Heavy-tail phenomena

Spatio-temporal extremal dependence

[1]

Thomas Mikosch

University of Copenhagen

www.math.ku.dk/∼mikosch/

[1]Leukerbad, September 2014

Contents

1. Heavy tails in real-life data

1.1. Finance.

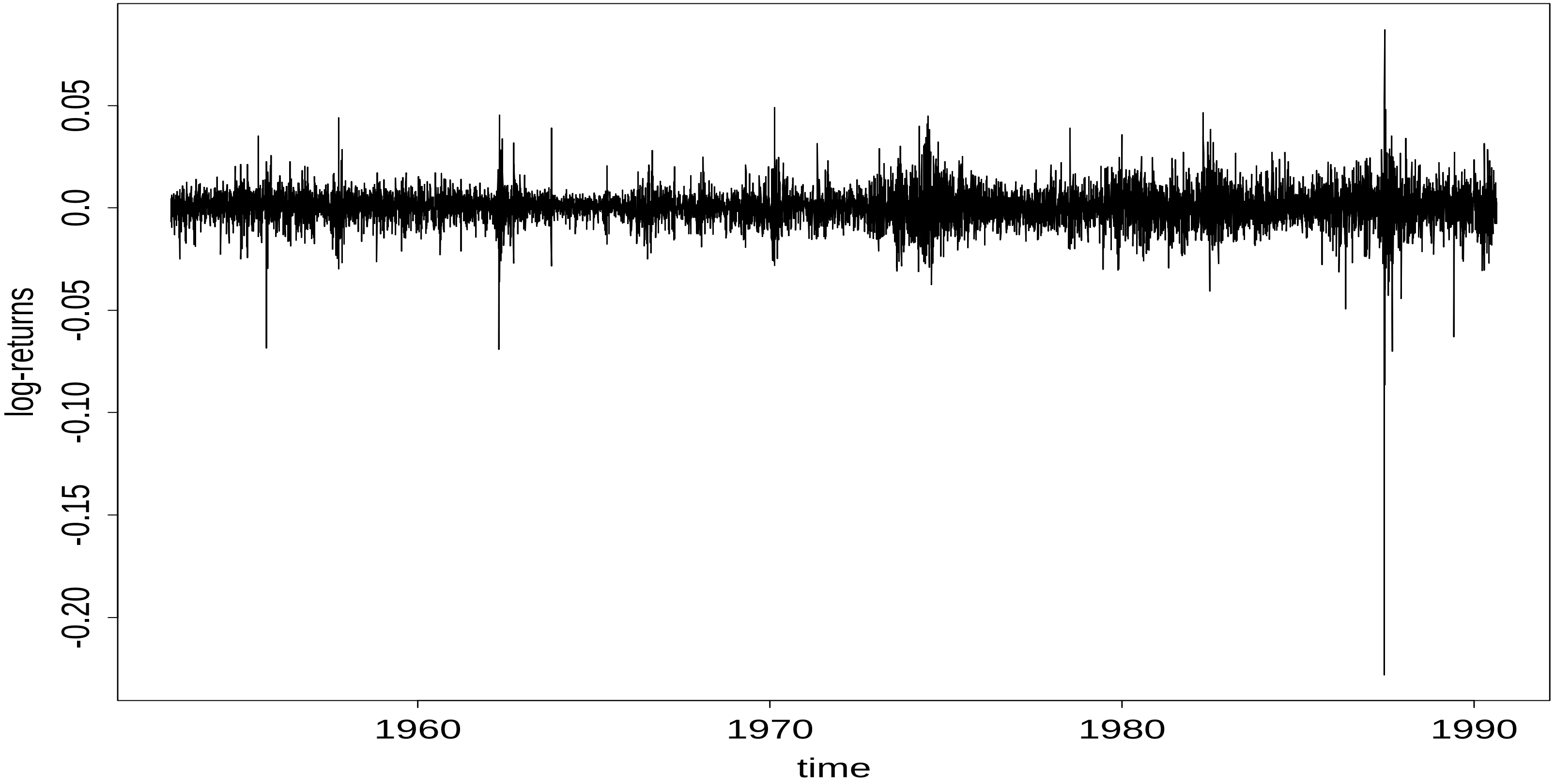

Figure 1. Plot of **9558** *S&P500* daily log-returns from January *2, 1953*, to December *31, 1990*. The year marks indicate the beginning of the calendar year.

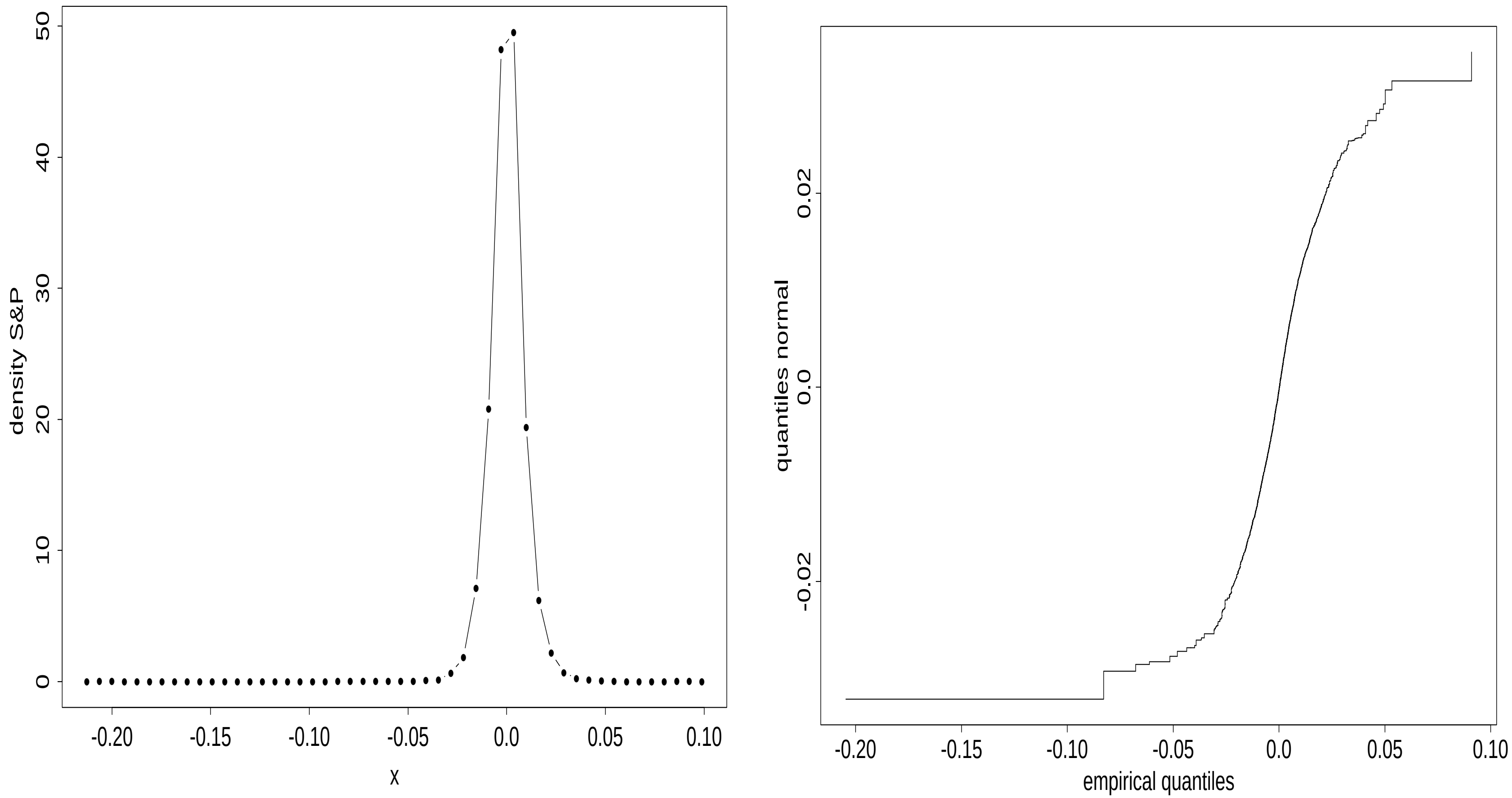

FIGURE 2. Left: Density plot of the *S&P500* data. The limits on the $\boldsymbol{x}$-axis indicate the range of the data. QQ-plot of the *S&P500* data against the normal distribution.

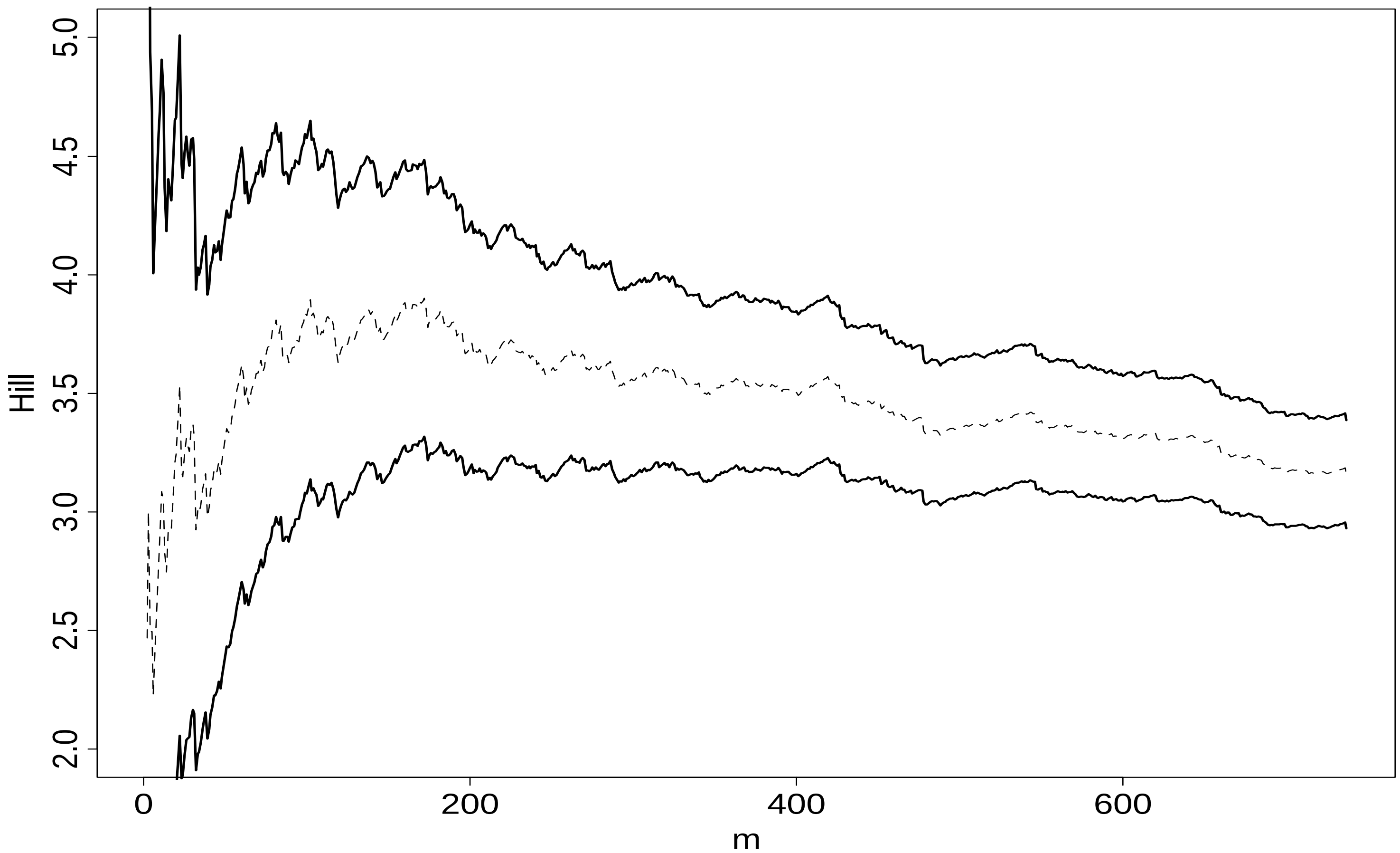

FIGURE 3. Hill plot (dotted line) for the *S&P500 data* with **95%** asymptotic confidence bounds. The Hill estimator approximates the tail index $\boldsymbol{\alpha}$ in the model $\mathbb{P}(\boldsymbol{X_1} > \boldsymbol{x}) \sim \boldsymbol{c}\,\boldsymbol{x}^{-\boldsymbol{\alpha}}$ as a function of the $\boldsymbol{m}$ upper order statistics in the return sample.

1.2. **Insurance.**

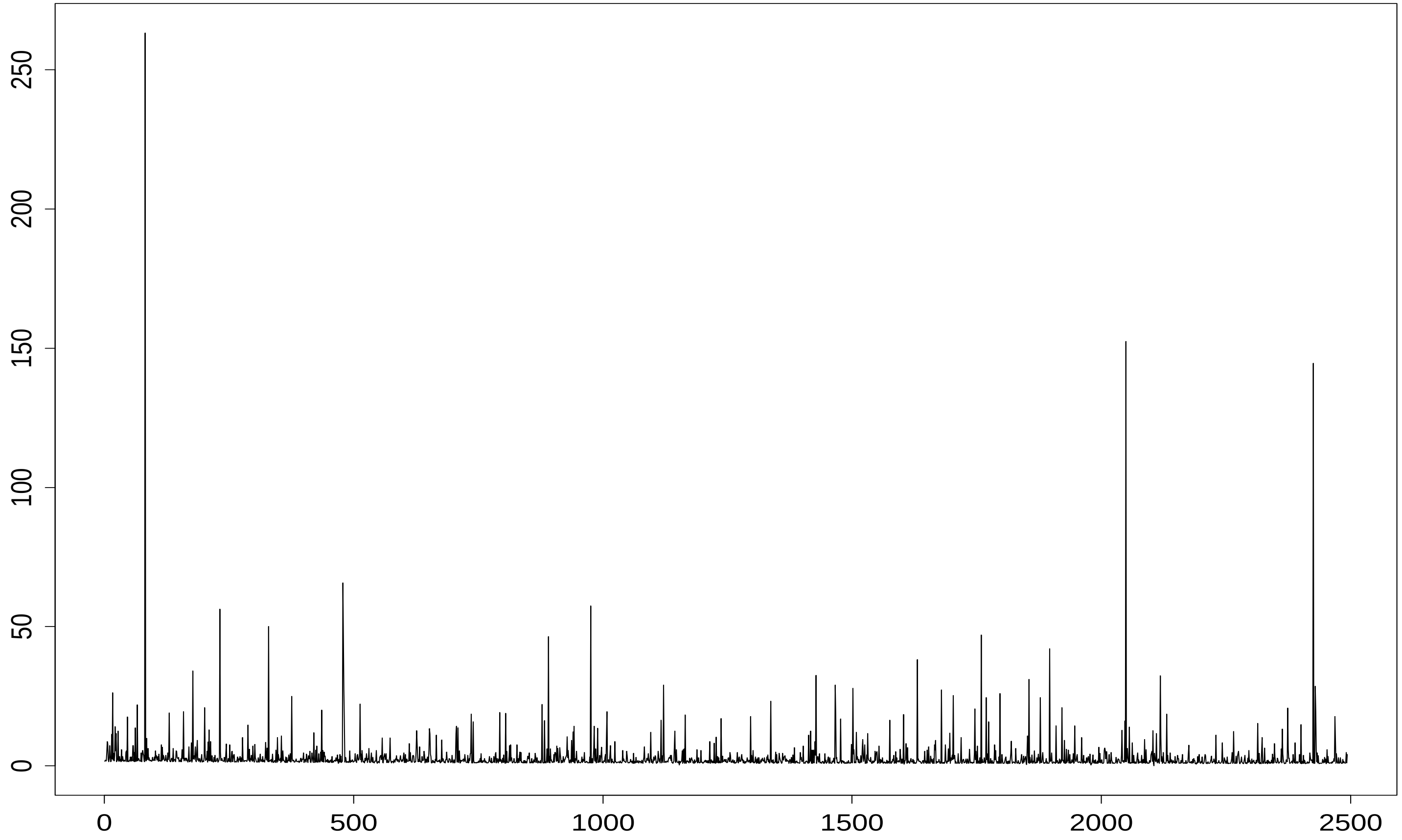

FIGURE 4. Danish fire insurance data.

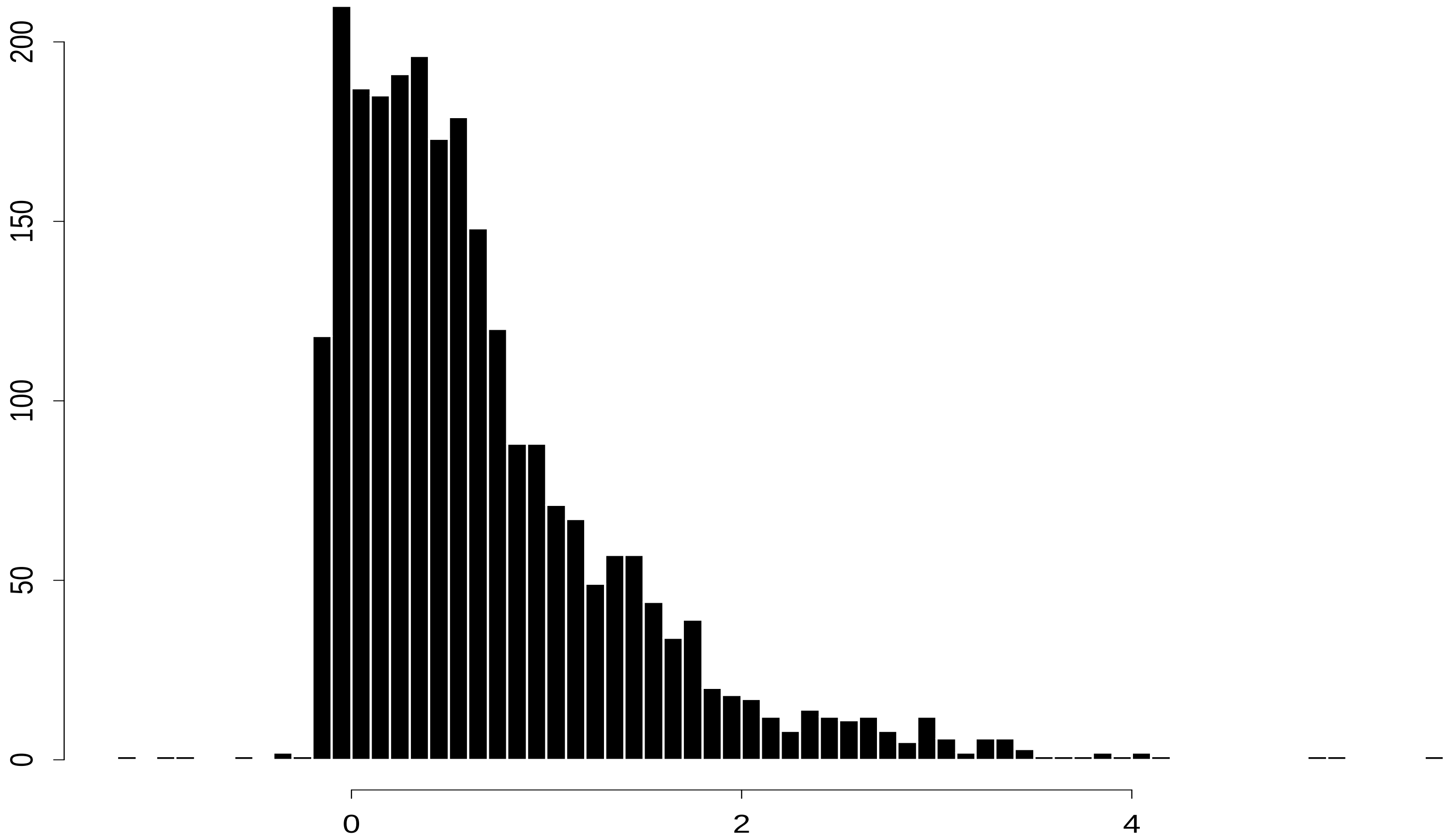

Figure 5. Histogram of the logarithmic Danish fire insurance data.

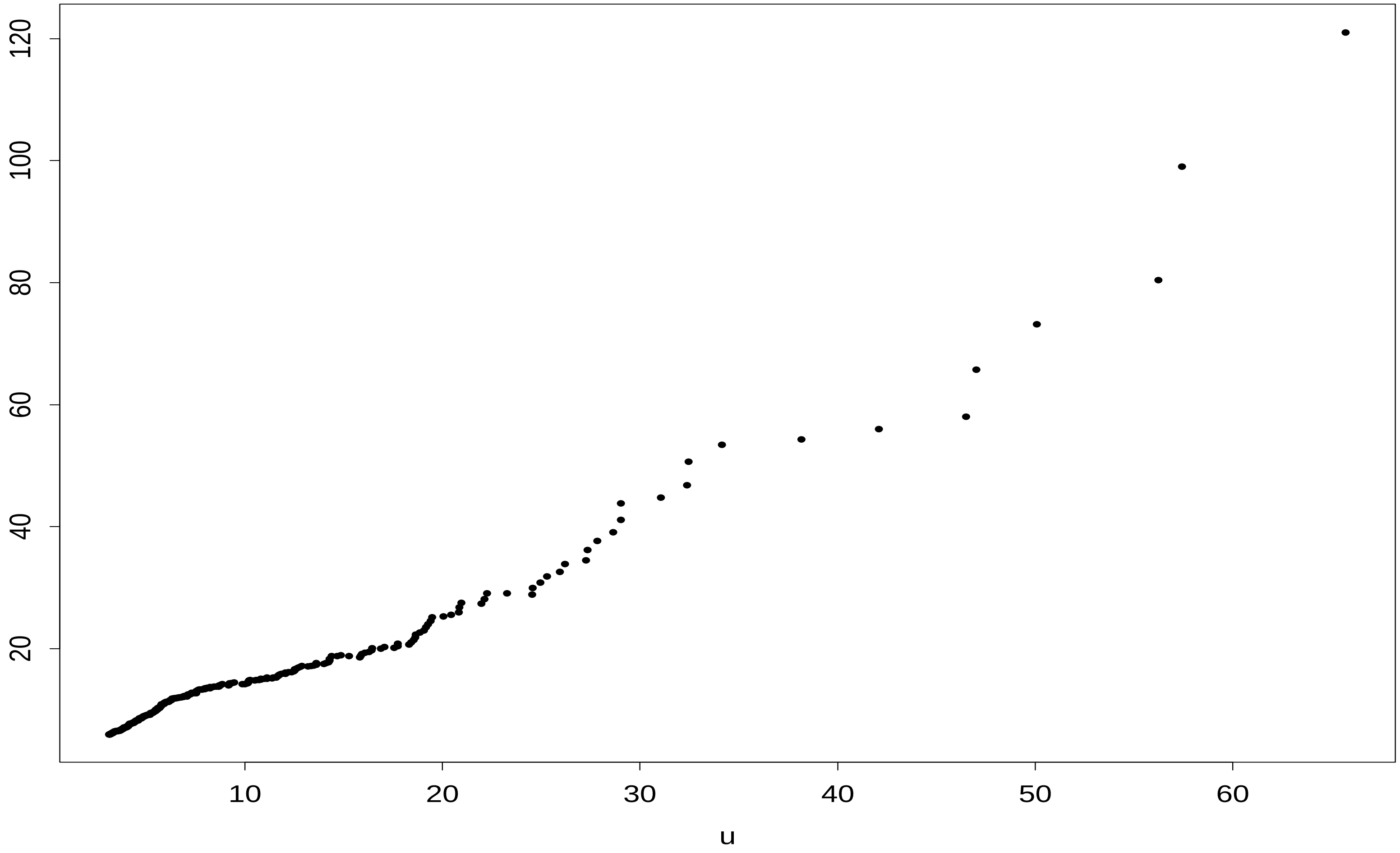

Figure 6. Empirical mean excess function of the Danish fire insurance data.

