January 22, 2013

MEASURES OF SERIAL EXTREMAL DEPENDENCE AND THEIR ESTIMATION

RICHARD A. DAVIS, THOMAS MIKOSCH, AND YUWEI ZHAO

ABSTRACT. The goal of this paper is two-fold: 1. We review classical and recent measures of serial extremal dependence in a strictly stationary time series as well as their estimation. 2. We discuss recent concepts of heavy-tailed time series, including regular variation and max-stable processes.

Serial extremal dependence is typically characterized by clusters of exceedances of high thresholds in the series. We start by discussing the notion of extremal index of a univariate sequence, i.e. the reciprocal of the expected cluster size, which has attracted major attention in the extremal value literature. Then we continue by introducing the extremogram which is an asymptotic autocorrelation function for sequences of extremal events in a time series. In this context, we discuss regular variation of a time series. This notion has been useful for describing serial extremal dependence and heavy tails in a strictly stationary sequence. We briefly discuss the tail process coined by Basrak and Segers to describe the dependence structure of regularly varying sequences in a probabilistic way. Max-stable processes with Fréchet marginals are an important class of regularly varying sequences. Recently, this class has attracted attention for modeling and statistical purposes. We apply the extremogram to max-stable processes. Finally, we discuss estimation of the extremogram both in the time and frequency domains.

1. INTRODUCTION

Measuring and estimating extremal dependence in a time series is a rather challenging problem. Since many real-life time series, especially those arising in finance and environmental applications, are non-Gaussian their dependence structure is not determined by their autocorrelation function. Correlations are moments of the observations and as such not well suited for describing the dependence of extremes which typically arise from the tails of the underlying distribution.

1.1. **The extremal index as reciprocal of the expected extremal cluster size.** Extremal dependence in a real-valued strictly stationary sequence (X_t) can be described by the phenomenon of extremal clustering. Given some sufficiently high threshold $u = u_n$, we would expect that exceedances of this threshold should occur according to a homogeneous Poisson process. if (X_t) is iid. On the other hand, for dependent (X_t) exhibiting extremal dependence, exceedances of u should cluster in the sense that an exceedance of a high threshold is likely to be surrounded by neighboring observations that also exceed the threshold. Although the notion of extremal clustering is intuitively appealing, a precise formulation is not so easy.

The intuition about extremal clusters in a time series can be made precise using point process theory. In the classical monograph by Leadbetter, Lindgren and Rootzén [40] the *point process of exceedances* of u was used to describe clusters of extremes as an asymptotic phenomenon when the threshold u_n converges to the right endpoint of the distribution F of X. (Here and in what follows, Y denotes a generic element of any strictly stationary sequence (Y_t).) To be more precise, (u_n) has to satisfy the condition $n\,\overline{F}(u_n) = n\,(1 - F(u_n)) \to \tau$ for some $\tau \in (0, \infty)$. Under this condition and

mixing assumptions, the point processes of exceedances converge weakly to a compound Poisson process (see Hsing et al. [31]):

$$(1.1) \qquad N_n = \sum_{i=1}^{n} \varepsilon_{i/n}\, I_{\{X_i > u_n\}} \xrightarrow{d} N = \sum_{i=1}^{\infty} \xi_i\, \varepsilon_{\Gamma_i}\,,$$

where the state space of the point processes is $(0,1]$, the points $0 < \Gamma_1 < \Gamma_2 < \cdots$ constitute a homogeneous Poisson process with intensity $\theta\tau$ on $(0,1]$ which is independent of an iid positive integer-valued sequence (ξ_i). Here $\theta \in [0,1]$ is the *extremal index* of the sequence (X_t). Thus, in an asymptotic way, a cluster of extremes is located at the Poisson points Γ_i with corresponding size ξ_i. The cluster size distribution $P(\xi = k)$, $k \geq 1$, contains plenty of information about the distribution of the extremal clusters. However, most attention has been given to determine the expected cluster size $E\xi$ which can be interpreted as reciprocal of θ as the following heuristic argument illustrates. Applying (1.1) on the set $(0,1]$ and taking expectations on both sides of the limit relation, we observe that

$$EN_n(0,1] = n\overline{F}(u_n) \to \tau = EN(0,1] = E\xi\, E\#\{i \geq 1 : \Gamma_i \leq 1\} = E\xi\, (\theta\,\tau)\,, \quad n \to \infty.$$

Thus $\theta = 1/E\xi$ with the convention that $E\xi = \infty$ for $\theta = 0$. In the case of an iid sequence, $\xi_i \equiv 1$ a.s., i.e. N collapses to a homogeneous Poisson process and $\theta = 1$.

Writing $M_n = \max(X_1, \ldots, X_n)$, we also observe that

$$P(M_n \leq u_n) = P(N_n(0,1] = 0) \to P(N(0,1] = 0) = P(\#\{i \geq 1 : \Gamma_i \leq 1\} = 0) = \mathrm{e}^{-\theta\,\tau}\,,$$

while for an iid sequence $(\widetilde{X}_t)$ with the same marginal distribution F as for (X_t) and $\widetilde{M}_n = \max(\widetilde{X}_1, \ldots, \widetilde{X}_n)$ we have

$$P(\widetilde{M}_n \leq u_n) = F^n(u_n) = \mathrm{e}^{-n\overline{F}(u_n)(1+o(1))} \to \mathrm{e}^{-\tau}\,.$$

If F belongs to the *maximum domain of attraction of an extreme value distribution H* ($F \in \mathrm{MDA}(H)$) there exist constants $c_n > 0, d_n \in \mathbb{R}$, $n \geq 1$, such that $P(c_n^{-1}(\widetilde{M}_n - d_n) \leq x) \to H(x)$ for every $x \in \mathrm{supp}(H)$ (the support of H); cf. Embrechts et al. [20], Chapter 3. Thus, writing $u_n(x) = c_n\, x + d_n$ and $\tau = \tau(x) = -\log H(x)$ for $x \in \mathrm{supp}(H)$, the existence of an extremal value index θ of the sequence (X_t) implies that

$$(1.2) \qquad P(c_n^{-1}(M_n - d_n) \leq x) \to H^{\theta}(x)\,, \quad x \in \mathbb{R}\,.$$

The concrete form of the extremal index is known for various standard time series models, including linear processes with iid subexponential noise (cf. [20], Section 5.5), Markov processes (see Leadbetter and Rootzén [41], Perfekt [53]) and financial time series models such as GARCH (generalized autoregressive conditionally heteroscedastic) and SV (stochastic volatility) models; cf. [11, 13, 14]. Expressions of the extremal index for regularly varying sequences (X_t) (see Section 1.2.2 for a definition) in terms of the points of the limiting point process were given in Davis and Hsing [8] and in terms of the limiting tail process in Basrak and Segers [2]; see (1.10) below. However, for most models these concrete expressions of θ are too complex to be useful in practice.

An exception are Gaussian stationary sequences (X_t). Writing $\gamma_X(h) = \mathrm{cov}(X_0, X_h)$, $h \geq 0$, for the covariance function of (X_t), this sequence has extremal index $\theta = 1$ under the very weak condition $\gamma_X(h) = o(1/\log h)$ as $h \to \infty$ (so-called *Berman's condition*); see Leadbetter et al. [40], cf. Theorem 4.4.8 in Embrechts et al. [20]. Notice that Berman's condition is satisfied for fractional Gaussian noise and fractional Gaussian ARIMA processes (see Chapter 7 in Samorodnitsky and Taqqu [64], and Section 13.2 in Brockwell and Davis [5]). Subclasses of the latter processes exhibit long range dependence in the sense that $\sum_h |\gamma_X(h)| = \infty$.[1]

---

[1]This remark also indicates that long range dependence for extremes should not be defined via the covariance function γ_X. As explained above (see (1.2)) the existence of a positive extremal index θ ensures that the type of

We conclude that any Gaussian stationary sequences which are relevant for applications do not exhibit extremal clustering in the sense that $\theta = 1$. If $\theta = 1$ one often says that (X_t) exhibits asymptotic independence of its extremes. However, the notion of *asymptotic independence* is not well defined and may have rather different meanings in the extreme value context, as we will observe later.

Due to the complexity of expressions for the extremal index it has been recognized early on that θ needs to be estimated from real-life or simulated data. Various estimators were proposed in the literature. Among them, the blocks and runs estimators are the most popular ones. These estimators are non-parametric estimators of θ which, in different ways, define and count clusters in the sample and use this information to build estimators of θ under mixing conditions. In addition to the delicate choice of a threshold u_n, these estimation techniques also involve the construction of blocks of constant (but increasing with the sample size n) length or of flexible length depending on the local extremal behavior. These estimators often exhibit a rather large uncertainty.

In Figures 1.1 and 1.2 we illustrate the estimation of θ for real and simulated data. We choose the simple *blocks estimator* $\widehat{\theta} = K_n/N_n$ of the extremal index θ, where N_n is the number of exceedances of the threshold $u = u_n$ in the sample $X_1, \ldots, X_n$ and K_n is the number of blocks of size $s = s_n$, $X_{(i-1)s+1}, \ldots, X_{is}$, $i = 1, \ldots, [n/s]$, with at least one exceedance of u.

Aspects of bias, variance and optimal choice of blocks for the estimation of θ were discussed in Smith and Weissman [65]. In a series of papers, Hsing [27, 28, 29, 30] studied the extremes of stationary sequences, including the asymptotic behavior of their extremal index estimators. The recent papers Robert [59, 60], Robert et al. [58], in particular [60], give historical accounts of estimation of θ and some new technology for the estimation of θ and the cluster size distribution $P(\xi = k)$, $k \geq 1$. The paper of Robert [60] is devoted to inference on the cluster size distribution. The literature on this topic is sparse; Robert [60] mentions Hsing [28] as a historical reference.

1.2. The extremogram: an asymptotic correlogram for extreme events.

Davis and Mikosch [13, 17] introduced another tool for measuring the extremal dependence in a strictly stationary $\mathbb{R}^d$-valued time series (X_t): the *extremogram* defined as a limiting sequence given by

$$(1.3) \qquad \gamma_{AB}(h) = \lim_{n \to \infty} n \operatorname{cov}(I_{\{a_n^{-1}X_0 \in A\}}, I_{\{a_n^{-1}X_h \in B\}}), \quad h \geq 0.$$

Here (a_n) is a suitably chosen normalization sequence and A, B are two fixed sets bounded away from zero. The events $\{X_0 \in a_n A\}$ and $\{X_h \in a_n B\}$ are considered as extreme ones and $\gamma_{AB}(h)$ measures the influence of the time zero extremal event $\{X_0 \in a_n A\}$ on the extremal event $\{X_h \in a_n B\}$, h lags apart. The choice of (a_n) depends on the situation at hand. To avoid ambiguity, we later assume that (a_n) satisfies the relation $n P(|X| > a_n) \sim 1$. With this choice of (a_n), $\gamma_{AB}(h) = \lim_{n \to \infty} n P(a_n^{-1}X_0 \in A, a_n^{-1}X_h \in B)$. Motivating examples of extremograms are the limiting conditional probabilities $\lim_{n \to \infty} P(a_n^{-1}X_h \in B \mid a_n^{-1}X_0 \in A)$ in Davis and Mikosch [13, 17].

---

the limiting extreme value distribution H remains the same as in the iid case. This is easily checked since the only possible non-degenerate limit distributions H are the types of the *Fréchet distribution* $\Phi_\alpha(x) = e^{-x^{-\alpha}}$, $x, \alpha > 0$, the *Weibull distribution* $\Psi_\alpha(x) = e^{-(-x)^\alpha}$, $x < 0, \alpha > 0$, and the *Gumbel distribution* $\Lambda(x) = e^{-e^{-x}}$, $x \in \mathbb{R}$. This is a consequence of the Fisher-Tippett theorem; cf. Embrechts et al. [20], Theorem 3.2.3. The notion of long range dependence in an extreme value sense would be reasonable if in (1.2) a limit distribution occurred which does not belong to the type of any of the three mentioned standard extreme value distributions. This, however, can only be expected if a given stationary sequence (X_t) with $F \in \mathrm{MDA}(H)$ does not have an extremal index or if $\theta = 0$. Examples of sequences with zero extremal index are given in Leadbetter et al. [40] and Leadbetter [39], but such examples are often considered pathological; see also the discussion in Samorodnitsky [63] who studied infinite variance stable stationary sequences with zero extremal index and the boundary between short and long range extremal dependence for these sequence.
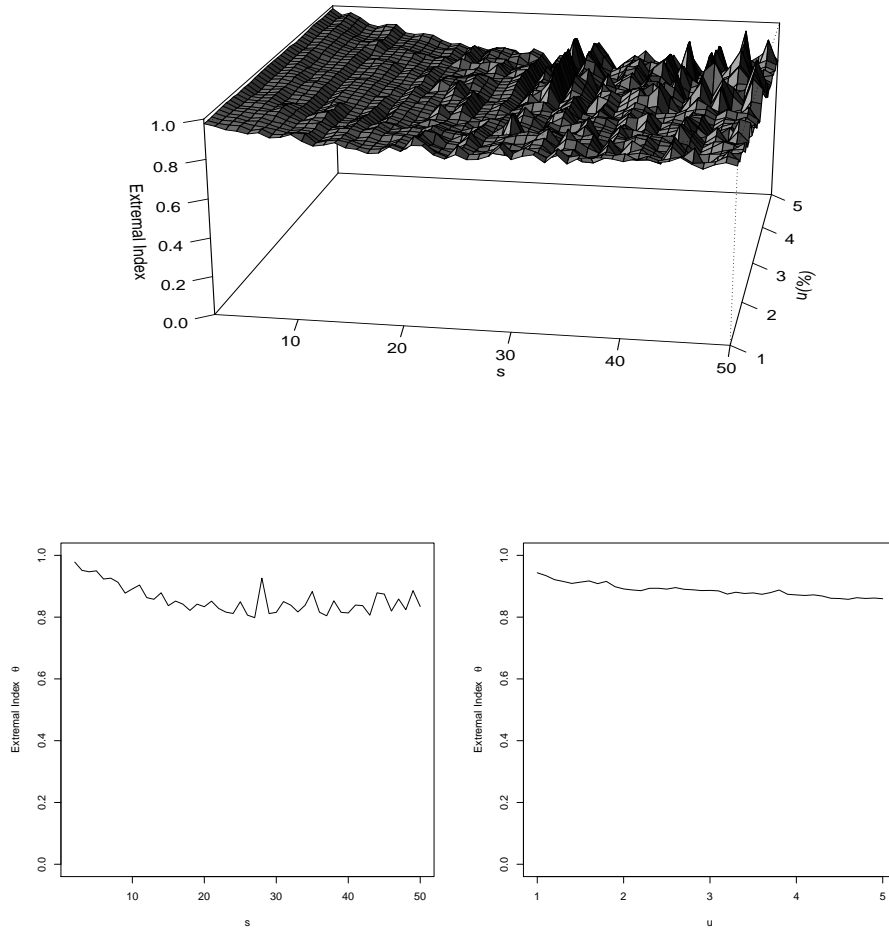
Figure 1.1. Blocks estimator $\widehat{\theta}$ of the extremal index θ for $31,757$ 5-minute log-returns of Bank of America stock prices. The blocks estimator as a function of the block size s and the $u\%$ upper order statistics (top), for fixed $u = 1.9\%$ and running s (bottom left) and for fixed $s = 10$ and running u (bottom right).

