Fourier analysis of extreme events

[1]

Thomas Mikosch
University of Copenhagen
www.math.ku.dk/∼mikosch/

[1]PhD Course an Extremes in Space and Time, 27-30, May, 2013

1. 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 $k \geq 1$,**

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

The measures μ_k determine the extremal dependence structure of the finite-dimensional distributions.

- **Notice: Normalization $P(|X_0| > x)$ does not depend on k.**

Examples of regularly varying stationary sequences

Linear processes.

- 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.

Solutions to stochastic recurrence equation.

- For an iid sequence $((A_t, B_t))_{t\in\mathbb{Z}}$, $A > 0$, 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 $E\log A < 0$, $E|\log|B|| < \infty$.

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

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

- **The GARCH(1, 1) process**[2] **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 $E(\alpha_1 Z^2 + \beta_1)^{\alpha/2} = 1$.

Other examples of regularly varying sequences.

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

[2]Bollerslev (1986)

2. Some facts from classical time series analysis Brockwell, Davis (1991,1996)

- Classical time series analysis deals with the covariance structure of second order stationary processes (X_t). We assume $X_t \in \mathbb{R}$.
- In the time domain, the autocovariance (ACVF) and autocorrelation functions (ACF) are of major interest:

$$\gamma_X(h) = \text{cov}(X_0, X_h)\,, \quad h \in \mathbb{Z}\,,$$
$$\rho_X(h) = \text{corr}(X_0, X_h) = \frac{\text{cov}(X_0, X_h)}{\text{var}(X_0)}\,, \quad h \in \mathbb{Z}\,.$$

- For a mean-zero Gaussian stationary sequence, they completely describe the dependence structure of (X_t).

- **BUT:** The extremal behavior of a Gaussian stationary sequence is similar to the extremal behavior of an iid sequence:

  No extremal clustering: Extremal index $\theta_X = 1$ if $\gamma_X(h) = o(1/\log h)$ as $h \to \infty$.

  Zero extremogram: $\lim_{x\to\infty} P(X_h > x \mid X_0 > x) = 0\,, \quad h \neq 0\,.$

- The ACVF/ACF of non-linear/non-Gaussian time series often tells little about the general dependence structure: For example, models for log-returns $X_t = \log P_t - \log P_{t-1}$, $t \in \mathbb{Z}$, are of the form $X_t = \sigma_t Z_t$, $\sigma_t > 0$ and Z_t with $EZ_t = 0$ are independent for every t, and often $\rho_X(h) = 0$, $h \neq 0$. (GARCH, SV models).

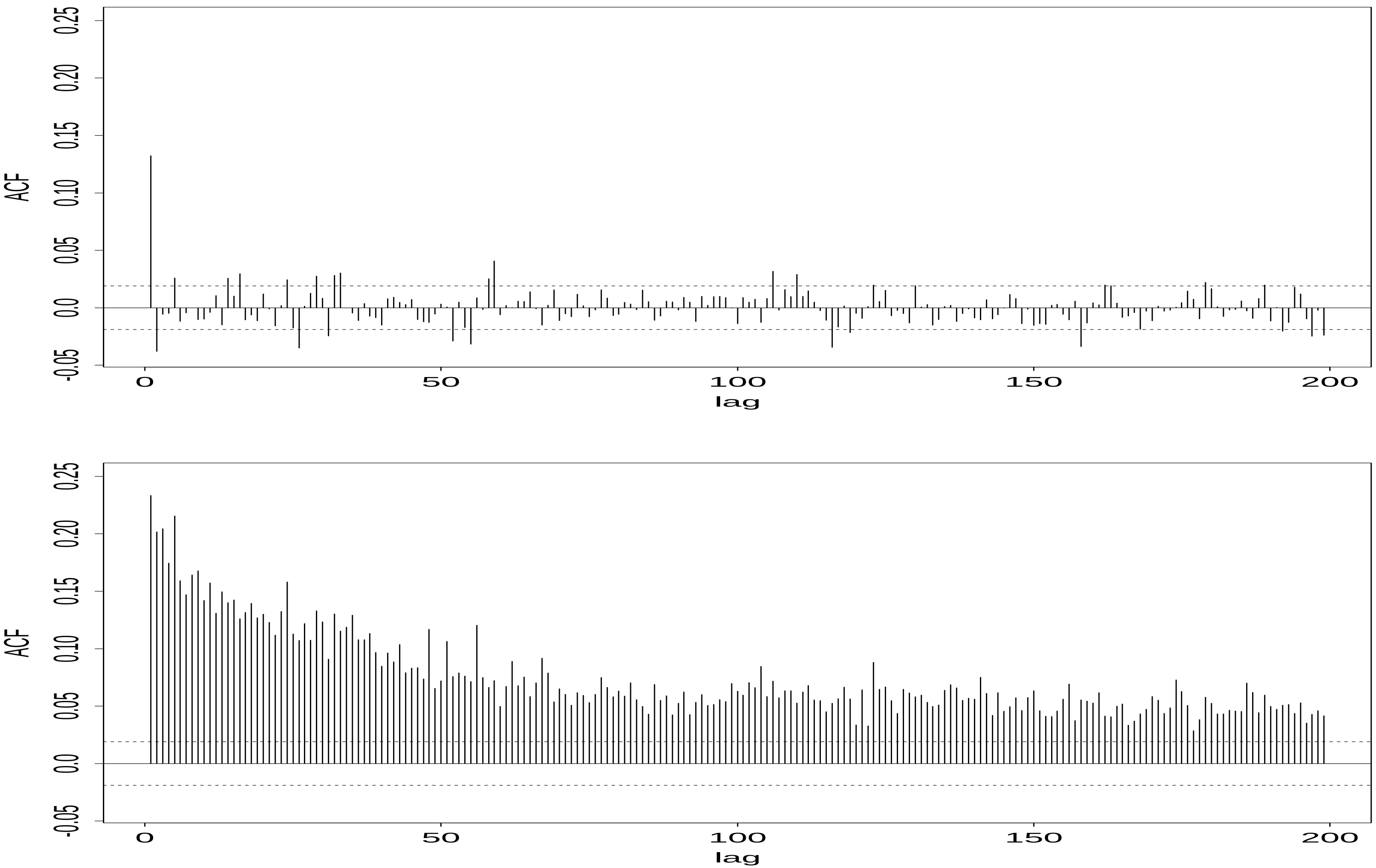

FIGURE 1. Sample ACFs for the log-returns *(top)* and absolute log-returns *(bottom)* of the *S&P500*. Here and in what follows, the horizontal lines in graphs displaying sample ACFs are set as the **95%** confidence bands $(\pm 1.96/\sqrt{n})$ corresponding to the ACF of iid Gaussian noise.