1.3. Telecommunications.

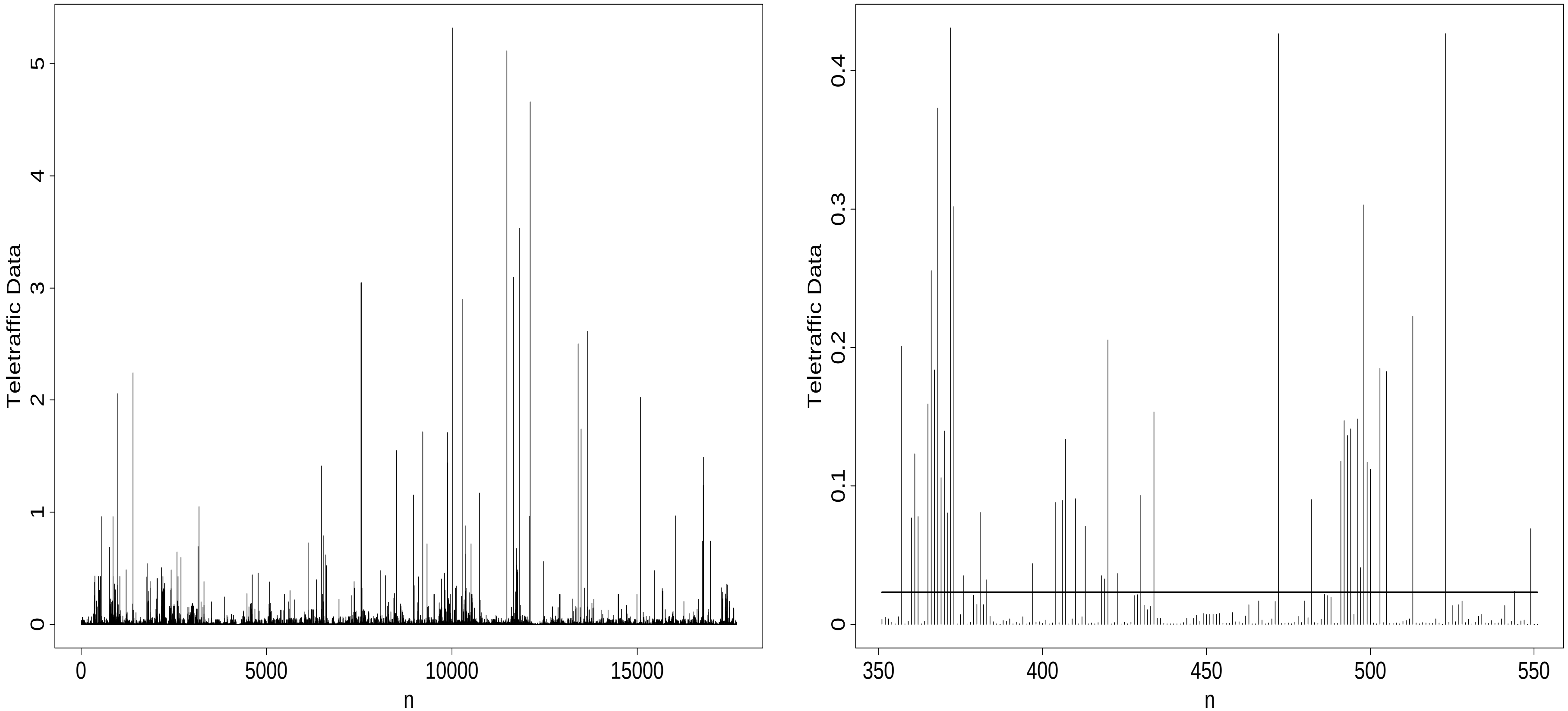

Figure 7. Time series of transmission durations (BU data).

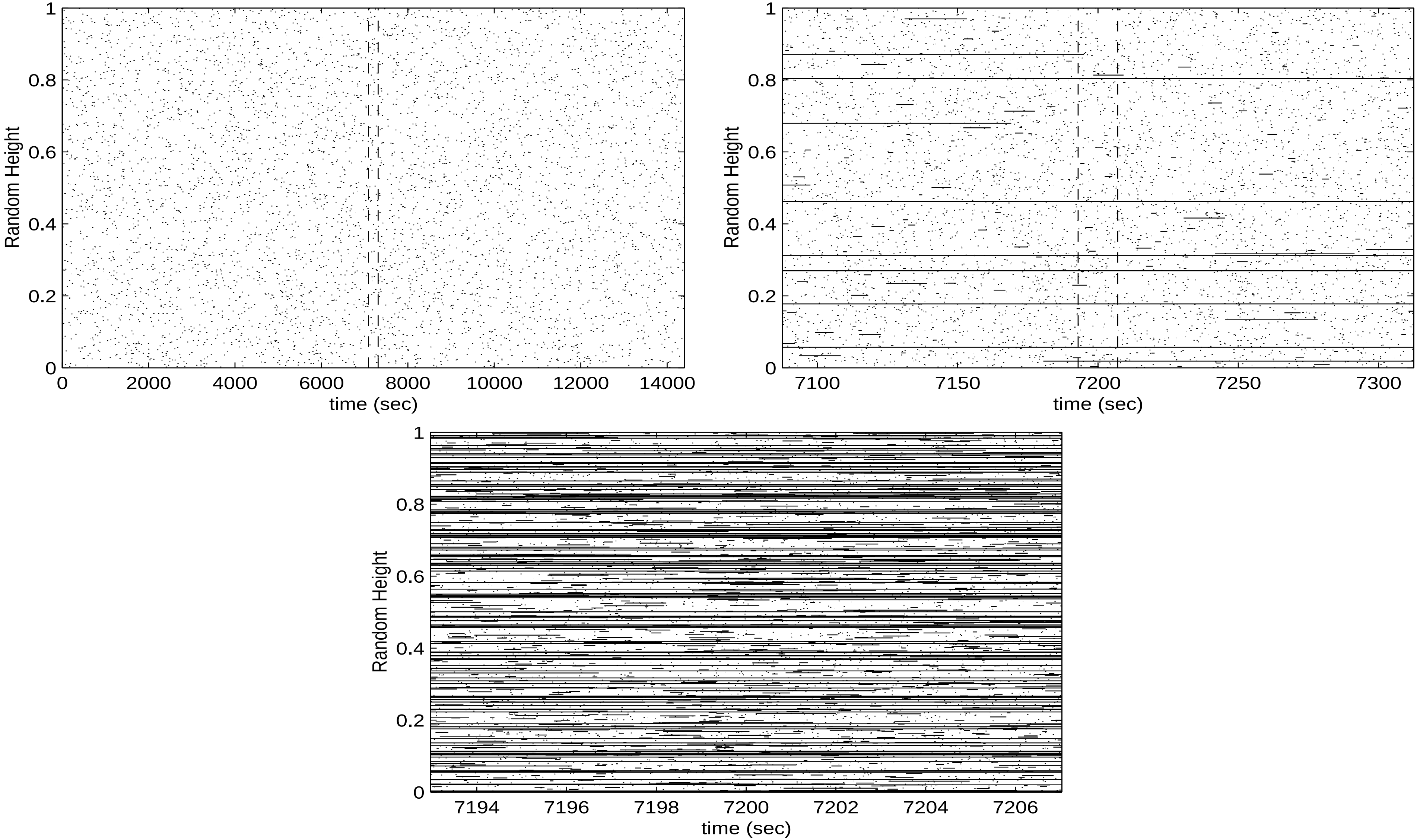

Figure 8. Mice and elephants plots (S. Marron).

2. Extremal dependence/independence in real-life data

2.1. Independence in insurance data.

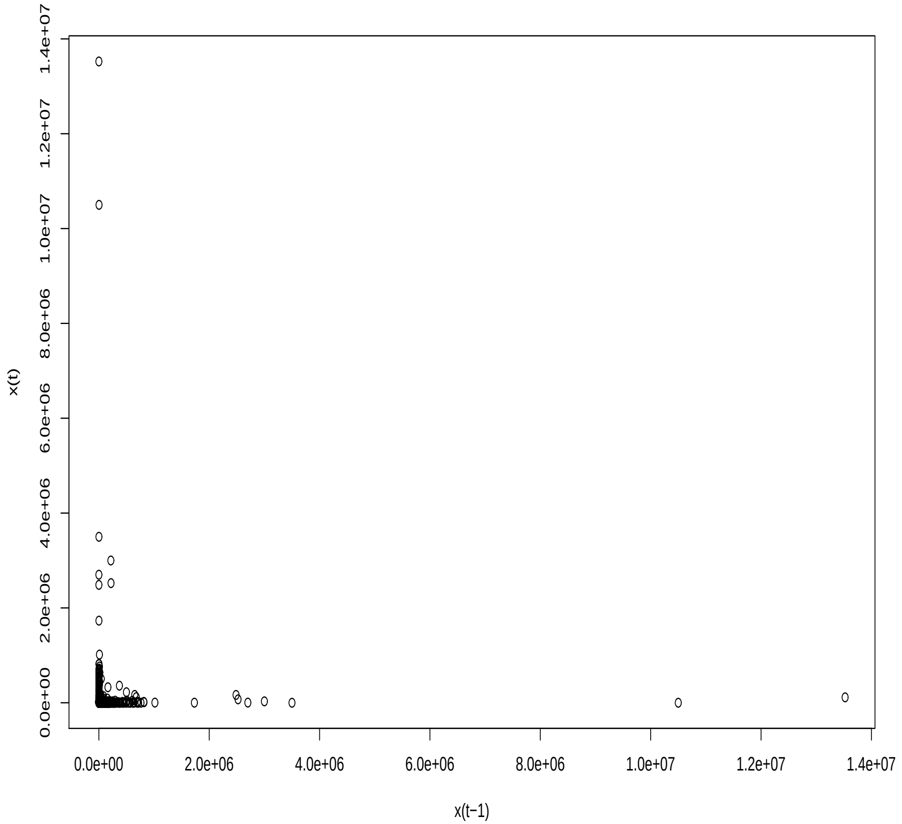

Figure 9. Scatterplot of US fire insurance losses - independence.

2.2. Extremal independence in telecommunication data.

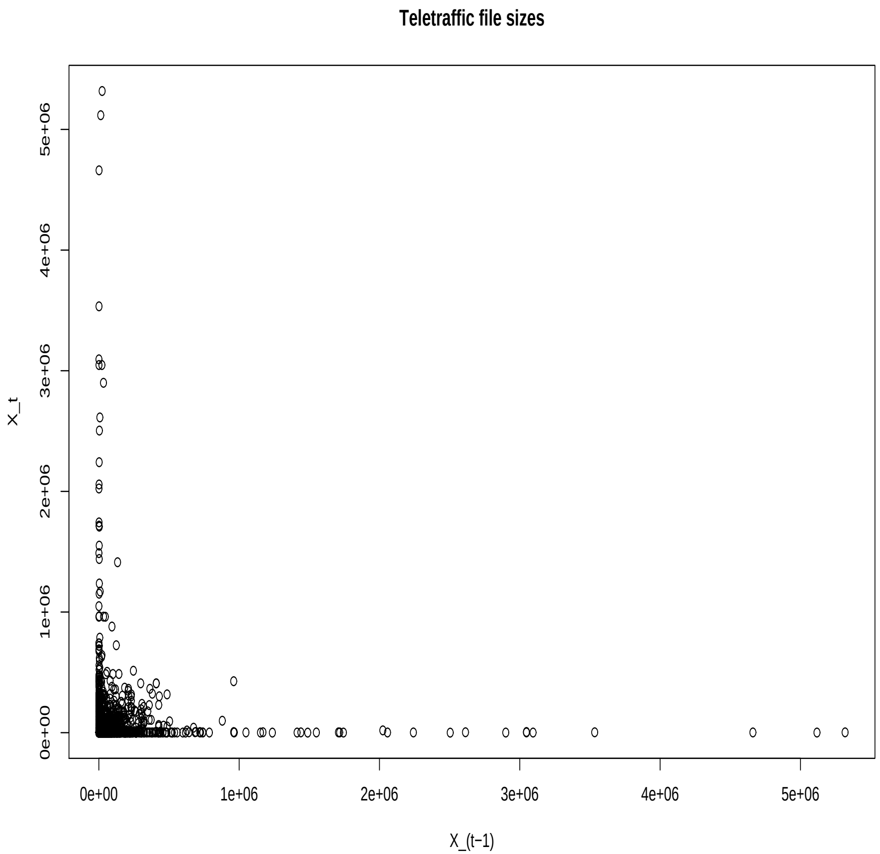

FIGURE 10. Scatterplot of file sizes of teletraffic data - extremal independence

2.3. Extremal dependence in financial data.

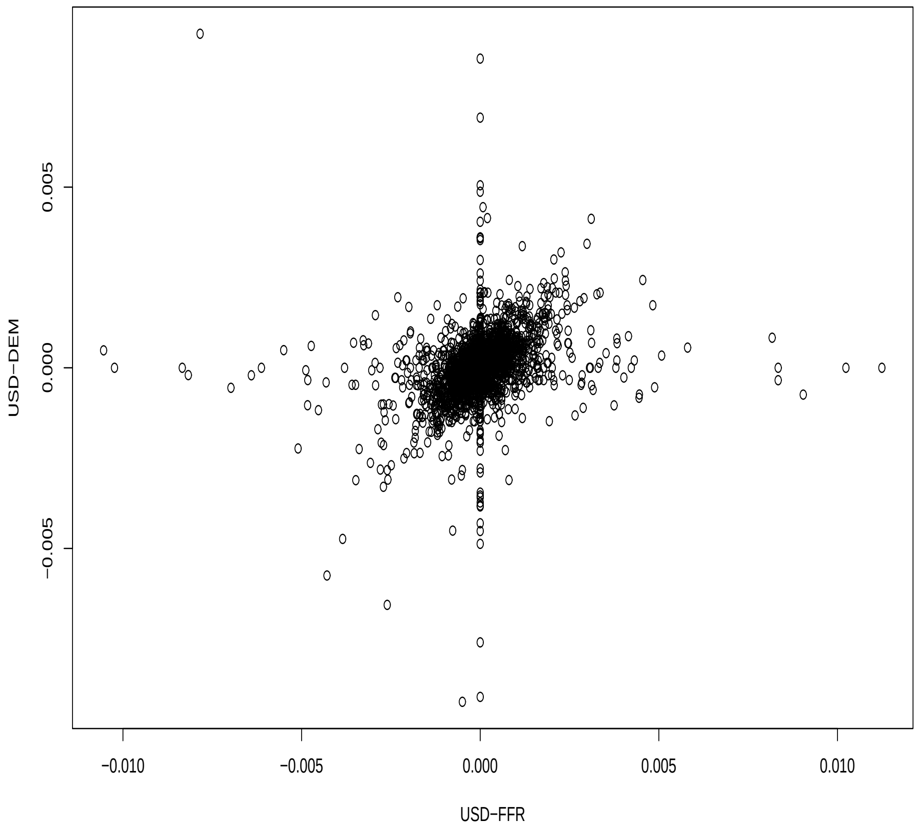

FIGURE 11. Scatterplot of 5 minute foreign exchange rate log-returns, USD-DEM against USD-FRF.

3. Extreme value theory for iid sequences Leadbetter et al. (1983), Resnick (1987), Embrechts et al. (1997), de Haan and Ferreira (2006)

3.1. Max-stable distributions (extreme value distributions).

- **A random variable X and its distribution F are max-stable if for every $n \geq 2$ there exist $c_n > 0$, $d_n \in \mathbb{R}$, such that for iid copies (X_i) of X,**

$$c_n^{-1}(M_n - d_n) = c_n^{-1}\big(\max_{i=1,\ldots,n} X_i - d_n\big) \stackrel{d}{=} X\,.$$

- **Any max-stable distribution belongs to the location/scale family of one of the three standard max-stable distributions (also called extreme value distributions):**

$$\begin{aligned}
\Phi_\alpha(x) &= \mathrm{e}^{-x^{-\alpha}}, \quad x > 0, \quad \alpha > 0 \quad \text{Fréchet}\\
\Psi_\alpha(x) &= \mathrm{e}^{-|x|^{\alpha}}, \quad x < 0, \quad \alpha > 0 \quad \text{Weibull}\\
\Lambda(x) &= \mathrm{e}^{-\mathrm{e}^{-x}}, \quad x \in \mathbb{R}, \quad \text{Gumbel}\,.
\end{aligned}$$

- **The max-stable distributions are the only possible non-degenerate weak limits for standardized maxima of an iid sequence (Fisher-Tippett Theorem 1928, Gnedenko (1943)).**
- **The 3 max-stable types can be written as one parametric family (generalized extreme value distribution (GEV)).**

- **Transformation of max-stable random variables .**

  If $X > 0$ has a Φ_α distribution,

  - **$\log X^\alpha$ has distribution Λ**
  - **$-X^{-1}$ has distribution Ψ_α.**

3.2. Maximum domains of attraction (MDA).

- The distribution F of X is in the maximum domain of attraction of the max-stable distribution $G \in \{\Phi_\alpha, \Psi_\alpha, \Lambda\}$ ($F \in \mathrm{MDA}(G)$) if there exist constants $a_n > 0$, $b_n \in \mathbb{R}$ such that

$$\lim_{n\to\infty} \mathbb{P}(a_n^{-1}(M_n - b_n) \le x) \to G(x)\,, \quad x \in \mathbb{R}\,.$$

- $F \in \mathrm{MDA}(\Phi_\alpha)$: Regular variation of the right tail

$$\overline{F}(x) = 1 - F(x) = \mathbb{P}(X > x) = x^{-\alpha} L(x)\,, \quad x > 0\,,$$

for a slowly varying function L.

Then the moments $\mathbb{E}[X^{\alpha+\delta}]$, $\delta > 0$, are infinite.

- $F \in \mathrm{MDA}(\Psi_\alpha)$: F has finite right endpoint x_F.
- $F \in \mathrm{MDA}(\Lambda)$: Moderately heavy $\to$ light tails.

- **Examples:**

  MDA(Φ_α): Student with α degrees of freedom,

  Cauchy ($\alpha = 1$),

  infinite variance α-stable distributions,

  Pareto $\overline{F}(x) = x^{-\alpha}$, $x > 1$,

  log-gamma distribution.

  MDA(Ψ_α): uniform, β-distribution.

  MDA(Ψ_α): log-normal distribution,

  Weibull $\overline{F}(x) = \mathrm{e}^{-x^\tau}$, $x > 0$, $\tau > 0$,

  gamma distribution,

  normal distribution.

4. The extremal index – a measure of the extremal cluster size

- **Let $(X_t)_{t\in\mathbb{Z}}$ be a strictly stationary real-valued time series.**
- **Its autocovariance and autocorrelation functions do in general not contain information about extremal dependence.**

4.1. Definition.

- **The *extremal index* θ_X is a standard measure of extremal dependence in a sequence:[2] for $M_n = \max_{t=1,\ldots,n} X_t$ and a suitable sequence $u_n \uparrow x_F$**

$$\mathbb{P}(M_n \le u_n) \approx [\mathbb{P}(X_1 \le u_n)]^{n\,\theta_X}\,.$$

- **$\theta_X \in [0,1]$ has the interpretation as reciprocal of the expected cluster size above high thresholds.**

[2] See Leadbetter, Lindgren, Rootzén (1983); cf. Embrechts et al. (1997), Section 8.1

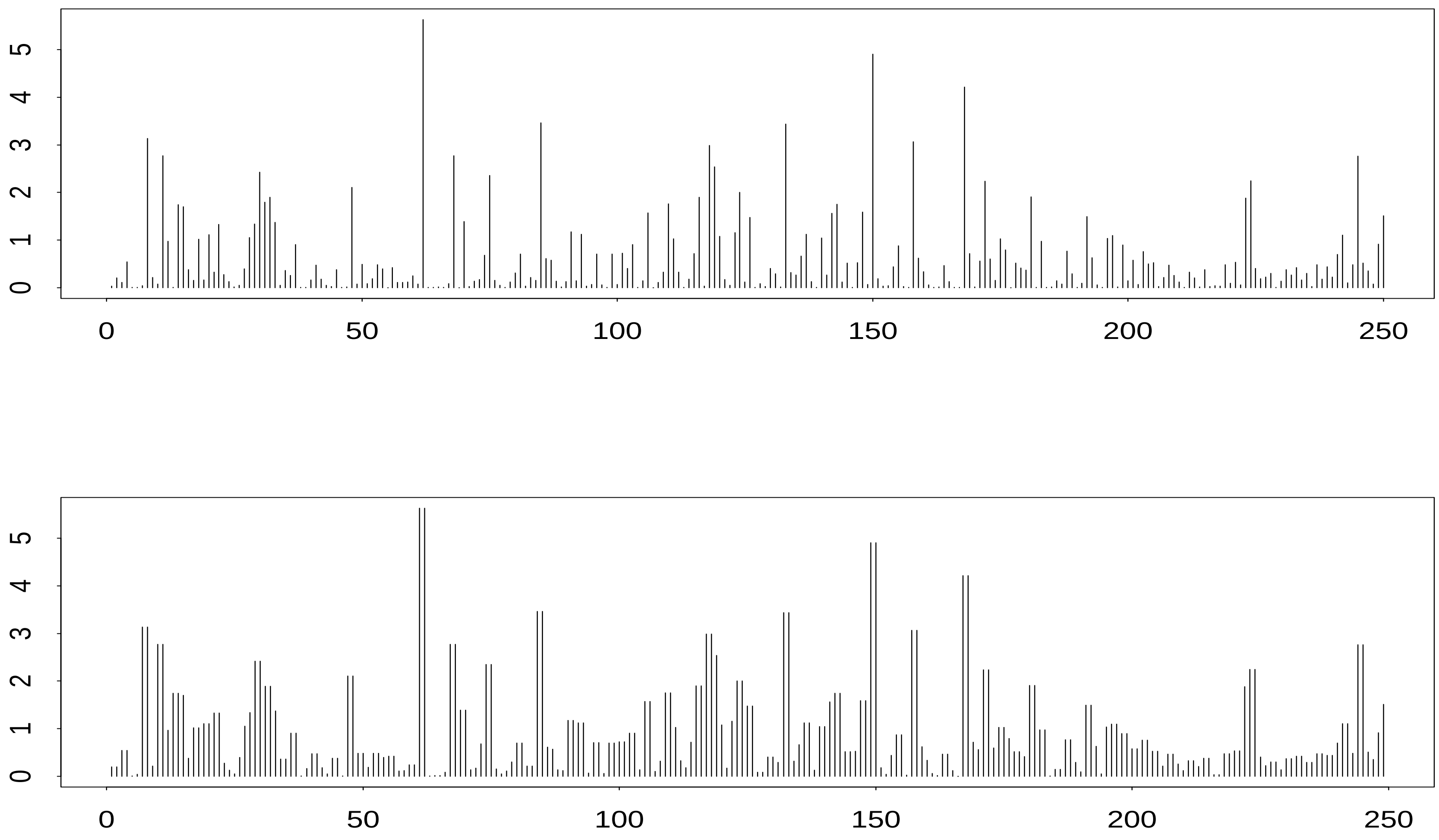

FIGURE 12. A sequence of iid random variables Y_i (Top) with distribution function $\sqrt{F}$, where F is standard exponential. Bottom: the sequence of pairwise maxima $\max(Y_i, Y_{i+1})$ with distribution F. By construction, extremes appear in clusters of size 2. The extremal index is $\theta = 1/2$.

4.2. Examples.

- A Gaussian stationary sequence (X_t) with autocorrelation function $\rho_X(h) = o(1/\log h)$, $h \to \infty$:

  $\theta_X = 1$. No extremal clustering.

- AR(1) model $X_t = \phi\, X_{t-1} + Z_t$, $\phi \in (-1,1)$, (Z_t) iid student with α degrees of freedom:

  $\theta_X = 1 - |\phi|^\alpha$.

- Models for log-returns $X_t = \log P_t - \log P_{t-1}$:

  $$X_t = \sigma_t\, Z_t, \quad (Z_t) \text{ iid}, \quad \sigma_t > 0$$

- The simple stochastic volatility model: $(\log \sigma_t)$ linear Gaussian, independent of iid student (Z_t):

  $\theta_X = 1$ Davis, Mikosch (2001ab,2009ab) No extremal clustering.