A motivating example for $d = 1$ with $A = B = (1, \infty)$ is the so-called *(upper) tail dependence coefficient* of the vector (X_0, X_h) given as the limit

$$(1.4) \qquad\qquad \rho(h) = \lim_{x \to \infty} P(X_h > x \mid X_0 > x).$$

(Here we assume that X has infinite right endpoint.) These pairwise tail dependence coefficients have attracted some attention in the literature on quantitative risk management; see for example McNeil et al. [42]. Notice that $\rho(h)$ coincides with $\gamma_{AA}(h)$ if we choose (a_n) such that $nP(X_0 > a_n) \sim 1$ as
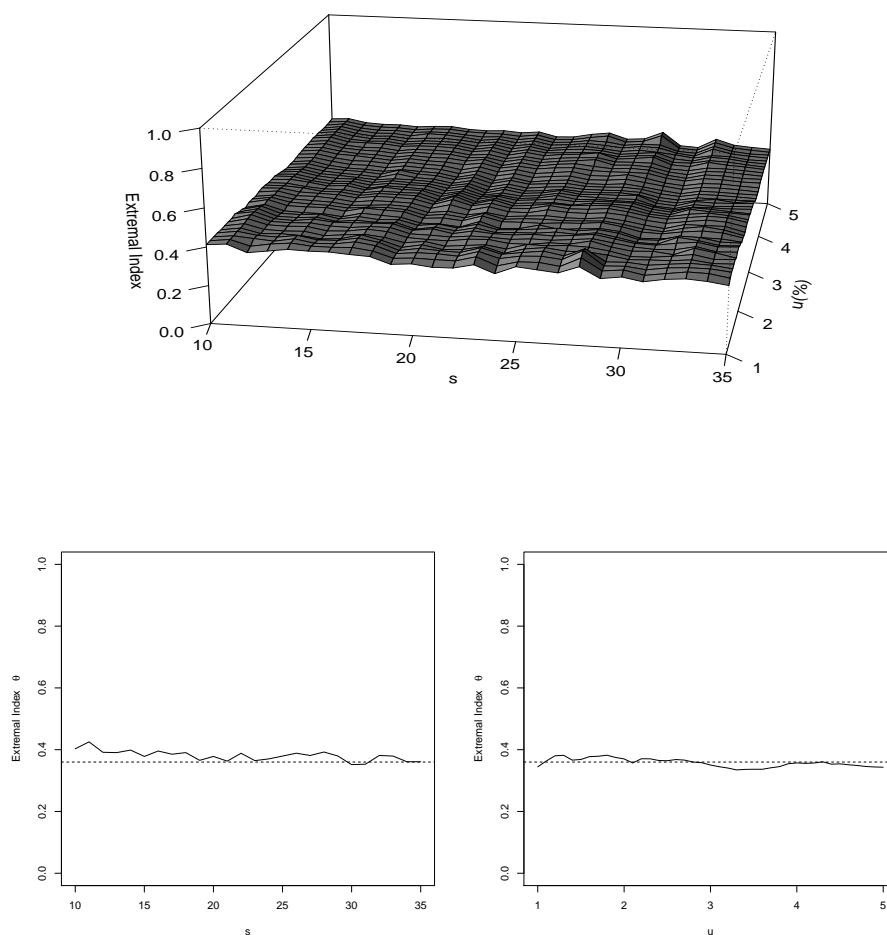
Figure 1.2. Blocks estimator $\widehat{\theta}$ of the extremal index θ for a sample of size $20\,000$ from the AR(1) process $X_t = 0.8\,X_{t-1} + Z_t$. The iid noise (Z_t) has a common student distribution with $\alpha = 2$ degrees of freedom. The extremal index $\theta = 0.37$ is known (indicated by dashed line); see [20], Section 8.1. The blocks estimator as a function of the block size s and $u\%$ of the upper order statistics (top), for fixed $u = 2\%$ and running s (bottom left) and for fixed $s = 24$ and running u (bottom right).

$n \to \infty$. Indeed,

$$
\begin{aligned}
n\,\mathrm{cov}(I_{\{X_0 > a_n\}}, I_{\{X_h > a_n\}}) \quad &\sim \quad \frac{P(X_h > a_n, X_0 > a_n) - (P(X_0 > a_n))^2}{P(X_0 > a_n)} \\
&\sim \quad P(X_h > a_n \mid X_0 > a_n).
\end{aligned}
$$

A similar calculation for any dimension d and suitable sets A, B shows that the limiting sequence

$$
(1.5) \qquad \left(\begin{array}{cc} \gamma_{AA}(h) & \gamma_{AB}(h) \\ \gamma_{BA}(h) & \gamma_{BB}(h) \end{array} \right), \quad h \geq 0,
$$

inherits the properties of a matrix covariance function. Notice that the entries of these matrices cannot be negative. The interpretation of (1.5) as covariance function allows one to use the classical notions of time series analysis *in an asymptotic sense*. For example, notions such as long or short range dependence of extremal events can be made precise by specifying the rate of decay of (1.5) as $h \to \infty$. Davis and Mikosch [16], Mikosch and Zhao [51] introduced an analog of the spectral density as a Fourier transform of the sequence (1.5). They showed that the periodogram of the sequence of indicators $I_{\{a_n^{-1} X_t \in A\}}$ of the extremal events $\{a_n^{-1} X_t \in A\}$, $t \in \mathbb{Z}$, has properties similar to the classical periodogram of a stationary sequence. In particular, weighted averages of the periodogram are consistent estimators of the spectral density.

In the literature, the pairwise tail dependence coefficients (1.4) are mostly considered for concrete examples of distributions, such as elliptical ones, including the multivariate t- and Gaussian distributions; see e.g. McNeil et al. [42]. In these cases, one can verify that the limits in (1.4) exist. For a Gaussian stationary sequence, $\rho(h) = 0$, $h \geq 1$, unless $X_t = X$ a.s. for all $t \in \mathbb{Z}$. The case $\rho(h) = 0$ for some $h \geq 1$ is (again) referred to as *asymptotic extremal dependence* in the vector (X_0, X_h) although no extremal index is in view.

In general, it is not obvious whether the limits $\rho(h)$ and, more generally, $\gamma_{AB}(h)$ for $h \geq 0$ exist. In this paper, we will use the notion of a *regularly varying stationary sequence*. It is a sufficient condition for the existence of the limits $\gamma_{AB}(h)$. Roughly speaking, a regularly varying sequence of random variables (X_t) has power law tails for every lagged vector $(X_1, \ldots, X_h)$, $h \geq 1$. In what follows, we make precise what regular variation means.

1.2.1. *Regularly varying random vectors.* The notion of regular variation is basic in extreme value theory and limit theory for partial sums of iid random variables. In multivariate extreme value theory, regular variation with index $\alpha > 0$ of the d-dimensional iid random vectors X_t, $t \in \mathbb{R}$, with values in $(0, \infty)^d$ is necessary and sufficient for the fact that the normalized sequence of component-wise maxima $(a_n^{-1} \max_{t \leq n} X_t^{(i)})_{i=1,\ldots,d}$, $t = 1, 2, \ldots$, converges in distribution to a d-dimensional extreme value distribution H on $(0, \infty)^d$ whose marginal distributions are Fréchet Φ_α-distributed; see Resnick [56] for a general theory of multivariate extremes for iid sequences. Similarly, for a general $\mathbb{R}^d$-valued iid sequence (X_t), the sequence of suitably normalized and centered partial sums $a_n^{-1}(X_1 + \cdots + X_n - b_n)$ converges in distribution to an infinite variance α-stable limit if and only if the distribution of X is regularly varying with index α. The index α is then necessarily in the range $\alpha \in (0, 2)$. We refer to Rvačeva [62] and Resnick [57] for proofs of this fact.

Various definitions of a d-dimensional regularly varying vector X exist; we refer to Resnick [55, 56, 57]. We start with a definition in terms of spherical coordinates. We say that X is regularly varying with index $\alpha > 0$ and spectral measure $P(\Theta \in \cdot)$ on the Borel σ-field of the unit sphere $\mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$ if [2] the following weak limits exist for every fixed $t > 0$:

$$
(1.6) \qquad \frac{P(|X| > tx, X/|X| \in \cdot)}{P(|X| > x)} \xrightarrow{w} t^{-\alpha} P(\Theta \in \cdot), \quad x \to \infty.
$$

Relation (1.6) can be written in an equivalent form as a pair of conditions:

---

[2]The choice of the norm $|\cdot|$ is relevant for defining the corresponding unit sphere and the spectral measure on it, but the notion of regular variation of a vector does not depend on a particular choice of norm. In this paper, $|\cdot|$ will stand for the Euclidean norm.

(1) The norm $|X|$ is regularly varying in the classical sense, i.e. $P(|X| > tx)/P(|X| > x) \to t^{-\alpha}$, $t > 0$, or, equivalently, $P(|X| > x) = x^{-\alpha}L(x)$, $x > 0$, for a slowly varying function L; cf. Bingham et al. [3] for an encyclopedia on regularly varying functions.

(2) The angular component $X/|X|$ is independent of $|X|$ for large values of $|X|$ in the sense that

$$(1.7) \qquad P(X/|X| \in \cdot \mid |X| > x) \xrightarrow{w} P(\Theta \in \cdot), \quad x \to \infty.$$

In any of these limit relations, it is possible to replace the converging parameter x by a sequence (a_n) such that $P(|X| > a_n) \sim n^{-1}$. Then (1.6) and (1.7), respectively, read as

$$n\,P(|X| > ta_n, X/|X| \in \cdot) \xrightarrow{w} t^{-\alpha}P(\Theta \in \cdot) \quad \text{and} \quad P(X/|X| \in \cdot \mid |X| > a_n) \xrightarrow{w} P(\Theta \in \cdot).$$

The convergence relation (1.6) can be understood as convergence on the particular Borel sets $\{\mathbf{x} \in \mathbb{R}^d : |\mathbf{x}| > t, \mathbf{x}/|\mathbf{x}| \in S\}$ for Borel sets $S \subset \mathbb{S}^{d-1}$ with a smooth boundary. This convergence can be extended to the Borel σ-field on $\overline{\mathbb{R}}_0^d = \overline{\mathbb{R}}^d \setminus \{\mathbf{0}\}$, $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty, -\infty\}$:

$$(1.8) \qquad \mu_x(\cdot) = \frac{P(x^{-1}X \in \cdot)}{P(|X| > x)} \xrightarrow{v} \mu(\cdot), \quad x \to \infty.$$

Here $\xrightarrow{v}$ refers to vague convergence of measures on the Borel σ-field on $\overline{\mathbb{R}}_0^d$, i.e. $\int_{\overline{\mathbb{R}}_0^d} f\,d\mu_x \to \int_{\overline{\mathbb{R}}_0^d} f\,d\mu$ as $x \to \infty$ for any continuous and compactly supported f on $\overline{\mathbb{R}}_0^d$; see Kallenberg [36], Resnick [56]. This means in particular, that the support of f is bounded away from zero. In view of (1.6), $\mu(\{\mathbf{x} \in \overline{\mathbb{R}}_0^d : |\mathbf{x}| > t, \mathbf{x}/|\mathbf{x}| \in S\}) = t^{-\alpha}P(\Theta \in S)$, and therefore μ is a Radon measure (i.e. finite on sets bounded away from zero) satisfying $\mu(tA) = t^{-\alpha}\mu(A)$, $t > 0$. In particular, μ does not charge points containing infinite components. Again, the parameter x in (1.8) can be replaced by a sequence (a_n) satisfying $P(|X| > a_n) \sim n^{-1}$ and then we get

$$n\,P(a_n^{-1}X \in \cdot) \xrightarrow{v} \mu(\cdot), \quad n \to \infty.$$

For an iid sequence (X_t) with generic element X, the latter condition is equivalent to the convergence of the point processes

$$N_n = \sum_{t=1}^{n} \varepsilon_{a_n^{-1}X_t} \xrightarrow{d} N,$$

where N is a Poisson random measure with mean measure μ and state space $\overline{\mathbb{R}}_0^d$; see Resnick [55, 56]. Since point process convergence is basic to extreme value theory, the notion of multivariate regular variation is very natural in the context of extreme value theory for multivariate observations with heavy-tailed components; see also the recent monograph by Resnick [57] who stresses the importance of the notion of regular variation as relevant for many applications in finance, insurance and telecommunications.[3]

1.2.2. *Regularly varying stationary sequences.* A strictly stationary sequence (X_t) is regularly varying with index α if its finite-dimensional distributions are regularly varying with index α, i.e. for every $h \geq 1$, there exist non-null Radon measures μ_h on the Borel σ-field of $\overline{\mathbb{R}}_0^h$ and a sequence (a_n) such that $a_n \to \infty$ and

$$(1.9) \qquad n\,P(a_n^{-1}(X_1, \ldots, X_h) \in \cdot) \xrightarrow{v} \mu_h(\cdot), \quad n \to \infty.$$

---

[3]The notion of regular variation is essentially dimensionless; see for example relation (1.6) which immediately extends to normed spaces and, more generally, to metric spaces. An account of the corresponding theory can be found in Hult and Lindskog [32]. Applications of regular variation in function spaces to extreme value theory can be found in de Haan and Tao [26], Davis and Mikosch [12], Meinguet and Segers [43], to large deviations in Hult et al. [33], Mikosch and Wintenberger [49, 50], and to random sets in Mikosch et al. [44, 45].