- In the frequency domain, the spectral density is of major interest

$$f_X(\lambda) = \frac{1}{2\pi} \sum_{h \in \mathbb{Z}} \gamma_X(h)\, e^{-i\lambda h}\,, \quad \lambda \in [0, \pi]\,.$$

- Then we have the representation

$$X_t = \int_{(-\pi,\pi]} e^{it\lambda} Z_X(d\lambda)\,, \quad t \in \mathbb{Z}\,, \tag{2.1}$$

with respect to a process with orthogonal increments Z_X and such that

$$E|Z_X(d\lambda)|^2 = f_X(\lambda)\, d\lambda\,.$$

- If we interpret (2.1) as superposition of trigonometric functions with random amplitude, X_t is strongly influenced by the contribution of $e^{it\lambda} Z_X(d\lambda)$ for "large" values $f_X(\lambda)$.

- **If $f_X(\lambda)$ is "large relative to f_X at other frequencies" we expect to see periodic cycles of length $2\pi/\lambda$ in the data.**

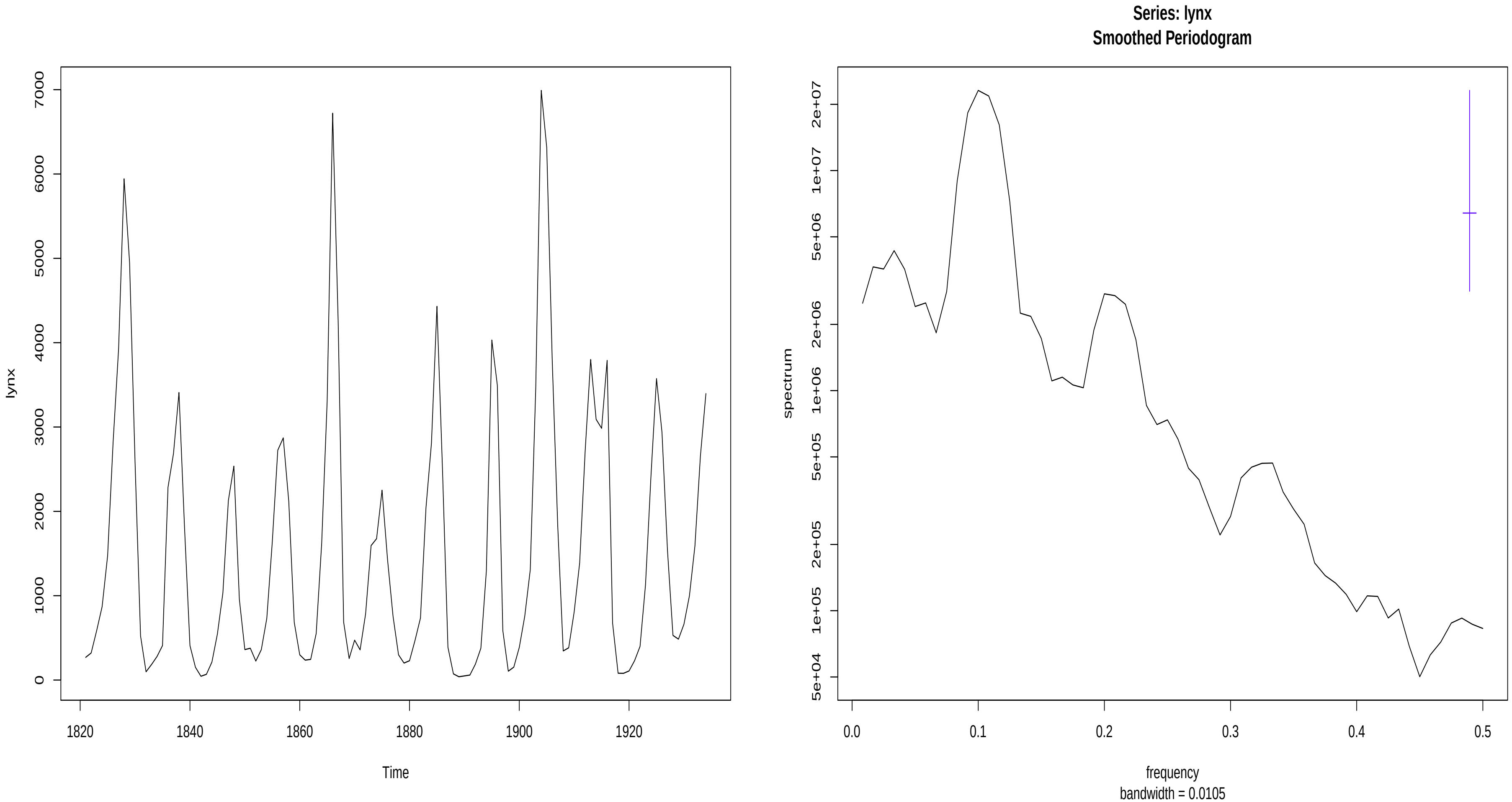

Figure 2. Monthly Canadian Lynx data (left) and estimated log-spectral density (right, the $\boldsymbol{x}$-axis corresponds to frequencies $\boldsymbol{\lambda} \in (\mathbf{0}, \boldsymbol{\pi})$). The highest peak of the density corresponds to a 10-year cycle.

- **The ACVF/ACF is estimated by the sample ACVF/ACF:**

$$\gamma_{n,X}(h) = \frac{1}{n}\sum_{t=1}^{n-|h|}(X_t - \overline{X}_n)(X_{t+|h|} - \overline{X}_n)\,, \quad h \in \mathbb{Z}\,,$$

$$\rho_{n,X}(h) = \frac{\gamma_{n,X}(h)}{\gamma_{n,X}(0)}\,, \quad h \in \mathbb{Z}\,.$$

- **Under mild assumptions (ergodicity, mixing conditions or concrete dependence structure) these sample versions are consistent and asymptotically normal estimators of their deterministic counterparts.**

- The spectral density is estimated by the periodogram

$$I_{n,X}(\lambda) = \frac{1}{2\pi n}\Big|\sum_{t=1}^{n} e^{-it\lambda} X_t\Big|^2, \quad \lambda \in (0,\pi)\,.$$

- It is not a consistent estimator of $f(\lambda)$. For example, for a linear process, at distinct frequencies $\lambda_i \in (0,\pi)$.

$$\big(I_{n,X}(\lambda_i)\big)_{i=1,\ldots,m} \stackrel{d}{\rightarrow} \Big(f_X(\lambda_i)\,E_i\Big)_{i=1,\ldots,m}$$

for an iid exponential sequence (E_i).