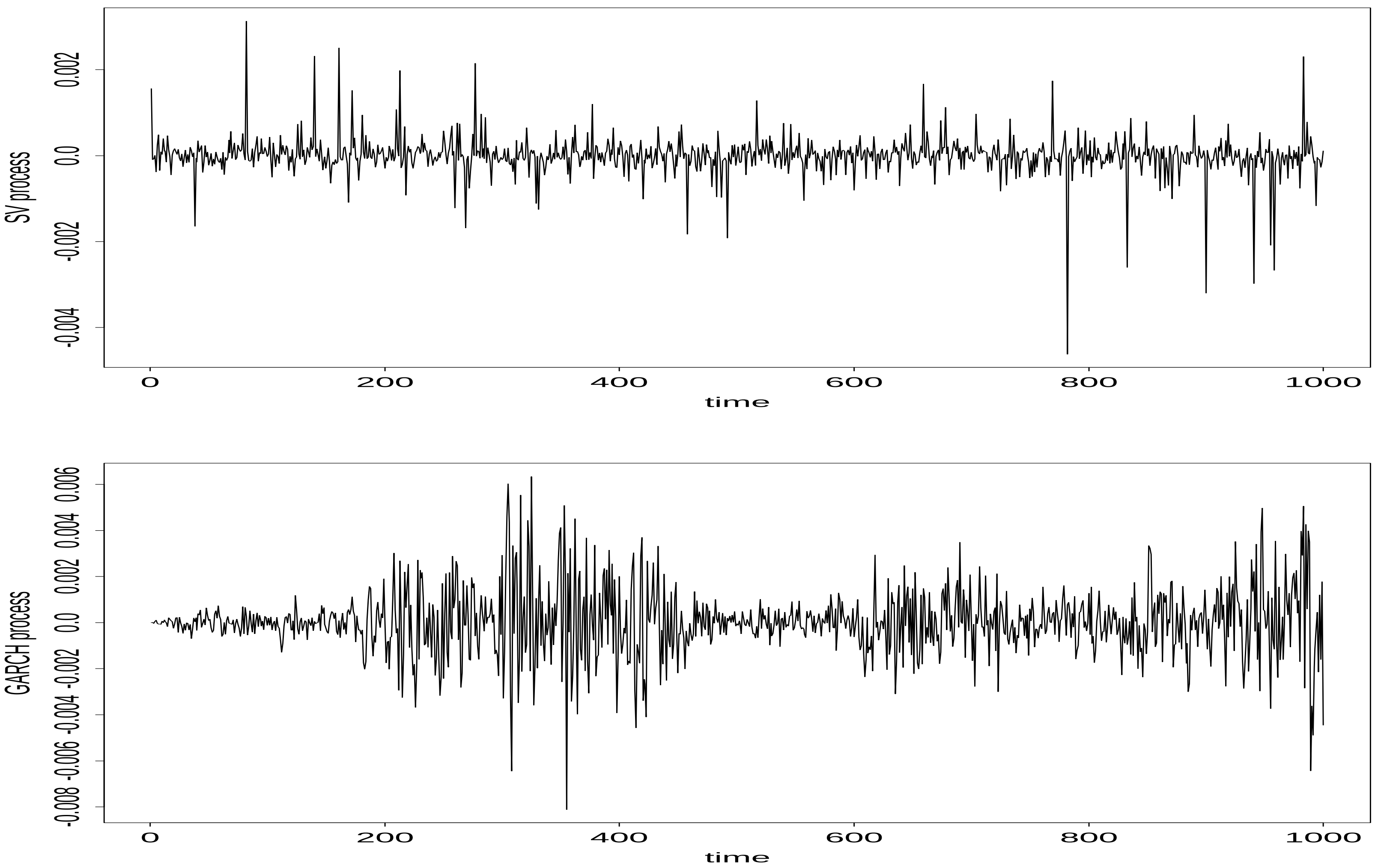

FIGURE 13. *Top:* Stochastic volatility process $\boldsymbol{X_t = \sigma_t \, Z_t}$ for iid student $\boldsymbol{(Z_t)}$ with $\mathbf{4}$ degrees of freedom, Gaussian ARMA(1,1) process $\boldsymbol{\log \sigma_t = 0.5 \log \sigma_{t-1} + 0.3\eta_{t-1} + \eta_t}$. *Bottom:* GARCH$\boldsymbol{(1,1)}$ process $\boldsymbol{X_t = (0.0001 + 0.1X_{t-1}^2 + 0.9\sigma_{t-1}^2)^{0.5} Z_t}$ for iid standard normal $\boldsymbol{(Z_t)}$.

- The GARCH(1,1) model:[3] $X_t = \sigma_t Z_t$,

$$\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1)\,\sigma_{t-1}^2\,, \quad (Z_t) \text{ iid } N(0,1).$$

There exists $\alpha > 0$ such that $\mathbb{E}(\alpha_1 Z_1^2 + \beta_1)^{\alpha/2} = 1$ and[4]

$$\frac{\alpha}{2}\int_1^\infty \mathbb{P}\left(\max_{n\geq 1}\prod_{t=1}^n(\alpha_1 Z_t^2 + \beta_1) \leq y^{-1}\right) y^{-\frac{\alpha}{2}-1}\,dy = \theta_\sigma \in (0,1)\,.$$

[3] Bollerslev (1986)

[4] de Haan, Resnick, Rootzén, de Vries (1989)

- The GARCH(1,1) model: $X_t = \sigma_t Z_t$,

$$\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1)\, \sigma_{t-1}^2 \,, \quad (Z_t) \text{ iid } N(0,1).$$

There exists $\alpha > 0$ such that $\mathbb{E}(\alpha_1 Z_1^2 + \beta_1)^{\alpha/2} = 1$ and

$$\frac{\alpha}{2} \int_1^\infty \mathbb{P}\left(\max_{n \geq 1} \prod_{t=1}^n (\alpha_1 Z_t^2 + \beta_1) \leq y^{-1}\right) y^{-\frac{\alpha}{2}-1}\, dy = \theta_\sigma \in (0,1)\,.$$

- Expressions for the extremal index of a stationary process are often complicated.

- The GARCH(1,1) model: $X_t = \sigma_t Z_t$,

$$\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1)\,\sigma_{t-1}^2\,, \qquad (Z_t) \text{ iid } N(0,1).$$

  There exists $\alpha > 0$ such that $\mathbb{E}(\alpha_1 Z_1^2 + \beta_1)^{\alpha/2} = 1$ and

$$\frac{\alpha}{2}\int_1^\infty \mathbb{P}\left(\max_{n\geq 1}\prod_{t=1}^n(\alpha_1\,Z_t^2 + \beta_1) \leq y^{-1}\right)\,y^{-\frac{\alpha}{2}-1}\,dy = \theta_\sigma \in (0,1)\,.$$

- Expressions for the extremal index of a stationary process are often complicated.

- Monte Carlo simulation is not straightforward.

- The GARCH(1,1) model: $X_t = \sigma_t Z_t$,

$$\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1)\,\sigma_{t-1}^2\,, \qquad (Z_t) \text{ iid } N(0,1).$$

There exists $\alpha > 0$ such that $\mathbb{E}(\alpha_1 Z_1^2 + \beta_1)^{\alpha/2} = 1$ and

$$\frac{\alpha}{2}\int_1^\infty \mathbb{P}\left(\max_{n\geq 1}\prod_{t=1}^n(\alpha_1 Z_t^2 + \beta_1) \leq y^{-1}\right) y^{-\frac{\alpha}{2}-1}\,dy = \theta_\sigma \in (0,1)\,.$$

- Expressions for the extremal index of a stationary process are often complicated.

- Monte Carlo simulation is not straightforward.

- Estimation of the extremal index and extremal cluster size distribution is non-trivial; see C. Y. Robert (2009)

5. Regular variation - univariate and multivariate

5.1. Univariate regularly varying distributions. .[5]

- Recall that $F \in \text{MDA}(\Phi_\alpha)$ for some $\alpha > 0$ if and only if

$$\overline{F}(x) = \mathbb{P}(X > x) = x^{-\alpha} L(x)\,, \quad x > 0\,,$$

for some slowly varying function L, (i.e. $L(cx)/L(x) \to 1$, $x \to \infty$, $c > 0$)

- We call a random variable $X \in \mathbb{R}$ and its distribution F **regularly varying with index $\alpha > 0$** if there exist constants $p, q \geq 0$ such that $p + q = 1$ and

$$F(-x) \sim q\, x^{-\alpha} L(x) \quad \text{and} \quad \overline{F}(x) \sim p\, x^{-\alpha} L(x)\,, \quad x \to \infty\,.$$

[5] See Bingham et al. [5], for an encyclopedia on regularly varying functions, Resnick [49, 50] for a modern theory of regular variation for the purposes of applied probability and statistics.

- If e.g. $p = 0$: $\overline{F}(x) = o(x^{-\alpha} L(x))$, $x \to \infty$.
- Equivalently, $|X|$ is regularly varying with index $\alpha > 0$ and

$$\frac{\mathbb{P}(X \le -x)}{\mathbb{P}(|X| > x)} \to q \quad \text{and} \quad \frac{\mathbb{P}(X > x)}{\mathbb{P}(|X| > x)} \to p\,, \qquad x \to \infty\,.$$

- **Examples.** Pareto, student, Cauchy, α-stable, $\alpha \in (0, 2)$, Burr, log-gamma, Fréchet.

- **Two fundamental operations.**[6]

Convolution. Feller (1971) Let $X_1 > 0$ be regularly varying with $\alpha > 0$. Assume X_2 regularly varying with index α and independent of X_1 **OR** $\mathbb{P}(|X_2| > x) = o(\mathbb{P}(X_1 > x))$.

Then $X_1 + X_2$ is regularly varying with index α and

$$\mathbb{P}(X_1 + X_2 > x) \sim \mathbb{P}(X_1 > x) + \mathbb{P}(X_2 > x)\,, \quad x \to \infty\,.$$

Products. Breiman (1965) $\sigma > 0$, $X > 0$ independent and $\mathbb{E}\sigma^{\alpha+\delta} < \infty$ for some $\delta > 0$, X regularly varying with index α.

Then, as $x \to \infty$,

$$\mathbb{P}(\sigma X > x) \sim \mathbb{E}\sigma^{\alpha}\,\mathbb{P}(X > x)\,.$$

[6] See Resnick (2007)

- **Examples.**

Stochastic volatility model. $X_t = \sigma_t Z_t$, $t \in \mathbb{Z}$, σ_t log-normal, (Z_t) iid regularly varying with index α, (σ_t) and (Z_t) independent. Then, as $x \to \infty$,

$$\mathbb{P}(X_t > x) \sim \mathbb{E}\sigma_0^\alpha \, \mathbb{P}(Z_0 > x)\,,$$

$$\mathbb{P}(X_t \le -x) \sim \mathbb{E}\sigma_0^\alpha \, \mathbb{P}(Z_0 \le -x)\,.$$

Moving average. $X_t = \theta_0 Z_t + \theta_1 Z_{t-1} + \cdots + \theta_m Z_{t-m}$, $t \in \mathbb{Z}$, $m \ge 1$, $Z_t > 0$ iid regularly varying with index α.

$$\mathbb{P}(X_t > x) \sim \sum_{i=1}^{m} \mathbb{P}(\theta_i \, Z_i > x) \sim \mathbb{P}(Z_0 > x) \sum_{i=0}^{m} \theta_i^\alpha \, I_{\theta_i > 0}\,, \quad x \to \infty\,.$$

How can one model spatio-temporal extremal dependence and heavy tails?

- **One needs to model both the size and the direction of extremes.**

Asymptotic extremal independence

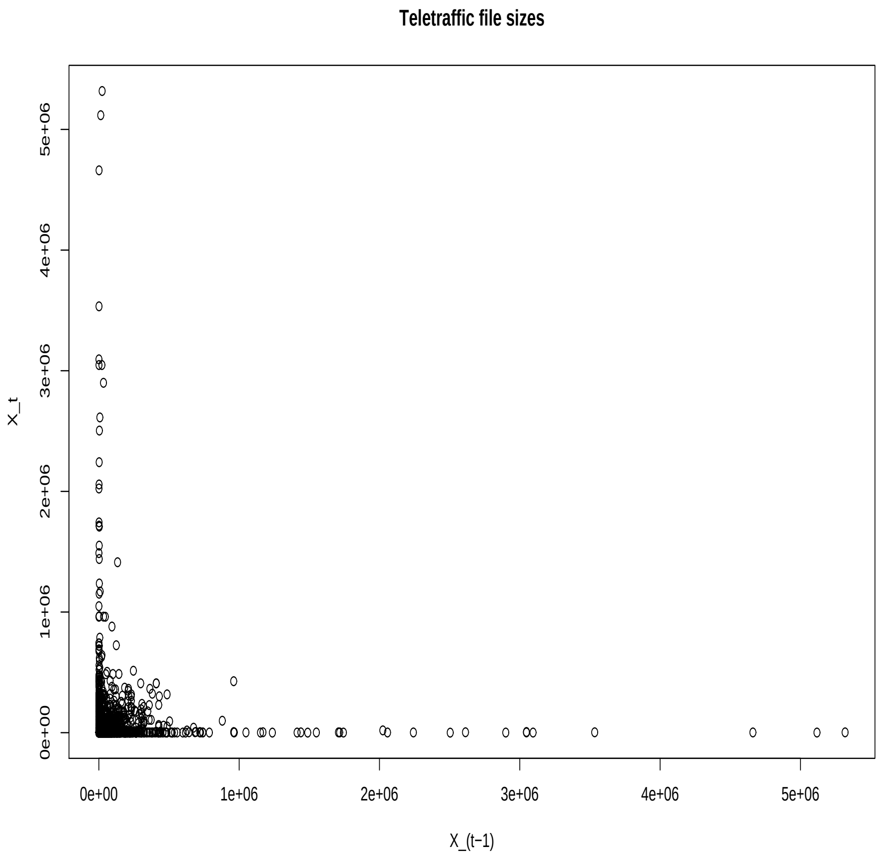

FIGURE 14. Scatterplot of file sizes of teletraffic data.

Asymptotic extremal dependence

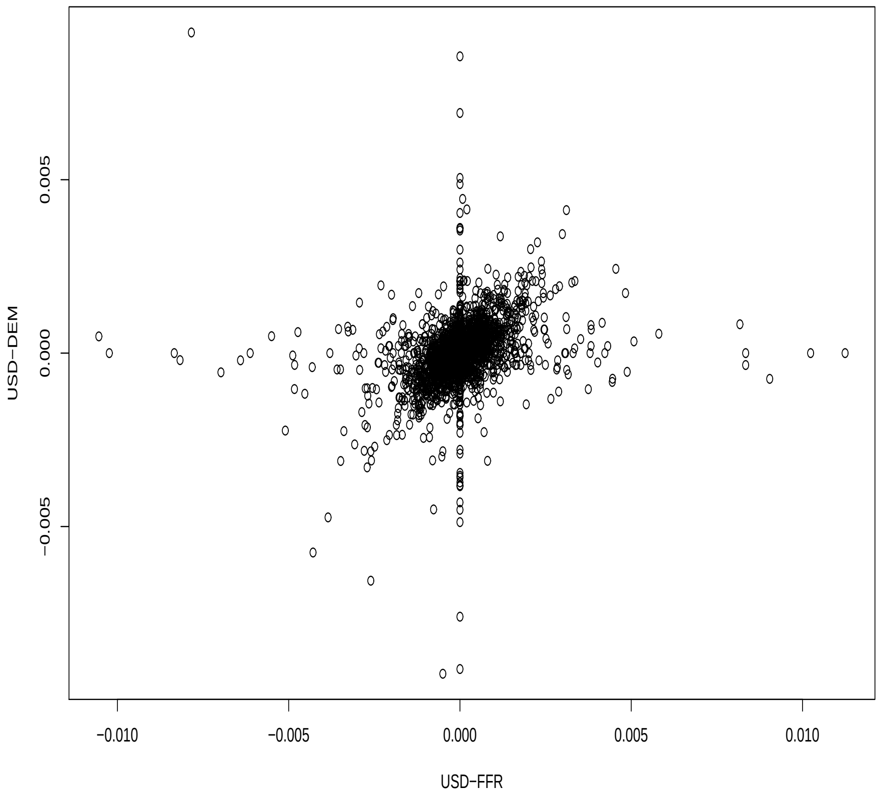

Figure 15. Scatterplot of 5 minute foreign exchange rate log-returns, USD-DEM against USD-FRF.

5.2. Multivariate regular variation Resnick (1987,2007).

- A random vector $\mathbf{X} \in \mathbb{R}^d$ and its distribution are regularly varying with index $\alpha > 0$: there exists a random vector $\boldsymbol{\Theta} \in \mathbb{S}^{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$ such that for $t > 0$:

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

The distribution of $\boldsymbol{\Theta}$: spectral measure of $\mathbf{X}$.

- Equivalently,

$$\frac{\mathbb{P}\left(x^{-1}\mathbf{X} \in \;\cdot\right)}{\mathbb{P}(|\mathbf{X}| > x)} \xrightarrow{v} \mu(\cdot)\,,$$

for a non-null Radon measure μ on the Borel σ-field of $\overline{\mathbb{R}}_0^d = \overline{\mathbb{R}}^d \backslash \{0\}$ with $\mu(tA) = t^{-\alpha}\mu(A)$, $t > 0$.

- **Equivalently:** as $x \to \infty$,

$$\frac{\mathbb{P}(|\mathbf{X}| > tx)}{\mathbb{P}(|\mathbf{X}| > x)} \to t^{-\alpha}\,, \quad t > 0\,, \quad \text{and}$$

$$\mathbb{P}\Big(\frac{\mathbf{X}}{|\mathbf{X}|} \in \cdot \mid |\mathbf{X}| > x\Big) \xrightarrow{w} \mathbb{P}(\Theta \in \cdot)\,.$$

- **A toy example.** R, θ independent, θ distributed on $[0, 2\pi)$, $\mathbb{P}(R > r) = r^{-\alpha}$, $r > 1$, Pareto.

$$\mathbf{X} = R\,(\cos\theta, \sin\theta) = R\,\Theta,$$

Then

$$|\mathbf{X}| = R \quad \text{and} \quad \mathbf{X}/|\mathbf{X}| = \Theta = (\cos\theta, \sin\theta)\,,$$

$$\mathbb{P}(R > tx) = t^{-\alpha}\mathbb{P}(R > x)\,, \quad tx > 1\,,$$

$$\mathbb{P}(\Theta \in \cdot \mid R > x) = \mathbb{P}(\Theta \in \cdot)\,.$$

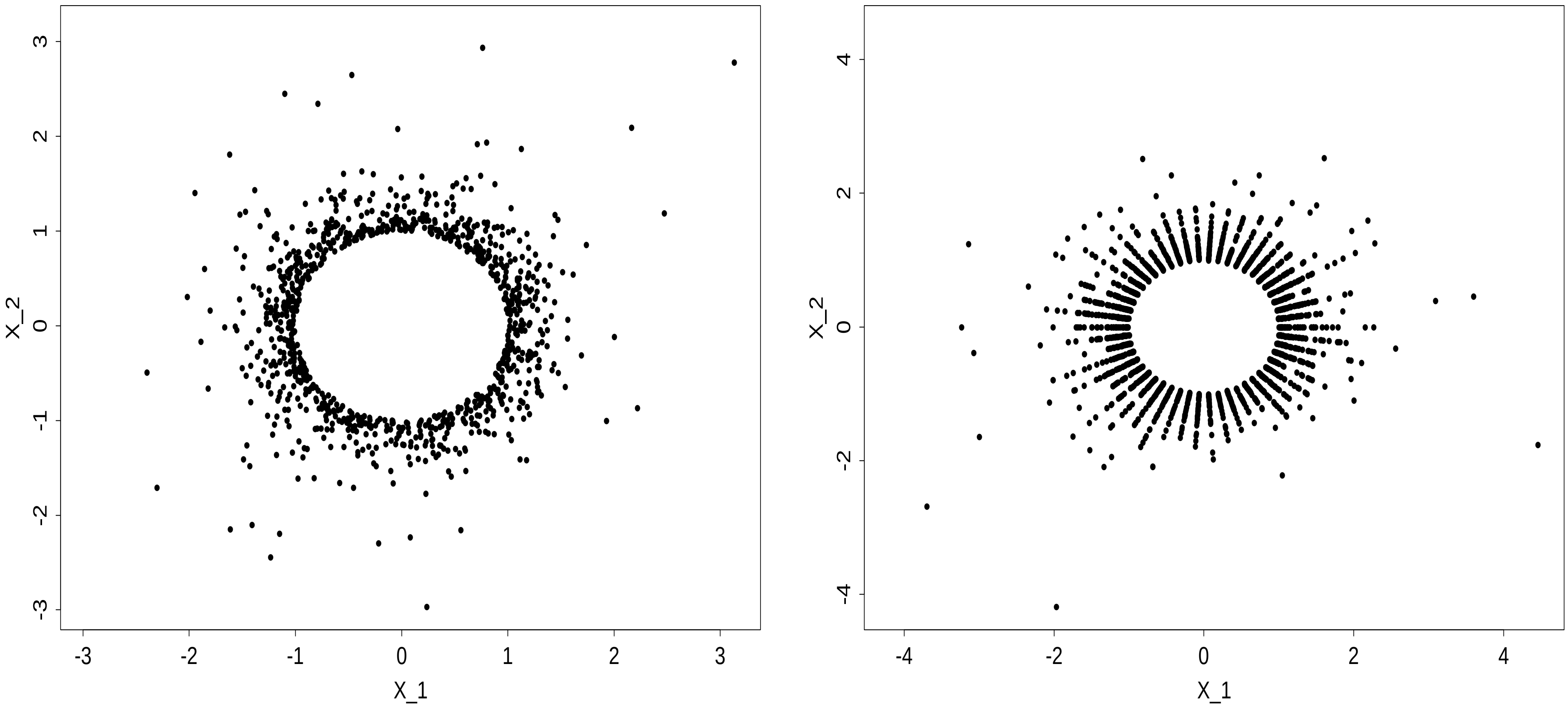

FIGURE 16. IID vectors $\mathbf{X}_i$ from the toy model with tail index $\alpha = 5$. Left: θ is uniform on $[0, 2\pi)$. Right: θ has a discrete uniform distribution on the points $2\pi i/50$.

- **Example**

Assume $\mathbf{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2) \in \mathbb{R}_+^2$ **iid regularly varying with index** $\alpha > 0$**. Choose** $|\mathbf{X}| = \max(\boldsymbol{X}_1, \boldsymbol{X}_2)$**. Then, for any** $\varepsilon > 0$**,**

$$\begin{aligned}\frac{\mathbb{P}(x^{-1}\mathbf{X} \in (\varepsilon, \infty)^2)}{\mathbb{P}(|\mathbf{X}| > x)} &= \frac{\mathbb{P}(\boldsymbol{X}_1 > \varepsilon x, \boldsymbol{X}_2 > \varepsilon x)}{\mathbb{P}(\max(\boldsymbol{X}_1, \boldsymbol{X}_2) > x)} \\ &= \frac{[\mathbb{P}(\boldsymbol{X}_1 > \varepsilon x)]^2}{2\mathbb{P}(\boldsymbol{X}_1 > x) - [\mathbb{P}(\boldsymbol{X}_1 > x)]^2} \\ &\to \mu((\varepsilon, \infty)^2) = 0\,.\end{aligned}$$

μ **does not charge sets bounded away from the axes.**

Hence μ **and the spectral measure** $\mathbb{P}(\boldsymbol{\Theta} \in \cdot)$ **are concentrated on the axes.**

- **In general, if a regularly varying vector X has independent components the measures μ and $\mathbb{P}(\boldsymbol{\Theta} \in \cdot)$ are concentrated on the axes.**
- **This means: It is very unlikely that two distinct components of X are large (extreme) at the same time.**

Extremes occur close to the axes.

- **For a vector X with *dependent* components, this property is sometimes referred to as asymptotic extremal independence.**

- **Example.** The stochastic volatility model: for $s \neq t$, iid regularly varying (Z_t), independent of strictly stationary log-normal (σ_t),

$$
\begin{aligned}
\mathbb{P}(|X_t| > \varepsilon x, |X_s| > \varepsilon x) &= \mathbb{P}(\min(\sigma_t |Z_t|, \sigma_s |Z_s|) > \varepsilon x) \\
&\leq \mathbb{P}(\max(\sigma_t, \sigma_s) \min(|Z_t|, |Z_s|) > \varepsilon x) \\
&\sim \mathbb{E} \max(\sigma_t^{2\alpha}, \sigma_s^{2\alpha})\, \mathbb{P}(\min(|Z_t|, |Z_s|) > \varepsilon x) \\
&= \mathbb{E} \max(\sigma_t^{2\alpha}, \sigma_s^{2\alpha})\, [\mathbb{P}(|Z_t| > \varepsilon x)]^2 \\
&= o(\mathbb{P}(|X_t| > x)) \,, \quad x \to \infty \,.
\end{aligned}
$$

Here we used Breiman's result. Hence for any $\varepsilon > 0$, as $x \to \infty$,

$$
\frac{\mathbb{P}(|X_t| > \varepsilon x, |X_s| > \varepsilon x)}{\mathbb{P}(|(X_t, X_s)| > x)} \to 0 = \mu(\{(x_1, x_2) : \min(|x_1|, |x_2|) > \varepsilon\}) \,.
$$

This means:

the limit measures μ and $\mathbb{P}(\Theta \in \cdot)$ of the regularly varying vector (X_t, X_s) are concentrated on the axes.

or

there is asymptotic extremal independence between X_t and X_s

although X_t, X_s are dependent.

- **Further examples**

  - **X has iid student(α) distributed components. $\mathbb{P}(\boldsymbol{\Theta} \in \cdot)$ is concentrated on the intersection of unit ball and axes.**

  - **X has a multivariate student(α) distribution. $\mathbb{P}(\boldsymbol{\Theta} \in \cdot)$ is supported on the whole unit ball.**

  - **X is obtained from a Gaussian vector by transforming the marginals to student(α). Then $\mathbb{P}(\boldsymbol{\Theta} \in \cdot)$ is concentrated on the intersection of unit ball and axes.**

    X exhibits asymptotic extremal independence.