Here and in what follows, we will choose the normalizing sequence (a_n) such that $P(|X| > a_n) \sim n^{-1}$, where we use the notation (a_n) in a way different from Section 1.2.1. Indeed, in the latter section we defined the normalization $(a_n^{(h)})$ such that $P(|(X_1, \ldots, X_h)| > a_n^{(h)}) \sim n^{-1}$, but then the normalization would depend on the dimension h. This is not desirable. However, notice that

$$1 \quad = \quad \lim_{n \to \infty} \frac{P(|(X_1, \ldots, X_h)| > a_n^{(h)})}{P(|X| > a_n)} \,.$$

Therefore, by the properties of regularly varying functions, there exist positive constants $c_h^{1/\alpha} = \lim_{n \to \infty} a_n/a_n^{(h)}$, $h \geq 1$. Hence for sets A bounded away from zero such that $\mu_h(\partial A) = 0$ we have

$$
\begin{aligned}
n\, P((a_n^{(h)})^{-1}(X_1, \ldots, X_h) \in A) &= \quad n\, P(a_n^{-1}(a_n/a_n^{(h)})(X_1, \ldots, X_h) \in A) \\
&\sim \quad c_h\, [n\, P(a_n^{-1}(X_1, \ldots, X_h) \in A)] \\
&\to \quad c_h\, \mu_h(A) \,,
\end{aligned}
$$

i.e. the limit measures of regular variation under the different normalizations only differ by some positive constants.

The condition of regular variation on the sequence (X_t) seems to be a severe restriction since the tails of the marginals are power laws. However, following Resnick [56], Proposition 5.10, any multivariate distribution (with continuous marginals) in the maximum domain of attraction (MDA) of a d-dimensional extreme value distribution can be transformed to a distribution G with common Fréchet or Pareto marginals. Then G is in the MDA of an extreme value distribution with Fréchet marginals or, equivalently, G is regularly varying.

For example, transforming the marginals of a Gaussian stationary sequence to unit Fréchet, the resulting sequence is regularly varying with index $\alpha = 1$. We mentioned before that the tail dependence coefficient $\rho(h) = 0$, $h \geq 1$, for any non-trivial Gaussian stationary sequence. The quantities $\rho(h)$ remain invariant under monotone increasing transformations of the marginals. Hence, the transformed Gaussian distribution with unit Fréchet marginals exhibits asymptotic independence in the sense that the limit measures μ_h are concentrated on the axes.

The tail process. An insightful characterization of an $\mathbb{R}^d$-valued regularly varying stationary sequence (X_t) was given in Theorem 2.1 of Basrak and Segers [2]: there exists a sequence of $\mathbb{R}^d$-valued random vectors $(Y_t)_{t \in \mathbb{Z}}$ such that $P(|Y_0| > y) = y^{-\alpha}$ for $y > 1$ and for any $h \geq 0$,

$$P(x^{-1}(X_{-h}, \ldots, X_h) \in \cdot \mid |X_0| > x) \xrightarrow{w} P((Y_{-h}, \ldots, Y_h) \in \cdot), \quad x \to \infty \,.$$

The process (Y_t) is the *tail process* of (X_t). Writing $\Theta_t = Y_t/|Y_0|$ for $t \in \mathbb{Z}$, one also has for $h \geq 0$,

$$P(|X_0|^{-1}(X_{-h}, \ldots, X_h) \in \cdot \mid |X_0| > x) \xrightarrow{w} P((\Theta_{-h}, \ldots, \Theta_h) \in \cdot), \quad x \to \infty \,.$$

The process (Θ_t) is independent of $|Y_0|$ and called the *spectral tail process* of (X_t). Notice that $P(\Theta_0 \in \cdot)$ is the spectral measure of X.

Basrak and Segers [2] also gave an expression for the extremal index in terms of the spectral tail process:

$$(1.10) \qquad \theta = E\Big[\sup_{t \geq 0} |\Theta_t|^\alpha - \sup_{t \geq 1} |\Theta_t|^\alpha\Big] \,.$$

1.2.3. *The extremogram revisited.* Now consider an $\mathbb{R}^d$-valued regularly varying stationary (X_t). Then the extremogram $\gamma_{AB}(h)$, $h \geq 0$, is well defined. Indeed, for every $h \geq 0$, the vector $(X_1, \ldots, X_{h+1})$ is regularly varying with limit measure μ_{h+1}. Then, with normalization (a_n) such that $P(|X| > a_n) \sim n^{-1}$,

$$n\, P(a_n^{-1} X_h \in B, a_n^{-1} X_0 \in A) \to \mu_{h+1}(A \times \mathbb{R}^{d(h-1)} \times B) = \gamma_{AB}(h), \quad h \geq 0 \,,$$

provided $A \times \mathbb{R}^{d(h-1)} \times B$ is a continuity set with respect to the measure μ_{h+1}. Similarly, for $d = 1$ and $A = B = (1, \infty)$,

$$\rho(h) = \frac{\mu_{h+1}(A \times \mathbb{R}^{h-1} \times A)}{\mu_{h+1}(A \times \mathbb{R}^h)}, \quad h \geq 0.$$

These limits can also be expressed in terms of the tail process. In the former case, assuming that A is bounded away from zero, there exists $\delta > 0$ such that $A \subset \{x \in \mathbb{R}^d : |x| > \delta\}$. Hence

$$\frac{P(a_n^{-1}X_h \in B, a_n^{-1}X_0 \in A)}{P(|X| > a_n)}$$
$$= \frac{P(a_n^{-1}X_h \in B, a_n^{-1}X_0 \in A, |X_0| > \delta a_n)}{P(|X| > a_n)}$$
$$= \frac{P((\delta a_n)^{-1}X_h \in \delta^{-1}B, (\delta a_n)^{-1}X_0 \in \delta^{-1}A, |X_0| > \delta a_n)}{P(|X| > \delta a_n)} \frac{P(|X| > \delta a_n)}{P(|X| > a_n)}$$
$$\to \quad P((Y_0, Y_h) \in \delta^{-1}(A \times B)) \delta^{-\alpha}$$
$$= \quad P((Y_0, Y_h) \in A \times B) = \gamma_{AB}(h).$$

Similarly, for $d = 1$ and $A = B = (1, \infty)$, assuming that $\lim_{x \to \infty} P(X > x)/P(|X| > x) = E(\Theta_0)_+^\alpha = P(\Theta_0 = 1) > 0$,

$$\frac{P(X_h > a_n, X_0 > a_n)}{P(X > a_n)} = \frac{P(X_h > a_n, X_0 > a_n, |X_0| > a_n)}{P(|X| > a_n)} \frac{P(|X| > a_n)}{P(X > a_n)}$$
$$\to \quad P(Y_h > 1 \mid Y_0 > 1)$$
$$= \quad \frac{P(|Y_0| \min(\Theta_0, \Theta_h) > 1)}{P(|Y_0| \Theta_0 > 1)}$$
$$= \quad \frac{E(\min(\Theta_0, \Theta_h))_+^\alpha}{E(\Theta_0)_+^\alpha} = \rho(h).$$

1.2.4. *Examples of regularly varying sequences and their extremograms.* In this section, we will introduce some important classes of real-valued strictly stationary regularly varying stationary sequences with index $\alpha > 0$. We will also give the values of the extremogram $\rho(h)$, $h \geq 1$, in (1.4). For the calculation of ρ in these examples, we refer to [16, 17, 51].

IID sequence. An iid sequence (Z_t) is regularly varying with index α if and only if Z is regularly varying with the same index; the limit measures μ_h are concentrated on the axes and $\rho(h) = 0$, $h \geq 1$.

Linear process. Historically, the class of linear processes with regularly varying iid real-valued noise (Z_t) has attracted attention in extreme value theory and in time series analysis. A (causal) linear process

$$(1.11) \qquad\qquad X_t = \sum_{t=0}^{\infty} \psi_j Z_{t-j}, \quad t \in \mathbb{Z},$$

inherits regular variation under conditions on the deterministic sequence (ψ_i) which are close to those dictated by the 3-series theorem, ensuring the a.s. convergence of the series in (1.11). This fact was proved in Mikosch and Samorodnitsky [47] for the distribution of X. The regular variation of the finite-dimensional distributions of (X_t) follows since regular variation is preserved under affine transformations of regularly varying vectors. The class (1.11) includes causal ARMA processes which are relevant for applications. We refer to Chapter 7 of Embrechts et al. [20] for various applications of regularly varying linear processes.

Under the tail balance condition $P(Z > x) \sim p\,P(|Z| > x)$, $P(Z \leq -x) \sim q\,P(|Z| > x)$, as $x \to \infty$, for some $p, q \geq 0$ with $p + q = 1$,

$$\rho(h) = \frac{\sum_{i=0}^{\infty}\left[p\,(\min(\psi_i^+, \psi_{i+h}^+))^\alpha + q\,(\min(\psi_i^-, \psi_{i+h}^-))^\alpha\right]}{\sum_{i=0}^{\infty}\left[p\,(\psi_i^+)^\alpha + q\,(\psi_i^-)^\alpha\right]}, \quad h \geq 1\,.$$

Stochastic recurrence equations. Next to linear processes, solutions to the stochastic recurrence equation

$$(1.12) \qquad\qquad X_t = A_t\,X_{t-1} + B_t\,, \quad t \in \mathbb{Z}\,,$$

have attracted some attention. Here (A_t, B_t), $t \in \mathbb{Z}$, is an iid $\mathbb{R}^2$-valued sequence. An a.s unique causal solution to (1.12) exists under the moment conditions $E \log A^+ < 0$ and $E \log |B| < \infty$. It follows from work by Kesten [37] and Goldie [37] that X is regularly varying in the precise sense that

$$P(X > x) \sim c_+ x^{-\alpha} \quad \text{and} \quad P(X \leq -x) \sim c_- x^{-\alpha}\,, \quad x \to \infty\,,$$

for constants $c_+, c_- \geq 0$ such that $c_+ + c_- > 0$ provided the equation

$$(1.13) \qquad\qquad E|A|^\kappa = 1$$

has a positive solution α (which is unique due to convexity), $EB^\alpha < \infty$ and further regularity conditions on the distribution of A are satisfied. Iteration of (1.12) shows that the finite-dimensional distributions of (X_t) are regularly varying with index α. This fact is rather surprising since the distributions of A and B do not need to be heavy-tailed, in contrast to linear processes , where the noise (Z_t) itself has to be heavy-tailed to ensure regular variation of (X_t). We mention that the case of multivariate B and matrix-valued A has also been studied, starting with Kesten [37]; see the recent paper Buraczewski et al. [7].

Assuming $A > 0$ a.s., similar calculations as in the proof of Lemma 2.1 in [16] yield

$$(1.14) \qquad\qquad \rho(h) = E[\min(1, A_1 \cdots A_h)^\alpha]\,, \quad h \geq 1\,.$$

Models for returns. Log-returns $X_t = \log P_t - \log P_{t-1}$, $t \in \mathbb{Z}$, of a speculative price series (P_t) are often modeled of the form $X_t = \sigma_t Z_t$, where (σ_t) is a strictly stationary sequence of non-negative *volatilities* and (Z_t) is an iid multiplicative noise sequence. The feedback between (σ_t) and (Z_t) can be modeled in a rather flexible way.

Stochastic volatility models. The most simple approach is to assume that (σ_t) and (Z_t) be independent. The resulting time series model is frequently referred to as *stochastic volatility model*. Its probabilistic properties are rather simple; see Davis and Mikosch [15]. In particular, regular variation of (X_t) results if $E\sigma^{\alpha+\delta} < \infty$ for some $\delta > 0$ and (Z_t) is iid and regularly varying with index α. The corresponding limit measures μ_h in (1.9) are then concentrated on the axes; see Davis and Mikosch [11, 14],[4] and then also $\rho(h) = 0$, $h \geq 1$, as in the iid case. The situation changes if $E|Z|^{\alpha+\delta} < \infty$ for some $\delta > 0$ and (σ_t) is regularly varying with index α. Then (X_t) is regularly varying with index α and extremal clustering for this sequence is possible; see Mikosch and Rezapur [46].