- The periodogram needs to be smoothed for consistent estimation of f_X.

3. Time series analysis for regularly varying linear processes and beyond

- **In the early 1980s,** Hannan; Davis, Resnick (1985), **and others, discovered that the sample ACF of the linear process**

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

with iid regularly varying noise (Z_t) **and index** $\alpha \in (0,2)$ **consistently estimates the function**

$$\rho_X(h) = \frac{\sum_j \psi_j \psi_{j+h}}{\sum_j \psi_j^2}, \quad h \in \mathbb{Z},$$

roughly at the rate $n^{1/\alpha}$**.**

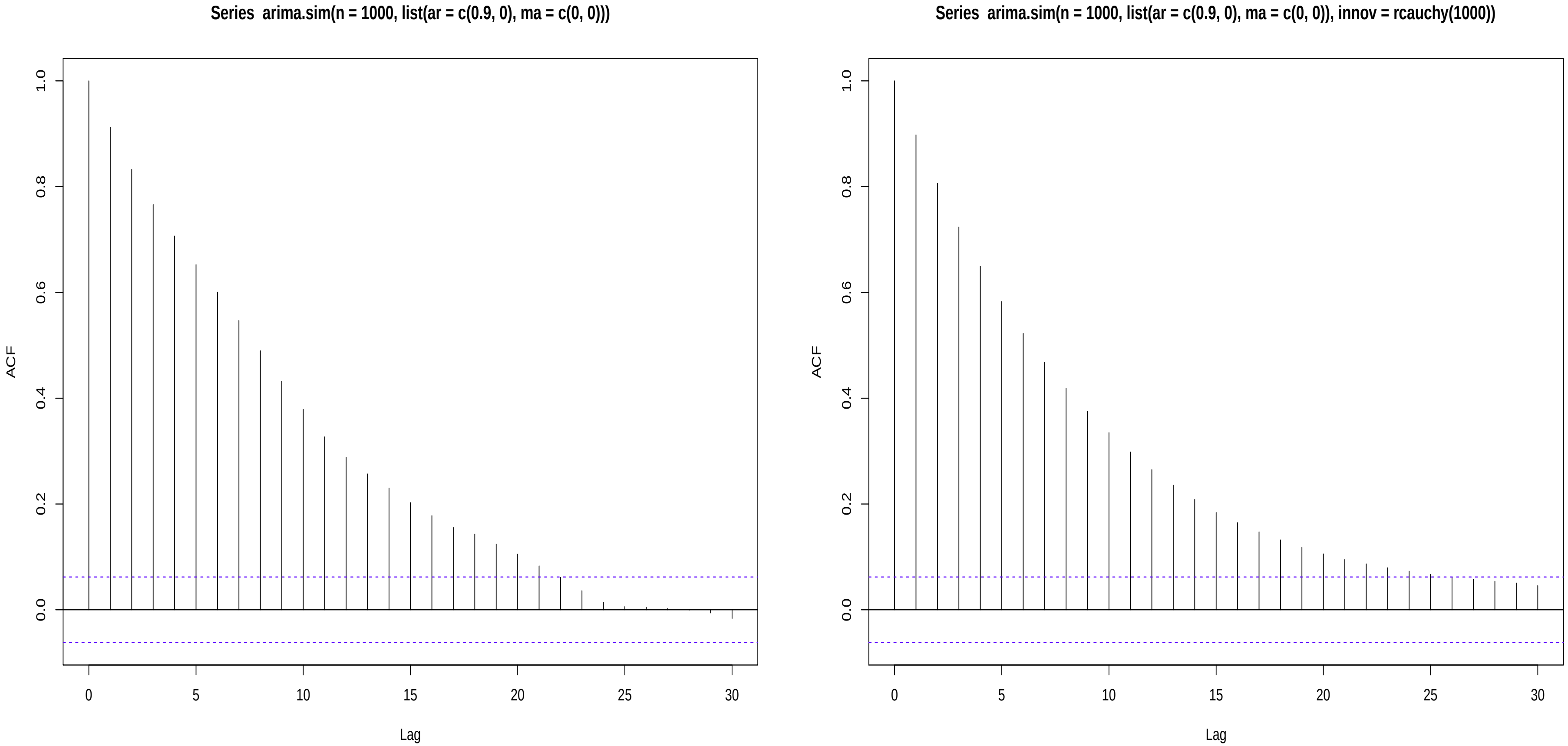

Figure 3. Sample ACF of AR(1) process $\boldsymbol{X_t = 0.9X_{t-1} + Z_t}$. Left: IID normal noise. Right: Cauchy noise.

- In this case, correlations and covariances of (X_t) are not defined and ρ_X is not the ACF of the data, BUT it is the ACF of a finite variance linear process with the same coefficients (ψ_j).

- Moreover, various estimation procedures for finite variance linear processes work for infinite variance ones with rates faster than $\sqrt{n}$, see Embrechts et al. (1997), Chapter 7, for an overview e.g.
  - Yule-Walker estimation for AR-processes,
  - Whittle estimation for general FARIMA processes Mikosch, Gadrich, Klüppelberg, Adler (1995), Kokoszka, Taqqu (1995),
  - goodness-of-fit tests based on the integrated periodogram. Klüppelberg, Mikosch (1996)

