}}))\big)\mathbf{1}_{[R_i/b(\frac{n}{k}) \in [t^{-\gamma}, s^{-\gamma})]}$$

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

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

$$\begin{aligned}
&\sigma^4 \mathrm{E}\big(|W_n(t) - W_n(s)|^2 |W_n(s) - W_n(r)|^2\big) \\
&= \frac{1}{k^2}\mathrm{E}\left(\sum_i \alpha_i\right)^2 \left(\sum_i \beta_i\right)^2 \\
&= \frac{n(n-1)}{k^2}\mathrm{E}\big(\alpha_1^2\beta_2^2\big) + \frac{n(n-1)(n-2)}{k^2}\mathrm{E}\big(\alpha_1^2\beta_2\beta_3\big) \\
&\quad + \frac{n(n-1)(n-2)}{k^2}\mathrm{E}\big(\alpha_1\alpha_2\beta_3^2\big) + \frac{n(n-1)(n-2)(n-3)}{k^2}\mathrm{E}\big(\alpha_1\alpha_2\beta_3\beta_4\big).
\end{aligned}$$

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

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

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 $\mathrm{E}\big(\alpha_1^2\big) = \mathrm{E}g(R_1/b(n/k), \boldsymbol{\Theta}_1)$, we get

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

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

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

by assumption (19). Thus $\frac{n^2}{k}\big(\mathrm{E}(\alpha_1)\big)^2 = \big(\sqrt{k}\frac{n}{k}\mathrm{E}(\alpha_1)\big)^2 \to 0$, and similarly we also get $\frac{n^2}{k}(\mathrm{E}(\beta_1))^2 \to 0$. Combining these results yields

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

We conclude that Eq. 21 holds. □

References

Basrak, B., Segers, J.: Regularly varying multivariate time series. Stoch. Process. Their Appl. **119**, 1055–1080 (2009)

Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J.: Statistics of Extremes: Theory and Applications. Wiley, New York (2004)

Beran, J.: Statistics for Long-Memory Processes. Chapman & Hall, New York (1994)

Billingsley, P.: Convergence of Probability Measures, 2nd edn. Wiley Series in Probability and Statistics. Wiley, New York (1999)

Cline, D.: Estimation and Linear Prediction for Regression, Autoregression and ARMA with Infinite Variance Data. Thesis, Colorado State University. ProQuest LLC, Ann Arbor, MI (1983)

D'Auria, B., Resnick, S.I.: The influence of dependence on data network models. Adv. Appl. Probab. **40**, 60–94 (2008)

Davis, R.A., Resnick, S.I.: Limit theory for moving averages of random variables with regularly varying tail probabilities. Ann. Probab. **13**, 179–195 (1985)

Davis, R.A., Resnick, S.I.: Basic properties and prediction of max-arma processes. Adv. Appl. Probab. **21**(4), 781–803 (1989)

Davis, R.A., Mikosch, T.: The extremogram: a correlogram for extreme events. Bernoulli **15**(4), 977–1009 (2009)

Fasen, V., Klüppelberg, C., Schlather, M.: High-level dependence in time series models. Extremes **13**(1), 1–33 (2010)

Heffernan, J.: A directory of coefficients of tail dependence. Extremes **3**, 279–290 (2000)

Hernandez-Campos, F., Jeffay, K., Park, C., Marron, J.S., Resnick, S.I.: Extremal dependence: internet traffic applications. Stoch. Models **21**(1), 1–35 (2005)

Hsing, T.: Extreme value theory for suprema of random variables with regularly varying tail probabilities. Stoch. Process. Their Appl. **22**, 51–57 (1986)

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

Kulik, R., Soulier, P.: The tail empirical process for long memory stochastic volatility sequences. Stoch. Process. Their Appl. **121**(1), 109–134 (2011)

Leadbetter, M.: Extremes and local dependence in stationary sequences. Z. Wahrscheinlichkeitstheor. Verw. Geb. **65**(2), 291–306 (1983)

Ledford, A., Tawn, J.: Statistics for near independence in multivariate extreme values. Biometrika **83**(1), 169–187 (1996)

Ledford, A., Tawn, J.: Diagnostics for dependence within time series extremes. J. R. Stat. Soc. Ser. B **65**(2), 521–543 (2003)

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

Resnick, S.I.: Heavy-Tail Phenomena: Probabilistic and Statistical Modeling. Springer Series in Operations Research and Financial Engineering. Springer, Berlin (2007)

Resnick, S.I., Stărică, C.: Smoothing the Hill estimator. Adv. Appl. Probab. **29**(1), 271–293 (1997)

Samorodnitsky, G.: Long range dependence. Found. Trends Stoch. Syst. **1**(3), 163–257 (2006)

Samorodnitsky, G., Taqqu, M.: Stable Non-Gaussian Random Processes: Stochastic Models with Infinite Variance. Chapman & Hall, New York (1994)

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

Zhang, Z., Smith, R.L.: The behavior of multivariate maxima of moving maxima processes. J. Appl. Probab. **41**(4), 1113–1123 (2004)