---

[4]The fact that μ_h, $h \geq 1$, is concentrated on the axes is also referred to as *asymptotic extremal independence.* Recall that, in an extreme value context, various other situations are also referred to as asymptotic extremal independence, among them the cases of unit extremal index and zero tail dependence coefficient. Asymptotic independence in the sense of the limiting measures μ_h is much more complex than the other notions which are just numerical characteristics. The fact that μ_h is concentrated on the axes means that it is very unlikely that any two values X_t and X_s, $s \neq t$, are big at the same time, just as for independent random variables. On the other hand, this kind of asymptotic independence heavily relies on the notion of multivariate regular variation.

GARCH model. Among the models for returns $X_t = \sigma_t Z_t$, $t \in \mathbb{Z}$, the GARCH family gained most popularity. The simplest model of its kind (ARCH) was introduced by Engle [21] and the more sophisticated GARCH model by Bollerslev [4]. For simplicity, we consider the GARCH$(1,1)$ case given by $\sigma_t^2 = \alpha_0 + \sigma_{t-1}^2(\alpha_1 Z_{t-1}^2 + \beta_1)$, $t \in \mathbb{Z}$, where $\alpha_0, \alpha_1, \beta_1$ are positive constants with certain restrictions on the values of $\alpha_1 + \beta_1, \beta_1 < 1$, to ensure strict stationarity. Typical choices are standard normal or unit variance t-distributed Z. Notice that we can write $\sigma_t^2 = A_t \sigma_{t-1}^2 + B_t$, where $B_t = \alpha_0$ and $A_t = \alpha_1 Z_{t-1}^2 + \beta_1$, $t \in \mathbb{Z}$. Therefore regular variation of (σ_t) follows from the corresponding theory for stochastic recurrence equations, and (X_t) inherits the same property; see Davis and Mikosch [10] for the ARCH(1) case, Mikosch and Stărică [48] for the GARCH$(1,1)$, Basrak et al. [1] for GARCH(p,q) and the review paper Davis and Mikosch [13]. Real-life log-returns are typically heavy-tailed. The GARCH model captures this property and this was one of the reasons that it became a benchmark model in financial time series analysis from which numerous other models where derived. An expression of ρ for (X_t^2) is given by (1.14) with $A_t = \alpha_1 Z_t^2 + \beta_1$.

Infinite variance stable sequence. Stable processes with infinite variance have become popular due to their attractive theoretical and modeling properties; see Samorodnitsky and Taqqu [64]. The finite-dimensional distributions of an α-stable process are jointly α-stable, hence they are regularly varying with index $\alpha \in (0,2)$. The class of infinite variance stationary stable processes has been intensively studied; see Rosiński [61]. An expression of ρ is given in [16], Section 2.4.

Max-stable processes. This class of processes has recently attracted some attention since it is a flexible class for modeling heavy tails and spatio-temporal dependence. Since the finite-dimensional distributions of max-stable processes are explicitly given it is often simple to verify properties (such as regular variation) and to calculate certain quantities (e.g. mixing coefficients, extremal index). We will use this class of regularly varying processes to illustrate the general theory.

Following de Haan [23], a real-valued process $(\xi_t)_{t \in T}$, $T \subset \mathbb{R}$, is α-*max-stable* for some $\alpha > 0$ if its finite-dimensional distributions satisfy the relation

$$P(\xi_{t_1} \le x_1, \ldots, \xi_{t_d} \le x_n) = \exp\left\{-\int_{\mathbb{S}^{d-1} \cap \mathbb{R}_+^d} \max_{i \le d} \left(\frac{s_i}{x_i}\right)^\alpha \Gamma_{\mathbf{t}_d}(d\mathbf{s})\right\},$$
$$t_i \in T, i = 1,2,\ldots,d, \quad x_i > 0, \quad d \ge 1,$$

where $\Gamma_{\mathbf{t}_d}$ are finite measures on the unit sphere. This means in particular that the marginal distributions of the process ξ have a Fréchet distribution with parameter α given by[5]

(1.15)
$$\Phi_\alpha(x) = \mathrm{e}^{-x^{-\alpha}}, \quad x > 0.$$

De Haan [23] also introduced the notion of α-*max-stable integral*. Given a σ-finite measure space $(E, \mathcal{E}, \nu)$, consider a Poisson random measure $\sum_{i=1}^\infty \varepsilon_{(\Gamma_i, Y_i)}$ with $0 < \Gamma_1 < \Gamma_2 < \cdots$ on the state space $\mathbb{R}_+ \times E$ with mean measure $\mathbb{LEB} \times \nu$. For $f \ge 0$ with $f \in L^\alpha(E, \mathcal{E}, \nu)$ the max-stable integral is defined as

$$\int_E^\vee f \, dM_\nu^\alpha = \sup_{i \ge 1} \Gamma_i^{-1/\alpha} f(Y_i).$$

---

[5]The choice of Frechét Φ_α marginals is for convenience only; then results on regular variation are applicable. Since Gumbel or Weibull distributed random variables can be obtained by suitable increasing transformations of a Fréchet random variable any result for max-stable processes with Fréchet marginals can be formulated in terms of the transformed processes with Gumbel or Weibull marginals; see for example Kabluchko et. al [35] who formulated their results in terms of Gumbel distributions. Since the choice of the parameter α is also arbitrary in this context, most results in the literature are formulated for processes with unit Fréchet Φ_1 marginals.

Using the order statistics property of the homogeneous Poisson process with points (Γ_i), one obtains

$$(1.16) \qquad P\Big(\int_E^\vee f\, dM_\nu^\alpha \le x\Big) = \exp\Big\{ - x^{-\alpha} \int_E f^\alpha(y)\, \nu(dy)\Big\}, \quad x > 0\,.$$

Moreover, for any non-negative $f_i \in L^\alpha(E, \mathcal{E}, \nu)$, $i = 1, \ldots, d$, by (1.16),

$$
\begin{aligned}
P\Big(\int_E^\vee f_i\, dM_\nu^\alpha \le x_i\,, i = 1, \ldots, d\Big) &= P\Big(\int_E \max_{i=1,\ldots,d} \frac{f_i}{x_i}\, dM_\nu^\alpha \le 1\Big) \\
(1.17) \qquad\qquad &= \exp\Big\{ - \int_E \max_{i=1,\ldots,d} \Big(\frac{f_i(y)}{x_i}\Big)^\alpha \nu(dy)\Big\}, \quad x_i > 0\,, i = 1, \ldots, d\,.
\end{aligned}
$$

The notions of max-stable process and integral bear some resemblance with the corresponding α-stable ones; see Stoev and Taqqu [67], Kabluchko [34].

We will focus on *stationary ergodic* max-stable processes with integral representation

$$(1.18) \qquad X_t = \int_E^\vee f_t\, dM_\nu^\alpha\,, \quad t \in \mathbb{R}\,, \quad f_t \ge 0\,, \quad f_t \in L^\alpha(E, \mathcal{E}, \nu)\,.$$

As in the case of α-stable stationary ergodic processes (Rosiński [61]), the choice of (f_t) is rather sophisticated; see Stoev [66], Kabluchko [34] for details. De Haan [23] showed that any max-stable process with countable index set $T \subset \mathbb{R}$ and stochastically continuous sample paths has representation (1.18) and Kabluchko [34] proved this fact for any max-stable process on T for sufficiently rich measure spaces $(E, \mathcal{E}, \nu)$. In what follows, we will always assume that the considered max-stable processes have representation (1.18).

Next we give some basic properties of a stationary max-stable process.

Proposition 1.3. *The following statements hold for the skeleton process $(X_t)_{t\in\mathbb{Z}}$ of the process* (1.18).

(1) *The finite-dimensional distributions of (X_t) are regularly varying with index α and the limit measures μ_h of the finite-dimensional distributions are given by its values on the complements of the rectangles $(\mathbf{0}, \mathbf{x}] = \{\mathbf{y} \in \mathbb{R}^h : 0 < y_i \le x_i\,, i = 1, \ldots, h\}$, $h \ge 1$, $\mathbf{x} = (x_1, \ldots, x_h)$ with $x_i > 0$, $i = 1, \ldots, h$:*

$$(1.19) \qquad \mu_h\big((\mathbf{0}, \mathbf{x}]^c\big) = \frac{\int_E \max_{i=1,\ldots,h} \big(\frac{f_i(y)}{x_i}\big)^\alpha \nu(dy)}{\int_E f_0^\alpha(y)\, \nu(dy)}\,.$$

(2) *The sequence (X_t) has extremal index θ if and only if the limit*

$$(1.20) \qquad \theta = \lim_{n\to\infty} \frac{1}{n} \frac{\int_E \max_{t=1,\ldots,n} f_t^\alpha(y)\, \nu(dy)}{\int_E f_0^\alpha(y)\, \nu(dy)}$$

exists.

(3) *The extremogram for the sets $A = (a, \infty)$ and $B = (b, \infty)$, $a, b > 0$, is given by*

$$(1.21) \qquad \gamma_{AB}(h) = \frac{\int_E f_0^\alpha(y) \wedge \big(\frac{a}{b} f_h(y)\big)^\alpha \nu(dy)}{a^\alpha \int_E f_0^\alpha(y)\nu(dy)}\,, \quad h \ge 0\,,$$

and, for $a = b = 1$,

$$(1.22) \qquad \rho(h) = \frac{\int_E f_0^\alpha(y) \wedge f_h^\alpha(y)\, \nu(dy)}{\int_E f_0^\alpha(y)\, \nu(dy)}\,, \quad h \ge 1\,.$$

(4) Let S_1, S_2 be finite disjoint subsets of $\mathbb{Z}$ and $\sigma(C)$ the σ-field generated by $(X_t)_{t \in C}$ for any $C \subset \mathbb{Z}$. Recall the α-mixing coefficient relative to the sets S_1, S_2.

$$\alpha(S_1, S_2) = \sup_{A \in \sigma(S_1), B \in \sigma(S_2)} |P(A \cap B) - P(A) P(B)| .$$

and for $S_{-\infty}^0 = \{\ldots, -1, 0\}$, $S_h^\infty = \{h, h+1, \ldots\}$, $h \geq 1$, introduce the mixing rate function

$$\alpha_h = \alpha(S_{-\infty}^0, S_h^\infty), \quad h \geq 1 .$$

Then there exists a universal constant $c > 0$ such that

(1.23)
$$\alpha_h \leq c \sum_{s_1=-\infty}^{0} \sum_{s_2=0}^{\infty} \int_E f_0^\alpha(y) \wedge f_{h+s_2}^\alpha(y) \, \nu(dy), \quad h \geq 1 .$$

Remark 1.4. If the limit in (1.20) exists it belongs to the interval $[0, 1]$. Indeed, by stationarity of (X_t),

$$\int_E \max_{t=1,\ldots,n} f_t^\alpha(y) \, \nu(dy) \leq \sum_{t=1}^{n} \int_E f_t^\alpha(y) \, \nu(dy) = n \int_E f_0^\alpha(y) \, \nu(dy) .$$

Remark 1.5. Part (4) is a consequence of Corollary 2.2 in Dombry and Eyi-Minko [19] proved for β-mixing. If $\int_E f_0^\alpha(y) \wedge f_h^\alpha(y) \, \nu(dy) \leq c_0 \, e^{-c_1 h}$, $h \geq 1$, for some constants $c_0, c_1 > 0$, then we conclude that $\alpha_h \leq C \, e^{-c_1 h}$, for some $C > 0$.

Proof. **Part (1)** Since the integrals $\int_E^\vee f_i dM_\nu^\alpha$, $i = 1, \ldots, h$, are supported on $(0, \infty)$ it suffices to show that there exists a non-null Radon measure μ_h on the Borel σ-field of $\overline{\mathbb{R}}_0^d \cap (0, \infty]^d$ such that

(1.24)
$$n P\left(a_n^{-1}(X_1, \ldots, X_h)\right) \in [0, \mathbf{x}]^c) \to \mu_h\left([0, \mathbf{x}]^c\right) ,$$

where $\mathbf{x}$ is chosen such that $[0, \mathbf{x}]^c$ is a μ_h-continuity set and

$$P(X > a_n) = 1 - \exp\left\{ - a_n^{-\alpha} \int_E f_0^\alpha(x) \nu(dx) \right\} \sim n^{-1} ,$$

see Resnick [57], Theorem 6.1. A Taylor expansion argument shows that we can always choose

$$a_n = n^{1/\alpha} \left(\int_E f_0^\alpha(x) \, \nu(dx) \right)^{1/\alpha} .$$

An application of (1.17) and a Taylor expansion yield (1.24) with limit as specified in (1.19).
Part (2) Applying (1.17) for $x_i = x > 0$, we obtain

$$P(a_n^{-1} M_n \leq x) \quad = \quad \exp\left\{ - a_n^{-\alpha} x^{-\alpha} \int_E \max_{t=1,\ldots,n} f_t^\alpha(y) \nu(dy) \right\} .$$

By definition of the extremal index, the right-hand side must converge to $\Phi_\alpha^\theta(x)$ for some $\theta \in [0, 1]$. Equivalently, the limit θ in (1.20) exists.
Part (3) As regards the extremogram for sets $A = (a, \infty)$, $B = (b, \infty)$, $a, b > 0$, we have the relation

$$
\begin{aligned}
P(X_h > b\,x, X_0 > a\,x) \quad = \quad & P(X_h > b\,x) + P(X_0 > a\,x) - P((X_h/b) \vee (X_0/a) > x) \\
= \quad & 1 - \exp\left\{ - x^{-\alpha} \int_E \left(\frac{f_0(y)}{a}\right)^\alpha \nu(dy) \right\} - \exp\left\{ - x^{-\alpha} \int_E \left(\frac{f_h(y)}{b}\right)^\alpha \nu(dy) \right\} \\
& + \exp\left\{ - x^{-\alpha} \int_E \left(\frac{f_0(y)}{a}\right)^\alpha \vee \frac{f_h(y)}{b}\right)^\alpha \nu(dy) \right\} .
\end{aligned}
$$

In view of stationarity, $\int_E f_h^\alpha(y) \, \nu(dy) = \int_E f_0^\alpha(y) \, \nu(dy)$. Using a Taylor expansion as $x \to \infty$, we obtain the desired formulas (1.21) and (1.22)

Part (4) We obtain from Corollary 2.2 in Dombry and Eyi-Minko [19] for any disjoint closed countable subsets S_1, S_2 of $\mathbb{R}$

$$\alpha(S_1, S_2) \leq c \sum_{s_1 \in S_1} \sum_{s_2 \in S_2} \int_E f_0^\alpha(y) \wedge f_{|s_1 - s_2|}^\alpha(y)\, \nu(dy)\,.$$

Then (1.23) is immediate.   □

Next we consider two popular models of max-stable processes.