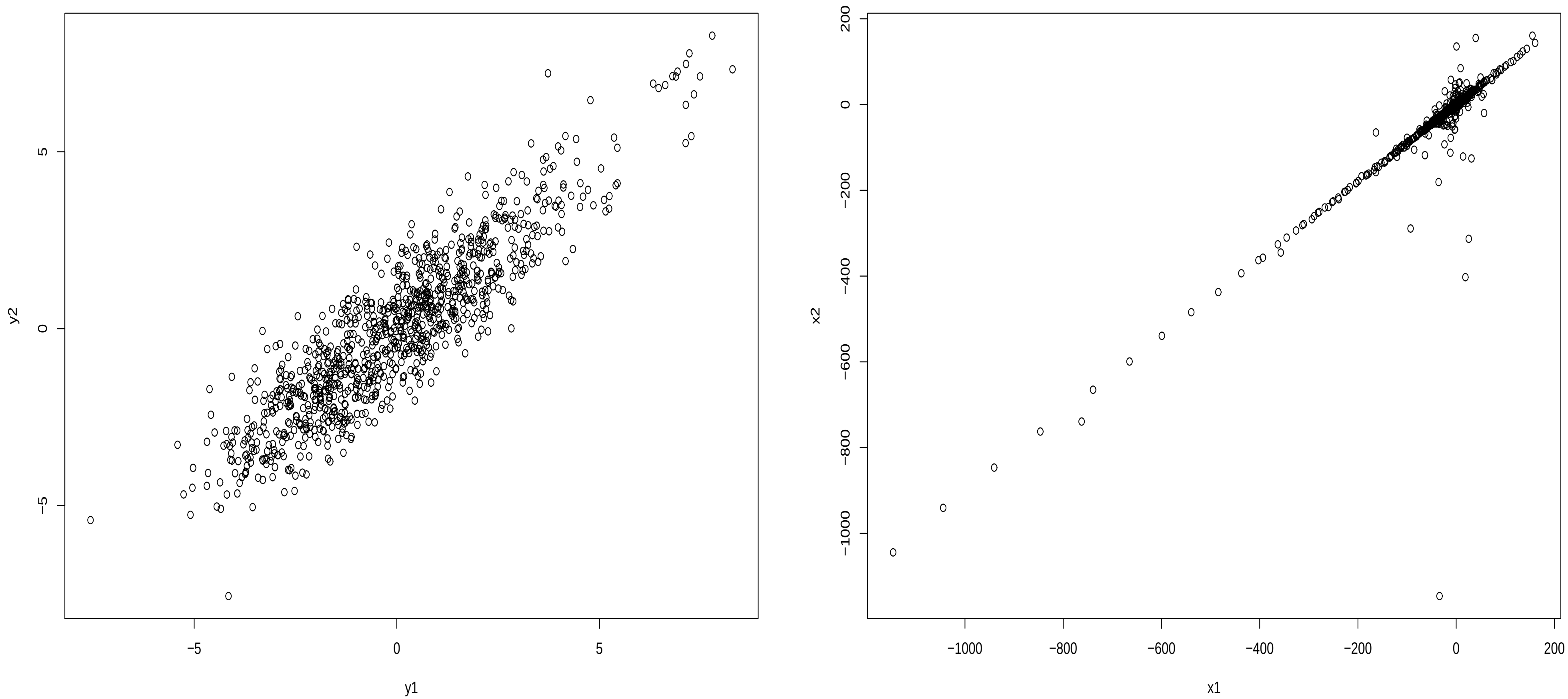

FIGURE 4. Scatterplot of AR(1) process $\boldsymbol{X_t = 0.9X_{t-1} + Z_t}$ with iid standard normal (left) and Cauchy (right) noise.

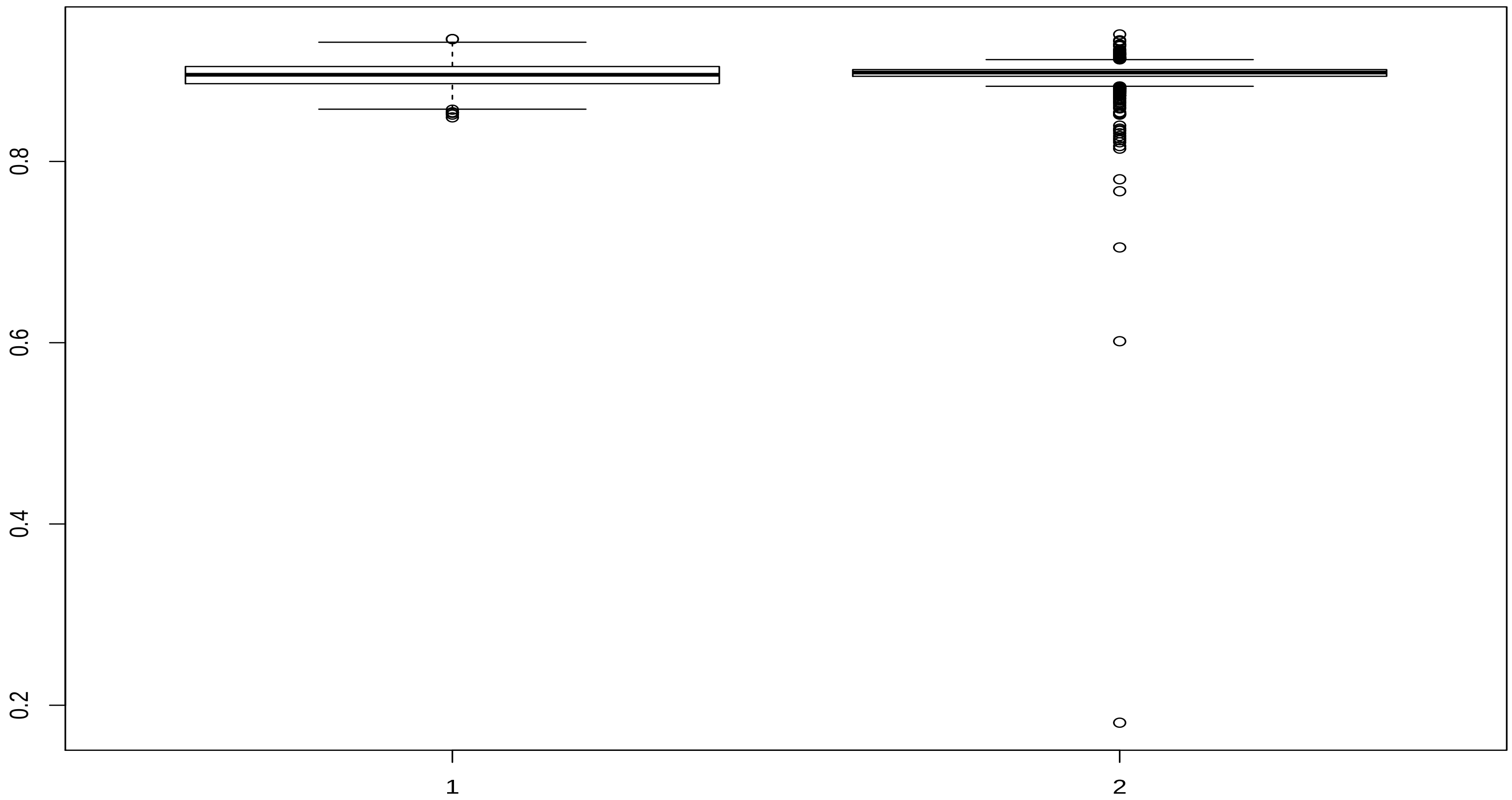

FIGURE 5. Boxplots for the Yule-Walker estimator of the AR(1) parameter. Normal noise (left) and Cauchy noise (right)