5.3. Operations on regularly varying vectors.

- **Lemma.** Let $\mathbf{X} \in \mathbb{R}^d$ be $\mathrm{RV}(\mu_{\mathrm{X}}, \alpha)$ and $g : \mathbb{R}^d \to \mathbb{R}^{d'}$ be continuous, positive homogeneous of order 1. Then

$$\frac{\mathbb{P}(x^{-1} g(\mathrm{X}) \in \cdot)}{\mathbb{P}(|\mathrm{X}| > x)} \xrightarrow{v} \mu_{\mathrm{X}}(g^{-1}(\cdot)) .$$

If $\mu_{\mathrm{X}}(g^{-1}(\cdot))$ is non-null $g(\mathrm{X})$ is regularly varying with index α.

- In particular, if the inverse image of the null set in $\mathbb{R}^{d'}$ under the mapping g is the null set in $\mathbb{R}^d$ the limit measure $\mu_{\mathrm{X}}(g^{-1}(\cdot))$ is non-null and $g(\mathrm{X})$ is regularly varying with index α.

- **Examples.** Let X be an $\mathbb{R}^d$-valued random vector which is RV$(\alpha, \mu_{\mathrm{X}})$.
  - If the support of X is the positive quadrant and $g(\mathrm{X}) = \max(X_1, \ldots, X_d)$, then the inverse image of the null set is the null set and $g(\mathrm{X})$ is regularly varying.
  - The sum $g(\mathrm{X}) = X_1 + \cdots + X_d$ is regularly varying if $\mu_X(g^{-1}(\cdot)) \neq o$. In this case, for $z > 0$,

$$\begin{aligned}\mu_X(g^{-1}(z,\infty]) &= \mu_X(\{y \in \overline{\mathbb{R}}_0^d : y_1 + \cdots + y_d > z\}) \\ &= z^{-\alpha}\mu_X(\{y \in \overline{\mathbb{R}}_0^d : y_1 + \cdots + y_d > 1\}),\end{aligned}$$

$$\begin{aligned}\mu_X(g^{-1}(-\infty,-z]) &= \mu_X(\{y \in \overline{\mathbb{R}}_0^d : y_1 + \cdots + y_d \leq -z\}) \\ &= z^{-\alpha}\mu_X(\{y \in \overline{\mathbb{R}}_0^d : y_1 + \cdots + y_d \leq -1\}).\end{aligned}$$

The right-hand side can be zero, e.g. if $\mathbf{X} = Y(1,-1)'$ and $Y > 0$ is regularly varying. Then $X_1 + X_2 = 0$.

– **The function $g(\mathbf{X}) = \theta'\mathbf{X}$ is positive homogeneous for $\theta \in \mathbb{S}^{d-1}$ (relative to the Euclidean norm). If θ has positive components and $\mathbf{X}$ is supported on the positive quadrant, $\theta'\mathbf{X} = 0$ a.s. implies $\mathbf{X} = 0$ a.s. In this case, $\theta'\mathbf{X}$ is regularly varying.**

– **Any norm $g(\mathbf{X}) = |\mathbf{X}|$ is positive homogeneous and $|\mathbf{X}| = 0$ implies $\mathbf{X} = 0$ a.s. Hence $|\mathbf{X}|$ is regularly varying.**

- **Convolution.** Let X_i be $\mathbb{R}^d$-valued random vectors which are $\mathrm{RV}(\mu_{\mathrm{X}_i}, \alpha)$, $i = 1, 2$. Assume that

$$\lim_{x \to \infty} \frac{\mathbb{P}(|\mathrm{X}_2| > x)}{\mathbb{P}(|\mathrm{X}_1| > x)} = c_0 \tag{5.1}$$

for some non-negative constant c_0 and

$$\mathbb{P}(|\mathrm{X}_1| > x\,, |\mathrm{X}_2| > x) = o(\mathbb{P}(|\mathrm{X}_1| > x))\,, \quad x \to \infty\,.$$

Then

$$\frac{\mathbb{P}(x^{-1}(\mathrm{X}_1 + \mathrm{X}_2) \in \cdot)}{\mathbb{P}(|\mathrm{X}_1| > x)} \xrightarrow{v} \mu_{\mathrm{X}_1} + c_0\, \mu_{\mathrm{X}_2}\,.$$

The result remains valid for regularly varying X_1 and any X_2 if (5.1) holds with $c_0 = 0$.

- **Example.** **Assume $(\mathbf{X}_i)$ iid and $\mathrm{RV}(\mu_X, \alpha)$. Then for**

  $\mathbf{S}_n = \mathbf{X}_1 + \cdots + \mathbf{X}_n$,

$$\frac{\mathbb{P}(x^{-1}\mathbf{S}_n \in \cdot)}{\mathbb{P}(|\mathbf{X}| > x)} \xrightarrow{v} n\,\mu_X(\cdot)\,.$$

- **Products.**[7] Let A be a random $d' \times d$ matrix, independent of the $\mathbb{R}^d$-valued random vector X which is $\mathrm{RV}(\mu_{\mathrm{X}}, \alpha)$ for some $\alpha > 0$ and $\mathbb{E}\|A\|^{\alpha+\varepsilon} < \infty$ for some $\varepsilon > 0$. Here $\|A\|$ is any norm of A. Then

$$\frac{\mathbb{P}(x^{-1}A\,\mathrm{X} \in \cdot)}{\mathbb{P}(|\mathrm{X}| > x)} \xrightarrow{v} \mathbb{E}\mu_X(\{y \in \mathbb{R}^d : Ay \in \cdot\}) = \mathbb{E}\mu_X(A^{-1}\cdot)\,.$$

- If the limit measure is non-null then $A\mathrm{X}$ is regularly varying with index α.
- The result remains valid if $\mathbb{E}\|A\|^{\alpha} < \infty$ and $\mathbb{P}(|\mathrm{X}| > x) \sim c\,x^{-\alpha}$ as $x \to \infty$.

[7] See Basrak et al. [1].

6. Regularly varying stationary sequences

- **A real-valued stationary sequence (X_t) is regularly varying with index $\alpha > 0$ if its finite-dimensional distributions are regularly varying with index α.**
- **Equivalently, for every $k \geq 1$, there exists a non-null Radon measure on $\overline{\mathbb{R}}_0^k$ such that**

$$\frac{\mathbb{P}(x^{-1}(X_1, \ldots, X_k) \in \cdot)}{\mathbb{P}(|X_0| > x)} \xrightarrow{v} \mu_k(\cdot)\,.$$

  The measures μ_k determine the extremal dependence structure of the finite-dimensional distributions.
- **Notice: Normalization $\mathbb{P}(|X_0| > x)$ does not depend on k.**

6.1. Examples.

Linear process.

- Examples of linear processes are ARMA processes with iid noise (Z_t), e.g. the AR(p) and MA(q) processes

$$X_t = Z_t + \varphi_1 X_{t-1} + \cdots + \varphi_p X_{t-p}\,,$$
$$X_t = Z_t + \theta_1 Z_{t-1} + \cdots + \theta_q Z_{t-q}\,.$$

- A linear process

$$X_t = \sum_j \psi_j Z_{t-j}, \quad t \in \mathbb{Z},$$

is regularly varying with index $\alpha > 0$ if the iid sequence (Z_t) is regularly varying with index α.

- **Under mild conditions on (ψ_j),**[8]

$$\frac{\mathbb{P}(X_0 > x)}{\mathbb{P}(|Z_0| > x)} \sim \sum_j |\psi_j|^\alpha (p\, I_{\psi_j>0} + q\, I_{\psi_j<0}) = \|\psi\|_\alpha^\alpha\,, \quad x \to \infty\,.$$

- **Expressions for μ_k are complicated.**

8 See Resnick (1987); cf. Embrechts et al. (1997), Appendix

Solutions to stochastic recurrence equations.

- For an iid sequence $((A_t, B_t))_{t\in\mathbb{Z}}$, $A, B > 0$, the stochastic recurrence equation

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

has a unique stationary solution

$$X_t = B_t + \sum_{i=-\infty}^{t-1} A_t \cdots A_{i+1} B_i\,, \quad t \in \mathbb{Z},$$

provided $\mathbb{E}\log A < 0$, $\mathbb{E}\log^+ |B| < \infty$.

- The sequence (X_t) is regularly varying with index α which is the unique solution to $\mathbb{E}[A^\kappa] = 1$, $\kappa > 0$, (given this solution exists) Kesten (1973), Goldie (1991) and

$$\mathbb{P}(X_0 > x) \sim c_+ x^{-\alpha}\,, \quad x \to \infty\,.$$

- **Regular variation of the finite-dimensional distributions.**

  Notice that with $\Pi_i = A_1 \cdots A_i$, $i \geq 1$,

$$(X_1, \ldots, X_n) = (\Pi_1, \ldots, \Pi_n)\, X_0 + R_n$$

  We assume $\mathbb{E}[|B|^\alpha] < \infty$ hence $\mathbb{E}[|R_n|^\alpha] < \infty$ and therefore

$$\mathbb{P}(|R_n| > x) = o(\mathbb{P}(|X_0| > x))\,, \qquad x \to \infty\,.$$

- Applications of the convolution and product rules for regularly varying vectors (p. 46 and 48) yield the joint regular variation of $(X_1, \ldots, X_n)$: for smooth sets $C \in (0,\infty)^n$,

$$\frac{\mathbb{P}(x^{-1}(X_1, \ldots, X_n) \in C)}{\mathbb{P}(X_0 > x)}$$
$$\to \mathbb{E}\mu_{X_0}(\{y \in (0,\infty) : (1, \Pi_1, \ldots, \Pi_{n-1})y \in C\})\,.$$

- **But for $b > 0$,**

$$\frac{\mathbb{P}(x^{-1}X_0 > b)}{\mathbb{P}(X_0 > x)} \to b^{-\alpha} = \mu_{X_0}((b,\infty))\,, \quad x \to \infty\,.$$

Hence

$$\begin{aligned}&\mathbb{E}\mu_{X_0}(\{y \in (0,\infty) : (\Pi_1,\ldots,\Pi_n)y \in C\})\\ &= \int_0^\infty \alpha y^{-\alpha-1}\,\mathbb{P}((\Pi_1,\ldots,\Pi_n)y \in C)\,dy\end{aligned}$$

- **Then we also have as $x \to \infty$,**

$$\begin{aligned}&\mathbb{P}(x^{-1}(X_1,\ldots,X_n) \in C \mid X_0 > x)\\ &= \frac{\mathbb{P}(x^{-1}(X_1,\ldots,X_n) \in C, x^{-1}X_0 > 1)}{\mathbb{P}(X_0 > x)}\\ &\to \int_1^\infty \alpha y^{-\alpha-1}\,\mathbb{P}((\Pi_1,\ldots,\Pi_n)y \in C)\,dy\\ &= \mathbb{P}(Y\,(\Pi_1,\ldots,\Pi_n) \in C)\,,\end{aligned}$$

where Y has a Pareto distribution with density $\alpha y^{-\alpha-1}$, $y > 1$, independent of (Π_i).

- **The GARCH(1, 1) process**[9] **satisfies a stochastic recurrence equation: for an iid standard normal sequence (Z_t), positive parameters $\alpha_0, \alpha_1, \beta_1$,**

$$\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1)\sigma_{t-1}^2 \,.$$

The process $X_t = \sigma_t Z_t$ is regularly varying with index α satisfying $\mathbb{E}(\alpha_1 Z^2 + \beta_1)^{\alpha/2} = 1$.

- **Then, by Breiman's result (see p. 30), as $x \to \infty$,**

$$\mathbb{P}(X_t > x) \;=\; \mathbb{P}(\sigma_t Z_t > x) \sim \mathbb{E}(Z_0)_+^{\alpha}\,\mathbb{P}(\sigma_0 > x) \sim \mathbb{E}(Z_0)_+^{\alpha}\,c_+ x^{-\alpha}\,,$$

$$\mathbb{P}(X_t \le -x) \;=\; \mathbb{P}(\sigma_t Z_t \le -x) \sim \mathbb{E}(Z_0)_+^{\alpha}\,\mathbb{P}(\sigma_0 > x) \sim \mathbb{E}(Z_0)_+^{\alpha}\,c_+ x^{-\alpha}\,.$$

- **Compare with the tails of a stochastic volatility model on p. 31.**

[9] Bollerslev (1986)

Other examples of regularly varying sequences.

- **α-stable stationary processes** are regularly varying with index $\alpha \in (0, 2)$. Samorodnitsky and Taqqu (1994)

- **Max-stable stationary processes** with Fréchet (Φ_α) marginals are regularly varying with index $\alpha > 0$; see Section 8.

- **The simple stochastic volatility model** (see p. 31) with iid regularly varying noise.

6.2. Limiting representation of a regularly varying stationary sequence.

- Basrak and Segers (2009) **found a useful representation of the limiting measures of the finite-dimensional distributions of a regularly varying sequence.**

- **Theorem.** Assume that (X_t) is an $\mathbb{R}^d$-valued strictly stationary sequence. TFAE to regular variation of (X_t) with index $\alpha > 0$.

(1) There exists a sequence (Y_t) of $\mathbb{R}^d$-valued random vectors with $\mathbb{P}(|Y_0| > x) = x^{-\alpha}$, $x > 1$, such that the following conditional limit relation holds for every $h \geq 0$:

$$\mathbb{P}\big(x^{-1}(X_0, \ldots, X_h) \in \cdot \mid |X_0| > x\big) \xrightarrow{w} \mathbb{P}\big((Y_0, \ldots, Y_h) \in \cdot\big)\,.$$

(2) $|X_0|$ is regularly varying with index α and there exists an $\mathbb{R}^d$-valued process (Θ_t) such that for every $h \geq 0$:

$$\mathbb{P}\big(|X_0|^{-1}(X_0, \ldots, X_h) \in \cdot \mid |X_0| > x\big) \xrightarrow{w} \mathbb{P}\big((\Theta_0, \ldots, \Theta_h) \in \cdot\big)\,.$$

Moreover, the process (Y_t) has representation $Y(\Theta_t)$, where $Y \stackrel{d}{=} |Y_0|$, and Y is independent of (Θ_t).

- Basrak and Segers (2009) refer to (Y_t) and (Θ_t) as the **tail process** and **tail spectral process**, respectively.
- As a particular consequence we observe that the distribution of Θ_0 is the spectral measure of X_0. Indeed, we have

$$\mathbb{P}\big(|X_0|^{-1}X_0 \in \cdot \mid |X_0| > x\big) \xrightarrow{w} \mathbb{P}(\Theta_0 \in \cdot)\,.$$

Also, the distributions of Θ_k, $k \neq 0$, and Θ_0 are in general not the same. In general, $|\Theta_k| \neq 1$ a.s. for $k \neq 0$. Also notice that, in contrast to the sequence (X_t), (Y_t) and (Θ_t) are not strictly stationary.

- **Example.** Assume (X_t) iid regularly varying. Then, for $k \neq 0$, $\varepsilon > 0$,

$$\mathbb{P}(x^{-1}|X_k| > \varepsilon \mid |X_0| > x) = \mathbb{P}(x^{-1}|X_k| > \varepsilon) \to 0\,, \quad x \to \infty\,.$$

Thus $x^{-1}X_k \xrightarrow{P} 0$ and

$$(Y_1, \ldots, Y_k) = (\Theta_1, \ldots, \Theta_k) = 0 \quad \text{a.s.} \quad k \geq 1\,.$$

- **Example.** Assume (X_t) positive strictly stationary regularly varying. Then the following limits exist (extremogram).

$$\begin{aligned}
\rho(h) &= \lim_{x\to\infty} \mathbb{P}(X_h > x \mid X_0 > x) \\
&= \mathbb{P}(Y_h > 1) \\
&= \mathbb{P}(Y\,|\Theta_h| > 1) = \int_1^\infty \alpha y^{-\alpha-1}\,\mathbb{P}(|\Theta_h| > y^{-1})\,dy \\
&= \mathbb{E}\min(1, |\Theta_h|^\alpha)\,, \quad h \geq 1\,.
\end{aligned}$$

6.3. The extremal index of a regularly varying sequence revisited.

- The notion of extremal index originates from Newell (1964), Loynes (1965), O'Brien (1974) and was made precise by Leadbetter (1983).
- The notion of extremal index depends on the definition of a cluster of high level exceedances in the sequence (X_t).
- Although it is "intuitively clear" what an extremal cluster means (many unusually large values appear roughly at the same time) an exact definition is not easy. Various probabilistic definitions exist, depending on some asymptotic theory, in particular on point process convergence Leadbetter et al. (1983), Falk et al. (2004), Embrechts et al. (1997), Section 8.1, while for statistical purposes

(estimation of θ) one has to be pragmatic since one cannot rely on the form of a limiting point process.

- One way of defining a cluster in a sample $X_1, \ldots, X_n$ is to split the sample into blocks of equal size m:

$$X_1, \ldots, X_m\,, X_{m+1}, \ldots, X_{2m}\,, \ldots, X_{(k_n-1)m+1}, \ldots, X_n\,.$$

$k_n = [n/m]$ for the number of (full) blocks.

- For asymptotic theory, it will be important that $m = m_n \to \infty$ and $m = o(n)$.
- Now we simply assume that a block constitutes an extremal cluster if there is at least one exceedance of the high threshold $u = u_n \uparrow x_F$ in this block, and the expected cluster size is given

by

$$\mathbb{E}\Big(\sum_{t=1}^{m} I_{\{X_t>u_n\}} \mid M_m > u_n\Big)$$
$$= \sum_{t=1}^{m} \frac{\mathbb{P}(X_t > u_n\,, M_m > u_n)}{\mathbb{P}(M_m > u_n)}$$
$$= \sum_{t=1}^{m} \frac{\mathbb{P}(X_t > u_n)}{\mathbb{P}(M_m > u_n)} = \frac{m\,\mathbb{P}(X > u_n)}{\mathbb{P}(M_m > u_n)} = \theta_n^{-1}\,.$$

- **We define the extremal index θ as the limit of the reciprocal of the expected cluster size of exceedances of u_n in a block of size m:**

$$\theta = \lim_{n\to\infty} \theta_n = \lim_{n\to\infty} \frac{\mathbb{P}(M_m > u_n)}{m\,\mathbb{P}(X > u_n)}\,,$$

provided this limit exists, and then $\theta \in [0,1]$.[10]

---

[10] For positive θ, Leadbetter (1983) showed that this definition is independent of the particular choice of a threshold sequence $u_n \uparrow x_F$.

- Our next goal is to reduce the calculation of θ to a problem for finitely many random variables $X_1, \ldots, X_l$. We use an argument from Segers (2005); cf. Lemma 2.8 in Davis, Hsing (1995).
- Lemma. Assume $u_n \uparrow x_F = \infty$ and the anti-clustering condition

$$\lim_{l\to\infty} \limsup_{n\to\infty} \mathbb{P}(M_{l,m} > u_n \mid X_0 > u_n) = 0\,,$$

where $M_{s,t} = \max_{s\le i\le t} X_i$ for $s \le t$. Then the following relation holds

$$\lim_{l\to\infty} \limsup_{n\to\infty} \Big|\theta_n - \mathbb{P}(M_l \le u_n \mid X_0 > u_n)\Big| = 0\,,$$

and $\liminf_{n\to\infty} \theta_n > 0$.

- The anti-clustering condition prevents the sequence (X_t) from staying too long above the threshold u_n. “No extremal long-range dependence”.
- We conclude that if $\theta = \lim_{n\to\infty} \theta_n$ exists,

$$\lim_{l\to\infty} \limsup_{n\to\infty} \Big|\theta - \mathbb{P}(M_l \le u_n \mid X_0 > u_n)\Big| = 0\,.$$

in agreement with O’Brien’s (1987) characterization of the extremal index as the limit

$$\theta = \lim_{n\to\infty} \mathbb{P}(M_{l_n} \le u_n \mid X_0 > u_n)$$

for a sequence (l_n) with $l_n = o(n)$, thresholds $u_n \uparrow x_F$ such that $n\overline{F}(u_n) \sim 1$ as $n \to \infty$, provided an asymptotic independence and other conditions hold.

Example. Basrak and Segers (2009) For a positive strictly stationary regularly varying sequence (X_t), satisfying the anti-clustering condition, we know that the limits exists:

$$\lim_{n\to\infty} \mathbb{P}(M_l \le u_n \mid X_0 > u_n) = \mathbb{P}\big(\max_{t=1,\dots,l} Y_t \le 1\big)$$
$$\downarrow \mathbb{P}\big(\sup_{t\ge 1} Y_t \le 1\big)\,, \quad l \to \infty\,.$$

Hence the extremal index is given by

$$\begin{aligned}\theta &= \mathbb{P}\Big(\sup_{t\ge 1} Y_t \le 1\Big) = \mathbb{P}\Big(Y \max_{t\ge 1} \Theta_t \le 1\Big)\\ &= \int_1^\infty \alpha\, y^{-\alpha-1}\, \mathbb{P}\Big(\sup_{t\ge 1} \Theta_t \le y^{-1}\Big)\, dy\\ &= \mathbb{E}\Big[1 - \sup_{t\ge 1} \Theta_t^\alpha\Big]_+\\ &= \mathbb{E}\Big[\sup_{t\ge 0} \Theta_t^\alpha - \sup_{t\ge 1} \Theta_t^\alpha\Big]\,.\end{aligned}$$

Example. Recall that the solution to the stochastic recurrence equation $X_t = A_t X_{t-1} + B_t$, $t \in \mathbb{Z}$, is regularly varying (see p. 52) and the finite-dimensional distributions satisfy

$$\mathbb{P}(x^{-1}(X_1, \ldots, X_n) \in C \mid X_0 > x) \to \mathbb{P}(Y\,(\Pi_1, \ldots, \Pi_n) \in C)\,,$$

where Y has a Pareto distribution with density $\alpha y^{-\alpha-1}$, $y > 1$, independent of $\Pi_i = A_1 \cdots A_i$, $i \geq 1$. Hence

$$(\Theta_0, \Theta_1, \ldots, \Theta_n) \stackrel{d}{=} (1, \Pi_1, \ldots, \Pi_n)$$

and

$$\theta = \mathbb{E}[\sup_{t \geq 0} \Pi_t^\alpha - \sup_{t \geq 1} \Pi_t^\alpha]\,.$$

7. The extremogram - an analog of the autocorrelation function

- **Goals.**

  Find measure of serial extremal dependence in a strictly stationary sequence (X_t).

- **Is there an analog of the sample autocorrelation function for serial extremal dependence?**

- **Estimation of this function should be "uncomplicated" and graphical visualization should be possible.**

7.1. A motivating example: the tail dependence coefficient.

- In risk management,[11] the (upper) *tail dependence coefficient* gained some popularity:

$$\lambda_h = \lim_{x\to\infty} \mathbb{P}(X_h > x \mid X_0 > x)\,, \quad h \geq 1.$$

- If $\lambda_h = 0$: no upper tail dependence as, for example, in a non-trivial Gaussian stationary sequence.

- The definition of λ_h requires that the limit exists.

- A sufficient condition is *regular variation* of (X_t), i.e., the finite-dimensional distributions of (X_t) are regularly varying.

---

[11] For example, McNeil et al. [40].