Example 1.6. The *Brown-Resnick process* (see [6]) has representation

$$(1.25) \qquad\qquad X_t = \sup_{i \geq 1} \Gamma_i^{-1/\alpha}\, e^{\,W_i(t) - 0.5\,\sigma^2(t)}\,, \quad t \in \mathbb{R}\,,$$

where (Γ_i) is an enumeration of the points of a unit rate homogeneous Poisson process on $(0, \infty)$ independent of the iid sequence (W_i) of sample continuous mean zero Gaussian processes on $\mathbb{R}$ with stationary increments and variance function σ^2. The max-stable process (1.25) is stationary (Theorem 2 in Kabluchko et al. [35]) and its distribution only depends on the variogram $V(h) = \mathrm{var}(W(t + h) - W(t))$, $t \in \mathbb{R}, h \geq 0$. It follows from Example 2.1 in Dombry and Eyi-Minko [19] that the functions (f_t) in representation (1.18) satisfy the condition

$$\int_E f_0^\alpha(y) \wedge f_h^\alpha(y)\nu(dy) \leq c\, \overline{\Phi}(0.5\sqrt{V(h)})\,,$$

where Φ is the standard normal distribution. For example, if W is standard Brownian motion, $V(h) = h$, $\overline{\Phi}(0.5\sqrt{h}) \sim c\, e^{-h/8} h^{-0.5}$, as $h \to \infty$. An application of Remark 1.5 shows that (α_h) decays at an exponential rate.

Recently, the Brown-Resnick process has attracted some attention for modeling spatio-temporal extremes; see [34, 35, 66, 52]. The processes (1.25) can be extended to random fields on $\mathbb{R}^d$. These fields found various applications for modeling spatio-temporal extremal effects; see Kabluchko et al. [35]. For further spatio-temporal applications of max-stable random fields, see also Davis et al. [9].

Example 1.7. We consider de Haan and Pereira's [25] *max-moving process*

$$(1.26) \qquad\qquad X_t = \sup_{i \geq 1} \Gamma_i^{-1/\alpha} f(t - U_i)\,, \quad t \in \mathbb{R}\,,$$

where f is a continuous Lebesgue density on $\mathbb{R}$ such that $\int_\mathbb{R} \sup_{|h| \leq 1} f(x+h)\, dx < \infty$ and $\sum_{i=1}^\infty \varepsilon_{(\Gamma_i, U_i)}$ are the points of a unit rate homogeneous Poisson random measure on $(0, \infty) \times \mathbb{R}$.

The resulting process (X_t) is α-max-stable and stationary. According to Example 2.2 in Dombry and Eyi-Minko [19],

$$\int_E f_0^\alpha(y) \wedge f_h^\alpha(y)\nu(dy) \leq c \int_\mathbb{R} \min(f(-x), f(h - x))\, dx\,, \quad h \geq 0\,.$$

For example, if f is the standard normal density, this implies that (α_h) decays to zero faster than exponentially, i.e. the memory in this sequence is very short. In Figure 1.2.4 a simulation of the corresponding process (1.26) for $\alpha = 5$ is shown.

2. Estimation of the extremogram

2.1. **Asymptotic theory.** Natural estimators of the extremogram are obtained by replacing the probabilities in the limit relations (1.3) and (1.4) with their empirical counterparts. In this context, one works with quantities which are derived from the *tail empirical process*; see the monographs de Haan and Ferreira [24], Resnick [57] for the underlying theory. For the introduction of the sample
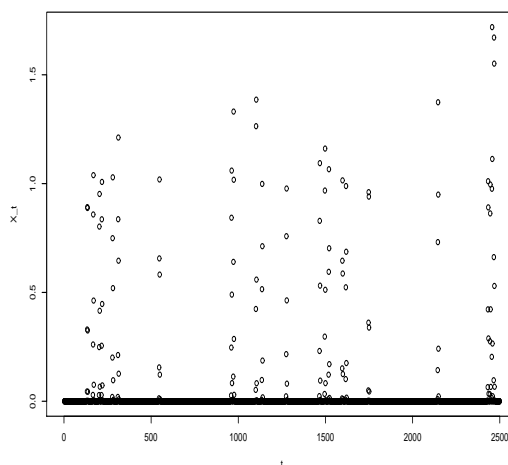
Figure 1.8. Max-stable process (1.26) where f is the standard normal density and $\alpha = 5$. Extremal clusters are clearly visible.

extremogram, consider an $\mathbb{R}^d$-valued strictly stationary regularly varying process (X_t) and a Borel set $C \subset \overline{\mathbb{R}}_0^d$ bounded away from zero. Then, for any sequence $m = m_n \to \infty$ with $m_n/n \to 0$ as $n \to \infty$, we define the following estimator of $P_m(C) = m P(a_m^{-1} X \in C)$:

$$\widehat{P}_m(C) = \frac{m}{n} \sum_{t=1}^{n} I_{\{a_m^{-1} X_t \in C\}}$$

A possible choice of (a_m) is given by $P(|X| > a_m) \sim m^{-1}$. By definition of regular variation of X, for any μ_1-continuity set C,

$$E[\widehat{P}_m(C)] = m P_m(C) \to \mu_1(C) \,.$$

Here the condition $m_n \to \infty$ as $n \to \infty$ was crucial for asymptotic unbiasedness. For the calculation of the asymptotic variance of $\widehat{P}_m(C)$ we assume the following condition:

(M) The sequence (X_t) is α-mixing with rate function (α_h) and there exists a sequence $r_n \to \infty$ such that $r_n/m_n \to 0$ as $n \to \infty$,

(2.1) $$\lim_{n \to \infty} m_n \sum_{h=r_n}^{\infty} \alpha_h = 0$$

and for every $\epsilon > 0$,

(2.2) $$\lim_{k \to \infty} \limsup_{n \to \infty} m_n \sum_{h=k}^{r_n} P(|X_h| > \epsilon \, a_m \,, |X_0| > \epsilon \, a_m) = 0 \,.$$

This condition is technical: (2.1) imposes some rate on the mixing function (α_h) and (2.2) avoids "extremal long range dependence"; (2.2) is an asymptotic independence condition in the spirit of the classical condition D'; see Leadbetter et al. [40], Embrechts et al. [20]. The quantities m_n and r_n have some straightforward interpretation as size in large-small block scheme: the sample $X_1, \ldots, X_n$ consists of roughly $[n/m_n]$ large disjoint blocks of size m_n. After chopping off the first

r_n elements in each large block one aims at ensuring the asymptotic independence of the resulting large blocks.

If C is a μ_1-continuity set and $C \times \overline{\mathbb{R}}_0^{d(h-1)} \times C$ are μ_{h+1}-continuity sets for every $h \geq 1$, regular variation of X implies

$$(2.3) \qquad \qquad \mathrm{var}[\widehat{P}_m(C)] \sim \frac{m}{n} V(C) \,,$$

where

$$V(C) = \mu_1(C) + 2 \sum_{h=1}^{\infty} \tau_h(C) \quad \text{and} \quad \tau_h(C) = \mu_{h+1}(C \times \overline{\mathbb{R}}_0^{d(h-1)} \times C) \,, \quad h \geq 1 \,,$$

and we also assume that the infinite series is finite. The asymptotic relation (2.3) indicates that the condition $m_n/n \to 0$ is needed to ensure the consistency of the estimator $\widehat{P}_m(C)$. Under additional conditions, $(\widehat{P}_m(C))$ is asymptotically normal and this property also holds jointly for finitely many sets $C_1, \ldots, C_h$. The complicated form of the asymptotic variance in (2.3) suggests that it is difficult to apply this central limit theorem for constructing asymptotic confidence bands.

The motivating examples of extremograms are limits of conditional probabilities

$$\rho_{AB}(h) = \lim_{x \to \infty} P(x^{-1} X_h \in B \mid x^{-1} X_0 \in A) \,, \quad h \geq 0 \,,$$

for sets A, B bounded away from zero. In Section 1.2 we mentioned the close relationship of ρ_{AB} with a cross-correlation function. Replacing the probabilities in these conditional probabilities by estimators of the type $\widehat{P}_m(C)$ and applying the corresponding central limit theory from [16], Section 3, and the continuous mapping theorem, one obtains an asymptotic theory for the *ratio estimators*

$$\widehat{\rho}_{AB}(h) = \frac{\sum_{t=1}^{n-h} I_{\{a_m^{-1} X_t \in A, a_m^{-1} X_{t+h} \in B\}}}{\sum_{t=1}^{n} I_{\{a_m^{-1} X_t \in A\}}} \,, \quad h \geq 0 \,.$$

The latter estimators only depend on the high threshold a_m which one typically chooses as a high empirical quantile of the data. These estimators can be interpreted as a sample cross-correlation function.

We recall a central limit theorem for these estimators; see Corollary 3.4 in [16] and its correction Theorem 4.3 in [18].

Theorem 2.1. *Let (X_t) be an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence with index $\alpha > 0$. Assume that the following conditions are satisfied.*

- *The Borel sets $A, B \subset \overline{\mathbb{R}}_0^d$ are bounded away from zero and $\mu_1(A) > 0$.*
- *The sets A, B are continuous with respect to μ_1.*
- *Condition* (M), $(n/m_n)\alpha_{r_n} \to 0$ *as* $n \to \infty$.
- $m_n = o(n^{1/3})$ *or*

$$(2.4) \qquad \frac{m_n^4}{n} \sum_{j=r_n}^{m_n} \alpha_j \to 0 \quad \text{and} \quad \frac{m_n r_n^3}{n} \to 0 \quad \text{as } n \to \infty.$$

Then the following central limit theorem holds for $h \geq 0$

$$(2.5) \qquad \sqrt{\frac{n}{m_n}} \left[\widehat{\rho}_{AB}(h) - \rho_{AB,m}(h) \right]_{h=0,\ldots,m} \xrightarrow{d} N(\mathbf{0}, (\mu_1(A))^{-4} \mathbf{\Sigma}) \,.$$

for some matrix $\mathbf{\Sigma}$,[6] where $\rho_{AB,m}(h) = P(a_m^{-1} X_h \in B \mid a_m^{-1} X_0 \in A)$.

Some comments are here in place.

---

[6]This matrix is complicated and irrelevant for our purposes; see [16], (3.15) and (3.16) for its value.

- The conditions of this result are rather technical but the mixing and anti-clustering conditions can be verified for standard time series models. For example, if (X_t) is α-mixing with geometric rate, then one can simply choose sequences $r_n = [C \log n]$ for a large constant $C > 0$ or $r_n = n^\epsilon$, and $m_n = n^{2\epsilon}$ for suitable small $\epsilon > 0$.

- The asymptotic variance is not of practical use. Therefore Davis et al. [17] suggest an alternative way of constructing confidence bands for $\widehat{\rho}_{AB}(h)$, by using the stationary bootstrap introduced by Politis and Romano [54].

- Central limit theory and bootstrap consistency for the sample extremogram do not follow from standard results for sequences of mixing stationary sequences. This is due to the fact that we deal with sequences of indicator functions $(I_{\{a_m^{-1}Y_t \in C\}})$ for certain strictly stationary sequences (Y_t) and sets C bounded away from zero. The sequences $(I_{\{a_m^{-1}Y_t \in C\}})$ constitute a *triangular array of row-wise strictly stationary sequences* for which, to the best of our knowledge, standard asymptotic theory is not available.

- The convergence rate $\sqrt{n/m}$ in the central limit theorem (2.5) is due to the triangular array structure; it can be significantly slower than standard $\sqrt{n}$-rates.

- We call $\rho_{AB,m}$ in (2.5) the *pre-asymptotic extremogram* since, in general, one cannot replace the centering constants $\rho_{AB,m}(h)$ by their limits $\rho_{AB}(h)$; see Example 2.2 below. Moreover, it is in general very difficult to show that

$$(2.6) \qquad \sqrt{\frac{n}{m_n}}|\rho_{AB,m}(h) - \rho_{AB}(h)| \to 0\,, \quad \text{as } n \to \infty,$$

  even for "nice" models such as the GARCH$(1,1)$. For this well studied model one lacks precise information about the tail behavior. The central limit theorem (2.5) (and related bootstrap procedures) are then used to approximate the conditional probabilities $\rho_{AB,m}(h)$. These have a very concrete interpretation in contrast to their less intuitive limits $\rho_{AB}(h)$.