- There are certain situations when the results for finite variance linear processes do not transfer to the infinite variance case.
- Write the squared GARCH$(1,1)$ $X_t^2 = \sigma_t^2 Z_t^2$ with $\sigma_t^2 = \alpha_0 + (\alpha_1 Z_{t-1}^2 + \beta_1)\sigma_{t-1}^2$ as an ARMA(1,1) process

$$X_t^2 - EX_0^2 = \varphi_1(X_{t-1}^2 - EX_0^2) + \nu_t - \beta_1 \nu_{t-1}\,, \quad t \in \mathbb{Z}\,,$$

  where $\nu_t = \sigma_t^2(Z_t^2 - 1)$, $\varphi_1 = \alpha_1 + \beta_1 < 1$, $\beta_1 < 1$, (Z_t) is iid standard normal.
- If $\sigma_\nu^2 = \text{var}(\nu_0) < \infty$, (ν_t) is a white noise sequence and the ARMA(1,1) process $(X_t^2 - EX_0^2)$ has spectral density, with $\theta = (\varphi_1, \beta_1) = (\alpha_1 + \beta_1, \beta_1)$,

$$f(\lambda, \theta) = \frac{\sigma_\nu^2}{2\pi} g(\lambda, \theta) = \frac{\sigma_\nu^2}{2\pi} \frac{|1 - \beta_1 \mathrm{e}^{-i\lambda}|^2}{|1 - \varphi_1 \mathrm{e}^{-i\lambda}|^2}\,.$$

- The Whittle estimator $\widehat{\theta}$ of θ (which is the Yule-Walker estimator if $\beta_1 = 0$) is the minimizer of the objective function

$$\frac{1}{n}\sum_j \frac{I_{n,X^2}(2\pi j/n)}{g(\lambda_j,\theta)}, \qquad \{\theta : 0 \le \beta_1 < 1, \beta_1 \le \varphi_1 \le 1\}.$$

Then for the tail index $\alpha \in (0, 8)$ Mikosch, Straumann (2002)

$$n^{1-4/\alpha}(\widehat{\theta} - \theta_0) \xrightarrow{d} \xi_{\alpha/4}$$

where θ_0 is the true parameter (of the data), $\xi_{\alpha/4}$ is an $\alpha/4$-stable random vector.

- Notice: For $\alpha \le 4$, $\widehat{\theta}$ is not a consistent estimator of θ_0. For $\alpha \in (4, 8)$, rates of convergence are slower than $\sqrt{n}$.

- **Main reason for the failure of the Whittle estimator for the squared GARCH$(1,1)$ process: Slow convergence rates of the sample ACF/ACVF.**

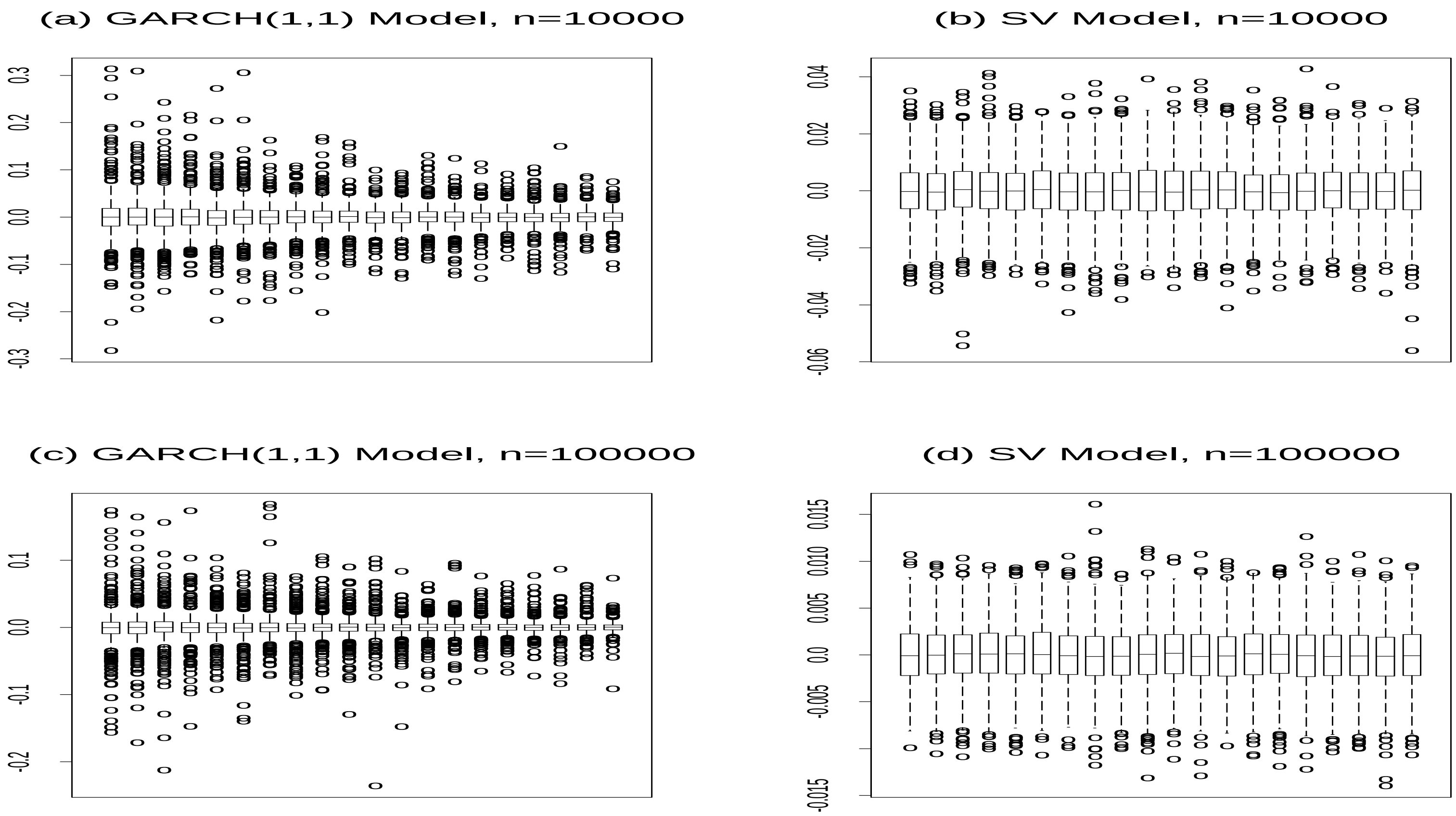

Figure 6. Boxplot comparison of sample ACFs of GARCH$(\mathbf{1,1})$ and stochastic volatility models with tail parameter $\boldsymbol{\alpha = 3}$.