Recall: Regularly varying strictly stationary time series

- A strictly stationary $\mathbb{R}^d$-valued sequence $(\mathbf{X}_t)$ is regularly varying with index $\alpha > 0$: there exist limit measures $\mu_h \neq o$, $h \geq 1$, such that

$$\frac{\mathbb{P}(x^{-1}(X_1, \ldots, X_h) \in \cdot)}{\mathbb{P}(|X_0| > x)} \xrightarrow{v} \mu_h(\cdot)\,, \quad x \to \infty\,.$$

- Equivalently, for any sequence (a_n) satisfying $\mathbb{P}(|\mathbf{X}_1| > a_n) \sim n^{-1}$ there exist limit measures $\mu_h \neq o$, $h \geq 1$, such that

$$n\,\mathbb{P}(a_n^{-1}(\mathbf{X}_1, \ldots, \mathbf{X}_h) \in \cdot) \xrightarrow{v} \mu_h(\cdot)\,.$$

The upper tail dependence coefficient revisited

- Let (X_t) be a strictly stationary real-valued sequence which is regularly varying with index $\alpha > 0$. Let $A = B = (1, \infty)$ and $\mathrm{Y}_h = (X_0, \ldots, X_h)$. Then

$$\lambda_h = \lim_{x \to \infty} \mathbb{P}(X_h > x \mid X_0 > x)$$

The upper tail dependence coefficient revisited

- Let (X_t) be a strictly stationary real-valued sequence which is regularly varying with index $\alpha > 0$. Let $A = B = (1, \infty)$ and $\mathbf{Y}_h = (X_0, \ldots, X_h)$. Then

$$\lambda_h = \lim_{x\to\infty} \mathbb{P}(X_h > x \mid X_0 > x)$$
$$= \lim_{x\to\infty} \frac{\mathbb{P}(x^{-1}\mathbf{Y}_h \in A \times \mathbb{R}_0^{h-1} \times B)}{\mathbb{P}(x^{-1}\mathbf{Y}_h \in A \times \mathbb{R}_0^{h})}$$

The upper tail dependence coefficient revisited

- Let (X_t) be a strictly stationary real-valued sequence which is regularly varying with index $\alpha > 0$. Let $A = B = (1, \infty)$ and $\mathrm{Y}_h = (X_0, \ldots, X_h)$. Then

$$\begin{aligned}
\lambda_h &= \lim_{x\to\infty} \mathbb{P}(X_h > x \mid X_0 > x) \\
&= \lim_{x\to\infty} \frac{\mathbb{P}(x^{-1}\mathrm{Y}_h \in A \times \mathbb{R}_0^{h-1} \times B)}{\mathbb{P}(x^{-1}\mathrm{Y}_h \in A \times \mathbb{R}_0^{h})} \\
&= \lim_{x\to\infty} \frac{\mathbb{P}(x^{-1}\mathrm{Y}_h \in A \times \mathbb{R}_0^{h-1} \times B)/\mathbb{P}(|X_0| > x)}{\mathbb{P}(x^{-1}\mathrm{Y}_h \in A \times \mathbb{R}_0^{h-1})/\mathbb{P}(|X_0| > x)}
\end{aligned}$$

The upper tail dependence coefficient revisited

- Let (X_t) be a strictly stationary real-valued sequence which is regularly varying with index $\alpha > 0$. Let $A = B = (1, \infty)$ and $\mathbf{Y}_h = (X_0, \ldots, X_h)$. Then

$$\begin{aligned}
\lambda_h &= \lim_{x\to\infty} \mathbb{P}(X_h > x \mid X_0 > x) \\
&= \lim_{x\to\infty} \frac{\mathbb{P}(x^{-1}\mathbf{Y}_h \in A \times \mathbb{R}_0^{h-1} \times B)}{\mathbb{P}(x^{-1}\mathbf{Y}_h \in A \times \mathbb{R}_0^{h})} \\
&= \lim_{x\to\infty} \frac{\mathbb{P}(x^{-1}\mathbf{Y}_h \in A \times \mathbb{R}_0^{h-1} \times B)/\mathbb{P}(|X_0| > x)}{\mathbb{P}(x^{-1}\mathbf{Y}_h \in A \times \mathbb{R}_0^{h-1})/\mathbb{P}(|X_0| > x)} \\
&= \frac{\mu_{h+1}(A \times \mathbb{R}_0^{h-1} \times B)}{\mu_{h+1}(A \times \mathbb{R}_0^{h})}.
\end{aligned}$$

7.2. **Definition** Davis, M. (2009).

- For an $\mathbb{R}^d$-valued strictly stationary regularly varying sequence (X_t) and Borel sets A, B bounded away from zero the *extremogram*[12] is the limiting function

$$
\begin{aligned}
n\,\mathbb{P}(a_n^{-1}\mathrm{X}_0 \in A\,,\, a_n^{-1}\mathrm{X}_h \in B) &\rightarrow \mu_{h+1}(A \times \mathbb{R}_0^{d(h-1)} \times B) \\
&= \gamma_{AB}(h)\,, \quad h \geq 0\,.
\end{aligned}
$$

- Extremograms of the type

$$
\begin{aligned}
\rho_{AB}(h) &= \lim_{n\to\infty} \mathbb{P}(a_n^{-1}\mathrm{X}_h \in B \mid a_n^{-1}\mathrm{X}_0 \in A) \\
&= \lim_{n\to\infty} \frac{\mathbb{P}(a_n^{-1}\mathrm{X}_0 \in A\,, \quad a_n^{-1}\mathrm{X}_h \in B)}{\mathbb{P}(a_n^{-1}\mathrm{X}_0 \in A)}\,, \quad h \geq 0\,,
\end{aligned}
$$

for sets A bounded away from zero are of main interest.

[12] A possible choice of (a_n) is $\mathbb{P}(|X_0| > a_n) \sim n^{-1}$.

- Since

$$n\,\mathrm{cov}(I_{\{a_n^{-1}\mathbf{X}_0\in A\}}, I_{\{a_n^{-1}\mathbf{X}_h\in B\}})$$

$$= n\,[\mathbb{P}(a_n^{-1}\mathbf{X}_0 \in A\,, a_n^{-1}\mathbf{X}_h \in B) - \mathbb{P}(a_n^{-1}\mathbf{X}_0 \in A)\,\mathbb{P}(a_n^{-1}\mathbf{X}_0 \in B)]$$

$$\to \gamma_{AB}(h) \geq 0\,,$$

the matrix function $\begin{pmatrix} \gamma_{AA}(h) & \gamma_{AB}(h) \\ \gamma_{BA}(h) & \gamma_{BB}(h) \end{pmatrix}$ is a *covariance matrix function.*

- One can use the notions of time series analysis to describe the extremal dependence structure in a strictly stationary sequence.

- One can define *long/short range dependence* in some meaningful way or the *spectral distribution* for extremal events in a strictly stationary sequence, or one can use the extremogram to justify the *selection of a particular time series model.*

- For example, the autocorrelation functions of a GARCH$(1,1)$ process and a stochastic volatility model can be very similar, the extremogram of a stochastic volatility model vanishes while it does not for a GARCH$(1,1)$ process.

7.3. **Examples.** **Take $A = B = (1, \infty)$.**

- **The AR(1) process $X_t = \phi X_{t-1} + Z_t$ with iid symmetric regularly varying noise (Z_t) with index α and $\phi \in (-1, 1)$ has the extremogram**

$$\gamma_{AA}(h) = \max(0, (\text{sign}(\phi))^h |\phi|^{\alpha h}) \,.$$

Short serial extremal dependence

- **The extremogram of a GARCH(1,1) process is not very explicit, but $\gamma_{AA}(h)$ decays exponentially fast to zero. This is in agreement with the geometric β-mixing property of GARCH.**

Short serial extremal dependence

- **The stochastic volatility model with stationary Gaussian ($\log \sigma_t$) and iid regularly varying (Z_t) with index $\alpha > 0$ has extremogram $\gamma_{AA}(h) = 0$ as in the iid case.**

  No serial extremal dependence

7.4. The sample extremogram.

- Let (X_t) be strongly mixing (possibly vector-valued) regularly varying.
- Assume $m = m_n \to \infty$ and $m_n/n \to 0$, and $a_m \to \infty$ satisfies $\mathbb{P}(|\mathrm{X}_0| > a_m) \sim m^{-1}$. Then

$$\widehat{P}_m(C) = \frac{m}{n} \sum_{t=1}^{n} I_{\{\mathrm{X}_t/a_m \in C\}}$$

is a consistent estimator of

$$\mu_1(C) = \lim_{m \to \infty} m\, \mathbb{P}(\mathrm{X}_0/a_m \in C)\,.$$

- In particular,

$$\mathbb{E}\widehat{P}_m(C) \;\rightarrow\; \mu_1(C)\,,$$

$$\mathrm{var}(\widehat{P}_m(C)) \;\sim\; \frac{m}{n}\sigma^2(C) = \frac{m}{n}\Big[\mu_1(C) + 2\sum_{h=1}^{\infty}\tau_h(C)\Big]$$

for

$$\tau_h(C) = \mu_{h+1}(C \times \mathbb{R}_0^{d(h-1)} \times C)\,.$$

- For μ_1-continuity sets C bounded away from zero,

$$\Big(\frac{n}{m}\Big)^{1/2}\,[\widehat{P}_m(C) - m\,\mathbb{P}(a_m^{-1}\mathrm{X}_0 \in C)] \stackrel{d}{\rightarrow} N(0,\sigma^2(C))\,.$$

(pre-asymptotic central limit theorem).

- An analogous result holds for finitely many sets $C_1,\ldots,C_h$.

- The ratio sample extremogram

$$\widehat{\rho}_{AB}(h) = \frac{\frac{m}{n}\sum_{t=1}^{n-h} I_{\{a_m^{-1}\mathrm{X}_{t+h}\in B, a_m^{-1}\mathrm{X}_t\in A\}}}{\frac{m}{n}\sum_{t=1}^{n} I_{\{a_m^{-1}\mathrm{X}_t\in A\}}}$$

$$= \frac{\sum_{t=1}^{n-h} I_{\{a_m^{-1}\mathrm{X}_{t+h}\in B, a_m^{-1}\mathrm{X}_t\in A\}}}{\sum_{t=1}^{n} I_{\{a_m^{-1}\mathrm{X}_t\in A\}}}, \quad h \geq 0,$$

estimates

$$\rho_{AB}(h) = \lim_{n\to\infty} \mathbb{P}(a_n^{-1}\boldsymbol{X}_h \in B \mid a_n^{-1}\boldsymbol{X}_0 \in A)$$

$$= \frac{\mu_{h+1}(A \times \overline{\mathbb{R}}_0^{d(h-1)} \times B)}{\mu_{h+1}(A \times \overline{\mathbb{R}}_0^{dh})}, \quad h \geq 0.$$

- **Pre-asymptotic** limit theory for the ratio estimator follows from the previous central limit theory

$$\Big(\frac{n}{m}\Big)^{1/2}\Big(\widehat{\rho}_{AB}(i) - \rho_{AB:m}(i)\Big)_{i=0,\dots,h} \stackrel{d}{\rightarrow} N(0,\Sigma)\,,$$

where $\rho_{AB:m}(h) = \mathbb{P}(a_m^{-1}X_h \in B \mid a_m^{-1}X_0 \in A)$.

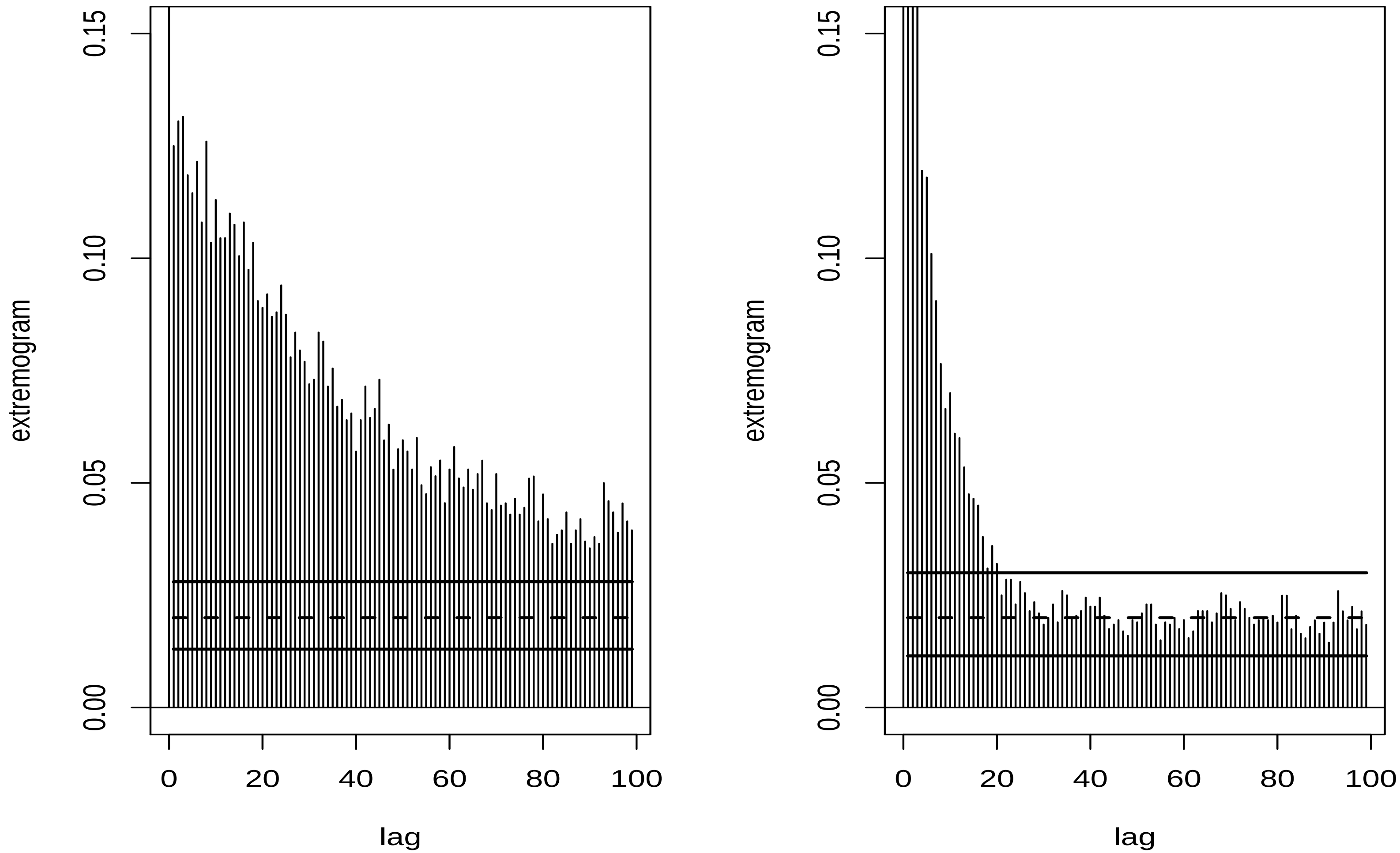

FIGURE 17. The ratio sample extremogram with sample size $\boldsymbol{n = 100,000}$ for the GARCH$\boldsymbol{(1,1)}$ (left) and stochastic volatility processes (right). $\boldsymbol{A = B = (1,\infty)}$.

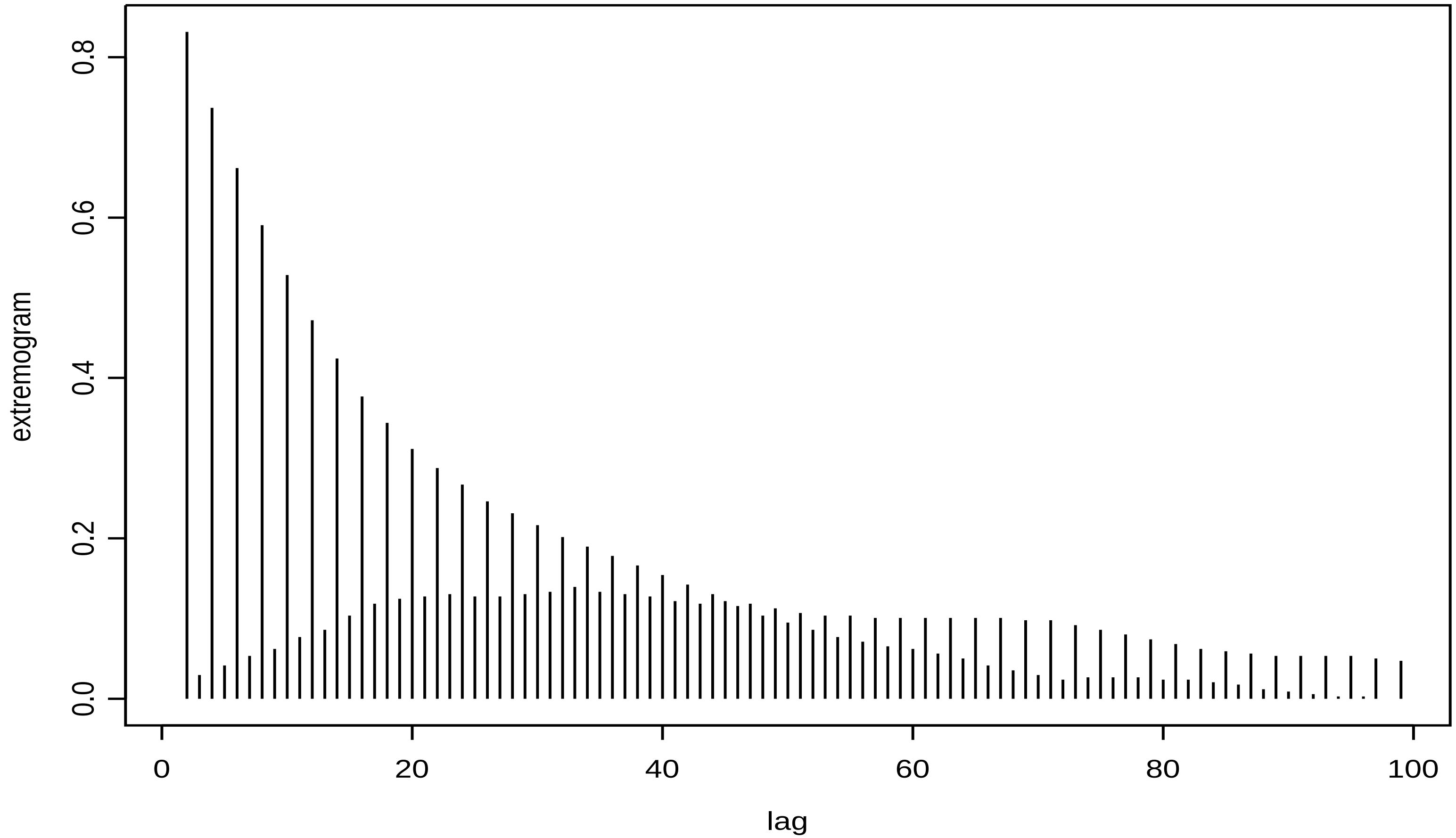

FIGURE 18. Ratio sample extremogram with $\boldsymbol{A} = \boldsymbol{B} = (\mathbf{1}, \boldsymbol{\infty})$ for $\mathbf{5}$ minute returns of USD-DEM foreign exchange rates. The extremogram alternates between large values at even lags and small ones at odd lags. This is an indication of AR behavior with negative leading coefficient.

Problems

(1) The central limit theorem for the ratio sample estimator is pre-asymptotic. (For applications, the pre-asymptotic centering $\rho_{AB:m}(h) = \mathbb{P}(a_m^{-1} X_h \in B \mid a_m^{-1} X_0 \in A)$ is more relevant than its limit $\rho_{AB}(h)$.)

(2) The asymptotic variance-covariance structure of the ratio sample estimator depends on expressions which are unknown.

Two methods to overcome (2):

random permutations and stationary bootstrap.

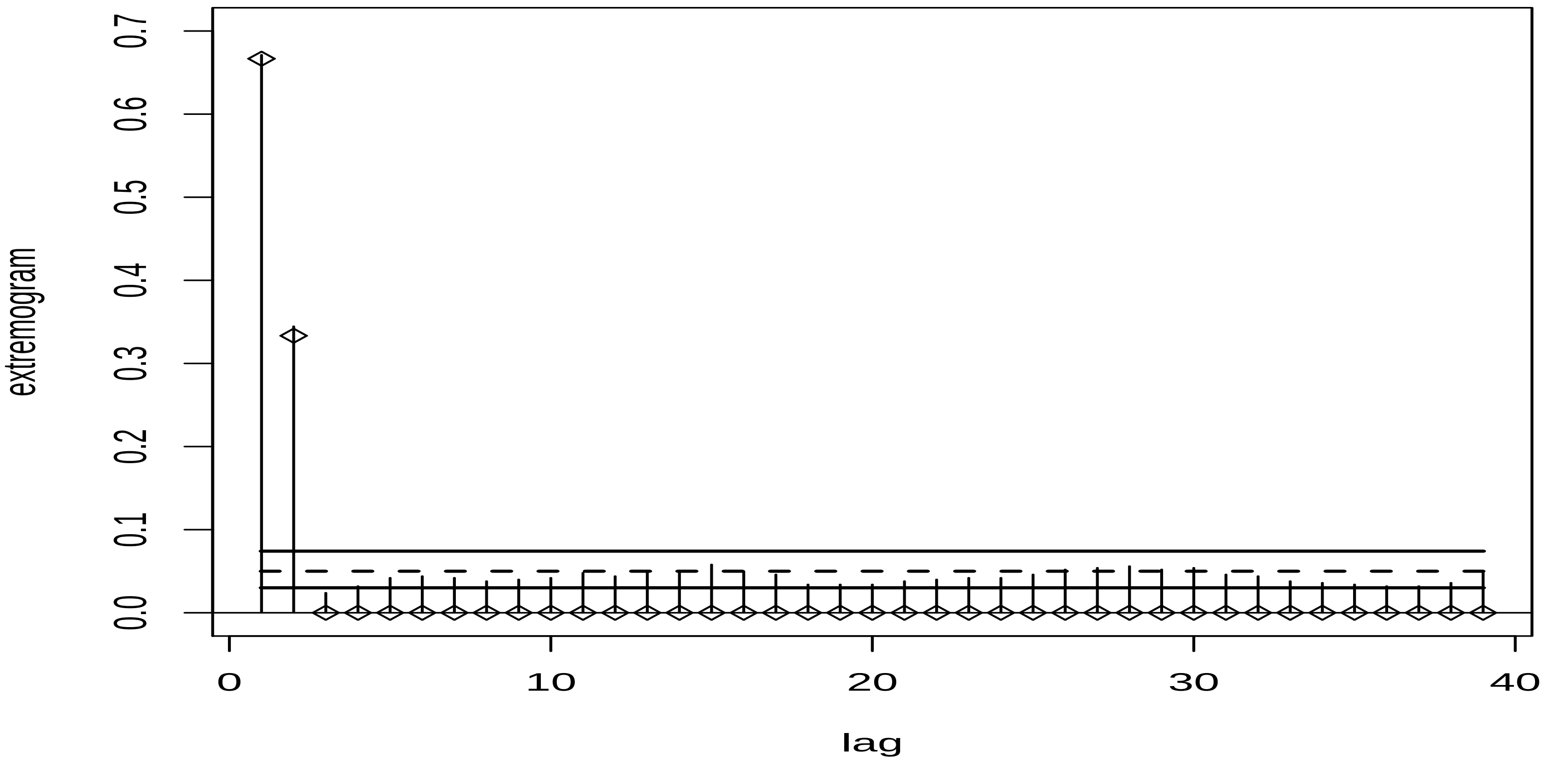

Figure 19. Sample extremogram for the max-moving average MMA(2) process $\boldsymbol{X_t} = \mathbf{max}(\boldsymbol{Z_t}, \boldsymbol{Z_{t-1}}, \boldsymbol{Z_{t-2}})$ for an iid positive regularly varying sequence $(\boldsymbol{Z_t})$, $\boldsymbol{A} = \boldsymbol{B} = (\mathbf{1}, \boldsymbol{\infty})$. The diamonds superimposed on the figure represent the population extremogram values. Confidence bands are based on random permutations of the data.

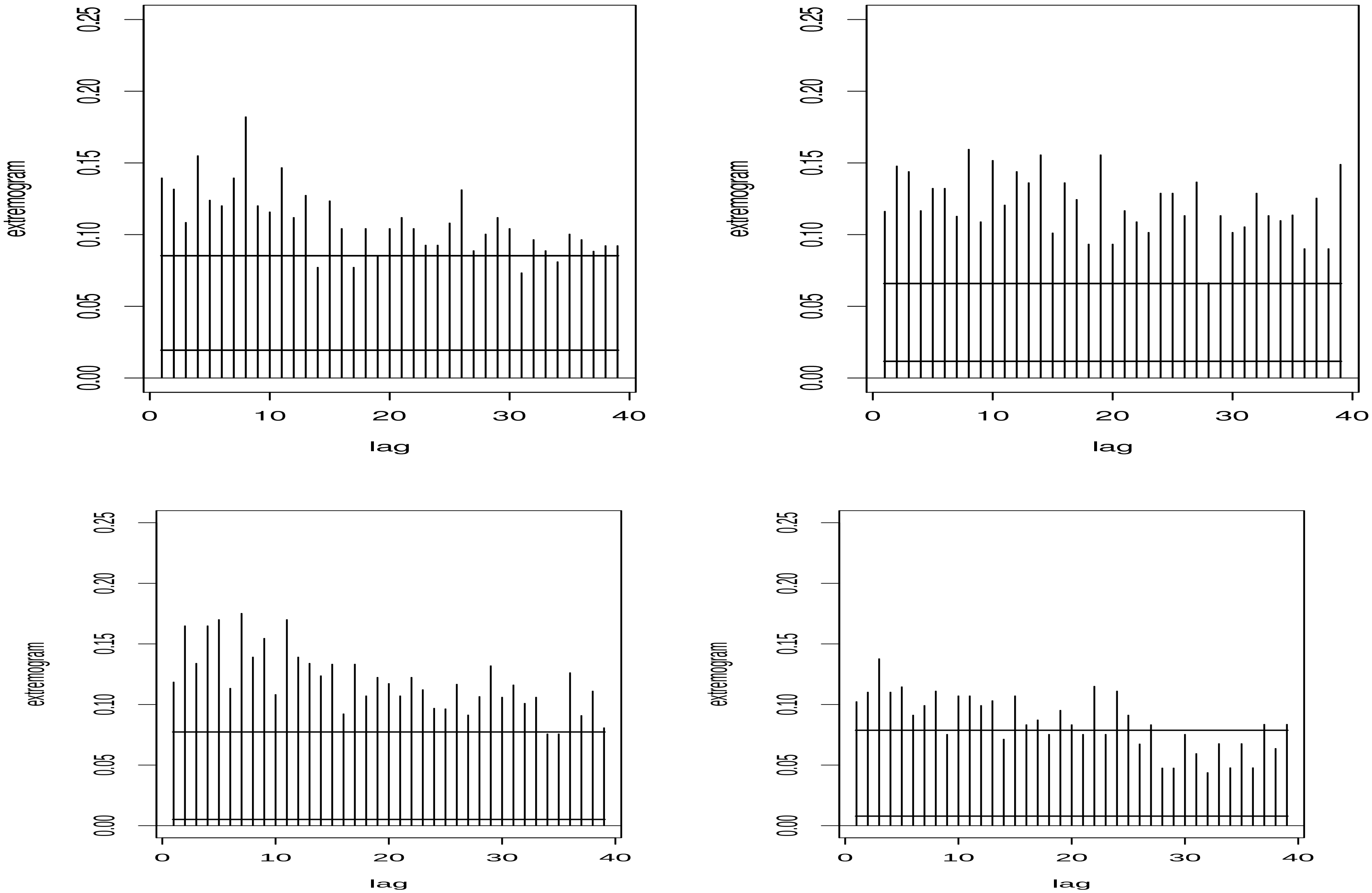

Figure 20. The sample extremogram for the lower tail of the FTSE (top left), S&P500 (top right), DAX (bottom left) and Nikkei. The bold lines represent sample extremograms based on random permutations of the data.