- If (2.6) holds with a rather slow convergence rate one faces a bias problem. This problem can be observed e.g. for a simulated stochastic volatility process (X_t) and $A = B = (1, \infty)$. Then $\rho(h) = \rho_{AA}(h) = 0$ for $h \geq 1$. If (X_t) is α-mixing with geometric rate it can be verified that (2.6) and (2.4) hold for $m_n = n^\gamma$, $\gamma \in (1/3, 1)$, and then (2.5) applies with $\rho_{AA,m}(h)$ replaced by $\rho(h) = 0$; see [16], Section 4.2. Also notice that $\widehat{\rho}_{AA}(h)$ is of the order $1/m$ in the iid case and hence greater than zero.

- The formulation of the results in [16, 17, 18] related to Theorem 2.1 involve various other continuity conditions on sets. These conditions can be avoided as a close inspection of the proofs shows: these conditions follow from continuity of A and B with respect to μ_1. Indeed, one needs that sets of the form $\bigotimes_{i=1}^k C_i$ are μ_k-continuity sets, where $C_i \in \{A, B, \overline{\mathbb{R}}_0^d, A \cap B\}$, $i = 1, \ldots, k$, and at least one of the sets C_i does not coincide with $\overline{\mathbb{R}}_0^d$. Let S be the set of indices i such that C_i does not coincide with $\overline{\mathbb{R}}_0^d$. Then

$$\partial\Big(\bigotimes_{i=1}^k C_i\Big) \subset \bigcup_{i \in S} (\overline{\mathbb{R}}_0^d)^{i-1} \times \partial C_i \times (\overline{\mathbb{R}}_0^d)^{k-i-1}\,.$$

  The sets in the union have μ_k-measure zero. For the sake of argument assume that $i = 1 \in S$. Then

$$\lim_{n \to \infty} n\, P(a_n^{-1}(X_1, \ldots, X_k) \in \partial C_1 \times (\overline{\mathbb{R}}_0^d)^{k-1}) = \lim_{n \to \infty} n\, P(a_n^{-1}X_1 \in \partial C_1)$$
$$= \mu_1(\partial C_1) = 0\,.$$

- In applications one needs to choose the threshold a_m in some reasonable way. The choice of a threshold is inherent to extreme value statistics to which no easy solution exists. In [16] we

advocated to choose a_m as a fixed high/low empirical quantile of the absolute values of the data and to experiment with several quantile values. If the plot of the sample extremogram is robust for a range of such quantiles one can choose a quantile from that region. The theory in [16, 17, 51] is based on *deterministic* values a_m. The heuristic method described above advocates the choice of a *data dependent* threshold. Recent work by Kulik and Soulier [38] yields a theory for a modified sample extremogram with data dependent threshold in the case of short and long memory stochastic volatility processes.

Example 2.2. In [16] we did not provide a concrete example of a sequence (X_t) for which (2.6) does not hold. Such counterexamples can easily be constructed from max-stable strictly stationary processes with Fréchet marginals; see Section 1.2.4. We assume the conditions of Proposition 1.3. For simplicity choose $\alpha = 1$, $A = B = (1, \infty)$ and $a_m = m \int_E f_0(x) \nu(dx)$. The function ρ is given in (1.22). By Taylor expansion,

$$
\begin{aligned}
\rho_{AA,m}(h) &= \frac{P(\min(X_0, X_h) > a_m)}{P(X > a_m)} = \frac{1 - \mathrm{e}^{-a_m^{-1} \int_E f_0(x) \wedge f_h(x)\, \nu(dx)}}{1 - \mathrm{e}^{-a_m^{-1} \int_E f_0(x)\, \nu(dx)}} \\
&= \rho(h) - m^{-1} c_h \left(1 + o(1)\right),
\end{aligned}
$$

for some constant $c_h \neq 0$. Hence

$$
(n/m_n)^{0.5} |\rho_{AA,m}(h) - \rho(h)| \sim c_h \, (n/m^3)^{0.5},
$$

and the right-hand side converges to zero if and only if $n^{1/3} = o(m)$ and the rate of convergence to zero can be arbitrarily small. The latter condition is, of course, in contradiction with $m = o(n^{1/3})$ which is one possible sufficient condition for (2.5). Fortunately, the other sufficient condition (2.4) can still be satisfied if $m = o(n^{1/3})$ does not hold. For example, if α_h decays to zero at a geometric rate and one chooses $r_n = [C \log n]$ for some $C > 0$ and $m_n = n^\gamma$ for some $\gamma \in (1/3, 1)$. Particular cases with geometric decay of (α_h) were mentioned in Examples 1.6 and 1.7.

2.2. **Cross-extremogram for bivariate time series.** While the definition of the extremogram covers the case of multivariate time series, it is of limited value if the index of regular variation is not the same across the component series. For example, consider two regularly varying univariate strictly stationary time series (X_t) and (Y_t) with tail indices $\alpha_X < \alpha_Y$. Then, assuming $((X_t, Y_t))_{t \in \mathbb{Z}}$ stationary, this bivariate time series would be regularly varying with index α_X, and for Borel sets A, B bounded away from zero, $\widetilde{A} = A \times \mathbb{R}$ and $\widetilde{B} = \mathbb{R} \times B$,

$$
\rho_{\widetilde{A}\widetilde{B}}(h) = \lim_{x \to \infty} P(Y_h \in xB \mid X_0 \in xA) = \lim_{x \to \infty} P((X_h, Y_h) \in x\widetilde{B} \mid (X_0, Y_0) \in x\widetilde{A}) = 0, \quad h \in \mathbb{Z}.
$$

The asymptotic theory of Section 2.1 is applicable to the sets $\widetilde{A} = A \times \mathbb{R}$ and $\widetilde{B} = \mathbb{R} \times B$. In this case, no extremal dependence between the two series would be measured. To avoid these rather uninteresting cases and obtain a more meaningful measure of extremal dependence, we transform the two series so that they have the same marginals. In extreme value theory, the transformation to the unit Fréchet distribution is standard. For the sake of argument, assume that both X_t and Y_t are positive so that the focus of attention will be on extremal dependence in the upper tails. The case of extremal dependence in the lower tails or upper and lower tails is similar. Under the positivity constraint, if F_1 and F_2 denote the distribution functions of X_t and Y_t, respectively, and are continuous, then the two transformed series, $\widetilde{X}_t = G_1(X_t)$ and $\widetilde{Y}_t = G_2(Y_t)$ with $G_i(z) = -1/\log(F_i(z))$, $i = 1, 2$, have unit Fréchet marginals Φ_1; see (1.15). Now assuming that the bivariate time series $((\widetilde{X}_t, \widetilde{Y}_t))_{t \in \mathbb{Z}}$ is regularly varying, we define the *cross-extremogram* by

$$
\rho_{\widetilde{A}\widetilde{B}}(h) = \lim_{x \to \infty} P(\widetilde{Y}_h \in x\,B \mid \widetilde{X}_0 \in x\,A), \quad h \in \mathbb{Z},
$$

At first glance, this may seem inconvenient since transformations to unit Fréchet marginals are required. If one restricts attention to sets A and B that are intervals bounded away from 0 or finite unions of such sets the transformation simplifies: if a_n denotes the $(1 - n^{-1})$-quantile of Φ_1, then by monotonicity of G_i, $\{\widetilde{X}_h \in a_n A\} = \{X_h \in a_{X,n} A\}$ and $\{\widetilde{Y}_h \in a_n B\} = \{Y_h \in a_{Y,n} B\}$, where $a_{X,n}$ and $a_{Y,n}$ are the respective $(1 - n^{-1})$-quantiles of the distributions of X_t and Y_t. For sets A and B of the required form, the cross-extremogram becomes

$$(2.7) \qquad \rho_{\widetilde{A}\widetilde{B}}(h) \;\; = \;\; \lim_{n \to \infty} P(Y_h \in a_{Y,n} B \mid X_0 \in a_{X,n} A)\,.$$

Thus we do not actually need to find the transformations converting the data to unit Fréchet, only the component-wise quantiles, $a_{X,n}$ and $a_{Y,n}$, need to be calculated. Clearly, this notion of extremogram extends to more than two time series.

3. AN EXAMPLE: EQUITY INDICES

We consider daily log-returns equity indices of four countries: S&P 500 for US, FTSE 100 for UK, DAX for Germany, Nikkei 225 for Japan. Figure 3.1 shows the sample extremogram $\widehat{\rho} = \widehat{\rho}_{AA}$ for the negative tails ($A = B = (-\infty, -1)$ with a_m estimated as the 96% empirical quantiles of the absolute values of the negative data) applied to 6,443 log-returns of the FTSE and S&P (April 4, 1984 to October 2, 2009), to 4,848 log-returns of the DAX (November 13, 1990 to October 2, 2009) and to 6,333 log-returns of the Nikkei (August 23, 1984 to October 2, 2009).[7] The solid horizontal lines in the plots represent 98% confidence bands. They correspond to the maximum and minimum of the sample extremogram at lag 1 based on 99 random permutations of the data. If the data were independent random permutations would not change the dependence structure: values $\widehat{\rho}_{AB}(h)$ which are outside the confidence bands indicate that there is significant extremal dependence at lag h. The sample extremograms for all four indices decay rather slowly to zero, with S&P the slowest. Among the four indices, the Nikkei displays the least amount of extremal dependence as measured by the extremogram. The top graphs in Figure 3.1 indicate extremal dependence in the lower tail over a period of 40 days.

We assume that the log-return series are modeled by a GARCH$(1, 1)$ process (see Section 1.2.4 for its definition). Then we can estimate its parameters, calculate the volatility sequence $(\widehat{\sigma}_t)$ and the *filtered* sequence $\widehat{Z}_t = X_t/\widehat{\sigma}_t$, $t = 1, \dots, n$. Figure 3.2 shows the sample extremograms $\widehat{\rho}$ for the filtered FTSE and S&P sequences. These plots confirm that much of the extremal dependence (as measured by the extremogram) has been removed. Hence the extremal dependence in the log-returns is due to the volatility sequences (σ_t), as suggested by the GARCH model.

For a bivariate time series $((X_t, Y_t))_{t \in \mathbb{Z}}$ the sample cross-extremogram is given by

$$\widehat{\rho}_{\widetilde{A}\widetilde{B}}(h) = \frac{\sum_{t=1}^{n-h} I_{\{Y_{t+h} \in a_{Y,m} B, X_t \in a_{X,m} A\}}}{\sum_{t=1}^{n} I_{\{X_t \in a_{X,m} A\}}}\,,$$

where $a_{X,m}$ and $a_{Y,m}$ are the $(1 - m^{-1})$-quantiles of the marginal distributions of X and Y, respectively. For applications, they need to be replaced by the corresponding empirical quantiles.

We calculate the sample cross-extremograms for the pairs of the log-returns series, again for the negative tails, i.e. $A = B = (-\infty, -1)$, and $a_{X,m}$ and $a_{Y,m}$ are chosen as the 96% empirical quantiles of the negative values of the corresponding component samples. Since the samples have different sizes we consider those periods for which we have observations on both indices.

---

[7]As noted in the literature, the lower tails of returns tend to be heavier than the upper tails. Similar plots (not shown here) of the sample extremogram for the upper tails also reveal extremal dependence, but to a lesser extent than seen in the lower tails.
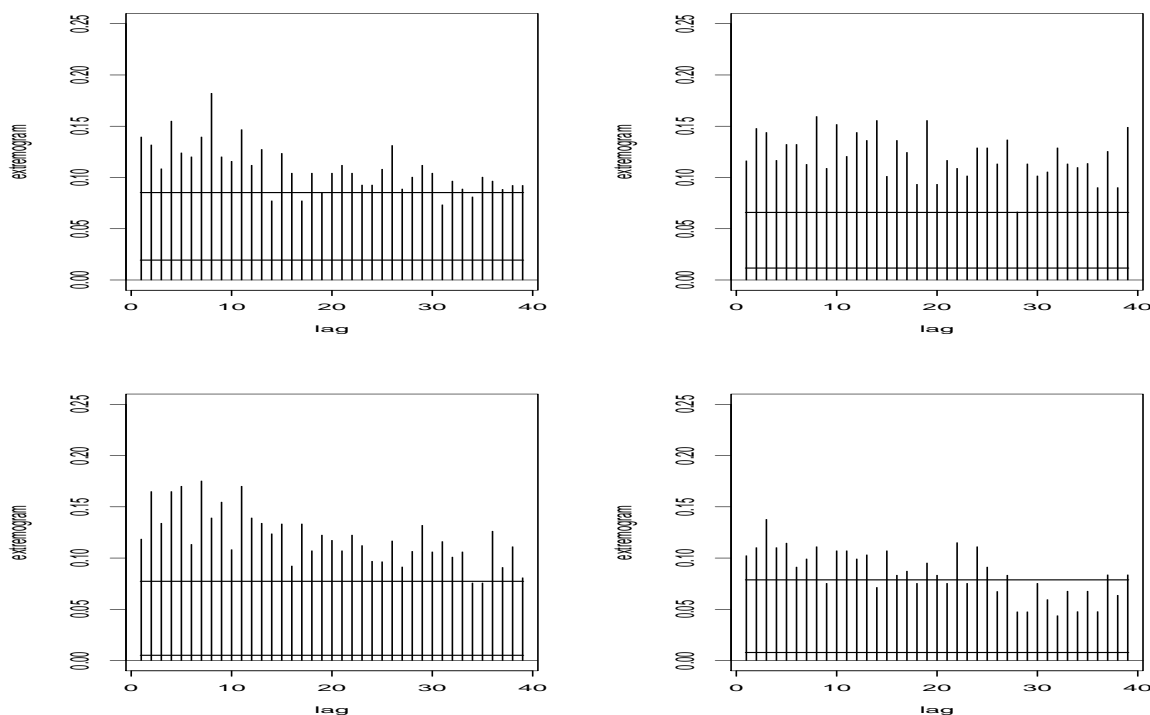
Figure 3.1. *The sample extremogram for the lower tails of the FTSE* (top left)*, S&P* (top right)*, DAX* (bottom left) *and Nikkei. The solid lines represent 98% confidence bands based on* 99 *random permutations of the data.*

The sample cross-extremogram of any pair of series exhibits a similar pattern of slow decay as seen in the univariate sample extremograms (we do not include these figures). Figure 3.3 shows the sample cross-extremograms for the filtered series. For example, in the first row of graphs, (X_t) is the filtered FTSE and (Y_t) are the filtered S&P, DAX and Nikkei, respectively. There are signs of various types of cross-extremal dependence in the filtered series. The spikes at lag zero (except between the Nikkei and S&P) indicate the strong extremal dependence of the multiplicative shocks, In the second row, there is evidence of significant extremal dependence at lag one for each sample cross-extremogram: given the S&P has an extreme left tail event in a shock at time t there will be a corresponding large left tail shock in the FTSE, the DAX and the Nikkei at time $t = 1$. Given the dominance of the US stock market, one might expect a carry-over effect of the shocks on the other exchanges on the next day. Since only marginal GARCH models were fitted to the data, it may not seem all that surprising that the filtered series exhibit serial dependence. We should note, however, that the dependence in the shocks does not appear to last beyond one time lag.