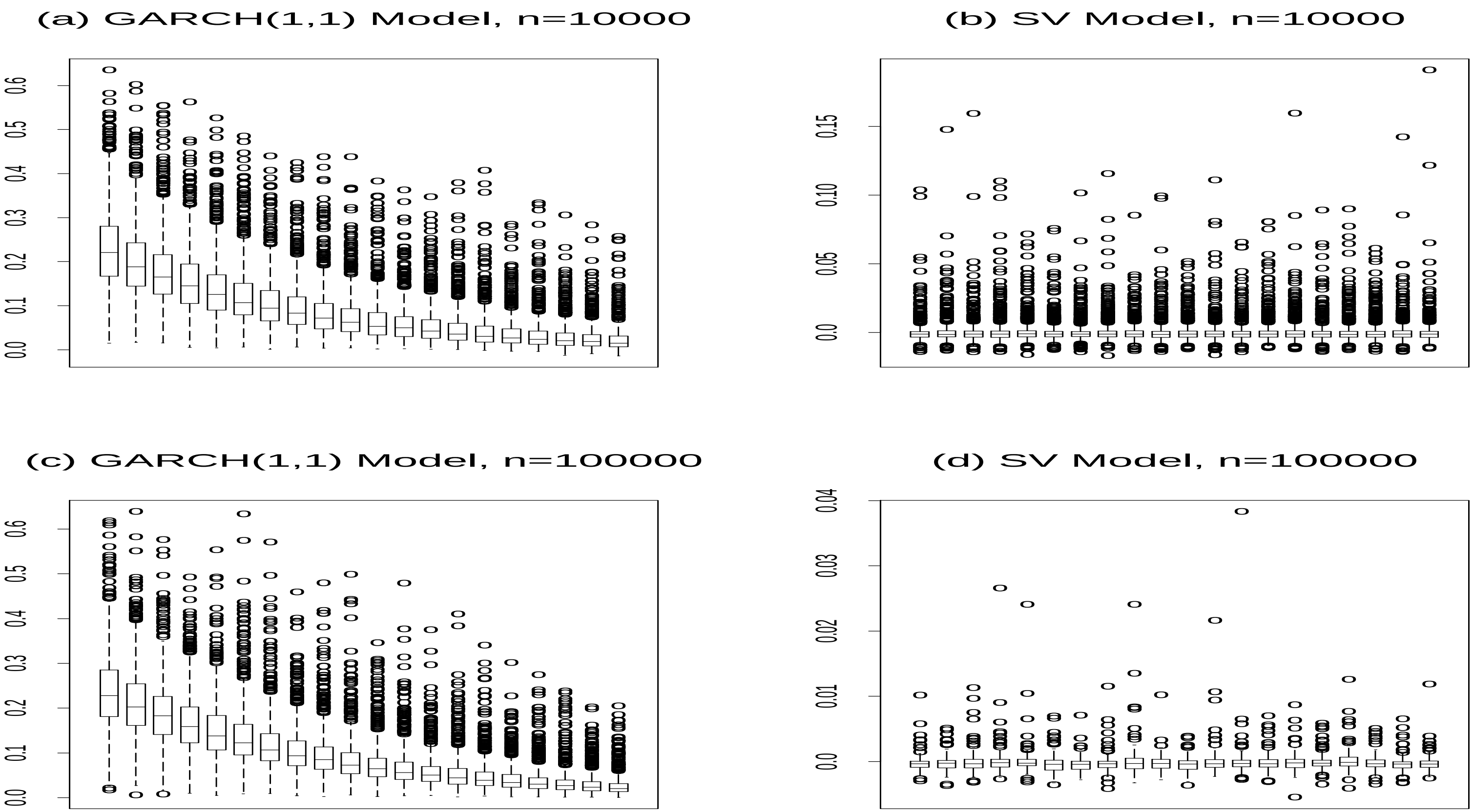

Figure 7. Boxplot comparison of sample ACFs of squared GARCH$(1,1)$ and stochastic volatility models with tail parameter $\alpha = 1.5$.

4. Frequency domain analysis of extreme events Mikosch and Zhao (2012)

- We assume (X_t) is a strictly stationary $\mathbb{R}^d$-valued sequence, regularly varying with index $\alpha > 0$.
- The extremogram for a given set A bounded away from zero

$$\begin{aligned}\rho_A(h) &= \lim_{n\to\infty} P(x^{-1}X_h \in A \mid x^{-1}X_0 \in A)\\ &= \operatorname{corr}\big(I_{\{x^{-1}X_0\in A\}}, I_{\{x^{-1}X_h\in A\}}\big)\,, \quad h \ge 0\,,\end{aligned}$$

is the (non-negative) ACF of a stationary process.

- 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 self-normalized periodogram for $\lambda \in (0,\pi)$:

$$\widehat{f}_{nA}(\lambda) = \frac{I_{nA}(\lambda)}{I_{nA}(0)} = \frac{\left|\sum_{t=1}^{n} \mathrm{e}^{-it\lambda} I_{\{x_n^{-1}\mathrm{X}_t \in A\}}\right|^2}{\sum_{t=1}^{n} I_{\{x_n^{-1}\mathrm{X}_t \in A\}}},$$

for a threshold sequence $x_n \to \infty$.