7.5. Bootstrapping the sample extremogram.

- The stationary bootstrap Politis and Romano (1994) is well suited for the sample extremogram.
- Stationary bootstrap setup. Have data $X_1, \ldots, X_n$ and construct a bootstrap (BS) sample as follows: Generate
  $K_1, K_2, \ldots$ iid uniform on $1, \ldots, n$
  $L_1, L_2, \ldots$ iid geometric(p_n)
- The BS sample is given by the first n values (in the circular sense) in the sequence

$$X_{K_1}, \ldots, X_{K_1+L_1-1}, X_{K_2}, \ldots, X_{K_2+L_2-1}, \ldots.$$

- Mean block size $1/p_n \to \infty$, mean number of blocks $np_n \to \infty$.

- **The stationary bootstrap ratio sample extremogram is consistent:**

$$\mathbb{P}^*\Big(\big(\frac{n}{m}\big)^{1/2}\Big(\widehat{\rho}^*_{AB}(i) - \widehat{\rho}_{AB}(i)\Big)_{i=0,\ldots,h} \in A\Big) \stackrel{P}{\to} \Phi_{0,\Sigma}(A)\,,$$

for continuity sets A of the limit distribution $\Phi_{0,\Sigma}$.

David, M., Cribben (2012)

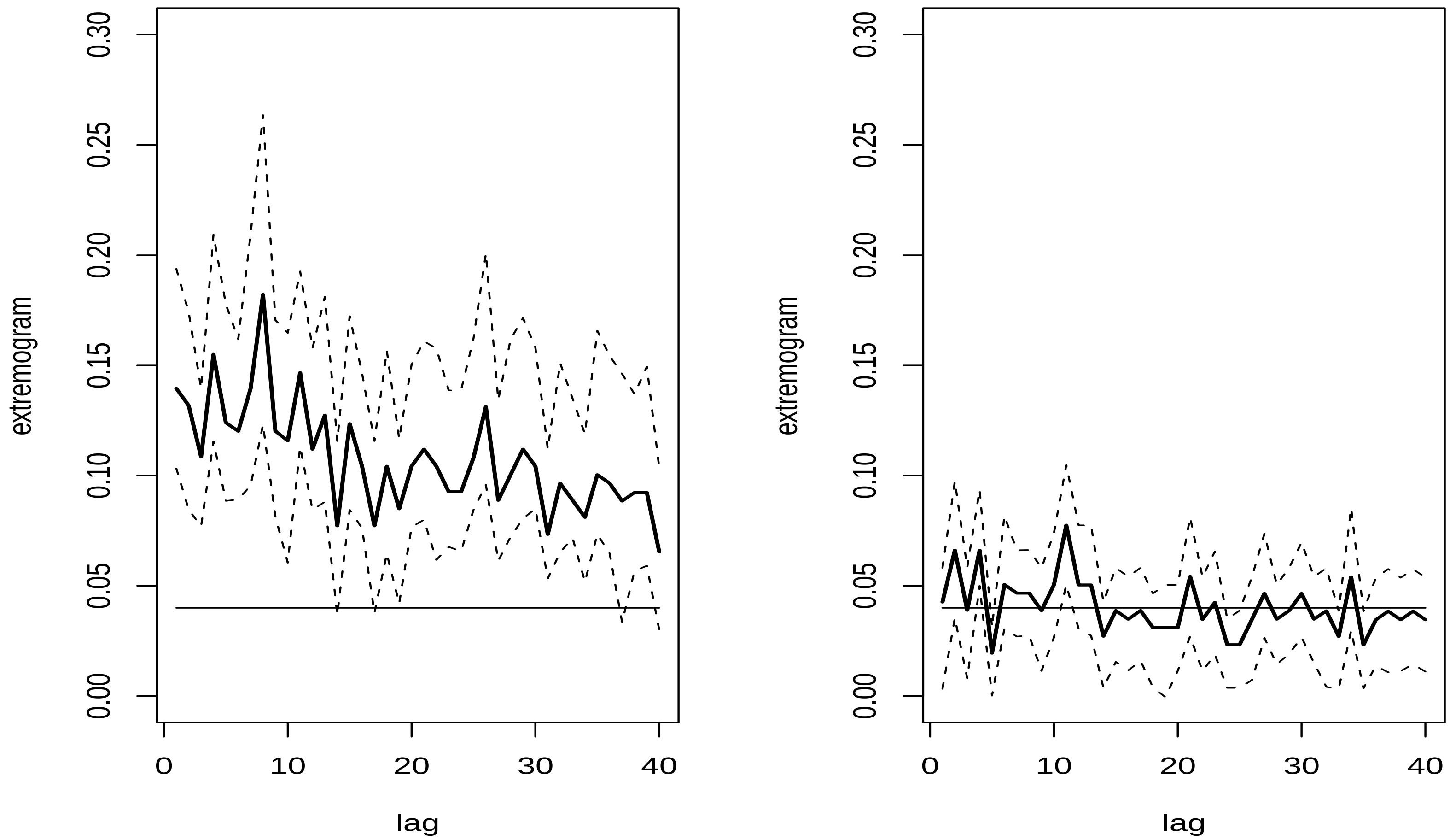

FIGURE 21. Left: 95% bootstrap confidence bands for pre-asymptotic extremogram of 6440 daily FTSE log-returns. Mean block size 200. Right: For the residuals of a fitted GARCH$(\mathbf{1},\mathbf{1})$ model.

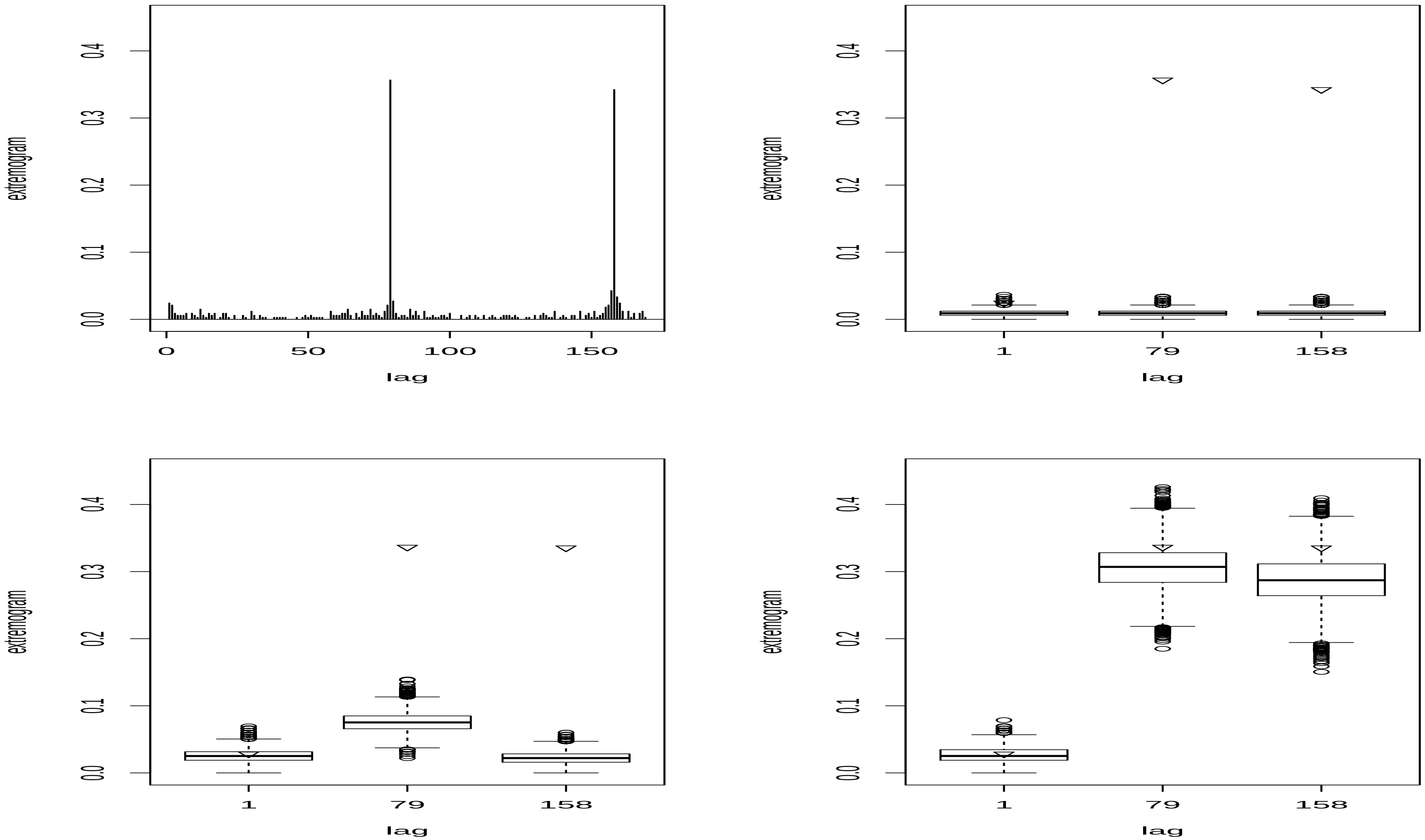

FIGURE 22. Top left: Sample extremogram for **5**-minute GE log-returns. Boxplots at lags 1, 79, 158 using permutations (top right) and bootstrap with mean block size **50** and **200** (bottom).

7.6. Variations on the theme Davis, M., Cribben (2012ab), Davis, M., Zhao (2013).

Cross-extremogram Consider a strictly stationary bivariate regularly varying time series $((X_t, Y_t))_{t\in\mathbb{Z}}$.

For two sets A and B bounded away from 0, the cross-extremogram

$$\gamma_{AB}(h) = \lim_{x\to\infty} \mathbb{P}(Y_h \in x\,B \mid X_0 \in x\,A)\,, \quad h \geq 0\,,$$

is an extremogram based on the two-dimensional sets $A \times \mathbb{R}$ and $\mathbb{R} \times B$.

The corresponding sample cross-extremogram for the time series $((X_t, Y_t))_{t\in\mathbb{Z}}$:

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

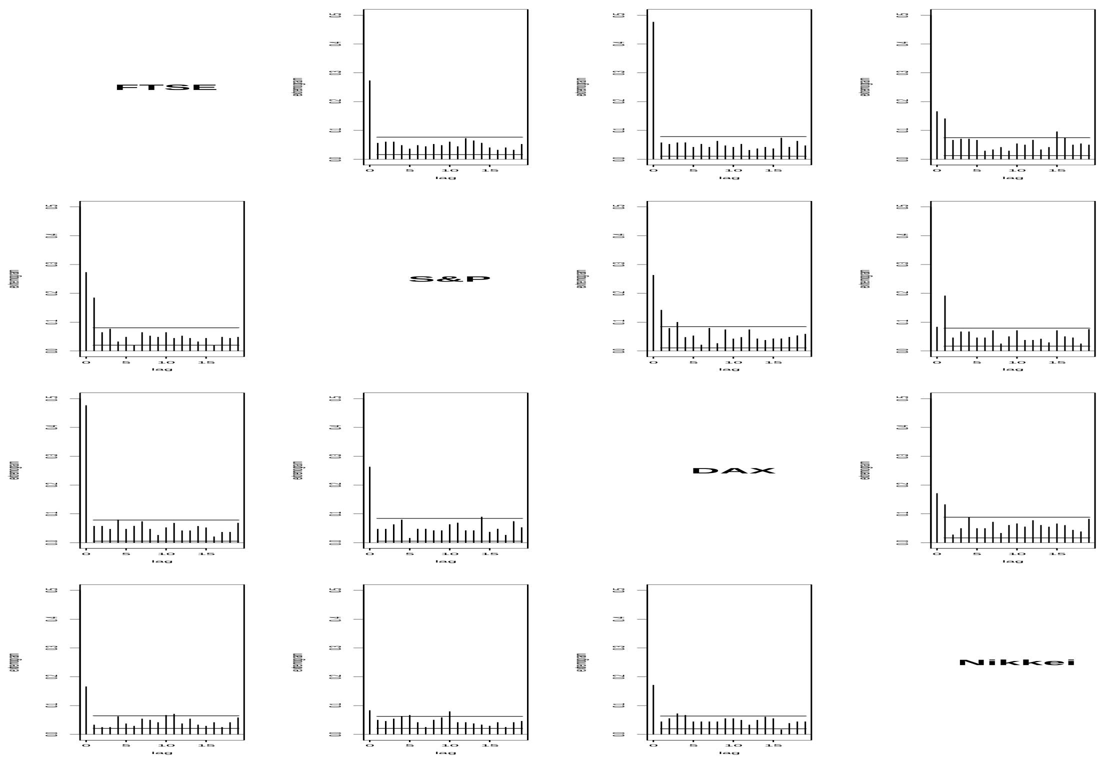

FIGURE 23. The sample cross-extremograms for the filtered FTSE, S&P, DAX and Nikkei series. For the first row, $(\boldsymbol{X_t})$ is the filtered FTSE and $(\boldsymbol{Y_t})$ are the filtered S&P, DAX and Nikkei (from left to right). For the second, third and fourth rows, the $\boldsymbol{X_t}$'s are the filtered S&P, DAX and Nikkei series, respectively.

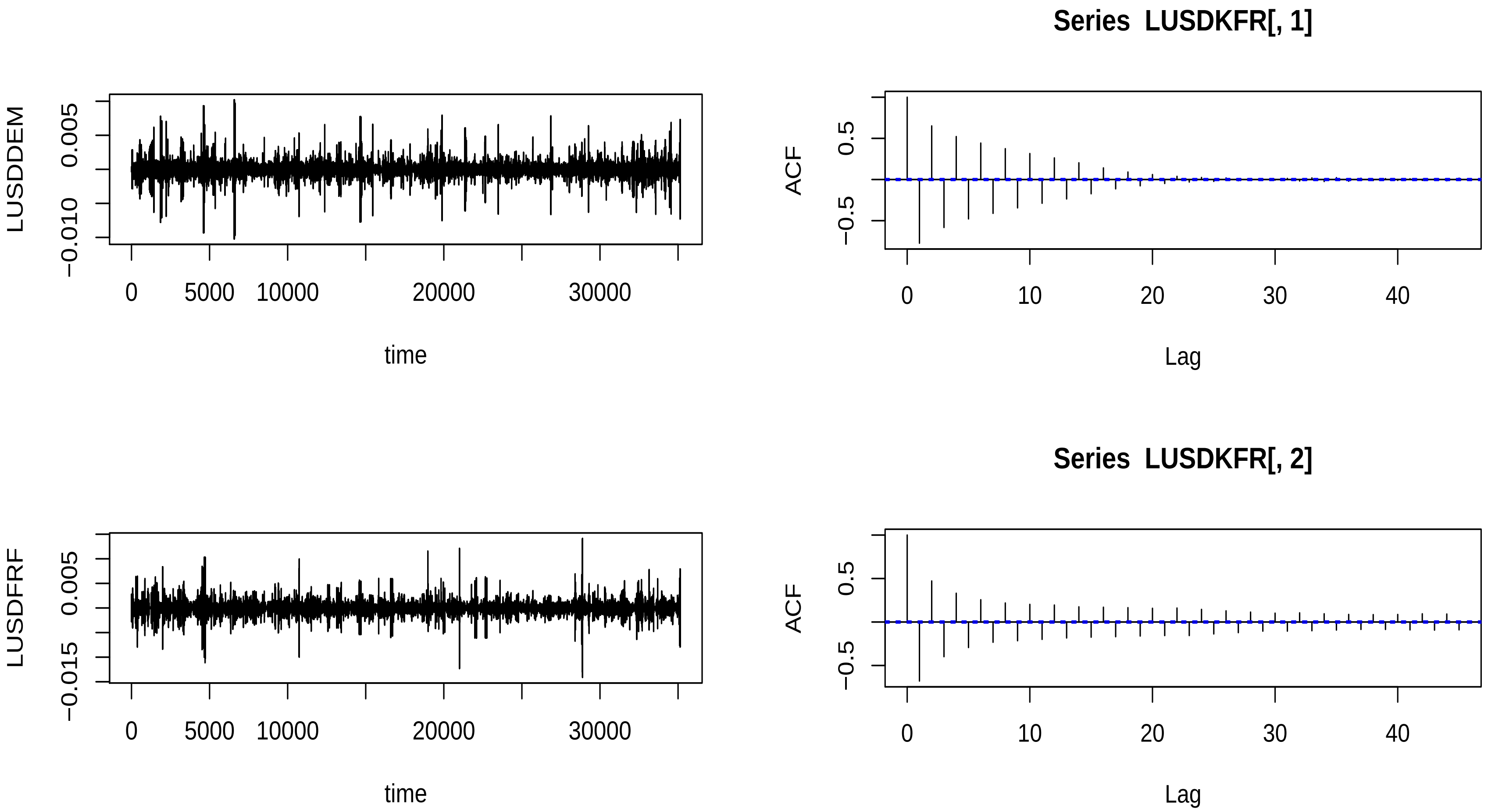

FIGURE 24. 5-minute log-return series of FX data and their sample autocorrelation functions: USD-FRF (top), USD-DEM (bottom).

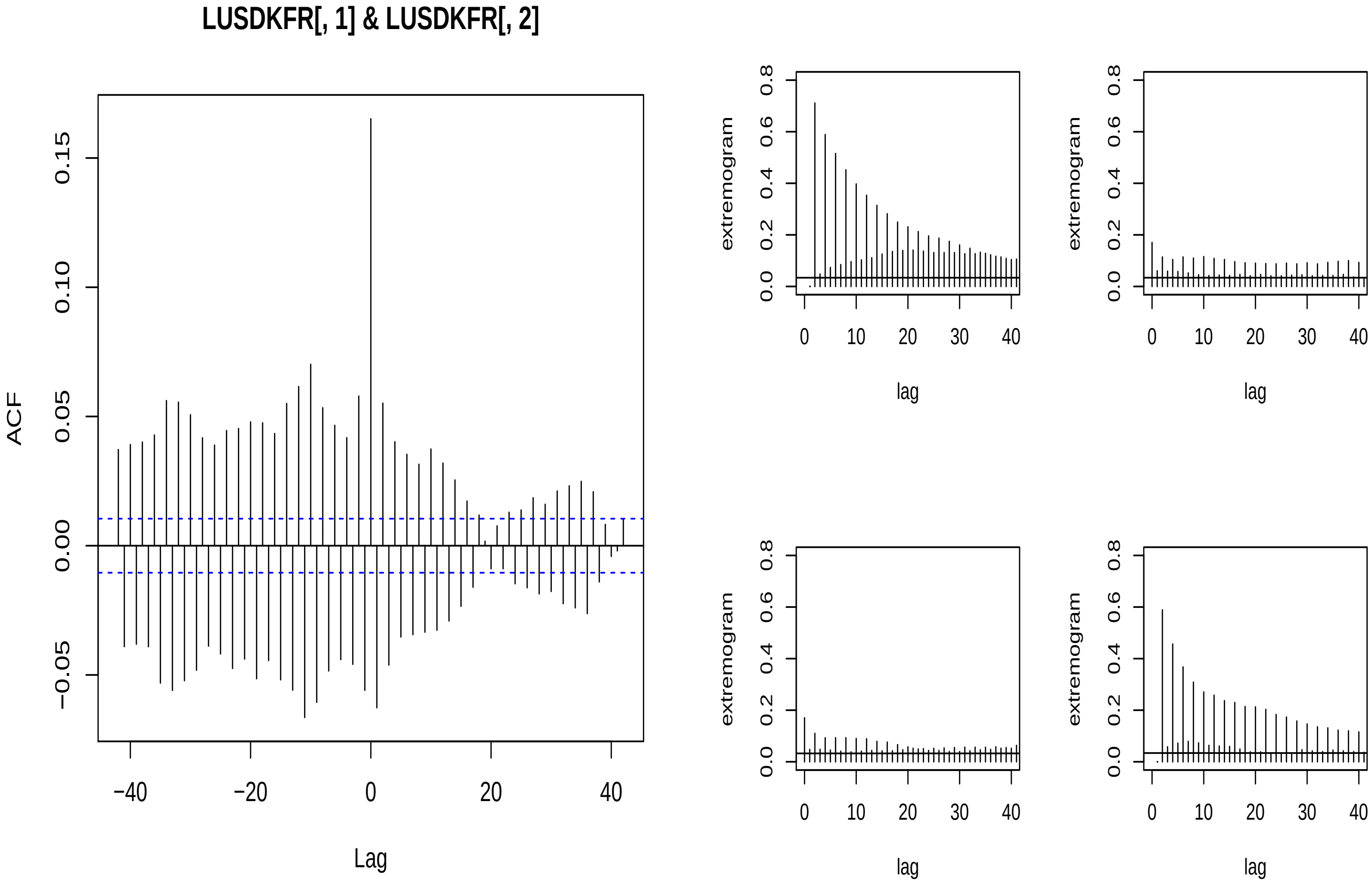

FIGURE 25. The sample cross-correlation function of the FX data (left) and their extremograms (right)

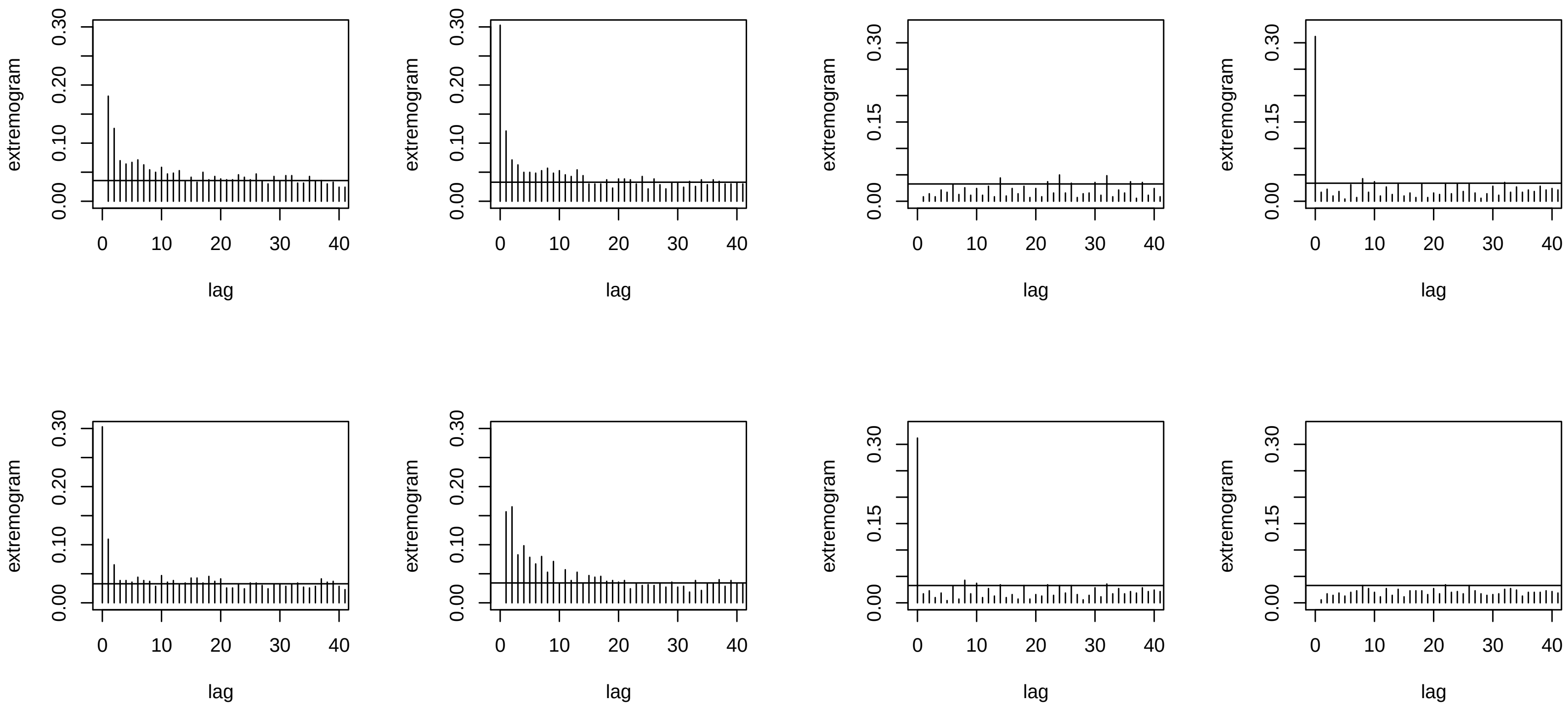

Figure 26. Left: Extremograms for residuals after an AR fit. Right: Extremograms of residuals after an AR-GARCH fit.