4. A FOURIER ANALYSIS OF EXTREME EVENTS

Classical time series analysis studies the second order properties of stationary processes in the time and frequency domains. The latter approach refers to spectral (or Fourier) analysis of the
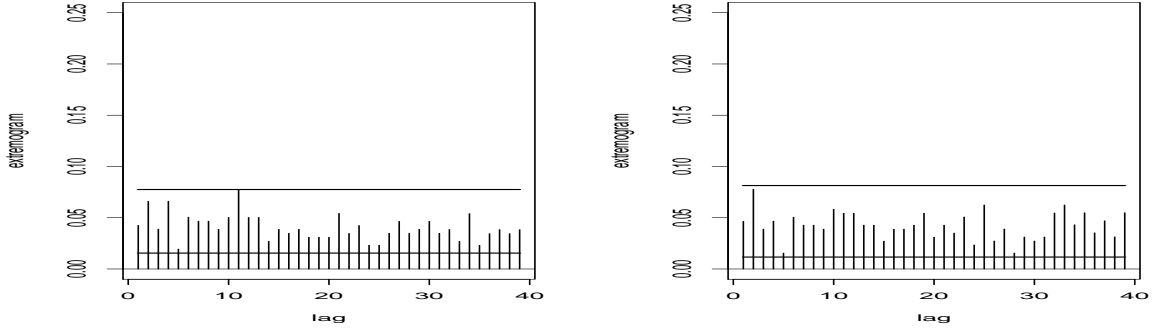
Figure 3.2. *The sample extremograms for the filtered FTSE* (left) *and filtered S&P* (right) *series. The bold lines represent 98% confidence bounds based on* 99 *random permutations of the series.*

time series. We mentioned in Section 1.2 that the extremograms $\gamma_A = \gamma_{AA}$ and $\rho_A = \rho_{AA}$ for a set A bounded away from zero are covariance and correlations functions of some stationary sequence. Then it is possible to study the corresponding spectral properties of the extremogram. Research in this direction was started in [16] and continued in [51]. We recall some of the results. Throughout we assume that (X_t) is a $\mathbb{R}^d$-valued strictly stationary regularly varying sequence with index $\alpha > 0$ and A is a μ_1-continuity set. In the comments following Theorem 2.1 we mentioned the latter property also implies that the sets $A \times (\overline{\mathbb{R}}_0^d)^h \times A$ are μ_h-continuity sets for $h \geq 0$, so the limits $\rho_A(h)$ exist.

Assuming that ρ_A is square summable, we consider the corresponding spectral density corresponding to ρ_A (see Brockwell and Davis [5], Chapter 4):

$$f_A(\lambda) = 1 + 2 \sum_{h=1}^{\infty} \rho_A(h) \, e^{-i h \lambda}, \quad \lambda \in [0, \pi].$$

A natural estimator of the spectral density is obtained if we replace the quantities $\rho_A(h)$ by the sample versions

$$\widetilde{\rho}_A(h) = \frac{\frac{m}{n} \sum_{t=1}^{n-h} (I_{a_m^{-1} X_t \in A} - p_0)(I_{a_m^{-1} X_{t+h} \in A} - p_0)}{\widehat{P}_m(A)}, \quad h \geq 1.$$

where $p_0 = P(a_m^{-1} X \in A)$:[8]

$$\widetilde{I}_{nA}(\lambda) = 1 + 2 \sum_{h=1}^{\infty} \widetilde{\rho}_A(h) \, e^{-i h \lambda} = \frac{\frac{m}{n} \left| \sum_{t=1}^{n} (I_{\{a_m^{-1} X_t \in A\}} - p_0) \, e^{i h \lambda} \right|^2}{\frac{m}{n} \sum_{t=1}^{n} I_{\{a_m^{-1} X_t \in A\}}} = \frac{I_{nA}(\lambda)}{\widehat{P}_m(A)}.$$

We will refer to I_{nA} and its standardized version $\widetilde{I}_{nA}$ as *periodogram of the extreme event* $a_m A$. Indeed, if we replaced the normalization m/n by $1/n$, I_{nA} is the periodogram of the sequence of centered indicators $(I_{\{a_m^{-1} X_t \in A\}} - p_0)$. These sequences constitute a triangular array of row-wise strictly stationary sequences for which standard asymptotic theory for the periodogram does not apply; for an asymptotic theory of the periodogram of a strictly stationary linear process, see

---

[8]The centering of the indicator functions with their expectation p_0 is crucial for deriving asymptotic theory. In applications, $I_{nA}(\lambda)$ is typically evaluated at the Fourier frequencies $\omega_j(n) = 2\pi j/n \in (0, \pi)$ and since $\sum_{h=1}^{n} e^{-i h \omega_j(n)} = 0$, centering in $I_{nA}(\omega_j(n))$ is not needed.
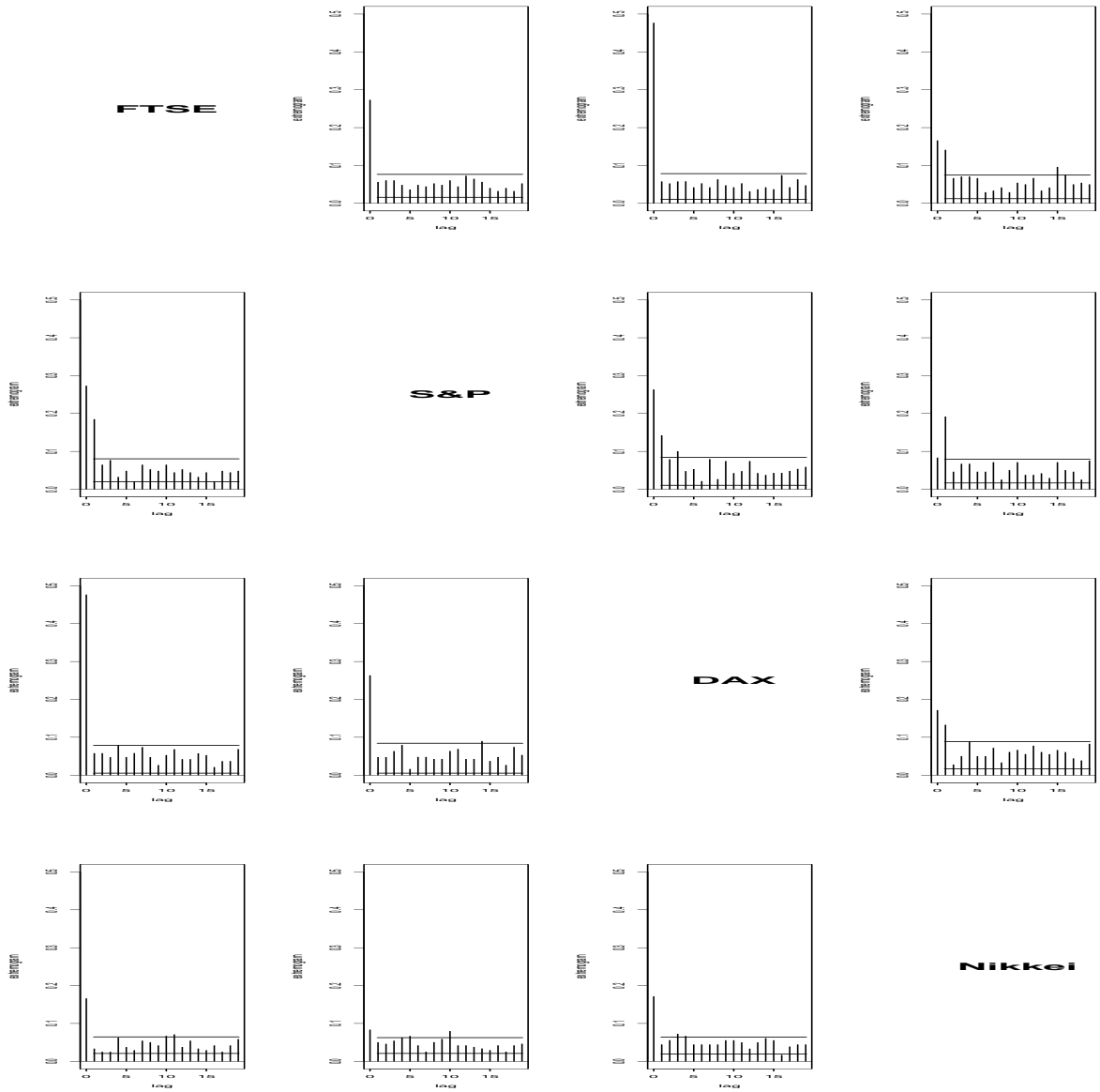
Figure 3.3. *The sample cross-extremograms for the filtered FTSE, S&P, DAX and Nikkei series. For the first row, (X_t) is the filtered FTSE and (Y_t) are the filtered S&P, DAX and Nikkei (from left to right). For the second, third and fourth rows, the X_t's are the filtered S&P, DAX and Nikkei series, respectively.*

Brockwell and Davis [5], Chapter 10. However, the periodogram of extreme events shares some of the basic properties of the periodogram, as the following results from Mikosch and Zhao [51] show.

Theorem 4.1. *Let (X_t) be an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence with index $\alpha > 0$ satisfying condition* (M), $A \subset \overline{\mathbb{R}}_0^d$ *be a μ_1-continuity set and $\sum_{h \geq 1} \rho_A(h) < \infty$.*

- *Assume $\lambda \in (0, \pi)$ is fixed and $\omega_n = 2\pi j_n/n$, $j_n \in \mathbb{Z}$, is any sequence of Fourier frequencies such that $\omega_n \to \lambda$. Then*

$$\lim_{n \to \infty} EI_{nA}(\lambda) = \lim_{n \to \infty} EI_{nA}(\omega_n) = \mu_1(A) f_A(\lambda),$$

- *Assume in addition that the sequences (m_n), (r_n) from (M) also satisfy the growth conditions $(n/m)\alpha_{r_n} \to 0$, and $m_n = o(n^{1/3})$. Let (E_i) be a sequence of iid standard exponential random variables. Consider any fixed frequencies $0 < \lambda_1 < \cdots < \lambda_N < \pi$ for some $N \geq 1$. Then the following relations hold:*

$$\left(I_{nA}(\lambda_i)\right)_{i=1,\dots,N} \quad \overset{d}{\to} \quad \mu_1(A)\left(f_A(\lambda_i)E_i\right)_{i=1,\dots,N}, \quad n \to \infty,$$

$$\left(\widetilde{I}_{nA}(\lambda_i)\right)_{i=1,\dots,N} \quad \overset{d}{\to} \quad \left(f_A(\lambda_i)E_i\right)_{i=1,\dots,N}, \quad n \to \infty.$$

Consider any distinct Fourier frequencies $\omega_i(n) \to \lambda_i \in (0, \pi)$ as $n \to \infty$, $i = 1, \dots, N$. The limits λ_i do not have to be distinct. Then the following relations hold:

$$\left(I_{nA}(\omega_i(n))\right)_{i=1,\dots,N} \quad \overset{d}{\to} \quad \mu_1(A)\left(f_A(\lambda_i)E_i\right)_{i=1,\dots,N}, \quad n \to \infty,$$

$$\left(\widetilde{I}_{nA}(\omega_i(n))\right)_{i=1,\dots,N} \quad \overset{d}{\to} \quad \left(f_A(\lambda_i)E_i\right)_{i=1,\dots,N}, \quad n \to \infty.$$

These properties are very similar to those of a strictly stationary weakly dependent sequence. The asymptotic independence of the periodogram of the extreme event $a_m A$ and consistency in the mean of $I_{nA}(\lambda)$ give raise to the hope that pointwise consistent smoothed periodogram estimation of the spectral density $f_A(\lambda)$ is possible.