- One has $\widehat{f}_{nA}(\lambda) \stackrel{P}{\to} f_A(\lambda)$ 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 frequencies λ_j, and iid standard exponential E_j,

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

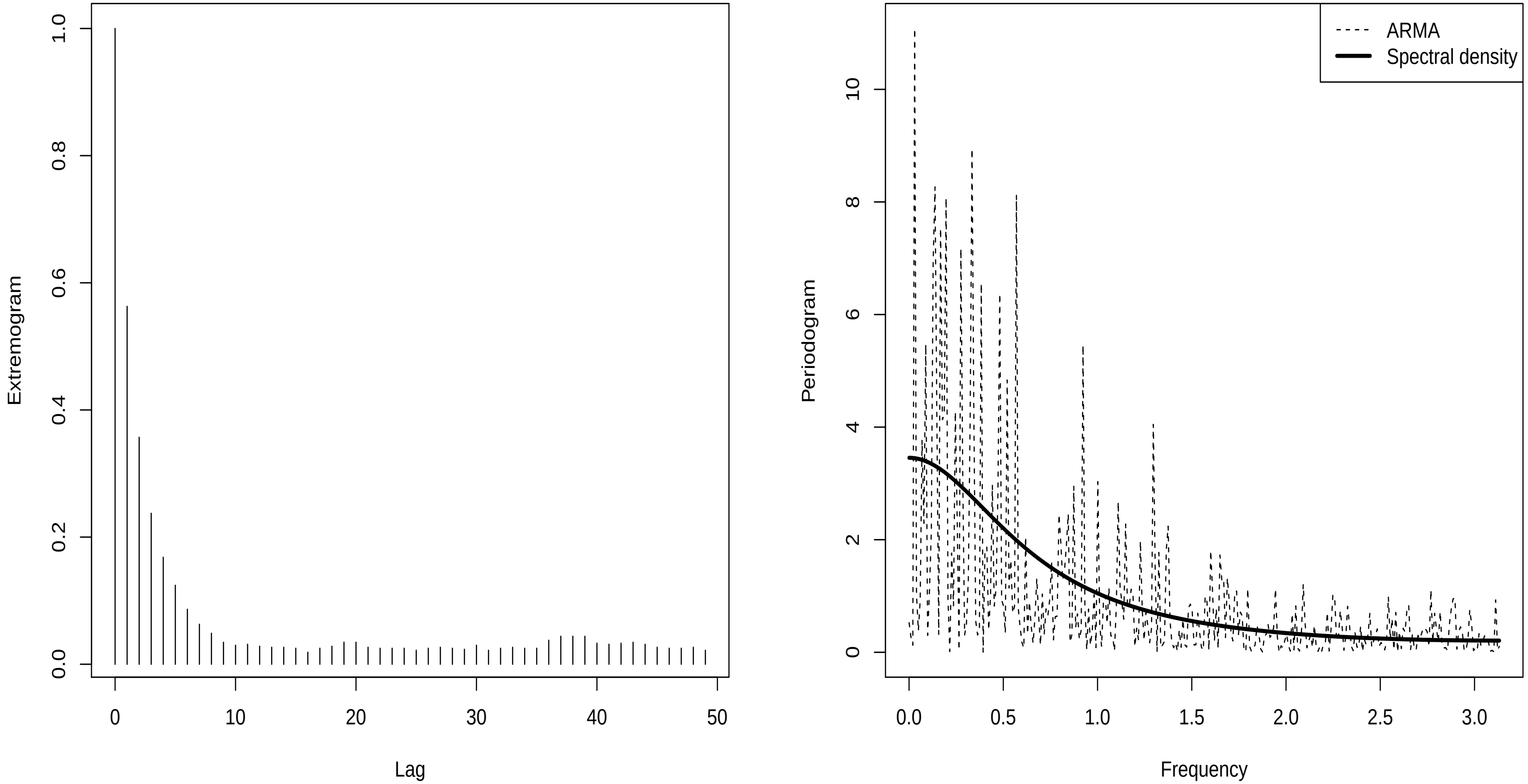

FIGURE 8. Sample extremogram and periodogram for ARMA(1,1) process with student(4) noise. $\boldsymbol{A = (1, \infty)}$