The extremogram of return times between rare events.

- We say that X_t is extreme if $X_t \in xA$ for a set A bounded away from zero and large x.
- If the return times were truly iid, the successive waiting times between extremes should be iid geometric.
- Using the histogram of waiting times, this hypothesis can be verified.
- The corresponding return times extremogram is given by

$$\rho_A(h) = \lim_{x\to\infty} \mathbb{P}(X_1 \notin xA, \ldots, X_{h-1} \notin xA, X_h \in xA \mid X_0 \in xA)$$

$$= \frac{\mu_{h+1}(A \times (A^c)^{h-1} \times A)}{\mu_{h+1}(A \times \overline{\mathbb{R}}_0^{dh})}, \quad h \geq 0\,.$$

- **The return times sample extremogram is then defined as**

$$\widehat{\rho}_A(h) = \frac{\sum_{t=1}^{n-h} I_{\{X_{t+h}\in a_m A, X_{t+h-1}\notin a_m A, \ldots, X_{t+1}\notin a_m A, X_t\in a_m A\}}}{\sum_{t=1}^{n} I_{\{X_t\in a_m A\}}}, \quad h < n.$$

- **It is consistent, asymptotically normal, and the stationary bootstrap version has the same properties.**

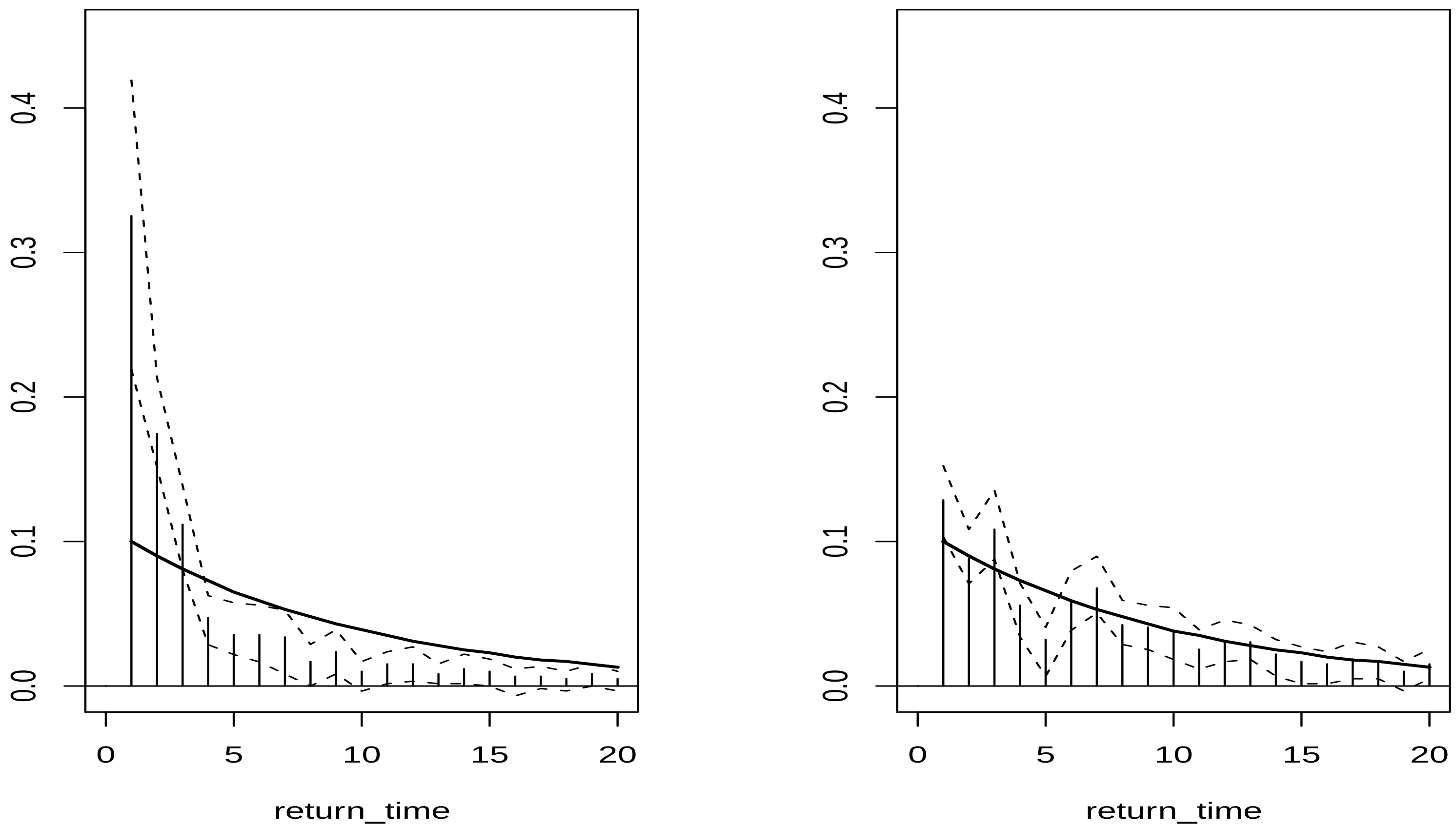

FIGURE 27. Left: Return times sample extremogram for extreme events with $\boldsymbol{A} = \mathbb{R}\backslash[\boldsymbol{\xi_{0.05}}, \boldsymbol{\xi_{0.95}}]$ for the daily log-returns of BAC using bootstrapped confidence intervals (dashed lines), geometric probability mass function (light solid). Right: The corresponding extremogram for the residuals of a fitted GARCH$(\mathbf{1}, \mathbf{1})$ model (right).

7.7. Frequency domain analysis M. and Zhao (2012).

- The extremogram for a given set A bounded away from zero

$$\rho_A(h) = \lim_{n\to\infty} \mathbb{P}(a_n^{-1} X_h \in A \mid a_n^{-1} X_0 \in A)\,, \quad h \geq 0\,,$$

is a correlation function.

- Therefore one can define the spectral density

$$f_A(\lambda) = 1 + 2\sum_{h=1}^{\infty} \cos(\lambda\, h)\, \rho_A(h) = \sum_{h=-\infty}^{\infty} \mathrm{e}^{-i\,\lambda\, h}\, \rho_A(h)\,.$$

- We also introduce the periodogram for $\lambda \in (0,\pi)$:

$$\widehat{f}_{nA}(\lambda) = \frac{I_{nA}(\lambda)}{\widehat{P}_m(A)} = \frac{\frac{m}{n}\Big|\sum_{t=1}^{n} \mathrm{e}^{-it\lambda} I_{\{a_m^{-1}\mathrm{X}_t \in A\}}\Big|^2}{\frac{m}{n}\sum_{t=1}^{n} I_{\{a_m^{-1}\mathrm{X}_t \in A\}}}\,.$$

- One has $\mathbb{E} I_{nA}(\lambda)/\mu_1(A) \to f_A(\lambda)$ for $\lambda \in (0, \pi)$.
- As in classical time series analysis, $\widehat{f}_{nA}(\lambda)$ is not a consistent estimator of $f_A(\lambda)$: for distinct (fixed or Fourier) frequencies λ_j, and iid standard exponential E_j,

$$(\widehat{f}_{nA}(\lambda_j))_{j=1,\dots,h} \xrightarrow{d} (f_A(\lambda_j) E_j)_{j=1,\dots,h}\,.$$

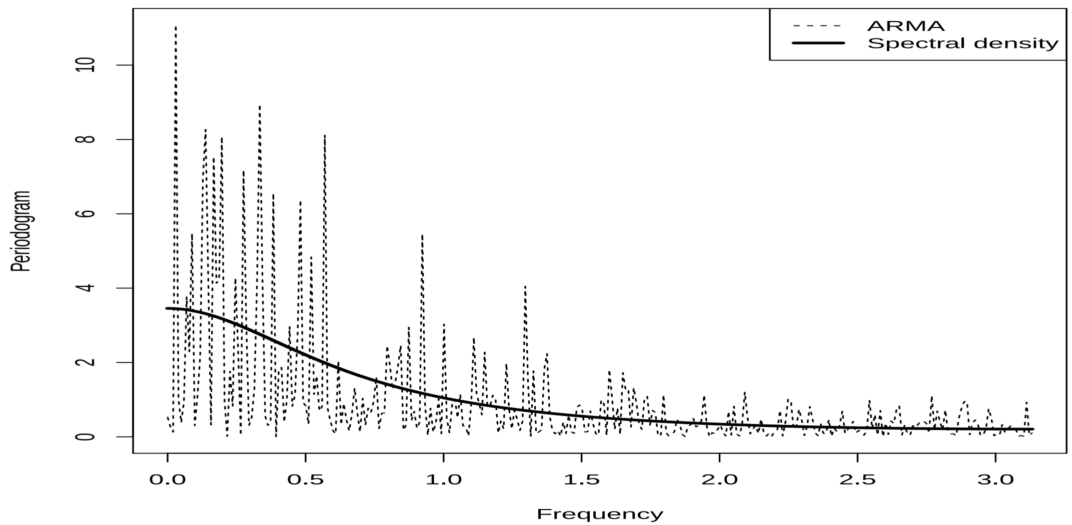

FIGURE 28. Sample extremogram and periodogram for ARMA(1,1) process with student(4) noise. $A = (1, \infty)$

- **Smoothed versions of the periodogram converge to $f(\lambda)$:**

  If $w_n(j) \geq 0$, $|j| \leq s_n \to \infty$, $s_n/n \to 0$, $\sum_{|j|\leq s_n} w_n(j) = 1$ and $\sum_{|j|\leq s_n} w_n^2(j) \to 0$ (e.g. $w_n(j) = 1/(2s_n+1)$) then for any distinct Fourier frequencies λ_j such that $\lambda_j \to \lambda$,

$$\sum_{|j|\leq s_n} w_n(j)\,\widehat{f}_{nA}(\lambda_j) \xrightarrow{P} f_A(\lambda)\,, \quad \lambda \in (0,\pi)\,.$$

- **These results do not follow from classical time series analysis: the sequences $(I_{\{a_m^{-1}\mathrm{X}_t \in A\}})_{t\leq n}$ constitute a triangular array of rowwise stationary sequences.**

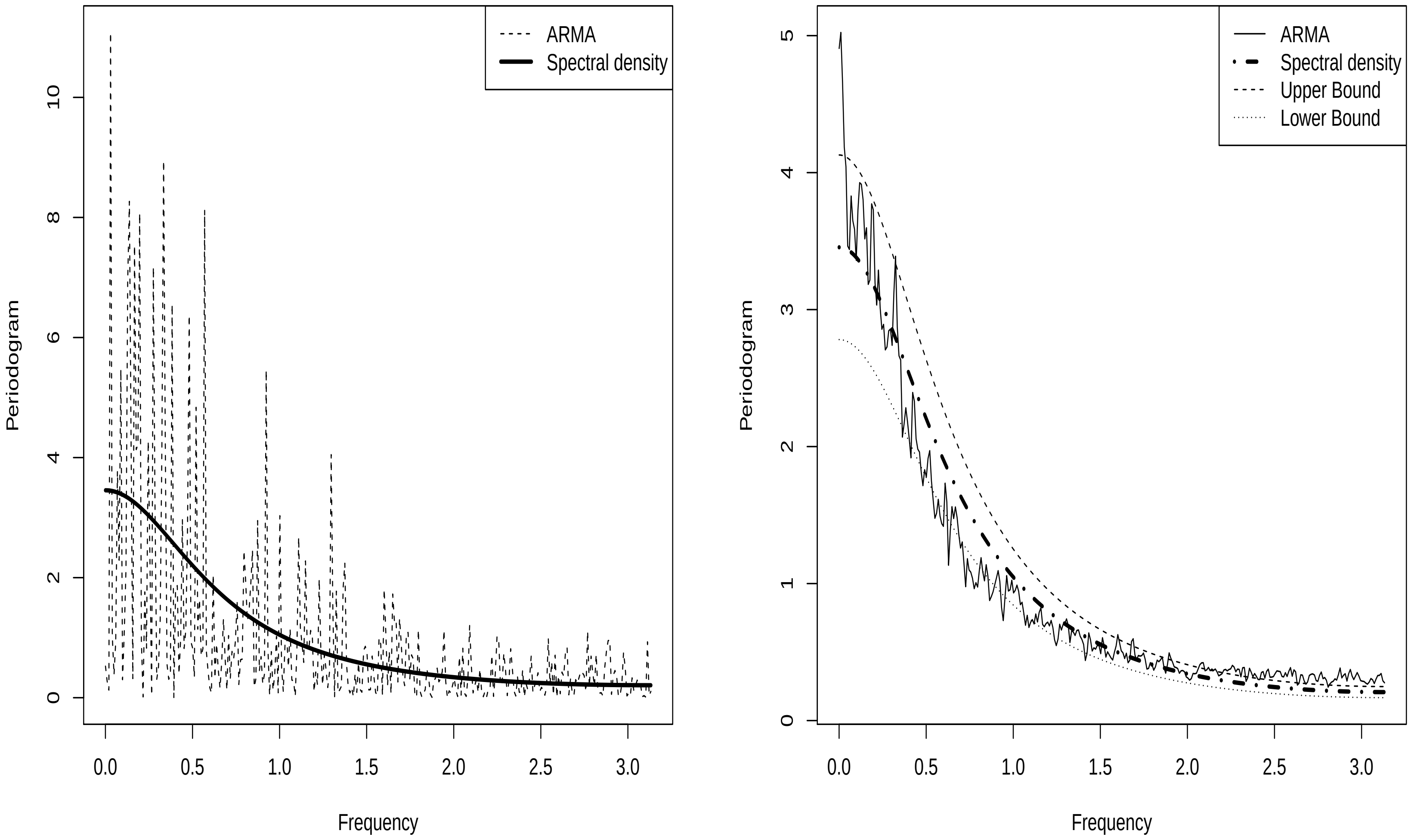

Figure 29. Raw and smoothed periodogram for ARMA(1,1) process with student(4) noise. $\boldsymbol{A = (1, \infty)}$

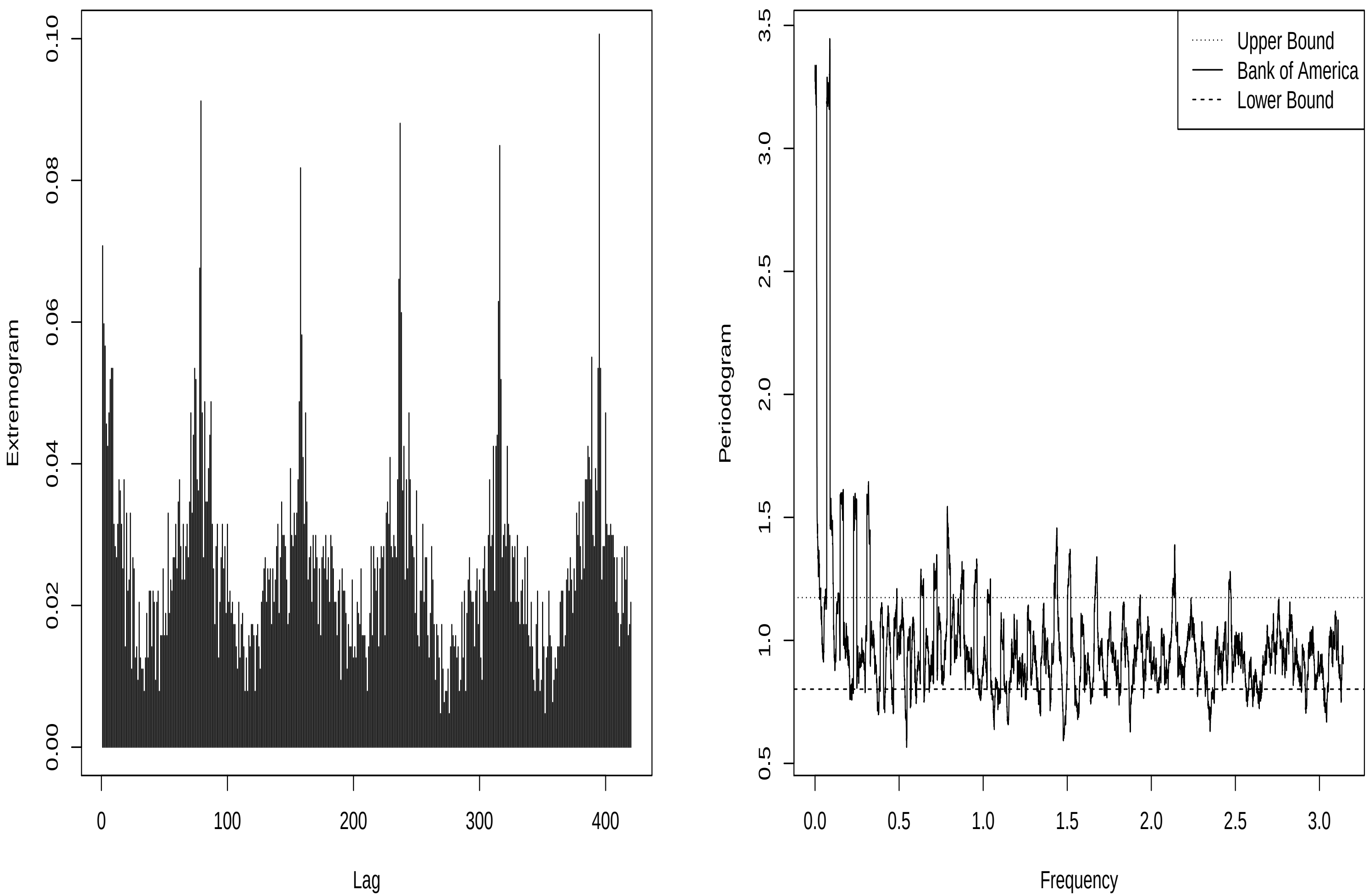

FIGURE 30. Sample extremogram and smoothed periodogram for BAC 5 minute returns. The end-of-the day effects cannot be seen in the corresponding sample autocorrelation function.

7.8. Problems and possible extensions.

- **Frequency based tools for distinguishing between time series models based on their extremes. Goodness-of-fit tests.** Zhao (2014) in progress
- **Choice of the threshold (a_n) (depending on the data)** Holger Drees; Rafal Kulik, Philippe Soulier
- **The extremogram for spatio-temporal data, estimating max-stable processes with the extremogram,...** Richard A. Davis and co-workers, Claudia Klüppelberg, Christina Steinkohl,....
- **Extremogram for functional data.**

8. Max-stable processes with Fréchet marginals

- **Max-stable processes and random fields have recently attracted some attention for modeling spatio-temporal extremal phenomena.**
- **Recall that a max-stable random variable X satisfies**

$$c_n^{-1}(M_n - b_n) \stackrel{d}{=} X\,, \quad n \geq 1\,,$$

for suitable constants $c_n > 0$ and $d_n \in \mathbb{R}$, iid copies (X_i) of X.

- **We will assume that X has a Fréchet distribution function $\Phi_\alpha(x) = \mathrm{e}^{-x^{-\alpha}}$, $x > 0$. Then $c_n = n^{1/\alpha}$ and $d_n = 0$.**
- **A Fréchet random variable has the following useful representation.**

- **Lemma.** Let

$$\Gamma_i = E_1 + \cdots + E_i\,, \quad i \geq 1,$$

be an iid standard exponential sequence (E_i), independent of an iid sequence (V_i) of positive random variables with $\mathbb{E}[V^\alpha] < \infty$ for some $\alpha > 0$. Then $\sup_{i\geq 1} \Gamma_i^{-1/\alpha} V_i$ has a Fréchet $\Phi_\alpha^{\mathbb{E}[V_1^\alpha]}$ distribution.

- **Note.** The counting process

$$N(t) = \#\{i \geq 1 : \Gamma_i \leq t\}\,, \quad t \geq 0\,,$$

is a unit rate homogeneous Poisson process. Given $N(t) = k$,

$$(\Gamma_1, \ldots, \Gamma_k \mid N(t) = k) \stackrel{d}{=} t\,(U_{(1)}, \ldots U_{(k)})$$

for the order statistics of an iid uniform sample on $(0, 1)$.

- **Proof.** Let (U_t) be iid uniform on $(0,1)$, independent of N and (V_t). Using the order statistics property of N, for $x>0$,

$$
\begin{aligned}
\mathbb{P}\Big(\sup_{i\geq 1}\Gamma_i^{-1/\alpha}V_i \leq x\Big) &= \lim_{t\to\infty}\mathbb{E}\Big[\mathbb{P}\Big(\sup_{i\leq N(t)}\Gamma_i^{-1/\alpha}V_i \leq x \mid N(t)\Big)\Big]\\
&= \lim_{t\to\infty}\mathbb{E}\Big[\mathbb{P}\Big(\sup_{i\leq N(t)}(tU_{(i)})^{-1/\alpha}V_i \leq x \mid N(t)\Big)\Big]\\
&= \lim_{t\to\infty}\mathbb{E}\Big[\mathbb{P}\Big(\sup_{i\leq N(t)}(tU_{i})^{-1/\alpha}V_i \leq x \mid N(t)\Big)\Big]\\
&= \lim_{t\to\infty}\mathbb{E}\Big[\mathbb{P}^{N(t)}\Big((tU_1)^{-1/\alpha}V_1 \leq x\Big)\Big]\\
&= \lim_{t\to\infty} \mathrm{e}^{-t\,\mathbb{P}(V_1^\alpha > x^\alpha t U_1)}\\
&= \lim_{t\to\infty} \mathrm{e}^{-x^{-\alpha}\int_0^{tx^\alpha}\mathbb{P}(V_1^\alpha > y)\,dy}\\
&= \mathrm{e}^{-x^{-\alpha}\,\mathbb{E}[V_1^\alpha]} = \Phi_\alpha^{\mathbb{E}[V_1^\alpha]}(x)\,.
\end{aligned}
$$

8.1. Definition de Haan (1984).

- A (positive) **max-stable process** $(Y_t)_{t\in T}$, $T \subset \mathbb{R}$ with Fréchet marginals: for iid copies $(Y_t^{(i)})_{t\in T}$, $i = 1, 2, \ldots$, of $(Y_t)_{t\in T}$,

$$n^{-1/\alpha}\big(\max_{i=1,\ldots,n} Y_t^{(i)}\big)_{t\in T} \stackrel{d}{=} (Y_t)_{t\in T}\,, \quad n \geq 1\,.$$

- Then, in particular, all one-dimensional marginals of the process $(Y_t)_{t\in T}$ are Fréchet distributed, i.e. Y_t has distribution $\Phi_\alpha^{c(t)}$ for some function $c(t) \geq 0$, $t \in T$.
- **Example** (de Haan (1984)). $0 < \Gamma_1 < \Gamma_2 < \cdots$, independent of the iid sequence (U_i) uniform on $(0,1)$, (f_t) non-negative functions with $\mathbb{E}[f_t^\alpha(U_1)] < \infty$.

- Then the process

$$Y_t = \sup_{i\geq 1} \Gamma_i^{-1/\alpha} f_t(U_i)\,, \quad t \in T\,,$$

is max-stable with Fréchet marginals.

- We have to show (max-stability): For any $t_i \in T$, $x_i > 0$,

$i = 1, \ldots, m$, $n \geq 1$,

$$\mathbb{P}\big(Y_{t_1} \leq x_1, \ldots, Y_{t_m} \leq x_m\big) = \mathbb{P}^n\big(Y_{t_1} \leq x_1 n^{1/\alpha}, \ldots, Y_{t_m} \leq x_m n^{1/\alpha}\big)\,.$$

- In view of the Lemma:

$$\begin{aligned}\mathbb{P}\big(Y_{t_1} \leq x_1, \ldots, Y_{t_m} \leq x_m\big) &= \mathbb{P}\Big(\sup_{i\geq 1} \Gamma_i^{-1/\alpha} \max_{1\leq t\leq m} \big(f_t(U_i)/x_t\big) \leq 1\Big)\\ &= \mathrm{e}^{-\mathbb{E}\max_{1\leq t\leq m}(f_t(U)/x_t)^\alpha}\\ &= \mathrm{e}^{-\int_0^1 \max_{1\leq t\leq m}(f_t(u)/x_t)^\alpha\,du}\,.\end{aligned}$$

- Then the process

$$Y_t = \sup_{i\ge 1} \Gamma_i^{-1/\alpha} f_t(U_i)\,, \quad t \in T\,,$$

is max-stable with Fréchet marginals.