For a fixed frequency $\lambda \in (0, \pi)$ define

$$\lambda_0 = \min\{2\pi j/n : 2\pi j/n \geq \lambda\}, \quad \text{and} \quad \lambda_j = \lambda_0 + 2\pi j/n, \quad |j| \leq s.$$

(We suppress the dependence of λ_j on n.) Assume that $s = s_n \to \infty$ and $s_n/n \to 0$ as $n \to \infty$. Consider the non-negative weight function $(w_n(j))_{|j| \leq s}$ satisfying the conditions

$$\sum_{|j| \leq s} w_n(j) = 1 \quad \text{and} \quad \sum_{|j| \leq s} w_n^2(j) \to 0 \text{ as } n \to \infty.$$

Introduce the corresponding *smoothed periodogram*

$$\widehat{f}_{nA}(\lambda) = \sum_{|j| \leq s_n} w_n(j) I_{nA}(\lambda_j),$$

Under the conditions of the second item in Theorem 4.1 and some further restrictions on the growth of (m_n) and (α_h) the following limit relations hold for a fixed frequency $\lambda \in (0, \pi)$,

$$\widehat{f}_{nA}(\lambda) \overset{L^2}{\to} \mu_1(A) f_A(\lambda) \quad \text{and} \quad \widehat{f}_{nA}(\lambda)/\widehat{P}_m(A) \overset{P}{\to} f_A(\lambda).$$

In Figure 4 we illustrate how the smoothed periodogram $\widehat{f}_{n,(-\infty,-1)}$ works for 5-minute log-returns of Bank of America stock prices. We choose a simple Daniell window with $w_n(j) = 1/(2s_n + 1)$ and $s_n = 52$.

References

[1] BASRAK, B., DAVIS, R.A. AND MIKOSCH. T. (2002) Regular variation of GARCH processes. *Stoch. Proc. Appl.* **99**, 95–116.

[2] BASRAK, B. AND SEGERS, J. (2009) Regularly varying multivariate time series. *Stoch. Proc. Appl.* **119**, 1055–1080.

[3] BINGHAM, N.H., GOLDIE, C.M. AND TEUGELS, J.L. (1987) *Regular Variation.* Cambridge University Press, Cambridge.

[4] BOLLERSLEV, T. (1986) Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**, 307–327.
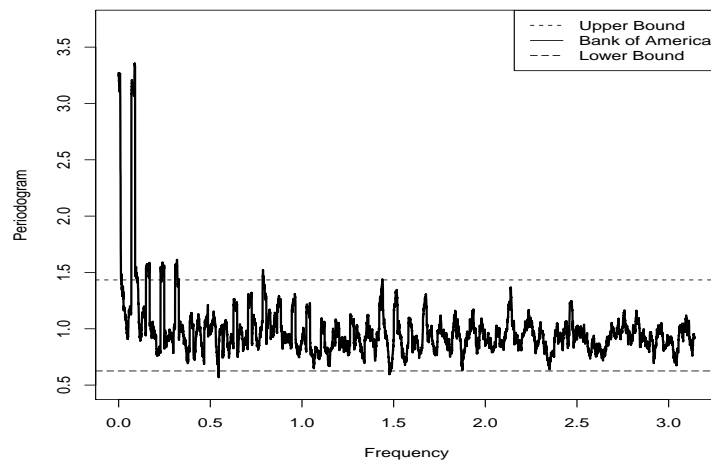
Figure 4.2. The smoothed periodogram (corresponding to the losses) for $31,757$ 5-minute log-returns of Bank of America stock prices with Daniell window, $s_n = 52$. The simultaneous confidence bands are constructed by taking the 97.5% quantile of the maxima and the 2.5% quantile of the minima over the Fourier frequencies calculated from the smoothed periodograms of 10 000 random permutations of the data. If the data were iid, permutations would not change the dependence structure. The fact that the periodogram is outside the confidence bands at various frequencies indicates that there is significant extremal dependence in the data. The peaks at various frequencies show that there are cycles of extremal behavior in the data. These cycles cannot be detected by autocorrelation plots of the data, their absolute values or squares.

[5] BROCKWELL. P. AND DAVIS., R.A. (1991) *Time Series: Theory and Methods.* 2nd Edition. Springer, New York.
[6] BROWN, B. AND RESNICK, S.I. (1977) Extreme values of independent stochastic processes. *J. Appl. Probab.* **14**, 732–739.
[7] BURACZEWSKI, D., DAMEK, E., GUIVARC'H, Y., HULANICKI, A. AND URBAN, R. (2009) Tail-homogeneity of stationary measures for some multidimensional stochastic recursions. *Probab. Th. Rel. Fields* **145**, 385–420.
[8] DAVIS, R.A. AND HSING, T. (1995) Point process and partial sum convergence for weakly dependent random variables with infinite variance. *Ann. Prob.* **23**, 879–917.
[9] DAVIS, R.A., KLÜPPELBERG, C. AND STEINKOHL, C. (2012) Statistical inference for max-stable processes in space and time. *J. Royal Statist. Soc., Series B*, to appear.
[10] DAVIS, R.A. AND MIKOSCH, T. (1998) Limit theory for the sample ACF of stationary process with heavy tails with applications to ARCH. *Ann. Statist.* **26**, 2049–2080.
[11] DAVIS, R.A. AND MIKOSCH, T. (2001) Point process convergence of stochastic volatility processes with application to sample autocorrelation. *J. Appl. Probab* **38A**, 93–104.
[12] DAVIS, R.A. AND MIKOSCH, T. (2008) Extreme value theory for space-time processes with heavy-tailed distributions. *Stoch. Proc. Appl.* **118**, 560–584.
[13] DAVIS, R.A. AND MIKOSCH, T. (2009) Extreme value theory for GARCH processes. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *Handbook of Financial Time Series.* Springer, pp. 187–200.
[14] DAVIS, R.A. AND MIKOSCH, T. (2009) Extremes of stochastic volatility models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *Handbook of Financial Time Series.* Springer (2009), pp. 355–364.
[15] DAVIS, R.A. AND MIKOSCH, T. (2009) Probabilistic properties of stochastic volatility models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (Eds.) *Handbook of Financial Time Series.* Springer (2009), pp. 255–268.
[16] DAVIS, R.A. AND MIKOSCH, T. (2009) The extremogram: a correlogram for extreme events. *Bernoulli* **15**, 977–1009.

[17] Davis, R.A., Mikosch, T. and Cribben, I. (2012) Towards estimating extremal serial dependence via the boostrapped extremogram. *J. Econometrics* **170**, 142–152.

[18] Davis, R.A., Mikosch, T. and Cribben, I. (2012) Estimating extremal dependence in univariate and mutivariate time series via the extremogram. arXiv:1107.5592v1 [stat.ME]

[19] Dombry, C. and Eyi-Minko, F. (2012) Strong mixing properties of max-infinitely divisible random fields. *Stoch. Proc. Appl.* **122**, 3790–3811.

[20] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[21] Engle, R.F. (1982) Autoregressive conditonal hetereoscedastic models with estimates of the variance of United Kingdom inflation. *Econometrica* **50**, 987–1007.

[22] Goldie, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1**, 126–166.

[23] Haan, L. de (1984) A spectral representation for max-stable processes. *Ann. Probab.* **12**, 1194–1204.

[24] Haan, L. de and Ferreira, A. (2006) *Extreme Value Theory. An Introduction.* Springer, New York.

[25] Haan, L. de and Pereira, T.T. (2006) Spatial extremes: models for the stationary case. *Ann. Statist.* **34**, 146–168.

[26] Haan, L. de and Tao, L. (2003) Weak consistency of extreme value estimators in $C[0,1]$. *Ann. Statist.* **31**, 1996–2012.

[27] Hsing, T. (1988) On the extreme order statistics for a stationary sequence. *Stoch. Proc. Appl.* **29**, 155–169.

[28] Hsing, T. (1991) Estimating the parameters of rare events. *Stoch. Proc. Appl.* **37**, 117–139.

[29] Hsing, T. (1991) On tail index estimation using dependent data. *Ann. Statist.* **19**, 1547–1569.

[30] Hsing, T. (1993) On some estimates based on sample behavior near high level excursions. *Probab. Th. Rel. Fields* **95**, 331-356.

[31] Hsing, T., Hüsler, J. and Leadbetter, M.R. (1988) On the exceedance point process for a stationary process. *Probab. Th. Rel. Fields* **78**, 97–112.

[32] Hult, H. and Lindskog, F. (2006) Regular variation for measures on metric spaces. *Publ. Inst. Math. (Beograd) (N.S.)* **80(94)**, 121-???140.

[33] Hult, H., Lindskog, F., Mikosch, T. and Samorodnitsky, G. (2005) Functional large deviations for multivariate regularly varying random walks. *Ann. Appl. Probab.* **15**, 2651–2680..

[34] Kabluchko, Z. (2009) Spectral representations of sum- and max-stable processes. *Extremes* **12**, 401–424.

[35] Kabluchko, Z., Schlather, M. and Haan, L. de (2009) Stationary max-stable fields associated to negative definite functions. *Ann. Probab.* **37**, 2042–2065.

[36] Kallenberg, O. (1983) *Random Measures,* 3rd edition. Akademie–Verlag, Berlin.

[37] Kesten, H. (1973) Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.

[38] Kulik, R. and Soulier, P. (2011) The tail empirical process for long memory stochastic volatility sequences. *Stoch. Proc. Appl.* **121**, 109–134.

[39] Leadbetter, M.R. (1983) Extremes and local dependence of stationary sequences. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **65**, 291-306.

[40] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes.* Springer, Berlin.

[41] Leadbetter, M.R. and Rootzén, H. (1988) Extremal theory for stochastic processes. *Ann. Probab.* **16**, 431–478.

[42] McNeil, A., Frey, R. and Embrechts, P. (2005) *Quantitative Risk Management: Concepts, Techniques, and Tools.* Princeton University Press, Princeton.

[43] Meinguet, T. and Segers, J. (2010) Regularly varying time series in Banach spaces. Technical report, see arXiv:1001.3262v1.

[44] Mikosch, T., Pawlas, Z. and Samorodnitsky, G. (2011) Large deviations for Minkowski sums of heavy-tailed generally non-convex random compact sets. *Vestnik St. Petersburg Univ., Series Mathematics* 70–78.

[45] Mikosch, T., Pawlas, Z. and Samorodnitsky, G. (2011) A large deviation principle for Minkowski sums of heavy-tailed random compact convex sets with finite expectation. Special volume of *J. Appl. Probab. New Frontiers in Applied Probability - Festschrift in Honour of Søren Asmussen* **48A**, 133–144.

[46] Mikosch, T. and Rezapur, M. (2012) Stochastic volatility models with possible extremal clustering. *Bernouli*, to appear.

[47] Mikosch, T. and Samorodnitsky, G. (2000) The supremum of a negative drift random walk with dependent heavy-tailed steps. *Ann. Appl. Probab.* **10**, 1025–1064.

[48] Mikosch, T. and Stărică, C. (2000) Limit theory for the sample autocorrelations and extremes of a GARCH(1,1) process. *Ann. Statist.* **28**, 1427–1451.

[49] Mikosch, T. and Wintenberger, O. (2012) Precise large deviations for dependent regularly varying sequences. *Probab. Rel. Fields*, to appear.

[50] Mikosch, T. and Wintenberger, O. (2012) The cluster index of regularly varying sequences with applications to limit theory for functions of multivariate Markov chains. Technical report.

[51] Mikosch, T. and Zhao, Y. (2012) A Fourier analysis of extreme events. Technical report.

[52] Oesting, M., Kabluchko, Z. and Schlather, M. (2012) Simulation of Brown-Resnick processes. *Extremes* **15**, 89–107.

[53] Perfekt, R. (1994) Extremal behaviour of stationary Markov chains with applications. *Adv. Appl. Probab.* **4**, 529–548.

[54] Politis, D.N. and Romano, J.P. (1994) The stationary bootstrap. *J. Amer. Statist. Assoc.* **89**, 1303–1313.

[55] Resnick, S.I. (1986) Point processes, regular variation and weak convergence. *Adv. Appl. Prob.* **18**, 66–138.

[56] Resnick, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[57] Resnick, S.I. (2007) *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling.* Springer, New York.

[58] Robert, C.Y., Segers, J. and Ferro, C.A.T. (2009) A sliding blocks estimator for the extremal index. *Electron. J. Stat.* **3**, 993-???1020.

[59] Robert, C.Y. (2009a) Asymptotic distributions for the intervals estimators of the extremal index and the cluster-size probabilities. *J. Statist. Plann. Inference* **139**, 3288-???3309.

[60] Robert, C.Y. (2009b) Inference for the limiting cluster size distribution of extreme values. *Ann. Statist.* **37**, 271-???310.

[61] Rosiński, J. (1995) On the structure of stationary stable processes. *Ann. Probab.* **23**, 1163–1187.

[62] Rvaceva, E.L. (1962) On domains of attraction of multi-dimensional distributions. *Select. Transl. Math. Statist. and Probability of the AMS* **2**. 183–205.

[63] Samorodnitsky, G. (2004) Extreme value theory, ergodic theory and the boundary between short memory and long memory for stationary stable processes. *Ann. Probab.* **32**, 1438–1468.

[64] Samorodnitsky, G. and Taqqu, M.S. (1994) *Stable Non-Gaussian Random Processes.* Chapman & Hall, New York.

[65] Smith, R.L. and Weissman, I. (1994) Estimating the extremal index. *J. Roy. Statist. Soc. Ser. B* **56**, 515–528.

[66] Stoev, S. (2008) One the ergodicity and maxing of max-stable processes. *Stoch. Proc. Appl.* **118**, 1679–1705.

[67] Stoev, S. and Taqqu, M.S. Extremal stochastic integrals: a parallel between max-stable processes and α-stable processes. *Extremes* **8**, 237–266.

Richard A. Davis, Department of Statistics, Columbia University, 1255 Amsterdam Ave. New York, NY 10027, U.S.A.
  E-mail address: rdavis@stat.columbia.edu

Thomas Mikosch, University of Copenhagen, Department of Mathematics, Universitetsparken 5, DK-2100 Copenhagen, Denmark
  E-mail address: mikosch@math.ku.dk

Yuwei Zhao, University of Copenhagen, Department of Mathematics, Universitetsparken 5, DK-2100 Copenhagen, Denmark
  E-mail address: zywmar@gmail.com