- 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 $\lambda_j \to \lambda$,

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

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

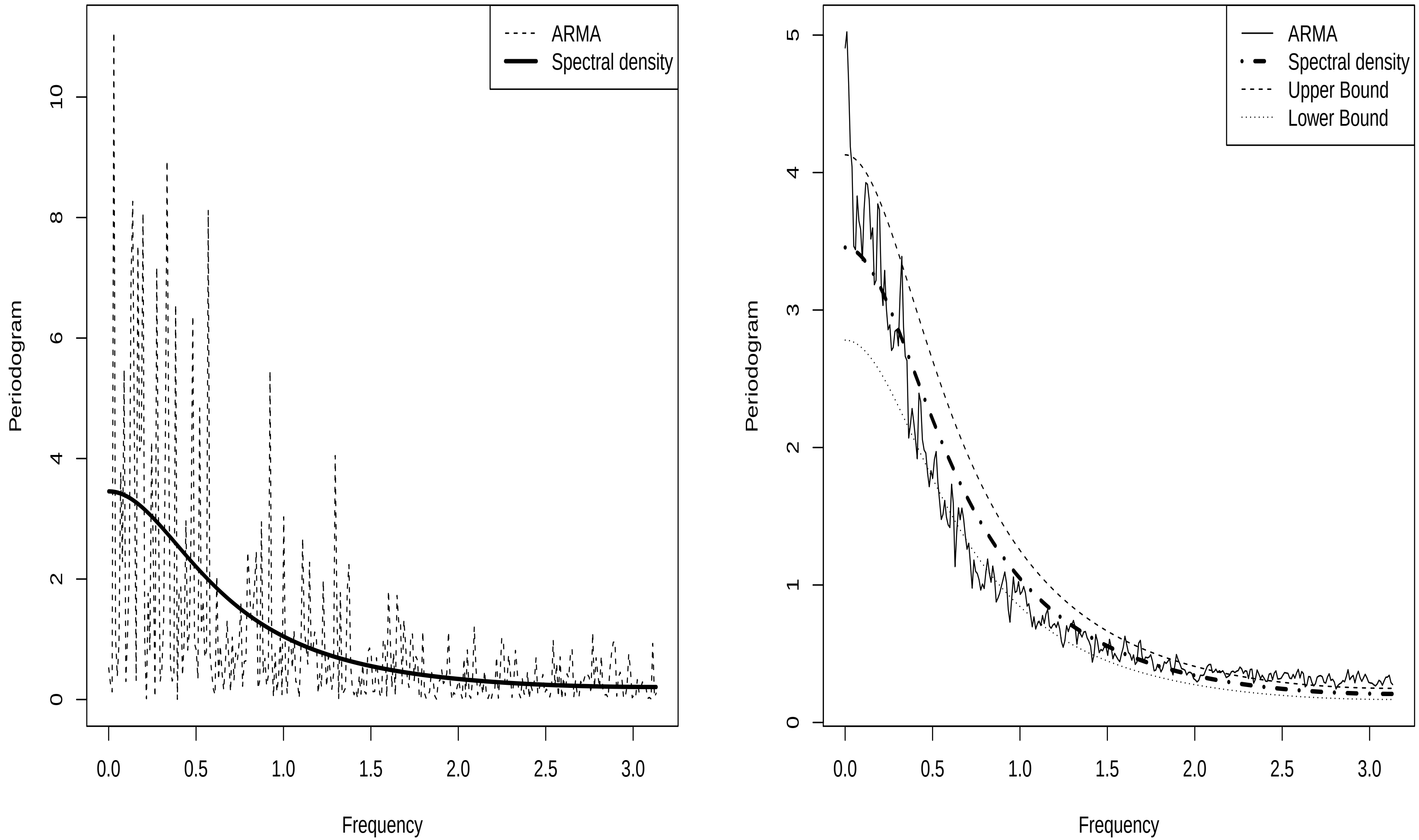

FIGURE 9. Raw and smoothed periodogram for ARMA(1,1) process with student(4) noise. $\boldsymbol{A = (1, \infty)}$

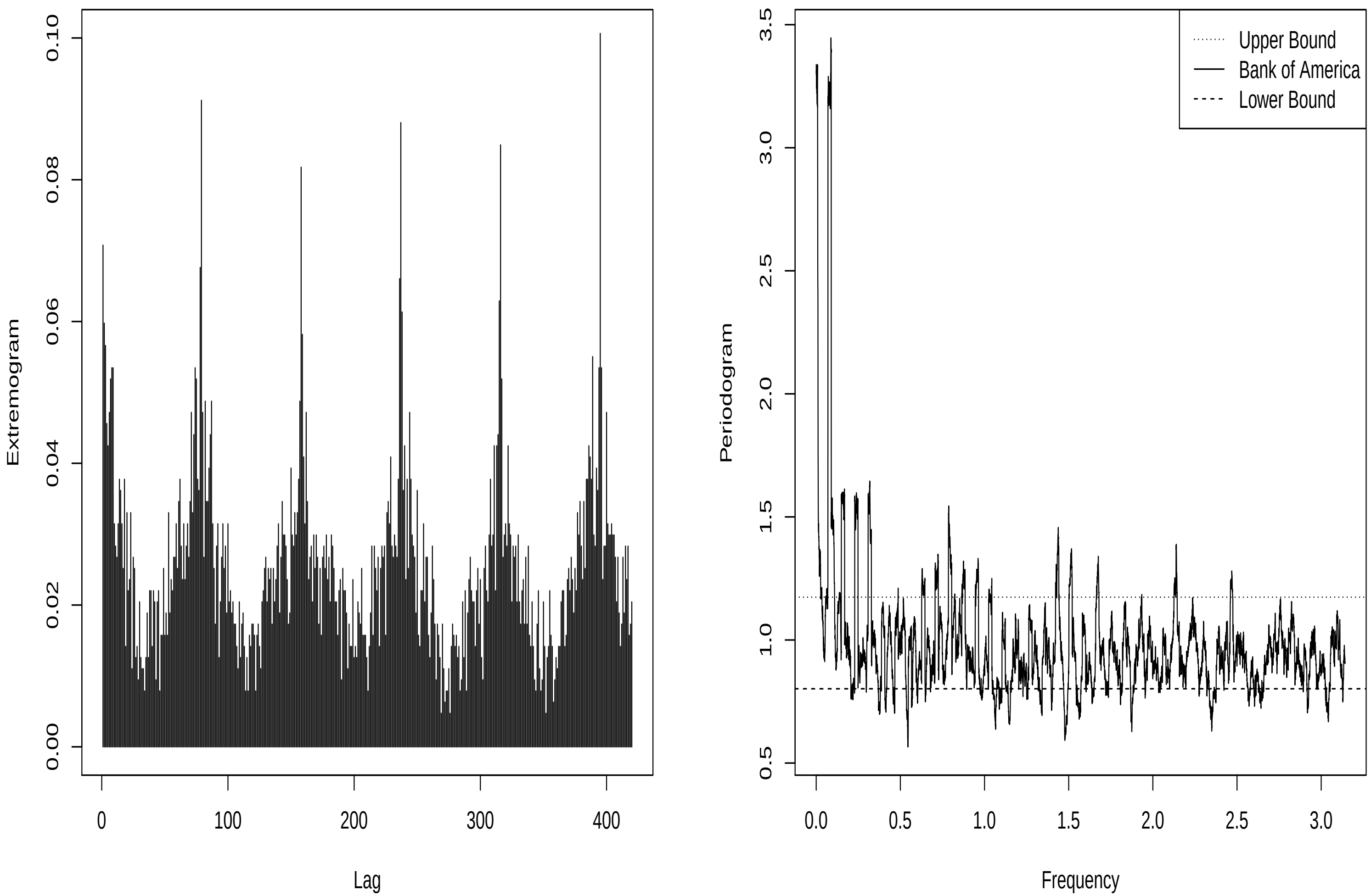

Figure 10. Sample extremogram and smoothed periodogram for BAC 5 minute returns. The end-of-the day effects cannot be seen in the corresponding sample autocorrelation function.

References

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

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

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

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

[5] 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.).

[6] 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.

[7] 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.

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

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

[10] Feller, W. (1971) *An Introduction to Probability Theory and Its Applications.* Vol. II. Second edition. Wiley, New York.

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

[12] Haan, L. de, Fereira, A. (2006) *Extreme Value Theory: An Introduction.* Springer, Berlin, New York.

[13] 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.

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

[15] Klüppelberg, C. and Mikosch, T. (1996) The integrated periodogram for stable processes. *Ann. Statist.* **24**, 1855–1879.

[16] Kokoszka, P.S. and Taqqu, M.S. (1996) Parameter estimation for infinite variance fractional ARIMA. *Ann. Statist.* **24**, 1880–1913.

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

[18] Mikosch, T., Gadrich, T., Klüppelberg, C. and Adler, R.J. (1995) Parameter estimation for ARMA models with infinite variance innovations. *Ann. Statist.* **23**, 305–326.

[19] Mikosch, T. and Straumann, D. (2002) Whittle estimation in a heavy-tailed $\mathbf{GARCH(1,1)}$ model. *Stoch. Proc. Appl.* **100** (2002), 187–222.

[20] Mikosch, T. and Zhao, Y. (2013) A Fourier analysis of extreme events. *Bernoulli* to appear.

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

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

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

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