- We have to show (max-stability): For any $t_i \in T$, $x_i > 0$,

$i = 1, \ldots, m$, $n \ge 1$,

$$\mathbb{P}\big(Y_{t_1} \le x_1, \ldots, Y_{t_m} \le x_m\big) = \mathbb{P}^n\big(Y_{t_1} \le x_1 n^{1/\alpha}, \ldots, Y_{t_m} \le x_m n^{1/\alpha}\big)\,.$$

- In view of the Lemma:

$$\begin{aligned}\mathbb{P}\big(Y_{t_1} \le x_1, \ldots, Y_{t_m} \le x_m\big) &= \mathbb{P}\Big(\sup_{i\ge 1} \Gamma_i^{-1/\alpha} \max_{1\le t\le m} (f_t(U_i)/x_t) \le 1\Big)\\ &= \mathrm{e}^{-\mathbb{E}\max_{1\le t\le m}(f_t(U)/x_t)^\alpha}\\ &= \mathrm{e}^{-n\int_0^1 \max_{1\le t\le m}(f_t(u)/(n^{1/\alpha}x_t))^\alpha\,du}\,.\end{aligned}$$

- Max-stability is immediate. $\square$

8.2. Characterization of α-Fréchet max-stable processes.

- This Example already yields an almost complete characterization of the finite-dimensional distributions of a max-stable process.
- De Haan (1984) gave a complete characterization (we consider the case $T = \mathbb{Z}$ only).
- **Theorem.** The finite-dimensional distributions of a max-stable sequence $(Y_t)_{t\in\mathbb{Z}}$ with Fréchet marginals with index $\alpha > 0$ satisfy the relation for $x_i > 0$, $i = 1, \ldots, m$, $m \geq 1$,

$$\mathbb{P}(Y_1 \leq x_1, \ldots, Y_m \leq x_m) = \mathrm{e}^{-\int_{\mathbb{R}_+^n} \max_{t\leq n}(y_t/x_t)^\alpha G_m(dy)},$$

where G_m is the m-dimensional restriction to $\mathbb{R}_+^m$ of a finite measure on $\mathbb{R}_+^\infty$.

- Moreover, there exists a finite measure ρ on $[0,1]$ such that (Y_t) has representation

$$Y_t = \sup_{i\ge 1} \Gamma_i^{-1/\alpha} f_t(T_i)\,, \quad t \in \mathbb{Z}\,, \tag{8.2}$$

where $((\Gamma_i, T_i))_{i=1,2,\ldots}$ is an enumeration of PRM(*Lebesgue* $\times\, \rho$) on $(0,\infty)\times[0,1]$, (f_t) are suitable non-negative measurable functions on $[0,1]$ such that $\mathbb{E}[f_t^\alpha(T_1)] = \int_0^1 f_t^\alpha(x)\rho(dx) < \infty$.

- Kabluchko (2009): any max-stable process $(Y_t)_{t\in T}$, $T\subset\mathbb{R}$, with α-Fréchet marginals has representation (8.2), $f_t \in L_+^\alpha(\mathbb{E},\mathcal{E},\nu)$, $t\in T$, ν a σ-finite measure on the Borel σ-field $\mathcal{E}$ of the state space $\mathbb{E}$, $((\Gamma_i, T_i))_{i\geq 1}$ are the points of a PRM(*Lebesgue* $\times\, \nu$) on the state space $\mathbb{R}_+ \times \mathbb{E}$.

- Using the same notation, one can introduce de Haan's (1984) extremal integral

$$\int_{\mathbb{E}}^{\vee} f \, dM_\nu^\alpha = \sup_{i\geq 1} \Gamma_i^{-1/\alpha} f(T_i) \,,$$

where, as above f is a non-negative function in $L^\alpha(\mathbb{E}, \mathcal{E}, \nu)$, and M_ν^α is an α-Fréchet random sup-measure with control measure ν.

- Stoev (2008) proved that $\int_{\mathbb{E}}^{\vee} f dM_{\nu}^{\alpha}$ has various properties similar to the α-stable integrals; see Samorodnitsky and Taqqu (1994). A proof similar to the Example above yields

$$\mathbb{P}\Big(\int_{\mathbb{E}}^{\vee} f dM_{\nu}^{\alpha} \le x \Big) = \exp\left\{ -x^{-\alpha} \int_{\mathbb{E}} f^{\alpha} d\nu \right\} = \Phi_{\alpha}^{\int_{\mathbb{E}} f^{\alpha} d\nu}(x) \,.$$

- The integral representation of a max-stable process is convenient. For example, for any $f_t \in L^{\alpha}(\mathbb{E}, \mathcal{E}, \nu)$, $x_t > 0$, $t = 1, \ldots, m$, $m \ge 1$,

$$\begin{aligned} & \mathbb{P}\Big(\int_{\mathbb{E}}^{\vee} f_t dM_{\nu}^{\alpha} \le x_t, t = 1, \ldots, m \Big) \\ = \ & \mathbb{P}\Big(\int_{\mathbb{E}}^{\vee} \max_{t=1,\ldots,m} (f_t / x_t) dM_{\nu}^{\alpha} \le 1 \Big) \\ = \ & \exp\left\{ - \int_{\mathbb{E}} \max_{t=1,\ldots,m} (f_t / x_t)^{\alpha} d\nu \right\} . \end{aligned}$$

- We also have for $\mathbf{x} = (x_1, \ldots, x_m) > 0$ and $y \to \infty$,

$$
\begin{aligned}
& y\Big[1 - \mathbb{P}\Big(\int_{\mathbb{E}}^{\vee} f_t dM_\nu^\alpha \le y^{1/\alpha} x_t, t = 1, \ldots, m \Big)\Big] \\
&= y\,\mathbb{P}\Big(y^{-1/\alpha}\Big(\int_{\mathbb{E}}^{\vee} f_t dM_\nu^\alpha \Big)_{t=1,\ldots,m} \notin [0, \mathbf{x}] \Big) \\
&= y\Big(1 - \exp\Big\{ -y^{-1} \int_{\mathbb{E}} \max_{t=1,\ldots,m} (f_t/x_t)^\alpha d\nu \Big\}\Big) \\
&\to \int_{\mathbb{E}} \max_{t=1,\ldots,m} (f_t/x_t)^\alpha d\nu = \mu_{m,\alpha}([0, \mathbf{x}]^c)\,. \qquad (8.3)
\end{aligned}
$$

Thus the finite-dimensional distributions of a max-stable process $(Y_t)_{t\in T}$ are regularly varying with index α and limiting measure $\mu_{m,\alpha}$ given by (8.3).

8.3. Stationary max-stable processes.

- Recently, strictly stationary max-stable processes $(Y_t)_{t\in T}$ for $T = \mathbb{Z}$ or $T = \mathbb{R}$ have attracted some attention. Such a process has again integral representation

$$Y_t = \int_{\mathbb{E}}^{\vee} f_t dM_{\nu}^{\alpha}\,, \quad t \in T\,,$$

where the family of functions (f_t) has to satisfy some particular conditions to ensure strict stationarity, ergodicity, mixing, and other desirable properties; see Kabluchko (2009), Stoev (2008).

- **Example.** Since (Y_t) is regularly varying with index α one can define its extremogram. For example, the extremogram with respect to the set $(1,\infty)$ is given by

$$
\begin{aligned}
\rho(h) &= \lim_{x\to\infty} \mathbb{P}(x^{-1}Y_h > 1 \mid x^{-1}Y_0 > 1) \\
&= \frac{\mathbb{P}(x^{-1}\min(Y_0, Y_h) > 1)}{\mathbb{P}(Y_0 > x)} \\
&= \lim_{x\to\infty} \frac{1-\exp\big\{-x^{-\alpha}\int_{\mathbb{E}}\min(f_0^\alpha, f_h^\alpha)d\nu\big\}}{1-\exp\big\{-x^{-\alpha}\int_{\mathbb{E}} f_0^\alpha d\nu\big\}} \\
&= \frac{\int_{\mathbb{E}}\min(f_0^\alpha, f_h^\alpha)d\nu}{\int_{\mathbb{E}} f_0^\alpha d\nu}.
\end{aligned}
$$

- It is also straightforward to calculate the extremal index of (Y_t) provided it exists.
- Consider (a_n) satisfying

$$\mathbb{P}(Y_0 > a_n) = 1 - \mathrm{e}^{-a_n^{-\alpha} \int_{\mathbb{E}} f_0^\alpha d\nu} \sim n^{-1}$$

i.e. $a_n \sim n^{1/\alpha} \big(\int_{\mathbb{E}} f_0^\alpha d\nu \big)^{1/\alpha}$. Then, for $x > 0$,

$$\begin{aligned} \mathbb{P}\Big(a_n^{-1} \max_{t=1,\ldots,n} Y_t \le x\Big) &= \exp\Big\{-a_n^{-\alpha} x^{-\alpha} \int_{\mathbb{E}} \max_{t=1,\ldots,n} f_t^\alpha d\nu\Big\} \\ &= \Big[\Phi_\alpha(x)\Big]^{n^{-1} \int_{\mathbb{E}} \max_{t=1,\ldots,n} f_t^\alpha d\nu \Big/ \int_{\mathbb{E}} f_0^\alpha d\nu (1+o(1))} . \end{aligned}$$

If the limit

$$\theta_Y = \lim_{n\to\infty} \frac{1}{n} \frac{\int_{\mathbb{E}} \max_{t=1,\ldots,n} f_t^\alpha d\nu}{\int_{\mathbb{E}} f_0^\alpha d\nu}$$

exists it is the extremal index of (Y_t).

The Brown-Resnick process Brown, Resnick (1977).

- Has representation

$$(8.4) \qquad Y_t = \sup_{i\geq 1} \Gamma_i^{-1/\alpha} \, \mathrm{e}^{W_i(t) - 0.5\,\sigma^2(t)}\,, \quad t \in \mathbb{R}\,,$$

where $0 < \Gamma_1 < \Gamma_2 < \cdots$ are the points of a unit rate homogeneous Poisson process on $(0, \infty)$, independent of an iid sequence (W_i) of sample continuous zero-mean Gaussian processes on $\mathbb{R}$ with stationary increments and variance function σ^2, e.g. Brownian motions or fractional Brownian motions.

- The max-stable process (8.4) is stationary; see Kabluchko, Schlather, de Haan (2009). In this paper, the authors also consider the case of max-stable random fields, i.e. W is a mean zero Gaussian random field with stationary increments. They show that its distribution only depends on the variogram

$$V(h) = \mathrm{var}(W(t+h) - W(t)), \quad t \in \mathbb{R}^d, \quad h \in \mathbb{R}^d.$$

- The Brown-Resnick process has attracted some attention; see e.g. Kabluchko (2009), Kabluchko et al. (2009), Stoev (2008), Oesting, Kabluchko, Schlather (2012). The processes (8.4) can be extended to random fields on $\mathbb{R}^d$. These fields found various applications for modeling spatio-temporal extremal effects; see e.g. Kabluchko et al. (2009), Davis, Klüppelberg, Steinkohl (2013), Davison, Padoan, Ribatet (2012).
- The Brown-Resnick process cannot be simulated in a naive way by mimicking the formula (8.4) and replacing the supremum over an infinite index set by a finite one.

- For example, assume that W is standard Brownian motion. Then $(\mathrm{e}^{W(t)-0.5t})_{t\ge 0}$ is a martingale with expectation 1. On the other hand, by virtue of the law of the iterated logarithm, $\mathrm{e}^{W(t)-0.5t}\longrightarrow 0$ a.s. exponentially fast as $t\to\infty$. For every finite m, $\sup_{1\le i\le m}\Gamma_i^{-1/\alpha}\,\mathrm{e}^{W_i(t)-0.5\,\sigma^2(t)}\longrightarrow 0$ exponentially fast as $t\to\infty$. This fact turns the simulation of (Y_t) into a complicated problem; see e.g. Schlather (2002), Oesting et al. (2012), M., Dieker (2014)

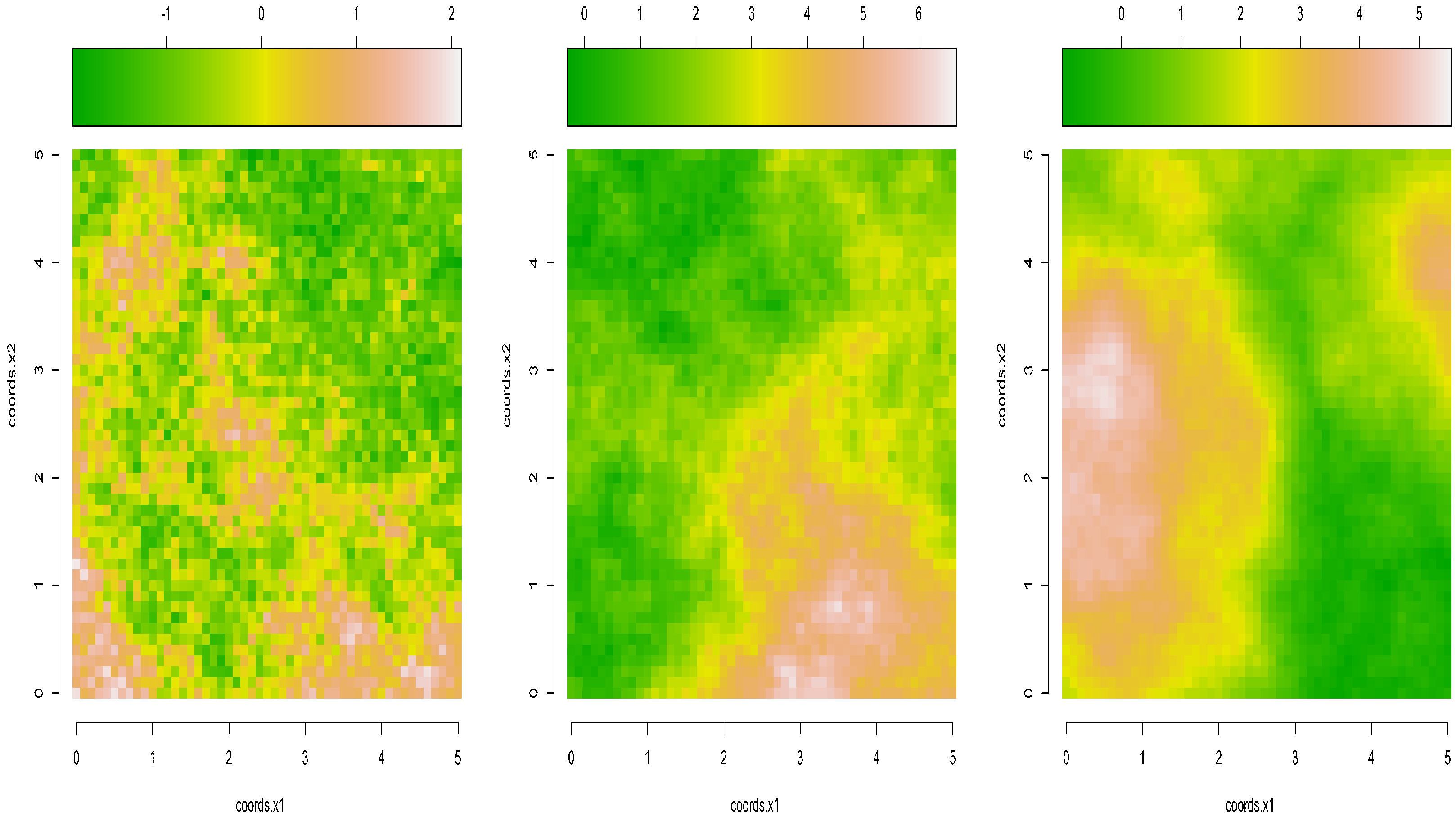

FIGURE 31. Sample of a Brown-Resnick random field on $[0,5]^2$ with variogram $\gamma(t) = |t|^\alpha/2$ for $\alpha = 1/2$, $\alpha = 1$, $\alpha = 3/2$ from left to right, respectively. The grid mesh is 0.1.

9. Concluding remarks

- **Over the last 10-15 years, multivariate regular variation has become a major tool for dealing with extremes in time series and spatial data.**
- **Regularly varying structures (such as regularly varying time series, max-stable processes and random fields) are flexible models for the extreme part of the data.**
- **This means that regular variation is well suited for describing both tails and dependence of extremal events in space and time beyond second order characteristics (covariances).**

- **Important parts of the theory were left out in this course:**

(1) **Asymptotic theory for point processes, partial sums, maxima, large deviations,... acting on dependent regularly varying structures.** e.g. Davis, Hsing (1995), Basrak, Segers (2009), M., Wintenberger (2013a,b)

(2) **The statistical theory of these processes is far from being complete.**

(3) **Functional regular variation.** e.g. de Haan, Tao (2003), Davis, M. (2006)

(4) **Regularly varying random matrices.** e.g. Soshnikov (2006), Ben Arous and Guionnet (2008), Belinschi, Dembo, Guionnet (2009), Bose, Hazra, Saha (2010), Davis, M., Pfaffel (2013)

(5) Simulation of max-stable and other regularly varying structures. e.g. Schlather (2002), Oesting et al. (2012)

(6) Beyond the regularly varying case, the literature on moderately heavy-tailed multivariate structures and time series is sparse (e.g. natural generalisations of subexponentiality to higher dimensions).

References

[1] Basrak, B., Davis, R.A. and Mikosch. T. (2002) A characterization of multivariate regular variation. *Ann. Appl. Probab.* **12**, 908–920.

[2] Basrak, B. and Segers J. (2009) Regularly varying multivariate time series. *Stoch. Proc. Appl.* **119**, 1055–1080.

[3] Belinschi, S., Dembo, A. and Guionnet, A. (2009) Spectral measure of heavy tailed band and covariance random matrices. *Comm. Math. Phys.* **289**, 1023–1055.

[4] Ben Arous, G. and Guionnet, A. (2008) The spectrum of heavy tailed random matrices. *Comm. Math. Phys.* **278**, 715–751.

[5] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987) *Regular Variation.* Cambridge University Press, Cambridge.

[6] Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31**, 307–327.

[7] Bose, A., Subhra Hazra, R. and Saha, K. (2010) Spectral norm of circulant type matrices with heavy tailed entries. *Electron. Commun. Probab.* **15**, 299–313.

[8] Breiman, L. (1965) On some limit theorems similar to the arc-sin law. *Theory Probab. Appl.* **10**, 323–331.

[9] Brockwell, P.J. and Davis, R.A. (1991) *Time Series: Theory and Methods,* 2nd edition. Springer, New York.

[10] Brockwell, P.J. and Davis, R.A. (1996) *Introduction to Time Series and Forecasting.* Springer, New York.

[11] Brown, B. and Resnick, S.I. (1977) Extreme values of independent stochastic processes. *J. Appl. Probab.* **14**, 732–739.

[12] 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.

[13] Davis, R.A., Klüppelberg, C. and Steinkohl, C. (2013) Statistical inference for max-stable processes in space and time. *J. Royal Statist. Soc., Series B*, to appear.

[14] Davis, R.A. and Mikosch, T. (2001a) Point process convergence of stochastic volatility processes with application to sample autocorrelations. *J. Appl. Probab.* Special Volume **38A**, A Festschrift for David Vere-Jones, 93–104.

[15] Davis, R.A. and Mikosch, T. (2001b) The sample autocorrelations of financial time series models. In: Fitzgerald, W.J., Smith, R.L., Walden, A.T. and Young, P.C. (Eds.) *Nonlinear and Nonstationary Signal Processing,* pp. 247–274. Cambridge University Press, Cambridge (U.K.).

[16] Davis, R.A. and Mikosch, T. (2008) Extreme value theory for space-time processes with heavy-tailed distributions. *Stoch. Proc. Appl.* **118**, 560–584.

[17] Davis, R.A. and Mikosch, T. (2009a) Fundamental properties of stochastic volatility models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. *Handbook of Financial Time Series.* Springer, Berlin.

[18] Davis, R.A. and Mikosch, T. (2009b) Extremes of stochastic volatility models. In: Andersen, T.G., Davis, R.A., Kreiss, J.-P. and Mikosch, T. (2007) *Handbook of Financial Time Series.* Springer, Berlin.

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

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

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

[22] Davis, R.A., Mikosch, T. and Pfaffel, O. (2013) Asymptotic theory for the sample covariance matrix of heavy-tailed linearly dependent double arrays. Work in progress.
[23] Davis, R.A., Mikosch, T. and Zhao, Y. (2013) Measures of serial extremal dependence and their estimation. *Stoch. Proc. Appl.* **123**, 2575–2602.
[24] Davison, A.C., Padoan, S.A. and Ribatet, M. (2012) Statistical modeling of spatial extremes. *Statist. Sci.* **27**, 161-186.
[25] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.
[26] Falk, M., Hüsler, J. and Reiss, R.-D. (2004) *Laws of Small Numbers: Extremes and Rare Events.* 2nd Edition. Birkhäuser, Basel.
[27] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications.* Vol. II. Second edition. Wiley, New York.
[28] Gnedenko, B.V. (1943) Sur la distribution limité du terme d'une série aléatoire. *Ann. Math.* **44**, 423–453.
[29] Goldie, C.M. (1991) Implicit renewal theory and tails of solutions of random equations. *Ann. Appl. Probab.* **1**, 126–166.
[30] Haan, L. de (1984) A spectral representation for max-stable processes. *Ann. Probab.* **12**, 1194–1204.
[31] Haan, L. de, Fereira, A. (2006) *Extreme Value Theory: An Introduction.* Springer, Berlin, New York.
[32] Haan, L. de, Resnick, S.I., Rootzén, H. and Vries, C. de (1989) Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stoch. Proc. Appl.* **32**, 213–224.
[33] Haan, L. de and Tao, L. (2003) Weak consistency of extreme value estimators in $C[0,1]$. *Ann. Statist.* **31**, 1996–2012.
[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] Kesten, H. (1973) Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131**, 207–248.
[37] Leadbetter, M.R. (1983) Extremes and local dependence of stationary sequences. *Z. Wahrscheinlichkeitstheorie verw. Gebiete* **65**, 291–306.
[38] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983) *Extremes and Related Properties of Random Sequences and Processes.* Springer, Berlin.
[39] Loynes, R.M. (1965) Extreme values in uniformly mixing stationary stochastic processes. *Ann. Math. Statist.* **36**, 993–999.
[40] McNeil, A., Frey, R. and Embrechts, P. (2005) *Quantitative Risk Management.* Princeton University Press.
[41] Mikosch, T. and Zhao, Y. (2013) A Fourier analysis of extreme events. *Bernoulli*, to appear.
[42] Mikosch, T. and Wintenberger, O. (2013a) Precise large deviations for dependent regularly varying sequences. *Probab. Rel. Fields*, to appear.
[43] Mikosch, T. and Wintenberger, O. (2013b) The cluster index of regularly varying sequences with applications to limit theory for functions of multivariate Markov chains. *Probab. Rel. Fields*, to appear.
[44] Newell, G.F. (1964) Asymptotic extremes for $m$dependent random variables. *Ann. Math. Statist.* **35**, 1322–1325.
[45] Oesting, M., Kabluchko, Z. and Schlather, M. (2012) Simulation of Brown-Resnick processes. *Extremes* **15**, 89–107.
[46] O'Brien, G.L. (1974) Limit theorems for the maximum term of a stationary process. *Ann. Probab.* **2**, 540–545.
[47] O'Brien, G.L. (1987) Extreme values for stationary and Markov sequences. *Ann. Probab.* **15**, 281–291.
[48] Politis, D.N. and Romano, J.P. (1994) The stationary bootstrap. *J. Amer. Statist. Assoc.* **89**, 1303–1313.
[49] Resnick, S.I. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

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

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

[52] Samorodnitsky, G. and Taqqu, M.S. (1994) *Stable Non–Gaussian Random Processes. Stochastic Models with Infinite Variance.* Chapman and Hall, London.

[53] Schlather, M. (2002) Models for stationary max-stable random fields. *Extremes* **5**, 33–44.

[54] Segers, J. (2005) Approximate distributions of clusters of extremes. *Stat. Probab. Letters* **74**, 330–336.

[55] Soshnikov, A. (2006) Poisson statistics for the largest eigenvalues in random matrix ensembles. In: *Mathematical Physics of Quantum Mechanics.* Lecture Notes in Physics **690**. Springer, Berlin, pp. 351–364.

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

[57] Stoev, S. and Taqqu, M.S. (2005) Extremal stochastic integrals: a parallel between max-stable processes and $\boldsymbol{\alpha}$-stable processes. *Extremes* **8**, 237–266.